

Chanjun Park

NATURAL LANGUAGE PROCESSING & ARTIFICIAL INTELLIGENCE LAB · KOREA UNIVERSITY

311 Aegineung Student Center, College of Informatics, Korea University, 145 Anam-ro, Seongbuk-gu, Seoul, 02841, Korea

☎ (+82) 10-2580-9066 | ✉ bcj1210@naver.com | 📷 parkchanjun | 📧 bcj1210

Research Interests

I am a MS/Ph.D student in the Department of Computer Science and Engineering at Korea University. My advisor is prof. Heuiseok Lim, and I am part of the Natural Language Processing & Artificial Intelligence Lab.

- Neural Machine Translation
- Domain Specialized Neural Machine Translation
- Ancient Korean Neural Machine Translation
- Simultaneous Speech Translation
- Automatic Post Editing
- Quality Estimation
- Korean Spelling Correction
- Grammar Error Correction
- Speech Recognition Post-Processor
- AI for Accessibility
- AI for Education

Education

Korea University

M.S/PH.D IN COMPUTER SCIENCE AND ENGINEERING, GPA: 4.5 / 4.5

- Advisor: Heuiseok Lim
- Natural Language Processing & Artificial Intelligence Lab

Seoul, South Korea

Sep. 2019 - Current

Busan University of Foreign Studies(BUFS)

B.S. IN NATURAL LANGUAGE PROCESSING & CREATIVE CONVERGENCE, GPA: 4.2 / 4.5

- Advisor: Pummo Ryu
- Full Scholarship

Busan, South Korea

Mar. 2015 - Exp. Aug. 2018

Honors & Awards

2020.11	1st Place in Flitto Hackathon (Team Lead), 2020 , Flitto	South Korea
2020.10	Best Paper Award, 2020 , HCLT 2020	South Korea
2020.05	Best practices for using NIA AI training data(Korean-English Neural Machine Translation model) , AI Hub	South Korea
2019.10	Best Paper Award, 2019 , HCLT 2019	South Korea
2019.10	1st Place Microsoft AI Accessibility Hackathon in Korea (Team Lead) , Microsoft	South Korea
2019.03	Graduate School Associate Scholarship , Sungkyunkwan University	South Korea
2018.10	Next Generation Information Processing NLP Competition 2018: Participation Award , Next-generation information computing technology development business	South Korea
2017.06	Bit Computer Excellence Award (President Award) , Bit Computer	South Korea
2015-2019	Full Scholarship , BUFS	South Korea
2016-2017	Scholarship for academic excellence , Sooyoungro Church	South Korea

Selected Publications

Should we find another model?: Improving Neural Machine Translation Performance with ONE-Piece Tokenization Method without Model Modification

Chanjun Park, Sugyeong Eo, Hyeonseok Moon, Heui-Seok Lim

Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2021): Human Language Technologies: Industry Papers, 2021

BTS: Back TranScripton for Speech-to-Text Post-Processor using Text-to-Speech-to-Text

Chanjun Park, Jaehyung Seo, Seolhwa Lee, Chanhee Lee, Hyeonseok Moon, Sugyeong Eo, Heui-Seok Lim

ACL 2021 - WAT(Workshop on Asian Translation) 2021, 2021

Exploring the Data Efficiency of Cross-Lingual Post-Training in Pretrained Language Models

Chanhee Lee, Kisu Yang, Taesun Whang, **Chanjun Park**, Andrew Matteson, Heuseok Lim

Applied Sciences 11.5 (2021) p. 1974. Multidisciplinary Digital Publishing Institute, 2021

Neural Spelling Correction: Translating Incorrect sentences to Correct sentences for Multimedia

Chanjun Park, Kuekyeng Kim, YeongWook Yang, Minho Kang, HeuiSeok Lim

Multimedia Tools and Applications (2020) pp. 1–18. Springer, 2020

Ancient Korean Neural Machine Translation

Chanjun Park, Chanhee Lee, YeongWook Yang, Heuseok Lim

IEEE Access 8 (2020) pp. 116617–116625. 2020

Comparison of the evaluation metrics for Neural Grammatical Error Correction with Overcorrection

Chanjun Park, YeongWook Yang, Chanhee Lee, Heuseok Lim

IEEE Access 8 (2020) pp. 106264–106272. 2020

Decoding Strategies for Improving Low-Resource Machine Translation

Chanjun Park, Yeongwook Yang, Kinam Park, Heuseok Lim

Electronics 9.10 (2020) p. 1562. Multidisciplinary Digital Publishing Institute, 2020

Selected Research Experience

Subtitle production and editing technology development for colloquial broadcasting contents

Seoul, South Korea

RESEARCHER IN KOREA UNIV

April. 2021. - December, 2024.

