Research Article | Vol. 27, No. 17 | 19 Aug 2019 | OPTICS EXPRESS 24808
Optics EXPRESS

Optical spectrum feature analysis and recognition for optical network security with machine learning

YANLONG LI,[1,2] NAN HUA,[1,2] JIADING LI,[1,2] ZHIZHEN ZHONG,[1,2] SHANGYUAN LI,[1,2] CHEN ZHAO,[1,2] XIAOXIAO XUE,[1,2] AND XIAOPING ZHENG[1,2,*]

[1]*Beijing National Research Center for Information Science and Technology (BNRist), Beijing 100084, China*
[2]*Department of Electronic Engineering, Tsinghua University, Beijing 100084, China*
**xpzheng@mail.tsinghua.edu.cn*

Abstract: Physical layer attacks threaten services transmitted through optical networks. To detect attacks, we present an investigation of optical spectrum feature analysis (OSFA) and recognition. By analyzing the spectral features of optical signals, recognition and detection of unauthorized signals can be realized. In this paper, (1) we theoretically analyzed factors influencing optical spectrum (OS) features and simulated these factors. OSs collected from the simulation are quantitatively analyzed, spectral features are extracted by principal component analysis, and the theoretical derivation is validated. (2) We proposed support vector machine (SVM) and one-dimensional convolutional neural network (1D-CNN) machine-learning OSFA methods. (3) Experimentally collected OSs from commercial small form-factor pluggable modules are used to verify the performance of the SVM and 1D-CNN methods, which achieved 98.54% and 100% recognition accuracies, respectively, demonstrating that the methods are promising solutions for optical network security.

1. Introduction

1.1. Background of optical network security issues

Driven by massive amounts of data and a large number of provided services, and an increasingly complex network architecture, current optical networks face multiple security threats, rendering it impossible to use traditional means to ensure the security of network provided services. Eavesdropping and multiple attack methods in optical networks are constantly evolving, and the security threats the network is exposed to are increasingly sophisticated. Optical network security threats such as transmission quality reduction, communication interruption and services blocking will continue to cause significant impact, making optical network security a pressing research topic.

Due to certain vulnerabilities, optical networks are easily threatened by eavesdropping and multiple attacks [1–5]. Different from the information level attacks or data leakage in the Internet, the eavesdropping and attack of optical networks are mainly considered to occur at the physical layer [1,3]. With increased research in recent years, physical layer defects of optical networks have drawn researchers' attention. Generally, threats faced by optical networks at the physical layer can be divided into three main categories: (1) eavesdropping, (2) masquerade attacks in which unauthorized users gain access to optical nodes incognito, (3) jamming attacks preventing optical networks from providing services [1,2]. The threats to the optical physical layer are difficult to solve with packet level security methods, thus, optical layer security approaches are required.

Optical network security approaches focused on attack prevention, detection, and reaction have been studied and proposed. Some researchers have proposed security approaches

#369168 https://doi.org/10.1364/OE.27.024808
Journal © 2019 Received 6 Jun 2019; revised 13 Jul 2019; accepted 2 Aug 2019; published 16 Aug 2019

Optics EXPRESS

focusing on attack prevention, and reaction mechanisms such as attack-aware route wavelength allocation (RWA) [6], RWA in multi-domain/multi-core fiber environments [7,8], and lightpath hopping [9]. Meanwhile, few physical-layer based encryption techniques have been proposed such as the chaotic constellation transformation technique [10], and time-frequency domain encryption [11] which encrypt the signal to mitigate eavesdropping.

1.2. Problem description of signal insertion attacks

The problem investigated in this paper is focusing on the security issue that attackers or unauthorized users gain access to the network and insert their own signals to the network, which is called signal insertion attacks.

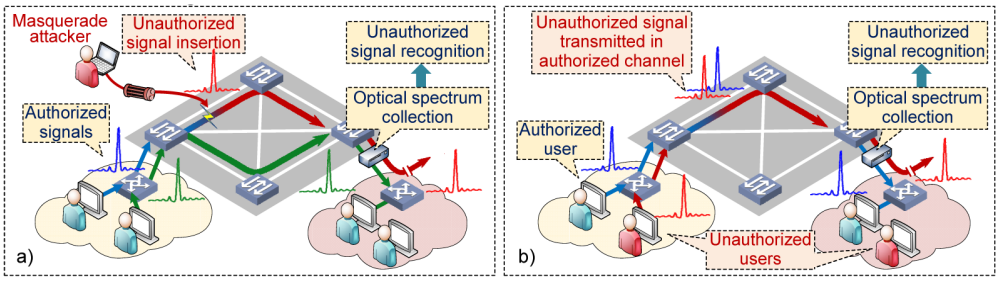

Fig. 1. (a) Masquerade attacker gain access to network incognito and insert signals; (b). Unauthorized users transmit unauthorized signals in authorized channels

Figure 1 shows two possible attack methods. Figure 1(a) describes the masquerade attack. The attacker intercepts and blocks the original authorized signal in the fiber link, and at the same time, gain access to the network as an authorized user (or impersonating one) or by otherwise breaching into the network [2,5]. Figure 1(b) describes the unauthorized users illegally gain access to the optical nodes and transmit unauthorized signals in authorized channels [5,12,13].

Detection and localization of physical layer attacks are difficult to realize [1,5], due to the transparency of optical networks, because the attacks can propagate to many different nodes or links in the optical network, making the localization of optical attacks complicated. In transparent optical networks, the traffics are not processed electronically and attacks cannot be noticed at the information level; thus, detection of attacks relies on accurate and reliable physical layer monitoring, which also increases the difficulty. On the other hand, masquerade attackers can gain access to the network by sending unauthorized signals in lieu of legitimate users, instead of or in addition to the authorized signals, making it difficult to detect the attacks with conventional methods. Secure communications over a public host channel is proposed [14], which reveals the possibility of malicious users exploiting these channels to attack. Thus, detection and localization methods of signal insertion attacks (masquerade attacks, and power jamming) are requisite. Optical time domain reflectometer (OTDR) is a conventional method for attack detection [15]. However, due to the OTDR dead-zone (event dead-zone and attenuation dead-zone) effects, attackers can evade OTDR detection. Some localization and detection mechanisms are proposed to solve the problem with monitored physical layers parameters [16–19]. Optical layer monitoring provides operators with the ability to perceive, and supports network optimization via analytics [20], because the network status is hidden in the optical layer parameters. Thus, it is feasible in principle to detect physical layer attacks by monitoring, and analyzing optical layer parameters and performance.

Facing the physical layer signal insertion attacks, we hope to be able to recognize the attack signals from the physical layer. By monitoring the physical layer parameters (in this paper, we utilize optical spectrum), a database of physical layer impairment and parameters is formed, and finally perform analysis of the network security environment to realize unauthorized signal detection and recognition.

Optics EXPRESS

1.3. Motivation of optical spectrum analysis for attack detection

Due to the difficulty of detection physical layer attacks with conventional methods, analyzing optical layer parameters and performance with more useful information, is a potential method. The optical spectrum (OS) is a significant indicator of optical signal performance [21–23], which contains accurate information pertaining to the laser, modulator, etc. (collectively referred to as source features) and the transmission and switching environments. Each of the different signal source features is unique. When a signal insertion attack occurs, an abnormal unauthorized signal is inserted to the network. Thus, detection of the attacks can be achieved if the unauthorized signal is recognized. By analyzing the spectrum, the feature information of the light source can be used to identify the optical signal, thereby realizing the recognition of the authorized signal/unauthorized signal, as shown in Fig. 1.

However, this information is hard to extract without proper analytical tools. Conventional signal processing methods like correlation coefficient is attempted to recognize OS [17], however the recognition accuracy is only 64.59%, which is not satisfying for security issues.

Machine learning (ML) and artificial intelligence technologies have become a hot topic in optical networking [24,25]. ML methods have been extensively studied in analyzing and predicting the quality of transmitted signals in optical networks [26,27] (such as crosstalk [28], phase noise [29], optical signal-to-noise ratio (OSNR) [30]) and QoT prediction of unestablished lightpaths [31], as well as optimization channel reorganization of EDFAs [32], module format classification [33], and mitigation of linear and nonlinear distortions [34], ML-based soft-failure detection and identification [35]. ML-based monitoring database and network abstraction are proposed in [36,37], which assist optical network management and planning. ML technology provides a powerful statistical method to solve optical network problems, especially to analyze and monitor the physical layer parameters of optical networks.

Therefore, OS analysis can be achieved with ML methods. An ideal reference OS retrieval method is proposed in [23], in which high resolution OS are used for feature extraction and classification of different types of optical signals. A filter failure detection and identification method using support vector machine (SVM) for OS analysis is proposed in [25]. We proposed optical spectrum feature analysis (OSFA) methods for optical network security using an SVM and a one-dimensional convolutional neural network (1D-CNN), to identify unauthorized light sources and rogue ONUs, in [17] and [38], respectively, which have already gained attention as available solutions to attack detection and accurate localization. These works demonstrate the processing ability of ML methods for OS analysis and recognition problems.

While, during the life-cycle of the network, several factors would change the OS features. Thus, OS samples should be periodically collected and the OS database should be updated to retrain the ML model, to adapt to different network environments and hardware conditions.

