

國立政治大學圖書資訊與檔案學研究所

碩士論文

Master Thesis

Graduate Institute of Library, Information and Archival Studies

National Chengchi University

基於自動情境標註之圖像檢索工具發展與數位

人文應用研究

Developing an Image Retrieval Tool based on Automatic Context

Annotation for Digital Humanities Research

指導老師：陳志銘 博士

Adviser: Dr. Chih-Ming Chen

研究生：趙映翔

Author: Ying-Xiang Zhao

中華民國一一〇年七月

July, 2021

## 謝辭

兩年的碩士生涯歲月如梭，初次步入政大校園至今遇到的人事物，使我成長許多，如今碩士論文能如期完成，要感謝的人太多太多。

感謝我的父母支持我的決定，使我能在學業上無後顧之憂地繼續精進，求學過程中，我不曾知道學習的意義為何，每次學習都是為了考試，如今我開始了解學習的意義在於自我成長，學習前人的智慧，提出自己的想法及聆聽他人的建議並改進，當學習不再是為了考試時，我才開始懂得學習。

感謝指導教授陳志銘教授，從第一次踏入老師辦公室與老師的談話開始，就深深地感受老師對於研究的熱忱，或許就是這個契機，促使我積極開始提升自我學習能力，從當初對於深度學習地懵懂無知，至今開始有了自己的見解。此外，非常感謝老師在我在論文方向出現迷惘時，仍不辭辛勞地細心指導，老師的支持與鼓勵是我能在兩年之內順利完成碩士學位論文的最大動力。

感謝口試委員，張道行教授與張素玠教授，於計畫書口試時，給予了許多論文的修改建議，並且在最後口試時，教授們也點出了許多論文目前不足的地方，讓我能進一步修正，以完善更為全面的論文內容。

感謝 DLLL 實驗室的大家，張鐘學長、明娟學姊、志帆學長及布丁學長，對於我的系統設計與實驗設計上提供了相當多的協助。感謝張鐘學長，在系統設計上的幫助，總是在我陷入困難時拉我一把，而我也在與學長討論的過程中學習到很多。感謝明娟學姊，在我實驗設計遇到困難時，提出犀利的建議使我頓悟。感謝志帆學長，與學長聊天的過程中，使我對系統架構上的思緒能夠逐漸清晰。感謝布丁學長，在情境分類模型上，學長總是不遺餘力地教導我，從資料整理到模型評估上，使我建立一個嚴謹的模型建構能力。

感謝同窗的夥伴們，冠霖、仙姁、宇涵、佩淳、雅竹、俊佑與阿貴，碩一時與大家一起打籃球閒聊的日子，回想起來總是那麼地簡單快樂，唯一的遺憾是遇上疫情，無法一同到日本參加 IIAI，但後來疫情趨緩，實驗室舉辦的金山兩天一

夜遊以及單車木柵到淡水行，創造了不少我們的回憶，之後的日子裡忙於系統開發與論文寫作，在這過程中夥伴們幾個簡單的關心，使我有動力繼續撐到最後，很感謝大家。感謝嘉培學長與永慈學妹，標註圖片的過程很枯燥，但多虧有你們的幫助下，才能完成如此龐大的標註量。

感謝政大 Google Developer Student Club 社團的朋友們，陳俊憲、唐嘉偉、黃瑜萍與陳德璋，那段熱血的日子裡，每周二的社課結束後，與大家在社辦中標註鳥的照片，跟著唐嘉偉建構模型，為了參與 Google 舉辦 Solution Challenge，甚至還做了一個網站，在程式學習的旅途中遇到你們，是我最大的幸運。

寫到這裡，碩士兩年的日子將畫上句點，衷心感謝一路上給予我支持的大家，因為如此，讓我在碩士生涯中能夠專注地進行研究與學習，順利完成碩士學位，期許能夠在未來繼續與大家共患難，誠心感謝。

趙映翔 謹誌於政大

中華民國 110 年 7 月



## 摘要

「圖像檢索」在資訊蓬勃發展的現代，已經成為數位人文研究的重要方式之一。而影響傳統「基於文本的圖像檢索工具(Text-Based Image Retrieval, TBIR)」之圖像檢索效能的主要問題，為人工所注入代表圖像後設資料(metadata)與使用者所下檢索詞之間的語意鴻溝(semantic gap)。隨著電腦視覺技術快速發展而衍伸出的自動圖像標註(automatic image annotation)，由機器為其自動添加後設資料以降低的語意鴻溝。然而自動圖像標註的物件標註僅能找到具有該物件特徵，對於使用者的圖像檢索及圖像理解的幫助有限，進而促成本研究探索「自動情境標註」為減少圖像情境與人之間的語意鴻溝，並發展出得以有效輔助人文學者進行圖像檢索及圖像解讀之數位人文工具。

因此，本研究發展出「基於自動情境標註之圖像檢索工具(Image Retrieval Tool Based on Automatic Context Annotation, IRT-ACA)」。該系統的核心技術採用 Mask R-CNN、TF-IDF 及 SVM，主要目的為圖像中的實體物件識別，以及抽象的情境識別，並將所得之數據以標籤化形式提供使用者用於圖像檢索與瀏覽，讓使用者得以可以快速萃取數位圖像中的實體物件以及抽象情境之訊息。進而促進人文學者更有效率地解讀圖像情境。

為驗證本研究發展之 IRT-ACA 是否有助於人文學者進行圖像解讀，本研究採用實驗研究法之對抗平衡設計，將實驗對象分為兩組，根據不同的系統使用順序來依次操作「IRT-ACA」與「基於文本的圖像檢索工具(Text-Based Image Retrieval, TBIR)」來完成檢索任務學習單。並透過行為歷程記錄來完整記錄實驗對象的系統操作行為、科技接受度問卷來反映實驗對象對於系統的實際感受，以及半結構式訪談來瞭解實驗對象的想法與建議，透過多種方法進行交互驗證，以瞭解本研究發展之 IRT-ACA 與 TBIR 在自動情境標註之準確度、解讀圖像情境之成效以及科技接受度上的差異。

研究結果發現：第一，IRT-ACA 的自動情境標註準確度已足以有效輔助使用者解讀圖像情境；第二，使用 TBIR 與 IRT-ACA 在解讀圖像情境之成效上達顯著差異，並且 IRT-ACA 顯著優於 TBIR；第三，使用 TBIR 與 IRT-ACA 在整體科技接受度上達顯著差異，並且 IRT-ACA 顯著優於 TBIR，但其中的系統易用性未達顯著差異。從訪談分析中顯示，實驗對象對於兩個系統的操作難意度及使用流暢性上均感到滿意，因此給予系統易用性分數差異不大；第四，IRT-ACA 的標籤型檢索比起自由下達檢索詞的檢索更能促進實驗對象的檢索意願；第五，使用 IRT-ACA 高分組使用者之檢索行為更充分使用到所有檢索功能；第六，IRT-ACA 使用者之查看圖像至筆記紀錄之轉移率高於 TBIR。

關鍵字：數位人文、圖像檢索、語意鴻溝、物件標註、情境標註、深度學習、自動圖像標註、Mask R-CNN、TF-IDF、SVM、行為分析



## Abstract

Image retrieval has become one of the significant approaches in digital humanities research in the digital age. The main problem affecting the performance of Text-Based Image Retrieval (TBIR) is the semantic gap between the manually determined metadata for images and the users' search terms or keywords. With the rapid development of computer vision technology in recent years, automatic image annotation developed by machine learning schemes can reduce the semantic gap between humans through automatically adding metadata based on the identified image objects' tags. However, the object tags determined by automatic image annotation can only find the characteristics of image objects, which is of little help to users' image retrieval and image comprehension because they are still too low level from human's perspectives. It prompted this research to develop automatic context annotation as a digital humanities tool that can effectively assist humanities scholars in image context interpretation by reducing the semantic gap between the subject of the image context and humans.

Therefore, this research developed an Image Retrieval Tool Based on Automatic Context Annotation (IRT-ACA). The core technology of the tool is Mask R-CNN, TF-IDF, and SVM, which aims to identify physical objects and abstract contexts hidden in images and provide users with more rich metadata in the form of object and contextual tags for image retrieval and browsing, so that users can quickly extract needed information from images, thus facilitating the more efficient interpretation of image contexts by humanists.

To verify whether IRT-ACA developed in this research is beneficial to humanities scholars in image interpretation, this research utilized a counterbalanced design of the experimental research method to examine the research questions. Users were divided

into two groups and operated the IRT-ACA and the TBIR tool alternately to complete the two designed image retrieval tasks. Besides, the behavioral history recorder was used to record the system operation behavior of the users using the IRT-ACA and the TBIR completely. The technology acceptance questionnaire was used to reflect the actual feelings and perception of the experimental subjects towards the two systems. The semi-structured interview was used to understand the thoughts, ideas, and suggestions of the users who alternately used the two image retrieval systems.

The research results are summarized as follows. First, the accuracy of the automatic context annotation of IRT-ACA was sufficient to interpret the image context effectively. Second, there was a significant difference in the effectiveness of interpreting the image context between TBIR and IRT-ACA as well as IRT-ACA is significantly superior to TBIR. Third, there was a significant difference in the overall technology acceptance and perceived usefulness between TBIR and IRT-ACA as well as IRT-ACA is significantly superior to TBIR, but there was no significant difference in the perceived ease of use between TBIR and IRT-ACA. The analysis results show that users were satisfied with the ease of operation and the smoothness of using the two systems, so the difference in the scores of perceived ease of use was not significant. Fourth, the labeled retrieval provided by the IRT-ACA promoted the research subjects' willingness to retrieve more than the free retrieval of using keywords. Fifth, the retrieval behaviors of the IRT-ACA users with high image interpretation performance made full use of all retrieval functions, including contextual tags' search, object tags' search, full-text search, and title research. Sixth, the transfer rate from viewing images to take notes of the users who used IRT-ACA is higher than that of users who used TBIR.



Keywords: Digital humanities, Image retrieval, Semantic gap, Object annotation, Contextual annotation, Deep learning, Automatic image annotation, Mask R-CNN, TF-IDF, SVM, Behavioral analysis





# 目次

目次.....	i
圖目次.....	iii
表目次.....	iv
第一章 緒論.....	1
第一節 研究背景與動機.....	1
第二節 研究目的.....	3
第三節 研究問題.....	4
第四節 研究範圍與限制.....	4
第五節 名詞解釋.....	5
第二章 文獻探討.....	7
第一節 圖像檢索系統發展.....	7
第二節 自動圖像標註.....	8
第三章 系統設計.....	11
第一節 系統設計理念.....	11
第二節 系統架構.....	12
第三節 系統使用者介面.....	14
第四節 系統開發環境與工具.....	17
第五節 系統操作說明.....	19
第四章 研究方法與實驗設計.....	21
第一節 研究架構.....	21
第二節 研究方法.....	22
第三節 研究對象.....	24
第四節 研究工具.....	24
第五節 實驗流程.....	26
第六節 資料處理與分析.....	28
第七節 研究步驟.....	32
第五章 實驗結果分析.....	34
第一節 實驗對象基本資料.....	34
第二節 發展之 IRT-ACA 自動情境標註之準確率分析.....	36
第三節 使用者使用 TBIR 與 IRT-ACA 解讀圖像情境成效之差異分析.....	41

第四節 使用者使用 IRT-ACA 與 TBIR 系統之科技接受度差異分析.....	43
第五節 使用者使用 IRT-ACA 與 TBIR 之系統操作行為歷程紀錄分析.....	44
第六節 半結構式訪談質性資料分析.....	52
第七節 綜合討論.....	55
第六章 結論與建議.....	61
第一節 結論.....	61
第二節 IRT-ACA 之系統改進建議.....	64
第三節 未來研究方向.....	65
參考文獻.....	67
附錄.....	71
附錄一 訪談大綱.....	71
附錄二 IRT-ACA 科技接受度問卷.....	72
附錄三 TBIR 科技接受度問卷.....	74



## 圖目次

圖 1-1 圖像內容分層 .....	3
圖 2-1 圖像檢索系統發展史 .....	7
圖 3-1 具自動情境標註之圖像檢索工具系統架構圖 .....	12
圖 3-2 具自動情境標註之圖像檢索工具使用者介面 .....	15
圖 3-3 點石筆記使用者介面 .....	16
圖 3-4 檢視圖像詳細資料之使用者介面 .....	17
圖 4-1 本研究之研究架構圖 .....	21
圖 4-2 本研究之實驗流程圖 .....	26
圖 4-3 研究實施步驟圖 .....	32
圖 5-1 採用 IRT-ACA 之低分組使用者之系統操作行為轉移序列圖 .....	50
圖 5-2 採用 IRT-ACA 之高分組使用者之系統操作行為轉移序列圖 .....	51

## 表目次

表 5-1 研究對象背景說明 .....	34
表 5-2 IT、TI 兩組實驗參與人數.....	35
表 5-3 IT 組依據題目分組及實驗參與人數 .....	35
表 5-4 TI 組依據題目分組及實驗參與人數.....	35
表 5-5 訓練集之圖像情境分類與數量 .....	36
表 5-6 訓練集之圖像物件分類與數量 .....	37
表 5-7 不同機器學習演算法在 K-fold Cross-Validation=10 之下分類結果比較表 .....	39
表 5-8 使用 SVM 進行情境分類之 Confusion Matrix .....	40
表 5-9 圖像檢索任務學習單之評分標準 .....	40
表 5-10 分別使用 IRT-ACA 及 TBIR 系統完成兩種圖像檢索任務之解讀圖像情 境成效敘述統計分析.....	41
表 5-11 使用兩系統解讀圖像情境成效之曼惠特尼 U 檢定 .....	41
表 5-12 使用 TBIR 與 IRT-ACA 系統使用者科技接受度敘述統計分析.....	42
表 5-13 使用 TBIR 與 IRT-ACA 系統使用者科技接受度之曼惠特尼 U 檢定分析 結果.....	43
表 5-14 實驗對象使用「IRT-ACA」及「TBIR」之系統操作行為說明.....	44
表 5-15 使用 TBIR 進行圖像檢索任務之各項行為統計表 .....	45
表 5-16 使用 IRT-ACA 進行圖像檢索任務之各項行為統計表.....	46
表 5-17 IRT-ACA 高、低分組實驗對象之組別人數統計表.....	48
表 5-18 IRT-ACA 高、低分組實驗對象解讀圖像情境成效之曼惠特尼 U 檢定 .....	48
表 5-19 IRT-ACA 高、低分組之任務單結果之敘述統計.....	48

表 5-20 IRT-ACA 高、低分組實驗對象之圖像檢索行為歷程統計表.....49

表 5-21 使用兩系統進行任務檢索之檢索次數行為統計表 .....55

表 5-22 TBIR 與 IRT-ACA 使用者在檢索任務下的圖像檢索成效分析結果歸納56

表 5-23 使用兩系統進行主題檢索之筆記次數行為統計表 .....56

表 5-24 採用 TBIR 與 IRT-ACA 輔以圖像檢索使用者之科技接受度分析結果歸  
納.....58



# 第一章 緒論

本章分成五個小節，第一節為研究背景與動機；第二節為研究目的；第三節為研究問題；第四節為研究範圍與限制；第五節為名詞解釋，依序說明如下：

## 第一節 研究背景與動機

近年來隨著資訊科技的快速發展，過去依賴紙本的人文研究方法產生改變，人文學者對於透過數位資料庫進行數位人文研究的資訊檢索需求與日俱增，基於資訊科技輔助人文研究的數位人文開始蓬勃發展，數位人文儼然已成為一門新興學科(Schreibman, 2012)。基於數位人文研究需求，人文學者對於歷史文本及圖像研究，會採用具文字及圖像瀏覽及檢索功能的典藏系統輔以研究，因此促使越來越多的博物館、圖書館及學術研究機構開始將典藏的紙本進行數位化，並透過網際網路提供研究者進行研究使用，例如數位化文化遺產典藏、歷史報紙典藏等(Agosti, Ferro, Orio, & Ponchia, 2014; Chen & Tsay, 2017; Lorang, Soh, Datla, & Kulwicki, 2015)。值得重視的是，除了文字檢索之外，隨著圖像資料輔以數位人文研究需求的與日俱增，圖像檢索的重要性開始受到重視(Beaudoin & Brady, 2011)。以使用者的角度出發來探討圖像檢索系統(image retrieval system)如何支援圖像檢索的研究中，Beaudoin (2014)解釋了使用者的圖像使用(image use)係指專業領域之群體，在檢索和選擇圖像之後如何使用圖像，明確定義了專業人員在工作中使用圖像的幾個原因，包括拓展知識、激發靈感和回憶資訊。

影響傳統「基於文本的圖像檢索工具(Text-Based Image Retrieval, TBIR)」之圖像檢索效能的主要問題，為人工所注入代表圖像後設資料(metadata)與使用者所下檢索詞之間的語意鴻溝(semantic gap) (Hare, Lewis, Enser, & Sandom, 2006; Smeulders, Worring, Santini, Gupta, & Jain, 2000)。因此，許多研究開始採用機器學習來降低語意鴻溝，包括使用支援向量機(Wan & Chowdhury, 2003)、決策樹(Liu,

Zhang, & Lu, 2008)及 K-means (Murthy, Vamsidhar, Kumar, & Rao, 2010)行圖像分類來輔助檢索，良好的圖像分類準確率，大幅提升圖像檢索系統的效能。此外，近幾年來圖像檢索研究開始專注於自動圖像標註(automatic image annotation)，Zhang 等人(2012)定義自動圖像標註是基於分析圖像的內容，由機器為其自動添加後設資料的過程，目的是為了讓描述圖像的後設資料更具客觀性，以及降低人工注入圖像低階特徵後設資料與使用者認知之間的語意鴻溝。此舉改善了過去手動標註需要大量人力投入、成本高昂、專業知識背景及思維模式所產生的認知差異(Cheng, Zhang, Fu, Tu, & Li, 2018)等缺點。因此，基於卷積神經網路的物件辨識(object recognition)描述圖像的自動圖像標註技術(automatic image annotation)，成為了這幾年的主要研究方向(Murthy, Maji, & Manmatha, 2015)，透過物件標註為圖像注入後設資料，也促使基於文字的圖像檢索可以進行圖像上的物件檢索。

然而，數位人文更注重於透過抽取圖像中的高階資訊進行圖像解讀(Llamas, Lerones, Zalama, & Gómez-García-Bermejo, 2016)，因此可藉由 Bradshaw (2000)提出了語意為基礎的圖像檢索(Semantic-Based Image Retrieval, SBIR)的概念，逕行機率計算來提取圖像中的語意。此外，Zhang (2006)的研究提出 SBIR 的語意分層概念，將圖像的語義分為原始資料層、特徵層，以及語義層，而本研究以數位人文角度，把圖像資訊分成特徵層(feature layer)、物件層(object layer)，以及情境層(context layer)，如圖 1-1 所示。其中特徵層為圖像的最低階特徵，例如顏色、紋理、空間關係及形狀等；物件層為圖像中的單一物件，例如人、動物或物品等；情境層則呈現圖像整體所表達的抽象主題性。



