

Glass breaks detection system using deep auto-encoders with fuzzy rules induction algorithm



Wai Yan Nyein Naing*, Zaw Zaw Htike, Amir Akramin Shafie

Mechatronic Engineering Department, International Islamic University Malaysia (IIUM), Gombak, Malaysia

ARTICLE INFO

Article history:

Received 25 July 2018

Received in revised form

27 November 2018

Accepted 5 December 2018

Keywords:

Glass break detection

Deep auto-encoder neural network

Fuzzy rule induction algorithm

ABSTRACT

Main uses of glass windows in commercial and residential buildings are prevalent. While a glass-based material has its advantages, it also poses security risks. Therefore, glass break detectors play an important role in security protection for offices and residential buildings. Conventional vibration-based and acoustic-based glass break detectors are designed to detect predetermined temporal and frequency feature thresholds of glass breakage sound signals. This leads to the inability to differentiate glass break from environmental sounds (such as the sound of striking objects, heavy sounds and shouted sounds) that are similar in their amplitude threshold and frequency pattern. Machine learning based acoustic audio classification has been popular in security surveillance applications. Researchers are interested in this research area, and different approaches have been proposed for anomaly event detection (such as gunshots, glass breakage sounds, etc.). This paper proposes a new design of a glass break detection algorithm based on Fuzzy Deep Auto-encoder Neural Network. The algorithm reduces false alarms and improves detection accuracy. Experimental results indicate that proposed fuzzy deep auto-encoder network system attained 95.5% correct detection for the proposed audio dataset.

© 2019 The Authors. Published by IASE. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Nowadays, the primary uses of glass windows in commercial and residential buildings are prevalent. Light, comfort, style and energy efficiency are among the benefits of today's high-performing glass-based materials (Stelzer, 2010). While a glass-based material has its advantages, it also poses security risks. At night or when the owner is not present, an intruder can easily break the glass of a door, then reach inside and open the latch lock (Sidhu, 2005). Therefore, glass break detectors play an important role in security protection for commercial and residential buildings. A Glass Break detector is a sensor used in electronic burglar alarms that are designed to detect if a pane of glass is shattered or broken (Venkat, 2010).

Several electronic sensor-based techniques and machine learning based models are being considered

to improve detection accuracy. Smith (2004) studied vibration signals and acoustic waves resulting from contact force applied on the glass surface when shocking or a glass breaking event occurred. Eskildsen (2005) proposed a temporal feature based glass break detection method, in which a detection alert triggers if the zero crossing rate of input audio that exceeds the predetermined threshold of glass breaking the zero crossing rate. Richard (2011) proposed a dual channel based glass break detection algorithm that contains two different audio transducers, called omnidirectional and highly directional transducers. Their proposed control circuit analyses amplitude and frequency thresholds from both the first and second transducers, then determines whether or not a glass breakage profile is present. Conventional sensor-based glass break detector is triggers a detection alarm if the input audio has the same or higher than the pre-defined vibration and frequency threshold of the glass breaking sounds. A door slamming, hitting, and falling objects can cause similar vibrations and frequency patterns of glass breaking sounds. Therefore, electronic sensor based glass break detectors can also repeatedly trigger an alarm by mistakenly identifying a high-pitched tone from the surrounding environmental (such as dogs barking,

* Corresponding Author.

Email Address: waiyan.nn18@email.com (W. Y. N. Naing)

<https://doi.org/10.21833/ijaas.2019.02.006>

Corresponding author's ORCID profile:

<https://orcid.org/0000-0001-5639-9899>

2313-626X/© 2019 The Authors. Published by IASE.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

children screaming and motorcycles driving noise) as glass breaking sounds (Lopatka et al., 2016). False alarms can trigger costly fines and pull emergency professionals away from real crises and be an inconvenience for the environment and homeowner.