1.4. Contributions

We summarized the contributions of this paper: We present a comprehensive investigation of optical spectrum feature analysis and unauthorized signal recognition using ML methods enhancing optical network security. (1) We theoretically analyzed the influencing factors of optical spectrum features and set up simulations to verify the theoretical derivation. We quantitatively analyzed OS features and extracted OS features using principal component analysis (PCA). (2) Two machine-learning-based methods (SVM and 1D-CNN) for OS recognition and unauthorized signal detection are proposed and verified with the OSs collected from the simulations. We investigated the OS recognition accuracy, false alarm rate (FAR), and miss alarm rate (MAR), to evaluate the performance of the methods. (3) An experimental demonstration is set up, and actual OSs are collected from eight commercial SFP modules. Experimental results verified the ability of SVM and 1D-CNN to recognize OS features and detect unauthorized signals, with 98.54% and 100% OS recognition accuracies

Optics EXPRESS

respectively, proving the method to be a promising solution of attack detection and recognition for optical network security.

The paper is organized as follows: In Section 2, we mainly discuss the spectral features and theoretically analyze the non-ideal factors of transmitters influencing the OS. We built a simulation to collect OSs with different transmitters and quantitatively analyze the spectral features. Section 3 proposes SVM-based and 1D-CNN-based algorithms for OSFA and unauthorized signal detection. The methods are verified with simulated OS data. In, Section 4 we built an experimental platform to acquire OS samples and verify the validity of the ML methods for actual OS recognition. Section 5 summarizes and concludes the present work.

2. Optical spectrum features and influencing factors

In this section, theoretical analysis of the influencing factors of OS features, simulation setup of OS generation, OS feature evaluation and extraction with PCA are presented. The simulation results support the theoretical derivation.

The optical spectrum carries a lot of information about the transmitter and the transmission environment. The transmitter will have different features due to differing lasers, modulators, signal sources, modulation formats, signal rates, etc. In addition, non-ideal factors of the transmitter also reflect some subtle features of the spectrum, collectively referred to herein as light source features. The transmitters deployed in optical networks have distinctive features due to non-ideal characteristics, resulting in optical spectra with different features. Similar to fingerprints identification, it is possible to use these OS features to recognize light sources.

2.1. Theoretical analysis of optical spectrum features

The optical signal is affected by the laser, signal source, modulator, amplifier, modulation format, and signal rate of the transmitter, which will be reflected in the spectral features. To simplify the problem, the combination of laser, data source, and modulator is now collectively considered as the light source. The OS features of different modulation formats and bitrates are simulated in the literature [23]. However, even if the same batch and the same type of transmitter equipment and devices emit optical signals of the same rate and modulation mode, due to their unique non-ideal features, there are differences in their respective spectral features which can be used to identify a particular transmitter. When an attack signal occurs in the network, the unauthorized light source can be screened because it has its own unique features.

It is assumed that a transmitter consists of a laser, a signal source, and a Mach-Zehnder Modulator (MZM) that modulates an on-off keying (OOK) signal. Assuming that the output light field of the laser is $E_{in}(t)$, the output light field modulated by the push-pull MZM is:

$$E_{\text{out}}(t) = E_{\text{in}}(t)\cos\frac{\pi}{2V_\pi}[V_1(t) - V_2(t)]e^{j\frac{\pi}{2V_\pi}V_1(t)+V_2(t)} \tag{1}$$

$V_1(t)$ and $V_2(t)$ are the signals loaded on the two arms of MZM, and V_π is the half-wave voltage of MZM. Assuming $V_1(t) = V_{\text{bias1}} + V_{\text{drive}}u_1(t)$, $V_2(t) = V_{\text{bias2}} + V_{\text{drive}}u_2(t)$, for the push-pull MZM, the two-arm loaded signal $u_1(t) = -u_2(t)$, $u_1(t)$ consists of the signal portion $g(t)$ and the noise portion $N(t)$. Let $V_{\text{bias}} = V_{\text{bias1}} - V_{\text{bias2}}$, which is the bias voltage of the MZM. The output light field of MZM can be derived as:

$$E_{\text{out}}(t) = E_{\text{in}}(t)\cos\frac{\pi}{2V_\pi}\left\{V_{\text{bias}} + 2V_{\text{drive}}\left[g(t) + N(t)\right]\right\}e^{j\frac{\pi}{2V_\pi}(V_{\text{bias1}}+V_{\text{bias2}})} \tag{2}$$

Thus, the output intensity of the MZM is:

Optics EXPRESS

$$P_{\text{out}}(t) = E_{\text{out}}(t)E_{\text{out}}^{*}(t) = P_{\text{in}}(t)\cos^2\frac{\pi}{2V_{\pi}}\left\{V_{\text{bias}} + 2V_{\text{drive}}\left[g(t)+N(t)\right]\right\} \tag{3}$$

Therefore, it can be inferred that the spectrum of the output of the optical transmitter composed of the MZM is:

$$P_{\text{out}}(\omega) = \frac{1}{2}\underbrace{P_{\text{in}}(\omega)}_{(a)} + \frac{1}{4\pi}\underbrace{P_{\text{in}}(\omega)}_{(a)} * \mathcal{F}\left\{\cos\frac{\pi}{V_{\pi}}\left\{\underbrace{V_{\text{bias}}}_{(b)} + 2\underbrace{V_{\text{drive}}}_{(c)}\left[\underbrace{g(t)}_{(d)} + \underbrace{N(t)}_{(e)}\right]\right\}\right\} \tag{4}$$

In the case of ideal OOK signal modulation: $P_{\text{in}}(\omega)=2\pi P_0\cdot\delta(\omega\text{-}\omega_0)$, the bias voltage $V_{\text{bias}}=V_{\pi}$, $V_{\text{drive}}=V_{\pi}/2$, and noise are partially ignored ($N(t)=0$). However, according to Eq. (4), in the non-ideal case, the following aspects will affect the spectrum of the output of the optical transmitter:

(a). The optical power and linewidth of the actual laser's $P_{\text{in}}(\omega)$ introduces non-ideal features;

(b). Mismatch of the bias voltages V_{bias} and V_{π} introduces non-ideal features to the spectrum;

(c). Mismatch of the driving voltages V_{drive} and $V_{\pi}/2$ introduces non-ideal features to the spectrum;

(d). Changes of the source signal $g(t)$ (e.g. changes in transmission rate) affect spectrum features;

(e). Noise $N(t)$ caused by amplifiers or sources introduces non-ideal features to the spectrum.

In addition to the above factors, the non-ideal factors of the transmitter include insufficient bandwidth, and extinction ratio degeneration of modulators and the like. The difference in non-ideal factors or the combination of different factors, results in a spectrum with certain features. Using these spectral features to identify transmitters, requires: 1) that the OS features of a given transmitter should be similar; 2) the OS features of different transmitters should be distinguishable and recognizable.

2.2. Simulation configuration and optical spectrum generation

In order to verify the “similarity” and “recognizability” of the spectral features resulting from the influencing factors of a transmitter, we set up simulations using the commercial software VPI TransmissionMaker V.7.6, to generate and collect OSs from different transmitters. In the simulation, we use the common modules, such as continuous wave (CW) laser, differential MZM (DMZM), etc., to build the simulation environment. According to Eq. (4), seven groups of different transmitter parameters are set, the first group is the reference transmitter, and the other six groups are control groups. As shown in Fig. 2(b), the transmitter generates a random signal from a pseudo-random sequence generator (PRBS), which is driven by the code driver into a sequence of non-return-to-zero (NRZ) codes. A continuous wave laser generates DC light, and the NRZ sequence is input to the arms of the differential MZM modulator. The DMZM operates in a push-pull state to modulate the light with the NRZ signal. Two channels of a DC voltage source control the bias voltage of the DMZM’s arms independently.

Optics EXPRESS

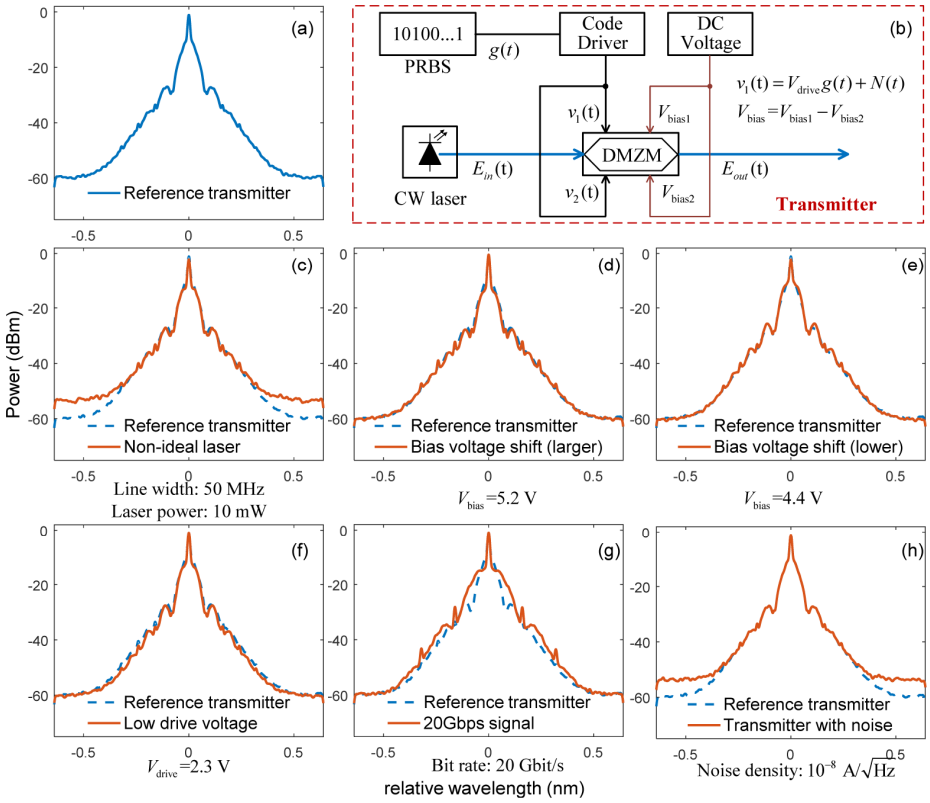

Fig. 2. (a) OS of reference transmitter; (b) Internal structure of the transmitters; (c–h) OS of simulated transmitters, each with one parameter differing from the reference as indicated.

