Yu Zeng Ph.D.

✉ zengxianyu18@gmail.com   🏠 https://zengxianyu.github.io/   🎓 Google Scholar   ⭕ GitHub

Research Interest

My research interest lies in computer vision and deep learning. My research advances artificial intelligence through two primary directions: generative AI and label-efficient learning for computer vision. My PhD research focused on generative models for visual synthesis with multimodal and hierarchical inputs. My most recent research focuses on advancing embodied AI using deep generative models. By combining these areas, I aim to build scalable AI models that bridge human creativity and machine intelligence while minimizing the need for intensive human supervision.

Education

Johns Hopkins University

Ph.D. in Electrical and Computer Engineering. 2024

Dalian University of Technology

M.S. in Information Engineering.

B.S. in Electronic and Information Engineering.

Employment History

2024 – present	**Research Scientist,** NVIDIA.
2023 – 2024	**Research Intern,** NVIDIA.
2022 – 2023	**Research Scientist Intern,** Adobe.
2020 – 2021	**Researcher,** Tencent.
2019 – 2019	**Research Scientist Intern,** Adobe.

Research Publications

Under Review/Preprints

1. NVIDIA, "Cosmos world foundation model platform for physical ai," *arXiv preprint arXiv:2501.03575,* 2025. 🔗 URL: https://github.com/NVIDIA/Cosmos/.

2. NVIDIA, "Cosmos-transfer1: Conditional world generation with adaptive multimodal control," *arXiv preprint arXiv:2503.14492,* 2025.

3. J. Gu, X. Liu, **Y. Zeng**, Y. Fan, K. Zhou, Q. Yan, F. Zhu, A. Nagarajan, D. Hong, M.-Y. Liu, and X. E. Wang, "Divephygen: A multidimensional benchmark for evaluating physical realism in video generation models," *Under Review.,* 2025.

4. H. Zhao, Y. Zhuge, H. Lu, L. Wang, and **Y. Zeng**, "Learning universal features for generalizable image forgery localization," *Under Review.,* 2024.

5. Z. Wang, Z. Li, A. Mandlekar, Z. Xu, J. Fan, Y. Narang, L. Fan, Y. Zhu, Y. Balaji, M. Zhou, M.-Y. Liu, and **Y. Zeng**, "One-step diffusion policy: Fast visuomotor policies via diffusion distillation," *arXiv preprint arXiv:2410.21257,* 2024. 🔗 URL: https://research.nvidia.com/labs/dir/onedp/.

6. NVIDIA, "Edify image: High-quality image generation with pixel space laplacian diffusion models," *arXiv preprint arXiv:2411.07126,* 2024. 🔗 URL: https://research.nvidia.com/labs/dir/edify-image/.

7. L. Ling, C.-H. Lin, T.-Y. Lin, Y. Ding, **Y. Zeng**, Y. Sheng, Y. Ge, M.-Y. Liu, A. Bera, and M. Li, "Scenethesis: Combining language and visual priors for 3d scene generation and reasoning," *Under review*, 2024.

8. **Y. Zeng**, M. Zhou, Y. Xue, and V. M. Patel, "Securing deep generative models with universal adversarial signature," *arXiv preprint arXiv:2305.16310*, 2023.

9. **Y. Zeng**, Z. Lin, and V. M. Patel, "Shape-guided object inpainting," *arXiv preprint arXiv:2204.07845. Under Review.*, Jul. 2022.

10. **Y. Zeng**, H. Lu, and A. Borji, "Statistics of deep generated images," *arXiv preprint arXiv:1708.02688*, 2017.

Published

1. H. Zhao, **Y. Zeng**, H. Lu, and L. Wang, "Large occluded human image completion via image-prior cooperating," *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024.

2. **Y. Zeng**, V. M. Patel, H. Wang, X. Huang, T.-c. Wang, M.-Y. Liu, and Y. Balaji, "Jedi: Joint-image diffusion models for finetuning-free personalized text-to-image generation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Rocognition*, 2024. 🔗 URL: https://research.nvidia.com/labs/dir/jedi/.

3. Y. Mei, **Y. Zeng**, H. Zhang, Z. Shu, X. Z. Zhang, S. Bi, J. Zhang, H. Jung, and V. M. Patel, "Holo-relighting: Controllable volumetric portrait relighting from a single image," *Proceedings of the IEEE Conference on Computer Vision and Pattern Rocognition*, 2024.

4. **Y. Zeng**, Z. Lin, J. Zhang, Q. Liu, J. Collomosse, J. Kuen, and V. M. Patel, "Scenecomposer: Any-level semantic image synthesis," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (Hightlight, top 2.5%)*, 2023. 🔗 URL: https://zengyu.me/scenec/.

5. J. Shang, **Y. Zeng**, X. Qiao, X. Wang, R. Zhang, G. Sun, V. Patel, and H. Fu, "Jr2net: Joint monocular 3d face reconstruction and reenactment," *Proceedings of the AAAI Conference on Artificial Intelligence (Oral presentation)*, 2023.

6. **Y. Zeng**, Z. Lin, and V. M. Patel, "Sketchedit: Mask-free local image manipulation with partial sketches," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5951–5961, 2022. 🔗 URL: https://zengyu.me/sketchedit/.

