INTERNATIONAL JOURNAL OF INTEGRATED ENGINEERING
ISSN: 2229-838X e-ISSN: 2600-7916

IJIE

Vol. 17 No. 2 (2025) 54-65
https://publisher.uthm.edu.my/ojs/index.php/ijie

LSTM: Anomaly Activity Type Classification Using Distributed Acoustic Sensing Based on MFCC Features

Nurul Ain Abdul Aziz[1], Hong Yeap Ngo[1], Kan Yeep Choo[1]*, Tee Connie[2], Hafiz Zulhazmi Jabidin[1], Sithi Vinayakam Muniandy[3]

[1] *Faculty of Engineering,*
Multimedia University, Persiaran Multimedia, 63100 Cyberjaya, Selangor, MALAYSIA
[2] *Faculty of Information Science & Technology,*
Multimedia University, Jalan Ayer Keroh Lama, 75450 Bukit Beruang, Melaka, MALAYSIA
[3] *Department of Physics, Faculty of Science,*
University of Malaya, 50603 Kuala Lumpur, MALAYSIA

*Corresponding Author: kychoo@mmu.edu.my
DOI: https://doi.org/10.30880/ijie.2025.17.02.005

Article Info

Received: 10 November 2024
Accepted: 24 June 2025
Available online: 18 July 2025

Keywords

Distributed acoustic sensing, fibre optic sensor, deep learning, long short-term memory, Mel-frequency cepstral coefficients, construction tools.

Abstract

The integrity and connectivity of the fibre optical network are important in preserving the quality of services and internet reliability between providers and end users. However, these networks are vulnerable to disruptions due to unintentional breaks and damage caused by physical disturbances such as construction activity. An accurate classification of anomalous activity at surrounding area plays a crucial role in monitoring the buried fibre optical network from harm which can lead to denial of services. Distributed acoustic sensing (DAS) with combination of deep learning-based technique have potential in targeting this issue, by leveraging the unique pattern of vibration signal measured by the DAS to classify and identify anomalous activities. This work demonstrated utilization of dark fibre buried along the road until the server room, then connected to the DAS interrogator unit (IU). The vibration signals induced by construction hand tools, including hoe, shovel, and sledgehammer, which are used to mimic anomalous activity, are measured by DAS IU and underwent pre-processing before the Mel frequency cepstral coefficient (MFCC) features extraction for long short-term memory (LSTM) model training. The average accuracy score using 13 MFCCs resulted up to 88%, indicating that the proposed method has great potential for anomalous activity monitoring for fibre break prevention.

1. Introduction

In the revolution of telecommunication industry, fibre optics have been used widely as physical medium for data transmission and expand the communication networks over a long range with low signal degradation [1]. Fiber optics offers high bandwidth, fast transmission, low-cost material, longer lifespan and anti-electromagnetic interference compared traditional copper-based network that can oxidize over time and frequently reported stolen issue [2], [3]. Unfortunately, the networks are vulnerable to the unintentional damage caused by environment harm and human activity, such as construction work near to the buried fibre optics [4]. An unsupervised work that involves dig-up and excavation, such as in property development, drainage installation, or railway construction, can potentially damage any fibre optic cable buried nearby [5]. The data transmission will be affected when fibre break happens within the network, which can lead to low quality of services and

communication outages. Conventionally, faults in fibre optic networks are monitored using Optical Time-Domain Reflectometry (OTDR) instrument but can only detect connectivity disruptions after they have been reported by users [6]. As a result, this method is considered time-consuming and labour-intensive, as it requires manual monitoring, detects fibre optic faults only after a break occurs, and involves significant time for troubleshooting and repairs [7]. Thus, fibre optic network monitoring system plays a crucial role in real time detection and localization of anomalous activities that could potentially cause damage near the fibre optic cables, then can minimize delays in maintaining work.

Recently, distributed acoustic sensing (DAS) has attracted significant attention as a promising solution for long-distance disturbance detection and localization, with a capability extending over 40 to 50 km of fibre length [8]. DAS has been utilized in several applications, such as structural health monitoring in road networks and infrastructures [9], [10], pipeline systems [11], [12], perimeter security [13], [14], and railway constructions [15], [16]. The utilization of the DAS system for event detection becomes more reliable by combining deep learning techniques for the intelligent classification of events and real-time identification of anomaly that happens along the optical fibre [17]. However, interpreting DAS data is challenging due to its large size and superimposed noise, necessitating advanced signal processing and feature extraction techniques for accurate pattern recognition.

Current advancements in deep learning, specifically in recurrent neural networks (RNNs), offer excellent performance in the analysis and pattern handling of time-series or sequential data recognition, such as acoustic signals, including audio [18], vibration, and speech [19]. For instances, long short-term memory (LSTM) networks have shown exceptional performance in capturing long-term dependencies and recognizing complex patterns in sequential data, achieving test accuracy exceeding 85% with the LSTM network and up to 90% with the bidirectional long short-term memory (BLSTM) model [19]. Long short-term memory (LSTM) networks can be utilized for pattern recognition, and when combined with appropriate feature extraction, they enable accurate detection and classification of anomalous events for effective optical fibre network health monitoring. Besides, Mel-frequency cepstral coefficients (MFCCs) have been widely used in sound analysis owning to their ability to extract key spectral features while less affected by amplitude variations and noise of the signal [20], [21].

This work evaluates the effectiveness of several proposed MFCC-LSTM models for feature extraction and classification of anomalous activities detected via distributed acoustic sensing, aiming for fibre optic network health monitoring and break prevention.

2. Methodology

The core principle of the DAS system is based on the OTDR technique, where a repetitive light pulse is generated by a laser source, amplified by an erbium-doped fibre amplifier (EDFA), and then injected into the fibre under test (FUT) via an optical circulator, as illustrated in Fig. 1.