Machine learning algorithms advanced in the past few years and become competitors to traditional statistical models and electronic sensors for various security surveillance fields. Wang et al. (2017) proposed a new audio event detection and classification approach based on a state-of-the-art fully convolutional network framework. They extracted two spectrogram feature representations of glass break audio as input. Their proposed approach contains two stages: the first detects whether there are audio events by sliding convolutional kernel in the time axis, and then proposals which possibly contain audio events generated by Region Proposal Networks (RPN). In the second stage, time and frequency domain information is integrated to classify these proposals and refine their boundaries. They attained 94.6 % accuracy for detecting glass breaking events and outperform traditional based line approach. Alain et al. (2005) proposed a complete detection and recognition system for impulsive sounds (such as glass breaks, human screams, gunshots, explosions or door slams) using a Hidden Markov Model (HMM) classifiers under variant background noise conditions. They attained 98% correct detection accuracy at 70dB signal-to-noise ratio (SNR) and above 80% for 0 dB SNR. Their result found that machine learning based glass break detection models have better accuracy and fewer false alarms than conventional electronic sensors. The long computation process and complex training architecture is a problem for traditional machine learning models to attain higher accuracy at various noise levels. In this paper, we propose a new approach for a glass break detection algorithm based on deep learning architecture to improve detection performance and to reduce wasteful false alarms.

The remaining outline of the paper organizes as follows: Section 3.1 describes the Fast Fourier Transform (FFT) algorithm for the pre-processing process. Section 3.2 discusses the dimension reduction process by using the deep Auto Encoder Neural Network approach. Session 3.3 explains the concept of the fuzzy rules-based data extraction model referred to as Fuzzy Rule Induction Algorithm (FURIA) for the detection of glass breaking events. Section 4 describes the results of the experimental simulation model, and Section 5 summarizes results and discussion of proposed glass break detection system. Section 6 describes conclusion.

2. Methodology

2.1. Fast Fourier transform (FFT)

Every audio wave contains different amplitude and frequency values. Time domain of input audio signal cannot detect the exact periodic frequency of

breaking glass sound in the data. A Fourier transform is a mathematical function allows understanding of the frequency representation of a glass breaking event inside a signal. Therefore, a FFT algorithm converts audio input from the time domain into a frequency spectral representation data make the data more accessible for computational analysis.

The FFT algorithm is an efficient method for subtracting the Discrete Fourier Transform (DFT). DFT is a mathematical operation that precisely identifies and extracts both the frequency and the amplitude of the partial waveform and reduction of the harmonics. To overcome the slowest processing time of DFT, highly efficient computing algorithms, namely the FFT algorithm, is applied on DFT (Van Loan, 1992). The Fast Fourier Transform Algorithm of DFT defines by the formula:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{j2\pi kn}{N}}, k = 0, \dots, N-1. \quad (1)$$

Where N denotes the number of time domain samples in the dataset, n denotes current samples under consideration (0 to $N-1$), x_n denotes values of signals at time n , k denotes current frequency spectrum of considering (0 to $N-1$ Hz), and X_k denotes the amount of amplitude and frequency spectrum k in the signals (Mohlenkamp, 1999). Fig. 1 shows the pattern of the converted frequency domain over time, and Fig. 2 denotes the transformed frequency domain pattern of non-glass break signals over time (ms) determined by using the FFT algorithm.

2.2. Deep auto-encoder neural network

Auto-encoder is the unsupervised learning artificial neural network model for efficient coding. An auto-encoder neural network can learn a set of imprecise data into compressed dimensionality form (encoding), and force reconstruction (decoding) to original data form by learning itself without significant loss in information. A deep auto-encoder neural network is required to find the correlated feature of each input signal and convert it as reduced dimensionality representation vectors before classification of audio signals.

If the input x to an auto-encoder is a vector $x \in R^{D_x}$, then the encoder maps the vector x to another reduced dimensionality vector $z \in R^{D^{(1)}}$ as follows:

$$z^{(1)} = h^{(1)}(W^{(1)}x + b^{(1)}). \quad (2)$$

Where $h^{(1)}$ denotes sigmoid non-linear activation function, $W^{(1)}$ indicates the weight of network, x is the input feature vector, and $b^{(1)}$ indicates as a bias vector. The simple form of an auto-encoder is similar to the multilayer perceptron (MLP) neural network that includes an input layer, an output layer and one or more hidden layers connecting them. However, one difference is that the output layer of auto-

encoder has the same number of nodes as the input layer. Yet, the problem of conventional auto-encoder neural network is the limited performance of processing a very complex non-linear unstructured

dataset (like audio, images). Therefore, we propose deep learning architecture with an auto-encoder neural network for dimensionality reduction tasks.

