

# Haoran Xu

☎ 443-453-8801 | ✉ hxu64@jhu.edu | 🏠 haoranxu.com | 🌐 github.com/haoranxu

## EDUCATION

### Johns Hopkins University (JHU)

Master of Science in Engineering - Computer Science; GPA: 3.91/4.0  
Concentration on Human Language Technology (HLT)

Baltimore, USA

Aug 2019 – Present

### University of Illinois at Chicago (UIC, 3+2 exchange program)

Master of Science - Electronic and Computer Engineering; GPA: 3.93/4.0

Chicago, USA

Aug 2017 – May 2019

### East China University of Science and Technology (ECUST)

Bachelor of Engineering - Information Engineering; Major Ranking: 1/90

Shanghai, China

Sep 2014 – July 2018

**Main coursework:** Nature Language Processing, Machine Learning, Human Language Technology, Algorithms, Pattern Recognition, Neural Networks, Computer Vision

**Current Research areas:** Multilingual Word Embedding, Cross-Lingual Transfer Learning, Machine Translation

## PUBLICATIONS

- [1] Gradual Fine-Tuning for Low-Resource Domain Adaptation  
**Haoran Xu**, Seth Ebner, Mahsa Yarmohammadi, Aaron Steven White, Benjamin Van Durme and Kenton Murray  
*in Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2021, **Under Review**
- [2] Zero-Shot Cross-Lingual Dependency Parsing through Contextual Embedding Transformation  
**Haoran Xu** and Philipp Koehn  
*in Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2021, **Under Review**
- [3] Cross-Lingual Contextual Embedding Space Mapping  
**Haoran Xu** and Philipp Koehn  
*in Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, 2021, **Under Review**
- [4] Efficient Quadratic Programming for Peak-to-Average Power Ratio Reduction in Communication Systems  
**Haoran Xu**, Shahin Khobahi and Mojtaba Soltanalian  
[Preprint Version]

## RESEARCH EXPERIENCE

### Better Extraction from Text Towards Enhanced Retrieval Program [1]

Center for Language and Speech Processing, mentors: Benjamin Van Durme, Mark Dredze, JHU

Research Intern

May 2020 – Present

- Cooperated with the research group in cross-lingual transfer learning for the Information Extraction (IE) task.
- Proposed a novel data augmentation approach for domain adaptation, which improved the performance of the original IE system by 37%.
- Investigated the general applicability of the augmentation approach and achieved state-of-the-art performance in dialogue state tracking and event extraction tasks.

### Zero-Shot Cross-Lingual Dependency Parsing [2]

Center for Language and Speech Processing, advised by Philipp Koehn, JHU

Research Assistant

May 2020 – Sep 2020

- Developed a zero-shot approach for dependency parsing by building a multilingual shared semantic space.
- Investigated the properties of embedding spaces across various languages.
- Surpassed state-of-the-art methods by 2.82% in 6 languages from 4 language families on average.

## Cross-Lingual Contextual Embedding Space Mapping [3]

Center for Language and Speech Processing, advised by Philipp Koehn, JHU

Research Assistant

Sep 2019 – May 2020

- Developed a brand-new method of sense-level and dictionary-free contextual embedding mapping.
- Outperformed static embedding alignment approach by approximate 10% accuracy on bilingual dictionary induction task.
- Revealed the tight relationship of isotropy, isometry and isomorphism in contextual embedding spaces, and explained the strong correlation between the quality of mapping and them.

## Efficient Quadratic Programming in Wireless Communication [4]

Waveform Optimization Lab, advised by Mojtaba Soltanalian, UIC

Research Assistant

Nov 2018 – Mar 2019

- Developed a new algorithm to reduce the peak energy of data transmission based on *Unimodular Quadratic Programming (UQP)*, and achieved higher stability and faster convergence rate than existing methods.
- Provided more mature and general techniques for solving the *UQP* problem in radar code design scenarios as well as other active sensing and communication applications.

## PROJECTS

---

### Neural Machine Translation between Arabic and English

Center for Language and Speech Processing, JHU

Research Assistant

Oct 2020 - Present

- Built neural machine translation models based on *fairseq* framework.
- Developed an ensemble model composed of transformer translation model and *GigaBERT*, an English-Arabic bilingual pre-trained model which is used for embedding layer.
- Achieved state-of-the-art *BLEU* scores from Arabic to English.

### Image Expansion with GANs

Science and Engineering Laboratory, UIC

Team Leader

Jan 2019 - May 2019

- Built a deep learning method based on *GAN* to naturally expand the boundaries of incomplete images.
- Designed an encoding-decoding hybrid CNN composed of dilated convolution and the normal one, rendering the extended image to acquire better realism and conform to the semantics of the whole image.
- Replaced traditional global discriminator with local discriminator to reduce the blur of the vertical part of the image and enhance the authenticity.

### Invisible Signature Security System

Science and Engineering Laboratory, UIC

Team Leader

Jan 2018 - May 2018

- Created GUI interface and local database for human-computer interaction, which allowed users to sign their names in the air to implement signature recognition.
- Proposed high-dimension dynamic signature features extracted from the signature process to promise the uniqueness of the signature and the security of the system.
- Utilized *Fast-DTW (Fast Dynamic Time Warping)* and statistical information of high-dimension dynamic features to recognize the signature with high precision.
- Received the Winner Award in 2018 Expo at UIC and the Best Research Paper Award at ECUST.

## HONORS AND AWARDS TOP 5

---

- Best Research Paper Award ECUST - 2018
- Expo 2018 Best in Category Award UIC - 2018
- Special Scholarship (top 2%-ranked student) ECUST - 2015, 2016, 2018
- Social Work Award (received for community contribution) ECUST - 2015, 2016, 2018
- Excellent Student Award (4 selected among 1000 students) ECUST - 2015

## SKILLS

---

**Programming Languages:** Python, MATLAB, C, bash, C++, Java,  $\text{\LaTeX}$ , SQL

**Toolkits and Libraries:** PyTorch, Allennlp, fairseq, Moses, NLTK, Sklearn, Keras, TensorFlow, PyQt, MySQLdb