

結合機器視覺與深度學習之金屬圓柱表面缺陷檢測系統

Machine Vision and Deep Learning Based Defect Inspection System for Cylindrical Metallic Surface

蘇詠靖、游原璋、何昭慶*

Eugene Su, Yuan-Wei You, Chao-Ching Ho*

自 2012 年加拿大 Alex Krizhevsky 所領軍的團隊贏得 ImageNet 的 Large Scale Visual Recognition Competition 冠軍開始，深度學習領域掀起了新一波人工智慧的浪潮，相關的產業與研究團隊都競相投入此深度學習的研究領域。本論文著墨於探討深度學習是否適用於自動光學檢測 (AOI) 領域與產業、分析光學檢測的問題是否可直接套用現行的深度學習模型、及分析相較於語音識別領域、自然語言處理領域、醫療或其它電腦視覺領域，應用深度學習方法於光學瑕疵檢測的研究論文仍相對欠缺不足的原因。本文並舉金屬圓柱表面 (如高爾夫球桿) 的瑕疵檢測應用為實例，來探討深度學習應用於自動光學檢測瑕疵之方法，以及當結合機器視覺與深度學習技術時，可能遭遇的問題與克服方案。

Since the Alex Krizhevsky team won the championship of ImageNet Large Scale Visual Recognition Competition in 2012, deep learning has set off a new wave of artificial intelligence field, with all industries and research groups competing for deep learning field. However, the open question is whether the deep learning is the magic solution in a fairy tale or not. In this paper, we discuss how to employ the deep learning in the field of automated optical inspection industry, and analyze the skills of directly applying the current deep learning model to the defect detection based on the optical inspection method. Compared to the deep learning in the fields of speech recognition, natural language processing, medical or other computer vision applications, few literatures are found in discussing the defect detection based on deep learning. Therefore, this paper offered the solution of applying the deep learning in the field of defect detection for golf clubs based on the optical inspection method.

一、前言

自動光學檢測領域 (automated optical inspection, AOI) 結合了包含光學、機械、電機和資訊等多種技術領域，其廣泛的應用於各種產業或相關的產品，但由於每個產業或產品都有不同的檢測特性，

例如集成電路及電子產業要求極低的漏判率 (false negative rate)、而應用於金屬鋼鐵產品表面時，則易受照射光源的影響產生表面反光，採光源設計可以加強某些特徵，但同時也消除了一些特徵，使得圖像包含的特徵更少，這些檢測應用都是現行電腦視覺的深度學習模型沒有考慮到的特性。

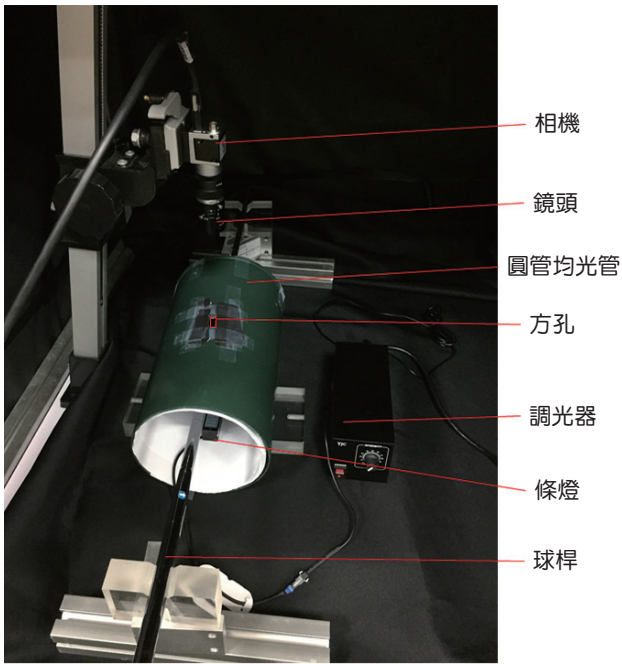


圖 1. 光源系統架構。

綜觀現行深度學習於光學檢測的研究有，Daniel Weimer⁽¹⁾ 等人使用深度卷積神經網路訓練和測試 Weakly Supervised Learning for Industrial Optical Inspection 資料集⁽²⁾，測試結果的平均準確率達到 99.2%，其檢測效果遠比其它的非卷積神經網路方法等優越。

Xavier Gibert⁽³⁾ 等人使用深度卷積神經網路來辨識鐵軌的各個元件材質，包含道渣、木頭軌枕、三種不同表面風化程度的混凝土軌枕、及兩種不同破裂程度的混凝土軌枕、潤滑器、鋼軌和扣件等十種類別，其準確率為 93.35%，但採用其它非神經網路的方法最高準確率為 82.05%。Xavier Gibert⁽⁴⁾ 等人在此網路架構上加入了丟棄法的正規化方法 (dropout regularization) 和權值衰減的技巧 (weight decay) 來防止過擬合，可進一步將判斷準確率提高到 95.02%。

以圖像識別的深度學習模型為例，絕大部分都是以前 ImageNet 的資料集為基礎，來開發設計衍生的模型，並再將衍生的模型直接套用在光學檢測的應用上，但模型的套用上難免會衍生映射不吻合的情況發生，因此亦需去探討如何將根據深度學習的基礎資料所開發設計的模型衍生到 AOI 的應用領域。上述四個瑕疵檢測應用⁽¹⁻⁴⁾，都是基於深

度卷積神經網路的方法，檢測效果確實優於其它的方法，因此對於圖像複雜，傳統方法 (如模板法等) 處理不好的問題，深度卷積神經網路是可以嘗試的方向。

本文使用 Yuan-Wei You 與 Chao-Ching Ho⁽⁵⁾ 等人的高反射高爾夫球桿為檢測資料，其檢測系統採用圓管光源，並以 101 層的 ResNet⁽⁶⁾ 訓練此資料集，此檢測結果顯示，本系統能辨識背景與六種高爾夫球桿的瑕疵，並且可標示出瑕疵的種類與影像位置。

二、系統架構

1. 光源

本研究選用條狀光源與圓管均光板作為燈源設置。如圖 1 所示為本研究採用的球桿瑕疵檢測系統，此系統採 2592 × 1944 像素大小的相機、與鏡頭焦距 50 mm 組成，使用光學倍率為 0.52、相機視野大小為 11 mm × 8 mm，影像解析度為 4.23 μm。圖 1 中的方孔是為了在圖像中產生黑色的條紋遮罩。

圖 2 為本文實驗所使用的藍色條狀光，圖 3 為進行本實驗時條狀燈源架設的示意圖；而相關的檢測瑕疵有：圖 4(a) 為條狀光照射於球桿的刮傷瑕疵、圖 4(b) 為球桿的擦傷瑕疵、圖 5(a) 為球桿的連續孔瑕疵、圖 5(b) 為球桿的白霧瑕疵、圖 6(a) 為球桿的針孔瑕疵，圖 6(b) 為球桿的殘膠瑕疵。可以從圖 4 到圖 6 看出每個球桿瑕疵區域像素值皆

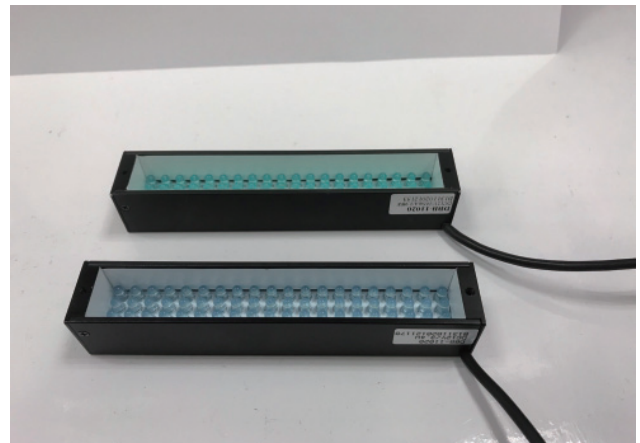


圖 2. 條狀光源。

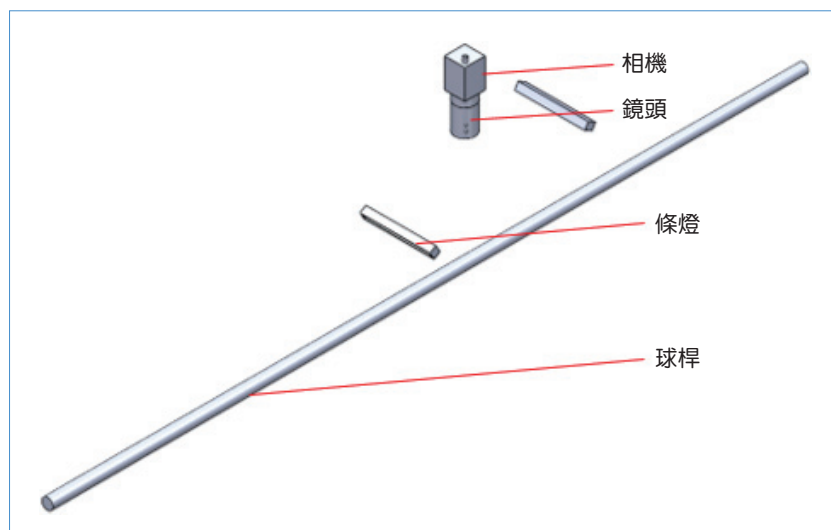


圖 3. 實驗示意圖。

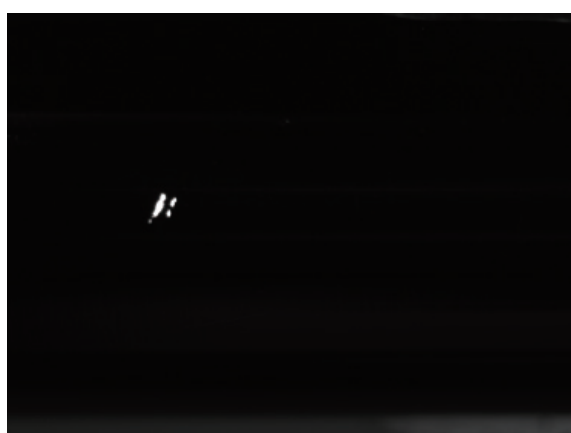


(a)

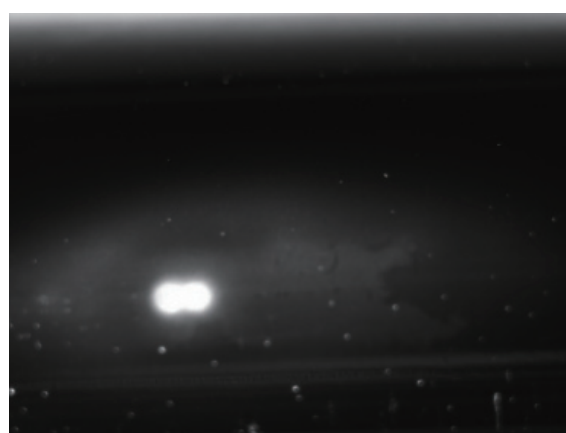


(b)

圖 4. 條狀光源於球桿瑕疵實驗：(a) 刮傷圖、(b) 擦傷圖。

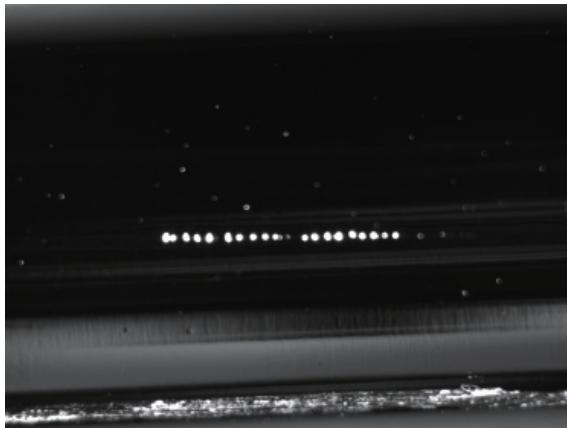


(a)

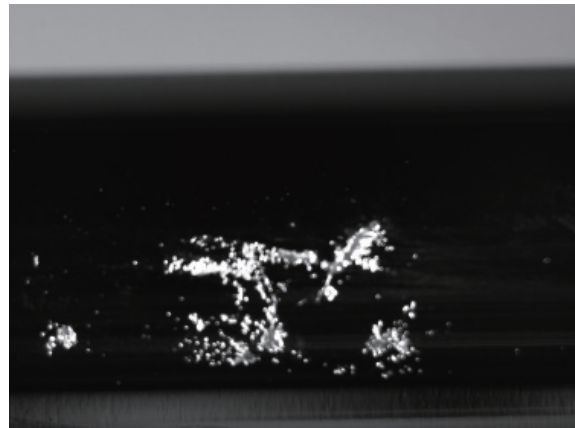


(b)

圖 5. 條狀光源於球桿瑕疵實驗：(a) 連續孔圖、(b) 白霧圖。



(a)



(b)

圖 6. 條狀光源於球桿瑕疵實驗：(a) 針孔圖、(b) 殘膠圖。

比鄰近區域高，然而採用條狀光源的缺點會有，如圖 6(a) 中具有一條帶狀高反光區域，在高反光區域中無法判別此處為瑕疵造成的高反光或是球桿鏡面所造成的高反光。

因此本研究為解決條狀光源實驗中條狀高反光問題，使用一全白的圓管內壁，可將條狀光源放置於圓管內，使光源能在圓管內壁中可均勻反光。圖 7 為圓管經均光設計後的實驗示意圖，並在圓管上方開一個相機視野大小的方孔，條狀光會在圓管內均勻反射，散射光線會通過方孔穿出，因此在圖像上會有一條狀區域為黑條紋遮罩，這時當照射到瑕疵時，會呈現一高反光時圖像，其與黑條紋遮罩呈現高對比，此黑條紋遮罩大小跟方孔與球桿的相對距離有關，當相對距離愈近時，此灰條紋遮罩越大，然而當球桿為錐形圓管時，此黑條紋遮罩會隨檢測球桿的不同位置，亦會呈現有不一樣的尺寸。

