HyoJung Han

한 효 정

Ph.D. in Computer Science at University of Maryland, College Park (UMD)

hjhan@cs.umd.edu / Google Scholar / Homepage / LinkedIn

EDUCATION

University of Maryland, College Park (UMD)**,** Maryland, United States — Aug. 2021 – **Present**
Department of Computer Science (Ph.D.)
- Computational Linguistics and Information Processing (CLIP) Lab
- Graduate Assistant *(full-time)**, advised by Prof. Marine Carpuat and Prof. Jordan Boyd-Graber***

KAIST (Korea Advanced Institute of Science and Technology), Daejeon, Korea — Mar. 2016 – Feb. 2018
School of Electrical Engineering (M.S.)
- Artificial Intelligence & Machine Learning (AIM) Lab

Korea University, Seoul, Korea — Mar. 2011 – Feb. 2016
School of Electrical Engineering (B.S.)

PROFESSIONAL EXPERIENCE

Samsung Research, R&D hub of Samsung Electronics, Seoul, Korea *(full-time)* — Jul. 2018 – Jul. 2021
Efficient Neural Machine Translation Team, Natural Language Processing Lab
- ♡ Simultaneous Machine Translation
- ☆ End-to-End Speech Translation
- ♢ Low-Resource Machine Translation
- ♧ Quality Estimation Task / Automatic Post-Editing
- Development of In-house Neural Machine Translation Model

Naver, Seongnam, Korea *(full-time)* — Apr. 2018 – Jul. 2018
Clova AI Research
- Automatic Advertisement Evaluation/Generation with Computer Vision & Reinforcement Learning

PUBLICATION (* as same contribution)

HyoJung Han, Marine Carpuat and Jordan Boyd-Graber. *SimQA: Detecting Simultaneous MT Errors through Word-by-Word Question Answering*. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2022, Online and Abu Dhabi, United Arab Emirates, Association for Computational Linguistics. **To Be Appear in **EMNLP 2022 main conference** and Arxiv**

♥ **HyoJung Han***, Seokchan Ahn*, Yoonjung Choi, Insoo Chung, Sangha Kim and Kyunghyun Cho. *Monotonic Simultaneous Translation with Chunk-wise Reordering and Refinement*. In Proceedings of the Sixth Conference on Machine Translation (WMT), 2021. Association for Computational Linguistics. [pdf] [video]

★Sathish Indurthi, Mohd Abbas Zaidi, Nikhil Kumar Lakumarapu, Beomseok Lee, **Hyojung Han**, Seokchan Ahn, Sangha Kim, Chanwoo Kim and Inchul Hwang. *Task Aware Multi-Task Learning for Speech to Text Tasks*. In ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 7723-7727, 2021. [pdf]

♥ **HyoJung Han***, Sathish Indurthi***,** Mohd Abbas Zaidi, Nikhil Kumar Lakumarapu, Beomseok Lee**,** Sangha Kim, Chanwoo Kim and Inchul Hwang**.** *Faster Re-translation Using Non-Autoregressive Model for Simultaneous Neural Machine Translation*. arXiv preprint arXiv:2012.14681. [pdf]

(**HyoJung Han since July 2020, Previously HouJeung Han.)

♥★**Hou Jeung Han**, Mohd Abbas Zaidi, Sathish Reddy Indurthi, Nikhil Kumar Lakumarapu, Beomseok Lee and Sangha Kim**.** *End-to-End Simultaneous Translation System for IWSLT2020 Using Modality Agnostic Meta-Learning*. In Proceedings of the 17th International Conference on Spoken Language Translation (IWSLT), pages 62-68, Online, July 2020. Association for Computational Linguistics. [pdf] [video]

★Nikhil Kumar Lakumarapu, Beomseok Lee, Sathish Reddy Indurthi, **Hou Jeung Han**, Mohd Abbas Zaidi and Sangha Kim**.** *End-to-End Offline Speech Translation System for IWSLT 2020 using Modality Agnostic Meta-Learning*. In Proceedings of the 17th International Conference on Spoken Language Translation (IWSLT), pages 73-79, Online, July 2020. Association for Computational Linguistics. [pdf] [video]

★Sathish Indurthi, **Houjeung Han**, Nikhil Kumar Lakumarapu, Beomseok Lee, Insoo Chung, Sangha Kim and Chanwoo Kim**.** *End-end Speech-to-Text Translation with Modality Agnostic Meta-Learning*. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pages 7904-7908, 2020. [pdf] [video]

◆ **HouJeung Han***, Sathish Indurthi* and Sangha Kim**.** *Exploiting General Purpose Sequence Representations for Low Resource Neural Machine Translation*. Practical ML for Developing Countries Workshop at International Conference on Learning Representations (PML4DC@ICLR), 2020. [pdf] [slide] [video]

