OPTICAL FIBER FAULTS DETECTION USING DEEP LEARNING MODEL

Thuy Tran Thi Thanh, **Tho Nguyen Van**, **Cao Dung Truong**
Posts and Telecommunication Institute of Technology

***Abstract*:** Backhaul and access communication networks are heavily reliant on effective optical network management to prevent service interruptions and ensure Quality of Service (QoS) compliance. New types of failures in optical networks present challenges that could expose the network to various risks. Traditional detection techniques are increasingly inadequate in addressing these issues. In contrast, Deep Learning (DL) has emerged as a promising approach for fault identification and prevention. This study introduces several key contributions that distinguish it from conventional fault management systems. The first contribution is the development of a Long Short-Term Memory (LSTM) model, integrated with the Mutual Information (MI) technique, to assess their combined effectiveness in detecting normal optical fibers and seven distinct fault types, including fiber cutting, fiber eavesdropping (fiber tapping), dirty connectors, bad splices, bending, reflectors, and PC connectors, achieving an accuracy rate of up to 93.23%. Finally, the proposed model is benchmarked against other deep learning models, such as BiLSTM, CNN, DNN, and RNN, to evaluate critical performance metrics of the AI model.

***Keywords*:** fiber faults detection, deep learning model, LSTM, MI.

I. INTRODUCTION

In contemporary society, there has been a substantial increase in the demand for widespread access to high-speed information across various platforms, encompassing both fixed and wireless services [1],[2]. As a result, optical fiber technology has gained popularity as a vital component of information infrastructure [3]. To meet the demands of rapid data transmission in wireless networks such as 4G, 5G, and beyond [4],[5], optical fiber-based information systems have been widely adopted.

Optical fiber serves as the primary medium for transmitting vast amounts of data across the Internet, mobile backhaul, and core networks. A single fiber link can support thousands of customers and businesses, carrying a mix of personal, corporate, and public information. Consequently, any disruption to the fiber can have significant repercussions, necessitating immediate action.

Contact author: Tran Thi Thanh Thuy,
Email: thuyttt@ptit.edu.vn
Manuscript received: 12/9/2024, revised: 8/11/2024, accepted: 15/11/2024.

Optical fibers are susceptible to several issues, including physical damage like fiber cuts and security breaches such as eavesdropping, both of which can impact network availability and data confidentiality. Identifying and diagnosing these problems manually requires specialized knowledge and time. Optical Time Domain Reflectometer (OTDR) is widely used for monitoring fibers, as it can measure characteristics and detect faults. However, due to noise interference, analyzing OTDR data can be complex and time-consuming with conventional methods. Therefore, an automated system that can accurately and swiftly identify fiber issues would help reduce operational costs and ensure timely restoration of services.

Furthermore, nowadays, Artificial Intelligence (AI) has made significant contributions to the prediction of anomalies in a number of fields, including e-commerce, healthcare, IoT, vehicular networks (VANETs), and power monitoring systems, among others. In optical communication systems, current research has widely applied machine learning (ML) and deep learning (DL) techniques, such as linear regression (LR) methods used for signal amplification [6]. In particular, machine learning approaches have been developed for flaw detection. For example, to anticipate the location of a fiber cut in an underground cable, the Single-Layer Perceptron Neural Network (SLP NN) approach was developed based on the fundamental LR technique [7]. Fiber optics also incorporates other methods, such as Autoencoder (AE) and Bidirectional Gated Recurrent Unit (BiGRU) algorithms, for anomaly detection [8].

In the quest for more accurate fault detection and localization in optical fiber networks, various advanced machine learning techniques have been explored. Among these, Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have been proposed as methods to identify and localize reflecting fiber events caused by mechanical splices and connections [9-11]. The BiLSTM-CNN is a hybrid machine learning (ML) architecture that combines CNNs with a Bidirectional Long Short-Term Memory (BiLSTM) network to identify, localize, and distinguish between reflective, non-reflective, and merged events [12].