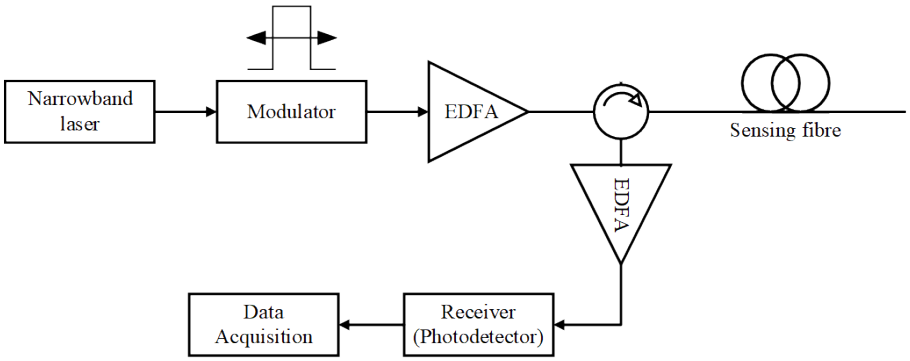


Fig. 1 *Diagram of a DAS interrogator unit*

Backscattered light is generated by deformations in the FUT, separated from the incident light by the optical circulator, amplified by another EDFA, and ultimately detected by a photodetector. The intensity of backscattered light is then transformed into electrical signal, which represents the vibration information. Environmental vibrations surrounding the fibre, such as underground movement, vehicle, or human activities, can introduce noise into the backscattered light.

Fig. 2 presents the flowchart of the optical fibre network health monitoring and break prevention system proposed in this work. The process began with data collection using the DAS interrogator unit, as shown in Fig. 1, and ended with the training and evaluation of the LSTM models for classifying anomalous activities that could affect the health of the optical fibre network.

Int. Journal of Integrated Engineering Vol. 17 No. 2 (2025) p. 54-65

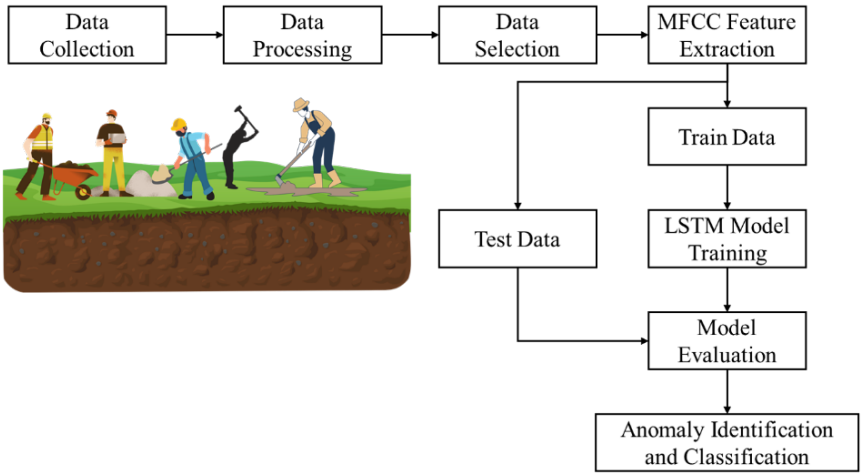


Fig. 2 *Flowchart of the optical fibre network health monitoring and break prevention system*

2.1 Data Collection

Fig. 3 depicts the path of a 1.67 km-long fibre-under-test (FUT) buried beneath the pavement and experiment sites. The fibre is a single-mode optical communication fibre connecting a commercial DAS system in the data centre room at the Faculty of Engineering, MMU Cyberjaya, as represented by the black icon, and extending toward the student hostel.

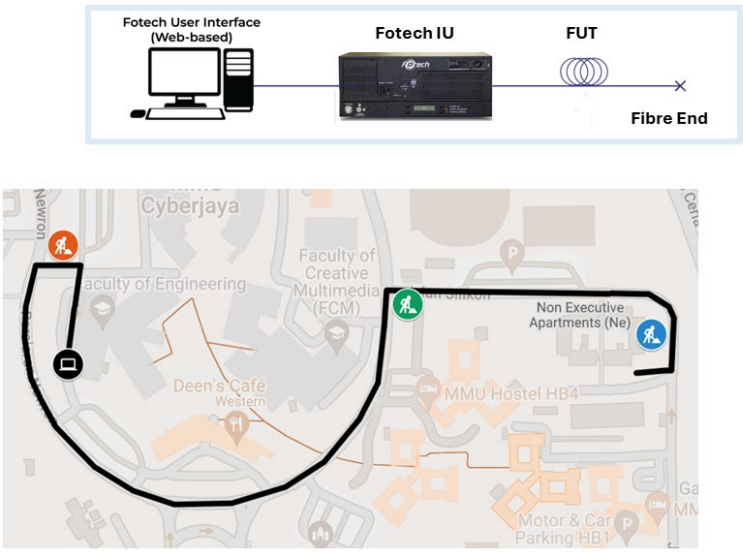


Fig. 3 *Path of the FUT and experiment sites*

The DAS system was set at a sampling rate of 55.6 kHz, which allows a wide range of detection of vibration frequencies depending on the activity. Meanwhile, the analog-to-digital converter (ADC) rate was limited to 150 mega samples per second, resulting in a spatial sampling interval of 0.68 meters. This enabled high-resolution detection of vibration disturbances along the length of the fibre. Controlled experiments were carried out to gather data for proof of concept and assess the effectiveness of the proposed model in distinguishing between four activities: hammering, hoeing, shovelling, and no event (environment noise due to human or vehicle activities). Three distinct sites were selected for the experiments, represented by the orange, green, and blue icons, corresponding to locations approximately 500 m, 1000 m, and 1500 m away, respectively. These sites were chosen for their relatively quiet environment with minimal human and vehicle activity, reducing potential external noise in the collected data.

2.2 Data Pre-Processing

Figure 4 depicts the raw signal, containing a DC offset and noise, obtained from the DAS interrogator unit at an approximate location of 1505 m. Therefore, it is essential to eliminate the DC offset and unwanted noise to ensure optimal signal quality prior to model training.

