

商家與圖片配對研究

凌杰甫

國立中央大學資訊工程學系
Kingjason0414@hotmail.com

張嘉惠

國立中央大學資訊工程學系
chiahui@g.ncu.edu.tw

摘要

近年來，地理位置服務(location-based services)的使用人口伴隨著行動化裝置的普及而快速成長。因此地圖中的 POI(Points of Interest)資料庫更新尤為重要。相較單純以文字敘述店家，透過商品或是店家實際圖片的呈現更能引起使用者的興趣。由於大部分 POI 圖片需要以人工的方式去新增，為了減少時間與人力消耗，本文提出新穎的作法可從網路上的資源尋找可以與商家 POI 配對的圖片與相關資料。我們利用網路爬蟲收集店家資訊，並從中萃取店家的相關特徵資訊，再從包含店家名稱的網頁中，收集相關圖片，最後從圖片中挑選相關資訊最符合特徵者作為店家配對的圖片。

關鍵詞: crawler; Image Retrieval; POI and image pairing

1. 緒論

根據思科視覺網路指標行動預測 2020 年全球將有七成人口成為行動用戶[1]，事實上在 2015 的夏天，Google 在全球行動平台搜尋次數已正式超越 PC 平台[2]，這些訊息指出行動裝置正在快速改變著我們生活的世界。因應這樣人手一機的年代，手機的相關的服務與應用也越來越受歡迎，其中最常被使用的就是地理位置服務(location-based services)。

利用 POI(Point Of Interest)配合電子地圖，這些以使用者為主的地理位置服務(如 Google map、百度地圖等)逐漸取代過去紙本地圖，也讓人們的生活更加便利。使用者只要輕鬆地輸入關鍵字，即可透過行動裝置隨時隨地獲取周遭的資訊，快速地安排行程。但隨著各家的服務如雨後春筍般地不斷推出，要如何能吸引使用者的目光成了一個重要的課題。除了良好的使用者體驗與使用者介面外，最重要的就是豐富的訊息。以往 POI(Point of Interest)的資料大多以文字為主，包含商家名稱、地址、電話和營業時間等，若能再加上圖片等資訊，則更能帶給使用者另一種感官上的接觸。相較單純以文字敘述店家，透過商品或是店家實際圖片的呈現更能引起使用者的興趣。

由於目前 POI 資料庫的建立多仰賴於人工標記或群眾外(Crowd sourcing)為主，因此我們希望透過網路上廣大的使用者生產的資源(User Generated

Content)如部落格和社群媒體來處理這個問題。同樣一個 POI，它可能會有十幾個不同網站在介紹它。我們藉由分析這些網站獲取相關的資訊，並減少只從單一網站抓取但卻是錯誤資料的機會。

除了 POI 文字資訊(地址、電話和營業時間)的部份，我們也嘗試從網頁中擷取與 POI 相符的圖片一般會利用電腦視覺技術(Computer Vision)，透過 OCR(Optical character recognition)辨識圖片中的商標或是文字訊息，又如 Google 的相簿功能，辨識照片內的人物、並加上標籤、自動分類，但是許多照片其實並沒有包括文字，且對於商家或商品的辨識，尚無相關的影像辨識方法。如何不經由影像上的分析，找到與可以代表商家的圖片是本文的目標。

事實上，人工判斷可以代表商家的圖片，如果僅以圖片沒有圖片周圍的輔助文字做為判斷依據，也並非容易。以圖 1 (a)而言由於有商家商標圖案人工判斷相對容易，但圖 1 (b)只有商品或是店家內部裝潢卻容易產生混淆。因此照片旁的文字佐證，是尋找商家相關圖片的重要資訊。



(a) 有商標圖片



(b) 無商標圖片

圖 1 網頁中常出現的圖片樣式

由於黃頁資料庫必有的資料只有店家名稱和地址，因此我們希望利用搜尋引擎來擴增商家的描述，並且以每個圖片旁的文字做為圖片描述，最後利用這些描述資料給予圖片分數，作為排列的依據。

