

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

- ▶ **Tianyu Song** introduced Multi-task- Densely Convolutional Networks (Multi-DenseNet), Pyramid Split Attention- Densely Convolutional Networks (PSA-DenseNet), and Multi-task Pyramid SplitmAttention- Densely Convolutional Networks (MPSA-DenseNet) for English accent classification.
- ▶ **Das and Bhattacharjee** proposed a method to identify Assamese dialects using acoustic and prosodic features. A vowel speech corpus from four dialect regions is analyzed with classifiers such as Random Forest, Extreme Random Forest (ERF), and Extreme Gradient Boosting (XGB). They introduced two features— Filter Bank (FBank) and MFCC and developed a 1D CNN-BiGRU-Attention model.

State-of-the-Art Methods

- ▶ **Humayun** developed an acoustic accent profiling model that uses convolutional networks with global average pooling to analyze sound features.
- ▶ **Ariadne** evaluated accent classification for spontaneous speech using CNN followed by Long short-term memory (LSTM) networks and the Wav2vec2 model.

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.