POLITECNICO DI TORINO
Repository ISTITUZIONALE

Deep Learning Based Early Earthquake Detection through Terrestrial Optical Networks

Original
Deep Learning Based Early Earthquake Detection through Terrestrial Optical Networks / Usmani, Fehmida; Awad, Hasan; Straullu, Stefano; Bratovich, Rudi; Virgillito, Emanuele; Aquilino, Francesco; Proietti, Roberto; Curri, Vittorio. - (2025), pp. 1-6. (2025 25th Anniversary International Conference on Transparent Optical Networks (ICTON) Barcelona (Spa) 06-10 July 2025) [10.1109/ICTON67126.2025.11125067].

Availability:
This version is available at: 11583/3002588 since: 2025-09-09T10:15:04Z

Publisher:
IEEE

Published
DOI:10.1109/ICTON67126.2025.11125067

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2025 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

26 April 2026

Deep Learning Based Early Earthquake Detection through Terrestrial Optical Networks

Fehmida Usmani
National University of Computer & Emerging Sciences Pakistan
fehmida.usmani@polito.it

Hasan Awad
DET, Politecnico di Torino
Torino, Italy
hasan.awad@polito.it

Stefano Straullu
LINKS Foundation
Torino, Italy
stefano.straullu@linksfoundation.com

Rudi Bratovich
SM-Optics
Cologno Monzese, Italy
rudi.bratovich@sm-optics.com

Emanuele Virgillito
DET, Politecnico di Torino
Torino, Italy
emanuele.virgillito@polito.it

Francesco Aquilino
LINKS Foundation
Torino, Italy
@linksfoundation.com

Roberto Proietti
Politecnico di Torino
Torino, Italy
roberto.proietti@polito.it

Vittorio Curri
Politecnico di Torino
Torino, Italy
vittorio.curri@polito.it

***Abstract*—We demonstrate the use of existing terrestrial optical networks as a smart sensing grid for early earthquake detection we integrate real ground displacement data from seven earthquakes, magnitudes ranging from four to six, to simulate the strains within fiber cables and collect large set of light's polarization evolution data to train the model to detect Primary Waves (P waves) arrivals that precede earthquakes' destructive Surface waves. . The main idea of our approach is to deploy a fast, accurate and reliable trained deep learning model. We evaluated the performance of LSTM and GRU models on experimentally emulated data collected from a 38 km deployed fiber link in Turin, Italy. Our results demonstrate that the GRU model consistently outperformed the LSTM model with 99% recall for P-wave detection.**

***Index Terms*—Fiber optic sensing, deep learning, early earthquake detection, smart sensing grid.**

I. Introduction

Significant earthquakes have continued to cause widespread devastation in recent years, particularly in seismically active regions. Recently, the Turkey-Syria earthquake in 2023 was one of the deadliest earthquakes that occurred on February 6, 2023, with a magnitude of 7.8. More than 50,700 people lost their lives and over 107,000 were injured, and cities in southern Turkey and northern Syria suffered significant damage. This incident highlights the ongoing threat that earthquakes pose to human life and infrastructure. An early earthquake warning (EEW) system is crucial in order to reduce the destructive effects of earthquakes. By detecting seismic activity in real-time and providing seconds to minutes of early warning, EEW systems enable people and organizations to take life-saving measures, such as evacuating buildings, stopping trains, or turning off hazardous equipment. This short timeframe can drastically lower the number of fatalities, damage to infrastructure, and financial losses, particularly in heavily populated urban areas. In seismically active areas, EEW system is essential for improving resilience and hazard preparedness. The global distribution of seismic stations is sparse, resulting in inadequate spatial resolution for effectively analyzing seismic activity. This problem is more severe in remote and under-monitored places. Expanding and maintaining a more broad and consistent network of these stations has substantial cost challenges, making it difficult to adopt even for economically developed countries. This strategy is not feasible on a broader scale due to the high installation and maintenance expenses. Consequently, there is an increasing demand for more effective and economical solutions to improve seismic monitoring without incurring the significant costs of installing extra physical sensors.

