Zeren Jiang

Email: zeren.jiang99@gmail.com | Tel: (+41) 0794731854 | Zurich, Switzerland

EDUCATIONAL BACKGROUND

- **Beihang University** *09/2016 - 07/2020*
  Bachelor's Degree, Major: Software Engineering, Dual-Degree: Mathematics and Applied Mathematics
  Cumulative GPA: 3.85/4.00, Rank: 3/157
- **Hong Kong University of Science and Technology (HKUST)** *01/2019 - 05/2019*
  Exchange Program, Major: Computer Science and Engineering
  Cumulative GPA: 3.9/4.3
- **Eidgenössische Technische Hochschule Zürich (ETH Zurich)** *09/2021 - Present*
  Master Degree Candidate, Major: Computer Science (Machine Intelligence)
  Cumulative GPA: 5.83/6.00

PUBLICATION

Xing Dai*, **Zeren Jiang***, et al. "General Instance Distillation for Object Detection", (*co-first author in alphabetical order)
2021 Internaltional Conference on Computer Vision and Pattern Recognition (CVPR 2021)**, Accepted**

Shangzhe Di*, **Zeren Jiang***, et al. "Video Background Music Generation with Controllable Music Transformer",
2021 ACM International Conference on Multimedia (ACM MM 2021)**, Accepted** (*co-first author) **Best Paper Award**

Luting Wang, Xiaojie Li, Yue Liao, **Zeren Jiang**, et al. "HEAD: HEtero-Assists Distillation for Online Knowledge Transfer among Heterogeneous Object Detectors", 2022 European Conference on Computer Vision (ECCV 2022)**, Accepted**

Zeren Jiang*, Chen Guo*, et al. "MultiPly: Reconstruction of Multiple People from Monocular Video in the Wild",
2024 Internaltional Conference on Computer Vision and Pattern Recognition (CVPR 2024)**, Accepted** (*co-first author)

RESEARCH EXPERIENCE

Multi-person reconstruction from monocular video in the wild *04/2023 - 01/2024*
ETHz Semester Project *Advisor: Prof. Otmar Hilliges*

- Proposed the first approach based on layered VolSDF for the multi-person reconstruction task from monocular video.
- Designed a confidence-guided optimization to alternatively optimize the human pose and shape/texture network based on the pose confidence. This strategy allows the gradual construction of a canonical human model from frames with reliable pose, while alleviate the negative influence of frames with unreliable pose.
- Combined the self-supervised segmentation in 3D and the promptable 2D segmentation module to achieve clean separations between people even under close interaction. Specifically, we progressively update the instance segmentation masks from SAM by leveraging our progressively evolving human surfaces in deformed space as input prompts.
- My related co-first-authored paper was **accepted by CVPR 2024**

Automatic Video Background Music Generation with Controllable Music Transformers *11/2020 - 04/2021*
Dept. of Computer Science and Engineering, Beihang University *Advisor: Prof. Si Liu*

- Used linear transformer to generate a background music in MIDI format for a given video.
- Proposed a new representation of MIDI music based on compound words, including note density and strength of simu-notes as additional features.
- By connecting the optical flow and visual beats from the given video with density and strength of music respectively, together with user-defined music genre, we are able to control the music generation process.
- My co-first-authored paper was **accepted by ACM MM 2021** and got **best paper award** (1 over 542 accepted papers).

Knowledge Distillation for Object Detection *11/2019 - 09/2021*
Face ++, Beijing Megvii Co., Ltd. | *Face Detection Team* | *Research intern*

- Designed a novel knowledge distillation scheme for detection model, which adaptively selects the most discriminative patch to distill and takes advantage of feature-based, relation-based, and response-based knowledge for distillation. My related co-first authored paper was **accepted by CVPR 2021.**
- Applied our method to the actual face detection business scenario, the AP of student models increased by 2% on average (tested on Megvii Crowdv1 dataset). For experiments in public dataset MSCOCO, the student model (Retinanet-Res50) was improved from 36.2% mAP to 39.1% mAP and even surpassed the teacher model (Retinanet-Res101) with 38.1% mAP by a large margin.
- Explored the relationship between teacher and student model and designed a non-convergence teacher for knowledge distillation which is proven to be effective on both classification and detection tasks.

COURSE PROJECTS

Capturing Multi-person Interactions from Videos *02/2023 - 05/2023*

ETHz Digital Human Course Project

- Investigated the SOTA multi-person multi-view 3D key point estimation pipeline, including MVPose & MVMP
- Integrated traditional (OneEuro Filter) and neural-based smoothing methods (SmoothNet) as a motion prior, achieved lower MPJPE and PAMPJPE.
- Fitted a parametric human model LISST to the estimated 3D key point.

Point-based 3D Object Detection *04/2022 - 07/2022*

ETHz Deep Learning for Autonomous Driving Course Project

- Implemented different 3D target representations based on geometry and scaling. Achieved remarkable improvements with residual prediction on displacement field with canonical transformation.
- Investigated and experimented task-specifc regression and classification loss functions (encoder loss, corner loss and bin-based loss, IOU loss) to improve the performance on accuracy, robustness and small objects.
- Introduced feature fusion module to enhance original point features with embedded low-dimensional data. Used positional encoding function for absolute position embedding.
- Ranked Top 5 among 40 teams.

Human body estimation by RGB image *02/2022 - 06/2022*

ETHz Machine Perception Course Project

- Estimated human body pose and shape given RGB images centered around humans. Predicted the joint rotations and shape coefficients of statistical body model SMPL.
- Improved the SOTA method SPIN by investigating different backbones for feature extracting and adopting test-time data augmentation.
- Ranked Top 3 among 30 teams.

Fast Implementation of Triangle Listing Algorithm *02/2022 - 06/2022*

ETHz Advanced Systems Lab Course Project

- Analyzed the characteristics of three classic triangle listing algorithms and proposed optimizations, including branch elimination, blocking, unrolling, and vectorization.
- Validated the theoretical analysis and show consistent improvement of our optimization with different generated and real-world graphs on Intel and ARM processors.
- Achieved up to ten times speedup and performance improvement.

Cloth Simulation by Position Based Dynamics *09/2021 - 12/2021*

ETHz Physically-based Simulation Course Project

- Simulated a cloth soft body to interact with itself and static object. Based on Taichi and Open3D tools.
- Implemented the simulation pipeline from scratch according to the paper “Position Based Dynamics”.

AWARDS & HONORS

Best Paper Award of ACM MM 2021 10*/2021*

Best Video Award of IJCAI 2021 Video Competition 06*/2021*

National Scholarship *10/2019*

- Awarded to top 2% students based on academic performance by Ministry of Education

Dean’s Honors List, Hong Kong University of Science and Technology *06/2019*

The Samsung Scholarship *11/2018*

Beihang University Merit Student *2016-2018*

- Awarded to top 10% students based on academic performance by Beihang University

The First Prize of the 9th ”Lan Qiao Cup” Contest *04/2018*

- Software Category Beijing Division C/C++ Programming University Group A

PROFESSIONAL SKILLS

Reviewer of CVPR, ICCV, ECCV, 3DV , Information Fusion, and Pattern Recognition

Programming Languages: C/C++, Python (PyTorch, OpenCV)

Miscellaneous: TOEFL-106, GRE 327+3.5, Piano