

應用深度影像學習辨識醫院盤餐食物

郭雅婷^{1†} 范掀裕^{2†} 簡國龍^{2,3} 朱瑩悅⁴ 洪士林⁵ 張智星⁶ 陳珮蓉^{1*}

Application of deep image learning for recognizing hospital tray meals

Ya-Ting Kuo^{1†}, Hsien-Yu Fan^{2†}, Kuo-Liong Chien^{2,3}, Ying-Yueh Chu⁴, Shih-Lin Hung⁵,
Jyh-Shing Roger Jang⁶, Pey-Rong Chen^{1*}

¹Department of Dietetics, National Taiwan University Hospital, Taipei, Taiwan

²Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan

³Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

⁴Department of Nutrition and Health Science, Chinese Culture University, Taipei, Taiwan

⁵Department of Civil Engineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

⁶Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan

†Y.T. Kuo and H.Y. Fan contributed equally to this work and share first authorship.

(Received: February 6, 2024. Accepted: May 2, 2024.)

Abstract Background: Deep learning models for image recognition have been widely applied, including the identification of simple food items. However, the accuracy of recognition in the complex dishes served in hospital tray meals remains underexplored.

Objective: Building upon a foundation of regular hospital diets, this study utilizes hospital tray meals as test subjects. Employing artificial intelligence deep learning models, the aim is to investigate the recognition accuracy of various dishes within tray meals. The results will serve as a basis for future development of nutritional assessment software.

Methods: This research collected 1,754 photos of regular hospital diet tray meals. Standard tray patterns, based on hospital menu content, were used to segment each photo into five sections labeled as cereals, appetizer, side dishes, main dish, and vegetables. A total of 8,770 food photo segments were collected. The Residual Neural Network (ResNet) was chosen to execute the food photo recognition model. Eighty percent of the data was randomly assigned for training and validation, with the remaining 20% used for testing.

Results: The automatic recognition model was tested on 8,770 food photo segments, achieving an overall recognition accuracy of 98.2%. Further analysis revealed specific recognition accuracies from highest to lowest: cereals (99.7%), appetizer (99.1%), side dishes (98.7%), vegetables (97.4%), and main dish (96.4%).

Conclusion: The image recognition capability of deep learning applied to dish identification on standard meal trays demonstrates high accuracy.

Key words: deep learning, convolutional neural networks, hospital tray meals

* Corresponding author: Pey-Rong Chen, RD, PhD

Address: No.7, Chung Shan S. Rd. (Zhongshan S. Rd.), Zhongzheng Dist., Taipei City 100225, Taiwan (R.O.C.)

Tel: +886-2-2356-2789

Fax: +886-2-2341-2690

E-mail: prchen@ntuh.gov.tw

前 言

飲食控制在住院病人疾病照護中佔有很重要的一部份。在執行營養照護計畫和監控時，營養評估

為首要步驟。飲食評估常見的做法是由營養師以 24 小時飲食回憶法詢問病友，或請病友提供書面或拍照飲食紀錄來進行，然而這些方法受限於時間、病友對食物份量的認知程度以及配合度等因素影響⁽¹⁾，不僅過程耗時且正確性令人存疑。因此，若開發人工智慧輔助的膳食評估工具，將能加速營養師對飲食攝取量評估之效率⁽²⁾。

深度學習 (deep learning) 是人工智慧 (artificial intelligence, AI) 領域中的一種學習技術，其主要特色在於能夠自動學習特徵 (feature)。深度學習法是以卷積神經網路 (convolutional neural network, CNN) 為主要核心⁽³⁾，進行影像分類模型訓練的方式。卷積神經網路是參考人的大腦視覺組織來建立的深度學習模型，它的組成主要包含幾個主要部分，分別為卷積層 (convolution layer) – 用於提取辨識的主要特徵、池化層 (pooling layer) – 過濾掉不重要的特徵，降低計算的複雜性，以及全連接層 (fully connected layer) – 用於最後整合特徵與輸出辨識分類結果⁽⁴⁾。透過這些分層功能，CNN 能夠自動學習並捕捉輸入圖像中的階層特徵，因此比起傳統的機器學習方法需手動設計，需要由專家不斷提供特定問題的特徵相比，CNN 深度學習法在給予足夠大量的學習樣本後即可透過系統自主學習提升辨識準確率⁽⁵⁾。

