

國立交通大學

電控工程研究所

碩士論文

一個使用脈衝傳遞時間及相位差與身體質量指數的
基於影像的低幀率連續血壓量測系統

An image-based and low frame rate continuous blood pressure
measurement system based on pulse transit time, phase difference, and
body mass index

研究 生：方宇凡

指 導 教 授：吳炳飛 教授

中 華 民 國 一 百 零 八 年 八 月

一個使用脈衝傳遞時間及相位差與身體質量指數的
基於影像的低幀率連續血壓量測系統

An image-based and low frame rate
continuous blood pressure measurement system based on
pulse transit time, phase difference, and body mass index

研究 生：方宇凡

Student : Yu-Fan Fang

指導 教授：吳炳飛

Advisor : Bing-Fei Wu

國 立 交 通 大 學
電 控 工 程 研 究 所
碩 士 論 文

A Thesis

Submitted to Institute of Electrical and Control Engineering

College of Electrical Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

In

Electrical and Control Engineering

August 2019

Hsinchu, Taiwan, Republic of China

中華民國一百零八年八月

一個使用脈衝傳遞時間及相位差與身體質量指數的基於影像的 低幀率連續血壓量測系統

學生：方宇凡

指導教授：吳炳飛 教授

國立交通大學電控工程研究所

摘要

自動與非接觸式的連續血壓量測近年來受到相當程度的關注。高血壓會導致冠狀動脈疾病、心臟疾病、短暫性腦缺血病變、中風、腎衰竭以及腦退化等相關疾病與併發症，並且血壓是生理健康的重要指標。因此，近年來開始有許多無袖套式以及進一步的非接觸式血壓量測演算法相關的研究。這當中，包含使用心電圖、光體積訊號描述儀的演算法，以及以純影像與訊號處理為基礎的影像光體積訊號之演算法。血壓量測可以透過基於血液流動模型的脈衝傳遞時間特徵來達成，因為該傳遞時間與血壓有反向關係。透過訊號處理技術，可獲得本論文所需要之心跳脈衝訊號波形。然而近期其他研究使用的主要仍然是基於光體積描述儀的血壓量測方法，此接觸式的方法會導致皮膚較為敏感或脆弱的新生兒與燒燙傷病患量測時的困難。本論文提出之基於影像式光體積描述訊號的血壓量測演算法可以改善此問題。透過色度訊號以及皮膚正交平面等演算法結合有限脈衝響應帶通濾波器，本論文得以獲得雜訊受抑制的脈衝訊號以量測血壓。

本論文使用之資料庫為透過多名不同受測者長期蒐集而得，包含使用者之臉手影像以及對應的收縮壓以及舒張壓等資訊，以 90 幀/秒的幀率進行錄製，本論文改善了特徵擷取的方式以避免更高幀率的設備需求。而由於不同受測者會因為血管順應性以及血管彈性係數的不同而有不同的血壓偏移量，因此許多論文會使用一台血壓機每隔一段時間進行校正以補足此偏差。為了達成純影像式的血壓量測，本論文提出了一種基於身體質量指數的血壓偏移量量測方法。使用受測者臉部的影像特徵，並藉由卷積神經網路迴歸預測出其對應的身體質量指數，用做血壓偏移量的估測以去除使用額外的血壓機校正的需求。此外，由於基於臉手的影像式血壓量測仍有其限制，例如若用於車輛駕駛之血壓量測時，駕駛無法將手空出來進行血壓量測，故血壓量測系統只能獲得其臉部影像。因此，本論文亦提出了一種全臉影像式的基於頻域相位差特徵的演算法用以量測血壓。

An image-based and low frame rate continuous blood pressure measurement system based on pulse transit time, phase difference, and body mass index

student : Yu-Fan Fang

Advisor : Prof. Bing-Fei Wu

**Institute of Electrical and Control Engineering
National Chiao Tung University**

ABSTRACT

Automatic and non-contact continuous blood pressure measurement has attracted great attention recently. Hypertension leads to coronary artery, heart failures, transient ischemic attack, stroke, kidney failure, mild cognitive impairment, related disease, and complications. Blood pressure is an important indicator for health. Thus, many research focused on cuffless or even image-based blood pressure measurement these years. Electrocardiograph and photoplethysmography (PPG) are popular contact-based method, some image-based methods with high frame rates are also attempted. With pulse transit time (PTT) features based on vascular flow model, blood pressure measurement is made available due to the inverse relationship between PTT and blood pressure. With signal processing techniques, this thesis retrieves the remote PPG for blood pressure estimation, which deals with the difficulty that neonates or patients with empyrosis may face during contact-based measurement.

The thesis collected the dataset from different subjects in a long term basis, including images of face and palm recorded with camera at a rate of 90 frames/second and the corresponding systolic and diastolic blood pressure ground truth. To notice, due to the intrinsic difference of blood vessel compliance and elastic modulus, different subjects will have different blood pressure offsets. The thesis proposed an algorithm to deal with such offsets based on body mass index (BMI) so as to eliminate the need for extra calibration equipment and develop a system based on image only. Moreover, with transfer learning model based on Resnet-50 convolutional neural network, BMI can be estimated from facial images. However, blood pressure measurement based on face and palm is still limited in some cases, such as measurement for automobile drivers, who cannot put their hands in front of the camera. Hence, the thesis also proposed a frequency-based algorithm using phase difference features from forehead and cheek for face-only blood pressure measurement.

誌謝

在擔任研究生的這兩年研究生涯，得到許多人的幫助而得以完成此篇論文的研究。首先，需要感謝指導教授吳炳飛的諄諄教誨，以及其對於論文研究與實作上提供的寶貴經驗與建議。不厭其煩地給予許多叮嚀與指導，也在每次開會時提供非常明確的方向。更感謝老師提供良好的研究設備與學習環境，老師的學者風範讓學生永銘於心。

其次，要感謝黃柏維學長對於研究方向的許多建議。並且時常願意共同討論相關的訊號處理與神經網路訓練等演算法與模型；從不吝於分享如何編寫高效率的程式碼，也提供學弟妹相當良好的經驗與技術協助；對於整體資料庫的錄製規畫也付出了相當多的辛勞共同籌備。同時也要感謝陳冠宏學長，每次的開會都能從企業面以及實做的角度給予許多非常實質的建議與幫助，而且有相當全面的領域專業知識能夠為研究的方向指引一條明確的道路；學長對於企業以及實際應用面的了解相當多方位，不吝於分享經驗，亦提供了學弟妹非常良好的學習資源，冠宏學長的指導以及進度規畫對於研究有莫大幫助。

再者，非常需要感謝的是鍾孟良學長，為了實驗室的設備以及資源付出了莫大的辛勞，相當感謝您對於整個研究的支持與協助；學長亦提供了許多比賽的經驗技術分享以及上台簡報時的專業風範；對硬體設備的深入了解，使研究時需要的硬體資源都能順利獲得，讓整體的研究沒有後顧之憂，孟良學長的支持對於此篇論文的研究有著莫大的幫助；同時學長也相當關心學弟妹的狀況以及研究狀態，是不可多得的好學長。

接著，要感謝測試工程師謝豫芳以及劉容蓁協助資料庫的錄製，讓血壓整個龐大資料庫的錄製能順利進行，兩位測試工程師時刻確保實驗環境的穩定並謹慎的錄製資料庫，為血壓資料庫以及血壓的相關研究提供了莫大的支持，使得血壓的研究能夠順利進行；同時，也相當感謝兩位於資料庫錄製時期，每日辛苦的一起陪同搬移錄製設備，在此表達相當大的感謝之意。再者，在此感謝實驗室的所有同窗以及研究夥伴，在遇到實驗壓力以及挑戰時，共同分擔一起解決；相當喜歡實驗室同儕之間的熱情互助以及相互交流、共同努力的氛圍。

在此亦要感謝口試委員。對於此篇論文的指教以及給予明確的研究相關的建議，受益良多。在口試過程當中獲得了相當寶貴的經驗以及與研究相關的啟發與引導。

最後要感謝母親與家人的一路的支持以及養育，有了家庭作為後盾，便能夠放心的著手進行研究，並且順利的完成此篇論文，達成人生求學階段相當重大的理想與目

標。在此非常感謝所有上述提到的研究夥伴、朋友以及家人，有了他們的協助，此篇論文才能順利的完成。如若此篇論文有幸能夠對社會產生微薄的貢獻，請將榮耀歸給他們。

目錄

中文摘要.....	I
英文摘要.....	II
誌謝.....	III
目錄.....	V
圖目錄.....	VII
表目錄.....	IX
第一章、導論.....	1
1.1、背景與研究動機.....	1
1.2、相關歷史文獻.....	4
1.3、論文貢獻.....	10
1.4、論文架構說明.....	12
第二章、臉手式血壓量測系統.....	13
2.1、血液循環模型與血壓.....	13
2.2、感興趣區域(Region of interest)的選擇.....	15
2.3、訊號提取與抗雜訊演算法.....	17
2.4、訊號峰谷值偵測演算法及特徵選取與擷取.....	22
2.5、血壓資料庫及基於脈衝傳遞時間特徵量測血壓的機器學習模型.....	27
2.5.1、K 最近鄰居演算法模型.....	29
2.5.2、深度信念網路-深度神經網路(DBN-DNN).....	30
2.5.3、人工神經網路(ANN).....	32
2.6、臉手式血壓實驗結果以及與相關論文之比較.....	33
2.6.1、K 最近鄰居法(KNN)之血壓方均根誤差.....	33
2.6.2、深度信念網路-深度神經網路(DBN-DNN)之血壓方均根誤差.....	33
2.6.3、人工神經網路(ANN)之血壓方均根誤差.....	33
2.6.4、與相關論文[2][31]之比較.....	33
2.6.5、基於本論文實作之血壓量測系統結合身體質量指數的誤差表現.....	37
2.7、臉手影像式血壓模型之總結與討論.....	40
2.8、未來展望.....	42
第三章、臉部影像身體質量指數量測及血壓偏移量校正.....	43
3.1、背景.....	43
3.2、基於人臉特徵之身體質量指數量測模型及使用之資料庫.....	44
3.2.1、人臉影像特徵擷取.....	44
3.2.2、線性迴歸(LR).....	46
3.2.3、支援向量迴歸(SVR).....	46
3.2.4、高斯過程迴歸(GPR).....	47
3.3、基於 Resnet-50 卷積神經網路的遷移式學習模型及資料庫擴增法.....	48
3.4、身體質量指數量測模型之表現與相關論文比較.....	53

3.5、人臉影像身體質量指數量測系統及血壓偏移量之總結與討論.....	55
3.6、未來展望.....	56
第四章、全臉影像式頻域血壓量測模型.....	57
4.1、研究動機與介紹.....	57
4.2、影像光體積訊號處理以及頻率域特徵擷取.....	59
4.3、訊號擷取區域之濁海偵測與警示機制.....	61
4.4、全臉影像式血壓量測資料庫.....	65
4.5、全臉影像式血壓量測演算法與模型.....	66
4.5.1、線性迴歸法(Linear regression).....	66
4.5.2、K 最近鄰居法(KNN).....	66
4.6、全臉影像式血壓實驗結果以及與相關專利技術之比較.....	68
4.6.1、線性迴歸模型(LR)之血壓方均根誤差.....	68
4.6.2、K 最近鄰居法(KNN)之血壓方均根誤差.....	68
4.6.3、與相關專利技術之比較.....	68
4.6.4、基於本論文實作之全臉影像式血壓量測系統的表現.....	71
4.7、全臉影像式血壓模型之總結與討論.....	75
4.8、未來展望.....	77
第五章、總結與未來展望.....	78
5.1、臉手影像式血壓量測演算法之總結.....	78
5.2、Resnet-50 與臉部影像量測身體質量指數進行血壓校正之總結.....	79
5.3、全臉影像式血壓量測演算法之總結.....	80
5.4、未來展望.....	81
參考文獻.....	82

圖 目 錄

圖 1、接觸式血壓量測所用生理訊號之波形.....	1
圖 2、接觸式光體積描述訊號量測儀器.....	2
圖 3、影像式光體積訊號量測原理(漫反射模型)[32].....	4
圖 4、尤拉放大法強化心跳脈波訊號.....	5
圖 5、EX-FH20 420 帖/秒相機.....	6
圖 6、Morph II 人臉影像資料庫.....	8
圖 7、臉手影像式血壓的訊號擷取區域.....	15
圖 8、Logitech Brio 攝影機.....	17
圖 9、未經過帶通濾波器的影像光體積訊號.....	20
圖 10、經過帶通濾波器後的影像光體積訊號.....	20
圖 11、使用動態門檻擷取脈波訊號之峰谷值.....	22
圖 12、影像式光體積訊號之脈衝傳遞時間之定義.....	24
圖 13、傳統的脈衝傳遞時間定義.....	24
圖 14、峰值強度之 25% 與 75% 的特徵點.....	25
圖 15、臉手影像式血壓簡要流程圖.....	26
圖 16、血壓資料庫錄製環境.....	27
圖 17、歐姆龍血壓計.....	28
圖 18、臉手影像式血壓完整流程圖.....	40
圖 19、用以量測身體質量指數的 7 個臉部特徵.....	44
圖 20、Dlib 的 68 個臉部特徵點的模型.....	45
圖 21、Viola Jones 的人臉偵測功能.....	48
圖 22、運用仿射變換進行資料擴增.....	48
圖 23、運用隨機抹除技術進行資料擴增.....	50
圖 24、加入身體質量指數的臉手影像式血壓量測流程圖.....	52
圖 25、全臉影像式血壓之訊號擷取區域.....	57
圖 26、血壓資料庫的資料錄製流程.....	65
圖 27、全臉影像式血壓的流程圖.....	75
圖 28、臉手影像式血壓量測介面.....	40

圖 29、全臉影像式血壓量測介面.....	76
圖 30、異常狀態下之臉手影像式血壓量測介面.....	41
圖 31、影像式血壓資料庫錄製時對應之幀率記錄檔.....	29
圖 32、臉部影像式脈波檢測技術[100].....	69
圖 33、濁海阻擋樣例以及相關的偵測機制.....	61
圖 34、Morph II 資料庫之人臉影像額頭區域標記.....	61
圖 35、Morph II 資料庫之人臉影像遮擋問題.....	49
圖 36、臉部計算脈衝傳遞時間相關專利[107].....	70

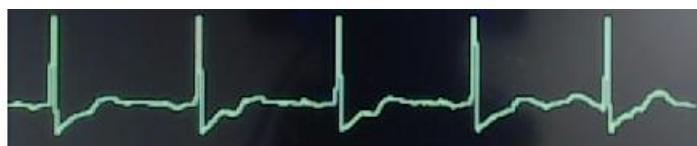
表 目 錄

表 1、臉手血壓所使用之特徵.....	26
表 2、身體質量指數模型與相關論文之比較.....	53
表 3、影像式血壓資料庫之錄製規格表.....	28
表 4、Likert 疲勞狀況量表.....	42
表 5、Jeong 所提出之臉手影像式血壓[2]論文的資料庫資訊.....	34
表 6、Sean 所提出之臉手影像式血壓[31]論文的資料庫資訊.....	35
表 7、臉手影像式血壓與相關論文之比較.....	36
表 8、本論文之全臉影像式血壓模型量測結果.....	71
表 9、結合身體質量指數特徵之臉手影像式血壓系統的實際量測結果.....	39
表 10、結合身體質量指數特徵之全臉影像式血壓系統的實際量測結果.....	72

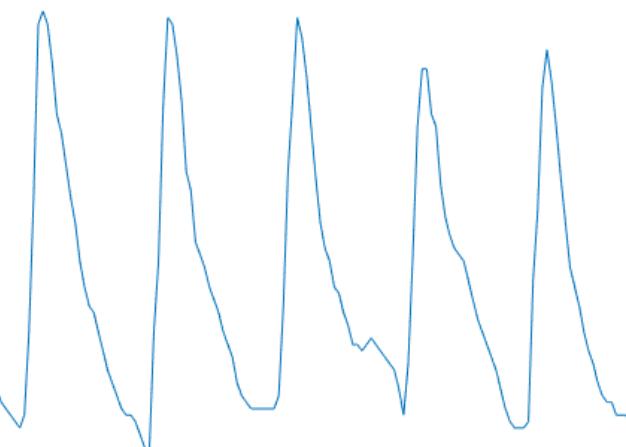
第一章、導論

1.1、背景與研究動機

血壓是相當重要的生理健康指標，因此血壓的即時量測與監控一直相當受到關注。相關的血壓量測法舉凡使用震盪法的袖套式血壓機、以及非袖套式的接觸式光體積描述儀法[1]量測血壓，及本論文所使用的影像式光體積描述儀訊號[2]，皆是當前血壓研究與量測的重要方法。高血壓會造成中風[3]、冠狀動脈疾病[4]、短暫性腦缺血病變[5]、心臟疾病[6]、腎衰竭[7]以及腦退化[8]等相關疾病與併發症。袖套式血壓[9]的主要缺點在於其利用血壓氣囊壓縮動脈阻絕血液流動，並慢慢洩壓，過程中由壓力感測器量測出受測者的收縮壓與舒張壓，因此整個血壓量測過程約需 35~40 秒的時間才能輸出一組離散的血壓數值[9]，因而無法支援連續性的血壓量測。此外，震盪法之血壓量測須確保受測者的肘部和心臟在同一個水平線，且其衣服之厚度不可超過 0.5 公分。許多量測的細節與技術使得震盪法之血壓量測應由專業之醫護人員為之，此亦為震盪法血壓量測的缺點。再者，袖套式血壓由於會對受測者的上臂進行加壓，因此會造成受測者一定程度的不適[10]，也會造成血流不順的問題。而非袖套式的血壓量測由於需要計算受測者的脈衝傳遞時間特徵用做血壓預測[11]，因此會使用心電圖與光體積描述儀的訊號[12]來進行血壓的量測。心電圖如圖 1(a)是由 P、Q、R、S 和 T 波所構成，其中 P 波對應的是心房去極化的過程，意即心房收縮之前進行去極化所引起



(a)心電圖的 P、Q、R、S 和 T 波的週期(來源為示波器)



(b)光體積描述儀的訊號

圖 1、接觸式血壓量測儀器所使用之訊號

的電位差，而 QRS 波則對應到心室的去極化，也就是心室開始收縮的時間。T 波則對應到心室的再極化，也就是心肌細胞極化後，將會有一段時間不接受任何刺激，並進行再極化的過程。其中最重要的便是 R 波，也是 QRS 波群當中的第一個正向波。心電圖當中的 RR 間距對應到心跳的週期，因此 R 波常常被用來配合光體積描述儀訊號以計算脈衝傳遞時間[12]，R 波的波峰則是圖 1(a)當中波形出現最高點的位置。

而光體積描述儀訊號如圖 1(b)則是利用手指夾式的脈搏血氧儀來量測脈搏的訊號。如本篇論文便是使用 CMS-50E 血氧儀[98]如圖 2 來量測光體積描述儀訊號。透過脈衝傳遞時間特徵包含心電圖訊號的 R 波波峰到光體積描述儀脈波訊號的週期低點、

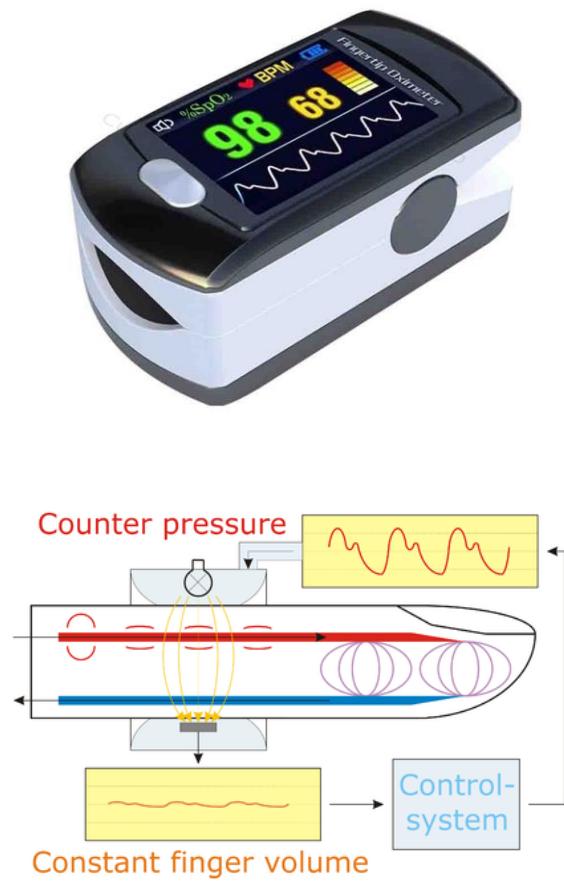


圖 2、CMS-50E 手指夾式脈波血氧儀[98]

以及一階導數的最大值點和週期高點等特徵[12]進行血壓的量測。然而使用手指夾式接觸式血氧儀進行量測也有其明顯缺點，由於需要接觸到受測者的肌膚，因此對於新生兒以及皮膚燒燙傷的患者[2]的血壓量測有其應用困境。

因此，本篇論文使用影像式光體積描述儀[2]訊號進行血壓的量測。影像式光體積描述儀脈波訊號可以透過綠色通道訊號法[18]、色度訊號法[14]、皮膚正交平面法[15]或是抗晃動影像式光體積訊號法[16]來獲得。色度訊號法和皮膚正交平面法通常用於較為穩定的環境，而抗晃動光體積描述儀法則多用於有雜訊以及晃動的環境。近年許多的論文皆著重於提升心跳脈波訊號的品質以及抗雜訊的能力。這些研究也提供了本論文能擷取較為乾淨的脈波訊號的技術基礎。結合帶通濾波器後，此論文得以獲得品質足夠且可以用以量測血壓的影像式光體積描述訊號，並且把心跳範圍以外的高頻雜訊以及基準線飄移的直流偏移量去除[17]。此外，由於每個受測者的血管順應性以及血管彈性係數不同，所以每個人對應的血壓偏移量也會不同。因此許多論文需要在以演算法量測血壓前，預先以符合醫療標準的血壓機針對不同的受測者進行血壓量測的校正[19]。本論文為了避免額外血壓機設備的需求，因此使用身體質量指數和血壓的正向關係[20]來進行血壓的量測。透過遷移式學習[21]以及 Resnet-50 卷積神經網路[22]，本論文得以利用受測者的臉部影像進行身體質量指數的迴歸預測並利用此結果進行血壓偏移量的量測。結合受測者的臉手影像算出的脈衝傳遞時間特徵，本論文得以算出受測者的收縮壓以及舒張壓。

然而，利用臉手影像進行血壓量測仍有其侷限性，因為在許多情況下，受測者無法將手置於攝影機前方進行量測，例如當汽車駕駛欲使用臉手血壓量測系統量測血壓時，駕駛將無法空出手進行量測。因此，本論文亦提出了一種全臉式的基於頻域特徵以及相位差的影像式血壓量測演算法進行血壓量測，期許日後將此系統與車用駕駛生理資訊系統結合以進行即時性的血壓量測，保護駕駛與用路人的安全。同時，此系統亦可用於日常家居的高齡照護以及高血壓患者的實時血壓監控。

1.2、相關歷史文獻

袖套式的血壓計無法準確量測受測者於運動狀態下或甫結束運動時的血壓，因為受測者上臂會因為運動而有所起伏[36]。而除了袖套式血壓外，傳統而言，多使用接觸式的方式進行血壓的量測[12]。例如使用胸導程心電圖的電極貼片貼於受測者的胸壁，由於接近受測者心臟的位置，因此胸導程可以從水平面觀察心臟的活動，也因此可以獲得對應的 QRS 波群的 R 波波峰，R 波波峰可以用作 RR 間距的計算[13]，也可以用來配合光體積描述儀訊號計算出脈衝傳遞時間特徵用作血壓的回歸[12]預測。心電圖導程的技術是利用貼在胸部位置的電極貼片偵測心肌的電位變化以獲得心電圖訊號。而光體積描述儀運用的原理則是在環境光進入受測者皮膚時，皮膚當中的血紅素會吸收一部分光線，而其餘光線則會反射由光體積描述儀接收，結合接收到的光，得以獲得受測者的脈波訊號如圖 3[23][24][25][30]。由於胸導程貼片的方便性以及手指夾式光體積描述儀的便攜性，使得此種血壓量測的方式成為許多血壓相關學術研究用以擷取脈衝傳遞時間特徵的方式。這些論文如[23][24][25]便是使用心電圖以及光體積描述儀訊號，利用 R 波波峰以及光體積描述儀的脈衝訊號的波谷、波峰及最大斜率點發生的時間差以取得脈衝傳遞時間特徵，進行血壓的預測。而亦有如同[30]，並非使用脈衝傳遞時間等相關特徵的研究，而是以訊號的 Kurtosis、歪斜度等波形指標利用人工神經網路迴歸進行血壓的預測。

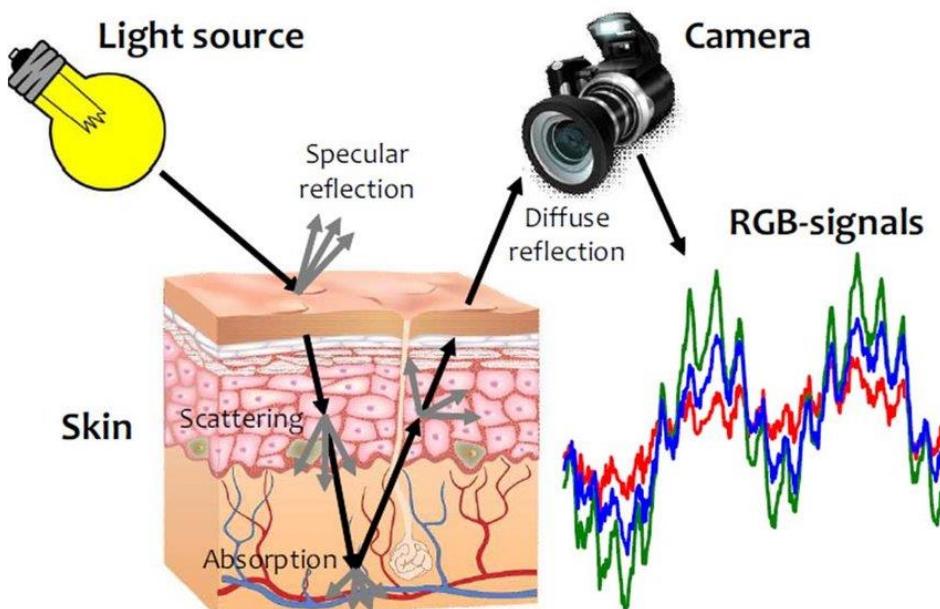


圖 3、利用漫反射光量測影像式脈衝訊號的原理 [32]

然而，此種量測仍有其侷限性，特別是心電圖，受測者進行訊號量測時必須貼上電極貼片 會產生一定程度的不適，心電圖電極貼片導程及方向貼法的專業性，也造成一般居家照護不容易操作。因此，許多論文亦僅使用光體積描述儀訊號進行脈衝傳遞時間特徵的擷取[26][27][28][29][30]。其中，最早的便是 Jones 於 1992 年首先提出的[26]，並且申請了美國專利。而以光體積描述儀為訊號擷取特徵的血壓量測演算法亦分為 2 個發展方向，大部分是使用人體的 2 個位置之脈波訊號計算脈衝傳遞時間[27][29]；而也有如[30]利用單一位置的光體積描述儀訊號進行血壓量測的演算法，利用單一位置的訊號結合血流理論估計整體的血壓特徵，此法的優勢是可以透過只量測 1 個位置的光體積描述儀訊號進行血壓的量測；而缺點則是如此一來將無法獲得精確的脈衝傳遞時間[11]，因為許多論文所定義與研究的脈衝傳遞時間特徵皆是來自受測者 2 個部位的脈波特徵的時間差[11][26][27][28][29]。

