ITU Journal Future and evolving technologies
ITU Journal on Future and Evolving Technologies, Volume 7, Issue 1, March 2026
ARTICLE

Toward AI-enabled autonomous underwater acoustic networking: Challenges, opportunities, and research directions

Amani Alshawabka[1], Kerem Enhos[1], Emrecan Demiros[1,2], Daniel Uvaydov[2], Deniz Unal[2], Tommaso Melodia[1,2]

[1] Hydronet Inc., Burlington, MA 01803, USA, [2] Institute for Intelligent Networked Systems, Northeastern University, Boston, MA 02115, USA

Corresponding author: Emrecan Demirors, emrecan@hydronet.tech

Underwater acoustic networks enable critical applications including environmental monitoring, offshore infrastructure inspection, maritime security, and Autonomous Underwater Vehicle (AUV) operations. Their design, however, is fundamentally constrained by the acoustic propagation environment, characterized by severe bandwidth limitations, long and variable delays, strong Doppler effects, complex multipath, non-Gaussian noise, and the absence of widely accepted statistical channel models. These challenges significantly limit the effectiveness of traditional model-driven approaches. In terrestrial Radio Frequency (RF) systems, Artificial Intelligence (AI) has successfully enabled data-driven inference and control across the protocol stack. Motivated by these advances, recent work has begun exploring AI-driven underwater acoustic communications, yet adoption remains limited due to strong non-stationarity and environmental dependence. Using motivating examples based on real underwater data, we show that learned models exhibit limited generalization across environments. This paper presents a research roadmap toward AI-enabled autonomous underwater acoustic networks and sensing applications, identifying key challenges and outlining directions for robust, adaptive inference and control under stringent energy, computational, and environmental constraints.

Keywords: Acoustics, artificial intelligence, underwater

1. INTRODUCTION

Artificial Intelligence (AI) has emerged as a foundational technology reshaping the design, operation, and optimization of modern communication networks. By enabling systems to learn directly from data generated across multiple layers of the protocol stack, AI offers a powerful alternative to traditional model-based approaches, which often rely on closed-form analytical models that fail to capture the full complexity, non-stationarity, and heterogeneity of real-world network dynamics. Instead, data-driven methods offer adaptive, context-aware solutions capable of performing complex inference, prediction, and control tasks with increased flexibility and robustness. This paradigm shift is fundamentally redefining how communication networks are designed, from static, pre-engineered systems to adaptive platforms capable of autonomous operation in highly dynamic and uncertain environments.

Building on these capabilities, AI has been successfully adopted across a broad range of classification, prediction, and control tasks in RF wireless networking systems. Data-driven techniques have been applied throughout the protocol stack, enabling spectrum usage identification, signal and modulation classification, and device identification at the physical layer, including deep learning approaches for radio fingerprinting [1, 2, 3, 4, 5] and physical-layer feature extraction [6, 7]. In higher layers, AI has been used for traffic profiling and demand inference [8], as well as the prediction of user mobility, beamforming patterns, and traffic dynamics [9, 10, 11].

Beyond inference, AI has also demonstrated strong potential as a control mechanism, supporting adaptive load balancing, scheduling, and radio resource management strategies [12, 13]. Only a limited number of studies have begun to explore AI-driven solutions specifically tailored to underwater acoustic communications. Among these, real-time deep learning approaches for underwater acoustic spectrum sensing and modulation classification have been proposed that leverage convolutional neural networks to localize and identify physical-layer protocols directly from raw I/Q samples [14, 15].

These early efforts highlight both the promise of AI in the underwater domain and the substantial gap that remains relative to the maturity achieved in terrestrial RF systems.

This imbalance is not due to a lack of opportunity, but rather to fundamental structural differences between underwater networks and their terrestrial RF counterparts.

These differences arise not only from distinct propagation physics, but also from the absence of universally accepted statistical channel models applicable across diverse environments. While RF communication benefits from standardized and widely validated fading models that support reproducible and transferable system design, underwater acoustic channels remain strongly environment-dependent and difficult to generalize across locations and conditions [16, 17, 18].

These limitations suggest that future underwater communication systems must move beyond static, model-driven designs toward architectures capable of adapting to highly dynamic channel conditions and operational environments. Achieving this level of adaptability requires mechanisms that can extract actionable information directly from observed signals and environmental measurements. In this context, artificial intelligence provides a powerful framework for enabling data-driven inference, adaptive control, and intelligent decision-making across multiple layers of the underwater communication stack.

1.1 Motivation and contributions

In this context, AI can provide the foundational tools required to design, develop, and prototype the next generation of underwater communication transceivers and networking systems. At the physical layer, AI-driven techniques can be used to continuously monitor the acoustic spectrum, detect and classify incoming signals, infer their bandwidth and modulation characteristics, and enable robust decoding under severe channel impairments. Beyond inference, AI can support adaptive control of transmission operations by dynamically adjusting key parameters in response to time-varying channel conditions, topology, and traffic patterns. These include the selection of Modulation and Coding Schemes (MCS), transmission bandwidth and power, pilot placement, as well as higher-layer decisions such as routing and resource allocation. By tightly integrating inference and control across layers, AI-enabled systems can move beyond inflexible, preconfigured designs toward flexible and autonomous underwater networks capable of learning from, and adapting to, the environments in which they operate.

The goal of this paper is to identify and systematically analyze the key challenges and opportunities associated with the use of Artificial Intelligence (AI) in underwater acoustic communication and networking systems, and to articulate a research roadmap toward the realization of AI-enabled autonomous underwater networks. To this end, we first examine the fundamental characteristics of the underwater environment that distinguish it from terrestrial wireless systems and impose unique constraints on learning-based solutions. We then review and synthesize existing work on AI for underwater communications and networking, highlighting both successful applications and current limitations. Finally, we outline a set of promising use cases and open research directions where AI-driven inference, prediction, and control can play a transformative role in enabling more adaptive, efficient, and resilient underwater networking systems.

The remainder of this paper is organized as follows. Section 2 provides technical background on deep learning and underwater communications. Section 3 introduces the proposed architectural framework. Section 4 reviews the key challenges and characteristics of underwater acoustic environments. Section 5 presents AI-driven inference and control mechanisms across the protocol stack. Section 6 illustrates representative application scenarios and case studies. Section 7 discusses open challenges in enabling underwater AI. Finally, Section 8 concludes the paper and outlines directions for future research.

2. TECHNICAL BACKGROUND

2.1 Deep learning

Deep Learning (DL) techniques have been widely used in computer vision, speech processing, and natural language processing due to their ability to learn complex mappings between inputs and outputs directly from data [2, 3]. DL utilizes parameterized models, typically composed of multiple layers of neural networks, that approximate nonlinear relationships through learned numerical weights. These models are trained by optimizing a loss function using gradient-based methods such as backpropagation.

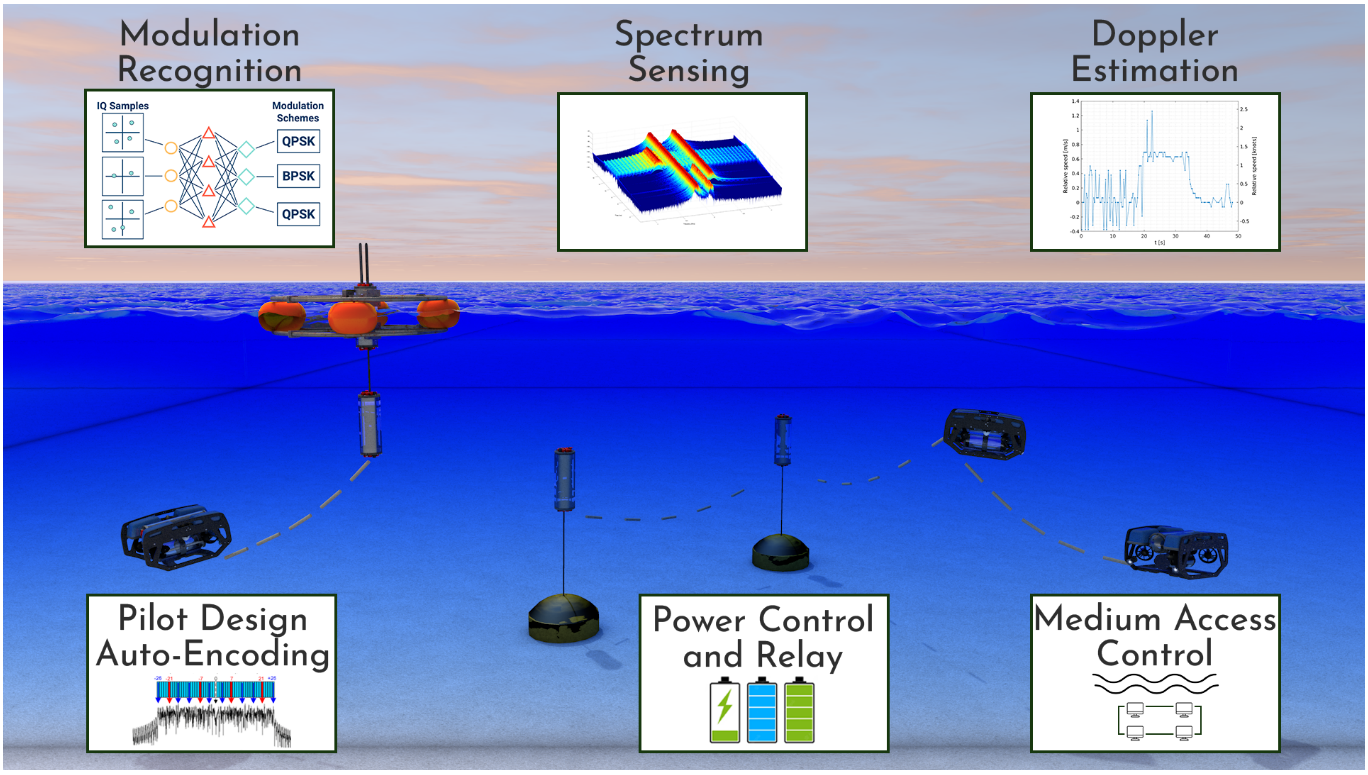


Figure 1 – AI-enabled inference applications in underwater wireless communications

Depending on the application, learning can occur under different paradigms, including supervised, unsupervised, and reinforcement learning. Supervised learning relies on labeled datasets in which the model learns to map inputs to known outputs. In contrast, unsupervised learning extracts patterns or structure from unlabeled data through objectives such as clustering or representation learning.

Reinforcement Learning (RL), and its deep learning variant Deep Reinforcement Learning (DRL), address sequential decision-making problems where an agent interacts with an environment and learns policies that maximize long-term rewards [19]. DRL combines neural networks with reinforcement learning to enable decision-making in complex environments with high-dimensional observations.

