Kai Wang

Email: wangkai881617@gmail.com
Mobile: +438-927-6776

Research Interests:

My research interests are in the areas of machine learning and deep learning. I am particularly interested in speech and natural language processing, computer vision, self-supervised learning and transfer learning.

Education

Concordia University, Montreal, Canada 2019-present

MASc, Electrical and Computer Engineering (GPA: 3.7/4.0)

Advisor: Prof. Wei-Ping Zhu

North University of China (NUC), Shanxi Province, P.R. China 2012-2016

Bachelor of Engineering in Electronic Information Engineering

Advisor: Prof. Liming Wang

Publications

1. ***TSTNN: Two-stage Transformer based Neural Network for Speech Enhancement in the Time Domain***
   Kai Wang, Bengbeng He, Wei-Ping Zhu, accepted by *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021.
2. ***CAUNet: Context-Aware U-Net for Speech Enhancement in Time Domain***
   Kai Wang, Bengbeng He, Wei-Ping Zhu, accepted by *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2021.

Academic Experiences

Academic experience in graduate:

★ Speech Separation based on Developed Deep Generative Models

Research Assistant, Concordia University

Advisor: Prof. Wei-Ping Zhu, collaborated with Microchip in Ottawa, Canada

- Explored Generative Adversarial Networks (GANs) based models to build a mapping from the noisy speech to clean speech.
- Optimized the model structure to achieve the light model parameters compared with existing the architectures.

★ Transformer based Neural Network for Speech Denoising

Research Assistant, Concordia University

Advisor: Prof. Wei-Ping Zhu

- Proposed a two-stage transformer module (TSTM) to extract the local and global information of long-range speech feature.
- Built a sequence-to-sequence denoising model based on the TSTM by masking method in time domain, which outperforms most of the state-of-the-art methods in the benchmark dataset with the lowest model parameters.
- Paper has been accepted by *ICASSP 2021,* one of top conferences in signal processing.

★ Attention Mechanism based on Encoder-Decoder Architecture for Speech Separation

Research Assistant, Concordia University

Advisor: Prof. Wei-Ping Zhu

- Explored the encoder-decoder structure incorporating with attention mechanism for speech separation.
- Proposed a context-aware U-Net based on the transformer to map the noisy speech sequence into clean speech. The proposed architecture introduces the dilated-dense block into the encoder and decoder of U-Net to enlarge the receptive field of feature. Paper has been accepted by *ISCAS 2021.*

★ Self-Supervised Learning for Speech Representation

- Built the upstream model based on the BERT to learn the speech feature representation in unlabeled data.
- Explored different pretext tasks to build a strong pre-trained model. For example, adding random Gaussian Noise with various SNRs into clean speech to generate noisy speech, then the task is to recover the clean speech by training the model. The other

method is masking some input speech feature on time-frequency direction or channel direction, then the target is to recover those masked speech feature.

- Applied the pre-trained model into downstream tasks such as speech recognition, speech denoising, voice conversion, speaker identification, etc.

★ Self-Supervised Audio-Visual Representation Learning

- Built self-supervised audio-visual model to learn the cross-modal representation between audio and vision by solving some pretext tasks including audio-visual synchronization, vision-guided audio predication or audio-guided vision generation, etc.
- Applied the pre-trained audio-visual model into different downstream tasks speaker separation, sound source localization and action recognition.
- Learned cross-modal information could be utilized into improving performance of single-modal task under the guidance of the other modality.

★ Heart-Sound Detection based on the Machine Learning

Research Assistant, Concordia University

Advisor: Prof. Wei-Ping Zhu

- Built convolutional neural network based on U-Net structure to classify the normal and abnormal heart sound signal.
- Built sequence models based on LSTMs or GRUs to achieve the segmentation of heart sound (S1, S2) on PCG signal.

★ Object Detection based on Supervised and Self-Supervised Learning

- Explored the Faster R-CNN and YOLO architecture for pedestrian detection.
- Explored the self-supervised model to learn the visual representation by solving pretext tasks such as add gaussian noise, image rotation, Jigsaw Puzzles, etc. Then, applying the pre-trained model to help improve the performance of downstream tasks including object detection, image classification and image segmentation.
- Explored the visual transformer on vision tasks.