- Development of Colloquial Aware Neural Machine Translation (CANMT) model that is robust to colloquial expression based on broadcast content
- Consortium: LLsoLLu, Korea Univ, Hanyang Univ, ON THE IT, SBS Contents Hub, CBS
- Supported by IITP
- Korean Colloquial-Sentence-Splitter(Ko-CSS) for simultaneous speech translation
- Neural Architecture Search(NAS) based NMT Research
- Speech recognition post-processor research
- Cross lingual adaptation research
- Skills: Neural Machine Translation, Automatic Speech Recognition, simultaneous speech translation, Pytorch, Huggingface, Fairseq
- Ko-CSS Demo: <http://nlp1lab.iptime.org:32285/>
-

BTS: Speech recognition post-processor research

Seoul, South Korea

RESEARCHER IN KOREA UNIV

January. 2021. - Present

- Proposed BackTranScripton(BTS)
- Proposed Text to Speech to Text (TST) methodology
- Skills: Neural Machine Translation, Pytorch, Huggingface, Fairseq
- Publication of research paper in WAT 2021 (ACL Workshop): <https://lotus.kuee.kyoto-u.ac.jp/WAT/WAT2021/index.html>

Cross Lingual Pretrained Language Model Research

Seoul, South Korea

RESEARCHER IN KOREA UNIV

January. 2021. - Present

- Research for Exploring the Data Efficiency of Cross-Lingual Post-Training in Pretrained Language Models
- Research for combining BART and Parallel Corpus Filtering
- Skills: Neural Machine Translation, Pytorch, Huggingface, Fairseq
- Publication of research paper in Applied Science (SCIE): <https://www.mdpi.com/2076-3417/11/5/1974>

Automatic Post Editing Research

Seoul, South Korea

RESEARCHER IN KOREA UNIV

January. 2021. - Present

- APE research based on Cross Lingual Pretrained Language Model
- APE data augmentation technique research
- Skills: Neural Machine Translation, Pytorch, Automatic Post Editing, Huggingface, Fairseq

Quality Estimation Research

Seoul, South Korea

RESEARCHER IN KOREA UNIV

January. 2021. - Present

- XLM-R-based quality estimation development and research
- Q.E data augmentation technique development
- Skills: Neural Machine Translation, Pytorch, Quality Estimation, OpenKiwi, Huggingface, Fairseq

ONE-Piece: A new tokenization method for Neural Machine Translation

Seoul, South Korea

RESEARCHER IN KOREA UNIV

September. 2020. - January. 2021.

- Our paper is accepted in NAACL 2021 Industrial Track: "Should we find another model?: Improving Neural Machine Translation Performance with ONE-Piece Tokenization Method without Model Modification"
- How to improve the performance of NLP application software without changing the model through data pre- and post-processing
- Skills: Neural Machine Translation, Pytorch, Python, Subword Tokenization, Parallel Corpus Filtering
- Publication of research paper in NAACL 2021: <https://www.aclweb.org/anthology/2021.naacl-industry.13/>

COVID-19 Machine Translation Platform

Seoul, South Korea

RESEARCHER IN KOREA UNIV

May. 2020. - November.2020

- Neural Machine translation specialized for COVID-19 domain
- Skills: Neural Machine Translation, Pytorch, Python
- Publication of research paper in Journal of the Korea Convergence Society(KCI): <https://bit.ly/2QgQKLK>

Domain Specialized Neural Machine Translation (Security)

Seoul, South Korea

RESEARCHER IN KOREA UNIV

May. 2020. - November.2020

- Research on Neural Machine Translation specialized for security domain
- Skills: Neural Machine Translation, Pytorch, Python
- Supported by "National Security Research Institute"
- Project Manager, Tech Lead
- Development of the latest translation model based on Transformer/BERT, AI-based Korean-English translation model suitable for the security field, and research trend of the latest AI-based translation technology research
- We propose Cross Communication Method (CCM), a new DS-NMT training methodology, inspired by cognitive scientific interpretation of cross language speech perception. This is the first study to reflect cognitive science theory in DS-NMT training methodology.

A Neural-Symbolic Model for Knowledge Acquisition and Inference Techniques

Seoul, South Korea

RESEARCHER IN KOREA UNIV

April. 2020. - Present

- Neural-Symbolic Model Research and Development
- Skills: Neural Symbolic Computing, Knowledge base, Pytorch, Python
- Supported by IITP
- Demo: <http://nlplab.iptime.org:32245/>

Ancient Korean Neural Machine Translation

Seoul, South Korea

RESEARCHER IN KOREA UNIV

March. 2020. - Present

- Ancient Korean Neural Machine Translation Research
- Bilingual Word Embedding for Ancient Korean Neural Machine Translation
- Priming Ancient Korean Neural Machine Translation Research
- Skills: Pytorch, Python, REST API, Flask, Neural Machine Translation
- Demo: <http://nlplab.iptime.org:32242/>
- Publication of research paper in IEEE Access(SCIE): <https://ieeexplore.ieee.org/document/9125904/>

