

研究課題別事後評価結果

1. 研究課題名： 言語表現の階層構造に基づく End-to-End 音声認識の研究

2. 個人研究者名

樋口 陽祐（早稲田大学大学院基幹理工学研究科 大学院生）

3. 事後評価結果

本研究では、End-to-End 音声認識において言語情報を利用することで精度を向上させる手法を開発した。音声信号から高精度で単語列を認識するためには、音声と単語との対応関係だけでなく、単語列の構造的特徴を捉える必要がある。この問題に対し、サブワード分割に基づいた階層的な音声認識手法では、音声信号に近い細かい単位のサブワードから、単語列に近い粗い単位のサブワードを段階的に認識するモデルおよびその学習方法を提案し、その有効性を示した。さらに、事前学習済み言語モデルから得られる情報を用いた音声認識手法では、BERT 等の言語モデルの出力を音声認識モデルに組み込むことで、文脈を考慮した高精度な音声認識が実現されることを示した。これらの提案手法およびその応用について、音声処理や自然言語処理のトップ国際会議に複数の論文が採択されており、国際的に高く評価されている。ニューラルネットワークにより音声認識精度は飛躍的に向上したものの、ノイズがある環境下や言い淀みなどが現れる自然な会話での音声認識はいまだ十分な精度とは言えず、提案手法をさらに発展させて実用的な音声認識技術を確立することが期待される。