However, the aforementioned works utilized limited ML/DL techniques for fault detection, with a narrow range of fault types and complex models that combined two algorithms. Moreover, these models required lengthy training times, partly due to processing all the features of the OTDR data. In contrast, this research proposes and

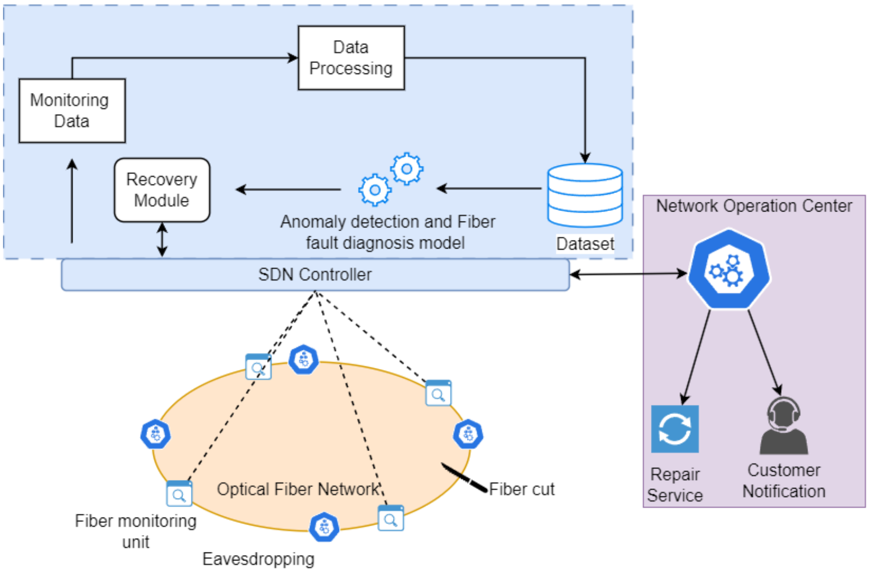


Fig.1 Optical network monitoring model.

evaluates the detection of eight major faults using an LSTM classification model combined with the Mutual Information (MI) feature selection technique, referred to as MI-LSTM. Additionally, this paper compares the performance of the proposed model with other deep learning (DL) models, including BiLSTM, CNN, DNN, and RNN.

II. PROPOSED MODEL

The four main stages of the fiber monitoring model are fiber monitoring and monitoring from the optical fiber network; data processing; anomaly detection and fault diagnosis; and fiber problem mitigation and recovery, as shown in Fig.1. The Optical Time Domain Reflectometer (OTDR) plays a pivotal role in regularly examining optical fibers within the network. By leveraging light backscattering, the OTDR assesses parameters such as attenuation, fault locations, and connection integrity. Data is collected by transmitting light pulses through the fibers and capturing reflected signals from disruptions like cuts, bends, or poor connections. The process involves replicating real-world conditions, gathering signals, labeling them according to fault types or normal operation, and applying noise reduction techniques. The resulting OTDR traces, or monitoring data, are sent to the software-defined networking (SDN) controller managing the optical network. This data is segmented into fixed lengths and normalized before being fed into a deep learning model for identifying fiber anomalies or diagnosing defects. Once a fault is detected, predefined recovery protocols are triggered, and the SDN controller informs the network operation center. The center then alerts the maintenance team and relevant customers about the detected issue. This paper emphasizes the stages of data processing, anomaly detection, and fault diagnosis using a deep learning-based approach.

The main components of our proposed model include Minmax scaler normalization method, Mutual Information feature selection technique and LSTM model, illustrated in Fig. 2. In this research, selecting key features, by Mutual Information technique, from the OTDR dataset plays a crucial role in our research. This approach helps identify factors that significantly influence prediction outcomes while avoiding the use of unnecessary features, which can waste training time and even lead to inaccurate results. Additionally, we focus on developing a deep learning, LSTM model, based on OTDR monitoring data from optical networks to analyze fiber faults such as fiber cuts and optical eavesdropping attacks, which are considered major incidents. Since these fault patterns and other anomalies like bad splices, dirty connectors, bending, and back reflection have similar characteristics, especially under very low signal-to-noise ratio (SNR) conditions, we have integrated them into the deep learning model's training phase for fault diagnosis, ensuring accurate fault identification and reducing false alarms. Furthermore, we also compare the LSTM model with other popular deep learning models such as BiLSTM, CNN, DNN, and RNN to assess the effectiveness of each method.

