# Kaili Zhao

Assistant Professor
Beijing University of Posts and Telecom.
No. 10 Xitucheng Road, Haidian District
Beijing, China 100876

Cell: (+86) 15201646826
Email: kailizhao@bupt.edu.cn
kaili.zhao0619@gmail.com
Github: https://github.com/zkl20061823

## Research Interests

My research interests are in **computer vision** and **machine learning**. I have developed techniques spanning over structured multi-task learning, weakly-supervised learning and deep learning. I also lead research projects in facial expression analysis, AU detection and image segmentation.

## Degrees

Ph.D. (2016) in Information and Signal Processing, Beijing University of Posts and Telecom.
B.S. (2010) in Automation, Hefei University of Technology

## Research Grants

- Principal Investigator, "Street-view based Crowd Counting analysis," with DOCOMO Beijing Communications Laboratories Co., Ltd., 03/2019–08/2019. CNY ¥250, 000 (USD $40,000)
- Principal Investigator, "Drone-based Image Processing for Crack Inspection," with DOCOMO Beijing Communications Laboratories Co., Ltd., 06/2018–08/2019. CNY ¥200, 000 (USD $30,000)
- Principal Investigator, "Weakly-supervised Spectral Clustering and Its Application to Recognize Facial Expressions in 1 Million Facial Images," with National Natural Science Foundation of China, 09/2017–12/2020. CNY ¥300, 000 (USD $40,000)
- Co-Principal Investigator, "Depression Detection Model based on Automated Facial Action Unit Analysis," with Institute of Mental Health, Beijing Anding Hospital (top psychiatry and mental health hospital in China), 11/2017–11/2021. CNY ¥1000, 000 (USD $156,000)
- Principal Investigator, "Multi-label Learning for Facial Action Unit Detection," with Fundamental Research Funds for the Central Universities, 07/2017–10/2018. USD CNY ¥150, 000 (USD $20,000)

## Research Experience

- **Street-view based Crowd Counting** 03/2019–08/2019
  DOCOMO Beijing Labs, Beijing
  - Developed a crowd counting module for "AI Box" product of DOCOMO Labs.
  - Filed one patent in China (in confidential status)
- **Drone-based Image Processing for Crack Inspection** 06/2018–08/2019
  DOCOMO Beijing Labs, Beijing
  - Developed a deep U-shape model for inhomogeneity pixel-level crack detection.
  - The proposed deep model showed that the encoder-decoder model with skip layers fit the crack inspection problem very effective (one paper is published in [J1]).
- **Scalable Weakly-Supervised Learning for Facial AU Detection** 11/2016–04/2017
  Dept. of Electrical and Computer Engineering, The Ohio State University (OSU)

- Developed a scalable weakly-supervised spectral clustering algorithm for learning facial action unit from 1 million web images (published in [C1])
- Results showed that the proposed method can identify and correct inaccurate labels in weakly annotated images, further improving the performance of existing CNN-based models.

- **End-to-End Supervised Facial AU Detection** 04/2015–12/2015
  Institute of Computing Technology, Chinese Academy of Sciences
  - Developed deep region and multi-label learning (DRML) for facial AU detection (published in [C2])
  - The proposed region layer demonstrated effectiveness over standard conv layer (i.e., shared kernels over an entire image) and locally connected layer (i.e., DeepFace by Facebook).
  - DRML has become as a benchmark approach for CNN-based AU detection. Our implementation on GitHub has been forked ∼27 times (`https://github.com/zkl20061823/DRML`).

- **Joint Learning of Facial Region and AU Relations for Expression Analysis** 03/2014–03/2015
  Robotics Institute, Carnegie Mellon University (CMU)
  - Developed a joint patch and multi-label learning (JPML) framework to jointly learn sparse facial patches and a multi-label classifier, with consideration of AU relations (published in [J2, C3])
  - Results showed that the identified local patches are related to active facial regions of different expressions. Considering label relations improves performance for AUs with low base-rate.

- **Structured Multi-task Learning** 09/2013–02/2014
  Statistics Institute, The Ohio State University (OSU)
  - Developed structured multi-label learning algorithms to categorize compound expressions [C4, C5]
  - Results showed that learning multiple expressions with group constraints improves performance for compound expression recognition.

