# Interpretable Optical Fiber Fault Detection Using FFT and Kolmogorov-Arnold Networks

Jiafang Fu [1, a)], Weikuan Qiu[2, b)] and JiaSheng Wang [3, c)]

[1]*School of Physics and Optoelectronic Engineering, Ocean University of China, Qingdao, China*
[2]*School of Integrated Circuits, Shandong University, Jinan, China*
[3]*Communication Engineering, Xidian University Xi'an, China*

[a)] Corresponding author: 23070001018@stu.ouc.edu.cn
[b)] 202300400175@mail.sdu.edu.cn
[c)] 23012100065@stu.xidian.edu.cn

**Abstract.** In the era of the Internet of Everything, optical network fault detection plays a pivotal role in maintaining high service quality. Traditional machine learning methods for optical performance monitoring require extensive data quality and quantity, often resulting in limited interpretability and low efficiency. This study introduces an innovative fault detection and migration training approach based on a Fast Fourier Transform and Kolmogorov-Arnold Network (FFT-KAN) model. The proposed method employs the FFT algorithm for effective feature extraction from 16QAM constellation diagrams, capturing frequency domain characteristics and image-based features such as edge density, contrast, and correlation. These features are subsequently processed by the KAN model, which combines the advantages of multilayer perceptrons and splines to simplify network structure and enhance interpretability while maintaining computational efficiency. Experimental results demonstrate that the FFT-KAN model achieves perfect performance in binary classification tasks, with accuracy, precision, recall, and F1 scores reaching 100%. Moreover, in five-class fault detection, the model attains an overall accuracy of 99%, outperforming conventional techniques such as support vector machines and isolation forests. The integration of FFT and KAN offers a robust solution for fault mode detection in optical networks. Future work will explore the incorporation of big data analytics, cloud computing, and dynamic model optimization to further enhance the applicability of the proposed approach in intelligent maintenance and network management. These results confirm the FFT-KAN model's potential as a tool for diagnostics.

## INTRODUCTION

With the advent of the Internet of Everything and the rapid deployment of 5G networks, the complexity and data volume within optical networks have significantly increased. These developments have brought about new challenges in maintaining high service quality, where optical network fault detection plays a pivotal role in ensuring reliable data transmission [1]. As the backbone infrastructure for modern communication, optical networks require robust fault detection and diagnostic techniques to promptly identify and address performance degradations that could lead to service disruptions.

Traditional fault detection approaches in optical performance monitoring (OPM) have predominantly relied on classical machine learning methods such as support vector machines (SVM), ridge regression (RR), and artificial neural networks (ANN) [2–4]. While these methods have contributed to early advancements in fault identification and localization, they typically demand large volumes of high-quality data and often exhibit limited interpretability and efficiency. The inherent complexity of optical networks, combined with the dynamic behavior of network components, necessitates more sophisticated approaches that can extract critical features from noisy signals while simplifying the model structure for enhanced interpretability.

In recent years, advances in machine learning have opened new avenues for addressing these challenges. In 2024, Liu Ziming and Max Tegmark introduced the Kolmogorov-Arnold Network (KAN), a computational model derived

from the Kolmogorov-Arnold theorem, which promises improved interpretability by reconstructing complex multivariable functions through finite layers of univariate computations [5]. The KAN model, when combined with efficient feature extraction methods, offers a promising solution for optical network fault detection.

This paper proposes an innovative fault detection and migration training approach based on the integration of the Fast Fourier Transform (FFT) and the KAN model, herein referred to as the FFT-KAN model. The FFT algorithm is utilized for the extraction of frequency-domain features from 16QAM constellation diagrams, capturing both signal frequency characteristics and image-based properties such as edge density, contrast, and correlation. By processing these features through the KAN model—which synergistically combines the strengths of multilayer perceptrons and splines—the proposed approach simplifies network structure while achieving high computational efficiency and interpretability.

The following sections summarize previous work in optical performance monitoring using FFT, KAN, and traditional machine learning techniques, and then provide a detailed comparison of OPM performance using KAN versus multilayer perceptrons (MLPs). This work not only highlights the superior performance of the FFT-KAN model in both binary and multi-class fault detection tasks but also discusses its potential for future integration with advanced technologies such as big data analytics, cloud computing, and fog computing to further enhance intelligent maintenance and network management [6-7].

# RESEARCH METHODS

## FFT algorithm

Fast Fourier Transform (FFT) is an efficient Discrete Fourier Transform (DFT) algorithm widely used in signal processing, image processing, communication systems, etc. The core idea of the FFT algorithm is to use the symmetry and periodicity of the Discrete Fourier Transform (DFT) to decompose a long DFT sequence into multiple shorter DFT sequences, thereby reducing the computational complexity and improving the efficiency of machine learning.

The FFT algorithm is used for feature extraction. It transforms the signal from the time domain to the frequency domain, which allows for the extraction of the signal's frequency characteristics. In this experiment, the FFT algorithm is employed for image feature extraction, and the frequency characteristics are utilized to describe the image's edges, shapes, and other features.

