Yuxin Yao

+44(0)7784316157 | yy561@cam.ac.uk | yuxinyao620.github.io/ | https://github.com/YuxinYao620

Education

University of Cambridge — Cambridge, UK
PhD in Engineering — September 2023 - Present
- Research about Computer Vision and Geometric Deep Learning.

University College London — London, UK
MEng Mathematical Computation — September 2019 - June 2023
- First Class Honours

Technical Skills

- **Programming**: Proficient in Python (7 years), PyTorch (6 years), vLLM, Lightning, HuggingFace, Scikit-learn, and TensorFlow.
- **Deep Learning**: Experienced with VLM, Diffusion Models, Transformers, Autoencoders.
- 3D **Vision**: Skilled in 4D reconstruction, Gaussian splatting, and Colmap for SLAM tasks. Proficient in working with the SMPL-X human body model and tools like Blender.

Work Experience

EPIC Games — May 2025 - Dec 2025
Machine Learning Engineer — London, UK

- **IP Infringement Detection**: Evaluated state-of-the-art **VLMs** including Qwen, ChatGPT and Gemini for IP violation detection. benchmarked on 100k+ images and improved precision/accuracy by over 40%. Reduced inference time from 13 hours to 30 minutes using **vLLM** and distributed multiprocessing.
- **Automated Detection Pipeline**: Developed a scalable detection pipeline using YOLO-World to identify objects potentially infringing Epic's IP across large-scale datasets. This service is deployed and integrated into EPIC's IP detection service.
- **Texture Generation for 3D Assets**: Built a re-texturing pipeline for 3D meshes using Hunyuan3D with text/image prompts; familiar with modern text-to-3D and image-to-3D systems used in procedural content pipelines.
- **VLM Finetuning for PCG**: Collected in-engine datasets in Unreal Engine and designed a finetuning-ready data schema. Contributed to AI-driven procedural building generation within the Unreal Engine Procedural content generation ecosystem.

Research Experience

Feed-Forward 3D Object Articulation — June 2024 – Dec 2025
University of Cambridge — Department of Engineering

- **Feed-forward Transformer Architecture**: Designed a transformer with learned embeddings and Mixture-of-Experts modules for predicting part segmentation and articulation structure. Improved previous baselines by **2x in overall accuracy**, with strong generalization to unseen object categories.
- **Synthetic Data Generation**: Processed USD/URDF assets; used Blender for rendering and DINOv3 for feature extraction; applied IsaacSim to obtain ground-truth articulations.

- **Benchmark Creation**: Built evaluation benchmark; proposed improved metrics aligned with human perception.

Gaussian Splatting based Camera Pose Estimation June 2024 - March 2025
University of Cambridge Department of Engineering

- **Camera Pose Estimation**: Engineered efficient batch camera pose estimation using Gaussian splattings, achieving SOTA performance in handling videos with moving objects in specific scenarios.
- **Small-Baseline Video**: Enhanced camera pose estimation in videos with minimal camera movement to support existing scene reconstruction pipelines.
- **Pretrained Models**: Utilized pretrained models (e.g., DINOv2, AnyDepth) for feature extraction, depth estimation, and semantic segmentation; familiar with current foundation models.
- **Dynamic Motion Modeling**: Modelled dynamic object motions with motion field applied with deep neural network using PyTorch. Optimized the motion field with prior knowledge of the semantics of the video.

Human Motion Generation and Control with Keyframes May 2024 - Jan 2025
University of Cambridge Department of Engineering

- **Apply Diffusion Models**: Employed Diffusion model with different backbones for conditioned autoregressive human motion generation.
- **Design Algorithm**: Disigned keyframe selection algorithms, choosing the keyframe timing with respect to the user inputs.
- **Classifier Guidance**: Applied Skeleton-based gradient guidance to control the generated keyframe while maintaining the motion consistency.
- **Human Motion Controlling**: Add user indicated control point using Blender and regenerate the keyframes based on the control points, generating novel human motions.

Publications and Talk

- Li, R.*,**Yao, Y***, ... **Particulate: Feed-Forward 3D Object Articulation** . CVPR under reviewing
- **Yao, Y.**, ...**SmallGS: Gaussian Splatting Based Camera Pose Estimation for Small-Baseline Videos**. CVPR workshop 2025
- Zheng, B., Chen, K., **Yao, Y.**, ... **AutoKeyframe: Autoregressive Keyframe Generation for Human Motion Synthesis and Editing**. Siggraph 2025
- **Yao, Y.**, Hockey, C., Lasenby, J. **Simplifying and Generalising Equivariant Geometric Algebra Networks**. The 9th conference on Applied Geometric Algebras in Computer Science and Engineering
- Yan, Y., Schaffter, T., Bergquist, T., ...**Yao, Y.**.,... DREAM Challenge Consortium. (2021). **A Continuously Benchmarked and Crowdsourced Challenge for Rapid Development and Evaluation of Models to Predict COVID-19 Diagnosis and Hospitalization.** JAMA network open, 4(10), e2124946-e2124946.
- **Contributed Talk**: Simplifying and Generalising Equivariant Geometric Algebra Networks 9th conference on Applied Geometric Algebras in Computer Science and Engineering (Amsterdam, NL)

Award

- Award of College Senior Scholarships, Fitzwlliams College, University of Cambridge.