



◎ 黃能富

電腦也會 選咖啡豆

一句「電腦也會選土豆」曾經是 80 年代的經典廣告台詞，如今透過人工智慧及影像分析技術，電腦已經能應用在各行各業中，且其準確度及速度也可超越肉眼的辨識了。



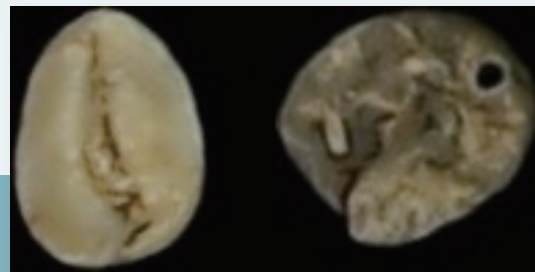
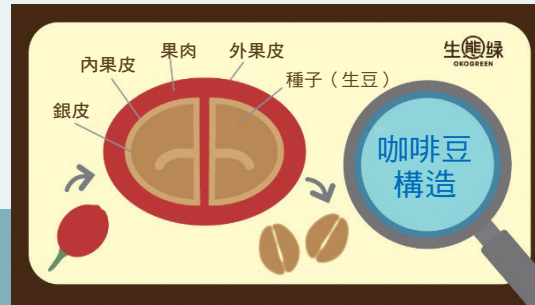
挑選目的

咖啡是流行的飲料之一，也是重要的經濟作物。咖啡樹的種植遍布全球七十餘個國家，其果實最初呈綠色，成熟後逐漸轉為紅色、黃色或橘色。市面上看到已烘焙好的咖啡豆，其實是咖啡果實內部的種子，通常每一顆咖啡果實內有兩粒種子，種子由外而內的構造分別是外果皮、果肉、果膠層、內果皮（種子外殼）、銀皮、種子。咖啡果透過日曬或水洗的加工處理，就是為了去除這些外部的構造，以取得內部的咖啡種子。

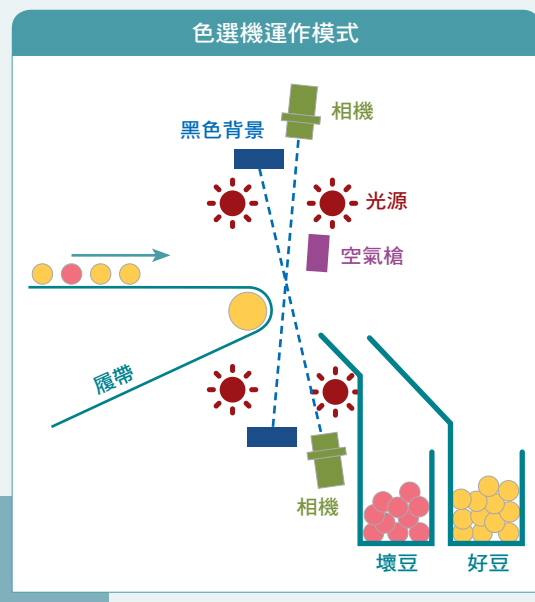
初經加工處理後的咖啡豆稱為生豆，生豆需經過篩選才能夠進入烘焙機烘烤。而挑選的標準是由世界上兩大精品咖啡協會—美國精品咖啡協會及歐洲精品咖啡協會—共同擬定的，主要分為 5 級標準，即精選級（specialty grade）、高品質（premium grade）、基準級（exchange grade）、低於標準級（below standard grade）、級外（off grade）。

其分類標準是 300 公克中的完全瑕疵指標數量，精選級是 5 個以下且無主要瑕疵、高品質是 8 個以下、基準級是 9 至 23 個、低於標準級是 24 至 86 個、級外是超過 86 個以上的完全瑕疵指標。透過這些標準把瑕疵豆一一挑出，留下的就是可烘焙的生豆。

生豆若未經過挑選就烘焙，在烘焙過程中會因受熱不均勻，且不健康、有害的物質如真菌、毒素等會汙染整體的豆子，俗話說「一粒老鼠屎壞了一鍋粥」，一顆瑕疵豆不僅會壞了一整杯咖啡，甚至影響飲用人的身體健康。現今已有許多自動化機械可進行大米、雜糧、茶葉等物料的挑選，原理是當物料經由履帶送至相機處時，可透過光學技術如紅外線或 X 光技術等即時

