SHUYANG ZHAO

PhD Candidate | +358449732012

DETAILS

+358449732012
shuyang.zhao@tuni.fi

DATE OF BIRTH

09.09.1987

LANGUAGES

Chinese

English

Finnish

PROFILE

My main proficiency is in machine learning and audio content analysis. In addition, I also gained basic experience with node servers.

EDUCATION

Ph.D., Tampere University, Tampere

January 2016 — July 2020

I have been working for nearly seven years in the audio research group of Tampere University of Technology (recently merged into Tampere University). One of the main topics of the research group is detection and classification of acoustic scenes and events (DCASE). We are one of the main organizers of DCASE challenge.

My own research topic is active learning for DCASE problems. The target is to optimize the performance of learned DCASE models with a limited annotation effort, by the means of recommending the most representative sounds for annotation.

In addition to my own research topic, I have been working on a few collaborative research projects with partners including VTT, Nokia and DSP Group. My main role in the projects is to develop DCASE models. The details of the projects are shown in the next Section.

M.Sc., Tampere University of Technology

August 2010 — June 2014

Major in Signal Processing. The curriculum includes digital linear filtering, digital image processing, knowledge mining, speech recognition, and algorithm analysis.

B.Sc., Huazhong University of Science and Technology

September 2005 — June 2009

Major in Biomedical Engineering. The curriculum includes organic chemistry, anatomy, physiology, medical imaging, and linear systems.

PROJECTS

SmartSound

January 2018 — Present

The project goal is to introduce acoustic sensor networks into elderly care centers for non-intrusive health monitoring. I am responsible for the development of the prototype system. Commercial sensor nodes (Minut Point device) are used to stream on-site audio to a node server on cloud, where sound event detection is performed.

INCA

January 2017 — December 2017

The project goal is to develop a real-time acoustic scene classifier, determining whether the location of the microphone is in a driving car. Demos are made with speakers playing back the sounds recorded inside cars. Impulse responses are measured in office environments to match the demo condition.

VoiceActive

January 2016 — December 2016

The goal of the project is to develop a speech/singing classification for close microphone scenario. The system is used for automatic transcription of stage performances.

InteractiveAudio

January 2015 — December 2015

The project goal is to develop a system, where it learns sound classification model interactively from an end-user. My colleague develops the GUI for the interaction part. I am responsible for the selection of sounds that are presented to end-users for annotation, and the training of the acoustic model.

My Ph.D. topic comes from the idea formed in this project.

AKU

April 2014 — December 2015

The project goal is to classify noise sources in an environmental noise monitoring sensor node. The aim is to assign the sound pressure level to the most prominent noise sources and find a time pattern for the noise sources.

PUBLICATIONS

An active learning method using clustering and committee-based sample selection for sound event classification

In proc. 16th International Workshop on Acoustic Signal Enhancement (IWAENC), 2018

Environmental Noise Monitoring Using Source Classification in Sensors

Applied Acoustics, Volume 129, Pages 258–267, 2018.

Active learning for sound event classification by clustering unlabeled data

In proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), p. 751--755, 2017.

The paper is nominated as the best student paper award finalist.

Learning vocal mode classifiers from heterogeneous data sources

In proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), p. 16--20, 2017.

Active learning for sound event detection

IEEE/ACM Transactions on Audio, Speech, and Language Processing, in progress