Recently, the use of already deployed optical fiber networks as large-scale sensors has attracted a lot of attention from the research community. Traditionally deployed for high-capacity data transmission, wavelength division multiplexing (WDM)-based optical networks are now being investigated for their potential to serve as vast, cost-effective sensing grids. This novel method uses the current telecommunications infrastructure to carry out sensing by examining physical parameters, such as frequency, intensity, phase, and state of polarization (SOP) fluctuations, currently monitored by network nodes. Most fiber optic sensing (FOS) applications utilize specialized equipment and commonly depend on distributed acoustic sensing (DAS), phase-sensitive optical time-domain reflectometry (OTDR), or interferometric configurations. A more cost-effective approach is to analyze SOP fluctuations in the existing optical channels rather than relying on specialized equipment. This approach provides a scalable, cost-effective solution that enables sensing over large terrestrial regions without the need for specialized equipment, even if it is not as spatially accurate as dedicated distributed fiber optic sensing systems. SOP monitoring via intensity-modulated and coherent optical channels has demon-

979-8-3315-9777-1/25/$31.00 © 2025 IEEE

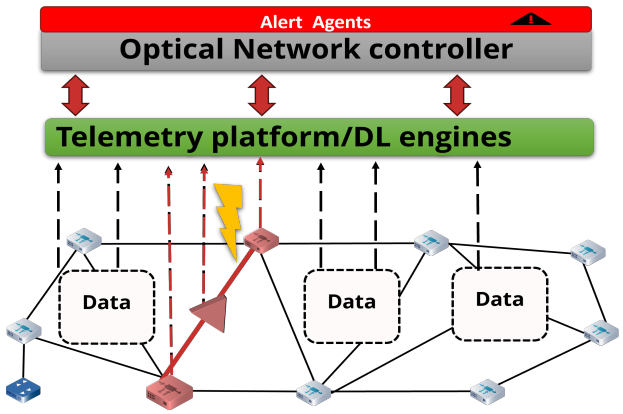

Fig. 1: Architecture of smart sensing network

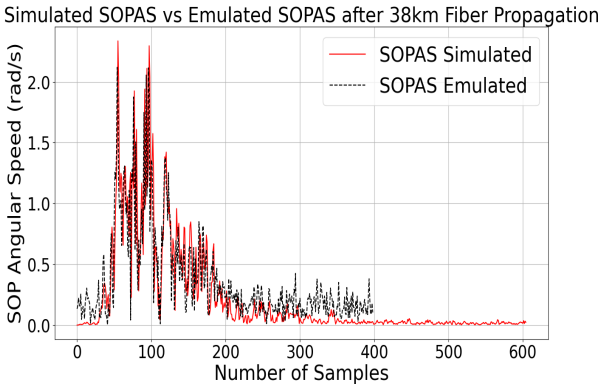

Fig. 2: Emulated vs Simulated SOPAS

strated potential for identifying mechanical stress induced by anthropogenic activities. SOP-based solutions are presented in [1] and [2] for underwater earthquake detection, while [3] illustrates their application in road traffic monitoring. Although phase-based monitoring has been widely applied, a challenge remains in developing reliable methods to detect specific events, like those caused by mechanical stress, across various monitored quantities.

Recently, the application of artificial intelligence techniques has greatly enhanced the sensing capabilities of FOS, allowing for the identification and segmentation of specific events by leveraging an understanding of their defining characteristics. Deep learning (DL) methods are well-suited to address this challenge by learning the critical features that characterize such events, enabling accurate and scalable event detection within fiber sensing networks. Several studies have highlighted the advantages of integrating machine learning with FOS, such as vehicle detection and classification [1], real-time intruder detection in railway systems [2], and earthquake detection [3]. Building on these advancements, our previous work [4] focused on road traffic detection using a Long Short Term Memory (LSTM) autoencoder to analyze SOP variations in deployed metropolitan fiber cables. In [5], we have pioneered a smart grid fiber sensing approach that employs a Bidirectional Gated Recurrent Unit (Bi-GRU) model with an attention mechanism for earthquake early warnings, enabling reliable primary waves (P-wave) detection promptly based on SOP measurements and leveraging terrestrial optical networks with realistic synthetic earthquake waves. The P-wave of an earthquake travels through the Earth's interior and arrives before the destructive surface waves that propagate along its surface. Surface waves carry the most energy and are typically the main contributors to destruction. For this work, We synthetically generated realistic ground motion displacement waves based on a 4.9 magnitude earthquake that occurred in the Marradi region of central Italy. Furthermore, in [6], we extracted actual ground motion data from the same 4.9 magnitude earthquake earthquake in Marradi. By applying the same methodology, we successfully detected the arrival of primary waves with equivalent accuracy. In [7], we expand our focus to treat the entire network as a sensing grid by leveraging six fiber cables as substitutes for six seismic stations and training a temporal convolutional network (TCN) on four nodes and testing it on two nodes, thereby enhancing the model's predictive capabilities. Furthermore, we utilize the interconnected optical mesh network in [8] across three municipalities as a comprehensive sensing and epicenter localization grid. By leveraging real ground displacement data from seven earthquakes with magnitudes ranging from 4 to 6, we employed a Temporal Fusion Attention Network (TFAN) based on a neural network architecture that integrates TCN, LSTM, and an attention mechanism for detecting the arrival of P-Waves.

