



新北

AI

微課程讀本

局長序

人工智慧 (Artificial Intelligence, 簡稱 AI) 是當今最具潛力的發展領域之一，也是新北市科技教育關鍵的發展方向。為了讓更多師生能夠接觸和了解 AI，並提升師生對 AI 技術的認識與應用，教育局積極推動 AI 相關的課程教材讀本計畫。本次推出的「新北 AI 微課程讀本」，正是希望透過簡潔易懂的語言，讓師生可以更輕鬆地踏進 AI 的新世界。

新北市政府教育局於 112 年委由錦和高中做為 AI 教育發展基地，邀集臺灣師範大學、臺灣大學、臺灣科技大學及臺北科技大學，四所大學教授組成學者團隊指導，並由新北市教師們擔任編輯群，投入教材研發及課堂教學，積極推動 AI 科技教育。同時也透過產官學合作的方式，以現今較為普及的設備來進行課程設計，希冀能夠更有效地落實教學端之教學推動。

本書作為銜接國高中的教材，涵蓋了 AI 的基礎概念、應用範例和實作活動等內容，並以專題課程的方式來啟發學生的創造力和創新思維，鼓勵他們進一步探索 AI 技術在各個領域的應用，將所學應用於實際問題的解決中。透過這樣的學習，學生們不僅可以提高自己的技能和能力，更好地迎接未來社會和職場的挑戰，也能夠為社會和經濟發展做出積極的貢獻。

最後，我要感謝所有為本書付出心力的人員，新課綱科技領域帶給學生全新的學習視野，AI 教育敲開了數位時代通往世界的大門，希望新北市所有師生們可以從本書獲得啟發和收穫。



新北市政府教育局

局長 **張明文**



主編序

你今天使用了 ChatGPT 了嗎？自從美國的 OpenAI 於 2022 年底推出人工智慧 (Artificial Intelligence, 簡稱 AI) 聊天器 ChatGPT 後，對於教學現場帶來許多新的可能與挑戰。人工智慧應用是下個世代的主流，也將全面改變日常生活的樣態。然而，在學校既有課程中，如何藉由科技領域課程融入相關素養，引領學生調適與學習，廣泛運用在各個領域層面，會是新一波學習革命。

當未來面臨充滿易變性 (volatility)、不確定性 (uncertainty)、複雜性 (complexity)、模糊性 (ambiguity) 的 VUCA 的世界，高中階段的學生需要高度的自主學習，「把生活情境當作一個大型專題」，培養系統思考與解決問題的態度，同時，也更期待他們擁有運用人工智慧的素養與能力。新課綱課程中的彈性學習時間規劃提供學生很好的機會，而本書即可讓學生運用彈性學習時間，自主學習體驗人工智慧的堂奧。

本書是專為國高中學生打造的 AI 教材讀本，內容包含 AI 的基礎概念、生活範例和實作活動等。透過本書，學生們可以在專題課程的引導下發揮想像力與創新思維，更進一步探索 AI 技術在各個領域中的應用，並學以致用地解決生活中的問題。

謝謝臺灣師範大學科技應用與人力資源發展學系張玉山教授與臺灣大學聯盟李坤彥教授與陸敬互教授及北科大李俊賢教授的指導，帶領著新北市國高中跨校教師編輯群，包括錦和高中、三重商工、江翠國中、永和國中、福和國中與、北市大同高中科技教師，共同編輯本書。感謝新北市政府教育局中等教育科的支持，讓科技教育具體而微地落實在高中階段，學生能依著自己的興趣與天賦，嘗試各種學習體驗。讓學生長出科技的翅膀，面對未來人工智慧的挑戰，我們準備好了！

錦和高中

校長

指導教授簡介

(依姓名筆劃數順序排列)

李坤彥

- 現職** 國立臺灣大學工程科學及海洋工程學系教授
- 學歷** 美國普渡大學博士
- 經歷** 臺大與工研院合設奈米科技中心副主任
國立臺灣大學工程科學及海洋工程學系副教授
國立臺灣大學工程科學及海洋工程學系助理教授

李俊賢

- 現職** 國立臺北科技大學電機工程學系教授
- 學歷** 國立交通大學電機與控制工程博士
- 經歷** 美國紐澤西理工學院 (New Jersey Institute of Technology) 電機與電腦研究所 訪問學者
工業技術研究院資訊與通訊研究所研究員

張玉山

- 現職** 國立臺灣師範大學科技應用與人力資源發展學系教授
- 學歷** 國立臺灣師範大學博士
- 經歷** 國立臺灣師範大學科技應用與人力資源發展學系系主任
國立臺灣師範大學師培處組長
教育部科技領域教學研究中心主持人

陸敬互

- 現職** 國立臺灣科技大學電機工程系教授
- 學歷** 國立臺灣大學資訊工程博士
- 經歷** 元智大學資訊傳播學系 助理教授
元智大學大數據與數位匯流創新中心物聯網計畫共同主持人
Intel-NTU Connected Context Computer Center, M2M 節能計畫共同主持人
臺大智慧生活科技整合與創新研究中心博士後研究

作者簡介

(依姓名筆劃數順序排列)

汪殿杰

- 現職** 臺北市立大同高級中學
- 學歷** 實踐大學產品與建築設計研究所
- 經歷** 科技領綱課綱委員
高中生活科技學科中心研究教師
臺北市生活科技輔導團

林易民

- 現職** 新北市立三重高級商工職業學校
- 學歷** 國立臺灣師範大學資訊教育所
- 經歷** 新北市立三重高級商工職業學校資料處理科科主任

林欽鴻

- 現職** 新北市立錦和高級中學
- 學歷** 國立臺灣師範大學科技應用與人力資源發展學系
- 經歷** 新北市立錦和科技中心組長
新竹縣立竹東科技中心主任

高靖岳

- 現職** 新北市立永和國民中學
- 學歷** 國立臺灣師範大學科技應用暨人力資源發展學系科技教育學碩士
- 經歷** 新北市科技領域輔導團輔導員
教育部中學科技領域研究中心計劃執行秘書
教育部國教署 111 年度科技領域中央課程與教學輔導諮詢教師

陳筠雅

- 現職** 新北市立福和國民中學
- 學歷** 國立臺灣師範大學科技應用與人力資源發展學系
- 經歷** 新北市立福和國中生活科技教師

蔡志聰

現職 新北市立江翠國民中學

學歷 國立清華大學特殊教育研究所

經歷 新北市立江翠國中設備組長

蔡勝安

現職 新北市立江翠國民中學

學歷 國立臺灣師範大學工業科技教育研究所

經歷 新北市立江翠科技中心組長

新北市科技領域輔導團

新北市立江翠國中設備組長

蔡釋鋒

現職 新北市立樟樹國際實創高級中學

學歷 國立高雄師範大學工業科技教育學系研究所

經歷 新北市立樟樹科技中心組長

臺東縣立初鹿科技中心主任

盧麗如

現職 新北市立福和國民中學

學歷 國立臺灣海洋大學教育研究所

經歷 新北市立福和國中資訊組長

國立華南高商電子商務科

新北市立三重商工資處科

目錄

loading...



第一篇 AI 基礎：大數據及演算法

單元 1	前言	2
單元 2	認識大數據	3
單元 3	認識演算法	7
單元 4	處理大數據的流程與應用	10

第二篇 AI 原理：影像辨識及語音辨識

壹、影像辨識

單元 1	人工智慧與影像辨識	28
單元 2	生活中影像辨識的應用	29
單元 3	影像辨識的原理	31
單元 4	線上圖像創作產生的 AI 應用	37
單元 5	影像辨識的挑戰	41

貳、語音辨識

單元 1	語音辨識的原理—認識聲音訊號與處理	42
單元 2	語音辨識的運作流程與類型	47
單元 3	語音辨識的升級—以文字為主的自然語言處理技術	49
單元 4	語音辨識的挑戰	53

第三篇 AI 應用與實作

單元 1	AI 語音辨識的應用與思考	56
單元 2	生活中的語音辨識	60
單元 3	AI 視覺辨識的智慧新生活	83
單元 4	生活中的影像辨識	104

第一篇



AI 基礎 — 大數據及演算法



單元 1

前言

作者：盧麗如·陳筠雅—福和國中 | 林易民—三重商工

現在當你到國外，可以使用語言翻譯機跟外國人愉快地聊天。在未來，你可以使用汽車的自動駕駛技術，你可以同時玩著遊戲，然後安全到達目的地。這一切都是人工智慧（Artificial Intelligence，簡稱 AI）所帶來的未來影響。什麼是人工智慧？1956 年計算機科學家約翰·麥卡錫（John McCarthy）提出人工智慧¹的概念，主要想法是要讓機器看起來就像人類一樣有智慧，可以思考解決問題。

大數據是人工智慧成長需要的糧食 JS

日常生活中人類遇到問題時，會運用大腦思考解決方法，為了加快處理問題的流程，人類發明了電腦來協助解決問題。但是如果沒有人類的協助，電腦也無法發揮任何功能，因為當電腦遇到問題時，它必須使用人類提供的策略方法來解答問題。這些由人類提供給電腦解決問題的方法策略稱為演算法（Algorithm）。但是如果需要解決的問題太複雜或者是沒有明確的解題策略，就無法使用演算法來解決。此時只能使用人工智慧中的機器學習（Machine Learning，簡稱 ML）來建立模型（Model），然後使用模型來解決問題。機器學習會透過大量資料，模仿人類大腦解決問題的能力，建立模型。越多的學習資料，就可以獲得更好的機器學習解決方案。因此大數據是 AI 的基礎。如何處理大量的資料及數據，就是機器學習的基礎。

1 參考資料：維基百科 - 約翰·麥卡錫。網址：<https://reurl.cc/gQKo7p>



單元 2

認識大數據

作者：盧麗如·陳筠雅—福和國中 | 林易民—三重商工

以下先對大數據的相關概念進行介紹。

一·什麼是資料？

資料 (Data, 又稱數據)

在我們日常生活中產生的文字、數字、符號、語音、影像、視訊等沒有經過整理或不具意義的資料都算是 Data。這些資料必須經過處理才會產生有意義的資訊 (Information)，資訊經過分析後可以提供推論或決策判斷的訊息，就稱為知識 (Knowledge)，最後逐漸形成智慧 (Wisdom)。



▲圖 1-1 資料的分類²

2 參考資料：陳宗和王雅惠 (2022) 《108 版高中資訊科技 (全)》。臺灣：全華圖書有限公司。

結構化資料 (Structured data) JS

傳統上收集資料時，為了方便整理，會用表格的方式來呈現，即先固定欄位和格式後，將性質相似的資料放在一起，然後按照順序將資料一筆一筆記錄到表格中。由於表格的欄位是固定的，所以非常容易管理和儲存。例如：薪資會計系統、員工出缺勤紀錄、商品進出貨記錄等都可以用結構式的方式來儲存。

非結構化資料 (Unstructured data) JS

隨著數位科技進步，電腦、平板、手機等設備載具連上網路後產生的數位內容更多元，包含文字、圖片、聲音、影片等這些格式不同的資料難以用表格方式來呈現。這些資料就可以稱為非結構化資料。舉例：FB、Instagram、Twitter 等社群互動數據、影音圖片、網路搜尋索引記錄等。

資料 (Data) 與資訊 (Information) 的差異 JS

我們日常生活中，產生的文字、圖形或聲音等訊息可統稱為資料 (Data)，這些訊息代表某種現象或事實的紀錄。將蒐集來的資料整理成有意義可以決策或判斷的情報就可稱為資訊 (Information)。經過計算、分析、彙整的過程又稱為資料處理 (Data Processing)。

舉例：1. 資料 → 2. 處理 → 3. 資訊

1. 資料：同學們的身高、體重（甲：179cm、59kg、乙：156cm、59kg、丙：167cm、60kg）
2. 處理：BMI 計算公式 ($BMI = \text{kg}/\text{m}^2$)， $BMI < 18.5$ （過輕）； $18.5 \leq BMI < 24$ （健康）； $BMI \geq 24$ （過重）
3. 資訊：判斷是否過重？（甲： $BMI = 18.41$ （過輕），乙： $BMI = 24.24$ （過重），丙： $BMI = 21.51$ （健康）



▲圖 1-2 資料和資訊的差異

電腦中常用的資料儲存單位

電腦的開機／關機需要電源「開 ON」、「關 OFF」，因此，當資料輸入電腦時，電腦需要將電子訊號轉成數位訊號「0」或「1」，這樣才能辨識或儲存資料。我們會將一個「0」或一個「1」稱做是一個「位元」(bit)，「位元」是電腦資料儲存的最小單位。

▼表 1-1 儲存單位的名稱

單位	中文名稱	換算
bit	位元	最小單位
Byte	位元組	1Byte = 8 bits
KB	千位元組	1KB = 2^{10} Bytes = 1024Bytes $\approx 10^3$ Bytes
MB	百萬位元組	1MB = 2^{20} Bytes = 1024 KB $\approx 10^6$ Bytes
GB	吉位元組	1GB = 2^{30} Bytes = 1024 MB $\approx 10^9$ Bytes
TB	兆位元組	1TB = 2^{40} Bytes = 1024 GB $\approx 10^{12}$ Bytes
PB	拍位元組	1PB = 2^{50} Bytes = 1024 TB $\approx 10^{15}$ Bytes
EB	艾位元組	1EB = 2^{60} Bytes = 1024 PB $\approx 10^{18}$ Bytes
ZB	皆位元組	1ZB = 2^{70} Bytes = 1024 EB $\approx 10^{21}$ Bytes

二．什麼是大數據 (Big Data) ？

大數據是開啟人工智慧的鑰匙

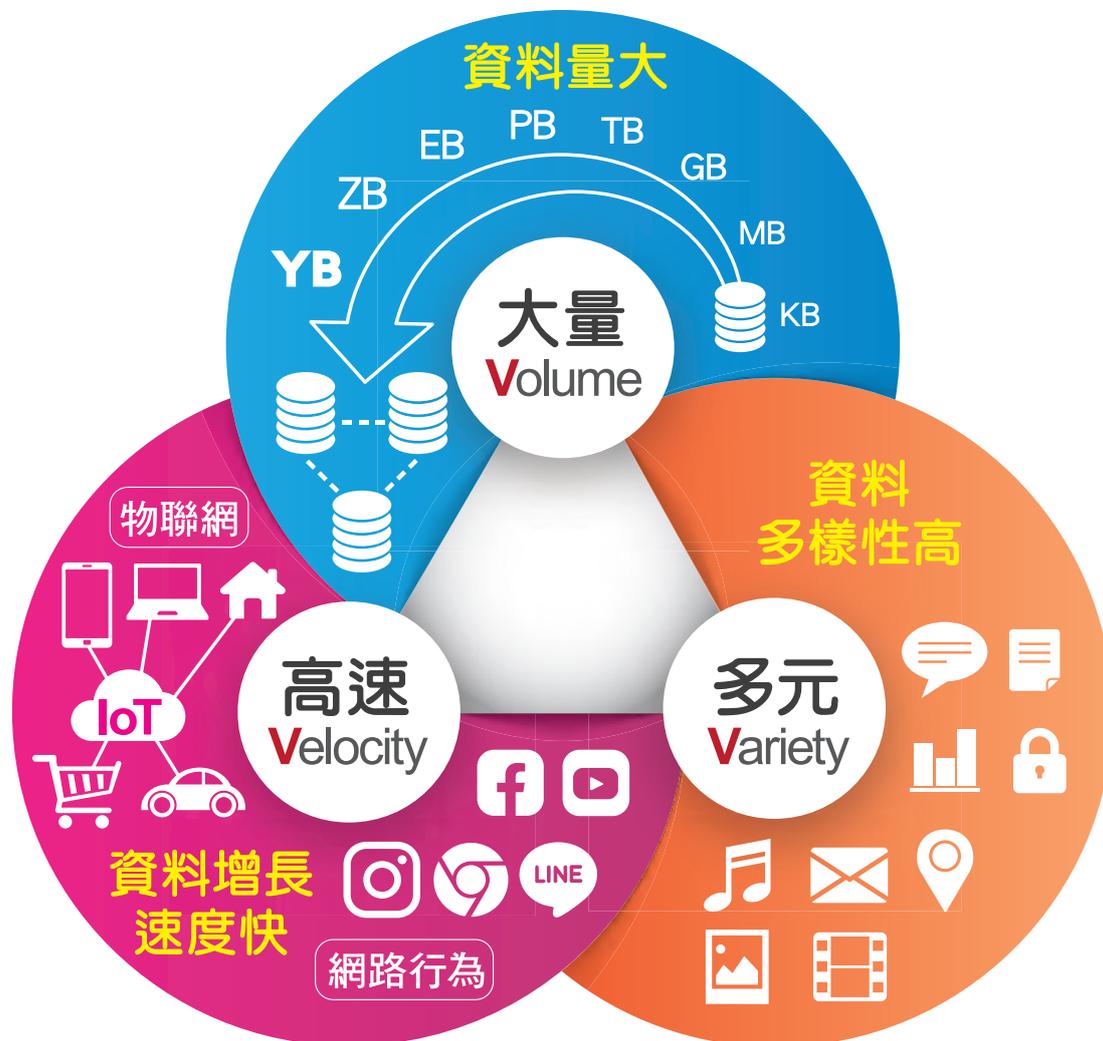
由於網際網路的出現，人們對於資料的使用量越來越大，舉例來說，臉書從創立以來就是用戶最多、廣告營收最高、且最具主導地位的平臺，在 2022 年底為止擁有著高達將近 30 億的活躍用戶³，如果 30 億的使用者，每個人都按一個讚、貼一張圖片、傳送一則訊息，那麼每天系統就需要處理 30 億個讚、30 億則訊息及 30 億張相片。傳統的資料處理的演算法已經不敷使用，也因此催生了大數據這個概念。

2011 年資料分析師萊尼⁴ (Doug Laney) 在研究報告中提出大數據的概念，大數據 (Big Data，又稱巨量資料) 是指龐大複雜資料的集合體，具有以下 3 個特徵：①大量 (Volume)：一般電腦設備無法儲存、無法分析的資料量。舉例：若產生 1TB 以上資料，相當於需要 128 個 8GB 的隨身碟儲存資料。②多元 (Variety)：資料呈現種類多元。舉例：

3 參考資料：2023 年 Facebook 用戶分佈與成長前景如何？主要趨勢為何？ | OOSGA。網址：<https://zh.oosga.com/facebook-overview/>

4 參考資料：Laney, Douglas. "The Importance of 'Big Data': A Definition". Gartner. Retrieved 21 June 2012.

文字、圖片、聲音、影片。③高速（Velocity）：資料增長速度快。舉例：物聯網（IoT）無數的感測器，隨時隨地都在產生大量感測資料紀錄。



▲圖 1-3 3V 概念圖

三．現實生活中的大數據

現今社會，人們常常使用各式各樣的科技產品與網路服務，而這些工具所產生的數據量皆非常龐大。例如：通訊軟體（LINE）上傳貼文、社群軟體（FB、IG）直播影片、搜尋引擎（Google、Yahoo）搜尋關鍵字、網路地圖（Google Maps）上傳景點照片、電商平臺（Electronic Commerce）下單資料、網路相簿（Google 相片）上傳分享照片等，都是現實生活中的大數據例子。



單元 3

認識演算法

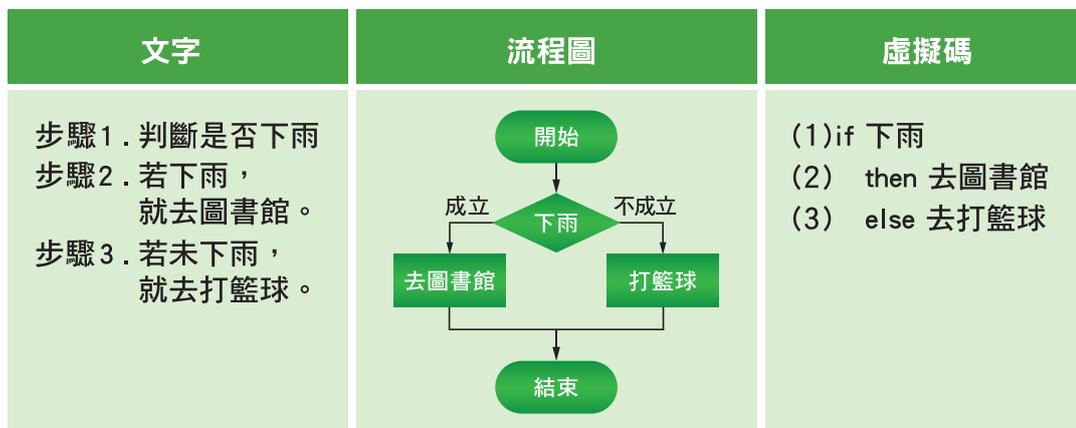
作者：盧麗如·陳筠雅—福和國中 | 林易民—三重商工

一·演算法與電腦解題⁵

演算法 (Algorithm) 定義

人類使用電腦的目的是要幫助我們解決問題，電腦是機器不像人類有大腦可以思考判斷，電腦需要人類給定明確的指令或步驟才能執行。解決問題的方法有很多種，設計出來的演算法可能不同，所以同一個問題的解題方法可能有多種不同的演算法，雖然方法不同，但都可以得到相同結果。演算法可以用文字敘述、流程圖 (flowchart) 或虛擬碼 (pseudocode) 的方式來表示解決問題的步驟，這些操作步驟要符合以下特性：

1. 輸入 (Input)：可以有 1 個或多個輸入值，也可以無輸入值。
2. 輸出 (Output)：至少要有 1 個輸出結果。
3. 明確性 (Definiteness)：輸出結果或答案要正確。
4. 有限性 (Finiteness)：有限時間可以結束完成。
5. 有效性 (Effectiveness)：可以具體執行。



▲圖 1-4 3 種表示法比較⁶

5 參考資料：蔡志敏和張國恩 (2010)。普通高中資訊科技概論。臺灣：松崗資訊有限公司。

6 參考資料：國中資訊科技第六冊 (3 下)，臺灣：康軒文教事業。

日常生活中的演算法

統一發票對獎是搜尋演算法 (Searching Algorithm)

當我們在核對統一發票是否中獎時，必須逐一檢視每張發票的號碼，然後再與中獎號碼進行比對，以此判斷是否中獎。假設有 100 張發票需要對獎，最少要看 100 張發票上的末三碼，總共 300 個數字。

開學時老師依同學身高安排座位可以用插入排序 (Insertion Sort) 演算法

先挑選同學 A 排在第一位，再挑選同學 B 與同學 A 比較身高，若同學 B 比同學 A 高，同學 B 排在第二位；再挑選同學 C 比較，同學 C 比目前排在第一位的同學 A 矮，把同學 C 排在第一位，原本序列中的同學每人都往後移一個位置。依照此規則將所有同學插入適當的位置，就可完成身高排列的順序，在資料量少的時候十分有效率。

隨著處理的資料量不同，我們需要使用不同的演算法來解決問題。舉例學校公告注意事項為例來說，如果只要通知某個班級的幾位同學，學校可以請班長一一轉告。但是如果快速通知全校的所有師生，請人一一通知，會太慢而且可能有疏漏，這時如果使用廣播系統，就可以快速通知所有師生。大數據由於處理的資料量非常大，所以處理資料的演算法，也與處理少量資料的演算法有所不同。

二．處理少量資料的演算法 JS

傳統處理少量資料的演算法包括循序搜尋、二分搜尋、泡沫排序、插入排序、選擇排序等，這些演算法簡單高效，不需要平行和分散運算的功能，因此適用於處理少量資料。舉例來說：班級教室每個同學都有一個置物櫃（假設每個櫃子都一樣，無標記任何記號，沒有照座號排序），要找到（座號 1 號）同學的置物櫃，就需要一個接著一個打開確認，「一個接著一個」就是循序的動作，當找到（座號 1 號）時就停止動作，否則就繼續下一個尋找目標，最好的結果是第一個找的置物櫃就是（座號 1 號）同學的，最差的結果是全部都找過一遍，但沒有（座號 1 號）同學的櫃子。

以下說明循序搜尋法（Sequential Search）其演算法的操作步驟：

1. 定義一個待搜尋的資料。
2. 定義一個待查找的項目（目標）。
3. 從資料的第一項開始進行搜尋，並逐一與目標項目進行比較。
4. 如果找到了目標項目，則結束搜尋，並返回目標項目的索引。
5. 如果沒有找到目標項目，則繼續搜尋下一項，直到搜尋完整個資料。
6. 如果找遍了整個資料仍然未找到目標項目，則返回一個特殊的值（例如 -1）表示未找到。

三．處理大數據的平行演算法

團結力量大，電腦也是一樣。「分而治之」的策略解決傳統電腦技術的困境。早期的電腦只有單一處理器，所有機器工作時是以「指令」為單位，細分成一連串的基本運算，然後單顆處理器「循序」，一個一個處理指令。處理器的時脈和功能有極限，不可能只靠加快電子信號傳輸速度，或擴充單一處理器的運算單元規模，就可無限提升單一處理器的運算效能。因此，電腦科學中出現「平行性」的概念，可利用「多顆處理器」或「多臺電腦」，同時完成「單一工作不同部份」或「幾件不同工作」。目前常用來處理大數據的平行演算法包括：分散式運算、MapReduce 等。這些演算法通過分散計算和分布式儲存，以解決大數據中的計算問題和分散儲存問題。

舉例來說：有一間廚房只有一位廚師需要獨自處理完成 10 道菜，因為只有一位廚師，所以第 1 道菜煮好，才能煮第 2 道菜，然後一直到最後一道菜完成。如果廚房中同時有 10 位廚師，每一位廚師都準備 1 道菜，同時間可以馬上做好 10 道菜，速度快 10 倍。



▲圖 1-5 1 位廚師做料理



▲圖 1-6 10 位廚師做料理



單元 4

處理大數據的流程與應用

作者：盧麗如·陳筠雅—福和國中 | 林易民—三重商工

一· 處理大數據的流程 JS

1. 資料的蒐集→ 2. 資料的處理→ 3. 資料的儲存→ 4. 資料分析



▲圖 1-7 處理大數據的流程⁷

1. 資料的蒐集

(1) 網路爬蟲 (Web Crawler)

解析網頁內容的程式，可以自動抓取網頁資料的技術。以 Google 搜尋引擎 (Search Engine) 為例：演算法是透過機器學習、大數據分析等方式找出使用者搜尋關鍵字背後的意圖，然後針對不同使用者提供個人化的搜尋結果。

(2) 感測器 (Sensor) 蒐集數據

感測器是電子元件，用以將收集到的資料轉換成設備能處理的訊號。舉例：運動員身上配戴或是運動器材上的傳感器技術，可以了解該運動員的運動姿勢及角度，透過大數據資料分析提出各種改善建議，讓運動員有更好的運動表現。

7 資料來源：陳宗和王雅惠 (2022) 《108 版高中資訊科技 (全)》。臺灣：全華圖書有限公司。

(3) 開放資料 (Open Data) 平臺

開放資料可以是大數據的一種，免費提供大眾使用的資料，開放資料的來源可以是個人、私人企業、政府機關等任何單位，這些資料不受著作權控制，大家都可以使用。舉例：蒐集政府公開的資料，轉換成易懂的圖表介紹。



▲圖 1-8 Google 搜尋引擎運作方式⁸



▲圖 1-9 世足賽科技資訊融合⁹



▲圖 1-10 用數據看臺灣¹⁰



▲圖 1-11 政府資料開放平臺¹¹

8 參考資料：Google 演算法運作。網址：<https://reurl.cc/Ydy92x>

9 參考資料：2022 世足賽科技資訊融合 - 運動科學網。網址：<https://reurl.cc/VRkj2N>

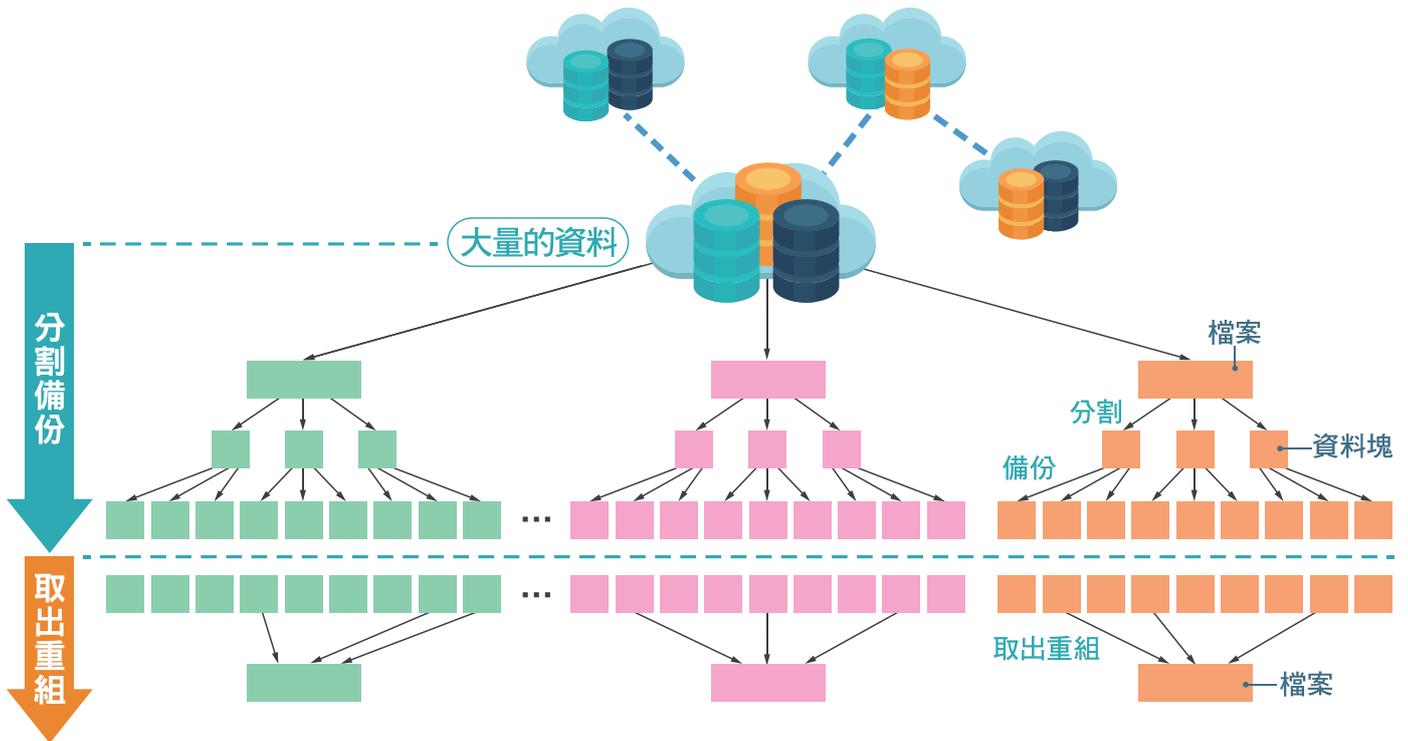
10 參考資料：用數據看臺灣，<https://www.taiwanstat.com/>

11 參考資料：政府資料開放平臺，<https://data.gov.tw/>

2. 資料的儲存

(1) MapReduce 軟體系統

由 Google 公司開發，是一種「分割重組」巨量資料的系統，主要用來將巨量資料分散儲存至多臺電腦，將分散的資料經由資料分析處理後再重組成一份完整的資料。

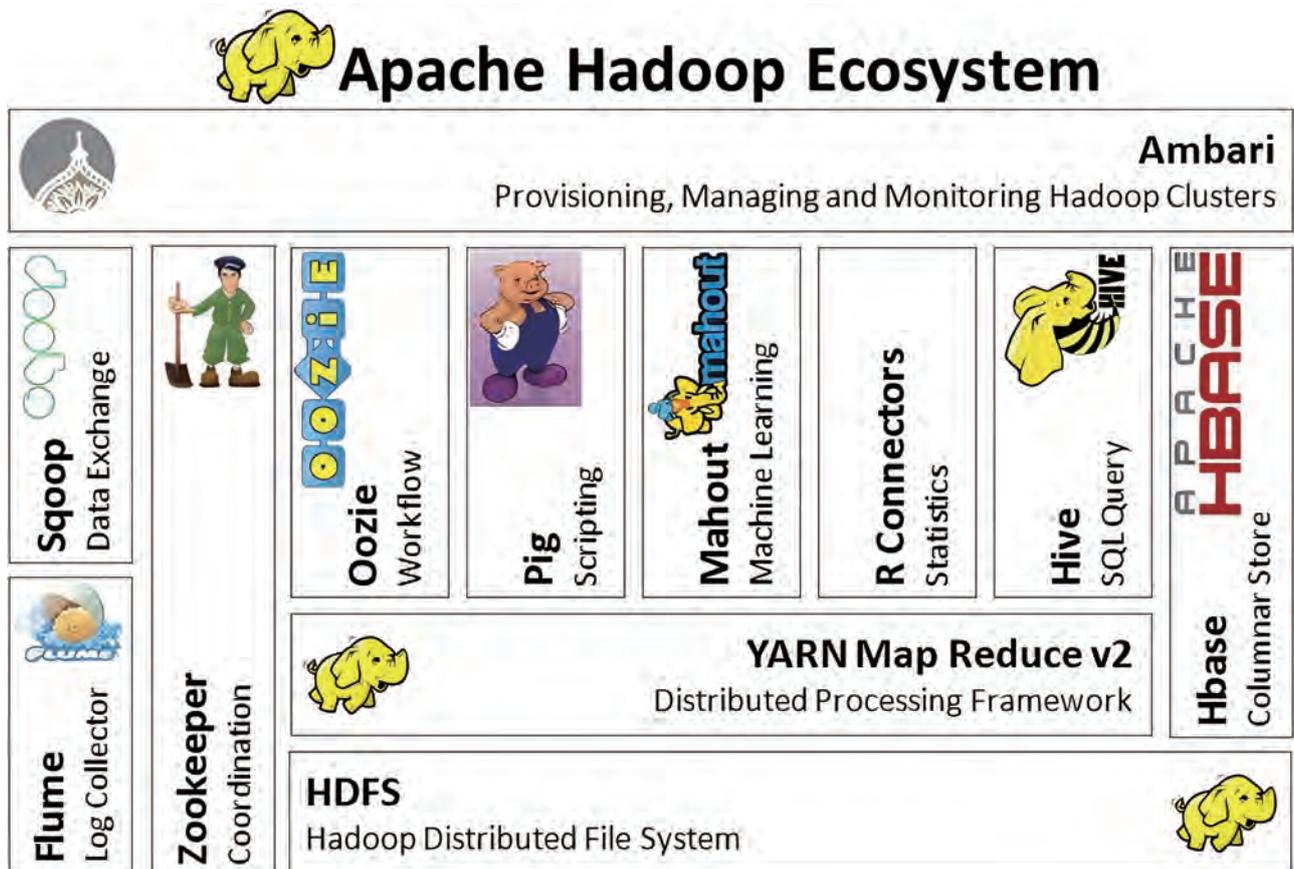


▲圖 1-12 資料分割備份與重組¹²

12 參考資料：施威銘、傅玗華、旗立資訊研究室。普高 - 資訊科技（全）。臺灣：旗立資訊。

(2) Hadoop 軟體系統 ¹³

由 Apache 軟體基金會開發，開放原始碼且免費下載，除了具有分割和重組資料的功能外，Hadoop 還支援分析和統計大量資料的能力，因此衍生出了許多的服務和功能。



▲圖 1-13 Hadoop 安裝教學與介紹 ¹⁴

3. 資料的處理

一般而言，要將資料轉化為資訊，必須將各種不同來源的資料進行整合，再進行清理、轉換等前處理工作，最後再以統計、歸納等方法進行資料分析與解釋。若無確認資料是否正確，可能導致最後分析結果失真。

(1) 資料整合

不同來源的資料分散於不同的檔案，而且每個檔案的欄位可能不一致，因此必須調整後匯集在同一檔案中，以利後續工作。

¹³ 參考資料：Apache Hadoop，<https://hadoop.apache.org/>

¹⁴ 參考資料：Hadoop 3.2.0 安裝教學與介紹，<https://reurl.cc/4XAdn2>

(2) 資料清理

資料整合後，接著要預先處理資料中常見的問題，以確保後續統計、分析的正確性。常見問題包含 (a) 資料是否完整：資料漏失，缺少分析時需要的屬性資料。（漏填了性別、生日、年齡等資料。）(b) 資料是否有雜訊：資料錯誤造成的不合理或不正常。（身分證字號多 1 碼、電話多 1 碼、年齡 200 歲等。）常用的清理方法：①直接刪除法：當資料量很龐大，而遺缺資料只佔少數，則可直接刪除此筆資料。②填補法：以人工方式加以確認、修正資料，或以平均值、中位數等數值取代遺缺資料。若遺缺資料過多，則會浪費許多時間和人力。

(3) 資料轉換

將資料轉換成相同單位、格式，避免分析時產生錯誤。

4. 資料的分析

常用的兩種資料分析技術

(1) 統計學方法

利用統計學的方法，對資料進行描述性和預測性分析，常見的分析工具有迴歸分析、相關分析等。常見的經典案例：超市的尿布和啤酒。美國著名的 Walmart 超市曾經用統計分析，發現一個現象：星期五大部分買尿布的客人同時也會買啤酒。其中可能的原因是美國年輕爸爸們，星期五晚上常會奉太太之命去買尿布，到了超市後，一想到晚上要看球賽，又順便拿手啤酒回家配。因此 Walmart 超市將啤酒與尿布擺放在相同的區域，讓顧客可以快速同時找到這兩件商品，進而同時提升尿布和啤酒的銷售量。



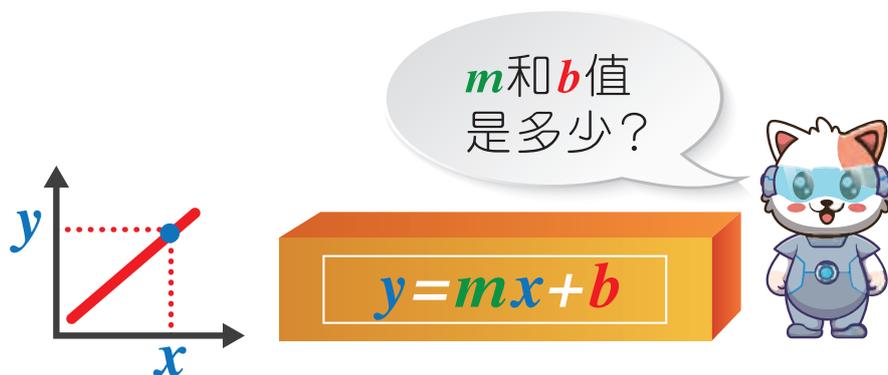
▲圖 1-14 資料探勘

(2) 機器學習 (Machine Learning)

是目前較常使用的方法，指電腦可以運用演算法自我學習，改進演算法效能可以建立最佳模型 (Model)。假設「模型」是直線方程式 $y = mx + b$ ，知道 m 、 b 輸入 x 就可以得到 y ，其中 m 和 b 需要不斷嘗試修改才能得到最佳值，修改過程就是學習。



▲圖 1-15 機器學習建立模型



▲圖 1-16 找出趨勢直線

常見的资料分析演算法

1. 線性迴歸分析 (Linear Regression)：趨勢預測¹⁵

統計學工具中分析數據的方法，常被用來做趨勢預測。趨勢預測是指預測一個隨著時間經過而改變的變數。做法是在平面上畫出一條直線，這條直線可以代表點 (資料) 的變化趨勢。舉例：如何在散佈圖上，找出一條直線，使得這條直線讓我們可以根據當天氣溫來預測冰淇淋銷售數量。說明如下：