- **Automatic Backlight Detection** 03/2012–07/2013
  Huawei Technologies Co. Ltd, Beijing
  - Led a team to build a scene classification and backlight detection system
  - Filed two patents in China (in confidential status)

## Teaching & Mentoring

**Beijing University of Posts and Telecommunications**
Autumn (17′, 18′), Digital Image Processing, 2-credit course, 3 classes (190 undergraduate students)

**Doctoral Student Advisor**
Xiaolin Song (17′): Automated Depression Detection Based on Facial Action Units
Shi Pu (15′), Jin Feng (17′): Tracking by Temporal Propagation

**Undergraduate Research Projects**
Danyi Zhang, Zichen Tian (17′): End-to-end Skin Lesion Segmentation

**Undergraduate Thesis Advisor**
Danyi Zhang etc., 10 students (17, 18′): Facial Expression Recognition, Skin Lesion Analysis, and Concrete Crack Inspection.

## Honors & Awards

Excellent Dissertation Award, CNY ¥10, 000 07/2016
National Scholarship for Excellent Ph.D. Students, CNY ¥20, 000 12/2015
Distinguished Doctorate Student Award (top 3%, 10/300 Ph.D. students), CNY ¥50,000 10/2015
National Scholarship for Abroad Visiting Ph.D. Students, USD $24, 000 03/2013

# Publications

**Journal**

[J1] **Kaili Zhao**, Anxin Li, Bo Xiao, Jun Guo, "DUNet: Dilated U-shape Convolutional Neural Network for Pixel-level Inhomogeneity Concrete Crack Detection," *Pattern Recognition (***PR***)*, preprint

[J2] **Kaili Zhao**, Wen-Sheng Chu, Fernando De la Torre, Jeffrey F. Cohn, and Honggang Zhang, "Joint Patch and Multi-label Learning for Facial Action Unit and Holistic Expression Recognition," *IEEE Trans. on Image Processing (***TIP***)*, 2016.

[J3] **Kaili Zhao**, Honggang Zhang, Zhanyu Ma, Yi-Zhe Song, and Jun Guo, "Multi-label Learning with Prior-knowledge for Facial Expression Analysis" *Neurocomputing*, 2014.

[J4] Honggang Zhang, **Kaili Zhao**, Yi-Zhe Song, and Jun Guo, "Text Extraction from Natural Scene Image: A Survey," *Neurocomputing*, 2013.

**Conference**

[C1] **Kaili Zhao**, Wen-Sheng Chu, and Aleix M. Martinez, "Learning Facial Action Units from Web Images with Scalable Weakly Supervised Clustering," in *International Conference on Computer Vision and Pattern Recognition* (**CVPR**), 2018.

[C2] **Kaili Zhao**, Wen-Sheng Chu, and Honggang Zhang, "Deep Region and Multi-label Learning for Facial Action Unit Detection," in *International Conference on Computer Vision and Pattern Recognition* (**CVPR**), 2016.

[C3] **Kaili Zhao**, Wen-Sheng Chu, Fernando De la Torre, Jeffrey F. Cohn, and Honggang Zhang, "Joint Patch and Multi-label Learning for Facial Action Unit Detection," in *International Conference on Computer Vision and Pattern Recognition* (**CVPR**), 2015.

[C4] **Kaili Zhao**, Honggang Zhang, and Jun Guo, "An adaptive Group Lasso based Multi-label Regression Approach for Facial Expression Analysis," in *International Conference on Image Processing* (**ICIP**), 2014.

[C5] **Kaili Zhao**, Honggang Zhang, Mingzhi Dong, and Jun Guo "A Multi-label Classification Approach for Facial Expression Recognition," in *IEEE Visual Communications and Image Processing*, 2013.

# Professional Services

- **Journal reviews**
  - IEEE Transactions on Affective Computing (**TAFFC**) (1)
  - IEEE Transactions on Multimedia (**TMM**) (1)
  - IEEE Transactions on Image Processing (**TIP**) (3)
  - Pattern Recognition (1)
  - Journal of Visual Communication and Image Representation (1)
  - Pattern Recognition Letters (2)
  - Frontiers of Computer Science (1)
- **Conference reviews**
  - International Conference in Computer Vision (**ICCV**) (19′)
  - IEEE/CVF Computer Vision and Pattern Recognition (CVPR 19′)
  - IEEE Winter Conf. on Applications of Computer Vision (WACV 18′)
  - International Conference on Network Infrastructure and Digital Content (IC-NIDC 18′)

## Invited Talks

**'Computer Vision Crash Course"**, DOCOMO Beijing Labs, Beijing, 12/2018
**"Learning Facial Action Units from Web Images with Scalable Weakly Supervised Clustering"**, Microsoft Research Asia, Beijing, 05/2018
**"Automatic Facial Action Unit Detection"**, China Academy of Chinese Medical Science, Beijing, 03/2018
**"Text Extraction from Natural Scene Image: A Survey"**, Canon Information Technology (Beijing) Co., LTD, 09/2011

## Skills

- **Language**: Python, Matlab, C/C++, Caffe, Tensorflow, LaTeX
- **Platforms**: Unix, Linux