

Article

Reinforcement Learning Technique for Self-Healing FBG Sensor Systems in Optical Wireless Communication Networks

Rénauld A. Dellimore, Jyun-Wei Li, Hung-Wei Huang, Amare Mulatie Dehnaw , Cheng-Kai Yao , Pei-Chung Liu and Peng-Chun Peng *

Department of Electro-Optical Engineering, National Taipei University of Technology, Taipei 10608, Taiwan

* Correspondence: pcpeng@ntut.edu.tw; Tel.: +886-2-2771-2171 (ext. 4671)

Abstract

This paper proposes a large-scale, self-healing multipoint fiber Bragg grating (FBG) sensor network that employs reinforcement learning (RL) techniques to enhance the resilience and efficiency of optical wireless communication networks. The system features a mesh-structured, self-healing ring-mesh architecture employing 2×2 optical switches, enabling robust multipoint sensing and fault tolerance in the event of one or more link failures. To further extend network coverage and support distributed deployment scenarios, free-space optical (FSO) links are integrated as wireless optical backhaul between central offices and remote monitoring sites, including structural health, renewable energy, and transportation systems. These FSO links offer high-speed, line-of-sight connections that complement physical fiber infrastructure, particularly in locations where cable deployment is impractical. Additionally, RL-based artificial intelligence (AI) techniques are employed to enable intelligent path selection, optimize routing, and enhance network reliability. Experimental results confirm that the RL-based approach effectively identifies optimal sensing paths among multiple routing options, both wired and wireless, resulting in reduced energy consumption, extended sensor network lifespan, and improved transmission delay. The proposed hybrid FSO–fiber self-healing sensor system demonstrates high survivability, scalability, and low routing path loss, making it a strong candidate for future services and mission-critical applications.

Keywords: sensor network; fiber Bragg grating; self-healing ring-mesh structures; artificial intelligence; deep learning; reinforcement learning; optical wireless communication network



Academic Editor: Alessandro Lo Schiavo

Received: 21 November 2025

Revised: 13 January 2026

Accepted: 16 January 2026

Published: 19 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Modern sensor systems include numerous sensors and can serve as key components of Internet of Things (IoT) infrastructure, performing multipoint measurements across diverse applications such as structural health monitoring, renewable energy systems, and transportation networks [1,2]. In modern IoT deployments, sensing nodes are often distributed across large geographical areas and must operate continuously for extended periods. These sensors, including strain, motion, moisture, temperature, light, and air quality sensors, continuously gather data from the environment, enabling more intelligent decision-making, condition-based maintenance, and the early warning of potential failures [1,3,4]. The ability to reliably collect and transmit sensing data from multiple locations is therefore a fundamental requirement for next-generation intelligent monitoring systems.

Among these sensing technologies, fiber Bragg grating (FBG) sensors are particularly favored due to their lightweight structure, immunity to electromagnetic interference, high

multiplexing capability, and suitability for remote sensing in harsh environments [2,5–8]. Compared with conventional electrical sensors, FBG sensors offer superior resistance to electromagnetic noise and corrosion, as well as high sensitivity and long-term stability. These advantages make FBG sensors especially suitable for large-scale deployments in civil infrastructure, aerospace structures, and energy facilities, where harsh environmental conditions and limited accessibility pose significant challenges [6–9].

FBG sensors operate by detecting shifts in the Bragg wavelength in response to changes in strain or temperature, enabling accurate, real-time monitoring of physical parameters. However, in large-scale FBG sensor networks, system survivability becomes increasingly complex due to the potential for fiber link or node failures [8–12]. Such failures are particularly problematic in environments such as dams, aircraft, tunnels, and offshore structures, where physical access for repair is restricted, and maintenance costs are high. In these cases, sensing is often blocked beyond the fault point, especially in bus or linear topologies, critically impacting system reliability and data availability [6,10,11].

To address these challenges, self-healing FBG sensor networks have been proposed to maintain sensing functionality even when faults occur [5,13–15]. In conventional self-healing solutions, optical switches (OSs) are deployed at remote nodes (RNs) to reconfigure the sensing path upon link failure dynamically. By rerouting optical signals through alternative paths, these networks can restore connectivity without manual intervention. However, the use of optical switches increases system cost, control complexity, and power consumption. Furthermore, the use of optical splitters introduces additional insertion loss, which limits network scalability as the number of sensing points increases [13–18]. In addition, learning-assisted fault detection and recovery have been investigated to improve adaptability and support shortest-path restoration in optical fiber sensor networks, enhancing survivability under dynamic failure conditions [19–22]. Physical obstacles to fiber installation, such as rough terrain, long distances, and temporary deployment requirements, also remain persistent issues in many practical applications.

Related optical networking studies have shown that wireless and hybrid optical technologies can enhance coverage and redundancy without requiring additional fiber deployment [23–26]. In this context, the free-space optical (FSO) links in this work are primarily intended to extend coverage and enable deployment in geographically challenging or hard-to-wire areas [23,24,26]. High-capacity FSO links have been demonstrated for mobile fronthaul and backhaul applications, providing high data rates and low latency while effectively bypassing physical obstacles [23,24,26]. Moreover, hybrid millimeter-wave and FSO radio-over-fiber systems have been reported to improve link reliability through coordinated mapping and signal combining techniques, thereby increasing resilience to atmospheric disturbances and alignment issues [23–25].

Fault-protected optical access networks that integrate fiber and FSO channels further enhance system survivability by providing alternative transmission paths when fiber links are damaged [27–30]. Bidirectional fiber-FSO architectures have also been demonstrated to support real-time data transmission and multipoint sensor interrogation, enabling efficient sensing data collection in distributed environments [27–29]. To extend coverage and introduce greater flexibility in distributed deployments, FSO links can be integrated as optical wireless backhaul connections between central monitoring hubs and remote monitoring sites [26–30]. These high-speed, line-of-sight FSO links complement existing fiber infrastructure by interconnecting multiple remote sensor subsystems in applications such as structural health monitoring, renewable energy systems, and transportation sensing platforms.

In parallel with survivable networking, learning-based optimization has been explored in optical and photonic systems to improve sensing quality and demodulation performance

in FBG-based networks [31–35]. Such approaches include deep learning for enhancing multiplexing capacity and measurement accuracy, self-supervised learning for separating distorted overlapping spectra, and machine-learning-assisted hybrid fiber–FSO sensing systems, as well as cloud-based deep learning and meta-learning methods for spectral processing and multichannel sensing [31–35].

Routing efficiency becomes a critical issue in hybrid fiber–FSO networks. Reinforcement learning (RL), a type of machine learning in which agents interact with the environment and learn optimal actions through trial and error, offers a robust solution for adaptive routing and decision-making [36]. Early RL-based routing studies demonstrated that learning agents can avoid congestion and optimize path selection in dynamic sensor and communication networks [37,38]. Subsequent research further refined RL-based routing approaches to improve convergence speed and routing stability [39,40]. More recent investigations have applied deep reinforcement learning (DRL) to complex networking problems and have developed DRL-based planning and control frameworks for network optimization [41–44]. Building upon these advances, this paper proposes a novel large-scale, self-healing FBG sensor network that integrates a ring-mesh fiber topology with optical switching and an FSO backhaul. An RL-based routing mechanism is introduced to optimize sensing path selection and dynamically respond to failures, achieving improved survivability, reduced transmission delay, enhanced scalability, and extended system lifespan. The proposed system concept is validated through simulation, and its key contributions are summarized as follows:

- A reliable ring-mesh fiber architecture designed for large-scale, multipoint FBG sensing with inherent self-healing capability.
- Integration of FBG sensors with FSO technology to enhance network scalability and address challenges caused by physical obstacles in fiber installation.
- An RL-based intelligent path selection mechanism that utilizes efficient routing techniques to improve system survivability, scalability, and transmission efficiency.

The remainder of this paper is organized as follows. Section 2 details the conceptual structure and operational principles of the multipoint sensing system. Section 3 presents the proposed methodology. Section 4 discusses the experimental and simulation results, and Section 5 concludes the study.