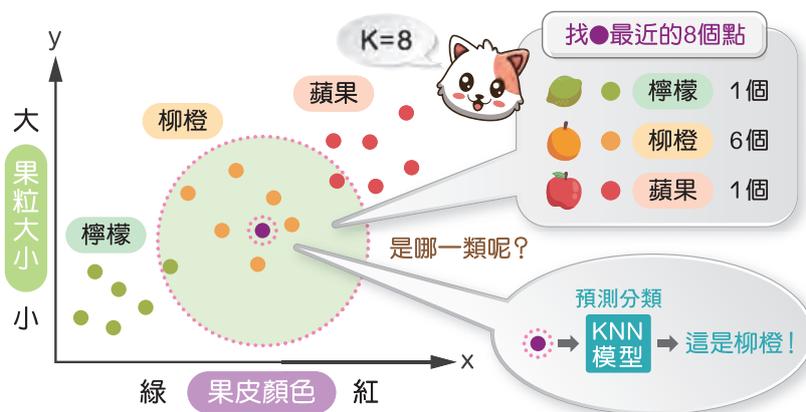
15 參考資料：陳宗和王雅惠 (2022)。108 版高中資訊科技 (全)。臺灣：全華圖書有限公司。

▼表 1-2 散佈圖與趨勢線說明

圖片	說明
	<p>先觀察數據資料之間的關係¹⁶。例如：左圖可以先觀察每天最高氣溫和冰淇淋銷售量的關係。</p>
	<p>找出儘可能靠近數據資料的一條線，讓數據資料的點和這一條線的距離總和最小，也就是讓「誤差」最小¹⁷。例如：左圖必須找出一條通過所有點的直線，因為並沒有一條線可以通過所有點，故僅能找出一條最接近所有點的直線，再透過該直線來預測冰淇淋銷售量。</p>

2. K 最近鄰居法 (K Nearest Neighbor, 縮寫 KNN)

此分類 (Classification) 的方法是先找出距離最近的 K 個鄰居，然後再去觀察鄰居所屬類別，最後再決定自己的類別。



▲圖 1-17 參考資料¹⁸

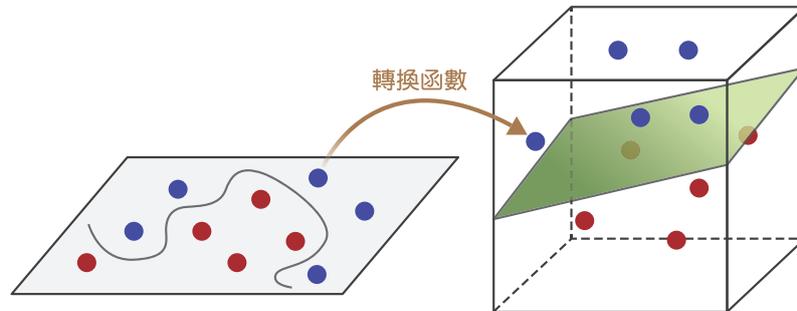
16 參考資料：108 版高中資訊科技（全），全華。

17 參考資料，108 版高中資訊科技（全），全華。

18 參考資料，108 版高中資訊科技（全），全華。

3. 支援向量機 (Support Vector Machine, 縮寫 SVM)

SVM 是常用來進行對事物進行分類的技巧。例如我們要對於下圖中左邊平面找到一條直線將紅色及藍色的點隔開是不可能的。但是 SVM 透過一個轉換的函數，將平面的資料轉換到一個立體的空間，就可以尋找一個將紅色及藍色的點分割的平面。



▲圖 1-18 轉換函數

4. 類神經網路 (Neural Network, 縮寫 NN)

即使現在的資訊科技進步神速，但是仍然無法完全取代人類大腦的能力。因此類神經網路嘗試著模仿人類的腦神經系統，來達到與人類大腦相近的功能。類神經網路包含許多神經元，神經元之間平行且用分散的方式來進行運算，透過從使用者輸入的資料中學習並持續改善，並進而解決複雜的問題。



▲圖 1-19 類神經網路示意圖

二·大數據的應用

透過資料分析找出資料之間的關聯性、趨勢預測、分類、分群等，大數據的應用範圍廣泛。日常生活中利用穿戴的裝備，像是智能手錶或智能手環等所生成的數據，傳回大數據雲端（大數據資料庫），進行自身身體狀況的追蹤及分析，可以知道自身的身體狀況（熱量消耗、睡眠品質、壓力高低等）做出應對的策略。或是像智慧車聯網透過共享機車安裝 PM2.5 感測器搭配 GPS，可以即時蒐集偵測城市各街道的空氣品質並回傳雲端，蒐集空氣大數據。



▲圖 1-20 移動式空品預測平臺¹⁹



▲圖 1-21 感測器²⁰

我們也可以從自己的生活經驗中發現問題，然後利用大數據分析技術來思考解決策略。舉例：放學回家想借用校門口的 Ubike 單車，卻無半臺 Ubike 可借用！如果可以知道各站點車輛的歷史數據，加上附近學校學生人數、捷運站數量、公車站牌數量等等相關資料，或許可以依據這些數據預測出 Ubike 單車站點的供給與需求量，然後建議相關單位提供更好服務，讓我們生活更便利。

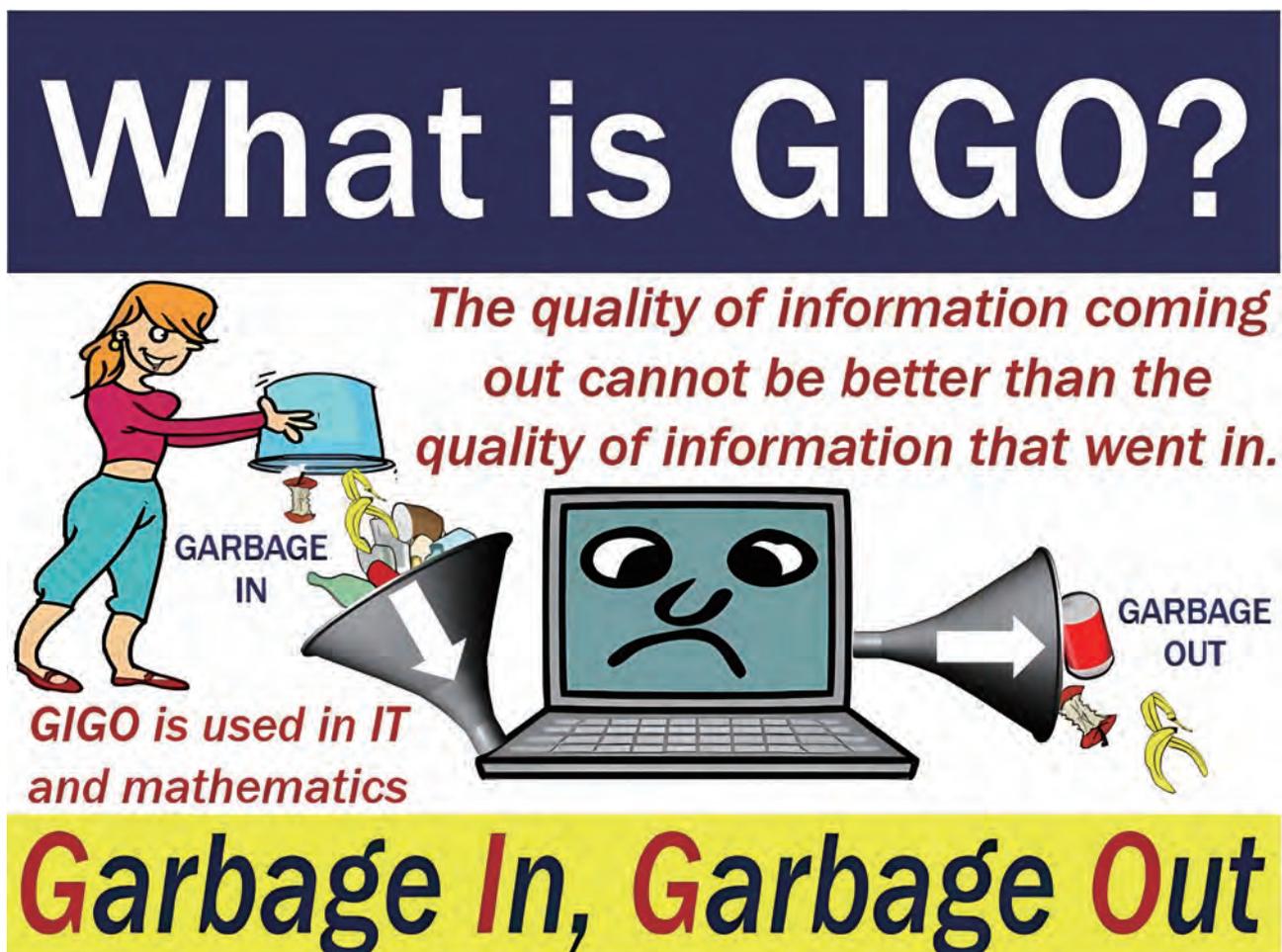
19 參考資料：WeMo Scooter 攜手臺灣大哥大打造移動式空品預測平臺。網址：<https://reurl.cc/kqD1Rx>

20 參考資料，智慧城市感測器。網址：<https://www.kocpc.com.tw/archives/298121>

(三) GIGO

預測分析首要收集正確數據開始。數據多並不一定意味著數據好。資料收集是資料處理的前步驟，如果資料準備的細節做好，可以讓資料處理更加的精準且有效率，因此處理資料前就要做好規劃，避免垃圾進垃圾出（Garbage In, Garbage Out，縮寫 GIGO）的情況發生。

會導致垃圾進垃圾出（GIGO）常見錯誤包含以下情形：①收集超出必要的數據、②構建比必要更複雜的模型、③在分析中選擇錯誤的預測變量或特徵、④使用缺乏足夠質量和相關性的數據。²¹



▲圖 1-22 垃圾進垃圾出（GIGO）示意圖²²

21 參考資料：使用關聯資料來預測分析：避免 (Garbage In, Garbage Out, GIGO)，網址：<https://reurl.cc/4XZMnV>

22 What is GIGO (garbage in, garbage out)? - Market Business News. ◦
網址：<https://marketbusinessnews.com/financial-glossary/gigo-garbage-in-garbage-out/>



S

➤ 軟體：雲端平臺，我的程式筆記本 Google Colab

➤ 連結：<https://colab.research.google.com/>

➤ 主題：輸入身高體重計算 BMI 值。

➤ 說明：

1. 程式語言有很多種（C、C++、Python 等），Python 是最近非常受歡迎的程式語言，語法簡潔、容易學習。
2. 在 Google Colab 雲端平臺透過瀏覽器編寫 Python 程式，手機或個人電腦都不需要下載安裝軟體，程式碼會自動存到 Google 雲端硬碟，非常方便。

1 開啟瀏覽器→登入 Google 帳號→搜尋「Google Colab」



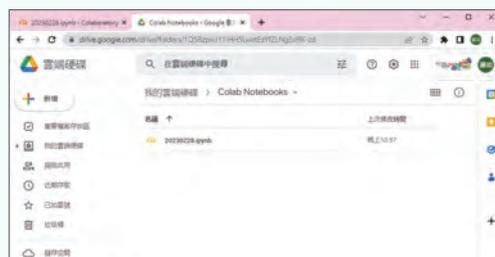
2 點選→「新增筆記本」



3 在筆記本編寫程式



4 寫好的程式會存在雲端硬碟→資料夾「Colab Notebooks」（圖 1-26）



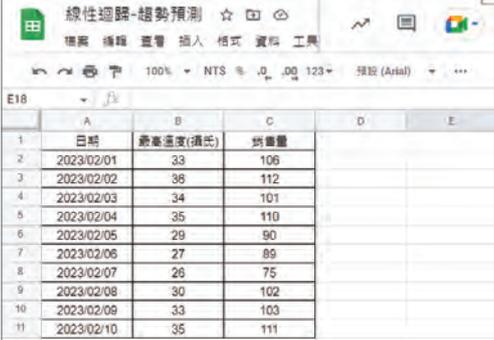


體驗課程

J

- 主題：預測商品銷售量
- 軟體：Google Sheets 試算表
- 說明：線性迴歸分析是在座標平面上畫出一條直線，這條線可以代表點（資料）的趨勢變化。我們可以用試算表軟體求出這一條趨勢線。
- 任務：店長每天記錄溫度與冰淇淋銷量的資料。
 1. 溫度越高，銷售量是否增加？
 2. 假設氣象局預測明天溫度是 31 度，請預測冰淇淋銷售量。
- 實作：

1 先輸入少量資料。



	A	B	C	D	E
1	日期	最高溫度(攝氏)	銷售量		
2	2023/02/01	33	106		
3	2023/02/02	36	112		
4	2023/02/03	34	101		
5	2023/02/04	35	110		
6	2023/02/05	29	90		
7	2023/02/06	27	89		
8	2023/02/07	26	75		
9	2023/02/08	30	102		
10	2023/02/09	33	103		
11	2023/02/10	35	111		

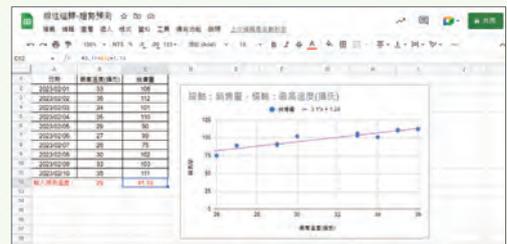
2 選取資料範圍 [B1:C11] → 插入圖表 → 散佈圖。



3 按自訂→按圖表標籤和軸標題→勾選趨勢線。以產生點（資料）的趨勢線和其方程式。



4 輸入預測分析資料→使用公式預測。



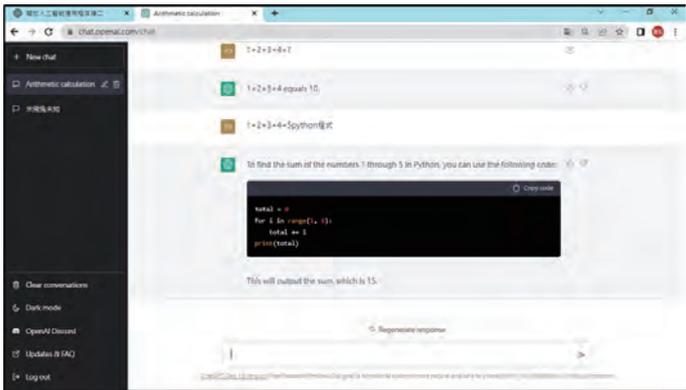


J S

主題：聊天機器人 (Chatbot)

傳統商店為了與顧客互動，會聘僱專人在電話或是通訊平臺等待回答消費者疑問，這樣的方式耗費人力，也無法處理大量問題。聊天機器人 (Chatbot) 是一種自動即時問答系統，可以全天提供即時服務，是一種可以增加與訪客互動的強大工具。

▼表 1-3 聊天機器人 (Chatbot) 介紹

序號	項目	說明
1	<p>ChatGPT</p> 	<p>OpenAI 開發的 ChatGPT 聊天機器人 (Chatbot)，使用者輸入問題就會提供答案互動。步驟：①第一次登入需要註冊。②註冊後需要驗證信箱。③預設是英文介面，可以按滑鼠右鍵選繁體中文。</p> 
2	<p>CHATISFY</p> 	<p>臉書 (FB) 的聊天機器人，全中文介面，整合網路開店功能。註冊後登入管理後臺，可以建立貼文回覆、關鍵字、上架商品等功能。</p> 



體驗課程

J S

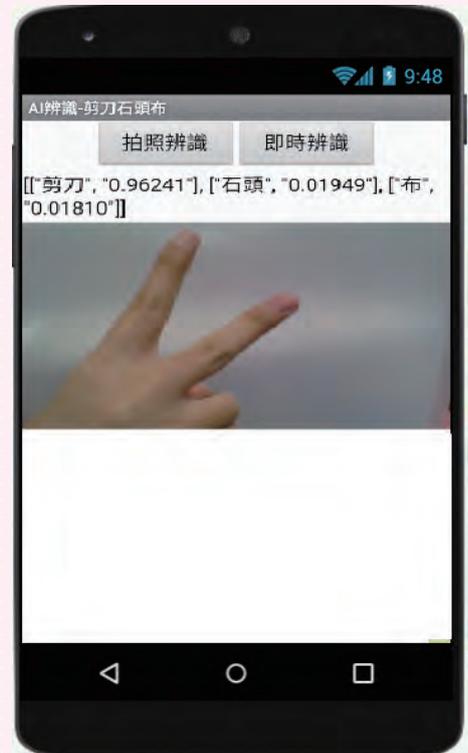
 主題：圖像辨識 app²³

 軟硬體：App Inventor + PIC (Personal Image Classifier) + BlueStacks 模擬器 + webcam

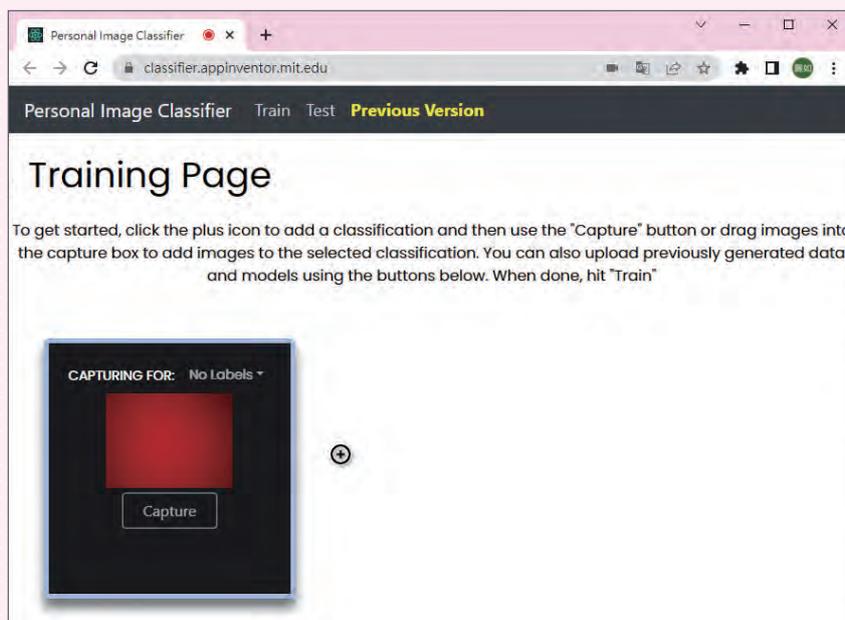
 說明：PIC 是由 MIT App Inventor 所開發，提供人工智慧的訓練程式，使用者可在雲端訓練辨識模型，並下載模型、應用在 App Inventor 中設計圖像辨識的 app。

 功能：

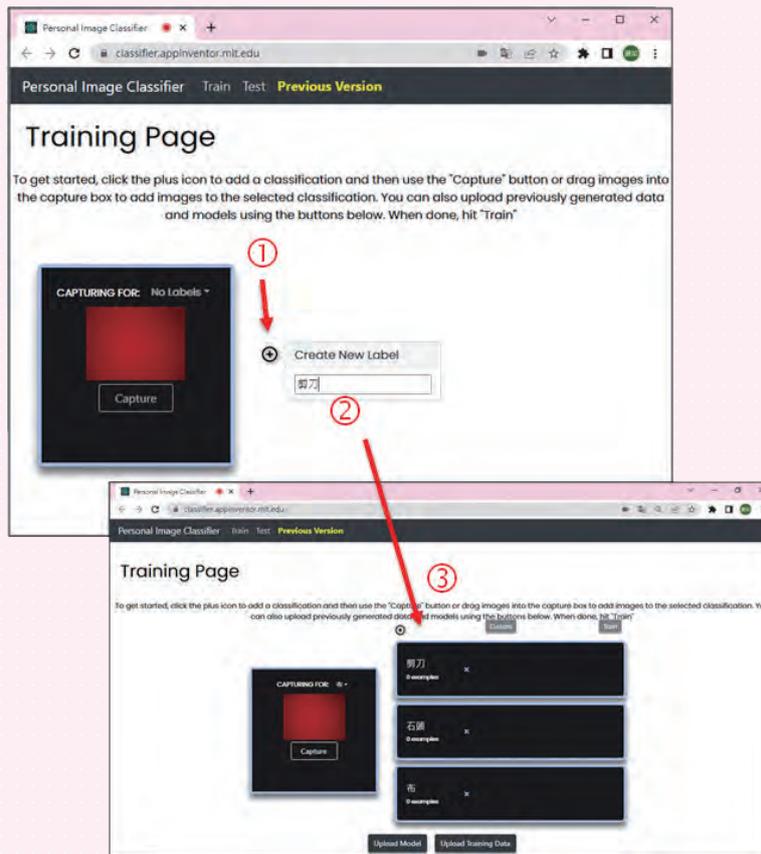
1. 點擊【拍照辨識】鈕：開啟相機，拍攝照片後呈現於 app 中，並進行辨識。
2. 點擊【即時辨識】鈕：首次點擊時，開啟相機鏡頭，再次點擊時便會進行即時辨識。
3. 【辨識結果】標籤：呈現辨識結果，數字代表可能性。



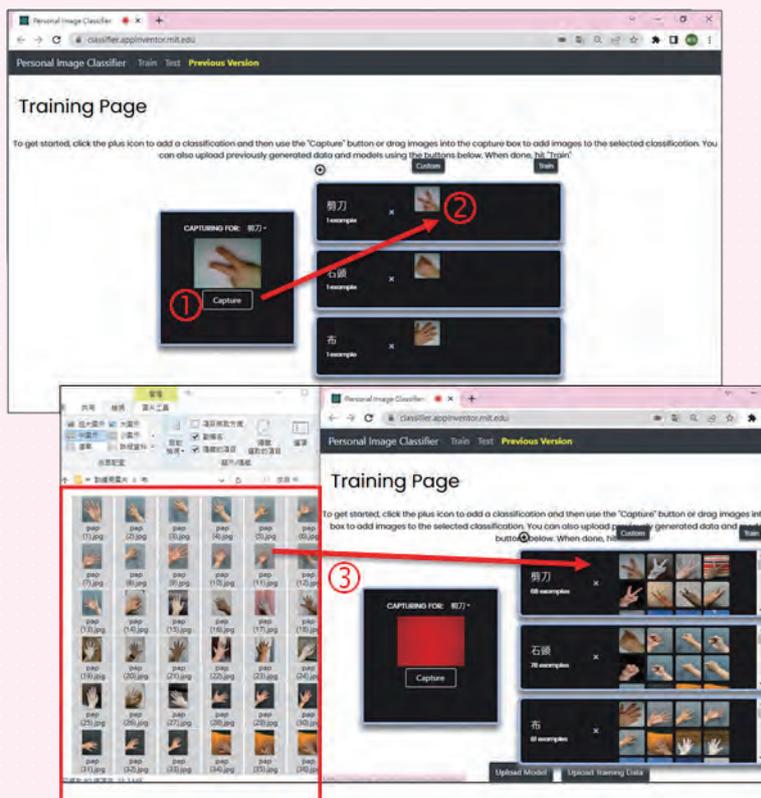
1 登入 PIC 首頁 (<https://classifier.appinventor.mit.edu>)



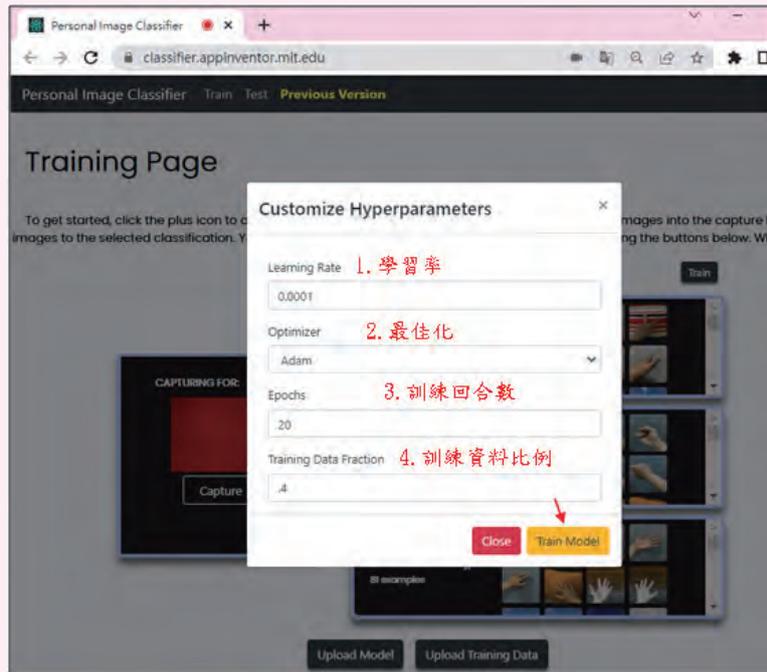
2 訓練辨識模型



▲圖 1-23 設定分類



▲圖 1-24 收集資料



▲圖 1-25 訓練、測試、下載模型

3 點選【匯入擴充套件】→加入 PIC 元件



▲圖 1-26 加入擴充元件 PIC

4 圖像辨識 app 程式設計



▲圖 1-27 圖形辨識 app 程式碼

第二篇



AI 原理 — 影像辨識及語音辨識

壹 · 影像辨識



單元 1

人工智慧與影像辨識

作者：汪殿杰—大同高中 | 蔡志聰—江翠國中

一 · 人工智慧與影像辨識 JS

人工智慧是指電腦透過程式、算法和數據，使電腦具有學習、判斷和解決問題的能力。常見應用包括自然語言處理、計算機視覺、機器學習等。人工智慧在視覺與影像辨識 (Image Recognition) 的應用，是電腦程式透過特徵提取、分類和判斷，自動辨識影像中的物體或人物。這類技術再結合自動化控制的硬體設備，可以在日常生活中有許多應用，產生便利的自動化服務，是一種非常實用的 AI 技術，同時也為許多領域帶來了巨大的潛力和價值。





單元 2

生活中影像辨識的應用

作者：汪殿杰—大同高中 | 蔡志聰—江翠國中

人工智慧與影像辨識結合起來，可以實現許多有用的功能。例如，通過人工智慧與影像辨識技術，可以實現自動駕駛車輛、自動停車系統、物品辨識系統等。

以下是幾個生活常見的案例：

1. 自動停車系統：利用影像辨識技術，可以自動偵測停車位是否有車輛，並指示進入停車場的駕駛如何停車。
2. 安全監控：影像辨識技術可以用來自動偵測場所內的異常行為，如有人越界或攜帶危險物品，並通知相關人員做出回應。
3. 生物辨識：影像辨識技術可以用來辨識人臉，這在某些場所或設施，會用來作為身份驗證的方法。
4. 圖像搜尋：影像辨識技術可以用來搜尋圖像中的特定物體或圖案，例如在網路上搜尋同樣款式的衣服或物品，協助操作者找尋更多相關訊息。
5. 自動驗證：影像辨識技術可以用來自動驗證文件的真偽，例如身份證、駕照或信用卡的辨識。
6. 辨識圖書：影像辨識技術可以用來幫助盲人辨識圖書中的文字或圖像，讓盲人可以閱讀書籍或報章雜誌。

下面我們舉例兩個在生活當中，目前已經非常普及的人工智慧影像辨識應用：

一· 車牌辨識的具體操作程序

車牌辨識是一種計算機視覺技術，用於自動辨識車輛的車牌號碼。這種技術通常用於停車場、高速公路收費站和其他車輛管理系統。車牌辨識的具體操作程序可以分為以下幾個步驟：

1. 影像擷取：首先，使用攝像頭或其他影像採集設備擷取車輛的影像。
2. 影像預處理：然後，對影像進行預處理，通常包括去雜訊、二值化、形態學處理等步驟。

3. 車牌定位：接下來，使用車牌定位算法尋找車牌所在的位置。這一步驟通常使用尺度不變特徵及其他方法來確定車牌的位置。
4. 字符分割：然後，將車牌分割成若干個單獨的字符。這一步驟可以使用分水嶺算法或形態學處理等方法。



5. 字符辨識：最後，使用字符辨識算法對每個字符進行辨識，得出車牌號碼。這一步驟可以使用支持向量機、神經網絡等機器學習算法來實現。

二．人臉影像辨識的具體操作程序 JS

人臉影像辨識的人工智慧應用，電腦會先進行人臉檢測，找出影像中的人臉區域。然後將人臉區域轉換為數字特徵，並與一個已知人臉庫中的特徵進行比對。如果找到了匹配的特徵，則會確定該人臉的身份。在進行人臉影像辨識時，需要完成以下一些步驟：

1. 收集人臉圖像數據：首先需要收集大量已標記的人臉圖像數據。這些數據用於訓練人臉影像辨識模型，以便讓它學習人臉的特徵。
2. 建立人臉特徵模型：接下來，可以使用機器學習方法建立一個人臉特徵模型。通過對已標記的人臉圖像進行訓練，可以讓模型學習人臉的特徵，並建立一個能夠識別人臉的模型。
3. 提取人臉特徵：當系統處理新的人臉圖像時，需要對圖像進行特徵提取。特徵提取是指將圖像轉換為一組數值，這些數值能夠代表圖像中人臉的特徵。
4. 預測人臉類別：最後，可以利用人臉特徵模型對圖像進行預測。模型會根據圖像中提取的人臉特徵，預測該圖像中人臉的類別。這類技術可以幫助電腦自動辨識照片中的人，或者用於其他應用。





單元 3

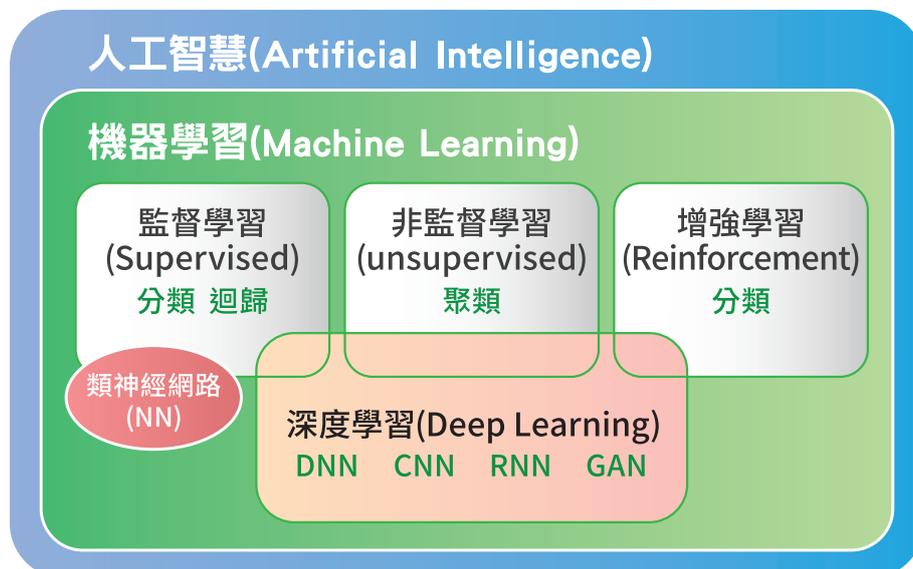
影像辨識的原理 J S

作者：汪殿杰—大同高中 | 蔡志聰—江翠國中

在進一步說明影像辨識的人工智慧應用前，我們先分別說明在視覺應用的兩種類型，當我們觀看的畫面是靜止時，這樣的單個畫面稱為圖像 (image)，當一個個畫面連續出現，且更換的頻率夠快，讓人眼覺得是連續的動畫，則是視頻 (video)。這兩種類型影像資料作為辨識的素材，分別為靜態與動態的型態。這意味著，在圖像辨識中，單個圖像可以被解釋為一個狀態，而在視頻辨識中，多個連續的影像可以被解釋為一個動態的過程。在視頻辨識中需要考慮時間因素，並對連續的影像進行分析，以找出在時間軸上的變化，所以視頻仰賴電腦運算處理部分會更為繁重。

一 · 影像辨識的 AI 原理 J S

影像辨識是指使用人工智慧透過機器學習技術，識別影像中的物體或背景的過程。機器學習技術是人工智慧的一個分支，它使電腦能夠自動“學習”和做出預測或決策，這是通過訓練電腦使用大量的資料與演算法來優化實現的。機器學習可分為三大類：監督學習、非監督學習和增強學習。監督學習是通過提供有標籤的訓練資料來訓練辨識模型；這裡所謂的標籤是指預先知道的關於訓練數據的正確解決方案。這些標籤可以是分類標籤，例如給定一個圖像，標籤可以是「貓」或「狗」；也可以是迴歸標籤，例如給定一個數據的實際值。模型使用這些標籤進行學習，並且在訓練後，可以對新的未標記數據進行預測。

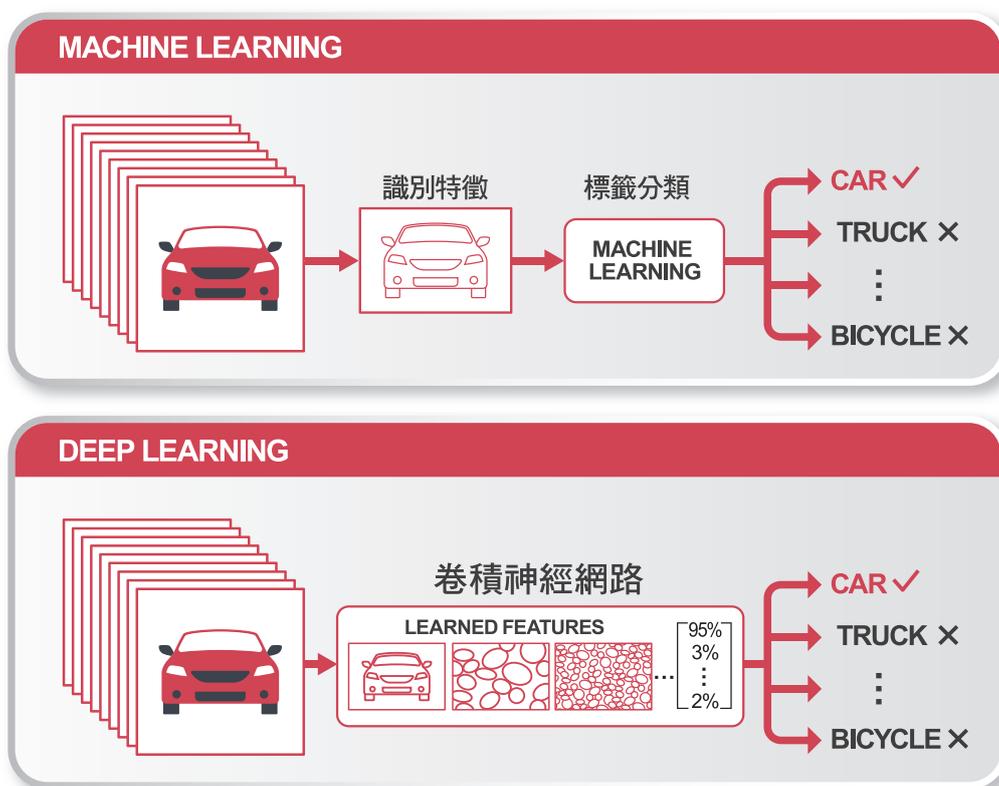


非監督學習是不需要標籤資料，只需要輸入資料。增強學習是通過讓模型在特定環境中不斷嘗試和學習，透過正向增強的方式來改善其性能。影像辨識過程是通過將圖像轉換為數據，並使用訓練過的模型來識別特徵進行的。

深度學習是目前用於影像辨識較為進階的方法，可以減少人工資料前期處理的建構程序，與提高辨識的效率。深度學習是由神經網路（NN）的方式進行，這些神經網路又會因為採用不同的函數學習方式分別命名（例如 DNN、CNN、RNN、GAN），並且使用大量的訓練數據和許多層次來學習如何識別影像中的特徵。在訓練過程中，模型會被提供大量的圖像和標籤，其中標籤指出圖像中的物體或背景。模型通過不斷地對圖像進行分類，並根據其預測結果與實際標籤的匹配情況來調整其參數，以便在未來的預測中更精確地識別圖像。當模型被訓練好之後，它就可以使用來識別新圖像中的物體或背景。

二 · 影像辨識的機器學習方法 JS

影像辨識技術通常通過機器學習來實現。機器學習是一種人工智慧技術，它可以讓電腦自動學習並改進其預測能力。在影像辨識中，通常會使用監督式學習方法，提高辨識系統建置的效率。在這種方法中，系統會被提供大量已標記的圖像，其中每個圖像都有正確的物體標籤。然後，系統會自動學習圖像中物體的特徵，並建立一個能夠識別這些物體的模型。當系統被提供新的未標記圖像時，它可以利用所學習的模型來預測圖像中物體的類別。

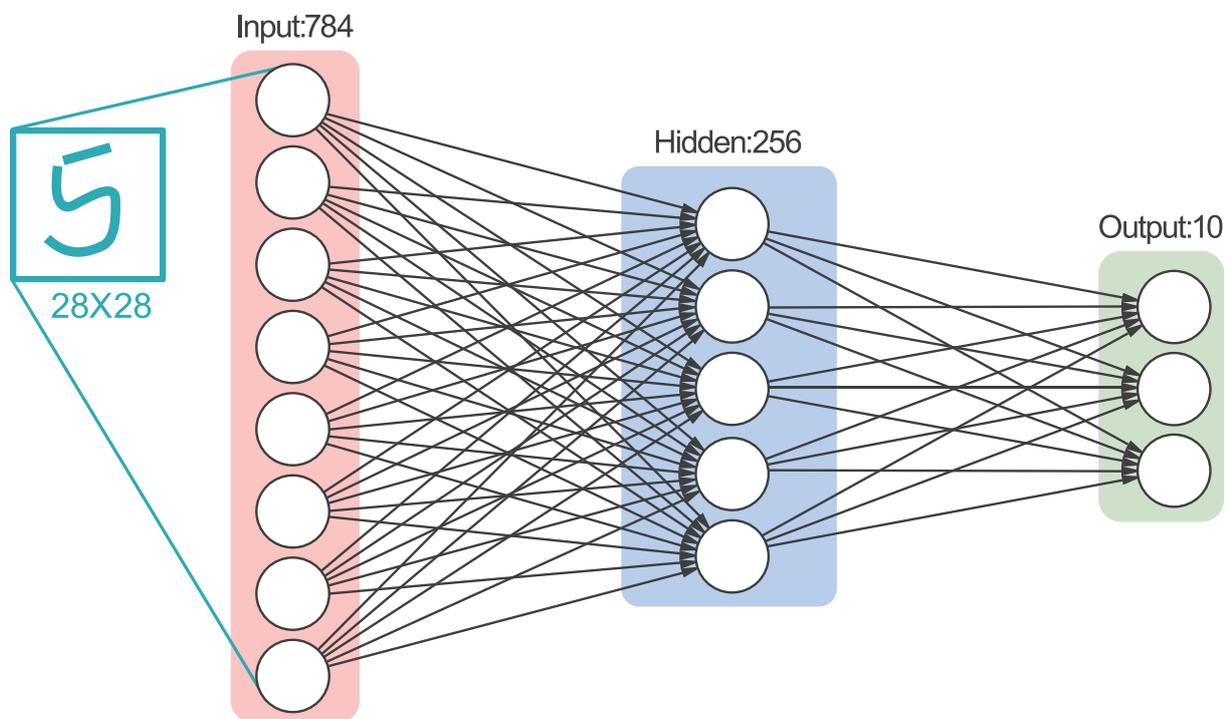


▲圖 2-1

三 · 影像辨識使用的深度學習

深度學習是另一種進階的機器學習方法，它可以讓電腦自動學習現實生活中複雜的非線性關係；例如溫度、濕度與氣候變化，有許多交互影響的變動因素，這些複雜的非線性關係，讓天氣難以精準的預測。深度學習通過神經網絡的模型來實現這一目標，神經網絡由許多輸入、隱藏和輸出層組成，每層都包含多個神經元。神經元接收輸入信息，並通過計算來產生輸出信息。神經網絡可以被訓練，以便它可以自動學習如何將輸入轉換為正確的輸出。在影像辨識中，深度學習可以用來建立一個能夠識別圖像中物體的神經網絡模型。通過提供大量已標記的圖像，可以訓練神經網絡模型，以便它能夠學習圖像中物體的特徵。一旦訓練完成，神經網絡模型就能夠接收新的未標記圖像，並預測圖像中物體的類別。

