

利用遠程光體積變化描記圖法於喉部吞嚥動作偵測之研究

¹林聖德, ²羅群智, ³謝欽旭, ⁴洪盟峰

國立高雄科技大學電子工程系

E-mail: [F112152174](mailto:F112152174@nkust.edu.tw), [georgelo](mailto:georgelo@nkust.edu.tw), [csshieh](mailto:csshieh@nkust.edu.tw), [mfhorng](mailto:mfhorng@nkust.edu.tw)@nkust.edu.tw

摘要

遠程光體積變化描記圖法 (Remote Photoplethysmography, rPPG) 是一種透過分析皮膚表面顏色微小變化來捕捉血液動態的非接觸式生理監測技術。過去, 該技術主要應用於心率、呼吸率和血氧等方面, 並展現了創新價值。然而, 吞嚥動作涉及喉部多種精確協調的肌肉運動, 這些運動會對喉部血管造成短暫的壓迫或形變, 導致血流量的變化。這些血流變化會引起喉部皮膚表面反射光強度的細微波動, 使得利用 rPPG 技術捕捉吞嚥過程中的血液流動變化成為可能。本研究探討了 rPPG 技術在偵測喉部吞嚥動作方面的應用。我們利用 Mediapipe 平台、離散小波轉換 (DWT) 和 Yolov8 模型, 開發了一種即時且非接觸性的吞嚥功能評估技術。實驗結果顯示, 此技術能以 92.5% 的成功率捕捉吞嚥過程中的血液流動變化。這不僅提供了一種嶄新的非接觸式吞嚥困難評估方法, 還有望為臨床早期識別和診斷吞嚥障礙提供新的能力。

關鍵字: 遠程光體積變化描記圖法、吞嚥偵測、Mediapipe、離散小波轉換、非接觸性評估

Abstract

Remote Photoplethysmography (rPPG) is a non-contact physiological monitoring technology that captures blood dynamics by analyzing subtle color changes on the skin's surface. Historically, this technology has been primarily applied in areas like heart rate, blood oxygen, and blood pressure, demonstrating innovative value. However, swallowing involves precise and coordinated muscle movements in the throat, which can cause temporary compression or deformation of throat blood vessels, leading to changes in blood flow. These blood flow changes induce subtle fluctuations in the intensity of light reflected from the throat's skin surface, making it possible to use rPPG

technology to capture blood flow variations during the swallowing process.

This study explores the application of rPPG technology in detecting throat swallowing movements. We utilized the Mediapipe platform, Discrete Wavelet Transform (DWT), and the Yolov8 model to develop a real-time and non-contact swallowing function evaluation technology. Experimental results show that this technique can capture blood flow changes during swallowing with a success rate of 92.5%. This not only provides a novel non-contact method for assessing swallowing difficulties but also holds promise for offering new capabilities in the early clinical identification and diagnosis of swallowing disorders.

Keywords: Remote Photoplethysmography, Swallowing Detection, Mediapipe, Discrete Wavelet Transform, Non-Invasive Assessment

1. 前言

近年來, 非接觸式生理訊號監測技術在醫療科技中扮演著重要角色。傳統的生理訊號監測方法, 如心電圖 (ECG) 和血氧飽和度 (SpO₂) 測量, 通常需要直接接觸患者皮膚, 這可能引起不適和抗拒, 特別是在需要長期監測的情況下。因此, 開發非接觸式的生理訊號監測技術成為研究與應用的焦點 [1]。遠程光體積變化描記圖法 (Remote Photoplethysmography, rPPG) 是一種透過分析皮膚表面顏色微小變化來捕捉血液流動的非接觸式技術, 可提取心率、呼吸率和血氧等重要生理資訊 [2]。這項技術不僅提升了患者的舒適度, 還能廣泛應用於遠距醫療、健康監測等需要非接觸監測的領域 [3]。吞嚥是一個高度複雜且精密的生理過程, 涉及喉部多種精確協調的肌肉運動, 這些運動會對喉部血管造成短暫的壓迫或形變, 導致血流量的變化, 進而引起喉部皮膚表面反射光強度的細微波動 [5]。本研究主要是在驗證 rPPG 技術在偵測喉部吞嚥動作方面的可行性。我們開發了一套基於 Mediapipe 平台