儘管只以光體積描述儀進行訊號的量測可以大幅降低量測過程的繁瑣性，此種方法仍必須透過量測儀器和使用者的皮膚進行接觸[27]，會造成受測者一定程度的不適。而且對於新生兒、皮膚病患以及燒燙傷病患而言，此種接觸式的血壓量測較不可行[2]。因此，近年亦有研究提出基於影像式的光體積描述儀訊號血壓量測演算法[2][31]。透過攝影機進行影像的擷取以便利用抗雜訊與抗晃動的訊號擷取演算法[14][15][16]獲得影像式光體積脈衝訊號用作脈衝傳遞時間的特徵擷取，其中 De Haan 提出的色度訊號(Chrominance trace)[14]以及 Wenjin Wang 提出的皮膚正交平面法(Plane orthogonal to skin)[16]是利用三個顏色通道的線性組合降低雜訊的方式，也是本論文選擇使用的訊號擷取演算法。



圖 4、尤拉放大法(EVM)處理影像式光體積訊號[33]

有了乾淨的訊號後，許多論文開始研究影像式光體積訊號能進一步改善的方法。包含了 MIT 首先提出可使用尤拉放大法(Euler video magnification)[33]，對心跳頻帶範圍的影像訊號能進行放大如圖 4，將影像的綠色通道的訊號能變化強化進而達到訊號能的改善，Tasli 亦使用演算法改善影像光體積描述儀訊號能用作生理資訊監控[34]。然而，Wenjin Wang 提到尤拉放大法固然強化了心跳頻帶範圍的訊號能，雜訊卻同時也被強化了[35]，無法有效的抵抗雜訊並提升訊號能品質。因此，基於色度訊號能或皮膚正交平面的訊號能擷取技術配合帶通濾波器仍是較多論文所採用的演算法。

而針對血壓相關特徵的研究，起於 1981 年 La Geddes 首先提出使用脈波傳播速度(pulse wave velocity)和血壓之間的相關性預測血壓的模型，此研究奠定了以脈波傳播速度的倒數，即脈衝傳遞時間量測血壓的基礎。而 Smith 則於 1999 年針對脈波傳播速度和血壓的關係進行進一步的研究，並提出血管壓力造成的脈波的會在樹狀的血管網路進行傳播，並且實驗發現其傳播速度和實際的血壓有相當高的相關性[37]，此理論也進一步引領日後[38][39]提出了脈波傳遞速度與血壓相關公式的模型。此也奠定了使用影像式光體積訊號能量測血壓的理論基礎[2][31][40]。值得說明的是，此處影像式光體積訊號能儀所觀察到的時域特徵實際為脈衝到達時間，而非脈衝傳遞時間。兩者的差別為脈衝到達時間額外包含了心臟周期當中的射血前期(pre-ejection period)。然而由於脈衝傳遞時間只能以心震儀(seismocardiography)進行量測，且脈衝傳遞時間及脈衝到達時間相當接近，因此當今影像式血壓研究的相關論文皆以脈衝到達時間取代脈衝傳遞時間，並沿用脈衝傳遞時間之名詞，本文中之脈衝傳遞時間，亦即脈衝到達時間[41]。



圖 5、EX-FH20 420 帧/秒相機

Jeong 於 2016 年首先提出了使用影像式光體積描述儀訊號進行血壓量測的技術 [2]，此篇論文使用高幀率的攝影機，即日本卡西歐公司生產的 EX-FH20 如圖 5，以 420 幀/秒的幀率進行臉部及手部影像光體積描述儀訊號的擷取，並結合線性回歸模型以用於血壓的量測。此論文使用的是臉部訊號最大導數點(FdBV)到右手訊號最大導數點(HdBV)之間的時間差作為脈衝傳遞時間的特徵用於血壓的預測，針對 7 位受測者建立”個人化”的模型。此篇使用 420 幀/秒進行血壓量測的論文的 7 位受測者的舒張壓變動範圍為 9.8 ± 3.7 毫米汞柱，相對來說變化不大，在此情況下，影像脈衝傳遞時間特徵和舒張壓的相關性亦僅有 0.361 ± 0.254 ；而收縮壓的變動範圍較大為 24.4 ± 10.7 毫米汞柱，而相關性為 -0.8 ± 0.11 [2]，此論文僅針對收縮壓進行評比。且此論文並無直接提供收縮壓之方均根誤差，而是以脈衝傳遞時間與血壓的相關性作為演算法評估的基準。此外，此論文較大的限制便是其個人化的模型，實用上須先收集該受測者以血壓機量測出的真實血壓建立模型才能為其進行預測[2]，無法實際達成針對非資料庫當中的不同使用者直接進行影像式血壓量測的功能。且此論文於應用面最大的挑戰為其使用的幀率過高，為 420 幀/秒，無法用來進行血壓的即時量測，需先進行影像錄製才能離線結合演算法進行血壓量測，對設備的規格要求相對也高許多，高速攝影機造成的記憶體負擔較大，且此論文基於影像量測血壓時也需要額外使用其他血壓儀器進行校正[2]。更低幀率且無須額外校正的非個人化的血壓量測模型有其必要性。

因此，2017 年 Sean 提出了以 5 個脈衝傳遞時間相關特徵的基於影像式血壓量測的模型[31]。此論文使用較低的 75 幀/秒的幀率結合更多的脈衝傳遞時間相關特徵並以 K 最近鄰居法進行血壓的預測。該論文使用的是其自行錄製的資料庫，而其限制在於收集的受測者之收縮壓範圍為 84~135 毫米汞柱，缺乏較高血壓(140 毫米汞柱以上)的受測者樣本，且受測者人數亦偏少，唯以 13 位受測者進行血壓量測。且針對血壓為 84~135 毫米汞柱的資料庫範圍以 K 最近鄰居演算法算出收縮壓的方均根誤差為 16.50 毫米汞柱。以可預測範圍為 84~135 毫米汞柱而言，此誤差仍過高。此外，此篇論文使用之加護病房多參數智慧監測資料庫(MIMIC-II)當中部分訊號有錯誤紀錄，即心電圖與光體積描述訊號和受測者對應的血壓有資料配對錯誤問題[42]。而此論文並無注意到此問題，仍使用了 MIMIC-II 資料庫協助模型的建立，因此亦會受到上述提及的資料庫配對錯誤問題影響。

而由於不同血壓受測者之血壓會受到血管順應性以及血管彈性係數的影響，因此會有不同的血壓偏移量[19]，許多論文以校正的方式去除此偏移值[2][19]，然而如此便

需要額外的血壓儀器針對演算法及影像量測系統進行頻繁的校正。本論文為除去此需求，因此採用身體質量指數和血壓之間的相關性進行血壓偏移量的量測，彌補需要額外血壓校正的需求。本論文結合了透過臉部影像量測身體質量指數的神經網路模型，進以用作血壓偏移量的量測。因此，以下亦列舉了臉部影像量測身體質量指數的相關研究[43][44]。



圖 6、Morph II 縱向數據庫

以臉部影像量測身體質量指數相關的研究始於 2013 年 Lingyun Wen 利用臉部特徵點擷取出人臉的 7 個特徵包含顴寬額寬比、臉寬臉高比、周長面積比、眼睛大小、下臉臉高比、臉寬下臉高比、平均眉眼距進行身體質量指數的量測[43]。使用了包含線性迴歸、支援向量迴歸以及高斯程序回歸等較為傳統的機器學習模型。而最終使用支援向量法得出身體質量指數的平均絕對誤差為 3.14 公斤/公尺^2 。值得一提的是，此論文用作建立模型的是公開資料庫 Morph II 縱向數據庫學術版[44]，包含收集自約 13000 人的總共 55000 張的人臉影像(圖 6)以及對應的身高體重資訊，和本論文所使用的是同一個數據庫。

而 2018 年則有 Antitzza 提出了基於卷積神經網路的透過臉部影像量測身體質量指數的模型[45]。使用的資料庫為其自行自網路使用爬蟲技術下載的名人資料庫，包含來自 513 位男性及 513 位女性的共 1026 張臉部影像。此論文使用大小為 256x256 的影像結合卷積神經網路進行量測，是首先使用深度學習的方式透過人臉影像來進行身體質量指數的量測的論文。

而全臉影像式血壓量測目前則無相關的論文，此篇論文首先嘗試使用頻域相位差特徵相關的演算法進行血壓的量測。

1.3、論文貢獻

1. 大部分非袖套式量測血壓的研究皆是用接觸式光體積描述儀進行，而基於純影像式的血壓研究則相對少了許多。且目前基於純影像式的血壓研究主要使用的是相當高的幀率(420 幀/秒)[2]的攝影機，並且採用個人化血壓模型，無法適用於每一位受測者。且沒辦法處理每個人血管順應性與彈性係數造成血壓偏移量不同的問題，仍需要額外的血壓設備進行校正。本論文則是以通用的血壓模型為基礎且結合身體質量指數進行血壓的校正以抵銷血壓偏移量的影響，且本論文使用的是幀率為 90 幀/秒的攝影機進行純影像式血壓量測，設備的規格及運算需求較低。通用化的模型以及不需要額外標準袖套式血壓量測設備校正的特性使得此血壓量測系統之實用性提升，可進一步達成長期健康照護的功能。
2. 本論文使用的資料庫蒐集自 27 位受測者的共 1033 筆錄製，每筆錄製包含 45 秒的受測者影像以及對應的實際收縮壓與舒張壓，是近期血壓論文研究使用的資料庫當中較大型的。如此一來，較不會有可預測的血壓範圍小的問題，本論文使用之資料庫當中最高的受測者收縮壓達到 166 毫米汞柱，較[31]的最高血壓高了 31 毫米汞柱；而最高的舒張壓為 105 毫米汞柱，也較[31]高了 14 毫米汞柱。而最低的收縮壓為 83 毫米汞柱，也較[31]低了 1 毫米汞柱；最低的舒張壓為 48 毫米汞柱，較[31]低了 9 毫米汞柱。使用此更全面的資料庫使得此模型的通用性更佳，可預測的血壓範圍也較大。
3. 許多研究仍較著重在理論面，如使用較高幀率的[2]，高幀率的需求使得對設備以及運算需求的成本較高，實用與商品化上成本相對大了許多。本論文使用了較低的幀率，且以身體質量校正的方式抵銷血壓偏移量的影響，不需要額外的血壓設備。僅需要以單一一個支援 90 幀/秒幀率的攝影機便可以透過影像式光體積訊號擷取脈衝傳遞時間特徵，配合身體質量指數進行血壓的量測。不僅是離線的研究，亦可用以進行即時血壓量測系統的實作。
4. 本論文針對用作臉部影像量測身體質量指數的 Morph II 縱向數據庫[44]進行了資料擴增，包含將原影像進行水平、垂直方向的平移、旋轉以及翻轉等仿射變換。並且結合了[46]的影像隨機抹除技術(random erasing)進行資料擴增以強化對抗雜訊的能力。如此一來，所訓練而得的模型對於過擬合的抗性亦可提升。並且，本論文使用了基於 Resnet-50 卷積神經網路的遷移式學習模型，並將最後的全連接層分類器變

更為迴歸層以求得受測者對應的身體質量指數，較使用傳統機器學習模型的[43]有更小的誤差。

5. 本論文首先提出了僅以受測者臉部的影像進行血壓量測的方式。如此一來，針對許多應用諸如車用駕駛的生理監測，本論文實作之系統皆有其實用性。受測者不必如[2][31]將手掌空出來，便得以進行血壓的量測。本論文透過頻域相關的特徵，得以克服時域上臉部的額頭與臉頰間的脈衝傳遞時間過短的問題。

1.4、論文架構說明

1. 第一章：導論

說明本論文的研究動機與目的，並且進行相關研究與文獻的回顧，以及提出本論文的主要貢獻。

2. 第二章：臉手式血壓量測系統

說明臉手式血壓量測的前置訊號處理機制、特徵選取演算法、量測模型以及所使用的資料庫，並針對所使用的模型與相關論文[2][31]進行結果比較。

3. 第三章：臉部影像身體質量指數量測及血壓偏移量校正

說明身體質量指數和血壓偏移量之間的關係，並透過臉部影像量測身體質量指數。介紹所使用的特徵與機器學習演算法，以及另一種基於深度學習與遷移式學習的卷積神經網路模型。並且介紹所使用的資料庫，以及詳細說明本論文所採用的包含仿射變換(affine transformation)以及影像隨機抹除(random erasing)等資料擴增技術。並將身體質量指數量測的結果和相關研究[43][45]進行比較。

4. 第四章：全臉影像式頻域血壓量測模型

說明全臉式血壓量測系統所使用的訊號、頻域相關的相位差特徵、基於機器學習的量測模型以及所使用的資料庫，並與其他全臉影像式量測血壓的相關專利進行演算法技術的比較。

5. 第五章：總結與未來展望

針對臉手影像式血壓、基於臉部影像的身體質量指數量測以及全臉式血壓量測系統相關的演算法進行統整，並總結與討論本論文的研究成果。以及分析本論文的題目於未來可以繼續研究與發展的方向，並提出進一步完善血壓量測系統所需要解決的問題。

第二章：臉手式血壓量測系統

本章節為臉手影像式血壓量測系統的介紹。包含臉部與手部選取作為訊號擷取區域位置的介紹、訊號擷取與抗雜訊的演算法、特徵擷取與訊號峰谷值偵測演算法以及預測的模型。並且會和比較論文[2][31]進行模型準確度的比較。

在每一個脈波的週期循環當中，心血管系統的活動始自心臟的泵血，接著便是心肌的去極化。整個過程形成了脈波的訊號，並且能透過影像與訊號處理技術獲得此心跳週期相關的脈波訊號。透過色度訊號[14]及皮膚正交平面法[15]與有限脈衝響應帶通濾波器結合，所擷取出的訊號有抗雜訊的能力。本論文得以使用此訊號結合特徵擷取演算法獲得脈衝傳遞時間[11]的特徵以進行收縮壓與舒張壓的量測。

2.1、血液循環模型與血壓

接著，本節將介紹血液循環模型以及心臟活動的完整週期。並且透過 Moens-Korteweg 方程式[47]獲得血壓和脈衝傳遞時間的關係。

心跳週期當中，首先會先由心房接收到靜脈血，為含氧量較低的血液。而心臟接著會收縮並且施加壓力將右心室的血液送出[48]。此時，右心室會舒張而形成負壓，因此右心房的血液將經三尖瓣進入右心室。進行肺循環後，高含氧量的血液變會經入左心房，接著通過二尖瓣進入左心室。當心臟接著收縮到一定程度後，心室壓力大於大動脈壓力，心瓣會打開，便進入了射血期將血液注入動脈。因此，整個心臟跳動的週期包含了心房與心室各自的收縮與舒張，未運動狀態下的正常人心跳週期約在 0.5~1.1 秒。

在心臟循環過程進行當中，當血液流經動脈時，脈衝的壓力波也會以一定的速度在血管中傳播，此速度稱作脈波傳遞速度(pulse wave velocity)[50]。脈波傳遞速度則和動脈與血管的彈性係數高度相關[51]。此相關性可以由 Moens 及 Korteweg 提出的著名的 Moens-Korteweg 方程式了解。此方程式定義了脈波傳遞速度、血管和血流參數與血壓之間的相關性[47]。而脈波傳遞速度為脈衝傳遞時間的倒數，因此便得以獲得脈衝傳遞時間和血壓之間的關係式[12][47]如下式(1):

$$PWV = \frac{l}{PTT} = \sqrt{\frac{E \cdot th}{2 \cdot rad \cdot \sigma}} \quad (1)$$

其中，PWV 是脈波傳遞速度(pulse wave velocity)，l 是血管距離的長度，而 PTT 則是脈衝傳遞時間(pulse transit time)。E 則是血管管壁的彈性係數，而 th 是血管管壁的厚度， rad 則是血管的內徑， σ 為血液的密度。其中，Hughes 提出血管管壁彈性參數 E 與血壓有相當高的相關性[50]如下式(2):

$$E = E_{org} e^{rp} \quad (2)$$

當中， E_{org} 是在血管管壁不受壓力之下的原始彈性係數，因此每個不同的受測者皆有不同的 E_{org} ，此亦為造成血壓偏移量的原因。而 r 則是比例常數 0.017，p 為血管中的血壓。因此，在假設其他參數不變的情況下，血壓 p 可以透過脈衝傳遞時間 PTT 被算出。而其他參數則會造成血壓偏移量，對應的校正將在第 3 個章節詳細討論並提出解決方案。而血壓 p 與脈衝傳遞時間 PTT 的關係也可以得出如下式(3):

$$p = \frac{1}{r} [2 \ln \left(\frac{1}{PTT} \right) + \ln \left(\frac{2 \cdot rad \cdot \sigma}{E_{0,th}} \right)] = \frac{1}{r} [-2 \ln(PTT) + \ln \left(\frac{2 \cdot rad \cdot \sigma \cdot l^2}{E_{0,th}} \right)] \quad (3)$$

此關係式即提供了以脈衝傳遞時間量測血壓的理論基礎。

2.2、感興趣區域(Region of interest)的選擇

感興趣區域的選擇對於血壓量測模型的準確度有相當的影響。本論文採用的是臉頰與手掌[2][31]作為訊號擷取的區域。由於臉頰與額頭是臉部訊號較強烈的區域，且臉頰微血管分布密集[52]，而手掌靠近虎口的位置則為受測者的橈動脈，血管分布也相當密集，有較好的訊號[53]。此外，臉頰與手掌位置之間的血管距離較遠，足夠用作脈衝傳遞時間的計算[2]。而選擇的兩個感興趣的區域若距離太近，則以普通幀率的攝影機將無法成功提取出對應的脈衝傳遞時間特徵，且會因為取樣率的不足而造成脈衝傳遞時間特徵的誤差。因此，本論文最終使用作臉手影像式血壓量測系統的感興趣區域為受測者的臉頰與手掌如圖 7。

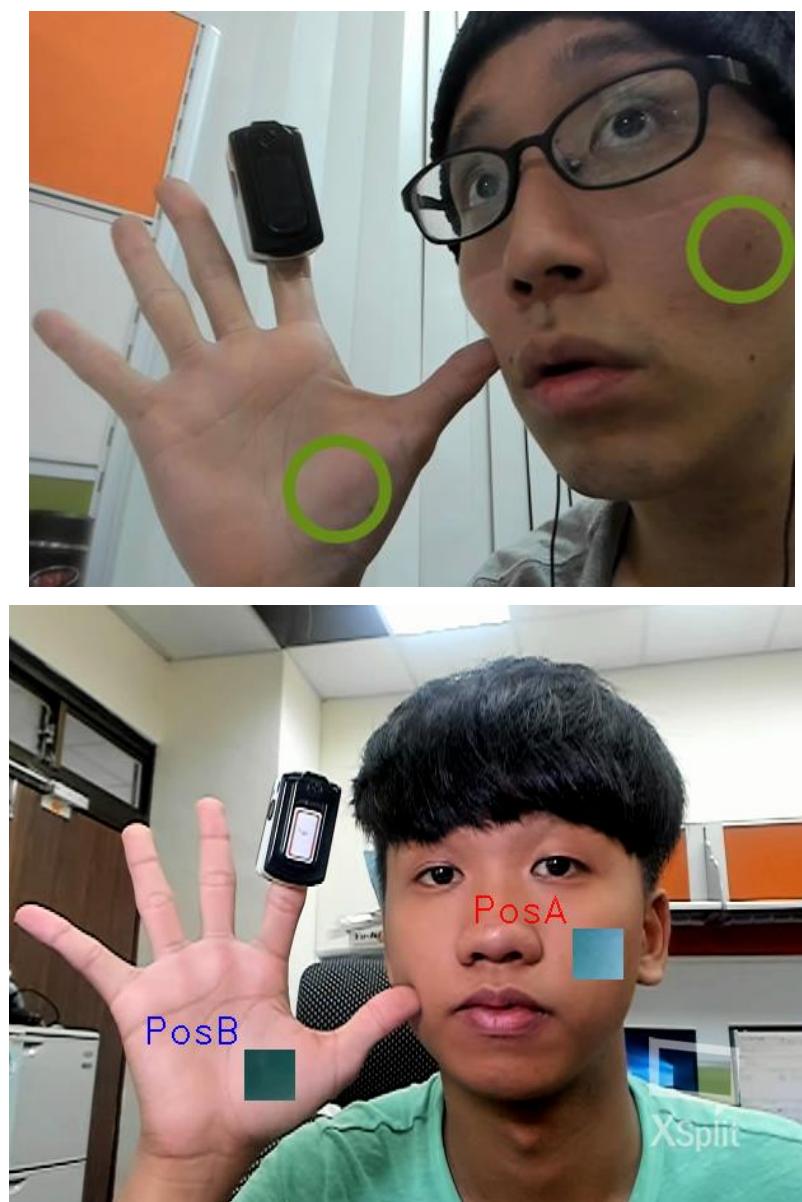


圖 7、選擇受測者的臉頰與手掌橈動脈區域作為血壓量測的感興趣區域

選擇了感興趣區域後便可以開始進行訊號的擷取，接下來將介紹本論文如何利用色度訊號[14]與皮膚正交平面[15]等演算法降低雜訊，以獲得能用於擷取特徵與量測血壓的訊號。

2.3、訊號提取與抗雜訊演算法

本論文使用的是 Logitech Brio 90 帧/秒的攝影機如圖 8，所擷取的臉手影像為由紅色、綠色與藍色三個顏色通道所構成。量測時須將攝影機聚焦在受測者的皮膚，由於有一部份的入射光會被受測者的血紅素吸收，因此能從漫射光獲得血液體積脈衝的生理訊號。其中，綠色通道的生理訊號是最強的[54]，而藍色與紅色通道對應的生理訊號較弱，但 3 個通道都會受到雜訊的影響[16]。因此若直接使用綠色通道作為訊號擷取的來源(Gtrace)，仍會受到較多雜訊的影響。基於此，本論文使用與[31]的 Gtrace 不同的訊號來源擷取方式量測血壓以改善訊號的品質。



圖 8、Logitech Brio 90 帧/秒的 RGB 攝影機

De Haan 於 2013 年首先提出了色度訊號法[14]，將紅色、綠色以及藍色的通道進行線性組合以消除雜訊。而雜訊的產生則是由於在環境光源的影響下，皮膚會產生鏡面反射以及漫反射，而鏡面反射不會受到血液體積脈衝的影響，只會隨著環境光源變化以及相機、光源與皮膚之間的入射角而改變；僅有漫反射光會隨著血液體積脈衝而變化。然而，最終的攝影機收到的影像卻是由漫反射與鏡面反射共同組成的，因此需處理鏡面反射光影響的問題，而不宜僅使用 Gtrace 作為訊號。此外，收到的光源會隨著受測者之於相機的角度而變化，因此需使用色度訊號法處理此問題。

此降雜訊演算法初步的前提假設為環境無色光。在無色光的狀況下(即白光)，鏡面反射對於所有顏色通道當中的漫反射光源的影響皆相同，因此可以透過三個顏色通道(RGB)的線性組合將鏡面反射的成分抵銷，此即為色度訊號法的核心概念[14]。透過兩個正交的色度訊號如下式(4)(5)相除(式(6))便得以消除鏡面反射光的影響：

$$X_i = R_i - G_i \quad (4)$$

$$Y_i = 0.5R_i + 0.5G_i - B_i \quad (5)$$

$$S_i = \frac{X_i}{Y_i} - 1 \quad (6)$$

然而，環境時常會受到色光的影響，也會和受測者的膚色相關。因此需針對膚色的影響進行標準化，而根據 De Haan 的研究，平均的標準化膚色為：

$$[R_s, G_s, B_s] = [0.7682, 0.5121, 0.3841] \quad (7)$$

以膚色作為校正後，可以得出新的 2 個正交的色度訊號，進行相除以獲得較不受色光影響的影像式光體積脈衝訊號(8)：

$$S = \frac{X_s}{Y_s} - 1 \quad (8)$$

其中，(8)當中的 X_s 及 Y_s 為膚色校正後的顏色通道訊號的線性組合如(9)

$$\begin{aligned} X_s &= 3R_n - 2G_n \\ Y_s &= 1.5R_n + G_n - 1.5B_n \end{aligned} \quad (9)$$

由於相機與皮膚之入射角不會太大，因此可以使用泰勒展開式進行近似如(10)

$$\log(1+S) = \log\left(\frac{X_s}{Y_s}\right) = \log(X_s) - \log(Y_s) \quad (10)$$

可以得出(8)的近似為(11)以簡化運算：

$$S \approx X_s - Y_s = 1.5R_n - 3G_n + 1.5B_n \quad (11)$$

然而，在膚色標準化的機制下 X_s 和 Y_s 的變化強度大小可能會不同，因此需進一步進行 α 校正以確保線性組合後的訊號能投影出漫射光的成分，並且能確實的消除鏡面反射光的影響(12)。

$$S = X_f - \alpha Y_f \quad (12)$$

其中， X_f 以及 Y_f 分別為經過帶通濾波作用後的 X_s 及 Y_s 訊號。而用作校正的 α 即為 X_f 的標準差與 Y_f 的標準差的比例如式(13)，可用以消除 X_f 及 Y_f 變化強度不同的影響。色度訊號法最大的好處便是對於非白色的光照訊號有很好的穩健性。並且透過膚色的統計選擇投影方向，以減少需要選取的參數。

$$\alpha = \frac{\sigma(X_f)}{\sigma(Y_f)} \quad (13)$$

透過色度訊號法[14]，本論文得以獲得較不受鏡面反射雜訊影響的訊號。而皮膚正交平面法[15]則是 Wenjin Wang 於 2016 年提出的另一種脈衝訊號擷取演算法。本質上同色度訊號法，皆是透過三色通道(RGB)的線性組合以達成降低雜訊的目的。前提假設為照射肌膚的光源由固定的頻譜成分組成，而強度則會隨著照相機感光元件以及光源

與皮膚的距離而改變，受測者的動作亦會造成強度以及顏色的改變。主要的心跳訊號來自肌膚組織的血紅蛋白和黑色素以及血液體積脈衝的影響。

皮膚正交平面法是基於空間子區域旋轉法(2SR)[55]所延伸得到的訊號降雜訊演算法。2SR 透過基於不同的受測者的肌膚顏色通道 RGB 成分向量隨著時間的改變來量測脈衝訊號。膚色的色調可以透過皮膚像素在圖像當中的統計分布而決定。使用色度作為脈衝擷取的基本參數，便能使用不同色彩空間進行脈衝訊號的量測。然而 2SR 對於雜訊抗性較低。

皮膚正交平面法透過將 2SR 的特徵[55]引入模型，獲得了對抗使用者動作造成的能力。並且進一步透過投影的方式將鏡面反射雜訊和血液體積脈衝漫反射訊號區分開來。投影的訊號包含心跳的同相位以及反相位的成分[15]。此處同色度訊號法[14]透過 α 校正將訊號投影到同相位的方向以獲得正確的心跳訊號。因此，皮膚正交平面法的訊號可表示為(12):

$$R(t) = S_1(t) + \alpha S_2(t), \text{ 其中 } \alpha = \frac{\sigma(S_1(t))}{\sigma(S_2(t))} \quad (12)$$