影像資料運算的過程採用神經網路（如下圖所示）的結構。它和大腦類似，由節點（腦細胞）連接和權重組成，並根據梯度下降原理運作。梯度下降是影像辨識中常用訓練深度學習的演算法，用於訓練神經網絡模型。通過對模型的預測進行評估，並使用梯度下降演算法計算損失函數的梯度，演算法可以自動更新模型的權重，使模型的預測更精確，是一種最優化演算法。梯度下降的過程通常需要大量的訓練數據，並需要通過不斷調整模型的權重，來使模型的預測結果與實際數據的差距逐漸減小，提高精準辨識準確度。



▲ 圖 2-2

四 · 影像辨識如何與 AI 結合應用 **S**

影像辨識是人工智慧應用的一個分支，主要利用電腦視覺技術來辨識圖像中的物體或人物。主要應用於自動驗證、監控、生物辨識等領域，是近年來人工智慧非常熱門的研究領域之一。

影像辨識的主要原理是利用深度學習技術建立神經網路模型，透過大量訓練圖像和標籤（即圖像中的物體或人物），來學習辨識圖像的能力。在訓練過程中，模型會不斷地對訓練資料進行預測，並比較預測結果和實際標籤的差異，然後再根據錯誤率來調整模型的參數，使得模型能夠更加準確地辨識圖像。當訓練過程結束後，模型就會具有辨識圖像的能力，可以用來辨識新的圖像。

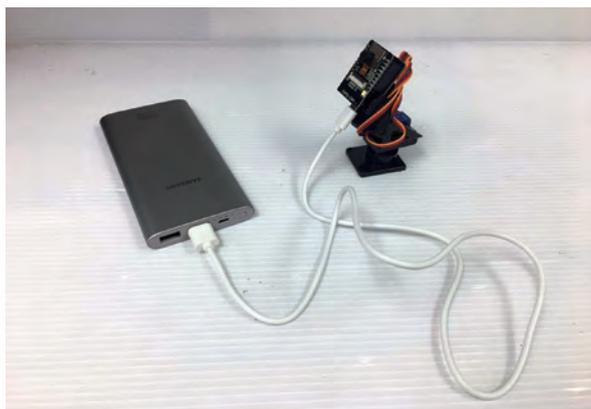
影像辨識深度學習技術通常使用卷積神經網絡 (CNN) 來訓練，因為它們對於圖像數據具有很高的效率。CNN 通過運用卷積運算對圖像進行特徵提取，並通過多層網絡結構來進行高層次的特徵表示。這些表示可以用於辨識不同的物體或人物。除了 CNN 之外，還有其他一些技術，如深度學習中的類似卷積神經網絡的技術，如基於深度學習的模型如 YOLO、Faster R-CNN、RetinaNet 和 Mask R-CNN 等，這些技術都可以用於影像辨識。此外，影像辨識技術還可以利用電腦圖像處理技術來幫助辨識，例如利用特徵檢測和描述、圖像分割、圖像質量評估等方法，來提升辨識的準確度。

五 · 影像辨識與 AI 結合應用的實作案例 **S**

這裡我們使用一個便宜簡單的方法，示範說明 AI 影像辨識的實際操作方式。本案例採用 ESP32CAM 的控制模板，和 YOLO 深度學習的人工智慧在影像辨識開發的程式模型，示範說明車輛識別的應用方法。YOLO (You Only Look Once) 是一種實時物體影像辨識的演算法，用於對圖像中的物體進行定位和分類。YOLO 通過對整張圖像進行單次掃描來實現這一目標，因此作為影像辨識使用具有較高的速度和效率，因此能夠在多種應用中進行物體檢測，例如自動駕駛、安全監控等。

STEP 1

準備 ESP32CAM 控制板，這是一種帶有小尺寸攝像頭的 WiFi 控制板，可廣泛應用於各種物聯網應用。使用前需安裝 Arduino 驅動程式，並匯入需要的函式庫 (WebServer.h WiFi.h esp32cam.h)



STEP 2

上傳程式至 ESP CAM 控制板，使其能夠將影像儲存至網頁伺服器，方便之後電腦連線讀取視訊資料進行分析。

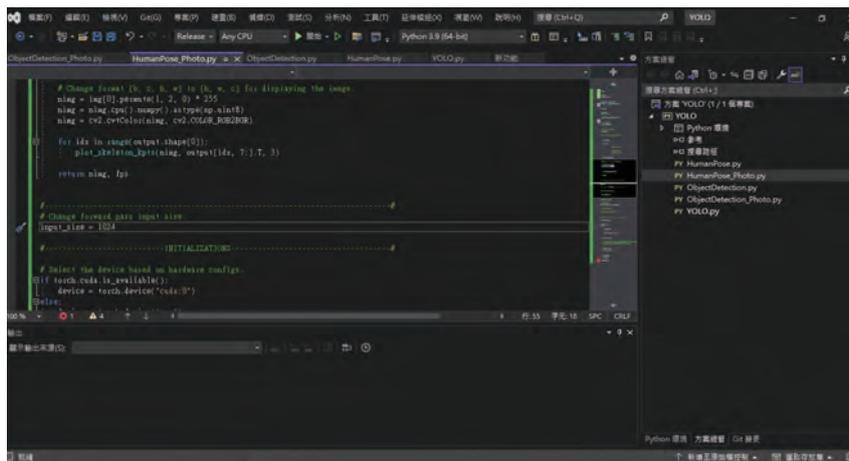
```

ESP_Cam_Streaming [Arduino 1.8.15]
ESP_Cam_Streaming
1 #include <WebServer.h>
2 #include <WiFi.h>
3 #include <esp32cam.h>
4
5 const char* WIFI_SSID = "ESP32CAM";
6 const char* WIFI_PASS = "12345678";
7
8 WebServer server(80);
9
10
11 #static auto i0Res = esp32cam::Resolution::find(320, 240);
12 #static auto i1Res = esp32cam::Resolution::find(1920, 530);
13 #static auto i2Res = esp32cam::Resolution::find(640, 480);
14 #static auto i3Res = esp32cam::Resolution::find(640, 480);
15 #static auto i4Res = esp32cam::Resolution::find(640, 480);
16
17 void setup() {
18   server.on("/", HTTP_GET, [](WebServer &server) {
19     server.send(200, "text/html", "ESP32CAM Streaming");
20   });
21   server.on("/capture", HTTP_GET, [](WebServer &server) {
22     auto frame = esp32cam::capture();
23     if (frame == nullptr) return;
24     Serial.println("CAPTURE FAIL");
25     server.send(500, "", "");
26     return;
27   });
28   Serial.println("CAPTURE OK " + frame->width() + "x" + frame->height());
29   server.setContentLength(frame->size());
30 }
31
32 void loop() {
33   server.poll();
34 }
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

STEP 3

在 Visual Studio 2022 軟體的編輯環境中，下載需要的 YOLO 程式碼及權重檔，並在 Python 環境中安裝如 PyTorch 的深度學習框架。（本範例 YOLO 程式碼下載連結 <https://github.com/WongKinYiu/yolov7>）



```

# Change format (B, G, R, A) to (B, G, R) for displaying the image.
img = img[0].permute(2, 0, 1) * 255
img = img.cpu().numpy().astype(np.uint8)
img = cv.cvtColor(img, cv.COLOR_BGR2RGB)

for idx in range(output.shape[0]):
    plot_skeleton_3pt(img, output[idx, 7, 2, 3])

return img, fps

# Change forward pass input size
input_size = 1024

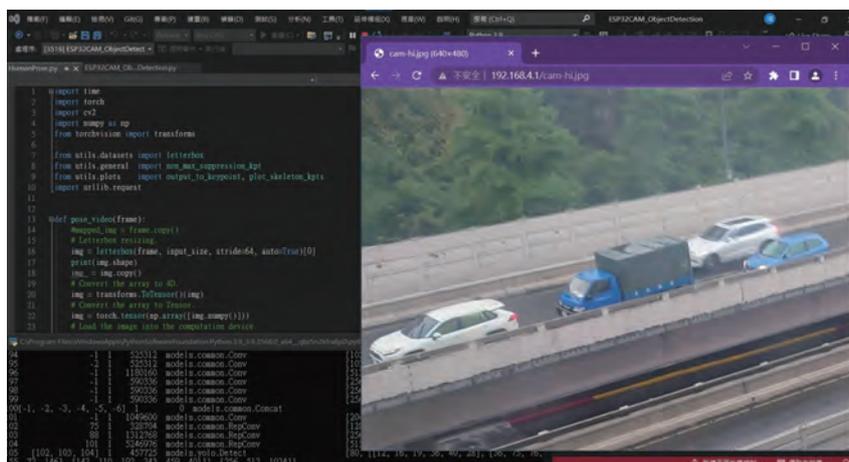
# INITIALIZE/LOAD

# Select the device based on hardware configs.
if torch.cuda.is_available():
    device = torch.device('cuda:0')
else:
    device = torch.device('cpu')

```

STEP 4

使用電腦透過 Wi-Fi 方式連接 ESP CAM，網址 <http://192.168.4.1>，透過鏡頭拍攝畫面並存取伺服器中的影像資料。



```

1 import time
2 import torch
3 import cv2
4 import numpy as np
5 from torchvision import transforms
6
7 from utils.datasets import letterbox
8 from utils.general import non_max_suppression_3pt
9 from utils.plots import plot_output_to_image, plot_skeleton_3pt
10 import urllib.request
11
12
13 def pose_video(frame):
14     imgsz_in = frame.shape[0]
15     # Letterbox resizing
16     img = letterbox(frame, input_size, stride=4, auto=True)[0]
17     print(img.shape)
18     img = img.cpu()
19     # Convert the array to 3D
20     img = transforms.ToTensor()(img)
21     # Convert the array to tensor
22     img = torch.tensor(img.unsqueeze(1))
23     # Load the image into the computation device
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

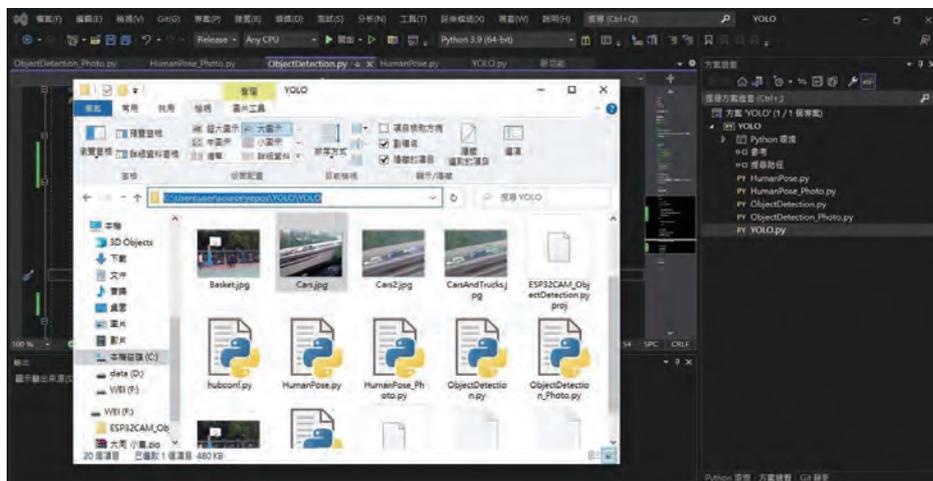
STEP 5

完成動態的車輛辨識系統的影像辨識功能，分辨攝像頭拍攝高架橋上車輛的種類。



補充說明

實作應用 AI 動態影像辨識的分析，硬體設備需要比較好的電腦，萬一學校的電腦設備不夠好，也可以用照片的方式匯入電腦分析，照片需要先匯入到這個地方，採取靜態的圖像辨識分析方法 C:\Users\user\source\repos\YOLO\YOLO(固定的位置)，再設定程式的照片尺寸設定，就可以進行分析了。





單元 4

線上圖像創作產生的 AI 應用

作者：汪殿杰—大同高中 | 蔡志聰—江翠國中

圖片或圖畫的生成，目前已經有許多通過使用人工智慧技術來完成的線上程式，特別是使用生成對抗網絡（Generative Adversarial Networks, GAN）來完成圖像的創作。生成對抗網絡是一種人工神經網絡架構，由兩個網絡組成：一個生成器和一個判別器；例如：生成器好比一個假鈔偽造者，他不斷製造假鈔。然而，判別器好比是警察，不斷從偽造者那邊拿到假鈔，並判斷真假。然後，偽造者就根據警察判斷結果的回饋，不斷改良，直到最後，假鈔變成真假難辨，讓警察束手無策。在圖像創作中，生成器的任務是根據輸入隨機噪聲生成新圖像，而判別器的任務則是辨別生成的圖像是否真實。這兩個網絡通過對抗訓練來學習，以便生成器能夠生成看起來像真實圖像的東西，而判別器則能夠更好地辨別真實圖像和生成的圖像。

下面這些網站均是使用人工智慧技術來產生圖像，並且在這個領域中有著比較高的知名度。

1. DALL-E：可以將文字描述轉換成圖像的網站，由 OpenAI 所提供。
2. DeepDream：可以將照片中的圖像轉換成各種不同風格的圖像的網站，由 Google 所提供。
3. ThisPersonDoesNotExist：可以產生看似真實的人臉照片的網站，由 Nvidia 所提供。
4. Generative Pretrained Transformer (GPT) Image Generation：可以將文字描述轉換成圖像的網站，由 OpenAI 所提供。
5. BigGAN：可以生成高解析度的圖像的網站，由 DeepMind 所提供。

一 · DALL-E AI 圖像創作方法 JS

DALL-E 是一個可以通過文本描述中生成圖像的人工智慧程序。能夠生成各種風格的圖像，從照相寫實主義圖像到繪畫和表情符號。它還可以「操縱和重新排列」圖像中的對象，例如，當被要求畫一個白蘿蔔擤鼻涕、喝拿鐵或騎獨輪車時，DALL-E 通常能在合理的位置畫出擤鼻涕用的手帕、人體或白蘿蔔的手和腳。

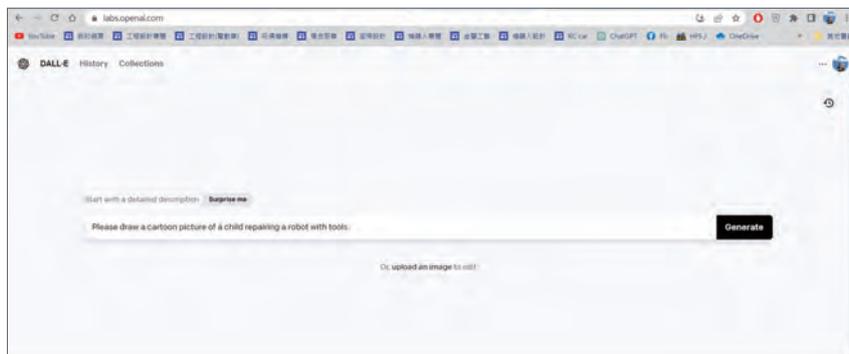
STEP 1

進入 DALL-E 的官方操作網站 (<https://openai.com/dall-e-2/>) 註冊並登入操作畫面。



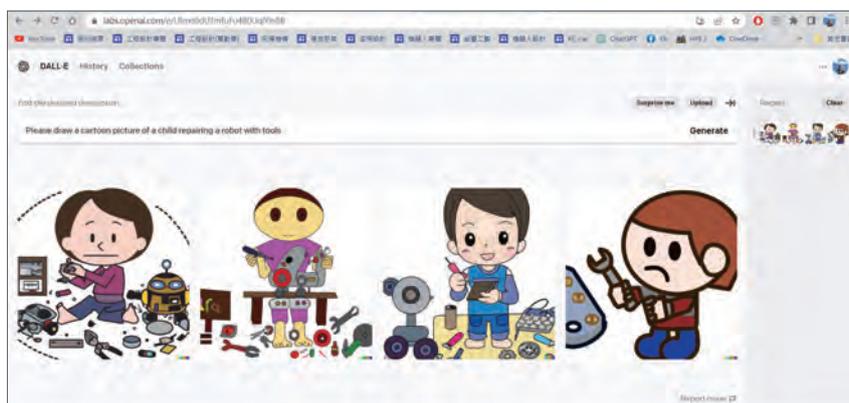
STEP 2

輸入生成圖片的文字描述，例如：請畫一張小朋友使用工具修理機器人卡通風格的圖片，指令的輸入建議使用英文，Please draw a cartoon picture of a child repairing a robot with tools。



STEP 3

產生並選擇自己喜歡的圖片下載儲存。

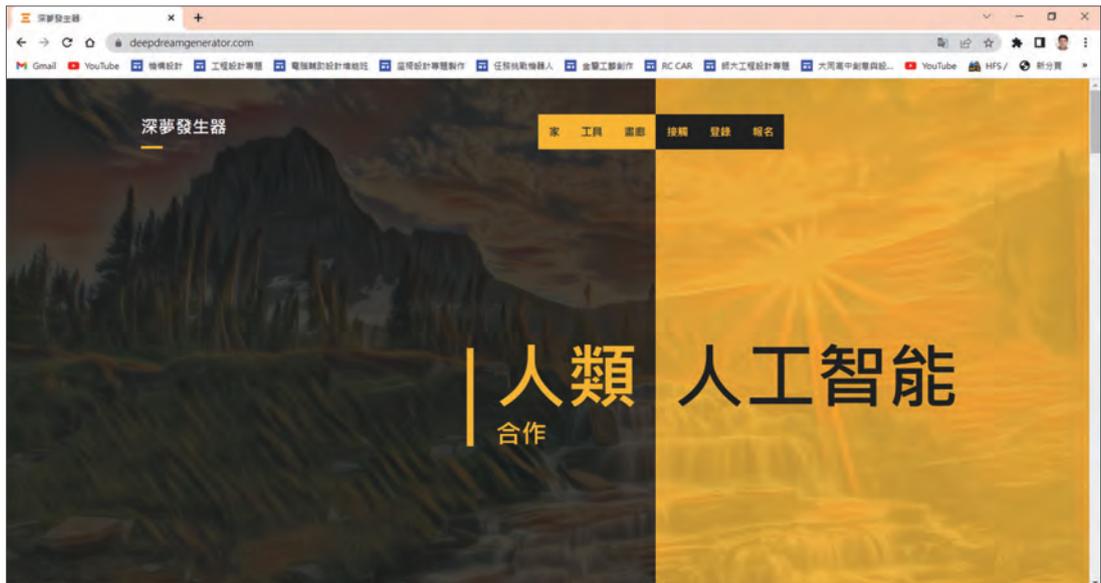


二 · DeepDream AI 圖像創作方法 JS

DeepDream 是 Google 工程師亞歷山大·莫爾溫采夫 (Alexander Mordvintsev) 創建的 AI 電腦視覺程序，它使用卷積神經網絡方式，通過演算法來查找和增強圖像中的特徵，故意過度處理圖片的特徵效果，自動產生人工智慧生成的圖片。

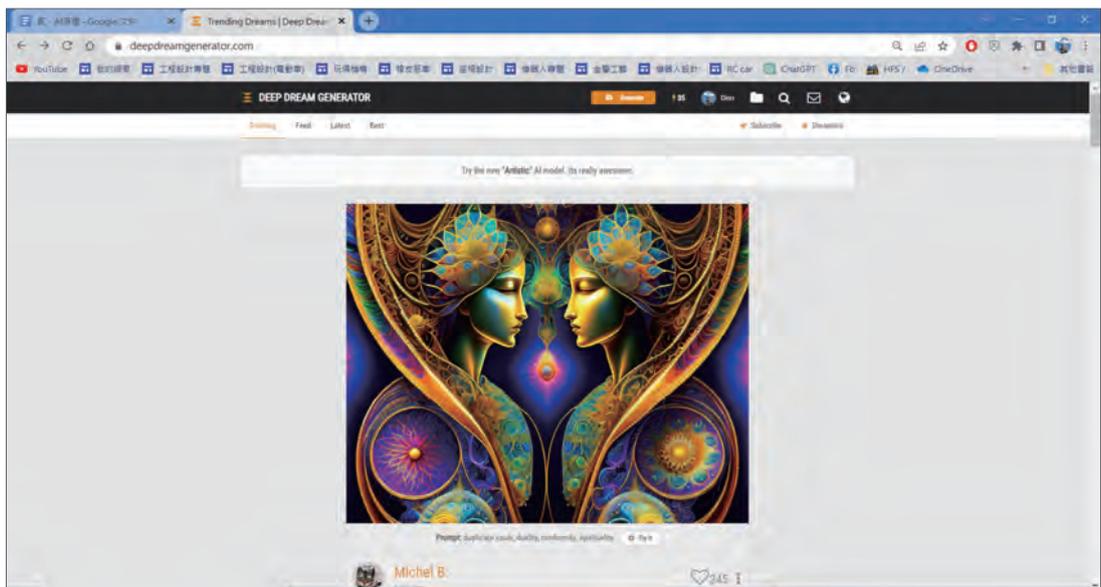
STEP 1

進入 DeepDream 的官方操作網站 (<https://deepdreamgenerator.com/>) 可直接用 google 直接登入操作畫面。



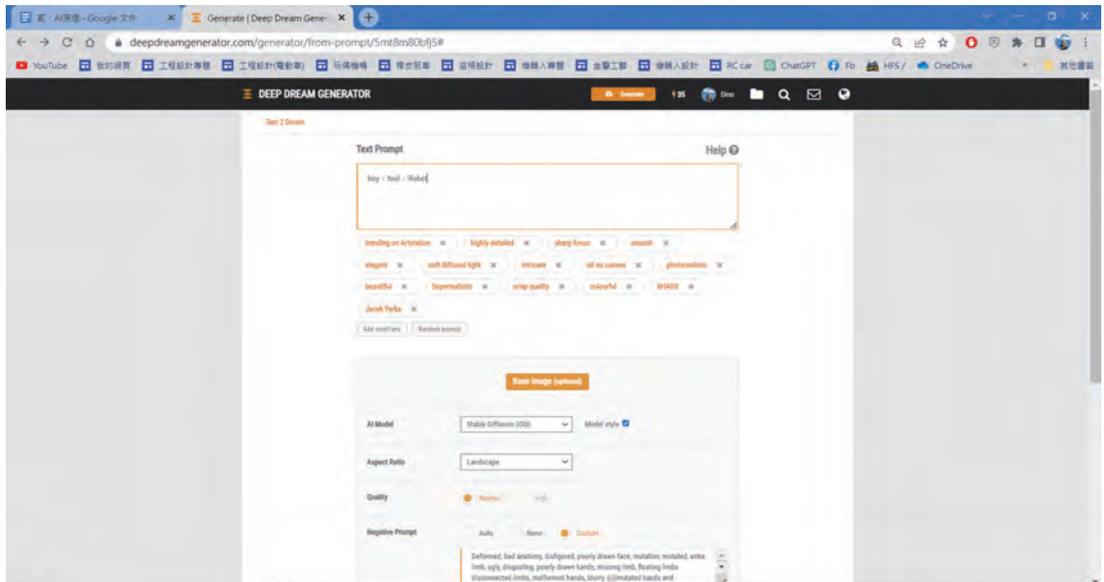
STEP 2

可直接選擇官方推薦的圖片生成風格樣式。



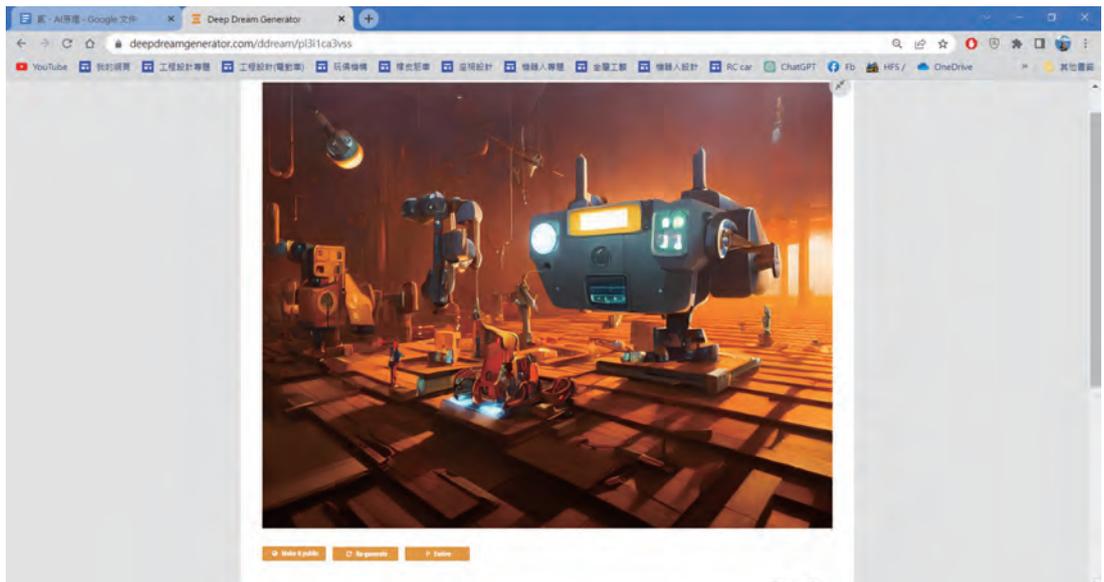
STEP 3

輸入簡單生成的主題詞彙，並使用逗號區隔不同詞彙，例如 boy, tool, Robot。



STEP 4

電腦自動生成圖片後，下載喜歡的圖片到電腦。





單元 5

影像辨識的挑戰 JS

作者：汪殿杰—大同高中 | 蔡志聰—江翠國中

目前影像辨識應用 AI 的技術還在持續努力發展中，不論在硬體設備的效率與人工智慧的運算處理，都還是存在許多的困難，需要更有創新的研究突破去解決。例如：

1. 處理遮蔽：遮蔽物體的存在會使得影像辨識難以辨識目標物。
2. 處理變形：目標物的變形會使得影像辨識難以辨識。
3. 處理光線變化：光線變化會導致影像顏色、對比度和亮度的變化，這會使得影像辨識難以辨識。
4. 處理小物體：較小的物體往往比較難以辨識，因為它們在影像中所佔的面積較小。
5. 處理多樣化的圖像：影像辨識技術需要處理許多不同類型的圖像，包括不同的顏色、尺寸、形狀和背景。
6. 處理動態圖像：辨識動態圖像會比辨識靜態圖像困難得多，因為它們中的目標物會持續變化，難以快速聚焦在辨識的目標。
7. 處理高畫質圖像：高畫質圖像中有許多細節需要辨識，這會使得影像辨識變得更困難。

雖然如此，可以預期未來的生活中，AI 影像辨識的功能服務會更多也會更好，改善生活的便利，創造更多科技服務的可能性。同時，我們也需要警惕人工智慧和影像辨識技術帶來的倫理和法律問題，以確保這些技術能夠為人類帶來真正的利益。同學們如果有興趣，可以參考我們提供的 AI 影像辨識體驗活動，和一系列的課程模組，實際動手做學習 AI 影像辨識的相關應用。

貳 · 語音辨識



單元 1

語音辨識的原理——認識聲音訊號與處理

作者：汪殿杰——大同高中 | 蔡志聰——江翠國中

前言：語音辨識是什麼？

語音辨識 (speech recognition) 是一種讓電腦能自動將人類的語音內容轉換為相對應的文字的技術，通常也稱為自動語音辨識 (Automatic Speech Recognition, ASR) 或是語音轉文字識別 (Speech To Text, STT)。在手機上的聽寫功能或是使用電腦文件的語音輸入，都是常見的語音識別應用。

近年隨著電腦運算功能的提升以及人工智慧、自然語言處理技術 (NLP) 的應用，語音辨識除了能將語音轉換為文字外，更可使電腦透過分析內容與回應來和我們進行各種人機交流與互動。例如你我身邊的智慧型手機語音助理、語音翻譯及智慧音箱 (如 Siri、Amazon Alexa、Google Home) 等，都已經成為許多人日常不可或缺的生活工具。

《嘿！Siri 背後的技術》——Siri 怎麼知道你在呼叫它呢？

在 iPhone 裡裝有一個語音辨識裝置，第一次使用前，Siri 會要你先講幾次「嘿！Siri」並記住你的聲音波形。接下來只要你一說出「嘿！Siri」，它就會以約萬分之一秒的速度把這個聲音轉換為一個聲波，並計算出這個聲音是不是你本人的。如果計算後的符合度分數夠高，Siri 就啟動開始和你互動。要是分數不夠，它就不理你了。



想一想

有沒有辦法讓二個、或三個人同時可以呼叫同一個 Siri 呢？

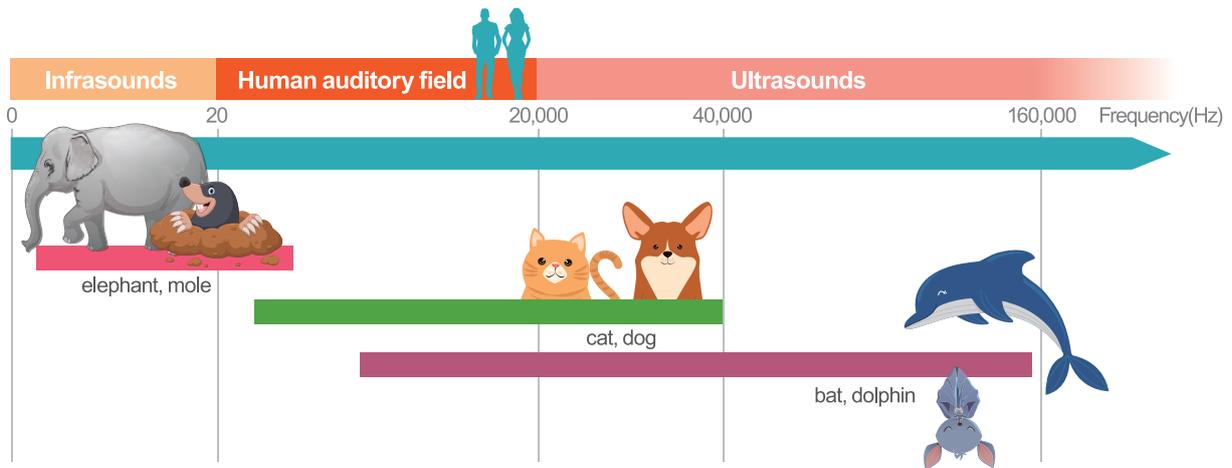
資料來源：<https://machinelearning.apple.com/research/hey-siri>

一 · 認識聲音訊號與處理

(一) 什麼是音訊

「音訊」就是聲音訊號 (audio signals) 的簡稱，通常指一般人耳朵能夠聽到的各種聲音，例如講話的聲音、好聽的音樂、環境的噪音等。

人的耳朵可以聽到的頻率範圍是從 20Hz (赫茲) 到 20,000Hz；不同的生物可以聽到的聲音頻率範圍也不同。有些動物可以聽到更高頻的聲音，有些則更低。有些頻率超過 20,000 Hz 的聲音，則是人耳無法聽到的聲音，稱為超音波。



▲圖 2-7¹

《音訊小秘密》

因為人耳所能聽到的聲音，會隨著年紀而改變。年紀越大，能聽到的聲音頻率範圍就越窄。如果你的上課老師年紀已經大於 50 歲，你可以使用蚊子鈴聲 (mosquito ringtone) 來做為你的手機鈴聲，由於其基本頻率很高，當上課時你手機響起，只有你和你的同學聽得到，而臺上的老師可能就聽不到了！

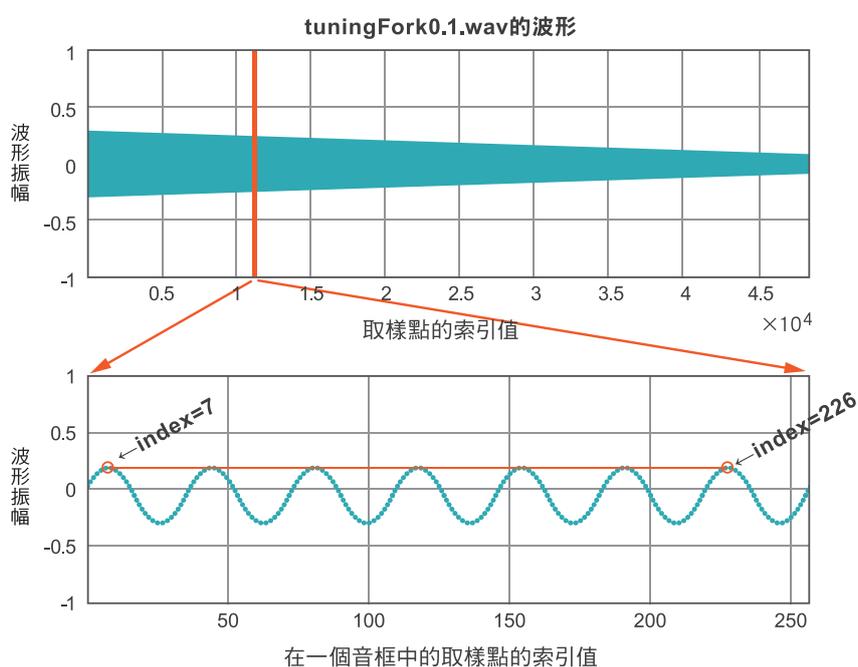
資料來源：Jyh-Shing Roger Jang, "Audio Signal Processing and Recognition," (in Chinese) available at the links for on-line courses at the author's homepage at <http://www.cs.nthu.edu.tw/~jang>.

1 圖片來源：<https://www.vapor.com.tw/%E8%81%BD%E5%8A%9B%E5%B0%8F%E7%99%BE%E7%A7%91/%E5%93%AA%E4%BA%9B%E5%8B%95%E7%89%A9%E6%9C%89%E8%91%97%E7%B5%95%E4%BD%B3%E7%9A%84%E8%81%BD%E5%8A%9B/>

(二) 音訊的分析

當我們的耳朵聽到一段音訊後，通常可以立刻感受與分辨出這個聲音的三個重要特徵：包含音量（volume）的大小、音調的高低（pitch）和音色（timbre）的區別。

如果我們想要用電腦來分析一段音訊，就必須將它切成比較短的單位（約 20-30ms），這些單位我們稱為音框（frame），一個音框必須包含數個基本週期，才能讓電腦擷取出其音訊的特徵，並使用數學公式來描述它的特性，以「逼近」人耳聽到的感覺。這些由每一個音框所抽取出來的數值或向量就稱為聲學特徵（acoustic features），也是構成聲音的三個重要元素。



▲圖 2-4²

1. 音量 (volume) J S

音量又稱為「響度」，代表音訊的強度（intensity）或能量（energy），通常可以使用音訊的震（振）幅來類比，震（振）幅越大，音量越大，音量的單位是分貝（decibel）。

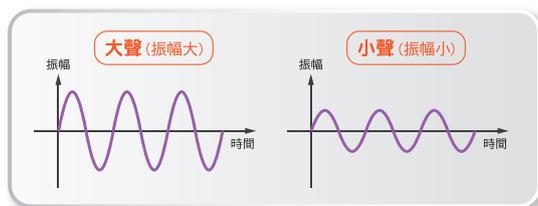


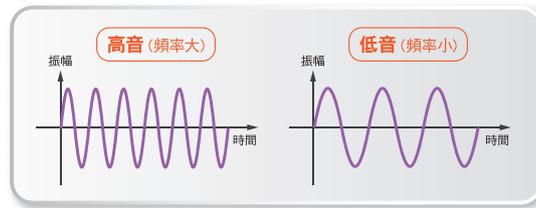
圖 2-5³

2 圖片來源：鴻海人工智慧導論教材簡報

3 圖片來源：110 學年度康軒版九年級資訊科技課本

2. 音高 (pitch)

音高又稱「音調」，代表聲音的高低表現，當聲波的振動頻率越高（密集），則音高越高；反之則越低。例如女生的聲音一般來說會比較高，而男生的聲音會比較低。通常我們會依據每一秒內出現的聲波基本週期的個數，來代表音高。



▲圖 2-6⁴

3. 音色 (timbre)

音色就是指「聲音的特色」，不同的聲波波型就會出現不同的音色，這也讓我們能區別不同的發音、不同的人聲、不同樂器的聲音等。例如「ㄅ」和「ㄆ」的發音方式不同，就會產生不同的音色；二個同學，一樣是說「你好」二字，產生的音色也會不同。



▲圖 2-7⁵

(三) 聲音的數位化

前面提到聲音的分析可以使用電腦來作業，但在這之前則需要先將聲音轉化為數位檔案。使用電腦軟體進行錄音並將音訊錄製成數位格式時，必須使用可以偵測空氣波動的感測器進行收音（通常是使用麥克風），收音後將壓力波轉換為電壓訊號，最後將此電壓訊號以數位的方式儲存下來。而在這個記錄音訊的過程中，通常我們必須記錄下列三項參數：

1. 取樣率 (sample rate)

取樣率是指在一秒內的取樣數量，通常以 Hz (Hertz) 為單位。取樣率越高，所儲存的音訊品質越好，但所佔用的儲存空間也會越大。一般常見的取樣率如下：

4 圖片來源：110 學年度康軒版九年級資訊科技課本

5 圖片來源：110 學年度康軒版九年級資訊科技課本

▼表 2-1 聲音的取樣率

項目	取樣率	品質
一般電話的通話	8 kHz：每秒鐘取樣 8000 點	低
一般語音辨識的錄音	16 kHz：每秒鐘取樣 16000 點	中
音樂 CD	44.1 kHz：每秒鐘取樣 44100 點	高

2. 位元解析度 (bit resolution) J S

對於每一個取樣點，我們要使用的位元 (bits) 數來代表此點。一般常用的位元數是 8 或 16 個位元，位元數越高，音訊品質越好，但是需要的儲存空間也會越大。

3. 聲道數 (number of channels) J S

同時要用幾個麥克風來錄音，一般是單聲道 (mono) 或是雙聲道 (stereo，又稱為立體聲)。

當我們將聲音數位化後，就可以利用上列三項參數來進行語音的分析外，也可以藉由各種聲音編輯軟體進行編修，以取得我們需要的聲音檔案資料。



單元 2

語音辨識的運作流程與類型

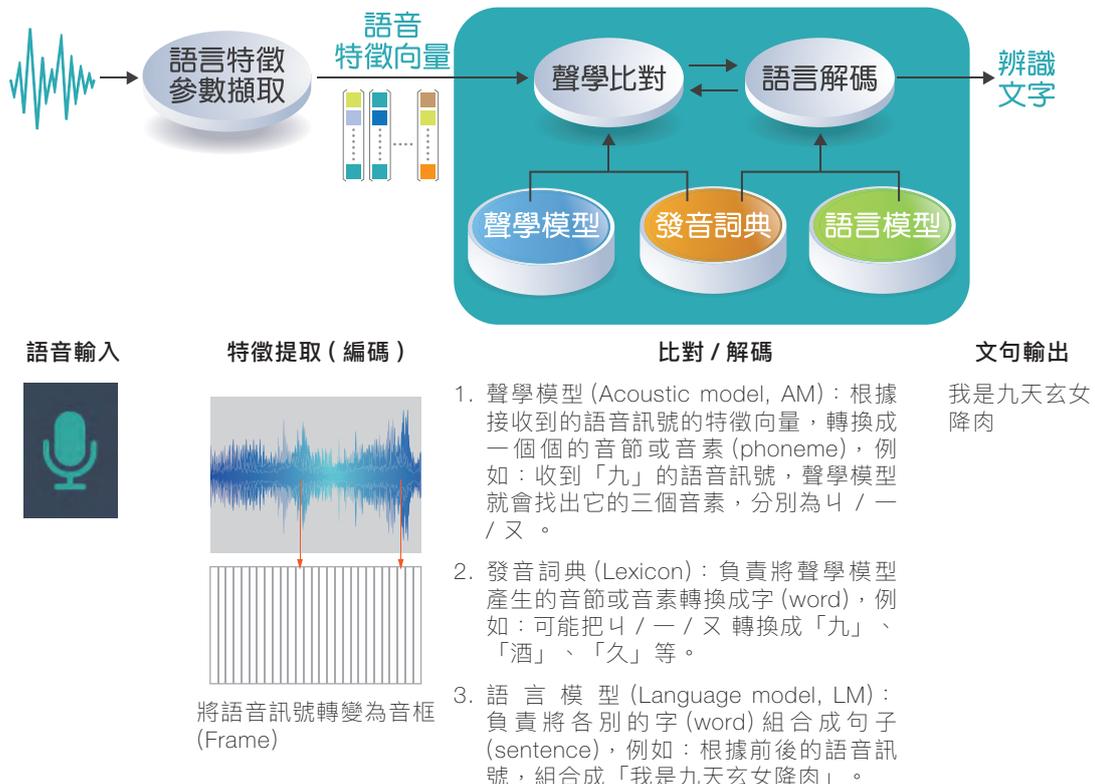
作者：汪殿杰—大同高中 | 蔡志聰—江翠國中

一．把語音轉變成文字——語音辨識的運作流程

語音辨識主要的目的是希望電腦聽懂人類說話的聲音，進而命令電腦執行指定的工作。這個過程的第一步，需要確認電腦是不是真的聽見以及能辨識出它所接收到的語音。

當電腦接收到一段語音後，會先將這段語音由類比訊號轉換為數位訊號，並將這段聲音訊號和已儲存好的聲音樣本（這些聲音樣本都是我們事先蒐集和標註好的）進行語音特徵比對。比對完成後，電腦即找出一個最「接近」的聲音樣本，並執行這項工作。

然而，即使是同樣一個語詞或一段話，由不同的人說出來，聽起來也都不會一模一樣。但是電腦還是可以辨識出你說的話，並轉換成為文字，主要就是進行語音識的比對和解碼過程中，使用到聲學模型、發音詞典及語言模型這三個語音處理功能。運作示意流程如下：

▲圖 2-18⁶

6 圖片來源：改編自「現階段大詞彙連續語音辨識研究之簡介」

二 · 語音辨識系統的分類

有關語音辨識的各種應用，可以根據不同的使用者和運作執行難度來分類。

第 1 種類別：根據語音辨識系統的不同使用者來分類

1. 特定說話者 (speaker dependent) 的語音辨識：系統使用者只限定在特定的人，也就是這個語音辨識系統，只能辨識固定的人聲，換作他人可能就無法使用。
2. 非特定說話者 (speaker independent) 語音辨識：系統的使用者可以通用於每一個人。

第 2 種類別：是根據語音辨識系統的功能與執行難度來區分

1. 語音命令 (voice command)：使用者下達一句語音指令，系統從已儲存的指令資料庫中，找出最佳符合程度的指令，並執行相關動作。
2. 關鍵詞偵測 (keyword spotting)：使用者下達一句語音（例如「請幫我查詢今天天氣如何？」），系統可以偵測這一句語音是否含有特定的內容（如「今天」及「天氣」）。
3. 聽寫 (dictation)：使用者下唸出一段語音內容（例如一段新聞播報），系統可以自動產生正確逐字稿，例如 youtube 的字幕產生系統。
4. 對話 (dialog)：使用者可以直接和電腦對話，電腦收到一段語音後，能夠瞭解使用者的意圖 (intention)，並以語音或文字進行合理且正確的回覆。一般的語音助理或聊天機器人 (chatbot)，就是屬於這種系統。

根據上面的分類，最簡單的語音辨識系統就是「特定說話者的語音命令辨識系統」，其原理是「用自己的聲音比對自己的聲音」，例如早期的手機（如 Sony Ericson T18）的語音撥號功能，可以預錄幾組語音，每一個語音對應到一組電話號碼，用「拉麵」的語音對應到拉麵店的電話，因此當你對手機喊「拉麵」時，系統會對你輸入的聲音以及已經預錄好的聲音進行比對，若比對正確，手機就會自行撥電話到拉麵店。（但是，對電腦而言，人類的語音變化度極大，若是金城武的妹妹對金城武的手機喊「拉麵」，不見得有效，因為內部用來比對的錄音是金城武的聲音，而不是他妹妹的聲音。）

最複雜的語音辨識系統，就是「非特定說話者的對話系統」，例如蘋果手機的 Siri 語音助理、亞馬遜的 Alexa 智慧音箱，以及 Google Home 語音助理，這些系統就像是虛擬助理，都可以和人們進行簡單的對話，同時藉由瞭解使用者的意圖，幫忙人們做一些簡單事情，如預定車票、查詢天氣或電影等。

「特定說話者的語音命令辨識系統」—Sony Ericson T18 的語音撥號功能影片連結：<https://youtu.be/V4RiDUAqajM>

引用資料來源：Jyh-Shing Roger Jang, "Audio Signal Processing and Recognition," (in Chinese) available at the links for on-line courses at the author's homepage at <http://www.cs.nthu.edu.tw/~jang>.



