



# 東華AI通訊報

雙月刊

發行 | 人工智慧辦公室  
連絡電話 | (03)890-3792

發行人 | 林信鋒  
傳真號碼 | (03)890-0225

主編 | 江政欽  
網址 | <https://aioffice.ndhu.edu.tw>

## 【AI 科技萬花筒】

- [達人分享：「深度學習於水下影像之物件偵測」-----1](#)
- [達人分享：「AI 輔助醫學影像改變了醫療照護模式」----3](#)
- [「全球第一個真正的 AI 規範即將到來！來看看它們長什麼樣子」-----5](#)
- [「2 步驟用 AI 為家裡裝潢找靈感！在 RoomGPT 上傳空間照就能搭出 9 種風格」-----5](#)
- [「快遞公司 FedEx 愛上 AI，聘請 DexR 機器人搬運包裹，擺放整齊不會亂丟客戶貨物！」-----5](#)
- [「英國、印度都在做！AI 攝影機助力科技執法，道路安全大加分」-----5](#)
- [「員工今天心情好嗎？AI 心理健康工具都知道，掌握是否有憂鬱，必要時提供建議！」-----6](#)
- [「誠品生活新店首展「AI 靈感大師：澳洲 3D 光影觸動樂園」，登場！AI 科技打造沉浸式藝術體驗」-----6](#)
- [「【AI 時代人類的新工作】AI 威脅了一些職業，但有 4 類機會正在增加」-----6](#)
- [「ChatGPT 最強對手「Claude」開放台灣註冊、兩步驟免費用！能讀整本書？支援中文？」-----6](#)

## 【AI 元宇亮東華】

- [亮點研發「台灣手語辨識—神奇的深度學習」-----6](#)

## 【AI 小學堂開講】

- [《Claude 2—AI 聊天機器人的新選擇》-----8](#)

## 【AI 跨域徵人帖】

- [《AI 人才能力檢定 2024 年上路，半導體、資通、醫療、金融、政府、製造、資安七產業先行》-----10](#)

## 【AI 科技萬花筒】-達人分享

### 《深度學習於水下影像之物件偵測》

臺灣師範大學電機工程學系 葉家宏教授  
國立中山大學電機工程學系 林楚涵  
臺灣師範大學電機工程學系 何鈺漢

隨著自然資源日益匱乏和全球經濟發展，水下環境探索逐漸受到人們的重視。水下探索研究是個非常重要而充滿挑戰的領域，它具有廣泛的應用，包括測量水下地形、探索水下生態系統、進行考古學研究和資源開發等。水下探索技術不僅是國內外研究熱點，更是全球性的發展趨勢。

水下影像處理[1][2]在探索水下環境方面具有巨大的潛力，該技術已廣泛應用於多個領域，包括水下地形掃描和自主水下載具 (Autonomous Underwater Vehicles, AUVs) 等。可以執行許多重要任務，包括繪製水下地形圖、執行搜救行動及水下生態系統觀測等。水下載具通常配備有攝像設備以捕捉水下影像。

高質量的水下影像對水下探索任務至關重要，然而，水下環境充滿了挑戰，因為水的物理特性對影像品質有顯著影響。如圖 1 所示；光在水中會被吸收、散射和反射，使水下影像常有模糊、雜訊和顏色失真等問題，從而降低了物件偵測的效能。因此，有許多改善水下影像品質的方法被提出，其中深度學習 (Deep Learning) 在這方面的表現出色。此外，由於水下能源有限，因此開發低複雜度或低功耗的演算法來執行相關任務，例如物件偵測 (Object Detection) 和物件追蹤 (Object Tracking)，變得極為重要。

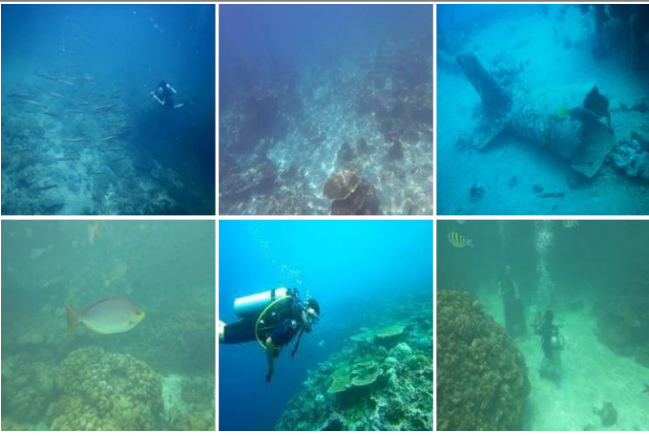


圖 1 水下影像樣本範例

為了提升水下物件偵測的準確度並降低計算複雜度，本文基於深度學習提出一個輕量化的水下物件偵測網路。透過水下影像色彩轉換和物件偵測的聯合學習，將目標放在單張水下影像的物件偵測，此網路的特點是利用色彩轉換模組將彩色水下影像轉換為相應的強化灰階影像，以解決水下顏色失真的問題，因而不需要提前恢復影像品質，也因此能以較低的計算複雜度增強物件偵測效能。實驗共蒐集了 12400 張水下影像，並將影像尺寸縮放為 320×320 像素。資料集分為訓練集（8,000 張）、驗證集（2,000 張）和測試集（2,400 張）。在資料集中，魚類占 20%、碎片殘骸占 50%、潛水員占 30%。此外我們在 Raspberry Pi 3 B+ 設備上實現模型，以證明這個基於深度學習的物件偵測器可在低計算複雜度情況下有最大工作效能。

本研究應用影像處理與深度學習技術於水下物件偵測 [3] 的整體系統流程如圖 2 所示。首先，將輸入的水下影像進行尺寸縮放，再採用色彩轉換網路將彩色影像轉為相應的強化灰階影像；此方法不需要提前恢復影像，即能以較低的計算複雜度有效增強物件偵測效能。研究中提出的色彩轉換網路由 10 個輕量級卷積層組成；前 5 個卷積層旨在從輸入影像中提取特徵，而後 5 個卷積層的設計則受到 MobileNets[4] 的啟發，透過深度可分離卷積（Depth-wise Separable Convolution），將標準卷積分解為深度卷積（Depth-wise Convolution）和逐點卷積（Point-wise Convolution）。這種分解方式顯著降低了計算複雜性和模型大小。在色彩轉換深度網路中，深度可分離卷積的應用，為通過逐點卷積減少前一層輸出的維度，之後再利用

