

# Youngjoon Jang

Postdoctoral Researcher, VGG, University of Oxford.

Email : youngjoon@robots.ox.ac.uk

Webpage: <https://art-jang.github.io>

## RESEARCH INTEREST

- My research aims to effectively train deep neural networks with multi-modality (vision, audio and text). In particular, I am interested in automatic sign language recognition for helping deaf people.

## EDUCATION

<b>Korea Advanced Institute of Science and Technology (KAIST)</b> Ph.D. in Division of Future Vehicle, School of Electrical Engineering; Advisor: Joon Son Chung	Daejeon, South Korea Sep. 2022 – Present
<b>Korea Advanced Institute of Science and Technology (KAIST)</b> M.S. in Division of Future Vehicle, School of Electrical Engineering; Advisor: In So Kweon ◦ Thesis: Learning Methodology According to Characteristics of Continuous Sign Language Recognition Dataset	Daejeon, South Korea Mar. 2020 – Feb. 2022
<b>Kwangwoon University</b> B.S. in Division of Robotics; GPA: 4.3/4.5	Seoul, South Korea Mar. 2014 – Feb. 2020

## WORK EXPERIENCE

<b>University of Oxford</b> Postdoctoral Researcher, VGG group	Oxford, UK Nov. 2025 - current
<b>Korea Advanced Institute of Science and Technology (KAIST)</b> Postdoctoral Researcher, Multimodal AI Lab.	Daejeon, South Korea Sep. 2025 - Nov. 2025
<b>University of Oxford</b> Visiting student, VGG group	Oxford, UK May. 2024 - Nov. 2024
<b>Korea Advanced Institute of Science and Technology (KAIST)</b> Researcher, Multimodal AI Lab.	Daejeon, South Korea Mar. 2022 - Aug. 2022

## PUBLICATION

### International Journals

- [J1] Deep Understanding of Sign Language for Sign to Subtitle Alignment.  
*Youngjoon Jang\**, *Jeongsoo Choi\**, *Junseok Ahn*, and *Joon Son Chung*  
◦ IEEE Transactions on Multimedia (TMM), 2025.

### International Conferences

- [C18] LP-CFM: Perceptual Invariance-Aware Conditional Flow Matching for Speech Modeling.  
*Doyeop Kwak*, *Youngjoon Jang*, and *Joon Son Chung*  
◦ International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2026.
- [C17] FastAV: Efficient Token Pruning for Audio-Visual Large Language Model Inference.  
*Chaeyoung Jung*, *Youngjoon Jang*, and *Joon Son Chung*  
◦ International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2026.
- [C16] AVCD: Mitigating Hallucinations in Audio-Visual Large Language Models through Contrastive Decoding.  
*Chaeyoung Jung\**, *Youngjoon Jang\**, and *Joon Son Chung*  
◦ Neural Information Processing Systems (NeurIPS), 2025.
- [C15] Test-Time Augmentation for Pose-invariant Face Recognition.  
*Jaemin Jung\**, *Youngjoon Jang\**, and *Joon Son Chung*  
◦ International Conference on Automatic Face and Gesture Recognition (FG), 2025.
- [C14] Lost in Translation, Found in Context: Sign Language Translation with Contextual Cues.  
*Youngjoon Jang\**, *Haran Raajesh\**, *Lilane Momeni*, *Gül Varol*, and *Andrew Zisserman*  
◦ Computer Vision and Pattern Recognition Conference (CVPR), 2025.
- [C13] VoiceDiT: Dual-Condition Diffusion Transformer for Environment-Aware Speech Synthesis.  
*Jaemin Jung\**, *Junseok Ahn\**, *Chaeyoung Jung*, *Tan Dat Nguyen*, *Youngjoon Jang*, and *Joon Son Chung*  
◦ International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2025.
- [C12] Let Me Finish My Sentence: Video Temporal Grounding with Holistic Text Understanding.  
*Jongbin Woo*, *Hyeonggon Ryu*, *Youngjoon Jang*, *Jae Won Cho*, and *Joon Son Chung*  
◦ ACM International Conference on Multimedia (ACMMM), 2024.

- [C11] Faces that Speak: Jointly Synthesising Talking Face and Speech from Text.  
*Youngjoon Jang\**, *Ji-Hoon Kim\**, *Junseok Ahn*, *Doyeop Kwak*, *Hongsun Yang*, *Yooncheol Ju*, *ILHWAN KIM*, *Byeong-Yeol Kim*, and *Joon Son Chung*  
◦ Computer Vision and Pattern Recognition Conference (CVPR), 2024.
- [C10] Slowfast Network for Continuous Sign Language Recognition.  
*Junseok Ahn\**, *Youngjoon Jang\**, and *Joon Son Chung*  
◦ International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2024.
- [C9] VoxMM: Rich Transcription of Conversations in the Wild.  
*Doyeop Kwak*, *Jaemin Jung*, *Kihyun Nam*, **Youngjoon Jang**, *Jee-weon Jung*, *Shinji Watanabe*, and *Joon Son Chung*  
◦ International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2024.
- [C8] FreGrad: Lightweight and Fast Frequency-aware Diffusion Vocoder.  
*Tan Dat Nguyen*, *Ji-Hoon Kim*, **Youngjoon Jang**, *Jaehun Kim*, and *Joon Son Chung*  
◦ International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2024.
- [C7] TalkNCE: Improving Active Speaker Detection with Talk-aware Contrastive Learning.  
*Chae young Jung*, *Suyeon Lee*, *Kihyun Nam*, *Kyeongha Rho*, *You Jin Kim*, **Youngjoon Jang**, and *Joon Son Chung*  
◦ International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2024.
- [C6] Seeing Through the Conversation: Audio-visual Speech Separation based on Diffusion Model.  
*Suyeon Lee*, *Chae young Jung*, **Youngjoon Jang**, *Jaehun Kim*, and *Joon Son Chung*  
◦ International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2024.
- [C5] That's What I said: Fully-Controllable Talking Face Generation.  
**Youngjoon Jang\***, *Kyeongha Rho\**, *Jongbin Woo*, *Hyeongkeun Lee*, *Jihwan Park*, *Youshin Lim*, *Byeong-Yeol Kim*, and *Joon Son Chung*  
◦ ACM International Conference on Multimedia (ACMMM), 2023.
- [C4] Self-Sufficient Framework for Continuous Sign Language Recognition.  
**Youngjoon Jang**, *Youngtaek Oh*, *Jae Won Cho*, *Myungchul Kim*, *Dong-Jin Kim*, *In So Kweon*, and *Joon Son Chung*  
◦ International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2023. (**Top 3% Paper Recognition**)
- [C3] Metric Learning for User-Defined Keyword Spotting.  
*Jaemin Jung\**, *Youkyum Kim\**, *Jihwan Park*, *Youshin Lim*, *Byeong-Yeol Kim*, **Youngjoon Jang**, and *Joon Son Chung*  
◦ International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2023.
- [C2] Signing Outside the Studio: Benchmarking Background Robustness for Continuous Sign Language Recognition.  
**Youngjoon Jang**, *Youngtaek Oh*, *Jae Won Cho*, *Dong-Jin Kim*, *Joon Son Chung*, and *In So Kweon*  
◦ British Machine Vision Conference (BMVC), 2022.
- [C1] KSL-Guide: A Large-scale Korean Sign Language Dataset Including Interrogative Sentences for Guiding the Deaf and Hard-of-Hearing.  
*Soomin Ham*, *Kibaek Park*, **Youngjoon Jang**, *Youngtaek Oh*, *Seokmin Yun*, *Sukwon Yoon*, *Chang Jo Kim*, *Han-Mu Park*, and *In So Kweon*  
◦ International Conference on Automatic Face and Gesture Recognition (FG), 2021.

## AWARDS & HONORS

---

### International Competitions

- *1st Place*, R-BIZ Challenge TURTLEBOT3 AUTORACE, Nov. 2018
- *5th Place*, Sumo Robot, International Robot Contest (IRC), Oct. 2018
- *3rd Place*, R-BIZ Challenge TURTLEBOT3 AUTORACE, Sep. 2017
- *5th Place*, RoboCup Iran Open Rescue, Apr. 2017

### National Competitions

- *2nd Place*, Science and Technology Specialized University Startup Competition (GIST), Nov. 2021
- *2nd Place*, App Startup Support Program Contest (KAIST), Apr. 2021
- *1st Place*, RoboCup Korea Open Rescue, Feb. 2017

## TEACHING

---

### Teaching Assistance (TA) at FV, KAIST

- \* PD513: Future Vehicle Capstone Design (Spring, 2023)
- \* PD513: Future Vehicle Capstone Design (Fall, 2022)
- \* PD806: Automobile Special Topics in Mechanical Engineering (Fall, 2021)

## TECHNICAL SKILLS

---

**Programming:** C, C++, Python, Pytorch