A key distinction of this work is our analysis of the performance of various deep learning models, initially trained on a simulated dataset and validated on experimentally emulated data over a 38 km deployed fiber link deployed in the city of Turin, Italy, for early earthquake detection in the presence of real-world propagation challenges.

II. Dataset Collection and Preprocessing

A. Simulated Dataset

We obtained real ground displacement data for seven local earthquakes in the Modena region, Italy, from the Italian National Institute of Geophysics and Volcanology (INGV) [9], with magnitudes ranging from 4 to 6. Using a Python-based waveplate model [8], we simulated the strain effects on fiber cables of 10 km each, positioned at the same geographical coordinates as the seismic stations. The waveplate model converted earthquake displacement values into nano strain according to the standard iDAS method [10], where 116 nm of ground displacement corresponds to 11.6 nano strain on the fiber. Polarized light at 45 degrees was used for 50 simulations per earthquake, each with randomized waveplate orientations. The State of Polarization (SOP) evolution was captured for each configuration, transformed into SOPAS values, and compiled into 350 simulation files, providing a comprehensive dataset (210,000 data samples) for training and validating our DL models across all earthquake events.

B. Emulated Dataset

We generate emulated SOP Stokes parameters using a tunable laser operating at a wavelength of 1550 nm and an

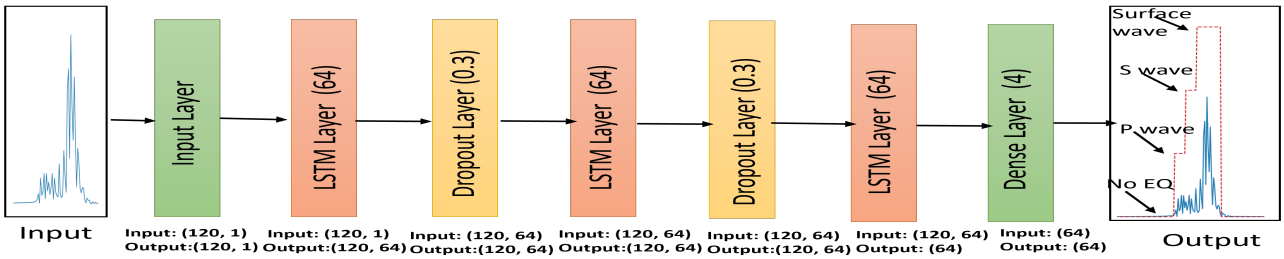

Fig. 3: Architecture of LSTM model

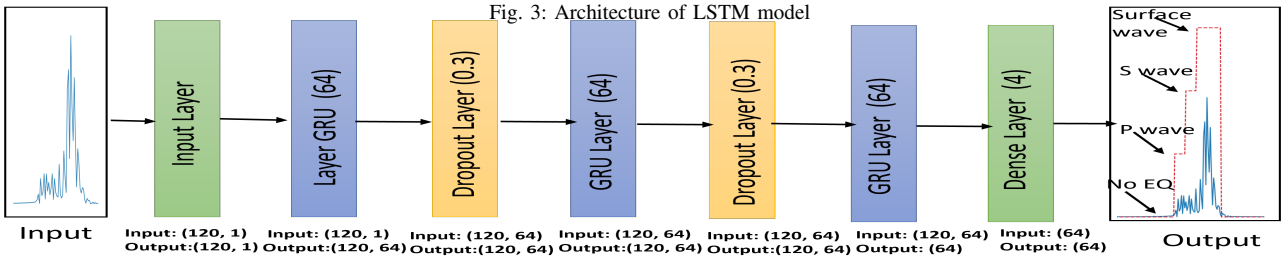

Fig. 4: Architecture of GRU model

output power of 6 dBm, which effectively replicates those produced by the Python-based Waveplate model. This optical signal is directed into an optical scrambler consisting of seven plates with randomly oriented configurations. The orientations of these plates are adjusted based on the voltage applied to each, aligning them with the target SOP specified for the scrambler. A polarimeter measures the SOP after the signal has been scrambled, giving feedback on the difference between the measured and target SOP. After that, the scrambler adjusts the plate orientations using a minimization function to progressively align the plates with the desired SOP. This layout is known as a Back-to-Back (B2B) system.The scrambler successfully learns to generate the intended SOP evolution driven by earthquake strain values. The seven plates are then subjected to the identical voltage settings as the scrambler to replicate the B2B-emulated Stokes characteristics. These are then enabled to spread over a distance of 38 kilometers via single-mode fiber (SMF) that has been deployed in Turin, Italy. To simplify the analysis, we further calculate the State of Polarization Angular Speed (SOPAS) from the SOP. The simulated and B2B-emulated signals are shown in Fig. 2. We generated 19,950 data samples for the emulated SOPAS dataset, emulating the wave patterns of an earthquake.

III. Deep Learning-based Smart Sensing Grid for Early Earthquake Detection

