



應用 影像語意分割 技術 於 鋼筋間距查驗

紀乃文／國立臺灣科技大學營建工程系 專案助理教授
莊仕杰／國立臺灣大學土木工程學系電腦輔助工程組 碩士
陳翊翔／國震中心與台大土木合設AI中心 研究助理
陳鵬元／原國震中心與台大土木合設AI中心 研究助理
陳俊杉／國立臺灣大學土木工程系 教授

鋼筋檢驗在鋼筋混凝土構造物的施工過程當中是一個重要的環節，其檢查要項包含鋼筋之號數、間距、形式（含彎鉤與搭接）等等。傳統鋼筋檢驗仰賴人力完成，並且以局部抽驗的方式為主。但許多鋼筋籠必須在施工現場組立，使得查驗人員無法避免在危險的工作場域對鋼筋進行查驗。在人工智慧蓬勃發展的當下，本研究提出一種基於影像語意分割的技術，配合景深攝影機所取得的 RGB-D 深度影像，能夠以鋼筋籠上縱橫交錯的結點為辨識目標，並因此計算出結點與結點之間的距離來代表鋼筋間距。結果顯示其誤差可達工地現場實用之等級，並提供鋼筋檢驗自動化的可行參考方向及示範。

背景

在台灣的建築物與大型構造物（如橋樑、隧道、水壩等等）當中，鋼筋混凝土構造物（以下簡稱為 RC 構造物）佔了相當大的比例。而在 RC 構造物的施工過程當中，鋼筋組立的品質影響結構安全甚鉅。鋼筋組立的成果在混凝土澆置之後便無法再加以檢視及修正，因此鋼筋查驗（Rebar inspection）便成了 RC 構造物施工過程當中重要的檢查程序。鋼筋組立必須依循著施工圖當中所標註的各項參數來加以完成，包含鋼筋號數、形狀（如彎鉤或搭接）、尺寸、間距等重要資訊。而鋼筋的組立為了配合結構安全之需求，又或預留機電管線之空間，對應著不同的功能需求，會衍生出不同的複雜形式。而在實務上，鋼筋查驗的程序多半是透過查驗人員對鋼筋籠進行「局部抽測」，並且輔以拍照記錄，將抽測結果之時間、地點、抽測位置、

檢查人員及參數（如前述之間距、形狀、尺寸等等）之測量值，如圖 1 所示。

即便基於抽測的方式已經降低了鋼筋查驗的困難度與繁雜性，然而在實務的操作上，參與鋼筋籠組立及檢驗的工作人員並無法完全避免危險的工作場所。以深基礎的開挖、高架橋墩柱或高層建築為例，檢查人員與施工人員在作業期間皆會面臨相同的墜落風險。而國內外的職災統計數據一致都顯示，在營建產業當中，佔據比例最高的重大傷亡事件都是「墜落與滾落」^[1,2]，防範工安意外的重要方法除了減少不安全的操作行為以外，也需要避免不安全的工作場域，這也是營建自動化的重要目標之一。因此，開發自動化的鋼筋查驗方法，不但能夠降低查驗人員在危險環境作業的頻率，也有機會對鋼筋籠進行更完整及全面的檢視，因為目前鋼筋查驗使用局部抽測乃是一種權宜

