

多光譜影像技術及著色 AI 模組於脊椎內視鏡之應用

陳郁瑜	林郁庭	陳兆煒	吳孟晃	張正春
國立臺北科技大學 電機工程系	國立臺北科技大學 電機工程系	衛生部萬芳醫院 神經內科 主治醫師	臺北醫學大學 醫學系骨科學科 專任助理教授 maxwutmu@tmu.edu.tw	國立臺北科技大學 電機工程系 副教授 ccchang@ntut.edu.tw

國科會計畫編號：112-2823-8-027 -003

摘要

內視鏡手術是在醫療診斷上平凡出現的手術之一。以脊椎內視鏡手術為例，當中常見的適應症包含椎間盤突出及半月板受損等。然而，這樣的侵入式手術，經常會在手術過程中發生組織，或是血管出血等情況。其中，出血是內視鏡診斷和治療過程中最常見的併發症之一，而出血引起的血視覺屏障 (blood vision barrier, BVB) 是內視鏡手術中的主要嚴重問題，同時也經常難以準確評估和檢測血視覺屏障中的出血區域，若是無法有效處置血液屏障的根源，便會嚴重提高手術的失敗機率。因此，我們嘗試利用血液在不同光波段下擁有不同穿透度之特性，尋找在血視覺屏障環境下，依舊可以使醫師觀察到影像之近紅外波段，並開發相對應之 AI 著色軟體，即可提供醫師在血視覺屏障環境下繼續進行手術之條件。

關鍵詞：多光譜、光譜成像、內視鏡影像系統、血液屏障、AI 著色。

1. 前言

內視鏡手術常應用於消化系統、泌尿系統、婦科、關節和脊椎等領域。出血是內視鏡診斷和治療過程中最常見的併發症之一[1],[2]。此外，包括出血在內的併發症發生率隨著全球預期壽命的增長而增加，預計將進一步增加。血液視覺屏障 (BVB) 是由於血液積聚而在內視鏡檢查中阻礙視線的現象，也是內視鏡手術的主要嚴重問題[3]。這導致了併發症的風險增加，降低了手術效率，也影響了手術時間、出血量和臨床結果[2]。由於難以看清解剖結構，血液視覺屏障 (BVB) 使外科醫生準確診斷和治療變得具有挑戰性。根據專家的經驗和技術不同，解決血視覺屏障的時間也會有所不同。儘管內視鏡手術對於大多數治療和檢測都是一種出色的工具，但上述挑戰仍然存在。我們相信，可以利用血液的光學特性來減輕解決血視覺屏障的困難。

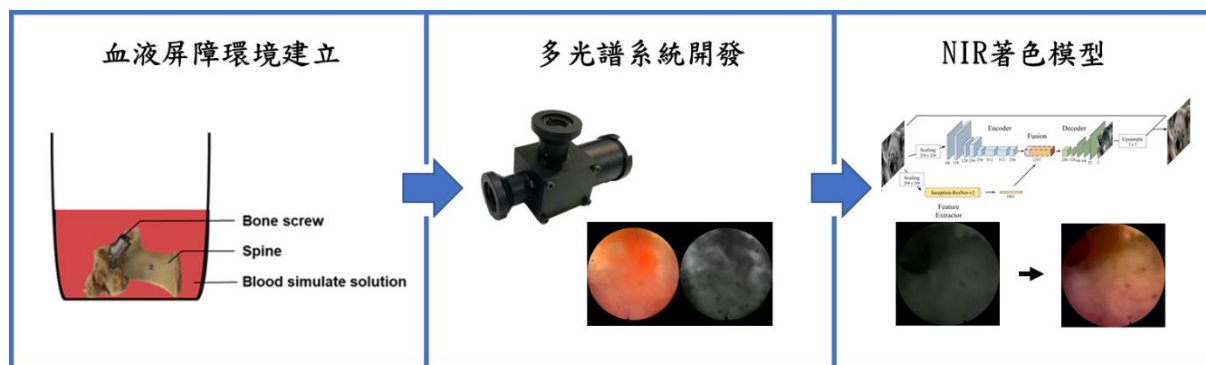
血液屏障是在脊椎內視鏡手術中，醫師經常遇到的一個問題。其發生的原因有許多，包含手術過程中的血管破裂、肌肉出血，或是骨頭出血等。一旦血液濃度到達一定程度，便會造成醫師在脊椎手術的視野模糊不清，失去當前手術方