圖 8 為圓管經均光實驗所量測出的刮傷成像圖，可以看出刮傷與黑條紋遮罩呈現高對比，然而擷取的圖像顯示比採用條狀燈源實驗圖像呈現較少的雜點。圖 9(a) 為圓管均光實驗的擦傷成像圖，圖 9(b) 為圓管均光實驗白霧成像圖，圖 10(a) 為圓管均光實驗針孔成像圖，當採用 8 位元的亮度編碼時，針孔瑕疵其亮度值接近於 255，故與採黑條紋遮罩，其像素值為 0 時，可形成高對比。圖 10(b) 為圓管均光實驗所呈現的殘膠缺陷成像圖。觀察圖 8 到圖 10 中圖像的複雜度遠比採用條狀光源實驗簡單，採用圓管均光也一併解決了條狀光源實驗中所呈現的高反光區域問題。

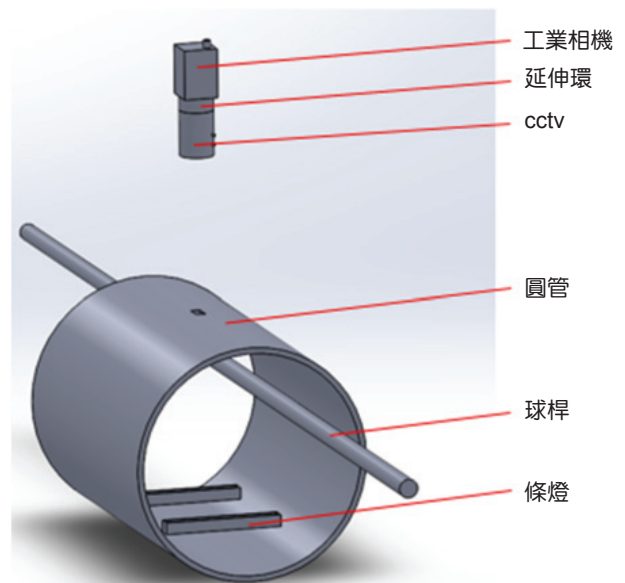


圖 7. 圓管均光示意圖。

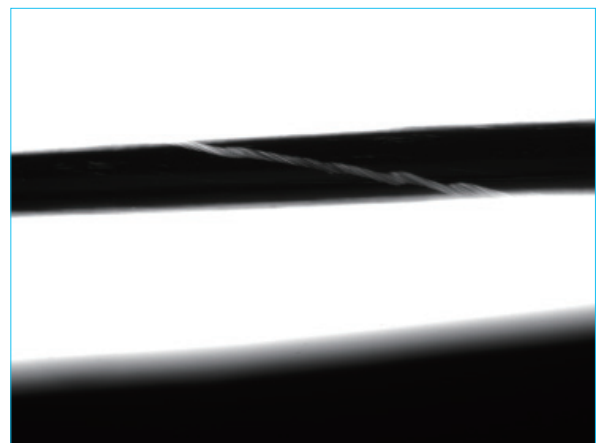
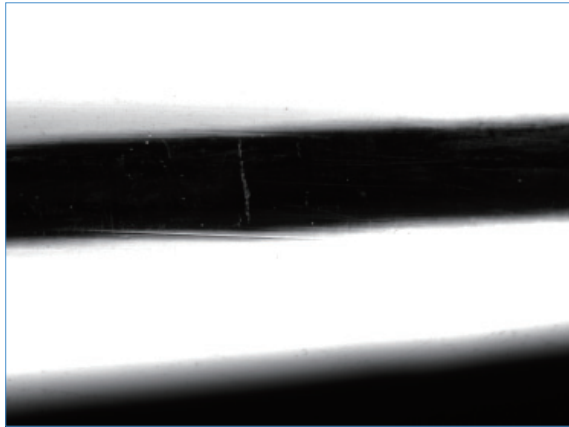
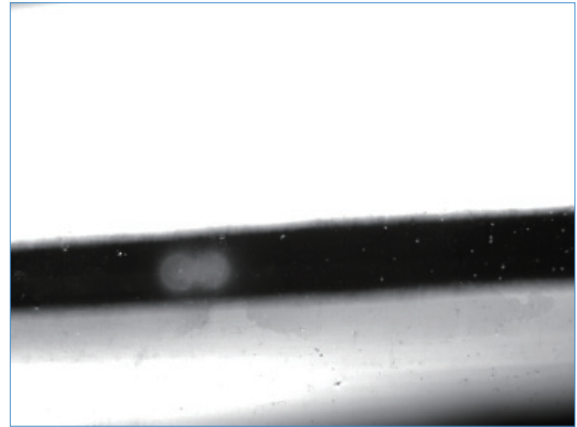


圖 8. 均光實驗刮傷圖。



(a)



(b)

圖 9. 圓管均光於球桿瑕疵實驗：(a) 擦傷圖、(b) 白霧圖。

2. 瑕疵分類與檢測

在為數眾多用於圖像分類的深度卷積神經網路中，本研究選擇 2015 年 ImageNet 圖像識別競賽的冠軍 ResNet，ResNet 的殘差架構可有效解決梯度消失 (vanishing gradients) 與梯度爆炸 (exploding gradients) 問題、Google 在 GitHub 的開源專案：Models 中提供完整的 ResNet 預訓練模型與高階函數庫、許多論文都有提供 ResNet 的比較數據，這讓 ResNet 形成一個完善的生態圈，即使是 ResNet 的改良版 DenseNet，各項效能都比 ResNet 好，但目前為止，也無法像 ResNet 可提供完整的檢索資源。

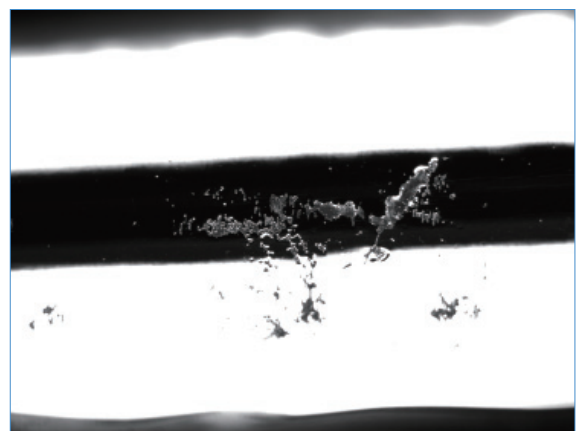
圖 11 是 Google 開源專案 Models 的 Resnet 實作架構，與 ResNet 的原作者在論文⁽⁶⁾ 裡提供的網

路架構有一些差異，原作者版本的最後一層神經網路是使用全連接層，但 Google 開源專案 Models 版本的最後一層是使用 1×1 的卷積神經網路，來達到全連接層的效果，其參數比全連接層少，因此其實作技巧的表現相對良好。

圖 12 是本文使用於高爾夫球桿瑕疵分類的網路架構，與圖 11 的唯一差別是在最後一層卷積神經網路之後，再增加向上採樣 (up-sampling) 的設計，向上採樣是由一組放大、填充和卷積的運算組合，目的是希望瑕疵檢測時能更精準地標記瑕疵的位置。雖然反卷積 (deconvolution) 亦是一種向上採樣的方法，但本文不是像素級別 (pixel-wise) 的圖像分割 (segmentation)，也不是要重建構原圖，增加反卷積層會增加硬體資源的成本與檢測時間，所



