

# 基於深度殘差學習理論的改良式快速區域卷積神經網絡之夜間市區車種行人辨識系統

專題編號：107-CSIE-S002

執行期限：106年第1學期至107年第1學期

指導教授：陳彥霖 教授

專題參與人員：104360098 梁皓鈞

104590043 吳建毅

## 一、摘要

使用大部份用於物件偵測的公開數據集所訓練出來的模型在夜間環境下或燈光嚴重不足的情況下會有非常不滿意的表現。本研究提出的解決方案包括收集夜間車種行人數據、標記數據及針對夜間及燈光不足所改良的 Faster R-CNN 模型用於夜間車種行人辨識系統。本系統可以達到 15-17 FPS 及 0.8377 mAP 的辨識表現。

**關鍵詞：**載具辨識、深度學習、黑暗環境、卷積神經網絡、殘差結構、照明不足

## 二、目前本專題研究成果

(科技部大專學生研究計劃及研討會論文)

本實務專題目前研究成果取得科技部大專學生研究計劃，編號為 MOST-107-2813-C-027-017-E，亦已於 2018年8月19日在中華民國影像處理與圖形識別學會 (Chinese Image Processing and Pattern Recognition Society, IPPR) 的第31屆電腦視覺、圖學暨影像處理研討會 (Conference on Computer Vision, Graphics, and Image Processing, CVGIP) 進行口頭發表。

## 三、緣由與目的

神經網絡及深度卷積神經網絡 (DCNN, CNN) [1, 2] 帶領我們在各項機器學習任務取得非常優秀的成果。可是我們很少會看到是在夜間或照明嚴重不足的情況下的展示。在測試過用 PASCAL VOC 2007 [3] 所訓練的 Faster R-CNN [4]，我們發現這些表現十分不理想。這些可能是由於物件在夜

間沒辦法很明顯地顯示出他們的外型及重要特徵難以被捕捉。這些問題都在前人的研究有提出過 [5, 6]。

## 四、數據標記及夜間車種行人辨識系統

目前我們已收集約一萬筆夜間數據，並且根據 PASCAL VOC Challenge [3] 的慣例進行標記。本研究的類別包括腳踏車 (bike)、公車 (bus)、轎車 (car)、機車 (motorbike)、人 (person)、箱型車 (van) 及貨車 (truck)。數據都是在台北市夜間進行收集的，有些是在平常市區照明環境下，有些是在嚴重缺乏照明的黑暗環境下收集。其中針對重疊物件、佔有較少像素的物件及模糊物件有特別的標記處理。但主要離不開考量有關合理的大小及單獨可接受的模糊物件。

本研究使用一個典型的 Faster R-CNN [4] 用於系統中，由於我們所收集的數據量沒有十分大，不足以訓練整個神經網絡，因此我們採用經由 ILSVRC-2012-CLS [8] 預訓練的 VGG16 [9] 及 ResNet101 [10] 進行實驗及改良。根據圖1，我們固定了紅色部份的參數，並且只訓練藍色部份的參數。

## 五、實驗、比較與總結

如表2所示，使用 VGG16 [9] 的模型可以達到 0.6570 mAP 而使用 ResNet101 [10] 的模型可以達到 0.8377 mAP。圖2與圖3的左直排是由 PASCAL VOC 2007 [3] 所訓練的表現，而右直排是由本研究系統的表現。可以看到本研究的系統的表現在平常的夜

間市區及極端的黑暗環境下辨識表現都十分優秀。本研究發現及驗證了一些夜間偵測辨識的一些特性及行為。我們的系統可以在正常夜間市區照明下的嚴重模糊物件，而在極端黑暗環境下，甚至連車頭尾燈都沒有開的情況下，需要相對比較清楚的照片進行偵測與辨識。此外，透過深度卷積神經網絡所開發的夜間車種行人偵測辨識系統比起傳統極度依賴車頭尾燈的方法 [5, 6] 更加可以捕捉到物件的外型。

