MA3113\* Topics in Mathematical Image Processing

**Second Written Report**

Wei-Zhi Chen & Po Jui Huang

Department of Mathematics

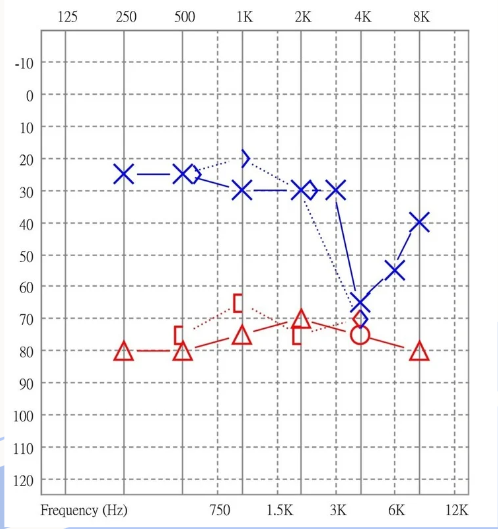
National Central University

**2024.05.17**

1. **前情提要**

問題介紹

我們的主題是分析的醫院聽力檢查所產生的圖表結果，藉由圖形的辨識，輸出成一個數據形式的excel檔案。



圖一：使用圖表的範例

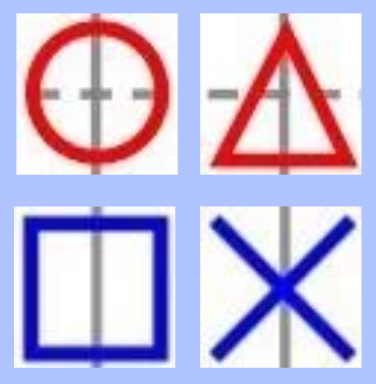
圖表內容一共有紅藍2色，隨著不同形式的檢查，圖表上具有8種symbols，而symbols之間會具有排擠、重疊、特殊標記的清況發生，排擠會發生於滿足特定條件下symbol可能不會出現在它該出現的位置， 而是出現至左右兩邊，重疊則是會發生於symbol位於相同座標點上時，特殊標記則為醫生在做測試時，患者沒有反應的紀錄。



圖二：輸出格式

演算法

先透過濾色器使圖表減少其他顏色的干擾，再將我們需要偵測的圖樣已截圖的形式製作成對照的素材，並利用MATLAB把素材做成矩陣，再以逐點的方式一一對照，讓素材不同顏色之處作為特徵，如果特徵小於某一特定數值，代表我們找出了圖表上的圖樣。



圖三：圖樣的樣本

1. **會議紀錄**

2024/03/15

使用最初版的演算法，透過圓形的特徵，在圖表上找出圓形，但受限於偵測方式，只能使用特定的圖表，因此在後續討論中廢除。和醫師討論方面，確定了我們所需要尋找的八種symbols和醫師希望能呈現出來的結果，並告知我們所遇到的問題。

2024/03/28

更新至當前最新的演算法，將演算法進行了模組化，讓其他人能夠簡單的操作，其辨識產生的結果還不錯，繼續將其他的symbol以此方法辨識。

和醫師討論上，考量到圖表的樣式，預計將輔助線的干擾去除，以求更好的品質，醫生有要求盡量將準確率控制在九成五以上，數據上控制在正負五分貝以內。

2024/04/26

我們發現輔助線對於干擾上幾乎微乎其微，繼續投入其他symbols的辨識。和醫師討論目前市面上或學術上的研究成果，開始閱讀醫生所提供的論文資料，以及學習在撰寫paper上的模式。

2024/05/10

將程式結果輸出為Excel檔，並交至醫師判讀是否符合他們的要求。並另外和醫生統計圖表的準確率。

2024/05/13

和其他醫師討論後續的事宜，將剩餘的問題排除。時間上這學期目前已辨識完成為主。和醫生另外提及海報比賽的事情，並取得了他們的同意。

1. **文獻參考**

我們參考了以下這兩篇paper作為參考依據:

