Dr. Zhenyu Wu

Personal Website: https://wuzhenyusjtu.github.io/
Email: wuzhenyusjtu@gmail.com
Phone: +1(614)-721-9374
Google Scholar: *XOjjlREAAAAJ*.

EDUCATION

Texas A&M Unversity, College Station, Texas, the United States 08/2017-09/2021
Doctor of Philosophy in Computer Science, Advised by Prof. *Zhangyang (Atlas) Wang*

The Ohio State University, Columbus, Ohio, the United States 08/2015-05/2017
Master of Science in Computer Science, Advised by Prof. *Han-Wei Shen*

Shanghai Jiao Tong University, Shanghai, China 09/2011-06/2015
Bachelor of Engineering in Information Security, Advised by Prof. *Cunqing Hua*

RESEARCH INTERESTS

1. Building practical perception systems that are: (a) **robust** to perturbations/degradations and distribution shifts[**C1,C4,C9**]; (b) **efficient** for both inference and training[**C5,C6,C13,J1**]; (c) **ethical** for individual privacy and group fairness[**C7,C12,J3,J4**].

2. Solving **low-level** and **high-level** vision problems that include: (a) dehazing[**J2**] and image enhancement[**C10**]; (b) object detection[**C9**] and action detection[**C1,C2,C3**]; (c) text spotting[**C6**]; (d) pose estimation[**C7**]; (e) tracking[**C4**].

PUBLICATION

[C1] Y. Zhai, Z. Liu, **Z. Wu**, Y. Wu, C. Zhou, D. Doermann, J. Yuan, and G. Hua, *"SOAR: Scene-debiasing Open-set Action Recognition"*, **ICCV**, 2023.

[C2] K. Kahatapitiya, Z. Ren, H. Li, **Z. Wu**, M. Ryoo, and G. Hua, *"Self-supervised Pretraining with Classification Labels for Temporal Activity Detection"*, **AAAI**, 2023.

[C3] **Z. Wu**, Z. Ren, Y. Wu, Z. Wang, and G. Hua, *"TxVAD: Improved Video Action Detection by Transformers"*, ACM **MM**, 2022.

[C4] X. Hu, **Z. Wu**, H. Miao, S. Fan, T. Long, Z. Hu, P. Pi, Y. Wu, Z. Ren, Z. Wang, and G. Hua, *"E^2TAD: An Energy-Efficient Tracking-based Action Detector"*, Arxiv Preprint, 2022.

[C5] X. Hu, ..., **Z. Wu**, ... et. al. *"The 2020 Low-Power Computer Vision Challenge"*, **AICAS**, 2021.

[C6] Z. Hu, P. Pi, **Z. Wu**, Y. Xue, J. Shen, J. Tan, X. Lian, Z. Wang, and J. Liu, *"E^2VTS: Energy-Efficient Video Text Spotting From Unmanned Aerial Vehicles"*, **CVPR UG2+ Workshop**, 2021.

[C7] **Z. Wu***, S. Hoang*, S. Lin, Y. Xie, W. Fan, Y. Lin, and Z. Wang, *"3D-Aware Multi-modal Guided Hand Generation for 3D Hand Pose Synthesis"*, ACM **MM**, 2020.

[C8] Y. Zheng, **Z. Wu**, Y. Yuan, T. Chen, and Z. Wang, *"PCAL: A Privacy-Preserving Intelligent Credit Risk Modeling Framework based on Adversarial Learning"*, Arxiv Preprint, 2020.

[C9] **Z. Wu**, K. Suresh, P. Narayanan, H. Xu, H. Kwon, and Z. Wang , *"Delving into Robust Object Detection from Unmanned Aerial Vehicles: A Deep Nuisance Disentanglement Approach"*, **ICCV**, 2019.

[C10] P. Uplavikar, **Z. Wu**, and Z. Wang, *"All-In-One Underwater Image Enhancement using Domain-Adversarial Learning"*, **CVPR UG2+ Workshop**, 2019.

[C11] P. Narayanan, **Z. Wu**, H. Kwon, Z. Wang, and R. Rao, *"Overview of Machine Learning-based Perception Algorithms for Unstructured and Degraded Visual Environments"*, Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications, 2019.

[C12] **Z. Wu**, Z. Wang, Z. Wang, and H. Jin, *"Towards Privacy-Preserving Visual Recognition via Adversarial Training: A Pilot Study"*, **ECCV**, 2018.

[C13] J. Wu, Y. Wang, **Z. Wu**, Z. Wang, A. Veeraraghavan, and Y. Lin, *"Deep k-Means: Re-Training and Parameter Sharing with Harder Cluster Assignments for Compressing Deep Convolutions"*, **ICML**, 2018.

[J1] X. Hu, Z. Jiao, A. Kocher, **Z. Wu**, J. Liu , J. Davis , G. Thiruvathukal, and Y. Lu *"Evolution of Winning Solutions in the 2021 Low-Power Computer Vision Challenge"*, IEEE **Computer**, 2023.

