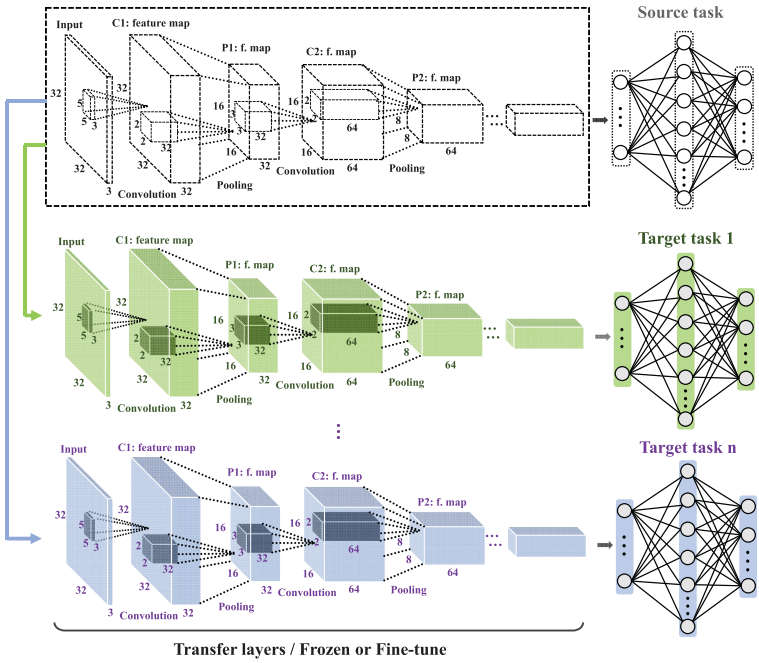
IEEE Photonics Journal
An IEEE Photonics Society Publication

Open Access

Comprehensive Eye Diagram Analysis: A Transfer Learning Approach

Volume 11, Number 6, December 2019

Danshi Wang, *Member, IEEE*
Yilan Xu
Jianqiang Li
Min Zhang
Jin Li
Jun Qin
Cheng Ju
Zhiguo Zhang
Xue Chen



DOI: 10.1109/JPHOT.2019.2947705

Comprehensive Eye Diagram Analysis: A Transfer Learning Approach

Danshi Wang,[1] ***Member, IEEE***, **Yilan Xu**,[1] **Jianqiang Li**,[1] **Min Zhang**,[1] **Jin Li**,[1] **Jun Qin**,[2] **Cheng Ju**,[3] **Zhiguo Zhang**,[1] **and Xue Chen** [1]

[1]State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China
[2]State Key Laboratory of Advanced Optical Communication Systems and Networks, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China
[3]College of Electronic Information, Qingdao University, Qingdao 266071, China

DOI:10.1109/JPHOT.2019.2947705

Manuscript received September 27, 2019; accepted October 13, 2019. Date of publication October 16, 2019; date of current version November 13, 2019. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Project 61705016, in part by Fundamental Research Funds for the Central Universities 2019RC12, and in part by BUPT Excellent Ph.D. Students Foundation CX2019313. Corresponding author: Danshi Wang (e-mail: danshi_wang@bupt.edu.cn).

Abstract: A deep transfer learning (TL)-based comprehensive eye diagram analysis and diagnosis scheme that can output essential eye diagram parameters, estimate fiber link length, calculate Q-factor, and diagnose device imperfection-induced impairments is proposed. TL can be used to extract system information and optical signal characteristics contained in eye diagrams and apply the learned knowledge and extracted features obtained from source tasks to related target tasks. As a source task, the proposed method estimates the transmission distance of a fiber link using convolutional neural network (CNN)-based eye diagram recognition. The feature extraction layers of the CNN are transferred to six target tasks involving the recognition of cross percentage, levels "0" and "1," eye height and width, and Q-factor. Using TL reduces the total training times for on-off keying (OOK) and pulse amplitude modulation (PAM4) formats by $>$95% and 60%, respectively. We also investigated six common PAM4 impairments caused by transmitter imperfection by setting the impairment category identification as source task and the impairment-degree diagnoses as target tasks. The TL methods consistently outperformed non-TL methods, with higher accuracies and significantly reduced training times. The proposed impairment diagnosis technique should be useful in impairment healing and fault correction.

Index Terms: Transfer learning, eye diagram analysis, impairment diagnosis.

1. Introduction

Optical signal measurement and analysis are significant tools for optical performance monitoring (OPM), signal quality assessment, impairment diagnosis, and fault detection [1]. In intensity modulation-direct detection (IM-DD) systems, eye diagrams are a primary important analysis object for comprehensively modeling the quality of optical signals and intuitively displaying the characteristics of various impairments occurring in optical communication systems [2]. Eye diagram analysis primarily focuses on on-off keying (OOK)-and pulse amplitude modulation (PAM)-formatted signals, which are widely used in optical access, metropolitan area, and data center optical networks [3]–[5].

However, conventional eye diagram analysis methods rely on professional expertise and engineering skills and are therefore prone to unavoidable human error and subject to higher labor costs.

To address this problem, convolutional neural network (CNN)-based deep learning techniques have recently been applied in optical signal analysis. Such techniques include optical signal-to-noise (OSNR) estimation [6], [7], modulation format recognition (MFR) [8], bit-rate identification [9], and atmospheric turbulence detection [10]. CNN-based eye and constellation diagram analysis schemes have also been proposed to perform joint MFR and OSNR estimation for multiple formats [11], [12]. However, such schemes can only simultaneously execute a small number of tasks. This is because in multi-task learning, the multiple tasks are simultaneously implemented in one whole neural network, as multi-task learning algorithms require increasingly complex network structures, such as deeper network layers and more neurons in the output layer, as the number of tasks increases. When the number of tasks is larger than three, the algorithmic complexity becomes much less tractable and the processing accuracy sharply decreases [13]. Although single-task learning algorithms can process multiple tasks, to do so they must assemble multiple algorithms to achieve each target separately [14], making it necessary to train each algorithm separately from the original state for each task, which takes substantial training time and consumes significant computational resources. Eye diagrams contain abundant information and plentiful optical signal characteristics, making it useful to develop more powerful eye diagram analysis techniques that can execute as many tasks as possible with limited training times and computational resource use, and without the extra cost of added complexity and accuracy.

As artificial intelligence (AI) techniques have become more widespread, transfer learning (TL) has become a rapidly expanding field of research for making deep learning easier, faster, and more efficient, especially in multi-task cases [15]. TL is a learning process that focuses on storing knowledge gained from a source task and transferring it to different but related tasks [16]. While most deep learning algorithms assume that the training and testing data have the same distribution, under TL, the domains, tasks, and distributions used in training and testing can differ—a condition that is much more compatible with real-world situations [17]. TL techniques have been applied in learning text data among different domains, cross-language processing, human activity classification, and multi-class image recognition [18]. In the field of optical signal analysis, TL can be used to obtain a variety of characteristics and information from eye diagram data through knowledge transfer among different task domains.

In this study, a TL approach was used to carry out comprehensive eye diagram analysis and impairment diagnosis. To perform image processing, a CNN was used to recognize eye diagram images and the TL algorithm was then used to effectively transfer the learned knowledge and extracted features from a source task (fiber length estimation) to execute six target tasks—the recognition of cross percentage, levels “0” and “1,” eye height and width, and Q-factor—for OOK- and PAM4-formatted signals. Both fine-tuning and frozen TL approaches were investigated. Compared to a TL-free approach, the TL algorithm achieved higher accuracies with significantly reduced training time. We then investigated six common impairments caused by transmitter imperfection—clipping and pattern effect, overshoot, eye-skew effect, insufficient extinction ratio, and mismatching power—using an impairment category identification as the source task and the corresponding degree-of-impairment diagnoses as the target tasks. Once again, higher accuracies and sharply reduced training time were achieved by the TL methods, demonstrating the feasibility of the proposed eye diagram analysis scheme.

2. Operating Principle

An eye diagram is an image based on the afterglow effect within an oscilloscope, which accumulatively produces an overlapping symbol waveform of a scanned signal. A wealth of information can be retrieved from an eye diagram, including essential characteristic parameters (such as high and low levels, cross percentages, and eye height and width), overall signal performance indicators (including modulation format, Q-factor, and OSNR), link properties and information (such as

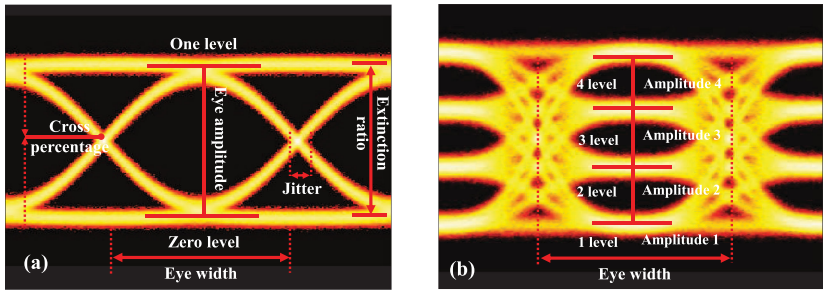


Fig. 1. Essential characteristic parameters of eye diagrams: (a) OOK and (b) PAM4 formats.

fiber transmission distance), and device imperfection-induced impairment features. The information obtained from eye diagram analysis is important for signal quality evaluation, OPM, impairment diagnosis and recovery, and the provision of reliable guidance for subsequent signal processing.

