

# 中華民國第 59 屆中小學科學展覽會 作品說明書

---

高級中等學校組 電腦與資訊學科

052508

基於深度學習卷積神經網路的垃圾分類裝置研究

學校名稱：新北市立林口高級中學

作者：  高二 李丞揚  高二 牟代君  高二 葉建德	指導老師：  胡裕仁
---	------------------

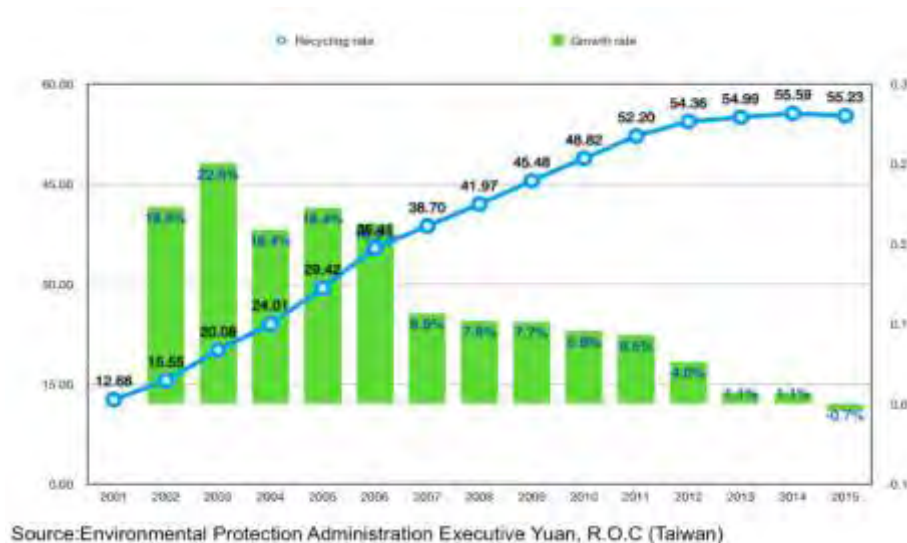
關鍵詞：深度學習、卷積神經網路、垃圾分類

## 摘要

垃圾分類是環保重要的議題，本研究提出一個自動辨識垃圾分類的裝置，透過深度學習卷積神經網路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，均有良好的效果。經測試，本系統已能有效分辨垃圾物件，在神經網路模型的評價中，平均 mAP 為 0.973，而針對實際垃圾物件進行辨識，在固定外型的垃圾物件中辨識率達 98%；而在包括形變的垃圾物件中辨識率仍具 95%；本研究證明利用卷積神經網路直接辨識垃圾物件具備一定的辨識率，若進一步發展有機會優化網路效能及辨識程式運作效率，可大幅度降低硬體的成本，擴大提升垃圾桶具備垃圾辨識的能力並進一步普及化並結合物聯網、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[8]，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(圖一)，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、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。

## 參、 研究方法

### 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(圖二)

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

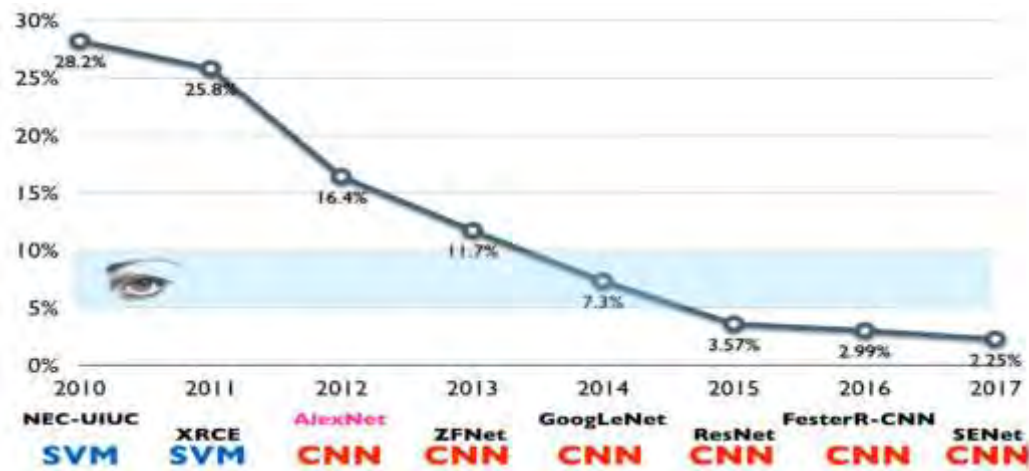
圖二、全國博碩士論文網站上查詢「垃圾」與「辨識」關鍵字的結果

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[9]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[10]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[11]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[12]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；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)：

年度	國家	公司/組織	名稱	主要辨識技術	垃圾形變辨識
2016	德國	Sielaff	Leergut 回收桶	辨識回收標誌	不具備
2107	瑞典	Eskilstuna	自動垃圾分類廠	辨識垃圾袋顏色	不具備
2018	韓國	SuperBin	AI 回收 Nephron	辨識垃圾物件	不具備
2018	台灣	大豐回收	zero zero 無人回收	辨識條碼	不具備
2019	台灣	本研究	辨識垃圾物件分類裝置	辨識垃圾物件	具備

表一：智能化垃圾分類設備的主要辨識技術整理比較

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(圖三)，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。