和離散小波轉換(Discrete Wavelet Transform, DWT)的訊號處理方法。實驗中，系統透過攝影機捕捉受試者喉部區域的連續影像，並對影像進行 RGB 分量的提取與平均處理，得到皮膚表面反射光強度變化的 RGB 波形。這些波形數據隨後經過 DWT 處理，以提取吞嚥過程中的關鍵特徵，並進一步應用於即時吞嚥動作的偵測，使用 YOLOv8 模型判定是否為吞嚥。

實驗結果顯示，此技術能以 92.5% 的成功率捕捉吞嚥過程中的血液流動變化，從而準確判斷是否正在進行吞嚥。這意味著，這項非接觸式的吞嚥功能評估方法有望成為臨床早期識別和診斷吞嚥障礙的全新工具，進一步提升醫療品質和患者的生活品質。

2. 文獻探討

2.1. 非接觸式生理訊號監測技術

光體積變化描記圖法 (Photoplethysmography, PPG) 是一種基於光學原理的接觸式生理訊號監測技術。其原理是當光源照射皮膚時，光線會穿透多層組織，包括靜脈血、動脈血及其他組織。動脈血的脈動部分隨著心臟的收縮和舒張產生變化，對應於 PPG 訊號中的交流 (AC) 部分，而靜脈血和其他組織則形成穩定的直流 (DC) 部分[1]。如圖 1 所示，PPG 技術通常需要將光源和感測器直接接觸皮膚，以測量血液的脈動訊號，廣泛應用於心率、呼吸率和血氧等生理指標的監測。心率的檢測基於皮膚反射光的微小變化，這些變化主要由毛細血管內血液容積的變化引起。當心臟收縮時，血液被泵送至全身，導致毛細血管內的血液量增加，皮膚的反射光量因此減少；在心臟舒張期間，血液量減少，反射光量隨之增加[1][2]。rPPG 是在 PPG 技術基礎上發展而來的非接觸式生理訊號監測技術。rPPG 透過攝影機檢測皮膚表面的顏色微小變化，無需接觸即可提取心率、呼吸率等生理參數[2]。基於 PPG 的原理，rPPG 利用攝影機捕捉影像中的微小顏色變化來推測血液流量，這些變化反映了心臟的脈動活動。

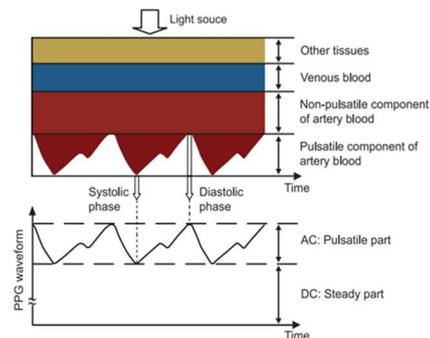


圖 1. 各組織對光衰減的變化[1]

rPPG 技術已被多項研究證實為有效的非接觸式生理訊號監測工具，特別是在心率和呼吸率的測量中展現了廣泛的應用前景[1]。該技術在遠距醫療、個人健康監測以及低光照環境中取得了明顯成果[2]。然而，環境光的變化和人體運動對訊號的干擾仍是其應用面臨的主要挑戰，這些因素可能影響訊號的準確率和穩定性[3][4]。

