

國立臺東大學 資訊管理學系

115(級)畢業專題報告書

人臉情緒辨識之 台東美食推薦系統

指導教授：陳志全 教授

學生：何品萱(11112115)

蔡承哲(11112120)

邱炫瑀(11112125)

陳郁婷(11112130)

張惟鈞(11112140)

劉柏毅(11112144)

中 華 民 國 114 年 11 月

摘要

隨著數位科技的快速發展，人臉辨識與情緒分析技術已廣泛應用於安全監控、智慧客服、健康照護及教育等領域。近年研究指出，個體的情緒狀態會顯著影響飲食偏好與食慾表現，例如在壓力或焦慮情境下，人們更傾向選擇高糖或高碳水化合物食物 (Kaneko, 2018)。然而，將情緒辨識技術與餐飲推薦系統結合的研究仍相對稀少，尤其針對特定地區美食進行個人化推薦的實作更為罕見。因此，本研究提出一套情緒導向美食推薦系統，以台東地區的在地美食為資料基礎，結合人臉偵測、情緒辨識與年齡預測技術，透過影像即時分析提供個人化餐點建議，旨在提升飲食推薦的精準度與互動性。

系統流程方面，首先利用 OpenCV 進行人臉偵測與區域擷取，以確保輸入影像能精準定位目標臉部；接著採用 DeepFace 進行情緒辨識，將表情分類為高興、悲傷、憤怒、驚訝、中性等七種情緒。年齡預測則使用適合亞洲臉部特徵的 InsightFace 模型進行推論，並以 ± 5 歲誤差範圍作為準確判定標準。為提升實務應用效能，系統針對推論參數進行優化，以兼顧辨識速度與準確率。前端呈現部分採用 Gradio 建構互動介面，支援即時影像上傳與拍攝，並即時顯示情緒與年齡分析結果，同步推薦對應的台東美食建議。推薦邏輯基於文獻中所提出的「情緒—食物對應關係模型」(Egon P. Köster, Jozina Mojet, 2015)，例如在偵測到壓力或焦慮情緒時，系統會推薦具有放鬆效果的甜點或溫熱飲品；當偵測到愉悅情緒時，則推薦具有活力感的餐點。

整體系統最終部署於 Hugging Face Spaces 平台，使用者可透過瀏覽器直接體驗辨識與推薦流程，無需額外安裝軟體，進一步提升研究成果的可及性與展示性。

系統測試採用 500 張影像進行效能評估，情緒辨識的整體準確率達 73%，年齡預測的正確率則依 ± 5 歲誤差範圍計算。結果顯示，本系統在單人即時辨識與餐點推薦情境下，能提供穩定且流暢的互動體驗，證明以情緒導向的美食推薦在實務應用上的可行性。

關鍵詞：人臉情緒辨識、美食推薦、深度學習、卷積神經網路

目 錄

摘要.....	i
目錄.....	ii
圖目錄.....	iii
表目錄.....	iv
第一章、導論.....	1
1.1 研究背景.....	1
1.2 研究目的.....	2
第二章、文獻探討	3
2.1 美食推薦依據之食物心理學	3
2.2 深度學習.....	5
2.3 卷積神經網路（Convolutional Neural Network, CNN）	7
2.4 DeepFace.....	9
2.5 OpenCV	10
2.6 Tensorflow	12
2.7 InsightFace.....	13
2.8 年齡差距正確錯誤依據	14
2.9 Gradio.....	15
第三章、研究方法	17
3.1 實驗流程.....	17
3.2 系統架構.....	18
3.3 年齡系統與模型設定.....	20
3.4 研究流程.....	22
第四章、實驗結果與分析	24
4.1 資料蒐集與儲存.....	24
4.2 圖像與人臉處理.....	25
4.3 情緒與屬性辨識模型.....	27
4.4 美食推薦系統.....	31
4.5 系統部署與應用介面設計：Hugging Face Spaces.....	33
5.1 研究成果.....	38
5.2 問題探討.....	39
5.3 研究限制.....	39
5.4 未來展望.....	40
參考文獻.....	42

圖目錄

圖 1 流程架構圖	18
圖 2 美食推薦系統架構	19
圖 3 研究方法流程圖	23
圖 4 人臉採取示意圖	25
圖 5 各情緒分別對應的年齡標籤以及美食推薦	32
圖 6 前端網頁呈現	36
圖 7 前端網頁呈現	36
圖 8 前端網頁呈現	36

表目錄

表格 1《食物情緒大解密》中各類食物對應的含意	5
表格 2、各情緒對應準確率表格 2 各情緒對應準確率.....	30

第一章、導論

1.1 研究背景

隨著人工智慧（Artificial Intelligence, AI）與深度學習（Deep Learning）技術的迅速發展，人臉辨識與情緒分析已成為眾多智慧應用的重要技術基礎。近年來，情緒辨識技術在行銷分析、智慧客服、教育輔助系統、安全監控以及心理健康評估等領域均展現了廣泛的應用潛力。有研究探討出情緒辨識技術在醫療／家居環境中的應用，並指出「智能裝置 + 情緒辨識」已發展出應用系統與模型 (Guo, R et al., 2024)。其核心價值在於能夠即時掌握使用者的情緒狀態，進而為互動系統引入情境感知能力（Context Awareness），提供更具個人化與動態適應性的使用體驗。

然而，在智慧餐飲與飲食推薦領域，目前多數系統仍以使用者過往的點餐紀錄、評分資料或靜態偏好設定作為主要依據，缺乏對使用者當下情緒的即時感知與響應。(Yildirim, Akbulut, & Catal, 2023)，有研究指出「目前尚無系統在用餐時完全達成即時情緒偵測、且多數仍基於傳統問卷或靜態評價」。文中對於即時情緒偵測與用餐行為／推薦系統的結合，提出了尚待加強的方法 (Yildirim, Akbulut, & Catal, Analysis of facial emotion expression in eating occasions using deep learning, 2023)，並且傳統的餐廳推薦系統「沒有考慮使用者的情緒狀態」，因此研究者提出將情緒智慧（emotional intelligence）納入餐廳推薦流程 (Sureliya, S, et al., 2024)。這種方法雖在一定程度上能反映長期喜好，但無法捕捉情緒波動對飲食選擇的即時影響。例如，同一位使用者在心情低落時與情緒愉悅時，其對食物種類、口味與用餐環境的需求可能截然不同。

心理學與行為科學的研究已多次證實情緒與飲食行為之間的密切關聯。有研究者提出的系統回顧分析成人情緒（包括正／負情緒）與飲食行為（如食量增加、不健康選擇）的關係。指出壓力、憂鬱、悲傷等負情緒常導致飲食量增加或營養不佳的選擇 (Devonport, Nicholls, & Fullerton, 2019)。若能將即時情緒偵測融入餐飲推薦系統，不僅能讓推薦結果更契合當下心理狀態，還有助於提升使用者的滿意度、互動感與信任度。

近年來，隨著攝影設備普及與邊緣運算技術的進步，透過即時人臉影像進行情緒分析已不再是高成本或高延遲的作業。結合人臉辨識與情緒分析的智慧餐飲服務，能夠在使用者與系統互動的時間點即時完成影像擷取、表情分析與餐飲建議的生成，實現從「靜態推薦」到「動態情境化推薦」的轉變。本研究基於此背景，運用深度學習技術進行人臉辨識與情緒分析，並將分析結果與飲食心理關聯模型結合，提供對應的台東美食推薦，期望能開啟智慧餐飲服務的新應用場景與商業模式。

1.2 研究目的

本研究旨在結合人臉辨識與情緒分析技術，建構一套即時感知使用者情緒並依此推薦美食的智慧系統。首先，系統透過攝影鏡頭擷取使用者臉部影像特徵，利用深度學習模型精準辨識當前情緒狀態，並依據情緒與食物偏好之間的心理與行為關聯，提供符合即時情境與心理需求的飲食建議。

傳統的美食推薦系統多依賴歷史訂單紀錄、評分資料或熱門排行進行篩選，雖在功能上具備一定準確性，但忽略了使用者當下的情緒與心理狀態。情緒作為影響人類飲食選擇的重要因素，在心理學與行為科學中已被廣泛證實：例如，正向情緒可能促使人選擇甜食或創新口味，而負向情緒則可能偏好高碳水或高熱量的慰藉性食物。將情緒作為推薦決策的核心依據，有助於系統在提供建議時更貼近使用者的即時心理需求，實現更高層次的個人化體驗。因此，本研究的目標在於突破傳統推薦方法的侷限，匯入情緒辨識作為動態決策因素，提升推薦的情境感知能力與準確度。透過整合人臉影像處理、深度學習模型訓練、情緒分類與食物資料庫映射等技術，本研究將開發出一套可即時運作的情緒導向美食推薦系統，並透過實驗評估其在準確性、使用者滿意度及實際應用便利性方面的表現。

最終，本研究期望驗證情緒導向推薦系統的可行性與實用性，並探討其在智慧餐飲服務、行銷推廣及個人化健康管理等領域的應用潛力，為未來智慧化餐飲與人機互動介面的發展提供參考依據。

第二章、文獻探討

2.1 美食推薦依據之食物心理學

在現代社會中，食物已成為影響人類情緒的重要媒介。當情緒不穩定或是不安定時，就會希望自己可以改變這種狀態，因此人們常常利用吃來排不好的情緒。根據國立中興大學（許捷鳴，2016）以及《食物情緒大解密》（朵琳·芙秋，2012）一書中的說明，人們在情緒低落或壓力大時，常會特別想吃某些特定的食物，這種行為並非單純的口腹之欲，而是一種潛意識中對心理安慰的需求。書中指出，許多能帶來滿足感與幸福感的安慰食品，普遍具有高糖、高脂肪或富含碳水化合物的特性，這些成分能刺激大腦釋放多巴胺，使人感到愉悅與放鬆。

BBC 中文網 (BBC, 2021) 的報導《飲食與健康：我們對安慰食品依賴背後的科學道理》也指出，人類在壓力或不安情緒出現時，會特別想吃某種食物，因為希望透過吃東西改善壞情緒或壓力，雖然人們喜愛的安慰食品千差萬別，但那些能帶給人們滿足和幸福感的食物都具有一個共同特點，即高糖、高脂肪以及富含碳水化合物，因為這些食物能快速激發味蕾的滿足感，並在短時間內改善情緒。

此外，在 (Rozanne Kruger, et al., 2007) 所提及到，研究者更深入探討了味覺與情緒之間的聯想關係。該研究透過實驗方式，使用代表主要味道品質的四種溶液（甜、鹹、酸、苦），測試參與者在體驗這些味道後所聯想到的七種基本情緒（快樂、悲傷、生氣、驚訝、厭惡、恐懼與中立）。結果發現，甜味與快樂和驚訝呈現出高度相關，而苦味則與憤怒與厭惡顯著相關。鹹味與酸味則顯示出與多種情緒的聯想關係，反映其在情緒調節上的多樣性。這些結果說明，味覺感受會激發不同的情緒意見反應，而人們在不同情緒狀態下，確實會傾向選擇與其當下心理狀態相符或能對應的食物。

再者，根據《Marie Claire 美麗佳人》(Weng, 2023) 的報導〈科學證實這 6 種食物最能有效提升快樂感！〉指出，食物的味道能夠直接影響大腦情緒反應。例如，當心情低落時，攝取辣椒或洋蔥等