深度卷積進行特徵提取，最終輸出每個顏色通道的轉換參數，並可透過公式(1)將彩色水下影像轉換成強化灰階影像。

$$I_{Gray} = \frac{p_{\alpha}}{p_{\alpha} + p_{\beta} + p_{\gamma}} I_R + \frac{p_{\beta}}{p_{\alpha} + p_{\beta} + p_{\gamma}} I_G + \frac{p_{\gamma}}{p_{\alpha} + p_{\beta} + p_{\gamma}} I_B \quad (1)$$

其中  $I_{Gray}$  為強化灰階影像， $I_R$ 、 $I_G$ 、 $I_B$  分別為輸入彩色影像之 R、G、B 通道， $p_{\alpha}$ 、 $p_{\beta}$ 、 $p_{\gamma}$  分別為經色彩轉換網路後所得到的各通道之轉換參數。經過色彩轉換網路後，將得到的強化灰階影像輸入至物件偵測網路中。物件偵測網路中包含了特徵提取模組（Feature Extraction Module）與特徵聚合模組（Feature Aggregation Module）。特徵提取模組由多個特徵提取層、感知模組與輸出層所組成。其中，特徵提取層由 1 個輸入層、1 個最大池化層、3 個卷積層和 1 個輸出層組成，其主要功能將輸入特徵圖的空間大小減半，以提取初級特徵。為降低計算負擔，特徵提取模組僅使用了有限的特徵圖。感知模組由 8 個卷積層與 1 個最大池化層所組成，目的是為了彌補可能存在的特徵提取不充分的問題，進一步細化先前特徵提取層的特徵。輸出層則為 1 個卷積層，將得到的特徵輸出至特徵聚合模組中。特徵聚合模組是基於特徵金字塔網路（Feature Pyramid Network, FPN）[5] 架構而設計，用於連接特徵提取模組，通過具有橫向連接的自上而下架構，提取不同尺度的高層次語義特徵圖。研究中，特徵提取模組的第三、第四和第五階段的特徵圖被輸入至特徵聚合模組。特徵聚合網路最終會產生不同層級的 3 個輸出，用於不同尺度的物件偵測，可以獨立地進行各個層級的預測。

為了證實色彩轉換網路的必要性，研究中設計了消融實驗，分別對未使用色彩轉換網路、使用本研究提出之色彩轉換網路，和傳統 BT.601 色彩轉換進行物件偵測效能評估，最終證明使用所提出的色彩轉換網路結合物件偵測網路所得到的效能最佳。

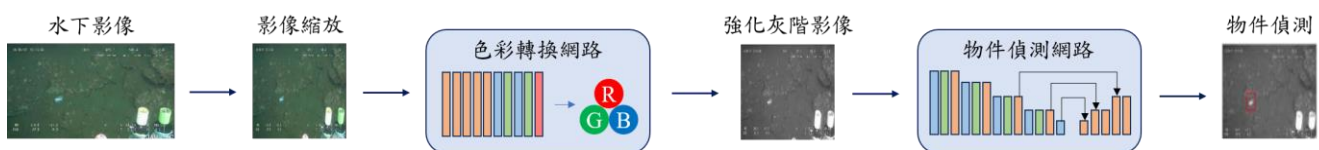


圖 2 應用影像處理與深度學習技術於水下物件偵測之整體系統流程圖



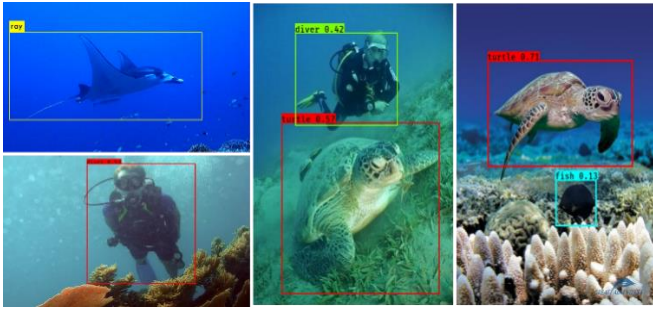


圖 3 水下物件偵測實驗結果圖

實驗結果如圖 3 所示；本研究所提出的方法在水下物件偵測上的準確度達到 0.8956，與其他現有的物件偵測方法 Fast R-CNN[6]、SSD[7]、Tiny-YOLO[8]相較，準確度提升了 5-10%，而在計算量方面僅有 5.06 GFLOPs，減少了 50-80%。不同於一般基於深度學習需先將水下影像還原，本研究透過強化後的灰階影像解決了水下影像色彩失真的問題，同時也降低了運算複雜度，使其能在運算能力有限的設備上，依舊能夠快速地進行任務。

本研究結合深度學習與影像處理技術，實現高度精確且輕量化的水下物件偵測系統。我們的目標是將這一技術應用於自主水下載具，以快速而準確地分析攝像設備所捕獲的水下影像，並在任務執行中實現更即時的反應，從而提高工作效率。這一技術將對多個領域產生重大影響，包括水下探索、環境監測、海洋研究和搜索救援。透過在自主水下載具上的應用，我們可以實現更快速和更即時的反應，提高執行任務的效率，並有望開闢出新的水下技術應用領域。

#### 參考文獻

- [1] C. H. Yeh, C. H. Lin, L. W. Kang, C. H. Huang, M. H. Lin, C. H. Chang, and C. C. Wang, "Lightweight Deep Neural Network for Joint Learning of Underwater Object Detection and Color Conversion," IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 11, pp. 6129-6143, 2021.
- [2] C. H. Yeh, C. H. Huang, and L. W. Kang, "Multi-scale deep residual learning-based single image haze removal via image decomposition," IEEE Transactions on Image Processing, vol. 29, pp. 3153-3167, 2020.
- [3] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, "Deep learning for generic object detection: A survey," International journal of computer vision, pp. 261-318, 2020.
- [4] A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient

Convolutional Neural Networks for Mobile Vision Applications," arXiv preprint arXiv: 1704.04861, 2017.