The parameters of the reference transmitter are set to: V_{π}=5 V , laser line width = 10 MHz, laser output power = 10 mW, V_{bias} = 5 V, V_{drive} = 2.5 V. The bit rate is 10 Gbit/s, and the noise spectral density is 10^{-12} A/$\sqrt{\text{Hz}}$, as shown in Fig. 2(a). In order to investigate the influence of different parameters on the OS features, for each of the six compared transmitters, only one parameter is different to the reference transmitter as shown below the spectra in Figs. 2(c)–2(h).

The spectral samples are sampled with 1.28 nm resolution around the center wavelength of the laser (192.1 THz, 1560.1 nm), and the number of sampling points is 16382. Therefore, the spectral sample can be regarded as a feature vector of 16382 dimensions, and the resolution of the optical spectrum analyzer (OSA) is set to 0.01 nm in the simulation. By comparing the spectra of the six control groups, it can be found that the differences between the spectra and the reference spectrum are somewhat visible to the eye.

However, during transmission, the features of the transmitter itself may be affected by factors such as channel attenuation, dispersion, ASE noise introduced by EDFA, and filter bandwidth limitations. In order to verify the changes in the spectral features of the transmitter under different transmission conditions, we built six different transmission environments of differing lengths in the VPI simulation, with differing transmission conditions of, standard single mode fiber (SSMF) length, prescence or absence and number of EDFAs, length of dispersion of compensating fiber (DCF), and prescence or absence of optical band pass filter (OBPF). The SSMF is set to different lengths, so that it will produce different attenuations. In the case of transmission distances above 20 km, an EDFA is added to compensate for the increased attenuation. At a distance of 50 km or more, EDFAs are deployed at the transmitting and receiving ends, and an OBPF is included for filtering. Further, since the

Optics EXPRESS

dispersion is proportional to the transmission distance, DCF of the appropriate length is introduced to compensate for the dispersion, when the transmission distance is 50 km or more. The details of the specific transmission conditions are shown in Fig. 3.

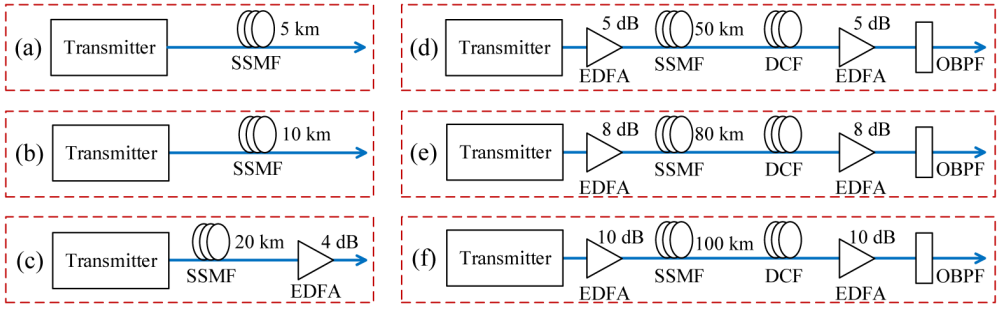

Fig. 3. The six groups of different transmission distances and conditions in the simulation.

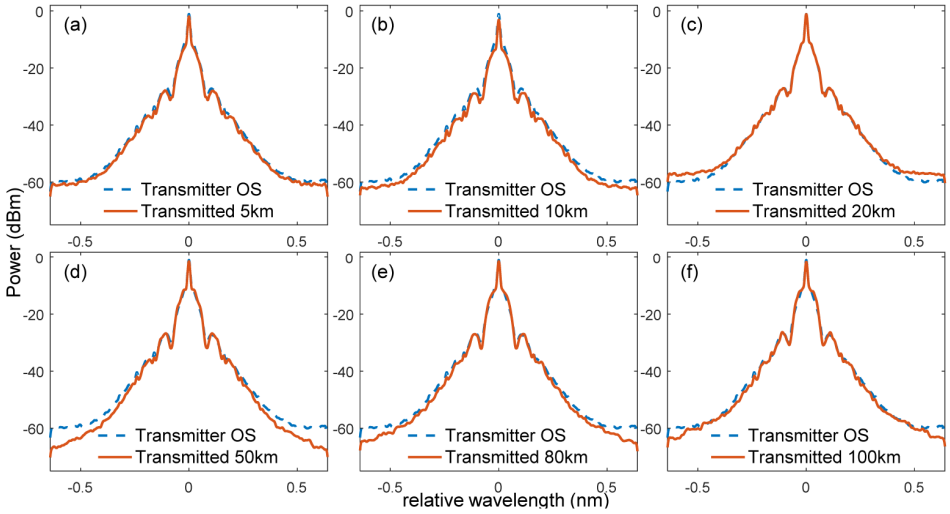

Fig. 4. Optical spectra of the reference transmitter under the six conditions shown in Fig. 3.

After each transmitter's signal is transmitted through these six transmission paths, we collect the spectrum of the signal at the receiving end. In Fig. 4, the spectrum of the reference transmitter's optical signal after six paths is shown.

In the simulation, we collected 350 output spectra produced by the above seven optical transmitters, with 50 samples for each. Meanwhile, 2100 OS samples of seven transmitters under the six transmission conditions were collected, with 50 samples collected for each transmitter in each transmission condition. In total, 2450 OS samples (350 output ones and 2100 tranmitted ones) were collected. These OS samples were used for feature analysis, as described in the next section.

2.3. Optical spectrum feature evaluation and extraction

In order to analyze the features of the spectra acquired in the simulation efficiently and quantitatively, we first evaluate the features of the output spectra. In this problem, the spectra of transmitters that are the same should belong to the same category, whereas, the spectra of transmitters that are different should belong to a different category. In the ideal case, the distribution of OS samples of the same categories should be concentrated in the high-dimensional feature space, while the distribution of OS samples of different categories should be scattered. In order to measure the class separability of OS samples, we introduce the

Optics EXPRESS

concept of the within-class and between-class scatter matrices, S_w and S_b, respectively [39]. S_w mainly measures the within-class distance of a set of sample points, which is the distance from each sample point to the center point of the category. S_b mainly measures the distance between classes of different sets of sample points, which is the distance between the center points between the classes. S_w and S_b are defined as Eq. (5) and Eq. (6), where μ is the mean of all samples, μ_i is the mean of the sample of class i, and x_j is the sample $x_j = (x_{j1}, x_{j2}, ..., x_{jd})$ which belong to set $\{X_i\}$.

$$S_w = \sum_{i=1}^{N} S_{wi} = \sum_{i=1}^{N} \sum_{x_j \in X_i} (x_j - \mu_i)(x_j - \mu_i)^T \tag{5}$$

$$S_b = \sum_{i=1}^{N} N_i (\mu_i - \mu)(\mu_i - \mu)^T \tag{6}$$

To quantify the class separability measure, we define the between/within-class distance ratio (BWCDR), as shown in Eq. (7), where matrix H is a transformation such that the projection of the sample point in this direction has the features of large distance between classes and small distance within the class, and which maximizes the *BWCDR*. $tr(\bullet)$ is the trace of the matrix. The larger the *BWCDR* is, the larger the difference between the spectral sample classes is, the more similar the features within the class are, and the more separable the sample is.

$$BWCDR = \frac{tr(H^T S_B H)}{tr(H^T S_W H)} \tag{7}$$

We analyzed the spectra of the direct output of the transmitter and at the output of the six transmission configurations shown in Fig. 3, and calculated the BWCDR of each group. The result is shown in Fig. 5. It can be seen that with increased transmission distance and complication of the configuration, the BWCDR of the spectral samples decreases, indicating that the separability of the samples is weakened, because the OS features are degraded.

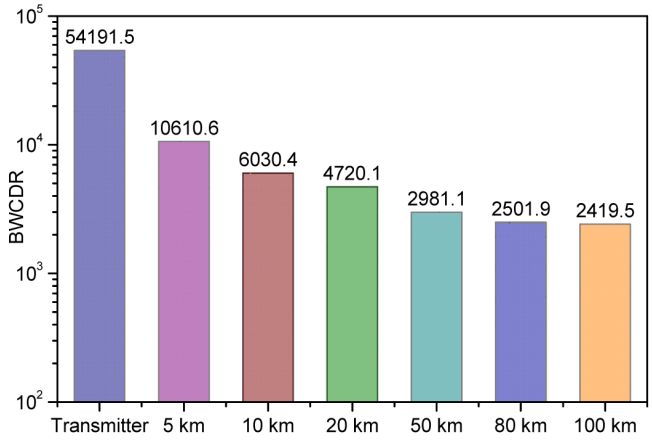

Fig. 5. BWCDR calculated from OS samples with different transmission conditions corresponding to those shown in Fig. 3.