A. Smart Sensing Grid

The architecture of our proposed smart grid optical sensing network for early earthquake detection is illustrated in Fig 1. We leverage the software-defined network (SDN) based approach to exploit the streaming telemetry paradigm for real-time data transmission monitoring. The centralized Optical network controller (ONC) manages the network by coordinating with network elements (NEs) to enable streamlined and efficient control. The NEs are continuously collecting the telemetry data and communicating with ONC for network control and management purposes. The integrated telemetry platform employs a deep learning model to analyze variations in SOP data, enabling the ONC to issue early earthquake warnings upon detecting the arrival of P waves.

B. Deep Learning Models

This section presents our proposed framework for analyzing time series data to detect and quantify seismic events using the smart sensing grid approach. The framework involves using LSTM and GRU models separately, enabling a comparative evaluation of their individual detection accuracy performance.

Input Data Sequence: Preparing the earthquake dataset for LSTM and GRU models requires converting the original time-series data into a compatible 3-dimensional format. Both LSTM and GRU models require input in the shape (samples, time steps, features). The original univariate (single-feature) seismic data is divided into fixed-length sequences, represented as $(X_1, X_2, X_3, \ldots, X_n)$. For each sequence X, a designated time window contains observations $(x_1, x_2, x_3, \ldots, x_t)$, where each observation $x_t \in \mathbb{R}^1$ represents a single feature value (SOPAS) at a given time instance t. This 3D format, (240645, 120, 1), ensures the data meets the input requirements for both LSTM and GRU models, enabling them to learn sequential patterns effectively. Each dimension in the array represents a specific aspect of the input. The sample dimension provides the total number of sequences for training, Time steps represent the length of each sequence, and Features indicate the single feature value in each step.

Long Short Term Memory (LSTM): Recurrent Neural Networks (RNNs) are commonly used for modeling time series data due to their capability to capture temporal dependencies. However, their effectiveness for learning long-term patterns is often hindered by the vanishing gradient problem. To address this limitation, LSTMs networks were developed, incorporating a memory cell and gating mechanism that mitigate the vanishing gradient issue. The LSTM unit relies on three gates—the input gate, output gate, and forget gate—to control information flow [11]. Specifically, the input and output gates manage the activation signals entering and leaving the cell, while the forget gate resets the cell memory when certain

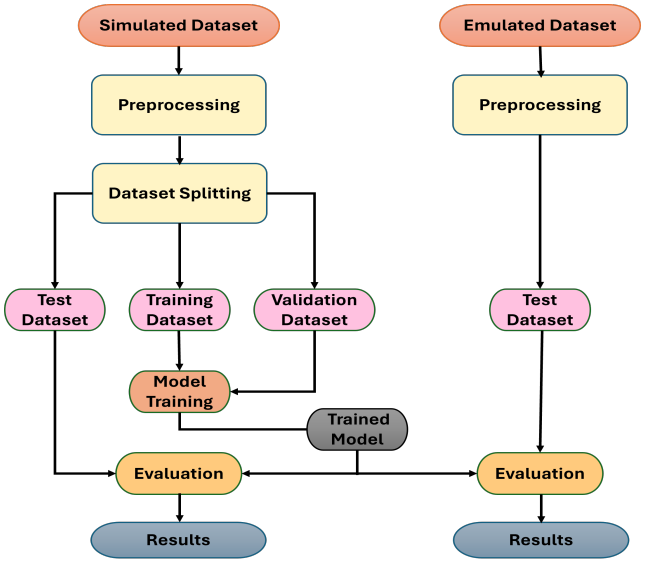

Fig. 5: Flow Diagram for DL models training, validation and testing

information is no longer needed. Let $\mathbf{x}_t$ represent the input time-series data at time t. The values of the input gate (i_t), output gate (o_t), and forget gate (f_t) are computed as follows:

