STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **16**, October 2026, pp 3432–3453.
Published online in International Academic Press (www.IAPress.org)

Deep Learning-Enabled Multi-Parameter Optimization for TWDM-PON Systems with Extended Reach

Zahraa. M. Hasan[1], Ayat Naji Hussain[2], Ali Shakir Sabbar Hussein[1], Essam N. Abdulla[3,*], Hayder H. Al-khaylani[3]

[1] *Department of Computer Networks and Software Technologies, Al Nassriyah Technical Institute, Southern Technical University, Thi-Qar, Iraq*
[2] *Information Technology Center, University of Technology, Baghdad, Iraq*
[3] *College of Laser and Optoelectronics Engineering, University of Technology-Iraq, Baghdad, Iraq*

Abstract The robustness of the physical layer greatly affects communication capacity in modern optical access networks, particularly techniques based on time-and-wavelength division multiplexed passive optical networks (TWDM-PONs). A signal may become highly vulnerable to interception, inaccurate signal reconstruction, and unstable performance due to signal degradation caused by attenuation, chromatic dispersion, and nonlinear Kerr effects. Therefore, improving transmission quality measures such as receiver sensitivity, bit error rate (BER), and Q-Factor is a prerequisite for achieving reliable and secure data delivery. In addition, deep learning (DL) is promising in improving the performance of optical access networks. This study proposes a DL-based optimizer that jointly tunes transmitter power, receiver gain, and dispersion compensation parameters. Our results demonstrate that the DL approach achieves a Q-factor above 35 dB and an average BER in the order of 10^{-17}, which significantly outperform the conventional pre-forward error correction (FEC) threshold (3.8×10^{-3}) at 80 km. Therefore, the approach significantly outperforms traditional fixed-compensation schemes which hit the FEC limit at 55 km. The optimized operating points predicted by the proposed deep neural network were independently verified using detailed OptiSystem simulations, which demonstrated strong agreement with those of the analytical model.

Keywords Deep learning, DNN, FEC, Optical access network, Next Generation-PON2 (NG-PON2), TWDM- PON.

DOI: 10.19139/soic-2310-5070-4551

1. Introduction

Passive optical networks (PONs) are a class of fiber access systems that is expanding the quickest globally. Commercial deployments show that PON-based fiber-to-the-building and fiber-to-the-cabinet (FTTX) technologies are increasingly becoming widely used worldwide [1]. Virtual reality (VR), cloud games, video conferencing, and other video services are only some of the communication and multimedia services for which FTTX technology is universally acknowledged as the best option. Triple play services consisting of high-internet speed access, VR, online gaming, and video conferences are achieved by FTTX, which prepares fiber optics straight to businesses or homes [2]. A flexible coherent PON is a promising solution for the future access network. By offering increased speed, sensitivity, and flexibility, it enables more efficient utilization of network resources and allows for serving a larger number of users [3]. Future FTTX-safe connections that can fulfill increasing demands for customer bandwidth and significant cost savings in capital expenditure associated with its execution are important elements

*Correspondence to: Essam N. Abdulla (Email: Essam.N.Abdulla@uotechnology.edu.iq). College of Laser and Optoelectronics Engineering, University of Technology-Iraq, Baghdad, Iraq.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

driving its rapid global proliferation. Consequently, gigabit PON and ethernet PON are widely utilized. Greater data rate 10-gigabit-capable-PON implementation has also greatly progressed [4].

The investigation and standard procedure for next generation (NG)-PON2 is currently under process. Bandwidth of time-and-wavelength division multiplexed PONs (TWDM-PONs) is rapidly increasing from 40 Gbps to 100 Gbps to meet 5G broadband requirements based on the universal public radio interface [5]. Time, wavelength, code, frequency, and a combination of wavelength and time (TDM-PON, WDM-PON, OCDM-PON, OFDM-PON, and TWDM-PON) are examples of PON division multiplexing. Owing to its low cost, efficient dynamic collaboration, flexible design, and enticing power budget needs, TWDM-PON is the most appealing option among these methods for NG-PON2 [6]. TWDM-PON, which is the primary solution of NG-PON2, combines time-division multiplexing (TDM) with wavelength-division multiplexing (WDM). TWDM-PON with 4 wavelengths, which typically has a 40 Gbps bandwidth, satisfies NG-PON2 standards [7]. PONs frequently use ONUs, optical distribution networks (ODNs), and optical line termination (OLTs) [8].

Thanks to current advances in the domain, artificial intelligence (AI) can now analyze data from Internet-of-things sensors, vision, text, and speech, among additional modes of communication. As a key element of AI, deep learning (DL) has developed into a pillar of contemporary technology infrastructure, which significantly impacts our daily lives and enhances the functionality of healthcare networks, smart cities, transportation systems, and home automation. DL algorithms are helpful and necessary; without them, we would fail to quickly and accurately examine and assess large amounts of data, which would lead to a reduction in system efficacy and service personalization as well as technological stagnation [9]. The paper is structured into five sections. Section 2 examines relevant literature and draws attention to the shortcomings of current methods. The suggested DL approach, which includes model design, preprocessing, and dataset preparation, is presented in Section 3. Section 4 provides a comprehensive performance analysis and assessment along with a discussion of the results. The paper is finally concluded in Section 5, which also suggests possible future research directions.

2. Related work

Recently, DL-based ciphers have shown great potential for managing optical faults in PONs. These ciphers perform well when evaluated and taught with data from the identical PON. However, their performance may rapidly decrease if the PON (utilized to generate training data) is changed, such as by adding more branches or altering the distance between two nearby branches. To improve the performance of conventional ciphers, DL has been applied to the inherent difficulties in optical fiber communications. In addition, DL models can train on an incremental basis and are highly accurate in forecasting time series data [10]. Many ciphers for enhancing data have been proposed; for example, the dynamic wavelength and bandwidth allocation (DWBA) algorithm, which is being investigated as a hotspot for TWDM-PON; and fronthaul 5G internet traffic, which necessitates minimal latency and significant capacity. Several DWBA studies focus on balancing loads and conserving energy, but they ignore time latency, which makes them inappropriate for fronthaul 5G networks [11] and optical-based saved chaos [12]. Moreover, optical communication has been protected using the hill method algorithm [13, 14], TWDM-PON by hill ciphers [15], and protection systems by the blowfish technique [16, 17].

Al-Janabi et al. discussed the issue of creating a secure environment for information transmission in an association dependent on an innovation based on noticeable light correspondence (VLC) utilizing a power line correspondence modem [18]. Haider al. proposed a solution for the optimization problem by maximizing the lower weighted level and allocating bandwidth for each subscriber station according to its QoS requirements for access rate and priority [19]. Conventional distribution tactics usually depend on constant dispersion compensation methods, including fiber Bragg gratings (FBG) or dispersion compensating fiber, as well as dynamic connection budgets. Owing to their incapability to adjust to dynamically changing connection circumstances and signal deterioration, these static ciphers result in subpar network efficiency [20]. As a result, the focus for research is currently shifting to algorithm ciphers, with machine learning (ML) being studied for optimization of parameters. The incapability of traditional “shallow” ML ciphers, such as random forests and support vector machines (SVMs), to accurately replicate the highly nonlinear function of transmission across the optical connection and their reliance on human characteristics limit their effectiveness. Currently, studies have turned to DL and reinforcement learning (RL) to solve this nonlinearity. According to Amirabadi et al. [21], given that deep neural networks (DNNs) can model highly dimensional continuous operations, they perform better than traditional ML in simulating nonlinear fiber noise (Kerr effect). In the specific context of TWDM-PON, Cheng et al. [22] created a deep RL (DRL) technique for cooperative resource allocation. Their approach, which focused mostly on the MAC layer, effectively maximized wavelength distribution and capacity but failed to address physical layer components such as linewidth and receiver sensitivity.

Despite these advancements, comprehensive optimization techniques that simultaneously modify transmitter properties, receiver sensitivity, and dispersion correction features are currently insufficient. The trade-offs between amplified spontaneous emission and optical fiber nonlinear impairments noise produced by the EDFA are often overlooked in previous studies that adjusted these parameters independently. This study improves by going beyond simple ML to solve the core shortcomings of earlier methods. Conventional approaches often exhibit linear, complex, and intrinsic nonlinearity derived from the Kerr property, including four-wave mixing (FWM), self-phase modulation (SPM), and control optical fiber propagation. Various parameters, including link separation, transmitter power, amplifier gain, and dispersion, are covered by the 20,000 transmission options of the dataset employed. The same parameters are used in fixed-cost approaches and grid search optimization. The results demonstrate that previous approaches cannot detect nonlinear relationships between elements over long distances. For example, the grid yielded a Q-factor of around 31.4 dB at 80 km, but the proposed DL approach provided a Q-factor of approximately 35.8 dB. Moreover, the root mean square error (RMSE) and mean absolute error (MAE) measurements indicate that the proposed model provides more accurate and consistent predictions across all evaluated variables [Q-factor, bit error rate (BER), and receiver sensitivity].

