

Liu He

919-928-2396 | riverliuhe71@gmail.com | Portfolio (arking1995.github.io) | LinkedIn

EDUCATION

Purdue University (*Ph.D. Candidate in Computer Science*, GPA: **3.94/4.0**) 06/2019 – est.2024, West Lafayette, IN

- RA in CGVLab (Computer Graphics and Visualization Lab) since 2019
- Honors: 2024 Purdue Merit Recognition Award (\$1500, for high-profile research)

University of North Carolina at Chapel Hill (*M.A. in Envir. Engineering*) 08/2017 – 05/2019, Chapel Hill, NC

- RA in Remote Sensing and Ecological Modeling Lab

Wuhan University (*B.E. in Electr. Info. Sci. and Tech.*, GPA: **3.75/4.0**, Top: **2%**) 09/2013 – 06/2017, Wuhan, P.R.C.

- Honors: Microsoft Scholarship (1/248); Best graduation thesis (Top: 1%)

INTERNSHIP

Advancing MLLMs by Large-Scale 3D Visual Instruction Dataset Generation 06/2024 – 09/2024, Amazon

- Proposed an unlimited VQA data generator focusing camera-object relation, keeping photorealistic image quality.
- Provided *Ultimate3D* dataset (240K) and benchmark (7K) for finetuning and evaluation of camera-object relation perception.
- Improved LLaVA-1.6 and Llama3.2-Vision to outperform GPT-4o/Claude-V3.5 by 33.4% on prediction accuracy.

Video Generation by MLLM Agent Collaborations (Project: Kubrick) 03/2024 – 05/2024, Baidu Research USA

- Proposed multi-modal LLM agentic workflow for 3D generation, simulation, and animation given multi-modal prompts.
- Designed **multi-agent reflection** and **collaboration** for complex instructions of 3D Engine tool usage (Blender, etc.).
- Finetuned advanced MLLM agents for video, image, and text understanding and comprehensive evaluation of synthesis.

Diffusion-Based Document Layout Generation (Project: DocDiff) 05/2022 – 08/2022, Azure AI, Microsoft Research

- Proposed a **diffusion** model (**DM**) with **Transformer** backbone for document layout synthesis.
- Designed extendable model structure for simultaneous generation of layout and text content for multilingual verticals.
- Discovered the user guidance by prompts obtained by pre-trained CLIP for controllable generation given real document images.

RESEARCH

Refine Generative Artifacts by Semantic Alignment (Project: RefineATF) 05/2024 – 09/2024, Adobe Research (Remote)

- Innovated an automatic artifact localization method by cross-attention on training-free Stable Diffusion feature priors.
- Outperformed SOTAs (Paint-by-Example, AnyDoor, etc.) on generative artifacts refinement for broad image personalization.
- Provided a comprehensive benchmark (*GenArtifactBench*) for generative artifacts detection and refinement.

Scalable Urban Layout Synthesis (Project: GlobalMapper, Project: COHO) 06/2021 – 03/2024, Purdue Univ.

- Established graph-based canonical **Vector Quantized** representation for arbitrary-shaped urban layout with scalable hierarchy.
- Introduced Transformer/Graph Attention Network (**GAT**)/Masked Autoencoder (**MAE**) of infinite 3D urban layout synthesis.
- Implemented city-scale urban 3D modeling and social-climate risk prediction for 330+ cities in the North America.

Globalwise Styled-Controlled Building Modeling by Staged GANs (Project) 06/2019 – 05/2021, Purdue Univ.

- Designed staged GANs for large-scale building segmentation with extreme upsampling refinement (**10x**).
- Utilized learned priors as style control to generate footprints with plausible instance-level metric. Beat SOTA by at least **15%**.

Multi-Modal Continental Land Cover Segmentation 08/2017 – 05/2018, UNC at Chapel Hill

- Accomplished U-Net-based segmentation to continental-scale land cover monitoring of entire Southeastern U.S.
- Implemented dense segmentation across prevailing satellite constellations, achieved over **80%** average precision.

Rule-Based Recognition and 3D modeling of Pole-Shaped Objects (Project Leader) 05/2015 – 06/2017, Wuhan Univ.

- Implemented multi-modal calibrations of camera, LiDAR, and IMU for mobile **SLAM** system. (Patent No. 201511000640.3)
- Integrated in **Unity** GUI to enable “one-click” **rule-based** 3D reconstruction of urban poles with over 90% accuracy.

HIGHLIGHTED PUBLICATIONS

He, L., Xiao, Z., ... (2024) Advancing Multimodal LLMs by Large-Scale 3D Visual Instruction Dataset Generation. (Under Reviewing)

Song, Y., He, L., ... (2024). Refine-by-Align: Refinement of Generative Artifacts for Personalized Image Generation (Under Reviewing) (**Project**)

He, L., Song, Y., Huang, H., Zhou, X. (2024). Kubrick: Multimodal Agent Collaborations for Video Generation. (Under Reviewing) (**Project**)

He, L., & Aliaga, D. (2024, **Oral**). COHO: Context-Sensitive City-Scale Hierarchical Urban Layout Generation. In *European Conference on Computer Vision*. Cham: Springer Nature Switzerland (**Project**)

He, L., & Aliaga, D. (2023). GlobalMapper: Arbitrary-Shaped Urban Layout Generation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 454-464). (**Project**)

He, L., Lu, Y., Corring, J., Florencio, D., Zhang, C. (2023, **Oral**). Diffusion-Based Document Layout Generation. *Document Analysis and Recognition - ICDAR 2023. Lecture Notes in Computer Science*, vol 14187. Springer, Cham. (**Project**)

He, L., Shan, J., Aliaga, D. (2023). Generative Building Feature Estimation from Satellite Images. *IEEE Transactions on Geoscience and Remote Sensing*.

Kamath, H. G., Singh, M., Malviya, N., Martilli, A., **He, L.**, Aliaga, D., ... & Niyogi, D. (2024). GLOBal Building heights for Urban Studies (UT-GLOBUS) for city-and street-scale urban simulations: Development and first applications. *Scientific Data*, 11(1), 886.

Patel, P., Kalyanam, R., **He, L.**, Aliaga, D., & Niyogi, D. (2023). Deep Learning based Urban Morphology for City-scale Environmental Modeling. *PNAS Nexus*, pgad027.

Zhang, X., Ma, W., Varinlioglu, G., Rauh, N., **He, L.**, & Aliaga, D. (2022). Guided pluralistic building contour completion. *The Visual Computer*, 1-12.

Bhatt, M., Kalyanam, R., Nishida, G., **He, L.**, May, C., Niyogi, D., & Aliaga, D. (2020). Design and Deployment of Photo2Building: A Cloud-based Procedural Modeling Tool as a Service. In *Practice and Experience in Advanced Research Computing* (pp. 132-138).

Wang, L., Huang, Y., Shan, J., & **He, L.** (2018). MSNet: Multi-Scale Convolutional Network for Point Cloud Classification. *Remote Sensing*, 10(4), 612.

SKILL

Languages: C++ | Python | C | JAVA | Matlab | R **Libraries:** Pytorch | OpenCV | OpenGL | Qt | Tensorflow | Pthread
Tools & OS: Linux | Git | LLVM | Google Cloud | Unity | Google Earth Engine