2.1 Essential Eye Diagram Characteristic Parameters

Conventional eye diagram analysis focuses on essential characteristic parameters that can be directly observed or measured, including "0"/"1" level, eye height and width, and cross percentage. As shown in Fig. 1(a), for an OOK signal the "0" and "1" levels in the eye diagram correspond to voltage reference values representing logic "0" and "1," respectively. The corresponding levels for a PAM4 signal are called the "1"–"4" level. Eye height corresponds to the degree of opening in the vertical direction at the middle of the eye diagram and, to a certain extent, reflects signal noise tolerance. Eye width corresponds to the extent to which the eye expands horizontally, i.e., to the time difference between the intersection of the upper and lower edges of the signal. Because signal jitter causes symbol widening, eye width can effectively reflect the overall jitter state of the signal. The cross percentage represents the relationship between the cross point amplitude and the "0"/"1" level. In general, a standard cross percentage of 50% indicates that the logic bits "1" and "0" each account for half of the signal. In other words, signals with different cross percentages have different level "0" and "1" transmission capacities.

In addition to these partial and microscopic parameters, eye diagrams also contain overall performance parameters such as modulation format and Q-factor. The modulation format of a signal is a performance parameter that can be directly observed from the overall eye diagram. The Q-factor, a performance index that is related to OSNR and can be obtained by eye diagram measurement, indicates the ratio of signal to noise power in a receiver under the optimal decision threshold. As the Q-factor in an optical interconnection system is affected by power, noise, and linear and nonlinear factors, it can be used to comprehensively represent system performance: larger Q-factors correspond to higher optical signal quality and better system performance.

2.2 Fiber Link Information Reflected in Eye Diagram

In addition to the parameters discussed above, eye diagrams also contain abundant transmission link information, reflecting the inherent impairments in the optical fiber such as dispersion, loss, nonlinear effects, etc. Such parameters cannot be obtained directly through observation or effectively through quantitative mathematical modeling, but machine learning algorithms can be used to mine hidden important performance parameters to enable qualitative analysis. As shown in Fig. 2, the eye diagrams produced by signals transmitted through fiber will differ by fiber length because the impacts of dispersion, loss, and nonlinearity in the fiber vary with transmission distance. Therefore, eye diagram analysis can be used to measure transmission link length, a valuable input for compensating the corresponding link impairments for digital signal processing at the receiver side.

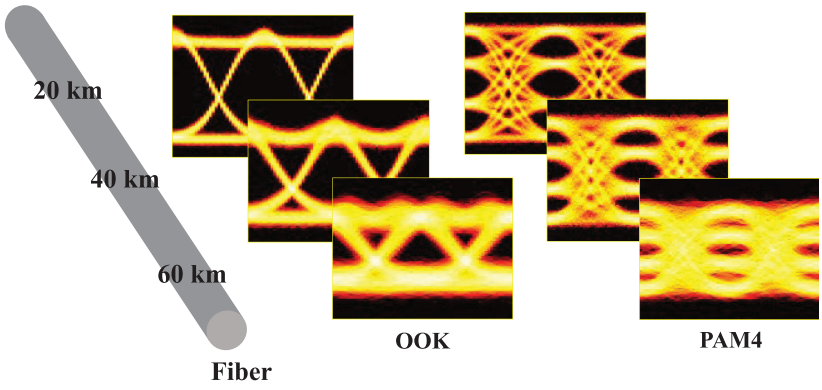


Fig. 2. Fiber transmission distance as reflected by eye diagrams for three OOK and PAM4 signals captured following transmission through 20-, 40-, and 60-km fibers.

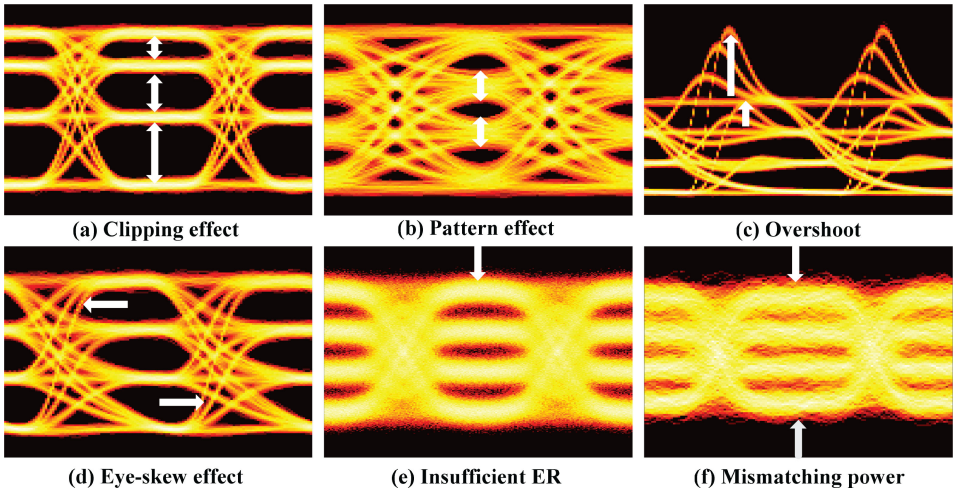


Fig. 3. Eye diagrams corresponding to six device imperfection-induced impairments: (a) clipping effect; (b) pattern effect; (c) overshoot; (d) eye-skew effect; (e) insufficient ER; and (f) mismatching power.

2.3 Impairment Characteristics Caused by Imperfect System Components

When device imperfection-induced impairment occurs, it is essential to diagnose the impairment, find its cause and the degree of impairment, and provide reliable guidance for distortion correction and system recovery. To enable impairment analysis, information on system impairment characteristics can be extracted from eye diagrams.

Here, we examine six common physical impairments originating from imperfect devices or improper operation in an IM-DD system and their mechanisms (Fig. 3):

a) Clipping effect: To provide a Mach-Zehnder modulator (MZM) with a maximum undistorted optical modulation amplitude, the modulation amplitude should be located in the linear region of the modulation curve while the optimal bias voltage point should be located at the quadrature point of the cosine modulation curve [19]. When bias voltage drifts as a result objective factors, the modulated electrical signal amplitude will leave the linear region, resulting in a non-equidistant signal output level, or clipping effect [20].

b) Pattern effect: The electro-optic (EO) bandwidth of a modulator is determined by the material constituting the active region [21]. When the bitrate exceeds the EO bandwidth, the electrical signal suffers from an insufficient recovery time to reach the reference level before the next

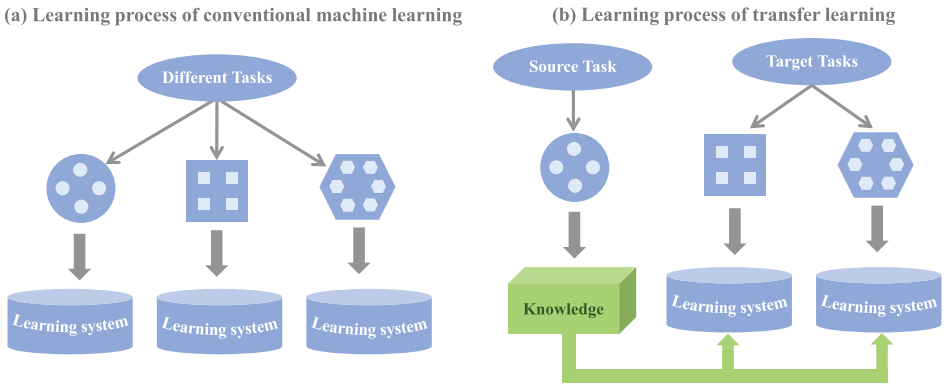


Fig. 4. Difference between conventional machine and transfer learning approaches for multiple tasks: (a) conventional machine learning and (b) transfer learning.

pulse arrives, resulting in a pattern effect that produces an eye diagram with a thickened eyelid and blurred and sharply changed traces [22].

c) Overshoot: The output signal from a directly modulated laser (DML) will be distorted at high modulation frequencies as a result of the intrinsic relaxation resonance in the laser cavity [23]. The overshoot occurs at the rising edge of the optical pulse because the current pulse injected into the laser induces a change in the carrier concentration of the active region.

d) Eye-skew effect: Direct-current modulation of semiconductor lasers, especially for the vertical cavity surface emitting lasers (VCSELs), is always accompanied by high-frequency chirps caused by the dependence of refractive index on carrier density [24]. Different injection currents change the refractive index, leading to changes in the frequency of the output optical signal and ultimately inducing a level-dependent skew in the eye diagram called the eye-skewed effect [25].

e) Insufficient extinction ratio (ER): ER is an important parameter that affects the quality of an output signal. At higher signal rates, it is more difficult to control the signal-to-noise ratio, which in turn significantly reduces the ER as a result of optical signal attenuation and transmission loss [26]. This finally results in an increased bit error rate and a compressed eye diagram.

f) Mismatched power: As a transmitter ages and its usage time grows, it becomes increasingly difficult to reach the rated laser output power. The resulting signal can be easily masked by noise, leading to a blurred and shrunken eye diagram [27]. Conversely, if the laser power is so high that it exceeds the capacity of the modulators, signal performance will also be impaired.

2.4 Transfer Learning

As noted in the discussion above, information on impairment errors can be acquired from the eye diagram produced by the signal. Although such analysis requires a multi-task learning procedure with multiple learning models, all of the tasks will focus on the same study object, namely, the eye diagrams, which means that different target tasks can learn from similar stored knowledge gained from a single source task. TL can be used to skillfully solve this problem by utilizing similarities between data, tasks, or models to transfer a trained model from a source domain to a new target domain [18], [28]. In conducting multi-task learning, conventional machine learning and TL differ most significantly in that approaches using the former must be retrained for different target tasks, whereas TL approaches do not require retraining (see Fig. 4).