位，因此稱為血液視覺屏障，又稱血液屏障(blood vision barrier, BVB)。當中血液造成的視覺屏障及視野混濁，其發生率高達 89%。脊椎內視鏡中遇到出血及視野混濁將造成手術時間拉長、無法完成手術、神經血管損傷，甚至嚴重併發症。而組織辨識困難，影響醫師對組織判斷，構造組織損傷以及嚴重副作用產生。

近紅外光(Near-Infrared, NIR)是介於紅外光和可見光之間的電磁波，其波長範圍通常為 700 到 2500 奈米。近紅外光技術已被廣泛應用於化學、生物、農業、食品和環境等領域，其主要用途是進行非破壞性和非接觸式的分析測量。近紅外光技術可以應用於分析物質的組成、結構和性質。由於每種物質的分子結構都具有不同的吸收能力，因此近紅外光技術可以通過分析物質的吸收光譜來識別和量測物質的成分。

以血液為例，血液的吸收光譜在不同的波長下，其中對光的吸收程度也有著差異。例如在 700 到 900 納米的波長範圍內，血液的吸收能力較低，穿透能力較高，可以用來進行非侵入性的生物醫學檢測。因此，紅外光在醫學應用中常常被使用，例如近紅外光譜學技術就是利用紅外光的穿透能力來測量人體組織中的化學成分。

卷積神經網路(Convolutional Neural Network, CNN)是一種特殊的神經網路模型，專門用於處理視覺類任務。卷積神經網路模型基於神經網路模型，通過卷積運算和池化運算等方式，從圖像中提取出特徵，並在此基礎上進行分類、標記、識別等操作。卷積神經網路模型的主要特點在於對圖像中的局部信息進行建模，能夠有效地捕捉圖像中的紋理、形狀、邊緣等特徵，並進一步提取高層次的抽象特徵。卷積神經網路模型通常由多個卷積層(Convolutional Layer)、池化層(Pooling Layer)、全連接層(Fully Connected Layer)等組成。在卷積層中，通過一些卷積核對圖像進行卷積運算，從而提取出局部特徵。在池化層中，通過對局部特徵進行降維操作，從而減少模型的參數數量和計算量。在全連接層中，將所有的特徵融合起來，進行分類或其他操作。卷積神經網路透過這些組成部分學習圖像的特徵和類別，它可以對圖像中的物體進行準確識別。CNN 的應用非常廣泛，在現實生活中，卷積神經網路被用於網路安全、交通監控、疾病診斷等方面，有著非常重要的應用價值。



2. 研究方法

2.1 系統架構

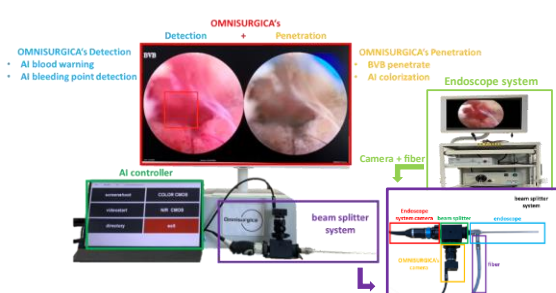


圖 1 整體系統架構

我們在現有的內視鏡系統下，加裝了由我們自己設計分光鏡模組，分光鏡模組可以將原先的內視鏡畫面一分为二，使得兩台相機捕捉到相同的畫面，在分光鏡上我們會安裝 NIR 相機以及原先的內視鏡相機，我們的系統會將兩台相機的畫面整合再一起，並提供 AI 出血點偵測、AI 血液屏障警示等功能。

2.2 全血、血漿和血清的光學性質

全血、血漿和血清都是生物體內重要的液體成分，它們的光學性質對於生物醫學領域中許多光學技術的應用都具有重要意義。全血是由紅血球細胞、白血球細胞、血小板和血漿等成分混合而成的褐紅色液體，血漿是全血在去除細胞成分後得到的液體部分，它主要是由水、蛋白質、鹽類和其他生物分子等組成的。血清是全血在凝固後得到的液體部分，它與血漿相比缺少凝血因子和細胞成分。

