

Statistical Features and Traditional SA-SVM Classification Algorithm for Crack Detection

Azadeh Noori Hoshyar, Sergey Kharkovsky, Bijan Samali

Centre of Infrastructure Engineering, Western Sydney University, Sydney, Australia
Email: Azadeh.H@westernsydney.edu.au

How to cite this paper: Hoshyar, A.N., Kharkovsky, S. and Samali, B. (2018) Statistical Features and Traditional SA-SVM Classification Algorithm for Crack Detection. *Journal of Signal and Information Processing*, 9, 111-121.
<https://doi.org/10.4236/jsip.2018.92007>

Received: September 1, 2017

Accepted: May 28, 2018

Published: May 31, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In recent years, the interest in damage identification of structural components through innovative techniques has grown significantly. Damage identification has always been a crucial concern in quality assessment and load capacity rating of infrastructure. In this regard, researchers focus on proposing efficient tools to identify the damages in early stages to prevent the sudden failure in structural components, ensuring the public safety and reducing the asset management costs. The sensing technologies along with the data analysis through various techniques and machine learning approaches have been the area of interest for these innovative techniques. The purpose of this research is to develop a robust method for automatic condition assessment of real-life concrete structures for the detection of relatively small cracks at early stages. A damage identification algorithm is proposed using the hybrid approaches to analyze the sensors data. The data obtained from transducers mounted on concrete beams under static loading in laboratory. These data are used as the input parameters. The method relies only on the measured time responses. After filtering and normalization of the data, the damage sensitive statistical features are extracted from the signals and used as the inputs of Self-Advising Support Vector Machine (SA-SVM) for the classification purpose in civil Engineering area. Finally, the results are compared with traditional methods to investigate the feasibility of the hybrid proposed algorithm. It is demonstrated that the presented method can reliably detect the crack in the structure and thereby enable the real-time infrastructure health monitoring.

Keywords

Signal Processing, Machine Learning, Statistical Features, Feature Extraction, Classification, Crack Detection

1. Introduction

The main concern in civil infrastructures is the condition-based maintenance

which requires the consequent monitoring of critical structures in order to minimize the probability of failure leading major effects on national assets and public safety in every country [1]. Therefore, the structural health monitoring is absolutely an essential concept which can potentially provide effective solutions to assess the health state of infrastructure. It can reduce the asset management costs, effectively prolongs operational lifetime and ensures the public safety. In this regard, getting access to a robust paradigm to deal with aforementioned concerns is of high priority.

The considerable literature [2]-[10] on structural health monitoring has highlighted the crucial importance of damage detection in civil engineering systems in the earliest time.

One of the effective monitoring techniques in determining the health state of structures is an automatic diagnosis system using sensing data which is based on the principle that damage in a structure, e.g. crack or a loosened connection, can modify the dynamic structural responses [11]. The automated condition assessment systems can interpret the large volumes of inspection data to detect and prevent potential structural failure in early stages by minimizing errors to ensure effective risk management, while reducing the asset management costs.

Therefore, the main objective of this paper is to propose a more robust scheme of automated condition assessment system to monitor and evaluate the health state of the structure. For the purpose of this detection system, non-destructive testing (NDT) method which has attracted lots of attention in structural health monitoring for characterizing and assessing the materials and structures is used to characterize the concrete members. Thereafter, the signal processing and artificial intelligent approaches are employed to analyze the data and enhance the accuracy of detection. However, the critical challenge here is the huge number of data which are obtained. The high priority is proposing the potential algorithm to reduce the loss of information while reducing the noise effect. The second concern is the proper extraction of features which can effectively determine the health state of the structure. These features need to be sensitive enough to detect any discontinuity happens in the structure. These features are directly affecting the performance of the classifier. The third challenge is choosing the appropriate classifier that identifies the crack as happened. The question raised here is what features and which classification technique are appropriate to determine the structural condition; what can be measured that correlates to damage, and how to measure and employ the raw measurements to determine the structural condition.

In this paper, the mounted SA based approach is utilized to detect and monitor the cracks in concrete members under loading for simple concrete beams. The collected data are analysed through different steps of pre-processing, feature extraction and classification. After pre-processing through Wavelet signal analyzer, a set of features (feature vector) are proposed through statistical analysis in time domain. In most of researchers the different features are utilized as the fea-

ture set to feed the machine learning approaches, while in this paper the combination of the three damage indexes which are obtained between the healthy and target signal are considered as the feature set of the system. The Self-Advising Support Vector Machine [12] is used to classify and detect the damages in classification step. This machine learning approach has not been ever used in the application of crack detection in civil engineering to determine its effectiveness in the area.