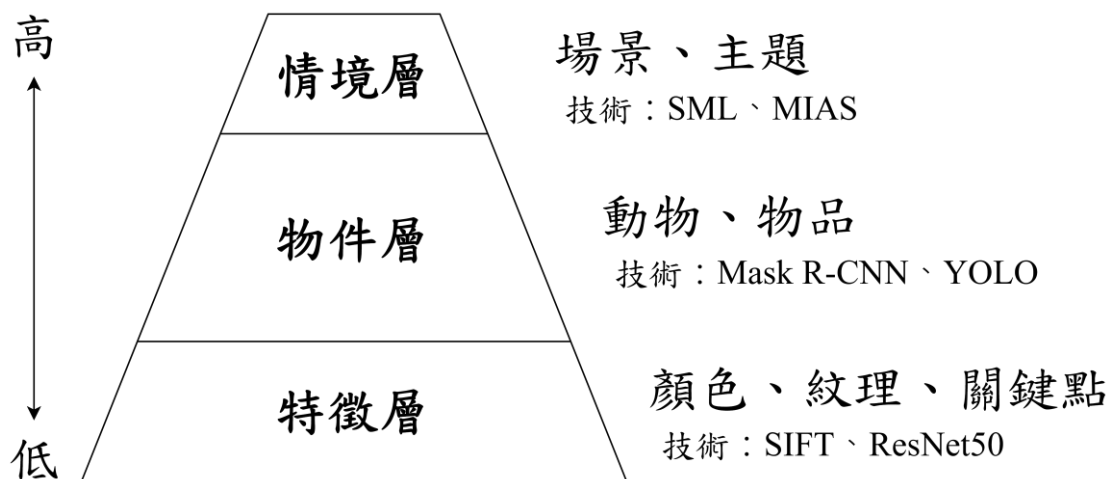


圖 1-1 圖像內容分層

(資料來源：本研究整理)

為了彌補自動圖像標註的物件檢索方式僅止於找到具有該物件特徵的圖像限制，以及輔助使用者對於圖像理解的幫助有限。近幾年來，藉由對於圖像情境分類發展之自動情境標註技術(Li, Tang, 2015; Burdescu, Mihai, Stanescu, & Brezovan, 2013)，除了可以基於既有的圖像物件標註進行圖像檢索外，亦可以進而依據圖像情境識別結果，檢索並分析龐大的圖像資料庫，為人文學者提供更多的解讀圖像有用資訊(Terras, 2012)。

## 第二節 研究目的

基於研究背景與動機，本研究旨在發展支援數位人文研究之「基於自動情境標註之圖像檢索工具(Image Retrieval Tool based on Automatic Context Annotation, IRT-ACA)」，輔助人文學者提高圖像檢索與圖像情境脈絡分析的效能。據此，本研究的研究目的如下：

- 一、基於圖像物件與情境識別技術，發展支援數位人文研究之「基於自動情境標註之圖像檢索工具」，提升圖像檢索與情境脈絡解讀成效。
- 二、評估人文學者使用「基於自動情境標註之圖像檢索工具」輔以圖像檢索與

圖像情境脈絡解讀，在解讀圖像情境、科技接受度上是否具有顯著的效益提升。

- 三、輔以半結構式訪談與系統操作行為歷程分析，了解人文學者對於採用「基於自動情境標註之圖像檢索工具」支援數位人文研究的感受，以及有效之系統操作行為模式為何。

### 第三節 研究問題

基於上述的研究目的，本研究之研究問題如下：

- 一、本研究發展之「基於自動情境標註之圖像檢索工具」，其情境分類的 (precision rate)、召回率(recall rate)及 F1-measure 分別為何？是否足以支援數位人文研究？
- 二、人文學者使用本研究發展之「基於自動情境標註之圖像檢索工具」，以及「基於文本的圖像檢索工具」支援數位人文研究，在探索解讀圖像情境的成效上是否具有顯著的差異？
- 三、人文學者使用本研究發展之「基於自動情境標註之圖像檢索工具」，以及「基於文本的圖像檢索工具」支援數位人文研究，在科技接受度上是否具有顯著的差異？
- 四、人文學者採用「基於自動情境標註之圖像檢索工具」支援數位人文研究的感受與建議，以及有效之系統操作行為模式為何？

### 第四節 研究範圍與限制

由於研究人力與時間等限制，本研究之研究範圍與限制如下：

- 一、本研究以數位人文出發，但礙於具有研究價值、開放公開存取且具有解讀意義的圖像數位人文資料庫取得困難，故本研究所採用的圖像來源係以既

有之典藏系統—「點石齋畫報」為驗證系統成效之圖像集來源，未來希望能找到更多具有數位人文研究價值之數位圖像集進行驗證。因此，本研究之研究結果能否推論至其他類型圖像集，需要再做進一步研究，不宜作過度推論。

二、本研究因經費有限，難以邀請足夠具「點石齋畫報」圖像解讀經驗之資深人文學者進行實驗，故以懂得古文且對於清末民初畫報具有背景知識之大學生、研究生及研究畢業生為研究對象，其研究結果能否推論至專門從事數位人文圖像研究的人文學者，需要再做進一步研究，不宜作過度推論。

## 第五節 名詞解釋

### 一、數位人文(Digital Humanities)

數位人文源自於人文計算(humanities computing)，主要係由於資訊科技的快速發展，促使人們將電腦的計算與分析應用於人文學研究上，企圖透過資訊科技從事過去紙本閱讀難以觀察、無法想像或難以探索的議題(項潔、陳麗華，2014)。本研究聚焦於基於檢索並解讀圖像情境脈絡之數位人文研究，關注人文學者如何使用本研究開發之「基於自動情境標註之圖像檢索工具」輔助解讀及探索圖像情境，並希望能藉由此一工具的發展，支援人文學者更有效率地從事基於圖像解讀之數位人文研究。

### 二、監督式學習(Supervised Learning)

監督式學習為機器學習(machine learning)的一種，主流的監督式機器學習的分類演算法包含支持向量機(support vector machine)、隨機森林(random decision forests)及單純貝氏分類器(naïve bayes)，此類型分類演算法經具有已標註之資料集進行訓練後，可對於未知資料進行預測，本研究對於圖像的情境分類標註後，

進行訓練，取最佳結果作為自動情境標註工具之情境標註所用。

### 三、圖像檢索(Image Retrieval)

圖像檢索為資訊檢索(information retrieval)的子領域，一般的圖像檢索方式是以注入具有描述圖像特徵或描述內容的後設資料方法實作(Hyvönen, Saarela, Styrman & Viljanen, 2003)，並據此透過關鍵字進行檢索。近年來則開始將自動圖像標註、語意網路及知識圖譜等方式納入其中(Huang et al., 2018; Su, Wang, Yeh, & Tseng, 2009; Vijayarajan, Dinakaran, Tejaswin, & Lohani, 2016)，進而發展更貼近使用者需求的圖像檢索系統。

### 四、點石齋畫報

《點石齋畫報》於 1884 年發行之初乃是隨著《申報》派送之畫報，《申報》原先全稱《申江新報》，清同治十一年創刊於上海，創辦人為英商安納斯脫·美查(Ernest Major)，係近代中國首次利用石版印刷技術大量發行之刊物。《點石齋畫報》的畫報風格特別強調內容的新奇、轟動性、娛樂性、即時性(Wang, Gu, 2007)，並且畫風精細。清光緒二十四年出至第 528 號後停刊。參與畫報繪製的畫家包含吳友如、金蟾香、王釗、周慕橋及張志瀛等 17 人，畫風採用西方透視畫法，構圖嚴謹，線條流暢簡潔，使得畫報圖像生動地展現晚清各階層人群的中西方思想衝擊和社會變化(Xiaoqing, 2003)。發行的 15 年間，共計有圖 4,658 幅、文 120 萬字(Henningsmeier, 1998)。

## 第二章 文獻探討

本章針對與本研究相關之文獻進行探討，共分成二個小節，第一節為圖像檢索系統發展；第二節為自動圖像註解與檢索，依序說明如下：

### 第一節 圖像檢索系統發展

圖像檢索技術最早可追溯至 1970 年代，過去的研究發展了三種型態的圖像檢索方法，包括基於文本之圖像檢索(Text-Based Image Retrieval, TBIR)、基於內容的圖像檢索(Content-Based Image Retrieval, CBIR)，以及兩者混合型圖像檢索，如圖 2-1 所示。1970 年代之初，主要以發展基於文本之圖像檢索系統為主，此一方法採用文字描述的方式描述圖像的特徵，並據此注入為描述圖像的後設資料，作為檢索詞，例如圖像的關鍵詞、作者、年代、收錄及尺寸等。1990 年代以後，發展出了基於內容的圖像檢索，採用對圖像的高低階特徵提取方式抽取圖像特徵，常見的特徵有顏色分度、紋理、形狀及空間關係等，並依靠其建立索引，然後透過計算比較這些特徵與查詢條件之間的相似程度來達成圖像檢索。

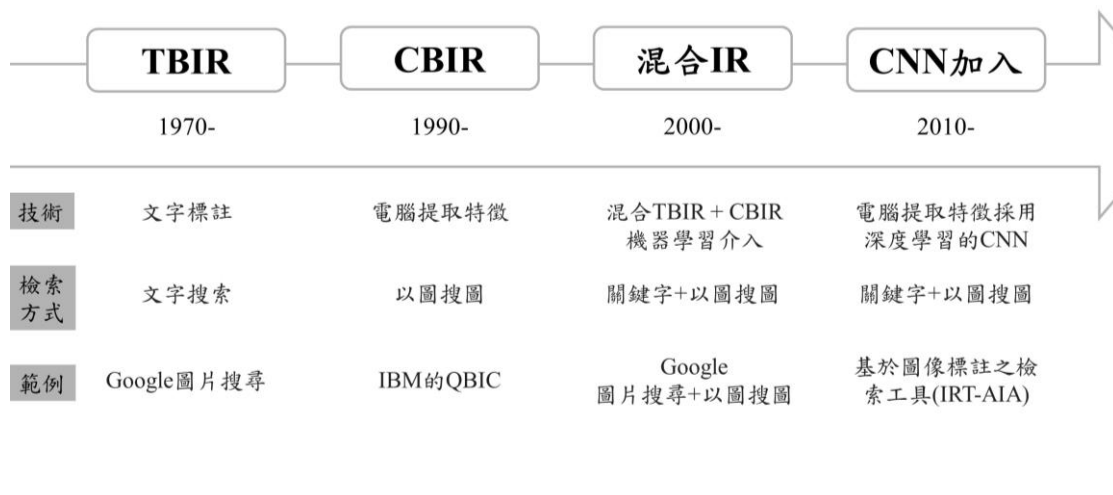


圖 2-1 圖像檢索系統發展史

(資料來源：本研究整理)



TBIR 採用描述圖像的關鍵字為圖像檢索的後設資料，雖然這樣的方式最貼近人們的觀點，但也容易受到人工注入關鍵字之主觀性或認知差異，而產生人工注入關鍵字與使用者觀點差異之語意鴻溝(semantic gap)問題。所以近年來關於 TBIR 的研究多著墨於改善語意鴻溝問題，例如有研究加入語意網路(semantic network)(Yin & Chen, 2009; Huang et al., 2018)，以及 Google 圖片搜尋加入知識圖譜(knowledge graph)來改善此一問題。

CBIR 透過對於圖像的特徵提取方式，亦存在一個底層特徵和上層理解之間的差異——語意鴻溝問題，而要達到提取圖像高階語意的方法包括物件本體(object ontology)、監督式及非監督式機器學習、關聯反饋(relevance feedback)，以及語意模板(semantic template)等(Liu, Zhang, Lu, & Ma, 2007)，其中「物件本體」係描述物件本身的特徵，例如穿著紅色條紋襯衫的男子，紅色跟條紋為顏色與紋理，依照定義此為低階特徵，但用於描述物件本體即變成為此一物件的高階語意特徵；「監督式及非監督式機器學習」則採用監督式學習的支援向量機(Support Vector Machine, SVM) (Chen, Zhou & Huang, 2001)、決策樹(decision tree) (Liu, Zhang, & Lu, 2008)，以及非監督式學習的 K-means (Murthy, Vamsidhar, Kumar, & Rao, 2010) 等法進行圖像分類及分群，以提高圖像檢索效能；「關聯反饋」則是透過人機互動過程學習來改進檢索效能，人機互動流程大致可區分為「輸入關鍵字或圖像」、「呈現檢索結果」、「使用者標註結果是否相關」，以及「回傳資料庫進行學習」等四個步驟(Rui & Huang, 2000)，也有研究透過決策樹將圖像分成相關與不相關(Shyu, 2000)，用以訓練決策樹與人機互動的結果；「語意模板」係指採用自然語言處理的 WordNet 進行語意(Chen & Chen, 2017)分析，對於使用者所輸入之關鍵字進行概念分類後，再傳回相對應之圖像檢索結果。

## 第二節 自動圖像標註

2012 年，Hinton 與兩名學生 Krizhevsky 與 Sutskever (2012)提出 AlexNet 圖

像分類模型，此一模型係基於深度學習之卷積神經網路(Convolutional Neural Network, CNN)發展而成，並使用硬體 GPU (Graphics Processing Unit)進行訓練，刷新了圖像分類挑戰賽 ImageNet 的紀錄，人們開始意識到電腦性能已足以模擬人類的神經網路，達到模擬人類視覺的效果，也開啟了近幾年機器視覺領域的快速發展。基於卷積神經網路之深度學習對於圖像標註的準確率，因為卷積層具有自動特徵選取功能，而被大幅提高，也促使圖像檢索領域有了長足的進步。而本研究對自動圖像標註如何抽取圖像的特徵，達成有效的檢索，整理後將其區分為三層，說明如下：

1. 情境層：如海邊、戰爭、嚴肅及街頭表演等，可以大致分為場景與主題。
2. 物件層：如人、動物與物品，包含的圖像中可以觀察到的所有物件，可以透過目前的電腦視覺技術進行識別。
3. 特徵層：如顏色、紋理及關鍵點等，是圖像中最基礎的特徵，也是電腦視覺判別物件的基礎。

諸如 Mask R-CNN (He, Gkioxari, Dollár & Girshick, 2017)、YOLO (Redmon, Divvala, Girshick & Farhadi, 2016)等自動圖像標註技術的發展，都是對於圖像的物件層進行自動標註，基於這些技術也誕生許多圖像檢索工具，例如 Lorang 等人(2015)透過識別歷史報紙圖像，建立效能相當優異的數位圖像檢索工具，以增進對於歷史學及文學的研究支援，也對於數位圖像的取用及檢索有了更多的選擇。而對於情境層進行自動標註的研究發展則較少，Vasconcelos 等人(2007)提出了監督式的多概念標註模型(Supervised Multiclass Labeling, SML)，將每個圖像情境定義成圖像分類，並讓模型進行分類。Ivasic-Kos, Ipsic 與 Ribaric (2015)提出基於模糊知識的多層圖像標註系統(Multi-layered Image Annotation System, MIAS)，將圖像辨識推往更高階語意的方向，也使得圖像檢索有更好的效能。

綜合以上所述，本研究認為圖像檢索方式對於促進數位人文研究的關鍵，除了在於擁有圖像物件層的元素後，應該再嘗試如何使之獲取圖像更高階的情境語



義。然而至今尚缺乏研究嘗試從數位人文學者角度，發展有助於提升圖像檢索與圖像脈絡解讀工具之相關研究，本研究正好可以彌補這個缺口。



## 第三章 系統設計

本章分為四個小節：第一節為系統設計理念；第二節為系統架構；第三節為系統使用者介面；第四節為系統開發環境與工具，依序說明如下：

### 第一節 系統設計理念

現今主流圖像檢索方法之一為基於文本的圖像檢索(Text-Based Image Retrieval, TBIR)技術，其概念係透過描述圖像的後設資料來查找圖像，而其後設資料通常係透過人工以關鍵詞方式注入，其優點是易於實現，缺點則是這些基於關鍵字描述圖像的後設資料會因人而異，並且充滿著個人的主觀性，這對於數位人文領域的圖像檢索而言，形成了難以準確決定檢索關鍵詞的問題。因此，如何更精確且客觀的決定描述圖像的後設資料，甚至給予不同檢索層次的圖像檢索後設資料，是影響圖像檢索準確率(precision rate)及召回率(recall rate)的關鍵。

為了提升人文學者在檢索圖像輔以數位人文研究上的效能，本研究發展「基於自動情境標註之圖像檢索工具(Image Retrieval Tool based on Automatic Context Annotation，以下簡稱 IRT-ACA)」，其核心設計理念為藉由自動圖像物件及情境標註提升影像檢索成效，進而輔助人文學者得以使用 IRT-ACA 的物件檢索及情境檢索功能，增進其對於圖像之檢索效能，以克服傳統基於人工注入之圖像後設資料跟使用者之間對於檢索詞認知差異，所產生的圖像檢索問題。

近年來在數位人文領域，已開始有研究採用 Mask R-CNN 深度學習技術進行圖像物件辨識(Huang, Zhong, Sun & Huo, 2019)，並據此結合基於機器自動辨識之物件後設資料(object metadata)及人工注入之主觀後設資料(subjective metadata)，提升影像檢索效能，進而提升人文學者對於圖像解讀之成效。當機器對圖像的內容擁有了完整的物件標註後，我們透過文字轉成向量且採用 TF-IDF 進行詞向量加權，接著使用多種機器學習演算法進行情境分類的訓練得到模型，為所有實驗圖片添加其適當的情境標籤使其具有自動物件及情境標註功能，並據此發展兼具

圖像物件及情境之圖像檢索功能，以增進人文學者解讀圖像之效能。

## 第二節 系統架構

本研究發展之「自動情境標註之圖像檢索工具」系統架構如圖 3-1 所示，整體系統架構可以區分為「使用者介面」、「搜尋處理」，以及「自動情境標註 API」三個部分。其中「使用者介面」包括圖像檢索工具、圖像詳細資訊檢視，以及點石筆記；「搜尋處理」提供使用者多種檢索方法，爾後將使用者所輸入情境、物件、全文及標題關鍵字，傳送至「點石齋標註資料庫」進行檢索後，並以情境相關性由高至低方式進行排序，回傳搜尋結果；「自動情境標註 API」會基於專家提供之圖像物件及情境標註資料，經由 SVM 之模型進行訓練，產出一個可提供情境識別之模型，並使用此一模型為每張資料庫中的圖像進行情境識別，最後將識別結果傳回「點石齋標註資料庫」進行儲存。

