JUNE YOUNG YI

Audio ML Researcher & Engineer

Seoul, Republic of Korea
julianyi0417@gmail.com | +82-10-9909-6392
github.com/Rick-McCoy | LinkedIn | Google Scholar

Summary

Audio ML Researcher & Engineer with **4+ years of industry experience** and **NeurIPS 2025** authorship. Specializes in **Generative Audio** and **Data-Centric AI**, engineering the synthetic pipelines and fine-tuning infrastructure that enabled the open-source release of the Supertonic TTS system. Proven ability to bridge academic research with production-grade deployment.

Education

Seoul National University (SNU) *Feb 2019 – Current*
B.S. in Computer Science and Engineering Seoul, Korea
GPA: 3.94 / 4.3 (Major), 3.91 / 4.3 (Overall) (≈96.1/100)

Note on Timeline: Served 3 years as **Alternative Research Personnel** (in lieu of mandatory military service in South Korea), working full-time in ML R&D at NGINE Studios (2021–2024).

Publications & Preprints

- **SAO-Instruct: Free-form Audio Editing using Natural Language Instructions**. Ungersböck, Grötschla, Lanzendörfer, **Yi**, Choi, Wattenhofer. *NeurIPS 2025 (main track, poster). arXiv:2510.22795. Contribution:* manual-edit synthetic data pipeline (categories + code).
- **Training Flow Matching Models with Reliable Labels via Self-Purification**. Kim, Yu, **Yi**, Lee. *arXiv:2509.19091. Submitted to ICASSP 2026 (Main Track, under review).*
- **Robust TTS Training via Self-Purifying Flow Matching**. **Yi** et al. *arXiv:2512.17293. ICASSP 2026 SPGC (WildSpoof Challenge) Submission.*

Experience

Researcher (Part time, while senior at SNU) *Apr 2025–Dec 2025*
Supertone (HYBE) — Research Team Seoul, Korea

- **Supertonic (Open-Source Release):** Engineered the synthetic data pipeline and fine-tuning framework that cleared the path for public release; enabled the model to handle complex real-world edge cases (GitHub, HF, PyPI).
- **Model Robustness:** Resolved critical pronunciation failures (abbreviations, scientific units) in the 66M-parameter architecture via category-specific synthetic generation + ASR validation; validated on internal hard-case sets.
- **Neural Codec:** Prototyped a high-fidelity 44.1kHz speech codec operating at an ultra-low **12.5Hz framerate**, enabling highly efficient tokenization for downstream generation tasks.

Machine Learning Engineer (On leave from SNU) *Sep 2024–Dec 2024*
OptimizerAI (Text-to-Audio Startup) Seoul, Korea

- End-to-End Ownership: Architected and implemented the V2 text-to-audio system (SFX-focused); managed the full lifecycle from data synthesis and codec training to evaluation and deployment.
- Trained at scale on 8×H100 for ∼2 weeks; optimized sampling (5s generated in 1–2s) and managed a six-figure GPU-credit budget.

Team Leader / ML Engineer (Technical Research) *Aug 2021–Sep 2024*
NGINE Studios (NEXON subsidiary) Seoul, Korea

- Sparse Tabular Modeling: Replaced dense feature representations with sparse embeddings, reducing memory footprint by 100–1000× and enabling training on ultra-high-dimensional user features previously impossible to fit in memory.
- Internal TTS: rebuilt mel-based TTS + vocoder pipeline and increased training throughput by ≥10×; supported internal narration and an accessibility deployment in a live title.
- Led a ∼15-person team: project prioritization, stakeholder communication with game teams, and transition-to-support for delivered systems.

Awards & Scholarship

ICASSP 2026 SPGC (WildSpoof Challenge TTS Track) *2026*
Achieved best WER metric; 2nd in MOS
ElevenLabs × a16z Worldwide Hackathon (Voice Guardian) *2024*
Seoul Region 1st Place
Samsung Humantech Paper Award (Bronze; National research competition) *2018*
KFAS Undergraduate Scholarship (Highly selective, full merit-based support) *2019–Current*

Skills

Languages & Frameworks: Python, PyTorch, PyTorch Lightning
Deployment & Tools: Docker, Kubernetes, Linux/Bash, Git, CI/CD
Audio: TTS, ASR, Neural Codecs, Audio LMs
Languages: Korean (native), English (fluent)