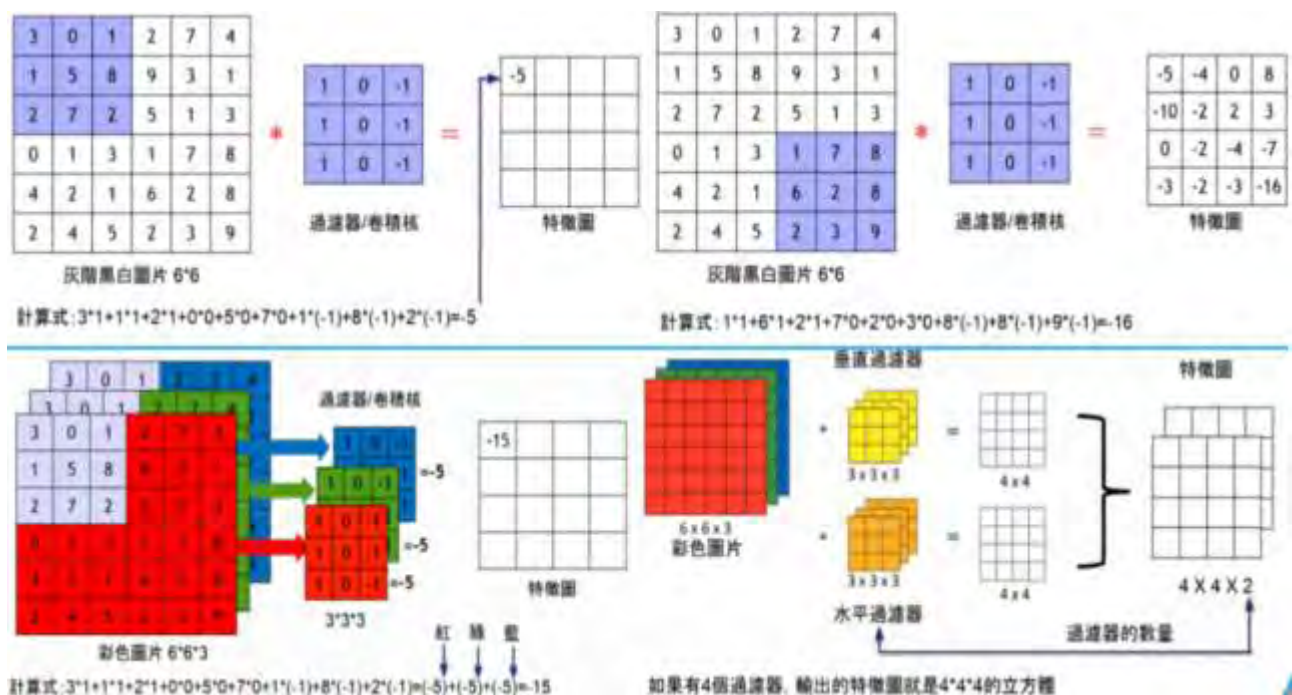


圖三:ILVRC 歷屆冠軍錯誤率與採用的主要方法圖(本研究整理)

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，如下圖四。

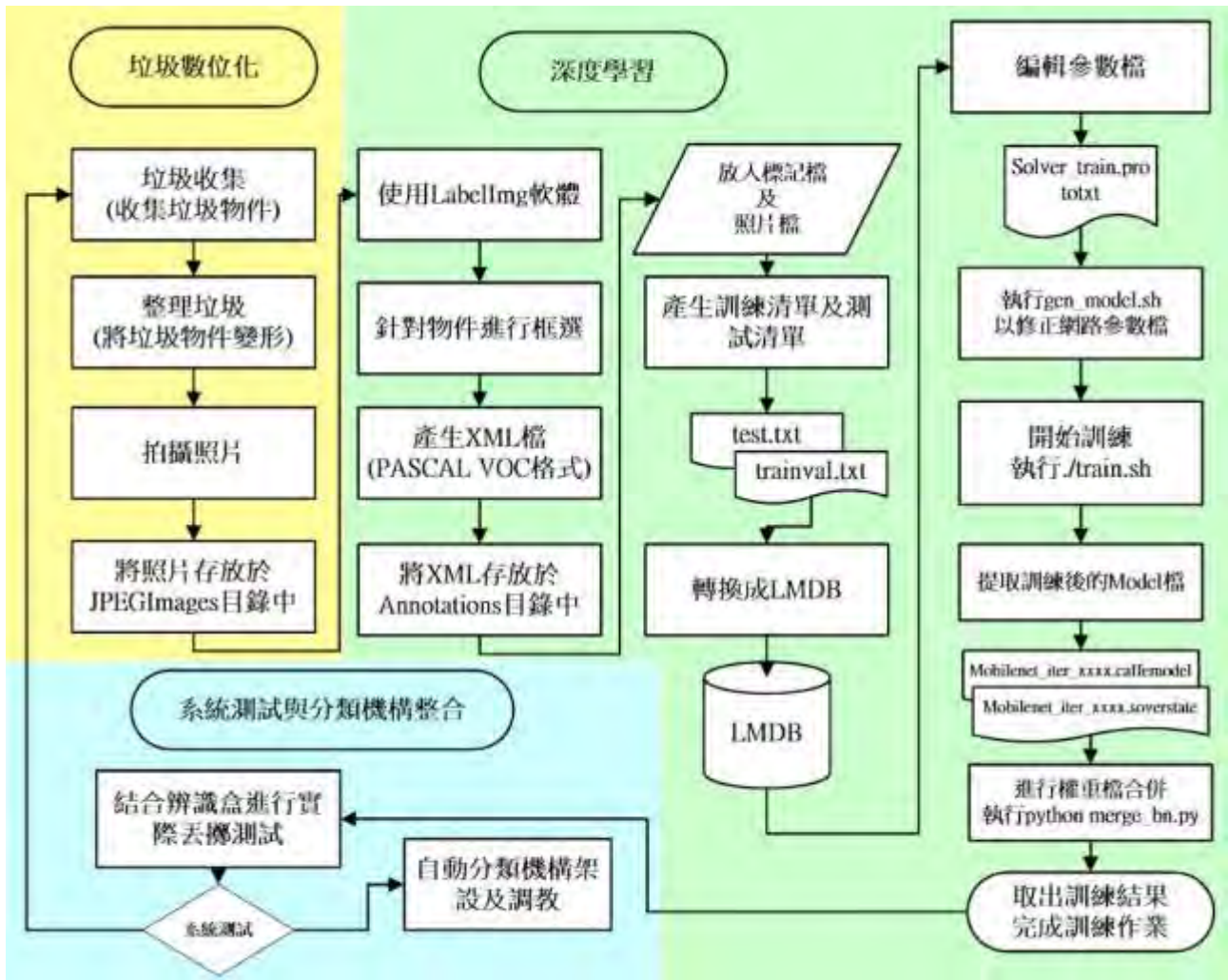


圖四:卷積運算示意圖(本研究整理)

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  $3 \times 3 \times 3$  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

由於缺乏前人直接透過影像辨識垃圾物件(包括形變)類別的公開研究可供參考，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(圖五)



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、整理垃圾以及拍攝照片，以下分別針對個步驟進行說明：

#### 1. 垃圾蒐集：

眾所週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，因此本研究先鎖定八類垃圾物件做為代表，分別是鐵鋁罐類、廢紙容器類、塑膠瓶類、廢燈炮類、廢電池類、玻璃瓶類、紙盒類、廢紙類等八種，這八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. 整理垃圾：

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，其餘五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，對於這五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，本研究一共完成 15,871 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「標記作業」、「訓練作業」及「模型評價」三步驟，以下進一步說明：

### 1. 標記作業

本研究將採用 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 [13]來針對相片進行標記，如下圖（圖六）；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 15,871 張照片，進行標記，標記後區分為訓練樣本及測試樣本，相關數量如下表二。

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	2,572	2,186	386
廢紙容器類	4,894	4,160	734
塑膠瓶類	4,287	3,644	643
廢燈炮類	635	540	95
廢電池類	739	628	111
玻璃瓶	901	766	135
紙盒類	1,266	1,076	190
廢紙類	577	491	86
小計	15,871	13,491	2,380

