

應用影像深度學習於人形機器人 足球賽之即時物件偵測

Utilizing Visual Deep Learning for Large-range Object Detections of Humanoid Robot Soccer Games

林昆鋒、郭重顯

Ken-Fun Lin, Chung-Hsien Kuo

RoboCup 為全球人形機器人之重要比賽之一，其中雙足人形機器人足球賽除了需考慮到雙足動態步行之穩定性外，也需進行即時物件偵測、辨識與定位，以作為自主決策與運作之依據。所探討之人形機器人搭載 NVIDIA Jetson TX2 嵌入式人工智慧運算器實現機器人視覺即時偵測、辨識之運算，並以深度學習卷積神經網路開發，導入 You Only Look Once 網路進行足球與機器人類別之偵測與辨識。此一成果在 RoboCup 2017 賽事進行實際驗證，並獲得中型組人形機器人足球賽第二名成績。

RoboCup is one of the most important humanoid robotic competitions in the world. The development of humanoid soccer robot considers not only the bipedal locomotion stability, but also the real-time object detection, identification and positioning which help the autonomous decision-making and operation during competition. The proposed humanoid robot is equipped with a NVIDIA Jetson TX2 embedded artificial intelligence controller to realize the aforementioned real-time object detection based on the deep convolutional neural network (CNN). Practically, the classes of soccer ball and humanoid robots are detected and recognizes in terms of the You Only Look Once (YOLO) model. The results are validated in the RoboCup 2017, and our team won the 2nd place of the TeenSize league.

一、前言與文獻回顧

RoboCup 世界盃足球機器人競賽為全球人形機器人之重要比賽之一，其以機器人能夠在 2050 年達到與人類足球員一樣之能力為發展宗旨。因此，大會每年更動賽制內容促使機器人之工程與技術發展，並逐步納入各式不同的功能或效能要求，以逐年提高足球機器人在機構設計、系統架構、控制理論、影像處理之技術挑戰。RoboCup 各種不同機器人聯盟中，雙足人形機器人足球賽具備高技術挑戰賽，其除了需考慮到雙足動態步行之穩定性

外，也需要進行比賽場上之即時物件偵測、辨識與定位，以作為機器人自主決策與運作，使得在比賽中能有好的表現。

從 2017 年開始，RoboCup 足球賽不再使用特定顏色足球作為影像辨識標的；其使用標準足球，如此一來使得以顏色特徵為辨識的方法變得不再適用。因此，以深度學習所發展之物件偵測技術變成一個可行的方案。有鑑於此，本論文研究提出一基於深度學習的機器視覺系統架構，以中型雙足人形機器人足球影像追蹤與機器人識別為目的，達到自主足球追蹤與帶球之功能，將實際應用於 RoboCup 國際性機器人競賽之中⁽¹⁾。

由於本文重點在於闡述人形機器人於比賽過程中，機器視覺、物件偵測與辨識、影像定位之技術實現方法，因此首先針對機器人足球賽之相關影像技術進行探討。Awaludin 等學者針對 RoboCup 競賽進行研究⁽²⁾，並探討守門員之影像辨識及防守策略。在球的位置測量的過程中，第一次進行辨識三個不同的對象，包含足球，球門與場地線。利用 HSV 色彩空間針對上述三種物件進行辨識，而球場上球位置的量測，則是通過神經網路模型對其在圖像中的位置進行訓練而得到，而訓練結束後，其平均誤差小於 1 厘米，得到相當準確之足球位置。

此外，Sudin 等學者針對足球機器人競賽進行研究⁽³⁾，提出了一種藉由一個機器人和一個球門進行初始自我定位演算法來估測機器人之實際位置，透過幾何學，估測出機器人的真實世界坐標位置。Minakata 等學者亦針對 RoboCup 機器人足球世界盃進行位置估測研究⁽⁴⁾，使用單攝影機定位方法並捕捉 2 個地標包含球門與定位柱。而每一個地標的世界座標位置皆為固定，使機器人可以透過三角測量回推自己的絕對位置。Chiang 等學者對 RoboCup 機器人足球世界盃進行研究⁽⁵⁾，透過過濾綠色場地，使用線性迴歸色彩模型之方法，建立一更精準之顏色模型，從場地上之線段及多重物件中找到足球。以擴展型卡爾曼濾波器建立一動態足球影像子視窗，達到更準確有效之偵測與追蹤結果。

在另一方面，深度學習影像辨識與導航亦是本文之技術重點。Chen 等學者於 2014 年發展出一套對於機器人自主導航的深度學習方法⁽⁶⁾；選擇深度學習中的卷積神經網絡建模技術來提取特徵，其模型具有多層特徵，其中環境場景也可以被識別，可以識別門的位置等資訊。所取得的資訊可以用於機器人導航，使得機器人可以準確地接近目標。同時，Contreras 等研究學者於 2016 年提出利用深度學習改進室內移動式機器人探索方法的兩個部份⁽⁷⁾：一部分是採用車載攝影機所取得圖像來識別物理結構；另一方面則是判斷離障礙物的距離，以便確定機器人的方向和移動速度。

二、系統架構

1. 視覺影像相關研究

本文中提及與使用之人形足球機器人平台，承用了本實驗室歷年之研究成果，所提出機器視覺深度學習系統應用於既有之中型人形機器人平台之上。其系統整體架構可以區分為深度學習視覺影像系統、核心系統、機器人頭部馬達控制，以及動作控制等四個部分，其中深度學習視覺影像辨識系統涵蓋線下監督式學習使用訓練平台部分，而機器人頭部馬達控制與動作控制分離，直接接收來自核心系統控制指令，加速影像資訊對物件之追蹤反應速度，整體架構如圖 1 所示之系統架構圖。

本論文中將著重於深度學習視覺影像系統物件追蹤之開發，本深度學習視覺影像系統包括：

1. Nvidia 所推出 Pascal™ 架構系列產品 Jetson TX2 為人形足球機器人上影像系統運作單元。
2. 配羅技科技 (Logitech) C920 相機做為影像感測元件。
3. 英特爾 (Intel Co.) 所推出的 Mini PC (NUC Kit D54250WYKH)，做為核心系統運作單元。
4. 動作控制系統使用 32 位元微處理單元 (Teensy 3.2) 接收來自六軸感測單元 (Uranus-Lite) 輸出之 I2C 信號，並透過 UART 傳輸介面做為與核心系統運作單元與伺服馬達之溝通管道。