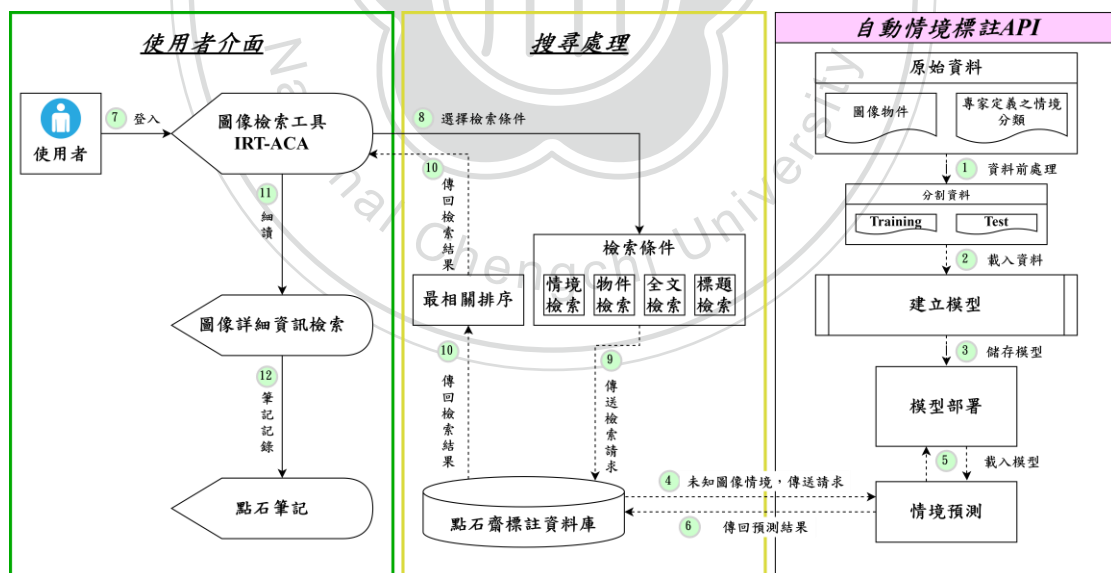


圖 3-1 具自動情境標註之圖像檢索工具系統架構圖

以下依據系統架構圖中的流程說明整個系統的運作機制：

1. 資料前處理：包括資料清理(data cleaning)、資料整合(data integration)及資料轉換(data transformation)，其中資料清理用以填補遺失值、處理 Outliers

及 Noise，修正資料的不一致等問題；資料整合係將不同資料集的資料進行合併；資料轉換係為了讓資料的數值符合資料分析所需之數值範圍或去單位化，常見的方法為標準化。

2. 載入資料：把資料前處理後的資料進行訓練集及測試集分割後，作為圖像物件及情境識別模型建立圖像識別模型之用。
3. 儲存模型：將已經訓練好之物件及情境識別模型，儲存並佈署到 IRT-ACA 的自動情境標註 API 中，當系統提出請求時，可以利用此一模型，為其識別符合之圖像之情境。
4. 未知圖像情境，傳送請求：當資料庫中的圖像尚未情境分類時，向自動情境標註 API 傳送請求。
5. 載入模型：「自動情境標註 API」接受情境預測請求後，即針對系統傳送過來的檢索請求，使用已經訓練好之情境識別模型，進行符合檢索條件之物件及情境識別運算。
6. 傳回預測結果：將預測結果傳回資料庫中，並合併回原始 Metadata 中。
7. 登入：使用者登入此一圖像檢索工具，準備進行圖像檢索。
8. 選擇檢索條件：使用者透過圖像檢索工具之檢索功能，可分為情境檢索、物件檢索、全文檢索與標題檢索。情境檢索使用機器預測之情境標籤進行檢索；物件檢索為機器標註之物件標籤及人工標註之物件標籤進行檢索；全文檢索採輸入關鍵字的方式對點石齋畫報之內文進行檢索；標題檢索也採輸入關鍵字的方式對點石齋畫報之標題進行檢索。
9. 傳送請求：使用者提供欲尋找圖像的情境、物件、全文及標題的關鍵字組合後，系統傳送此一請求至點石齋標註資料庫中，進行條件比對。
10. 傳回檢索結果：在系統條件比對後，將符合其檢索條件之結果，進行情境相關度排序後，回傳到使用者的檢索頁面，供使用者查看。
11. 細讀：使用者在圖像檢索工具中得到了檢索結果圖像後，可點擊展開其顯

示圖像詳細內容，系統會帶出「圖像詳細資訊檢視」介面，以便使用者進行圖像資訊細讀，頁面將提供更詳細的圖像資料，包含其出版年份、作者、繪師及文章內文等。

12. 筆記記錄：在檢索圖像後，若使用者感興趣或認為該圖像有用，便可將此一圖像加入「點石筆記」中，以利於日後再次進行檢視。

### 第三節 系統使用者介面

本節介紹本研究發展之「自動情境標註之圖像檢索工具」的各個使用者介面，包括圖像檢索介面，以及點石筆記，依序說明如下：

#### 一、圖像檢索工具

圖像檢索工具為 IRT-ACA 的核心功能頁面，如圖 3-2 所示。其中，左側為檢索工具，使用者可以透過多個圖像情境、物件、全文及標題的組合詞彙，向點石齋標註資料傳送檢索請求，以取得最符合現在關鍵字組合之情境圖像識別結果至右側呈現。此外，所有頁面都有點石筆記的按鈕，使用者可依不同的研究主題，進行所需圖像蒐集。再則，使用者介面左側提供篩選功能，可針對所需圖像做進一步的篩選。





圖 3-2 具自動情境標註之圖像檢索工具使用者介面

## 二、筆記

筆記為 IRT-ACA 提供作為紀錄使用者個人註解與作為回溯之用的筆記工具，讓使用者得以在利用「具自動情境標註之圖像檢索工具」找到目標圖像時，將需求的圖像紀錄進點石筆記中，以便日後再次進行檢視，或者紀錄檢索過程所獲得之線索，使用者介面如圖 3-3 所示。其中，點石紀錄有新增與刪除功能，以便使用者進行圖像檢索過程進行所需之檢索過程記錄。

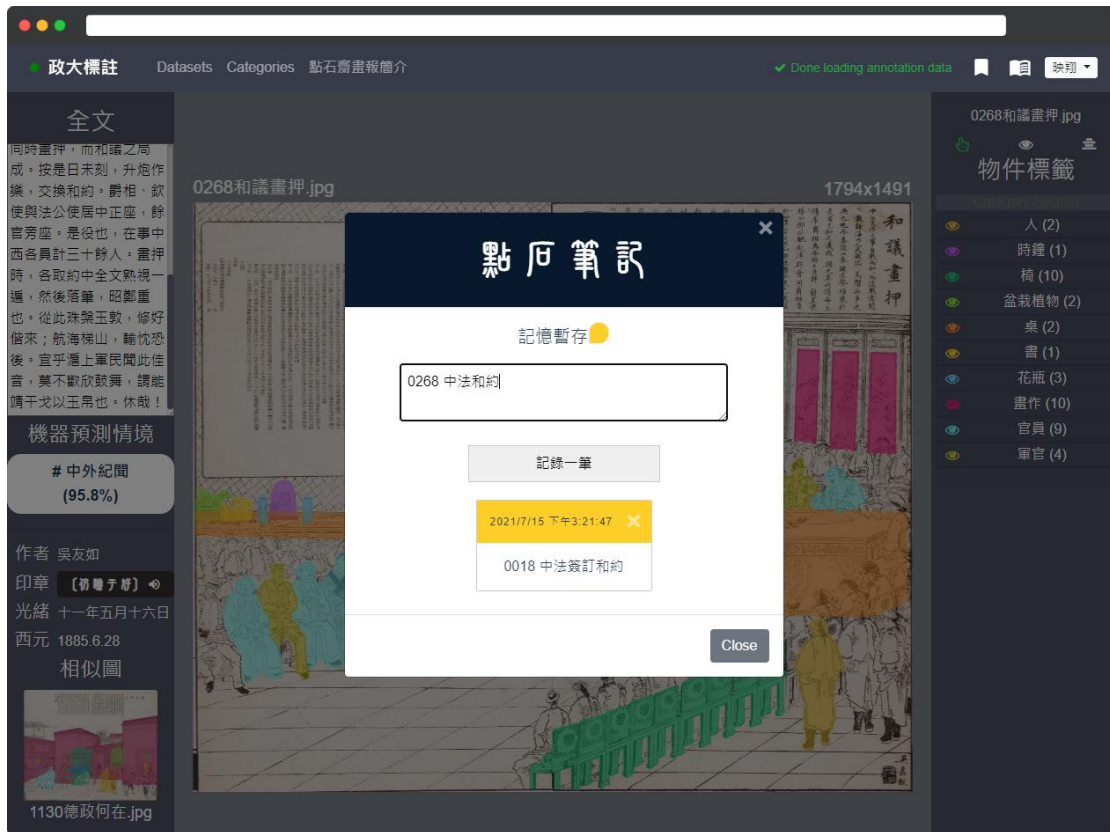


圖 3-3 點石筆記之使用者介面

### 三、圖像詳細資訊

圖像詳細資訊之使用者介面如圖 3-4 所示，此一介面提供使用者進行透過系統自動辨識後的圖像物件及情境後設資料解讀，其中左側呈現該圖像的情境標註、全文內容、出版日期，以及作者等後設資料給使用者閱讀，中間為圖像展示區域，右側則為此圖中所有被系統自動識別出來的物件標註及其數量。



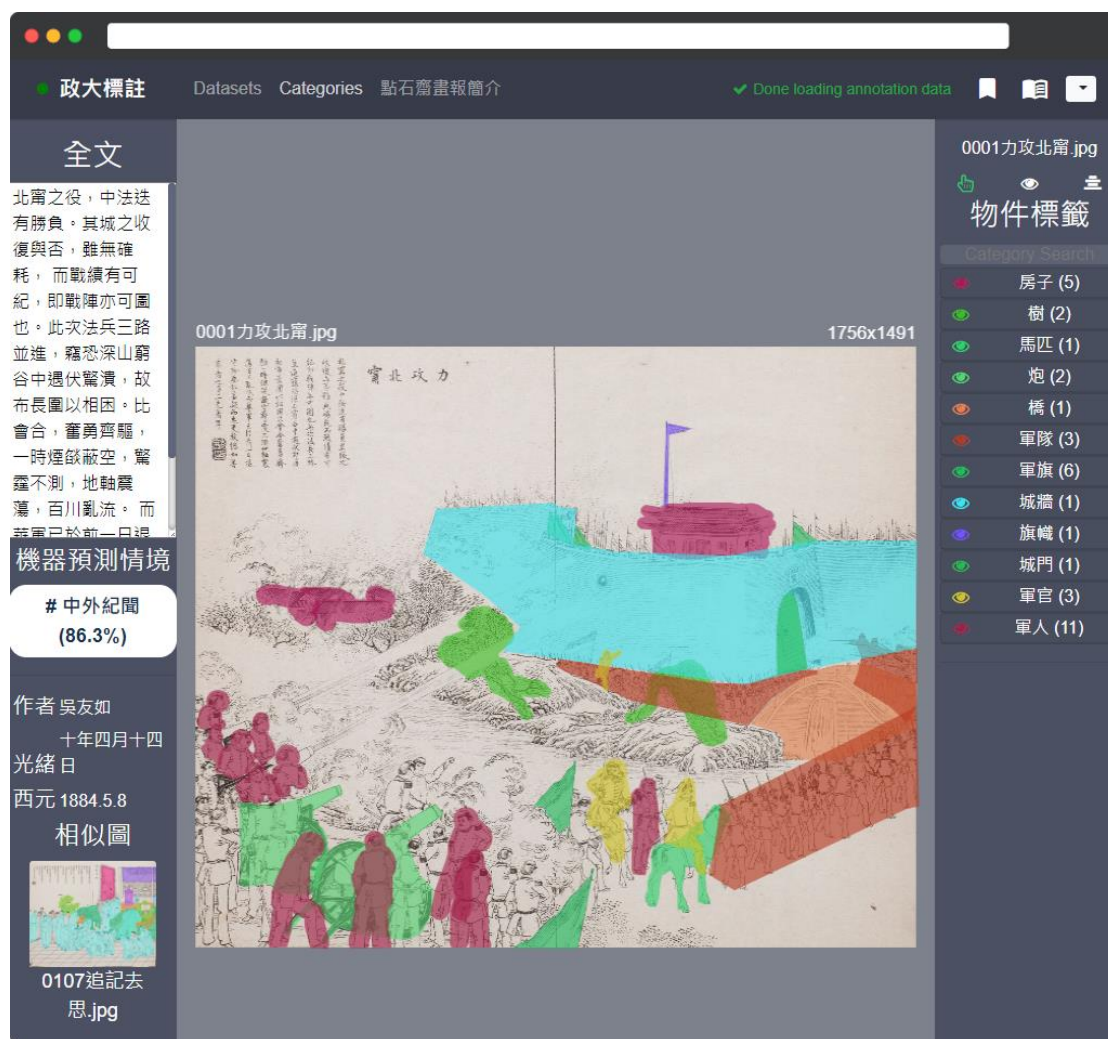


圖 3-4 檢視圖像詳細資料之使用者介面

#### 第四節 系統開發環境與工具

本節說明本研究發展「自動情境標註之圖像檢索工具」的開發工具，包括使用的程式語言、技術，以及資料庫，說明如下：

##### 一、Scikit-learn

Scikit-learn 又稱為 Sklearn，是一個基於 Python 程式語言發展的開放原始碼機器學習工具包，此一工具包含透過 NumPy、SciPy 和 Matplotlib 等 Python library 實現之大量主流機器學習演算法，並包含可實現多種電腦視覺演算法之 Scikit-image。本研究利用其所提供的機器學習與電腦視覺演算法，進行「自動情境標

註之圖像檢索工具」之功能開發。

## 二、Pytorch

PyTorch 前身為 torch，torch 是 2002 年用 Lua 程式語言撰寫發布的開源機器學習函式庫。Facebook AI 研究團隊將其改寫支援 Python，於 2017 年開源，支援 GPU 硬體加速及動態神經網路，語法簡潔直觀，並具有很好的靈活性與易用性，因而本研究採取此一框架，以實現圖像分類模型。

## 三、Docker

Docker 是個管理容器(container)的服務，提供「應用層級」的虛擬化技術，此一技術不同於 OS 層級虛擬化的虛擬機，係藉由 Linux 內核的 cgroup、namespace，以及 UnionFS 等技術，來達到應用程式的資源限制與封裝隔離。本研究資料處理及模型訓練時需使用容器佈署實驗環境，並應用在發展的 IRT-ACA 系統上。

## 四、Vue.js

Vue.js 又簡稱為 Vue，是一個用於開發互動式 web 界面的 JavaScript 前端框架，和其他整體性的框架不同，Vue 可從底層開始逐漸的擴展其框架。特別的是，Vue 的底層核心專注在 View 上，可以輕易的和其他 library 或現有專案整合。本研究採用此一框架，發展 IRT-ACA 的使用者介面。

## 五、COCO Annotator

本研究採用 Justin Brooks (2019)以 Vue.js、Python 撰寫，並於 Github 釋出之一款 Web-based 的開源軟體圖像標註工具 COCO Annotator，作為蒐集基於人工

標註圖像物件與情境之工具。此一工具支援多種標註方式，包括了定界框(bounding box)、圖像分割(object segmentation or instance segmentation)，並且支援線上多人合作標註，有利於進行標註策略之標準化方法與原則制定，其標註結果可以標準化之標註格式—MS COCO 匯出，以作為圖像物件及情境識別之機器學習方法訓練使用。

## 第五節 系統操作說明

本節說明本研究開發之 IRT-ACA 及 TBIR 的系統操作說明，包括使用的動機、過程及使用到的功能，說明如下：

### 一、IRT-ACA 之操作說明舉例

若使用者想了解晚清與外國有哪些的互動關係，因為歷史常識而得知晚清的戰爭頻繁，想從《點石齋畫報中》找出與戰爭相關的圖像。IRT-ACA 的「圖像檢索工具」中檢索功能有「情境檢索」、「物件檢索」、「全文檢索」及「標題檢索」，使用者想先從「情境檢索」嘗試找到所有戰爭圖像，使用者查看了情境檢索中的四個情境標籤之說明，認為「中外紀聞」是符合其概念的，進而用此檢索功能進行篩選，IRT-ACA 回傳了中外紀聞的相關圖像，使用者發現符合中外紀聞所涵蓋之圖像仍然非常多，不僅僅有戰爭，也包括了戰後所簽訂條款及國家衝突等圖像，而使用者目前只想聚焦於戰爭上，於是乎想再透過物件檢索進一步做篩選，物件檢索中提供大量的物件標籤，使用者聯想到戰爭在圖像中可能出現的物件，而點擊了物件標籤「槍」進行檢索篩選，系統此時回傳了符合中外紀聞及槍之檢索結果給使用者，目前檢索結果已滿足使用者預期，接下來可以依序點開圖像，進入「圖像詳細資訊」頁面進行察看，包括了該圖像的 Metadata、內文及圖像的細節等等，進而對該圖有更深刻的了解，再透過「筆記」功能，將其用文字敘述記錄下來。

## 二、TBIR 之操作說明舉例

若使用者想了解晚清與外國有哪些的互動關係，因為歷史常識而得知晚清的戰爭頻繁，想從《點石齋畫報中》找出與戰爭相關的圖像。TBIR 的「圖像檢索工具」中檢索功能有「全文檢索」及「標題檢索」，「全文檢索」提供了可以對其圖像中的內文進行檢索，使用者想找到與戰爭有關，嘗試輸入戰爭、中法、戰等檢索詞彙進行查找，若檢索結果已滿足使用者預期，接下來可以依序點開圖像，進入「圖像詳細資訊」頁面進行察看，包括了該圖像的 Metadata、內文及圖像的細節等等，進而對該圖有更深刻的了解，再透過「筆記」功能，將其用文字敘述記錄下來。



## 第四章 研究方法與實驗設計

本章分為七個小節，第一節為研究架構，說明本研究的自變項以及依變項彼此之間的關係；第二節為研究方法，說明本研究所採用的研究方法；第三節為研究對象，說明本研究的研究對象及其對於點石齋畫報的知識背景；第四節為研究工具，說明本研究所採用的研究工具及其用途；第五節為實驗流程，說明本研究的實驗設計與流程；第六節為資料處理與分析，說明本研究如何針對所蒐集資料進行分析；第七節為研究步驟，說明本研究的研究流程，依序說明如下：

### 第一節 研究架構

本研究旨在探討使用者利用不同圖像檢索工具，輔以進行《點石齋畫報》之圖像檢索與解讀的任務過程中，採用「自動情境標註之圖像檢索工具(IRT-ACA)」輔助的實驗組，以及採用「基於文本的圖像檢索(TBIR)」輔助的控制組，在解讀圖像物件與情境關係的成效，以及科技接受度上是否具有顯著的差異，本研究之研究架構如圖 4-1 所示。