2.2 Deep learning for RF wireless communications

Recently, DL and DRL have gained increasing attention in the wireless networking community due to their ability to support both inference and control tasks [19, 20]. From classifying physical layer characteristics to allocating network resources, DL and DRL have proven highly conformable to the RF domain [20]. For wireless communications, the inputs to the parameterized models can take on many shapes. Generally the inputs to the models are some representation of the wireless environment or channel (e.g., raw IQ samples, power spectral density, network requests, etc.) that map to outputs representing aspects of the channel or the transmitted signal for classification tasks (e.g., modulation, coding, device identifier, etc.) or actions to be taken by devices for control tasks (e.g., network resources to allocate, modulation to transmit, frequency to use, etc.). A more in depth survey of the current applications of DL in wireless communications and networking can be found in [20, 21].

2.3 Underwater wireless communications

Even though the Underwater Wireless Communications (UWC) have common features, which are already present in the RF domain, it is clear that the UWC presents many unique aspects and challenges that can be addressed with data-driven DL algorithms. This section summarizes these unique aspects and challenges that should inspire novel research on the use of DL for UWC.

- *Acoustic Waves for Communication.* A wireless communications area is dominated by the usage of RF waves for terrestrial, cellular or sensor networks. However, RF waves suffer from high attenuation underwater, where long range communication can only be achieved with very low frequency (30-300 Hz) transmission, which requires large antennas and high transmission power while also impacting the data rate severely. Optical communication is an alternative to RF signals but alignment requirements and medium communication range prohibits this modality to be commonly used. Thus, acoustic wireless communication is the dominantly used modality for underwater wireless communication, mainly because of its low attenuation and long range signal transmission [18]. However, the Underwater Acoustic (UW-A) channel poses formidable challenges that are usually not present in any other modality. High path loss, colored noise, multipath, Doppler spread, and long propagation delays are some of these effects which characterizes this bandwidth limited, temporally and spatially varying communication channel.

- *No Accurate Channel Model.* Unlike in the RF domain where a number of models for both the probability distribution (e.g., Rayleigh fading) and the power spectral density of the fading process (e.g., the Jakes' model) are well-accepted and even standardized, as of today there is no consensus on statistical characterization of acoustic communication channels. Indeed, current models involve experimental measurements that are often made to assess the statistical properties of the channel in a particular deployment site at a particular time frame which limits their ability to generalize due to the space and time variability of the UW-A channel. Overall, the absence of accurate channel models significantly complicates the design of optimal communication systems and algorithms based on traditional block-based design principles [22].

- *Scarcity of Spectrum Resources.* Unlike in RF-based terrestrial communication systems where tens or even hundreds of MHz bandwidths per user and tens of GHz of total bandwidth are available and even standardized, in underwater acoustic communication channel available spectrum resources are exceedingly scarce. Underwater communication systems operating over long-range links (i.e., 10 - 100 km) may have available bandwidths of only a few kHz, while systems operating over medium-range links (i.e., 1 - 10 km) and over short range links (i.e., 0.1 - 1 km) have available bandwidths of on the order of 10 kHz and few tens of kHz, respectively. Only systems operating over very short-range links (less than 100 m) may have more than a 100 kHz of bandwidth [18].

- *Lack of Spectrum Regulations.* One of the unique aspects of underwater communications is the lack of regulation of underwater acoustic spectrum. Unlike the RF spectrum, underwater devices do not require licenses to operate at certain underwater acoustic spectrum bands. There is no spectrum allocation plan or power limitations, therefore underwater nodes can utilize any portion or the entire spectrum at any time with virtually unlimited transmit power levels. In addition to the lack of frequency allocation regulations, the spectrum is not exclusively used for communication links and various scientific and maritime applications. These applications utilize the underwater acoustic medium and are used in seismology, profiling, navigation, and marine biology areas. Apart from these applications, marine life also uses certain parts of the spectrum. For instance, some marine mammals emit sounds for navigation and communication while some species generate noise. The coexistence of these users results in an unpredictable environment for communication links, where spectrum resources are, may experience strong interference both in frequency and time.

3. TOWARD INTELLIGENT UNDERWATER ACOUSTIC SYSTEMS

AI-driven underwater systems blend sensing, inference, communication, and adaptive control into a cohesive processing architecture capable of operating under the challenging characteristics of underwater acoustic propagation. Unlike traditional pipelines, typically built upon fixed model assumptions and manually tuned Digital Signal Processing (DSP) blocks, AI-enabled platforms exploit learned representations that can generalize across deployments, hardware variations, and environmental conditions. These capabilities enable underwater systems to extract structure from complex acoustic observations, infer operational context, achieve situational awareness of the surrounding environment, and adapt communication strategies and system configurations in real time.

AI-enabled underwater applications can be categorized based on their functional role within the system. Specifically, we distinguish between inference functions that focus on extracting semantic or physical information from acoustic observations, control functions that adapt system parameters and operational behavior based on available information, and joint inference–control functions in which perception and adaptation are tightly coupled through closed-loop operation. This categorization provides a conceptual framework for the architectural views and application examples discussed throughout this section. The applications summarized in Table 1 are representative of those examined in this paper and are not intended to be exhaustive.

Fig. 2 illustrates the sensing and perceptual front end of an intelligent underwater system, summarizing the acoustic sources and environmental factors that must be captured and interpreted. These include marine mammals, surface and underwater vessels, Unmanned Underwater Vehicles (UUVs), anthropogenic activities, and communication waveforms transmitted by nearby nodes. All such sources are embedded within an acoustic field shaped by Doppler shifts, multipath propagation, and non-stationary noise, which collectively influence the structure and reliability of the received waveform. The figure highlights how the DSP front end of platforms such as *HydroNet Nexus*, a software-defined modem and edge platform [36], receives raw acoustic signals and

Table 1 – Data-driven underwater acoustic applications, organized by functional role and system layer.
Note: The learning methods listed represent examples used in the referenced work and are not exhaustive.

Domain	Type	Layer	Application	Learning Tool	Ref
Communication	Inference	PHY	Spectrum sensing	DL (CNN, autoencoder, multi-modal feature learning)	[14, 23]
			Modulation recognition	DL (RNN, CNN)	[15, 24]
			Doppler estimation / tracking	DL, Machine Learning (ML-based)	[25, 26]
			Channel estimation	DL	[27, 28]
	Control	PHY	Modulation and coding selection	DL, Reinforcement Learning	[29]
			Pilot design / placement	Reinforcement Learning	[30]
			Power control	DL	[31]
			Adaptive TX/RX pair	DL, Reinforcement Learning	[29, 15, 24]
		MAC	Medium access control	Reinforcement Learning	[32]
Sensing	Inference	APP	Marine mammal detection / classification	DL	[33]
			Vessel detection and classification	DL	[34, 35]

converts them into suitable signal representations, such as spectrograms, wavelet coefficients, or time-domain IQ sequences, that preserve the temporal–spectral characteristics required for both sensing and communication tasks and then applies necessary AI inference and control algorithms at the edge.

These representations expose discriminative patterns in the waveform that drive the neural inference process, enabling the system to associate observed patterns with corresponding acoustic sources, generate characteristic signatures, or inform decisions on subsequent actions aimed at improving overall system performance. Deep models based on convolutional, recurrent, or transformer architectures, among others, operate on these representations to learn compact feature embeddings that allow HydroNet AI to disentangle overlapping acoustic sources, recognize event signatures, and infer contextual information about the surrounding environment. Beyond supporting high-level semantic outputs such as detection and classification, these learned embeddings also serve as internal features that guide adaptive communication functions at the physical layer and, when applicable, at higher layers.

Fig. 3 extends this architectural view to a broader landscape of AI-enabled underwater applications, illustrating how learned representations support a wide range of sensing and communication functions. Building upon reliable event detection and representative datasets, AI-powered underwater systems enable sensing applications that are well established in underwater acoustics, including marine mammal detection and species classification, vessel detection and classification, UUV acoustic signature characterization, and monitoring of anthropogenic activities such as shipping and pile-driving operations. While some of these applications have traditionally relied on handcrafted signal processing pipelines, others remain challenging due to environmental variability and limited generalization across deployment conditions. Data-driven approaches provide a path toward more robust and scalable realizations by learning acoustic structure directly from observations, allowing systems to interpret complex and non-stationary underwater soundscapes across diverse environments.

At the physical layer, AI enables adaptive transmitter and receiver architectures capable of coping with the highly time-varying nature of underwater acoustic channels, where accurate analytical models are often unavailable. Data-driven methods for tasks such as channel estimation, Doppler compensation, and equalization learn channel characteristics directly from observations and can support adaptive communication strategies including modulation, coding, and waveform selection. At higher layers, AI can also assist network coordination by leveraging

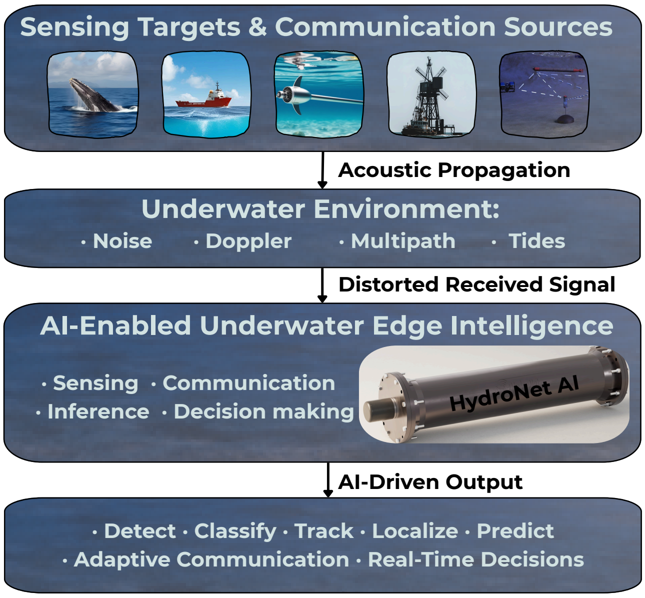


Figure 2 – AI-enabled sensing and communication front end of an underwater system. The figure illustrates representative sensing targets and acoustic sources, environmental effects shaping the received signal, and an AI-driven edge platform capable of joint sensing, inference, decision-making, and adaptive communication.

locally inferred channel and link information to support routing and other network-level decisions. Detailed examples of joint inference–control mechanisms are discussed later in Section 6. A central challenge underlying these capabilities is generalization across environments. To address this, AI-based techniques such as transfer learning, domain adaptation, self-supervised representation learning, and continual adaptation enable models to recalibrate using limited or unlabeled data collected in new deployments.

Figures 2 and 3 illustrate a unified learning-driven architecture where sensing, inference, and adaptive communication co-evolve. In this framework, the HydroNet Nexus modem acts as an intelligent processing unit that interprets the acoustic environment, adapts transmitter and receiver configurations, and enables reliable communication under dynamic underwater conditions while supporting edge AI analytics for inference and control. This approach marks a shift from rigid model-based pipelines to data-driven underwater communication systems capable of resilient operation across diverse missions. The following sections explore applications in inference and control and present a case study of joint inference–control operation.