其中， $S_1(t)$ 與 $S_2(t)$ 是投影的訊號如式(13)，根據肌膚色調進行投影以消除鏡面反射對應的非心跳相位成分。

$$\begin{aligned} S_1(t) &= G_n(t) - B_n(t) \\ S_2(t) &= G_n(t) + B_n(t) - 2R_n(t) \end{aligned} \quad (13)$$

於受測者沒有大幅運動的情況下，色度訊號法[14]及皮膚正交平面法[15]皆有較穩定的表現。本論文最終使用的是皮膚正交平面法所擷取的影像訊號，以用作特徵擷取以及血壓的量測。

有了影像式光體積描述儀的原始訊號後，本論文透過 FIR 有限脈衝響應帶通濾波器(finite impulse response)[56]對其進行濾波。而濾波範圍則為根據美國國家標準的心跳量測儀器規格所設計，為 0.5 Hz 到 3.35Hz(即 30 bpm~201 bpm)[57]。透過濾波器，得以去除高頻雜訊以及低頻的基準線飄移[58]。使用截止角頻率為 $\omega_{c_{low}}$ 與 $\omega_{c_{high}}$ 的帶通濾波器如(14):

$$\begin{aligned} \omega_{c_{low}} &= \frac{0.5}{FPS} 2\pi \\ \omega_{c_{high}} &= \frac{3.35}{FPS} 2\pi \end{aligned} \quad (14)$$

其中，FPS 為幀率 90 幀/秒。而此論文使用之帶通濾波器為 M=192 階，離散的 IIR 帶通濾波器公式 $h[n]$ 如下(15):

$$h[n] = \begin{cases} \frac{(\omega_{chigh} - \omega_{clow})}{\pi} & , \text{當 } n = \frac{M}{2} \\ \frac{\sin(\omega_{chigh}(n-\frac{M}{2})) - \sin(\omega_{clow}(n-\frac{M}{2}))}{\pi(n-\frac{M}{2})} & , \text{當 } n \neq \frac{M}{2} \end{cases} \quad (15)$$

並且為避免濾波器的 Gibbs 效應[59]，本論文使用了窗函數以將濾波器之截止頻率附近的不連續特性平滑化以提升訊號的品質[31]。為了減少旁瓣(sidelobe)的能量並避免高頻干擾，因此採用了 Hamming 窗函數 $ham[n]$ 如(16):

$$ham[n] = 0.54 - 0.46 \cos\left(\frac{2\pi n}{M}\right) \quad (16)$$

透過將(15)與(16)相乘，可得到 FIR 濾波器如(17):

$$f[n] = h[n]ham[n], \quad \forall n \quad (17)$$

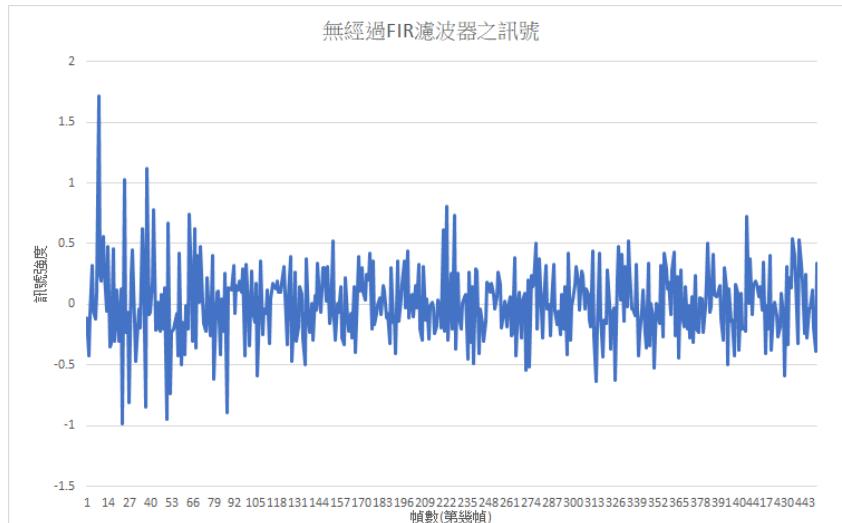


圖 9、未經過 FIR 帶通濾波器之原始訊號

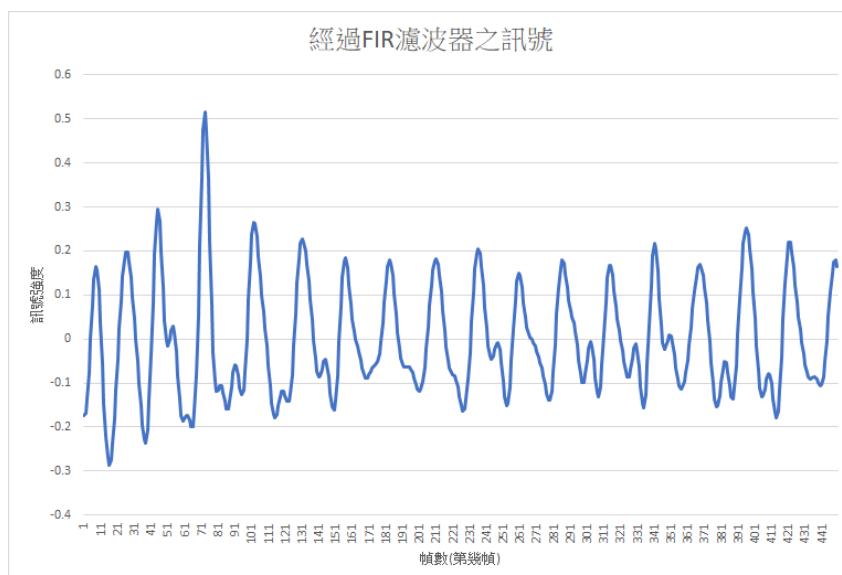


圖 10、經過 FIR 帶通濾波器之訊號

而帶通濾波器的功能為去除高頻的雜訊，以及低頻的基準線飄移。由於攝影機的高幀率，因此原始訊號易受到高頻雜訊影響(圖 9)。由圖 10 的結果可看出透過上述的 0.5 Hz~3.35Hz 的 FIR 帶通濾波器結合 Hamming 窗函數得以擷取心跳範圍內頻率的訊號，將雜訊濾除。

2.4、訊號峰谷值偵測演算法及特徵選取與擷取

獲得了濾波後的影像式光體積描述訊號後，便可開始進行訊號特徵的擷取。為偵測波峰與波谷，本論文採用一階導數法進行極值偵測(式 18)。即偵測一階導數的零點並結合原訊號之鄰近點以判斷其為極大值或極小值。

$$\text{signal}^{(1)}[n] = \text{signal}[n] - \text{signal}[n-1], \forall n \quad (18)$$

然而，由於環境光等雜訊影響，因此訊號的平均值會有所改變，造成偵測到的心跳訊號最大值與最小值不一定是對應的波峰與波谷，而可能是雜訊造成的區域極值。因此，無法使用普通的峰谷值偵測法，而必須使用動態門檻閾值法(dynamic threshold)[31][60]。

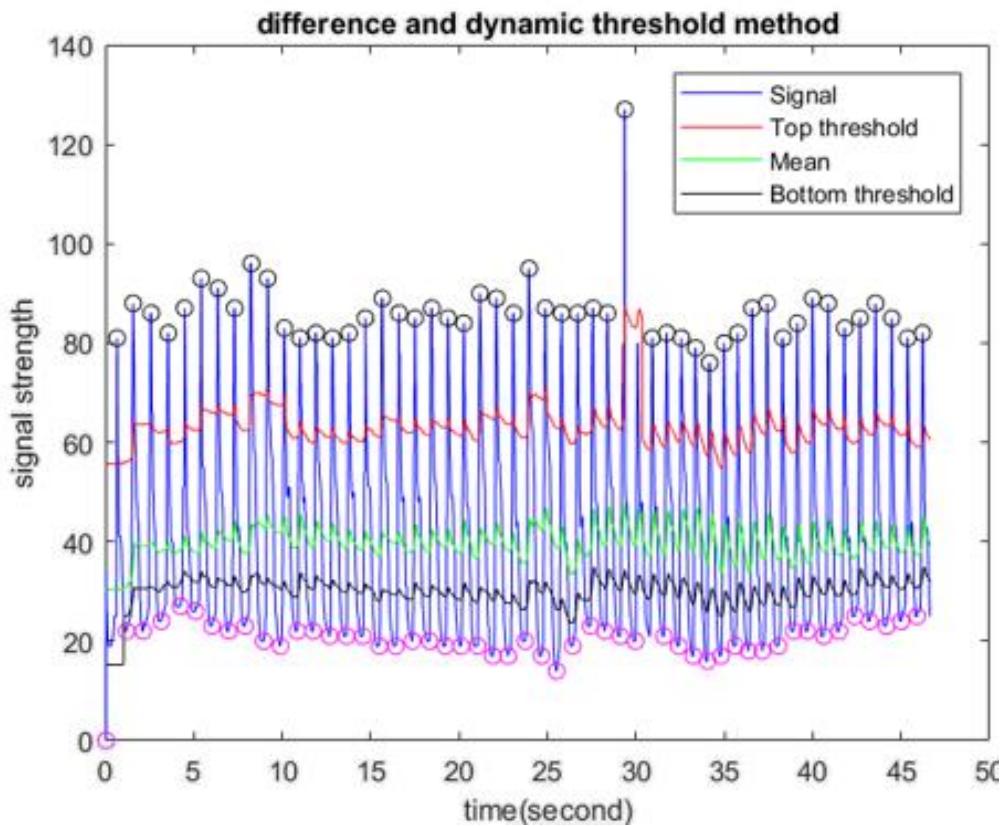


圖 11、使用動態門檻閾值法進行波峰與波谷之偵測。圖中藍色線段為原始訊號，而黑色圓圈為脈波訊號之峰值，而粉紅色圓圈則為谷值。紅色、綠色以及黑色線段則分別為上門檻、平均值與下門檻。

動態門檻閾值法是一種即使遇到訊號之基準線變動，仍能正確地找出實際心跳訊號的最大值與最小值，而不會產生假正(false positive)偵測之演算法。實作動態門檻閾值演算法時，首先須定義動態窗格之大小，由於本論文使用之 Logitech Brio 攝影機幀

率為 90 帖/秒，因此窗格大小定義為 90 帖。首先，當影像累積超過 90 帖時，開始計算窗格當中訊號 $s[i]$ 值的平均 μ 如下(19):

$$\begin{aligned}\mu[N] &= \frac{1}{N} \sum_{i=1}^N s[i], \quad \text{當 } N < 90 \\ \mu[N] &= \frac{1}{90} \sum_{i=N-90+1}^N s[i], \quad \text{當 } N \geq 90\end{aligned}\quad (19)$$

其中， $\mu[N]$ 代表在窗格為第 $N-90+1$ 到第 N 個帖時的訊號之平均值，而在尚未收集到 90 帖的影像前，則是將窗格內所有影像訊號進行平均所獲得的平均值。同理，找出窗格內對應的極大值 M 和極小值 m 如下(20):

$$\begin{aligned}M[i] &= \max(s[i], s[i-1], \dots, s[i-90+1]), \quad \forall i \\ m[i] &= \min(s[i], s[i-1], \dots, s[i-90+1]), \quad \forall i\end{aligned}\quad (20)$$

至此，已獲得在每一個帖當前對應窗格的平均值、最大值以及最小值。可用以計算動態門檻閾值當中所需的上門檻 th_top 與下門檻 th_bom 如下(21):

$$\begin{aligned}th_top[i] &= \mu[i] + (M[i] - \mu[i]) \times 0.25 \\ th_bom[i] &= \mu[i] - (\mu[i] - m[i]) \times 0.25\end{aligned}\quad (21)$$

獲得上門檻與下門檻後，便可以開始進行動態門檻閾值法的峰谷值搜尋。脈波訊號的波峰即為大於對應之上門檻的訊號極大值，而脈波訊號的波谷則為小於對應之下門檻的訊號極小值。將波峰與波谷發生之時間存取下來即可用以後續計算脈衝傳遞時間特徵。

有了峰值與谷值發生之時間特徵後，便可用以計算脈衝傳遞時間特徵。然而，由於環境雜訊的影響，因此極值點仍有可能因為訊號受雜訊干擾變動而偵測錯誤。故本論文採用有效特徵組[31]的技術進行特徵點可靠度的判斷。有效特徵組演算法基於臉部微血管與手部橈動脈訊號特徵發生時間的先後順序為固定的特性，以進行特徵品質的判斷。即臉部訊號前一週期之峰值點、手部訊號當前週期之谷值點、手部訊號當前週期之峰值點、臉部訊號當前週期之谷值點、臉部訊號當前週期之最大導數點，此 5 個特徵點應當照順序發生才會是合理的峰谷值發生順序如圖 12。因此，若此 5 個特徵未照此時間順序發生，則不可構成有效特徵組的條件，亦即無法被使用做後續血壓量測之特徵，以避免特徵受到訊號當中的雜訊影響而造成血壓量測的偏差。血壓量測的模型則會由有效特徵組當中的此 5 個特徵，結合其他相關的脈衝傳遞時間特徵所建立。

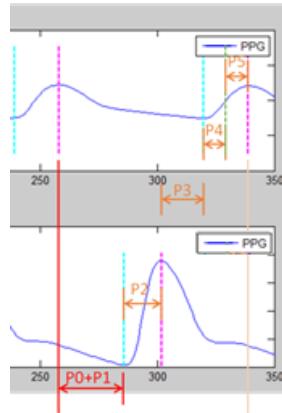


圖 12、上圖為臉部訊號，而下圖為手部訊號。臉部訊號前一週期之峰值點、手部訊號當前週期之谷值點、手部訊號當前週期之峰值點、臉部訊號當前週期之谷值點、臉部訊號點前週期之最大導數點，此 5 個特徵點應當照順序發生才會是合理的峰谷值發生順序

而進行脈衝傳遞時間特徵計算時，將使用到心跳速率、幀率以及心跳間隔時間之間的關係式如下(22):

$$HR(\text{bpm}) = \frac{60 \times \text{fps}}{\text{IBI}} \quad (22)$$

其中，HR 代表心跳，單位是每分鐘脈搏數(beats per minute)；而 fps 為幀率，單位為幀/秒；IBI 則為心跳間隔(inter-beat interval)，單位為間隔之幀數。此式描述了心跳與心跳間隔之間的換算關係。

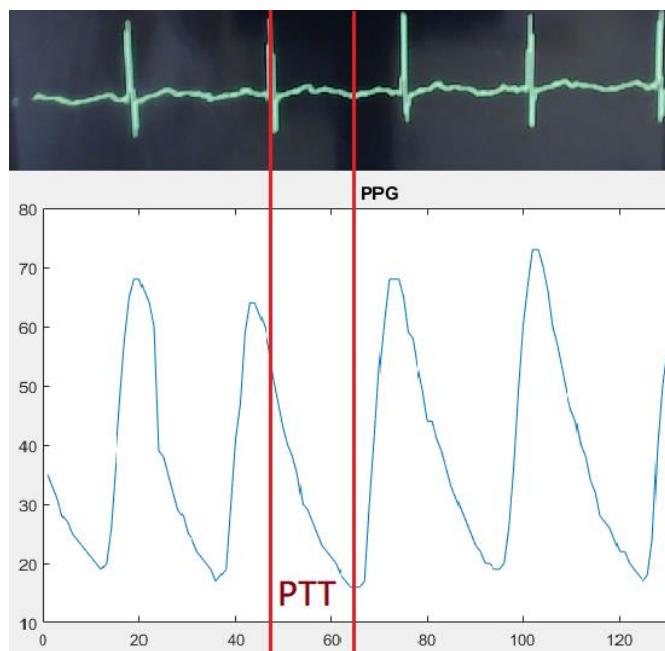
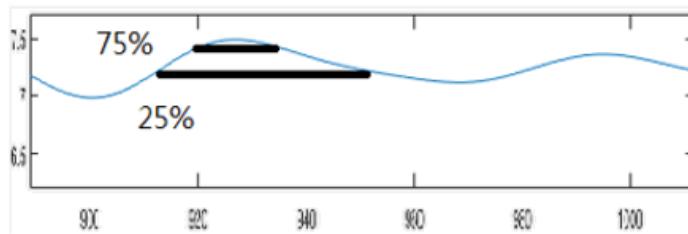


圖 13、傳統脈衝傳遞時間(pulse transit time)的定義為心電圖的 R 波波峰至光體積描述儀訊號的谷值點之間的時間差

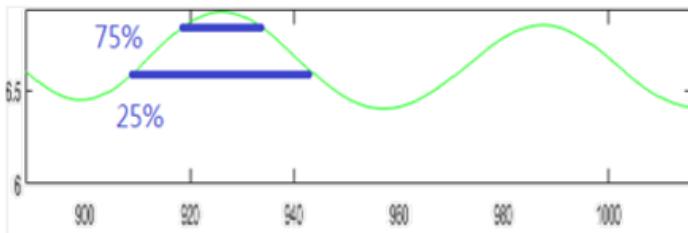
傳統而言，脈衝傳遞時間的定義為心電圖訊號 R 波的波峰至光體積描述儀訊號的谷值或最大導數點之間的時間差如圖 13。然而，本論文使用的為影像式光體積脈衝訊號，因此更改脈衝傳遞時間特徵擷取的方式為人體的臉部與手部影像訊號之峰谷值等特徵的時間差。

本論文使用包含臉部訊號前一週期之峰值點、手部訊號當前週期之谷值點、手部訊號當前週期之峰值點、臉部訊號當前週期之谷值點、臉部訊號當前週期之最大導數點作為特徵，且此五者須照順序發生。此外，亦包含手部訊號最大導數點、臉部訊號對應週期振幅的 25%強度點、臉部訊號對應週期振幅的 75%強度點、手部訊號對應週期振幅的 25%強度點以及手部訊號對應週期振幅的 75%強度點(如圖 14)共九個特徵。此九個特徵便是用以計算脈衝傳遞時間的核心資訊，透過詳細的時域特徵資訊，本論文得以降低對高幀率攝影機的需求[2]，因此能以 90 幀/秒的幀率進行血壓量測。

Artery:25%、75%



Microvascular:25%、75%



Plot with both the signals

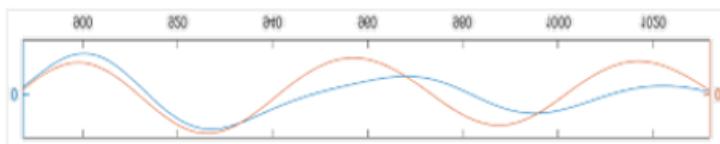


圖 14、臉(微血管)手(橈動脈)訊號振幅之 25%以及 75%之特徵點

而本論文進行特徵擷取前，先將窗格內訊號進行 Z 標準化，亦即將訊號減去其窗格範圍之平均值並除以標準差。採取此步驟可除去臉部及手部訊號強度上的差異性，並取得能用以擷取特徵的訊號。

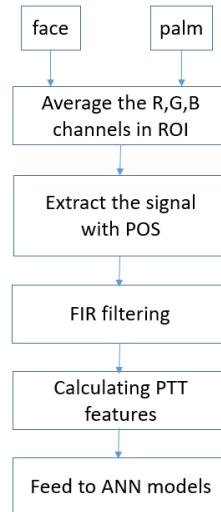


圖 15、臉手影像式血壓量測流程圖

本論文之血壓系統用以量測受測者之收縮壓與舒張壓的流程如圖 15。使用做比較的模型包含 K 最近鄰居法(K nearest neighbors)、深度信念網路-深度神經網路(deep belief network-deep neural network)以及人工神經網路(artificial neural network)，而所採用之脈衝傳遞時間特徵共 9 個如表 1。

Features	Description
p0~p1	the bottom between artery signals
p2	the bottom of the artery signal to the top of it
p3	the bottom of the microvascular signal to the max slope of it
p4	the max slope of the microvascular signal to the top of it
p5	the inter-beat interval
p6	the top of previous microvascular signal to the 25% peak with negative slope of the current microvascular signal
p7	Artery bottom to the 25% peak with positive slope
p8	Artery bottom to the 75% peak with positive slope
p9	the 75% peak with positive slope of the microvascular signal to the 75% peak with positive slope of the artery signal

表 1、臉手影像式血壓量測系統所使用之脈衝傳遞時間特徵

2.5、血壓資料庫及基於脈衝傳遞時間特徵量測血壓的機器學習模型

本論文使用之影像式光體積描述訊號皆為由 Logitech Brio 90 帧/秒的攝影機所取得。血壓資料庫亦為同一攝影機所錄製，而用做比較之實際收縮壓與舒張壓則由標準血壓量測儀器同步錄製。

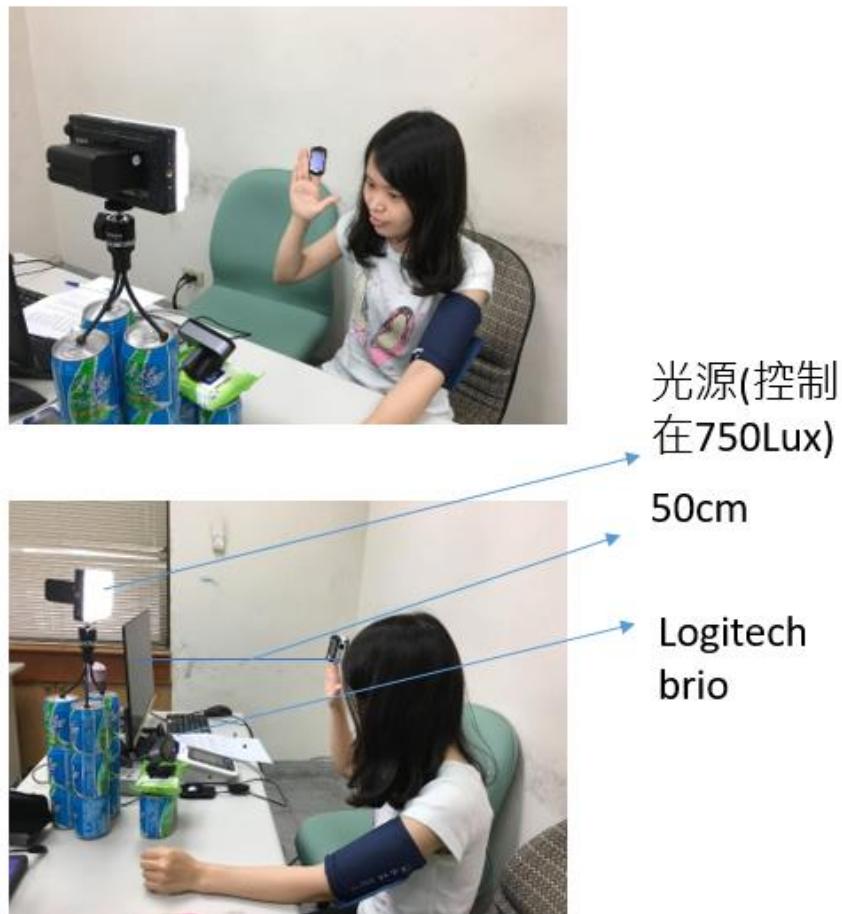


圖 16、臉手式血壓資料庫之錄製環境，包含補光燈、Logitech Brio 攝影機以及綁於受測者上臂之歐姆龍血壓計

採用之血壓資料庫收集自 27 位年齡自 20 歲至 45 歲的受測者，當中包含 19 名男性與 8 名女性，皆簽名同意接受實驗。總共包含 1033 筆資料，每筆資料包含 45 秒以 Logitech Brio 攝影機進行之影像錄製，幀率為 90 帧/秒，而影像解析度則為 640x480。受測者採靜態坐姿，而環境亮度皆以照度計控制在 750Lux，距離螢幕 50 公分進行錄製(如圖 16)。而實際的收縮壓與舒張壓則由符合美國食品藥物管理局(Food and Drug Administration)認證之歐姆龍(Omron)袖套式血壓量測儀器(如圖 17)提供，量測過程恪遵歐姆龍血壓計之標準使用規則。將袖套綁於受測者之上臂，並且袖套上之箭頭標示對準受測者肘窩肱動脈處[61]，血壓資料庫量測之規格如下表 3。

項目	說明
相機	Logitech Brio
解析度	640x480
幀數	90 fps
顏色深度	3 通道 24bits
儲存型態	. bmp
系統平台	Windows-vivo book，CPU:i7 7700，使用 16G 的 RAM 進行錄製(用 Ramdisk 軟體將 RAM 用作存取媒介來錄製 90 fps 影像)
Ground Truth	歐姆龍血壓機
亮度(臉前)	750 lux
姿勢	靜態坐姿
光源	3500k/5000k/6500k/8000k 色溫燈
距離	50cm

表 3、影像式血壓資料庫之錄製規格表



圖 17、歐姆龍血壓機

本論文自上述錄製之血壓資料庫當中的每筆資料以 2.4 節說明之移動窗格法蒐集受測者的 9 個脈衝傳遞時間相關特徵進行血壓量測。不同於比較論文[2]使用之個人化模型，本論文採用由血壓資料庫當中之所有受測者訓練出之通用模型進行血壓量測。如此，便能提升血壓量測系統之實用性與方便性，因不必針對不同受測者獨立訓練不同的模型。

此外，本論文使用之資料庫當中最高的受測者收縮壓達到 166 毫米汞柱，較[31]的最高血壓高了 31 毫米汞柱；而最高的舒張壓為 105 毫米汞柱，也較[31]高了 14 毫米汞柱。而最低的收縮壓為 83 毫米汞柱，較[31]低了 1 毫米汞柱；最低的舒張壓為 48 毫米汞柱，也較[31]低了 9 毫米汞柱。由於血壓資料庫的範圍較大，因此所建構出的模型，對於不同範圍之血壓的預測能力也較佳。而本論文之資料庫的錄製過程皆嚴格紀錄對應之幀率，以確保所錄製之影像幀率(fps)皆介於 89 幀/秒至 90 幀/秒之間。圖 31 為本論文實作之血壓資料庫錄製程式於錄製時所產生的幀率相關紀錄檔。