(a)



(b)

圖 10. 圓管均光於球桿瑕疵實驗：(a) 針孔圖、(b) 殘膠圖。

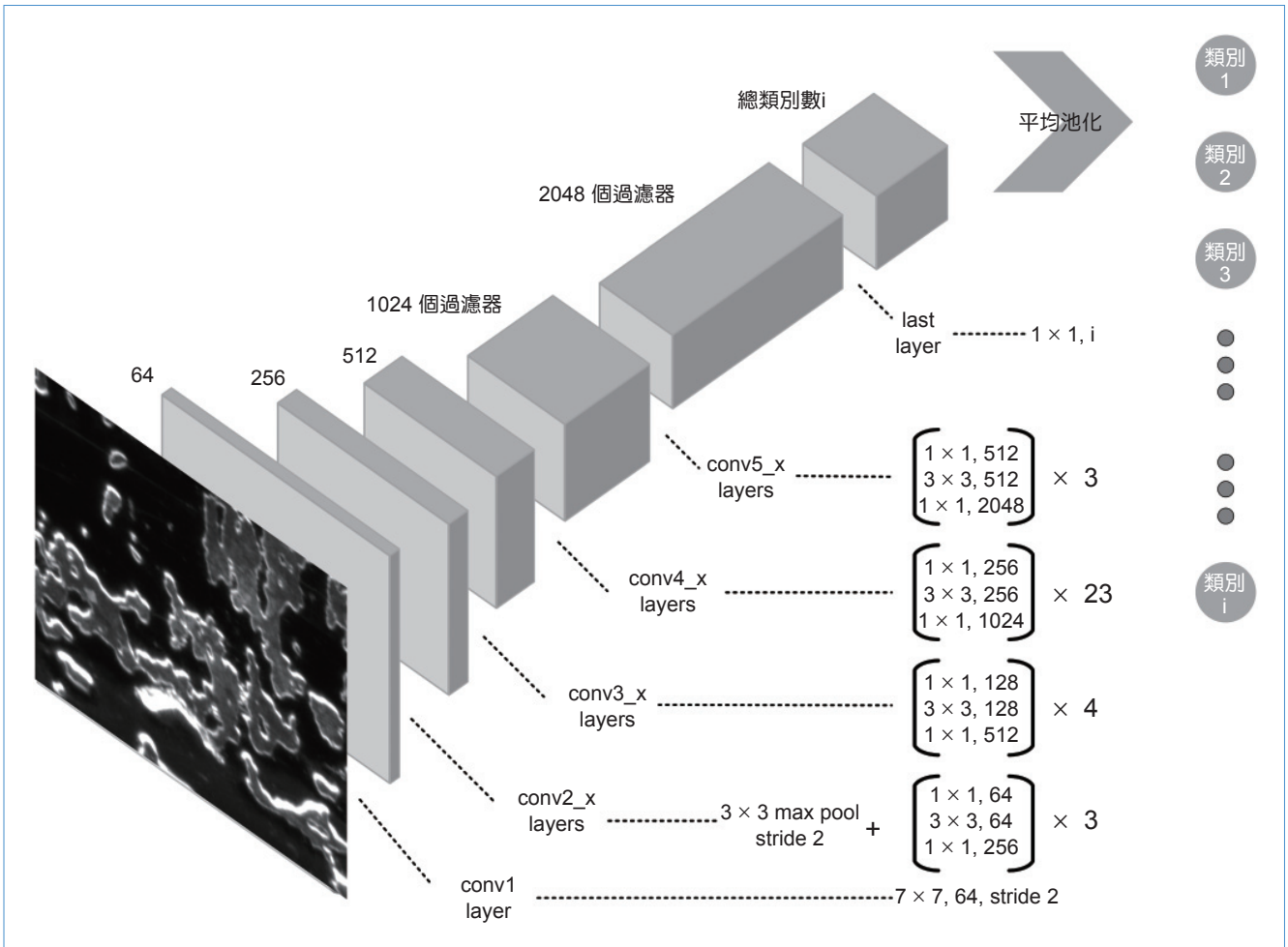


圖 11. ResNet 網路架構。

以反卷積不是本文考慮的方法。

Augustus Odena 等人⁽⁸⁾ 實驗發現以插值法放大圖片，再加上一層卷積運算的效果接近於反卷積運算，也就是本論文裡所提到的放大與卷積 (resize-convolution) 的運算組合。本論文的插值法放大特徵圖運算，乃是在最後一層卷積神經網路之後，採過濾器個數等於類別個數，因此瑕疵檢測的瑕疵類別個數不會像 ImageNet 有 1000 類多的個數，因此記憶體的增加量還在容許的範圍內。

本論文的瑕疵檢測不使用 RCNN 系列⁽⁹⁻¹¹⁾、SSD⁽¹²⁾ 或 YOLO⁽¹³⁾ 之類的架構，主要原因尚有：

1. 邊界範圍 (bounding box) 的標記 (labelling) 方式運用到本論文的資料集有其難度。
2. 主流目標檢測 (object detection) 架構的執行時間不符合需求，以 NVIDIA 1080Ti 顯示卡推論

(inference) 一張 2592×1944 的圖片，其運算時間會超過 10 秒。

所以本文的瑕疵檢測的策略，採取將大圖切割為小圖的分段檢測策略，切割方式如圖 13 所示。

三、實驗方法

1. 資料集

高爾夫球桿條狀與圓管光源資料集⁽⁵⁾ 共包含了 14 種瑕疵，本論文僅訓練與測試連續孔、刮傷、擦傷、白霧、針孔與殘膠等六種瑕疵，並且在高爾夫球桿條狀與圓管光源資料集內，只選取包含這六種瑕疵的原圖共 155 張。且每張原圖可能包含不只一種瑕疵，有些區域甚至重疊兩種瑕疵，當採用這些原圖上裁剪感興趣區域 (ROI)，當作瑕疵分