2.2. 訊號處理技術

為了提高 rPPG 訊號的穩定性和可靠性，各種訊號處理技術被應用於去除環境光和運動雜訊。過濾技術如帶通濾波器、適應性濾波器和卡爾曼濾波器被廣泛應用於此，有效提升了訊號品質。帶通濾波器用於消除低頻和高頻雜訊，保留目標訊號的頻段資訊，適應性濾波器能依據環境的動態變化自動調整參數，從而提高過濾效果，而卡爾曼濾波器則通過估計和校正訊號狀態，有效減少運動干擾，提升訊號品質。此外，訊號增強技術被用來提高皮膚顏色變化的可見性。這包括使用近紅外攝影機在低光照環境中增強光量，應用直方圖均衡化和 Gamma 校正來改善影像對比度，以及採用如 Savitzky-Golay 平滑濾波器減少高頻噪聲，進一步提升訊號的穩定性[3]。

深度學習技術如卷積神經網路 (CNN) 也被引入訊號處理領域，用於學習皮膚反射模式。模型如 PhysNet 和 rPPGNet 能在不同條件下提升訊號處理效果，使 rPPG 更加穩健，適用於日常健康監測。這些技術的應用不僅有效降低了監測過程中的噪聲影響，還提升了訊號的可靠性[2]，同時在不影響患者舒適度的情況下進行監測[6]。而根據過去研究

顯示，吞嚥動作的分析可使用高密度表面肌電(HD-sEMG)訊號來理解吞嚥過程中的肌肉協同活動。通過非負矩陣分解(NNMF)重建肌肉激活模式，結合心率訊號，可以更精確地識別吞嚥動作[5]。

3. 吞嚥偵測設計

3.1. 數據收集

本研究使用 MediaPipe 與 YOLOv8 模型進行即時影像處理與吞嚥動作偵測。透過攝影機捕捉受試者的面部和喉部區域，並聚焦於甲狀軟骨部位。由於吞嚥時甲狀軟骨部位的血管會出現明顯的壓迫和血流變化，這有助於提高偵測的準確率。隨後，透過離散小波轉換(DWT)對 RGB 波形進行過濾和分析，以判定吞嚥事件。研究情境如圖 2 所示。系統能夠即時顯示受試者的影像，數據收集完成後，會生成波形圖，並且可以保存影像數據供後續分析使用。



圖 2.系統應用情境圖

3.2. 系統架構與開發環境

本研究的系統架構由兩個部分組成：即時影像處理與儲存，以及 YOLOv8 吞嚥檢測。前者負責處理即時影像資料並進行儲存，後者則對波形進行吞嚥檢測與標記。在即時影像處理與儲存階段，首先需要輸入受試者的姓名，以便後續驗證。完成後，系統透過 OpenCV 進行影像捕捉，並結合 MediaPipe 進行面部偵測及面部網格定位。當攝影機啟動時，解析度自動設置為 2K (2560x1440)，並依據環境自動調整曝光與白平衡，以獲得最佳影像品質。接著，使用 MediaPipe 偵測面部關鍵點，成功識別面部並鎖定喉部區域後，從該區域提取 RGB 數據進行後續分析。為了確保測試品質，採用

離散小波轉換(DWT)對提取的 RGB 數據進行濾波處理，以消除雜訊並減少干擾。最終，生成吞嚥波形圖並記錄偵測結果。整個過程涵蓋影像捕捉與處理、面部偵測與喉部區域識別、RGB 數據提取以及 DWT 濾波處理，實現了吞嚥事件的偵測與記錄。系統架構如圖 3 所示。