全血和血漿的吸收光譜從 590 nm 到 980 nm 的趨勢較平坦且吸收較低，能夠實現較高的光通量傳輸和較低的能量損失於血液視覺屏障中。相反地，血清的吸收光譜在 10%血清濃度組顯示出一個特徵峰在 415 nm，而在較低的血清濃度中影響較小。此外，我們觀察到在 300 至 380 nm 範圍內出現上升的曲線。血清的吸收光譜在 450 nm 至 980 nm 範圍內呈現平坦的趨勢，能夠在此範圍內實現高亮度光通過和較低的能量損耗。由於全血中有 60%是血漿，因此可以推測全血的吸收光譜

與血漿的吸收光譜有關，而血清的吸收光譜則呈現平坦的曲線，其中血漿是血液視覺屏障中的主要因素。

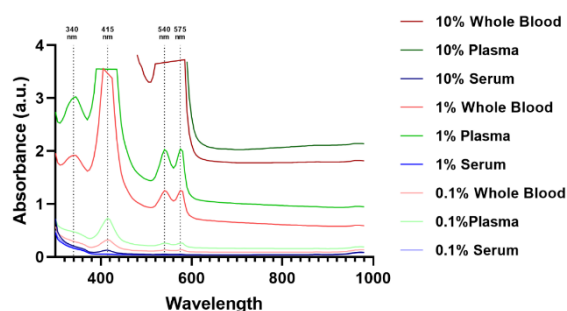


圖 2 血液吸收光譜圖

2.3 血液屏障模型

為了可以有效評估不同光譜對於血液視覺屏障在內視鏡成像造成的影響，因此針對多光譜實驗設計了一套臨時血液視覺屏障模型，用於模擬內視鏡手術過程中遇到的血液視覺屏障。實驗評估使用的為帶有骨釘的豬脊椎骨擬真內視鏡手術環境，其中骨釘為此次實驗主要拍攝之目標物。

此外，為了模擬不同情況下的血液視覺屏障環境，因此血液比水濃度亦為此次實驗重要的操縱因子，其中包括 3%、2%、1%、0.5%及 0%以上五種血液比水濃度，模擬血液視覺屏障環境。

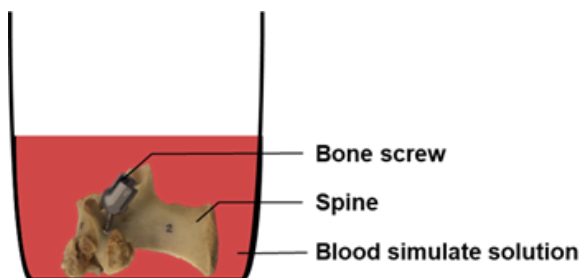


圖 3 血液視覺屏障模型示意圖

2.4 分光鏡及近紅外光相機

為了在不破壞現有的內視鏡系統模組的前提下增加更多的相機接口，以此提供醫師更多的設備輔助，我們設計了一分光鏡系統 TranScope，TranScope 可以使得兩台相機可以同時觀測到相同的手術畫面。

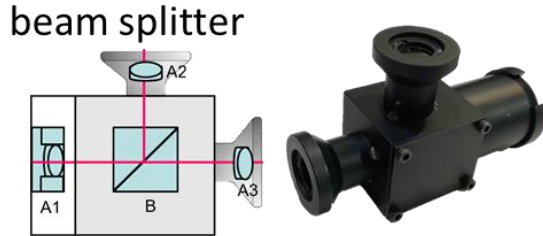


圖 4 分光鏡系統 TranScope

在保留原始內視鏡相機的情況下，我們額外添加了一近紅外光相機 UI-3240LE-NIR-GL，下圖是 UI-3240LE-NIR-GL 的 Quantum efficiency，以此用來於實驗中採集在近紅外光情況下的影像資訊，使得醫師可以在血液屏障狀態下繼續執行手術。

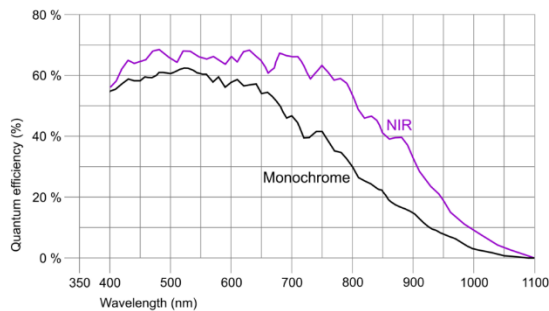


圖 5 近紅外光相機 UI-3240LE-NIR-GL 之 Quantum efficiency

2.5 著色模型

在 NIR 相機所捕捉到的影像擁有較多的細節，並且在血液屏障的環境下擁有更高的穿透度，但是 NIR 影像依舊有其缺點，其黑白的影像，對醫師來說無法直觀的判斷其組織，有可能產生誤傷病人的情況，因此我們開發了一著色模型，致力於將原先黑白的 NIR 影像進行著色。著色模型主要由四個部分組成：Encoder、Feature Extractor、Fusion Layer 和 Decoder。