人工智慧輔助的膳食評估工具已陸續被用於協

助膳食評估，辨識食物種類、大小和特徵^(6,7)。這些工具利用深度學習演算法和各種資訊來源，如食物照片、條碼掃描和文字輸入等，追蹤個案的飲食情況^(2, 8-10)，使個案的膳食資訊能夠即時分析，並依據個案的飲食攝取量提供相關的反饋和建議。許多國外研究已提出了各種人工智慧輔助評估工具，包括食物圖像方法，但由於飲食文化與料理呈現方式不同的緣故，這類工具在台灣醫院飲食以及中式餐點應用上的可行性還有待評估。深度影像學習雖然在辨識單一食物，例如：麵包、水果、蛋等圖像辨識技術上已經有很高的正確性，然而中式的菜餚具有較多如菜肉拌炒的混合菜餚，因此深度影像辨識技術在這類混合多種食材菜餚的辨識正確性還有待進一步探討。

方 法

一、資料收集

本研究係以某醫學中心住院病人普通飲食盤餐為測試物，使用智慧型手機作為拍照工具，針對同一款式的日式餐盒進行照片拍攝，圖一為日式餐盒的樣式。拍攝 16 天循環菜單的午晚餐，包含了 32 套餐點 (包括主食在內約 138 道菜)。最終本研究共收集 1,754 張普通飲食的日式餐盒照片。



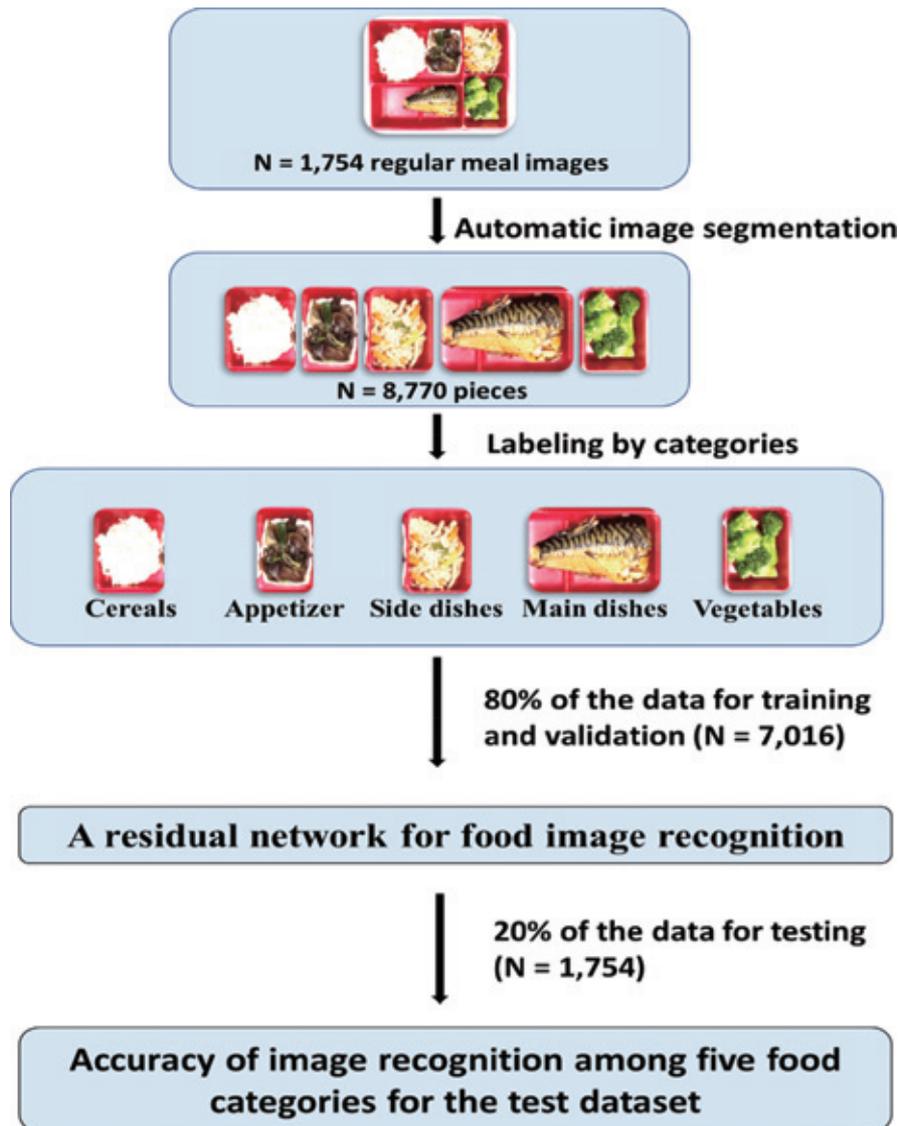
圖一 盤餐拍攝使用之日式便當盒。此圖像被分割成五個類別，分別標記為 A 至 E。主食 (A)、小菜 (B)、副菜 (C)、主菜 (D) 和青菜 (E)

