Yilin PAN

LinkedIn: Yilin Pan
Email: yilin.pan@sheffield.ac.uk
yilinpan.hit@gmail.com
GitHub: YilinSpeechandNLP

Education

University of Sheffield, *Doctor of Engineering*, Computer Science 2018–now | Sheffield, UK

- Research interest: Pathological Speech Processing; Speech Recognition; Natural Language Processing
- Supervisor/Research Group: Heidi Christensen/Speech and Hearing Research Group
- Thesis: Linguistic- and Acoustic-based Automatic Dementia Detection using Deep Learning Methods
- Scholarship: 2020 Horizon Marie Skłodowska-Curie Actions (TAPAS)

Harbin Institute of Technology, *Master of Engineering*, Computer Science (ESI top 1%) 2015–2017 | China

- Research interest：Speaker Verification; Speech Processing
- Supervisor/Research Group: Tieran Zheng/Speech Processing Lab
- Thesis: I-vector based probabilistic linear discrimination analysis for speaker verification research
- Scholarship: First Level Scholarship (twice)

Northeastern University, *Bachelor of Engineering*, Computer Science（National key Uni.） 2011–2015 | China

- Research interest：Pedestrian Detection；Signal Processing
- Supervisor/Research Group: Hai Zhao/Embedded Technology Lab
- Thesis：Multi-pedestrian detection method based on HOG and improved ViBe
- Scholarship: University Second-level Scholarship (twice); University third-level Scholarship (three times)

Work Experience

Philips, *Intern Researcher*, Pathological Speech Processing 2020.06–2020.09 | Eindhoven, Netherlands

- Research topic: Speech based dementia disease detection and longitudinal tracking

Iflytek, *Early Stage Researcher*, Natural Language Processing 2017.08–2017.10 | Beijing, China

- Research topic: Composition scoring and ancient poetry recommendation based on deep learning

Ruiming, *Intern Researcher*, Multi-modal data processing 2021.11–2022.01 | Chengdu, China

- Research topic: Multimodal Information Based Speech Processing

Challenges

- **Alzheimer's Dementia Recognition through Spontaneous Speech** 2020.06–2020.10
  The work was accepted by Interspeech 2020.
- **Alzheimer's Dementia Recognition through Spontaneous Speech only** 2021.02–2021.04
  The work has been accepted by Interspeech 2021 (joint winners).

Projects

- **I-vector/PLDA based speaker verification system** 2016.10–2017.05
  There is a mismatch between the distribution of the i-vector and it's Gaussian hypothesis in the back-end Probabilistic Linear Discriminant Analysis (PLDA) while doing the speaker verification. In this situation, Kullback-Leibler divisive is adopted for regularizing the distribution of the i-vector for minimizing the distribution distance between the extracted i-vector and the Gaussian hypothesis. After regularization, the verification accuracy on the NIST SRE 2008 dataset was improved by 2% with the i-vector/PLDA system. The work was accepted by IEEE Global SIP 2017.
- **Composition score and ancient poetry recommendation** 2017.08–2017.10
  Existing essay scoring systems generally only aim to score essays but cannot provide polishing that consider the content of the essay. To help students polish their essays, a deep learning model based on the attention mechanism was adopted to extract the topic of the essay and match it with a pre-stored library of ancient poems and quotes to make the recommendations. This project was carried out at the Iflytek, Beijing Research Institute.
- **Linguistic based dementia detection** 2018.10–2019.04
  According to research, the speech of people living with Alzheimer's Disease has deficits at both the word and the sentence level. Inspired by this phenomenon, we proposed an end-to-end hierarchical attention structure for extracting the linguistic information from transcripts at both the word and the sentence level, which is similar to the analysis that the clinician would perform. The proposed method achieved the state of the art performance on a publicly available dataset: DementiaBank. The work was accepted by Interspeech 2019.
- **Acoustic based dementia detection** 2019.03–2021.02
  - Speech collected from recordings with a noisy background environment and unclear pronunciations caused by Alzheimer's Disease decreases the Automatic Speech Recognition (ASR) System's performance. This encouraged us to explore a method for extracting high-quality features from the audio recordings. To this end, the time alignment information and confidence scores estimated by the ASR system were used for rhythm feature designation and high-quality feature extraction. The work was accepted by Interspeech 2020.
  - The raw waveform, which embeds more information than the extracted acoustic features, can be inputted into the end-to-end system for Alzheimer's Disease detection. However, the end-to-end system is un-interpretable by non-specialists, which is vital for medical applications. SincNet is used as the first layer of the end-to-end acoustic feature extractor to create an interpretable system. Our method is the first end-to-end structure used in acoustic-based Alzheimer's Disease detection. The work was accepted by Interspeech 2020.
  - A set of classification models that integrate the two tasks as mentioned above were designed. The research results will be submitted to IEEE/ACM Transactions on Audio Speech and Language.
- **Speech based dementia tracking** 2020.06–2020.10
  The properties of speech of an individual may change over time due to ageing, but changes may also occur as a consequence of cognitive decline. In response to this phenomenon, a theoretical analysis of the impact of age and illness on the patient' s voice revealed the relationship between age and cognitive status. Multi-task regression models were proposed. The work was accepted by ICASSP 2021. This project was carried out at the Philips Research Center in the Netherlands.
- **Speech based acoustic and linguistic feature extraction for dementia detection** 2021.01–2021.04
  Acoustic features can be extracted directly from the audio recordings. However, linguistic features in fully automatic systems need to be extracted from transcripts generated by an ASR system. We explore two state-of-the-art ASR

