Minguk Kang

minguk@pika.art | mingukkang1994@gmail.com | https://mingukkang.github.io | Google Scholar | GitHub | X

Education

Ph.D. in Graduate School of AI — Pohang, Republic of Korea
*Pohang University of Science and Technology (**POSTECH**)* — *02/2020 – 02/2026*
Advisors: Prof. Suha Kwak (2023–2026) and Prof. Jaesik Park (2020–2023)
Thesis: Efficient Deep Generative Models for Visual Content Generation
GPA: 4.09/4.30

B.S. in Engineering — Busan, Republic of Korea
Pusan National University — *03/2013 – 08/2019*
Major: Mechanical Engineering, Minor: Statistics
Graduated summa cum laude, ranked 1st among 394 students in the College of Engineering.

Experience

Founding Research Scientist — South Korea (Remote)
Pika Labs — *11/2024 – Present*
Core contributor to video generation models, an audio-driven performance model, and PikaStream1.0;
Developed tokenizers, distillation pipelines, fast super-resolution methods, and real-time video agent system components.

Research Scientist Intern — Korea (Remote) / Palo Alto, USA
Pika Labs — *06/2024 – 10/2024*
Worked with Chenlin Meng on video generation research.

Research Scientist Intern — Korea (Remote) / San Francisco, USA
Adobe Research — *07/2022 – 05/2024*
Mentors: Taesung Park, Connelly Barnes, Eli Shechtman, Jun-Yan Zhu, Richard Zhang, Sylvain Paris
Our GigaGAN (CVPR 2023) contributed to the development of Adobe Firefly.

Undergraduate Research Student — Busan, Republic of Korea
Vision and Intelligent System Laboratory — *08/2017 – 01/2020*
Advisor: Prof. Dongjoong Kang.

Research Interest

My research focuses on efficient generative AI systems for real-time content generation across video, audio, and multimodal settings. I develop high-compression, low-latency tokenizers, few-step diffusion distillation methods, and fast super-resolution pipelines for interactive generation. My broader interests include scalable training of GANs and diffusion models, tokenizer design and diffusibility, and multimodal generative modeling.

Publications

Conferences

1. Minguk Kang and Suha Kwak
   FlashDecoder: Real-Time Latent-to-Pixel Streaming Decoder with Transformers
   CVPR, 2026
2. Minguk Kang, Richard Zhang, Connelly Barnes, Sylvain Paris, Suha Kwak, Jaesik Park, Eli Shechtman, Jun-Yan Zhu, Taesung Park
   Distilling Diffusion Models into Conditional GANs
   ECCV, 2024
3. Seoyeon Kim, Minguk Kang, Dongwon Kim, Jaesik Park, Suha Kwak
   Extending CLIP's Image-Text Alignment to Referring Image Segmentation
   NAACL, 2024
4. Joonghyuk Shin, Minguk Kang, Jaesik Park
   Fill-Up: Balancing Long-Tailed Data with Generative Models
   arXiv:2306.07200, 2023
5. Tony Lee, Michihiro Yasunaga, Chenlin Meng, Yifan Mai, Joon Sung Park, Agrim Gupta, Yunzhi Zhang, Deepak Narayanan, Hannah Benita Teufel, Marco Bellagente, Minguk Kang, Taesung Park, Jure Leskovec, Jun-Yan Zhu, Li Fei-Fei, Jiajun Wu, Stefano Ermon, Percy Liang
   Holistic Evaluation of Text-to-Image Models

NeurIPS Datasets and Benchmarks Track, 2023 (Spotlight)

