



## Call for Papers

# Special Issue on Cross-Media Analysis for Visual Question Answering

### 1. Scope:

Visual Question Answering (VQA) is a recent hot topic in multimedia analysis, computer vision, natural language processing, and even a broad perspective of artificial intelligence, which has attracted a large amount of interest from the deep learning, computer vision, and natural language processing communities. Given an image (or a video clip) and a question in natural language, VQA requires grounding textural concepts to visual elements so as to infer the correct answer. The challenge lies in that, in most cases, it requires reasoning over the connecting between visual content and languages as well as the external knowledge. Towards general applications, besides the understanding of visual content, potential abilities of VQA largely come from the leveraging of different kinds of data (e.g., visual, audio, and text etc.) across multiple sources (e.g., social-media sites, surveillance videos, and Wikipedia etc.) for knowledge discovering and QA reasoning, which is recognized as cross-media analysis in multimedia scope. Though booming recently, VQA still remains in its infant as there are so many challenges. Recently proposed methods are always specific to certain datasets that lack generalization ability to some extent, as different benchmark may have different characteristics of its own. Though reasoning ability is always claimed, most ‘reasoning’ simply uncovers latent connections between visual elements and textual/semantic facts during the training on a large amount of images/videos annotated with QA pairs. As reasoning depends on knowledge, it still remains implicit how to appropriately and effectively incorporate the knowledge-based reasoning into the data-driven learning framework of VQA. Towards a broad perspective of applications, cross-media knowledge discovering and reasoning has yet been well exploited in existing VQA research.

This special issue serves as a forum for researchers all over the world to discuss their works and recent advances in VQA methods and its applications. Both state-of-the-art works, as well as literature reviews, are welcome for submission. Especially, to provide readers of the special issue with a state-of-the-art background on the topic, we will invite one survey paper, which will undergo peer review. This special issue seeks to present and highlight the latest developments on practical methods of VQA via cross-media reasoning. Papers addressing interesting real-world applications are especially encouraged. Topics of interest include, but are not limited to,

- Cross-media analysis for knowledge discovering
- Cross-media reasoning on the cross-media knowledge graph
- Knowledge-embedded for visual question answering
- Co-attention model for textual questions and visual content
- Sequential deep learning for visual question answering
- Reinforcement learning for visual question answering

- Temporal reasoning for video question answering
- Visual captioning for visual question answering
- Large-scale video question answering benchmark
- Real-world applications of cross-media analysis for visual question answering, e.g., human-computer interaction, intelligent cross-media search, visual description for blinded, incident report for surveillance, and seeing chat bot etc.

## **2. Important Dates:**

- Paper Submission Deadline: ~~June 30, 2018~~ **Extended to July 31th, 2018**
- First Notification: August 31, 2018
- Notification of Acceptance: December 1, 2018
- Final Manuscript Due: December 31, 2018

## **3. Review Process:**

The review process will comply with the standard review process of the ACM Transactions on Multimedia Computing, Communications and Applications (ACM TOMM) journal. Each paper will receive at least three reviews from experts in the field.

## **4. Submission Instructions:**

Prospective authors are invited to submit their manuscripts electronically after the “open for submissions” date, adhering to the ACM Transactions on Multimedia Computing, Communications and Applications (ACM TOMM) journal guidelines (see <http://tomm.acm.org/authors.cfm>). Please submit your papers through the online system and be sure to select the special issue. Manuscripts should not be published or currently submitted for publication elsewhere. Submitted manuscripts should not have been published previously, nor be under consideration for publication elsewhere. If the submission is an extended work of a previously published conference paper, please include the original work and a cover letter describing the changes that have been made. According to ACM TOMM publication policy previously published conference papers can be eligible for publication provided that at least 25% new material is included in the journal version.

## **3. Guest Editors**

- Richang Hong, Hefei University of Technology, Email: hongrc@hfut.edu.cn
- Yahong Han, Tianjin University, China, Email: yahong@tju.edu.cn
- Tat-Seng Chua, National University of Singapore, Email: chuats@comp.nus.edu.sg