Figure 1. The Japanese bento box used for plate meal photography. This image is segmented into five categories, labeled A to E. These include the cereals (A), appetizer (B), side dishes (C), main dishes (D), and vegetables (E).

二、預處理

圖二顯示研究的流程。我們將 1,754 張盤餐照片的影像分割成 5 個類別 (A-E)，分別標註為主食 (A)、小菜 (B)、副菜 (C)、主菜 (D) 和青菜 (E)。總共收集的 8,770 個食物圖像被隨機分派 80% 的數據用於訓練和驗證，剩餘的 20% 用於測試。對於大數據集，訓練集和驗證集的理想比例 8:2 應該是合適的⁽¹¹⁾。圖像分割先於預處理階段執行，以減少手動標註的工作。本研究使用物體檢測模型 (object detection model)，此模型功能為定位

圖像中每道菜的位置並辨識這道菜屬於的類別。當這個過程完成後，Inception 網絡被用來作為神經網絡特徵提取的骨幹結構⁽¹²⁾。Inception 網絡是深度學習的一種模型結構，主要用來處理圖片相關的任務。由於在辨識一張圖片時，主要辨識的特徵可能有大小不一致的情形，而 Inception 網絡的特點就是同時使用了多個不同尺寸的卷積核 (用來提取圖片特徵的工具)，可以同時納入考慮這些不同大小的特徵，從而更好地辨識圖片中的內容，這有助於縮短訓練時間。



圖二 使用殘差網絡分類的模型研究流程圖

Figure 2. Flowchart of the proposed residual neural network based classification model⁽¹³⁾.

三、影像辨識步驟

針對每個類別進行了 5 個步驟的食物識別。首先，使用方形框標註餐盤中食物的位置以及該食物的類別。其次，將數據分為 3 個數據集，用於後續訓練、驗證和測試使用。第三，使用訓練和驗證數據集來訓練一個深度學習 CNN 模型。在這個步驟中，神經網絡模型以殘差網絡 (residual neural network, ResNet) 為其骨幹結構 (使用 Adam 優化器的 ResNet-50)⁽¹⁴⁾，採取 3 通道 (RGB) 的方式。殘差網絡的功能是允許部分運算輸入跳過一個或多個神經層、卷積層直接傳到輸出層。這樣的設計幫助我們在神經網路很多層下，也能夠更容易訓練並計算最後輸出物體的方框位置及個別分類類別的預測機率。然後，循環訓練並參考驗證資料集結果進行超參數優化。第四，評估測試數據集的模型性能，並評估其在分類類別方面的準確率。最後，基於關鍵績效指標 (key performance indicators, KPI) 進行模型性能測試。

