
저자 (Authors)	정다혜, 김세빈, 정초롱, 김장원 Jeong Da-Hye, Kim Se-Bin, Jeong Cho-Rong, Gim Jang-Won
출처 (Source)	Proceedings of KIIT Conference , 2018.11, 302-304(3 pages)
발행처 (Publisher)	한국정보기술학회 Korean Institute of Information Technology
URL	http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07577957
APA Style	정다혜, 김세빈, 정초롱, 김장원 (2018). 인스타그램에서 해시태그와 이미지를 활용한 감정 분석 방법. Proceedings of KIIT Conference, 302-304
이용정보 (Accessed)	가천대학교 203.249.***.201 2019/09/29 19:00 (KST)

저작권 안내

DBpia에서 제공되는 모든 저작물의 저작권은 원저작자에게 있으며, 누리미디어는 각 저작물의 내용을 보증하거나 책임을 지지 않습니다. 그리고 DBpia에서 제공되는 저작물은 DBpia와 구독계약을 체결한 기관소속 이용자 혹은 해당 저작물의 개별 구매자가 비영리적으로만 이용할 수 있습니다. 그러므로 이에 위반하여 DBpia에서 제공되는 저작물을 복제, 전송 등의 방법으로 무단 이용하는 경우 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

Copyright Information

Copyright of all literary works provided by DBpia belongs to the copyright holder(s) and Nurimedia does not guarantee contents of the literary work or assume responsibility for the same. In addition, the literary works provided by DBpia may only be used by the users affiliated to the institutions which executed a subscription agreement with DBpia or the individual purchasers of the literary work(s) for non-commercial purposes. Therefore, any person who illegally uses the literary works provided by DBpia by means of reproduction or transmission shall assume civil and criminal responsibility according to applicable laws and regulations.

인스타그램에서 해시태그와 이미지를 활용한 감정 분석 방법

정다혜(*), 김세빈(*), 정초롱(*), 김장원(*)[†]

(*) 군산대학교 소프트웨어융합공학과, {gwg03391, bhy401, jeongcr01603, jwgim}@kunsan.ac.kr

An Emotion Analysis Method Using Hash Tags and Images in Instagram

Jeong Da-Hye(*), Kim Se-Bin(*), Jeong Cho-Rong(*), and Gim Jang-Won(*)[†]

(*) Kunsan national University, Department of Software convergence engineering

요약

최근 스마트폰 보급률이 높아짐에 따라 SNS는 생활 속 많은 부분에 영향을 미친다. 많은 사람들은 자신의 의견을 표현하거나 감정을 표출하는데 이러한 SNS를 활용한다. 따라서 SNS에서 사용자들의 감정분석 연구가 활발하게 진행 중이다. 기존에는 인스타그램의 해시태그를 기반으로 감정분석을 하였지만 하나의 게시물에 한 개 이상의 해시태그가 부여되기 때문에 감정을 나타내는 해시태그가 다수가 포함될 경우 정확한 감정을 분류하기 어렵다. 이를 해결하기 위해 본 연구에서는 해시태그와 게시물의 이미지를 활용한 감정분석 방법을 제안한다. 게시물의 대표 감정을 나타내는 유사감정용어를 도출하고 이미지에서 표현되는 감정의 분석을 통해 다중해시태그가 포함된 게시물의 대표 감정을 분류할 수 있음을 보였다.

1. 서론

최근 소셜 네트워크 서비스(Social Network Service, SNS)가 스마트폰 이용자의 증가와 무선 인터넷 서비스의 확장으로 인해 생활 속 일부분으로 자리 잡았다. 이러한 시대 속에서 다양한 SNS 서비스들 중 인스타그램(Instagram)은 사진 공유 및 소셜 네트워킹 서비스를 통해 다른 사용자와 피드를 주고받을 수 있으며 자신의 의견을 표현하거나 감정을 표출하는 데 활용한다[1, 2]. 최근 SNS(페이스북(facebook), 트위터(Twitter), 인스타그램(instagram))에 등록된 게시물에 대한 감정을 분석하는 연구가 활발히 진행되고 있다. 그 중 SNS 게시물

단어가 자살징후가 있는지 아닌지 진단하는 연구는 SNS 감정분석이 얼마나 중요한지를 보여준다[3]. 그렇지만 인스타그램에서는 사용자의 게시물에 한 개 이상의 해시태그가 포함되고 그 중에 여러 개의 감정을 표현하는 해시태그들이 있다. 그러므로 한 개의 게시물의 대표적인 감정을 분류함에 있어 해시태그만으로는 어려움이 발생한다. 따라서 본 연구에서는 인스타그램 게시물에 포함된 해시태그와 이미지를 활용한 감정분석을 제안한다.