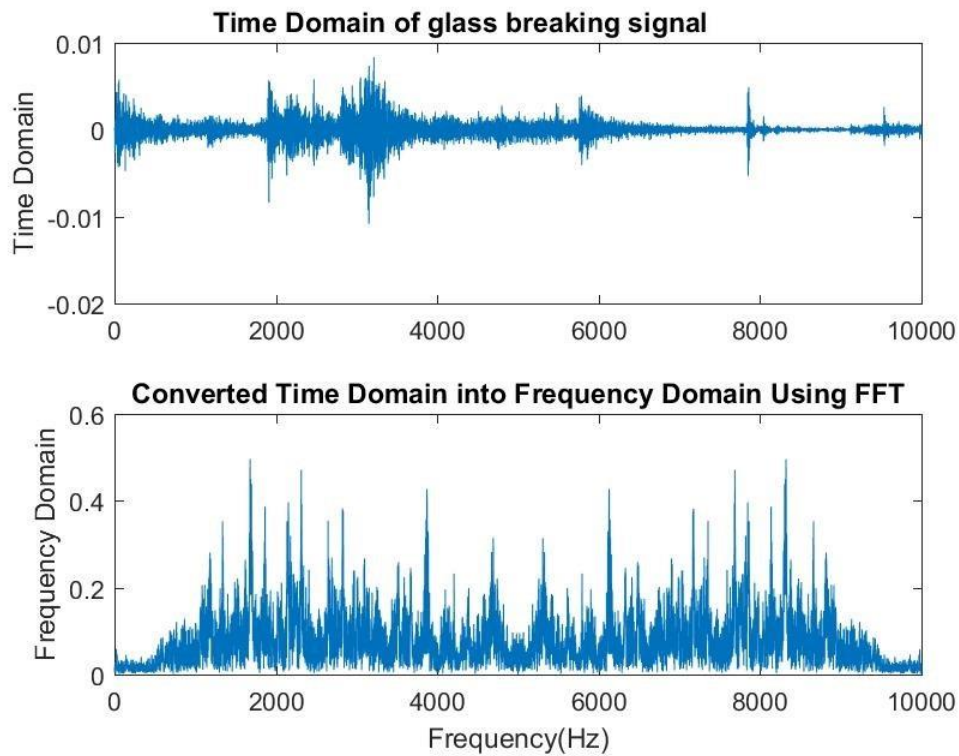


Fig. 1: Converted time domain to frequency domain of glass break signals using the FFT algorithm