本系統區分各運作單元，分別進行並行與各自獨立處理流程，並由核心系統運作單元連接各分散式系統單元通訊管道，以平行處理的運作方式，降低個別單元內處理量。分散式模組的方式能夠針對不同運算需求，進行個別硬體平台上之選用，除增加系統整體之效能與穩定性外，能使系統有彈性調整可能，以及針對各部分進行問題除錯的速度。

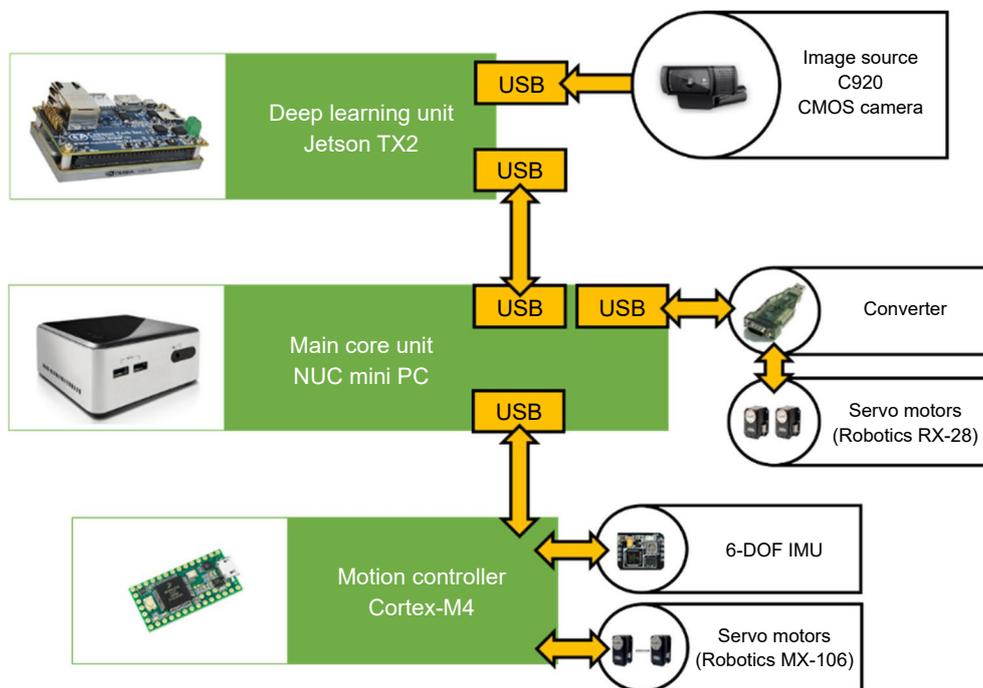


圖 1. 系統架構圖。

本文之系統開發於實驗室既有中型人型機器人 (HuroEvolutionTN) 平台之上。該機器人係以符合 RoboCup Humanoid League 競賽之規則而設計，機器人之實體與機構示意圖如圖 2 所示。機器人全身總共具有 16 個自由度，分別包含位於機器人下肢的 14 個自由度，以及頸部水平與垂直控制的兩個自由度。機器人使用的動作步態可以分為兩種：全向移動的動作步態，以及非移動時的特殊步態等兩種步態模式。在全向移動的控制中，利用軌跡規劃與運動學等方式，以達到機器人的動作控制；並具有前進、側移、旋轉 (X, Y, Z) 等三個分量的向量控制。在非移動時的特殊步態中，包含左踢、右踢、撲倒與站立等非移動進行之動作，利用固定的軌跡以單次播放的方式執行，並且在執行的過程中無法被中斷，以確保動作完整地執行。

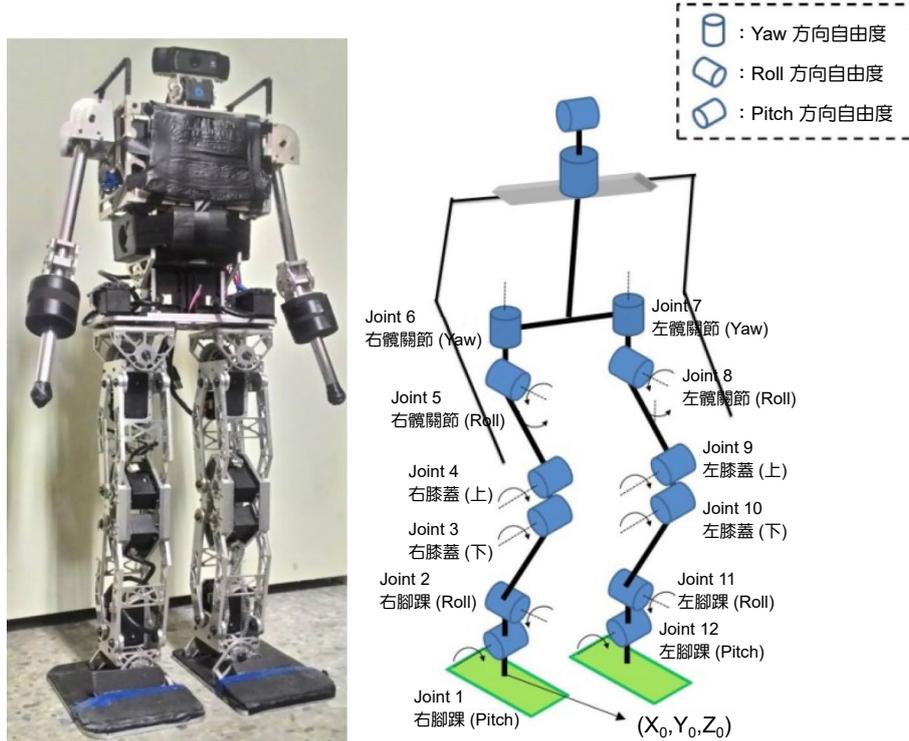


圖 2. 中型人型機器人 (HuroEvolutionTN)。

本論文之系統運作流程，以視覺影像處理使用深度學習之物件辨識開始，視覺影像由嵌入式深度學習平台擷取自位於機器人頭部之攝影機，攝影機初始視角與地面相交 45 度，機器人平台靜止站立時高度 90 公分，所視範圍約落於機器人前方 30 公分至 120 公分處，使用影像解析度大小為 640×480 像素，之後進入卷積類神經網路中，逐層解析特徵，並輸出影像中可信度高之物件類別與位置框選，經由機器人頭部二維的角度回推，可算得該所識別物件之相對機器人座標位置，以此定義所識別物件距離機器人之遠近。隨機器人行走時會產生左右晃動視野，產生所識別物件小幅度晃動現象，加入一階的數位低通濾波器來抑制數值變化量，得到穩定的物件距離資訊，最後依已建立之機器人追蹤行為資料庫轉化為機器人實際移動指令靠近足球。

