DOI: 10.6653/MoCICHE.202406_51(3).0010

淺談智能影像在勞安科技的應用

陳哲仲*／泛亞工程建設股份有限公司 BIM中心 經理

近年來，工程單位常透過「金安獎」、「金質獎」等競賽，展現工地導入創新科技之成果，此類創新科技如：人臉辨識、車牌辨識、安全帽/反光背心穿戴辨識、AI 電子圍籬 … 等，可謂琳瑯滿目，但競賽後還有多少單位實際使用？如果硬是要將不合規的產品置放到不合宜的場域，只能在競賽時跟評審委員做展示而已。目前各式競賽，委員也不是僅看是否應用創新科技，而是著重在參賽廠商是否用心、盡心與決心？同時也關注細節當中是否積極投入「科技減災」的應用。

本文將探討智能影像技術在營建工地的實際運用情形，並分享筆者六年來與多家智能影像廠商進行概念驗證（Proof of concept, POC）的心得，提供同業參考，以期在「營建數位轉型」的浪潮下，善用科技減災，降低職災風險。

前言

門禁與員工出勤考核導入人臉辨識系統已廣泛應用於各行各業。然而，若未採用高水準之設備與演算法，仍可能存在漏洞。同理，「AI 攝影機電子圍籬」系統若未考量第三維度（景深），亦可能影響人員入侵告警的判斷準確性。

上面兩個例子不是否定各家智能影像廠商的產品，但「隔行如隔山」，營建業與資訊業對彼此領域的認知差異甚大。若貿然將不適用之產品置於不合宜的場域，競賽時或可爭取佳績，實際運用則事倍功半。因此，營建業者需積極投入「科技減災」，並深入瞭解各項細節，方能降低職災發生之風險。本文將分享筆者對智能影像技術的鑽研心得，探討其在營建工地的實際應用。

基本技術介紹

在深入探討理論前，本文先引導讀者以最簡便的方式，利用 Teachable Machine 網頁工具建置物件模型，並使用筆電攝影鏡頭進行影像辨識，體驗智能影像的基本應用。透過此過程，可初步瞭解模型訓練與應用驗證的流程，為後續學習奠定基礎。

Teachable Machine [1] 是 Google 推出的網頁訓練模型的工具之一，簡單易用，不需專業知識和撰寫程式的經驗即可替網站應用程式訓練機器學習模型，支援辨識圖片、辨識姿勢和分類聲音。