Deep Learning based Grammar Error Correction

Seoul, South Korea

RESEARCHER IN KOREA UNIV

June. 2019. - June.2020

- Deep learning based English grammar corrector development
- Skills: Pytorch, Python, REST API, Flask, Neural Machine Translation
- Demo: <http://nlplab.iptime.org:32292/>
- Publication of research paper in IEEE Access(SCIE): <https://ieeexplore.ieee.org/document/9102992>

Deep Learning based Korean Spelling Correction

Seoul, South Korea

RESEARCHER IN KOREA UNIV

June. 2019. - April.2020

- Development of Korean spelling corrector based on deep learning
- Skills: Pytorch, Python, REST API, Flask, Neural Machine Translation
- Demo: <http://nlplab.iptime.org:32288/>
- Publication of research paper in multimedia tools and applications(SCIE): <https://link.springer.com/article/10.1007/s11042-020-09148-2>

Multilingual Neural Machine Translation

Seoul, South Korea

RESEARCHER IN KOREA UNIV

June. 2019. - January. 2021.

- Deep learning based Korean-English and English-Korean, Korean-Japanese machine translator development
- Research for Parallel Corpus Filtering, Data Augmentation, Various decoding strategies
- Skills: Pytorch, Python, REST API, Flask, Neural Machine Translation
- Demo: <http://nlplab.iptime.org:32296/>
- Best practices for using NIA AI training data: <http://www.aihub.or.kr/node/4525>
- Publication of research paper in ELECTRONICS(SCIE): <https://www.mdpi.com/2079-9292/9/10/1562#cite>

Development of English Free-talking system and grammar correction feedback system using NAO robot

Seoul, South Korea

RESEARCHER IN KOREA UNIV

April. 2019. - Dec.2019

- Team Leader
- Demo: <https://www.youtube.com/watch?v=pSBXouica-A>
- Skills: Pytorch, Python, Google Cloud Platform, REST API, Nao Toolkit

AI for Accessibility Research

Seoul, South Korea

RESEARCHER IN KOREA UNIV

June. 2019. - Present

- PickTalky: Design of augmentative alternative communication (AAC) software for children with developmental disabilities
- 1st Place Microsoft AI Accessibility Hackathon in Korea (Team Lead)
- Award: <https://bit.ly/2vPaZ91>
- PickTalky Demo: <https://pictalky.df.r.appspot.com/>

Research and development of human-inspired multiple intelligence

Seoul, South Korea

RESEARCHER IN KOREA UNIV

September. 2019. - Present

- Human-inspired language understanding model
- Funded by the Korea Ministry of Science and ICT (MSIT)
- Site: <http://hi.ai.co.kr/>

Next-generation information computing technology development business

Seoul, South Korea

RESEARCHER IN KOREA UNIV

September. 2019. - January. 2021.

- Development of Core Application Techniques for Korean Intelligent Systems
- Research and development of Korean natural language processing core technology

Core technology development of the real-time simultaneous speech translation based on knowledge enhancement

Seoul, South Korea

RESEARCHER IN SYSTRAN

June. 2018. - July.2019

- Collaboration with the Korea Electronics and Telecommunications Research Institute (ETRI) Language Intelligence Research Group
- Director: Dr. Young-Gil Kim (ETRI)
- 3rd year: Real-time lecture interpretation solution development
- 4th year: Busan University of Foreign Studies Field Group Test (FGT) progress and software maintenance
- Skills: JAVA, JAVA FX, Kaldi

Neural Machine Translation service for SK Telecom's in-house education program

Seoul, South Korea

RESEARCHER IN SYSTRAN

April. 2019. - April.2019

- Development of automation software to provide translation service to SK Telecom's in-house education program
- Skills: Python

LCloud-API Platform Service Development

Seoul, South Korea

RESEARCHER IN SYSTRAN

January. 2019. - June.2019

- Planning and development of natural language processing API Platform Service
- Gravitee IO-based platform design, planning, marketing materials production, strategy establishment
- REST API management using Swagger
- Transcription API, Translation API, Address Correction API, Speaker Verification API
- Learned Microservices architecture(MSA), IT Infrastructure
- Skills: Gravitee IO, Microservices architecture(MSA), IT Infrastructure, Mockup, Marketing, Strategic Planning, Technology management, Amazon AWS

Speech recognition transcription tool development

Seoul, South Korea

RESEARCHER IN SYSTRAN

January. 2019. - July. 2019

- Development and maintenance of speech recognition transcription tool, modification and addition of various functions
- Skills: Front End, Spring Boot, Flask, AJAX, MySQL, Kaldi,

CJ Logistics Speech Recognition Agent

Seoul, South Korea

RESEARCHER IN SYSTRAN

Aug. 2018. - July. 2019

- CJ Logistics artificial intelligence (AI) virtual assistant project
- Skills: Quality Assurance, Quality control, Mockup, Strategic Planning, Kaldi

Automatic subtitle generation platform service development and planning

Seoul, South Korea

RESEARCHER IN SYSTRAN

June. 2018. - June.2019

- Development and maintenance of automatic subtitle generation platform
- Establish platform service strategies such as B2B and B2C
- Various company contacts and SWOT analysis
- Skills: JAVA, JAVA FX, Kaldi, Mockup, Strategic Planning, Technology management, Marketing

