The Journal of the Acoustical Society of America

JUNE 04 2026

Advancing passive acoustic ship detection in the Western Canadian Arctic: Signal processing and deep learning approaches

Farid Jedari-Eyvazi ; Fabio Soares Frazao; William D. Halliday; Romina Gehrmann ; Laurelie Menelon; Jacob T. Dingwall; Chris Whidden; Michael Dowd

Check for updates

J. Acoust. Soc. Am. 159, 4843–4861 (2026)
https://doi.org/10.1121/10.0043937

View Online
Export Citation

Articles You May Be Interested In

Enhancing voice biometric security: Evaluating neural network and human capabilities in detecting cloned voices

Proc. Mtgs. Acoust. (November 2024)

A practical implementation of a lucky covariance estimator using cumulative coherence

Proc. Mtgs. Acoust. (July 2024)

Comparison of conventional and adaptive acoustic beamforming algorithms using a tetrahedral microphone array in noisy environments

Proc. Mtgs. Acoust. (October 2024)

04 August 2026 15:12:24

ASA
LEARN MORE
Advance your science and career as a member of the
Acoustical Society of America

ASA
ACOUSTICAL SOCIETY OF AMERICA

JASA ARTICLE

Advancing passive acoustic ship detection in the Western Canadian Arctic: Signal processing and deep learning approaches

Farid Jedari-Eyvazi,[1,a)] Fabio Soares Frazao,[2] William D. Halliday,[3,4] Romina Gehrmann,[5] Laurelie Menelon,[3] Jacob T. Dingwall,[4] Chris Whidden,[2] and Michael Dowd[1]

[1]*Department of Mathematics & Statistics, Dalhousie University, 6316 Coburg Road, Halifax, Nova Scotia, B3H 1Z9, Canada*

[2]*Institute for Big Data Analytics, Faculty of Computer Science, Dalhousie University, 6050 University Avenue, Halifax, Nova Scotia, B3H 1W5, Canada*

[3]*Wildlife Conservation Society Canada, 169 Titanium Way, Whitehorse, Yukon Territory, Y1A 0E9, Canada*

[4]*School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, V8P 3E6, Canada*

[5]*Defence Research and Development Canada, 9 Grove Street, Dartmouth, Nova Scotia, B3A 3C5, Canada*

ABSTRACT:

The retreat of Arctic sea ice is driving an increase in vessel traffic and associated underwater noise, which interferes with the frequency bands used by Arctic marine mammals. Detecting co-occurring vessel noise and marine mammal vocalizations in passive acoustic monitoring (PAM) data can help to assess their adverse impacts and guide mitigation strategies. This paper proposes two ship noise detection techniques: a modified variant of the Frequency Amplitude Variation (FAV) method, MFAV, which integrates signal processing with a simple statistical threshold to enhance both interpretability and detection performance; and a convolutional neural network (CNN) model specifically trained to advance ship detection in the Canadian Arctic. Comparative analysis of our PAM test dataset from the western Canadian Arctic, based on peak F1-scores, demonstrates that the CNN model generalizes well to unseen sites and, with one exception, consistently outperforms both MFAV and FAV by 1%–8%, maintaining scores above 91%. Furthermore, MFAV improves the detection of boats by up to 22% and of larger ships by 6%. The developed methods are publicly available as an open-source tool on GitHub, contributing to the advancement of acoustic vessel monitoring techniques in Canadian Arctic waters in support of conservation efforts aimed at protecting Arctic marine mammal habitats.

© *2026 Crown. All article content, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International (CC BY-NC-ND) license (https://creativecommons.org/licenses/by-nc-nd/4.0/).* https://doi.org/10.1121/10.0043937

(Received 10 December 2025; revised 8 May 2026; accepted 9 May 2026; published online 4 June 2026)

[Editor: Jianlong Li] Pages: 4843–4861

04 August 2026 15:12:24

I. INTRODUCTION

Sound plays a pivotal role in communication, navigation, echolocation, and perception of the underwater environment for marine mammals (Erbe *et al.*, 2018; Erbe *et al.*, 2019; Halliday *et al.*, 2020a). They rely on sound across a wide spectrum of frequencies, ranging from a few hertz to more than 100 kHz (Erbe *et al.*, 2018), which overlaps with the frequency range of common anthropogenic noise sources, including vessel traffic (Halliday *et al.*, 2017; Hildebrand, 2009; Wittekind and Schuster, 2016). Anthropogenic noise has been documented to interfere with marine mammal communication, impair foraging efficiency, and even cause hearing damage or mortality (Erbe *et al.*, 2019; Gomez *et al.*, 2016; Halliday *et al.*, 2017, 2020a; Jain-Schlaepfer *et al.*, 2018; McCauley *et al.*, 2017; Merchant *et al.*, 2014; Rolland *et al.*, 2012; Southall *et al.*, 2019). Given these concerns, advancing our understanding of the adverse effects of underwater anthropogenic noise remains an active area of research (Halliday *et al.*, 2017). This has also urged policymakers to advocate for monitoring underwater sound levels and establishing limits on anthropogenic underwater noise (Colbert, 2020; Halliday, 2021).

Multiple studies indicate a rising trend in underwater ambient noise levels across various regions worldwide (Andrew *et al.*, 2002; Chapman and Price, 2011; McDonald *et al.*, 2006; McDonald *et al.*, 2008; Miksis-Olds *et al.*, 2013; Miksis-Olds and Nichols, 2016; Ross, 2005), driven by both anthropogenic and natural sources (Halliday, 2021; Hildebrand, 2009; Matsumoto *et al.*, 2014; Stroeve *et al.*, 2007). Commercial shipping is the main source of underwater noise below frequencies of 500 Hz, with an increase in as much as 20 dB re. 1 μPa in low-frequency ambient noise levels since preindustrial times (Hildebrand, 2009). In the 500–25000 Hz range, in addition to natural sources, small vessels and sonar activity constitute an important anthropogenic contribution to underwater noise (Hildebrand, 2009; Stanistreet *et al.*, 2022).

[a)]Email: Farid.JE@dal.ca

 ©Crown 2026.

In recent years, the Arctic underwater soundscape has received increased attention. Between 1990 and 2015, ship traffic in the Canadian Arctic nearly tripled, with cargo vessels and government icebreakers (including research ships) accounting for the highest traffic, while pleasure craft also experienced rapid growth (Dawson *et al.*, 2018). Recent findings indicate that baseline ambient noise levels in Canadian Arctic waters typically range from 70 to 127 dB re. 1 μPa within the 50–1000 Hz frequency band, largely influenced by wind speed and sea ice concentration, with a median value of approximately 91 dB re. 1 μPa (Halliday, 2021). As the loss of sea ice continues, it is expected to open new shipping routes across the Arctic Ocean, further increasing vessel traffic and introducing substantial acoustic disturbances into Arctic ecosystems (Dawson *et al.*, 2018; Halliday *et al.*, 2017; Halliday *et al.*, 2018; Halliday, 2021; Arctic Council, 2019; Pizzolato *et al.*, 2016; Stephenson *et al.*, 2011; Stroeve and Notz, 2018). Reliable vessel monitoring techniques are necessary for assessing the long-term cumulative impact of anthropogenic noise on Arctic marine mammals and informing effective conservation and management strategies (Erbe *et al.*, 2018; Halliday, 2021).

Passive acoustic monitoring (PAM) technology has emerged as a key approach for monitoring vessel traffic and evaluating its ecological impacts, such as generating noise in the sound frequencies used by Arctic whale species (Pollara *et al.*, 2017). PAM systems use hydrophones to measure sound produced by vessels, including noise from engines, propeller blades, and hydrodynamic flows (Hildebrand, 2009; Pollara *et al.*, 2017; Smith and Rigby, 2022). Vessel-generated sounds consist of both broadband and narrow-band signals with dominant frequencies typically below 1 kHz. However, sounds from small boats in close proximity to a receiver can be detected at frequencies exceeding 5 kHz (Hildebrand, 2009).

The broadband signals primarily originate from propeller cavitation, modulated at the blade pass rate and fundamental frequency (Kudryavtsev *et al.*, 2003; Pollara *et al.*, 2017). This modulation is exploited by the detection of envelope modulation on noise (DEMON) algorithm to detect vessel presence (McDonald *et al.*, 1963; Slamnoiu *et al.*, 2016). DEMON methods extract the amplitude envelope of the raw acoustic waveform and apply a Fourier transform to reveal periodic components corresponding to the propeller blade and shaft rotation rates (Pollara *et al.*, 2016). However, they have notable limitations, including their reliance on manually selected bandpass filters, which are required to target specific vessel types (Hanson *et al.*, 2008). This reliance, in particular, hinders the automation of DEMON-based approaches and makes them impractical for application to large-scale PAM datasets. Additionally, they have a restricted detection range due to their dependence on rapidly attenuating high-frequency components of broadband signals (Chung *et al.*, 2011; Pollara *et al.*, 2017; Reis *et al.*, 2019). To improve the effectiveness of DEMON-based vessel noise detection methods, a variety of signal processing techniques have been proposed to enhance the vessel noise signature (Averbuch *et al.*, 2011; Hanson *et al.*, 2008; Leal *et al.*, 2015; Li *et al.*, 2024). However, these approaches can be less effective in environments where broadband noise from biological sources mimics vessel emissions, leading to potential misclassification (Hanson *et al.*, 2008; Hildebrand, 2009).

A different class of vessel detection techniques is based on the low-frequency analysis and recording (LOFAR) approach, focusing on frequency components typically below 1 kHz (Urick, 1983). These methods target the distinct narrow-band components of vessel noise (i.e., tonal noise) in spectrograms, typically produced by rotational machinery and reciprocating engine parts (Hildebrand, 2009). An automated approach for identifying vessel tonal noise was proposed by (Reis *et al.*, 2019), namely, the frequency amplitude variation (FAV) technique (see Sec. III for a detailed description). Compared to DEMON-based methods, FAV operates with lower computational cost, and once it is fine-tuned, it requires no calibration or human intervention, enabling automatic detection in underwater acoustic recordings (Reis *et al.*, 2019). This method has demonstrated superior performance under noisy conditions compared to DEMON-based methods (Reis *et al.*, 2019). The FAV method may, however, struggle to distinguish the tonal noise produced by small-engine boats from fish choruses, as both can exhibit similar frequency peaks in spectrograms (Reis *et al.*, 2019).

Extracted acoustic features from DEMON and LOFAR are also commonly used as input to various machine learning (ML) classifiers. Among the most successful classical ML techniques is the support vector machine (SVM) method and its variants (Hummel *et al.*, 2024; Prasad and Gurugopinath, 2022; Vieira *et al.*, 2019; Yan *et al.*, 2017). However, the manual selection of features necessary to identify signals of interest requires substantial expertise in signal processing and may inadequately represent the intricate and variable time-frequency signature of vessel acoustic signals (Jiang *et al.*, 2019; White *et al.*, 2022). In addition, processing and analyzing large-scale PAM datasets is a highly demanding task. To manage these complexities, recent research increasingly leverages deep learning (DL) methods. Adoption of DL methods has grown steadily, driven by their robustness to noise and ability to automatically learn discriminative features from raw or minimally processed audio, consistently outperforming traditional signal processing and classical ML techniques (Frazao *et al.*, 2025; Hummel *et al.*, 2024; Ji *et al.*, 2024; Topple *et al.*, 2024; White *et al.*, 2022; Yang and Chen, 2017; Zak, 2008).

In this paper, we address the growing need to monitor the impacts of anthropogenic vessel noise that overlaps with the frequency range of Arctic marine mammals' calls by focusing on the development and application of automated ship-noise detection techniques. We propose a modified variant of the FAV method, referred to as MFAV. The proposed approach provides a solution for detecting small boats, which are typically difficult to automatically identify due to their highly variable noise characteristics and lack of automatic identification system (AIS) transmission. Additionally, we trained a deep convolutional neural network (CNN) ship detector using a predefined Curriculum Learning strategy (Wang *et al.*, 2022)

04 August 2026 15:12:24

https://doi.org/10.1121/10.0043937

informed by the MFAV results. To the best of our knowledge, there is currently no publicly available open-source DL model for ship detection using PAM data that is tailored to the Western Canadian Arctic, largely due to the region's remoteness and the limited availability of annotated datasets. The performance of the trained CNN model and MFAV method was evaluated against the original FAV method using seven representative acoustic datasets collected from six sites in the western Canadian Arctic. We further evaluated the performance of the developed ship noise detection methods for identifying small boat noise, which may exhibit different acoustic signatures than larger ships, and proposed a set of best practices to support effective vessel noise detection in Arctic waters. Due to increasing shipping activity in the Canadian Arctic, where vessel noise overlaps with frequencies used by resident marine mammals, we have released the original FAV, the MFAV, and the trained CNN model implementation as an open-source ship-noise detection tool, publicly available on GitHub to support broader adoption within the Arctic PAM community.

In the remainder of this paper, we first describe the PAM datasets used in this study in Sec. II. We then elaborate on the ship-noise detection methodologies that we have developed and used in this work in Sec. III. Next, in Sec. IV, we evaluate and compare the performance of all methods across the available test datasets under varying environmental and operational conditions, and provide justification for the observed results. Finally, we provide an interpretation of the results, discuss the contribution of this work, and provide the best practices for applying the detectors in Sec. V. In Sec. VI, we conclude the key outcomes of this study.

II. DATA

The underwater acoustic datasets used to develop and evaluate ship noise detection methods in this study were collected from seven hydrophone deployments across six sites in the western Canadian Arctic, including Ulukhaktok (Halliday *et al.*, 2020b), Pearce Point, two locations at Cape Bathurst (CB) (Dingwall *et al.*, 2024), two deployments at Cape Parry (CP) during 2022–2023 and 2023–2024, and one deployment in the Prince Albert Sound waters (PAS) (Fig. 1; Table I). Together, these datasets capture a diverse acoustic environment in the western Canadian Arctic, spanning multiple years, sites, hydrophone models, depths, seabed and shoreline morphologies, and distances from shipping lanes. This breadth enables a comprehensive evaluation of ship noise detection methods under varying noise conditions and sound propagation characteristics relevant to the region.