首先在瀏覽器搜尋「teachable machine」，按 [Get Started] 開始建立新專案（圖 1）。

接下來會看到下面這個頁面（圖 2）：

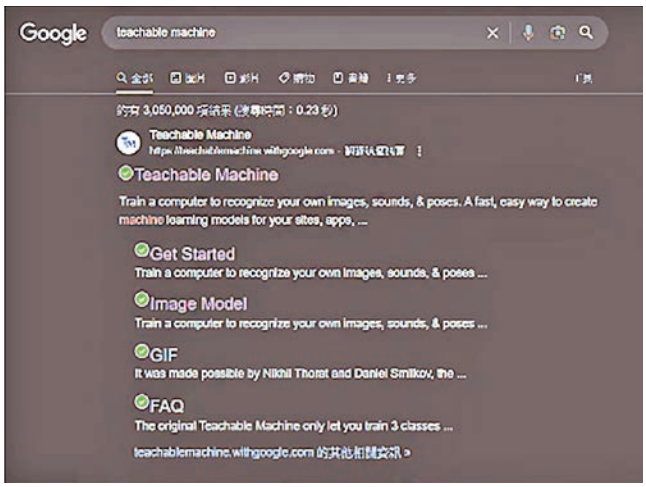

圖 1　利用網頁瀏覽器搜尋「teachable machine」畫面

* 通訊作者，jejung.chen@e-panasia.com

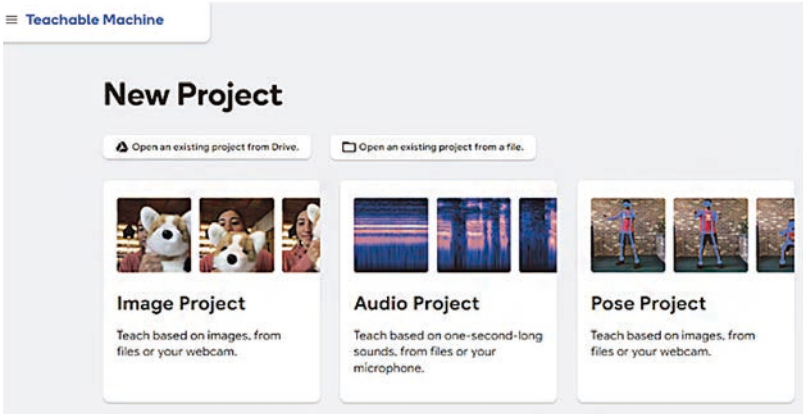

圖 2　Teachable Machine 建立專案之畫面

選擇第一個「Image Project」辨識圖片（第二、三項分別是聲音與姿態辨識），在「New Image Project」下有兩個選項，選擇「Standard image model」進入下一個畫面，更改第一個類別名稱（Class 1）為接下來所要學習辨識的東西（臉、布偶、杯子 … 等，當然 Class 1 也要改成你想識別的物件名稱，例如此例直接改名為「多啦 A 夢」），接下來圖框中會有兩個選項（圖 3），一個是用筆電攝影機建立模型庫，第二個選項是你提供大量的照片供電腦學習。此處選擇左邊第一個利用筆電攝影機來建立模型庫。

按住「Hold to Record」後會產出一次大約 4 連拍的照片（圖 4），轉動玩偶不同角度（可增加辨識率與可信任度）拍攝，至少產出個 4 ~ 50 張左右（原則上愈多愈好）。接下來下面「Add a Class」可以增加想訓練的模型樣本。

等圖片大致都已建置完成，就可以按第二個節點（Training）的「Train Model」進行模型訓練（圖 5），過了些許時間模型訓練完畢，接下來第三節點（Preview）就可以實際來驗證你所訓練的成果？