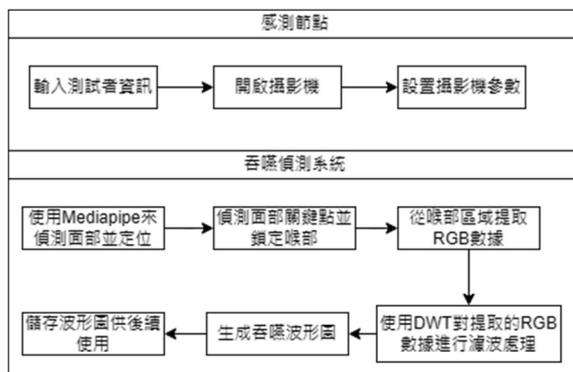


圖 3.吞嚥偵測系統架構圖

在 YOLOv8 吞嚥檢測階段，模型對波形圖進行處理，以便後續辨識吞嚥動作。當將波形圖輸入 YOLOv8 模型時，模型會自動識別並選取一系列重要特徵來標記吞嚥事件。這些特徵通常包括波形的顯著下陷(谷值)、特定的振幅變化以及短時間內的快速波形變化，這些特徵對應於典型的吞嚥動作。在吞嚥過程中，喉部的肌肉運動導致血液流動的變化，這些變化反映在波形中的局部下降和恢復。因此，經過預先訓練的模型能夠學習吞嚥動作的獨特波形模式，例如波谷的深度、波形的週期性以及周圍區域的波動特徵。模型將這些局部特徵視為吞嚥的標誌，並在波形中自動生成目標框來標註偵測到的吞嚥事件。如圖 4 所示，這些目標框清楚地標記出模型認為發生吞嚥的區域。如圖 5 所示，模型的整體流程分為訓練階段與檢測階段。在訓練階段，首先進行數據標註，標記出吞嚥相關的特徵，接著進行模型設置與訓練，使用標註數據來優化 YOLOv8 模型參數，並在模型驗證後保存訓練好的模型以供後續使用。在檢測階段中，首先將波形圖輸入系統，並透過 YOLOv8 模型進行處理生成目標框。結果會被顯示並保存，以便後續分析。最後，針對波形圖進行詳細分析，評估檢測結果的準確率，以確保能夠準確地偵測到吞嚥動作。

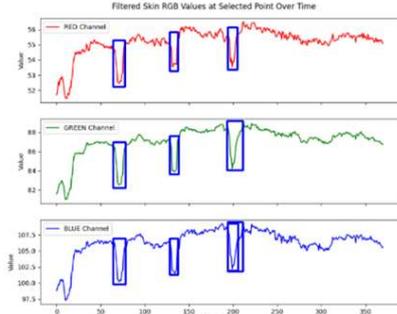


圖 4. 吞嚥標註目標框

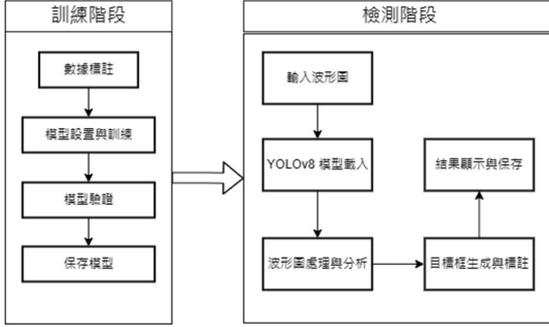


圖 5. Yolov8 模型流程圖

3.3. 波形濾波設計

本研究的波形濾波設計主要是從攝影機捕捉到的 RGB 數據中去除雜訊，提升訊號的可靠性，以便更準確地進行吞嚥事件的偵測與標記。由於攝影機在不同光照條件下會受到環境光變化的影響，可能導致 RGB 數據中雜訊的增加，因此需要採用有效的濾波技術來確保訊號品質。為此，本研究選用了離散小波轉換(DWT)技術進行濾波。DWT 是一種有效的時頻分析方法，可以將原始 RGB 波形分解為不同尺度的分量，從而分離出訊號中的高頻雜訊和低頻成分。在 DWT 分解中，使用了以下公式：

$$X(t) = \sum_k c_{j_0,k} \phi_{j_0,k}(t) + \sum_{j=j_0}^J \sum_k d_{j,k} \psi_{j,k}(t) \quad (1)$$