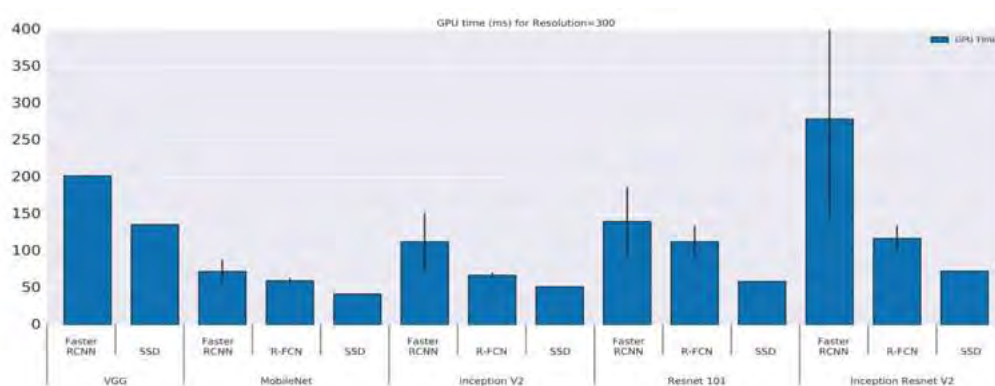
表二:垃圾數位化數量及訓練及測試樣本數量表

## 2. 訓練作業

本研究採用 Caffe (Convolutional Architecture for Fast Feature Embedding) 深度學習框架[4] 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 [14]環境作為 Host 並在 Linux 環境中安裝 Docker[15]以及 Nvidia-Docker[16]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(圖七)，



圖七、不同網路下 GPU 運行的時間比較圖[7]

因此本研究選擇以 MobileNet-SSD 作為深度學習卷積神經網路作為基礎，針對標記的資料進行訓練。

### (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-SSD 的基礎網路，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)  $\alpha$  來減少輸入輸出的通道數(channels)，輸入層的通道個數  $M$  變成  $\alpha M$ ，輸出層的通道個數  $N$  變成  $\alpha N$ ，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)  $\rho$  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 15,871 張標記完成的照片放到訓練環境中，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 LMDB(Lightning Memory-Mapped Database)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 120,000 次得到 loss 約 0.3-0.5 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 目錄中我們得到每千次的權重檔，具體結果如圖八：

```

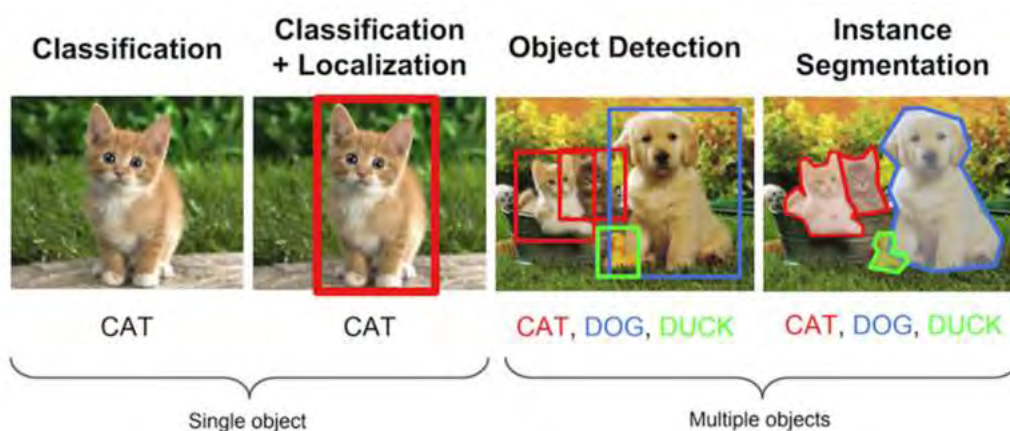
root@yang: /opt/caffe/examples/yang/snapshot
File Edit View Search Terminal Help
root@yang:/opt/caffe# cd examples/yang/snapshot/
root@yang:/opt/caffe/examples/yang/snapshot# ls mobilenet_iter_90*
mobilenet_iter_9000.caffemodel  mobilenet_iter_90000.caffemodel
mobilenet_iter_9000.solverstate  mobilenet_iter_90000.solverstate
root@yang:/opt/caffe/examples/yang/snapshot# ls mobilenet_iter_10*
mobilenet_iter_1000.caffemodel  mobilenet_iter_104000.caffemodel
mobilenet_iter_1000.solverstate  mobilenet_iter_104000.solverstate
mobilenet_iter_10000.caffemodel  mobilenet_iter_105000.caffemodel
mobilenet_iter_10000.solverstate  mobilenet_iter_105000.solverstate
mobilenet_iter_100000.caffemodel  mobilenet_iter_106000.caffemodel
mobilenet_iter_100000.solverstate  mobilenet_iter_106000.solverstate
mobilenet_iter_101000.caffemodel  mobilenet_iter_107000.caffemodel
mobilenet_iter_101000.solverstate  mobilenet_iter_107000.solverstate
mobilenet_iter_102000.caffemodel  mobilenet_iter_108000.caffemodel
mobilenet_iter_102000.solverstate  mobilenet_iter_108000.solverstate
mobilenet_iter_103000.caffemodel  mobilenet_iter_109000.caffemodel
mobilenet_iter_103000.solverstate  mobilenet_iter_109000.solverstate
root@yang:/opt/caffe/examples/yang/snapshot# ls mobilenet_iter_120*
mobilenet_iter_12000.caffemodel  mobilenet_iter_120000.caffemodel
mobilenet_iter_12000.solverstate  mobilenet_iter_120000.solverstate
root@yang:/opt/caffe/examples/yang/snapshot#

```

圖八:本研究訓練結果產生權重檔示意圖

### 3. 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，如下圖九。



圖九、常見的計算機視覺問題(Stanford's CS231n 課程資料)

因此，本研究採用 PASCAL Visual Object Classes(VOC) Challenge 所定義的 11-point interpolated average precision[17]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三。

類別	鐵鋁罐類	廢紙容器類	塑膠瓶類	廢燈炮類	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	0.973

表三、各類垃圾物件 AP 值及 mAP 值





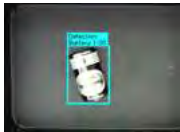







訓練結果得到各類垃圾物件約平均 0.907 ~ 0.998 的 AP 值，此外，mAP 高達 0.973，算是得到不錯的成果，執行合併權重檔的程式後(merge\_bn.py) 我們會得到卷積類神經網路的模型檔 (MobileNetSSD\_deploy.caffemodel )在提取出訓練好的模型後，就可以放在辨識和環境中進行測試。

