

# 基於光學字元辨識之智慧閱卷模型

## Intelligent Grading Model Based on Optical Character Recognition

賴家煜

國立臺中科技大學 資訊工程系  
a0909956502@gmail.com

黃馨蝶

國立臺中科技大學 資訊工程系  
alisa991200@gmail.com

劉嘉雯，黃祈勝\*

國立臺中科技大學 資訊工程系  
{ daisy716, vcshuang}@nutc.edu.tw

【摘要】在教育評量的需求不斷擴增下，傳統紙本試卷的批改過程往往費時且易受主觀判斷所左右。隨著近年電腦視覺與深度學習技術逐漸成熟，使用像是 YOLOv7 的物件偵測模型，來自動辨識試卷上的題目、圖表以及手寫作答區等不同區塊，展現出相當可觀的應用潛力。為了因應試卷影像的多元性與品質差異，本研究計畫結合資料增強及遷移學習策略，評估 YOLOv7 在多樣化試卷影像的偵測表現，並設計一套初步的光學字元辨識流程，期望為數位化評量在效率與準確度上帶來進一步的助益。

【關鍵詞】YOLOv7、物件偵測、試卷辨識、深度學習、教育科技

**Abstract:** With the increasing demand for educational assessments, the traditional manual grading of paper-based tests often proves time-consuming and susceptible to subjective judgment. In recent years, advances in computer vision and deep learning have led to promising applications of object detection models, such as YOLOv7, for automatically identifying various sections of a test paper, including questions, diagrams, and handwritten response areas. To address the diverse quality and formats of test images, this study integrates data augmentation and transfer learning strategies to evaluate the detection performance of YOLOv7 on heterogeneous test-paper images. Additionally, a preliminary optical character recognition (OCR) workflow is designed in the hope of further enhancing the efficiency and accuracy of digital assessments.

**Keywords:** YOLOv7, Object Detection, Test Paper Recognition, Deep Learning, Educational Technology

## 1. 前言

在教育評量的需求不斷擴增下，傳統紙本試卷的批改過程往往費時且易受主觀判斷所左右。隨著測驗規模的擴大，教育工作者需投入大量人力與時間進行校對與批改，且容易因個人主觀意識或生理疲勞等因素，導致誤判或漏批的情形。近年來，隨著電腦視覺（Computer Vision）與深度學習（Deep Learning）技術的迅速發展，物件偵測模型，如：YOLOv7（You Only Look Once Version 7）[4]逐漸成熟，為自動化辨識與分析試卷內容帶來了新的契機。YOLOv7 以其高效的偵測速度與準確度，在車牌辨識（Automatic Number-Plate Recognition, ANPR）、公共場所安全監控及疫情期間的口罩配戴偵測等多種應用場景中展現出顯著成效。本研究旨在結合 YOLOv7 模型、資料增強（Data Augmentation）與遷移學習（Transfer Learning）技術，自動辨識試卷中的各類區塊，包括題目、文字段落、作答區域與圖表等。透過先行的區域偵測與精準的區塊切割，再輔以（Optical Character Recognition, OCR）[5]技術進行文字內容的解析，本研究期望能有效提升數位化評量的效率與準確度，減少人工作業的時間與潛在誤差，推動教育評量的數位化與自動化發展，為各教育單位提供可靠且高效的評量工具，促進教育科技（Educational Technology, EdTech）的應用與發展。

## 2. 相關研究與文獻探討

### 2.1. 試卷影像辨識（Image Recognition）在教育領域的應用

試卷影像辨識技術旨在透過自動化工具快速且準確地處理和分析紙本試卷，從而提高評量效率並降低人工作業錯誤。過去的研究多聚焦於運用光學字元辨識之技術將印刷文字轉為數位文本[1]，然而，這些研究往往忽略了試卷中手寫筆跡、圖表及多樣化題型等複雜元素的辨識。部分研究嘗試結合影像處理技術來分割一份文檔中的不同區域[6]，許多處理方法僅適用於結構簡單或排版固定的試卷，對於多樣化版面配置的試卷則表現較為不佳。因此，現有的試卷影像辨識技術在面對實際教育場景中多元且複雜的試卷格式時，仍存在顯著的局限性。

