

第十八屆旺宏科學獎

成果報告書

參賽編號：SA18-157

作品名稱：以雙卷積神經網路模型辨識固定及形變之垃圾物件及自動分類裝置之研究

姓名：李丞揚

關鍵字：垃圾分類、深度學習、卷積神經網路

摘要

垃圾分類一直是環保重要的議題，本研究提出一個具備辨識垃圾物件分類的裝置，透過深度學習卷積神經網路(Convolutional Neural Network, CNN)來進行分類辨識演算，特別是直接透過影像辨識垃圾物件外觀，不論是固定外型垃圾物件類別，甚或是具有形變的垃圾物件類別，辨識裝置對於經過形變的垃圾物件具有良好的辨識效果，再透過系統整合將垃圾類別資訊傳遞到移動翻蓋機構，能正確自動移動到該類別垃圾桶進行翻蓋丟擲垃圾的動作。

之前在創意說明書的階段，本系統採用「單一物件檢測(object detection)卷積神經網路模型」來作為研究的核心技術，經測試，已能有效分辨垃圾物件，針對實際垃圾物件進行辨識，在固定外型的垃圾物件中(包括廢燈泡類、廢電池類以及玻璃瓶類)辨識率達 98%；而在包括形變的垃圾物件中，包括鐵鋁罐類、廢紙容器類、塑膠瓶類、紙盒類、廢紙類、辨識率仍高達 95%。

不過採用此方法最重要的缺點，就是需要收集大量垃圾物件的數位資料作為神經網路訓練的素材，然後再經過幾萬次的訓練迭代，才能夠讓模型具備良好的辨識能力，這對未來落實的過程應該有不少的困難度，因此，本研究針對這個問題進行改良，目前採用「雙卷積神經網路模型」來針對垃圾物件進行辨識，第一個神經網路模型只針對物件進行辨識(不分類)，第二個神經網路再針對第一個辨識到物件的影像進行分類辨識，這個架構的好處是對第一個物件偵測網路模型而言，只需要辨識影像中是否存在一個或多個物件，而且我們可以累積物件偵測的能力，將這個神經網路的模型應用在其他需要偵測物件的應用中，例如除了垃圾物件以外，我們可以偵測商品物件、車輛物件、人形物件等等；而第二個分類辨識的神經網路只負責針對第一個神經網路傳遞過來的物件內容進行分類辨識，在分類的神經網路中訓練的時間遠低於單一物件檢測神經網路的訓練時間，而且如果我們收集到的新垃圾物件(或新商品)資訊有限，仍然有機會在很少的影像資訊下能夠成功的辨識物件照片中屬於哪一類，因此本研究新推出的深度學習影像辨識架構應用在固定或形變的垃圾物件辨識中，將有助於降低垃圾物件數位化資訊的收集以及神經網路模型訓練的時間，經測試，新架構在實際丟擲測試中，在固定外型的垃圾物件辨識率約為 98.3%，而在形變垃圾物件的辨識率約為 95.1%，效果均有提升。

此外針對辨識系統對光源的強弱影響也進一步進行測試(本次新增)，參考 CNS 國家照度標準，嘗試針對辨識系統模擬戶外狀況、室內狀況以及地下室狀況三種垃圾桶常見的擺放場所，對垃圾物件辨識系統對光源的敏感程度再作進一步實驗分析辨識系統之強健性，實驗證明辨識系統在照度 10,000 流明至 10 流明左右仍可正確辨識垃圾物件。

透過本研究證明利用卷積神經網路直接辨識垃圾物件具備良好的辨識率，若進一步發展有機會優化網路效能及辨識程式運作效率，大幅度降低辨識盒的成本，擴大提升垃圾桶具備垃圾物件辨識的能力，並進一步降低硬體成本，讓智能垃圾桶普及，並且可以結合物聯網、雲端技術及智動技術，擴大垃圾分類自動化的可行性，為地球環保盡一份貢獻。

壹、 研究動機

有一次在國家地理雜誌的網站上看到一則「吸管怎麼會跑進海龜的鼻子裡？」的文章報導與影片[11]，讓我非常的震驚！所以就想找跟環保議題相關的方向來研究；在查詢相關文獻的過程中，看到一篇國內的學者：范宏偉[1]，以苗栗縣國民中學學生對於學校垃圾分類的研究發現，高達 78.9%的中學生，是完全沒有垃圾分類的習慣；此外，根據行政院環保署的統計資料顯示，我國歷年的垃圾回收率雖然逐年提高，但每年的成長率增長的幅度也逐漸趨於緩和，如下圖，所以顯然在垃圾回收的部分也呈現瓶頸，足見垃圾分類的重要性。



圖一：台灣歷年垃圾回收率與垃圾回收增長率圖，環保署(2001-2015)

而全球科技發展中，資訊科技突飛猛進，包括物聯網、AI 人工智慧、大數據分析等新創技術的蓬勃發展，從 AlphaGo 完勝棋王柯潔到 AI 人臉辨識超越人類的辨識能力等，但相對於大家平常所見，垃圾桶始終就是傳統置放垃圾的桶子，顯然有很大改進的空間；資訊科技促進各行各業改變甚至翻轉或讓某些行業消失，與其期待人們的公德心與對垃圾分類的認知來達到垃圾正確分類的目的，不如善用資訊科技的手段來幫助人們進行垃圾分類的協助，也相信這是全球共同的題目，因此本研究希望藉由深度學習的技術來進行垃圾分類辨識，協助一般人對垃圾分類的落實與提升有效性，對科技運用於垃圾分類的領域盡一份貢獻。

貳、 研究目的

思考透過影像直接辨識垃圾物件的類別，最少會有下列幾項困難點需要被克服及突破：

一、垃圾物件的多樣性：

由於垃圾物件是由商品的包裝物轉變為垃圾，商品的包裝物每天都有新的設計及不同的變化，因地因時隨時都在增加，因此垃圾物件充滿了不同的多樣性，垃圾物件的外觀其豐富的程度及辨識的難度均超過現有手寫文字辨識、車牌辨識、撲克牌辨識等固定範圍的辨識。

二、垃圾物件的相似性：

商品的包裝經常會出現不同的包裝材質但風格印刷設計雷同的包裝物，而變成垃圾物件後就會有相似的特徵存在於不同材質的包裝中，例如舒跑這個商品，有紙盒鋁箔包裝也有塑膠瓶包裝，雖然兩者的包裝尺寸及材質都不相同，但上面的印刷圖案卻都有雷同的相同圖案及商標特徵，因此垃圾物件的相似性有別於其他物件的辨識，存在相似的比例很高。

三、垃圾物件的形變：

垃圾物件辨識比商品物件辨識更加困難，商品為了銷售會保護外觀外型以增加購買的可能性，而垃圾物件相反，因為是拋棄的垃圾，因此對於垃圾不會愛惜保護，會對垃圾壓扁、丟棄、踩踏，因此實務中的垃圾物件經常是不可預期變形的外觀，增加辨識上的困難度。

本研究目的如下：

- 一、使用深度學習卷積神經網路作為垃圾類別辨識的核心技術基礎。
- 二、具備解決垃圾物件的多樣性、相似性與形變的特性，能有良好且強健的辨識能力。
- 三、比較單一物件偵測卷積神經網路與雙卷積神經網路對固定及形變垃圾物件的辨識能力及優缺點比較。
- 四、設計一個自動移動翻蓋的機構，驗證辨識盒能整合機構，達到自動分類的目的。
- 五、針對辨識系統進行環境光源測試，以了解對於環境光源變化範圍的辨識能力。

參、 研究過程

一、 現有技術及方法及探討

在全國博碩士論文網站上進行檢索，尋找過去研究者對於垃圾及影像辨識方面的研究發現：透過影像直接辨識垃圾物件的研究確實很少，不論以關鍵字「垃圾」and 「辨識」或「垃圾」and 「影像辨識」或是「垃圾」and 「卷積神經網路」或「垃圾」and 「分類」and 「辨識」的搜尋，在已授權全文的 47 萬 6 千多筆以及書目及摘要的 112 萬 1 千多筆中均沒有查到相關類似的研究，如下圖：



編號	檢索歷史	檢索推薦	筆數
<input type="checkbox"/> 4	"垃圾 and 分類 and 辨識"(精準)		0
<input type="checkbox"/> 3	"垃圾 and 卷積神經網路".ti(精準)		0
<input type="checkbox"/> 2	"垃圾 and 影像辨識".ti(精準)		0
<input type="checkbox"/> 1	"垃圾 and 辨識"(精準)		0

