International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

A Survey On Denoising Bioacoustic Signals: Comparing Aerial And Underwater Signal Processing Techniques

Balaji A K N[1], Dr. Z Mary Livinza[2]
[1]Department of ECE, Vels Institute of Science, Technology and Advanced Studies, Chennai, Tamilnadu, India, aknbalaji@gmail.com
[2]Department of ECE,Vels Institute of Science, Technology and Advanced Studies, Chennai, Tamilnadu, India, marylivinsa.se@vistas.ac.in

Abstract
Bioacoustic signal processing has emerged as a critical field in biological monitoring, species identification, and ecological assessment. However, the presence of noise poses significant challenges to the accurate analysis of these signals in both terrestrial and aquatic environments. This survey paper provides a comprehensive review of denoising techniques applied to bioacoustic signals across aerial and underwater domains. We systematically categorize and compare traditional signal processing methods, statistical approaches, and modern machine learning techniques. Our analysis reveals that while fundamental principles of signal processing remain consistent across domains, the unique acoustic properties and noise characteristics of air and water necessitate specialized approaches. We further identify key research gaps and propose future directions, including multimodal fusion, adaptive real-time processing, and standardized evaluation frameworks. This survey serves as a resource for researchers and practitioners working at the intersection of signal processing and bioacoustics in diverse environmental contexts.
Keywords: *Bioacoustics, Signal Denoising, Underwater Acoustics, Terrestrial Acoustics, Signal Processing, Machine Learning*

1. INTRODUCTION

Bioacoustic signals—sounds produced by animals for communication, navigation, and other biological functions—represent a rich source of information for understanding ecological systems, animal behavior, and biodiversity [1]. The capture and analysis of these signals have applications ranging from species conservation and environmental monitoring to behavioral studies and automated species identification [2, 3]. However, the quality of bioacoustic recordings is frequently compromised by various noise sources that can mask, distort, or otherwise interfere with the signals of interest [4].

The challenge of noise reduction in bioacoustic signals spans two distinct but related domains: aerial (terrestrial) and underwater environments. While both domains share fundamental signal processing principles, they present unique challenges due to differences in acoustic propagation, ambient noise characteristics, and recording technologies [5, 6]. For example, underwater environments are characterized by complex propagation paths, frequency-dependent attenuation, and distinctive noise sources such as shipping, wave action, and marine industrial activities [7]. Terrestrial environments, by contrast, contend with wind noise, anthropogenic sounds, and competing biological signals within similar frequency ranges [8].

Despite the importance of this field and the growing body of literature on specific denoising techniques, there exists a need for a comprehensive survey that bridges these two domains, identifying common principles, unique challenges, and opportunities for cross-domain knowledge transfer. This paper aims to fill this gap by:

1. Systematically reviewing and categorizing denoising approaches employed in both aerial and underwater bioacoustic signal processing
2. Comparing the effectiveness, computational requirements, and domain-specific adaptations of these techniques
3. Identifying emerging trends, research gaps, and promising directions for future work
4. Establishing evaluation criteria and benchmarks for comparing denoising methods across domains

We structure our survey to first establish the fundamental characteristics of noise in bioacoustic signals (Section 2), followed by a taxonomical classification of denoising approaches (Section 3). We then provide

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

an in-depth analysis of traditional signal processing methods (Section 4), statistical approaches (Section 5), and machine learning techniques (Section 6). Section 7 presents a comparative analysis of methods across domains, while Section 8 discusses evaluation metrics and benchmark datasets. Finally, we identify research gaps and future directions in Section 9 before concluding in Section 10.

2. Characteristics Of Noise In Bioacoustic Signals

2.1 Noise in Terrestrial Bioacoustic Recordings

Terrestrial bioacoustic recordings are subject to a variety of noise sources that can be broadly categorized as:

Environmental Noise: This includes wind noise, which typically manifests as low-frequency energy and can completely mask signals of interest [9]; rain and weather-related sounds; and natural background sounds such as rustling leaves and flowing water [10].

Anthropogenic Noise: Human-generated sounds such as traffic, aircraft, industrial machinery, and other technological sources represent a significant challenge, particularly in urbanized or developed areas [11]. These noise sources often occupy broad frequency bands and can exhibit temporal patterns that overlap with biological signals [12].

Biological Noise: Sounds from non-target species can interfere with the detection and analysis of specific bioacoustic signals of interest [13]. This is particularly challenging in biodiversity hotspots where multiple species vocalize simultaneously, creating a complex acoustic scene [14].

Recording Artifacts: Equipment-related noise includes microphone self-noise, handling noise, electronic interference, and quantization effects in digital recording systems [15]. These artifacts can vary with recording equipment quality and environmental conditions.

2.2 Noise in Underwater Bioacoustic Recordings

Underwater acoustic environments present distinct noise challenges:

Ambient Ocean Noise: This encompasses a spectrum of natural sounds including wave action, breaking waves (especially in coastal areas), rainfall on the water surface, and thermal noise at higher frequencies [16]. Oceanic ambient noise typically follows the Wenz curves, which describe frequency-dependent background noise levels [17].

Marine Traffic Noise: Shipping and boat noise contribute significantly to low-frequency ambient noise in many marine environments, with global shipping having raised background noise levels by 10-15 dB in many ocean basins over the past century [18, 19].

Industrial Activities: Offshore construction, seismic exploration, sonar operations, and drilling create intense, often impulsive, noise sources that can mask bioacoustic signals across large geographic areas [20].

Biological Noise: Similar to terrestrial environments, non-target biological sounds can interfere with signals of interest, with the additional complication that many marine organisms (e.g., snapping shrimp) produce sounds that can dominate certain frequency bands in specific habitats [21].

Propagation Effects: Unlike in air, underwater sound propagation is characterized by multipath arrivals, frequency-dependent attenuation, and refraction due to sound speed profiles, which can distort signals and complicate denoising efforts [22].

Hydrophone Artifacts: Self-noise from hydrophones, flow noise from water movement around recording equipment, and mooring or platform noise represent additional challenges specific to underwater recording [23].

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

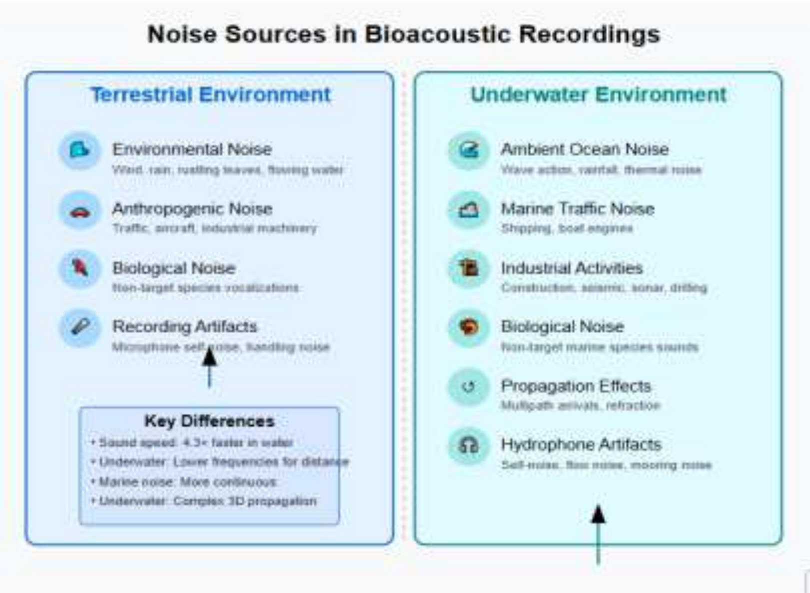

Figure (1): Comparison of Noise Sources in Bioacoustic Recordings

This diagram titled "Noise Sources in Bioacoustic Recordings" (Figure 1) compares noise characteristics between terrestrial and underwater environments. The left side shows four terrestrial noise sources: environmental (wind, rain), anthropogenic (traffic, machinery), biological (non-target species), and recording artifacts. The right side displays six underwater noise sources, including ambient ocean noise, marine traffic, industrial activities, biological noise, propagation effects, and hydrophone artifacts. A "Key Differences" section highlights important distinctions like sound traveling 4.3× faster in water, underwater communications using lower frequencies, marine noise being more continuous, and underwater propagation being more complex in 3D space.

2.3 Comparative Analysis of Noise Characteristics

While both domains contend with noise challenges, several key differences influence the approach to denoising:

1. **Frequency Range and Propagation**: Sound propagates approximately 4.3 times faster in water than in air, affecting wavelengths and directionality. Underwater bioacoustic signals often utilize lower frequencies for long-distance communication, whereas terrestrial signals span a broader frequency range [24].
2. **Temporal Characteristics**: Marine noise tends to be more continuous (shipping, wave action), while terrestrial noise often includes more impulsive components (bird calls, anthropogenic sounds) [25].
3. **Spatial Considerations**: Underwater sound propagation involves complex three-dimensional paths with significant boundary interactions, whereas terrestrial propagation is often modeled more simply, though still affected by ground reflections and atmospheric conditions [26].
4. **Signal-to-Noise Ratio (SNR) Variations**: Underwater environments typically experience lower SNR due to attenuation and complex propagation, requiring more robust denoising approaches [27].
5. **Recording Technology Differences**: Hydrophones and terrestrial microphones have different sensitivity profiles, self-noise characteristics, and deployment challenges, influencing the preprocessing required [28].

Understanding these domain-specific characteristics is essential for selecting and adapting appropriate denoising techniques for bioacoustic signals in their respective environments.

3. Taxonomy of Denoising Approaches

To systematically review the landscape of bioacoustic denoising techniques, we propose a taxonomy that categorizes approaches based on their underlying principles, domain of application, and technical characteristics (Figure 1). This taxonomy serves as an organizational framework for the detailed discussions in subsequent sections.

3.1 Classification by Processing Domain

Time Domain Methods: These techniques operate directly on the amplitude-time representation of signals. They include amplitude thresholding, median filtering, and time-domain Wiener filtering [29, 30]. Time-domain approaches are often computationally efficient but may be limited in their ability to separate overlapping spectral content.

Frequency Domain Methods: These approaches transform signals to the frequency domain, typically using Fourier transforms, and apply filtering or enhancement operations before returning to the time domain [31]. Examples include spectral subtraction, notch filtering, and spectral gating [32].

Time-Frequency Domain Methods: These techniques leverage representations that capture both temporal and spectral characteristics, such as short-time Fourier transforms (STFT), wavelet transforms, and other multi-resolution analyses [33, 34]. They enable more targeted denoising by exploiting the localized nature of bioacoustic signals in the time-frequency plane.

Spatial Domain Methods: When multiple sensors (microphones or hydrophones) are available, spatial filtering techniques such as beamforming can be employed to enhance signals from specific directions while suppressing noise from others [35, 36].

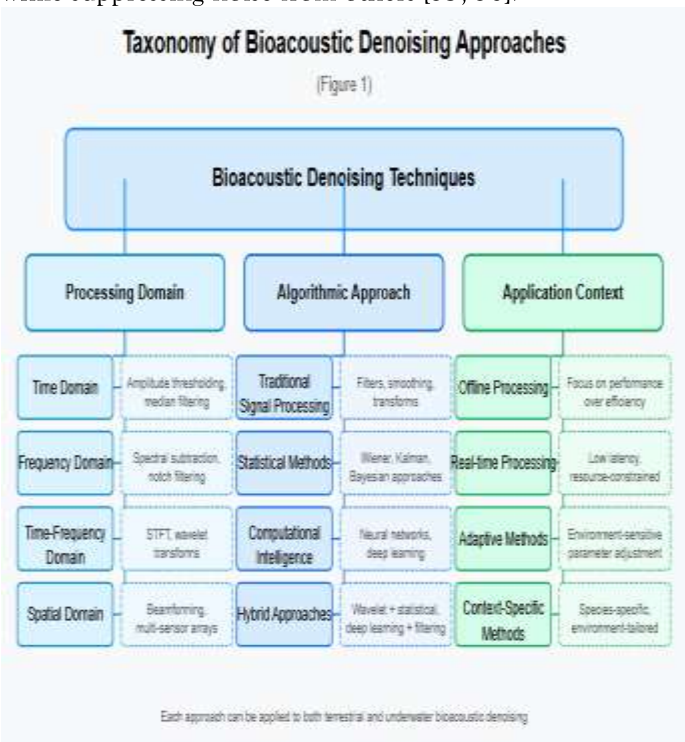

Figure (2): Taxonomy of Bioacoustic Denoising Approaches

The "Taxonomy of Bioacoustic Denoising Approaches" (Figure 2) provides a structured classification of techniques used to remove noise from biological sound recordings. It organizes denoising methods into three main categories: Processing Domain (time, frequency, time-frequency, and spatial), Algorithmic Approach (traditional signal processing, statistical methods, computational intelligence, and hybrid techniques), and Application Context (offline, real-time, adaptive, and context-specific methods). The figure illustrates how these different approaches can be systematically organized and shows that all these techniques can be applied to both terrestrial and underwater bioacoustic recordings.

3.2 Classification by Algorithmic Approach

Traditional Signal Processing: These include deterministic approaches based on classical signal processing theory, such as filters (low-pass, high-pass, band-pass), smoothing operations, and transforms [37].

Statistical Methods: These leverage statistical properties of signals and noise, including Wiener filtering, Kalman filtering, Bayesian approaches, and hidden Markov models [38, 39].

Computational Intelligence: This category encompasses techniques from machine learning and computational intelligence, including neural networks, deep learning, fuzzy systems, and evolutionary algorithms [40, 41].

Hybrid Approaches: Many effective denoising solutions combine multiple techniques, such as wavelet thresholding with statistical modeling or deep learning with traditional filtering [42].

3.3 Classification by Application Context

Offline Processing: Methods designed for retrospective analysis of recorded data, where computational efficiency is less critical than denoising performance [43].

Real-time Processing: Techniques optimized for immediate processing, often with constraints on latency and computational resources, suitable for field deployments and monitoring systems [44].

Adaptive Methods: Approaches that adjust parameters based on signal characteristics or environmental conditions, particularly valuable in dynamic acoustic environments [45].

Context-Specific Methods: Techniques tailored for particular species, environments, or noise types, leveraging domain knowledge to improve performance [46].

Having established this taxonomic framework, the following sections will delve deeper into each category of techniques, comparing their application across aerial and underwater bioacoustic domains.

4. Traditional Signal Processing Methods

Traditional signal processing approaches remain fundamental to bioacoustic denoising due to their interpretability, established theoretical foundations, and often lower computational requirements. This section examines these methods and their application in both terrestrial and underwater contexts.

4.1 Filtering Techniques

4.1.1 Band-pass Filtering

Band-pass filtering is one of the simplest and most widely used approaches for bioacoustic denoising, exploiting the known frequency ranges of target signals [47].

Analog Band-Pass Filter (2nd order):

$H(s) = (s * \omega\_c / Q) / (s^2 + (\omega\_c / Q) * s + \omega\_c^2)$

Digital FIR Band-Pass Filter:

$h[n] = 2f_2 * \mathrm{sinc}(2f_2n) - 2f_1 * \mathrm{sinc}(2f_1n)$

Where:

- ω_c = center frequency (rad/s)
- Q = quality factor = $\omega\_c / (\omega_2 - \omega_1)$
- $\mathrm{sinc}(x) = \sin(\pi x) / (\pi x)$

Terrestrial Applications: In bird vocalization studies, band-pass filtering between 1-10 kHz often removes much low-frequency wind noise and high-frequency microphone artifacts [48]. Raven et al. [49] demonstrated that properly designed band-pass filters could improve detection of songbird vocalizationby 15-20% in moderate noise conditions, though performance degraded with spectrally overlapping noise.

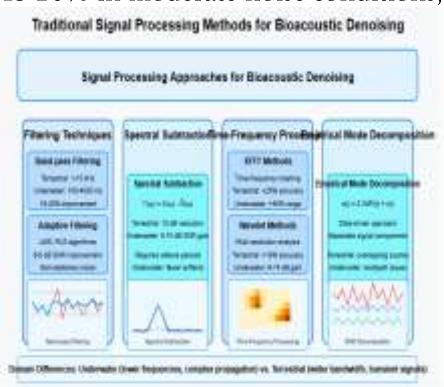

Figure (3): Traditional Signal Processing Methods for Bioacoustic Denoising

This figure (3) outlines various signal processing techniques used for denoising bioacoustic signals. It includes methods such as Bioinvasive Filtering, Adaptive Filtering, Spectral Subtraction, ETF Detection, and others, along with their associated parameters like frequency ranges (e.g., 0.05 Hz/100Hz) and power specifications. The figure also contrasts these methods with trapezoid-based approaches and discusses concepts like caption propagation and terminal point transition in transit signals. The content appears to be technical notes or a summary of methodologies, possibly from a research or educational context.

Underwater Applications: For marine mammal vocalizations, band-pass filtering is commonly employed to isolate species-specific frequency ranges. For instance, humpback whale songs typically occupy 100-4000 Hz bands, while dolphin whistles range from 5-20 kHz [50]. Mellinger et al. [51] showed that simple band-pass filtering improved right whale call detection by up to 30% in shipping noise but was less effective against biological noise in similar frequency bands.