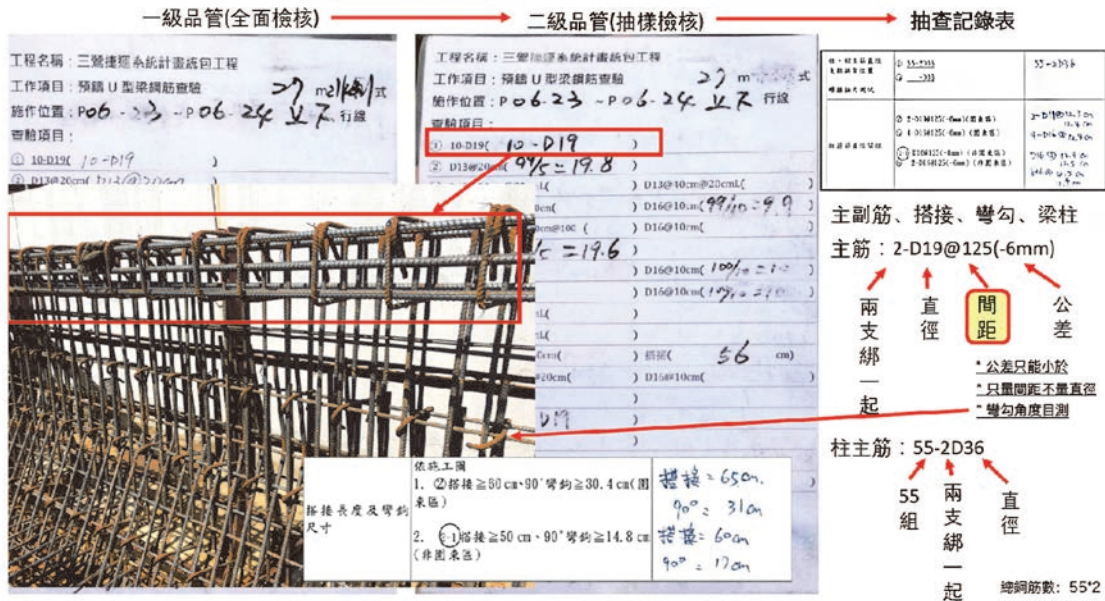


圖 1 鋼筋查驗項目及查驗表單

的作法，因為要對為數可觀且密集的鋼筋進行逐根查驗是難以透過有限人力來達成的。但自動化的檢測方式在成熟之後卻有機會克服這個問題。

隨著電腦視覺 (Computer Vision) 及人工智慧 (Artificial Intelligence) 應用的成熟與發達，鋼筋查驗有機會往自動化的方向進化。鋼筋查驗有許多不同的面相，包括間距、型式、號數等等，本研究以鋼筋的間距為例，展示一種基於深度學習與影像語意分割 (Image Semantic Segmentation) 的鋼筋間距辨識方法。從深度影像的蒐集與人工標記、機器學習的歷程，最後展現出可以符合實務標準的誤差範圍。其成果雖然只展現了鋼筋檢驗的其中一個環節 (即間距)，卻提供了其他的查驗項目一個可行的參考方向。

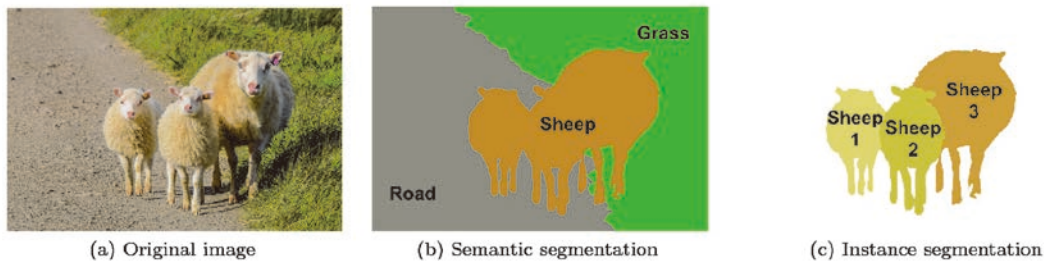
鋼筋查驗技術發展現況

以智慧科技輔助鋼筋查驗並不是一個全新的議題，但是既有研究的適用情境與需求與本研究略有不同。過往的許多科技輔助鋼筋查驗應用方法反而是作為非破壞性檢測來使用，亦即在混凝土已澆置完成的狀態下探測 RC 構造物當中的鋼筋的健康狀態，如鋼筋之直徑或是銹蝕情形。其多半需要透地雷達 (Ground Penetrating Radar, GPR) [3-5] 或是電感式接近感測器 (Inductive sensor) [6] 等硬體設備的輔助。雖然其應用情境與本研究提及的「鋼筋籠綁紮階段的鋼筋查驗」有所區別，但其中某些基於影像處理的技巧亦值得本研