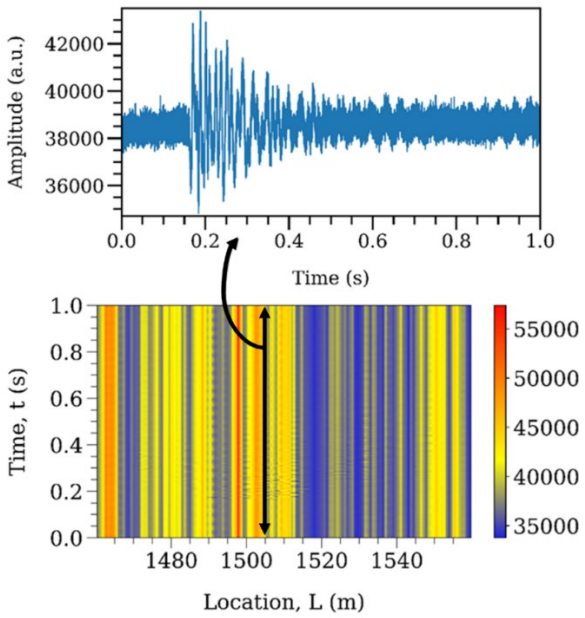


Fig. 4 *Raw signal from the unfiltered data collected using the DAS system*

Fig. 5 illustrates a typical example of a raw and filtered vibration signal with zero DC offset in the time and frequency domains. The noise is eliminated using a 21-order bandpass Butterworth filter with a passband frequency range of 20 Hz to 300 Hz, as the dominant frequencies fall within this range. Fig. 6 shows the filtered waterfall plot of the vibration patterns produced by the sledgehammer, with noise effectively removed, making the signature patterns clearly visible.

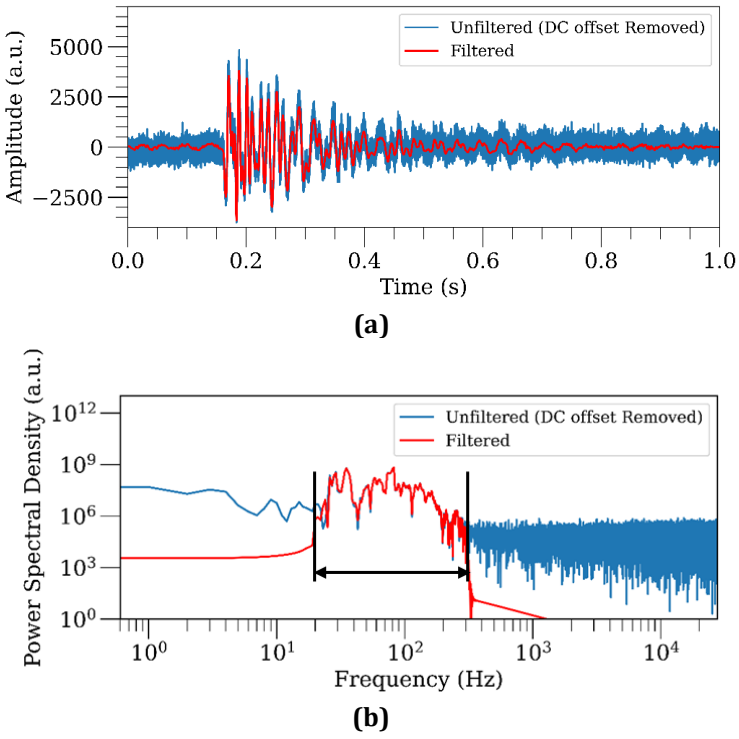


Fig. 5 *Typical example of a raw vibration signal and filtered vibration signal using a bandpass Butterworth filter plotted in the (a) Time-domain; and (b) Frequency-domain*

Int. Journal of Integrated Engineering Vol. 17 No. 2 (2025) p. 54-65

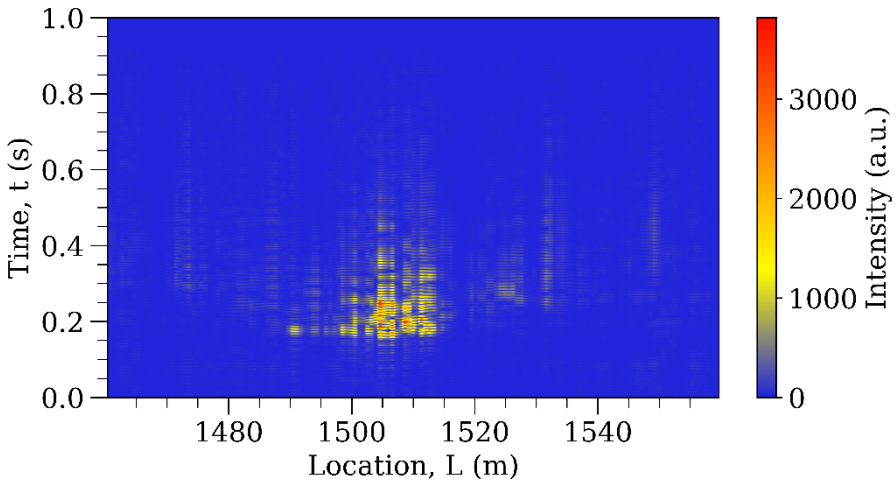


Fig. 6 *Filtered waterfall plot of the vibration patterns produced by the sledgehammer*

2.3 MFCC Feature Extraction

The key spectral features of the vibration signal are captured by the MFCC, which are obtained by applying the discrete Fourier transform to the Mel spectrogram. The flowchart of the feature extraction process is shown in Fig. 7.