Table I summarizes the deployment information. The low-noise HTI 92-WB hydrophone (High Tech Inc., Gulfport, MS), deployed at the Ulukhaktok site, had a relatively flat sensitivity of –155 dB re. 1 V/μPa between 200 Hz and 6 kHz, with a steep decline below 50 Hz. The recordings also exhibited recorder self-noise at frequency channels of 75, 125, 131, 132, 220, 375, 440, 660, 750, and 1124 Hz, which resembled the ship tonal noise and thus were replaced

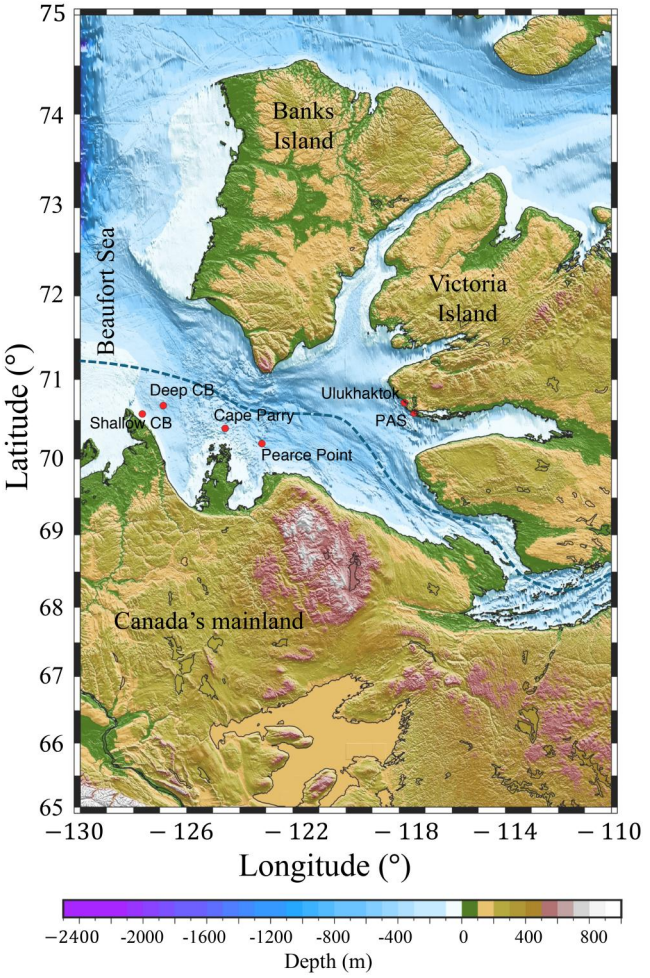

FIG. 1. Location of hydrophone datasets (red circles) in the western Canadian Arctic, used for developing and evaluating the performance of the ship noise detection methods. The background Bathymetry map is derived from the GEBCO 2024 Grid provided by the GEBCO Compilation Group (2024). The abbreviations "CB" and "PAS" stand for CB and PAS, respectively. The blue dashed line shows the Northwest Passage within the study area (obtained from Arctic Portal, 2025).

with local background energy levels prior to analysis. ST500 and ST600 STD hydrophones (Ocean Instruments, Auckland, New Zealand) exhibited a flat frequency response of about –174 and dB re. 1 V/μPa respectively between 20 Hz and 60 kHz, with short non-vessel tonal artifacts, lasting approximately 0.1 s every 30 to 50 s, across the 15–25 kHz band (Dingwall *et al.*, 2024). These artifacts were negligible and outside the frequency range of interest, and so were not removed. For the ST500 deployment at CP during 2022–2023, hydrophone serial information was missing; thus, a provisional sensitivity value of –175 dB re. 1 V/μPa, based on the similar hydrophone model ST500 used in PAS, was applied for sound pressure level (SPL) processing.

In total, 17 665 audio files, each five minutes long and collected during the shipping season, were annotated at the file level for the presence or absence of vessel noise (Table II). Of these, less than 7% (1151) were annotated as vessel noise, and about 93% (16 514) were annotated as background noise. The annotated data show a median

04 August 2026 15:12:24

TABLE I. Summary of hydrophone deployments at the study sites. "Hyd." stands for hydrophone, and the duty cycle represents the ratio of active recording time to the total period over which the recording cycle repeats. The labels (22–23) and (23–24) denote the 2022–2023 and 2023–2024 deployment periods, respectively.

Site	Hydrophone model	Deployment period	Water/Hyd. depth (m)	Duty cycle (min/min)	Sampling rate (kHz)
Ulukhaktok[a]	HTI 92-WB	Jul 2017–Mar 2018	24.4/21.4	5/30	48
PAS	ST500	Jul 2023–Jul 2024	50/48	5/30	48
Pearce Point[b]	ST500	Oct 2018–Sep 2019	351/348	5/60	48
Shallow CB[b]	ST500	Oct 2018–Sep 2019	49/43.5	5/60	48
Deep CB[b]	ST500	Oct 2018–Sep 2019	295/289.5	5/60	48
CP (22-23)	ST500	Oct 2022–Sep 2023	150/140	5/30	48
CP (23-24)	ST600 STD	Sep 2023–Aug 2024	153/150	5/30	48

[a]Halliday *et al*. (2020b).
[b]Dingwall *et al*. (2024).

background SPL of 92.8 dB re. 1 μPa within the 50–1000 Hz frequency band (the dominant range for vessel noise; frequencies below 50 Hz were excluded due to the hydrophones' steep decline in sensitivity; see Dingwall *et al*., 2024; Halliday *et al*., 2020b) during the shipping season (see Sec. IV A for site specific values). This is consistent with the median SPL of 91 dB re. 1 μPa reported in Halliday (2021) for the same frequency range. The Ulukhaktok dataset was used for method development and calibration, and datasets from the remaining study sites, recorded using different hydrophone models, were reserved exclusively for independent performance evaluation. Additionally, the Ulukhaktok annotations, which differentiate between ships and boats, provide an opportunity to evaluate the performance of ship detection methods on a dataset that includes modulating tonal signatures, characteristic of unstable small boats, as illustrated in Figs. 2(c) and 2(d). This was complemented by results from the PAS site, which also predominantly experienced boat noise, but under different conditions, such as source distance and speed.

Figure 2 shows four examples of vessel noise signatures in spectrograms that are used to detect vessel presence. They highlight the constant tonal noise produced by a large ship at constant engine and propeller rate [Figs. 2(a) and 2(b)] at the Pearce Point and Ulukhaktok sites, and the modulating tonal noise from a small-engine boat [Figs. 2(c) and 2(d)] at the PAS and Ulukhaktok sites. The high-amplitude, parabola-like pattern observed in Figs. 2(c) and 2(d) is characteristic of the Lloyd's Mirror effect (Bachman and Kaye, 1983). These patterns result from constructive and destructive interference between the direct and reflected (from seabed or sea surface) sound paths from a moving ship or boat passing near a hydrophone. The observable moderate differences in boat tonal modulation and Lloyd's Mirror patterns between Ulukhaktok and PAS are due to the duty cycling of boats near the Ulukhaktok hydrophone versus more distant passing boats at PAS (internal communication).

III. METHODS

In this study, we developed two ship detection methods and evaluated their performance against the original FAV method using the F1-score, the harmonic mean of precision and recall. F1-score is particularly important in imbalanced datasets such as ours, where negative instances (background noise) greatly outnumber positive instances (vessel noise). The F1-score is calculated using the following equation:

$$\mathrm{F1} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}. \quad (1)$$

Precision is defined as the proportion of true positive classifications among all positive classifications, and Recall is the proportion of true positive classifications among all actual positives, as shown in Eq. (2) and (3), respectively,

$$\mathrm{Precision} = \frac{TP}{TP + FP}, \quad (2)$$

TABLE II. Summary of annotations across the study sites. For each location, the table reports the number of background, ship, and boat noise annotations, site water and hydrophone deployment depths, along with the recording periods of the annotated data. Annotations were performed at the file level. All recordings are 5 min in duration and sampled at 48 kHz.

Site	Background annotations	Ship annotations	Boat textbf annotations	Annotated deployment period
Ulukhaktok	2627	313	290	Jul–Sept 2017
PAS	2884	0	64	Jul–Sept 2023
Pearce Point	1486	70	0	Oct 2018, Aug–Sept 2019
Shallow CB	1382	55	0	Oct 2018, Aug–Sept 2019
Deep CB	1552	81	0	Sept–Oct 2018, Aug–Sept 2019
CP (22-23)	4146	191	0	Jul–Sept 2023
CP (23-24)	2437	87	0	Sept 2023, Jul–Aug 2024

04 August 2026 15:12:24

https://doi.org/10.1121/10.0043937

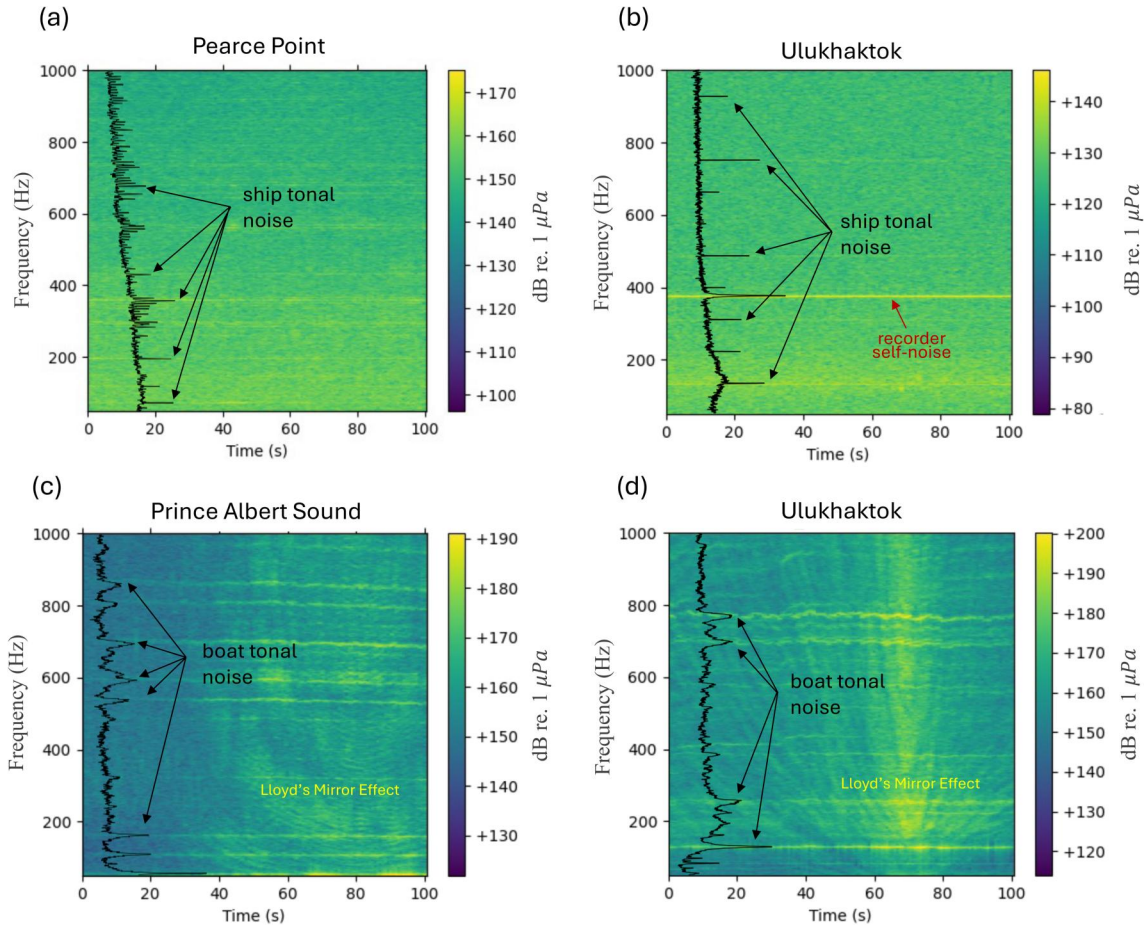

FIG. 2. Spectrogram plots showing examples of stable and modulating tonal noise from large ships at the Pearce Point (a) and Ulukhaktok (b) sites, and small boats recorded at the PAS (c) and Ulukhaktok (d) sites. Each panel includes the temporally averaged spectrum overlaid on the spectrogram (black line). Note the differences in the width of the averaged spectra (broad tonal peaks for the boat and narrow tonal peaks for the ship), as well as the presence of the Lloyd's Mirror effect in (c) and (d).

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{3}$$

where, TP is the number of true positives, FP is the number of false positives, and FN is the number of false negatives.

The proposed methods include a MFAV technique, which builds on and enhances the original FAV method, and a trained CNN model. These approaches were assessed for their ability to automatically detect ship (and boat) noise signatures in passive acoustic recordings from western Canadian Arctic.

A. FAV

The FAV algorithm is an LOFAR method and detects ship noise by identifying stationary narrow-band spectral signatures (i.e., tonals) in spectrograms. While it can operate across a broad frequency range, it is generally applied below 1 kHz, where ship noise dominates.

The original FAV method consists of four main steps:

Step 1: The detection process begins by computing the spectrogram of an acoustic signal based on the short-time Fourier transform (STFT). The spectrogram is computed using a 1.5-s window with a 0.75-s step size (50% overlap) and a Hamming taper to reduce spectral leakage. This configuration provides an effective frequency resolution of 0.67 Hz. This amplitude spectrogram is then divided into multiple time segments (Fig. 3), with segment lengths informed by a heuristic estimate of expected tonal noise duration (i.e., the minimum timing of detectable vessel noise). Within each segment, the spectrogram is averaged over time to obtain the corresponding frequency spectrum. The averaging process enhances the persistent ship noise signature while suppressing random background noise. The resultant spectrum is smoothed using a Blackman window filter (black curve in Fig. 3).

