基於PPG之瘻管功能異常分析

Analysis of Abnormal Fistula Function Based on Photoplethysmography

Po-Kai Yang, Ming-Chao Wu, Wen-Fong Wang\*

# 摘要

現今在台灣末期腎病的病患約90%會接受血液透析治療。當病患接受長期血液透時必須建立長期血管通路(vascular access)方便上針。自體動靜脈瘻管(arteriovenous fistula, AVF)及人工動靜脈瘻管(arteriovenous graft, AVG)為病患接受血液透析時併發症較低的血管通路(vascular access)選項，但使用一段時間後，血管可能產生狹窄、血栓、或形成假性動脈瘤(pseudoaneurysm)，若血管完全阻塞會導致病患多餘的水分及毒素無法及時自體內清除，甚至造成生命危險。

2020KDOQI 血管通路照護指引指出，目前仍無足夠證據顯示常規用儀器監測血管通路狹窄對血管通路通暢度維持有幫助。目前有許多針對血管通路功能的感測器，然而這些感測器大多以監測(surveillance)血管狹窄程度作為目標，較少有文獻針對預測血管通路的通暢率及預後直接進行著墨。光電容積圖(photoplethysmography, PPG)以非侵入性方式獲得血管隨脈衝壓引起的體積變化，有較高的準確率去分類瘻管狀態。

因此，本研究將受試者分為「易阻塞病患」、「不易阻塞的病患」兩個群體，利用ADInstruments PowerLab醫學級研究用感測儀器採集且濾波之訊號數據(PPG值)，分析群體間之差異，監測透析瘻管健康程度。

關鍵字:血液透析、瘻管、光體積變化描記圖法、互相關運算、機器學習

# ABSTRACT

In Taiwan, about 90% of end-stage renal disease(ESRD) patients received hemodialysis(HD) treatment. Arteriovenous fistula(AVF) and arteriovenous graft(AVG) are the preferred access for most patients receiving maintenance HD due to the lower complication rate. However, it is often complicated with thrombus formation, vascular access dysfunction, and failure in the long term. These complications prevent patients from receiving regular dialysis on schedule until an emergent thrombectomy is performed. Delay hemodialysis may cause excessive water accumulation or electrolyte imbalance, which becomes a life-threatening condition.

Currently, care of vascular access can be categorized into “surveillance” and “monitoring”. The 2020 update of the KDOQI guideline does not recommend using these surveillance methods as the sole indicator to do a vascular intervention. Many works on noninvasive sensors for evaluating AVF were published but mostly focused on the degree of stenosis. For direct predicting of patency of AVF, there’s still limited work on it. To acquire the hemodynamic status of a certain vessel, as a technique of optical measurement, photoplethysmography(PPG) is an appropriate option.

The study population is divided into two groups: “recurrent access dysfunction”,” good access patency”. We use ADInstruments PowerLab sensors to obtain PPG signal, and analyze differences between groups. This study aims to investigate whether the PPG signal can be a predictor for the patency of AV fistula.

Keywords:Hemodialysis、Fistula、Photoplethysmography、Auto-/Cross-Correlation,   
Machine Learning.

# 介紹

## 研究背景

在台灣，目前末期腎病(End-stage renal disease, ESRD)的病患約80000人1，治療末期腎病最常用的方法為血液透析(Hemodialysis, HD)。現今在台灣末期腎病的病患約90%會接受血液透析治療，當病患接受長期血液透析時必須建立長期血管通路(vascular access)方便安裝輸血針頭進行血液透析治療。一般來說，血管外科醫師會根據病患自身血管狀況作評估，決定血管通路的方式。

自體動靜脈瘻管(arteriovenous fistula, AVF)及人工動靜脈瘻管(arteriovenous graft, AVG)因為併發症少、感染率低，相較於其他血管通路(如：中心靜脈導管(central venous catheter))等方法，是較好且較常見的選項2,3。然而，不論自體動靜脈瘻管或人工動靜脈瘻管，都有其缺點。在早期階段，瘻管可能在手術後無法順利成熟，成為可用的血管通路。隨著時間推移，血管可能產生狹窄、血栓、或形成假性動脈瘤(pseudoaneurysm)等現象。在進行血液透析治療時，透析護理師需對病患之瘻管進行反覆穿刺上針，導致瘻管的血管內膜組織增生或血管壁重塑，進而形成血管狹窄。而血管狹窄容易導致瘻管通路流速下降，從而造成血栓或瘻管通路功能異常等現象4。若瘻管通路功能異常甚至完全阻塞時，病患將無法接受常規的血液透析治療，必須先接受緊急血栓移除手術(thrombectomy)將血管打通，才能繼續接受血液透析治療。有些病患可能因此導致多餘的水分及毒素無法及時自體內排出，甚至造成生命危險。此外，血管通路反覆阻塞也被認為跟心血管疾病事件及死亡率相關5。因此，能提前找出這些反覆阻塞的病患並予以介入處理十分重要

## 相關研究

目前有許多針對血管通路功能的感測器，然而這些感測器大多以監測(surveillance)血管狹窄程度6作為目標。但2020KDOQI血管通路照護指引指出，目前仍無足夠證據顯示，使用常規用儀器來監測血管通路的狹窄化，對血管通路通暢度之維持有幫助7。這其中的疑問包括：1. 偵測工具是否真的能可靠地偵測血管內狹窄 2. 目前對於何種程度的狹窄日後會造成血管通路阻塞仍無定論 3. 除了狹窄之外，還有其他原因(例如：血栓形成或通路周圍有血管瘤發生)可能造成血管通路阻塞。

由許多研究顯示，我們可以根據從雙手同樣的手指或是有瘻管之手臂等部位，擷取出心搏活動在動脈和微血管所引發之心搏脈衝訊號變化，例如：以光體積變化描記圖法(Photoplethysmography, PPG)所產生的光電訊號，進行該類訊號的比對和分析，以研究血管通路之結構或功能。P.-Y. Chiang等人透過自製的無線PPG感測器，利用採集瘻管之手臂的PPG訊號，計算出PPG振幅、收縮壓、舒張壓、心律等參數，透過類神經網路的訓練與辨識瘻管狀態所獲得的結果，並與醫院常規使用於觀察瘻管狀態的超音波多普勒血流儀器之檢查結果作比對，證明彼此之間具有高度的相關性8,9。另外，為了能夠有較高的準確率去分類出瘻管窄化等級之結果，J.-X. Wu等人利用雙側手指PPG訊號的上升時間和振幅的差異，以便反映出瘻管狹窄程度10。Y.-C. Du等人也利用雙側手指PPG訊號，計算洗腎前/後的上升斜率與下降斜率對瘻管窄化進行分類11,12。上述這些研究都是針對瘻管窄化進行分析，但因瘻管窄化無法與血管通路阻塞的機率畫上等號，因此距離臨床直接應用仍有一段距離。

## 研究動機(問題)及解決方案

PPG是通過LED光線照射皮膚下的微血管，並測量光吸收的變化量，是一種以非侵入性方式，獲得血管隨脈衝壓引起的體積變化。因此，本研究將受試者分為接受血液透析治療之瘻管通路是「易阻塞病患」和「不易阻塞的病患」等兩個群體，利用ADInstruments PowerLab儀器系統之PPG感測模組採集實驗訊號，分析兩個群體間之差異，並監測透析瘻管狀態，進而預測瘻管阻塞的機率。