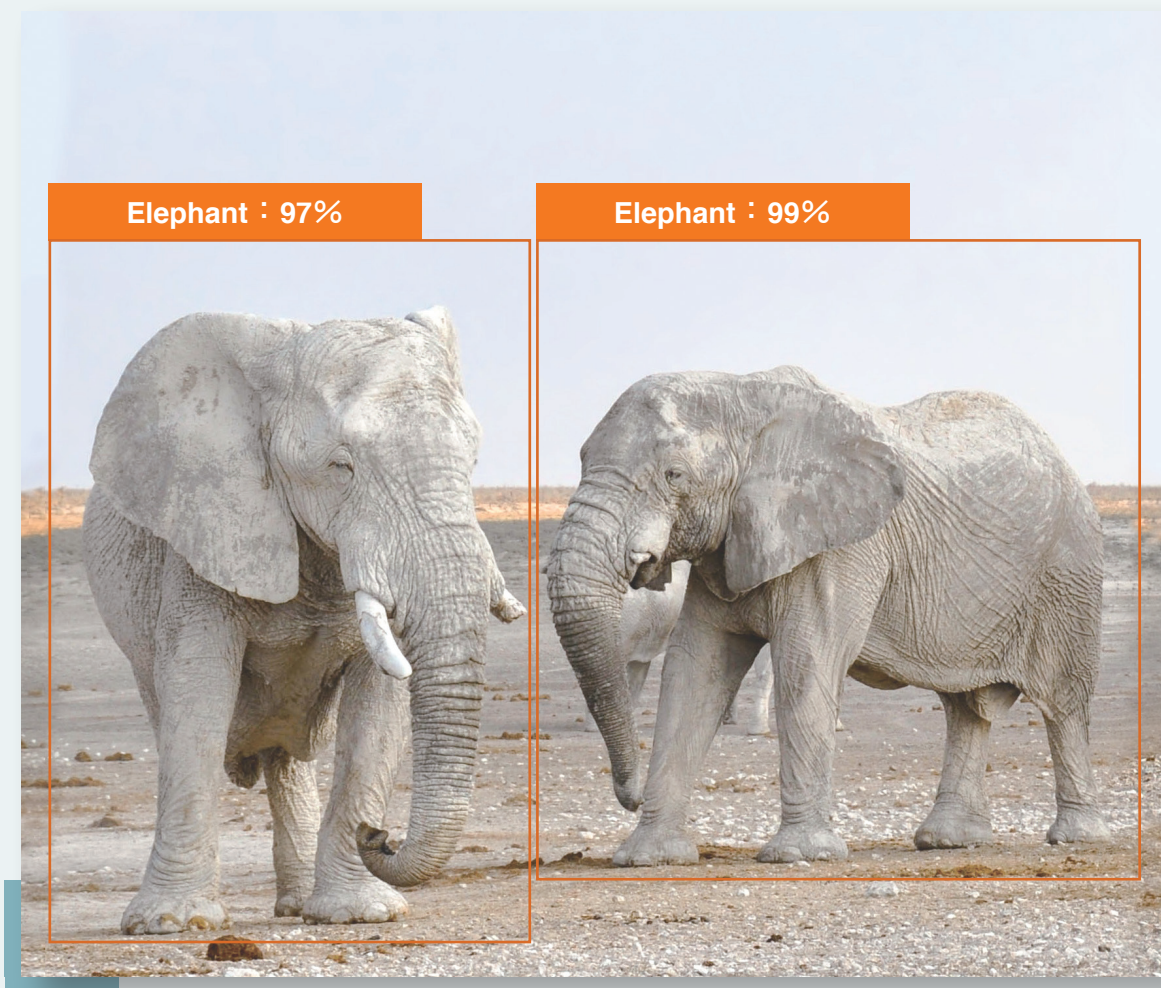


好咖啡豆（左）和壞咖啡豆（右）。





現今已有許多自動化機械可進行物料的挑選，
原理是透過光學技術即時分析物料的颜色及光學特性以判斷其好壞。



使用物件偵測辨識出畫面中物件的種類，並給出信心指數。

分析物料的颜色及光學特性以判斷其好壞，再經由空氣噴嘴把優劣品分開，完成自動化挑選物料的動作。這方法相較於人工挑選更為省工省時且效率高。

深度學習

早在 1956 年人工智慧這個名詞就已出現，但因早期電腦的效能有限，所以只能解決一些簡單的問題，無法解決現實生活的問題。因此理論雖然一直都在發展，主題卻一直被限制住，無法蓬勃發展。

