Reinforcement Learning to Enhance Digital Twin Based EDFA Fault Prediction in Optical Networks

(Article begins on next page)

Reinforcement Learning to Enhance Digital Twin Based EDFA Fault Prediction in Optical Networks

Mashboob Cheruvakkadu Mohamed, Renato Ambrosone, Muhammad Umar Masood, Gulmina Malik, Stefano Straullu, Antonino Nespola, Sai Kishore Bhyri, Antonio Napoli, João Pedro Sasipim Srivallapanondh, Gabriele Maria Galimberti, Walid Wakim and Vittorio Curri

***Abstract*—This study proposes a novel two-stage framework that integrates Digital Twin (DT) with Reinforcement Learning (RL) for failure prediction in optical communication networks. The framework focuses on optical amplifier failures, one of the most critical elements impacting network reliability. In the first stage, DTs are employed using the GNPy (Gaussian Noise model in Python) open-source optical network simulation framework to emulate amplifier behavior across diverse operational scenarios. This enables the generation of a comprehensive training dataset to train the Machine Learning (ML) model based on Long Short-Term Memory (LSTM). The LSTM model is able to achieve an accuracy of 98%. In the second stage, RL is applied to enhance the predictive accuracy to 99.5% and to improve the adaptability of the model. The LSTM model is rewarded for accurate predictions using external feedback, allowing it to iteratively refine its performance. This feedback-driven learning mechanism strengthens the robustness and decision-making capabilities of the model under dynamic network conditions. The proposed fine-tuning of the DT-based ML model using RL with external feedback enables proactive and adaptive fault management in optical networks.**

***Index Terms*—Deep Learning, Digital Twin, Fault Detection, Network Reliability, Optical Amplifiers, Proactive Failure Management, Reinforcement Learning, Soft Failures.**

I. Introduction

The rapid growth of high bandwidth services and advancements in Artificial Intelligence (AI) have intensified the demand for ultra high speed data transmission. To meet this demand, optical fiber networks, the core communication infrastructure are being driven to operate at reduced margins to maximize capacity. Such a narrow margin along with static operational management increases vulnerability to failures, emphasizing the need for intelligent, adaptive failure management. Machine Learning (ML) has gained prominence for automating and improving fault detection and localization [1]–[5], particularly in identifying soft failures undetectable by conventional systems [6]–[9]. However, their performance critically depends on comprehensive and balanced training datasets encompassing diverse fault conditions [10]. In practice, the scarcity and uneven distribution of real fault data particularly rare or transient failures cause class imbalance and hinder generalization in deployment. Data augmentation techniques [11] partially address this issue but often fail to capture the full complexity of operational networks. Digital Twins (DTs) offer an alternative by generating synthetic fault data under controlled conditions [12], yet models trained solely on DT data may under perform in real environments due to noise and distributional shifts. These limitations highlight the necessity of adaptive learning frameworks that can fine-tune pretrained ML models to enhance robustness and reliability in optical network failure prediction.

Over the past few years, there has been a notable surge in research focused on Reinforcement Learning (RL) in the domain of optical networks. This emerging trend is driven by the increasing complexity and dynamic nature of modern optical communication systems, which demand intelligent, adaptive, and real-time decision-making mechanisms. RL offers a powerful framework for enabling autonomous network management by allowing agents to learn optimal policies through interaction with the environment. Current research efforts are predominantly centered on network parameter optimization issues [13], resource allocation problems such as routing, modulation, wavelength or spectrum assignment [14]–[19]. Collectively, these studies emphasize the applicability and practical utility of RL in optical networks. By continuously learning from network conditions and performance feedback, RL-based approaches have shown promising results in improving network efficiency, reducing operational costs, and enhancing overall system resilience.

