2021 5th International Conference on Information Systems and Computer Networks (ISCON)
GLA University, Mathura, India. Oct 22-23, 2021

Human Action Classification using seismic sensor and machine learning techniques

1st Mainak Chakraborty
School of Mechatronics and Robotics
Indian Institute of Engineering Science And Technology
Shibpur, India
mainakchakraborty92@gmail.com

2nd Srinivasan A
Robotics and Automation Group
CSIR- CMERI
Durgapur, India
srinivaspsg@gmail.com

3rd Srinivasa Reddy
Robotics and Automation Group
CSIR-CMERI
Durgapur, India
reddynsrinivasa@gmail.com

4th Sanjib Kumar Mandal
Jagadish Chandra Bose Centre of Advanced Technology
Defence R and D Organization (DRDO)
Kolkata, India
sanjibkumar.mandal@jcbcat.drdo.in

5th Subhasis Bhaumik
Aerospace Engineering and Applied Mechanics
Indian Institute of Engineering Science and Technology
Durgapur, India
subhasis@aero.iiests.ac.in

2021 5th International Conference on Information Systems and Computer Networks (ISCON) | 978-1-6654-0341-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISCON52037.2021.9702317

***Abstract*—In this work, we propose a novel method for identifying human activities like walking and running, utilising ground vibration obtained from seismic sensor and using machine learning techniques. The proposed methodology is based on statistical feature extraction. It is grounded on the idea, that each of the activity generate distinctly unique seismic signatures. We curated a unique dataset of various activites using off-the-shelve geophones and 16 bit analog to digital converter. The seismic data were sampled at 10000Hz. The datasets we recorded is 2 min long each. Single and 6-channels sensors are used to record the dataset. The data recorded is denoised using bandpass filter. The denoised data is used to detect peaks by comparing with neighbouring values. Peak based segementation is done and various statistical features are extracted from the dataset. The effectiveness of feature extraction is increased by converting the data in spectral domain, by extracting the power spectrum. The feature extracted are labelled and is used to train machine learning models. We have explored different machine learning algorithms and Random Forest algorithm gives the accuracy of** 91.62%**.**

