ORIGINAL ARTICLE

FFEMS | Fatigue & Fracture of Engineering Materials & Structures   WILEY

Feature selection of surface topography parameters for fatigue-damage detection using Pearson correlation method and neural network analysis

Hassan Alqahtani[1] [ORCID] | Asok Ray[2,3]

[1]Department of Mechanical Engineering, Taibah University, Medina, Saudi Arabia

[2]Department of Mechanical Engineering, Pennsylvania State University, University Park, Pennsylvania, USA

[3]Department of Mathematics, Pennsylvania State University, University Park, Pennsylvania, USA

Correspondence
Hassan Alqahtani, Department of Mechanical Engineering, Taibah University, Medina, KSA 42353, Saudi Arabia.
Email: hhq1408@gmail.com

Abstract

The global objective of this study was to investigate the best features of the surface topography for fatigue-damage detection and classification. The presence of the stress concentration in valleys of the surface topography causes a grain slip and a crack initiation at the surface of the machined structure and finally leads to fatigue failures. Therefore, the surface topography has a major influence on the fatigue strength of the machined structure. An optical confocal measurement system (Alicona) was applied to measure six surface topography parameters. In this paper, feature selection using the Pearson correlation method was adopted to select the best surface textures that provide best the neural network (NN) model performance. The NN model is capable of detecting and classifying the damage with an accuracy of up to $\sim 94.4\%$.

KEYWORDS
artificial neural network (ANN), critical surface roughening, damage fatigue, surface crack

1 | INTRODUCTION

Fatigue damage is the main source of the deterioration of a component or material, and it is caused by a large number of applied stress cycles, maximum tensile stress, and fluctuations in applied stress. The fatigue life of materials is determined by two main periods, fatigue crack initiation and fatigue crack propagation. The fracture mechanics approach has been applied to study the growth of the fatigue crack propagation which includes both long and short cracks. However, studying the fatigue crack initiation is vital because without crack initiation, there is no crack propagation. Therefore, studying fatigue crack initiation has received considerable critical attention from scientists and engineers. The mechanism of the fatigue crack initiation in heterogeneous materials is defined by two methods: (i) The crack initiates from pre-existing crack (e.g., defect) or bonded second phase particles (e.g., inclusions), and (ii) the crack initiates from natural elementary defects such as twin and grain boundaries, point defects, stacking faults, and dislocations.[1–4]

Practically, the fatigue-damage risk increases as the fatigue damage moves from one regime to another, where the fatigue-damage risk in the crack-initiation regime is very low because the crack growth rate is not significant (e.g., ~ 1 nm per cycle). According to Paris' law, the log of crack growth rate is linear with respect to the log of stress intensity factor range; hence, the fatigue-damage risk increases as the crack growth rate is increased. While the physics of fatigue crack propagation is modeled fairly accurately by the Paris' law and its modifications, the modeling of the physics of fatigue crack initiation is apparently an open issue in the current state of the art of fracture mechanics, because the mechanics-based models of crack initiation are grossly inaccurate and the grain-based models do not have well-formulated mathematical structures. Therefore, it is difficult to accurately determine the number of cycles for the crack initiation. In

general, the fatigue life of a component is influenced by some external factors such as temperature and surface finish.[1,2,5,6]

It is well known that the last stage of the manufacturing process is the surface finishing, and it determines the surface topography of a machined part. The surface finishing is known as a fingerprint of the manufacturing process because it plays a major role in lubricant retention, sealing, wear resistance, friction, and fatigue. In fact, the strength design and life prediction of a machined part are significantly affected by the surface finishing, because the local microscopic stress and strain concentration at the surface defects are major factors for crack initiation and propagation, also generally that fatigue cracks initiate from the free surface. Therefore, the surface topography is considered a major factor in the fatigue behavior of the product.[7–10]

The surface roughness provides a global view of the surface topography, and it is generally accounted for in many fatigue life models. A great deal of previous research into fatigue life has focused on the effect of surface roughness on fatigue strength. Most of these studies conclude that the fatigue strength decreases as the surface roughness increases. A qualitative study by Taylor and Clancy[11] defined the consequences of machining methods on the fatigue strengths of four distinct machined parts. The results show that the fatigue strength of the machined part decreases with increasing surface roughness. Another investigation by Itoga et al[12] provides a similar conclusion to the study of Taylor and Clancy, which implies that the fatigue life is decreased when the surface roughness of the component increased. The surface roughness is extremely affected by several factors such as temperature and stress level; (e.g., low temperature and high stress) that promote the influence of the surface roughness on the fatigue life. Zuluaga-Ramirez et al[13] used a confocal microscope to measure the roughness of carbon fiber-reinforced polymers (CFRP); their study points out that the magnitude of the surface roughness is significantly affected by the fatigue loads. Generally, surface roughness acts as a stress concentrator and creates microcracks at the free surface.[13–15]

