

JCDL 2018 Workshop Proposal - Image Collections: Creation, Organization, Access, and Use

Wei Lu
Wuhan University
Wuhan, Hubei, China
weilu@whu.edu.cn

Krystyna Matusiak
University of Denver
Denver, Colorado, USA
Krystyna.Matusiak@du.edu

Jevin West
University of Washington
Washington, USA
jevinw@uw.edu

Qinghua Zhu
Nanjing University
Nanjing, Jiangsu, China
qhzh@nju.edu.cn

Brian O'connor
University of North Texas
Denton, Texas, USA
Brian.O'Connor@unt.edu

Jason Lu
Cincinnati Children's Hospital
Medical Center
Cincinnati, Ohio, USA
long.lu@cchmc.org

Jiangping Chen
University of North Texas
Denton, Texas, USA
Jiangping.Chen@unt.edu

ABSTRACT

We propose to have a full day workshop at JCDL 2018. This workshop will provide an opportunity for participants to exchange research ideas on image collections, including the creation, organization, access and use (COAU) of various image datasets. We expect to discuss various theories, methods, techniques, challenges, and new research directions as related to image's COAU. Especially we would like to explore innovative ideas on image annotation, retrieval, use behavior personas, processing of different types of images, and visual image metrics. The workshop will allow researchers to communicate with their peers on projects and develop new ideas through presentation and discussion. We hope to establish a community of researchers from related disciplines and explore questions critical to the future development of image's COAU. Participants of this workshop will be invited to submit a full paper to a special issue at *The Electronic Library* (<http://www.emeraldinsight.com/journal/el>) on Image Collections.

CCS CONCEPTS

• **Information systems** → Image search; • **Computing methodologies** → Image representations;

KEYWORDS

image collections; image annotation; image retrieval; image organization

ACM Reference Format:

Wei Lu, Jevin West, Brian O'connor, Krystyna Matusiak, Qinghua Zhu, Jason Lu, and Jiangping Chen. 2018. JCDL 2018 Workshop Proposal - Image Collections: Creation, Organization, Access, and Use. In *JCDL '18: The 18th ACM/IEEE Joint Conference on Digital Libraries, June 3-7, 2018, Fort Worth, TX, USA*, Jennifer B. Sartor, Theo D'Hondt, and Wolfgang De Meuter (Eds.). ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3197026.3200208>

1 INTRODUCTION

The evolution of image generating and sharing device leads to explosion of image resources. Consequently, how to effectively manage and use these image resources and to satisfy requirement of users from different areas have been challenging issues for both academia and business. There have been growing interests in image related research areas, such as image annotation and retrieval, clustering and classification, and image understanding. In addition, image research in different subject domains, such as medical disease recognition, culture image understanding, showing features of institution, and visual image metrics have drawn great attention. Major conferences such as the IEEE International Conference on Image Processing, and ACM Multimedia Conference have advanced image processing technologies.

Images are important digital objects that carry much information. Libraries and museums have been collecting and organizing images and preserving them as cultural heritages. While image collections have been developed and used by different communities. There are few opportunities for researchers and practitioners in digital libraries to interact or learn from each other on the creation organization, access and use of image collections.

We propose this workshop to bring together researchers and information practitioners on image collections, especially image organization/tagging, retrieval diversification, and understanding to exchange research ideas and promote collaboration. Our workshop is well aligned with the theme of JCDL 2018: "the fusion of different research communities, different methods, different research directions." We hope participants could develop interesting research

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).
JCDL '18, June 3-7, 2018, Fort Worth, TX, USA
© 2018 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-5178-2/18/06
<https://doi.org/10.1145/3197026.3200208>

ideas and collaborate with each other to better organize and use image collections for various purposes.

2 PROPOSED WORKSHOP FORMAT

We plan a full day workshop with the following components:

- Short introduction by the organizers.
- Presentations of papers submitted and accepted in response to an open call for research and position papers. Papers will be presented as short oral format. The number of oral presentations will be limited to at most 12 and allow for the maximum of time for interactive activities.
- Working groups focusing on relevant topics on image collections. Solicited topics are listed in Section 3.

3 POSSIBLE TOPICS

Creation of large image collections. High-quality and large-scale image datasets (LabelMe and ImageNet) play vital roles to push relevant research forward. However, existing datasets cannot cover every aspect of people's daily life. Thus, automatic dataset construction methods have emerged to create several large-scale datasets [2-4, 7-9]. Experience from the development of image collections will be valuable to develop future high-quality and large-scale image collections.

Theories and/or framework for image annotation (image tagging). Image annotation could be divided into two parts: manual annotation and automatic annotation. The former one, also named "image tagging", represents the traditional way to annotate image, it is usually processed by experts and social media users of images; The latter one represents state-of-the-art way for image annotation, that is, use computer to annotate image automatically. Currently, as the development of artificial intelligence, machine learning has been widely used in automatic image annotation. We would like to call for submission in this topic to advance research in both manual annotation and automatic annotation, and to explore new direction for development.