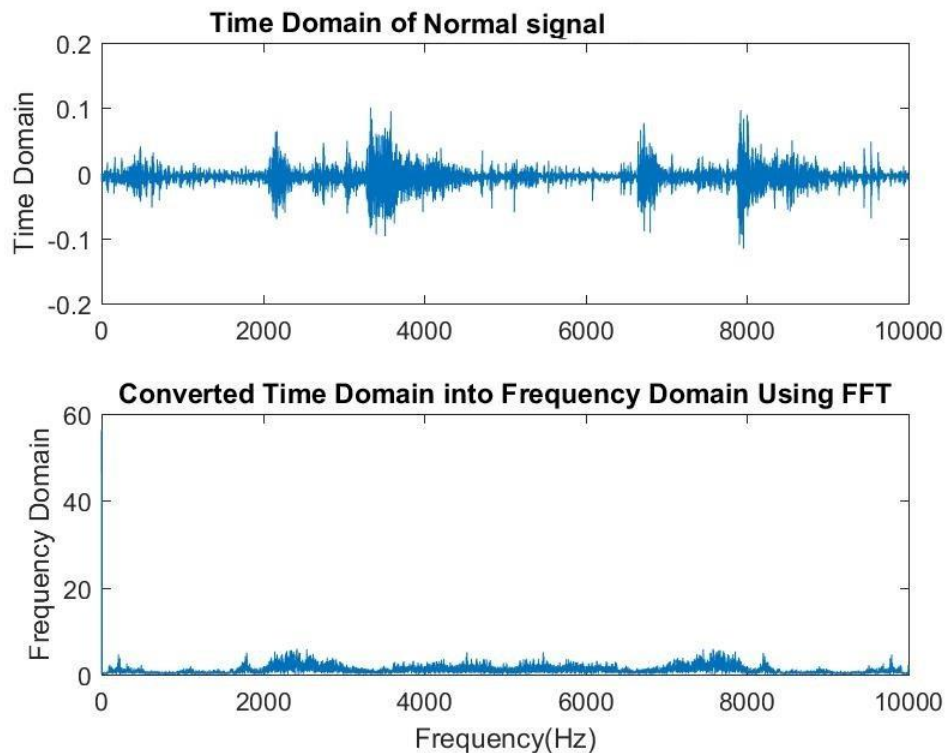


Fig. 2: Converted time domain to frequency domain of non-glass break signals using the FFT algorithm

The basic architecture of the proposed deep auto-encoder neural network is composed of two, symmetrical deep-belief networks (DBN). The encoding part (i.e., data compressed) contains three shallow hidden layers to obtain encoded bottleneck

feature vectors, and the decoding part (i.e., data reconstructed) includes three hidden layers to reconstruct (i.e., original input) as output ([Hinton and Salakhutdinov, 2006](#)).

The learning structure of DBN is that the first encoded output of current layer is considered as the input vector for the next hidden layer until the final expected bottleneck layer of reduced dimensionality is attained. For optimization and reduction in cost error of the encoded and decoded results, the network uses a back propagation learning algorithm with sigmoid activation function in every layer of the training model. As a result, the deep auto-encoder neural network algorithm can produce a useful

representation of encoded data without any information loss. If the data can be represented in fewer dimensions without loss of information, then the classifier can get faster computation performance of analyzing the data and better detection accuracy in the classification step. Fig. 3 denotes encoding and decoding part of symmetrical deep auto-encoder neural network for dimensionality reduction of the input audio feature vector (Tüske et al., 2014).

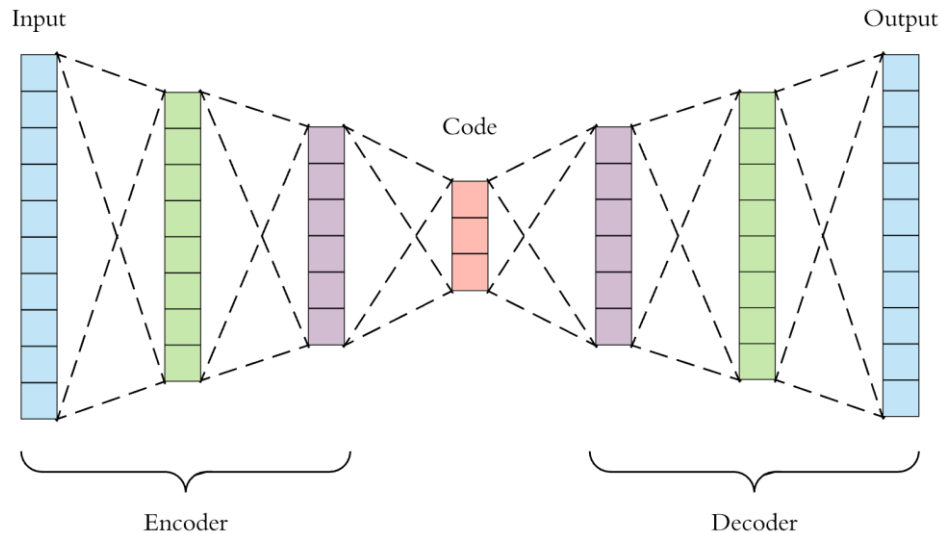


Fig. 3: Combining two symmetrical encoding and decoding part of deep belief network (DBN) using the deep auto encoder neural network algorithm for dimensionality reduction of feature vector (Hinton and Salakhutdinov, 2006)