[5] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2117-2125, 2017.

[6] S. Ren, K. He, R. B. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," Advances in neural information processing systems, pp. 91-99, 2015.

[7] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, "SSD: Single Shot MultiBox Detector," in Proceedings of the European Conference on Computer Vision, pp. 21-37, 2016.

[8] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779-788, 2016.

### 《AI 輔助醫學影像改變了醫療照護模式》

門諾醫院 游冬齡副院長/核醫科主任

#### 醫學影像的發展

醫學影像是現代醫療照護中不可或缺的一環。醫學影像提供了疾病的診斷、治療結果評估、疾病癒後等臨床上重要的資訊，成為醫療人員診療病人的重要依據。

倫琴 (W. Roentgen) 於 1895 年發現了 X 光，自此開啟了醫學影像臨床應用的序幕。在醫學影像發展過程中，大約每 10 年就會有突破性的設備或技術出現，例如 1950 年代的超音波影像、1960 年代的核醫同位素影像、1970 年代的電腦斷層影像 (CT)、1980 年代的磁振影像 (MRI)、1990 年代的 3-D 超音波影像、2000 年代正子電腦斷層掃描影像 (PET-CT) 都是醫學影像發展史上非常重要的里程碑。

早在 1990 年代初期，AI 的技術就已經被應用在輔助醫學影像上。由於軟硬體的進步，近年來 AI (Artificial Intelligence) 及 ML (Machine Learning) 技術蓬勃發展，已經陸續被應用於醫學影像上。美國食品藥物管理局 (U.S. The Food and Drug Administration, FDA) 於 1995 年核准第一個 AI 應用醫療領域的產品上市，此後 AI 在各醫療領域被 FDA 核准的應用產品漸增。2018 年，首個 AI 輔助醫學影像的應用產品也獲得 FDA 核准上市。依據 FDA 於 2023 年 10 月的報告指出，截至

2023 年為止，總計有 692 個 AI 相關的醫療產品獲得美國 FDA 核准。2023 單年便有 171 個 AI 醫療應用產品取得 FDA 核准，數量較 2022 年增加 30%，未來可預見也將會有更快速的成長。在 FDA 核准的 692 個 AI 醫療應用產品中，AI 輔助醫療影像相關的項目占了 87% 之多，突顯 AI 在醫療影像扮演相當重要的角色。因此，2020 年代，AI 有機會在醫學影像領域接棒成為下一個突破性發展的項目。

### AI 醫學影像可提升診斷效率

現代醫學影像技術發展越來越精細及多元，所產生的資料量隨之也越來越龐大，這對於從事醫學影像診療工作的醫師們而言，意味著更準確的影像判讀要求及巨大的工作量。透過 AI 輔助，正好可以解決此問題。

國醫療機構一項研究報告指出，透過 AI 的應用，能更快且精準的從大量的醫學影像中找到病灶，讓醫師撰寫診斷報告的時間可由 11.2 天大幅縮短到 2.7 天，可以縮短了病人的焦慮等待[1]。

臺北榮總於 2019 年用 AI 輔助醫師進行腦部 MRI 醫學影像診斷腦腫瘤，一般情況下，醫師進行影像判讀時要從一系列的數百張 MRI 影像中找出腦腫瘤，過程平均要花費 30 分鐘以上。然而，使用臺北榮總和台灣人工智慧實驗室 (Taiwan AI Lab) 共同訓練的臨床人工智慧腦轉移瘤自動判讀系統，醫師 30 秒內就能得到初步判讀結果。這套系統除了大幅縮短判讀時間，也提升了判讀的正確性。節省下來的時間，醫師可以用在和其他醫療人員及病人溝通及討論上，從而提升醫療照護的品質。

辨別腫瘤是良性或惡性在癌症治療上至關重要，然而臨床上要區別惡性或良性有一定的難度。現在透過 DL (Deep Learning) 及 CNNs (Convolutional Neural Networks) 的演算法分析，已經能夠精準判斷腫瘤是良性或惡性[2]。

### AI 輔助醫學影像可早期發現疾病

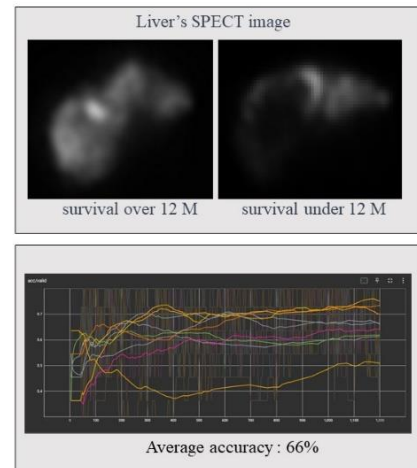
疾病治療最佳原則是早期發現早期治療。美國麻省總醫院暨布萊根婦女醫院 (Mass General Brigham) 醫學影像部門擁有大約 50 個 AI 輔助醫學影像系統，能夠早期發現癌症、腦血管瘤、中風等疾病早期的病變，並且預測疾病的可能進展，以便提早治療預防疾病進一步的發展，也可提高疾病的治癒率及避免併發症的發生。這些成就沒有 AI 的輔助是很難實現的。透過 AI 的輔助，突顯了醫學影像的關鍵角色，進而改善醫療治療的成效。

以腦梗塞性中風為例，腦梗塞性中風發生初期不容易用電腦斷層診斷出來，一般至少要臨床症狀出現 24 小時之後，才有機會經由電腦斷層診斷，很容易因此錯失治療的先機。研究顯示，透過 CNNs 演算法，就可以在早期階段透過電腦斷層診斷出來，讓病人在急診時即可迅速給予血栓溶解劑治療，避免腦部細胞進一步受傷害[3]。在國外已經有很多醫院的急診室都建置有這種 AI 輔助醫學影像系統。心肌梗塞也是國人重要的疾病之一。研究顯示，應用 ML 模式及 CNNs 演算法輔助的超音波影像，可以偵測到早期冠狀動脈鈣化的程度，避免心肌梗塞的發生[4]。