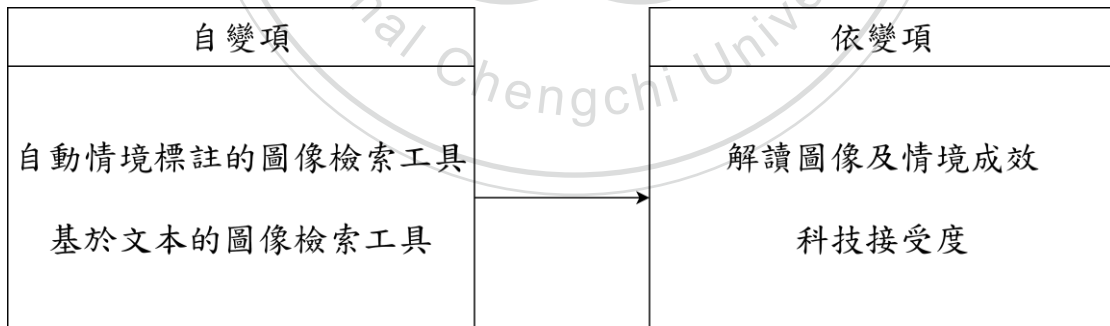


圖 4-1 本研究之研究架構圖

#### 一、研究變項

以下針對圖 4-1 研究架構圖中的各個變項說明如下：



### 1. 自變項

本研究之自變項為使用者使用不同圖像檢索工具輔以進行點石齋畫報圖像檢索任務，其中實驗組採用「自動情境標註之圖像檢索工具」輔以解讀點石齋畫報圖像，而控制組則採用「基於文本的圖像檢索工具」輔以解讀點石齋畫報圖像。

### 2. 依變項

本研究之依變項為解讀點石齋畫報之圖像及情境成效與科技接受度，依序說明如下：

#### (1) 解讀點石齋畫報圖像及情境成效

解讀點石齋畫報之圖像及情境成效係以兩組研究對象在實驗過程中，分別採用兩個不同圖像檢索工具，進行點石齋畫報圖像檢索之正確率、召回率，以及所撰寫解讀圖像情境摘要進行評分方式，作為採用兩個不同圖像檢索工具解讀圖像及情境成效，是否具有顯著的差異分析依據。

#### (2) 科技接受度

本研究於研究對象完成實驗後，邀請其填寫該系統之科技接受度問卷，探討研究對象對於兩個不同圖像檢索工具之「整體」科技接受度，以及科技接受度之「認知有用性」與「認知易用性」感受，是否具有顯著的差異。

### 3. 控制變項

本研究採用「自動情境標註之圖像檢索工具」及「基於文本的圖像檢索工具」輔以解讀圖像情境的兩組研究對象，其實驗過程之解讀圖像時間及主題均控制為相同。

## 第二節 研究方法

本研究使用之研究方法包括準實驗研究法之對抗平衡設計、問卷調查法、半

結構式訪談法，以及系統操作行為歷程分析，說明如下：

### 一、準實驗研究法之對抗平衡設計

為了驗證本研究發展之「自動情境標註之圖像檢索工具」是否有助於人文學者進行圖像情境解讀，本研究採用準實驗研究法之對抗平衡設計(counterbalanced design)，將研究對象分為先使用「基於文本的圖像檢索工具」再使用「自動情境標註之圖像檢索工具」之 TI 組，以及先使用「自動情境標註之圖像檢索工具」再使用「基於文本的圖像檢索工具」之 IT 組。以比較使用兩種不同圖像檢索工具進行圖像情境解讀之研究對象，在圖像檢索正確率、召回率、解讀圖像情境成效，以及科技接受度上是否具有顯著的差異。準實驗研究法之對抗平衡設計旨在利用輪換方式，將每種實驗處理都呈現給每一組研究對象，藉以平衡每組之間的實驗誤差，其實驗實施細節於本章節第五節中詳述。

### 二、問卷調查法

本研究使用科技接受度問卷，來調查研究對象對「自動情境標註之圖像檢索工具」與「基於文本的圖像檢索工具」輔以進行圖像情境解讀的科技接受度感受。分別為在研究對象完成所採用工具輔以進行圖像情境解讀的當下，邀請研究對象填寫對於該工具的科技接受度量表，以瞭解研究對象對於本研究發展之「自動情境標註之圖像檢索工具」與「基於文本的圖像檢索工具」輔以進行圖像情境解讀上的科技接受度感受，希望藉此了解兩組使用不同工具支援圖像情境解讀研究對象，在「整體」、「認知有用性」及「認知易用性」科技接受度上是否具有顯著的差異。

### 三、半結構式訪談法

本研究於研究對象完成整個實驗流程後，進行半結構式深度訪談，來補充量



化分析上的不足。訪談內容依照研究對象操作兩個不同工具進行「檢索任務」，以瞭解研究對象對於兩個工具支援不同使用情境圖像解讀時的想法、感受和建議。訪談資料以錄音方式收錄後轉換成逐字稿，並進行逐字稿之編碼分析，最後再與量化分析結果進行交互驗證，以多方角度分析兩組研究對象在考量之依變項上產生顯著差異的主要原因。

#### 四、系統操作行為歷程分析

本研究為進一步瞭解研究對象在使用「自動情境標註之圖像檢索工具」與「基於文本的圖像檢索工具」時的有效解讀圖像情境之工具操作行為，在兩個工具中嵌入 Google Analytics 行為歷程記錄器，以紀錄研究對象使用工具進行圖像檢索與情境解讀時的行為操作序列，再透過各行為的敘述統計分析，並輔以陳勇汀（2017）開發之滯後序列分析(Lag Sequential Analysis)工具，探討研究對象使用兩個工具輔以圖像情境解讀之有效行為模式。

### 第三節 研究對象

由於本研究的圖像情境解讀之標的為「點石齋畫報」，為驗證本研究發展之「自動情境標註之圖像檢索工具」是否在輔助研究對象進行「點石齋畫報」圖像檢索與情境解讀成效及科技接受度上，顯著優於「基於文本的圖像檢索工具」。本研究徵求對於「點石齋畫報」具有一定認識與熟悉度，並且有意願配合實驗之大專院校學歷以上為研究對象。

### 第四節 研究工具

本研究所使用的研究工具包含「自動情境標註之圖像檢索工具」、「基於文本的圖像檢索工具」、科技接受度問卷，以及半結構式訪談大綱，分別說明如下：

## 一、自動情境標註之圖像檢索工具(IRT-ACA)

本研究發展之「自動情境標註之圖像檢索工具」具有「情境檢索」、「物件檢索」、「全文檢索」及「標題檢索」之檢索工具，以彌補「基於文本的圖像檢索工具」基於人工決定關鍵字描述圖像的後設資料輔以檢索，常常會因人而異，並且具有個人主觀性的問題。更重要的是，這對於數位人文領域的圖像檢索而言，形成了難以準確決定檢索關鍵詞的問題。因此，本研究評估分別使用 IRT-ACA 及 TBIR 在輔助人文學者檢索及解讀圖像情境成效，以及科技接受度上是否具有顯著的差異。

## 二、基於文本的圖像檢索工具(TBIR)

「基於文本的圖像檢索工具」為透過文字自由輸入之方式進行圖像之 Metadata 檢索，本研究為探討 TBIR 與 IRT-ACA 的檢索差異是否帶來最佳的檢索體驗，除了開發 IRT-ACA，也對 TBIR 的功能進行了開發，提供使用者「全文檢索」及「標題檢索」兩種檢索方式，使之能進行實驗中之檢索任務。

## 三、科技接受度問卷

為了解研究對象對於「自動情境標註之圖像檢索工具」及「基於文本的圖像檢索工具」輔以圖像檢索與情境解讀之整體科技接受度，以及科技接受度之「認知有用性」、「認知易用性」感受，本研究在實驗結束後邀請研究對象填寫科技接受度問卷。本研究採用 Hwang, Yang 及 Wang (2013)所編製之科技接受度問卷，此一問卷採用李克特五點量表，包含兩個科技接受度面向，分別為「認知有用性」，共 5 題；及「認知易用性」，共 3 題，總計 8 題。在問卷信度上，「認知有用性」面向之 Cronbach's  $\alpha$  值為 0.95，「認知易用性」面向之 Cronbach's  $\alpha$  值為 0.94。Devellis (1991)指出若問卷之 Cronbach's  $\alpha$  值達 0.7 以上，則具有可接受之最小信

度，本研究採用之科技接受度問卷具有良好的信度。訪談大綱詳見附錄二、三。

#### 四、半結構式訪談大綱

為了解研究對象對於「自動情境標註之圖像檢索工具」及「基於文本的圖像檢索工具」之使用感受，本研究於實驗後對於使用兩個不同圖像檢索工具輔以圖像檢索與情境解讀的研究對象進行半結構式深度訪談，以深入瞭解實驗過程中研究對象對於包含「檢索任務」、「圖像情境解讀摘要撰寫」，以及系統的「整體」感受上的想法、感受與建議。訪談大綱詳見附錄一。

### 第五節 實驗流程

本研究之實驗流程如圖 4-2 所示，詳細說明如下：



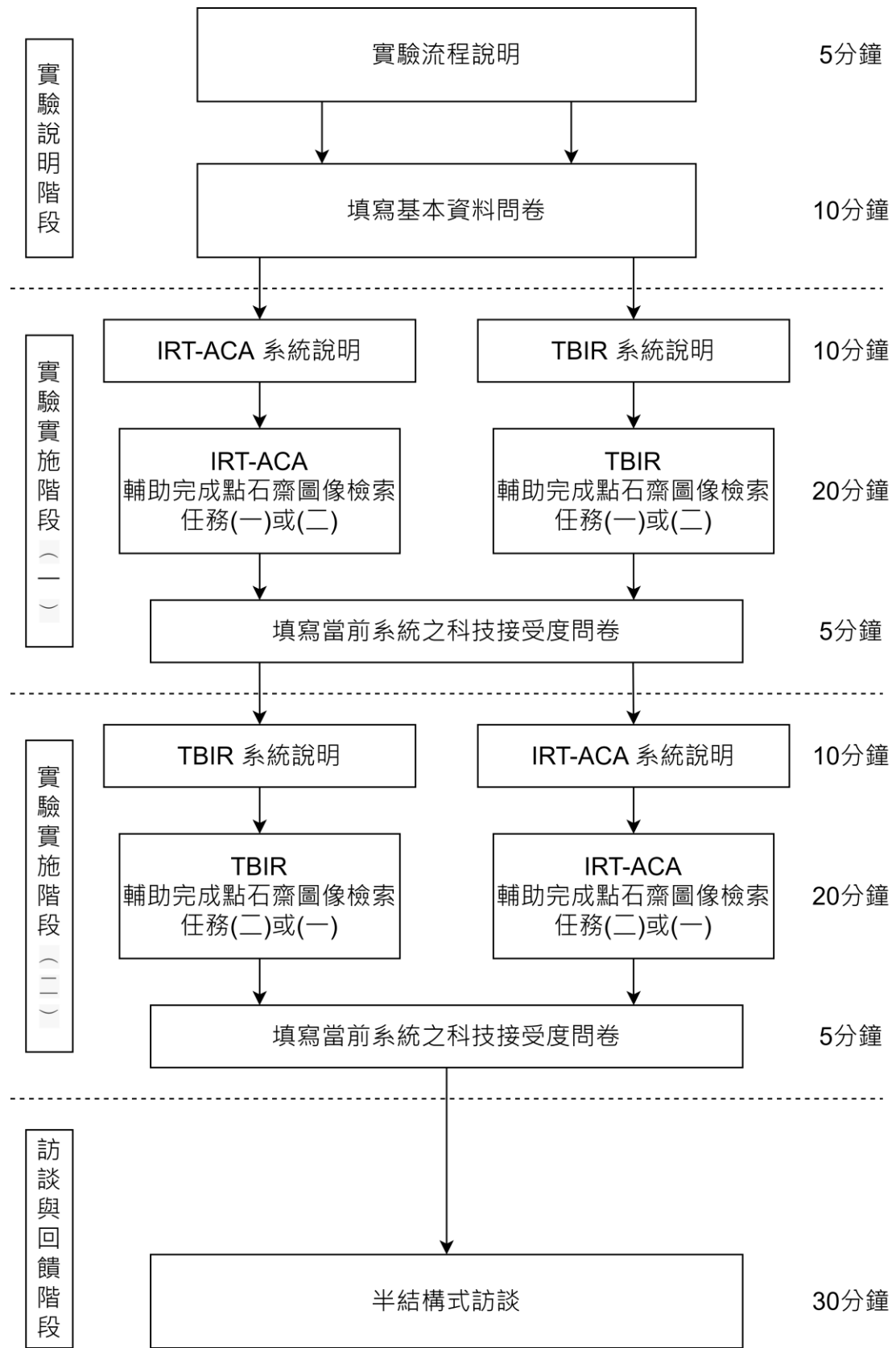


圖 4-2 本研究之實驗流程圖

## 一、實驗說明階段

首先，在正式進行實驗前先進行實驗說明，讓研究對象了解本研究所進行的圖像檢索與圖像情境解讀任務的實驗流程與目的，接著請研究對象填寫先備知識測驗單，此一階段規劃時間為 15 分鐘。

## 二、實驗實施階段

實驗實施階段將分為階段一與階段二，在兩個階段中，為減少任務的難易度影響到系統的使用成效，任務皆平均分配任務給同一系統的實驗對象。任務內容包含任務一為檢索任務「晚清與外國的互動關係」，任務二為檢索任務「晚清的平民百姓生活」，兩組研究對象分別透過「自動情境標註之圖像檢索工具」及「基於文本的圖像檢索工具」進行圖像檢索，以及檢索到的圖像敘述撰寫。其中實驗實施階段一與二，皆包括系統說明、檢索任務說明、操作系統完成任務單，以及填寫科技接受度問卷，共需 70 分鐘；

## 三、訪談與回饋階段

實驗結束後，邀請參與研究對象接受半結構式深度訪談，進而了解研究對象在使用不同圖像檢索工具進行檢索圖像與情境解讀過程之歷程與感受，特別是對於「自動情境標註之圖像檢索工具」輔以圖像檢索及情境解讀的感受與建議，以作為未來持續改善系統功能的方向。

## 第六節 資料處理與分析

本研究旨在探討分別採用「自動情境標註之圖像檢索工具」及「基於文本的圖像檢索工具」來輔助研究對象完成解讀「點石齋畫報」圖像情境任務，在圖像情境解讀成效及科技接受度上是否有顯著的差異，以下說明本研究針對所蒐集資

料的分析方法：

### 一、自動情境標註之準確率分析

為了評估自動情境標註模型的好壞，本研究採用「交叉驗證 Cross-Validation」的方法來做驗證，而 Cross-Validation 的方法有很多種，本研究採用其最經典的 K-fold Cross-Validation，設定 K=10，將資料集分割成十份，其中九份為訓練集、一份為驗證集，如此反覆直到每一份資料都當過驗證集，總共執行十回，得到十個 Validation Error，再將其 Validation Error 取平均數，做為評估自動情境標註模型好壞的指標。

此外，本研究以準確率(precision rate)、召回率(recall rate)以及 F1-measure 評估演算法的準確率，當 F1-measure 值越接近 1 時，表示該演算法有較佳的精準度，其計算方法說明如下：

基本參數

TP (True Positive)：事實為真，且被我們的方式判斷為真，其情境識別成功。

TN (True Negative)：事實為假，且被我們的方式判斷為假，其情境識別成功。

FP (False Positive)：事實為假，且被我們的方式判斷為真，其情境識別失敗。

FN (False Negative)：事實為真，且被我們的方式判斷為假，其情境識別失敗。

準確率(precision rate)

準確率為衡量系統檢索圖像結果的正確性測度，如公式(1)所示。準確率越高時，表示檢索的圖像結果越具有相關性；反之，則表示檢索的圖像結果較不具有相關性。

$$Precision = \frac{TP}{TP+FP} \quad (1)$$



召回率(recall rate)

召回率為衡量系統檢索到所有相關圖像的完整性，如公式(2)所示，召回率越高時，代表資料庫未檢索到的相關圖像越少；反之，則表示資料庫中未檢索到的相關圖像越多。

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

F1-measure

準確率關注於檢索結果有多少是真，召回率關注於檢索漏掉多少沒找到。而 F1-measure 則可表示此工具同樣的注重準確率及召回率兩個指標，可以反映出該演算法的精確度，如公式(3)所示。

$$F1\text{-measure} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3)$$

## 二、解讀圖像情境成效差異分析

本研究根據《圖像晚清-點石齋畫報》之四大主題進行圖像檢索及情境解讀任務設計，讓分別使用「自動情境標註之圖像檢索工具」及「基於文本的圖像檢索工具」的研究對象進行兩個圖像檢索及情境解讀任務，並基於任務完成後所作答之圖像及針對圖像之描述進行評分，作為使用者探索解讀圖像情境成效之分析依據。本研究採用曼惠特尼(Mann-Whitney)U 檢定，檢測兩組使用者在分別使用 IRT-ACA 及 TBIR 兩個系統時，解讀圖像情境成效是否具有顯著的差異。

## 三、科技接受度

本研究在實驗過程中，在兩組研究對象分別完成兩個不同圖像檢索工具輔以圖像檢索與情境解讀任務後，各分別填寫一份使用工具之科技接受度問卷，以了解使用不同圖像檢索工具之科技接受度感受。本研究採用曼惠特尼(Mann-Whitney)U 檢定，檢測兩組使用者在分別使用 IRT-ACA 及 TBIR 兩個系統輔以解

讀圖像情境下，在「整體科技接受度」、「認知有用性」及「認知易用性」科技接受度上是否具有顯著的差異。

#### 四、訪談質性分析

本研究於實驗結束後，採用內容分析法針對基於半結構式訪談所蒐集研究對象，對於使用兩個不同圖像檢索工具在包含任務過程中的「檢索任務」、「圖像情境解讀摘要撰寫」以及系統的「整體」感受上的想法與建議進行逐字稿分析，希望透過訪談內容之質性資料分析，補充量化分析上的不足。

#### 五、系統操作行為歷程分析

本研究為了了解兩組研究對象使用兩個不同圖像檢索工具輔以圖像檢索及情境解讀之有效工具操作行為模式，因此基於研究對象所撰寫之圖像情境解讀內容成績的平均分數，將研究對象區分為高、低不同成效研究對象，以滯後序列分析(Lag Sequential Analysis)探討高檢索成效研究對象在使用工具輔以圖像檢索與情境解讀之顯著操作行為序列轉移，以找出有效之工具操作行為模式。

## 第七節 研究步驟

本研究之研究步驟如圖 4-4 所示，說明如下：

- (一) 擬定研究方向與主題。
- (二) 確定研究問題與目的。
- (三) 蒐集相關文獻，進行文獻探討與整理歸納。
- (四) 選擇研究方法與研究規劃。
- (五) 進行演算法設計、訓練模型與建立模型。
- (六) 系統設計與實作。
- (七) 徵求研究對象進行實驗，評估系統支援圖像檢索與情境解讀的成效。
- (八) 整理實驗所得之數據與工具操作之行為歷程資料，進行 SPSS 統計及滯後序列分析。
- (九) 針對研究成果進行歸納，提出研究結論與相關建議。
- (十) 撰寫論文。