4. APPLICATIONS IN INFERENCE

4.1 Spectrum sensing

As detailed in Section 2.3, UW-A channel spectrum is scarce and can most times experience strong interference due to number of man-made and natural noise sources that are frequency-dependent and may vary based on the location and time. Moreover, the spectrum access is not under regulation for frequency bands, which means users can access any part of the spectrum at any time with any transmission power level. In practice, some nodes such as battery-powered sensors may use little transmit power as they use energy sparingly to prolong their lifetimes while some others like SONAR systems on commercial vessels may transmit narrow-band signals with orders of magnitude more transmission power causing severe interference.

In the case of communication networks, centralized scheduling is uncommon since large delay spreads make time synchronization very challenging. Unlike the RF spectrum, the underwater acoustic spectrum can be unpredictable due to a variety of noise and interference sources that are dependent on environment and time. Usually, in underwater channels transmissions are spaced apart in time due to synchronization and delay spread constraints and may even appear infrequent compared to the RF channels. Unlike spectrum sensing in the RF domain, detection latency is not crucial in this setting. AI-enabled underwater nodes can sense the spectrum and adapt to the dynamic environment to determine the set of spectrum resources that satisfy the link requirements [14, 23].

4.2 Modulation recognition

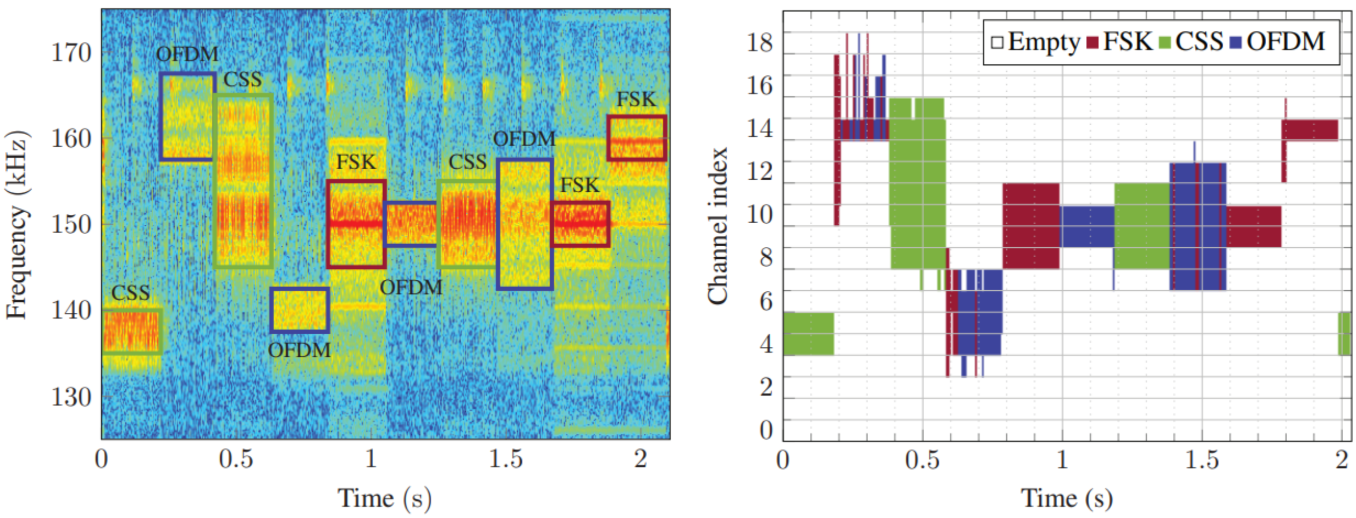
Similar to spectrum sensing, modulation recognition for underwater acoustic communications can be a crucial feature for efficient and adaptive communication. Due to high delay spread and low bit rate capability, decreasing the header length or avoiding a handshake mechanism can be accomplished with the help of AI-enabled modulation recognition. By eliminating the usage of information that does not contain actual data, the data rate can be increased. Additionally, automatic modulation recognition systems can be used to design polymorphic receivers. In this approach, the transmitter can send information bits without requiring acknowledgments using multiple modulation schemes and randomly varying packet parameters, making it significantly more difficult to eavesdrop on the communication channel.

Recently, several studies have been presented for modulation recognition using AI, which mostly exploits Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) network structures for automatic modulation recognition at the receiver chain [37, 38, 39, 40, 14, 15]. However, implementation of real-time AI-enabled automatic modulation recognition systems, feasible on computational complexity, energy efficiency, and adaptive to the time-varying channel conditions employed

Underwater AI-enabled applications

Sensing & perception	**Environmental awareness**	**Communications & networking**
• Marine mammal detection • Species classification • Vessel detection • Object monitoring & tracking	• Noise level estimation • Environmental safety monitoring • Ocean variability awareness	• Adaptive PHY (equalization, modulation) • AI channel estimation/ Doppler tracking • Jamming detection, secure routing
Security & system integrity		

Figure 3 – Landscape examples of AI-enabled underwater applications and system capabilities.



(a) Spectrogram of the received signal

(b) Joint localizer and classifier output

Figure 4 – Illustrative example: real-time deep learning–based modulation classification and localization of underwater acoustic signals [14]

on underwater sensor, mobile or stationary nodes is still an open challenge for underwater acoustic communication. Fig. 4 shows the results of underwater experiments using two acoustic modems, where one modem acts as the transmitter (Tx) and the AI-powered receiver (Rx) classifies the received signal modulation and localizes the channel indices and indices [14].

Another work [15] investigated the trade-off between input size (i.e., the number of received IQ samples) to the deep learning–based modulation classifier, where smaller input sizes imply a more efficient system, and overall classification performance. Fig.5 illustrates how classification accuracy and overall system latency vary with the input size. These results are based on real underwater experiments conducted using Hydronet modems deployed at the MArena testbed [41], as reported in [15].

4.3 Doppler estimation

An important physical layer limitation that affects underwater acoustic communication systems is the Doppler effect. The effect is caused by the relative motion between the transmitter and receiver, and it is an inherent problem for network nodes on mobile platforms such as underwater vehicles (ROVs or AUVs)[42, 43]. On the other hand, anchored nodes experience unintentional motion as they are not completely stationary and drift over time. In addition to node mobility, environmental processes such as waves can also affect propagation speed and paths. The resulting relative motion causes frequency shifts and spreading at the receiver. Relative motion in the order of a few m/s can result in non-negligible frequency-dependent distortion. The resulting signal suffers heavy performance degradation without compensation, especially in multi-carrier communication systems.

Therefore, estimating and compensating for frequency shift and spread caused by the Doppler effect is a critical

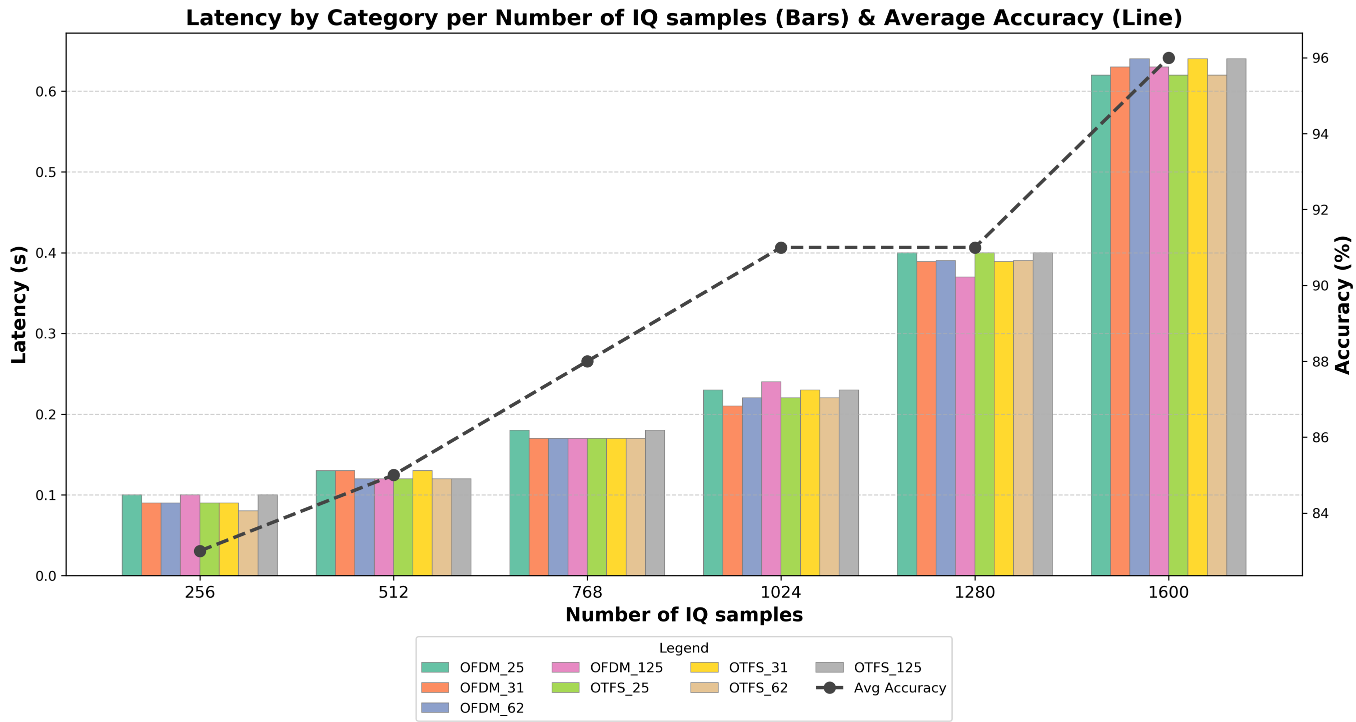


Figure 5 – Latency per category (modulation and bandwidth) and average accuracy versus the number of IQ samples. Bars show latency, while the dashed line indicates overall classification accuracy, highlighting the latency–performance trade-off on the Hydronet SDM-EP platform [36, 15]

challenge in the realization of high data rate communication systems in mobile settings. Conventional approaches to Doppler compensation systems typically insert training signals into the packets to estimate the duration of the packet at the receiver. From these estimations, time dilation/compression is calculated in the form of the Doppler scale factor for the received packet. Combined with frequency offset detection methods, non-uniform and uniform Doppler that result in spreading and shifting effects on the signal are estimated at the receiver [44, 45]. The efficacy of compensation depends on the accuracy of the estimates, and compensation algorithms are usually computationally intensive.

Doppler estimation can be implemented with AI to improve efficiency. Intentional motion of mobile platforms such as AUVs can be characterized in real time with trajectory tracking. This information can then be used to improve accuracy in the estimation phase in a closed-loop feedback system. Similarly, channel utilization can be improved by reducing the overhead used for Doppler estimation.