## 論文大綱

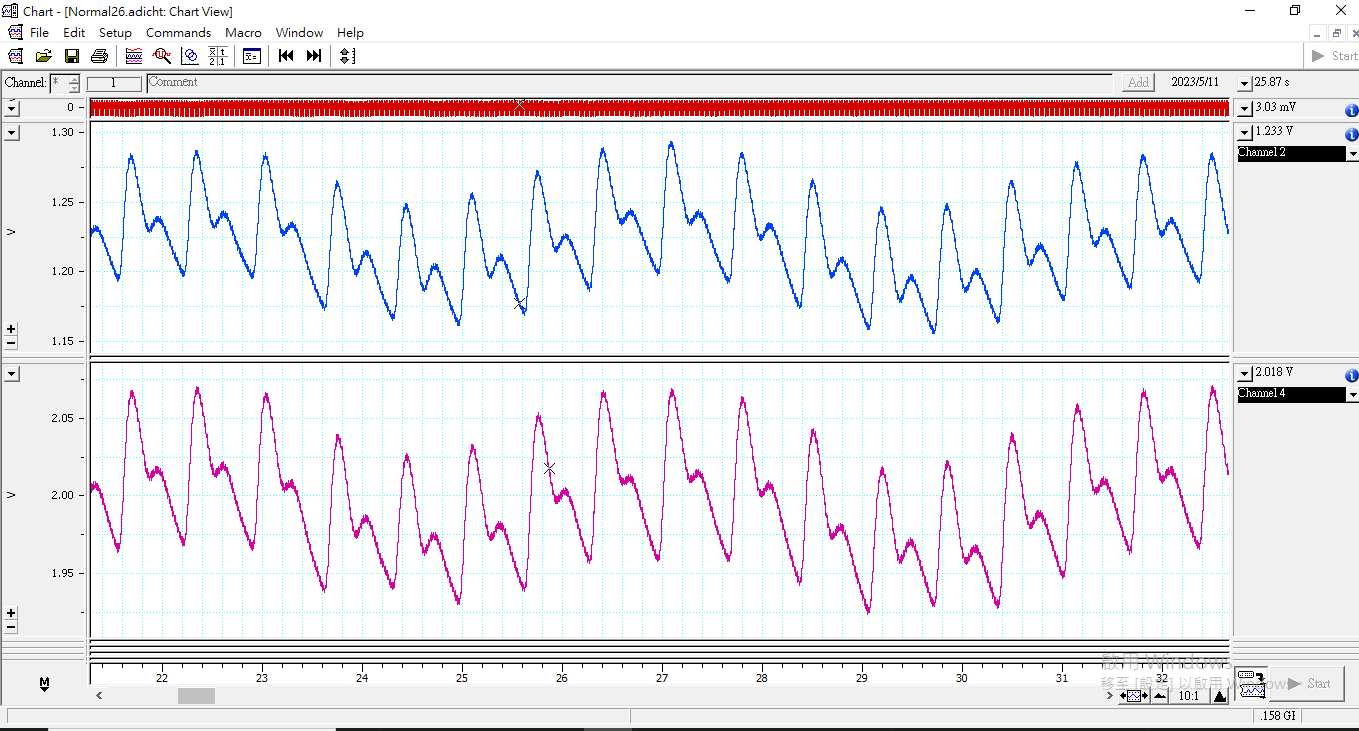
第二章會介紹我們使用的儀器、實驗參與者、實驗方法以及PPG訊號特徵參數等的介紹。第三章說明訊號的前處理、特徵篩選和統計分析評估、以及與瘻管狀態有關的訊號辨識演算法。第四章說明和討論分類器分類出來的結果。第五章總結分類出來的結果，以及未來可以改進的地方。

# 材料

## 研究儀器

1. (b)



(c)

圖一、PPG訊號感測儀器：(a) PowerLab系統、(b) 指夾式PPG感測器模組、(c) LabChart軟體操作畫面

在本研究中，我們使用PowerLab系統(如圖一(a)所示)進行PPG訊號的擷取，此系統是由ADIstruments公司所生產。PowerLab可擷取和收集如PPG、ECG等生理訊號，同時也提供其他類比/數位生理訊號(例如：EMG等)的輸入通道，以及支援其他類生理訊號之類比數位轉換（Analog-to-digital conversion, A/D）與訊號調變。本研究實驗使用了指夾式的PPG感測器模組(如圖一(b)所示)，此時PPG感測器的採樣率設為1000Hz。另外，PowerLab系統可搭配LabChart軟體，對PPG的原始訊號(如圖一(c)所示)進行即時顯示和觀察，並將PPG訊號儲存成.adicht的檔案格式。而後續PPG訊號的分析及處理，則是利用Python程式語言撰寫PPG訊號處理程式，來進行後續的訊號分析和應用。

## 受測者

本研究已通過國立成功大學醫學院附設醫院之醫學倫理審查委員會之醫學倫理審查，IRB醫學倫理審查通過之核准文件編號為A-ER-112-135，本研究所執行之實驗程序與實驗內容皆屬於非侵入式的行為，符合赫爾辛基宣言之精神。

在本研究中，將接受血液透析治療之病患依其瘻管狀態區分為「易阻塞病患」、「不易阻塞病患」兩組。另外，關於「易阻塞病患」及「不易阻塞病患」群組的實驗參與者，徵召來自國立成功大學醫學院附設醫院斗六分院的洗腎中心之病患。本研究實驗共有18位已接受瘻管手術的患者參與實驗，其中「易阻塞」的病患5位，「不易阻塞」的病患13位。

實驗參與者的選擇基於一系列納入或排除條件，詳列於表一中。選定的參與者需符合以下條件：年齡需大於等於20歲，而性別可以是男性或女性。此外，被選入實驗的參與者須在左手或右手處接受過自體動靜脈瘻管(arteriovenous fistula, AVF)及人工動靜脈瘻管(arteriovenous graft, AVG)，並且不得有心血管疾病的病史。另外，實驗參與者將依照瘻管介入手術的頻率，被區分為兩種類型：介入手術頻率高者，定義為三個月內接受過兩次以上瘻管介入手術，被歸類為「易阻塞病患」；介入手術頻率低者，被歸類為「不易阻塞病患」5。

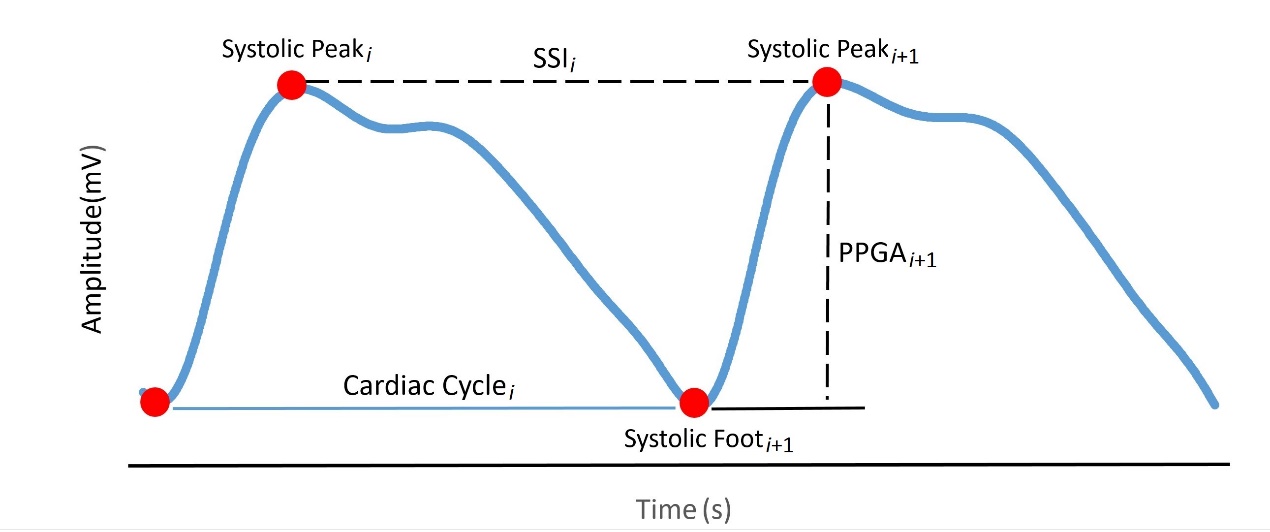
表一 受試者納入條件

|  |  |  |
| --- | --- | --- |
|  | 易阻塞病患 | 不易阻塞病患 |
| 年齡 | 大於等於20歲 | 大於等於20歲 |
| 心血管疾病 | 無 | 無 |
| 是否安裝瘻管 | 是 | 是 |
| 介入頻率 | 三個月內兩次以上 | 三個月內小於兩次 |