***Index Terms*—seismic sensor, Human activity classification, machine learning, classification**

I. Introduction

Human Action Classification is primarily focused on the detection of activities through vision and sensor based methods. However, in vision based there are limitations such as low illumination, line of sight and privacy. Wearable sensor needs to be in contact with the user. On the other hand, Seismic sensor does not have such limitations and has an added advantage of being low cost and passive in nature [1]. Seisimic sensor is historically used for macroseismic event detections. In 2000, it has been observed that a person walking and vehicle transversing can be modeled as a series of vertical impacts made on the ground. A comparison with acoustic and seismic data obtained of same event, shows identical variation [2]. Slow Adaptive Thresold and Quick Adaptive Theresold has been proposed as a effective algorithm for footstep detction [3]. Where, Slow adaptive theresold is the average of last 1000 data points, and quick adaptive threshold is the average of last 30 datapoints. A comparison between the two is used to detect footstep. UREDT has been proposed as a four step algorithm, based on clustering technique, for event detection. 135 features are extracted in frequency domain and 4 in time domain. However, it is computationally high in nature [4]. A method for peak based segmentation, for personalisation of classification models is proposed. The dataset is seperated into segments based on peaks. 18 handcrafted features are extraction to train the learning algorithms. They implemented and calculated accuracy on various open source Human Activity Recognition Datasets [5]. Another method is proposed by seperating the dataset into basic and translational activity, by windowing techique. K-means clustering algorithm is used for clustering and Random Forest classification technique achieves the highest accuracy [6]. A classification of human walking, human running, car and bike exploiting the vertical and horizontal channels. It is observed that the horizontal channel gives better analytical results for car and bike . Whereas the vertical channels gives good results for walking and running. Power Spectral density is proposed as a sensitive feature [7]. Symbolic dynamic filtering has been used alongwith learning techniques to identify various targets. A tree-based structure is proposed to classify activity at each node. A multi-class classification problem is transformed into multi binary-classification [8]. The combine use of time and frequency domain analysis by extracting statistical features like mean, standard deviation, skewness and kurtosis is done for personal detection. Hilbert Transform is used to obtain an envelope of the seismic signal. GMM based clustering is used to extract events. The majority of the footstep event is observed to be concentrated into 0-250Hz [1]. CNN has also been proposed for the classification of vehicle signature [9]. They modified the mel-frequency cepstral coefficient to a linear scale and used particle swarm optimisation to optimise. Spectral analysis has been shown as the effective way for various activity classification. Unconventional targets like helicopter, ship and train also generate a typical seismic data [10] [11]. A Kurtosis based methodology is used for detection of footsteps. It mainly distinguishes the walker detection or non-detection on the ratio of the fourth moment to the second moment squared. Once the detection criteria is satisfied i.e. Kurtosis threshold is reached, a regular cadence is investigated in human footstep data using a Cestrum based analysis. If the second criteria are established then it is classified as a footstep. The results were impressive as it managed to detect footsteps from long range [12] [13]. CNN is also projected as an effective way to analyse seismic signal, by compressing dataset and converting the signal-based problem into image based classification [14] [15] [16]. Indoor localisation and fall detection is done by wavelet denoising technique and 13 features are calculated in both temporal and spectral domain, multi-class SVM is used for classification [17]. Intrusion detection has been done using SVM-RBF has achieved an accuracy of 77% and prediction of the state of motion has achieved an accuracy of 86% [18]

The paper is organised as follows. In section II describes the complete methodology. In section III the experimental analysis and results are presented. Section IV concludes the paper.

II. Proposed Method

In this proposed method, we stressed on kurtosis as a single parameter to classify between no movement and human movement. Data is denoised by using bandpass filter. Denoised data is enhanced based on peak. Statistical features is then extracted on both time and spectral domain. It is then divided into training and testing data to train machine learning algorithms.

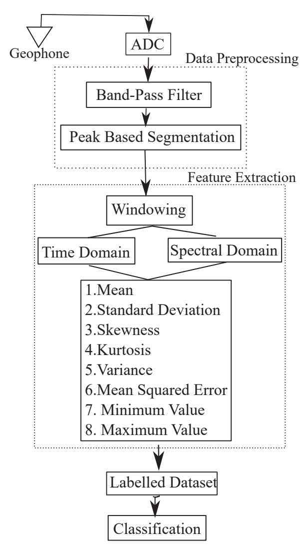

Fig. 1. Complete overview of the proposed human activity classification technique

A. Data Collection

Data is collected using a off-the-shelf geophones and measurement computing 16 bit analog to digital convertor(ADC) USB1608Fs plus is used. The sampling rate is maintained at 10000Hz. Each sensor is buried on the ground to a depth of 50cm. During data collection each individual is made to walk around within the sensing radius of 1.5m. Each dataset recorded is of duration 2 min and saved in .csv format. (Due to the sampling rate of 10000, in each sec 10000 data points are recorded. In 2 min (120*10000) points are recorded). During each experiment, other activities were intentionally supressed to acquire clean data. Both six channels and single sensors are used to record data. In case of six channels geophones, each geophone is kept at a distance of 10m from each other and arranged in the following fashion (0,5), (10,5), (20,5), (30,5), (40,5), (40,0). Three different locations for collection of data i.e. barren land, garden and parking. While the first two is chosen for its low level of seismic activities. The latter is chosen for its moderate to high level of seismic noises. Five individual volunteered for the experiments.