單元 3

語音辨識的升級 ——以文字為主的自然語言處理技術

作者：汪殿杰——大同高中 | 蔡志聰——江翠國中

在經過語音辨識系統的一連串運算比對後，電腦可以將出我們說的內容辨識出來，並以文字或語音輸出，但我們如果想知道電腦是否了解這些文字的意思，並作出適當的回應，這時就需要使用到人工智慧裡的「自然語言處理」(Natural Language Processing，簡稱 NLP) 技術。

一．什麼是「自然語言」？

自然語言就是「我們自然而然說出來的語言」，在一般的對話情境中，我們的對話都會有上下文可參考，所以能互相理解語意並作出適當的回應，但對電腦來說「理解」和「回應」自然語言可是很難的一件事。

自然語言的例句

夏天時，媽媽對小明說：衣服能穿多少，就穿多少。

冬天時，媽媽對小明說：衣服能穿多少，就穿多少。

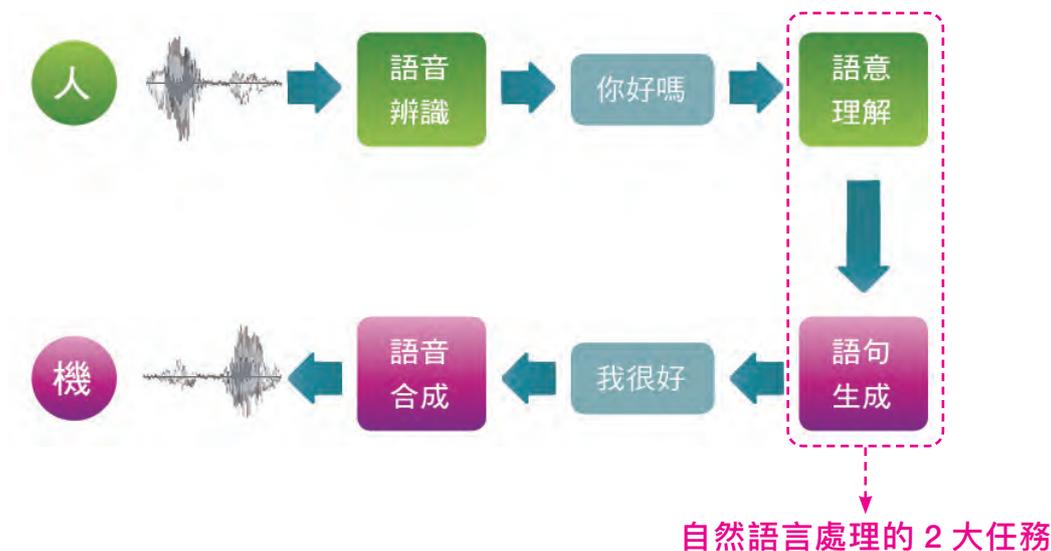
你覺得聰明的電腦可以理解這二段話的意思嗎？

二．「自然語言處理」可以做什麼？

自然語言處理 (NLP) 就是可以讓電腦能理解、產生及使用人類語言的一種技術，屬於人工智慧發展的其中一環。

在自然語言處理中有 2 個重要的任務，分別是「自然語言理解 (Natural Language Understanding，簡稱 NLU)」及「自然語言生成 (Natural Language Generation，簡稱 NLG)」兩大部份，第 1 部份是讓電腦理解人類所講的話；第 2 部份是讓電腦能生成出人類認為合理回應的一段話或是一段文句。

以智慧音箱為例，當使用者說：「現在是幾點？」首先機器要先判斷使用者的意圖，理解這個問題的重點是什麼？然後再以生成一段語音來說出：「現在是 7：30 分」。



▲圖 2-13 機器從聽到人說一句話到做出回應所需要的 4 項關鍵技術⁷

Google 助理向美髮沙龍預約剪髮

谷歌 2018 年 I/O 大會上，執行長 Sundar Pichai 展示了一段語音助理打電話給真實髮廊的預錄影片，他以預約剪頭髮為例，谷歌語音助理可以撥打電話給用戶指定的髮廊並向對方預約剪頭髮的日期時間，從頭到尾，用戶根本不必花時間打這通電話，只要交代語音助理去做就行了。

影片連結：<https://reurl.cc/vka1ka>



初期的自然語言處理主要以大量的統計資料去訓練模型，讓電腦閱讀大量類似字典的文章段落，再用演算法計算單字、句子出現的機率，但這樣的方式無法使系統很好地辨識複雜的文法。

1980 年代末期，自然語言處理引進機器學習 (Machine Learning) 的演算法，不再用程式語言命令電腦所有規則，而是建立演算法模型，讓電腦學會從訓練的資料中，尋找資料所含的特定模式和趨勢，自動歸納出語言的特性。

7 資料來源：科技大觀園 / 機器能聽懂人說話嗎？ / 李宏毅 網址：<https://scitechvista.nat.gov.tw/Article/c000003/detail?ID=f427e540-4e86-4344-881a-f0faf28b5982>

三 · 如何教會電腦認識字詞呢？

我們都知道構成文句的最小單位是「字詞」，字詞組成句子，小句子繼續組成大句子與文章，所以使用自然語言處理的首要的目標就是能「辨識字詞」。

為了讓電腦理解中文的句子，自然語言處理需要做到兩件事：

第一件是將句子「斷詞」以理解個別詞義；第二件是「分析句子」，包括文法和整個句子的語意解讀。

將句子斷詞舉例

我是一個天才 □ 「我 · 是 · 一個 · 天才」

我今天才遇到他 □ 「我 · 今天 · 才 · 遇到 · 他」

這兩句話雖然都有「天、才」，但兩個句子的斷句方式不同，「天、才」在這兩句話中的意思也不同。

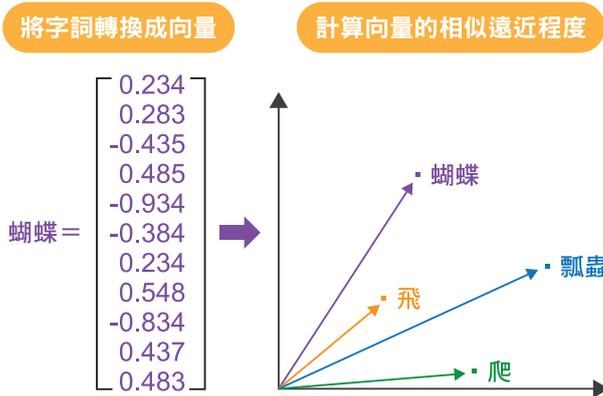
中央研究院中文詞知識庫小組線上斷詞系統
<https://ckip.iis.sinica.edu.tw/service/corenlp/>

就如同「我是一個天才」的這段句子，在 NLU 中的處理上首先就必須要先區分出各個不同的字詞，再進一步的理解字詞之間的關係。

四 · 字詞間的關係——詞向量 (Word Vector)

為了訓練電腦理解字詞間關係，於是將字詞轉換一個數學數值，稱為「詞向量」，過程中讓電腦閱讀大量文章，利用前後文的統計特性，慢慢學習出每一個詞彙的詞向量，而不必利用任何語言學知識。如此一來，字意比較為相關的字詞，其向量距離也就越接近，反之亦然。

例如，電腦在閱讀大量的文章、文句及字詞內容中，會發現蝴蝶和瓢蟲的描述都會寫到「飛舞」或「飛走」；但可能只有「瓢蟲」會提到在枝葉花朵上「爬行」，因此「蝴蝶」與「飛」的向量距離就會比較接近，而「蝴蝶」與「爬」的向量距離則較遠（這也和我們一般人的認知概念相符）。



▲圖 2-10 不同字詞的向量關係圖⁸

隨著訓練的文本越來越多，電腦可以自動調整各個詞彙的向量，以解決訓練資料不足的問題，並提升電腦的抽象化思考與理解字詞的能力。

當電腦可以理解人類輸入的訊息後，它就能依據從內建的資料庫或語言邏輯形式，生成自然語言後，來回應我們的問題或指令，達成人機互動的目標，並執行更多和生活結合的應用。

Chat-GPT(Chat Generative Pre-trained Transformer)

由 OpenAI 開發的聊天機器人程式，主要以文字方式互動，除了可以透過人類自然對話方式進行對談，還可以用於相對複雜的語言工作，包括自動文字生成、自動問答、自動摘要等在內的多種任務。推出後即獲得大量的關注與討論。



▲圖 2-11 與網路聊天機器人 Chat-gpt 的問答⁹

8 資料來源：如何斷開中文峰峰相連的詞彙鎖鍊，讓電腦能讀懂字裡行間的語意？(泛科學)，網址：<https://pansci.asia/archives/144346>

9 資料來源：<https://www.cool3c.com/article/186079>



單元 4

語音辨識的挑戰

作者：汪殿杰—大同高中 | 蔡志聰—江翠國中

從阿里巴巴的「芝麻開門」咒語開始，語音辨識就一直是人類長久以來的想像創意與渴望。發展至今，在我們生活週遭，結合人工智慧的語音辨識裝置與應用已隨處可見，如智慧居家、語音助理、車聯網、行動載具、語音客服等，這些裝置透過各種軟體或程式隨時接受使用者提出的需求並以語音或文字回應。只是從使用者經驗來看，仍可能有所限制與風險問題有待克服：

一．環境噪音的干擾

若是接收的音訊是清楚無雜訊的情形下，電腦辨識的結果就幾乎能和期望相符合。但在現實環境中會有許多干擾語音內容的雜訊、噪音，輕則像是人們之間的交談聲，重如施工工地或是飛機起降時的噪音，這些環境的噪音往往會造成語音辨識的效果變差。

二．語音的變化與混用

不同地區的人，說話的語調、語別、發音及方言使用時的混用等歧異性，可能遠大於語音辨識量能的負荷。閩語、客語、山東話、四川話等不同語別的發音與使用的混用對於語音辨識發展應用而言，則仍有相當的挑戰。

三．語音訊息遺漏或結構影響

對於老人、幼童或發音功能受限者，說話時的訊息可能就會有不完整或語句結構缺漏的問題，對機器而言，接受的訊息數量及品質，都會直接顯現於後續的回應結果中。

四．資安與隱私權的疑慮

目前多數語音裝置的運作，均是採雲端服務模式，因此使用者的聲音（或聲紋），都將上傳給提供服務的廠商，過程中是否有隱私權益及影響資訊安全的問題，仍有待進一步的討論。

第三篇



AI 應用與實作



單元 1

AI 語音辨識的應用與思考

作者：蔡勝安—江翠國中



體驗活動

1. 語音助理互動體驗
2. 真的是辨識「語音」嗎？

1. 語音助理互動體驗

手機語音助理的應用，就像是有個「聽話」的小幫手，可以依照你的吩咐來做動作。同學們或多或少都有使用語音助理的經驗，但語音助理的「聽話」程度，也就是能理解你說出來的話到怎樣的程度呢？就讓我們用手邊的平板來試試看吧。

1. 先跟它打招呼，看它會如何回應你。（可以嘗試不同語言、不同的說法，或是不斷重複）
2. 問它一些有具體答案的問題，例如：天氣、時間、數字計算等，看它會如何回應你。另外，試試看如果你依據它的回答繼續問下去，看它會如何回應你。
3. 問它一些沒有標準答案的問題，例如：心情、看法等，看它會如何回應你。另外，試試看如果你依據它的回答繼續問下去，看它會如何回應你。
4. 繼續和它互動，並觀察記錄看看，那些問題是它可以理解的，又有那些問題是它無法聽懂的。
5. 試著總結出語音助理可以聽懂的問題，並分析它能聽懂的原因以及聽不懂的可能原因。

2. 真的是辨識「語音」嗎？

現今有很多科技產品都號稱可以利用語音來做控制，但對語音辨識技術已有基本認識的你應該了解，機器要能理解出人們說出的話（自然語言），其實並沒有那麼簡單。因此，請同學們試著討論分析一下，這些宣稱可以藉由語音控制的產品，其真實的控制原理可能是甚麼？

1. 分組進行，小組利用平板搜尋具備語音控制功能的科技產品。
2. 仔細閱讀其功能介紹後，思考看看，該產品真的有用到語音(自然語言)辨識嗎？亦或只是「聲音」控制呢？
3. 試著思考一下，「聲音」控制和「語音」控制的差異為何？在使用上會有甚麼差異呢？
4. 各小組請輪流上臺分享一下自己的檢索與分析成果。

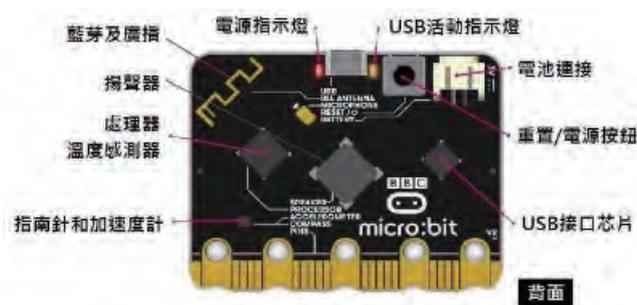


實作活動

聲控變幻燈

聲控燈是現今市面上常見的科技產品，它的聲控功能方便使用者在一定範圍，利用聲音即可對燈光的明滅與亮度做控制。而已經具備基本程式撰寫以及設計製作能力的你，應該也有能力設計出符合你需求的聲控燈，因此，在此將請同學利用「micro:bit」和「語音辨識模組」，來製作一個專屬於自己的聲控變幻燈。

指定材料 製作控制系統用，micro:bit V2，專用擴充板（悟空板），語音辨識模組，三色燈條（5V）。



(a) micro:bit V2



(b) 專用擴充板（悟空板）



(c) 語音辨識模組



(d) 三色燈條

▲圖 3-1 （缺圖）

自選材料 製作燈具外殼用，任何你想使用的，而且有能力加工的材料皆可。

- 要求**
1. 可以利用語音控制燈光的明滅與顏色變幻
 2. 利用自選材料設計並製作出燈具外殼，此外殼須將控制系統包覆在內，僅可露出聲音接收和充電孔位。
 3. 外型與主題無任何限制，但請設想預計將會擺在甚麼樣的「情境」下使用，最後成品需在真實情境下拍攝使用狀況之短影片。

進行步驟

- STEP 1** 情境設定，你預計將變幻燈放在哪裡呢？
- STEP 2** 評估需求，你希望變幻燈需要具備哪些變換功能？
- STEP 3** 外型設計，你希望變幻燈尺寸多大？造型？顏色？以及如何包覆控制系統？
- STEP 4** 決定材料，依據外型設計的要求，決定你想使用的材料。
- STEP 5** 控制系統設定，進程式撰寫與功能測試，確定無誤後再進行外殼製作。
- STEP 6** 外殼製作，依據設計圖進行製作，過程中如果需要調整設計也請留下紀錄。
- STEP 7** 組裝，將控制系統與外殼做組裝，並做最後的調整以確保功能正常。
- STEP 8** 拍攝短影片，將作品擺在預設的情境後，拍攝實際聲控使用的影片。



專題活動

解決生活中小問題

科技的發展協助人們解決所遭遇的問題，但除了追求便利之外，也應該利用科技來達到人文關懷的功能。舉例來說，聲音控制技術的發展，讓肢體不方便的人們，可以做到以往不容易做到的事，例如打字、啟動開關或是移動等，讓人們在生理上得到方便，心理上也得到獨立自主的機會。

因此，在學會簡單的聲控技術之後，想請同學們嘗試思考看看，生活中有哪些事情，如果利用聲音來做控制，是會更加便利與適用所有人的呢？。

1. 請先設想你要協助的對象，例如：老年人、肢體不方便的人、家務繁忙的人、特定工作的人等。（為方便進行，同學們可以先將情境限縮在校園環境裡的人事物，後續比較能夠實現）
2. 以你設定的對象為出發點，利用同理心去設想它可能遭遇的狀況。
3. 針對設想的狀況，提出可能的解決方法。（雖然方法很多，但在此希望同學們先以「聲控」技術為主來做出發想）
4. 將同理的對象，設想的情境，想到的方法以及帶來的效益等四項元素，以四格漫畫的方式呈現出來，並公開發表給所有同學知道。
5. 發表完畢後，進行票選，由全班同學依據可行性、適切性、獨創性等要項，選擇出大家最喜歡的那個想法。
6. 以全班為一個大團隊，共同合作將該想法轉換成實際，共同去完成這個大專題。



單元 2

生活中的語音辨識

作者：林欽鴻—錦和高中



體驗活動 1

google 文件語音輸入

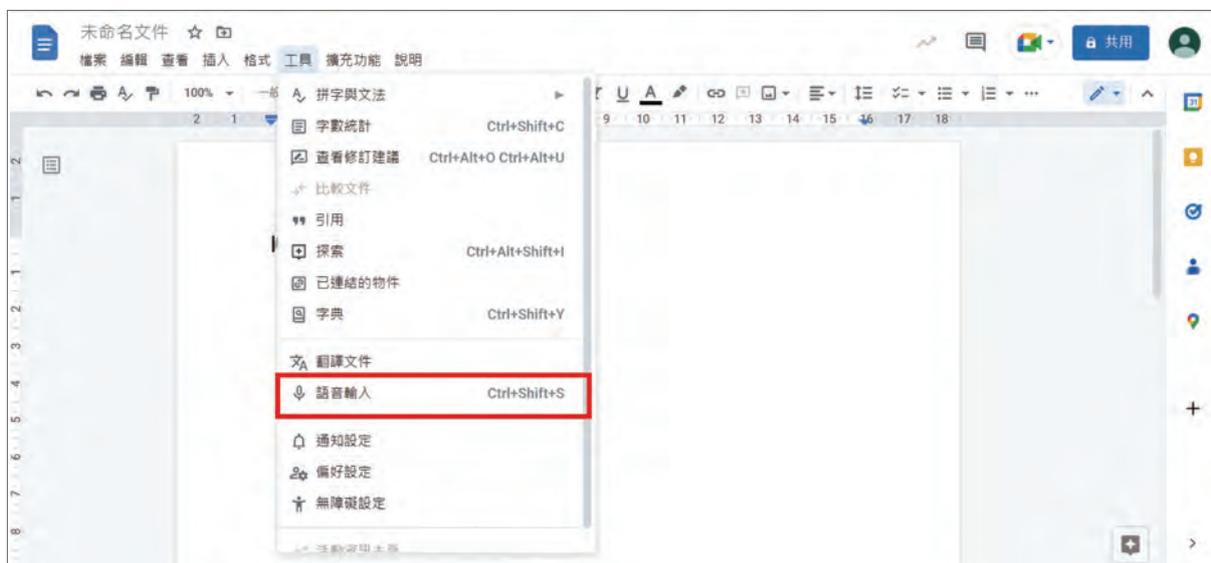
STEP 1

打開 google 文件

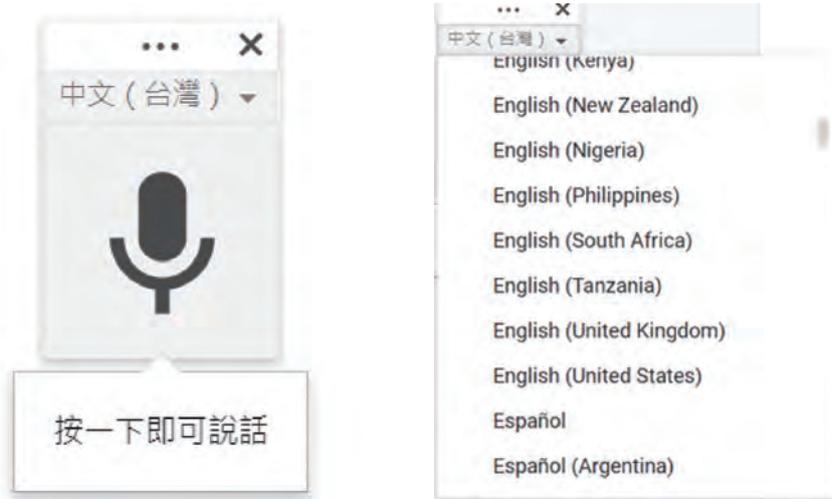


STEP 2

點擊上方工具選擇語音輸入



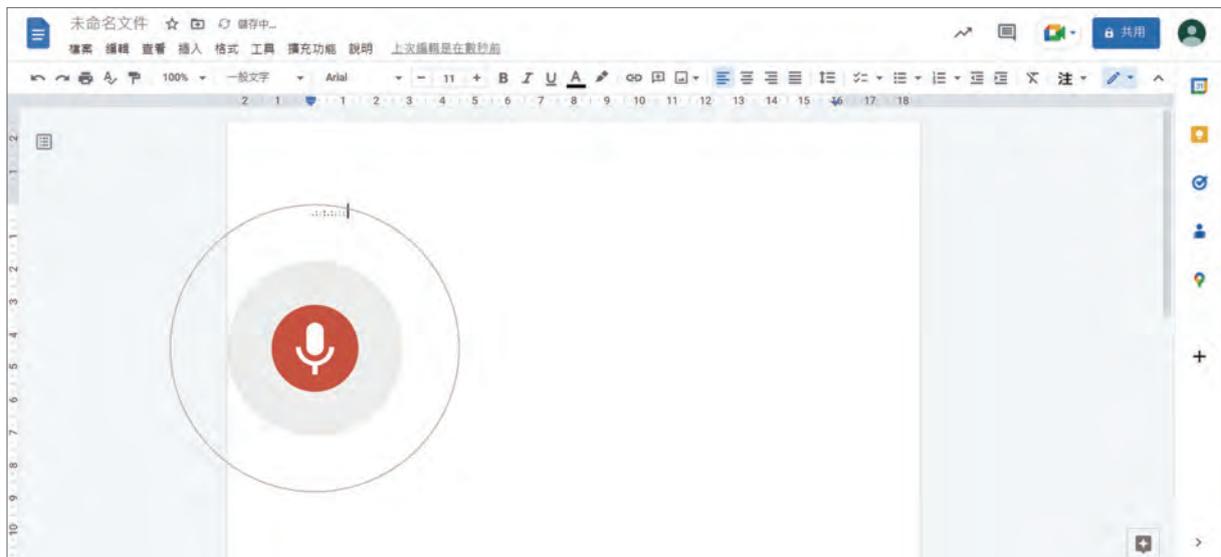
STEP 3 點擊下拉選單，選取要翻譯的語言



STEP 4 按下麥克風圖示並說出內容



STEP 5 將語音內容轉換成文字





體驗活動 2

Shazam 音樂辨識 APP

Shazam 有準確的音樂辨識能力，歸因於其有個龐大的音樂資料庫，而為了提升預測機器的準確度，Shazam 還會將使用者提供的回饋資料，特別是搜尋失敗的資料，幫助預測機器從回饋中學習，提升音樂辨識能力。

STEP 1 打開 APP 並開啟麥克風



▲ iOS



▲ Android

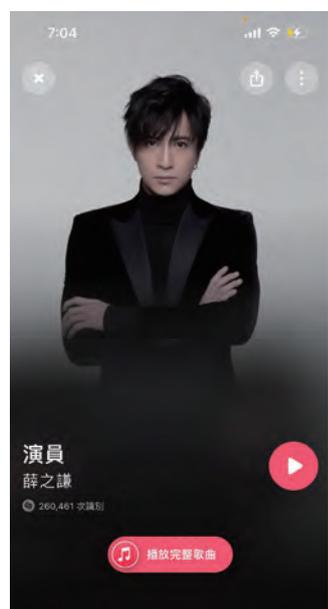
STEP 2 打開 APP 並開啟麥克風



STEP 3 點擊圖示，靠近音源即可開始辨識音樂



STEP 4 辨識成功，顯示歌曲名稱



操作範例



實作活動

Cyberpi 語音辨識操作

使用設備

1. cyberpi Go Kit 主控版。
2. 鋰電池擴展版。
3. 伺服馬達兩顆。
4. 雷切零件，作為點餐機構及出餐裝置。

cyberpi 硬體介紹



▲圖 3-2



▲圖 3-3



mBlock 使用教學

1. 下載 mBlock 軟體 (網站: <https://mblock.makeblock.com/zh-cn/>)



2. 註冊 mBlock 帳戶

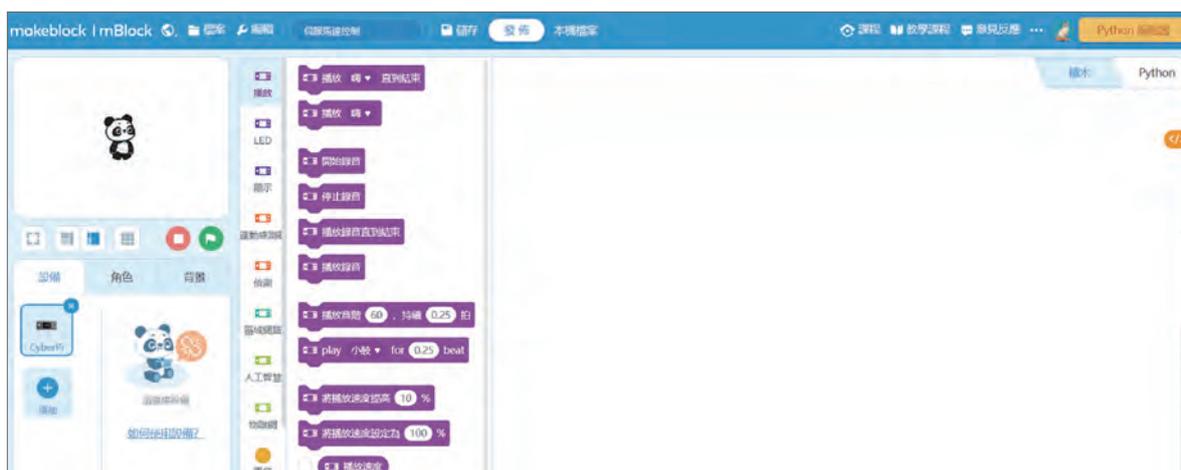
點選右上角登入帳號 (可直接使用 google 帳戶)

★使用 AI 語音辨識功能需登入帳戶，否則相關功能將受限制。

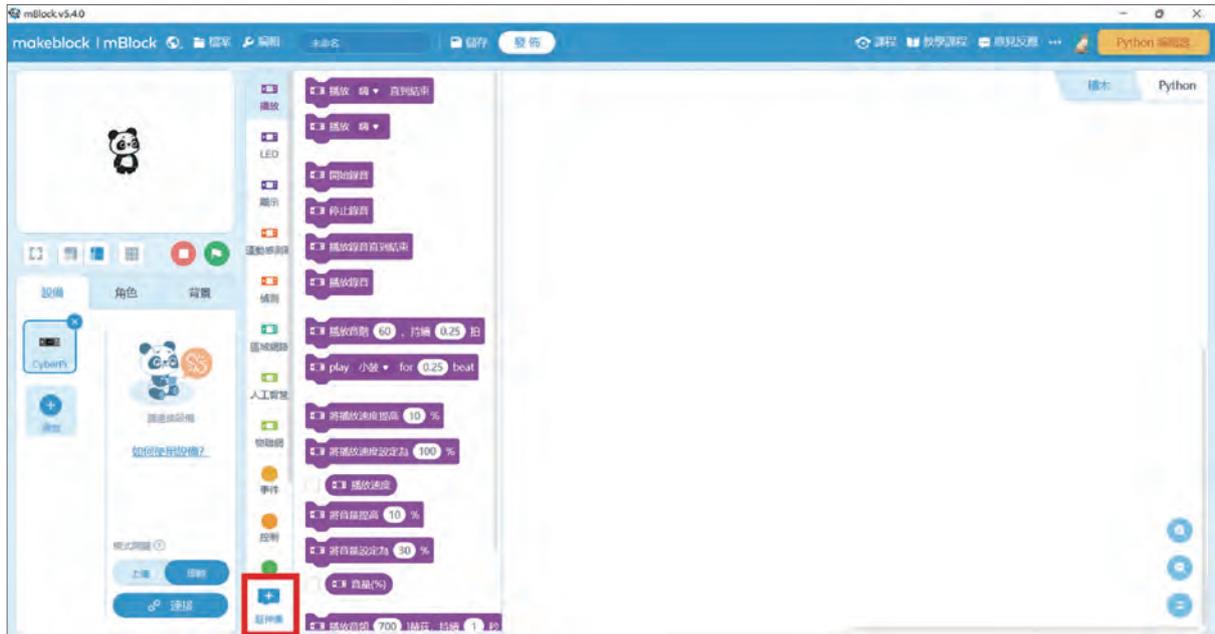


3. 開始使用

STEP 1 選擇設備：cyberpi (若找不到設備，則按添加)



STEP 2 新增擴充設備



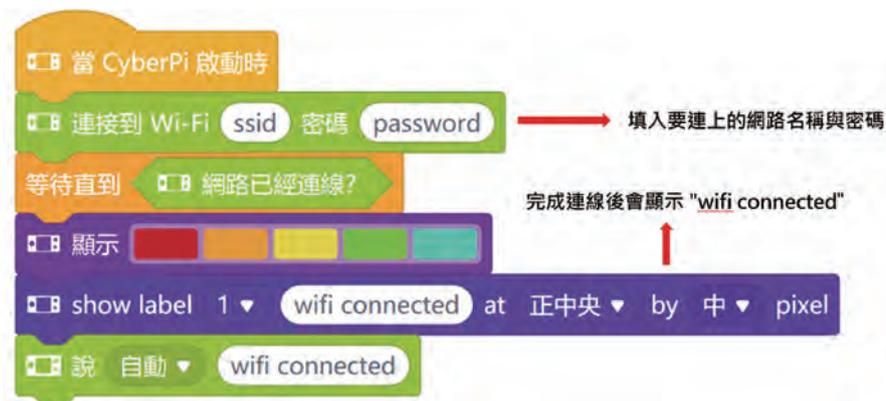
STEP 3 設備連接



STEP 4 更換至上傳模式（語音辨識模組需在上傳模式下才能啟用）



STEP 5 wifi 連線程式
拉出相對應的積木並輸入要連接的網路名稱 / 密碼



STEP 6 上傳至 cyberpi（上傳時 cyberpi 必須開機）



STEP 7 螢幕顯示 "wifi connected" 代表已連上網路



Cyberpi 語音辨識教學

1. 螢幕顯示並說出語音辨識內容

STEP 1 按壓搖桿，啟動語音辨識



STEP 2 選擇辨識語言種類



STEP 3 選擇中文 (繁體)



STEP 4 將語音辨識結果顯示於螢幕上



STEP 5 說出語音辨識結果

```
當搖桿 中間按壓  
show label 1 請說話 at 正中央 by 中 pixel  
在 3 秒後, 辨識 中文(繁體)  
show label 1 語音識別結果 at 正中央 by 中 pixel  
說 自動 語音識別結果
```

完整程式範例

```
當 CyberPi 啟動時  
說 自動 hello world  
連接到 Wi-Fi ssid 密碼 password  
等待直到 網路已經連線?  
顯示 [red, orange, yellow, green, blue]  
show label 1 wifi connected at 正中央 by 中 pixel
```

成果



```
當搖桿 中間按壓  
show label 1 請說話 at 正中央 by 中 pixel  
在 3 秒後, 辨識 中文(繁體)  
show label 1 語音識別結果 at 正中央 by 中 pixel  
說 自動 語音識別結果
```

2. 翻譯語言並說出

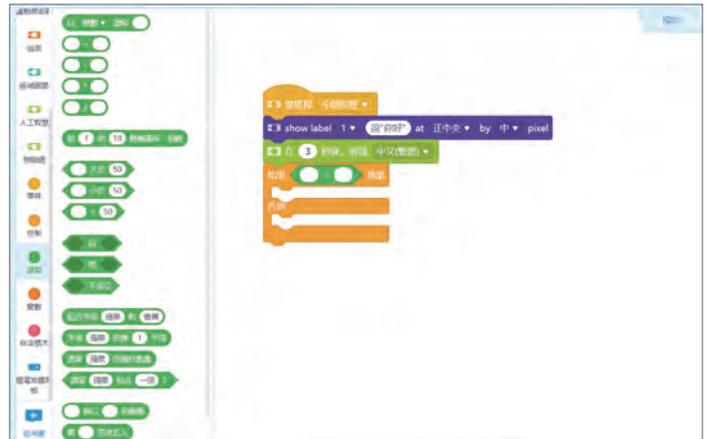
STEP 1 將相對應的積木拉出，並選擇中文（繁體）辨識

```
當搖桿 中間按壓  
show label 1 說你好 at 正中央 by 中 pixel  
在 3 秒後, 辨識 中文(繁體)
```

STEP 2 使用如果...否則的判斷式，如果辨識結果為"你好"

```
當搖桿 中間按壓  
show label 1 說你好 at 正中央 by 中 pixel  
在 3 秒後, 辨識 中文(繁體)  
如果 那麼  
否則
```

STEP 3 從運算類別中找出 " 等式 " 的積木



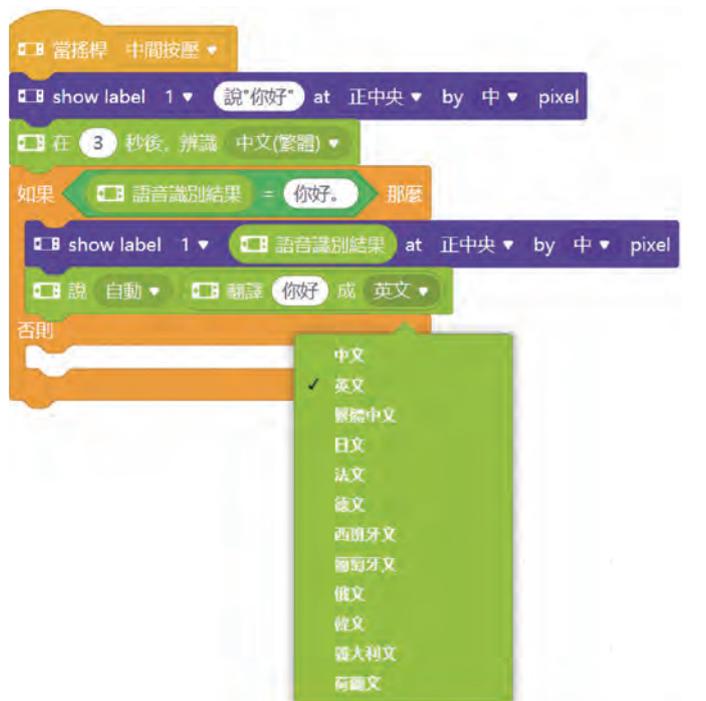
STEP 4 將語音辨識結果積木放入等式中



STEP 5 辨識正確，則螢幕顯示 " 你好 "