The results will be analyzed and compared with other traditional algorithms to evaluate the accuracy and effectiveness of the system.

2. The Proposed Damage Detection Algorithm

This section explains the proposed damage identification algorithm for SHM. This paradigm consists of four steps: Data Collection, Pre-Processing and Feature Extraction, Classification.

2.1. Data Collection and Pre-Processing

The Experimental analysis is performed on the data collected in Centre of Infrastructural Engineering laboratory (CIE) through using the piezoelectric-based transducers mounted on the concrete beams under static loading. These transducers are transforming the energy of transient stress wave to an electric waveform. The other transducer is used to receive the signal. Three standard concrete beams ($400 \times 100 \times 100$) mm³ have been examined subjected to three-point bending. The recorded signals are pre-processed through signal processing approaches.

2.2. Feature Extraction

After de-noising the signal through the Wavelet signal denoiser tool, the following features are extracted from the signal in time-domain.

1) Difference of gradient between healthy and non-healthy state. In gradient, the central difference for interior data points is calculated.

$$G_D = G_H - G_{NH} \quad (1)$$

$$G_H(:, j) = 0.5 \times (A_H(:, j+1) - A_H(:, j-1)) \quad (2)$$

$$G_{NH}(:, j) = 0.5 \times (A_{NH}(:, j+1) - A_{NH}(:, j-1)) \quad (3)$$

2) Difference of mean between healthy and non-healthy state.

$$M_D = M_H - M_{NH} \quad (4)$$

$$M_H = \frac{1}{N} \sum_{i=0}^n A^i \quad (5)$$

$$M_{NH} = \frac{1}{N} \sum_{i=0}^n A^i \quad (6)$$

3) Difference of root mean square error between healthy and non-healthy state.

$$RMS_D = RMS_H - RMS_{NH} \quad (7)$$

$$RMS_H = \sqrt{\frac{1}{N} \sum_{n=1}^N |A_2|^2} \quad (8)$$

$$RMS_{NH} = \sqrt{\frac{1}{N} \sum_{n=1}^N |A_2|^2} \quad (9)$$

2.3. Classification

The Support Vector Machine (SVM) and the traditional Self-Advising Support Vector Machine (SA-SVM) is used to classify the extracted features. The brief descriptions of these two learning machines are expressed as follows.

2.3.1. Support Vector Machine

Support Vector Machine has been proven as a powerful technique of classification in many research areas. This statistical machine learning approach was firstly presented in 1995 by Vapnik [13]. This learning algorithm prevents over fitting by selecting a particular hyperplane among many which can separate the data in feature space. SVM applies a linear separating hyperplane for creating a classifier to maximize the margin. The width of margin between the classes is considered as the optimization criterion. Margin is defined as the distance of optimal hyperplane and nearest training data points of a class. In cases of non-linearly separation of original input space, in 1992, Guyon, Boser and Vapnik presented an approach to generate nonlinear classifiers using kernel functions [14]. SVM firstly transforms the original feature to a higher dimensional feature space. The transformation may be obtained through different nonlinear mappings. The kernel function $K(x; y)$ may be chosen to suit the problem. After this transformation, the optimal hyperplane may easily found. The achieved hyperplane is the optimal case with respect to a maximal margin [14].

In this paper, SVM is selected as the method of choice as it linearly classifies data in a high dimensional feature space which is related nonlinearly to the input space using specific Radial Basis (RBF) kernel.

2.3.2. Traditional Self-Advising Support Vector Machine

The Self-Advising Support Vector Machine is a newly developed Support Vector Machine introduced in 2014 [12]. This machine uses the prior knowledge in its decision making to separate one class from the others. This classifier receive some additional information through finding the misclassified data in training phase and use this knowledge along with the training data to make the classification decision. This information are the weights assigned to the outliers.