Erbium Doped Fiber Amplifier (EDFA) aging and noise accumulation can induce gradual and masked gain degradation in Optical Line System (OLS) networks, resulting in latent Quality of Transmission (QoT) impairments. Such slow, progressive degradation remains undetected by conventional power monitoring schemes and static threshold based alarm mechanisms, thereby delaying fault identification and mitigation. In our previous research, we trained the ML model using DT to perform amplifier failure prediction in optical networks [12], [20]. To enhance the predictive performance and generalization capability of the model, we employed Transfer Learning (TL) techniques leveraging an experimental testbed dataset [10]. This approach allowed the pre-trained DT-based models to adapt to new network conditions and failure patterns with limited additional data. Building upon this foundation, this work advances the existing DT-based

Mashboob Cheruvakkadu Mohamed (mashboob.cheruvakkadu@polito.it), Renato Ambrosone, Muhammed Umar Masood, Gulmina Malik and Vittorio Curri are with Politecnico di Torino, Italy. Stefano Straullu and Antonino Nespola are with Links Foundation, Italy and Sai Kishore Bhyri, Antonio Napoli, João Pedro, Sasipim Srivallapanondh, Gabriele Maria Galimberti and Walid Wakim are with Nokia - Optical Networks. João Pedro is also with Instituto de Telecomunicações, Instituto Superior Técnico, Portugal. This publication has received funding from the projects PNRR-NGEU (MUR–DM117/2023), ALLEGRO (GA: 101092766) and SENSEI (GA: 101189545).

Manuscript received Nov xx, 2025; revised Nov xx, 2025.

framework through the integration of RL to further enhance ML model performance. This work represents, to the best of our knowledge, the first attempt to leverage RL for fine-tuning a DT pre-trained ML model for failure prediction in optical networks. The objective is to improve the adaptability and decision-making capabilities of the failure prediction system in dynamic network environments. By incorporating RL, the model can continuously learn from external feedback, update its predictions and optimize its performance over time. Failure prediction is conducted in a two-stage framework. In the first stage, the ML model is trained on a comprehensive dataset generated through the use of DT, which simulates diverse failure scenarios and system behaviors. In the second stage, RL is integrated to enhance the predictive accuracy and adaptability of the model. The overall concept of the DT-integrated RL framework for failure prediction in optical networks is depicted in Fig. 1. The remainder of this paper is organized as follows. Section II introduces DT-based ML model for optical amplifier failure prediction. Section III details the RL framework and algorithm for fine-tuning pre-trained model and evaluate the performance using standard performance metrics. Finally, Section IV concludes and outlines potential avenues for future work.

II. DT-based ML model for optical amplifier failure prediction

Optical amplifiers are susceptible to gradual performance degradation due to component issues such as gain medium aging, Amplified Spontaneous Emission (ASE) [21] noise increase, pump laser degradation, etc. Such impairments can reduce amplifier gain, degrade the Optical Signal-to-Noise Ratio (OSNR), and compromise the QoT. In an OLS with multiple amplifiers, early-stage degradation in one amplifier may remain undetected if its magnitude is below alarm thresholds. Downstream amplifiers may compensate for this loss, masking the fault from being identified using power monitors at the lightpath endpoints. This can result in latent/soft failures that persist undiagnosed, posing a risk to network reliability. In this work, the scenario of amplifier gain degradation due to aging effects is represented through failure modeling in DTs.

GNPy (Gaussian Noise model in Python) facilitates optical network design by modeling and simulating the behavior of key infrastructure elements [22]. The DT of the experimental test bed network topology shown in Fig. 2 is modeled using GNPy to evaluate the impact of physical-layer impairments on signal quality. This tool enables simulation of key optical signal degradation factors such as ASE noise, Non-Linear Interference (NLI) and dispersion accumulation along the transmission path. GNPy requires a comprehensive set of network descriptors, including the physical topology (e.g., fiber links, nodes, and inline optical elements), as well as detailed equipment specifications, including amplifier characteristics (e.g., gain and noise figure) and transponder configurations. Leveraging this input, the tool computes critical signal quality metrics such as Generalized Signal-to-Noise Ratio (GSNR), OSNR and other QoT indicators, facilitating informed network assessment and optimization. GNPy conducts signal propagation analysis by modeling the interaction between