我們從資料庫中隨機抽取 1000 筆 POI，並利用每筆 POI 在搜尋引擎所得的查詢網頁筆數分為兩組 Data set，每組 Data set 包含 500 筆 POI，分別代表網路資源的多寡。我們將這兩個 Data set 分別在系統與 Google image search 做測試。在系統方面，我們透過搜尋引擎所得到的 snippets 做斷詞，並

選擇高頻率者做為拓展詞，再利用這些詞去取得網頁中的圖片並給予分數，最後取出高分者並加上來自社群網路的資源。在資源豐富的 Data set 上系統有 8 成的成功率，比 Google image search 高 8%。在資源較少的 Data set 則有 5 成的成功率。

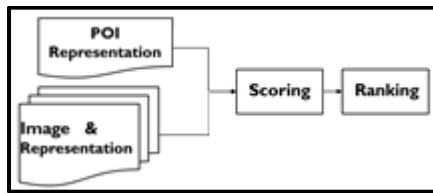


圖 2 構想 Idea

本篇論文架構共分五節，第一節為緒論，說明研究動機。第二節為相關研究，關於如何取的店家資訊與網頁圖片，第三節為方法，會詳細介紹系統內部模組如何運作，第四章為實驗環境，第五章為結論及未來方向。

2. 相關研究

圖像檢索可以分為透過關鍵字查詢的系統以及以圖像內容為基底(Content-based Information Retrieval)兩種[8]。前者利用使用者輸入的關鍵字與圖像的註解和描述進行文字上的比對，回傳可能的圖像，仰賴對圖片的註解和描述，通常需要大量人工去建立標記。而後者透過圖像的特徵值去進行比對，找出相似的圖像，但可能存在語意上的鴻溝，需要額外去處理。因為以上的問題與挑戰，也有人提出將兩者混和並應用於醫學上，利用文字標註圖像的意義並結合擷取圖像的特徵，讓系統能有更好的效能。

自動圖片標示(Automatic Image Annotation; AIA)是圖片檢索的重要一環[4]。自動標示的流程大致上可分為三個步驟，取得圖片語意、特徵萃取、與標示。它的優點在於提高效能與減少使用人工所花費時間，但因為圖片與文字之間的差異，在研究上仍然有不少的挑戰。

應用圖片周邊的文字來代表圖片內容則是運用網路資源的另一種做法，例如 Shen 等人提出一個新的圖像表示法 Weight ChainNet，針對圖像在網頁原始碼中的 Title、Alt、網頁上的相關文字與網頁 Title 建立的詞彙鍊，這些詞彙鍊將在查詢時被使用到[5]。

除了文字上的方式，Srihari 與 Burhans 在 1994 年發表的論文，針對圖像與周圍文字進行研究。首先對文字部分以 Lexical Database 進行語意分析與分群，圖像則以圖像辨識找出物件，最後搭配 CBCG(Caption-based Constraint Generator)，一種以物件之間關係、特徵與位置去選擇可代表語意的詞彙的方法，找出圖像的含意[6]。

另外，Zhou 等人在 2015 年提出對大量圖片文字配對進行檢索與找尋的實驗，它透過網路爬蟲

收集大量的圖文與周遭文字，而每一個圖片透過整理從文字部份萃取有意義的詞並形成詞的列表，最後利用圖像特徵進行分群，結果發現歸類於相同群的圖會有許多重複出現的詞[7]。Feng 與 Deschacht 也分別在針對 BBC 與 Yahoo 的新聞上的圖文配對進行研究[3]。這些研究也間接指出，雖然圖片與周圍文字並不一定有關係，但在標示上依然有很大的幫助。

3. 方法

本研究的目的是要建構一個自動化系統去更新 POI 資料庫，主要更新的內容為商家的圖片、電話和營業時間。系統分成 8 個模組，POI Information Search 與 POI Image-Bearing Pages crawling 為資料的收集和準備，從資料庫抓取商家相關資料並經由爬蟲收集各個模組會使用到的資料。POI Expansion 利用前面收集的資料作分析並找出更多可用的詞彙。Web page Ranking 會利用得到的新詞彙與地址等資訊來給予這些網站分數，並將分數運用到評分與排序的部份。Images Extraction 為圖片與文字的擷取，透過 DOM(Document Object Model)從網站找到圖片的節點並收集周圍文字訊息，最後生成多筆的圖文配對。Images Extraction from FB 則利用 FB ID 透過 API 對 Facebook 的收集資訊。TEL & Business hour Extraction 對網頁上的文字資訊擷取可能的電話與營業時間。Image Ranking 為圖文配對的評分與排序，並輸出前六名作為該筆 POI 的圖片，最後做資料更新。系統架構如圖 3 所示，接下來將依序介紹每個模組的內容。