The other contribution of this paper is to investigate the application of this algorithm in damage detection particularly crack detection in civil engineering area with the extracted features. In the SA-SVM, the outliers are defined in training phase as [12].

$$MD = \bigcup_{i=1}^n (x_i | y_i) \neq \text{sign} \left(\sum_{\alpha_j > 0} y_j \alpha_j k(x_i, x_j) + b \right) \quad (10)$$

Although the misclassified data can be null, different experiments has proven that it mostly includes data. The neighbourhood length (NL) for each MD is formulated as [12],

$$NL(x_i) = \text{minimum}_{x_j} (\|x_i - x_j\| | y_i \neq y_j) \quad (11)$$

The distance is computed in higher dimension

$$\|\theta(x_i) - \theta(x_j)\| = (k(x_i, x_i) + k(x_j, x_j) - 2k(x_i, x_j))^{0.5} \quad (12)$$

Finally, the proposed weighing for the data are computed as [12],

$$\begin{cases} 0, & \forall x_i \in MD, \|x_k - x_i\| > NL(x_i) \text{ or } MD = MUL \\ 1 - \frac{\sum_{x_j} \|x_k - x_j\|}{\sum_{x_j} NL(x_j)} & x_i \in MD, \|x_k - x_i\| \leq NL(x_i) \end{cases} \quad (13)$$

3. Results

In this section, the results of the proposed algorithm are investigated. **Figure 1** shows the de-noised signal through wavelet de-noiser. The red signal is the signal obtained through the sensor on concrete beam and the blue signal shows the de-noised signal through the Wavelet de-noiser.

The result shows the effectiveness of this de-noising tool on the experiment obtained data.

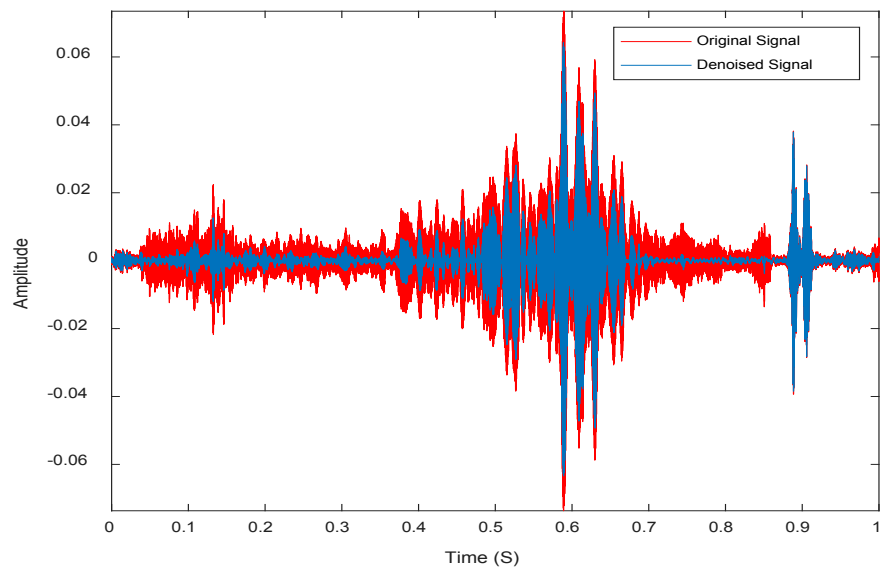


Figure 1. The original and de-noised signal.

Figure 2, **Figure 3** and **Figure 4** show the difference of gradient, mean and root mean square error of recorded signals between healthy and non-healthy state of concrete after 15 min of loading.

The peaks of differences between healthy and non-healthy state are the indicators of the crack.

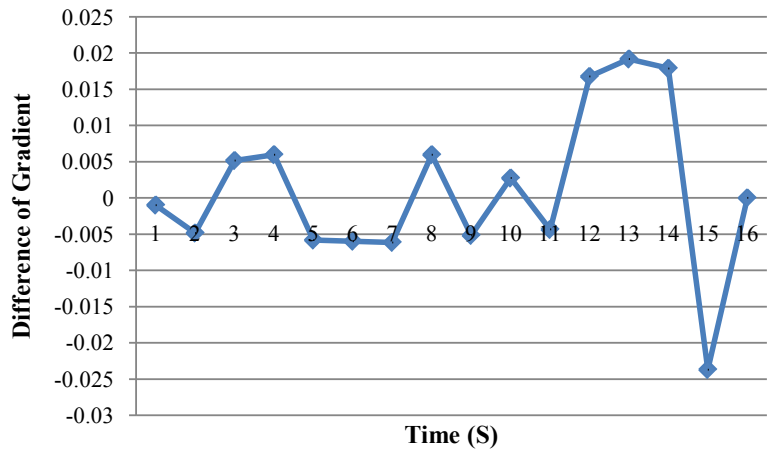


Figure 2. Difference of Gradient between healthy and non-healthy state of the concrete beam.