2. Operational Principles of the Multi-Point Sensing System

2.1. Conceptual Structure of Multi-Point Sensing System Using Reinforcement Learning

Figure 1 demonstrates a conceptual overview of how free-space optical links can be used to interconnect multiple distributed FBG sensor networks. A central monitoring hub serves as the control office (CO) and uses line-of-sight FSO backhaul links to reach remote nodes hosting ring-based FBG sensor subnets for different applications. This high-level view illustrates how disparate sensing systems, ranging from structural health monitoring to renewable-energy and transportation assets, could be brought under a single supervisory framework via optical wireless links, extending coverage and improving deployment flexibility.

The figure illustrates a multipoint sensing system based on the proposed quadrilateral mesh structure, which integrates mesh and ring configurations. The quadrilateral mesh structure FBG sensor system includes a CO and several FBG sensors distributed across different subnet ring sensing regions for strain sensing. The CO is responsible for supplying the light source and for monitoring sensing signals reflected from the sensor network in the sensing regions.

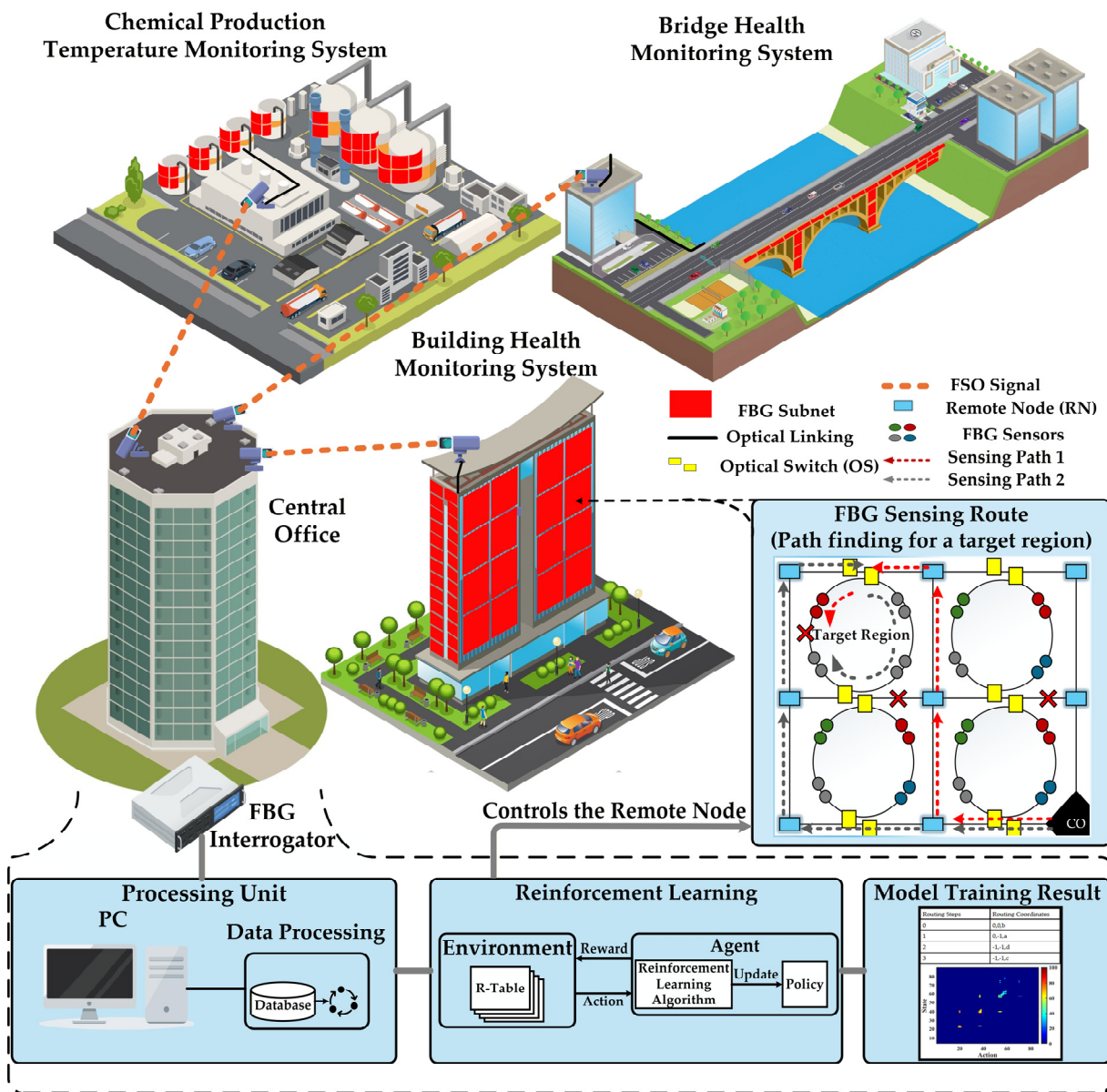


Figure 1. Multi-point sensing system based on the proposed quadrilateral mesh structure using RL.

The proposed sensing system comprises several sensing regions to improve the sensor system's survivability and capabilities. Within each fiber-sensing region, a ring subnet is deployed to serve multiple FBG sensors. As shown in Figure 1, eight FBG sensors are installed in each ring-region subnet, represented in different colors and connected to the mesh architecture via an RN. Two 2×2 optical switches are implemented in the RN. One optical switch has a parallel mode, and the second optical switch has a cross mode. The 2×2 optical switch in each ring subnet is controlled by a TDM signal to increase the number of sensors supported in the FBG sensor system.

2.2. Working Principle of the Experimental Setup

Figure 2a illustrates the structure of the remote node (RN), which consists of two cascaded 2×2 optical switches (OS) operating in either parallel or cross mode to control signal routing among four ports (1–4). Different combinations of these switch states define six distinct switching functions (Functions 1–6), each corresponding to a specific set of connected port pairs and establishing a unique transmission path through the RN.

These switching functions enable adaptive signal routing between the mesh and ring subnetworks according to network conditions. The correspondence between each function and its associated optical paths is explicitly shown in the right panel of Figure 2a, where red dashed lines indicate the enabled signal routes. The RL agent selects a switching function based on the current network state and reward feedback, ensuring reliable, cost-effective routing decisions.

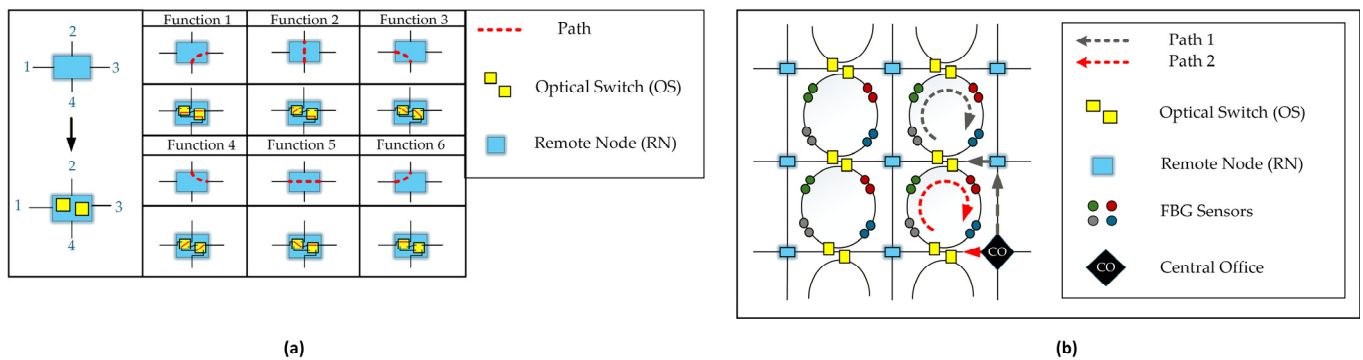


Figure 2. (a) Configuration of RN and all its possible switching states. (b) Implementation of the RN and switch for pathfinding.

Figure 2b illustrates how these switching functions are applied within the overall mesh-ring hybrid FBG sensor network. The network includes multiple ring subnets connected via RNs and OS, with a CO providing control and signal processing. Two alternative signal transmission paths, Path 1 (gray dashed line) and Path 2 (red dashed line), are shown as potential routing options from the CO to a given ring subnet.

