Trains Detection Using State of Polarization Changes Measurement and Convolutional Neural Networks

DEJDAR, P.; MYŠKA, V.; MÜNSTER, P.; BURGET, R.

Accepted manuscript

Trains Detection Using State of Polarization Changes Measurement and Convolutional Neural Networks

Petr Dejdar[1,2], Vojtech Myska[1], Petr Munster[1,2], Radim Burget[1]

[1]Brno University of Technology, FEEC, Dept. of Telecommunications, Brno, Czech Republic

[2]Czech Optical Solutions, Zlin, Czech Republic

Abstract—**Fiber optic infrastructure security is of growing interest. The current distributed sensor systems are robust and expensive solutions, and their practical applications are uncommon. Research into simple and cost-effective solutions based on changes in the state of polarization is crucial. This paper expands the use of a vibration sensor based on the sensing of rapid changes in the state of polarization (SOP) of light in a standard single-mode optical fiber by using a convolutional neural network to detect trains running along the optical fiber infrastructure. It is a simple system that determines ongoing events near the optical fiber route by simply determining the signal boundaries that define the idle state. By using a neural network, it is possible to eliminate the distortion caused by the temperature changes and, for example, to improve detection in the the zones where the vibrations are not strong enough for a simple threshold resolution.**

Index Terms—**artificial intelligence, machine learning, optical fiber sensor, state of polarization changes, vibration.**

I. Introduction

Most of distributed fiber optic sensors are based on back-reflection or interferometry. These systems are extremely costly and their sampling rate is low. Detection of the state of polarization changes of a signal is a highly effective and cost-effective solution. Therefore, it is necessary to design cost-effective systems for sensing optics. This field is currently relevant to a wide range of applications and non-traditional security options are now being sought.

The field of artificial intelligence has seen a huge boom in recent years that has affected many other industries, including optics. These two mentioned fields are combined in many works. For example, [1] deals with the detection of steps in the seismic data of cities using neural networks. Another work [2] combines the fields of optics and AI in order to create a system for monitoring pipeline safety.

However, the combination of these two disciplines is rather unique. For example, some applications can use artificial neural network to recognize the modulation format used in fiber optic transmission [3]. In the field of sensors, it is possible to find experiments detecting fiber bending in data networks [4] or the detection of various types of partial discharge signals in power cables using optical fibers [5]. In the field of detection of vibrations sensed by optical fibers, detection of the effect of wind on the perimeter security sensor has been investigated in order to obtain more accurate results from the sensor [6].

However, the application of a similar range using a polarimeter and artificial intelligence is rather rare.

In this paper, we present an algorithm based on convolutional neural networks (CNN) [7] for the detection and classification of idle state or train passage based on the measurement of a railroad using an optical fiber sensing system with a commercial 4-diode polarimeter. The polarimeter uses logic with electronic amplifiers and transistors to amplify the state of polarization.

II. Optical Fiber Sensing System

The basis of the sensor system is a commercial polarimeter that has its own designed hardware amplification capabilities using transistor logic. To generate the light wave, a Distributed Feedback Laser (DFB) and a 20 km long fiber optical route (connected as a loop) buried in the non-freezing area of the railroad are used as shown in Figure 1. There is only one exception where the fiber is attached directly to the bridge structure. A commercial 4-diode polarimeter was implemented to detect the state of polarization changes. It is necessary to choose a device capable of sampling at least 4 channels with a sampling frequency of at least 1 kS/s.

The myRIO-1950 device (National Instruments, USA) was chosen for this particular application. This device can scan up to 12 channels at a maximum sampling rate of 500 kS/s using only 1 channel. When using multiple channels, the maximum sample rate is divided by the number of channels, so with maximum optimization, we can reach up to 125 kS/s. However, in practice, this value will be significantly lower due to some pre-processing of the data by the myRIO itself. The maximum value that could be programmed was 6.5 kS/s and was reduced to 5 kS/s during the experiment due to a large amount of data. Thus, a total of 20,000 samples per second were recorded.
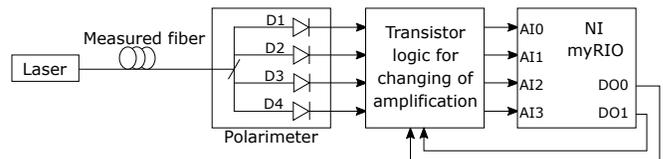

Fig. 1. Experimental setup for measuring the state of polarization changes.