## 實驗方法

在實驗開始前，會先向實驗參與者說明研究實驗流程，並向實驗參與者詢問瘻管位置。在本研究中，沒有安裝瘻管的手稱為健側手(Health Side: H)，有安裝瘻管的手稱為患側手(Affected Side: A)。而在本研究實驗所募集到的實驗參與者中，多數的患者之瘻管是安裝在左手，少數是安裝在右手。之後，將指夾式PPG感測器探頭夾住在實驗參與者的健側與患側雙手食指的指腹位置，並提醒實驗參與者放輕鬆。當手指頭被感測器夾住後，不要對感測器探頭作按壓動作。實驗時間一共進行5分鐘的PPG訊號擷取，如果實驗期間實驗參與者的瘻管有異常狀態，則標記此PPG訊號為異常訊號。

## PPG訊號與實驗特徵參數



圖二、PPG訊號和其特徵點

在圖二中，我們可以觀察到一個心搏週期內的PPG訊號及其擁有的特徵點，包括：收縮壓峰值(Systolic Peak)、收縮壓峰腳(Systolic Foot)、以及心搏週期長度(Cardiac Cycle Interval)等。Systolic Peak指的是PPG波形中收縮期的最大值，Systolic Foot是PPG波形中收縮期的起始點，心搏週期指的是從本次心跳開始到相鄰之下一次心跳開始之前，所經歷完整的過程，包含了一個收縮期與一個舒張期。而Cardiac Cycle Interval的衡量方式在許多的研究中，皆以Systolic Peak to Peak Interval (SSI)來代替，指得是從本次心跳之Systolic Peak時間點到下一次心跳Systolic Peak時間點之間的間隔。Systolic Foot、Systolic Peak、SSl等的量測與計算，有助於評估接受血液透析之病患的瘻管通路的良好狀態。

除了以上的參數之外，表二中顯示了在本研究中所提議的其他參數，例如：PPGA、PPGA-H/A、PPGA-A/H、PPG上升時間、PPG下降時間、SSI等PPG參數。關於PPGA值之計算，其方法是透過Systolic Peak的PPG振幅值，減去Systolic Foot的PPG振幅值。而PPGA-H/A的值，是透過將健側手指獲得到的PPGA值，除以患側手指獲得到的PPGA值。反之，對於PPGA-A/H的值，則是將患側PPGA值除以健側PPGA，所以PPGA-A/H和PPGA-A/H互為倒數關係。計算PPG上升時間時，其方法是透過當前Systolic Peak的時間點減去當前Systolic Foot的時間點，來獲得PPG上升時間數值。至於計算PPG下降時間，其方法是透過下一個Systolic Foot的時間點，減去當前Systolic Peak的時間點，以獲得PPG下降時間數值。

表二、實驗特徵參數定義

|  |  |
| --- | --- |
| **特徵(參數)** | **定義** |
| Systolic Foot | PPG波形中收縮壓的起始點。 |
| Systolic Peak | PPG波形中收縮壓最大值。 |
| Systolic Peak to Peak Interval (SSI) | 本次PPG訊號波峰到下一個PPG訊號波峰的時間間隔。 |
| PPGA | 波谷到波峰的垂直振幅差。 |
| PPGA-H/A、PPGA-A/H | PPGA-H/A：將健側PPGA值除以患側PPGA。  PPGA-A/H：將患側PPGA值除以健側PPGA。 |
| PPG上升時間 | PPG訊號從波谷到波峰時間。 |
| PPG下降時間 | PPG訊號從波峰到波谷時間。 |
| AVG | 整段資料的平均值。 |
| STD | 整段資料的標準差。 |
| VAR | 整段資料的變異度。 |

在本研究所使用的實驗特徵參數(如表二)中，還包括了3個統計特徵，例如：平均值(Average: AVG)、標準差(Standard Deviation: STD)、變異度(Variation: VAR)。有關於AVG和STD之計算公式如(公式一)和(公式二)所示，其中N是資料數值*Xi*的總數。

(公式一)

(公式二)

變異度(公式三)是醫學上常用於衡量一組數據(假設S={*X*1, *X*2, …, *X*N})分散程度的一種指標，其計算方式是將該組數據S之最大值減去同組數據之最小值，再除以同組全部數據之平均值，如(公式三)所示。

(公式三)

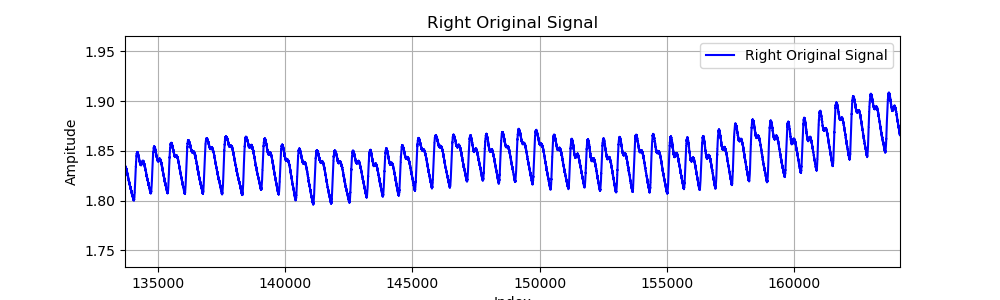
其中Max()和Min()是求集合S中最大值和最小值的函數。

# 研究方法

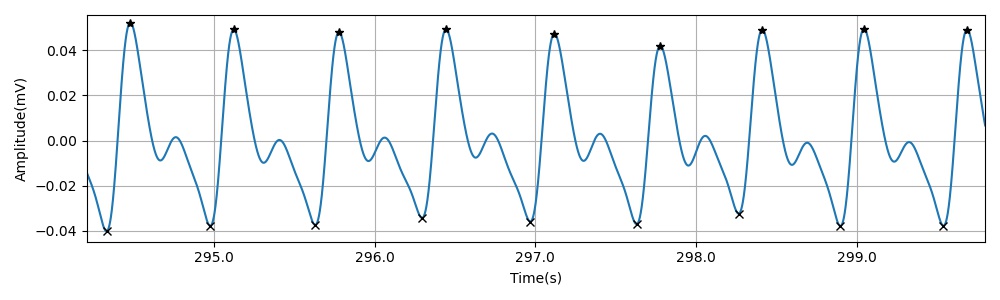
在實驗中所擷取到的原始訊號，將進行訊號濾波處理工作，以濾除不必要的雜訊。之後的訊號處理，將會使用訊號特徵峰值檢測演算法標記出收縮壓腳(Systolic Foot)和收縮壓峰值點(Systolic Peak)，以及利用收縮壓腳值(Systolic Foot)或收縮壓峰值(Systolic Peak)的這兩個特徵點進行Cardiac Cycle的切割，以便找出表二所示之各個PPG特徵值。再利用切割之後的一個Cardiac Cycle與整段PPG訊號資料進行Auto-/Cross-Correlation運算，來計算不同的PPG訊號來源之間的相關性，並提取後續要進行統計推論和智慧辨識的相關性特徵。