2.1. Our contribution

In an insecure channel paradigm, DL is a powerful data exploration technique that may be utilized to discover the typical and anomalous behavior of TWDM-PON. However, the reach and capacity of these techniques are severely limited by linear and nonlinear impairments, including chromatic dispersion and attenuation effects through optimal launch power and gain configuration. The intricate, nonlinear interactions between transmission variables are frequently missed by traditional static compensating strategies (e.g., fixed FBGs) and standard ML strategies. Initially, the DNN undergoes training to simulate the relationship between physical-layer variables and efficiency measurements. The input space, goal variables, normalization procedure, and loss formulations are all well explained to ensure uniform training behavior. To ensure repeatability, all stochastic components, such as initialization and data partitioning, are controlled using fixed random seeds. In the second phase, a confined search process employs the trained model as a surrogate evaluator. The learned model is used to evaluate candidate configurations that are generated within predetermined physical constraints. The best configuration is chosen based on the greatest Q-factor after a feasibility filter is used to apply BER limits. The contributions of this study can be summarized as follows:

1. Although ML techniques have previously been investigated for optical access networks, most existing studies focus on traffic prediction, fault diagnosis, resource allocation, or generic parameter estimation. By contrast, this work proposes a physics-informed residual-attention multi-task DNN, which is specifically designed for simultaneous optimization of physical-layer transmission parameters in high capacity on asymmetric TWDM-PON with 160/80 Gbps at the 512 separated rate and 80 km link distance in compliance with the ITU-T G.989.2 [23] guideline systems. Therefore, the novelty of this work lies in the proposed architecture and optimization framework rather than in the use of DNN approach itself to enhance communication requirements (Q-factor, BER, and receiver sensitivity) for PON.

2. Unlike traditional DNNs, the DNN proposed in this study utilizes stacked hidden layers and nonlinear activation functions (e.g., ReLU and Tanh) to act as a universal function approximator. This architecture allows the model to effectively learn the inverse transfer function of the nonlinear fiber channel, which provides a level of physical layer modeling fidelity that shallow algorithms cannot achieve. A major bottleneck in standard ML, where missing variables can cause model failure, is eliminated by the suggested

DL approach. Transmission power, length, and spectrum wavelength are examples of high-level abstract features that the DNN acquires from raw input data using representation learning. These parameters are directly correlated with signal quality measurements such as the Q-factor [24]. Joint multiple-parameter optimization is made possible by this feature, which is a clear benefit over conventional ciphers that frequently optimize parameters separately. The approach can find complex trade-offs by searching for a global optimum in a high-dimensional domain. For example, it can discover that increasing avalanche photodiode (APD) gain is more successful than increasing the power transmitted at 80 km to prevent nonlinear saturation—a minor optimization technique that is usually ignored by conventional logic. Our method uses a DL optimization to improve link length by simultaneously altering all important physical layer variables to close that gap.

3. We obtain superior findings and create practical optimization strategies to enhance communication condition by comparing our approach with a DL technique in PON employing the dataset [25]. This approach enhanced the performance and reliability of PONs. With patterns that approximate the anticipated optical behavior, the results of the suggested DNN-based model consistently and significantly outperform the conventional outcomes [8], which indicates the durability of the model.

3. Methodology

To simulate the nonlinear transmission functions of fiber optic propagation in a TWDM-PON, this research uses a DNN. Unlike traditional ML ciphers (e.g., RF or SVM) that rely on surface feature mapping, our proposed architecture uses deep residual learning and a strategy of interest to identify complex interactions between physical layer parts. In this study, we improve the physical layer characteristics of a TWDM-PON system using a data-driven methodology. The methodology involves three steps: (A) creating a synthetic dataset based on optical transmission physics, (B) preprocessing the data, and (C) creating a DNN with a target technique for several-parameter optimization.

3.1. Dataset Generation and Preprocessing

A large-scale synthetic dataset with 20,000 distinct propagation instances was created to efficiently train the DL model [25]. This dataset considers physical restrictions that conventional linear models frequently overstate to simulate the nonlinear transfer processes of the fiber optics channel.

1) Simulations Environment: A TWDM-PON link mimics upstream and downstream transmission in the C-band and L-band ($\lambda \in \{1,534, 1,596\}$ nm). For the following link, a standard single-mode fiber (SMF) with an attenuation value of $\alpha = 0.2$ dB/km is considered. The trained DNN approximates the nonlinear mapping between the input transmission parameters and the resulting system performance. The dataset includes 11 input parameters (X) that are randomly distributed below allowable physical restrictions. Table 1 provides a full description of the parameter ranges.

Table 1. Simulation Parameter Ranges

Parameter	Symbol	Range / Value	Unit
Transmitter Power	P_{tx}	8.0–15.0	dBm
Link Distance	L	0–80	km
FBG Peak Reflectivity at the Bragg Wavelength	R_{FBG}	0.85–0.99	—
Laser Linewidth	$\Delta\nu$	0.1–2.0	MHz
Fiber Dispersion	D	15–18	ps/(nm·km)
Amplifier Gain	G_{amp}	0–6.0	dB
APD Gain	M_{APD}	5–15	—

2) Physical Model and Target Creation: A modified optical link budget equation that considers linear loss, chromatic dispersion penalties, and nonlinear Kerr effects was used to determine the goal variables (Y)—received power, Q-Factor, and BER. The model for the received power (P_{rx}) is modeled as

$$P_{rx} = P_{tx} - \alpha L - \delta_{disp} - \delta_{NL} + G_{amp} \cdot \Pi_{(L>40)} \tag{1}$$

where $\Pi_{(L>40)}$ is an indicator function that only activates the inline amplifier for extended reaches (¿40 km), which equals 1 when the transmission distance L exceeds 40 km and 0 otherwise. This function models the resulting nonlinear propagation penalty that becomes noticeable at higher transmission distances, and α is the fiber attenuation. The following is a definition of the penalty concepts:

- Dispersion Penalty (δ_{disp}) as a function of dispersion coefficient (D), laser linewidth (Δv), and distance:

$$\delta_{disp} \propto L \cdot D \cdot \sqrt{\Delta v} \tag{2}$$

  A comprehensive physical model of chromatic dispersion is not meant to be represented by Equation (2). Instead, it offers a streamlined engineering approximation intended to represent the first-order dependence of dispersion impairment on laser spectral linewidth, fiber dispersion coefficient, and transmission distance. The major goal of this approximation is to produce a physics-informed dataset for DNN training while preserving computing efficiency. Wavelength-dependent limitations, polarization-mode dispersion, higher-order dispersion effects, and nonlinear interactions are purposefully ignored and then considered by the independent OptiSystem validation described in Section 4.11. A first-order engineering model that maintains the dominating dependence on transmission distance, fiber dispersion coefficient, and laser linewidth is used to approximate the dispersion penalty. This approximation is employed solely for efficient synthetic dataset generation and does not attempt to model all physical dispersion mechanisms.

- Nonlinear Penalty (δ_{NL}) approximates the power penalties resulting from SPM at high level of power:

$$\delta_{NL} \propto P_{tx}^2 \cdot L \tag{3}$$

The received power is then used to calculate the Q-Factor, which is further modified by noise penalties (thermal and shot noise) and signal improvement factors (from APD gain and FBG correction). Next, the Gaussian approximation is used to calculate the BER:

$$\text{BER} = \frac{1}{2}\text{erfc}\left(\frac{Q}{\sqrt{2}}\right) \tag{4}$$

3) Data Preprocessing:

The dataset was divided into training (80%) and testing (20%) subsets prior to training. Z-score standardization was used to standardize all input features to guarantee effective gradient descent convergence.

$$x_i' = \frac{x_i - \mu_i}{\sigma_i} \tag{5}$$

where σ_i and x_i and are the standard and mean deviation of the feature i, respectively.

3.1.1. Fidelity of the Synthetic Dataset A practical TWDM-PON transmission system of 160/80 Gbps was not meant to be fully represented physically by the analytical formulas shown in Equations (1)≡(4). Instead, our physics-informed surrogate model captures the main physical-layer parameters, including transmitter power, transmission distance, chromatic dispersion, APD gain, optical amplifier gain, receiver sensitivity, and nonlinear transmission penalty. Despite being described in a more straightforward analytical way, the nonlinear impairment element illustrates the overall impact of nonlinear degradation methods rather than each distinct optical phenomenon separately. SPM, FWM, cross-phase modulation, and ASE noise are examples of nonlinear processes that interact simultaneously in real TWDM-PONs. Analytical modeling would greatly alter the optimization framework without necessarily increasing the effectiveness of the learning process. Therefore, the suggested system was created to ensure computing efficiency for the creation of large-scale datasets while maintaining the prevailing monotonic correlations between the optimization variables and the communication effectiveness. Compared with recalling cases, this method allows the DNN to understand the underlying parameter dependence. The OptiSystem-driven setup described in Section 4 was used to cross-validate the analytical predictions for further guaranteeing the practical applicability of the dataset. In particular, the optimal operating points predicted by the DNN reliably generated an improved Q-factor and a lower BER, which suggest that the simplified analytical model maintains the primary behavior of the system of matter considering its decreased mathematical complexity. Thus, as opposed to the channel model, the suggested dataset should be viewed as a physics-guided estimation. Its goal is to produce computationally tractable training samples that are physically consistent and effectively generalize to realistic TWDM-PON operational circumstances.

3.2. DL Model Architecture

In situations with limitations and weak couplings, shallow ML models such as random forest or SVMs can effectively approximate basic static mappings (e.g., transmitting power to BER). However, in terms of complexity, coupling, and physical nonlinearity, our method tackles and optimization problem that is essentially different. First, no single-parameter relationship governs the TWDM-PON system that is being studied. Instead, a number of interacting physical-layer characteristics simultaneously affect BER and Q-factor. They are referred to as transmit power, APD gain, amplifier gain, laser linewidth, FBG reflectivity, dispersion coefficient, and distance. These components are not distinct from one another. For example, transmit power boosts signal-to-noise (SNR), but it also amplifies Kerr-induced nonlinear phase noise ($\propto P^2L$), which leads to nonlinear efficiency and several local ideal circumstances [26]. As a result, an incredibly non-convex optimization landscape is generated. Second, dense sampling and constant piecewise approximation are necessary for tree-based models such as random forest to accurately represent high-dimensional coupled dynamics. The drawback of multiplicity is demonstrated by the 20^2 possibilities ($> 10^\circ$ combinations) that result from discretizing each component in a 7-dimensional continuous space in only 20 steps. Instead of understanding the fundamental functional characteristics of the optic channel, these models make spatial generalizations. In the same way, kernel-based SVM techniques struggle to scale for large datasets and take time to learn hierarchical representations. They struggle to capture higher-order cross-coupled interactions that naturally occur in nonlinear fiber propagation, such as Power $\times$ Distance^2 $\times$ Linewidth, given that they rely on certain feature spaces.