2. 系統運作流程

本文之系統運作流程如圖 3 所示，其以視覺影像處理使用深度學習之物件辨識開始，視覺影像由嵌入式深度學習平台擷取自位於機器人頭部之攝影機，攝影機初始視角與地面相交 45 度，機器人平台靜止站立時高度 90 公分，所視範圍約落於機器人前方 30 公分至 120 公分處，使用影像解析度大小為 640×480 像素，之後進入卷積類神經網路中，逐層解析特徵，並輸出影像中可信度高之物件類別與位置框選，經由機器人頭部二維的角度回推，可算得該所識別物件之相對機器人座標位置，以此定義所識別物件距離機器人之遠近。隨機器人行走時會產生左右晃動視野，產生所識別物件小幅度晃動現象，加入一階的數位低通濾波器來抑制數值變化量，得到穩定的物件距離資訊，最後依已建立之機器人追蹤行為資料庫轉化為機器人實際移動指令靠近足球。在深度學習卷積類神經網路物件辨識之中，將其設定為

一可獨立運作之系統，影像於此系統中直接擷取自攝影機，經過深度學習卷積類神經網路獲得物件類別與框選資訊，並透過串列埠將識別出之物件資訊傳出，後進入到機器人控制核心系統，達到深度學習機器視覺系統獨立運作，並能已轉至別的系統上使用之優點。

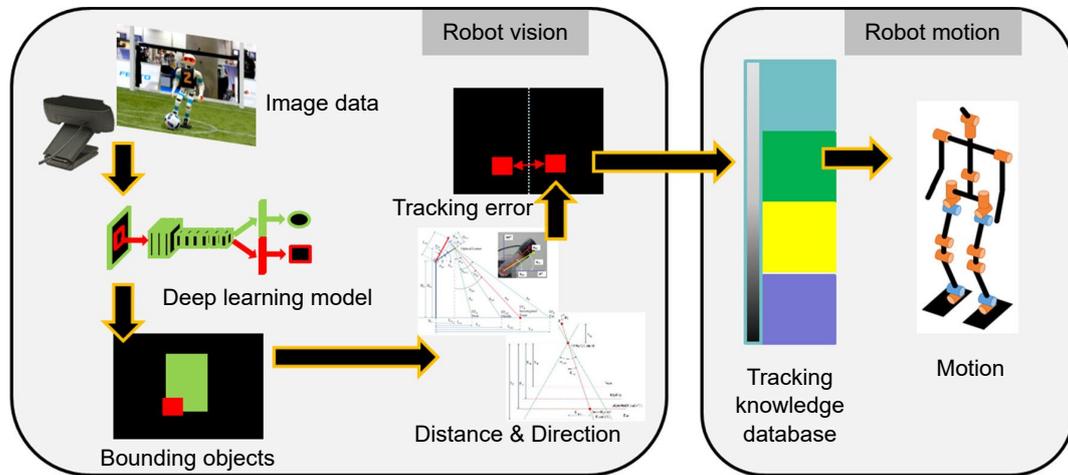


圖 3. 人形足球機器人運作流程圖。

三、人形足球機器人視覺深度學習網路

1. 深度學習卷積類神經網路架構

目前在即時運算的系統中，常見的目標物偵測包含下列幾種方式，如利用目標物的顏色、形狀等特徵進行判斷，以及背景相減法、樣本比對法等方法。然而在目前常見之目標物偵測演算法中，皆無法有效地解決機器人對於複雜的背景環境的影響；對於在偵測物件程序中，物件區塊上被環境陰影遮蔽或光源分布不平均時的影響，例如來自機器人之陰影或燈光下與偏燈光下之亮度強弱區域，對於一般常見影像演算中，這種局部差異，會影響到程序上的抉擇，產生不穩定現象；在賽用足球物件上，可以從賽事規則上期待其外態，但不能確定實際賽場上的足球物件在程序上的適應情形。

因此，為了解決上述所提及之問題，以及滿足系統對於即時運算之需求，本文以卷積類神經網路架構進行物件認知學習，其中經由卷積 (convolution) 與池化 (pooling) 過程，將圖片大量數據進行特徵的精簡化，形成一特徵資訊層，以精簡化的特徵資訊進入最後全鏈結之類神經網路學習層，通過類神經網路各權重對特徵進行數據調整，最後加總其輸出數據進行物件分類評分，此一卷積類神經網路架構如圖 4 所示。

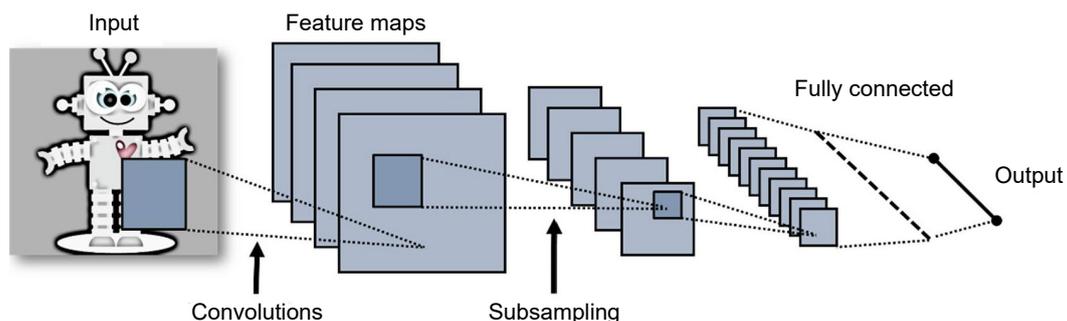


圖 4. 卷積類神經網路架構圖。

2. 網路學習架構與足球賽應用

本文之卷積類神經網路學習架構中應用了 Redmon 等人提出的 You Only Look Once (YOLO) 的網路學習架構，透過一次通過與簡化重複卷積與池化過程，而降低網路運算量，並以分批隨機調整圖片尺度的方式防止過擬合 (overfitting) 結果。YOLO 為可一次性預測多個物件框選位置與類別的卷積類神經網路，同時也是端到端 (end to end) 的一個深度學習架構，運作方式如其命名，以一次性通過卷積類神經網路，並無在卷積過程後之特徵圖中，使用候選區域 (region proposal) 的過程，而是以是先將圖片切割成網格區塊，進行不同區塊中的物件識別學習與位置框選，其最大的優勢在於運行速度快與使用較少的訓練資源。本文所用之深度學習架構，採用此架構來節省運行資源，同時提高即時影像速率，達到即時辨識之效果，其原始模型如圖 5 所示。