Planning of English speaking evaluation platform service

RESEARCHER IN SYSTRAN

- Establishing a service strategy using English speaking evaluation technology
- English speaking evaluation technology Q.A & Q.C
- Skills: Quality Assurance, Quality control, Mockup, Strategic Planning, Technology management, Marketing

Seoul, South Korea

June. 2018. - December.2018

Customer-specific recognizer service platform service planning

RESEARCHER IN SYSTRAN

- Customer-specific recognizer service app service design
- Establish platform service strategies such as B2B and B2C
- Skills: Mockup, Marketing, Strategic Planning, Technology management

Seoul, South Korea

June. 2018. - April.2018

Speech Recognition Research

RESEARCHER IN SYSTRAN

- Research for Automatic Speech Recognition Technology
- Acoustic Model, Language Model
- Speaker Verification
- Kaldi
- English speaking evaluation
- Real-time speech recognition system
- Skills: Kaldi

Seoul, South Korea

Aug. 2018. - July. 2019

Neural Machine Translation Research

RESEARCHER IN SYSTRAN

- Research for Neural Machine Translation
- Skills: OpenNMT Tensorflow, OpenNMT Pytorch

Seoul, South Korea

Aug. 2018. - July. 2019

Development of multilingual word translation service based on Korean semantics analysis for multicultural family

RESEARCHER IN BUFS

- Han-Tong-E: Multilingual Word Translation Service based on Word Semantic Analysis
- Collaboration with Professor Ok Cheol-young's laboratory at Ulsan University
- Project homepage development
- In-depth understanding of Utagger and seminar on WSD
- Build project-related data and conduct surveys
- Project Page: <https://bufsconv.blogspot.kr/>
- Funded by National Research Foundation of Korea(NRF)

Seoul, South Korea

Sep. 2016. - June. 2018

Establishment of Gold Standard language resource bank for Korean machine learning and evaluation(Next-generation information computing technology development business)

RESEARCHER IN BUFS

- For the purpose of developing Korean Natural Language Toolkit
- Development of automatic Korean word spacing and Korean compound noun decomposer
- Skills: Python

Seoul, South Korea

Sep. 2017. - June. 2018

Local industry development project in connection with public institutions

RESEARCHER IN BUFS

- ICT VAEMS: Vertical Aquaculture Energy Management
- Funded by the Korea Institute for Advancement of Technology(KIAT)

Seoul, South Korea

Sep. 2016. - Aug. 2017

Basic Research on Semantic Frame Generation for Short Questions (ExoBrain)

RESEARCHER IN BUFS

- Exobrain Project

Seoul, South Korea

Oct. 2015. - Feb. 2016

Professional Experience

SYSTRAN, Automatic Speech Recognition(ASR) Team

Seoul, South Korea

RESEARCH ENGINEER

Sep. 2018 - July. 2019

- Core technology development of the real-time simultaneous speech translation based on knowledge enhancement (ETRI)
- Participated in NTT Docomo speech Recognition Solution Project
- Participated in CJ Logistics speech recognition agent development project
- Participated in Samsung SDS speech recognition project
- Speech recognition service planning and development (Real-time lecture interpretation, Automatic subtitle generation service, English speaking evaluation, Customer-specific recognizer service, Web RTC-based Real-time video interpretation)
- Deep learning based speech recognition research (Acoustic Model, Language Model Training)
- Machine Translation research
- IaaS and PaaS service planning, API Platform service operation planning, B2B B2C planning and development
- Micro Service Architecture Design and Successful API Platform Service Launching Experience (LCloud)
- Governmental project proposal experience
- Reason for leaving: Graduate school

SYSTRAN, Automatic Speech Recognition(ASR) Team

Seoul, South Korea

RESEARCH INTERN

June. 2018 - August. 2018

- Participated a project "Core technology development of the real-time simultaneous speech translation based on knowledge enhancement (ETRI)"
- Developed Simultaneous Speech Translation (implemented using JAVA, Python and Kaldi)

Skills

Programming	JAVA, Python, C / C++
Machine Learning & Deep Learning	Pytorch, Keras, Tensorflow, OpenNMT, Kaldi
Natural Language Processing	Machine Translation, Speech Recognition, Neuro-Symbolic Computing
Full Stack	Html, CSS, Javascript, Java FX, Oracle, MySQL, Spring Boot, VertX, AWS
Strategy & Planning	API Platform Service, Platform Service Planning, Gravitee IO, PaaS, IaaS, Q.A, Q.C, Mockup, Swagger
Etc	MS Office, Redmine, Linux, Docker, Web RTC, Latex
Language Skills	Korean, English, Japanese, Thai, Myanmar

