

A Green Coffee Beans Inspection System Using Deep Learning

Jiann-Shu Lee ^{1*}, Ya-Lun Tsai ¹, Shian Lin ¹, Chun-Chi Lin ², Yao-Xian Ma ¹

¹ Department of Computer Science and Information Engineering, National University of Tainan, Tainan, 70005, Taiwan

² Dachu Coffee Estate, Tainan, 733016, Taiwan

*E-mail : cslee@mail.nutn.edu.tw

Abstract

The so-called defective coffee beans refer to coffee beans of poor quality. These blemish beans can give a cup of coffee an uncomfortable taste. At present, most of selecting defective beans is still artificially, and the selection is made through eye examination. However, this manual classification method not only consumes manpower and time, but also often causes the problem of unstable screening quality due to human negligence and inconsistent judgment standards of different people. These problems can be overcome by automatic detection mechanism. Therefore, this research proposes a model for automatically classifying coffee beans. This model uses VGG16 as the main architecture to improve the classification accuracy of coffee beans through transfer learning. Experimental results show that this model can achieve 99% accuracy.

Keywords: Coffee beans inspection, machine learning, transfer learning.

* Corresponding author: cslee@mail.nutn.edu.tw
DOI : 10.3966/222344892021101102005

植基於深度學習之咖啡原豆瑕疵檢測系統

李建樹^{1,*}, 蔡亞倫¹, 林萱¹, 林俊吉², 馬耀賢¹

¹ 國立臺南大學資訊工程學系 ² 大鋤花間咖啡生態農場

摘要

所謂的瑕疵豆就是指品質不好的咖啡豆，這些瑕疵豆會讓一杯咖啡產生令人不舒服的味道，現在挑選瑕疵豆方法大多還是以人為方式為主，通過眼睛檢查進行挑選。然而這種人工分類方式不僅耗費人力與時間，而且常常會因為人為的疏失以及不同人的判斷標準不一致，造成篩選品質不穩定的問題。這些問題可藉由自動檢測機制來加以克服，因此，本研究提出一個自動分類咖啡豆的模型，此模型以 VGG16 為主架構，透過轉移學習提升咖啡豆的分類準確度，實驗顯示本模型能得到 99% 的準確率。

關鍵詞：咖啡豆檢測、監督式學習、轉移學習

1. 緒論

就咖啡原豆而言，所謂的瑕疵豆就是指品質不好的豆子，這些瑕疵豆會讓咖啡產生令人不舒服的味道，像是霉味、臭味、尖銳且刺激的酸等等味道及口感，因此，咖啡農或者加工廠商必須設法儘量剔除這些瑕疵豆。目前雖然有一些自動化剔除瑕疵豆的挑選機，但其價格高昂，動輒上百萬元，不是一般咖啡農所能負擔，而且挑選機的準確率還不是非常高，導致使用者常常面臨：如果將系統設定成嚴格剔除瑕疵豆，會伴隨著許多好豆也被犧牲掉；相反地，如果將系統設定成儘量不要犧牲好豆，則會有許多瑕疵豆蒙混過關，讓使者頭疼不已。造成有非常多咖啡農或業者寧願使用人力方式來篩選咖啡豆，導致耗費許多時間以及人力成本，影響其競爭力。因此，市場上極度需要高準確度且低單價的咖啡原豆挑選機，基於此需求，本研究與咖啡小農合作，由咖啡小農提供咖啡原豆的正負樣本，研究團隊設計視覺檢測模組，希望能朝向設計出滿足實際需求的咖啡原豆檢測設備邁出第一步。由於製作一部咖啡原豆挑選機的技術成份涵蓋：機構模組、機電控制模組以及視覺檢測模組，橫跨多種專門技術領域，團隊成員是資工系師生，對機構設計以及機電控制較不熟悉，因此將重點先擺在視覺檢測模組的研發上。

隨著機器學習技術的進步，以及深度學習的蓬勃發展，本研究採用深度學習技術，以機器學習中的監督式學習方式來訓練深度類神經網路，進行咖啡原豆檢測分類。本系統將咖啡原豆分成三類：正常豆、病蛀豆(病豆與蟲蛀豆)以及破碎豆，其中病蛀豆與破碎豆都屬於壞豆(即 NG 豆)。相較於只分正常豆與 NG 豆，本系統提供 NG 豆的來源細分類，讓咖啡農更能掌握自己豆子的瑕疵樣態，進而能擬定田間管理的精進方向，極為實用。期望本系統的研發能提升咖啡原豆篩選的效能，為提升台灣的咖啡產業發展貢獻一份心力。