### (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，結果如下表四：

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
鐵鋁罐類(含變形)	120	113	7	94.1%
正常物件辨識效果		變形物件辨識效果		
				
廢紙容器類(含變形)	120	109	11	90.8%
正常物件辨識效果		變形物件辨識效果		
				
塑膠瓶類(含變形)	120	116	4	96.6%

正常物件辨識效果		變形物件辨識效果		
				
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%
				
紙盒類(含變形)	120	115	5	95.8%
正常物件辨識效果		變形物件辨識效果		
				
廢紙類(變形)	120	117	3	97.5%
				
小計	960	923	37	96.1%

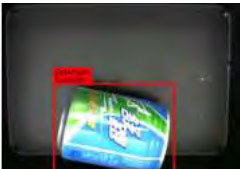
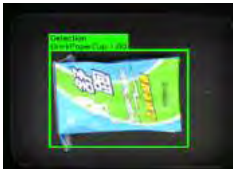

表四、辨識盒安裝在垃圾桶機構上進行實際垃圾丟擲辨識率統計表

分析發現鐵鋁罐類垃圾物件在總數 120 次的測試中，正確辨識出 113 次，而錯誤辨識為 7 次，辨識率為 94.1%；廢紙容器類垃圾物件在總數 120 次的測試中，正確辨識出 109 次，而錯誤辨識為 11 次，辨識率為 90.8%，塑膠瓶類垃圾物件在總數 120 次的測試中，正確辨識出 116 次，而錯誤辨識為其他類別數為 4 次，辨識率為 96.6%，廢燈泡類垃圾物件在總數 120 次

的測試中，正確辨識出 119 次，錯誤辨識僅有 1 次，辨識率 99.2%，廢電池類垃圾物件在總數 120 次的測試中，正確辨識出 116 次，錯誤辨識 4 次，辨識率為 96.6%，玻璃瓶類在總數 120 次的測試中，正確辨識出 118 次，錯誤辨識 2 次，辨識率為 98.3%，紙盒類在總數 120 次的測試中，正確辨識出 115 次，錯誤辨識 5 次，辨識率為 95.8%，廢紙類在 120 次的測試中，正確辨識出 117 次，錯誤辨識 3 次，辨識率 97.5%。

整體測試的辨識率在八類總數 960 次的測試中，正確辨識該類別的垃圾物件共 923 次，整體辨識率為 96.1%。



而在物件的相似性上也具備辨識能力，如下表五：

相似性的辨識效果(不同材質但雷同的印刷圖案)		
鋁罐材質	廢紙容器材質	寶特瓶材質
		

表五、不同材質但雷同的印刷外觀的物件辨識效果表

## (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，提升效果如下表六：

改善前(約 1.04fps)	改善後(約 11fps)
	

表六、效能調整改進前與改進後效果表

### (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，經測試，本分類機構能正確移動垃圾放置盤到垃圾桶定位點再進行翻蓋丟擲，分類機構如下圖十：



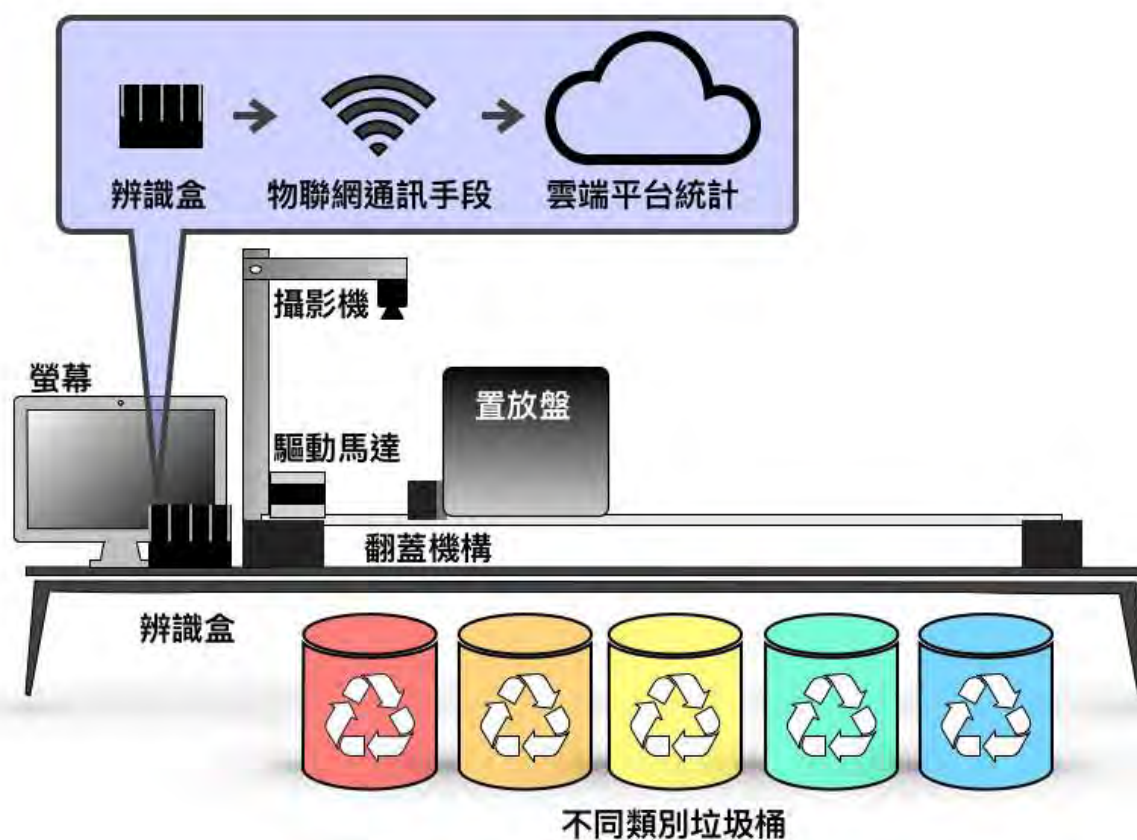
圖十、自動分類機構



## 肆、 研究結果與展望

本研究透過深度學習卷積神經網路來直接對垃圾物件進行辨識，具有良好的效果，其中三類固定外型(未形變)的垃圾物件，辨識率約達 98%，而五類包括變型的垃圾物件也具有約 95%的辨識率，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到該類垃圾桶定位點進行翻蓋動作。

此外，由於垃圾物件均屬於商品包裝的丟棄物，各地流行的商品及包裝方式均不相同，因此，垃圾物件屬於範圍一直在擴充及變化的目標檢測標的，若能透過物聯網手段將收集的照片集中在雲端中訓練，提升模型的辨識能力並透過推送的機制，同步提升分賽在各地垃圾端點辨識盒的辨識能力，如下(圖十一)，就有機會解決新垃圾物件生成而辨識率下降的問題，是值得未來持續研究的方向。