近年隨著 CPU 效能提升以及圖形處理器 (GPU) 技術的發展，透過 GPU 平行運算的技術訓練學習模型，可以使得處理時間大幅度縮短。最著名的深度學習案例是手寫數字辨識，透過所謂的卷積神經網路 (convolutional neural networks, CNN)，可辨別阿拉伯數字 0 至 9 的手寫圖片。過程中需透過資料蒐集、資料預處理、資料強化等繁瑣但重要的工作，達到充足且完善的訓練資料，才能夠作為類神經網路的輸入，中間再透過訓練與學習產生辨識手寫阿拉伯數字的模型。



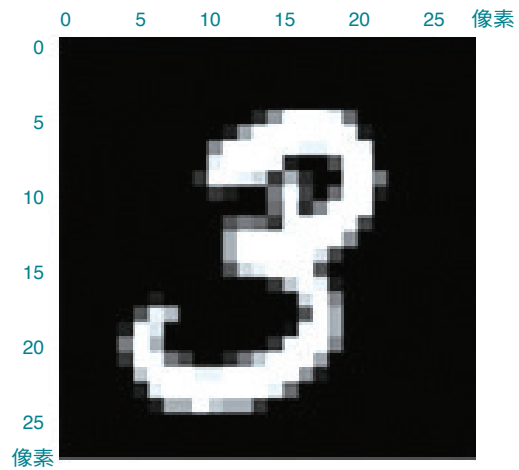
拍攝咖啡豆的環境設置

藉由 CNN 技術於圖形及影像辨識領域解決問題，包含本文所提及咖啡豆優劣的辨識、停車場的車牌辨識系統等，都可以透過 CNN 更有效率地解決問題。

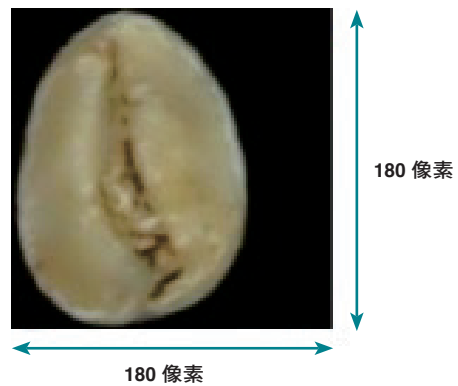
資料蒐集

在深度學習中，若要訓練出一個不錯的模型，大量且有效的資料是非常重要的。在訓練咖啡豆辨識模型的時候，好的咖啡豆及壞的咖啡豆都需進行資料蒐集，且圖片數目不可過於懸殊。舉例來說，若好咖啡豆的資料占 80%，壞咖啡豆的資料只有 20%，則電腦在每次預測時只會猜測是好咖啡豆，因為不論結果如何，準確度都有不錯的 80%。在這樣不平衡的資訊下，容易造成神經網路單方面地猜測較多數量資料的一方。為了避免這樣偏頗的現象發生，必須盡量平衡好壞咖啡豆的比例。

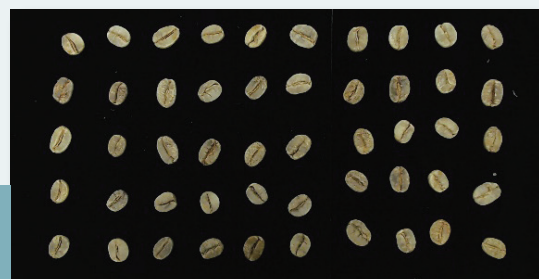
手寫辨識圖片



手寫數字 3 範例圖：藉由大量資料訓練，可以讓模型學習到每個手寫數字的特徵，達到辨識的目的。



固定大小的咖啡豆



咖啡豆拍攝結果



資料預處理

在資料預處理中，需先進行資料分割，當然把拍攝圖片切割的工作不可能交給人手動切割。Open source computer vision library (OpenCV) 是一個跨平台的電腦視覺函式庫，可用於開發即時的圖像處理、電腦視覺，以及圖形辨識程式。

透過 OpenCV 中所提供的圖像分割技術，程式會自動地把一張有多顆咖啡豆的圖片切成一顆一張的圖片。然而，咖啡豆的大小本來就長寬不一，切割出來的圖片也會長寬不一致。因此需要透過圖片調整，把所有的圖片重新縮放，但是不可更改原先圖片的比例，否則圖片一失真，資料就無效。透過圖片調整可把所有的圖片縮放為 180×180 像素的大小。

最後，為了讓每張圖片可以充分被學習，需要透過去除背景的方式把圖像其餘的像素去除。儘管圖中所看到的背景是黑的，但並非是色碼表所謂的黑色 (RGB (0, 0, 0), #000000)，即使以人眼看起來都是黑色的，對電腦來說卻是黑的不一樣。因此需要透過去除背景的方式把多餘的雜訊去除，以提升資料訓練的品質與準確度。

背景去除的方式，通用的方法有兩種：邊緣檢測、顏色檢測。前者是把圖片轉為灰階後，再透過高斯濾波模糊化的方式去除影像雜訊，最後透過分析像素的梯度強度判斷是否是邊緣，設定上閾值 (upper threshold) 與下閾值 (lower threshold)，把之間的數值去除。後者則透過顏色的偵測設定閾值，把範圍內接近黑色的部分去除。

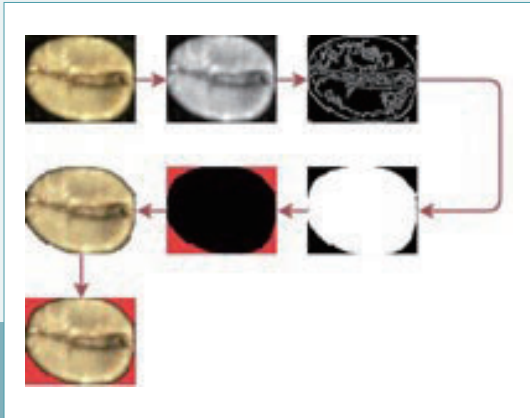
資料增強

若原始資料集比較小，就無法滿足網路模型的訓練，進而影響模型的準確度及



藉由高斯濾波，並調整模糊程度，去除影像中的雜訊。

效能，甚至可能導致過度擬合的情狀發生。過度擬合，顧名思義就是過度學習訓練的資料時，電腦會無法順利地預測或分辨不在訓練資料內的其他資料。舉例來說，當準備某一門課的考試時，練習了太多相同且類似的題目，但是當考試的題目作了一點變化時，



利用邊緣檢測的方式去除背景

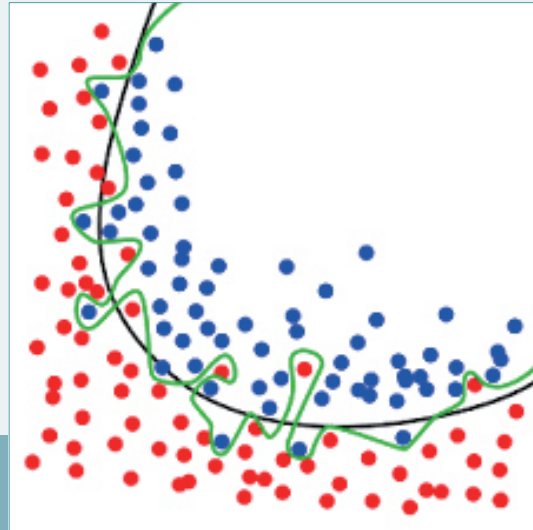
會發現卻無法應付考卷上的題目，這就是機器學習上的過度擬合的情況。

然而當一張照片經過旋轉、調整大小、比例尺寸，或者改變亮度色溫、翻轉等處理後，雖然人眼仍能辨識出來那是相同的相片，但是對機器而言那可是完全不同的新圖像。因此，資料增強就是把資料集內的圖片予以修改變形，創造出更多的圖片讓神經網路學習，彌補原先資料量不足的困擾。

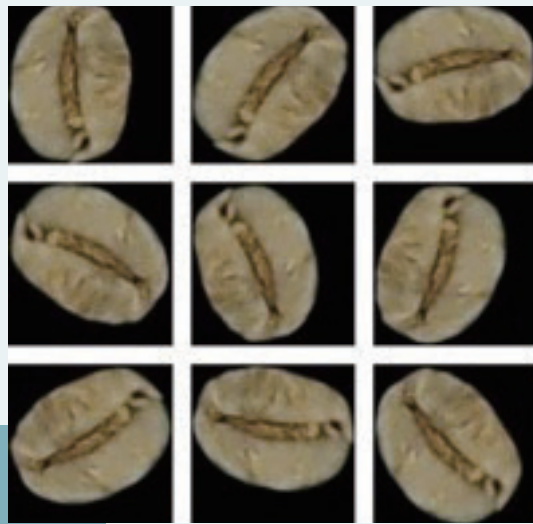
不過對於咖啡豆圖片增強的選擇，不能夠透過調整大小或改變原圖的亮度色溫來增強資料，如此會破壞圖片原有的特徵值。因此，咖啡豆圖片可以透過旋轉及翻轉的方式做資料的增強。例如可以每 40 度的翻轉及上下左右的翻轉方式來做資料增強，最後能夠產生 36 倍的圖片數據。