B. Kurtosis Thersolding

Kurtosis is the ratio of fourth central moment to the fourth power of standard deviation. Kurtosis verifies whether the dataset has extreme values or not [19]. It requires a pre-defined theresold, in order to classify between activities [13]. Kurtosis of no movement is calculated and then averaged, to obtain the theresold value. After averaging various kurtosis value of no activity regions recorded on varied time and location, the thersolding value we decided is 4.5. During validation, once the thersold value is crossed the algorithm identifies it as a activity.

$$Kurtosis = \frac{\widehat{\mu}}{\sigma^4} \tag{1}$$

where, $\widehat{\mu}$ = fourth central moment and σ= standard deviation

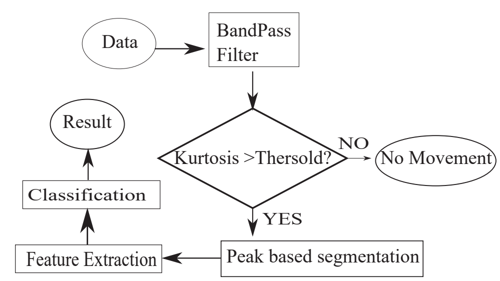

Fig. 2. Validation method based on Kurtosis thersolding

C. Data Preprocessing

The data is collected and pass through bandpass filter. Butter band-pass filter (4-250Hz) of fourth order is used. It is observed that the majority of the human footstep is concentrated between 0-250Hz [1]. The data is processed and local maximas of the entire dataset are detected by comparing with neighbouring values. The instances detected are segmented and appended with preceding 50 datapoints. The number 50 is chosen emperically. Increasing the number will increase the size of dataset. Peak based segmentation has been proven to increase the overall efficiency of the model [5]. The steps are given below :

Step 1: Take seismicdata as an input array in 1D format
Step 2: if Seismicdata (i+1)> Seismicdata (i)
Step 3: peak ← Seismicdata (i+1)
Step 4: Resultant ← Seismicdata(from peak−50 to peak)
Step 5: Resultant is the output array in 1D format

D. Feature Extraction

Before extracting feature the dataset is divided by windowing technique. Window length is chosen 7500 datapoints at a time, with 50% overlap. Features from each window is calculated and included in a seperate column which is stored in a seperate .csv file. Total of 8 handcrafted features are extracted from both domains. The features are skewness, kurtosis, mean of squared error, mean, standard-deviation, variance, minimum value and maximum value. In spectral domain, power spectrum is calculated by sub-dividing the data into overlapping segments and computing a modified periodogram for each divisions. Afterwards averaging the periodograms, as per welch's formula [20]. Data is seperated into 256 length segments and the spectrum of each section is computed. Hanning window is used with 50% overlap. This is a reasonable trade-off between power accuracy and excluding over counting of data. The spectral data is divided by 8 datapoints, to calculate the features with the help of sliding window technique. Figure 3, summarises the entire process.

$$P_{x_m,M(w_k)} = \frac{1}{M} \mid FFT_{N,k}(x_m) \mid^b \tag{2}$$

where, $P_{x_m,M(w_k)}$ denotes the periodogram of the mth block, and $\widehat{S}_x^W$ is the welch estimate of power spectral density

$$\widehat{S}_x^W = \frac{1}{K} \sum_0^k P_{x_m,M(w_k)} \tag{3}$$

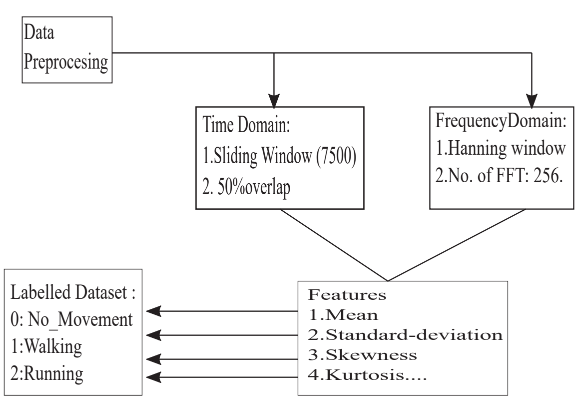

Fig. 3. Feature Extraction Methodology

The primary features extracted are listed below:

