# Beilei Cui

beileicui@link.cuhk.edu.hk / beileicui@gmail.com / nickyemail@126.com
(86) 19250180314 | https://beileicui.github.io

## Education Experience

**The Chinese University of Hong Kong，Department of Electronic Engineering** 2022.08 - 2027.07(Estimated)

- ✧ Degree & Major: Doctor's degree，Electronic Engineering（The Vice Chancellor scholarship of CUHK）
- ✧ Supervisor：Prof.Hongliang Ren & Prof.Jiewen Lai
- ✧ Research Interests：Depth Estimation and 3D Reconstruction, Medical Image Analysis, Medical Robot Perception and Navigation

**Imperial College London，Department of Electrical and Electronic Engineering** 2021.10 - 2022.07

- ✧ Degree & Major: Master's degree，Communication and Signal Processing
- ✧ GPA：Second Class Degree（Pass with Merit）

**University of Electronic Science and Technology of China，School of Information and Communication Engineering** 2017.09 - 2021.06

- ✧ Degree & Major: Bachelor's degree，Communication Engineering
- ✧ GPA：3.9/4

Representative honors and awards：

[1] 2025, MICCAI (CCF-B) Young Scientist Award
[2] 2025, MICCAI (CCF-B) Best Paper Candidate
[3] 2024, IJCARS Journal IPCAI Conference Special Issue Best Paper Candidate
[4] 2024, ICBIR Second place in the Best Paper on Medical Robots
[5] 2022, The Vice Chancellor scholarship of CUHK

## Interns Experiences

**Alibaba, DAMO Academy, Medical AI Lab** 2025.05 - 2025.09

- ✧ Construct a segmentation-based liver enhanced CT tumor diagnosis model, using CT image prediction combined with various blood test history data for comprehensive diagnosis. Evaluate various indicators of the enhanced CT liver tumor diagnostic model on internal and external centers, and explore the ceiling of the enhanced CT liver tumor diagnostic model.

## Research Interests

**Surgical Robot Scene Understanding and Navigation - Building Efficient Surgical Navigation**

- ✧ **Depth estimation and 3D reconstruction:** Existing depth estimation models perform well in natural scenes, but their zero-shot performance is still poor in medical scenes. We efficiently fine tune Depth Anything using LoRA to obtain high-precision depth estimation with minimal training costs [2]; Meanwhile, current self-supervised pose and depth estimation schemes still require camera intrinsic parameters to obtain accurate reconstruction results. We design an intrinsic parameter estimation decoder and a more efficient dynamic vector LoRA, which can achieve high-precision pose and depth estimation without the information of camera intrinsic parameters [3]; We further propose the gate controlled dynamic vector LoRA to enable a unified model for estimating depth, pose, and camera intrinsic parameters, and designed a 3D reconstruction optimization algorithm to achieve higher accuracy in surgical scene 3D reconstruction [9]; We also introduce vibration signals to construct a vibration visual multimodal self-supervised visual pose depth estimation framework [8]; We further apply foundation models for the

initialization of 4D Gaussian splashing and introduce confidence calibration strategies to enhance the reconstruction performance in deformable scenes [5] and SLAM algorithm performance [6]. We also conducted the promotion and algorithm design of 4D Gaussian splashing in practical applications for low light scenes [10] and text driven 3D segmentation scenes [11]. We also estimate scene scales by combining image and text descriptions, efficiently transforming relative depth into metric depth across multiple domains using a lightweight model [1].

✧ **Surgical video segmentation based on noise label learning:** Noise labels have a significant impact on the performance of model training. By comparing the feature space similarity of the adjacent frames, we filtered out normal labels and noise labels, and further controlled the loss function at the pixel level, image level, and video level to construct a surgical video segmentation model that is more robust to noisy labels [7].