### AI 輔助醫學影像可更完整監測病人

化學治療是癌症治療的重要方法，觀察化學治療後細胞的早期反應，和病人的癒後有相當密切的關係。然而，傳統醫學影像技術必須在治療後經過一段時間，才可以看到治療的結果，這對於治療後的早期病變很難偵測，可能延誤治療的契機。但是經過 AI 的協助，就有可能儘早發現病變，並預測結果[2]。

門諾醫院和東華大學合作研究 AI 輔助肝臟血流灌注攝影，預測晚期肝細胞癌經肝動脈灌注化學治療結果，預測病人治療後是否可存活 12 個月的準確度大約在 66% (見下圖)。未來病人接受透過 AI 輔助肝臟血流灌注影像的分析來預測病人的存活期，若預測病人的存活期可能少於 12 個月，醫師可以考慮進一步合併其他治療，來提升病人存活期。



### AI 輔助醫學影像未來的發展

目前已經有相當多 AI、虛擬實境 (Virtual Reality, VR)、擴增實境 (Augmented Reality, AR) 互相結合的輔助系統廣泛應用在醫學影像的領域，醫師或病人可以透過 AI 輔助的虛擬實境了解疾病並模擬治療過程，讓醫師及病人共同參與疾病的

診斷及治療。這方面的應用未來必能開創醫學影像臨床應用的另一新紀元。

2022 年，ChatGPT 橫空出世，可預見生成式 AI 應用於醫學影像將會有巨大且正面的影響。另外，Google 在 2016 年提出了一個「聯盟學習 (Federated Learning)」的概念。主要是醫療機構都使用自己的資料進行 AI 的應用研究，然而這些資料來源受到病人數統計、使用的儀器設備或臨床專業等因素的影響，可能會出現偏差。聯盟學習的概念是多個參與醫療機構的醫學影像資料，能在不共享資料的情況下建立一個共同的機器學習模型，從而解決資料隱私、資料安全、資料存取權限和異構資料訪問等問題。聯盟學習 AI 演算法可以從不同來源的大量資料中獲得經驗，讓多家機構可以合作開發 AI 模型，又不用彼此直接分享機密的臨床資料，成為理想的 AI 應用開發合作模式。

由於技術、概念快速進展及不斷創新，各式 AI 輔助醫學影像系統應用於臨床各方面已獲得顯著成果[5][6][7][8]。這些成果提升醫療影像流程及服務效率、增進醫療影像診斷的品質、提供更精準追蹤的依據，大大提高了整體醫療照護的效能及品質，確保病人獲得最佳照顧。

AI 輔助醫學影像已經明顯地改變了醫療照護模式，標誌著醫學影像領域已經正式邁入 AI 輔助醫療影像的新世代。

#### 參考文獻

- [1] Annarumma M, Withey SJ, Bakewell RJ, Pesce E, Goh V, Montana G. Automated Triaging of Adult Chest Radiographs with Deep Artificial Neural Networks. *Radiology* 2019, 291:196–202.
- [2] Luchini C, Pea A, Scarpa A. Artificial intelligence in oncology: Current applications and future perspectives. *Br. J. Cancer* 2022, 126:4–9.
- [3] Soun JE, Chow DS, Nagamine M, Takhtawala RS, Filippi CG, Yu W, Chang PD. Artificial Intelligence and Acute Stroke Imaging. *Am. J. Neuroradiol.* 2021, 42:2–11.
- [4] Zhou J, Du M, Chang S, Chen Z. Artificial intelligence in echocardiography: Detection, functional evaluation, and disease diagnosis. *Cardiovasc. Ultrasound* 2021,19:29.
- [5] Reabal N. Redefining radiology: A review of artificial intelligence integration in medical imaging. *Diagnostics* 2023,13:2760.
- [6] Artificial intelligence in medical imaging: switch from radiographic pathological data to clinical meaningful endpoints. *Lancet Digital Health* 2020, 2: 486-488.

[7] Mayo RC, Kent D, Sen LC, Kapoor M, Leung JWT, Watanabe AT. Reduction of False-Positive Markings on Mammograms: A Retrospective Comparison Study Using an Artificial Intelligence-Based CAD. *J. Digit. Imaging* 2019, 32:618–624.

[8] Lakhani P, Sundaram B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology* 2017, 284:574–582.

### 【AI 科技萬花筒】-AI 科技動態

**全球第一個真正的 AI 規範即將到來！來看看它們長什麼樣子**

<https://user106763.pse.is/5dpxqw>

生成式 AI 迅速竄升為全球話題，而它的進步速度之快，也讓大家開始討論如何避免遭到濫用、對社會帶來危害，例如讓 AI 掌握研發藥物工具製造出生化武器等。日前，OpenAI 執行長在內的 350 位專家便共同連署，呼籲正視 AI 滅絕人類的風險。（TechOrange，廖紹伶，2023/10/13）

**2 步驟用 AI 為家裡裝潢找靈感！在 RoomGPT 上傳空間照就能搭出 9 種風格**

<https://user106763.pse.is/59e7a3>

想自己重新設計、裝飾個人居住空間的你，是否曾碰過這個問題—因為無法實際看到成果，因此不知道該買什麼風格的家具、該挑什麼顏色的窗簾、該怎麼進行擺設。

現在，AI 能為你解決這個痛點！（TechOrange，蘇郁晴，2023/10/19）

**快遞公司 FedEx 愛上 AI，聘請 DexR 機器人搬運包裹，擺放整齊不會亂丟客戶貨物！**

<https://user106763.pse.is/5e25ax>

美跨國運輸公司 FedEx 最近推出一款結合 AI 技術的雙臂機器人 DexR，能將不同大小的包裹在最大化空間利用的情況下，井然有序地堆疊在運輸卡車中，為員工減輕工作負擔。（TechOrange，蘇郁晴，2023/10/11）