5. APPLICATIONS IN CONTROL

5.1 Modulation and coding

A typical problem in underwater networking is dynamical adaptation of the Modulation and Coding Scheme (MCS) to reduce the BER under varying SNR conditions [46]. In this context, AI can play a major role as a practical tool to ensure reliable and high performance communications by, for example, using a DRL agent capable of dynamically computing the optimal MCS that maximizes a set of performance metrics under varying channel conditions and changing performance requirements.

In time-varying wireless communication channels, adaptive communication schemes are conventionally used to achieve maximum throughput for changing channel conditions. The majority of currently available adaptive modulation schemes, however, heavily rely on feedback mechanisms which, due to the increased latency and overhead, have a detrimental effect on the performance of the acoustic communication systems. Moreover, due to the high delay spread, channel state information fed back from the receiver can rapidly become obsolete by the time it reaches the transmitter node, thus resulting in inaccurate channel measurements. In this context, there is an increasing need for solutions capable of recognizing modulations at the receiver side and adapting MCS at the transmitted side without any feedback mechanism in place, which is where AI can play a huge role in paving the way towards next-generation high-throughput and low-latency underwater acoustic networks.

Future research efforts should be targeted at designing zero-feedback DRL-based solutions that are capable of adapting MCS in real time while ensuring high performance and energy efficiency, as well as low computational complexity.

5.2 Pilot design

Channel estimation and equalization is one of the most critical challenges in underwater networks. Recently, pilot-assisted estimation and equalization techniques for underwater communications, especially in Orthogonal Frequency Division Multiplexing (OFDM) systems [30], have gained increasing momentum. By leveraging DRL and generative learning techniques, it is possible to develop data-driven adaptive pilot schemes where the amount of pilot symbols and their location in frequency is updated on a packet or frame basis. This adaptive form of pilot allocation, as opposed to static pilot allocation, can help avoid highly chaotic frequency selective interference, as well as increase resiliency as the underwater channel becomes worse. Moving pilots to avoid interference and increasing the number of pilots in poorer channel conditions.

For example, a transmitter node may estimate the channel and select the optimal pilot configuration for a specific channel condition based on feedback from the receiving node (e.g., throughput and error rate measurements). Indeed, pilots require a certain level of coordination between transmitter and receiver such that the latter can detect their location and number. In this context, such coordination can be achieved by either (i) allowing the transmitter to choose from a limited pool of pre-shared pilot configurations known to both the transmitter and receiver; or (ii) using deep learning techniques capable of classifying packet structures that identify and characterize pilots. For example, a practical approach is that of using a neural network to determine which I/Qs belong to preambles and which correspond to payload data or pilots.

5.3 Medium access control

Due to the large delay spread and channel variability, scheduling and Medium Access Control (MAC) in the underwater network poses a much greater challenge than in the RF domain. For instance, Carrier-Sense Multiple Access (CSMA)-based MAC protocols, which are widely used in RF systems, may sense the channel as idle while a transmission is ongoing, as the signal may not have reached the receiver yet due to the large propagation delay. On the other hand, ALOHA or Time-Division Multiple Access (TDMA) based protocols can only offer limited channel utilization and consequently reduced throughput.

On the contrary, AI techniques can offer a paradigm shift in the design of medium access control mechanisms. By utilizing deep learning mechanisms such as DRL, adaptive or dynamic MAC, strategies can be developed based on observed channel conditions and network states. Recent work has explored DRL-based approaches for wireless networking and resource allocation, demonstrating the potential of learning-based protocols to optimize spectrum access and transmission scheduling in dynamic environments [19, 47]. Unlike traditional MAC protocols that rely on predefined models or static channel assumptions, learning-based approaches can adapt their policies through interaction with the environment, enabling improved robustness under uncertain or time-varying channel conditions. Developing environment-adaptive control mechanisms suitable for ad hoc deployments is particularly important for underwater networks, where node mobility, intermittent connectivity, and limited prior channel knowledge make traditional protocol design challenging.

To this end, mechanisms that rely on distributed MAC solutions (as opposed to centralized schedulers) may be easier to integrate with AI-enabled control mechanisms, particularly when considering feedback-based learning approaches such as DRL. Recent work on distributed underwater robotic systems and multi-AUV coordination highlights the importance of decentralized communication and control architectures [48, 49].This is because feedback may not be easily achieved in the underwater setting, especially if the wrong scheduling or control action is taken, which may result in a complete link loss or packet drop. Indeed, a centralized solution with a unique node providing feedback to multiple nodes introduces a unique point of failure that can prevent the proper execution of networking functionalities, especially in the case of link failures and lost packets. So focusing on more distributed AI-enabled MAC topologies may be easier to establish in the underwater channel and have higher resilience to chaotic channel conditions.

5.4 Power control and relay

Long-distance underwater communications will generally require optimization of power control and relay mechanisms. Due to the high absorption of acoustic signals by water over large distances and the limited battery power of nodes, maximizing SNR and minimizing communication distance are heavily constrained. The energy limitation for underwater networks, as it, determines the lifetime. Depending on the topology, depleted nodes may result in premature loss of network connectivity, and therefore, lifetime is a crucial metric for routing protocols. By using deep reinforcement learning, nodes can tweak their transmission powers and, subsequently, their relay paths based on their interactions or changes with their environment while keeping battery constraints in mind.

In practice, these techniques are required to be lightweight in the sense that full network awareness (or network topology) will not fully be known or even available in many cases. Therefore, much of the challenge in effectively deploying power control/relay control agents will be in minimizing the overhead of the deep reinforcement learning feedback or reward system to allow the node to change with its environment. Future work should focus not only on high-performing DRL agents in this space but feedback protocols that are conducive to a DRL agent and consider DRL agents as a natural part of the TX/RX chain.

Energy efficiency of underwater acoustic communication systems can also be increased with recognizing the modulation at the receiver side without supplying any header information. Compared to traditional RF wireless devices, underwater modems consume much more energy for transmission in order to maintain long distance communication. Instead of the traditional energy consumption structure $E_T > E_R = E_L > E_P$, where E_T, E_R, E_L and E_P represent the energy consumption for transmitting, receiving, idle listening and processing respectively, the energy consumption structure in underwater communication system can be represented as $E_T >> E_R = E_L > E_P$ [50]. Also due to the fact that underwater modems are battery powered devices with limited energy storage and hardly accessible (deep ocean, remote location deployments) for recharging or switching batteries, energy efficiency is a major performance parameter for these systems. Hence, reception and idle listening is a more energy efficient task to operate for power constrained underwater devices. Even with the implementation of energy harvesting mechanisms or AI-enabled modulation recognition capable wake-up receiver structures, underwater sensor nodes can potentially operate as battery-less platforms [51]. In such architectures, lightweight inference models executed on ultra-low-power hardware can enable event-driven activation of the main communication and processing subsystems, significantly reducing idle listening and energy consumption. This approach highlights the importance of designing energy-efficient AI mechanisms tailored for resource-constrained underwater sensing platforms[52].

6. CASE STUDY: JOINT INFERENCE AND CONTROL

This case study illustrates how inference and control operate jointly within an AI-enabled underwater communication system. We consider a representative deployment scenario in which multiple underwater nodes perform local acoustic sensing and inference while adapting their transmission and reception strategies in response to changing channel conditions and environmental dynamics. The objective is to demonstrate how the architectural principles discussed earlier can be realized in practice, rather than to provide a comprehensive performance benchmark.

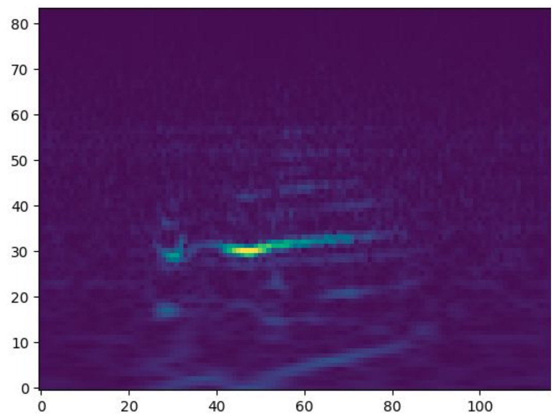


Figure 6 – CQT representation of a North Atlantic Right Whale (NARW) vocalization, illustrating characteristic spectral–temporal structure.

6.1 Marine mammal acoustic inference

Marine mammal acoustic inference is a challenging problem due to the diversity of vocalization patterns and the highly variable underwater environment in which they are produced and recorded. Vocalizations differ significantly across species and behaviors, and are often corrupted by environmental noise and interference from other natural and anthropogenic sources. Reliable classification of these signals is nevertheless important for applications such as species monitoring, behavioral analysis, and assessing the impact of human activities on marine ecosystems.

Fig. 6 shows a representative Constant-Q Transform (CQT) time–frequency representation of a North Atlantic right whale (NARW) vocalization. The CQT provides a compact description of spectral–temporal structure across a wide frequency range, capturing harmonic structure and temporal–spectral patterns characteristic of this species. While shown here for illustrative purposes, such representations highlight the signal structure typically exploited for downstream inference tasks in marine mammal monitoring.

The development and validation of marine mammal inference methods are further complicated by challenges associated with data availability and labeling. Widely-used reference datasets, such as the Watkins Marine Mammal Sound Database (WMMD) [53], provide valuable recordings spanning multiple decades and species, but also exhibit variability in recording conditions, background noise, and class imbalance. In addition, some species remain underrepresented, limiting the ability to construct uniformly robust inference methods across different environments and deployment scenarios.

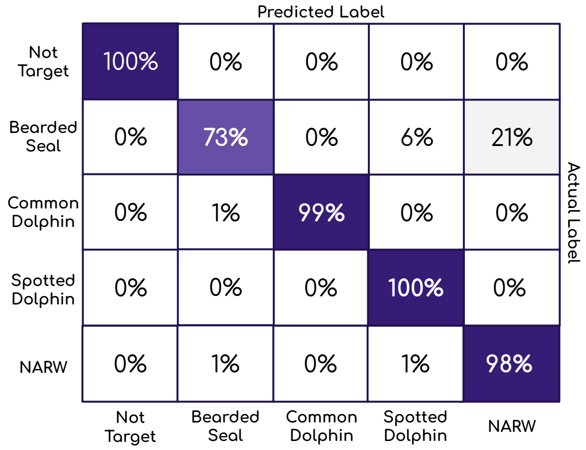


Figure 7 – Marine mammal acoustic inference results for a small dataset of multi-class species classification. The confusion matrix is obtained using a CNN-based classifier operating on time–frequency representations of the acoustic signals.

