ML-Based Detection and Categorization of Complex Mechanical Vibrations via State of Polarization Analysis in Optical Networks

Leyla Sadighi[1], Stefan Karlsson[2], Marco Ruffini[1], Marija Furdek[3]

[1] *Department of Computer Science and Statistics, CONNECT center, Trinity College Dublin, Ireland*
[3]*Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden*
[2] *AB Micropol Fiberoptic, Halmstad, Sweden.*
E-mail: sadighil@tcd.ie

ABSTRACT Modern optical networks form the critical backbone of global communications, enabling high-speed data transmission for a wide range of applications. Despite their inherent advantages in bandwidth and scalability, these networks are not immune to physical-layer vulnerabilities. Mechanical disturbances, both accidental and intentional, can compromise service quality or serve as gateways for more severe cyber-physical attacks. Thus, there is a growing need for intelligent, real-time monitoring solutions capable of detecting and interpreting subtle anomalies in optical fiber infrastructures. This paper presents a Machine Learning (ML)-based State of Polarization (SOP) monitoring approach for the identification and classification of complex mechanical vibrations in optical fiber networks. We address the real-world challenge of mixed-frequency and overlapping vibration signatures, arising from benign activities, malicious attacks, or simultaneous events, by collecting 14 distinct polarization signatures under various physical scenarios. A diverse set of supervised ML classifiers is evaluated, with Histogram Gradient Boosting (HGB) achieving the highest performance at 88.33% accuracy.
Keywords: Optical Fiber Monitoring, Mechanical Vibrations, State of Polarization (SOP), Perturbation, Supervised Machine Learning (ML), Anomaly Detection, Eavesdropping, Classification.

1. Introduction

Optical networks constitute the core infrastructure for high-capacity and long-distance data transmission, delivering low-loss and high-bandwidth connectivity that is essential to modern communication networks. Their role is critical in a wide range of domains, including telecommunications, healthcare, defense, and data center interconnects. Despite their physical robustness, optical fibers remain susceptible to external mechanical disturbances that can degrade signal quality or compromise security. In particular, construction activities and heavy machinery, such as excavators operating near buried cables, generate low-frequency vibration patterns that may damage the infrastructure and pose a significant risk of accidental fiber cuts. More alarmingly, deliberate security breaches such as covert eavesdropping, enabled by deliberately bending the fiber to a specific degree [1] to extract optical signals, raise serious concerns about the confidentiality of transmitted information. Such acts may go undetected without fine-grained monitoring.

In light of these vulnerabilities, recent real-world incidents involving sabotage and tampering of fiber-optic infrastructure have underscored the urgent need for advanced threat detection mechanisms [2]. Parallel to this, emerging research has demonstrated the potential of utilizing existing optical fiber infrastructure for environmental sensing, with proven effectiveness in capturing both natural and human-induced activities [3]. A key enabler of such sensing is the State of Polarization (SOP), which is particularly effective in detecting subtle and complex vibration patterns caused by physical tampering. SOP-based sensing offers key advantages over traditional vibration detection methods due to its inherent sensitivity to minute physical perturbations in the optical fiber. Unlike conventional methods, such as Distributed Acoustic Sensing (DAS) or Optical Time Domain Reflectometry (OTDR), which often require specialized infrastructure, backscattering techniques, or high power levels, SOP-based analysis leverages the intrinsic polarization variations of light propagating through standard fiber [4]. Furthermore, when combined with Machine Learning (ML) techniques, SOP-based sensing enables real-time classification and anomaly detection, offering a scalable and robust solution for protecting critical optical communication infrastructure. Recent studies have extensively investigated ML-based analysis of SOP variations induced by physical disturbances such as eavesdropping attempts and mechanical vibrations [5]–[11]. These works demonstrate the effectiveness of SOP for anomaly detection frameworks in reliably identifying and characterizing disruptive events under experimental and real-world deployment scenarios. However, these research works did not consider the challenge of overlapping or mixed-frequency vibration patterns.

In a real-life installation, a normal event such as traffic passing close to the installation causes signatures with a broad-frequency spectrum content. This spectrum could be related to signatures caused by an eavesdropping event, or it can be mixed with other broad-spectrum signatures from normal traffic or from a malicious excavator threatening to cut the fiber optical cable. In order to avoid false alarms, it is therefore essential to classify and separate the frequency spectrum contents from signatures representing potential harmful events. In practical deployments, capturing clean and isolated signatures of real-world mechanical disturbances, such as those caused by heavy vehicles or trains passing near the installation, is a time-consuming and operationally challenging task.