The myRIO device was programmed in LabVIEW 2018 with the myRIO toolkit. Analog outputs from an inline polarimeter are connected to analog inputs of myRIO through transistor logic. Two digital output pins have been set for myRIO that can be used to control four different gain states ($10\times$, $5\times$, $1\times$ and $0.9\times$). This setting is carried out using server variables. The program reads the data from analog inputs, and these 4 values are then inserted into the array and sent to the operating PC using the TCP protocol. Communication with the operating PC is possible via USB that emulates an Ethernet card. For later use, it is also planned to eliminate the service PC and send data directly to the server.

The program on the operating PC is also created in Lab-VIEW and only reads the values from myRIO and saves them continuously to the Technical Data Management Solution (TDMS) file. It is also possible to set the gain, sample rate, acquisition time and name of the file using the software. The PC uses server variables to communicate with myRIO. Thus, preprocessing and subsequent sending of data is ensured on the myRIO side and subsequent capture and storage of data in files for later processing on the PC side. A full description of the programs is provided in [8].

Fig. 2. Example of polarization states of 2 of 4 channels for long-distance train R972 in first direction (used in data set).
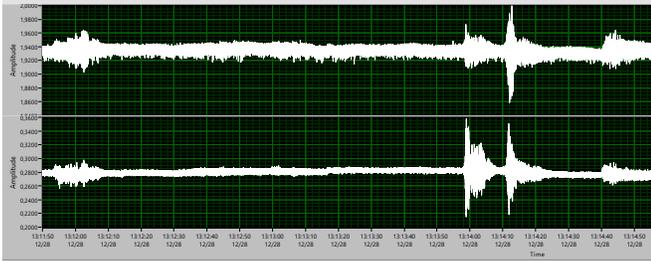

Fig. 3. The same R979 train in the opposite direction.

An example of a two-hour elevation is shown in Figure 8. From the figure, it is possible to determine several amplitude deviations that show the capture of vibrations of passing trains. Initially, it was assumed that it would be possible to detect the passage of a train based on tresholding. However, the polarization value changes depending not only on the vibrations but also on the temperature. Therefore, it is necessary to perform at least normalization, but we can determine only the passages on the bridge, because of the largest deviation value in this case. By contrast, using artificial intelligence, it is possible to use the properties of signals where the entire passage of the

TABLE I
AN OVERVIEW OF THE SIZE OF THE INDIVIDUAL DATA SETS.

Data set Class	Training	Validation	Testing
Idle	758	110	221
Train	2 051	355	771

train is recorded. Therefore, it is possible to determine whether a train is currently passing through a fiber section, or whether the track is free. Figures 2 and 3 shows the passages of the same long-distance train in both directions. As can be seen by the naked eye, the direction of the movement of the train can be determined.

III. NEURAL NETWORK

An overview of the machine learning pipeline of the proposed detection system is shown in Figure 4.
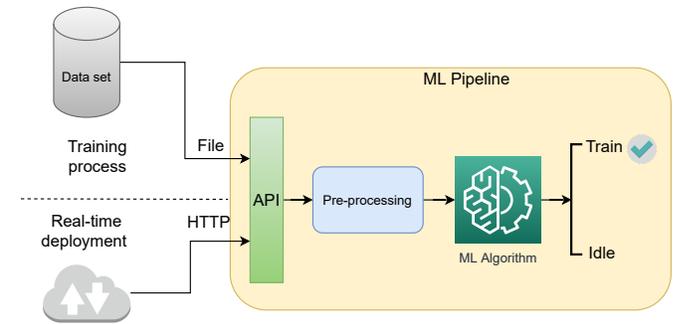

Fig. 4. Overview of the data flow in the train detection system.

This system is designed for easy real-time deployment. It is clear that the only change from the training process will consist of the use of another end point of the system. Rather than retrieving the data from the data set, the data will be received from the Internet.