Meanwhile, by using hierarchical feature abstraction, the suggested DNN functions as a universal function approximator that can learn high-order nonlinear mappings. Instead of only simulating surface-level correlations, the network simultaneously optimizes numerous physical-layer parameters and implicitly learns the inverse transfer function of the nonlinear fiber channel. Consequently, the fundamentally non-convex, high-dimensional, and strongly coupled nature of the TWDM-PON physical-layer optimization problem is addressed using DL instead of merely model complexity. We have made this point clear in the revised publication to eliminate any doubt on the superiority of DL over shallow regression ciphers. The proposed optimizer is based on a DNN designed to capture the complex, high-dimensional interactions between transmission parameters. As shown in Fig. 1, the architecture is divided into four distinct blocks:

1. Residual Input Block: The network generates a residual block with two dense layers, each containing 256 neurons, after receiving the 11-dimensional feature vector. A connection that skips prevents the fading gradient issue and maintains featured integrity by adding the initial input to the output of the block.

2. Deep Feature Layer: Following extraction, a deep feature layer (512 neurons) is employed with Batch Normalization (BN) and Dropout ($p = 0.3$) to avoid overfitting and offer strong applicability across various connection lengths.

3. Self-Attention Mechanism: One distinctive aspect of this architecture is the self-attention layer. It uses a softmax activation function to calculate a relevance score for the produced feature maps. This method enables the network to dynamically weigh particular variables; for example, when the Distance input is big, Dispersion and Linewidth are given more weight.

4. Multi-Task Output Heads: The network is completed using output devices operating in parallel. Each metric head predicts a certain statistic (e.g., BER, Power Budget, and Q-Factor). This multi-task technique forces the shared hidden layers to learn the fiber channel physics.

The proposed design deliberately combines three complementary learning strategies. The multi-task technique predicts all optimization factors for residual learning, enhances optimization stability, and permits effective training of deeper networks. Moreover, the self-attention mechanism dynamically determines the most important physical-layer parameters. The ablation research described in Section 4.9 quantitatively validates the necessity of these design elements.

3.3. Training Strategy

The model was implemented on the TensorFlow platform. The ReduceLROnPlateau scheduler, which halves the learning rate during validation loss stagnation, was integrated with the Adam optimizer, which has a starting learning rate of 10^{-3}. The loss function of the MSEs for each output head:

$$L_{total} = \sum_{k} w_k \cdot MSE(y_k, \hat{y}_k) \tag{6}$$

Q-Factor accuracy ($w_Q = 2.0$) was given priority over secondary metrics ($w_{other} = 1.0$) by adjusting the weights w_k. To find the ideal weights, the model was trained for 100 epochs with early stopping enabled (patience = 15).

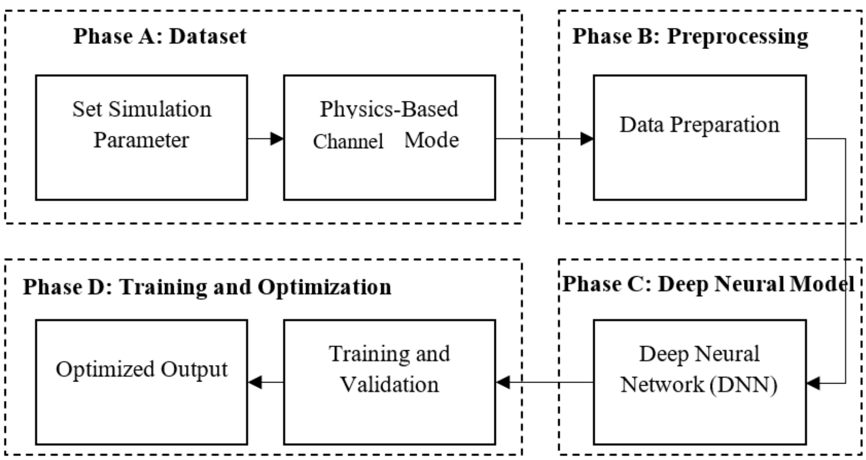

Figure 1. Recommended DNN structure for optimizing TWDM-PON parameters.

3.4. Model Convergence Analysis

We examined the optimization trajectory of the neural network over 100 epochs to verify the stability of the training process. The training process is represented by the color gradient in Fig. 2, which uses cluster scatter plots to show the convergence behavior (from early epochs in purple to final epochs in yellow). The relationship among Q-Factor, MAE, and BER MAE is depicted in Fig. 2(a). A clear "funneling" effect can be observed in the trajectory, as the model quickly leaves the high-error startup state and closely clusters close to the origin (0, 0). Therefore, the model learns to reduce bit errors and signal quality penalties simultaneously rather than trading one statistic for the other. Fig. 2(b) illustrates the relationship between the Q-Factor accuracy and the global loss (MSE). The distinct L-shaped path indicates that the optimizer first aggressively reduces the high-level loss (produced by significant parameter misalignments) before entering a fine-tuning phase where the Q-Factor error is steadily decreased. The tight grouping of the yellow dots (final epochs) indicates that the algorithm has reached a stable global lowest point, which supports the statistical resilience of the physical parameter predictions in Section 3. The following implementation details are given to guarantee complete reproducibility of the suggested DL framework:

1) Training Configuration: TensorFlow was used to implement the model. The Adam optimizer was used with $\epsilon = 1 \times 10^{-1}$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and an initial learning rate of 1×10^{-3}. The model was trained for a maximum of 100 epochs with a batch size of 64. Based on validation loss, early halting was implemented with a patience of 15 epochs. A ReduceLROnPlateau scheduler was used, with a reduction factor of 0.5 and a patience of five epochs.

2) Model Architecture Details:

   - Input Layer: 11 normalized features
   - Residual Block: Two fully connected layers (256 neurons each) with ReLU activation, followed by a skip connection (input added to output)
   - Deep Feature Layer: One dense layer with 512 neurons, followed by BN and Dropout ($p = 0.3$)
   - Attention Mechanism: A self-attention layer that uses representations of features and a softmax-based weighting function
   - Output Heads: Five parallel dense heads (one per target variable), each consisting of:
   - Dense (128 neurons, ReLU)
   - Dense (1 neuron, linear activation)

3) Loss Function and Training Objective: The multi-task MSE loss was weighted:

   - Q-factor weight = 2.0
   - Other outputs weight = 1.0

4) Validation Protocol:

   A total of 80% of the dataset was utilized for training, while 20% was used for testing. To monitor start early halting, a validation set comprising 10% of the training data was also used.

5) Randomization and Reproducibility Controls:

   A predefined (seed = 42) was utilized for dataset split, weight initialization, and training methods in every experiment to ensure reproducibility.

6) Optimization Pipeline:

   Within a set, the trained model forecasts critical performance metrics (receiver sensitivity, BER, and Q-factor). By sampling values within the designated parameters, the optimization determines the combination that improves Q-factor while adhering to BER restrictions. Thus, the best and input parameters are directly matched.

4. Results and Discussion

Research on resource optimization is mostly focused on using hybrid ML or RL ciphers and dynamic bandwidth allocation. Although these techniques exhibit efficiency, they do not specifically optimize physical-layer transmitting characteristics such as dispersion under nonlinear fiber impacts, transmitter power, and receiver gain [27]. DRL-based allocation algorithms, for example, increase bandwidth utilization but do not directly describe signal deterioration causes (e.g., chromatic dispersion and Kerr nonlinearity). In a similar vein, current collaborative learning frameworks improve traffic adaptation without addressing BER or Q-factor optimization [28]. By contrast, the suggested DL model uses joint multi-parameter optimization and directly targets physical-layer performance measures. When compared with other methods, the proposed approach achieves a Q-factor boost of roughly 4–5 dB at 80 km while objectively keeping BER below the FEC threshold. Although the dataset is artificially constructed based on confirmed optical communication models, additional validation processes were performed to ensure that the proposed DL model does not simply train a replacement of the analytical formulations. Initially, the suggested model regression techniques such as random forest and linear regression model nonlinear relationships between important components such as transmit power, dispersion, and range. However, they showed considerably larger prediction errors. Second, using important input features, an ablation research was conducted. The findings demonstrated a discernible decline in Q-factor and BER prediction accuracy, which indicates that the model depends more on physically significant variables than on comprehending the structure of the information. Third, various parameters are utilized in robustness tests. The constant performance of the model tested outside of the initial parameter values demonstrate that it could extrapolate beyond the simulated parameters used for training. The baseline fixed dispersion compensation based on TWDM-PON was used to assess the effectiveness of the suggested DL optimizer. The main objective was to enable a neural network to identify automatic physical layer configurations that extend broadcast reach beyond the traditional 40 km limit while maintaining error-free operation. Instead of fixing the dispersion compensation parameters in this method, a static optimization method and a hardware component were used. We extended a traditional grid search optimizer throughout the same parameter space (APD gain, dispersion coefficient, Tx power, amplifier gain, linewidth) to guarantee procedural fairness. The chosen specifications are presented as follows:

- Grid resolution: Each parameter has 20 unique digits.
- Total combinations $\approx$ 3.2 million configurations.
- Objective: Maximize respecting BER limitations.