A. Dataset and Data preprocessing

The generated OTDR traces [13] have 8 labels (0-normal, 1-fiber cutting, 2-fiber eavesdropping (fiber tapping), 3-dirty connector, 4-bad splice, 5-bending, 6-reflector, 7-PC connector). They are segmented into sequences of length 30 together with SNR resulting in 31 features. This approach ensures that each segment contains relevant information for anomaly detection, while the inclusion of the SNR feature helps assess the quality of the signal, improving the model's ability to detect faults accurately. The dataset is normalized using the MinMax scaler technique. MinMax scaler is a commonly used data normalization technique in machine learning, which transforms data values to the range [0, 1]. This technique maintains the relative scale between values, helping machine learning and deep learning algorithms perform

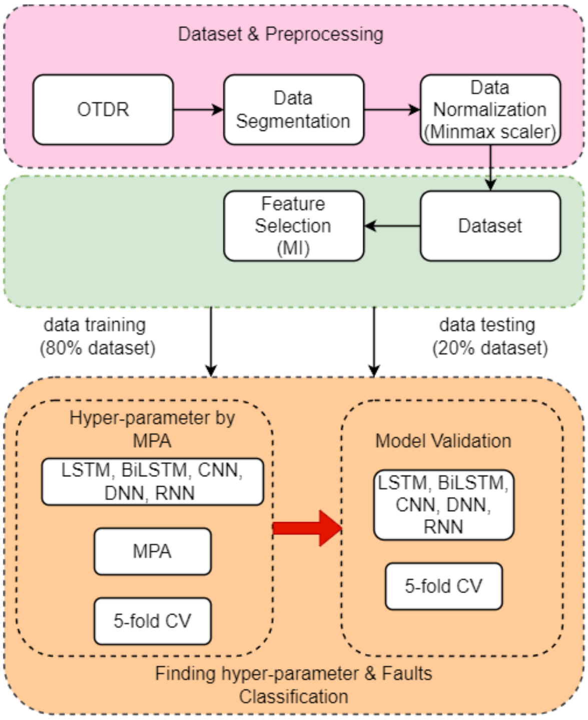


Fig.2. The proposed model and operation processes.

more effectively, especially those sensitive to value ranges such as neural networks. By normalizing data to the [0, 1] range, the impact of outliers is reduced, which enhances the accuracy of the models.

B. Feature selection

Once the data preprocessing is complete, the dataset is fed into a feature selection mechanism to identify and filter out the most crucial features. This feature selection process helps eliminate unnecessary attributes, reduce model complexity, and enhance the performance and accuracy of machine learning algorithms by focusing on the factors that most significantly impact prediction outcomes.

The feature selection technique employed here is Mutual Information (MI) [14]. MI is a feature selection technique used to measure the degree of dependency between two variables. It assesses the amount of shared information between features and the target variable, helping to identify the most crucial features in a dataset. Unlike methods that rely on linear relationships, MI can detect nonlinear dependencies, making it effective for revealing significant features that may not be linearly related to the target variable. By filtering out less relevant features, MI reduces model complexity and improves performance, focusing on factors with the strongest influence on prediction outcomes. Additionally, MI does not require specific distribution assumptions, making it versatile and effective in various scenarios.

C. LSTM-based optical fiber fault detection

After selecting the important features or using all features from the dataset, the data is divided into two parts: 80% is used for training, while the remaining 20% is reserved for testing the performance of the trained model.