簡易查詢 | 進階查詢 | 論文瀏覽 | 熱門排行 | 我的研究室

本系統(Web7)共收集：論文已授權全文：476928 筆、書目與摘要：1121438 筆

圖二、全國博碩士論文網站上查詢「垃圾」與「辨識」關鍵字的結果

此外，針對目前智能化垃圾分類設備的作法進行資料收集，德國 Sielaff 公司[12]發表了 Leergut 回收桶，可以針對垃圾上印刷的資源回收標誌進行辨識，當無法辨識到回收標誌時，回收桶會將垃圾物件退出，其主要的辨識技術是識別垃圾上的回收標誌，另瑞典 Eskilstuna 公司[13]所設計的自動垃圾分類廠則是透過辨識垃圾袋的顏色進行分類，而韓國的 SuperBin 公司[14]發表的 AI 回收桶 Nephron 主要是針對瓶狀類進行辨識，可針對寶特瓶、鐵鋁罐及玻璃瓶三類的垃圾物件進行回收，在辨識技術上算是比較接近本研究的方向，不過仍然沒有針對形變的垃圾進行辨識，國內的大豐回收公司[15]推出的 ZeroZero 無人回收站則是以半自動回收的方式，其辨識的核心技術主要是以瓶罐類透過讀取垃圾物件上的條碼進行辨識；清華大學[21]推出 AI 資源回收機器人(智能垃圾桶)，可針對鐵鋁罐、寶特瓶、玻璃瓶及紙盒四類垃圾物件進行辨識及分類，唯丟擲垃圾時需按鈕啟動辨識機制，辨識時間需花 3-5 秒進行辨識；將收集的這幾家公司的辨識技術進行整理，與本研究的方向進行比較，主要的共通點都是透過影像辨識，但差異點則在於直接辨識垃圾物件以及直接辨識形變後的垃圾物件，整理如下表：

表一：智能化垃圾分類設備的主要辨識技術整理比較

年度	國家	公司/組織	名稱	辨識標的	形變垃圾	辨識時間
2016	德國	Sielaff	Leergut 回收桶	回收標誌	不具備	約 1-2 秒
2107	瑞典	Eskilstuna	自動垃圾分類廠	垃圾袋顏色	不具備	約 1 秒
2018	韓國	SuperBin	AI 回收 Nephron	垃圾物件	不具備	不詳
2018	台灣	大豐回收	zero zero 無人回收	條碼	不具備	不詳
2019	台灣	清華大學	AI 資源回收機器人	垃圾物件	不詳	3-5 秒(註)
2019	台灣	本研究	辨識垃圾物件分類裝置	垃圾物件	具備	約 0.5 秒

註：3-5 秒資訊來自新聞報導中表示的辨識時間

此外，不僅直接透過影像辨識垃圾物件的相關研究發表的論文不多，針對形變物件的相關研究及探討也較少，查找國內有關形變辨識的相關研究包括：龔純瑩[8]，以普通空間圖形與富氏空間特徵結合四組雙向性組合式記憶矩陣，再透過最小均方法 (LSM) 進行學習訓練，以改變記憶矩陣內的元素值運用在程式辨識引表機的大寫英文字母與阿拉伯數字；鄭萬旗[9]，針對手寫數字的變形待測物利用每個手寫數字的樣本圖與待測標的進行比較，依照形變量來衡量匹配程度，對手寫數字進行分類；李丞揚、牟代君、葉建德、***[10]，利用深度學習卷積神經網路來辨識垃圾物件並進行分類，比較本研究與相關文獻演算法及模型差異如下表：

表二：形變物辨識相關辨識技術整理比較

年度	研究者	辨識標的	種類	使用方法	模型類別
1991	龔純瑩	大寫英文字母與數字	36 類 (A-Z,0-9)	最小均方法	機器學習方法
2000	鄭萬旗	手寫數字	10 類(0-9)	以形變量來分類	依照變形程度
2019	李丞揚 牟代君 葉建德 ***	垃圾物件 (包含固定及形變)	3 類 (鐵鋁罐、廢紙容器類、塑膠瓶類)	利用單一物件辨識卷積神經網路進行垃圾物件影像辨識	深度學習-卷積神經網路
2019	本研究	垃圾物件 (包含固定及形變)	9 類 固定外型 3 類：廢燈泡類、廢電池類、玻璃瓶類 可變形外型 6 類：鐵鋁罐類、鋁箔包(廢紙容器類)、塑膠瓶類、紙盒類、廢紙類、紙杯類)	利用兩個卷積神經網路(物件偵測 + 物件分類)進行垃圾物件影像辨識	深度學習-卷積神經網路

而在視覺領域競賽的 ImageNet Large Scale Visual Recognition Competition (ILSVRC)在 2012 年由來自多倫多大學的學生[2]Alex Krizhevsky 設計了一個 Deep Learning 模型-AlexNet，在百萬

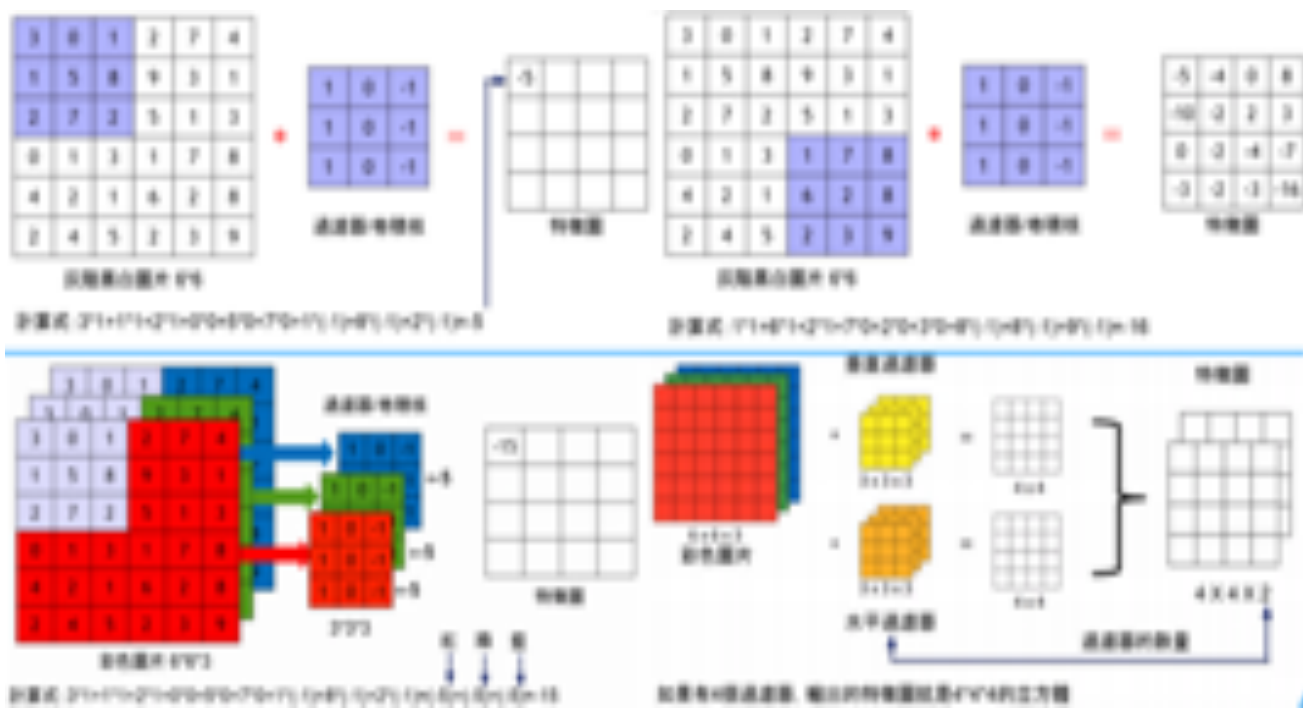
量級的 ImageNet 資料庫上大幅超越了傳統方法，奪下了當年度冠軍，top-5 的錯誤率降低到 16.4%，相較第二名成績的 26.2% 錯誤率有了顯著的提升，而隨後每一年的冠軍得主，都是採用卷積類神經網路的技術，尤其在 2015 年這年，來自微軟北京研究院的團隊 top-5 的錯誤率降低到 3.57%，這項成就超越了人類辨識的錯誤率，也證明了卷積類神經網路成為影像辨識的主流技術方法[3]，如下圖，因此本研究採用卷積類神經網路作為影像辨識的核心方法。



圖三:ILVRC 歷屆冠軍錯誤率與採用的主要方法圖(本研究整理)

這類的方法主要是應用卷積類神經網路的卷積運算，透過滑窗(Sliding Windows)方式將圖像與卷積核(kernel)或稱過濾器(filter)進行運算，卷積核可以設計成各種不同的強化特徵，例如強化垂直線條的垂直卷積核、強化水平線條的水平卷積核等等，簡單說明一下卷積運算的方式。

首先我們使用 $3 \times 3 \times 1$ 的卷積核去過濾一張 $6 \times 6 \times 1$ 的黑白圖片，以數字的大小來表示圖片的深淺，數字越大顏色越深，數字越小顏色越淺。我們會先把卷積核放在圖片的左上角去進行卷積運算，透過圖片與卷積核上對應到的數字相乘再相加，得到的結果會放在特徵圖(feature map)的左上角。我們可以設定滑窗移動的大小，稱作步長(Stride)，若設定步長為 1，接下來會把卷積核往右移一格進行運算，重複這些動作之後會得到特徵圖的第一列。當完成第一列後會往下移一格，同樣是從最左邊開始做運算，一樣重複這些動作之後，就可以得到一張完整的特徵圖，如下圖：