To extract the features carried in the high-dimensional spectral samples, we used principal component analysis (PCA) [40] to process the spectral samples collected in the simulation. As shown in Eq. (8), PCA reduces the dimension of the original OS data x_i by an orthogonal linear transformation W and converts it into a set of linearly uncorrelated principal components (PCs) z_i, so that they retain as much original variable information as possible, to

Optics EXPRESS

reveal the correlation and difference between OSs by a few PCs. $x_i \in X_d$ is a vector of dimension d, $z_i \in Z_k$ is a vector of dimension k, and W is a matrix of $d \times k$.

$$z_i = (z_{i1}, z_{i2}, ..., z_{ik}) = x_i \times W^T \tag{8}$$

We define the contribution of the first k PCs to the spectral samples as C_k, where λ_i is the eigenvalue of matrix $X_d{}^T X_d$:

$$C_k = \sum_i^k \lambda_i \Big/ \sum_j^d \lambda_j \tag{9}$$

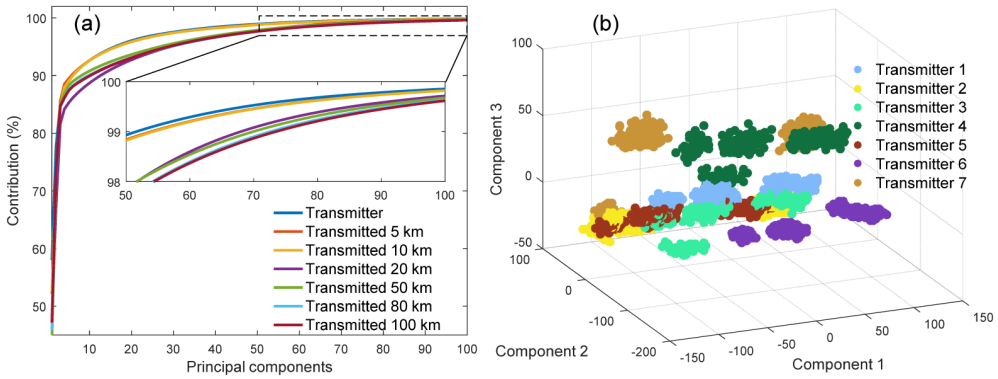

Fig. 6. (a) Contributions of PCs as the number of PCs grows under different transmission conditions. (b) Distribution of 2450 OS samples in the first 3 PC spaces (colors represent different transmitters; different transmission conditions form clusters of points).

According to the definition above, we performed PCA on the OSs of the transmitters and the OSs of signals transmitted under six conditions, then calculated the contribution C_k of each group under a different number of PCs, as shown in Fig. 6(a). According to the curves in the figure, we can see that after the PCA of 7 groups of OSs, the contributions of the first 15 PCs reach more than 90% and the first 80 PCs reach more than 99%, which thus, degrading a 16384-dimensional OS sample to 80 dimensions can characterize 99% of its features. Additionally, as the transmission distance increases and the transmission conditions are complicated, the contribution of the PCs to the sample features is lower, thus, a given features information requires more dimensions, which means that the OS features of the transmitter are weakened

Figure 6(b) shows the distribution of 2450 OS samples in the feature space composed of the first three PCs. Color indicates the transmitter emitting the spectrum. Clusters of the same color represent differences in features arising from differing transmission conditions. It can be seen that the distribution of the OSs of the same transmitter in the feature space is relatively close, indicating that with the influence of the transmission environment, the spectrum still includes the OS features originating from the transmitter. While, some samples for different transmitter are overlapped in Fig. 6(b), but in higher dimension (with more principal components), more features can be used to distinguish the differences between different categories of samples, thus to achieve OS recognition (The recognition result will be given in Section 3).

3. Machine learning algorithms for optical spectrum recognition

In this section, we propose two supervised ML methods based on SVM and 1D-CNN to learn the features of the transmitter carried by the OS and recognize the inserted unauthorized signals.

In the OSFA problem, the OSs of different transmitters are grouped and given the same label, and a large amount of labeled data is used to train the ML model. After feature

extraction using PCA, the features of the first k PCs are selected so that the model can learn the features of the spectrum. In the recognition procedure, the ML model recognizes the learned features and matches the test OS to a certain class of trained OSs, so that the spectral recognition problem becomes a classification problem. The untrained unauthorized OSs cannot be classified as an authorized class, thus they are recognized as unauthorized signals.

During the life-cycle of the network, several factors would affect the features of the transmitters as well as the output OS features. Over time, the optical transmitter will be aging. For examples, the modulator ages, causing the DC-drift influencing the phase to drift; the laser ages, causing the center wavelength to drift and output power deterioration. The features of the output optical signal of the transmitter will change over time. To solve the problem, in our proposed method, the OS database can be updated periodically (e.g. weekly or monthly) or after the network being upgraded. Then the ML model can be retrained with latest OS samples to adapt to changes in the network environment and hardware devices. Therefore, even if the OS collected from the actual networks would change during the life-cycle of the network, the proposed methods can still realize the recognition of the OS features.

3.1. Support vector machine for OSFA

We propose a spectral feature recognition method based on support vector machine (SVM). When the SVM deals with linearly inseparable data, the data is transformed to a high-dimensional space through a nonlinear transformation $\phi(x)$ (this nonlinear transformation is called a kernel function), and a "classified hyperplane" in high-dimensional space is found to separate data belonging to different categories in different regions.

$$f(x) = \omega^T \phi(x) + b \tag{10}$$

The training process of SVM is to optimize the appropriate hyperparameters to determine the hyperplane and maximize the samples to the hyperplane margin. The hyperparameters are determined to accommodate the different data features in the training set to form an SVM model. The spectral feature recognition problem is a multi-classification problem. For multi-classification problems, the problem can be solved with K binary classifiers, using error correcting output codes (ECOC). ECOC encodes the results of the K binary classifiers and expresses the results of the C-class classification problem.

The main research goal of our study of spectral feature analysis is to detect attack signals in the network and determine whether the received signal is an unauthorized attack signal. Because unauthorized attack signals cannot be trained and the features cannot be learnt, the SVM model cannot accurately classify attack signals. Thus, we use the loss function of the SVM model in the classification process. The loss function is used to define how accurate the classification is and quantify the error introduced by the classification. For unauthorized OSs, the loss function is larger than for authorized OSs, which can be used as a judgement criterion for unauthorized OSs. The loss function is shown in Eq. (11).

$$Loss(i,c) = \frac{\sum_k g_{kic}}{K} = \frac{\sum_k \left(\max(0, 1 - y_{kic} s_{kic})\right)}{2K} \tag{11}$$

$Loss(i,c)$ is the average binary loss for sample x_i, corresponding to class c and binary classifier k, where y_{kic} is a class label for a particular binary classifier (in the set {-1,1,0}), s_{kic} is the classifier score for sample x_i, and g_{kic} is the hinge loss function commonly used in SVM [41]. $Loss(i,c)$ represents the classification score that determines how well the classifier classifies sample x_i into class c. The smaller the $Loss(i,c)$, the more accurate the classification. $Loss_i = \min_c (Loss(i,c))$ can be used to determine which category the classification result of the OS sample x_i belongs to. When the spectrum x_i is an authorized

light source, it can be accurately classified as a certain light source. If the loss minimum value $Loss_i$ is smaller than the threshold *Th*, the classification is sufficiently accurate, and the class *c* corresponding to the minimum value is selected as the classification result. If the spectrum x_i is unauthorized, the model cannot accurately classify it, which means $Loss_i$ is larger than the threshold *Th,* and the spectrum x_i will be identified as an unauthorized spectrum:

$$Label_i = \begin{cases} SVM(x_i) & \min\limits_c (Loss(i,c)) < Th \\ unauthorized & \min\limits_c (Loss(i,c)) > Th \end{cases} \tag{12}$$

Based on the above method, we analyzed the 2450 spectral samples collected in the VPI simulation. Firstly, the SVM is used to classify the spectrum to verify the identifiability of the spectral samples. Secondly, the signal of Transmitter 2 is regarded as an attack signal, and it is verified whether the above method can recognize the unauthorized signal.

In the first experiment, in order to verify the identifiability of the spectral samples, we evaluated the spectral recognition accuracy under different numbers of PCs and the number of different training samples. We randomly divided all samples into training and test sets in a number of proportions and performed normalized preprocessing. After PCA processing, 3–80 PCs are selected to extract OS features, and polynomial kernels are used to train the SVM model. In Fig. 8(a), the horizontal axis is the main component format from 3 to 80 with an interval of 2, and the vertical axis is the recognition accuracy of the trained SVM model on the entire test set. Different curves represent different proportions of training/test samples. It can be seen from the accuracy curve that when the number of PCs is 13, the recognition accuracy is the highest. If the number of PCs is less than 13, the spectral features cannot be fully expressed. If the number of PCs is larger than 13, the spectral features are over expressed and unnecessary noise components are introduced to over-fit the training process. According to the different curves, it can be seen that the higher the recognition ratio of the training set, the higher the recognition accuracy of the test set. This is because with sufficient training samples the SVM model can fully learn the spectral features of a transmitter. The features of the spectrum are incorporated into the model so that it will more accurately match the actual spectrum features.

In the second experiment, we simulated an attack in the network and selected a transmitter (Transmitter 2) as a signal for an unauthorized attack. The attack signal does not participate in any training process. During the test process, the unauthorized attack signal is included in the authorized normal signal to participate in the test to verify whether the SVM model can distinguish between authorized and unauthorized signals. The method of determining the unauthorized signal by using the loss function requires that the loss threshold is set. When the loss value is less than the threshold, the sample is judged to be an unauthorized signal. However, when the loss function is used to judge whether the sample is unauthorized, an error is introduced, resulting in the authorized sample being recognized as an unauthorized one, and vice versa. Therefore, we define two parameters to measure the accuracy of authorized and unauthorized signal determination, as shown in Fig. 7: the miss alarm rate (MAR) and false alarm rate (FAR). The MAR is the proportion of unidentified unauthorized signals in the unauthorized samples, and the FAR is the proportion of the authorized sample that is identified as unauthorized. The closer the values of the MAR and FAR are to 0, the more accurately the model recognizes the unauthorized signal.