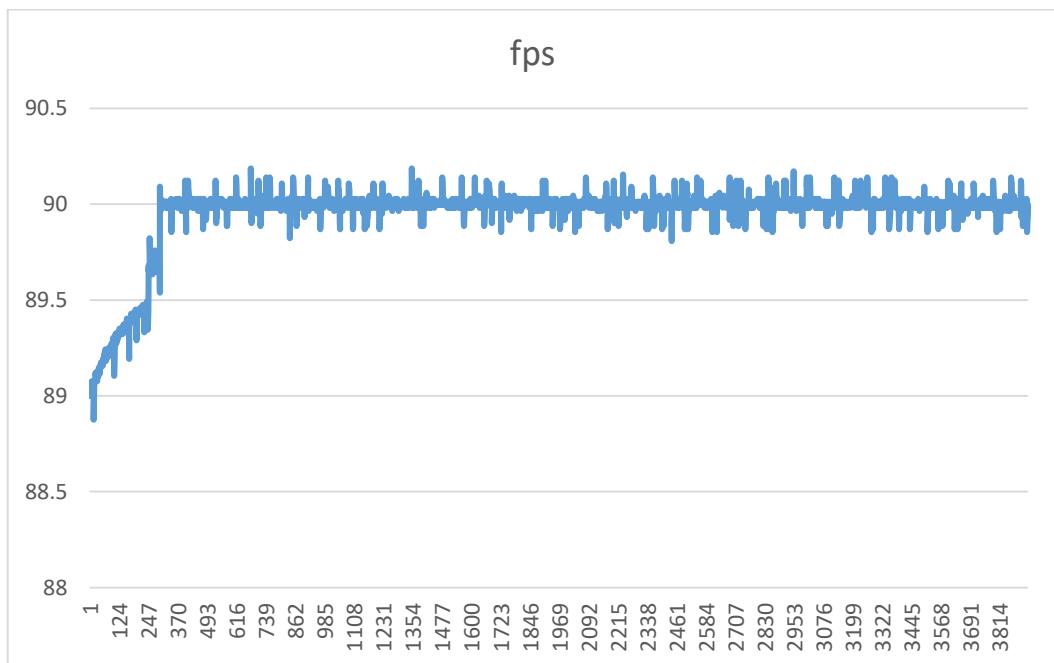


圖 31、影像式血壓資料庫錄製時所產生之幀率記錄檔，確保幀率皆介於 89 幀/秒至 90 幀/秒之間。X 軸為當前幀數，而 Y 軸則為當下對應之幀率

2.5.1、K 最近鄰居演算法模型

為進行模型表現之分析，本論文實作 K 最近鄰居法以進行上述血壓資料庫之收縮壓以及舒張壓的量測。K 最近鄰居法相較於神經網路而言為較傳統之監督式學習演算法。此演算法首先建立包含 9 個脈衝傳遞時間特徵以及對應的收縮壓及舒張壓的資料庫。並計算測試資料當中的特徵向量 x_{test} (9×1 的向量)與資料庫當中其他所有特徵向量 x_{train} 之間的歐式幾何距離(L_2 norm)如下(23):

$$\text{distance}(x_{train}, x_{test}) = \sqrt{\sum_{i=1}^9 (x_{train_i} - x_{test_i})^2} \quad (23),$$

接著，找出資料庫當中 $K=10$ 個距離測試資料最接近的特徵向量，並將其對應之血壓進行平均以得出最終的測試資料血壓量測如下(24):

$$y_{\text{test}} = \frac{1}{10} \sum_{i=1}^{10} y_{\text{train}_i} \quad (24)$$

其中， y_{test} 為最終根據測試資料之脈衝傳遞時間特徵量測出的血壓；而 y_{train_i} 則為與 x_{test} 特徵向量距離第*i*近之 x_{train_i} 特徵向量所對應之實際血壓。透過此10筆血壓值之平均便得以獲得K最近鄰居法血壓之量測。

K最近鄰居法的優點為精確度較高，且對資料當中的離群值較不敏感，也不須針對輸入資料作出假設。然而，K最近鄰居法仍是基於訓練資料庫的整體資料進行量測，對訓練集的依賴程度較高而不可丟棄任何資料。因此時間複雜度和空間複雜度皆會隨著訓練集規模的增加而增加。且K最近鄰居法之模型表現除了受資料庫之影響外，僅取決於K值的選擇以及特徵距離的計算函數(舉例而言，本論文之K值為10，而距離函數為L2 norm)。

為提升血壓量測的表現，本論文亦引入深度神經網路之模型，進行脈衝傳遞時間特徵對血壓的量測。

2.5.2、深度信念網路-深度神經網路(DBN-DNN)

深度信念網路[63]是Geoffrey Hinton於2006年提出之基於受限波茲曼機[64]的神經網路模型。透過機率的概念訓練網路當中神經元的權重，並使信念網路能透過機率模型產生訓練資料，進而識別輸入特徵和對應血壓之間的關係。深度信念網路是由許多層的神經元構成，其中包含顯性以及隱性神經元。顯性神經元用於模型前級以處理輸入的特徵資料向量，而隱性神經元則用於神經網路中間層，進行特徵的提取與偵測。

深度信念網路之組成基底為受限波茲曼機(restricted Boltzmann machines)。信念網路的每一層皆透過輸入之特徵向量以推斷隱形層，並繼續利用生成之隱形層特徵作為下一個隱形層之資料向量輸入，進而建構完整的深度網路[65]。受限波茲曼機的功能相當於聚類器，由一層顯性神經元及一層隱性神經元構成。且該顯性層與隱性層各自內部之神經元並無相互連接，連接線僅存在於層間，因此得以形成條件獨立。即在已知所有顯性神經元對應的數值的條件下，隱性神經元彼此之間的取值並無關聯；而反之亦然，在已知所有隱性神經元對應的數值的條件下，顯性神經元彼此之間的取值亦無關聯如下(25)：

$$\begin{aligned} p(H|V) &= \prod_{i=1}^N p(H_i|V) \\ p(V|H) &= \prod_{i=1}^M p(V_i|H) \end{aligned} \quad (25)$$

其中，H 代表隱性層之神經元，而 V 則為顯性層之神經元，i 代表對應層當中的神經元編號。條件獨立的性質大大簡化了神經網路的訓練，得以同時計算整層神經元，而不必各自運算。針對每一筆特徵向量，受限波茲曼機會透過訓練過程產生的權重決定開啟或關閉隱性神經元的機率值，並據此機率隨機的激活對應的隱性神經元；同理，給定隱性層，計算顯性層神經元是否開啟之方法亦同。因此，受限波茲曼機整體訓練過程之目的便是訓練出隱性層神經元，使其能產生出最接近原輸入特徵的機率分布。藉此，受限波茲曼機便能透過隱性層提取顯性神經元層之特徵，還原顯性層並且達到特徵提取的目的。

因此，由多層受限波茲曼機堆疊組成的深度信念網路便是一個生成模型與迴歸預測模型，一層一層擷取特徵，並作為下一級的輸入，進行預訓練以獲得初始權重。此特徵擷取過程屬於非監督式學習，為將此模型用於有標記之資料庫，本論文在最末層的受限波茲曼機訓練時，加入對應的標籤進行訓練，配合梯度下降法[66] 以回授的方式修改每層間的權重，進而降低誤差值。

以下將介紹本論文之受限波茲曼機的計算數學推導，使用 9 個脈衝傳遞時間(PTT)特徵，並設定隱性層神經元數為 10 個，首先定義其能量 E(v, h)如下(26):

$$E(PTT, H) = -\sum_{i=1}^9 PTT_i a_i - \sum_{j=1}^{10} H_j b_j - \sum_{i=1}^9 \sum_{j=1}^{10} PTT_i H_j W_{i,j} \quad (26)$$

其中， PTT_i 為輸入之第 i 個脈衝傳遞時間特徵，而 a_i 則為對其對應的偏移量； H_j 則為對應的第 j 個隱性層神經元， b_j 為其對應的偏移量；而 $W_{i,j}$ 則為顯性層神經元 PTT_i 以及隱性層神經元 H_j 之間的權重。此能量定義來自物理學的 Ising 模型[67]。將此能量函數進行歸一化後可以得到受限波茲曼機當中隱性層和顯性層之間的聯合機率分布如下(27):

$$P(PTT, H) = \frac{1}{\sum_{PTT,H} e^{-E(PTT,H)}} e^{-E(PTT,H)} \quad (27)$$

而顯性層的邊緣分布(marginal distribution)可透過對所有隱性層之神經元求和而得如下(28):

$$p(PTT) = \frac{1}{\sum_{PTT,H} e^{-E(PTT,H)}} \sum_H e^{-E(PTT,H)} \quad (28)$$

對應之各個隱性層及顯性層神經元被激活的機率分別如下(29):

$$\begin{aligned} p(H_j = 1 | PTT) &= \sigma(b_j + \sum_{i=1}^9 W_{i,j} PTT_i) \\ p(PTT_i = 1 | H) &= \sigma(a_i + \sum_{j=1}^{10} W_{i,j} H_j) \end{aligned} \quad (29)$$

其中， PTT_i 、 H_j 、 a_i 、 b_j 以及 $W_{i,j}$ 之定義同上，而 σ 則為邏輯函數(sigmoid function)。Ising 模型的能量形式使得深度信念網路當中每一層的運算皆可獨立，單獨訓練，而不會影響其他層的權重。亦即，欲將整體能量或所產生特徵之機率分布最大化，可直接將能量逐層最大化。此即為 Ising 模型的優勢，此能量表示法得以將脈衝傳遞時間的特徵擷取以及訓練過程模組化，便於模型的建構。本論文之深度信念網路為避免過擬合之問題，採用 0.005 的較小的學習率。並透過 Geoffrey Hinton 於 2002 年提出的對比性散度法[68]，結合梯度下降[66]以及基於馬爾克夫蒙地卡羅 (MCMC) 的吉布斯採樣[69]進行反向傳播以完成權重的更新，建立最終模型。並以此模型結合脈衝傳遞時間特徵進行血壓之量測。

2.5.3、人工神經網路(ANN)

人工神經網路[70]為一種基於回饋與遞歸的神經網路模型。透過人工神經元之間的聯繫進行計算，透過反饋的資訊改變內部神經元的權重與偏移量以達到學習的功能。由於每一層皆有非線性之激活函數作用於神經元，因此人工神經網路得以用作非線性相關之特徵資料的建模。

監督式神經網路模型主要包含三大可調整的要素，即模型架構、激活函數以及目標函數[71]。模型架構決定了神經網路當中神經元之間的連結關係，以及每層間的權重矩陣的大小。而激活函數則為網路引入了非線性的性質，且須為可微分的，因為模型使用梯度下降演算法進行誤差反向傳遞時，需對此函數進行微分。而本論文使用之目標函數則為損失函數，亦即神經網路模型訓練過程所欲最小化之目標。透過訓練過程當中的誤差反向傳播(back propagation)[72]，資料庫實際血壓得以透過損失函數反饋，並利用此函數之一次微分來對前級各層之權重進行校正。

由於血壓相關的時域特徵有 9 個，因此本論文並無採用深層的人工神經網路，以避免過擬合的問題。採用之架構包含 3 層隱藏層(hidden layer)，而每個隱藏層包含之神經元數為 10 個，以進行模型的訓練。

2.6、臉手式血壓實驗結果以及與相關論文之比較

於此小節，本論文將分別呈現 K 最近鄰居法、深度信念網路以及人工神經網路三種臉手式血壓資料庫的預測模型之結果(純使用脈衝傳遞時間特徵)，並最終與相關論文[2][31]進行比較。

同時亦將於 2.6.5 小節呈現基於本論文實作之血壓系統針對不同受測者即時量測之誤差結果，值得注意的是，此系統之功能包含第 3 章將提及之身體質量指數與血壓偏移量的校正(同時使用脈衝傳遞時間特徵以及身體質量指數進行血壓量測)。

2.6.1、K 最近鄰居法(KNN)之血壓方均根誤差

首先，將針對 K 最近鄰居法進行比較。K 最近鄰居法使用血壓資料庫當中，與當前所得特徵之間的 L2 norm 最接近的 K 個特徵分別對應的 K 個收縮壓以及脈壓進行平均以得到最終的收縮壓，及將收縮壓與脈壓相減得到的舒張壓量測結果。本論文用於測試之 K 值為 1 到 10，透過 2.5 節提及之血壓資料庫進行結果比較。其中，收縮壓之方均根誤差於 K=10 時為 12.33 毫米汞柱，而舒張壓之方均根誤差則為 9.83 毫米汞柱。

2.6.2、深度信念網路-深度神經網路(DBN-DNN)之血壓方均根誤差

而基於深度信念網路所訓練之臉手式血壓模型，針對 2.5 節之血壓資料庫以 8 比 2 的訓練及測試資料庫分割比例產生之測試資料集收縮壓的方均根誤差為 11.72 毫米汞柱，而舒張壓的方均根誤差則為 8.63 毫米汞柱，相較於 2.6.3 的人工神經網路表現略差。可能的原因為血壓資料庫使用的特徵量相對較少，因此使用深度信念網路進行預測時，產生了過擬合之問題。

2.6.3、人工神經網路(ANN)之血壓方均根誤差

基於人工神經網路所訓練之臉手式血壓模型，針對 2.5 節之血壓資料庫以 8 比 2 的訓練及測試資料庫分割比例產生之測試資料集收縮壓的方均根誤差為 11.13 毫米汞柱，而舒張壓之方均根誤差則為 7.78 毫米汞柱。為此 3 種演算法當中表現最好之結果。本論文之人工神經網路的優勢在於使用的層數不深，不會如 2.6.2 之深度信念網路發生過擬合的問題，使得血壓量測的誤差結果對於訓練集以及測試集較為接近。

2.6.4、與相關論文[2][31]之比較

Jeong 於 2016 年所提出的比較論文[2]，為使用純影像非接觸式的演算法進行之血壓量測。此論文之特色為使用高幀率(420 幀/秒)之攝影機進行 10 秒鐘的影像錄製以擷取相關訊號特徵，並進而利用單一的脈衝傳遞時間特徵建立模型，以進行血壓的量

測。所使用之血壓模型則為針對每個受測者，以線性迴歸演算法所客製化訓練的個人化模型。而由於僅有收縮壓與脈衝傳遞時間之間的線性相關性較高，因此該論文僅能針對收縮壓做預測，並無建構能準確預測舒張壓之模型。此外，由於每個受測者的影像式脈衝傳遞時間特徵與收縮壓之相關性都不同，且在不同的光照及膚色的條件下，訊號及特徵也會有變化，因此該論文針對每個不同的受測者之血壓偏移量皆以血壓儀器進行個別校正以設計個人化模型。此法需使用經醫療器材認證之血壓機首先量測實際的血壓，校正系統及演算法參數，才可繼續用之以量測收縮壓，實用上之方便性相對較低。而該論文使用之特徵為綠色通道之訊號(GTrace)，對於雜訊及非心跳相位訊號之光影變化等干擾的抗性亦較低。此外，該期刊僅使用包含 7 位身體健康之受測者進行血壓量測結果的評比，包含年齡皆介於 20 歲至 50 歲之間的 3 位女性及 4 位男性，並且要求受測者於量測過程中閉上雙眼。此外，該論文使用之血壓資料庫之總資料筆數僅有 21 筆，分別為各 3 筆來自共 7 位受測者的資料。而此論文所使用之血壓資料庫的年齡及不同受測者的血壓資訊如下表 5。

性質\統計量	最大值	最小值	平均值	標準差
年齡(歲)	皆為 20 至 50 歲			
受測者 1 的 收縮壓(毫米汞柱)	129	105	116.8	8.1
受測者 2 的 收縮壓(毫米汞柱)	139	119	129.1	8.4
受測者 3 的 收縮壓(毫米汞柱)	126	107	114.2	5.8
受測者 4 的 收縮壓(毫米汞柱)	104	92	98.9	4.1
受測者 5 的 收縮壓(毫米汞柱)	149	113	125.2	11.7
受測者 6 的 收縮壓(毫米汞柱)	170	128	149.3	14.8
受測者 7 的 收縮壓(毫米汞柱)	151	133	140.3	7.0

表 5、Jeong 所提出之臉手影像式血壓資料庫之相關資訊

而 Sean Lin 於 2017 年所提出之比較論文[2]，則使用較低之幀率(75 幀/秒)進行影像式血壓之量測。該論文使用之資料庫為收集自 13 位受測者，包含 10 位男性以及 3 位女性，而每位受測者皆配合進行 10 次血壓之量測，因此總資料數為 130 筆。選擇之訓練集與測試集之比例亦為 8 比 2。而所有受測者的血壓資訊如下表 6。該論文所建立之血壓量測模型為通用化的模型，並無針對不同受測者客製化設計。此外，Sean Lin 所使用之血壓量測模型為基於 K 最近鄰居法，得以描述脈衝傳遞時間特徵與血壓之間的非線性關係，因此該論文所提出之模型可對於收縮壓以及舒張壓進行量測。然而，Sean Lin 於論文中所使用之影像式光體積訊號和 Jeong 相同，為綠色通道訊號(Gtrace)，因此對於雜訊以及環境光變化等影響之抗性較低。此外，每位受測者之影像式脈衝傳遞時間特徵與血壓之相關性皆不同，然而此論文並無針對此血壓偏移量進行校正，因此會導致血壓的偏差，使得此論文之演算法所建構的通用化模型所量測出的血壓誤差較大。且該論文使用之血壓資料庫當中的受測者血壓範圍較小，收縮壓的範圍為 84 毫米汞柱至 135 毫米汞柱，而舒張壓之範圍則為 57 毫米汞柱至 91 毫米汞柱。因此，基於該資料庫所建構之模型將無法準確量測此血壓範圍以外的血壓。

性質\統計量	最大值	最小值	平均值	標準差
年齡(歲)	33	21	23.5	2.7
收縮壓(毫米汞柱)	135	84	110.03	11.05
舒張壓(毫米汞柱)	91	57	72.78	7.50

表 6、Sean Lin 所提出之臉手影像式血壓資料庫之相關資訊

上述之包含 Jeong 以及 Sean Lin 的兩篇論文為近期唯二僅使用影像進行血壓量測之論文。其他相關的血壓論文皆是使用心電圖配合接觸式光體積訊號，或是純接觸式光體積訊號進行血壓量測。部分論文則是使用超聲波感測器或是壓電感測器擷取脈波以進行血壓之量測。因此，本論文最終選擇之比較論文為 Jeong[2]與 Sean Lin[31]的研究。其中，Sean Lin 的論文使用之模型與本論文同為通用化模型，且使用之攝影機幀率較接近，因此為主要之比較對象。

本論文使用之資料庫範圍較廣，資料庫當中最高的受測者收縮壓達到 166 毫米汞柱，較[31]的最高血壓高了 31 毫米汞柱；而最高的舒張壓為 105 毫米汞柱，也較[31]高了 14 毫米汞柱。而最低的收縮壓為 83 毫米汞柱，也較[31]低了 1 毫米汞柱；最低的舒張壓為 48 毫米汞柱，較[31]低了 9 毫米汞柱。使用了更全面的資料庫使得此模型的通用性更佳，可預測的血壓範圍也較大。且本論文採用之影像式光體積描述訊號擷取

之演算法為色度訊號法[14]及皮膚正交平面法[15]。因此，對於雜訊以及環境光等的影響抗性較強。且本論文共擷取 9 個脈衝傳遞時間特徵(見表 1)，因此可保有較多的時域相關資訊。對於不同受測者血壓偏移量之校正，本論文亦有於第三章提出基於身體質量指數特徵的解決方案，而不需以額外的血壓儀器輔助。而本論文使用之影像擷取裝置為 Logitech Brio 90 帖/秒的攝影機，成本較 Jeong 使用的 EX-FH20 420 帖/秒的攝影機低了許多。此外，由於使用包含 K 最近鄰居法、深度信念網路模型以及人工神經網路等非線性模型，因此本論文得以較佳的描述脈衝傳遞時間與血壓之間的關係。而本論文與此兩篇比較論文[2][31]之間的結果比較如表 7。

性質\論文	本論文	Sean Lin 的研究[31]	Jeong 的研究[2]
資料庫受測者人數	27 位	13 位	7 位
總資料筆數	1033 筆	130 筆	21 筆
每筆資料錄製的影像包含	以 90 帖/秒所錄製之 45 秒影像	以 75 帖/秒所錄製之 40 秒影像	以 420 帖/秒所錄製之 10 秒影像
資料庫收縮壓範圍	83~166 毫米汞柱	84~135 毫米汞柱	個人化模型 (不討論)
資料庫舒張壓範圍	48~105 毫米汞柱	57~91 毫米汞柱	個人化模型 (不討論)
使用之攝影機以及幀率	Logitech Brio 90 帖/秒	Sony ps3 eye 75 帖/秒	EX-FH20 420 帖/秒
個人化/通用化	通用化模型	通用化模型	個人化模型
需額外血壓機校正	不須校正	不須校正	需校正
使用之模型	KNN/DBN-DNN/ANN	KNN	LSM(最小平方法)
可量測之血壓	收縮壓及舒張壓	收縮壓及舒張壓	收縮壓
訊號擷取區域	臉頰、手掌	臉頰、手掌	臉頰、手掌
基準血壓量測儀器	歐姆龍血壓機	歐姆龍血壓機	Carescape V100
收縮壓方均根誤差 (毫米汞柱)	11.13	16.50	15.08 (使用相同特徵)
舒張壓方均根誤差 (毫米汞柱)	7.78	8.20	

表 7、本論文使用時域特徵進行血壓量測之實驗相關資訊與數據，以及與相關論文
[2]、[31]之比較

本論文透過 K 最近鄰居法及深度學習之模型，得以建構較準確之血壓量測模型。且資料庫的大小及其中受測者的血壓範圍較大，因此所用以訓練之模型亦較全面。此外，使用之幀率 90 幀/秒的攝影機相較於 Jeong 等人的 420 幀/秒的攝影機低了許多。而通用化的模型，應用面也較廣，不須於每次量測時都針對不同受測者一一訓練模型。此外，本論文將於第三章提出一種經由人臉影像量測身體質量指數之演算法，並可將該身體質量指數特徵結合入血壓量測系統，以用做血壓偏移量的校正，因此除了額外血壓儀器校正的需求，提升了實用性。

2.6.5、基於本論文實作之血壓量測系統結合身體質量指數的誤差表現

本論文實作之血壓量測系統，除了使用 2.4 節所提及之 9 個脈衝傳遞時間之特徵外，亦於第三章實作了身體質量指數量測模型以處理不同受測者之血壓偏移量的問題。2.6.4 小節所呈現之比較表，為僅透過脈衝傳遞時間特徵，並基於血壓資料庫所進行的血壓量測的誤差之結果。而此小節將呈現之比較表格則是除了脈衝傳遞時間外，亦結合了身體質量指數特徵所訓練之模型。此外，此小節對應的量測結果及誤差為實際使用血壓量測系統，實時針對共 14 位不同受測者各進行 3 次血壓量測所統計出的方均根誤差，而非以血壓資料庫之現有的資料進行的模型準確率評比，因此更能體現血壓模型的量測能力。而使用作為比較基準的血壓量測儀器為歐姆龍血壓機。結合身體質量指數特徵的臉手影像式血壓量測系統的實際實驗結果如下表 9。

受測者編號	真實收縮壓 (mmHg)	真實舒張壓 (mmHg)	量測收縮壓 (mmHg)	量測舒張壓 (mmHg)
Subject 1	106	71	114.0	80.0
	110	68	119.3	75.5
	105	64	113.5	73.4
Subject 2	110	70	115.8	68.8
	112	73	120.1	69.9
	109	65	117.2	62.8
Subject 3	111	65	120.3	61.2
	116	68	121.9	65.8

	116	69	126. 8	69. 1
Subject 4	140	73	122. 1	63. 2
	142	75	120. 7	61. 2
	143	77	122. 5	63. 8
Subject 5	114	69	104. 4	73. 1
	114	69	102. 7	66. 3
	108	64	111. 6	61. 2
Subject 6	104	67	109. 2	61. 5
	107	68	99. 7	61. 3
	109	70	119. 2	62. 7
Subject 7	105	69	116. 9	60. 7
	101	66	109. 8	61. 9
	101	65	109. 2	60. 5
Subject 8	124	66	130. 5	60. 7
	118	62	128. 3	62. 9
	122	62	133. 5	61. 8
Subject 9	127	67	134. 8	73. 1
	121	59	111. 6	70. 3
	124	62	112. 4	72. 4
Subject 10	105	66	117. 0	75. 8
	103	61	112. 5	68. 1
	103	62	111. 0	67. 5
Subject 11	100	65	111. 8	62. 1
	98	64	102. 2	57. 5
	101	63	101. 2	62. 2
Subject 12	125	85	131. 1	72. 6
	135	81	140. 9	77. 1
	117	79	118. 2	70. 3
Subject 13	100	65	100. 8	64. 6

	99	53	97.4	51.4
	102	61	109.6	60.9
Subject 14	121	68	112.4	66.7
	115	68	107.1	68.7
	110	68	118.2	66.4

表 9、結合身體質量指數特徵之臉手影像式血壓量測系統的實驗結果

上表為實驗結果，而收縮壓之方均根誤差為 9.63 毫米汞柱，舒張壓之方均根誤差則為 6.62 毫米汞柱。本論文實作的血壓量測系統之方均根誤差相較於比較論文[31]有所下降。

2.7、臉手影像式血壓模型之總結與討論

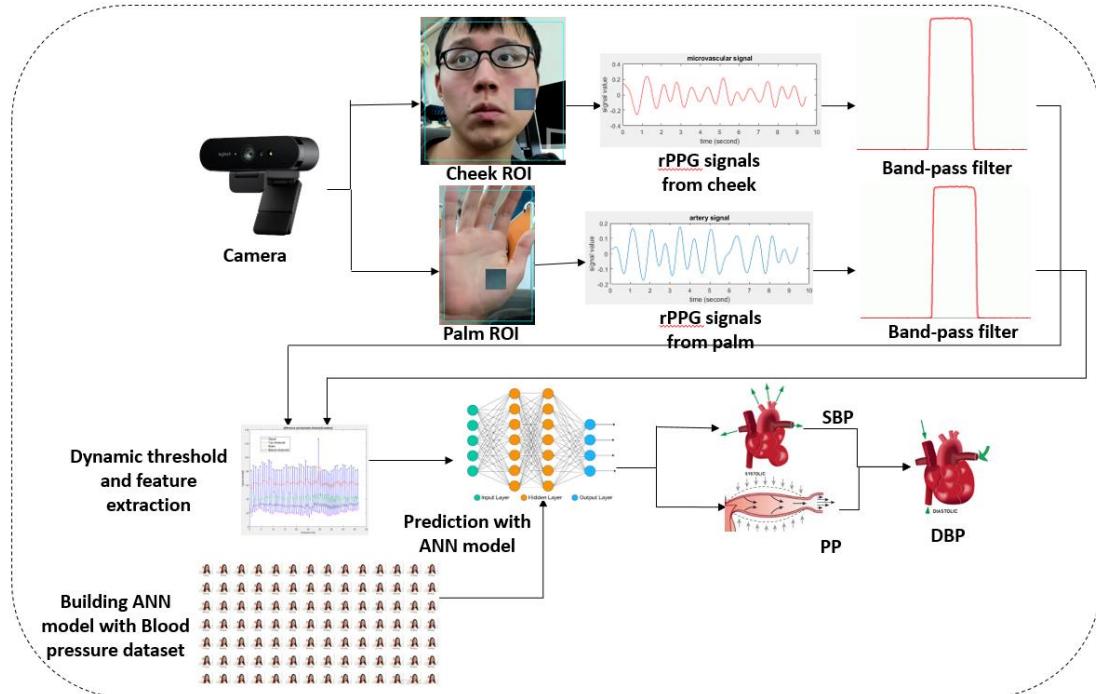


圖 18、臉手影像式血壓量測之整體流程架構圖

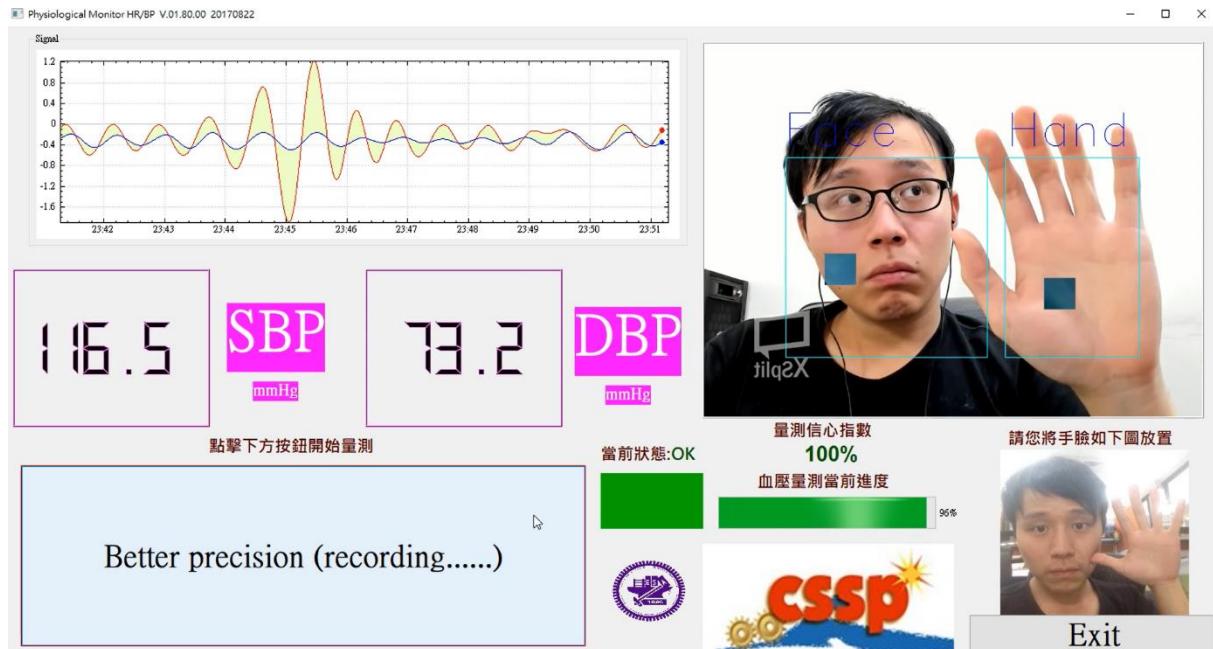


圖 28、臉手影像式血壓量測介面

本論文提出之血壓量測模型首先透過受測者之臉頰與手部之間的影像進行訊號擷取。使用了包含色度訊號[14]以及皮膚正交平面法[15]提升訊號品質，並結合濾波範圍為 0.5 Hz 到 3.35 Hz 之有限脈衝響應帶通濾波器將高頻雜訊除去。並且使用動態門檻閾值法選取出正確的脈波訊號峰、谷值，並透過有效特徵組演算法去除受雜訊影響。

