1505: CCI-MOBILE: COMPARATIVE ANALYSIS OF CNN-BASED MODELS VS HUMAN SOUND RECOGNITION AMONG COCHLEAR IMPLANT AND NORMAL HEARING SUBJECTS

Ram Charan Mudha Chandra Shekar, John H.L. Hansen
Department of Electrical Engineering, The University of Texas at Dallas, Richardson, TX, USA
School of Brain and Behavioral Sciences, The University of Texas at Dallas, Richardson, TX, USA

It has been previously shown that advantages in auditory processing exist when the following situational context traits or subject/system properties are present: (i) availability of a wider radial range up to 360 degrees, (ii) intolerance towards acoustic or visual obstruction, (iii) distant event horizon, (iv) availability for quick neural processing, (v) association of human attention and emotion, and (vi) activity during sleep. Sound Recognition (SR) is important in creating awareness regarding imminent environment danger and its role becomes even more pronounced in ensuring safety, enabling autonomy, and day-to-day ease among hearing-impaired individuals. Cochlear Implants (CIs) have been widely used as a solution to restore auditory function in hearing impaired individuals. SR among CI users is also an important measure used in the assessment in hearing-related quality of life. However to date, environmental SR among CIs has received very little attention and relatively few studies have investigated both assessment and impact relative to normal hearing (NH) subjects. These studies have reported large variations in SR performance and although, many of the findings from these studies are relatively inconsistent in methodology, experimental factors, and evaluation, the results broadly seem to suggest that there is a clear deficit in the existing CI sound processing to effectively process environmental sounds. In this study, a comparative analysis of NH and CI listeners is carried out to determine SR using classifiers trained on learned sound representations from a CNN-based sound event model. Audio files from the ESC-50 database were used as the sound battery to evaluate SR, where NH listeners were provided simulated CI listening conditions. Stimuli was provided via the CCi-MOBILE Research Platform to CI listeners and the Braeker Vocoder was used to auralize electric stimuli generated by CCi-MOBILE and synthesize the listening experience among CI subjects. Natural and auralized audio from ESC-50 were used to extract NH and simulated CI sound representations from a pre-trained CNN. A comparative analysis was performed to investigate the effect of machine models and subjective performance on SR performance. Metrics such as classification accuracy, F1-score, t-SNE based sound confusability feature-space analysis and others were used to perform a comparison of SR machine models for NH and CI listeners. It is suggested that findings from this study could be used to develop novel sound processing algorithms, identify optimal CI electrical stimulation characteristics for enhanced sound perception, and other key performance markers and characteristics necessary for advancing CI based environmental sound recognition (SR).

[1] Shekar, R.C., Belitz, C., & Hansen, J.H.L. (2021). Development of CNN-Based Cochlear Implant and Normal Hearing Sound Recognition Models Using Natural and Auralized Environmental Audio. IEEE SLT-2021: Spoken Language Technology Workshop, pp. 728-733.

This work was supported by the grant R01 DC010494-01A from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.