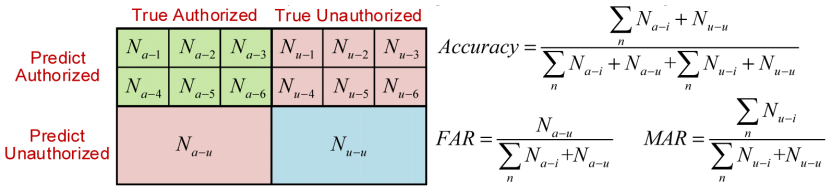

Fig. 7. True samples, predicted samples, and the calculation of accuracy by FAR and MAR.

Optics EXPRESS

To assess the effect of the selection of different thresholds on the recognition accuracy, we set the threshold to a value of 0.005–0.025 in increments of 0.001. The horizontal axis of Figs. 8(b)–8(d) is the threshold, and the vertical axis is the number of selected PCs. Figure 8(b) shows the value of the recognition accuracy. When the threshold is set to 0.02 and the number of PCs is 17, the overall accuracy is 93.37%. At this time, the MAR is 3.73% and the FAR is 23.36%. Figures 8(c) and 8(d) show the MAR and the FAR as a function of the threshold value and number of PCs. It can be seen that when the threshold is set to 0.005 and the number of PCs is 37, the MAR is the minimized value of 0.47%. When the threshold is set to 0.025 and the number of PCs is 5, the FAR is the minimized value of 1.89%. Due to the features of the spectral sample data set and the processing capability of the SVM method, the overall recognition accuracy, FAR, and MAR cannot be simultaneously optimized. Therefore, in practical applications, a compromise is required. In attack detection issues, the main goal is to recognize unauthorized signals and to eliminate the influence of attacks, thus, in order to prevent omission of the suspicious attack signal, MAR should be optimized as small as possible, and on this basis, the other two indicators should be optimized. In the figure, the area with an accuracy larger than 90%, and the areas with FAR and MAR smaller than 10% are marked. The loss threshold *Th* and the number of PCs can be selected from the overlap of these three regions to select an appropriate threshold and number of PCs.

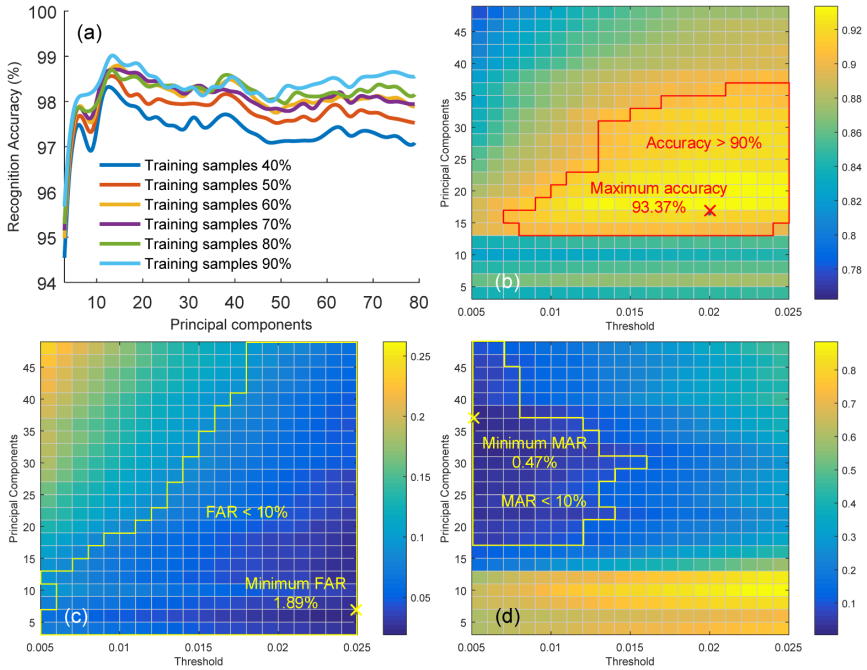

Fig. 8. Recognition results with the SVM. (a) Recognition accuracy with different training samples as the number of PCs grows in the first experiment; (b–d) Accuracy, FAR, and MAR of authorized/unauthorized OS recognition for different number of PCs and different threshold of loss in the second experiment.

3.2. One-dimensional convolutional neural network (1D-CNN) for OSFA

We also propose a spectral feature recognition method based on a 1D-CNN. In literature [38], we use 1D-CNN to realize the analysis of overlapped OS features and the identification of rogue ONUs. The neural network consists of a large number of neuron nodes connected to each other. The operation of each node can be represented as $y = f(W \cdot x + b)$, where x is the input of the neuron node, y is the output, W and b determine a linear transformation, and $f(x)$ is called the activation function, which is capable of introducing nonlinearity into the

Optics EXPRESS

calculation of neuron nodes (usually use ReLU function in CNN [42]). CNNs include at least a convolutional layer and a pooling layer to extract local features of the data samples. In the convolutional layer, a convolution kernel (also known as a convolution template) is used. The input layer features are convoluted to process local features in the input data. The pooling layer, also known as the down sampling layer, uses the mean (mean pooling) or maximal (max pooling) value of the features in a region to represent the entire region feature, thereby reducing the complexity of the model. CNN has the advantage of reducing the computational complexity, and has the features of fast convergence and avoiding over-fitting.

In the OSFA problem, since the OS sample is a one-dimensional data vector, we use a 1D-CNN algorithm to process OS features. The input layer, convolutional layer, pooling layer, and output layer are all 1D. Before using 1D-CNN for training or light source recognition, the spectral data needs to be centralized and standardized for preprocessing, to make the loss function less sensitive to neural network parameters, and increase the convergence rate. Next, the identity of the optical transmitter is used as a training label. For an *n*-class problem, the CNN's output layer outputs the result as one-hot code, which is a vector representation of length *n*. When the element *i* of the vector is 1 and the other elements are 0, it represents a CNN. The classification result points to the class *i*.

In our research, the main goal is to detect attack signals in the network and determine whether the received signal is an unauthorized attack signal. Therefore, similarly to the SVM method, we also use the loss value of the CNN output layer to judge the accuracy of the sample classification and quantify the error introduced in the classification. A commonly used loss function for neural networks is the categorical cross entropy:

$$loss_{\mathrm{CCE}} = -\frac{1}{N}\sum_{i=1}^{N}\hat{y}_i \log(\hat{y}_i), \sum_{i}^{N}\hat{y}_i = 1 \tag{13}$$

Where $\hat{y}_i$ is the output of the softmax function of the neural network output layer, representing the probability that a sample is classified into the *i*-th class, and the sum of all $\hat{y}_i$ is equal to 1. The smaller the absolute value of $loss_{\mathrm{CCE}}$, the more accurate the classification. If $loss_{\mathrm{CCE}}$ is less than the threshold *Th*, then the classification is sufficiently accurate that it can be accurately classified as a specific type of light source.

If the spectrum x_i is unauthorized, the 1D-CNN model cannot accurately classify it, thus, the loss function $loss_{\mathrm{CCE}}$ is larger than the threshold *Th*, and the spectrum x_i can be recognized as an unauthorized spectrum.

$$Label_i = \begin{cases} CNN(x_i) & loss_{CCEi} < Th \\ unauthorized & loss_{CCEi} > Th \end{cases} \tag{14}$$

To verify the performance of the 1D-CNN-based spectral recognition method, we analyzed the 2450 spectral samples collected in the VPI simulation. Similarly, to the purpose and conditions of the SVM test, two sets of test experiments were performed.

In the first experiment, in order to verify the identifiability of the spectral samples, we evaluated the spectral recognition accuracy under different numbers of PCs and of different training samples. According to Fig. 9(a), the CNN model is more sensitive to the number of training samples compared with the SVM method (the CNN requires a larger training sample to learn the features than the SVM). For 70–90% training samples, the recognition accuracy is relatively high. The CNN's recognition accuracy fluctuates as a function of number of PCs more obviously, however, the overall trend is the same as that of the SVM. There is also an optimal PC number. When the number of PCs increases, the over-sufficient expression feature introduces unnecessary noise by the components, making the training process over-fitting.

Optics EXPRESS

In the second experiment of unauthorized signal detection, the signal of Transmitter 2 is regarded as an unauthorized signal, so that OSs of Transmitter 2 are not trained but are tested. The CNN model is trained and learned the features of authorized OSs. Figures 9(b)–9(d) show the recognition accuracy, FAR and MAR under different number of PCs and different thresholds. In the figure, the area with the accuracy rate larger than 90%, and the areas of the FAR and MAR smaller than 10% are marked. When the number of PCs is 5 and the threshold is 1.2, the highest recognition accuracy is 99.86%. At this time, FAR = 0.07%, MAR = 0. Compared with the SVM results in Fig. 8, the respective areas are larger, indicating that using CNN to identify the spectrum allows for a wider threshold and number of PCs to achieve higher accuracy, and lower FAR, and MAR of OS recognition. Among the FAR and MAR results, there are a large number of parameter values that can be equal to 0 for both FAR and MAR, which is the ideal case.