卷積神經網路

卷積神經網路是一種多層的神經網路，其中包含卷積層、池化層、全連接層，這一結構使得卷積神經網路在圖像及語音辨識方面可以有更好的能力。

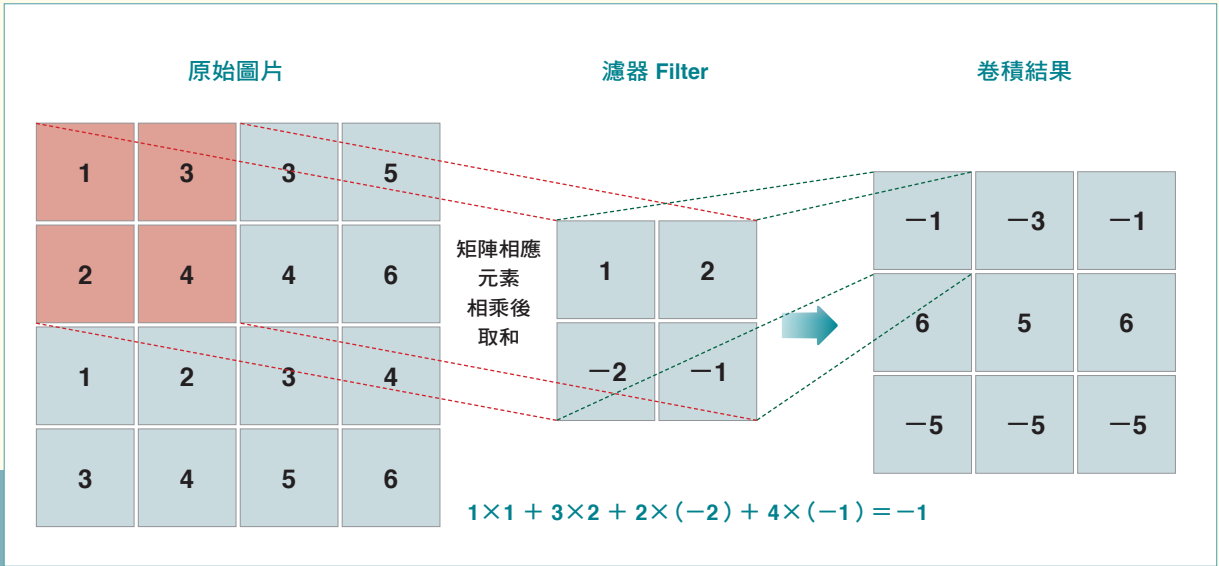


黑線表示理想分類情況；綠線表示過度擬合分類情況。

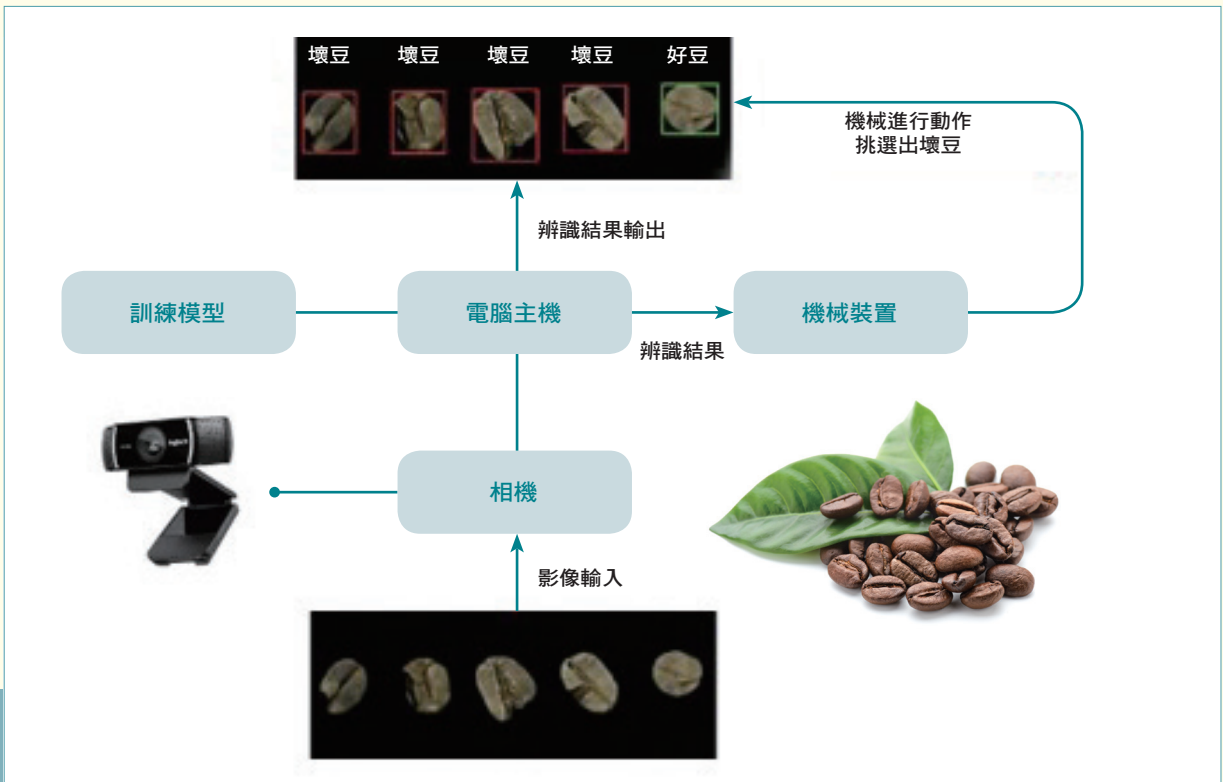


透過旋轉 40 度得到的圖片資料

卷積層中把原始圖片與特定的濾器進行卷積運算，目的在於幫助使用者萃取出所訓練圖片中的特徵，例如形狀、顏色等都是可以透過卷積層強化及萃取的特徵。



卷積的運算方法



咖啡豆辨識系統的架構

透過連結相機與電腦以及訓練出的模型，
可即時把辨識結果顯示在螢幕上，
並利用機器把有瑕疵的咖啡生豆挑選出來。

瑕疵豆標準

主要瑕疵		次要瑕疵	
項目	出現次數等於一個完全瑕疵指標	項目	出現次數等於一個完全瑕疵指標
全黑色	1	帶殼豆	2 ~ 3
全發酵	1	豆殼 / 果皮	2 ~ 3
莢 / 果實	1	破裂 / 屑片	5
大石頭	2	蟲害	2 ~ 5
中石頭	5	局部黑	2 ~ 3
大枝條	2	局部發酵	2 ~ 3
中枝條	5	異物	5
		貝殼豆	5
		小石頭	1
		小枝條	1
