Auditory Attention Decoding Approach to Cocktail Party Problem Using Deep Learning

Poondi, Govardhan (School: Plano West Senior High School)
Rho, John (School: Plano West Senior High School)

[Problem] For individuals with hearing loss, distinguishing between various voices in noisy environments is difficult. Current hearing aids are unable to identify and attenuate background voices due to software limitations and a single-microphone design. [Approach] The current audiological approach to this issue (the cocktail party problem) is auditory attention decoding, a process that harnesses the natural ability of the brain to identify and attenuate background noise. Non-invasive EEG data produced a predicted audio file through an existing state-of-the-art AAD framework, incorporating deep learning. A Deep Attractor Network was then employed to separate this file into its component voices through the use of attractor points. To compare the attended audio file and its component voices, a novel methodology implementing a Fast Fourier Transform deconstructed these files into their constituent frequencies. A correlation analysis then determined the component voice with the highest match to the attended audio file. [Testing] The degree of correlation between the component voices and the predicted audio file was measured by a Perceptual Evaluation of Speech Quality index and Mean Opinion Score index using a Tensorflow program. Both indices demonstrated correlation at ~90%, indicating excellent audio quality and comprehension from both the human and computer perspective. Thus, the integrated AAD approach developed is suitable for EEG-based speech separation and attenuation. [Applications] Since the algorithm functions in real-time and trains on unclean speech sources, it is ideal for implementation into hearing aids. The developed AAD technology also shows potential in military environments with heavy background noise and existing smart home devices.