

Dissertation: Automatic Face Understanding via Big Data

Overview

My research topic is applied machine vision with an emphasis on automatic face understanding. Specifically, with intent to advance security systems, affective computing, and human-computer interactive technologies as a whole. We built many face- and motion-based image and video collections to support our research efforts. Most notably, the Families in the Wild (FIW) Database [4, 5, 8]. FIW, the largest labeled image set for automatic kinship recognition, has been focal to many efforts through the building of my dissertation, providing opportunity for both technological and service-based contributions to the communities of FG. For instance, the annual challenge in kinship recognition—namely, Recognizing Families In the Wild (RFIW)—starting as a Data Challenge Workshop in conjunction with ACM MM (RFIW2017 [6]), to then be hosted as an FG challenge for the next three consecutive years (RFIW2018, 19, and 20). Having published several works in automatic kinship recognition, in parallel to providing the data, organizing the evaluations, and calling for papers for proceedings, we hosted several tutorials on the topic at top-tier conferences (i.e., FG, ACM-MM, and CVPR). Furthermore, Kaggle invited us to host a challenge as a premier competition. To our complete delight, press coverages also were released as of recently (ACM News and RSIP Vision). [9]

Early on in my Ph.D., I led a joint SMILE Lab/MIT-Lincoln Lab team in our TRECVID debut, placing 3rd highest in Multimedia Event Detection [1, 3]. This experience taught me to handle big data and apply deep learning to the constructing, benchmarking, and evaluating big data (i.e., as later done for the FIW database soon after that). Then, we created a multi-modal motion capture database for an automatic understanding of human actions and poses. Later on, I published the work done as part of my internship at Snap Inc during the summer of 2018, which involved automatic face detection and tracking [2]. This past summer, I worked at ISMConnect, with the task of investigating bias in automated facial recognition systems to be published in this year’s IEEE FG conference. See Fig. 1 for an end-to-end timeline of the research plan. [11]

Working Plan & Progress

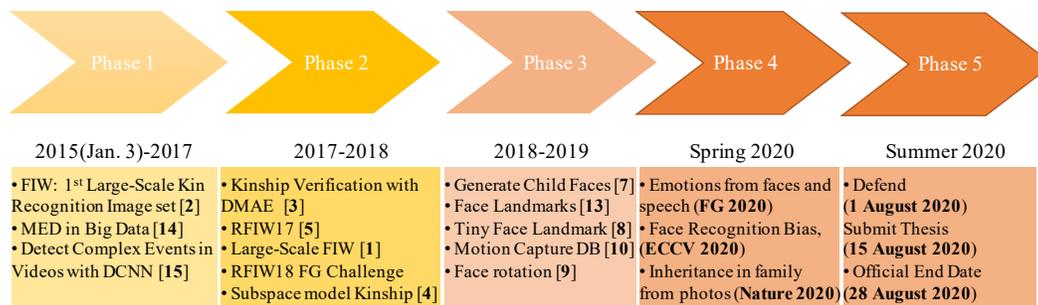


Figure 1: Progress of my research work. There are four phases, and almost one year one phase.

[Summary]: My research progress and plans can be broken down in four-phases: *Visual understanding via deep neural networks (Phase 1)*; *Building, labeling, and organizing image and video databases (Phase 2)*; *Experimental design: benchmarks and Evaluations (Phase 3)*; and soon to be *Nature-based study (Phase 4 (Final))*.

[Phase 1]: My thesis mainly is on face understanding. Specifically, kinship recognition (FG 2017, ACM MM 2017, FG 2018, TPAMI 2018, ICCV 2019). Additionally, video understanding: Complex event detection (TRECVID 2015 and SPIE 2016). Most lately we have deep dived in the problem of bias in facial recognition systems (FG 2019).

[Phase 2]: I have built several data resources. Most notably, was the first large-scale kinship-based image-set, FIW (ACM MM 2016, TPAMI SE 2018). Besides, I built a motion-captured human action [7]. Our current work in the bias study is supported by a face dataset that I built (to be released upon publication).