Fig. 7 presents representative classification results obtained for marine mammal acoustic inference using spectral–temporal representations derived from the CQT. In this implementation, the time–frequency inputs are processed by a CNN classifier that extracts time–frequency features and maps them to species labels through a fully connected softmax classification head. The model is trained using supervised learning with a cross-entropy loss objective. The results demonstrate that meaningful separation between classes can be achieved using time–frequency structure alone, while also revealing sensitivity to signal variability, environmental conditions, and dataset composition.

Within the broader system architecture, this marine mammal inference stage serves as a perceptual front end that provides situational awareness of biologically relevant acoustic activity. The inferred information can be used to detect the presence of protected species and to trigger appropriate system-level responses, such as modifying operational behavior or enforcing mitigation measures aimed at reducing disturbance to marine life. These results are included to illustrate system capability rather than to provide a comprehensive performance evaluation. Additional data collection and broader experimental coverage are required to obtain a more complete characterization of performance across species and environmental conditions; nevertheless, the observed behavior indicates a promising direction for integrated underwater sensing and inference within the proposed system framework.

For the inference task, spectral–temporal representations derived from the CQT are used as inputs to a CNN classifier. The network extracts hierarchical time–frequency features from the input representation and feeds them to a fully connected classification head with a softmax output layer to perform multi-class species identification. The model is trained using supervised learning with a cross-entropy loss function, where labeled recordings from the dataset are used to optimize the network parameters. This architecture provides a compact yet effective approach for learning discriminative acoustic features associated with different marine mammal vocalizations.

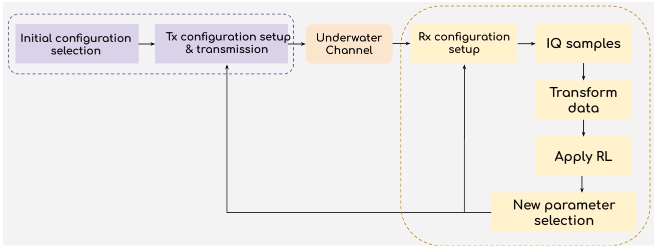


Figure 8 – High-level conceptual diagram of adaptive transmitter and receiver.

These results are included to illustrate the feasibility of the proposed inference pipeline rather than to provide a comprehensive benchmark comparison. A detailed performance evaluation against alternative models and larger datasets is left for future work.

6.2 Adaptive transmitter and receiver pair

Deep learning-assisted inference and control tasks described in the previous sections can be integrated into an adaptive communication scheme. The nodes use spectrum sensing for interference avoidance and discovering high SNR regions to decide on the spectrum resources to use. The transmitter can select the optimal modulation and coding based on the perceived channel conditions as trained. Whereas the receiver can identify this selection from the incoming packet without the need for an explicit header; Fig. 8 shows a high-level conceptual diagram of the deployment. A similar mechanism can be implemented for channel estimation with the use of adaptive pilot design techniques. If the transmitter expects a severe multipath channel based on measurements and feedback, the complexity of the pilots can be increased to improve estimation accuracy and reduce packet loss. Transmit power control complements these decisions to further optimize the link performance.

Overall, these methods aim to sustain near optimal link performance. In addition to increasing throughput, they can also increase energy efficiency and reduce outage probability and latency by preventing costly retransmissions due to avoidable packet loss. Similarly, an excess number of transmissions in the case of underutilized link performance are avoided. Models can be further trained to avoid or mitigate the effects of jamming.

This adaptive link structure can work with or without feedback from the receiver. If a feedback link is present, performance metrics such as signal-to-noise ratio or bit error rate can be reported to the transmitter as a reward. The inference tasks are not latency sensitive as they are in the RF domain. The signal bandwidths used in acoustic channels are significantly narrow compared to RF channels, thus a packet with a reasonable payload size would last more than tens of milliseconds.

A sensor node may choose to defer transmission if channel conditions require the use of a low code rate along with high transmission power. Instead, to preserve energy, it can wait and monitor the spectrum and transmit the accumulated data with a higher rate scheme at an instance when packet error probability is estimated to be lower. If either intentional or unintentional platform motion is detected at the transmitter, it can switch to a Doppler-resistant physical layer scheme to retain a robust communication link.

6.3 AI-based modulation classification example

To illustrate the role of AI-powered inference in underwater communication systems, we consider a modulation classification example using Orthogonal Frequency Division Multiplexing (OFDM) and Orthogonal Time Frequency Space (OTFS) waveforms collected from a real underwater acoustic deployment using Hydronet Nexus platforms **[17]**. Two Hydronet Nexus modems are used as the transmitter and receiver. The transmitter operates under predefined configuration settings, while the receiver captures and labels the corresponding signals to build a dataset for training a deep learning model. During operation, the pretrained model should run at the receiver to infer the configuration of the incoming signal, such as waveform type, bandwidth, and FFT size, and automatically adjust the receiver parameters accordingly.

The experiments evaluate whether a classifier can infer key transmission parameters, such as waveform type, bandwidth, and FFT size, from only a limited portion of the received payload. Two input lengths, 1024 and 1600 In-phase and Quadrature (IQ) samples, are evaluated.

Fig. 9 presents the confusion matrices for the AI-based modulation classification example using two input lengths. Increasing the observation window from 1024 to 1600 IQ samples improves overall classification accuracy from approximately 91% to 96%, indicating that larger input segments capture more discriminative signal features within a single deployment. However, even a well-performing CNN classifier trained in one environment can experience significant performance degradation when applied across different underwater conditions. This generalization gap highlights the challenges of deploying learning-based receivers in dynamic underwater environments and underscores the need for adaptive and environment-aware AI-driven communication architectures.

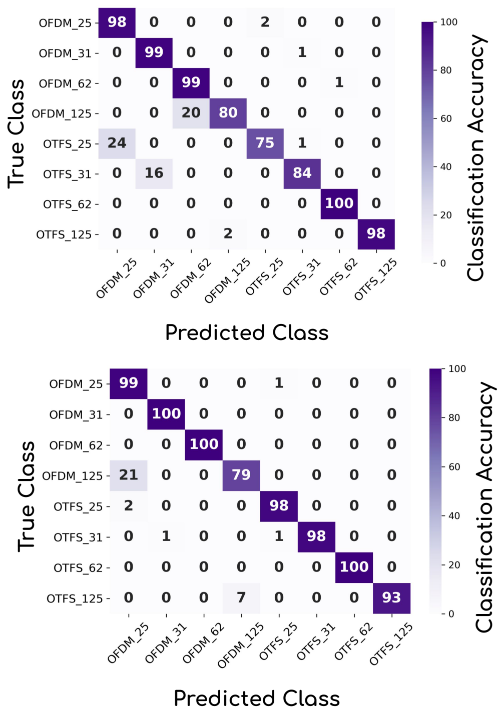


Figure 9 – Confusion matrices illustrating classification performance across eight waveform categories defined by waveform type (OFDM or OTFS) and bandwidths with FFT size fixed to 4096. Results are shown using portions of the received payload waveform: (a) (top) input size of 1024 IQ samples; (b) (bottom) input size of 1600 IQ samples [15]

.

6.4 End-to-end AI enabled wireless system

In addition to utilizing AI to help automate individual aspects of the encoding or decoding process (i.e. block code size, MCS, pilot design), there exists work in the RF domain to replace the entire encoding or decoding process with deep learning [54, 55, 56]. This involves a model at the transmitter that encodes m bits into k encoded symbols and a model at the receiver that does the inverse mapping. The advantage of such systems being the data-driven approach allows for the adaptation of nodes to their wireless channel rather than relying on assumed channel models or channel statistics. This has a clear advantage for the underwater domain as the channel is ever-changing and consistent channel models are unavailable for the underwater domain. The challenge here lies in back-propagation which is more difficult as the channel is considered to be part of the DL model, therefore updating the weights of the transmitter is not rudimentary if the channel model is unknown. This incurs training overhead which many works in the RF domain address and something that underwater implementations would need to make use of as well. Finding alternative ways to update the gradients in the transmitter while jointly reducing training overhead or deployment can prove very effective in overcoming channel stochasticity, especially if deployed in an online setting, as discussed in more detail in Section 7.2. Fig. 10 illustrates a complete example of an AI-based, real-time, end-to-end system deployment

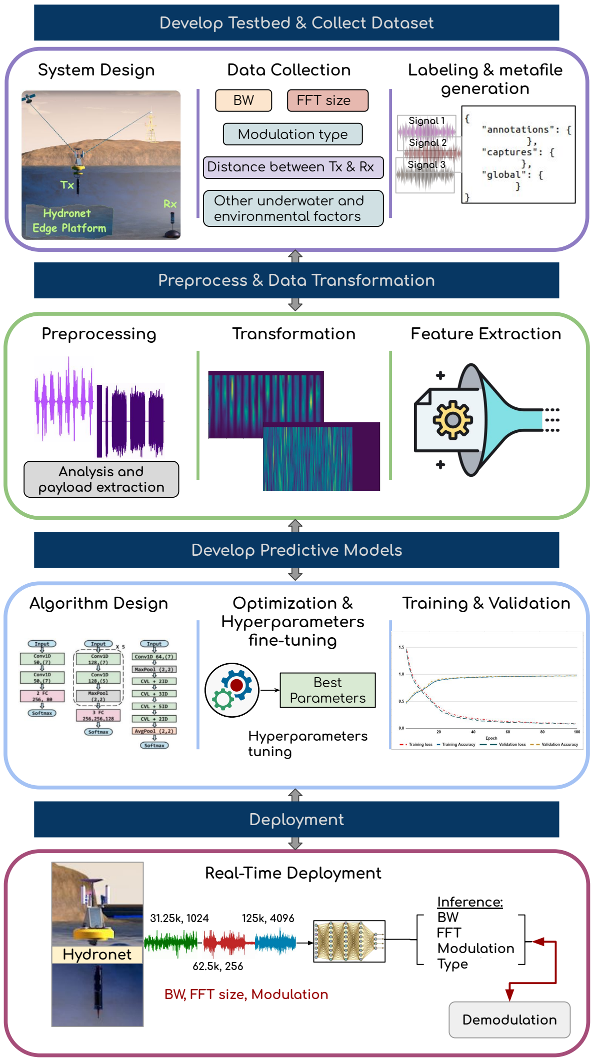


Figure 10 – Example of DL-based automatic modulation classification: End-to-end system pipeline [15]

pipeline. While the example is instantiated for modulation classification, the same processing pipeline can be applied to a wide range of underwater applications, including adaptive transmitter and receiver operation, vessel detection and classification, among others. By tailoring the underlying AI or deep learning models to specific system requirements and available observations, this architecture enables multiple inference and control tasks to be supported within a unified framework.

7. OPEN CHALLENGES IN ENABLING UNDERWATER AI

7.1 Dataset generation

Dataset generation is a demanding task as it typically requires a deployment setup with multiple vessels. The process is time-consuming and costly and therefore existing research often utilize datasets collected in past experiments. Although on-site dataset recordings are authentic and provide the best experimental evaluation method for AI models, it is impractical to record large datasets that capture temporal variations at a wide range of locations. Though not all networks require labeled datasets, such as unsupervised networks, these can be harder to train or give poor performance or generalization.

