

How can deep learning technique be leveraged to enhance ASR systems for dysarthric speech recognition?

Correspondence Author

Licence

Leon Starr

CC BY 4.0

New Vistas

Volume 12, Issue 1, 2026, UWL Annual Doctoral Students' Conference 2025

Abstract

Dysarthria is a neurological motor speech disorder that impacts an individual's ability to control speech production muscles, leading to reduced speech intelligibility. Therefore, individuals with dysarthria often struggle with interacting and communicating with daily life communication. The current state of Automatic Speech Recognition (ASR) for dysarthric speech is limited, especially for severe cases of dysarthria.

This research aims to expand the available data and explore novel approaches in developing robust ASR systems tailored for dysarthric speech. The data expansion phase will see the collaboration with specialist organisations and speech therapy centres to recruit volunteers with dysarthria and collect a substantial amount of dysarthric speech data. The dataset will include recordings from individuals with varying degrees of severity, covering a wide range of speech patterns and characteristics. The data will go through a preprocessing stage leveraging signal processing techniques to extract relevant speech features and enhance the quality of the recordings, ensuring the suitability of the data for ASR training.

The model development phase will use deep learning techniques such as convolutional neural networks and recurrent neural networks to create automated speech recognition models fitted for dysarthric speech. It will explore enhancing existing model architectures and creating new ones that can capture the distinctive acoustic and phonetic features of dysarthric speech.

Transfer learning will be used to make use of existing models trained on large-scale datasets of non-dysarthric speech, refining them for dysarthric speech usage. These models will be optimised for real-world application by addressing computational efficiency and real-time processing constraints. The performance of the developed ASR models will be evaluated using standard evaluation metrics such as word error rate on both dysarthric and non-dysarthric speech datasets. This research is expected to have significant impact in various domains, such as healthcare, education and accessibility.

Keywords

Dysarthria, Deep learning technique, Automatic Speech Recognition