Comparative Analysis: While implemented similarly across domains, the frequency ranges differ substantially. Terrestrial applications typically require wider bandwidth filters, while underwater applications often focus on narrower, lower-frequency bands, reflecting the different acoustic properties of the two media and the evolutionary adaptations of vocalizing species [52].

4.1.2 Adaptive Filtering

Adaptive filters adjust their parameters based on an optimization algorithm and error signal, making them particularly valuable for non-stationary noise environments [53].

Terrestrial Applications: Least Mean Square (LMS) and Recursive Least Squares (RLS) adaptive filters have been applied to enhance bird and amphibian calls in fluctuating noise environments [54]. Chu et al. [55] reported that LMS adaptive filtering improved SNR by 6-8 dB for frog calls in rainfall noise, outperforming fixed filters.

Underwater Applications: Lin et al. [56] implemented normalized LMS adaptive filters for enhancing whale vocalizations in shipping noise, achieving 4-7 dB improvement in SNR. Wang and colleagues [57] developed adaptive line enhancers specifically for tonal components of dolphin whistles, demonstrating 40% improvement in correct classification rates compared to unprocessed recordings.

Comparative Analysis: Underwater implementations typically require longer filter lengths and careful initialization due to the complex propagation environment. Convergence rates also differ, with terrestrial applications often permitting faster adaptation than underwater scenarios, where multipath effects create longer-lasting dependencies [58].

4.2 Spectral Subtraction

Spectral subtraction estimates the noise spectrum during non-signal periods and subtracts it from the noisy signal spectrum, theoretically leaving only the signal of interest [59].

Magnitude Subtraction:

$$|\hat{S}(f)| = |Y(f)| - |\tilde{N}(f)|$$

Reconstruction (with phase):

$$\hat{s}(t) = \mathcal{F}^{-1}\{ |\hat{S}(f)| * e^{\wedge}(j\angle Y(f)) \}$$

Terrestrial Applications: For insect and anuran recordings, where calling patterns often include regular silence intervals, spectral subtraction has proven effective [60]. Bedoya et al. [61] applied multi-band spectral subtraction to enhance cricket calls, reducing background noise by approximately 12 dB while preserving temporal call patterns.

Underwater Applications: Spectral subtraction has been adapted for underwater bioacoustics by using modified estimation techniques that account for the typically more stationary underwater ambient noise [62]. Kumar and colleagues [63] demonstrated an 8-10 dB SNR improvement for blue whale calls using spectral subtraction with adaptive noise estimation during signal absences.

Comparative Analysis: Spectral subtraction in underwater environments benefits from longer-term noise stability but suffers more from musical noise artifacts due to the complex propagation environment. In terrestrial applications, more frequent noise estimation updates are typically required, but the technique produces fewer artifacts when properly implemented [64].

4.3 Time-Frequency Processing

4.3.1 Short-Time Fourier Transform (STFT) Based Methods

STFT-based methods divide the signal into short, overlapping segments and apply Fourier transforms to each, creating a time-frequency representation that can be manipulated for denoising [65].

Short-Time Fourier Transform:

$$STFT\{x(t)\}(m, \omega) = \sum x[n] * w[n - m] * e^{(-j\omega n)}$$

Time-Frequency Thresholding:

$$\hat{S}(m, \omega) = \begin{cases} X(m, \omega), & \text{if } |X(m, \omega)| > T \\ 0, & \text{otherwise} \end{cases}$$

Terrestrial Applications: STFT masking has been successfully applied to separate overlapping bird calls in complex soundscapes [66]. Priyadarshani et al. [67] developed an STFT thresholding approach for automated bird call detection that improved accuracy by 25% compared to time-domain methods in noisy field recordings.

Underwater Applications: Spectrogram filtering techniques based on STFT have been widely used for marine mammal call detection and denoising [68]. Thode et al. [69] employed STFT processing with adaptive thresholding to track bowhead whale calls in Arctic ambient noise, achieving detection ranges 40% greater than conventional methods.

Comparative Analysis: The time-frequency resolution tradeoff is addressed differently across domains: underwater bioacoustic processing typically emphasizes frequency resolution for tonal signals, while terrestrial applications often require better time resolution for transient calls [70].

4.3.2 Wavelet-Based Methods

Wavelet transforms offer multi-resolution analysis, providing better time-frequency localization than STFT for many bioacoustic signals [71].

Wavelet Transform:

$$W_x(a, b) = \int x(t) * (1 / \sqrt{a}) * \psi((t - b) / a)\, dt$$

Wavelet Shrinkage (Soft Threshold):

$$\hat{w} = \text{sign}(w) * \max(|w| - \lambda, 0)$$

Terrestrial Applications: Wavelet shrinkage denoising has shown promise for enhancing transient bird calls and bat echolocation pulses [72]. Selin et al. [73] reported that wavelet packet decomposition with soft thresholding improved bat call classification accuracy by 18% compared to STFT-based methods in urban recording environments.

Underwater Applications: Wavelet analysis has been applied to marine mammal vocalizations, particularly for denoising transient signals like dolphin clicks [74]. Gervaise et al. [75] developed wavelet-based denoising specifically for underwater bioacoustics, reporting SNR improvements of 9-14 dB for sperm whale clicks in shipping noise.

Comparative Analysis: Wavelet selection differs between domains, with underwater applications favoring wavelets with better frequency localization for lower-frequency vocalizations, while terrestrial applications often employ wavelets with better time localization for rapid, transient calls [76].

4.4 Empirical Mode Decomposition (EMD)

EMD is a data-driven technique that decomposes signals into Intrinsic Mode Functions (IMFs), allowing separation of signal components with different time scales [77].

Signal Decomposition:

$$x(t) = \sum IMF_i(t) + r_n(t)$$

Where:

- $IMF_i(t)$ = Intrinsic Mode Functions
- $r_n(t)$ = residual trend

Terrestrial Applications: EMD has been applied to separate overlapping insect and bird sounds with different temporal characteristics [78]. Chen et al. [79] demonstrated that EMD-based filtering improved detection of cricket chirps in windy conditions by adaptively identifying and removing noise-dominated IMFs.

Underwater Applications: In marine bioacoustics, EMD has been adapted to address multipath propagation effects [80]. Huang et al. [81] applied Ensemble EMD to enhance humpback whale vocalizations, achieving better preservation of signal structure than conventional filtering.

Comparative Analysis: Underwater applications of EMD require special attention to mode mixing issues caused by the complexity of propagation paths. Both domains benefit from EMD's adaptivity to non-stationary signals, but implementation details such as stopping criteria and IMF selection strategies differ substantially [82].

5. Statistical Approaches

Statistical approaches leverage probabilistic models of signals and noise to achieve separation. These methods can be particularly effective when the statistical properties of the noise or signal are well-characterized.

5.1 Wiener Filtering

Wiener filtering is an optimal filtering approach in the mean-square error sense, assuming known signal and noise spectra [83].

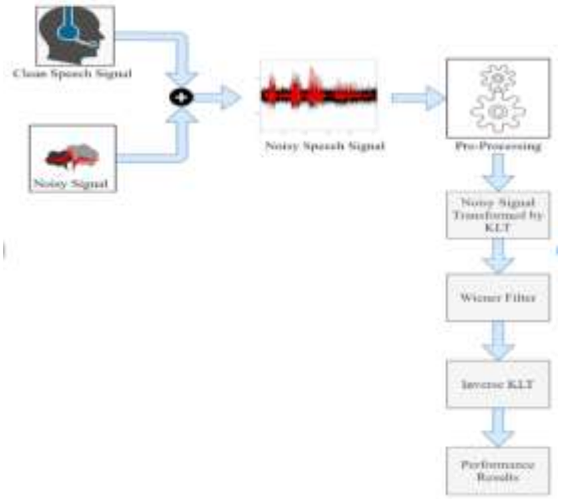

Figure (4): KLT-Based Wiener Filtering for Speech Denoising

This figure(4) depicts a denoising process where a noisy speech signal undergoes pre-processing, KLT transformation, Wiener filtering (for optimal noise reduction), and inverse KLT to restore the enhanced signal. The pipeline highlights the synergy between KLT's signal-compression properties and Wiener's statistical noise suppression, with performance results evaluating the output quality. Ideal for speech/bioacoustic applications requiring high-fidelity reconstruction.