## Kolmogorov-Arnold Network

Derived from the Kolmogorov-Arnold theorem—a mathematical contribution by Kolmogorov and Arnold in the 1950s—KANs exemplify a neural network design that operationalizes the theorem's key insight: complex multivariable functions can be reconstructed through finite layers of single-variable computations. Although the mathematical explanation of KANs is elegant, they are nothing more than a combination of splines and multilayer perceptrons (MLPs), taking advantage of their respective strengths and avoiding their respective weaknesses [5]. To accurately learn functions, the model should not only learn the combinatorial structure (external degrees of freedom) but also approximate the single-variable functions (internal degrees of freedom) well [5]. KANs are such models because they have MLPs on the outside and splines on the inside. Therefore, KAN can not only learn features (due to their external similarity to MLP) but also optimize these learned features very accurately (due to their internal similarity to splines) [5]. Based on this theory, KAN significantly simplifies the network structure. The model enhances its interpretability and computational efficiency by breaking down intricate high-dimensional functions into combinations of simpler univariate functions.

# EXPERIMENTAL PROCEDURE

## Feature Extraction

The laser within the optical transponder might malfunction or be improperly configured, leading to signal power that exceeds expectations and causing the channel center frequency to drift. These failures can cause additional interference to adjacent channels, affecting the overall performance of the optical network. This paper proposes using constellation diagrams to detect anomalies in optical transponder signals, focusing on the 16QAM constellation at the

channel receiver under test (CUT). Dataset is the constellation diagram of the 16QAM receiver at the channel under test (CUT), including both normal operation and failure (Figures 1, Figures 2 ) [6]. A binary label system is introduced during fault detection: 0 represents normal data, and 1 represents failure data. A five-level label system is constructed to distinguish the degree of failure during data classification, with 0, 1, 2, 3, and 4 corresponding to the five progressive levels from normal operation to failure, thereby achieving an accurate characterization of the data status.

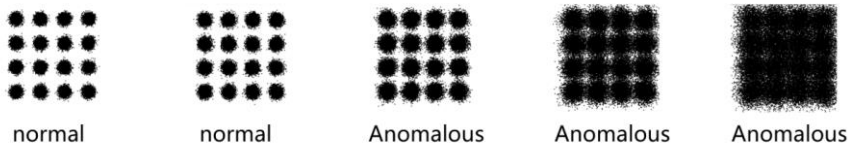



**FIGURE 1.** Legend of normal and abnormal signals in a dataset (Picture credit: Original).

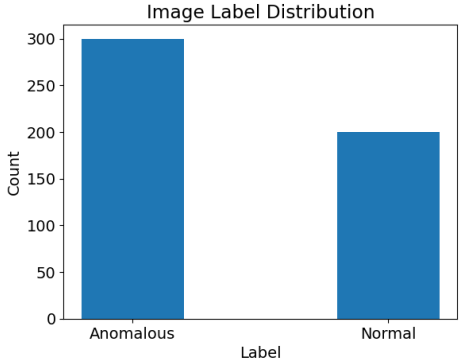



**FIGURE 2.** The number of normal signals and abnormal signals in the dataset (Picture credit: Original).

First, read the image, convert it to a grayscale image, and normalize it for subsequent feature extraction. Then, the FFT algorithm is used to calculate the Fourier transform of the constellation image and flatten the spectrum into one-dimensional data. Using the processed data, calculate Mean Frequency, Std Frequency, Frequency Variance, Mean Frequency_row, and Mean Frequency_col (as shown in Figure 3 and Figure 4). Next, re-read the image, obtain the Edge Density of the data, and use the Gray Level Co-occurrence Matrix (GLCM) to extract the image's Contrast, Correlation, Energy, and Homogeneity features. Finally, save the extracted features to a CSV file. Figure 5 shows the Feature Extraction Flow [7,8].

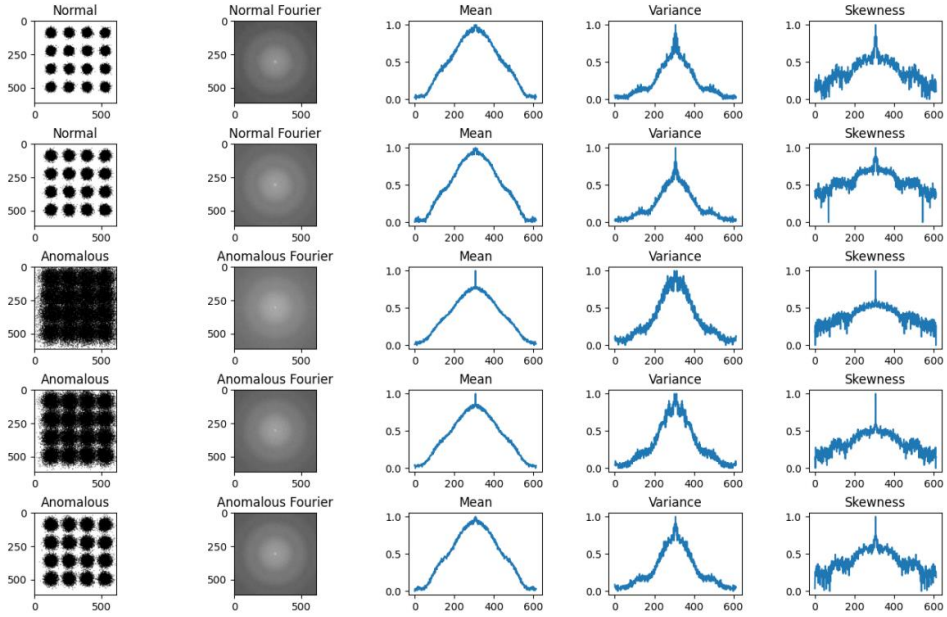



**FIGURE 3.** Visualization of features extracted from constellation images by FFT algorithm (Part 1) (Picture credit: Original).