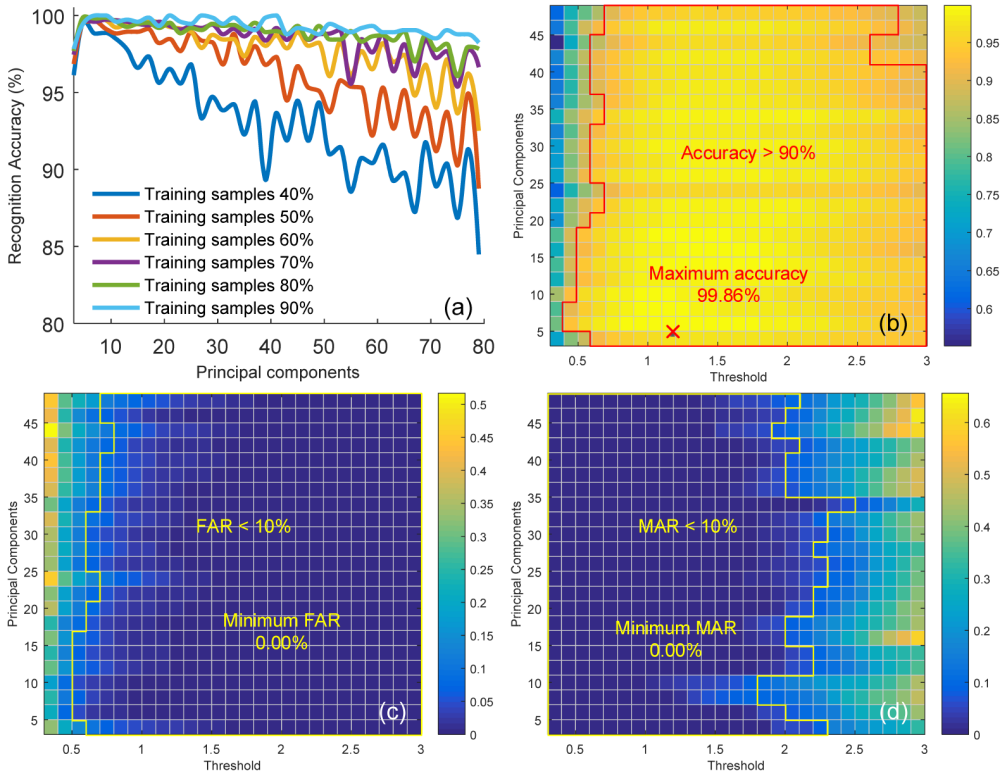

Fig. 9. Recognition results with 1D-CNN: (a) Recognition accuracy with different training samples as the number of PCs grows in the first experiment; (b–d) Accuracy, FAR, and MAR of authorized/unauthorized OS recognition with different number of PCs and different threshold of loss in the second experiment.

After using the SVM and 1D-CNN methods to recognize the OSs, it can be seen that SVM has the advantages of requiring smaller amount of training data, which meets the needs of OS feature recognition and unauthorized signal detection to some extent. In contrast, the disadvantage of 1D-CNN is that the training process requires more data, but it is not sensitive to the choice of the number of PCs or the threshold. 1D-CNN can achieve OS feature recognition and unauthorized signal detection with accuracy approaching 100%, and MAR and FAR close to 0, for various parameter combinations, outperforming SVM.

Optics EXPRESS

4. Experiment and optical spectrum recognition results

In order to verify the ability of the machine learning-based spectral feature recognition method to recognize an actual signal, we performed experiments using commercial small form-factor pluggable (SFP) modules as signal sources to collect OSs after transmission under different conditions. The SVM and 1D-CNN methods are used to process OS data and realize unauthorized signal detection, and the experimental results are analyzed.

4.1. Experimental setup

In the experiment, a full-mesh 4-node prototype network is established with 8 commercial SFPs, 4 erbium-doped optical fiber amplifiers (EDFA), 6 optical fibers with different lengths, and several 2 × 2 magneto-optic switches controlled by FPGAs for selecting lightpaths, as shown in Fig. 10. 6 different lightpaths are chosen in the experiment for transmission with different transmission conditions.

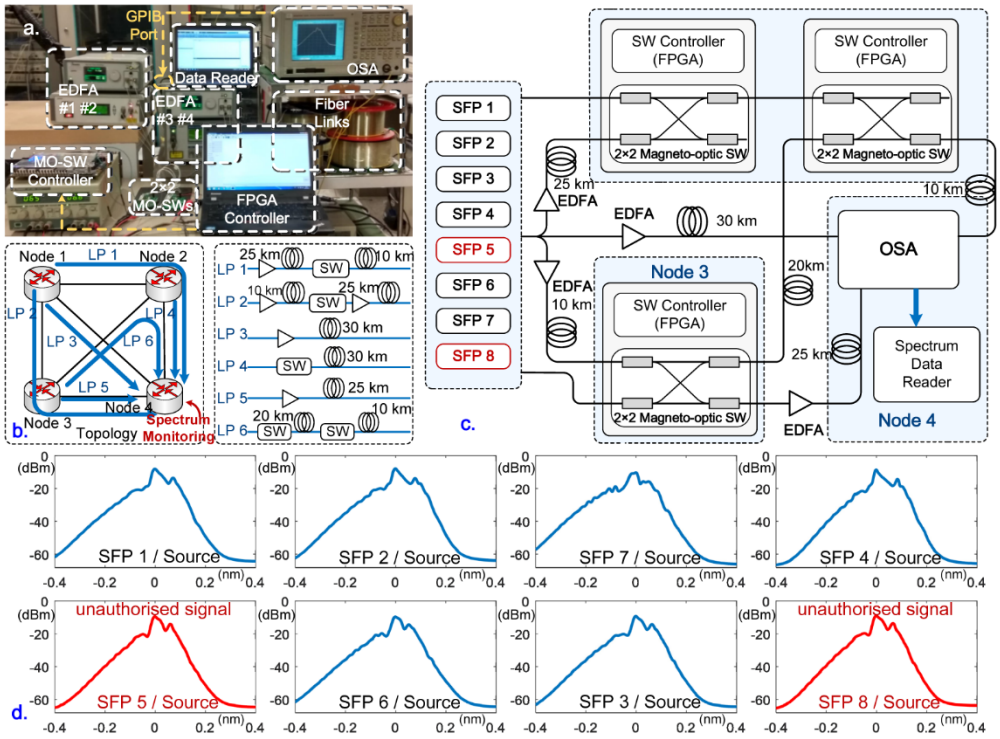

Fig. 10. (a) and (c) experimental setup, (b) experiment topology and transmission conditions, (d) optical spectra of SFP 1–8 light sources.

4800 OSs of 8 SFP transmitted along 6 lightpaths are collected by an optical spectrum analyzer (OSA). The OS are 0.8-nm wide with 10 pm spectral resolution, some of which are shown in Fig. 10(d).

In the experiment, according to the settings of the OSA, the acquired spectral data may be considered to be 501 dimensional. The SFP used was a 1 Gbps bandwidth module with OOK modulation. Since commercial SFPs generally adopts the direct modulation method, the spectrum obtained by the OSA and the VPI simulation are significantly different. It can be seen from the observation of Fig. 10(d) that the spectral shapes are generally similar in the 0.8-nm range we have collected, but there are some subtle differences in the dimensions. We consider these differences to be the features of the spectra emitted by the SFP. After passing through 6 paths, a certain degree of change will be introduced to the SFP features. In the

Optics EXPRESS

experiment, we assume that SFP 5 and SFP 8 are unauthorized anomalous spectra, verifying the ability of SVM and 1D-CNN to identify the actual acquired spectral data.

We first processed 4800 spectral samples using PCA and calculated the contribution of each group under different PC quantities, as shown in Fig. 11(a). According to the curve in the figure, we can see that the contribution of the first 10 PCs of all of the OS samples reached 97.89%, indicating that the difference of OSs is sufficiently obvious for the experimentally collected OS data. Figure 11(b) shows the distribution of 4800 OS samples in the feature space composed of the first 3 PCs. Color is used to indicate the SFP. It can be seen that the spectra corresponding to different SFPs are relatively localized. Spectra after transmission of different lightpaths exhibit similar spectral features.

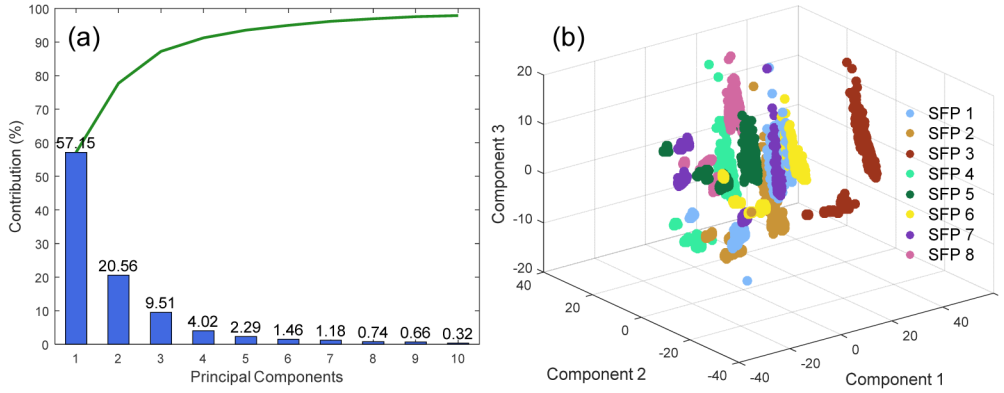

Fig. 11. (a) Contributions of PCs as the number of PCs is increased from 1 to 10. (b) Distribution of 4800 OS samples in the first 3 PC spaces (colors represents the SFP and different transmission conditions form distinct clusters of points).

4.2. Recognition results and analysis

In order to learn the spectral features of SFP to achieve OS recognition, we divided the 4800 OS samples into a training set and test set. We randomly selected 80% of the samples as training samples, and the remaining 20% of the samples were considered test samples. Since the experiment sets SFP 5 and SFP 8 as unauthorized signals, the spectral samples of the two SFPs are eliminated during model training. Then we performed regularization, centralization, and other pre-processing operations on the OS samples. PCA is used to extract the features in the spectrum and select different numbers of PCs. Using the extracted features, we trained the SVM and 1D-CNN models separately.