In this study, at first, the LSTM model [15] was constructed with four layers: one input layer, two hidden layers, and one output layer. The selection of 128 and 64 nodes for the two hidden layers was based on a balance between computational efficiency and model performance [16]. Specifically, the initial layer with 128 nodes captures the complexity of the input features, while the subsequent layer with 64 nodes reduces dimensionality and prevents overfitting without significant loss of information. This architecture was determined through multiple trials and was found to provide optimal accuracy for the OTDR dataset. The dropout rate of 5% was selected after testing various values ranging from 2% to 20%. This rate effectively minimized overfitting while preserving the model's ability to learn from the data. The "ReLU" activation function was applied in the hidden layers to introduce non-linearity, enhancing the model's ability to learn complex patterns. The output layer, containing eight nodes corresponding to the eight fault classes, uses the "softmax" function for classification. To ensure efficient training, the model employs the "Adam" optimizer with a learning rate of 0.001, which is dynamically reduced by a factor of 0.2 if the validation error stagnates or worsens over three consecutive epochs. Early stopping is also utilized to avoid overfitting, ending training when no improvement is observed in validation error after a specified number of epochs. These architectural choices

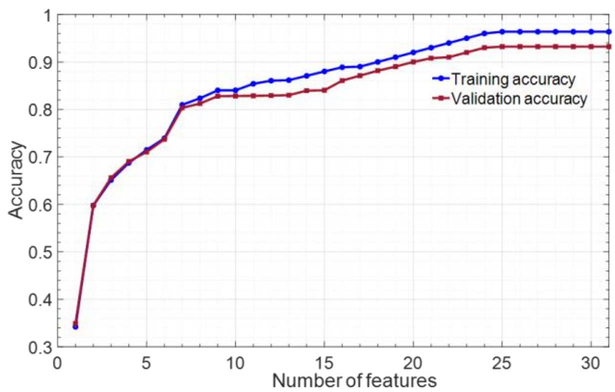


Fig.3. Accuracy of MI-LSTM model dependence on the number of features is swept in the range from 1 to 31

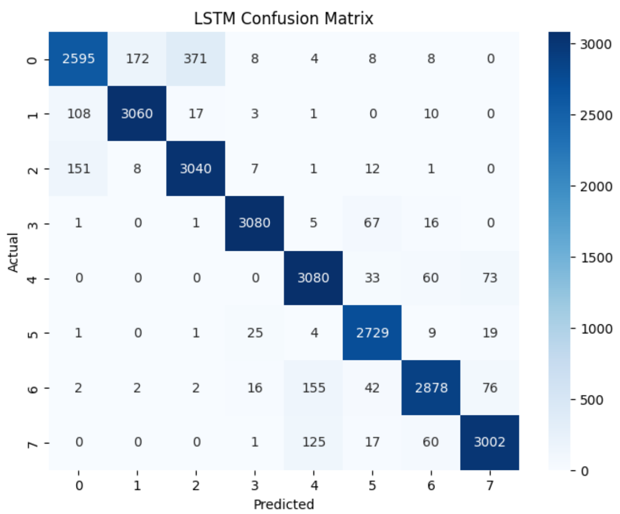


Fig.4. Confusion matrix of MI-LSTM model.

were thoroughly tested and demonstrated superior performance across key evaluation metrics.

The hyperparameter optimization process involves using the Marine Predators Algorithm (MPA) [17] in conjunction with 5-fold Cross-Validation (CV5) [18]. MPA, inspired by the hunting strategies of marine predators, systematically explores the hyperparameter space to identify the optimal values for key parameters, such as the learning rate, dropout rate, and layer sizes. This approach ensures that the model achieves a balance between performance and complexity. Additionally, CV5 divides the dataset into five subsets, using four subsets for training and one for testing in each iteration. This process evaluates the model's performance across multiple configurations and prevents overfitting by ensuring the selected hyperparameters generalize well to unseen data. These methods together ensure that the optimized model is robust and performs reliably across various scenarios. Finally, the proposed model is evaluated using the test dataset, with accuracy, recall, and F1-score serving as performance metrics, which are calculated as follows:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F1\text{-}score = \frac{TP}{TP+\frac{1}{2}(FP+FN)}$$