Meaningful dataset generation poses as one of the most crucial challenges in enabling underwater learning. More specifically, curating datasets that not only utilize real underwater signals, but which have high diversity or generalization and can in turn lead to generalizing learning algorithms. Dataset generalization here means the dataset contains signals gathered from a wide variety of settings and parameters (i.e., locations, distances, scenarios, signal-to-noise ratios). This is especially important when considering inference tasks as they are trained using past data and new scenarios may cause classification inaccuracies. Therefore, a community effort to curate more datasets for underwater learning is critical.

However, it is clear to see that achieving generalization is a largely time-encompassing task that may not be feasible in underwater. For example one can approximate the time, T it takes to collect a dataset of size N samples with M IQs per sample and a sampling rate of f_s for training/testing as $T = \frac{NM}{f_s}$. In Table 2 we show the difference in time it takes to collect datasets of the same size in RF and underwater. We can see underwater dataset collection is largely bottle-necked by the small sampling rates. Therefore, there is a need to curate datasets in a more intelligent manner, or utilize learning algorithms designed for small imbalanced datasets. A simple way to help DL algorithms deal with unknown scenarios that are also indifferent to the classifier is by collecting data for an "other" class. This means not only collecting data for the classes of interest but also adding an extra class that contains data representing useless information. Generally, this involves collecting data from the underwater channel during periods when no active communication is present. The collected data therefore contains ambient noise and interference typical of the underwater environment, as well as other undesired signals that may be present in the deployment area.

Table 2 – Duration of data collection for datasets of different sizes in RF and underwater

# of Dataset Samples	# of IQs	RF Time (s) $f_s = 20MSps$	UW Time (s) $f_s = 200kSps$
100,000	256	1.28	128
500,000	512	12.8	1280
1,000,000	1024	512	51,200

Furthermore, data augmentation has some existing work in the RF domain for dealing with sparse datasets and can easily translate to the underwater realm. This can involve a combination of recorded and augmented data. For example, collecting a small "in the wild" dataset which can be extended through augmentation techniques that modify SNR, multipath, and interference. The advantage here, is that the dataset is not entirely synthetic. Another way of augmenting data is to train a DL model such as a Generative Adversarial Network (GAN) to emulate or mimic the underwater channel (think "deep fake" as it is colloquially termed), as done in RF [57, 58]. This offers a data-driven method of modeling the underwater channel which can be used for generating new semi-augmented datasets without the need for further deployment. The drawback here is a sort of catch-22 as in order to train a good channel emulator you need to generate a dataset to train it initially. Another drawback is that the GAN channel model will most likely be a good emulator of the channel with which it was trained in, and not generalize into completely different bodies of water.

7.2 Online learning

Online learning can enable constant learning and keep AI-enabled underwater nodes adaptable to new environments. This is an especially attractive quality as it minimizes the need for large dataset collection and furthermore can achieve a high level of generalization. The current challenge in developing online learning for underwater nodes mainly deals with increased complexity and overhead. Underwater nodes have limited battery lifetime and constant learning can constrain mobile underwater devices. Future work must focus on jointly optimizing hardware constraints with performance and AI complexity to be conducive with mobile underwater devices. For example, embedded deep learning is a vastly studied field that can hugely benefit the underwater community, using techniques such as dimensionality reduction and pruning to develop computationally efficient DL and DRL algorithms. Such models can help ease processing latencies in already lengthy training times.

In addition to reducing computational complexity of the DL or DRL agent itself, to be able to deploy online in an effective way, feedback needs to be easily obtained. This can apply to both cases as a DRL agent will generally need some form of feedback from the environment or the receiving device. Also for the DL network, as in the end-to-end case mentioned in Section 6.4, where a transmitter must update its weights based on the performance of the receiver. This poses a challenge in the underwater domain as it may incur higher communication overhead in an otherwise unstable channel. An even greater challenge exists if the underwater network topology takes on a centralized structure due to the potential single point of failure in sending out feedback to multiple nodes. Developing protocols or mechanisms that can mitigate the high communication overhead for deploying underwater nodes is therefore crucial to its functionality. Some ideas to consider may be to employ nodes that aggregate multiple feedbacks or observations before sending feedback for training the node, such as many on policy methods in the DRL case. Although the feedback packet size is larger there is less frequent communication needed.

7.3 Federated learning

Another option for avoiding extraneous dataset collection is to utilize federated learning. Federated learning involves training nodes locally, at the edge, then transmitting the model weights, rather than any data, to be aggregated and updated in the cloud or at a server. Federated learning is rarely considered in the underwater space [59]. This is most likely due to: (i) the instability of online learning which generally predicates the ability to achieve federated learning and (ii) the centralized topology that federated learning generally assumes for model aggregation (in a cloud or base station). Although a centralized topology may be feasible in the underwater space, it provides a single point of failure in an already difficult communication environment. Further work can focus on enabling federated learning in underwater environments that is conducive to an ad hoc or decentralized topology, as well as centralized topology. A decentralized form of federated learning is referred to as gossip learning [60]. This involves nodes communicating amongst each other to aggregate and update the model shared amongst them, without the need for a central server. For federated learning in the reinforcement learning case, methods such as A3C which involve agents learning in parallel and aggregating every so often combined with a decentralized aggregation can be very powerful.

7.4 Simulation-based training and digital twins

Beyond traditional dataset collection and augmentation techniques, simulation-based training has emerged as a promising approach for addressing data scarcity in underwater AI systems. Physics-informed acoustic propagation models, such as ray-tracing simulators (e.g., Bellhop), can generate large volumes of synthetic data representing diverse channel conditions and environmental configurations [61]. These simulation environments enable controlled exploration of propagation phenomena such as multipath, Doppler spread, and environmental variability, providing valuable training data for learning-based inference and control algorithms in underwater communication systems.

However, models trained purely on simulated data often face a domain gap when deployed in real underwater environments. This discrepancy arises because even sophisticated simulators cannot fully capture the complexity, environmental heterogeneity, and non-stationarity of real acoustic channels. To address this challenge, techniques from the sim-to-real transfer literature, such as domain adaptation and domain randomization, can be leveraged to improve model robustness and generalization [62]. These approaches have been successfully applied in robotics and autonomous systems, where models are first trained in simulation and later adapted to real-world conditions using limited experimental data [62].

An emerging concept that further extends this paradigm is the development of *digital twins* for underwater communication networks. A digital twin represents a virtual counterpart of a physical deployment that integrates physical models, environmental sensing data, and operational history to reproduce system behavior in a simulated environment [63, 64]. In underwater settings, such digital twins could incorporate acoustic propagation models, environmental observations, and network-level simulations to create a continuously evolving virtual environment that mirrors real deployments. This environment can be used to test new algorithms, evaluate network protocols, and support AI model development between costly at-sea experimental campaigns.

Hybrid training pipelines combining simulation and real-world data offer a particularly promising direction. In these pipelines, models are first pre-trained using large volumes of simulated data and subsequently fine-tuned using smaller datasets collected during real deployments [62, 65]. This strategy enables learning algorithms to capture general physical relationships in simulation while adapting to environment-specific characteristics during field operation. Developing accurate and scalable digital twin frameworks for underwater networks therefore represents an important research direction for enabling robust, adaptive, and continuously evolving AI-driven underwater communication systems [63].

8. CONCLUSIONS

The discussion in this paper underscores the growing need for underwater acoustic systems to move beyond rigid, model-driven designs toward architectures capable of adapting to environmental variability and operational uncertainty. The examples and system perspectives presented illustrate how data-driven processing can be employed not as a replacement for established acoustic principles, but as a complementary mechanism for advancing underwater applications toward more robust, reliable, scalable, and adaptive system designs. In doing so, such approaches support faster and more informed decision-making while strengthening operational resilience and system integrity.

Across sensing and communication tasks, the results highlight both the promise and the limitations of learned approaches. While data-driven methods enable improved perception and adaptive operation in highly non-stationary environments, their effectiveness remains closely coupled to data availability, environmental diversity, and deployment-specific constraints. In particular, generalization across sites and missions emerges as a central challenge, reinforcing the importance of adaptive, context-aware system design rather than static or site-specific model deployment.

The case study further demonstrates how acoustic inference can provide actionable situational awareness under realistic operating conditions, offering insight into how perception-driven processing can inform higher-level system responses. Although these results are preliminary, they illustrate the feasibility of integrating inference and adaptation within a unified underwater processing framework and motivate further exploration across broader datasets, longer deployments, and more diverse operational scenarios.

Looking ahead, continued progress in underwater acoustic systems will depend on deeper integration between sensing, inference, and adaptive communication, guided by practical considerations such as energy efficiency, computational constraints, and robustness to environmental change. By articulating how sensing, inference, and adaptive operation interact within a unified architectural perspective, this work provides a structured path forward for future research and for system development efforts aimed at enabling resilient, secure, and autonomous underwater platforms capable of sustained operation in complex and dynamic acoustic environments.

REFERENCES

[1] Francesco Restuccia and Tommaso Melodia. “Deep learning at the physical layer: System challenges and applications to 5G and beyond”. In: *IEEE Communications Magazine* 58.10 (2020), pp. 58–64.

[2] Amani Al-Shawabka, Francesco Restuccia, Salvatore D’Oro, Tong Jian, Bruno Costa Rendon, Nasim Soltani, Jennifer Dy, Stratis Ioannidis, Kaushik Chowdhury, and Tommaso Melodia. “Exposing the Fingerprint: Dissecting the Impact of the Wireless Channel on Radio Fingerprinting”. In: *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*. 2020, pp. 646–655. DOI: 10.1109/INFOCOM41043.2020.9155259.

[3] Francesco Restuccia, Salvatore D’Oro, Amani Al-Shawabka, Mauro Belgiovine, Luca Angioloni, Stratis Ioannidis, Kaushik Chowdhury, and Tommaso Melodia. “DeepRadioID: Real-time channel-resilient optimization of deep learning-based radio fingerprinting algorithms”. In: *Proceedings of the twentieth ACM international symposium on mobile ad hoc networking and computing*. 2019, pp. 51–60.

[4] Amani Al-Shawabka, Philip Pietraski, Sudhir B Pattar, Pedram Johari, and Tommaso Melodia. "SignCRF: Scalable Channel-Agnostic Data-Driven Radio Authentication System". In: *IEEE Transactions on Mobile Computing* 24.10 (2025), pp. 9383–9394. DOI: 10.1109/TMC.2025.3564556.

[5] Amani Al-Shawabka, Philip Pietraski, Sudhir B Pattar, Francesco Restuccia, and Tommaso Melodia. "DeepLoRa: Fingerprinting LoRa Devices at Scale Through Deep Learning and Data Augmentation". In: *Proceedings of the Twenty-second International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*. 2021.

