ARIA
APPLIED RESEARCH IN ACTION

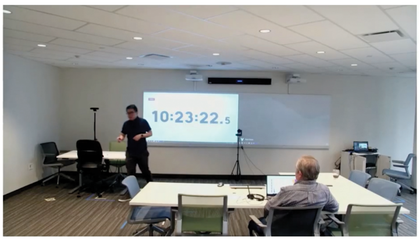
Sound Event Localization and Detection

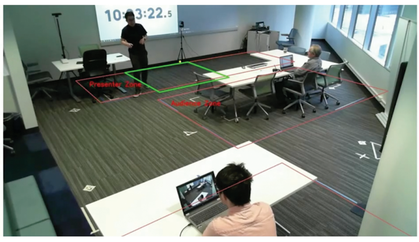
Leveraging CRNNs in SELD: Elevating Precision in Sound Source Localization by using linear arrays of microphones.

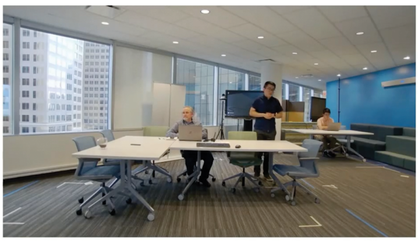
Miaoqi Zhang

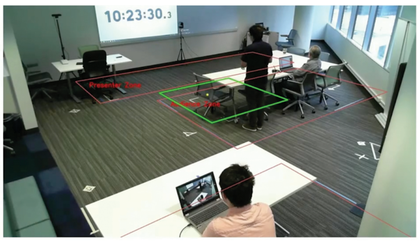
Gerald Penn
ACADEMIC SUPERVISOR

Paul Ferry
INDUSTRY SUPERVISOR









PROJECT SUMMARY

In the rapidly advancing landscape of auditory technology, the precision in detecting and localizing sound events has become a linchpin for numerous modern applications. We delved into the capacities of Convolutional Recurrent Neural Networks (CRNNs) to optimize Sound Event Localization and Detection (SELD). Our methodology began with meticulous audio data acquisition from varied environments, utilizing advanced microphone arrays. This data underwent rigorous processing with the GCC-PHAT algorithm, which facilitated the extraction of pivotal audio features.

Subsequently, these refined features were integrated into CRNNs. The SELD model we developed, backed by CRNN architecture, showcased remarkable prowess: it was adept at localizing sound sources within an error margin of just 4.3 degrees. This precision is especially crucial in applications where pinpointing sound origins is key, such as camera switching in classrooms, and audio pick up in meeting rooms.

Furthermore, we didn't stop at initial success. Our team engaged in continuous refinement processes, expanding the dataset to encapsulate a wider range of sound environments. This iterative approach ensures that our model remains versatile and adaptable, reinforcing its significance in various real-world applications. In essence, our research underscores the transformative role of CRNNs and deep learning techniques in the nuanced field of acoustic signal processing.

nureva

Computer Science
UNIVERSITY OF TORONTO
Master of Science in Applied Computing