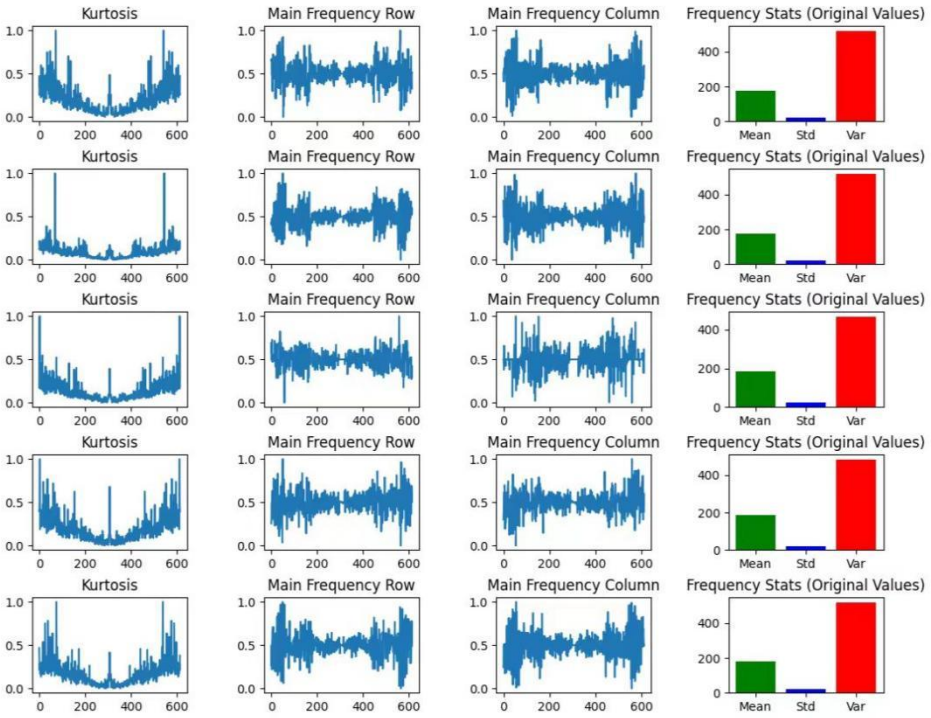



**FIGURE 4.** Visual image of features extracted from constellation image by FFT algorithm (Part 2) (Picture credit: Original).

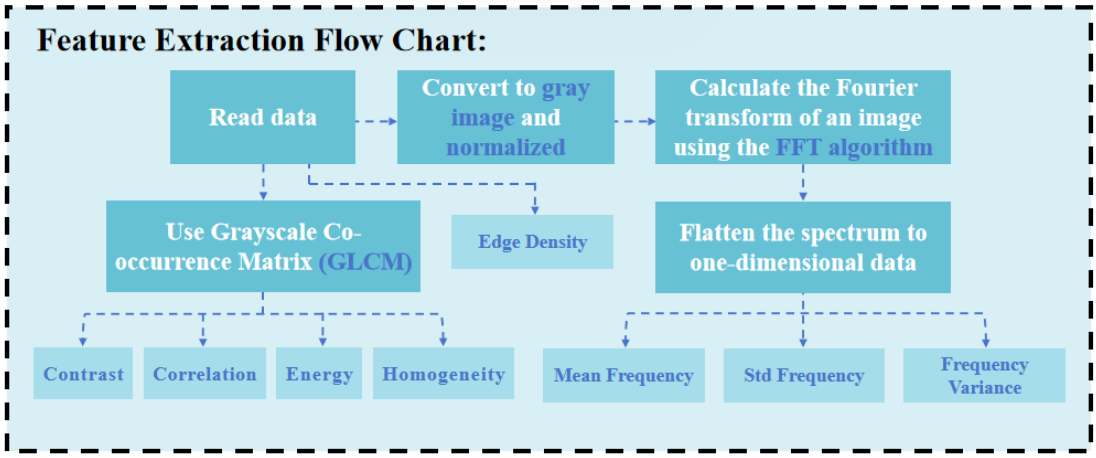
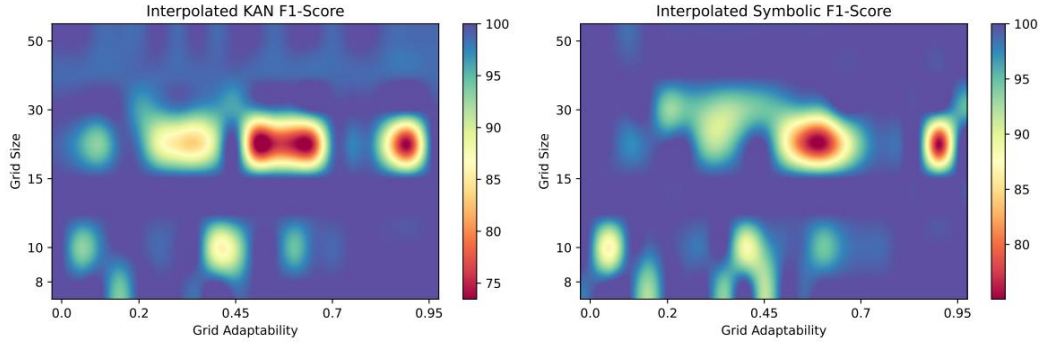