Through the use of convolutional layers composed of sets of kernels, CNN-based deep learning can apply feature extraction and self-learning to process eye diagram images in their raw form without knowing other features. During a forward pass, each kernel convolves with pixel points across the width and height of an input image to calculate the dot product between the kernel

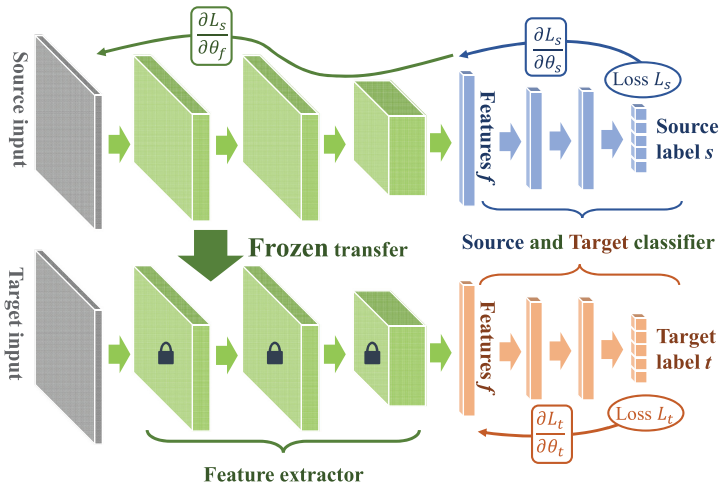


Fig. 5. Schematic of frozen transfer learning. The structure and connection parameters transferred from the source domain are retained while the parameters of the fully connected layers are adjusted through back-propagation.

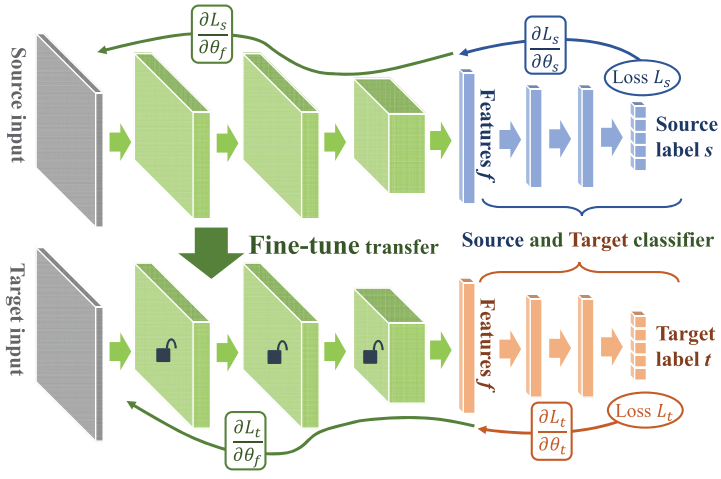


Fig. 6. Schematic of fine-tuning transfer learning. Both fully connected and transferred layers are adjusted by back-propagation.

entries and the input. The output units are organized into a two-dimension plane called the feature map [29]. In general, building an effective model requires several kernels to detect multiple features and produce multiple feature maps in a convolutional layer. Following feature extraction in the convolution layer, a pooling layer is used to merge semantically similar features into one feature. The features extracted from the convolutional and pooling layers are then sent to fully connected layers to execute a specific recognition task.

Convolutional and pooling layers trained from a source task can be transferred to other target tasks to reuse extracted general features and initialize a model from a half-finished state. Based on this pre-trained model, the target model can be further trained. To minimize the errors between target and actual output labels, the parameters in the networks are adjusted gradually via back-propagation using gradient descent methods [29]. TL has two commonly used parameter adjustment approaches–frozen and fine-tuning TL, which are shown in Figs. 5 and 6, respectively.

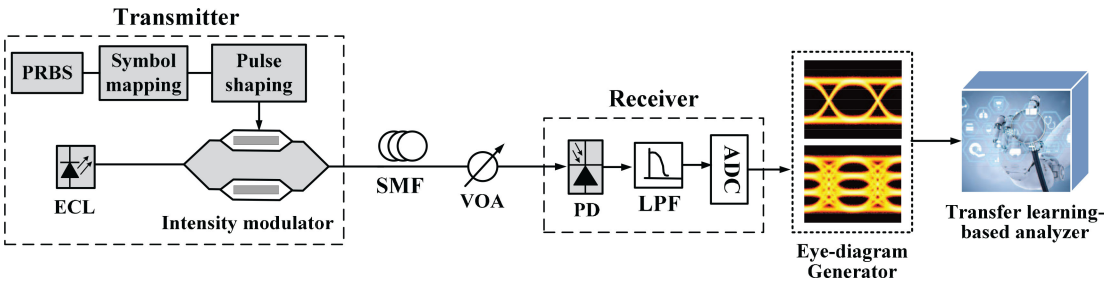


Fig. 7. Setup schematic. ECL: external cavity laser; PRBS: pseudo-random binary sequence; SMF: single mode fiber; VOA: variable optical attenuator; PD: photodetector; LPF: low pass filter; ADC: analog-to-digital converter.

Frozen TL involves the freezing of the front convolutional and pooling layers transferred from the source domain, including their network structure and connection parameters, to serve as a feature extractor and then adjusting only the parameters of the fully connected layers through back-propagation [30]. The loss function ∂L_t is used to assess the error magnitude, with $\partial L_t/\partial\theta_t$ used to adjust the weight value between the fully connected layers, as shown in Fig. 5. Under fine-tuning TL, both the network structure and connection parameters of the pre-trained model from the source domain are used [31]. Thus, in addition to adjusting the parameters of the fully connected layers, the parameters of the transferred components are fine-tuned by continued back-propagation in the target domain, as shown in Fig. 6.

3. System Setup and Demonstration

To demonstrate the feasibility of the proposed eye diagram analysis approach for IM-DD systems, we set up a simulation system based on VPI Transmission 8.6 software, as shown in Fig. 7. IM-DD systems for 10 Gb/s OOK and 20 Gb/s PAM4, respectively, were investigated. At the transmitter, an external cavity laser with the power of 0 dBm was used to provide an optical carrier to an intensity modulator (e.g., an MZM, DML, or VCSEL) driven by a pseudo-random binary sequence with length 2^{13}, symbol mapping, and a pulse shaper to generate OOK and PAM4 optical signals. In this simulation of transmitter, sample rate is 32 $\times$ 10 GHz, sample mode bandwidth is 128 $\times$ 10 GHz, rise time is 0.25$\times$ symbol period, and the greatest prime factor limit is 2. To simulate the transmission of a real optical link, we applied standard single-mode fiber (SMF) parameter settings, where dispersion is 16 $\times$ 10^{-6} s/m^2, nonlinear index is 2.6 $\times$ 10^{-20} m^2/W, and attenuation is 0.2 dB/km. At the receiver, the optical signal was directly detected by a photodetector, where detector type is PIN, electrical LPF filter type is 4-order Bessel filter, and filter bandwidth is 0.75 $\times$ 10 GHz. Following synchronous sampling, the eye diagram of the obtained digital electrical signal was drawn using a specialized oscilloscope-hosted eye diagram generation module that could convert the digital signal into the corresponding eye diagram images (in ".jpg" format). The generated eye diagrams were then fed to a TL-based eye diagram analyzer. Using this system, we separately collected eye diagrams produced under the two modulation formats over a transmission range of 0–80 km. Here, 70% of the diagrams were used as a training set to train the CNN model, while the remaining 30% were used as a test set to evaluate the accuracy of the results.

3.1 Transfer Learning for Fiber Link, Essential Parameters, and Q-Factor Estimation

The priority in implementing TL is to designate the source task, which is vital to original feature extraction and knowledge storing. In general, learning from the source task should enable the extraction of as many effective features as possible. In these experiments, we set fiber link analysis as the source task because it requires a more comprehensive recognition of eye diagram features. The structure of the CNN for eye diagram recognition—which comprises three pairs of convolutional