Image classification. Image classification mainly utilizes features of image, such as color, texture, shape and spatial relationship, to classify images into different categories according to the similarity degree of these features. The technologies in this area refer to image feature extraction, feature clustering, and so on. Image analysis, one of the vital parts of image classification, is aiming to transform digital image features to easy-to-read forms to assist image classification.

Image retrieval and image results diversification. Subtopics can be: image retrieval system construction and evaluation, content-based image retrieval, image retrieval based on semantic and visual features, and relevance feedback. As for image retrieval results diversification, information users often want to retrieve images that are both semantically relevant and visually diversity. For example, users submitting a query term like "apple" into a system may want to retrieve more diverse search results, i.e., images of apples with different shapes, colors, and textures.

Medical images. The use and role of medical imaging technologies in clinical oncology has greatly expanded from primarily a diagnostic tool to include a more central role in the context of individualized medicine over the past decade since people are paying

more and more attention to healthcare [5]. However, the automatic analysis of medical images resource is still a very challenging task because the main objective of this task is not only to reach high accuracy but also to identify which parts of human body are infected by the disease[6].

Social images. The evolution of image sharing websites has led to the explosion of image resources on the Internet. Social media user often uploads their images for sharing and communication, which aims to record and transmit information. Besides, they process tagging on image for the purpose of management and description[1]. Research in social image includes social image tags, tag classification, user motivation for sharing image, user behavior and persona.

Innovative technologies on image processing and analysis. Image processing is a sub-category of signal processing, and it is one of the technologies of image classification. It also assists in feature extraction and pattern recognition. Techniques in this area include image compression, enhancement and recovery, matching, description and recognition. It has been widely used in satellite image processing, medical image processing, face detection, feature detection, face identification, microscope image processing and car barrier detection.

4 ORGANIZERS

Wei Lu, Professor, Wuhan University, China

Jevin West, Assistant Professor, University of Washington, USA

Brian O'Connor, Professor, University of North Texas, USA

Krystyna Matusiak, Associate Professor, University of Denver, USA

Qinghua Zhu, Professor, Nanjing University, China

Jason Lu, Cincinnati Children's Hospital Medical Center, USA

Jiangping Chen, Professor, University of North Texas, USA

REFERENCES

- [1] Morgan Ames and Mor Naaman. 2007. Why we tag: motivations for annotation in mobile and online media. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 971-980.
- [2] Yalong Bai, Kuiyuan Yang, Wei Yu, Chang Xu, Wei-Ying Ma, and Tiejun Zhao. 2015. Automatic image dataset construction from click-through logs using deep neural network. In *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 441-450.
- [3] Duc-Tien Dang-Nguyen, Cecilia Pasquini, Valentina Conotter, and Giulia Boato. 2015. Raise: A raw images dataset for digital image forensics. In *Proceedings of the 8th ACM Multimedia Systems Conference*. ACM, 219-224.
- [4] Jyoti Deshmukh and Udhav Bhosle. 2016. Image mining using association rule for medical image dataset. *Procedia Computer Science* 85 (2016), 117-124.
- [5] Philippe Lambin, Emmanuel Rios-Velazquez, Ralph Leijenaar, Sara Carvalho, Ruud GPM van Stiphout, Patrick Granton, Catharina ML Zegers, Robert Gillies, Ronald Boellard, André Dekker, et al. 2012. Radiomics: extracting more information from medical images using advanced feature analysis. *European journal of cancer* 48, 4 (2012), 441-446.
- [6] Eka Miranda, Mediana Aryuni, and E Irwansyah. 2016. A survey of medical image classification techniques. In *Information Management and Technology (ICIMTech), International Conference on*. IEEE, 56-61.
- [7] Konstantin Pogorelov, Kristin Ranheim Randel, Carsten Griwodz, Sigrun Losada Eskeland, Thomas de Lange, Dag Johansen, Concetto Spampinato, Duc-Tien Dang-Nguyen, Mathias Lux, Peter Thelin Schmidt, et al. 2017. Kvasir: a multi-class image dataset for computer aided gastrointestinal disease detection. In *Proceedings of the 8th ACM on Multimedia Systems Conference*. ACM, 164-169.
- [8] Fisher Yu Yinda Zhang Shuran Song and Ari Seff Jianxiang Xiao. 2015. Construction of a Large-scale Image Dataset using Deep Learning with Humans in the Loop. *arXiv preprint arXiv:1506.03365* (2015).
- [9] Yan Xia, Xudong Cao, Fang Wen, and Jian Sun. 2014. Well begun is half done: Generating high-quality seeds for automatic image dataset construction from web. In *European Conference on Computer Vision*. Springer, 387-400.