LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

MIT

LLSC News
### Lincoln Laboratory Supercomputing Center

November 14 2023

Combining Neural Networks and Histogram Layers for Underwater Target Classification

Division 3: Air, Missile, And Maritime Defense Technology

Sound waves don't travel under water in straight lines; they reflect, diffract, and refract as they encounter boundaries and objects in their path. Their dynamic behavior complicates methods to classify targets for applications like search and rescue, port security, and ocean floor mapping. To detect targets, sea vessels largely rely on passive sonar, which records sounds with an underwater microphone. However, techniques for processing and analyzing passive sonar data often struggle to disentangle the complex patterns in target recordings.

Audio signals contain many features that represent different aspects of sound. Machine learning (ML) algorithms can extract such features to reveal patterns relevant for detecting and classifying targets. However, some features rank higher in importance than others for distinguishing between sound sources (e.g., a cargo ship vs. a tug boat). Features can interact such that signals from one source mask those of another (think of a megaphone speaker drowning out someone whispering). For audio data, many ML approaches are unable to directly model statistical features, which describe the sound intensity (amplitude) of signals. Amplitude, along with the positions of sounds across frequency and over time, can help distinguish sound. But, when trying to learn features, many approaches fail to consider local positions in the frequency-time domain or sound intensity.
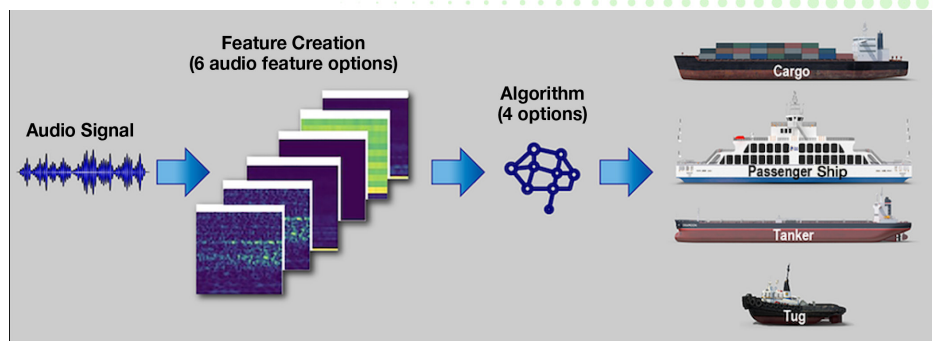
Audio signals are processed to create features for deep learning models capable of differentiating between signals from cargo, passenger, tanker, and tug ships.

To better capture statistical features within passive sonar data, a team from Lincoln Laboratory and the Advanced Vision and Learning Lab at Texas A&M University are adding local histogram layers into neural network architectures. Neural networks teach computers to process information using interconnected nodes, similar in concept to neurons of the human brain. This project employs two types of neural networks for automated feature learning that together can capture local relationships within audio signals while incorporating signal time dependencies. Histograms are bar graph-like visual representations showing how often each value in a dataset occurs—in this case, the distribution of frequencies in the audio data. Incorporated within the neural network architecture, the histogram layers directly compute how the values of audio features are locally distributed. Some vessels only make sound in certain frequency bands, so where the signal appears in the frequency-time domain can indicate what kind of vessel the signal may be coming from.

The team's histogram layer-integrated neural network architectures have demonstrated improved target classification over non–histogram layer baseline models. These demonstrations have leveraged both speech recordings and sonar data to encompass signal characteristics across different acoustic data sources. To further improve architecture performance, the team is incorporating state-of-the-art techniques such as knowledge distillation, in which information is transferred from a larger to smaller model for applications demanding higher computational efficiency. They are also implementing academia- and industry-adopted acoustic data features, many of which transform audio signals from the amplitude-time to frequency-time domain, to look for significant signal behaviors relevant to operational sonar datasets.

MIT LINCOLN LABORATORY
SUPERCOMPUTING CENTER