

2023 年度年次報告書

社会課題を解決する人間中心インタラクションの創出

2023 年度採択研究代表者

高島 遼一

神戸大学 都市安全研究センター

准教授

障害者のための省学習データ音声・音響認識技術の開発

## 研究成果の概要

第一年次は構音障害者音声認識のタスクを中心に研究を進めた。音声認識モデルを学習するためには、音声と発話内容(テキスト)のペアデータが必要である。従来は、事前に用意した台本を障害者の方が読み上げ、その収録音声と台本のテキストをペアデータとする方式が用いられていたが、この方法は障害者の方への負担が大きいため、音声認識モデルの学習に十分なデータを収録することが困難である、すなわち学習データ不足の問題があった。そこで本研究では、台本読み上げ音声ではなく日常会話(自由会話)音声进行収録し、それをモデル学習に活用する方法を検討した。日常会話音声は台本読み上げ音声に比べて障害者ユーザへの負担が小さいため、比較的容易に音声を収録することが可能である。一方で、聞き取りが困難な構音障害者音声から、その発話内容を書き起こすことは非常に困難なため、発話内容のラベルが得られないという新たな課題が生じる。この問題を解決するため、教師ラベルの存在しないデータをモデル学習に活用する技術である「自己教師あり学習」を障害者音声認識に適用することを検討した。健常者音声をを用いて自己教師あり学習を行う研究は多く存在するが、構音障害者音声をを用いた研究は本研究が世界初の試みである。自己教師あり学習の方式として、Autoregressive predictive coding (APC) を用いて評価を行い、従来の台本読み上げ音声のみを用いた学習よりも音声認識率の向上を確認できた。また、ラベルの無い音声に対して音声認識を行い、その認識結果を教師ラベルの代わり用いる疑似ラベリングの方式と組み合わせることで、さらなる性能向上を確認できた。さらに、APC よりも新しい自己教師あり学習方式として wav2vec2.0 での評価も行い、その有効性を確認した。wav2vec2.0 は APC よりも高いベース認識性能が得られる一方で、モデルが巨大なため、より多くの日常会話音声が必要という課題も明らかとなったため、今後はこの課題解決にも取り組んでいく。

### 【代表的な原著論文情報】

- 1) Ryoichi Takashima, Yuya Sawa, Ryo Aihara, Tetsuya Takiguchi, and Yoshie Imai, “Dysarthric Speech Recognition Using Pseudo-Labeling, Self-Supervised Feature Learning, and a Joint Multi-Task Learning Approach,” IEEE Access, vol. 12, pp. 36990-36999, March 2024, doi: 10.1109/ACCESS.2024.3374874.