1.Mean: $= \frac{1}{N}\sum_1^N(x_i)$
2.Standard Deviation: $\sigma^2 = \frac{1}{N}\sum_1^N(x_i - \mu_x)^2$
3.Minimum Value: $\min(x)$
4.Maximum Value: $\max(x)$
5.Variance: $\sigma = \frac{1}{N}\sum_1^N(x_i - \mu_x)^2$
6.Mean Sqaure Error: $mse = \frac{1}{N}\sum_1^N((x_i - \widehat{x}_i)^2)$
7.Skewness : S= $\frac{\frac{\sum_1^n((x_i-\overline{x})^3)}{N}}{\sigma^3}$
8. Kurtosis : k= $\frac{\widehat{\mu}}{\sigma^4}$

Here, x_i represents the datapoints, $\widehat{\mu}$ represents fourth central moment, N represents number of points σ represents standard deviation and μ_x represents mean of the dataset

E. Labelling and train-test split

Each activity is labelled as a seperate class. The eight features of each window is calculated and labelled according to the activity. Each Columns represents single feature and rows represent windowed activity. If total number of features are f_i and the total number of windowed segment are w_i. The features extraction process will generate $w_i * f_i$ matrix in one particular domain. As shown in figure 4, the data is seperated into 80:20 for training and testing on various algorithms. The algorithms used are Random Forest [21], Support Vector Machine (linear) [22], Support Vector Machine (rbf) [22], Nearest Neighbour [23], multilayer perceptron [24] [25]. Hereafter, two classification algorithm is described briefly Random Forest and Support Vector Machine. In validation phase, the random datasets are processed through the same process, features are extracted and cross-checked with the model prepared beforehand.

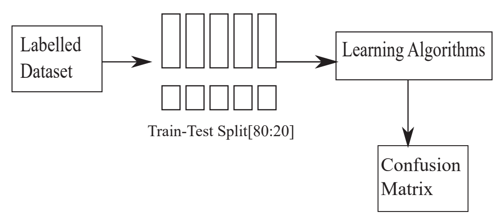

Fig. 4. Training process of various classification algrithms

F. Random Forest

Random Forest is an extension of decision tree algorithm [21]. Like decision tree, it is also split on the basis of entropy, chi-sqaure or gini index. In Random Forest , it takes bootstrap sample from the entire dataset and create multiple trees. Since datasets selected are different, results of each decision tree is also different. The decision of each tree, is then averaged for a regression and majority is chosen for a classification problem. Random Forest shuffles through the features, with some portions repeating more than once. This makes Random Forest relatively robust to outliers and noise, which is crucial in seismic data [26]. It is trained on different parts of the same data, with a primary objective to reduce variance. Unfortunately, this increases the bias. However, it boosts the performance of the entire model compared to a single decision tree [27]. Entropy is used as our decision criteria. Random Forest achieves the highest accuracy of 91.6%, with a depth of 8 and number of trees 100.

$$Entropy = -\sum_{1}^{n}(p_i \star \log_2 p_i) \tag{4}$$

where, p_i is the probabillity of belonging to ith class.

G. Support Vector Machine

Support Vector machine is a simple algorithm that divides the datapoint based on a n-dimensional hyperplane. Kernel functions creates different hyperplane. The hyperplane that maximises the width or difference between points of seperate class reduces the error function [28]. We have used both linear and Radial Basis Function Kernel to calculate accuracy. When the regularization hyperparameter C=1, L-SVM gives 86.4% accuracy and RBF-SVM achieved an accuracy of 80.2% for C=100. A multi-class Support Vector Classification is implemented by using libsvm library [22].

$$minL_p = \frac{1}{2} \parallel w^2 \parallel - \sum_{1}^{l} a_i y_i (x_i.w + b) + \sum_{1}^{l} a_i \tag{5}$$

where, w is the "maximum margin" parameter, that is being optimised.