Step 2: The difference between adjacent frequency bins in the smoothed spectrum is computed and then cubed (where the power of 3 was selected empirically, as suggested by Reis *et al.*, 2019). This cubed difference curve yields a detrended transformed spectrum that emphasizes rapid spectral variations and hence highlights tonal features [Fig. 4(a)]. It is important to note that with this detrending approach (i.e., using a differencing operator), every single tonal feature produces two corresponding positive and negative peaks (peak and trough) in the resultant curve.

04 August 2026 15:12:24

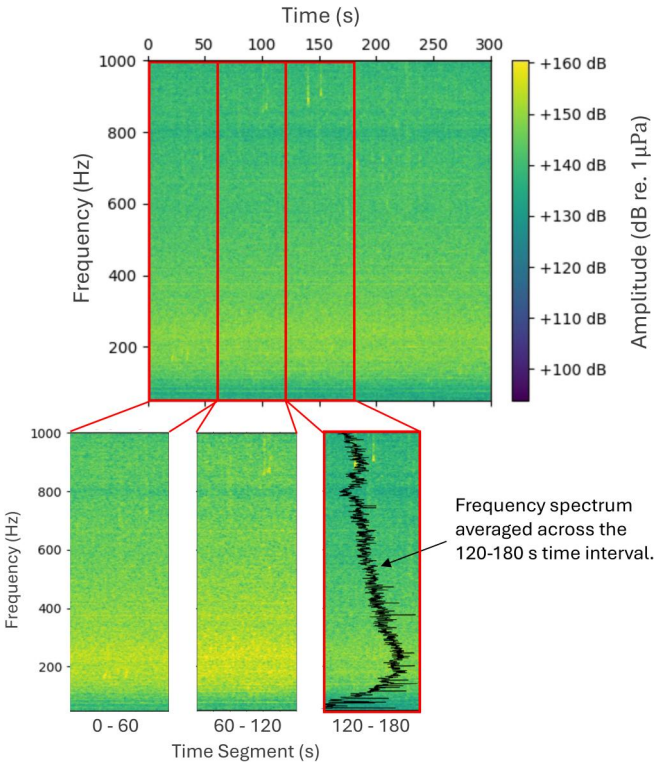

FIG. 3. Spectrogram representation of a 5-min hydrophone recording over a 50–1000 Hz frequency band, segmented into 60-s time windows. The black curve represents the averaged frequency spectrum of the 120–180 s segment.

Step 3: To retain only the positive peaks, the cubed differenced curve is first inverted (multiplied by −1) and subsequently shifted by three frequency bins so that the corresponding peaks align. This transformed (i.e., negated and time-shifted) version is subsequently multiplied element-wise with the original cubed difference signal, effectively squaring the positive portion (i.e., the peaks) of the detrended signal while nullifying the negative components (i.e., the troughs) [Fig. 4(b)].

Step 4: To reject weak variations, a thresholding operation is applied to the resultant curve shown in Fig. 4(b). Frequency bins with values exceeding the threshold are assigned a value of 1, indicating the presence of a tonal peak, while those below the threshold are assigned 0. The threshold is typically set to 1.5 times the squared standard deviation of the smoothed spectrum obtained in Step 1 (Reis *et al.*, 2019), hereafter referred to as the "FAV sensitivity threshold," but may be adjusted as needed. The resulting curve, shown in Fig. 4(c), highlights prominent tonal peaks that are potentially associated with ship noise.

B. MFAV

We developed the MFAV ship detection method with two primary objectives in mind. The first is to establish a more interpretable, statistically based framework that provides analysts with more intuitive control over the ship noise detection process. The second objective is to address the limitations of FAV in scenarios where boat noise is present (Reis *et al.*, 2019), by providing greater control over the frequency bandwidth of tonal noise to be retained. As previously discussed, the FAV algorithm is effective at detecting large commercial vessels (e.g., tankers), which generate persistent tonal components in spectrograms. However, its performance tends to degrade when detecting small boats

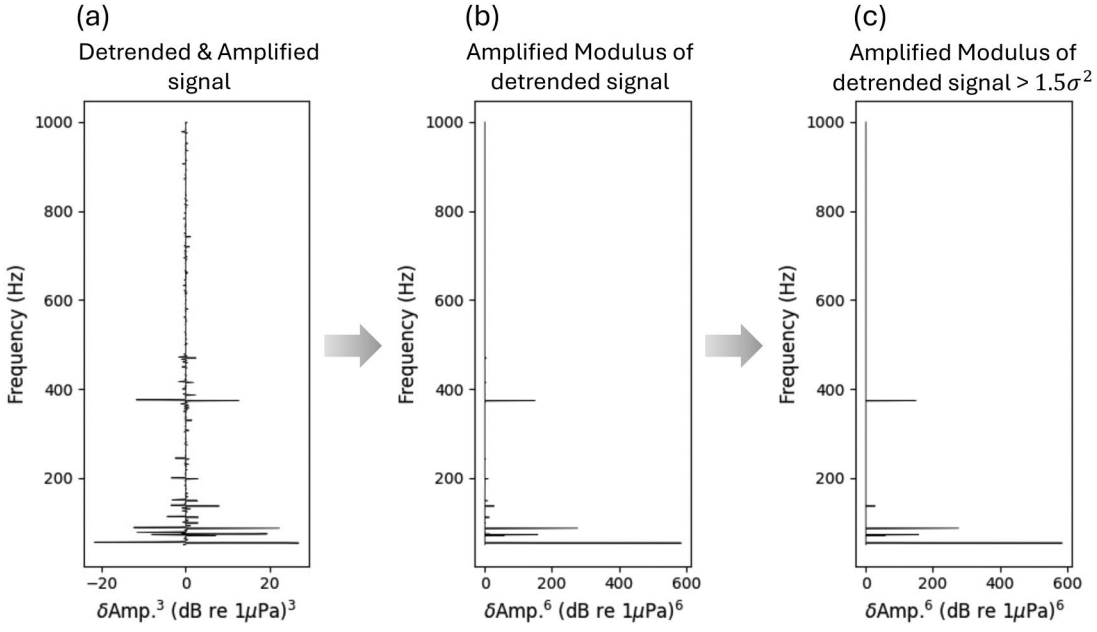

FIG. 4. Schematic illustration of the FAV method used to detect ships' noise signals. (a) Shows the detrended and cubed plot of the averaged spectrum (the black curve in Fig. 3), obtained by raising the differences over consecutive frequency bins to the power of 3. (b) Displays the squared positive portion of the detrended signal, with negative components set to zero. (c) Shows a filtered version of the curve from (b), where only values exceeding 1.5 times the squared standard deviation of the averaged spectrum are retained.

04 August 2026 15:12:24

https://doi.org/10.1121/10.0043937

(often outboard motors), which typically exhibit modulated tonal signatures [e.g., Figs. 2(c) and 2(d)].

The MFAV procedure comprises the following steps:

Step 1: As in the original FAV method, the smoothed average frequency spectrum is computed for each time segment following spectrogram segmentation [Fig. 5(a)].
Step 2: A 5th-order high-pass Butterworth filter with a "pseudo cutoff frequency" of 5 kHz^{-1} (empirically determined) is applied to the spectrum to isolate rapid amplitude fluctuations across frequency, effectively detrending the spectrum [Fig. 5(b)] (see Sec. IV E for more details). Note that the "pseudo" prefix reflects the filter's application in the frequency domain, resulting in units of Hz^{-1}. This approach enhances flexibility in preserving the modulated tonal signatures characteristic of small boats, thereby facilitating the detection of boat tonal noise signatures.
Step 3: The histogram of the high-pass filtered spectrum resembles a normal distribution [Fig. 5(c)], characteristic of random background noise. Values in the high-pass filtered spectrum that exceed 3.5 times its standard deviation (default value), hereafter referred to as the "MFAV sensitivity threshold," are here classified as statistical outliers [Fig. 5(c)]. The outliers in the estimated random background noise distribution can be potentially attributed to tonal vessel noise peaks.

We defined the default MFAV sensitivity of 3.5 standard deviation from the mean threshold through a trial-and-error process; however, this can be adjusted to suit different detection needs. Lower threshold values will increase the recall metric by enabling detection of vessel noise with weaker amplitudes close to background levels; however, this also increases the risk of falsely identifying background noise as vessel noise. Conversely, higher threshold values enhance the precision metric by restricting detections to stronger vessel noise signals, typically originating from closer vessels, thereby reducing false positives. Also, note that in Fig. 5(c), fitting a normal distribution to the high-pass filtered spectrum is unnecessary (it merely serves to show that using the sample standard deviation is reasonable); only the sample standard deviation is required for subsequent analysis.

Note that, although our data from various sites exhibited Gaussian-like background noise, strongly non-Gaussian noise remains possible. In such cases, percentiles may provide a more robust criterion than the empirical standard deviation.

C. DL model

As an alternative to the signal-processing approach, we trained a CNN model for binary ship classification. CNN models, widely recognized for their success in image recognition tasks (Khan *et al.*, 2020), have also demonstrated strong performance in processing complex audio data by capturing patterns in spectrograms and other acoustic representations (Frazao *et al.*, 2025; LeCun *et al.*, 1989; LeCun *et al.*, 2015). These models tend to generalize well across diverse acoustic environments and often match or exceed the performance of traditional signal processing methods (Hummel *et al.*, 2024). Several alternative CNN architectures can be used to improve both accuracy and computational efficiency. VGG architecture (Simonyan and Zisserman, 2014) enables the use of deeper networks (i.e., more layers) but is limited by high computational costs.

04 August 2026 15:12:24

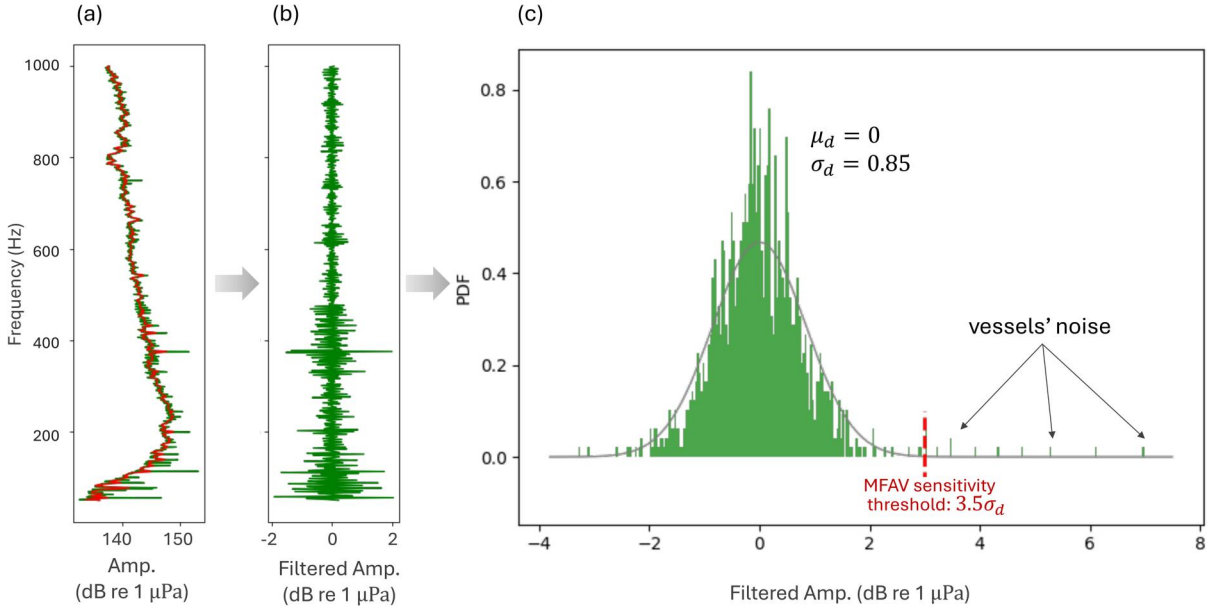

FIG. 5. Schematic illustration of the MFAV method. (a) Shows the averaged spectrum of the 120–180 s time segment of Fig. 3 (green curve) overlain by the low-pass filtered spectrum (red curve), with a pseudo cutoff frequency of 5 kHz^{-1}. (b) Shows the filtered frequency spectrum after the application of a 5 kHz^{-1} high-pass filter to the averaged spectrum. (c) Shows the scaled histogram plot (empirical pdf) of the high-pass filtered spectrum. The green lines show the data, and the gray curve represents a normal distribution fit to the data. The parameters μ_d and σ_d denote the estimated mean and standard deviation of the data, respectively. The $3.5\sigma_d$ threshold used to define the vessel noise outliers is shown by the red dashed line.

GoogLeNet (Szegedy *et al.*, 2015) improves the computational efficiency of deeper and wider networks (i.e., more nodes per layer) but struggles with very deep networks due to the vanishing gradient problem. ResNet (He *et al.*, 2016) allows the training of substantially deeper networks and achieves higher accuracy. EfficientNet (Tan and Le, 2019) similarly offers the potential for both efficiency and accuracy improvements. However, we selected ResNet because it is a simpler and more transparent architecture, offering memory efficiency and modular building blocks that enable flexible customization and fine-tuning in transfer learning, including adaptation to new environments, features that are particularly suitable for our study.

