

FlashSpeech and Diffusion-Based Models for Robust Speaker Verification Under Acoustic Distortions

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 8 peer-reviewed papers addressing the following research question: What is the comparative robustness of FlashSpeech versus diffusion-based baselines for speaker verification when tested on noisy subsets of the VoxCeleb benchmark with simulated real-world acoustic. 0 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 1.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: A Framework for Robust Speaker Verification in Highly Noisy Environments Leveraging Both Noisy and Enhanced Audio. Research question: What is the comparative robustness of FlashSpeech versus diffusion-based baselines for speaker verification when tested on noisy subsets of the VoxCeleb benchmark with simulated real-world acoustic distortions?.

2 Methodology

Systematic literature search across multiple databases yielded 8 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 1.8/10.

3 Results

8 papers retrieved. 0 claims extracted; 0 independently verified. Quality review score: 1.8/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

References

- <http://arxiv.org/abs/1705.03670v1>
- <http://arxiv.org/abs/2508.18913v1>
- <http://arxiv.org/abs/2109.02052v3>