Patents

2019.12	Device and method for correcting Korean spelling , Apply for a patent	South Korea
2020.09	APPARATUS AND METHOD FOR OUTPUTTING IMAGE CORRESPONDING TO LANGUAGE , Apply for a patent	South Korea
2020.09	METHOD OF TRANSLATING ANCIENT KOREAN USING MACHINE TRANSLATION , Apply for a patent	South Korea
2020.09	Correction performance evaluation metrics of neural network machine translation and method of constructing the same , Apply for a patent	South Korea
2020.09	METHOD OF BUILDING TRAINING DATA OF MACHINE TRANSLATION , Apply for a patent	South Korea
2021.03	Improving speech recognition performance using TTS in domain-specific environment , Apply for a patent	South Korea
2021.05	Method For Generating Training Data And Method For Post-Processing Of Speech Recognition Using The Same , Apply for a patent	South Korea

Demo

2019.12	Deep Learning based English Grammar Error Correction , http://nlplab.iptime.org:32292/	South Korea
2019.12	Deep Learning based Korean Spelling Correction , http://nlplab.iptime.org:32288/	South Korea
2020.01	Korea University Neural Machine Translation Platform , http://nlplab.iptime.org:32296/	South Korea
2020.02	Neural Symbolic Inference , http://nlplab.iptime.org:32245/	South Korea
2020.02	English Free Talking and Grammar Error Correction Feed Back System using NAO Robot , https://www.youtube.com/watch?v=R16dG_4cnFk	South Korea
2020.03	Korean Ancient Neural Machine Translation Platform , http://nlplab.iptime.org:32242/	South Korea
2020.03	PicTalky: Text to Pictogram Demo for Children with Developmental Disabilities , https://pictalky.df.r.appspot.com/	South Korea
2020.05	COVID-19 Machine Translation Platform , http://nlplab.iptime.org:32250/	South Korea
2020.07	Ancient Korean Translation & Machine Reading Comprehension with NAO , https://www.youtube.com/watch?v=nXBZe-t9a-4&t=8s	South Korea
2020.11	Bilingual Education Using NAO , https://www.youtube.com/watch?v=4kCMMa62LDY&t=9s	South Korea
2021.04	Ko-CSS (Korean Colloquial-Sentence-Splitter) for simultaneous speech translation , http://nlplab.iptime.org:32285/	South Korea
2021.06	BTS (BackTranScription) for STT Post-Processor , http://nlplab.iptime.org:32260/	South Korea

Invited Talk & Teaching Experience

Korea University

HEAD TEACHING ASSISTANT

- (BDC101) Introduction to Natural Language Processing in Big Data
- (DFE610) Natural Language Processing for Digital Finance Engineering

South Korea

Autumn 2020

Korea University

TEACHING ASSISTANT

- (COSE461) Natural Language Processing
- (DFC615) Artificial Intelligence and Natural Language Processing
- Lecture materials: <https://github.com/Parkchanjun/KU-NLP-2020-1>

South Korea

Spring 2020

LLsoLLu

PRESENTER

- Latest natural language processing trends
- Language Representation, Neural Symbolic Computing
- Invited by LLsoLLu Global Strategy and Planning Team

South Korea

March. 2020

NC SOFT

PRESENTER

- Technology Transfer Seminar
- Korean Spelling Corrector Using Automatic Noise Generation Methodology
- Invited by NC SOFT Language AI Team

South Korea

February. 2020

Dongguk University

PRESENTER

- A.I - NLP - MT for Liberal Arts
- Introduction to Natural Language Processing, RBMT, SMT, NMT, OpenNMT
- Invited by Dongguk University, Department of English Language and Literature professor. Myung-Kwan Park
- Lecture materials: https://github.com/Parkchanjun/Dongguk_AI_NLP_MachineTranslation

South Korea

January. 2020

SKC

INSTRUCTOR

- Deep Learning and Natural Language Processing Course
- Introduction of Text Preprocessing and NLP for Linguistics
- Introduction of Word Embedding
- Introduction of Machine Translation
- Lecture materials(Text Preprocessing): https://github.com/Parkchanjun/SKC_Text_Preprocessing
- Lecture materials(Introduction of Machine Translation): https://github.com/Parkchanjun/SKC_MachineTranslation

South Korea

Oct. 2019 - Nov. 2019

Naver Tech Talk

PRESENTER

- Machine Translation for Everyone
- Invited by Sung Kim who are the Head of Naver Clova AI
- Lecture videos: <https://tv.naver.com/v/9906991>

South Korea

Aug. 2019

SK T Academy

INSTRUCTOR

- Machine Translation for Everyone
- Invited by SK Planet
- Introduction to Natural Language Processing, RBMT, SMT, NMT, OpenNMT
- Lecture videos: <https://tacademy.skplanet.com/live/player/onlineLectureDetail.action?seq=165>

South Korea

Aug. 2019

Sooyoungro Church

INSTRUCTOR

- Sooyoungro Church Dream Bowl School Instructor (Free Service)
- Basic English Grammar and English Conversation Education

