

# 크라우드소싱 기반 이미지 태깅 시스템 구축 연구\*

## Development of an Image Tagging System Based on Crowdsourcing

이 혜 영 (Hyeyoung Lee)\*\*

장 윤 금 (Yunkeum Chang)\*\*\*

### 초 록

본 연구는 이미지에 대한 접근 및 검색을 향상시키고, 이미지에 대한 설명 제공 도구로서의 태그를 효과적으로 생성하기 위한 방안을 모색하는데 목적이 있다. 이를 위해 이미지 태그를 생성하는 두 가지 방법인 휴먼 태깅과 머신 태깅의 특징을 조사하고 휴먼 태그와 머신 태그의 속성을 비교 분석하였다. 머신 태그는 일반적 속성이 가장 높았으며, 특정적 속성과 시각적 요소는 일부 나타났고, 추상적 속성은 거의 나타나지 않았다. 휴먼 태그도 일반적 속성이 가장 높았으나 휴먼 태그 생성자가 명칭을 알 수 있는 객체 및 장면에 대해서는 특정적 속성의 비중이 높았으며, 감정과 정서, 추상적 개념의 주제뿐 아니라 사건, 장소, 시간, 관계 등이 다양한 태그로 표현되었다. 본 연구를 통해 생성된 태그 집합은 머신러닝 알고리즘을 개선하기 위한 트레이닝 데이터셋을 구성하는데 기초 자료로 활용될 수 있을 것이다.

### ABSTRACT

This study aims to improve the access and retrieval of images and to find a way to effectively generate tags as a tool for providing explanation of images. To do this, this study investigated the features of human tagging and machine tagging, and compare and analyze them. Machine tags had the highest general attributes, some specific attributes and visual elements, and few abstract attributes. The general attribute of the human tag was the highest, but the specific attribute was high for the object and scene where the human tag constructor can recognize the name. In addition, sentiments and emotions, as well as subjects of abstract concepts, events, places, time, and relationships are represented by various tags. The tag set generated through this study can be used as basic data for constructing training data set to improve the machine learning algorithm.

키워드: 이미지, 태깅, 이미지 태깅 시스템, 크라우드소싱, 휴먼 태그, 머신 태그

Image, Tagging, Image Tagging System, Crowdsourcing, Human Tag, Machine Tag

\* 이 논문은 숙명여자대학교 대학원 박사학위논문 "크라우드소싱 기반 이미지 태깅 시스템 구축 연구" 중 일부를 발췌·요약·수정한 것임.

\*\* 숙명여자대학교 문헌정보학과 강사(juejong@gmail.com) (제1저자)

\*\*\* 숙명여자대학교 문헌정보학과 교수(yunkeum@sookmyung.ac.kr) (교신저자)

논문접수일자 : 2018년 8월 21일 논문심사일자 : 2018년 8월 27일 게재확정일자 : 2018년 9월 7일  
한국비블리아학회지, 29(3): 297-320, 2018. [<http://dx.doi.org/10.14699/kbiblia.2018.29.3.297>]

## 1. 서론

정보통신기술 발달로 웹 서비스 및 모바일 기기의 사용이 확산되면서 정보 자원으로서 이미지, 동영상 등 시각정보에 대한 이용자의 정보요구와 이용이 급증하고 있다. 사람들은 개인의 디지털 기기로 손쉽게 시각정보를 생산, 저장하게 되었으며, 검색엔진을 통해 방대한 인터넷 자원을 검색할 수 있게 되었다. 또한 페이스북(Facebook)이나 플리커(Flickr)와 같은 소셜네트워크서비스를 사용하여 이용자가 직접 이미지, 동영상, 웹페이지 등에 키워드를 추가함으로써 타인과 정보 공유 및 검색이 더욱 용이하게 되었다. 이미지는 텍스트에 비해 내용을 직관적으로 전달하는 형태로서 가치와 의미를 지닌다. 이미지 정보 요구와 사용 비중이 확대되면서 누구나 손쉽게 이미지를 검색하는 것으로 생각할 수 있지만 이미지 이용자는 이미지를 찾기 위해 대부분 텍스트를 사용하며 검색 질의어를 생성하는데 어려움을 겪고 있다.

이미지는 형태적 속성, 내용적 속성, 의미적인 속성 등 다층적 속성을 가지며, 이용자의 정보요구에 따라 이미지의 복합적인 속성에 대한 검색 우선순위와 색인의 범주가 달라질 수 있다. 이용자 생성 콘텐츠가 급증하고 소셜미디어를 통한 공유 및 커뮤니케이션이 확대됨에 따라 이미지에 효과적으로 접근하기 위하여 시각적 객체를 기술하는 방법으로 태깅과 태깅 시스템의 필요성도 점차 증대하고 있다. 이미지 태깅을 생성하는 방식으로는 사람에 의해 이루어지는 수작업 태깅과 컴퓨터에 의해 자동으로 태깅을 생성하는 자동 태깅이 지속적으로 연구되

어 왔다. 최근에는 기계학습(machine learning)을 통해 이미지를 인식하는 인공지능 기술이 급속도로 발전하면서 사람이 이미지를 인식하는 과정과 유사한 방식으로 기계가 이미지를 인식하여 태그를 부여하고, 분류 및 검색하는 등 자동 방식의 태깅 기술이 빠르게 진화하고 있다. 이에 사람의 수작업에 의한 태깅과 기계에 의한 자동 태깅이 어떻게 다르며 향후 어떠한 방향으로 이미지 자료를 기술하고, 검색하도록 할 것인지에 대한 논의가 요구된다.

본 연구는 이미지에 대한 접근 및 검색을 향상시키기 위한 방법으로서 이미지 태깅을 효과적으로 생성할 수 있는 방안을 모색하기 위하여 동일한 이미지에 대해 사람이 생성한 휴먼 태그와 기계가 자동으로 생성한 머신 태그의 특징을 비교하고자 하였다. 본 연구의 목적을 상세히 기술하면 다음과 같다. 첫째, 이미지의 다층적 속성이 휴먼 태깅을 통해 어떻게 기술되는지 살펴보고자 분류 기준에 따라 휴먼 태그의 속성을 분석하고자 하였다. 둘째, 기계학습 기반으로 이미지를 인식하고 태그를 생성하는 머신 태깅 시스템을 선정하여 특징을 검토하고 머신 태깅의 결과로 생성된 머신 태그의 속성을 분석하고자 하였다. 셋째, 이미지 각각에 대한 머신 태그와 휴먼 태그의 빈도 및 속성을 비교함으로써 이미지 태깅에 요구되는 속성 기준 및 태깅 프로세스를 수립하고자 하였다. 넷째, 머신 태그와 휴먼 태그가 결합된 태그 데이터 집합을 대상으로 이용자의 검색 질의어로서 태그의 적합성을 살펴봄으로써 이미지의 검색 및 접근을 향상시키기 위한 이미지 태깅 시스템 모델을 제시하고자 하였다.

## 2. 이론적 배경 및 관련 연구

### 2.1 이미지 검색과 이미지 태깅

이미지는 사진, 지도, 그림, 포스터, 그래프, 차트, 일러스트, 클립아트 등 2차원으로 시각화된 정보를 의미한다(DCMI 2012). 이미지는 텍스트에 비해 내용을 직관적으로 전달하는 형태로서 가치와 의미를 지니며, 그 의미는 이미지 자체뿐만 아니라 이용자의 정보요구, 분석 시간이나 장소, 문화적·사회적 특징 등에 따라 달라질 수 있다. 이미지 검색은 내용 기반 검색과 의미 기반 검색으로 구분할 수 있다. 내용 기반 검색(CBIR: Content-Based Image Retrieval)은 이미지에서 시각적으로 보이는 것, 즉, 이미지 자체에서 추출되는 이미지의 색상, 형태 및 질감 등을 분석하여 검색하는 방법이다. 의미 기반 검색(CBIR: Concept-Based Image Retrieval)은 이미지가 무엇에 관한 것인가에 대한 기술어 기반(description-based), 또는 텍스트 기반(text-based) 이미지 검색을 말하며, 주제어나 키워드, 캡션, 자연어 텍스트 등을 통한 이미지 색인 및 검색의 개념을 가진다. 이미지 검색 행동 연구 결과를 보면, 이용자들은 이미지 검색을 위해 대부분 검색엔진을 사용하고 있고, 대체로 짧은 텍스트 키워드를 질의어로 이미지를 검색하지만 이미지 검색 실패율이 높은 것으로 나타났다(Yoon 2011; 정은경 2012). 이미지 검색과 일반적 웹 검색을 비교하면 이미지나 비디오를 찾는 이용자가 질의어 제안과 같은 시스템 지원을 더 많이 요구하였다(Spink and Jansen 2006). 이러한 연구들은 이미지 검색이 다른 유형의 검색보다 복잡하며, 시스템 보조기능을

요구한다는 특징을 보여주었다.

이미지 검색에 폭소노미를 적용하여 이용자에 의해 자유롭게 생성, 또는 선택된 주제 표목을 이용자 태그(user tag)라고 하며, 소셜 미디어 환경에서 폭소노미에 의한 색인 과정을 소셜 태깅(social tagging)이라 정의할 수 있다(Jørgensen, Stvilia, and Wu 2014). 시각적 자료에 대한 접근 가능성을 높이고 이미지에 대한 설명을 폭넓게 제공하는 태그의 가치가 평가되면서 플리커와 같은 이미지 공유 사이트를 비롯하여 도서관, 미술관, 박물관 등에서도 이용자 태깅의 유용성을 적극적으로 연구하고 시스템에 적용하였다(Smith 2011). 한편 이미지의 폭발적 증가에 따라 컴퓨터를 활용하여 이미지의 내용정보를 인식하고 이미지에 대한 어노테이션, 즉 태그를 자동으로 추출하는 자동 태깅(automatic tagging) 방식에 대한 연구도 지속적으로 증가해왔으나 태그의 품질이나 내용 면에서 만족스런 성능을 얻기는 쉽지 않았다(Wang et al. 2012). 최근 딥러닝(deep learning) 기반의 이미지 인식 방법인 CNN(Convolutional Neural Network)을 사용하여 이미지에 자동으로 태그를 붙임으로써 이미지 전체를 고려한 풍부한 내용의 태깅으로 발전해가고 있음을 보여주었다(장현웅, 조수선 2016; Ewerth et al. 2017).

