

代东洋

✉ accum.dai@gmail.com · ☎ (+86) 188-1787-0965

🎓 教育背景

清华大学, 计算机系 - 计算机技术

硕士, 2020 年毕业

同济大学, 软件学院 - 软件工程

学士, 2017 年毕业

👨‍💻 工作经历

字节跳动 – AI Lab | 智能创作实验室

2020 年 1 月 – Now

音色克隆 (Voice Cloning)

- 提出 Hieratron 声音克隆框架 (Text -> ASR bottleneck -> Mel), 使音色和其他语音特征解耦¹。以此为基础, 实现:
 - 极少量数据音色克隆 (Few-shot voice cloning)
 - 跨风格音色克隆 (Cross style voice cloning)
 - 跨语言音色克隆 (Cross language voice cloning)
- 在多说话人的 Tacotron 框架下, 实现噪声鲁棒的个性化语音合成², 情感可控的个性化语音合成。

腾讯 – AI Lab (Intern)

2019 年 1 月 – 2019 年 8 月

多音字发音预测、语音合成等工作

- 使用预训练的 BERT 模型提取多音字的语义特征, 提高多音字发音预测模型的性能⁴; 基于决策树的多音字发音预测模型, 上线使用; 基于 Tacotron-2 模型的情感语音合成、多说话人语音合成

👨‍🔬 科研经历

清华大学深圳研究生院 – 人机语音交互实验室

2017 年 9 月 – 2020 年 6 月

课题: 基于神经网络的语音中的情感特征表示学习

- 语音情感识别模型中引入 Center Loss, 从变长语谱图中提取更有效的情感特征⁵
- 使用 MMD、对抗学习等方法进行跨域语音情感识别
- 基于跨域语音情感识别模型, 在无情感标签数据 (BC2013) 上进行情感可控的语音合成³

📄 研究成果

1. Cloning One's Voice Using Very Limited Data In The Wild (arxiv) **D. Dai**, Y. Chen, L. Chen, M. Tu, L. Liu, R. Xia, Q. Tian, Y. Wang and Y. Wang

2. Noise Robust TTS for Low Resource Speakers using Pre-trained Model and Speech Enhancement (arxiv) **D. Dai***, L. Chen*, Y. Wang, M. Wang, R. Xia, X. Song, Z. Wu and Y. Wang

3. Emotion Controllable Speech Synthesis using Emotion-unlabeled Dataset with the Assistance of Cross-domain Speech Emotion Recognition (ICASSP 2021) X. Cai*, **D. Dai***, Z. Wu, X. Li, J. Li and H. Meng

4. Disambiguation of Chinese Polyphones in an End-to-End Framework with Semantic Features Extracted by Pre-trained BERT (INTERSPEECH 2019) **D. Dai**, Z. Wu, S. Kang, X. Wu, J. Jia, D. Su, D. Yu and H. Meng

5. Learning Discriminative Features From Spectrograms Using Center Loss For Speech Emotion Recognition (ICASSP 2019) **D. Dai**, Z. Wu, R. Li, X. Wu, J. Jia and H. Meng

📌 其他

- 荣誉与奖项: 清华大学计算机系优秀硕士毕业生
- 了解机器学习基本知识、推荐系统主要模型和流程
- Google Scholar 主页: <https://scholar.google.com/citations?hl=en&user=nhzY1rsAAAAJ>