圖四:卷積運算示意圖(本研究整理)

而彩色圖片因為有 RGB 所以分成三層，不過運算方法與黑白圖片相同，首先是使用 3*3*3 的卷積核從圖片左上角開始運算，因為圖片有三層，所以我們會把三層得出來的結果去做相加，一樣放到特徵圖的左上角，重複這些運算同樣會得到運算後的特徵圖。卷積運算主要在於獲得物件的特徵，也是最新的物件偵測方法，我們將以卷積神經網路作為垃圾物件分類辨識的核心技術。

二、 研究架構

本研究定義下列的研究架構，概分為「垃圾數位化」、「深度學習」、「系統測試與分類機構整合」三個部分，研究架構圖如下圖：



圖五、研究架構圖

(一) 垃圾數位化：

垃圾物件數位化的目的是蒐集垃圾物件，然後透過數位影像的手段來達到垃圾物件數位化的目的，因此，我們透過三個步驟來進行，分別是「垃圾蒐集」、「垃圾整理」以及「拍攝照片」，以下分別針對個步驟進行說明：

1. 垃圾蒐集：

眾所周知，垃圾物件種類繁多，而且隨著新產品上市在包裝、印刷、外觀、材質等等均會隨著時間推陳出新，因此數量龐大、種類繁多，礙於研究有其時間限制以及資源有限，因此本研究先鎖定九類垃圾物件做為代表，分別是鐵鋁罐類、鋁箔包(原廢紙容器類)、塑膠瓶類、廢燈泡類、廢電池類、玻璃瓶類、紙盒類、廢紙類、紙杯類等九種，這九種類別已涵蓋可回收的資源項目中主要代表的垃圾類別，其型態與種類的多樣性也很複雜，足以作為代表垃圾物件分類議題探討中包含多樣性、相似性與形變的三種主要特徵的代表。

2. 垃圾整理：

針對收集的垃圾種類，其中除了廢燈泡、廢電池及玻璃瓶三類在一般垃圾型態下較少變形外，其餘六類一般人們對於這類垃圾的處理經常是不溫柔的，因此這類垃圾物件因為丟棄而造成形變也是非常普遍的，垃圾物件的整理，主要就是希望模擬接近垃圾物件被丟棄的樣貌，因此，對於這六類蒐集到的垃圾物件，會針對垃圾物件進行外力的處理，例如擠壓、揉捏、踩踏等等，讓垃圾物件的外型趨於接近一般我們常見的垃圾樣貌。

3. 拍攝照片：

為了進一步進行深度學習的採樣，本研究需要對垃圾物件蒐集大量的照片以作為訓練的基礎，卷積神經網路對於物件特徵在縮放及位移上的辨識能力會有良好的效果，但對於物件特徵的旋轉的識別能力較弱，因此，本研究針對垃圾物件進行拍攝時就需要蒐集相當多物件翻轉及不同角度的照片，讓神經網路訓練時能夠得到不同角度的特徵資料，本研究在創意說明階段一共完成 15,871 張照片，另外在成果報告階段重新整理垃圾數位化照片，將原廢紙容器類別依照分類的習慣修正成鋁箔包類，並將非鋁箔包的物件剔除並新增第九類紙杯類，整理後，最終完成九類共 20,337 張照片的垃圾數位化工作。

(二)深度學習：

深度學習的步驟按照依照「模型架構」、「標記作業」、「訓練作業」及「模型評價」四步驟，以下進一步說明：

1. 模型架構

本研究在創意說明階段使用：單一深度學習神經網路物件檢測(Object Detection)技術，來進行神經網路模型的訓練及透過模型來進行垃圾物件的辨識，經測試對於固定外型及變形垃圾物件，均能有良好的辨識準確率，但考量商品不斷推陳出新垃圾的種類日新月異，原採用的方法需要大量的垃圾數位化照片進行標記，標記的過程需要人工進行分類，除了需要投入大量的標記人工成本，在訓練的階段也需要經過十萬以上數量級別的迭代訓練及調教才能夠讓模型本身具備優良的辨識能力，物件相似性也需要經過不斷調教才能夠避免互相干擾，因此，一直在思考及尋找改進的解決之道；目前找到的方向是利用「雙卷積神經網路」(兩個網路)來進行辨識模型

的改進，並發表在本次的成果報告階段，相較於與創意說明階段的模型架構差異及優缺點，整理如下表說明：

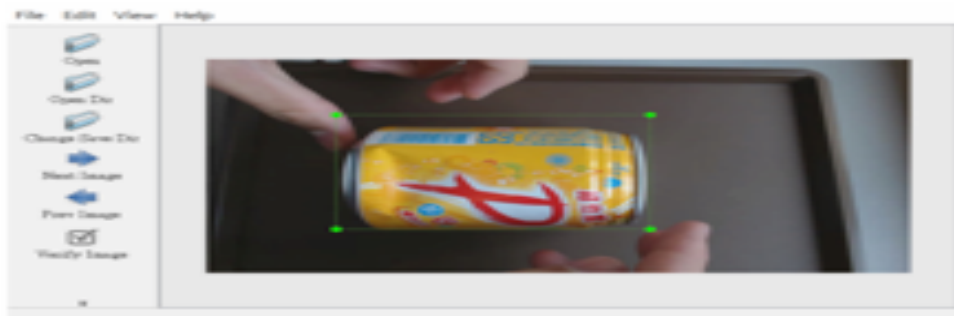
表三：新研究成果:網路模型架構差異比較表

項目	創意說明階段	成果報告階段	
模型架構	單一物件檢測卷積神經網路	雙卷積神經網路	
垃圾物件種類	8類(固定3類,可形變5類)	9類(固定3類,可形變6類)	
卷積神經網路-主幹	MobileNet+SSD	物件偵測網路	物件分類網路
		MobileNet+SSD	MobileNet
標記照片數量	15,871張分8類	20,337張分1類	20,337張分9類
標記難度	中	低	最低
訓練迭代次數	12萬次	10萬次	20萬次
loss	0.3~0.5	0.02~0.15	0.3~0.7
一次訓練時間	約84小時	約40小時	約6小時
增加物件時	需重新標記及訓練	幾乎不需要重標	需重新訓練
模型參數調教	越多類別越複雜	簡單	簡單
相似物件干擾性	較高	低	中
應用在其他物件	較低	較高	較低
整體運作效能	高(辨識時間約0.1秒)	低(辨識時間約0.5秒)	

本研究主要希望應用在垃圾物件辨識的場景應用，特別希望能夠降低硬體效能的需求，因此會特別考慮輕量級別的卷積神經網路作為辨識核心的骨幹，採用雙卷積神經網路模型架構，分成「物件偵測網路模型」以及「物件分類網路模型」兩個步驟，步驟一透過物件偵測網路模型進行偵測影片中的垃圾物件，再進行步驟二，透過物件分類網路模型進行物件分類，最終得到垃圾物件的類別。

2. 標記作業

本研究因為採用雙卷積神經網路的架構，因此，在物件偵測網路模型，仍須針對垃圾物件照片進行標記，本研究採用 PASCAL VOC 標記規範的格式進行標記，透過開源免費的標記軟體-LabelImg [16]來針對相片進行標記，如下圖；標記軟體可以將標記的內容輸出成 PASCAL VOC 規範的 xml 的檔案格式。



圖六：使用 LabelImg 軟體進行標記示意圖

標記作業就是針對照片上的物件進行框選，針對垃圾數位化照片共 20,337 張照片，進行標記，標記後區分為訓練樣本及測試樣本，相關數量如下表。

表四:垃圾數位化數量及訓練及測試樣本數量表

蒐集的垃圾物件類別	數位化照片數	訓練樣本數	測試樣本數
鐵鋁罐類	4,305	18,303	2,034
鋁箔包(原廢紙容器類)	5,749		
塑膠瓶類	4,369		
廢燈泡類	1,411		
廢電池類	739		
玻璃瓶	901		
紙盒類	1,266		
廢紙類	584		
紙杯類(新增)	913		
小計	20,337	18,303	2,034

由於我們改變了模型架構，因此在「物件偵測」模型的標記作業，標記的物件類別，我們統一僅標記「object」這一類，這樣的做法大幅降低標記作業中容易標錯的事件，尤其越多類別時很容易不小心框選錯誤，一但框選錯誤的標記資料一並進行訓練時，容易造成訓練資料的混淆，可能會造成辨識率的降低，尤其是單張照片多物類別物件的標記，特別容易框選錯誤，而在「物件分類」網路模型中，我們不需要針對照片的內容進行標記，而是透過整理檔案清單並加以加註類別來進行物件分類的標記文件作業，因此，相較於物件檢測的標記難度及所耗時間均有大幅度的節省。

3. 訓練作業

本研究採用 Caffe (Convolutional Architecture for Fast Feature Embedding) 深度學習框架[14]作為我們深度學習訓練的框架，訓練的作業系統選擇以 Ubuntu 18.4 [17]環境作為 Host 並在 Linux 環境中安裝 Docker[18]以及 Nvidia-Docker[19]來簡化 Caffe 安裝的作業，藉由下載 Opensource 中已經分享出來的 Caffe Docker 環境來進行訓練，本研究採用 Docker 來進行訓練主要的原因是為安裝步驟並非是本研究的重點，過去在系統安裝準備上經常耗費太多時間進行除錯或者修正，利用 Docker 下載 image 能有效簡化安裝步驟，而且萬一因為調整的關係，把環境弄壞了可以很容易地重新啟用恢復或迅速新增環境，是目前對於環境建置的有效手段，也是符

合目前資訊科技運用的主流，本研究透過網路上分享已建置好的環境，下載後進行變更，能快速運行及測試本研究所需要的環境。

由於未來應用的場景是希望辨識程式運作在相對便宜低成本的硬體上，因此在神經網路的設計上要思考較輕量的網路，最好能夠兼具辨識率與效能上的優勢，參照 Wei Liu 等人在 2016 年所提出的 SSD- Single Shot MultiBox Detector [5]及 Andrew G[6]等人在 2017 年發表的 MobileNet 網路架構以及 Jonathan Huang[7]等人在 2017 年發表針對不同神經網路的效能比較，顯示以 MobileNet 結合 SSD 在效能及辨識效果上具有良好的表現，如下圖：

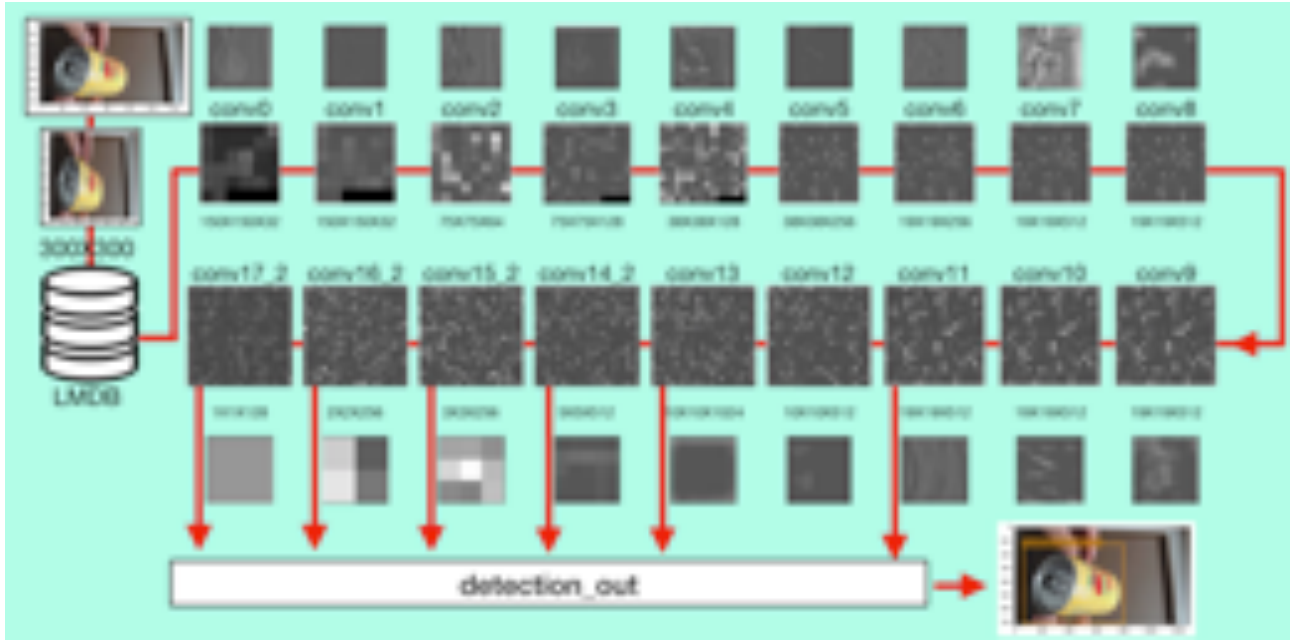


圖七、不同網路下 GPU 運行的時間比較圖[16]

因此本研究選擇以 MobileNet-SSD 作為步驟一物件偵測模型的卷積神經網路，此外 MobileNet 網路本身就是輕量級別的卷積神經網路，應用在物件分類模型上，亦兼具效能與效果上的優勢，因此我們亦選擇採用 MobileNet 來作為本研究步驟二的物件分類模型的卷積神經網路。

(1) MobileNet-SSD 卷積神經網路

MobileNet-SSD 卷積神經網路是結合 MobileNet 與 SSD 兩個神經網路，其中 MobileNet 取代原先的 VGG16 的基礎網路，當影像輸入到卷積神經網路過程中，從 conv0 到 conv13 為 MobileNet 神經網路部分，conv13 之後就是 SSD 網路部分，本研究將 MobileNet-SSD 卷積可視化整理如下圖：



圖八:MobileNet-SSD 卷積神經網路卷積過程可視化示意圖

MobileNet-SSD 主要特色之一是使用深度可分離卷積(Depthwise Separable Convolution)結合點卷積(Pointwise Convolution)取代標準卷積，將標準卷積核進行分解計算，來達到減少計算量，例如，假設輸入為 M 個通道的特徵圖，卷積核大小為 $D_K * D_K$ ，輸出通道為 N ，則標準卷積核即為 $M * D_K * D_K * N$ ，例如，輸入特徵圖為 $m * n * 16$ ，輸出 32 通道(channels)則標準的卷積計算量為：

$$m * n * 16 * 3 * 3 * 32 = m * n * 4,608 \quad (2.2.2.1)$$

而使用深度可分離卷積的計算量為：

$$m * n * 16 * 3 * 3 + m * n * 16 * 1 * 1 * 32 = m * n * 656 \quad (2.2.2.2)$$

因此深度可分離卷積核與標準卷積核相比計算量比率為：

$$\frac{D_K * D_K * D_F * D_F * M + D_F * D_F * M * N}{D_K * D_K * M * N * D_F * D_F} = \frac{1}{N} + \frac{1}{D_K^2} \quad (2.2.2.3)$$

其二是引入寬度乘數(Width Multiplier) α 來減少輸入輸出的通道數(channels)，輸入層的通道個數 M 變成 αM ，輸出層的通道個數 N 變成 αN ，引入寬度乘數後總計算量如下：

$$D_K * D_K * \alpha M * D_F * D_F + \alpha M * \alpha N * D_F * D_F \quad (2.2.2.4)$$

其三是引入解析度乘數(Resolution Multiplier) ρ 來降低圖片的解析度，即作用在特徵圖(feature map)的大小上，所以引入解析度乘數後的總計算量如下：

$$D_K * D_K * \alpha M * \rho D_F * \rho D_F + \alpha M * \alpha N * \rho D_F * \rho D_F \quad (2.2.2.5)$$

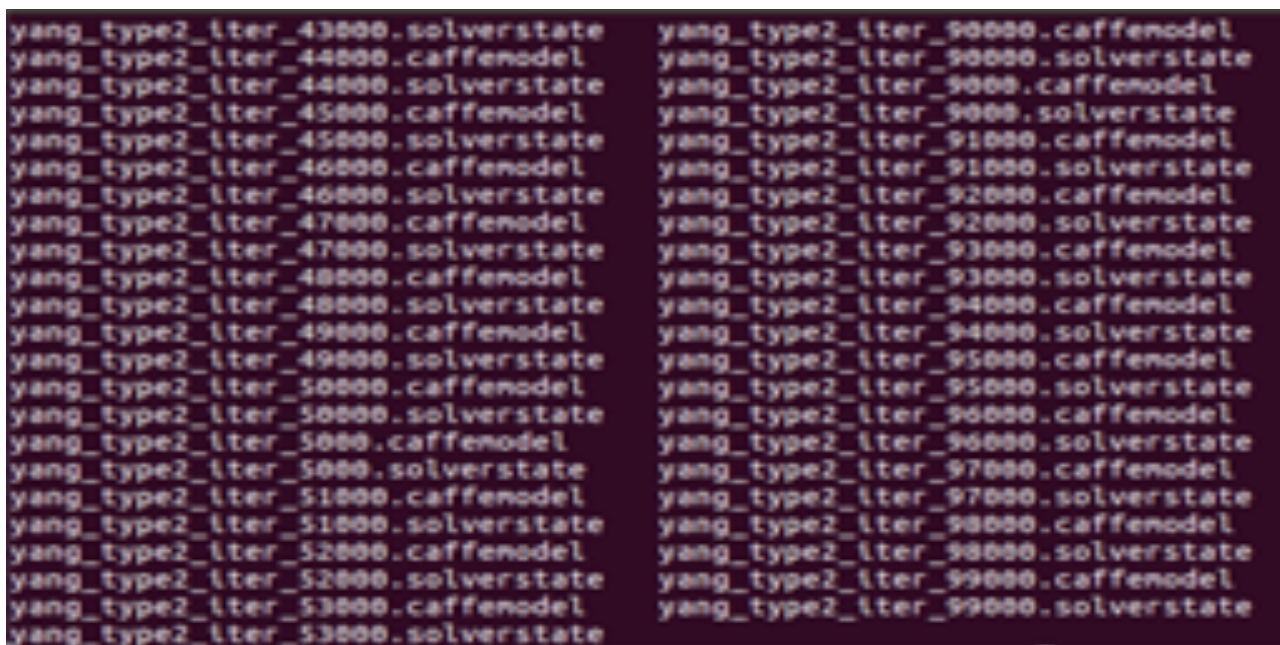
MobileNet-SSD[18]在 conv13 後面添加了 8 個卷積層，然後總共抽取 6 層作為檢測，提取預設框(default box)，該 6 層為 conv11,conv13,conv14_2,conv15_2,conv16_2,conv17_2，其後連接用於類別得分的 3*3 卷積層，層名為 conv11_mbox_conf, conv13_mbox_conf, conv14_2_mbox_conf, conv15_2_mbox_conf, conv16_2_mbox_conf, conv17_2_mbox_conf 輸出是以我們訓練的類別數乘上 3 個預設框。