본 연구에서는 이미지 태깅에 있어 사람(human)과 기계(machine)의 성능을 비교한 Ewerth et al.(2017)의 연구와 두 가지 태깅 방식의 결합을 제안한 Li, Snoek, and Worring (2009)의 연구와 같이 사람과 기계의 태깅을 비교해 보고자 하였다. 이를 위하여 사람에 의한 수동 방식의 태깅은 '휴먼 태깅(human tagging)'으로, 기계를 활용한 자동 방식의 태깅은 '머신

태깅(machine tagging)'으로 용어를 정의하고 두 가지 방식으로 생성되는 이미지 태그의 속성과 태깅의 특성이 어떠한 양상으로 나타나는지 살펴보았다.

## 2.2 이미지 태그의 속성

이미지 태그의 속성 범주를 구분하기 위해서는 이미지 질의어에 나타난 이미지 속성 분류 연구를 살펴볼 필요가 있다. Panofsky(1962)는 예술 이미지의 속성을 전도상학 기술, 도상학 분석, 도상학적 해석의 세 단계로 제시하였다. Shatford(1986)는 이용자 질의어에 나타난 정보 요구를 통해 이미지의 의미적 내용 기술을 특정적(specific), 일반적(generic), 추상적(abstract) 요구로 분류하고, 누가(who), 무엇을(what), 어디서(where), 언제(when)라는 4개의 패킷으로 이미지 속성 분류 프레임워크를 제시하였다. Jörgensen(1998)은 사람, 사물, 색상, 이야기, 지리적 장소, 설명, 시각적 요소, 미술사학 정보, 사람과 관련된 정보, 외부와의 관계, 감상자의 반응, 추상적인 정보 등 12개로 이미지의 속성을 구분하였다. Jörgensen et al.(2001)은 이미지의 시각적 내용에 대한 개념적 구조를 피라미드 형태로 표현하고 이미지 속성 범주를 10개의 수준으로 구분하였다. Klavans, LaPlante, and Golbeck(2014)은 Panofsky(1962)와 Shatford(1986)의 패킷 분류를 기반으로 Armitage와 Enser(1997)가 개발한 매트릭스를 사용하여 주제의 유형(누가, 무엇을, 어디서, 언제)과 특이성(일반적, 특정적, 추상적)의 두 차원으로 구성하였다. Golbeck, Koepfler, and Emmerling(2011)도 Panofsky(1962)와 Shatford(1986)

의 두 가지 프레임워크를 활용하여 12가지 유형에 '시각적 요소' 및 '알 수 없음' 분류를 추가하여 14가지 유형의 매트릭스를 제시하였다. Sigurbjörnsson과 Zwol(2008)은 태그 속성의 범주를 장소, 사물, 사람/그룹, 활동/사건, 시간 등으로 분류하였고, Jörgensen, Stvilia, and Wu(2014)는 이용자가 이미지에 어떤 종류의 용어를 태그를 할당하는지 구문론적 측면과 의미와의 관계를 연구하여 사물, 스토리, 예술역사, 사람 속성, 색상, 사람, 그룹, 기술, 위치, 추상, 시각적 요소, 외부관계, 이용자응답으로 속성을 분류하였다.

## 2.3 이미지 태깅 행태 분석

휴먼 태깅에 대한 연구는 이용자 태그, 소셜 태깅 등 이미지 태깅 행태를 중심으로 2000년대 후반부터 지속되어 왔다. 태깅 행태 분석 결과를 살펴보면, Hollink et al.(2004)의 연구에서는 추상적 용어보다 일반적 용어나 특정적 용어가 태그로 더 자주 사용되는 것으로 분석하였고, Huang(2006)은 플리커 인기 태그에 대한 문법적 분석을 통해 이미지를 설명하는 태그로 명사나 형용사가 주로 사용되며, 행동이나 과업을 설명하는 동사는 사용 빈도가 낮은 것을 보여주었다. Beaudoin(2007)은 플리커 사용자 14명이 생성한 140개의 이미지 태그를 18개의 범주로 분류하여 태깅의 패턴을 분석한 결과, 지리적 위치(예: 뉴욕, 중국)를 나타내는 장소 명칭이 가장 자주 사용되었으며, 두 개 이상의 용어를 결합한 복합 태그가 두 번째 자주 사용되는 범주로 분류되었다. 또한 단일 개념에 대해 가능한 많은 태그를 추가하거

나, 태그를 단수형과 복수형으로 작성하는 것도 태깅 패턴으로 주목하였다. 김현희와 김민경(2009)은 플리커 이미지의 태그를 분석한 결과, 이미지 태깅자의 사건/활동, 사람, 객체와 같은 논리적 태그의 비중이 높다는 것을 분석하고 폭소노미(folksonomy)와 내용 기반 이미지 검색 방법을 결합하여 이미지 검색의 효율성을 높이는 방안을 제안하였다.

Ransom과 Rafferty(2011)는 사람과 사물 속성이 태그로 가장 자주 사용된다고 분석하였고, Huang과 Jörgensen(2013)은 플리커 인기 태그에는 일반적 사물이나 특정 위치에 관련된 태그가 다수이며 특정 사물이나 시간 속성이 태그로 사용되는 빈도가 높다고 하였다. Jörgensen, Stvilia, and Wu(2014)는 태그의 의미론적 측면과 구문론적 측면을 밝히기 위해 플리커에 등록된 미국 의회도서관의 역사적 사진에 대한 참여자 태깅 행동을 살펴본 결과, 사물, 사람, 위치가 빈번한 태그로 분석되어 이전 연구들의 결과와 일치하는 경향을 보여주었다. 다수의 연구결과와 같이 이미지 태그의 속성은 일반적, 또는 특정적 속성이 높게 나타났으며, 공간/장소, 사건/활동과 관련된 태그가 이용자에 의해 다수 생성되는 것을 확인할 수 있었다.

#### 2.4 클라우드소싱과 태깅 시스템

태깅 시스템은 이미지 검색의 효율성을 높이는 방안으로 태깅 시스템 성능 개선, 또는 내용 기반 이미지 검색과 이용자 태깅을 결합하는 등 여러 방안이 제시되었다. 태깅 시스템 간의 차이가 태깅 결과 및 정보 생성 동기에 큰 영향을 주며(Marlow et al. 2006), 이미지 태그에

있어 상호작용의 효과의 이해를 위해 태그의 다양성, 태그 집합의 크기, 태그의 정확도와 일치도 등을 비교하는 연구(Bar-Ilan et al. 2010)가 수행되었다. 또한 사용자가 태그 생성 시 요구하는 태깅 시스템 인터페이스를 설계하는 방안이나 태깅에 클라우드소싱 방식을 적용한 시스템이 제안되었다(Nowak and Rürger 2010; Lin et al. 2015).

클라우드소싱(crowdsourcing)은 네트워크로 연결된 사람들이 공동 작업을 수행하는 형태로 문제를 해결하고 콘텐츠를 생산하는 개념의 모델이라 할 수 있다. 클라우드소싱 태그는 공동주석(collaborative annotation)의 의미로 소셜태그라고도 하며, 비디오, 이미지, 웹페이지와 같은 온라인 디지털 객체를 구성하고 색인할 때 인터넷 사용자가 수행하는 공동작업의 결과이다. 클라우드소싱을 적용한 연구로 Nowak과 Rürger(2010)는 아마존의 MTurk(Amazon Mechanical Turk)를 통해 수집된 클라우드소싱 태그가 비용을 감소시키며 전문가 수준의 태그만큼 신뢰할 수 있다고 보고 아마존 MTurk 클라우드 워커(worker)로부터 이미지 태그 수집을 실시하였다. 이때 이미지 설명 유무를 비교하여 이미지 검색 성능을 비교하였다. Lin et al.(2015)은 MTurk 플랫폼을 사용하여 이미지 설명이 사용자의 태깅에서 사용자에게 어떤 영향을 주는지 연구하였다. 태깅에 있어 이미지 설명의 영향에 대한 경험적 데이터를 수집 및 분석하여 태깅 프로세스를 이해하고, 이미지 태그의 수집에서 클라우드소싱의 향상 방안을 모색하였다.

머신 태깅은 기계학습을 활용한 시각 인식(visual recognition) 기술로 이미지를 인식하여

객체의 클래스를 분류하고 태그를 추출하는 시스템을 의미한다. 학습 데이터가 대량으로 축적되고 기계학습 기반의 이미지 인식 기술이 급격히 성장함에 따라 머신 태깅 시스템을 자체 개발하지 않고도 API(Application Programming Interface) 서비스로 공개된 이미지 인식 기술을 활용하여 머신 태깅을 수행할 수 있게 되었다. API는 응용프로그램에서 사용할 수 있도록 운영 체제나 프로그래밍 언어가 제공하는 기능을 제어할 수 있게 만든 인터페이스를 뜻하며, 대표적인 기계학습 기반의 머신 태깅 시스템으로 구글의 'Google Cloud Vision', 아마존의 'Amazon Rekognition', 아이비엠의 'IBM Visual Recognition' 등이 있다(〈표 1〉 참조).

### 3. 연구 설계

#### 3.1 연구 설계 개념도

본 연구는 이미지의 접근성 및 검색을 향상

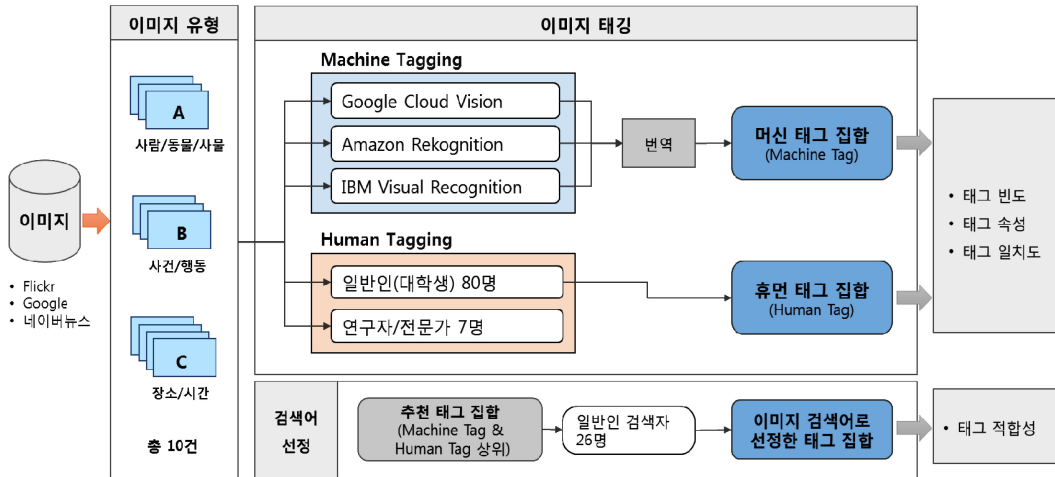
하기 위한 방안으로 휴먼 태깅과 머신 태깅의 두 가지 방식을 적용하여 이미지에 대한 휴먼 태그와 머신 태그를 수집 분석함으로써 공동작업의 개념을 실현하는 클라우드소싱 기반 '이미지 태깅 시스템(image tagging system)' 구축의 기초를 마련하는데 목적이 있다. 이러한 연구 목적을 달성하기 위하여 〈그림 1〉과 같이 연구 설계 방안을 수립하고 연구절차를 구조화하였다.