III. Experimental Analysis and Discussion

Each sample is recorded with a length of 2min, sampling frequency 10000Hz (12 lakh datapoints each). Each 10000 datapoints physically represents 1 second. Since, we are calculating 7500 data points, we are calculating features for 750 millisecond at a time. No movement, walking, running are labelled as seperate classes. The entire experiment is divided into three parts. Training, testing and validation. Training and testing is done by separating a single dataset into 80:20. Whereas, in validation, we are using completely different datasets, recorded in different location. We are using no movement data as a seperate class. The kurtosis theresolding is the first checkpoint, to avoid unnecessary calculations. Figure 5 shows the experimental setup with laptop and Data Acquisition unit.

The accuracy of various models are shown in Table I. The number of trees for random forest is 100. Max depth is kept at 6. These hyperparameters are decided emperically. The entropy criterion is used for calculation. In Support vector machine we have used two kernels 'linear' and 'radial basis function'. L-SVM achieves 86.4% accuracy and RBF-SVM 80.2%. Decision tree with maximum depth 8 and maximum leaf nodes 20, achieved an accuracy of 87%. For nearest neighbour, we iterated for various cluster values from 3 to 10. It was found out that optimum cluster size is found out as the number of classes or activity used to train the model, which was intuitive. In our case 3. The accuracy achieved

Fig. 5. Experimental Setup

by the KNN is 89.7%. We have also trained a multi-layer perceptron, which achieved an accuracy of 87%, when trained with lbfgs optimiser. However, multi-layer perceptron is a data hungry classification algorithm, which might provide better result when trained with more data. Due to highly uneven dataset (presence of more noise than actual activity), alongside test accuracy we are also calculating True Positive(TP), False Positive (FP), True Negative(TN), False Negative (FN). These parameters are used to calculate the F1 score, Precision, Recall. As shown in figure 6, comparison with data enhancement and raw processed data (output from bandpass filter), it is found out that the data enhancement technique achieves more accuracy.

Prediction outcome

actual value		**p**	**n**	**total**
	p′	True Positive	False Negative	P′
	n′	False Positive	True Negative	N′
	total	P	N	

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$F1score = \frac{2 \times (Precision \times Recall)}{Precision + Recall}) \tag{8}$$

Within a recorded dataset of 2 minute, we are using to validate our model, an individual may walk, walk fast, run or simply stand-still. We are trying to predict each activity by extracting features from single window. Once the prediction is made, we are taking count of which prediction is made. If the dominant activity is running, then the running count will be higher than walking count. Hence, the entire activity

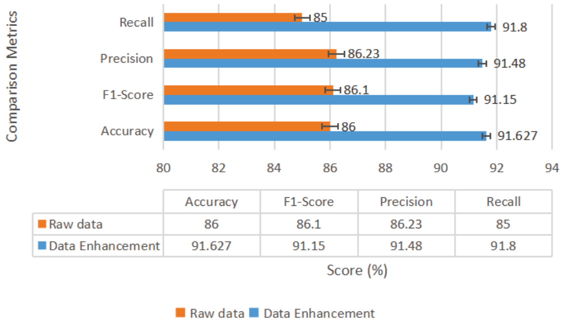

Fig. 6. Comparison of random forest classifier with and without data enhancement technique