刺激性食物能帶來微弱痛感，大腦為了舒緩這種不適會分泌腦內啡，使人感受到平衡與穩定。當人處於憤怒狀態時，適量攝取甜點則能促使多巴胺分泌，帶來愉悅與舒適感。而在焦慮或心神不寧的情緒下，酸性食物如醋或檸檬，能透過刺激味蕾的酸感，緩解膠著的情緒，協助穩定心情。此外，鹹味食物中所含的鈉元素具有調節細胞與血液功能的效果，能促進新陳代謝，特別適合早起無精神或經常熬夜者，藉由如鹹粥或熱湯麵等溫熱料理，達到舒緩疲勞的目的。至於苦味食物，雖普遍不被偏好，卻能刺激神經系統，有助於提神醒腦，恢復精力。此篇文章再次印證味覺與情緒間的密切關聯，強調適當攝取對應味道的食物可成為調節情緒的有效途徑。

因此，本研究正是建構在此心理學基礎上，結合情緒與味覺聯想的概念，並進一步以台東地區的在地美食為素材來源，根據上述不同文獻與報導中的實證研究與專家觀點，整理歸納出各類食物所具有的味覺特性與情緒連結，並參考多方對於不同情緒對應食物的論述，將不同食物特性與七種基本情緒（快樂、悲傷、生氣、驚訝、厭惡、恐懼與中立）進行配對。並根據既有文獻中所統整的「情緒—食物」對應邏輯，推薦最適合當下情緒的在地料理或點心。

在實務應用上，當系統判定使用者處於悲傷情緒時，可能推薦具有甜味與溫暖感受的在地甜品，如紅豆湯或地瓜球；若使用者感到憤怒，則建議具有酸味或鹹味的料理，如檸檬魚或海鮮冷盤，幫助降低壓力賀爾蒙、緩和焦躁心情。又例如，在面對驚訝或恐懼情緒時，推薦具有安撫作用且易於消化的食物，如雞湯、粥類或花草茶；而處於快樂情緒時，則推薦能延續愉悅感的酥炸點心如炸米糕或鹹酥雞。此模式不僅提升推薦的個人化程度，也結合了行為心理與在地文化，進一步拓展智慧化飲食在情緒療癒與健康生活中的應用潛力。

食物類型	代表含意
巧克力	渴望愛
乳製品	渴望抗憂鬱
鹹零嘴	渴望舒緩壓力、憤怒、焦慮
香辣食物	渴望驅動刺激
飲品	渴望安定起伏不定的能量循環
堅果與花生醬	渴望樂趣
麵包、飯、麵	渴望安慰與平靜
餅乾、蛋糕、派	渴望擁抱、喜悅、安心
糖果	渴望快樂、獎勵、娛樂
高脂肪食物	渴望填補空虛

表格 1《食物情緒大解密》中各類食物對應的含意

本研究也將年齡納入食物推薦依據的考量，將使用者區分為兩大年齡群組，60 歲以下（青年族群）與 60 歲以上（銀髮族群），並根據不同年齡層推薦適合的美食。本研究主要聚焦於比較老年與非老年族群之情緒與飲食偏好差異，因此採用兩類區分方式，此分類方式符合世界衛生組織的定義，其指出年滿 60 歲即被視為老年人口（Organization, 無日期）。

2.2 深度學習

深度學習（Deep Learning）是人工智慧（Artificial Intelligence, AI）與機器學習（Machine Learning, ML）的重要分支，其核心思想源於類神經網路（Artificial Neural Networks, ANN）的多層結構。透過大量神經元的層疊與非線性轉換，深度學習模型能自動從原始資料中學習多層次的抽象特徵（LeCun, et al., 2015）。

深度學習的方法多樣，依照應用領域可分為影像處理、語音辨識、自然語言處理、醫學影像分析、自駕車感知系統等。其中，常見的模型包含卷積神經網路（Convolutional Neural Network, CNN）、循環神經網路（Recurrent Neural Network, RNN）、生成對抗網路（Generative Adversarial Network, GAN）以及變壓器（Transformer）等架構。其中，CNN 對於影像任務最具代表性，能自動學習局部與全域的視覺特徵，因此被廣泛應用於影像分類、物件偵測與人臉分析等任務。本研究即屬於深度學習在電腦視覺領域中的應用，聚焦於人臉影像的情緒與年齡辨識，並進一步延伸至美食推薦的互動式應用。

在人臉分析領域，深度學習的應用尤其廣泛，包括人臉檢測（face detection）、人臉辨識（face recognition）、情緒分析（facial emotion recognition）、年齡與性別預測（age and gender estimation）等任務 (Deng, 2020)。其中，DeepFace 是由 Meta AI 團隊(Facebook AI research group)開發的一套深度學習人臉辨識系統，能夠透過九層神經網路結構與三維對齊（3D alignment）技術，將人臉辨識準確率提升至接近人類水準 (Yaniv Taigman, et al., 2014)。DeepFace 能在輸入影像的過程中自動完成臉部檢測、特徵向量提取與身份比對，並可擴展應用於性別、年齡及情緒分析等多種場景。

此外，深度學習取代了傳統手工特徵（如 HOG、LBP、SIFT）成為主流方法，具備以下優勢：（1）高精度：在 LFW（Labeled Faces in the Wild）等資料集上，深度學習模型如 DeepFace (W. Zhao, et al., 2003) 準確率可達 99% 以上，接近甚至超越人類水準；（2）強健性：對光照、姿態與表情變化具有良好適應能力；（3）多任務整合：可同時進行身份辨識、年齡估計、性別分類與情緒分析等多種任務。

再者，深度學習技術在近年來的人臉辨識研究中佔據了核心地位（陳映志，2020）。相較於傳統方法，深度學習能自動從大量資料中學習到更具判別力的特徵，顯著提升了辨識準確率。卷積神經網路（CNN）特別適合處理影像任務，已被廣泛應用於人臉檢測、特徵提取與身份驗證等多個階段。許多研究證實，透過深層網路結構，

系統能在光照、姿勢變化及表情不同的情況下，依然保持穩定表現。這使得深度學習不僅推動了人臉辨識技術的進步，也擴展了其在安防監控、智慧門禁及人機互動等領域的應用潛力。國防大學中一篇為植基於深度學習之快速人臉辨識機制（林宗儒，2024）之研究也說明到，深度學習技術對人臉辨識領域的貢獻主要表現在其強大的資料處理能力和自動特徵提取能力。這些技術克服了傳統算法在面對複雜變具、部分遮擋、不同角度和光照條件下的局限性。深度學習模型，如卷積神經網路(CNN)，通過深層次的非線性轉換，能夠學習到從局部紋理到整體結構的多層次特徵（LeCun, et al., 2015），這些特徵對於提升辨識的準確性至關重要。

最後，為了讓深度學習模型在實務中更易部署，本研究同時採用 OpenCV（Open Source Computer Vision Library）作為影像處理與前處理工具。OpenCV 是一套跨平台、開源的電腦視覺與影像處理函式庫，內建高效的人臉檢測演算法（如 Haar Cascades、DNN 模型），並提供影像縮放、灰階轉換、直方圖均衡化等功能 (Bradski, 2000)。在本研究中，OpenCV 負責對影像進行預處理，確保輸入 DeepFace 的人臉影像具有穩定的大小與清晰度，進而提升模型的辨識與分類精度。

綜上所述，深度學習為人臉分析任務提供了高效且自動化的解決方案，而透過結合 DeepFace 與 OpenCV，本研究能在影像資料中自動完成臉部檢測、特徵提取與年齡分類，並將資料以及模型導入建構完成的美食系統呈現最終型態。

2.3 卷積神經網路（Convolutional Neural Network, CNN）

卷積神經網路（Convolutional Neural Network, CNN）是一種具有層級結構的深度學習模型，特別適用於處理具有空間結構特性的資料類型，例如影像與語音等。CNN 模型的設計靈感源自生物視覺皮質的機制，其透過模仿人類視覺對局部區域的感知方式，自動學習影像中的空間層次特徵，並有效捕捉物體的形狀、邊緣、紋理等資訊。

CNN 的核心組成包括卷積層（Convolutional Layer）、池化層（Pooling Layer）以及全連接層（Fully Connected Layer）等。其中，卷積層負責以不同的卷積核（Filter）對輸入資料進行局部特徵的提取；池化層則透過最大池化（Max Pooling）或平均池化（Average Pooling）等方式進行資料降維，以減少參數量並提升模型的容錯能力；而全連接層則擔任分類或迴歸等任務的決策角色。這種由淺入深的結構設計，使得 CNN 能夠自動學習從低階特徵（如邊緣、角點）到高階語義特徵（如人臉、表情）之間的層級關係，無需人工設計複雜的特徵工程。

自 2012 年 AlexNet (Goodfellow, 2017) 於 ImageNet 大型視覺辨識競賽中大放異彩以來，CNN 在電腦視覺領域的應用快速拓展。AlexNet 引入深層卷積結構與 GPU 平行運算，成功突破以往模型在大規模資料處理上的限制，顯著提高了影像分類的準確率。其後，VGGNet (Karen Simonyan, Andrew Zisserman, 2014) 進一步優化網路架構，採用多個 3×3 的小卷積核堆疊以增加深度，同時保持結構簡潔一致，展現了模型設計模組化的潛力。ResNet (Kaiming He, et al., 2016) 則引入殘差學習（Residual Learning）機制，以跳躍連接（Skip Connection）的方式緩解深層網路的梯度消失與退化問題，使得網路深度得以擴展至上百甚至上千層，而不犧牲模型的訓練穩定性與性能表現。

隨著 CNN 架構的日益成熟，其應用也逐漸拓展至情緒辨識領域，特別是在視覺情緒分析方面。情緒辨識旨在判斷個體於特定時間內所表現的情緒狀態，對於人機互動、心理健康評估、智慧監控等領域皆具有高度實用價值。CNN 在此類任務中的優勢在於其能自動從人臉影像中學習區域性與全域性的情緒表徵。例如，眉毛上揚、嘴角下垂、眼角皺縮等細微的面部肌肉變化，均可透過多層卷積學習進行有效捕捉，進而提升模型在複雜環境下的辨識準確度。

相較於傳統方法（如主成分分析 PCA、局部二值模式 LBP、方向梯度直方圖 HOG 等）所需的人工特徵設計流程，CNN 所帶來的端對端學習能力，不僅減少了人為干預的需求，亦能在不同的資料集與應用情境中展現出更強的泛化能力。許多研究指出，CNN 模型

在各大情緒影像資料集（如 FER2013、CK+、AffectNet 等）上皆能達到高水準的辨識表現 (Ali Mollahosseini,et al., 2017)。

除了基礎的 CNN 架構外，近期研究亦嘗試將其與其他深度學習模型結合，以提升情緒辨識的整體效能。其中，遷移學習（Transfer Learning）是常見的策略之一。由於情緒資料普遍受限於標註成本高與樣本數不足的問題，遷移學習可利用在大規模通用資料集（如 ImageNet）上預訓練的模型參數作為初始權重，再透過少量目標任務資料進行微調（Fine-tuning），從而加快模型收斂速度並提升辨識準確率 (Jason Yosinski,et al., 2014)。

此外，針對影片或即時影像中的動態情緒分析，研究者也提出將 CNN 與循環神經網路（Recurrent Neural Network, RNN）結合的多模組架構。CNN 用於擷取每一幀影像的空間特徵，而 RNN（特別是長短期記憶網路 LSTM）則能處理序列資料中的時間依賴性，有效捕捉情緒變化的動態趨勢。這類混合模型已廣泛應用於影像對話系統、虛擬助理、教學陪伴機器人等多種情境中，展現出良好的應用潛力 (Zhanpeng Zhang,et al., 2014)。

綜合以上，卷積神經網路在電腦視覺與情緒辨識領域展現出高度的表現與靈活性，無論是在靜態影像的情緒分類，抑或動態影片中的情緒軌跡分析，皆能透過結構設計與技術整合持續突破現有瓶頸。隨著硬體資源與資料集規模的持續擴展，CNN 預期將在未來情緒計算與多模態感知領域中，扮演核心與關鍵的角色。上述研究成果不僅為本研究提供堅實的理論基礎，也驗證了深度學習方法於情緒辨識應用的可行性與前瞻性。