四、準確率評估方法

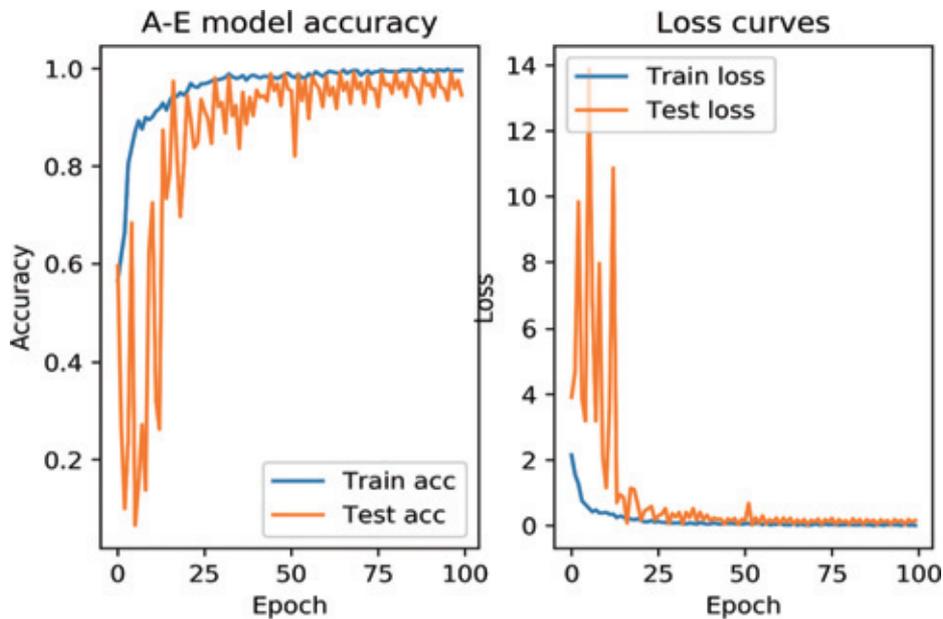
使用混淆矩陣 (confusion matrix) 來評估對菜品的識別準確率。混淆矩陣是運用於評估一個模型

準確率有多高，垂直軸表示正確答案，水平軸表示模型的預測值。如果模型預測的結果完全正確，對角線上的值將最大化，其他區域將為零。所有分析都使用 Python 3.6 作為開發語言進行。開發工具採用 Vim 8、TensorFlow 1.14、CUDA 10、CuDNN 7.5。測試環境則採用 Ubuntu 16.04 和搭載 4 個 NVIDIA-TESLA V100 GPU 的 DGX-station。

結 果

一、總體準確率

圖三顯示 CNN 模型在食物辨識方面的準確率曲線。藍線代表使用訓練集的準確率，黃線代表使用測試集的準確率。在準確率模型中，測試數據集的初始準確率低於 0.80，但經過 25 個 Epochs (訓練輪數) 後，測試數據集的準確率增加到接近 0.90。模型的損失值 (loss value) 是用來度量模型預測輸出與實際目標之間差異的指標。它是一種衡量模型性能的方式，通常越低代表模型越準確；損失值越高，表示模型預測的輸出與實際目標之間的差異越大。測試結果中，損失值在神經網絡訓練的



圖三 殘差神經網絡在每個訓練輪數的準確率 (左圖) 和損失值 (右圖) 的性能圖

Figure 3. Graphs of the performance of the accuracy (left plot) and the loss function (right plot) of the residual neural network in each epoch. The abbreviation "acc" represents "accuracy".

早期階段迅速收斂。大約經過 25 個 Epochs 後，損失曲線趨於平緩，顯示模型在 25 個 Epochs 後已經收斂。此模型最終達到了 97% 的高準確率，顯示 CNN 模型在食物識別方面的優越性。進一步對每個類別的準確率進行分析（詳見表一），主食為 99.7%，小菜為 99.1%，副菜為 98.7%，主菜為 96.4%，青菜為 97.4%。

二、主食類的辨識準確率

在主食照片的詳細辨識準確率上除了糙米薏仁飯的準確率為 0% 外，其他主食圖像，包括：白飯、胚芽飯、燕麥飯、紫米飯、五穀飯、堅果雜糧飯、山藥飯和地瓜飯的準確率都達到了 100%。