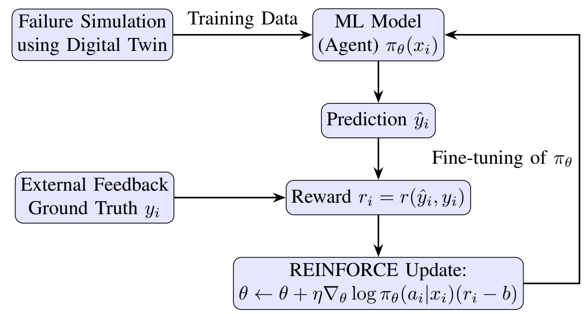


Fig. 1: RL based fine tuning of a pretrained agent π_θ, originally trained using data generated by the DT. The policy network $\pi_\theta(a_i|x_i)$ predicts optical amplifier failure conditions given input state of transmission parameters x_i, and receives reward $r_i = r(\hat{y}_i, y_i)$ based on prediction accuracy with respect to ground truth y_i. The policy parameters θ are updated using the REINFORCE rule with learning rate η: $\theta \leftarrow \theta + \eta \nabla_\theta \log \pi_\theta(a_i|x_i)(r_i - b)$, where b is a moving average baseline.

optical signals and various network components, such as fiber spans, optical amplifiers, and Reconfigurable Optical Add-Drop Multiplexers (ROADMs). Utilizing the Gaussian Noise (GN) model as its analytical foundation, GNPy simulates the transmission path to estimate the progressive accumulation of signal impairments. This approach enables a realistic approximation of end-to-end transmission quality by capturing both linear and nonlinear degradations across the network. Across multiple GNPy simulation iterations, amplifier gain was incrementally reduced in small steps, while the Variable Optical Attenuator (VOA) initially set to its maximum was concurrently decreased to maintain output power, thereby modeling gradual, silent degradation that remains undetectable through conventional power monitoring. The generated dataset comprises $\approx$ 20k samples with detailed amplifier attributes and associated transmission parameters, such as GSNR and OSNR, characterized under both normal operating conditions and amplifier failure scenarios. This dataset is used to train the ML model based on the Long Short-Term Memory (LSTM) architecture. The model was shown to achieve a 98% accuracy in predicting optical amplifier failures [12], [20].

III. Reinforcement Learning Framework and Performance Evaluation

The aforementioned fine-tuning of the LSTM model through RL, as illustrated in Fig. 1, employs the REINFORCE policy gradient algorithm to adapt a pretrained ML agent for optical amplifier failure prediction. The Algorithm 1 outlines the RL framework applied to the pretrained LSTM model for improving its effectiveness in the prediction of optical amplifier failures. The pretrained policy network $\pi_\theta(a|x)$ outputs class probabilities for each input sample. During training, actions (predicted classes) are selected according to an ϵ-greedy strategy to balance exploitation and exploration. Correct classifications yield positive, class-weighted rewards $W[y_i]$, whereas incorrect predictions incur a small negative penalty. The algorithm computes the advantage function, $A_i = r_i - b$, where b is an exponentially moving average baseline used to

This article has been accepted for publication in IEEE Photonics Technology Letters. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/LPT.2026.3681516

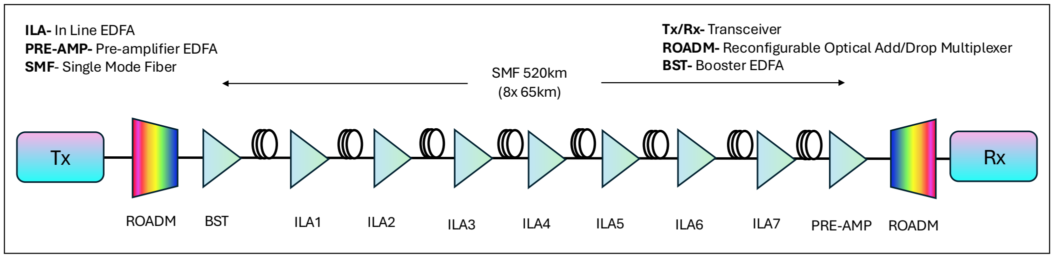


Fig. 2: Optical transmission line test bed with Erbium Doped Fiber Amplifiers (EDFA), spanning over 520 kms.