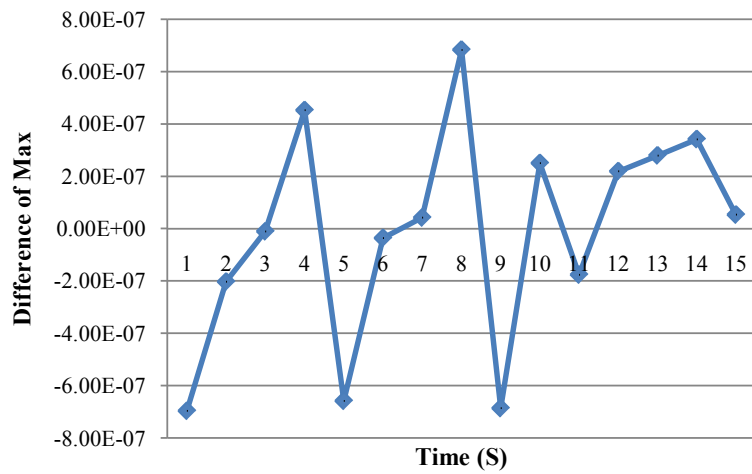


Figure 3. Difference of Max between healthy and non-healthy state of the concrete beam

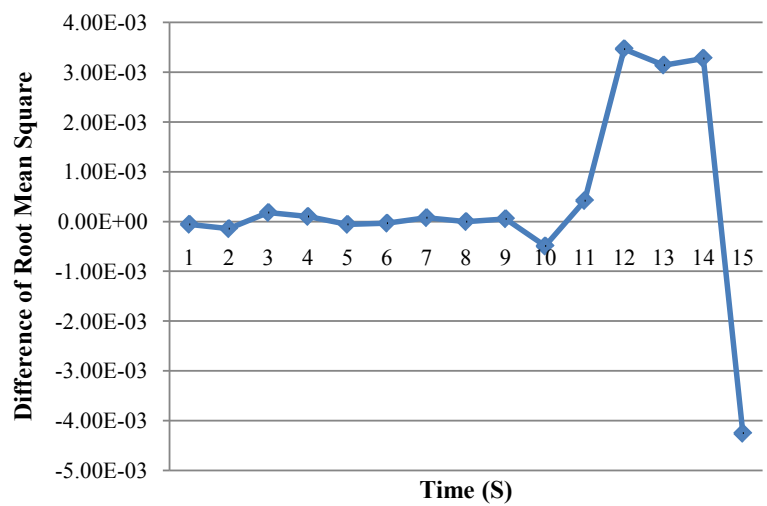


Figure 4. Difference of Mean square error between healthy and non-healthy state of the concrete beam.

However, each of the differences of gradient, mean and mean square error demonstrated in above figures can be separately considered as the damage index to provide the health state of the concrete. They show the sharp changes when the damage occurred. However, the difference of gradient and the difference of mean square error are able to show the main occurrence of the damage, while the difference of max is able to determine the micro cracks.

Therefore, in this paper to take the advantage of all three damage indexes, the combination of three damage indexes has been taken into account to provide the more affective damage recognition. In this regard, the extracted features through Equation (1) and Equation (4) and Equation (7) which computes these damage indexes between the healthy and non-healthy state of the concrete beam are considered as the feature vector. This feature vector is sensitive to any change that may happen in the structure.

Figure 5 shows the comparison of gradient, mean and mean square error between healthy and non-healthy concrete beam in one graph. The comparison shows that the mean square error difference has the highest value than the other two mean and gradient differences.

The green curve which represents the mean square error reaches the highest value more than 0.02, while the other two curves values are less than 0.02. It shows that the mean square error shows the differences between the healthy and non-healthy concrete beam more accurately.