**英國、印度都在做！AI 攝影機助力科技執法，道路安全大加分**

<https://user106763.pse.is/5cxtq4>

為了取締違停、超速等交通違規事項，台灣於 2019 年開始推行「科技執法」，上路至今已設置 265 處，儘管如此，還是得倚靠大量人力進行違規取締。據內政部警政署統計，今年



1 月到 7 月，科技執法取締件數僅占整體的 6.8%。（未來商務，Ariel Chang，2023/10/5）

**員工今天心情好嗎？AI 心理健康工具都知道，掌握是否有憂鬱，必要時提供建議！**

<https://user106763.pse.is/5dvtpw>

美國新創 Kintsugi 開發能透過人聲辨識憂鬱和焦慮狀況的 AI 工具 Kintsugi Voice。Kintsugi 共同創辦人暨執行長 Grace Chang 接受《Forbes》採訪時提到：「根據我們進行的調查顯示，企業越來越重視員工的心理健康狀況，因為若每一位員工每年因此請假 12 個工作天，將使美國經濟損失 476 億美元。」（TechOrange，蘇郁晴，2023/10/5）

**誠品生活新店首展「AI 靈感大師：澳洲 3D 光影觸動樂園」登場！AI 科技打造沉浸式藝術體驗**

<https://user106763.pse.is/5dvlw>

由澳洲團隊 ENESS 親自前來打造的亞洲首展「AI 靈感大師：澳洲 3D 光影觸動樂園」，9/29 全面開放；藉由 AI 科技打造沉浸式藝術之旅，由 21 件作品組成，部份剛在法國結束展出，此次還有為誠品打造的獨家角色-樂園中的至高神「太陽神」、身著光影變幻舞裙的「森之舞者」終於亮相；還能用手機解開「靈感大師」的科技魔咒，獲得 ENESS 團隊為台灣觀眾特意安排的中文限定詩籤，開放帶回品讀紀念；「沉睡山靈」的低吟密語也具有震撼臨場感！（ZEEK 玩家誌，Irene，2023/9/30）

**【AI 時代人類的新工作】AI 威脅了一些職業，但有 4 類機會正在增加**

<https://user106763.pse.is/5djgln>

當 OpenAI 去年底發布 ChatGPT，AI 搶奪人類工作的疑慮、討論便不曾停止。根據《CNN》，最近這幾個月，有越來越多科技公司將 AI 當作裁員和重新考慮新員工的原因。高盛日前報告更指出，生成式 AI 將搶走 3 億個工作，其中律師和行政等白領工作面臨的風險最大。（TechOrange，廖紹伶，2023/10/24）

**ChatGPT 最強對手「Claude」開放台灣註冊、兩步驟免費！能讀整本書？支援中文？**

<https://user106763.pse.is/5dgs49>

Claude 開放台灣註冊了！被稱為 ChatGPT 最強對手的 Claude，是由 Anthropic 推出的一款聊天機器人，具有強大的自然語言處理能力，用戶可以和它自然的交流。（虎嗅網，2023/7/12）

## 【AI 元宇亮東華】- 亮點研發

### 〈台灣手語辨識-神奇的深度學習〉

資訊工程學系 羅壽之教授  
陳鼎勳

全世界有超過 4000 萬聽障人士，而台灣約有 12 萬名，不論是聽障者與聽力正常者間或是聽障者彼此間，手語扮演著重要的溝通管道。因此，建置一個能夠學習或是辨識台灣手語的系統是相當重要的。在國外，美國手語、印度手語、德國手語、中國手語均有人投入手語辨識甚至是翻譯的研究，而台灣手語辨識的研究相對於此，更多的是類似指拼式手語（注音手語）的辨識，針對一般台灣手語單詞的辨識相對缺乏。有鑑於此，我們希望建構一個台灣手語辨識系統。

台灣手語組成多元，照地域可區分為北、中、南三類主要手語。北部手語為台灣手語使用上的大宗，其語系主要為東京手語及中國手語；南部手語組成主要為大阪手語及中國手語；而中部手語組成主要來自南部手語。北、中、南三部手語皆屬於自然手語，自然手語為聽障人士的母語，語法與聽人所用的慣用語法不大相同。手語的詞彙組成基本可以拆成四種更小單位：手形、位置、動作與朝向。手形指的是雙手的形狀跟樣態，位置為手形所比的位置，動作為手勢上的變化及轉動，朝向則是手心或指尖所指的方向。目前，台灣手語手形多達 62 種，如果兩個詞彙的位置、動作及朝向相同，手形不同就會導致語意的完全不同。此外，非手部信號也為台灣手語表達中重要的一環，諸如臉上的表情、眼神甚至到整個軀體方向都可能對語意產生影響。

手語識別通常可以分為兩大主流：手語辨識及手語翻譯。手語辨識是建立在影片逐幀的動作辨識，就本質上較接近動作辨識；手語翻譯是建立多個單位的手語辨識，考慮的是連續動作辨識，並結合語義間的文法架構。理論上不管在資料集的規模或是模型的複雜度，手語翻譯皆比手語辨識來得困難。

手語辨識為動作識別的支類，受益於深度學習（Deep Learning）的崛起，以及聲音序列模型上的成功，手語辨識方法也變得更為多元。近年的發展都主要以 RGB 相機提供圖像資料流，其他借助 Kinect 體感設備或是手部傳感器提供輔助辨識的資料流，藉此提高整體辨識度。而辨識的方法也百花齊放，例如使用光流（optical flow）追蹤運動軌跡，將 RGB 深度圖搭配 3D 卷積神經網路（Convolutional Neural Network，CNN），或是使用關節點座標，與注意力（attention）機制的

時序網路相結合等。能肯定的是，深度神經網路在動作或是手語辨識上已變成不可缺少的技術。

動作辨識一直是電腦視覺的重要領域，由於關注範圍為整個影片序列的資料特徵，所以動作辨識相較於單一圖片辨識變得更為複雜。手語辨識的方法其實與動作辨識相差無幾，差別在於手語辨識的主角在於手。因此，我們先使用物件偵測的方法找出左右手位置並進行裁切，而裁切的手部圖像亦會搭配整張圖像作為輔助特徵，如圖 4。我們將整張圖像裁切出人像主體，畢竟背景圖像對於辨識毫無幫助，且可能對模型造成冗餘的計算負擔。