## 訊號前處理

透過PowerLab系統之PPG感測模組所擷取到的左/右雙手PPG原始訊號(圖三(a))，這些原始訊號含有大量的雜訊，例如：呼吸、環境光源、手部運動造成的基線飄移、運動偽影等。為了要去除這些雜訊，我們使用了Butterworth帶通濾波器進行雜訊處理13。這些原始訊號將會經過4階Butterworth帶通濾波器的訊號處理，該濾波器之導通頻帶設定為0.5HZ到7HZ之間，這有助於去除呼吸、環境光源、手部運動造成的基線飄移、運動偽影等，使得PPG訊號更加可靠且有利於後續的瘻管阻塞狀態之分析和辨識，如圖三(b)所示。



(a)



(b)

圖三、PPG訊號—(a)原始訊號，(b)去除雜訊後的訊號

## 訊號處理

### 峰值特徵標記

經過Butterworth濾波器處理後，原始訊號已經被轉換為濾波過後的PPG訊號。在接下來步驟中，我們將對經過濾波的訊號進行峰值標記，以便於後續進行PPG訊號週期分割與訊號特徵計算。在PPG訊號峰值檢測方法中，我們使用了Python程式語言的Scipy函式庫中的find-peaks函式，來進行濾波過後PPG訊號中的峰值偵測，這些峰值標記結果如圖三(b)中之波峰和波谷的標記點所示。

### Cardiac Cycle Segmentation

透過PPG訊號峰值檢測方法處理後，我們可將收縮期峰值點(Systolic Peak)與收縮期谷值點(Systolic Foot)標記出來，再透過這兩種標記方式，來分割出各別獨立的心搏週期(Cardiac Cycle)。在本研究中，我們採用Systolic Peak作為心搏週期的分割點，以獲得心搏週期間隔的數值。當偵測到第i個Systolic Peak時，第i個Systolic Peak與第i+1個Systolic Peak之間的區間資料，可視為是第i個Cardiac Cycle，利用此方法可以完成各個單獨的心搏週期之分割(Cardiac Cycle Segmentation)。

基於以上心搏週期之分割方法，在本研究的後續特徵資料計算過程中，皆以Systolic Peak作為心搏週期的分割點，從而產生與PPG心搏週期訊號相關的各種瘻管阻塞狀態之特徵資訊。

### Auto-/Cross-Correlation運算

自相關/互相關運算(Auto-/Cross-Correlation: CC)是一種於訊號處理中常用的數學運算方法(公式四)，透過滑動點積計算的方式來衡量兩個數值序列間之相似性或關聯性的一種方法。假設一組經濾波處理後的PPG訊號數列為*Qi* (其中1≤ *i* ≤ M，M是一個整數)，則將*Q*設為執行自相關/互相關運算的左手/右手PPG訊號資料(Data)集合(如圖三(b))。利用前述的Cardiac Cycle分割方法，將Data集合中的資料進行Cardiac Cycle分割，並從分割後的Data集合中，隨機找出一個完整的Cardiac Cycle集合，將它當作是執行自相關/互相關運算所需要的滑動視窗訊號之左手/右手PPG訊號資料數列*Pj* (其中1≤ *j* ≤ N，N是一個整數)，也就是當作滑動視窗的Pattern集合，則可實行自相關/互相關計算的公式如下所示。

(公式四)

其中Data的長度為M，Pattern的長度為N，且M會大於或等於N。另外，t為整數指標變數，且1≤ *t* ≤ (M-N+1)， 是集合*P*集合的平均值， 則是*Q*集合的部分集合(例如：(*Q*(*t*+1), ..., *Q*(*t*+n)))之平均值，對於每個CCr(*t*)的計算值，其正規化後範圍為−1 到 1之間。

當進行完自相關/互相關運算後，相關數列 *CCr*(*t*) (其中1≤ *t* ≤ (M-N+1)) 儲存了在*Q*數列上滑動*P*數列作自相關/互相關計算的結果(公式四)，則這段相關數列 *CCr*(*t*) 的資料長度為M-N+1。

### Auto-/Cross-Correlation資料模型

Auto-/Cross-Correlation資料模型之建立，可分成兩個階段。第一階段是使用Auto-/Cross-Correlation運算，建立健側手指與患側手指的PPG訊號之二種自相關和二種互相關等運算組合，分別以HH、AA代表自相關運算組合，以及以HA、AH代表互相關運算組合。第二階段則是根據HH、AA、HA、AH等自相關/互相關係數數列所形成之係數曲線，計算其曲線下的面積，提供後續對瘻管通路狀態之統計推論和智慧辨識。透過上述這兩個階段所建立的Auto-/Cross-Correlation資料模型，可以用來揭露和評估血液透析用瘻管通路狀態的良好程度。

本研究之詳細工作流程描述如下：

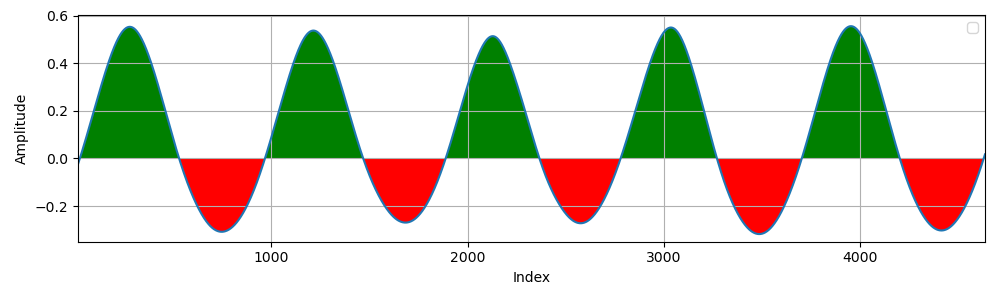
**第一階段**：將左/右雙手經訊號處理濾波後之PPG訊號，利用前述對Cardiac Cycle分割的方法，分割出相對於實驗參與者健側與患側的PPG訊號之滑動視窗訊號數列之H\_Pattern與A\_Pattern。然後，將分割完成後的PPG Pattern與PPG Data進行匹配組合運算，使其可以產生H\_Data-H\_Pattern (HH)、A\_Data-A\_Pattern (AA)、H\_Data-A\_Pattern (HA)與A\_Data-H\_Pattern (AH)等四種自相關/互相關係數運算組合。自相關/互相關係數運算之執行步驟，敘述如下：

步驟一：利用Cardiac Cycle分割出H-Pattern和A-Pattern。

步驟二：利用分割出來的H-Pattern與A-Pattern和H-Data與A-Data，進行配對組合之自相關/互相關係數運算。

步驟三：經自相關/互相關係數運算後，產生HH、AA、HA、AH等四種組合，再分別計算出HH、AA、HA、AH等四種組合各別的自相關/互相關係數數列之平均數(AVG)、標準差(STD)、變異度(VAR)等數據。

**第二階段**：我們將計算出的HH、AA、HA、AH等四種自相關/互相關係數數列，對這些數列的曲線進行曲線下面積之計算(如圖四所示)。根據曲線數值，我們可以進行以下判斷；當曲線的數值大於0時，計算所得之曲線下面積稱為正區域面積(Poitive area: P-area)，該區域在圖九中以綠色部分表示。而當曲線數值小於0時，則稱為負區域面積(Negative area: N-area)，該區域以紅色部分表示。