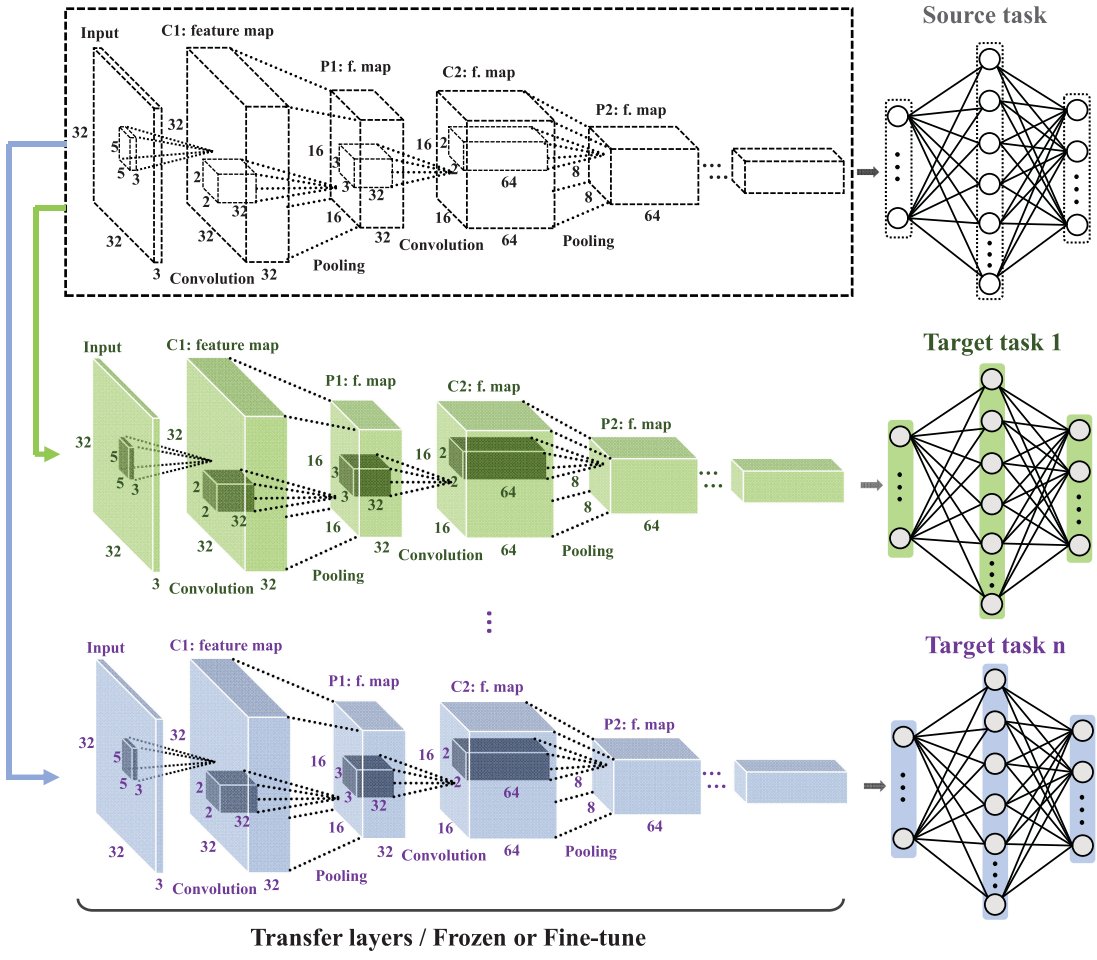


Fig. 8. CNN-based transfer learning network structure. Input layer: Eye-diagram images with pixel size of 32 × 32 × 3. Convolution layer 1(C1): 32 × 32 × 32 feature maps generated by 32 × 5 × 5 kernels. Pool layer (P1): 32 × 16 × 16 feature maps after subsampling from 2 × 2 region. C2: 64 × 16 × 16 feature maps generated by 64 × 3 × 3 kernels. P2: 64 × 8 × 8 feature maps after subsampling from 2 × 2 region. C3: 128 × 8 × 8 feature maps generated by 128 × 3 × 3 kernels. P3: 128 × 4 × 4 feature maps after subsampling from 2 × 2 region. The input layer and hidden layer of the fully connected layers consist of 2048 and 1024 neurons, respectively. The number of neurons in the output layer is task-dependent.

and pooling layers and three-layer fully connected layers—is displayed in Fig. 8. The corresponding parameter details are also presented in the caption of Fig. 8. We first used the two formats to measure the accuracies of estimation of fiber link transmission length at different epochs, as shown in Fig. 9. It is apparent from the figure that accuracy increased with the number of epochs. The accuracy of the OOK signal reached 100% at the minimum 0.5-km distance interval, while the accuracy of the PAM-4 signal remained relatively low at 0.5 km (81.73%) before finally reaching 100% at 1 km. It is because PAM-4 has a larger number of signal levels than the OOK format, leading to more features and the need to use a deeper network structure to improve the accuracy.

To demonstrate the advantages of CNN, five other well-known and widely-used machine learning algorithms [32]–[34]—k-nearest neighbor (KNN), random forest (RF), gradient boosting decision tree (GBDT), support vector machine (SVM), and artificial neural network (ANN)—were used to perform fiber link length estimation at 0.5- and 1-km intervals, as shown in Fig. 10. The key parameters of each algorithm are depicted in the caption of Fig. 10. It is seen from the histogram that the CNN obviously outperformed the other algorithms under both the OOK and PAM4 formats.

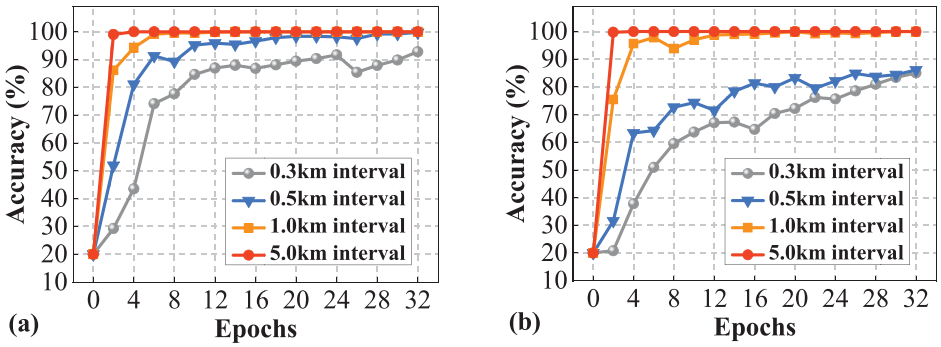


Fig. 9. Accuracy of fiber link length estimation at four different link length intervals: 0.3, 0.5, 1.0, and 5.0 km for (a) OOK and (b) PAM4.

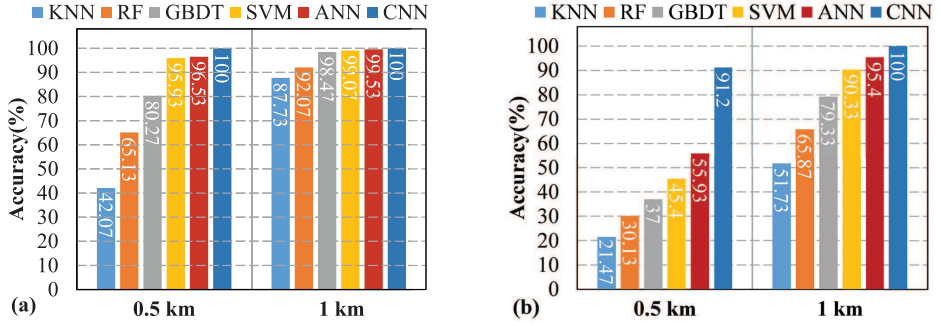


Fig. 10. Comparison between CNN and other algorithms in link length estimation at 0.5- and 1-km intervals for (a) OOK and (b) PAM4. The key parameters of each algorithm: KNN: number of neighbor is 5; RF: number of trees is 10; GBDT: number of estimators is 100; SVM: kernel function is radial basis function and degree of the polynomial kernel function is 3; ANN: number of neurons in hidden layer is 1024 and the maximum number of epochs is 200.

This result is attributable to the ability of the CNN to automatically extract and exploit more deep features in image applications.

We then transferred all of the CNN-trained feature extraction layers from the source task (link length estimation at 1-km intervals) to six target tasks, i.e., recognition of cross percentage, levels "0" and "1," and eye height and width, and estimation of Q-factor. Note that the first five tasks were processed as classification issues, while Q-factor estimation was processed as a regression issue to obtain more accurate analog values. The eye diagram images are divided into a 30×40 meshed grid (i.e., made up of 1200 finely subdivided regions). The positions where eye lips and cross points are located and the scopes that eye height and width cover correspond to the specific grid coordinates and regions. Through detecting the located coordinates and regions of eye local features, CNN can implement recognition tasks for different targets. As shown in Fig. 8, all of the CNN structures were similar to the source task structure. The feature extraction layers (i.e., convolutional and pooling layers) were completely transferred to the six target models as pre-trained feature extraction layers, while the fully connected layers were separately designed for the respective target tasks and retrained from their original state. For the purpose of comparison, we studied three cases for each target task: without TL, using fine-tuning TL, and using frozen TL. For the case without TL, both the convolutional and fully connected layers were retrained from the beginning and all of the parameters were adjusted from a randomly initialized state; for the fine-tuning TL case, three convolutional layers and one fully connected layer were finely adjusted

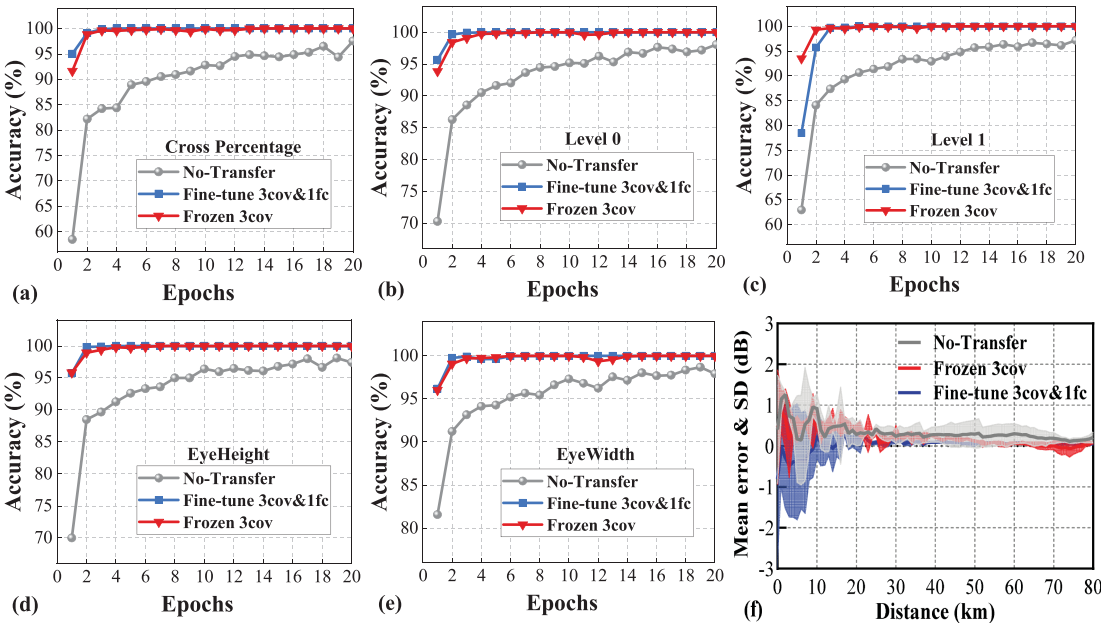


Fig. 11. Test results for six target tasks using OOK format with fine-tuning, frozen. and non-TL: (a) cross percentage; (b) level "0"; (c) level "1"; (d) eye height; (e) eye width; and (f) Q-factor.

from the pre-trained state of the source task; for the frozen TL case, three convolutional layers were retained and only the fully connected layers were adjusted.

