Wuwei Lan

Department of CSE
The Ohio State University
Columbus, OH-43210

✉ lan.105@osu.edu
☎ +1-352-281-8103
🌐 https://lanwuwei.github.io/
G+ https://scholar.google.com/citations?user=rpOgHRMAAAAJ&hl=en

RESEARCH	Natural Language Processing, Machine Learning, Deep Learning.

EDUCATION

The Ohio State University (OSU)   05/2015-05/2021
◇ Ph.D. candidate in Computer Science and Engineering
◇ Research Topic: Sentence Level Semantics, Paraphrase
◇ Advisor: Wei Xu, GPA: 3.82/4.0

University of Science and Technology of China (USTC)   08/2010-06/2014
◇ B.E. in Computer Science and Engineering
◇ Graduated with the highest Guo Moruo scholarship

EXPERIENCE

Graduate Research Assistant, OSU *Advisor:* Wei Xu

◇ Neural Semi-CRF aligner for monolingual word alignment (2019-2020) [9]
- Proposed a neural semi-CRF model for token-level and phrase-level alignments
- Achieved SOTA performance for in-domain and out-of-domain evaluations
- Annotated the largest monolingual word alignment corpus to date

◇ Bilingual BERT pre-training for English and Arabic (2020) [8]
- Pre-trained a customized bilingual BERT for Arabic NLP and English-to-Arabic zero-shot transfer learning
- Outperforms mBERT, XLM-RoBERTa, and AraBERT in both the supervised and zero-shot transfer settings

◇ Deep neural networks for sentence pair modeling (2018) [5]
- Implemented and analyzed several neural network designs (DecAtt, ESIM, PWIM) for sentence pair modeling tasks
- Systematically compared the strengths and weaknesses of different network designs and provided insights for the model architecture

◇ Subword-based embeddings for paraphrase identification (2018) [4]
- Proposed a multi-task framework for RNN/CNN based subword models and deep sentence pair models
- Studied how effective subword-level (character and character n-gram) representations are in paraphrase identification

◇ Large-scale paraphrase collection from Twitter (2016-2017) [3]
- Presented the largest human-labeled paraphrase corpus to date of 51,524 sentence pairs from Twitter
- More than 30,000 new sentential paraphrases can be easily and continuously captured every month at 70% precision

Undergrad Research Assistant, USTC *Advisor:* Shangfei Wang

◇ Facial expression recognition with deep boltzmann machine (2013) [1]
- Proposed to use the Deep Boltzmann Machine to learn thermal features for expression recognition from thermal long wavelength infrared images
- Comparison experimental results on the NVIE database demonstrate that our approach outperforms other approaches

INTERNSHIPS

Research Intern, Tencent AI Lab *Mentor:* Jia Cui, Dong Yu

◇ State-of-the-art language models for automatic speech recognition (2019)
- Implemented and analyzed several state-of-the-art language models for ASR system, including ON-LSTM, Transformer, Transformer-XL and XLNet

Research Intern, Microsoft Research *Mentor:* Haitao Wu

◇ PAC: Taming TCP incast congestion using proactive ACK control (2014) [2]
- Presented PAC, a simple yet very effective design to tame TCP in cast congestion via Proactive ACK Control at the receiver.

PUBLICATIONS

[9] *Neural semi-Markov CRF for Monolingual Word Alignment*
Wuwei Lan and Wei Xu (under submission)

[8] *An Empirical Study of Pre-trained Transformers for Arabic Information Extraction*
Wuwei Lan, Yang Chen, Wei Xu and Alan Ritter
Proceedings of EMNLP 2020 (short paper) (pdf)

[7] *Neural CRF Model for Sentence Alignment in Text Simplification*
Chao Jiang, Mounica Maddela, **Wuwei Lan**, Yang Zhong and Wei Xu
Proceedings of ACL 2020 (pdf)

[6] *Travel Time Estimation without Road Networks: An Urban Morphological Layout Representation Approach*
Wuwei Lan, Yanyan Xu and Bin Zhao
Proceedings of IJCAI 2019 (pdf)

[5] *Neural Network Models for Paraphrase Identification, Semantic Textual Similarity, Natural Language Inference, and Question Answering*
Wuwei Lan and Wei Xu
Proceedings of COLING 2018, **Best Paper Award** (pdf)

[4] *Character-based Neural Networks for Sentence Pair Modeling*
Wuwei Lan and Wei Xu
Proceedings of NAACL 2018 (short paper) (pdf)

[3] *A Continuously Growing Dataset of Sentential Paraphrases*
Wuwei Lan, Siyu Qiu, Hua He and Wei Xu
Proceedings of EMNLP 2017 (pdf)

[2] *PAC: Taming TCP Incast Congestion Using Proactive ACK Control*
Wei Bai, Kai Chen, Haitao Wu, **Wuwei Lan** and Yangming Zhao
Proceedings of ICNP 2014 (pdf)

[1] *Facial Expression Recognition using Deep Boltzmann Machine from Thermal Infrared Images*
Shan He, Shangfei Wang, **Wuwei Lan**, Huan Fu, and Qiang Ji
Proceedings of ACII 2013 (pdf)

TECHNICAL SKILLS

◇ Tools: PyTorch, TensorFlow, Keras, Torch
◇ Language: C/C++, Python, Java, Matlab, HTML/CSS, Javascript

AWARDS

◇ Best Paper Award for COLING 2018.
◇ Guo Moruo Scholarship, the best scholarship of USTC, only 32 award winners, 2014.
◇ CCF Outstanding Undergraduate Award, 100 award winners in China, 2013.
◇ Google Excellence Scholarship, 100 award winners in China, 2013.
◇ National Scholarship, 2 out of 109 in CSE, USTC, 2012.

TALKS	⋄ Neural Network Models for Sentence Pair Modeling, COLING 2018
	⋄ Automatic Paraphrase Collection and Identification in Twitter, MSLD 2018
	⋄ Automatic Paraphrase Collection and Identification in Twitter, OSU 2017
	⋄ A Continuously Growing Dataset of Sentential Paraphrases, MASC-SLL 2017
SERVICE	⋄ PC Member for Conference on ACL, 2020, 2021
	⋄ PC Member for Conference on EMNLP, 2018, 2020
	⋄ PC Member for Conference on NAACL, 2018, 2019, 2021
	⋄ PC Member for Conference on COLING, 2018, 2020
	⋄ PC Member for workshop of ACL-SRW, 2018
	⋄ PC Member for workshop of MASC-SLL, 2017
	⋄ PC Member for workshop of WNUT at EMNLP, 2017-2020
TEACHING	⋄ CSE-3521 (Instructor): Artificial Intelligence I: Basic Techniques
	⋄ CSE-3521 (Instructor): Introduction to Artificial Intelligence
	⋄ CSE-5522 (TA): Artificial Intelligence II: Advanced Techniques
	⋄ CSE-5525 (TA): Speech and Language Processing
	⋄ CSE-2111 (TA): Modeling and Problem Solving with Database
AFFILIATIONS	The Association for Computational Linguistics (ACL)
REFERENCES	**Wei Xu**
	Assistant Professor
	School of Interactive Computing
	Georgia Institute of Technology
	✉ wei.xu@cc.gatech.edu