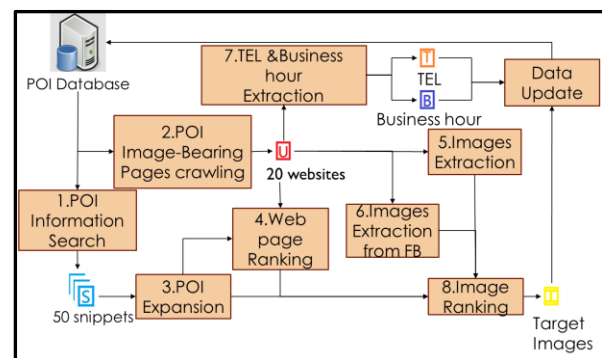


圖 3 系統架構圖

3.1 資料收集與準備

系統每次從資料庫讀取一筆 POI 的資料，內容包含該筆 POI 的名稱、地址、當初收集資料的來源網址與類別。透過組合可以得到「名稱」、「名稱+地址」、「名稱+類別」三種不同的查詢字。

● POI Information Search

為了收集與 POI 相關的資料，我們以「名稱+類別」為查詢字透過網路爬蟲收集 snippets 回來，每個 POI 我們收集前 50 筆 snippets。採用此方法主要原因為搜尋結果受到關鍵字影響很大，在只用「名稱」的情況下很容易得到錯誤資料，以中央大學附近的餐廳「七色鳥」(已歇業)為例，直接使用「名稱」作為查詢字時，得到的資訊為介紹七色鳥的動物生態內容，在判斷圖片上無法提供幫助。如果只使用「名稱+地址」則為因搜尋結果大多為黃頁導致無圖片資訊可以使用。

● POI Image-Bearing Pages Search

在搜尋有圖片網站部份採用「名稱+地址」與「名稱」為查詢字，目的為區別店家擁有分店的情況，同時也可以提高精確度。除此之外，店名是常見的詞彙時，例如：壽司屋，地址能給於相當大的幫助。但有時只使用「名稱+地址」也會有因為條件太嚴苛而沒有網站使用，這時「名稱」的搜尋結果可作為補充。

為了加速系統的效能，每一個查詢字我們使用前 10 筆查詢的網站作為含有圖片的網頁(Image-Bearing Pages)。這些網頁可能來自於結構化的網站(中華黃頁、愛評網)或是非結構化的網站(Facebook、部落格文章)。前者的資料具有固定的格式比較方便收集，而後者寫作方式較為隨意，但圖片含量通常較為豐富。這是因為網路的生態與訊息的快速擴散導致有如此差別，有時候第一手的訊息不是來自於官方網站，反而是網路寫手的業務配合文章。這些文章為了吸引眾人的目光，內容通常有大量的圖片並配合介紹的文字，而這一類型的網站也是我們最好的目標。

3.2 POI Expansion

為了要得到更多 POI 的相關資訊，我們使用 NER 斷詞將 snippets 切成大量的詞彙並計算 TF(Term Frequency)。計算完的詞彙在經過準備好的字典(內容為常出現的無意義的詞彙)過濾後並重新排序，最後再加入 POI 的名稱作為額外的詞彙。透過這樣的方式我們可以取得與 POI 名稱同等重要的詞彙，舉例來說，當 POI 為「Costco」時，我們可以得到「好市多」或「大賣場」等額外資訊，範例如圖 4 所示。因為現在部落格文章寫作方式較為隨意，不太可能每張照片旁都有 POI 名稱的出現，但透過拓展出來的詞彙我們可以收集到更多的圖片。



圖 4 Costco 的拓展詞 好市多

3.3 Web page ranking

我們使用模組 3 所得到的拓展詞和 POI 的地址來幫網站排序，目的是為了給予這些網站區別，有越多相關資訊的網站應該與 POI 有越大的關連性。透過計算網頁中文字部份所包含的拓展詞、店名與地址，每一個拓展詞的一分，店名與地址則為五分。得分越高排序越前面，最後分數 S_{web} 為排序的倒數。

