Heng Wang hengwang00@gmail.com
Staff Research Scientist, Facebook AI Research https://hengcv.github.io/

Summary

- 12+ years of combined research and engineering experience with deep expertise in computer vision, deep learning, machine learning, *etc.*
- Lead and build teams, mentor and supervise researchers, engineers and interns; drive product initiatives for content understanding and content creation, solve real-world computer vision problems.
- Invented the most effective and widely used video features named "Dense Trajectories" for video understanding, and won multiple international competitions and challenges.
- Proven track record of publications on top computer vision conferences/journals with over 10,000 citations.

Professional Experience

- **Facebook AI Research** Menlo Park, CA
  Staff Research Scientist *April 2017 - Now*
  - Lead and drive the research efforts on video understanding across Facebook. Invent next generation deep learning technologies for video understanding. Make impact in both research community and industry.
  - Transfer the latest research results to different product applications, such as content understanding (FB, IG, Ads), integrity (detect violence, nudity, hate speech, misinformation), infra (video search/retrieval, deduplication/copy detection), multi-modality (text, speech, audio, video), *etc.*
  - Tech lead projects, drive major technical bets and define future directions of video applications. Mentor and supervise researchers and engineers.
  - Lead open source projects for video understanding, increase the adoption of deep learning frameworks such as Caffe2 and PyTorch.
- **Amazon Go** Seattle, WA
  Research Scientist *November 2014 - April 2017*
  - Early member of the research team which develops complex computer vision systems that can replace human cashiers in a new form of grocery stores named "Amazon Go".
  - Design and implement the "Just Walk Out" feature using a network of cameras for Amazon Go. Push the boundary of computer vision and machine learning systems to achieve human-level accuracy for automatically understanding human behavior during shopping.
  - Lead the efforts on RGB-D cameras, 3D point cloud construction, human detection/tracking/ReID, action recognition/gesture association, *etc.*
  - Lead and collaborate with hardware/firmware/software engineers and product managers, and deliver results with limited time and resource.
- **LEAR Team, INRIA Rhône-Alpes** Grenoble, France
  Postdoc Researcher *July 2012 - April 2014*
  - **Advisor**: Cordelia Schmid
  - Improved the "dense trajectories" features to better handle camera motion. Published the **1st** mostly cited paper of ICCV 2013.
  - Won two major video classification competitions: THUMOS action recognition challenge 2013 and TRECVID Multimedia Event Detection challenge 2013.

  - Code available at: `http://lear.inrialpes.fr/people/wang/improved_trajectories`, which is the best video features and generates the state of the art results for video understanding.

- **LEAR Team, INRIA Rhône-Alpes** Grenoble, France
  Research Intern *March - December 2010*
  - **Advisor**: Cordelia Schmid
  - Invented the "dense trajectories" features and achieved groundbreaking results in action recognition. On the popular HMDB51 dataset, the accuracy was improved from 26.9% to 48.3%.
  - The original paper is the **2nd** mostly cited among all 438 papers of CVPR 2011.
  - Code available at: `http://lear.inrialpes.fr/people/wang/dense_trajectories`, which is among the most widely used video features for action recognition.

- **LEAR Team, INRIA Rhône-Alpes** Grenoble, France
  Research Intern *February - August 2009*
  - **Advisor**: Cordelia Schmid
  - First extensive evaluation of different video feature detectors/descriptors and their combinations for action recognition.
  - Proposed "dense sampling" instead of sparse feature detector and demonstrated its superior performance.
  - Established a new state of the art and published a paper in BMVC 2009, which is highly considered as a standard baseline for comparison and receiving over 1500 citations.

Education

- **Chinese Academy of Sciences** Beijing, China
  National Laboratory of Pattern Recognition *2006 - 2012*
  - **PhD** in Pattern Recognition and Intelligent Systems
  - **Advisor**: Cheng-Lin Liu & Cordelia Schmid
  - **Thesis**: Human Tracking and Action Recognition in Video

- **Harbin Institute of Technology** Harbin, China
  School of Electrical Engineering and Automation *2002 - 2006*
  - **BSc** in Electrical Engineering

Awards

Winner of action recognition, THUMOS workshop with ICCV 2013
Winner of TRECVID Multimedia Event Detection . . . . . 2012 and 2013
PanDeng Scholarship, Chinese Academy of Sciences 2011
1st Prize, China Undergraduate Mathematical Contest in Modeling 2004

Professional Services

- Area Chair: BMVC'21
- Conference Reviewer: CVPR'13-21, ICCV'13-19, ECCV'14-20, BMVC'17, ICPR'12.
- Journal Reviewer: T-PAMI, IJCV, T-IP, CVIU, T-NNLS, PR, T-CSVT, IVC, PRL, SPL, *etc.*
- PhD Thesis Examiner & Research Grant Reviewer.