**FIGURE 5.** Feature Extraction Flow Chart (Picture credit: Original).

## Fault Detection and Classification

Firstly, the data sample is divided into training set, verification set and test set according to 7:1.5:1.5. Then, feature selection is performed by minimizing the loss function containing multiple regularization terms. λ represents the opposing hyperparameter, τ represents the domain value that determines feature selection, and [λ,τ] represents the opposing important feature selection. Finally, eight eigenvalues are selected to describe the data. Then, the hyperparameters of KANs are adjusted, including the grid G and the grid adaptation factor $g_e$ ($0 \leqslant g_e \leqslant 1$), and the function model $G = g_e G_u + (1 - g_e) G_a$, $G_u$ is a pure uniform grid parameter, and Ga is a pure adaptive grid parameter) is established to make each layer of grid adapt to the basic data structure of the input to improve the performance of the model. The KAN model is trained in the range of grid resolution (8-50) and adaptability parameter (0-1.0), combined with a dynamic grid update strategy (adjusting the B-spline distribution every 10 steps), and the optimal combination of balancing accuracy and complexity (such as G=15, $g_e$=0.2) is selected through Pareto analysis

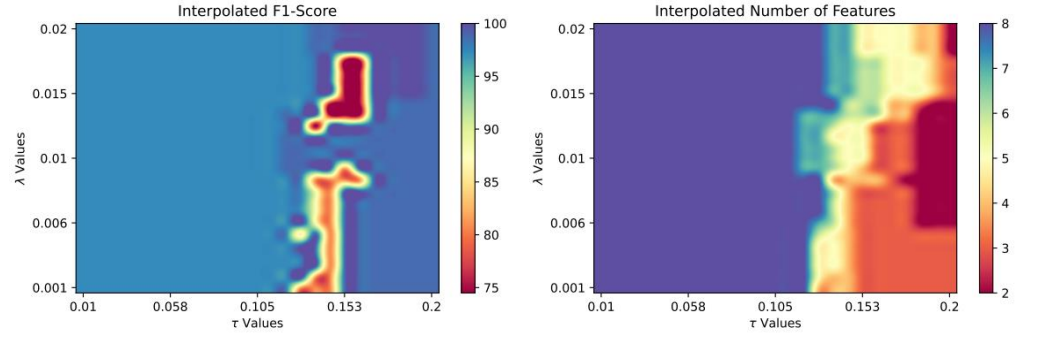
(as shown in Figure 6). The model undergoes refinement via symbolic fitting, with its predictive performance quantified by the determination coefficient $R^2$. The model's explanatory power improves as the coefficient of determination ($R^2$) nears 1.0. The cost function is defined as $C(c, R^2) = \exp(\alpha c) + \beta \ln(1 - R^2)$, where c represents the specified complexity,α is the factor that regulates the complexity, and β is the parameter that influences the quality of the fit. This objective function measures the deviation of the model's outputs from ground truth data. A lower magnitude of this metric corresponds to improved alignment between predictions and observations, reflecting enhanced model efficacy. Parameter optimization during training is achieved by iteratively reducing this function. After selecting several sets of model parameters, each trained model is evaluated on the training set to determine the most important features and the best model parameters. The KAN model is used to traverse 25 sets of thresholds (0.01-0.2) and regularization coefficients (0.001-0.02) to evaluate feature importance and generate the Pareto frontier (as shown in Figure 7). Finally, $\leqslant$4 key features and the optimal regularization parameter (such as λ=0.005, τ=0.156) were screened out, and the high-contribution features were retained for subsequent modeling. Finally, eight important features were determined: Contrast, Correlation, Energy, Homogeneity, Edge Density, Mean Frequency, Std Frequency, and Frequency Variance. Then, the training and validation sets were merged to train the final model parameters. The symbolic expression of KAN was extracted by the auto_symbolic method, and the threshold of $R^2 \geqslant 0$ was set to retain valid items (such as $0.12 * \tanh(1.5x)$). The F1 score of the symbolic model in the test set was verified to be close to that of the original model, ensuring interpretability and no significant loss of accuracy. Finally, the debugged KAN model was evaluated on an independent test set, including accuracy, precision, recall, and F1 score, and the evaluation results of binary tags and quintuple tags were output, respectively.




(a) (b)

**FIGURE 6.** (a) Grid Size vs. Grid Adaptability for KAN F1-Score (b) Grid Size vs. Grid Adaptability for Interpolated Symbolic F1-Score (Picture credit: Original).

By analyzing the two heatmaps, researchers can avoid the red low-performance region and choose a point in the blue Pareto efficient region (such as G=15, $g_e$=0.2), thereby determining a KAN model structure that is both accurate and easy to translate into symbolic formulas.