圖 4 人像與左右手偵測與裁切

手語辨識需要使用整個時間序列影像作為其辨識輸入，對於手部的序列影像，我們採用 3D-CNN 進行特徵提取，但對於人體影像，我們則是隨機採樣並搭配 2D-CNN 進行特徵提取。如此設計是考量對於手語辨識而言，手部的連續動作是不可忽略且變化相對較大，而對於整張圖像或人體而言，主要負責提供表情及整體架構，其變化相對較少，且相對於手部而言是輔助特徵。因此，對於手部圖像我們使用運算量高但準確度高的 3D-CNN 作為特徵提取器，而對於人像我們則使用運算量低的 2D-CNN 作為特徵提取器。

台灣手語資料集不論是在辨識或是翻譯上資源都不多，因此我們收集網路上的手語資源，加上自行拍攝的手語影片，創建了一個包含 15 種類別（不同背景與人物）、52 個單詞共 780 部短片的台灣手語資料集。整體手語辨識包含兩階段的網路模型運作，第一階段採用 YOLO 網路模型作為人體及手部的物件偵測模型，第二階段則為動作辨識網路模型。

YOLO 網路模型執行物件偵測的原理為透過影像單元格對影像中的物件進行捕捉，一旦物件中心落在單一單元格中，則單元格會負責對物件的邊框進行估測並對物件分類。目前

YOLO 已有多代的更新，每代的 YOLO 都不斷地融合添加一些當今圖像辨識模型的優點，使其準確度及推理速度更加優越。我們的實驗測試採用 YOLOv5。

第二階段的動作辨識採用 3D-ResNet（基於殘差的一種 3D-CNN 架構）當作雙手的特徵提取器，對於人像特徵則是使用 EfficientNet（同時兼顧速度及準確度的一種 2D-CNN 架構）。這些特徵值經過合併後送入兩層以 GRU（Gate Recurrent Unit）架構的時序模型進行動作的辨識。

如表 1，實驗結果可以發現大部分手語單詞辨識的精準度（precision）及召回率（recall）表現都不錯，平均精準度為 0.89 且平均召回率為 0.85。精確率表示判斷為真的案例中，有多少比率是正確的，召回率則為所有為真的案例中，有多少比率被正確判斷。

表 1 台灣 52 個手語動作之辨識精確度及召回率

Sign name	Pre/Recall	Sign name	Pre/Recall	Sign name	Pre/Recall	Sign name	Pre/Recall
arrange	0.89 1.0	good	1.0 1.0	Look for	0.73 1.0	place	0.71 0.95
be	1.0 1.0	government	1.0 0.71	lousy	0.82 0.3	please	0.67 0.81
body	1.0 0.84	handicapped	1.0 0.88	Make up	1.0 0.8	prove	0.93 1.0
Cant hear	1.0 0.48	happy	1.0 1.0	man	0.64 1.0	self	1.0 0.26
classmate	0.58 0.9	he	1.0 1.0	may	1.0 1.0	she	1.0 1.0
deaf	0.68 0.4	He/she	1.0 0.64	May not	0.81 1.0	signlanguage	0.85 1.0
degree	1.0 0.95	hearing	0.95 1.0	medical	1.0 1.0	society	0.93 1.0
do	0.93 1.0	heavy	1.0 0.33	meeting	1.0 1.0	speak	0.92 1.0
female	0.92 1.0	help	0.93 1.0	mouth	0.49 0.75	student	0.76 1.0
few	0.95 1.0	i	1.0 0.23	My pleasure	0.9 1.0	study	0.9 1.0
free	0.5 1.0	investigate	0.68 0.94	name	1.0 0.63	thank	1.0 0.38
friend	1.0 0.64	know	0.86 1.0	Not be	1.0 0.9	they	0.69 1.0
go	0.9 1.0	light	0.68 1.0	often	1.0 1.0	you	1.0 0.84

未來我們期望在手語種類上能夠更加多元，資料的收集建構能夠更加廣泛。由於模型本身具有實時辨識的潛力，未來也希望能夠建立一個能供使用者即時辨識的系統，提供一般使用者學習對照或作簡略的手語翻譯功能。最後希望以手語辨識網路為基底，以更長的時間序列影像，收集網路上的手語短片、手語新聞為資料集，構建更加複雜的台灣手語翻譯模型，實現實時翻譯的目標。

#### 參考文獻

陳鼎勳，基於深度學習的台灣手語辨識研究，資訊工程學系碩士論文，國立東華大學，2022。

## 【AI 小學堂開講】

## 〈Claude 2—AI 聊天機器人的新選擇？〉

通識中心 陳文盛助理教授

本期我們先介紹來自美國人工智慧新創公司 Anthropic 的最新對話 AI 系統—Claude 2。在 ChatGPT 和 Google 的 Bard 問世後不久，Claude 2 悄然登場，被視為與這兩大平台競爭的有力選擇。

根據 Anthropic 的聲明，Claude 2 的主要特色包括：

## ● 核心技術：

Claude 2 是 Anthropic 推出的進階自然語言處理模型。它能提供更為詳細的回答，且在程式設計、數學問題解決及邏輯推理方面有卓越的表現。此模型特別擅長處理長篇文本，因為它具備更大的上下文理解能力並接受了最新知識的訓練。相較於 OpenAI 的 GPT 模型，Claude 2 在保持相近生成品質的同時，展現出更高的計算效率和安全性。

## ● 內容控制功能：

Claude 2 具有強大的「內容控制」功能，可根據用戶需求生成符合特定主題、風格甚至道德標準的回覆，相較於 ChatGPT，提供了更好的可控性。

## ● 擴展文件處理能力：

Claude 2 允許用戶上傳多個 10MB 大小的文件，並能理解文件內容以提供相應的回覆，大大擴展了其文件處理能力。

## ● 高度適應性：

Claude 2 的語言模型允許 100k token 的輸入，可根據不同情境和特定需求進行個性化設定，提供了更多的輸入容量以描述和定制 Claude 2 的行為。