#### 3.2 데이터 수집 방법

본 연구를 위한 데이터 수집 절차는 다음과 같다. 첫째, 휴먼 태그 및 머신 태그를 생성하기 위한 이미지의 선정기준을 수립하였고, 이미지 공유 서비스인 플리커, 구글 이미지 검색, 네이버 뉴스를 통해 연구 대상 이미지 후보군을 수집하였다. 둘째, 휴먼 태그는 특정 분야 전문가나 색인 전문가가 아닌 일반인을 대상으로 하고자 대학생 80명을 태그 생성자로 참여하도록 하여 대상 이미지 10건의 휴먼 태그를 수집하

〈표 1〉 머신 태깅 시스템의 주요 기능

기능	Google Cloud Vision	Amazon Rekognition	IBM Visual Recognition
이미지 레이블 인식(객체, 장면)	√	√	√
이미지 속성 감지(Dominant Colors, Crop Hints)	√		
웹 감지(유사 이미지 검색)	√		
얼굴 인식/분석(감정, 성별, 연령, 안경 착용 등)	√	√	√
유명인사 인식	√	√	√
얼굴 비교		√	
로고 감지(제품 로고)	√		
랜드마크 감지(자연, 건축물)	√		
Safe 감지(폭력, 성인 이미지 감지)	√	√	
문서 텍스트 인식(광학문자인식)	√		
시각적 훈련(Image classifier)			√



〈그림 1〉 연구 설계도

였다. 수집된 휴먼 태그는 이미지별로 태그 속성을 코딩하였다. 셋째, 기계학습 기반의 이미지 인식 기술이 적용된 머신 태깅 시스템을 선정하고 대상 이미지 10건에 대한 머신 태그를 수집하였다. 이를 위해 이미지 인식 API를 제공하는 머신 태깅 시스템으로 Google Cloud Vision, Amazon Rekognition, IBM Visual Recognition 등 3종을 선정하였으며, 이미지 샘플에 대한 태깅 테스트를 실시한 후 연구 대상 이미지 10건에 대한 머신 태그를 수집하였다. 넷째, 수집된 휴먼 태그 및 머신 태그에 대한 속성과 출현 빈도, 출현빈도 상위 태그를 살펴보고, 머신 태그와 휴먼 태그의 일치도 및 속성을 비교하였다. 다섯째, 휴먼 태그와 머신 태그가 혼합된 태그 집합을 일반인 검색자 26명에게 제시하여 주고, 이미지 검색 질의어를 1순위부터 5순위까지 선택하게 하여 태그의 적합성을 살펴보았다. 여섯째, 군중이 공동으로 지식의 생성에 기여하는 클라우드소싱의 개념을 사람에 의한 태깅 뿐만 아니라 기계학습 기반의 머신 태깅 시스템으로

확장하여 사람과 기계의 협업 방식을 통해 이상적 태그 집합의 생성을 가능하도록 하는 태깅 시스템 모델을 제시하고자 하였다.

### 3.3 연구 대상 이미지 선정

이미지 태깅 분석을 위한 대상 이미지는 이미지 공유 서비스인 플리커, 구글 이미지, 네이버 뉴스를 통해 수집하였다. 플리커는 선행 연구에서 태깅 분석을 위해 자주 사용된 이미지 공유 사이트로 일반인이 특정 분야의 지식과 경험 없이도 친숙하고 쉽게 태그를 생성할 수 있으므로 이미지를 선정하는 대상 정보원으로 사용하였고 사회적 주제나 역사적 사건에 대한 이미지 선정을 위해서는 구글 이미지 검색과 네이버 뉴스 검색을 활용하였다. 이를 통해 1차 이미지 후보군을 선정하며 본 연구의 목적과 이미지 선정기준에 부합하는지 문헌정보학 전공자 2인의 검토를 거쳐 최종 10건의 이미지를 선정하였다.

또한 이미지의 유형별 특성을 살펴보기 위하여 Shatford(1986)의 Shatford 패킷 분류를 토대로 '누가(사람, 동물, 사물), 무엇을(사건, 행동, 상태, 감정), 어디서(위치, 장소), 언제(시간)'라는 4개 패킷을 포함하도록 10건의 이미지를 세 개의 그룹으로 구분하였다. A그룹은 객체(인물, 동물, 사물)를 포함한 이미지 3건, B그룹은 사건이나 행동(활동)이 나타나 있는 이미지 3건, C그룹은 장소와 시간이 포함되어 있는 이미지 4건을 각각의 유형으로 구분하였다(〈표 2〉 참조).

### 3.4 이미지 태그 수집 방법

이미지에 대한 휴먼 태그 생성을 위하여 2017년 10월 11일부터 13일까지 A대학교 대학생

80명이 참여하였다. 참여자 전공은 인문, 사회, 자연, 예체능 분야가 다양하게 구성될 수 있도록 하였다. 연구자가 참여자들에게 연구 목적을 설명하고 10건의 이미지와 함께 태그를 기입하는 조사지를 배포한 후 태그를 생성하도록 하였다. 머신 태그 생성을 위해서는 Google Cloud Vision과 Amazon Rekognition, IBM Visual Recognition 총 3종의 머신 태깅 시스템을 사용하여 휴먼 태깅과 동일한 이미지 10건에 대한 머신 태그를 수집하였다(2017년 10월). 3종의 머신 태깅 시스템에서는 이미지 인식을 통해 추출된 태그 목록과 함께 신뢰도 점수를 확인할 수 있었다. 신뢰도 점수는 0(신뢰 없음)에서 1(매우 높은 신뢰도)의 범위를 의미하며, 본 연구에서는 신뢰도 50% 이상의 머신 태그들을 분석에 사용하였다.

〈표 2〉 연구 대상 이미지(총 10건)

구분	이미지					
A그룹	1번		2번		3번	
	B그룹	4번		5번		6번
C그룹		7번		8번		
		9번		10번		

### 3.5 이미지 태그의 속성 분류

본 연구에서는 이미지 태그 속성 분류를 위하여 Shatford(1986)의 Shatford 패킷 분류를 기초로 Klavans, LaPlante, and Golbeck(2014)이 이미지 태그의 속성 분류에 사용한 Panofsky-Shatford 매트릭스를 기준으로 사용하였다. 이 매트릭스는 이미지의 의미적 내용을 특정적, 일반적, 추상적 범주로 분류하고, '누가, 무엇을, 어디서, 언제'라는 4개의 패킷으로 하여 총 12개의 범주로 구성되었다. 여기에 Golbeck, Koepfler, and Emmerling(2011)이 추가한 '시각적 요소'와 '알 수 없음' 항목을 포함하여 본 연구에서는 일반적(G), 특정적(S), 추상적(A), 시각적 요소(V), 알 수 없음(U)으로 이미지 태그의 속성 분류 기준을 설정하였다 (<표 3> 참조).

### 3.6 이상적 태그 집합의 구성

휴먼 태그와 머신 태그가 이미지의 어떠한 속성을 나타내고 있으며, 휴먼 태깅과 머신 태깅은 어떠한 특성을 가지고 있는지 분석하기 위하여 이미지의 다층적 속성을 포함하는 이상적 태그 집합을 구성하고자 하였다. 이상적인 태그 집합의 범주를 구성하기 위한 이론적 근거로는 Jörgensen et al.(2001)의 이미지 속성 분류 피라미드와 Shatford(1986)의 이미지 패킷 분류를 토대로 하였다. 이상적 태그 집합의 범주는 일반적, 특정적, 추상적, 시각적 요소의 4가지 범주를 가로축으로 하고, 시각적 인식 단계와 개념적 의미 단계를 세로축으로 구성하였다. 시각적 인식 단계는 이미지의 유형 및 기법, 전체적인 분포, 부분적인 구조, 전체적 구도를 세부 속성으로 구분하였다. 개념적 의미 단계는 객체와 장면에 대한 일반적 속성, 특정적 속

<표 3> 이미지 태그의 속성 분류 기준

속성 범주	설명 및 예시
일반적(G) General	사람, 사물, 동물의 유형 (예: 개, 나무, 할머니) 사건, 행동, 존재하는 상태 (예: 시위, 죽음, 게임) 위치의 유형 (예: 해변, 지하) 주기적인 시간 (예: 가을, 저녁)
특정적(S) Specific	명칭이 있는 사람, 객체 (예: 한국인, 은행나무) 특정 사건, 행동, 상태 (예: 세월호사건) 특정 위치 (예: 서울, 동작대교) 특정 기간, 날짜 (예: 11월 17일)
추상적(A) Abstract	신화적, 가공의 존재 (예: 유니콘, 인어) 감정, 추상화 (예: 행복, 신비, 쓸쓸함) 상징적 장소 (예: 천국) 상징화 된 시간 (예: 짧은 시절, 옛날)
시각적 요소(V) Visual Element	색상, 방향, 모양, 질감 등 시각적 구성 요소 (예: 검은색, 원형, 강한 빛놀이)
알 수 없음(U) Unknown	어떤 속성으로도 분류되지 않은 태그 (예: 오타자, 욕설, 의미를 알 수 없는 문장 등)

성, 추상적 속성의 세 단계로 구분하였다. 그중 일반적 속성과 특정적 속성은 각각 객체, 장면, 사건, 장소, 시간, 관계로 세부 속성을 구분하였고, 추상적 속성은 상징, 주제, 감정으로 하위 속성을 구분하였다. 객체는 사람, 사물, 동물, 식물 등 다양하므로 객체 특성에 따라 전체 및 부분으로 속성을 세분화하도록 하였다.