[6] Shilian Zheng, Shichuan Chen, Peihan Qi, Huaji Zhou, and Xiaoniu Yang. "Spectrum sensing based on deep learning classification for cognitive radios". In: *China Communications* 17.2 (2020), pp. 138–148.

[7] Fan Meng, Peng Chen, Lenan Wu, and Xianbin Wang. "Automatic modulation classification: A deep learning enabled approach". In: *IEEE Transactions on Vehicular Technology* 67.11 (2018), pp. 10760–10772.

[8] Ren-Hung Hwang, Min-Chun Peng, Chien-Wei Huang, Po-Ching Lin, and Van-Linh Nguyen. "An unsupervised deep learning model for early network traffic anomaly detection". In: *IEEE Access* 8 (2020), pp. 30387–30399.

[9] Yaohua Sun, Mugen Peng, Yangcheng Zhou, Yuzhe Huang, and Shiwen Mao. "Application of machine learning in wireless networks: Key techniques and open issues". In: *IEEE Communications Surveys & Tutorials* 21.4 (2019), pp. 3072–3108.

[10] Hao Huang, Yang Peng, Jie Yang, Wenchao Xia, and Guan Gui. "Fast beamforming design via deep learning". In: *IEEE Transactions on Vehicular Technology* 69.1 (2019), pp. 1065–1069.

[11] Bomin Mao, Fengxiao Tang, Zubair Md Fadlullah, Nei Kato, Osamu Akashi, Takeru Inoue, and Kimihiro Mizutani. "A novel non-supervised deep-learning-based network traffic control method for software defined wireless networks". In: *IEEE Wireless Communications* 25.4 (2018), pp. 74–81.

[12] Jinglin Li, Guiyang Luo, Nan Cheng, Quan Yuan, Zhiheng Wu, Shang Gao, and Zhihan Liu. "An end-to-end load balancer based on deep learning for vehicular network traffic control". In: *IEEE Internet of Things Journal* 6.1 (2018), pp. 953–966.

[13] Francesco Davide Calabrese, Li Wang, Euhanna Ghadimi, Gunnar Peters, Lajos Hanzo, and Pablo Soldati. "Learning radio resource management in RANs: Framework, opportunities, and challenges". In: *IEEE Communications Magazine* 56.9 (2018), pp. 138–145.

[14] Daniel Uvaydov, Deniz Unal, Kerem Enhos, Emrecan Demirors, and Tommaso Melodia. "Sonair: Real-time deep learning for underwater acoustic spectrum sensing and classification". In: *2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)*. IEEE. 2023, pp. 9–16.

[15] Amani Al-Shawabka, Kerem Enhos, Emrecan Demirors, and Tommaso Melodia. "Real-Time Modulation Classification in Underwater Acoustic Channels Using Deep Learning". In: *OCEANS 2025-Great Lakes*. IEEE. 2025, pp. 1–7.

[16] Milica Stojanovic. "On the relationship between capacity and distance in an underwater acoustic communication channel". In: *ACM SIGMOBILE Mobile Computing and Communications Review* 11.4 (2007), pp. 34–43.

[17] James Preisig. "Acoustic propagation considerations for underwater acoustic communications network development". In: *ACM SIGMOBILE Mobile Computing and Communications Review* 11.4 (2007), pp. 2–10.

[18] T. Melodia, H. Kulhandjian, L. Kuo, and E. Demirors. "Advances in Underwater Acoustic Networking". In: *Mobile Ad Hoc Networking: Cutting Edge Directions*. Ed. by S. Basagni, M. Conti, S. Giordano, and I. Stojmenovic. 2nd. Inc., Hoboken, NJ: John Wiley and Sons, 2013, pp. 804–852.

[19] Nguyen Cong Luong, Dinh Thai Hoang, Shimin Gong, Dusit Niyato, Ping Wang, Ying-Chang Liang, and Dong In Kim. "Applications of Deep Reinforcement Learning in Communications and Networking: A Survey". In: *IEEE Communications Surveys Tutorials* 21.4 (2019), pp. 3133–3174.

[20] Chaoyun Zhang, Paul Patras, and Hamed Haddadi. "Deep learning in mobile and wireless networking: A survey". In: *IEEE Communications surveys & tutorials* 21.3 (2019), pp. 2224–2287.

[21] Qian Mao, Fei Hu, and Qi Hao. "Deep learning for intelligent wireless networks: A comprehensive survey". In: *IEEE Communications Surveys & Tutorials* 20.4 (2018), pp. 2595–2621.

[22] Linglong Dai, Ruicheng Jiao, Fumiyuki Adachi, H. Vincent Poor, and Lajos Hanzo. "Deep Learning for Wireless Communications: An Emerging Interdisciplinary Paradigm". In: *IEEE Wireless Communications* 27.4 (2020), pp. 133–139. DOI: 10.1109/MWC.001.1900491.

[23] Yufang Li, Liliang Zhang, Kai Wang, Lingwei Xu, and T Aaron Gulliver. "Underwater acoustic intelligent spectrum sensing with multimodal data fusion: An Mul-YOLO approach". In: *Future Generation Computer Systems* (2025), p. 107880.

[24] Weilong Zhang, Xinghai Yang, Changli Leng, Jingjing Wang, and Shiwen Mao. "Modulation Recognition of Underwater Acoustic Signals Using Deep Hybrid Neural Networks". In: *IEEE Transactions on Wireless Communications* 21.8 (2022), pp. 5977–5988. DOI: 10.1109/TWC.2022.3144608.

[25] Pushpa Kotipalli, Adi Surendra Mohanraju M., and Praveena Vardhanapu. "Frame Boundary Detection and Deep Learning-Based Doppler Shift Estimation for FBMC/OQAM Communication System in Underwater Acoustic Channels". In: *IEEE Access* 10 (2022), pp. 17590–17608. DOI: 10.1109/ACCESS.2022.3148410.

[26] Yung-Ting Hsieh, Zhuoran Qi, and Dario Pompili. "ML-based Joint Doppler Estimation and Compensation in Underwater Acoustic Communications". In: *Proceedings of the 16th International Conference on Underwater Networks & Systems*. WUWNet '22. Boston, MA, USA: Association for Computing Machinery, 2022. ISBN: 9781450399524. DOI: 10.1145/3567600.3568139. URL: https://doi.org/10.1145/3567600.3568139.

[27] Rongkun Jiang, Xuetian Wang, Shan Cao, Jiafei Zhao, and Xiaoran Li. "Deep Neural Networks for Channel Estimation in Underwater Acoustic OFDM Systems". In: *IEEE Access* 7 (2019), pp. 23579–23594. DOI: 10.1109/ACCESS.2019.2899990.

[28] Huaiyin Lu, Ming Jiang, and Julian Cheng. "Deep Learning Aided Robust Joint Channel Classification, Channel Estimation, and Signal Detection for Underwater Optical Communication". In: *IEEE Transactions on Communications* 69.4 (2021), pp. 2290–2303. DOI: 10.1109/TCOMM.2020.3046659.

[29] Lianyou Jing, Chaofan Dong, Chengbing He, Wentao Shi, and Hongxi Yin. "Adaptive Modulation and Coding for Underwater Acoustic Communications Based on Data-Driven Learning Algorithm". In: *Remote Sensing* 14.23 (2022). ISSN: 2072-4292. DOI: 10.3390/rs14235959. URL: https://www.mdpi.com/2072-4292/14/23/5959.

[30] Yinsheng Liu, Zhenhui Tan, Hongjie Hu, Leonard J Cimini, and Geoffrey Ye Li. "Channel estimation for OFDM". In: *IEEE Communications Surveys & Tutorials* 16.4 (2014), pp. 1891–1908.

[31] A. Ur Rehman, Laura Galluccio, and Giacomo Morabito. "AI-Driven Adaptive Communications for Energy-Efficient Underwater Acoustic Sensor Networks". In: *Sensors* 25.12 (2025). ISSN: 1424-8220. DOI: 10.3390/s25123729. URL: https://www.mdpi.com/1424-8220/25/12/3729.

[32] Yihao Zhao, Yougan Chen, Wenxiang Zhang, Xuchen Wang, Yi Tao, Chao Li, and Xiaomei Xu. "A Multiagent Reinforcement Learning-Based MAC Protocol for Clustered Multihop Underwater Acoustic Sensor Networks". In: *IEEE Sensors Journal* 25.15 (2025), pp. 30034–30046. DOI: 10.1109/JSEN.2025.3581990.

[33] Juan Li, Wenkai Xu, Limiao Deng, Ying Xiao, Zhongzhi Han, and Haiyong Zheng. "Deep learning for visual recognition and detection of aquatic animals: A review". In: *Reviews in Aquaculture* 15.2 (2023), pp. 409–433.

[34] Sai Sree Nathala, Rakesh Reddy Yakkati, Aveen Dayal, M. Sabarimalai Manikandan, Jing Zhou, and Linga Reddy Cenkeramaddi. "Vessel Type Classification Utilizing Underwater Acoustic Data and Deep Learning". In: *2024 IEEE 19th Conference on Industrial Electronics and Applications (ICIEA)*. 2024, pp. 1–6. DOI: 10.1109/ICIEA61579.2024.10665252.

[35] Muhammad Irfan, ZHENG Jiangbin, Shahid Ali, Muhammad Iqbal, Zafar Masood, and Umar Hamid. "DeepShip: An underwater acoustic benchmark dataset and a separable convolution based autoencoder for classification". In: *Expert Systems with Applications* 183 (2021), p. 115270.

[36] https://www.hydronet.tech/.

[37] Xuesong Yu, Li Li, Jingwei Yin, Mengqi Shao, and Xiao Han. "Modulation Pattern Recognition of Non-cooperative Underwater Acoustic Communication Signals Based on LSTM Network[*]". In: *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP)*. 2019, pp. 1–5. DOI: 10.1109/ICSIDP47821.2019.9173133.

[38] Yan Wang, Yiheng Jin, Hao Zhang, Qian Lu, Conghui Cao, Zhanliang Sang, and Mei Sun. "Underwater Communication Signal Recognition Using Sequence Convolutional Network". In: *IEEE Access* 9 (2021), pp. 46886–46899. DOI: 10.1109/ACCESS.2021.3067070.

[39] Ding Li-Da, Wang Shi-Lian, and Zhang Wei. "Modulation Classification of Underwater Acoustic Communication Signals Based on Deep Learning". In: *2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)*. 2018, pp. 1–4. DOI: 10.1109/OCEANSKOBE.2018.8559101.

[40] Weilong Zhang, Xinghai Yang, Changli Leng, Jingjing Wang, and Shiwen Mao. "Modulation Recognition of Underwater Acoustic Signals using Deep Hybrid Neural Networks". In: *IEEE Transactions on Wireless Communications* (2022), pp. 1–1. DOI: 10.1109/TWC.2022.3144608.