圖四、自相關/互相關係數數列曲線下之面積

## 參數之統計分析

在本研究中，進行統計分析的自變量參數種類，可以分為三種類別。首先，是由自相關/互相關運算經計算而得的自相關/互相關係數數列(例如：HH、AA、HA、AH等四種組合)的平均數(AVG)、標準差(STD)、變異度(VAR)，這些參數組合形成第一類參數集合。其次，透過峰值檢測方法，識別出患側手/健側手之PPG訊號的波峰與波谷，進而計算得到PPGA、PPGA-H/A、PPGA-A/H、PPG上升時間、PPG下降時間、以及SSI等，這些參數構成了第二類參數集合。最後，透過自相關/互相關運算所獲得的相關係數數列之曲線下面積，亦即P-area和N-area，形成了第三類參數集合。

而統計分析的因變量，則設定為接受血液透析治療之瘻管通路，是「易阻塞病患」或是「不易阻塞病患」等兩個群體。為了評估以上三類自變量參數在統計分析上，針對因變量參數的差異顯著性，我們利用獨立樣本T檢定來進行分析和評估不同自變量參數在後續智慧辨識之應用上的差異14。在T檢定中，我們需要設定虛無假設H\_0與對立假設H\_1。虛無假設通常表示2個母體間的平均值是相等的，這意味著兩個母體是相同的母體。而對立假設則代表2個母體間的平均值具有顯著差異，也意味者這兩個母體是不相同的母體。

根據樣本數據來判斷虛無假設是否成立，當p值大於0.05時，我們接受虛無假設，反之，當p值小於0.05時，表示樣本數據之間具有足夠的差異性，我們將拒絕虛無假設，並接受對立假設。表三至表五顯示了三類自變量參數，在「易阻塞病患」和「不易阻塞病患」等兩個群體中分佈差異的顯著性，以及進行獨立樣本T檢定後得到的P-value結果。

表三、第一類參數T-test之P-value值

|  |  |
| --- | --- |
| **Parameter** | **P-value** |
| HH-AVG | 0.106☆ |
| HH-STD | 0.96 |
| HH-VAR | 0.77 |
| AA-AVG | 0.99 |
| AA-STD | 0.62 |
| AA-VAR | 0.76 |
| HA-AVG | 0.82 |
| HA-STD | 0.37 |
| HA-VAR | 0.30★ |
| AH-AVG | 0.87 |
| AH-STD | 0.80 |
| AH-VAR | 0.83 |

表四、第二類參數T-Test之P-value值

|  |  |  |
| --- | --- | --- |
| **Parameter** | **P-value(健側手)** | **P-value(患側手)** |
| PPGA | 0.24★ | 0.77 |
| PPGA-H/A(*a*), PPGA-A/H(*b*) | 0.06(*a*)☆ | 0.20(*b*)★ |
| PPG上升時間-AVG-H(*a*), PPG上升時間-AVG-A(*b*) | 0.46(*a*) | 0.65(*b*) |
| PPG上升時間-STD-H(*a*), PPG上升時間-STD-A(*b*) | 0.25(*a*)★ | 0.46(*b*) |
| PPG上升時間-VAR-H(*a*), PPG上升時間-VAR-A(*b*) | 0.25(*a*)★ | 0.409(*b*) |
| PPG下降時間-AVG-H(*a*), PPG下降時間-AVG-A(*b*) | 0.85(*a*) | 0.05(*b*)\* |
| PPG下降時間-STD-H(*a*), PPG下降時間-STD-A(*b*) | 0.2(*a*)★ | 0.29(*b*)★ |
| PPG下降時間-VAR-H(*a*), PPG下降時間-VAR-A(*b*) | 0.26(*a*)★ | 0.29(*b*)★ |
| SSI-AVG-H(*a*), SSI-AVG-A(*b*) | 0.06(*a*)☆ | 0.06(*b*)☆ |
| SSI-STD-H(*a*), SSI-STD-A(*b*) | 0.06(*a*)☆ | 0.06(*b*)☆ |
| SSI-VAR-H(*a*), SSI-VAR-A(*b*) | 0.15(*a*)★ | 0.15(*b*)★ |

註：上註標(a)和(b)分別延伸標註各個參數是定位於健側手(-H)或是患側手(-A)

表五、第三類參數T-Test之P-value值

|  |  |
| --- | --- |
| **Parameter** | **P-value** |
| HH-P-area | 0.06☆ |
| HH-N-area | 0.828 |
| AA-P-area | 0.361 |
| AA-N-area | 0.45 |
| HA-P-area | 0.087☆ |
| HA-N-area | 0.292★ |
| AH-P-area | 0.346★ |
| AH-N-area | 0.419 |

## 智慧辨識

從表三至表五中P-value所顯示的結果，可以看出幾乎所有的自變量參數，都無法展示出對於兩種瘻管狀態的因變量，是否具有顯著性的差異。因此，我們嘗試使用表三至表五中的自變量參數，依其P-value大小之值選擇兩個自變量特徵集合，再利用K-Nearest Neighbors與Support Vector Machine等監督式學習演算法，進行兩種瘻管狀態的智慧辨識。其中，第一個自變量特徵集合(named as Feat.Set-1)為表三至表五中P-value小於0.1的所有自變量參數所組成之集合，第二個自變量特徵集合(named as Feat.Set-2)為表三至表五中P-value小於0.35的所有自變量參數所組成之集合。另外，我們參考了表三至表五中所列舉的P-value，設定篩選自變量參數進入這兩個特徵集合的P-value為0.1 (高辨識率條件)和0.35 (低辨識率條件)，其動機是為了分析自變量參數的P-value範圍，對於實施智慧辨識運算下會影響辨識結果的關鍵。這兩個特徵集合列出如下：

Feat.Set-1={HH-AVG, PPGA-H/A, PPG下降時間-AVG-A, SSI-AVG-H, SSI-AVG-A, SSI-STD-H, SSI-STD-A, HH-P-area, HA-P-area},

Feat.Set-2={HH-AVG, HA-VAR, PPGA-A/H, PPG上升時間-STD-H, PPG上升時間-VAR-H, PPG下降時間-AVG-A, PPG下降時間-STD-H, PPG下降時間-STD-A, PPG下降時間-VAR-H, PPG下降時間-VAR-A, SSI-AVG-H, SSI-AVG-A, SSI-STD-H, SSI-STD-A, SSI-VAR-H, SSI-VAR-A, HH-P-area, HA-P-area, HA-N-area, AH-P-area}.

### kNN

K個最近鄰演算法(k-Nearest Neighbors, kNN)15,16是一種監督式學習演算法，可應用於分類和回歸問題。kNN通過比較待分類樣本和其他樣本之間的特徵距離，來進行預測。在本研究中，我們將使用曼哈頓距離(Manhattan distance，如公式五所示)17，來計算不同樣本之間的距離。這種距離度量方式可以表示為*d*(*x*, *y*)，其中*xi*屬於一個訓練數據集合，而*yi*則為一個查詢點。透過公式五，找出*xi*與*yi*兩點之間的距離。

公式五

在分類問題上，kNN根據最相鄰的K個鄰居中所屬的最大多數標籤之類別，來確定新的查詢點所屬的標籤之類別；在回歸問題中，kNN使用鄰居的數值平均值來預測新的查詢點的標籤值。

kNN的主要優點在於演算步驟簡單易懂，易於實現與解釋，它不需要複雜數學模型。然而，kNN也有缺點，對於大型數據集，kNN的計算成本相對較高。因為每次預測都需要計算待分類樣本與所有訓練樣本之間的距離，這增加了計算的負擔。此外，選擇不適合的K值可能導致預測結果的有偏差，因此選擇適合的K值對於獲得較高準確率的結果很重要。