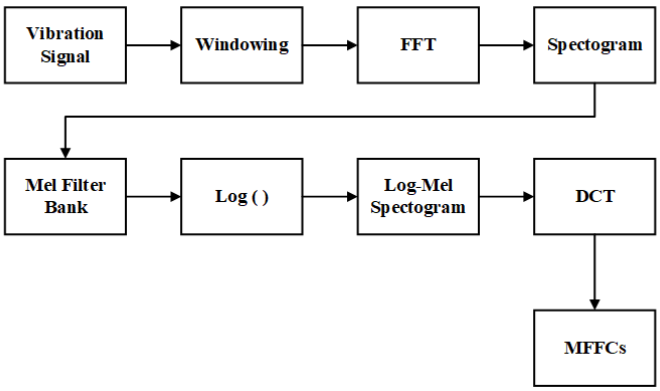


Fig. 7 *Flowchart of feature extraction process*

The vibration signal was downsampled to 4 kHz, as the key frequencies ranged from 20 Hz to 300 Hz. It was then segmented into frames of 256 to 2048 samples with a 50% overlap to examine the effect of frame size on LSTM model performance. In each frame, the vibration signal was processed using a Hamming window to reduce power leakage before being transformed from the time domain to its spectrogram. The spectrogram was processed using Mel-scale filters, a set of triangular bandpass filters, to generate the Mel spectrogram, which was then converted into the log-Mel spectrogram using a logarithm function. The Discrete Cosine Transform (DCT) is applied to the log-Mel spectrogram to decorrelate spectral features and extract MFCCs, which more effectively capture the spectral envelope. Thirteen MFCC coefficients were extracted from each vibration signal for each frame size, as illustrated in Fig. 8, and were used as feature data to train various LSTM models, as detailed in Table 1.

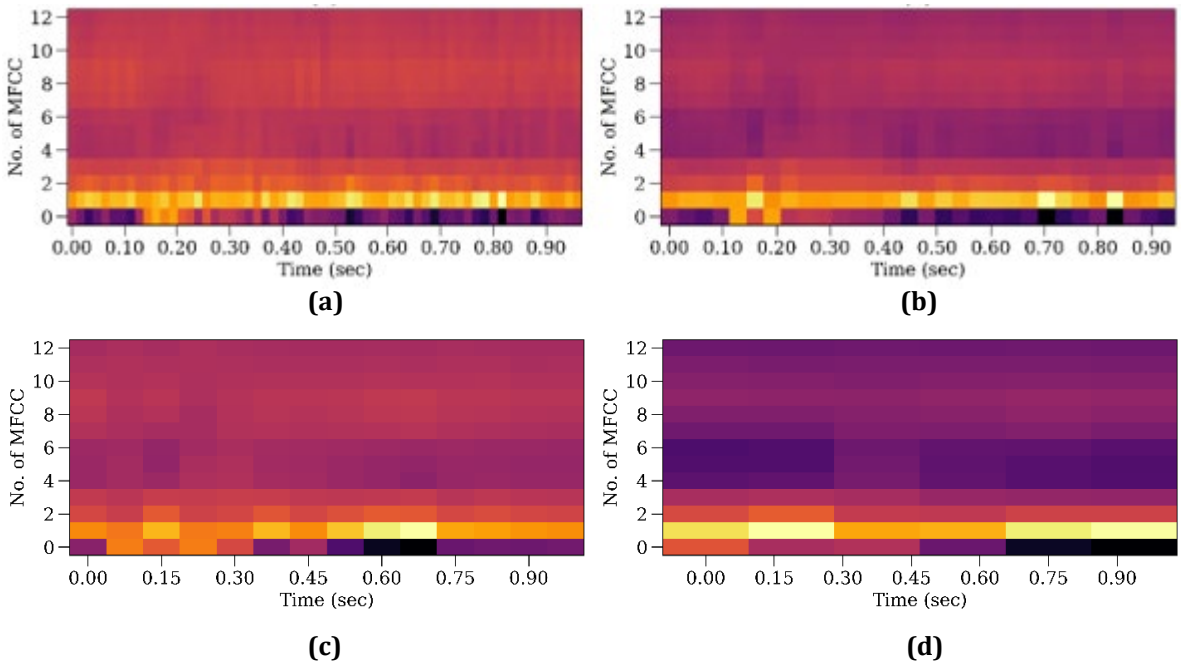


Fig. 8 *MFCCs of a vibration signal obtained with different frame sizes (a) 128; (b) 256; (c) 512; and (d) 1024*

Table 1 *Input data dimension of a vibration signal based on different MFCC frame sizes*

Frame Size	Input data dimension (*sequence* × *feature*)
128	62 × 13
256	32 × 13
512	14 × 13
1024	6 × 13

2.4 Long Short-Term Memory (LSTM)

LSTM network is considered as an extension of the recurrent neutral network (RNN) that is designed to overcome the short-range dependency limitation of RNN. Based on [22], LSTM network is superior for prediction and classification of time series with long range dependency. The architecture of LSTM networks incorporates memory cells and feedback loops, where data from previous time steps, whether short- or long-range, is used as input for the current time step. This allows the network to capture the memory effects of the signal and determine how they impact the current state of the signal [23]. For this work, several LSTM networks were designed based on the Vanilla LSTM and stacked LSTM, as described in Table 2.

Table 2 *Descriptions of various LSTM models*

Model	Layers Description
LSTM 1	A block of LSTM with 128 neurons and a dropout of 0.5
	1 Dense layer of 128 unit and a ReLU activation
	1 Dense layer with 4 classes and a SoftMax activation
LSTM 2	Same as LSTM 1, but after the LSTM layer has a Batch Normalization
LSTM 3	3 blocks of LSTM with 128 neurons and a dropout of 0.5
	1 Dense layer of 128 neurons and a ReLU activation
	1 Dense layer with 4 classes and a SoftMax activation
LSTM 4	Same as LSTM 3, but after the LSTM layer has a Batch Normalization