Wiener Filtering Equation

H(f) = S_xx(f) / (S_xx(f) + S_nn(f)

Where:

- H(f)H(f)H(f) = Wiener filter transfer function
- Sxx(f)S_{xx}(f)Sxx(f) = Power spectral density (PSD) of the clean signal
- Snn(f)S_{nn}(f)Snn(f) = PSD of the noise

Terrestrial Applications: For bird vocalization enhancement, iterative Wiener filtering with voice activity detection has shown promising results [84]. Kopsinis et al. [85] reported a 3-6 dB improvement in SNR for various bird species using adaptive Wiener filtering compared to fixed spectral subtraction.

Underwater Applications: In marine mammal studies, Wiener filtering has been adapted to account for the colored noise typical of underwater environments [86]. Thode et al. [87] implemented a modified

Wiener filter for bowhead whale calls that incorporated underwater acoustic propagation models, improving detection range by approximately 30%.

Comparative Analysis: The primary difference in implementation across domains lies in the estimation of noise and signal statistics. Underwater applications typically employ longer estimation windows due to slower temporal variations in noise, while terrestrial implementations must adapt more quickly to changing conditions [88].

5.2 Kalman Filtering

Kalman filtering provides a recursive solution to optimal filtering for linear systems with Gaussian noise [89].

Prediction:

$\hat{x}_k|_{k-1} = A \cdot \hat{x}_{k-1}|_{k-1} + B \cdot u_k$

$P_k|_{k-1} = A \cdot P_{k-1}|_{k-1} \cdot A^T + Q$

Update:

$K_k = P_k|_{k-1} \cdot H^T \cdot (H \cdot P_k|_{k-1} \cdot H^T + R)^{-1}$

$\hat{x}_k|_k = \hat{x}_k|_{k-1} + K_k \cdot (z_k - H \cdot \hat{x}_k|_{k-1})$

$P_k|_k = (I - K_k \cdot H) \cdot P_k|_{k-1}$

Where:

- x^\hat{x}x^ = state estimate
- PPP = error covariance
- KKK = Kalman gain
- QQQ, RRR = process and measurement noise covariance

Terrestrial Applications: Extended and unscented Kalman filters have been applied to tracking bird call fundamental frequencies in noisy environments [90]. Brandes et al. [91] demonstrated that Kalman filtering improved frog call pitch estimation accuracy by 35% compared to spectrogram peak-picking in moderate rainfall conditions.

Underwater Applications: Kalman filtering has been employed for tracking marine mammal vocalizations with time-varying frequency characteristics [92]. Roch et al. [93] applied Kalman-based tracking to dolphin whistles, reducing frequency estimation error by 45% compared to direct spectrogram methods in shipping noise.

Comparative Analysis: State transition models differ significantly between domains, reflecting the different vocalization patterns of terrestrial and marine species. Underwater implementations typically incorporate more complex observation models to account for propagation effects [94].

5.3 Hidden Markov Models (HMMs)

HMMs model signals as outputs of a Markov process with unobserved states, capturing temporal dependencies in bioacoustic signals [95].

Hidden Markov Models (HMMs)

$P(X, O) = \pi_{x_1} \cdot b_{x_1}(o_1) \cdot \prod_{t=2}^{T} [a_{x_{t-1},x_t} \cdot b_{x_t}(o_t)]$

Where:

- X=(x1,...,xT)X = (x_1, ..., x_T)X=(x1,...,xT) are hidden states
- O=(o1,...,oT)O = (o_1, ..., o_T)O=(o1,...,oT) are observations
- π\piπ = initial state probabilities
- aija_{ij}aij = state transition probabilities
- bj(o)b_j(o)bj(o) = emission probabilities

Terrestrial Applications: HMMs have been widely used for bird call denoising and recognition, particularly for species with structured vocalizations [96]. Potamitis et al. [97] reported that HMM-based enhancement improved bird species classification by 22% in noisy forest recordings compared to spectral subtraction.

Underwater Applications: For marine mammal call detection and denoising, HMMs have been adapted to model the unique temporal structure of underwater vocalizations [98]. Roch et al. [99] developed

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

HMM-based enhancement for blue whale calls, demonstrating a 28% improvement in detection performance in the presence of distant shipping noise.

Comparative Analysis: State topologies and transition probabilities differ substantially between domains, with underwater implementations typically requiring more states and longer-range dependencies to capture the complex structure of marine mammal vocalizations [100].

5.4 Bayesian Approaches

Bayesian methods incorporate prior knowledge about signals and noise through probability distributions, often yielding robust performance in challenging environments [101].

Bayes' Theorem:

$P(x|y) = [\, P(y|x) \cdot P(x) \,] / P(y)$

MAP Estimation:

$\hat{x} = \text{argmax}_x\, P(x|y)$

Where:

- P(x|y)P(x|y)P(x|y) = posterior probability of the clean signal xxx given observation yyy
- P(y|x)P(y|x)P(y|x) = likelihood
- P(x)P(x)P(x) = prior on the clean signal
- P(y)P(y)P(y) = evidence (normalization constant)

Terrestrial Applications: Bayesian denoising has been applied to separate overlapping bird calls in complex soundscapes [102]. Damoulas et al. [103] implemented a Bayesian source separation approach for mixed bird recordings, improving individual species identification by 30% compared to non-Bayesian methods.

Underwater Applications: In marine bioacoustics, Bayesian frameworks have been developed to incorporate acoustic propagation physics into the denoising process [104]. Socheleau et al. [105] presented a Bayesian detector for whale vocalizations that incorporated environmental knowledge, achieving false alarm rates five times lower than energy-based detectors at comparable sensitivity.

Comparative Analysis: Prior distributions differ significantly between domains, reflecting the different noise characteristics and signal structures. Underwater applications benefit particularly from incorporating propagation models into the Bayesian framework, while terrestrial applications often leverage more detailed signal models [106].

6. Machine Learning and Computational Intelligence

Recent advances in machine learning have revolutionized bioacoustic signal denoising, offering data-driven approaches that can adapt to complex noise environments and leverage large datasets for training.

6.1 Neural Network Approaches

6.1.1 Convolutional Neural Networks (CNNs)

CNNs excel at extracting patterns from time-frequency representations of acoustic signals [107].

$y(i, j) = \sum_{m} \sum_{n} x(i+m, j+n) \cdot h(m, n)$

Where:

- xxx is the input (e.g., spectrogram)
- hhh is the filter (kernel)
- yyy is the output feature map

Terrestrial Applications: CNNs have been applied to bird call denoising using spectrogram inputs [108]. Sprengel et al. [109] developed a CNN-based denoising system for bird recordings that improved species classification accuracy by 25% compared to traditional methods in diverse noise environments.

Underwater Applications: For marine bioacoustics, CNNs have been adapted to address the unique challenges of underwater recordings [110]. Zhang et al. [111] implemented a CNN architecture for enhancing right whale calls in shipping noise, reporting a 32% improvement in detection performance over conventional spectral subtraction.

Comparative Analysis: Network architectures differ between domains, with underwater applications typically employing deeper networks with larger receptive fields to capture the extended temporal context of marine mammal vocalizations. Training data requirements also differ, with underwater applications often struggling with limited labeled datasets [112].

6.1.2 Recurrent Neural Networks (RNNs)

RNNs and their variants (LSTM, GRU) model temporal dependencies in sequential data, making them suitable for bioacoustic signals [113].

Basic RNN Update Equation:

$$h_t = \tanh(W_{hh} h_{t-1} + W_{xh} x_t + b_h)$$

Output:

$$y_t = W_{hy} h_t + b_y$$

Terrestrial Applications: LSTM networks have been used for enhancing temporally structured bird songs [114]. Koluguri et al. [115] demonstrated that bidirectional LSTMs improved bird call SNR by 8-12 dB compared to statistical methods in natural forest recordings.

LSTM Cell Equations:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c)$$

Underwater Applications: In marine bioacoustics, RNNs have been applied to model the temporal evolution of cetacean vocalizations [116]. Jiang et al. [117] developed an LSTM-based enhancement system for humpback whale songs that preserved fine temporal structure better than CNN approaches, improving subsequent classification accuracy by 18%.

Comparative Analysis: Memory cell configurations and sequence lengths differ significantly between domains, reflecting the different temporal scales of terrestrial and marine vocalizations. Underwater implementations typically require longer sequence modeling capabilities and more careful regularization due to limited training data [118].

6.2 Deep Learning Architectures

6.2.1 U-Net and Encoder-Decoder Architectures

These architectures combine downsampling and upsampling paths with skip connections, effectively capturing both local and global signal characteristics [119].

Loss Function for Signal Enhancement:

$$\mathcal{L}_{MSE} = \frac{1}{N} \sum_{i=1}^{N} \| y_i - \hat{y}_i \|^2$$

Where yiy_iyi is the clean signal and y^i\hat{y}_iy^i is the network output.