These baseline paths demonstrate the network's ability to direct sensing signals through different RN configurations selectively. By controlling the RNs' switching states, the system can dynamically select the optimal data transmission path based on application requirements, network load, or signal priority. This architecture enables high configurability and scalability while supporting efficient resource use across the sensor network. The flexibility of switching paths via RNs enables better performance in sensing coverage, latency management, and network traffic balancing, without requiring physical rewiring or manual intervention.

Furthermore, to evaluate the self-healing performance of the proposed mesh-ring hybrid topology, various breakpoint scenarios are tested, as illustrated in Figure 3. By utilizing multiple redundant paths through mesh-interconnected ring subnetworks, the system maintains connectivity and uninterrupted sensing even when one or more links fail. These scenarios demonstrate robust fault recovery and routing flexibility while avoiding the excessive hardware and control complexity associated with full-mesh architectures. Supported by the RL-based routing mechanism, the network dynamically adapts to failures by identifying cost-effective alternative paths, thereby ensuring reliable operation, enhanced survivability, and efficient self-healing performance.

Each FBG ring subnet supports multiple sensors and employs time-division multiplexing (TDM) to distinguish sensing signals across different wavelengths. As illustrated in Figure 3, a partial failure occurs during signal transmission from the central office (CO) to the designated ring subnet, resulting in the loss of some sensing data. In this scenario, the CO and the target ring subnet serve as the source and destination, respectively. The initial sensing attempt along Path 2 (red line) is disrupted within the target ring after only two FBG sensors are detected. To restore sensing continuity, local ring-level reconfiguration is first applied; when this is insufficient due to link or sensor breakpoints, mesh-level reconfiguration between remote nodes is invoked through optical switching. Consequently,

the system reroutes the signal via Path 1 (gray line) and successfully detects the remaining six FBG sensors. This hierarchical reconfiguration process demonstrates the network's self-healing and adaptive behavior, while highlighting the challenge of identifying the most efficient routing path in large-scale mesh topologies. The above-mentioned challenge of identifying the most efficient route within a large-scale mesh topology is addressed in this study through the use of RL.

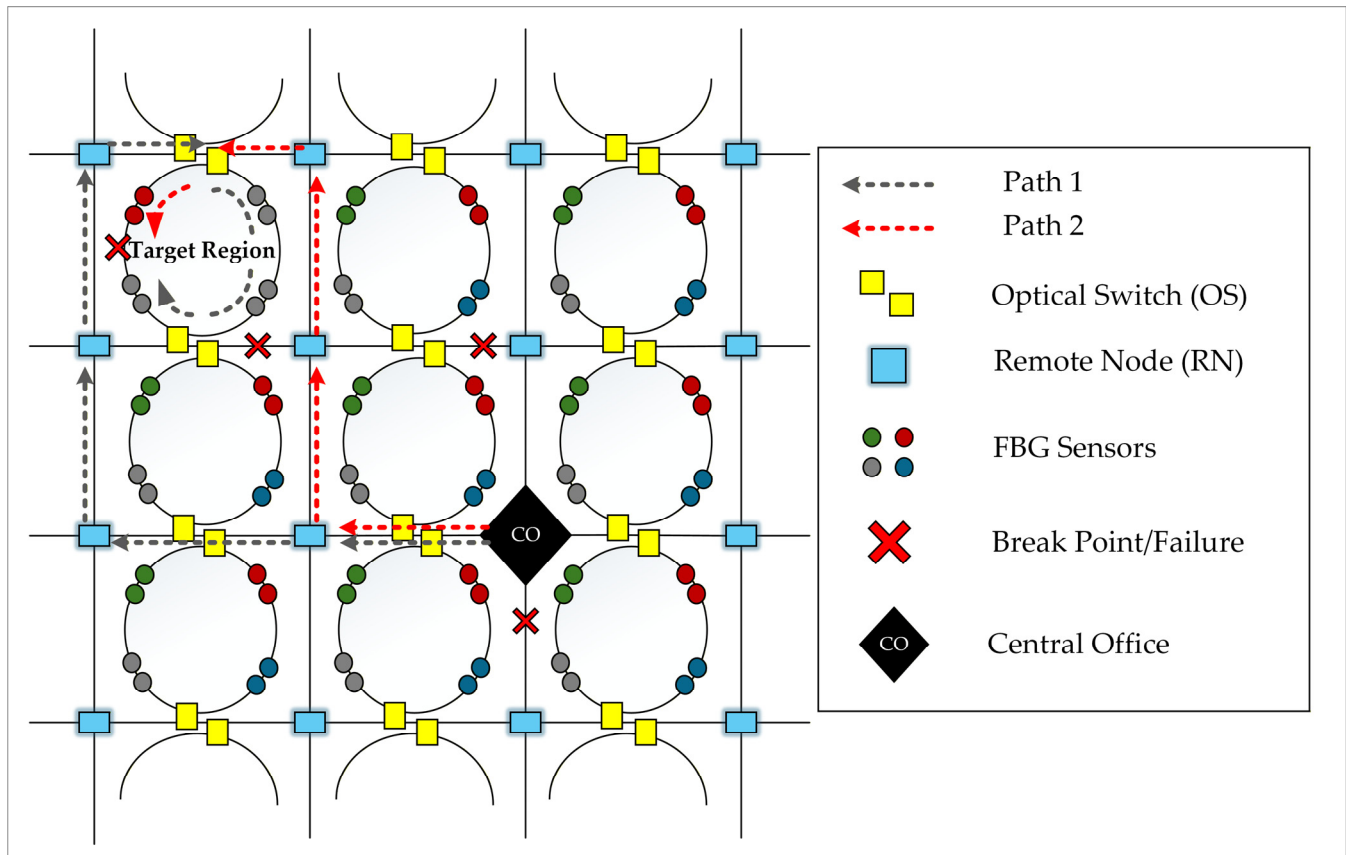


Figure 3. Situation indicated by line, when a target sensing region is selected by a time-division multiplexing (TDM) signal.

Figure 4 illustrates the experimental setup, where a quadrilateral mesh-based FBG sensing network is combined with an RL framework to identify optimal routing paths under multiple link faults. The sensor region contains interconnected FBG ring subnets that form a redundant mesh architecture, allowing self-healing when several links fail. Network states collected by the FBG interrogator are sent to the processor and then to the RL agent, which learns to bypass breakpoints and select the most reliable path to the target region. In this architecture, the CO is responsible for providing the light source and monitoring the sensing signals reflected from the sensor network. The CO includes an FBG interrogator, processor, and access to the distributed FBG sensors deployed across the mesh. The FBG interrogator consists of an optical source, an optical interpretation module, and a measurement unit. The FBG interrogator allows for dynamic configuration of the sensing paths toward different ring-based subnets in the sensor region.

To overcome the significant challenge of determining the most efficient route within a large-scale mesh topology, an RL algorithm is introduced to identify cost-effective routing paths from the CO to specific sensing regions. The RL framework consists of two main components: the agent and the environment. The agent perceives the environment (i.e., the

mesh sensor network) and learns how to act to improve sensing performance. Three core RL elements are defined:

- State: The condition of the network as perceived after each action (e.g., path success, failures, or partial sensing).
- Action: The decision to select a specific path or reconfigure a node.
- Reward: Feedback given based on the success of the action, often derived from cost metrics such as loss rate, energy consumption, or sensing delay.

The agent updates its routing policy over time to maximize rewards and minimize costs. This enables the sensor system to autonomously determine optimal sensing paths, even under failure conditions or varying sensing demands.