損失函式分為兩個部分：定位損失和分類損失，公式如下：

$$L(x, c, l, g) = \frac{1}{N}(L_{conf}(x, c) + \alpha L_{loc}(x, l, g)) \quad (2.2.2.6)$$

其中，c 代表分類可信度，l 代表預測框，g 代表真實框，conf 代表分類損失，loc 代表定位損失。

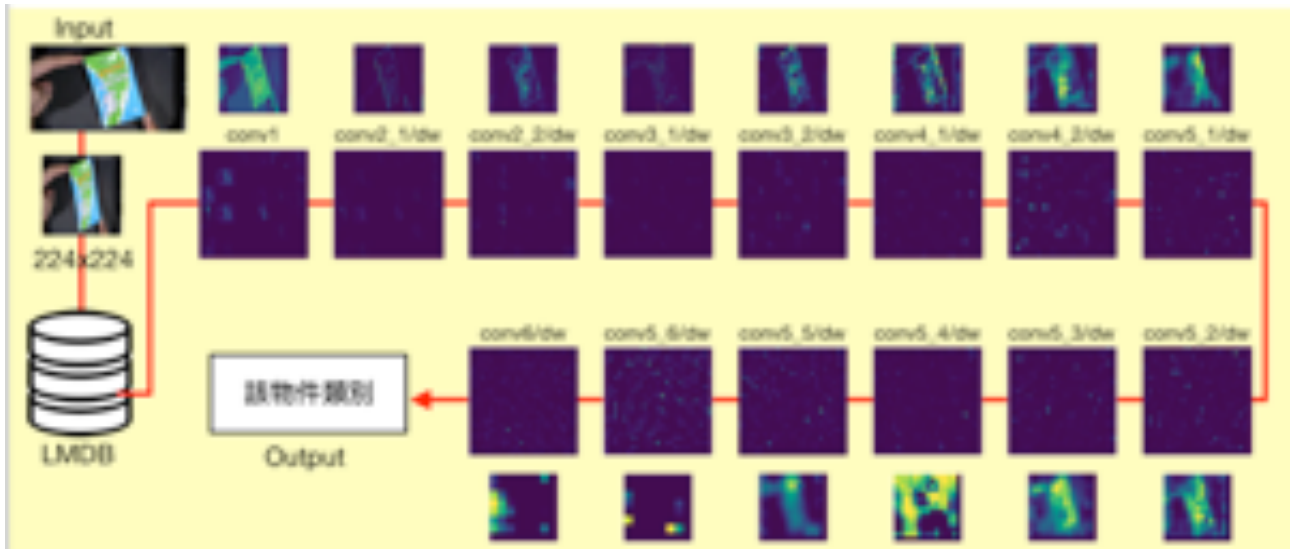
針對 20,337 張標記完成的照片放到步驟一的訓練環境中，並執行 create_list.sh 產生訓練集(trainval.txt)及測試集(test.txt)的文字檔，並執行 create_data.sh 將 JPEGImages 目錄中的照片轉成 lmbd 資料庫，並在 Docker 下的 Caffe 深度學習框架中以 Mobilenet-SSD 卷積神經網路進行訓練(執行./train.sh)，經過迭代 10 萬次得到 loss 約 0.02-0.05 左右的訓練成果，本研究設定每 1,000 次迭代將產生一份權重檔在 Snapshot 目錄中，訓練完成後在 Snapshot 目錄中我們得倒每千次的權重檔，具體結果如下圖：



圖九:本研究訓練結果產生權重檔示意圖

(2) MobileNet 卷積神經網路

MobileNet 卷積神經網路的特色，如前面描述，就是透過深度可分離卷積(Depthwise Separable Convolution)結合點卷積(Pointwise Convolution)取代標準卷積，進而達到降低運算量的特點，本研究採用 MobileNet 作為步驟二物件分類的卷積神經網路，同樣針對卷積層可視化，整理如下圖表示 MobileNet 的卷積過程。



圖十:MobileNet 卷積神經網路卷積過程可視化示意圖

4. 模型評價

在二元分類中，常使用精確度(Precision)及召回率(Recall)來衡量及評價的統計，不過目標檢測(Object Detection)應用在影像辨識上除了分類的正確與否，還包括圖像中物體的定位，而且目標檢測除了單一物件外，也包括多物件偵測，如下圖。



圖十一、常見的計算機視覺問題(Stanford's CS231n 課程資料)

因此，本研究在創意說明階段採用 PASCAL Visual Object Classes(VOC) Challenge 所定義的 11-point interpolated average precision[20]來作為衡量模型的評價，其公式如下：

$$AP = \frac{1}{11} \sum_{r \in \{0,0.1, \dots, 1\}} P_{interp}(r) \quad (2.2.3.1)$$

在訓練階段我們開啟 solver_tran.prototxt 中的 show_per_class_result: True 的參數，並定義每迭代 1,000 次進行一次 AP 及 mAP 的測試，統計 120 次的評價平均結果如下表：

表五、各類垃圾物件 AP 值及 mAP 值(創意說明階段)

類別	鐵鋁罐類	廢紙容器類	塑膠瓶類	廢燈泡類	廢電池類	玻璃瓶類	紙盒類	廢紙類	mAP
AP 值	0.959	0.977	0.998	0.907	0.994	0.991	0.933	0.998	0.969

訓練結果得到各類垃圾物件約平均 0.907~0.998 的 AP 值，此外，mAP 高達 0.969，算是得到不錯的成果。

不過本次成果報告階段改變模型架構，我們在步驟一採用物件偵測模型的訓練，所有標記的 20,337 張照片均僅標記「object」一種類別物件，最終訓練迭代 10 萬次完成訓練的 AP 及 mAP 均為 1（100%），如下圖所示：

```

18985 12:05:32.813568 3486 solver.cpp:259] Train net output #0: mbox_loss =
0.167391 (* 1 = 0.167391 loss)
18985 12:05:32.813513 3486 sgd_solver.cpp:138] Iteration 99970, lr = 0.000125
18985 12:05:46.673804 3486 solver.cpp:243] Iteration 99980, loss = 0.041827
18985 12:05:46.673825 3486 solver.cpp:259] Train net output #0: mbox_loss =
0.0227683 (* 1 = 0.0227683 loss)
18985 12:05:46.673828 3486 sgd_solver.cpp:138] Iteration 99990, lr = 0.000125
18985 12:06:00.545885 3486 solver.cpp:243] Iteration 99990, loss = 0.0488104
18985 12:06:00.545969 3486 solver.cpp:259] Train net output #0: mbox_loss =
0.0238257 (* 1 = 0.0238257 loss)
18985 12:06:00.545974 3486 sgd_solver.cpp:138] Iteration 99990, lr = 0.000125
18985 12:06:13.278344 3486 solver.cpp:596] Snapshotting to binary proto file sn
snapshot/yang_type2_iter_100000_caffemodel
18985 12:06:13.311591 3486 sgd_solver.cpp:387] Snapshotting solver state to bin
ary proto file snapshot/yang_type2_iter_100000_solverstate
18985 12:06:13.413807 3486 solver.cpp:332] Iteration 100000, loss = 0.0506427
18985 12:06:13.413825 3486 solver.cpp:433] Iteration 100000, Testing net (#0)
18985 12:06:13.413944 3486 net.cpp:693] Ignoring source layer mbox_loss
18985 12:06:13.426895 3486 solver.cpp:540] class1: 1
18985 12:06:13.426166 3486 solver.cpp:546] Test net output #0: detection_ev
al = 1
18985 12:06:13.426110 3486 solver.cpp:337] Optimization Done.
18985 12:06:13.426111 3486 caffe.cpp:254] Optimization Done.
38484_1 Bot
    
```