的特徵點。擷取 9 個脈衝傳遞時間相關之時域特徵用於臉手影像式血壓的量測。並使用包含 K 最近鄰居法、深度信念網路-深度神經網路以及人工神經網路 3 種不同的模型，透過蒐集之血壓資料庫進行血壓誤差結果的評比，同時並與相關之影像式血壓量測論文[2][31]進行比較。本論文之收縮壓及舒張壓的方均根誤差均較低，並且使用的為通用化模型，相較於[2]之個人化且需校正的模型有更大之應用面。整體臉手影像式血壓量測之流程圖如圖 18。而圖 28 為基於本論文所實作之臉手影像式血壓量測系統的使用者介面。本系統亦基於人臉偵測以及訊號振幅之變動實作出保護機制，當受測者之人臉於影像中無法被偵測，代表受測者的臉部受到遮擋，圖 28 當中的當前狀態提示便會由綠色轉為紅色，而介面中的量測信心指數值也會跟著下降。同時，若臉頰或手掌之訊號在一個心跳週期當中的振幅過大，超過上門檻，則代表受測者處於晃動的狀態，信心水準亦會下降如圖 30。而同理，若臉頰或手掌之訊號在一個週期當中的振幅過小，低於下門檻，代表受測者的臉頰或手掌並無在訊號擷取區域內，攝影機僅擷取到背景訊號，此時量測信心指數值也會下降。

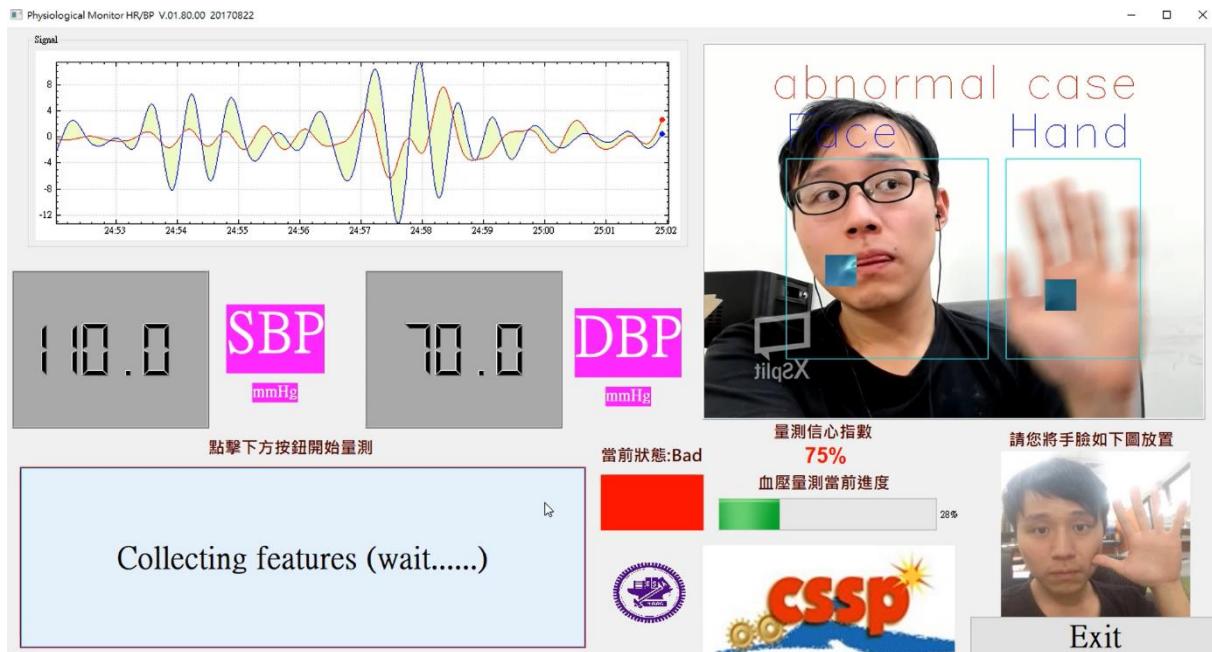


圖 30、異常狀態如臉部及手部晃動時，使用者介面將作出提示，包含調整當前狀態為紅色警示，及調降量測信心指數

2.8、未來展望

透過本論文提及之訊號處理、特徵擷取以及量測模型，本論文已實作出時域即時臉手影像式血壓量測系統，以實時量測收縮壓以及舒張壓。未來將持續修改此系統並進一步收集來自醫院病患的血壓資料庫，用以訓練並更加完善影像式血壓的量測模型。以期能將此系統應用於駕駛生理資訊監測、老人健康長照等場域，以因應病患之突發狀況並提供長期之連續非接觸式血壓監測的解決方案。同時，本論文於血壓資料庫錄製時，亦基於 Likert 量表[99]收集了受測者之疲勞相關狀況如表 4。未來將與相關的疲勞偵測技術結合，以建構完整的車用駕駛生理資訊監控系統。

問題描述/程度	從不	偶爾	普通	經常	總是
1. 我有疲累問題	1	2	3	4	5
2. 我忙一下就會累	1	2	3	4	5
3. 我每天完成不了多少事	1	2	3	4	5
4. 我每天的生活都缺乏能量	1	2	3	4	5
5. 我常常在身體上筋疲力竭	1	2	3	4	5
6. 我很難開始投入一件事	1	2	3	4	5
7. 我沒辦法清晰的思考問題	1	2	3	4	5
8. 我不想做任何事，缺乏動力	1	2	3	4	5
9. 我常常心靈上筋疲力竭	1	2	3	4	5
10. 當我做某件事時，我無法集中	1	2	3	4	5

表 4、Likert 疲勞狀況量表

第三章、臉部影像身體質量指數量測及血壓偏移量校正

脈衝傳遞時間的時域相關特徵對於影像式血壓量測相當重要。然而，不同受測者的血壓仍然受到其他因素，包含受測者的血管順應性、血管彈性係數、動脈管壁厚度等生理參數之影響。因此，不同的受測者會有不同的血壓偏移量[19]。而比較論文如[2][19]皆透過美國食品藥物管理局認證之血壓儀器，於量測前預先進行血壓模型的校正。此額外血壓設備校正的需求，降低了該系統的實用性與方便性。而本論文則透過結合其他的生理參數-即身體質量指數，解決了血壓偏移量校正的問題。

身體質量指數能描述受測者之身形特徵，並且相當程度的影響其健康狀況[73]。同時，身體質量指數與血壓之間也有高度的相關性[74][75]。因此，本論文提出基於臉部影像的卷積神經網路遷移式學習模型，以量測受測者的身體質量指數。並採用身體質量指數相關的特徵，對不同受測者的血壓偏移量進行校正。

3.1、背景

隨著近年來電腦視覺以及相關的影像處理技術的發展，人臉影像式身體質量指數量測成為影像辨識領域新興的主題[43][45]。身體質量指數對健康、慢性病、高血壓皆有相當程度的影響[73][74][75][76]。傳統而言，研究人員欲獲得受測者之身體質量指數，需要求其提供精確的身高與體重資訊進行換算。然而，受測者之身體質量指數資訊在許多情況下無法準確獲得，例如受測者提供之體重資訊可能已過時，或是受測者基於隱私考量不願提供相關資訊。本論文則分別採用了基於7個臉部影像特徵的機器學習模型，以及基於Resnet-50卷積神經網路[22]的遷移式學習[21]模型，透過人臉影像進行身體質量指數的量測，並將結果應用於血壓偏移量的校正。

3.2、基於人臉特徵之身體質量指數量測模型及使用之資料庫

本論文使用作人臉影像式身體質量指數量測之資料庫為北卡羅萊那大學威名頓分校(UNCW)提供之 Morph II 縱向數據庫(Morph II longitudinal dataset)[44]學術版。此資料庫包含蒐集自 13000 餘人的共 55000 張人臉影像，以及受測者對應的身高、體重及換算的身體質量指數等生理資訊。此節將首先介紹基於 7 個人臉影像特徵進行身體質量指數量測的模型。

3.2.1、人臉影像特徵擷取

本論文使用的共 7 個臉部影像特徵包含顴寬額寬比 CJWR (cheekbone to jaw width ratio)、臉寬臉高比 WHR (width to upper facial height ratio)、周長面積比 PAR (perimeter to area ratio)、眼睛大小 ES (eye size)、下臉臉高比 LF/FH (lower face to face height ratio)、臉寬下臉高比 FW/LFH (face width to lower face height ratio) 以及平均眉眼距 MEH (mean of eyebrow height)如圖 19。

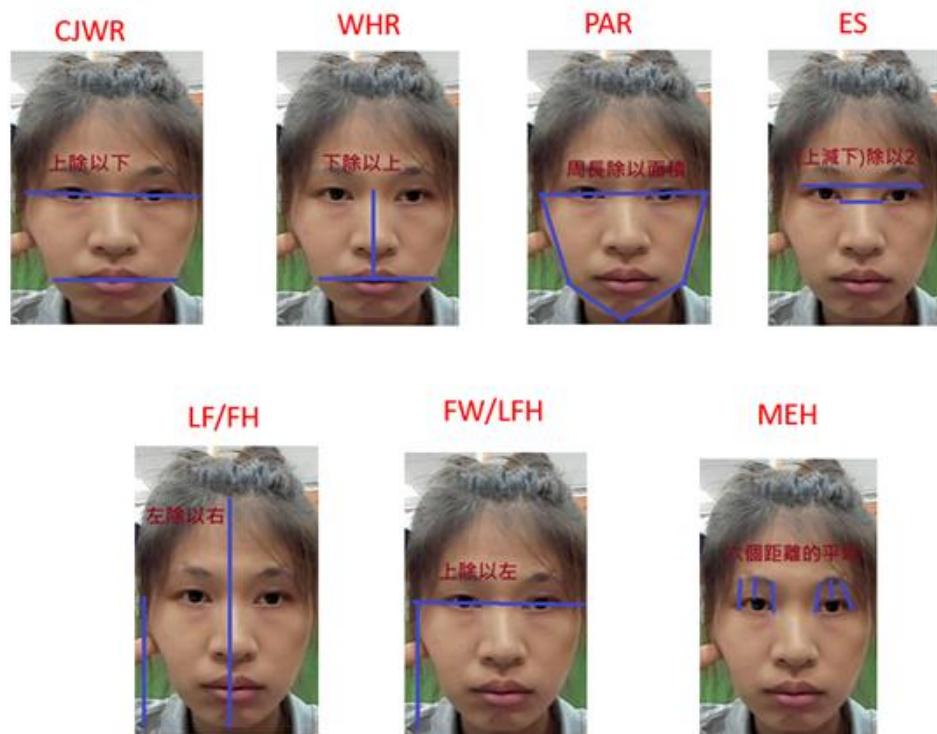


圖 19、7 個身體質量指數相關的人臉影像特徵(CJWR、WHR、PAR、ES、LF/FH、FW/LFH、MEH)

這些特徵為使用 dlib 之臉部 68 個特徵點模型[77]擷取而得如圖 20。使用特徵點座標以及幾何的內差與外差法便得以標出如圖 19 之 7 個臉部特徵用於進行身體質量指

數之計算。其中，顴寬額寬比 CJWR 特徵當中的顴寬為圖 20 當中特徵點 1 及特徵點 17 之間之距離；而額寬則為特徵點 5 及特徵點 13 之間的距離。臉寬臉高比 WHR 當中

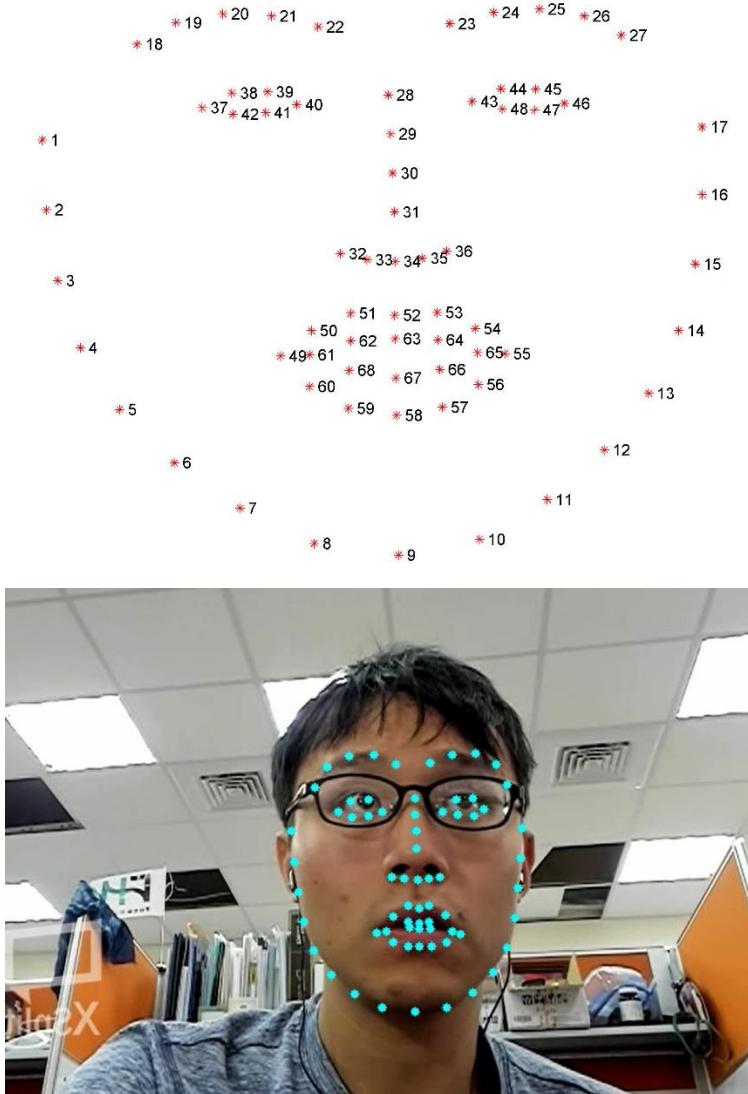


圖 20、dlib 人臉 68 個特徵點模型，以及使用其偵測人臉特徵點的結果

的臉寬則為特徵點 5 及特徵點 13 之間的距離；而臉高則為特徵點 28 與”特徵點 63、67 之中心點”的距離。周長面積比 PAR 則為特徵點 1、5、9、13、17 所形成的五邊形之周長與面積的比例。眼睛大小 ES 則為”特徵點 37 與特徵點 46 之距離”及”特徵點 40 及特徵點 43 之距離”兩者的差值。下臉臉高比 LF/FH 中的下臉高為特徵點 1 與特徵點 9 於垂直方向的距離；而臉高則為”特徵點 41 與特徵點 22”兩點延伸之延長線以及”特徵點 48 與特徵點 23”兩點延伸之延長線此兩條線的交點，與特徵點 9 之間的距離。此處用於計算臉高的額頭點由於沒有對應的 dlib 特徵點，因此採用上述的幾何方式獲得。臉寬下臉高比 FW/LFH 當中的臉寬為特徵點 1 與特徵點 17 之距離；而下臉高則為特徵點 1 與特徵點 9 於垂直方向的距離。平均眉眼距 MEH 則為”特徵點 37 與 18 之距

離”、”特徵點 38 與 39 之中心點，與特徵點 20 之距離”、”特徵點 40 與特徵點 22 之距離”、”特徵點 43 與特徵點 23 之距離”、”特徵點 44 與 45 之中心點，與特徵點 25 之距離”以及”特徵點 46 與特徵點 27 之距離”這 6 者的平均。而擷取 7 個臉部特徵後，便得以透過機器學習模型進行迴歸以量測身體質量指數。

本論文採用之身體質量指數機器學習模型包含線性迴歸 LR(linear regression)、支援向量迴歸 SVR(support vector regression)以及高斯過程迴歸 GPR(Gaussian process regression)，基於上述之 7 個臉部特徵進行迴歸預測。

3.2.2、線性迴歸 (LR)

其中，線性迴歸為較傳統的演算法，透過最小平方迴歸法[78]描述此 7 個特徵與身體質量指數之間的線性關係。此處假設對應的臉部特徵向量(大小為 Nx7)為 x ，而與特徵對應的訓練資料集之身體質量指數為 y (大小為 Nx1)。則可透過最小平方法計算出權重矩陣 W (大小為 7x1)如下(30):

$$W = (x^T x)^{-1} x^T y \quad (30)$$

獲得權重矩陣 W 後即可透過式(31)進行身體質量指數的量測：

$$y = xW + \delta \quad (31)$$

其中，測試資料集之預測結果則為 $y' = xW$ ，而 δ 為誤差項。使用此演算法針對 Morph II 縱向數據庫算出之身體質量指數的平均絕對誤差為 3.35 kg/m^2 。

3.2.3、支援向量迴歸 (SVR)

支援向量機迴歸法[79]則為 Drucker 於 1997 年提出的新興機器學習演算法。是一種監督式學習的演算法，透過不同的核函數(kernel function)將輸入的特徵投影至高維空間進行迴歸。假設對應的臉部特徵向量(大小為 Nx7)為 x ，而與特徵對應的訓練資料集之身體質量指數為 y (大小為 Nx1)。支援向量迴歸之目標為學習 1 個函數 $f(x)$ ，利用人臉影像特徵 x 以預測與特徵對應的訓練資料集中的身體質量指數 y ，並且確保誤差值小於敏感度 δ 。並使得權重之範數盡可能小以避免過擬合之問題。其中，過高的敏感度 δ 會造成模型的不準確，而過低的 δ 值則會導致過擬合的問題。本論文使用之敏感值 δ 為 0.005。假定預測函數 $f(x)$ 如下式(32):

$$f(x) = w \cdot x + b \quad (32)$$

其中， w 為權重矩陣， b 為偏移量，而”.”運算子代表內積。為了使模型能滿足對敏感度 δ 的要求，並將權重矩陣的範數最小化，此模型之訓練過程可透過凸函數最佳化[80]的方式進行如式(33):

$$\text{最小化 } \frac{1}{2} |w|^2, \text{ 並滿足} \begin{cases} y - w \cdot x - b \leq \delta \\ w \cdot x + b - y \leq \delta \end{cases} \quad (33)$$

欲解決此凸函數最佳化問題，可以透過二次規畫演算法[81]獲得出權重矩陣 w 與偏移量 b 。訓練完成後，即可使用此組參數，針對測試資料集利用人臉影像特徵求得身體質量指數的量測結果。而本論文使用此演算法針對 Morph II 縱向數據庫算出的身體質量指數之平均絕對誤差為 3.22 kg/m^2 。

3.2.4、高斯過程迴歸 (GPR)

高斯過程迴歸[82]是一種能擬合複雜的特徵與目標之間關係的演算法，使用了多維高斯分布的概念。假設對應的臉部特徵向量(大小為 $N \times 7$)為 x ，而與特徵對應的訓練資料集之身體質量指數為 y (大小為 $N \times 1$)。高斯過程將 $y = [y_1 \ y_2 \ \dots \ y_n]$ 當作 n 維高斯分布當中隨機抽取出的 1 個點。透過高斯核函數將人臉影像特徵從低維空間投影到高維空間進行身體質量指數的量測。高斯過程的目標即為訓練一預測函數 $f(x) = x^T w$ ，利用輸入特徵求得對應之輸出 $y = f(x) + \delta$ 。其中， $\delta \sim N(0, \sigma_n^2)$ 是變異數為 σ_n^2 的高斯分布雜訊，而 w 為權重且滿足 $w \sim N(0, \Sigma p)$ 。此處的 Σp 為權重對應的共變異數矩陣。透過訓練集之輸入特徵 x 與輸出結果 y 之間的關係建模，並進一步利用測試集之輸入特徵 x' 求出其對應的輸出 f 之機率分布 $p(f|x', x, y)$ 如下式(34):

$$p(f|x', x, y) = \int p(f|x', w)p(w|x, y)dw = N\left(\frac{1}{\sigma_n^2} x'^T A^{-1} x y, x'^T A^{-1} x'\right)$$

$$A = \frac{1}{\sigma_n^2} x x^T + \sum p^{-1} \quad (34)$$

式(34)當中的第一個等式的含義為首先透過訓練集的輸入特徵 x 與對應輸出 y 訓練出權重矩陣 w ，並進一步透過測試集之輸入特徵 x' 配合前述獲得的權重 w 預測出該特徵對應的輸出。本論文使用此演算法針對 Morph II 縱向數據庫算出的身體質量指數之平均絕對誤差為 3.25 kg/m^2 。

3.3、基於 Resnet-50 卷積神經網路的遷移式學習模型及資料庫擴增法

由於 Morph II 縱向數據庫[44]當中之原始資料影像皆是受測者的正臉，並無旋轉或平移的臉部影像。因此，基於此資料庫建置的模型於實用情況下，若遇到受測者之臉部轉動等狀況將無法準確的預測。為處理此問題，本論文基於原 Morph II 縱向數據庫之影像進行資料擴增。而進行資料擴增前，本論文首先針對資料庫影像使用 Viola Jones 人臉偵測演算法[83]進行人臉偵測如圖 21。並擷取臉部區域影像將大小縮放為 224x224 以用作後續的資料擴增，因為大小為 224x224 的影像適宜用作 Resnet-50 卷積神經網路模型的輸入[84]。



圖 21、使用 Viola Jones 演算法對 Morph II 縱向數據庫影像進行人臉偵測



圖 22、使用仿射變換(旋轉、平移、放大、縮小、反轉)技術針對 Morph II 縱向數據庫之影像進行資料擴增

資料擴增的演算法採用仿射變換(affine transformation)[85]，對原圖進行包含旋轉、平移、放大、縮小以及反轉等操作。而本論文所採用之資料擴增參數如下，包含旋轉人臉影像(轉角介於 -40° 至 40°)、水平平移(平移範圍為影像寬度的0~20%)、垂直平移(平移範圍為影像高度的0~20%)、尺寸變換(放大或縮小影像至原圖的0.8~1.2倍)與水平反轉如圖 22。

透過仿射變換進行資料擴增得以增加資料庫之規模，並為訓練的模型增添對於被旋轉及平移的人臉的處理能力。然而，Morph II 縱向數據庫的影像仍受到其他的雜訊與遮擋的影響。例如有部分的受測者之臉部影像包含了鬍子以及眼罩、口罩等遮擋如圖 35。當影像受到上述之影響時，身體質量指數量測表現的準確度也會下降。這是因為許多和身體質量指數相關的重要影像特徵被這些遮擋所損害。因此，若僅使用透過仿射變換進行資料擴增之資料庫進行身體質量指數的模型建構，則仍會受到此些雜訊及遮蔽的影響，降低模型的預測能力及穩健性。



圖 35、Morph II 縱向數據庫的人臉影像，有部分受到受測者配戴之眼罩、口罩或鬍子的影響

因此，本論文進行資料擴增時，亦採用了隨機抹除技術(random erasing)[46]。隨機抹除技術是 Zhong Zhun 於 2017 年提出的一種可以提升模型對於遮擋的穩健性以及泛化能力(generalization ability)的資料擴增技術。由於若用於訓練模型的資料庫包含的遮擋樣本較少，則經過訓練後得到的模型對於遮擋等情境的處理能力也會比

較差，然而模型的泛化能力受到遮擋的影響甚大。隨機抹除技術可以大幅的降低預測模型受遮擋影響的程度，並且是一種常用於處理卷積神經網路模型的過擬合問題的資料擴增方法。同時亦能使模型更好的描述資料影像特徵的分布，並降低模型受影像雜訊的影響。透過隨機抹除技術，本論文得以確保若所訓練之模型於訓練資料集能取得好的預測結果，則其於測試資料集亦能取得好的預測結果。