1. DL model architecture and optimization configuration

The ResNet architecture used in this study comprised an initial two-dimensional (2D) convolutional layer with 16 kernels of size 3×3 and a stride of 1, followed by three residual block sets. Each set contained two blocks, and each block had two or three 2D convolutional layers, depending on whether a skip connection was included. All convolutional layers within the blocks also used 3×3 kernels with a stride of 1, but the number of kernels doubled with each block (16, 32, and 64). The network had a total of 15 convolutional layers. Finally, a fully connected layer with two output neurons was used for classification. The network was trained using the Adam optimizer (Kingma and Ba, 2014) with a learning rate of 0.001, a decay factor of 0.01, $\beta_1 = 0.9$, and $\beta_2 = 0.999$. The learning rate defines the initial step size for weight updates during backpropagation, while the decay factor gradually reduces this step size at each epoch, allowing for smaller updates and helping to avoid overshooting global minima. The β_1 and β_2 parameters are the exponential decay rates for the moving averages of the gradient and the squared gradient, respectively. This helps to smooth out short-term fluctuations and enhance overall training stability. FScoreLoss was used as the objective function. This is the inverse of a differentiable approximation of the F1-score that guides the model updates to improve both precision and recall simultaneously for imbalanced datasets.

2. Training dataset

The "Ulukhaktok" audio files, acquired in 2017 during the shipping traffic period from July to September, were used to train the model (Table II). The DL model was developed specifically to detect ship noise. A total of 524 audio files were selected, 265 positive (with ship noise) and 259 negative (no ship noise), to maintain a balanced training (and validation) set. An additional 56 audio files (26 positive and 30 negative) were retained as the test set (in addition to the test datasets from the remaining sites). Each recording is 5 min in duration and is annotated for the presence or absence of ship noise at the file level. Data augmentation was performed by dividing each file into 12 non-overlapping 25-s segments, with each segment inheriting the annotation label of its parent audio file, based on the assumption of persistent tonal noise throughout the 5-min recordings. Visual inspection of recordings confirmed that this assumption is generally valid, supporting the pragmatic choice of this approach over a more precise but less efficient fine-scale annotation approach. These 25-s segments were converted into magnitude spectrograms using a 1.5-s Hamming window, overlapped by 0.75 s. The selected window length provides sufficient frequency resolution (i.e., <1 Hz) to capture ship tonal noise, while the 0.75-s overlap ensures a reasonably smooth spectrogram. Note that, due to the temporal consistency of ship tonal noise, high temporal resolution is unnecessary. The frequency range was limited to 50–1000 Hz. Furthermore, we normalized the spectrograms to enable the model to learn relative patterns rather than absolute signal intensities, thereby enhancing its ability to generalize across recordings made with hydrophones placed at varying distances from ship or under different background noise conditions.

3. Training strategy

We employed a stage-wise, predefined curriculum learning strategy, wherein the model was initially trained on simpler examples (25-s segments) and subsequently exposed to increasingly complex ones. This improves generalization and facilitates convergence of the optimization algorithm (i.e., backpropagation process) in the presence of non-convex objective functions (Bengio *et al.*, 2009; Wang *et al.*, 2022). The choice of data complexity levels was informed by the insight that the MFAV method performs better when more tonal frequencies are present. The positively annotated 25-s segments were therefore subdivided according to the number of tonal peaks present (as defined by the MFAV approach) using thresholds of 7 and 11 peaks. These thresholds were chosen to ensure an approximately equal distribution of positively annotated recording segments across the subgroups. The negatively annotated pairs were selected from those negatively annotated samples in which the MFAV method also returned zero detected peaks. The model was trained in three stages by sequentially introducing sub-datasets with decreasing numbers of peaks, as follows:

Stage 1: 25-second segments with more than 11 peaks, representing strong ship-noise signatures (comprising 91 positive and 88 negative annotations).
Stage 2: 25-second segments with more than seven but fewer than or equal to 11 peaks (comprising 94 positive and 88 negative annotations).
Stage 3: 25-second segments with seven or fewer peaks, considered the most challenging cases due to weaker or less distinct ship-noise features (comprising 80 positive and 83 negative annotations).

At each stage, a tenfold cross-validation was conducted, with each fold using a randomly selected 90% of the data

04 August 2026 15:12:24

https://doi.org/10.1121/10.0043937

for training and the remaining 10% for validation. The training progress of this binary classifier was evaluated using accuracy, precision, and recall metrics. The network was trained for 30 epochs with a batch size of 8, and the best-performing model from each stage was used as the starting model for the subsequent stage. Model performance was evaluated by applying all ten trained models at each stage to the Ulukhaktok test set data and selecting the model that achieved the highest F1-score, thereby implying stronger generalization capacity.

The performance of the trained DL model on the remaining 56 audio files (Ulukhaktok test set) is presented in terms of F1-score, recall, and precision in Fig. 6. The model achieved its best performance at the classifier's confidence score threshold of 0.9, where all three metrics were near 98%.

IV. RESULTS

The results of our study are presented as follows. We first characterize and compare the acoustic environments at each location. This is followed by an analysis of tonal peak characteristics detected by the MFAV and FAV methods, focusing on the number of peaks and their signal-to-noise level differences (SNLDs). Subsequently, we assess the performance of each method across the study sites based on the highest achieved F1-score and compare the results to those obtained from the original FAV method. Further analysis is conducted to examine the performance of methods by vessel type (ship versus boat), leveraging the Ulukhaktok dataset where ship and boat noise annotations are provided separately (Table II). Additionally, we examine the impact of varying the pseudo cutoff frequency in the MFAV method on boat detection performance. All analyses are confined to the 50–1000 Hz frequency band, within which vessel noise is typically dominant. Finally, it should be noted that the reported performance metrics for all cases are averages over multiple test subsets, carried out to address the imbalance between positive and negative annotations. Each subset includes all positive annotations from a given site, paired with a uniquely selected (i.e., non-overlapping) set of negative samples to maintain class balance.

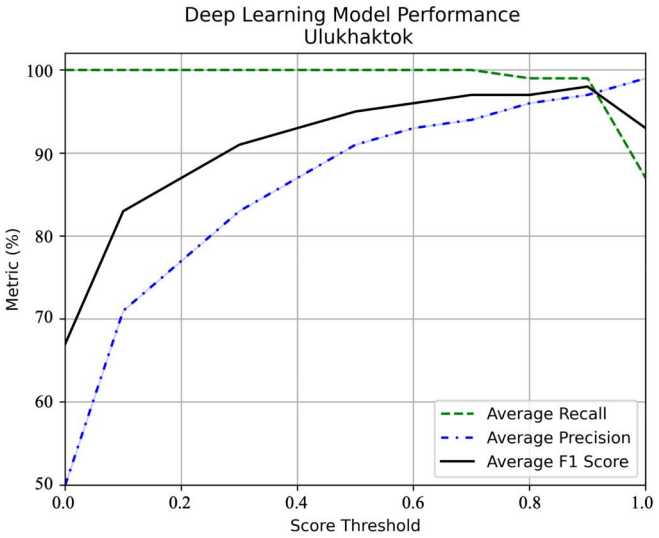

FIG. 6. The performance of the trained DL model on the Ulukhaktok test dataset after the third curriculum learning stage is shown. The black curve represents the F1-score, the green dashed line indicates recall, and the blue dashed-dotted line shows precision. The model achieved its highest performance at a score threshold of 0.9.

A. Acoustic characteristics of study sites

To examine potential site-specific influences on detection performance, we characterized the acoustic environment during the seven hydrophone deployment periods at the study sites using root mean square (RMS) SPLs in dB re. 1 μPa (hereafter referred to as dB, unless otherwise specified). Figure 7 shows violin plots of SPL distributions in the 50–1000 Hz frequency band at each site, comparing recordings with vessel noise (gray) to those containing only background noise (green). As shown in Fig. 7, with the exception of PAS and CP (23–24), the median SPL across the remaining study sites increases by up to 3 dB within the 50–1000 Hz band, whereas the 75th percentile and the maximum SPL values increase by up to 8 and 25 dB in the presence of vessel noise, respectively. The slight elevation in median SPL can be attributed to the composition of the dataset, in which a few recordings containing vessel noise are outweighed by the majority representing only background noise (Table II). On the other hand, the maximum SPL values are prone to overestimation caused by non-vessel sources such as wind, considering the fact that the recorded values (130–150 dB re. 1 μPa) are within the same order of magnitude of those reported by Halliday (2021), i.e., 127 dB re. 1 μPa. Therefore, the 75th percentile SPL change seems to be a more reliable indicator of vessel noise presence. Both CP (23–24) and Pearce Point exhibit the greatest increase in the 75th percentile. In contrast, the CP (22–23) exhibits the smallest increase in 75th percentile SPL. Contrary to the overall trend, the PAS site exhibits a 6% reduction in the median SPL and a 8% reduction in the 75th-percentile SPL in the presence of vessel noise. The available weather data from the nearest station at Ulukhaktok show a negative correlation between boat activity and wind speed for the PAS data (see Appendix A, Fig. 11), which could help to explain the higher SPL without vessel noise presence. However, this negative correlation between wind speed and boat activity is also evident in the Ulukhaktok dataset (see Appendix A, Fig. 12), which shows an increase in SPL in the presence of vessel noise. This could be attributed to an inadequate sampling strategy, in which only a few boat-noise recordings were made during relatively quiet periods, while many background-noise samples were collected under high wind conditions, and therefore may not accurately reflect the expected increase in SPL in the presence of boat noise. Another possibility is that some vessel-noise events were missed by the annotator during windier periods. These observations highlight the importance of interpreting SPL data within a well-defined

04 August 2026 15:12:24

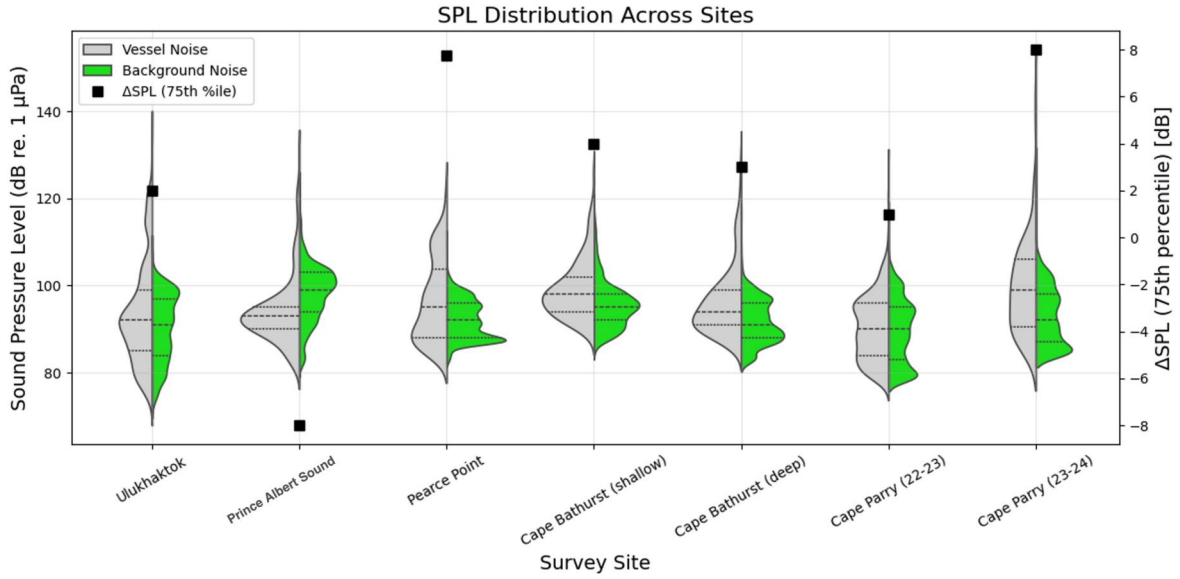

FIG. 7. Violin plots of SPLs (50–1000 Hz) recorded during shipping seasons across seven hydrophone deployments in the Western Canadian Arctic. Green indicates the distribution of background noise only (i.e., no vessel noise), while gray represents the distribution for data including vessel noise. Each violin illustrates the distribution of SPL values, with the width representing the density of measurements. The internal dotted lines indicate the lower and upper quartiles, and the dashed line marks the median SPL. Differences between 75th percentile SPL with and without vessel noise at each site are indicated by black square markers.

environmental context and considering potential sampling bias, rather than relying solely on raw SPL values.

B. Characteristics of detected tonal noise by MFAV and FAV

The characteristics of the tonal peaks detected by the MFAV and FAV methods across the study sites are analyzed in terms of the number of tonal peaks detected per instance (spectrogram segment) and their SNLDs in decibels [Eq. (4)], defined as the difference between peak and local background SPLs in the vicinity of the tonal (i.e., how much the tonal stands out from background),

$$\mathrm{SNLD\,(dB)} = \mathrm{SPL}^{\mathrm{tonal_peak}}\,[\mathrm{dB}] - \mathrm{SPL}^{\mathrm{local_background}}\,[\mathrm{dB}], \quad (4)$$

where $\mathrm{SPL}^{\mathrm{tonal_peak}}$ and $\mathrm{SPL}^{\mathrm{local_background}}$ denote the detected tonal peak SPL and the local background SPL in the vicinity of the tonal peak, respectively. This analysis offers insight into the relative performance of the two methods under the given parameter settings. Figure 8 shows the distribution of SNLD values for tonal peaks detected by MFAV and FAV, indicating that MFAV consistently identifies more instances of tonal noise with lower SNLD (i.e., weaker tonal noise) compared to the original FAV approach. On the other hand, the box plot in Fig. 9 shows that the FAV method generally detects a higher number of tonal peaks per recording. Based on the median number of tonal peaks, the MFAV and FAV methods have detected the highest number of ship tonal peaks at Pearce Point and CP (23–24) sites, respectively, while both methods have detected the fewest ship tonal peaks at CP (22–23) and deep CB. Additionally, both methods have returned the fewest tonal peaks at the PAS site, where only boat noise is present.

C. Method performance across study sites

Table III summarizes the highest achieved F1-scores (hereafter referred to as F1-score) for the DL, MFAV, and FAV methods across all study sites.

All methods perform well at the shallow and deep CB sites, CP (23–24), and Pearce Point, with F1-scores ranging from 91%–95%. At the remaining sites, Ulukhaktok, PAS, and CP (22–23), the DL model achieved F1-scores of 84%, 91%, and 94%, respectively, while MFAV and FAV yielded scores between 83% and 89% and between 67% and 85%, respectively. The superior performance at CP (23–24) and Pearce Point, with F1-scores from 93% to 97%, can be attributed to the larger number of detected tonal peaks and the more pronounced increases in the 75th percentile SPLs due to vessel noise (Figs. 7 and 9), which likely enhanced the detectability of vessel noise. In contrast, the lowest performance is observed at CP (22–23), where both the increase in the 75th percentile SPL and the number of detected tonal peaks are negative or relatively low. Nonetheless, no consistent general correlation is observed between performance and either site characteristics or tonal noise features.

Overall, the DL model outperforms the other two methods at all sites, except at Ulukhaktok, where the F1-scores of all methods range from 84% to 88%. This lower performance is attributable to the high number of boat noise instances (290 audio files) in the test set, nearly matching the number of ship noise instances (313 audio files) (Table II). As shown in Table IV, when evaluation at Ulukhaktok is restricted to ship noise detections, performance improves

04 August 2026 15:12:24

https://doi.org/10.1121/10.0043937

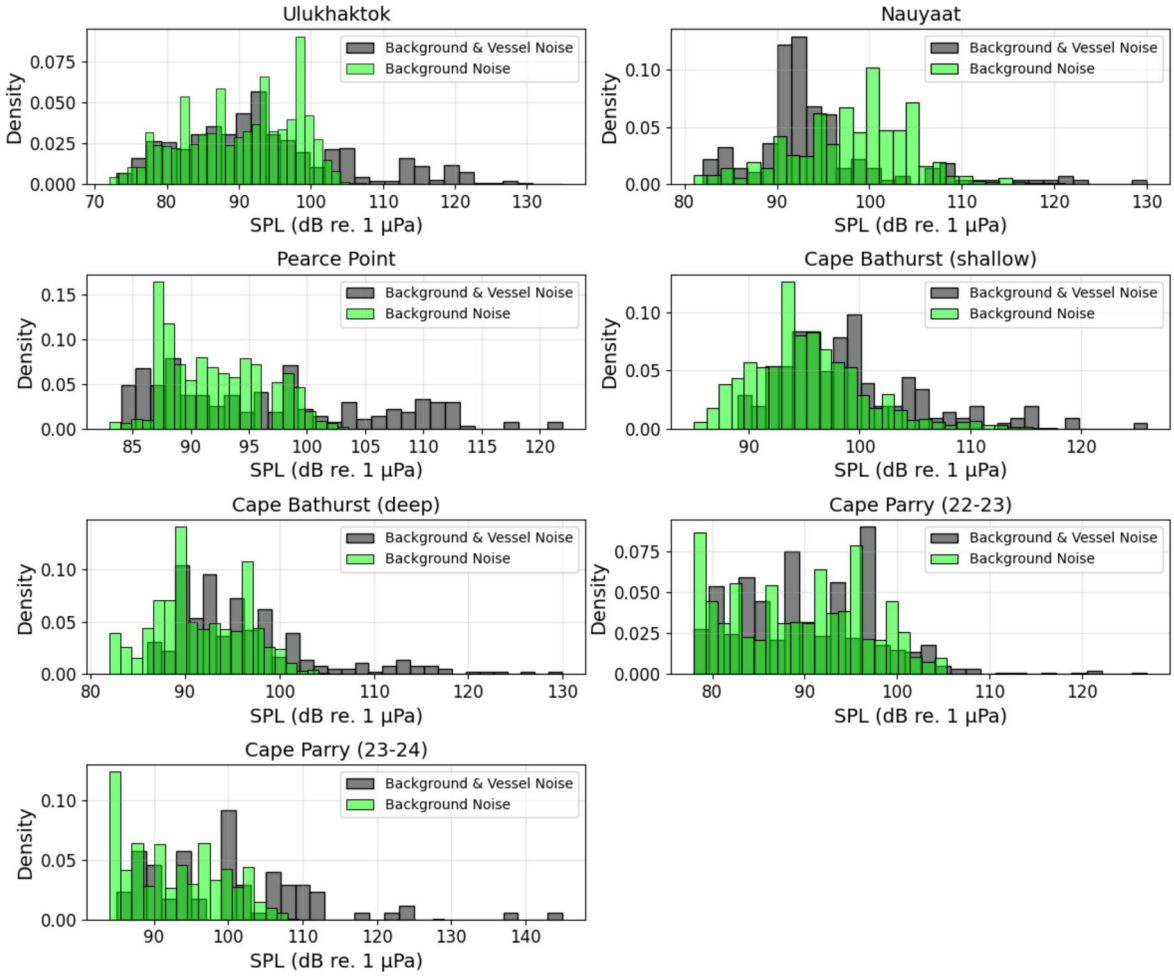

FIG. 8. Distribution of SNLD (in dB) of tonal peaks (i.e., the difference between peak and local background noise levels) detected by MFAV (gray) and by FAV (cyan) methods across study sites.

substantially, with F1-scores rising to 92–94%. MFAV generally outperforms FAV, except at Pearce Point, where it falls behind by just 1%. At PAS, where only boat noise is present, MFAV surpasses FAV by 22% (Table IV), while at other sites its margin of improvement is up to 6%.

D. Method performance by source type

The performance of the DL model, MFAV, and FAV methods by source type (i.e., ship versus boat) is compared using the Ulukhaktok dataset (Table IV), which includes

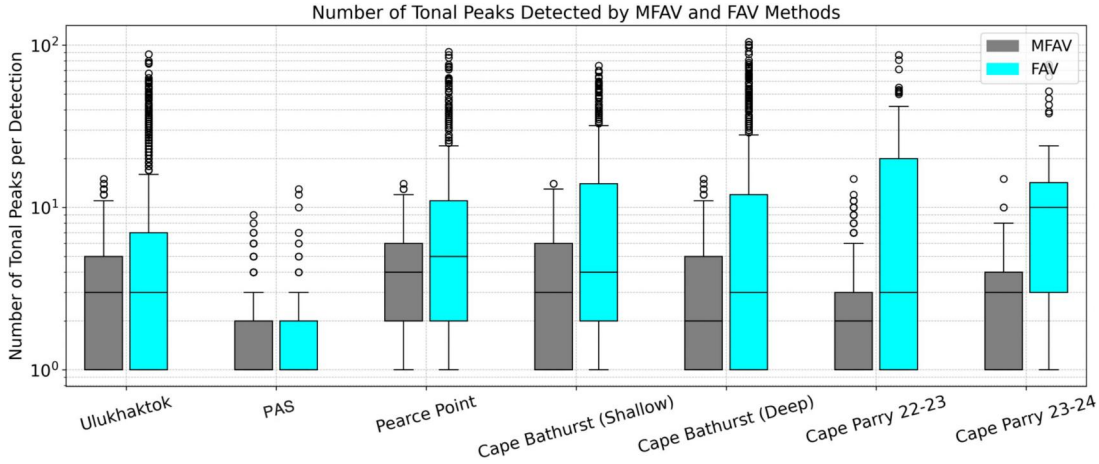

FIG. 9. Box plot of the number of detected tonal peaks using the MFAV and FAV methods across the study sites. Each box shows the interquartile range (IQR), with the horizontal line indicating the median SPL. The whiskers extend to 1.5 times the IQR, while points beyond the whiskers show outliers.

04 August 2026 15:12:24

TABLE III. Overview of the performance of FAV, MFAV, and DL methods across study sites based on the highest achieved F1-score metric. Labels (22–23) and (23–24) denote the 2022–2023 and 2023–2024 deployment periods, respectively.

		Study Site						
		Ulukhaktok	PAS	Pearce point	Shallow CB	Deep CB	CP (22–23)	CP (23–24)
Method	**FAV**	0.85	0.67	0.95	0.91	0.92	0.77	0.95
	MFAV	0.88	0.89	0.94	0.92	0.92	0.83	0.97
	DL	0.84	0.94	0.95	0.93	0.94	0.91	0.93

313 audio files annotated as ship noise and 290 audio files annotated as boat noise (Table II), all recorded under consistent hydrophone and environmental context. Since the DL model is trained specifically to detect ship noise, its performance is notably higher when evaluated on audio recordings containing only ship noise. MFAV outperforms the original FAV method by 10% in detecting tonal noise from boats (Table IV). This can be attributed to the modulating nature of tonal noise from small boats [Fig. 2(d)], which compromises the performance of the FAV method. In contrast, the proposed MFAV method demonstrates more robust performance in extracting modulating tonal signatures. This observation is consistent with the 22% increase in F1-score obtained by MFAV at the PAS site, where only boat noise is present, albeit with moderate tonal modulation [Fig. 2(c)] due to differences in boat operation and under a different environmental context. The effect of moderate modulation is also evident in the performance of the DL model applied to the PAS dataset, which yielded an F1-score of 91% (Table III), compared to only 76% when the model is applied to the Ulukhaktok data restricted to the boat noise instances (Table IV). MFAV method also demonstrates a marginal improvement (i.e., 2%) in detecting large ships at the Ulukhaktok site (Table IV). This improvement can be attributed to the MFAV's enhanced sensitivity to weak tonal noise as evident in Fig. 8. These observations imply that method selection may depend on the specific characteristics of the target vessel type.

E. Sensitivity of MFAV performance at varying high-pass pseudo frequency threshold

The MFAV technique applies a Butterworth high-pass filter of 5th-order to detrend an average spectrum derived from the corresponding spectrogram of an audio segment.

TABLE IV. The highest achieved F1-score of each method at the Ulukhaktok site, grouped by vessel type: ship, boat, and both.

		Annotated label		
		Ship noise	Boat noise	Ship & boat noise
Method	**FAV**	0.92	0.76	0.85
	MFAV	0.94	0.86	0.88
	DL	0.94	0.76	0.84

The pseudo cutoff frequency of this high-pass filter determines the width of tonal peaks retained for further analysis. That is, higher pseudo cutoff frequencies preserve only narrower peaks, whereas lower pseudo cutoff frequencies allow wider peaks to remain. Unlike the FAV method, this detrending approach offers flexibility to retain broader peaks in the temporally averaged spectrum, which result from modulating tonal noise produced by small-engine boats, like the ones observable in Fig. 2(c) and 2(d). The effect of adjusting the pseudo cutoff frequency on the filtered spectrum, and consequently on the preservation of tonal peaks, is illustrated in Fig. 10 for three pseudo cutoff frequencies of $5\,\mathrm{kHz}^{-1}$, $2\,\mathrm{kHz}^{-1}$, and $100\,\mathrm{Hz}^{-1}$. We have evaluated two high-pass pseudo cutoff frequencies, the default $5\,\mathrm{kHz}^{-1}$ (referred to as "ship mode") and a lower $2\,\mathrm{kHz}^{-1}$ (referred to as "boat mode"), which is intended to more effectively capture modulating tonal components from boats. As shown in Table V, this adjustment results in a minor decrease in F1-score for both ship and boat noise detection at the Ulukhaktok site. This is because in the "boat mode," the recall metric improves at the cost of precision, resulting in a negligible net change in overall performance as measured by the F1-score. See Appendix B for further details.

V. DISCUSSION

In this study, we developed two automated ship noise detection techniques tailored for application in Canadian Arctic waters. Their performance was assessed using the highest achieved F1-scores and compared against the baseline performance of the original FAV method (Reis *et al.*, 2019) across multiple acoustic monitoring sites in the region of interest. The methods are developed to contribute to the growing need for effective monitoring of vessel noise interference with marine mammal vocalization, due to increased vessel traffic in the region (McWhinnie *et al.*, 2018). A key contribution of this work is the introduction of a robust variant of the FAV ship detection method, the MFAV approach. This novel approach offers a more intuitive analytical framework and appears to enhance performance and the detection of small boats. The second major contribution is the development of a ResNet-based CNN ship detector, specifically tailored for use in western Canadian Arctic waters, leveraging in-house acoustic recordings from the region. CNN-based algorithms have been reported to demonstrate

04 August 2026 15:12:24

https://doi.org/10.1121/10.0043937

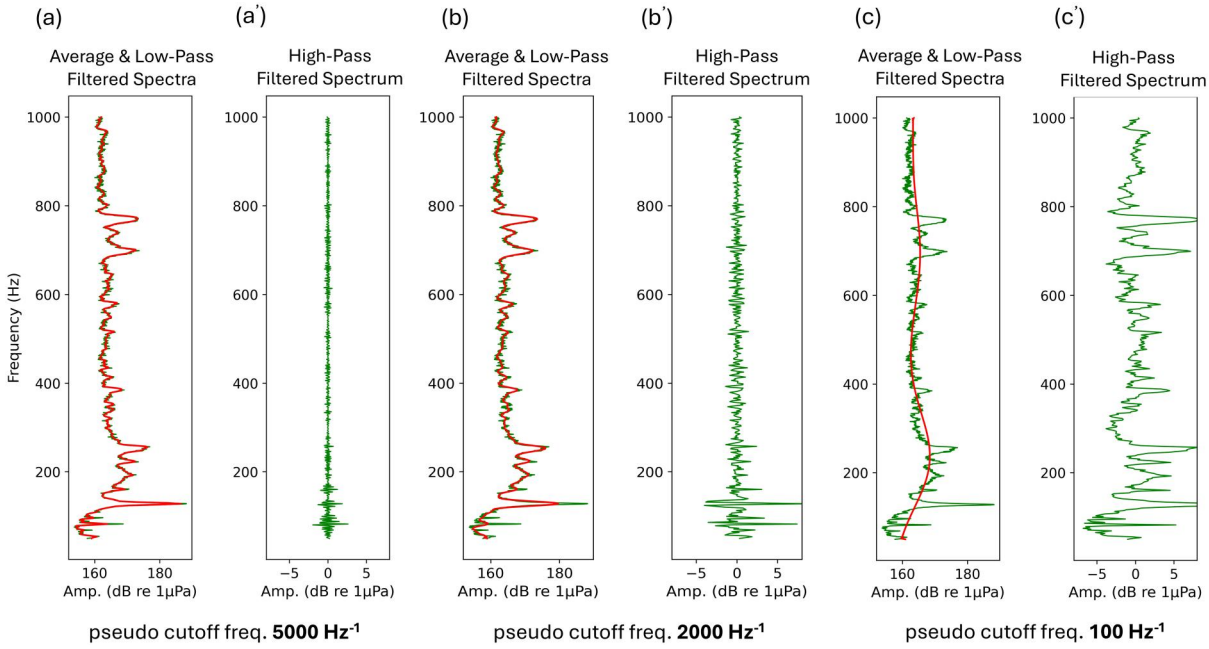

FIG. 10. Effect of adjusting pseudo cutoff frequency on the high-pass filtered spectrum of a spectrogram. (a), (b), and (c) show the average spectra (green curves) corresponding to the spectrogram in Fig. 2(d), overlaid with low-pass filtered spectra (red curves) using pseudo cutoff frequencies of $5\,\mathrm{kHz}^{-1}$, $2\,\mathrm{kHz}^{-1}$, and $100\,\mathrm{Hz}^{-1}$, respectively. The difference between each average spectrum and its corresponding low-pass filtered version yields the high-pass filtered spectrum shown in (a′), (b′), and (c′). Note that, in practice, the high-pass filtered spectra are obtained by directly applying a high-pass filter, rather than by subtracting the low-pass filtered signal from the original spectrum.