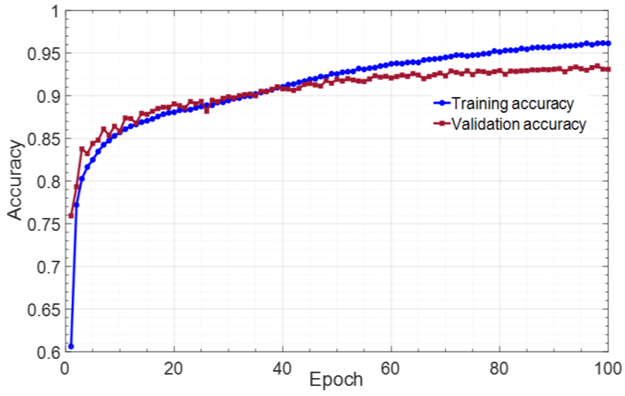
where *TP*, *TN*, *FN*, and *FP* stand for true positive, true negative, false negative, and false positive in confusion matrix, respectively.

Similarly, we conducted experiments on four models (BiLSTM, CNN, DNN, and RNN) to compare them with the proposed LSTM model.

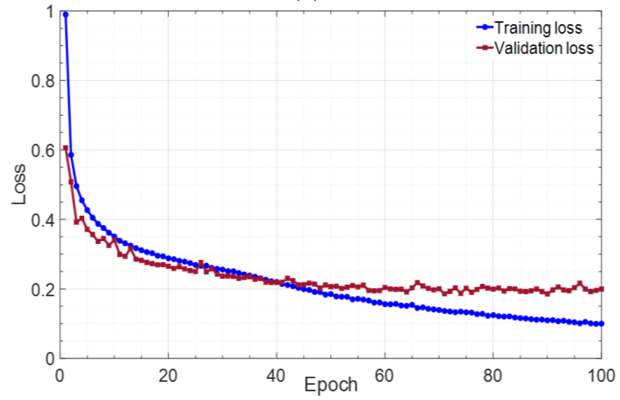
III. RESULTS AND DISCUSSION

Initially, we employ the Mutual Information (MI) feature selection technique in conjunction with the LSTM model to reduce the dimensionality of the dataset, selecting a smaller, practical set of features. This approach improves processing speed without significantly compromising fault detection accuracy. Figure 3 illustrates the relationship between the number of selected features and the accuracy of the MI-LSTM model. Notably, with only 20 features, the model achieves an accuracy of 90%. Furthermore, by using just 25 features, the proposed LSTM model reaches the same accuracy-approximately 97% on the training set and 94% on the test set-as when utilizing all 31 OTDR features. This demonstrates the model's efficiency in maintaining high accuracy while using fewer features, thereby reducing both computational load and training time.

Next, the confusion matrix is utilized as a key method to evaluate the model's performance. A high number of correct predictions (the cells along the main diagonal) indicates that the model is performing well, while the off-diagonal cells represent the number of incorrect predictions. Fig. 4 presents the confusion matrix for the test



(a)



(b)

Fig.5. Training graph the proposed model with: (a) accuracy; and (b) loss

dataset of the proposed MI-LSTM model with 25 features. This result demonstrates that the proposed model effectively classifies all 8 labels (0-normal, 1-fiber cutting, 2-fiber eavesdropping (fiber tapping), 3-dirty connector, 4-bad splice, 5-bending, 6-reflector, 7-PC connector), including those that are challenging to distinguish, such as bad splice and dirty splice.

The training process, illustrated in Fig. 5, indicates that the model's performance is quite satisfactory. Both training and validation graphs reveal that the model is learning efficiently, with high accuracy and low loss after multiple epochs. The close alignment of accuracy and loss metrics between the training and test sets further suggests that the model exhibits strong generalization ability and is not overfitting.

Finally, Tab. 1 compares the performance of the LSTM model against other deep learning models such as BiLSTM, CNN, DNN, and RNN. Based on evaluation metrics such as accuracy, recall, and F1-score, the proposed model demonstrates superior performance, achieving over 93.23% accuracy, outperforming the other models. Although the BiLSTM model shows comparable performance, its more complex architecture results in a longer training time.

