Normal to Dysarthric Voice Conversion for Data Augmentation

• Dysarthric speakers have impaired motor skills and produce distorted speech
• Current ASR trained on normal speech can’t recognize dysarthric speech
• Need to create dysarthric speech from normal speech to increase amount of training data
• Unique problems to dysarthric speech such as no phone alignments and high variability in speech characteristics from speaker to speaker
Dysarthric to Normal Speech Conversion

• Rather than perform normal to dysarthric speech conversion for data augmentation, we might go for the other way around for the same purpose, i.e., speech recognition

• Train WaveNet to generate normal speech using dysarthric auxiliary features (mcep, vuv, F0, aperiodicity, etc.)

• In testing time, generate waveform using dysarthric auxiliary features only

• May need special processing to account for high variability of dysarthric speech