### SVM

支持向量機(Support Vector Machine，SVM)18,19是一種監督式學習演算法，可用於分類和回歸問題中，並在智慧辨識領域具有廣泛的應用。SVM的主要思想是在一個高維度的空間中，找到一條最佳的決策邊界(Decision Boundary)，如公式六所示。

公式六

其中，*y*(*x*)表示分類模型的輸出函數，對於輸入的特徵向量*x*，模型預測的輸出值，針對2個類別的分類函數*y*(*x*)會有+1與-1的兩個輸出數值，這用於表示數據的類別。ω*T*代表權重向量，用於表示分類超平面的方向，使其方便與特徵向量*x*進行內積運算。*x*代表是我們輸入到分類模型的特徵，*b*為一個常數，調整*b*的數值可以使回歸線在空間中作平移動作。

在分類模型中的決策邊界，可以將不同類別的數據有效地分隔開來。從決策邊界(Decision Boundary)上/下移動1個單位距離，以延伸出兩條邊界，如公式七所示。在這兩條邊界之間，可視為是一個緩衝區，當緩衝區越大時，分類結果的可信度就越高。

公式七

SVM可使用不同的Kernel Function，例如：線性核函數(Linear Kernel)、多項式核函數(Polynomial Kernel)和高斯核函數(Gaussian Kernel)，這些核函數可將低維度的數據轉映射至高維度空間，從而獲得更高的分類準確率。另外，當處理大型數據集並使用非線性核函數時，SVM的計算成本可能會相對提高，這也是需要考慮的一環。

# 結果分析

在本研究中，自變量參數經過獨立樣本T檢定後，將會以統計分析結果所產生的兩個自變量特徵集合(Feat.Set-1和Feat.Set-2)，進行kNN與SVM兩種監督式學習演算法的訓練和智慧辨識。

本研究實驗共有18位安裝廔管的患者參與實驗，其中「易阻塞」的病患5位，「不易阻塞」的病患13位，臨床及血管通路基本狀態如表六。

表六

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Frequent occlusion** |  | **Non-frequent occlusion** |  |
|  | n(mean) | %(SD) | n(mean) | %(SD) |
| Total | 5 |  | 13 |  |
| ***Demographics*** |  |  |  |  |
| Age, mean(SD) | 66.6 | 11.19 | 68.54 | 13.23 |
| Gender, n(%) |  |  |  |  |
| Male | 0 | 0 | 7 | 54% |
| Female | 5 | 100% | 6 | 46% |
| **Vascular character** |  |  |  |  |
| Access Type, n(%) |  |  |  |  |
| AVF | 0 | 0 | 9 | 69% |
| AVG | 5 | 100% | 4 | 31% |
| Vascular access inner diameter (mm), mean(SD) | 6.8 | 0.46 | 6.85 | 0.58 |
| Access duration (years), mean(SD) | 2.8 | 0.84 | 5.69 | 4.85 |
| Aneurysm formation, n(%) |  | 4 |  |  |
| Yes | 0 | 0 | 6 | 46% |
| No | 5 | 100% | 7 | 54% |
| Stent deployment, n(%) |  |  |  |  |
| Yes | 4 | 80% | 4 | 31% |
| No | 1 | 20% | 9 | 69% |

為了增加訓練模型的可靠性，我們採用了K-Fold交叉驗證策略，在K-Fold方法中會將資料集合切割分成K個子集合。本次實驗K-Fold的K值設為5，也就是將資料集合切割成五等份。意味著辨識模型訓練要進行五次，每次會從這五份隨機挑選其中四份當作是訓練資料集，另外一份當作是驗證集。最終，再取每次辨識結果的平均值，作為模型評估的效能指標值。為了評估模型效能，我們可以透過分類完成後產生的混淆矩陣(表七)進行準確率(Accuracy)、靈敏度(Sensitivity)、特異度(Specificity)、精確度(Precision)與F1-score(公式八至公式十二)等，進行辨識效果的評估。

表七、混淆矩陣示意圖

|  |  |  |  |
| --- | --- | --- | --- |
|  | | 實際值 | |
| 類別一 | 類別二 |
| 預測值 | 類別一 | TP | FP |
| 類別二 | FN | TN |

公式八

公式九

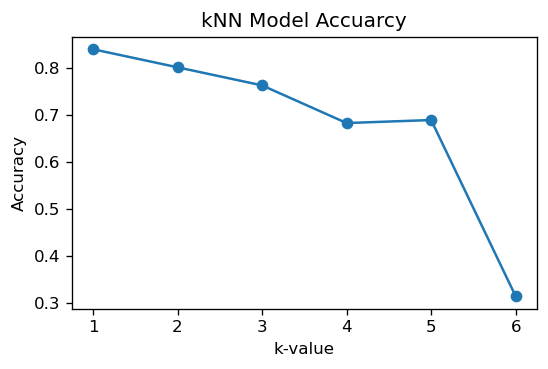
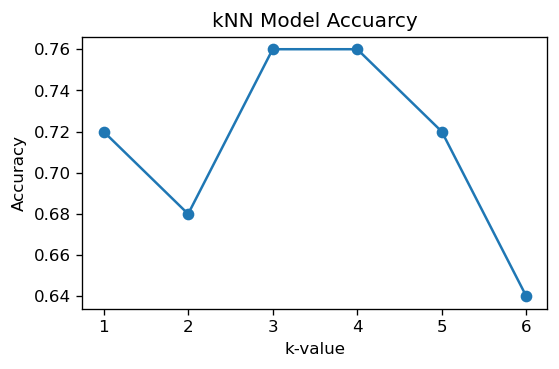
公式十

公式十一

公式十二

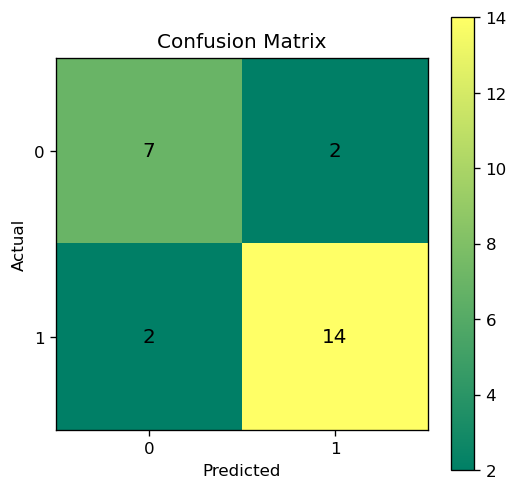
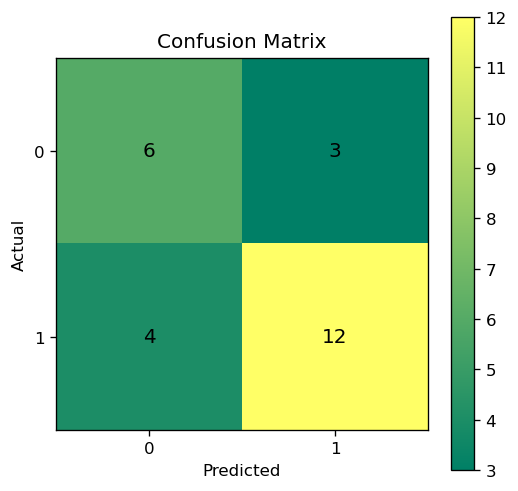
## kNN分析

當使用Feat.Set-1和Feat.Set-2等兩個特徵集進行kNN模型訓練時，我們發現在不同的k值下，獲得了不同的結果，如圖五所示。在使用Feat.Set-1特徵集進行kNN模型訊練時，當我們將k設置為1，我們可得到的最高辨識準確率為0.84。當我們使用Feat.Set-2特徵集進行訊練，且將k設置為3時，我們取得的最佳辨識準確率為0.72。