圖十一、未來結合雲端及物聯網技術發展架構圖

## 伍、 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
- [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.

### 網站參考資料

- [8] 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>
- [9] 看德國如何回收各種瓶子。資料取自 <https://www.youtube.com/watch?v=3J3PbeY7ENo>
- [10] 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>
- [11] 韓國資源回收公司。資料取自 <http://superbin.co.kr/new/contents/product.php>
- [12] 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>
- [13] GitHub 影像圖形標記技術論壇。資料取自 <https://github.com/tzotalin/labelImg>

- [14] Ubuntu Linux Kernel 。資料取自 <https://www.ubuntu.com/>
- [15] 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/>
- [16] Github nvidia-docker 技術論壇。資料取自 <https://github.com/NVIDIA/nvidia-docker>
- [17] The PASCAL Visual Object Classes (VOC) Challenge 。資料取自 [http://homepages.inf.ed.ac.uk/ckiwi/postscript/ijcv\\_voc09.pdf](http://homepages.inf.ed.ac.uk/ckiwi/postscript/ijcv_voc09.pdf)
- [18] Github-io-netscope 技術論壇。資料取自 <http://ethereon.github.io/netscope/#/gist/5b1ffa5d22f6ac1f25f30fbd24a716be>
-

## 陸、附錄

### 研究設備及器材

電腦設備		
裝置	型號	規格
MacBook Pro13 蘋果筆記型電腦	A1278	Intel Core 2 Due CPU @2.2GHz 4GB 1066MHz DDR3 RAM ADATA SU650 960GB SSD /13.3 吋 1280 X 800 / Nvidia GeForce 360
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	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

## 【評語】 052508

此延續作品提出一個自動辨識垃圾分類的裝置，透過深度學習卷積神經網路來進行分類辨識，實驗結果顯示採用透過影像來辨識垃圾物件外觀，有良好的效果。此作品使用類神經網路方法來做模型訓練和影像辨識的應用，對類神經網路方法的技術沒有改進而只是使用，在技術深度和探討上不足。

使用影像辨識技術進行垃圾分類，對一些類別的分辨會有盲點，例如要辨別同形狀同顏色但使用金屬或塑膠材質的物品，就不易分類。建議未來在這部分多加改進。

# 基於深度學習卷積神經網路的垃圾分類裝置之研究

## 研究動機

### 1. 將近8成的中學生完全沒有垃圾分類習慣

根據國內學者范宏偉(2012)[1]以苗栗線國民中學學生對於學校的垃圾分類的研究發現高達78.9%的中學生完全沒有垃圾分類習慣。

### 2. 深度學習人工智慧超越人類的技術逐漸成熟

從Alpha Go戰勝圍棋棋王以及人臉辨識超越人類辨識能力都顯示深度學習技術值得驗證是否適合作為垃圾物件辨識的核心技術。

我們希望透過科技的方式來進行垃圾物件的辨識，驗證辨識的可行性後

進一步協助一般人對於垃圾分類的落實與提升其有效性，對於科技運用於垃圾分類的領域盡一份貢獻。

### 3. 以人工智慧或影像辨識來融合垃圾分類的研究相對較少

從台灣博碩士論文網站上48萬3230篇中查詢「垃圾」關鍵字僅972筆其中沒有針對垃圾物件進行影像辨識的研究。

圖一: 在台灣博碩士論文網站查詢關鍵字的示意圖。

## 研究目的

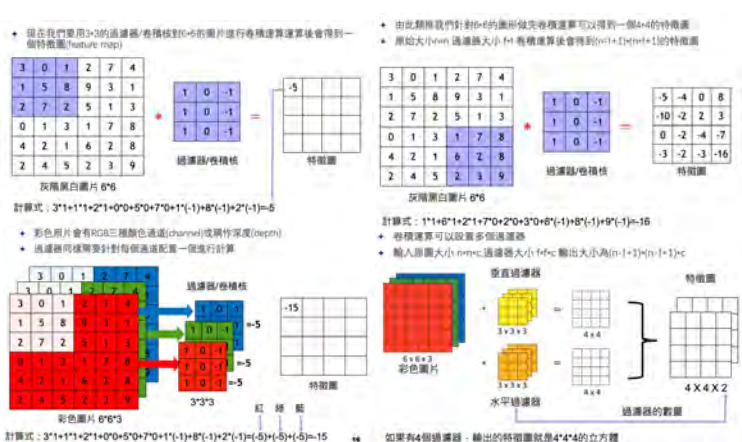
1. 希望結合資訊科技能賦予垃圾桶具備辨識垃圾類別的能力。
2. 希望透過深度學習卷積神經網路作為垃圾類別辨識的核心技術基礎，確認是否適合應用於垃圾物件分類辨識。
3. 具備解決垃圾物件的多樣性、相似性與形變的特性，能有良好且強健的辨識能力
4. 設計一個自動移動翻蓋的機構，驗證辨識盒能整合機構，達到自動分類的效果。

## 現有技術的探討及比較

我們針對目前智能化垃圾分類設備資料整理如下表：

表一：智能化垃圾分類設備的主要辨識技術整理及比較

年度	國家	公司/組織	名稱	主要辨識技術	垃圾形變辨識
2016	德國	Sielaff	Leergut回收桶	辨識回收標誌	不具備
2017	瑞典	Eskilstuna	自動垃圾分類廠	辨識垃圾袋顏色	不具備
2017	韓國	SuperBin	AI回收桶Nephron	辨識垃圾物件	不具備
2018	台灣	大豐回收	ZeroZero無人回收站	辨識條碼	不具備
2019	台灣	本研究	垃圾辨識分類裝置	辨識垃圾物件	具備