圖十二、步驟一物件偵測模型訓練圖

而在步驟二，物件分類模型訓練的模型評價，透過 MobileNet 分類卷積神經網路訓練 20 萬次後最終得到了 0.959 準確度(accuracy) 的訓練成果，訓練結果如下圖。

```

10909 21:19:31.831743 28957 solver.cpp:259] Train net output #0: loss = 1.64
915e-05 (* 1 = 1.64915e-05 loss)
10909 21:19:31.831746 28957 sgd_solver.cpp:138] Iteration 199970, lr = 0.01
10909 21:19:33.068761 28957 solver.cpp:243] Iteration 199980, loss = 3.91873e-05
10909 21:19:33.068779 28957 solver.cpp:259] Train net output #0: loss = 7.09
760e-05 (* 1 = 7.09760e-05 loss)
10909 21:19:33.068783 28957 sgd_solver.cpp:138] Iteration 199980, lr = 0.01
10909 21:19:34.305748 28957 solver.cpp:243] Iteration 199990, loss = 3.69648e-05
10909 21:19:34.305766 28957 solver.cpp:259] Train net output #0: loss = 7.55
495e-06 (* 1 = 7.55495e-06 loss)
10909 21:19:34.305769 28957 sgd_solver.cpp:138] Iteration 199990, lr = 0.01
10909 21:19:35.516394 28957 solver.cpp:506] Snapshotting to binary proto file DW
-2-snapshot/yang_step2_rm_dw2_iter_200000.caffemodel
10909 21:19:35.542598 28957 sgd_solver.cpp:387] Snapshotting solver state to bin
ary proto file DW-2-snapshot/yang_step2_rm_dw2_iter_200000.solverstate
10909 21:19:35.576629 28957 solver.cpp:332] Iteration 200000, loss = 4.42679e-05
10909 21:19:35.576643 28957 solver.cpp:358] Iteration 200000, Testing net (#0)
10909 21:19:41.182582 28957 solver.cpp:425] Test net output #0: accuracy = 0
.9595
10909 21:19:41.182518 28957 solver.cpp:425] Test net output #1: loss = 0.144
435 (* 1 = 0.144435 loss)
10909 21:19:41.182528 28957 solver.cpp:337] Optimization Done.
10909 21:19:41.182523 28957 caffe.cpp:254] Optimization Done.

```

圖十三、步驟二物件分類模型訓練完成圖

(三) 系統測試與分類機構整合：

(1) 系統測試

系統測試主要是了解本研究的效果，首先將訓練好的模型檔放在辨識盒中，為了盡量降低成本，本研究選擇研揚公司的 UP 系列產品，該系列產品主要是應用在 IoT 領域，作為較輕量運算的設備，系統並支援 Linux 與 Windows 作業系統，將辨識盒安裝在模擬的機構上進行實際垃圾物件的丟擲辨識測試及統計辨識率；在創意說明階段，本研究使用單一物件檢測卷積神經網路進行固定外型及形變垃圾物件的辨識，得到辨識結果結果如下表：

表六、創意說明階段實際垃圾丟擲辨識率統計表

蒐集的垃圾物件類別		實體垃圾丟擲次數	辨識正確數	辨識錯誤數	辨識率
固定外型	廢燈泡類	120	119	1	99.2%
	廢電池類	120	116	4	96.6%
	玻璃瓶類	120	118	2	98.3%
包含形變外型	鐵鋁罐類	120	113	7	94.1%
	廢紙容器類	120	109	11	90.8%
	塑膠瓶類	120	116	4	96.6%
	紙盒類	120	115	5	95.8%
	廢紙類	120	117	3	97.5%
小計		960	923	37	96.1%

使用單一物件檢測卷積神經網路模型，得到的整體辨識率約為 96.1%。



而在成果報告階段使用雙卷積神經網路進行固定及形變垃圾物件的辨識，使同樣方式在機構上進行實際垃圾物件的丟擲辨識測試及統計辨識率，得到辨識結果如整理如下表：

表七、成果報告階段實際垃圾丟擲辨識率統計表

蒐集的垃圾物件類別		實體垃圾丟擲次數	辨識正確數	辨識錯誤數	辨識率
固定外型	廢燈泡類	120	119	1	99.2%
	廢電池類	120	118	2	98.3%
	玻璃瓶類	120	117	3	97.5%
包含形變外型	鐵鋁罐類	120	113	7	94.2%
	鋁箔包類	120	112	8	93.3%
	塑膠瓶類	120	115	5	95.8%
	紙盒類	120	116	4	96.7%
	廢紙類	120	117	3	97.5%
	紙杯類	120	112	8	93.3%
小計		1,080	1,039	41	96.2%

整體辨識效果與單一物件偵測卷積神經網路相近的辨識水準，整體辨識率約 96.2%，不過雙卷積神經網路模型在多物件表現，明顯優於單一物件偵測卷積神經網路模型，下表是創意說明階段與成果報告階段採用兩種模型架構在多物件的比較表：

表八、兩種模型架構在多物件辨識比較表

創意說明階段(先前)	成果報告階段(本次)
單一物件偵測卷積神經網路	雙卷積神經網路
	
兩物件太靠近時會錯誤辨識為一個物件	兩物件太靠近時仍能分別偵測不同物件



而在物件的相似性上也具備辨識能力，如下表：

表九、不同材質但雷同的印刷外觀的物件辨識效果表

相似性的辨識效果(不同材質但雷同的印刷圖案)		
鋁罐材質	鋁箔包材質	寶特瓶材質

(2) 效能調整

效能調整的部分主要是希望盡量優化辨識盒的辨識效能，本研究以程式及系統效能優化進行調教，由於硬體支援 OpenCL 加速，可以透過 OpenCL 讓 GPU 協助運算，因此針對硬體首先安裝支援 OpenCL 的顯示卡的驅動程式，再透過 OpenCV 驅動令 GPU 進行運算，讓程式運作的效能提升，先前在創意說明階段成功透過 GPU 將效能從 1.09fps(frame per Second)提升至 12fps 左右，下表：

表十、效能調整改進前與改進後效果表

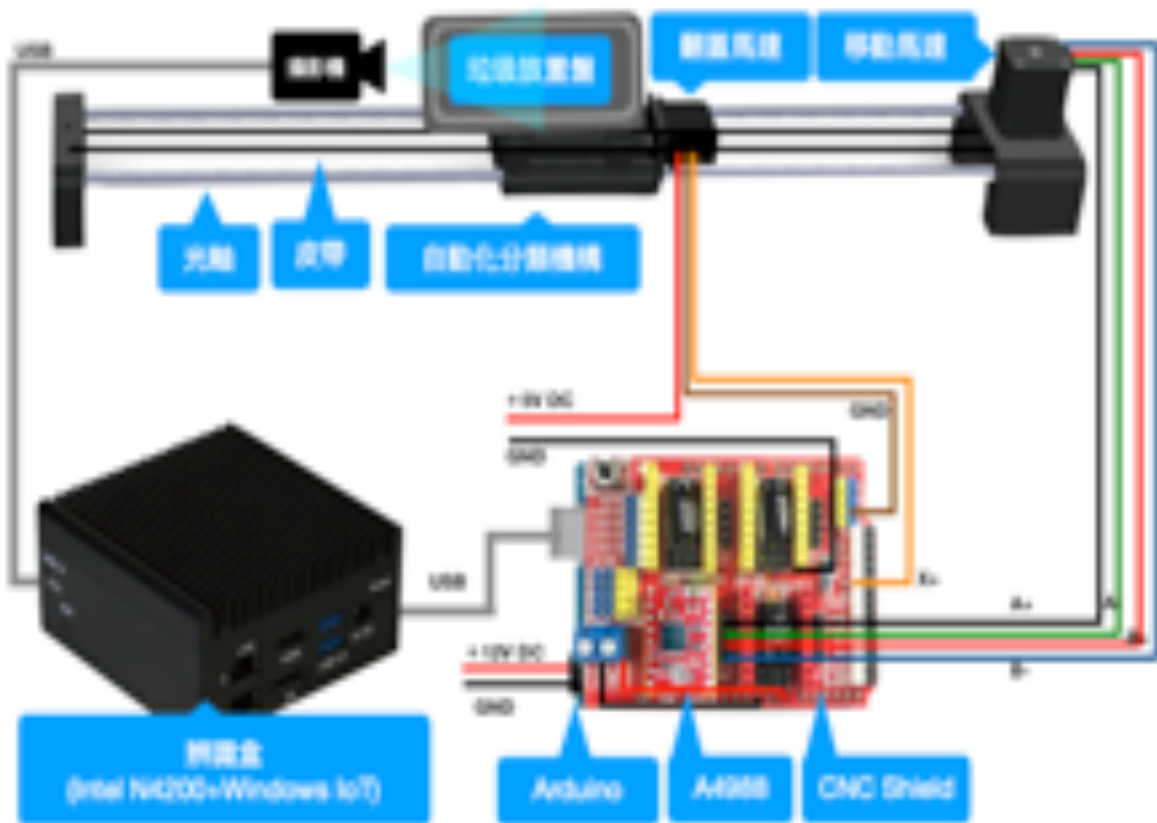
改善前(約 1.09fps)	改善後(約 12fps)

不過本研究在成果報告階段採用雙卷積神經網路架構，因此，相較之前單一卷積神經網路需要增加另一個神經網路的運作，即便透過 GPU 優化，在 N4200 低階電腦執行仍僅有 2-3fps 左右，雖然在整體架構上勉強能夠應付，不過這也是未來值得進一步努力優化的方向。