表2: 驗證集的 mAP 表現。

| Classes    | Average Precision (AP)        |                                   |
|------------|-------------------------------|-----------------------------------|
|            | VGG16<br>VOC2007  <br>VOC2010 | ResNet101<br>VOC2007  <br>VOC2010 |
| Bike       | 0.6031 0.6206                 | 0.7738 0.7993                     |
| Bus        | 0.7474 0.7747                 | 0.8630 0.8970                     |
| Car        | 0.7879 0.8018                 | 0.8172 0.8618                     |
| Motor-bike | 0.6827 0.6901                 | 0.7243 0.7774                     |
| Person     | 0.4877 0.4920                 | 0.6799 0.6839                     |
| Truck      | 0.7243 0.7522                 | 0.9039 0.9236                     |
| Van        | 0.5656 0.5725                 | 0.9091 0.9212                     |
| Mean       | 0.6570 0.6720                 | 0.8102 0.8377                     |

## 參考文獻

- [1] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Neural Information Processing Systems 2012*, 2012.
- [2] Y. LeCun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel, "Backpropagation Applied to Handwritten Zip Code Recognition," *Neural Computation*, vol. 1, no. 4, pp. 541-551, 1989.
- [3] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn and A. Zisserman, "The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Results," 7 April 2007. [Online]. Available: <http://www.pascal-network.org/challenges/VOC/voc2007/workshop/index.html>. [Accessed 21 June 2018].
- [4] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2017.
- [5] Y. L. Chen, C. J. Lin, C. J. Fan, C. M. Hsieh and B. F. Wu, "Vision-based nighttime vehicle detection and range estimation for driver assistance," in *IEEE International Conference on Systems, Man and Cybernetics*, Singapore, 2008.
- [6] Y. L. Chen, B. F. Wu, H. Y. Huang and C. J. Fan, "A Real-Time Vision System for Nighttime Vehicle Detection and Traffic Surveillance," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 2030-2044, 2011.
- [7] M. Everingham, S. M. Ali Eslami, L. Van Gool, C. K. I. Williams, J. Winn and A. Zisserman, "The Pascal Visual Object Classes Challenge: A Retrospective," *International Journal of Computer Vision*, vol. 111, no. 1, pp. 98-136, 2015.
- [8] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C.

Berg and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *arXiv:1409.0575*, 2014.

- [9] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *International Conference on Learning Representations*, San Diego, 2015.
- [10] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 2016.

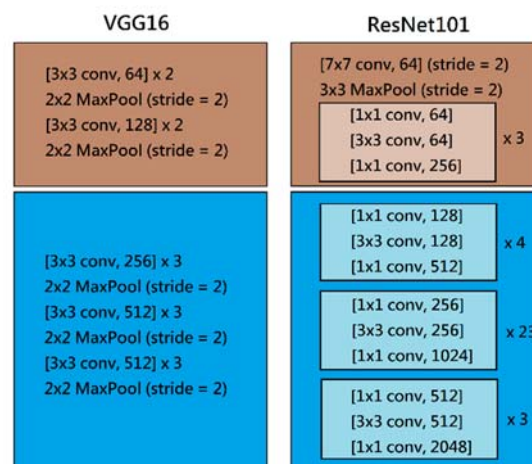


圖1: 本研究所使用的預訓練 CNN。

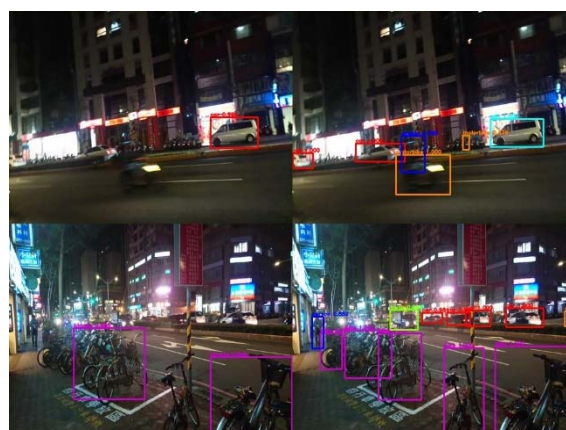


圖2: 平常市區夜間的辨識比較。



圖3: 黑暗環境下的辨識比較。