究借鏡與參考。此外亦有相關的鋼筋查驗技術應用在混凝土澆置之前，例如利用光達 (Lidar) 建立鋼筋籠的點雲資料 [7]，再據以計算鋼筋之間距。比起影像處理，利用點雲資料的最大優勢在於較易於判定鋼筋的深度，因為鋼筋的組立是立體的，並且往往有分層的配置。若要利用自動化的方式來判讀鋼筋間的間距，其中一大障礙便在於容易將不同層次的鋼筋誤判為同一層次。然而，光達建立的點雲資料雖然能夠有效克服此障礙，在實用性上卻也有著設備昂貴不易普及，或是掃描需要較長時間的問題。本研究同樣取得了影像的深度資訊，卻是使用較新穎並且較普及的景深攝影機來取代光達。

影像語意分割

影像語意分割 (Image Semantic Segmentation) 與影像實體分割 (Image Instance Segmentation) 是近年深度學習領域的熱門議題，影像語意分割乃是企圖在圖像當中辨識出特定類型的物件，並且進一步界定物件邊界的技術 (以圖 2 為例 [8]，比方說在一張風景圖像當中區別哪些區塊屬於路面、哪些區塊屬於草地)。而影像實體分割則是更進一步地區別同一個類別當中的不同實體 (以圖 2 為例 [8]，影像語意分割僅能確定影像中哪些部份是「羊」，卻無法將羊的區塊進一步區隔為不同隻羊，而影像實體分割即是能夠將不同隻羊視為不同物件而界定邊界的方法)。本研究由於辨識

圖 2 影像語意分割及影像實體分割示意圖^[8]

目標為彼此區隔而互不重疊的鋼筋結點，使用影像語意分割已能滿足需求。而相較於一般比較為人所熟知的影像分類「將整張圖片對應到預先定義好的類別」，影像語意分割則是「將圖片中的每一個像素預測為預先定義好的類別」^[9]，換言之它實現的是像素等級的預測。在深度學習蔚為風潮之後，基於深度學習的影像語意分割方法也大放異彩，其中最為人所熟知的有 FCN (Fully Convolutional Networks)^[10]，SegNet^[11]，Dilated Convolutions^[12]及 DeepLab^[13-15]等等。要進行影像語意分割的監督式機器學習，使用者需要先製造數量足夠並且帶有遮罩的訓練樣本，亦即透過人工標記的方式在影像上精確以色塊標記物件的範圍，如圖 2 所示^[8]。

研究方法

本研究為了將深度學習方法套用在鋼筋查驗的自動化之上，需要自資料的蒐集開始，逐步建立監督式學習所需要的訓練樣本，其操作流程圖如圖 3 所示。

為了克服鋼筋影像當中不同層次的鋼筋可能造成互相干擾並且影響間距計算的問題，本研究除了取得鋼筋籠原始的 RGB 影像以外，還利用基於 Intel

RealSense 技術的 D415 及 D435 景深攝影機同時取得鋼筋籠的深度影像及點雲資料，如圖 4 所示。深度影像當中的每個像素顏色記錄著該像素的深度，以紅黃藍綠的順序代表深度的距離大小。景深攝影機本身並能配合開發套件輸出點雲檔案。

而本研究自正在施工階段的捷運車站及預鑄場拍攝約 37,000 張（共計 1,040 GB）的原始 RGB 影像，每張影像皆有對應之深度影像及點雲資料，其資料集之概況如表 1 所示。

在完成了初步的影像資料蒐集之後，下一個步驟便是據以製作訓練樣本，亦即為影像加上人工的標記。針對要計算鋼筋間距用的訓練樣本，本研究將其鋼筋縱橫交錯的結點以 Labelme 軟體標註於 RGB 影像之上，如圖 5 左半部所示。由於訓練資料的標註需要大量的人力投入，本研究僅選取 169 張鋼筋影像加以標記，其中 149 張作為訓練資料，10 張作為驗證資料，10 張作為測試資料。而圖 5 右半部則展示了作為測試資料的圖像，右上半為經過標記的結點位置，右下半為語意分割模型預測出來的結點位置。