We first measured the OOK-format results, as shown in Fig. 11. It can be seen from Figs. 11(a)–(e) that at epoch four the accuracies of the tasks based on fine-tuning and frozen TL are close to 100%, while the case without TL performs with poor accuracy up to epoch 20, demonstrating the superior performance of TL. To assess the regression of Q-factor estimation, the mean error and standard deviation (SD) were used to develop a performance evaluation index (Fig. 11(f)). The solid lines in the figure represent the mean errors of the estimated Q-factor values at epoch 20 at different transmission distances, with the colored envelopes indicating the corresponding SDs reflecting the respective degrees of mean error. It can be seen that the two cases with TL achieve performances similar to the non-TL case within 20 km and outperform it from 20 to 80 km. The TLs achieve a smaller degree of jitter, particularly in long-range transmission.

The training time, one of the most important performance indices, was also measured (Fig. 12). It is apparent from the figure that applying TL—frozen TL in particular—significantly reduces the training times for the six target tasks. This reduction occurs because the pre-trained model and features extracted from the eye diagram by TL save an enormous amount of time and shorten the parameter adjustment procedure. Frozen TL has fewer parameters and adjusts fewer layers, resulting in an extremely shortened training time.

We then conducted similar performance assessments of the six target tasks using the PAM4 format (Fig. 13) with the highest and lowest of the four PAM4 levels identified as token levels. Here, the cross percentage refers to the middle cross point ratio and eye height refers to the height between the highest and lowest levels. It is seen from Fig. 13 that, by epoch 30, the accuracies of each task performed using fine-tuning and frozen TL are close to 100%. PAM4 requires a longer training period than OOK and is therefore more difficult for the CNN to recognize: the format produces a larger number of eyes and levels than OOK and therefore represents more features, which in turn leads to a more complicated feature detection process and a longer learning time. Despite this, the TL-based methods still performed with an apparent comparative advantage to the non-TL approach, which could achieve accuracies of only 95% for each target task at epoch

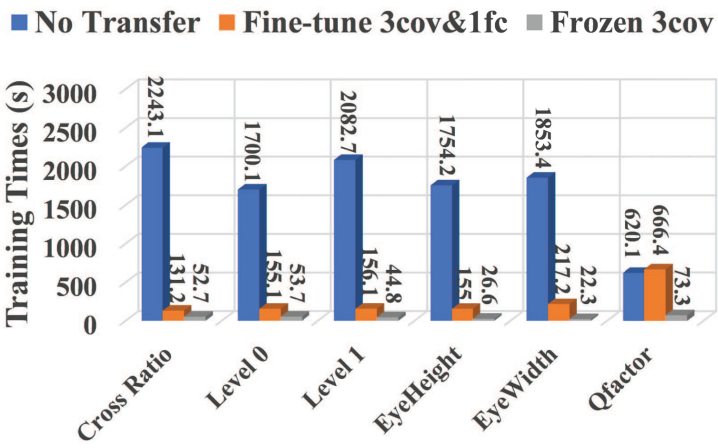


Fig. 12. Comparison of training times for six target tasks required using OOK format and fine-tuning, frozen, and non-TL.

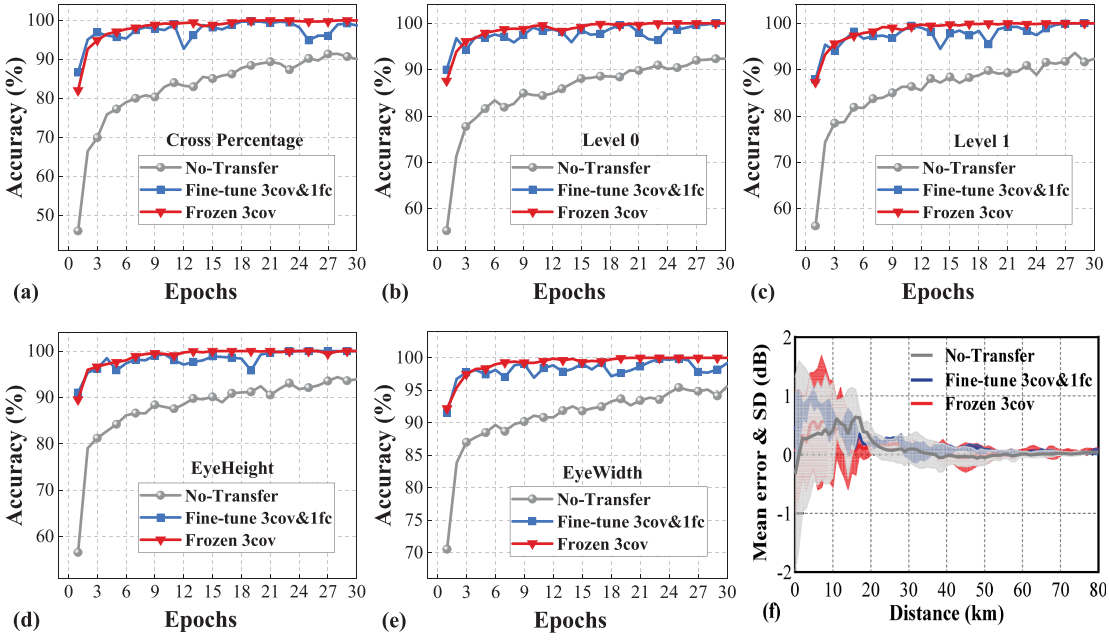


Fig. 13. Test results for six target tasks carried out using PAM4 format with fine-tuning, frozen, and non-TL: (a) cross percentage; (b) level "0"; (c) level "1"; (d) eye height; (e) eye width; and (f) Q-factor.

30. Frozen TL performed more stably than fine-tuning TL throughout the training process. The training times at which the respective methods achieved their optimal results are shown in Fig. 14. Compared to the non-TL case, training time could be reduced by an average of 90% by frozen TL, a result that can significantly reduce training cost and avoid repetitive learning.

To figure out how sources task features are utilized by the specific target tasks in TL, we also study how different source tasks affect the TL performance. Here, eye width recognition as the example is selected as target task and other three different tasks (fiber length, eye height, Q-factor) are set as source task separately. Three source models are trained by the corresponding source tasks, and after 100 epochs, the whole convolutional layers from these three source models are frozen and transferred to the target task for eye width recognition. The test results are displayed

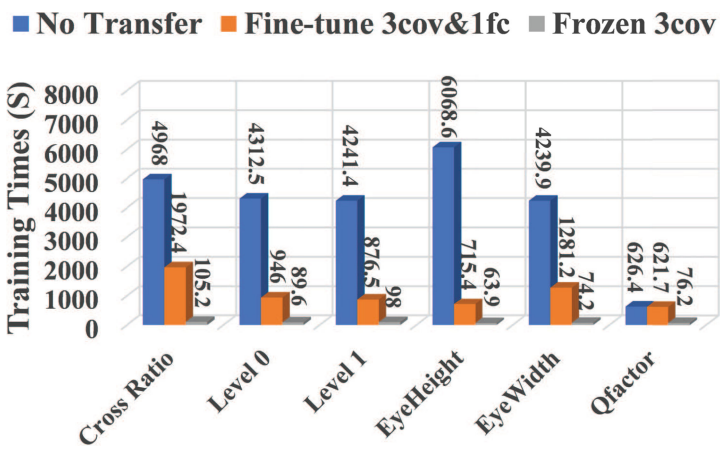


Fig. 14. Comparison of training times among fine-tuning, frozen, and non-TL for six target tasks using PAM4 format.

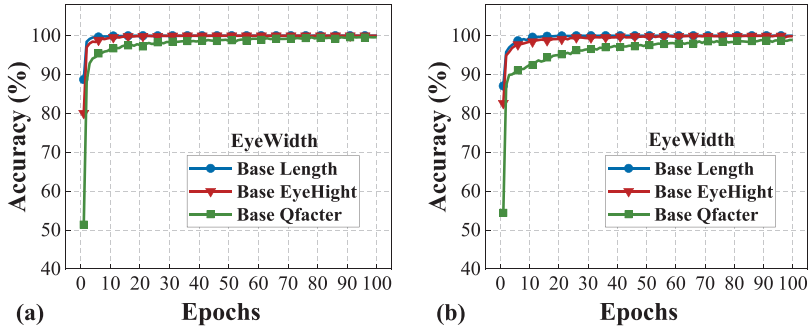


Fig. 15. The test results for eye width recognition transferred from different source tasks: (a) OOK and (b) PAM4.

in Fig. 15. It is seen that the source task “fiber length recognition” performs the optimal results, illustrating the features extracted from “fiber length recognition” is the most effective and worthy to be transferred.

