

# Jeongsoo Choi

Ph.D. candidate, Image and Video Systems Lab.  
Korea Advanced Institute of Science and Technology (KAIST)

## Contact

---

Email	jeongsoo.choi@kaist.ac.kr
Webpage	<a href="https://choijeongsoo.github.io">https://choijeongsoo.github.io</a>
Phone	+82-10-2106-5748

## Research Interests

---

**Deep Learning:** multi-modal learning, self-supervised learning

**Speech Processing:** speech recognition, speech synthesis

**Computer Vision:** generative modeling, video synthesis

## Education

---

<b>Ph.D.</b> in Electrical Engineering, <b>KAIST</b> - Advisor: Prof. Joon Son Chung	Sep. 2020 - Present
<b>B.S.</b> in Electrical Engineering, <b>KAIST</b>	Mar. 2015 - Aug. 2020

## Publications

---

### Preprints

P2. Minsu Kim\*, Jeong Hun Yeo\*, **Jeongsoo Choi**, Se Jin Park, Yong Man Ro  
“Multilingual Visual Speech Recognition with a Single Model by Learning with Discrete Visual Speech Units”  
Under review, **arXiv 2024**

P1. Minsu Kim\*, **Jeongsoo Choi\***, Dahun Kim, Yong Man Ro  
“Many-to-Many Spoken Language Translation via Unified Speech and Text Representation Learning with Unit-to-Unit Translation”  
Under review, **arXiv 2023**

### Peer-Reviewed Conferences

C9. **Jeongsoo Choi\***, Se Jin Park\*, Minsu Kim\*, Yong Man Ro  
“AV2AV: Direct Audio-Visual Speech to Audio-Visual Speech Translation with Unified Audio-Visual Speech Representation”  
**CVPR 2024**

C8. **Jeongsoo Choi**, Minsu Kim, Se Jin Park, Yong Man Ro  
“Text-driven Talking Face Synthesis by Reprogramming Audio-driven Models”  
**ICASSP 2024**

C7. Minsu Kim, **Jeongsoo Choi**, Soumi Maiti, Jeong Hun Yeo, Shinji Watanabe, Yong Man Ro  
“Towards Practical and Efficient Image-to-Speech Captioning with Vision-Language Pre-training and Multi-modal Tokens”  
**ICASSP 2024**

C6. Se Jin Park, Minsu Kim, **Jeongsoo Choi**, Yong Man Ro  
“Exploring Phonetic Context-Aware Lip-Sync For Talking Face Generation”  
**ICASSP 2024**

C5. **Jeongsoo Choi\***, Joanna Hong\*, Yong Man Ro  
“DiffV2S: Diffusion-based Video-to-Speech Synthesis with Vision-guided Speaker Embedding”  
**ICCV 2023**

C4. Minsu Kim\*, Jeong Hun Yeo\*, **Jeongsoo Choi**, Yong Man Ro  
“Lip Reading for Low-resource Languages by Learning and Combining General Speech Knowledge and Language-specific Knowledge”  
**ICCV 2023**

C3. **Jeongsoo Choi**, Minsu Kim, Yong Man Ro  
“Intelligible Lip-to-Speech Synthesis with Speech Units”  
**Interspeech 2023**

C2. Joanna Hong\*, Minsu Kim\*, **Jeongsoo Choi**, Yong Man Ro  
“Watch or Listen: Robust Audio-Visual Speech Recognition With Visual Corruption Modeling and Reliability Scoring”  
**CVPR 2023**

C1. Se Jin Park, Minsu Kim, Joanna Hong, **Jeongsoo Choi**, Yong Man Ro  
“SyncTalkFace: Talking Face Generation with Precise Lip-Syncing via Audio-Lip Memory”  
**AAAI 2022**

#### **Peer-Reviewed Journals**

J1. Jeong Hun Yeo, Minsu Kim, **Jeongsoo Choi**, Dae Hoe Kim, Yong Man Ro  
“AKVSR: Audio Knowledge Empowered Visual Speech Recognition by Compressing Audio Knowledge of a Pretrained Model”  
**TMM 2024: IEEE Transactions on Multimedia**

#### **Programming Skills**

---

##### **Languages**

- Python, C/C++, and MATLAB

##### **Deep Learning Libraries**

- Pytorch and Tensorflow

#### **References**

---

##### **Prof. Joon Son Chung**

- School of Electrical Engineering, KAIST  
- Email: joonson@kaist.ac.kr  
- Homepage: <https://mmai.io/joon>  
- Relationship: Ph.D. advisor at KAIST