不管在第一、二甚至第三節點進行運作前，網頁

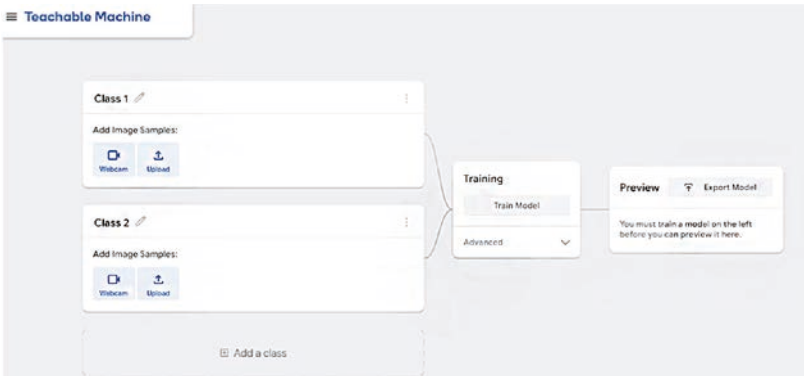

圖 3　Teachable Machine 節點式模型訓練預覽畫面

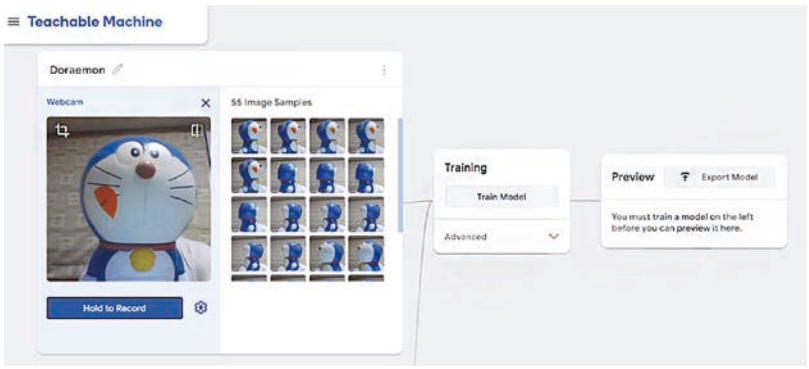

圖 4　利用玩偶進行影像模型取樣，取樣數愈多，準確率相對提高

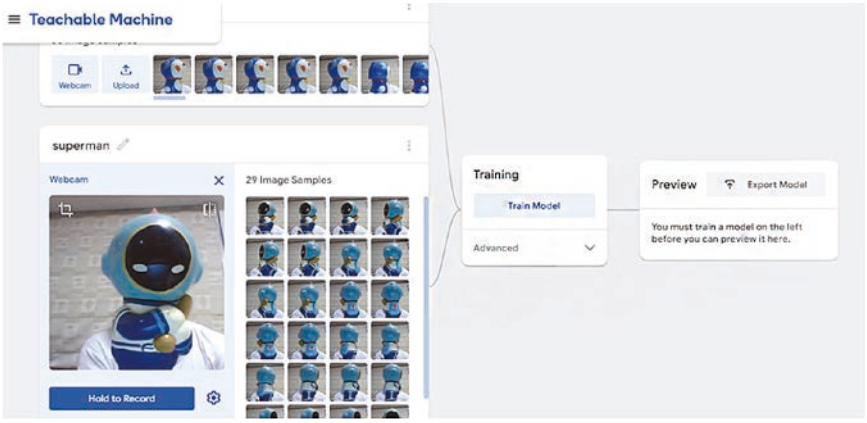

圖 5　增加第二種類型物件樣本供電腦學習，可一直增加類別

都會有一段示範影片提供您練習前參考，所以操作過程不致出現什麼問題。但到第三節點最後有一個「Export Model」的選項，點擊後會出現三個選項供選擇。因為牽扯到程式鏈接的部分，稍後會再介紹另一種簡易拖拉式節點的程式語言（Node-RED）。但如果能再有一些些 JavaScript 程式基礎，會對日後自行開發屬於自已甚至公司工地專屬場域環境的資料庫會更具優勢。

回到剛剛所談的第三個節點「Preview」（圖 5）。事實上這個節點就可以開始幫我們驗證剛剛所建立的模型庫是否能精準辨識物件？例如（圖 6）拿「多啦 A 夢」玩偶給電腦辨識，大家可以看到一條橘色的指示線條（Bar）一直在動，並呈現百分比（這邊我稱做「信心度」），本例玩偶愈接近攝影鏡頭，百分比愈高（這關係到另一個物件影像畫素取樣問題，愈遠畫素愈低。一般智能影像系統開發商通常會有最小成像值至少 100 × 100 或 50 × 50 Pixel 以上之成像畫素才做影像辨識等要求，營造廠在採購這方面技術與設備時，必須特別注意。）

