**Digi+Talent跨域數位人才加速躍升計畫**

**實務專題成果報告**

**實務專題：麻將Mentor**

|  |  |
| --- | --- |
| 團隊成員： | 郭建國、唐梓宸 |
| 研習生姓名： | 郭建國、唐梓宸 |
| 研 習 部 門： | 系統所智慧體感系統中心 |
| 研習期間： | 107年6月-107年12月 |
| 研習業師： | 翁明昉、吳雅萍 |

中華民國107年11月

1. **研究動機**
2. **研究目標**
3. **功能**
4. **架構**
5. **問題和挑戰**
6. **研究方法**
7. **研究成果**
8. **心得**
9. **未來建議**

**十一、參考文獻**

1. **研究動機**

**智慧眼鏡**

研習的部門主要發展的是智慧眼鏡。智慧眼鏡有兩個特色：**解放雙手、第一視角。**一般在使用智慧型手機或相機時，我們需要至少一隻手來進行操作，有時候一邊使用智慧型手機、一邊維修設備、還要一邊紀錄設備的數據，很容易忙不過來。而解放雙手的意思是智慧眼鏡不需要雙手即可使用，我們擁有更多的自主性可以去進行機器設備的維修或檢測。同理，在打麻將的時候若使用手機拍照進行判斷聽牌或胡牌進行輔助，即代表少一隻手來打麻將。

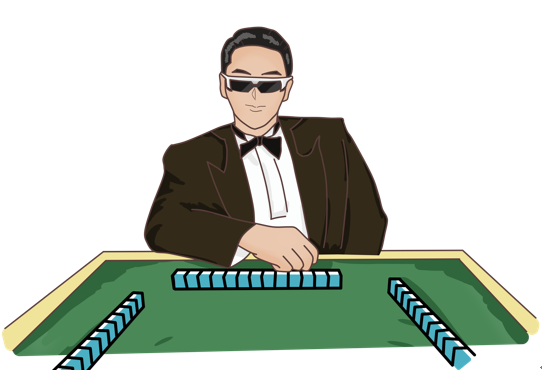
第一視角即是智慧眼鏡使用時，可以用第一視角觀看影像。傳統上在維修設備或有操作問題時觀看的說明文件，不是第一視角，是以第三者的角度去看文件，較無法讓自己身歷其境，效率也較低。



圖一：Epson智慧眼鏡

**輔助判斷（人工智慧）**

麻將玩法複雜且變化多端，尤其當一副牌內同一花色的牌數量很大時，不易辨別是否胡牌或聽牌。本組實作一個裝在智慧眼鏡上的APP - 麻將Mentor，結合影像辨識、人工智慧與演算法，當麻將玩家遇到是否聽牌或是胡牌的狀況時，將結果顯示於智慧眼鏡上，輔助麻將玩家迅速判斷。



圖二：麻將Mentor示意圖

1. **研究目標**

* 偵測與辨識麻將牌，包括牌的位置及牌的內容
* 判斷麻將牌組，並提供輔助資訊，包括：胡牌和聽牌；若可胡牌則智慧眼鏡顯示“胡”，若聽牌則智慧眼鏡顯示“聽”與聽哪幾張牌。

**1. 準確度**

在判斷聽什麼牌與是否胡牌方面，本組目標是至少達到98%以上的準確度。若系統準確度低於90%，則會發生使用智慧眼鏡的玩家“詐胡”的狀況。本組希望完成實務上可行的專案，因次希望麻將Mentor系統是真正可應用在真實世界之上。

**2. 軟硬體整合**

本組硬體使用Epson的智慧眼鏡，軟體則是胡牌聽牌的演算法、影像處理、微軟的Azure Custom Vision中的Classification Project。整個系統流程大致是：智慧眼鏡輸入影像，經過影像處理後傳入Custom Vision進行預測，預測的結果放入演算法進行計算胡牌聽牌結果，最後再傳回智慧眼鏡顯示於螢幕上。本組目標是整合整個系統，達到一定的完成度，並能實際產出成品進行展示。

**3. 使用流暢性**

使用麻將Mentor判斷聽牌或胡牌時，從輸入系統到判斷得出結果顯示到螢幕上，中間需要資料傳輸處理與轉換的時間，本組希望能讓玩家使用的流暢性越高越好，目標是在1秒之內得到判斷的結果。