(3) 分類機構整合

礙於研究時間有其限制加上學生的資源有限，分類機構的設計主要的目的在於驗證辨識盒對垃圾物件辨識後，將類別傳遞到自動分類機構，用以驗證自動化分類的可行性，因此在機構設計上先不追求美感，而以功能為主。

機構的固定件透過 FreeCAD 進行繪製，並利用 3D 印表機列印相關機構固定件；而在控制系統的部分，本研究利用 Arduino Uno 控制馬達，並驅動皮帶，帶動垃圾置放盤，當 Arduino 收到辨識盒辨識置放在垃圾置放盤上的垃圾類別後，會移動垃圾置放盤到垃圾桶定位點，再驅動馬達進行翻蓋丟擲垃圾的動作，本研究採用 Arduino 作為馬達及機構控制的控制板，主要是已經有很多成功案例透過 Arduino 控制板來進行自動化機構驗證，在相關學習的資源較容易取得，另一方面 Arduino 簡易且直觀的程式語法可以有效縮減本研究準備相關部件控制的時間，此外在成本上也較便宜，因此本研究選擇 Arduino 作為本研究機構控制的主要控制板，整體機構的線路連接方式如下圖：



圖十四、機構電路連接示意圖

另外整理本研究機構設計使用的材料及組裝後樣式如下表：

表十一、自動分類機構使用材料表(含組裝後樣式)

自動分類裝置示意模型(本研究自製)			
裝置	型號	實體照片	規格
Arduino 控制板	Uno		ATMEGA328P-U
Arduino CNC Shield	V3		Arduino CNC Shield 數控機擴展板
步進馬達控制器	A4988		A4988 步進馬達控制晶片組
伺服馬達	MG995		MG995 伺服馬達(舵機)
步進馬達	NEMA17		42 型步進馬達 L40 2 相 4P
同步帶輪	2GT-20T		內孔 5mm 齒面 7.3mm AF10.5mm
同步帶輪	2GT-20T		內孔 5mm 齒面 7.3mm CF16mm
線性承軸	LM8UU		標準線性承軸 X4
訂製長度光軸	None		直徑 8mm X 600mm
同步帶	2GT		UNITTA 2GT 寬 6mm X 長 600mm
電源供應器	LRS-50-12		100V-240V 輸入 12V 4.2A 輸出
USB 線	None		Type-A(Type-B)
鋁擠	None		鋁擠 X5
托盤	None		260mmX190mm
小型垃圾桶	None		小型垃圾桶 X1
3D 列印機構元件	None		自行繪製機構元件並透過 3D 列印
3D 列印機構元件	None		自行繪製機構元件並透過 3D 列印
3D 列印機構元件	None		自行繪製機構元件並透過 3D 列印
3D 列印機構元件	None		自行繪製機構元件並透過 3D 列印
3D 列印機構元件	None		自行繪製機構元件並透過 3D 列印
組裝後樣式			
			

分類機構的移動方式是透過馬達帶動皮帶牽引垃圾放置盤下的機構，再透過垃圾放置盤旁的伺服馬達來進行翻蓋倒垃圾的動作，機構的優點是動作簡單，透過 Arduino 控制第一個步進馬達進行移動到達定位點，之後再透過伺服馬達(舵機)來執行翻蓋的任務。

第一個步進馬達所使用的是 NEMA17 從規格上得知走 1 步是 1.8 度，透過公式換算可以知道一圈是 360 度，除以 1.8 度等於 200，表示步進馬達走 200 步等於 1 圈，而我們透過 Arduino 上安裝 CNC Shield 數控擴展板上插 A4988 的馬達控制模組來控制步進馬達，所以要特別注意 CNC Shield 板上在 A4988 下的跳線，分別有全步進、1/2 步進、1/4 步進、1/8 步進、1/16 步進的模式可以選擇，如下表：

表十二、透過 A4988 控制 NEMA17 步進馬達步數表

MS1	MS2	MS3	Microstep Resolution		一圈的步數
L	L	L	Full Step	1	200
H	L	L	Half Step	1/2	400
L	H	L	Quarter Step	1/4	800
H	H	L	Eight Step	1/8	1,600
H	H	H	Sixteenth Step	1/16	3,200

所以在 Arduino 的程式控制上要注意這個差異，另外利用 A4988 第四腳位(pin4) 上的高電位控制步進馬達正轉也就是移動垃圾放置盤前進，低電位來控制步進馬達反轉也就是垃圾放置盤回到原位，第五腳位(pin5)則用來控制步進馬達移動的步數，一個 High 加上一個 Low 信號代表一步，可以依照垃圾桶位置計算好距離讓垃圾放置盤移動到定位點，再透過控制伺服馬達 MG995 來進行翻蓋的動作，程式啟動後會先宣告伺服馬達 MG995 初始角度為 180 度，也就是水平，到達定位點後再讓伺服馬達 MG995 轉到 0 度，也就是垂直方向，後停留 500 毫秒，再從 0 度轉回 180 度。

分類機構的限制是水平移動的速度不可以太快，否則放置在垃圾放置盤上的垃圾物件有可能有機會掉出來，其次是翻蓋透過伺服馬達舉升的能力有限，對於較重的玻璃瓶垃圾物件有時會比較吃力，使用多次受力容易崩壞。

本研究之自動分類機構主要是表達自動分類的可行性，想像丟垃圾只要把垃圾物件丟在垃圾放置盤後，透過攝影機辨識到垃圾物件的類別，進而控制垃圾放置盤移動到該類垃圾的垃圾桶位置，然後進行翻蓋(倒垃圾)的動作，因此，一次只能丟擲一個類別的垃圾也算是自動分類機構的限制，實際上如果一次放至兩種類別的垃圾物件在垃圾放置盤上，自動分類機構會以第一個通知 Arduino 移動到定位點的資訊為主，另一個物件就會跟著進垃圾桶中，這是後續還是要繼續研究及思考的部分。

肆、 結論

本研究成功改進創意說明階段所提出-透過深度學習卷積神經網路模型來直接對固定及形變垃圾物件進行辨識，在創意說明階段，本研究採用單一物件檢測卷積神經網路模型來作為影像辨識的核心，針對收集的垃圾物件直接透過影像進行辨識，不論是固定外型或是形變的

垃圾物件，均具有良好的辨識效果；本次成果報告階段更進一步提出「雙卷積神經網路模型」來針對固定及形變垃圾物件直接進行影像辨識。

在雙卷積神經網路模型中，首先針對物件偵測卷積神經網路進行模型訓練，獲得物件偵測的能力，其次透過物件分類卷積神經網路進行模型訓練，來針對物件進行分類的辨識，在此架構中，物件偵測的模型能夠跨越不同的物件類別，重複使用性大幅提升，而在物件分類的模型訓練中，亦能獲得訓練時間大幅縮短的好處，新架構透過程式進行整合，在實際測試的辨識效果中，有良好的辨識能力，其中三類固定外型(未形變)的垃圾物件，辨識率約達 98.3%，而六類包括變型的垃圾物件也具有約 95.1%的辨識率，整體辨識率約達 96.2%，辨識效果良好；另外整合自動化分類機構能透過 Arduino Uno 控制馬達移動垃圾放置盤，正確移動到該類垃圾桶定位點，進行翻蓋動作。

雙卷積神經網路模型架構對比單一物件檢測卷積神經網路模型雖然優點不少，但運作的效能卻降低將近三倍，雖然硬體效能的提升及改朝換代的速度極快，但這部分應仍可作為未來延伸研究的重點項目。

伍、 討論與應用













影像辨識常受光線影響，因此在不同光照強度下經常影響辨識率，屬於較常見且不可避免的要害之一，本研究希望提升垃圾桶的智能，具備辨識垃圾物件的能力，因此想到垃圾桶放置的位置，大致包括戶外、室內及室內角落的場所，這些場所光照程度不一，很可能影響實際辨識的結果，於是想進一步驗證目前所作基於深度學習卷積神經網路的固定及形變的垃圾物件辨識系統，對於光照的影響及辨識範圍，因此參考中華民國國家標準 CNS 照度標準(總號：12112 類號：Z1044)，此標準規範不同場所的光照標準，本研究參考此標準作了簡單的實驗設計並加以探討，

- 一、參考標準將場地狀況大致區分為「戶外」、「室內」以及「地下室」三個場景的光照狀況來加以探討。
- 二、到各個場景尋找垃圾桶置放位置並透過流明計紀錄場域實際照度，照度的定義為被照體單位面積所受的光通量，其單位為勒克斯(LUX)。
- 三、在辨識分類裝置上模擬各種光線並進行實際測試驗證辨識效果。

經實際量測不同場景，發現光照變化巨大，戶外場景下的垃圾桶位置所量測到的光照強度約 5,000~10,000 流明(Lux)，室內場景下在垃圾桶位置所量測到的光照強度約 200~500 流明(Lux)，而地下室或大樓資源回收處垃圾桶位置所量測到的光照強度約 10~50 流明(Lux)，比較 CNS 照度標準，實際場景約略低於 CNS 所規範的場域光照標準，依據選定的三個場景進行實際量測並記錄，並在分類裝置上藉由流明計模擬各種場景下的光照強度，並選定寶特瓶來作為實驗的測試物件，主要是因為寶特瓶的材質較容易受光線的反射而產生的特徵樣貌材質感