Algorithm 1 Fine-tuning pretrained ML model (agent) with REINFORCE algorithm for optical amplifier failure prediction

Require: Pretrained policy network π_θ (pretrained ML agent with parameters θ, which outputs class probabilities $\pi_\theta(a|x))$, dataset X, labels Y, class weights W, exploration rate ϵ, number of episodes N

Ensure: Fine-tuned agent with updated policy parameters, θ

Initialize optimizer (Adam)
Set moving average baseline $b \leftarrow 0$
Set smoothing factor α
for each episode $= 1$ to N **do**
  Initialize total reward to 0
  for each sample x_i, label y_i in dataset **do**
    Predict action probabilities: $\pi_\theta(a|x_i)$
    Select action a_i:
- With probability $1-\epsilon$: $a_i \leftarrow \arg\max \pi_\theta(a|x_i)$
- With probability ϵ: choose a_i randomly

    Compute log probability: $\log \pi_\theta(a_i|x_i)$
    if $a_i = y_i$ **then**
      $r_i \leftarrow W[y_i]$ positive weighted reward
    else
      $r_i \leftarrow -0.1$ Penalty for wrong action
    end if
    Compute advantage: $A_i = r_i - b$
    Compute loss: $L_i = -\log \pi_\theta(a_i|x_i) \cdot A_i$
    Update model parameters using gradient descent on L_i with learning rate η: $\theta \leftarrow \theta - \eta \nabla_\theta L_i = \theta + \eta \nabla_\theta \log \pi_\theta(a_i|x_i) \cdot A_i$
    Accumulate reward $R \leftarrow R + r_i$
  end for
  Update baseline: $b \leftarrow (1-\alpha) \cdot b + \alpha \cdot \left(\frac{R}{|X|}\right)$
end for
Evaluate model using accuracy, precision, recall, and F1-score

reduce variance in the gradient estimates. Model parameters are updated via gradient ascent on the expected return, using the REINFORCE update rule $\theta \leftarrow \theta + \eta \nabla_\theta \log \pi_\theta(a_i|x_i) \cdot A_i$. Across multiple episodes, the baseline is iteratively refined to track the mean reward, stabilizing the learning process. This reinforcement-based fine-tuning allows the pretrained model to adjust its decision policy based on task-specific reward signals rather than relying solely on supervised loss, thereby enhancing predictive robustness and class discrimination in optical amplifier fault detection tasks.

The LSTM model is trained and subsequently fine-tuned using RL with the GNPy simulated dataset to predict faulty amplifiers along with their respective Fault Levels (FL), as defined in Fig. 3. The model input includes amplifier attributes and transmission parameters. The fault thresholds represent soft failures up to 3 dB, consistent with link budget engineering practices that include over 3 dB of VOA attenuation per span to ensure operational margin [23]. The classes are categorized according to the position of the In-Line Amplifier (ILA) and the corresponding FL – $\{0, 1, 2\}$. The notation ILAx_y denotes the ILA at position x with fault level y; for example, ILA1_0 represents the first ILA with fault level 0, corresponding to a gain degradation of less than 2 dB (i.e., no fault).

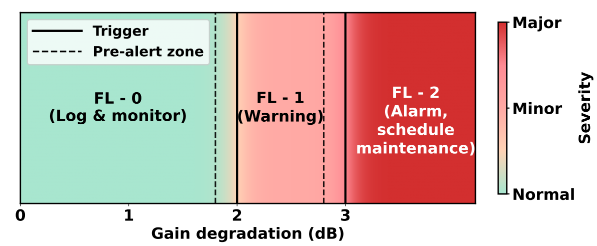


Fig. 3: Amplifier fault levels based on gain degradation

In the classification task, each model prediction is categorized into one of four outcomes: True Positive (TP), True Negative (TN), False Positive (FP), or False Negative (FN). These outcomes serve as the foundation for computing key performance metrics such as accuracy, precision, recall, and F1-score, which quantitatively evaluate the model's classification capability and reliability. Fig. 4 presents a comparative analysis of the pre-trained and RL fine-tuned model on unseen data, demonstrating improvements in overall accuracy as well as in weighted-average (WA) precision, recall, and F1-score, reflecting the model's superior balance between sensitivity and specificity. The class-wise performance depicted in Fig. 5a