**三、功能**

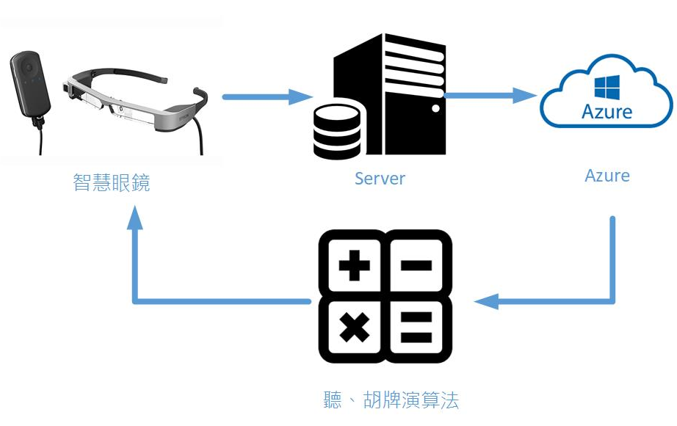
本組的情境是使用廣東13張麻將的規則，玩家在尚未摸牌的階段有13張牌，在摸牌之後會有14張牌。因此我們分成兩大功能：聽牌與胡牌。胡牌的規則是，要有一對 “眼”（兩張相同的牌）以及四對 “刻”（三張相同的牌）或是 “順”（連號的牌）。以下分別對胡牌與聽牌進行實例說明。

胡牌：舉例來說，玩家手上有很多張的條子，分別是11233445556677條，則此副牌可以湊成一對 “11”、四對順子“234”、“345”、“567”、“567”，這樣一來就是可以胡牌。反之，若湊不出一對再加上四對刻或是順，則不能胡牌。

聽牌：玩家手上有13張麻將牌，我們要知道再加上哪一張麻將牌，就能使整副牌胡牌。例如，1112345678999 這副牌，又稱“九寶連燈”，1到9都可以聽。又如，11333344455578聽的是6與9。

本組專注解決一副牌有很多同樣花色麻將牌的問題，比如整副牌都是條子，要計算聽什麼牌或是否胡牌需要大量的計算；這種複雜的計算也是人類相對機器而言較不善長的，因此我們聚焦於“條子”的聽牌與胡牌的功能實作，期望能帶給麻將玩家更棒的遊戲體驗。

**四、架構**



1. 智慧眼鏡：本專案使用EPSON BT-300智慧眼鏡做為應用軟體的載具，具備執行開發程式所需的硬體及作業系統。
2. Server（以Python撰寫）：智慧眼鏡透過相機取得影像後將資料傳至Server以進行下一步的辨識。
3. Azure：Server接到影像資料後，前置處理單元將取出多個感興趣的局部影像傳至Azure Custom Vision進行預測。
4. 聽、胡牌演算法（以C++撰寫）：取得Azure的預測後結果後，此演算法可計算出是否胡牌、聽牌等，最後將牌組的判斷結果回傳至智慧眼鏡。

**五、問題和挑戰**

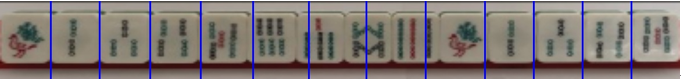
1. **影像辨識：**

因為捕捉麻將影像的時候，不同的距離、角度會對辨識結果造成偏誤，尤其是當一副麻將牌黏在一起的時候，無法準確的辨識出一張一張麻將牌的範圍（region）。如下圖，紅色的框框代表預測的結果，不符合實際麻將牌的位置。



* **裁切牌：**

為了解決黏在一起的牌辨識率不佳的問題，本組將整副麻將牌裁切成一張一張的麻將牌，然而切牌時有時候會切得不準確，如下圖：



* **高辨識率：**

若要增加辨識率，則要能準確的進行影像處理，將麻將牌裁切正確，並丟進Azure Custom Vision，利用Classification進行訓練，否則辨識的成功率沒辦法提升。

1. **聽/胡牌演算法：**

要使得智慧眼鏡能夠在接收麻將牌影像之後，將聽/胡牌的結果顯示出，必經的一段處理就是演算法。這個演算法的計算效能需要比人類快，並且每次餵牌進去時，輸出的結果需要有穩定且一致的結果。