2.4 DeepFace

DeepFace 是由 Meta AI 團隊(Facebook AI research group)於 2014 年所提出的人臉辨識系統，亦是首個在人臉辨識任務中達到接近人類準確率（97.35%）的深度學習架構 (Yaniv Taigman,et al., 2014)。其核心概念在於以模組化方式整合臉部偵測、對齊、特徵抽取、特徵比對與情緒分類等流程，並以開放式深度學習架構（如 TensorFlow/Keras）作為後端支援。這樣的設計降低了開發門檻，

同時兼顧彈性與可擴充性，使得開發者能以相對簡單的方式完成複雜的人臉辨識任務。

在應用層面上，DeepFace 支援多種臉部偵測模組（如 OpenCV、SSD、Dlib、MTCNN 與 RetinaFace），使用者可依需求在速度與準確率之間進行權衡。例如即時互動或串流應用通常偏好速度較快的模組，而金融與門禁安全等高準確度場域則多採用精度較高的偵測方法。

此外，DeepFace 亦整合了多種預訓練卷積神經網路（CNN）架構，如 VGG-Face、FaceNet、ArcFace、SFace 與 MobileFaceNet，以支援不同的效能與資源配置需求。這些設計使其在人臉辨識與情緒分析領域提供了一個低門檻且高擴展性的實務解決方案。

然而，DeepFace 的訓練資料多半來自光線良好、正臉構圖的自拍影像，與真實世界場域（如逆光、低解析度、部分遮擋等情境）存在顯著差異。當模型遷移至這些場域時，準確率常出現明顯下降，此現象被稱為 domain gap。因此，若要將 DeepFace 應用於更廣泛的實務場景，仍需透過資料擴增、自建資料集或其他方法來縮小此落差。整體而言，DeepFace 的提出不僅代表深度學習在人臉辨識上的重大突破，也為後續多模態人機互動、醫療輔助與安防應用提供了重要基礎。

2.5 OpenCV

OpenCV（Open Source Computer Vision Library）是一個開源且跨平台的電腦視覺與機器學習函式庫，由 Intel 於 1999 年發起，目前由 OpenCV.org 維護與持續更新。根據官方說明，OpenCV 旨在提供一個統一的電腦視覺基礎架構，支援多種程式語言（如 C++、Python、Java）及多平台環境（包括 Windows、Linux、macOS、Android 與 iOS），以促進即時影像處理與人工智慧應用的發展。（OpenCV.(n.d.), 無日期）

隨著深度學習技術的興起，OpenCV 亦整合了 DNN（Deep Neural Network）模組，允許使用者載入與執行，如：Caffe、

TensorFlow、ONNX 等格式的訓練模型，進一步拓展其在影像辨識與分析領域的應用範圍與深度。OpenCV 提供豐富的圖像處理功能，包括顏色空間轉換，如：RGB、HSV、YCrCb、濾波、Gaussian、Median、Bilateral，直方圖處理、影像增強（如亮度與對比調整）、邊緣檢測等。這些為進階電腦視覺任務提供了堅實基礎。進一步地，OpenCV 支援多種特徵提取與比對技術，如：SIFT（Scale-Invariant Feature Transform）、SURF（Speeded-Up Robust Features）、ORB（Oriented FAST and Rotated BRIEF）等，廣泛應用於圖像比對、拼接與目標辨識等任務。此外，OpenCV 也內建多種機器學習演算法，如：KNN、SVM、決策樹等。使其能執行分類、回歸與預測等工作。

在物體偵測與辨識方面，OpenCV 提供多種預訓練模型與演算法，如基於 Haar cascade 的人臉偵測演算法，其以良好的準確度與執行效率至今仍被廣泛應用。而在影片分析方面，OpenCV 支援運動偵測、物體追蹤，如：KCF、CSRT、MOSSE、背景建模、光流估計與影片穩定等功能，適用於監控、行為分析等場景。其他功能如仿射與透視變換、圖像旋轉、縮放與剪裁等，也構成圖像前處理流程的重要組成。

與其他的人臉偵測後端相比。如：SSD、MTCNN，OpenCV 的主要優勢在於其運行效率高、資源需求低且易於整合。雖然在偵測準確率方面可能略遜於深度學習型模型，但其輕量化特性使其在本研究中的即時情境下展現高度實用性。此外，OpenCV 與其他深度學習架構，如：TensorFlow、PyTorch。之間亦具有良好的整合能力，使得在 DeepFace 架構中的前處理與後處理流程得以順利執行。整體而言，其不僅擔任人臉偵測與圖像預處理的角色，也作為整體辨識流程的起點，協助系統在影像中快速、準確地進行情緒、性別與年齡等多維度的分析，為後續智慧應用，奠定堅實基礎。

此外，OpenCV 能夠迅速將影像從 RGB 格式轉換為模型所需的灰階或 BGR 格式，以符合各類深度學習模型（如 DeepFace）輸入格式要求。總結來說，研究透過 OpenCV 的 Haar Cascade 分類器偵測影像中的人臉位置，再將偵測到的區域作為 DeepFace 進行情緒分析的輸入。此外，OpenCV 也被用來進行影像色彩空間的轉換（如 RGB 與 BGR、灰階），以配合不同模型對輸入格式的需求。由於

OpenCV 運行效率高、資源占用低，特別適合本研究辨識時的即時反饋應用情境，且能與其他深度學習模型高效整合。

2.6 Tensorflow

隨著人工智慧技術的迅速發展，電腦視覺與情緒辨識已成為當代智慧系統中的重要應用方向。人臉辨識技術作為電腦視覺中的一個核心分支，在智慧監控、人機互動、身份驗證等領域發揮關鍵作用。而人臉情緒辨識則結合了人臉特徵分析與深度學習模型，用以解讀影像中的情緒狀態，廣泛應用於教育科技、情緒輔療、行銷分析等場域。(Martín Abadi,et al., 2016)

TensorFlow 由 Google Brain 團隊於 2015 年開源的第二代深度學習系統，其前身為 DistBelief 系統。為因應當時日益增長的神經網路計算需求，Google 開發出 TensorFlow，致力於提供一個靈活、可擴展、跨平台的深度學習解決方案。其主要採用資料流圖（Data Flow Graph）作為其計算基礎，透過節點（Node）定義操作，邊（Edge）定義資料（張量）傳遞的關係。此架構提供強大的視覺化能力與可擴展性，特別適合需要大量並行計算與分散式訓練的深度神經網路。自 TensorFlow 2.0 起，官方將 Keras 完整整合為主要的高階 API，讓使用者能以更簡潔、模組化的方式建立神經網路。這對於進行影像分類、人臉辨識或情緒預測任務的研究者而言，提供了高度便利性。此外，TensorFlow 支援 GPU 加速與多平台部署，無論是使用 NVIDIA CUDA 進行本地加速訓練，或透過 TensorFlow Lite 將模型部署至行動裝置，皆能因應各類型應用需求，尤其適合即時辨識系統之開發。(GoogleBrain, 無日期)

在人臉辨識任務中，TensorFlow 常與 OpenCV 搭配使用，進行臉部偵測與特徵提取。研究者藉由使用 TensorFlow 實現卷積神經網路（Convolutional Neural Networks, CNN）架構來分類不同的臉部特徵，進行身份比對，或透過遷移學習（Transfer Learning）使用。如：FaceNet、VGGFace 等預訓練模型。舉例來說，FaceNet 模型利用三重損失函數（Triplet Loss）來學習人臉嵌入（Embedding），使得同一個人的人臉特徵向量彼此接近，而不同人的人臉特徵向量彼此遠

離，從而達到高準確率的人臉辨識效果。(Martin Abadi,et al., 2016)

本研究發現，情緒辨識通常使用 CNN 或 CNN-RNN 混合結構，輸入為人臉圖像，輸出為情緒分類，例如：快樂、悲傷、憤怒、驚訝、中立、噁心、恐懼。TensorFlow 提供多種方式來建構這樣的網路模型，結合遞迴神經網路（Recurrent Neural Networks, RNN）或長短期記憶網路（Long Short-Term Memory, LSTM）可以捕捉時間序列中的情緒變化，提升連續情緒辨識的準確率。

雖然 Tensorflow 可使用多種前端程式語言，但對 Python 的支援是最佳的，且 Python 具有程式碼簡明、易學習、高生產力的特質，物件導向、函數式的動態語言，應用非常廣泛 (GoogleBrain, 無日期)。故本研究使用 Python 執行各項任務已達成美食推薦系統模型。

2.7 InsightFace

InsightFace 為一套開源人臉辨識系統，其架構建置於 MXNet 與 PyTorch 等深度學習框架之上，廣泛應用於人臉辨識、人臉驗證及人臉屬性預測等領域（如年齡辨識）。其中，InsightFace 的 buffalo_l 模型以多族群人臉資料集（如 MS1MV3、Glint360K）訓練而成，其中包含部分亞洲臉孔樣本，因此在亞洲族群的臉部特徵辨識與年齡預測上仍具有良好準確度，能更精確預測亞洲面孔之年齡區間，進而有效提升整體辨識準確率，特別適用於台灣等華人為主之應用情境 (Jiankang Deng,et al., 2022)。

InsightFace 系統之技術核心為 ArcFace 損失函數（ArcFace loss），該方法透過在特徵嵌入空間中引入角度邊界限制（angular margin），增強類別間的可區辨性，同時保持同一類別內特徵的一致性。此技術在多項國際人臉辨識標竿測試中皆展現優異表現，具備世界級水準。在實作方面，InsightFace 模型通常將輸入影像標準化，並透過多層卷積神經網路萃取臉部深層特徵，輸出 512 維向量以進行後續之分類或回歸任務。在年齡預測應用中，InsightFace 提供已完成預訓練之回歸模型，可直接輸出估計年齡數值。

2.8 年齡差距正確錯誤依據

本研究的年齡依據採取估計值與實際年齡誤差在 ± 5 歲之內視為正常範圍的標準，表示當圖片進行辨識時，只要與實際辨識的年齡相差 ± 5 歲之內，都能算是正確的執行。此誤差設定並非主觀憑據，而是根據多篇學術文獻與國際評估報告所提出之實證結果，為目前人臉年齡預測領域中最常見且具可行性的誤差容忍標準。

首先，Han、Otto 與 Jain (2013) 於第六屆國際生物特徵辨識會議 (ICB) 發表之論文《Age Estimation from Face Images: Human vs. Machine Performance》中，採用「CS (Cumulative Score within 5 years)」作為主要評估指標，亦即預測年齡落在實際年齡 ± 5 歲之內即視為正確。該研究針對 FG-NET、MORPH II 與 PCSO 三個資料集進行測試，CS 命中率分別為 74.8%、72.4%、與 64.0%，皆顯示 ± 5 歲為合理且可接受的模型誤差範圍。此外，研究也顯示人類在 FG-NET 上的平均絕對誤差 (MAE) 為 4.7 歲，顯示人為估算亦多落在 ± 5 歲之內 (Hu Han, et al., 2013)

再者，美國國家標準與技術研究院 (NIST) 於 2014 年與 2024 年各發表一份人臉年齡辨識技術評估報告。2014 年的《Automated Facial Age Estimation Report》指出，最精確的模型達成 4.3 歲的 MAE，並有 67% 的預測結果落在 ± 5 歲誤差範圍之內，顯示 ± 5 歲為業界採用的主要評估基準之一 (S. E. Bekhouche, et al., 2024)。2024 年的《FATE AEV Report》進一步強調，超過三分之二的系統預測結果均在 ± 5 歲之內，並明確說明其核心績效指標即為落在 ± 5 歲內的預測比例 (S. E. Bekhouche, et al., 2024)。該報告同時指出人臉老化表徵易受表情、族群、環境等多重因素影響，因此 ± 5 歲作為容忍範圍更具實際意義。