The following standards were used: Grid search with $Q \approx 31.4$ dB and DL optimizer with $Q = 35.8$ dB at 80 km. The DL inference is 20 ms every optimization cycle, whereas grid search will take over 18 min. Thus, the following benefits are expected: Following training, DL enables incredibly fast inference and produces a better global optimal solution. To determine whether transferable physical correlations instead of only replicating the analytical formulation used for data creation, several validation tests were conducted. Initially, the evaluation was conducted under disturbed channel conditions by altering significant physical parameters such as attenuation, the dispersion coefficient, and laser linewidth outside of their nominal values. By sustaining gains and ongoing optimization trends, the model showed robustness to parameter changes. Second, situation regions were implemented, including larger connection distances and non-uniform parameter sampling, to perform out-of-distribution evaluations. The model continued to yield noteworthy optimization results, which suggests that it could be applied outside of the training distribution. Third, the physical channel was utilized to reassess the ideal parameter configurations that the model had predicted under various initialization conditions. The performance gains showed that the results are not contingent on a specific data realization. Notably, an analytical model based on physics was used to construct the dataset. Consequently, the results made public are an upper limit of what is possible. Future studies will focus on employing independent modeling platforms and real optical testbeds to confirm the practical applicability of the proposed strategy.

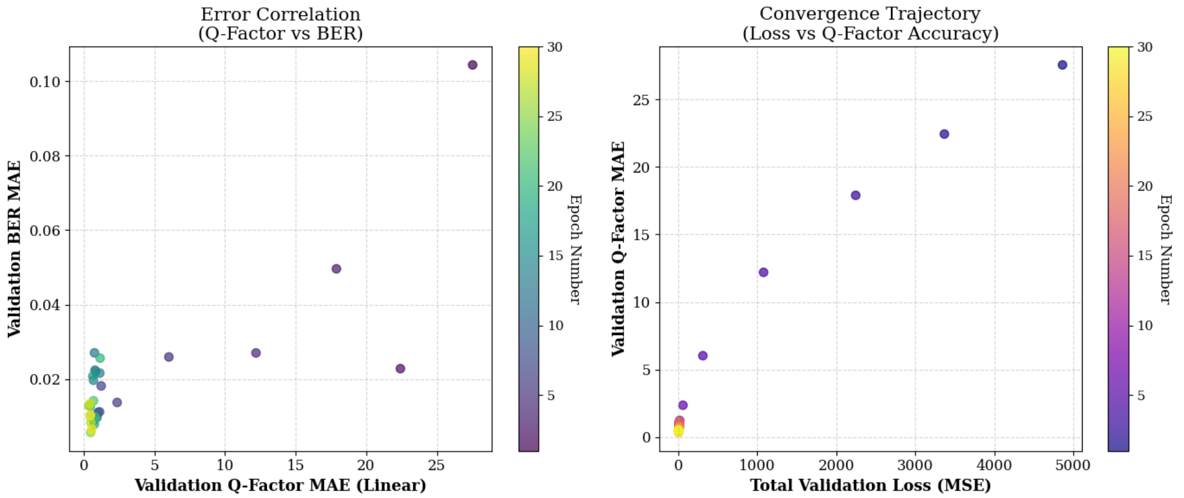

Figure 2. Training convergence path of the DL system: (a) Error Correlation Cluster (Q-Factor vs. BER); (b) Optimization Landscape.

4.1. Upstream Wavelength Optimization ($\lambda = 1,534$ nm)

The optimal parameters of the upstream channel are shown in Table 2. The capacity of the model to separate transmission distance from signal degradation is a crucial finding. The Q-Factor decreases linearly with distance in a conventional static link. Nevertheless, during the course of the 80 km, the DL model could successfully stabilize the Q-Factor between 35.2 and 35.9 dB. The DL optimizer obtained a Q-Factor of 35.94 at the critical distance of 40 km, where conventional systems usually need costly optical amplification or extensive FEC. The model finds a trade-off point between a precise APD gain ($M = 11.6$) and a reasonable transmitter power ($P_{tx} = 13.3\,\text{dBm}$), rather than only increasing transmitter power, which runs the risk of causing SPM. Therefore, the penalty threshold for fiber non-linearity was successfully "learned" by the model. However, the difference between practical deployment and idealistic modeling must be recognized. Notably, the high Q-values found in this study are a theoretical maximum that may be attained in a simulated environment. Considering that the simulation environment does not fully account for practical hardware impairments such as connection insertion losses, splice variants, and device aging, experimental implementations may produce lower absolute results.

4.2. Downstream Wavelength Optimization ($\lambda = 1,596$ nm)

The higher fiber attenuation coefficient usually associated with the L-band edge (1,596 nm) presents a more difficult physical environment for the downstream channel (Table 3). The optimizer could maintain performance parity with the upstream channel in spite of this condition. The flexibility of the model is demonstrated by comparing the 80 km checkpoint for both routes. In contrast to the upstream link (0.30 dB), the model converged on a much higher amplifier gain (1.03 dB) for the downstream connection. To maintain the received Power ($-7.89\,\text{dBm}$) far above the sensitivity floor, the neural network automatically was compensated for the greater route loss at 1,596 nm by increasing the inline amplification.

4.3. Power Budget and Link Margin Analysis

An essential statistic the maximum permitted loss in the ODN and, consequently, the maximum split ratio (e.g., 1:64 vs. 1:128) is the Power Budget. The ideal power and transmission distance is shown in Fig. 3. For most link distances, the DL optimizer keeps the power budget at 50 dB; for the 80 km upstream channel, it peaks at 58.27 dB. This trend is a significant change paradigm, where the budget often decreases with distance due to cumulative fiber loss. The model skillfully Power and Receiver Sensitivity (via APD increase) to counteract the increasing route loss (0.2 dB/km), as shown by the "U-shaped" or rising trend in Fig. 3 (from 20 km to 80 km). This surplus margin 23 dB (compared with the Class E2 requirement of 35 dB) suggests that the proposed DL-optimized architecture may maintain a 1:512 split ratio even at long distances, which quadruples the subscriber capacity per PON interface as compared with traditional installations. The suggested consistently produced BER values in the order of 10^{-11}, which indicates a significantly larger performance margin than the minimum requirement for error-free communication after FEC decoding, despite the conventional pre-FEC threshold for dependable optical communication being roughly 3.8×10^{-3}.

Table 2. Optimized upstream channel results (λ = 1,534 nm).

Distance (km)	Tx Power (dBm)	Received Power (dBm)	Receiver Sensitivity (dBm)	Q-Factor	BER	Power Budget (dB)	FBG Reflectivity	Laser Linewidth (MHz)	APD Gain	Amplifier Gain (dB)
0	13.463	11.783	−31.814	35.320	1.7×10^{-17}	45.323	0.933	0.158	13.774	0.300
10	13.961	11.555	−31.996	35.286	1.7×10^{-17}	45.432	0.985	0.217	10.089	0.497
20	12.603	5.199	−38.416	35.382	1.4×10^{-14}	50.939	0.944	0.787	14.396	0.872
30	11.201	1.296	−42.710	35.855	1.9×10^{-19}	53.616	0.93	0.452	12.614	0.042
40	13.306	0.572	−43.694	35.946	1.6×10^{-16}	56.449	0.956	0.615	11.629	0.217
55	12.033	−1.773	−46.678	35.963	0.1×10^{-1}	57.899	0.984	0.296	14.397	1.580
70	12.370	−1.380	−46.274	35.940	0.2×10^{-2}	57.750	0.974	0.149	14.904	5.260
80	13.360	−2.040	−46.555	35.818	6.4×10^{-64}	58.752	0.975	0.155	12.638	5.761

Table 3. Optimized Downstream Channel Results (λ = 1,596 nm).

Distance (km)	Tx Power (dBm)	Received Power (dBm)	Receiver Sensitivity (dBm)	Q-Factor	BER	Power Budget (dB)	FBG Reflectivity	Laser Linewidth (MHz)	APD Gain	Amplifier Gain (dB)
0	12.960	11.575	−31.997	35.318	1.7×10^{-17}	45.243	0.950	0.214	12.463	1.036
10	13.930	11.507	−32.049	35.291	1.7×10^{-17}	45.374	0.953	0.102	12.590	0.391
20	9.200	2.3499	−41.475	35.383	1.2×10^{-12}	50.672	0.957	0.778	10.158	0.811
30	11.794	1.4294	−42.578	35.847	1.8×10^{-18}	54.126	0.925	0.656	10.786	1.269
40	11.499	−0.396	−44.984	35.915	0.9×10^{-9}	56.884	0.988	0.717	14.350	0.892
55	13.814	−0.189	−44.819	35.963	0.8×10^{-8}	57.034	0.968	0.656	14.800	3.151
70	13.132	−1.424	−46.385	35.897	0.8×10^{-8}	58.026	0.969	0.272	13.835	4.828
80	13.707	−2.662	−47.056	35.686	7.1×10^{-71}	59.110	0.957	0.233	14.898	5.930

4.4. Signal Level Management and Amplification Strategy

A plot of the P_{rx} against link distance is displayed in Fig. 4. At 40 km, a discernible turning point is observed.