The critical problem of fatigue damage is that fatigue damage typically occurs suddenly, without prior notice. Therefore, most petroleum refineries apply periodic inspections to certify the quality of their assets. The common technique that is used in inspection is visual inspection (VT). According to the Mainblades inspection company, more than 75% of all inspections on large transport aircraft are done by VT, and it is widely used because it is easy to train, easy to prepare, and inexpensive. On the other hand, VT can identify large flaws, and it is possible to have misinterpretation flaws. To detect tiny surface cracks or internal defects, other nondestructive testing (NDT) techniques such as ultrasonic and eddy current are applied.[16–20]

The fatigue-damage risk assessment is a fundamental task in these refineries to overcome personal, environmental, and refinery plants hazard situations. However, in most cases, it is a time-consuming task, and it depends on the inspector's experience. Therefore, building an automated crack detection system is deemed very necessary. Recent years have witnessed a growing academic interest in using machine learning techniques to detect fatigue damage. The machine learning approaches have shown the ability to provide models for fatigue-damage detection because of their excellent nonlinear approximation and multivariable learning ability.[21,22]

Machine learning is a set of algorithms that are applied to data-driven systems such as artificial neural network (ANN), fuzzy logic, convolution neural network (CNN), and support vector machine (SVM).[23,24] Machine learning techniques have been applied in fatigue fields for different purposes because of their capability of modeling the internal connections and tendencies from complicated or imprecise data. For example, Mohanty et al[25] applied an ML method called the genetic program for fatigue life prediction of the 2024-T3 aluminum alloy; and Wang et al[26] successfully utilized the relevance vector machine (RVM) to estimate the remaining useful life of a structure.

Several feature extraction methods were discussed by D'Angelo and Rampone[27] to classify different defects by an ANN-based method of detecting fatigue cracks propagating in the load-carrying structure of aircraft. Dworakowski et al[28] utilized an ANN for detecting fatigue cracks propagating in aircraft structure. Hu and Deng[29] successfully applied the nearest-neighbor SVM and CNN to detect the fatigue damage in bearings. Alqahtani and Ray[30] applied a CNN technique for fatigue-damage detection and classification of a polycrystalline alloy (Al7075-T6), where these CNN models show good applicability for classifying the fatigue damage. However, most of these studies have not used the surface topography measurements for fatigue-damage detection and classification.

This paper investigates the evolving fatigue damage in machinery structures using surface topography measurements. The surface topography illustrates generally the deviations of a real surface from its ideal form, and the crack initiation has a significant effect on the deviations of a real surface. Hence, the surface topography measurements were used to detect fatigue damage. All surface topography measurements were taken by an optical meteorology device (Alicona). The investigated surface textures are the arithmetical mean height S_a, the

root-mean-square height S_q, the maximum peak height S_p, the maximum valley depth S_v, the maximum height S_z, and 10-point height S_{z10}. In addition, this paper shows the most significant surface texture that is affected by the fatigue damage, which can be determined by obtaining the best ANN model performance accuracy.

The main contributions of this investigation are delineated as follows:

1. *Data classification*: The surface topography measurements have been used to classify the state of the fatigue damage into two classes, a healthy class and an unhealthy class. The healthy class presents all measurements before the crack initiation, while the unhealthy class presents all measurements after the crack initiation.
2. *Best surface-index selection by ANN-based pattern classification*: The performance of the ANN model has been used to select the best surface index, individually, where the model performance refers to the best surface index and vise versa.
3. *Development of the performance of the ANN model*: A correlation-based method (CBM) has been applied to select the best features that improve the ANN model performance of the best surface index.

The overall structure of this paper takes the form of five sections including the present section. The second section describes the experimental apparatus utilized in fatigue-damage detection. The third section is concerned with the methodology employed for this study, which will include an overview of (i) tested surface textures, (ii) a CBM, and (iii) ANNs. The fourth section discusses the feature selection using CBM and the performance of each surface-index ANN model. The fifth section summarizes the main points of the paper with conclusions and recommendations for future research.

2 | DESCRIPTION OF THE EXPERIMENTAL APPARATUS

This section includes the investigated material and the experimental apparatus, as shown in Figure 1. The main goal of this experiment is to obtain the surface topography parameters for the evaluation of the damaged state on the free surface. Ten typical experiments have been conducted to build an automated system for fatigue-damage detection based on surface texture data.

2.1 | Materials

The material selected to evaluate the methods of this paper is a polycrystalline alloy (made of Al7075-T6 alloy). Figure 2 shows a CAD drawing of a notched test specimen. The dimensions of the tested specimens are 3 mm thickness, 50 mm width, and (1 mm * 3.5 mm) slot cut at the edge.

2.2 | The computer-controlled fatigue testing machine

A hydraulic test machine was used for the fatigue tests of this investigation, and all conducted experiments were under load-controlled conditions. The tension–tension load cycles at 60 Hz was selected to study the influence of the surface topography parameters on surface cracks. The maximum tensile load is 11,000 N, and the minimum tensile load is 6000 N. All tested experiments were investigated in the laboratory environment (25°C and 1 atm).