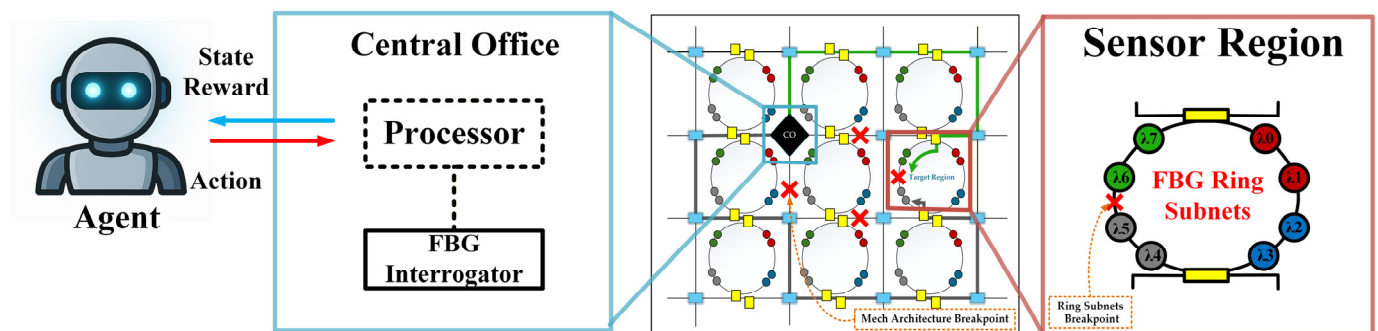


Figure 4. Reinforcement learning-based adaptive routing system in a large-scale FBG mesh-ring sensor network.

3. The Proposed Methodology

In a mesh-structured sensing network, the primary challenge is to identify routing paths that minimize optical power loss while ensuring reliable signal transmission. This routing problem can be effectively addressed using reinforcement learning (RL), which enables an agent to learn optimal decisions through continuous interaction with the network environment [39,42–44]. The problem is formulated as a Markov Decision Process (MDP) defined by the tuple (S, A, R) . Each state S represents the current node position, the availability of surrounding optical paths from the central office to the targeted sensing area, and the status of fiber breakpoints, thereby capturing network connectivity and accumulated routing cost. The action set A consists of admissible routing and switching decisions that guide the agent toward neighboring nodes. Based on the current state and the selected action, the agent transitions deterministically to the next state and receives feedback from the environment.

The reward function R plays a critical role in guiding the learning process by assigning positive rewards for successfully progressing toward or reaching valid sensing regions, and negative rewards for encountering fiber breakpoints, invalid paths, or incorrect optical switch entries [39,42]. These reward assignments encourage the selection of routes with lower fiber and switching losses while discouraging unfavorable paths. Over repeated interactions, the agent learns an optimal routing policy that maximizes cumulative reward. When all state-action-reward relationships are collected, they are integrated into a reward table (R -table), from which the corresponding Q -values are computed and stored in a Q -table. This Q -table records the accumulated access loss between neighboring sensing regions and provides a reference for selecting the optimal routing path. Consequently, the collaboration between the proposed RL-based routing algorithm and the mesh-structured sensing network can be optimized in terms of routing accuracy, computational efficiency, and overall system complexity [39,42].

The Q-learning-based optimal routing pathfinding technique enhances the survivability of the proposed self-healing mesh-structured sensor network under unexpected physical or environmental faults.

When fiber or sensor breakpoints occur, the mesh-structured network dynamically recomputes routing paths, while the Q-learning algorithm continuously updates routing information through reward feedback. Breakpoints in the surrounding ring subnets are assigned negative rewards, guiding the agent to avoid faulty states. Although this scenario is presented for clarity, the state-action formulation and reward design are independent of absolute node coordinates, allowing the framework to generalize to different mesh sizes, destinations, and failure patterns.

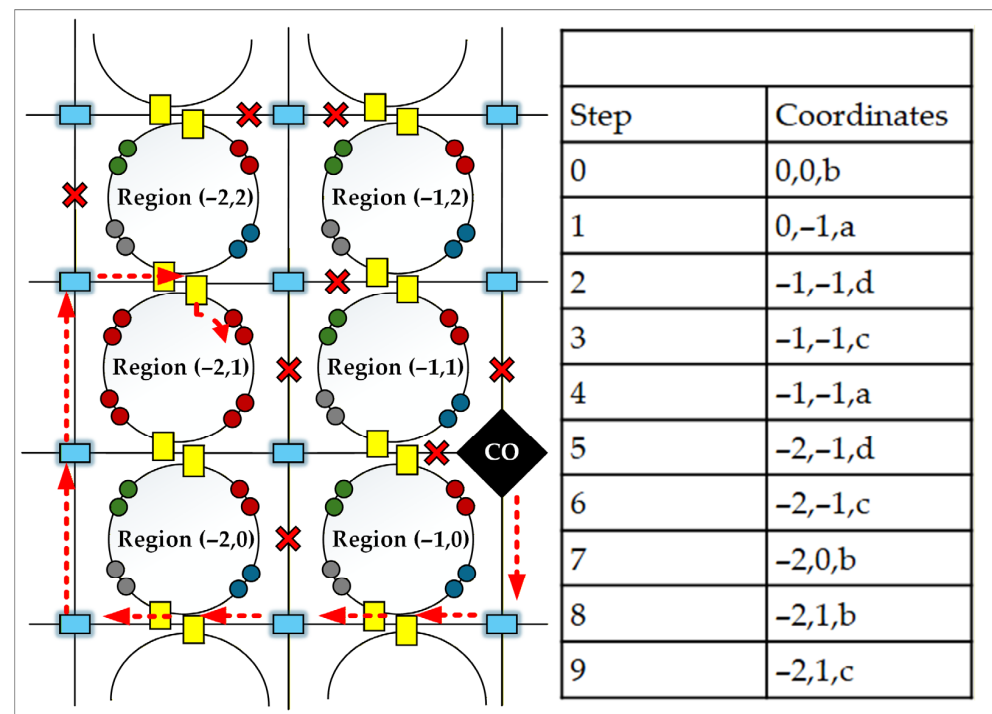


Figure 6. Conceptual diagram of the damaged sensing network and route visit intention with corresponding values.

As illustrated in Figure 6, broken links between the upper-left sensor subnet and the CO prevent access to the destination sensing region via those routes, resulting in negative rewards that discourage their selection during training. Consequently, the Q-learning algorithm identifies a viable alternative path to reach the target sensing region. The optimal routing path, summarized in table of Figure 6 from step 0 to step 9, is derived based on the action definitions in Figure 5 and the network connectivity shown in Figure 6. This learned path successfully avoids breakpoints and converges to the shortest available route to the destination sensing region at coordinates $(-2, 1)$, demonstrating the adaptive and topology-agnostic nature of the proposed routing framework.

Figure 7 illustrates the training process of the proposed RL-based method for optimal routing path selection in the FBG sensor network. The process begins with the data preprocessing phase, during which the sensing area and breakpoint information are defined. Based on this, reward adjustment is applied to different states and actions, which are integrated into an environmental reward table (R-table).

Additionally, environmental settings such as possible initialization states and action sets are adjusted. Once the environment is defined, the Q-learning algorithm is initiated. The state space is modeled using a Cartesian coordinate system, with the central office (CO) as the origin and the target sensing region at $(0, 0)$. Each RN is assigned a rectangular coordinate, and its connecting optical fibers are divided into sections a, b, and c, while the links between optical switches are referred to as section d. This structure helps the system avoid paths containing damaged RNs or fibers. Because FBG sensors reflect light, signal paths that would reflect to the CO before reaching the destination are excluded. After all valid states are defined, positive rewards are assigned near the destination to encourage shorter paths. Negative rewards are applied to paths involving backtracking or undesirable switch loops to discourage inefficient routing. This reward structure enhances path flexibility and supports failure scenarios, such as missing wavelengths caused by open circuits in ring subnets. When such break-points occur, the system evaluates four different paths, allowing the Q-learning agent to find the most cost-effective

sensing route. During training, the Q-learning loop begins by initializing the state and defining available actions. An ϵ -greedy policy balances exploration and exploitation when selecting the next state. The agent receives a reward and updates the Q-table and current state based on the calculated value function. This process continues until the convergence condition is met, indicating that the optimal path has been learned. The resulting routing path is then exported. To fine-tune the learning process, several parameters are defined:

- Number of iterations.
- Learning rate (α): A value of 0 means no learning occurs (Q-values remain unchanged), while $\alpha = 1$ allows full updates based on the most recent reward.
- Attenuation factor (γ): If $\gamma = 1$, future rewards are weighted equally with current ones, which can hinder present-focused decision-making. If $\gamma = 0$, only immediate rewards are considered, limiting long-term planning.
- Greedy coefficient (ϵ): Set initially to 1 for full exploration, and exponentially decayed (0.9 per interaction) to gradually increase exploitation.