이에 따라 <표 4>와 같이 이상적 태그 집합의 속성을 범주화한 후, 3종의 머신 태깅 시스템에서 생성된 머신 태그와 대학생(80명)이 생성한 휴먼 태그를 세부 속성 범주에 따라 분류하였고, 추가적으로 연구자를 포함하여 디자인(2명), 인문학(2명), 시스템설계(2명) 분야 총 7명의 전문가로부터 휴먼 태그를 추가하여 이

미지 10건에 대한 이상적인 태그 집합을 구성하였다.

## 4. 이미지 태깅 결과 분석

### 4.1 머신 태그 분석 결과

이미지 10건에 대하여 Google Cloud Vision, Amazon Rekognition, IBM Visual Recognition 등 3종의 태깅 시스템에서 생성된 머신 태그는 총 385개였다. 태그 수를 시스템 별로 비교해보면, Amazon Rekognition이 154건(40.0%), Google Cloud Vision이 141건(36.6%)이었으며,

<표 4> 이상적 태그 집합의 범주

단계	속성 구분	속성 정의	
시각적 인식 단계 Perceptual level	Type/Technique	사진, 일러스트, 그림, 흑백사진 등 생산된 이미지의 양식 및 기법	
	Global Distribution	색상, 질감, 명암, 채도 등 이미지의 전체적인 시각적 분포	
	Local Structure	부분적인 세부사항의 시각적 요소(점, 선, 그림자, 밀도, 모양 등)	
	Global Composition	전체적 구조 및 구도, 형태 등(평행, 대칭, 원형, 사선 등)	
개념적 의미 단계 Conceptual level	일반적 General	Object	전체적인 이미지로 설명될 수 있는 객체의 일반적인 속성
		Scene	전체적인 장면에 대한 일반적인 속성
		Event	객체의 활동, 행동 및 사건에 대한 일반적인 속성
		Place	객체나 장면에 대한 일반적인 장소
		Time	객체나 장면에 대한 일반적인 시간
		Relation	객체들 간, 또는 객체와 장면의 관계에 대한 일반적인 속성
	특정적 Specific	Object	객체에 대한 구체적 지식을 통해 식별되거나 명명, 설명되는 속성
		Scene	장면에 대한 구체적 지식을 통해 식별되거나 명명, 설명되는 속성
		Event	객체의 활동, 행동 및 사건에 대한 구체적, 특정적 속성
		Place	객체나 장면에 대한 구체적, 특정적 장소
		Time	객체나 장면에 대한 구체적, 특정적 시간
		Relation	객체와 객체 간, 또는 객체와 장면의 관계에 대한 특정적 속성
	추상적 Abstract	Symbol	객체 및 장면, 행동, 사건, 장소, 시간에서 표현되는 것에 대한 상징적 의미
		Subject	객체 및 장면, 행동, 사건, 장소, 시간에서 표현되는 것에 대한 내재적, 추상적 개념
		Emotion	객체 및 장면, 행동, 사건, 장소, 시간에서 표현되는 것에 개인적 감정 및 정서

IBM Visual Recognition은 이보다 훨씬 적은 90건(23.4%)이었다. 태그수로만 보면 Amazon Rekognition이 가장 많은 태그를 생성하였다. 머신 태그를 이미지별로 살펴보면, 9번 이미지에 가장 많은 태그가 생성되었다. 머신 태그 전체 총 385개의 출현 빈도는 일반적 속성(85.7%)이 월등하게 많았으며, 특정적 속성은 6.8%, 시각적 요소는 6.8% 추상적 속성은 0.5%에 그쳤다. 중복 제거 고유 태그(unique tag)는 총 266개로 출현 빈도는 일반적 속성(84.6%), 특정적 속성(9.4%), 시각적 요소(5.3%) 순으로 전체 태그의 속성과 유사한 비중을 보였다(〈표 5〉 참조).

세 개의 그룹으로 나누어 태그의 속성을 비

교한 결과, 장소/시간에 대한 C그룹 이미지 유형의 경우에 생성된 머신 태그가 211개로 가장 많았으며, 그중 일반적 속성은 90.0%에 달하였다. A그룹과 B그룹은 C그룹에 비해 시각적 요소에 대한 머신 태그의 비중이 높았으며, 세 그룹에서 공통적으로 추상적 속성의 머신 태그는 거의 나타나지 않았다(〈표 6〉 참조).

#### 4.2 휴먼 태그 분석 결과

총 80명의 참여자로부터 수집한 이미지 10건에 대한 휴먼 태그 수는 〈표 7〉과 같다. 생성된 태그 수는 총 7,725개였고, 중복 제외 고유 태그(unique tag)는 총 2,120개, 이미지 1건당







〈표 5〉 머신 태그의 속성

속성	전체 태그		고유 태그	
	빈도(개)	비중(%)	빈도(개)	비중(%)
일반적(G)	330	85.7	225	84.6
특정적(S)	26	6.8	25	9.4
추상적(A)	2	0.5	1	0.4
시각적 요소(V)	26	6.8	14	5.3
알 수 없음(U)	1	0.3	1	0.4
합계	385	100.0	266	100.0

〈표 6〉 이미지 그룹별 머신 태그의 속성

이미지 유형	머신 태그 속성					전체(개)
	일반적	특정적	추상적	시각요소	알수없음	
그룹A(사람/동물)	77	10	0	12	1	100
	77.0%	10.0%	-	12.0%	1.0%	
그룹B(사진/행동)	63	2	0	9	0	74
	85.1%	2.7%	-	12.2%	-	
그룹C(장소/시간)	190	14	2	5	0	211
	90.0%	6.6%	1.2%	2.4%	-	
전체	330	26	2	26	1	385
	85.7%	6.8%	0.5%	6.8%	0.3%	

〈표 7〉 이미지 번호별 휴먼 태그 수

이미지 번호	1번	2번	3번	4번	5번	6번	7번	8번	9번	10번	계
											
총 태그 수 (개)	798	648	742	797	792	761	756	778	742	911	7,725
고유 태그 수 (개)	227	210	176	226	194	235	192	203	203	254	2,120
이미지 1건당 평균 태그수 (개)	9.98	8.10	9.28	9.96	9.90	9.51	9.45	9.73	9.28	11.39	9.66

평균 태그수는 9.66개로 분석되었다. 고유 태그 수는 3번 이미지가 176개로 가장 적었고, 10번 이미지가 254개로 가장 많이 수집되었다. 이미지 1건 당 평균 이용자의 태그 수 역시 10번 이미지가 11.39개로 가장 많았다(〈표 7〉 참조). 휴먼 태그의 출현빈도를 보면, 고유 태그 2,120개 중 1,195개(56.4%)가 1회 출현한 것으로 나타났다. 2회 출현한 태그는 337개(15.9%), 3회 출현한 태그는 151개(7%)였다. 이는 Choi와 Syn(2016) 연구에서 1,540개의 고유 태그 중 823개인 53.44%가 한 번만 사용된 태그로 나타난 결과와 비슷한 경향을 보였다.

대상 이미지 10건에 대한 휴먼 태그(고유 태그) 총 2,120개의 속성을 살펴보면, 일반적 속성(64.7%), 특정적 속성(14.8%), 추상적 속성(13.8%), 알 수 없음(3.8%), 시각적 요소(3.0%)의 순으로 나타났다(〈표 8〉 참조). 미술관의 이미지에 적용된 태그에 대하여 주제별 분류를 살펴본 Klavans, LaPlante, and Golbeck(2014)의 연구 결과에서는 태그의 66%가 일반적 속성으로 분류되었고, 추상적 속성은 12%, 특정적 속성은 6%로 분류되었다. 대학생이 생성한 태그를 분석한 본 연구에서도 일반적 속성이 다수를 차지한 것은 이미지에 대한 전문적인 지식이 없는 일반인이 생성한 태그 집합에서 나타나는 양상과 비슷한 경향을 보였다. 총 2,120개의

휴먼 태그 중 이미지별로 출현빈도가 높은 상위 10개 태그(총 100개) 속성의 분석 결과는 일반적 속성(56.0%), 특정적 속성(29.0%), 추상적 속성(11.0%), 시각적 요소(4.0%) 순으로 나타났다. 일반적 속성은 출현빈도 상위 태그에서 비중이 더 낮았으며, 특정적 속성은 전체 태그(14.8%)와 비교하여 상위 태그가 29.0%로 높았다. 태그를 생성하는 휴먼 태거(human tagger)가 식별 가능하며, 명칭을 알고 있는 사람이나 사물, 사건이나 장소에 대해서는 특정적 속성의 태그가 다수 나타났으며, 감정과 정서, 추상적 개념의 주제 등도 다양한 태그로 표현되었다. 이는 이미지에 대한 이용자 태그에서 사건/활동, 사람, 객체와 같은 태그의 비중이 높다고 분석한 선행연구(김현희, 김민경 2009; Jörgensen, Stvilia, and Wu 2014)와도 유사한 경향을 보였다.

이미지 그룹별로 휴먼 태그의 속성을 살펴본 결과, 세 그룹 모두 일반적 속성의 비중이 높게 나타났으며 그중 C그룹이 가장 높았다. A그룹과 B그룹에서는 추상적 속성이 19% 내외로 나타나 C그룹에 비해 정서나 추상 개념, 감정에 대한 태그가 많았으며, B그룹의 경우 A그룹 및 C그룹과 비교해 볼 때 일반적 속성의 비중은 상대적으로 낮고 특정적 속성의 비중이 높은 것으로 분석되었다(〈표 9〉 참조).