$$i_t = \sigma(W_i[\mathbf{h}_{t-1}, \mathbf{x}_t] + b_i), \tag{1a}$$

$$o_t = \sigma(W_o[\mathbf{h}_{t-1}, \mathbf{x}_t] + b_o), \tag{1b}$$

$$f_t = \sigma(W_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + b_f), \tag{1c}$$

where W_i, W_o, and W_f are the weight matrices for the input, output, and forget gates, respectively, and b_i, b_o, and b_f are their respective bias terms. Here, $\mathbf{h}_{t-1}$ denotes the hidden state from the previous time step, and $\sigma(\cdot)$ is the sigmoid activation function. The hidden state $\mathbf{h}_t$ and cell state $\mathbf{c}_t$ at time t are then updated as:

$$\mathbf{c}_t = i_t \odot \tilde{\mathbf{c}}_t + f_t \odot \mathbf{c}_{t-1} \tag{2}$$

$$\mathbf{h}_t = o_t \odot \tanh(\mathbf{c}_t) \tag{3}$$

where $\odot$ denotes the element-wise (Hadamard) product, and $\tanh(\cdot)$ is the hyperbolic tangent function. The candidate cell state $\tilde{\mathbf{c}}_t$ is given by

$$\tilde{\mathbf{c}}_t = \tanh(W_c[\mathbf{h}_{t-1}, \mathbf{x}_t] + b_c),$$

with W_c and b_c being the weight matrix and bias for the cell state update. The architecture of our LSTM model is shown in the Fig 3.

We employ three LSTM layers with 64 units each to effectively capture the temporal dependencies in the seismic data. We carefully place two dropout layers with a dropout rate of 0.3 after the first and second LSTM layers to reduce overfitting and improve generalization. This configuration helps prevent the model from relying too heavily on specific training data patterns. Finally, a dense layer with a softmax activation function is used to classify the data into four categories: no earthquake, P wave, S wave, and surface wave. Grid search and empirical testing were combined to get the best hyperparameters for our model. In order to systematically explore combinations that would best capture the temporal patterns of seismic activity, grid search was first used to evaluate a range of potential values for the number of LSTM units, dropout rates, and the number of layers. Iterative experimentation led to the selection of three LSTM layers with 64 units each, striking a balance between computing efficiency and model complexity.The dropout rate of 0.3 was chosen after testing various values to identify the point where generalization improved without significant loss in model performance.

Gated Recurrent Neural Network (GRU): GRU was introduced as a simplified alternative to the LSTM network, aimed at reducing the architectural complexity and minimizing the number of parameters involved. Unlike LSTM, the GRU uses only two gates: the update gate and the reset gate. The update gate combines the roles of the input and forgets gates found in LSTM, allowing control over how much of the previous memory should be retained in the new state. Meanwhile, the reset gate functions similarly to LSTM's forget gate, determining which parts of past information can be discarded when they are no longer needed [12].

The update gate u_t and the reset gate r_t in the GRU are defined as follows:

$$u_t = \sigma(W_u[\mathbf{h}_{t-1}, \mathbf{x}_t] + b_u) \tag{5}$$

$$r_t = \sigma(W_r[\mathbf{h}_{t-1}, \mathbf{x}_t] + b_r) \tag{6}$$

Here, W_u and W_r represent the weight parameters for the update and reset gates, respectively. The parameters b_u and b_r in equations (5) and (6) denote the corresponding bias terms. The term $\mathbf{h}_t$ refers to the hidden state vector or output state vector. At time step t, the hidden state $\mathbf{h}_t$ is expressed as follows:

$$\mathbf{h}_t = u_t \odot \tilde{\mathbf{h}}_t + (1 - u_t) \odot \mathbf{h}_{t-1} \tag{7}$$

where $\tilde{\mathbf{h}}_t$ is given by:

$$\tilde{\mathbf{h}}_t = \tanh(W_h[\mathbf{h}_{t-1}, \mathbf{x}_t] + b_h) \tag{8}$$

In these equations, W_h and b_h represent the weight and bias parameters, respectively. where $\tilde{\mathbf{h}}_t$ is given by:

$$\tilde{\mathbf{h}}_t = \tanh(W_h[\mathbf{h}_{t-1}, \mathbf{x}_t] + b_h) \tag{8}$$

In these equations, W_h and b_h represent the weight and bias parameters, respectively. The architecture of our proposed GRU model is shown in the Fig 4.