2.3. Fuzzy unordered rule induction algorithm

A novel fuzzy rule-based classification model called the Fuzzy Unordered Rule Induction Algorithm is an extension and adaptation of the conventional simple rule learner RIPPER algorithm (Hühn and Hüllermeier, 2010). Besides, it includes several modifications and extensions. Particularly, FURIA learns unordered fuzzy rule sets instead of general rule lists. Moreover, to handle uncovered examples, FURIA makes use of an efficient rule stretching method. FURIA automatically analyzes whether there is a statistical difference between different frequency features of glass break and non-glass break signals in the proposed dataset, and extract appropriate membership parameter values for each class. Then, the proposed algorithm creates the fuzzy rules sets by using significance membership parameters of two signals. Classification output of signals decides on a support value for each fuzzy rule set. Fig. 4 denotes the principle Learning Algorithm of FURIA, Fig. 5 explains the antecedent fuzzification algorithm for a single rule pruning, and Fig. 6 extracts rule examples from glass break and non-glass break audio data using fuzzy Rule induction algorithm as follows (Hühn and Hüllermeier, 2009):

3. Experiments

For data acquisition, input audio signals are recorded from an acoustic sensor microphone. Then,

acquired signals are passed through the 5V Amplifier to amplify the input signals.

1. For generating a rule directly from training data (Generate the rules form of fuzzy membership function set from training examples by using covering approach of Ripper Rule Learning algorithm)
2. Rule Pruning (Select the best rules induction of training example (learn-one-rule), using a general to specific (Greedy Search) algorithm for rule pruning)
3. Classification (Define a support value for rules in each signal and decide for classification)

Fig. 4: Learning structure of FURIA (Fuzzy rule induction algorithm) (Hühn and Hüllermeier, 2010)

If the amplitude of the amplified input signal is higher than a threshold, then the output of the comparator is HIGH. The comparator activates the 555 timer which goes on for 5 seconds and wakes up a Micro Controller Processing Unit (MCU) and input audio signals are received. The collected audio signal for glass break detection system has a 44100 sample rate per second over 5 sec time frames. We collected 1000 audio (.wav) samples for data acquisition processing at various noise level environments in Fig. 7. This dataset made of two types sound classes consist of 500 samples of breaking glass sounds data

(breaking glass sounds with different noise level) and 500 non-breaking glass data (combining of shouted sounds, cars horn, household, alarm, animals, farm and child playing, and peoples' conversations).

Algorithm 1 The antecedent fuzzification algorithm for a single rule r

```

1: Let A be the set of numeric antecedents of  $r$ 
2: while  $A \neq \emptyset$  do
3:  $amax \leftarrow \text{null}$  { $amax$  denotes the antecedent with the highest purity}
4:  $purmax \leftarrow 0$  { $purmax$  is the highest purity value, so far}
5: for  $i \leftarrow 1$  to  $\text{size}(A)$  do
6: compute the best fuzzification of  $A[i]$  in terms of purity
7:  $purA[i] \leftarrow$  be the purity of this best fuzzification
8: if  $purA[i] > purmax$  then
9:  $purmax \leftarrow purA[i]$ 
10:  $amax \leftarrow A[i]$ 
11: end if
12: end for
13:  $A \leftarrow A \setminus amax$ 
14: Update  $r$  with  $amax$ 
15: end while

```

Fig. 5: The antecedent fuzzification algorithm for a single rule pruning (Hühn and Hüllermeier, 2009)

Rule Example from Fuzzy Rule Induction of a Non-Glass Break

If 10th feature of membership parameter is [0.049444, 0.049511, ∞ , ∞]
AND (20th feature is [$-\infty$, $-\infty$, 0.050749, 0.050753])
AND (26th features is [0.049497, 0.049512, ∞ , ∞])
AND (17th feature is [$-\infty$, $-\infty$, 0.049749, 0.049751])
Then Label is Non Glass Break (Certainty Factor (CF) = 0.99)

Rule Example from Fuzzy Rule Induction of a Glass Break

If 45th feature of membership parameter is [0.049824, 0.049831, ∞ , ∞]
AND (40th feature is [$-\infty$, $-\infty$, 0.049825, 0.049826])
Then Label is Glass break (CF = 0.92)

Fig. 6: Rule examples from fuzzy rule induction algorithm of glass break and non-glass-break signals (Hühn and Hüllermeier, 2009)

