

Breaking Boundaries: Harnessing Unrelated Image Data for Robust Risky Event Classification with scarce State of Polarization Data

Khoulood Abdelli⁽¹⁾, Matteo Lonardi⁽²⁾, Jurgen Gripp⁽³⁾, Samuel Olsson⁽³⁾, Fabien Boitier⁽⁴⁾, and Patricia Layec⁽⁴⁾

⁽¹⁾ Nokia Bell Labs, Germany Khoulood.Abdelli@nokia.com

⁽²⁾ Nokia Bell Labs, Vimercate, Italy, ⁽³⁾ Nokia, Murray Hill, NJ, 07974 USA, ⁽⁴⁾ Nokia Bell Labs, France

Abstract We present an innovative transfer learning method for classifying risky events with scarce state of polarization (SOP) data, utilizing a deep convolutional neural network pre-trained on unrelated images. Achieving a 96.3% accuracy on just 400 samples, this approach offers a robust solution for data-limited scenarios. ©2023 The Author(s)

Introduction

Fiber cuts are notable issues within optical communication systems, often leading to service disruptions and financial implications for network operators. The conventional approaches for detecting fiber outages are reactive based on predefined thresholds [1], or employing monitoring devices such as optical time domain reflectometry [2]. Recently, machine learning (ML) approaches for identifying precursor events to fiber breaks by leveraging insights from a state of polarization (SOP) data have been proposed [3-5]. However, such ML-based approaches necessitate a large amount of data to be trained for good performance, limiting their easy deployment in operational networks. Overfitting can occur when training a deep neural network on a small amount of training data. This results in low generalization when tested with unseen data. Therefore, the scarcity of meaningful data pose challenges in training ML models for SOP event recognition.

In this work, we propose a transfer learning approach for event classification given a small amount of SOP data. We investigate the transferability of the knowledge learned from millions of images in the ImageNet [6] dataset to

the SOP event domain by adopting a pre-trained deep neural network (MobileNet) [7] as a feature extractor. We then experimentally validate our approach. Our method helps to accelerate training and enhance model performance in cases when training data is limited.

Transfer Learning Approach

We propose to use a pre-trained model on a source dataset D_s and task. The learnt features (the network's weights) are then transferred to a fine-tuned model re-trained on a target dataset D_t and task. The proposed transfer learning for SOP event recognition is illustrated in Fig. 1 (a).

We consider as a source data the ImageNet dataset [6] (a generic image dataset) which is derived from an entirely different domain with almost no similarity to the target investigated images. D_s consists of over 14 million images belonging to 1000 classes including animals and objects. We use MobileNet, which has been pre-trained on ImageNet classes, as a feature extractor to transfer the learnt feature maps reflecting complex representations in generic images to the SOP image domain. MobileNet is a type of convolutional neural networks (CNNs) designed for mobile and embedded vision