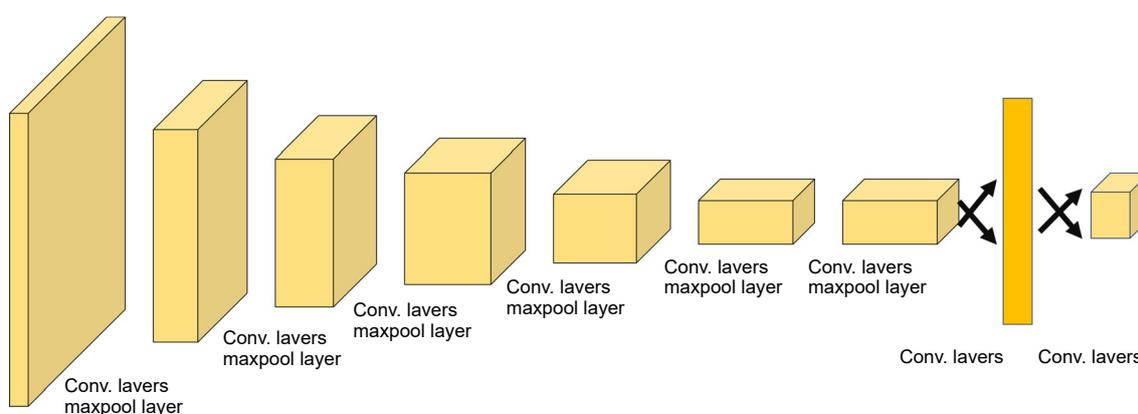


圖 5. You Only Look Once 初版卷積網路架構。

本文 YOLO 模型架構採以 9 層卷積層與 6 層池化層進行反覆卷積，並使用 Leaky ReLU 激活函數在卷積後激活單元與最大池化方式進行池化，最後過濾出學習之特徵數據，進入最後一層卷積層中進行學習，用網格分區資訊傳遞方式，使之與初版進入完全鏈結類神經網路學習略有不同，最後使用從 Sigmoid 函數發展而來的 Softmax 對事前所定義之物件，進行個別單元數據總和之計算與結果輸出。YOLO 採用固定特徵框 (anchor box) 的方式，進行類神經網路訓練與結果輸出的。

為了達到資料一次性通過就能獲得辨識解結果，在圖片的前處理上，先對圖片進行區域的劃分，以分區提取資料屬性 (attribute) 的方式進行區域化資料分析，如圖 6 所示；其中資訊又分為以區域中得到最大可能之物件可能與針對每一區域特徵資訊評算而出的框定範圍，最後以評選之閾值將符合程度高者保留，如圖 7 之流程所示。

為了測試，深度學習之成效，使其得以有較大彈性運用在機器人足球的活動上，以廣泛的足球測試樣本，進行不同樣貌之足球物件認知測試；以往年賽場上之場景與所使用之足球物件測試臨場認知測試；最後測試其物件框選之誤差，經由測試後所得結果可予以此深度學習網路一常態分佈學習成效。為了檢驗深度學習後卷積類神經網路所訓練之成效，以訓練樣本外的數據來進行測試。其中設計有：(1) 所收集之樣本對於該物件的認知情況，本論文稱之為認知廣泛程度測試；(2) 實際場景上之表現效果的應用場景測試；(3) 物件框選遺漏比例的誤差統計。

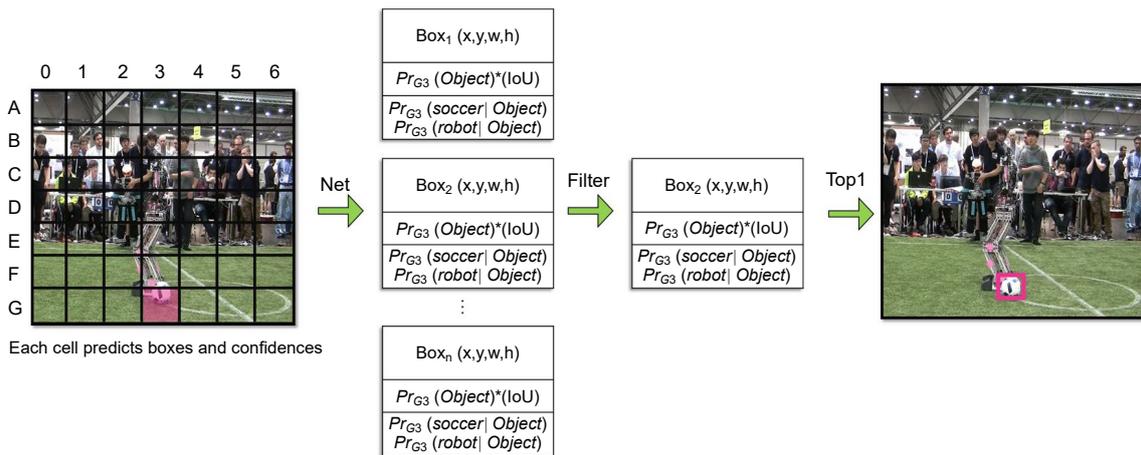


圖 6. 分區提取資料屬性。

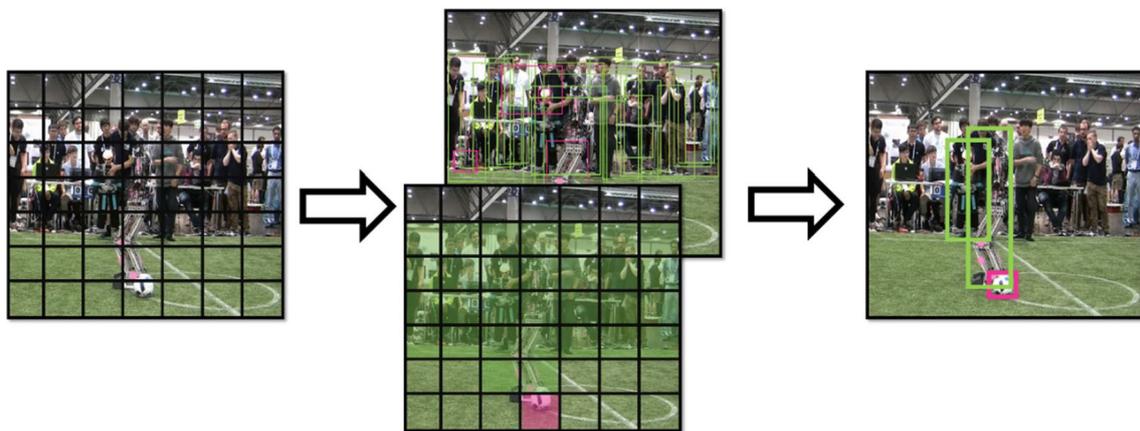


圖 7. 一次性通過物件偵測流程。

就同樣的物件而言，其色彩變化上可能有極多的變化，不同出產地、材質選擇或者外觀設計上之差異，可能都對同樣一個物品在視覺模型上，產生了不同屬性的資訊。為了測試訓練後模型對於不同樣貌的同名物件認知情形，本論文測試中，以收集物件樣貌及背景具差異的測試圖集進行測試，如圖 8 所示。其中測試以所收集圖集總數 N_{vt} ，模型框選狀態與實際框選框重疊程度 $Overlap$ (%)，採用每張圖中第一相似率框選 $Top 1$ ，計算其認知廣泛程度結果 $Result_{vt}$ (%)，如公式 (1)–(2)。

