

# Social Multimedia Sentiment Analysis

Jiebo Luo  
Department of Computer Science  
University of Rochester  
Rochester, New York 14627  
jluo@cs.rochester.edu

Damian Borth  
German Research Center for Artificial  
Intelligence  
Kaiserslautern, Germany 67663  
damian.borth@dfki.de

Quanzeng You  
Department of Computer Science  
University of Rochester  
Rochester, New York 14627  
qyou@cs.rochester.edu

## ABSTRACT

Social multimedia refers to the multimedia content (text, images, and videos) generated by social network users for social interactions. The increasing popularity of online social networks leads to a significant amount of multimedia content generated by online social network users. Researchers from both the industrial and academic have been working on a broad range of projects related to the analyzing and understanding the online multimedia content, including real world activity prediction and content recommendation. Particularly, understanding online users' opinions or sentiments is a fundamental task that can benefit many applications, such as political campaigning and commercial marketing. We present a few recent advances in social multimedia sentiment analysis. Specifically, this tutorial consists of three parts. The first part is on visual sentiment analysis. We will introduce the task of visual sentiment, its main challenges, and the state-of-the-art approaches. We will include several representative approaches to manually designing visual features for this task as well as some approaches using deep neural networks. The second part is on building multimedia sentiment analysis datasets. We will introduce the challenges, the solutions in the construction of different large-scale datasets for sentiment analysis. The final part is mainly on multimodality model for sentiment analysis. We will introduce some recent research projects on multimodality designing and learning. In addition, we will also share some applications of sentiment analysis, as well as thoughts on current challenges and future directions.

## CCS CONCEPTS

• **Information systems** → **Multimedia information systems; Information systems applications**; • **Computing methodologies** → **Computer vision**; • **General and reference** → *Measurement*;

## KEYWORDS

social multimedia, sentiment analysis, multimodality

### ACM Reference format:

Jiebo Luo, Damian Borth, and Quanzeng You. 2017. Social Multimedia Sentiment Analysis. In *Proceedings of MM '17, Mountain View, CA, USA, October 23-27, 2017*, 2 pages.  
<https://doi.org/10.1145/3123266.3130143>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

MM '17, October 23-27, 2017, Mountain View, CA, USA

© 2017 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-4906-2/17/10.

<https://doi.org/10.1145/3123266.3130143>

## 1 INTRODUCTION

Today, we live a life in online social networks. Every minute, there is an enormous amount of multimedia data being generated, shared and propagated across different online social network platforms. Online social networks have become one of the most important parts of our life. They have reshaped the way of communication and the acquirement of information. Consequently, they offer a new approach and platform for people to easily express their opinions by posting different multimedia content related to a topic. These enormous, unprecedented opinions could be obtained by collecting and analyzing related user generated multimedia content, which can provide signals for many other applications, such as predicting box-office revenues for movies, political elections, and economic indicators.

In this tutorial, we pay attention to multimedia sentiment analysis [15]. There are many studies on textual sentiment analysis. However, sentiment analysis on visual and audio content has been less studied. Compared with text, multimedia content, especially image and videos offer some unique advantages which consequently make them the most pervasive media formats on the Internet: they are more attractive and salient than plain text. Thus, they can grab users' attention instantly and carry much more information that can be comprehended more quickly - just like an old saying, "a picture is worth a thousand of words". Compared with traditional multimedia, "social multimedia" is by nature social as it contains rich context information besides visual content. More importantly, social multimedia contents have become the primary and promising sources for expressing opinions or sentiments by online social network users.

However, for a very long time, little has been done on multimedia sentiment analysis. Researchers have been mainly focusing on textual sentiment analysis, where the data sources can be easily identified and collected from online review forums or online shopping websites. Compared with text, there is no such a domain for collecting data of visual sentiment. However, more recently, there are some studies on collecting visual data for sentiment analysis. Thus, researchers from the multimedia community have made significant progress in this direction.

## 2 TUTORIAL DESCRIPTION

### 2.1 Structure

This tutorial aims at 1) introducing, reviewing and summarizing the recent advances in multimedia sentiment analysis, and 2) presenting insight into the challenges and future research directions in this topic. In particular, this tutorial consists of the following topics:

1. **Introduction** We introduce the task of multimedia sentiment analysis. Current status of research on this topic will

be discussed and summarized. At the same time, we also present the key research challenges in this area.