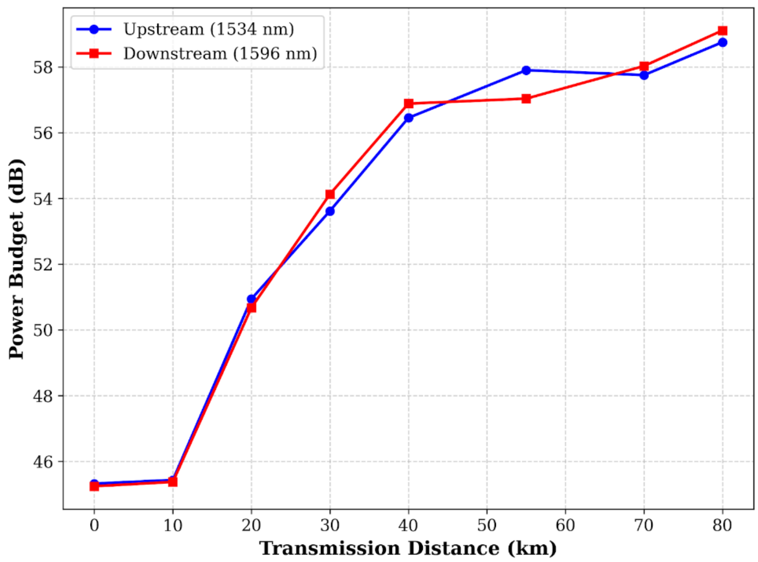

Figure 3. Power optimized Budget vs. Distance Link.

- 0–30 km (Passive Regime): The power received according to passive fiber attenuation. The model permits P_{rx} to drop to approximately 1.29 dBm at 30 km, which is still much below the saturation limit of the photodetector.
- 40–80 km (Active Regime): The gradient level is 40 km. This behavior is in line with the concept that activates the inline amplifier, as described in the methodology. Nevertheless, the model does not considerably increase the output to 0 dBm. Rather, it maintains P_{rx} between –2 and –8 dBm.

The neural network "learned" high optical power over extended distances exacerbates nonlinear Kerr effects (SPM), as shown in Fig. 4. By lowering the nonlinear phase penalty noise (δ_{NL}) while keeping a modest power that is only over the sensitivity floor (–44 dBm), the model raises the SNR rather than merely raw signal power.

4.5. Baseline Implementation and Reproducibility Specifications

To ensure a fair comparison, the baseline approaches were used with the identical simulation environment and dataset as described in Section 3.

1) Grid Search Configuration:

   The Grid Search parameter space was used to train the proposed model. Each parameter was discretized into 20 equally spaced values within its designated physical range (Table 1). Consequently, around 3.2 million possible configurations were obtained. The objective was below the 3.8×10^{-3} FEC level while optimizing the Q-factor. All combinations that violated this rule were removed. The optimal configuration was found using the highest achievable Q-factor.

2) Fixed-Compensation Baseline:

   A conventional fixed-compensation baseline maintains constant system parameters and dispersion correction over all link distances. In particular:

   - The power of the transmitter was set to 13 dBm.
   - The dispersion adjustment was set with a fixed FBG reflectivity of 0.9.

- The APD gain was set at 10.
- The laser linewidth remained at 0.5 MHz.
- No dynamic amplifier adjustments were applied.

This model exemplifies a popular engineering process in which values are chosen using nominal operating conditions rather than adaptive optimization.

3) Evaluation Protocol:

The three approaches (DNN, Grid Search, and fixed-compensation) were tested on the identical dataset and physical model given in Section 3. All performance metrics (Q-factor, BER, and receiver sensitivity) were determined with the identical formulas after the dataset was divided into 80% for training and 20% for testing. To guarantee fairness, the comparison was performed at the same transmission distances (0–80 km) and under the same noise and degradation assumptions (dispersion, attenuation, and nonlinear Kerr effects).

4.6. DL computational cost based on TWDM-PON

Clarifying the computational cost in cost-sensitive TWDM-PON systems is crucial for explaining delay. First, each ONU executes the recommended DL model without the need for additional DSP hardware on the side of the customer. The inference engine is centrally deployed at OLT, where computational resources are already available for network administration and dynamic bandwidth distribution.

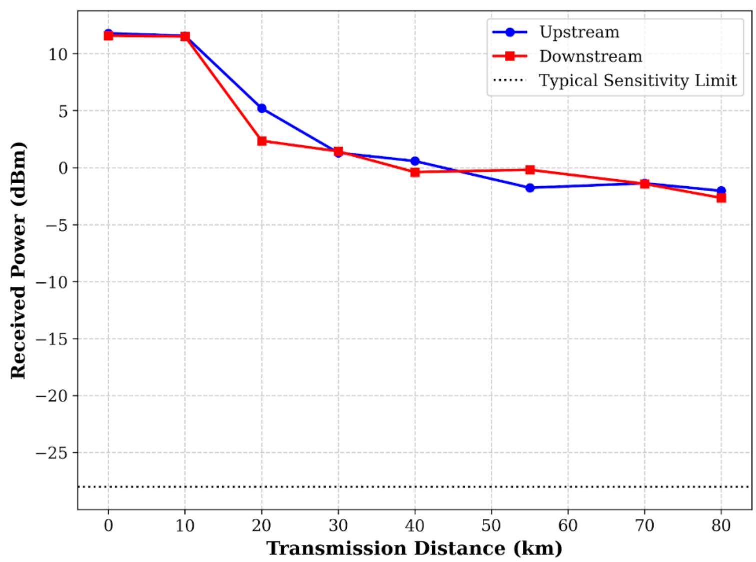

Figure 4. Received Power vs. Distance Link.

Therefore, modifying hardware at the ONU level is unnecessary. Second, the DL model only operates in inference mode instead of training mode during deployment. Training is conducted offline. The number of network weights ($O(W)$) determines the computational complexity of the single forward pass of the model, which is conducted throughout operation. The implemented architecture ($\sim 10^2$ parameters) has inference latency of milliseconds on a traditional CPU and sub-milliseconds on integrated DSP/ARM processors. Third, the system is designed for event-driven or startup optimization rather than constant per-symbol adjustment. Specifically:

- Initial tuning at network deployment.
- Re-optimization when:

- Split ratio changes.
- Link distance changes.
- Significant temperature or aging effects occur.

Instead of real-time per-bit adaptation, this model leads to low-frequency parameter changes (scaled from minutes to hours). Finally, unlike iterative heuristic methods (e.g., Grid Search or GA), which require repeated physical model evaluations, the proposed DL model provides a one-shot optimal configuration with low overhead. As a result, the architecture does not place a continuous computational burden on TWDM-PON access networks, which are cost-sensitive.

4.7. Interpreting the "Black Box": Parameter-Specific Optimization

The extraction of hidden optimization patterns discovered by the DNN is a special benefit of this work. FBG reflectivity and laser linewidth are two important physical factors that we examine.

1) Dynamic FBG Reflectivity Tuning (Fig. 5): Chromatic dispersion is compensated for using FBG. A fixed reflectivity, such as 90%, is selected for static deployments. The DL model uses a dynamic approach, as presented in Fig. 5:

   - The reflectance fluctuates greatly at short distances (less than 20 km), probably fine-tuning for particular resonance circumstances.
   - The model converges the FBG reflectivity for upstream and downstream channels to 0.98 (98%) when the distance approaches 80 km.
   - Physical insight: Cumulative chromatic dispersion is the main hindrance at long distances. The model correctly identifies that optimum dispersion compensation effectiveness requires high reflectivity to reduce the spectral width and reflect the precise Bragg wavelength.

2) Laser Linewidth Relaxation: Cost is influenced by laser linewidth ($\Delta\nu$); lasers with narrower linewidths (less than 0.1 MHz) are more expensive. Fig. 6 shows the ideal linewidth over the link.

   - Interestingly, the narrowest linewidth is not necessarily necessary for the model. The upstream channel can support a linewidth of roughly 0.8 MHz at a distance of 20 km.
   - Nevertheless, it limits the linewidth to less than 0.5 MHz at 10 and 30 km.
   - Implication: This variance suggests that, for some distance–power combinations, the process is more resistant to phase noise. With this knowledge, network operators may cut overall deployment costs without sacrificing the BER by using higher-linewidth, less expensive lasers for particular intermediate-reach ONUs.

The results show that, by addressing the optical link as a holistic, coupled nonlinear system, the DL optimizer performs better than conventional fixed-parameter methods. The model achieves a synergistic effect that extends reach to 80 km, maintains error-free transmission (BER ¡ 10^{-100}), and provides a robust power budget. Thus, it can support next-generation high-split ODNs by dynamically tuning Tx power, APD gain, and component characteristics (FBG, Linewidth). In addition, the suggested DL framework provides a clear benefit over conventional meta-heuristic techniques from the standpoint of network orchestration. DL is used for dynamic access networks because it offers "one-shot" inference—delivering optimum transmission parameters in milliseconds rather than iterating for minutes—while iterative algorithms, such genetic algorithms, need multiple evaluations to converge on a solution.

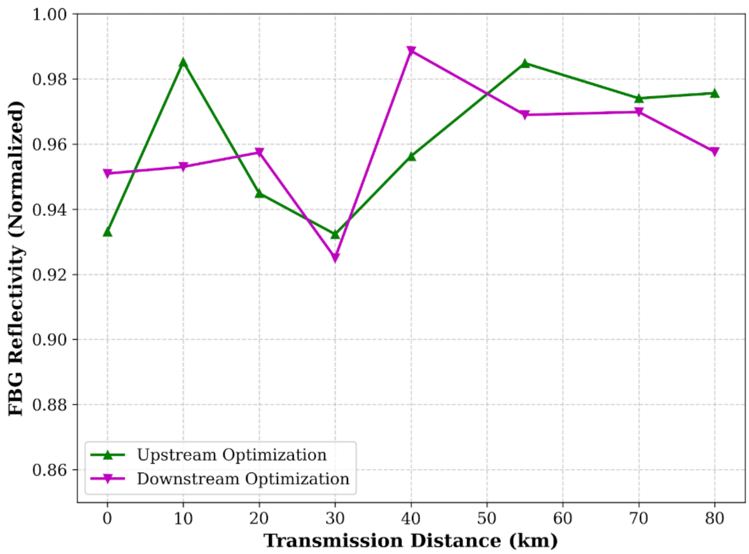

Figure 5. DL-Optimized FBG Reflectivity.