These naturally occurring events often produce highly variable and overlapping vibration patterns that are difficult to reproduce under controlled conditions. To address this, we simulate disturbance environments by generating two complex vibration signatures. These synthetic signatures feature pseudo-random frequency content ranging from 0 to 2000 Hz, effectively emulating the spectral characteristics observed in real-world disturbances. To evaluate the classifier's robustness in differentiating between benign and potentially malicious activities, we further combine these complex vibration signatures with known attack scenarios, including fiber eavesdropping and soft patch cable bending. This paper addresses the gap in the literature by proposing an ML classification technique capable of separating and classifying complex, co-occurring mechanical signatures using SOP-based signatures analysis.

2. Experimental Setup

Our experimental setup, depicted in Figure 1, is designed to generate and analyze polarization signatures resulting from mechanical perturbations applied to an optical fiber transmission line. A stabilized 1310 nm Continuous Wave Distributed Feedback (CW-DFB) laser serves as the optical source, injecting constant polarized light into a 1 km coupling fiber. This coupling fiber connects to the central sensing region, where deliberate physical manipulations are applied to either a bare fiber or a patch cable segment. The perturbed signal then propagates through a 20 km fiber spool, bringing the total optical transmission length to 21 km and emulating a real-world long-distance deployment. Mechanical perturbations acting on the fiber installation induce measurable SOP variations, whose magnitude and frequency content are characteristic of the specific external activity, resulting in distinct polarization signatures. To simulate realistic and challenging monitoring conditions in this research work, we introduce two complex vibration patterns, referred to as complex vibration A (A) and complex vibration B (B), each lasting over 10 and 20 minutes, respectively. These patterns are composed of pseudo-random frequency components spanning the 0–2000 Hz range, mimicking naturally occurring disturbances such as persistent background noise generated by heavy traffic or trains passing near fiber installations. To simulate overlapping effects, we mix complex vibration A and complex vibration B with known events like eavesdropping, soft bending, and potentially harmful vibrations.

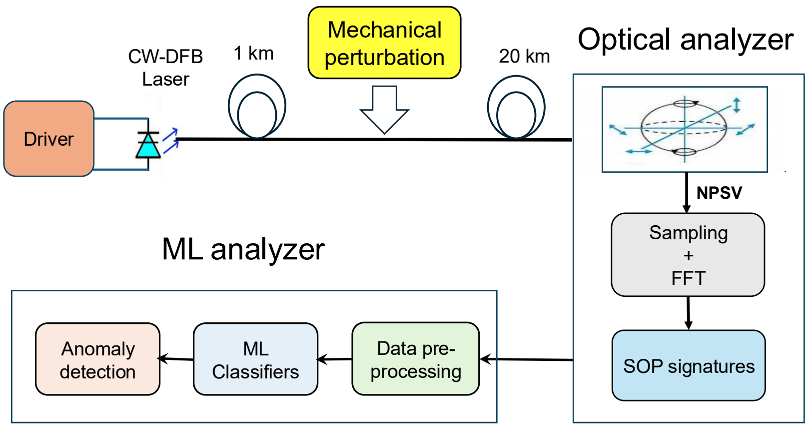

Figure 1: Schematic of the experimental testbed used for polarization signature analysis. Mechanical disturbances are applied between the 1 km coupling fiber and the 20 km fiber spool.

To analyze the impact of mechanical disturbances and generate distinct SOP variation signatures, we follow the approach introduced in [7]. As illustrated in Figure 1, an optical analyzer computes the numerical variation of the SOP, referred to as Numerical Polarization State Variation (NPSV), as time-sequenced samples on the Poincaré sphere at 0.5 ms intervals over 10-minute and 20-minute durations for complex vibrations A and B, respectively, yielding approximately 1.2 million and 2.4 million data points per event. The NPSV is then segmented into windows of 1000 samples and transformed into the frequency domain via an Fast Fourier Transform (FFT) with 512 bins. This results in a time-frequency data matrix (SOP signature) with the shape $[1200, 512]$ for events containing complex vibrations A and $[2400, 512]$ for complex vibrations B, forming the basis of our ML analysis.

The ML analysis is grounded in a comprehensive dataset generated through controlled combinations of complex vibrations and targeted mechanical perturbations applied to both bare fiber and patch cable segments. We consider three representative tampering scenarios: soft bending (sb), eavesdropping (eav), and malicious vibrations at 80 Hz (*80vb*). The *sb* condition emulates benign maintenance activity, where fibers are gently bent to a radius of approximately 2 cm at 10-second intervals. The *eav* scenario simulates intentional tapping by introducing a 4 mm-radius, 25-degree bend using a precision coupler, as described in [12]. The *80vb* scenario replicates ground-borne vibrations from nearby excavation equipment, modeled by applying an 80 Hz, 60 dBA sinusoidal tone from a loudspeaker placed 5 cm from the fiber. These events are overlaid with either complex vibration A or B to emulate challenging real-world environments.

3. Signatures and Data Collection

We collected the following 14 distinct SOP signature involving two types of complex mechanical vibrations, denoted as A and B, applied to both bare fiber and patch cable under various tampering conditions:

- Complex A/B on bare fiber: A_{br}, B_{br}
- Complex A/B on bare fiber + eavesdropping: A_{br+eav}, B_{br+eav}
- Complex A/B on bare fiber + 80 Hz vibration: $A_{br+80vb}$, $B_{br+80vb}$
- Complex A/B on bare fiber + soft bending: A_{br+sb}, B_{br+sb}
- Complex A/B on patch cable: A_{pc}, B_{pc}
- Complex A/B on patch cable + soft bending: A_{pc+sb}, B_{pc+sb}
- Complex A/B on patch cable + eavesdropping: A_{pc+eav}, B_{pc+eav}

The ML analyzer, illustrated in Figure 1, comprises three main stages: data preprocessing, classification, and anomaly detection. In the preprocessing phase, the 14 collected signatures are merged to construct training and testing datasets. Each event involving complex vibration A consists of 1,200 points (8,400 in total across seven scenarios), while those involving vibration B yield 2,400 points (16,800 total). An 80/20 split is applied to each class, resulting in 960 training and 240 testing points per A scenario, and 1,920 training and 480 testing points per B scenario. After aggregation, the final dataset comprises 20,160 training points and 5,040 testing points. This dataset is then analyzed using supervised ML models to detect anomalies indicative of potentially harmful or malicious events, enabling robust identification of overlapping and complex disturbances in real-world optical fiber installations.

4. Results

We conduct a comprehensive evaluation of multiple supervised ML classifiers to identify the most suitable model for detecting and categorizing mechanical vibration signatures in optical fibers. A diverse range of classifiers from the Scikit-learn library is evaluated, including ensemble-based methods (Histogram Gradient Boosting (HGB), eXtreme Gradient Boosting (XGBoost), Random Forest (RF), Extra Trees Classifier (ETC)), kernel-based models (Support Vector Machine (SVM)), linear models (Logistic Regression (LR), Linear Discriminant Analysis (LDA)), instance-based learning (k-Nearest Neighbors (KNN)), and Decision Tree (DT). These models are selected for their proven applicability in similar time-frequency classification tasks.

TABLE I: Performance comparison of various ML classifiers

Classifier Name	**Accuracy**	**Precision**	**Recall**	**F1-score**	**Training Time (s)**	**Inference Time (s)**
Hist Gradient Boosting	0.8833	0.8856	0.8833	0.8828	47.1555	0.1349
XGBoost	0.8724	0.8741	0.8724	0.8716	24.9517	0.0368
Random Forest	0.8387	0.8450	0.8387	0.8363	41.6140	0.0859
Gradient Boosting	0.8171	0.8211	0.8171	0.8157	3640.1562	0.1015
SVM Classifier	0.8163	0.8289	0.8163	0.8151	36.5793	25.3881
Extra Trees Classifier	0.7972	0.8076	0.7972	0.7906	7.6353	0.1281
Logistic Regression	0.7014	0.6992	0.7014	0.6991	53.7773	0.0183
K-Nearest Neighbors	0.6671	0.7297	0.6671	0.6467	0.0187	0.5907
Linear Discriminant Analysis	0.6609	0.6586	0.6609	0.6517	1.0038	0.0062
Decision Tree	0.6583	0.6592	0.6583	0.6585	16.2349	0.0041

Table I presents a detailed comparison of classifier performance in terms of accuracy, precision, recall, F1-score, training time, and inference latency. The HGB classifier delivers the highest overall performance, achieving an accuracy of 88.33% and an F1-score of 0.8828. It also offers a favorable trade-off between predictive accuracy and computational efficiency. XGBoost and RF also demonstrate good performance, with accuracies of 87.24% and 83.87%, respectively. Although SVM yields comparable accuracy (81.63%), its inference time of over 25 seconds makes it impractical for real-time applications. Meanwhile, Gradient Boosting (GB), despite being conceptually robust, requires a substantially longer training time (over 3,600 seconds) without a corresponding performance gain. Simpler classifiers such as LR, KNN, LDA, and DT show significantly lower accuracy and are less effective at capturing the complex patterns inherent in SOP-based vibration signatures.

The confusion matrix in Figure 2 further demonstrates the strength of the HGB classifier in accurately identifying the 14 mechanical disturbance scenarios. Perfect and near-perfect classification was achieved for the $A_{br+80vb}$ and $B_{br+80vb}$ scenarios (100.0% and 99.4% accuracy), while high accuracy was also observed for baseline conditions like A_{br} (95.4%) and B_{br} (99.4%). Misclassifications were mostly concentrated in complex overlapping scenarios that contain *sb* and *eav*. For example, B_{br+eav} showed confusion with B_{pc+eav}. Likewise, A_{pc+sb} and B_{pc+sb} were misclassified with related classes such as A_{pc+eav} and B_{br+sb}, likely due to overlapping frequency content and compounded mechanical effects. Despite these challenges, the classifier consistently preserved the dominant spectral cues that distinguish each class, confirming its strong effectiveness in SOP-based ML classification for recognizing subtle and overlapping disturbances in optical network environments.

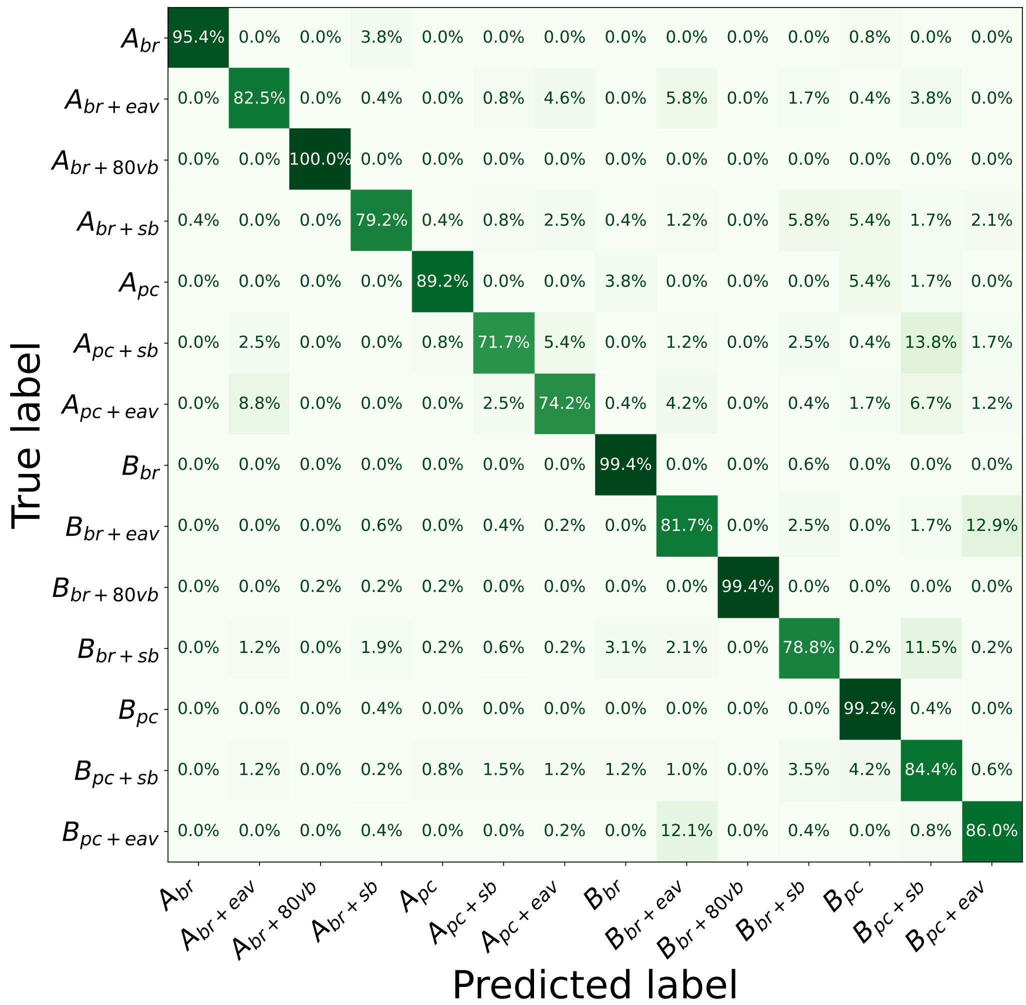

Figure 2: Results of Confusion Matrix for Histogram Gradient Boosting classifier

5. Conclusion

This paper proposed an ML-based classification of SOP signatures analysis for identifying and categorizing complex mechanical disturbances in optical fiber networks. Considering that real-world optical infrastructures are increasingly exposed to both benign and malicious mechanical activities, often occurring simultaneously and with overlapping spectral characteristics, we proposed a robust sensing approach leveraging SOP sensitivity and advanced supervised ML classification. A comprehensive dataset comprising 14 distinct SOP signatures was collected under controlled experimental conditions. Multiple supervised ML classifiers were evaluated for their performance in classifying these events, with Histogram Gradient Boosting (HGB) emerging as the most accurate and computationally efficient model, achieving an accuracy of 88.33%. Overall, the results validate the effectiveness of SOP-based spectral analysis combined with ML in accurately distinguishing and categorizing co-occurring mechanical disturbances, paving the way for robust monitoring and threat detection in future optical communication networks.

REFERENCES

[1] S. Karlsson, R. Lin, L. Wosinska, and P. Monti, “Eavesdropping g.652 vs. g.657 fibres: a performance comparison,” in *International Conference on Optical Network Design and Modeling (ONDM)*, 2022, pp. 1–3.

[2] BBC News, “France investigates after suspected sabotage hits fibre-optic cables,” https://www.bbc.com/news/articles/cn52rglxr62o, 2025.