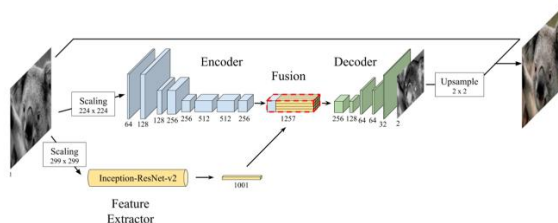


圖 6 著色模型架構

首先是 Encoder，Encoder 的主要任務是處理輸入的灰階圖像，提取其低層次特徵，由多層卷積層組成，結構如下表所示。

Layer	Kernels	Stride
conv	64x(3x3)	2x2
conv	128x(3x3)	1x1
conv	128x(3x3)	2x2
conv	256x(3x3)	1x1
conv	256x(3x3)	2x2
conv	512x(3x3)	1x1
conv	512x(3x3)	1x1
conv	256x(3x3)	1x1

表 1 Encoder 架構

再來是 Feature Extractor，Feature Extractor 是使用預訓練的深度卷積神經網絡(Inception-ResNet-v2) 來提取高層次特徵，這些特徵包含了豐富的信息。是一種結合了 Inception 模塊和殘差網絡 (ResNet) 的深度模型。Inception 模塊能夠並行提取多尺度特徵，而殘差網絡通過跳躍連接 (skip connections) 來緩解梯度消失問題。

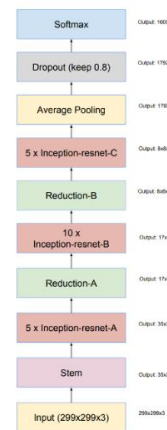


圖 7 Inception-ResNet-v2 架構

而 Fusion Layer 則是將 Feature Extractor 和 Encoder 的輸出進行融合，並提供給後續 Decoder 的反卷積層逐步恢復特徵圖的空間分辨率，生成最終的彩色圖像。

Layer	Kernels	Stride
conv	128x(3x3)	1x1
upsamp	-	-
conv	64x(3x3)	1x1
conv	64x(3x3)	1x1
upsamp	-	-
conv	32x(3x3)	1x1
conv	2x(3x3)	1x1
upsamp	-	-

表 2 Decoder 架構

2.6 分析方法

2.5.1 signal-to-noise ratio(SNR)

Signal-to-Noise Ratio 是一個評估訊號強度和噪聲干擾程度的參數，通常用於影像、音頻和通訊等領域中。在影像處理中，SNR 表示圖像中訊號與噪聲的比率，該比率越高，圖像的質量就越好。

為有效評估多光譜血液穿透試驗實驗影像輸出，此次實驗採用一種廣泛利用於磁共振影像(MRI)評估醫學影像雜訊比之方法，其方法定義於 NEMA 標準當中。此方法訊號為平均測量興趣區域(ROI)的像素值，如(2.1)式所示。雜訊定義為影像 4 個角落大小為 16*16 方形區域，重新組合成 32*32 方形區域取標準差，如(2.2)式所示。而後，取圖像測量興趣區域訊號與四周雜訊比值，作為最後 SNR 輸出，如(2.3)式所示。

$$S = \frac{1}{N_{ROI}} \sum_{N_{ROI}} l_s(r, k) \quad (2.1)$$

$$SD = std(l_n(r, k)) \quad (2.2)$$

$$SNR = 0.66 * \frac{S}{SD} \quad (2.3)$$

- S：訊號強度
- N_{ROI} ：測量興趣區域的像素數量
- $l_s(r, k)$ ：於 ROI 中位置 (r, k) 的像素值
- SD：雜訊強度
- $l_n(r, k)$ ：於 32*32 訊方塊中位置 (r, k) 的像素值
- SNR：雜訊比