## ● 大型檔案處理分析能力：

Claude 2 具備強大的大型檔案處理和分析功能。它可以快速篩選法律文件中的關鍵內容，並生成摘要報告，協助律師掌握案件要點。Claude 2 的智能分析算法，能準確抓取文件中的法律概念和要素，大幅節省人工閱讀時間。使用 Claude 2，律師可以更專注在高層次的法律工作，提升工作效率。

Claude 2 和 ChatGPT 在用戶端的操作相似，但兩者在 AI 工具的內部運作上有所不同。Claude 2 專注於創建“有幫助、無害、誠實”的 AI，而 OpenAI 的最新產品 GPT-4 則有不同的重點。表 2 提供了 Claude 2 與 ChatGPT 的詳細比較。

表 2 Claude 2 和 ChatGPT 的比較

特徵	Claude 2	ChatGPT 3.5
開發商	Anthropic	OpenAI
訓練資料	高品質人工標註資料	網絡爬蟲資料
模型規模	數十億參數	數百億參數
資料內容長度	100K token	16K token
能力	溝通能力佳 安全性高	技術知識廣 計算能力強
優點	自我監控 理解力好，對話連貫	知識量大 計算快，開源代码
缺點	知識有限 計算能力弱	易被誤導 上下文聯繫差
費用	免費使用	免費使用

Claude 2 的使用方式和 ChatGPT 很類似。

您只需前往 Claude 2 網站：

<https://www.anthropic.com/index/claude-2>（如圖 5），選擇「Talk to Claude」。

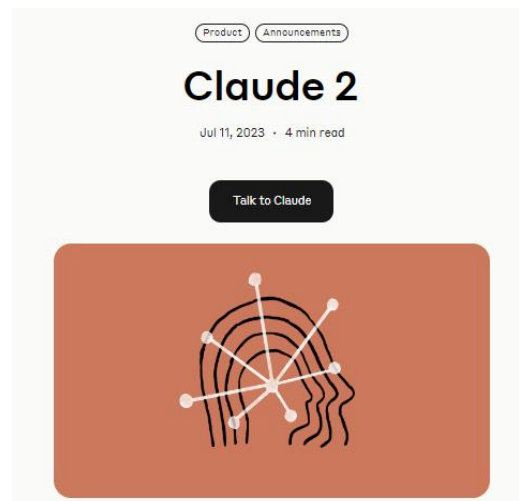
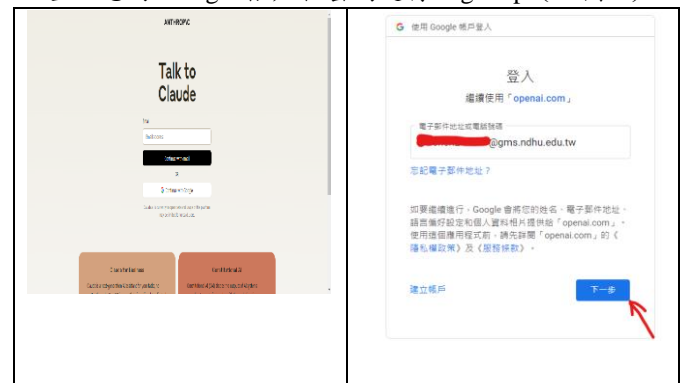


