ZIXIN GUO

📞 +1 6473896610   ✉ zixin.guo@mail.utoronto.ca   ⦿ github.com/charlieguo0610   🌐 zixinguo.me

EDUCATION

University of Toronto   *Sept. 2020 – Jun. 2025*
Honours Bachelor of Science   Toronto, ON
- **Program:** Computer Science Specialist
- **cGPA:** 3.98/4.0, Avg 91.5%
- **Research Interest:** I'm currently interested in researching Computer Vision and Generative AI related projects.

TECHNICAL SKILLS

Technical Skills: Python, C/C++, Java, HTML, CSS, JavaScript, Shell, MySQL, LaTeX, R
Toolkits: PyTorch, CUDA, Diffusers, OpenCV, React, Django, Git, Vim, Slurm

PUBLICATIONS

Z. Guo, J. Zhang, "FastTalker: Jointly Generating Speech and Conversational Gestures from Text," in *European Conference on Computer Vision Workshop (ECCVW)*, 2024, arXiv: 2409.16404.

D. Huo, **Z. Guo**, X. Zuo, Z. Shi, J. Lu, P. Dai, S. Xu, L. Cheng, Y.-H. Yang, "TexGen: Text-Guided 3D Texture Generation with Multi-view Sampling and Resampling," in *European Conference on Computer Vision (ECCV)*, 2024, arXiv: 2408.01291.

Z. Guo, R. Yang, "A Channel Attention and Feature Manipulation Network for Facial Expression Recognition," in *3rd International Conference on Signal Processing and Machine Learning*, 2023 (Oral Presentation), DOI: 10.54254/2755-2721/6/20230751.

Z. Guo, Y. Bai, W. Zhang, T. Mei, B. Zhou, "A Method on Industrial Defect Detection Using Self-Attention Mechanism and Memory Networks, " CN Patent application no. 202210255817.8. Filed Mar 15, 2022, Under Confidentiality Agreement.

EXPERIENCES

Machine Learning and Computational Healthcare Lab, Vector Institute   *Sept. 2024 - Present*
Undergraduate Student Researcher. Supervisor: Prof. Rahul G. Krishnan   Toronto, ON
- Conducted a literature review on various model architectures, spanning from early designs like LSTM, through transformers, to large language models (LLMs), in the context of medical time series.
- Developed a codebase and conducted benchmark studies on various time series forecasting models and data imputation methods, applying them to datasets such as MIMIC and HiRID.
- Demonstrated explainability in patient vital sign predictions and experimented with the performance impact of data imputation using LLM as a zero-shot predictor.

Dynamic Graphics Project Lab, University of Toronto   *May 2024 - Nov. 2024*
UTEA Undergraduate Student Researcher. Supervisor: Prof. David Lindell   Toronto, ON
- Proposed a novel text-to-image diffusion-based optimization method that learns age-aware component-level concepts given a collection of person-specific images spanning years.
- Performed comprehensive quantitative analysis and user studies to evaluate multiple applications of the proposed method, such as recontextualization, component-level swapping, age regression, and progression. Demonstrated enhanced age-specific component control in the personalization process over baseline methods.
- Work resulted in a first-author poster presentation at Applied Research in Action (ARIA) 2024 [1] and a conference paper submission to CVPR 2025 (under review).

Noah's Ark Lab, Huawei Technologies Canada   *May 2023 - May 2024*
Computer Vision Research Intern. Supervisor: Dr. Juwei Lu   Toronto, ON
- Developed and implemented a 3D scene editing pipeline using 3D Gaussian Splatting (3DGS) and diffusion models for smart factory modeling as an alternative to the LiDAR-based method. This proof of concept was successfully transitioned to the production department.
- Proposed and initiated a follow-up project on 3D object removal and view-consistent inpainting in 3DGS, which aligns with both research and industry domains. The work resulted in a US patent submission.
- Contributed to a research project on diffusion-based 3D mesh texturing by developing the Attention-Guided Multi-View Sampling module, leading all baseline benchmark studies, and drafting the initial paper. The work resulted in an ECCV 2024 conference paper.