圖 6 與圖 7 兩例的信心值都到 100%，信心值取決於物件畫素大小、模型採樣數量多寡 … 等因素。以工

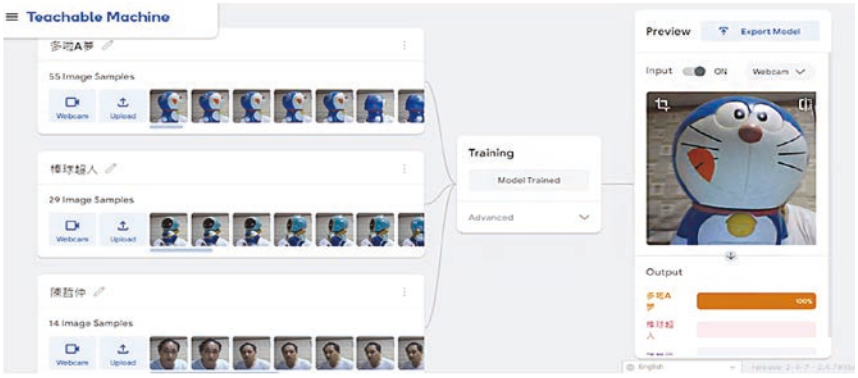

圖 6　拿「多啦 A 夢」玩偶對準鏡頭辨識，其橘色線條變長，信心度達 100%

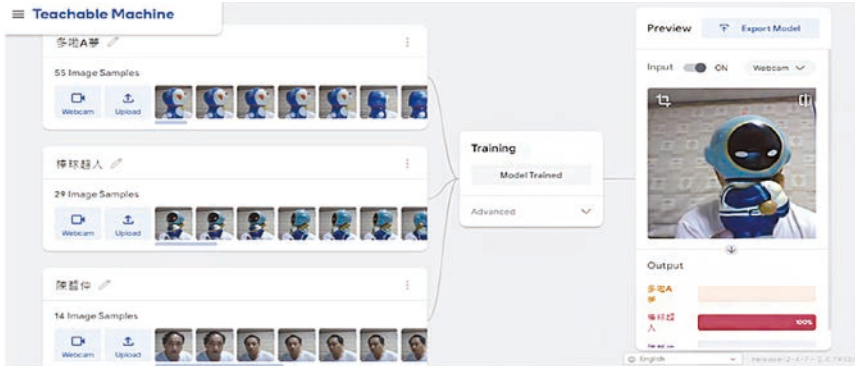

圖 7　拿「棒球超人」玩偶對準鏡頭辨識，其粉紅色線條變長，信心度達 100%

地安全帽為例，如果訓練的安全帽以白色居多，日後移至現場做安全帽穿戴辨識時，如果有人戴別種顏色安全帽，辨識信心值可能會大大降低，甚至可能無法辨識。

經由本例的基本介紹後，讀者應對開發智能影像辨識有深一層體驗，也了解其基本原理。當然影像辨識還是有其複雜的理論基礎，本文後半部分會再介紹一些理論並利用低程式碼（Low-Code）編寫工具來做模型庫的實際應用與嘗試。

TensorFlow 與 Node-RED

坊間相關智能影像產品多數以建置在攝影機本體（一般通稱為「前端」）或影像集中處理端（即網路影像錄影機、Network Video Recorder 簡稱：NVR）之後加設伺服器（這種方式通稱為「後端」）來進行 AI 辨識處理，一般前端智能影像辨識多由攝影機開發商直接寫入攝影機晶片裡居多，也有在攝影機近端加邊緣運算裝置輔助影像偵測。但目前市面上一般多採後端辨識居多。前端辨識做法就暫不討論，然就後端辨識部分；目前廠家多半是以 Python 或 C# 程式語言撰寫居多。本文的目的是希望在儘可能在不去撰寫艱澀的程式語言下，輕鬆將智能影像技術導入公司施工場域應用，且不需花費動輒數佰萬元導入一套不適合工地使用的工具。

在講到如何真正落實影像辨識應用之前，先介紹兩個名詞（工具）：TensorFlow [2]：是 Google Brain Team 開發，在 2017 年所推出的第一個正式版本，之所以稱為 TensorFlow，其因為輸入 / 輸出的運算資料是向量、矩陣等多維度的數值資料，稱之為張量（Tensor）。當機械學習模型需要使用流程圖來描述訓練過程中所有數值運算操作，稱之計算圖（Computational Graphs），這是一種低階運算描述的圖形，Tensor 張量就是經過這些流程（Flow）的數值運算來產生輸出結果，稱為：Tensor + Flow = TensorFlow。

Node-RED [3]：是 IBM 以 Node.js 為基礎，開發出來的 IoT 視覺化開發工具，因為純粹透過流程圖的方式，所以不需要會 Node.js 也可以透過 Node-RED 完成了許多後端才能做的事情。