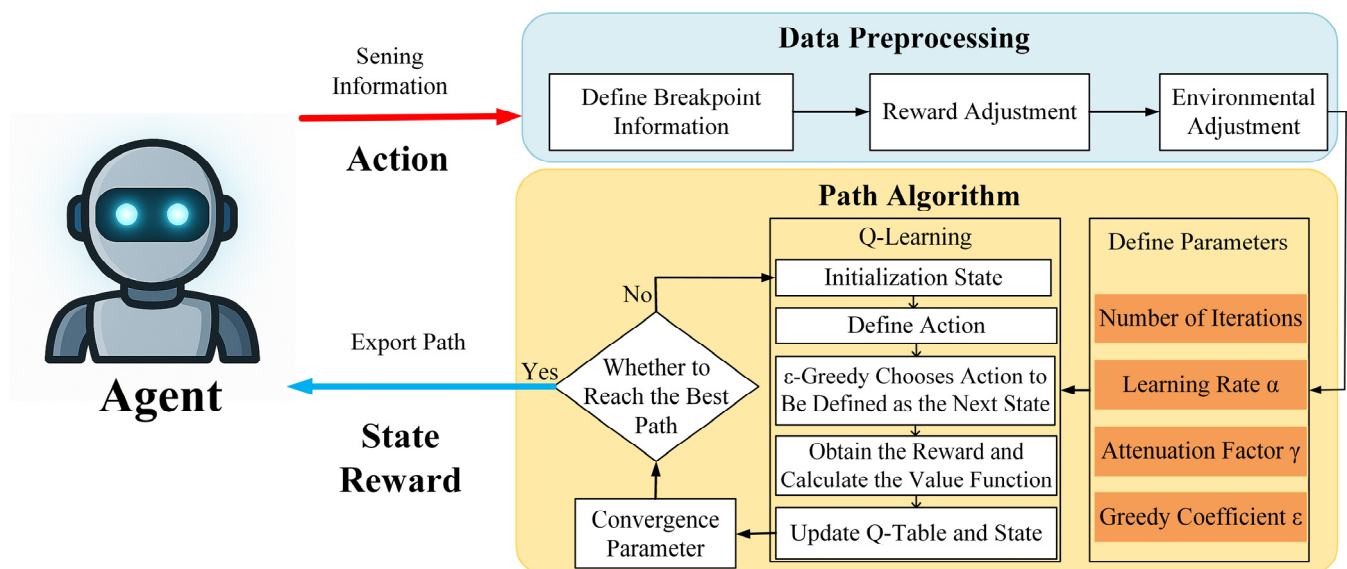


Figure 7. Flowchart of the Q-learning-based optimal routing pathfinding process.

Additionally, the initial learning rate is set to 0.7 and reduced over time using an attenuation factor of 0.9. A γ value of 0.7 is used to mitigate shortsightedness, balancing current and future rewards. In summary, the updated formula of the Q table is as follows:

$$Q(S_t, A_t) = Q(S_t, A_t) + \alpha [R_t + \gamma \max_a Q(S_t, A_t) - Q(S_t, A_t)], \quad (1)$$

where $Q(S_t, A_t)$ on the left is the updated Q value, $Q(S_t, A_t)$ on the right has not yet updated the previous Q value, $\max_a Q(S_t + 1, a)$ is the maximum action of the next state (state), R_t is the reward of the action made by the current state, α is learning rate in a value between 0 and 1, and γ is the attenuation factor.

Algorithm 1 presents the pseudocode for the Q-learning algorithm used in the routing pathfinding process.

The algorithm begins by initializing the Q-table and setting key parameters: exploration rate ($\epsilon = 0.8$), learning rate ($\alpha = 0.9$), and attenuation factor ($\gamma = 0.8$). For each episode, the agent starts from an initial state and selects actions using an ϵ -greedy policy that balances exploration and exploitation. Based on the chosen action, the agent transitions to the next state, receives a reward from the R-table, and updates the Q-value using the updated formula (1).

After each step, ε and α are decayed by a factor of 0.6 to shift from exploration to exploitation gradually. This process continues until the optimal routing path is learned.

Algorithm 1. Q-learning Algorithm Applied to this System

Step	Description
1	Initialize Q table: $\varepsilon = 0.8, \alpha = 0.9, \gamma = 0.8$
2	Repeat (for each episode):
3	Initialize s
4	Repeat (for each step in the episode):
5	Define A from R table
6	If rand () < ε , then:
7	Select a randomly from A
8	Else:
9	Choose a from s in Q table
10	Next s = a
11	Define reward from R table
12	Next a = $\arg \max_a Q(\text{next } s, \text{next } a)$
13	$Q(s, a) = Q(s, a) + \alpha [\text{reward} + \gamma Q(\text{next } s, \text{next } a) - Q(s, a)]$
14	s = next s
15	$\varepsilon = \varepsilon \times 0.6$
16	$\alpha = \alpha \times 0.6$

4. Experimental Results and Discussion

Figures 8–11 illustrate the evolution of the Q-table during Q-learning training with fixed hyperparameters ($\alpha = 0.9, \varepsilon = 0.8, \gamma = 0.8$). In this study, time is defined as the number of routing steps (state transitions or hops) undertaken by the agent during training rather than physical transmission delay. The vertical axis represents the 85 discrete network states, while the horizontal axis denotes the available routing actions. Color intensity reflects the magnitude of the Q-values, with warmer colors indicating more favorable routing decisions and cooler colors representing less preferred actions.

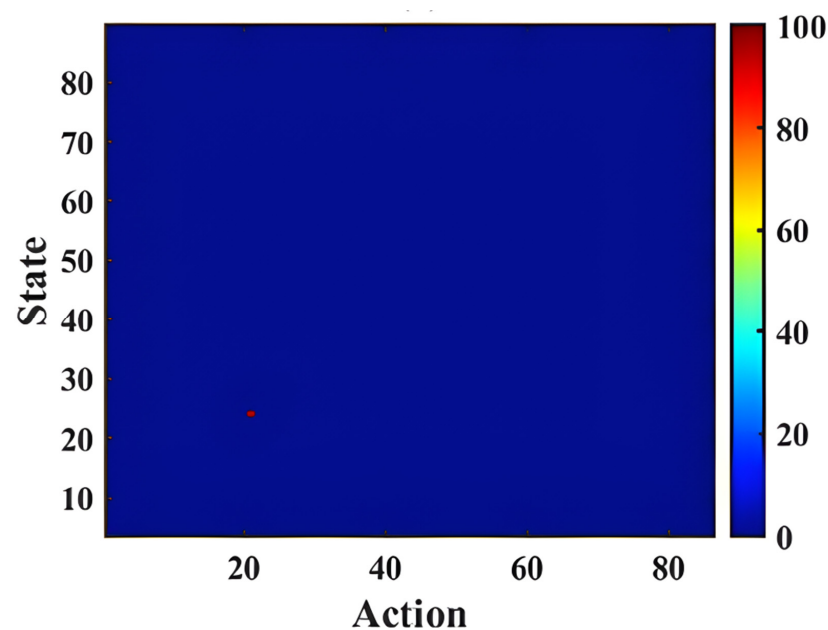


Figure 8. Q-table at 10 iterations (initial learning).