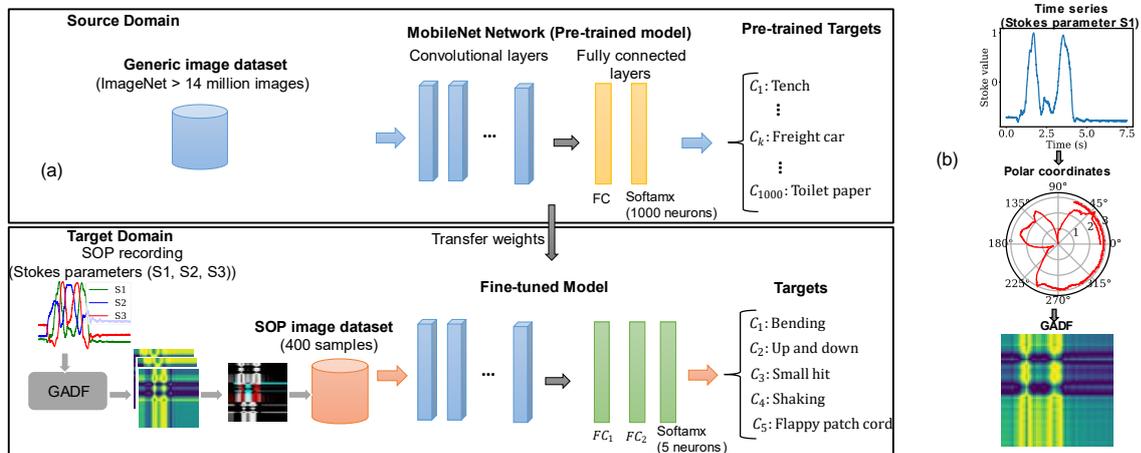


Fig. 1: (a) Illustration of the proposed transfer learning approach. The pre-trained model is developed on generic image dataset, then fine-tuned with 2D SOP images, (b) Steps for time series encoding into images.

applications. It is built on a streamlined convolutions [7] to develop lightweight deep neural networks. The MobileNet architecture is composed of 27 convolutions layers which includes 13 depth-wise convolutions, 1 average pooling layer, 1 fully connected layer, and 1 Softmax layer [7].

The target domain uses D_t , composed of SOP images. As shown in Fig. 1 (a), the SOP images are generated by (1) transforming the time series of the Stokes parameters (modelling the SOP variations) into multiple images using Gramian angular difference field (GADF) method [8] and (2) appending them to build one image, as will be discussed later. D_t incorporates the patterns or the signatures of 5 mechanical events (simulating precursor events to fiber outages) namely C_1 : bending, C_2 : shaking, C_3 : small hit, C_4 : up and down, and C_5 : fan ventilation. Given D_t , MobileNet is fine-tuned by replacing the fully connected layer (FC) with two fully connected layers (FC_1, FC_2) with 50 and 20 neurons respectively, followed by a Softmax layer whose number of neurons is equal to the number of classes in D_t (5 in our case). These three layers are in green in Fig. 1 (a). The convolution layers are kept frozen to avoid deleting any of the extracted feature information, i.e., learned weights that characterize typical features in images.

SOP Data Encoding as Images

Each SOP recording consists of a multivariate time series modelling the variations of the Stokes parameters (S_1, S_2, S_3). First, the times series of the Stokes parameters are converted into colored images using the GADF method [8]. Then, the transformed images are concatenated to produce a single image.

Let X denote a time series of a Stokes parameter $S_{i \in \{1,2,3\}}$, made of n observations. First, X is

rescaled between -1 and 1. The rescaled time series \tilde{X} is then represented in polar coordinates by encoding the value as the angular cosine and the time stamp as the radius by applying the following equation:

$$\begin{cases} \varphi = \arcsin(\tilde{x}_i), & -1 \leq \tilde{x}_i \leq 1, \tilde{x}_i \in \tilde{X} \\ r = \frac{t_i}{N}, & t_i \in \mathbb{N} \end{cases} \quad (1)$$

where t_i denotes the time stamp and N represents a constant factor to regularize the span of the polar coordinate system.

The polar coordinate representation has two advantages: (i) a one-to-one mapping of the time series to the polar coordinate system (i.e., it is bijective), and (ii) a preservation of absolute temporal correlations. Afterwards, the trigonometric difference of the inverse sine of each point is computed using the GADF method to identify the temporal correlation within different time intervals. Thus, a GADF matrix of size $(n \times n)$, representing the temporal features in the form of an image, is constructed. GADF is defined as follows:

$$GADF = [\sin(\varphi_i - \varphi_j)]. \quad (2)$$

The concept of transforming a time series of a Stokes parameter into an image using GADF is illustrated in Fig. 1 (b).

Three GADF images are produced, one for each Stokes parameter time series. These images must be appended to be fed into the CNN model as a single image. Each GADF image of size $(n \times n)$ is first divided into three monochrome channel images: red, green, and blue (RGB). The derived monochrome images are then concatenated together to build a bigger RGB image of size $(3n \times n)$.