A. Data set

Considering the unique focus of this experiment, it was necessary to create our own data set, because no similar data set is currently available. The result is a new privately collected data set containing the two monitored categories. It contains 18 and 58 data files with the idle state, respectively, with the movement of the train. Every file contains the results of the measurements from 4 channels. The values given in Table I correspond to the number of data samples after division into smaller parts.

B. Data pre-processing

For further post-processing, it was necessary to modify the structure of the acquired data because the raw measurement data obtained from the polarimeter were not suitable for use as the input for the neural network. This is mainly due to the following three reasons: the length of the raw data stream, the high sampling frequency and the range of the values that the input data can take. The process is illustrated in Figure 5. The acquired measurement data signals were too long and the

sampling frequency was high for classification process. First, the raw data signal stream was divided into smaller parts (2 000 samples) and then subsampled 4 times. To maintain a certain context, individual parts of the data overlap by 25 % (500 samples). This is a technique of the so-called sliding window that enables the prevention of overtraining of the model and also increases the generalization ability. It is often used for the classification of time series using recurrent or convolutional neural networks [9]. The last operation in data pre-processing is the scaling of the data in the range from 0.1 to 0.9.
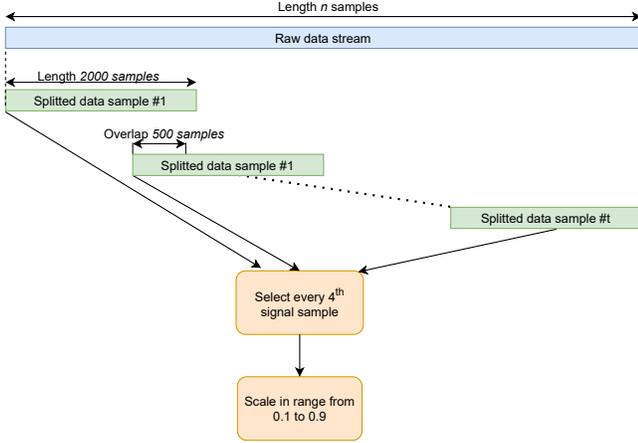

Fig. 5. Raw data stream pre-processing. The input data stream is marked in blue. In the next step, this stream is split into smaller parts marked in green. Furthermore, these parts are processed by the functions marked in orange.

C. Neural network architecture

This proposed approach tries to learn a mapping function F between the input data $\vec{d} = (l, n)$, where l is the length of the samples (500) and n is the number of channels (4) and ground truth y. This function is realized by a neural network based on the CNN. Although recurrent neural networks (RNN) are used to process time series, the decision whether to use an RNN or a CNN was influenced by the size of the dataset and the fact that CNN are suitable not only for image processing but also for text or time series such as our data. The proposed simple model of the NN is illustrated in Figure 6. In the network, 2D convolutional layers that need input with shape $channels, rows, cols$ are used. Thus, the first layer is a reshape layer that adds a channel to the input data. Our model can be divided into two main parts. The first one is three consecutive 2D convolutional layers. All of these have the same kernel, stride and features map setting. The second part of the model is composed of three dense layers between which are inserted dropout layers, the aim of which is to prevent the overfitting of the model. The last dense layer that has 2 units with a combination of the softmax activation function is the output layer.

IV. RESULTS

Because this is a preliminary study in which a data set of relatively limited size was collected, it is necessary to use
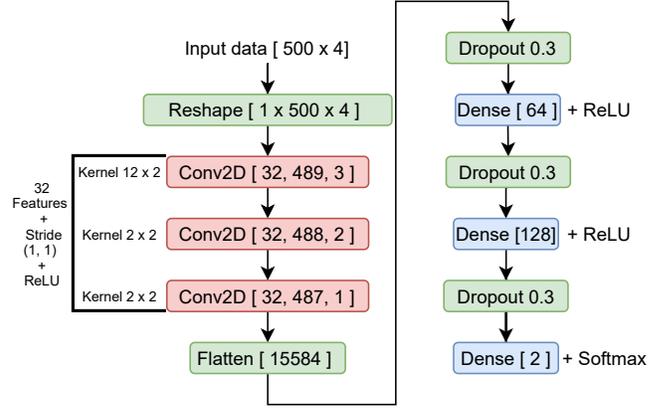