(a) (b)

**FIGURE 7.** (a) Interpolated F1-Score for λ Values vs. τ Values (b) Interpolated Number of Features for λ Values vs. τ Values (Picture credit: Original).

By analyzing these two heatmaps, researchers can accurately select λ and τ (such as λ=0.005, τ=0.156), so as to screen out the most core features with the minimum performance cost, which greatly enhances the simplicity and interpretability of the model.

The identical dataset was independently processed using conventional machine learning models, after which the precision, accuracy, recall, and F1 scores were compared statistically. The process of fault detection is shown in the Figure 8.

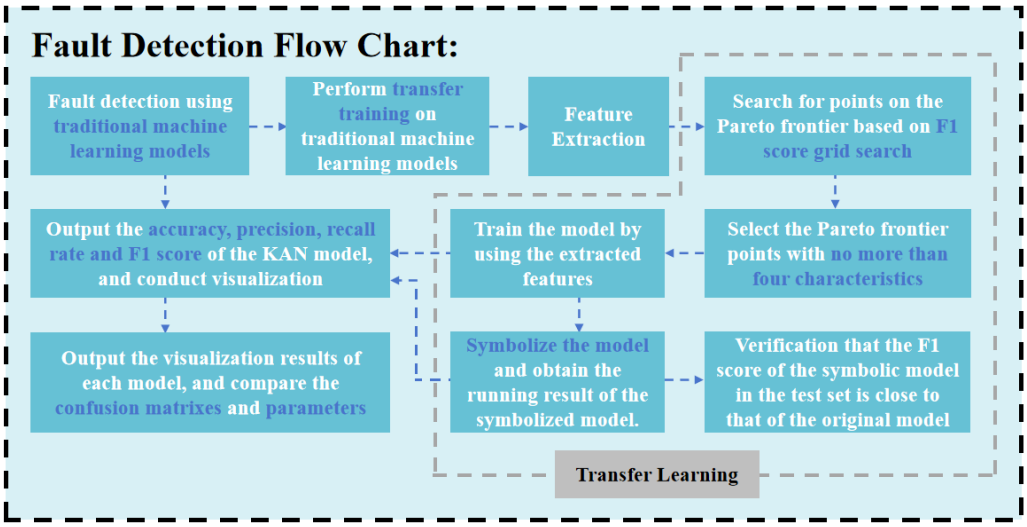



**FIGURE 8.** Fault Detection Flow Chart (Picture credit: Original).

## RESULTS

In the ultimate assessment of the model, the KAN-FFT model demonstrated remarkable capability in fault detection. In the fault detection task for binary labeled data, the accuracy, precision, recall, and F1 score of the KAN model and the model after transfer learning reached 1.0, indicating that it can identify faults and normal states with extremely high accuracy. As shown in Table 1 and Figure 9~Figure 12, under the same data set test environment, the KAN-FFT model performed significantly better than traditional machine learning models, including the support vector machine (SVM) and isolation forest models [9,10].

**TABLE 1.** Execution effect of KAN network on data of binary classification label.

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 1.00 | 1.00 | 1.00 | 46 |
| 1 | 1.00 | 1.00 | 1.00 | 30 |
| accuracy | | | 1.00 | 76 |
| macro avg | 1.00 | 1.00 | 1.00 | 76 |
| weighted avg | 1.00 | 1.00 | 1.00 | 76 |

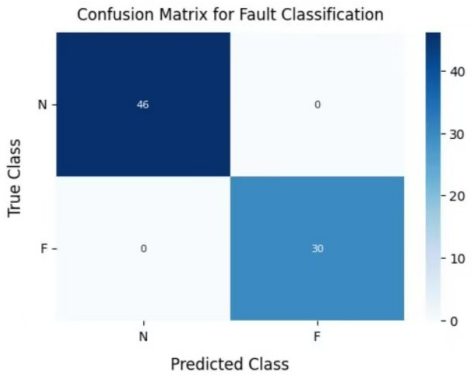



**FIGURE 9.** Execution effect of KAN network on data transfer of binary classification label (Picture credit: Original).

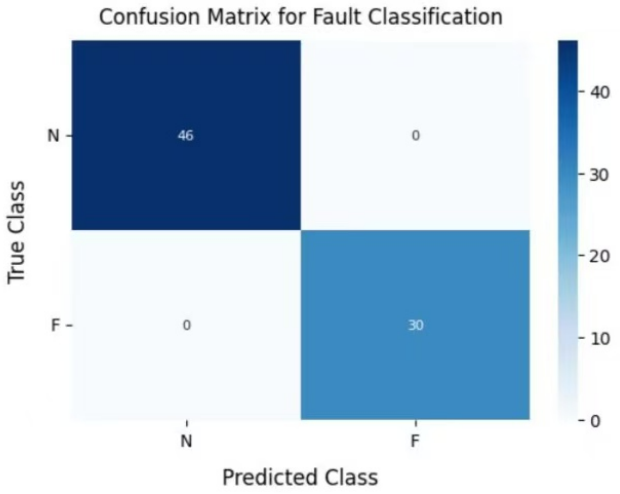



**FIGURE 10.** Execution effect of KAN network after learning of data of binary classification labels (Picture credit: Original).

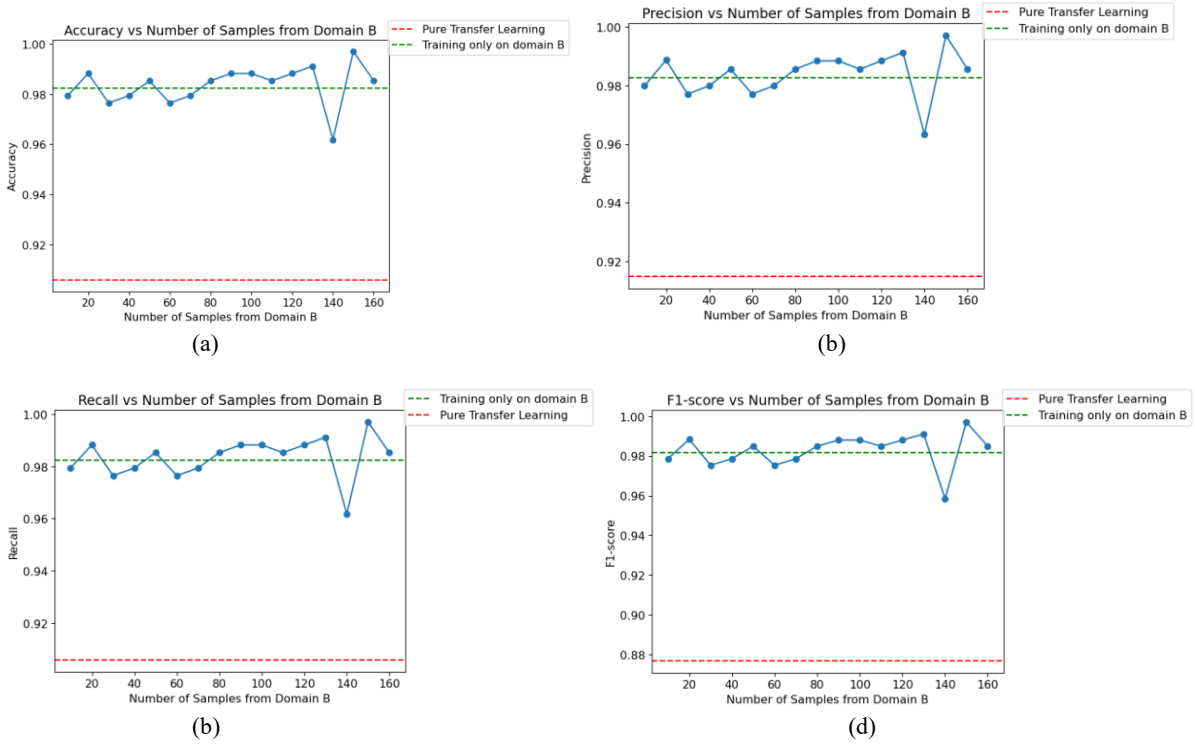



**FIGURE 11.** Results of isolated forest model execution (a) Accuracy (b) Precision (c) Recall (d) F1-Score (Picture credit: Original).

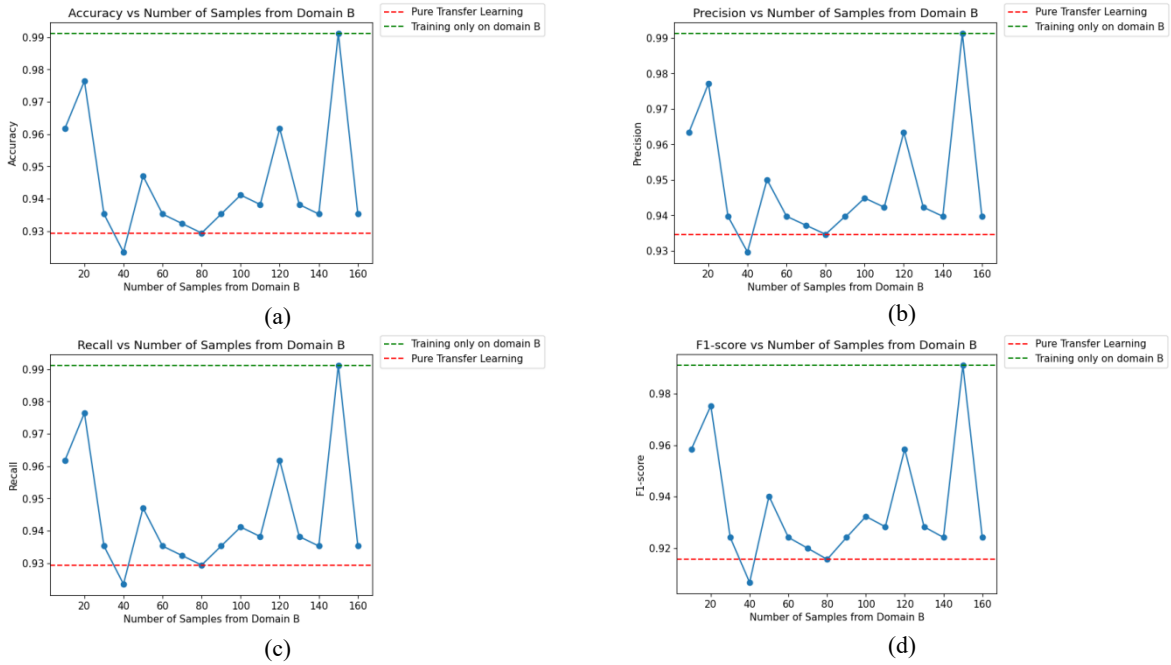



**FIGURE 12.** Results of SVM model execution (a) Accuracy (b) Precision (c) Recall (d) F1-Score (Picture credit: Original).