STEP 6 選擇要 Cyberpi 翻譯並說出的語言，辨識成功則說出



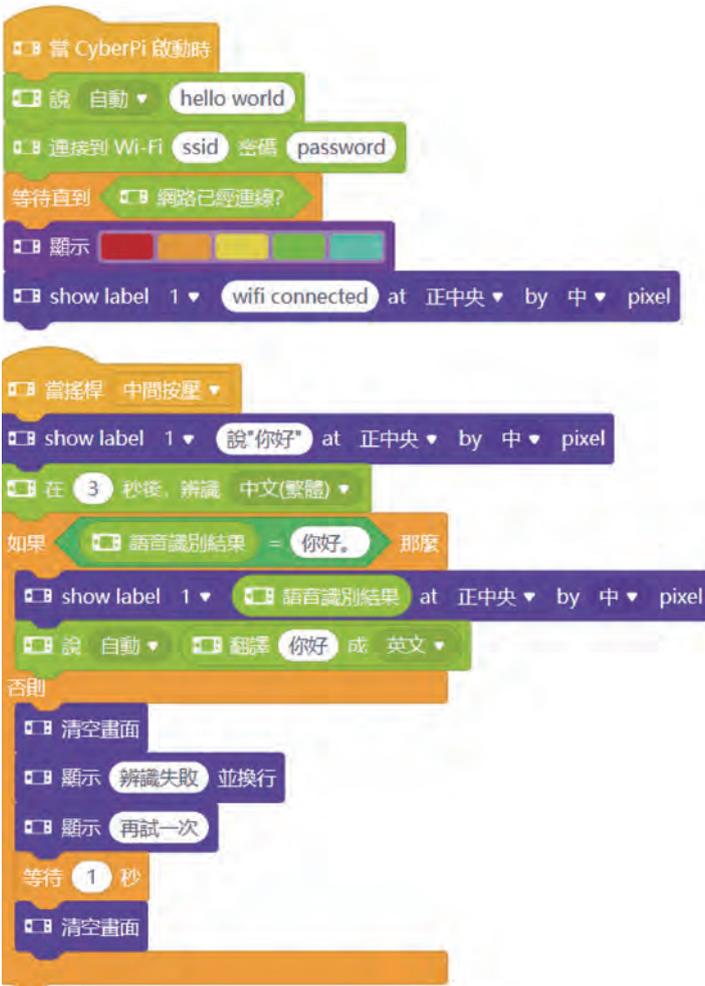
STEP 7 辨識失敗，顯示 " 辨識失敗 再試一次 "



當搖桿 中間按壓

- show label 1 說"你好" at 正中央 by 中 pixel
- 在 3 秒後, 辨識 中文(繁體)
- 如果 語言識別結果 = 你好. 那麼
 - show label 1 語言識別結果 at 正中央 by 中 pixel
 - 說 自動 翻譯 你好 成 英文
- 否則
 - 清空畫面
 - 顯示 辨識失敗 並換行
 - 顯示 再試一次
 - 等待 1 秒
 - 清空畫面

完整程式範例



當 CyberPi 啟動時

- 說 自動 hello world
- 連接到 Wi-Fi ssid 密碼 password
- 等待直到 網路已經連線?
- 顯示 [red, orange, yellow, green, blue]
- show label 1 wifi connected at 正中央 by 中 pixel

當搖桿 中間按壓

- show label 1 說"你好" at 正中央 by 中 pixel
- 在 3 秒後, 辨識 中文(繁體)
- 如果 語言識別結果 = 你好. 那麼
 - show label 1 語言識別結果 at 正中央 by 中 pixel
 - 說 自動 翻譯 你好 成 英文
- 否則
 - 清空畫面
 - 顯示 辨識失敗 並換行
 - 顯示 再試一次
 - 等待 1 秒
 - 清空畫面

成果





練習活動

使 CyberPi 辨識「歡迎光臨」並翻譯成 3 國語言



參考答案

```

    當 CyberPi 啟動時
    說 自動  hello world
    連接到 Wi-Fi  ssid 密碼 password
    等待直到 網路已經連線?
    顯示  [red, orange, yellow, green, blue]
    show label 1  wifi connected  at 正中央  by 中  pixel
    等待 1 秒
    不停重複
    show label 1  說"歡迎光臨"  at 正中央  by 中  pixel
    在 3 秒後 辨識 中文(繁體)
    如果  語言識別結果 = 歡迎光臨  那麼
        show label 1  語言識別結果  at 正中央  by 中  pixel
        說 自動  翻譯 歡迎光臨 成 英文
        說 自動  翻譯 歡迎光臨 成 日文
        說 自動  翻譯 歡迎光臨 成 法文
    否則
        清空畫面
        顯示 辨識失敗 並換行
        顯示 再試一次
        等待 1 秒
        清空畫面
    
```

3. 語音控制亮燈

STEP 1 按壓搖桿來啟動程式，使用迴圈積木



STEP 2 螢幕顯示 "說出要亮的顏色"



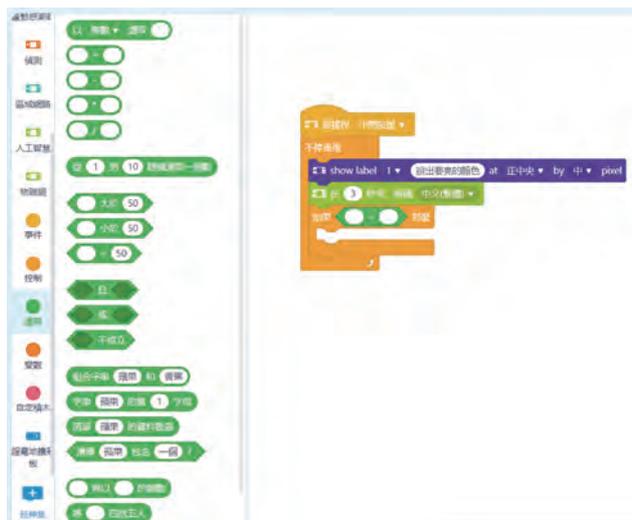
STEP 3 辨識中文 (繁體)



STEP 4 使用如果...判斷式



STEP 5 從運算類別中找出 "等式" 的積木

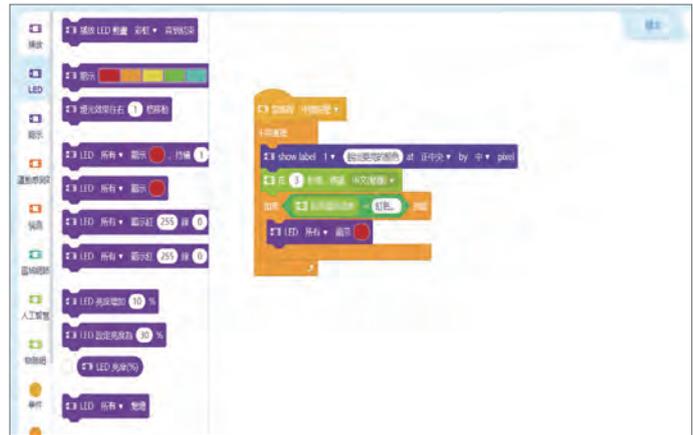


STEP 6 將語音辨識結果積木放入等式中。



辨識內容設定時，需在結尾加上句號

STEP 7 從 LED 類別中找出對應的積木。



STEP 8 LED 顯示的顏色可自行選擇，找出要亮燈的顏色。



STEP 9 依照相同方式，重複多次可使 cyberpi 顯示不同顏色的燈。



完整程式範例

成果



```
當 CyberPi 啟動時
說 自動  hello world
連接到 Wi-Fi  ssid  密碼  password
等待直到 網路已連線?
顯示 [red, orange, yellow, green, blue]
show label 1  wifi connected at 正中央 by 中 pixel

當搖桿 中間按壓
不停重複
show label 1  說出要亮的顏色 at 正中央 by 中 pixel
在 3 秒後, 辨識 中文(繁體)
如果 語音識別結果 = 紅色 那麼
LED 所有 顯示 [red]
如果 語音識別結果 = 藍色 那麼
LED 所有 顯示 [blue]
如果 語音識別結果 = 綠色 那麼
LED 所有 顯示 [green]
如果 語音識別結果 = 黃色 那麼
LED 所有 顯示 [yellow]
```



4. 語音控制馬達轉動

STEP 1 將伺服馬達接至 Cyberpi 電池擴充版的 S1 腳位



STEP 2

從鋰電池擴充版中找出伺服馬達的積木，並將伺服馬達設定為 90 度。



STEP 3

顯示說 "旋轉" 於螢幕上，並辨識中文（繁體）



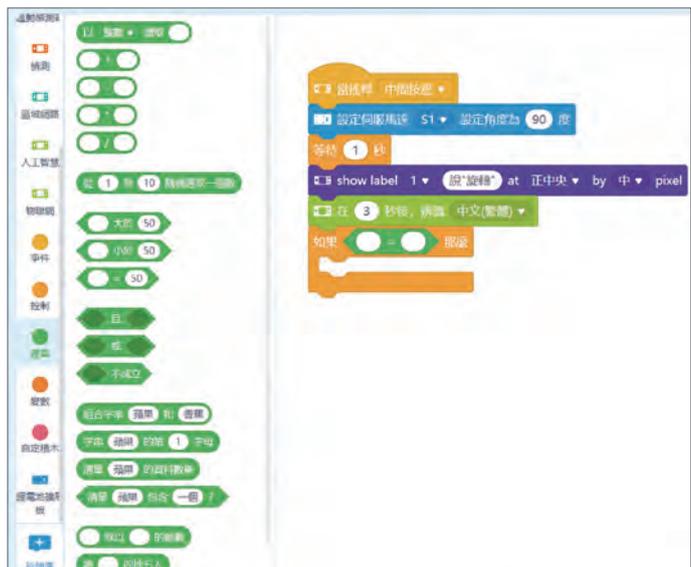
STEP 4

使用如果 ... 判斷式



STEP 5

從運算類別中找出 "等式" 的積木



STEP 6 將語音辨識結果積木放入等式中。

辨識內容設定時，需在結尾加上句號

STEP 7 辨識成功，則螢幕顯示 "旋轉"，且伺服馬達增加 10 度

完整程式範例

成果



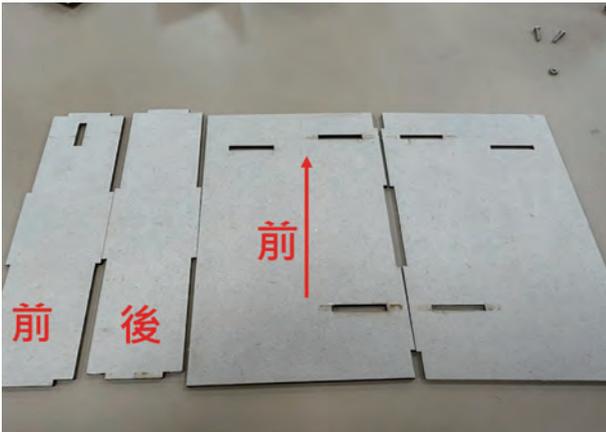


實作活動

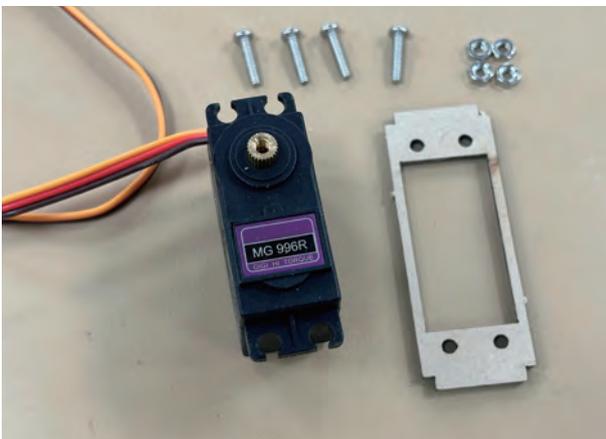
語音點餐系統

機構組裝

STEP 1 組裝側板，並用白膠接合



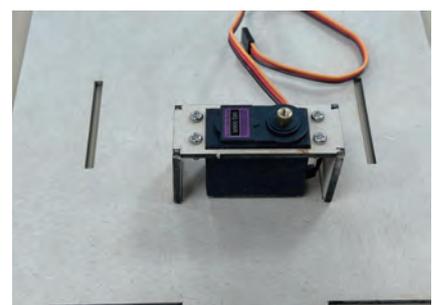
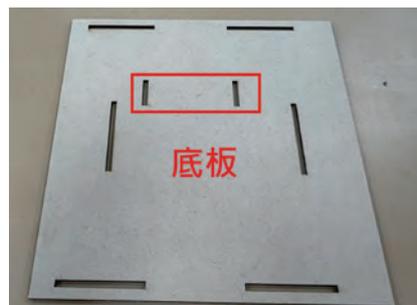
STEP 2 組裝大伺服馬達



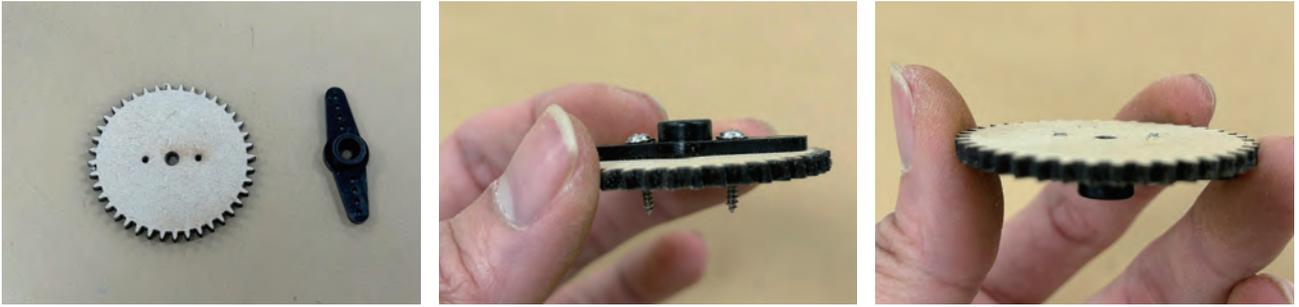
STEP 3 四個孔皆用 3mm 螺絲固定



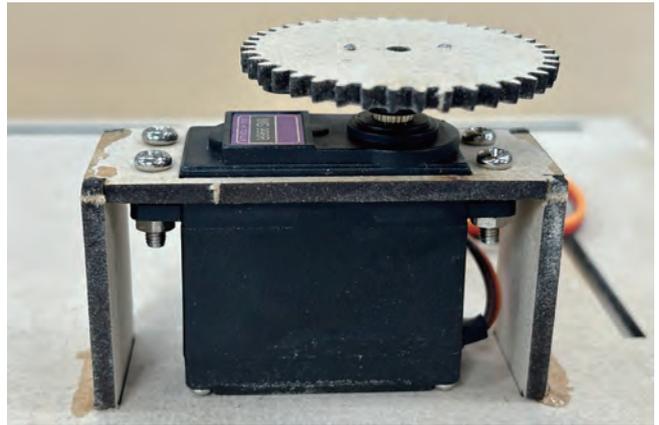
STEP 4 接合以下部件，並接合至底板



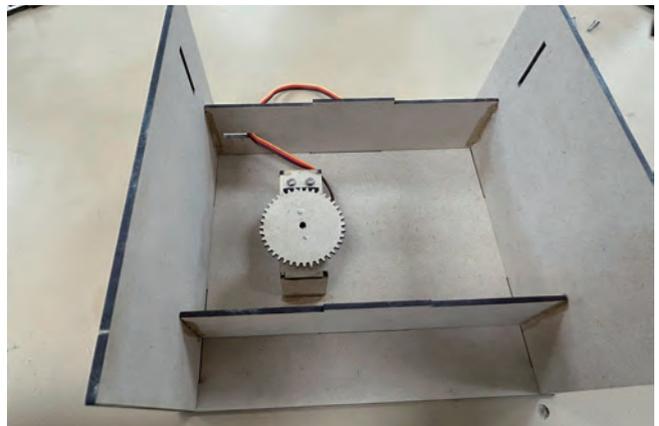
STEP 5 組裝大齒輪，並將螺絲突出部分剪斷



STEP 6 將其安裝至大伺服馬達上

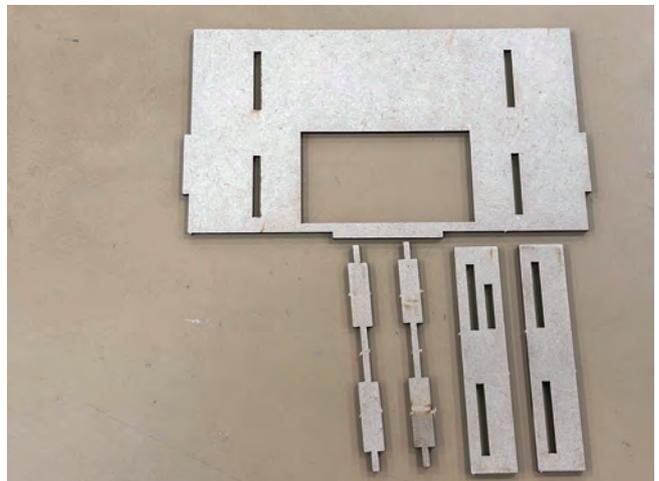


STEP 7 將側板安裝至底板上，且將大伺服馬達的線穿過前側板孔位。

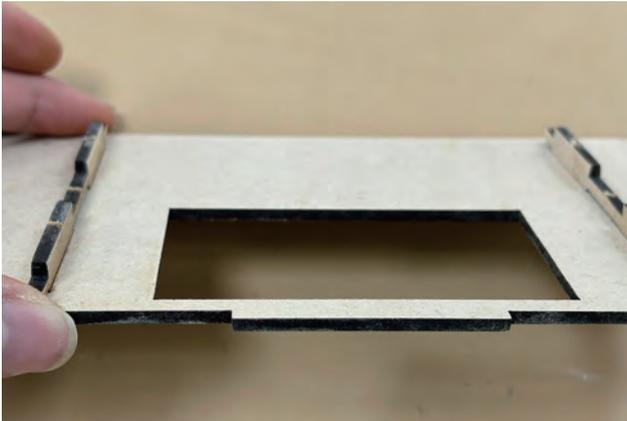


STEP 8 組裝取餐門，

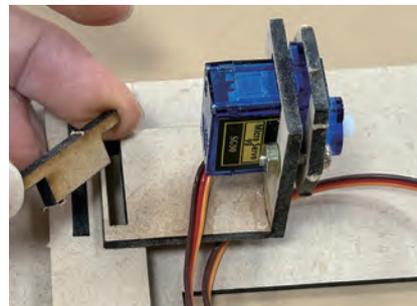
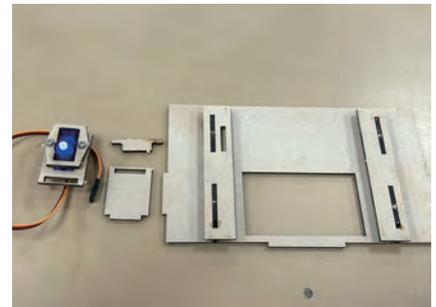
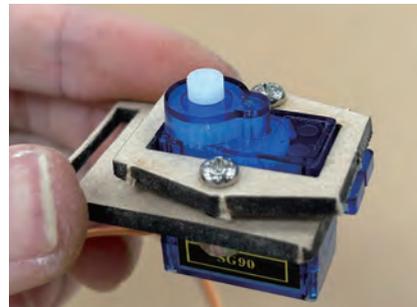
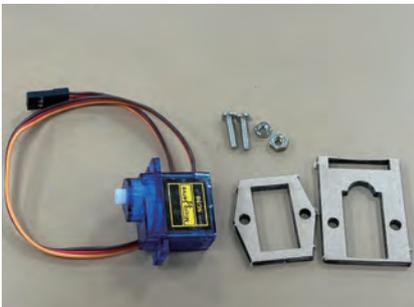
取出以下部件



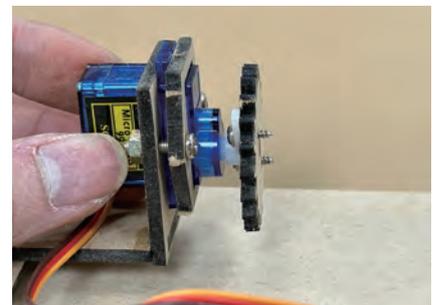
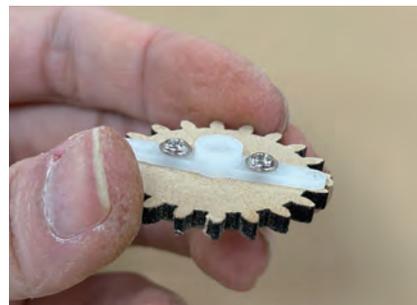
STEP 9 插入相對應孔位，並接合



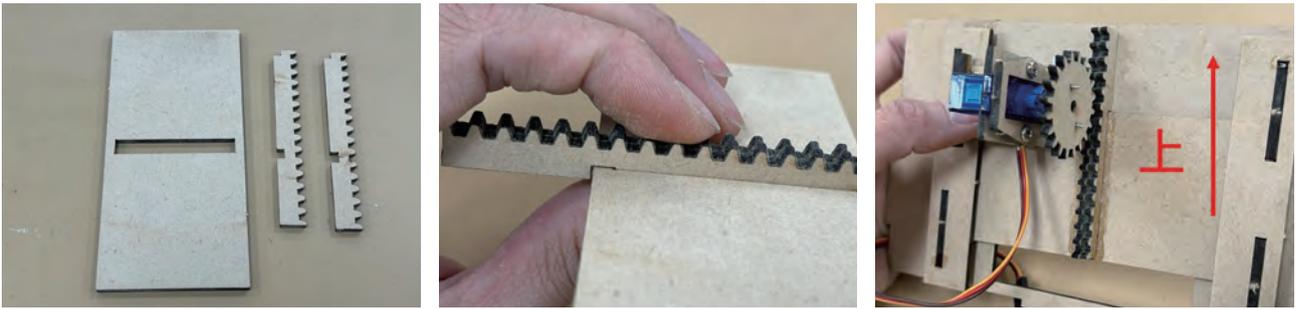
STEP 10 組裝小伺服馬達



STEP 11 組裝小齒輪，並安裝到小伺服馬達上



STEP 12 組裝出餐門，並組裝到門框上



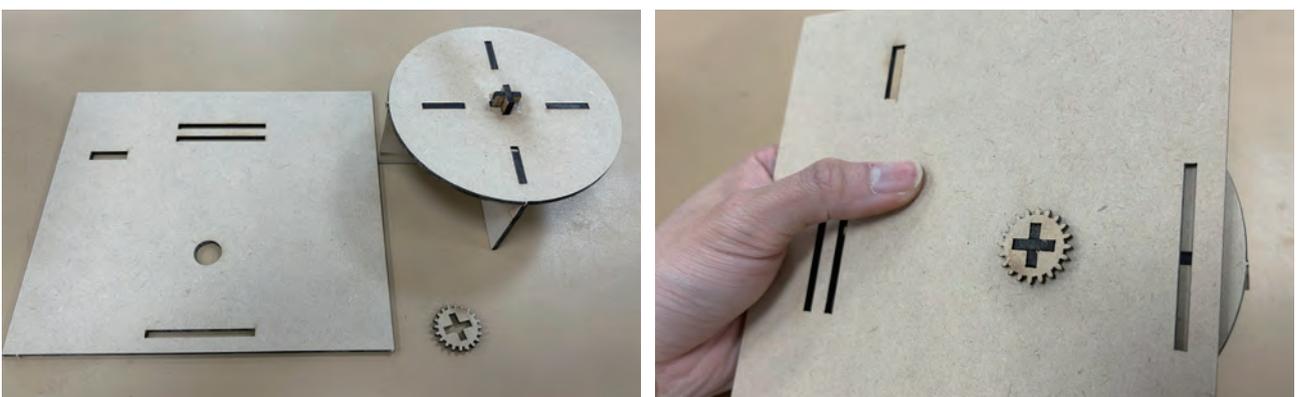
STEP 13 組裝旋轉餐檯，注意組裝方向



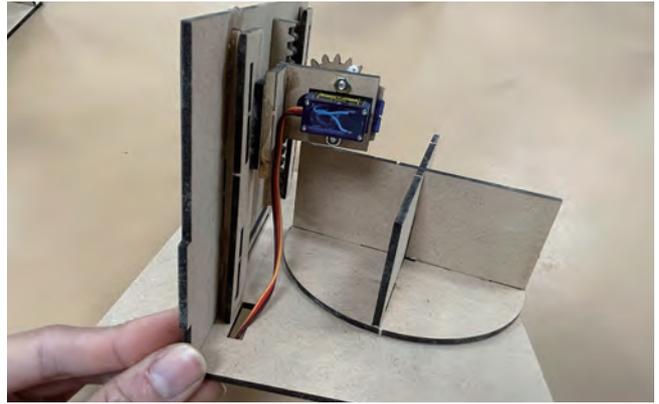
STEP 14 組裝旋轉柱，並與圓形轉盤接合



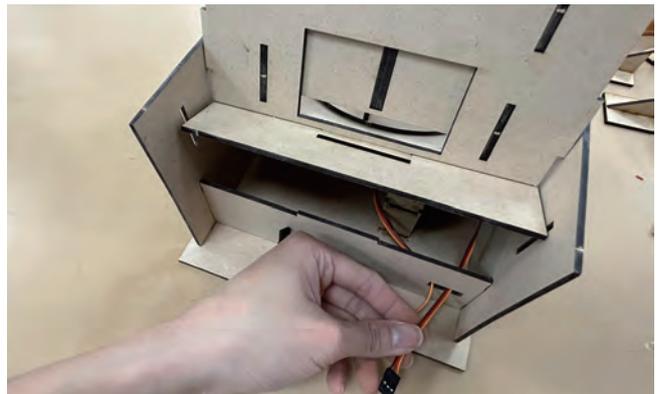
STEP 15 組裝頂板，上到下順序為：轉盤→頂板→頂板→十字齒輪



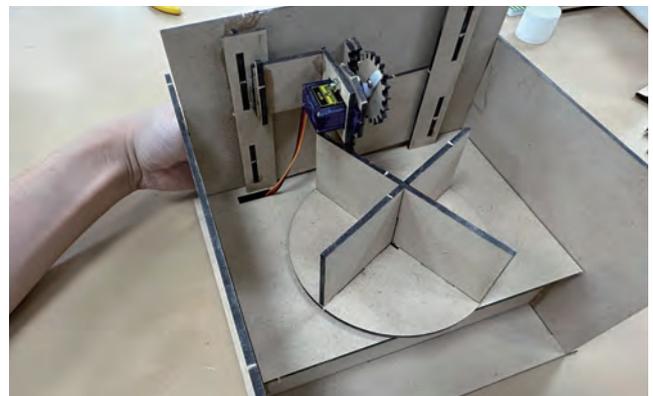
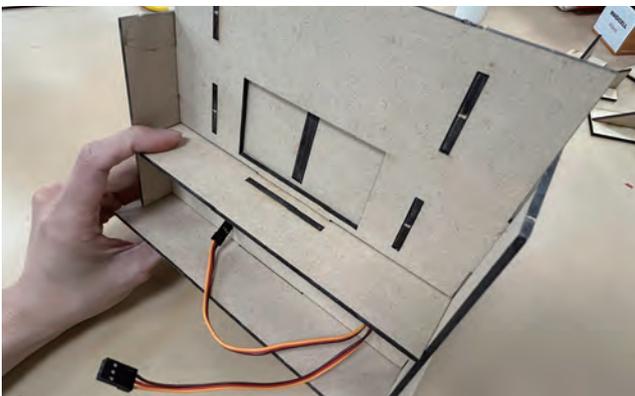
STEP 16 組裝頂板與取餐門，將小伺服馬達線穿過頂板



STEP 17 組裝頂板與底板，並將小伺服馬達線穿出前版的洞

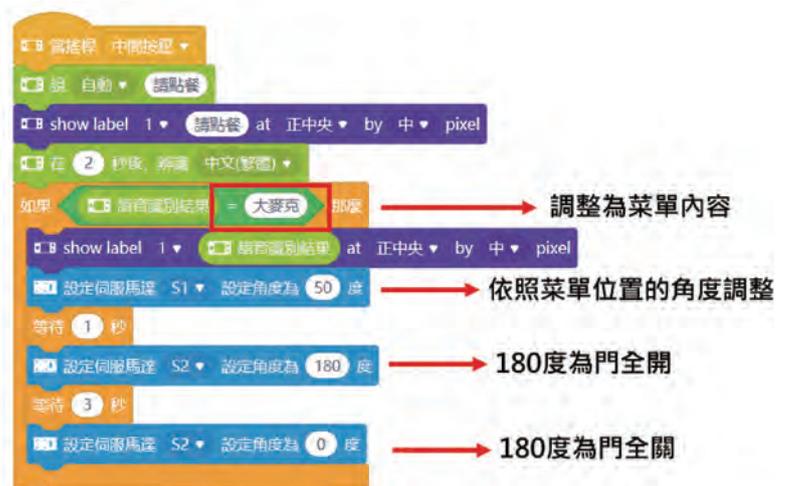


STEP 18 完成組裝



程式編寫

S1 為大伺服馬達，S2 為小伺服馬達，角度可照菜單擺放位置調整。



完整程式範例

```
當 CyberPi 啟動時
說 自動 hello world
連接到 Wi-Fi ssid 密碼 password
等待直到 網路已經連線?
顯示 [LEDs]
show label 1 wifi connected at 正中央 by 中 pixel
設定伺服馬達 全部 設定角度為 0 度
```

```
當按鈕 中間按壓
說 自動 請點餐
show label 1 請點餐 at 正中央 by 中 pixel
在 2 秒後, 辨識 中文(繁體)
如果 語音識別結果 = 大麥克. 那麼
  show label 1 語音識別結果 at 正中央 by 中 pixel
  設定伺服馬達 S1 設定角度為 50 度
  等待 1 秒
  設定伺服馬達 S2 設定角度為 180 度
  等待 3 秒
  設定伺服馬達 S2 設定角度為 0 度
如果 語音識別結果 = 麥香魚. 那麼
  show label 1 語音識別結果 at 正中央 by 中 pixel
  設定伺服馬達 S1 設定角度為 0 度
  等待 1 秒
  設定伺服馬達 S2 設定角度為 180 度
  等待 3 秒
  設定伺服馬達 S2 設定角度為 0 度
如果 語音識別結果 = 鮮蝦堡. 那麼
  show label 1 語音識別結果 at 正中央 by 中 pixel
  設定伺服馬達 S1 設定角度為 100 度
  等待 1 秒
  設定伺服馬達 S2 設定角度為 180 度
  等待 3 秒
  設定伺服馬達 S2 設定角度為 0 度
如果 語音識別結果 = 牛肉堡. 那麼
  show label 1 語音識別結果 at 正中央 by 中 pixel
  設定伺服馬達 S1 設定角度為 140 度
  等待 1 秒
  設定伺服馬達 S2 設定角度為 180 度
  等待 3 秒
  設定伺服馬達 S2 設定角度為 0 度
```

成果



教學資源



專題活動

語音辨識專題

觀察生活周遭並思考語音辨識可以如何應用，使生活更加便利，以此出發點來設計製作一個專題！



單元 3

AI 視覺辨識的智慧新生活

作者：高靖岳—永和國中



體驗活動

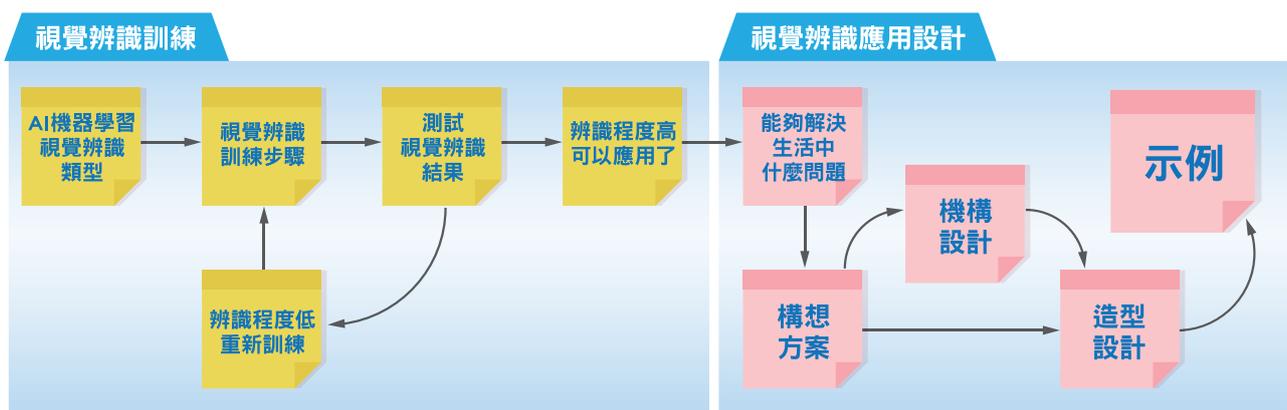
智高 webAI 影像辨識

實作活動：

1. 辨識開關
2. 口罩檢查哨
3. 無人公車
4. 智慧回收分類機



1. 前言



▲圖 3-4 視覺辨識應用設計流程圖

這個單元，我們就利用智高 #1206 積木箱 (AI 人工智慧積木組) 來練習 AI 視覺辨識吧！

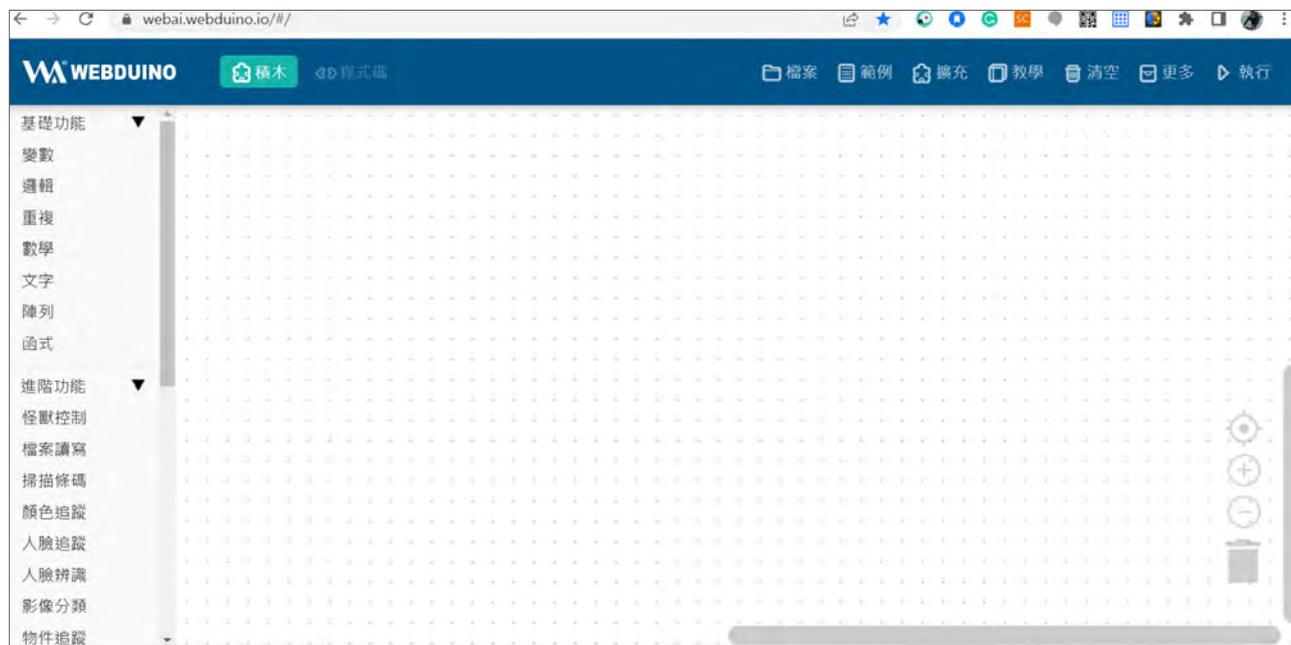
為了讓機器學習辨識，首先，我們需要提供要辨識的類別樣本，作為機器學習的材料，機器會將圖片透過濾鏡的方法將圖片的特徵分辨，變成數字，再透過卷積運算的方式去找到圖片的特徵，從一系列的圖片當中的特徵，給予一個屬性命名，下回機器再次看到這個圖片，就知道它是甚麼名稱，如此一來，我們就能透過程式語言，讓機器在判斷之後，做出對應的自動化行為，接下來我們就來實作練習。

2. 硬體準備

1. 智高 AI 主控盒，請參考積木箱教材手冊說明 p.4。
2. 行星齒輪馬達。
3. 18650 電池座，需自行準備 18650 鋰電池。
4. 積木零件，作為機體支架與連動機構的構成零件。

3. 視覺辨識訓練步驟

STEP 1 使用瀏覽器進入 Web:AI 程式積木頁面，網址：<https://webai.webduino.io/#/>



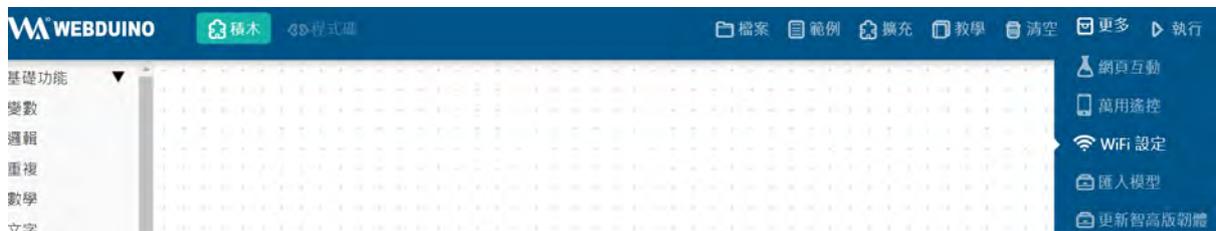
▲圖 3-5 Web:AI 程式積木編輯環境 (模糊)

STEP 2 擴充 :#1206-CN AI 人工智慧積木組



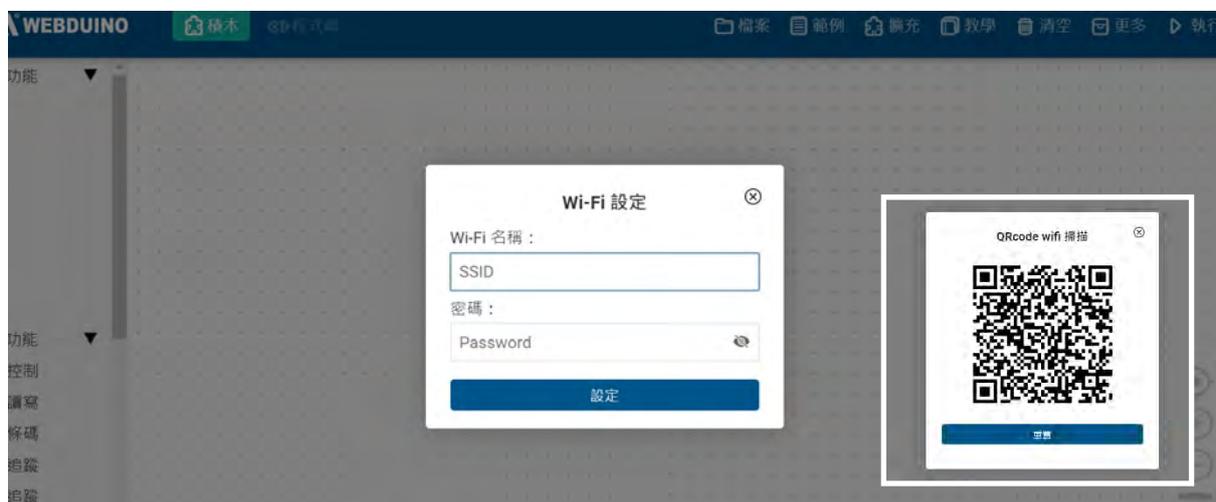
▲圖 3-6 智高人工智慧物件程式積木模組 (模糊)

STEP 3 智高 AI 主控盒 Wi-Fi 設定



▲圖 3-7 設定 Wi-Fi 讓主控盒自動連上網際網路 (模糊)