〈표 8〉 휴먼 태그의 속성

속성	전체 태그		출현빈도 상위 태그	
	빈도(개)	비중(%)	빈도(개)	비중(%)
일반적(G)	1,371	64.7	56	56.0
특정적(S)	313	14.8	29	29.0
추상적(A)	292	13.8	11	11.0
시각적 요소(V)	64	3.0	4	4.0
알 수 없음(U)	80	3.8	0	0
합계	2,120	100.0	100	100.0

〈표 9〉 이미지 그룹별 휴먼 태그의 속성











이미지 유형	휴먼 태그 속성					전체(개)
	일반적	특정적	추상적	시각요소	알수없음	
A그룹(사람/동물)	368	80	118	30	17	613
	60.0%	13.1%	19.2%	4.9%	2.8%	
B그룹(사건/행동)	375	138	121	7	14	655
	57.3%	21.1%	18.5%	1.1%	2.1%	
C그룹(장소/시간)	628	95	53	27	49	852
	73.7%	11.2%	6.2%	3.2%	5.8%	
전체	1,371	313	292	64	80	2,120
	64.7%	14.8%	13.8%	3.0%	3.8%	

### 4.3 머신 태그와 휴먼 태그의 속성 비교

이미지별로 휴먼 태그와 머신 태그의 속성을 비교한 결과는 〈표 10〉과 같다. 머신 태그의 속성 비중을 보면, 이미지 10건 모두 일반적 속성이 80% 이상으로 높게 나타났고, 2번(골든 리트리버)과 9번(가을) 이미지에서 특정적 속성이 다른 이미지에 비해 상대적으로 높게 나타났다. 2번(골든 리트리버)의 경우, 특정적 속성이 32.1%로 가장 높았는데, 동물 중 개의 '종'에 해당하는 명칭인 '골든 리트리버'가 3종의 시스템에서 공통적인 태그로 생성되었다. 9번(가을) 이미지에 대한 태그도 특정적 속성이 17.6%로 상대적으로 높았는데, 특히 식물에 대

한 특정 명칭이 태그로 다수 생성되었다. 1번(스티브잡스), 3번(할머니), 5번(노란 리본), 6번(5.18민주화운동) 등 4건의 이미지는 시각적 요소가 10% 이상으로 나타난 것이 특이적이었는데, 이러한 이미지들의 공통점은 흑백 사진이거나 특정 색상이 주를 차지하는 경우였다. 휴먼 태그의 속성 비중을 보면, 머신 태그와 같이 10건의 이미지에서 일반적 속성이 가장 많았으나, 머신 태그의 일반적 속성이 80% 이상인 것과 비교해 보면 낮은 비중을 보였다. 그중 1번(스티브잡스) 이미지를 비롯하여(39.2%), 5번(노란 리본) 46.4%, 4번(촛불집회) 51.8%로 일반적 속성이 52% 미만이었고, 특정적 속성과 추상적 속성이 각각 25% 내외로 나타났다.

〈표 10〉 이미지 번호별 머신 태그와 휴먼 태그의 속성

이미지	머신 태그 속성						휴먼 태그 속성					
	일반적	특정적	추상적	시각요소	알수없음	전체(개)	일반적	특정적	추상적	시각요소	알수없음	전체(개)
1번 	21	0	0	5	0	26	89	56	65	17	0	227
	80.8%	-	-	19.2%	-	-	39.2%	24.7%	28.6%	7.5%	-	-
2번 	17	9	0	2	0	28	155	10	25	10	10	210
	60.7%	32.1%	-	7.1%	-	-	73.8%	4.8%	11.9%	4.8%	4.8%	-
3번 	39	1	0	5	1	46	124	14	28	3	7	176
	84.8%	2.2%	-	10.9%	2.2%	-	70.5%	8.0%	15.9%	1.7%	4.0%	-
4번 	20	0	0	2	0	22	117	55	50	3	1	226
	90.9%	-	-	9.1%	-	-	51.8%	24.3%	22.1%	1.3%	0.4%	-
5번 	20	2	0	3	0	25	90	51	45	2	6	194
	80.0%	8.0%	-	12.0%	-	-	46.4%	26.3%	23.2%	1.0%	3.1%	-
6번 	23	0	0	4	0	27	168	32	26	2	7	235
	85.2%	-	-	14.8%	-	-	71.5%	13.6%	11.1%	0.9%	3.0%	-
7번 	46	1	1	1	0	49	138	14	22	9	9	192
	93.9%	2.0%	2.0%	2.0%	-	-	71.9%	7.3%	11.5%	4.7%	4.7%	-
8번 	49	1	1	2	0	53	132	48	7	9	7	203
	92.5%	1.9%	1.9%	3.8%	-	-	65.0%	23.6%	3.4%	4.4%	3.4%	-
9번 	55	12	0	1	0	68	155	16	15	8	9	203
	80.9%	17.6%	-	1.5%	-	-	76.4%	7.9%	7.4%	3.9%	4.4%	-
10번 	40	0	0	1	0	41	203	17	9	1	24	254
	97.6%	-	-	2.4%	-	-	79.9%	6.7%	3.5%	0.4%	9.4%	-
전체	330	26	2	26	1	385	1371	313	292	64	80	2,120
	85.7%	6.8%	0.5%	6.8%	0.3%	-	64.7%	14.8%	13.8%	3.0%	3.8%	-

4.4 이미지 검색어로서 태그의 적합성

휴먼 태그와 머신 태그가 혼합된 이미지 태그 집합에 대하여 일반인 검색자로부터 이미지 검색 질의어로 적합한 태그를 1순위부터 5순위 까지 선택하도록 하여 이미지 검색어로서 태그의 적합성을 살펴본 결과, 각 이미지별로 우선 순위가 높은 태그(상위 5개)는 〈표 11〉과 같다.

A그룹(인물/동물/사물: 1, 2, 3번), B그룹(사건/활동: 4, 5, 6번), C그룹(위치/시간: 7, 8, 9, 10번)으로 구분하여 검색어로 선정되는 순위가 높은 태그를 분석한 결과, A그룹은 휴먼 태그가 머신 태그에 비해 검색어로 선택되는 비중이 높게 나타났으며, 머신 태깅을 통해서 동물의 특정한 종과 명칭이 태그로 생성되

고 이용자들이 이를 검색 태그로 선택하게 되는 것을 알 수 있었다. B그룹은 대한민국의 사회적 사건 및 활동을 다룬 이미지들로 검색어로 선택된 태그 대부분은 머신 태그보다는 휴먼 태그에서 다수 선택되었다. 반면 C그룹은 일반적 속성의 머신 태그 다수가 검색어로 선정되었으며, B그룹과 비교하여 해당 이미지에 포함된 객체 및 시간, 장소에 대하여 일반적, 특정적 속성의 태그가 생성되고, 이용자에 의해 검색어로 선정되었다. 이는 Chung과 Yoon(2010)이 웹검색 엔진 로그 데이터에 대한 이미지 검색 질의어에서 살펴본 것과 같이 특정어와 지칭어가 다양한 정보 탐색 단계에서 지속적으로 나타나는 경향과 유사한 결과로 분석할 수 있었다.

〈표 11〉 이미지 검색어로 선택된 태그(상위 5개)

 1번	 2번	 3번	 4번	 5번
스티브잡스 (HT)	골든 리트리버 (HT/MT)	할머니 (HT)	촛불집회 (HT)	노란 리본 (HT)
애플 (HT)	리트리버 (HT/MT)	6.25전쟁 (HT)	촛불시위 (HT)	세월호 (HT)
아이폰 (HT)	개 (HT/MT)	옛날사진 (HT)	광화문 (HT)	잊지않겠습니다 (HT)
잡스 (HT)	대형견 (HT)	1950-60년대 (HT)	민주주의 (HT)	단원고 (HT)
스티브잡스 패션 (HT)	강아지 (HT)	흑백사진 (HT/MT)	탄핵 (HT)	4.16 (HT)
 6번	 7번	 8번	 9번	 10번
5.18민주화운동 (HT)	무인도 (HT)	남산타워 (HT)	가을 (HT/MT)	수능 (HT)
군사정권 (HT)	외딴섬 (HT)	서울야경 (HT)	단풍 (HT/MT)	고3 (HT)
광주민주화운동 (HT)	바닷가 (HT/MT)	야경 (HT)	낙엽 (HT/MT)	수험생 (HT)
5.18 (HT)	모래사장 (HT)	한강 (HT)	산책로 (HT/MT)	대학입시 (HT)
탄압 (HT)	나룻배 (HT)	한강 다리 (HT)	단풍나무 (HT/MT)	모의고사 (HT)

HT(휴먼 태그), MT(머신 태그)

#### 4.5 머신 태깅과 휴먼 태깅의 특성 비교

머신 태깅 시스템 3종에서 생성된 머신 태그와 대학생에 의해 생성된 휴먼 태그를 비교하여 각각의 이미지별로 태그와 그 속성을 비교해 본 결과, 머신 태그의 경우 모든 이미지에서 일반적 속성이 대다수를 차지하였고, 추상적 속성은 거의 드물었다. 1번, 3번, 5번, 6번 이미지는 시각적 요소에 대한 속성이 타 이미지에 비해서 상대적으로 높았으며, 이는 흑백이나 특정 색상이 강조되는 이미지에서 나타나는 머신 태그의 특징으로 분석되었다. 이미지 10건에 대한 태그를 A그룹(사람/동물), B그룹(사건/행동), C그룹(장소/시간)으로 나누어 속성을 분석한 결과, 머신 태그는 일반적 속성이 현저하게 높았고 특정적 속성은 낮은 비중으로 나타났으며, 추상적 속성은 거의 나타나지 않았다. 반면, 휴먼 태그는 일반적 속성이 다른 속성에 비해 상대적으로 높기는 하였으나 A그룹과 B그룹은 추상적 속성도 20% 내외로 나타났

다. 출현빈도가 높은 휴먼 태그의 경우 특정적 속성의 비중이 더욱 커지는 것으로 분석되었다.

이미지 태그 집합의 특징을 종합적으로 비교하여 머신 태깅과 휴먼 태깅의 강점을 속성별로 살펴보고자 '이상적 태그 집합 범주(〈표 4〉 참조)'에 따라 이미지에 대한 모든 태그를 분류하였다. 각각의 이미지에는 검색어 우선순위가 가장 높았던 태그를 이미지의 제목으로 부여하였다.

인물 및 동물에 대한 1번(스티브잡스)와 2번(골든 리트리버) 이미지의 경우를 보면, 머신 태그에는 일반적 속성과 시각적 요소가 많이 나타났고, 휴먼 태그에는 사람의 이름이나 사물의 구체적 명칭이 다수 나타났으며, 혁신이나 웃음, 외로움이나 기다림 등 추상적, 감정적 태그도 많았다. 3번(할머니)은 흑백 사진인 점을 고려할 때 휴먼 태거는 정확한 시간 및 장소를 알 수 없는 막연한 과거 시점으로 인식하여 일반적 장소와 시간에 대한 태그를 비롯하여 사람의 신분 및 관계(예: 할머니, 어린이, 가족), 사람의 신체

(예: 흰머리, 주름살), 의상(예: 한복, 저고리) 등 전체적, 또는 부분적으로 나타나는 객체에 대해 상세하게 태그를 생성하는 경향을 보였다.