$$Result_{vt} = \frac{\sum Correct_{vt}}{N_{vt}} \times 100\% \quad (1)$$

$$Correct_{vt} = \begin{cases} 1, & Overlap(Top1) > 50\% \\ 0, & otherwise \end{cases} \quad (2)$$

由於本文之應用場景為 RoboCup 足球比賽場景，因比賽規則上有定義其足球大致樣貌，足球樣貌在所測試之測試集中相近，如圖 9 所示；故測試的重點著落在模型框選與實際框選重疊程度上，此處提升框選重疊程度為測試結果數值調整如公式 (3)–(4)。

$$Result_{at} = \frac{\sum Correct_{at}}{N_{at}} \times 100\% \quad (3)$$

$$Correct_{at} = \begin{cases} 1, & \text{Overlap}(Top1) > 90\% \\ 0, & \text{otherwise} \end{cases} \quad (4)$$



圖 8. 認知廣泛程度測試圖集。



圖 9. RoboCup 足球比賽場景測試圖集。

而之後段章節將提及之足球定位方法，其涉及足球物件之框選，使用訓練完之模型測試其框選之狀態，可以獲得有用之修正框選數據，此處框選結果之評分標準以物件實際框選區域 $Object_{truth}$ ，訓練後模型框選區域 $Object_{pred}$ ，計算未覆蓋比例 $Err_{overlap}$ 。

四、足球機器人足球追蹤系統

1. 足球定位系統

足球機器人足球追蹤系統需基於前章節中所獲得深度學習機器視覺物件辨識之結果，應用於實際比賽足球場上目標物件定位，控制機器人進行目標物件之方向誤差計算與追蹤行為資訊庫之建立。整體運作係從影像擷取到嵌入式的深度學習辨識系統中識別物件，並進行物件之定位計算，最後執行機器人靠近目標物件 (如足球) 之任務，整體流程如圖 10 所示。

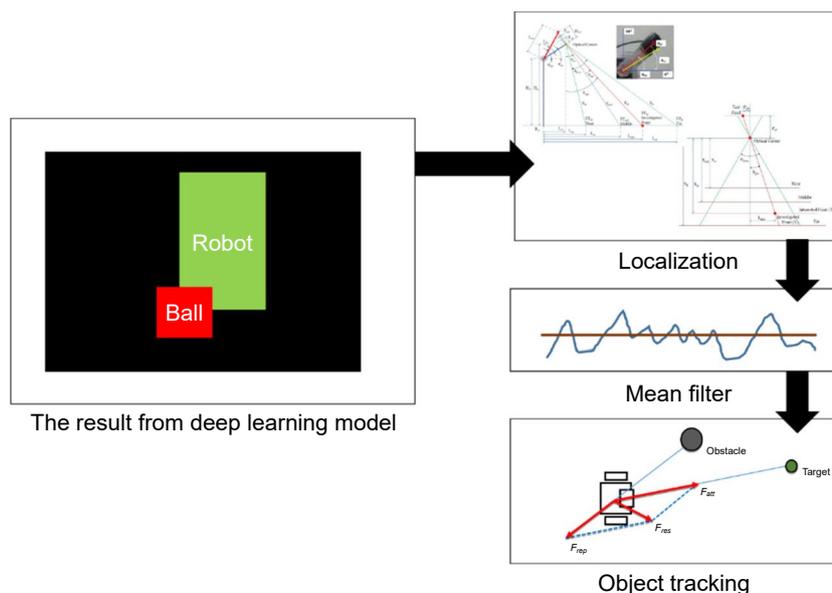


圖 10. 人形足球機器人追蹤流程。

本文所提出之視覺定位系統係利用攝影機與目標物體的底部位置之光學幾何關係建立而成。根據定位方向的不同，本論文定義機器人之前方方向為 y 方向；機器人之側邊方向為 x 方向。其定位可以由機器人的側視圖以及俯視圖進行說明，如圖 11 與圖 12 所示。在機器人視覺定位的側視圖中，主要量測的是目標物位於機器人前方的距離。其中的基本參數包含， P_{vf} 為攝影機之成像距離； θ_{mt} 為攝影機之傾角，由一垂直方向之馬達控制； H_{ro} 為機器人之高度； H_{obj} 為目標物之高度； L_{crc} 為馬達轉動軸到感光元件之間的距離。

當在影像中取得了目標物框選的中心位置之後，計算該位置與影像中心位置 (optical center) 的距離，分別求得在影像中 x 、 y 方向上的 P_{ipx} 與 P_{ipy} 之像素距離。並且利用該像素距離 (P_{ipx} 、 P_{ipy}) 與攝影機的成像距離 (P_{vf})，計算求得該物體 (FP_{ip}) 與視野中心 (FP_{md}) 之夾角角度 (θ_{ipx} 、 θ_{ipy})。將馬達傾角 (θ_{mt}) 轉為實際攝影機之傾角 (θ_{ca}) 後，加入上述所求得之角度 (θ_{ipy})，得出機器人與目標物實際夾角的角度 (θ_{ripy})。

若直接將公式帶入將產生距離誤差，因此利用已知足球之半徑高度，對攝影機高度進行補償，使攝影機高度 (H_{ca}) 為機器人高度 (H_{ro}) 減掉物體高度 (H_{obj}) 來補償因目標物之半徑高度所產生之投射誤差，最後利用三角函數求得目標物位於機器人前方之距離 (L_{ripy})，以及攝影機 (optical center) 至目標物於 y 方向上之投影位置 (FP_{ip}) 的直線距離 (S_{ip})。在求得攝影機與目標物於 y 方向上之投影點的距離 (S_{ip}) 後，利用如圖 12 之光學幾何關係，以該距離

