

# Juhwan Choi

✉ gold5230@cau.ac.kr | 🌐 c-juhwan | 🏠 c-juhwan.github.io

## Research Interest

Deep Learning, Natural Language Processing, Large Language Models, Data-Centric Approach

## Education

### Chung-Ang University

M.S. in Artificial Intelligence

GPA: 4.33 / 4.5

Advisor: [Youngbin Kim](#)

Seoul, Korea

Mar. 2023 - Feb. 2025 (Expected)

### Chung-Ang University

B.S. in Electrical and Electronics Engineering

GPA: 3.94 / 4.5 [[Certificate](#)] [[Transcript](#)]

Seoul, Korea

Mar. 2017 - Feb. 2023

## Research Experience

### Research Assistant

Advisor: [Youngbin Kim](#)

Intelligent Information Processing Lab @ Chung-Ang Univ.

Mar. 2023 - Present

### Undergraduate Research Intern

Advisor: [Youngbin Kim](#)

Intelligent Information Processing Lab @ Chung-Ang Univ.

Apr. 2021 - Feb. 2023

### Undergraduate Intern

Advisor: [Hohyun Park](#)

Big Data & AI Lab @ Chung-Ang Univ.

Jan. 2021 - Feb. 2021

## International Publications

### Korean Offensive Language Detection

Under review.

### Curriculum Data Augmentation for Image Classification

Under review.

### Semantic-preserving Rule-based Text Data Augmentation

Under review.

### Constructing Sport Game Conversation Dataset

Under review.

### Text Data Augmentation Based on Decision Boundary

Under review.

### AutoAugment Is What You Need: Enhancing Rule-based Augmentation Methods in Low-resource Regimes

[Juhwan Choi](#), Kyohoon Jin, Junho Lee, Sangmin Song and Youngbin Kim

EACL 2024 Student Research Workshop

### GPTs Are Multilingual Annotators for Sequence Generation Tasks

[Juhwan Choi](#), Eunju Lee, Kyohoon Jin and Youngbin Kim

EACL 2024 Findings

### SoftEDA: Rethinking Rule-Based Data Augmentation with Soft Labels [[Paper](#)] [[Review](#)] [[Code](#)]

[Juhwan Choi](#), Kyohoon Jin, Junho Lee, Sangmin Song and Youngbin Kim

ICLR 2023 Tiny Papers

Accepted as notable work with oral presentation.

## Generative Data Augmentation via Wasserstein Autoencoder for Text Classification [Paper] [Code]

Kyohoon Jin, Junho Lee, [Juhwan Choi](#), Soojin Jang and Youngbin Kim  
The 13th International Conference on ICT Convergence

## Domestic Publications

---

### 소프트 라벨을 적용한 규칙 기반 텍스트 데이터 증강 기법 [Paper] [Poster]

[Juhwan Choi](#), Junho Lee, Sangmin Song, Kyohoon Jin and Youngbin Kim  
2023 Summer Annual Conference of IEIE

### 감정 간의 관계를 고려한 지도 대조 학습 기반 감정 인식 [Paper]

Dongje Yoo, Kyunghoon Jeon, [Juhwan Choi](#) and Youngbin Kim  
2023 Summer Annual Conference of IEIE

### KITE: 한국어 고의 오타자를 활용한 텍스트 데이터 증강 방법론 [Paper]

Seungkuk Yu, [Juhwan Choi](#), Heejae Suh, Kyohoon Jin and Youngbin Kim  
Proceedings of HCI Korea 2023

### 낯선 데이터를 활용한 과잉신뢰 완화 텍스트 증강 기법 [Paper] [Presentation]

Junho Lee, Sangmin Song, [Juhwan Choi](#), Juhyoung Park, Kyohoon Jin and Youngbin Kim  
Proceedings of HCI Korea 2023

### Semantic Preservation and Natural Language Data Augmentation via Variational Autoencoder [Paper]

Yuchul Shin, Kyohoon Jin, [Juhwan Choi](#), Junho Lee, Soojin Jang and Youngbin Kim  
TechArt

### Variational Autoencoder 기반 의미 보존 자연어 데이터 증강 기법 [Paper] [Presentation]

[Juhwan Choi](#), Junho Lee, Kyohoon Jin, Yehoon Jang, Soojin Jang and Youngbin Kim  
2022 Summer Annual Conference of IEIE

## Research Projects

---

Multimodal Dataset Generation	Nov. 2023 - Present
Domain Generalization through Dataset Generation	Nov. 2023 - Present
Multilingual Data Annotation using LLMs Submitted a paper and accepted to EACL 2024 Findings.	Apr. 2023 - Jan. 2024
Text Data Augmentation with Soft Label Submitted a paper and accepted to ICLR 2023 Tiny Papers.	Feb. 2023 - May. 2023
Interactive News Summarization	Jan. 2023 - Present
Decision Boundary-Based Data Augmentation Submitted a paper - under review	Aug. 2022 - Jun. 2023
Deep Learning Seminar Based on Stanford CS224n and CS231n	Feb. 2022 - Mar. 2022
Text Data Augmentation using Wasserstein Autoencoder Related work research, Implemented text classification model.	Aug. 2021 - Jan. 2022
Implementing Image Captioning Model using Transformers	Apr. 2021 - Jun. 2021
Korean Hate Speech Detection Model Used pretrained <a href="#">KcELECTRA</a> . Ranked 11/88 (in Dec. 2022) at <a href="#">[Kaggle competition]</a>	Apr. 2021 - May. 2021
Deep Learning Seminar Based on CMU 11-785	Jan. 2021 - Feb. 2021

## Skills

---

**Programming Languages** Familiar: Python | C | C++ Experienced: Matlab  
**Libraries & Tools** Pytorch | HuggingFace | Numpy | Pandas | Git | CLI  
**Language Proficiency** Native: Korean Fluent: English [TOEIC 955] [TEPS 409]

## Services

---

**Reviewer** ACL Rolling Review (From 2023 August)  
EACL 2024 (Through ACL Rolling Review 2023 October)  
NAACL 2024 (Through ACL Rolling Review 2023 December)  
ICLR 2024 Tiny Papers

## Relevant Coursework

---

### Graduate Level

**Artificial Intelligence** Deep Unsupervised Learning | Deep Generative Model | Graph Neural Networks | Advanced Natural Language Processing | Artificial Intelligence Literacy | Ethics and Fairness in Artificial Intelligence

### Undergraduate Level

**Computer Science** Algorithms | Artificial Intelligence | Basic Computer Programming | Computer Architecture | Data Structure | Digital Image Processing | Internet of Intelligent Things | Logic Circuit | Microcomputer Systems | Mobile Network | Numerical Analysis | Object-Oriented Programming | Operating Systems

**Mathematics** Calculus 1 | Calculus 2 | Electric Mathematics | Engineering Mathematics | Linear Algebra | Random Variables