

利用敲擊試驗找出芋頭音波傳遞與質地之相關性

指導老師：趙于翔 教授

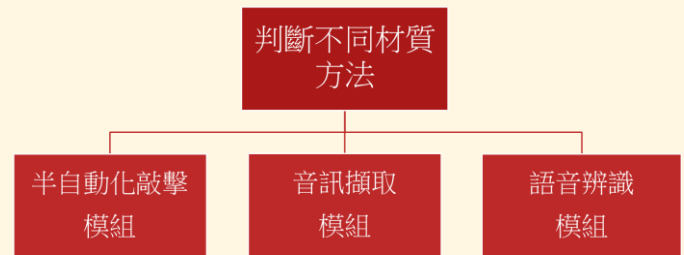
學生：鍾瑄、王子衡、廖明志、張仲恩

前言

近年來，人工智慧成為電腦科學中突飛猛進的一門領域，語音辨識是其中的一種常見的應用，例如：音樂資訊檢索、SIRI、GOOGLE 語音助理等等，於是我們就想利用這次專題的機會來研究語音辨識的相關技術，剛好食品系也提出相關的計畫：使用聲音來辨識出芋頭的好壞，由於目前並非芋頭之產季，我們就決定先從基本的演算法開始研究起。

系統架構

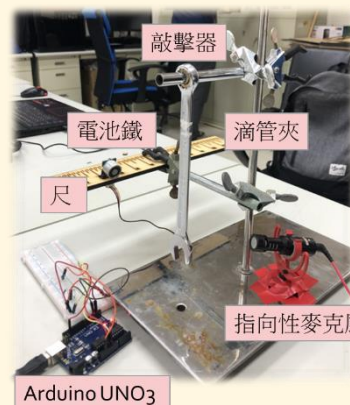
在此計畫中，我們提出一套判斷材質的方法，針對這個方法我們提出了三個主要的模組，分別是自動化敲擊模組、音訊擷取模組以及深度學習辨識模組，如圖一所示。



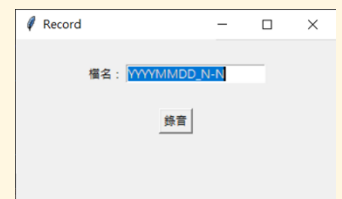
圖一、系統架構圖

簡介

整體的操作模式是，首先使用半自動化敲擊模組如圖二所示，敲擊實驗物品，並使用音訊擷取模組如圖三所示錄製不同材質的敲擊聲音，透過深度學習辨識模組做音訊前處理後使用隱藏馬可夫模型 (Hidden Markov Model, HMM) 得到辨識結果，最後將最佳辨識結果經過交叉驗證 (K-Fold Cross Validation) 找出，並簡易的顯示出來。



圖二、半自動化敲擊模組



圖三、音訊擷取模組

實際展示

之前再跑辨識時，發現當狀態數為 1 和 7 時的辨識率最高，但這不太正常，因為我們的樣本種類單純，但不至於單純到狀態數為 1，更不可能複雜到狀態數為 7。於是，我們使用了 k-fold 對不同的狀態數做交叉驗證，得到在狀態數為 4 的情況下，模型的辨識率最高。圖四圖五為實驗結果。

```

dataSet/k_fold_3/
iron stone woodBox softBottle plasticBottle
20 0 0 0 0
0 20 0 0 0
0 0 20 0 0
0 1 0 17 2
0 0 0 1 19
  
```

圖四、混淆矩陣結果圖

```

Final recognition rate is 97.00 %
Final recognition rate is 95.00 %
Final recognition rate is 91.00 %
Final recognition rate is 78.00 %
Final recognition rate is 98.00 %
  
```

圖五、k-fold 各資料集結果

未來展望

這次使用了傳統的 HMM 模型以及少許的樣本數作為研究的資料就已經達到 92.16% 的正確率，但是我們認為整體辨識率可以再提高，所以未來會增加更多的樣本，以及使用更多不同的神經網絡來嘗試提高辨識的準確度，並且設計一套便於民眾使用的行動應用程式以將此研究結果推廣到一般大眾的手中，讓大眾可以不需要經過訓練就可以買到品質較好的芋頭。