此外，Bekhouche 等人於 2024 年發表於 Electronics 的研究中，針對 MORPH2、CACD 與 AFAD 三個公開資料集進行多階段深度神經網路 (MSDNN) 之測試，結果顯示在 ± 5 歲範圍內的命中率分別為 86.66%、65.31% 與 80.17%，進一步驗證該誤差範圍具跨資料集的穩定性與普遍性 (S. E. Bekhouche, et al., 2024)。

綜合來看，從人類的辨識能力、學術研究結果，到國際機構的

實測報告，都把 ± 5 歲當作年齡預測是否正確的重要標準。這也證明這個範圍在技術上是可行的，在實際應用上也為合理。因此本研究也採用 ± 5 歲作為判斷是否辨識正確的依據，當系統預測出年齡後，只要和實際年齡的差距在 ± 5 歲以內，就會被視為正確結果。

2.9 Gradio

Gradio 為開源的 Python 函式庫，最初由 Abubakar Abid 等人於 2019 年開發，旨在降低機器學習模型的展示門檻，使開發者能夠快速建立圖形化的使用者介面（Graphical User Interface, GUI），並與模型進行即時互動 (Gradio, 無日期)。Gradio 對於研究人員與開發者而言，提供另一種更輕便、直觀的方式來展示模型成果，特別適合需要視覺化回饋與即時操作的應用情境。如：影像分類、語音辨識與自然語言處理等任務。(Abubakar Abid, et al., 2019)

再者，Gradio 最大的特色在於其極簡的介面設計與高度整合性。開發者僅需透過幾行程式碼，便可將一個函式包裝成可互動的網頁介面。其支援多種輸入與輸出類型，包括文字（Text）、圖片（Image）、音訊（Audio）、影片（Video）與滑桿（Slider）等，使其特別適合處理多媒體資料。相較於傳統的 Web 架構如 Flask 或 Django，Gradio 不僅降低了開發門檻，更將部署與展示時間大幅縮短，使機器學習原型設計（prototyping）變得更為迅速與視覺化。(Abid, 2021)

此外，Gradio 也提供與 Hugging Face 平台的高度整合，開發者可直接從 Hugging Face Model Hub 匯入現有模型，並一鍵部署於 Hugging Face Spaces 上。這項特性使得模型分享更加便利，並加速了人工智慧模型的開源與社群交流。在學術研究中，透過 Gradio 建立互動式展示，有助於提升論文的可重現性與可操作性，亦促進非技術背景的使用者能夠更直觀地理解模型行為。

在現今快速迭代的 AI 應用開發流程中，Gradio 的角色不再只是展示工具，更可作為前後端整合的橋樑。其支援即時上傳圖片、即時回傳模型預測結果，結合資料視覺化與資料查詢模組，建構出具備完整使用者互動體驗的應用系統。根據 Gradio 官方統計資料，截

至 2024 年底，已有超過數十萬個互動介面透過 Hugging Face Spaces 被建立並公開上架，可見其在實務與學術間的廣泛應用。(GradioTeam, 無日期)

本研究所開發之「情緒辨識與台東美食推薦系統」即採用 Gradio 作為前端互動架構。主要考量為其具備以下優勢：一、部署簡便：可直接搭配 Hugging Face Spaces 無需自行架設伺服器，適合學生專題或小型專案；二、介面親和：使用者可透過拖曳上傳圖片，快速檢視辨識結果與推薦內容；三、程式碼整合度高：可與 DeepFace、Pandas 等常見套件共同使用，降低整合成本；四、跨平台存取：使用者無需安裝額外軟體，只需透過網頁瀏覽器即可操作，提升系統可近性與使用彈性。

綜合上述，Gradio 不僅具備技術深度的開發工具，更推動 AI 模型開放與應用民主化的重要媒介。使其在教育、研究與開發三者之間取得平衡。對於本研究而言，Gradio 提供了理想的平台，讓情緒辨識系統能夠以視覺化、互動性的方式呈現，不僅強化使用者體驗，也提升系統的展示與推廣價值。

第三章、研究方法

3.1 實驗流程

本研究系統主要由五個核心模組組成，依序為：資料蒐集與儲存、圖像預處理與人臉偵測、情緒與屬性辨識、美食推薦建構以及使用者互動與輸入模組。整體研究流程，如圖 2 所示。

首先，透過雲端平台（如 Google Drive）蒐集並儲存包含不同情緒標註的人臉影像資料，作為模型辨識與系統測試所需的輸入來源。接著，系統使用 OpenCV 進行圖像的預處理，包含格式統一與尺寸標準化，並透過人臉偵測技術（本系統採用 OpenCV Haar Cascade 及 InsightFace 模型）自圖像中擷取出人臉區域。

在本研究中，OpenCV 被用來進行人臉偵測與圖像前處理，包含色彩空間轉換、臉部區域擷取與灰階化等操作。首先，利用 OpenCV 所內建的 Haar Cascade 分類器進行人臉偵測，其核心為級聯式機器學習模型，藉由大量的正負樣本學習出 Haar-like 特徵後，再透過 AdaBoost 提升分類強度，被裁切出的人臉區域（ROI）將作為 DeepFace 模組的輸入，進行情緒辨識。

在人臉擷取完成後，進一步使用 DeepFace 套件進行情緒分類，同時結合 InsightFace 模型擷取年齡與性別屬性。本系統以情緒辨識結果為主，搭配使用者預測年齡作為輔助資訊。最後，系統根據辨識結果查詢預先建立的美食推薦資料庫，並回傳與情緒相符的餐點資訊。推薦邏輯以規則式比對為主，並整合食物名稱、推薦理由與外部連結等欄位，提供使用者具體且具參考價值的美食建議。

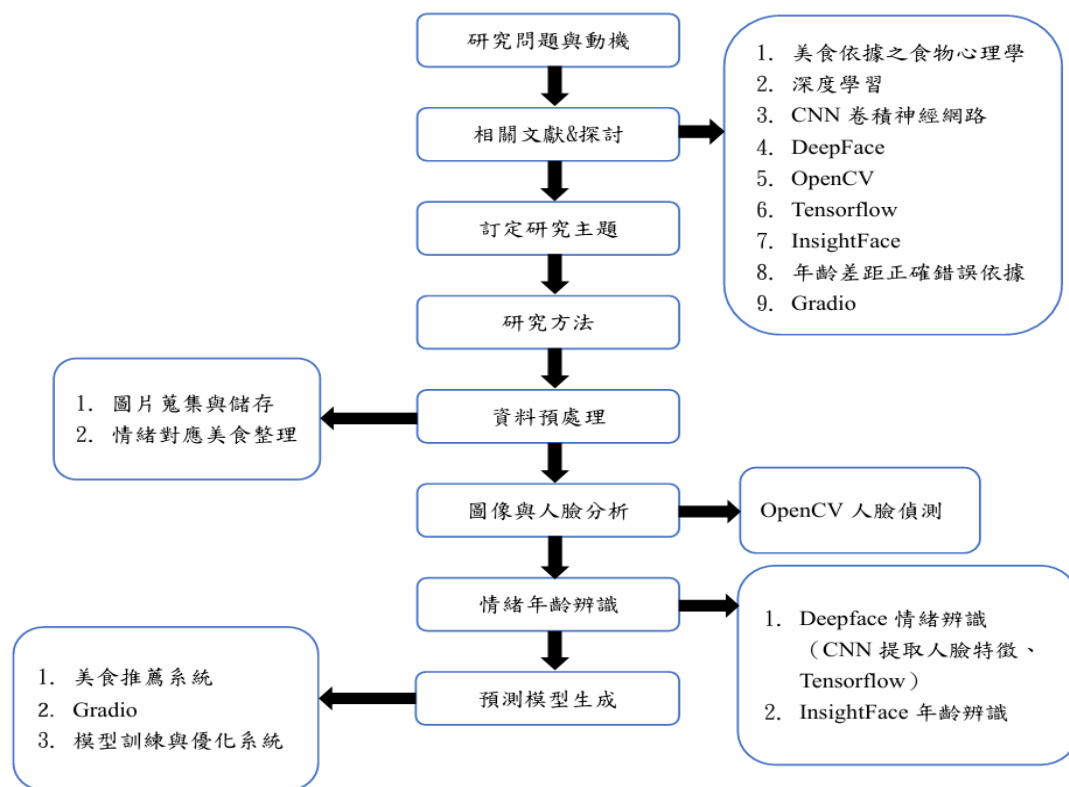


圖 1 流程架構圖

3.2 系統架構

本研究所開發之「情緒導向美食推薦系統」，整體架構如圖 3 所示。系統由五大模組構成，涵蓋影像接收與前處理、人臉偵測與辨識、情緒與年齡預測、推薦邏輯引擎、以及前端互動展示模組。以下為各模組說明：

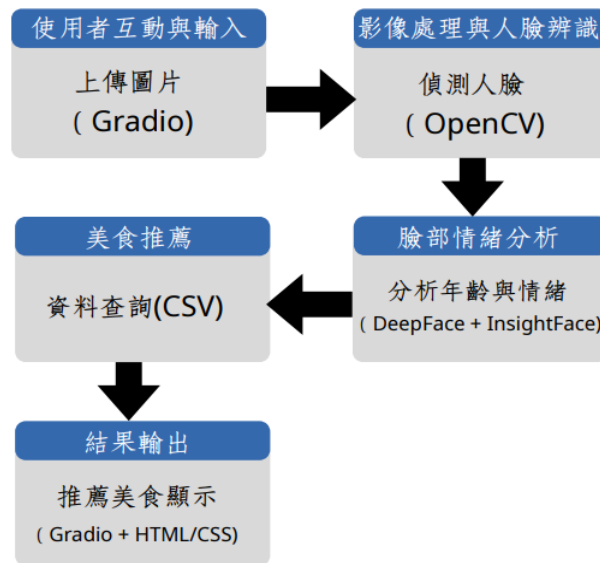


圖 2 美食推薦系統架構

一、使用者互動與輸入模組（Gradio）

- 提供圖形化介面，讓使用者能夠上傳包含人臉的靜態圖片。
- 支援拖曳或點擊上傳，並即時將圖片傳送至後端分析模組。
- 前端介面同時提供辨識結果與推薦資訊，具備即時意見反應功能。

二、影像處理與人臉偵測模組（OpenCV）

- 將圖片轉換為 OpenCV 可辨識格式（BGR、灰階等）。
- 使用 Haar Cascade 技術進行人臉定位與裁切，取得 ROI 區域。
- 執行圖像標準化作業，使輸入影像統一尺寸與格式，利於模型辨識。

三、臉部屬性分析模組（DeepFace + InsightFace）

- DeepFace 負責辨識七種基本情緒（快樂、悲傷、驚訝、憤怒、噁心、恐懼、中立）。

- InsightFace 用於推估使用者年齡，並進行 ± 5 歲容差範圍內的結果修正。
- 此模組為整體推薦機制之核心，決定最終推薦條件。

四、美食推薦引擎模組（CSV 資料查詢）

- 根據辨識出的情緒與年齡結果，查詢建立好的美食推薦資料庫（Excel / CSV 格式）。
- 資料庫中依據年齡層與情緒狀態分類各種適合的在地美食。
- 每筆推薦包含：美食名稱、推薦原因、地點資訊與網站連結。

五、結果輸出與視覺化模組（Gradio + HTML/CSS）

- 將辨識結果與推薦資訊整合，回傳至前端網頁展示。
- 使用自訂化的 HTML/CSS 設計美食卡片、按鈕、超連結等資訊模組。
- 提供清楚的圖文並茂推薦資訊，強化互動與使用者體驗。

