

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Towards Self-Supervised Learning based Acoustic Modeling for Non-Native Mispronunciation Verification
著者(和文)	YANGLongfei
Author(English)	Longfei Yang
出典(和文)	学位:博士(工学), 学位授与機関:東京科学大学, 報告番号:甲第284号, 授与年月日:2025年3月26日, 学位の種別:課程博士, 審査員:篠崎 隆宏,奥村 学,中山 実,船越 孝太郎,長谷川 晶一
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Institute of Science Tokyo, Report number:甲第284号, Conferred date:2025/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	Longfei Yang		
論文審査 審査員		氏名	職名		氏名	職名
	主査	篠崎 隆宏	教授	審査員	長谷川 晶一	准教授
	審査員	奥村 学	教授			
		中山 実	教授			
船越 孝太郎		准教授				

論文審査の要旨 (2000 字程度)

本論文は「Towards Self-Supervised Learning based Acoustic Modeling for Non-Native Mispronunciation Verification」と題し、英文全5章から構成されている。

第1章「Introduction (序論)」では、本研究の背景と目的を述べている。発音誤り検証 (Mispronunciation Verification, MV) は、言語学習支援システム (CAPT) において重要な技術であり、学習者の発話の誤りを自動的に検出し、適切なフィードバックを提供することを目的としている。従来のMVシステムは、自動音声認識 (ASR) を基盤としているが、非母語話者の音声データの不足やアノテーション作業の負担が課題となっていることを指摘している。本研究では、自己教師あり学習を活用し、大規模な未ラベル音声データを用いた音響モデルを構築することで、これらの課題の克服を目指している。

第2章「Language Representation Learning for Mispronunciation Verification (発音誤り検証のための言語表現学習)」では、発音誤り検証における言語表現学習の手法について述べている。本研究では、自己教師あり学習を用いて大規模な中国語および日本語の音声データから音響表現を学習し、それを非母語話者の音響モデルへ応用する枠組みを提案している。特に、言語逆学習 (Language Adversarial Training, LAT) を導入し、異なる言語間の特徴分布を整合させることで、非母語話者の発音誤り検証に有用な表現を獲得できることを示している。実験の結果、提案手法が従来の非母語話者音声のみに基づいたモデルと比較して、音素認識および発音誤り検証の精度向上に寄与することを報告している。

第3章「Formant Augmented Language Adversarial Representation Learning for Non-Native Acoustic Modeling of Mispronunciation Verification (発音誤り検証のためのフォルマント拡張型敵対的表現学習)」では、非母語話者の発音誤り検証におけるフォルマント情報を活用した敵対的表現学習手法について述べている。本研究では、先行研究で提案した自己教師あり事前学習フレームワークを拡張し、音響モデルの入力層に従来の畳み込み層の代わりに sinc フィルターを導入することで、フォルマントに類似した特徴を抽出する手法を提案している。このアプローチにより、発音誤りの識別に重要な調音位置や調音方法に関する情報を効果的に捉えられることを指摘している。実験結果に基づき、提案手法が非母語話者の音素認識および発音誤り検証の精度向上に寄与し、特に精度の向上に有効であることを報告している。

第4章「Self-Supervised Learning with Multi-Target Contrastive Coding for Non-Native Acoustic Modeling of Mispronunciation Verification (発音誤り検証のためのマルチターゲット対照コーディングを用いた自己教師あり学習)」では、非母語話者の発音誤り検証に向けた自己教師あり学習の新たな手法について述べている。近年の深層学習技術の発展により、発音誤り検証の精度は向上しているが、大規模なラベル付きデータの収集とアノテーション作業が課題となっていることを指摘している。本研究では、自己教師あり学習に基づくマルチターゲット対照コーディング (multi-target contrastive coding) を導入し、異なる言語や話者における音韻構造の相違を捉えるモデルを構築している。さらに、入力音声の再構成を行う補助項を導入し、より抽象度の高い音響表現の獲得を試みている。実験の結果、日本語話者の中国語学習音声データを用いた検証において、本手法が非母語話者の音素認識および発音誤り検証の性能向上に寄与することを報告している。

第5章「Conclusions and Future Works (結論および今後の課題)」では、本研究の結論と今後の課題について述べている。本研究では、自己教師あり学習を活用し、非母語話者の発音誤り検証の精度向上を図る手法を提案し、その有効性を実験により確認している。今後の課題として、より高精度な音響モデリング、多様な母語背景の学習者への適用、モデルの解釈性向上、実環境におけるノイズ耐性の強化を挙げている。これらの課題に取り組むことで、発音学習支援システムの実用性向上が期待されることを指摘している。

以上を要するに、本論文は、自己教師あり学習を活用し、データ不足の課題を克服することで、非母語話者の発音誤り検証の精度向上を実現するものである。提案手法の有効性は実験により確認され、発音学習支援システムの高度化に貢献する工学的価値が大きい。よって博士 (工学) の学位を授与するに十分な価値を持つものと認められる。

注意: 「論文審査の要旨及び審査員」は、東京科学大学リサーチリポジトリ (T2R2) にてインターネット公表されますので、公表可能な範囲の内容

で作成してください。