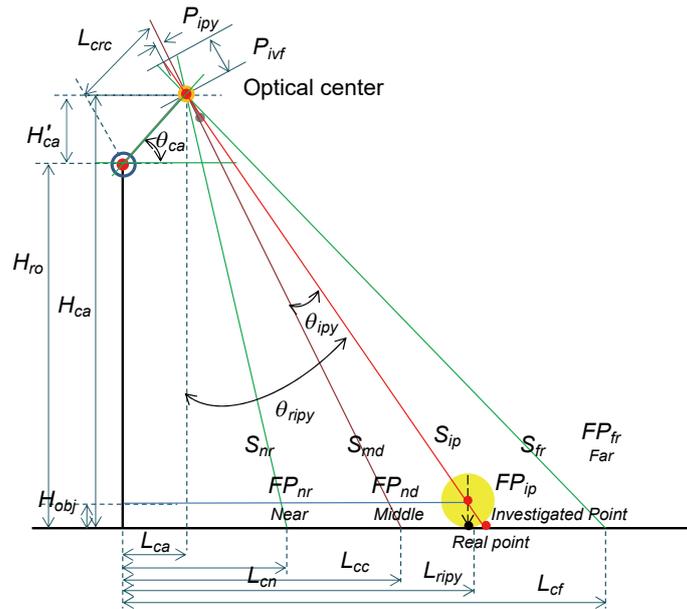


圖 11. 視覺定位之側視面示意圖⁽⁸⁾。

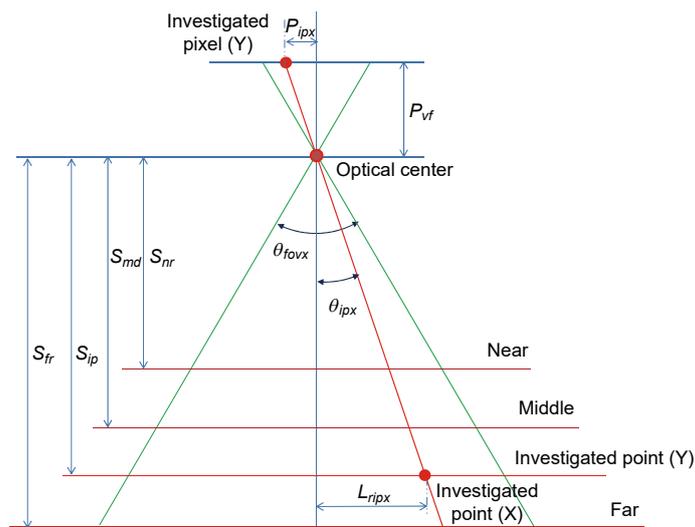


圖 12. 視覺定位之俯視面示意圖⁽⁸⁾。

(S_{ip}) 與目標物至影像中心之夾角角度 (θ_{ipx})，即可計算求得目標物位於機器人側邊 x 方向上之距離 (L_{ripx})。上述之定位公式可參閱我們發表於期刊之論文⁽⁸⁾。

由於機器人在行進間所產生的晃動，會使上節視覺定位的結果隨之產生漂移，並造成後續決策判斷上的困難。因此本系統在透過光學幾何求得物體定位的結果之後，加入均值濾波器對於定位的結果進行濾波，以取得較為穩定之視覺定位，其公式如 (5) 所示；其中，為時間區間 T 的平均輸出值；為時間區間 T 的每單位時刻歷經值； T 為固定之時間間隔區間。

$$\rho_{avg} = \sum_{t=0}^T \frac{\rho_t}{T} \quad (5)$$

2. 機器人足球追蹤系統

本文所使用之人形機器人平台採雙足全向移，其動控制流程如圖 13 所示，在目標物件進行定位與降低移動造成的誤差後，會給予機器人移動控制指令 $(x_{cmd}, y_{cmd}, \theta_{cmd})$ ，執行雙足機器人場域中移動追蹤的實現。

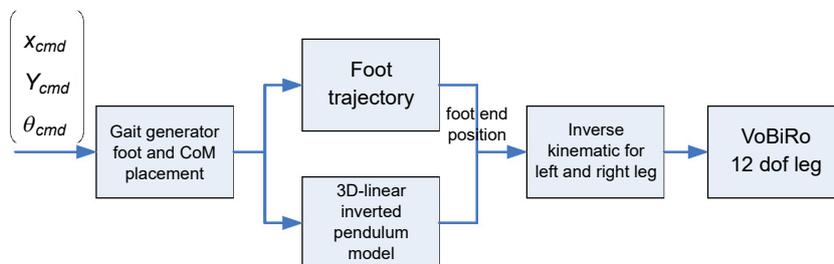


圖 13. 全向移動控制流程。

在雙足機器人在追蹤的實現上，本論文以建立人工位能場 (artificial potential field, APF)，來調整其從不同遠近程度，進行追蹤目標物與靠近目標物件的移動控制。移動行為從目標定位遠近程度，進行機器人對目標物的靠近行為。人工位能場是藉由計算目的地與障礙物之分佈情形而產生的導航控制系統。其中利用位能場的概念，對於目的地給予具有吸引力的位能場 (U_{att})；並且根據障礙物的分佈產生具有排斥力的位能場 (U_{rep})。在該位能場中，吸引力 (F_{att}) 會將機器人導引至所要到達的目的地，同時排斥力 (F_{rep}) 則會依照場中障礙物的分佈讓機器人遠離周圍的障礙物，藉以達到避障之目的。因此在上述所提之位能場環境中，吸引力與排斥力最後將形成一合力向量 (F_{res})，該向量即為機器人在當下位置的運動趨勢方向，如圖 14 所示。相關技術細節請參閱我們先前發表⁽⁸⁾。

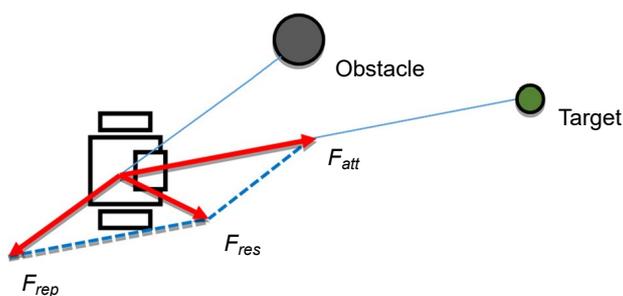


圖 14. 人工位能場應用於避障導航之合力示意圖。

五、實驗結果與分析

1. 足球物件框選結果與機器人物件認知測試

本實驗設計依照不同資料量的訓練集，來比對訓練用資料量增加時，提升辨識率不同之結果，訓練集依序從 100 張隨機足球圖片開始，每次進行 100 次隨機照片之追加，並使用不與訓練集重複之測試集足球圖片 200 張，來測試訓練集隨機圖片增加呈現之訓練後適用結果，此實驗結果涉及第三章節所提之認知廣泛程度驗證、應用場景正確度以及訓練後模型框