3.3 年齡系統與模型設定

本研究以 InsightFace 進行影像年齡推估，採用 InsightFace 中的 buffalo_l 模型，並配合 OpenCV 模組進行人臉偵測。於實驗設計上，針對年齡辨識之準確度，本研究容許 ± 5 歲範圍內之預測誤差視為正確預測，並以此標準進行正確率統計分析。InsightFace 預設的 FaceAnalysis 類別會回傳屬性 age，其值為一個連續型預測年齡。根據實測，該模型雖表現穩定，但仍可能會出現年齡偏高現象（over-estimation）。為了修正這類偏差，本研究採用一組簡化線性修正公式，將原始預測年齡 raw_age 進行校正後再使用。本研究採用 Python 實作並使用官方類別完成模型初始化與推論。程式實作如下

(程式碼 1)：

程式碼 1. InsightFace 年齡推估

```
import insightface # 模型初始化與推論

insight_model = insightface.app.FaceAnalysis(name='buffalo_l',
providers=['CPUExecutionProvider'])

insight_model.prepare(ctx_id=0)

faces = insight_model.get(image) # image 為輸入影像

raw_age = faces[0].age # 取得模型預測年齡
```

3.3.1 年齡校正策略（線性回歸）

為了修正模型於特定年齡層的高估偏差，本研究根據實際測試樣本，使用線性回歸擬合校正參數，本研究以線性回歸 $Y = aX + b$ 擬合出校正參數，推導出以下修正公式，如公式(1)：

$$Y = aX + b \quad (1)$$

另外希望提升模型預測結果之準確性，本研究根據實際測試樣本設計一組修正公式，以兩組樣本為依據：

- (1) 實際 20 歲，模型預測 24 歲。
- (2) 實際 60 歲，模型預測 69 歲。

由此可得斜率

$$a = \frac{60-20}{69-24} = \frac{8}{9} \quad (1)$$

並代回求得截距

$$b = 20 - \frac{8}{9} \times 24 = -\frac{4}{3} \quad (1)$$

因此，本研究採用之校正公式為(2)

$$adjusted_age = round\left(\frac{8}{9} \cdot raw_age - \frac{4}{3}\right) \quad (2)$$

線性回歸的應用不僅具備補正能力，其公式簡單易懂，亦為後處理階段帶來高度可解釋性。最終年齡輸出值經四捨五入後，分類為不同族群（如：「60 歲以上」為銀髮族、「60 歲以下」為青年族），並輸入推薦系統作為食物推薦條件之一。依據相關文獻與經驗判斷，若模型預測與實際年齡相差在 ± 5 歲以內，則視為有效預測範圍（Cole et al., 2017；Tang et al., 2021）。

3.4 研究流程

本研究的整體流程自影像輸入開始，依序經過人臉偵測、情緒辨識、年齡推估與校正，最終完成美食推薦並由前端介面輸出。圖 4 為研究方法流程圖。首先，系統利用 OpenCV 將輸入影像進行格式與尺寸的標準化，再透過 Haar Cascade 與 InsightFace 模組定位並裁切臉部區域，確保後續辨識不受背景干擾。完成臉部擷取後，系統將影像輸入 DeepFace 進行情緒分類，模型輸出各情緒的分數，並以 Softmax 函數，如公式(3)。

$$Softmax(z_i) = \frac{e^{z_i}}{\sum_{j=1}^k e^{z_j}} \quad (3)$$

將其轉換為機率分布，進而確定最有可能的情緒類別。另一方面，年齡則由 InsightFace 模型進行推估，考量到模型在部分年齡段容易產生高估偏差，本研究依據測試樣本建立線性回歸校正公式，將預測年齡進一步修正，再依結果區分為「60 歲以下」與「60 歲以上」兩大群體。若需比對不同樣本間的相似程度，則會額外使用歐式距離公式，如公式(4)。

$$d(x_1, x_2) = \sqrt{\sum_{i=1}^n (x_{1i} - x_{2i})^2} \quad (4)$$

來衡量臉部特徵向量間的差異。最終，情緒與年齡的辨識結果會傳回至 Gradio 介面，系統即時呈現分析結果，並從資料庫中查詢符合條件的在地美食，回饋至前端給使用者參考。整體流程不僅驗證了 DeepFace 在情緒分類上的實用性，也展現 InsightFace 在年齡推估上的價值，搭配前端互動與推薦機制，形成一個可即時應用的情緒導向美食推薦系統。

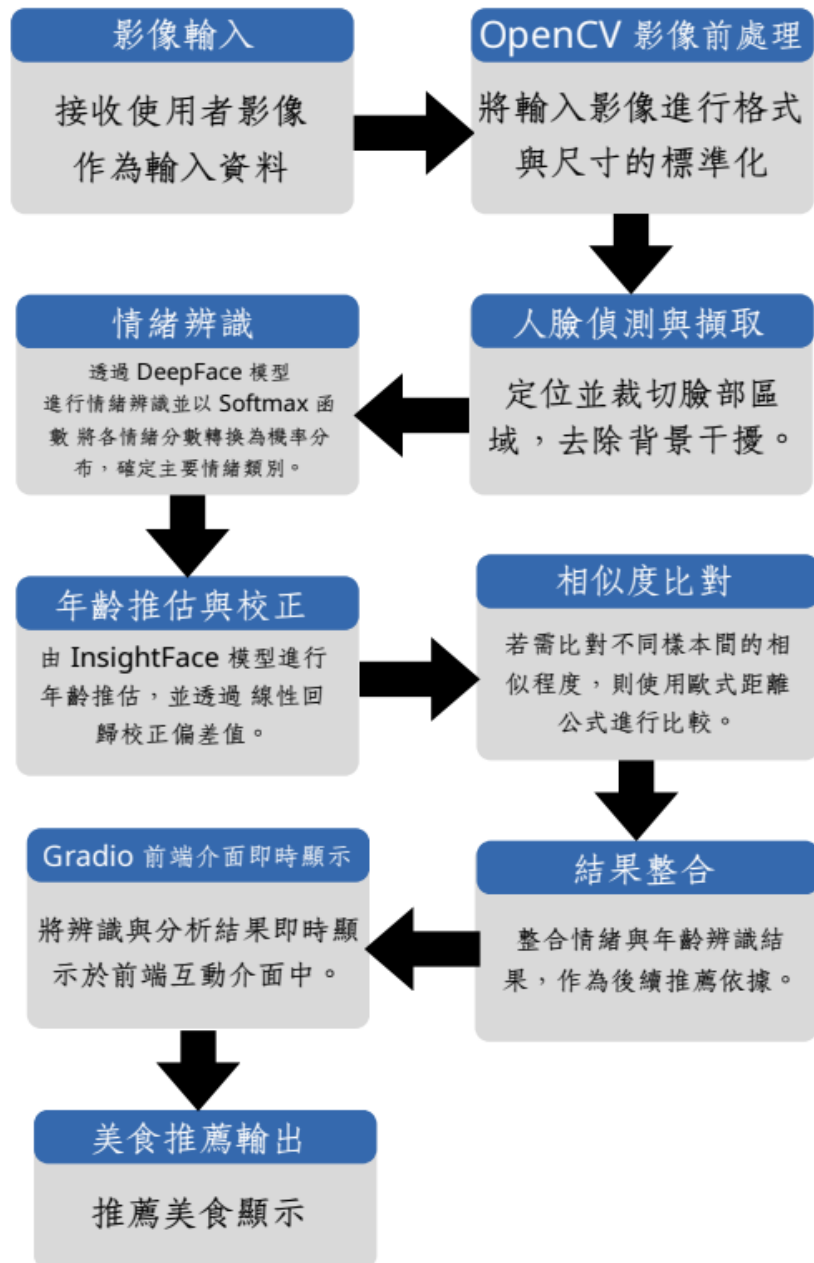


圖 3 研究方法流程圖

第四章、實驗結果與分析

本研究將針對研究方法的實際應用進行說明，並呈現各階段的實驗過程與成果。首先，將介紹資料蒐集、標註與整理的方式，以確保後續模型訓練所需的資料品質。接著，說明影像前處理與人臉偵測的技術流程，藉此提升辨識的準確性與穩定度。在模型構建部分，本研究將展示情緒與屬性辨識模型的設計與參數調整，並討論準確率提升的方式與成效。同時，也會探討年齡推估演算法的應用與結果分析，以驗證模型在不同任務下的表現。最後，本章將介紹本研究所建置的美食推薦系統，並展示其在 Hugging Face Spaces 上的系統部署與應用介面設計。透過上述實驗與分析，本研究旨在驗證研究方法的可行性，並展現於實際應用上的潛力。

4.1 資料蒐集與儲存

4.1.1 蒐集、標註與儲存

資料來源主要來自三個途徑：一、公開可用的人臉表情資料集（如 FER2013、CK+ 等），二、網路搜尋（包含具授權的開源影像資源），以及三、透過使用者自行上傳或拍攝的影像資料。在蒐集過程中，嚴格篩選具清晰臉部特徵與明顯情緒表現的圖片，並確認圖片格式以 JPG、PNG 為主，確保能夠被模型正常讀取與處理。

在雲端硬碟中建立資料庫，用來儲存所有人臉表現情緒的圖片。先針對每種人臉情緒標籤獨立出一個資料夾(生氣、害怕、噁心、開心、難過、驚訝、正常)，再將蒐集到的資料集(檔案類型為 jpg、png)依據情緒做分類，放置於資料夾儲放。

4.1.2 資料整理

將所收集的資料集（人臉圖片）依照情緒標籤做分類之後（生氣、害怕、噁心、開心、難過、驚訝、正常），再經由人工篩選，篩選掉情緒呈現不明顯、非正面角度、輪廓不顯著或不具代表性的

圖片，此舉動目的為提升資料品質，避免出現結果不準確的情況。圖片皆參考網路上公開相片。圖 4 為人臉示意圖（舉例情緒為開心）。左圖為正確參考：圖片解析度清楚，且人臉角度為正面；右圖則為錯誤依據(原資料集圖片將此標籤設定為生氣，但肉眼看相片則偏向開心，故不符合)



圖 4 人臉採取示意圖

此外，在整理過程中也注意每個情緒分類下圖片數量的平衡性。發現某些情緒類別（如「噁心」）資料較少，導致模型在訓練時容易出現嚴重的類別不平衡問題，因此選擇移除這些類別，確保模型能在主要情緒類別上達到較佳的辨識準確率與穩定性。

4.2 圖像與人臉處理

4.2.1 影像前處理與標準化

在本研究系統中，影像處理流程的第一步為影像前處理與格式標準化，此步驟是後續人臉偵測與情緒分析準確性的基礎。使用者上傳的圖片或透過攝影機取得的即時畫面。

首先，本研究以 Gradio 介面接收為 Pillow 格式（PIL.Image 物件），再透過 NumPy 轉換為 OpenCV 可處理的陣列格式（NumPy array），並進一步進行色彩空間與資料型態的轉換，以符合後續模型處理需求。影像會從預設的 RGB 色彩空間轉換為 OpenCV 使用之 BGR 格式（cv2.COLOR_RGB2BGR），此轉換為深度學習模型提供

一致的輸入標準。隨後，為利於 OpenCV 執行 Haar Cascade 人臉偵測，圖像再被轉換為灰階格式（cv2.COLOR_BGR2GRAY）。此步驟可降低資料維度與雜訊干擾，提升偵測效率，其程式碼如下。

```
# 將圖片轉換為灰階，降低資料維度，方便後續人臉偵測
gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)

# 將圖片調整為固定大小 (400x400)，避免不同解析度影響偵測結果
img = resize_with_padding(img, (400, 400))
```