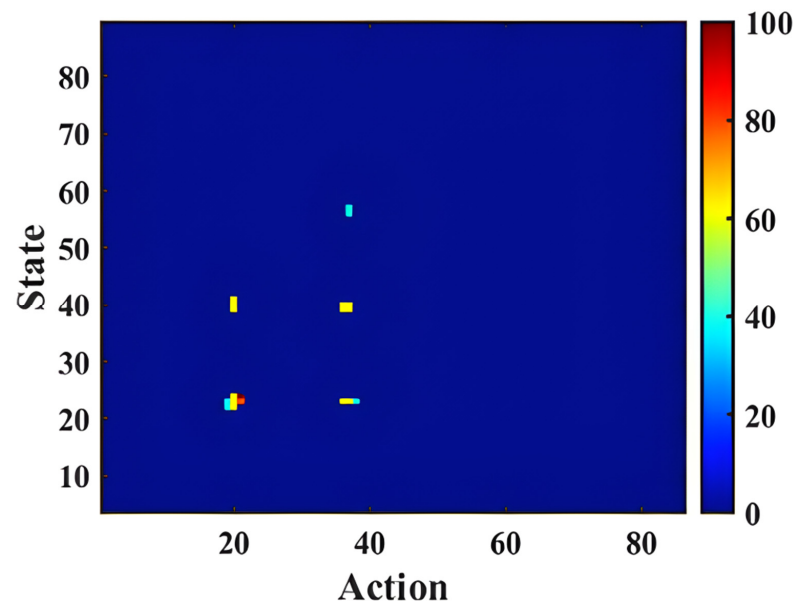


Figure 9. Q-table at 20 iterations (early improvement).

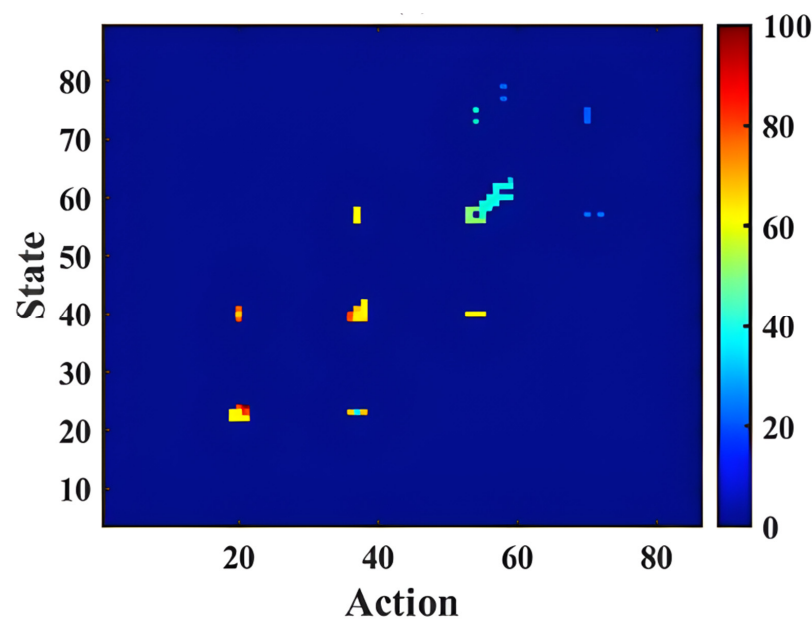


Figure 10. Q-table at 30 iterations (structured learning).

As training progresses over routing steps, the Q-table gradually converges. Initially, Q-values are relatively uniform due to extensive exploration encouraged by the high ϵ value. With continued learning, higher Q-values concentrate around specific state-action pairs, indicating that the agent has identified efficient routing paths requiring fewer hops to reach the target sensing region. In the later stages, the Q-table becomes more structured, confirming stable policy formation and demonstrating that Q-learning effectively learns optimal routing behavior using routing-step-based time dynamics.

In Figure 8, after 10 training iterations, the Q-table still shows no meaningful pattern. The Q-values remain relatively close and scattered, indicating that the agent has not yet learned which state-action pairs result in efficient routing. During this early stage, the agent continues to explore widely, resulting in longer routing times to reach the target node.

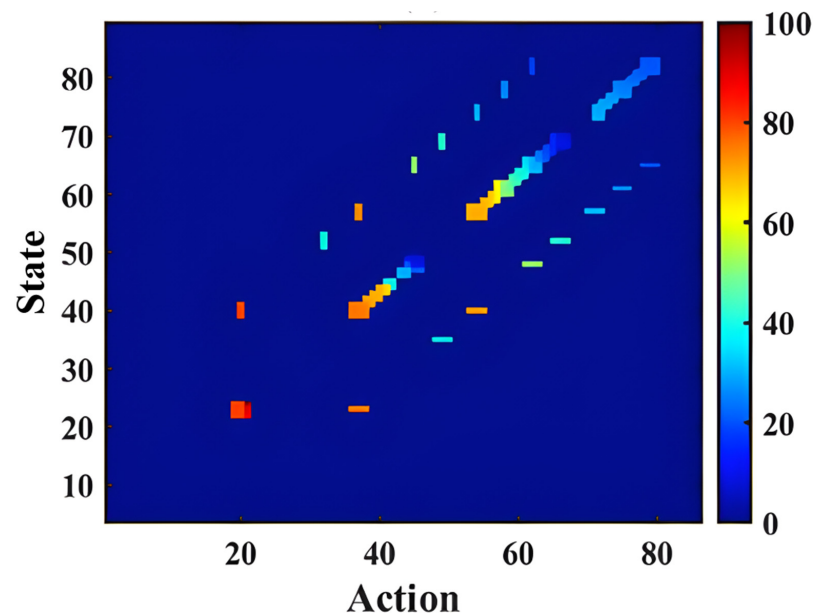


Figure 11. Q-table at 40 iterations (converged behavior).

Figure 9 shows the Q-table after 20 iterations. At this stage, initial warm-colored regions begin to emerge, suggesting that the agent is starting to recognize some beneficial routing choices. Routing performance improves compared to Figure 8, as the agent starts reinforcing more effective decisions. However, the Q-table remains only partially structured, meaning learning is still in progress. These emerging high-value regions indicate that the agent has begun associating specific routing actions with lower cumulative loss and shorter sensing time, even though full convergence has not yet been achieved.

In Figure 10, following 30 iterations, the Q-table becomes more organized and structured. Warm-colored clusters become more prominent, indicating that the agent is learning to consistently choose higher-value actions. The time required to reach the target node continues to decrease, demonstrating that the agent is transitioning from exploration to exploitation, relying more on accumulated experience. At this stage, routing decisions increasingly avoid breakpoints and inefficient switch configurations, leading to improved routing stability and reduced control overhead.

Finally, Figure 11 shows the Q-table after 40 iterations, where the warm-colored patterns are now clearly defined. This indicates that the Q-values have stabilized and the agent has successfully converged toward optimal routing behavior. At this stage, the routing path requires the shortest time to reach the target node, demonstrating that the agent has become highly familiar with the environment through repeated learning. Such convergence behavior is consistent with reinforcement-learning-based routing frameworks reported in the literature, where stable Q-value distributions indicate reliable policy learning in dynamic network environments [39].

Unlike deterministic shortest-path algorithms based on static link-loss weights which are still a common and widely used routing solution in optical sensor and communication networks [20,39] the learned policy allows for quick adaptation to link failures without the need for repeated global recomputation. From a comparative perspective, traditional non-reinforcement-learning routing strategies perform well when link conditions are static and fully understood.

However, in self-healing FBG sensor networks where fiber breaks, attenuation variations, or sensor faults may occur unpredictably, deterministic routing approaches require frequent recalculation and centralized control. Reinforcement-learning-based routing schemes have been shown to outperform conventional shortest-path methods in such

dynamic scenarios by continuously updating routing decisions based on environmental feedback rather than fixed cost metrics [39,42].

It should be noted that tabular Q-learning introduces scalability limitations as the number of network states and actions increases with larger mesh sizes or additional sensing regions. While the present study focuses on a moderate-scale topology to ensure stable convergence and interpretability, prior studies have demonstrated that these limitations can be mitigated through function approximation techniques, such as deep reinforcement learning and graph-based learning models, which enable scalable routing optimization in large optical and communication networks [42]. Overall, the experimental results confirm that the proposed reinforcement-learning-based routing mechanism effectively improves routing efficiency, reduces sensing delay, and enhances survivability in self-healing mesh-ring FBG sensor networks, making it a promising solution for long-term and large-scale sensing applications.

