KAI WANG

Concordia University ◇ Montreal, Canada
(+1) 438 927 6776 ◇ wangkai881617@gmail.com ◇ https://key2miao.github.io/

EDUCATION

Concordia University, Montreal, Canada *Jan, 2019 – Present*

- **MASc**, Electrical and Computer Engineering, GPA: 3.7
- Research Assistant, Advisor: Prof. Wei-Ping Zhu

North University of China (NUC), Shanxi Province, P.R. China *Sep, 2012 – Jul, 2016*

- **Bachelor** of Engineering in Electronic Information Engineering
- Research Assistant, Advisor: Prof. Liming Wang

RESEARCH INTEREST

Fields Signal Processing, Computer Vision, Multi-Modal Learning, Unsupervised Learning

Methods Deep Learning, Self-Supervised Learning, Reinforcement Learning, Graph Neural Networks

PUBLICATIONS & MANUSCRIPTS

1 **Kai Wang**, Bengbeng He, Wei-Ping Zhu. CPTNN: Cross-Parallel Transformer Neural Network for Speech Enhancement in the Time Domain. Submitted to *IEEE/ACM Transactions on Audio, Speech, and Language Processing (TASLP)*. Under review.

2 **Kai Wang**, Bengbeng He, Wei-Ping Zhu. SE-MLP: An MLP Neural Network for Speech Enhancement. Submitted to *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022*. Under review.

3 **Kai Wang**, Bengbeng He, Wei-Ping Zhu. TSTNN: Two-stage Transformer based Neural Network for Speech Enhancement in the Time Domain. Accepted by *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021*. [arXiv]

4 **Kai Wang**, Bengbeng He, Wei-Ping Zhu. CAUNet: Context-Aware U-Net for Speech Enhancement in Time Domain. Accepted by *IEEE International Symposium on Circuits and Systems (ISCAS) 2021*. [IEEE Xplore]

RESEARCH EXPERIENCE

Huawei Noah's Ark Laboratory, Montreal, Canada *Jul, 2021 – Present*
Research Intern, Advisor: Dr. Chao Xing, Dr. Mehdi Rezagholizadeh and Dr. Anderson Avila
Project: Self-supervised Speech Pre-training based on Feature Disentanglement

- Applied feature disentanglement into speech pre-training to distangle environment audio features and speech features
- Proposed an encoder-decoder based architecture with time reduction, where the decoder is adopted to recover speech waveform to evaluate the generated feature embeddings from encoder
- Achieved competitive performance in downstream tasks including spoken language understanding (SLU) and keyword spotting (KWS) compared with state-of-the-art methods

- Working on paper submission to IJCAI'2022

Project: Disentangled Audio-Visual Representation Learning

- Applied feature disentanglement and contrastive function to extract mutual information between audio and visual modalities
- Achieved impressive performance in downstream tasks including visual speech enhancement, speaker separation and audio-visual retrieval
- Working on paper submission to ICML'2022

Concordia University, Montreal, Canada *Sep, 2019 – Present*
Research Assistant, Advisors: Prof. Wei-Ping Zhu

Project: Transformer based Neural Network for Speech Enhancement in the Time Domain

- Proposed a simple and efficient segmentation approach to process long-range speech sequences
- Proposed a two-stage transformer neural network (TSTNN) to extract the local and global information of long-range speech sequences, achieving competitive performance compared with state-of-the-art methods while having low model parameters. Accepted by ICASSP'2021
- Proposed a cross-parallel transformer neural network (CPTNN) to parallelly extract the local and global information, which are fused by proposed cross-attention based transformer to generate the contextual representation. Proposed CPTNN outperforms most of the state-of-the-art methods with the lowest model complexity. In the submission to TASLP

Project: Attention Mechanism based Encoder-Decoder Architecture for Speech Separation

- Applied the dilated-dense block into the encoder and decoder to enlarge the receptive field of features
- Proposed a context-aware U-Net (CAUNet) based on the self-attention transformer to map the noisy speech sequence into clean speech. Proposed CAUNet obtains a comparable performance with existing methods. Accepted by ISCAS'2021
- Proposed TransUNet, an extension of CAUNet that incorporate transformer structure into encoder and decoder to solve the limited receptive field of CNN structure. This works will be submitted to IEEE Signal Processing Letters