promising results in non-Arctic marine environments by other authors, e.g., Ferguson *et al.* (2017); however, their applicability under Arctic conditions has been largely unexplored. In addition, the use of a curriculum learning strategy in this context appears to be successful in spite of the relatively small number of training data samples. Despite being trained on data from a single site, the DL model demonstrates strong generalization across recordings from other hydrophone deployments in the region, requiring no further fine-tuning. To facilitate adoption of the developed methods by the Arctic PAM community, the implementations have been made publicly available as an open-source ship-noise detection tool on GitHub, accompanied by a comprehensive user guide documentation.

It is worth mentioning that we also tested the DEMON approach, but encountered similar limitations to those reported in the literature. These challenges prevented us from automating the DEMON method, ultimately leading to its exclusion from further use in this study.

TABLE V. Highest F1-scores achieved by the MFAV method at the Ulukhaktok site for each vessel type (ship, boat, and combined), under different operational modes ("ship mode" vs "boat mode"), as defined by the high-pass pseudo cutoff frequency parameter used in the MFAV method.

		Annotated label		
		Ship noise	Boat noise	Ship & boat noise
MFAV run mode	**Ship Mode**	0.94	0.86	0.88
	Boat Mode	0.93	0.85	0.88

A. Site-specific underwater acoustic characteristics

To ensure a fair comparison of the methods and to facilitate the synthesis of findings across the study sites, we analyzed background (no vessel noise) and vessel-induced elevated RMS SPLs against the methods' performance. While background SPLs (green violin plots in Fig. 7) reflect variations in the acoustic environment and hydrophone characteristics across the study sites, the vessel-induced elevated SPLs (i.e., the difference between corresponding gray and green violin plots in Fig. 7) provide insight into source-related differences such as vessel type and proximity to the hydrophones.

In general, the measurements revealed minor variations in background SPLs (less than 5 dB), which may be partly attributable to differences in sample size (tested but not shown). Regardless of the source, environmental context or numerical, these differences fall within the variation ranges within individual sites, thus, can be reasonably disregarded for method performance comparison across study sites. This is further supported by the consistent performance of all methods across sites, particularly the DL model, which achieved stable results without site-specific fine-tuning (Tables III and IV). However, while this may suggest similar acoustic conditions across sites, it could also be a reflection of the DL model's strong generalization capacity.

04 August 2026 15:12:24

In the analysis of vessel-induced SPL elevations, the elevation in the 75th percentile SPL appeared to be a more robust indicator of vessel noise presence compared to the elevated median SPLs, likely due to the limited number of recordings. Across the study sites, the elevated 75th percentile SPLs differ by less than 4 dB (black square markers in Fig. 7), with the exceptions of Pearce Point and CP (23–24), which exhibited an 8 dB increase, and PAS, which showed a reduced SPL. Although the largest increases in the 75th percentile SPL at Pearce Point and CP (23–24) align with the superior performance of the methods, we could not confirm a consistent correlation between method performance (Table III) and the acoustic characteristics of the study sites (Fig. 7). Notably, elevated SPLs are similar (differ by 1 dB) at both shallow and deep CB sites with comparable performance (differ by 1%), suggesting that shallow deployments might be sufficient for vessel noise detection within the 50–1000 Hz band, offering logistical and cost advantages.

B. Method performance on ship noise detection

Overall, all methods demonstrated satisfactory performance in detecting ship noise across all study sites, with the DL model outperforming the others at all sites, except for CP (23–24), achieving F1-scores of 91%–95%, compared to 83%–97% for MFAV and 77%–95% for the original FAV method (Tables III and IV). The best performance of the DL model was achieved at Pearce Point site, with F1-score of 95%. F1-scores reported in similar studies employing CNN-based models for vessel detection in non-Arctic regions typically fall within the range of 85%–90% (e.g., Kim *et al.*, 2020; White *et al.*, 2022). However, direct comparisons across studies should be interpreted with caution, as differences in datasets, environmental conditions, and evaluation criteria can significantly influence performance metrics. The strong performance of the DL model might be attributed, in part, to the use of the curriculum learning strategy. In this study, the curriculum learning begins by presenting the model with predefined simpler (more easily classifiable) spectrograms containing a higher number of tonal peaks, which serves as a potential proxy for difficulty measure. It must be noted that some samples considered difficult from the DL model's perspective may still exhibit many tonal peaks, or *vice versa*, potentially leading to misgrouping. However, we expect such cases to be rare given the observed correlation between MFAV performance and tonal peak counts. This limitation is inherent to predefined curriculum learning schemes used in this study, as the exact features used by DL models for classification cannot be fully interpreted or controlled (Wang *et al.*, 2022). Nevertheless, this approach leads to a smoother cost function during the early stages of learning, thereby increasing the likelihood of converging to a global minimum. As training progresses, gradually more complex samples are introduced, which effectively increases the complexity of the cost function hierarchically, while assuring the starting point at each learning stage is in the vicinity of the global minimum. This gradual exposure enables the model to learn the robust features of the ship noise in spectrograms in a structured way, effectively compensating for the limited size of the training dataset. We speculate that the commonly occurring tonal noise signatures produced by ships might be the key feature learned by the model. As a result, the trained DL model would exhibit strong generalization capabilities, achieving high performance on unseen datasets with no fine-tuning. Additional obvious features of ship presence, such as Lloyd's Mirror effect [see Figs. 2(c) and 2(d)], may have also contributed to the model learning. However, due to their infrequent occurrence, being limited primarily to shallow hydrophone deployments and typically strongest near the vessel's closest point of approach, the role of such features is not obvious in enhancing the model's overall generalization capacity. To gain insight into the features learned by the DL model, post-processing techniques such as Grad-CAM (Gradient-weighted Class Activation Mapping) (Selvaraju *et al.*, 2017) can reveal influential acoustic features learnt by the DL model. While the DL model performed well on the datasets from the study sites, such insights may inform further fine-tuning and adaptation to improve ship noise detection in more complex and acoustically diverse environments. Nevertheless, the MFAV and the original FAV methods may both serve as a viable approach for ship noise detection, offering only a minor compromise in performance. Yet, we favor the MFAV method over the original FAV, as it provides greater flexibility and a more user-interpretable analytical framework.

C. Method performance on boat noise detection

In applications involving boat noise detection, the MFAV method demonstrates superior performance compared to the DL model and FAV method (Tables III and IV). At the Ulukhaktok site, both the original FAV method and the DL model yield F1-scores approximately 10% lower than that achieved by MFAV. This outcome is expected for the DL model, as it was not specifically trained for boat noise detection. We believe that the characteristic pattern of the Lloyd's Mirror effect observed in boat annotations from the Ulukhaktok and PAS recordings may serve as a strong feature in detecting boat noise that can be learned by the CNN model. This can be explored in future works using the developed DL model in this study, through a transfer learning approach (Michalski, 1983). At the PAS site, however, where only boat noise is present, the DL model shows the best performance, with an F1-score of 94%, and the MFAV method shows a 22% improvement in F1-score compared to the original FAV method. This difference is likely due to the more moderate tonal modulation of boat noise at PAS [Fig. 2(c)], which produces tonal signatures closer to ship noise that are more consistent and easier to detect. The low performance of the original FAV method can be attributed to the suppression of broad tonal peaks, characteristic of modulating tonal noise of boats, during its spectral

04 August 2026 15:12:24

https://doi.org/10.1121/10.0043937

detrending step. In contrast, the MFAV method not only avoids the elimination of broad tonal peaks during its detrending process, but also demonstrates enhanced sensitivity to tonal peaks with lower SNLD (Fig. 8). Although MFAV typically detects fewer peaks per instance (often fewer than 20 in this study), it is more effective at identifying faint tonal peaks from distant ships and boats (Fig. 8). These low SNLD peaks are frequently missed by the original FAV approach due to its peak amplification scheme, where the power transformation disproportionately emphasizes stronger tonal components, thus, suppressing weaker ones, and the subsequent thresholding criteria excludes them.

D. The MFAV sensitivity threshold

We introduced an alternative detrending approach in the MFAV method, where the high-pass pseudo cutoff frequency is a user-defined parameter and can be calibrated to better detect broad tonal noise, typically associated with the presence of a boat. Lowering the high-pass pseudo cutoff frequency from $5\,\mathrm{kHz}^{-1}$ (ship mode) to $2\,\mathrm{kHz}^{-1}$ (boat mode) improved the recall for boat noise detection (Fig. 12); however, this adjustment resulted in a modest 1% decrease in the overall F1-score, highlighting the inherent trade-off between precision and recall during parameter tuning. The high-pass filtered spectrum reveals that, while the broad tonal peaks resulting from the modulating tonals are preserved, numerous spurious peaks are also detected, which reduces the precision of the method. This is most evident in the extreme case when the pseudo cutoff frequency is reduced to $100\,\mathrm{Hz}^{-1}$, as shown in Fig. 10(c). Nevertheless, if maximizing recall is prioritized, such as in applications where missing a detection is more critical than generating false positives, lowering the pseudo cutoff frequency may be advantageous. Otherwise, retaining the default pseudo cutoff frequency of $5\,\mathrm{kHz}^{-1}$ is recommended. Notably, the results also show that reducing the pseudo cutoff frequency to $2\,\mathrm{kHz}^{-1}$ had a negligible impact on ship noise detection by the MFAV method (Table V). Therefore, the pseudo cutoff frequency can be confidently adjusted to better capture boat noise in recall-focused use cases without adversely affecting ship noise detection.

E. Implications for marine conservation in the Canadian Arctic

This study was initiated primarily in recognition of the urgent need to investigate the impacts of increasing ship traffic and associated underwater noise on native Arctic marine mammals, such as narwhals and beluga whales, which have evolved in a characteristically quiet acoustic environment in Canadian Arctic waters (Halliday *et al.*, 2017). These species depend heavily on sound for critical life functions. The ship noise detectors developed and tailored for Arctic conditions in this study can be integrated with automated marine mammal vocalization detectors to assess behavioral responses in the presence of vessel noise. Such integrated monitoring is essential for informing effective mitigation strategies and guiding policymakers in implementing evidence-based regulations to protect Arctic marine ecosystems (Halliday *et al.*, 2017; Halliday *et al.*, 2020a).

While the primary motivation and focus of this study is on conservation in the Canadian Arctic, the proposed methods are broadly applicable and can serve other purposes in maritime surveillance beyond their original scope. These may include the detection of illegal fishing activities (Domingos *et al.*, 2022; Neasham and Lowes, 2021), enhancement of harbor security (Chung *et al.*, 2011; Sorensen *et al.*, 2010), protection of offshore infrastructure (Piskur and Szymak, 2017), and various military and defense applications. The MFAV method, in particular, may be advantageous in scenarios where small vessels, such as fishing or recreational boats, are of interest, thanks to its enhanced sensitivity to weak and/or broad tonal peaks. In addition, method selection should be aligned with the primary objectives of the monitoring effort. For example, in studies examining the impact of elevated underwater noise from large commercial ships on marine mammals, both the MFAV and DL model approaches are suitable. Conversely, in defense or security applications where the presence of a vessel may merely be sought, regardless of its SPL, the DL model may be a more appropriate choice. This is due to its ability to learn complex features, which may not produce high energy but are still indicative of vessel activity. A particular strength of the DL model is its suitability for transfer learning (Yosinski *et al.*, 2014), allowing it to be adapted to new tasks or environments using limited additional training data. It is important to note that while these methods show promise for broader use, they have not yet been validated for these specific applications. Dedicated datasets, additional calibration, and further fine-tuning would be required for operational deployment.

F. Future work

1. Hybrid MFAV-CNN workflow

To merge the physics-based tonal analysis with the spectrogram-based DL approach, a hybrid neural network architecture could be employed. Frazao *et al.* (2025) demonstrated the potential of leveraging alternative acoustic representations of training data to improve model performance. In this regard, one possibility is a dual-branch feature fusion network, where one branch uses a 2D CNN with the spectrogram, and the other branch comprises one-dimensional-convolutional layers for the physics-based tonal features. The learned embeddings can be concatenated at the feature level before being connected.

2. Alternative CNN Architecture

Exploring EfficientNet as an alternative architecture is recommended due to its state-of-the-art computational efficiency and accuracy (Tan and Le, 2019). While ResNet was a reasonable choice for exploring the data and CNN-based

04 August 2026 15:12:24

model development, EfficientNet may offer a more accurate model for the production phase.