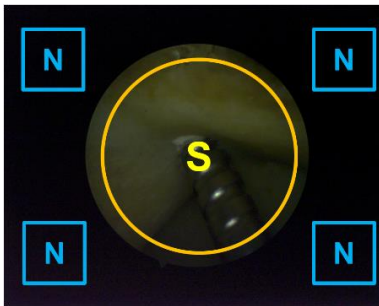


圖 8 SNR 計算取值範圍，S 為訊號區域，N 為雜訊區域

3. 實驗結果

3.1 血液屏障下於不同波長的可視化能力

為了評估不同波長在血液視覺屏障下的可視化能力，此次試驗選擇了三種血液視覺屏障模型 0%、0.5% 以及 1%，模擬內視鏡手術過程中血液屏障下的組織環境。如下圖實驗數據所示，可以發現當環境血液濃度增加，視覺化評分也隨之遞減。如圖所示，當環境中血液濃度從 0% 增加至 1% 時，視覺化關於 SNR 的評分則從 3.925 降至

0.817，這與血液吸收光譜結果相似，也表明視覺化與血液濃度間的負相關性。

於彩色相機影像中，波長在 629nm 環境之血液視覺屏障下，皆顯示出較高的 SNR 評分，能夠清楚觀察到目標物如骨釘等。而在波長 735nm 與 850nm 情況下，皆顯示出較低的 SNR 評分值。

而在近紅外光相機的畫面中，相較於彩色相機擁有較高可視化 SNR 評分。近紅外光相機在 629 nm 和 735 nm 的環境下在各種血液濃度情況下提供了比白光源更好的可視化效果。SNR 分數在白光源中為 4.909，在 629 nm 光源中為 13.1534，在 735 nm 光源中為 12.8931。數據顯示出，即使在極端的血液濃度環境下如 1% 的血液視覺屏障模型，依舊可以很好的觀測目標物。

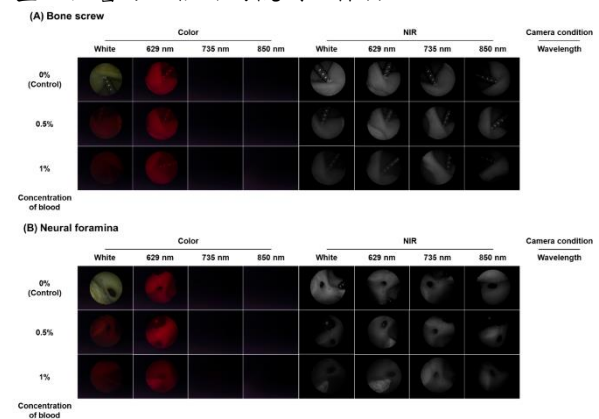


圖 9 實驗結果圖，圖 A 目標物為骨釘，圖 B 目標物為神經孔

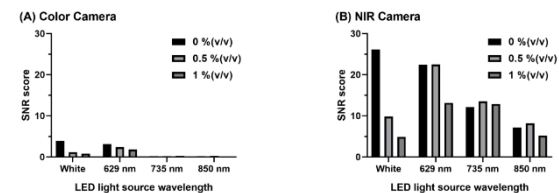


圖 10 實驗長條圖，圖 A 為彩色相機，圖 B 為近紅外光相機

3.2 血液屏障下於不同血液濃度的可視化能力

為了評估血液視覺屏障在不同血液濃度下的可視能力效果，此次實驗將在白色光源及 735nm 的環境下，分別分析在血液濃度 0%、0.5%、1%、2% 及 3% 的血液屏障模型的可視效果評分。

於彩色相機中，白色光源及 735nm 在極端血液濃度 2% 及 3% 環境下，已經無法完整的成像，完全遺失目標物的影資訊。

在近紅外光相機下，白色光源呈現的影像對比於彩色相機並沒有明顯的差異。然而，於 735nm 波段下，近紅外光相機在顯示上有著顯著的提升，使內視鏡影像有更高的可視化能力。在近紅外光相機下，735nm 波長於 2% 的血液濃度條件下畫面的 SNR 值評分為 6.155，於 3% 的血液濃度條件下的 SNR 值評分為 5.471。