3.4 圖片與相關資料擷取

● Image Extraction

在這一部份我們主要分成兩個部份，第一部份針對一般網頁，如：黃頁與部落格文章。首先使用 Jsoup 與文件物件模型(Document Object Model, DOM)來觀察網頁原始碼的架構，透過找尋節點名稱為 IMG 的節點來定位圖片所在的位置，並經過字串處理取出圖片的網址。然後從該節點開始前後擴張找尋兄弟節點並收集內部的文字資訊。網頁的畫面來看，這些文字大部份是存在圖片附近的。如圖 5 所示

我們替每一張圖片收集文字直到數量達到 100 個字，或是出現 POI 的拓展詞才會停止。當兄弟節點的文字量不足時，我們會往上跳一層到父節點重複收集的動作，以一層層的方式往外擴張。圖片在沒達到收集條件且往上達到一定次數時會被捨棄。

而除了文字的部份，也有一些資源(如 Title 和 Alt 的部份)是網頁上無法直接觀察到。這些文字內容通常比較短，但在圖片辨識上能給予相當大的幫助。不過因為是利用字串比較的方式來使用，有時難免因為一兩個字的不同而失敗。因此我們運用最小編輯距離(Minimum Edit Distance)來估算它與 POI 拓展詞的相似度，當相似度足夠時，我們視為具有同等效力。整理資訊如表 1 圖片相關資訊所示。

表 1 圖片相關資訊

POI Name	Alt	Title	文字資訊
九九滴水洞	九九滴水洞	九九滴水洞	好吃的九九滴水洞...

但並非網站上的圖片都是可以使用的，裡面也會存在網頁美工圖案(網站的按鍵或美編)或是商業廣告，我們可以透過偵測圖片的長寬與比例來予以過濾。當圖片的長與寬都無超過 40px 時，我們視之為美工圖案，而長寬比超過 10 或低於 0.1 則可能為網路廣告。除此之外，我們也整理出一些常出現的廣告網址以作為過濾使用

● Images Extraction from FB

這部份為針對 Facebook 的網頁，因社群網路的流行，許多的公司行號或公家團體它們沒有建立官方網站，但可能會有 Facebook 的社群或粉絲頁。因為是屬於動態網頁的部分，資料沒有辦法使用第一部份的方法，因此我們使用 Facebook 提供的 API 作為接口。

當 20 個 Image-bearing pages 裡存有 Facebook 網頁時會將 Facebook ID 擷取出來，並透過 API 去讀取相簿。當沒有 Facebook ID 時，則利用 POI 名稱在 API 上進行搜尋，取得可能的 Facebook ID，並透過臉書名稱與 POI 名稱是否相同來減少使用錯誤網頁的機率。

● TEL & Business hour Extraction

在這一部份，我們藉由分析網頁中常出現的形式並透過正則表達式(Regular Expression)來取得網頁中商家電話與營業時間。

每一筆 POI 使用抓取圖片的 20 個網頁去擷取電話與營業時間。我們先收集網頁上所有的文字部份，並利用正則表示法擷取資料，最後整理成相同格式方便統計。每一個網頁擷取一筆電話與營業時間。最後每筆 POI 結算所有網頁中擷取的資料並以頻率最高者為答案。測試範例如圖 6。

老媽精緻火鍋	老媽精緻火鍋
02-8772-4775 9↓	營業時間：今日 00:00~01:30、17:30~01:30 2
02-2739-6683 1↓	營業時間：週一至週日，17:30~01:30 3↓
02-2777-5608 1↓	今日 00:00~01:30、17:30~01:30 2↓
201302171750 2↓	17:00~01:00 2↓

圖 6 擷取範例

3.5 Image Ranking

在本節中，我們將給前面收集收集的圖文配對計算分數並排序。計算公式如下：

$$S_{image} = \sum_{i \in expansion} \left(\alpha \times (F_{alt}(i) + F_{title}(i)) + \frac{\beta \times F_{text}(i)}{D_{node(i)} * L_{text}(i)} \right) + S_{web}$$