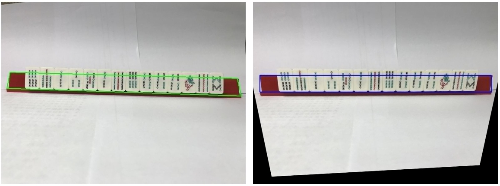
1. **系統整合效能**

每次進行整個輸入到輸出的系統流程，耗時最久的就是在Auzre Custom Vision的預測上面，因為每讀取一張麻將牌，就得進行一次結果的預測。如果一副牌有13張，那就須進行13次預測，可能耗時會超過3秒鐘。在分秒必爭的麻將賽場上可能會派不上用場。

**六、研究方法**

1. **裁切方法：**

本組預先對影像進行分析，利用輪廓偵測演算法以及透視轉換技術調整影像內容。利用輪廓偵測演算法，我們偵測出紅色牌架的四個頂點（corners），再利用透視轉換技術，將不同角度拍攝的一副麻將牌影像扭曲（warping）成與水平與垂直軸對齊的結果，如下圖：



接下來，因為我們只對麻將牌有興趣，其他不相關的背景我們想去除掉，以專注在麻將牌上，因此，我們裁取出麻將牌的感興趣區域（region of interest）。而為了修正因pitch夾角產生的形變，我們偵測麻將牌的實際位置，並再一次採用透視轉換技術將影像內容扭曲，如下圖：



我們已經去蕪存菁，處理後的影像只剩下一副麻將牌本身，接著我們想將一副麻將牌切割成一張一張的麻將牌，以方便進行後續的辨識。但因為鏡頭設計、3D空間拍攝角度差異、外框偵測不精準等因素，雖經過影像處理，相同大小的物體最後呈現在影像中的實際大小是不相等的。因此使用等距切割，會切得很不準確，造成物件辨識準確率低，如下圖：



為達成準確切割，我們設計一套調整切割線邊界的方法。利用“目標函式最小化”的技巧，我們先分析問題的本質以定義目標函式，並求取最佳解。

目標函式分為兩個部分，分別為資料項（data term）與平滑項（regularization term）：

* 資料項：物件和物件中間白色區域多、亮度高，因此傾向垂直線平均亮度越大越好的位置。
* 平滑項：物件和物件大小相差不大，因此傾向分割線與分割線的距離越均勻越好。

為求函式最佳解，我們隨機在起始的切割線附近尋找能量最小，最滿足目標函式的切割線，並調整至新的位置。待全部切割線皆調整過後，再重複此步驟數回合，即可得到準確切割一張一張麻將牌的切割線，如下圖：



1. **影像辨識：**

切割完麻將牌後，我們將一張一張的麻將牌放入Azure Custom Vision進行影像分類（Image Classification）訓練。我們使用影像分類取代物件偵測（Object Detection），一次辨識一張麻將牌，而非物件偵測的整副麻將牌，以增加辨識的成功率。本組使用Azure Custom Vision的平台進行影像分類訓練。Azure Custom Vision屬於最熱門的AI 機器學習服務系列產品之一，利用深度學習的技術對影像進行辨識與訓練，最後進行影像預測。它的特色是優化的使用者介面，可以在Azure網站上進行上傳照片、標記Tag、訓練模型、預測模型等等。此外，它提供訓練與測試的API，也支援許多程式語言如Python、C#、Java，方便使用者用程式執行。

為了提升影像辨識的準確率，我們蒐集了許多角度、光線、距離不同的影像作為樣本以增加樣本多元性，並使用Azure Custom Vision提供的k-fold cross verification交叉驗證，也就是將樣本拆成k份，第k份當作測試樣本，剩下的k-1份當作訓練樣本進行訓練。經過k次訓練得到的準確度取平均後，得到平均準確度來當作準確度衡量的標準。

在訓練模型時，我們使用漸進式訓練的方式，先將訓練樣本“候選人”丟入最新的模型進行測試，若測試結果是錯誤的（例如應該預測為3條，但是測試結果卻是4條），代表這錯誤的訓練樣本“候選人”無法被這個模型所準確預測，因此將錯誤的候選人們丟入模型進行訓練，利用機器學習的技術使模型從錯誤中學習，才能有效提升準確率。