差異較為多變，經過實驗測試本研究發展的固定及形變垃圾物件辨識系統，在光照強度 10Lux~10,000Lux 下，仍具備強健的影像辨識能力，實際測試整理如下表：

表十三、垃圾桶不同場域光照模擬辨識效果表

場景	實際量測照度	模擬量測照度	固定物件辨識效果	變形物件辨識效果
戶外				
	9,290Lux	9200Lux	辨識正確	辨識正確
室內				
	271.5Lux	287.2Lux	辨識正確	辨識正確
地下室				
	24.5Lux	10.8Lux	辨識正確	辨識正確

非常榮幸獲得評審給予作品評語及意見指導，在思考如何改善演算法及模型以達到較佳的固定及形變垃圾分類時，考量之前單一卷積神經網路模型在收集資料及訓練調教中的難度，因此改變作法更新以「雙卷積神經網路」來進行垃圾物件的辨識，也在本次成果報告中增加有關形變物件辨識的相關研究的資料查找及比較、增加清大 AI 回收機器人的相關比較、因應一般分類的習慣，將原本八類垃圾物件重新整理為九類，將廢紙容器類修正為鋁箔包類，並將其中的紙杯類獨立成新增的一類(紙杯類)加入訓練與辨識、增加相關垃圾數位化照片的拍攝與收集、也將自動分類的機械結構、電路架構以及軟體的運作方式進行補充並探討其限制。

在深度學習領域雖然是較為新興的資訊科技，但在研究固定及形變垃圾物件的辨識過程，發現新技術的浩瀚與全球資訊科技前輩不斷的創新與開源，使得能有機會躬逢其盛，並站在巨人的肩膀獲得些許的成果，相信在不斷努力前進的道路上，有機會整合全球分享的技術知識，來解決垃圾分類與資源回收的相關議題。

陸、 參考資料

- [1] 范宏偉-學生對於學校垃圾分類與公共關係組織效能關係之研究-以苗栗縣國民中學為例，2012 玄奘大學碩士論文。
- [2] Alex Krizhevsky,Ilya Sutskever,Geoffrey E. Hinton imagenet classification with deep convolutional neural networks 2012.
- [3] Jonathan Huang,Vivek Rathod,Chen Sun,Kevin Murphy,Menglong Zhu,Anoop Korattikara-Speed/accuracy trade-offs for modern convolutional object detectors-Google Research, 2017.
- [4] Yangqing Jia,Evan Shelhamer,Jeff Donahue,Sergey Karayev,Jonathan Long,Ross Girshick, Sergio Guadarrama,Trevor Darrell-Caffe Convolutional architecture for fast feature embedding-2014.
- [5] Wei Liu, Dragomir Anguelov,Dumitru Erhan,Christian Szegedy-SSD- Single Shot MultiBox Detector-2016.
- [6] Andrew G. Howard,Menglong Zhu,Bo Chen, Dmitry Kalenichenko-MobileNets- Efficient Convolutional Neural Networks for Mobile Vision Applications-2017.
- [7] Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Speed/accuracy trade-offs for modern convolutional object detector-2017.
- [8] 龔純瑩-雙向性組合式記憶模型輿圖行辨識之研究，1991 中原大學碩士論文。
- [9] 鄭萬旗-類神經網路於手寫數字辨識之應用，2000 淡江大學碩士論文。
- [10] 李丞揚、牟代君、葉建德、***-以人工智慧及物聯網方法探討垃圾分類自動辨識之研究，2019 TANET-臺灣網際網路研討會口頭論文發表。

網站參考資料

- [11] Jane J. Lee，吸管怎麼會跑進海龜鼻子裡？國家地理雜誌，2015。資料取自 <https://www.natgeomedia.com/environment/article/content-5899.html>
- [12] 看德國如何回收各種瓶子。資料取自 <https://www.youtube.com/watch?v=3J3PbeY7ENo>
- [13] 李天怡，瑞典寶特瓶回收系統超威!快速辨色分 100%再利用，消失的國界 PART1，三立新聞台，2017。資料取自 <https://www.youtube.com/watch?v=3t2A27ppWMI>

- [14] 韓國資源回收公司。資料取自 <http://superbin.co.kr/new/contents/product.php>
- [15] 王若樸，整合 NB-IoT 和 AI 排程，大豐要打造新型態資源回收物聯網，iThome，2018。資料取自 <https://www.ithome.com.tw/news/126386>
- [16] GitHub 影像圖形標記技術論壇。資料取自 <https://github.com/tzutalin/labelImg>
- [17] Ubuntu Linux Kernel。資料取自 <https://www.ubuntu.com/>
- [18] Mark Everingham · Luc Van Gool · Christopher K. I. Williams · John Winn · Andrew Zisserman，Enterprise Container Platform for High-Velocity Innovation。資料取自 <https://www.docker.com/>
- [19] Github nvidia-docker 技術論壇。資料取自 <https://github.com/NVIDIA/nvidia-docker>
- [20] The PASCAL Visual Object Classes (VOC) Challenge。資料取自 http://homepages.inf.ed.ac.uk/ckiw/postscript/ijcv_voc09.pdf
- [21] 郭正芬，清大團隊開發 AI 回收機器人，3 秒回收 4 物資。資料取自 <https://udn.com/news/story/11322/3879627>




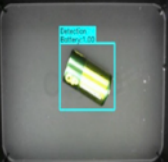





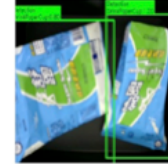

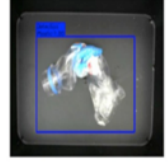




柒、附錄

研究設備及器材

電腦設備			
裝置	型號	實體照片	規格
MacBook Pro13 蘋果筆記型電腦	A1502		Intel Core i5 CPU @2.9GHz 8GB 1867MHz DDR3 RAM 512GB SSD 13-inch Intel Iris Graphics 1536MB
Logitech WebCam 羅技網路攝影機	C310		720P/30fps
AORUS X5 技嘉電競筆記型電腦	X5		Intel Core i7-7800K CPU @2.9GHz 16GB DDR4-2400MHz RAM / 256GB SSD+ 1TB HDD /15.6 吋 2880X1620 IPS 螢幕 /Nvidia GeForce GTX-1080 8GB DDR5
影像辨識盒			
裝置	型號	實體照片	規格
UP 系統	UPS-GWS01		Intel 奔騰 N4200/4 GB DDR4 /eMMC 32GB/RJ45X2/HDMI*1/USB3.0*3/ Windows 10 IoT
Logitech WebCam 羅技網路攝影機	C310		720P/30fps


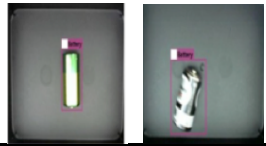
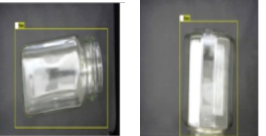

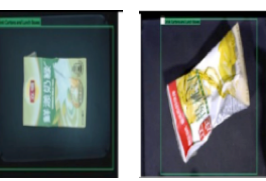



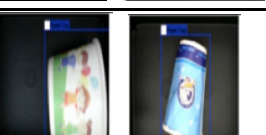
創意說明階段-實際垃圾丟擲辨識率統計表(含圖示)

8 類垃圾物件-單一卷積神經網路模型

蒐集的垃圾物件類別		實體垃圾 丟擲次數	辨識 正確數	辨識 錯誤數	辨識率	辨識效果
固定 外型 3 類	廢燈泡類	120	119	1	99.2%	 
	廢電池類	120	116	4	96.6%	 
	玻璃瓶類	120	118	2	98.3%	 
包含 形變 外型 5 類	鐵鋁罐類	120	113	7	94.1%	 
	廢紙容器類	120	109	11	90.8%	 
	塑膠瓶類	120	116	4	96.6%	 
	紙盒類	120	115	5	95.8%	 
	廢紙類	120	117	3	97.5%	 
小計		960	923	37	96.1%	

成果報告階段-實際垃圾丟擲辨識率統計表(含圖示)

9 類垃圾物件-雙卷積神經網路模型

蒐集的垃圾物件類別		實體垃圾 丟擲次數	辨識 正確數	辨識 錯誤數	辨識率	辨識效果
固定 外型 3 類	廢燈泡類	120	119	1	99.2%	
	廢電池類	120	118	2	98.3%	
	玻璃瓶類	120	117	3	97.5%	
包含 形變 外型 6 類	鐵鋁罐類	120	113	7	94.2%	
	鋁箔包類 (重新分類)	120	112	8	93.3%	
	塑膠瓶類	120	115	5	95.8%	
	紙盒類	120	116	4	96.7%	
	廢紙類	120	117	3	97.5%	
	紙杯類 (新增類別)	120	112	8	93.3%	
小計		1,080	1,039	41	96.2%	