Tab.1. Performance comparison between deep learning models

Model / Metrics	LSTM	BiLSTM	CNN	DNN	RNN
Accuracy on training set	0.9657	0.9745	0.9429	0.8962	0.2285
Accuracy on validation set	0.9323	0.9369	0.8945	0.9536	0.3112
Recall	0.9329	0.9372	0.8945	0.9534	0.3047
F1-score	0.9323	0.9371	0.8943	0.9528	0.2296

IV. CONCLUSION

The paper introduces a model for automatic detection of optical fiber anomalies or faults in optical networks, LSTM combined with MI feature selection technique. By building LSTM models in Python programming language and testing them on OTDR dataset, the study proves that the model can detect optical fiber faults well with an accuracy of up to more than 96%. The results show that with 25 features of the dataset, the model has the same performance as the model using all 31 features. In addition, the study also compares the proposed model with 4 other deep learning models (BiLSTM, CNN, DNN and RNN), thereby emphasizing the feasibility of the proposed model.

ACKNOWLEDGMENT

This research is sponsored by the Posts and Telecommunications Institute of Technology under project code 12–2024–HV–KTĐT1.

REFERENCES

[1] H. A. Chan, "Requirements of interworking WirelessLAN and PLMN wireless data network systems," IEEE AFRICON Conference, vol. 1. pp. 251–255, 2004.

[2] Z. Pi, J. Choi, and R. Heath, "Millimeter-wave gigabit broadband evolution toward 5G: Fixed access and backhaul," IEEE Communications Magazine, vol. 54, no. 4. pp. 138–144, 2016.

[3] A. Banerjee et al., "Wavelength-division-multiplexed passive optical network (WDM-PON) technologies for broadband access: A review," J. Opt. Netw., vol. 4, no. 11, pp. 737–758, 2005.

[4] G. Gampala and C. J. Reddy, "Massive MIMO - Beyond 4G and a basis for 5G," in 2018 International Applied Computational Electromagnetics Society Symposium in Denver, ACES-Denver 2018.

[5] D. Novak et al., "Radio-over-fiber technologies for emerging wireless systems," IEEE J. Quantum Electron., vol. 52, no. 1, 2016.

[6] Huang, Yishen, et al. "A machine learning approach for dynamic optical channel add/drop strategies that minimize EDFA power excursions." ECOC 2016; 42nd European Conference on Optical Communication. VDE, 2016.

[7] O. Nyarko Boateng, A. F. Adekoya, and B. A. Weyori, "Predicting the actual location of faults in underground optical networks using linear regression," Eng. Rep., vol. 3, no. 3, pp. eng212304, Mar. 2021.

[8] K. Abdelli, J. Y. Cho, F. Azendorf, H. Griesser, C. Tropschug and S. Pachnicke, "Machine-learning-based anomaly detection in optical fiber monitoring," J. Opt. Commun. Netw., vol. 14, no. 5, pp. 365–375, May 2022.

[9] K. bdelli et al., "Reflective fiber fault detection and characterization using long short-term memory," IEEE/OSA J. Opt. Commun. Netw., vol. 13, no. 10, October 2021.

[10] K. bdelli et al., "Reflective Event Detection and Characterization in Fiber Optical Links Given Noisy OTDR Signals," hotonic Networks; 22th ITG Symposium, 2021.

[11] W. Zhang et al., "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load", Mechanical Systems and Signal Processing, Volume 100, 2018.

[12] K. B. Delli et al., "A BiLSTM-CNN based Multitask Learning pproach for Fiber Fault iagnosis," OCF 2021.

[13] Khouloud Abdelli , Florian Azendorf, Carsten Tropschug, Helmut Griesser , Stephan Pachnicke , Joo Choo , October 17, 2022, "Dataset for Optical fiber faults ", IEEE Dataport.