選中未覆蓋比例，整體實驗結果表示於表 1 中。由於這些照片均由網路搜尋照片取得，因此在收集到 600 張時，在實務上已經到達一定程度之極限。以應用場景正確度達 92% 而言，其對於實際比賽中，使用平滑濾波器將可避免部分一小斷時間無法正確判斷到球位置之問題。同時，在導入足球定位系統之後，即可針對不同足球離機器人之距離 (0.3 公尺到 1.5 公尺) 進行定位誤差之驗證，此一驗證結果如表 2 所示。

表 1. 足球物件框選結果。

訓練集 (張)	100	200	300	400	500	600
認知廣泛程度 (%)	63.5%	68.0%	72.0%	75.5%	81.5%	83.5%
應用場景正確度 (%)	70.0%	71.5%	75.0%	84.5%	90.5%	92.0%
未覆蓋比例 (%)	27.60%	34.41%	45.88%	29.75%	18.28%	11.11%

表 2. 影像視覺定位之分析結果。

實際座標 (cm)			測量座標 (cm)			影像座標 (Pixel)		馬達角度		誤差 (cm)		
x	y	dis	x	y	dis	x	y	hor	ver	x	y	dis
-60	30	67.082	-61.069	31.416	68.676	91	75	665	597	-1.07	1.42	1.59
-30	30	42.426	-29.446	32.465	43.830	154	89	657	620	0.55	2.47	1.40
0	30	30.000	0.681	30.105	30.113	159	148	500	620	0.68	0.11	0.11
30	30	42.426	28.531	33.094	43.695	163	87	375	620	-1.47	3.09	1.27
60	30	67.082	61.956	30.678	69.135	236	99	375	595	1.96	0.68	2.05
-60	60	84.853	-62.348	61.906	87.861	162	83	665	561	-2.35	1.91	3.01
-30	60	67.082	-28.476	63.314	69.423	167	102	591	576	1.52	3.31	2.34
0	60	60.000	0.391	57.714	57.715	169	97	504	595	0.39	-2.29	-2.28
30	60	67.082	32.287	63.295	71.054	171	88	415	575	2.29	3.30	3.97
60	60	84.853	54.395	64.917	84.694	161	88	375	562	-5.61	4.92	-0.16
-60	90	108.167	-55.108	93.995	108.958	159	88	616	538	4.89	4.00	0.79
-30	90	94.868	-32.455	93.907	99.357	169	108	573	538	-2.46	3.91	4.49
0	90	90.000	1.745	86.413	86.431	162	100	506	554	1.75	-3.59	-3.57
30	90	94.868	34.617	89.672	96.122	157	83	442	554	4.62	-0.33	1.25
60	90	108.167	55.314	91.683	107.077	166	93	405	536	-4.69	1.68	-1.09
-60	120	134.164	-54.481	122.893	134.428	159	90	594	519	5.52	2.89	0.26
-30	120	123.693	-33.041	121.924	126.322	162	101	563	519	-3.04	1.92	2.63
0	120	120.000	2.344	121.258	121.281	159	107	509	519	2.34	1.26	1.28
30	120	123.693	34.306	120.013	124.820	154	103	455	519	4.31	0.01	1.13
60	120	134.164	59.845	116.053	130.575	164	95	418	519	-0.16	-3.95	-3.59
-60	150	161.555	-57.129	153.384	163.678	166	95	580	502	2.87	3.38	2.12
-30	150	152.971	-27.414	151.234	153.699	162	103	546	502	2.59	1.23	0.73
0	150	150.000	3.163	152.357	152.390	168	105	502	502	3.16	2.36	2.39
30	150	152.971	34.612	149.123	153.087	157	103	464	502	4.61	-0.88	0.12
60	150	161.555	58.357	147.142	158.292	163	99	437	502	-1.64	-2.86	-3.26

另一方面，機器人樣貌對於學習上較為複雜，同時在實際應用中所視機器人樣貌會有不同視角之差異，故欲對其增加訓練集圖片時，需考慮原始訓練集上缺乏之元素資訊，故本實驗分成以所視角度與元素角度出發，進行使用初始訓練集 1000 張機器人圖片訓練後之模型測試。其中因本論文所用機器人平台，並未進入到訓練樣本之中，以其為測試用資訊最為恰當，而在比賽中機器人視野所視，多以全身及下半身為常見之影像，故本小節實驗一以機器人全身以及下半身不同角度視野進行，每間隔 45 度蒐集資料一次，其結果如圖 15 所示。相關實驗結果彙整如表 3 所示。