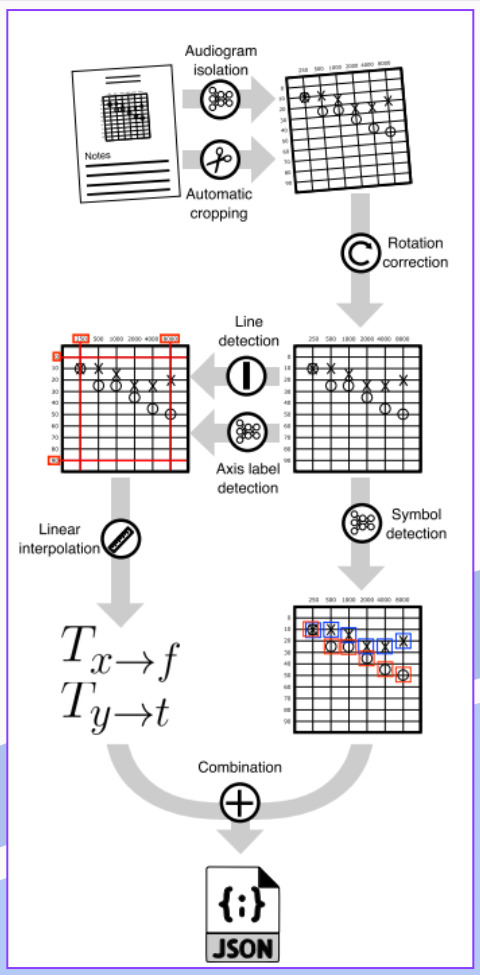
INTERPRETING AUDIOGRAMS WITH MULTI-STAGE NEURAL NETWORKS

Shufan Li, Congxi Lu, Linkai Li, Jirong Duan, Xinping Fu, and Haoshuai Zhou

Audiogram Digitization Tool for Audiological Reports

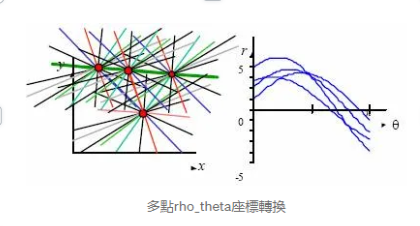
FRANÇOIS CHARIH AND JAMES R. GREEN , (Senior Member, IEEE)

其中第二篇是作為第一篇的延伸研究，流程圖大致如下：



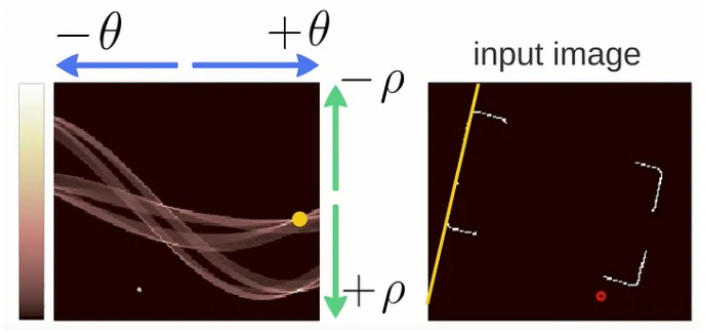
圖四：流程圖

在流程方面，先利用Hough transform 作為尋找圖表的工具，找出圖表在文件上的位置，並透過斜率來判斷是否將圖表進行轉正，透過這個處理方式來整理參差不齊的文件，隨後在進行座標線和symbols 的辨識，由於他們的資料來源是相機拍攝的，和我們的圖表來源都是同個機器和格式的產物不同，我們並不需要前述整理圖表的操作。不過我們曾經有想過利用Hough transform作為消除輔助線的依據，在此簡單敘述一下：



圖五：Hough transform

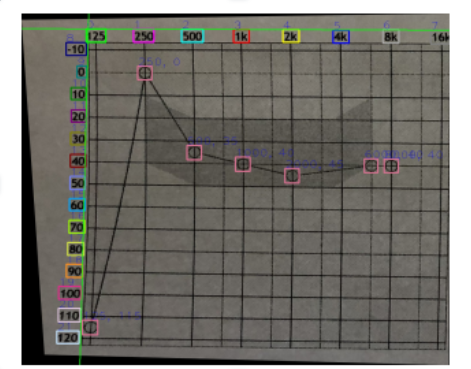
如圖片所示，一個點可以通過無限條的直線，我們可以將這些直線在極座標上記錄成一條正弦曲線，而多個點的圖像，將會形成具有數個交點的曲線，每個交點都是代表特定的直線，下圖則為例子，左圖中的黃點即代表右圖的黃色直線。



圖六：Hough transform例子

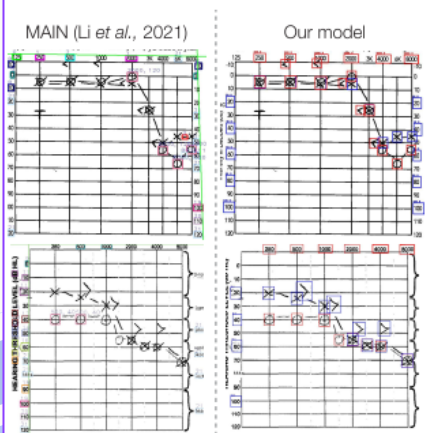
而在後續的辨識上，他們分別採用了YOLOv5和FASTER-RCNN他們都屬於多層的卷積神經網路，用於訓練辨識物件的工具，由於卷積神經網路並非我們所採用的的演算法，在此不多做太多的贅述，不過我們發現了一個有趣的訓練集，Common Objects in Context(COCO) Dataset，它以數以萬計日常生活中的照片，加上許多特徵標記，讓人們減少了許多收集資料，訓練資料的麻煩，只要再另外再加入一些針對專案需要的特殊圖形標記，就能快速的建立一個數量龐大且完善的偵測資料庫，算是以後可以參考的方法。