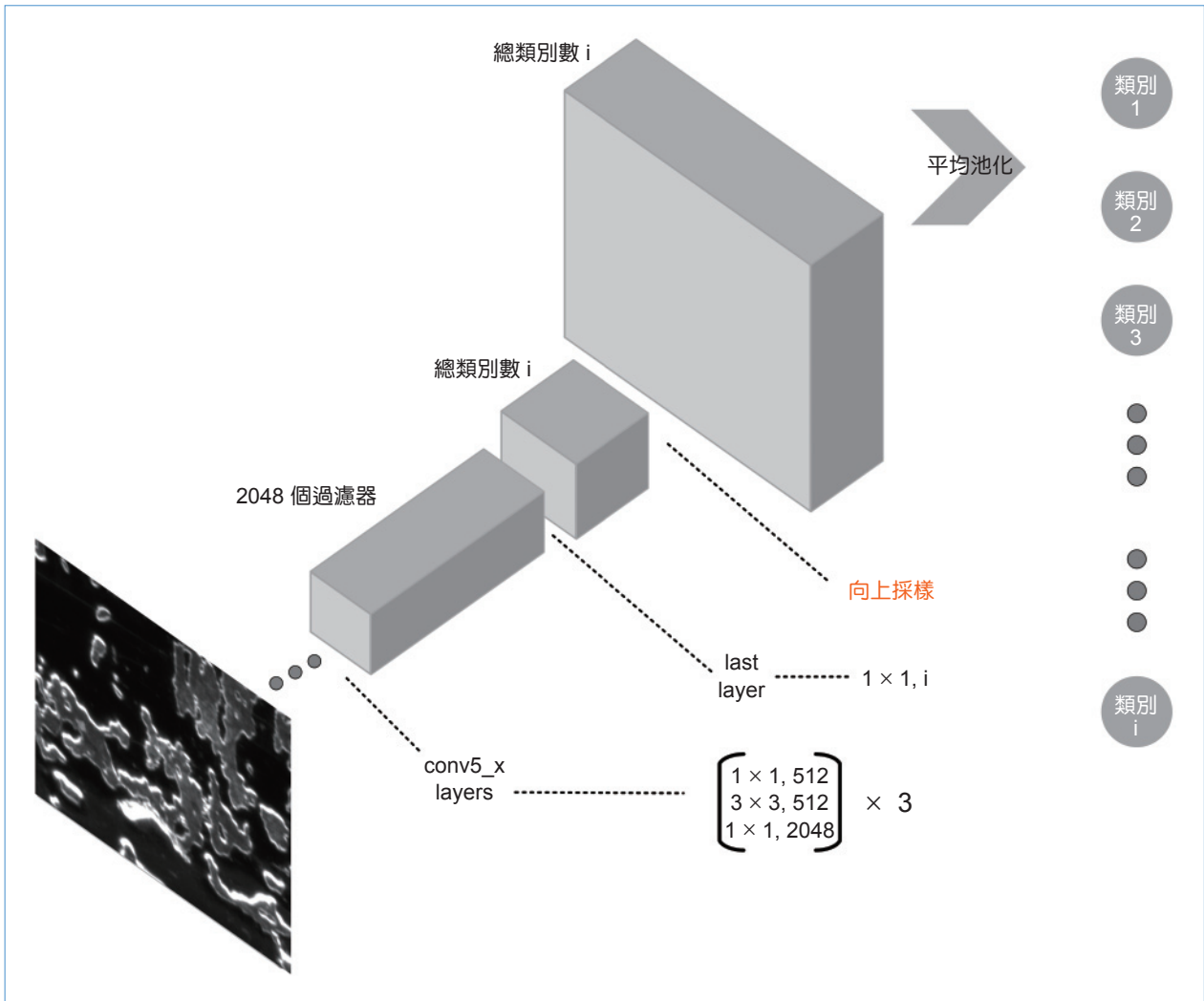


圖 12. 本論文使用的網路架構。

類器的訓練與測試資料時，會遭遇下列的問題：不同瑕疵種類的大小差距太懸殊、針孔瑕疵小到只有 25×25 像素、大範圍的殘膠瑕疵像潑墨畫一樣佈滿整張圖、區域為 2580×630 像素的刮傷瑕疵也很難標記。

引此本論文以直覺、簡單和系統化的方式，來處理訓練與測試資料標記問題，將寬與高為 $W \times H$ 像素的原圖，以步幅 S ，裁剪大小為 $P \times P$ 像素的片段 (patch)，組成共有 M 張片段，當裁剪時遇邊界問題的解決方法，其處理方式如圖 13 所示。從 M 張片段中挑選出包含瑕疵的片段共 N 張圖，其中每個瑕疵種類各有 $n_1, n_2, n_3, \dots, n_i$ 張圖，隨機選取 n_i 的 80% 當作訓練集 (train set) 資料，剩下的 20% 當作驗證集 (validation set) 資料。

本論文的原圖尺寸為 2592×1944 像素，並以步幅 56，裁剪大小為 224×224 的片段，並人工挑選出 12363 張有瑕疵的片段，以及 2887 張非瑕疵的片段，其中 12207 張片段做為訓練資料，3043 張片段當作驗證資料。為了避免實驗數據不公正，測試集 (test set) 資料不從這 15250 張挑選，另外以不同比例的長方形標記出 546 張，其中 326 張是有瑕疵的片段，220 張是非瑕疵的片段。

2. 訓練方法

本論文使用 TensorFlow 開源軟體庫開發深度學習網路，與對應的訓練的分類器，本方法除了能分類六種瑕疵，還能辨識非瑕疵的背景圖像，辨識背景圖像能幫助瑕疵檢測的運作。訓練過程分為

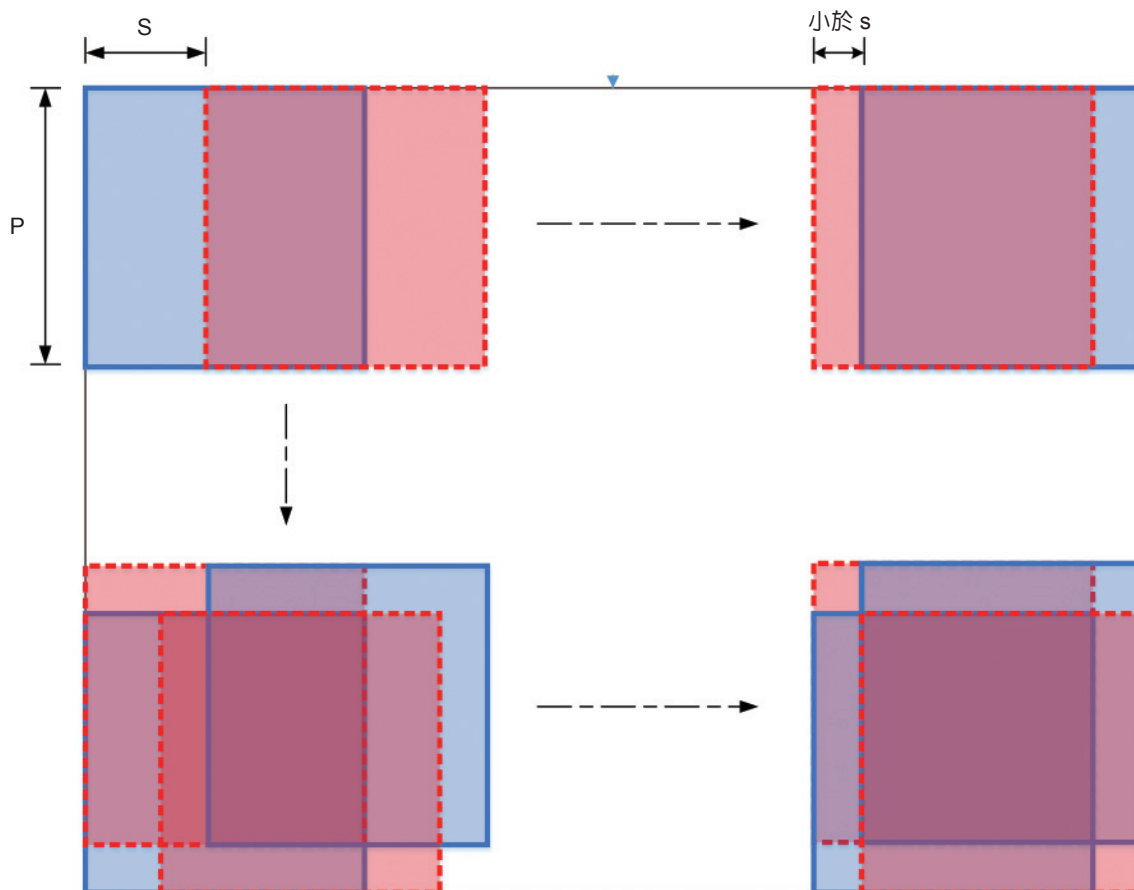


圖 13. 裁剪片段示意圖。

兩個步驟，第一個步驟使用 ResNet V1 101 的預訓練模型做遷移學習，只訓練與更新最後一層卷積神經網路層的參數，其餘網路層的參數都固定不動，批次處理大小 (batch size) 設定為 32，學習率設定為 0.005，權值衰減設定為 0.0004，總共訓練 10000 步。第二個步驟訓練與更新所有網路層的參數，批次大小 (batch size) 設為 32，學習率設定為 0.0002，權值衰減設定為 0.0004，總共訓練 40000 步。

