

Education

- 2019–Present **PhD, Multimedia Lab (MMLab)**, *Chinese University of Hong Kong (CUHK)*, HongKong, China.
Ph.D. student advised by Prof. Xiaogang Wang and Prof. Hongsheng Li.
- 2015–2017 **MS, Master in Electrical Engineering**, *BarcelonaTech School of Telecommunications Engineering (ETSETB)*, Barcelona, Spain.
- 2011–2015 **BS, Engineer's degree in Electrical Engineering**, *BarcelonaTech School of Telecommunications Engineering (ETSETB)*, Barcelona, Spain.

Publications

- 2022 **J. Pan**, Z. Lin, X. Zhu, Y. Wang, Y. Qiao, H. Li, Retrieving-to-Answer: Zero-Shot Video Question Answering via Training-free Adaption of Foundation Models. *In submission*
- 2022 Y. Wang, K. Li, Y. Li, Y. He, B. Huang, Z. Zhao, H. Zhang, J. Xu, Y. Liu, Z. Wang, S. Xing, G. Chen, **J. Pan**, J. Yu, Y. Wang, L. Wang, Y. Qiao, InternVideo: General Video Foundation Models via Generative and Discriminative Learning. *Technical Report*
- 2022 **J. Pan***, Z. Lin*, X. Zhu, J. Shao, H. Li, ST-Adapter: Parameter-Efficient Image-to-Video Transfer Learning for Action Recognition. *NeurIPS, 2022* (* indicates equal contribution.)
- 2022 **J. Pan**, A. Bulat, F. Tan, X. Zhu, L. Dudziak, H. Li, G. Tzimiropoulos and B. Martinez, EdgeViTs: Competing Light-weight CNNs on Mobile Devices with Vision Transformers. *ECCV, 2022*
- 2021 **J. Pan***, S. Chen*, Z. Shou, J. Shao and H. Li, Actor-Context-Actor Relation Network for Spatio-Temporal Action Localization. *CVPR, 2021*. **(1st place solution of ActivityNet Challenge 2020 – AVA-Kinetics Task)**
- 2020 L. Sheng*, **J. Pan***, J. Guo, J. Shao, C. C. Loy, High-Quality Video Generation from Static Structural Annotations. *IJCV* (* indicates equal contribution.)
- 2019 **J. Pan**, C. Wang, X. Jia, J. Shao, L. Sheng, J. Yan and X. G. Wang, Video generation from Single Semantic Label Map. *CVPR, 2019*.
- 2018 **J. Pan***, Z. Shou*, J. Chan, K. Miyazawa, H. Mansour, A. Vetro, X. Giro-i-Nieto, S. F. Chang. Online Detection of Action Start in Untrimmed, Streaming Video. *ECCV, 2018*.
- 2017 **J. Pan**, C. Canton, K. McGuinness, N. E. O'Connor, J. Torres, E. Sayrol, X. Giro-i-Nieto. SalGAN: Visual Saliency Prediction with Generative Adversarial Networks. *CVPR Scene Understanding Workshop (CVPR SUNw), 2017 (Spotlight)*
- 2016 **J. Pan***, K. McGuinness*, N. E. O'Connor, E. Sayrol, X. Giro-i-Nieto. Shallow and Deep Convolutional Networks for Saliency Prediction. *CVPR, 2016*
- 2015 **J. Pan**, X. Giro-i-Nieto. End-to-end convolutional network for saliency prediction. *CVPR Large Scale scene Understanding Workshop (LSUN)*

Research Experience

- Oct 2021–Feb 2022 **Research Intern**, *Samsung AI Center*, Cambridge, UK.
- 2022 - Work on efficient mobile vision transformer. (Work published in ECCV'22)
- We propose a family of lightweight ViTs from the practical on-device deployment perspective.
- We build a benchmark of the on-device performance of mainstream ViTs.

- Mar **Research Assistant**, *DVMM Lab at Columbia University*, New York, NY.
- 2017–Dec - Work on Video Modelling and Prediction tasks.
- 2017 - Propose novel methods for Online Action Detection that is to detect action in a timely manner and to recognize its action category. (Work published in ECCV'18)
- I mainly contributed in the core idea for the detection model and its implementation.
- Feb 2015–Feb 2017 **Research Assistant**, *Image Processing Group, Univesitat Politecnica de Catalunya*, Barcelona, Spain.
- Research on saliency prediction and generative models under supervision of Professor Xavier Giro-i-Nieto, Elisa Sayrol, Cristian Canton (Facebook, Seattle, WA) and Kevin McGuinness (Insight Centre DCU, Dublin, Ireland)
- We proposed a winning solution for LSUN Challenge in CVPR'16 which was lately published in CVPR'17
- We investigated the contribution of Adversarial loss in Saliency Prediction task. The model is now one of the most popular saliency prediction model (SalGAN).

Teaching Experience

- July 2016 **Teaching Assistant at Deep Learning for Computer Vision Summer Seminar**, *BarcelonaTech School of Telecommunications Engineering (ETSETB)*, Barcelona, Spain.
- Instructors Xavier Giro-i-Nieto, Elisa Sayrol, Jordi Torres, Kevin McGuinness, Amaia Salvador, Eva Mohedano
- March–June 2016 **Teaching Assistant at Introduction to Digital Communication**, *BarcelonaTech School of Telecommunications Engineering (ETSETB)*, Barcelona, Spain.
- Instructor Margarita Cabrera

Awards

- 2022 1st Place in Short Term Object Interaction Anticipation task of Ego4D Challenge 2022.
- 2020 1st Place in AVA-Kinetics Crossover task of ActivityNet Challenge 2020.
- 2020 1st Place in Trimmed Activity Recognition task of ActivityNet Challenge 2020.
- 2015 1st Place in Saliency Prediction task of CVPR 2015 LSUN Challenge.
- 2015 Outstanding academic achievement recognition (6.37%)
- 2013 2nd Better Academic Record (2/361)
- 2012 3rd Better Academic Record (3/361)

Professional Service

Reviewer

- Computer Vision and Pattern Recognition (CVPR 2019, 2020, 2021, 2022)
- Conference on Neural Information Processing Systems (NeurIPS 2022)
- International Conference on Computer Vision (ICCV 2019, 2021)
- European Conference on Computer Vision (ECCV 2020)
- International Journal of Computer Vision (IJCV)
- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)