本論文共分成四個章節，第一節介紹研究動機、背景及研究目的，第二節介紹系統架構，第三節介紹本研究如何進行咖啡豆偵測以及利用深度類神經網路實行咖啡豆分類，第四節為實驗說明並進行系統效能分析，第五節則對本研究做出結論及未來發展方向的描述。

2. 系統架構

本研究開發一個能區分咖啡原豆是好豆還是 NG 豆(再細分成病蛀豆與破碎豆)的分類系統，系統的運作流程如圖一所示，分為四個步驟：

1. 先檢測相機是否連接，如果成功連接則進入 UI 畫面，點擊開始，進入檢測畫面。
2. 將咖啡豆至於一張白紙上，以拖拉方式移動。
3. 攝影機取得連續影格，然後針對每張影格用影像二值化初步分割出咖啡原豆位置，再利用型態學的膨脹與侵蝕運算進行影像處理，分割出咖啡豆並框出其對應之 Bounding Box。
4. 將 Bounding Box 內的影像丟入事先訓練好的 VGG16[1]模型進行分類，並將分類結果顯示出來。



圖一、系統流程圖

3. 研究方法

3.1 咖啡豆偵測

本研究使用攝影機所取得之連續影格，並針對各個影格分別進行咖啡豆偵測。首先利用影像處理中大津二值化(OTSU)[2]初步分割出咖啡豆，再利用型態影像處理中的閉合運算(closing operator)，分割出較為完整的咖啡豆且框出其對應之外接框(bounding box)。

3.2 咖啡豆分類

系統將框出的咖啡豆，輸入深度類神經網路進行分類，本研究採用在影像分類與偵測中常被使用的網路模型 VGG16 作為分類模組的神經網路主幹，並使用遷移學習(Transfer Learning)技術[3]以提升特徵提取效果與縮短訓練的時間。

3.2.1 VGG16 模型

VGG 是英國牛津大學 Visual Geometry Group 的縮寫，主要貢獻是使用更多的隱藏層，能有效提高物件分類準確率。VGG16 結構共有 16 層，由 13 個卷積層及 3 個全連接層組合而成，結構圖如圖二所示。



圖二、VGG16 架構圖[1]

3.2.2 遷移學習(Transfer Learning)

為了提升特徵提取效果並縮短網路訓練時間，本研究使用以 ImageNet 預先訓練過的 VGG16 網路權重為起始權重來進行遷移學習。遷移學習一般做法是保留網路模型中捲積層的權重，並在其後加入新的全連接層進行，再利用新的任務訓練資料集進行訓練。而參數更動方式又能分為兩種不同方式，一種是在訓練過程中僅更動後方全連接層的權重，不會對特徵提取層做權重調整；另一種則會對全連接層與特徵提取層的權重都進行調整。後者作法適合當新任務有足夠多訓練樣本的情況使用，由於本研究所收集的訓練樣本數不夠多，所以採用前一種作法。

3.2.3 系統建構

本研究使用 VGG16 中預訓練過的卷積層權重進行初始化，並於後方加入 3 層新的全連接層(如圖三)，並使用咖啡豆訓練影像進行分類任務訓練，訓練過程透過倒傳遞算法調整網路中全連接層的參數。

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
Flatten (Flatten)	(None, 25888)	0
mydense (Dense)	(None, 512)	12845568
my_dense (Dense)	(None, 256)	131328
mydense (Dense)	(None, 3)	771

圖三、修改後之 VGG16

4. 實驗結果

4.1 實驗環境

本研究利用 Keras 套件建構機器學習模型，並以 Basler 相機(型號：acA1300-30uc)所拍攝的咖啡豆作為訓練樣本與測試樣本。取像架構如圖四所示，圖五為本研究所使用的相機。