We then evaluated the training results of SVM and 1D-CNN. In the experiment, we selected the loss threshold of different loss values in the VPI simulation data recognition to the unauthorized sample, because there are differences between the sample collected by the experiment and those from VPI simulation. At the same time, because the SVM and the 1D-CNN model calculate the loss value in different ways (see Eq. (11) and Eq. (13) for details), the loss of the two models cannot be compared.

In the spectral recognition experiment, we evaluate the value of the accuracy, FAR, and MAR with different loss thresholds (SVM: 0.0001–0.01, interval 0.0003; 1D-CNN: 0.1–3, interval 0.1) when the number of principal components extracted (i.e., feature dimension) is 3–50. The closer the value of MAR and FAR are to 0, the more accurately the model recognizes the unauthorized signal. In Fig. 12, we present the results of the SVM and 1D-CNN for the SFP OS recognition.

It can be seen from Fig. 12(a) that when the threshold is set to 0.0022 and the number of extracted PCs is 21, the recognition accuracy of the SVM is the highest, reaching 98.54%. At this time, the MAR and FAR are 3.23% and 0.83%, respectively. The area with an accuracy larger than 90% is bounded by red in the figure. The FAR and MAR of the SVM recognition

Optics EXPRESS

result are shown in Figs. 12(c) and 12(e), and the areas of the MAR and FAR smaller than 10% are indicated with yellow lines. The FAR reaches a minimum value of 0.08% under the condition that the number of PCs is 13 or 17, and the loss threshold is in the range of 0.0085–0.01. The minimum value of the MAR appears when the loss threshold is 0.0001 and the number of PCs is 21, and is 1.74%. From the results of the experiments, we can see that, the value of loss threshold is very significant to achieve higher spectral recognition accuracy, and lower MAR, and FAR. In general, an SVM recognition accuracy of 98.54% is satisfactory in application scenarios.

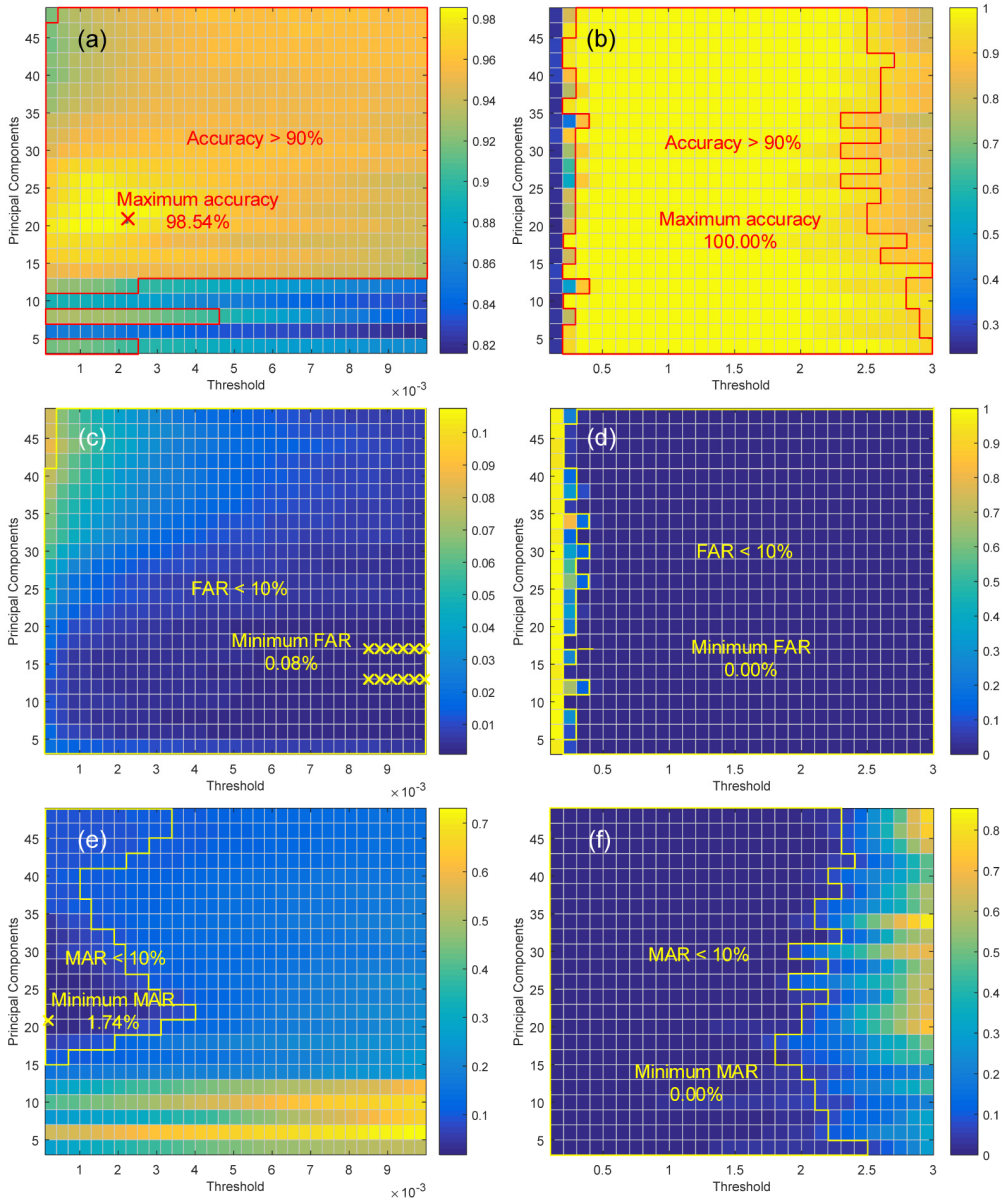

Fig. 12. Recognition results with SVM (a, c, e) and 1D-CNN (b, d, f): accuracy, FAR, and MAR of authorized/unauthorized OS recognition with different numbers of PCs and thresholds of loss.

Optics EXPRESS

Figures 12(b), 12(d) and 12(f) are the accuracy, FAR, and MAR of the 1D-CNN for actual SFP OS sample recognition. The number of PCs is 3–50, and the loss threshold is 0.1–3.0. As can be seen from the figures, the recognition results of 1D-CNN have obvious advantages compared with SVM. The overall recognition accuracy is higher. In the figures, the area over which the accuracy>90%, FAR<10%, and MAR<10% are larger than the corresponding areas of SVM, indicating that more values of loss threshold and number of PCs can be chosen. 1D-CNN also shows the advantage of high accuracy with low FAR and MAR at low numbers of PCs, indicating better performance in feature analysis and understanding, and the model can be further simplified more with fewer PCs. The accuracy of 1D-CNN reaches 100%, with corresponding FAR and MAR values of 0, under many combinations of loss threshold and number of PCs, effectively meeting the needs of unauthorized spectral recognition.

In general, PCA processing can reduce the dimensionality of spectral data, more efficiently extract spectral features, and eliminate extraneous components. The extracted features be used to simply the model. In this experiment, 1D-CNN exhibited better performance in feature analysis and unauthorized OS recognition than SVM, and has better adaptability to the selection of algorithm parameters. 1D-CNN is capable of detecting and recognizing unauthorized attack signals in complex networks.

5. Conclusions

In this paper, we mainly focus on OS features and feature analysis methods for OS recognition, which could be used for unauthorized attack signal detection in optical networks.

First, we theoretically analyzed the non-ideal influencing factors of OS features. Simulations were configured using commercial software VPI Transmission Maker and different transmitters with multiple non-ideal influencing factors were simulated to verify the theoretical derivation. The features of OSs collected from the simulation were quantitatively analyzed with the proposed BWCDR, which measures the separability of the samples. We also use PCA to extract spectral features and reduce the dimensionality of the OS samples. The simulation results support the theoretical derivation.

Secondly, we proposed SVM-based and 1D-CNN-based OSFA methods to learn the OS features and realize OS feature recognition and unauthorized signals detection. The methods are verified with the OSs collected from the simulations. We investigated the OS recognition accuracy, FAR, and MAR to evaluate the performance of the methods, and the result shows that the recognition accuracy of 1D-CNN (99.86%) is better than SVM (93.37%). Further, 1D-CNN has increased flexibility of model parameter optimization.

Thirdly, we set up an experimental demonstration to collect 4800 spectral samples from eight commercial SFPs transmitted through six lightpaths. The proposed unauthorized signal recognition methods were verified with the OSs collected from the experiments. The results show that both SVM and 1D-CNN can effectively realize feature learning and recognition of actual OS data. The accuracy of SVM can reach 98.54%, FAR = 0.83%, MAR = 3.23%, while 1D-CNN realized 100% recognition accuracy with FAR and MAR values of 0. In general, 1D-CNN has greater advantages than SVM in terms of accuracy and adaptability to different algorithm parameters. The investigation proves the proposed methods to be a promising solution for optical network security.

In future work, the effects of transmission distance and transmission configuration on OS features can be further studied. At the same time, when the spectrum of the attack signal and the transmitted signal are overlapped, methods to extract and recognize the OS features of the attack signal from the overlapped spectrum can be further studied.

The investigation is focusing on optical network security issues. Thus, how the method can be employed in a real field environment should be further studied. In this paper, we mainly study the OS features, the influencing factors and verified the proposed method in a small experimental network. In the future some experiments in real commercial network could be demonstrated. The working architecture and the workflow of the proposed method

Optics EXPRESS

will be further studied. The impact of the expansion of the real network scale, the complexity of the network environment and utilization of multiple transmission technologies, can be further studied.