5. Conclusions

In summary, this work shows that integrating fiber-based ring-mesh topologies with reinforcement-learning (RL) control and free-space optical (FSO) backhaul can significantly enhance the performance of large-scale sensor networks. By exploiting the high sensitivity and robustness of fiber Bragg grating (FBG) sensors, the proposed system achieves accurate and reliable measurements even in harsh environments such as aerospace structures and civil infrastructure. The self-healing architecture, supported by dual optical switches and intelligent fault-detection mechanisms, effectively isolates both soft and hard faults and dynamically reconfigures the network to maintain service continuity, thereby improving survivability and reducing downtime. RL further strengthens this resilience by continuously evaluating network conditions and selecting optimal routing paths, reducing the need for extensive experimental tuning and lowering operational costs. In addition, FSO links extend network coverage without extra cabling, offering flexible line-of-sight connections that complement the fiber mesh and reduce latency.

Overall, the combined use of self-healing mechanisms, FSO connectivity, and machine-learning-based routing results in a scalable and robust sensing architecture well-suited for next-generation structural health monitoring and other demanding applications. Future work will explore deep reinforcement learning techniques to address scalability limitations and enable more complex routing decisions, further strengthening the system's applicability to large and dynamic sensor networks.

Author Contributions: Conceptualization, R.A.D., J.-W.L., H.-W.H., A.M.D., C.-K.Y., P.-C.L. and P.-C.P.; methodology, R.A.D., J.-W.L., A.M.D. and P.-C.P.; software, R.A.D., J.-W.L., H.-W.H. and A.M.D.; model validation, R.A.D.; formal analysis, R.A.D., A.M.D., J.-W.L., C.-K.Y., H.-W.H. and P.-C.P.; investigation, R.A.D., J.-W.L., H.-W.H., A.M.D., C.-K.Y., P.-C.L. and P.-C.P.; resources, R.A.D., J.-W.L., H.-W.H., A.M.D. and P.-C.P.; data curation, R.A.D., J.-W.L., H.-W.H. and A.M.D.; writing—original draft preparation, R.A.D.; writing—review and editing, R.A.D., C.-K.Y., A.M.D. and P.-C.P.; visualization, R.A.D., J.-W.L., C.-K.Y., A.M.D. and P.-C.P.; supervision, R.A.D., A.M.D. and P.-C.P.; project administration, R.A.D., J.-W.L., A.M.D. and P.-C.P.; funding acquisition, P.-C.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Science and Technology Council, Taiwan, under Grant NSTC 112-2221-E-027-076-MY2 and Grant NSTC 114-2221-E-027-056.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Domingues, M.D.F.F.; Radwan, A. Optical Fiber Sensors in IoT. In *Optical Fiber Sensors for IoT and Smart Devices*; Springer: Cham, Switzerland, 2017; pp. 73–86.
2. Kok, S.P.; Go, Y.I.; Wang, X.; Wong, M.L.D. Advances in Fiber Bragg Grating (FBG) Sensing: A Review of Conventional and New Approaches and Novel Sensing Materials in Harsh and Emerging Industrial Sensing. *IEEE Sens. J.* **2024**, *24*, 29485–29505. [\[CrossRef\]](#)
3. Zhang, B.; Kahrizi, M. High-Temperature Resistance Fiber Bragg Grating Temperature Sensor Fabrication. *IEEE Sens. J.* **2007**, *7*, 586–591. [\[CrossRef\]](#)
4. Abedin, S.; Biondi, A.M.; Wu, R.; Cao, L.; Wang, X. Structural Health Monitoring Using a New Type of Distributed Fiber Optic Smart Textiles in Combination with Optical Frequency Domain Reflectometry (OFDR): Taking a Pedestrian Bridge as Case Study. *Sensors* **2023**, *23*, 1591. [\[CrossRef\]](#)
5. Jia, D.; Zhang, Y.; Chen, Z.; Zhang, H.; Liu, T.; Zhang, Y. A Self-Healing Passive Fiber Bragg Grating Sensor Network. *J. Light. Technol.* **2015**, *33*, 2062–2067. [\[CrossRef\]](#)
6. Yeh, C.-H.; Chow, C.-W.; Wu, P.-C.; Tseng, F.-G. A Simple Fiber Bragg Grating-Based Sensor Network Architecture with Self-Protecting and Monitoring Functions. *Sensors* **2011**, *11*, 1375–1382. [\[CrossRef\]](#)
7. Yeh, C.H.; Chow, C.W.; Wang, C.H.; Shih, F.Y.; Wu, Y.F.; Chi, S. A simple self-restored fiber Bragg grating (FBG)-based passive sensing ring network. *Meas. Sci. Technol.* **2009**, *20*, 43001. [\[CrossRef\]](#)
8. Vallejo, M.F.; Perez-Herrera, R.A.; Elosua, C.; Diaz, S.; Urquhart, P.; Bariaín, C.; Lopez-Amo, M. Resilient amplified double-ring optical networks to multiplex optical fiber sensors. *J. Light. Technol.* **2009**, *27*, 1301–1306. [\[CrossRef\]](#)
9. Lopez, O.G.; Schires, K.; Urquhart, P.; Gueyne, N.; Duhamel, B. Optical fiber bus protection network to multiplex sensors: Amplification by remotely pumped EDFAs. *IEEE Trans. Instrum. Meas.* **2009**, *58*, 2945–2951. [\[CrossRef\]](#)
10. Izquierdo, E.L.; Urquhart, P.; Lopez-Amo, M. Protection architectures for WDM optical fibre bus sensor arrays. *J. Eng. Sci. Int.* **2007**, *1*, 1–18.
11. Yuan, L.; Wang, Q.; Zhao, Y. A passive ladder-shaped sensor architecture with failure detection based on fiber Bragg grating. *Opt. Fiber Technol.* **2023**, *81*, 103540. [\[CrossRef\]](#)
12. Kuroda, K. A passive ladder-shaped fiber Bragg grating sensor network with fault detection using time-and wavelength-division multiplexing. *Sensors* **2025**, *25*, 4261. [\[CrossRef\]](#)
13. Wu, C.Y.; Feng, K.M.; Peng, P.C.; Lin, C.Y. Three-dimensional mesh-based multipoint sensing system with self-healing functionality. *IEEE Photonics Technol. Lett.* **2010**, *22*, 565–567. [\[CrossRef\]](#)
14. Feng, K.-M.; Wu, C.-Y.; Yan, J.-H.; Lin, C.-Y.; Peng, P.-C. Fiber Bragg Grating-Based Three-Dimensional Multipoint Ring-Mesh Sensing System with Robust Self-Healing Function. *IEEE J. Sel. Top. Quantum Electron.* **2012**, *18*, 1613–1620. [\[CrossRef\]](#)
15. Yeh, C.-H.; Tsai, N.; Zhuang, Y.-H.; Chow, C.-W.; Liu, W.-F. Fault self-detection technique in fiber Bragg grating-based passive sensor network. *IEEE Sens. J.* **2016**, *16*, 8070–8074. [\[CrossRef\]](#)
16. Yeh, C.-H.; Chang, Y.-J.; Huang, T.-J.; Yang, Z.-Q.; Chow, C.-W.; Chen, K.-H. A fiber Bragg grating-based passive semicircular sensor architecture with fault monitoring. *Opt. Fiber Technol.* **2019**, *48*, 258–262. [\[CrossRef\]](#)
17. Chang, C.-H.; Lu, D.-Y.; Lin, W.-H. All-Passive Optical Fiber Sensor Network with Self-Healing Functionality. *IEEE Photonics J.* **2018**, *10*, 7203310. [\[CrossRef\]](#)
18. Chang, C.-H.; Tsai, C.-H. A large-scale optical fiber sensor network with reconfigurable routing path functionality. *IEEE Photonics J.* **2019**, *11*, 6801811. [\[CrossRef\]](#)
19. Hu, J.; Hu, X.; Shen, Z.; Wang, Z.; Li, J.; Hu, J. Self-Healing FBG Sensor Network Fault-Detection Based on a Multi-Class SVM Algorithm. *Opt. Express* **2023**, *31*, 41313–41327. [\[CrossRef\]](#)
20. Hu, X.; Si, H.; Mao, J.; Wang, Y. Self-healing and shortest path in optical fiber sensor network. *J. Sens.* **2022**, *2022*, 5717041. [\[CrossRef\]](#)
21. Zhang, R.; Liu, H.; Wang, Y.; Li, Z.; Zhang, Y. Research on Self-Diagnosis and Self-Healing Technologies for Fiber Optic Sensing Networks in Complex Environments. *Sensors* **2025**, *25*, 1641. [\[CrossRef\]](#)
22. Zhang, Y.; Xin, J. Survivable deployments of optical sensor networks against multiple failures and disasters: A survey. *Sensors* **2019**, *19*, 4790. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Zhang, R.; Lu, F.; Xu, M.; Liu, S.; Peng, P.-C.; Shen, S.; He, J.; Cho, H.J.; Zhou, Q.; Yao, S.; et al. An Ultra-Reliable MMW/FSO A-RoF System Based on Coordinated Mapping and Combining Technique for 5G and Beyond Mobile Fronthaul. *J. Light. Technol.* **2018**, *36*, 4952–4959. [\[CrossRef\]](#)