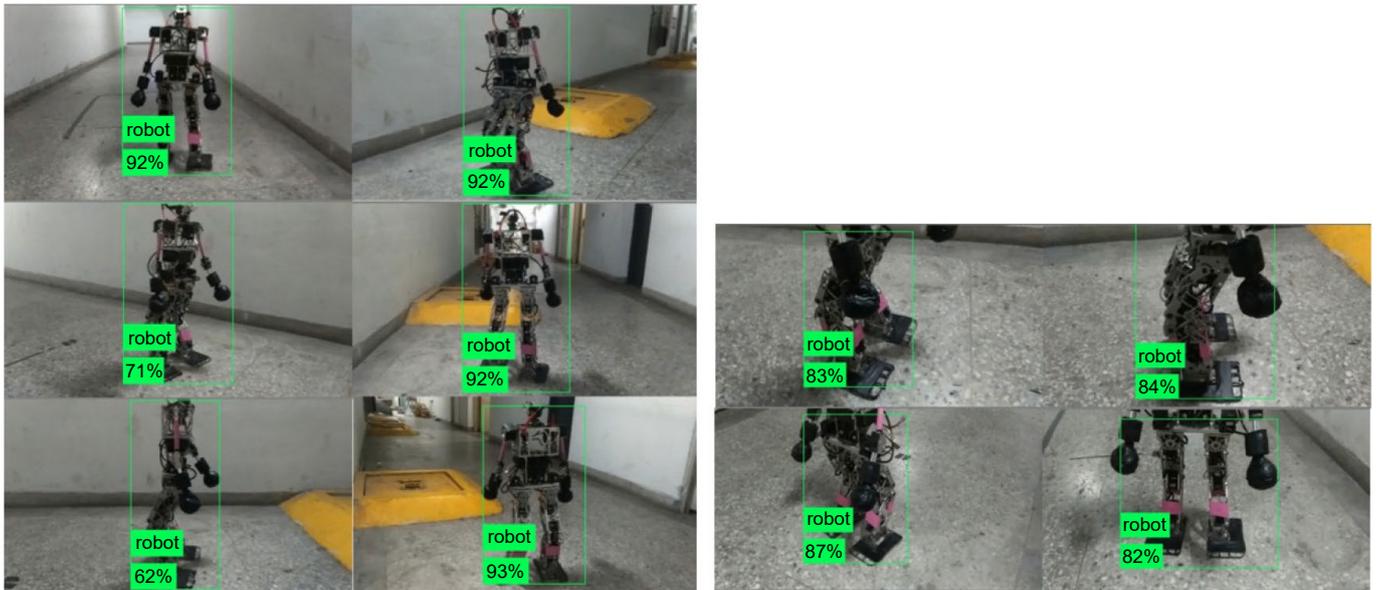


圖 15. 機器人全身不同視角實驗 (左)；機器人下半身不同視角實驗 (右)。

表 3. 全身以及下半身不同角度視野測試結果。

所視範圍	視角	相似度
機器人全身	機器人正後方	91.8%
機器人全身	機器人左側後方	71.3%
機器人全身	機器人左側	62.1%
機器人全身	機器人左前方	91.6%
機器人全身	機器人正前方	91.8%
機器人下半身	機器人右側後方	86.8%
機器人下半身	機器人正後方	82.3%
機器人下半身	機器人左側後方	83.1%
機器人下半身	機器人左側	82.5%

最後，此一系統採用 NVIDIA Jetson TX2 嵌入式人工智慧運算器實現機器人視覺即時偵測、辨識之運算，並以深度學習卷積神經網路開發，導入 You Only Look Once 網路進行足球與機器人類別之偵測與辨識。在實際運作測試下，其達到每秒 5-7 次之運算速度，其對於此一雙足人行機器人在 20 cm/s 行走速度下，影像處理速度預備運作可行性。

2. 機器人足球追蹤系統於 RoboCup 2017 之驗證

為了驗證機器視覺系統應用於機器人上之成果，將此系統整合於雙足人形機器人上，並參與 RoboCup 2017 機器人足球賽。其中獲得球權一方得以由中場進行開球，機器人被放置於進攻方場域圓弧內，自主找到被放置場中的賽用足球進行踢球，其實際比賽過程如圖 16 所示。RoboCup 2017 機器人足球賽比賽過程中，賽用足球會隨場地非均勻表面與機器人進攻或防守動作，落入任意區域中，使用本論文視覺系統之機器人可以穩定的找到賽用足球，並以人工引力場吸引，將目標足球順利踢進至球門區域。其實際賽中踢球過程如圖 17 以及圖 18 所示。



圖 16. RoboCup 2017 機器人足球賽進攻方開球。

六、結論

根據實驗結果證明，本文所提出的深度學習機器視覺系統，能夠正確且有效識別出足球目標物件，並能在較為複雜的視野環境中，檢視出一般影像處理流程，較難認知的機器人。最後結合定位系統，使機器人能根據影像判斷足球之位置，定位於實際場地中，運用在 RoboCup 機器人足球賽中，人形足球機器人辨識比賽物件能力提升，將有助於機器人對比



圖 17. RoboCup 2017 機器人足球賽場中踢球過程。

賽現場狀態更有掌握力。本研究結果實際應用於日本名古屋 RoboCup 2017 國際人形機器人足球賽之中，打敗多數以傳統顏色特徵偵測目標物支隊伍，並取得第二名之成績，得以驗證此視覺系統應用之成效。

參考文獻

1. K.F. Lin, "Development of Humanoid Soccer Robot Machine Vision System with Deep Learning," *Master Thesis, Department of Electrical Engineering, National Taiwan University of Science and Technology*, (2017).
2. I. Awaludin, P. Hidayatullah, J. Hutahaean and D. G. Parta, "Detection and object position measurement using computer vision on humanoid soccer," *2013 International Conference on Information Technology and Electrical Engineering (ICITEE)*, Oct. 07-08, pp. 88-92, (2013).
3. M. N. Sudin, M. F. Nasrudin and S. N. H. S. Abdullah, "Humanoid localisation in a robot soccer competition using a single camera", *IEEE International Colloquium on Signal Processing & its Applications*, March 7-9, 77 (2014).
4. H. Minakata, Y. Hayashibara, K. Ichizawa, T. Horiuchi, M. Fukuta, S. Fujita, H. Kaminaga, K. Irie and H. Sakamoto, "A method of single camera Robocup humanoid robot localization using cooperation with walking control", *IEEE International Workshop on Advanced Motion Control*, March 26-28, 50 (2008).
5. J. S. Chiang, C. H. Hsia, S. H. Chang, W. H. Chang, H. W. Hsu, Y. C. Tai, C. Y. Li and M. H. Ho, "An efficient object recognition and self-localization system for humanoid soccer robot", *SICE Annual Conference*, February 28, 2269 (2010).



圖 18. RoboCup 2017 機器人足球賽場中進球過程。

6. W. Chen, T. Qu, Y. Zhou, K. Weng, G. Wang and G. Fu, "Door recognition and deep learning algorithm for visual based robot navigation", *2014 IEEE International Conference on Robotics and Biomimetics*, December 5-10, 1793 (2014).
7. S. Contreras and F. D. L. Rosa, "Using deep learning for exploration and recognition of objects based on images", *2016 XIII Latin American Robotics Symposium and IV Brazilian Robotics Symposium*, October 08-12, 1 (2016).
8. C.H. Kuo, H.C. Chou, S.W. Chi, and Y.D. Lien, *International Journal of Humanoid Robotics*, **10** (3), 1350021-1 - 1350021-36, (2013).

作者簡介

林昆鋒先生為國立臺灣科技大學電機工程碩士，目前任職於建興儲存科技公司。

Mr. Ken-Fun Lin received his M.S. in Electrical Engineering from National Taiwan University of Science and Technology. He is currently an Engineer in Solid State Storage Technology Corporation (SSSTC) Co., Ltd.

郭重顯先生為國立臺灣大學機械工程博士，現為國立臺灣大學機械系教授。

Dr. Chung-Hsien Kuo received his Ph.D. in Mechanical Engineering from National Taiwan University. He is currently a Professor in the Department of Mechanical Engineering at National Taiwan University.