ZHENHE ZHANG

Principle Researcher
BEng, MRes, Ph.D.
At Huaiwei Noah's Ark Lab

zhenhezhang0601@sina.com
https://www.linkedin.com/in/zhenhezhang-0601/
Tel/WeChat: (+86) 13691651909
Huawei Ltd. Bantian F3 Building
Shenzhen, Guangdong Province
China

WORK EXPERIENCE

1. Principle Researcher at Huaiwei Noah's Ark Lab   03/2019 - Present

• *Data Engineering in LLM (Large Language Modeling)*
- *Leading the CommonCrawl filtering pipeline Design and Development. The pipeline consists of three steps:*
  stage1 is general cleaning (including the modification of documents)
  stage2 is language specific cleaning (filtering documents only)
  stage3 is quality filtering using classifiers
- *Stage1 and stage2 are mostly regex expressions, and stage3 uses classifiers, e.g. NgramLM filtering, logistic regression and text classification*
- *Data are stored on Apache Hadoop Platform. And algorithms are running in Hadoop Streaming applications*

• *Self-Supervised Perturbation Invariant Representation Learning For Speech Pre-training (SPIRAL)*
- *Supervision on a new approach for speech pre-training named SPIRAL, published on ICLR 2022*
- *Our approach works by learning de-noising representation of perturbed data in a teacher-student framework*
- *It differs from wav2vec2.0 by the latter performs MLM in intermediate latent space and performs target discretization with a differentiable quantization scheme*
- *It differs from HuBERT approach by the latter predicts hidden cluster assignments of the masked frames, which are generated by one or more iterations of k-means clustering*
- *Our approach does not need quantization scheme nor iterations of k-means clustering, but achieves competitive or better results compared to state-of-the-art with significant reduction of training cost (80% for BASE model, 65% for LARGE model)*

• *Productize End-to-End Speech Recognition based on RNNT and ConvTransformer-T*
- *Productized an End-to-End Speech Recognition System On-Device and Cloud Server*
- *The model is the integration of convolution and Transformer. And the streaming pattern is realized by the truncated size window of convolution and uni-directional attention*
- *Contexture biasing is applied in the fashion of ShallowFusion and Rescoring depending on the computing capability*
- *The terminology is realized by WFSTs supporting dynamic adding and removal*

• *Productize Keyword Spotting HeyCelia overseas on Huawei mobile*
- *Keyword Spotting based on LVCSR. The role is to optimize the acoustic model training in NNET3/NNET2 in Kaldi*
- *Design and develop the decoding HCLG graph of the specific keyword, in order to achieve high accuracy and low false acceptance*

2. Senior Engineer at Huaiwei ICT Business Group,   10/2017 - 03/2019

• *One year project management experience: Being the owner of the research project of AI speech and semantic in Hybrid Video department*
- *The research proposal is the on-device application on speech recognition, voiceprint recognition, keyword spotting and far-field speech techniques to the advanced set top box with limited computing resources*
- *The project has achieved outstanding performance beyond the expectation. All the AI applications have landed in the targeting service provider and provides the intelligent service to the users of the advanced set top box*

3. Research assistant at Birmingham Centre for Railway Research and Education (BCRRE),   09/2015 - 09/2017

• Being participated in two research projects funded by HITACHI Europe:

- *One is railway vision project involving the computer vision kernel of convolution (CNN) development of CCTV surveillance system on the embedded platform*
- *The difficulty is to make the algorithm running on-device in real-time*
- *The solution involves of ALU (Algorithm Logic Unit) design in order to provide the same experiment results simulated in Simulink.*
- *The other project is the condition monitoring on roller bearing. Apply the acoustic signal analysis and processing techniques to the real-world roller bearing data.*
- *The difficulty is to retrieve the diagnosis information of roller bearing from low SNR acoustic signal.*
- *The solution invovles the combination of EMD (Empirical Mode Decomposition) and MED (Minimum Entropy Deconvolution) to identify early stage roller bearing faults in noisier signals.*

4. Laboratory Assistant at University of Birmingham, U.K.   09/2015 - 09/2017

• The laboratory assistance covers the modules in BEng and Master course

- *BEng modules include Microprocessor Systems, Computer Hardware and Digital Design (FPGA platform based)*
- *Master module include Advanced Digital Design (Master module), Small Embedded Systems (Master module)*