### 2.2. 光學字元辨識（OCR）技術的應用與局限

光學字元辨識之技術在自動車牌辨識[2]及小型文字識別（如：點讀筆）等場景中表現穩定，特別是在處理規整的印刷字體時。於試卷中通常包含大量不規則的手寫字跡、複雜的文章段落及圖表，這些特徵超出了傳統 OCR 的最佳適用範疇，導致辨識精度下降。在部分研究中有嘗試運用影像處理演算法對文檔中的特定區域進行分割[7]，但這些方法僅適用於排版單一或結構簡明的卷面，當試卷版面格式多樣或含有大量圖表時，辨識效果顯著性較低。

### 2.3. 物件偵測模型（YOLOv7）的應用優勢

YOLO 系列模型作為一階偵測架構（One-Stage Detector），以其「一次性」完成目標分類與邊界框（Bounding Box）定位的特點，顯著降低了推論（Inference）時間，適用於

實時應用（Real-time Application）[3]。自初代 YOLO 至 YOLOv4、YOLOv5 及 YOLOv7，該系列模型在網路結構（Network Architecture）與訓練策略上不斷優化，提升了準確度與速度的平衡。YOLOv7 特別在高效偵測能力上表現優異，能在多種類別區塊的試卷中同時完成多物件的定位與分類，適應性強，且在中大型資料集上維持不錯的準確度。

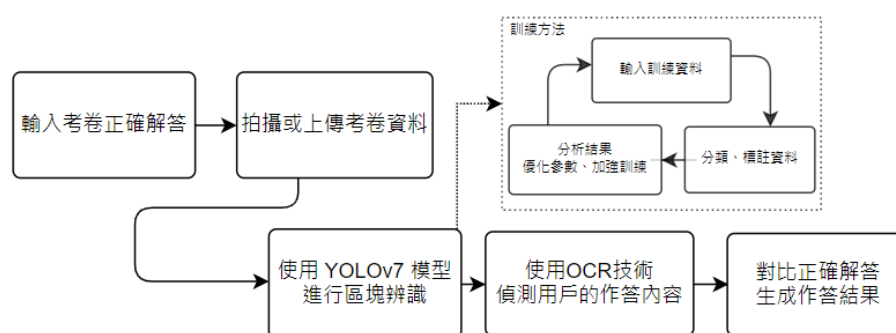
## 2.4. 資料增強與遷移學習技術的應用

本研究將結合資料增強（Data Augmentation）[8]與遷移學習（Transfer Learning）[9]之技術。透過將影像旋轉、翻轉、裁切及亮度調整等增強手段，模擬不同拍攝角度、紙張皺褶與筆跡變化，提升模型對多樣化影像的泛化能力。利用在大型資料集（如：ImageNet）上預訓練完成的權重，再在試卷影像上進行微調（Fine-tuning），可有效補足樣本數與資料品質參差的問題，以增強模型在實際應用中的穩定性與準確度。

## 3. 方法與研究結果

### 3.1. 整體架構與應用潛力

本研究提出一套完整的試卷影像自動辨識流程，結合 YOLO V7 在多樣化試卷影像中的高效偵測能力與 OCR 在文字解析中的精確性。透過先行的區域偵測與精準的區塊切割，再輔以 OCR 技術進行文字內容的解析，期望能有效提升數位化評量的效率與準確度，減少人工作業的時間與潛在誤差，推動教育評量的數位化與自動化發展。此整合式流程不僅為教育單位提供更為可靠與高效的評量工具，亦促進教育科技的進一步應用與發展，為未來的大規模自動批改與教學分析奠定堅實的技術基礎。