Validation of the Transfer Learning Approach

The experimental setup shown in Fig. 2 (a) is carried out to record SOP measurements with different event signatures to build D_t . An SOP monitoring algorithm embedded within a

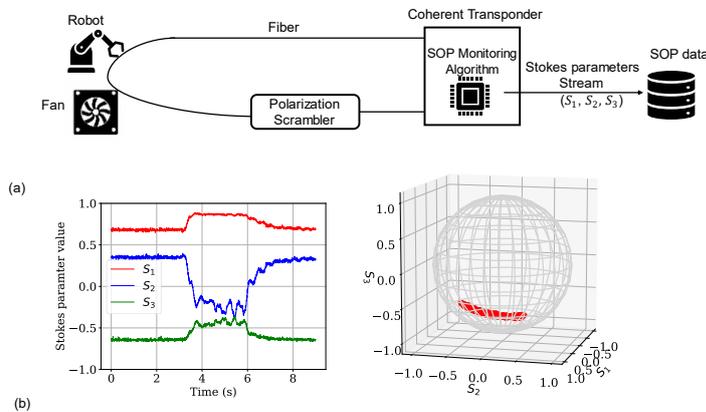


Fig. 2: (a) Experimental setup for SOP monitoring and recording. A robot arm and a fan simulate events along the fiber. (b) Example of Stokes parameters versus time for a "fan ventilation" event, also represented on the Poincaré sphere.

					1.000
C_5					
C_4			0.038	0.962	
C_3			0.923	0.077	
C_2		1.000			
C_1	0.941		0.059		
	C_1	C_2	C_3	C_4	C_5

Fig. 3: Confusion matrix achieved by our approach trained with a limited amount of data.

coherent transponder is used to measure SOP data. The Stokes parameters are measured periodically and stored in an external database. The robot arm controlled by an Arduino microcontroller is used to generate four different movements along the fiber including bending and shaking. The fan is controlled by an IoT power plug. The polarization scrambler is used to randomly change the initial SOP state prior to the event generation. Fig. 2 (b) shows the variation of the Stokes parameters as a function of time for a “fan ventilation” event, and its representation on the Poincaré sphere. The gathered SOP data is then converted into images using the previously described GADF and image concatenation methods before being fed into the pre-trained ML model for fine-tuning. D_t is composed of 400 SOP images (80 images for each investigated event (i.e., $C_i, \forall i \in \{1..5\}$)). The built data is split into a training (70%), validation (10%), and test (20%) set.

Our approach is tested with an unseen test dataset. The confusion matrix (Fig. 3) shows that our method (i) achieves a high average diagnostic accuracy of 96.3% despite being trained with a small dataset, and (ii) accurately identifies the different investigated events particularly the “shaking” and “fan ventilation” classes. The learning curves shown in Fig. 4 (a) demonstrate that our technique requires only 4 epochs to converge to a stable validation accuracy, which could speed up training.

We then compare in Fig. 4 (b) our approach to other methods namely gated recurrent unit (GRU), 1D-CNN, and 2D CNN, in terms of test accuracy for different sizes of the datasets. GRU and 1D-CNN methods take as input multivariate time series of Stokes parameters from SOP data after down-sampling them to reduce the computational complexity of the models. The 2D-CNN is fed with SOP images. GRU is composed of two GRU layers with 64, and 32 cells respectively, followed by a Softmax layer. The 1D-CNN consists of a convolutional layer with 32 filters followed by a max pooling layer, a flatten layer, and Softmax layer. The 2D-CNN is made of a convolutional layer with 16 filters followed by

a max pooling layer, a flatten layer, and a Softmax layer. Fig. 4 (b) convincingly demonstrates that our approach surpasses alternative methods, especially when dealing with extremely limited datasets, emphasizing its paramount importance in the context of scarce data. For a dataset of 400 samples, our transfer learning approach achieves improvements of 8%, 38%, and 43% when compared to 2D-CNN, 1D-CNN, and GRU, respectively. The GRU and 1D-CNN models trained with SOP time series require additional data to achieve good-enough performance as their accuracy is below 84% for 1000 samples. Converting the SOP time series into images helps to achieve better performance for event classification by providing an improvement of more than 10%.

We finally compare the MobileNet pre-trained model to different state-of-the-art pre-trained models namely VGG16 [9], Inception [10], and DarkNet [11], in terms of test accuracy and inference time. Fig. 4 (c) shows that MobileNet achieves better accuracy while consuming less time in inference due to its streamlined architecture. This helps to significantly reduce the number of the model parameters (variables of the model learned from the data during training including weights) hence computational complexity.