本論文中的圖片預處理，運用了隨機調整圖片大小、鏡射、上下翻轉、亮度、對比、色彩飽和度與色相 (hue) 等一般常見的數據增強 (data augmentation) 方法，但本文並未使用隨機裁剪 (crop)，隨機裁剪並不適合本文使用的資料集。由於相機與檢測物的距離是固定的，光源也是固定的，隨機調整圖片大小的範圍不需要太大，以免浪費圖形處理器的資源與拉長訓練時間；亮度、對比、色彩飽和度、色相的範圍不能跟處理 ImageNet

圖片時一樣，處理範圍要限制更小，以免抹除圖片上的瑕疵特徵。

3. 瑕疵檢測

瑕疵檢測的切割原理與標記資料集的原理一樣，但作法和參數略有不同，首先將需要檢測的大圖四周填充黑邊，讓原始大圖的長寬都是 224 的倍數，再將原始大圖以步幅 224 的切割方式，裁剪成許多 224×224 的片段小圖，並以訓練好的瑕疵分類器，來批次推論這些片段小圖，當檢測為背景的小圖時不需要後處理；但檢測為瑕疵的小圖，需將對應的瑕疵類別之特徵圖標準化 (normalization) 成灰階圖，如圖 14 的右上 (b) 是左上片段小圖 (a) 的灰階特徵圖，右下 (c) 是左下 (d) 的灰階特徵圖，兩者皆是二值化前的特徵圖。二值化灰階圖讓黑的像素更暗，白的更亮，亮點區域就是瑕疵的所在位置，再將灰階圖轉換成彩色圖，並且將亮點區域依不同瑕疵種類取代為不同顏色，最後將彩色片段小

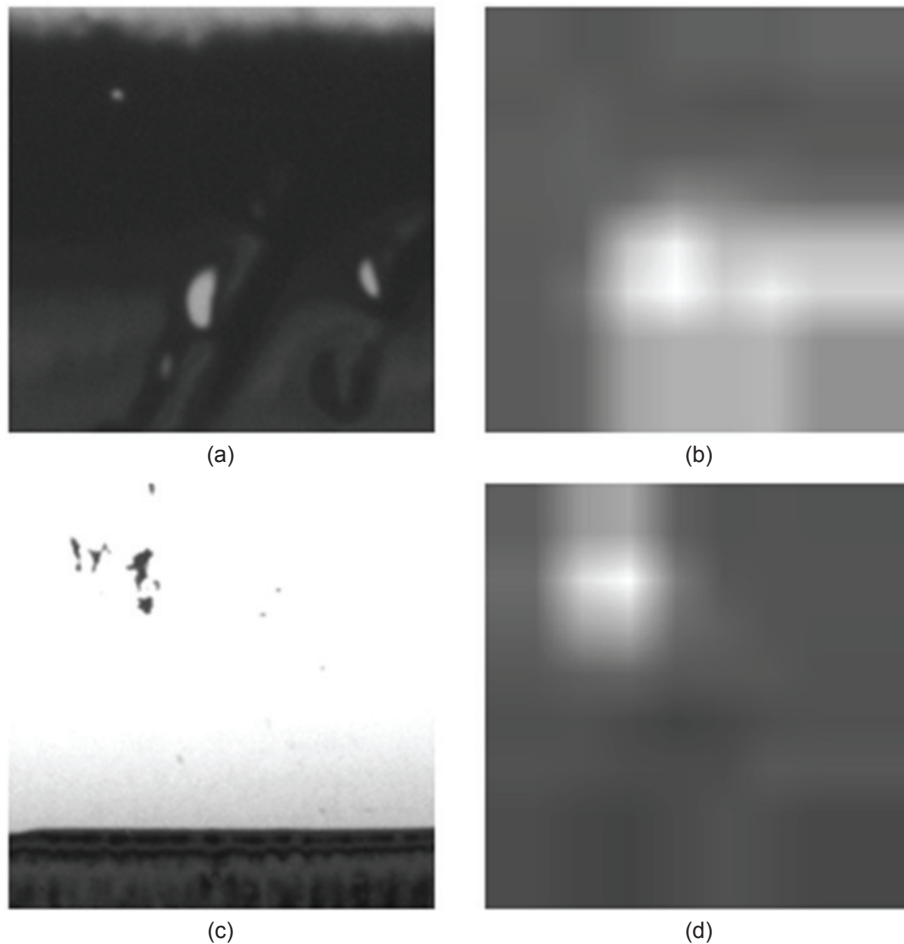


圖 14. 特徵圖。

圖放置回原位置，如圖 15 所示，許多片段小圖像拼圖一樣，拼疊出瑕疵的種類與位置圖。

四、實驗結果

1. 瑕疵分類

圖 16 是本文分類器效能的混淆矩陣 (confusion matrix)，實驗結果顯示：連續孔和針孔的錯誤率偏高，推估原因是連續孔和針孔相較於其它四類瑕疵，在面積上小很多，部分連續孔因存在的位置與拍攝的角度，造成反光而沒有凹陷感，因此容易與雜點混淆；部分殘膠瑕疵因殘膠量少，中間突起部分沒有產生陰影，變成一個實心的小亮點，不容易與針孔區分開。圖 17 是 ResNet 分類器效能的混淆矩陣，本文分類器的效能略勝於 ResNet。

2. 瑕疵檢測

圖 18 的右上 (b) 是左上原圖 (a) 的瑕疵種類與位置圖，本方法可正確檢測出瑕疵的種類與位置，淡綠色表示檢測出的瑕疵種類是刮傷；右下 (d) 的瑕疵種類與位置圖顯示的是淡灰色，表示瑕疵種類是殘膠。圖 19 右上 (b) 是 ResNet 架構的瑕疵種類與位置圖，淡黃色表示為白霧缺陷，左下 (c) 是有使用向上採樣的瑕疵種類與位置圖，很明顯左下 (c) 的效果比較好。

當步幅 S 等於片段小圖 P 的缺點是：當一個 $P \times P$ 的片段小圖若存在一種以上的缺陷，也只能檢測出其中一種缺陷；但優點是檢測速度快，若以 NVIDIA 1080 Ti 檢測一張 2592×1944 的大圖，檢測時間只需大約 0.9 秒就可完成。