(a) (b)

圖五、不同k值下kNN模型的辨識準確度—(a)使用Feat.Set-1特徵集、(b)使用Feat.Set-2特徵集

(a) (b)

圖六、基於Feat.Set-1和Feat.Set-2特徵集進行辨識之kNN混淆矩陣，(a)使用Feat.Set-1、(b)使用Feat.Set-2

針對Feat.Set-1特徵集進行kNN模型訓練的結果，其所對應的混淆矩陣，如圖六所示。其中，Label0代表是易阻塞病患，Label1為不易阻塞病患，對Feat.Set-1特徵集在k設置為1時，取得了最高的辨識準確率(0.84)和F1-score (0.77)。另外，針對使用Feat.Set-2特徵集，且在k設置為3時，可取得最高的辨識準確率(0.72)和F1-score (0.63)。

除了混淆矩陣外，表八顯示了使用兩個不同特徵集所計算出的效能指標，例如：Accuracy、Sensitivity、Specificity、Precision和F1-score。這些指標可提供分類器的性能進行評估，以便了解kNN模型之效能表現。

表八、針對Feat.Set-1和Feat.Set-2特徵集之kNN效能指標

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Accuracy** | **Sensitivity** | **Specificity** | **Precision** | **F1-score** |
| **Feat.Set-1** | 0.84 | 0.77 | 0.87 | 0.77 | 0.77 |
| **Feat.Set-2** | 0.72 | 0.66 | 0.75 | 0.60 | 0.63 |

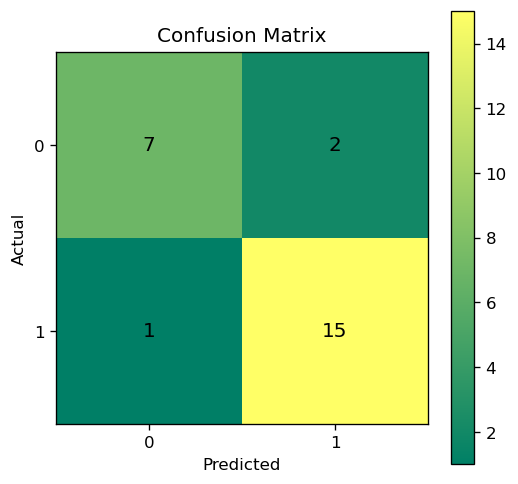
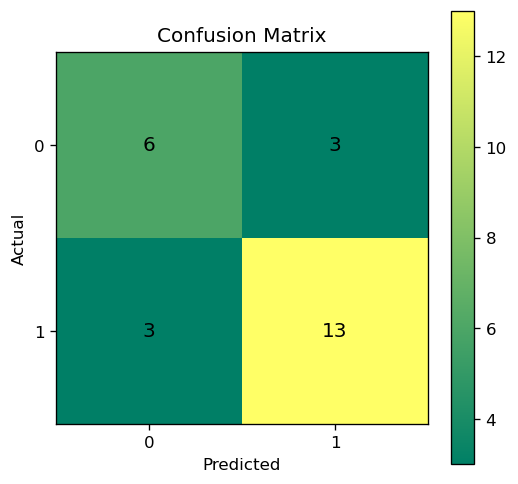
## SVM分析

核函數（kernel function）是SVM中非常重要的一個分類機制。SVM利用不同的核函數對數據進行非線性映射，非線性映射指得是將原始數據映射到高維空間中，從而將原本線性不可分的問題轉化為線性可分的問題。這種映射能夠減少機器學習的運算量，並避免維度災難（curse of dimensionality）。

當使用Feat.Set-1特徵集進行SVM模型訓練時，在使用Linear、Polynomial、Redius-based Function (RBF)等不同的Kernel Function下，其準確度如表八所示，使用Linear Kernel Function的辨識準確率最好，可達到0.88。圖九(a)為其SVM的混淆矩陣，其中Label0為易阻塞病患，Label1為不易阻塞病患。針對使用Feat.Set-2特徵集進行SVM模型訓練時，使用不同Kernel Function下的準確率，如表九中所示，而使用Polynomial Kernel Function的準確率最好，可達到0.76，圖九(b)為其混淆矩陣。另外，表十則展示了SVM在使用Feat.Set-1和Feat.Set-2特徵集時的各項效能指標。值得注意的是，在使用線性核函數對Feat.Set-1所進行的辨識之下，獲得了最好的辨識準確率(0.88)和F1-score (0.82)。

表九 使用Feat.Set-1和Feat.Set-2於SVM在不同核函數下之準確率

|  |  |  |  |
| --- | --- | --- | --- |
| **準確率** | **Linear** | **Polynomial** | **RBF** |
| **Feat.Set-1** | 0.88 | 0.80 | 0.64 |
| **Feat.Set-2** | 0.72 | 0.76 | 0.64 |

(a) (b)

圖九、基於(a) Feat.Set-1特徵集、(b) Feat.Set-2特徵集，進行辨識之SVM混淆矩陣

表十、針對Feat.Set-1和Feat.Set-2特徵集之SVM效能指標

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Accuracy** | **Sensitivity** | **Specificity** | **Precision** | **F1-score** |
| **Feat.Set-1** | 0.88 | 0.77 | 0.93 | 0.87 | 0.82 |
| **Feat.Set-2** | 0.76 | 0.66 | 0.81 | 0.66 | 0.66 |

## 討論

本研究的目標，是希望比較於兩手食指所採集的PPG訊號並進行探討，以辨識實驗受試者是屬於「瘻管易阻塞」或是「瘻管不易阻塞」的病患。這個研究目標設立的原因，是因為在血液透析病患的兩手食指部位採集PPG訊號，相對於在其他部位上進行訊號採集更為容易，利於臨床診斷或日常居家照護之應用。本研究的方法是採用統計分析t檢定技術，對表三至表五所羅列的自變參數做p值分析，找出具有顯著性優勢的訊號特徵。此時，這個t檢定的因變數是「瘻管易阻塞」和「瘻管不易阻塞」等兩種狀況。另外，我們再以監督式學習演算法進行智慧辨識，產生「瘻管易阻塞」或是「瘻管不易阻塞」病患的分類，提供病患於日常生活中的照護參考或是醫師於臨床門診中的醫療參考。癭管反覆阻塞也被認為是心血管疾病事件及死亡率的重要指標，因此提早辨識出這類病患並予適當介入十分重要。

在表三到表五中，我們發現只有表四的患側手之“PPG下降時間-AVG-A”這個自變數，對兩個因變數才具有統計檢測上的顯著性(p≦0.05)。其次在表四中，對於血液透析病患的健側手和患側手之PPGA-H/A、SSI-AVG-H、SSI-AVG-A、SSI-STD-H和SSI-STD-A等五個自變數，以及表五中的HH-P-area等自變數，它們同時對兩個因變數有統計檢測上的次佳顯著性(p≦0.06)。由此可見，如果有包含以上7個參數進行智慧辨識的話，將可能獲得比較好的辨識效能。

在本研究中，我們參考表三到表五中各個自變數的p值，並以p≦0.1和p≦0.35為條件，建構了Feat.Set-1和Feat.Set-2等兩個特徵集合，來進行智慧辨識。在使用Feat.Set-1為特徵集合時，以kNN和SVM進行分類實驗可以分別達到84%和88%的準確度，同時F1-score也可以分別達到77%和82%。然而，當使用Feat.Set-2為特徵集合時，kNN和SVM的辨識準確度則有很明顯的退化，分別是72%和76%。另外，它們的F1-score也是明顯地下滑，分別是63%和66%。在這裡，對於Feat.Set-2的辨識效能比Feat.Set-1退化的原因，我們推論最主要是因為在Feat.Set-2集合中，p值大於0.1且小於0.35的自變數，對於kNN和SVM的辨識準確度帶來很大的負面效應。因此，要提高對「瘻管易阻塞」或是「瘻管不易阻塞」的智慧辨識效果，我們所選取的自變數之p值不能太大。