South Korea

Aug. 2017 - Dec. 2017

Korea Press Foundation

INSTRUCTOR

- Reading mentoring

South Korea

Aug. 2015 - Dec. 2015

TOP CONFERENCE

Should we find another model?: Improving Neural Machine Translation Performance with ONE-Piece Tokenization Method without Model Modification

Chanjun Park, Sugyeong Eo, Hyeonseok Moon, Heui-Seok Lim

Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2021): Human Language Technologies: Industry Papers, 2021

BTS: Back TranScription for Speech-to-Text Post-Processor using Text-to-Speech-to-Text

Chanjun Park, Jaehyung Seo, Seolhwa Lee, Chanhee Lee, Hyeonseok Moon, Sugyeong Eo, Heui-Seok Lim

ACL 2021 - WAT(Workshop on Asian Translation) 2021, 2021

INTERNATIONAL JOURNAL

Exploring the Data Efficiency of Cross-Lingual Post-Training in Pretrained Language Models

Chanhee Lee, Kisu Yang, Taesun Whang, **Chanjun Park**, Andrew Matteson, Heuseok Lim

Applied Sciences 11.5 (2021) p. 1974. Multidisciplinary Digital Publishing Institute, 2021

Neural Spelling Correction: Translating Incorrect sentences to Correct sentences for Multimedia

Chanjun Park, Kuekyeng Kim, YeongWook Yang, Minho Kang, HeuiSeok Lim

Multimedia Tools and Applications (2020) pp. 1–18. Springer, 2020

Ancient Korean Neural Machine Translation

Chanjun Park, Chanhee Lee, YeongWook Yang, Heuseok Lim

IEEE Access 8 (2020) pp. 116617–116625. 2020

Comparison of the evaluation metrics for Neural Grammatical Error Correction with Overcorrection

Chanjun Park, YeongWook Yang, Chanhee Lee, Heuseok Lim

IEEE Access 8 (2020) pp. 106264–106272. 2020

Decoding Strategies for Improving Low-Resource Machine Translation

Chanjun Park, Yeongwook Yang, Kinam Park, Heuseok Lim

Electronics 9.10 (2020) p. 1562. Multidisciplinary Digital Publishing Institute, 2020

INTERNATIONAL CONFERENCE

PicTalky: Deep learning-based speech therapy platform for developmental disabilities

Chanjun Park, Yanghee Kim, Yoonna Jang, Umadevi G.R, HeuiSeok Lim

The International Conference on Big data, IoT, and Cloud Computing, 2020

Domain-specialize Neural Machine Translation Methodology

Chanjun Park, HeuiSeok Lim

International conference on Interdisciplinary research on Computer Science, Psychology, and Education, 2020

Toward High Quality Parallel Corpus Using Monolingual Corpus

Chanjun Park, Youngdae Oh, Jongkeun Choi, Dongpil Kim, HeuiSeok Lim

The 10th International Conference on Convergence Technology, 2020

Data Augmentation Method for Korean Neural Machine Translation

Chanjun Park, Kinam Park, HeuiSeok Lim

INTERNATIONAL CONFERENCE ON INFORMATION, SYSTEM AND CONVERGENCE APPLICATIONS, 2020

Self- Supervised Korean Spelling Correction via Denoising Transformer

Chanjun Park, Sungjin Park, HeuiSeok Lim

INTERNATIONAL CONFERENCE ON INFORMATION, SYSTEM AND CONVERGENCE APPLICATIONS, 2020

Can artificial intelligence translate the Annals of the Joseon Dynasty ?

Chanjun Park, Sungjin Park, Kinam Park, Jaechoon Jo, HeuiSeok Lim

The 10th International Conference on Convergence Technology, 2020

Design Neural Machine Translation Model Combining External Symbolic Knowledge in translation task

Sugyeong Eo, **Chanjun Park**, HeuiSeok Lim

International conference on Interdisciplinary research on Computer Science, Psychology, and Education, 2020

Free Talking NAO Robot for Children

Chanjun Park, Jaekue Choi, Sungjin Park, Yanghee Kim, HeuiSeok Lim

International conference on Interdisciplinary research on Computer Science, Psychology, and Education, 2019

Improved Machine Translation Performance Using Two Stage Subword Tokenization

Chanjun Park, Sora Choi, HeuiSeok Lim

International conference on Interdisciplinary research on Computer Science, Psychology, and Education, 2019

Quality before Quantity: Improved Machine Translation performance using parallel corpus filtering

Chanjun Park, Seunggyu Han, HeuiSeok Lim

International conference on Interdisciplinary research on Computer Science, Psychology, and Education, 2019

Improved Machine Translation Performance Using Relative Ratio of Original and Synthetic Corpus

Chanjun Park, Jang Mook Kang

International conference on Interdisciplinary research on Computer Science, Psychology, and Education, 2019

DOMESTIC JOURNAL

A study on performance improvement considering the balance between corpus in Neural Machine Translation