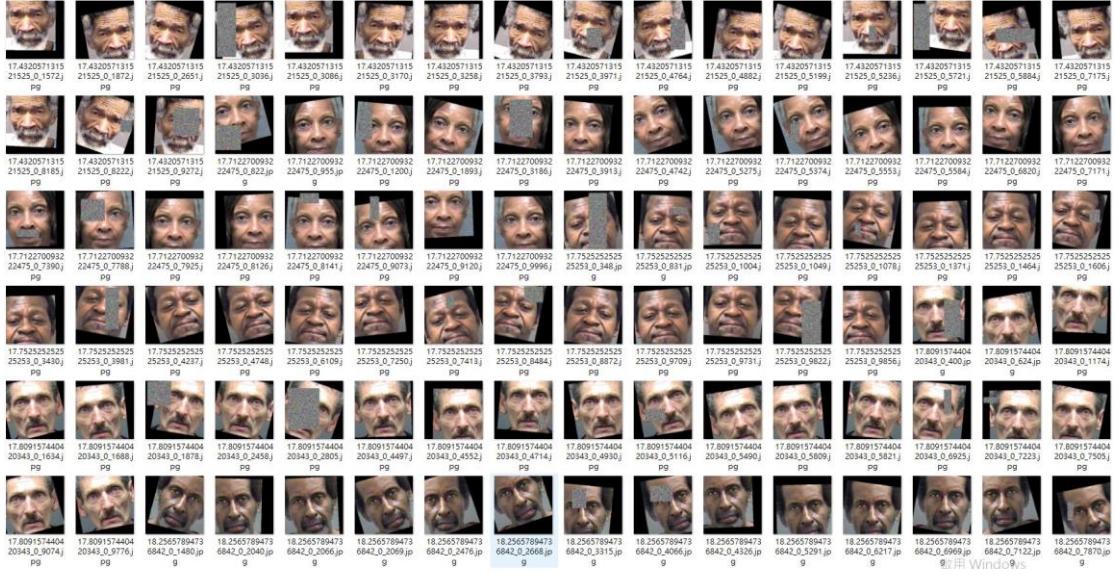


圖 23、結合仿射變換以及隨機抹除技術，針對 Morph II 縱向數據庫之人臉影像進行資料擴增所形成之資料庫

利用隨機抹除技術進行資料擴增時，需以 Morph II 縱向數據庫當中的人臉影像作為輸入，並選擇其中 1 塊長方形區域，將此區域當中的每一個像素的 R、G 和 B 三色影像通道對應的值逐個隨機重新設為 0 至 255 的數值。而本論文進行隨機抹除時採用的參數範圍如下，所選擇之欲抹除的長方形區域之大小為原影像的 0.02 至 0.4 倍；而該長方形之長寬比需介於 0.3 至 3.33 之間；對 1 張影像決定是否進行隨機抹除的機率則為 50%。本論文結合了隨機抹除以及仿射變換技術進行資料擴增，以形成足夠穩健且可以用做身體質量指數量測之人臉影像資料庫如圖 23。而最終生成之資料庫總共包含 80 萬張人臉影像。

獲得了擴增之資料庫後，本論文基於 Resnet-50 卷積神經網路[22]以遷移式學習[21]的方式建構身體質量指數量測的模型架構並進行訓練。而使用作訓練的輸入影像即為大小是 224x224 的由隨機抹除及仿射變換資料擴增法生成的人臉影像。其中，Resnet-50 是一個層數為 50 層的深度殘差網路[86]架構。透過殘差的結構，此模型對於深度神經網路容易產生之梯度爆炸[87]及梯度消失[88]的問題便有了較高的抗性。

殘差網路透過加入了層間的捷徑連結(shortcut connection)，使得原輸入影像包含的特徵資訊得以在層間資訊不斷前饋式的傳遞之後，仍被保留。殘差網路的捷徑連結使用的便是不需參數的恆等映射(identity mapping)的概念如下式(35)：

$$x_{L+1} = x_L + F\{x_L, W_L\} \quad (35)$$

其中， x_{L+1} 以及 x_L 分別代表殘差網路第 L 層的輸出及輸入，而 F 是殘差的函數， W_L 則為此模型第 L 層對應之權重。深度殘差網路的發展，來自 Zhang 和 He 於 2016 年提出的論文[89]，該論文指出欲訓練深度網路，可以先訓練較淺層的網路，並且在該淺層網路上堆上恆等映射層，便可以形成深層網路。由於堆上的層皆為恆等映射，因此深層網路之預測結果理論上不會比淺層網路差。且當此殘差連結架構當中的隱藏單元學到有幫助的資訊時，便能提升網路的表現。

遷移式學習可以將其他神經網路已經訓練好的模型參數轉換到新的模型進行訓練。如此一來，新的模型之訓練將更有效率，亦不用從頭訓練。而本論文基於 Resnet-50 的網路架構[22]進行遷移式學習時，將其頂層用作分類預測的全連接層 fc-1000 移除，替換成對於 1 個迴歸值進行預測輸出的全連接層，並且使用 ReLu[90]作為末層對應之激活函數。同時，使用平滑 L1 損失函數(smooth L1 loss)作為目標函數以建構實用的身體質量指數迴歸量測模型。

建構了整體的模型架構後，本論文使用 Morph II 縱向數據庫學術版所提供的實際身高與體重換算出身體質量指數，作為真實的資料標記值用以訓練迴歸模型。使用的模型之學習率為 10^{-4} ，而學習率的衰減速率則為 10^{-6} 以避免過擬合，同時並將動量(momentum)調至 0.9 以加速模型的收斂過程。而擴增之資料庫則以 8 比 2 的訓練集與測試集比率進行分割，用作身體質量指數量測模型的訓練與分析。而此模型基於擴增資料庫所量測出的身體質量指數之平均絕對誤差為 2.38 kg/m^2 ，相較於比較論文[43]中的 3.14 kg/m^2 表現較佳。

透過此模型預測身體質量指數，本論文得以將此生理資訊用作特徵以進行血壓偏移量的估計並將結果結合至第 2 章的血壓模型。由於身體質量指數和血壓之間的高相關性，因此本論文將其用作特徵，並以實際的血壓作為標記資料，以訓練一個線性回歸之模型以量測不同受測者的血壓偏移量。其中，線性迴歸模型的訓練方法可類比於 3.2.2 小節之說明。結合了身體質量指數進行量測後，血壓量測模型的方均根誤差值可以進一步下降，如 2.6.5 所呈現之完整血壓系統之結果比較。而結合了身體質量指數的整體的臉手影像式血壓量測系統的流程圖如圖 24。

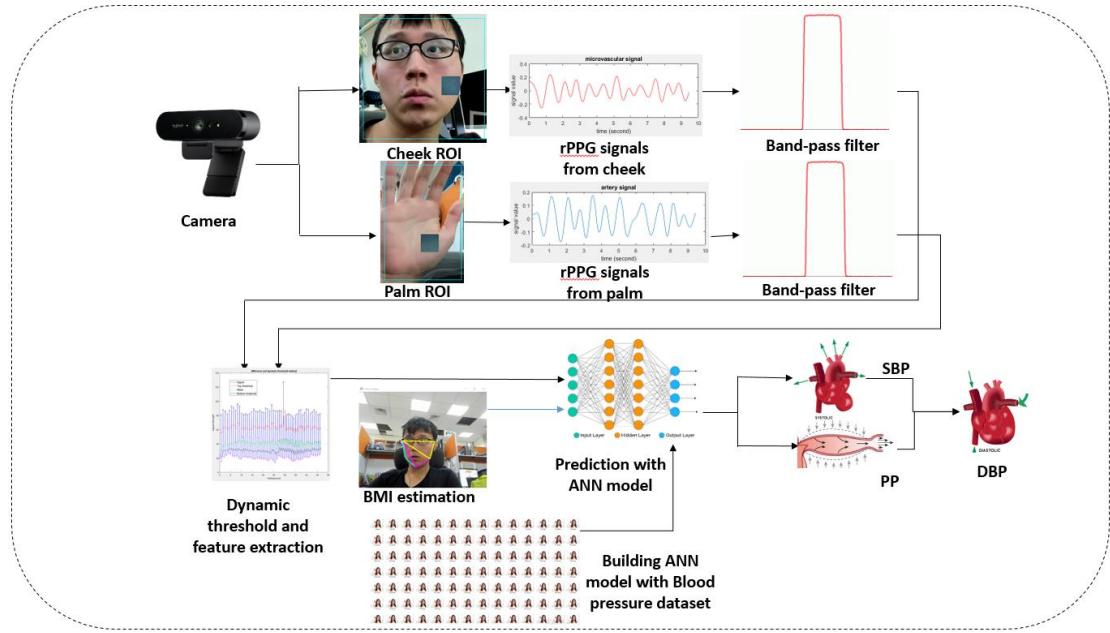


圖 24、將身體質量指數特徵用作血壓偏移量估計的臉手影像式血壓量測系統流程圖

3.4、身體質量指數量測模型之表現與相關論文比較

本論文使用之評比指標為平均絕對誤差(MAE)，即測試資料集當中的所有資料的實際身體質量指數以及模型利用人臉影像特徵所量測出的身體質量指數之間的誤差之絕對值的平均值。平均絕對誤差之公式如下(36):

$$MAE = \frac{\sum_{i=1}^N |y(i) - \hat{y}(i)|}{N} \quad (36)$$

其中，N 為測試資料集之資料筆數，而 $y(i)$ 為資料庫中之第 i 筆資料的實際身體質量指數， $\hat{y}(i)$ 則為模型所量測出之第 i 筆資料的身體質量指數。而本論文針對身體質量指數所選之比較論文為 Wen 和 Kuo 提出的[43]，該論文利用臉部的 7 個影像特徵建立迴歸模型，並使用 Cootes 所提出的主動形狀模型(active shape model)[91]進行面部輪廓以及特徵點的擷取。採用作訓練與測試之資料庫與本論文同為 Morph II 縱向數據庫。據本論文所知，此篇論文是唯一同樣使用 Morph II 縱向數據庫，且僅以人臉影像進行身體質量指數量測之論文。本論文與[43]之比較結果如下表 2:

	使用之 資料庫	該資料庫對應 之資料量	使用之模型	平均絕對誤差	
本論文	Morph II 縱向 數據庫	來自 13000 位 受測者的 55000 張影像	基於 Resnet- 50 卷積神經網 路的遷移式學 習模型	Resnet:	2.38 kg/m ²
身體質量指 數之比較論 文[43]	Morph II 縱向 數據庫	來自 13000 位 受測者的 55000 張影像	基於人臉 7 個 特徵，利用線 性迴歸(LR)、 支援向量迴歸 (SVR)及高斯迴 歸(GPR)所建構 的模型	LR:	3.22 kg/m ²
				SVR:	3.14 kg/m ²
				GPR:	3.20 kg/m ²

表 2、利用 Morph II 縱向數據庫進行身體質量指數量測的準確度，及其與比較論文 [43]之間的結果評比

使用之資料庫相同的狀況下，本論文通過隨機抹除技術及仿射變換，並結合 Resnet-50 卷積神經網路設計遷移式學習的模型，進一步降低了身體質量指數量測模型

之平均絕對誤差。較當前的最高水平的相關論文降低了 24.2%的誤差。儘管 Wen 和 Kuo 所提出之機器學習模型能夠透過人臉影像量測身體質量指數，其對抗人臉影像受遮擋及旋轉或平移的能力仍較低。本論文則透過隨機抹除等技術改善了模型在遮擋等情境下的量測能力。

3.5、人臉影像身體質量指數量測系統及血壓偏移量之總結與討論

本論文首先針對論文[43]所提出的演算法，使用 dlib 的人臉特徵點模型，配合幾何的公式進行臉部特徵擷取，進行實作與比較。使用的特徵包含顴寬額寬比 CJWR (cheekbone to jaw width ratio)、臉寬臉高比 WHR (width to upper facial height ratio)、周長面積比 PAR (perimeter to area ratio)、眼睛大小 ES (eye size)、下臉臉高比 LF/FH (lower face to face height ratio)、臉寬下臉高比 FW/LFH (face width to lower face height ratio) 與平均眉眼距 MEH (mean of eyebrow height)。透過此些人臉影像特徵，利用線性迴歸、支援向量迴歸以及高斯過程迴歸等機器學習演算法進行身體質量指數的量測。

同時，本論文亦提出了一種基於 Morph II 縱向數據庫的人臉影像資料擴增演算法。結合了仿射變換(旋轉、平移、放大、縮小、反轉)以及隨機抹除技術，提升了資料庫訓練出之模型對於遮擋與雜訊等影響的抗性。並且提出了基於 Resnet-50 卷積神經網路的架構進行遷移式學習的模型。透過將原架構最末的全連接分類層移除，並替換為迴歸用的全連接層，本論文得以利用 Resnet-50 的特徵擷取功能進行臉部影像特徵擷取，並且將擷取之特徵用於進行迴歸預測以獲得受測者之身體質量指數。與相關論文[43]進行比較後可看出本論文透過資料擴增之技術，對於人臉影像當中的遮擋與雜訊有較好的抗性，模型預測出的身體質量指數對應之平均絕對誤差也較[43]降低了 24.2%。而透過此身體質量指數特徵，本論文可量測血壓之偏移量以進行校正，並將此校正結果結合至血壓量測模型，進一步提升血壓量測的準確度。

3.6、未來展望

透過本論文所提出之基於深度殘差網路 Resnet-50 的遷移式學習模型，所實作出的系統得以提升人臉影像量測身體質量指數的準確度。身體質量指數是與健康相關的相當重要的指標，未來將進一步結合受測者之半身影像進行身高與體重的量測，使此生理資訊量測系統更加全面。而身體質量指數特徵除可用於輔助量測血壓，亦可用以提供飲食計畫相關之手機應用程式使用，即時將體重與身體質量指數的資訊回饋給使用者。因此，未來將繼續發展更全面的身體質量指數量測系統，並結合身高與體重等生理資訊提供使用者即時的體態管控功能，並且針對肥胖或過瘦的問題發出警訊，以期能協助妥善規畫飲食與治療，降低罹患高血壓等潛在疾病的風險。

第四章、全臉影像式頻域血壓量測模型

臉手影像式的血壓量測模型需透過受測者的臉頰(微血管)以及手掌(橈動脈)作為訊號擷取之來源。此模型之實用性於某些應用場域仍會受限。例如，當車輛駕駛駕車時，需手握方向盤，無法將其手掌靜置於攝影機前方進行血壓量測。因此，本論文亦提出了一個全臉影像式的血壓量測模型欲解決此問題，以提升血壓量測系統在不同情境下的實用性。

然而，進行全臉影像式血壓量測時，訊號擷取的來源則為受測者的額頭與臉頰，此兩處之間的脈衝傳遞時間相較於臉頰與手掌間的傳遞時間小了許多，因此擷取脈衝傳遞時間的時域相關特徵時將會比較困難。本論文於是提出了一種基於頻域特徵的全臉影像式血壓量測系統，得以擷取頻域特徵的資訊進行血壓量測，以取代對時域特徵的需求。本論文使用之頻域演算法所採用之主要特徵為受測者的額頭與臉頰訊號之間的相位差。並透過這些頻率域的特徵進行受測者的收縮壓及舒張壓的量測。

4.1、研究動機與介紹

血壓對於身體健康而言是相當重要的指標。高血壓會導致不同的疾病，包含冠狀動脈疾病[4]、心臟疾病[6]、短暫性腦缺血病變[5]、中風[4]、腎衰竭[7]以及腦退化[8]等相關疾病與併發症。而影像式的連續性血壓量測可以提供高血壓前期以及高血壓的病患方便且可靠的血壓量測功能，並協助其相關的長期照護。

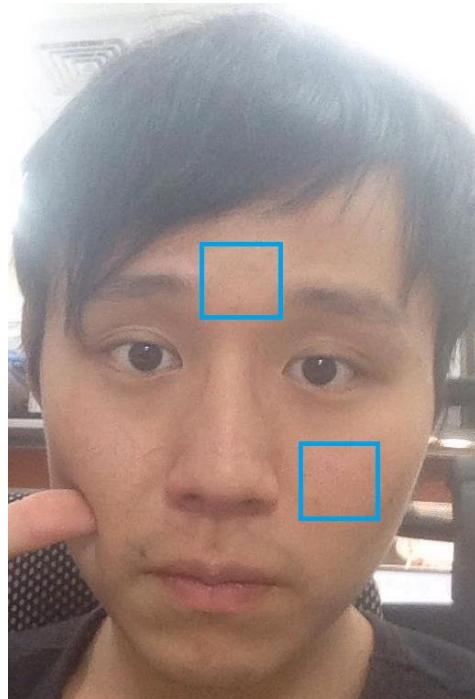


圖 25、全臉影像式血壓之訊號擷取區域，包含額頭與臉頰

而透過基於影像式光體積描述訊號的全臉影像式血壓量測系統，將可以持續監控使用者的血壓，而且僅需臉部影像便能計算血壓的特性也拓展了血壓量測系統的應用場域。受測者僅需端坐於 Logitech Brio 90 帖/秒的攝影機之正前方，並確保其額頭與臉頰出現在影像中，即可進行收縮壓以及舒張壓的量測。本論文使用作訊號擷取的區域如圖 25。

以頻域為基礎的全臉影像式血壓量測演算法主要通過移動窗格法[92]實作，本論文所選擇之窗格大小為 10 秒，此大小的窗格足以用來擷取足夠的血壓相關頻域相位差特徵。其中，頻率域的相位差類似於時域上脈衝傳遞時間的概念，脈衝傳遞時間的物理意義為額頭與臉頰之間波形訊號的時間差延遲，似於相位差的核心概念。本論文實作之全臉影像式血壓量測系統每次量測血壓的總時間為 45 秒。在獲得了每個移動窗格對應的頻率域相位差特徵後，使用線性迴歸及 K 最近鄰居法的模型，結合身體質量指數特徵進行全臉式血壓量測。

本論文用作擷取影像式光體積訊號的裝置為 Logitech Brio 90 帖/秒的攝影機，擷取來自受測者的額頭及臉頰的訊號。不須使用其他的生理訊號感測器或其他硬體，僅以攝影機即可完成整體的血壓量測。因此，此全臉影像式的血壓量測系統之應用場域相當廣。

4.2、影像光體積訊號處理以及頻率域特徵擷取

此處使用之影像式光體積訊號同於第二章所提及。使用了包含色度訊號(chrominance trace)[14]以及皮膚正交平面法(plane orthogonal to skin)[15]等訊號擷取技術，透過R、G、B三色通道之線性組合的方式將非心跳相關的雜訊消除，詳見第二章。獲得了原始的影像式光體積訊號後，本論文透過有限脈衝響應(finite impulse response)帶通濾波器對其進行濾波。所採用之通帶範圍為0.5 Hz~3.35 Hz(對應於30心搏/分鐘至201心搏/分鐘)，進而消除高頻的雜訊成分以及低頻的直流偏移量。

獲得了經過濾波的影像式光體積描述訊號後，本論文得以擷取額頭與臉頰訊號於頻率域之相位差特徵。而相位差特徵擷取的方式包含三種演算法，即交叉相關函數法(cross correlation)、能量譜密度法(energy spectral density)以及功率譜密度法(power spectral density)。接著，本論文將分別針對這三種頻率域相位差特徵擷取的演算法依序進行介紹。

首先，計算兩個部位的訊號之頻域相位差特徵時，交叉相關函數法會將兩個訊號的其中之一加上一個額外的相位 θ ，此處以臉頰訊號為例。接著，便計算額頭之訊號以及加上了額外相位 θ 之臉頰訊號之間的交叉相關函數。下一步，不斷的調整加至臉頰訊號的額外相位 θ ，並且找出在 θ 值為多大時，兩個部位的訊號之間的交叉相關函數值最大。則此時的 θ ，即為兩個部位的訊號之間的相位差。而用以計算兩個離散訊號 x_i 及 y_i 之間的交叉相關函數之公式如下(37):

$$(x * y)_i = \sum_j x_j^* y_{i+j} \quad (37)$$

其中， j 為窗格內之所有可能的訊號範圍對應之索引值。因此，此演算法的最終目標便是要將下式(38)針對相位 θ 最大化：

$$E_\tau[x(t)y(t-\tau+\theta)] \quad (38)$$

式中的 $E[\cdot]$ 代表期望值的運算，而 $x(t)$ 及 $y(t)$ 則分別是擷取自兩個不同部位(額頭及臉頰)的訊號，而 τ 則是計算交叉相關函數時的輔助變數， θ 則是加到了訊號 y (即臉頰訊號)的相位。透過不斷調整 θ ，將式38最大化，便可以得到實際的相位差。

而能量譜密度的演算法則是使用頻域的能量頻譜，計算受測者的額頭以及臉頰訊號之間的相位差特徵。而能量譜密度的公式是透過帕塞瓦爾恆等式(Parseval's identity)如下式(39)推導出來的：

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |X(f)|^2 df \quad (39)$$

其中， $x(t)$ 及 $X(f)$ 分別對應於訊號 x 在時間域以及頻率域的表示法。帕塞瓦爾恆等式說明了時域函數值平方之後的積分值等於該函數的傅立葉係數的平方和。因此，可以推導出頻率域的能量譜密度如下式(40):

$$S_{xx}(f) = |X(f)|^2 \quad (40)$$

其中， $S_{xx}(f)$ 是訊號 x 的能量譜密度，而且是與其頻率 f 相關的函數。透過此能量譜密度的定義式，本論文得以分別尋找出兩個部位訊號的傅立葉轉換頻譜強度之絕對值的最大值，及最大值對應的索引值。接著，將兩者的索引值相減便能換算出相位差特徵。

最後，本論文將討論如何使用功率譜密度法擷取受測者的額頭與臉頰之訊號間的相位差特徵。功率譜密度之定義為訊號的自相關函數之傅立葉轉換，而訊號 $x(t)$ 的自相關函數之定義提供如下式(41):

$$\gamma(\tau) = x(t)x^*(t+\tau) \quad (41)$$

其中， $\gamma(\tau)$ 為訊號 $x(t)$ 的自相關函數。因此，訊號 x 的功率譜密度可以定義如下式(42):

$$S_{xx}(\omega) = \int_{-\infty}^{\infty} \gamma(\tau) e^{-j\omega\tau} d\tau \quad (42)$$

其中， $S_{xx}(\omega)$ 是訊號 x 的功率譜密度，亦等於訊號 x 的自相關函數 $\gamma(\tau)$ 之傅立葉轉換。

導出了功率譜密度的公式後，便可以分別尋找出兩個部位訊號的功率譜密度強度之最大值，及最大值對應的索引值。接著，將兩者的索引值相減便能換算出相位差特徵。

4.3、訊號擷取區域之瀏海偵測與警示機制

而由於受測者於額頭訊號擷取過程中，可能發生瀏海遮擋的問題如圖 33，導致訊號品質大受影響，無法準確取得皮膚所反射的光體積訊號。因此本論文亦有針對此狀況提出解決方案。本論文首先針對 Morph II 資料庫[44]，透過 dlib 提供的人臉 68 個特徵點模型擷取出額頭區塊並進行標記如圖 34



圖 33、瀏海阻擋樣例以及相關的偵測機制

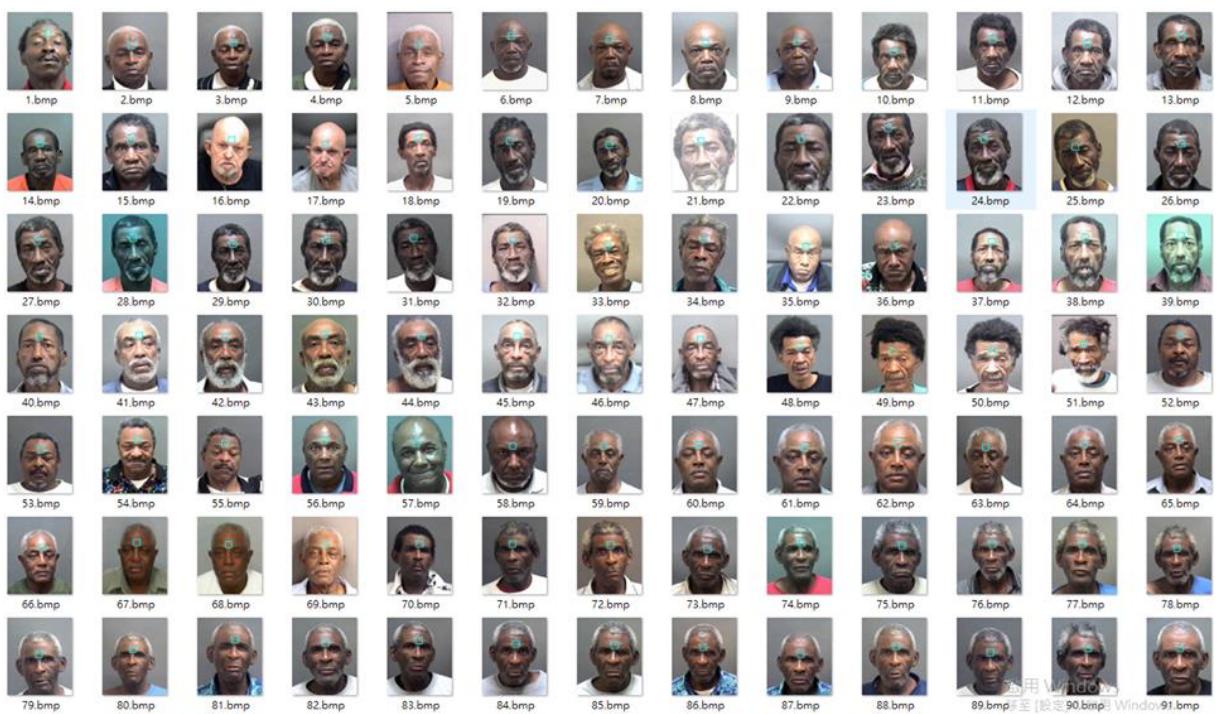


圖 34、將 Morph II 縱向數據庫之人臉影像當中的額頭區域以藍色方框進行標示
 撷取出額頭區塊後，本論文以人工的方式針對 55000 張影像一一判別該區域是否有瀏海或其他遮擋，並進行對應的人工標記。而獲得對應的人臉影像瀏海資料庫後，本論文設計核函數，用以與額頭區域之影像進行卷積運算以擷取出該區域之影像特徵。同時，亦擷取額頭區域影像當中 RGB 通道的平均值以及變異數等特徵用以進行瀏海偵測如下式 46：

$$\begin{aligned}\mu_{R,G,B} &= \frac{1}{SIZE(ROI)} \sum_{x,y \in ROI} ROI_{R,G,B}(x,y) \\ \sigma_{R,G,B}^2 &= \frac{1}{SIZE(ROI)} \sum_{x,y \in ROI} [ROI_{R,G,B}(x,y) - \mu_{R,G,B}]^2\end{aligned}\quad (46)$$

此處 $\mu_{R,G,B}$ 分別為額頭區域影像的 R、G、B 通道之平均值，而 $\sigma_{R,G,B}^2$ 則為額頭區域影像的 R、G、B 通道對應之變異數。而由於瀏海相對於影像而言屬於較高頻的雜訊，因而得以透過這些特徵對瀏海的樣態進行描述。而本論文所使用的核函數之大小為 3x3，包含 Laplacian 核函數如式 47：

$$\text{Laplacian 核函數: } \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad (47)$$

以及 Sobel 的水平、垂直、對角及反對角梯度核函數如式 48：

$$\left\{ \begin{array}{l} \text{Sobel 水平梯度核函數: } \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \\ \text{Sobel 垂直梯度核函數: } \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \\ \text{Sobel 對角核函數: } \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix} \\ \text{Sobel 反對角核函數: } \begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} \end{array} \right. \quad (48)$$

與 Scharr 核函數如式 49：

$$\left\{ \begin{array}{l} \text{Scharr 水平核函數: } \begin{bmatrix} -3 & 0 & 3 \\ -10 & 0 & 10 \\ -3 & 0 & 3 \end{bmatrix} \\ \text{Scharr 垂直核函數: } \begin{bmatrix} -3 & -10 & -3 \\ 0 & 0 & 0 \\ 3 & 10 & 3 \end{bmatrix} \end{array} \right. \quad (49)$$

$$\begin{cases} \text{Scharr 對角核函數: } \begin{bmatrix} 0 & 3 & 10 \\ -3 & 0 & 3 \\ -10 & -3 & 0 \end{bmatrix} \\ \text{Scharr 反對角核函數: } \begin{bmatrix} -10 & -3 & 0 \\ -3 & 0 & 3 \\ 0 & 3 & 10 \end{bmatrix} \end{cases} \quad (49)$$