3. Advancing CNN Detector

The proposed ship noise detector is currently trained to detect large vessels in the Western Canadian Arctic. A natural next step would be to adapt the model to incorporate boat noise using a transfer learning approach. Subsequently, the model could be extended to classify different vessel types, enhancing its operational utility.

VI. CONCLUSIONS

In this study, we developed and evaluated two automated ship noise detection methods for the Western Canadian Arctic waters: the MFAV algorithm and a ResNet-based CNN model, aiming to advance ship detection with PAM data in this remote region. Our results show that the trained CNN model outperforms both MFAV and FAV, while MFAV consistently exceeds FAV performance.

The MFAV method was introduced as an enhancement of the original FAV algorithm. It provides an intuitive analytical framework and improves the detection of modulating tonal noise from small-engine boats, achieving up to a 22% performance improvement over the original FAV method within our test datasets. MFAV is particularly effective at identifying faint tonal peaks from distant ships or boats, which are often missed by the original FAV due to its peak amplification and thresholding process. We found that lowering the pseudo cutoff frequency parameter (i.e., switching to the "boat mode") in the MFAV method enhances recall but reduces precision for boat noise detection, with minimal impact on ship detection. This offers flexibility for tuning the method based on specific use cases where either precision or recall is prioritized.

The CNN model is specifically trained to detect large ship noise in western Canadian Arctic waters. By incorporating MFAV results into a curriculum learning strategy, the CNN model was able to learn robust features in a structured manner, resulting in strong generalization performance, as evidenced by consistently high F1-scores ranging from 91% to 95%. However, it may need further fine-tuning before applying the model to more acoustically diverse environments. Where the CNN model cannot be used, the MFAV method serves as a robust alternative with negligible compromise in the detection performance.

Further testing is required to confidently correlate environmental characteristics with the performance of the ship detection methods discussed in this paper. While certain cases, such as CP (23–24), show a clear link between high performance and elevated SPL, such relationships are inconclusive. Therefore, larger datasets with sufficient variability are needed to establish statistically significant correlations. These potential correlations could lead to a more informed choice of a ship noise detection method for a specific case study.

ACKNOWLEDGMENTS

This work was funded by Fisheries & Oceans Canada, under its Ecosystems and Oceans Science Contribution Framework (MEQ-2022-23-003). Acoustic data were collected and provided by Wildlife Conservation Society Canada in the Inuvialuit Settlement Region under the authorization of Northwest Territories Research Licences 15996, 16330, 16451, 17308, and 17492 and through funding provided by the Weston Family Foundation, Fisheries and Oceans Canada, and the Fisheries Joint Management Committee. Thank you to S. Insley, A. Kudlak, K. Borg, M. Dempsey, A. Niemi, N. Diogou, A. Hare, N Shajahan, T. Alanak, R. Inuktalik, and J. Kitekudlak for assisting with acoustic data collection.

AUTHOR DECLARATIONS

Conflict of Interest

The authors declare that this research was carried out without any commercial or financial relationships that could be interpreted as a potential conflict of interest.

DATA AVAILABILITY

The resources that support the findings of this study are openly available in GitHub, Ref. (Meridian Analytics, 2026). The datasets supporting this study are available from the authors upon reasonable request, subject to permission from the Wildlife Conservation Society (WCS) Canada.

APPENDIX A—WIND SPEED RECORDS AT ULUKHAKTOK STATION

Figures 11 and 12 display the wind speed record at the Ulukhaktok station overlapped with boat operation timings within the PAS and Ulukhaktok hydrophone catchment, respectively.

APPENDIX B—EFFECT OF PSEUDO CUTOFF FREQUENCY ON MFAV PERFORMANCE

Figure 13 illustrates the recall, precision, and F1-score as functions of the number of peaks detected by the MFAV method, applied to the Ulukhaktok data, containing boat noise only, under both $5\,\mathrm{kHz}^{-1}$ and $2\,\mathrm{kHz}^{-1}$ pseudo cutoff frequency configurations. In "boat mode," recall metric improved at the cost of precision, resulting in a negligible net change in overall performance as measured by the F1-score.

A notable observation in Fig. 13 is the greater variability in the recall curve across different test subsets when the method is applied in "ship mode". Note that, to address the imbalance between positive and negative annotations, the data were subdivided into multiple balanced subgroups. Performance metrics were computed separately for each subgroup and then averaged for reporting. The observed variation in the recall curve reflects the variability among these

04 August 2026 15:12:24

https://doi.org/10.1121/10.0043937

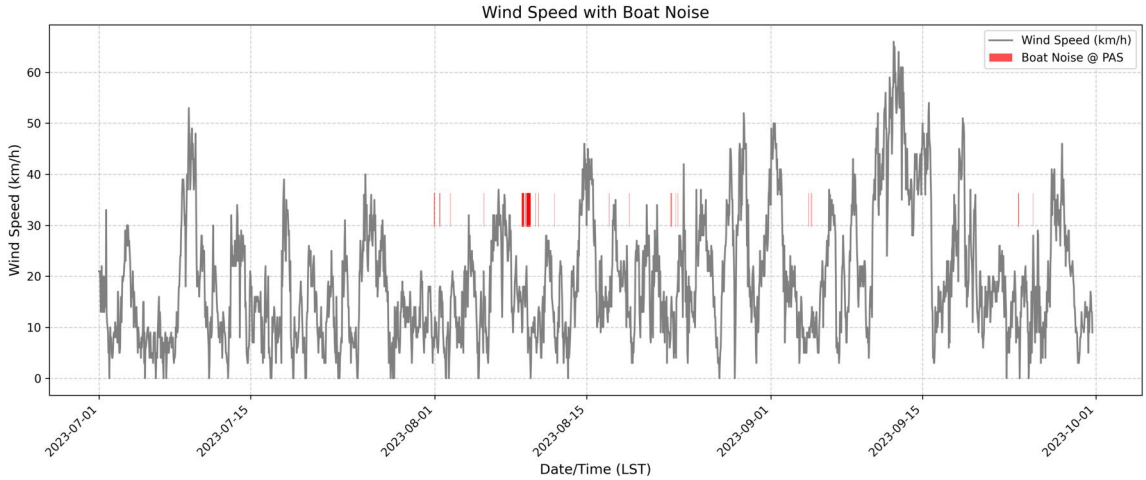

FIG. 11. The wind speed record at the Ulukhaktok station (gray line) overlapped with boat operation timings at the PAS site (red bars).

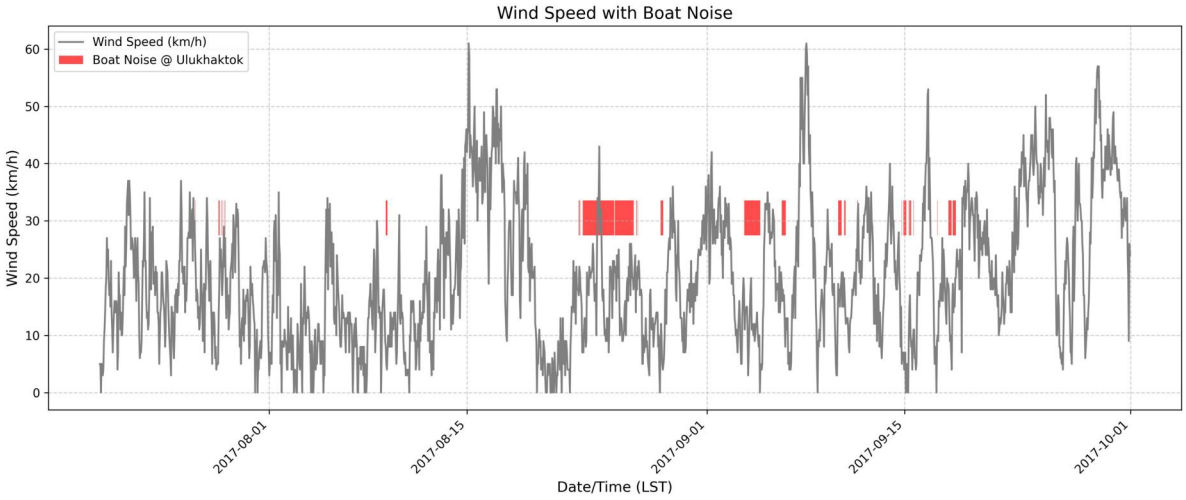

FIG. 12. The wind speed record at the Ulukhaktok station (gray line) overlapped with boat operation timings at the Ulukhaktok site (red bars).

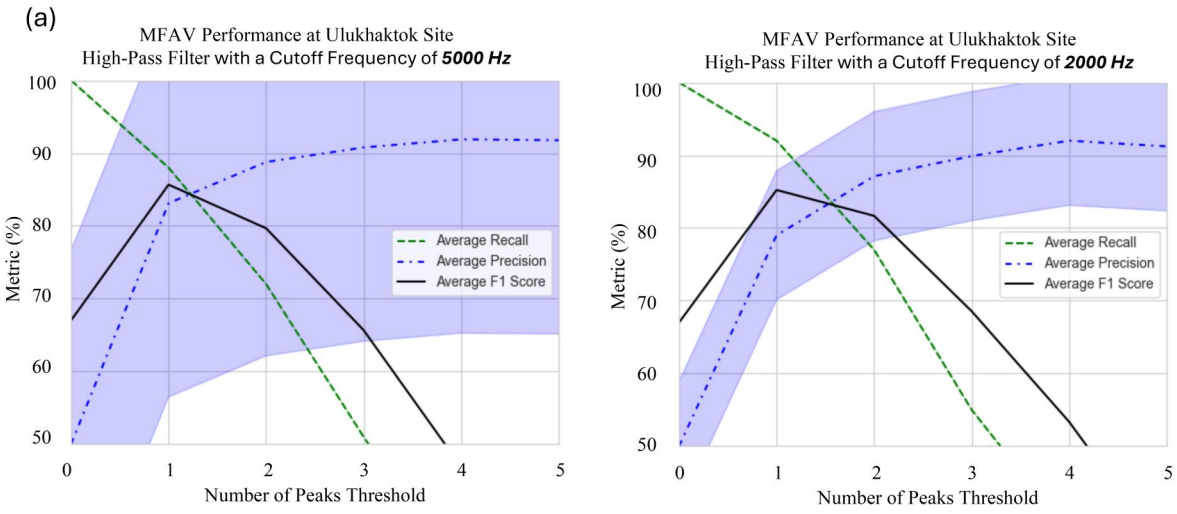

FIG. 13. Performance of the MFAV method for detecting boat noise at the Ulukhaktok site using two different high-pass filter pseudo cutoff frequencies: $2\,\mathrm{kHz}^{-1}$ (right) and $5\,\mathrm{kHz}^{-1}$ (left). Each plot displays the average recall (green dashed line), precision (blue dashed-dotted line), and F1-score (solid black line) as functions of the number of detected tonal peaks. Performance metrics were computed across multiple balanced subgroups and then averaged. Shaded regions of the corresponding colors represent variability across these subgroups for each metric.

04 August 2026 15:12:24

balanced subgroups, which can be attributed to the increased sensitivity of the MFAV method to tonal peaks that may vary in character across subgroups.

Andrew, R. K., Howe, B. M., Mercer, J. A., and Dzieciuch, M. A. (**2002**). "Ocean ambient sound: Comparing the 1960s with the 1990s for a receiver off the California coast," Acoust. Res. Lett. Online **3**(2), 65–70.

Arctic Council (**2019**). "Protection of the Arctic marine environment: Underwater noise in the Arctic—A state of knowledge report," Technical report (Arctic Council, Tromsø, Norway).

Arctic Portal (**2025**). "Arctic portal website," https://www.arcticportal.org (Last viewed August 26, 2025).

Averbuch, A., Zheludev, V., Neittaanmäki, P., Wartiainen, P., Huoman, K., and Janson, K. (**2011**). "Acoustic detection and classification of river boats," Appl. Acoust. **72**(1), 22–34.

Bachman, R. T., and Kaye, G. T. (**1983**). "Broadband interference patterns in shallow water," J. Acoust. Soc. Am. **74**(2), 576–580.

Bengio, Y., Louradour, J., Collobert, R., and Weston, J. (**2009**). "Curriculum learning," in *Proceedings of the 26th Annual International Conference on Machine Learning, ICML '09*, pp. 41–48.

Chapman, N. R., and Price, A. (**2011**). "Low frequency deep ocean ambient noise trend in the northeast Pacific Ocean," J. Acoust. Soc. Am. **129**(5), EL161–EL165.

Chung, K. W., Sutin, A., Sedunov, A., and Bruno, M. (**2011**). "Demon acoustic ship signature measurements in an urban harbor," Adv. Acoust. Vib. **2011**(1), 952798.

Colbert, B. R. (**2020**). "Trends and developments in international regulation of anthropogenic sound in aquatic habitats," J. Acoust. Soc. Am. **147**(5), 3100–3107.

Dawson, J., Pizzolato, L. A. V., Howell, S. E. L., Copland, L., and Johnston, M. E. (**2018**). "Temporal and spatial patterns of ship traffic in the canadian arctic from 1990 to 2015 + supplementary appendix 1: Figs. s1–s7 (see article tools)," Arctic **71**, 15–26–15–26.

Dingwall, J., Halliday, W., Diogou, N., Niemi, A., Steiner, N., and Insley, S. 05 (**2024**). "The Arctic marine soundscape of the Amundsen Gulf, western Canadian Arctic," Mar. Pollut. Bull. **204**, 116510.

Domingos, L. C. F., Santos, P. E., Skelton, P. S. M., Brinkworth, R. S. A., and Sammut, K. (**2022**). "An investigation of preprocessing filters and deep learning methods for vessel type classification with underwater acoustic data," IEEE Access **10**, 117582–117596.

Erbe, C., Dunlop, R., and Dolman, S. (**2018**). "Effects of noise on marine mammals," in *Effects of Anthropogenic Noise on Animals* (Springer, New York), pp. 277–309.