以上步驟處理完成後，系統將依據灰階圖像中偵測到的臉部區域，擷取相對應之 BGR 色彩範圍作為人臉 ROI（Region of Interest）。此區塊將提供給後續之 DeepFace 模組進行情緒辨識分析，確保模型輸入統一且具有辨識品質。整體前處理流程有效提升人臉辨識精度，也為年齡預測與食物推薦模組提供穩定輸入來源，其程式碼如下。

```
# 載入 Haar Cascade 分類器，用於人臉偵測
face_cascade = cv2.CascadeClassifier(cv2.data.haarcascades +
"haarcascade_frontalface_default.xml")

# 在灰階圖像中偵測人臉，設定縮放比例與最小臉部大小
faces_cv = face_cascade.detectMultiScale(gray, scaleFactor=1.05,
minNeighbors=4, minSize=(40, 40))
```

4.2.2 人臉偵測技術

完成影像前處理後，系統進入人臉偵測與情緒辨識階段。人臉偵測採用 OpenCV 提供的 Haar Cascade 分類器（Haar Cascade Classifier），透過 detectMultiScale() 函式在灰階圖像中偵測出可能的人臉區域。系統會選擇第一張偵測到的人臉進行後續處理，並擷取其位置座標進行裁切。

被裁切出的人臉區域（ROI）將作為 DeepFace 模組的輸入，進行情緒辨識。DeepFace 為一套開源的人臉分析工具，內部結合多種深度學習模型，能針對人臉圖像執行表情、年齡、性別等屬性推論。在本研究中，系統設定 DeepFace 以 actions=["emotion"] 模式分析情

緒，回傳結果中會標示出最具代表性的 dominant_emotion（如 happy、sad、angry 等），其程式碼如下。

```
# 使用 DeepFace 進行情緒辨識
emotion_results = DeepFace.analyze(
    face,                    # 輸入 ROI 區域 (人臉)
    actions=['emotion'],     # 僅進行情緒分析
    enforce_detection=False, # 若人臉特徵不足，仍允許進行分析

    detector_backend='opencv' # 指定 OpenCV 偵測器
)

# 取得主要情緒標籤 (如 happy, sad)
dominant_emotion = emotion_results[0]['dominant_emotion']

# 在圖像中標示辨識出的情緒 (轉換成中文顯示)
img = putText(img, x, y - 10, text_obj[dominant_emotion], size=25,
color=(255, 255, 255))
```

本研究的情緒辨識的結果不僅會在使用者介面中顯示出來，還會作為後續食物推薦模組的重要判斷依據。情緒辨識階段在整個系統的運作中扮演著關鍵角色，因為它的準確性將直接影響食物推薦的精確度。如果情緒辨識有誤，推薦的食物可能會與使用者的需求不符，進而降低整體系統的使用體驗。因此，前處理和臉部擷取的準確性在這個階段顯得尤為重要，因為它們直接決定了情緒辨識結果的可靠性，進而影響到後續推薦系統的運作效果。這不僅是技術層面的挑戰，更是提高使用者滿意度的關鍵。

4.3 情緒與屬性辨識模型

本研究於情緒辨識模組中採用 DeepFace 架構內建之 CNN 模型進行人臉特徵提取與表情分類。DeepFace 的情緒分析模型以 mini-Xception 為基礎，模型包含卷積層、池化層、全連接層，細節特徵如「嘴角的形狀」、「眼睛的角度」，並同時進行分類，幫助系統從整張圖片中分辨出人臉。特徵會被 CNN 的多層結構學習，最後進

行分類，輸出最可能的情緒結果。透過模型訓練後，系統能自動將輸入圖片中的人臉分類為七種基本情緒：生氣（Angry）、厭惡（Disgust）、害怕（Fear）、高興（Happy）、難過（Sad）、驚訝（Surprise）、與中性（Neutral）。情緒辨識的核心程式法如下。

```
# 使用 DeepFace 進行情緒分析
emotion_results = DeepFace.analyze(
    face,          # 輸入 ROI 區域 (人臉)
    actions=['emotion'], # 僅進行情緒分析
    enforce_detection=False, # 若人臉特徵不足，仍允許進行分
    detector_backend='opencv' # 指定 OpenCV 作為偵測後端
)

# 取得主要情緒標籤 (如 happy, sad)
dominant_emotion = emotion_results[0]['dominant_emotion']

# 將辨識結果顯示於圖片上 (轉換為中文顯示)
img = putText(img, x, y - 10, text_obj[dominant_emotion], size=25,
color=(255, 255, 255))
```

DeepFace 的 analyze() 方法內部結合多種 CNN 架構進行情緒分類，其 dominant_emotion 會輸出最佳預測結果，並以字典形式回傳七類情緒的機率分布。

4.3.1 參數調整

在實驗初期，本研究對影像尺寸與前處理方式進行了多次調整。例如將輸入圖片統一縮放至 400x400 像素，並保持原始寬高比，再以黑色邊框填補空白區域，以避免圖片變形。然而實驗顯示，影像縮放方式對整體準確率的提升效果有限，準確率約為 36.64%，且在多次測試中準確率的浮動僅在 2% 以內，以下為對應程式碼。

```
# 影像縮放並保持比例，填補黑邊，避免失真
```

```

def resize_with_padding(img, target_size):
    h, w = img.shape[:2]
    scale = min(target_size[0] / w, target_size[1] / h) # 計算縮放比例
    new_w, new_h = int(w * scale), int(h * scale) # 新的寬高
    resized_img = cv2.resize(img, (new_w, new_h)) # 縮放圖片
    result = np.zeros((target_size[1], target_size[0], 3), dtype=np.uint8) #
    建立黑色背景

    # 將縮放後的圖片置中
    result[(target_size[1] - new_h) // 2 : (target_size[1] - new_h) // 2 +
    new_h,
           (target_size[0] - new_w) // 2 : (target_size[0] - new_w) // 2 +
    new_w] = resized_img
    return result

```

此外，本研究也針對人臉偵測的參數進行了優化。以 OpenCV 的 `cv2.CascadeClassifier.detectMultiScale()` 為例，其中的 `scaleFactor` 參數決定影像金字塔每層的縮放比例，進而影響偵測的靈敏度與準確度。經過多次測試，本研究將 `scaleFactor` 由預設的“1.1”調整為“1.05”，結果顯示在模糊或人臉較小的圖片中能更有效地偵測人臉，但需注意偵測靈敏度提升可能導致誤檢率的增加，其程式碼如下。

```

# 調整後的人臉偵測參數
faces = face_cascade.detectMultiScale(
    gray, # 輸入灰階圖片
    scaleFactor=1.05, # 影像金字塔縮放比例 (較小值 → 偵測更靈敏)
    minNeighbors=4, # 偵測的最小鄰居數 (避免誤檢)
    minSize=(40, 40) # 最小人臉尺寸
)

```

4.3.2 準確率提升

本研究於辨識完人臉情緒後，將七種情緒，分別為生氣、噁心、害怕、開心、難過、驚訝、正常的各項準確率整理至以下分類。此研究步驟可持續修正，藉此提升模型優化的成效。結果如表格 2。

表格 2、各情緒對應準確率表格 2 各情緒對應準確率

情緒類別	樣本數	正確數	準確率 (%)
生氣 (Angry)	40	29	72.50
厭惡 (Disgust)	0	0	0.00
害怕 (Fear)	38	21	55.26
開心 (Happy)	101	93	92.08
難過 (Sad)	49	19	38.78
驚訝 (Surprise)	83	69	83.13
中性 (Neutral)	96	61	63.54

總正確率：73.00%（總樣本數：400）

結果分析：

一、高準確率情緒：

- 開心 (Happy) 準確率最高 (92.08%)，原因是「開心」表情在資料集中特徵明顯（嘴角上揚、眼部表情），且訓練樣本充足。
- 驚訝 (Surprise) 也達到 83.13%，因其特徵（眼睛張開、嘴巴大開）在圖像中辨識度高。

二、低準確率情緒：

- 難過 (Sad) 僅 38.78%，原因可能為表情變化細微，且與中性表情相似，導致模型混淆。
- 害怕 (Fear) 也偏低 (55.26%)，可能因資料集樣本不足，且與驚訝表情特徵有重疊。
- 厭惡 (Disgust) 無樣本 (0%)，表示資料集未涵蓋此情緒，無法評估。

三、整體觀察：

- 模型對於特徵明顯的情緒（開心、驚訝）表現良好，但對於特徵細微或樣本不足的情緒（難過、害怕）辨識效果較差。
- 總體準確率 73%，顯示系統具備一定辨識能力，但仍有改進空間。

4.4 美食推薦系統

系統利用 InsightFace 模型預測出的年齡資料，將使用者劃分為「年齡低於 60 歲」與「年齡高於 60 歲」兩個主要族群。此分群策略源自於不同年齡階段在飲食偏好、生理需求與健康考量上的差異，亦有助於使推薦內容更符合使用者的生活狀況與健康管理目標。

在資料庫的設計上，本研究將食物資料依據情緒類別及年齡層別（<60 歲、≥60 歲）進行分類與標註，使系統能根據使用者當下的情緒狀態與年齡分組，精確篩選出適合其需求的食物選項。這樣的年齡與情緒雙重條件篩選，強化了推薦的針對性與實用性。

4.4.1 情緒對應美食資料庫建立

本研究依據彙整結果，歸納出七種常見情緒狀態，分別為快樂、悲傷、生氣、驚訝、厭惡、恐懼與中立，並針對每一情緒推薦適合之在地台東小吃美食，期望以飲食作為情緒調節的輔助工具。為達成系統推薦功能之實作目標，本研究將情緒與對應之美食資訊整理為結構化資料，並輸入至 Excel 表單中，並轉呈 CSV 檔案。以利後續程式碼提取關鍵詞進行輸出。

4.4.2 基於情緒與年齡推薦規則設計

本研究參考多項文獻與資料，作為情緒與飲食關聯分析之理論依據。主要資料來源包括：國立中興大學出版之《情緒大解密》一書、BBC 中文網報導〈飲食與健康：我們對安慰食品依賴背後的科學道理〉、以及相關情緒與飲食之探討文章，如〈情緒穩定是可以吃出來的！焦慮、煩躁、心情低落時對應的 5 類食物〉等。上述文獻提供了不同情緒狀態與飲食之間關聯的理論背景，並有助於本研究後續推薦系統之設計，如圖 5 所示。

emotion	age_group	food_name	reason
angry	young	初鹿鮮奶冰淇淋	生氣時吃初鹿鮮奶冰淇淋，可以透過冰涼滑順的口感與天然奶香帶來感官上的愉悅，
angry	old	寶桑豆花	生氣時吃寶桑豆花，可以透過其綿密細緻的口感與淡淡豆香帶來溫和的安撫感。豆花
angry	young	冰菓鋪	生氣時吃冰菓鋪，繽紛多樣的新鮮水果與剉冰的冰涼刺激，能迅速轉移注意力，喚醒
angry	young	哩哩冰果室	生氣時來台東哩哩冰果室，獨特創意的冰品與在地新鮮食材結合，能帶來驚喜與愉悅
angry	young	陽山茶道	生氣時喝台東陽山茶道的手搖飲，清新的茶香搭配微甜滑順的口感，能迅速帶來味覺
angry	old	糖廠冰棒	生氣時吃台東糖廠冰棒，濃濃復古風味與單純天然的甜味，能勾起童年回憶，帶來安
angry	young	鐘花-村長的家	生氣時走進鐘花-村長的家，品嚐酥香可頌與細緻甜點，能透過濃郁奶香與層層口感
angry	young	Ho Li_dessert舒芙蕾鬆餅專賣	生氣時吃台東 Ho Li_dessert 的舒芙蕾鬆餅，綿柔蓬鬆的口感宛如雲朵入口即化，能帶
angry	young	雨田家馬卡龍	生氣時吃台東雨田家的馬卡龍，繽紛的色彩與酥脆外殼、柔滑內餡的雙重口感，能瞬
disgust	old	蘇天助素食麵	老字號純素麵，溫潤湯湯搭配炭火慢熬的素肉燥，口感濃郁卻不膩，熱湯與乾麵交
disgust	old	粥大福雞湯粥舖	各式稀飯粥品以雞湯熬煮，入口溫潤不刺激，配小菜清淡，特別適合胃敏感或噁心時
disgust	old	阿盟壹部落廚房	以水煮、燻烤、醃漬的排灣族原民料理呈現清爽甘香，像刺蔥皮蛋、地瓜球、小米粽
disgust	young	有時散步 Walkabout Café	其餐點以清爽、健康為主，像是德式香腸搭配蔬果沙拉與法國麵包，或是燻雞肉與佛
disgust	old	寶桑豆花	傳承逾40年，採用非基改黃豆手工製作，豆花滑嫩、糖水清香不膩。搭配紅豆、綠豆
disgust	young	秒殺水煎包	拳頭大小、高麗菜高纖水煎包，外皮金黃恰恰，內餡鮮甜多汁。雖為人氣破表的秒殺