[Phase 3]: We hosted several challenges at top conferences (ACM MM 2017, FG 2018, FG 2019) on kinship recognition (i.e., verification and classification) using FIW, RFIW 2017, 2018, and 2019. Additionally, I served as a chair for several workshops at top tier conferences (e.g., CVPR 2018 & 2019 and ICME 2018 & 2019). [10] **[Phase 4]:** [present to June 2020]. My focus is to be put on a nature-based study using our FIW dataset. For this, we concepts of inheritance as seen in imagery. Our goal is to see if appearance in photos in the wild line-up with studies done in

biology and nature. En route to Nature, we are summarizing the research in kinship recognition as a survey submitted to IJCAI. This will be supported with an open-source Python project for the research community to learn from, start on, build upon. Thus, the models trained for the benchmarks for the survey will be done using well developed code that will be released to the public, while producing the models needed for the nature-based study. With that, the complete package pertaining to visual kinship recognition (i.e., the database, research papers, survey, open-source Python API, and the nature-base study).

[Phase 5 (Final)]: [July-August 2020]. Complete thesis (i.e., update and finalize draft submitted as part of the dissertation proposal). The same for the slides presented as part of the proposal.

References

- [1] **Joseph P. Robinson** and Yun Fu. Pre-trained d-cnn models for detecting complex events in unconstrained videos. In *Sensing and Analysis Technologies for Biomedical and Cognitive Applications 2016*, volume 9871, page 987100. International Society for Optics and Photonics, 2016.
- [2] **Joseph P. Robinson**, Yuncheng Li, Ning Zhang, Yun Fu, and Sergey Tulyakov. Laplace landmark localization. *International Conference on Computer Vision (ICCV)*, 2019.
- [3] **Joseph P. Robinson**, E Scott, and Yun Fu. Trecvid 2015: Multimedia event detection by pre-trained cnn models. National Institute of Standards and Technology (NIST).
- [4] **Joseph P. Robinson**, Ming Shao, Yue Wu, and Yun Fu. Families in the wild (fiw): Large-scale kinship image database and benchmarks. In *ACM Conference on Multimedia (ACMMM)*, pages 242–246, 2016.
- [5] **Joseph P. Robinson**, Ming Shao, Yue Wu, Hongfu Liu, Timothy Gillis, and Yun Fu. Visual kinship recognition of families in the wild. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2018.
- [6] **Joseph P. Robinson**, Ming Shao, Handong Zhao, Yue Wu, Timothy Gillis, and Yun Fu. Recognizing families in the wild (rfiw): Data challenge workshop. In *ACM Conference on Multimedia (ACMMM)*, 2017.
- [7] Lichen Wang, Bin Sun, **Joseph P. Robinson**, and Yun Fu. Ev-action: Electromyography-vision multi-modal action dataset. *International Conference on Automatic Face and Gesture Recognition*, 2020.
- [8] Shuyang* Wang, **Joseph P. Robinson***, and Yun (*EQUAL CONTRIBUTION) Fu. Kinship verification on families in the wild with marginalized denoising metric learning. In *International Conference on Automatic Face and Gesture Recognition*, pages 216–221. IEEE, 2017.
- [9] Yue Wu, Zhengming Ding, Liu, **Joseph P. Robinson**, and Yun Fu. Kinship classification through latent adaptive subspace. In *International Conference on Automatic Face and Gesture Recognition*. IEEE, 2018.
- [10] Yu Yin, Songyao Jiang, **Joseph P. Robinson**, and Yun Fu. Dual-attention gan for large-pose face frontalization. In *International Conference on Automatic Face and Gesture Recognition*, 2020.
- [11] Yu Yin, **Joseph P. Robinson**, Yulun Zhang, and Yun Fu. Joint super-resolution and alignment of tiny faces. In *Conference on Artificial Intelligence (AAAI)*, 2020.