三、小菜類的辨識準確率

在小菜照片的詳細識別準確率中，有 3 樣小菜的準確率較低，包括：九層茄子一素（94.1%）、九層茄子一葷（95.2%）和三杯鮮菇（93.8%）。其他小菜圖像的準確率皆達到了 100%。

四、副菜類的辨識準確率

在副菜照片的詳細識別準確率中，大多數副菜照片的準確率大於 95%，但金針菇炒肉絲（80%）、野菇咖哩燉豬肉（94.4%）、木耳炒茼蒿（83.3%）的準確率較低。

五、主菜類的辨識準確率

於主菜照片的詳細識別準確率上，大多數主菜圖像的準確率大於 95%，但紅燒豬排（37.5%）、五香滷雞排（86.7%）、香蔥雞腿（50%）、香橙翅腿（75%）的準確率較低。

六、青菜類的辨識準確率

在青菜照片的詳細識別準確率中，3 樣青菜圖像的準確率低於 100%，包括枸杞高麗菜（80%）、炒高麗菜（97%）、大白菜炒胡蘿蔔（90.1%）和炒萵苣（93.2%）。其他青菜圖像的準確率均為 100%。

七、完整盤餐辨識準確率實測

本研究另外取某日普通盤餐影像 6 組，以確認系統模型對完整盤餐辨識準確率。共測試分析了 6 組盤餐照片，同樣每張照片切割為 5 個部分，總計 30 個片段分別測試。最終分析結果顯示，CNN 模型正確辨識了 29 個片段，準確率為 96.6%，只有一道小菜（炒大黃瓜）辨識錯誤。

討 論

一、主要發現

此模型使用 8,770 張食物圖像進行測試，整體識別準確率為 98.2%。對每個類別進行進一步分析發現，識別準確率由高到低依次為：主食（99.7%）、小菜（99.1%）、副菜（98.7%）、青菜（97.4%）和主菜（96.4%）。

儘管目前已有各種使用深度學習方式進行自動食物辨識的系統，但這是首次將卷積神經網絡應用於辨識醫院盤餐照片的研究。目前至少已有 17 項食物識別系統之相關研究發表⁽¹⁵⁾。其中，大多數系統都是使用 CNN 模型，並表現出整體準確率在 0.86 或以上⁽¹⁶⁻²⁵⁾。例如，Teng 等人提出了一個用於中國食物識別的 CNN 模型，包含來自 25 個飲食套餐

表一 測試數據集中五個食物類別的影像辨識準確率

Table 1. Accuracy of image recognition among five food categories for the test data set.

類別†	總張數	正確數	錯誤數	辨識準確率 (%)
主食 (A)	337	336	1	99.7
小菜 (B)	333	330	3	99.1
副菜 (C)	378	373	5	98.7
主菜 (D)	357	344	13	96.4
青菜 (E)	349	340	9	97.4

†每張圖像被分割為五個類別，標為 A 到 E。主食 (A)、小菜 (B)、副菜 (C)、主菜 (D) 和青菜 (E)

的 8,734 張圖像。他們的模型擁有 5 層深度的卷積神經網絡，達到了 0.99 的準確率⁽¹⁷⁾。本研究得到 0.98 的整體準確率，結果與其相似。然而，很少有研究的結果為針對每個食物類別進行準確率分析。透過切割模板的分類模式，使系統可以在學習或辨識前根據每個區域的類型特性預先搜尋特定目標菜餚，有助於避免菜餚錯置的問題，例如將青菜誤認為蔬菜為主的小菜之情形。此外，亦能有助於日後發展獨立類型智慧模型時，不需重新將資料庫中的影像進行分類標註。本次研究提供了每個食物類別的準確率，主食為 99.7%，小菜為 99.1%，副菜為 98.7%，主菜為 96.4%，青菜為 97.4%。由此結果顯示，相較其他食物圖像，主菜和青菜在此次研究模型識別準確率有待再進一步提升。

二、影響食物辨識的關鍵因素

影像辨識模型在學習不同類別食物時，會受到食材本身特徵與烹調方式影響，例如：青菜部分，因深色葉菜類不論是顏色或形狀方面相似度都非常高，系統不易找到明顯差異的特徵去做區分，因此影響準確率；主菜部分則受烹調完的成品色澤過於相似，例如：紅燒豬排與茄汁豬排皆是使用相同的豬里肌部位進行烹調，成品也因調味料中都使用深色醬料而呈現類似的色澤。而影響系統辨識率。