首先每一筆圖文配對利用之前收集到的拓展詞來計算分數。在文字的部份，我們有兩種來源。第一種為透過節點收集來的文字，第二種為網頁原始碼上的文字，即 Title 與 Alt。我們利用拓展詞在這些文字中有無出現作為評分的依據， $F_{alt}(i)$ 、 $F_{title}(i)$ 、 $F_{text}(i)$ 分別指在 Alt、Title、圖片周圍文字的部份，有的話則算 1 分，反之為 0 分。換言之，當拓展詞出現越多時分數就會越高。另外周邊文字也會因收集文字的多寡而受到影響， $L_{text}(i)$ 為收集文字的數量，若在周邊文字總數越少的情況下，拓展字越多的話，則這圖片應該越重要。 $D_{node}(i)$ 為往父節點移動的次數，移動的次數越多則文字與圖片的距離越遠。 S_{web} 則為模組所得的網頁分數。最後利用這個分數為圖片排序，並取出前六名。

3.6 Data Update

在更新資料方面，我們將檢查 POI 的原始資料來源的網站存活與否。如果已經不存在或沒有資料，則以新的圖片來源網站中最高分者取代。

接下來依序檢查是否有電話、營業時間等資料，若不存在則補上新的資料，最後將最高的前六張圖片作為可能的圖片存入資料庫。

4. 實驗與系統效能

本研究目的在於幫店家尋找可以配對之圖片與店家資料的更新，因此實驗部份也依照這兩個部份進行。

第一階段的實驗為商家圖片配對準確度。因為圖片是否有關連十分受到主觀的影響，因此我們設立一些條件做為判斷的基準。1.當圖片中有明確的商標 2.來源的網站能給於佐證，當以上條件有任意一項成立時，即是有關連。每筆 POI 從收集的圖片中選取最高分前 6 張圖片，並以 Precision at K (P@K) 和 Success at K (S@K) 作為判斷標準。我們也另外設計 UI 介面提供使用者快速標記圖片，如圖 7 標記 UI 介面。



圖 7 標記 UI 介面

第二階段為商家資料的更新，我們將實驗正規化表示法擷取電話與營業時間的成效。我們利用 1000 個 POI 所使用的 20000 個網站，並觀察成功擷取的電話與營業時間準確率與召回率。並觀察 1000 筆 POI 正確擷取的數量。

4.1 資料集

資料集主要來自於 APP「疾疾店家現身」所使用的 Solr 資料庫，總計約有 140 萬筆 POI，可分為 29 個不同的類別。我們從資料庫中挑選出 1000 筆沒有圖片資訊的 POI，並取出名稱、地址、類別與來源 URL 來使用。這 1000 筆 POI 我們以 Google search 所得到的搜尋結果筆數作為網路資源多寡的判斷依據，經過排序後取前 500 筆作為 Data set 1 為網路資源豐富的組別，其餘作為 Data set 2。

單使用資料庫的資料不夠，為了擴充 POI 的資料，我們透過網路爬蟲進行額外的收集。每一筆 POI 使用 3 種不同查詢詞「名稱」、「名稱類別」和「名稱 地址」在搜尋引擎上查詢。在系統上原先使用線上即時查詢以確保得到的是最新的訊息，但查詢會受到次數的限制，每 3 分鐘只能進行一次。為了加速實驗的流程，我們將查詢後的結果以文字保留下。內容包含網頁 Title、網址與 snippet，以方便多次實驗。

4.2 效能評比

在校能方面，我們使用 S@K(Success at rank K)和 P@K(Precision at rank K)來計算與 POI 的相關性。S@K 為對一個 POI 有 K 張圖，若有一張與 POI 有關連則為 1，反之為 0。P@K 為一個 POI 有 K 張圖片中與 POI 相關的圖片有幾張， $\delta(i)$ 為相關性的分數，有相關為 1 分，反之為 0。P@K 計算公式如下：

$$P@K = \frac{\sum_{i=1}^k \delta(i)}{k}$$

每一個 POI 分別計算 S@1、S@3、S@6 和 P@6。最後每一項將所有的 POI 取平均代表效能。

4.3 系統成效

我們以 Google image search 作為比較對象，為保持公平性以兩種不同的查詢字進行查詢。第一種為 POI 名稱，第二種為 POI 的名稱與地址。

• Data set 1

在 Data set 1 中大多數的 POI 多為人氣的餐廳、景點與商家為主。這一類的 POI 在網路上有相當多的資源，如：部落格文章、Facebook 粉絲頁等。這些都蘊藏很多圖片資源可以使用。