The collected audio signal of glass break detection dataset has different sizes and different audio lengths. Therefore, before analyzing and pre-processing the dataset, this work applied a normalization process. The collected time domain of the audio sample in the dataset normalizes into 16384 (14 bit) resample per 250 milliseconds of each signal to make the system invariant to audio length and size. Then, the audio data is converted into a frequency representation form by using the FFT algorithm. After pre-processing, to reduce model complexity and simplify of input audio data, we proposed a Deep Auto-encoder neural network algorithm for dimensionality reduction. Collected

normalized audio signals of the glass break detection dataset are large and complex such as, for example, 1 sample \times 16384 sampling frequency features vectors. To adequately handle this large and complex data in the training state, the proposed deep auto-encoder neural network algorithm is implemented for dimensionality reduction to compress vector size of audio data. For the classification process, the proposed fuzzy rule induction algorithm is applied on data of reduced dimensionality to detect glass breaking audio signals.

4. Result and discussions

For training and measuring accuracy, the audio input data can be divided into two randomly selected groups, such as a training group that corresponds to 70% of the samples and a test group that corresponds to 30% of the pattern. As a result, the generalization ability of the simulated model can be checked after the training phase. Besides, 30 audio signal data are select as invisible data. The True Positive Rate (TP), Error Positive Rate (FP), area under the ROC curve and correct detection rate were used as an accurate measure of the proposed classification criteria.

Fig. 7 denotes step by step classification progress of the proposed glass break detection system (red point denotes glass break, blue point denotes non-glass break signals). Table 1 represents an experimental result of the proposed glass break detection system that attains 95.5% total correct detection for 1000 samples in an audio data set. Table 2 summarizes the detection result of state of the art system performance during the intrusion process. Results indicate that in three test trials the GB sensor failed to detect the glass breakage event 10 times in the testing process yielding a detection accuracy of 65% (Zidan, 2015). Since different researchers in the literature use different datasets, it is not possible to perform any direct comparison.

5. Conclusion

This research proposed a new architecture for a robust glass break detection algorithm by using machine learning and signal processing to improve the detection accuracy for glass break sounds. A deep auto-encoder method is proposed to reduce the size and complexity of audio data signal (without any loss of useful information). An effective rule stretching technique, called the Fuzzy Rule Induction method, is applied to detect a glass breaking event based on membership parameters of input signals. The new fuzzy rule stretching algorithm is simpler than existing hitherto techniques and can be used more efficiently when comparing existing rules. According to experiments, classification result of the glass break dataset is the attainment of correct detection accuracy of 95.5% which indicates that proposed fuzzy deep auto-encoder neural network algorithm is significantly better than the former conventional machine learning based detection

system and other sensor-based glass break detection techniques. Future work may cover real-time protection of home and office security systems based

on proposed machine learning based glass break detection approaches integrated with other security devices.

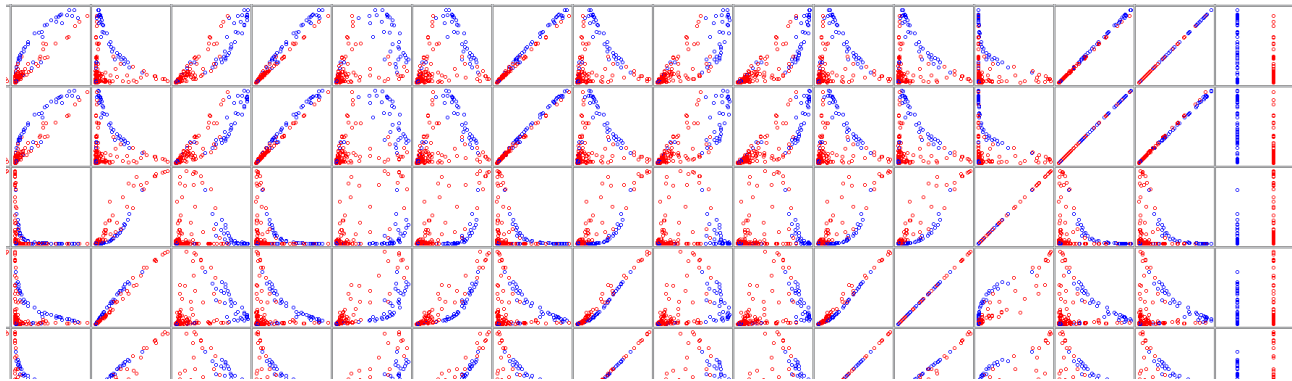


Fig. 7: Classification result of glass break detection system using deep fuzzy auto-encoder neural network (Red point denotes glass break, blue point indicates non-glass break signals)