除了食材本身因素外，系統學習時亦會受到幾個因素對識別準確率的影響。首先，辨識準確率受到拍攝距離的影響。在影像辨識的學習過程中，用於分割的模板圖案是固定的。因此，如果拍攝距離過遠，可能造成圖像周圍有過多的空白處或兩個菜品出現在同一個分割圖像內，而導致判斷錯誤。如需進行相關研究，可使用標準相機或手機固定器來確保照片拍攝時的比例，或者裁剪圖像周圍的空白處以降低干擾因素。未來在運用軟體開發時，操作界面可以提示使用者在特定框內拍攝，以確保圖像邊框比例的一致性。

其次，辨識準確率還受到菜名設計的影響。醫院飲食之菜單名稱會配合病友禁忌而有微小食材差異，同樣類型菜色可能會有不同名稱，例如：葷素的註記，因此系統模型會分類產生兩種不同學習類別，但在影像辨識上，同樣一道菜色的葷素版本可能在視覺上難以區分。例如，青菜類的葷素食材差異是在使用蒜末或薑絲，份量極少，且照片上未必會拍攝到特定材料，卻可能造成系統模型辨識時無

法判定其葷素而降低最終辨識準確率。因此後續應用上，是否有區分兩者的必要性或是選擇合併皆需要再考量。

再者，訓練樣本數量，也可能會對識別準確率產生影響。當學習樣本越多，準確率將越趨近於正確值。國外許多研究選擇以網路圖像作為研究基礎，以利於在短時間內獲取大量相似圖像^(26,27)。然而，本實驗所使用的影像來源是依賴實際醫院盤餐拍攝模式所獲取，因此，在照片獲取數量和效率方面，較難在短時間內獲得足夠的學習樣本。本研究中，許多辨識錯誤與樣本數量不足有關。從分析結果顯示，膳食照片數量高於 20 張的菜色其辨識率大部分可高於 95%，但若菜色型態相似度偏高，如：深綠色葉菜（例：炒萵苣），則需提供更多的張數作為訓練模型。

最後，這項研究受到照片解析度的限制。深度學習模型本身具有一定誤差的容忍度，但解析度不足可能導致辨識錯誤或學習效果不佳。然而，此缺陷可以透過增加訓練數據量來彌補。未來開發運用軟體，可要求使用者在拍攝前調整相機對焦，以確保影像品質與解析度，對整體系統辨識準確率有所助益。

運用深度學習模型辨識菜餚仍有許多挑戰，在 2023 年由 Shao, W 等人發表的一篇文獻中也提到，在現階段影像辨識技術上發現菜餚中的食用油，如：橄欖油、植物油、魚油等，在影像中很難被檢測⁽²⁶⁾。另外，拍攝過程中因為中式菜餚有許多葷素混合菜，而當多種食品堆疊在一起時，會導致部分食物被遮蔽和覆蓋，導致模型得到不正確的辨識結果。

三、限制性

由於本研究之照片來源為研究期間病人訂購的醫院普通飲食實際供應餐盒內容，因此各類別的食材樣本數容易受到病友個人飲食禁忌與喜好而有所落差。以主食類舉例，由於大多數病友主食類別選擇以白飯與胚芽飯居多，因此其他種類的主食樣本數相對不足。此外，本研究對菜餚之辨識只限於品項名稱，無法進行份量的估算。當運用於營養量評估時，必須搭配醫院標準菜單系統的營養成分資料庫。在影像辨識技術發展上，具有辨識具體食物份量大小的部分尚待努力。

