

# KAN CHEN

Apt. 6, 1210 W 27th Street, LA, CA, 90007

+1-213-880-8751 • [kanchen@usc.edu](mailto:kanchen@usc.edu)

---

## EDUCATION

**PhD of Electrical Engineering** • University of Southern California, LA, CA, USA 2013-2018  
**Master of Computer Science** • University of Southern California, LA, CA, USA 2016-2017  
**Major GPA:** 4.00/4.00 • **Rank:** 1/300  
**Selected Courses:** Machine Learning(A) · Algorithm Analysis(A) · Computer Vision(A) · Digital Image Processing(A) · Database(A)  
**Bachelor of Electronics Engineering** • Tsinghua University, Haidian, Beijing, China 2009-2013  
**Major GPA:** 3.94/4.00 (94.4/100) • **Rank:** 7/242

---

## INTERNSHIP

**Research Intern, Adobe System Inc.** (May - August.2016)

*Multimodal Image Ranking system*

- Constructed a multimodal image ranking system for retrieving related images based on users' queries
- Improved the ranking performance over 5% using [Torch](#)
- Incorporated in [Adobe Stock](#) search engine
- A paper published in IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017
- A patent for this technique (P6545-US/ADBS.263071)

**Research Intern, Institute of Deep Learning in Baidu USA LLC** (May - August.2015)

*Visual Question Answering*

- Proposed a configurable convolutional neural network to address visual question answering
- Realized *LSTM* model to extract semantic information from question sequence on [PaddlePaddle](#) platform
- Achieved 5% improvement on *Toronto-COCO-QA* dataset and 7% improvement on *DAQUAR* dataset
- A paper published in IEEE Conference on Computer Vision and Pattern Recognition Workshop 2016
- A patent for this technique (BN151030USN1)

**Research Intern, Toyota Technological Institute in Chicago** (July.2012)

*Recovering Layout of indoor Images*

- Combined Kinect's data with Geometric context and orientation map to get a more accurate estimation of Layout of single indoor picture.
- Combined Geometric context and Normal vectors features to boost performance
- Developed a labeling tool to generate sufficient training data
- A paper published in International Conference on Computer Vision (ICCV), 2013

---

## PROJECT

**Image Grounding, The Defense Advanced Research Projects Agency** (September.2015 - Current)

*Project: MEDIFOR image manipulation detection system*

- In charge of image grounding part
- Developed a multimodal vregression network for proposal generation in [TensorFlow](#)
- Applied reinforcement learning technique to leverage context information
- Achieved more than 14% and 17% improvement in accuracy on two academic datasets
- Two papers published for our system (*ICMR* 2017 and *ICCV* 2017)

**People Re-identification, The Defense Advanced Research Projects Agency** (February.2017 - Current)

*Project: MEDIFOR image manipulation detection system*

- Applied people re-identification technique to detect potential manual image manipulation
- Applied a Domain Guided Dropout Network (CVPR 2016) to extract visual features
- Achieved 15% improvement in accuracy on a project re-identification dataset
- Apply metric learning technique to further boost performance

**Face Recognition & Detection, Intelligence Advanced Research Projects Activity** (August.2013 - 2015)

*Project: JANUS face recognition system*

- In charge of face detection part
- Applied a modified algorithm in *face detection without bells and whistles* (ECCV 2014)
- Modified code to train neural network using R-CNN
- Ranked #2 in competition on behalf of the team of University of Southern California

---

## SKILLS

- Proficiency in software development in Python, C, C++, MATLAB,  $\text{\LaTeX}$ .
- Proficiency in deep learning platforms: [TensorFlow](#), [Torch](#), [Caffe](#), [PaddlePaddle](#).
- Familiarity with Android, Swift, Database Design, SQL.

- **To appear:** **Kan Chen**, Rama Kovvuri, Jiyang Gao, Ram Nevatia: MSRC: Multimodal Spatial Regression with Semantic Context for Phrase Grounding, *International Journal of Multimedia Information Retrieval (IJMIR)*, 2017
- **Published:** **Kan Chen**, Rama Kovvuri, Ram Nevatia: Query-guided Regression Network with Context Policy for Phrase Grounding, *International Conference on Computer Vision (ICCV)*, 2017
- **Published:** Jiyang Gao, Zhenheng Yang, Chen Sun, **Kan Chen**, Ram Nevatia: TURN TAP: Temporal Unit Regression Network for Temporal Action Proposals, *International Conference on Computer Vision (ICCV)*, 2017
- **Published:** **Kan Chen**, Rama Kovvuri, Jiyang Gao, Ram Nevatia: MSRC: Multimodal Spatial Regression with Semantic Context for Phrase Grounding, *ACM International Conference on Multimedia Retrieval (ICMR)*, 2017
- **Published:** **Kan Chen**, Trung Bui, Chen Fang, Zhaowen Wang, Ram Nevatia: AMC: Attention guided Multimodel Correlation Learning for Image Search, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017
- **Published:** **Kan Chen**, Jiang Wang, Wei Xu, Liang-Chieh Chen, Haoyuan Gao, Ram Nevatia: ABC-CNN: An Attention Based Approach for Visual Question Answering, *IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW)*, 2016
- **Published:** Song Cao, **Kan Chen**, Ram Nevatia: Abstraction Hierarchy and Self Annotation Update for Fine Grained Activity Recognition, *WACV*, 2016
- **Published:** Song Cao, **Kan Chen**, Ram Nevatia: Activity Recognition and Prediction with Pose based Discriminative Patch Model, *WACV*, 2016
- **Published:** **Kan Chen**, Jiang Wang, Wei Xu: Attention Based Model in Visual Question Answering, *Bay Learn Forum*, 2015
- **Published:** Jian Zhang, **Kan Chen**, Alexander G. Schwing, Raquel Urtasun: Estimating the 3D Layout of Indoor Scenes and its Clutter from Depth Sensors, *IEEE International Conference on Computer Vision (ICCV)*, 2013
- **Published:** Qiang Ning, **Kan Chen**, Li Yi, Chuchu Fan, Yao Lu, Jiangtao Wen: *Image super resolution via analysis sparse prior*, IEEE transaction of Signal Process, (2013) 1-1