Sang Hoon (Tony) Woo

+82 10-4801-0363 tonywoo.me tonyswoo@gmail.com linkedin.com/in/tonyswoo
github.com/tonyswoo scholar.google.com/citations?user=kt3_CnoAAAAJ

Education

University of California, Davis **Davis, CA**
Bachelor of Science in Computer Science *Sep. 2015 – Jun. 2019*

- GPA: 3.68/4.0 (Major GPA: 3.91/4.0)
- Dean's Honor List: Spring 2017, Fall 2017, Winter 2018, Spring 2018, Fall 2018, Winter 2019, Spring 2019

Research

Seoul National University Vision & Learning Lab **Seoul, South Korea**
Researcher *Jan. 2025 – Current*

- Conducted research on spoken dialogue systems and general audio processing under Professor Gunhee Kim.
- Co-authored research papers:
  - Created a benchmark to evaluate omnimodal language models' multimodal reference resolution capabilities.[1]
  - Developed a speech tokenization method optimized for downstream LLM applications.[2]
  - Designed a benchmark to assess low-level listening abilities of large audio-language models.[3]
  - Implemented incremental response rewriting framework to adapt LLM responses for spoken dialogue systems.[4]

UC Davis RUbiNet Lab **Davis, CA**
Undergraduate Research Student *Jan. 2019 – Jun. 2019*

- Contributed to machine learning powered clinical decision support system under Professor Chen-Nee Chuah.
- Developed a client-side application for iOS that monitors mechanical ventilator data and receives anomaly notifications.[10]

UC Davis DECAL Lab **Davis, CA**
Undergraduate Research Student *Apr. 2018 – Mar. 2019*

- Conducted experiments on integration of entropy scoring for genetic programming based automatic software repair program under the guidance of Professor Prem Devanbu.

Employment

SK Telecom **Seoul, South Korea**
Research Intern *Oct. 2025 – Current*

- Selected to contribute to South Korea's flagship Sovereign AI Foundation Model Project, developing a large-scale national model initiative.
- Designing and implementing the spoken language interface for multimodal large language models, for a more natural human-AI interaction.

Mindlogic Inc. **Seoul, South Korea**
NLP Engineer *Sep. 2022 – Nov. 2023*

- Initiated and developed a prototype of a dialogue system with persona-grounding based on historical chat data, enhancing user personalization.
- Established and utilized LLM-based modules, including prompt engineering techniques, to deliver high-quality business chatbot solutions.

MINDsLab Inc. **Seongnam, South Korea**
Senior AI Scientist *Jan. 2022 – Sep. 2022*

- Led the speech recognition research project, achieving a 68% relative error rate reduction compared to the previous internal model for Korean and English.
- Built and deployed a gRPC-based multilingual text-to-speech service compatible with the company's internal talking face generation system.[9]
- Developed an internal grapheme-to-phoneme process for Japanese and Chinese, enabling the company to expand its text-to-speech service to new markets.[8]

AI Scientist *Dec. 2020 – Jan. 2022*

- Built a text preprocessing pipeline, including web scraping and text cleaning, for language model pre-training.
- Initiated and developed a transformer-based speech recognition service, overseeing model prototyping, training, and evaluation.

University of California, Davis **Davis, CA**
Undergraduate Reader *Jan. 2019 – Mar. 2019*

- Graded coursework materials, including assignments and exams.
- Held office hours to provide guidance on assignments.

Skills

Language: English (Native), Korean (Native), Japanese (Fluent)
Programming: Python, Java, Kotlin, Swift, SQL
Tools: Git, Docker
Frameworks/Libraries: PyTorch, TensorFlow/Keras, Hugging Face Transformers, gRPC, FastAPI, LangChain, vLLM

Projects

EnCLAP: Automated Audio Captioning Model

Researcher *Aug. 2023 – Oct. 2024*

- Developed EnCLAP and EnCLAP++, state-of-the-art audio captioning models leveraging pretrained audio and language models, achieving top-tier performance in a competitive challenge.
- Co-authored a conference paper, technical report, and workshop paper documenting the models and improvements.[5,6,7]

Publications

[1] *DExTER: Can Omnimodal Language Models Resolve Audio-Visual Deixis?*
Sehun Lee, Yoonji Nam, **Sang Hoon Woo**, Gunhee Kim - *Under review*

[2] *SubAlign: Speech Tokenization Aligned with LLM Vocabularies for Spoken Language Modeling*
Kang-wook Kim, Sehun Lee, **Sang Hoon Woo**, Gunhee Kim - *To be submitted to ARR January 2026 Cycle*

[3] *WoW-Bench: Evaluating Fine-Grained Acoustic Perception in Audio-Language Models via Marine Mammal Vocalizations*
Jaeyeon Kim, Heeseung Yun, **Sang Hoon Woo**, Chao-Han Huck Yang, Gunhee Kim - *Submitted to ARR October 2025 Cycle (Reviewer Scores: 3, 3, 3.5)*

[4] *Think, Verbalize, then Speak: Bridging Complex Thoughts and Comprehensible Speech*
Sang Hoon Woo*, Sehun Lee*, Kang-wook Kim, Gunhee Kim - *EMNLP 2025*

[5] *EnCLAP++: Analyzing the EnCLAP Framework for Optimizing Automated Audio Captioning Performance*
Jaeyeon Kim, Minjeong Jeon, Jaeyoon Jung, **Sang Hoon Woo**, Jinjoo Lee - *DCASE Workshop 2024*

[6] *Expanding on EnCLAP with Auxiliary Retrieval Model for Automated Audio Captioning*
Jaeyeon Kim, Jaeyoon Jung, Minjeong Jeon, **Sang Hoon Woo**, Jinjoo Lee - *DCASE 2024 Task 6 (2nd Place)*

[7] *EnCLAP: Combining Neural Audio Codec and Audio-Text Joint Embedding for Automated Audio Captioning*
Jaeyeon Kim, Jaeyoon Jung, Jinjoo Lee, **Sang Hoon Woo** - *ICASSP 2024*

[8] *SANE-TTS: Stable And Natural End-to-End Multilingual Text-to-Speech*
Hyunjae Cho, Wonbin Jung, Junhyeok Lee, **Sang Hoon Woo** - *InterSpeech 2022*

[9] *Talking Face Generation with Multilingual TTS*
Hyoung-Kyu Song*, **Sang Hoon Woo***, Junhyeok Lee, Seungmin Yang, Hyunjae Cho, Youseong Lee, Dongho Choi, Kang-wook Kim - *CVPR 2022 (Demo)*

[10] *Leveraging IoTs and Machine Learning for Patient Diagnosis and Ventilation Management in the Intensive Care Unit*
Gregory B Rehm, **Sang Hoon Woo**, Xin Luigi Chen, Brooks T Kuhn, Irene Cortes-Puch, Nicholas R Anderson, Jason Y Adams, Chen-Nee Chuah - *IEEE Pervasive Computing 2020 (19.3)*

[11] *Mining Vehicle Failure Consumer Reports for Enhanced Service Efficiency*
Ali Khodadadi, Chen-Nee Chuah, **Sang Hoon Woo**, Ashish Dalal - *IEEE VTC2019-Fall*