사건/행동을 포함하는 이미지의 태그를 보면, 머신 태그와 휴먼 태그의 특성이 부각되었다. 4번(촛불집회)의 머신 태그에는 일반적 속성과 시각적 요소가 주로 나타났고, 휴먼 태그에는 촛불집회가 나타나게 된 사회적 맥락과 사건 관련 특정 인물과 용어를 비롯하여 해당 이미지에서 느껴지는 감정·정서(예: 따뜻함, 한마음), 추상 개념(예: 민주주의, 평화시위) 주제가 포함되었다. 촛불을 들고 있는 사람을 나타내는 태그는 ‘군중, 시민, 국민’ 등 비슷한 의미를 가진 다양한 용어로 표현되었다. 5번(노란 리본)의 머신 태그는 색깔(예: yellow color, lemon yellow color)이나 사물 명칭을 태그(예: flower, fruit, banana)로 생성하였으나, 적합하지 않은 태그가 다수 제시되었다. 반면, 휴먼 태그는 이미지에 내포된 사회적 맥락과 사건의 연관관계를 설명하는 태그(예: 세월호, 단원고, 4.16)가 많았고, 사회적 공동체 및 개인으로서 느끼는 감정(예: 슬픔, 미안함)에 대한 추상적 의미의 태그도 다수 나타났다. 6번(5.18민주화운동)은 ‘폭력, 폭행, 억압, 군인, 탄압’ 등 일반적 속성의 휴먼 태그가 다수 나타났으며 ‘5.18 광주민주화운동’을 기록한 사진으로 특정 사건의 장면임을 사전 인지하고 있는 휴먼 태거는 해당 사건과 관련한 배경, 인물, 장소를 토대로 특정적 속성의 태그들을 생성하였다. 한편, 3번과 6번과 같은 흑백사진이나 5번과 같이 주된 컬러가 광범위하게 나타나는 이미지의 경우, 머신 태그와 휴먼 태그에서 색상에 대한 태그(예: 흑백, 노란색)가 공통적인 특징으로

분석되었다.

장소를 나타내는 이미지에 대한 태그를 보면, 장소명이나 위치를 특정할 수 없는 이미지(7번 무인도)의 경우 휴먼 태그와 머신 태그에서 이미지에 포함된 객체의 명칭에 대해 다수의 태그가 생성되었다. 8번(남산타워)과 같이 이미지 상의 장소나 위치가 식별되는 휴먼 태그에는 특정 지역이나 장소의 명칭(예: 한강, 남산타워, 한강철교)과 일반적 용어(예: 도시, 다리, 강, 하늘)가 함께 나타난 반면, 머신 태그에는 ‘도시, 풍경, 하늘, 다리, 밤, 저녁’ 등 일반적 속성의 태그가 대부분을 차지하였다. 9번(가을)의 경우 머신 태그에서 다양한 식물(나무)의 일반적, 특정적 명칭을 태그로 제시하여 태그 수는 많았으나 이미지 상의 객체에 대한 정확성 여부를 판단하기 어렵다는 이슈가 있었다. 10번(수능)에서 머신 태그는 ‘학생, 시험, 학교, 교실’ 등 장면과 객체를 인식하여 일반적 속성의 태그가 주를 이루었으며, 휴먼 태그는 태그 생성자의 경험을 통한 다양한 단어가 일반적 속성의 태그(예: 수능, 모의고사, 시험, 고등학생)로 표현되었다.

이러한 결과를 통해 머신 태깅과 휴먼 태깅의 주요 특성을 몇 가지로 정리하면 다음과 같다.

첫째, 머신 태깅은 사람 및 사물을 나타내는 일반적 속성과 시각적 요소의 태깅에 강점을 가진 반면 사회문화적 배경을 반영한 특정적 속성이나 연관 주제나 감정, 정서 등 추상적 속성의 태그 생성은 미흡하였다. 휴먼 태깅으로는 이미지 내에 포함된 인물, 사건, 행동 등에 대한 특정적 속성을 가진 단어와 복합어, 유사어, 연관어 등을 다양하게 생성할 수 있었다.

둘째, 휴먼 태깅은 휴먼 태그 생성자의 지식

과 경험 내에서 식별되는 인물, 사진, 행동 등에 대해서 특정적 태그를 정확하게 생성하는데 강점을 나타냈다. 사회적 배경이 포함된 사건/행동을 포함하는 이미지의 경우, 머신 태깅은 시각적 요소와 같은 이미지 내용 기반의 태그를 생성하였고 그 신뢰도 점수도 높았다. 또한 일반적 속성의 태그도 다수 생성하였으나 이미지 상의 객체와 불일치하는 문제가 여전히 존재하였다. 반면, 휴먼 태깅은 객체의 전체와 부분에 대한 일반적, 특정적 태그와 함께 의미 기반의 태그를 다양하게 생성하였다. 즉, 이미지에 내포된 사회적 맥락과 연관 관계, 이미지가 표출하는 상징과 주제, 이미지에서 느껴지는 감정과 정서 등이 다양한 용어로 생성되는 특성에 강점을 가지고 있음을 보여주었다.

셋째, 머신 태깅은 사람이나 동물에 대한 이미지에 대해서는 공통적으로 'people, person, human, animal, mammal' 등을 가장 신뢰도 높은 태그로 추출하였고, 식물이나 자연 풍경과 같은 이미지에 대해서는 'nature, tree, plant' 등 생물을 나누는 상위 분류군에 대한 명칭을 태그로 추출하는 특성이 공통적으로 나타났다. Wang et al.(2012)이 분석한 다수의 연구에서 보듯이 기계학습 방법을 적용한 모델들은 대부분 ImageNet이나 PASCAL VOC 데이터세트 및 플리커, 인스타그램 등 소셜 태깅 데이터를 학습데이터로 활용하였다. 이를 통해 학습에 사용한 데이터가 어떤 것인지에 따라 태그 데이터의 수집 및 분류가 달라짐을 알 수 있었다.

본 연구의 결과로 나타난 휴먼 태깅과 머신 태깅의 특성은 이미지에 대한 정확하고 풍부한 태깅을 위하여 사람의 지적능력을 바탕으로 이미지-태그 데이터세트를 축적해나가고 지속적

인 기계학습의 실행으로 머신 태깅 시스템의 성능을 개선해가는 것이 필요함을 시사하였다.

## 5. 이미지 태깅 시스템 모델 제안

본 연구에서는 이미지에 대한 태그를 생성하기 위하여 클라우드소싱을 활용하는 접근법을 제안하고자 하였다. 최근 딥 러닝을 사용한 CNN 이미지 인식 기술이 이미지의 특정 객체뿐만 아니라 배경을 포함한 이미지 전체를 고려한 머신 태깅을 가능하게 하여 검색의 정확도가 증가하는 것을 보여준 연구(장현웅, 조수선 2016)와 같이 머신 태깅은 방대한 양의 이미지 데이터세트, 또는 소셜 태그 데이터를 사용한 기계학습을 통해 인간의 뇌에서 일어나는 의사결정과정을 모방하며 이미지 인식 품질을 빠르게 향상시켜 나가고 있다.

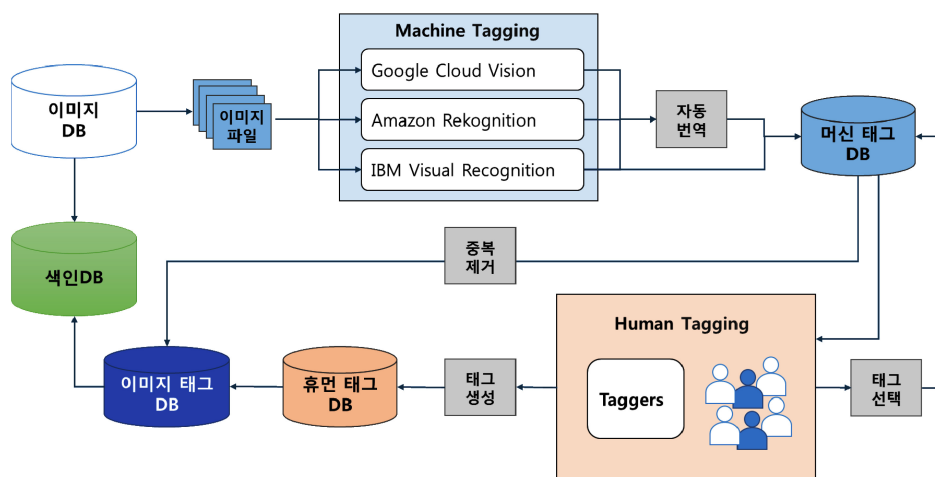
클라우드소싱은 네트워크로 연결된 사람들이 공동 작업을 수행하여 문제를 해결하고 정보를 생산하는 개념의 모델을 말한다. 본 연구에서는 태그 생성자의 범위를 군중에 의해 태그를 생성하는 휴먼 태거(human tagger)와 인공지능 기술이 적용된 머신 태거(machine tagger)가 협업하여 공동 태깅을 수행하는 것으로 개념화 하였다. 휴먼 태거는 일반인과 전문가를 포함하는 다수의 사람이며, 머신 태거는 기계학습 기반의 이미지 태깅 시스템이다 (Google Cloud Vision과 Amazon Rekognition, IBM Visual Recognition 등). 인간의 인지능력과 컴퓨터의 처리능력을 결합하는 협업방식 태깅 시스템을 제안한 Wang et al.(2012)은 수동 방식과 자동 방식의 장점을 결합하여 '데

이터의 수집, 태그 추천, 태그 처리'라는 세 가지 방식으로 태깅 시스템을 설명하였다. 본 연구에 적용되는 방식은 머신 태그를 생성하여 기초데이터를 구성하고, 휴먼 태그를 통해 태그 추천 및 추가정보를 보완하는 협업 태깅 모델이다. 사람과 기계의 공동 작업을 토대로 휴먼 태깅과 머신 태깅을 결합하는 상호보완적 접근을 통해 이미지 태그를 정교화하고 품질을 향상시키며, 이미지의 다층적 기술과 다면적 해석 정보를 축적해가는 클라우드소싱 기반 이미지 태깅 시스템 모델을 구성하였다(〈그림 2〉 참조).