2. **Sentiment ontology.** We discuss the current efforts on building large scale sentiment ontologies [4, 5, 10]. We believe that the constructions of image and audio sentiment ontologies [8, 9] promote the research on this topic by encouraging researchers proposing novel and effective models. More importantly, the principles of collecting such large scale datasets will again assist others participating the construction of more complete and larger datasets.
3. **Visual sentiment analysis.** We present several recent studies on visual sentiment analysis [1, 13, 17, 20], including the challenges and discussions of future directions. In this section, we further discuss the research on fine-grained visual sentiment analysis [3, 7, 18], which is also known as emotion analysis. Several widely used datasets are included to compare the performance of current algorithms.
4. **Multimodality sentiment analysis** Beyond the visual sentiment analysis, we include the multimodality sentiment analysis [11, 12, 16, 19]. Challenges and relevant studies will be summarized in this section. Meanwhile, affective image captioning [2, 6, 14], which is related sentiment synthesis, will also be discussed.
5. **Future directions** We conclude this tutorial by introducing some challenges and future research directions in this area.

## 2.2 Intended Audience

This tutorial is appropriate and timely for ACM MM, graduate students, researchers and industry practitioners working in the field of social multimedia, media understanding, as well as advertising. The course materials are mainly from the recent publications in this area.

## 3 PRESENTER INFORMATION

Jiebo Luo joined the University of Rochester in Fall 2011 after over fifteen years at Kodak Research Laboratories, where he was a Senior Principal Scientist leading research and advanced development. He has been involved in numerous technical conferences, including serving as the program co-chair of ACM Multimedia 2010 and IEEE CVPR 2012. He is the Editor-in-Chief of the Journal of Multimedia, and has served on the editorial boards of the IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Multimedia, IEEE Transactions on Circuits and Systems for Video Technology, Pattern Recognition, Machine Vision and Applications, and Journal of Electronic Imaging. He has authored over 300 technical papers and 90 US patents. Prof. Luo is a Fellow of the SPIE, IEEE, and IAPR, as well as a member of the ACM, AAAI and AAAS.

Damian Borth is the director of Deep Learning Competence Center, Head of Multimedia Analysis & Data Mining Deep Learning Competence Center Multimedia Analysis & Data Mining (MADM) German Research Center for Artificial Intelligence (DFKI). In 2014-2015, he was a post-doc with Dr. Gerlald Friedland (Int. Computer Science Institute) and Prof. Trevor Darrell (UC Berkeley). In 2012, he was also a visiting scholar with Prof. Shih-Fu Chang (Columbia University). His work has also been awarded with the ICMR 2012

Best Paper Award, the McKinsey Business Technology Award, and a Google Research Award.

Quanzeng You received both B.E. and M.E. from Dalian University of Technology. Since 2012, he has been a PhD candidate in the Department of Computer Science at the University of Rochester. His advisor is Prof. Jiebo Luo. His research focuses on social multimedia, computer vision, social networks and data mining. He is interested in developing effective machine learning algorithms that can help us understand all forms of data. His most recent research is on high level visual understanding, including image captioning and visual sentiment analysis.