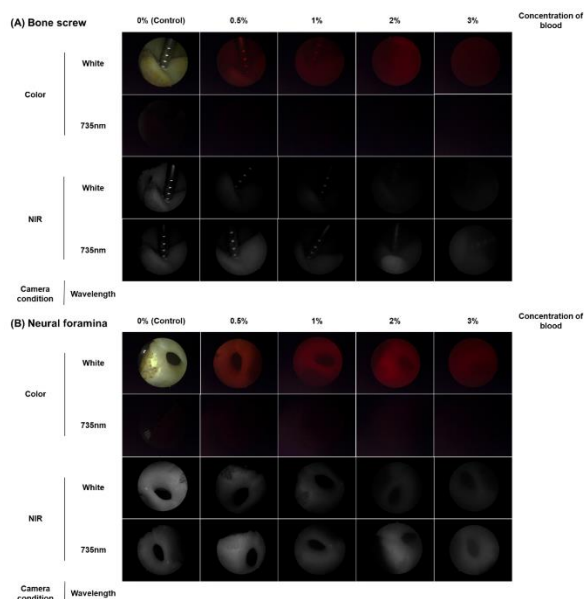


圖 11 實驗結果圖，圖 A 目標物為骨釘，圖 B 目標物為神經孔

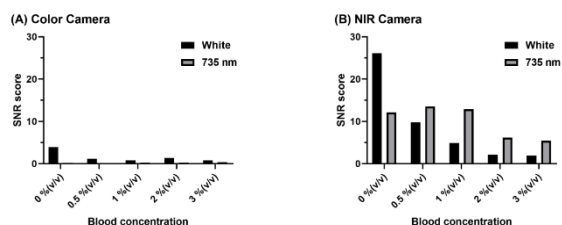


圖 12 實驗長條圖，圖 A 為彩色相機，圖 B 為近紅外光相機

3.3 近紅外相機影像著色

關於著色模型的資料收集，我們使用的資料集是透過，我們自行開發的分光鏡系統 TranScope 結合近紅外相機 UI-3240LE-NIR-GL 搭配豬博士之 Stryker 系統之彩色相機，進行資料蒐集，實驗環境為豬隻脊椎內視鏡手術試驗，以豬隻模擬脊椎減壓術式，共計進行了 3 組脊椎減壓術式，同時使用 Stryker 系統之彩色相機及近紅外相機 UI-3240LE-NIR-GL 採集手術錄像。

數據總共收集 11380 組影像，收集資料如下圖所示，我們將其中挑選出第一組關節減壓術式的 6904 組影像用於訓練，以上數據在實驗前皆由相關醫療領域專業人員反覆查驗確認其影像可用性，以證明這次實驗結果的可靠度。

訓練集總計 6214 筆數據，驗證集總計 690 筆數據，其驗證集與訓練集的數據比例為 1:10，當中驗證集不會參與模型的訓練。

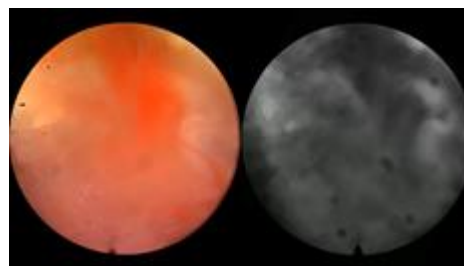


圖 13 收集資料，左側為彩色相機影像，右側為近紅外光相機影像

實驗設定訓練 300 個週期，批次大小設定為 47。根據實驗結果數據，可以確定訓練集與驗證集已經正確地收斂。最終訓練集準確率收斂在 0.9401，驗證集準確率收斂在 0.9516。而訓練集損失收斂在 0.00097，驗證集損失收斂在 0.00081。

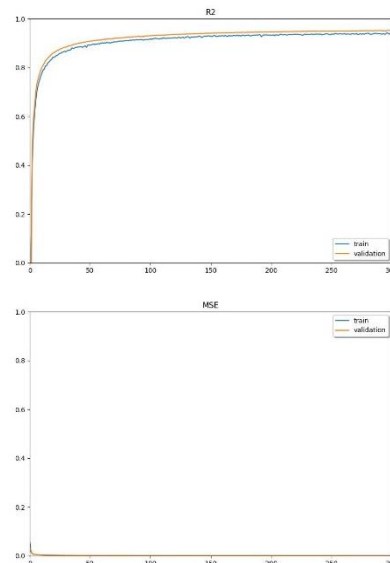


圖 14 模型訓練 acc、loss 曲線

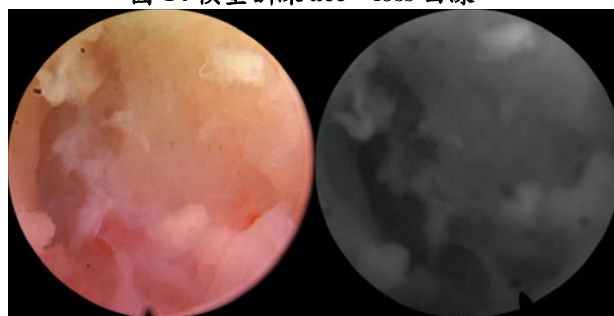


圖 15 著色前影像(右側近紅外光相機影像)