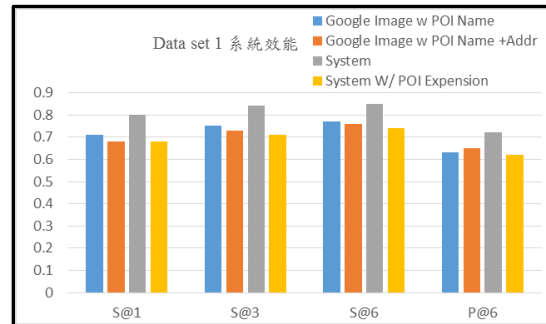


圖 8 Data set 1 系統成效

如圖 8，在 S@1 時，系統的成功率有 80%，在沒有使用拓展詞時成功率為 68%。會有如此差異主要是圖片的周圍不一定有完整 POI 名稱的出現，由其在部落格文章上時常發生。另一方面 Google Image Search 使用 POI 名稱當作查詢字時可以得到大量相關的圖片，但當 POI 名稱為常見或有特殊含意的詞時，如：壽司屋或白胡子時，容易抓錯圖片。

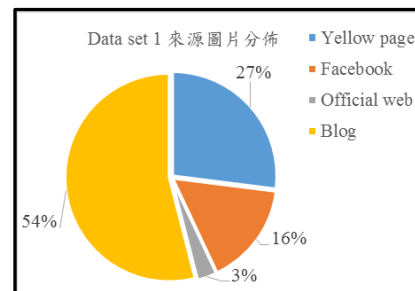


圖 9 Data set 1 圖片來源

從圖 9 來源分析來看最多的圖片來源是部落格文章，其次是黃頁與 Facebook。主要原因是現在社群媒體的流行與商業上的運用，人氣高的店家會有大量的相關網路資源，搜尋排序也較為優先。

• Data set 2

Data set 2 從 POI 的類別分析來看，這一類大多為公家單位、教學單位或公司行號為主。

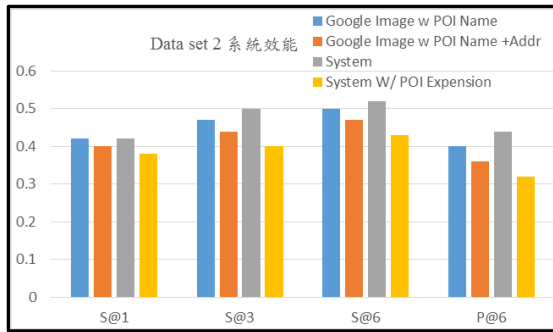


圖 10 Data set 2 系統成效

如圖 10，在 S@1 表現為 42%。這是因為教學與公家單位的網路資源比較少有部落格文章介紹，從圖 11 可以發現，它的網路資源主要來自於黃頁、官方網站以及 Facebook 粉絲頁，而一般在黃頁上多以文字訊息為主較難取得圖片。

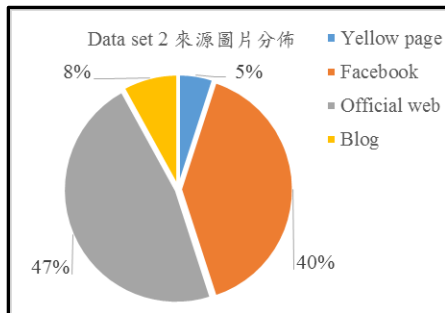


圖 11 Data set 2 圖片來源

另一項效能降低的原因為公司行號的商標圖片如圖 12，一般官方網頁上的商標圖片多為長條形，在判斷是否為廣告時容易被誤判。



圖 12 誤刪除的公司商標

4.4 系統模組比較

這個章節，我們將比較各個模組對於系統成效的影響。我們使用 Data set 1 與 Data set 2 共 1000 筆 POI，並使用 S@K 與 P@K 來測試效能，結果呈現如圖 13。

