

# 2019「中技社AI創意競賽」

## 2019 CTCI Foundation AI Innovation Competition

### AI 與創新服務 AI & Innovation Service



## 咖啡來了歐

國立臺灣師範大學 電機工程學系

許陳鑑教授

研究生 黃品叡博二、陳薪鴻碩二、蔡仁凱碩二



### 作品概述

從工廠生產延伸到居家生活，自動化已是現在的趨勢。隨著自動化的發展，越來越多的自動化已不再只是大量重複性動作，而這些動作的規劃需要專業人員來製作，這不只耗費時間，也需要一定程度的訓練學習。因此本團隊製作一套系統，只要透過人類的示範，就能使機器人學習並重現行為。為了適應多種環境中的變化，機器人不僅止於「模仿」，更必須「理解」示範的目的，才能完成人類的任務。若機器人只會模仿而無法認知環境，那環境稍有變化，機器人就只能無意義的作動。

本團隊的示範學習系統利用視覺影像與AI機器深度學習架構辨識環境中的物件，使用深度攝影機與時間關係定位並追蹤物件，再透過物件各種狀態推斷出人類的行為目的，並將整個示範流程切分成多個動作片段，透過分析將動作與互動之物件作結合製作動作庫，動作庫中記錄示範的流程與互動的物件，再經由路徑規劃，依序透過硬體中的機器手臂重現人類示範的目的。在此作品中我們以泡咖啡為例，動作包含拿取不同的杯子、倒入咖啡、啟動咖啡機等。

### 創意及核心技術

一台模仿機器人，只要記錄所有運動軌跡，重複動作即可，但要學習，就必須有理解的過程。我們以居家的泡咖啡為例，拿起杯子放入咖啡機中再按下咖啡機按鈕，如果沒有識別杯子或咖啡機的位置，一旦物件位置改變就無法重現人類的示範，因此機器人對環境的認知與識別是至關重要的。

機器人對環境的認知與識別的方法有很多種，本團隊使用RGB-D攝影機獲取環境中的三維場景，利用深度學習YOLO v3架構實現物件辨識與深度影像結合，取得物件的三維座標，藉由物件與時間的關係追蹤此物件，透過此物件追蹤模組可以推斷出物件的位置與各種狀態，進而推斷出人類演示的目的。

以泡咖啡為例，偵測到咖啡機撥桿被撥動，則可推斷出人類啟動了咖啡機，或是偵測到杯子在移動，則可推斷人類正移動杯子。當知道物件狀態與推斷出的人類行為後，我們可將整體示範流程切分出各種小行為，例如：拿、放、倒水...等，切分出動作片段後結合物件的位置製作動作庫。如圖1，在此例子中機器人學習到的是「拿杯子」、「到」、「某個地點」，值得注意的是「拿杯子」的部分，機器人是學習到「拿A杯子」而不是「到某個地點去拿」，如果是「到某個地點去拿」那就只是單純的模仿，當杯子的位置不同時就會出錯。並且依據不同動作做出對應的路徑規劃，再利用TCP/IP與機器手臂作溝通控制完成任務，即可重現人類示範的目的。

