# JUNYI ZHANG

junyi42zhang@gmail.com ◇ GitHub ◇ junyi42.com

## RESEARCH INTERESTS

My research interests mainly lie in **computer vision**, **deep generative models**, and **representation learning**. I am now focusing on the **application of diffusion models**, exploring its potential for **both generation and perception scenarios**. I also have a background in action recognition, self-supervised learning, and modalities such as video, 3D, human motion, and language.

## EDUCATION

**SEIEE, Shanghai Jiao Tong University** — Shanghai, China
Major in Computer Science and Technology; GPA: 3.92 (8/108) — Sept. 2020 - Present

**Zhiyuan College, Shanghai Jiao Tong University** — Shanghai, China
Major in Zhiyuan Honors Program of Engineering (Top 5%) — Sept. 2020 - Present

## SELECTED AWARDS

**Microsoft Research "Stars of Tomorrow" Certificate** (Top 10% interns) 2022
Huawei Fellowship (Top 1% in honors program) 2022
**National Scholarship** (Top 0.2% nationwide, highest honor in China) 2021
Outstanding Undergraduate Scholarship, A Level, SJTU (Top 1%) 2021

## PUBLICATIONS

**A Tale of Two Features: Stable Diffusion Complements DINO for Zero-Shot Semantic Correspondence**
Junyi Zhang, Charles Herrmann, Junhwa Hur, Luisa F. Polanía, Varun Jampani, Deqing Sun, Ming-Hsuan Yang. *In submission to NeurIPS*, 2023. (Received **all positive scores** with median rating of *Accept*)

**LayoutDiffusion: Improving Graphic Layout Generation by Discrete Diffusion Probabilistic Models**
Junyi Zhang, Jiaqi Guo, Shizhao Sun, Jian-Guang Lou, Dongmei Zhang. *In ICCV*, 2023.

**Bridging The Isolated Islands in Human Action Understanding**
Yong-Lu Li*, Xiaoqian Wu*, Xinpeng Liu, Yiming Dou, Yikun Ji, Junyi Zhang, Yixing Li, Xudong Lu, Jingru Tan, Cewu Lu. *Under review*, 2023.

**Mining Cross-Person Cues for Body-Part Interactiveness Learning in HOI Detection**
Xiaoqian Wu*, Yong-Lu Li*, Xinpeng Liu, Junyi Zhang, Yuzhe Wu, Cewu Lu. *In ECCV*, 2022.

## EXPERIENCE

**University of California, Merced - Vision Learning Lab** — *Merced, California*
Visiting Student – Supervisor: *Prof.* Ming-Hsuan Yang — June 2023 - Present

- **Diffusion Models for Semantic Correspondence**
  - Explore the internal representations of text-to-image diffusion models for semantic correspondence.
  - Discover the complementary of Stable Diffusion features and DINOv2 features, based on which achieve state-of-the-art performance under zero-shot setting on a variety of semantic correspondence benchmarks.
  - First-author paper submitted to **NeurIPS'23**. More details at project page.

**Microsoft Research Asia - Data, Knowledge and Intelligence Group** — *Beijing, China*
Research Intern – Mentor: *Dr.* Shizhao Sun and *Dr.* Jian-Guang Lou — July 2022 - Dec. 2022

- **Diffusion Models for Layout Generation, AI for Design Project**
  - Study the application of diffusion models in graphic design, especially the graphic layout generation.
  - Develop discrete diffusion models carefully tailored for layout data, achieve state-of-the-art performance in unconditional and conditional generation tasks for several public layout datasets.
  - First-author paper accepted by **ICCV'23**, with potential for product implementation.

**Shanghai Jiao Tong University - Machine Vision and Intelligence Group** *Shanghai, China*
Undergraduate Research Intern – Supervisor: *Prof.* Yong-Lu Li Oct. 2021 - Present

- **Diffusion Models for Interactive Motion Generation**
  - Design a general diffusion models framework for interactive motion generation, including Human-Object Interaction, Hand-Object Interaction, and Human-Human Interaction generation.
  - Introduce time-consistency and interaction-awareness to DM by decoupling the added Gaussian noise.
- **Unified Human Action Understanding Project**
  - Research on unifying multi-modal physical action spaces (2D, 3D, image, video, etc.) to a unified semantic space through the introduction of linguistic structure knowledge.
  - Design and conduct experiments to extend our method on video datasets and verify the performance.
- **Human Action Knowledge Engine Project**
  - Study on improving Human-Object Interaction (HOI) detection by incorporating human action knowledge.
  - Conduct experiments to examine the effect of image depth and multi-person interactions on HOI detection.
  - Research results were submitted as co-author in two papers, one of which was accepted by **ECCV'22**.

## SKILLS

- **Skills:** Python (Pytorch), C/C++, html, css, Linux, LaTeX, Assembly language
- **Languages:** English (TOEFL 105/120), Chinese (Native)