Int. Journal of Integrated Engineering Vol. 17 No. 2 (2025) p. 54-65

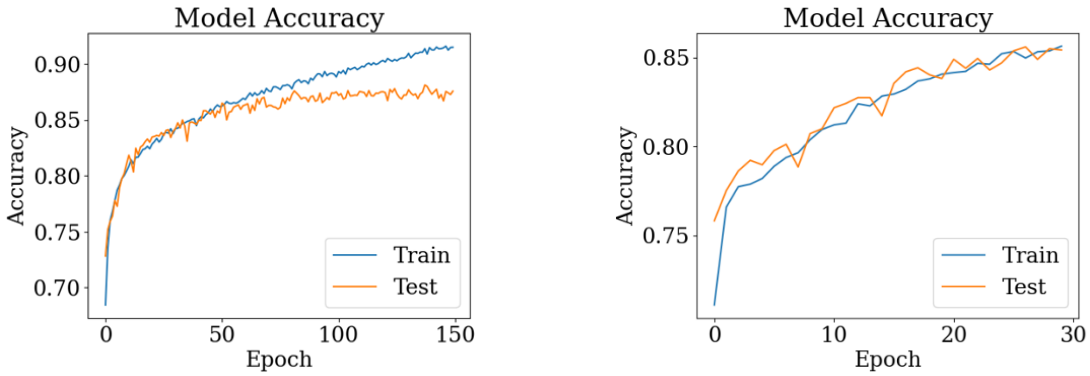
Each LSTM model received an input tensor with dimension of *batch size × sequence × feature*, where *batch size* is defined as the total number vibration signals divided by the number of batches, *sequence* is the number of frames per vibration signal, and *feature* is the number of features extracted from a vibration signal. The LSTM network was structured with an input layer, followed by three recurrent regularization blocks and a fully connected regularized output. Each LSTM layer, containing 128 neurons, was used to capture and learn the temporal variation of the spectral features. The batch normalization was used to standardise the input to each new layer. In addition, A dropout layer with a 50% rate was included to prevent overfitting by randomly deactivating neurons during training. After that, the weights were passed through two Dense layers, the first one containing 128 neurons with ReLU activation and the next one with 3 neurons using a SoftMax activation to generate the final classification output. The model was written using Python 3.7 and computed on multiple 12th-Generation Intel i9-12900 processors with an NVIDIA RTX A4000 graphic card and 32 GB of RAM. The dataset was divided into 60% for training, 20% for validation, and 20% for inference testing. The data distribution is presented in Table 3.

Table 3 *Distribution data for training, validation, and inference test*

Class	Dataset		
	Train	Validation	Test
Hoe	2986	990	1024
Shovel	2967	1029	1004
Sledgehammer	3031	992	977
No Event	3016	989	995
Total	12000	4000	4000

3. Results and Discussion

Initially, the LSTM models were trained for 150 epochs with a batch size of 32. However, during the first trail on the LSTM 1 model, it was observed that the model began to overfit after 50 epochs, as shown in Fig. 9(a). To address this, an early stopping callback with a patience value of 5 was implemented to halt the training before further overfitting occurred, resulting in the performance shown in Fig. 9(b).



(a) **(b)**

Fig. 9 *Train and test accuracies with (a) Callback; and (b) Without early stopping callback*

Table 4 summarizes the performance of four different LSTM models, with the LSTM 4 model achieving the highest accuracy of 88.05%, followed by the LSTM 3 model with a slight lower accuracy of 87.50%. The LSTM 4 model, illustrated in Fig. 10, includes a multi-LSTM block with each layer containing 128 neurons, followed by a dropout layer and batch normalization, in contrast to the LSTM 3 model, which does not include batch normalization. This configuration demonstrates that adding a batch normalization after the LSTM layers improve slightly accuracy as normalizing the data to a common scale which enhances the stability and convergence of the network. Besides, LSTM models with only a single-LSTM block, such as LSTM 2 and LSTM 1 models, resulted in lower accuracies of 85.05% and 84.57%, respectively. This confirms that the LSTM model designed with multiple blocks and batch normalization delivers optimal performance.

The performance of the LSTM models was further evaluated using different frame sizes, which influence the time and frequency resolutions of the time-frequency distribution of the vibration signal, specifically the spectrogram, Mel-spectrogram, and Mel-frequency cepstrum. LSTM 1 and 2 models with a frame size of 512

produces the highest accuracy compared to other frame sizes. Smaller frame sizes, like 128 and 256, reduce frequency resolution, leading to less distinguishable frequency information. On the other hand, a larger frame size of 1024 results in lower time resolution, which fails to capture fine temporal details of the Mel-frequency cepstrum. Both factors make learning less efficient in single-block LSTM models. As a result, a frame size of 512 appears to provide the optimal balance between capturing adequate temporal and frequency information of the Mel-frequency cepstrum in single-block LSTM models, like LSTM 1 and LSTM 2. Additionally, the use of batch normalization in the LSTM 2 model results in slightly better performance compared to the LSTM 1 model.

For multiblock models like the LSTM 3 model, a smaller frame size of 256 yields the highest accuracy, suggesting that more LSTM layers with higher time resolution provide sufficient temporal information of the Mel-frequency cepstrum for the model to learn more effectively. Among these models, the LSTM 4 model, which features three stacked LSTM blocks and batch normalization in each layer, achieves the highest accuracy. This shows that LSTM 4 model effectively balances the trade-off between time and frequency resolutions across different frame sizes, enabling it to capture both temporal (small frame size) and frequency (large frame size) patterns of the vibration signal, as represented by the MFCCs.

Table 4 *Validation accuracies of different LSTM models*

Model	Frame Size			
	128	256	512	1024
LSTM 1	82.78	83.97	84.57	84.13
LSTM 2	82.59	84.40	85.05	84.25
LSTM 3	85.47	87.50	87.08	86.00
LSTM 4	86.92	87.55	87.88	88.05