1. **演算法論文：**

藉由2017年Yuan Cheng教授等人的一篇探討麻將數學的論文[1]，得到麻將聽牌與胡牌的演算法，我們根據此演算法進行麻將聽/胡牌的計算。

再者，師大的郭君逸教授也統整了麻將胡牌的計算方式[2]，定理如下：

一副牌P，若把一個對子（俗稱眼睛）拿掉後，假設此時數字最小的牌是x，若x 的張數是3 張以上，則拿掉3 張x（一刻），剩下牌為Q。否則拿掉x、x+1、x+2（一順）之後，剩下的牌為Q。若無法拿，則P 沒胡。換言之，「P 胡」若且唯若「Q 胡」。

**七、研究成果**

**辨識數據：**

* 資料庫：本實驗蒐集13張麻將牌和14張麻將牌的影像各150張作為實驗資料庫，我們先將全部300張影像透過前置處理演算法自動裁切成一張一張的麻將牌，共得出4,050張切割後的影像，再以人工篩選方式，共取得3,931張有效的麻將牌，以此作為訓練麻將牌分類模型及驗證麻將牌分類模型準確度的資料庫，其中1條至9條的樣本數分佈整理如下表所示：

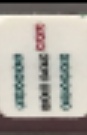
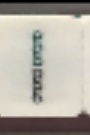
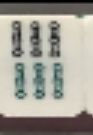
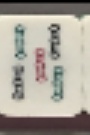
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1條 | 2條 | 3條 | 4條 | 5條 | 6條 | 7條 | 8條 | 9條 |
| 340 | 439 | 583 | 588 | 533 | 487 | 396 | 282 | 283 |

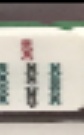
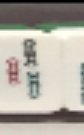
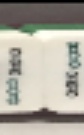
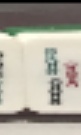
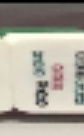
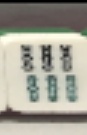
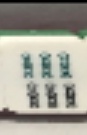
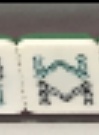
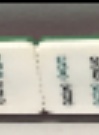
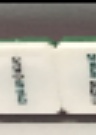
\*其中有119張麻將牌是切錯的，所以訓練時已預先將之挑除。

1. **13張/14張牌辨識結果：**本實驗的測試樣本以13張牌或14張牌的影像為主，若麻將牌的數量辨識正確，則代表我們知道要回答的問題類型（例如牌組有13張麻將牌則回答聽什麼，若有14張麻將牌則回答是否胡牌），反之則可能導致答非所問的錯誤狀況。由下表的實驗結果可知，目前我們開發的演算法可以100%準確的判斷所輸入影像的麻將牌數目。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 牌組類型 | 樣本數 | 正確數 | 錯誤數 | 準確度 |
| 13張 | 150 | 150 | 0 | 100.00% |
| 14張 | 150 | 150 | 0 | 100.00% |

1. **牌裁切結果：**牌裁切正確，代表影像中每張麻將牌是切齊的，且單張牌的內容是完整的。具體的說，我們以人工方式確認切割後的每一張麻將牌，如果可以判斷出是什麼牌，且不產生模擬兩可的情況，就是正確；反之若切到牌的中間、涵蓋到其他牌或者缺少自身牌的影像，則為錯誤。我們量化評估的計算方式是採用影像為單位，也就是說，如果一張影像中只要有一張牌或一張牌以上的裁切不正確，即代表牌裁切錯誤，反之，則需要若所有牌（也就是13張或14張）的裁切皆是完整且可辨認才代表牌裁切正確。由下表的實驗結果可知，在150個樣本，每個樣本皆有13張麻將牌的切割中，正確數為143張，錯誤數為7張，切割準確度為95.33%，與一個樣本有14張麻將牌的切割準確度相同。

****

****

上排圖：切正確。下排圖：切錯誤。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 牌組類型 | 樣本數 | 正確數 | 錯誤數 | 準確度 |
| 13張 | 150 | 143 | 7 | 95.33% |
| 14張 | 150 | 143 | 7 | 95.33% |

1. **牌辨識結果：**

本組的實驗數據是利用Azure Custom Vision[3]訓練模型時內部的k-fold cross-validation的結果，也就是訓練樣本切成k份，利用剩下的k-1份進行訓練去測試第k份的結果。

