Khandokar Md. Nayem

nayem.github.io

+1 (812) 369-6358

knayem@iu.edu

TECHNICAL EXPERTISE

• Speech Enhancement, Processing & Recognition • Machine Learning • Deep Learning • Natural Language Processing (NLP) • Classification & Regression • Deep Neural Network (DNN) • Recurrent Neural Network (RNN) • Long-Short Term Memory (LSTM) • Convolutional Neural Network (CNN) • Machine Learning Models (Bayesian Network, HMM, Clustering, Decision Tree, Ensemble Methods) • Deep Learning Libraries (TensorFlow, Keras) • NLP Libraries (NLTK, Scikit-Learn) • Human Survey Platform (Qualtrics, Amazon MTurk) • Computer Vision and Graphics Libraries (CImg, OpenCV, OpenGL) • Python • Matlab • C\C++ • Java • HTML, CSS & Javascript • SQL • R • Shell Script

EMPLOYMENT

Indiana University, Bloomington, IN, *Research Assistant, ASPIRE research lab* — Fall 2016 - Present

- Developed an attention-based monaural speech enhancement model that aims to maximize human perceptual rating of the enhanced speech by incorporating embedding vectors from a human Mean-Opinion Score (MOS) prediction model and jointly training the models on real-world noisy speech data. (INTERSPEECH-2021)
- Proposed and implemented a quantized speech prediction model that classifies speech spectra into a corresponding quantized class and applies a language-style model to ensure more realistic speech spectra. Acceptable quantization level is determined by a listener study ran in Amazon MTurk designed in Qualtrics. (ICASSP-2021, poster, slides, video)
- Designed a recurrent layer named Intra-Spectral Recurrent (ISR) layer that captures spectral dependencies within the magnitude and phase responses of the noisy speech using Markovian recurrent connections, and successfully deployed in a LSTM-based single-channel speech enhancement model. (ICASSP-2020, slides, video)
- Formulated a new type of recurrent output layer that enforces spectral-level dependencies within each spectral time frame modeling the Markovian assumption along the frequency axis in both uni-directional and bi-directional ways, and tested in a magnitude speech enhancement model. (MLSP-2019, poster)
- Engineered a deep architecture named Recurrent Stacked Generative Adversarial Network (RSGAN) which generates video clips based on a pre-condition like a sentence description, action classes, or fMRI signals. (IU-VISION-2017, poster)
- Designed a prediction model for nulliparous pregnant women to diagnose and prevent gestational diseases like diabetics, pre-eclampsia, and hypertension as part of the IU Public Health project. Also, developed a smart system that traces the physical activity of women on daily basis collected by wearable devices and helps to diagnose gestational complications.
- Investigated speech emotion detection deep architectures like RNN and CNN in the scale of arousal and valence.
- Developed an automatic classification of rhetorical questions with stress detection on our own collected dataset using Recurrent Neural Network (RNN) and Convolutional Recurrent Neural Network (CRNN) models. (code)
- Implemented an end-to-end speech recognition system for the English language using bi-directional recurrent neural network architecture without any frame wise labeling. (code)

BOSE Corporation, Boston, MA, *Machine Learning/Neural Signal Processing Intern* — Summer 2020

- Researched enhancing speech in remote microphone applications by self-speech removal to provide better quality sound with low latency to the hearing aids and voice-assistive wearable devices. LSTM-based architecture with speaker dependent d-vector is used for real-time operation.

United International University (UIU), Dhaka, *Lecturer, Department of CSE* — August 2016

- Taught courses of Computer Science curriculum, like C++ Programming language, Algorithms, Digital Logic Design and Pattern Recognition courses in classes of more than 90 undergrads.

REVE Systems, Dhaka, *Jr. Software Engineer, Team Media Gateway* — January 2015

- Programmed media gateway controller to facilitate both calls and faxes between the telephone network and VoIP network or another telephone network via Megaco 1.0 protocol. Also designed front-end panel by .JSP framework for VoIP administrators and customers for easy use.

EDUCATION

Ph.D. in Computer Science, Fall 2022 (Anticipated)
Indiana University, Bloomington, IN — Advisor: Prof. Donald S. Willamson

M.Sc. in Computer Science, December 2019
Indiana University, Bloomington, IN

B.Sc. in Computer Science & Engineering (CSE), July 2014
Bangladesh University of Engineering & Technology (BUET), Dhaka, Bangladesh