첫째, 머신 태깅을 통해 이미지에 대한 시각적 요소 및 일반적 속성에 대한 태그 집합을 우선 구성할 수 있을 것이다. 기계에 의한 내용 기반 접근은 이미지 자체에서 추출되는 이미지의 색상, 형태 및 질감 등을 분석하여 유사 이미지 영역을 식별함으로써 이미지 속성의 상당한 부분을 처리하고 검색할 수 있게 한다. 본 연구에서는 이미지에 동물이나 건축물 등이 포함되

어 있는 경우, 일반적 속성과 구체적 명칭에 대한 머신 태그가 상당히 정확한 수준으로 생성되는 결과를 보였다. 인공지능의 발전에 따라 Wang et al.(2012)이 제시한 컴퓨터 보조 태깅 방식인 태그 추천과 태그 처리 등에 있어 휴먼 태깅은 태그의 신규 생성보다는 머신 태깅 결과물에 대한 적합성을 판단하는 역할의 비중을 확대하게 될 것이다.

둘째, 휴먼 태깅을 통해서서는 이미지 내에 포함된 인물, 사건, 행동 등에 대한 특정적 속성과 이미지와 관련된 연관 주제어, 감정, 정서 등 추상적 속성에 대한 태그가 생성될 수 있다. 특히 우리나라의 역사적, 사회적 사건이나 장소, 건축물 이미지의 경우 머신 태깅을 통해 적합한 태그 생성이 현재로는 매우 미약하다. 따라서 이러한 이미지와 관련하여 휴먼 태깅을 통해 문화적, 사회적 맥락을 포함하는 데이터를 생성하고 추상적 감정이나 정서, 해석적 수준의 용어들도 휴먼 태깅을 통해 이미지-태그 데이터셋으로 구축해 가는 것이 필요할 것이다.



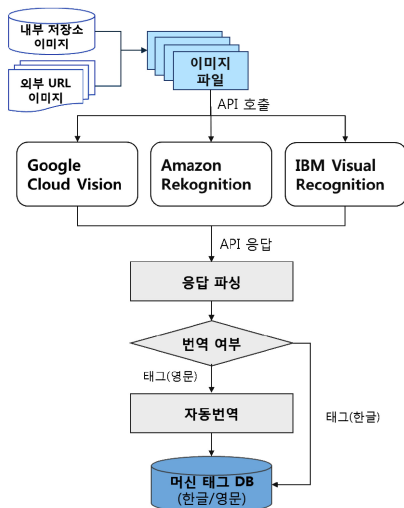
〈그림 2〉 이미지 태깅 시스템 구성도

셋째, 휴먼 태깅과 머신 태깅 방식의 강점을 상호보완적으로 사용하는 이미지 태깅 시스템을 통해 이미지의 복합적 속성을 포함할 수 있는 이상적인 태그 집합을 구성할 수 있을 것이다. 이러한 태깅 프로세스를 통해 생성된 태그와 이미지의 매칭 데이터는 머신러닝 알고리즘을 개선하기 위한 학습 데이터셋으로 활용함으로써 머신 태깅을 통해 생성 가능한 이미지 태그의 범위를 더욱 확장해가고 머신 태그의 정확성과 신뢰도를 높이는데 기여할 수 있을 것이다.

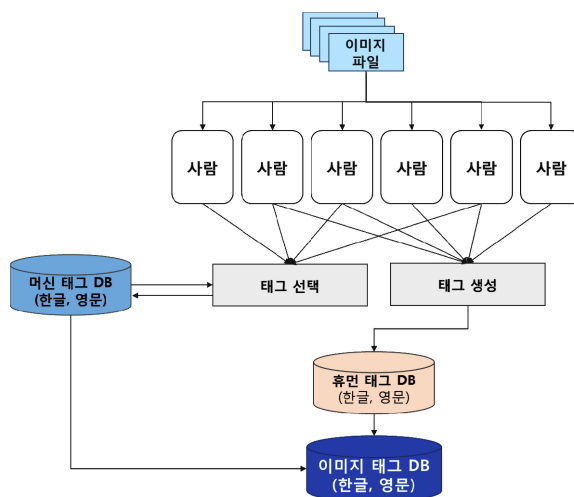
머신 태그 생성 과정은 <그림 3>과 같이 먼저 태깅을 생성할 이미지 파일을 내부저장소 및 외부 URL로부터 가져온다. 머신 태깅 시스템으로 사용할 대상을 선택하여 API를 호출하게 되는데 여러 이미지의 동시 호출이 가능하나 1회 또는 1일 호출하는 이미지 개수가 제한된다. 본 연구에서는 Google Cloud Vision, Amazon Rekognition, IBM Visual Recognition 등 3개

시스템을 활용하였으나, API를 제공하는 형태의 타 시스템을 추가하여 머신 태깅을 수행할 수 있다. 다음으로 API 호출 결과에 따라 태그 목록과 신뢰도 점수를 응답결과로 받게 된다. 신뢰도는 0(신뢰 없음)에서 1(매우 높은 신뢰도)의 범위를 의미하므로 휴먼 태깅 시 참고를 위하여 태그와 함께 제시하도록 한다. 머신 태깅 시스템에서는 응답결과를 대부분 영문으로 제공하고 일부 한글이 제공되므로 번역여부를 판단하여 한글 태그는 즉시 머신 태그 DB에 저장하며, 영문 태그는 자동번역 과정을 거쳐 생성된 한글 태그와 쌍으로 머신 태그 DB에 저장하게 된다.

다음으로 휴먼 태그 생성 과정은 <그림 4>와 같이 다수의 사람이 네트워크에 연결되어 태깅을 하게 되며, 태깅 태스크는 머신 태그의 적합성을 판단하는 태스크와 추가 태그를 생성하는 태스크가 제시된다. 휴먼 태그 생성은 두 가지 프로세스로 진행된다. 첫째, 제시된 이미지에



<그림 3> 머신 태그 생성 과정



<그림 4> 휴먼 태그 생성 과정

대하여 휴먼 태거가 머신 태그 리스트 중에서 이미지와의 내용 적합성을 판단하여 태그를 선택하도록 한다. 둘째, 제시된 이미지에 대하여 기존 태그 외에 신규 태그를 추가 생성하는 단계를 거친다. 이러한 과정을 통해 수집된 태그는 휴먼 태그 DB에 저장된다. 하나의 이미지에 대한 휴먼 태그 및 머신 태그는 태그 데이터셋으로 구성되어 최종적으로 이미지 태그 DB에 저장되며 색인 및 검색 과정이 진행된다.

## 6. 결론 및 제언

본 연구는 이미지 자료에 대한 접근 및 검색을 향상시키기 위한 방법으로서 이미지 태그를 효과적으로 생성할 수 있는 방안을 검토하고자 휴먼 태그와 머신 태그의 속성과 태그의 특성을 비교 분석하였다. 이를 토대로 이미지의 다층적 수준과 세부 속성을 포함하는 태그를 효과적으로 생성하기 위한 클라우드소싱 기반 이미지 태깅 시스템 모델을 제시하였다.

본 연구를 통한 결과를 요약하면 다음과 같다. 첫째, 수집된 이미지 태그를 분석한 결과, 10건의 이미지에 대하여 생성된 태그 수는 머신 태그가 총 385개(고유 태그 266개)였으며, 휴먼 태그는 총 7,725개(고유 태그 2,120개)이었다. 이미지 한 건당 머신 태깅으로는 평균 12.8개의 태그가 생성되었으며, 휴먼 태깅으로는 평균 9.66개의 태그가 생성되었다. 태그의 속성을 분석한 결과, 머신 태그는 일반적 속성이 가장 높았으며, 특정적 속성과 시각적 요소는 일부 나타났고, 추상적 속성은 거의 나타나지 않았다. 휴먼 태그는 일반적 속성이 가장 높

았으나 특정적 속성과 추상적 속성도 포함되었다. 둘째, 휴먼 태그 집합 중 출현 빈도가 높은 태그의 경우, 태그 전체와 비교할 때 일반적 속성의 비중은 상대적으로 낮고, 특정적 속성의 비중은 높은 것으로 나타났다. 한편, 휴먼 태그 생성자가 명칭을 알 수 있는 객체 및 장면에 대해서는 특정적 속성의 비중이 높았으며, 감정과 정서, 추상적 개념뿐만 아니라 사건, 장소, 시간, 관계 등이 다양한 태그로 표현되었다. 셋째, 태그 집합 중 일반인 검색자에 의해 검색어로 선정된 태그의 분석 결과, 인지되는 특정 인물이나 사물, 사건, 장소 등을 포함하고 있는 이미지는 검색어 역시 특정적 속성을 가진 태그의 순위가 높았다. 반면 막연한 인물이나 장소의 경우에는 이미지에 포함된 객체나 장면에 대하여 일반적 속성을 가진 태그의 순위가 높게 나타났다. 이는 이미지에서 인물이나 위치를 식별할 수 있다면 첫 번째 태그로 할당하는 경향을 보였던 Golbeck, Koepfler, and Emmerling (2011)의 결과와 이미지 검색질의에서 특정어와 지칭어 범주가 가장 많이 나타났던 Chung과 Yoon(2010)의 결과와도 유사한 경향을 나타냈다.

휴먼 태그와 머신 태그 속성에 대한 차이를 살펴보면, 휴먼 태그는 특정 사람이나 사건, 또는 인지되는 장소가 있는 경우 일반적 속성보다는 특정적 속성의 비중이 상대적으로 높게 나타났다. 사건 이미지의 경우 이미지에 나타난 객체나 장면에 대한 태그 외에도 사건을 둘러싼 사회적, 정서적 맥락과 관련된 휴먼 태그들이 다수 생성되었다. 머신 태그는 대체적으로 일반적 속성과 시각적 요소에 대한 태그의 비중이 높았는데, 동물, 식물, 건축물 등이 포함

된 이미지에 대해서는 객체의 명칭이 태그로 제시되었다. 또한 흑백 사진이나 특정 색상의 영역이 강하게 지각되는 이미지의 경우, 시각적 요소 중 색상에 대한 태그 수가 높게 나타났다. 이용자의 정보요구를 맥락적으로 이해하고자 한 정은경과 정선영(2012)의 이미지 속성 분석 결과를 보면, 보편성 속성이 가장 높은 비율로 나타났고, 특정적 속성과 추상적 속성도 유사한 비율로 나타난 것과 비교하여 구성적 요소는 비율이 낮게 나타났다. 이용자의 이미지 정보요구를 충족하기 위해서 머신 태그가 강점을 가지는 구성 측면의 접근점 외에 의미 측면과 비시각 측면의 접근점을 풍부하게 제공하는 휴먼 태그의 확장을 통해 다양한 접근점 제공이 필요하겠다.