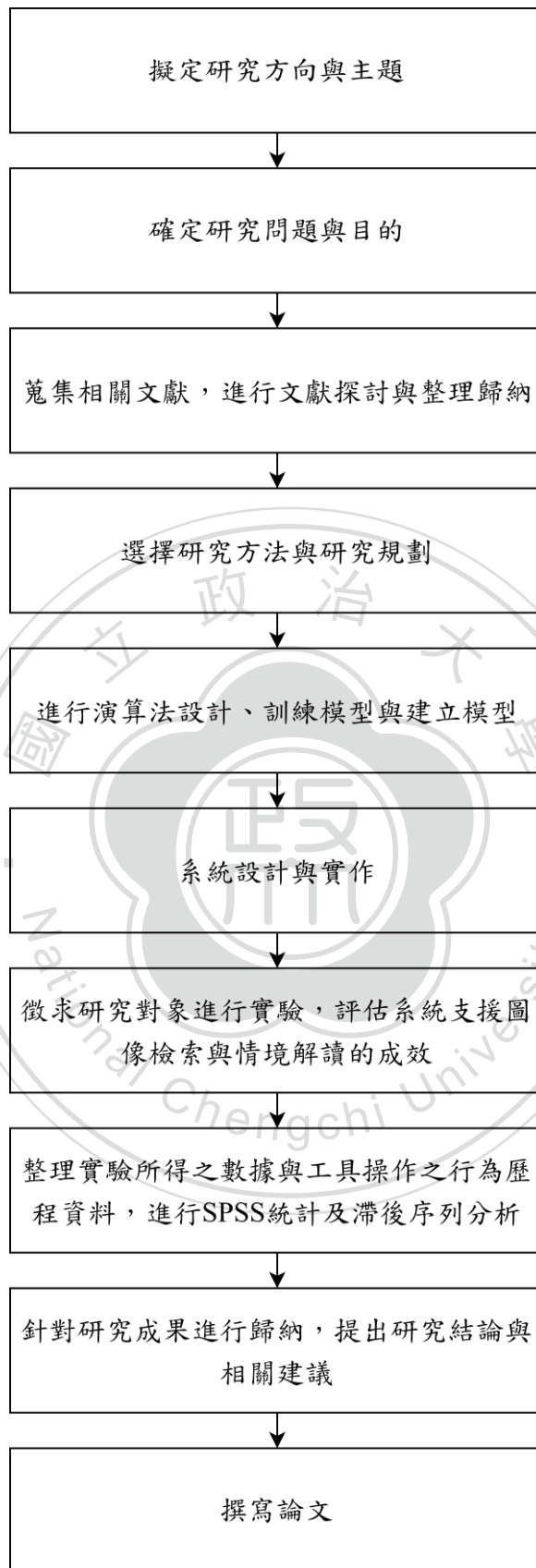


圖 4-3 研究實施步驟圖

## 第五章 實驗結果分析

本研究為探討實驗對象使用 IRT-ACA 及 TBIR 在解讀圖像情境成效，以及科技接受度上是否具有顯著的差異，因此採用對抗平衡設計分別讓實驗對象輪流使用 IRT-ACA 以及 TBIR 進行點石齋畫報之圖像檢索任務，並評估其圖像解讀任務成效。此外，本研究在實驗期間所蒐集兩組實驗對象填寫之科技接受度量表，作為科技接受度評估依據，以及完成圖像檢索任務過程所檢索獲得之圖像與系統操作行為紀錄予以量化分析，進而進行系統操作之行為滯後序列分析，再搭配半結構式訪談資料輔助，歸納本研究之結論。

本章分為七個小節：第一節為實驗對象基本資料，說明本研究之兩組實驗對象背景資料；第二節為發展之 IRT-ACA 自動情境標註系統準確率分析，以驗證本研究所發展之自動情境標註系統的準確率；第三節為實驗對象使用 TBIR 與 IRT-ACA 解讀圖像情境成效之差異分析，用以評估兩組實驗對象在不同的檢索工具下所獲得之圖像以及針對圖像進行之描述，以評估兩系統解讀圖像情境之成效差異；第四節為使用 TBIR 與 IRT-ACA 之實驗對象科技接受度差異分析，評估實驗對象使用兩個系統輔以點石齋畫報之圖像檢索任務之整體科技接受度，以及科技接受度之認知有用性、認知易用性之實際感受差異；第五節為實驗對象操作 TBIR 與 IRT-ACA 系統之行為紀錄分析，以探究實驗對象操作兩個系統來完成圖像檢索任務時的系統操作行為模式；第六節為半結構式訪談質性資料分析，歸納分析實驗對象操作兩系統完成圖像檢索任務時，對於系統提供之功能輔助各任務過程之感受與建議。第七節為綜合討論，針對本研究所獲得之實驗分析結果進行深入的討論。

### 第一節 實驗對象基本資料

本研究以透過網路徵求有意願參與本研究實驗之大學生以及碩博士生為主

要實驗對象，比較實驗對象使用 TBIR 或 IRT-ACA 輔以解讀點石齋畫報之圖像情境成效，以及科技接受度上是否具有顯著的差異，完成實驗活動之有效樣本共計 12 人，實驗對象基本資料如表 5-1 所示。從身分來看，大學生有 6 人、碩博士生的有 5 人、碩博士畢業生有 1 人；從系所來看，中文系背景有 3 人、非中文系之文科有 5 人、非文科有 4 人；從年齡來看，22 歲有 6 人、24 歲有 2 人、25 歲有 2 人、26 歲有 1 人、30 歲有 1 人。為了避免採用 TBIR 與 IRT-ACA 輔以點石齋畫報之圖像檢索任務的實施順序影響實驗結果，本研究採用對抗平衡設計，將實驗對象區分為兩組，依據實驗實施順序將其區分為先採用 IRT-ACA 後採用 TBIR 的 IT 組，以及先採用 TBIR 後採用 IRT-ACA 的 TI 組，其中先採用 IRT-ACA 後採用 TBIR 組的實驗對象有 7 人，先採用 TBIR 後採用 IRT-AC 的實驗對象亦有 7 人，其實驗實施順序與被分配之圖像檢索任務如表 5-2 所示。此外，為了避免 IT 組及 TI 組實驗對象完成不同圖像檢索之任務難度可能具有差異，本研究又再將 IT 組及 TI 組區分為具有 4 人及 3 人的兩個小組，並在使用系統解讀圖像情境過程中使用相同的系統，但是完成不同圖像檢索任務，以降低兩個任務難度不同可能影響圖像解讀任務成效之可能效益，其實驗實施順序與被分配之圖像檢索任務如表 5-3 及表 5-4 所示。

表 5-1 研究對象背景說明

身分	人數	系所	人數	年齡	人數
大學生	6	中文系	3	22 歲	6
碩博士生	5	文科	5	24 歲	2
碩博士畢業生	1	非文科	4	25 歲	2
				26 歲	1
				30 歲	1
總計	12	總計	12	總計	12



表 5-2 IT、TI 兩組實驗參與人數

組別	系統順序 1	系統順序 2	受測者人數
IT 組	IRT-ACA	TBIR	6
TI 組	TBIR	IRT-ACA	6

表 5-3 IT 組依據題目分組及實驗參與人數

組別	IRT-ACA	TBIR	受測者人數
組別 1	晚清與外國的互動關係任務單	晚清的平民百姓生活任務單	3
組別 2	晚清的平民百姓生活任務單	晚清與外國的互動關係任務單	3

表 5-4 TI 組依據題目分組及實驗參與人數

組別	TBIR	IRT-ACA	受測者人數
組別 1	晚清與外國的互動關係任務單	晚清的平民百姓生活任務單	3
組別 2	晚清的平民百姓生活任務單	晚清與外國的互動關係任務單	3

## 第二節 發展之 IRT-ACA 自動情境標註之準確率分析

為了能識別出得以反映「點石齋畫報」所收藏之圖像集特色之相應圖像情境標籤，本研究參考北京大學中文系陳平原教授所撰寫的《圖像晚清：點石齋畫報》一書中所分類的包括「中外紀聞」、「官場現形」、「格致彙編」、「海上繁華」四大

畫報描述情境，作為透過 TF-IDF 轉成向量後再用 SVM 進行文本分類之依據。該書中為其四大畫報情境共收錄了 160 餘幅畫報，然而因為時間限制，本研究僅採用其中 46 幅作為訓練集，其訓練集之情境分類與數量如表 5-5 所示。

表 5-5 訓練集之圖像情境分類與數量

編號	情境類別	樣本數
1	中外紀聞	11
2	官場現形	10
3	格致彙編	11
4	海上繁華	14
總計		46

為了訓練畫報圖像情境辨識模型，本研究針對圖像上的物件進行標註，再依據每張圖像所辨識出來的影像物件來辨識其圖像情境，此一過程可以分為三個步驟，步驟一先針對 Mask R-CNN 預訓練模型進行遷移訓練，該預訓練模型的原有訓練資料集為 Microsoft COCO 圖像資料集(Lin et al., 2014)，這些圖像集主要係從複雜的日常場景中擷取而來，總共有 80 類的物件類別，由於 Microsoft COCO 資料集中皆為現代照片，與《點石齋畫報》的風格差異過大。因此，本研究僅採用其中 22 個分類進行自動物件標註，在本研究實驗的 206 張圖中，自動物件標註總共計再額外標註了 1048 個物件作為訓練集；步驟二為人工標註校正，經與政大中文系熟悉點石齋畫報圖像的教授討論之後，本研究將物件標籤拓展至 112 個物件分類，其中「中外紀聞」以戰爭及談判和議為主題，因此圖像中包含了包括船、馬匹、炮、軍隊、官員等圖像物件；而「官場現形」則以人作為主體，因此人的身分區隔對於辨識此一情境類別畫報圖像就顯得額外的重要，包含了官員、女性、孩童、罪犯、棺材等圖像物件；此外「格致彙編」以介紹西方傳入晚清的

新事物為主軸，因此包含了炮、火車、電線桿、望遠鏡、西醫等圖像物件；最後，「海上繁華」以描述上海地區的平民百姓視角下，發生的各種奇聞軼事，因此包含了和尚、西妓、水煙斗、花瓶等圖像物件；本研究在聘請 3 位學生進行點石齋畫報圖像物件合作標註，總計透過人工標註了 4955 個圖像物件。圖像物件標籤經由上述兩個步驟之後，總計得到了 6003 個圖像物件標註，並區分成 112 分類，如表 5-6 所示；步驟三為情境分類，當機器對圖像的內容擁有了完整的物件標註後，我們透過文字轉成向量且採用 TF-IDF 進行詞向量加權，接著使用多種機器學習演算法進行情境分類的訓練得到模型，為所有實驗圖片添加其適當的情境標籤。為了評估模型的好壞，採用「K-fold Cross-Validation」的方法來進行驗證，並透過 F1-measure 比較不同機器學習分類演算法之好壞，結果如表 5-7 所示。最終本研究採用 SVM 進行情境分類，得到一個 Confusion Matrix，如表 5-8 表示。從表中可以看到「格致彙編」的 Precision rate 與 Recall rate 遠低於其他三個情境，導致拉低了整體的準度率。

表 5-6 訓練集之圖像物件分類與數量

編號	物件類別	樣本數	編號	物件類別	樣本數	編號	物件類別	樣本數
1	井	1	39	水龍	6	77	軍人	96
2	亭子	2	40	法國國旗	47	78	窗	71
3	人	1,930	41	炮	30	79	羊	1
4	人力車	23	42	煙囪	18	80	牛	10
5	傘	38	43	旗幟	50	81	十字架	4
6	報紙	2	44	日本國旗	8	82	船帆	12

7	壁燈	6	45	書	19	83	花瓶	76
8	城門	6	46	杯	79	84	英皇子	1
9	大水缸	7	47	水煙斗	8	85	蠟燭	32
10	大鐘	1	48	水稻	19	86	西妓	1
11	天文台	1	49	水雷	1	87	軍官	56
12	孩童	131	50	火車	1	88	軍旗	67
13	女性	290	51	煤	1	89	軍隊	35
14	守望台	5	52	牌位	1	90	轎子	11
15	屍體	5	53	時鐘	13	91	花窗玻璃	7
16	帳篷	23	54	當舖	1	92	門	45
17	庸醫	1	55	燈籠	83	93	雕像	4
18	官員	299	56	眼鏡	10	94	電報局	1
19	戎克船 /鯨	2	57	砲彈	11	95	電線桿	16
20	房子	476	58	盆栽植 物	97	96	鞭子	3
21	木馬	8	59	狗	36	97	鐵路	2
22	望遠鏡	3	60	猴子	2	98	領帶	4
23	桌	136	61	美國國 旗	1	99	馬匹	84
24	棺材	2	62	珊瑚	2	100	馬廐	4
25	城牆	29	63	船錨	2	101	香爐	5
26	樹	469	64	畫作	73	102	魚	2
27	槍	26	65	琵琶	3	103	鳥籠	5

28	氣球	9	66	畫像	4	104	鼓	2
29	消防員	14	67	碗	58	105	馬車	23
30	潛水艇	1	68	英國國旗	3	106	雞	13
31	濕版攝影相機	1	69	蒸汽機	1	107	船	127
32	椅	106	70	藤牌	5	108	十字鎬	1
33	橋	27	71	罪犯	21	109	匾額	41
34	水桶	29	72	西醫	3	110	公主	1
35	掛燈	30	73	警官	11	111	刀具	93
36	水瓶	41	74	象	6	112	儀表	1
37	扇	105	75	賽馬場	1			
38	教堂	2	76	路燈	24			
							總計	6,003

表 5-7 不同機器學習演算法在 K-fold Cross-Validation=10 之下分類結果比較表

方法	Accuracy	Precision	Recall	F1-measure
SVM	50.000%	0.481	0.498	0.487
NaiveBayes	40.625%	0.421	0.406	0.404
RandomForest	21.875%	0.146	0.219	0.173

表 5-8 使用 SVM 進行情境分類之 Confusion Matrix

實際 預測	中外紀聞	格致彙編	海上繁華	官場現形	總計	Precision
中外紀聞	9	2	1	2	14	0.643

格致彙編	1	3	3	3	10	0.300
海上繁華	0	5	7	1	13	0.538
官場現形	1	1	3	4	9	0.444
總計	11	11	14	10	16	
Recall	0.818	0.273	0.500	0.400		

### 第三節 使用者使用 TBIR 與 IRT-ACA 解讀圖像情境成效之差異分析

#### 一、使用 TBIR 與 IRT-ACA 系統解讀圖像情境成效之敘述統計分析

本研究將分別使用 IRT-ACA 及 TBIR 輔以圖像檢索之兩組使用者於兩階段圖像檢索任務學習單中，完成作答之圖像及針對圖像之描述，採取如表 5-9 所示之評分標準進行評分，作為使用者探索解讀圖像情境成效之分析依據。兩組使用者分別使用 IRT-ACA 與 TBIR，在兩種檢索任務下解讀之圖像情境成效敘述統計分析如表 5-10 所示。

表 5-9 圖像檢索任務學習單之評分標準

評分標準	0	1	2	3	4
以受測者找到的每一張圖像與檢索任務關聯進行評分	完全離題	正確描述此圖 但敘述為複製內文，並無自己的見解	正確描述此圖 但敘述過於抽象不完整	正確描述此圖 有著自己的見解且部分正確	正確描述此圖 有著自己的見解且完全正確

表 5-10 分別使用 IRT-ACA 及 TBIR 系統完成兩種圖像檢索任務之解讀圖像情



境成效敘述統計分析

項目	組別			
	實驗組 (n=12)		控制組 (n=12)	
	平均數	標準差	平均數	標準差
解讀圖像情境得分	31.83	10.29	22.83	11.51

二、使用 IRT-ACA 及 TBIR 系統解讀圖像情境成效之差異分析

本研究採用曼惠特尼(Mann-Whitney) U 檢定，檢測兩組使用者在分別使用 IRT-ACA 及 TBIR 兩個系統時，解讀圖像情境成效是否具有顯著的差異，結果如表 5-11 所示。結果顯示在兩個檢索任務下，IRT-ACA 與 TBIR 在解讀圖像情境成效上( $p=.046 > .05$ )達顯著的差異，並且採用 IRT-ACA 的實驗組顯著優於採用 TBIR 的控制組。

表 5-11 使用兩系統解讀圖像情境成效之曼惠特尼 U 檢定

項目	組別				U	p
	實驗組 (n=12)		實驗組 (n=12)			
	平均數	標準差	平均數	標準差		
解讀圖像情境 得分	31.83	10.29	22.83	11.51	37.50*	.046

\*表  $p<.05$

## 第四節 使用者使用 IRT-ACA 與 TBIR 系統之科技接受度差異分析

### 一、使用 TBIR 與 IRT-ACA 系統使用者之科技接受度差異分析

#### (一) 兩系統之科技接受度敘述統計分析

本研究針對全體使用者使用 IRT-ACA 以及 TBIR 進行圖像檢索，在科技接受度上是否具有顯著的差異進行分析。除了系統「整體」科技接受度之外，也進一步針對科技接受度的「認知有用性」及「認知易用性」兩個面向進行統計分析，三個面向之平均分數與標準差如表 5-12 所示。結果顯示使用 IRT-ACA 以及 TBIR 進行圖像檢索之兩組使用者科技接受度均高於中間值，科技接受度均屬良好。

表 5-12 使用 TBIR 與 IRT-ACA 系統使用者科技接受度敘述統計分析

項目	組別			
	IRT-ACA (n=12)		TBIR (n=12)	
	平均數	標準差	平均數	標準差
整體科技接受度	4.54	0.38	3.35	0.65
認知有用性	4.52	0.40	2.87	0.82
認知易用性	4.58	0.38	4.17	0.54

#### (二) 兩系統之科技接受度差異分析

為了瞭解分別使用 IRT-ACA 與 TBIR 進行圖像檢索使用者之科技接受度差異，本研究採用曼惠特尼(Mann-Whitney) U 檢定針對 TBIR 與 IRT-ACA 使用者的「整體」科技接受度、「認知有用性」及「認知易用性」等三個科技接受度面向進行檢定分析，結果如表 5-13 所示。結果顯示 IRT-ACA 與 TBIR 在「整體」科技接受度( $U=8.00, p=.001<.05$ )，以及「認知有用性」( $U=2.50, p=.001<.05$ )上達

顯著的差異。而「認知易用性」( $U=39.00, p=.051 > .05$ )則兩組無顯著的差異。

表 5-13 使用 TBIR 與 IRT-ACA 系統使用者科技接受度之曼惠特尼 U 檢定分析

結果

	組別				<i>U</i>	<i>p</i>
	IRT-ACA (n=12)		TBIR (n=12)			
	平均數	標準差	平均數	標準差		
整體科技接受度	4.54	0.38	3.35	0.65	8.00**	<.001
認知有用性	4.52	0.40	2.87	0.82	2.50**	<.001
認知易用性	4.58	0.38	4.17	0.54	39.00	.051