[J2] P. Narayanan, X. Hu*, **Z. Wu***, M. Thielke, J. Rogers, A. Harrison, J. Agostino, J. Brown, L. Quang, J. Uplinger, H. Kwon, and Z. Wang, *"A Multi-purpose Real Haze Benchmark with Quantifiable Haze Levels and Ground Truth"*, IEEE **TIP**, 2023.

[J3] **Z. Wu**, Z. Wang, Y. Yuan, J. Zhang, Z. Wang, and H. Jin, *"Black-Box Diagnosis and Calibration on GAN Intra-Mode Collapse: A Pilot Study"*, ACM **TOMM**, 2021.

[J4] **Z. Wu***, H. Wang*, Z. Wang, Z. Wang, and H. Jin, *"Privacy-Preserving Deep Action Recognition: An Adversarial Learning Framework and A New Dataset"*, IEEE **TPAMI**, 2020.

PROFESSIONAL EXPERIENCE

Researcher @ Wormpex AI Research, Seattle, WA 10/2021-Present

Report to Dr. Zhou Ren, Dr. Yi Wu, and Dr. Gang Hua

1. Efficient Multi-Task Integration on Edge Devices

- Model: designed an ultra-efficient multi-task model for one-stage detection and pose estimation
- Data: Built an OpenCV-based annotation tool for multi-objects' bounding box and pose keypoints with pseudo labels
- Deployment: deployed the multi-task model on edge devices using TensorRT (250 fps) and RKNN-toolkit (40 fps)
- Accuracy: combined heatmap-based and regression-based in one-stage way for better pose estimation
- Accuracy: collected detection and pose estimation related benchmarks for better model pretraining
- Robustness: designed a hard-sample mining strategy to acquire out-of-distribution samples
- Efficiency: proposed a multi-scale branch pruning strategy for acceleration ($\times$1.25 faster)
- Efficiency: implemented an ultra-efficient post-processing module in C++ ($\times$250 faster over Python)
- Efficiency: proposed a hardware-agnostic quantization workflow that uses both Quantization-Aware Training (QAT on multi-gpu training device) and Post-Training Quantization (PTQ on target edge device) to achieve minimal accuracy loss ($<$ 1 mAP), with de-quantization before PTQ and after QAT

2. Product Display Check by Text Segmentation

- Proposed a text segmentation-based "display check" model for products on the shelves
- Encapsulated the "display check" model as an http service in Docker
- Collect datasets from retail business for training and evaluation

3. E^{2}TAD: *An Energy-Efficient Tracking-based Action Detector* [**C3, J1, ICCV-21 LPCV 1st Prize**]

- Duty: served as the team leader (8 grad students) and main contributor
- Award: won the **1st Place out of** 31 **Teams** at *ICCV-21 Low-Power Computer Vision Challenge*
- System design: designed a tracking-based spatio-temporal video action detector
- Robustness: presented a view-invariant and occlusion-robust person Re-IDentification (ReID) model
- Efficiency: reduced the parameter storage and energy consumption of ReID by 40$\times$ and 60$\times$ respectively, via data-driven adaptive pruning and quantization-aware training with minor accuracy loss
- Efficiency: proposed two lightweight YOLO variants for human detection achieving 10 FPS on Raspberry Pi 3B+
- Efficiency: proposed a cache-friendly computation pipeline that traded memory with less data movement to save energy
- Efficiency: proposed two dynamic inference strategies to save energy consumption

Research Assistant @ Texas A&M University, College Station, TX 08/2017-09/2021
Advised by Dr. Zhangyang Wang

1. Privacy-Preserving Action Recognition via Adversarial Learning [**ECCV-18**, **TPAMI-20**]

- Fulfilled the privacy-preserving purpose by applying learnable active degradation on image/video data in smart home setting
- Formulated a three-party game among the utility, the privacy budget and the degradation module
- Proposed novel training strategies, evaluation protocols, and visualization methods
- Collected a benchmark by annotating privacy-related attributes on existing action recognition dataset

2. E^2VTS: Energy-Efficient Video Text Spotting from UAV [**C4, C5, CVPR-20 LPCV 2nd Prize**]

- Duty: served as the team leader (6 grad students) and main contributor
- Award: won the **2nd Place out of** 11 **Teams** at *CVPR-20 Low-Power Computer Vision Challenge*
- System design: a two-step text spotting pipeline that revisited RCNN's crop & resize training strategy
- System design: EAST model for detection and CRNN model for recognition on low-resolution images
- Efficiency: a multi-stage image processor to select the highest-quality frame in a sliding window manner by rejecting text-free frames, cropping non-text regions, and rejecting out-of-distribution frames
- Efficiency: an LBP and Canny edge-based algorithm to identify the key frames and crop text regions to save energy consumption
- Efficiency: structured pruning and quantization-aware training to reduce both model size and computation cost with minimal accuracy loss
- Efficiency: used L1 filter, ADMM pruner etc. methods in NNI tool to prune both models by about 75 % parameters
- Efficiency: used dynamic/QAT quantization methods in PyTorch to quantize both models

Research Intern @ Wormpex AI Research, Seattle, WA 05/2020-08/2020
Advised by Dr. Zhou Ren, Dr. Yi Wu and Dr. Gang Hua