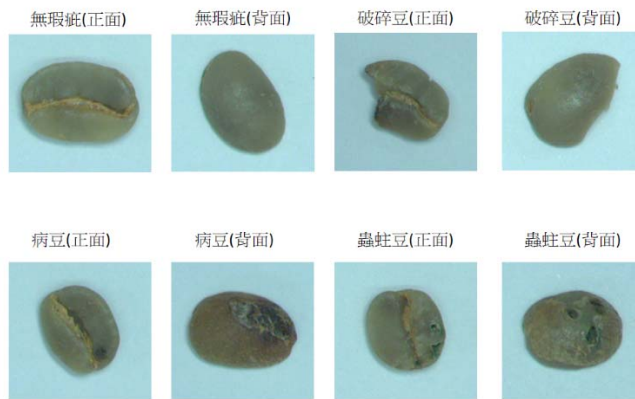
圖四、取樣架構



圖五、本研究所使用的相機

4.2 資料集創建與配置

本研究所拍攝的咖啡原豆影像資料集共包含 4 種類型，分別為：正常豆、破碎豆、蟲蛀豆與病豆，如圖六所示。本研究發現病豆和蟲蛀豆特徵有多處極為相似，所以選擇將兩者合為一類，標示為：病蛀豆。雖然咖啡豆皆有正反二面，但在訓練過程中發現，有沒有將咖啡豆正反二面分開訓練，對分類效能並沒有甚麼影響，考慮到實際應用上的方便性，本研究在訓練階段不對咖啡豆區分正面還是反面。咖啡原豆的影像資料集共有 5724 張影像，總共分成 3 類，格式為 PNG 檔。圖六為正常豆、蟲蛀豆、病豆及破碎豆的樣本豆的正面與反面影像。至於實驗所使用的訓練集、驗證集及測試集的配置則如圖七所示，以訓練集進行分類模型訓練，並以驗證集驗證模型當下效能，最終以測試集進行最後效能評估。



圖六、三類咖啡原豆影像樣本



圖七、資料集配置示意圖

4.3 實驗結果

實驗設計部分，本研究測試了三種不同的模型架構並進行分類效能比較，第一種模型乃採用文獻[4]所提出的咖啡豆分類模型，第二種模型乃採用 Google 所提出的 Inception-V3 模型[5]，第三種模型乃採用 VGG16 模型。表一是文獻[4]模型的測試結果，從結果發現此模型的分類效能並不佳，表示該模型並無法學習到具高分辨能力的特徵，導致模型無法精確分辨各類咖啡原豆。表二是 Inception-V3 模型的測試結果，從結果發現此模型的分類效能比起文獻[4]模型有大幅提升，表示該模型能透過學習找到具不錯辨能力的特徵，使其平均分類 accuracy 提升到 0.77。然而對於實際運用而言，這樣的 accuracy 仍然遠遠不夠高，從測試結果觀察，此模型對於破碎豆的 precision 達到 1，但是對應的 recall 則只有 0.47，表示該模型對於判別破碎豆的偽陰性率偏高；相反地，此模型對於病蛀豆的 recall 達到 1，但是對應的 precision 則只有 0.62，表示該模型對於判別病蛀豆的偽陽性率是偏高的。表三是 VGG16 模型的測試結果，從結果發現此模型的分類效能最佳，平均分類 accuracy 可到 0.99，表示 VGG16 模型能透過學習找到具高分辨能力的特徵，不論是正常豆、病蛀豆還是破碎豆，都能有極高的分辨能力。

表一、文獻[4]模型的分類效能

	precision	recall	F1
破碎豆	0.92	0.55	0.69
病蛀豆	0.88	0.35	0.50
正常豆	0.50	1.00	0.67
accuracy	0.63		

表二、Inception-V3 的分類效能

	precision	recall	F1
破碎豆	1.00	0.47	0.64
病蛀豆	0.62	1.00	0.77
正常豆	0.89	0.83	0.86
accuracy	0.77		

表三、VGG16 模型的分類效能

	precision	recall	F1
破碎豆	1.00	0.99	1.00
病蛀豆	0.97	1.00	0.98
正常豆	1.00	0.97	0.99
accuracy	0.99		

5. 結論與未來展望

本研究經過實作三不同模型，以相同訓練集、驗證集與測試集進行模型準確率測試，發覺採用 VGG16 模型當成主架構，再透過轉移學習方式所訓練出來的模型具有最佳的分類性能，準確率高達 99%，已經具有實用潛力，未來將結合機構設計與電控系統，讓此分類模型能進行整機測試，為咖啡產業貢獻一份心力。

6. 參考文獻

- [1] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556, 2014.
- [2] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Biomed. Eng., BME-9, pp. 63-66, 1979.
- [3] S. J. Pan and Q. Yang, "A survey on transfer learning," IEEE Trans. Knowl. Data Eng., Vol. 22, No. 10, pp. 1345-1359, Oct. 2010.
- [4] C. Pinto, J. Furukawa, H. Fukai and S. Tamura, "Classification of Green coffee bean images based on defect types using convolutional neural network (CNN)," Procs. of 2017 International Conference on Advanced Informatics Concepts Theory and Applications (ICAICTA), pp. 173-177, 2017.
- [5] Xia, X., Xu, C., Nan, B., "Inception-v3 for flower classification," Proc. Int. Conf. Image, Vision and Computing (ICIVC), Chengdu, China, June 2017, pp. 783-787.