Then we set “fiber length recognition” as source task and transfer different convolutional layers to the target task. Four cases are measured: transferring the first one convolutional layer, the first two convolutional layers, the whole three convolutional layers, the whole three convolutional layer and the first one fully-connected layer, as shown in Fig. 16. It is clearly seen that more layers transferred to the target task lead to the faster convergence and better performance. As discussed in [35], deep neural networks trained on natural images exhibit a common rule: on the first layer they learn features similar to Gabor filters and color blobs. Such first-layer features appear not to be specific to a particular dataset or task, but general in that they are applicable to many datasets and tasks. Features must eventually transition from general to specific by the last layer of the network. Therefore, in our scheme, the more specific features in last layer is effective and helpful for target task “eye width recognition”.

In conclusion, the target tasks can learn the valuable features from the more related source task with diverse characteristics. Meanwhile, if source task is appropriate and strongly related, compared with general features obtained in first few layers, the more specific features extracted from the last layer are more applicable to target task, but if source task is not weakly related, then the general features would be preferable for transferring.

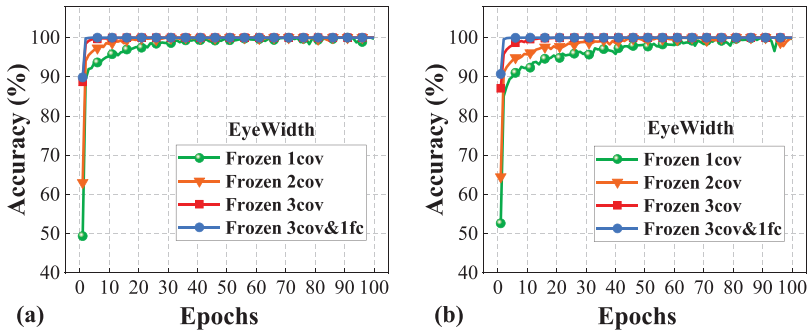


Fig. 16. Transferring different layers to target model of eye width recognition: (a) OOK and (b) PAM4.

3.2 Transfer Learning for Impairment Diagnosis

We then extended the TL technique to the task of impairment diagnosis based on eye diagram images, which to the best of our knowledge had not been reported to date. When impairment occurs, two types of tasks must be executed through impairment diagnosis: identifying the impairment category—that is, confirming what has caused the impairment—and, based on the determined impairment category, quantifying the degree of impairment to obtain the accurate diagnosis information for distortion correction and fault repair. We investigated six common impairments that may occur under the PAM4 format as a result of transmitter imperfection, including the clipping and pattern effects, overshoot, the eye-skew effect, insufficient ER, and mismatched power (see the preceding discussion of these effects and Fig. 3). The impairment category identification was set as the source task and diagnosing the degrees of impairment for the respective impairment types was set as the six target tasks.

The main driving factor of the clipping effect is bias voltage. Under the indirect modulation PAM4 system, 11 degrees (corresponding to classes) of clipping effect can be obtained by adjusting bias voltage at 0.25 V intervals over a half cycle (i.e., 2–4.5 V) of the MZM transmission function. As the initial phase of the transmission function in our system was -0.5 V, to ensure a wide scope of coverage we sampled signals corresponding to an offset point located within the second half of the cycle to obtain varying degrees of clipping effect. When the offset point is located at the quadrature point of transmission function (i.e., 3.25 V), the clipping effect will not occur. When the offset point is located at the highest or lowest points, the corresponding eye diagrams cannot be distinguished because of serious signal distortions, and therefore the eye diagrams sampled at these two points are regarded as one case. Based on these restrictions, the number of degrees of impairment are reduced to nine classes.

The primary factor driving the pattern effect is the EO bandwidth of the modulators. Because EO bandwidth is determined by the material properties in the modulator gain region and is therefore relatively fixed, it is possible to obtain different degrees of pattern effect by adjusting the signal transmission rate. As the EO bandwidth of the modulator used in our system was 36.5 GHz and the signals were generated at bitrates of 20, 40, 60, and 80 Gb/s, the impairment degree was divided into four classes. At signal bitrates exceeding 80 Gb/s the signal was seriously distorted and eye diagram quality was very poor, and therefore no results were considered for this extreme case. On the other hand, signal transmission rates far below the EO bandwidth produced ideal eye diagrams that required no additional sampling.

The primary driving factor of overshoot at the rising signal edge is the injection current. The effect can be suppressed and mostly ignored if the bias current is close to the current threshold. However, when current pulse modulation is applied and there is a strong, instantaneous change in injection current, the effect will be significant and impossible to ignore. For the experiments, we set the injection current at intervals of 10 mA over the range 25–55 mA and divided the degree of impairment into four classes.

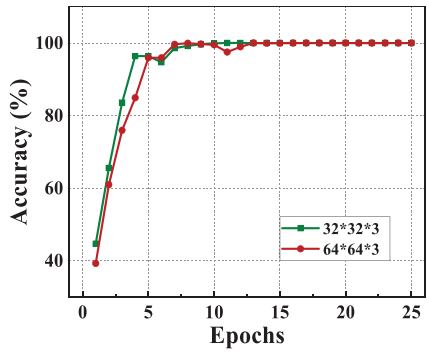


Fig. 17. Accuracy of impairment category identification by epoch at image sizes of $32 \times 32 \times 3$ and $64 \times 64 \times 3$.

The eye-skewing effect is primarily driven by the modulation current amplitude. As the VCSEL employed in the direct modulation PAM4 system had a threshold current of about 2 mA, a bias current close to the threshold value of 2.2 mA was chosen. Because the amplitude of the applied modulation current ranged from 2 to 20 mA, the impairment degree was divided into four classes by sampling at current intervals of 6 mA. At modulation currents below 2 mA, the modulated PAM4 signal was quite similar at all four levels and could easily be seriously distorted. At modulation currents above 20 mA, the overshoot effect occurred as a result of excessive modulation current.

Various degrees of signal impairment can be obtained by adjusting the ER parameter (insufficient ER). The lower end of the ER range was selected based on the general requirement that the signal ER should be above 10 dB; accordingly, we selected 8 dB as the lower measurement limit. Similarly, ERs above 15 dB have very little influence on the performance of an optical communication system and produce eye diagrams that differ only marginally; thus, we only chose an ideal measuring point at 20 dB. Based on this range, the impairment could be divided into four classes.

Different degrees of signal impairment from optical power mismatch can be obtained by adjusting the laser output power. For practical operation, we separately sampled signals at optical powers of 5, 10, and 50 mW and accordingly divided the mismatch into three classes: small, normal, and large.

From the preceding breakdown, there are a total of 28 classes of eye diagram data ($9 + 4 + 4 + 4 + 4 + 3$) covering six impairment categories with different degrees of impairment. We collected 100 eye diagram images for each class to produce a dataset containing 2,800 images with a 9:1 ratio of training to testing data. To conduct the source task of impairment category identification, half of the dataset samples were selected to train the CNN model to carry out the source task of impairment category identification with the goal of determining for each sample which of the six impairment categories it belonged to. We measured identification accuracy as a function of epoch for images of pixel-size $32 \times 32 \times 3$ and $64 \times 64 \times 3$, as shown in Fig. 17. It can be seen that, at both image sizes, an identification accuracy of 100% could be achieved by epoch 15. Although the larger eye diagram image size provided higher resolution and richer information, for the purposes of the study, the size of $32 \times 32 \times 3$ with smaller associated parameter count was sufficient and could converge faster. The following tests were therefore based on this size and consistent with the CNN structure in Fig. 8.