EDUCATION

University of Birmingham, Birmingham, U.K.   09/2012 - 07/2017

Ph.D. in Electrical, Electronics and Systems Engineering

- Research Areas: Improving railway vehicle structure inspection and monitoring through the integration of multiple monitoring technologies; Acoustic sensing technology; Non-destructive testing (NDT); Condition monitoring

University of Birmingham, Birmingham, U.K.   09/2011 - 12/2012

MRes (Master of Research) in Electrical, Electronics and Systems Engineering

- Research Areas: Developing an FPGA system for parallel processing of railway non-destructive testing data; FPGA; Hardware system design; Non-destructive testing (NDT); Condition monitoring

University of Birmingham - Birmingham, U.K.   09/2010 - 07/2011

B.Eng in the School of Electrical, Electronics and Systems Engineering

- Final Year Project: Music Genre classification through Pattern Recognition techniques, including GMM and HMM models

Huazhong University of Science and Technology - Wuhan, China   09/2007 - 07/2010

B.Eng in the School of Electronic Information and Communications

- Excellent graduate with GPA 80+/100

LANGUAGES & SKILLS

Languages: English (Full professional proficiency), Mandarin (Native)

Skills: wav2vec · Hubert · RNNT · Fairseq · Nemo · Transformer · OpenSeq2Seq · C/C++ · Kaldi · Openfst · Kenlm · Festival/Flite · MFA · Speech Recognition · Machine Learning · Artificial Intelligence · Deep Learning · Digital signal processing · Matlab · VHDL/Verilog · Assembly Language · Signal Propagation Model simulation · Embedded Systems · Xilinx ISE & ModelSim

Academic Skills: Data Mining, Academic writing, Self-driven, academic project management, academic team building

PUBLICATIONS

• **SPIRAL: Self-supervised Perturbation-Invariant Representation Learning for Speech Pre-Training**
ICLR 2022 · Jan 25, 2022
• **Enhanced fault diagnosis of roller bearing elements using a combination of Empirical Mode Decomposition and Minimum Entropy Deconvolution**
Proc IMechE Part C: J Mechanical Engineering Science · Nov 2015
• **Improvement of Axle Bearing Monitoring Systems Through the Use of High Speed Imaging for Directing Acoustic Beamforming**
Condition Monitor/The British Institute of Non-Destructive Testing (BINDT) · June 2015
• **Acoustic analysis techniques for condition monitoring of roller bearings**
The 6th IET Conference on Railway Condition Monitoring (RCM 2014) · Sep 2014
• **Development of an FPGA system for parallel processing of railway non-destructive testing data**
UBIRA eTheses · Dec 2014

AWARDS

Oversear High-Caliber Personnel (Level C) in Shenzhen   May 2018 - Present
Issued by Human Resources and Social Security Administration of Shenzhen Municipality
• Awarded the Innovation Sponsorship ¥1,600,000
"Shenlong High-level Talent", Category C, of Longgang District in Shenzhen   September 2018 - Present
Issued by Human Resources and Social Security Administration of Longgang district in Shenzhen
• Awarded the Innovation Sponsorship ¥800,000
Ph.D. Scholarship   September 2012
• Fully-funded Ph.D.: £15,150/year School Scholarship covering PhD tuition fee plus
   £13,863/year BCRRE Scholarship covering living costs through the Ph.D. course
Master Scholarship   September 2011
• £6,000 Scholarship: The European commission through the FP7 INERAIL project
Undergraduate Scholarship for Excellent Direct-Entry Oversea Students   September 2010
£2,500/year Scholarship by the administration of University of Birmingham
Scholarship for Excellent Direct-Entry Students   September 2010
• £2,500/year Scholarship for Excellent Direct-Entry Students: University of Birmingham.
The 1st prize in the 'TI cup' hardware design competition in 2011   June 2010
• The competition is the design of a waveform generator using "Texas Instruments" components
• The advances of our solution is the continues of resolution adaption, which applies a sine function
   to map the resolution steps to the user input
• The solution is commonly used in high frequency waveform transmission. In this project, it allows
   the waveform generator achieve high resolution with a small memory cost