Erbe, C., Marley, S. A., Schoeman, R. P., Smith, J. N., Trigg, L. E., and Embling, C. B. (**2019**). "The effects of ship noise on marine mammals—A review," Front. Mar. Sci. **6**, 2296–7745.

Ferguson, E. L., Ramakrishnan, R., Williams, S. B., and Jin, C. T. (**2017**). "Convolutional neural networks for passive monitoring of a shallow water environment using a single sensor," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2657–2661.

Frazao, F., Joy, R., and Dowd, M. 07 (**2025**). "Comparing acoustic representations for deep learning-based classification of underwater acoustic signals: A case study on orca (*orcinus orca*) vocalizations," Ecol. Inf. **90**, 103297.

Gomez, C., Lawson, J. W., Wright, A. J., Buren, A. D., Tollit, D., and Lesage, V. (**2016**). "A systematic review on the behavioural responses of wild marine mammals to noise: The disparity between science and policy," Can. J. Zoology **94**(12), 801–819.

Halliday, W. D. (**2021**). "Underwater sound levels in the Arctic: Filling knowledge gaps," Geophys. Res. Lett. **48**(15), e2021GL094607, https://doi.org/10.1029/2021GL094607.

Halliday, W. D., Insley, S. J., Hilliard, R. C., de Jong, T., and Pine, M. K. (**2017**). "Potential impacts of shipping noise on marine mammals in the western canadian arctic," Mar. Pollut. Bull. **123**(1), 73–82.

Halliday, W. D., Pine, M. K., and Insley, S. J. (**2020a**). "Underwater noise and arctic marine mammals: Review and policy recommendations," Environ. Rev. **28**(4), 438–448.

Halliday, W. D., Pine, M. K., Mouy, X., Kortsalo, P., Hilliard, C., and Insley, S. (**2020b**). "The coastal Arctic marine soundscape near Ulukhaktok, Northwest Territories, Canada," Polar Biol. **43**, 623–636.

Halliday, W. D., Têtu, P.-L., Dawson, J., Insley, S. J., and Hilliard, R. C. (**2018**). "Tourist vessel traffic in important whale areas in the western Canadian Arctic: Risks and possible management solutions," Mar. Policy **97**, 72–81.

Hanson, D., Brown, G., and Emslie, R. (**2008**). "Cyclostationarity for passive underwater detection of propellor craft: A development of demon processing," in *Proceedings of the Annual Conference of the Australian Acoustical Society, AAS'08*, pp. 360–365.

He, K., Zhang, X., Ren, S., and Sun, J. (**2016**). "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778.

Hildebrand, J. A. (**2009**). "Anthropogenic and natural sources of ambient noise in the ocean," in *Proceedings of the Marine Ecology-Progress Series*, pp. 5–20.

Hummel, H. I., van der Mei, R., and Bhulai, S. (**2024**). "A survey on machine learning in ship radiated noise," Ocean Eng. **298**, 117252.

Jain-Schlaepfer, S., Fakan, E., Rummer, J., Simpson, S., and McCormick, M. (**2018**). "Impact of motorboats on fish embryos depends on engine type," Conserv. Physiol. **6**, coy014.

Ji, Y., Marian, A. D., and Montie, E. W. (**2024**). "Deep-learning-based detection of recreational vessels in an estuarine soundscape in the May River, South Carolina, USA," PLoS One **19**(7), e0302497.

Jiang, J-j., Bu, L-r., Duan, F-j., Wang, X-q., Liu, W., Sun, Z-b., and Li, C-y (**2019**). "Whistle detection and classification for whales based on convolutional neural networks," Appl. Acoust. **150**, 169–178.

Khan, A., Sohail, A., Zahoora, U., and Qureshi, A. S. (**2020**). "A survey of the recent architectures of deep convolutional neural networks," Artif. Intell. Rev. **53**, 5455–5516.

Kim, S., Jung, S-k., Kang, D., Kim, M., and Cho, S. (**2020**). "Application of the artificial intelligence for automatic detection of shipping noise in shallow-water," J. Acoust. Soc. Korea **39**, 279–285.

Kingma, D. P., and Ba, J. (**2014**). "Adam: A method for stochastic optimization," arXiv:math-ph/1412.6980.

Kudryavtsev, A. A., Luginets, K., and Mashoshin, A. I. (**2003**). "Amplitude modulation of underwater noise produced by seagoing vessels," Acoust. Phys. **49**, 184–188.

Leal, N., Leal, E., and Sanchez, G. (**2015**). "Marine vessel recognition by acoustic signature," ARPN J. Eng. Appl. Sci. **10**, 9633–9639.

LeCun, Y., Bengio, Y., and Hinton, G. (**2015**). "Deep learning," Nature **521**(7553), 436–444.

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (**1989**). "Backpropagation applied to handwritten zip code recognition," Neural Comput. **1**(4), 541–551.

Li, G., Liu, F., and Yang, H. (**2024**). "A new detection model of ship-radiated noise signal," Ocean Eng. **297**, 117081.

Matsumoto, H., Bohnenstiehl, D. R., Tournadre, J., Dziak, R. P., Haxel, J. H., Lau, T.-K. A., Fowler, M., and Salo, S. A. (**2014**). "Antarctic icebergs: A significant natural ocean sound source in the southern hemisphere," Geochem. Geophys. Geosyst. **15**(8), 3448–3458.

McCauley, R. D., Day, R. D., Swadling, K. M., Fitzgibbon, Q. P., Watson, R. A., and Semmens, J. M. (**2017**). "Widely used marine seismic survey air gun operations negatively impact zooplankton," Nat. Ecology Evol. **1**(7), 0195.

McDonald, M. A., Hildebrand, J. A., and Wiggins, S. M. (**2006**). "Increases in deep ocean ambient noise in the northeast Pacific west of San Nicolas Island, California," J. Acoust. Soc. Am. **120**(2), 711–718.

McDonald, M. A., Hildebrand, J. A., Wiggins, S. M., and Ross, D. (**2008**). "A 50 year comparison of ambient ocean noise near San Clemente Island: A bathymetrically complex coastal region off southern California," J. Acoust. Soc. Am. **124**(4), 1985–1992.

McDonald, R. A., Schultheiss, P. M., Tuteur, F. B., and Usher, T. (**1963**). "Processing of data from sonar systems," Report No. U417-63-045 (General Dynamics, Groton, CT).

McWhinnie, L. H., Halliday, W. D., Insley, S. J., Hilliard, C., and Canessa, R. R. (**2018**). "Vessel traffic in the Canadian Arctic: Management solutions for minimizing impacts on whales in a changing northern region," Ocean Coastal Manage. **160**, 1–17.

Merchant, N. D., Pirotta, E., Barton, T. R., and Thompson, P. M. (**2014**). "Monitoring ship noise to assess the impact of coastal developments on marine mammals," Mar. Pollut. Bull. **78**(1), 85–95.

Meridian Analytics (**2026**). "Vessel-Noise-Detector Tool (VND)," GitHub. https://github.com/meridian-analytics/VND.

04 August 2026 15:12:24

Michalski, R. S. (**1983**). "A theory and methodology of inductive learning," in *Machine Learning* (Elsevier, Amsterdam), pp. 83–134.
Miksis-Olds, J. L., Bradley, D. L., and Maggie Niu, X. (**2013**). "Decadal trends in indian ocean ambient sound," J. Acoust. Soc. Am. **134**(5), 3464–3475.
Miksis-Olds, J. L., and Nichols, S. M. (**2016**). "Is low frequency ocean sound increasing globally?," J. Acoust. Soc. Am. **139**(1), 501–511.
Neasham, J., and Lowes, G. (**2021**). "Impact of United Kingdom lockdown on subsea noise from vessel propulsion," J. Acoust. Soc. Am. **149**(4), A28–A28.
Piskur, P., and Szymak, P. (**2017**). "Algorithms for passive detection of moving vessels in marine environment," J. Mar. Eng. Technol. **16**(4), 377–385.
Pizzolato, L., Howell, S. E. L., Dawson, J., Laliberté, F., and Copland, L. (**2016**). "The influence of declining sea ice on shipping activity in the Canadian Arctic," Geophys. Res. Lett. **43**(23), 12–146, https://doi.org/10.1002/2016GL071489.
Pollara, A., Sutin, A., and Salloum, H. (**2016**). "Improvement of the detection of envelope modulation on noise (demon) and its application to small boats," in *Proceedings of the OCEANS 2016 MTS/IEEE Monterey*, pp. 1–10.
Pollara, A., Sutin, A., and Salloum, H. (**2017**). "Passive acoustic methods of small boat detection, tracking and classification," in *Proceedings of the 2017 IEEE International Symposium on Technologies for Homeland Security (HST)*, pp. 1–6.
Prasad, S. K. R., and Gurugopinath, S. (**2022**). "Supervised learning techniques for detection of underwater acoustic sources," in *Proceedings of the 2022 2nd International Conference on Intelligent Technologies (CONIT)*, pp. 1–7.
Reis, C. D. G., Rodrigues Padovese, L., and de Oliveira, M. C. F. (**2019**). "Automatic detection of vessel signatures in audio recordings with spectral amplitude variation signature," Methods Ecol. Evol. **10**(9), 1501–1516.
Rolland, R. M., Parks, S. E., Hunt, K. E., Castellote, M., Corkeron, P. J., Nowacek, D. P., Wasser, S. K., and Kraus, S. D. (**2012**). "Evidence that ship noise increases stress in right whales," Proc. R. Soc. B: Biol. Sci. **279**(1737), 2363–2368.
Ross, D. (**2005**). "Ship sources of ambient noise," IEEE J. Oceanic Eng. **30**(2), 257–261.
Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (**2017**). "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 618–626.
Simonyan, K., and Zisserman, A. (**2014**). "Very deep convolutional networks for large-scale image recognition," arXiv:math-ph/1409.1556.
Slamnoiu, G., Radu, O., Roşca, V., Pascu, C., Damian, R. G., Surdu, G., Curcă, E., and Rădulescu, A. (**2016**). "Demon-type algorithms for determination of hydro-acoustic signatures of surface ships and of divers," IOP Conf. Ser.: Mater. Sci. Eng. **145**, 082013.
Smith, T. A., and Rigby, J. (**2022**). "Underwater radiated noise from marine vessels: A review of noise reduction methods and technology," Ocean Eng. **266**, 112863.
Sorensen, E., Ou, H. H., Zurk, L. M., and Siderius, M. (**2010**). "Passive acoustic sensing for detection of small vessels," in *Proceedings of OCEANS 2010 MTS/IEEE SEATTLE*, pp. 1–8.
Southall, B. L., Finneran, J. J., Reichmuth, C., Nachtigall, P. E., Ketten, D. R., Bowles, A. E., Ellison, W. T., Nowacek, D. P., and Tyack, P. L. (**2019**). "Marine mammal noise exposure criteria: Updated scientific recommendations for residual hearing effects," Aquatic Mammals **45**(2), 125–232.
Stanistreet, J., Beslin, W., Kowarski, K., Martin, B., Westell, A., and Moors-Murphy, H. (**2022**). "Changes in the acoustic activity of beaked whales and sperm whales recorded during a naval training exercise off eastern Canada," Sci. Rep. **12**, 1973.
Stephenson, S. R., Smith, L. C., and Agnew, J. A. (**2011**). "Divergent long-term trajectories of human access to the Arctic," Nat. Clim. Change **1**(3), 156–160.
Stroeve, J., Holland, M. M., Meier, W., Scambos, T., and Serreze, M. (**2007**). "Arctic sea ice decline: Faster than forecast," Geophys. Res. Lett. **34**(9), L09501, https://doi.org/10.1029/2007GL029703.
Stroeve, J., and Notz, D. (**2018**). "Changing state of arctic sea ice across all seasons," Environ. Res. Lett. **13**, 103001.
Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (**2015**). "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9.
Tan, M., and Le, Q. V. (**2019**). "Efficientnet: Rethinking model scaling for convolutional neural networks," arXiv:math-ph/1905.11946.
Topple, J., Thomson, D., Bates, J., Xu, J., Webstey, A., and Binder, C. (**2024**). "Passive acoustic monitoring of arctic surface ships with deep-pam," in *Proceedings of Oceans 2024*, pp. 1–6.
Urick, R. J. (**1983**). *Principles of Underwater Sound*, 3rd ed. (McGraw-Hill, New York).
Vieira, M., Amorim, M. C. P., Sundelöf, A., Prista, N., and Fonseca, P. J. (**2019**). "Underwater noise recognition of marine vessel passages: Two case studies using hidden Markov models," ICES J. Mar. Sci. **77**(6), 2157–2170.
Wang, X., Chen, Y., and Zhu, W. (**2022**). "A survey on curriculum learning," IEEE Trans. Pattern Anal. Mach. Intell. **44**(9), 4555–4576.
White, E. L., White, P. R., Bull, J. M., Risch, D., Beck, S., and Edwards, E. W. J. (**2022**). "More than a whistle: Automated detection of marine sound sources with a convolutional neural network," Front. Mar. Sci. **9**, 879145.
Wittekind, D., and Schuster, M. (**2016**). "Propeller cavitation noise and background noise in the sea," Ocean Eng. **120**, 116–121.
Yan, J., Sun, H., Cheng, E., Kuai, X., and Zhang, X. (**2017**). "Ship radiated noise recognition using resonance-based sparse signal decomposition," Shock Vib. **2017**, 1–6930605.
Yang, L., and Chen, K. (**2017**). "Performance comparison of two types of auditory perceptual features in robust underwater target classification," Acta Acust. united Ac. **103**(1), 56–66.
Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (**2014**). "How transferable are features in deep neural networks?," arXiv:math-ph/1411.1792.
Zak, A. (**2008**). "Ships classification basing on acoustic signatures," WSEAS Trans. Sig. Proc. **4**(4), 137–149.

04 August 2026 15:12:24