輸入環境中可用的 Wi-Fi 名稱與密碼，系統會轉換成 QRcode，再以主控盒攝影機掃描紀錄資訊，這樣主控盒可自行連上 Wi-Fi。



▲圖 3-8 手動輸入鄰近的 WiFi AP 的 SSID 和密碼會產生 QRcode, 透過鏡頭掃描輸入資訊

按著智高 AI 主控盒上的 R 按鈕 7 秒鐘，會進入 QRcode 掃描模式。把攝影機鏡頭對準 QRcode，進入 Wi-Fi 設定的自動流程。完成後，主控盒會重新開機，上面就會有 Wi-Fi 的連線資訊，並且請記錄下裝置的辨識碼（device ID）。

請寫下你的裝置辨識碼：_____。請自己留存，不可外流

▼表 3-1 Gigo X AI 安裝版編輯器下載連結



Gigo X AI 安裝版編輯器

從網頁版編輯器的「更多」選項，點選「下載智高安裝」的選項，取得安裝版本的編輯器，在智高提供的雲端空間連結取得範例程式。

<https://www.dropbox.com/sh/mfv7aaj6h0kd0ln/AADqAy6DZK1TJFptXrJVXX3Ra?dl=0>

4. 訓練影像分類

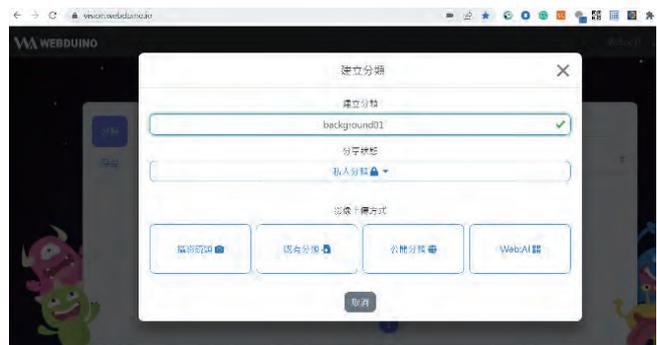
智高 AI 透過慶奇科技 Webduino 影像訓練平臺，所以首先要在訓練平臺完成影像辨識訓練，網址是：<https://account.webduino.io/>。第一次使用要先註冊帳號，或用 Google、Facebook 帳號授權登入訓練平臺。關於影像訓練的教學資料與更新內容說明，慶奇科技也設置說明網頁於 <https://md.webduino.io/s/mopjgVaZU>。影像訓練平臺網址是：<https://vision.webduino.io/>。

影像訓練平臺的操作步驟

STEP 1 左側有兩個標籤，一個是分類，另一個是模型，先建立分類，再從分類中挑選模型。



STEP 2 建立背景，輸入類別名稱，以英數命名，例如：background01，使用攝影鏡頭或使用 AI 主控盒上的攝影機。建立背景也是提供一個類別，作為判斷目前是否有物件或圖片的依據。



STEP 3

擷取 15-30 張照片，作為背景的照片盡量不要動作，或以單色底色較容易辨識物件。按下建立分類。此外，當下的光線也會影響拍照的效果，因此，對於辨識環境的區域範圍也要控制，就像車牌辨識的閘門，也需要在特定的區域。



STEP 4

訓練完成後，就會看到剛剛訓練的分類，在分類上至少要有 2 個以上。
 影像分類：最多 4 個分類物件追蹤：最多 5 個分類



STEP 5

再重複上述的方式，建立分類，右圖，在分類上中有 3 個類別，就可以提供給機器學習分辨，也就是要建立模型。



STEP 6

點選右邊「模型」標籤，再點選新增，就會跳出新增模型的畫面。



STEP 7

模型名稱是以英數命名，建立模型「istape01」，讓機器從圖片中學習，哪一個狀態是有膠帶的分類，哪一個狀態是沒有膠帶的分類。

同學也可以嘗試，剪刀、石頭、布的辨別分類。



STEP 8

挑選2個分類，這個練習，挑選背景的類別，名稱是 background01 和膠帶的類別，名稱是 tape。勾選之後，建立模型。

訓練完畢之後，再測試模型，有時候是因為類別圖片的相似度高，需要重新拍攝，取得明確辨別物件的特徵。



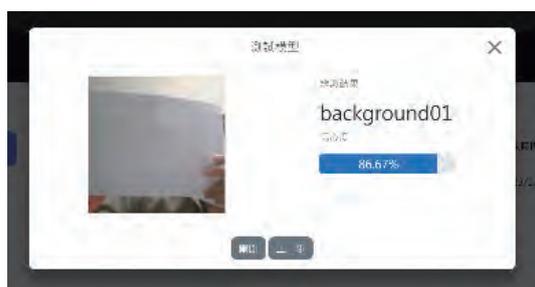
STEP 9

訓練好的模型先測試看看，辨識度若明確，就可以丟給機器直接應用，或者在機器上測試也可以，請記得，主控盒的鏡頭環境條件，要控制跟拍攝鏡頭的環境相同。



STEP 10

測試時，會顯示辨別的信心度，再撰寫自動控制程式的時候，利用信心度來做動作判別。



STEP 11

由於機器是從 2 個類別的特徵來分類判斷，所以人臉在這個分類裡頭，機器依據這 2 個類別的特徵原則，就會判定人臉是 tape，因此，機器學習的原則是，給機器甚麼資料，它就學習甚麼資料，當資料很相似，有可能會誤判或是信心度較低，這時候就是要回到資料的特徵來調整。或是多提供資料給機器學習。



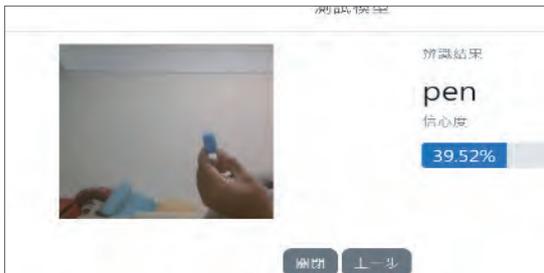
STEP 12

建立新的模型，這次選擇空白背景、原子筆、膠帶。訓練後，請同學再測試看看，機器是不是能分辨你手中的筆或是膠帶。



STEP 13

辨視的角度若與其他物件相似，或是受到背景干擾，所辨別的結果與信心度就會比較低，或是判斷有誤。



STEP 14

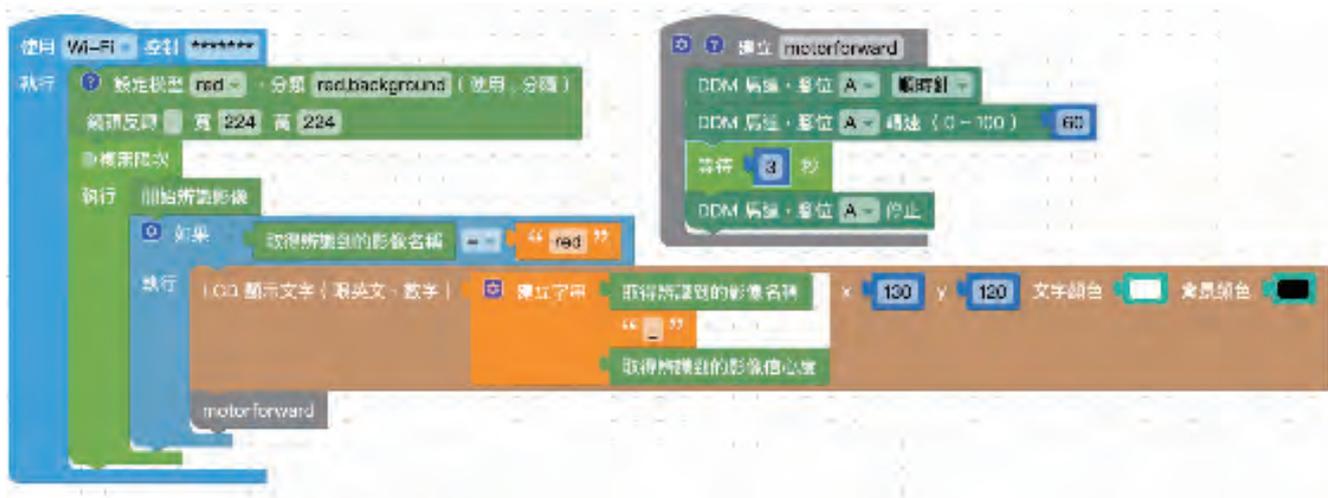
因此，進行辨視的時候，要減少背景的干擾，提供的影像需要多一些角度特徵，這樣機器學習透過影像資料的特徵經過演算處理產生一個判斷的神經元，這個神經元當接觸到背景、筆、膠帶，就會分辨背景、原子筆和膠帶，但它是靠運算的結果，所以提供參考的機率，若要讓機器更精準分辨原子筆、鋼筆、鉛筆，那就得提供三種筆的特徵，或是侷限辨視的特徵，例如針對筆頭的特徵來訓練機器分辨。



STEP 15

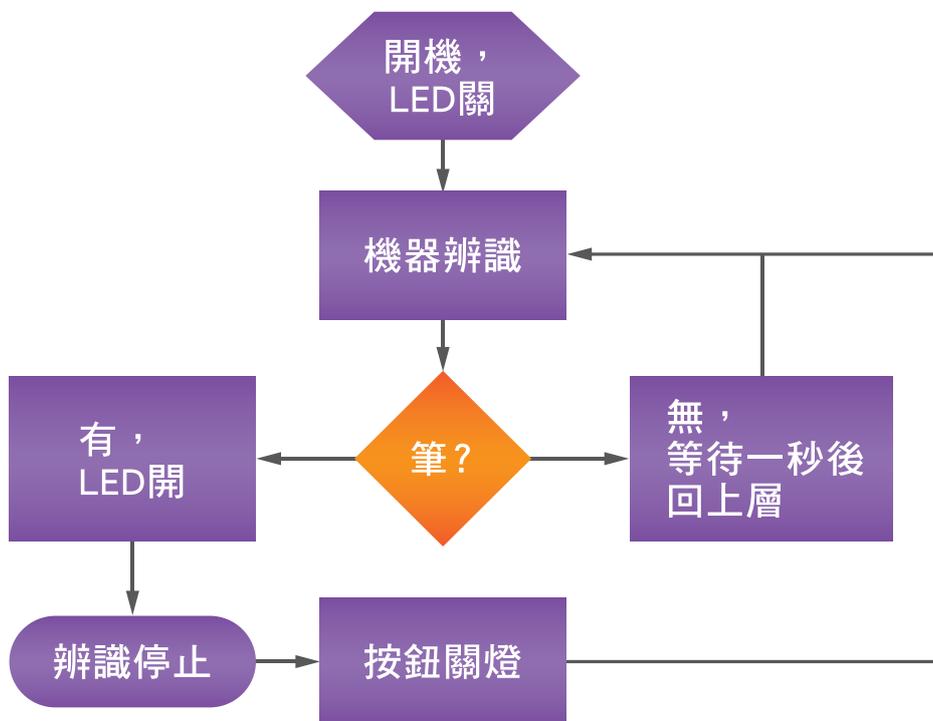
辨識訓練完成之後，在 Web:AI 程式積木平臺使用訓練完的模型，在自動控制的程式設計中加入視覺辨識的功能，讓機器辨識之後，做出反應。

<https://webai.webduino.io/#/>



學習任務

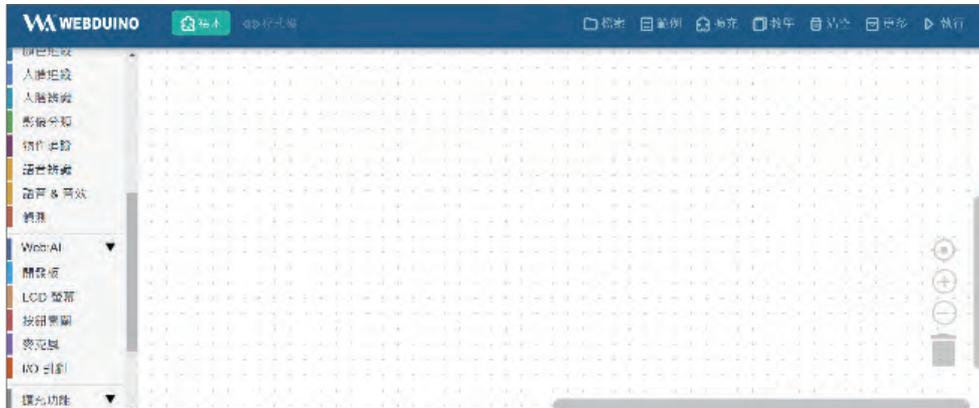
任務一，看到筆，打開燈。



▲圖 3-9 視覺辨識自動感應流程圖

操作步驟

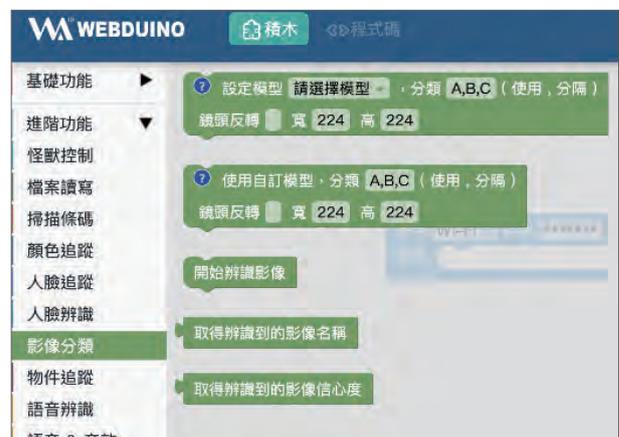
STEP 1 Web:AI 積木平臺 <https://webai.webduino.io/#/>



STEP 2 左側點 Web:AI 開發板，積木拖曳出來，輸入開發板的識別碼。



STEP 3 在進階功能中有影像分類。



STEP 4 將第一個積木拉到主控板積木，然後選擇辨識模型，此時後面的分類就會跟著顯示，由於辨識訓練的系統與積木的系統都用帳號連動，在載入模型方面是很方便的操作。



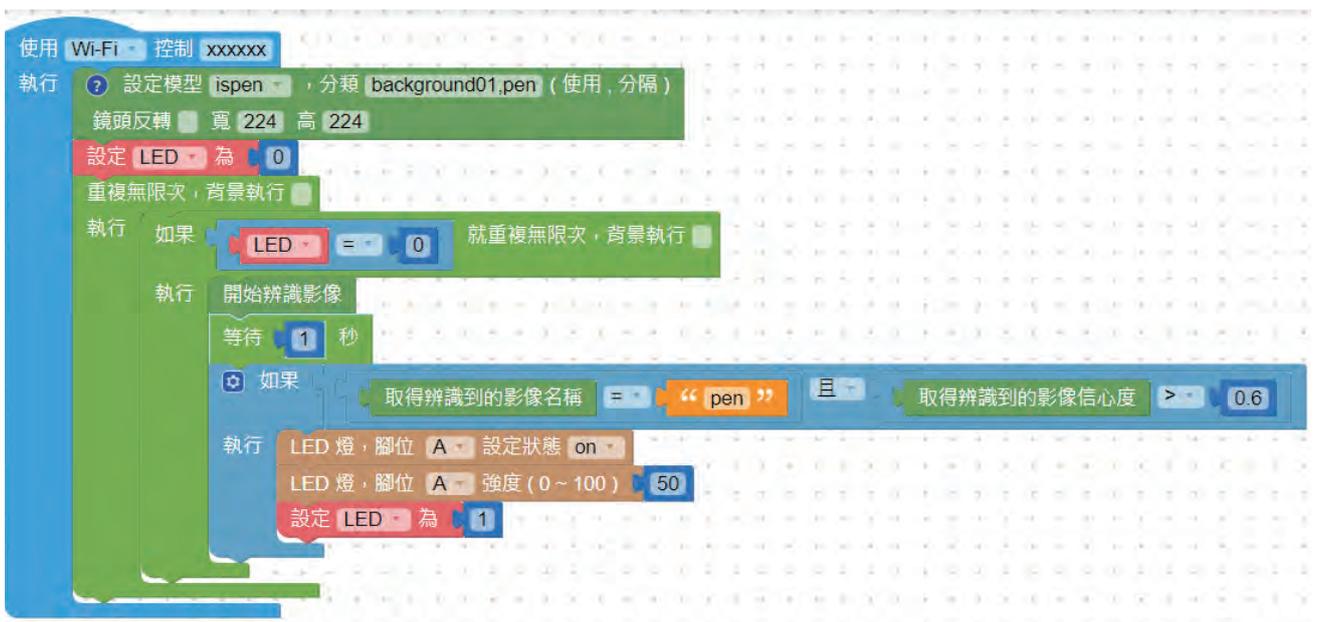
STEP 5 建立紀錄 LED 狀態的變數，預設值為 0，利用這個變數的狀態來控制 LED 的啟動與關閉。



STEP 6 從影像分類當中把開始影像辨識放到重複結構，在 LED=0 的條件下，執行開始辨識，每隔 1 秒鐘偵測一次。



STEP 7 在重複結構內，再做條件判斷，辨識結果是筆且信心度 60% 以上，LED 變數改變為 1，燈打開。
請注意，辨識環境需具備基本照明，否則太暗的環境機器也無法辨識。接著使用智高模組積木。



STEP 8

智高積木模組讓電路作業簡化，在主控板後面有 8 個腳位，以英文字母大寫命名，將智高積木 LED、按鈕開關、紅外線感測器、減速直流馬達、伺服馬達的接線作業簡化，不擔心接錯線燒毀，在這個練習作業我們用到 LED 燈和按鈕開關兩個積木，LED 燈設置在 A 腳位，按鈕開關設置在 B 腳位。



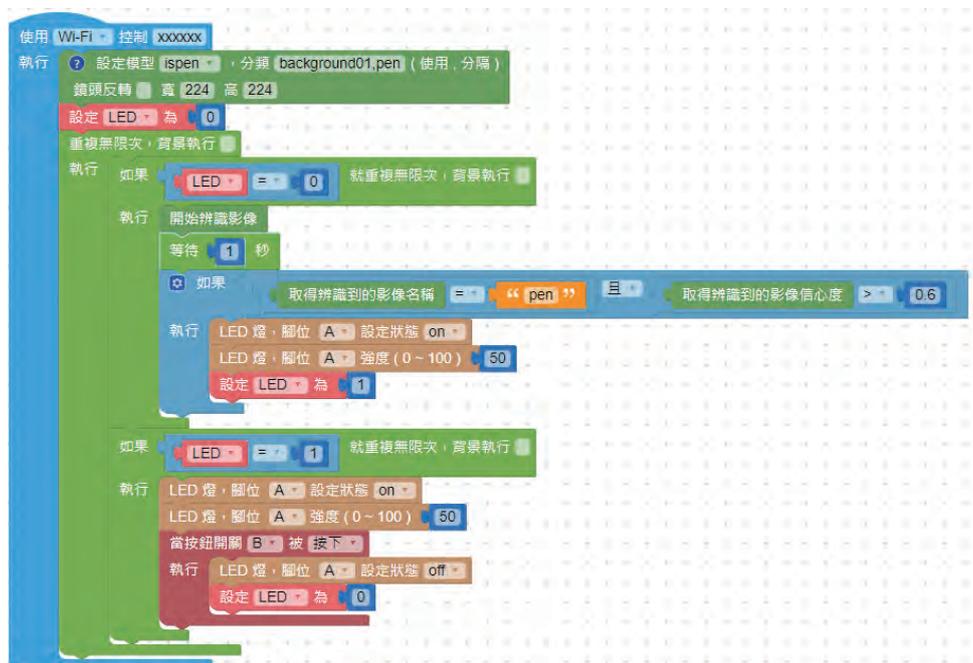
STEP 9

當燈已經打開，就不需要再辨識了，因為鏡頭也是很耗電，所以利用 LED 變數來離開重複結構，但燈還是要用，所以另外用一個重複結構讓燈維持，直到按鈕開關按下，才把燈關閉，然後再回到辨識準備狀態。



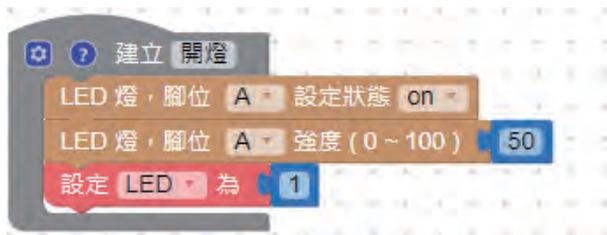
STEP 10

兩個條件重複結構，裝在開機就啟動的無限重複的結構中，換句話說，打開了主控板，就可以用筆來啟動電燈，再用按鈕關閉電燈，事實上，照明燈的啟動關閉有很多種方法，視覺辨識與程控板能發揮得更多，同學發揮創意再想想看！



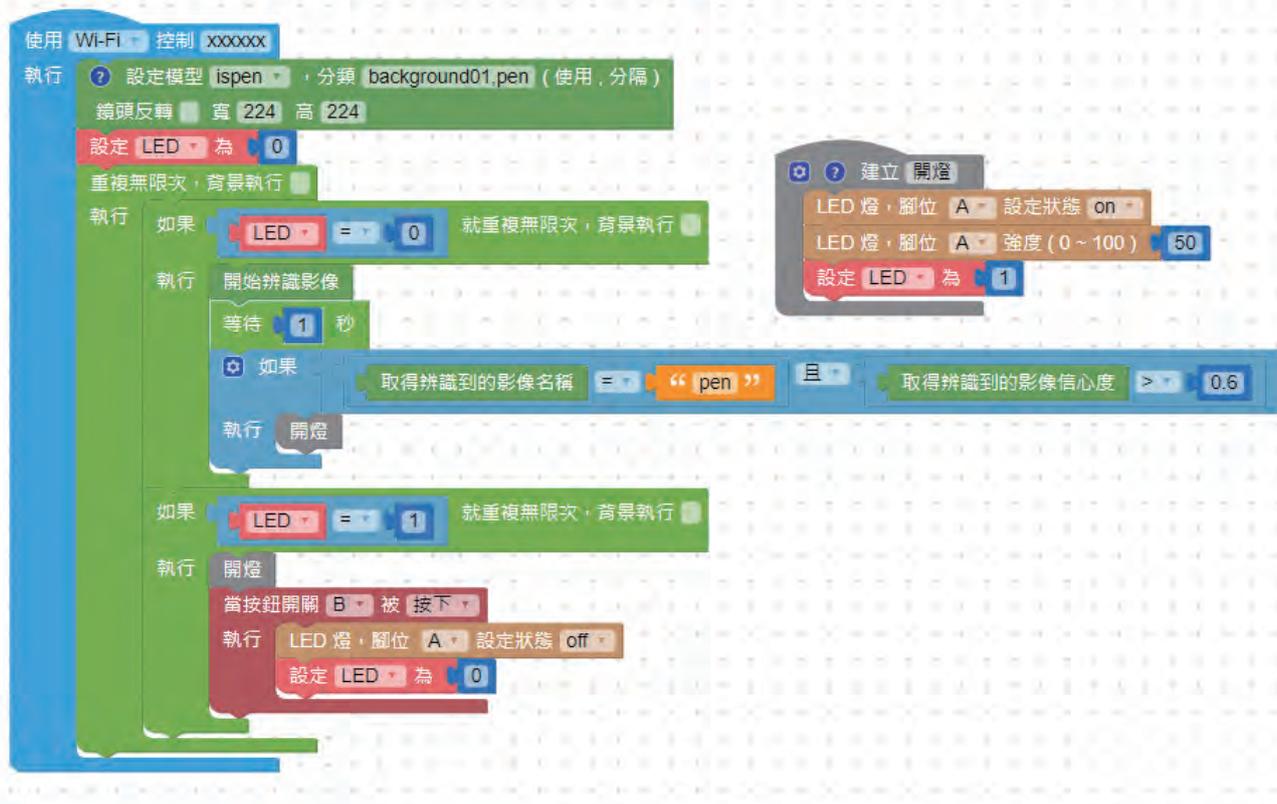
STEP 11

另外，運用函式積木，來簡化程式編輯。這個技巧在大的專題中很好用，例如，從這邊建立不同強度的燈，然後只要呼叫函式。



STEP 12

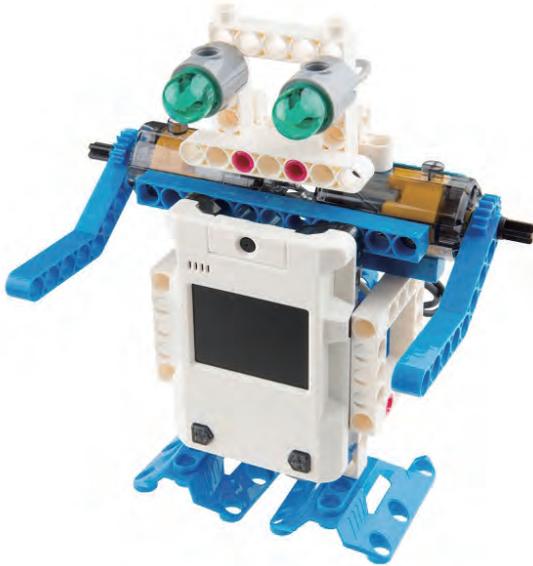
能簡化程式敘述，讓主程式的流程更接近流程圖。這樣就可以把程式碼燒錄在主控板上，然後再將 LED 插在主控板 A 腳位，按鈕開關插在 B 腳位，當攝影機辨識到筆出現，燈就會亮，然後按下按鈕，燈滅，再次重新辨識。



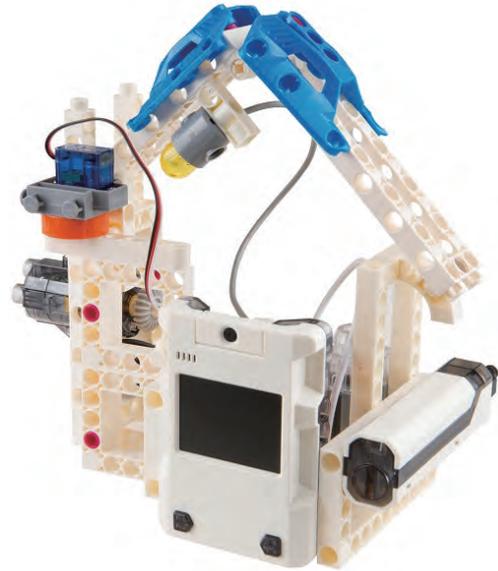
參考範例

積木結構參考

1. 智高課程手冊第一課：打招呼機器人
2. 智高課程手冊第十七課：物聯網智能家居



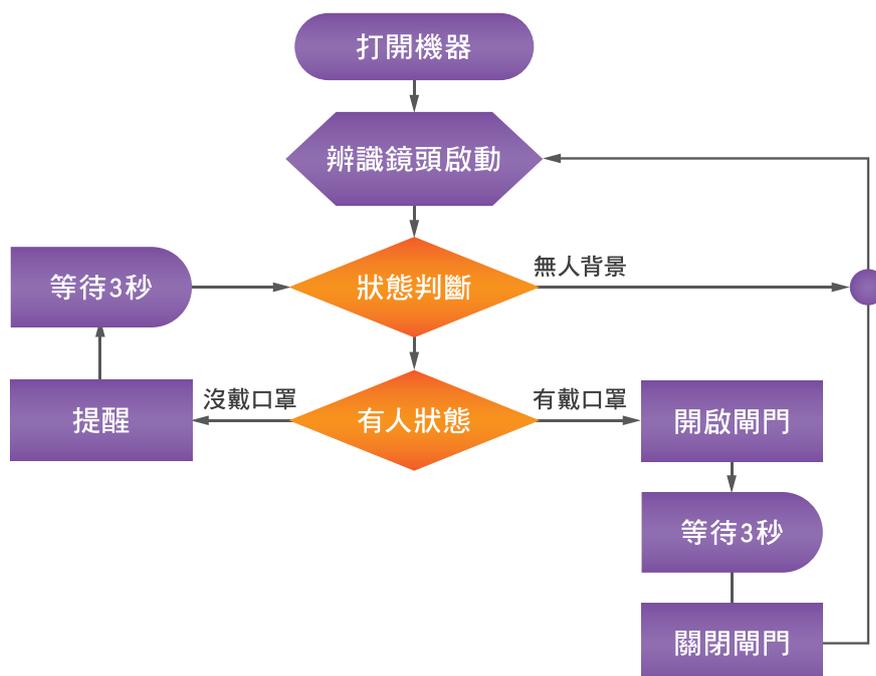
▲圖 3-10 打招呼機器人



▲圖 3-11 物聯網智能家居

學習任務

任務二，刷臉提示，你有沒有戴口罩。



▲圖 3-12 戴口罩檢查自動閘門流程圖

設計操作步驟

STEP 1 可參考智高積木 #1206 組的範例 8 組合運作機構，連上平臺進行類別訓練，建立辨識背景、有戴口罩、沒戴口罩三個類別。

STEP 2 將分類模型下載到主控板

STEP 3 連上積木程式平臺撰寫辨別程式，下載到主控板。

STEP 4 組裝積木，組裝一個由直流馬達驅動的自動門，並建立開門與關門的函式

STEP 5 將積木自動門與主控版連接測試結果

視覺辨識討論

請見表 3-5 欺騙機器的案例。

▼表 3-5 欺騙機器的案例

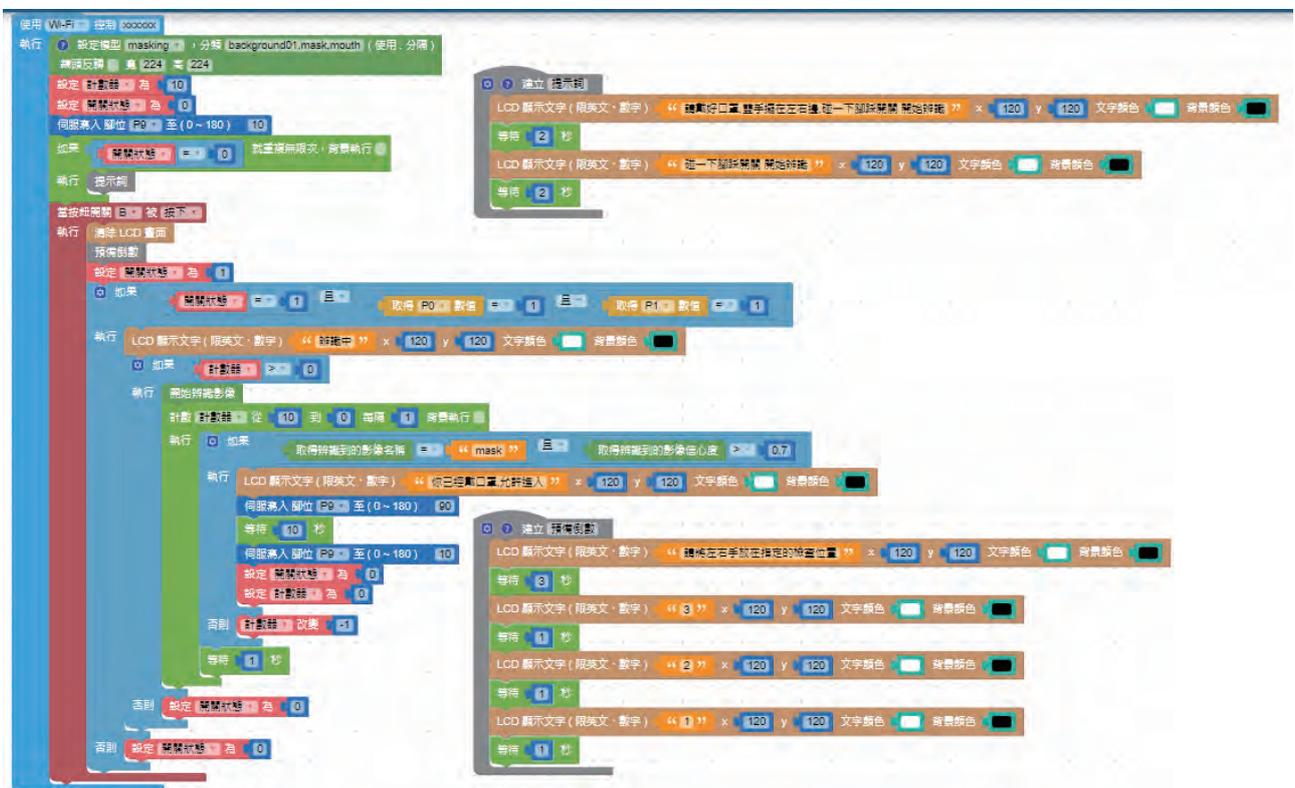
		
<p>依照之前的方法訓練辨識，讓機器學習如何分辨有戴口罩跟沒戴口罩，但只有拿著口罩，似乎無法確定有沒有戴在臉上？</p>		
		
<p>真神奇！迪士尼公主也能辨識戴口罩，拿膠帶遮著一部分畫面，好像也能過關？</p>		

機器的運作是依照規則與邏輯，定義範圍內的狀態都能被接受，因此，定義需要更多條件，因此，可加入其他的感測器元件，限制辨識的範圍與確認是真人在進行辨識，加入紅外線感測器、按鈕開關等，增加機器辨識的真實性。例如：進到辨識區，按下按鈕開關開始計時 30 秒，左手右手需要同時遮斷紅外線感測器，臉部朝向主控板攝影機，進行辨識，辨識結果為口罩狀態，才打開閘門。接下來，組裝原型機來測試這樣的想法是否可行。（也可以直接使用範例修改喔！範例中的積木已經是做過機器訓練，可以直接應用。）



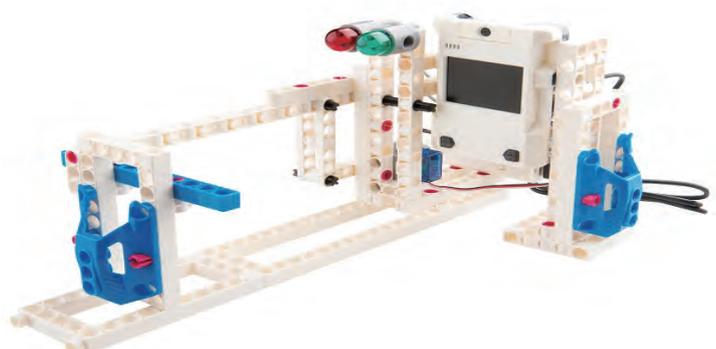
▲圖 3-13 已經訓練好的辨識模型，可以直接使用

硬體：AI 主控板、伺服馬達 x 1、智高 #1206 範例 8 零件、紅外線感測器 x 2
程式碼：



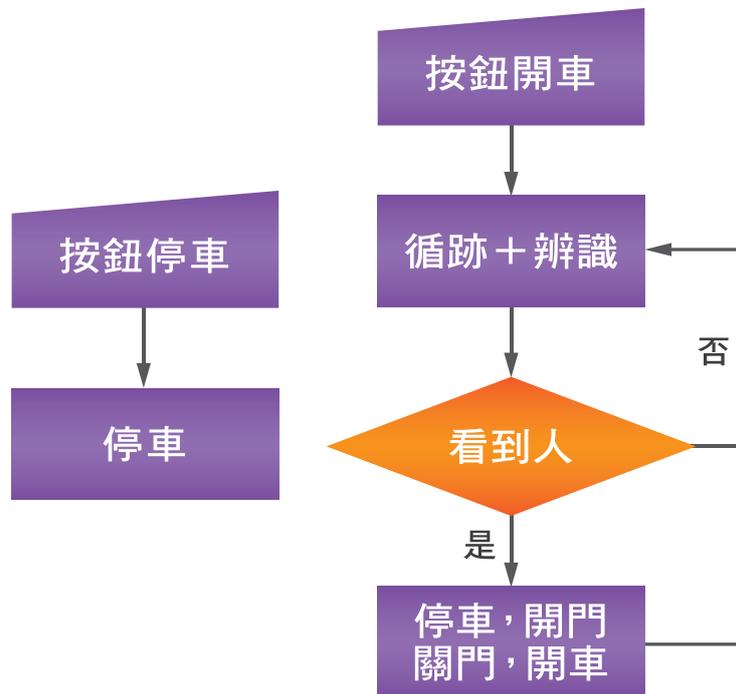
▲圖 3-14 口罩辨識閘門程式碼

組裝參考：
智高 #1206 範例 8
人臉識別門禁



▲圖 3-15 人臉識別門禁

任務三：循跡無人車，招手車停下來載人，紅燈停車



▲圖 3-16 循跡無人車流程圖

設計操作步驟

- STEP 1** 可參考智高積木 #1206 組的範例 3 和 18 組合車輛，但要確保鏡頭的視野在前進的方向。
- STEP 2** 在訓練平臺訓練辨識紅燈與行人，可利用智高與 web:AI 的小怪獸圖卡當作行人，另外以紅色圓形作為紅燈訓練。把分類模型下載到主控板。訓練的背景要盡量符合當時的空間。
- STEP 3** 在積木程式平臺撰寫辨識與馬達控制程式。將程式下載到主控板。
- STEP 4** 馬達前進的轉速不需要太快，先確認能夠有足夠的時間辨認與反應。
- STEP 5** 拿出紅色圓形圖卡測試車輛是否自動停止。並確認是否訓練的模型與程式有達到預期的效果，有必要就要回到上面的步驟修正。
- STEP 6** 拿出小怪獸圖卡測試車輛是否停止。並確認是否訓練的模型與程式有達到預期的效果，有必要就要回到上面的步驟修正。

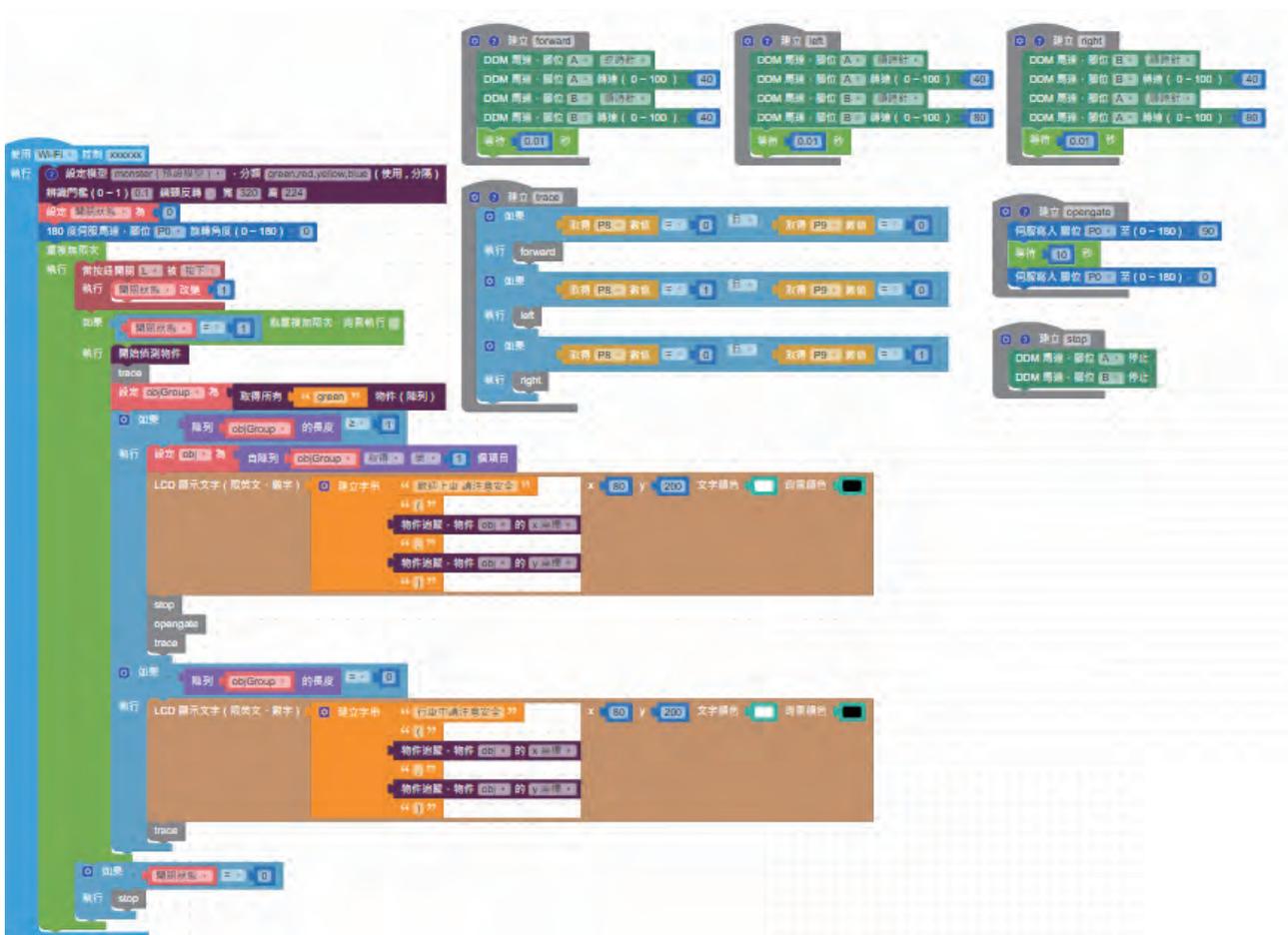
視覺辨識討論

車子要一邊走一邊察覺有沒有人招手攔車，這時候的視覺辨識，並非靜止的畫面，此時，要使用物件追蹤的方法來，讓機器覺察這個物件。視覺辨識，是針對攝影機的全螢幕所看到的畫面來識別，例如：辨識車牌，知道是哪一臺車進出停車場；而物件追蹤，會判斷畫面中的物件，物件在移動畫面中，機器透過攝影鏡頭擷取出物件的特徵值，因為物件遠近畫面的大小也會不同，但是機器學習的卷積神經網路，會把畫面中的物件做等比例分割，然後取分割區塊的平均值，得到物件的特徵，畫面中就會圈選出物件的位置與大小，某些攝影機保全系統，如果偵測到人影，就會發出警報，也是這種方法，但如果是貓跟狗就不會觸動警報，因為在特徵上，人類的特徵就跟貓與狗不一樣，但如果是黑猩猩，就可能觸發警報。物件追蹤大的特徵，例如：顏色、形狀、動作等。例如：在走廊上的攝影機偵測到有奔跑的動作，就會提醒安全人員注意發生了甚麼事情。這個學習作業，結合循跡車與物件追蹤，做一臺循跡無人車。