[14] Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. Physical Review E—Statistical, Nonlinear, and Soft Matter Physics, 69(6), 066138.

[15] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. Neural computation, 31(7), 1235-1270.

[16] Linming, Z. H. A. O., and W. E. I. Dehua. "The method to choose the optimal number of hide nodes of artificail nueral networks." Journal of North China Institute of Water Conservancy and Hydroelectric Power 20.4 (1999): 44-48.

[17] Faramarzi, A., Heidarinejad, M., Mirjalili, S., & Gandomi, A. H. (2020). Marine Predators Algorithm: A nature-inspired metaheuristic. Expert systems with applications, 152, 113377.

[18] Fushiki, T. (2011). Estimation of prediction error by using K-fold cross-validation. Statistics and Computing, 21, 137-146.

PHÁT HIỆN LỖI SỢI QUANG TRONG GIÁM SÁT SỢI QUANG SỬ DỤNG MÔ HÌNH HỌC SÂU

Tóm tắt: Mạng truyền thông truy cập và truyền ngược phụ thuộc rất nhiều vào quản lý mạng quang hiệu quả để ngăn ngừa gián đoạn dịch vụ và đảm bảo tuân thủ Chất lượng dịch vụ (QoS). Các loại lỗi mới trong mạng quang đặt ra những thách thức có thể khiến mạng phải đối mặt với nhiều rủi ro khác nhau. Các kỹ thuật phát hiện truyền thống ngày càng không đủ để giải quyết những vấn đề này. Ngược lại, Học sâu (DL) đã nổi lên như một phương pháp tiếp cận đầy hứa hẹn để xác định và ngăn ngừa lỗi. Nghiên cứu này giới thiệu một số đóng góp quan trọng giúp phân biệt nó với các hệ thống quản lý lỗi thông thường. Đóng góp đầu tiên là phát triển mô hình Bộ nhớ dài hạn ngắn (LSTM), tích hợp với kỹ thuật Thông tin tương hỗ (MI), để đánh giá hiệu quả kết hợp của chúng trong việc phát hiện các sợi quang thông thường và bảy loại lỗi riêng biệt, bao gồm cắt sợi, nghe trộm sợi (khai thác sợi), đầu nối bẩn, mối nối xấu, uốn cong, phản xạ và đầu nối PC, đạt tỷ lệ chính xác lên tới 93,23%. Cuối cùng, mô hình được đề xuất được so sánh với các mô hình học sâu khác, chẳng hạn như BiLSTM, CNN, DNN và RNN, để đánh giá các số liệu hiệu suất quan trọng của mô hình AI.

Từ khóa: phát hiện lỗi sợi quang, mô hình học sâu, LSTM, MI.

Thuy Tran Thi Thanh graduated from the Posts and Telecommunications Institute of Technology with a Bachelor's degree in 2021 and completed the Master's program in 2023. Currently, she is a lecturer in the Department of Electronics Engineering at the Posts and Telecommunications Institute of Technology. Her primary research interests include integrated optoelectronic circuits, high-speed optical information systems, IoT, and the application of artificial intelligence in IoT.

Tho Van Nguyen is a lecturer at the Vietnam-UK Institute for Research and Training, University of Danang and a researcher at the Institute of Advanced Science and Technology. He received the M.Eng degrees in Telecommunications Engineering from Ha Noi university of Science and Technology in 2013, respectively. Currently, he is a graduate student at the Posts and Telecommunications Institute. His current research interests include communication systems, image/video processing, machine learning, and deep learning.

Cao Dung Truong was born in Thanh Hoa, Vietnam, in 1980. He received the Engineer, Master of Science, and Ph.D. degrees from Hanoi University of Science and Technology, Hanoi, in 2003, 2006, and 2015, respectively. He is currently a Lecturer in the Department of Electronic Engineering1, Posts and Telecommunications Institute of Technology, Hanoi, Vietnam. His research interests include photonic integrated circuits, plasmonics, high-speed optical communication systems, smart IoT networks, applications of deep learning models for photonics design, and AI in photonics.