Funding

National Natural Science Foundation of China (61871448, 61427813, 61621064), Beijing Municipal Natural Science Foundation (4172029).

References

1. M. Medard, D. Marquis, R. A. Barry, and S. G. Finn, "Security issues in all-optical networks," IEEE Netw. **11**(3), 42–48 (1997).
2. M. Furdek, N. Skorin-Kapov, S. Zsigmond, and L. Wosinska, "Vulnerabilities and security issues in optical networks," in *Proceedings of International Conference on Transparent Optical Networks* (IEEE, 2014), 1–4.
3. A. Lazzez, "All-optical networks: Security issues analysis," J. Opt. Commun. Netw. **7**(3), 136–145 (2015).
4. K. I. Kitayama, M. Sasaki, S. Araki, M. Tsubokawa, A. Tomita, K. Inoue, K. Harasawa, Y. Nagasako, and A. Takada, "Security in photonic networks: Threats and security enhancement," J. Lightwave Technol. **29**(21), 3210–3222 (2011).
5. M. Furdek and N. Skorin-Kapov, "Physical-layer attacks in transparent optical networks," in *Optical Communications Systems* (InTech, 2012).
6. N. Skorin-Kapov, J. Chen, and L. Wosinska, "A new approach to optical networks security: Attack-aware routing and wavelength assignment," IEEE/ACM Trans. Netw. **18**(3), 750–760 (2010).
7. J. Zhu, B. Zhao, W. Lu, and Z. Zhu, "Attack-aware service provisioning to enhance physical-layer security in multi-domain EONs," J. Lightwave Technol. **34**(11), 2645–2655 (2016).
8. J. Zhu and Z. Zhu, "Physical-layer security in MCF-based SDM-EONs: Would crosstalk-aware service provisioning be good enough?" J. Lightwave Technol. **35**(22), 4826–4837 (2017).
9. Y. Li, N. Hua, Y. Song, S. Li, and X. Zheng, "Fast lightpath hopping enabled by time synchronization for optical network security," IEEE Commun. Lett. **20**(1), 101–104 (2016).
10. W. Zhang, C. Zhang, C. Chen, and K. Qiu, "Experimental demonstration of security-enhanced OFDMA-PON using chaotic constellation transformation and pilot-aided secure key agreement," J. Lightwave Technol. **35**(9), 1524–1530 (2017).
11. Y. Xiao, Z. Wang, J. Cao, R. Deng, Y. Liu, J. He, and L. Chen, "Time-frequency domain encryption with SLM scheme for physical-layer security in an OFDM-PON system," J. Opt. Commun. Netw. **10**(1), 46–51 (2018).
12. Computer Security Institute, "CSI computer crime and security survey" (2008).
13. N. Skorin-Kapov, M. Furdek, S. Zsigmond, and L. Wosinska, "Physical-layer security in evolving optical networks," IEEE Commun. Mag. **54**(8), 110–117 (2016).
14. B. B. Wu and E. E. Narimanov, "A method for secure communications over a public fiber-optical network," Opt. Express **14**(9), 3738–3751 (2006).
15. P. Saengudomlert, "Analysis and detection of jamming attacks in an all-optical network," Diss., Massachusetts Institute of Technology (1998).
16. C. Mas, I. Tomkos, and O. K. Tonguz, "Failure location algorithm for transparent optical networks," IEEE J. Sel. Areas Comm. **23**(8), 1508–1519 (2005).
17. N. Li, N. Hua, Y. Yu, Q. Luo, and X. Zheng, "Light Source and Trail Recognition via Optical Spectrum Feature Analysis for Optical Network Security," IEEE Commun. Lett. **22**(5), 982–985 (2018).
18. R. Rejeb, M. S. Leeson, and R. J. Green, "Multiple attack localization and identification in all-optical networks," Opt. Switching Networking **3**(1), 41–49 (2006).
19. T. Wu and A. K. Somani, "Cross-talk attack monitoring and localization in all-optical networks," IEEE/ACM Trans. Netw. **13**(6), 1390–1401 (2005).
20. S. Yan, A. Aguado, Y. Ou, R. Wang, R. Nejabati, and D. Simeonidou, "Multilayer network analytics with SDN-based monitoring framework," IEEE/OSA J. Opt. Commun. Netw. **9**(2), A271–A279 (2017).
21. D. Gariépy, S. Searcy, G. He, and S. Tibuleac, "Non-intrusive OSNR measurement of polarization-multiplexed signals with spectral shaping and subject to fiber non-linearity with minimum channel spacing of 37.5GHz," Opt. Express **24**(18), 20156–20166 (2016).
22. S. Preussler, A. Zadok, A. Wiatrek, M. Tur, and T. Schneider, "Enhancement of spectral resolution and optical rejection ratio of Brillouin optical spectral analysis using polarization pulling," Opt. Express **20**(13), 14734–14745 (2012).
23. H. Lu, S. Cui, C. Ke, and D. Liu, "Automatic reference optical spectrum retrieval method for ultra-high resolution optical spectrum distortion analysis utilizing integrated machine learning techniques," Opt. Express **25**(26), 32491–32503 (2017).
24. F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An Overview on Application of Machine Learning Techniques in Optical Networks," IEEE Comm. Surv. and Tutor. **21**(2), 1383–1408 (2019).
25. D. Rafique and L. Velasco, "Machine learning for network automation: Overview, architecture, and applications [invited tutorial]," J. Opt. Commun. Netw. **10**(10), D126–D143 (2018).

Optics EXPRESS

26. M. Bouda, S. Oda, O. Vasilieva, M. Miyabe, S. Yoshida, T. Katagiri, Y. Aoki, T. Hoshida, and T. Ikeuchi, "Accurate prediction of quality of transmission with dynamically configurable optical impairment model," in *Optical Fiber Communications Conference* (Optical Society of America, 2017), paper Th1J.4.
27. L. Barletta, A. Giusti, C. Rottondi, and M. Tornatore, "QoT estimation for unestablished lighpaths using machine learning," in *Optical Fiber Communications Conference* (Optical Society of America, 2017), paper Th1J.1.
28. F. Ye, J. Tu, K. Saitoh, K. Takenaga, S. Matsuo, and T. Morioka, "A new and simple method for crosstalk estimation in homogeneous trench-assisted multi-core fibers," in *Asia Communications and Photonics Conference* (Optical Society of America, 2014), paper AW4C.3.
29. D. Zibar, L. H. H. de Carvalho, M. Piels, A. Doberstein, J. Diniz, B. Nebendahl, C. Franciscangelis, J. Estaran, H. Haisch, N. G. Gonzalez, J. C. de Oliveira, and I. T. Monroy, "Application of machine learning techniques for amplitude and phase noise characterization," J. Lightwave Technol. **33**(7), 1333–1343 (2015).
30. F. N. Khan, K. Zhong, X. Zhou, W. H. Al-Arashi, C. Yu, C. Lu, and A. P. T. Lau, "Joint OSNR monitoring and modulation format identification in digital coherent receivers using deep neural networks," Opt. Express **25**(15), 17767–17776 (2017).
31. C. Rottondi, L. Barletta, A. Giusti, and M. Tornatore, "Machine-learning method for quality of transmission prediction of unestablished lightpaths," J. Opt. Commun. Netw. **10**(2), A286–A297 (2018).
32. Y. Huang, C. L. Gutterman, P. Samadi, P. B. Cho, W. Samoud, C. Ware, M. Lourdiane, G. Zussman, and K. Bergman, "Dynamic mitigation of EDFA power excursions with machine learning," Opt. Express **25**(3), 2245–2258 (2017).
33. J. Thrane, J. Wass, M. Piels, J. C. M. Diniz, R. Jones, and D. Zibar, "Machine Learning Techniques for Optical Performance Monitoring From Directly Detected PDM-QAM Signals," J. Lightwave Technol. **35**(4), 868–875 (2017).
34. L. Yi, T. Liao, L. Huang, L. Xue, P. Li, and W. Hu, "Machine Learning for 100 Gb/s/λ Passive Optical Network," J. Lightwave Technol. **37**(6), 1621–1630 (2019).
35. S. Shahkarami, F. Musumeci, F. Cugini, and M. Tornatore, "Machine-learning-based soft-failure detection and identification in optical networks," in *Optical Fiber Communications Conference* (Optical Society of America, 2018), paper M3A.5.
36. S. Yan, F. N. Khan, A. Mavromatis, D. Gkounis, Q. Fan, F. Ntavou, K. Nikolovgenis, F. Meng, E. H. Salas, C. Guo, C. Lu, A. P. T. Lau, R. Nejabati, and D. Simeonidou, "Field trial of machine-learning-assisted and SDN-based optical network planning with network-scale monitoring database," in *Proceedings of European Conference on Optical Communication* (IEEE, 2017), 1–3.
37. S. Yan, F. N. Khan, A. Mavromatis, Q. Fan, H. Frank, R. Nejabati, A. P. T. Lau, and D. Simeonidou, "Field trial of machine-learning-assisted and SDN-based optical network management," in *Optical Fiber Communication Conference* (Optical Society of America, 2019), paper M2E.1.
38. Y. Li, N. Hua, C. Zhao, H. Wang, R. Luo, and X. Zheng, "Real-Time Rogue ONU Identification with 1D-CNN-based Optical Spectrum Analysis for Secure PON," in *Optical Fiber Communication Conference* (Optical Society of America, 2019), paper Tu3B.3.
39. P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 711–720 (1997).
40. I. Jolliffe, *Principal Component Analysis* (Springer, 2011).
41. O. Chapelle, "Training a support vector machine in the primal," Neural Comput. **19**(5), 1155–1178 (2007).
42. A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems* (NIPS, 2012), pp. 1097–1105.