Node-RED 是一種編輯工具，它提供了一個基於瀏覽器的編輯器，在工具面板中提供了許多節點，我們可以使用這些節點，將各個節點連接起來形成流程圖，之後只要按下「部署」按鈕，即可完成部署與程式執行。

接下來本文會使用這兩種工具，Google 的 TensorFlow 與低程式碼（Low-Code）的程式編寫工具：Node-RED，來完成程式的撰寫。

首先安裝 Node.js 到電腦（下載網址：https://nodejs.org/en/），再安裝 Node-RED（切換至 CMD 輸入：npm install -g node-red），開啟瀏覽器，網址列輸入「http://127.0.0.1:1880/」，就可以打開 Node-RED。（因文章篇幅限制無法一一詳述，如對上面安裝過程遇任何疑問，可逕自相關網站查詢。）

接下來將直接把節點程式展示（圖 8、圖 9），有興趣讀者可逕搜尋相關網站學習佈設節點等相關要領。

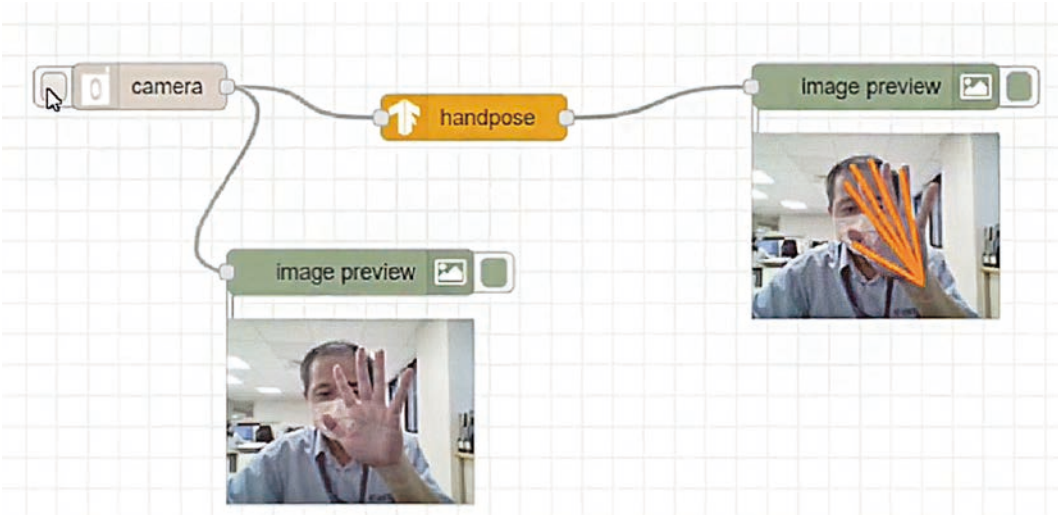

圖 8　利用幾個節點間的線連結，實現姿態（Pose）辨識

圖 9　影像姿態辨識（俗稱火柴人）

同理；Node-RED 本身就是 Intel 為 IoT 物聯網打造的連接應用工具程式，它除了可以做影像辨識外，還能把你想要連動的感測器、裝置一併導入程式內。有了這套工具對於初學者在學習物聯網與影像辨識會更加得心應手且事半功倍。

當然不是只有 TensorFlow 能做到這樣。目前影像辨識應用還有 Open CV 及 YOLO 等技術，其中 YOLO（YOLO 全名是「You Only Look Once」，意思是電腦只需要看一次，就能完成物件的影像辨識），是一種專門用於目標檢測的算法，因此它可以比 TensorFlow 或 OpenCV 能更有效地執行目標檢測任務，目前實施目標檢測較多還是以 YOLO 技術居多。而 YOLO 技術版本也從早期第一代到第八代，當然這些衍生出來的版本並未在科學期刊上有效揭露與評估，所以未被學術界所認可，因而一般人普遍認知技術多半都還停留在 V4 或 V7 版本。