Chanjun Park, Kinam Park, Sugyeong Eo, Hyeonseok Moon, HeuiSeok Lim

Journal of the Korea Convergence Society 00.00 (2021) pp. 0–0. Korea Convergence Society, 2021

Research on Subword Tokenization of Korean Neural Machine Translation and Proposal for Tokenization Method to Separate Jong-sung from Syllables

Sugyeong Eo, **Chanjun Park**, Hyeonseok Moon, HeuiSeok Lim

Journal of the Korea Convergence Society 12.3 (2021) pp. 1–7. Korea Convergence Society, 2021

Latest Researches on Quality Estimation

Sugyeong Eo, **Chanjun Park**, Hyeonseok Moon, Jaehyung Seo, HeuiSeok Lim

Journal of the Korea Convergence Society 00.00 (2021) pp. 0–0. Korea Convergence Society, 2021

Filter-mBART Based Neural Machine Translation Using Parallel Corpus Filtering

Hyeonseok Moon, **Chanjun Park**, Sugyeong Eo, JeongBae Park, HeuiSeok Lim

Journal of the Korea Convergence Society 00.0 (2021) pp. 0–0. Korea Convergence Society, 2021

Recent Automatic Post Editing and Multilingual Pretrained Language Model Survey

Hyeonseok Moon, **Chanjun Park**, Sugyeong Eo, Jaehyung Seo, HeuiSeok Lim

Journal of the Korea Convergence Society 00.00 (2021) pp. 0–0. Korea Convergence Society, 2021

Cross-Validated Ensemble Methods in Natural Language Inference

Kisu Yang, Taesun Whang, Dongsuk Oh, **Chanjun Park**, HeuiSeok Lim

Journal of KIISE 48.2 (2021) pp. 154–159. Korean Institute of Information Scientists and Engineers (KIISE), 2021

Neural Machine translation specialized for Coronavirus Disease-19 (COVID-19)

Chanjun Park, Kyeong-Hee Kim, Ki-Nam Park, Heui-Seok Lim

Journal of the Korea Convergence Society 11.9 (2020) pp. 7–13. Korea Convergence Society, 2020

An AI Service to support communication and language learning for people with developmental disability

Chanjun Park, Yanghee Kim, Yoonna Jang, Umadevi G.R, HeuiSeok Lim

Journal of the Korea Convergence Society 11.6 (2020) pp. 51–57. 2020

Automatic Post Editing Research

Chanjun Park, HeuiSeok Lim

Journal of the Korea Convergence Society 11.5 (2020) pp. 1–8. 2020

A Study on the Performance Improvement of Machine Translation Using Public Korean-English Parallel Corpus

Chanjun Park, HeuiSeok Lim

Journal of Digital Convergence 18.6 (2020) pp. 271–277. 2020

DOMESTIC CONFERENCE

Quality, not Quantity? : Effect of parallel corpus quantity and quality on Neural Machine Translation

Chanjun Park, Yeonsu Lee, Chanhee Lee, HeuiSeok Lim

The 32st Annual Conference on Human Cognitive Language Technology, 2020

Alleviation of Overcorrection Problem in Neural Korean Spelling Correction

Chanjun Park, Yeonsu Lee, Kisu Yang, HeuiSeok Lim

The 32st Annual Conference on Human Cognitive Language Technology, 2020

Design Neural Machine Translation Model Combining External Symbolic Knowledge

Sugyeong Eo, **Chanjun Park**, HeuiSeok Lim

The 32st Annual Conference on Human Cognitive Language Technology, 2020

Cross-Lingual Transfer of Pretrained Transformers to Resource-Scarce Languages (Best Paper Award)

Chanhee Lee, **Chanjun Park**, Gyeongmin Kim, Dongsuk Oh, HeuiSeok Lim

The 32st Annual Conference on Human Cognitive Language Technology, 2020

MMA: Multi-modal Message Aggregation for Korean VQA

Sungjin Park, **Chanjun Park**, Jaehyung Seo, HeuiSeok Lim

The 32st Annual Conference on Human Cognitive Language Technology, 2020

Korean Spell Correction based on Denoising Transformer

Chanjun Park, Sol Jeong, Kisu Yang, Sumi Lee, Jaechoon Joe, HeuiSeok Lim

The 31st Annual Conference on Human Cognitive Language Technology, 2019

Parallel Corpus Filtering and Korean-Optimized Subword Tokenization for Machine Translation

Chanjun Park, Gyeongmin Kim, HeuiSeok Lim

The 31st Annual Conference on Human Cognitive Language Technology, 2019

Optimization of Data Augmentation Techniques in Neural Machine Translation

Chanjun Park, Kuekyeng Kim, HeuiSeok Lim

The 31st Annual Conference on Human Cognitive Language Technology, 2019

Comparative Study of Sentiment Analysis Model based on Korean Linguistic Characteristics

Gyeongmin Kim, **Chanjun Park**, HeuiSeok Lim

The 31st Annual Conference on Human Cognitive Language Technology, 2019

Cross-Validated Ensemble Methods in Natural Language Inference (Best Paper Award)