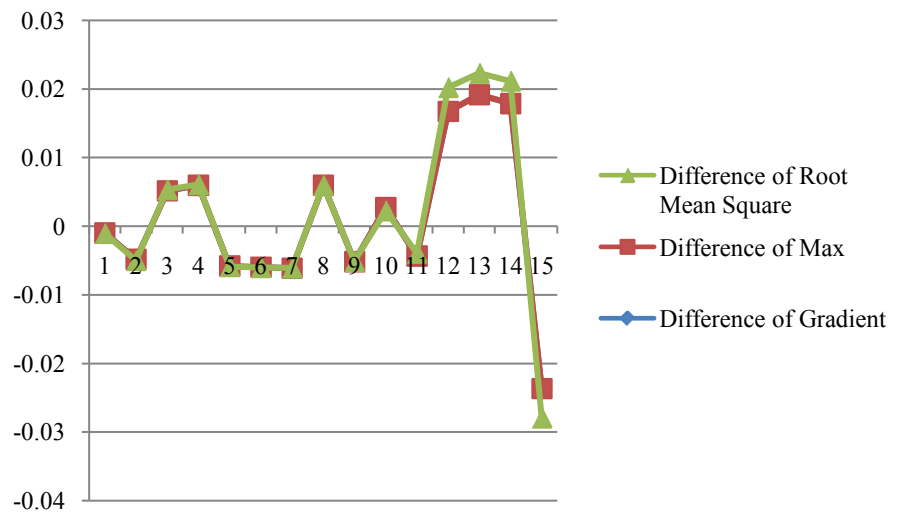


Figure 5. Comparison of features.

The obtained feature vector is trained through the SVM and traditional SA-SVM.

3.1. Experimental Result for “Feature Extraction and SA-SVM Classification”

Table 1 indicates the average performance of the features in detecting the dam-

age by our proposed algorithm “GMM/SA-SVM classification” in 10 experiments (validation for each experiment has performed by 100 times of running program).

Based on the results, average accuracy and F-score of “our proposed approach” are 83.54377 and 0.54242, respectively.

Table 1. Average performance of each set of features by our proposed algorithm.

Exp No.	Accuracy	F-Score
	GMM/SA-SVM classification	GMM/SA-SVM classification
1	83.4476	0.5442
2	83.7218	0.5664
3	82.2968	0.4199
4	83.7224	0.5177
5	83.2968	0.5367
6	84.8524	0.6231
7	83.9524	0.5450
8	84.2589	0.6280
9	82.6524	0.5179
10	83.2362	0.5253

3.2. Comparing the Results of “Feature Extraction and SA-SVM Classification” and “Feature Extraction and SVM Classification”

For the purpose of comparison, the SVM is also trained with the same set of features employed for our proposed algorithm results in **Table 1**. The results of “GMM/SVM classification” are shown in **Table 2** and compared with the results of “GMM/SA-SVM classification”. Based on the results, average accuracy and F-score of “GMM/SA-SVM classification” are 83.54377 and 0.54242, and in “GMM/SVM classification” approach are 83.14774 and 0.52281, respectively. The results show the better performance of our proposed algorithm of “GMM/SA-SVM classification” than another algorithm. Our achieved results are confirmed by the results achieved by the reference paper we used for SA-SVM [9]. Both the reference paper and our results show a similar pattern in performance of SVM and SA-SVM. The t-test is also carried out as another analysis factor to show the effectiveness of the proposed algorithm rather than the traditional one. **Table 3** shows the superiority of our proposed algorithm rather than the traditional one through the t-test.

Table 2. Average performance comparison between “GMM/SVM classification” and “GMM/SA-SVM classification” for each set of features.

Exp. No	Accuracy		F-score	
	GMM/SA-SVM classification	GMM/SVM classification	GMM/SA-SVM classification	GMM/SVM classification
1	83.4476	82.247	0.5442	0.5312
2	83.7218	82.1238	0.5664	0.5623
3	82.2968	81.2286	0.4199	0.4194
4	83.7224	82.9524	0.5177	0.5143
5	83.2968	82.9952	0.5367	0.4061
6	84.8524	83.9514	0.6231	0.5222
7	83.9524	83.3514	0.5450	0.5330
8	84.2589	83.15	0.6280	0.5146
9	82.6524	82.2524	0.5179	0.5009
10	83.2362	83.2252	0.5253	0.4241

Table 3. T-Test: two-sample assuming equal variances.

	GMM/SA-SVM classification	GMM/ SVM classification
Mean	83.54377	82.74774
Variance	0.552091	0.611033
Observations	10	10
Pooled Variance	0.581562	
Hypothesized Mean Difference	0	
df	18	
t Stat	2.334083	
P(T ≤ t) one-tail	0.015691	
t Critical one-tail	1.734064	
P(T ≤ t) two-tail	0.031381	
t Critical two-tail	2.100922	

The p-value is less than 0.05 which indicates that the difference between the obtained accuracy of the traditional and proposed algorithm in this paper are significant.