Open Source

- **PyTorchVideo** (A deep learning library for video understanding research): `https://pytorchvideo.org/`
- **VMZ** (A Caffe2 library for video classification): `https://github.com/facebookresearch/VMZ`
- **iDT** (Best hand-crafted video feature): `http://lear.inrialpes.fr/people/wang/improved_trajectories`
- **Dense Trajectories** (SOTA video feature): `http://lear.inrialpes.fr/people/wang/dense_trajectories`

Selected Publications

Full publication list: Semantic Scholar, Google Scholar.

- W. Wang, M. Feiszli, **H. Wang**, D. Tran. Unidentified Video Objects: A Benchmark for Dense, Open-World Segmentation. ICCV. 2021
- X. Gong, **H. Wang**, Z. Shou, M. Feiszli, Z. Wang, Z. Yan. Searching for Two-Stream Models in Multivariate Space for Video Recognition. ICCV. 2021.
- X. Wang, L. Zhu, **H. Wang**, Y. Yang. Interactive Prototype Learning for Egocentric Action Recognition. ICCV. 2021.
- G. Bertasius, **H. Wang**, L. Torresani. Is Space-Time Attention All You Need for Video Understanding? ICML. 2021.
- X. Yang, H. Fan, L. Torresani, L. Davis, **H. Wang**. Beyond Short Clips: End-to-End Video-Level Learning with Collaborative Memories. CVPR. 2021.
- Y.-T. Hu, **H. Wang**, N. Ballas, K. Grauman, A. Schwing. Proposal based Video Completion. ECCV, 2020.
- **H. Wang**, D. Tran, L. Torresani, M. Feiszli. Video Modeling with Correlation Networks. CVPR, 2020.
- L. Zhu, L. Sevilla-Lara, D. Tran, M. Feiszli, Y. Yang, **H. Wang**. FASTER Recurrent Networks for Efficient Video Classification. AAAI, 2020.
- D. Tran, **H. Wang**, L. Torresani, M. Feiszli. Video Classification with Channel-Separated Convolutional Networks. ICCV, 2019.
- D. Ghadiyaram, M. Feiszli, D. Tran, X. Yan, **H. Wang**, D. Mahajan. Large-scale Weakly-supervised Pre-training for Video Action Recognition. CVPR, 2019.
- J. Ray, **H. Wang**, D. Tran, Y. Wang, M. Feiszli, L. Torresani, M. Paluri. Scenes-Objects-Actions: A Multi-Task, Multi-Label Video Dataset. ECCV, 2018.
- D. Tran, **H. Wang**, L. Torresani, J. Ray, Y. LeCun, M. Paluri. A Closer Look at Spatiotemporal Convolutions for Action Recognition. CVPR, 2018. **(702 citations)**
- **H. Wang**, D. Oneata, J. Verbeek, C. Schmid. A Robust and Efficient Video Representation for Action Recognition. IJCV, 2015. **(256 citations)**
- **H. Wang**, C. Schmid. Action Recognition with Improved Trajectories. ICCV, 2013. **1st mostly cited paper (2745 citations)**

- **H. Wang**, A. Kläser, C. Schmid, C.-L. Liu. Dense Trajectories and Motion Boundary Descriptors for Action Recognition. IJCV, 2013. **(1537 citations)**
- **H. Wang**, A. Kläser, C. Schmid, C.-L. Liu. Action Recognition by Dense Trajectories. CVPR, 2011. **2nd mostly cited paper (2351 citations)**
- **H. Wang**, M. M. Ullah, A. Kläser, I. Laptev, C. Schmid. Evaluation of Local Spatio-temporal Features for Action Recognition. BMVC, 2009. **1st mostly cited paper (1585 citations)**

Technical skills

- C/C++, Python, Linux(bash), OpenCV, Caffe/Caffe2, PyTorch, CUDA, Matlab, OpenMP, Lapack, *etc.*