Dynamic Graphics Project Lab, University of Toronto   *Jan. 2023 - May 2023*
Undergraduate Research Assistant. Supervisor: Prof. David Lindell   Toronto, ON
- Developed and replicated a diffusion-based model from scratch for 3D-aware neural field generation, evaluating various 3D neural representations, such as SDFs, beyond the occupancy fields outlined in the original work.

- Evaluated the model's performance by benchmarking it against Chamfer-L1 and Volumetric IoU metrics to analyze the surface geometry and volumetric alignment of the generated 3D shapes.
- Work ended with a technical report: Evaluating 3D Neural Field Diffusion Model Using Triplanes [2] and served as a baseline for other research projects in the group.

CV Lab, JD.com   *May 2021 - Sept. 2021*
Computer Vision Research Intern. Supervisor: Dr. Tao Mei   Beijing, China
- Proposed a novel self-attention defect detection method integrating a Vision Transformer (ViT) with a memory module for industrial quality inspection.
- Conducted detailed experiments using the proposed method that achieved state-of-the-art results on the SDNET 2018 dataset using just 50% of the training data. This work subsequently led to a published patent.
- Competed in the CVPR 2021 workshop challenge on open-set image classification; identified optimal settings across dozens of combinations to enhance classification accuracy, securing 3rd place on the leaderboard.

PROJECTS

Object Editing with 3D Scene Representation and Rendering   *Jan. 2024 - Apr. 2024*
- Adopted the Segment Anything (SAM) on a single reference view for editing region segmentation, propagated the segmentation mask to remaining views via depth warping, and performed 3D semantic mask reconstruction to accurately localize edited regions within the 3D scene.
- Proposed utilizing 2D editing techniques and attention-guided control to achieve smooth and view-consistent 3D editing. The edited images are then utilized to fine-tune the 3D scene, represented by the neural radiance field (NeRF) or 3DGS.

"An Image is Worth One Sentence": Fast Textual Inversion with Supreme Initialization [3]   *Jan. 2023 - Apr. 2023*
- Improved textual inversion, a state-of-the-art image personalization method, by increasing its convergence speed from 5000 steps to 100 steps.
- Engineered the Image Personalization feature, dynamically customizing images to align with the user's prompt.
- Pioneered the multitoken initialization method and the class/caption initialization method.

Smart Oilfield   *May 2022 - Aug. 2022*
- Engaged oil and gas experts to acquire professional knowledge and annotated a large set of surveillance videos captured at the PetroChina Oilfield for anomaly detection.
- Fine-tuned YOLO series detection models on the annotated datasets for abnormal activity recognition at the construction site, such as PPE wearing, smoke detection, and illegal trespassing.
- Demoed and deployed the trained models on the Atlas 500 AI Edge Station and achieved a 30 FPS detection rate on the monitoring platform.

AWARDS AND HONOURS

University of Toronto Excellence Awards (UTEA) (Awarded $7500 CAD)	*2024*
Dean's List Scholar at the University of Toronto	*2021, 2022, 2023, 2024*
Star Intern Certificate [4] for Self-Attention Defect Detection by the VP Dr. Tao Mei of JD.com	*2021*
3rd Place [5] in Open-World Image Classification Challenge at CVPR 2021 Workshop	*2021*
President's Scholar of Excellence Program at the University of Toronto (Awarded $10000 CAD)	*2020*
Top 2% in the Senior Level Canadian Computing Competition	*2020*
Top 3% in the Euclid Contest and Canadian Senior Math Competition	*2020*

[1] https://drive.google.com/file/d/1q59ebZQiwBY9UgBKMJEUWeX32UDor3Ja/view?usp=sharing
[2] https://drive.google.com/file/d/1VXK9EWMUM-Mzs3i1IDi4z0Kqh7Iy-uvX/view
[3] https://github.com/jenci2114/csc413-project
[4] https://drive.google.com/file/d/1Kh5xj82v_ihqysz4ZNyljiYcowwQaKLJ/view?usp=sharing
[5] https://drive.google.com/file/d/1iDYAeq4svOW1rlZM59sz7bu7gQqT2jW7/view