Fig. 6. Architecture of the proposed convolutional neural network including the settings of the appropriate layers.

techniques and experimental settings that can be used to deal with this problem. During the training of the proposed model, it was necessary to adjust the settings of the hyperparameters of the neural network such as the learning rate and number of iterations in order to avoid significant overfitting of the network due to the ratio of its complexity and the amount of available data. In the experiment, we used k-fold cross-validation [10] with $k = 10$ splitting into 10 groups. The learning rate was reduced to 0.0006777 and the model was trained in only 30 iterations. Based on the graph shown in Figure 7, it is clear that nevertheless, the model showed a slight overfitting.
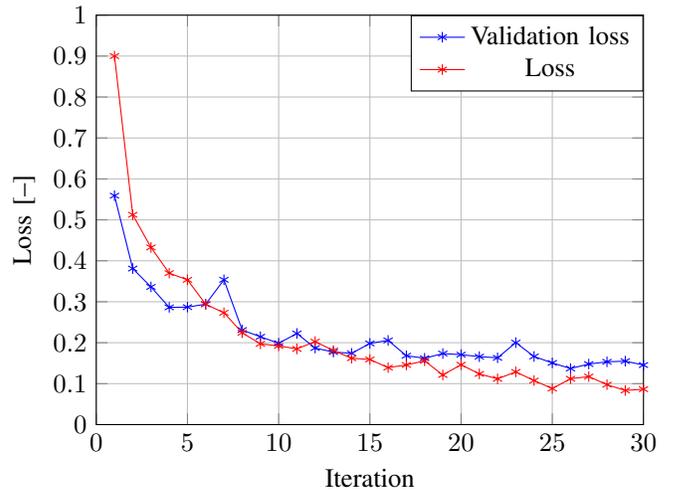

Fig. 7. Graphical course of the dependence of the loss function on the training process. The dependence shows a slight overfitting of the model.

The resulting trained model was tested on the data that were not used during the training process. Table II summarizes the achieved results for the task of binary classification of the data obtained from a polarimeter that was classified with a relatively high success rate of 98.71 %. This value was calculated by averaging all 10 cross-validation groups.
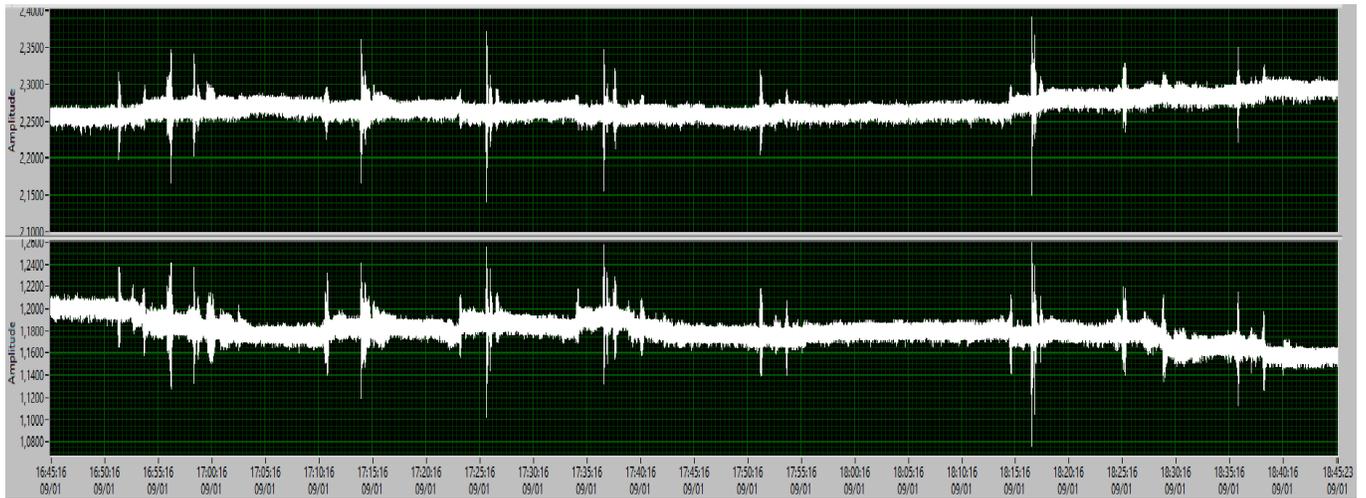
Fig. 8. Example of the polarization state in the measurement fiber for 2 hours.