上述不管是用何種技術，它的道理大多相同。白話一點說就是：「你有看過的，你就知道；你沒看過的，你永遠不知道。」也就是說訓練模型就如同小孩成長過程類似，你拿一隻狗的絨布娃娃跟小孩說：「這是『狗』」，下次再拿不同款式的狗娃娃跟他說：「這是『狗』」，小 Baby 的腦海中就已記下來，下次再拿同樣的狗娃娃給 Baby 看，相信他將不加思索地跟你講：「狗」。這個比喻就類比影像學習資料庫中數百種樣貌的模型，經由學習的方式加深影像辨識的信心度，以提高影像偵測的可靠度。有些科技公司引用國外既有模型庫來做工地安全帽穿戴辨識，因為學習樣本包含不同人種（如：白種人、黑人 …）不同裝扮與不同環境背景。或許用在臺灣還有 6、7 成的信心度，但若用在臺灣營建工地場域，如果工人戴上頭巾、頭罩甚至女性勞工安全帽內戴大草帽，在這些狀況下，影像辨識結果還能剩幾成？國外的模型庫可能也沒有學習到這些情況。

圖 10 就是利用已訓練完成的節點包（COCO SSD）進行物件偵測的範例，圖 11 使用 Node-RED 進行人形與物件（本例為手機）偵測的結果。重點是過程中不需要撰寫任何一行程式碼便可輕易達成 AI 影像識別，而該系統可植入的設備裝置僅需要一台相對高效的工業級樹莓派 4B 以上的機型，再加上一台 200 萬畫素的攝影機就可達成（按：目前 NVIDIA® JETSON Nano 亦可植入該相關應用，惟讀者在開發該相關應用時注意；切勿同一 Node-RED 系統內植入兩種不同技術；例如與 YOLO 節點包同時並存，容易造成系統當機或程式無法執行。）

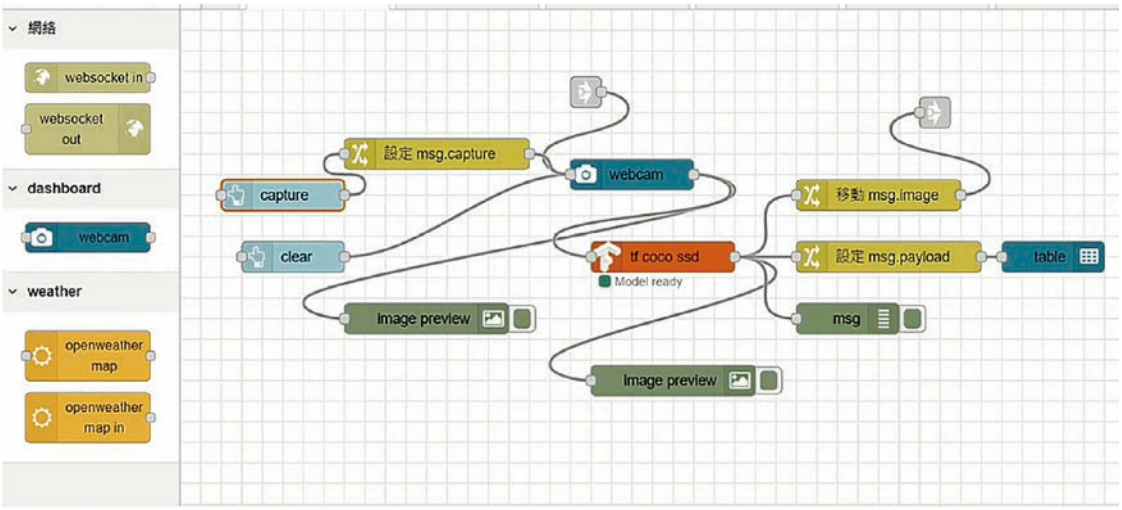

圖 10　利用 COCO SSD 節點包裡既有的模型庫，進行物件偵測的程式流（Flow）

圖 11　引用原節點包進行影像物件偵測：人與手機

談談智能影像原理

智能影像技術的三大支柱為人工智慧（AI）、機器學習（ML）與深度學習（DL）。人工智慧是廣義的概念，涵蓋模仿人類智慧的各種技術。機器學習為 AI 的子集，讓電腦在無需明確程式的情況下，透過數據學習與改進。深度學習則是機器學習的分支，以多層神經網路實現更複雜的學習模型。整個領域的架構如圖 12 所示。