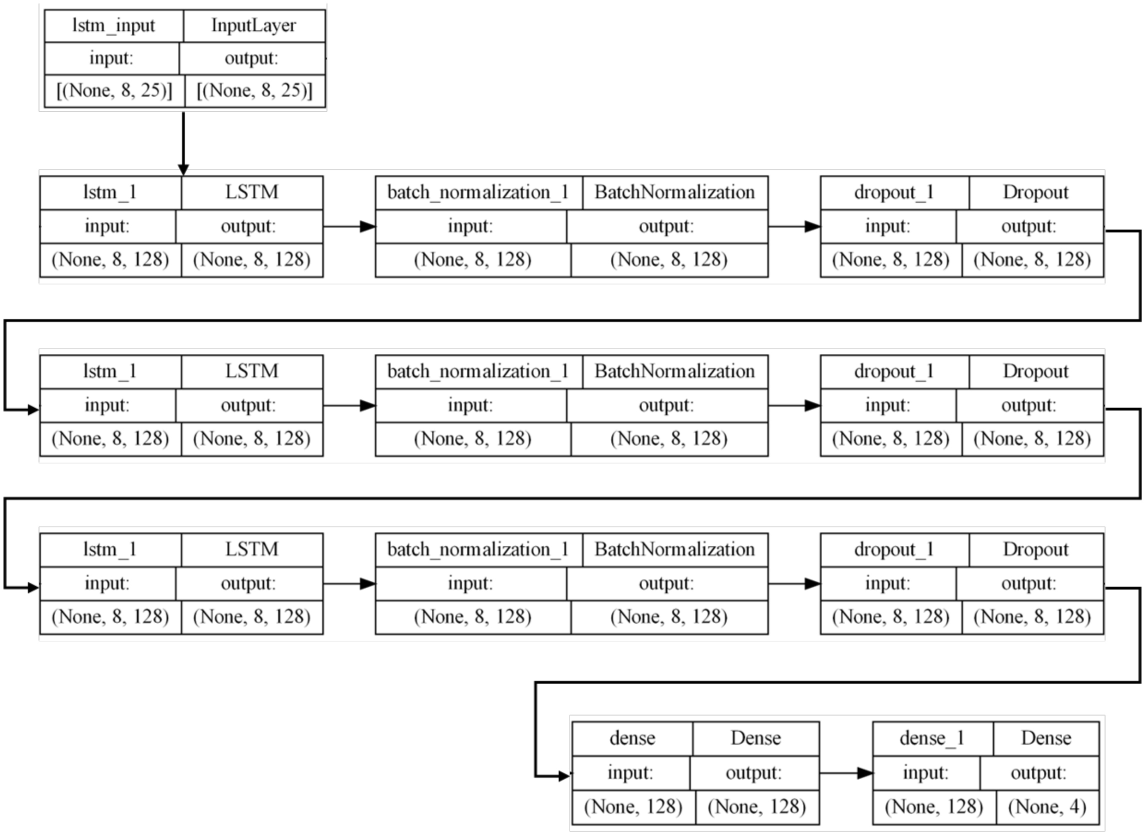


Fig. 10 *The network architecture of the best LSTM model, LSTM 4*

Int. Journal of Integrated Engineering Vol. 17 No. 2 (2025) p. 54-65

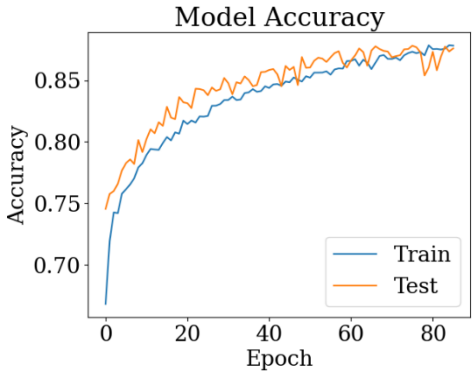


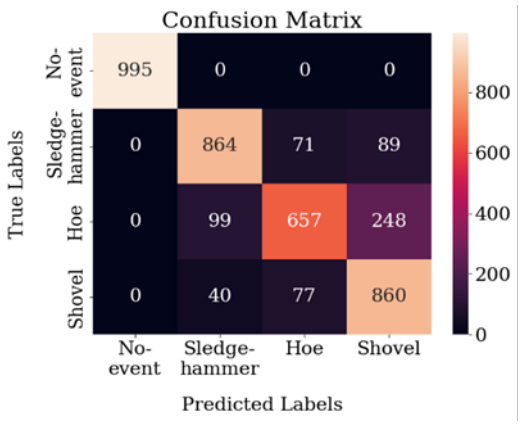
Fig. 11 *Train and test accuracies of LSTM 4 model with frame size of 1024*

Table 5 presents the classification report of the inference test using the best LSTM mode, LSTM 4, with various frame sizes. The LSTM 4 model accurately classifies all construction hand tools, achieving a precision score of approximately 80% for the hoe, 90% for the sledgehammer, and 100% for no event across different frame sizes. Additionally, the precision score for the shovel increases with larger frame sizes, rising from 72% to 78%. A similar trend is seen in the Recall and F1-score, with variations between 2% and 4%, except for the Recall score for the hoe, which increases from 74% to 80% as the frame size rises.

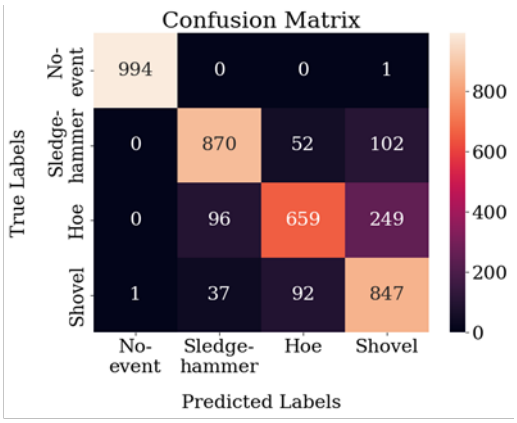
Table 5 *Classification report of the inference test using the LSTM 4 model with various frame sizes*

Frame Size	Precision				Recall				F1-score				Average Accuracy
	No Event	Hoe	Sledge-hammer	Shovel	No Event	Hoe	Sledge-hammer	Shovel	No Event	Hoe	Sledge-hammer	Shovel	
128	1.00	0.82	0.92	0.72	1.00	0.74	0.87	0.87	1.00	0.78	0.89	0.81	0.87
256	1.00	0.81	0.89	0.79	1.00	0.79	0.86	0.84	1.00	0.80	0.87	0.81	0.87
512	1.00	0.83	0.91	0.77	1.00	0.77	0.86	0.88	1.00	0.80	0.89	0.82	0.88
1024	1.00	0.82	0.92	0.78	1.00	0.80	0.86	0.86	1.00	0.81	0.89	0.82	0.88

Fig. 12 presents the confusion matrices for the four different frame sizes obtained using the LSTM 4 model. The LSTM 4 model seems to have some difficulty distinguishing between the shovel and hoe, as these tools have similar shape. The shovel has a triangular thin plate, while the hoe features a flat thin plate, which may produce nearly identical vibration signal patterns and key frequency components, as illustrated in Fig. 13.



