Sunghyo Chung

sunghyo.chung@gmail.com | sunghyo

Interests

- Multimodal Embedding Models, Multimodal Large Language Models, Human-Centered AI

Education

Korea University *Seoul, Korea*

M.Eng. in Computer Science and Engineering, GPA: 4.33 / 4.5 *Mar. 2017 – Aug. 2019*

- Data and Visual Analytics Lab (DAVIAN), *Advisor: Prof. Jaegul Choo*

Korea University *Seoul, Korea*

B.S. in Computer Science and Engineering, GPA: 4.09 / 4.5 *Mar. 2013 – Feb. 2017*

- Graduated with Great Honors
- Samsung SDS sGen Club, participated in the development of four Android applications.

Experience

Kakao *Seongnam, Korea*

Research Engineer, Kakao *Jun. 2024 - Present*

- **Kanana-v-embedding** (tech blog)
  - Led the development of a multimodal embedding model for cross-modal text-image retrieval.
  - Built KoEmbed, a Korean multimodal retrieval benchmark; outperformed Jina Embeddings v4 (3.8B) across all Korean-centric tasks (77.1 vs 36.7 on KoEmbed-Sent) and English benchmarks (Visual News, MSCOCO).
  - Deployed to Kakao's internal ad similarity and image analysis/review pipeline.
- **Reading.help** (arXiv) *w/ Prof. Naimul Hoque, Prof. Sungbok Shin*
  - Built an LLM-based reading assistant for EFL learners as a human-centered AI project, providing proactive vocabulary, grammar, and comprehension support.
  - Developed interpretable CEFR-level difficulty estimators, outperforming GPT-4o on lexical complexity prediction.

Research Engineer, Kakao Brain *Aug. 2023 - Jun. 2024*

- **Multimodal Large Language Models (MLLM)**
  - Developed a vision-language model for fashion and commerce product retrieval, generating image recommendations from text descriptions.
  - Built high-quality Korean training datasets for multimodal large language model development.

Machine Learning Engineer, Kakao / Kakao Enterprise *Jul. 2019 - Aug. 2023*

- **Optical character recognition (OCR)**
  - Developed scene text detection, recognition, and document understanding models for Kakao's OCR services.
  - Proposed CAFE-Net (arXiv), combining context-aware and context-free experts to improve scene text recognition for character-level long-tailed distributions in large character set languages such as Korean.

Naver Corp. *Seongnam, Korea*

Research Intern at CLOVA Vision (OCR team), Adviser: Dr. Hwalsuk Lee *Mar. 2018 - Aug. 2018*

- **Optical character recognition (OCR)**
  - Worked on scene text detection and script identification tasks.
  - Achieved first place on script identification task on ICDAR MLT 2017 dataset. (Jul. 2018, news article)

Publications

- **Reading.help: Supporting EFL Readers with Proactive and On-Demand Explanation of English Grammar and Semantics**
  Sunghyo Chung, Hyeon Jeon, Sungbok Shin, and Md Naimul Hoque. arXiv 2025
- **Character-Level Long-Tailed Distribution in Scene Text Recognition**
  Sunghyun Park,* **Sunghyo Chung,*** Jungsoo Lee, and Jaegul Choo (*: equal contributions). arXiv 2023

- **A Scanner Deeply: Predicting Gaze Heatmaps on Visualizations Using Crowdsourced Eye Movement Data**
  Sungbok Shin, **Sunghyo Chung**, Sanghyun Hong, and Niklas Elmqvist. IEEE Trans. on Visualization and Computer Graphics (**TVCG**), 2023 (Proc. IEEE VIS'22) (26.5% acceptance rate)
- **AnimeCeleb: Large-Scale Animation CelebHeads Dataset for Head Reenactment**
  Kangyeol Kim,* Sunghyun Park,* Jaeseong Lee,* **Sunghyo Chung**, Junsoo Lee, and Jaegul Choo (*: equal contributions). European Conference on Computer Vision (**ECCV**), 2022, WA (28.4% acceptance rate)
- **VATUN: Visual Analytics for Testing and Understanding Convolutional Neural Networks**
  Cheonbok Park,* Soyoung Yang,* Inyoup Na,* **Sunghyo Chung**, Sungbok Shin, Bum Chul Kwon, Deokgun Park, and Jaegul Choo (*: equal contributions). EG/VGTC Conference on Visualization (**EuroVis**), Short Paper, 2021
- **Exploring Unlabeled Faces for Novel Attribute Discovery**
  Hyojin Bahng, **Sunghyo Chung**, Seungjoo Yoo, and Jaegul Choo. IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**), 2020, WA (22.1% acceptance rate)
- **Learning to Focus and Track Extreme Climate Events**
  Sookyung Kim,* Sunghyun Park,* **Sunghyo Chung,*** Joonseok Lee, Yunsung Lee, Hyojin Kim, Mr Prabhat, and Jaegul Choo (*: equal contributions). British Machine Vision Conference (**BMVC**), 2019, Cardiff, UK, Accepted as Spotlight Presentation (6.9% acceptance rate for spotlight papers).
- **Coloring With Limited Data: Few-Shot Colorization via Memory-Augmented Networks**
  Seungjoo Yoo, Hyojin Bahng, **Sunghyo Chung**, Junsoo Lee, Jaehyuk Chang, and Jaegul Choo. IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**), 2019, CA (25.2% acceptance rate).
- **ReVACNN: Steering Convolutional Neural Network via Real-Time Visual Analytics**
  Sunghyo Chung, Cheonbok Park, Sangho Suh, Kyeongpil Kang, Jaegul Choo, and Bum Chul Kwon. Future of Interactive Learning Machines Workshop at NIPS (**NIPS'16-FILM**), 2016.

Honors & Awards

2017	**Great Honor,** Winter 2017 Graduation	*Korea University*
2015-16	**National Science and Engineering Scholarship,** 2015, 2016	*Korea Student Aid Foundation*
2013-16	**Semester High Honors,** 1st Semester 2013, 1st Semester 2014, 1st and 2nd Semester 2015, 1st Semester2016	*Korea University*
2015	**Dean's List,** 2nd Semester 2015	*Korea University*