A Physiological Self-Supervised Attention-Based Approach for Sound Source Separation in Auditory Scenes

Chen, Sophie (School: Caddo Parish Magnet High School)

Over 430 million individuals worldwide require rehabilitative technologies to address their disabling hearing loss. When untreated, hearing loss results in cognitive fatigue, strained communication and other consequences that significantly impact daily lives. Despite this, only around 20% of those in need utilize hearing aid devices due to frequent issues like overamplification and limitations in effectively distinguishing and processing sound. Computational methods for sound separation have been proposed, yet these approaches fall short in real world applications, lacking the nuanced adaptability inherent in biological auditory processes. This paper introduces a novel unsupervised-learning based approach for performing audio source separation in a physiological manner. Cochleagrams, biologically inspired representations of sound, were generated from mixed-audio signals. A recurrent neural network slot-attention based model (RNN-SA) was constructed and trained with cochleagrams to distinguish auditory stimuli within multisound environments. Upon evaluation, the model demonstrates promising abilities in accurately performing sound source separation, achieving high signal-to-distortion and scale-invariant signal-to-noise ratios of 6.43 and 14.35, respectively. Integration of this approach in hearing aid technologies could significantly enhance performance of such devices. Further, this modeling of auditory attention mechanisms through simulating sound source separation offers greater insight into current understanding of mid-level processing stages and auditory cognition.

Awards Won: