# An Efficient Failure Detection Model based on Semi-Supervised Algorithm for Optical Networks with Limited Labeled Data

Zhiming Sun<sup>(1)</sup>, Chunyu Zhang<sup>\*,(1)</sup>, Min Zhang<sup>(1)</sup>, Bing Ye<sup>(2)</sup>, Cheng Xing<sup>(1)</sup>, Danshi Wang<sup>(1)</sup>

<sup>(1)</sup> State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing, China, \*chunyuzhang@bupt.edu.cn
<sup>(2)</sup> State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen, China

**Abstract** We propose a semi-supervised OFE-VIME model for failure detection in optical networks with limited labeled data, which achieves detection F1 score and accuracy of 0.951 and 0.949, false negative and false positive rates of 0.018 and 0.085 at a labeled data ratio of 3.85%. ©2023 The Author(s)

## Introduction

Failure detection is a key technology to ensure optical network reliability and quality of service (QoS) [1]. In the early stages, optical network detection (ON-FD) technology failure set thresholds based on manual experience, which was difficult to adapt to large-scale and dynamic optical networks [2]. With the development of artificial intelligence technology, data-driven supervised machine learning has become a popular research topic in the field of ON-FD due to its ability to effectively mine potential patterns within data [3-5]. Nevertheless, it is difficult for acquiring sufficient labeled data owing to the high cost of manual data labeling, the long period needed for labeling data, and the low failure rate, and training models with limited labeled data usually affects the detection performance of the model [6]. Conversely, telemetry has significantly increased the reporting speed of optical network nodes, reaching the second-level reporting, and has generated a substantial quantity of unlabeled data [7]. Regrettably, supervised learning algorithms cannot leverage unlabeled data. Therefore, in the case of given limited labeled data, how to effectively use a large number of unlabeled real data to improve failure detection performance is the key problem of ON-FD.

best global decision boundary from a large number of labeled data and unlabeled data, which provides a solution for the utilization of limited labeled data and large amount of unlabeled data in optical network [8]. Value Imputation and Mask Estimation (VIME) based on a semi-supervised learning framework, has been proposed as a novel inference method for tabular data [9]. The algorithm can support tabular data, and can fully utilize the limited labeled data and large amount of unlabeled data in optical network for model training.

In this paper, an OFE-VIME semi-supervised learning model is proposed for failure detection under limited labeled data. Open Automated Feature Engineering (OFE) algorithm is used for automatic expansion of features [10], and VIME model is used to improve the performance of failure detection under limited labeled data by using a large amount of unlabeled data. Typical evaluation metrics are used to evaluate the detection performance of the proposed model, and the effectiveness of failure detection under limited data labels is verified on real optical network dataset. Meanwhile, to explore the learning ability of OFE-VIME model to unlabeled data, the influence of the amount of unlabeled data on the performance of the detection model is studied. Additionally, we leverage t-SNE



Fig. 1: ON-FD scenario based on data-driven AI algorithm: (a) Data collection; (b) Data annotation; (c) Supervised model; (d) Unsupervised model; (e) Semi-supervised OFE-VIME model.

visualization techniques to explore the detection mechanism of proposed model [11].

### **ON-FD Background**

The ON-FD based on a data-driven AI is shown in Fig. 1. The first stage is data pre-processing. As shown in Fig. 1(a), a large number of unlabeled data are obtained based on telemetry technology. Fig. 1(b) shows the process of identifying failure labels by manual method. It is usually time-consuming and expensive to mark the data through regular inspection and manual experience. Furthermore, due to the low failure rate of optical networks equipment, a limited amount of labeled data is available, which makes it both scarce and costly. Limited labeled data usually limit the detection performance of supervised learning algorithms (Fig. 1(c)), because supervised algorithms usually need a certain amount of labeled data for training. The unsupervised algorithm does not need labels in the learning process (Fig. 1(d)), but unsupervised algorithms usually need prior assumptions to understand the data distribution, which will affect the detection performance of unsupervised algorithms if the data distribution in optical networks is not clear.

Therefore, aiming at the limited labeled data and a large number of unlabeled data scenes obtained by optical network equipment, this paper proposes a semi-supervised learning method for failure detection under limited data (Fig. 1(e)). This method can make full use of limited labeled data to extract prior knowledge, and learn more comprehensive decision boundaries based on limited labeled data and unlabeled data, thus improving the detection performance under limited labeled data.

## **OFE-VIME Semi-Supervise ON-FD Model**

The OFE-VIME semi-supervised learning model is illustrated in Fig. 2, and the details of each component are elaborated below.

<u>OFE.</u> As shown in Fig. 2(a), this method uses an expansion-reduction framework with feature enhancement to measure the effectiveness of candidate features. Additionally, a two-stage filtering method is employed to reduce computational and memory burdens while ensuring precise and rapid feature selection.

<u>Pretext Generator.</u> The module uses a Mask generator to randomly generate multiple sets of mask data, which is used to increase data diversity and improve model robustness and generalization performance. These sets of mask data are then combined with the unlabeled data features to form multiple sets of Corrupted features. The process is illustrated in Fig. 2(b).

<u>Encoder.</u> As illustrated in Fig. 2(c), this module utilizes an encoder to extract deep representations from the labeled data features. Additionally, it also extracts deep representations of the corrupted features generated by the pretext generator module.

Predictor. As shown in Fig. 2(d), this part is the core of semi-supervised learning model. In the training stage, the encoder extracts features from the labeled data and the unlabeled data, obtaining supervised loss and consistency loss respectively, and the total loss is shown in Eq. (1), where  $\mathcal{L}_s$  and  $\mathcal{L}_u$  are supervised loss and consistency loss respectively. For the labeled data, the encoder extracts the features, uses the extracted features for predictor evaluation, compares the real labels with the predicted results, and calculates the supervised loss as shown in Eq. (2). For unlabeled data, the encoder extracts features from different data, and the extracted features are used for predictor evaluation, and different output results are compared to calculate the consistency loss of different output results, as shown in Eq. (3).

$$\mathcal{L}_{final} = \mathcal{L}_s + \beta \cdot \mathcal{L}_u \tag{1}$$

$$\mathcal{L}_s = \mathbb{E}_{(x,y) \sim p_{XY}}[l_s(y, f_e(x))]$$
(2)

$$\mathcal{L}_{u} = \mathbb{E}_{x \sim p_{X}, m \sim p_{m}, \tilde{x} \sim g_{m}(x, m)} [(f_{e}(\tilde{x}) - f_{e}(x))^{2}]$$
(3)

## **Experimental Setup and Results**

The experimental data comes from the optical backbone networks data in the real network, and its total data is 38082. The number of samples used for training is 26,000, and the number of



Fig. 1: The framework of OFE-VIME semi-supervised learning model: (a) OFE; (b) Pretext generator; (c) Encoder; (d) Predictor.



Fig. 3: (a) Loss function for the model; (b) Performance metrics for the model; (c) F1 and accuracy of various types of models; (d) Performance of model with varying amounts of unlabeled data; (e) Principle of model through t-SNE visualizations.

samples used for testing is 12,082. Among the training samples, there are 1,000 labeled samples and 25,000 unlabeled samples, and the ratio of labeled data is 3.85%.

During the training process of the model, the loss function is shown in Fig. 3(a), and the curves of the total loss function and the two sub-loss functions are gradually parallel to the X-axis, indicating that the model converges well. Fig. 3(b) shows the accuracy, F1, false positive rate (FPR), false negative rate (FNR), and confusion matrix of the proposed OFE-VIME model on the test data. Compared with VIME model, the detection metrics of the proposed OFE-VIME model have been improved.

Moreover, to evaluate the performance of OFE-VIME model, the detection performance of typical unsupervised and supervised models with limited labeled data is compared and analyzed. F1 score and accuracy are used to evaluate the model performance, and the results are plotted on the scatter plot, as shown in Fig. 3(c). The closer the points in the scatter plot are to the upper right corner, the better the detection performance of the model. From Fig. 3(c), it can be concluded that compared with the typical supervised and unsupervised algorithms, the proposed OFE-VIME model has the best accuracy and F1 score, which shows that the proposed semi-supervised OFE-VIME model can achieve better detection performance in limited labeled data scenarios.

To explore the influence of unlabeled data on the detection performance of semi-supervised OFE-VIME model, we have carried out a series of experiments to increase the unlabeled sample data from 5,000 to 25,000, and its detection performance index are shown in Fig. 3(d). As can be seen from Fig. 3(d), with the increase of the number of unlabeled samples, the F1 score and accuracy of detection gradually increase. When the number of unlabeled samples increases to 25,000, the F1 score and accuracy of detection are 0.951 and 0.949, and the false positive rate and false negative rate are 0.085 and 0.018, respectively. Therefore, it can be concluded that the proposed OFE-VIME model improves the detection performance of the model under limited labeled data by using unlabeled samples.