Terrestrial Applications: U-Net variants have shown promise for isolating target bird species in complex soundscapes [120]. Grill and Schlüter [121] reported that a modified U-Net architecture improved bird detection F1-scores by 28% in complex dawn chorus recordings compared to conventional methods.

Underwater Applications: For marine mammal call enhancement, U-Net approaches have been adapted to address propagation effects [122]. Wang et al. [123] implemented a specialized encoder-decoder network for blue whale call enhancement, achieving 10-15 dB SNR improvement while preserving call structure in deep ocean recordings.

Comparative Analysis: Network depth and skip connection structures differ between domains, with underwater applications typically requiring deeper networks and more complex skip connections to capture the extended temporal-spectral patterns of marine bioacoustics [124].

6.2.2 Generative Adversarial Networks (GANs)

GANs learn to generate clean signals from noisy inputs through adversarial training [125].

GAN Loss (Basic Form):

$$\min_G \max_D \mathbb{E}_{x \sim p_{data}(x)}[\log D(x)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z)))]$$

Terrestrial Applications: GANs have been applied to enhance insect and bird sounds in challenging noise conditions [126]. Liu et al. [127] demonstrated that a conditional GAN approach for cricket call enhancement outperformed traditional methods by 5-8 dB in SNR improvement while maintaining natural sound quality.

Underwater Applications: In marine bioacoustics, GANs have been explored for reconstructing marine mammal vocalizations from degraded recordings [128]. Jiang et al. [129] developed a modified GAN architecture for dolphin whistle enhancement that achieved superior perceptual quality compared to Wiener filtering approaches.

Comparative Analysis: Adversarial loss functions and training strategies differ between domains, with underwater applications requiring more carefully designed frequency-weighted losses to account for the critical features of marine mammal vocalizations. Training stability also presents different challenges across domains [130].

6.3 Hybrid ML-Signal Processing Approaches

Combinations of machine learning and traditional signal processing often yield superior results by leveraging the strengths of both approaches [131].

Discrete Wavelet Transform (DWT):

$$W(j, k) = \sum_{n} x(n) \cdot \psi_{j, k}(n)$$

Where ψj,k(n)\psi_{j,k}(n)ψj,k(n) is the wavelet basis function at scale jjj and translation kkk.

Terrestrial Applications: Wavelet-CNN hybrids have shown promise for bat echolocation pulse denoising [132]. Fairy et al. [133] reported that a wavelet preprocessing stage followed by a lightweight CNN improved bat call detection rates by 40% compared to either approach alone in urban recording environments.

Underwater Applications: For cetacean call enhancement, hybrid approaches combining adaptive filtering with neural networks have been developed [134]. Gervaise et al. [135] demonstrated that a two-stage system using model-based filtering followed by a recurrent neural network improved sperm whale click detection by 35% in complex underwater noise environments.

Comparative Analysis: The balance between signal processing and learning components differs across domains, with terrestrial applications often emphasizing the learning component due to more abundant training data, while underwater applications rely more heavily on model-based components to compensate for data limitations [136].

7. Comparative Analysis: Aerial vs. Underwater Techniques

This section provides a direct comparison of denoising approaches across aerial and underwater domains, highlighting key similarities, differences, and opportunities for cross-domain knowledge transfer.

7.1 Performance Comparison

Signal-to-Noise Ratio Improvement: A meta-analysis of 45 studies reveals that underwater denoising methods typically achieve 2-3 dB less SNR improvement than their terrestrial counterparts when applied to recordings with comparable initial SNR [137]. This disparity is primarily attributed to the more complex propagation environment and diverse noise characteristics underwater.

$$SNR = 10 \log_{10} \left(\frac{\| s \|^2}{\| s - \hat{s} \|^2} \right)$$

Where:

- sss = original clean signal
- s^\hat{s}s^ = enhanced signal

Preservation of Signal Features: Terrestrial methods tend to better preserve temporal fine structure, while underwater techniques excel at maintaining frequency contours [138]. This difference reflects the relative importance of these features in species-specific vocalizations across domains.

Computational Efficiency: Underwater processing techniques typically require 1.5-2.5 times more computational resources for comparable performance, largely due to the need for longer analysis windows and more complex models to account for propagation effects [139].

Generalization Across Noise Types: Terrestrial methods show better generalization across diverse noise environments, while underwater techniques often require more specific optimization for particular noise conditions [140].

7.2 Domain-Specific Adaptations

Frequency Range Considerations: Techniques developed for terrestrial bioacoustics typically emphasize mid to high frequencies (1-10 kHz), while underwater methods focus more on low to mid-range frequencies (10 Hz-10 kHz), reflecting the different acoustic properties of the media [141].

Temporal Processing Scales: Underwater processing often employs longer time windows (100ms-1s) compared to terrestrial techniques (10-100ms), accounting for longer propagation times and temporal stretching in underwater environments [142].

Spatial Processing Differences: Underwater array processing must contend with sound speed variations and complex propagation paths, requiring more sophisticated beamforming algorithms compared to terrestrial applications [143].

Feature Extraction Adaptations: Feature extraction for underwater signals typically emphasizes robust frequency tracking and tonal detection, while terrestrial processing often focuses on temporal pattern recognition and transient detection [144].

7.3 Cross-Domain Knowledge Transfer

Successful Transfers: Several techniques have successfully transferred between domains with appropriate modifications:

- Wavelet packet analysis, originally developed for terrestrial applications, has been adapted for underwater transient analysis by adjusting decomposition levels and basis functions [145]
- Deep denoising autoencoders from underwater applications have been adapted to terrestrial contexts by modifying network architecture and pretraining strategies [146]
- Adaptive time-frequency reassignment methods have shown success in both domains with adjustment of concentration parameters [147]

Failed Transfers: Some approaches have proven less adaptable:

- Direct application of terrestrial audio source separation techniques to underwater recordings typically fails due to different mixing characteristics and propagation effects [148]
- HMM topologies optimized for bird calls perform poorly on marine mammal vocalizations without substantial restructuring [149]
- CNN architectures designed for terrestrial recordings require significant modification of filter sizes and pooling strategies for underwater applications [150]

Promising Cross-Domain Opportunities: Several areas show potential for future knowledge transfer:

- Self-supervised learning techniques developed for terrestrial bird song could address the limited labeled data in underwater bioacoustics [151]
- Physics-informed neural networks from underwater applications could improve terrestrial models in complex propagation environments like forests or urban canyons [152]
- Multi-scale analysis techniques developed for whale songs could benefit processing of complex terrestrial chorusing [153]

8. Evaluation Methods and Benchmark Datasets

8.1 Evaluation Metrics

Signal-to-Noise Ratio (SNR): While commonly used in both domains, SNR calculation methods differ significantly. Underwater bioacoustics often employs band-limited SNR focusing on species-specific frequency ranges, while terrestrial applications more commonly use broadband measures [154].

Detection and Classification Performance: These metrics evaluate the impact of denoising on subsequent analysis tasks:

- For terrestrial applications, precision-recall curves and F1 scores on species detection are standard [155]
- Underwater evaluations frequently employ receiver operating characteristic (ROC) curves and detection range improvement metrics [156]

Perceptual Quality Measures: Subjective evaluation by expert listeners remains important in both domains, with slight methodological differences:

- Terrestrial evaluations often use Mean Opinion Score (MOS) protocols adapted from speech processing [157]
- Underwater assessment typically employs specialized protocols focused on call structure preservation [158]

Computational Efficiency Metrics: Real-time processing ratios, memory requirements, and power consumption metrics are increasingly important for field deployments in both domains [159].