圖 5 Claude 2 網站

登入您的 Google 帳戶和密碼進行 Sign up（如圖 6）。





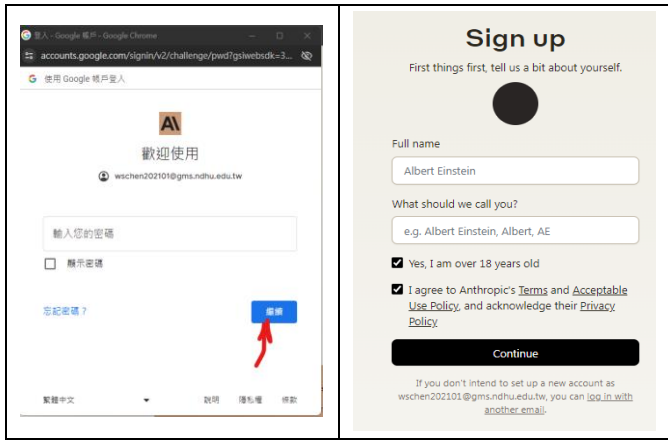


圖 6 登入 Google 帳戶進行 Sign up

並進行電話簡訊認證 (如圖 7)。

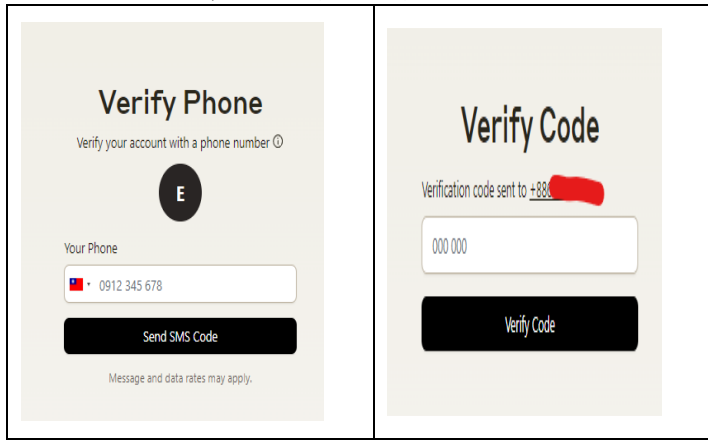


圖 7 電話簡訊認證

即可進入 Claude 2 歡迎和說明畫面 (如圖 8)。

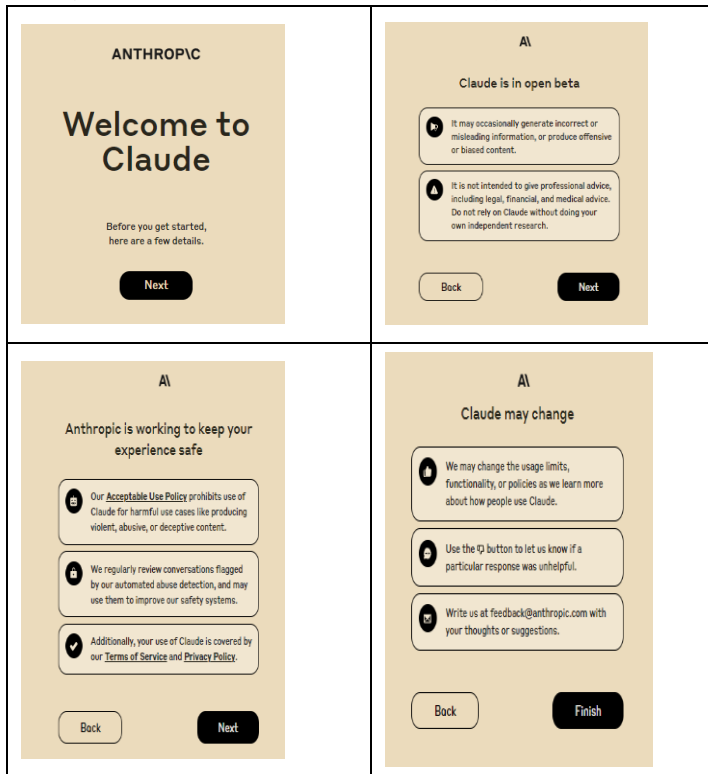


圖 8 Claude 2 歡迎和說明畫面

要使用 Claude 2，請在 Claude 2 首頁輸入問題或提示。Claude 2 會盡力以有意義的方式回答您的問題。Claude 2 首頁主要分為三個部分：左側的對話紀錄、下面的提問框和上方的工具框 (如圖 9)。

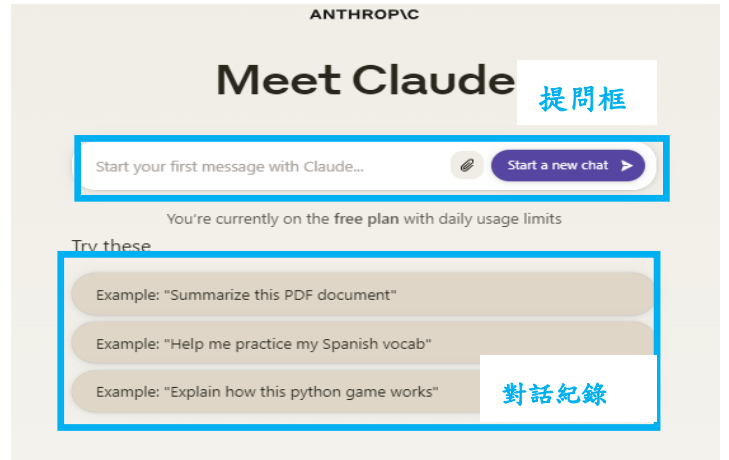


圖 9 Claude 2 首頁

Claude 2 可以直接上傳 PDF，CSV，TXT 等檔案作為輸入文檔，再請 Claude 2 根據文檔回答我們的提問。例如，我們可以在首頁中選擇迴紋針符號，上傳「性侵害犯罪防治法」的 PDF 文件給 Claude 2 (如圖 10)，然後在訊息框中輸入「請他做一個 100 字摘要」的 Prompt，再點選 ➤ 箭頭符號送出 (如圖 11 所示)。

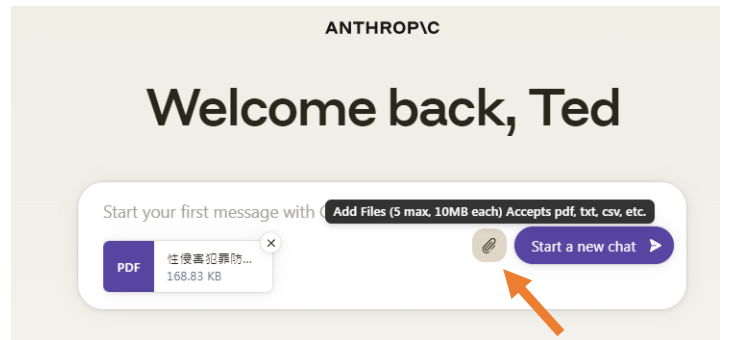


圖 10 上傳「性侵害犯罪防治法」的 PDF 文件



圖 11 送出 Prompt

訊息送出後，Claude 2 就會直接給我們所需要的回應（如圖 12）。回應區塊中可以選擇 Copy 複製所需要的訊息或是按 Retry 得到新的回應。

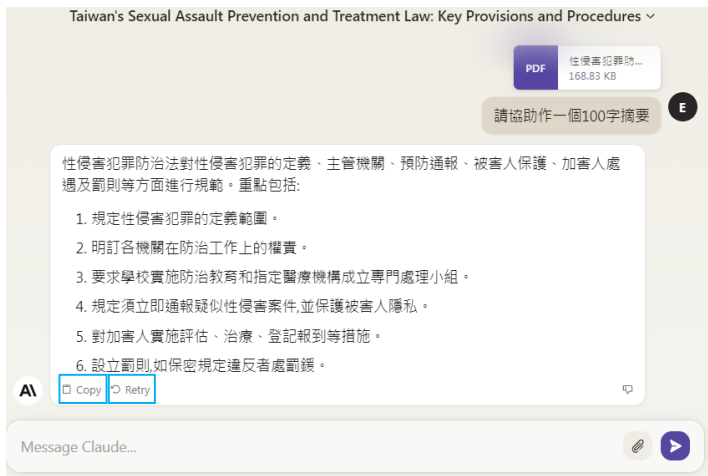


圖 12 Claude 2 回應區塊

## 【AI 跨域徵人帖】

AI 人才能力檢定 2024 年上路，半導體、資通、醫療、金融、政府、製造、資安七產業先行

<https://futurecity.cw.com.tw/article/3184>

台灣人工智慧學校於九月 15、16 日舉辦「台灣人工智慧年會」，宣布明年將推動「AI 能力檢定」，目前已選定 15 間企業、七大產業試行。以下為台灣人工智慧學校校務長蔡明順的演講精華。（未來城市，陳芳毓，2023/9/21）