Then, the principle of semi-supervised OFE-VIME model for failure detection is explored. After t-SNE visualization, the original data (left) and the deep feature representation based on OFE-VIME output (right) are obtained, as shown in Fig. 3(e). It can be concluded from Fig. 3(e) that the normal sample and the failure sample in the original data coincide, while the failure sample and the normal sample are separated in the deep representation obtained by the OFE-VIME model, which is a positive signal of failure detection. Therefore, it can be concluded that the OFE-VIME-based model is effective in enhancing the distinguishing ability of learning features.

#### Conclusions

This paper proposes a semi-supervised OFE-VIME model that is used for failure detection with limited labeled data. In the training stage, the scheme uses labeled and unlabeled data for training where the ratio of labeled data is 3.85%, and the detection F1 score and accuracy on the test data reach 0.951 and 0.949, and the FPR and FNR are 0.085 and 0.018 respectively. Moreover, the influence of the amount of unlabeled data on the detection performance of OFE-VIME model and the failure detection principle of OFE-VIME model are analysed.

#### Acknowledgements

This work was supported in part by National Natural Science Foundation of China (No. 61975020, 62171053).

#### References

- D. Wang, C. Zhang, W. Chen, H. Yang, M. Zhang, and A. P. T. Lau, "A review of machine learning-based failure management in optical networks," *Science China Information Sciences*, vol. 65, no. 11, pp. 1-19,2022. DOI: <u>10.1007/s11432-022-3557-9</u>
- [2] M. Zhang, C. You, H. Jiang, and Z. Zhu, "Dynamic and adaptive bandwidth defragmentation in spectrum-sliced elastic optical networks with time-varying traffic," *Journal* of Lightwave Technology, vol. 32, no. 5, pp. 1014– 1023,2014. DOI: <u>10.1109/jlt.2013.2296781</u>
- [3] C. Zhang, D. Wang, L. Wang, L. Guan, H. Yang, Z. Zhang, X. Chen, and M. Zhang, "Cause-aware failure detection using an interpretable XGBoost for optical networks," *Optics Express*, vol. 29, no. 20, pp. 31974-31992, 2021. DOI: <u>10.1364/OE.436293</u>
- [4] S. Varughese, D. Lippiatt, T. Richter, S. Tibuleac and S. E. Ralph, "Low complexity soft failure detection and identification in optical links using adaptive filter coefficients," *Optical Fiber Communication Conference* (*OFC*), San Diego, USA, 2020. DOI: <u>10.1364/ofc.2020.m2j.4</u>
- [5] L. Shu, Z. Yu, Z. Wan, J. Zhang, S. Hu, and K. Xu, "Dualstage soft failure detection and identification for lowmargin elastic optical network by exploiting digital spectrum information," *Journal of Lightwave Technology*, vol. 38, no. 9, pp. 2669-2679, 2019. DOI: <u>10.1109/jlt.2019.2947562</u>
- [6] X. Chen, B. Li, R. Proietti, Z. Zhu, and S. B. Yoo, "Selftaught anomaly detection with hybrid unsupervised/supervised machine learning in optical networks," *Journal of Lightwave Technology*, vol. 37, no. 7, pp. 1742-1749, 2019. DOI: <u>10.1109/jlt.2019.2902487</u>
- [7] A. Sadasivarao, S. Jain, S. Syed, K. Pithewan, P. Kantak, B. Lu, L. Paraschis, "High performance streaming telemetry in optical transport networks," *Optical Fiber Communication Conference (OFC)*, San Diego, USA, 2020. DOI: <u>10.1364/OFC.2018.Tu3D.3</u>
- [8] J.E. van Engelen, H.H. Hoos, "A survey on semisupervised learning," *Machine learning*, vol. 109, no. 2, pp. 373-440. DOI: <u>10.1007/s10994-019-05855-6</u>
- [9] J. Yoon, Y. Zhang, J. Jordon, M. van der Schaar, "VIME: Extending the success of self-and semi-supervised learning to tabular domain," *Advances in Neural Information Processing Systems (NeurIPS)*, New Orleans, USA, 2020. DOI: <u>neurips.cc/paper/2020/hash/7d9766</u>
- [10] T. Zhang, Z. Zhang, Z. Fan, H. Luo, F. Liu, W. Cao, J. Li, "OpenFE: Automated Feature Generation beyond Expertlevel Performance," *arXiv preprint*. DOI: <u>10.48550/arXiv.2211.12507</u>
- [11]L. Van der Maaten, G. Hinton, "Visualizing data using t-SNE," *Journal of machine learning research*, vol. 9, no. 11, pp. 2579-2605, 2008.
  DOI: <u>10.5555/2987189.2987463</u>