預測模型：測試樣本辨識單張麻將牌準確率皆超過九成以上，其中1、2、5、6、8、9條的準確率皆為100%；準確率尚未到達100%，但是接近100%的為3條與7條；其中比較不準的是4條，只有94.77%的準確度，其中最可能錯誤辨識成3條，錯誤辨識成3條的機率為3.68%。

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | 正確答案 | | | | | | | | |
|  |  | 1條 | 2條 | 3條 | 4條 | 5條 | 6條 | 7條 | 8條 | 9條 |
| 預測結果 | 1條 | 100.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| 2條 | 0.00% | 100.00% | 0.24% | 0.58% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| 3條 | 0.00% | 0.00% | 99.52% | 3.68% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| 4條 | 0.00% | 0.00% | 0.24% | 94.77% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| 5條 | 0.00% | 0.00% | 0.00% | 0.97% | 100.00% | 0.00% | 0.00% | 0.00% | 0.00% |
| 6條 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100.00% | 0.66% | 0.00% | 0.00% |
| 7條 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 99.34% | 0.00% | 0.00% |
| 8條 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100.00% | 0.00% |
| 9條 | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 0.00% | 100.00% |

1. **聽/胡牌系統結果：**

本組另外取了100張影像作為系統測試的樣本，分為兩類樣本。第一：單張影像有14張麻將牌的樣本有50張。其中能夠胡牌的影像有30張，沒有胡牌的影像有20張；第二：單張影像有13張麻將牌的樣本也有50張。其中聽6條的有14張、聽147條的有12張、沒聽的有24張。下圖橫軸的；“正確” 代表測試樣本經由本組所發展的麻將Mentor系統計算獲得的結果完全符合實際預期的結果，“錯誤”則表示電腦推薦的輔助結果與本組採用人工判斷的預期結果不一致。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 測試數量 | 正確 | 錯誤 | 正確率 |
| 胡 | 30 | 20 | 10 | 66.67% |
| 沒胡 | 20 | 19 | 1 | 95.00% |
| 聽6條 | 14 | 3 | 11 | 21.43% |
| 聽147條 | 12 | 1 | 11 | 8.33% |
| 沒聽 | 24 | 23 | 1 | 95.83% |
| 總計 | 100 | 66 | 34 | 66.00% |

上表是本組聽/胡牌系統初步實驗的結果，在100張測試樣本中，影像中的牌組是胡牌類型的正確率為66.67%，沒胡類型的牌組的正確率為95%，聽6條的正確率為21.43%，聽147條的正確率為8.33%，沒聽的正確率為95.83%，全部的平均正確率為66.00%。

以下是一個辨識正確的測試樣本的示意圖。第一張圖是原始影像；第二張圖是將第一張原始影像經過影像處理，切割成一張一張的麻將牌；第三張圖是麻將Mentor提供的建議。此例是成功辨識是否胡牌，系統顯示“恭喜!胡牌了!!”。







以下是一個辨識錯誤的測試樣本的示意圖。第一張圖是原始影像；第二張圖是將第一張原始影像經過影像處理，切割成一張一張的麻將牌；第三張圖是麻將Mentor提供的建議。此例是錯誤辨識是否胡牌，系統顯示“很抱歉!牌組無法辨識，請重試”3。因為切割錯誤，所以無法辨識，系統會告訴使用者此牌組無法辨識。[[1]](#footnote-1)



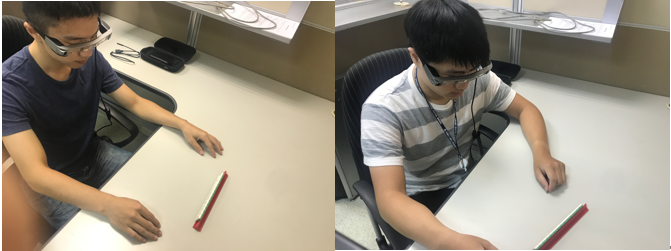




現階段麻將Mentor準確度不如預期的原因可能是因為單張麻將牌辨識的準確度仍不足，舉例來說，假設單張麻將牌辨識的準確度為98%，如果一副牌有14張牌，要全部辨識成功的機率為0.98的14次方（約等於0.75）。由於目前系統單張麻將的平均準確度不到那麼高，因此牌組所提供的輔助建議正確率尚有進步空間。針對此問題，本組預計採用漸進式模型訓練方法，將錯誤的測試樣本當作訓練樣本持續訓練模型，例如4條的錯誤率高，因此我們多放4條的樣本進行訓練。