圖一、系統架構流程圖

本研究旨在評估 YOLOv7 在試卷影像辨識中的效能，並透過資料增強與遷移學習技術來提升模型的準確度。實驗分為三個階段進行，以逐步優化模型性能。

### 3.1. 使用拍照的試卷資料進行訓練

初步作業是使用手機拍攝的試卷資料進行實驗，目的為測試模型是否能有效擷取題目和作答區域，並辨識手寫特徵進行判斷。實驗結果顯示，模型確實能夠初步識別題目區和

作答區，但準確度較低，僅在 0.02 至 0.15 之間，僅能捕捉到相關區塊，且定位並不精準。針對結果，推測可能的原因如下：

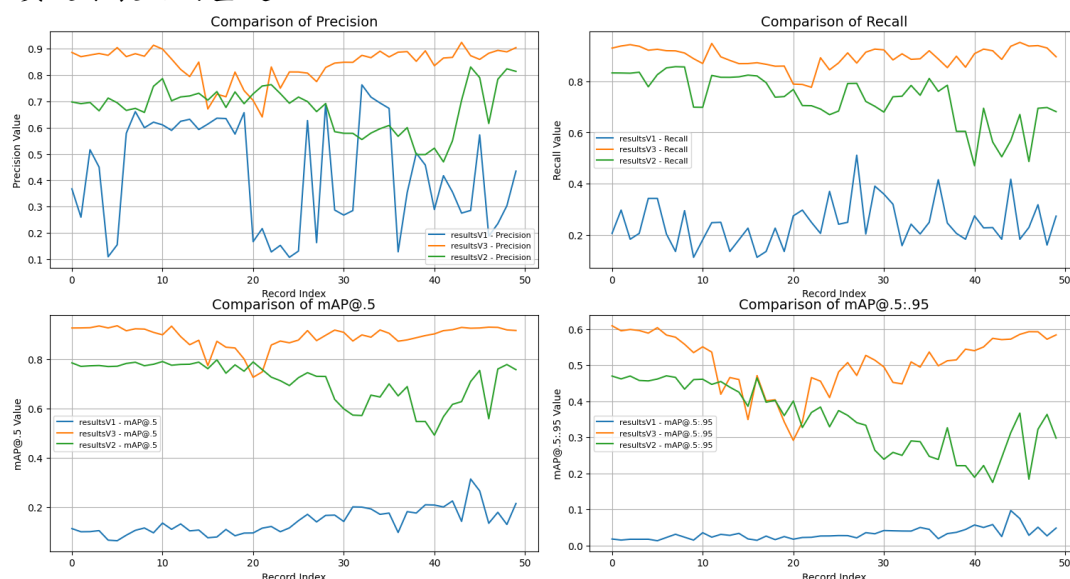
1. **資料過於複雜**：本次實驗使用的是手機拍攝的相片資料，受到光線、角度、距離等因素的影響，可能會使模型的判斷結果造成干擾。
2. **資料量不足**：訓練數據樣本有限，未能提供足夠的多樣性和數量，導致模型的判斷準確度無法提升。

### 3.2. 增強空白試卷的電子檔訓練

本次訓練結果顯著改善了第一次實驗中的不足，大幅提升了題目判斷的準確度。準確率穩定落在 0.3 至 0.8 之間，模型能準確抓取作答的括弧區塊、題目區塊、文章區塊和圖片區塊。特別值得肯定的是，填充題中的括弧區塊被判斷為作答區，以及簡答題的區塊也被成功抓取，這說明模型在辨識題型結構方面有了顯著進步。

然而，由於本次訓練僅使用空白電子檔，模型在手寫作答部分的辨識能力尚待提升。例如，手寫字母容易被判斷為圖像，顯示模型在手寫特徵辨識上的弱項。未來建議引入更多包含手寫作答的訓練數據，針對手寫文字進行專項優化，以進一步提升模型的綜合表現。