其中： $X(t)$ 是輸入訊號。 $c_{j_0,k}$ 是尺度係數，代表訊號中的低頻分量（趨勢）。 $d_{j,k}$ 是細節係數，代表訊號中的高頻分量（細節）。 $\phi_{j_0,k}(t)$ 是尺度函數。 $\psi_{j,k}$ 是小波函數。透過將原始 RGB 波形分解為不同尺度的分量，可以有效分離出訊號中的高頻和低頻雜訊。在本研究中，利用 DWT 對 RGB 波形進行多層分解，提取與吞嚥相關的主要訊號，同時去除不必要的高頻雜訊。DWT 能夠將 RGB 數據分解為細節（高頻）和趨勢（低頻）兩部分，主要保

留低頻部分以強調吞嚥過程中的緩慢變化，並去除高頻部分以消除環境光變動及其他瞬態雜訊。重建的過程使用了逆離散小波轉換（Inverse Discrete Wavelet Transform, IDWT）來合成低頻部分的訊號，其公式如下：

$$X(t) = \sum_{j=j_0}^J (c_{j,k} + d_{j,k}) \quad (2)$$

在重建過程中，僅保留對應於吞嚥特徵的低頻分量，從而獲得濾波後的訊號。相比其他濾波方法，如快速傅立葉轉換或傳統的低通濾波器，DWT 的優勢在於能夠同時在時域和頻域中進行分析，提供多尺度的訊號。主要是 DWT 在捕捉吞嚥動作的瞬態特徵方面更加有效，特別是在處理非穩態訊號時，DWT 能夠更好地描述訊號在不同時間點的頻率變化。相較之下，FFT 主要針對整體訊號進行頻率分解，難以應對吞嚥這類具有時間局部變化的訊號。此外，傳統的低通濾波器雖然可以減少高頻雜訊，但可能導致訊號的過度平滑，從而損失與吞嚥動作相關的重要細節。因此，可看出 DWT 在處理非穩態訊號及保留訊號細節方面具有明顯優勢，因此本次研究中選擇 DWT 作為濾波器的主要原因。

3.4. 數據分析與結果評估

在實驗結束後，本研究對收集到的數據進行了詳細分析。首先，將經過濾波處理的 RGB 訊號波形與使用 YOLOv8 模型偵測到的實際吞嚥事件進行對比，以評估系統的偵測準確率。為此，我們檢查 YOLOv8 模型在每次吞嚥事件中生成的目標框是否與波形的特徵變化相匹配。分析結果將以圖表形式呈現，展示不同情境下的吞嚥偵測結果，並對系統的穩定性和靈敏度進行綜合評估。此外，我們對每次實驗的準確率進行了量化分析，例如，透過統計所有實驗中成功偵測到的吞嚥次數與實際吞嚥次數，計算準確率。最後，將經過 YOLOv8 模型檢測並標記的影像資料保存，作為後續結果驗證和系統性能改進的依據。這些影像資料將有助於未來的模型優化和系統可靠性分析。圖 6 展示了整個吞嚥動作偵測的流程，從數據採集、濾波、模型偵測到結果分析的完整過程。

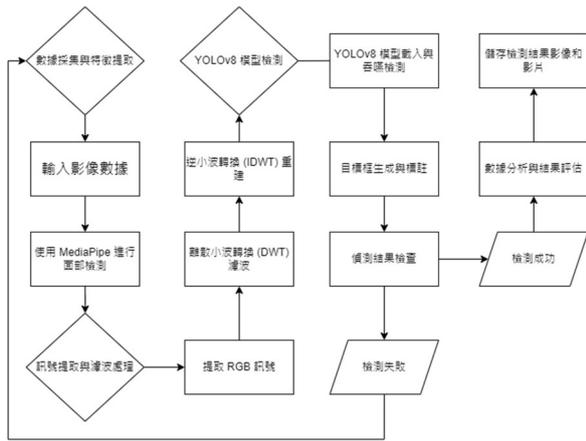


圖 6. 吞嚥動作偵測流程圖

4. 實驗結果

以下將介紹實驗過程中收集的數據分析結果，並評估吞嚥動作偵測的準確率和穩定性。經過訊號處理，得到了與吞嚥過程相關的關鍵訊號特徵。我們將詳細討論這些特徵在實驗中的表現，以及對結果的分析與解釋。實驗的詳細參數設置如表 1 所示。