人工智慧 [4] 是一個廣泛的概念，它涵蓋了任何能夠模仿人類智慧的技術。而機器學習是人工智慧的一個子集。它是一種基於數據的 AI 技術，可以讓計算機在沒有明確程式的情況下學習與改進。深度學習則又是機器學習的技術分支，用多層（深度）的神經網路技術來實現機械學習。

近年來硬體計算能力增強，許多演算法的精進，人們開始將傳統的單層神經網路擴增為更複雜的多層神經網路，這即為所謂的深度學習，在諸如影像辨識和自然語言處理方面都能得到很好的結果，如今我們口中的狹義型人工智慧，其核心其實就是由機械學習或深度學習技術所構成。而機械學習或深度學習經由圖 13 簡單的比較，這也就是為何深度學習大受歡迎的主要原因（人力消耗比起來較小）。

智能影像技術在勞安的應用

近年來已有很多營建廠商導入如安全帽 / 反光背心穿戴辨識、人員跌倒 / 昏厥、電子圍籬、高空作業安全帶掛勾與安全母索吊扣甚至還有專業廠商發展推高機環景人員入侵告警等應用。然而，工地的多變環境對影像辨識的精準度造成極大挑戰。

但事故發生前多半還是有其徵兆可循。例如：防止人員墜落，假設有人員進入尚未完全封版的屋頂，一

圖 12　人工智慧、機械學習及深度學習三者間的關係圖

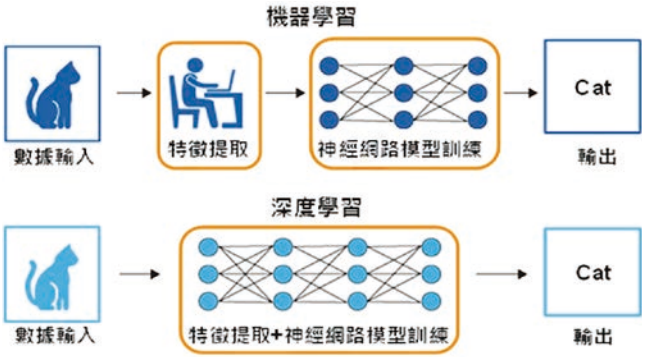

圖 13　機器學習、深度學習的運作概圖 [5]

旦施工人員接近開口，就近的攝影機能主動即時追蹤人員，並即時發出告警語音提醒。或者防止人員感電意外，若能在非專業電氣人員接近配電盤開啟盤門的同時，攝影機也能即時轉向該配電盤，借由感測元件主動發出告警語音或聲響，同時；CCTV 控制中心（或戰情中心）大屏幕即時顯示各種潛在可能發生危害事件之即時追蹤影像，授權控制中心人員循 SOP 回報並做即時因應處置，這才是發揮智能影像最積極有效的作法。

又如施工中的大型吊車翻覆案件時有所聞，吊車操作時，存在許多潛在危害因子，如吊車支撐腳伸出站立位置土質鬆軟，或墊底木塊、鋼板不確實放置、周邊阻隔三角錐防護區域未設置、吊車位於人員走動之通道 … 等，若攝影機能透過 AI 學習預判風險，適時發出告警語音，或由控制中心人員緊急通知工程師到場改正吊掛過程，這都有機會將智能影像的應用發揮至極致。

另外如道路管線埋設作業之應用 [6]，如果挖掘機挖掘深度過深，且未依規定設置擋土設施。此時人員進入管溝施作管路之風險極高。現今攝影機科技日新月異，倘若搭配景深攝影機將第三維度（也就是開挖深度）進行監控，一旦未施作擋土設施，即應發出告警訊息至控制中心，並將即時影像傳回，這才是避免災害一再發生的積極作為。