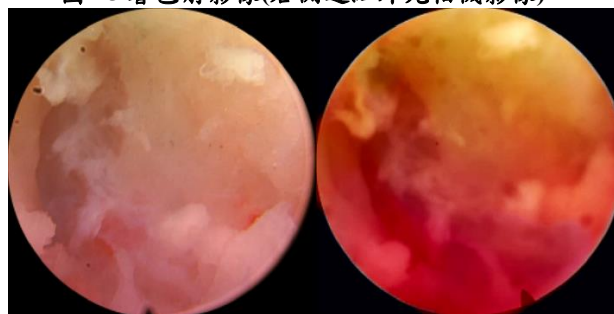


圖 15 著色後影像(右側近紅外光相機影像)

4. 結論

在本論文中，我們針對在內視鏡手術過程中血液視覺屏障的問題，提出利用血液在不同波段的光譜下擁有不同的穿透度之特性，設計出一款光學濾波系統，使得醫師在血液屏障的情況下，依舊可以提供足夠清晰的影像，讓醫師可以繼續進行手術，降低因為出血在手術過程中造成的影響。

在血液屏障下於不同波長的可視化能力的實驗中，我們實驗了在不同波段下影像的清晰度，以及彩色相機以及近紅外相機，於不同血液屏障環境中的成像效果，在 629nm 和 735nm 中影像的成像效果較好，並且在後續的血液屏障下於不同血液濃度的可視化能力實驗中，證實了當血液濃度提升後，在 735nm 波段的環境中擁有更好的可視化能力。

最後我們針對醫師提出的近紅外相機的黑白影像較不利於進行手術試驗，所以我們結合了 AI 著色功能，提供醫師彩色畫面，最終驗證集準確率收斂在 0.9516，損失收斂在 0.00081，呈現出效果不錯的著色功能。

假設能夠在出血情形下提供較為清晰的影像，使醫師繼續進行手術，便可以降低醫師的手術時間，以及提高醫師的手術成功機率。在未來的工作包括縮小整體分光鏡系統 TranScope 的體積，找尋穿透能力更強的波段，蒐集更多手術資料以此使得著色模型的顏色更加貼近實際情況，並降低模型運行時的延遲。

誌謝

感謝國科會計畫(計畫編號 112-2823-8-027 - 003)在研究經費上的補助。

參考文獻

- [1] Sairyo, K.; Sakai, T.; Higashino, K.; Inoue, M.; Yasui, N.; Dezawa, A. Complications of Endoscopic Lumbar Decompression Surgery. *Minim. Invasive Neurosurg.* 2010, 53, 175–178, doi:10.1055/s-0030-1262814.
- [2] Fjeld, O.R.; Grøvle, L.; Helgeland, J.; Småstuen, M.C.; Solberg, T.K.; Zwart, J.A.; Grotle, M. Complications, Reoperations, Readmissions, and Length of Hospital Stay in 34 639 Surgical Cases of Lumbar Disc Herniation. *Bone Joint J.* 2019, 101-B, 470–477, doi:10.1302/0301-620X.101B4.BJJ-2018-1184.R1.
- [3] Sun, D.D.; Lv, D.; Wu, W.Z.; Ren, H.F.; Bao, B.H.; Liu, Q.; Sun, M.L. Estimation and Influence of Blood Loss under Endoscope for Percutaneous Endoscopic Lumbar Discectomy (PELD): A Clinical Observational Study Combined with in Vitro Experiment. *J. Orthop. Surg. Res.* 2020, 15, 1–9, doi:10.1186/s13018-020-01797-1.
- [4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, Las Vegas, NV, USA, Jun 2016.
- [5] Deep Koalarization: Image Colorization using CNNs and Inception-Resnet-v2
- [6] Colorful Image Colorization Richard Zhang, Phillip Isola, Alexei A. Efros {rich.zhang, isola, efros}@eecs.berkeley.edu University of California, Berkeley