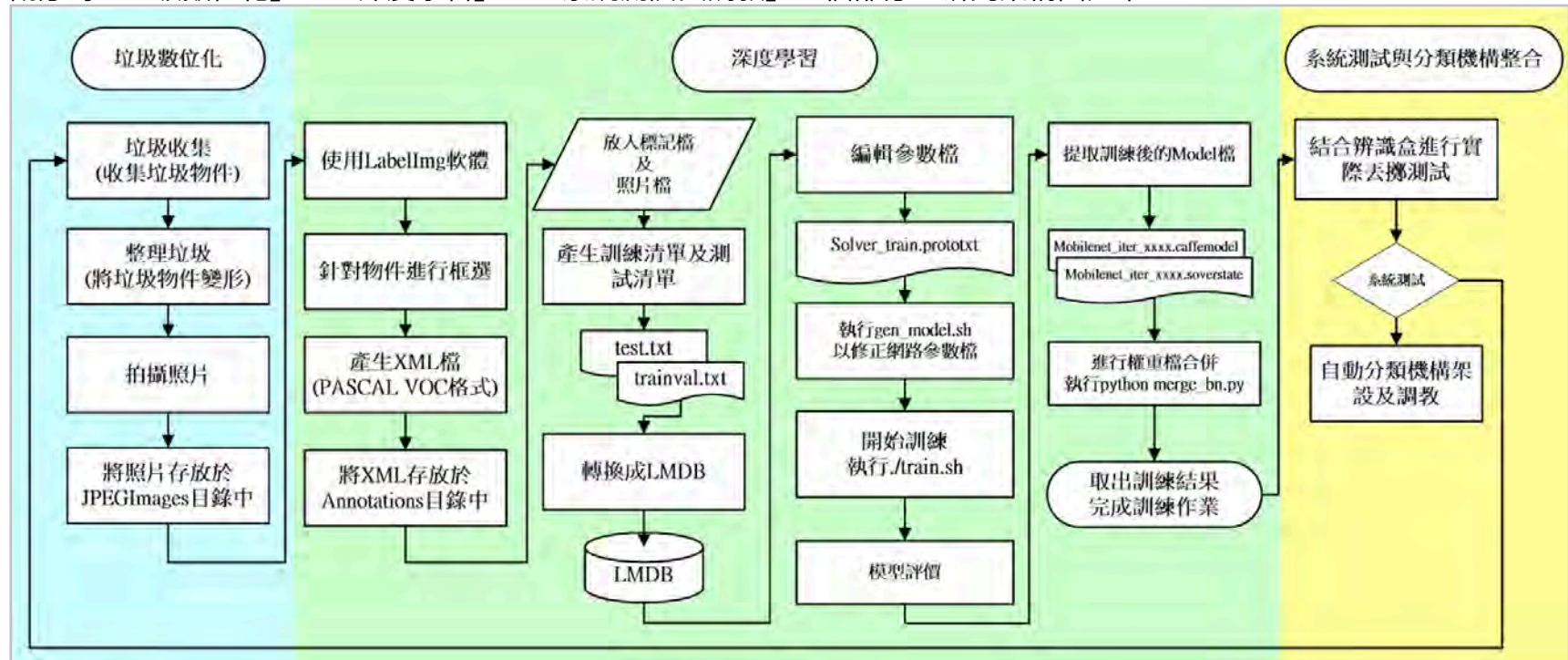
圖二：卷積運算示意圖 (本研究整理)

本研究採用卷積類神經網路作為影像辨識的核心方法，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。

## 研究方法

### 研究架構：

由於缺乏前人直接透過影像辨識垃圾物件(包括形變)類別的研究可供參考，本研究定義下列的研究架構，概分為「垃圾數位化」、「深度學習」、「系統測試與調教」三個部分，研究架構圖如下：



圖三：研究架構流程圖

由於辨識垃圾物件具備一定的困難度，最少需克服下列問題：一、垃圾物件的多樣性。二、垃圾物件的相似性。三、垃圾物件的形變。研究以三大步驟來進行(圖三)，(一)垃圾數位化，目的是透過垃圾物件數位化的過程來進行數位資料的蒐集。(二)深度學習，針對數位化資料進行標記，再經過神經網路進行訓練。(三)系統測試與分類機構整合，達到我們的目標後就進行最後階段的實體測試。再將訓練好的模型及程式安裝在辨識盒中，針對實際的垃圾進行辨識，並結合自動分類機構。

表二：本研究所選擇的垃圾類別

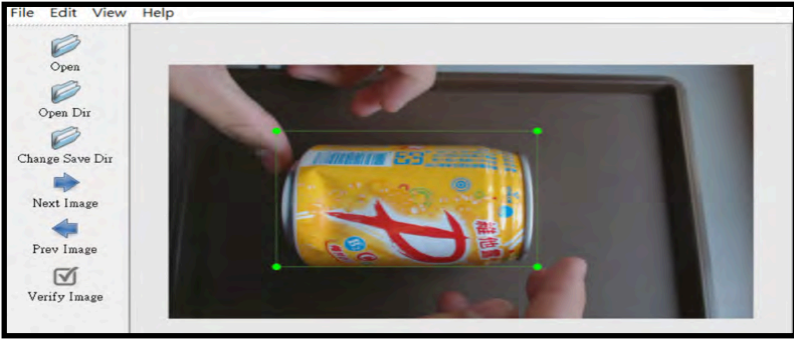
鐵鋁罐類	塑膠瓶類	廢紙容器類	廢燈泡類	廢電池類	紙盒類	玻璃瓶類	廢紙類

## 深度學習：

1.標記作業：本研究將採用 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 [13] 來針對相片進行標記，如下圖（圖四）；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。

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 15,871 張照片，進行標記，標記後區分為訓練樣本及測試樣本，相關數量如下表三。