End-to-End Video Action Detection with Transformers [**MM-22**]

- Proposed a Transformer-based Paradigm to do action detection in an end-to-end way
- Proposed a Person Localization Transformer (PTx) and an Action Classification Transformer (ATx)
- Proposed a Hardness-Aware Curriculum Learning strategy for better convergence
- Applied efficient Transformer solutions for training and inference acceleration

Research Intern @ Tencent AI Lab, Palo Alto, CA 05/2019-08/2019
Advised by Dr. Shih-Yao Lin, Dr. Yusheng Xie and Dr. Wei Fan

Hand Synthesis from Pose and Style: Data Augmentation for 3D Hand Pose Estimation [**MM-20**]

- Defined the problem of hand synthesis from pose and style
- Proposed a cGAN-based style transfer approach to synthesize hands from conditioned pose and style
- Proposed a geometry-based curriculum training and inference with nearest-neighbor match by hand pose similarity.

Research Intern @ Adobe Research, San Jose, CA 01/2019-04/2019
Advised by Dr. Zhaowen Wang, Dr. Jianming Zhang and Dr. Hailin Jin

Visual Privacy Shredder: Machine Unlearning for Privacy Protection in Generative Models [**TOMM-21**]

- Investigated the memorization issue of generative models on training data
- Explored GAN intra-mode collapse and calibrated the collapse in a black-box setting: no access to training data or model parameters
- Devised a set of sampling-based tools that can visualize, quantify, and rectify intra-mode collapse.

Research Intern @ USC Institute for Creative Technologies, Los Angeles, CA 05/2018-08/2018
Advised by Dr. Priya Narayanan and Dr. Heesung Kwon

Object Detection in Low-Quality Drone Imagery [**ICCV-19**]

- Formulated an adversarial learning pipeline to improve the robustness of drone-based object detection
- Utilized the free attributes of flying altitude, viewing angle and weather condition to learn domain-robust features via an adversarial training framework

Research Intern @ Siemens PLM Software, Cincinnati, OH 05/2016-08/2016
Advised by Dr. Pengcheng Liu
Visual Recognition using Deep Learning
- Built a 5-layer-ConvNet to classify images generated from CAD software using TensorFlow
- Leveraged LSTM+CNN architecture to localize multiple objects of interest in one image

Research Intern @ Siemens PLM Software, Shanghai, China 07/2014-02/2015
Advised by Dr. George Allen (the Chief Scientist)
1. Modeling with Curved Triangles
- Worked on a curved triangle algorithm to give better results in graphical display
- Derived the Curved Triangle as a triangular Bezier patch from a flat triangle with 3 normals to 3 points
- Implemented the curved triangles using NXOpen libraries, and tested on different geometric models
- Application: tessellation refinement for display, 3D printing and faster model transmission

2. Code Editor by Roslyn (Microsoft Open Compiler Technologies)
- Improved the code editor component in NX (CAD software) using Roslyn Code Analysis technology
- Implemented an editor prototype supporting Indenting, Syntax Highlighting, Code Completion, Intellisense and Verbosity Cleaning
- Developed the editor as a Windows Forms application supporting both Visual Basic and C# features

TEACHING EXPERIENCE

Teaching Assistant @ Texas A&M University, College Station, TX 08/2019-05/2020
- Instructor of CSCE 421: Machine Learning
- Prepared some course materials and served in some lectures

Teaching Assistant @ The Ohio State University, Columbus, OH 08/2016-12/2016
- Instructor of CSE 1223: Introduction to Programming in Java
- Prepared all the course materials and served in all the lectures

SERVICES & AWARDS

Challenge Awards
- 1st Prize of ICCV 2021 Low-Power Computer Vision Challenge: *Video Track*
- 2nd Prize of CVPR 2020 Low-Power Computer Vision Challenge: *Video Track*

Professional Service
- Workshop Organizing Committee: 5th and 6th UG2+ Workshop (CVPR 2022, CVPR 2023)
- Area Chair: ICIP 2022, 2023
- Conferences Reviewers: CVPR, ECCV, ICCV, NeurIPS, ICLR, ICML, AAAI, ICME, WACV, ICIP
- Journals Reviewers: IJCV, TPAMI, TIP, TCSVT, NCAA

Invited Talks
- Tutorial Talk: Efficient Computer Vision for Embedded Systems (DAC 2022)
- Tutorial Talk: Designing A Low-Power Video Recognition System (CVPR 2022: ECV Workshop)
- Winner's Talk: Low-Power Computer Vision (ICCV 2021: LPCVC Workshop)
- Winner's Talk: Low-Power Computer Vision (CVPR 2020: LPCVC Workshop)
- Winner's Talk: Chalearn LAP Inpainting Competition (ECCV 2018: LAPC Workshop)

Media Coverage
- Research @ Texas A&M: *New technique protects privacy from snoopers who hack smart devices*
- News of Texas A&M University Engineering: *New privacy filter for smart camera applications*
- News of Adobe Research: *Smart Home? Adversarial Machine Learning Could Protect Your Privacy*