Figure 11 consists of four subfigures, corresponding to the line charts of the four key indicators of the model in the binary classification task: Accuracy, Precision, Recall, and F1-Score, show the execution results of the isolation forest model, and Figure 12, with the same composition as Figure 11, shows the execution results of the SVM model with respect to the four key indicators.

Through the visualization of Figure 11 and Figure 12, it can be clearly seen that the index curves of the two traditional models show significant volatility, which is in sharp contrast with the stable performance of the FFT-KAN model approaching 1.0 (Figure 13, Figure 14), and clearly indicates that in the complex multi-classification task of optical fiber fault, The performance of the traditional isolation forest and SVM model is unstable and not good enough, which strongly proves the excellence and advanced of the FFT-KAN model in accuracy, precision and stability.

As shown in Table 2 and Table 3, Figure 13 and Figure 14, both the original and symbolic models classified the five types of data well, achieving an accuracy rate of 99%. The high accuracy, precision, recall, and F1 score indicate that the model can accurately classify the sample status.

**TABLE 2.** Execution effect of KAN network on data of five classification labels.

| | **precision** | **recall** | **f1-score** | **support** |
|---|---|---|---|---|
| 0 | 1.00 | 0.93 | 0.97 | 15 |
| 1 | 0.94 | 1.00 | 0.97 | 15 |
| 2 | 1.00 | 1.00 | 1.00 | 15 |
| 3 | 1.00 | 1.00 | 1.00 | 16 |
| 4 | 1.00 | 1.00 | 1.00 | 15 |
| accuracy | | | 0.99 | 76 |
| macro avg | 0.99 | 0.99 | 0.99 | 76 |
| weighted avg | 0.99 | 0.99 | 0.99 | 76 |

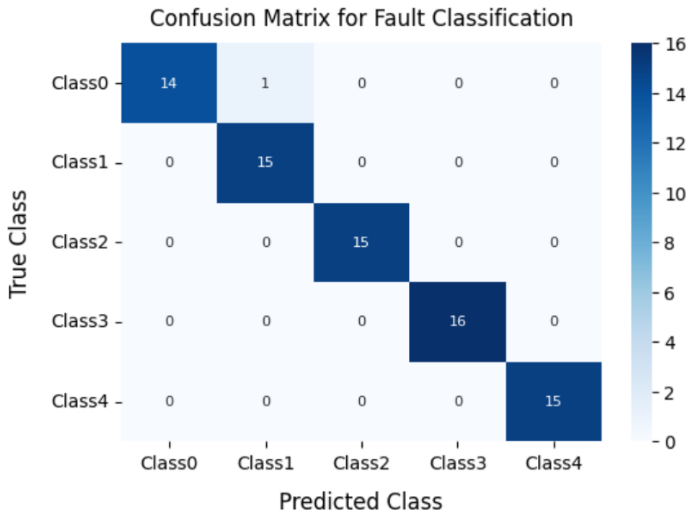



**FIGURE 13.** Execution effect of KAN network on data of five classification labels under the original model (Picture credit: Original).

**TABLE 3.** Execution effect of KAN network on data of five classification labels.

| | **precision** | **recall** | **f1-score** | **support** |
|---|---|---|---|---|
| 0 | 1.00 | 0.93 | 0.97 | 15 |
| 1 | 0.94 | 1.00 | 0.97 | 15 |
| 2 | 1.00 | 1.00 | 1.00 | 15 |
| … | | | | |
| accuracy | | | 0.99 | 76 |
| macro avg | 0.99 | 0.99 | 0.99 | 76 |
| weighted avg | 0.99 | 0.99 | 0.99 | 76 |

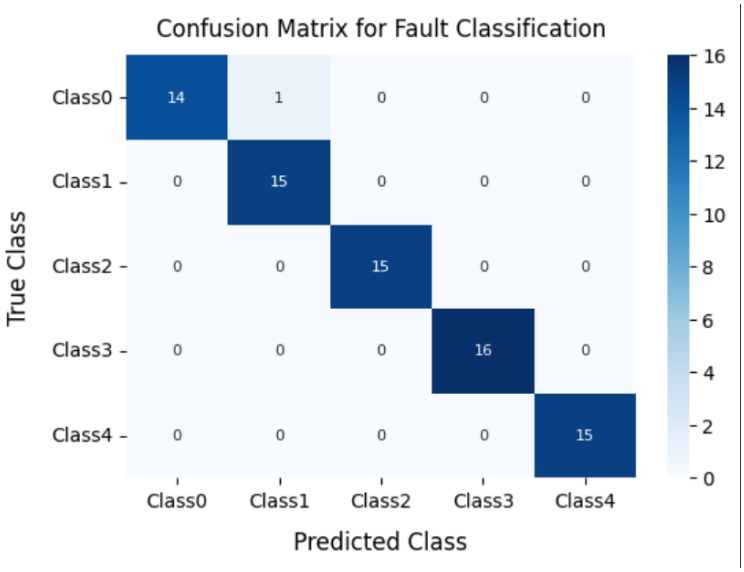



**FIGURE 14.** Implementation effect of KAN network on data with five classification labels under symbolic model (Picture credit: Original).

