

研究兴趣

语音处理; 病理语音检测; 声纹识别; 自然语言处理; 机器学习; 深度学习; 多模态融合

教育经历

- 谢菲尔德大学 (2018 QS 排名: 82), 哲学博士, 计算机科学与技术

2018-2022 | Sheffield, UK

  - 研究方向: 病理语音检测; 语音识别; 自然语言处理
  - 导师/课题组: Heidi Christensen/Speech and Hearing Research Group
  - 毕业论文: Linguistic- and Acoustic-based Automatic Dementia Detection using Deep Learning Methods
  - 奖学金: 欧洲玛丽居里奖学金
- 哈尔滨工业大学, 工学硕士, 计算机科学与技术

2015-2017 | 哈尔滨

  - 研究方向: 声纹识别, 语音信号处理
  - 导师/课题组: 郑铁然/语音处理实验室
  - 毕业论文: 基于 i-vector 特征规整的概率线性判别分析说话人确认方法研究
  - 奖学金: 哈尔滨工业大学一等奖学金 (两次); 国家一等助学金 (两次)
- 东北大学, 工学学士, 计算机科学与技术

2011-2015 | 沈阳

  - 研究方向: 行人检测; 信号处理
  - 导师/课题组: 赵海/嵌入式技术实验室
  - 毕业论文: 基于 HOG 和改进 ViBe 的多行人检测方法
  - 奖学金: 校级二等奖学金 (两次); 校级三等奖学金 (三次)

竞赛经历

- Alzheimer’s Dementia Recognition through Spontaneous Speech

2020.06-2020.10

主办单位: Interspeech2020 国际会议

竞赛简介: 第一届且唯一利用病理语音诊断阿尔兹海默症的国际竞赛, 包含病理语音和文本相关的分类和回归问题。

我的工作: 利用本人提出的层级注意力机制的端到端模型, 实现文本特征的提取, 语音和文本的特征融合工作。

论文成果: 排名: 团队全球第七名; 该工作被 Interspeech2020 收录。
- Alzheimer’s Dementia Recognition through Spontaneous Speech Only

2021.02-2021.04

主办单位: Interspeech2021 国际会议

竞赛简介: 第二届利用语音诊断阿尔兹海默症的国际竞赛, 包含病理语音相关的分类、回归和病情跟踪的问题。

我的工作: 利用近一年的研究成果进行语音到文本的转录、语音和文本的特征提取及特征融合。

论文成果: 排名: 团队全球第一名, 该工作被 Interspeech2021 收录。

## 工作经验

大连海事大学, 讲师, 多模态融合; 病理语音检测

2022.10–now | 大连

- 研究方向: 基于语音和文本的阿尔兹海默症病情检测与病情跟踪

飞利浦研究中心, 实习研究员, 病理语音检测

2020.06–2020.10 | Eindhoven, Netherlands

- 研究方向: 基于语音的阿尔兹海默症病情跟踪检测

科大讯飞北京研究院, 初级研究员, 自然语言处理

2017.08–2017.10 | 北京

- 研究方向: 基于深度学习的作文评分和古诗词推荐

## 研究课题

### • 作文评分和古诗词推荐

2017.08–2017.10

现存作文评分系统一般只就文章进行评分, 不能就具体内容提供有针对性的润色建议。针对这一缺陷, 利用基于注意力机制的深度学习模型对文章进行主题提取, 并与预存的古诗词名言库进行匹配, 针对性地给用户推荐用于文章润色的古诗词和名人名言。此项目开展于科大讯飞北京研究院。

### • 基于文本语义的阿尔兹海默症检测

2018.10–2019.04

研究发现, 阿尔兹海默症患者在语言表达时, 在词语层级和句子层级会有不同程度的退化。针对这一现象, 提出了用于识别阿尔兹海默症的基于层级注意力机制的端到端模型, 对语音的转录文本进行分层级地特征表示, 以提高分类效果。这是基于层级的注意力机制方法首次被应用在该领域。该方法在领域内最大的公开数据集 DementiaBank 上获得了当年最优效果。研究成果被 Interspeech 2019 (语音领域顶级会议) 收录并以 Poster 形式在大会上做报告。