Conclusions

To the best of our knowledge, this is the first effort to use a deep CNN, pre-trained on millions of images, as a feature extractor for images produced from SOP recordings. A 96.3% accuracy is attained with a small dataset of 400 samples utilizing a pre-trained model (MobileNet) as a feature extractor. The results show that (i) converting SOP recordings into images using GADF can preserve fine-grained details, (ii) feature mappings learned from a massive amount of generic data in a deep neural network can represent the images derived from SOP signals very effectively, and (iii) our approach could help improving performance and accelerating the training.

This work has been supported by the EU through the AINET-ANTILLAS research project.

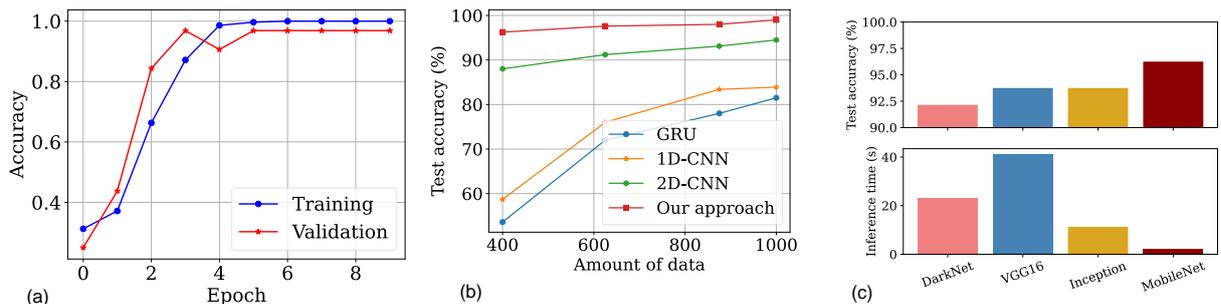


Fig. 4: (a) Training and validation curves, (b) Comparison of different approaches in terms of test accuracy for different amount of data, (c) Performance comparison of different pre-trained models.

References

- [1] J. Zweck and C.R. Menyuk, "Detection and Mitigation of Soft Failure due to Polarization-Mode Dispersion in Optical Networks," Optical Fiber Communication Conference (OFC), paper OFG5 (2006).
- [2] K. Abdelli, et al., "Machine-learning-based anomaly detection in optical fiber monitoring," in Journal of Optical Communications and Networking, vol. 14, no. 5, pp. 365-375 (2022).
- [3] F. Boitier et al., "Proactive Fiber Damage Detection in Real-time Coherent Receiver," European Conference on Optical Communication (ECOC), pp. 1-3, Gothenburg, Sweden (2017).
- [4] V. Lemaire, et al., "Proactive fiber break detection based on quaternion time series and automatic variable selection from relational data," International Workshop on Advanced Analysis and Learning on Temporal Data, pp 26–42 (2019).
- [5] K. Guan, et al., "Efficient Classification of Polarization Events Based on Field Measurements," Optical Fiber Communication Conference (OFC), paper Th3D.7 (2020).
- [6] J. Deng, et al., "ImageNet: A large-scale hierarchical image database," IEEE Conference on Computer Vision and Pattern Recognition, pp. 248-255, Miami, FL, USA (2009).
- [7] A. G. Howard, et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," ArXiv abs/1704.04861 (2017).
- [8] W. Zhiguang and T. Oates, "Encoding Time Series as Images for Visual Inspection and Classification Using Tiled Convolutional Neural Networks," In Workshop at the Twenty-Ninth AAAI Conference on Artificial Intelligence (2015).
- [9] S. Karen and Z. Andrew, "Very Deep Convolutional Networks for Large-Scale Image Recognition," CoRR abs/1409.1556 (2014).
- [10] C. Szegedy, et al., "Rethinking the Inception Architecture for Computer Vision," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2818-2826, Las Vegas, NV, USA (2016)
- [11] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," ArXiv abs/1804.02767 (2018).