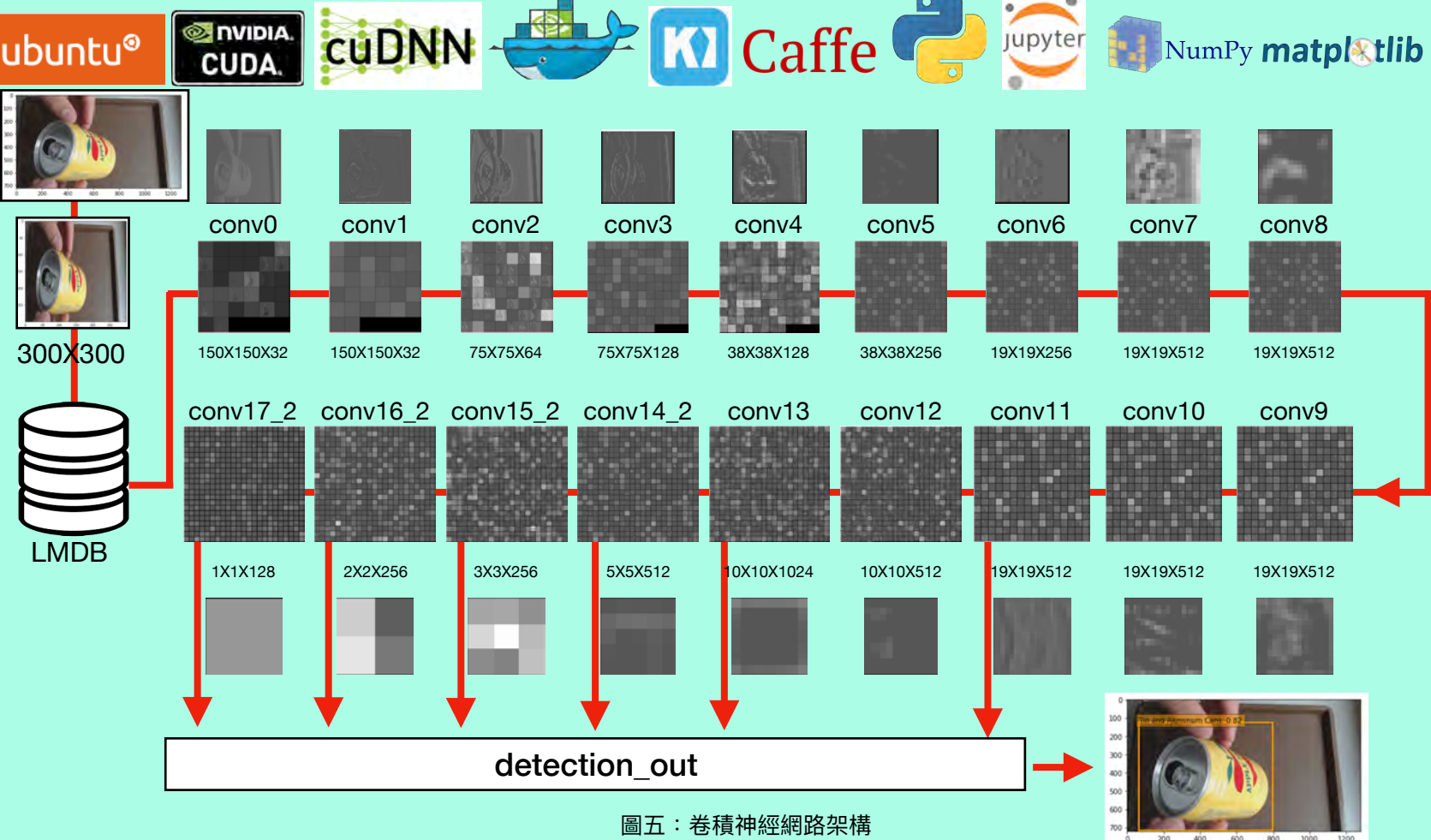


圖四：使用LabelImg進行標記的示意圖

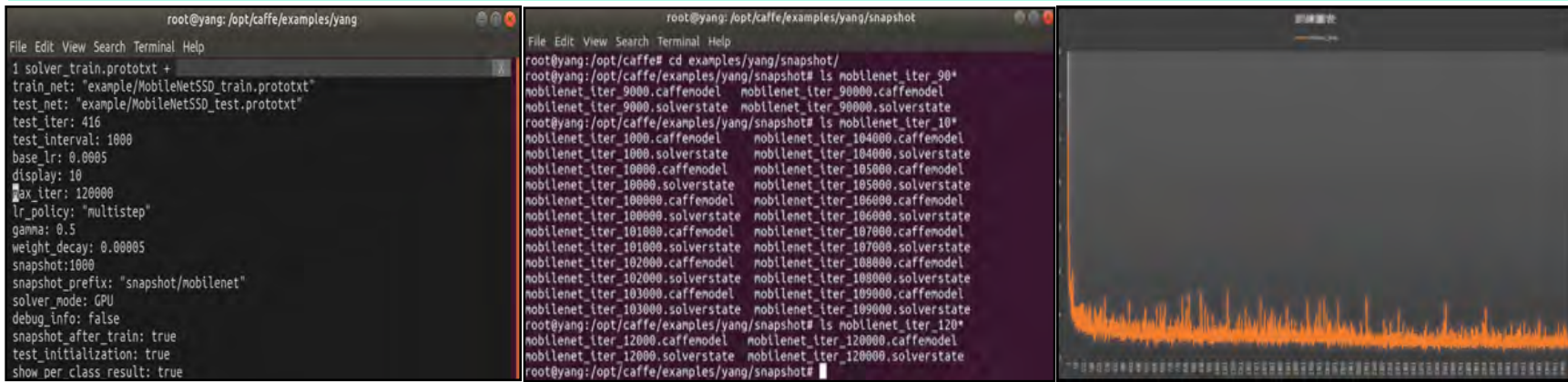
表三：訓練樣本及測試樣本相關數量表

搜集的垃圾物件類別	數位化照片數	訓練樣本數	測試樣本數
鐵鋁罐類	2,572	2,186	386
廢紙容器類	4,894	4,160	734
塑膠瓶類	4,287	3,644	643
廢燈泡類	653	540	95
廢電池類	739	628	111
玻璃瓶類	901	766	135
紙盒類	1,266	1,076	190
廢紙類	577	491	86
小計	15,871	13,491	2,380

2.本研究所採用的卷積神經網路架構：



圖五：卷積神經網路架構



圖六：設定訓練參數檔示意圖

圖七：訓練產生權重檔示意圖

圖八：訓練期間loss值變化示意圖

損失函數：

損失函式分為兩個部分：定位損失和分類損失，公式如下：

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。

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

## 系統測試-模型評價

模型評價：在二元分類中，常使用精確度(Precision)及召回率(Recall)來評價統計效果，不過目標檢測(Object Detection)在影像辨識除了分類的正確與否，還包括圖像中物體的定位，而且除了單一物件外，也包括多物件偵測因此不能單純透過精確度(Precision)及召回率(Recall)來評價。

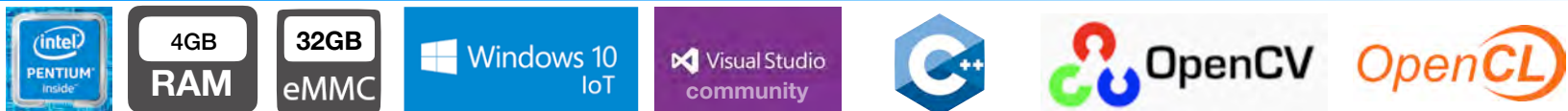
本研究採用PASCAL Visual Object Classes(VOC) Challenge 定義的11-point interpolated average precision[17]來衡量模型的評價，下表為各類垃圾物件AP(Average precision)值及mAP值(mean Average precision)

表三：各類垃圾物件AP(Average precision)值及mAP(mean Average precision)值

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，執行合併權重檔的程式後(merge\_bn.py) 我們會得到卷積類神經網路的模型檔(yang\_deploy.caffemodel)在提取出訓練好的模型後，就可以放在辨識和環境中進行測試。