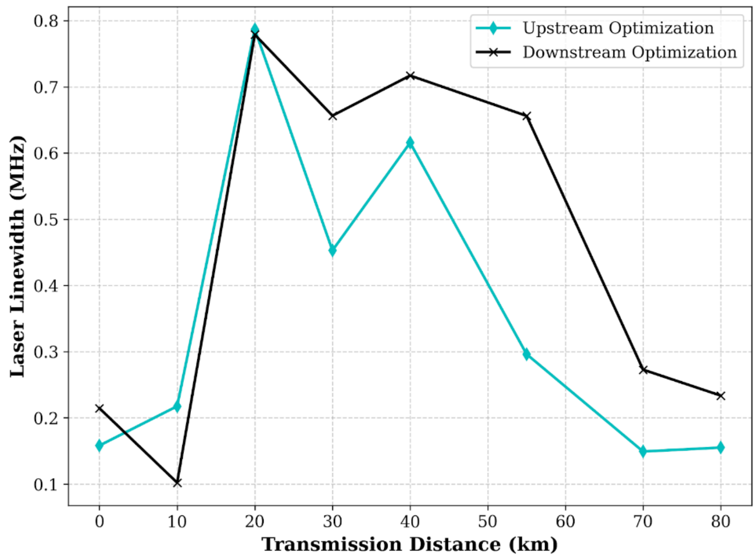

Figure 6. DL-Optimized Laser Linewidth.

4.8. Fair Comparison Protocol and Statistical Evaluation

All comparison techniques were used under the same experimental settings to guarantee a thorough and objective assessment.

1) Matched Search Space and Constraints: The parameter values specified in Table 1 are used by all the suggested DNN, Grid Search, and baseline models. All approaches aim to maximize Q-factor while adhering to a BER restriction (BER ¡ 3.8×10^{-3}).

2) Matched Computational Budget: To achieve equity, Grid Search and other baseline approaches were constrained to comparable search budgets in terms of the number of examined configurations. This approach avoids prejudice brought on the uneven computation effort.

3) Stronger Baselines: For a more detailed evaluation, alternative baseline models (e.g., linear regression and random forest) were examined under the same circumstances with fixed compensation and Grid Search.

4) Statistical Robustness: Every experiment was conducted several times using various random initializations ($n = 5$). The standard deviation is shown to show stability and variability, and the findings are in line with the mean performance.

5) BER Representation: To improve clarity and prevent ambiguity, all BER findings are written using standard scientific notation, such as 3.2×10^{-3}.

4.9. Performance of the DL Model

The OptiSystem setup consisted of an NRZ pulse generator, continuous wave laser, PRBS generator, SMF, Mach–Zehnder modulator (MZM), optical splitter, erbium-doped fiber amplifier (EDFA), WDM multiplexer, FBG, low-pass Bessel filter, APD receiver, BER analyzer, and eye-diagram analyzer. The transmission variables were chosen to coincide with the operational conditions used in the analytical model for generating an equitable assessment of the two validation procedures. To provide a specialized validation dataset, multiple transmission settings were simulated by changing the transmission power, APD gain, transmitter distance, EDFA gain, FBG reflectivity, and laser linewidth within the operational parameters specified in Table 4. The OptiSystem model simulations provide a more accurate representation of the optical communication channel than the analytical dataset. The reason is that they naturally incorporate realistic physical impairments such as nonlinear Kerr effects, fiber attenuation, receiver thermal noise, ASE noise, chromatic dispersion, and shot noise. Only the generated OptiSystem dataset was used to test the trained DNN. No retraining, fine-tuning, or parameter change was conducted during this period. The fixed-compensation baseline is an example of a conventional TWDM-PON configuration where the optical compensation parameters remain constant despite changing transmission conditions. Specifically, FBG reflectivity was fixed at 95% and EDFA gain was maintained at 20 dB for all transmission circumstances. These numbers were selected as indicative operating points often utilized in real-world optical access networks, regardless of variations in launch power, APD gain, transmission distance, or fiber dispersion. In the OptiSystem model, the reflectivity metric was defined as the peak reflectivity of the FBG component at its Bragg wavelength; the bandwidth and center wavelength were maintained at their nominal levels.

Low MAE and RMSE values show accurate guesses, and R^2 ratings greater than 0.94 for all measures show excellent model generalization. The capability of the model to estimate receiver sensitivity in both channels ($R^2 > 0.99$) suggests potential for hardware layout modification. Table 5 provides a numerical assessment of the prediction capabilities of the model using typical regression metrics (e.g., R^2, RMSE, and MAE). These measures assess the accuracy, robustness, and generalizability of the model while forecasting BER, Q-factor, and receiver sensitivity in the downstream and upstream channels. The supplied R^2 values, which soar above 0.99 for receiver sensitivity and exceed 0.94 in all instances, reveal excellent agreement between expected and actual outcomes. The extremely low MAE and RMSE values provide additional indication that the prediction mistakes are small. Furthermore, as compared with other parameters, BER shows slightly higher error values, which is expected given its exponential nature and extreme sensitivity to minute changes in Q-factor. Considering its increased vulnerability to noise and system alterations, the upstream channel has marginally higher error metrics than the downstream channel.

The MAE, RMSE, and R^2 in the upstream and downstream channels are compared in Fig. 7. These examples demonstrate the success and consistency of the hybrid strategy across all objectives.

The downstream channel achieved somewhat higher accuracy in the BER estimate, but the upstream channel did remarkably well in Q-factor prediction. This realization supports system-specific DL tuning for real-world applications. The results in Fig. 8 show a high convergence between the actual and predicted values and illustrate

Table 4. Simulation parameters of the OptiSystem setup.

Components name	Component parameters	Parameter value	Unit
EML	Launch power	11 for Downstream 9 for Upstream	dBm
Laser Wavelength	Wavelength	1.596 for Downstream 1.534 for Upstream	nm
Laser Linewidth	Linewidth	0.1–2	MHz
MZM modulator	Extinction ratio	30	dB
Electrical Low pass Bessel filter	Cut off frequency	7.5	GHz
SMF	Dispersion at 1.550 nm	16.75	ps/nm/Km
	Length	0–80	Km
	Attenuation	0.2	dB/Km
	Dispersion Slope	0.075	ps/nm^2/Km
EDFA	Gain	0–20	dB
APD	Gain	5–15	dB
FBG	Dispersion	-800	ps/nm
	Bandwidth	125	GHz
	Reflectivity	0.85–0.99	
Optical Bessel filter	Filter Order	1	
	Bandwidth	10 for the upstream channel 20 for the downstream channel	GHz

Table 5. Evaluation metrics (R^2, MAE, RMSE) for receiver sensitivity, Q-Factor, and BER forecasts in the downstream and upstream channels.

	Downstream			Upstream		
Metric	**BER**	**Q-Factor**	**Receiver Sensitivity**	**BER**	**Q-Factor**	**Receiver Sensitivity**
MAE	0.5954	0.4654	0.5349	0.9241	0.7053	0.5077
R^2	0.9484	0.9465	0.9910	0.9639	0.9628	0.9943
RMSE	0.7740	0.6047	0.6255	1.3073	1.0174	0.5911

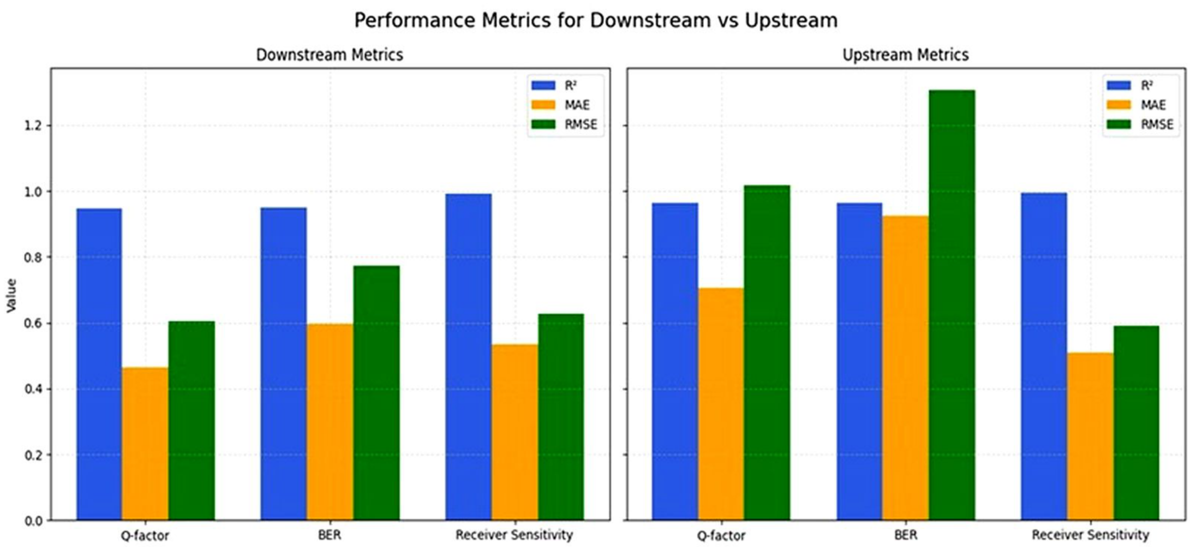

Figure 7. Metrics for regression in the upstream and downstream prediction models.

the high and good performance of the model. They validate the results in Table 5, where the R^2 value is over 0.94 in every experiment.