還有 Prewitt 核函數如式 50:

$$\begin{cases} \text{西北向 Prewitt 函數: } \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix}, \text{北向 Prewitt 函數: } \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \\ \text{東北向 Prewitt 函數: } \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}, \text{西向 Prewitt 函數: } \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \\ \text{東向 Prewitt 函數: } \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \text{西南向 Prewitt 函數: } \begin{bmatrix} 0 & -1 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix} \\ \text{南向 Prewitt 函數: } \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \text{東南向 Prewitt 函數: } \begin{bmatrix} -1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \end{cases} \quad (50)$$

以及 Kirsch 核函數如式 51:

$$\begin{cases} \text{西北向 Kirsch 函數: } \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix}, \text{北向 Kirsch 函數: } \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} \\ \text{東北向 Kirsch 函數: } \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix}, \text{西向 Kirsch 函數: } \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \\ \text{西南向 Kirsch 函數: } \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix}, \text{東向 Kirsch 函數: } \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} \\ \text{東南向 Kirsch 函數: } \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix}, \text{南向 Kirsch 函數: } \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} \end{cases} \quad (51)$$

與線段偵測核函數如式 52:

$$\begin{cases} \text{水平線段偵測核函數: } \begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix} \\ \text{45° 線段偵測核函數: } \begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix} \\ \text{垂直線段偵測核函數: } \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} \\ \text{135° 線段偵測核函數: } \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \end{cases} \quad (52)$$

其中，Laplacian 函數可以擷取出額頭區域影像的二次差分細節，因此能夠獲取出細部的瀏海髮絲等細邊緣特徵。而 Sobel 梯度核函數以及 Scharr 核函數則對應於一階

差分的運算，因此可以用以找出影像當中的粗邊緣特徵。Prewitt 函數則可以被用以取得包含北、東北、東、東南、南、西南、西、西北等 8 個不同方向的梯度特徵，用以針對不同方向的瀏海進行偵測。而 Kirsch 核函數則是一個非線性的邊緣偵測核函數，可用以獲取不同方向的邊緣特徵。而線段偵測核函數則可用以取得寬為單一像素的細節變化特徵。因此，本論文使用以上的 3x3 核函數與額頭區域的影像進行卷積以擷取瀏海相關的特徵以便進行建模。

而本論文所使用的資料庫為 Morph II 縱向數據庫，透過人工的標記結果配合由不同的核函數擷取出的額頭區域影像特徵，本論文得以建立以影像特徵為輸入，而瀏海標記為輸出的二元分類器模型。並使用隨機梯度提升(Stochastic gradient boosting)[107]演算法進行二元分類器模型的訓練。

Boosting 演算法為透過將多個弱分類器結合而形成強分類器的演算法。其中，弱分類器的模型複雜度低，訓練所花的時間亦較低，且低複雜度的特性使其不容易產生過擬合的問題，因此能夠做為 Boosting 演算法的基本學習單元。Boosting 透過訓練多個弱分類器進行組合以做預測。

而梯度提升法便是使用了 Boosting 的特性，並結合了不同的可微分的損失函數進行分類模型的建構，本論文使用的損失函數為 deviance 如下式 53:

$$L(h) = -\frac{1}{n} \sum_{i=1}^n [y^{(i)} \log h_\theta(x^{(i)}) + (1 - y^{(i)}) \log(1 - h_\theta(x^{(i)}))] \quad (53)$$

其中， $L(h)$ 為梯度提升模型對應的損失函數，而 n 則為總樣本數， h 為模型的預測函數， x 為對應的輸入特徵，而 y 則為對應的瀏海標記。透過隨機梯度提升演算法，本論文得以針對 Morph II 縱向數據庫的 55000 張人臉影像資料達到 96.8%的瀏海分類準確率。

藉由瀏海偵測及分類，本論文得以將對應的警示機制結合至全臉影像式血壓量測系統的使用者介面。如圖 33，於額頭區域受瀏海遮擋時，進行提示，以避免瀏海等干擾降低了全臉影像式血壓量測的準確度。

4.4、全臉影像式血壓量測資料庫

全臉影像式血壓使用之資料庫同 2.5 節所述。透過 Logitech Brio 90 帖/秒的 RGB 攝影機錄製 27 位受測者的人臉影像，其中包含 19 位男性，8 位女性，皆簽名同意接受實驗。而實際之收縮壓及舒張壓值則由 Omron 血壓量測儀器所提供之資料庫當中包含 1033 筆資料，其中每筆資料皆包含 45 秒的以 Logitech Brio 90 帖/秒攝影機錄製的影像。並且受測者每次量測血壓之間皆間隔至少 3 分鐘，血壓之量測流程如圖 26。

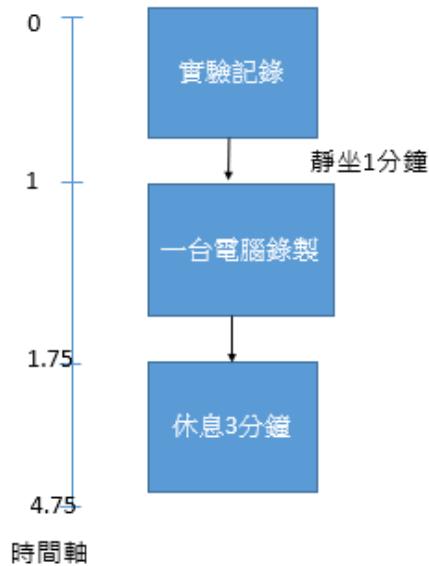


圖 26、血壓資料庫之每位受測者的錄製流程

本論文擷取資料庫當中受測者的額頭與臉頰區域的影像式光體積訊號，並使用有限脈衝響應帶通濾波器進行濾波。接著，便使用 4.2 節提出之演算法擷取兩個部位訊號於頻率域的相位差特徵。並針對所獲得的特徵向量進行 Z 標準化，以消除特徵尺度的差異所造成影響，用以進行血壓的量測。

4.5、全臉影像式血壓量測演算法與模型

本論文分別使用了線性迴歸法以及 K 最近鄰居法進行血壓的量測。相關的演算法分別介紹如下。

4.5.1 線性迴歸法 (Linear regression)

線性迴歸法可以透過輸入特徵與真實結果值之間的線性關係進行建模。本論文使用訓練集當中如 4.2 節所提及的相位差特徵作為輸入 x ，而以血壓資料庫之訓練集的實際血壓作為標記值 y 以訓練全臉影像式血壓量測的模型。此處血壓的線性迴歸關係式如下(43):

$$y = xA + \delta \quad (43)$$

其中， A 是透過最小平方法所訓練出的權重向量，而 δ 則是誤差項。透過最小平方法求出權重 A 的過程如下式(44):

$$A = (x^T x)^{-1} x^T y \quad (44)$$

而透過線性迴歸模型以及式(44)求得迴歸之權重 A 後，利用測試集之特徵 x' 量測血壓 \hat{y} 的方式則如下式(45):

$$\hat{y} = x' (x^T x)^{-1} x^T y \quad (45)$$

式 45 中， x' 為測試集之特徵， x 為訓練集的特徵，而 y 則為訓練集對應的實際血壓。透過由 x 與 y 訓練出之線性迴歸參數，本論文得以利用測試集的輸入特徵 x' 來量測測試集對應的血壓 \hat{y} 。

4.5.2 K 最近鄰居法 (KNN)

類似於本論文之 2.5.1 小節，全臉影像式血壓使用 K 最近鄰居法針對血壓資料庫進行收縮壓與舒張壓模型的建構。首先，計算當前所擷取之特徵向量與資料庫當中所有其他特徵向量的歐幾里得距離(L2 norm)。接著，便找出資料庫中的 K(此處 K=10)個與所擷取之特徵向量的 L2 距離最小的資料，並將這些資料特徵對應的實際收縮壓與舒張壓進行平均，以作為模型的最終量測結果，詳細說明見於 2.5.1 小節。

透過第三章的身體質量指數量測模型，得以估測不同受測者的血壓偏移量，並結合本章之頻率域特徵進行最終的血壓量測。而接著呈現的為僅使用頻率域特徵，而不包含身體質量指數特徵所進行血壓量測的結果。結合身體質量指數之血壓量測實驗結果將呈現於 4.6.4 小節。本論文使用線性迴歸法所量測出的收縮壓之方均根誤差為 14.73 毫米汞柱，而舒張壓的方均根誤差則為 9.72 毫米汞柱。使用 K 最近鄰居法量測

出的收縮壓之方均根誤差則為 13.68 毫米汞柱，而舒張壓之方均根誤差為 9.32 毫米汞柱。其中，本論文進行血壓量測模型的結果評比時，使用作比較的為血壓資料庫當中的 1033 筆資料，而訓練集以及測試集的比例為 8 比 2。

4.6、全臉影像式血壓實驗結果以及與相關專利技術之比較

本論文之臉手影像式血壓的結果以及與臉手影像式血壓相關論文[2]、[31]之比較可見於 2.6 節。而本章將探討全臉影像式血壓量測系統的量測表現與誤差，以及與相關技術進行比較。近年的研究皆是以接觸式光體積描述儀訊號配合心電圖、超音波感測器或是壓電感測器等接觸式的方式進行血壓量測為主。而臉手影像式的血壓量測相關的研究則僅有 2 篇[2][31]。並且目前並無其他相關的全臉影像式血壓量測之論文，因此本論文將與臉部影像式血壓相關的專利技術進行比較。

4.6.1 線性迴歸模型(LR)之血壓方均根誤差

本論文首先針對 4.4 節所提及之資料庫，利用線性迴歸演算法為全臉影像式血壓量測系統進行建模與比較，所量測出之收縮壓的方均根誤差為 14.73 毫米汞柱，而舒張壓之方均根誤差則為 9.72 毫米汞柱。此結果與 K 最近鄰居法相比較差，因為線性迴歸模型僅能描述頻率域特徵與血壓之間的線性關係，無法完整的利用特徵當中的非線性資訊輔助進行血壓的量測。

4.6.2 K 最近鄰居法(KNN)之血壓方均根誤差

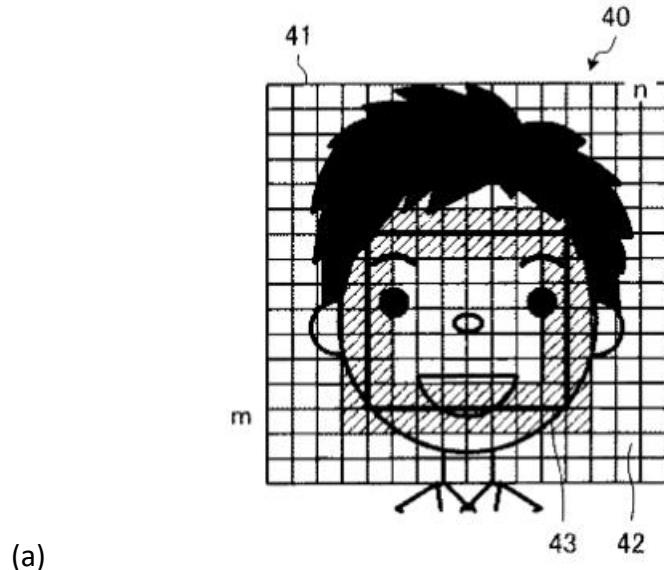
接著，針對基於 K 最近鄰居法之全臉影像式血壓表現進行比較。K 最近鄰居法使用影像式血壓資料庫中擷取之影像光體積描述訊號，所算出之頻率域特徵當中，與當前所得頻率域特徵之間的 L2 norm 最接近的 K 個特徵分別對應的 K 個收縮壓以及舒張壓進行平均以得到最終的收縮壓以及舒張壓的量測結果。本論文使用之 K 值為 10，透過 4.4 節所提及之血壓資料庫進行結果比較。其中，收縮壓之方均根誤差為 13.68 毫米汞柱，而舒張壓之方均根誤差則為 9.32 毫米汞柱。K 最近鄰居法的表現為兩者當中較好的，因為其利用了高維空間當中的距離相似度進行模型的建構，因此可以擷取特徵當中的非線性資訊以量測血壓。

4.6.3 與相關專利技術之比較

由於目前尚無其他僅使用全臉影像進行血壓量測的論文，因此本論文與相關專利所提出的血壓技術進行比較[100][101][102][103][104][105][106]。目前以全臉式影像進行血壓量測的公司僅有日本的富士通株式會社，其主要以發展臉部影像式脈波檢測技術為主，上述之比較專利即為其一系列的臉部影像血壓相關技術。該論文所使用之人臉影像如下圖 32。透過影像擷取裝置，持續擷取受測者之臉部影像，並且將臉部影像切分成 256 個子區域。透過峰值信噪比(peak signal-to-noise ratio)技術擷取

屬於臉部影像的區塊如圖 32 中之深黑色臉部方框 43，並將背景區域去除，達成訊號擷取區域選擇的目的。

計測領域の境界部分の脈波の計測を行う局所領域の一例を示す図



(a)

計測領域の変更の一例を示す図

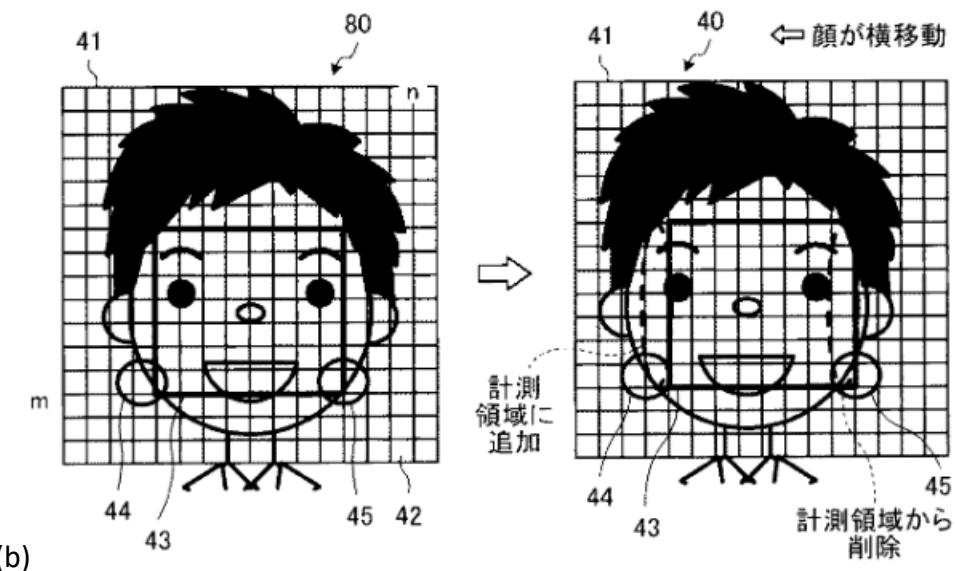


圖 32、(a)臉部影像式脈波檢測技術所使用之訊號擷取區域 (b)臉部訊號擷取區域之
微調與修正[100]

接著，並針對該臉部區域當中的所有子區塊進行脈波的擷取。並且同時擷取邊界區域如圖 32(b)當中的 44 及 45 之脈波訊號。透過邊界區域訊號的峰值信噪比計算，得以進一步微調臉部訊號擷取區域，將邊界屬於臉部影像之訊號納入計算，並去除邊界區域屬於背景的訊號。由於此技術可擷取受測者臉部影像當中的血液體積脈衝，因此

可以用以進行脈波擷取。透過肌膚區域隨著血流量變化所造成的光照變化，得以進行脈波的計算。由於光照的改變相對於脈波成分的影響較有限，因此，該專利透過使用較大的臉部區域進行脈波的擷取，並且提及需要額外的光源照射使用者的面部以提升訊號品質。並擷取如圖 36 的兩個矩形區域之臉部影像[107]對應之脈波訊號，以計算兩處脈波的延遲時間，進而求得脈衝傳遞時間，並且以量測模型透過脈衝傳遞時間的時域特徵量測血壓。然而，其亦有提及臉部眼睛範圍與下巴區域之間的訊號之脈衝傳遞時間約為 100 毫秒。因此，至少需要支援 30 帖/秒以上幀率的影像擷取裝置才能進行脈衝傳遞時間量測。此外，此專利亦提及，若僅使用 30 帖/秒的攝影機，對於脈衝時間週期的變化(低於 10 毫秒)將無法成功偵測，因此擷取之脈衝傳遞時間特徵準確度也會受到影響。而此專利並未提供基於其所實作之系統的準確率，亦未提供相關之數據以供比較，因此本論文僅與其提出的血壓量測技術進行比較。

顔画像の一例を示す図

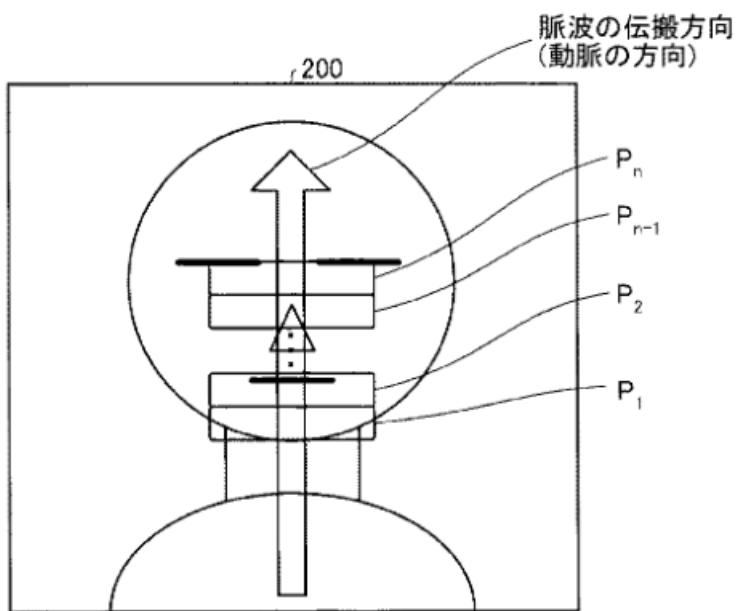


圖 36、偵測出臉部區域範圍後，擷取臉部兩個區域的影像訊號以計算對應的脈衝傳遞時間[107]

本論文之全臉影像式血壓量測為基於頻域特徵的方法，因此能避免時域上臉部之脈衝傳遞時間過短的問題。且仍可透過 4.2 提及的特徵擷取演算法獲得兩個部位訊號之相位差特徵用以進行血壓的量測。對於時域特徵擷取時，無法偵測脈波週期變化的問題也得以解決。且本論文使用結合身體質量指數特徵之演算法模型，得以進行血壓偏移量的校正，以去除不同受測者的血管順應性及血管彈性係數等不同所造成的血壓偏差影響。而加入身體質量指數特徵之量測結果呈現於 4.6.4 小節，本小節、4.6.1 小

節及 4.6.2 小節則為僅使用頻率域相位差特徵進行血壓量測的結果誤差評比。本論文使用來自 27 位受測者的共 1033 筆資料進行血壓量測系統的建模，使用於量測與錄製血壓資料的攝影機則為 90 帧/秒，避免了如[2]使用之高幀率的 420 帧/秒的攝影機，所引起的設備成本過高等實用性方面的問題。因此，本論文所實作之全臉影像式血壓量測系統為一種可持續性、低成本且可靠的血壓量測系統。基於本論文所錄製之血壓資料庫，分別以線性迴歸模型以及 K 最近鄰居法量測血壓，所形成的方均根誤差呈現於表 8。

性質\演算法	線性迴歸模型	K 最近鄰居法
資料庫受測者人數	27 位	
總資料筆數	1033 筆	
每筆資料錄製的影像包含	以 90 帧/秒所錄製之 45 秒影像	
資料庫收縮壓範圍	83~166 毫米汞柱	
資料庫舒張壓範圍	48~105 毫米汞柱	
使用之攝影機以及幀率	Logitech Brio 90 帧/秒	
個人化/通用化	通用化模型	
需額外血壓機校正	不須校正	
可量測之血壓	收縮壓及舒張壓	
訊號擷取區域	額頭、臉頰	
基準血壓量測儀器	歐姆龍血壓機	
收縮壓方均根誤差 (毫米汞柱)	14.73	13.68
舒張壓方均根誤差 (毫米汞柱)	9.72	9.32

表 8、本論文之全臉影像式血壓模型量測結果

4.6.4 基於本論文實作之全臉影像式血壓量測系統的表現

本論文實作之血壓量測系統，除了使用 4.2 節所提及之頻率域相位差特徵以外，亦於第三章實作了身體質量指數量測模型以處理不同受測者之血壓偏移量的問題。

4.6.3 小節所呈現之比較表，為僅透過頻率域相位差特徵，並基於血壓資料庫所進行的血壓量測的誤差之結果。而此小節將呈現之比較表格則是除了頻率域相位差特徵外，亦結合了身體質量指數特徵所訓練之模型。此外，此小節對應的量測結果及誤差為實際使用血壓量測系統，實時針對 14 位不同的受測者各進行 4 次血壓量測所統計出的方均根誤差，而非以血壓資料庫之現有的資料進行的模型準確率評比，因此更能體現全臉影像式血壓模型的量測能力。而使用作為比較基準的血壓量測儀器為歐姆龍血壓機。結合身體質量指數特徵的全臉影像式血壓量測系統實際實驗結果如下表 10。

受測者編號	真實收縮壓 (mmHg)	真實舒張壓 (mmHg)	量測收縮壓 (mmHg)	量測舒張壓 (mmHg)
Subject 1	120	70	107.8	61.0
	115	66	102.1	73.2
	102	67	112.8	72.2
	97	60	108.6	62.1
Subject 2	114	65	113.2	73.4
	112	63	111.7	68.2
	109	64	115.7	71.8
	108	66	113.0	72.5
Subject 3	143	80	121.1	66.5
	140	77	123.1	68.8
	136	78	120.3	64.2
	143	80	121.1	66.5
Subject 4	120	68	131.0	76.3
	115	68	107.0	61.2
	113	66	103.0	77.6
	116	64	102.0	55.5
Subject 5	115	69	123.0	76.4
	115	66	124.5	73.2
	113	67	119.3	76.2
	109	69	118.0	77.0

Subject 6	110	68	119. 8	73. 1
	109	71	118. 2	80. 4
	112	70	119. 1	78. 7
	102	65	115. 7	76. 7
Subject 7	108	66	118. 0	75. 0
	105	68	113. 1	77. 9
	103	64	112. 0	72. 7
	103	65	102. 6	63. 6
Subject 8	125	63	125. 4	74. 4
	122	56	129. 4	66. 4
	122	62	133. 6	72. 3
	126	59	118. 7	65. 2
Subject 9	125	62	128. 4	69. 1
	116	57	126. 8	65. 0
	118	55	105. 2	57. 3
	117	56	125. 4	56. 6
Subject 10	108	64	115. 1	61. 8
	99	60	112. 6	57. 7
	97	61	102. 7	63. 9
	103	62	97. 4	66. 4
Subject 11	102	59	113. 3	68. 3
	98	60	98. 4	60. 9
	101	64	108. 5	69. 3
	100	58	111. 4	58. 9
Subject 12	143	82	122. 6	86. 4
	128	82	136. 7	77. 4
	124	84	132. 3	82. 8
	128	82	137. 7	79. 6
Subject 13	110	62	110. 8	63. 8

	102	58	102.1	59.2
	100	59	100.5	58.9
	98	58	97.8	57.9
Subject 14	114	68	123.6	76.4
	119	74	113.4	81.0
	115	66	120.9	63.6
	121	73	133.8	66.9

表 10、結合身體質量指數特徵之全臉影像式血壓量測系統的實驗結果

上表為實驗結果，而收縮壓之方均根誤差為 10.12 毫米汞柱，舒張壓之方均根誤差則為 7.33 毫米汞柱。本論文透過頻率域進行特徵擷取，解決了時域上脈衝傳遞時間過短的問題。

4.7、全臉影像式血壓模型之總結與討論

本章節中，提出了一個基於全臉影像式的血壓量測系統。首先，透過 Logitech Brio 90 帧/秒的攝影機擷取受測者的額頭以及臉頰區域的訊號。並透過色度訊號法[14]及皮膚正交平面法[15]降低雜訊，以擷取對應的影像式光體積訊號。接著，利用通帶範圍為 0.5 Hz 到 3.35 Hz 的有限脈衝響應(FIR)帶通濾波器對訊號進行濾波，將高頻成分的雜訊以及低頻的直流偏移量濾除。獲得了濾波過的影像式光體積訊號後，便透過 4.2 節提到的包含交叉相關函數法、能量譜密度法以及功率譜密度法，對額頭與臉頰兩個部位的訊號進行頻率域相位差特徵的擷取。

接著，分別以 K 最近鄰居法和線性迴歸模型利用頻率相位差的特徵，針對血壓資料庫進行建模，以用作全臉影像式的收縮壓及舒張壓量測。本論文實作之 K 最近鄰居法之表現較線性迴歸法佳，對應之方均根誤差亦較小。這是因為線性迴歸法僅能呈現出特徵當中的線性資訊，因此量測能力上不如 K 最近鄰居法。K 最近鄰居法使用特徵向量於高維空間中的距離相似度進行量測，因此能將輸入特徵中的非線性資訊保留，用作血壓量測之表現亦較佳。而本論文採用之全臉影像式血壓量測演算法的整體流程圖如下圖 27。

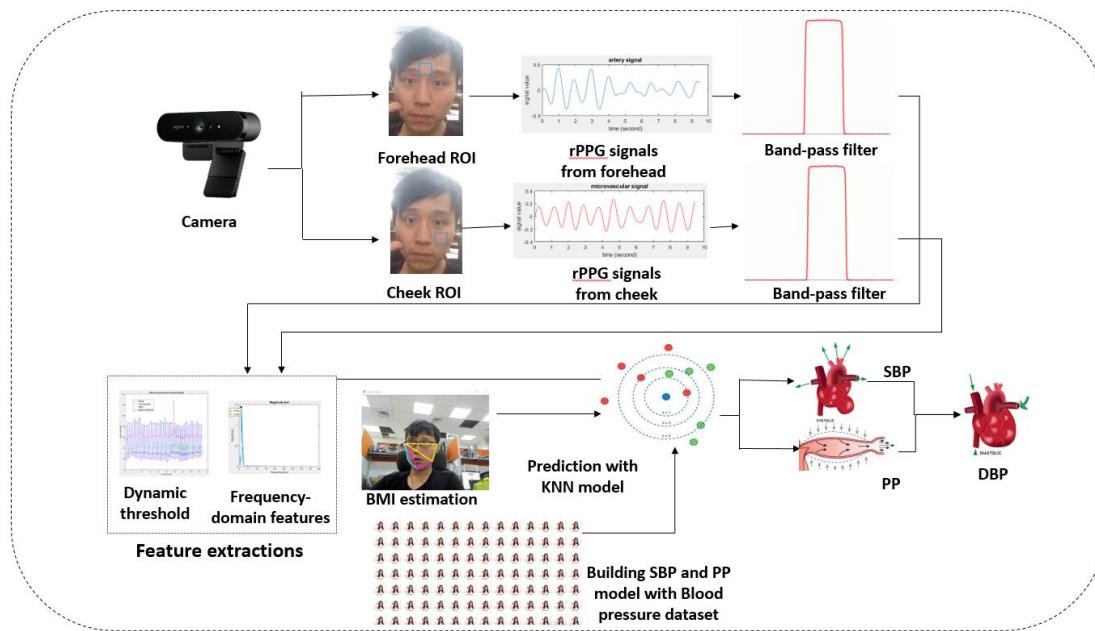


圖 27、全臉影像式血壓量測流程圖

基於此研究，本論文得以實作出即時的全臉影像式血壓量測系統，實時的量測使用者的收縮壓及舒張壓。而由於目前尚無其他全臉影像式血壓的相關論文，因此本論文亦與美國及日本的一些基於臉部影像的血壓相關專利技術進行比較。本論文之特點