# 系統測試及調教



鐵鋁罐		廢紙容器		塑膠瓶		廢燈泡		紙盒類		玻璃瓶		電池類		廢紙類					
120次		120次		120次		120次		120次		120次		120次		120次					
113	O	109	O	116	O	119	O	116	O	118	O	115	O	117	O				
7	X	11	X	4	X	1	X	4	X	2	X	5	X	3	X				
94.1%		90.8%		96.6%		99.2%		95.8%		98.3%		96.6%		97.5%					
累計				全部丟擲次數				辨識正確數				辨識錯誤數				平均辨識率			
				960				923				37				96.1%			

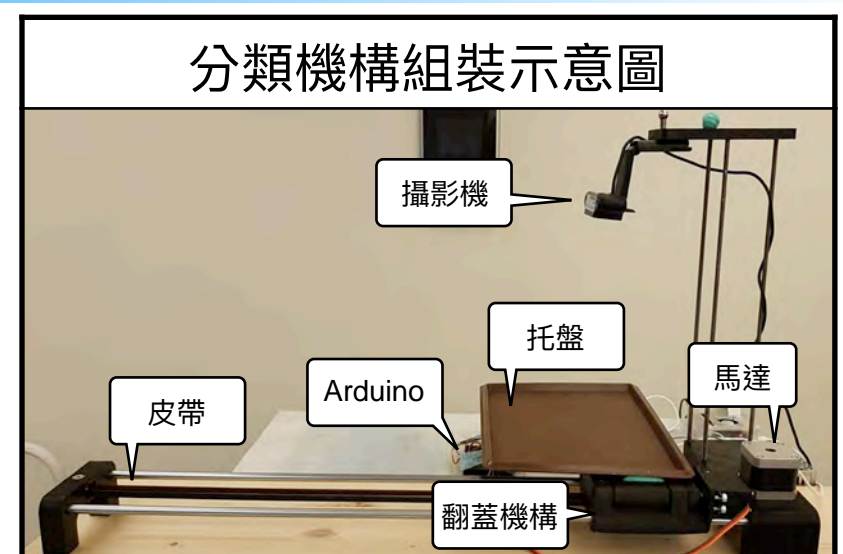
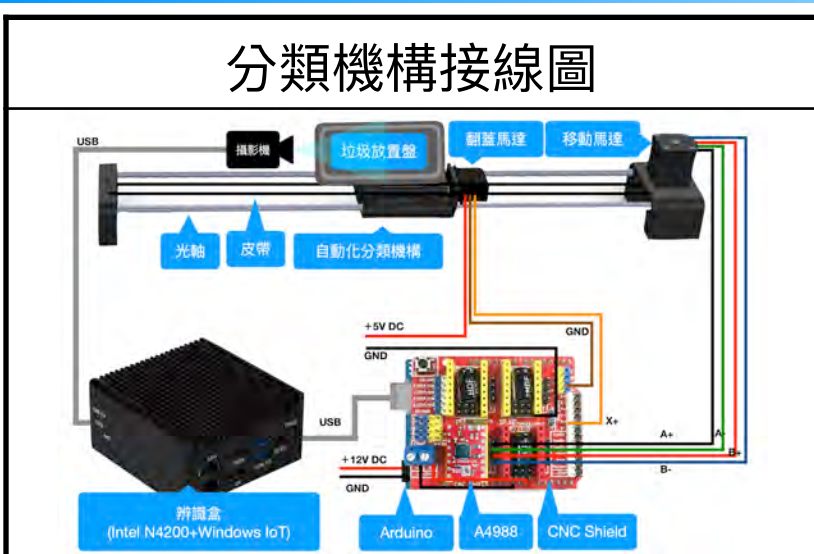
修改程式透過GPU加速提升辨識速度

CPU辨識(約1.04fps)	GPU辨識(約11fps)

不同材質但印刷雷同之物件辨識效果(相似性)

鐵鋁罐	廢紙容器	塑膠瓶

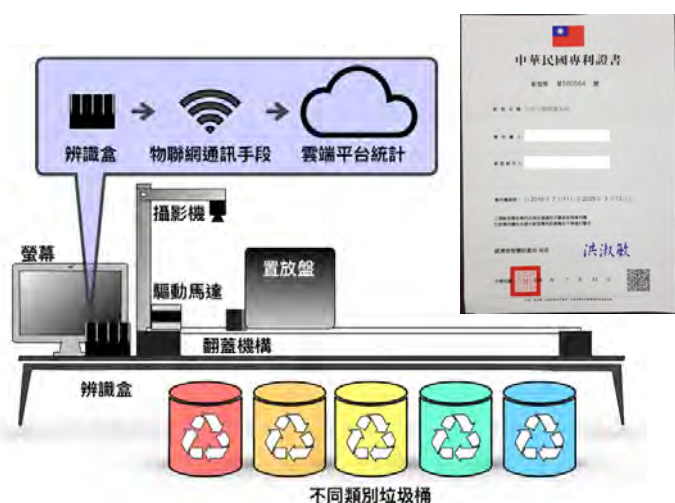
# 分類機構的整合



# 結論及未來展望

- 結論：
- 本研究證實卷積神經網路可作為垃圾物件辨識之核心技術。
  - 研究結果三類固定外型垃圾辨識率約98%，五類含變形垃圾辨識率約95%，平均辨識率約96.1%。
  - 整合自動化分類機構能控制馬達移動到該類定位點進行翻蓋動作。

- 未來展望：
- 結合IoT物聯網通信手段收集推陳出新的垃圾物件到雲端平台進行訓練再將新辨識模型派送到Edge端的辨識盒能解決垃圾多樣性的問題。
  - 申請「垃圾分類辨識系統」已獲中華民國專利，未來希望有機會落實技術授權或商品化。



# 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
- [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.
- 網站參考資料
- [8] 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>
- [9] 看德國如何回收各種瓶子. 資料取自 <https://www.youtube.com/watch?v=3i3PbeY7ENo>
- [10] 李天怡, 瑞典寶特瓶回收系統超威! 快速辨別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>
- [11] 韓國資源回收公司. 資料取自 <http://superbin.co.kr/new/contents/product.php>
- [12] 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>
- [13] GitHub 影像圖形標記技術論壇. 資料取自 <https://github.com/tzatalin/labelImg>
- [14] Ubuntu Linux Kernel. 資料取自 <https://www.ubuntu.com/>
- [15] 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/>
- [16] GitHub nvidia-docker 技術論壇. 資料取自 <https://github.com/NVIDIA/nvidia-docker>
- [17] The PASCAL Visual Object Classes (VOC) Challenge. 資料取自 [http://homepages.inf.ed.ac.uk/ckiw/posts/script/ijcv\\_voc09.pdf](http://homepages.inf.ed.ac.uk/ckiw/posts/script/ijcv_voc09.pdf)
- [18] GitHub-io-netscope 技術論壇. 資料取自 <http://ethereon.github.io/netscope/#/gist/5b1ffa5d22f6ac1f25f30fbd24a716be>