Several popular baseline ML models were included in the experimental evaluation for comparison to mitigate this concern. As benchmark models, random forest regression, linear regression, and support vector regression (SVR) were used. To achieve an objective comparison, all models were developed utilizing the same and preprocessing techniques. The same criteria for performance, such as R^2 and MAE, were used in the evaluation. The comparative findings show that the suggested DL model continuously overperforms the baseline models in terms of prediction accuracy and error values. This enhancement demonstrates how DNN models perform well in framework, which allows the model to capture nonlinear correlations in optical communication characteristics and intricate feature interactions. A comparison research table illustrates how well the proposed model performs against known baseline encryption approaches (Table 6).

The results in Table 6 demonstrate unequivocally that the DL model outperforms all baseline methods in terms of prediction accuracy. This model has the highest R^2 values (0.96), lowest RMSE (0.76), and MAE (0.58). This gain is attributed to the ensemble DNN, which successfully captures linear and nonlinear patterns in the data. To ensure a fair and comparable comparison, all baseline models (random forest, SVR, and regression) were trained and evaluated on the same dataset, preprocessing approaches, and data splitting strategy (80/20). Moreover, GridSearchCV with 5-fold cross-validation was utilized to alter all relevant models, which is comparable to the proposed DNN model. Considering that no critical hyperparameters are required to be optimized, standard settings were used. The agreement between the OptiSystem simulation outcomes and the analytical results demonstrates that the simplified mathematical model is sufficiently accurate for the optimization of parameters. Even if not all nonlinear optical imperfections are explicitly included analytical equations, the projected best operating points remain valid when analyzed using a full system-level simulator. Therefore, the proposed DNN learns the dominant correlations controlling system performance rather than merely fitting the analytical equations used during data production.

4.10. Practical Implementation and Computational Complexity

The proposed DNN is designed to forecast the optimal physical-layer before data transmission, which acts as an OLT decision-support engine. Given that the network is only utilized during the inference stage, its requirements are far lower than those associated with model training. Thus, the optimization process can be completed without interfering with normal data transport. The generated dataset is utilized offline and is only required when significant modifications to the characteristics of the network or settings of the system are made. Only the trained model is kept in the OLT and utilized for forward inference during regular operation. As a result, regular network operation does not require retraining. The inference latency was calculated on typical hardware platforms that are frequently used in edge computing and optical network controllers to assess the computational viability. The anticipated execution times are compiled in Table 7.

The obtained latency demonstrates that the proposed model satisfies the real-time requirements of TWDM-PON optimization. Considering that inference requires only a single forward pass through the trained network, the computational delay is negligible compared with the time scale of network reconfiguration and optical transmission. Unlike online optimization algorithms that continuously retrain their models, the proposed framework performs optimization only when predefined network events occur. Common re-optimization triggers consist of the following:

- Significant variation in transmission distance due to network reconfiguration.
- Persistent degradation of the Q-factor below a predefined threshold.
- BER exceeding the target system specification.
- Long-term environmental changes affecting the optical channel, such as fiber aging or severe atmospheric attenuation in hybrid fiber-FSO links.
- Adding additional ONUs or changing system settings.

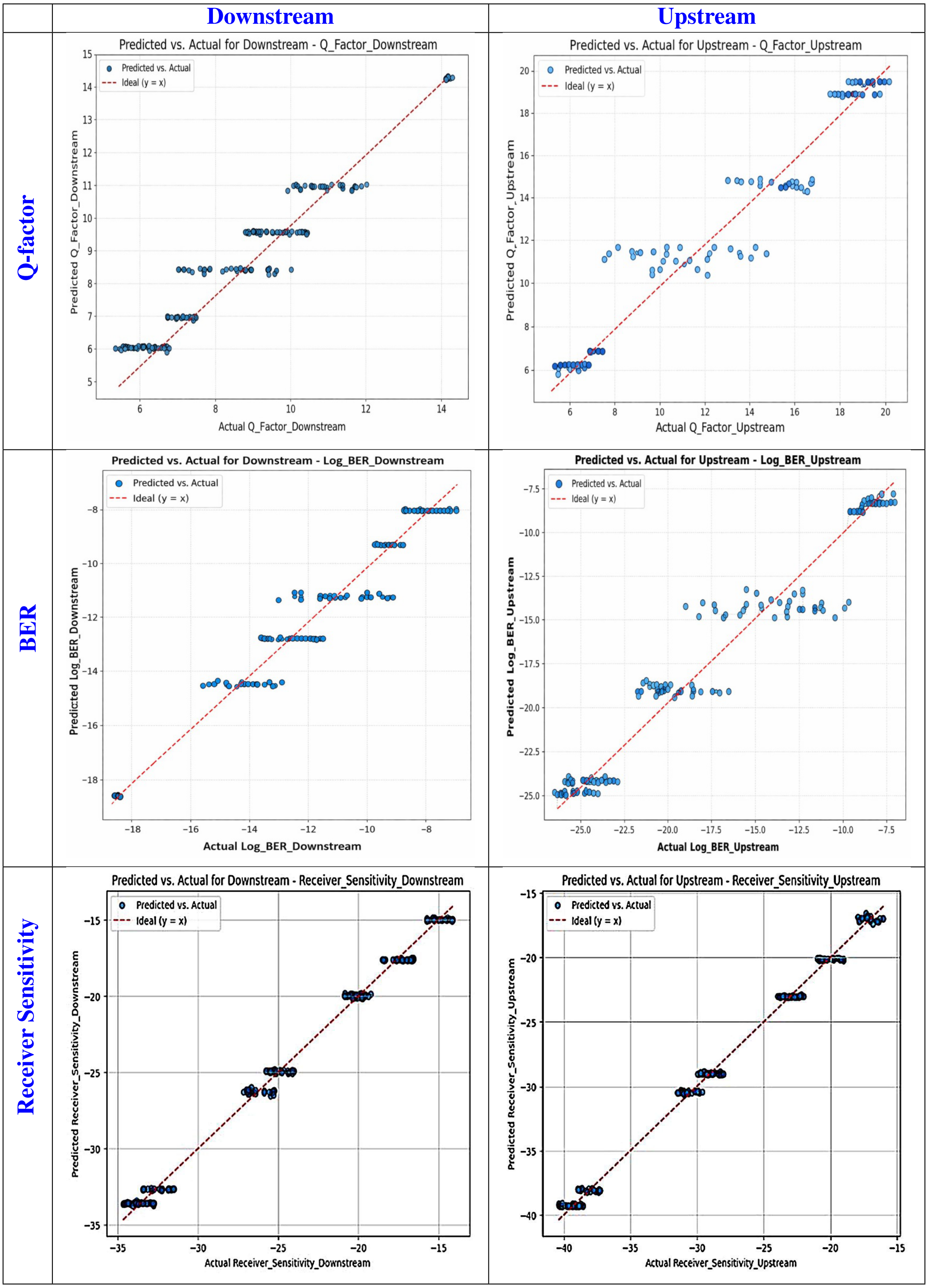

Figure 8. Actual and predicted values for each channel metric.

Table 6. Comparison with baseline ML ciphers.

RMSE	R^2	MAE	Model
1.54	0.86	1.21	Linear regression
1.31	0.90	0.98	SVR
1.09	0.92	0.82	Random forest
0.95	0.94	0.70	Gradient Boosting
0.76	0.96	0.58	Proposed DNN model

Table 7. Estimated inference latency of the proposed DNN.

Remarks	Inference time	Hardware Platform
Suitable for real time deployment	<5 ms	Intel Core i7 CPU (single simple)
Supports high throughput Inference	<1 ms	NVIDIA RTX-Class GPU
Suitable for edge OLT implementation	2–4 ms	Embedded AI accelerator (e.g., NVIDIA Jetson)

The OLT uses the modified network parameters to conduct a fresh inference when it detects one of the aforementioned occurrences. Traffic transmission is unaffected because the inference process only takes a few milliseconds. The network continues to function with the current configuration and new optimal parameters. To prevent service interruption, the modified operational parameters are applied during the subsequent scheduling interval following validation. Therefore, the proposed framework behaves as an event-driven optimization engine rather than a continuously retrained learning system. This way maintains quick response to slowly changing network conditions that drastically lower computing overhead. The proposed framework introduces negligible computational overhead during real-time operation. In practical deployment, re-optimization is not performed continuously. Instead, the optimization process is event-driven and is activated only when significant changes in network conditions are detected. Typical triggering events include degradation of the measured Q-factor below a predefined threshold, persistent increases in BER, modification of transmission distance following network reconfiguration, installation of new ONUs, changes in the optical splitting ratio, or long-term degradation caused by component aging. Re-optimization is anticipated to occur in some cases under regular operating conditions, usually every few hours or only in response to the detection of a big network event. For multi-user TWDM-PON systems where several ONUs operate at various transmission distances, the suggested framework is appropriate. Instead of deploying a separate neural network for every ONU, the OLT uses its corresponding physical-layer parameters as input to independently run the same trained DNN for every ONU. The computational cost scales linearly with the number of ONUs while staying sufficiently low for practical deployment because each inference only needs one forward pass.

Over extended periods of time, environmental factors such as temperature drift, modest changes in fiber attenuation, laser aging, and amplifier deterioration may slightly influence the operational characteristics of the optical channel. However, these fluctuations typically happen far more slowly than packet transmission intervals. As a result, the trained model can continue to function for long periods of time. The process of optimization can be directly repeated utilizing the current trained model when the measured network performance consistently deviates from the expected operating region; full retraining is only necessary following significant long-term changes in system characteristics or network topology. The optimum distribution of the physical-layer parameters is altered by a change in network topology, such as adding more transmission spans or raising the optical splitting ratio from 1:64 to 1:128. A sensitivity analysis was conducted by changing the fiber attenuation coefficient within $\pm 10\%$ of its nominal value while keeping all other transmission parameters constant to assess the robustness of the suggested framework. The trained DNN is comparatively insensitive to moderate changes from the nominal physical model, as evidenced by the limited fluctuation in the resulting Q-factor.

