Jianfeng Dong

Birthdate: February 4, 1991 Phone: (86)15068116106
Homepage: danieljf24.github.io Email: danieljf24@zju.edu.cn
Room 612, Caoguangbiao Building, 38 Zheda Road, Zhejiang University, Hangzhou, 310027, P. R. China

RESEARCH INTERESTS

Cross-media retrieval and deep learning. Specifically, I am interested in deep understanding of images, texts, videos, and their relations.

EDUCATIONS

Visiting student May 2015 - present
Renmin University of China (Advisor: Dr. Xirong Li)

M.S. & Ph.D. candidate in Computer Science and Technology September 2013 - present
Zhejiang University (Advisor: Prof. Duanqing Xu)

B.S. degree in Software Engineering September 2009 - June 2013
Zhejiang University of Technology

PULICATIONS

[1] **Jianfeng Dong**, Xirong Li, Weiyu Lan, Yujia Huo, and Cees G. M. Snoek. *Early Embedding and Late Reranking for Video Captioning*. In Proc. of ACM Multimedia, 2016 (Multimedia Grand Challenge Session).

[2] **Jianfeng Dong**, Xirong Li, Cees G. M. Snoek. *Word2VisualVec: Cross-Media Retrieval by Visual Feature Prediction*. arXiv preprint, 2016.

[3] Xirong Li, Weiyu Lan and **Jianfeng Dong.** *Adding Chinese Captions to Images.* In Proc. of ICMR 2016.

[4] **Jianfeng Dong**, Xirong Li, Shuai Liao, Jieping Xu, Duanqing Xu and Xiaoyong Du. *Image retrieval by cross-media relevance fusion*. In Proc. of ACM Multimedia, 2015 (Multimedia Grand Challenge Session).

[5] **Jianfeng Dong**, Xirong Li and Duanqing Xu. *Cross-Media Similarity Evaluation for Web Image Retrieval in the Wild*. under review.

COMPUTER SKILLS

- Programming proficiency
  - Good ability and extensive experience in python, C/C++, MATLAB
  - Extensive experience of using deep learning Framework : Theano, Keras
  - Familiar with Caffe, Tensorflow
- Have experience in academic writing and presentation

HONORS & AWARDS

- June 2016, Ranked the 4th among 22 teams in MSR Video to Language Challenge
- October 2015, Ranked the 1st in the image retrieval task of MSR-Bing Image Retrieval Challenge
- June 2013, Outstanding Graduate of Zhejiang Province (Top 1%)
- June 2013, University-wide outstanding bachelor thesis award of Zhejiang University of Technology
- August 2012, Third prize of China Students Service Outstanding Innovation Competition

ACADEMIC ACTIVITIES

- October 2015, Attending ACM MM 2015, Brisbane, Australia
- May 2014, Volunteer of ICML 2014, Beijing, China

PROJECTS

Video Captioning May 2016 - present

- Goal: Given a video, we aim to automatically generate a natural language sentence to describe main objects, scenes and events present in the video.
- We extend the popular ConvNet + LSTM model with two novel modules. One is early embedding, which enriches the current low-level input to LSTM by tag embeddings. The other is late reranking, which re-scores generated sentences in terms of their relevance to a specific video.
- This work led to a video captioning system that ranks 4th place in MSR Video to Language Challenge.

Cross-media Retrieval May 2015 - present

- Goal: Investigate how to estimate cross-media relevance between a given query and an unlabeled image.
- We propose cross-media relevance fusion, a conceptually simple framework that exploits the power of individual methods for cross-media relevance estimation. Four base cross-media relevance functions are investigated, and later combined by weights optimized on the development set.
- This work led to the top-performer solution for MSR-Bing Image Retrieval Challenge at ACM MM2015.

Novel Display of Multimedia Data on Android Mobile Devices 2013

- Goal: Design a unified framework to efficiently display high resolution images, panorama and 3D models on an Android mobile device.
- The high resolution images are divided into image patches (regular square parts) of different resolutions in advance. When users browse images, the device only needs to request some image patches that meet requirements of display rather than the whole image, which can reduce the waiting time at the client side.
- For panorama and 3D models, the client transmits the operations (such as rotation, zoom in and out) of the user to remote servers, and the servers transmitted the image shot of the rendered model back to the client in real time. Fast switching image shots of model can make the user have a feeling of browsing the panorama or 3D model. In this way, panorama and 3D models are rendered on the remote servers instead of the local mobile device (client), which not only improve the rendering efficiency but also save the battery power of mobile devise.
- This work was awarded university-wide outstanding bachelor thesis at Zhejiang University of Technology.

REFEREES

Dr. Xirong Li, Renmin University of China (xirong@ruc.edu.cn)
Prof. Duanqing Xu, Zhejiang University (xdq@zju.edu.cn)