The architecture of the proposed GRU is shown in Fig. 4. We have employed the same architecture for GRU as that used for LSTM networks to facilitate a fair comparison of their performance. A unified architectural framework enables us to assess the training dynamics, and convergence rates with deeper insights into the operational characteristics of each model.

IV. Performance Analysis

In this section, we outline the model training process, discuss the evaluation metrics used to assess performance, and present the results.

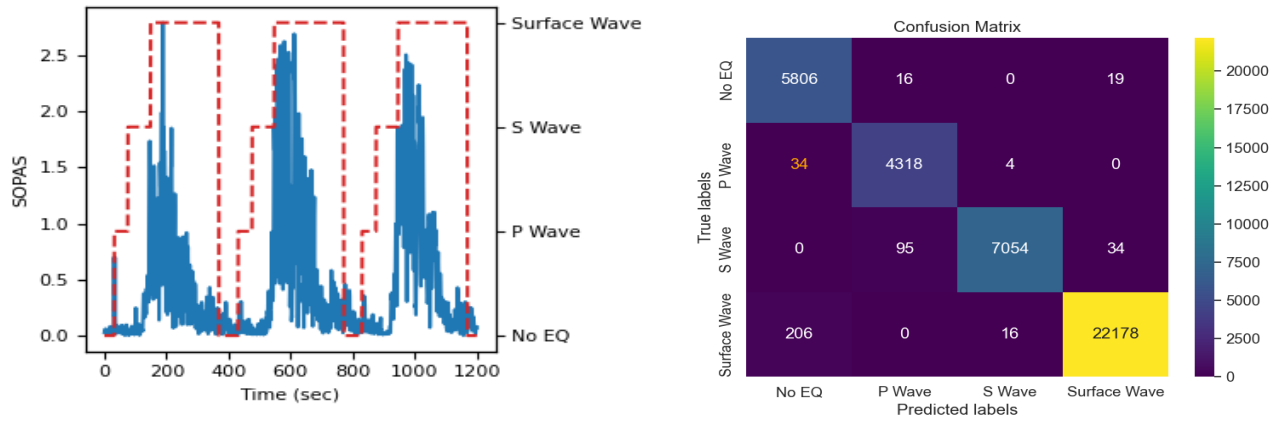

Fig. 6: (a) LSTM Predictions on Emulated Dataset, (b) LSTM Confusion Matrix for Emulated Dataset

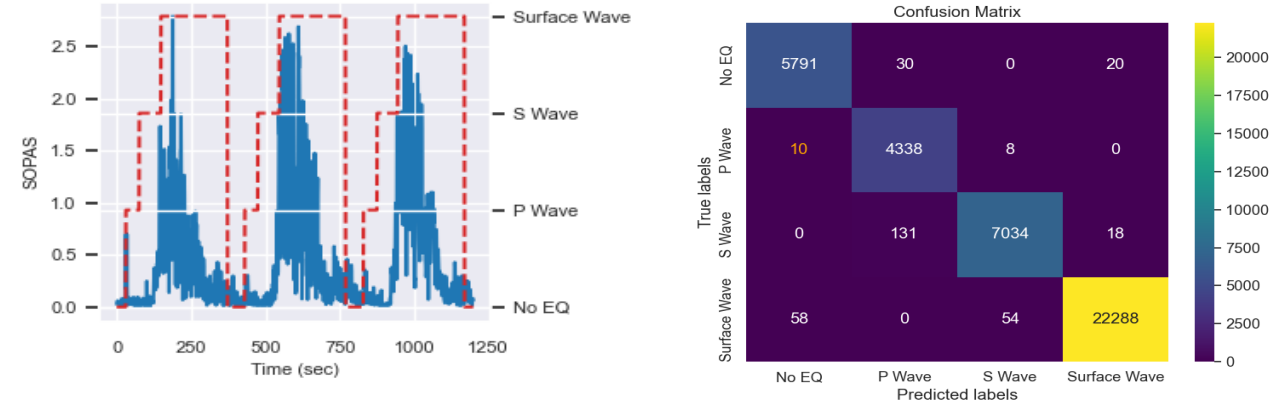

Fig. 7: (a) GRU Predictions on Emulated Dataset, (b) GRU Confusion Matrix for Emulated Dataset

