

Real-time Anomaly Detection with Distributed Acoustic Sensing and Machine Learning

Neubrex Co., Ltd.	Regular Member	○Thein Lin Aung
Central Japan Railway Company	Regular Member	Akinori Hasegawa
Neubrex Co., Ltd.	Regular Member	Tatsuo Kokubo
Neubrex Co., Ltd.	Regular Member	Yoshiaki Yamauchi

1. Introduction

Distributed Acoustic Sensing (DAS) uses optical fibers to measure small vibrations along the entire length (kilometer-scale) of the fiber. The fine spatial resolution ($\sim 1\text{m}$) of DAS produces a large amount of sensing data in both distance and time¹. Extracting useful information from such high-density acoustic signals is challenging with traditional methods but advances in machine learning (ML) and computing hardware make it possible to process big data in real-time. Combined with long-range coverage of DAS, ML-based approaches enable detection and early warning of hazardous events such as earthquakes, abnormal structural vibrations, rockfalls, landslides, and more.

A major difficulty in implementing ML systems in practice is the acquisition of enough high-quality training data. However, data from abnormal events that rarely occur are difficult to obtain. To address this hurdle, unsupervised anomaly detection models are trained exclusively on normal data which can be easily obtained from routine in-service measurements². This greatly accelerates development and deployment of DAS-based anomaly detection systems.

As an illustrative example of anomaly detection using DAS, a field experiment was conducted to detect rockfalls and human activity near the sensing fibers. An unsupervised autoencoder model was trained to distinguish anomalous signals from ambient environmental noise. This study focused on evaluating the practicality and effectiveness of the anomaly detection approach.

2. Experiment setup and measurement data

Three lines of 24m long optical fiber cables were laid out on the ground as shown in Fig. 1. The cables were connected with communication fibers, and one end of the fiber was connected to DAS interrogator (Neubrex NBX-S4100). The DAS acquisition parameters are sampling rate 500 Sps, spatial resolution 1 m, gauge length 1 m, and channel spacing 20 cm.

The recorded DAS data can be separated into three categories: ambient noise, rockfall, and human activity. The ambient noise (wind, rain, traffic, ground motion, etc.) was continuously measured. To simulate rockfall events, various rocks weighing from 4 kg to 52 kg were dropped from a nearby slope at 6.1 m height. In addition, human activity of walking and jumping near the fiber lines were also measured.

Two examples of DAS data from rockfall and human walk are shown as frequency band energy (FBE) heatmaps in Fig. 2. The total measurement data amounted to 14.8 hours of ambient noise, 22 minutes of rockfall, and 9 minutes of human activity.

3. Autoencoder training for anomaly detection

The goal of the autoencoder model is to learn the characteristics of normal ambient noise and identify

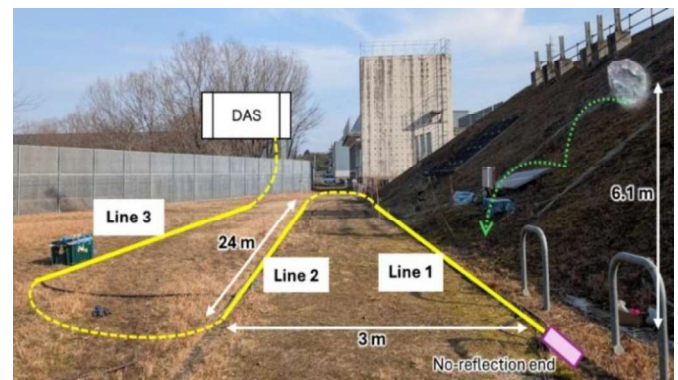


Fig. 1. DAS experiment with three lines of optical fiber cables. Solid yellow lines show sensing fibers and dotted yellow lines show communication fibers.

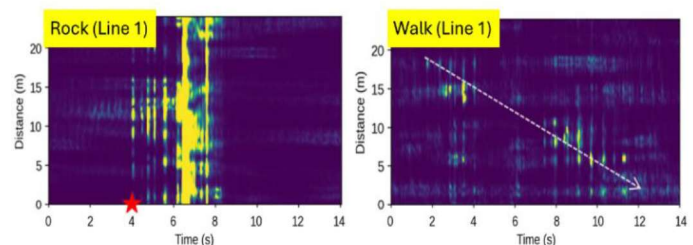


Fig. 2. Examples of DAS data (FBE) from rockfall and human walk. The red star denotes the start time of rockfall, and white arrow denotes the walking direction.

Keywords: DAS, optical fiber sensing, anomaly detection, rockfall, autoencoder

Contact address: Neubrex Co., Ltd. Sakaemachidori 1-1-24, Chuo-ku, Kobe, Hyogo, 650-0023, Japan, Tel: +81-78-335-3510

anomalous events when they occur. This is achieved by reconstructing the input data through an encoder-decoder network and calculating the mean-squared-error (MSE) between input and reconstruction as shown in Fig. 3. The autoencoder was trained only on ambient noise, and hence MSE is expected to be low for normal ambient noise, but high for anomalous (rockfall or human) signals.

The training was unsupervised, meaning that all ambient measurement data was used as training input without any manual labeling. The training dataset was prepared as follows: DAS data (strain rate) from each channel was cut into 3-second time series, then each time series was normalized by subtracting mean and dividing by standard deviation, and then converted to spectrograms with Short-Time Fourier Transform. Out of all ambient noise spectrograms, 20,000 samples were randomly selected as training data and 5,000 samples were selected as validation data to judge the training.