Table 1: Experimental result of proposed glass break detection system using deep auto-encoder with FURIA

Correct Classification Instances	Incorrect Classification Instances	Label
480	20	Glass-Break
475	25	Non-Glass-Break

Table 2: Experimental results of glass break detection system using conventional glass-break detector (Zidan, 2015)

Number of Trial	Success Detection Alarm	False Alarm	Accuracy
10	6	4	60%

Acknowledgement

This work was supported by the Ministry of Higher Education Malaysia under PRGS17-002-0042 and International Islamic University Malaysia under RIGS16-350-0514.

Compliance with ethical standards

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Dufaux A, Besacier L, Ansorge M, and Pellandini F (2005). Automatic sound detection and recognition for noisy environment. In the 10th European Signal Processing Conference, IEEE, Tampere, Finland: 1033-1036.
- Eskildsen KG (2005). Method and system for detecting a predetermined sound event such as the sound of breaking glass. U.S. Patent No. 7,680,283, Patent and Trademark Office, Washington, D.C., USA.
- Hinton GE and Salakhutdinov RR (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786): 504-507. <https://doi.org/10.1126/science.1127647> PMID:16873662
- Hühn J and Hüllermeier E (2009). FURIA: An algorithm for unordered fuzzy rule induction. *Data Mining and Knowledge Discovery*, 19(3): 293-319. <https://doi.org/10.1007/s10618-009-0131-8>
- Hühn JC and Hüllermeier E (2010). An analysis of the FURIA algorithm for fuzzy rule induction. In: Koronacki J, Raś ZW, Wierchoń ST, and Kacprzyk J (Eds.), *Advances in machine learning*: 321-344. Springer, Berlin, Germany. https://doi.org/10.1007/978-3-642-05177-7_16
- Lopatka K, Kotus J, and Czyzewski A (2016). Detection, classification and localization of acoustic events in the presence of background noise for acoustic surveillance of hazardous situations. *Multimedia Tools and Applications*, 75(17): 10407-10439. <https://doi.org/10.1007/s11042-015-3105-4>
- Mohlenkamp MJ (1999). A fast transform for spherical harmonics. *Journal of Fourier Analysis and Applications*, 5(2-3): 159-184. <https://doi.org/10.1007/BF01261607>
- Richard AS (2011). Highly directional glassbreak detector. U.S. Patent No. 13,105,026, Patent and Trademark Office, Washington D.C., USA.
- Sidhu AS (2005). The rise of crime in Malaysia: An academic and statistical analysis. *Journal of the Kuala Lumpur Royal Malaysia Police College*, 4: 1-28.
- Smith RA (2004). Method of eliminating impact/shock related false alarms in an acoustical glassbreak detector. U.S. Patent No. 7,388,487, Patent and Trademark Office, Washington D.C., USA.
- Stelzer I (2010). High performance laminated glass. In the *Challenging Glass Conference Proceedings*, Delft University of Technology, Delft, Netherlands, 2: 465-474.
- Tüske Z, Golik P, Schlüter R, and Ney H (2014). Acoustic modeling with deep neural networks using raw time signal for LVCSR. In the 15th Annual Conference of the International Speech Communication Association, Singapore: 890-894.
- Van Loan C (1992). Computational frameworks for the fast Fourier transform. Volume 10, Siam, Philadelphia, Pennsylvania, USA. <https://doi.org/10.1137/1.9781611970999>
- Venkat K (2010). Glass-breakage detector uses one microcontroller. *EDN*, 55(17): 39-45.
- Wang K, Yang L, and Yang B (2017). Audio event detection and classification using extended R-FCN Approach. Technical Report, Challenge, Detection and Classification of Acoustic Scenes and Events, Munich, Germany.
- Zidan WI (2015). Estimation of cluster sensors' probability of detection for physical protection systems evaluation. *Journal of Physical Security*, 8(1): 40-54.