4. Conclusion

The structuring of automatic damage detection system in civil structures has been proposed and developed. The experimental study is performed to identify cracks in a concrete beam under static loading in laboratory condition. The piezoelectric-based transducers have been used to present the discontinuities inside

the structure. The hybrid statistical feature analysis algorithm based on the combination of three damage indexes which are the differences of gradient, mean, mean square error between the healthy and non-healthy state of the structure along with the Self-Advising Support Vector Machine is carried out on the obtained signals to determine the health state of the structure. The result shows that the proposed algorithm is able to determine the state of the structure more accurately when compared with traditional algorithms. This algorithm can be applicable in other area of research. Further improvements on accuracy can be achieved through utilizing the more features in the feature set vector.

References

- [1] Stephan, P. and Salin, J. (2012) Ageing Management of Concrete Structure: Assessment of EDF Methodology in Comparison with SHM and AIEA Guides. *Construction and Building Materials*, **37**, 924-933.
- [2] Bornn, L., Farrar, C.R., Park, G. and Farinholt, K. (2009) Structural Health Monitoring With Autoregressive Support Vector Machines. *Journal of Vibration and Acoustics*, **131**.
- [3] Phinyomark, A., Nuidod, A., Phukpattaranont, P. and Limsakul, C. (2012) Feature Extraction and Reduction of Wavelet Transform Coefficients for EMG Pattern Classification. *Electronics and Electrical Engineering*, **122**, 1392-1215. <https://doi.org/10.5755/j01.eee.122.6.1816>
- [4] Wang, Q. and Deng, X. (1999) Damage Detection with Spatial Wavelets. *International Journal of Solids and Structures*, **36**, 3443-3468. [https://doi.org/10.1016/S0020-7683\(98\)00152-8](https://doi.org/10.1016/S0020-7683(98)00152-8)
- [5] Santos, J.B. and Perdigão, F. (2001) Automatic Defects Classification—A Contribution. *NDT & E International*, **34**, 313-318. [https://doi.org/10.1016/S0963-8695\(00\)00043-8](https://doi.org/10.1016/S0963-8695(00)00043-8)
- [6] Balageas, D., Fritzen, C.P. and Güemes, A., Wiley InterScience (Online Service), Structural Health Monitoring. ISTE, London; Newport Beach, 2006.
- [7] Frangopol, D.M. (2003) New Directions and Research Needs in Life-Cycle Performance and Cost of Civil Infrastructures. Structural Health Monitoring 2003, From Diagnostics & Prognostics to Structural Health Management, *Proceedings of the 4th International Workshop on Structural Health Monitoring*, Stanford, CA, Lancaster, PA, DEStech Publications, Inc, 53-63.
- [8] Chang, P., Flatau, A. and Liu, S. (2003) Review Paper: Health Monitoring of Civil Infrastructure. *Structural Health Monitoring*, **2**, 257-267. <https://doi.org/10.1177/1475921703036169>
- [9] Gres, S., Dalgaard Ulriksen, M., Döhler, M., Johansen, R.J. and Nielsen, S.A. (2017) Statistical Methods for Damage Detection Applied to Civil Structures. *Procedia Engineering*, **199**, 1919-1924. <https://doi.org/10.1016/j.proeng.2017.09.280>
- [10] Dang, X. (2015) Statistic Strategy of Damage Detection for Composite Structure Using the Correlation Function Amplitude Vector. *Procedia Engineering*, **99**, 1395-1406. <https://doi.org/10.1016/j.proeng.2014.12.675>
- [11] Shah, A.A., Alsayed, S.H., Abbas, H. and Al-Salloum, Y.A. (2012) Predicting Residual Strength of Non-Linear Ultrasonically Evaluated Damaged Concrete Using Artificial Neural Network. *Construction and Building Materials*, **29**, 42-50. <https://doi.org/10.1016/j.conbuildmat.2011.10.038>

- [12] Maali, Y. and Al-Jumaily, A. (2013) Self-Advising Support Vector Machine. *Knowledge-Based Systems*, **52**, 214-222.
<https://doi.org/10.1016/j.knosys.2013.08.009>
- [13] Blackledgey, J.M. and Dubovitskiyz, D.A. (2009) Texture Classification Using Fractal Geometry for the Diagnosis of Skin Cancers. EG UK Theory and Practice of Computer Graphics, 1-8.
- [14] Brieman, L. (1996) Bagging Predictors. *Machine Learning*, **24**, 123-140.
<https://doi.org/10.1007/BF00058655>