**Surgical Robot Imaging System - Improving Imaging Quality for Real Scene Data**

✧ **Non uniform rotational distortion correction in optical coherence tomography (OCT) imaging:** The imaging quality of our self-designed motorless remote robot OCT endoscope is poor due to non-uniform rotational distortion (NURD). We first establish a dataset of normal and distorted OCT images corresponding to scanning speeds based on a physical model, and then design a step-by-step training scheme to train two network models to recognize distorted image regions and estimate the scanning speed of distorted regions, respectively; Afterwards, we propose an inverse transform interpolation distortion correction algorithm to restore the OCT images and obtain higher quality OCT images and 3D reconstruction results [4].

## Excerpts of Publications [Full publication list to be found in Google Scholar]

[1] **B. Cui**, et al. *TR2M: Transferring Monocular Relative Depth to Metric Depth with Language Descriptions and Scale-Oriented Contrast.* The IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2026. (CCF-A， First Author)

[2] **B. Cui**, et al. *Surgical-DINO: adapter learning of foundation models for depth estimation in endoscopic surgery*. International Journal of Computer Assisted Radiology and Surgery (**IJCARS**)-IPCAI 2024. (IF: 3.1, Best Paper Short lists, First Author)

[3] **B. Cui**, et al. *EndoDAC: Efficient adapting foundation model for self-supervised depth estimation from any endoscopic camera.* Medical Image Computing and Computer Assisted Intervention (**MICCAI**), 2024. (CCF-B， First Author, Early Accept)

[4] S. Yuan*, C. Xu*, **B. Cui***, et al. *Motor-free telerobotic endomicroscopy for steerable and programmable imaging in complex curved and localized areas*. Nature Communications (**NC**). (IF:14.7, Co-first Author)

[5] Y. Huang*, **B. Cui***, L. Bai*, et al. *Endo-4DGS: Endoscopic Monocular Scene Reconstruction with 4D Gaussian Splatting.* Medical Image Computing and Computer Assisted Intervention (**MICCAI**), 2024. (CCF-B, Co-first Author)

[6] Y. Huang*, **B. Cui***, et al. *Advancing Dense Endoscopic Reconstruction with Gaussian Splatting-driven Surface Normal-aware Tracking and Mapping*. IEEE International Conference on Robotics and Automation (**ICRA**), 2025. (CCF-B, Co-first Author)

[7] **B. Cui**, et al. *Rectifying Noisy Labels with Sequential Prior: Multi-scale Temporal Feature Affinity Learning for Robust Video Segmentation.* Medical Image Computing and Computer Assisted Intervention (**MICCAI**), 2023. (CCF-B, First Author)

[8] L. Bai*, **B. Cui***, et al. $V^2$-SfMLearner: *Learning Monocular Depth and Ego-motion for Multimodal Wireless Capsule Endoscopy.* IEEE Transactions on Automation Science and Engineering (**T-ASE**). (IF:5.6, CCF-B, Co-first Author)

[9] **B. Cui**, et al. *Learning to Efficiently Adapt Foundation Models for Self-Supervised Endoscopic 3D Scene Reconstruction from Any Cameras.* Under review at Medical Image Analysis (**MedIA**). (IF:10.7, CCF-C, First Author)

[10] Y. Huang*, L. Bai*, **B. Cui***, et al. *Endo-4DGX: Robust Endoscopic Scene Reconstruction and Illumination Correction with Gaussian Splatting.* Medical Image Computing and Computer Assisted Intervention (**MICCAI**), 2025. (CCF-B, Co-first Author)

[11] Y. Huang*, L. Bai*, **B. Cui***, et al. *SurgTPGS: Semantic 3D Surgical Scene Understanding with Text Promptable Gaussian Splatting.* Medical Image Computing and Computer Assisted Intervention (**MICCAI**), 2025. (CCF-B, Best Paper Candidate, Young Scientist Award, Co-first Author)

## Academic Services

- ✧ **Program Committee：** MICCAI 2024 EARTH workshop, EMA4MICCAI 2025 workshop
- ✧ **Academic Reviews：** T-PAMI, T-MI, T-NNLS, MedIA, TVCG, RA-L, ICCV, AAAI, MICCAI, ICRA, IROS, IPCAI
- ✧ **Language Capability:** TOEFL：112/120，GRE：328/340，GRE writing 3.5