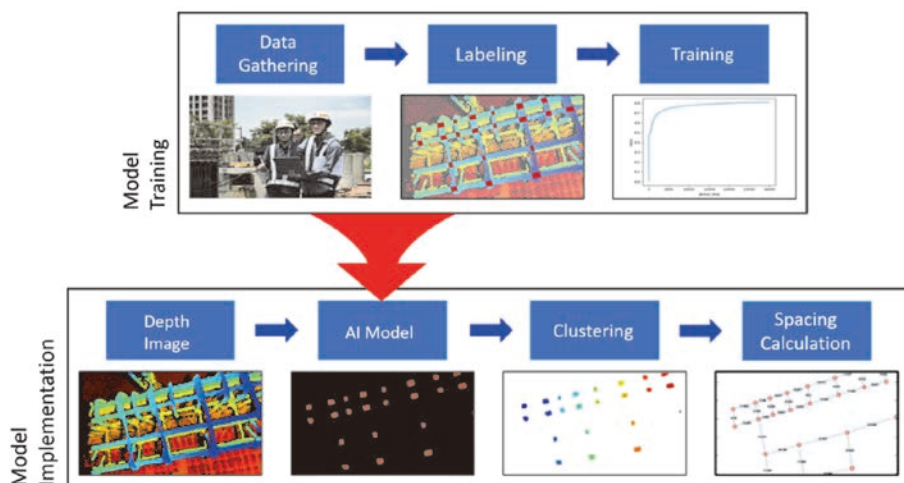


圖 3 人工智慧輔助鋼筋查驗流程圖

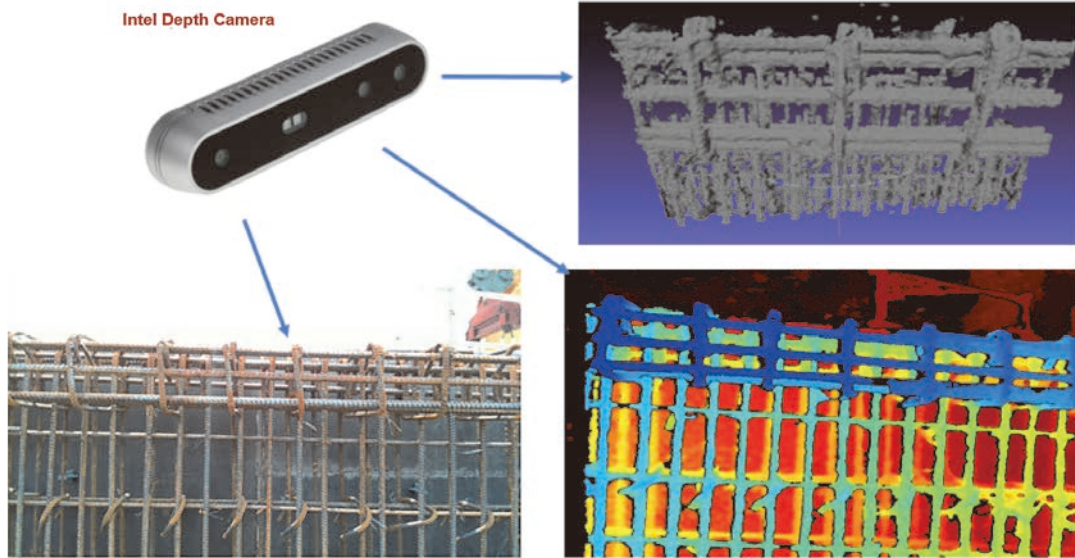


圖 4 景深攝影機運作示意圖

表 1 鋼筋影像訓練資料集概況

時間與地點	相機型號	蒐集影像總數量		
		RGB	深度影像	點雲
20190531 預鑄場 A	D415	13,528	13,580	13,580
20190710 預鑄場 A	D415	5,682	5,593	5,593
	D435	9,202	9,233	9,233
20191024 預鑄場 A	D415	3,579	3,593	3,593
20191220 捷運車站 B	D415	5,445	5,478	5,478
合計		37,436	37,477	37,477
共計 1,040 GB				