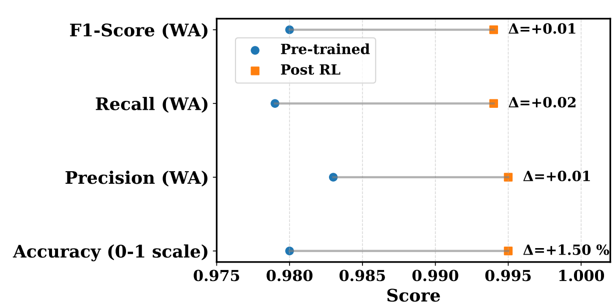
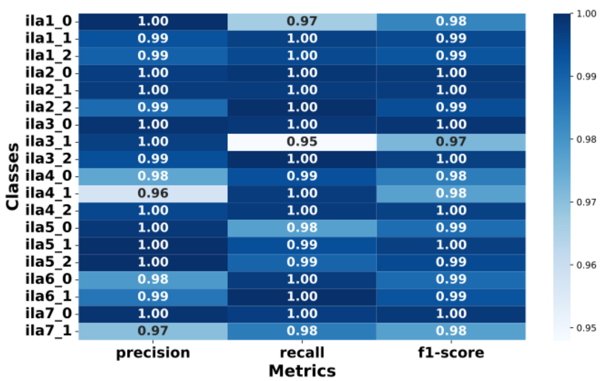
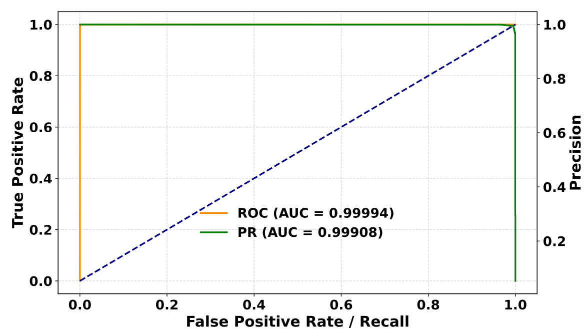


Fig. 4: Accuracy, precision, recall and F1 score - a comparison of pre-trained model and after applying RL.

demonstrates consistently high accuracy across all fault categories, with performance scores exceeding 0.95 for every class and approaching perfection (0.99–1.00) in most cases. Furthermore, the Receiver Operating Characteristic (ROC) and Precision Recall (PR) curves shown in Fig. 5b provide deeper insights into the model's decision behavior by illustrating the trade-offs between true positive and false positive rates, and between precision and recall, respectively. The Area Under the Curve (AUC) values of ≈1.00 obtained for both ROC and PR analyses indicate ideal classification performance, confirming the model's enhanced discriminative power. The RL process was executed over 30 episodes, and the total computational time remained within 10% of the LSTM training time. This indicates that while RL introduces additional steps for policy learning, the overall computational overhead is modest compared to LSTM training, highlighting its efficiency for adaptive, data driven fault prediction.



(a)



(b)

Fig. 5: Evaluation of ML classifier performance using (a) class-wise precision, recall, and F1-score, and (b) ROC, PR curves and AUC analysis.

IV. Conclusions

High prediction accuracy is fundamental to the reliable deployment of ML models for failure management in low-margin optical networks, where even minor prediction errors can lead to significant service degradation and operational costs. This work demonstrates that RL with external feedback can fine tune a DT trained model, achieving 99.5% classification accuracy, enhancing fault prediction and localization of optical amplifier failures. The proposed framework could be further extended to other network elements, such as ROADMs, enabling proactive mitigation and progress toward resilient, self-healing and autonomous optical networks.

References

[1] F. Musumeci et al. A tutorial on machine learning for failure management in optical networks. *Journal of Lightwave Technology*, 37(16):4125–4139, 2019.