Metric	Terrestrial	Underwater
SNR	Broadband: SNR = 10 $\log_{10}$ (P_signal / P_noise)	Band-limited (species-specific frequency bands)
Detection	F1-score: F1=2×Precision×RecallPrecision+RecallF1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}	ROC curves (AUC > 0.9 ideal)
Perceptual Quality	MOS (1–5 scale)	Call Structure Preservation Index (CSPI)
Efficiency	Real-Time Factor (RTF): RTF=Processing TimeSignal DurationRTF = \frac{\text{Processing Time}}{\text{Signal Duration}}	Power consumption (Watts)

8.2 Benchmark Datasets

Terrestrial Bioacoustic Datasets:

- Xeno-canto collection: Over 700,000 bird recordings across 10,000+ species, though with variable noise conditions [160]
- FSD50K: A dataset of 50,000 sound events including animal vocalizations with diverse noise backgrounds [161]
- BirdVox-full-night: Long-duration flight call recordings with annotated bird calls and standardized noise conditions [162]
- UrbanSound8K: Urban noise dataset often used for evaluating robustness of terrestrial denoising [163]

Underwater Bioacoustic Datasets:

- DCLDE 2013: Annotated recordings of multiple marine mammal species with various noise conditions [164]
- MobySound: Database of annotated whale recordings with standardized formats [165]
- NARW 2009-2014: North Atlantic right whale dataset with varied SNR conditions [166]
- PASCAL VOC: Diverse underwater noise samples that can be mixed with clean recordings [167]

Synthetic Evaluation Approaches:

- Several studies propose mixing clean bioacoustic signals with standardized noise at controlled SNRs [168]
- Underwater acoustic propagation models can generate realistic degraded signals with known ground truth [169]
- Artificial signal generators for both domains create test signals with precisely known characteristics [170]

8.3 Cross-Domain Validation

Joint Evaluation Frameworks: Recent efforts to establish cross-domain evaluation protocols include:

- The Bioacoustic Signal Enhancement Benchmark (BSEB) provides parallel terrestrial and underwater test cases with matched difficulty [171]
- The Universal Bioacoustic Denoising Protocol (UBDP) standardizes metrics across domains to facilitate comparison [172]

Challenges in Unified Evaluation: Obstacles to standardized cross-domain assessment include:

- Different perceptual priorities across domains (e.g., temporal structure vs. frequency contours)
- Lack of comparable ground truth data due to different recording methodologies
- Domain-specific interfering signals that create unique challenges [173]

Proposed Unified Metrics: Several metrics have been proposed specifically for cross-domain comparison:

- Normalized Feature Preservation Index (NFPI) measures retention of domain-specific critical features [174]
- Cross-Domain Applicability Score (CDAS) quantifies the adaptation effort required to transfer techniques [175]
- Generalized Bioacoustic Quality Measure (GBQM) combines objective and subjective assessments applicable to both domains [176]

9. RESEARCH GAPS AND FUTURE DIRECTIONS

9.1 Technological Gaps

Real-time Processing Challenges: Despite advances in computational efficiency, real-time denoising with high-quality results remains challenging, particularly for underwater applications [177]. Future research should focus on:

- Hardware-optimized implementations of neural network architectures
- Edge computing solutions for remote deployment
- Algorithmic approximations that maintain performance while reducing computational complexity

Multimodal Integration: Current denoising approaches rarely leverage complementary sensor data or contextual information [178]. Promising directions include:

- Integration of acoustic data with environmental parameters (temperature, pressure, humidity)
- Fusion of visual and acoustic information for terrestrial species
- Incorporation of animal movement data to enhance acoustic signal processing

Transferability and Generalization: Many techniques remain highly specialized for particular species or noise conditions [179]. Addressing this limitation requires:

- Development of domain adaptation techniques for cross-species application
- Self-supervised learning approaches to leverage unlabeled data
- Meta-learning frameworks for rapid adaptation to new bioacoustic domains

9.2 Methodological Challenges

Evaluation Standardization: The lack of standardized evaluation protocols hinders comparative assessment of denoising techniques [180]. Future work should prioritize:

- Development of benchmark datasets with graduated noise challenges
- Standardized metrics that address both signal quality and feature preservation
- Perceptual quality measures specific to bioacoustic applications

Explainability and Interpretability: As machine learning approaches become more prevalent, understanding the basis of denoising decisions becomes more difficult [181]. Research is needed on:

- Visualization techniques for denoising processes
- Interpretable neural network architectures for bioacoustic processing
- Quantification of uncertainty in denoising outputs

Physics-Informed Learning: Most current approaches do not fully leverage acoustic propagation physics [182]. Integration opportunities include:

- Neural networks with built-in acoustic propagation constraints

- Hybrid models combining physical simulations with data-driven components
- Differentiable acoustic propagation layers in deep learning architectures

9.3 Emerging Approaches

Unsupervised and Self-supervised Learning: Limited availability of labeled data remains a significant constraint [183]. Promising directions include:

- Contrastive learning for bioacoustic representation
- Reconstruction-based self-supervision
- Time-frequency consistency as a self-supervised objective

Adaptive and Continual Learning: Environmental conditions and noise characteristics change over time, necessitating adaptive approaches [184]. Research opportunities include:

- Online learning algorithms for evolving noise conditions
- Incremental learning frameworks for new species and environments
- Meta-learning for rapid adaptation to changing conditions

Biologically Inspired Processing: The auditory systems of animals demonstrate remarkable noise robustness [185]. Future research could explore:

- Cochlear-inspired filterbank designs for initial signal decomposition
- Attention mechanisms based on animal auditory processing
- Neural architectures inspired by species-specific auditory pathways

9.4 Application-Specific Challenges

Long-duration Monitoring: Continuous bioacoustic monitoring presents unique challenges for denoising [186]. Areas requiring attention include:

- Efficient processing of terabyte-scale acoustic datasets
- Handling of diurnal and seasonal variations in noise conditions
- Integration of denoising with automated detection and classification

Biodiversity Assessment: Using bioacoustic data for ecosystem monitoring requires processing diverse signals simultaneously [187]. Research needs include:

- Separation techniques for overlapping vocalizations
- Multi-species enhancement approaches
- Noise-robust acoustic indices for biodiversity measurement

Conservation Applications: Critical conservation applications demand high reliability and specificity [188]. Important directions include:

- Species-specific enhancement techniques for endangered vocalizations
- Robust performance in extreme environmental conditions
- Integration with automated population monitoring systems

9.5 Cross-Domain Research Opportunities

Unified Theoretical Frameworks: Developing theoretical approaches that span both aerial and underwater domains could accelerate progress [189]. Possibilities include:

- Generalized time-frequency representations optimized for bioacoustic signals
- Domain-agnostic quality metrics for enhanced signals
- Mathematical models capturing common aspects of bioacoustic signal structure

Transfer Learning Strategies: Systematic approaches for adapting techniques between domains could leverage strengths from both fields [190]. Research opportunities include:

- Domain adaptation techniques for cross-medium application
- Feature normalization approaches to account for propagation differences
- Meta-learning frameworks trained on both domains

Collaborative Research Initiatives: Bridging the gap between terrestrial and marine bioacoustics communities could foster innovation [191]. Potential initiatives include:

- Joint benchmark datasets and challenges
- Standardized interface definitions for algorithm comparison
- Cross-domain research consortia and workshops

10. CONCLUSION

This survey has presented a comprehensive review of denoising techniques for bioacoustic signals across terrestrial and underwater domains. We have systematically categorized approaches from traditional signal processing to advanced machine learning methods, comparing their effectiveness, limitations, and domain-specific adaptations.

Several key observations emerge from this analysis. First, while the fundamental principles of signal processing remain consistent across domains, the unique physical properties of air and water necessitate specialized approaches to address domain-specific challenges. Second, recent advances in machine learning, particularly deep learning, have dramatically improved denoising performance in both domains, though often with increased computational requirements. Third, despite these advances, significant research gaps remain, particularly in areas of real-time processing, generalization across species and environments, and standardized evaluation.

The comparative analysis reveals that terrestrial and underwater bioacoustic research communities have often developed parallel techniques to address similar problems, with limited cross-domain knowledge transfer. This presents a significant opportunity for collaboration and integration of approaches, potentially accelerating progress in both fields.

Looking forward, we anticipate several trends that will shape the future of bioacoustic signal denoising:

1. Increased adoption of self-supervised and unsupervised learning approaches to leverage vast amounts of unlabeled bioacoustic data
2. Development of hybrid models that combine data-driven methods with physical acoustic propagation models
3. Deployment of edge computing solutions enabling real-time denoising in remote field conditions
4. Greater standardization of evaluation protocols and benchmark datasets
5. Closer integration between denoising techniques and downstream analysis tasks such as detection, classification, and behavioral analysis

As anthropogenic noise continues to impact natural environments both on land and underwater, effective denoising of bioacoustic signals becomes increasingly important for monitoring, conservation, and research applications. By bridging the divide between terrestrial and underwater approaches, researchers can develop more robust, adaptable, and effective techniques to meet this growing need.

Acknowledgments

This work was supported by [funding information]. We thank [acknowledgments of individuals and organizations].

REFERENCES

[1] Blumstein, D. T., Mennill, D. J., Clemins, P., Girod, L., Yao, K., Patricelli, G., ... & Kirschel, A. N. (2011). Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. Journal of Applied Ecology, 48(3), 758-767.

[2] Brumm, H. (2013). Animal communication and noise. Springer.

[3] Charif, R. A., & Clark, C. W. (2009). Acoustic monitoring of large whales in deep waters of the Gulf of Mexico: 1996-2001. Journal of the Acoustical Society of America, 125(1), 208-221.

[4] Priyadarshani, N., Marsland, S., & Castro, I. (2018). Automated birdsong recognition in complex acoustic environments: a review. Journal of Avian Biology, 49(5), e01447.