四、實務應用

本研究結果顯示，運用深度學習之技術可以有 98% 以上的準確率辨識醫院盤餐中的每一道菜色，以此運用於開發智慧型手機應用軟體，透過拍照可得知飲食菜單內容，若配合將醫院標準菜單之營養成分資料庫連結建置，即可估算營養攝取量。然而，本次研究僅可處理完整餐盤內容，若只攝取 1/2 或 2/3 等情形，則需設計互動式資訊欄位，選取實際攝取比例，以估算實際營養攝取量。因此，依據本次的研究成果，在實務運用上，需要搭配食譜的營養成分資料庫，以及依據營養評估需求之相關應用軟體設計，這在醫院膳食由營養師設計標準化食譜且有營養量的情況下，將可實務運用。

未來研究需要再努力的方向，從辨識食物到估計份量大小，此為營養攝取評估下一個重要里程碑。由於 AI 技術的進步，3D 立體深度學習是未來需要精進的重要課題，同時，使用者相機的立體拍照功能也要能普及，方可整體優化以手機拍照評估營養攝取量之功能。

結 論

本次研究結果，深度影像辨識技術運用於餐點菜餚辨識效果上有優良表現，經過訓練的系統辨識準確率可達到 98.2%，後續搭配上營養成分資料庫的建立，不僅可加速營養師審閱病人提供的飲食照片資料的時間。若套用於 APP 上，更可以有利於將這份資料庫擴大應用至病人返家後自行評估營養攝取量，強化各類慢性疾病飲食控制成效。

致 謝

感謝雲像科技公司 (aetherAI) 對數據分析的協助。

資 金

本研究得到國立臺灣大學醫學院附設醫院 (NTUH.108-004279) 的補助。

倫理審查

不適用。

利益衝突

本文作者聲明無利益衝突。

參考文獻

1. Cade, J.E., et al., *DIETNET: Best Practice Guidelines for dietary assessment in health research*. 2017. **15**(1): p. 1-15.
2. Lu, Y., et al. *An Artificial Intelligence-Based System for Nutrient Intake Assessment of Hospitalised Patients*, in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019.
3. Chen, L., et al., *Review of image classification algorithms based on convolutional neural networks*. 2021. **13**(22): p. 4712.
4. Alzubaidi, L., et al., *Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions*. 2021. **8**: p. 1-74.
5. Janiesch, C., P. Zschech, and K.J.E.M. Heinrich, *Machine learning and deep learning*. 2021. **31**(3): p. 685-695.
6. Boushey, C., et al., *New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods*. 2017. **76**(3): p. 283-294.
7. Van Asbroeck, S. and C.J.J.f.r. Matthys, *Use of different food image recognition platforms in dietary assessment: comparison study*. 2020. **4**(12): p. e15602.
8. Nguyen, P.H., et al., *Relative validity of a mobile AI-technology-assisted dietary assessment in adolescent females in Vietnam*. 2022. **116**(4): p. 992-1001.
9. VijayaKumari, G., P. Vutkur, and P.J.G.T.P. Vishwanath, *Food classification using transfer learning technique*. 2022. **3**(1): p. 225-229.
10. Liu, Y.-C., et al., *Deep-Learning-Assisted Multi-Dish Food Recognition Application for Dietary Intake Reporting*. 2022. **11**(10): p. 1626.
11. Joseph, V., *Optimal ratio for data splitting*. *Statistical Analysis and Data Mining*, **15** (4), 531-538. 2022.
12. Elharrouss, O., et al., *Backbones-review: Feature extraction networks for deep learning and deep reinforcement learning approaches*. 2022.
13. Li, Y., et al., *KinNet: Fine-to-Coarse Deep Metric Learning for Kinship Verification*, in *Proceedings of the 2017 Workshop on Recognizing Families In the Wild*, 2017, Association for Computing Machinery: Mountain View, California, USA. p.13-20.
14. Li, Y., et al. *Kinnet: Fine-to-coarse deep metric learning*