Next, using the TL approach, the CNN model of the source task was used as a pre-trained model for the six target tasks of executing impairment degree diagnosis corresponding to the six impairment categories. Using frozen and fine-tuning TL structures, the pre-trained model was reused to train networks for diagnosing the impairment degree of each impairment category. To demonstrate the advantages of using TL, new networks were trained for the same targets from the original state as a comparison. As shown in Fig. 18, the TL-based methods started at more advanced stages and had better initial performance than the non-TL method; that is, the models

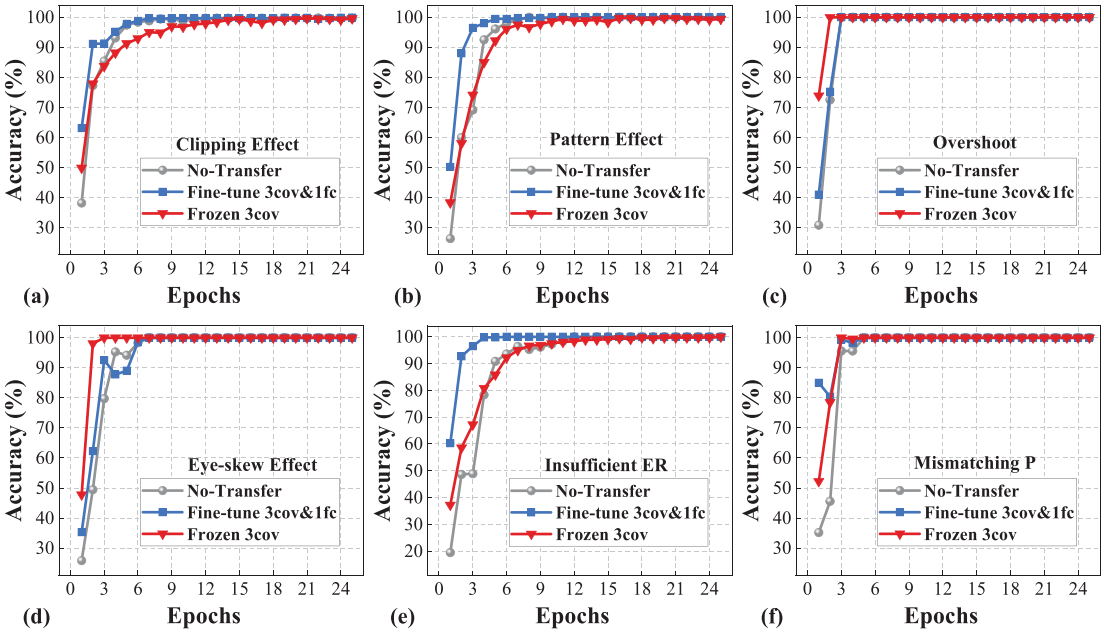


Fig. 18. Accuracies of six target tasks for impairment degree diagnosis: (a) clipping effect; (b) pattern; (c) overshoot; (d) eye-skew effect; (e) insufficient ER; and (f) mismatching optical power.

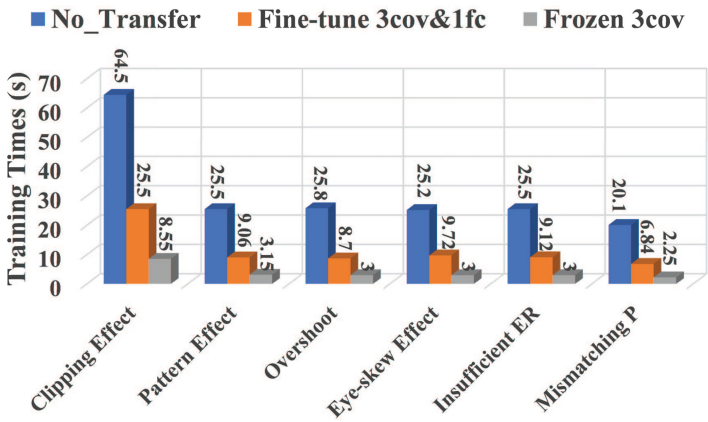


Fig. 19. Training times of fine-tuning, frozen, and non-TL approaches for six target tasks.

without TL started from zero knowledge and began initializing their parameters randomly, resulting in repetition of the learning procedure.

Figure 19 shows the training times of the three methods. The non-TL method had significantly longer training periods for each impairment category, with fine-tuning and frozen TL requiring the second-shortest and least training time, respectively, in each case. Specifically, the model training processes of the frozen and fine-tuning TLs were faster than the non-TL processes by factors of up to 20 and 2.5, respectively. By freezing the parameters of the transferred layers and only using the data to train the parameters of the fully connected layers, frozen TL reduced the number of model adjustment parameters significantly compared to the non-TL algorithm. Furthermore, most of the memory usage and time-consuming calculation in the CNN occur at the convolution layers, while the fully connected layers primarily identify eye diagram features and output classification results. The fine-tuning TL initializes its model parameters using the parameters obtained from the

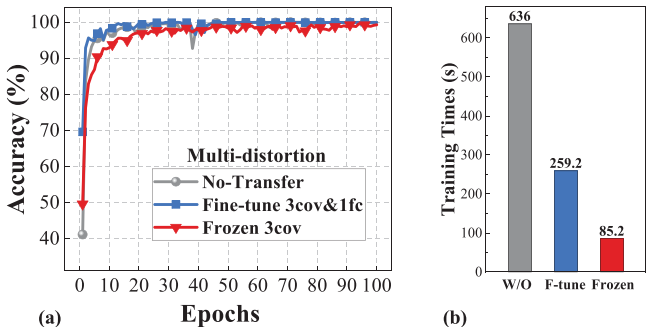


Fig. 20. Globally optimized scheme for multi-task learning with fine-tuning TL, frozen TL and no-TL: (a) accuracy under different epochs and (b) training time.

pre-trained model. This advancement of the starting point means that fewer epochs are needed to reach convergence, which reduces training time accordingly. However, because it is also necessary to fine-tune the parameters of the transfer layer through back-propagation, the training times of fine-tuning TL models are longer than those of models trained by freezing the transferred layer. In both cases, however, the application of deep TL significantly reduces the training time.

The six impairment degree diagnoses described above were conducted separately. Given the correlations among the six target tasks, the eye diagram analysis scheme could be globally optimized for multi-task operation using TL. In addition to a unique fully connected layer corresponding to each task, the other layers in the model would be hard parameter shared layers. It should be noted that, under such a multi-task globally optimized model, the transferred layer would not be the same as the shared layer but would instead coincide with the front part of the shared layer. The most significant difference between this scheme and an individually optimized model is that it would be able to execute each eye diagram analysis task simultaneously. That is, the model would have simple multiple outputs and therefore would be trained to minimize the sum of the error functions of each task, as the different tasks would correspond to different error functions.

Based on the above specifications, a pre-trained model was reused to train a composite network model for simultaneously diagnosing multi-impairment degree (Fig. 20(a)). As in the single-optimization cases, the transferred composite network demonstrated better initial performance than the non-transferred network. The accuracies of the non-transferred network and the transferred network using the fine-tuning method both reached 99.88%, while that of the transferred network using the frozen method reached 99.2%. Although in this case the accuracy of the transferred network using the frozen method was slightly lower than that of the other two networks, frozen TL still represents a worthwhile approach that can achieve good performance within a short training time (Fig. 20(b)). Furthermore, the transferred network converged faster than the network without TL.

3.3 Generalization

To demonstrate the generalization of the proposed method for practical eye diagrams, we collect some practical eye diagrams from experimental oscilloscope (Tektronix MSO 73304DX), which can execute real-time burst-mode eye diagram display. The optical signals at 10Gbaud are modulated by a MZM driven by an arbitrary waveform generator (AWG). After transmitting over 10 km fiber, OOK and PAM4 signals are detected by a PD. In addition, an ASE source and one variable optical attenuator (VOA) are added before receiver to adjust the optical signal-to-noise rate (OSNR) of optical signals so that generating various eye diagrams. We totally collect 20 different eye diagrams for OOK and PAM4 signals with OSNR ranging from 15 to 25 dB respectively. Here the eye diagrams of OOK and PAM4 in the case of back-to-back and at OSNR of 18 dB are selected as examples for display in Fig. 21. Similar to the simulation, all the images are transformed into gray-scale maps. Even there is a little chromatic aberration in color images between simulation and experiment, after

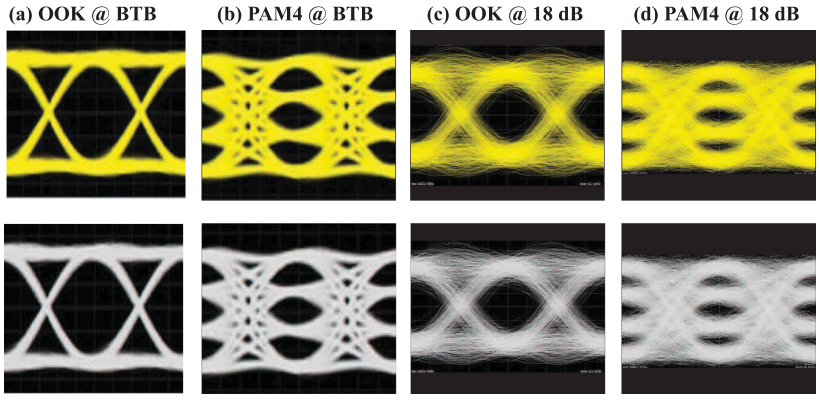


Fig. 21. The collected eye diagrams of OOK and PAM4 signals from oscilloscope: BTB and at OSNR of 18 dB.

TABLE 1
Summary of Misclassified Practical Eye Diagrams and Corresponding Error Rate

Format	Cross percentage	Level 0	Level 1	Eye height	Eye width	Q-factor
OOK	2 (10%)	0 (0%)	1 (5%)	1 (5%)	0 (0%)	ME<0.23
PAM4	2 (10%)	1 (5%)	1 (5%)	2 (10%)	3 (15%)	ME<0.31

gray transformation, the experimental eye diagrams look similar to the simulated ones. Then these experimental eye diagrams are directly sent into the transferred models from simulation to output the six essential characteristic parameters.