★ Applied Machine Learning Training

Advisor: Prof. William L. Hamilton at McGill University, MILA

- Analysis of Logistic Regression and Linear Discriminant Analysis for Two Binary Classification Tasks
- Explored Convolution Neural Networks (CNNs) in Multi-class Classification Image Problem on Modified MNIST Dataset with data augmentation methods, achieving the 98.49% accuracy.
- Developed supervised classification models to analyze text from the website Reddit, including Naïve Bayes classifier, Multinomial Naïve Bayes, Support Vector Machine, Decision Tree and Adaptive Boosting classifier.
- NeurIPS 2019 Reproducibility Challenge: reproduced paper is "*Adversarial Examples Are Not Bugs, They Are Features*".

Academic experience in undergraduate:

★ Omni-Directional Car based on Inertia Sensor (Provincial Level)

Advisor: Prof. Wen Hou

Aim: The project combined the gyroscope, accelerometer and inertial sensors to realize the navigation and positioning of the car in the GPS blind area.

★ Threshold Processing and Edge Detection in Image Segmentation based on Machine Learning

Advisor: Prof. Yingliang Zhao

Aim: Explored threshold processing and edge detection with different learning algorithms including SVM and AdaBoost.

Research and Work Experiences

Research Assistant, Microchip Technology Inc. 06/2020 - Present

- Implement deep learning based denoising methods for automatic speech recognition
- Develop novel deep learning based methods for speech and natural language processing, such as speech enhancement, keywords detection and voice activity detection.

Electronic Engineer Assistant, Research Department, Beijing Tsinghua Tongfang Co. Ltd 09/2017 - 07/2018

- The research of face and fingerprint recognition based on machine learning.

Research Intern, Beijing Becstrong Teaching Apparatus Co. Ltd 07/2016 - 08/2017

- The research of intelligent mobile robot based on machine learning.

Research Assistant, National Electrician and Electronics Experimental Center, NUC 09/2013 - 06/2016

Advisor: Prof. Liming Wang

- The research of positioning system based on underwater wireless sensor network.

Awards and Honors

✧	National Encouragement Scholarship | National Level, top 1%	09/2015
✧	Second Prize of National Electronic Design Contest | Provincial Level	09/2015
✧	Outstanding Participants of Summer Camp in Tianjin university | National Level, top 2%	07/2015
✧	Second Prize of the 6th "Lanqiao Cup" Contest, Embedded System | Provincial Level	04/2015
✧	Meritorious Award of Mathematical Contest in Modeling (MCM) | International Level, top 7%	01/2015
✧	Third Prize of National Mathematical Contest in Modeling | Provincial Level	09/2014
✧	First Prize of Mathematical Contest in Modeling | School Level	06/2014
✧	Second Prize of the 5th "Lanqiao Cup" Contest, Microcontrollers | Provincial Level	01/2014
✧	Comprehensive quality scholarship | School Level	06/2013
✧	Third Prize of "Yaqian Cup" Robot Design Contest | School Level	06/2013
✧	Honorable Mention of Program Design Contest | School Level	05/2013

Extracurricular Activities

Volunteer teacher of Dingxiang county Primary School| Shanxi	06/2016
Commissary in Charge of Studies| NUC	09/2012-06/2016
Member of Summer Camp| Xidian University	08/2015
Assistant Teacher in charge of class daily activities| NUC	09/2014-09/2015
Minister of the Practical Department of Election Association| NUC	10/2014
Vice Minister of the Practical Department of School Students' Union| NUC	09/2013-06/2014
Member of the Organization Department of School Students' Union| NUC	09/2012-06/2013

Professional Service

Reviewer for conferences: *ISCAS 2021*

Skills

Programming Language: Python, Matlab, C, C++ and Java

Deep Learning Framework: PyTorch, Tensorflow

Others: Linux, Cluster