6. Minguk Kang, Jun-Yan Zhu, Richard Zhang, Jaesik Park, Eli Shechtman, Sylvain Paris, Taesung Park
Scaling up GANs for Text-to-Image Synthesis
CVPR, 2023 (Highlight; review score 5/5/5; top 2.5% among 9,155 submissions)
7. Jinoh Cho, Minguk Kang, Vibhav Vineet, Jaesik Park
Context-Aware Image Completion
AI4CC **CVPR** Workshop, 2023
8. Minguk Kang, Woohyeon Shim, Minsu Cho, Jaesik Park
Rebooting ACGAN: Auxiliary Classifier GANs with Stable Training
NeurIPS, 2021
9. Minguk Kang and Jaesik Park
ContraGAN: Contrastive Learning for Conditional Image Generation
NeurIPS, 2020
10. Minguk Kang, Honghyun Kim, Dongjoong Kang
Finding a High Accuracy Neural Network for the Welding Defects Classification Using Efficient Neural Architecture Search via Parameter Sharing
ICCAS, 2018

Journals

1. Minguk Kang, Joonghyuk Shin, Jaesik Park
StudioGAN: A Taxonomy and Benchmark of GANs for Image Synthesis
IEEE **TPAMI**, 2023
2. Hyojung Ahn, Hanlim Choi, Minguk Kang, Sungtae Moon
Learning-Based Anomaly Detection and Monitoring for Swarm Drone Flights
Applied Sciences, 2019

Products & Software

1. **PikaStream1.0**: Core contributor to a real-time video agent system for group video chat, focusing on low-latency generation and multimodal capabilities.
2. **Audio-Driven Performance Model**: Developed generation and acceleration pipelines.
3. **Pika Video Generation Models**: Contributed to Pika 1.5, Pika 2.0, Pika 2.1, and Pika 2.2, with work on tokenizers, distillation, and fast super-resolution.
4. **E-LatentLPIPS** (★150+): Latent-space perceptual metric for Latent Diffusion Models, enabling regression tasks without pixel decoding.
5. **Adobe Firefly**: Adobe's visual generative AI product suite; my GigaGAN research contributed to its development.
6. **PyTorch StudioGAN** (★3500+): Open-source PyTorch library for representative GAN training and evaluation.

Selected Awards

1. Outstanding Reviewer, European Computer Vision Association, 2024
2. Graduate School Presidential Science Scholarship, Korea Student Aid Foundation, 2024
3. 2nd Prize, BK21 Outstanding Paper Awards, POSTECH Graduate School of AI, 2024
4. 1st Prize, BK21 Outstanding Paper Awards, POSTECH Graduate School of AI, 2022
5. Qualcomm Innovation Fellowship Korea, Qualcomm, 2021
6. Silver Prize, 16th Samsung Electro-Mechanics Paper Awards, 2020
7. National Science and Engineering Scholarship, Korea Student Aid Foundation, 2013–2019

Academic Services

Reviewer

- Conference Reviewer 2026: CVPR, NeurIPS
- Conference Reviewer 2024: ECCV, SIGGRAPH Asia
- Conference Reviewer 2023: ICML, ICCV, NeurIPS
- Conference Reviewer 2022: ICLR, CVPR, ECCV, NeurIPS
- Journal Reviewer: TPAMI, IJCV

Patents

Registered

1. Minguk Kang, Jaesik Park, "Method and Apparatus for Generating and Editing Images Using Contrastive Learning and Generative Adversarial Network," Korea Patent No. 1024777000000, Dec 2022.
2. Jinoh Cho, Jaesik Park, Minguk Kang, "Method for Restoring Damaged Images, Device for the Same, and Method for Image Reconstruction," Korea Patent No. 1028500230000, Aug 2025.

Filed

1. Taesung Park, Minguk Kang, Richard Zhang, Jun-Yan Zhu, Eli Shechtman, Sylvain Paris, "Text-Based Image Generation," US Patent Application No. 18170963, Feb 2023.
2. Taesung Park, Minguk Kang, Richard Zhang, Jun-Yan Zhu, Eli Shechtman, Sylvain Paris, "Super-Resolution on Text-to-Image Synthesis with GANs," US Patent Application No. 18171046, Feb 2023.

Proficiencies

- **Languages**: Korean (Native), English (Professional Working Proficiency)
- **Machine Learning Libraries**: PyTorch (Advanced), TensorFlow (Advanced)