Kaili Zhao

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

Email: kailizhao@bupt.edu.cn
kaili.zhao0619@gmail.com
Github: https://github.com/zkl20061823
Homepage: https://zkl20061823.github.io

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, semantic segmentation, crowd counting, and pedestrian detection.

Degrees

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

Research Grants

- Principal Investigator, "People Counting for Indoor Watching," with DOCOMO Beijing Communications Laboratories Co., Ltd., 10/2019–06/2020. CNY ¥250, 000 (USD $40,000)
- Principal Investigator, "Crowd Counting on Street-view Images," 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

- **People Counting for Indoor Watching** 10/2019–now
  DOCOMO Beijing Labs, Beijing
  - This is an ongoing project that aims to count how many people are attentive to a monitor or TV screen.
- **One-shot learning for Occluded Pedestrian Detection** 08/2019–now
  Pattern Recognition and Intelligent Labs, Beijing
  - Developed a progressive pedestrian detector to learn occluded pedestrian from visible parts by conditioning on occlusion statistics.
  - The results showed our model improved 4% miss rate in heavy occluded scenario and obtained comparable testing time with the popular SSD. (We have submitted one manuscript to CVPR 2020.)

- **Crowd Counting on Street-view Images** 03/2019–08/2019
  DOCOMO Beijing Labs, Beijing
  - Developed a light-weight crowd counting module for "AI Box" product of DOCOMO Labs.
  - The proposed light-weight deep model emphasizes learning on various head sizes. In about 200 people scenario, our model achieved absolute error of 10 and obtained 0.82 million parameters, which is 50% lighter than 16.2 millions included in the SoA methods.
  - Filing one patent in China (in confidential status)

- **Drone-based Image Processing for Crack Inspection** 06/2018–08/2019
  DOCOMO Beijing Labs, Beijing
  - Developed refinement fitting strategy and learned constraints of topology of cracks to jointly fit local details and capture global structure.
  - The results achieved 20% higher F1 score when comparing the SoA semantic segmentation models used in crack detection. (One paper is under review 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 [C2])
  - 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 [C3])
  - 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 [J3, C5])
  - 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 [C6, C7]
  - 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 one patent in China.

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′-now): Occluded Pedestrian Detection.
Shi Pu (15′-now), Jin Feng (17′-now): Tracking by Detection.

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

Undergraduate Thesis Advisor
Danyi Zhang etc., 15 students (17, 18′, 19′): Facial Expression Recognition, Skin Lesion Analysis, Concrete Crack Inspection, Head Pose Estimation, and Pedestrian Detection.

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, "Learning Spatial-Semantic Preserving Pixel-level Crack Detector by Refinement Fitting," *Under Review.*

[J2] Feng Jin, Peng Xu, Shi Pu, **Kaili Zhao**, Honggang Zhang, "Robust Visual Tracking by Embedding Combination and Weighted-Gradient Optimization," *Pattern Recognition*, 2020.

[J3] **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.

[J4] **Kaili Zhao**, Honggang Zhang, Zhanyu Ma, Yi-Zhe Song, and Jun Guo, "Multi-label Learning with Prior-knowledge for Facial Expression Analysis," *Neurocomputing*, 2014.

[J5] Honggang Zhang, **Kaili Zhao**, Yi-Zhe Song, and Jun Guo, "Text Extraction from Natural Scene Image: A Survey," *Neurocomputing*, 2013.

Conference

[C1] Jin Feng, **Shi Pu**, Kaili Zhao, Honggang Zhang, Tianming Du, "Enhanced Initialization with Multi-Stage Learning for Robust Visual Tracking," in *IEEE Visual Communications and Image Processing* (**VCIP**), 2019, **oral***.

[C2] **Kaili Zhao**, Wen-Sheng Chu, and Aleix M. Martinez, "Learning Facial Action Units from Web Images with Scalable Weakly Supervised Clustering," in *IEEE / CVF Conference on Computer Vision and Pattern Recognition* (**CVPR**), 2018.

[C3] **Kaili Zhao**, Wen-Sheng Chu, and Honggang Zhang, "Deep Region and Multi-label Learning for Facial Action Unit Detection," in *IEEE / CVF Conference on Computer Vision and Pattern Recognition* (**CVPR**), 2016.

[C6] Shi Pu, Honggang Zhang, **Kaili Zhao**, "Structure and Appearance Preserving Network Flow for Multi-object Tracking," in *International Conference on Pattern Recognition* (**ICPR**), 2016.

[C5] **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 *IEEE / CVF Conference on Computer Vision and Pattern Recognition* (**CVPR**), 2015.

[C6] **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.

[C7] **Kaili Zhao**, Honggang Zhang, Mingzhi Dong, and Jun Guo "A Multi-label Classification Approach for Facial Expression Recognition," in *IEEE International Conference on Visual Communications and Image Processing*, 2013.

Professional Services

- **Journal reviews**
  - Transactions on Pattern Analysis and Machine Intelligence (**TPAMI**)
  - IEEE Transactions on Affective Computing (**TAFFC**)
  - IEEE Transactions on Multimedia (**TMM**)
  - IEEE Transactions on Image Processing (**TIP**)
  - Pattern Recognition (**PR**)
  - Journal of Visual Communication and Image Representation
  - Pattern Recognition Letters
  - Frontiers of Computer Science
- **Conference reviews**
  - International Conference in Computer Vision (**ICCV**)
  - IEEE/CVF Computer Vision and Pattern Recognition (**CVPR**)
  - European Conference on Computer Vision (**ECCV**)
  - AAAI Conference on Artificial Intelligence (**AAAI**)
  - IEEE Winter Conf. on Applications of Computer Vision (**WACV**)
  - International Conference on Network Infrastructure and Digital Content (**IC-NIDC**)

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++, Caffe, Tensorflow, LaTeX
- **Platforms**: Unix, Linux