A Robust Accent Classification System Based on Variational Mode Decomposition

Darshana Subhash, Jyothish Lal G., Premjith B., Vinayakumar Ravi

Presented by: Harishankar Nagar   Vivek Khari

Introduction

- Speech recognition is widely used in real-world applications.
- Accents pose a significant challenge for ASR systems.
- Robust accent classification models improve ASR accuracy.

State-of-the-Art Methods

- **Variational Mode Decomposition (VMD) + MFCC:** Enhances speech signal clarity before feature extraction.
- **Support Vector Machines (SVM):** Effective for structured feature-based classification.
- **Convolutional Neural Networks (CNNs):** Captures spatial patterns from spectrogram representations.
- **Hybrid Models (CNN-LSTM):** Leverages both spatial and temporal features.

Evaluation Metrics

- **Accuracy:** Measures correctly classified instances.
- **Precision and Recall:** Evaluate model reliability across different accents.
- **F1-Score:** Balances precision and recall.
- **Confusion Matrix:** Analyzes model misclassifications.

Comparison of Methods

Method	Accuracy	Precision	Recall	F1-Score
MFCC + MPSA-DenseNet	72.9%	72.4%	72.7%	72.8%
MFCC + 1D CNN-BiGRU-Attention	80.5%	79.8%	78.5%	79.0%
MFCC + SVM	92.5%	91.6%	92.0%	91.8%
MFCC + CNN + LSTM	98.63%	97.0%	96.5%	96.7%
VMD + MFCC + SVM (Proposed)	99.3%	99.1%	99.5%	99.2%

Table: Performance comparison of accent classification models.

Conclusion and Future Directions

- VMD + MFCC significantly improves accent classification.
- Future work includes:
  - Enhancing model generalization to unseen accents.
  - Developing lightweight models for real-time applications.
  - Addressing ethical concerns in accent bias.