Kisu Yang, Taesun Whang, Dongsuk Oh, **Chanjun Park**, HeuiSeok Lim

Korean Spacing Module using Eojeol dictionary and Morpheme dictionary

Chanjun Park, Jin-hyuk Choi, Pummo Ryu

Korea Software Congress 2018, 2018

Knowledge-enhancing Real-time Lecture interpreter service Development

Chanjun Park, Minho Kang, Byunggun Park, Suntae Kim, Myeongjin Hwang, Dongpil Kim, Young-kil Kim

Korea Software Congress 2018, 2018

Voice recognition performance improvement using TTS system in address recognition system

Chanjun Park, Suntae Kim, Minho Kang, Myeongjin Hwang, Byunggun Park, Yeonsu Kim

Korea Software Congress 2018, 2018

Improved speech recognition performance in business meeting environments using domain-specific method

Chanjun Park, Suntae Kim, Minho Kang, Byunggun Park, Myeongjin Hwang, Yeonsu Kim

Korea Software Congress 2018, 2018

Spam/Ham Classification using Ensemble Technique of Machine Learning Algorithms

Chanjun Park, Pummo Ryu

Korea Computer Congress 2018, 2018

Two-Stage Korea Compound Noun Decomposer

Chanjun Park, Pummo Ryu

The 30st Annual Conference on Human Cognitive Language Technology, 2018

Dictionary based Compound Noun Decomposer using Resource extension

Chanjun Park, Hyejin Won, Pummo Ryu

Korea Computer Congress 2018, 2018

BOOK CHAPTERS

Natural Language Processing Bible

Heuiseok Lim, **Korea University NLPAL Lab**

Human Science, <https://www.aladin.co.kr/shop/wproduct.aspx?partner=rss&ISBN=K412637214>, 2020

Operate Naver's natural language processing blog

BLOGGER

- URL: <https://blog.naver.com/bcj1210>
- Linguistics for natural language processing
- Natural Language Processing (NLP)
- Machine Translation for Everyone
- Business, strategy, and planning for everyone
- Natural language processing with NLTK
- Software engineering
- Today's CCM

South Korea

2016.Dec-Present

OpenNMT Forum Members

MEMBERS

- Active as OpenNMT Forum Members
- Develop and share OpenNMT Pytorch, Tensorflow, Library Tutorial

South Korea

2019.July-2019.July

ETRI's machine translation quality evaluator

ESTIMATOR

- Quality evaluation of machine translation hosted by ETRI

South Korea

2019.October-2019.December

Completion of basic curriculum for natural language processing and language intelligence

STUDENT

- Korea University Human-Inspired AI & Computing Research Center

South Korea

2019.July-2019.July

Completed the natural language processing tutorial (2016-2019)

STUDENT

- Special Interest Group of Human and Cognitive Language Technology

South Korea

2016 - 2019

Reporter, Korea Electronics and Telecommunications Research Institute (ETRI)

REPORTER

- Served as 3rd and 4th reporters at the Korea Electronics and Telecommunications Research Institute

South Korea

2016.Feb-2018.Feb

Flitto Supporters

SUPPORTERS

- The first plito supporters

South Korea

2017.Dec-2017.Dec

Journalist of the Ministry of Science and ICT

REPORTER

- 8 articles published
- Visit Exobrain Quiz Show
- URL: <https://blog.naver.com/bcj1210/221147196551>

South Korea

2016.Feb-2017.Feb

Planning a video lecture for Natural Language Processing Basic at BUFS

PLANNER

- URL: <https://bit.ly/2NmFJV1>

South Korea

2016.July-2016.July

BIT Computer Advanced Course Completion

STUDENT

- Bet Advanced Course Completion and Bet President Award
- Bit project level 2 license
- Development of intelligent English learning APP

South Korea

2016.January-2016.August

BIT Computer Short Course Completion

STUDENT

- Short-term course completion (C, JAVA, Data Structure)

South Korea

2015.June-2015.August

Courses related to natural language processing and computer engineering

South Korea

2015-Present

STUDENT

- Computer Programming Basics 1 (C Language) (A+)
- Linguistics basics (A+)
- Basic language processing (A+)
- Computer and Information Society (A+)
- Introduction to computers (A+)
- Computer architecture (A+)
- Data structure (A)
- Concepts of programming language (A+)
- Operating system (A+)
- Database (A+)
- Natural language processing (BUFS) (A+)
- Advanced computer programming (System Programming, Machine Learning) (A+)
- Software Engineering (A+)
- Natural Language processing project (A+)
- Information extraction and text mining (A+)
- Machine learning application (A)
- Cognitive basis of linguistics (A+)
- Natural language processing (KU) (A+)
- Knowledge Representation Techniques (KU) (A+)
- Special Topics in Natural Language Processing (KU) (A+)
- Entrepreneurship and Technology Startups (KU) (A+)
- Smart Factory Security Seminar II (KU) (A+)