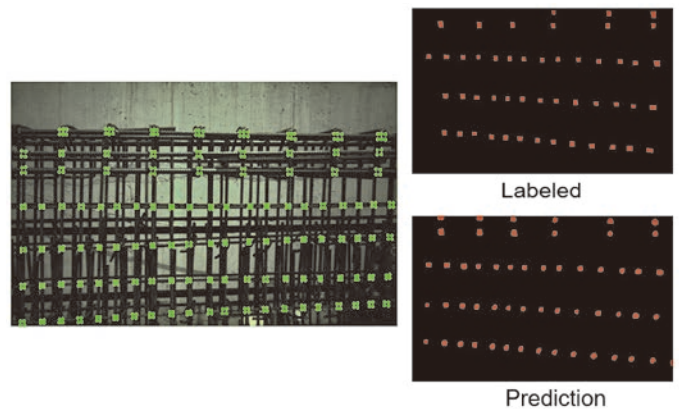


圖 5 標記鋼筋縱橫交錯之結點

本研究採用 DeepLab V3+ [15] 作為影像語意分割工具，並且以鋼筋縱橫交錯的節點為辨識目標。採用深度影像的一大優勢在於：在沒有考慮深度的情形下，不同層次的鋼筋的結點仍然都會被辨識出來，但是當不同層次的鋼筋結點都被投射在同一個影像平面上的時候，就無法保證「相鄰的兩個結點之間的距離」能夠代表鋼筋的間距，因為這兩個相鄰節點實際上可能屬於不同層次的鋼筋。然而，在深度影像有同時被導入的前提下，每個像素的深度資訊也被記錄在其中，而同時會被視為機器學習當中的特徵，因此當辨識的目標是「表層的鋼筋節點」的時候，屬於其他層次的結點不會造成干擾，而會被區隔開來。

結果與討論

透過 DeepLab V3+ 預測出結點的位置之後，下一個步驟是根據這些節點計算出鋼筋的間距。在大部份的場合，主筋與箍筋是整齊且互相垂直地交錯排列，因此本研究基於鋼筋整齊排列的假設，認定「在水平 / 垂直方向上相鄰的節點之間的距離可以代表間距」，繼而將間距計算出來並標記於圖上，如圖 6 所示。

本研究使用了 10 張鋼筋影像作為測試資料，每張影像如圖 6 所示，各自含有數十個水平間距及數十個垂直間距，每個實際的水平或垂直間距都已經經過測量，再進一步和「自預測結點所計算出來的間距」（亦即標記在圖 6 上頭的間距）相比較，繼而求算誤差，最後水平向與垂直向的平均誤差如表 2 所示。

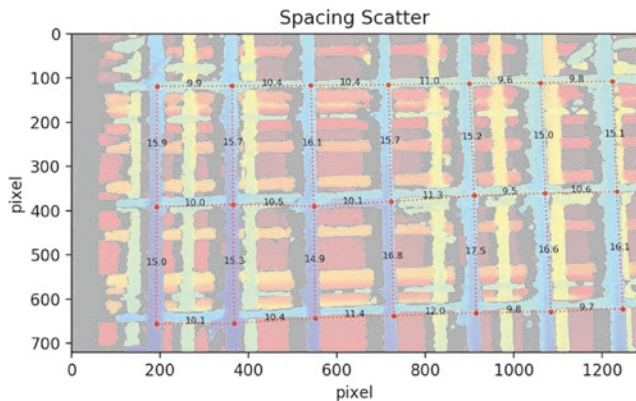


圖 6 將鋼筋籠結點連接取得鋼筋間距

表 2 水平及垂直向鋼筋間距之平均誤差

	平均計算間距 (cm)	平均實際間距 (cm)	平均誤差 (cm)	平均誤差 (%)
水平向	10.36	10.06	0.31	3.40%
垂直向	15.78	15.21	0.56	4.06%

結論

本研究提出基於深度學習與影像語意分割的方式，以發展自動化的鋼筋查驗方法。就預測結果而言，鋼筋間距的平均誤差值可以被鋼筋查驗的標準所接受，因此已具有投入實際場域應用的潛力。未來可能將其作為無人機上搭載的裝置，則可以減少檢驗人員接觸危險工作場域的機會，並有可能擴大鋼筋局部抽測的範圍，以達到更可靠的鋼筋查驗水準。