		水傷	2 ~ 5

池化層中透過輸入圖層劃分若干個區域，針對每個子區域進行運算，而運算分為最大池化及平均池化。最大池化的方式就是把區域的數值取出最大的數值作為該區段的代表數值，平均池化則是算出該區的平均數值。透過池化層可以不斷縮減資料的空間大小，降低參數的數量與計算量，以避免過度擬合的情況發生，並保持影像中的特徵。

全連接層把之前全部結果平坦化，依前面所學的特徵組合且從中分類，並完成卷積層網路的訓練。

咖啡豆辨識

把原始資料透過資料切割、資料預處理、資料增強等技術的運用，可以得到適合的圖片進行訓練。再藉由卷積網路層的訓練，過程中神經網路會把優良咖啡豆及瑕疵咖啡豆所具有的特徵抓取出來作為模型

的判別依據，訓練的結果就可以得到完善的訓練模型。

目前研究人員已可透過連結相機與電腦以及訓練出的模型，即時把辨識結果顯示在螢幕上，告訴使用者生咖啡豆的好壞。最後利用機器把有瑕疵的咖啡生豆挑選出來，使得挑選過後的咖啡豆符合美國精品咖啡協會所規範的精品咖啡標準，進而達成深度學習應用於咖啡豆辨識技術的目的。

上述技術的運用使得挑選咖啡豆的人力及時間成本都得以降低，並增加農民收入，也促使精品咖啡業更蓬勃發展。這項技術不僅可以辨識咖啡豆，未來也可以應用在多項作物上。

黃能富
清華大學資訊工程系