(a)



(b)

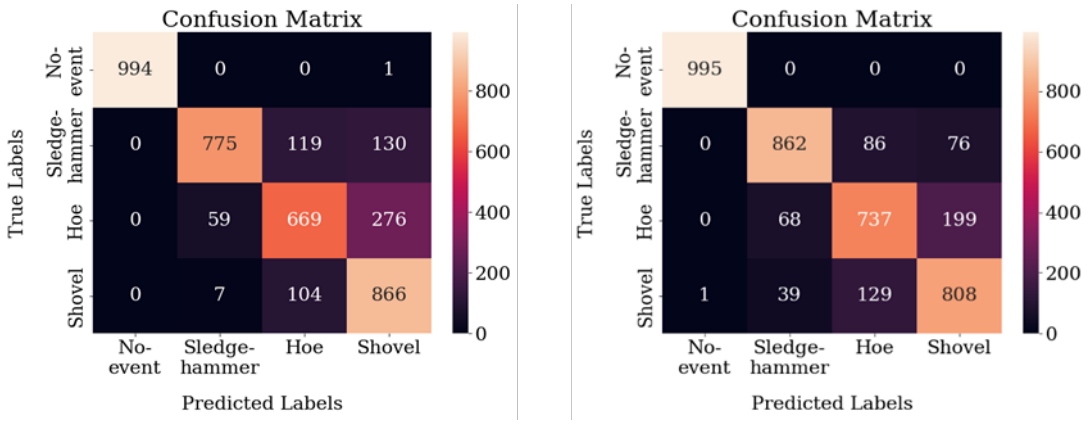


Fig. 12 *Confusion matrix of the inference test using the LSTM 4 model with various frame sizes of (a) 128; (b) 256; (c) 512; and (d) 1024*

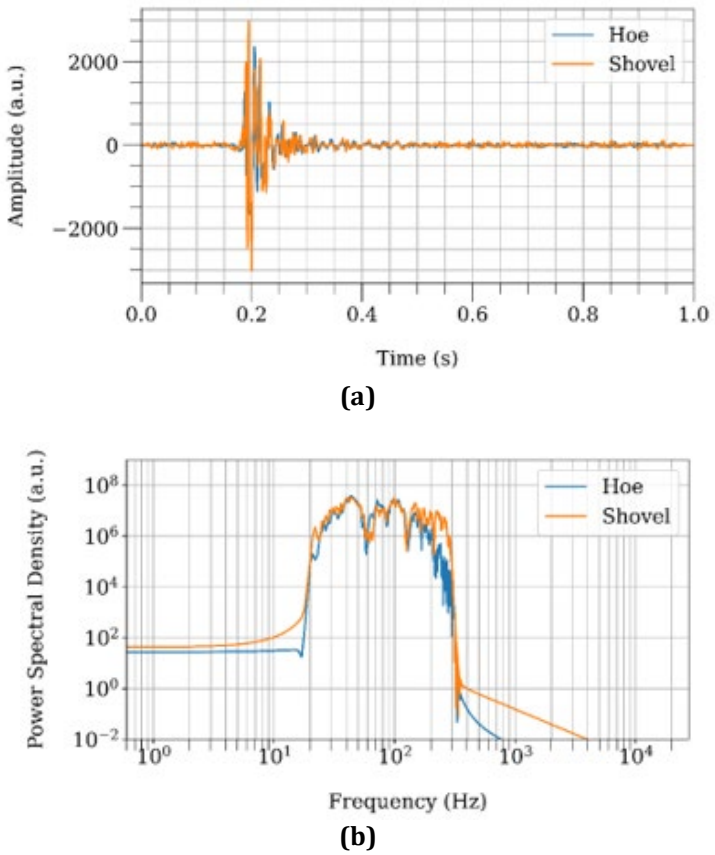


Fig. 13 (a) *Vibration signals; and (b) Power spectral densities of hoe and shovel*

4. Conclusion

Multi-LSTM block model with additional regularization, drop-out, and batch normalization in each layer (LSTM 4) demonstrated a better performance compared to single-LSTM block model. The multi-LSTM block model also effectively balances the trade-off between time and frequency resolutions of the Mel-frequency cepstrum, allowing it to capture both temporal (small frame size) and frequency (large frame size) patterns of a vibration signal. Furthermore, batch normalization improves the stability and convergence of the LSTM 4 model, enabling it to achieve optimal performance. Therefore, the proposed MFCC-LSTM model is highly effective in detecting and classifying anomalous activities generated by construction hand tools and can be easily applied to optical fibre network health monitoring and break prevention system.

Acknowledgement

The authors thank TM R&D Sdn. Bhd. for funding the research work with Project ID No. MMUE/220019.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***study conception and design:*** *Nurul Ain Abdul Aziz, Kan Yeep Choo;* ***data collection:*** *Ngo Hong Yeap, Hafiz Zulhazmi Jabidin, Nurul Ain Abdul Aziz;* ***analysis and interpretation of results:*** *Nurul Ain Abdul Aziz, Kan Yeep Choo, Connie Tee, Sithi V. Muniandy;* ***draft manuscript preparation:*** *Nurul Ain Abdul Aziz, Kan Yeep Choo. All authors reviewed the results and approved the final version of the manuscript.*

References

[1] X. Liu, "Evolution of Fiber-Optic Transmission and Networking toward the 5G Era," 2019, doi: 10.1016/j.isci.

[2] S. Kaur, P. Singh, V. Tripathi, and R. Kaur, "Recent trends in wireless and optical fibre communication," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 343–348, Jun. 2022, doi: 10.1016/j.gltp.2022.03.022.