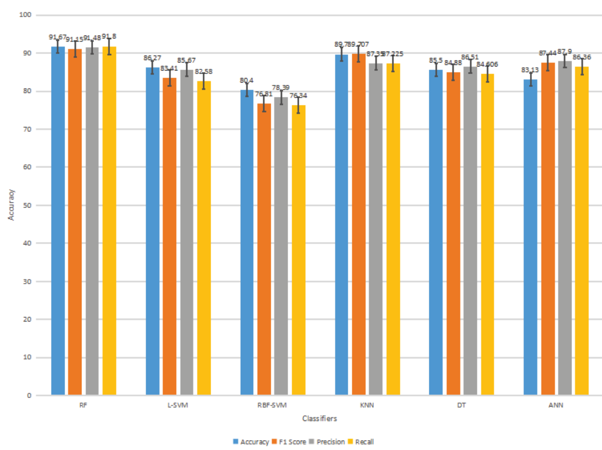

Fig. 7. Performance comparison of various classifiers

is predicted as running. A dataset containing various activity, may or may not contain one dominant activity. As shown in Table I, Random forest achieves the highest test accuracy of 91.62% validation accuracy of 89.88%.

TABLE I
PERFORMANCE MATRICS OF SIX CLASSIFICATION ALGORITHMS

S.No	Distance Method	Classification Accuracy (%)	F1 Score	Precision (%)	Recall (%)
1	Random Forest	91.627	91.15	91.48	91.8
2	SVM(linear)	86.27	83.41	85.67	82.58
3	SVM(rbf)	80.4	76.81	78.39	76.34
4	K-Nearest Neighbour	89.767	87.35	87.49	87.225
5	Decision Tree	85.581	84.88	86.159	84.606
6	ANN	87.44	86.5	87.9	86.36

IV. CONCLUSION

We propose a novel framework for activity recognition based on seismic sensor and learning techniques. Overall accuracy of random forest has surpassed all other machine learning techniques. Our methodology works well with less data. Several factors may affect the activity classification problem like a. Different brand of sensors, b. Different analog to digital convertor, c. Noise levels of region. We have tried to make the classification model as robust as possible, by recording data from different landscapes. Yet it is advisable to train the models on the local dataset. Our model works well in sparsely populated regions. We have used noise as a seperate class to train our models. However, it can be improved by training the model on noise data of the specific regions, where it is applied. It can have tremendous application in border and perimeter surveillance. If huge datasets of activity event are available, data enhancement technique can be skipped. Local maxima based segementation technique, works well when data available is less. We wish to test our work in real-time in future. Human activity classification has been an area of numerous research since last decade. However, the use of seismic sensor in this domain is yet to be explored in its entirety.

V. ACKNOWLEDGMENT

The work reported in this paper is a part of Grant-in-Aid Project (GAP-229812). The Authors would like to thank the sponsoring agency JC Bose Centre for Advanced Technology (JCBCAT), DRDO, Government of India.

REFERENCES

[1] S. Anchal, B. Mukhopadhyay, S. Kar, Person identification and imposter detection using footstep generated seismic signals, IEEE Transactions on Instrumentation and Measurement 70 (2020) 1–11.
[2] G. P. Succi, G. Prado, R. Gampert, T. K. Pedersen, H. Dhaliwal, Problems in seismic detection and tracking, in: Unattended Ground Sensor Technologies and Applications II, Vol. 4040, International Society for Optics and Photonics, 2000, pp. 165–173.
[3] G. Koç, K. Yegin, Footstep and vehicle detection using slow and quick adaptive thresholds algorithm, International Journal of Distributed Sensor Networks 9 (10) (2013) 783604.
[4] S. Anchal, B. Mukhopadhyay, S. Kar, Uredt: Unsupervised learning based real-time footfall event detection technique in seismic signal, IEEE sensors letters 2 (1) (2017) 1–4.
[5] A. Ferrari, D. Micucci, M. Mobilio, P. Napoletano, On the personalization of classification models for human activity recognition, IEEE Access 8 (2020) 32066–32079.
[6] J.-H. Li, L. Tian, H. Wang, Y. An, K. Wang, L. Yu, Segmentation and recognition of basic and transitional activities for continuous physical human activity, IEEE access 7 (2019) 42565–42576.
[7] L. Pucci, E. Testi, E. Favarelli, A. Giorgetti, Human activities classification using biaxial seismic sensors, IEEE Sensors Letters 4 (10) (2020) 1–4.
[8] X. Jin, S. Sarkar, A. Ray, S. Gupta, T. Damarla, Target detection and classification using seismic and pir sensors, IEEE Sensors Journal 12 (6) (2011) 1709–1718.
[9] G. Jin, B. Ye, Y. Wu, F. Qu, Vehicle classification based on seismic signatures using convolutional neural network, IEEE Geoscience and Remote Sensing Letters 16 (4) (2018) 628–632.
[10] K. M. Houston, D. P. McGaffigan, Spectrum analysis techniques for personnel detection using seismic sensors, in: Unattended Ground Sensor Technologies and Applications V, Vol. 5090, International Society for Optics and Photonics, 2003, pp. 162–173.
[11] A. Pakhomov, T. Goldburt, Seismic systems for unconventional target detection and identification, in: Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense V, Vol. 6201, International Society for Optics and Photonics, 2006, p. 62011I.