而本研究以鋼筋的間距為例，來展示影像語意分割的訓練方法，鋼筋查驗還有許多其他的查驗項目，諸如彎鉤、搭接、鋼筋號數（直徑）等等，在此應用框架之下，使用者僅需改變訓練樣本的標記方式，以及將來辨識目標的測量方式，便可將同樣的基礎應用到不同的鋼筋查驗項目之上。

誌謝

本研究感謝科技部與中興工程顧問股份有限公司在「以人工智慧輔助工地鋼筋查驗」產學合作研究案（MOST 109-2622-E-002-016-CC2）當中熱情贊助研究經費及慷慨提供實驗場域，並感謝國立臺灣大學土木工程系的韓仁毓教授、張書瑋副教授、張家銘副教授、陳柏華副教授對本研究的技術指導，與原國震中心與台大土木合設 AI 中心研究助理李啓民、國立台灣

大學土木工程學系電腦輔助工程組碩士生黃政維、國立台灣大學土木工程學系電腦輔助工程組暑期實習生邱仕鴻、黃湘娟、吳翊賢、李怡中對本研究案的協助。

參考文獻

- 行政院勞動部職業安全衛生署 (2020)。108 年勞動檢查年報。
<https://www.osha.gov.tw/1106/1164/1165/1168/29804/> (last access: 2021/01/25)
- OSHA (Occupational Safety and Health Administration) (2021). *Construction focus four training*. https://www.osha.gov/dte/outreach/construction/focus_four/index.html (last access: 2021/01/25)
- Hugenschmidt, J. and Mastrangelo, R. (2007). *The inspection of large retaining walls using GPR*. 4th International Workshop on Advanced Ground Penetrating Radar, pp. 267-271.
- Gehrig, M., Morris, D., and Bryant J. (2004). *Ground penetrating radar for concrete evaluation studies*. Technical Presentation Paper for Performance Foundation Association.
- Zhan, R. and Xie, H. (2009). GPR measurement of the diameter of steel bars in concrete specimens based on the stationary wavelet transform. *Insight*, 51, pp. 151-155.
- Zaid, M., Gaydecki, P., Quek, S., Miller, G., and Fernandes, B. (2004). Extracting dimensional information from steel reinforcing bars in concrete using neural networks trained on data from an inductive sensor. *NDT & E International*, 37, pp. 551-558.
- Han, K., Gwak, J., Golparvar-Fard, M., Saidi, K., Cheok, G., Franaszek, M., and Lipman, R. (2013). *Vision-based field inspection of concrete reinforcing bars*. 13th International Conference on Construction Applications of Virtual Reality, London, UK.
- Crous, M. (2018). *Combining weakly and strongly supervised segmentation methods for wind turbine damage annotation*. Bachelor Thesis, University of Amsterdam.
- Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., and Adam, H. (2018). *Encoder-decoder with atrous separable convolution for semantic image segmentation* [paper presentation]. ECCV (7).
- Long, J., Shelhamer, E. and Darrell, T. (2015). *Fully convolutional networks for semantic segmentation*. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3431-3440.
- Badrinarayanan, V., Kendall, A., and Cipolla, R. (2017). SegNet: A deep convolutional encoder-decoder architecture for image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39, pp. 2481-2495.
- Yu, F. and Koltun, V. (2015). *Multi-scale context aggregation by dilated convolutions*. Computing Research Repository (CoRR) in arXiv, 11/23.
- L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. Yuille (2014). *Semantic image segmentation with deep convolutional nets and fully connected CRFs*. Computing Research Repository (CoRR) in arXiv, 12/22.
- L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. Yuille (2016). *DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs*. IEEE Transactions on Pattern Analysis and Machine Intelligence.
- L. C. Chen, G. Papandreou, F. Schroff, and H. Adam (2017). *Rethinking atrous convolution for semantic image segmentation*. Computing Research Repository (CoRR) in arXiv, 06/17.