### • 基于语音的阿尔兹海默症检测

2019.03–2021.02

- 嘈杂的语音采集环境和模糊的患者语音将影响语音特征提取和自动语音识别 (ASR) 的效果, 这在公开数据集上尤为明显。为减少该现象的影响, 利用 ASR 系统输出的时间对齐信息和置信度评分信息, 设计了针对患者语音的韵律特征, 提高了提取的语音特征和语义特征的质量。这是首次将 ASR 输出的信息作为语音和语义特征质量的衡量标准。该方法在领域内最大的公开数据集 DementiaBank 上获得了当年最优效果。研究成果被 Interspeech 2020 收录并以 Online Presentation 形式在大会上作报告。
- 面对高质量的语音数据集, 虽然可以使用比语音特征包含更多信息的语音波形 (raw wave) 作为端到端模型的输入, 但模型不具有可解释性。针对这一现象, 利用 SincNet 设计了具有可解释性的基于语音的端到端语音特征提取模型。这是首次将基于语音波形的端到端的方法应用于阿尔兹海默症的识别领域。该方法在私有数据集上获得了优于以往特征的分类效果。研究成果被 Interspeech 2020 收录并以 Online Presentation 形式在大会上作报告。

### • 海岸电台自动报警识别

2023.06–now

基于人工智能的海上无人通信主要通过设计智能语音识别系统来构建进行自动指令识别, 用于提高海岸电台通信的可靠性, 对海上遇险指令进行及时识别和反馈, 提高船舶行驶和停靠的安全性。此任务主要包括以下三个技术任务: 1. 海上噪声比较强, 主要包括风声、海浪声、船舶行驶噪声及无线电传播噪音, 这类噪声会影响系统的识别效果。因此, 对于海上采集的语音, 需要对其进行降噪处理, 以提高系统的识别准确率; 2. 识别任务主要针对船舶特有指令词, 这与预训练语音识别模型所基于的语料库差别较大, 直接应用预训练模型应用于海上无人通信任务很难达到预期的效果。因此, 需要对预训练语音模型进行自适应训练, 以提高模型的识别准确率。

- [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, 56–60. (EI 检索).
- [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, 4105–4109. (CCF C, 语音顶会).
- [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, 4961–4965. (CCF C, 语音顶会).
- [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, 4806–4810. (CCF C, 语音顶会).
- [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, 7258–7262. (CCF B, 信号处理顶会).
- [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 detection through spontaneous speech”, in *INTERSPEECH*, 2021. (\* co-author; CCF C, 语音顶会).
- [7] **Pan, Yilin**, M. Lu, Y. Shi, and H. Zhang, “A path signature approach for speech-based dementia detection”, *IEEE Signal Processing Letters*, 2023. (CCF C, 语音高水平期刊).
- [8] **Pan, Yilin**, B. Mirheidari, D. Blackburn, and H. Christensen, “Tsac-att: A self-supervised speech-based dementia detection system”, *IEEE/ACM Transactions on Audio, Speech and Language Processing*, Major Correction, (CCF B, 语音顶刊).
- [9] **Pan, Yilin**, Y. Shi, Y. Zhang, and M. Lu, “Extracting the hierarchical information with swin transformer and bert-based system for speech-based dementia detection”, *IEEE Signal Processing Letters*, Under Review, (CCF C, 语音高水平期刊).
- [10] C. Chen, J. Han, and **Pan, Yilin**, “Speaker verification via estimating total variability space using probabilistic partial least squares.”, in *INTERSPEECH*, 2017, 1537–1541. (CCF C, 语音顶会).
- [11] 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.

- [12] 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.
- [13] 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, 2182–2186. (CCF C, 语音顶会).
- [14] B. Mirheidari, **Pan, Yilin**, D. Blackburn, R. O'Malley, and H. Christensen, "Identifying cognitive impairment using sentence representation vectors", in *INTERSPEECH*, 2021. (CCF C, 语音顶会).

<https://scholar.google.com/citations?user=Yrzc2wgAAAAJ&hl=zh-CN>