智能影像應用於工地職安衛的難處

營建工地的戶外環境、背光、陰暗等光線變化，以及工人配戴遮陽帽、頭巾、反光背心等影響辨識的因

圖 14　工人頭上披布巾再戴安全帽影像訓練成果

素，例如工人多半在頭上先戴個大草帽（草帽邊緣多半超出安全帽邊緣）或披條遮陽布後再戴安全帽（圖 14）；或者利用毛巾、頭巾包頭，然後只露出眼睛（圖 15）；而鋼筋工也經常打個赤膊穿個反光背心，或者是反光背心只是個交叉反光帶的形式，這些因素都會直接影響 AI 辨識的結果，這些影響辨識的因素，是智能影像廠商未必能夠考量周全之處。若能實際至工地採集樣本、標註分類，建構專屬的 AI 辨識系統，方能克服這些難題。

資訊業者受限於對營建環境的認知不足，恐難開發出真正符合工地需求的產品。隔行如隔山，真要說那方有錯，倒也不盡然，事實就是彼此認知不同罷了。傳統的系統整合商（System Integration，簡稱：SI），多半整合工業廠房自動化應用，對比營建業工作環境時時變化之情況要來得單純許多。建築物從無到有，再加上工作環境幾乎多在室外。對於電子產品來說，縱使設備具備防水、防塵等級。但相較於傳統廠房與工地環境，工地所需的防水、防塵等級似乎已超出安裝在室內環境下的工業廠房那種抗候等級，問題因此衍生開來。營造廠若能主動投入研發，深入探究影像辨識原理，定能創造出智慧工地管理系統，為傳統營建注入創新動能。

要真正把智能影像完全應用在營建產業，的確不是一件簡單的事。若能深入研究探知整個影像辨識的原理，甚至實際與廠商進行現場樣本採集、框圖、分類等過程，而建構屬於自己公司案場的 AI 辨識系統，才能真正輔助工地職安衛推行「科技減災」的最有效作法。

結論

智能影像技術為營建業的數位轉型帶來契機。藉由視覺化開發工具，營造廠可自行建置專屬的 AI 影像辨識模型，導入工安管理，避免職災發生。但科技產業中的使用

圖 15　泰藉勞工安全帽內包頭巾僅露出眼睛之訓練後辨識成果

模式，例如用在生產線不良品剔除，提高產線良率；或是利用智能影像監視人員操作設備之順序是否合規？更有利用熱顯像技術監測管線、器具是否有異常過熱現象？這些在廠房內使用的產品或技術，一旦移至戶外使用，效果可能大打折扣。因此長遠而言，營建業者應組成專案團隊，積極投入相關技術研發，惟有營造廠願意主動成立專案團隊抑注資金主動投入研發，才能創造屬於營建工地專屬的智慧工地管理系統。如此才能深化「科技減災」的應用層次，為打造安全、高效的智慧工地而努力。

參考文獻

1. 陳會安，「Node-RED 視覺化開發工具」，博碩文化，台北，P16-13~19（2021/9 初版）。
2. 陳會安，「RaspberryPi 樹莓派 Python x AI 超應用聖經」，旗標科技，台北，第 12-28 頁（2022/01 出版）。
3. 王進德，「工業 4.0 的物聯網智慧工廠應用與實作」，博碩文化，台北，第 106 頁（2021/04 初版三刷）。
4. 羅光志，「從 AI 到生成式 AI - 40 個零程式的實作體驗」，旗標科技，台北，第 2-29 頁（2023）。
5. 賴建宏，「人臉識別的原理與實作」，台北，研討會簡報資料檔第 5、7 頁（2020/08）。
6. 陳哲仲，「路面挖掘安全監視系統」，2023/02，「專利範圍」，專利號：M637835。