圖 5 各情緒分別對應的年齡標籤以及美食推薦

包含下列六項：

- emotion：情緒標籤（如：開心、難過、生氣等）
- age_group：年齡歸類，依據年紀歸類於不同標籤。（如：old 為銀髮族群，60 歲以上（含 60 歲）；young 為青年族群，60 歲以下）
- food_name：推薦美食名稱（如：牛牛鮮奶酪專賣店）
- reason：推薦該美食對應該情緒的原因與說明
- address：美食店家的實體地址
- website：美食店家的官方網站或相關介紹頁面之連結

每種情緒皆對應 7 至 9 種不同美食，且每項美食會根據對應情緒附上具說服力之推薦原因，以提升推薦系統之可信度與使用者接受度。當使用者透過後端進行情緒辨識時，系統將根據其所辨識出的情緒，自 Excel 資料中擷取對應之美食及推薦說明，並顯示於前端網頁介面上。例如，當系統辨識使用者情緒為「難過」時，推薦之美食可能為「牛牛鮮奶酪專賣店」，其推薦說明如：「傷心時，不妨來一份牛牛鮮奶酪專賣店的奶酪。奶酪中富含的色胺酸，是製造大腦快樂化學物質血清素的重要原料，能幫助穩定情緒、減輕焦慮。享受奶酪細膩滑順的口感，同時給身心帶來放鬆，讓悲傷慢慢淡去。」該店家地址以及網址也會隨之顯示於網頁呈現頁面。

為進一步提升推薦系統之個人化與實用性，本研究亦納入使用者年齡層作為推薦依據之一，將使用者區分為兩大年齡群組，60 歲以下（青年族群）與 60 歲以上（銀髮族群）。每個情緒類別下所對應之美食，將依據年齡層進行再次分類與調整：

- 青年族群（60 歲以下）：推薦項目偏向創新、有趣、具話題

性之美食，兼顧營養與感官刺激，例如療癒系甜點、異國料理、小眾手搖飲等，期望引起共鳴與興趣。

- 銀髮族群（60 歲以上，含 60 歲）：推薦項目以健康、溫和、低糖低鹽之美食為主，注重腸胃負擔、營養攝取與飲食安全，如溫補湯品、豆花、天然無添加甜品等，協助長者情緒調適同時兼顧健康。

本研究期望透過情緒與年齡並列考量之推薦機制，提供使用者一套兼具人性化、美食互動性與健康導向之推薦系統介面，進而促進情緒自我覺察、健康管理與生活品質提升。

4.5 系統部署與應用介面設計：Hugging Face Spaces

4.5.1 部署背景與選擇動機

隨著本研究在人臉情緒辨識與台東美食推薦系統的基礎功能開發完成後，開始著手於系統部署與可公開使用的互動介面設計。選擇適當的平台進行應用部署是系統落地的關鍵步驟，它不僅影響使用者的操作便利性，也關係到展示場合的應用可能性。因此，本研究綜合評估了目前主流的幾種部署方法，包括自行架設伺服器、在 Google Colab 結合 ngrok 暫時性公開連結、或是使用如 Hugging Face Spaces 等平台式部署方案。最終，本研究決定採用 Hugging Face Spaces 進行公開部署。

Hugging Face Spaces 由自然語言處理與開源機器學習社群 Hugging Face 所開發的平台，提供研究人員與開發者能夠快速部署、分享與展示模型與互動式應用。相較於其他方法，Spaces 的主要優勢包括：免安裝伺服器架構、無需額外申請 API 金鑰、支援 Gradio 應用架構、介面部署快速、且為免費服務。此外，其整合 Git 儲存庫與自動部署功能，使得開發流程更為自動化與穩定，特別適合用於學生專題、學術展示與教育推廣等情境。

再者，Spaces 提供的介面簡潔直觀，使用者僅需透過網頁即可上傳圖片、觸發模型運算並取得推薦結果，對於非技術背景的使用者極為友善。由於本研究最終需要進行成果發表與展示，因此選擇

一個穩定、高可用性、低進入門檻的平台具有實際與策略雙重考量。此選擇不僅強化了本系統的可實用性，也顯示出本研究在技術部署階段對於「開放性」與「延展性」的重視。

4.5.2 系統架構與部署流程

本系統部署至 Hugging Face Spaces 的架構主要由三個核心組成：情緒辨識模組、資料查詢模組與使用者互動介面模組，三者皆以 Python 為主要開發語言，並依賴於 Gradio 所提供的前端互動架構。

在情緒辨識部分，本研究使用了 DeepFace 函式庫中預訓練的模型進行推論分析。DeepFace 可辨識七種主要情緒，包括：Happy、Sad、Angry、Surprise、Fear、Disgust 與 Neutral。上傳圖片後，系統會先進行人臉偵測與裁切，接著進行情緒分析，並嘗試根據臉部特徵推估年齡區間，以便後續進一步的推薦精準化。

資料查詢模組則負責根據辨識結果，從本地 CSV 資料中篩選出對應的推薦項目。CSV 檔案目前包含六個欄位，分別為：emotion（情緒標籤）、food_name（美食名稱）、reason（推薦理由）、address（地址）、website（網站連結）以及新增的 age_group（年齡層）。其中，age_group 欄位的加入，使系統能根據使用者推估之年齡資訊與情緒結果進行交叉查詢，提升推薦內容的針對性與精準度並強化使用者體驗。

為了成功部署於 Hugging Face Spaces，本研究將程式主體整理為 app.py，並建立 requirements.txt 以列出所有相依套件，如 deepface, gradio, opencv-python, pandas 等。系統部署流程為：將上述檔案與資料一併推送至 Git 儲存庫，再由 Hugging Face Spaces 自動讀取並啟動。部署平台會根據 .huggingface.yml 設定檔進行應用初始化，包括指定執行指令、設定資源環境與端點導向，使得整體部署流程自動化，幾乎不需手動介入，大幅提升開發效率與版本一致性。

4.5.3 使用者互動流程與輸出內容

部署完成後，系統透過 Hugging Face 所提供的公開連結對外開放，使用者可直接透過瀏覽器訪問本系統介面，並進行互動操作。整體互動流程設計如下：

一、圖片上傳功能：使用者可拖曳或點擊介面中的上傳區塊，選擇一張包含清楚人臉的圖片檔案，支援常見格式如 JPG、PNG 等。

二、即時辨識處理：圖片上傳後，系統進行人臉偵測與情緒推斷，辨識結果包含主要情緒標籤與預測年齡層。

三、結果輸出與推薦美食顯示：系統會根據情緒與年齡層比對資料庫，篩選出一筆推薦美食資料，並同步顯示以下資訊：推薦美食名稱、推薦理由、美食地址、超連結按鈕(點擊後可連至該餐廳的官方網站或 Google 評論頁面)。

此設計不僅能讓使用者獲得即時的系統意見反應，也透過情緒與地方文化的結合，創造一種「情境式推薦」的互動體驗模式，進一步彰顯系統的地方創生應用價值。

4.5.4 自訂介面設計與視覺優化

為了強化系統在展示場域的使用性與視覺吸引力，本研究亦特別針對 Hugging Face Spaces 的介面進行客製化設計。Gradio 雖提供簡易的模組式介面，但透過 `gr.HTML()` 及 `gr.CSS()` 元件，本研究進一步加入了自訂的標題區、分段提示、說明文字與品牌配色設計。

具體設計包含：

一、頁首標題與說明：使用 `gr.HTML()` 嵌入標題文字「情緒美食推薦」，下方輔以簡單的使用說明，包含操作步驟與注意事項。

二、配色與排版：介面以天藍色為主，搭配邊框與陰影效果，提升整體視覺層次感。區塊排版採用分欄設計，使圖片上傳、辨識結果與美食資訊呈現在同一畫面上，不需滑動即可一覽。



我們是一群就讀於國立台灣大學的學生，致力於將 AI 與地方文化結合，讓科技與在地生活產生連結。

圖 6 前端網頁呈現

怎麼使用？

🔥 打開網頁

使用手機或電腦瀏覽器，進入本情緒辨識台東美食推薦系統的網頁。

📷 準備自拍照

你有兩種方式：

- 直接拍照：點選「選擇檔案」或「拍照」按鈕，手機會自動開啟前置鏡頭，讓你直接拍攝一張自拍照。
- 上傳照片：如果你手機或電腦裡已有自拍照，可以直接選擇該照片上傳。

✅ 確認照片

選擇或拍攝完照片後，網頁會顯示照片預覽，請確認照片清晰且是你想分析的自拍照。

📤 送出照片

按下「上傳並分析情緒」按鈕，系統會將照片送到 AI 進行情緒分析。

🕒 等待分析結果

只要幾秒鐘，AI 會根據你的臉部表情辨識出你當下的情緒（如開心、難過、驚訝等）。

🍽️ 查看美食推薦

系統將根據情緒，推薦今日最適合你的台東特色美食，結果會直接顯示在頁面上。

😊 享受美食

根據推薦，你可以安排今天的美食行程，讓情緒與味蕾都獲得滿足！

圖 7 前端網頁呈現



圖 8 前端網頁呈現

圖 8 為利用網路上公開的明星照進行辨識。首先，存取時該圖時已先確認此圖片中對象年齡為 41 歲，於系統辨識後顯示年齡為 42 歲，此分析符合本研究所設定的 ± 5 歲範圍。再者，辨識之情緒為開心，表示辨識正確，並依據開心的標籤，匯入對應的美時推薦以及

推薦原因。

此設計不僅提升了系統的美觀度，也讓使用者操作流程更有人性化，對於非科技背景的使用者而言更具親和力。在本研究進行內部展示測試時，亦有多位測試者反映介面設計清楚、資料呈現豐富而不繁瑣，顯示自訂化設計對提升使用者體驗具有明顯成效。

4.5.5 平台優勢與整合效益

綜合以上部署與介面設計成果，可歸納出 Hugging Face Spaces 作為應用展示平台所帶來的多項優勢。首先，其免費開放的特性讓本專題在資源有限的情況下，仍能實現公開可訪問的網路應用展示。其次，支援 Gradio 的高相容性與部署自動化流程，使開發者可專注於應用邏輯與資料設計，而無需處理繁瑣的伺服器與網路設定。

此外，Hugging Face Spaces 本身具有社群分享機制與搜尋曝光機制，當專案設為公開時，亦可被其他使用者搜尋與點閱，間接提升本研究的可見度與影響力。更重要的是，此平台的架構與互動設計高度模組化，便於未來擴充其他功能，如即時攝影機串接、使用者回饋模組、地理位置匯入等，為系統日後的擴展奠定良好基礎。最終，本研究透過本部署模組成功實現了從情緒辨識到食物推薦的完整閉環流程，並將其呈現在一個可即時互動、跨平台使用且易於展示的雲端介面中。這不僅完成了技術面與應用面的整合，更體現出本研究將人工智慧應用於地方文化推廣的核心理念與實踐能力。