7. S. Cai, **Y. Zeng**, S. Yang, X. Jia, H. Lu, and Y. He, "Deformable dynamic sampling and dynamic predictable mask mining for image inpainting," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

8. H. Zhang, **Y. Zeng**, H. Lu, L. Zhang, J. Li, and J. Qi, "Learning to detect salient object with multi-source weak supervision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

9. **Y. Zeng**, Z. Lin, H. Lu, and V. M. Patel, "Cr-fill: Generative image inpainting with auxiliary contextual reconstruction," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 14 164–14 173, 2021. 🔗 URL: https://github.com/zengxianyu/crfill.

10. **Y. Zeng**, Z. Lin, J. Yang, J. Zhang, E. Shechtman, and H. Lu, "High-resolution image inpainting with iterative confidence feedback and guided upsampling," *European conference on computer vision*, pp. 1–17, 2020. 🔗 URL: https://zengxianyu.github.io/iic/.

11. Y. Zhuge, **Y. Zeng**, and H. Lu, "Deep embedding features for salient object detection," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 9340–9347, 2019.

12. **Y. Zeng**, Y. Zhuge, H. Lu, L. Zhang, M. Qian, and Y. Yu, "Multi-source weak supervision for saliency detection," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 6074–6083, 2019. 🔗 URL: https://github.com/zengxianyu/mws.

13 **Y. Zeng**, Y. Zhuge, H. Lu, and L. Zhang, "Joint learning of saliency detection and weakly supervised semantic segmentation," *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 7223–7233, 2019. 🔗 URL: https://github.com/zengxianyu/jsws.

14 **Y. Zeng**, H. Lu, L. Zhang, M. Feng, and A. Borji, "Learning to promote saliency detectors," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1644–1653, 2018.

15 **Y. Zeng**, M. Feng, H. Lu, G. Yang, and A. Borji, "An unsupervised game-theoretic approach to saliency detection," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4545–4554, 2018.

Patents

1 **Y. Zeng**, Y. Balaji, T. Wang, X. Huang, and M.-Y. Liu, "Neural networks to generate objects within different images," US Patent App. 18/518,430, 2023.

2 **Y. Zeng**, Z. Lin, J. Zhang, Q. Liu, J. W. Y. Kuen, and J. Collomosse, "Multi-modal image generation," US Patent App. 18/057,857, Nov. 2022.

3 Z. Lin, **Y. Zeng**, J. Yang, J. Zhang, and E. Shechtman, "Iterative image inpainting with confidence feedback," US Patent 11,398,015, Jul. 2022.

4 Z. Lin, **Y. Zeng**, J. Yang, J. Zhang, and E. Shechtman, "Guided up-sampling for image inpainting," US Patent App. 16/864,388, Nov. 2021.

5 H. Lu, **Y. Zeng**, H. Zhang, J. Li, and L. Zhang, "Method for detection image salient object," CN110956185B, 2019.

6 Y. Wang, H. Ma, J. Ma, R. Yang, and **Y. Zeng**, "Coin separator," CN206097264U, 2016.

Teaching

Johns Hopkins University

Spring 2023	Deep Learning (EN.520.438), Guest Lecturer
Spring 2024	Machine Learning Meets Networks (EN.520.698), Course Assistant
	Deep Learning (EN.520.438), Guest Lecturer

Invited Talks & Presentations

Learning to synthesize images with multimodal and hierarchical inputs Baltimore County. April. 2024	University of Maryland,
Learning to synthesize images with multimodal and hierarchical inputs Science and Technology. Feb. 2024	King Abdullah University of
Learning to synthesize images with multimodal and hierarchical inputs Science and Technology, Guangzhou. Jan. 2024	Hongkong University of
Deep generative models for image synthesis and editing	Nanjing University. Dec. 2023
Deep generative models for image synthesis and editing Nov. 2023	Chinese University of Hongkong, Shenzhen.
SceneComposer: Any-level Semantic Image Synthesis.	TikTok. Jul. 2023
Joint Monocular 3d Face Reconstruction and Reenactment.	AAAI. Feb. 2023
Deep Learning for Dense Prediction Problems in Computer Vision.	Adobe. Oct. 2019

Skills

Coding	Python, C, C++, Matlab
Web Dev	Html, css, JavaScript
Misc.	Research on computer vision, image processing, machine learning, artificial intelligence

Awards and Achievements

2024	KAUST Rising Stars in AI
2023	DAAD AInet fellowship
2021	Johns Hopkins University ECE Kewei Yang and Grace Xin Fellowship
2020	Third place of AIM 2020 challenge on image extreme inpainting (ECCV 2020 Workshop)
2018	Third place of 2018 OPPO Top AI Competition on Portrait Segmentation (3rd from 456 teams)
2017	China National Scholarship
2016	Second Place of China National College Mechanical Design Contest

Services

Invited Reviewer for

2022,2023,2024	International Conference on Learning Representations (ICLR)
2020, 2021, 2022, 2023,2024	The IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR)
2022	The Conference on Neural Information Processing Systems (Neurips)
2019,2021	The International Conference on Computer Vision (ICCV)
2020,2022,2024	The European Conference on Computer Vision (ECCV)
2020,2021,2023	The AAAI Conference on Artificial Intelligence (AAAI)
2022	The ACM SIGGRAPH Asia Conference
	IEEE Transactions on Cybernetics
	IEEE Transactions on Image Processing
	IEEE Transactions on Pattern Analysis and Machine Intelligence
	Pattern Recognition
	International Journal of Computer Vision
	IEEE Transactions on Neural Networks and Learning Systems