

语音抑郁检测中的说话人泄露问题：一项控制研究

来源: arXiv

日期: 2026-04-15

DOI: -

链接: <https://arxiv.org/abs/2604.14354v1>

【中文标题】

语音抑郁检测中的说话人泄露问题：一项控制研究

【研究背景】

抑郁症的早期识别对于及时治疗至关重要。语音分析作为一种非侵入性方法，在抑郁症的早期诊断中显示出潜力。然而，基于说话人身份的识别可能成为混淆因素。

【研究方法】

本研究使用DAIC-WOZ数据集，提出了一种数据分割策略，在保持训练集大小不变的同时控制训练集和测试集之间的说话人重叠。实验评估了模型在说话人重叠和未见说话人上的性能。

【主要发现】

研究发现，说话人重叠显著提高了模型性能，但在未见过的说话人上准确率急剧下降。即使在采用域对抗神经网络的情况下，模型性能仍然受到说话人身份纠缠的影响，传统的评估协议可能因此高估了泛化能力和临床应用价值。

【临床意义】

这些发现强调了严格的无说话人依赖性评估的必要性，以更准确地评估基于语音的抑郁检测模型的临床应用潜力。