[2] B. Shariati et al. Learning from the optical spectrum: Failure detection and identification. *Journal of Lightwave Technology*, 37(2):433–440, 2019.

[3] M. F. Silva et al. Learning long- and short-term temporal patterns for ml-driven fault management in optical communication networks. *IEEE Transactions on Network and Service Management*, 19(3):2195–2206, 2022.

[4] F. N. Khan et al. An optical communication's perspective on machine learning and its applications. *Journal of Lightwave Technology*, 37(2):493–516, 2019.

[5] K. S. Mayer et al. Demonstration of ml-assisted soft-failure localization based on network digital twins. *J. Lightwave Technol.*, 40(14):4514–4520, Jul 2022.

[6] L. E. Kruse et al. Experimental validation of machine learning-based joint failure management and quality of transmission estimation. *IEEE Photonics Journal*, 15(6):1–9, 2023.

[7] C. Delezoide et al. Streamlined failure localization method and application to network health monitoring. *Journal of Lightwave Technology*, 41(19):6119–6125, 2023.

[8] D. Wang et al. When large language models meet optical networks: Paving the way for automation. *Electronics*, 13(13), 2024.

[9] H. Lun et al. A gan based soft failure detection and identification framework for long-haul coherent optical communication systems. *Journal of Lightwave Technology*, 41(8):2312–2322, 2023.

[10] M. C. Mohamed et al. Network data based transfer learning failure prediction agent pre-trained using digital twin. In *2025 International Conference on Optical Network Design and Modeling (ONDM)*, pages 1–6, 2025.

[11] C. Xing et al. Failure data augmentation for optical network equipment using time-series generative adversarial networks. In *Optical Fiber Communication Conference (OFC) 2023*, page M3G.4. Optica Publishing Group, 2023.

[12] M. C. Mohamed et al. Machine learning agents leveraging digital twins for failure prediction in optical networks. In *2025 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)*, pages 1–6, 2025.

[13] R. Koch et al. Reinforcement learning for generalized parameter optimization in elastic optical networks. *Journal of Lightwave Technology*, 40(3):567–574, 2022.

[14] J. W. Nevin et al. Techniques for applying reinforcement learning to routing and wavelength assignment problems in optical fiber communication networks. *Journal of Optical Communications and Networking*, 14(9):733–748, 2022.

[15] S. Nallaperuma et al. Interpreting multi-objective reinforcement learning for routing and wavelength assignment in optical networks. *Journal of Optical Communications and Networking*, 15(8):497–506, 2023.

[16] X. Chen et al. Deeprmsa: A deep reinforcement learning framework for routing, modulation and spectrum assignment in elastic optical networks. *Journal of Lightwave Technology*, 37(16):4155–4163, 2019.

[17] O. Ayoub et al. Towards explainable reinforcement learning in optical networks: The rmsa use case. In *Optical Fiber Communication Conference (OFC) 2024*, page W4I.6. Optica Publishing Group, 2024.

[18] Juan et.al P.R. Resource allocation in multicore elastic optical networks: A deep reinforcement learning approach. *Complexity*, 2023(1):4140594, 2023.

[19] Y. Teng et al. Deep-reinforcement-learning-based rmsca for space division multiplexing networks with multi-core fibers [invited tutorial]. *Journal of Optical Communications and Networking*, 16(7):C76–C87, 2024.

[20] M. C. Mohamed et al. Digital twin-integrated binary classifier ml model for edfa failure prediction. In *2025 International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, pages 1–6, 2025.

[21] A. W. S. Putra et al. Theoretical analysis of noise in erbium doped fiber amplifier. *IEEE Journal of Quantum Electronics*, 53(4):1–8, 2017.

[22] V. Curri. Gnpy model of the physical layer for open and disaggregated optical networking [invited]. *Journal of Optical Communications and Networking*, 14(6):C92–C104, 2022.

[23] A. Zhang et al. Field trial of 24-Tb/s (60×400 Gb/s) DWDM transmission over a 1910-km G.654.E fiber link with 6-THz-bandwidth C-band EDFAs,. *Opt. Express*, 29(26):43811–43818, Dec 2021.