\*\*表  $p < .001$

## 第五節 使用者使用 IRT-ACA 與 TBIR 之系統操作行為歷程紀錄分析

為探討本研究發展之自動情境標註功能所建立之「情境標籤」、自動物件標註功能所建立之「物件標籤」，以及傳統圖像檢索工具透過人力手動標註之圖像 Metadata，在輔助使用者完成圖像檢索任務學習單中的「晚清與外國的互動關係」、「晚清平民百姓生活」的歷程中，是如何與兩三標籤進行互動而有效完成任務。本研究透過使用者行為歷程記錄器，將使用者分別操作 IRT-ACA 與 TBIR 的歷程予以紀錄，行為的詳細事件內容說明如表 5-14 所示。透過此表可用以表達使用者在完成兩個任務階段所使用的各項系統操作行為，據此可採用滯後序列分析 (Lag Sequential Analysis, LSA) 進一步分析高成效使用者的顯著行為轉移序列，以了解實驗對象進行圖像檢索之有效行為模式。

表 5-14 實驗對象使用「IRT-ACA」及「TBIR」之系統操作行為說明

頁面	事件名稱	事件內容	
IRT-ACA	檢索初始化	使用者進入平台的平台首頁，進行檢索設定初始化，並傳回所有圖片。	
	進行檢索	使用者進行檢索。	
	情境檢索	可透過點擊機器自動分類的情境標籤，進而篩選。	
	物件檢索	可透過點擊機器與人合作標註的物件標籤，進而篩選。	
	全文檢索	可自由輸入檢索詞，用以檢索點石齋內文。	
	標題檢索	可自由輸入檢索詞，用以檢索點石齋標題。	
	頁面檢索篩選	使用者結束檢索行為，頁面刷新。	
	檢索清空	清除所有檢索條件。	
	滾動頁面	使用者查看檢索結果的行為。	
	查看圖片	可點擊檢索結果，跳轉到「ACA細看」頁面，用以查看全文、標註資訊及Metadata資訊等。	
ACA細看	縮放圖片	使用者縮放、拖曳圖片。	
	全物件開關	將其全部物件圖層開啟或關閉。	
	單物件開關	將指定物件塗層開啟或關閉。	
	文章閱讀	透過滑鼠位置，捕捉閱讀全文內容之行為。	
	Metadata 閱讀	透過滑鼠位置，捕捉閱讀Metadata之行為。	
筆記	紀錄一筆	輸入一筆紀錄。	
	刪除一筆	刪除一筆紀錄。	
TBIR	檢索初始化	使用者進入平台的平台首頁，進行檢索設定初始化，並傳回所有圖片。	
	進行檢索	使用者進行檢索。	
	全文檢索	可自由輸入檢索詞，用以檢索點石齋內文。	
	標題檢索	可自由輸入檢索詞，用以檢索點石齋標題。	
	頁面檢索篩選	使用者結束檢索行為，頁面刷新。	
	檢索清空	清除所有檢索條件。	
	滾動頁面	使用者查看檢索結果的行為。	
	查看圖片	可點擊檢索結果，跳轉到「TBIR細看」頁面，用以查看全文及Metadata資訊等。	
	TBIR細看	縮放圖片	使用者縮放、拖曳圖片。
		文章閱讀	透過滑鼠位置，捕捉閱讀全文內容之行為。
Metadata 閱讀		透過滑鼠位置，捕捉閱讀Metadata之行為。	

## 一、使用者操作兩系統進行圖像檢索任務之系統操作行為歷程分析

本研究於使用者進行任務學習單中「晚清與外國的互動關係」、「晚清平民百姓生活」任務中，提供使用者一個含有點石齋畫報四大情境的圖像 206 張，讓使用者可以根據不同的檢索任務主題自由地操作系統，以獲取使用者自己認為符合任務的圖像。面對相同的情境圖像檢索任務，TBIR 使用者只能使用全文檢索、標題檢索來檢索圖像，而 IRT-ACA 使用者則可以使用情境檢索、物件檢索、全文檢索及標題檢索來檢索圖像，使用 TBIR 和 IRT-ACA 系統之系統操作行為統計如表 5-15 及 5-16 所示。結果顯示 TBIR 使用者會使用檢索行為包括「全文檢索(5.70%)」、「標題檢索(0.50%)」來幫助查找相關的圖像內容；相較之下，IRT-ACA 使用者會使用包括「情境檢索(5.90%)」、「物件檢索(7.30%)」、「全文檢索(0.50%)」、「標題檢索(0.00%)」來幫助查找相關的圖像內容。此外，由於 IRT-ACA 同時提供「情境標籤」與「物件標籤」兩種類型的標籤，以擴充透過標籤進行圖像檢索的管道，相較於自由下達關鍵字(0.50%)的「全文檢索(0.50%)」及「標題檢索(0.00%)」，使用者更為依賴標籤型之「情境檢索(5.90%)」與「物件檢索(7.30%)」。

表 5-15 使用 TBIR 進行圖像檢索任務之各項行為統計表

TBIR (n=12)			
頁面	事件名稱	出現頻率	百分比
TBIR 細看	文章閱讀	380	25.40%
TBIR 細看	縮放圖片	328	21.90%
TBIR	查看圖片	248	16.60%
筆記	紀錄一筆	124	8.30%

TBIR 細看	Metadata 閱讀	101	6.70%
TBIR	頁面檢索篩選	94	6.30%
TBIR	全文檢索	85	5.70%
TBIR	進行檢索	60	4.00%
TBIR	滾動頁面	26	1.70%
TBIR	檢索初始化	12	0.80%
筆記	刪除一筆	11	0.70%
TBIR	檢索清空	9	0.60%
TBIR	標題檢索	8	0.50%
事件總數		1,486	100.00%

表 5-16 使用 IRT-ACA 進行圖像檢索任務之各項行為統計表

IRT-ACA (n=12)			
頁面	事件名稱	出現頻率	百分比
ACA 細看	縮放圖片	474	18.60%
ACA 細看	文章閱讀	421	16.50%
IRT-ACA	頁面檢索篩選	272	10.70%
ACA 細看	全物件開關	253	9.90%
IRT-ACA	查看圖片	206	8.10%
IRT-ACA	物件檢索	188	7.30%
IRT-ACA	進行檢索	171	6.70%
IRT-ACA	情境檢索	150	5.90%
筆記	紀錄一筆	141	5.50%
IRT-ACA	滾動頁面	69	2.70%



ACA 細看	單物件開關	59	2.30%
ACA 細看	Metadata 閱讀	55	2.10%
IRT-ACA	檢索清空	47	1.80%
IRT-ACA	全文檢索	14	0.50%
筆記	刪除一筆	14	0.50%
IRT-ACA	檢索初始化	12	0.50%
IRT-ACA	標題檢索	1	0.00%
事件總數		2,547	100.0%

## 二、IRT-ACA 高、低分組實驗對象進行圖像檢索任務之系統操作行為轉移模式

### 分析

#### (一) IRT-ACA 高、低分組實驗對象之系統操作行為敘述統計分析

本研究以實驗組實驗對象之檢索任務得分的平均分數作為圖像解讀成效高低分組之標準，將高於或等於平均分數者歸屬為 IRT-ACA 高分組，低於平均分數者歸屬為 IRT-ACA 低分組，IRT-ACA 高、低分組別實驗對象之組別人數統計如表 5-17 所示；IRT-ACA 高、低分組實驗對象之圖像檢索成效敘述統計如表 5-18 所示；IRT-ACA 高、低分組之任務單結果之敘述統計如表 5-19 所示；IRT-ACA 高、低分組高、低分組實驗對象之圖像檢索行為歷程統計表如表 5-20 所示。結果顯示，高分組的「檢索功能(19.4%)」行為比例雖遠低於低分組「檢索功能(30.5%)」，但更充分使用所有的檢索功能；在解讀圖片的面向，高分組的「點擊查看圖片(17.8%)」行為比例遠高於低分組「點擊查看圖片(8.1%)」，以及高分組的「瀏覽圖片資料(19.2%)」行為比例遠高於低分組「瀏覽圖片資料(8.5%)」。而高分組的最終的檢索圖像結果顯示，其「檢索得到圖像平均張數(11.13 張)」與低分組的「檢索得到圖像平均張數(11.00 張)非常相近；然而「解讀圖像情境平均總分」

卻遠高於低分組(37.75 > 20.00)，從這可以觀察出，高分組更擅長於檢索功能的運用，以促使更高成效的檢索後圖像之理解。

表 5-17 IRT-ACA 高、低分組實驗對象之組別人數統計表

組別	IRT-ACA 高分組	IRT-ACA 低分組	合計
人數	8	4	12

表 5-18 IRT-ACA 高、低分組實驗對象解讀圖像情境成效之曼惠特尼 U 檢定

項目	IRT-ACA				U	p
	高分組 (n=8)		低分組 (n=4)			
	平均數	標準差	平均數	標準差		
解讀圖像情境得分	37.75	5.06	20.00	6.98	.000*	.006

\*表 p<0.05

表 5-19 IRT-ACA 高、低分組之任務單結果之敘述統計

	高分組 (n=8)	低分組 (n=4)
檢索得到圖像平均張數	11.13	11.00
解讀圖像情境平均總分	37.75	20.00
平均每張圖片得分	3.39	1.82

表 5-20 IRT-ACA 高、低分組實驗對象之圖像檢索行為歷程統計表

檢索功能	組別			
	IRT-ACA 高分組 (n=8)		IRT-ACA 低分組 (n=4)	
	出現頻率	出現百分比	出現頻率	出現百分比
情境檢索	75	8.4%	75	12.7%
物件檢索	83	9.3%	105	17.8%
全文檢索	14	1.6%	0	0%
標題檢索	1	0.1%	0	0%
小計	173	19.4%	180	30.5%
頁面檢索篩選	138	15.5%	134	22.7%
進行檢索	96	10.8%	74	12.5%
瀏覽圖片資料	171	19.2%	50	8.5%
點擊查看圖片	158	17.8%	48	8.1%
紀錄	84	9.4%	46	7.8%
滾動頁面	32	3.6%	36	6.3%
檢索清空	29	3.3%	18	3.0%
檢索初始化	8	0.9%	4	0.7%
事件總數	889	100.0%	591	100.0%

(二) IRT-ACA 高、低分組別實驗對象之圖像檢索行為轉移分析

IRT-ACA 高、低分組實驗對象之行為轉移圖分別如圖 5-1 及圖 5-2 所示，結果顯示相較於採用 IRT-ACA 之低分組實驗對象，採用 IRT-ACA 之高分組實驗對象在行為轉移模式上最主要的差異為「情境檢索」到「全文檢索」( $Z=9.47 > 1.96$ )

出現顯著轉移，以及「物件檢索」到「標題檢索」( $Z=3.10 > 1.96$ )出現顯著的轉移。顯見採用 IRT-ACA 之高分組實驗對象善於應用「情境檢索」到「全文檢索」，以及「物件檢索」到「標題檢索」之檢索策略搭配，先透過「情境檢索」與「物件檢索」探索擬檢索圖像之可能具有情境與物件特徵下，再決定採用「全文檢索」或者「標題檢索」透過下達檢索詞來精準找到所需圖像。此一結果證明本研究所發展基於自動圖像情境與物件辨識技術所決定之「情境標籤」及「物件標籤」，有助於提升圖像檢索成效。

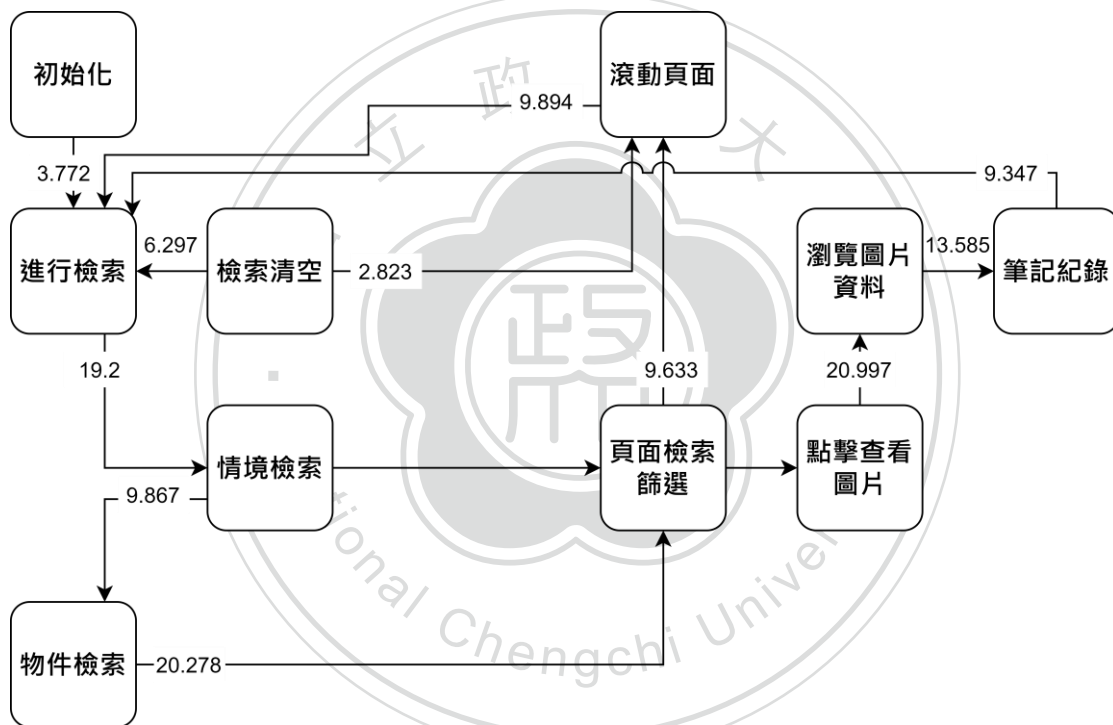


圖 5-1 採用 IRT-ACA 之低分組使用者之系統操作行為轉移序列圖

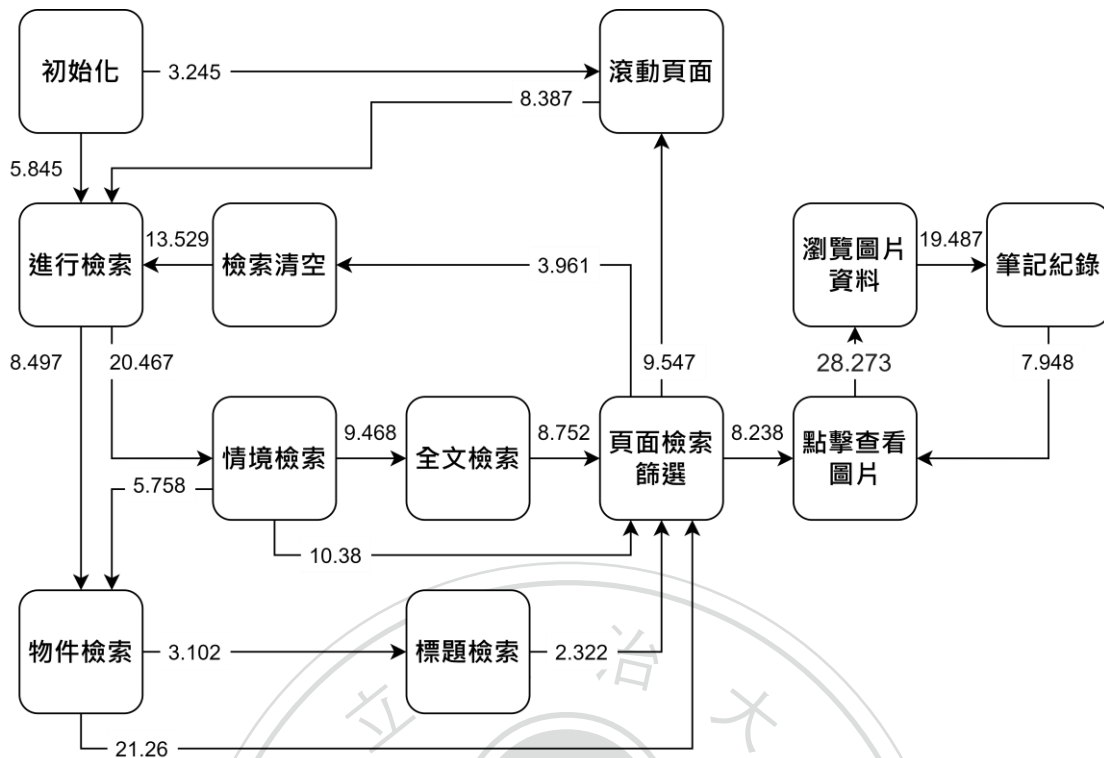


圖 5-2 採用 IRT-ACA 之高分組使用者之系統操作行為轉移序列圖

## 第六節 半結構式訪談質性資料分析

本研究除了採用統計方法分析有無自動情境及物件標註之檢索工具輔以實驗對象探索《點石齋畫報》之圖像解讀成效、科技接受度，以及系統操作行為歷程差異之外，亦透過訪談的方式瞭解實驗對象對於兩個系統輔以圖像脈絡探索上的看法、感受與建議。因此，本研究在實驗結束後邀請 12 名實驗對象進行半結構式深度訪談，藉此瞭解實驗對象的系統使用感受與對於系統後續功能改善的建議，以彌補量化資料分析之不足，並進行實驗結果之交互驗證。以下訪談資料係依照實驗對象的實驗順序進行編號，將其分別以 P1 至 P12 進行編號，以利識別。

### 一、兩個不同系統之圖像檢索功能輔以圖像情境解讀之使用感受比較

#### 1. IRT-ACA 提供之情境檢索使用感受

(1) 正面意見：情境檢索提供了主題探索的方向

「自動情境標註在檢索時，在點進去看全文前，因為有先篩選過，比較好找，不用一個一個去查看每張圖的內文。」(P4)

「在查看搜尋結果時，可以觀察圖片右下角的情境標籤，對圖片做初步判斷。」(P6)

「有時候全文檢索並不知道下什麼關鍵字，而情境檢索可以提供靈感，並且情境標籤可以先將主題上的限縮。」(P11)

「TBIR 全文和標題檢索難以輸入關鍵字，只能嘗試以酒、茶... 等等瑣碎的關鍵字進行檢索，總之情境標籤對於找抽象主題是有幫助。」(P12)

(2) 負面意見：情境檢索提供之主題不明確、準確率仍有疑問

「覺得有到六七成的準確度，但有一些不屬於四個情境的，應該再額外分類。比如說格致彙編，對於定義比較模糊，也有點隸屬於別的情境標籤，有點像是中外紀聞的子分類。」(P1)

「我覺得使用海上繁華標籤時比較困惑，因為似乎其餘三個情境標籤也都屬於海上繁華，所以感受上，海上繁華的主題性比較低。」(P3)

「官場現形，比較難以判斷，因為內文中比較會是以描述事件為主軸，比較難以用略覽的情況下，無從判斷是否與官場現形有關。整體而言，解讀艱澀。」(P6)