★Sathish Indurthi, **Houjeung Han**, Nikhil Kumar Lakumarapu, Beomseok Lee, Insoo Chung, Sangha Kim and Chanwoo Kim**.** *Data Efficient Direct Speech-to-Text Translation with Modality Agnostic Meta-Learning*. arXiv preprint arXiv:1911.04283. [pdf]

♣ Jaehun Shin, WonKee Lee, Youngkil Kim, **H. Jeung Han** and Jong-Hyeok Lee. *Research on the Decoder Attention Structure of Multi-encoder Transformer-based Automatic Post-Editing Model*. KIISE Transactions on Computing Practices (KTCP), Vol.26, No. 8. pp. 367-372, 2020. [pdf] (In Korean)

HouJeung Han and Chang D. Yoo. *Associative scene semantic segmentation*. Master's Thesis at Korea Advanced Institute of Science and Technology, 2018. [pdf]

HouJeung Han, Sunghun Kang and Chang D. Yoo. *Multi-view Visual Speech Recognition Based on Multi-task Learning*. In 2017 IEEE International Conference on Image Processing (ICIP), pages 3983-3987, 2017. [pdf] [slide] [code]

CHALLENGES

IWSLT 2020 The International Conference on Spoken Language Translation ACL - 17th IWSLT 2020

- ♥★***Top Score*** in low-latency regime Speech Track, for the Simultaneous Speech Translation Challenge
- Team SRSK. Sole End-to-End Speech Translation System among the participants

ImageNet 2016 Large Scale Visual Recognition Challenge 2016 (ILSVRC2016)

- Team KAIST-SLSP. Poster presentation in ECCV 2016 Workshop [poster]
- Object Detection (DET) : International Rank 7th, Domestic Rank 2nd
- Video Detection (VID) : International Rank 5th, Domestic Rank 1st

HONORS & AWARDS

Dean's Fellowship, University of Maryland 2021 - 2022

Korean Government Scholarship Program for Study Overseas 2021 - 2022

Tenure of two years granted by National Institute For International Education, Ministry of Education

Samsung Research Software Innovation Award, Samsung Research

★★★♥★'Speech Translation' project won the award in the 'Innovative Technology' June. 2021

Samsung Best Paper Awards 2020, Samsung Group Oct. 2020

Bronze, Artificial Intelligence Section

♥ **HyoJung Han***, Sathish Indurthi*, Beomseok Lee, Mohd Abbas Zaidi, Nikhil Kumar Lakumrapu and Sangha Kim. *Faster Re-translation using Non-Autoregressive Model for Simultaneous Neural Machine Translation. Samsung Best Paper Awards (SBPA). 2020.*

Government-Sponsored Scholarship during entire Master program at KAIST 2016 - 2017
National Science and Engineering Scholarship, Full Tuition 2013 - 2015
Academic Excellence Scholarship, Half Tuition Mar. 2012

TEACHING EXPERIENCE

Teaching Assistant, Object-Oriented Programming I (Java), CMSC131, UMD Aug. 2021 – Dec. 2021
Teaching Assistant, Introduction to Machine Learning, EE488, KAIST Sep. 2017 – Dec. 2017
Teaching Assistant, Statistical Learning Theory, EE531, KAIST Mar. 2017 – June. 2017
Tutor, Programming Structures for Electrical Engineering (C/C++), EE291, KAIST Mar. 2017 – June. 2017
Tutor, Programming Structures for Electrical Engineering (C/C++), EE291, KAIST Sep. 2016 – Dec. 2016

LANGUAGE SKILLS

- Language Proficiency : Korean (Native), English (Fluent), Japanese (Intermediate)
- Programming Language : Most experienced in Python (Pytorch, Tensorflow, Keras, Theano for deep learning), Java, Torch in Lua, C/C++
- *Professional Certification* for C/C++ in Samsung Electronics Software Competency Test

PATENT

- Korea Copyright Commission, two Software Patents
  - Multi-view Visual Speech Recognition Based on Multi-task Learning, C-2018-001350, *KAIST*
  - Associative Scene Semantic Segmentation, C-2018-001351, *KAIST*
- Electronic device and controlling method of electronic device, US20210097242A1, KR10-2019-0119791, *Samsung Research*
- Electronic apparatus and method for controlling electronic apparatus, US20220284018A1, *Samsung Research*

♣ Apparatus and Method for Training Model for Translation Quality Estimation, KR10-2020-0115333, *POSTECH*

★Electronic device and method for controlling the electronic device thereof, US20210065690A1, KR10-2019-0151482, *Samsung Research*

♥ Fixed prefix paraphrase for simultaneous latency reduction for Cascade Speech-to-Speech Translation system, US20220245364A1, *Samsung Research*

♥ Chunk-wise Decision Making in Neural Machine Translation, Under Application process, P2020-0184579, *Samsung Research*