[12] L. Peck, J. Lacombe, Seismic-based personnel detection, in: 2007 41st Annual IEEE International Carnahan Conference on Security Technology, IEEE, 2007, pp. 169–175.
[13] V. V. Reddy, V. Divya, A. W. Khong, B. P. Ng, Footstep detection and denoising using a single triaxial geophone, in: 2010 IEEE Asia Pacific Conference on Circuits and Systems, IEEE, 2010, pp. 1171–1174.
[14] K. Bin, J. Lin, X. Tong, Edge intelligence-based moving target classification using compressed seismic measurements and convolutional neural networks, IEEE Geoscience and Remote Sensing Letters (2021).
[15] Y. Yu, L. Zhang, J. Shen, Q. Wang, G. Liu, Seismic event detection via deep multi-task learning, in: 2020 International Joint Conference on Neural Networks (IJCNN), IEEE, 2020, pp. 1–6.
[16] S. Jakkampudi, J. Shen, W. Li, A. Dev, T. Zhu, E. R. Martin, Footstep detection in urban seismic data with a convolutional neural network, The Leading Edge 39 (9) (2020) 654–660.
[17] J. Clemente, F. Li, M. Valero, W. Song, Smart seismic sensing for indoor fall detection, location, and notification, IEEE journal of biomedical and health informatics 24 (2) (2019) 524–532.
[18] B. Mukhopadhyay, S. Anchal, S. Kar, Detection of an intruder and prediction of his state of motion by using seismic sensor, IEEE Sensors Journal 18 (2) (2018) 703–712. doi:10.1109/JSEN.2017.2776127.
[19] T.-H. Kim, H. White, On more robust estimation of skewness and kurtosis, Finance Research Letters 1 (1) (2004) 56–73.
[20] P. Welch, The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, IEEE Transactions on audio and electroacoustics 15 (2) (1967) 70–73.
[21] A. Liaw, M. Wiener, et al., Classification and regression by randomforest, R news 2 (3) (2002) 18–22.
[22] C.-C. Chang, C.-J. Lin, Libsvm: a library for support vector machines, ACM transactions on intelligent systems and technology (TIST) 2 (3) (2011) 1–27.
[23] L. E. Peterson, K-nearest neighbor, Scholarpedia 4 (2) (2009) 1883.
[24] G. E. Hinton, Connectionist learning procedures, in: Machine learning, Elsevier, 1990, pp. 555–610.
[25] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830.
[26] M. Pal, Random forest classifier for remote sensing classification, International journal of remote sensing 26 (1) (2005) 217–222.
[27] T. M. Hehn, J. F. Kooij, F. A. Hamprecht, End-to-end learning of decision trees and forests, International Journal of Computer Vision (2019) 1–15.
[28] L. Wang, Support vector machines: theory and applications, Vol. 177, Springer Science & Business Media, 2005.