## 2. IRT-ACA 提供之物件檢索使用感受

(1) 正面意見：有利於找出圖像中的細節

「它可以針對一些比較細微的東西，比如說國旗，很多時候國旗都很小，一時之間找不到，而物件標籤可以開關，把圖層顏色開啟後，比較容易看到。」(P7)

(2) 負面意見：可選擇的物件太多，建議再進行分類。此外，有些物件標籤不具檢索意義，例如「人」。



「在檢索頁面中，左邊物件數量太多，實驗中想要重新找到望遠鏡這個物件時，發現難以查找，希望能將物件做一點分類，比如物品、載具類。」

(P1)

「人這個標籤，幾乎每張圖都有，所以在使用時並無法篩選出有用訊息。」(P8)

### 3. 情境檢索、物件檢索、全文檢索及標題檢索之交叉檢索行為模式多元

「情境 → 物件。」(P1、P4、P7、P11、P12)

「情境 → 物件 → 全文 → 標題」(P2、P3、P10)

「情境 → 物件 → 標題」(P5)

「情境 → 全文 → 物件 → 標題」(P6)

「情境 → 標題 → 物件」(P8)

「情境 → 物件 → 全文」(P9)

### 4. 哪個系統使用後，會更有助於圖像解讀

(1) 選擇 IRT-ACA 使用者：檢索功能更強，容易找到圖片。此外，情境標籤及物件標籤對於初學者更友善，而自己下關鍵字檢索會很困難。

「應該還是 IRT-ACA，我對點石齋不了解，如果沒有這些標籤輔助，要完成題目有點太過困難。剛剛在使用 TBIR 解晚清平民百姓生活時，只能透過人、民、生、吃、喝... 等瑣碎的關鍵字做搜尋。如果剛剛是先用 IRT-ACA 的話，應該可以用一些賽馬場、望遠鏡、火車、海上繁華這些標籤去找。」(P1)

「剛剛在做晚清與外國的互動關係的時候，因為知道晚清的戰爭頻繁，所以可以用物件標籤，找關於水雷、槍、軍人，就可以很簡單找到相關的圖片。還有情境標籤在這種時候，因為軍人的標籤裡面一定還是參雜著非常

多圖片，所以用中外紀聞再把主題限縮，如果沒有這些功能，我只是想找個打仗的場景的時候就會要找很久。所以在不知道題目是什麼的情況下，我還是會選擇 IRT-ACA。」(P2)

「有提供有很多關鍵字(物件)，有點像是 hashtag，可以很快知道圖片的內容。」(P5)

「我會選擇 IRT-ACA，因為最重要的一點是，它可以很容易找到記憶中的圖，因為點石齋的圖名都比較文言文，比較不會記得，所以標題檢索、全文檢索能就派不上用場。」(P6)

## 二、系統改善建議

「全文的字有點小，希望能夠大一點。」(P1)

「希望有更多的情境標籤。」(P4)

「圖片比對功能，希望能夠一次看多張圖片的原圖。」(P5)

「筆記是蠻好用的，但還是希望有典藏功能。另外搜尋後，希望能夠自訂排序依據。」(P7)

「物件標註還有很多「人」這個標籤並沒有標出來。」(P8)

「Meatadata 全文中，目前仍有不少辨識不出來的字，會呈現「？」，希望能改正。」(P9)

「筆記現在只有新增跟刪除功能，可以添加編輯功能。」(P10)

「希望可以在各個物件標籤上，添加文字註解。」(P11)

## 第七節 綜合討論

本研究進一步將統計分析、系統操作行為序列分析、半結構式質性訪談內容等分析結果與四個研究問題進行對應，以交互驗證方式進行綜合性的研究結果討論。

## 一、自動情境標註功能準確率分析結果與討論

接著從行為分析來分析使用者檢索圖像的方式，結果歸納如表 5-21 所示，結果顯示 TBIR 使用者在「進行檢索(60 次)」的次數上遠低於 IRT-ACA「進行檢索(171 次)」的狀況下，TBIR 的「查看圖片(248 次)」的次數甚至還高於 IRT-ACA 的「查看圖片(206 次)」的次數，顯然透過 TBIR 的檢索功能(全文檢索、標題檢索)查詢到所需圖像的準確率並不高，大部分需要自己透過查看圖片的詳細資料頁面來尋找符合檢索任務的圖像。

表 5-21 使用兩系統進行任務檢索之檢索次數行為統計表

TBIR (n=12)				IRT-ACA (n=12)			
頁面	事件名稱	出現 頻率	百分比	頁面	事件名稱	出現 頻率	百分比
TBIR	進行檢索	60	4.00%	IRT-ACA	進行檢索	171	6.70%
TBIR	全文檢索	85	5.70%	IRT-ACA	物件檢索	188	7.30%
TBIR	標題檢索	8	0.50%	IRT-ACA	情境檢索	150	5.90%
				IRT-ACA	全文檢索	14	0.50%
				IRT-ACA	標題檢索	1	0.00%
TBIR	查看圖片	248	16.60%	IRT-ACA	查看圖片	206	8.10%

再從質性訪談分析來看，根據使用者對於檢索任務以及系統整體感受的訪談內容歸納，在全部 12 位實驗對象中，有一位提到在操作過程中，發現有部分圖片被分類錯誤的情況。此外，有兩人認為「海上繁華」標籤在檢索任務後，仍不能充分了解其涵蓋主題，因此對其定義感到困惑。因此，未來仍有繼續提升 IRT-ACA 在自動情境圖像標註功能準確率的必要性。

## 二、使用者使用 TBIR 與 IRT-ACA 之圖像檢索成效分析結果與討論

本研究基於對抗平衡設計，針對 TBIR 與 IRT-ACA 使用者在不同檢索主題下的圖像檢索成效結果歸納如表 5-22 所示。結果顯示採用採用 IRT-ACA 來檢索圖像情境的組別實驗對象顯著優於 TBIR 的組別實驗對象，顯見本研究在 IRT-ACA 系統上所發展的自動情境及物件標註技術所產生情境標籤及物件標籤，非常有助於提升圖像的檢索成效。

表 5-22 TBIR 與 IRT-ACA 使用者在檢索任務下的圖像檢索成效分析結果歸納

項目	平均分數比較	統計顯著性
解讀圖像情境得分	IRT-ACA > TBIR	*

\*表  $p < .05$

根據行為分析來觀察使用者在檢索任務下所進行筆記的狀況，歸納結果如表 5-23 所示，結果顯示使用者在使用 IRT-ACA 系統的「查看圖片(206 次)」行為略低於 TBIR 的「查看圖片(248 次)」次數，而使用者在 IRT-ACA 系統的筆記「紀錄一筆(141 次)」之行為次數略高於 TBIR 的「紀錄一筆(124 次)」的行為次數。此一結果顯示在使用者查看圖片並判斷為符合任務主題而將其放入筆記的轉移率上，IRT-ACA 優於 TBIR ( $68.45\% > 50\%$ )，其計算方式如公式(4)所示。換句話說，在相同的時間下，IRT-ACA 使用者解讀圖像情境之成效優於 TBIR 使用者。

$$\text{查看圖片並判斷為符合任務主題的轉移率} = \frac{\text{筆記一筆總次數}}{\text{查看圖片總次數}} \quad (4)$$

表 5-23 使用兩系統進行主題檢索之筆記次數行為統計表

TBIR (n=12)				IRT-ACA (n=12)			
頁面	事件名稱	出現	百分比	頁面	事件名稱	出現	百分比

		頻率				頻率	
TBIR	查看圖片	248	16.60%	IRT-ACA	查看圖片	206	8.10%
筆記	刪除一筆	11	0.70%	筆記	刪除一筆	14	0.50%
筆記	紀錄一筆	124	8.30%	筆記	紀錄一筆	141	5.50%

接著從質性訪談分析來看，根據使用者對於檢索任務以及系統整體感受的訪談內容歸納，使用者指出情境標籤與物件標籤可以有效支援所需圖像的檢索主題。例如，實驗對象提到「晚清與外國的互動關係」會使用到情境標籤之「中外紀聞」與「海上繁華」，以及物件標籤之「炮」、「船」、「軍隊」及「城牆」等；而「晚清的平民百姓生活」則會使用到「海上繁華」與「格致彙編」，以及物件標籤之「人」、「馬車」及「女人」等。而兩種標籤提供的幫助並不相同，其中情境標籤能傳達圖像整體的抽象主題，而物件標籤則能輔以更加具體的圖像內容，在交叉使用下更能找到符合任務的圖像，進而能提升圖像檢索的效能。

整體而言，情境標籤與物件標籤因擁有截然不同的特性，IRT-ACA 結合了情境檢索與物件檢索，使得系統能更全面地幫助使用者在不同狀況下進行圖像檢索，進而提高圖像檢索與解讀圖像情境之成效。

### 三、使用者使用 TBIR 與 IRT-ACA 之科技接受度差異分析結果與討論

本研究為探討使用 TBIR 與 IRT-ACA 使用者之科技接受度差異，以使用者在分別使用完 TBIR 與 IRT-ACA 輔以圖像檢索任務後所填寫之科技接受度問卷，作為評估分別使用兩個系統使用者的科技接受度差異分析之依據，結果歸納如表 5-24 所示。結果顯示使用者在使用完 IRT-ACA 當下的科技接受度，不論在「整體」、「認知有用性」、「認知易用性」上的平均數皆優於 TBIR 的科技接受度感受，並且在「整體」科技接受度、科技接受度之「認知有用性」上達到統計上的顯著

差異，並且 IRT-ACA 顯著優於 TBIR，而科技接受度之「認知易用性」上兩個系統未達統計上的顯著差異。

表 5-24 採用 TBIR 與 IRT-ACA 輔以圖像檢索使用者之科技接受度分析結果歸

納

科技接受度面向	平均數比較	統計顯著性
整體	IRT-ACA > TBIR	**
認知有用性	IRT-ACA > TBIR	**
認知易用性	IRT-ACA > TBIR	N.S.

\*\*表  $p < .001$

N.S.表示統計檢定結果不顯著

為了瞭解使用者對於 TBIR 與 IRT-ACA 科技接受度的「認知有用性」與「認知易用性」感受差異，本研究從質性訪談分析中進一步探討。根據使用者對於任務的訪談內容發現，多數使用者對於《點石齋畫報》了解不深，因此相較於以自由輸入關鍵字的方式進行圖像檢索，採用 IRT-ACA 的使用者更加依賴於包括情境標籤與物件標籤之標籤化的檢索。其中，使用者表示情境標籤可以傳達圖像的主題性，而物件標籤則可以傳達圖片的內容性，並藉由這兩種標籤輔助而快速的了解圖像的內容。因此，使用者對於 IRT-ACA 的「認知有用性」感受度顯著優於 TBIR。例如有使用者表達「我覺得使用【IRT-ACA】可以更快速地檢索到需要的圖像」，以及「我覺得使用【IRT-ACA】讓檢索到的圖像內容更豐富」。此外，由於兩個系統在操作層面上，僅在於檢索方式以及 Metadata 資訊量具有差異，其他的功能、操作體驗上都是完全一致。因此，多數使用者表示兩個系統在操作上並無太大的差異，這可能是使用者對於 IRT-ACA 及 TBIR 的「認知易用性」未達顯著差異的可能原因。



整體而言，儘管目前 IRT-ACA 仍存在一些需要進一步改進的問題，但使用者對於 IRT-ACA 提供的額外幫助與未來發展皆給予相當的肯定。因此，使用者也不吝給予了相當多的改進建議，以利於未來 IRT-ACA 系統功能的進一步優化與改進，進而發展成為一個可以有效輔助使用者解讀圖像情境的數位人文工具。



## 第六章 結論與建議

本研究旨在探討使用者採用 TBIR 與 IRT-ACA 輔助解讀圖像情境，對於機器在檢索圖像的情境分類準確率，以及使用者在解讀圖像情境成效、科技接受度之影響差異。此外，也探討使用者採用 TBIR 與 IRT-ACA 輔助解讀圖像之行為模式，並輔以深度訪談瞭解其使用感受與滿意度。本章歸納本研究之研究結論，共分為三小節：第一節為結論；第二節為提出能夠改善 IRT-ACA 的具體建議，以期許未來能發展為一個更全面輔助使用者解讀圖像情境之數位人文工具；第三節為未來研究方向。

### 第一節 結論

#### 一、自動情境標註功能的準確率已足以有效輔助使用者解讀圖像情境

本研究根據情境分類準確率之評估結果顯示，IRT-ACA 所發展之圖像辨識模型已具有一定水準之圖像情境分類準確率。而從使用者操作系統的使用行為分析得知，IRT-ACA 的使用者對於自動情境圖像標註功能所提供之「情境標籤」，對應其應用在系統之「情境檢索」的使用優先順位上，都較「物件檢索」、「全文檢索」，以及「標題檢索」高上許多。此外，使用「情境標籤」進行檢索時，多數使用 IRT-ACA 實驗對象會使用「情境標籤」與「物件標籤」來進行交互檢索，以獲得想要的檢索圖像。而從訪談質性分析顯示，使用者對於「情境標籤」提升檢索效能的幫助給予肯定，普遍覺得「情境標籤」具有呈現抽象主題的作用，以及可作為理解圖像情境時的參考。綜合以上，本研究所發展的自動情境標註功能準確率，已足以有效輔助使用者解讀圖像情境。

## 二、使用者使用 TBIR 與 IRT-ACA 解讀圖像情境之成效呈現顯著的差異

本研究根據曼惠特尼 U 檢定之分析結果顯示，使用 TBIR 與 IRT-ACA 解讀圖像情境之成效呈現顯著的差異，並且使用 IRT-ACA 輔助使用者解讀圖像情境的成效顯著優於 TBIR。而由訪談質性分析得知，「情境標籤」與「物件標籤」因為各自有獨特的性質，因此適用的檢索主題並不相同。抽象的「情境標籤」對於概念式的主題能提供更多訊息，而具體的「物件標籤」則對於圖像中出現的實體能提供更多訊息。因此，提供這兩種不同類型標籤的 IRT-ACA 系統，能同時滿足使用者各種不同檢索需求，有效提升圖像的檢索成效。

## 三、使用者使用 TBIR 與 IRT-ACA 的科技接受度呈現顯著的差異

本研究根據曼惠特尼 U 檢定之分析結果顯示，使用 TBIR 與 IRT-ACA 輔以圖像檢索任務的「整體」科技接受度與「認知有用性」呈現顯著的差異，並且 IRT-ACA 顯著優於用 TBIR，表示使用者對於具有自動情境標註的 IRT-ACA 系統，具有更高的科技接受度，特別是在「認知有用性」上。從訪談質性分析得知，實驗對象相當肯定具「情境標籤」之 IRT-ACA 提供的圖像內容性資訊，以及「物件標籤」提供了細緻的圖像內容訊息，以及因此提供了更為多元的檢索管道。因此，目前 IRT-ACA 作為輔助使用者解讀圖像情境的數位人文工具而言，已經提供許多具有幫助和發展性的功能，但仍有許多需要進一步改進的問題，未來才能成為既有用又易用的數位人文工具。

## 四、IRT-ACA 的標籤型檢索比起自由下達關鍵字的檢索更能促進實驗對象的檢

### 索意願

本研究根據行為歷程記錄分析顯示，實驗對象在 IRT-ACA 的檢索任務中，使用標籤型檢索的「情境檢索(150 次)」與「物件檢索(188 次)」次數，遠高於自

由下達關鍵字的「全文檢索(14次)」與「標題檢索(1次)」。

此外，在訪談中得知，多數實驗對象認為以自由下達關鍵字的方式進行圖像檢索，對他們而言是相當困難的只能嘗試以酒、茶...等瑣碎的關鍵字進行檢索，而標籤型檢索則可以提供他們靈感，進而提升他們的檢索意願。

#### 五、使用 IRT-ACA 高分組使用者之檢索行為更充分使用所有檢索功能

本研究根據 IRT-ACA 的行為歷程記錄分析顯示，使用 IRT-ACA 高分組實驗對象善於高效的應用各種檢索策略，特別是使用 IRT-ACA 高分組實驗對象在「情境檢索」到「全文檢索」，以及「物件檢索」到「標題檢索」皆出現顯著行為轉移，而低分組則無。顯示使用 IRT-ACA 高分組實驗對象會先透過「情境檢索」與「物件檢索」探索擬檢索圖像之可能具有情境與物件特徵下，再決定採用「全文檢索」或者「標題檢索」下達檢索詞來精準找到所需圖像。此一結果證明本研究所發展基於自動圖像情境與物件辨識技術所決定之「情境標籤」及「物件標籤」，有助於提升圖像檢索成效。

#### 六、IRT-ACA 使用者之查看圖像至利用筆記紀錄檢索結果之轉移率高於 TBIR

本研究根據行為歷程記錄分析結果顯示，在 IRT-ACA 與 TBIR 使用者查看圖片並判斷為符合任務主題而將其放入筆記的轉移率上，IRT-ACA 優於 TBIR (68.45% > 50%)，表示 IRT-ACA 使用者在檢索到的圖片中，更能找到符合自己預期之結果，因而解讀圖像情境之成效優於 TBIR 使用者。換句話說，在相同的時間下，IRT-ACA 能幫助使用者更快速捕捉到圖像所表達的情境與物件特徵，進而提升圖像探索成效。

## 第二節 IRT-ACA 之系統改進建議

根據本研究針對不同面向之實驗資料分析結果顯示，本研究發展之 IRT-ACA 對於使用者解讀圖像確實具有幫助，但仍存在許多需要進一步改善的地方。透過訪談，許多實驗對象提出了一些核心需要改善的問題與建議，包括自動情境標註功能以及研究功能面向的改善，具體歸納說明如下：

### 一、自動情境標註功能面向需要改善之功能

#### (一) 擴展可以自動辨識的情境標籤種類

目前本研究僅完成圖像檢索任務與解讀的研究初探，並不確定完整的 4000 餘幅《點石齋畫報》涵蓋了多少種類的抽象情境。透過實驗對象的訪談回饋，以及協助本研究進行圖像標註的人員反饋中，可以得知目前使用的中外紀聞、官場現形、格致彙編，以及海上繁華等四種圖像情境種類，尚不能完整覆蓋《點石齋畫報》的所有圖像情境主題，並有出現主題重疊的問題，例如有些歸屬於海上繁華的圖像，亦屬於中外紀聞。因此，未來系統應該擴展可以自動辨識的情境標籤種類，提供更完整的情境類別給使用者。

#### (二) 更進一步的圖像情境辨識模型訓練以提升圖像情境識別的準確率效能

有部分實驗對象於訪談中提及情境標籤的準確率尚須提升的問題，目前本研究發展的情境識別模型的識別能力仍相當有限，尚有許多方法可以提高其準確率。例如，使用圖像物件標註量更加完整的訓練集來進行圖像情境辨識模型的訓練，甚至應該考慮圖像物件在圖像中的空間分布來提高圖像情境識別的準確率。