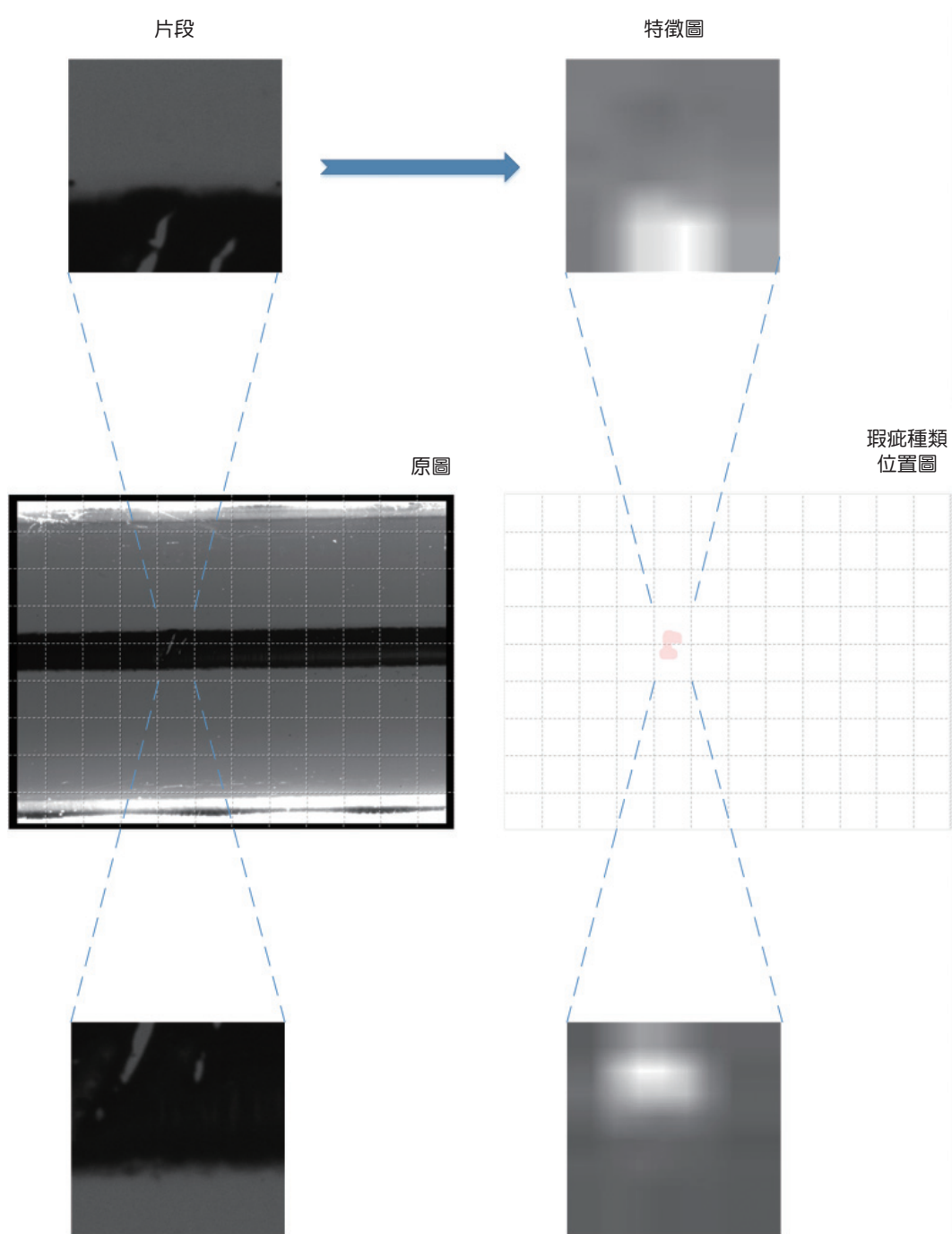


圖 15. 瑕疵檢測方法示意圖。

五、結論

由於金屬表面容易產生反光，即使是人工檢測的錯誤率也不低，這類的問題本身就很難處理，尤其還要機器做多種類的檢測與分類，當使用人工提取特徵和影像處理的方法解決這類的問題非常難，

這樣的難題很適合應用深度學習技術。深度學習顧名思義就是要學習，學習一定要有資料，瑕疵資料不易取得是自動光學檢測產業難解的問題之一，再厲害的深度學習網路架構，若沒有適當的資料也無用武之地，所以如何有效的取得更多訓練資料，亦是一需要研究思考的課題，但使用深度學習技術，

		預測的類別						
%		背景	連續孔	刮傷	擦傷	白霧	針孔	殘膠
真實的類別	背景	91.36	1.36	4.55	1.36	0	0	1.36
	連續孔	14.26	85.71	0	0	0	0	0
	刮傷	0	0	98.53	1.47	0	0	0
	擦傷	0	0	0	100	0	0	0
	白霧	3.12	0	0	0	96.88	0	0
針孔	0	0	1.49	1.49	0	77.61	19.4	
殘膠	2.16	0	0	4.35	0	0	93.04	

圖 16. 本文架構的實驗結果。

		預測的類別						
%		背景	連續孔	刮傷	擦傷	白霧	針孔	殘膠
真實的類別	背景	90	1.36	4.55	1.36	0	0.45	2.27
	連續孔	9.52	80.95	4.76	0	0	4.76	0
	刮傷	0	0	98.53	1.47	0	0	0
	擦傷	0	0	0	95.65	0	0	4.35
	白霧	6.25	3.12	0	0	90.62	0	0
針孔	8.96	0	0	0	0	83.58	7.46	
殘膠	5.22	0	0	0	0	0	94.78	

圖 17. ResNet 架構的實驗結果。

所需的資料量到底為何，是本論文探討的重點之一，若以 ImageNet 的資料集為例，每一個類別的訓練資料大約 1300 張圖片，這並非是難以達到的數量，但 ImageNet 提供了上千種類的資料集才是我們難以達到的目標，目前沒有像 ImageNet 一樣的機構可以提供質量好的瑕疵資料集，因此自動光學檢測應用深度學習技術的第一步，定是先要準備適當的資料集，其資料集的品質乃取決於光學設計架構，當自動光學檢測系統的取像能藉由光學架構來保留更多物理特徵時，定可達到更高的檢測正確率。

參考文獻

1. D. Weimer, B. Scholz-Reiter, M. Shpitalni, *CIRP Annals*, **65** (1), 417 (2016).
2. Please refer to the web site: <https://hci.iwr.uni-heidelberg.de/node/3616>.
3. X. Gibert, Vishal M. Patel, R. Chellappa, "Material classification and semantic segmentation of railway track images with deep convolutional neural networks", *IEEE International Conference*, 2015.
4. X. Gibert, V. M. Patel, R. Chellappa, "Deep multitask learning for railway track inspection", *IEEE Transactions on Intelligent Transportation Systems*, **18** (1), 153 (2017).