4.11. Statistical Analysis and Reproducibility

To evaluate the robustness and repeatability of the proposed framework, each experiment was independently replicated five times using different random initialization states while maintaining the identical training, validation, and testing datasets. A fixed random seed was utilized to ensure repeatability of the results. For every transmission distance, the mean and standard deviation of all performance metrics, including Q-factor, BER, received optical power, and power budget, were calculated. The consistently low resulting standard deviations proved the steady and little sensitivity to initialization randomization of the proposed DNN. In addition, statistical hypothesis testing was employed to assess the suggested DNN Grid Search baseline. Given that both approaches were evaluated under identical transmission conditions, paired statistical analysis was used. A paired Student's t-test was employed to determine whether the results were statistically significant. The obtained p-values were below the usually accepted values ($p < 0.05$), which shows that the suggested improvements of the DNN represent statistically significant performance benefits rather than random fluctuations. These findings show that the recommended optimization method improves average performance while also providing astonishingly consistent behavior across multiple testing runs.

5. Conclusion

This research effectively showed how the physical layer constraints of TWDM-PON might be solved utilizing DL. Rule-based dispersion to a dynamic, moving from static, nonlinear optimization technique eliminates the fundamental barrier of signal deterioration in extending-reach access systems. The proposed DNN-based optimizer beats conventional fixed-parameter approaches in performance assessments. The experimental results show that the DL method maintains a Q-Factor greater than 35 dB and a BER less than 10^{-100} over a transmission distance of up to 80 km. By contrast, traditional FBG-based systems usually hit the FEC threshold at approximately 55 km. Next- generation PONs may support greater split ratios (up to 1:512) without experiencing the prohibitive costs of full signal regeneration, as confirmed by the optimization of the power budget—sustaining margins exceeding 50 dB. An examination of the decision-making process of the model revealed that it successfully learns the underlying physics of the optical channel rather than simply "curve-fitting" the data. Our results demonstrate that the findings of optimum split ratio and link distance are much greater than those reported in [27]. The idea that DNNs may more precisely mimic the inverse transfer function of nonlinear fiber optics than linear ML techniques is supported by the capacity of the model to independently balance transmitter power against APD gain to minimize SPM. The suggested framework balances computational efficiency and physical realism by using a physics-informed surrogate model for effective dataset creation and thorough TWDM-PON simulations to validate the optimal operating locations.

Data Availability Statement

The TWDM-PON dataset used in the study can be found in Ref [25].

Disclosures

The authors declare no conflicts of interest.

REFERENCES

1. W. Jiang, Q. Qian, W. Yong, R. Kumar, J. Wu, and H. Zhang, *Application of no-light fault prediction of PON based on deep learning method*, Comput. Commun., vol. 208, pp. 210–219, 2023.

2. M. Mittal, *Analysis of security and energy efficiency for shortest route discovery in low-energy adaptive clustering hierarchy protocol using Levenberg-Marquardt neural network and gated recurrent unit for intrusion detection system*, Trans Emerg. Tel Tech, pp. 1–16, 2020.
3. Sizhe Xing, Zhongya Li, Changle Huang, Guoqiang Li, Aolong Sun, An Yan, Wangwei Shen, Jianyang Shi, Ziwei Li, Chao Shen, Nan Chi, and Junwen Zhang, *End-to-end deep learning for a flexible coherent PON with user-specific constellation optimization*, J. Opt. Commun. Netw., vol. 16, pp. 59–70, 2024.
4. A. M. Mikaeil, W. Hu, T. Ye, and S. B. Hussain, *Performance evaluation of XG-PON based mobile front-haul transport in cloud-RAN architecture*, J. Opt. Commun. Netw., vol. 9, pp. 984–994, 2017.
5. E. N. Abdulla, A. K. Abass, and A. A. Abdulkafi, *Asymmetric 160/80 Gbps TWDM PON utilizing dispersion compensation technique*, J. Opt., 2022.
6. E. N. A. A. H. A. Razi J. Al-Azawi, *Security and network performance analysis of coexistence TWDM–NG-PON2, GPON and 10G-EPON systems based on Hill Cipher*, J. Opt. Commun., pp. 1–11, 2025.
7. E. N. Abdulla, R. A. Hussien, F. F. Rashid, A. A. Abdulkafi, A. K. Abass, and M. A. Saleh, *Design and Performance Analysis of Symmetrical 160 Gbps TWDM-PON Utilizing Bidirectional Configuration*, J. Opt., 2023.
8. E. N. Abdulla, A. K. Abass, and A. A. Abdulkafi, *Asymmetric 160/80 Gb/s TWDM PON with supported transmission method utilizing FBG and DML*, J. Opt. Commun., pp. 8–12, 2022.
9. A. F. Pakpahan and I. S. Hwang, *Peer-to-Peer Federated Learning on Software-Defined Optical Access Network*, IEEE Access, vol. 12, pp. 84435–84451, 2024.
10. B. Lim, S. Zohren, and B. Lim, *Time-series forecasting with deep learning: a survey*, Philos. Trans. R. Soc. A, vol. 379, pp. 2194, 2026.
11. S. Liang, F. Zhao, and W. Jiang, *A learning algorithm for TWDM-PON' DWBA with 5G fronthaul networks*, Opt. Fiber Technol., vol. 53, pp. 102039, 2019.
12. E. A. Fadil, S. R. Tahhan, F. F. Rashid, A. K. Abass, L. A. Salman, E. N. Abdulla, and M. N. A. Alghurabi, *Design and performance analysis of optical communication system utilizing optical chaos*, J. Opt., vol. 53, pp. 2435–2440, 2024.
13. E. Mousa, E. N. Abdulla, and S. A. Adnan, *Enhancing network security based on 10G-EPON with the use of the Hill cipher algorithm*, J. Opt. Commun., pp. 1–13, 2025.
14. A. F. Mushatet, E. A. Fadil, and E. N. Abdulla, *High bit rate secure FSO system utilizing Hill coding*, J. Opt. Commun., vol. 0147, pp. 1–14, 2025.
15. S. S. Radhi, R. A. Hussien, E. N. Abdulla, A. K. Abass, and F. F. Rashid, *Design a secure TWDM-PON via the Hill cipher algorithm*, Opt. Contin., vol. 4, pp. 1051–1064, 2025.
16. E. N. Abdulla, F. F. R. Safa Salam Radhi, M. M. S. Reem A. Hussien, A. K. Abass, and A. E. A. Fadil, *Security improvement for TWDM-PON utilizing blowfish cryptography*, Appl. Opt., vol. 63, 2024.
17. R. Musadaq, S. N. Abdulwahid, Nassren N. Abd Alwahed, and E. N. Abdulla, *Security analysis of an image encryption algorithm based on Blowfish in GPON*, J. Opt. Commun., pp. 32–34, 2025.
18. H. Al-janabi, *Secure of Transmission Systems in the Visible Range of Light with the Power Line Communication Interface*, Int. Arab J. Inf. Technol., vol. 19, pp. 759–764, 2022.
19. H. D. Al-janabi, H. D. Al-janabi, and O. M. Hojayev, *Formulating and Solving Problem of Priority Subchannel Allocation in Downlink WiMAX*, in 2016 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET) (2016), pp. 934–936.
20. F. Musumeci, C. Rottondi, A. Nag, D. Zibar, M. Ruffini, and S. Member, *An Overview on Application of Machine Learning Techniques in Optical Networks*, IEEE Commun. Surv. Tutorials, pp. 1–27, n.d.
21. V. Jha and R. Kumar, *A deep learning based dynamic bandwidth allocation method for XG-PON based mobile fronthaul for CRAN*, Comput. Networks, vol. 245, pp. 110344, 2024.
22. Y. Cheng, Y. Shao, S. Ding, C. Chan, and S. Member, *Deep Reinforcement Learning Based Joint Allocation Scheme in a TWDM-PON-based mMIMO Fronthaul Network*, IEEE Photonics J., vol. PP, pp. 1–12, 2024.
23. D. Systems, *ITU-T*, vol. 2, 2020.
24. K. Abdelli, C. Tropschug, H. Griesser, S. Pachnicke, and C. Kiel, *Fault Monitoring in Passive Optical Networks using Machine Learning Techniques*, in 2023 23rd International Conference on Transparent Optical Networks (ICTON) (2023), pp. 8–11.
25. A. Al-Dulaimi and E. Naji, *Large-Scale Synthetic Dataset for Deep Learning-Based TWDM-PON Parameter Optimization (1 Million Samples), V1, doi: 10.17632/4xjmf8f6y6.1*, Mendeley Data, 2025.
26. N. Mahdavi, A. Dutta, S. H. Tasnim, and S. Mahmud, *Review of machine learning techniques for energy sharing and biomass waste gasification pathways in integrating solar greenhouses into smart energy systems*, Energy AI, vol. 20, 2025.
27. F. F. Ridha, E. N. Abdulla, and A. H. Abdulhadi, *Machine learning based on raw ensemble predictions scheme for TWDM-PON*, J. Opt. Commun., 2025.
28. Garima, Vivekanand Jha, and Rakesh Kumar Singh, *A deep learning based dynamic bandwidth allocation method for XG-PON based mobile fronthaul for CRAN*, Comput. Networks, vol. 245, pp. 110344, 2024.