與貢獻為引入了頻率域兩個部位訊號之間的相位差特徵，得以去除對時域特徵之高幀率的需求。對於彼此之間脈衝傳遞時間較短的額頭與臉頰訊號，仍可用以量測血壓。而基於本論文所實作之全臉影像式血壓量測系統的使用者介面如圖 29。



圖 29、全臉影像式血壓量測介面

4.8、未來展望

全臉影像式血壓量測系統的應用面廣，且不受臉手影像式血壓的諸多實用上的限制，方便性亦大幅提升。未來將繼續著重於全臉影像式血壓的研究，以期能達成與臉手影像式血壓相近的準確度。而本論文所開發之系統未來將結合嵌入式系統進行實作，並將其應用至車用電子，以協助駕駛之疲勞與生理相關資訊的偵測，期許能提升用路人的安全，減少事故的發生。同時，連續性的全臉影像式血壓量測，也使得高齡照護方面的生理資訊監控需求，得以更方便且持續的進行。未來，本論文實作之系統將針對不同領域之需求進行特化，以提供更全面的全臉影像式血壓量測服務。

第五章、總結與未來展望

5.1、臉手影像式血壓量測演算法之總結

本論文於第二章節提出了一個創新的基於時域特徵的臉手影像式血壓量測演算法，主要使用 9 個與脈衝傳遞時間相關的特徵。首先，針對受測者之臉部(微血管)以及手掌(橈動脈)影像進行訊號擷取。透過色度訊號法[14]以及皮膚正交平面法[15]降低訊號當中非心跳相位的雜訊，擷取出影像式光體積訊號。接著，透過有限脈衝響應(FIR)帶通濾波器進行濾波以去除高頻雜訊和低頻的直流偏移量，以提升訊號品質。獲得了臉頰與手掌的濾波過的影像式光體積描述訊號後，本論文便利用其來擷取時域相關的脈衝傳遞時間特徵。

透過 9 個時域特徵資訊的擷取，得以大程度的保留訊號中所蘊含的脈衝傳遞時間資訊。因此，本論文之連續性血壓量測得以透過 90 幀/秒的 Logitech Brio 攝影機達成。較 Jeong 提出的僅使用單一波峰間時間差特徵與個人化血壓模型的演算法[2]，所使用的 420 幀/秒之攝影機大幅提升了實用性。本論文提出之演算法相較於比較論文[31]能預測之血壓範圍較廣(見 2.5 節)，且所預測之收縮壓以及舒張壓的方均根誤差也較低。

本論文亦基於此演算法實作了一個連續性的臉手影像式血壓量測系統。其中，僅使用 Logitech Brio 90 幀/秒的攝影機即可進行血壓量測，不須額外的生理資訊儀器或接觸式感測器輔助。因此，本系統可支援長期性、低成本且可靠的非接觸式血壓量測功能。

5.2、Resnet-50 與臉部影像量測身體質量指數進行血壓校正之總結

本論文於第三章提出了一個能以受測者的身體質量指數，針對其血壓偏移量 [2][19] 進行校正的模型。首先，本論文使用了 dlib 的臉部特徵點模型，擷取來自受測者的 7 個不同的臉部特徵，以進行身體質量指數的量測。而所使用的臉部特徵包含顴寬額寬比 CJWR (cheekbone to jaw width ratio)、臉寬臉高比 WHR (width to upper facial height ratio)、周長面積比 PAR (perimeter to area ratio)、眼睛大小 ES (eye size)、下臉臉高比 LF/FH (lower face to face height ratio)、臉寬下臉高比 FW/LFH (face width to lower face height ratio) 與平均眉眼距 MEH (mean of eyebrow height)。獲得臉部影像特徵後，便分別透過訓練線性迴歸(LR)、支援向量迴歸(SVR)以及高斯過程迴歸(GPR)模型，進行身體質量指數的量測。

此外，本論文亦提出了一種基於 Resnet-50 卷積神經網路的遷移式學習模型，利用受測者的人臉影像量測其身體質量指數，且所量測出的身體質量指數之平均絕對誤差相較於比較論文[43]低了 0.76 kg/m^2 (降低了原誤差的約 24.2%)。而使用作評估以及測試的資料庫則為 Morph II 縱向數據庫。本論文針對原 Morph II 縱向數據庫之人臉影像，透過仿射變換以及隨機抹除等技術進行資料擴增，以降低遮擋以及雜訊等問題造成的影響，並且避免訓練過程發生過擬合的問題。介紹了身體質量指數量測演算法後，本論文亦提出了基於身體質量指數，對血壓偏移量進行校正的模型。並說明如何將身體質量指數特徵，結合進入血壓的量測系統，以提升血壓量測的準確度。

5.3、全臉影像式血壓量測演算法之總結

本論文於第四章提出了一種新穎的全臉影像式血壓量測演算法，僅透過輸入的 90 帖/秒的受測者臉部影像便得以量測受測者的血壓。首先，擷取來自受測者的額頭以及臉頰的訊號，並透過色度訊號法[14]以及皮膚正交平面法[15]，以 R、G、B 三色通道之線性組合的方式將非心跳的雜訊濾除，並獲得影像式光體積描述訊號。並利用通帶範圍為 0.5 Hz 至 3.35 Hz 的有限脈衝響應(FIR)帶通濾波器對其進行濾波。獲得了濾波過的影像式光體積描述訊號後，本論文利用其來擷取頻率域相關的相位差特徵，以進行全臉影像式血壓的量測。

在第四章，本論文提出基於頻率域的三種相位差特徵擷取演算法。包含交叉相關函數法(cross correlation)、能量譜密度法(ESD)以及功率譜密度法(PSD)。進行特徵擷取後，本論文分別使用線性迴歸法以及 K 最近鄰居法對全臉影像式血壓資料庫進行建模，以用作血壓之量測。此演算法是一個使用 Logitech Brio 90 帖/秒的攝影機所擷取之臉部影像，進行頻域特徵擷取以量測血壓的新穎的演算法。

基於此演算法，本論文實作出一個全臉影像式的連續性血壓量測系統，得以應用於許多情境，包含車用電子以及駕駛生理狀態的監控。本論文之優勢在於僅使用 Logitech Brio 90 帖/秒的攝影機擷取臉部影像，不需額外的生理監測儀器或接觸式感測器，亦不需要受測者配合將手靜置於攝影機前，大大提高了此系統的方便性與實用性。透過人臉影像即可完成血壓的量測。因此，此系統提供了持續性、低成本、方便且可靠的全臉影像式血壓量測功能。

5.4、未來展望

本論文將嘗試使用不同的機器學習模型，並調整參數以改善血壓量測的準確度。舉例而言，未來將研究並使用隨機森林演算法(random forest)[93]、隨機梯度下降法(stochastic gradient descent)[94]以及 XGBoost[95]訓練血壓模型以進行量測。此外，未來亦將嘗試使用循環神經網路(recurrent neural network)[96]的架構進行血壓模型的訓練。循環神經網路的記憶特性[97]，使其能較好的處理如脈衝傳遞時間等的時域相關的特徵資訊。此外，儘管本論文目前所提出的演算法仍需要受測者於血壓量測的過程中，靜坐於攝影機前，但未來將著力於發展對於小幅度的晃動具有抗性的更加穩健的血壓量測系統。期許能將全臉影像式的身體質量指數量測系統以及連續性影像式血壓量測系統的功能優化，並應用至不同場域。提供駕駛生理資訊監控及高齡長照等場域一種更加方便的生理資訊監控系統的解決方案。

參考文獻

- [1] Yoon, Y., Cho, J. H., & Yoon, G. (2009). Non-constrained blood pressure monitoring using ECG and PPG for personal healthcare. *Journal of medical systems*, 33(4), 261-266
- [2] Jeong, I. C., & Finkelstein, J. (2016). Introducing contactless blood pressure assessment using a high speed video camera. *Journal of medical systems*, 40(4), 77
- [3] Brunner, H. R., Laragh, J. H., Baer, L., Newton, M. A., Goodwin, F. T., Krakoff, L. R., ... & Bühler, F. R. (1972). Essential hypertension: renin and aldosterone, heart attack and stroke. *New England Journal of Medicine*, 286(9), 441-449
- [4] Weber, T., Lang, I., Zweiker, R., Horn, S., Wenzel, R. R., Watschinger, B., ... & Metzler, B. (2016). Hypertension and coronary artery disease: epidemiology, physiology, effects of treatment, and recommendations. *Wiener klinische Wochenschrift*, 128(13-14), 467-479
- [5] Whisnant, J. P., Niall, E. C., & Elveback, L. R. (1978). Carotid and vertebral-basilar transient ischemic attacks: Effect of anticoagulants, hypertension, and cardiac disorders on survival and stroke occurrence—A population study. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 3(2), 107-115
- [6] Levy, D., Larson, M. G., Vasan, R. S., Kannel, W. B., & Ho, K. K. (1996). The progression from hypertension to congestive heart failure. *Jama*, 275(20), 1557-1562
- [7] Atkins, R. C. (2005). The epidemiology of chronic kidney disease. *Kidney international*, 67, S14-S18
- [8] Cacciatore, F., Abete, P., Ferrara, N., Paolisso, G., Amato, L., Canonico, S., ... & Rengo, F. (1997). The role of blood pressure in cognitive impairment in an elderly population. *Journal of hypertension*, 15(2), 135-142
- [9] O'Brien, E., Waeber, B., Parati, G., Staessen, J., & Myers, M. G. (2001). Blood pressure measuring devices: recommendations of the European Society of Hypertension. *Bmj*, 322(7285), 531-536
- [10] Yang, B. H., Zhang, Y., & Asada, H. H. (2002). U.S. Patent No. 6,413,223. Washington, DC: U.S. Patent and Trademark Office
- [11] Geddes, L. A., Voelz, M. H., Babbs, C. F., Bourland, J. D., & Tacker, W. A. (1981). Pulse transit time as an indicator of arterial blood pressure. *psychophysiology*, 18(1), 71-74

- [12] Yoon, Y., Cho, J. H., & Yoon, G. (2009). Non-constrained blood pressure monitoring using ECG and PPG for personal healthcare. *Journal of medical systems*, 33(4), 261-266
- [13] Kleiger, R. E., Miller, J. P., Bigger Jr, J. T., & Moss, A. J. (1987). Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *The American journal of cardiology*, 59(4), 256-262
- [14] De Haan, G., & Jeanne, V. (2013). Robust pulse rate from chrominance-based rPPG. *IEEE Transactions on Biomedical Engineering*, 60(10), 2878-2886
- [15] Wang, W., den Brinker, A. C., Stuijk, S., & de Haan, G. (2016). Algorithmic principles of remote PPG. *IEEE Transactions on Biomedical Engineering*, 64(7), 1479-1491
- [16] Feng, L., Po, L. M., Xu, X., Li, Y., & Ma, R. (2014). Motion-resistant remote imaging photoplethysmography based on the optical properties of skin. *IEEE Transactions on Circuits and Systems for Video Technology*, 25(5), 879-891
- [17] Thamarai, P., & Adalarasu, K. (2018). Denoising of EEG, ECG and PPG signals using wavelet transform. *Journal of Pharmaceutical Sciences and Research*, 10(1), 156-161
- [18] Li, X., Chen, J., Zhao, G., & Pietikainen, M. (2014). Remote heart rate measurement from face videos under realistic situations. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4264-4271)
- [19] Yoon, Y., Cho, J. H., & Yoon, G. (2009). Non-constrained blood pressure monitoring using ECG and PPG for personal healthcare. *Journal of medical systems*, 33(4), 261-266
- [20] Dyer, A. R., & Elliott, P. (1989). The INTERSALT study: relations of body mass index to blood pressure. INTERSALT Co-operative Research Group. *Journal of human hypertension*, 3(5), 299-308
- [21] Torrey, L., & Shavlik, J. (2010). Transfer learning. In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques* (pp. 242-264). IGI Global
- [22] Akiba, T., Suzuki, S., & Fukuda, K. (2017). Extremely large minibatch SGD: training resnet-50 on imagenet in 15 minutes. *arXiv preprint arXiv:1711.04325*
- [23] Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. *Physiological measurement*, 28(3), R1
- [24] Kamal, A. A. R., Harness, J. B., Irving, G., & Mearns, A. J. (1989). Skin photoplethysmography—a review. *Computer methods and programs in biomedicine*, 28(4), 257-269

- [25] Crowe, J., Grubb, M., Hayes-Gill, B., & Miles, N. (2014). U.S. Patent No. 8,768,424. Washington, DC: U.S. Patent and Trademark Office
- [26] Jones, P. H., & Wang, W. M. (1992). U.S. Patent No. 5,140,990. Washington, DC: U.S. Patent and Trademark Office
- [27] Teng, X. F., & Zhang, Y. T. (2003, September). Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach. In Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No. 03CH37439) (Vol. 4, pp. 3153-3156). IEEE
- [28] Nitzan, M. (2009). U.S. Patent No. 7,544,168. Washington, DC: U.S. Patent and Trademark Office
- [29] Ding, X. R., Zhang, Y. T., Liu, J., Dai, W. X., & Tsang, H. K. (2015). Continuous cuffless blood pressure estimation using pulse transit time and photoplethysmogram intensity ratio. *IEEE Transactions on Biomedical Engineering*, 63(5), 964-972
- [30] Huang, S. C., Hung, P. H., Hong, C. H., & Wang, H. M. (2014). A new image blood pressure sensor based on PPG, RRT, BPTT, and harmonic balancing. *IEEE sensors Journal*, 14(10), 3685-3692
- [31] Huang, P. W., Lin, C. H., Chung, M. L., Lin, T. M., & Wu, B. F. (2017, November). Image based contactless blood pressure assessment using Pulse Transit Time. In 2017 International Automatic Control Conference (CACS) (pp. 1-6). IEEE
- [32] Fouad, R & Omer, Osama & M. Ali, Abdul-Magid & Aly, Moustafa. (2019). Refining ROI selection for real-time remote photoplethysmography using adaptive skin detection
- [33] Wu, H. Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., & Freeman, W. (2012). Eulerian video magnification for revealing subtle changes in the world
- [34] Tasli, H. E., Gudi, A., & den Uyl, M. (2014, October). Remote PPG based vital sign measurement using adaptive facial regions. In 2014 IEEE International Conference on Image Processing (ICIP) (pp. 1410-1414). IEEE
- [35] Wang, W., Stuijk, S., & De Haan, G. (2014). Exploiting spatial redundancy of image sensor for motion robust rPPG. *IEEE transactions on Biomedical Engineering*, 62(2), 415-425
- [36] Pickering, T. G., Hall, J. E., Appel, L. J., Falkner, B. E., Graves, J., Hill, M. N., ... & Roccella, E. J. (2005). Recommendations for blood pressure measurement in humans and experimental animals: part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the

American Heart Association Council on High Blood Pressure
Research. Circulation, 111(5), 697-716

- [37] Smith, R. P., Argod, J., Pépin, J. L., & Lévy, P. A. (1999). Pulse transit time: an appraisal of potential clinical applications. Thorax, 54(5), 452-457
- [38] cheol Jeong, I., & Finkelstein, J. (2013, July). Potential value of electrocardiogram and photoplethysmogram for non-invasive blood pressure estimation during exercise. In 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 2304-2307). IEEE
- [39] Pollonini, L., Padhye, N. S., Re, R., Torricelli, A., Simpson, R. J., & Dacso, C. C. (2015). Pulse transit time measured by photoplethysmography improves the accuracy of heart rate as a surrogate measure of cardiac output, stroke volume and oxygen uptake in response to graded exercise. Physiological measurement, 36(5), 911
- [40] cheol Jeong, I., Wood, J., & Finkelstein, J. (2013). Using individualized pulse transit time calibration to monitor blood pressure during exercise. Inf Manag Technol Healthc, 190, 39
- [41] Escobar Restrepo, B., Torres Villa, R., & Kyriacou, P. (2018). Evaluation of the Linear Relationship Between Pulse Arrival Time and Blood Pressure in ICU Patients: Potential and Limitations. Frontiers in physiology, 9, 1848
- [42] Saeed, M., Villarroel, M., Reisner, A. T., Clifford, G., Lehman, L. W., Moody, G., ... & Mark, R. G. (2011). Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC-II): a public-access intensive care unit database. Critical care medicine, 39(5), 952
- [43] Wen, L., & Guo, G. (2013). A computational approach to body mass index prediction from face images. Image and Vision Computing, 31(5), 392-400
- [44] Yip, B., Bingham, G., Kempfert, K., Fabish, J., Kling, T., Chen, C., & Wang, Y. (2018, December). Preliminary Studies on a Large Face Database. In 2018 IEEE International Conference on Big Data (Big Data) (pp. 2572-2579). IEEE
- [45] Dantcheva, A., Bremond, F., & Bilinski, P. (2018, August). Show me your face and I will tell you your height, weight and body mass index. In 2018 24th International Conference on Pattern Recognition (ICPR) (pp. 3555-3560). IEEE
- [46] Zhong, Z., Zheng, L., Kang, G., Li, S., & Yang, Y. (2017). Random erasing data augmentation. arXiv preprint arXiv:1708.04896

- [47] Newman, D. L., & Greenwald, S. E. (1978). Validity of the Moens-Korteweg equation. In *The Arterial System* (pp. 109-115). Springer, Berlin, Heidelberg
- [48] Safar, M. E., Henry, O., & Meaume, S. (2002). Aortic pulse wave velocity: an independent marker of cardiovascular risk. *The American journal of geriatric cardiology*, 11(5), 295-304
- [49] Cruickshank, K., Riste, L., Anderson, S. G., Wright, J. S., Dunn, G., & Gosling, R. G. (2002). Aortic pulse-wave velocity and its relationship to mortality in diabetes and glucose intolerance: an integrated index of vascular function?. *Circulation*, 106(16), 2085-2090
- [50] Hughes, D. J., Babbs, C. F., Geddes, L. A., & Bourland, J. D. (1979). Measurements of Young's modulus of elasticity of the canine aorta with ultrasound. *Ultrasonic Imaging*, 1(4), 356-367
- [51] Wang, R., Jia, W., Mao, Z. H., Sclabassi, R. J., & Sun, M. (2014, October). Cuff-free blood pressure estimation using pulse transit time and heart rate. In *2014 12th international conference on signal processing (ICSP)* (pp. 115-118). IEEE
- [52] Makita, S., Fabritius, T., & Yasuno, Y. (2008). Quantitative retinal-blood flow measurement with three-dimensional vessel geometry determination using ultrahigh-resolution Doppler optical coherence angiography. *Optics letters*, 33(8), 836-838
- [53] Pauca, A. L., Wallenhaupt, S. L., Kon, N. D., & Tucker, W. Y. (1992). Does radial artery pressure accurately reflect aortic pressure?. *Chest*, 102(4), 1193-1198
- [54] Aarts, L. A., Jeanne, V., Cleary, J. P., Lieber, C., Nelson, J. S., Oetomo, S. B., & Verkruyse, W. (2013). Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit—A pilot study. *Early human development*, 89(12), 943-948
- [55] Wang, W., Stuijk, S., & De Haan, G. (2015). A novel algorithm for remote photoplethysmography: Spatial subspace rotation. *IEEE transactions on biomedical engineering*, 63(9), 1974-1984
- [56] Saramäki, T., Mitra, S. K., & Kaiser, J. F. (1993). Finite impulse response filter design. *Handbook for digital signal processing*, 4, 155-277
- [57] Association for the Advancement of Medical Instrumentation. (1984). *American National Standard for Cardiac Monitors, Heart Rate Meters, and Alarms (EC13-1983)*. Arlington, VA, ANS1/AAMI

- [58] Lin, Y. C., Wang, Y. J., Cheng, J. C. H., & Lin, Y. H. (2018). Contactless Monitoring of Pulse Rate and Eye Movement for Uveal Melanoma Patients Undergoing Radiation Therapy. *IEEE Transactions on Instrumentation and Measurement*, 68(2), 474-482
- [59] Burrus, C. S., Soewito, A. W., & Gopinath, R. A. (1992). Least squared error FIR filter design with transition bands. *IEEE Transactions on Signal Processing*, 40(6), 1327-1340
- [60] Lin, Y. C., Lin, G. Y., Lin, J. W., Tsai, Y. L., & Lin, Y. H. (2017, December). An image-based pulse rate detection system using an FPGA platform. In *2017 IEEE/SICE International Symposium on System Integration (SII)* (pp. 481-486). IEEE
- [61] El Assaad, M. A., Topouchian, J. A., & Asmar, R. G. (2003). Evaluation of two devices for self-measurement of blood pressure according to the international protocol: the Omron M5-I and the Omron 705IT. *Blood pressure monitoring*, 8(3), 127-133
- [62] Cover, T. M., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1), 21-27
- [63] Hinton, G. E. (2009). Deep belief networks. *Scholarpedia*, 4(5), 5947
- [64] Sutskever, I., Hinton, G. E., & Taylor, G. W. (2009). The recurrent temporal restricted boltzmann machine. In *Advances in neural information processing systems* (pp. 1601-1608)
- [65] Lee, H., Grosse, R., Ranganath, R., & Ng, A. Y. (2011). Unsupervised learning of hierarchical representations with convolutional deep belief networks. *Communications of the ACM*, 54(10), 95-103
- [66] Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010* (pp. 177-186). Physica-Verlag HD
- [67] Glauber, R. J. (1963). Time-dependent statistics of the Ising model. *Journal of mathematical physics*, 4(2), 294-307
- [68] Hinton, G. E. (2002). Training products of experts by minimizing contrastive divergence. *Neural computation*, 14(8), 1771-1800
- [69] Gilks, W. R., & Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 41(2), 337-348
- [70] Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer*, (3), 31-44
- [71] Hush, D. R., & Horne, B. G. (1993). Progress in supervised neural networks. *IEEE signal processing magazine*, 10(1), 8-39

- [72] Goh, A. T. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, 9(3), 143-151
- [73] Stommel, M., & Schoenborn, C. A. (2010). Variations in BMI and prevalence of health risks in diverse racial and ethnic populations. *Obesity*, 18(9), 1821-1826
- [74] Jones, D. W., Kim, J. S., Andrew, M. E., Kim, S. J., & Hong, Y. P. (1994). Body mass index and blood pressure in Korean men and women: the Korean National Blood Pressure Survey. *Journal of hypertension*, 12(12), 1433-1437
- [75] Mungreiphy, N. K., Kapoor, S., & Sinha, R. (2011). Association between BMI, blood pressure, and age: study among Tangkhul Naga tribal males of Northeast India. *Journal of Anthropology*, 2011
- [76] Tirosh, A., Shai, I., Afek, A., Dubnov-Raz, G., Ayalon, N., Gordon, B., ... & Rudich, A. (2011). Adolescent BMI trajectory and risk of diabetes versus coronary disease. *New England Journal of Medicine*, 364(14), 1315-1325
- [77] Koestinger, M., Wohlhart, P., Roth, P. M., & Bischof, H. (2011, November). Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization. In 2011 IEEE international conference on computer vision workshops (ICCV workshops) (pp. 2144-2151). IEEE
- [78] Cochrane, D., & Orcutt, G. H. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *Journal of the American statistical association*, 44(245), 32-61
- [79] Drucker, H., Burges, C. J., Kaufman, L., Smola, A. J., & Vapnik, V. (1997). Support vector regression machines. In *Advances in neural information processing systems* (pp. 155-161)
- [80] Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Mathematics of operations research*, 23(4), 769-805
- [81] Frank, M., & Wolfe, P. (1956). An algorithm for quadratic programming. *Naval research logistics quarterly*, 3(1-2), 95-110
- [82] Rasmussen, C. E. (2003, February). Gaussian processes in machine learning. In *Summer School on Machine Learning* (pp. 63-71). Springer, Berlin, Heidelberg
- [83] Viola, P., & Jones, M. J. (2004). Robust real-time face detection. *International journal of computer vision*, 57(2), 137-154

- [84] Wu, Z., Shen, C., & Van Den Hengel, A. (2019). Wider or deeper: Revisiting the resnet model for visual recognition. *Pattern Recognition*, 90, 119-133
- [85] Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621
- [86] Zagoruyko, S., & Komodakis, N. (2016). Wide residual networks. arXiv preprint arXiv:1605.07146
- [87] Pascanu, R., Mikolov, T., & Bengio, Y. (2012). Understanding the exploding gradient problem. CoRR, abs/1211.5063, 2
- [88] Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(02), 107-116
- [89] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778)
- [90] Li, Y., & Yuan, Y. (2017). Convergence analysis of two-layer neural networks with relu activation. In Advances in Neural Information Processing Systems (pp. 597-607)
- [91] Cootes, T. F., Edwards, G. J., & Taylor, C. J. (1998, June). Active appearance models. In European conference on computer vision(pp. 484-498). Springer, Berlin, Heidelberg
- [92] Macias, J. R., & Exposito, A. G. (1998). Efficient moving-window DFT algorithms. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 45(2), 256-260
- [93] Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, 2(3), 18-22
- [94] Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In Proceedings of COMPSTAT'2010 (pp. 177-186). Physica-Verlag HD
- [95] Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794). ACM
- [96] Graves, A., Mohamed, A. R., & Hinton, G. (2013, May). Speech recognition with deep recurrent neural networks. In 2013 IEEE international conference on acoustics, speech and signal processing (pp. 6645-6649). IEEE

- [97] Sak, H., Senior, A., & Beaufays, F. (2014). Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In Fifteenth annual conference of the international speech communication association
- [98] Saugel, B., Cecconi, M., & Hajjar, L. A. (2019). Noninvasive cardiac output monitoring in cardiothoracic surgery patients: available methods and future directions. *Journal of cardiothoracic and vascular anesthesia*, 33(6), 1742-1752
- [99] Lee, K. A., Hicks, G., & Nino-Murcia, G. (1991). Validity and reliability of a scale to assess fatigue. *Psychiatry research*, 36(3), 291-298
- [100] Akira Fujii (2014). Japanese Patent No. JP2014198200A. Retrieved from
<https://patents.google.com/patent/JP2014198200A/>
- [101] Daisuke Uchida (2014). Japanese Patent No. JP2014193201A. Retrieved from
<https://patents.google.com/patent/JP2014198201A/>
- [102] Masato Sakata (2015). United States Patent No. US9986922B2. Retrieved from
<https://patents.google.com/patent/US9986922B2/>
- [103] Nakata Mikiko (2014). World Patent No. WO2014002276A1. Retrieved from
<https://patents.google.com/patent/WO2014002276A1>
- [104] Tatsuya Mori (2017). World Patent No. WO2017085894A1. Retrieved from
<https://patents.google.com/patent/WO2017085894A1>
- [105] Hidenori Sekiguchi (2014). Japanese Patent No. JP2014200389A Retrieved from
<https://patents.google.com/patent/JP2014200389A>
- [106] Akita Inomata (2014). Japanese Patent No. JP2014188237A Retrieved from
<https://patents.google.com/patent/JP2014188237A>
- [107] Friedman, J. H. (2002). Stochastic gradient boosting. *Computational statistics & data analysis*, 38(4), 367-378.