# **Baifeng Shi**

baifeng\_shi@berkeley.edu | (+1) 510-495-7418 https://bfshi.github.io/

## **EDUCATION BACKGROUND**

University of California, Berkeley	
Ph.D. student, Computer Science	08/2021 - now
Advisor: Prof. Trevor Darrell	
Research Field: Computer Vision and Robotics	
Awards and Honors:	
BAIR Ignition Reward, UC Berkeley	09/2021
Peking University	
B.S., Computer Science	09/2017 - 06/2021
Awards and Honors:	
Gold Medal (3 / 360), Chinese Physics Olympiad final contest	10/2016
EECS Dean Scholarship, Peking University	09/2017
Merit Student, Peking University	09/2018 & 09/2020

### **RESEARCH INTERNSHIP**

• NVIDIA Research, Research Intern Mentor: Dr. Hongxu (Danny) Yin & Dr. Pavlo Molchanov 03/2024 - now

#### **CONFERENCE PUBLICATIONS**

- Baifeng Shi, Boyi Li, Han Cai, Yao Lu, Sifei Liu, Marco Pavone, Jan Kautz, Song Han, Trevor Darrell, Pavlo Molchanov, Hongxu Yin, *Scaling Vision Pre-Training to 4K Resolution*, CVPR 2025, conference highlight
- Zhijian Liu, Ligeng Zhu, Baifeng Shi, ..., Hongxu Yin, Song Han, Yao Lu, NVILA: Efficient Frontier Visual Language Models, CVPR 2025
- Ilija Radosavovic, Bike Zhang, **Baifeng Shi**, Jathushan Rajasegaran, Sarthak Kamat, Trevor Darrell, Koushil Sreenath, Jitendra Malik, *Humanoid Locomotion as Next Token Prediction*, NeurIPS 2024, spotlight
- Dantong Niu, Yuvan Sharma, Giscard Biamby, Jerome Quenum, Yutong Bai, **Baifeng Shi**, Trevor Darrell, Roei Herzig, *LLARVA: Vision-Action Instruction Tuning Enhances Robot Learning*, **CoRL 2024**
- Baifeng Shi, Ziyang Wu, Maolin Mao, Xin Wang, Trevor Darrell, *When Do We Not Need Larger Vision Models?*, ECCV 2024
- Jiaxin Ge, Sanjay Subramanian, Baifeng Shi, Roei Herzig, Trevor Darrell, *Recursive Visual Programming*, ECCV 2024
- Ilija Radosavovic, **Baifeng Shi**, Letian Fu, Ken Goldberg, Trevor Darrell<sup>\*</sup>, Jitendra Malik<sup>\*</sup>, *Robot Learning with Sensorimotor Pre-training*, **CoRL 2023, oral presentation**
- Long Lian<sup>\*</sup>, Baifeng Shi<sup>\*</sup>, Adam Yala, Trevor Darrell, Boyi Li, *LLM-Grounded Video Diffusion Models*, ICLR 2024
- Baifeng Shi, Trevor Darrell, Xin Wang, *Top-down Visual Attention from Analysis by Synthesis*, CVPR 2023, conference highlight
- Baifeng Shi, Yale Song, Neel Joshi, Trevor Darrell, Xin Wang, Visual Attention Emerges from Recurrent Sparse Reconstruction, ICML 2022
- Baifeng Shi, Qi Dai, Judy Hoffman, Kate Saenko, Trevor Darrell, Huijuan Xu, *Temporal Action Detection with Multi-level Supervision*, ICCV 2021
- Baifeng Shi, Judy Hoffman, Kate Saenko, Trevor Darrell, Huijuan Xu, *Auxiliary Task Reweighting for Minimum-data Learning*, NeurIPS 2020
- Zhekun Luo, Devin Guillory, **Baifeng Shi**, Wei Ke, Fang Wan, Trevor Darrell, Huijuan Xu, *Weakly-Supervised* Action Localization with Expectation-Maximization Multi-Instance Learning, ECCV 2020
- Baifeng Shi<sup>\*</sup>, Dinghuai Zhang<sup>\*</sup>, Qi Dai, Zhanxing Zhu, Yadong Mu, Jingdong Wang, Informative Dropout for Robust Representation Learning: A Shape-bias Perspective, ICML 2020
- Baifeng Shi, Qi Dai, Jingdong Wang, Yadong Mu, *Weakly-Supervised Action Localization by Generative Attention Modeling*, CVPR 2020

#### JOURNAL PUBLICATIONS

• Letian Fu, Long Lian, Renhao Wang, **Baifeng Shi**, Xudong Wang, Adam Yala, Trevor Darrell, Alexei A Efros, Ken Goldberg, *Rethinking Patch Dependence for Masked Autoencoders*, **TMLR**, 2024

#### PREPRINTS

• Baifeng Shi, Siyu Gai, Trevor Darrell, Xin Wang, TOAST: Transfer Learning via Attention Steering, 2023

## **INVITED TALKS**

- Scaling Vision Pre-Training to 4K Resolution, Boston University, Apr 2025
- Scaling Vision Pre-Training to 4K Resolution, Princeton University, Apr 2025
- Scaling Vision Pre-Training to 4K Resolution, Google Deepmind, Apr 2025
- Scaling Up Visual Pre-Training, What's Next?, AI Tea Talk Singapore, Jun 2024
- Scaling Up Visual Pre-Training, What's Next?, VGG Group, University of Oxford, Apr 2024
- Scaling Up Visual Pre-Training, What's Next?, Prof. Yi Ma's group, UC Berkeley, Mar 2024
- Principles and Applications of Bottom-Up and Top-Down Visual Attention, Peking University, Oct 2023
- Principles and Applications of Bottom-Up and Top-Down Visual Attention, TechBeat, Jun 2023