硬體：主控板、減速馬達 x 2、紅外線感應器、伺服馬達、按鈕開關

程式碼：

這段程式碼結合了循跡自走車以及辨識，再辨識物件上，直接使用訓練好的小怪獸物件追蹤角色作為行人角色，同學們也能到影像訓練平臺自己訓練。



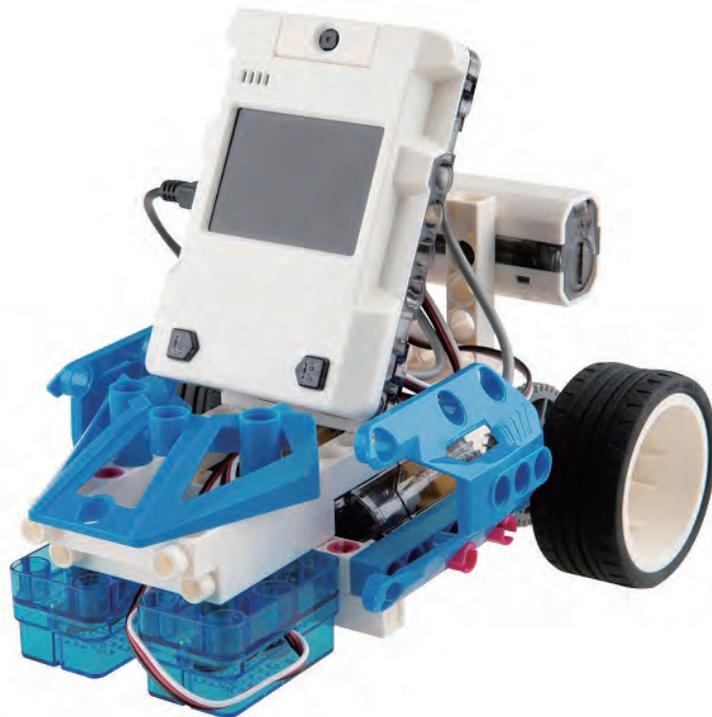
▲圖 3-17 循跡無人車程式碼

組裝參考

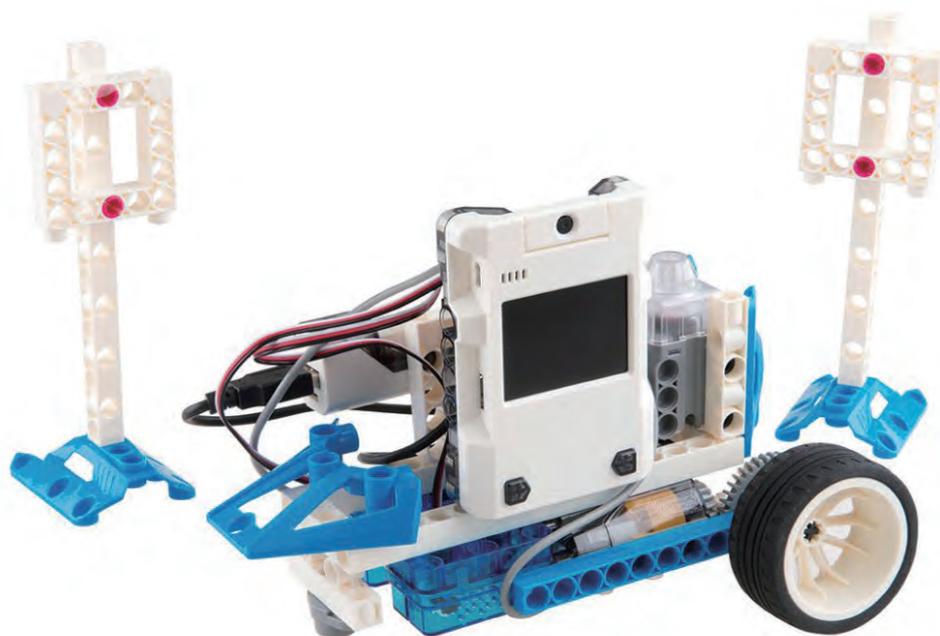
智高 #1206

範例 3 循跡無人車

範例 18 無人公車



▲圖 3-18 循跡無人車



▲圖 3-19 無人公車

任務四：紙杯、鐵鋁罐、鋁箔包、利樂包、寶特瓶分類器



▲圖 3-20 分類流程圖

設計操作步驟

- STEP 1** 可參考使用智高積木 #1206 組的範例 14，要確保鏡頭辨識的視野。
- STEP 2** 設計動作階段：1. 開啟丟入口閘門，鏡頭開始辨識 2. 置放到回收簍 關閉閘門 3. 判斷物件 4. 將回收簍轉到對應的回收桶 5. 打開回收簍的閘門 6. 搖晃回收簍使回收物自然掉落回收桶 7. 關閉回收簍閘門 8. 回到原位歸位，打開丟入口閘門。
- STEP 3** 依據動作選擇動作零件與機構，可由智高積木其他範例綜合改良，例如定位轉向、兩個閘門口都使用伺服馬達。
- STEP 4** 依現有的回收桶，設計回收簍，使用塑膠積木快速結構，以及打開閘門口的機構。
- STEP 5** 設置智慧分類回收簍，考量電力與硬體配置方式。
- STEP 6** 依條件訓練辨識模型，若桶內的光線不足以辨識，由主控板，再增加輔助照明，或紅外線感測器做非接觸式開關。
- STEP 7** 撰寫程式，並依照反應狀況調整程式內容與步驟。
- STEP 8** 完成原型，繼續實驗與測試修正，提出最終方案。

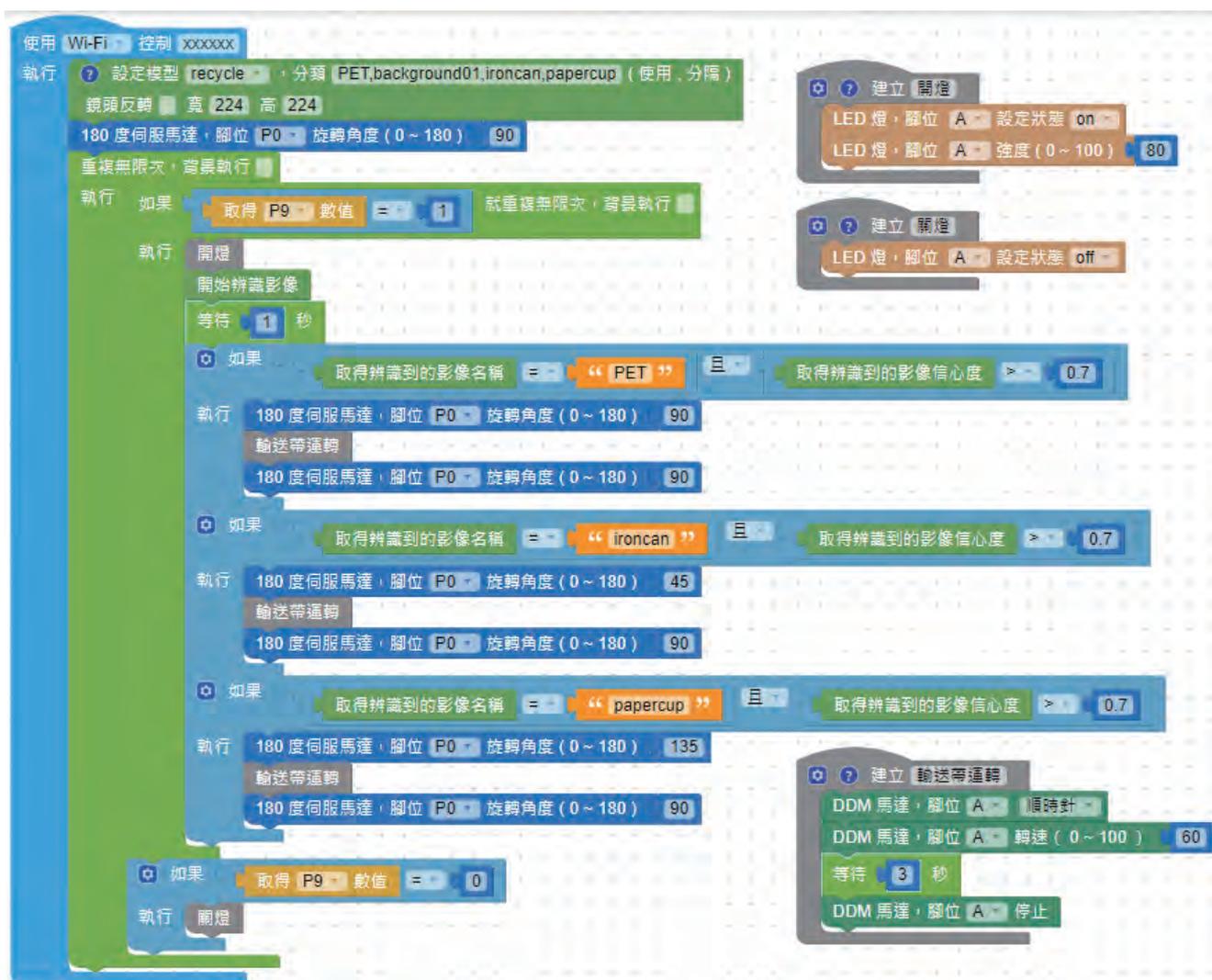
影像辨識討論

若資源回收桶是密閉的，那辨識的光線肯定不足夠，因此要啟動自動辨識之前也需要設置輔助照明，其次，回收物的包裝繁多，在訓練上，可多找一些不同包裝，但是同樣外型的容器來訓練，這樣子較容易讓機器辨認出容器的外型，在訓練辨識之前，要先確認攝影機的視野並且讓辨識物的背景盡量是白底，要有擺正的與擺反的照片，這樣子的辨識訓練結果會比較精確。

由於各類容器有大有小，應限制容器體積、高度、直徑或是對角線長度，在專題練習上的複雜度比較低。若辨識的精確度不高，也可以增加其他感測器，去增加判斷的準確度。在回收產業上，有在輸送帶上裝設辨識鏡頭，利用機器手臂挑出回收物的應用，各位同學也可以思考看看怎麼設計，能讓回收作業自動化更便利。

硬體：主控板、角度伺服馬達、減速直流馬達、紅外線感測器

程式碼：

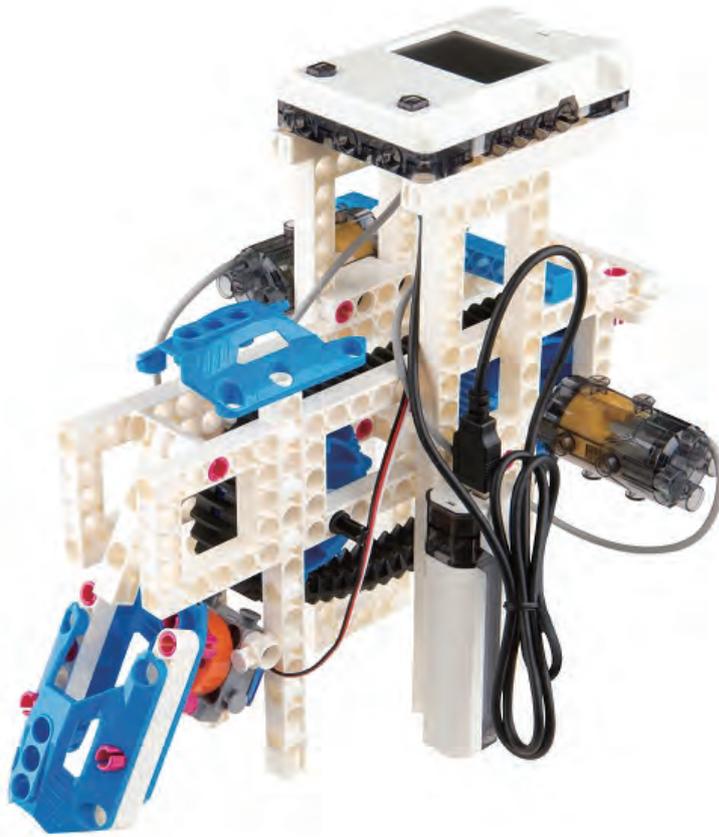


▲圖 3-21 分類器程式碼

組裝參考

智高 #1206

範例 14 資源回收裝置



▲圖 3-22 資源回收裝置

參考資料：

1. AI 人工智慧積木組，智高實業股份有限公司，<https://www.gigotoys.com/products/1206-tw.html>。
2. 影像辨識，慶奇科技股份有限公司，<https://md.webduino.io/s/mopjgVaZU>。



單元 4

生活中的影像辨識

作者：蔡釋鋒—樟樹實中



體驗活動 1

Google 即時拍照翻譯



▲操作實例



體驗活動 2

Google Lens 智慧鏡頭



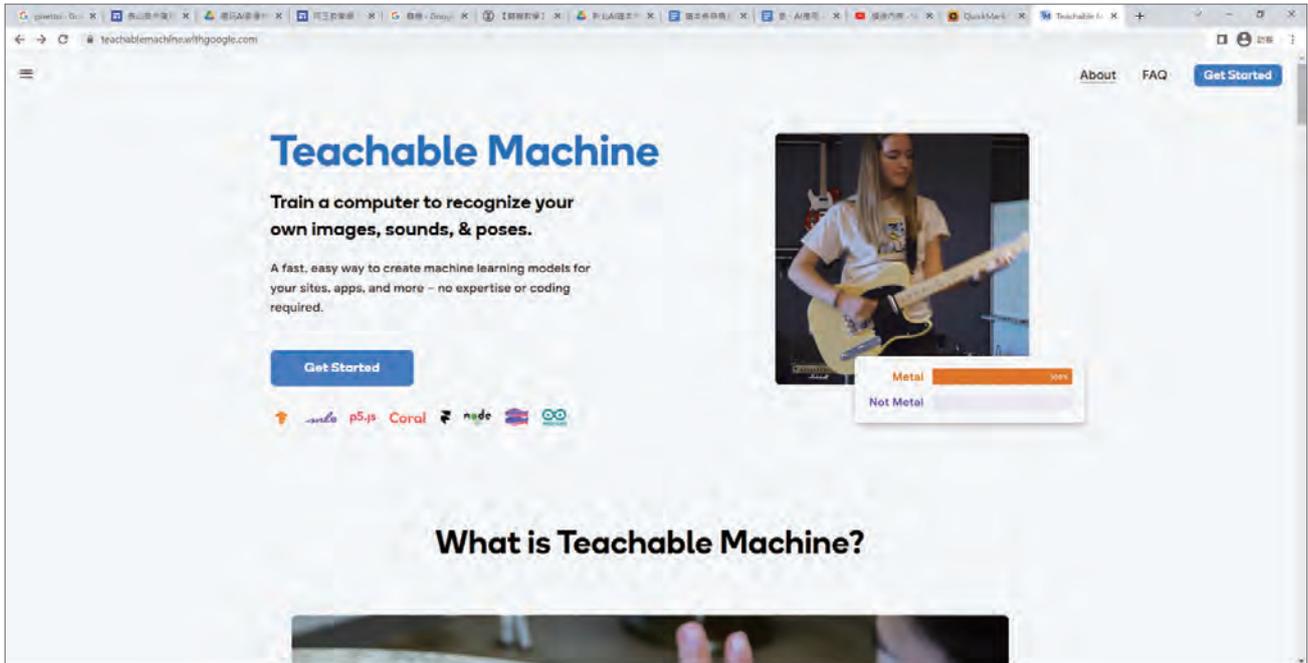
▲操作實例





體驗活動 3

Teachable Machine 影像辨識



議題：AI 人工智慧真的有這麼聰明嗎？

可先以單筆 A、B 資料作為辨識測試成果，失敗後帶出處理大數據與處理一般少量資料的差異。

活動準備

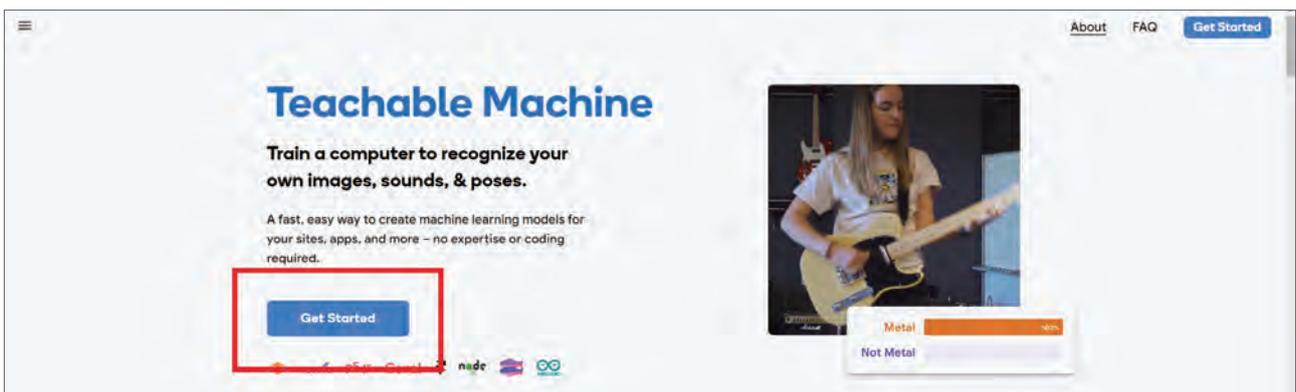
1. 蒐集素材：請同學分別在白紙上寫上英文字 A、B。
2. 視訊鏡頭。



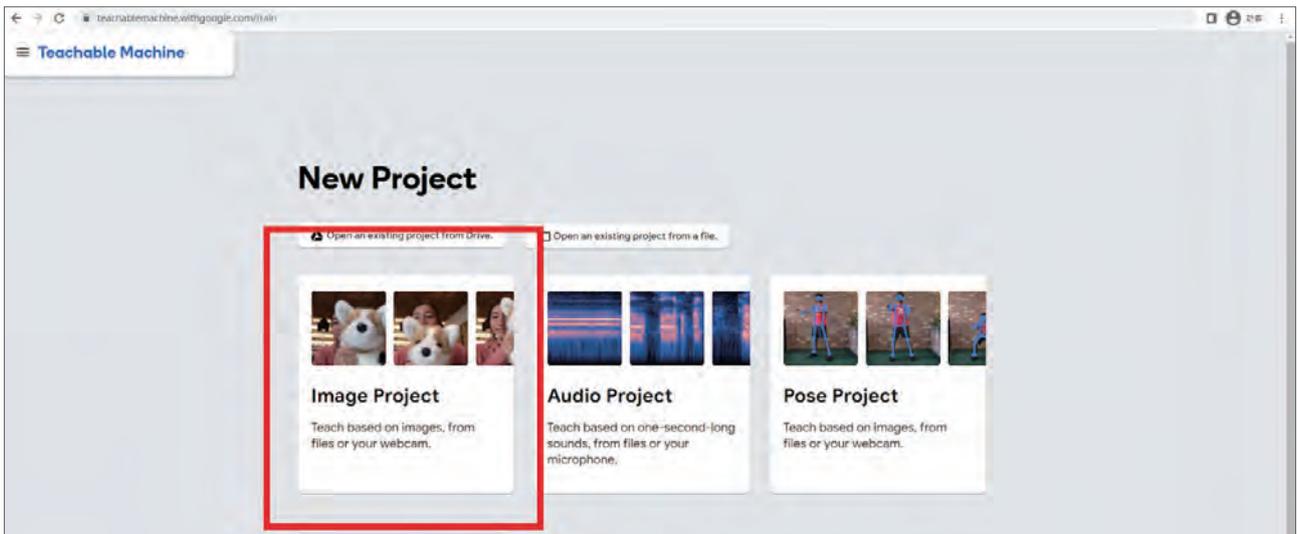
活動進行

STEP 1

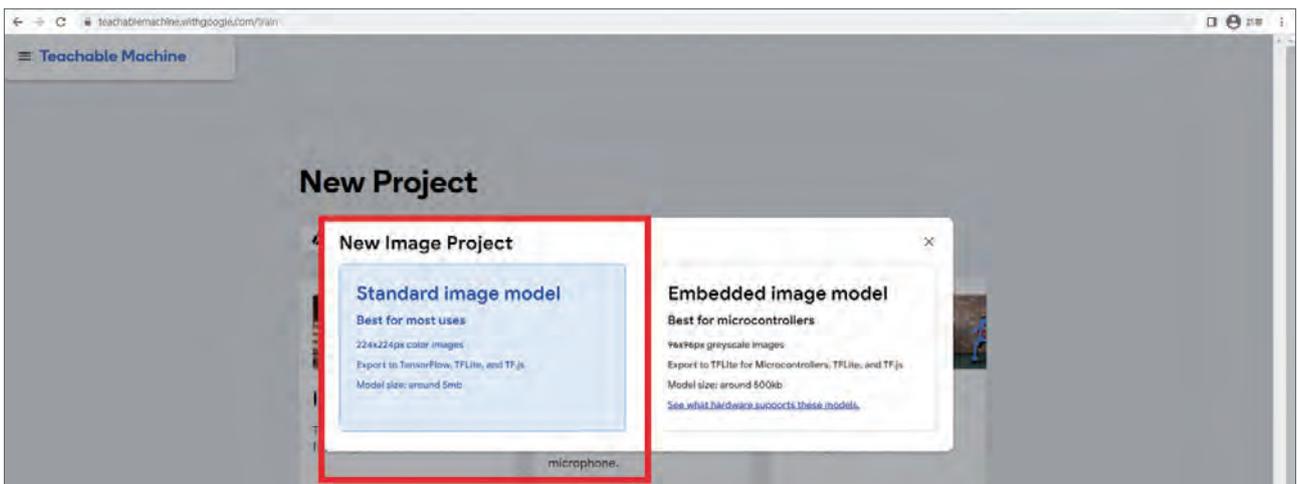
點選 Get Started



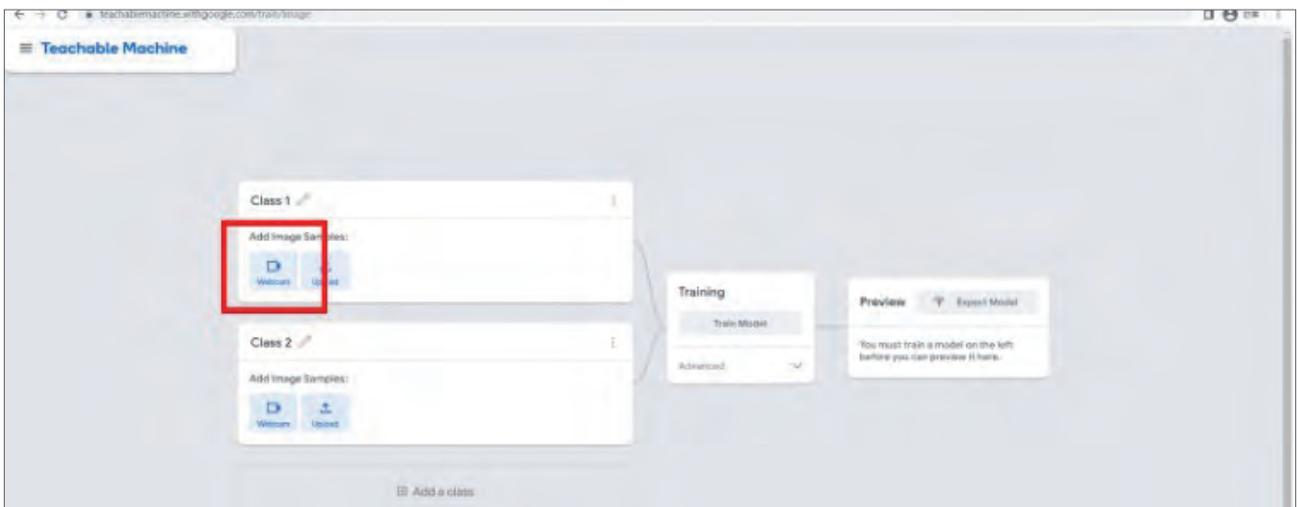
STEP 2 點選 Image Project



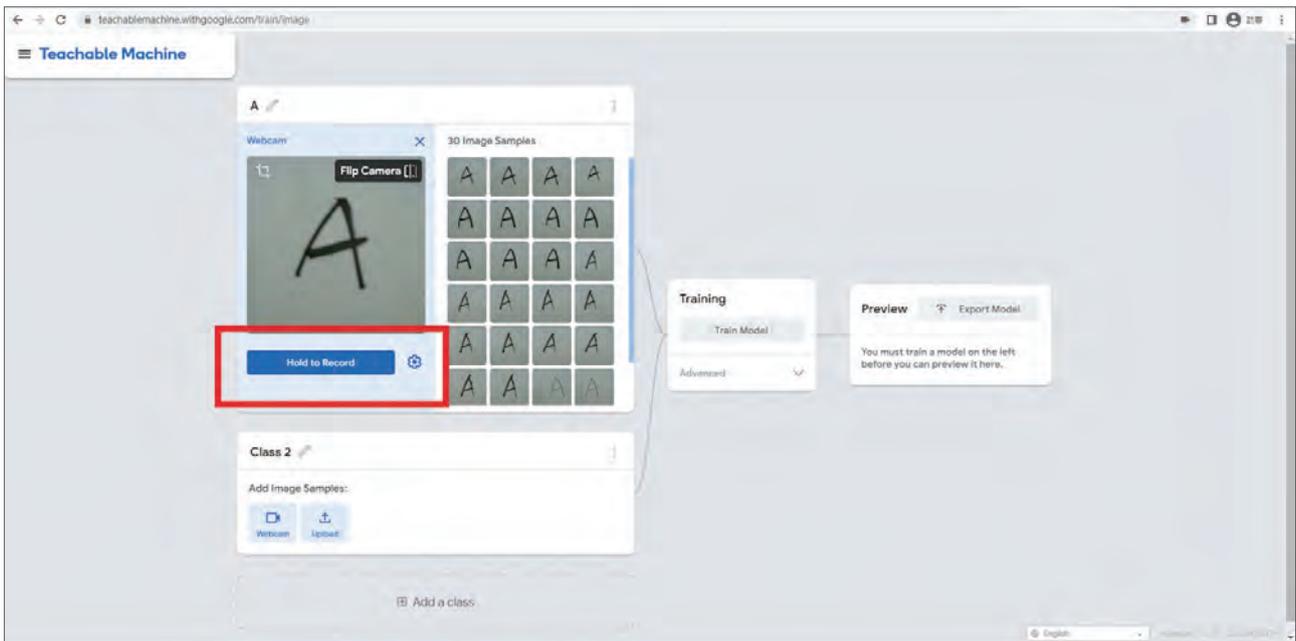
STEP 3 點選 Standard image model



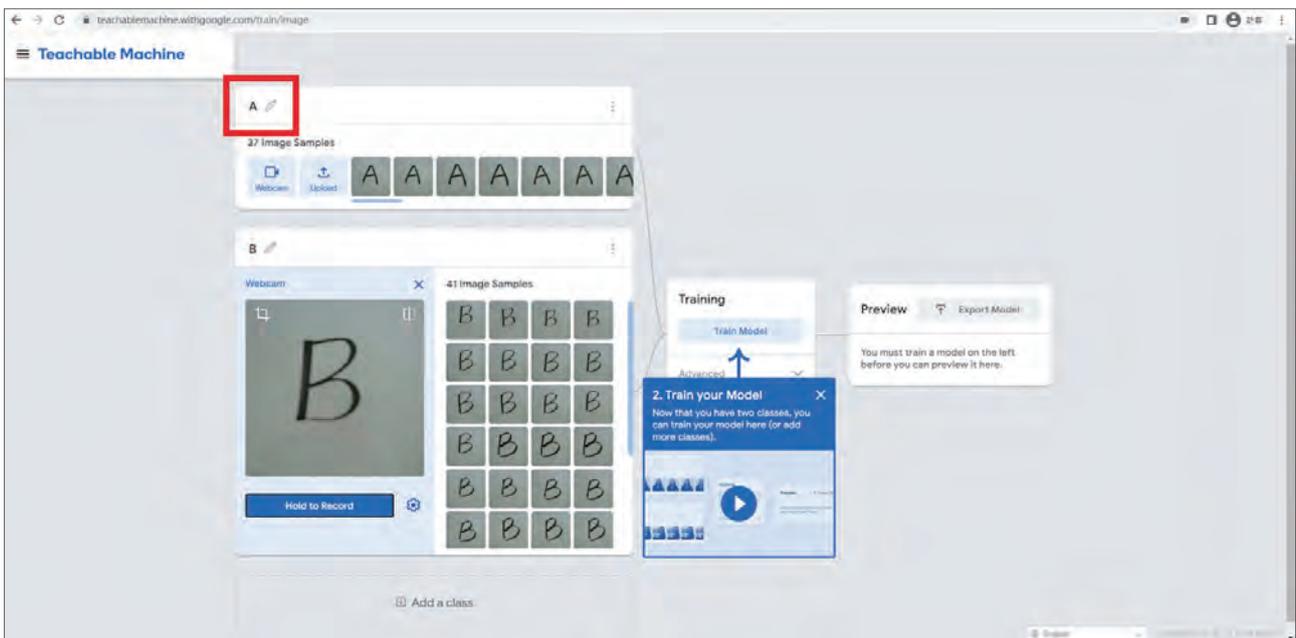
STEP 4 點選 Webcam



STEP 5 點選 Hold to Record3-33.tif



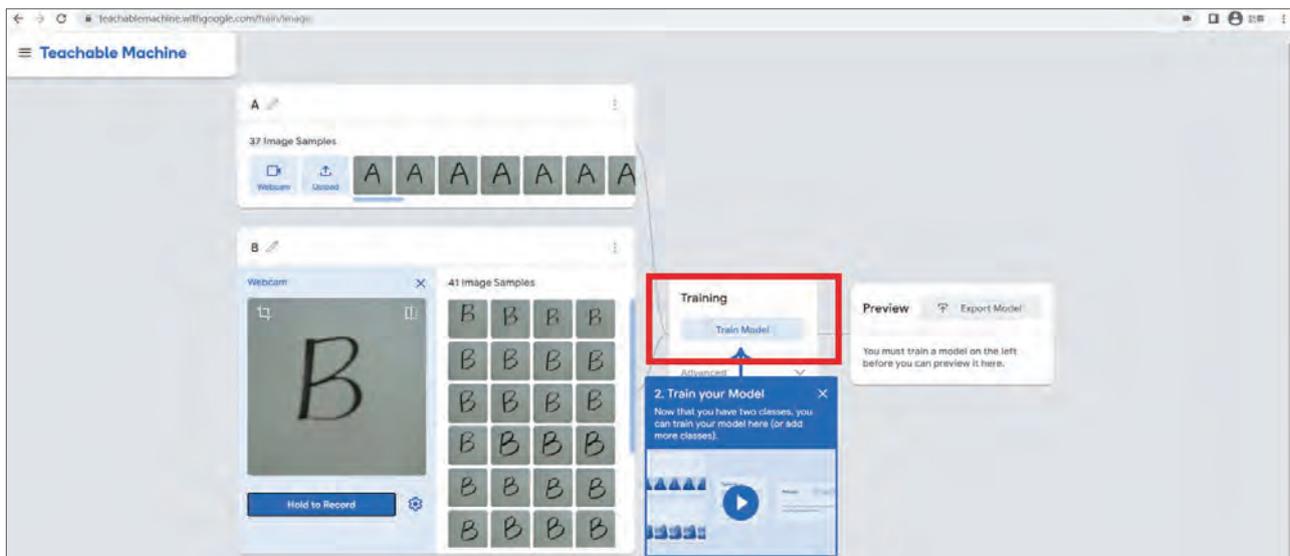
STEP 6 可將類別自行命名



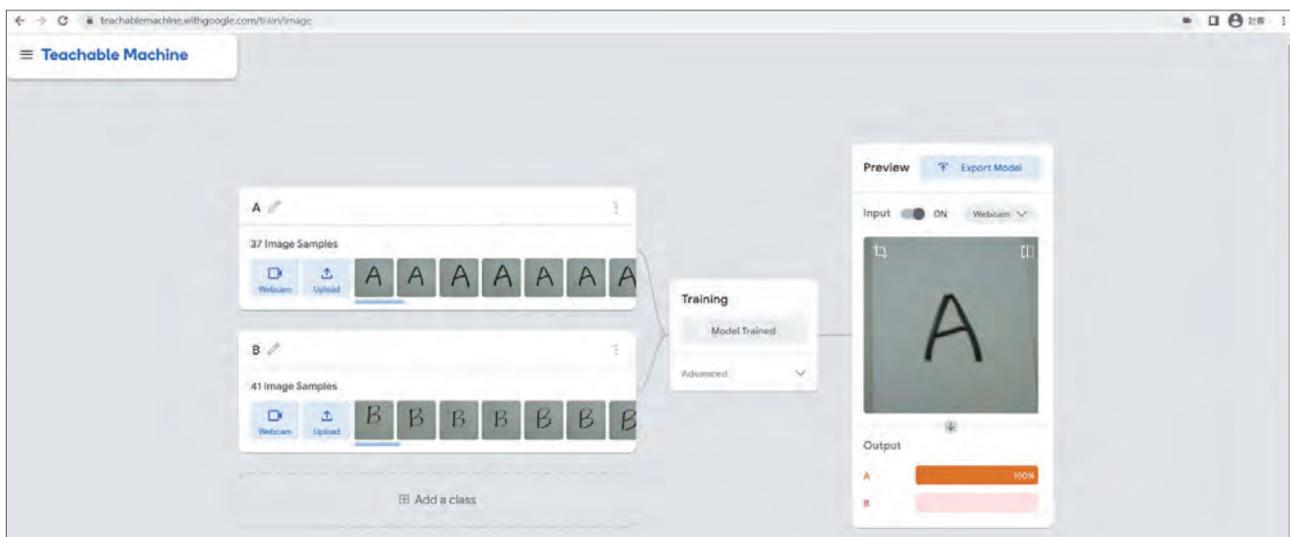
※ 注意：錄製的時候可 " 故意 " 將鏡頭拉遠拉近及翻轉角度製造素材的多樣性，以增進辨識的準確率。