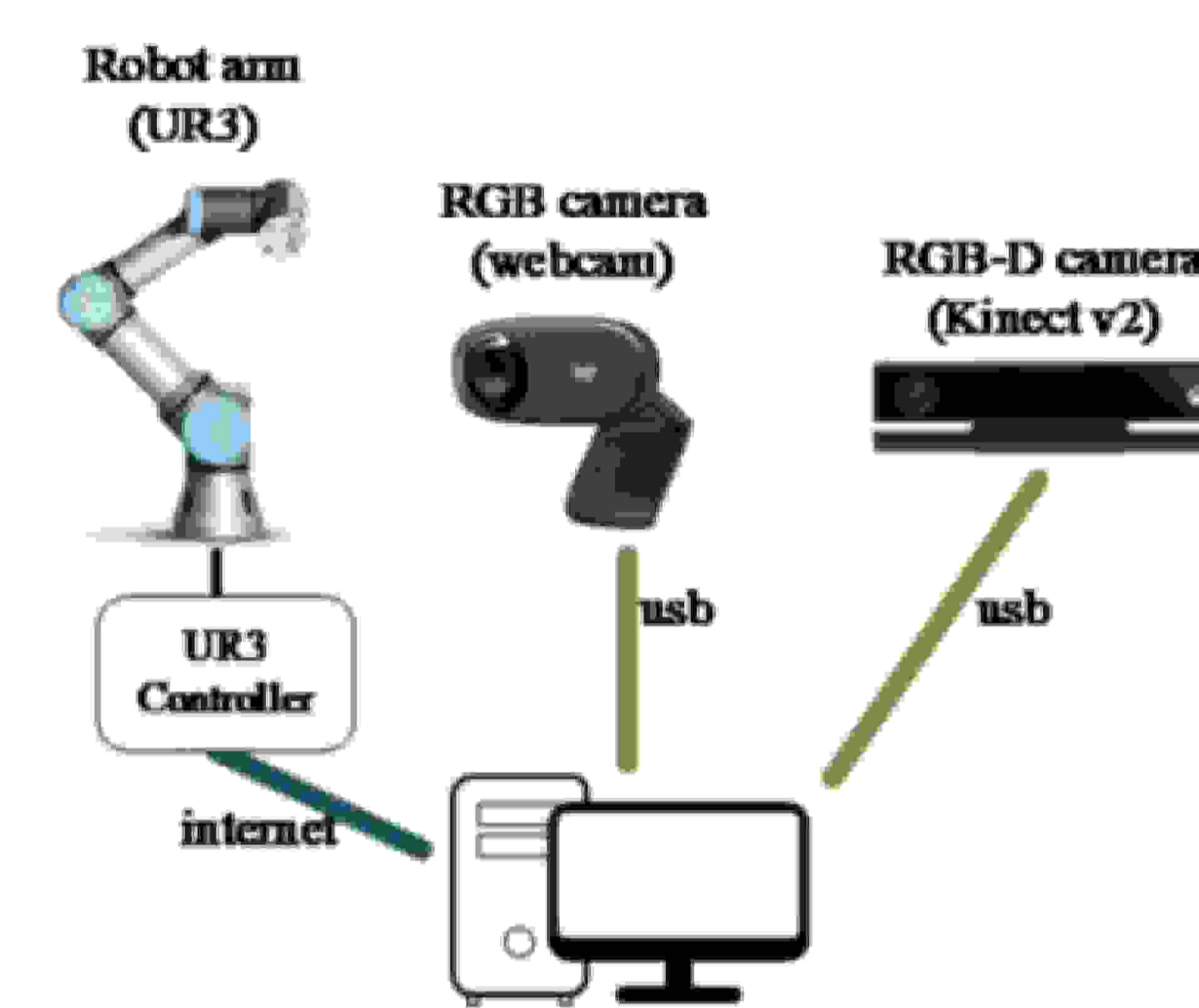


圖1、系統硬體架構圖

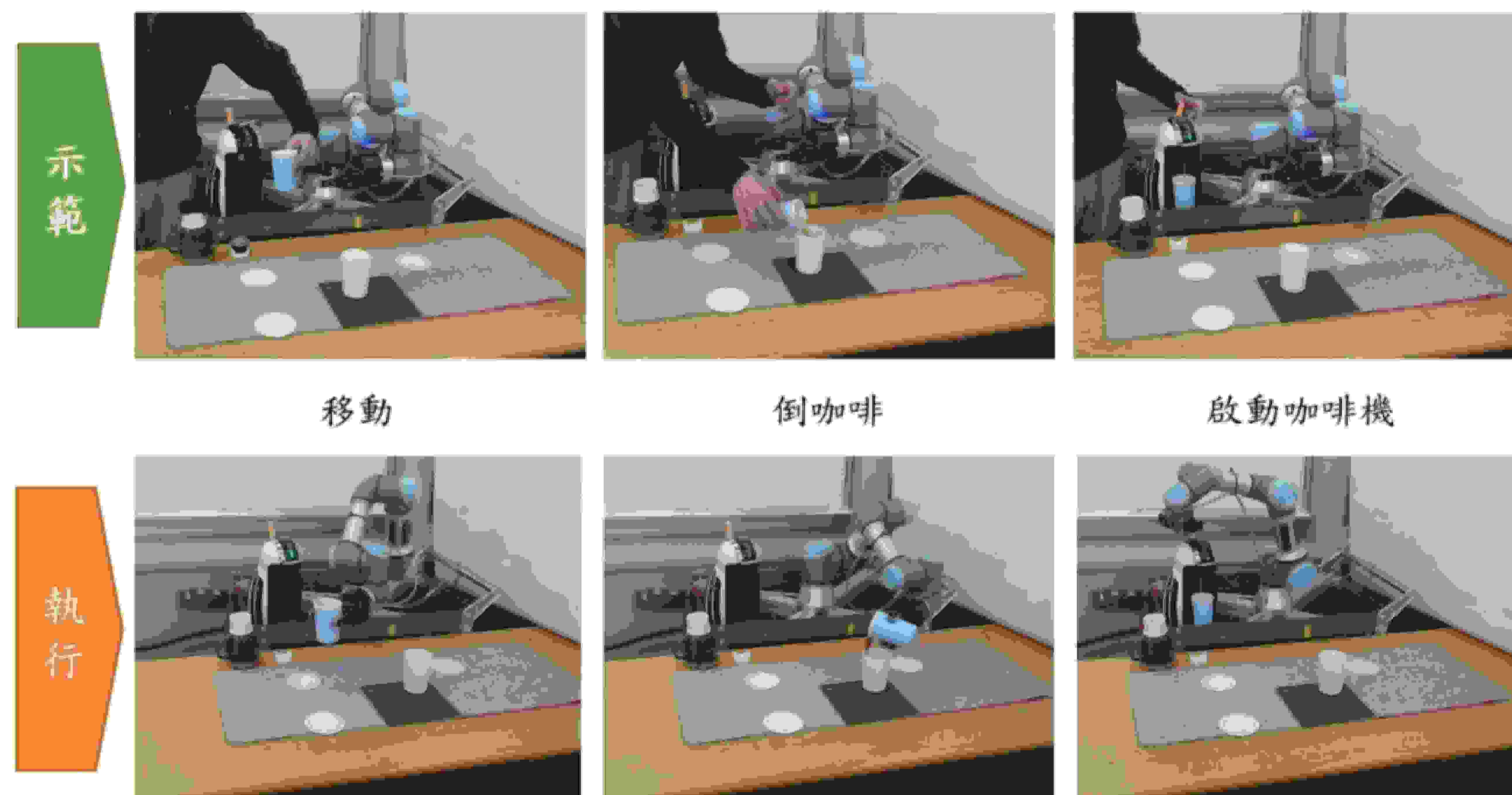


圖2、實際範例與執行成果

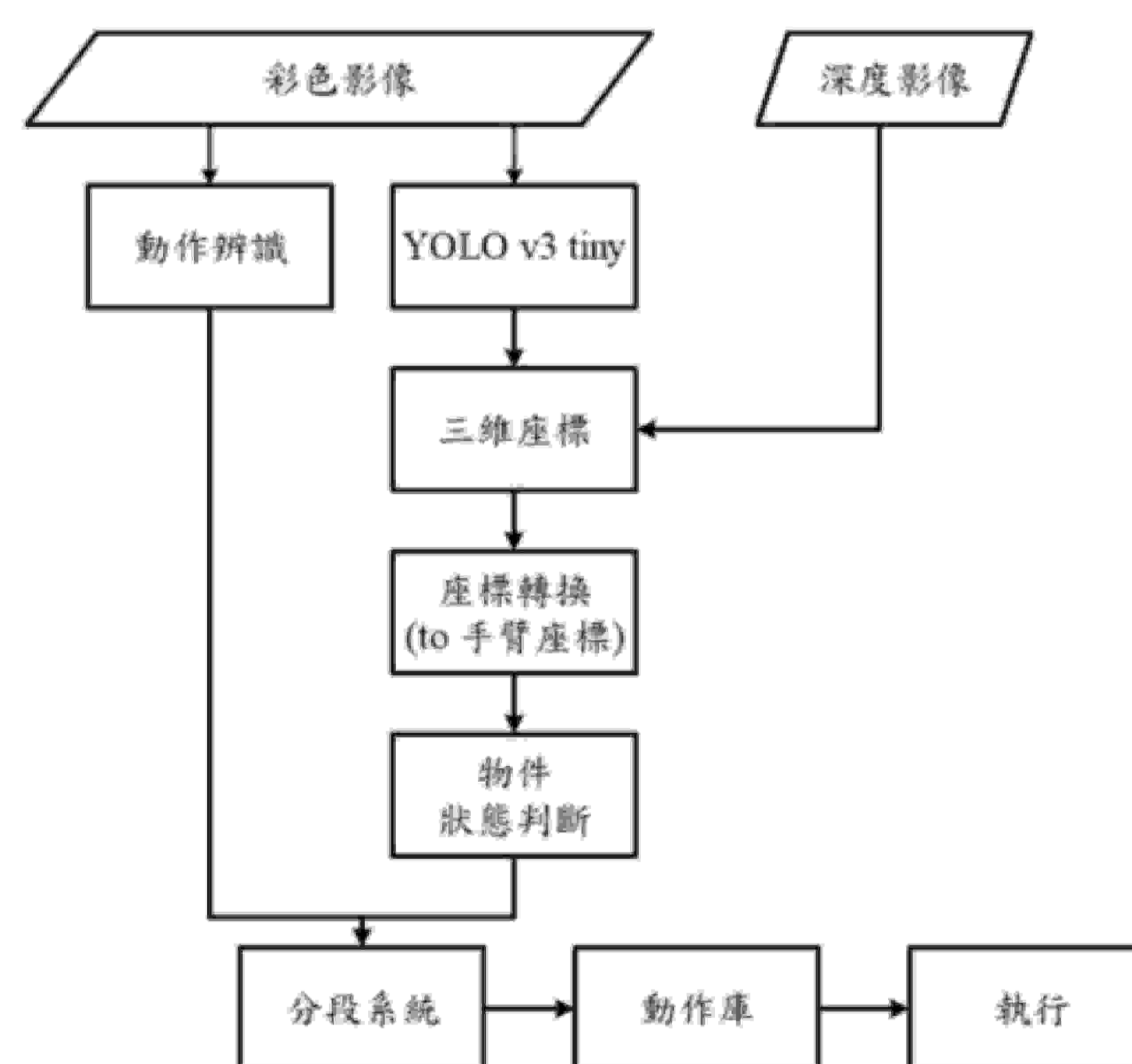


圖3、系統流程圖

### 產業及社會貢獻

示範學習系統可應用在生產線與居家，在生產線上已有用手牽引手臂並錄製手臂末端點路徑的方法，此方法的最大優點是不用撰寫程式，且不需要專業人員就可以快速規劃手臂運動流程，但此種方法無法容許任何環境的變動，工件的位置等都必須與示範時的位置一樣，若導入示範學習系統，就能藉由理解示範的目的與環境的偵測來適應許多變動，可以大大減少設計生產流程的時間。

在居家的部分，因為設計者不可能設計出能因應所有狀況的固定程式，因此透過人在家中示範，讓機器人可以學習人類的目的與流程，才更能適應多變的環境。



財團 中技社  
法人  
CTCI FOUNDATION