表 1：實驗參數

參數	值
攝影機解析度	2K
距離	60 公分
視窗大小	1280x720
儲存影片格式	AVI
訊號處理方法	DWT
光照強度	77.3~81.5 Lux

本研究使用的評估指標是吞嚥偵測的準確率，其公式如下：

$$Accuracy = \frac{No.of\ Detected\ Swallows}{Total\ No.of\ Actual\ Swallows} * 100\% \quad (3)$$

No. of Detected Swallows 為模型正確識別的吞嚥事件數量，*Total No. of Actual Swallows* 為受試者實際進行的吞嚥次數。透過比較這兩者，我們可以確定系統對吞嚥動作的識別準確程度，從而評估喉部吞嚥偵測的準確率。

表 2 展示了本實驗的測試結果。實驗中，我們對測試者喉部皮膚反射光強度的 RGB 值進行了詳細記錄和分析。經過數據處理，每個時間點的 RGB 值被平均，生成了 $Mean_{RGB(t)}$ 值。結果顯示，在每次吞嚥事件發生時， $Mean_{RGB(t)}$ 出現明顯變動，這與吞嚥動作引起的血流變化一致。完成數據處理

和離散小波轉換 (DWT) 後，我們進一步分析了吞嚥動作的偵測情況。

表 2 測試結果

性別	年齡	測試次數	吞嚥次數	第一次	第二次	第三次	準確率
男	23	9	3	3	3	3	100%
男	32	9	3	3	3	3	100%
男	22	9	3	3	3	3	100%
男	22	9	3	3	3	3	100%
男	22	9	3	3	2	3	88.8%
女	23	9	3	3	2	2	77.7%
男	26	9	3	3	3	2	88.8%
男	23	9	3	3	3	3	100%
男	22	9	3	3	1	3	77.7%
總體準確率							92.5%

進一步的數據分析顯示，波形可能受到測試者吞嚥力道和頭部抬起角度的影響。不同測試者在進行吞嚥時，力道的大小直接影響喉部對血管的壓迫程度，如圖 7 所示。力道較小的吞嚥動作通常導致血流變化較微弱，這種變化可能不足以超過設定的偵測閾值，從而影響波形的偵測準確率，如圖 8 所示。這種情況在喉部較為厚實或皮膚顏色較深的測試者中更為明顯，因為這些因素會進一步減弱反射光強度的變化。



