

Yuan-Hang Zhang

CONTACT INFORMATION

IIP Lab, Institute of Computing Technology
Chinese Academy of Sciences
6 Kexueyuan South Rd
Haidian, Beijing 100080
P. R. China
Homepage: www.sailorzhang.com

E-mail: zhangyuanhang15@mailsucas.ac.cn or yuanhang.zhang@vip1.ict.ac.cn

EDUCATION

**Institute of Computing Technology, Chinese Academy of Sciences (CAS),
Beijing, China**

M.Eng. in Computer Application Technology 09/2020 to present

- Supervisor: Prof. Shiguang Shan

University of Oxford, Oxford, UK

Visiting student in Computer Science 10/2018 to 03/2019

University of Chinese Academy of Sciences (UCAS), Beijing, China

B.Eng. in Computer Science and Technology 09/2015 to 07/2019

- GPA: 3.75/4.0
- Supervisor: Prof. Shiguang Shan and Dr. Shuang Yang
- Bachelor thesis: *Key Technologies for Large-scale Visual Speech Recognition (Outstanding Bachelor Thesis award)*

ACADEMIC EXPERIENCE

Institute of Computing Technology, CAS, Beijing, China

Research Assistant in Computer Vision 05/2017 to 06/2019

Automated Lip Reading of Chinese in the Wild

Supervisor: Prof. Shiguang Shan and Dr. Shuang Yang

- Investigated methods in shot boundary detection, visual tracking, and multi-modal learning for speaker identification and speech recognition
- **Primary contributor** to the construction of the first large-scale Chinese lip reading database, *LRW-1000*, which covers large variations in pose, age and other speaker attributes, using broadcast news video collected over more than a year
- Helped prepare the 1st Mandarin Audio-Visual Speech Recognition Challenge (MAVSR) at ACM ICMI 2019
- **Runner-up** at AVA Active Speaker Detection track of the ActivityNet Challenge 2019, in collaboration with Jingyun Xiao (Google Announcement)

Department of Foreign Languages, UCAS, Beijing, China

Research Student 12/2016 to 03/2018

Developing Coherence by Machine Readable Items In EAP Speaking Activities

Supervisor: Prof. Qun Zheng

- Constructed an interpretable scoring mechanism for the speaking section of Internet-based TOEFL (iBT) tests
- Accomplished by ranking the importance of different coherence metrics in linguistics with various machine learning techniques, such as LASSO regression

- Preliminary results presented at the “Faces of English 2” conference talk “*Developing Coherence by Machine Readable Items In EAP Speaking Activities*”

ADDITIONAL
EXPERIENCE

SeetaTech Technology Limited, Beijing, China

R&D Intern

04/2018 to 07/2018

- Integrated lip reading models into a fully functional demo system, trained on an internal Mandarin dataset to facilitate in-car voice control
- System demonstrated at VALSE 2018 and F&G 2019

Qingfan MOOP (now EduBrain Limited), Beijing, China

Project Participant

08/2015 to 09/2015

- *Project Name*: Mining emotion patterns in online social networks
- Natural language processing project with Dr. Hong Huang from Computer Networks (NET) Group at Institute of Computer Science, University of Göttingen
- Performed text sentiment analysis on Twitter datasets using NLTK and machine learning toolkits

AWARDS

CAS Undergraduate Student Scholarship

11/18, 12/17

UCAS Student Scholarship

09/21, 11/18, 11/17, 11/16

Merit Student Award, UCAS

12/18, 11/17, 11/17

Runner-up, AVA Active Speaker Detection track at ActivityNet Challenge 2019

06/2019

3rd prize, “Gulian Cup” Classical Chinese Literature Named Entity Recognition Task

10/2020

First place, AVA Active Speaker Detection track at ActivityNet Challenge 2021

06/2021

PUBLICATIONS

Y. Zhang, S. Liang, S. Yang, X. Liu, Z. Wu, S. Shan and X. Chen, *UniCon: Unified Context Network for Robust Active Speaker Detection*, ACM MM 2021 (**oral**). Co-first author. [\[link\]](#)

Y. Zhang, S. Liang, S. Yang, X. Liu, Z. Wu and S. Shan, *ICTCAS-UCAS-TAL Submission to the AVA-ActiveSpeaker Task at ActivityNet Challenge 2021*, technical report for AVA Challenge 2021. Co-first author. [\[link\]](#)

Y. Zhang, S. Yang, J. Xiao, S. Shan and X. Chen, *Can We Read Speech Beyond the Lips? Rethinking RoI Selection for Deep Visual Speech Recognition*, IEEE F&G 2020 (**oral**). First author. [\[link\]](#)

J. Xiao, S. Yang, **Y. Zhang**, S. Shan and X. Chen, *Deformation Flow Based Two-stream Network for Lip Reading*, IEEE F&G 2020 (poster). Third author. [\[link\]](#)

Y. Zhang, R. Huang, J. Zeng, S. Shan and X. Chen, *M³T: Multi-Modal Continuous Valence-Arousal Estimation in the Wild*, technical report for ABAW Challenge 2020. First author. [\[link\]](#)

Y. Zhang, J. Xiao, S. Yang and S. Shan, *Multi-Task Learning for Audio-Visual Active Speaker Detection*, technical report for AVA Challenge 2019. Co-first author. [\[link\]](#)

S. Yang, **Y. Zhang**, D. Feng, M. Yang, C. Wang, J. Xiao, K. Long, S. Shan and X. Chen, *LRW-1000: A Naturally-Distributed Large-Scale Benchmark for Lip Reading in the Wild*, IEEE F&G 2019 (**oral**). Co-first author. [\[link\]](#)

SKILLS &
MISCELLANEOUS

- Programming languages: Python, C++, C, MATLAB, Bash
- Machine learning frameworks: PyTorch, TensorFlow, JAX
- Proofread two chapters of Chinese translation for the textbook “Deep Learning” (Y. Bengio et al., Posts and Telecom Press, Beijing, 2017).
- Translated part of the textbook “Applied Machine Learning” (David Forsyth, China Machine Press, Beijing, 2021).