2. 관련 연구

SNS에서 사용자들이 게시한 게시물 등을 이용하여 감정을 분석하는 많은 연구들이 진행되어 왔다. 기존에는 SNS에서 키워드를 추출한 후 감정사전을 구축하여 긍정 및 부정을 의미하는 극성 키워드들을 정의해 분

[†] 교신저자

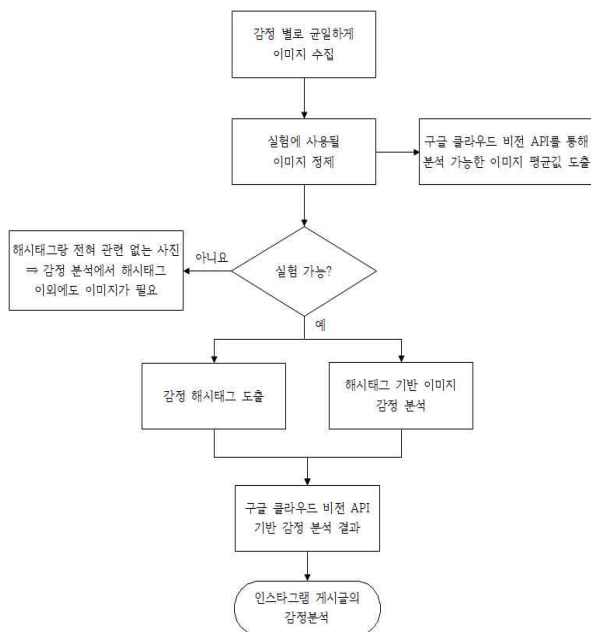
이 성과는 2018년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임
(No. NRF-2018R1C1B6008624).

류[4]한다. 그러나 SNS를 자주 사용하는 젊은 세대의 경우 이제 축약어로 대표되는 인터넷 은어를 트위터, 페이스북 등의 SNS 상에서 많이 사용되고 이모티콘, 철자 오류와 같은 특징을 고려하지 않아 좋은 성능을 보이지 못하고 있다. 그러므로 해시태그를 활용하여 분석을 하는 연구들이 제안되었다. 인스타그램의 해시태그를 이용해서 감정 분류 방법을 제안하고 기쁨, 화남, 지루, 평온의 감정을 나타내는 카테고리에 대한 평균 정확률이 90.7%의 성능을 보이고 있다[5]. 그렇지만 인스타그램에서 사용자들은 한 개의 게시물에 한 개 이상의 해시태그를 부여하기 때문에 다중 감정이 혼합되어 표현되는 경우가 발생한다. 따라서 본 연구에서는 인스타그램에서 해시태그와 함께 게시된 이미지를 기반으로 감정 분석 방법을 제안한다.

3. 실험 방법

그림 1은 본 실험 방법의 전체적인 순서도를 보인다. 먼저 감정 해시태그를 도출해 실험에 사용될 이미지를 정제하고 실험가능 유무에 따라 다르게 분석한다.

(그림 1) 감정 분석 제안모델



(Figure 1) Proposal model of emotional analysis

3.1 감정 해시태그 도출 방법

3.1은 사람이 가질 수 있는 대표 4가지 감정인 행복(Happy), 화남(Angry), 평온(peaceful), 지루(Bored)에 대한 감정 분석을 위해 인스타그램에서 대표 감정을 나타내는 게시글 빈도수가 높은 상위 유사 감정 해시태그를 도출하였고 이는 <표 1>에서 보인다.

<표 1> 대표 감정 및 감정을 나타내는 유사 해시태그

대표 감정	유사 감정 해시태그
행복	#행복, #행복해, #행복하자, #행복스타그램, #신남, #개신남, #감동
화남	#화남, #짜증, #짜증나, #개짜증, #긴장
평온	#평온, #차분한, #차분, #여유, #여유로움
지루	#지루, #지루해, #슬픔, #졸림

<Table 1> Representative emotion and Similar hash tag to represent emotion

3.2 해시태그 기반 이미지 분석

실험을 위해 인스타그램에서 각각의 대표 감정마다 게시글을 200건씩 수집하여 총 800건을 대상으로 실험을 실시하였다. 그리고 각 해시태그와의 연관성을 기준으로 실험에 사용가능한 사진들을 정제하였다. 이미지로부터 감정분석을 위해 구글 클라우드 비전 API (Google Cloud Vision API)를 사용하여 이미지에 포함된 객체 및 얼굴을 감지해 감정을 분석할 수 있다. 따라서 다중 감정이 포함된 게시글의 감정 분석을 위해 구글 클라우드 비전 API를 이용해 게시글의 이미지의 감정을 분석한다. 구글 클라우드 비전 API 기반의 감정 분석 효율성을 평가하기 위해 정확률(Precision)을 이용하였으며 <수식 1>와 같이 계산된다. 정확률은 실제로 대표 감정을 나타내는 정답셋 데이터 중 게시글의 해시태그와 구글 클라우드 비전 API결과가 일치하는 이미지의 수로 정확률을 판단하였다.