圖 7. 測試者力道較小時

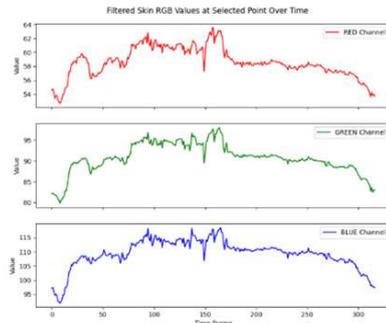


圖 8. 測試結果波形起伏較小

如圖 9 所示，當測試者以較大的力道進行吞嚥

時，喉部肌肉對血管的壓迫較強，導致血流變化更加明顯。這在光反射中產生更為顯著的波峰，從而提高偵測的準確度，如圖 10 所示。



圖 9. 測試者力道較大時

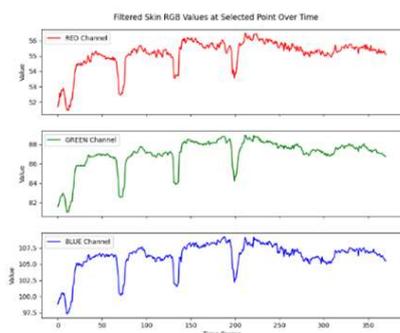


圖 10. 測試結果波形起伏較大

使用 YOLOv8 模型進行吞嚥辨識，可區分真實的吞嚥動作與訊號雜訊的影響。模型識別出吞嚥事件後，會以藍色矩形框進行標註，如圖 11 所示。

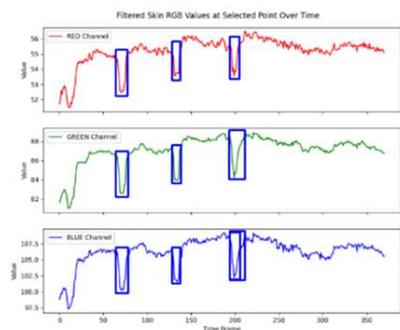


圖 11. 使用 Yolov8 模型來進行吞嚥辨識

5. 結論

本研究提出了一種利用遠程光體積變化描記圖法 (rPPG) 進行喉部吞嚥動作偵測的非接觸式技術。此技術結合了攝影機與訊號處理方法，能夠準確地偵測吞嚥動作。實驗結果顯示，在多次測試中均可取得 92.5% 的準確率。雖說在部分案例中準確率略有下降，例如某些測試者的準確率為 88.8%，一位女性測試者的準確率為 77.7%，這些結果可作為未來進行技術改良時的參考依據。此外，透過實

驗中也發現測試者的喉部結構、吞嚥力道以及外部光線環境等因素亦會影響偵測的準確率，尤其在準確率較低的案例中，這些因素的影響較為明顯。這也表示在實際應用中，系統的穩定性和精確度可能面臨多方面的挑戰，需要進一步的技術優化。未來將著重於提高此技術在不同環境和測試者條件下的適應性，進一步提升系統的可靠性和實用性。

參考文獻

- [1] Y. Sun and N. Thakor, "Photoplethysmography Revisited: From Contact to Noncontact, From Point to Imaging," IEEE Transactions on Biomedical Engineering, vol. 63, no. 3, pp. 463-477, Mar. 2016.
- [2] Z. Yang, H. Wang and F. Lu, "Assessment of Deep Learning-Based Heart Rate Estimation Using Remote Photoplethysmography Under Different Illuminations," IEEE Transactions on Human-Machine Systems, vol. 52, no. 6, pp. 1236-1246, Dec. 2022.
- [3] S. B. Park, G. Kim, H. J. Baek, J. H. Han and J. H. Kim, "Remote Pulse Rate Measurement From Near-Infrared Videos," IEEE Signal Processing Letters, vol. 25, no. 8, pp. 1271-1275, Aug. 2018.
- [4] Fan, Xuanhe and Liu, Fangwu and Huang, Zhijie and Xue, Wei and Gao, Tong and Fu, Jia and Zhang, Jingjing., "Robust Heart Rate Measurement by Adaptive RoI Selection for Head-Rotation Conditions," IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-12, Oct. 2023.
- [5] M. Zhu, O. W. Samuel, Z. Yang, W. Lin, Z. Huang, P. Fang, J. Tan, P. Li, M. C. Tong, K. K. Leung, S. Chen, G. Li., "Using Muscle Synergy to Evaluate the Neck Muscular Activities during Normal Swallowing," 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, , pp. 2454-2457, Oct. 2018.
- [6] P. Y. Chan, J. McNeil, T. Nguyen, N. Ryan, I. Hopper., "Novel wearable and contactless monitoring devices to identify deteriorating patients in the clinical setting: a systematic review protocol," Systematic Reviews, vol. 9, no. 1, May .2020.

誌謝

感謝國家科學及技術委員會(NSTC 113-2622-8-992 -009 -TD1 與 NSTC 112-2221-E-992 -057 -MY3 對本研究之支助，在此誌謝。