The correct values of these six parameters are measured by using the built-in dividing ruler of oscilloscope and determined by manual method, which are also the artificially estimated values but can be referenced as the comparison values for performance evaluation. Then we count the amount of misclassified experimental samples and the corresponding error rate for each parameter, as summarized in Table 1. For Q-factor, the mean error (ME) is calculated. This is a small sample set composed of 20 samples for each format, and thus error rate is equal to amount of misclassified samples/20. It is seen from Table 1, the misclassified samples for each parameter are not larger than 2 and 3 for OOK and PAM4 respectively, and the corresponding error rates are 10% and 15%. The mean errors of Q-factor are less than 0.23 and 0.31 for OOK and PAM4. Comparatively speaking, performance of PAM4 is worse than OOK. This is because the similarity of OOK eye diagrams between simulation and experiment is better than PAM4. It is observed from Fig. 21 that experimental eye diagram of PAM4 signal presents the more gentle rising edges (i.e., longer rising time caused by longer response time of a practical modulator) than simulation one, leading to the narrower eye width. But even so, all the misclassified samples are still limited within 3, demonstrating the acceptable generalization of the simulation-generated models in a certain extend. The performance can be further improved by adding the sufficient experimental data into the training data set, or applying technique of transferring learning to transfer the knowledge from simulation to the experiment, which are interesting and worthy of investigating in the further work.

Additionally, in our demonstration, the signal speed is 10 Gbaud and transmission distance is 0–80 km. If longer distance is deserved, the amplifier of EDFA is necessary at the intermediate node. Under this situation, eye diagram is amplified and accompanied with ASE noise from EDFA and other accumulated fiber impairments, meaning that the proposed method could be feasible to recognize these information for longer distance. Meanwhile, when the speed of optical signals grows from 10 Gbps to 40 Gbps or even more, the higher-rate signals are more sensitive to system impairments (like interference, dispersion, receiver bandwidth, etc), which can be visually displayed

in eye diagrams. Thus, we think the proposed method could be feasible for higher rate and longer distance with amplification

4. Conclusion

This paper proposed a comprehensive and multifunctional eye diagram analysis scheme based on deep TL. Using the estimation of transmission distance through the processing of eye diagram images via a CNN as a source task, six target tasks, including cross percentage, levels "0" and "1," eye height and width, and Q-factor estimation, were trained. Both fine-tuning and frozen TL-trained models achieved 100% accuracies within a small number of epochs, with training times reduced by over 95% for OOK and 60% for PAM4. We further investigated intelligent impairment diagnosis through the recognition of eye diagram features using impairment category identification as a source task and impairment degree diagnosis for six common impairments as target tasks. In all cases, the TL-based methods performed with higher accuracy and with a faster convergence speed. The training time relative to non-TL training was decreased by more than 60%, particularly for frozen TL-trained models. Finally, a composite CNN model to simultaneously diagnose multi-impairment degree was constructed. Even under its complicated network structure, accuracies of 99.88% and 99.2% were achieved using fine-tuning and frozen TL approaches, respectively. All the results suggest that TL is a feasible tool for eye diagram analysis and impairment diagnosis with reduced learning periods and lower computational overhead.

In conclusion, in our scheme, TL can not only speed up the training process, but also bring down the training data size. In addition, benefitting from the diverse data source, the models obtained by TL have the better generalization and stronger robustness.

References

[1] Z. Dong, F. N. Khan, Q. Sui, K. Zhong, C. Lu, and A. P. T. Lau, "Optical performance monitoring: A review of current and future technologies," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 525–543, Jan. 2015.

[2] P. D. Hale *et al.*, "A statistical study of de-embedding applied to eye diagram analysis," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 2, pp. 475–488, Feb. 2012.

[3] Y. Matsui *et al.*, "28-Gbaud PAM4 and 56-gb/s NRZ performance comparison using 1310-nm Al-BH DFB laser," *J. Lightw. Technol.*, vol. 34, no. 11, pp. 2677–2683, Jun. 2016.

[4] L. Chang *et al.*, "A 50 gb/s-pam4 CDR with on-chip eye opening monitor for reference-level and clock-sampling adaptation," in *Proc. Opt. Fiber Commun. Conf.*, 2018, Paper Tu2C–4.

[5] N. Eiselt *et al.*, "Performance comparison of 112-gb/s DMT, Nyquist pam4, and partial-response pam4 for future 5G ethernet-based fronthaul architecture," *J. Lightw. Technol.*, vol. 36, no. 10, pp. 1807–1814, May 2018.

[6] T. Tanimura, T. Hoshida, T. Kato, S. Watanabe, and H. Morikawa, "Convolutional neural network-based optical performance monitoring for optical transport networks," *J. Opt. Commun. Netw.*, vol. 11, no. 1, pp. A52–A59, Jan. 2019.

[7] D. Wang *et al.*, "Cost-effective and data size–adaptive OPM at intermediated node using convolutional neural network-based image processor," *Opt. Exp.*, vol. 27, no. 7, pp. 9403–9419, 2019.

[8] W. Zhang *et al.*, "Identifying modulation formats through 2D stokes planes with deep neural networks," *Opt. Exp.*, vol. 26, no. 18, pp. 23 507–23 517, 2018.

[9] X. Fan *et al.*, "Joint optical performance monitoring and modulation format/bit-rate identification by CNN-based multi-task learning," *IEEE Photon. J.*, vol. 10, no. 5, Oct. 2018, Art. no. 7906712.

[10] J. Li, M. Zhang, D. Wang, S. Wu, and Y. Zhan, "Joint atmospheric turbulence detection and adaptive demodulation technique using the CNN for the OAM-FSO communication," *Opt. Exp.*, vol. 26, no. 8, pp. 10 494–10 508, 2018.

[11] D. Wang *et al.*, "Modulation format recognition and OSNR estimation using CNN-based deep learning," *IEEE Photon. Technol. Lett.*, vol. 29, no. 19, pp. 1667–1670, Oct. 2017.

[12] D. Wang *et al.*, "Intelligent constellation diagram analyzer using convolutional neural network-based deep learning," *Opt. Exp.*, vol. 25, no. 15, pp. 17 150–17 166, 2017.

[13] S. Ruder, "An overview of multi-task learning in deep neural networks," 2017, *arXiv:1706.05098*.

[14] W. Liu *et al.*, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11–26, 2017.

[15] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[16] M. Long, Y. Cao, J. Wang, and M. I. Jordan, "Learning transferable features with deep adaptation networks," 2015, *arXiv:1502.02791*.

[17] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 1717–1724.

[18] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *J. Big Data*, vol. 3, no. 9, pp. 1–40, 2016.

[19] L. R. Hofer, D. Schaeffer, C. Constantin, and C. Niemann, "Bias voltage control in pulsed applications for Mach–Zehnder electrooptic intensity modulators," *IEEE Trans. Control Syst. Technol.*, vol. 25, no. 5, pp. 1890–1895, Sep. 2017.
[20] M.-H. Kim, B.-M. Yu, and W.-Y. Choi, "A Mach–Zehnder modulator bias controller based on OMA and average power monitoring," *IEEE Photon. Technol. Lett.*, vol. 29, no. 23, pp. 2043–2046, Dec. 2017.
[21] A. Rao *et al.*, "High-performance and linear thin-film lithium niobate Mach–Zehnder modulators on silicon up to 50 GHz," *Opt. Lett.*, vol. 41, no. 24, pp. 5700–5703, 2016.
[22] M. Zhang, D. Wang, Z. Cao, X. Chen, and S. Huang, "Suppression of pattern dependence in 10 gbps upstream transmission of WDM-pon with RSOA-based onus," *Opt. Commun.*, vol. 308, pp. 248–252, 2013.
[23] A. S. Karar, J. C. Cartledge, J. Harley, and K. Roberts, "Electronic pre-compensation for a 10.7-gb/s system employing a directly modulated laser," *J. Lightw. Technol.*, vol. 29, no. 13, pp. 2069–2076, Jul. 2011.
[24] R. S. Tucker, "High-speed modulation of semiconductor lasers," *IEEE Trans. Electron Devices*, vol. 32, no. 12, pp. 2572–2584, Dec. 1985.
[25] J. Zhou, C. Yu, and H. Kim, "Transmission performance of OOK and 4-pam signals using directly modulated 1.5-μ m vcsel for optical access network," *J. Lightw. Technol.*, vol. 33, no. 15, pp. 3243–3249, Aug. 2015.
[26] A. Napoli, M. M. Mezghanni, D. Rafique, V. A. Sleiffer, B. Spinnler, and M. Bohn, "Novel digital pre-distortion techniques for low-extinction ratio Mach–Zehnder modulators," in *Proc. Opt. Fiber Commun. Conf.*, 2015, Paper Th3G–1.
[27] H. Wenzel, P. Crump, A. Pietrzak, X. Wang, G. Erbert, and G. Tränkle, "Theoretical and experimental investigations of the limits to the maximum output power of laser diodes," *New J. Phys.*, vol. 12, no. 8, 2010, Art. no. 085007.
[28] L. Xia, J. Zhang, S. Hu, M. Zhu, Y. Song, and K. Qiu, "Transfer learning assisted deep neural network for osnr estimation," *Opt. Exp.*, vol. 27, no. 14, pp. 19 398–19 406, 2019.
[29] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, 2015, Art. no. 436.
[30] S. Christodoulidis, M. Anthimopoulos, L. Ebner, A. Christe, and S. Mougiakakou, "Multisource transfer learning with convolutional neural networks for lung pattern analysis," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 76–84, Jan. 2017.
[31] N. Tajbakhsh *et al.*, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016.
[32] C. M. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Germany: Springer, 2006.
[33] F. Musumeci *et al.*, "An overview on application of machine learning techniques in optical networks," *IEEE Commun. Surv. Tut.*, vol. 21, no. 2, pp. 1383–1408, Secondquarter 2019.
[34] F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "An optical communication's perspective on machine learning and its applications," *J. Lightw. Technol.*, vol. 37, no. 2, pp. 493–516, Jan. 2019.
[35] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 3320–3328.