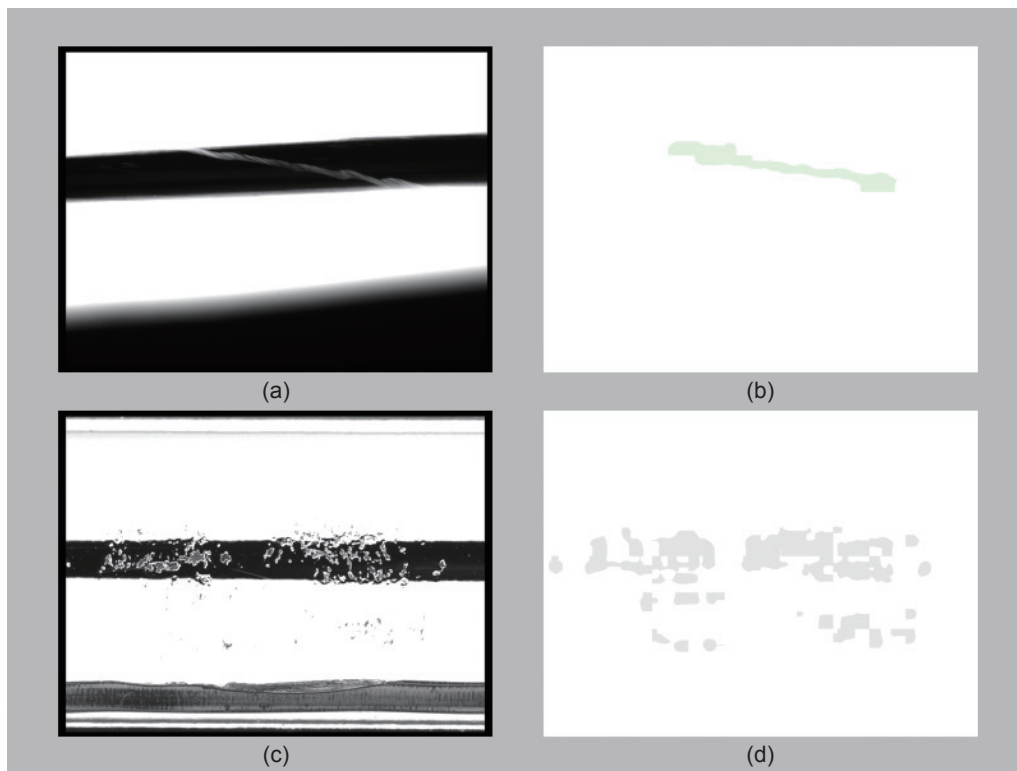


圖 18. 瑕疵種類與位置圖。

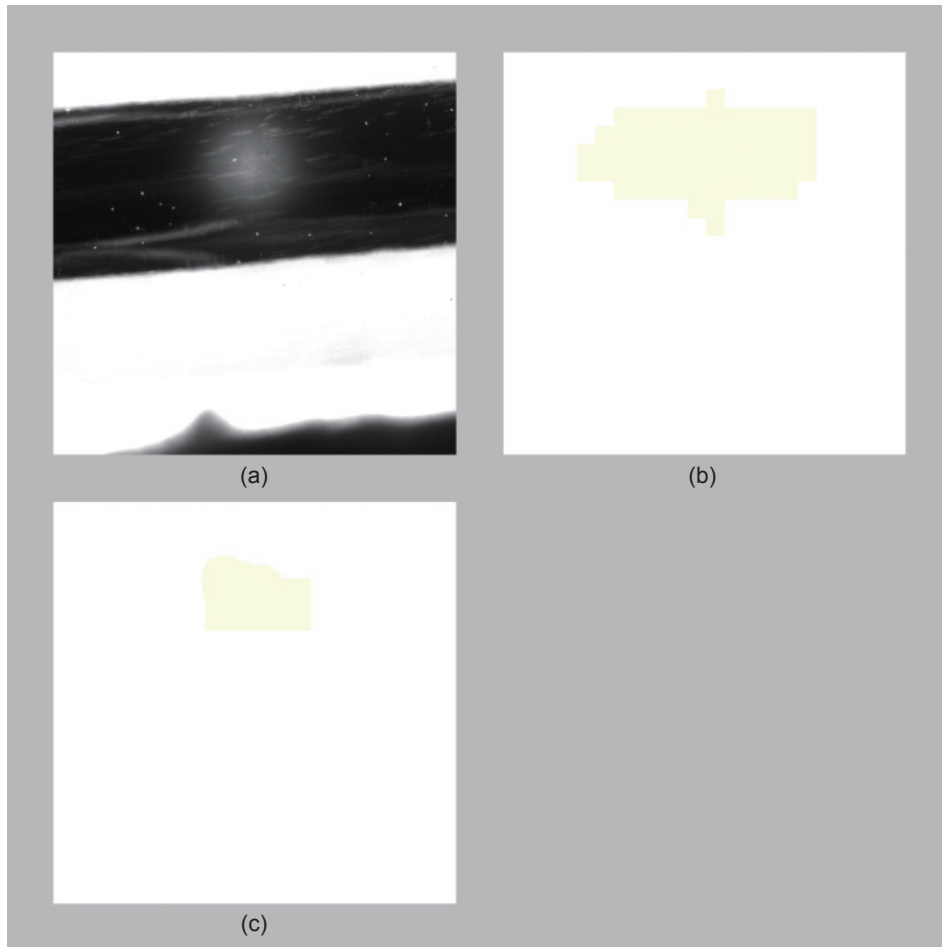


圖 19. 瑕疵位置圖比較。

5. 游原璋, 何昭慶, “基於機器視覺之高爾夫球桿瑕疵檢測”, 中國機械工程學會第三十四屆全國學術研討會, 2017 年.
6. K. He, X. Zhang, S. Ren, J. Sun, “Deep Residual Learning for Image Recognition”, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
7. Please refer to the web site: <https://github.com/tensorflow/models>.
8. A. Odena, V. Dumoulin, C. Olah, “Deconvolution and Checkerboard Artifacts,” *Distill*, 2016.
9. R. Girshick, J. Donahue, T. Darrell, J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation”, *IEEE conference on computer vision and pattern recognition*, 2014.
10. R. Girshick, “Fast R-CNN”, *IEEE International Conference on Computer Vision*, 2015.
11. S. Ren, K. He, R. Girshick, J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks”, *Advances in neural information processing systems*, 2015.
12. W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu,

A. C. Berg, “SSD: Single Shot MultiBox Detector”, *European Conference on Computer Vision (ECCV)*, 2016.

13. J. Redmon, A. Farhadi, “YOLO9000: Better, Faster, Stronger”, *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.



蘇詠靖先生為臺灣科技大學電子工程系碩士，現為臺灣大學先進自動化光學檢測設備研發聯盟軟體顧問。

Eugene Su received his M.S. in electronic and computer engineering from National Taiwan University of Science and Technology. He is currently a software consultant of Advanced Automated Optical Inspection Equipment Association at National Taiwan University.



游原瑋先生為國立臺北科技大學機械
北科技大學機械工程所的碩士班學
生。

Yuan-Wei You is currently a M.S. student
in the Department of Mechanical
Engineering at National Taipei University of Technology.



何昭慶先生為臺灣科技大學電機工程
博士，現為國立臺北科技大學製造科
技研究所暨機械工程系副教授。

Chao-Ching Ho received his Ph.D. in
electrical engineering from National
Taiwan University of Science and Technology. He is
currently an associate professor in the Graduate Institute
of Manufacturing Technology and Department of
Mechanical Engineering at National Taipei University of
Technology.