整體而言，第二次訓練不僅補足了第一次的不足，還實現了多項重要突破，為後續進一步優化奠定了良好的基礎。



圖二、實驗結果準確度對比

### 3.3. 增強手寫筆跡訓練資料

於第三次實驗中，本研究專注於優化手寫文字的辨識能力，並引入了更多手寫資料作為訓練樣本。這些樣本涵蓋了多樣化的字跡特徵，以增強模型對手寫內容的解析能力。

實驗結果顯示，此次優化有效解決了先前模型將手寫文字誤判為圖片的問題，辨識準確度有了明顯提升。然而，在面對較為複雜的情況時，模型仍有改進空間。例如，對於字

跡凌亂的內容、手寫計算過程中的符號和數字、橡皮擦未擦乾淨的痕跡，以及紙張透光造成的干擾等情況，模型的判斷仍顯不足。

### 3.4. 結論

經過三次模型訓練與優化，本研究在題目與作答區域的定位及手寫文字的辨識上取得了顯著進步，並且在四個關鍵指標（Precision、Recall、mAP@.5 和 mAP@.5:.95）上也顯示出明顯的提升。尤其是在第三次的訓練結果中，所有指標的表現都達到較高且穩定的水準，顯示模型在準確性、召回率及精度方面均有所突破，為實際應用奠定了堅實的基礎。Precision 和 Recall 分別穩定維持在 0.80 至 0.9 之間，表現出色，確保了題目區域與作答內容的精準識別；而 mAP@.5 和 mAP@.5:.95 也顯示出該模型在多種 IoU 範圍下的優越性能，分別穩定在 0.80.9 和 0.5 至 0.6 之間，顯示模型在處理大範圍紙本數據時的高效性與準確度。

未來，本研究將結合硬體設備，將這些優化過的演算法實踐於硬體平台，發展成自動化的閱卷批改系統。該系統將結合掃描模組與硬體架構，並利用本系統的題目定位與手寫辨識技術，實現自動定位題目區域與作答內容，進一步提升閱卷效率與準確性。類似於傳統 2B 鉛筆閱卷技術，但不同的是，該系統不僅僅進行識別，還能加入自動化批改功能，顯著減少錯誤率，並大幅減輕教育工作者在批改試卷上的工作負擔。

### 參考文獻

1. 彭星瑋 (2020)。一個基於 OCR 之文字辨識系統。碩士論文。國立高雄科技大學。
2. 劉振隆、劉恩睿 (2023)。深度學習 YOLOv4 物件偵測技術建立即時動態車牌辨識系統。臺東大學綠色科學學刊，13 (1)，1–18。
3. Kuo, E. W., (2018)。YOLO 物體偵測 (Object Recognition) 演算法概念與介紹。GitHub。
4. Wong, K. Y., & Yiu, W. K. (2023, November 3). *WongKinYiu/yolov7*. GitHub.
5. Jaied ai. (2023, September 24). *JaiedAI/EasyOCR*. Github.
6. HumanSignal. (2022, September 23). *HumanSignal/labelImg*. GitHub.
7. Birgit, P., Christoph, A., Michele, D., Ahmed S. N., Peter W. J. S. (2022, June). *DocLayNet: A Large Human-Annotated Dataset for Document-Layout Analysis*. Paper presented at ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (28<sup>th</sup>). (2022). Washington DC, U.S. Association for Computing Machinery.
8. Suorong Yang, Weikang Xiao, Mengchen Zhang, Suhan Guo, Jian Zhao & Furao Shen (2023, November). *Image data augmentation for deep learning: A survey*. Paper presented at Laboratory for Novel Software Technology, Nanjing University, China.
9. Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong & Qing He (2020, July). *A Comprehensive Survey on Transfer Learning*. Paper published in Proceedings of the IEEE (2021).