TABLE I: Performance comparison of LSTM and GRU approach on Emulated Dataset

Model	Parameters count	Training time (hours)	Prediction time (seconds)	Accuracy	Precision	Recall	F1 Score
LSTM	198,404	6.7	109	98.93%	98.25%	98.94%	98.58%
GRU	149,828	3.1	88	99.17%	98.55%	99.04%	98.79%

A. Model training

Both models are trained using the adaptive learning rate optimization technique (Adam), which adjusts the learning rate for each parameter based on approximations of the first and second moments of the gradients. Training is done using the sparse categorical cross-entropy loss function, specified as follows:

$$\text{Sparse Categorical Crossentropy}(y_{\text{true}}, y_{\text{pred}}) = -\frac{1}{N} \sum_{i=1}^{N} \log(y_{\text{pred}}^{(i)}[y_{\text{true}}^{(i)}]) \,, \quad (1)$$

where:

- N is the total number of samples in the dataset.
- $y_{\text{true}}^{(i)}$ represents the true label for the i^{th} sample, which is an integer index corresponding to the correct class.
- $y_{\text{pred}}^{(i)}$ is the predicted probability distribution over all classes for the i^{th} sample.
- The term $\log(y_{\text{pred}}^{(i)}[y_{\text{true}}^{(i)}])$ computes the logarithm of the predicted probability assigned to the true class for each sample, contributing to the overall loss.

Initially, we configure the number of epochs to 100 with a batch size of 32 and implement early stopping to further reduce the risk of overfitting. Both models are trained on a simulated dataset comprising 210,000 samples, using a learning rate of 0.0001. For the training process, 60% of this dataset is allocated for model training, while 20% is reserved for validation and the remaining 20% for testing. After evaluating the models on the 20% testing portion of the simulated dataset, we then assess their performance on an additional emulated dataset of 19,950 samples collected from our laboratory setup. The flow chart of our DL model's training, validation and testing is depicted in Fig. 5. To evaluate the performance of

our proposed model, we utilize accuracy, precision, recall, and F1 score metrics for four classes: No Earthquake, P-Wave, S-Wave, and Surface Waves.

B. Results

In this section, we evaluate the performance of LSTM (Fig 6) and GRU (Fig 7) on emulated dataset for earthquake detection. Key performance metrics, including accuracy, precision, recall, F1-score, and computational efficiency, are evaluated. Furthermore, the performance of LSTM and GRU is summarized in Table I. In Fig. 6, the prediction graph, where blue lines represent the SOP signal and the red dashed lines represent the model's predictions for earthquake waves. LSTM provides reasonably good predictions. Furthermore, the confusion matrix reveals the LSTM model's accuracy of 98.93%, with the correct classification of 22,047 Surface Waves and 7,012 S Waves. However, it misclassified some P Waves as No EQ, which slightly affected its recall for this particular class. The LSTM model's accuracy of 98.93%, with the correct classification of 22,047 Surface Waves and 7,012 S Waves. In Fig. 7, the GRU model's prediction graph and confusion matrix demonstrate very good performance. The P-wave class is a critical seismic wave, and its accurate detection and higher recall are essential for timely earthquake detection. The LSTM model achieved a recall of 98.94%, meaning that it correctly identified 98.94% of all actual P-waves in the dataset. There were some instances where the model missed P-waves, misclassifying them as No EQ or other classes, which led to a slight decrease in recall. The GRU model achieved a recall of 99.04%, which is higher than the LSTM model's recall. This indicates that the GRU model performs better at detecting true P-waves, even in cases of complex overlapping seismic events. The GRU confusion matrix demonstrates its superior performance, achieving an accuracy of 99.17%. The GRU correctly classified 22,178 Surface Waves and 7,054 S Waves, with fewer false positives and false negatives than the LSTM. The detailed evaluation of precision, recall, and F1-score for both models is provided in Table I. The computational efficiency of the models is assessed based on the number of parameters, training time, and prediction time (Table I). The GRU model, with 149,828 parameters, required significantly less time for training (3.1 hours) and prediction (88 seconds). Its lower parameter count and faster processing times, combined with its higher performance, make it a more efficient choice for earthquake monitoring. While the LSTM model delivered commendable results, its slightly higher misclassification rate and increased computational demands make it less optimal for real-time earthquake detection. In contrast, the GRU model provides a more efficient and reliable solution for seismic event classification, demonstrating its ability to accurately detect and classify seismic activity with minimal resource usage.