본 연구를 통해 살펴본 바와 같이 휴먼 태깅과 머신 태깅은 각각의 특성을 지니고 있으며, 태그 속성 분석을 통해 얻게 된 결과는 이미지 태그 집합을 구성하기 위한 태그의 속성 분류 기준을 수립하며, 효과적인 태깅 프로세스 및 시스템을 구축하는데 기초자료로 활용될 수 있을 것이다. 본 연구는 머신 태깅과 휴먼 태깅을 비교 분석하여 태그 속성의 차이와 강약점을 살펴봄으로써 이상적인 이미지 태그 집합을 구성하는 태깅 프로세스 및 시스템 모델을 제시하고자 하였는데 의의가 있다. 기계학습을 기반으로 이미지 인식 및 객체 감지를 통해 머신 태깅을 하는 시스템은 인공지능 알고리즘의 특성과 학습 데이터의 변화에 따라 머신 태깅 결과가 달라질 수 있다. 본 연구를 통해 수립된 태그 속성 기준 및 태그 집합은 머신러닝 알고리즘을 개선하기 위한 트레이닝 데이터 세트를 구성하는데 기초 자료로 활용될 수 있을 것이다.

본 연구에서 수행한 연구를 더욱 발전시키기 위한 후속 연구로 첫째, 이미지 유형에 따른 태그의 속성 기준을 적용한 클라우드 컴퓨팅 기반의 이미지 태깅 플랫폼 구축에 대한 상세 연구가 필요하겠다. 머신 태깅 시스템에 대해서는 지속적인 테스트와 태그 검증을 통해 신뢰도가 낮은 부분을 개선해 나가야 하며, 방대하고 정확한 태그와 이미지를 데이터셋으로 구축해야 한다. 둘째, 사람만이 가능한 휴먼 태깅에 대한 연구 축적이 필요하다. 감정과 정서, 연관 관계, 지식과 경험을 토대로 하는 해석적 이미지 기술과 사회문화적 맥락에 대한 태깅은 사람에 의해 가능한 부분이다. 휴먼 태깅을 위해서는 이미지의 속성 범주 및 태깅 기준을 정의하고 그에 따른 태깅이 가능하도록 프로세스를 구축하는 것이 필요하며, 태깅이 용이한 프로세스 설계와 직관적인 태깅 인터페이스 제공, 작업에 따른 보상 등 시스템 개발 및 정책적 측면의 연구가 필요할 것이다. 셋째, 다양한 유형의 대량 이미지에 대하여 휴먼 태깅 참여자를 확대하고 이미지 자원을 다양화하는 방안 연구가 필요하겠다. 본 연구는 대학생으로부터 태그를 수집하였으나, 다층적 속성의 태그 생성을 위해서는 연령이나 직업을 다양화하고 특정 분야 전문가의 휴먼 태깅을 수집하는 등 태깅 참여자를 확대하는 것이 필요할 것이다. 또한 현재 머신 태깅 시스템은 대부분 영문으로 태그를 제공하고 있으므로 태그 번역의 제한점을 개선해 나가야 하며, 한글 태그의 생성을 확대하기 위해 공공 부문에서 보유하고 있는 우리나라 역사, 사회, 문화에 대한 대량의 이미지를 활용한 한글 태그 데이터셋 구축 연구가 필요할 것이다.

## 참 고 문 헌

- 김현희, 김민경. 2009. 플리커 이미지 자료에 대한 이용자 태깅 행태 분석과 활용 방안. 『정보관리연구』, 40(2): 71-94.
- 장현웅, 조수선. 2016. CNN을 이용한 소셜 이미지 자동 태깅. 『정보과학회논문지』, 43(1): 47-53.
- 정은경, 정선영. 2012. 일상생활 맥락 정보요구 기반의 이미지 접근점 확장에 관한 연구. 『정보관리학회지』, 29(4): 273-294.
- 정은경. 2012. 멀티미디어 정보요구와 검색행태에 관한 탐색적 연구. 『한국문헌정보학회지』, 46(3): 251-270.
- Armitage, L. H. and P. G. B. Enser. 1997. "Analysis of user need in image archives." *Journal of Information Science*, 23(4): 287-299.
- Bar-Ilan, J., M. Zhitomirsky-Geffet, Y. Miller, and S. Shoham. 2010. "The Effects of Background Information and Social Interaction on Image Tagging." *Journal of the American Society for Information Science and Technology*, 61(5): 940-951.
- Beaudoin, J. 2007. "Folksonomies: Flickr image tagging: Patterns made visible." *Bulletin of the American Society for Information Science and Technology*, 34(1): 26-29.
- Choi, Y. and S. Y. Syn. 2016. "Characteristics of Tagging Behavior in Digitized Humanities Online Collections." *Journal of the American Society for Information Science and Technology*, 67(5): 1089-1104.
- Chung, E. and J. Yoon. 2010. "Examining Categorical Transition and Query Reformulation Patterns in Image Search Process." *Journal of the Korean Society for Information Management*, 27(2): 37-60.
- Dublin Core Metadata Initiative(DCMI). 2012. DCMI Type Vocabulary - DCMI Metadata Terms [online]. [cited 2017.9.12]. <<http://dublincore.org/documents/dcmi-type-vocabulary>>.
- Ewerth, R., M. Springstein, L. A. Phan-Vogtmann, and J. Schütze. 2017. "Are Machines Better Than Humans in Image Tagging? - A User Study Adds to the Puzzle." *Advances in Information Retrieval, ECIR 2017, LNCS*, 10193: 186-198.
- Golbeck, J., J. Koepfler, and B. Emmerling. 2011. "An experimental study of social tagging behavior and image content." *Journal of the American Society for Information Science and Technology*, 62(9): 1750-1760.
- Hollink, L., A. Schreiber, B. J. Wielinga, and M. Worring. 2004. "Classification of user image descriptions." *International Journal of Human-Computer Studies*, 61(5): 601-626.

- Huang, H. 2006. "Tag distribution analysis using the power law to evaluate social tagging systems: A case study in the Flickr database." *17th ASIS&T SIG/CR Classification Research Workshop*, 14-15.
- Huang, H. and C. Jørgensen. 2013. "Characterizing user tagging and co-occurring metadata in general and specialized metadata collections." *Journal of the American Society for Information Science and Technology*, 64(9): 1878-1889.
- Jørgensen, C. 1998. "Attributes of images in describing tasks." *Information Processing and Management*, 34(2/3): 161-174.
- Jørgensen, C., A. Jaimes, A. B. Benitez, and S.-F. Chang. 2001. "A conceptual framework and empirical research for classifying visual descriptors." *Journal of the American Society for Information Science and Technology*, 52: 938-947.
- Jørgensen, C., B. Stvilia, and S. Wu. 2014. "Assessing the relationships among tag syntax, semantics, and perceived usefulness." *Journal of the American Society for Information Science and Technology*, 65(4): 836-849.
- Klavans, J. L., R. LaPlante, and J. Golbeck. 2014. "Subject matter categorization of tags applied to digital images from art museums." *Journal of the Association for Information Science and Technology*, 65(1): 3-12.
- Li, X., C. G. M. Snoek, and M. Worring. 2009. "Annotating images by harnessing worldwide user-tagged photos." *IEEE International Conference on Acoustics, Speech, and Signal Processing 2009*.
- Lin, Y., C. Trattner, P. Brusilovsky, and D. He. 2015. "The impact of image descriptions on user tagging behavior: A study of the nature and functionality of crowdsourced tags." *Journal of the American Society for Information Science and Technology*, 66(9): 1785-1798.
- Marlow, C., M. Naaman, D. Boyd, and M. Davis. 2006. "HT06, tagging paper, taxonomy, Flickr, academic article, to read." *Proceedings of the seventeenth conference on Hypertext and hypermedia (HYPERTEXT '06)*, 31-40.
- Nowak, S. and S. Rüger. 2010. "How reliable are annotations via crowdsourcing: a study about inter-annotator agreement for multi-label image annotation." *Proceedings of the international conference on Multimedia information retrieval (MIR '10)*, 557-566.
- Panofsky, E. 1962. "Chapter I: Introductory. In Studies in iconology." *Humanistic themes in the art of the Renaissance*, 3-31.
- Ransom, N. and P. Rafferty. 2011. "Facets of user-assigned tags and their effectiveness in

- image retrieval.” *Journal of Documentation*, 67(6): 1038-1066.
- Shatford, S. 1986. “Analyzing the subject of a picture: A theoretical approach.” *Cataloging & Classification Quarterly*, 6(3): 39-62.
- Sigurbjörnsson, B. and R. V. Zwol. 2008. “Flickr tag recommendation based on collective knowledge.” *Proceedings of the 17th International Conference on World Wide Web (WWW '08)*, 327-336.
- Smith, M. K. 2011. “Viewer tagging in art museums: Comparisons to concepts and vocabularies of art museum visitors.” *Advances in Classification Research Online*, 17(1): 1-19.
- Spink, A. and B. J. Jansen. 2006. “Searching multimedia federated content web collections.” *Online Journal Review*, 30(5): 485-495.
- Wang, M., B. Ni, X. Hua, and T. Chua. 2012. “Assistive tagging: A survey of multimedia tagging with human-computer joint exploration.” *ACM Computing Surveys*, 44(4).
- Yoon, J. 2011. “Searching images in daily life.” *Library & Information Science Research*, 33: 269-275.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- Chung, EunKyung 2012. “An Exploratory Investigation on Multimedia Information Needs and Searching Behavior among College Students.” *Journal of the Korean Society for Library and Information Science*, 46(3): 251-270.
- Chung, EunKyung and SunYoung Chung. 2012. “An Approach Toward Image Access Points Based on Image Needs in Context of Everyday Life.” *Journal of the Korean society for information management*, 29(4): 273-294.
- Jang, Hyunwoong and Soosun Cho. 2016. “Automatic Tagging for Social Images using Convolution Neural Networks.” *Journal of KIISE*, 43(1): 47-53.
- Kim, Hyun-Hee and Min-Kyung Kim. 2009. “Investigating the End-User Tagging Behavior and its Implications in Flickr.” *Journal of information management*, 40(2): 71-94.