[3] A. A. Zhirnov *et al.*, "Fiber-Optic Telecommunication Network Wells Monitoring by Phase-Sensitive Optical Time-Domain Reflectometer with Disturbance Recognition," *Sensors*, vol. 23, no. 10, May 2023, doi: 10.3390/s23104978.

[4] K. Abdelli, H. Grieser, C. Tropschug, and S. Pachnicke, "Optical Fiber Fault Detection and Localization in a Noisy OTDR Trace Based on Denoising Convolutional Autoencoder and Bidirectional Long Short-Term Memory," *Journal of Lightwave Technology*, vol. 40, no. 8, pp. 2254–2264, Apr. 2022, doi: 10.1109/JLT.2021.3138268.

[5] T. Hayford-Acquah and B. Asante, "Causes of Fiber Cut and the Recommendation to Solve the Problem," *IOSR Journal of Electronics and Communication Engineering*, vol. 12, no. 01, pp. 46–64, Jan. 2017, doi: 10.9790/2834-1201014664.

[6] K. Abdelli, J. Y. Cho, F. Azendorf, H. Griesser, C. Tropschug, and S. Pachnicke, "Machine-learning-based anomaly detection in optical fibre monitoring," *Journal of Optical Communications and Networking*, vol. 14, no. 5, pp. 365–375, May 2022, doi: 10.1364/JOCN.451289.

[7] J. Y. Cho *et al.*, "DeepALM: Holistic Optical Network Monitoring based on Machine Learning," 2022.

[8] Y.-J. Rao, Z.-L. Ran, Y. Gong, A. Güemes, and J. Sierra Perez, "An Introduction to Distributed Optical Fiber Sensors," Boca Raton, FL, USA, 2017.

[9] M. F. Bado and J. R. Casas, "A review of recent distributed optical fibre sensors applications for civil engineering structural health monitoring," Mar. 01, 2021, *MDPI AG*. doi: 10.3390/s21051818.

[10] J. Braunfelds *et al.*, "Road Pavement Structural Health Monitoring by Embedded Fiber-Bragg-Grating-Based Optical Sensors," *Sensors*, vol. 22, no. 12, Jun. 2022, doi: 10.3390/s22124581.

[11] J. Zuo *et al.*, "Pipeline Leak Detection Technology Based on Distributed Optical Fiber Acoustic Sensing System," *IEEE Access*, vol. 8, pp. 30789–30796, 2020, doi: 10.1109/ACCESS.2020.2973229.

[12] P. Zhang *et al.*, "Machine learning data analytics based on distributed fibre sensors for pipeline feature detection," SPIE-Intl Soc Optical Eng, Jun. 2023, p. 11. doi: 10.1117/12.2663225.

[13] C. Lyu *et al.*, "Robust Intrusion Events Recognition Methodology for Distributed Optical Fiber Sensing Perimeter Security System," *IEEE Trans Instrum Meas*, vol. 70, 2021, doi: 10.1109/TIM.2020.3048521.

[14] P. Guo, T. Shi, Z. Ma, and J. Wang, "Human intrusion detection for high-speed railway perimeter under all-weather condition," *Railway Sciences*, vol. 3, no. 1, pp. 97–110, Feb. 2024, doi: 10.1108/rs-11-2023-0043.

[15] A. Venketeswaran *et al.*, "Recent Advances in Machine Learning for Fiber Optic Sensor Applications," *Advanced Intelligent Systems*, vol. 4, no. 1, Jan. 2022, doi: 10.1002/aisy.202100067.

[16] T. Wu, G. Liu, S. Fu, and F. Xing, "Recent progress of fibre-optic sensors for the structural health monitoring of civil infrastructure," Aug. 02, 2020, *MDPI AG*. doi: 10.3390/s20164517.

[17] Y. Xie, M. Wang, Y. Zhong, L. Deng, and J. Zhang, "Label-Free Anomaly Detection Using Distributed Optical Fiber Acoustic Sensing," *Sensors*, vol. 23, no. 8, Apr. 2023, doi: 10.3390/s23084094.

[18] S. Dwivedi, N. Srivastava, V. Rawal, P. Deshwal, and D. Dev, "Analysing the Impact of LSTM and MFCC on Speech Emotion Recognition Accuracy," in *2023 International Conference on Sustainable Emerging Innovations in Engineering and Technology, ICSEIET 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 55–58. doi: 10.1109/ICSEIET58677.2023.10303464.

[19] A. A. Alemu, M. D. Melese, and A. O. Salau, "Towards audio-based identification of Ethio-Semitic languages using recurrent neural network," *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-46646-3.

[20] A. Hamza *et al.*, "Deepfake Audio Detection via MFCC features using Machine Learning," *IEEE Access*, 2022, doi: 10.1109/ACCESS.2022.3231480.

[21] S. Bouagina, M. Naouara, S. Hafsi, and S. Djaziri-Larbi, "MFCC-Based Analysis of Vibratory Anomalies in Parkinson's Disease Detection using Sustained Vowels," in *2023 1st IEEE Afro-Mediterranean Conference on Artificial Intelligence, AMCAI 2023 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/AMCAI59331.2023.10431494.

[22] J. Menegazzo and A. von Wangenheim, "Road surface type classification based on inertial sensors and machine learning: A comparison between classical and deep machine learning approaches for multi-contextual real-world scenarios," *Computing*, vol. 103, no. 10, pp. 2143–2170, Oct. 2021, doi: 10.1007/s00607-021-00914-0.

[23] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: Lstm cells and network architectures," Jul. 01, 2019, *MIT Press Journals*. doi: 10.1162/neco_a_01199.

[24] H. Li, Z. Wang, X. Yue, W. Wang, H. Tomiyama, and L. Meng, "An architecture-level analysis on deep learning models for low-impact computations," *Artif Intell Rev*, vol. 56, no. 3, pp. 1971–2010, Mar. 2023, doi: 10.1007/s10462-022-10221-5.