另外，我們再針對Feat.Set-1特徵集合所包含的自變數之p值，再細分成p≦0.06和p≦0.1為條件，再次進行kNN和SVM的智慧辨識實驗，我們發現辨識準確度和F1-score等指標都相同。這說明了在Feat.Set-1特徵集合中，HH-AVG和HA-P-area這兩個自變數(p > 0.06)，對於提升智慧辨識的效能沒有帶來額外的幫助。

透過以上kNN和SVM之智慧辨識實驗結果，以及表三到表五中各個自變數的p值來觀察，我們發現在Feat.Set-1特徵集合中，可以看到以下三個事實。首先，兩組實驗參與者的患側手的“PPG下降時間-AVG-A”、“SSI-AVG-A”、“SSI-STD-A”等參數之參數值，在歐式空間上的分佈存在著明顯的資料叢聚特性，利於使用智慧辨識演算法來進行瘻管阻塞問題的分類。而且，這些參數對瘻管阻塞問題的兩個因變數，在統計分析上也具有相對比較顯著的差異性，有助於以t檢定對瘻管阻塞問題進行統計分類。其次，在Feat.Set-1中的“HH-AVG”、“SSI-AVG-H”、“SSI-STD-H”、“HH-P-area”等參數，呈現出兩組實驗參與者自其健側手所量測到的PPG訊號，再作進一步處理所衍生出的參數值，在瘻管阻塞問題上應用智慧辨識的結果具有很好的準確率。最後，以健側手的PPG訊號之衍生參數值對比患側手的同類參數值，例如：Feat.Set-1中的“PPGA-H/A”和“HA-P-area”等參數，我們發現針對瘻管阻塞問題，使用智慧辨識演算法可以獲得很好的辨識準確率。

本研究的主要研究貢獻有如下三點：第一，利用在雙手食指部位容易量測到的PPG訊號，以統計方法找出可有效辨識瘻管阻塞問題的訊號及其衍生特徵，並建立特徵集合以進行智慧辨識。第二，在表三到表五中，我們使用自相關/互相關運算，建立PPG訊號之相關性特徵(如表三所列)，以及PPG訊號相關性頻譜特徵(如表五所列)，再加上PPG訊號在形態學上的特徵(如表四所列)，在本研究中一共產生42個特徵參數。我們發掘出其中9個有智慧辨識價值的參數(如Feat.Set-1所示)，可對血液透析患者的瘻管阻塞問題進行智慧分類，獲得的研究成果可應用於提升醫療照護的品質。第三，透過kNN和SVM對Feat.Set-2的智慧辨識實驗，我們發現p > 0.1的特徵參數對於瘻管阻塞問題的智慧分類，沒有任何的助益。

# 結論

在本研究中，我們從不同的實驗參與者中，收集他們左/右手的PPG訊號，並依研究目標將其分為不易阻塞與易阻塞等兩類型研究對象。我們的研究目標，是通過處理和分析這兩類型研究對象的PPG訊號，對實驗參與者之瘻管阻塞問題作準確的辨識。

在收集到的PPG訊號上，我們進行了訊號前處理並建立了三類型獨立參數，並透過統計分析方法過濾出具有資料顯著性意義的獨立參數，組成可應用於監督式學習之智慧辨識的特徵向量(例如：Feat.Set-1特徵子集合)。透過kNN與SVM進行智慧分類，分類結果是使用Feat.Set-1特徵向量的SVM，獲得了最佳的準確度0.88。

儘管我們透過SVM可達到0.88的最佳準確率，成功區分了不易阻塞和易阻塞的病患，但我們認為目前成果仍存在進一步改進的空間。未來，我們將專注於預測瘻管阻塞的時間，特別是針對不易阻塞和易阻塞的病患區別。

# 參考文獻

1. Hsu CC, Liao CT, Lin MY, et al. Summary of the 2021 Annual Report on Kidney Disease in Taiwan. *Zhongzheng Rd, Zhonghe Dist*. 2022;36(4):886-888.

2. Ravani P, Palmer SC, Oliver MJ, et al. Associations between Hemodialysis Access Type and Clinical Outcomes. *Journal of the American Society of Nephrology*. 2013;24(3):465-473.

3. Ravani P, Gillespie BW, Quinn RR, et al. Temporal Risk Profile for Infectious and Noninfectious Complications of Hemodialysis Access. *Journal of the American Society of Nephrology*. 2013;24(10):1668-1677.

4. Anvari E, Vachharajani TJ. The Hemodialysis Access Surveillance Controversy Continues. *Kidney Int Rep*. 2020;5(11):1848-1850.

5. Kim HJ, Lee H, Kim DK, et al. Recurrent vascular access dysfunction as a novel marker of cardiovascular outcome and mortality in hemodialysis patients. *Am J Nephrol*. 2016;44(1):71-80.

6. Chiang, Chao, Tu, et al. Machine Learning Classification for Assessing the Degree of Stenosis and Blood Flow Volume at Arteriovenous Fistulas of Hemodialysis Patients Using a New Photoplethysmography Sensor Device. *Sensors*. 2019;19(15):3422.

7. Lok CE, Huber TS, Lee T, et al. KDOQI Clinical Practice Guideline for Vascular Access: 2019 Update. *American Journal of Kidney Diseases*. 2020;75(4):S1-S164.

8. Chao PCP, Chiang PY. Theoretical development with proper approximation and the corresponding clinical experiments for PPG sensor monitoring blood flow volume of hemodialysis patients with arteriovenous fistula. In: *2017 International Conference on Applied System Innovation (ICASI)*. IEEE; 2017:311-314.

9. Chiang PY, Chao PCP, Tarng DC, Yang CY. A Novel Wireless Photoplethysmography Blood-Flow Volume Sensor for Assessing Arteriovenous Fistula of Hemodialysis Patients. *IEEE Transactions on Industrial Electronics*. 2017;64(12):9626-9635.

10. Wu JX, Lin CH, Wu MJ, Li CM, Lim BY, Du YC. Bilateral photoplethysmography analysis for arteriovenous fistula dysfunction screening with fractional-order feature and cooperative game-based embedded detector. *Healthc Technol Lett*. 2015;2(3):64-69.

11. Du YC, Stephanus A. A novel classification technique of arteriovenous fistula stenosis evaluation using bilateral PPG analysis. *Micromachines (Basel)*. 2016;7(9).

12. Du YC, Stephanus A. Levenberg-Marquardt Neural Network Algorithm for Degree of Arteriovenous Fistula Stenosis Classification Using a Dual Optical Photoplethysmography Sensor. *Sensors*. 2018;18(7):2322.

13. Liang Y, Elgendi M, Chen Z, Ward R. An optimal filter for short photoplethysmogram signals. *Sci Data*. 2018;5(1):180076.

14. Kim TK. T test as a parametric statistic. *Korean J Anesthesiol*. 2015;68(6):540.

15. Zhang Z. Introduction to machine learning: k-nearest neighbors. *Ann Transl Med*. 2016;4(11):218-218.

16. Cover T, Hart P. Nearest neighbor pattern classification. *IEEE Trans Inf Theory*. 1967;13(1):21-27.

17. Short R, Fukunaga K. The optimal distance measure for nearest neighbor classification. *IEEE Trans Inf Theory*. 1981;27(5):622-627.

18. Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273-297.

19. Schölkopf B, Smola AJ. *Learning with Kernels*. The MIT Press; 2018.