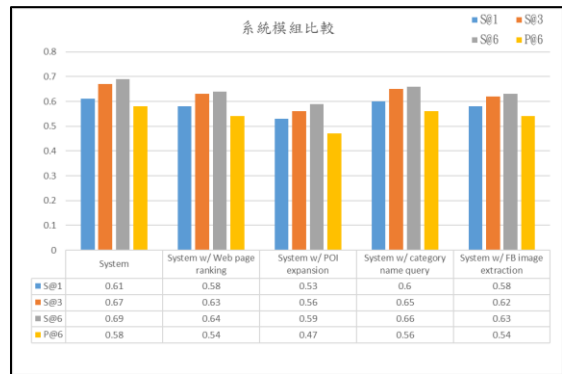


圖 13 系統模組比較

這些模組主要影響兩種方面，第一種為圖片的排序，第二種為圖片抓取。Web page ranking 主要影響第一種，POI expansion 的使用與搜尋有無使用類別則第一種與第二種都有影響，最後有無使用 FB 影響第二種。每一次我們移除一組模組，並重新收集 1000 筆 POI 的圖片，並觀察效能。

4.5 Images-Bearing pages 數量選取

系統在採取圖片時，當 images-bearing pages 越多，我們所能採取到的資料就越多。但每多使用一個網頁則需要消耗更多的時間，但效果提升卻是不一定的。而這一部份使用 100 筆 POI 對網頁的數量與系統的效能做一個測試，並以 S@6 與 P@6 評比。

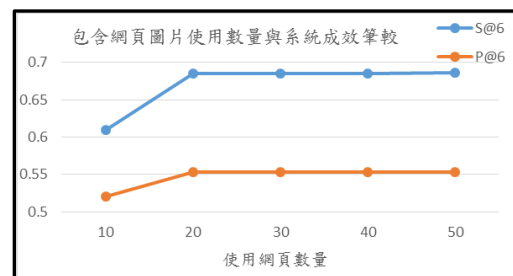


圖 14 包含網頁圖片使用數量與系統成效比較

從圖 14 可以發現當 images-bearing pages 選取 Top 10 個網頁時，已能擷取出足夠好的圖片，使用更多的網站效能也不會有顯著的提升。

4.6 公式的參數調整

在排序的公式中， α 代表為圖片相關資訊的部分 (Title 與 Alt) 的參數， β 為周圍文字的部分的參數，同時 $\alpha + \beta = 1$ 。

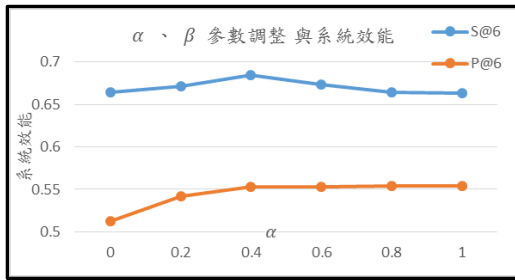


圖 15 α 、 β 參數調整與系統效能

總共實驗 100 個 POI，並逐次調整 α 以觀察 S@6 與 P@6 的變化。從圖 15 可以發現當把 α 的比重拉高時，系統準確率的會上升，但沒有圖片相關資訊的圖片分數則會降低，這樣反而會喪失一些重要的圖片。另一種情況為網頁中非相關的圖片的也有可能誤植相關資訊(Alt, Title)，這樣也會使得非正確圖片排序太前面。效果最好為 $\alpha=0.4$ 、 $\beta=0.6$ 。

4.7 電話與營業時間擷取實驗

我們將先前 1000 個 POI 所使用的 Image-bearing pages，總共 20000 個網頁，用來擷取電話與營業時間。首先，因為網頁上不一定有電話或營業時間，所以我們利用網頁上的文字來預估有這些資料的網站數量，當文字中有「電話」、「營業時間」或「開放時間」時，我們便判斷該網站有資料，並嘗試擷取。從全部的網頁中，總共取得 14502 筆有電話的網頁並擷取出 12987 筆 POI 與電話的配對。在營業時間方面，共取得 3621 筆網頁與 1596 筆配對。我們運用 precision 與 Recall 來評比。