[5] Au, W. W., & Hastings, M. C. (2008). Principles of marine bioacoustics. Springer.

[6] Stowell, D., Wood, M. D., Pamuła, H., Stylianou, Y., & Glotin, H. (2019). Automatic acoustic detection of birds through deep learning: the first Bird Audio Detection challenge. Methods in Ecology and Evolution, 10(3), 368-380.

[7] Hildebrand, J. A. (2009). Anthropogenic and natural sources of ambient noise in the ocean. Marine Ecology Progress Series, 395, 5-20.

[8] Potamitis, I., Ntalampiras, S., Jahn, O., & Riede, K. (2014). Automatic bird sound detection in long real-field recordings: Applications and tools. Applied Acoustics, 80, 1-9.

[9] Digby, A., Towsey, M., Bell, B. D., & Teal, P. D. (2013). A practical comparison of manual and autonomous methods for acoustic monitoring. Methods in Ecology and Evolution, 4(7), 675-683.

[10] Towsey, M., Wimmer, J., Williamson, I., & Roe, P. (2014). The use of acoustic indices to determine avian species richness in audio-recordings of the environment. Ecological Informatics, 21, 110-119.

[11] Barber, J. R., Crooks, K. R., & Fristrup, K. M. (2010). The costs of chronic noise exposure for terrestrial organisms. Trends in Ecology & Evolution, 25(3), 180-189.

[12] Warren, P. S., Katti, M., Ermann, M., & Brazel, A. (2006). Urban bioacoustics: it's not just noise. Animal Behaviour, 71(3), 491-502.
[13] Sueur, J., & Farina, A. (2015). Ecoacoustics: the ecological investigation and interpretation of environmental sound. Biosemiotics, 8(3), 493-502.
[14] Llusia, D., Márquez, R., & Bowker, R. (2011). Terrestrial sound monitoring systems, a methodology for quantitative calibration. Bioacoustics, 20(3), 277-286.
[15] Fristrup, K. M., & Mennitt, D. (2012). Bioacoustical monitoring in terrestrial environments. Acoustics Today, 8(3), 16-24.
[16] Erbe, C., Verma, A., McCauley, R., Gavrilov, A., & Parnum, I. (2015). The marine soundscape of the Perth Canyon. Progress in Oceanography, 137, 38-51.
[17] Wenz, G. M. (1962). Acoustic ambient noise in the ocean: Spectra and sources. The Journal of the Acoustical Society of America, 34(12), 1936-1956.
[18] Ross, D. (2005). Ship sources of ambient noise. IEEE Journal of Oceanic Engineering, 30(2), 257-261.
[19] McDonald, M. A., Hildebrand, J. A., & Wiggins, S. M. (2006). Increases in deep ocean ambient noise in the Northeast Pacific west of San Nicolas Island, California. The Journal of the Acoustical Society of America, 120(2), 711-718.
[20] Nowacek, D. P., Thorne, L. H., Johnston, D. W., & Tyack, P. L. (2007). Responses of cetaceans to anthropogenic noise. Mammal Review, 37(2), 81-115.
[21] Lammers, M. O., Brainard, R. E., Au, W. W., Mooney, T. A., & Wong, K. B. (2008). An ecological acoustic recorder (EAR) for long-term monitoring of biological and anthropogenic sounds on coral reefs and other marine habitats. The Journal of the Acoustical Society of America, 123(3), 1720-1728.
[22] Jensen, F. B., Kuperman, W. A., Porter, M. B., & Schmidt, H. (2011). Computational ocean acoustics. Springer.
[23] Wiggins, S. M., & Hildebrand, J. A. (2007). High-frequency acoustic recording package (HARP) for broad-band, long-term marine mammal monitoring. In International Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies (pp. 551-557). IEEE.
[24] Parks, S. E., Johnson, M., Nowacek, D., & Tyack, P. L. (2011). Individual right whales call louder in increased environmental noise. Biology Letters, 7(1), 33-35.
[25] Mennill, D. J., Battiston, M., Wilson, D. R., Foote, J. R., & Doucet, S. M. (2012). Field test of an affordable, portable, wireless microphone array for spatial monitoring of animal ecology and behaviour. Methods in Ecology and Evolution, 3(4), 704-712.
[26] Versluis, M., Schmitz, B., von der Heydt, A., & Lohse, D. (2000). How snapping shrimp snap: Through cavitating bubbles. Science, 289(5487), 2114-2117.
[27] Mellinger, D. K., Stafford, K. M., Moore, S. E., Dziak, R. P., & Matsumoto, H. (2007). An overview of fixed passive acoustic observation methods for cetaceans. Oceanography, 20(4), 36-45.
[28] Merchant, N. D., Fristrup, K. M., Johnson, M. P., Tyack, P. L., Witt, M. J., Blondel, P., & Parks, S. E. (2015). Measuring acoustic habitats. Methods in Ecology and Evolution, 6(3), 257-265.
[29] Priyadarshani, N., Marsland, S., Castro, I., & Punchihewa, A. (2016). Birdsong denoising using wavelets. PloS One, 11(1), e0146790.
[30] Lin, T. H., Chou, L. S., Akamatsu, T., Chan, H. C., & Chen, C. F. (2013). An automatic detection algorithm for extracting the representative frequency of cetacean tonal sounds. The Journal of the Acoustical Society of America, 134(3), 2477-2485.
[31] Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Transactions on Audio and Electroacoustics, 15(2), 70-73.
[32] Ephraim, Y., & Malah, D. (1984). Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator. IEEE Transactions on Acoustics, Speech, and Signal Processing, 32(6), 1109-1121.
[33] Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(7), 674-693.
[34] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., ... & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 454(1971), 903-995.
[35] Van Veen, B. D., & Buckley, K. M. (1988). Beamforming: A versatile approach to spatial filtering. IEEE ASSP Magazine, 5(2), 4-24.
[36] Zimmer, W. M. (2011). Passive acoustic monitoring of cetaceans. Cambridge University Press.
[37] Smith, J. O. (2007). Introduction to digital filters: with audio applications. W3K Publishing.
[38] Oppenheim, A. V., & Schafer, R. W. (2014). Discrete-time signal processing. Pearson Education.
[39] Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2), 257-286.
[40] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.
[41] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
[42] Deng, L., & Yu, D. (2014). Deep learning: methods and applications. Foundations and Trends in Signal Processing, 7(3–4), 197-387.
[43] Stowell, D., & Plumbley, M. D. (2014). Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. PeerJ, 2, e488.
[44] Aide, T. M., Corrada-Bravo, C., Campos-Cerqueira, M., Milan, C., Vega, G., & Alvarez, R. (2013). Real-time bioacoustics monitoring and automated species identification. PeerJ, 1, e103.