24. Alfadhli, Y.; Peng, P.-C.; Cho, H.; Liu, S.; Zhang, R.; Chen, Y.-W.; Chang, G.-K. Real-time FPGA demonstration of hybrid bi-directional MMW and FSO fronthaul architecture. In Proceedings of the Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA, 3–7 March 2019; pp. 1–3.
25. Chen, Y.-W.; Zhang, R.; Hsu, C.-W.; Chang, G.-K. Key Enabling Technologies for the Post-5G Era: Fully Adaptive, All-Spectra Coordinated Radio Access Network with Function Decoupling. *IEEE Commun. Mag.* **2020**, *58*, 60–66. [\[CrossRef\]](#)
26. Jeon, H.-B.; Kim, S.-M.; Moon, H.-J.; Kwon, D.-H.; Lee, J.-W.; Chung, J.-M.; Han, S.-K.; Chae, C.-B.; Alouini, M.-S. Free-space optical communications for 6G wireless networks: Challenges, opportunities, and prototype validation. *IEEE Commun. Mag.* **2023**, *61*, 116–121. [\[CrossRef\]](#)
27. Yeh, C.-H.; Lin, W.-P.; Jiang, S.-Y.; Hsieh, S.-E.; Hsu, C.-H.; Chow, C.-W. Integrated Fiber-FSO WDM Access System with Fiber Fault Protection. *Electronics* **2022**, *11*, 2101. [\[CrossRef\]](#)
28. Hayle, S.T.; Manie, Y.C.; Dehnaw, A.M.; Hsu, Y.-T.; Li, J.-W.; Liang, H.-C.; Peng, P.-C. Reliable self-healing FBG sensor network for improvement of multipoint strain sensing. *Opt. Commun.* **2021**, *499*, 127286. [\[CrossRef\]](#)
29. Yeh, C.-H.; Ko, H.-S.; Liaw, S.-K.; Liu, L.-H.; Chen, J.-H.; Chow, C.-W. A Survivable and Flexible WDM Access Network by Alternate FSO- and Fiber-Paths for Fault Protection. *IEEE Photonics J.* **2022**, *14*, 1–5. [\[CrossRef\]](#)
30. Wu, T.-H.; Liao, C.-Y.; Yeh, C.-H.; Chen, Y.-W.; Kao, Y.-H.; Lin, S.-Y.; Lin, Y.-H.; Liaw, S.-K. A Self-Healing WDM Access Network with Protected Fiber and FSO Link Paths Effective Against Fiber Breaks. *Photonics* **2025**, *12*, 323. [\[CrossRef\]](#)
31. Manie, Y.C.; Peng, P.-C.; Shiu, R.-K.; Hsu, Y.-T.; Chen, Y.-Y.; Shao, G.-M.; Chiu, J. Enhancement of the Multiplexing Capacity and Measurement Accuracy of FBG Sensor System Using IWDM Technique and Deep Learning Algorithm. *J. Light. Technol.* **2020**, *38*, 1589–1603. [\[CrossRef\]](#)
32. Sun, Y.; Zeng, W.; Shen, H.; Chen, W.; Chen, Y.; Liu, J.; Fan, Z. Separation of distorted overlapping spectra in fiber Bragg grating sensor networks using self-supervised contrastive learning. *Opt. Express* **2025**, *33*, 44654–44670. [\[CrossRef\]](#)
33. Arockiyadoss, M.A.; Dehnaw, A.M.; Manie, Y.C.; Hayle, S.T.; Yao, C.-K.; Peng, C.-H.; Kumar, P.; Peng, P.-C. Self-healing fiber Bragg grating sensor system using free-space optics link and machine learning for enhancing temperature measurement. *Electronics* **2024**, *13*, 1276. [\[CrossRef\]](#)
34. Arockiyadoss, M.A.; Yao, C.-K.; Liu, P.-C.; Kumar, P.; Nagi, S.K.; Dehnaw, A.M.; Peng, P.-C. Spectral demodulation of mixed-linewidth fiber Bragg grating sensor networks using cloud-based deep learning for land monitoring. *Sensors* **2025**, *25*, 5627. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Tefera, M.A.; Manie, Y.C.; Yao, C.-K.; Fan, T.-P.; Peng, P.-C. Meta-Learning for Boosting the Sensing Quality and Utility of FSO-Based Multichannel FBG Sensor System. *IEEE Sens. J.* **2023**, *23*, 31506–31512. [\[CrossRef\]](#)
36. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*; MIT Press: Cambridge, MA, USA, 2018.
37. Huang, R.; Guan, W.; Zhai, G.; He, J.; Chu, X. Deep Graph Reinforcement Learning Based Intelligent Traffic Routing Control for Software-Defined Wireless Sensor Networks. *Appl. Sci.* **2022**, *12*, 1951. [\[CrossRef\]](#)
38. Jin, Z.; Zhao, Q.; Su, Y. RCAR: A Reinforcement-Learning-Based Routing Protocol for Congestion-Avoided Underwater Acoustic Sensor Networks. *IEEE Sens. J.* **2019**, *19*, 10881–10891. [\[CrossRef\]](#)
39. Mammeri, Z. Reinforcement Learning Based Routing in Networks: Review and Classification of Approaches. *IEEE Access* **2019**, *7*, 55916–55950. [\[CrossRef\]](#)
40. Ottoni, A.L.C.; Nepomuceno, E.G.; Oliveira, M.S.d.; de Oliveira, D.C.R. Reinforcement learning for the traveling salesman problem with refueling. *Complex Intell. Syst.* **2022**, *8*, 2001–2015. [\[CrossRef\]](#)
41. Wang, Y.; Chen, L.; Zhou, H.; Zhou, X.; Zheng, Z.; Zeng, Q.; Jiang, L.; Lu, L. Flexible Transmission Network Expansion Planning Based on DQN Algorithm. *Energies* **2021**, *14*, 1944. [\[CrossRef\]](#)
42. Chen, X.; Li, B.; Proietti, R.; Lu, H.; Zhu, Z.; Yoo, S.J.B. DeepRMSA: A deep reinforcement learning framework for routing, modulation and spectrum assignment in elastic optical networks. *J. Light. Technol.* **2019**, *37*, 4155–4163. [\[CrossRef\]](#)
43. Nevin, J.W.; Nallaperuma, S.; Shevchenko, N.A.; Shabka, Z.; Zervas, G.; Savory, S.J. Techniques for applying reinforcement learning to routing and wavelength assignment problems in optical fiber communication networks. *J. Opt. Commun. Netw.* **2022**, *14*, 733–748. [\[CrossRef\]](#)
44. Kafaei, P.; Cappart, Q.; Chapados, N.; Pouya, H.; Rousseau, L.-M. Dynamic routing and wavelength assignment with reinforcement learning. *Inf. J. Optim.* **2023**, *6*, 1–18. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.