V. Conclusions

This study highlights a significant advancement in earthquake early warning systems through the use of existing terrestrial optical networks, enhanced with deep learning models, for rapid and accurate seismic detection. By simulating strain in fiber cables and training on real displacement data from multiple earthquakes, our approach successfully detects the initial P-waves that precede more destructive surface waves, achieving early detection accuracy of 98.93% with LSTM and 99.17% with GRU. These results are validated on an emulated dataset using a 38 km deployed fiber link in Turin, Italy, demonstrating the effectiveness of using State of Polarization (SOP) measurements in optical fibers, combined with deep learning, for reliable earthquake detection. Our results indicate that the GRU model is more effective at detecting seismic events with higher accuracy, making it a more suitable choice for real-time earthquake detection applications. This research advances the field of intelligent, fiber-optic-based early warning systems, contributing to enhanced public safety and disaster preparedness in earthquake detection.

References

[1] H. Liu, J. Ma, T. Xu, W. Yan, L. Ma, and X. Zhang, "Vehicle detection and classification using distributed fiber optic acoustic sensing," IEEE Transactions on Vehicular Technology **69**, 1363–1374 (2020).

[2] Z. Li, J. Zhang, M. Wang, Y. Zhong, and F. Peng, "Fiber distributed acoustic sensing using convolutional long short-term memory network: a field test on high-speed railway intrusion detection." Optics express **28 3**, 2925–2938 (2020).

[3] P. D. Hernández, J. A. Ramírez, and M. A. Soto, "Deep-learning-based earthquake detection for fiber-optic distributed acoustic sensing," Journal of Lightwave Technology **40**, 2639–2650 (2022).

[4] F. Usmani, A. D'Amico, R. Bratovich, M. R. Fransisco, S. Straullu, E. Virgillito, F. Aquilino, R. Pastorelli, and V. Curri, "Road traffic detection with a lstm autoencoder using state of polarization on deployed metropolitan fiber cable," in *2023 International Conference on Optical Network Design and Modeling (ONDM),* (2023), pp. 1–3.

[5] F. Usmani, H. Awad, E. Virgillito, R. Bratovich, S. Straullu, F. Aquilino, R. Proietti, R. Pastorelli, and V. Curri, "Earthquake early warning through terrestrial optical networks: A bi-gru attention model approach on sop data," in *Optical Fiber Communication Conference (OFC) 2024,* (Optica Publishing Group, 2024), p. Tu3J.2.

[6] H. Awad, F. Usmani, E. Virgillito, R. Bratovich, R. Proietti, S. Straullu, R. Pastorelli, and V. Curri, "Seismic detection through state-of-polarization analysis in optical fiber networks," in *Optical Fibers and Sensors for Medical Diagnostics, Treatment, and Environmental Applications XXIV,* , vol. 12835 I. Gannot and K. Roodenko, eds., International Society for Optics and Photonics (SPIE, 2024), p. 128350L.

[7] H. Awad, F. Usmani, E. Virgillito, R. Bratovich, R. Proietti, S. Straullu, R. Pastorelli, and V. Curri, "A machine learning-driven smart optical network grid for earthquake early warning," in *2024 24th International Conference on Transparent Optical Networks (ICTON),* (2024), pp. 1–6.

[8] H. Awad, F. Usmani, E. Virgillito, R. Bratovich, R. Proietti, S. Straullu, F. Aquilino, R. Pastorelli, and V. Curri, "Environmental surveillance through machine learning-empowered utilization of optical networks," Sensors **24** (2024).

[9] Italian National Institute of Geophysics and Volcanology (INGV), "Italian National Institute of Geophysics and Volcanology (INGV)," (2023). Available online: http://ismd.mi.ingv.it/ismd.php?tipo=lista (accessed on 20 April 2024).

[10] D. Fratta, "Overview and preliminary results from the porotomo project at brady hot springs, nevada: Poroelastic tomography by adjoint inverse modeling of data from seismology, geodesy, and hydrology," (2017).

[11] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation **9**, 1735–1780 (1997).

[12] K. Cho, B. van Merrienboer, Çaglar Gülçehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder–decoder for statistical machine translation," in *Conference on Empirical Methods in Natural Language Processing,* (2014).