第五章、結論與問題探討

5.1 研究成果

本研究旨在建構一套結合人臉辨識與情緒分析技術的即時美食推薦智慧系統，期望能根據使用者的臉部表情與情緒狀態，提供個人化且具情境感知能力的飲食建議。系統整合影像偵測、深度學習模型、年齡推估與規則式推薦機制，讓使用者能透過攝影鏡頭或影像上傳的方式，獲得即時且符合心理狀態的美食推薦結果。

在系統的實作層面，主要採用了 OpenCV 進行人臉偵測與區域擷取，並以 DeepFace 架構中的 CNN 模型（包含 VGG16 與 mini-Xception）進行情緒分類分析。年齡預測則使用 InsightFace 之 buffalo_l 模型進行推估，並透過線性回歸修正公式，使預測值更貼近實際年齡。整體架構透過 Gradio 前端介面與後端模型串接，讓使用者可於網頁環境中直接操作，系統最終部署於 Hugging Face Spaces 平台，使其能跨裝置運行、免安裝即可體驗。

從測試結果來看，情緒辨識準確率達 73%，其中以「快樂」與「驚訝」的辨識效果最為穩定，顯示模型在正向與高表情強度的類別上表現良好。年齡預測經回歸修正後，大部分結果可落於 ± 5 歲的誤差範圍內，符合國際年齡推估標準。整體運作流程平均反應時間短，具備即時應用潛力，且系統介面設計簡潔、操作容易，提升了使用者體驗。

在內容應用面，本研究亦結合心理學文獻與地方文化特性，建立「情緒－食物」關聯表，以台東在地美食為核心推薦項目。透過不同情緒對應合適的餐點，如「開心」時推薦甜食、「難過」時建議溫暖湯品等，使推薦結果更具心理依據與文化親和力。綜上可知，本研究成果已具備初步可行性，展現出人工智慧技術在日常飲食與情緒互動領域的應用潛力。

5.2 問題探討

雖然本系統成功實現即時辨識與推薦的功能，但在開發與測試過程中，仍面臨若干技術與設計上的挑戰。綜合觀察與分析後，主要問題可分為兩項：情緒辨識準確度不足與即時顯示不穩定。

5.2.1 情緒辨識準確度不高

即使採用了深度學習的 CNN 模型結構，並於訓練階段進行資料前處理與增強，模型在部分情緒類別仍易出現誤判現象。例如「驚訝」與「害怕」、「悲傷」與「厭惡」等情緒，其臉部表情特徵極為相似，差異主要呈現在微表情層面。由於影像資料中個體表達方式存在差異，加上光線、角度、膚色與表情幅度等環境因素的干擾，導致模型難以有效辨別這些細微特徵。此外，資料集中部分情緒樣本不足，使得模型在學習時產生偏倚，進而降低整體辨識準確度。

5.2.2 即時畫面顯示不穩定

在鏡頭即時偵測模式下，系統能即時擷取使用者臉部影像並進行情緒分析，但由於人類表情變化頻繁，系統在每一幀畫面進行推論時，輸出的情緒結果會快速變動，造成標籤閃爍與畫面不穩定的情形。這種狀況會讓使用者難以清楚理解系統當下的判斷結果，也可能影響整體體驗感受。本研究後期為改善此問題，改以靜態截圖方式進行辨識，以提升穩定度。然而，此做法雖可避免標籤閃爍問題，卻削弱了系統「即時互動」的特性，導致實際應用層面仍需進一步優化。

5.3 研究限制

本研究在模型訓練與系統實作過程中，仍存在若干限制，主要可分為「準確度限制」與「鏡頭顯示限制」兩個面向。

5.3.1 準確度限制

情緒辨識準確率受多重因素影響。首先，資料來源的多樣性不足導致模型泛化能力受限，尤其在真實場域下，臉部表情會受到光

線、角度、遮蔽物（如眼鏡、口罩）等外在因素影響，使模型難以穩定擷取關鍵特徵。其次，不同性別、年齡與文化背景的使用者，其情緒表達方式存在顯著差異，使得模型無法以單一準則應對所有個體，造成辨識偏誤。此外，部分情緒（如「厭惡」）的樣本數量過少，導致訓練結果不平衡。綜合而言，準確度限制源於資料品質、個體差異與環境干擾的交互影響。

5.3.2 鏡頭辨識顯示限制

即時辨識功能在畫面顯示上仍有改進空間。系統目前未設置穩定化機制（例如時間平滑或信心閾值控制），導致畫面上情緒標籤頻繁變換，影響使用者理解。雖然採用靜態截圖能提升顯示穩定性，但此方式限制了系統的互動性與即時應用價值。若未來能在畫面更新間加入信心度判斷或動態平滑策略，將有助於提升人機互動的自然性與可讀性。

5.4 未來展望

綜合本研究成果與限制，未來研究可從以下幾個方向持續深化與拓展：

5.4.1 強化模型訓練資料的多樣性

未來可擴充更大規模、更多族群的臉部表情資料集，並加入自然情境下的表情樣本，以提升模型對環境變化與個體差異的適應能力，增強其在真實場域的準確度與穩定性。

5.4.2 導入學習型推薦演算法

現階段推薦邏輯以規則式為主，未能根據使用者行為動態調整。未來可導入協同過濾或強化學習等演算法，使系統能從使用者的互動中自我學習，不斷優化推薦結果，提升個人化與準確性。

5.4.3 結合多模態資料分析

除了臉部影像外，可整合語音語調、姿態行為、甚至環境光線與聲音等多模態資訊，讓系統能更全面理解使用者情境，朝向情感計算（Affective Computing）與智慧人機互動發展。

5.4.4 優化即時顯示與互動機制

未來可於情緒顯示端加入平滑演算法、信心閾值控制與結果時間加權機制，平衡辨識速度與穩定性。同時可設計更友善的互動介面，增進使用者理解與參與感，提升系統實務應用價值。

整體而言，本研究已成功展現人工智慧於情緒辨識與個人化推薦整合的可能性。若未來能持續改善模型效能並拓展多模態應用，將可使系統在智慧服務、心理輔助、健康照護及文化推廣等領域發揮更大的社會影響力。

參考文獻

- [1] AbidAbubakar. (2021 年 5 月 25 日). Using & Mixing Hugging Face Models with Gradio 2.0. 擷取自 Using & Mixing Hugging Face Models with Gradio 2.0: <https://huggingface.co/blog/gradio>
- [2] Abubakar Abid,et al. (2019). Gradio: Hassle-Free Sharing and Testing of ML Models in the Wild. 擷取自 Gradio: Hassle-Free Sharing and Testing of ML Models in the Wild: https://arxiv.org/abs/1906.02569?utm_source.com
- [3] Ali Mollahosseini,et al. (2017). AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. IEEE Transactions on Affective Computing, 18-31.
- [4] BBC. (2021 年 1 月 14 日). 飲食與健康：我們對安慰食品依賴背後的科學道理. 擷取自 BBC Chinese Traditional: <https://www.bbc.com/zhongwen/trad/chinese-news-55665921>
- [5] BradskiGary. (2000 年 11 月). The OpenCV Library. Dr. Dobb' s Journal of Software Tools, 頁 120-126.
- [6] DengLi、 WeihongShan. (2020 年 3 月 16 日). Deep Facial Expression Recognition: A Survey. IEEE Transactions on Affective Computing, 頁 1195-1215.
- [7] DevonportT. J., NichollsW., & FullertonC. (2019). A systematic review of the association between emotions and eating behaviour in normal and overweight adult populations. Journal of Health Psychology.
- [8] Egon P. Köster, Jozina Mojet. (2015). From mood to food and from food to mood: A psychological perspective on the measurement of food-related emotions in consumer research. From mood to food and from food to mood: A psychological perspective on the measurement of food-related emotions in consumer research.
- [9] GoodfellowJ.Ian. (2017). Generative Adversarial Networks: An Overview. ACM Computing Surveys (CSUR), 10.
- [10]GoogleBrain. (無日期). TensorFlow: An end-to-end platform for machine learning. 擷取自 TensorFlow: An end-to-end platform for machine learning: <https://www.tensorflow.org/>
- [11]Gradio. (無日期). Gradio documentation. 擷取自 Gradio: <https://www.gradio.app/>
- [12]GradioTeam. (無日期). Gradio – Build Machine Learning Web Apps. 擷取自 Gradio – Build Machine Learning Web Apps: <https://www.gradio.app/>
- [13]Guo, R et al. (2024). Development and application of emotion

- recognition technology — a systematic literature review. BMC Psychology.
- [14]Hu Han,Charles Otto,Anil K. Jain. (2013 年 6 月). Age Estimation from Face Images: Human vs. Machine Performance. Madrid, Spain.
 - [15]Jason Yosinski,et al. (2014). How transferable are features in deep neural networks?, (頁 10).
 - [16]Jiankang Deng,et al. (2022). ArcFace: Additive Angular Margin Loss for Deep Face Recognition.
 - [17]Kaiming He,et al. (2016). Deep Residual Learning for Image Recognition., (頁 770 – 778).
 - [18]KanekoDaisuke. (2018). Methods for Evaluating Emotions Evoked by Food Experiences: A Literature Review. Methods for Evaluating Emotions Evoked by Food Experiences: A Literature Review.
 - [19]Karen Simonyan, Andrew Zisserman. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition., (頁 7).
 - [20]LeCun,et al. (2015). Deep learning. Nature, 頁 436-444.
 - [21]Martín Abadi,et al. (2016). TensorFlow: A System for Large-Scale Machine Learning., (頁 21).
 - [22]OpenCV.(n.d.). (無日期). OpenCV. (n.d.). About. OpenCV.org. Retrieved. 擷取自 OpenCV. (n.d.). About. OpenCV.org. Retrieved: <https://opencv.org/about/>
 - [23]OrganizationHealthWorld. (無日期). 擷取自 World Health Organization: https://www.who.int/health-topics/ageing#tab=tab_1
 - [24]Rozanne Kruger,et al. (2007 年 11 月). Validation of a food frequency questionnaire to assess macronutrient and calcium intake in New Zealand toddlers. Appetite, 頁 490-495.
 - [25]S. E. Bekhouche,et al. (2024 年 8 月 16 日). Facial Age Estimation Using Multi-Stage Deep Neural Networks. Electronics (MDPI).
 - [26]Sureliya, S,et al. (2024). EmoCuisine: Emotion-Based Restaurant Recommendation System. International Journal of Progressive Research in Science and Engineering.
 - [27]W. Zhao,et al. (2003). Face Recognition: A Literature Survey.
 - [28]Wengfish. (2023 年 11 月 20 日). 〈情緒穩定是可以吃出來的！焦慮、煩躁、心情低落時對應的 5 類食物，可減少煩憂 提升情緒管理能力！〉. 擷取自 Marie Claire 美麗佳人: <https://www.marieclaire.com.tw/beauty/health-fitness/76101>
 - [29]Yaniv Taigman,et al. (2014). DeepFace: Closing the Gap to Human-Level Performance in Face Verification.
 - [30]YildirimE., AkbulutF. P., & CatalC. (2023). Analysis of facial emotion expression in eating occasions using deep learning. Multimedia Tools and Applications.

- [31]Zhanpeng Zhang,et al. (2014). Learning Deep Representation for Face Alignment with Auxiliary Attributes., (頁 14).
- [32]Doreen Virtue.(2012)。食物情緒大解密：你愛吃的，透露了什麼祕密。
- [33]林宗儒。(2024)。植基於深度學習之快速人臉辨識機制之研究。碩士論文。國立國防大學。
- [34]許捷鳴(2016)。以情緒為導向之飲食建議行動應用程式設計。碩士論文。國立中興大學。
- [35]陳映志。(2020)。基於深度學習之人臉辨識系統。碩士論文。國立台北科技大學。