[41] Kerem Enhos, Deniz Unal, Joe Turco, Emrecan Demirors, and Tommaso Melodia. "Marena: Sdr-based testbed for underwater wireless communication and networking research". In: *Proceedings of The 17th ACM Workshop on Wireless Network Testbeds, Experimental evaluation & Characterization*. 2023, pp. 80–87.

[42] Xiaojing Tian, Xiujuan Du, Xiuxiu Liu, Lijuan Wang, and Lei Zhao. "A Low-Delay Source-Location-Privacy Protection Scheme With Multi-AUV Collaboration for Underwater Acoustic Sensor Networks". In: *IEEE Sensors Journal* (2025).

[43] Andrea Tiranti, Paolo Di Lillo, Francesco Wanderlingh, Enrico Simetti, Marco Baglietto, Gianluca Antonelli, and Giovanni Indiveri. "Motion Optimization Strategy for Passive Acoustic Monitoring With a Team of AUVs Considering Intermittent Communication". In: *IEEE Journal of Oceanic Engineering* (2025).

[44] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett. "Multicarrier Communication Over Underwater Acoustic Channels With Nonuniform Doppler Shifts". In: *IEEE Journal of Oceanic Engineering* 33.2 (Apr. 2008), pp. 198–209.

[45] Patrick J. Donnelly, Kerem Enhos, Emrecan Demirors, Deniz Unal, and Tommaso Melodia. "Comparison of Doppler Estimation Methods in Mobile Underwater Acoustic Communication". In: *OCEANS 2024 - Halifax*. 2024, pp. 1–10. DOI: 10.1109/OCEANS55160.2024.10753822.

[46] E. Demirors, G. Sklivanitis, G. E. Santagati, T. Melodia, and S. N. Batalama. "A High-Rate Software-Defined Underwater Acoustic Modem with Real-Time Adaptation Capabilities". In: *IEEE Access* PP.99 (2018), pp. 1–1. DOI: 10.1109/ACCESS.2018.2815026.

[47] Hongzi Mao, Mohammad Alizadeh, Ishai Menache, and Srikanth Kandula. "Resource management with deep reinforcement learning". In: *Proceedings of the 15th ACM workshop on hot topics in networks*. 2016, pp. 50–56.

[48] Jundi Ding, Wanlong Zhao, and Weixiao Meng. "Distributed Switching MAAC-Based Energy-Efficient Underwater Data Collection for Multi-AUV System". In: *IEEE Internet of Things Journal* (2025).

[49] Andrea Tiranti, Francesco Wanderlingh, Enrico Simetti, Marco Baglietto, Giovanni Indiveri, and Antonio Pascoal. "A distributed motion planning approach to cooperative underwater acoustic source tracking". In: *Ocean Engineering* 344 (2026), p. 123305.

[50] Albert F. Harris, Milica Stojanovic, and Michele Zorzi. "When Underwater Acoustic Nodes Should Sleep with One Eye Open: Idle-Time Power Management in Underwater Sensor Networks". In: *Proceedings of the 1st ACM International Workshop on Underwater Networks*. WUWNet '06. Los Angeles, CA, USA: Association for Computing Machinery, 2006, pp. 105–108. ISBN: 1595934847. DOI: 10.1145/1161039.1161061. URL: https://doi.org/10.1145/1161039.1161061.

[51] Raffaele Guida, Emrecan Demirors, Neil Dave, and Tommaso Melodia. "Underwater Ultrasonic Wireless Power Transfer: A Battery-less Platform for the Internet of Underwater Things". In: *IEEE Transactions on Mobile Computing* (2020).

[52] En Li, Liekang Zeng, Zhi Zhou, and Xu Chen. "Edge AI: On-Demand Accelerating Deep Neural Network Inference via Edge Computing". In: *IEEE Transactions on Wireless Communications* 19.1 (2020), pp. 447–457.

[53] Laela Sayigh, Mary Ann Daher, Julie Allen, Helen Gordon, Katherine Joyce, Claire Stuhlmann, and Peter Tyack. "The watkins marine mammal sound database: An online, freely accessible resource". In: *Proceedings of Meetings on Acoustics*. Vol. 27. 1. AIP Publishing. 2016.

[54] Hao Ye, Le Liang, Geoffrey Ye Li, and Biing-Hwang Juang. "Deep Learning-Based End-to-End Wireless Communication Systems With Conditional GANs as Unknown Channels". In: *IEEE Transactions on Wireless Communications* 19.5 (2020), pp. 3133–3143. DOI: 10.1109/TWC.2020.2970707.

[55] Alexander Felix, Sebastian Cammerer, Sebastian Dörner, Jakob Hoydis, and Stephan ten Brink. "OFDM-Autoencoder for End-to-End Learning of Communications Systems". In: *CoRR* abs/1803.05815 (2018). arXiv: 1803.05815. URL: http://arxiv.org/abs/1803.05815.

[56] Vishnu Raj and Sheetal Kalyani. "Backpropagating Through the Air: Deep Learning at Physical Layer Without Channel Models". In: *IEEE Communications Letters* 22.11 (2018), pp. 2278–2281. DOI: 10.1109/LCOMM.2018.2868103.

[57] Yang Yang, Yang Li, Wuxiong Zhang, Fei Qin, Pengcheng Zhu, and Cheng-Xiang Wang. "Generative-adversarial-network-based wireless channel modeling: Challenges and opportunities". In: *IEEE Communications Magazine* 57.3 (2019), pp. 22–27.

[58] Hao Ye, Geoffrey Ye Li, Biing-Hwang Fred Juang, and Kathiravetpillai Sivanesan. "Channel agnostic end-to-end learning based communication systems with conditional GAN". In: *2018 IEEE Globecom Workshops (GC Wkshps)*. IEEE. 2018, pp. 1–5.

[59] Dohyun Kwon, Joohyung Jeon, Soohyun Park, Joongheon Kim, and Sungrae Cho. "Multiagent DDPG-Based Deep Learning for Smart Ocean Federated Learning IoT Networks". In: *IEEE Internet of Things Journal* 7.10 (2020), pp. 9895–9903. DOI: 10.1109/JIOT.2020.2988033.

[60] István Hegedűs, Gábor Danner, and Márk Jelasity. "Gossip learning as a decentralized alternative to federated learning". In: *IFIP International Conference on Distributed Applications and Interoperable Systems*. Springer. 2019, pp. 74–90.

[61] Michael B Porter. "The bellhop manual and user's guide: Preliminary draft". In: *Heat, Light, and Sound Research, Inc., La Jolla, CA, USA, Tech. Rep* 260 (2011), p. 70.

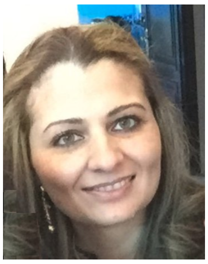
[62] Josh Tobin, Rachel Fong, Alex Ray, Jonas Schneider, Wojciech Zaremba, and Pieter Abbeel. "Domain randomization for transferring deep neural networks from simulation to the real world". In: *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2017, pp. 23–30.

[63] Adil Rasheed, Omer San, and Trond Kvamsdal. "Digital twin: Values, challenges and enablers from a modeling perspective". In: *IEEE access* 8 (2020), pp. 21980–22012.

[64] Zhihan Lv, Dongliang Chen, Hailin Feng, Wei Wei, and Haibin Lv. "Artificial intelligence in underwater digital twins sensor networks". In: *ACM Transactions on Sensor Networks (TOSN)* 18.3 (2022), pp. 1–27.

[65] Wenshuai Zhao, Jorge Peña Queralta, and Tomi Westerlund. "Sim-to-real transfer in deep reinforcement learning for robotics: a survey". In: *2020 IEEE symposium series on computational intelligence (SSCI)*. IEEE. 2020, pp. 737–744.

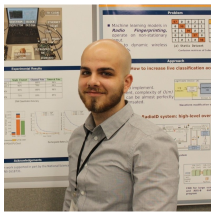
AUTHORS



Amani Alshawabka Amani Alshawabka is an AI and wireless systems engineer at Bionet Sonar. She received her Ph.D. in Computer Engineering from Northeastern University, Boston, MA, USA, in 2024. She also holds a Master's degree in Computer Engineering from Northeastern University (2019), an MBA in Marketing, and a Bachelor's degree in Communication Engineering. Her research interests focus on leveraging artificial intelligence and deep learning techniques to enhance wireless communication systems, with applications spanning underwater acoustic sensing and communication, as well as biomedical signal analysis. Her research contributions were recognized with the Northeastern ECE Excellence in Research Award in 2023.

Kerem Enhos received the Ph.D. degree in Electrical and Computer Engineering from Northeastern University in 2023, and the B.S. and M.S. degrees in Electrical and Electronics Engineering from Bilkent University in 2017 and 2019, respectively. He is currently a Senior Research and Development Engineer at Hydronet, specializing in underwater wireless communications, networking, and systems. His expertise spans acoustic waves and devices, transducers, visible light communications, and software-defined networking.

Emrecan Demirors received the Ph.D. degree in Electrical and Computer Engineering from Northeastern University, Boston, MA, USA, in 2017. He is Director of R&D at Bionet Sonar and a part-time Research Assistant Professor with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA. Dr. Demirors' current research interests include Internet of Things, software-defined wireless networks, cognitive radio networks, underwater communications and networks, and body-area networks. Dr. Demirors is an Associate Editor for the IEEE Access.



Daniel Uvaydov received a Ph.D. and M.S. degree in Electrical and Computer Engineering from Northeastern University, Boston, MA in 2023, and a B.S. degree in Electrical Engineering from University at Buffalo, Buffalo, NY in 2018. He is currently a Senior Modem Systems Engineer at Qualcomm working on developing commercial modems. Dr. Uvaydov has expertise at the intersection of AI, machine learning, signal processing and wireless communications systems.

Deniz Unal received the Ph.D. degree in electrical engineering from Northeastern University, Boston, MA, USA, in 2025, and the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2014 and 2018, respectively. He is currently with Abside Networks, where he develops DSP algorithms and software for Massive MIMO O-RAN radio units. His research interests include wireless networks, underwater acoustic communications, software-defined networking, and Open RAN architectures.

TOMMASO MELODIA received the Ph.D. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 2007. He is the President, Chief Scientist, and Founder of Hydronet. He is also the William L. Smith Professor with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, USA, and the Founding Director of the Institute for Intelligent Networked Systems. His research has been supported by over $100M in funds by U.S. federal agencies, including the National Science Foundation, the Air Force Research Laboratory, the Office of Naval Research, DARPA, the Army Research Laboratory, the Office of the Secretary of Defense. He is the Editor in Chief of Computer Networks, a Fellow of the IEEE (for contributions to acoustic communications), a Fellow of the National Academy of Inventors (NAI) and a Distinguished Member of the ACM. He sits in the Board of Directors of the AI-RAN Alliance and of the OpenAirInterface Software Alliance.