## REFERENCES

- [1] Mohammad Al-Naser, Seyyed Saleh Mozafari Chanijani, Syed Saqib Bukhari, Damian Borth, and Andreas Dengel. 2015. What Makes a Beautiful Landscape Beautiful: Adjective Noun Pairs Attention by Eye-Tracking and Gaze Analysis. In *Proceedings of the 1st International Workshop on Affect & Sentiment in Multimedia*. ACM, 51–56.
- [2] Philipp Blandfort, Tushar Karayil, Damian Borth, and Andreas Dengel. 2016. Introducing concept and syntax transition networks for image captioning. In *ICMR*. ACM, 385–388.
- [3] Damian Borth. 2014. Visual Learning of Semantic Concepts in Social Multimedia. *KI-Künstliche Intelligenz* 28, 4 (2014), 333–336.
- [4] Damian Borth, Tao Chen, Rongrong Ji, and Shih-Fu Chang. 2013. Sentibank: large-scale ontology and classifiers for detecting sentiment and emotions in visual content. In *ACM MM*. ACM, 459–460.
- [5] Damian Borth, Rongrong Ji, Tao Chen, Thomas Breuel, and Shih-Fu Chang. 2013. Large-scale visual sentiment ontology and detectors using adjective noun pairs. In *ACM MM*. ACM, 223–232.
- [6] Tushar Karayil, Philipp Blandfort, Damian Borth, and Andreas Dengel. 2016. Generating Affective Captions using Concept And Syntax Transition Networks. In *ACM MM*. ACM, 1111–1115.
- [7] Takuya Narihira, Damian Borth, Stella X Yu, Karl Ni, and Trevor Darrell. 2015. Mapping Images to Sentiment Adjective Noun Pairs with Factorized Neural Nets. *arXiv preprint arXiv:1511.06838* (2015).
- [8] Sebastian Sager, Damian Borth, Benjamin Elizalde, Christian Schulze, Bhiksha Raj, Ian Lane, and Andreas Dengel. 2016. Audiosentibank: Large-scale semantic ontology of acoustic concepts for audio content analysis. *arXiv preprint arXiv:1607.03766* (2016).
- [9] Christian Schulze, Dominik Henter, Damian Borth, and Andreas Dengel. 2014. Automatic detection of CSA media by multi-modal feature fusion for law enforcement support. In *ICMR*. ACM, 353.
- [10] Marco Stricker, Syed Saqib Bukhari, Mohammad Al-Naser, Seyyed Saleh Mozafari Chanijani, Damian Borth, and Andreas Dengel. 2017. Which Saliency Detection Method is the Best to Estimate the Human Attention for Adjective Noun Concepts?. In *ICAARTI*. 185–195.
- [11] Quanzeng You. 2016. Sentiment and Emotion Analysis for Social Multimedia: Methodologies and Applications. In *ACM MM*. 1445–1449.
- [12] Quanzeng You, Liangliang Cao, Hailin Jin, and Jiebo Luo. 2016. Robust Visual-Textual Sentiment Analysis: When Attention meets Tree-structured Recursive Neural Networks. In *ACM MM*. 1008–1017.
- [13] Quanzeng You, Hailin Jin, and Jiebo Luo. 2017. Visual Sentiment Analysis by Attending on Local Image Regions. In *AAAI*. 231–237.
- [14] Quanzeng You, Hailin Jin, Zhaowen Wang, Chen Fang, and Jiebo Luo. 2016. Image captioning with semantic attention. In *CVPR*. 4651–4659.
- [15] Quanzeng You and Jiebo Luo. 2013. Towards social imagematics: sentiment analysis in social multimedia. In *Proceedings of the Thirteenth International Workshop on Multimedia Data Mining*. ACM, 3.
- [16] Quanzeng You, Jiebo Luo, Hailin Jin, and Jianchao Yang. 2015. Joint visual-textual sentiment analysis with deep neural networks. In *ACM MM*. ACM, 1071–1074.
- [17] Quanzeng You, Jiebo Luo, Hailin Jin, and Jianchao Yang. 2015. Robust Image Sentiment Analysis Using Progressively Trained and Domain Transferred Deep Networks.. In *AAAI*. 381–388.
- [18] Quanzeng You, Jiebo Luo, Hailin Jin, and Jianchao Yang. 2016. Building a Large Scale Dataset for Image Emotion Recognition: The Fine Print and The Benchmark.. In *AAAI*. 308–314.
- [19] Quanzeng You, Jiebo Luo, Hailin Jin, and Jianchao Yang. 2016. Cross-modality consistent regression for joint visual-textual sentiment analysis of social multimedia. In *WSDM*. ACM, 13–22.
- [20] Jianbo Yuan, Sean Mcdonough, Quanzeng You, and Jiebo Luo. 2013. SentiBrow: image sentiment analysis from a mid-level perspective. In *International Workshop on Issues of Sentiment Discovery and Opinion Mining*. ACM, 10.