錄製的時候若採用單一背景，盡量以簡潔乾淨背景為主，或是提供多樣的背景，避免訓練模組時一併將背景列為辨識物的特徵。

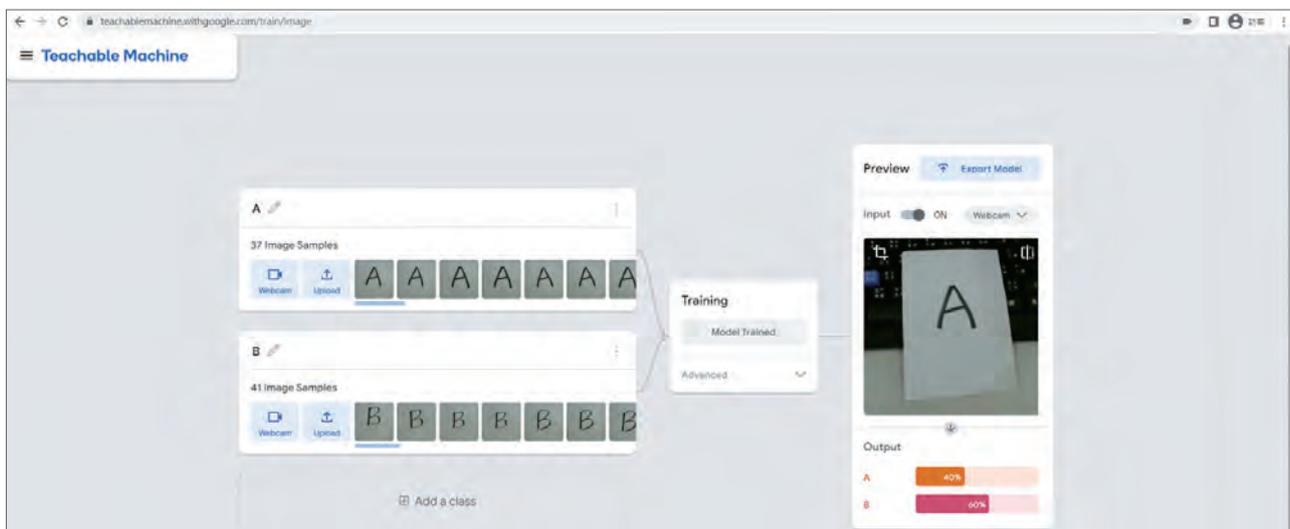
STEP 7 點選 Train Model



STEP 8 辨識成果



想想看：為什麼變換一下場景，辨識結果就不正確了？





實作活動

Pixetto- 影像辨識實作

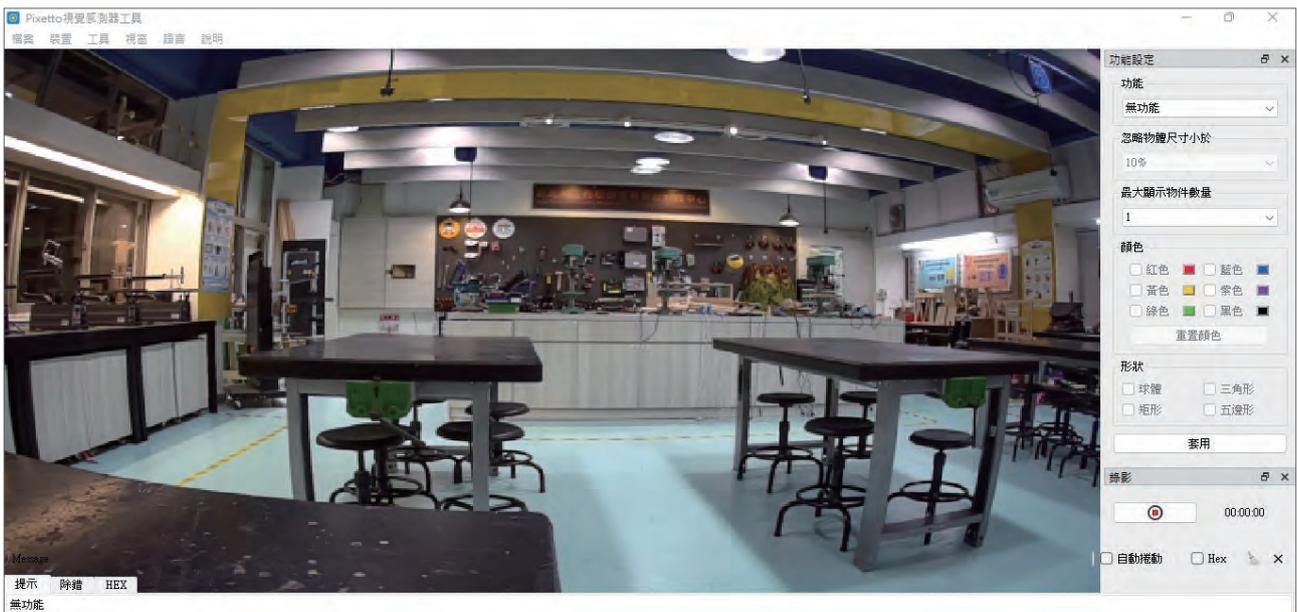
硬體介紹



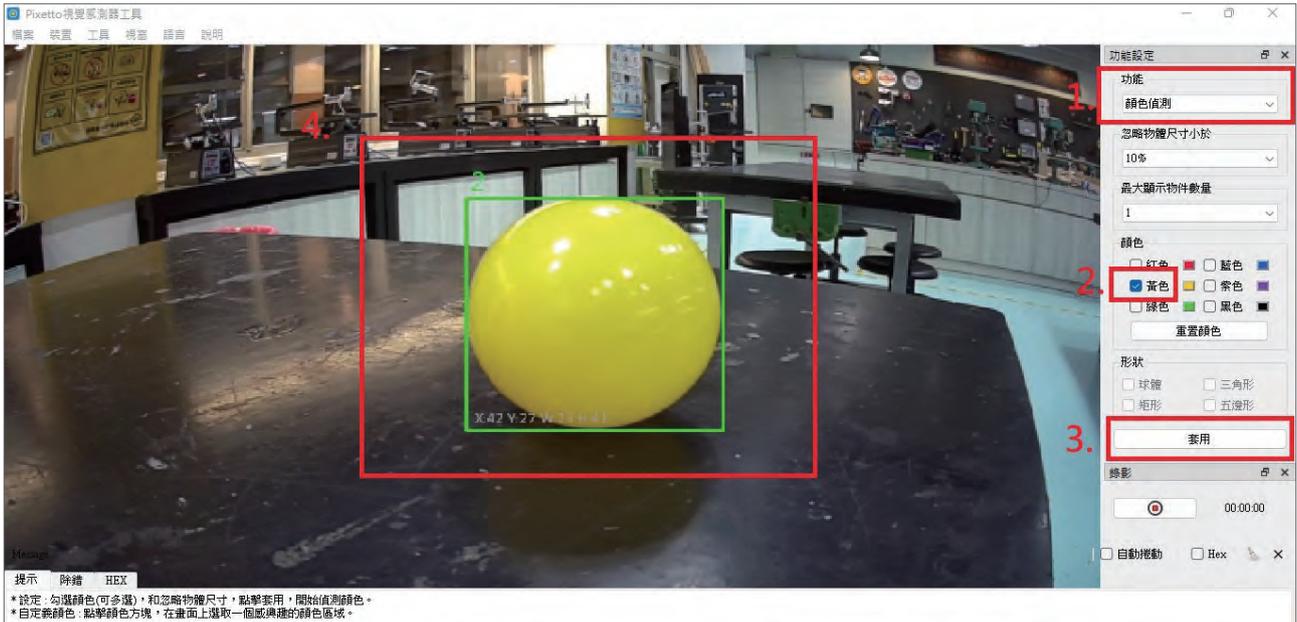
顏色辨識步驟

STEP 1 用 MicroUSB 線連接 Pixetto 並接上電腦，開啟 Pixetto Utility

STEP 2 啟動後畫面會呈現鏡頭對照的影像



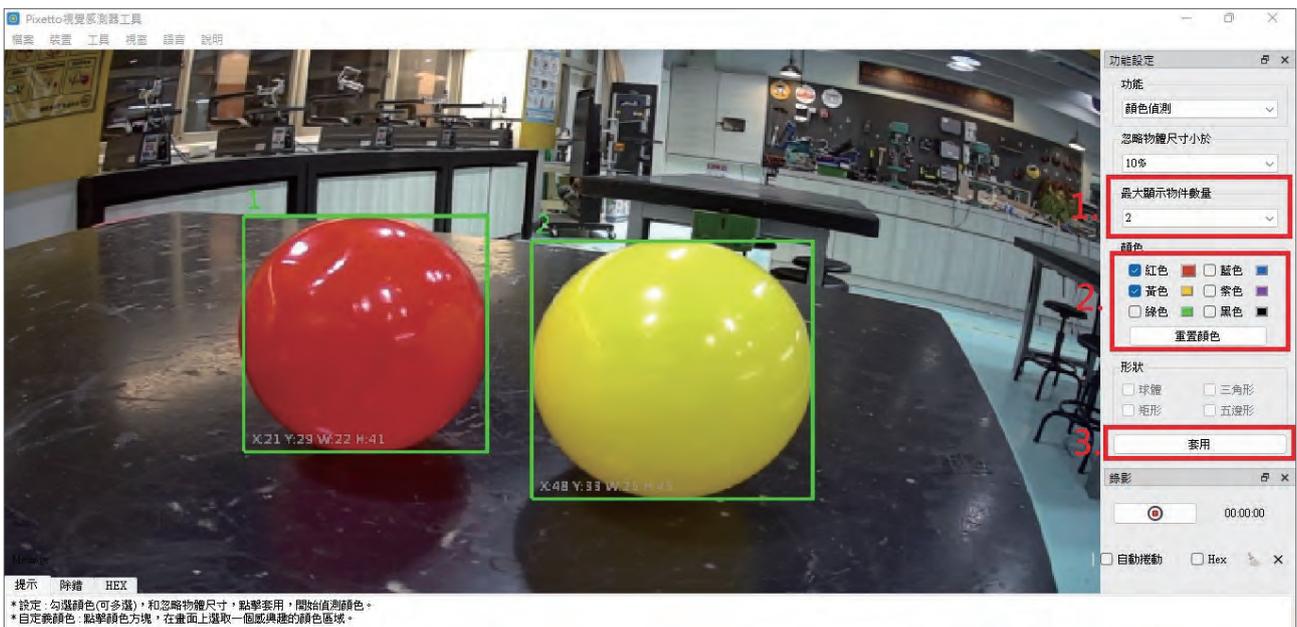
STEP 3 顏色偵測設定



- ① 功能切換成 " 顏色偵測 "
- ② 勾選辨識物體的顏色
- ③ 點選 " 套用 "
- ④ 可看到圖中變色物體 - 黃球被框起 (左上角數字為顏色代碼，紅黃綠藍紫黑依序為 1-6)

Tips：可根據辨識物體在鏡頭中的大小調整 " 忽略物體尺寸小於 " 的參數來過濾雜訊，讓演算法以拍攝較大的主體為主。

STEP 4 辨識不同顏色的多個物體



- ①視辨識物體更改 " 最大顯示物體數量 "
- ②勾選辨識物體的顏色
- ③點選 " 套用 "

可看到執行成果可以順利的偵測框起紅球與黃球

狀況：當顏色無法被偵測時該怎麼辦？

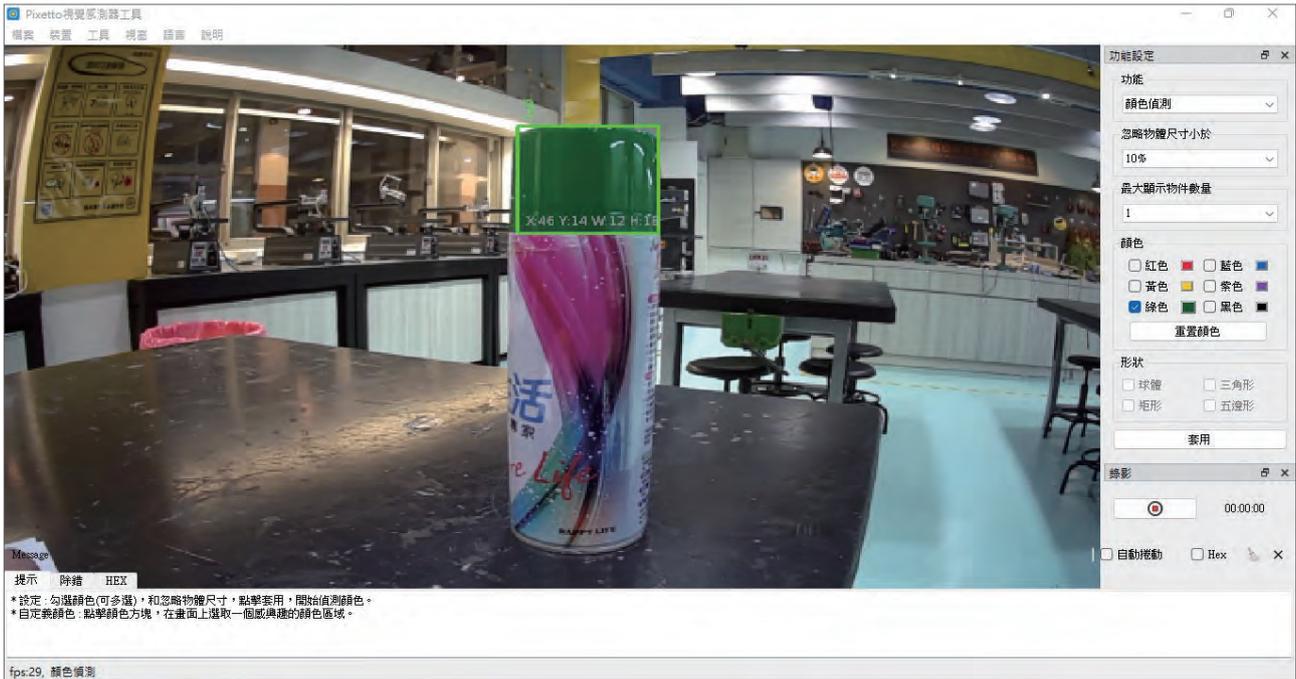
說明：如圖中照片勾選了綠色，但綠色漆罐依然無法被辨識



STEP 5 顏色校正



- ①點選顏色的小方框□（此時小方框會稍微變大□）
- ②用滑鼠框選辨識物體的顏色（此功能類似小畫家滴管 - 色彩選擇器的功能）
- ③點選 " 套用 "



執行成果可發現漆罐瓶蓋可順利被框起

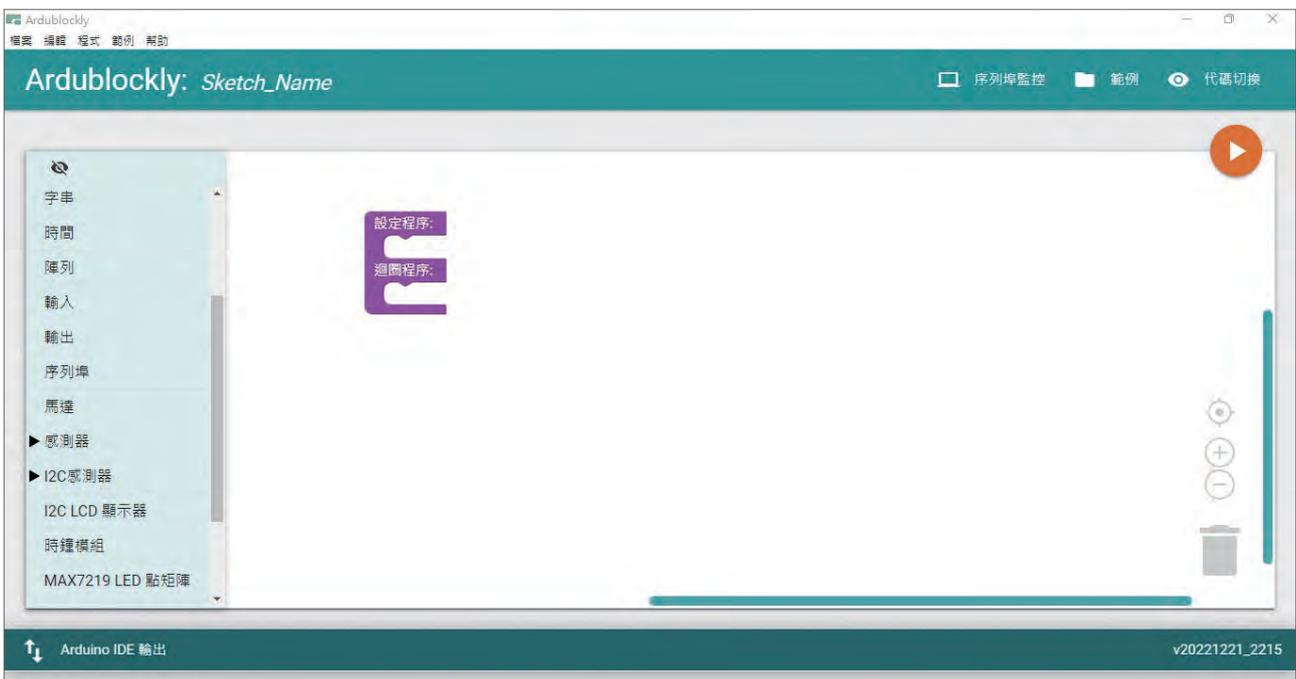


專題製作

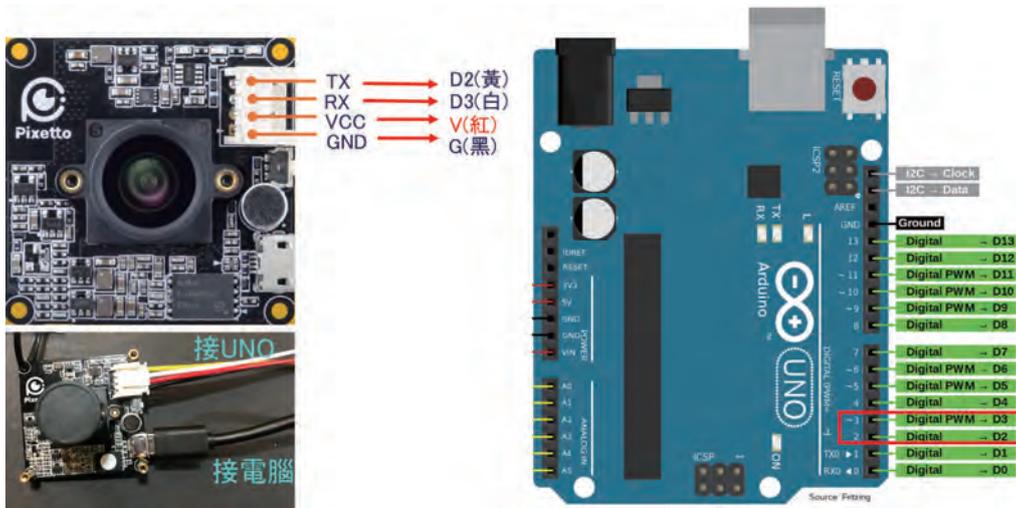
將 Pixetto 辨識結果整合 Arduino 製作專題

前置作業

1. 下載 ArduBlockly

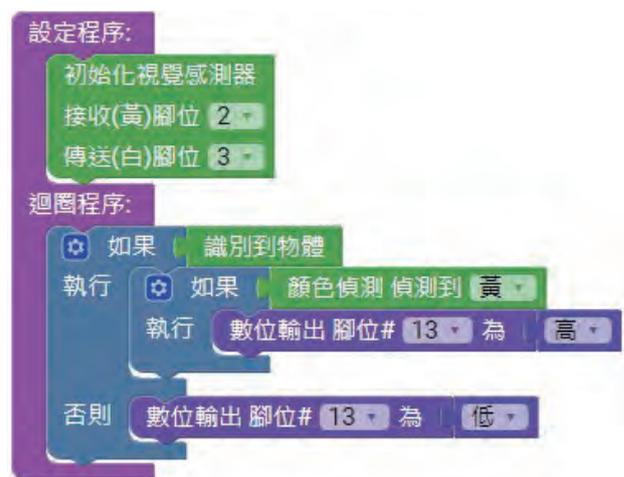


2.Pixetto 與 Arduino UNO 連接方式

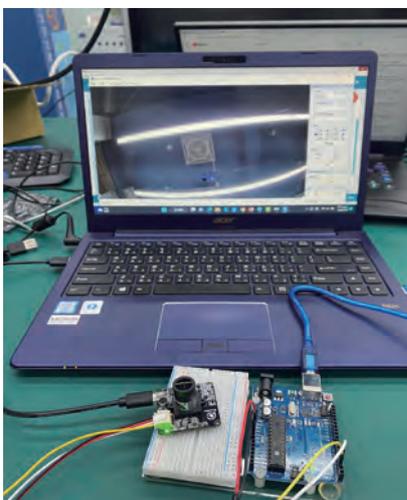


Tips : TX 與 RX 沒有一定要接在 D2 與 D3 腳位，硬體連接在哪，程式定義有對應到即可。

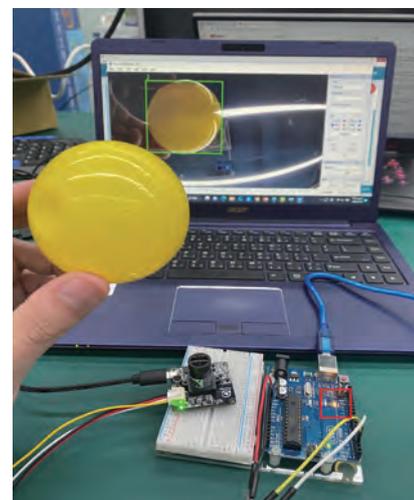
完整程式範例 辨識到黃球，LED 燈亮起



執行成果



▲ 當沒有黃球時 (燈暗)



▲ 偵測到黃球時 (燈亮)

Tips : 此專題沒有額外接 LED 燈，直接採用內建 D13 腳位的 LED 燈 (如紅框指示處)



練習活動

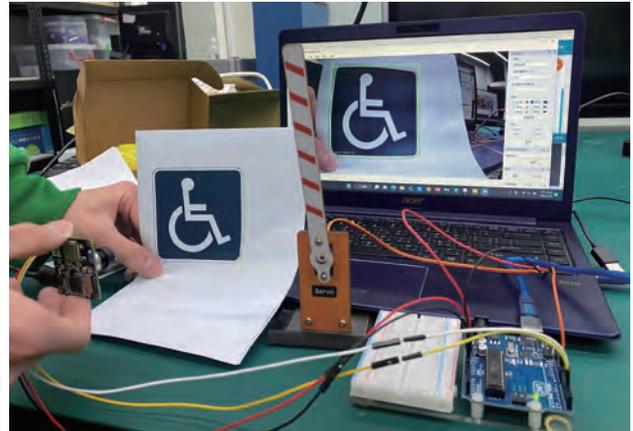
辨識殘障車位停車證後開啟
伺服馬達柵欄



參考答案

```
設定程序:  
初始化視覺感測器  
接收腳位 2  
傳送腳位 3  
設定伺服馬達腳位 servo_9, 腳位 9  
伺服馬達 servo_9 旋轉到 180 度  
迴圈程序:  
如果 識別到物體  
執行 如果 顏色偵測 偵測到 藍  
執行 伺服馬達 servo_9 旋轉到 90 度  
等待 3000 毫秒  
伺服馬達 servo_9 旋轉到 180 度
```

執行成果

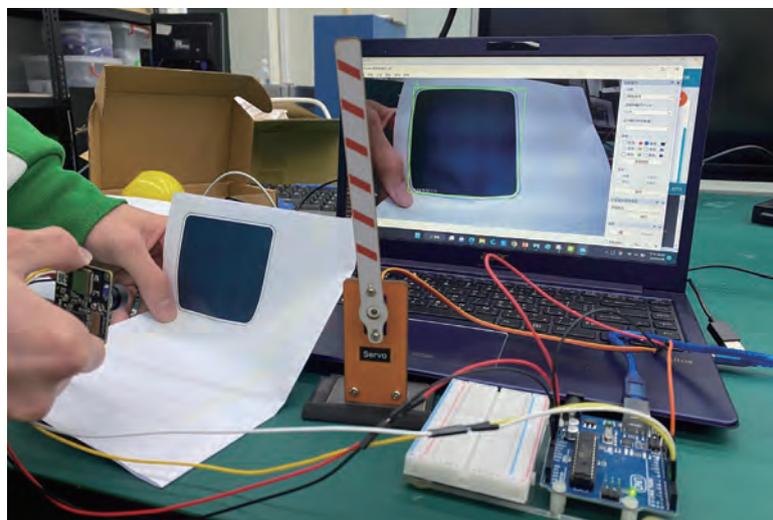


偵測到殘障車位標誌後，伺服馬達柵欄抬起 3 秒後放下。

狀況：以顏色辨識來完成這個專題有 Bug

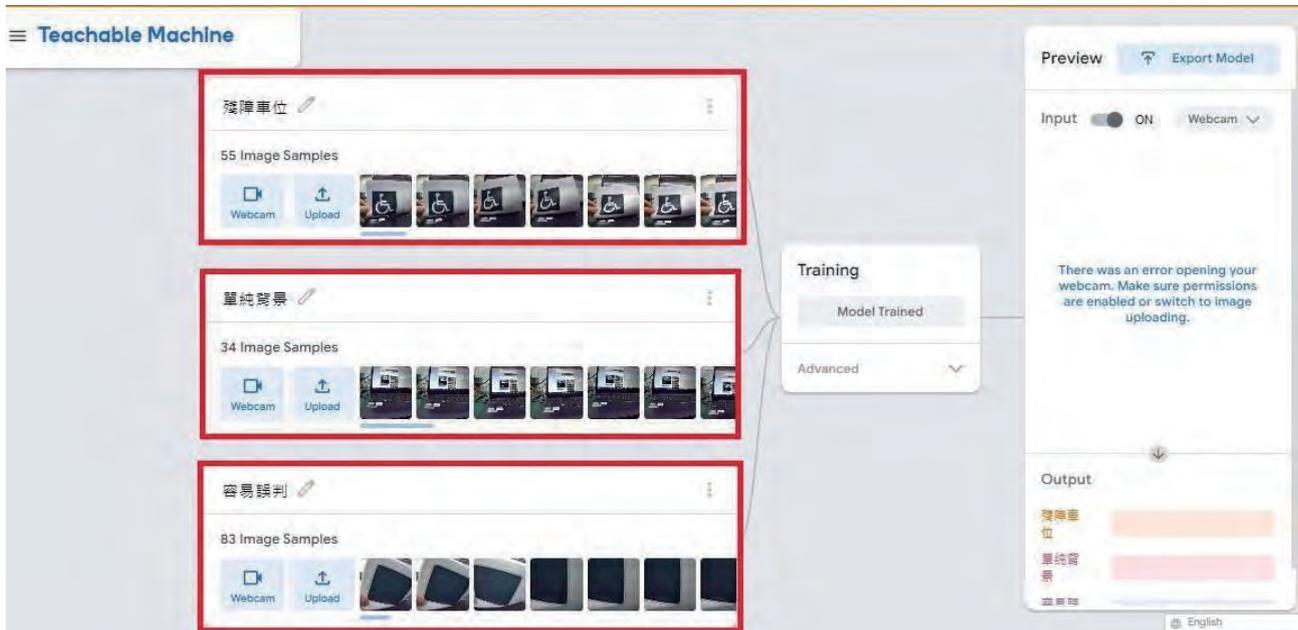
說明：因為是以 " 辨識藍色 " 作為伺服馬達柵欄抬起的條件，所以隨便拿一個顏色相近的物品就能啟動伺服馬達柵欄。

解決方法：試試以特徵參數作為判別的神經網路



神經網路辨識

1. 以 Teachable Machine 影像辨識訓練模型

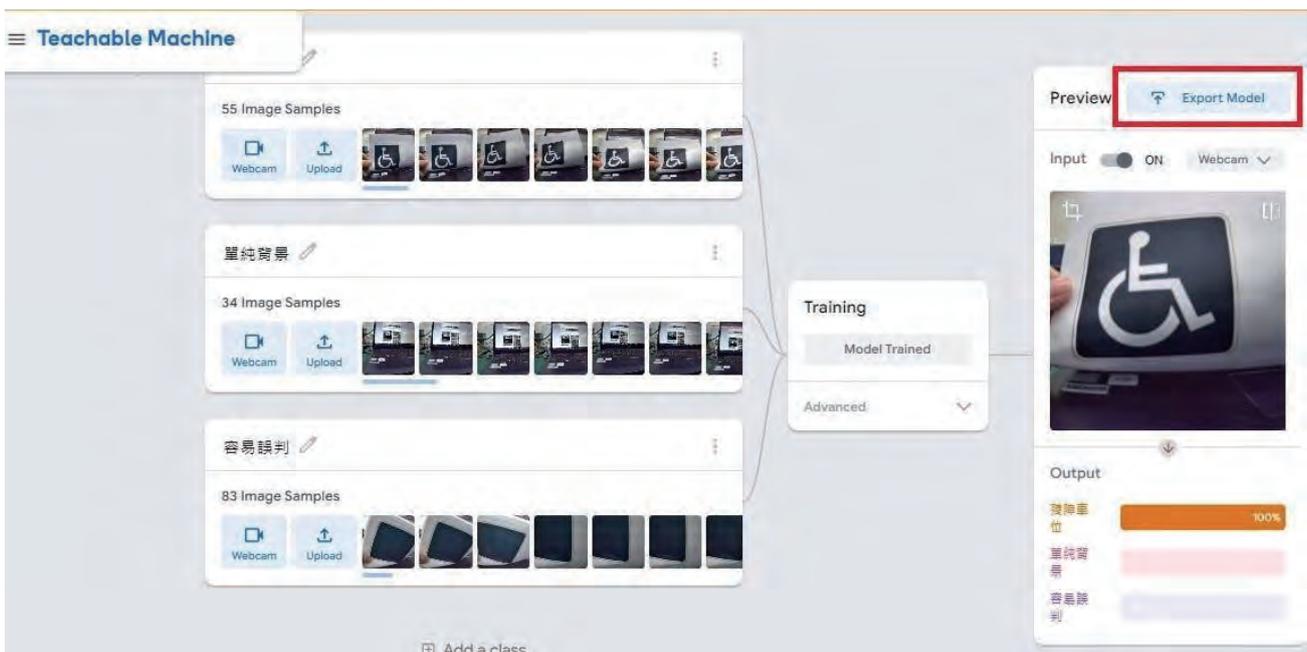


可以之前教過的方式分別記錄下列三類影像資料：

- ①殘障車位：資料盡量多元（不同角度、大小、遠近）
- ②背景：無辨識物在鏡頭前時的影像
- ③容易誤判的資料：如顏色相近的物體（如圖示）

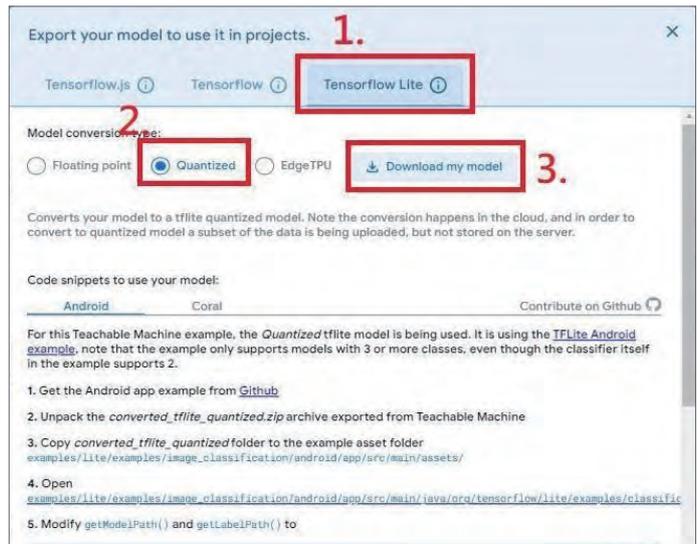


2. 下載訓練模型



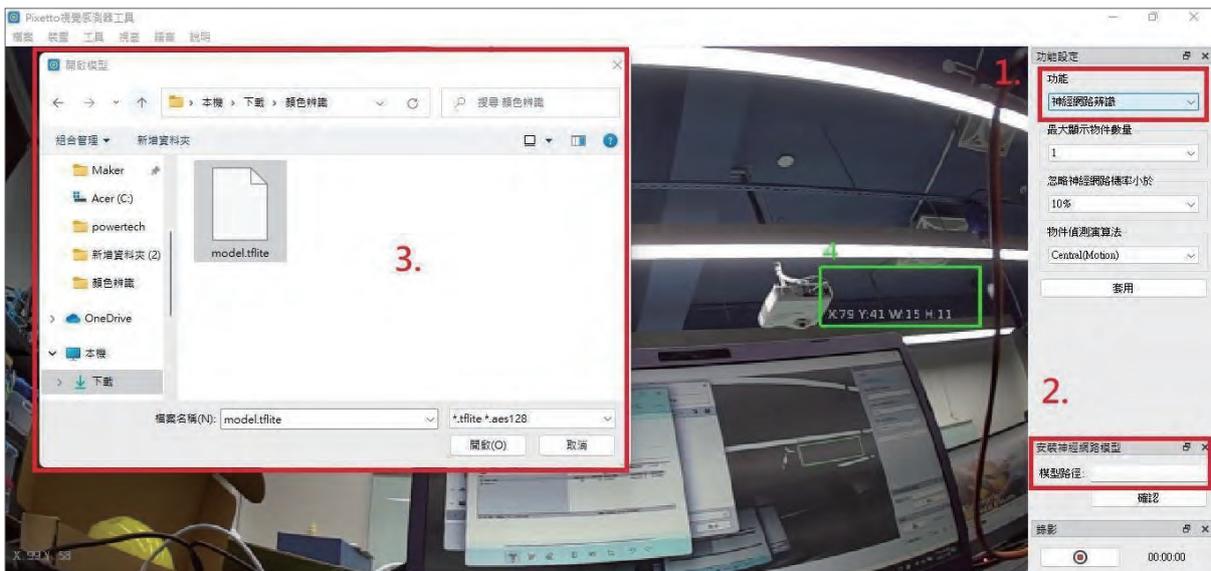
點選 "Export Model"

- ①點選 "Tensorflow Lite"
- ②點選 "Quantized"
- ③點選 "Download my model"

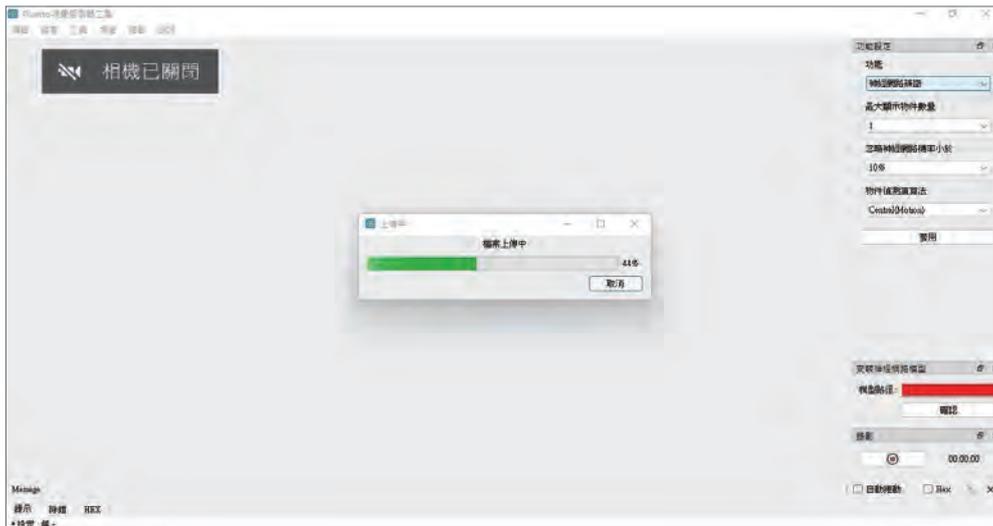


需先將 model 檔解壓縮至電腦

3.Pixetto 神經網路辨識



- ①功能選單更改為 "神經網路辨識"
- ②點選模型路徑的空白處
- ③點選剛剛解壓縮的訓練模型檔並開啟



點選確定後會將模型上傳

辨識測試：

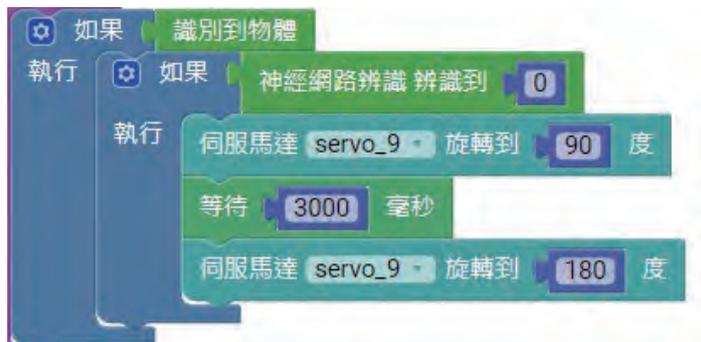


Pixetto 序號編列從 0 開始，可以看到目前辨識到的畫面為我們在 Teachable Machine 影像辨識訓練的第一類 " 殘障車位 "



①點選 " 除錯 "

②可看到目前 Type 的回傳值為 "0"，而 "0" 也會是我們連接 Arduino 上，神經網路辨識條件式判斷的值

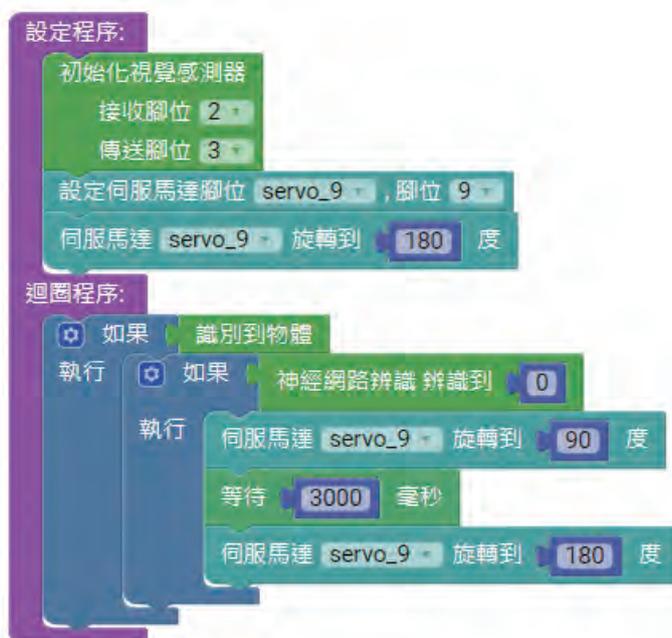


```
if (ss.isDetected()) {  
  if (ss.getFuncID() = Pixetto::FUNC_NEURAL_NETWORK && ss.getTypeID() = 0) {
```

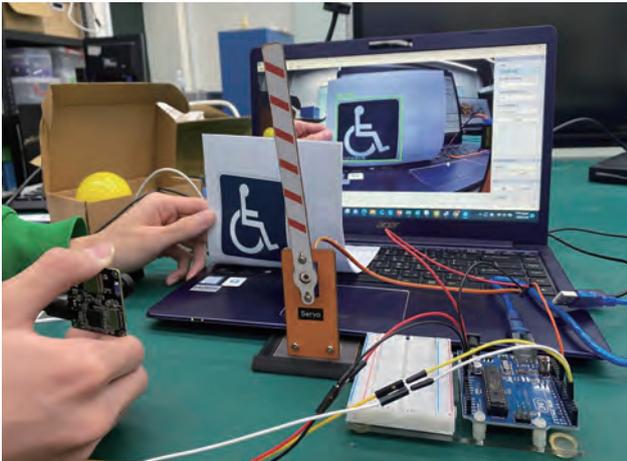


這時候顏色辨識拿來誤導的藍色色卡就會被排除到其他類

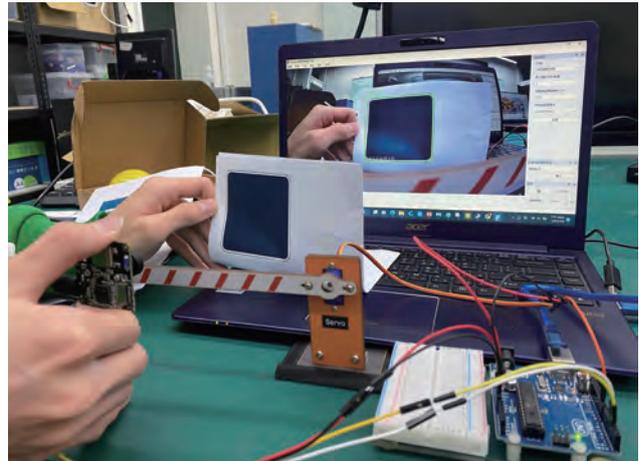
程式範例



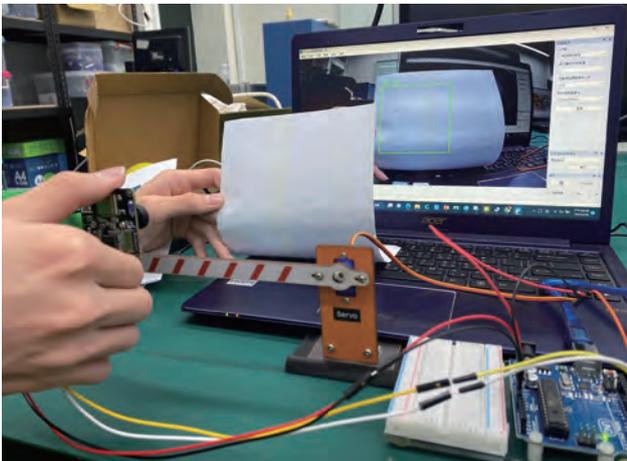
執行成果



▲ 殘障車位標示 (伺服馬達柵欄抬起 3 秒後放下)



▲ 其他類別 (無動作)



▲ 車牌辨識

小試身手

因應疫情時期，請運用學習過的 Pixetto 與 Arduino 製作一口罩辨識專題

任務

辨識民眾有無攜戴口罩，若無攜帶口罩經過則觸發蜂鳴器警報（可視零件設備自行調整）

自由創作

觀察一下生活周遭，想想所學影像辨識可以怎麼應用在生活中，設計製作一個專題吧

新北市政府教育局

新北 AI 微課程讀本

出版機關 —— 新北市政府教育局
發行人 —— 張明文
主辦單位 —— 新北市政府教育局中教科
協辦單位 —— 新北市政府新課綱行動協作平臺
承辦單位 —— 新北市立錦和高級中學
指導委員 —— 中等教育科吳佳珊科長·林育澄股長

主編 —— 張純寧
指導教授 —— 李坤彥·李俊賢·張玉山·陸敬互（依姓名筆劃數順序排列）
作者 —— 汪殿杰·林易民·林欽鴻·高靖岳·陳筠雅·蔡志聰
蔡勝安·蔡釋鋒·盧麗如（依姓名筆劃數順序排列）

承印單位 —— 台科大圖書股份有限公司
地址：24257 新北市新莊區中正路 649 號 7 樓
電話：02-2908-0313 傳真：02-2908-0112

版面構成 —— 陳依婷
封面設計 —— 魏怡茹
責任編輯 —— 楊清淵

出版日期 —— 中華民國 112 年 4 月