<수식 1> 정확률 수식

$$\text{정확률} = \frac{\text{감정이 올바르게 분류된 이미지 수}}{\text{정답셋 데이터 수}}$$

<formula 1> Precision

4. 실험 결과

인스타그램에서 수집한 게시글을 해시태그를 이용하여 대표 감정별로 분류하고 분류된 게시글들에 대하여 이미지를 통한 감정 분석을 수행하였다. 실험 데이터 집합에서 대표 감정별로 도출된 이미지 개수는 <표 2>와 같다. 많은 인스타그램 사용자가 앞서 분류했던 해시태그와 관련없는 이미지를 게시하고 있었다. 수치에서 확인할 수 있듯이, 사용자가 실제로 자신의 게시글에 #행복이라고 해시태그를 부여했어도 실제로 행복한 건지 의미없는 해시태그인지 알 수 없다. 그럴 경우 그 사용자의 다른 게시물들의 이미지를 통해서 감정분석의 지표가 될 수 있다.

<표 2> 데이터집합에서 대표 감정별로 도출된 이미지 개수

구분	개수	감정키워드가 도출된 개수	퍼센트율(%)
행복	200	32	16
화남	200	43	21.5
평온	200	45	22.5
지루	200	23	11.5
계	800	143	17.875

<Table 2> Number of images derived by representative emotion in data set

구글 클라우드 비전 API를 활용하여 이미지를 감정 분석할 경우 6가지 감정키워드(anger, joy, surprise, Exposed, Blurred, Sorrow)로 분류될 수 있다. 구글 클라우드 비전 API기반의 이미지 감정분석 정확률(Precision)을 <표 3>에서 보인다.

<표 3> 구글 클라우드 비전 API 기반의 이미지 감정분석 정확률(Precision)

구분	정답셋개수	해시태그와 이미지감정분석이 일치하는 이미지 개수	퍼센트율(%)
행복	23	14	60.87
화남	16	0	0
평온	21	14	66.67
지루	28	6	21.43
계	88	34	38.64

<Table 3> Precision of Image emotion analysis based Google cloud vision API

구글 클라우드 비전 API 기반의 이미지 감정 분석 정확률이 저조한 이유는 인스타그램 게시글 인물의 표정이 해시태그에 맞게 뚜렷하지 않은 경우 다른 감정 키워드가 도출되기 때문이다. 그렇지만 다양한 감정 해시태그가 부여되었을 때의 문제점을 이미지가 대표적 인 감정을 이끌어낼 수 있는 해결방안이 된다는 점을 확인할 수 있었다.

5. 결 론

본 연구에서는 인스타그램 게시글의 감정분석에 있어서 해시태그를 기반으로 이미지 활용이 가능할지 실험해보았다. 게시글이 다중 해시 태그일 경우 해시태그만으로 대표 감정을 도출할 수 없다는 문제가 발생한다. 이를 이미지 감정분석을 통해서 해결할 수 있음을 보인다. 향후에는 대표 감정이 나타나지 않을 경우 다양한 대체 감정키워드를 도출하는 연구로 이어져야 할 것이다. 또한 더 많은 표본을 확보하여 표정이 아닌 다른 특징을 활용하여 이미지 감정 분석의 정확도를 높일 수 있는 방법을 연구해 볼 예정이다.

참고 문헌

- [1] 위키백과, “인스타그램”, 2018.11.18. , <https://ko.wikipedia.org/wiki/인스타그램>
- [2] 권순모, 이종혁, “감정분석에서의 문맥적 정보로써 인스타그램 과거 게시글의 해시태그 활용 가능성 제안”, 한국정보과학회 학술발표논문집 논문집, pp. 649-651, June 2018.
- [3] 정영재, 서원준, 변예림, 강태원, “단순 베이지 분류기를 이용한 SNS 자살 징후 분류”, 한국정보과학회 학술발표논문집, pp.2129-2131, June 2015.
- [4] 오평화, “트위터의 감정 분석을 통한 실시간 장소 추천 시스템”, 한국전자거래학회지, Vol.21, no.3, pp.15-28, February 2016.
- [5] 남민지, 이은지, 신주현, “SNS 해시태그를 이용한 사용자 감정 분류 방법에 관한 연구”, 멀티미디어 학회 논문지, Vol.18, no.11, pp.1391-1399, August 2015.