最後，透過以上方法，第一篇的作者群成功找出圖表，並發表作法MAIN(Multi-stage Audiogram Interpretation Network)，但受到誤差和圖表的侷限性，也只能找出特定兩種圖形，並不支援多功能偵測，如下圖所示。



圖七：Multi-stage Audiogram Interpretation Network 結果

而另一團隊，則是挑戰了手寫圖表的辨識，並以MAIN做為參考及比較依據，可以發現，就準確率而言比MAIN還清楚，但手寫的複雜度，可以預見它會更加的不穩定，有時會直接錯過需要尋找的圖形。



圖八 輸出結果

1. **目前演算法**

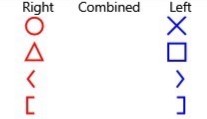
Notations:

整個的圖形：圓形、三角形、正方形、叉叉（附圖的上面兩排）。

半個的圖形：大於、小於、左中括弧、右中括弧（附圖的下面兩排）。

著色率：目標圖有顏色的像素數量除以樣本圖中有顏色的像素數量。

誤差率：不應該有顏色卻有顏色的像素除以樣本的著色像素量。

****

圖九：附圖

目前演算法版本已經推移到了4.0版本，改良了分辨依據且已經完善了整個的圖形的偵測。

分辨依據改良：

根據著色率的公式，若是著色率的值非常接近1，代表目標圖中含有的有顏色的像素數量相當接近樣本圖中有顏色的像素數量。

另外根據誤差率的公式，若是誤差率非常接近0，則代表「跑出去樣本外面的像素」不多，也就是扣除樣本的所有像素之後不應該存在的像素不多，結合著色率和誤差率就可以大致知道目標和樣本的吻合程度高低，進而推估目標圖形的存在可能性。

目前以經驗法則算出來，著色率18%以內，誤差率25%以內就真的存在目標圖形。

輸出步驟改良：

3.0版本之前是輸出一個判定矩陣，接下來轉譯該矩陣變成頻率跟分貝，然後再轉換格式變成Table，最後再輸出成.xlsx檔案。新版輸出會直接以目標格式輸出，再以MATLAB內建的array2table函式進行轉換，再輸出.xlsx檔案即可。

Key Pseudo Code:

for (every target point):

detect blue shape

detect red shape

if (blue shape detected):

record

move right and detect blue shape

else if (red shape detected):

record

move left and detect red half shape

else (no shape detected):

detect blue half shape

detect red half shape

if (blue half shape detected):

record

if (red half shape detected):

record

//move to next target point

1. **未來計畫**

預計於下週五（5/24）與醫生進行第一次測試，測試內容是整個的圖形的偵測，不過目前預計會在下週之前完成八種圖形的偵測功能以及重疊圖形的誤差分析，至於「帶有箭頭」的圖形偵測則預計在5月底或是6月初左右的時段完工。

本學期要發的Paper將會著重在影像處理的部分，我們主要負責method的撰寫，若是需要推進到統計的階段則會於下學期進行。

另外我們希望可以再整理程式碼，應該會是判斷標準算法改良，或是以可能性大小去判別圖形的存在性或種類，或許出來的結果會比現在的更加準確。

1. **後話（by 陳）**

時間過超快已經第十三周了我的天，前面（好像）（有點）混過頭了，雖然早就看到所有事情都會在擠在期末的未來但結果還是一天混過一天，現在報應來了。

很多方面來說真的很感謝柏睿幫我處理那些我不是很想做的事情，例如額外的research跟~~禮拜五早上起床meeting~~，讓我省下了很多心力去做其他的事情~~（例如睡飽一點）~~。

Code的剩餘部分主要是因為我還沒截圖剩下的樣本，而且半個的圖形在原始位置的時候其實是有偏移的，我還在思考樣本該截一個正方形區域還是長方形區域。不過主架構都已經結束了，如果我沒在混的話或許還可以再改良演算法變成權重分數去計算吻合程度，不過那個部分比較麻煩就是了，雖然想法有是有。

總之，計畫內的事情（以現有演算法對指定圖形進行偵測）我一定會如期完成，可以的話應該會調控在6月7日之前就完成，畢竟Paper某種程度上是我要寫，而且也要final presentation了。