[3] G. Marra, D. Fairweather, V. Kamalov, P. Gaynor, M. Cantono, S. Mulholland, B. Baptie, J. Castellanos, G. Vagenas, J.-O. Gaudron *et al.*, “Optical interferometry–based array of seafloor environmental sensors using a transoceanic submarine cable,” *Science*, vol. 376, no. 6595, pp. 874–879, 2022.

[4] S. Pellegrini, L. Minelli, L. Andrenacci, G. Rizzelli, D. Pilori, G. Bosco, L. D. Chiesa, C. Crognale, S. Piciaccia, and R. Gaudino, “Overview on the state of polarization sensing: application scenarios and anomaly detection algorithms,” *J. Opt. Commun. Netw.*, vol. 17, no. 2, pp. A196–A209, Feb 2025.

[5] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, “Detection and classification of eavesdropping and mechanical vibrations in fiber optical networks by analyzing polarization signatures over a noisy environment,” in *ECOC 2024; 50th European Conference on Optical Communication*, 2024, pp. 527–530.

[6] L. Sadighi, S. Karlsson, L. Wosinska, and M. Furdek, “Machine learning analysis of polarization signatures for distinguishing harmful from non-harmful fiber events,” in *Proc. ICTON*, 2024, pp. 1–5.

[7] L. Sadighi, S. Karlsson, C. Natalino, and M. Furdek, “Machine learning-based polarization signature analysis for detection and categorization of eavesdropping and harmful events,” in *2024 Optical Fiber Communications Conference and Exhibition (OFC)*, 2024, pp. 1–3.

[8] L. Sadighi, S. Karlsson, C. Natalino, L. Wosinska, M. Ruffini, and M. Furdek, “Deep learning for detection of harmful events in real-world, noisy optical fiber deployments,” *Journal of Lightwave Technology*, pp. 1–9, 2025.

[9] K. Abdelli, M. Lonardi, J. Gripp, S. Olsson, F. Boitier, and P. Layec, “Breaking boundaries: harnessing unrelated image data for robust risky event classification with scarce state of polarization data,” in *ECOC Conf.*, vol. 2023. IET, 2023, pp. 924–927.

[10] K. Abdelli, M. Lonardi, J. Gripp, D. Correa, S. Olsson, F. Boitier, and P. Layec, “Anomaly detection and localization in optical networks using vision transformer and sop monitoring,” in *OFC Conf.*, 2024, pp. 1–3.

[11] K. Abdelli, M. Lonardi, F. Boitier, D. Correa, J. Gripp, S. Olsson, and P. Layec, “Vision transformers for anomaly classification and localization in optical networks using sop spectrograms,” *Journal of Lightwave Technology*, vol. 43, no. 4, pp. 1902–1913, 2025.

[12] S. Karlsson, R. Lin, L. Wosinska, and P. Monti, “Eavesdropping G. 652 vs. G. 657 fibres: a performance comparison,” in *International Conference on Optical Network Design and Modeling (ONDM)*, 2022, pp. 1–3.