Project: Lightweight and Efficient MLP Architecture for Speech Processing

- Explored attention-free architecture for speech signal processing
- Proposed a simple and efficient MLP neural network (SE-MLP) to extract multiscale temporal informaiton and global frequency information of spectrogram features
- Achieved a competitive performance compared with attention or transformer based neural networks. This work has been submitted to ICASSP'2022

Project: Exploring Efficient Vision Transformer based on Self-Supervised Learning

- Explored the differences between contrastive and reconstruction based vision pre-training schemes
- Explored the efficient vision transformer with sparse attention mechanism. Also Applied convolutional module into image tokenization to make it flexible for different image size
- Combined the masked auto-encoder and contrastive function to learn semantic information of image by using vision transformer. Improved the performance of downstream vision tasks such as image classification, image segmentation and object detection

North University of China, Taiyuan, P.R. China *Sep, 2012 – Jul, 2016*

Project: Omni-Directional Car based on Inertia Sensor ((Provincial Level))
Research Assistant, Advisors: Prof. Wen Hou

- Implemented the gyroscope, accelerometer and inertial sensors to realize the navigation and positioning of the car in the GPS blind area

Project: Threshold Processing and Edge Detection in Image Segmentation based on Machine Learning
Research Assistant, Advisors: Prof. Yingliang Zhao

- Explored threshold processing and edge detection with different learning algorithms including SVM and AdaBoost

Project: Efficient Positioning System based on Underwater Wireless Sensor Network
Research Assistant, Advisors: Prof. Liming Wang

- Explored efficient positioning system based on underwater wireless sensor network and correlation algorithms by using MATLAB software

WORK EXPERIENCE

Microchip Technology Inc, Ottawa, Canada

Research Intern *Jun, 2020 – Mar, 2021*

- Developed novel deep learning algorithms for speech enhancement and speech separation
- Explored lightweight and efficient network structures for real-time speech-related applications such as automatic speech recognition (ASR), keyword spotting (KWS) and voice activity detection (VAD)

Beijing Tsinghua Tongfang Co. Ltd, Beijing, P.R. China

Electronic Engineer Assistant in **Research Department** *Sep, 2017 – Jul, 2018*

- Developed efficient deep learning algorithms for face and fingerprint recognition
- Explored model compression algorithm of face recognition for the application used in mobile devices

Beijing Becstrong Teaching Apparatus Co. Ltd, Beijing, P.R. China

Research Intern in **Research Department** *Jul, 2016 – Aug, 2017*

- Built Microprogrammed Control Unit (MCU) hardware system for mobile robot
- Developed machine learning algorithm for intelligent mobile robot to achieve automatic navigation and obstacle-avoidance

SCHOLARSHIPS & AWARDS

- **Concordia University Conference and Exposition Award** *2021*
- **Meritorious Award** of **International Mathematical Contest in Modeling** (top 5%) *2015*
- **National Encouragement Scholarship** (top 1%) *2015*
- **2th place** in **National Electronic Design Contest** (top 5%) *2015*
- **2th place** in the 6th **Provincial "Lanqiao" Cup Contest, Embedded System** *2015*

- **Outstanding Award** of **Summer Camp in Tianjin university** (top 2%) *2015*
- **3th place** in **National Mathematical Contest in Modeling** (top 5%) *2014*
- **1th place** in **Mathematical Contest in Modeling** (top 2%) *2014*
- **2th place** in the 5th **Provincial "Lanqiao Cup" Contest, Microcontrollers** *2014*
- **3th place** in **"YaqianCup" Robot Design Contest** *2013*

EXTRACURRICULAR ACTIVITIES

- **Volunteer teacher** of **Dingxiang county Primary School** in Shanxi Province *2016*
- **Commissary in Charge of Studies** in the North University of China *2012-2016*
- **Member of Summer Camp** in **Xidian University** *2015*
- **Member of Summer Camp** in **Tianjin University** *2015*
- **Assistant Teacher in charge of class** in the North University of China *2014-2015*
- **Minister** of the **Practical Department of Election Association** *2014*
- **Vice Minister** of the **Practical Department of School Students' Union** *2013-2014*
- **Member** of the **Organization Department of School Students' Union** *2012-2013*

PROFESSIONAL SERVICE

Reviewer for conferences ISCAS'2021, NEWCAS'2021

PROGRAMMING SKILLS

Proficient Python, PyTorch, MATLAB, Java, LaTeX

Familiar Linux, C/C++, TensorFlow, Keras, etc.