# Yuan Gong

Updated August 9, 2021

Cambridge, MA · +1(574)4010-833 · yuangong@mit.edu · https://yuangongnd.github.io

## Work Experience

**Massachusetts Institute of Technology** — Cambridge, MA, USA
Postdoctoral Researcher at CSAIL — Aug 2020 - present
Host: Dr. James Glass
Working on speech and natural language processing including automatic speech recognition, audio event detection, and self-supervised representation learning.

**Amazon Web Service** — Seattle, WA, USA
Research Scientist Intern at Comprehensive Medical Team — May 2019 - Aug 2019
Supervisor: Mohammed Khalilia *and* Parminder Bhatia
Investigated contextualized representation (e.g., BERT) for extracting relationships and entities from clinical text in multiple languages and improved the model performance.

**University of Notre Dame** — Notre Dame, IN, USA
Research Assistant at Mobile Computing Lab — Aug 2015 - Jul 2020
Advisor: Christian Poellabauer
Led the speech healthcare project and the speech security project, won the AVEC 2017 depression detection challenge, and mentored 10+ undergrade student for research. One paper was nominated for Interspeech 2019 best student paper award.

**Philips Healthcare** — Shanghai, China
Summer Intern — Jul 2012 - Aug 2012
Coordinated maintenance information collection and conducted on-site maintenance for CT/MR equipment places in 50+ hospitals throughout Shanghai.

## Education

**University of Notre Dame** — Notre Dame, IN, USA
Ph.D. in Computer Science — 2015-2020
Advisor: Prof. Christian Poellabauer, GPA: 4.0
Thesis: Healthcare Applications and Security Concerns of Speech Processing Systems.

**Fudan University** — Shanghai, China
B.S. in Electrical Engineering (*Biomedical Engineering Major*) — 2011-2015
Advisor: Prof. Yuanyuan Wang and Prof. Yuedong Xu, GPA Rank: 1/15
Thesis: Wavelet Transform Based Ultrasound Image Denoising.

## Service

Reviewer — Sep 2018 - present
Reviewed over 20 papers for IEEE TASL, TPAMI, TDSC, SPL, Interspeech 2021, and DCASE 2021.

ACM-BCB 2019 Travel Grant Committee Member — Aug 2019
Undergraduate Research Advisor, University of Notre Dame — Aug 2015 - Jul 2020
Dissertation Committee Member of Marisa Cameron (Master Student) — Apr 2017

## Technical Skills

**Math**
Real Analysis, Optimization, Neural Networks, Statistical Computing Methods.

**Computer Science**
Data Science, Machine Learning, Deep Learning, Speech Processing, Natural Language Processing, Algorithms, Computer Architecture, Operating Systems.

**Programming Languages**
Python, MATLAB, Java, C, Shell.

**Deep Learning Tools**
PyTorch, TensorFlow, MXNet.

**Languages**
English, Chinese, German.

## Selected Awards

| | |
|---|---|
| Interspeech 2019 Best Student Paper Award Nomination | Jul 2019 |
| AVEC 2017 Depression Detection Challenge Winner | Oct 2017 |
| Travel Award from IJCAI, NSF, ISCA, ICHI ($2500+) | Aug 2017 - Jul 2020 |
| Fudan First Prize Scholarship (Top 3%) and Outstanding Graduates | Apr/Jul 2015 |
| First Prize, TI National Biomedical Electronics Innovation Design Contest | Oct 2014 |

## Selected Publications (Google Scholar)

1. Yuan Gong, Yu-An Chung, and James Glass, *AST: Audio Spectrogram Transformer*, Interspeech 2021.

2. Yuan Gong, Yu-An Chung, James Glass, *PSLA: Improving Audio Tagging with Pretraining, Sampling, Labeling, and Aggregation*, IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2021 (under review).

3. Yuan Gong, Jian Yang, Christian Poellabauer, *Detecting Replay Attacks Using Multi-Channel Audio: A Neural Network-Based Method*, IEEE Signal Processing Letters, 2020.

4. Yuan Gong, Boyang Li, Christian Poellabauer, Yiyu Shi, *Real-time Adversarial Attacks*, The 28th International Joint Conference on Artificial Intelligence (IJCAI), 2019.

5. Yuan Gong, Jian Yang, Jacob Huber, Mitchell MacKnight, and Christian Poellabauer, *ReMASC: Realistic Replay Attack Corpus for Voice Controlled Systems*, Interspeech 2019 (*best student paper award nomination*).

6. Ning Xia, Yuan Gong, Yizhe Zhang, and Christian Poellabauer, *Non-local Second-order Attention Networks for Person Re-identification*, International Conference on Computer Vision (ICCV), 2019.

7. Yuan Gong and Christian Poellabauer, *Crafting Adversarial Examples For Speech Paralinguistics Applications*, The DYNAMICS Workshop in conjunction with the Annual Computer Security Applications Conference (ACSAC), 2018.

8. Yuan Gong and Christian Poellabauer, *Impact of Aliasing on Deep CNN-Based End-to-End Acoustic Models*, Interspeech 2018.

9. Yuan Gong, Hasini Yatawatte, Christian Poellabauer, Sandra Schneider, and Susan Latham, *Automatic Autism Spectrum Disorder Detection Using Everyday Vocalizations Captured by Smart Devices*, The 9th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM-BCB), 2018.

10. Yuan Gong, Kevin Shin, and Christian Poellabauer, *Improving LIWC Using Soft Word Matching (Poster)*, The 9th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (ACM-BCB), 2018.

11. Yuan Gong and Christian Poellabauer, *Protecting Voice Controlled Systems Using Sound Source Identification Based on Acoustic Cues*, The 27th International Conference on Computer Communications and Networks (ICCCN), 2018.

12. Yuan Gong and Christian Poellabauer, *An Overview of Vulnerabilities of Voice Controlled Systems*, 1st International Workshop on Security and Privacy for the Internet-of-Things, 2018.

13. Yuan Gong and Christian Poellabauer, *Topic Modeling Based Multi-modal Depression Detection*, The 7th Audio/Visual Emotion Challenge and Workshop (AVEC) in conjunction with ACM Multimedia (ACM-MM), 2017 (*challenge winner*).

14. Yuan Gong and Christian Poellabauer, *Continuous Assessment of Children's Emotional States using Acoustic Analysis*, The 5th IEEE International Conference on Healthcare Informatics (ICHI), 2017.

15. Yuan Gong, Jin Cao, Zehui Luo, and Guohui Zhou, *A Smart Low-Power-Consumption ECG Monitor Based on MSP430F5529 and CC2540*, Chinese Journal of Medical Instrumentation, 2015 (*Corresponding Project won TI National Biomedical Device Design Contest*).

## **Open-source Code** (GitHub)

1. Audio Spectrogram Transformer.

2. Neural Network Based Multi-channel Audio Antispoofing.

3. Real-time Adversarial Attacks.

4. ReMASC: Realistic Replay Attack Corpus for Voice Controlled Systems.