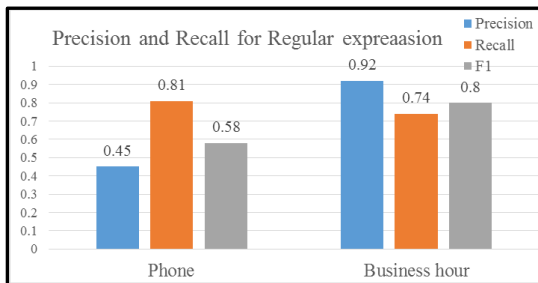


圖 16 電話與營業時間擷取成效

如圖 16 所示，電話的準確率只有 0.45，但召回率有 0.81。而營業時間準確率有 0.92，召回率為 0.71。會造成這樣的差別主要因為電話的格式比營業時間較為多樣化，因此抓取回來的資訊量比較多，但也容易發生錯誤。相對的營業時間格式較為單純，準確率較高，但有相關資料的網站筆營業時間少很多。除此之外，會碰到的問題有一些網頁開始將資料以圖片方式呈現，而造成再判斷上認為該網站有資料卻無法取得，或網站上有「電話」與「營業時間」的文字但沒有資訊。最

後統計 1000 筆 POI 成功配對狀況，成功擷取電話有 642 筆，營業時間有 423 筆。

5. 結論與未來工作

在使用電子地圖找尋 POI 時，若得到 POI 資訊能包含圖片能更增進使用者的體驗。但要能得到商家的相關圖片並不容易，大部份方法為透過圖片搜尋引擎的結果或透過人工採集，前者無法確定結果，而後者相當費時與費力。因此我們希望能建立一個系統並透過自動化的方式來替 POI 配對可以信任的圖片。

在本論文，我們提供一個方法來結合網路上眾多的資源來幫 POI 配對合適的圖片。透過從網頁中直接擷取與社群媒體的幫助，在一般 POI 上能與 Google Image Search 結果接近，而在受歡迎的 POI 上，我們能得到更精確的圖片。

在運行上仍有需要改進的地方。首先關於商家商標的部份，若能使用 OCR (Optical character recognition) 來幫忙辨識圖片，則更能找出圖片與商家的關係。在圖片抓取部份，我們可以在擷取得同時觀察圖片所在的位置，通常在 Head 部分會是商標的部份，而 Body 則是主要內容部分，會是重要圖片來源，最後 footer 則可能多廣告，藉由分析來給圖片更多的 feature 來使用。而在圖文配對的，我們也可以採用透過 cosine similarity 或 TF-IDF 來進行運算，並觀察排名的變化。最後，我們希望能建立一個有用系統，重續增進其效能。

參考文獻

- [1] 《思科視覺網路指標》行動預測未來全球有 7 成人口將成為行動用戶，http://www.cisco.com/c/zh_tw/about/news-center/20160218.html
- [2] DESCHACHT, Koen, et al. Text analysis for automatic image annotation. In: *ACL*. 2007. p. 1000-1007.
- [3] FENG, Yansong; LAPATA, Mirella. Automatic Image Annotation Using Auxiliary Text Information. In: *ACL*. 2008. p. 272-280.
- [4] GHOSH, Sayantani; BANDYOPADHYAY, Samir Kumar. A Tutorial Review Of Automatic Image Tagging Technique Using Text Mining. *International Journal Of Research In Engineering And Technology* ISSN, 2319-1163.
- [5] SHEN, Heng Tao; OOI, Beng Chin; TAN, Kian-Lee. Giving meanings to WWW images. In: *Proceedings of the eighth ACM international conference on Multimedia*. ACM, 2000. p. 39-47.

- [6] SRIHARI, Rohini K.; BURHANS, Debra T. Visual semantics: Extracting visual information from text accompanying pictures. In: *AAAI*. 1994. p. 793-798.
- [7] ZHOU, Ning; FAN, Jianping. Automatic image–text alignment for large-scale web image indexing and retrieval. *Pattern Recognition*, 2015, 48.1: 205-219.
- [8] 圖 像 檢 索
<https://zh.wikipedia.org/wiki/%E5%9C%96%E5%83%8F%E6%AA%A2%E7%B4%A2>