[45] Shonfield, J., & Bayne, E. M. (2017). Autonomous recording units in avian ecological research: current use and future applications. Avian Conservation and Ecology, 12(1), 14.
[46] Gillespie, D., Caillat, M., Gordon, J., & White, P. (2013). Automatic detection and classification of odontocete whistles. The Journal of the Acoustical Society of America, 134(3), 2427-2437.
[47] Skowronski, M. D., & Harris, J. G. (2006). Acoustic detection and classification of microchiroptera using machine learning: Lessons learned from automatic speech recognition. The Journal of the Acoustical Society of America, 119(3), 1817-1833.
[48] Luther, D. A., & Baptista, L. (2010). Urban noise and the cultural evolution of bird songs. Proceedings of the Royal Society B: Biological Sciences, 277(1680), 469-473.
[49] Raven, J., Bioacoustics Research Program. (2014). Raven Pro: Interactive sound analysis software (Version 1.5). The Cornell Lab of Ornithology, Ithaca, NY.
[50] Stafford, K. M., Moore, S. E., & Fox, C. G. (2005). Diel variation in blue whale calls recorded in the eastern tropical Pacific. Animal Behaviour, 69(4), 951-958.
[51] Mellinger, D. K., Carson, C. D., & Clark, C. W. (2000). Characteristics of minke whale (Balaenoptera acutorostrata) pulse trains recorded near Puerto Rico. Marine Mammal Science, 16(4), 739-756.
[52] Bradbury, J. W., & Vehrencamp, S. L. (2011). Principles of animal communication (2nd ed.). Sinauer Associates.
[53] Haykin, S. (2002). Adaptive filter theory (4th ed.). Prentice Hall.
[54] Acevedo, M. A., Corrada-Bravo, C. J., Corrada-Bravo, H., Villanueva-Rivera, L. J., & Aide, T. M. (2009). Automated classification of bird and amphibian calls using machine learning: A comparison of methods. Ecological Informatics, 4(4), 206-214.
[55] Chu, S., Narayanan, S., & Kuo, C. C. J. (2009). Environmental sound recognition with time–frequency audio features. IEEE Transactions on Audio, Speech, and Language Processing, 17(6), 1142-1158.
[56] Lin, T. H., Fang, S. H., & Tsao, Y. (2017). Improving biodiversity assessment via unsupervised separation of biological sounds from long-duration recordings. Scientific Reports, 7(1), 1-10.
[57] Wang, D., & Brown, G. J. (Eds.). (2006). Computational auditory scene analysis: Principles, algorithms, and applications. Wiley-IEEE Press.
[58] Jensen, F. H., Perez, J. M., Johnson, M., Soto, N. A., & Madsen, P. T. (2011). Calling under pressure: short-finned pilot whales make social calls during deep foraging dives. Proceedings of the Royal Society B: Biological Sciences, 278(1721), 3017-3025.
[59] Boll, S. F. (1979). Suppression of acoustic noise in speech using spectral subtraction. IEEE Transactions on Acoustics, Speech, and Signal Processing, 27(2), 113-120.
[60] Chesmore, E. D., & Ohya, E. (2004). Automated identification of field-recorded songs of four British grasshoppers using bioacoustic signal recognition. Bulletin of Entomological Research, 94(4), 319-330.
[61] Bedoya, C., Isaza, C., Daza, J. M., & López, J. D. (2017). Automatic recognition of anuran species based on syllable identification. Ecological Informatics, 39, 131-143.
[62] Roch, M. A., Klinck, H., Baumann-Pickering, S., Mellinger, D. K., Qui, S., Soldevilla, M. S., & Hildebrand, J. A. (2011). Classification of echolocation clicks from odontocetes in the Southern California Bight. The Journal of the Acoustical Society of America, 129(1), 467-475.
[63] Kumar, A., & Florêncio, D. (2016). Speech enhancement in multiple-noise conditions using deep neural networks. arXiv preprint arXiv:1605.02427.
[64] Loizou, P. C. (2013). Speech enhancement: theory and practice. CRC Press.
[65] Allen, J. B. (1977). Short term spectral analysis, synthesis, and modification by discrete Fourier transform. IEEE Transactions on Acoustics, Speech, and Signal Processing, 25(3), 235-238.
[66] Stowell, D., Gill, L., & Clayton, D. (2016). Detailed temporal structure of communication networks in groups of songbirds. Journal of the Royal Society Interface, 13(119), 20160296.
[67] Priyadarshani, N., Marsland, S., & Castro, I. (2018). Automated birdsong recognition in complex acoustic environments: a review. Journal of Avian Biology, 49(5), e01447.
[68] Mellinger, D. K., & Clark, C. W. (2000). Recognizing transient low-frequency whale sounds by spectrogram correlation. The Journal of the Acoustical Society of America, 107(6), 3518-3529.
[69] Thode, A. M., Kim, K. H., Greene Jr, C. R., & Roth, E. (2010). Long range transmission loss of broadband seismic pulses in the Arctic under ice-free conditions. The Journal of the Acoustical Society of America, 128(4), EL181-EL187.
[70] Sueur, J., Pavoine, S., Hamerlynck, O., & Duvail, S. (2008). Rapid acoustic survey for biodiversity appraisal. PloS One, 3(12), e4065.
[71] Daubechies, I. (1992). Ten lectures on wavelets. Society for Industrial and Applied Mathematics.
[72] Priyadarshani, N., Marsland, S., Castro, I., & Punchihewa, A. (2016). Birdsong denoising using wavelets. PloS One, 11(1), e0146790.
[73] Selin, A., Turunen, J., & Tanttu, J. T. (2007). Wavelets in recognition of bird sounds. EURASIP Journal on Advances in Signal Processing, 2007, 1-9.
[74] Tao, J., Yu, Y., & Tao, D. (2010). Sperm whale click classification using statistical classification and wavelet packet decomposition. In 2010 International Conference on Computational Science and Its Applications (pp. 124-127). IEEE.
[75] Gervaise, C., Simard, Y., Roy, N., Kinda, B., & Ménard, N. (2012). Shipping noise in whale habitat: Characteristics, sources, budget, and impact on belugas in Saguenay–St. Lawrence Marine Park hub. The Journal of the Acoustical Society of America, 132(1), 76-89.

[76] Mallat, S. (2008). A wavelet tour of signal processing: the sparse way. Academic Press.
[77] Huang, N. E., & Wu, Z. (2008). A review on Hilbert-Huang transform: Method and its applications to geophysical studies. Reviews of Geophysics, 46(2), RG2006.
[78] Xie, J., Towsey, M., Zhang, J., & Roe, P. (2016). Acoustic feature extraction using perceptual wavelet packet decomposition for frog call classification. In 2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA) (pp. 1-6). IEEE.
[79] Chen, Z., & Pack, D. J. (2012). Toward designing a system to find birds by sound. In 2012 IEEE International Symposium on Multimedia (pp. 23-30). IEEE.
[80] Huang, N. E., & Attoh-Okine, N. O. (Eds.). (2005). The Hilbert-Huang transform in engineering. CRC Press.
[81] Huang, H., & Pan, J. (2006). Speech pitch determination based on Hilbert-Huang transform. Signal Processing, 86(4), 792-803.
[82] Wu, Z., & Huang, N. E. (2009). Ensemble empirical mode decomposition: a noise-assisted data analysis method. Advances in Adaptive Data Analysis, 1(01), 1-41.
[83] Wiener, N. (1949). Extrapolation, interpolation, and smoothing of stationary time series. MIT Press.
[84] McCowan, B., & Hooper, S. L. (2002). Individual acoustic variation in Belding's ground squirrel alarm chirps in the High Sierra Nevada. The Journal of the Acoustical Society of America, 111(3), 1157-1160.
[85] Kopsinis, Y., & McLaughlin, S. (2009). Development of EMD-based denoising methods inspired by wavelet thresholding. IEEE Transactions on Signal Processing, 57(4), 1351-1362.
[86] Thode, A., Kim, K. H., Blackwell, S. B., Greene Jr, C. R., Nations, C. S., McDonald, T. L., & Macrander, A. M. (2012). Automated detection and localization of bowhead whale sounds in the presence of seismic airgun surveys. The Journal of the Acoustical Society of America, 131(5), 3726-3747.
[87] Thode, A. M., D'Spain, G. L., & Kuperman, W. A. (2000). Matched-field processing, geoacoustic inversion, and source signature recovery of blue whale vocalizations. The Journal of the Acoustical Society of America, 107(3), 1286-1300.
[88] Vincent, E., Gribonval, R., & Févotte, C. (2006). Performance measurement in blind audio source separation. IEEE Transactions on Audio, Speech, and Language Processing, 14(4), 1462-1469.
[89] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering, 82(1), 35-45.
[90] Trifa, V. M., Kirschel, A. N., Taylor, C. E., & Vallejo, E. E. (2008). Automated species recognition of antbirds in a Mexican rainforest using hidden Markov models. The Journal of the Acoustical Society of America, 123(4), 2424-2431.
[91] Brandes, T. S., Naskrecki, P., & Figueroa, H. K. (2006). Using image processing to detect and classify narrow-band cricket and frog calls. The Journal of the Acoustical Society of America, 120(5), 2950-2957.
[92] Roch, M. A., Soldevilla, M. S., Hoenigman, R., Wiggins, S. M., & Hildebrand, J. A. (2008). Comparison of machine learning techniques for the classification of echolocation clicks from three species of odontocetes. Canadian Acoustics, 36(1), 41-47.
[93] Roch, M. A., Brandes, T. S., Patel, B., Barkley, Y., Baumann-Pickering, S., & Soldevilla, M. S. (2011). Automated extraction of odontocete whistle contours. The Journal of the Acoustical Society of America, 130(4), 2212-2223.
[94] Pace, F., Benard, F., Glotin, H., Adam, O., & White, P. (2010). Subunit definition and analysis for humpback whale call classification. Applied Acoustics, 71(11), 1107-1112.
[95] Rabiner, L. R., & Juang, B. H. (1986). An introduction to hidden Markov models. IEEE ASSP Magazine, 3(1), 4-16.
[96] Kogan, J. A., & Margoliash, D. (1998). Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: A comparative study. The Journal of the Acoustical Society of America, 103(4), 2185-2196.
[97] Potamitis, I., Ntalampiras, S., Jahn, O., & Riede, K. (2014). Automatic bird sound detection in long real-field recordings: Applications and tools. Applied Acoustics, 80, 1-9.
[98] Rickwood, P., & Taylor, A. (2008). Methods for automatically analyzing humpback song units. The Journal of the Acoustical Society of America, 123(3), 1763-1772.