TABLE II
CONFUSION MATRIX. RESULTS OF THE MODEL ON THE TEST DATA SET.

		Predicted	
	Class	Idle	Train
Actual	Idle	218	3
	Train	9	702
Accuracy		98.71%	

V. CONCLUSION

Deep learning is an ever-evolving sub-area of artificial intelligence that has been applied in a wide range of disciplines. In this preliminary study, the fields of artificial intelligence and optics were combined for the purpose of detection and classification of trains on the railroad from the data acquired using a polarimeter.

Based on the obtained data, the use of a commercial 4-diode polarimeter to detect the state of polarization appears to be a good solution. Cost-effective myRIO was used for data acquisition and was fully implemented. Based on a closer examination of the dataset, it is also possible to determine whether the train in a given case is a regional, long-distance, or freight train. This is a very good basis for expanding this work and possibly designing a sensor system that can be used in practice to strengthen the security of the railway tracks.

In this experiment, a data preprocessing process and a CNN-based neural network model were designed. Furthermore, this model was trained on its own data set and tested on independent data. The obtained model can classify idle state and train passage with 98.71% accuracy. Our method was designed in consideration of its possible deployment in real time. The main aim was therefore to allow classification with fast reaction time. This is achieved by dividing the input data stream into short sequences.

REFERENCES

[1] S. Jakkampudi, J. Shen, W. Li, A. Dev, T. Zhu, and E. R. Martin, "Footstep detection in urban seismic data with a convolutional neural network," *The Leading Edge*, vol. 39, no. 9, pp. 654–660, 2020.

[2] R. Peng, Z. Liu, and S. Li, "Perimeter monitoring of urban buried pipeline subject to third-party intrusion based on fiber optic sensing and convolutional neural network," in *Eleventh International Conference on Information Optics and Photonics (CIOP 2019)*, H. Wang, Ed., vol. 11209, International Society for Optics and Photonics. SPIE, 2019, pp. 464 – 471.

[3] Q. W. Zhang, M. Liu, H. Zhou, J. Chen, B. Y. Cao, Y. X. Song, J. J. Zhang, Y. C. Li, and M. Wang, "Artificial neural network based modulation identification for elastic optical networks," in *2018 23rd Opto-Electronics and Communications Conference (OECC)*, 2018, pp. 1–2.

[4] T. Tanaka, W. Kawakami, S. Kuwabara, S. Kobayashi, and A. Hirano, "Intelligent monitoring of optical fiber bend using artificial neural networks trained with constellation data," *IEEE Networking Letters*, vol. 1, no. 2, pp. 60–62, 2019.

[5] Q. Che, H. Wen, X. Li, Z. Peng, and K. P. Chen, "Partial discharge recognition based on optical fiber distributed acoustic sensing and a convolutional neural network," *IEEE Access*, vol. 7, pp. 101 758–101 764, 2019.

[6] Wu Liang, "Study on the fiber-optic perimeter sensor signal processor based on neural network classifier," in *IEEE 2011 10th International Conference on Electronic Measurement Instruments*, vol. 1, 2011, pp. 93–97.

[7] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[8] P. Dejdar, "Continuous data acquisition by myrio used for measurement of polarization transient effects," in *Proceedings I of the 26th Conference STUDENT EEICT 2020*, 2020, conference paper, pp. 499–503.

[9] V. Suresh, P. Janik, J. Rezmer, and Z. Leonowicz, "Forecasting solar pv output using convolutional neural networks with a sliding window algorithm," *Energies*, vol. 13, no. 3, p. 723, 2020.

[10] F. J. Anscombe, *Frederick Mosteller and John W. Tukey: A Conversation*. New York, NY: Springer New York, 2006, pp. 647–660.