另外，在影像切割處理時有機率切割錯誤，造成麻將牌組無法辨識的狀況，像是切割到兩張麻將牌的中間白色部分、將一張麻將牌從中間切開等等。若此情形發生，則後面的辨識必發生錯誤。因此若能事先將辨識錯誤的情形放在模型中進行訓練，則可避免作出錯誤的建議，造成使用者嚴重的損失。本組預計未來兩周完成此項工作。

**智慧眼鏡顯示：**本組開發麻將Mentor，使用智慧眼鏡顯示辨識麻將牌的結果，並加入是否胡牌與聽牌的輔助訊息，讓使用者能在一邊用雙手進行麻將遊戲的情況下，一邊使用智慧眼鏡協助玩家做決策。如下圖所示，玩家不需空出額外的一隻手使用手機拍攝照片，使用智慧眼鏡能使玩家擁有更棒的遊戲體驗。



**八、心得**

**梓宸部分：**

很開心能夠參加這次的研習計畫，讓我能夠認識明昉、建國、雅萍等等高手，明昉在我們專題進行期間給我們許多指導也給予很多幫助；建國則是跟我一起做專題，他數學方面比較厲害，所以在演算法上面出了很多力；雅萍姐則是細心得提醒我們要交一些文件，也都會幫我們檢查再交出去。

這次專題學到了很多之前學不到的事物，例如模型訓練、團隊合作等等，這些都是非常難得的經驗，希望這份情誼大家可以繼續維持，在往後的日子裡依舊互相幫助。

**建國部分：**

我在資策會學到很多，我的好夥伴梓宸是非常值得信賴的好隊友，在做專題時遇到問題的時候，我們會一起討論，並思考如何解決。他很善於思考問題的本質，也很願意花時間研究解決問題的方法。他寫程式的資歷比我多，在系統整合、伺服器等方面幸虧有了他，才能讓專題能夠順利完成。雅萍在研習期間給了我們相當多的協助，包括電腦設備的架設、如何寫雙週誌、專題報告，讓我與梓宸無後顧之憂的專心進行專題研究。明昉是我們團隊的明星球員，在我們遇到物件辨識辨識率過低的問題時，跳出來解決這個難題。因為我之前是讀財務金融，沒有做過資訊相關的專題，明昉手把手帶著我們規劃整個專題的進度與流程，並在我們遇到問題時熱心的指導我們。十分感謝有這個機會可以參與這次的DIGI+計畫，我在研習期間收穫滿滿，希望這個計畫在未來可以繼續下去。

**九、未來建議**

* 牌的多元性：這次的專題只針對條子做計算，為更加貼近真實的麻將遊戲，我們可以加入萬子、筒子、字、花等麻將牌進行辨識。
* 系統流暢度：目前測試的結果，我們的辨識流程需要數秒鐘才能傳回是否聽牌/胡牌的答案，未來可以讓系統更加流暢，迅速地告訴玩家辨識的結果。

**十、參考文獻**

1. Yuan Cheng, Chi-Kwong Li and Sharon H. Li, Mathematical aspect of the combinatorial game "Mahjong"
2. <https://sites.google.com/a/g2.nctu.edu.tw/unimath/2018-07/mj2>
3. <https://docs.microsoft.com/en-us/azure/cognitive-services/custom-vision-service/python-tutorial>

1. 3 本組將實驗3牌辨識的類別由原本的9類擴增為10類（除了1條至9條之外，新增無法辨識類別），此類別將切割不正確的麻將牌當作訓練樣本。當麻將牌進行預測時，若牌組中有一張或一張以上的麻將牌的辨識結果顯示為無法辨識的話，則Mentor就會提供無法辨識，請重試的建議，以避免因提供錯誤建議而造成玩家嚴重損失的窘境。除此之外，本組也將針對實驗4聽/胡牌系統的部分重新統計量化結果，提供電腦推薦錯誤的案例中屬於無法辨識的數量和比例。 [↑](#footnote-ref-1)