## CONCLUSION

This research presents a novel approach to optical fiber fault detection and migration training by integrating the Fast Fourier Transform (FFT) with the Kolmogorov-Arnold Network (KAN), herein referred to as the FFT-KAN model. The proposed method addresses several challenges inherent in traditional optical performance monitoring techniques, such as the heavy dependency on high-quality and abundant data, limited model interpretability, and suboptimal computational efficiency. By leveraging the FFT algorithm for feature extraction, the model effectively transforms 16QAM constellation diagrams from the time domain into the frequency domain, capturing critical frequency-domain characteristics and image-based features such as edge density, contrast, and correlation. This transformation facilitates the accurate identification of fault patterns, which is crucial in detecting subtle anomalies in optical networks.

The KAN model, derived from the Kolmogorov-Arnold theorem, significantly enhances the interpretability and efficiency of the diagnostic process. Unlike conventional neural networks, KAN decomposes complex multivariable functions into simpler univariate functions by combining the strengths of multilayer perceptrons and splines. This unique combination simplifies the network structure while ensuring that essential features are optimized accurately. As demonstrated by our experimental results, the FFT-KAN model achieved 100% accuracy, precision, recall, and F1 scores in binary classification tasks, indicating its robust capability in distinguishing between normal and faulty states. Furthermore, in more complex five-class fault detection scenarios, the model maintained an impressive overall accuracy of 99%, outperforming established conventional techniques such as support vector machines (SVM) and isolation forest models.

The integration of FFT and KAN in the proposed model not only provides high diagnostic accuracy but also offers significant improvements in model interpretability and computational efficiency. The Pareto analysis conducted during model training ensured that the balance between accuracy and model complexity was optimized, resulting in a diagnostic tool that is both powerful and practical for real-time fault detection in optical networks. The symbolic fitting process further validates that the derived symbolic expression of the KAN model retains the performance of the original model, ensuring that interpretability is not sacrificed for accuracy. In addition to demonstrating superior performance compared to traditional machine learning models, the FFT-KAN model's ability to adapt to varying data distributions and dynamically update its internal parameters positions it as a promising candidate for future integration with emerging technologies. The potential to incorporate big data analytics, cloud computing, and fog computing into the framework opens up new avenues for scaling and enhancing the model's applicability in intelligent network management and maintenance. This adaptability is critical for addressing the challenges posed by the increasing complexity and data volume in modern optical networks.

## CONTRIBUTION

All the authors contributed equally and their names were listed in alphabetical order.

## REFERENCES


1. X. Li, "Research on optical network fault detection and location technology based on machine learning," Ph.D. thesis, Beijing University of Posts and Telecommunications, 2024. DOI: 10.26969/d.cnki.gbydu.2024.000614.
2. H. Li, W. Liu, and M. Luo, "Optical network monitoring and optimization method based on machine learning," Opt. Commun. Res., vol. 03, pp. 5–14, 2024.
3. N. Kaur, "Hybrid image splicing detection: Integrating CLAHE, improved CNN, and SVM for digital image forensics," Expert Syst. Appl., vol. 273, p. 126756, 2025.
4. R. Montañana, A. J. Gámez, and M. J. Puerta, "ODTE—An ensemble of multi-class SVM-based oblique decision trees," Expert Syst. Appl., vol. 273, p. 126833, 2025.
5. X. Zhu, Z. Liu, E. Cambria, X. Yu, X. Fan, H. Chen, and R. Wang, "A client–server based recognition system: Non-contact single/multiple emotional and behavioral state assessment methods," Comput. Methods Programs Biomed., vol. 260, p. 108564, 2025.
6. R. Wang, J. Zhu, S. Wang, T. Wang, J. Huang, and X. Zhu, "Multi-modal emotion recognition using tensor decomposition fusion and self-supervised multi-tasking," Int. J. Multimed. Inf. Retr., vol. 13, no. 4, pp. 39, 2024.
7. F. Wang, M. Ju, X. Zhu, Q. Zhu, H. Wang, C. Qian, and R. Wang, "A geometric algebra-enhanced network for skin lesion detection with diagnostic prior," J. Supercomput., vol. 81, no. 1, pp. 1–24, 2025.

8. A. A. M. Shannaq, A. Nasrawi, K. R. A. A. Bsoul, et al., "Abnormal heart sound recognition using SVM and LSTM models in real-time mode," Sci. Rep., vol. 15, no. 1, p. 9129, 2025.
9. M. Bergström, S. Lundberg, and M. Cederblad, "Funktionell familjeterapi (FFT) på ett socialkontor: En kvalitetsgranskning och en forskningsutvärdering," Research Reports in Social Work, no. 7, 2021.
10. H. Zhang, H. Zhang, Z. Liu, et al., "Visualization of soil freezing phase transition and moisture migration using polymer optical fibers," Measurement, vol. 229, p. 114402, 2024.