4. Anomaly detection results

For testing the trained model, 2,000 spectrograms were randomly selected from each event type (ambient noise, rockfall, walk, jump). Examples of autoencoder output and absolute error for each event are shown in Fig. 4, where rockfall and walk events result in notably higher MSE values. The MSE distributions for all events are presented in Fig. 5, revealing a clear separation between ambient noise and anomalous signals, as anomalies shift toward larger MSE. The decision threshold for anomalies was defined as the 99th percentile MSE of validation data indicated by the vertical red line in Fig. 5. Using this threshold, each sample was classified as normal or anomaly, and the overall results are summarized in Table 1. The model achieved false positive rate of 0.1% and maximum false negative rate of 1.75%.

5. Conclusions

An unsupervised autoencoder model was trained to detect rockfalls and human activity in DAS signals. The model achieved high sensitivity (>98% detection rate) and low false positive rate (0.1%). This approach is also practical for field deployment because new training data can be readily collected under normal operating conditions. Although the current method cannot differentiate between different types of anomalies, it remains effective as a screening or early warning system for further inspections.

References

(1) Kishida, K.; Guzik, A.; Nishiguchi, K.; Li, C. H.; Azuma, D.; Liu, Q.; He, Z. Development of Real-Time Time Gated Digital (TGD) OFDR Method and Its Performance Verification. *Sens. 2021 Vol 21 Page 4865* **2021**, 21 (14), 4865. <https://doi.org/10.3390/S21144865>.
(2) Xie, Y.; Wang, M.; Zhong, Y.; Deng, L.; Zhang, J. Label-Free Anomaly Detection Using Distributed Optical Fiber Acoustic Sensing. *Sensors* **2023**, 23 (8). <https://doi.org/10.3390/s23084094>.

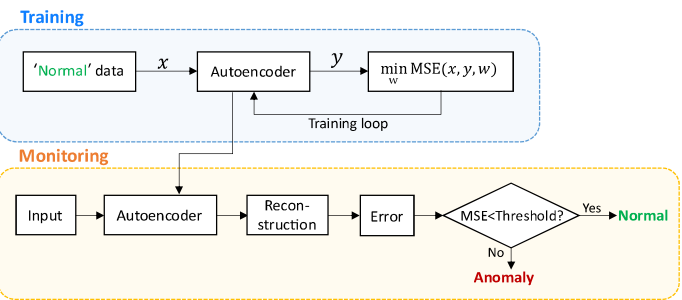


Fig. 3. Schematics of autoencoder for anomaly judgement

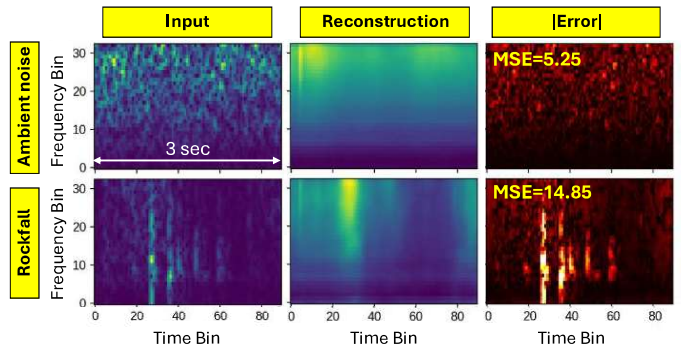


Fig. 4. (Top to bottom) Sample spectrograms from ambient noise and rockfall; (Left to right) original input, autoencoder reconstruction and absolute error.

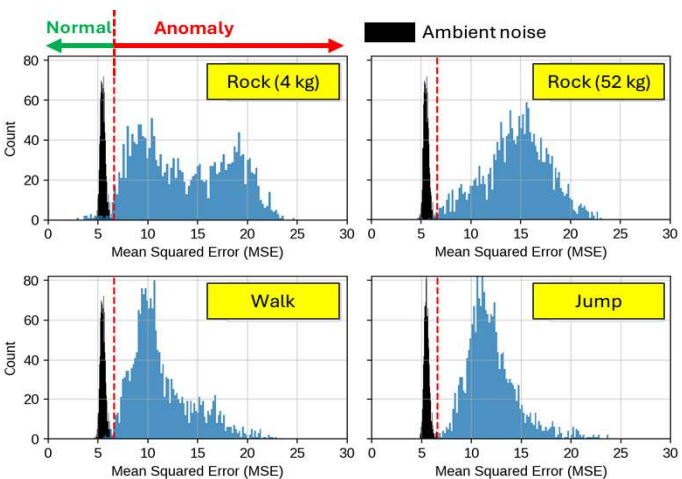


Fig. 5. Histogram distributions of ambient noise, rockfall (small & big) and walk. (n=2,000 samples per event)

Table 1. Summary of anomaly detection in DAS data.

Event	Total	Anomalies	Error
Ambient	2000	2 (0.10%)	2 (0.10%)
Rock (small)	2000	1965 (98.25%)	35 (1.75%)
Rock (large)	2000	1999 (99.95%)	1 (0.05%)
Walk	2000	1984 (99.20%)	16 (0.80%)
Jump	2000	1999 (99.95%)	1 (0.05%)