- for kinship verification. in *Proceedings of the 2017 workshop on recognizing families in the wild*. 2017.
15. Salim, N.O., et al. *Study for food recognition system using deep learning*. in *Journal of Physics: Conference Series*. 2021. IOP Publishing.
 16. Salim, N.O.M., et al., *Study for Food Recognition System Using Deep Learning*. *Journal of Physics: Conference Series*, 2021. **1963**(1): p. 012014.
 17. Teng, J., et al., *Recognition of Chinese food using convolutional neural network*. *Multimedia Tools and Applications*, 2019. **78**(9): p. 11155-11172.
 18. Srigurulekha, K. and V. Ramachandran. *Food image recognition using CNN*. in *2020 International Conference on Computer Communication and Informatics (ICCCI)*. 2020.
 19. Azizah, L.M., et al. *Deep learning implementation using convolutional neural network in mangosteen surface defect detection*. in *2017 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*. 2017.
 20. Pandey, P., et al., *FoodNet: Recognizing Foods Using Ensemble of Deep Networks*. *IEEE Signal Processing Letters*, 2017. **24**(12): p. 1758-1762.
 21. Martinel, N., G.L. Foresti, and C. Micheloni. *Wide-Slice Residual Networks for Food Recognition*. in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2018.
 22. Zhang, Y.-D., et al., *Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation*. *Multimedia Tools and Applications*, 2019. **78**(3): p. 3613-3632.
 23. Williams, H.A.M., et al., *Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms*. *Biosystems Engineering*, 2019. **181**: p. 140-156.
 24. Reddy, V.H., et al. *Food Recognition and Calorie Measurement using Image Processing and Convolutional Neural Network*. in *2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*. 2019.
 25. Alajrami, M.A. and S.S. Abu-Naser, *Type of Tomato Classification Using Deep Learning*. *International Journal of Academic Pedagogical Research (IJAPR)*, 2020. **3** (12): p. 21-25.
 26. Singla, A., L. Yuan, and T. Ebrahimi. *Food/non-food image classification and food categorization using pre-trained googlenet model*. in *Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management*. 2016.
 27. Rich, J., H. Haddadi, and T.M. Hospedales. *Towards bottom-up analysis of social food*. in *Proceedings of the 6th international conference on digital health conference*. 2016.

應用深度影像學習辨識醫院盤餐食物

郭雅婷^{1†} 范掀裕^{2†} 簡國龍^{2,3} 朱瑩悅⁴ 洪士林⁵ 張智星⁶ 陳珮蓉^{1*}

¹國立台灣大學醫學院附設醫院營養室

²國立台灣大學公共衛生學院流行病學與預防醫學研究所

³國立台灣大學醫學院附設醫院內科部

⁴中國文化大學保健營養學系

⁵國立陽明大學土木工程學系

⁶國立台灣大學資訊工程學系

(收稿日期：113 年 02 月 06 日。接受日期：113 年 05 月 02 日)

摘要 背景：深度影像學習 (deep learning) 模式對圖像進行辨識已廣泛被應用，包含簡單食物辨識。然而在醫院盤餐的食物，因複雜度較高，其辨識的正確率尚待研究。

目的：以醫院普通飲食為基礎，將盤餐作為測試物，運用人工智慧深度學習模型，旨在探討盤餐中各類菜餚的辨識準確率，並將結果作為未來研發營養評估軟體之依據。

方法：本研究收集 1,754 張醫院普通飲食盤餐照片。將標準樣式之餐盤根據醫院菜單內容，將每張照片分割成 5 個片段，分別標註為主食、小菜、副菜、主菜和青菜。最終共收集 8,770 張食物照片片段。選擇殘差神經網絡 (residual neural network, 簡稱 ResNet) 執行辨識食物照片的模式，將片段隨機分派 80% 數據用於訓練和驗證，剩餘 20% 用於測試。

結果：自動辨識模型經 8,770 張食物照片測試，整體辨識正確率達 98.2%。進一步分析各類辨識準確率從高到低分別為：主食 (99.7%)、小菜 (99.1%)、副菜 (98.7%)、青菜 (97.4%) 和主菜 (96.4%)。

結論：深度影像學習的影像辨識能力運用於標準餐盤之菜餚辨識上有良好的準確率。

關鍵字：深度影像學習、卷積神經網路、醫院盤餐食物

* 通訊作者：陳珮蓉博士

通訊地址：台北市中正區中山南路 7 號

電話：+886-2-2356-2789

傳真：+886-2-2341-2690

電子郵件：prchen@ntuh.gov.tw