#### (三) 改善圖像物件分類過多的問題

有部分實驗對象於訪談中提及 IRT-ACA 所提供的物件標籤種類過於龐大且複雜，而降低他們檢索時查找特定物件標籤的效率。因此，未來在眾多的標籤類別中進行類別分類時，應將諸如人物相關的標籤與物品相關的標籤分開，以

增進實驗對象對於物件標籤查找時的便利性與有用性。

## 二、支援數位人文研究功能面向需要改善之功能

### (一) 改善筆記的功能

目前 IRT-ACA 的圖像標註頁面所附帶提供的筆記功能尚過於少，僅能作新增及刪除功能，未來應修改為能讓使用者進行自由修改的功能，以利兩種不同類型標籤在輔助圖像檢索上能夠更加相輔相成。

### (二) 添加圖片比對功能

有部分使用者於訪談中提及希望能有圖像與圖像之比對功能，如此在比對或者是檢索同個主題的數位人文研究中，可提供另一個不同的圖像檢索管道。

## 第三節 未來研究方向

### 一、完善情境標籤，更精準的將情境標籤分類出來

本研究旨在藉由機器的自動判讀與分類，為圖像添加情境與物件之後設資料標籤，研究結果顯示抽象的情境標籤可帶給使用者更好的檢索體驗。目前本研究選用的四個圖像情境標籤仍然無法完整涵蓋所有《點石齋畫報》的圖像情境內涵，而透過訪談所獲得的回饋發現，未來可以再添加諸如「道德故事」與「怪力亂神」兩個圖像情境標籤。因此，未來可考慮發展具有人機合作的自動情境分類功能。

### 二、改善圖像情境識別效能

本研究旨在發展一個能有效輔助數位人文學者解讀圖像情境之數位人文工具，研究過程並無法獲得足夠數量的具不同圖像情境之物件標註圖像作為訓練集，故以人機合作的方式開始進行圖像物件標註，但所獲得之圖像訓練集仍然不



足。未來應該以目前所獲得的 206 張圖像的標註內容為基礎，持續擴展其圖像的訓練集，以提高圖像情境分類的準確率。

### 三、完善神經網路架構

本研究分為兩階段完成物件及情境標註，第一階段採用 Mask R-CNN 完成物件標註，第二階段採用 TF-IDF 與 SVM 完成情境標註。但是因為研究限制的關係，物件辨識的多樣性不如預期，僅對於人、碗、椅子等常見物件有較好的效果，而目前圖像物件標註的資料量可以重新訓練模型，提高物件辨識的效果後，未來希望情境辨識能整合進 Mask R-CNN 的網路架構，將物件辨識與情境辨識一起於神經網路中完成。



## 參考文獻

### 一、中文文獻

- 陳勇汀 (2017)。行為順序檢定：滯後序列分析/ Behavior Analysis: Lag Sequential Analysis。檢自 <https://pulipulichen.github.io/HTML-Lag-Sequential-Analysis/>
- 項潔、陳麗華 (2014)。數位人文－學科對話與融合的新領域。數位人文研究與技藝 (頁 9-23)

### 二、英文文獻

- Agosti, M., Ferro, N., Orio, N., & Ponchia, C. (2014). CULTURA outcomes for improving the user's engagement with cultural heritage collections. *Procedia Computer Science*, 38, 34-39. doi:10.1016/j.procs.2014.10.007
- Beaudoin, J. E. (2014). A framework of image use among archaeologists, architects, art historians and artists. *Journal of Documentation*, 70(1), 119-147. doi:10.1108/JD-12-2012-0157
- Beaudoin, J. E., & Brady, J. E. (2011). Finding visual information: a study of image resources used by archaeologists, architects, art historians, and artists. *Art Documentation: Journal of the Art Libraries Society of North America*, 30(2), 24-36. doi:10.1086/adx.30.2.41244062
- Bradshaw, B. (2000). Semantic based image retrieval: a probabilistic approach. In *Proceedings of the eighth ACM international conference on Multimedia*, 167-176. doi:10.1145/354384.354456
- Brooks, J. (2019). *COCO Annotator*. Jsbroks/coco-annotator. <https://github.com/jsbroks/coco-annotator/>
- Burdescu, D. D., Mihai, C. G., Stanescu, L., & Brezovan, M. (2013). Automatic image annotation and semantic based image retrieval for medical domain. *Neurocomputing*, 109, 33-48. doi:10.1016/j.neucom.2012.07.030
- Carneiro, G., Chan, A. B., Moreno, P. J., & Vasconcelos, N. (2007). Supervised learning of semantic classes for image annotation and retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(3), 394-410. doi:10.1109/TPAMI.2007.61
- Chen, C.-M., & Tsay, M.-Y. (2017). Applications of collaborative annotation system in digital curation, crowdsourcing, and digital humanities. *The Electronic Library*, 35(6), 1122-1140. doi:10.1108/EL-08-2016-0172
- Chen, S. H., & Chen, Y. H. (2017). A content-based image retrieval method based on the google cloud vision api and wordnet. *Intelligent Information and Database Systems*, 651-662. doi:10.1007/978-3-319-54472-4\_61
- Chen, Y., Zhou, X. S., & Huang, T. S. (2001). One-class SVM for learning in image

- retrieval. In *Proceedings 2001 International Conference on Image Processing, IEEE 2001*, 34-37. doi:10.1109/ICIP.2001.958946
- Cheng, Q., Zhang, Q., Fu, P., Tu, C., & Li, S. (2018). A survey and analysis on automatic image annotation. *Pattern Recognition*, 79, 242-259. doi:10.1016/j.patcog.2018.02.017
- Eakins, J., & Graham, M. (1999). Content-based image retrieval.
- Gordo, A., Almazán, J., Revaud, J., & Larlus, D. (2016). Deep image retrieval: learning global representations for image search. *Computer Vision – ECCV 2016*, 9910, 241-257. doi:10.1007/978-3-319-46466-4\_15
- Hare, J. S., Lewis, P. H., Enser, P. G., & Sandom, C. J. (2006). Mind the gap: another look at the problem of the semantic gap in image retrieval. In *Multimedia Content Analysis, Management, and Retrieval 2006*, 6073, 607309.1-607309.12. doi:10.1117/12.647755
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision, CVPR 2017*, 2961-2969. doi:10.1109/ICCV.2017.322
- Henningsmeier, J. (1998). The foreign sources of Dianshizhai huabao 點石齋畫報, A nineteenth century Shanghai illustrated magazine. *Ming qing yanjiu*, 7(1), 59-91. doi:10.1163/24684791-90000374
- Huang, C., Xu, H., Xie, L., Zhu, J., Xu, C., & Tang, Y. (2018). Large-scale semantic web image retrieval using bimodal deep learning techniques. *Information Sciences*, 430-431, 331-348. doi:10.1016/j.ins.2017.11.043
- Huang, Z., Zhong, Z., Sun, L., & Huo, Q. (2019). Mask R-CNN with pyramid attention network for scene text detection. In *2019 IEEE Winter Conference on Applications of Computer Vision, IEEE 2019*, 764-772. doi:10.1109/WACV.2019.00086
- Hwang, G. J., Yang, L. H., & Wang, S. Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education*, 69, 121-130. doi:10.1016/j.compedu.2013.07.008
- Hyvönen, E., Saarela, S., Styrman, A., & Viljanen, K. (2003). Ontology-based image retrieval. In *Proceedings of XML Finland Conference*, 27-51. Retrieved from <https://seco.cs.aalto.fi/publications/2002/hyvonen-styrman-saarela-ontology-based-image-retrieval-2002.pdf>
- Ivasic-Kos, M., Ipsic, I., & Ribaric, S. (2015). A knowledge-based multi-layered image annotation system. *Expert Systems with Applications*, 42(24), 9539-9553. doi:10.1016/j.eswa.2015.07.068
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105. <https://doi.org/10.1145/3065386>
- Li, Z., & Tang, J. (2015). Weakly supervised deep metric learning for community-contributed image retrieval. *IEEE Transactions on Multimedia*, 17(11), 1989-1999.

doi:10.1109/TMM.2015.2477035

- Liu, Y., Zhang, D., & Lu, G. (2008). Region-based image retrieval with high-level semantics using decision tree learning. *Pattern Recognition*, 41(8), 2554-2570. doi:10.1016/j.patcog.2007.12.003
- Liu, Y., Zhang, D., Lu, G., & Ma, W. Y. (2007). A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40(1), 262-282. doi:10.1016/j.patcog.2006.04.045
- Llamas, J., Leronés, P. M., Zalama, E., & Gómez-García-Bermejo, J. (2016). Applying deep learning techniques to cultural heritage images within the inception project. *Progress in Cultural Heritage: Documentation, Preservation, and Protection*, 25-32. doi:10.1007/978-3-319-48974-2\_4
- Lorang, E., Soh, L.-K., Datla, M. V., & Kulwicki, S. (2015). Developing an image-named classifier for detecting poetic content in historic newspaper collections. *D-Lib Magazine*, 21(7/8). doi:10.1045/july2015-lorang
- Lowe, D. G. (1999). Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision, IEEE 1999*, 1150-1157. doi:10.1109/ICCV.1999.790410
- Murthy, V. N., Maji, S., & Manmatha, R. (2015). Automatic image annotation using deep learning representations. In *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, 603-606. doi:10.1145/2671188.2749391
- Murthy, V. S. V. S., Vamsidhar, E., Kumar, J. S., & Rao, P. S. (2010). Content based image retrieval using Hierarchical and K-means clustering techniques. *International Journal of Engineering Science and Technology*, 2(3), 209-212.
- Nguyen, H. V., & Bai, L. (2011). Cosine similarity metric learning for face verification. In *Computer Vision – ACCV 2010*, 709-720. doi:10.1007/978-3-642-19309-5\_55
- Özyurt, F., Tuncer, T., Avci, E., Koç, M., & Serhatlioğlu, İ. (2019). A novel liver image classification method using perceptual hash-based convolutional neural network. *Arabian Journal for Science and Engineering*, 44(4), 3173-3182. doi:10.1007/s13369-018-3454-1
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition, CVPR 2016*, 779-788. doi: 10.1109/CVPR.2016.91
- Rui, Y., & Huang, T. (2000). Optimizing learning in image retrieval. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2000*, 236-243. doi: 10.1109/CVPR.2000.855825
- Schreibman, Susan. (2012). Digital humanities: centres and peripheries. *Historical Social Research-Historische Sozialforschung*, 37(3), 46-58.
- Shyu, C. R. (2000). Relevance feedback decision trees in content-based image retrieval. In *2000 Proceedings Workshop on Content-based Access of Image and Video Libraries*, 68-72.

- Su, J. H., Wang, B. W., Yeh, H. H., & Tseng, V. S. (2009). Ontology-based semantic web image retrieval by utilizing textual and visual annotations. In *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, 425-428. Milan, Italy: IEEE. doi:10.1109/WI-IAT.2009.317
- Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 1349-1380. doi:10.1109/34.895972
- Terras, M. (2012). Image processing and digital humanities. *Digital Humanities in Practice*, 71-90. Facet. Retrieved from <http://discovery.ucl.ac.uk/1327983/>
- Vijayarajan, V., Dinakaran, M., Tejaswin, P., & Lohani, M. (2016). A generic framework for ontology-based information retrieval and image retrieval in web data. *Human-Centric Computing and Information Sciences*, 6(1), 18. doi:10.1186/s13673-016-0074-1
- Wan, H. L., & Chowdhury, M. (2003). Image semantic classification by using SVM. *Journal of software*, 14(11), 1891-1899.
- Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., ... & Liu, W. (2018). Cosface: Large margin cosine loss for deep face recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition, CVPR 2018*, 5265-5274. doi:10.1109/CVPR.2018.00552
- Wang, L. Z., & Gu, X. E. (2007). Dian-Shi-Zhai pictorial suiting both refined and popular tastes. *Journal of Shanxi Normal University (Social Science Edition)*, 4.
- Weinberger, K. Q., & Saul, L. K. (2009). Distance metric learning for large margin nearest neighbor classification. *Journal of machine learning research*, 10(2).
- Xiaoqing, Y. (2003). *The Dianshizhai pictorial: Shanghai urban life, 1884-1898* (No. 98). Ann Arbor, MI: University of Michigan Press.
- Yin, S., Chen, W., & Qin, X. (2009). Research on semantic network image retrieval method. In *2009 International Conference on Future BioMedical Information Engineering*, 449-452. doi: 10.1109/FBIE.2009.5405823
- Zhang, D., Islam, Md. M., & Lu, G. (2012). A review on automatic image annotation techniques. *Pattern Recognition*, 45(1), 346-362. doi:10.1016/j.patcog.2011.05.013
- Zhang, Y. J. (2006). *Semantic-based visual information retrieval*. Pennsylvania, PA: IGI Global.



## 附錄

### 附錄一 訪談大綱

#### (任務一) 晚清與外國的互動關係

1. 在 IRT-ACA 或 TBIR 中，都提供了圖像檢索功能，請問在晚清與外國的互動關係的圖像檢索過程中，兩個系統是否都能幫助您找到您想找到的圖像？為什麼？
2. 您的檢索歷程為何？
3. 在任務一中，任務希望您找到關於晚清與外國的互動關係的圖像，請說明您如何使用目前的系統完成之，使用了什麼情境標籤、物件標籤以及如何下達關鍵字？
4. 對於該系統的使用感受？

#### (任務二) 晚清的平民百姓生活

1. 在 IRT-ACA 或 TBIR 中，都提供了圖像檢索功能，請問在晚清的平民百姓生活的圖像檢索過程中，兩個系統是否都能幫助您找到您想找到的圖像？為什麼？
2. 您的檢索歷程為何？為什麼？
3. 在任務二中，任務希望您找到關於平民百姓各種生活面向的圖像，請說明您如何使用目前的系統完成之，使用了什麼情境標籤、物件標籤以及如何下達關鍵字？
4. 對於該系統的使用感受？

#### 整體

1. 整體而言，在兩個圖像檢索任務中，IRT-ACA 與 TBIR 何者給予您更多的幫助？為什麼？若重新給予您一個圖像檢索任務，您會更傾向選擇使用哪一個系統來輔助您完成？
2. 情境標籤是由機器標註，詢問是否對情境標籤有什麼改善上的建議？
3. 對於本研究所開發的 IRT-ACA 有什麼使用者介面及功能改善上的建議？



## 附錄二 IRT-ACA 科技接受度問卷

姓名：\_\_\_\_\_ 填寫日期：西元\_\_\_\_\_年\_\_\_\_月\_\_\_\_日

您好：

這份「IRT-ACA 科技接受度量表」問卷，主要的目的在瞭解您使用 IRT-ACA 來完成任務單時的感受或感想，填答時請依照您實際的想法或經驗填寫。本研究收集之資料僅供學術研究使用，所有資料將會被匿名保密，不做其他用途，請您放心填答。

國立政治大學圖書資訊與檔案學研究所

研究生：趙映翔

指導教授：陳志銘博士

填答說明：

■ 每個問題有五個選項：

- 5 表示您「非常同意」這個敘述。
- 4 表示您「同意」這個敘述。
- 3 表示您對這個敘述「同意和不同意程度相當」。
- 2 表示您「不同意」這個敘述。
- 1 表示您「非常不同意」這個敘述。

■ 請圈選出您覺得最適合的選項：

依據您個人的實際想法進行填答，例如：如果您對於「我覺得點石齋數位資料庫是很棒的網站」這個敘述表示「非常同意」，則將「非常同意」的數字 5 圈起來，⑤，如下所示。

### 科技接受度

編號	題目	非常不同意	不同意	程度和相當	同意	非常同意
例題	我覺得點石齋數位資料庫是很棒的網站。	1	2	3	4	⑤

### 科技接受度

編號	題目	非常不同意	不同意	程度和相當	同意	非常同意
1	我覺得使用【IRT-ACA】讓建置的圖像後設資料內容更豐富。	1	2	3	4	5
2	我覺得使用【IRT-ACA】讓檢索到的圖像內容更豐富。	1	2	3	4	5
3	我覺得使用【IRT-ACA】可以更快速地檢索到需要的圖像。	1	2	3	4	5
4	我覺得使用【IRT-ACA】讓解讀到的圖像情境內容更豐富。	1	2	3	4	5
5	我覺得使用【IRT-ACA】可以更快速地解讀圖像情境。	1	2	3	4	5
6	我很快便瞭解【IRT-ACA】的操作方式。	1	2	3	4	5
7	我覺得【IRT-ACA】的系統介面很容易使用。	1	2	3	4	5
8	我覺得【IRT-ACA】的操作過程順暢沒有困難。	1	2	3	4	5

資料來源：本研究參考 Hwang、Yang 與 Wang (2013) 修改。

### 附錄三 TBIR 科技接受度問卷

姓名：\_\_\_\_\_ 填寫日期：西元\_\_\_\_\_年\_\_\_\_月\_\_\_\_日

您好：

這份「TBIR 科技接受度量表」問卷，主要的目的在瞭解您使用 TBIR 來完成任務單時的感受或感想，填答時請依照您實際的想法或經驗填寫。本研究收集之資料僅供學術研究使用，所有資料將會被匿名保密，不做其他用途，請您放心填答。

國立政治大學圖書資訊與檔案學研究所

研究生：趙映翔

指導教授：陳志銘博士

填答說明：

■ 每個問題有五個選項：

- 5 表示您「非常同意」這個敘述。
- 4 表示您「同意」這個敘述。
- 3 表示您對這個敘述「同意和不同意程度相當」。
- 2 表示您「不同意」這個敘述。
- 1 表示您「非常不同意」這個敘述。

■ 請圈選出您覺得最適合的選項：

依據您個人的實際想法進行填答，例如：如果您對於「我覺得點石齋數位資料庫是很棒的網站」這個敘述表示「非常同意」，則將「非常同意」的數字 5 圈起來，⑤，如下所示。

### 科技接受度

編號	題目	非常不同意	不同意	程度相當	同意和不同意	同意	非常同意
例題	我覺得點石齋數位資料庫是很棒的網站。	1	2	3	4	5	⑤

### 科技接受度

編號	題目	非常不同意	不同意	程度相當	同意和不同意	同意	非常同意
1	我覺得使用【TBIR】讓建置的圖像後設資料內容更豐富。	1	2	3	4	5	
2	我覺得使用【TBIR】讓檢索到的圖像內容更豐富。	1	2	3	4	5	
3	我覺得使用【TBIR】可以更快速地檢索到需要的圖像。	1	2	3	4	5	
4	我覺得使用【TBIR】讓解讀到的圖像情境內容更豐富。	1	2	3	4	5	
5	我覺得使用【TBIR】可以更快速地解讀圖像情境。	1	2	3	4	5	
6	我很快便瞭解【TBIR】的操作方式。	1	2	3	4	5	
7	我覺得【TBIR】的系統介面很容易使用。	1	2	3	4	5	
8	我覺得【TBIR】的操作過程順暢沒有困難。	1	2	3	4	5	

資料來源：本研究參考 Hwang、Yang 與 Wang (2013) 修改。