paradigms, Wav2vec2.0 (for transcription and feature extraction) and time-delay neural networks (TDNN), on the ADReSSo dataset containing recordings of people describing the Cookie Theft picture. The work has been accepted by Interspeech 2021.

Publications

[1] **Pan, Yilin**, T. Zheng, and C. Chen, "I-vector kullback-leibler divisive normalization for plda speaker verification", in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2017, pp. 56–60.

[2] **Pan, Yilin**, B. Mirheidari, M. Reuber, A. Venneri, D. Blackburn, and H. Christensen, "Automatic hierarchical attention neural network for detecting AD.", in *INTERSPEECH*, 2019, pp. 4105–4109.

[3] **Pan, Yilin**, B. Mirheidari, M. Reuber, A. Venneri, D. Blackburn, and H. Christensen, "Improving detection of Alzheimer' s Disease using automatic speech recognition to identify high-quality segments for more robust feature extraction", in *INTERSPEECH*, 2020, pp. 4961–4965.

[4] **Pan, Yilin**, B. Mirheidari, Z. Tu, R. O' Malley, T. Walker, A. Venneri, M. Reuber, D. Blackburn, and H. Christensen, "Acoustic feature extraction with interpretable deep neural network for neurodegenerative related disorder classification", in *INTERSPEECH*, 2020, pp. 4806–4810.

[5] **Pan, Yilin**, V. S. Nallanthighal, D. Blackburn, H. Christensen, and A. Härmä, "Multi-task estimation of age and cognitive decline from speech", in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2021, pp. 7258–7262.

[6] **Pan, Yilin***, Mirheidari, Bahman*, J. M. Harris, J. C. Thompson, M. Jones, J. S. Snowden, D. Blackburn, and H. Christensen, "Exploring using the outputs of different automatic speech recognition paradigms for acoustic- and bert-based Alzheimer' s dementia detectionthrough spontaneous speech", in *INTERSPEECH*, 2021 (*co-author).

[7] C. Chen, J. Han, and **Pan, Yilin**, "Speaker verification via estimating total variability space using probabilistic partial least squares.", in *INTERSPEECH*, 2017, pp. 1537–1541.

[8] N. Cummins, **Pan, Yilin**, Z. Ren, J. Fritsch, V. S. Nallanthighal, H. Christensen, D. Blackburn, B. W. Schuller, M. Magimai-Doss, H. Strik, *et al.*, "A comparison of acoustic and linguistics methodologies for Alzheimer' s dementia recognition", in *INTERSPEECH*, 2020, pp. 2182–2186.

[9] B. Mirheidari, **Pan, Yilin**, D. Blackburn, R. O'Malley, and H. Christensen, "Identifying cognitive impairment using sentence representation vectors", in *INTERSPEECH*, 2021.

[10] B. Mirheidari, **Pan, Yilin**, T. Walker, M. Reuber, A. Venneri, D. Blackburn, and H. Christensen, "Detecting Alzheimer's disease by estimating attention and elicitation path through the alignment of spoken picture descriptions with the picture prompt", *arXiv preprint arXiv:1910.00515*, 2019.

[11] B. Mirheidari, **Pan, Yilin**, D. Blackburn, R. O'Malley, T. Walker, A. Venneri, M. Reuber, and H. Christensen, "Data augmentation using generative networks to identify dementia", *arXiv preprint arXiv:2004.05989*, 2020.