FIGURE 1 The experimental apparatus [Colour figure can be viewed at wileyonlinelibrary.com]
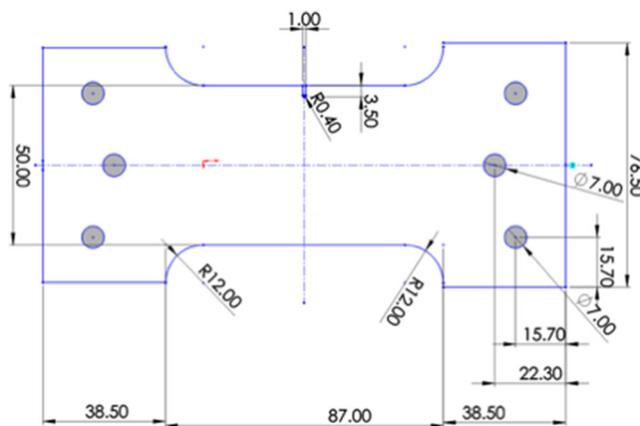
FIGURE 2 Specimen geometry used for fatigue testing [Colour figure can be viewed at wileyonlinelibrary.com]

2.3 | The optical metrology

3D surface measurements using an optical metrology device (Infinite-Focus, Alicona) illustrate the surface topography of a tested surface. Figures 3 and 4 show Alicona 3D surface topography of the notch surface of the tested specimen for both the nondamaged state and damaged state, respectively.[31,32] In this study, the measured surface topography parameters are the arithmetical mean height (S_a), root-mean-square height (S_q), maximum peak height (S_p), maximum valley depth (S_v), maximum height (S_z), and 10-point height (S_{z10}).

In the Focus-Variation system of Alicona, the topographical and color information are created from the variation of focus where the small depth of focus of an optical system is combined with vertical scanning. The vertical resolution of the Infinite-Focus system reaches 20 nm.[31,32]

3 | METHODOLOGY OF DAMAGE ANALYSIS

This section discusses the specific approaches by which the research and analyses were conducted, and it is subdivided into three subsections:

FIGURE 3 Typical 3D topography of undamaged specimens [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 4 Typical 3D surface topography of damaged specimen [Colour figure can be viewed at wileyonlinelibrary.com]

3.1 | Surface topography

The size of the investigated surface using Alicona is $0.4\text{mm} \times 0.4\text{mm}$ which is a very small area, as shown in Figure 5. However, the number of pixels in this tiny area is huge (4,161,600 pixels). Thus, the ability of Alicona to detect microcracks is significant. Amplitude parameters of the surface topography provide information for the areal height deviation of the surface topography. Six parameters in the amplitude group were selected to study the fatigue damage.

1. *Arithmetical mean height, S_a*: it is the arithmetic mean of the absolute value of the height within a sampling area. The following equation shows how to calculate S_a:

$$S_a = \frac{1}{A} \int_A |z(x,y)| dxdy, \tag{1}$$

where A is the sampling area, xy. The scalar-valued $z(x,y)$ is a measured point in the sampling area.

2. *Root-mean-square height, S_q*: it is the standard deviation value of the height distribution of the surface departures ,$z(x,y)$, within the sampling area, and it is defined by the following equation:

$$S_q = \sqrt{\frac{1}{A} \int_A |z^2(x,y)| dxdy}. \tag{2}$$

3. *Maximum peak height, S_p*: it is the largest peak height value from the mean plane within the sampling area.
4. *Maximum valley depth, S_v*: it is the largest pit value from the mean plane within the sampling area.
5. *Maximum height, S_z*: it is the sum of the largest peak height value and largest pit value within the sampling area, thus
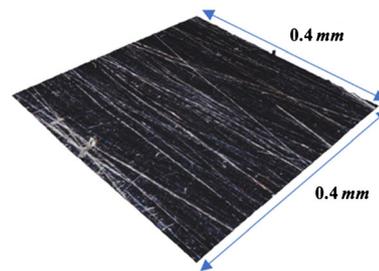
FIGURE 5 3D surface generated using Infinite-Focus deceive [Colour figure can be viewed at wileyonlinelibrary.com]

$$S_z = S_p + S_v. \tag{3}$$

6. *Ten-point height*, S_{z10}: it is the average height of the five highest maximums plus the average height of the five lowest minimums within the sampling area. The following equation expresses S_{z10}: where z_{p_i} and z_{v_i} are the height of the ith highest maximums of the sampling area and the ith lowest minimums of the sampling area, respectively.[33,34]

3.2 | Feature selection

Feature selection has been applied in ANN applications to develop a predictive model by reducing the number of input variables. Hence, minimizing the computational cost of molding and improving the performance of the model are achieved by applying feature selection methods.[35,36] This paper presents one of the feature selection methods called the Pearson correlation method. The correlation is used to determine the association between the features, where the correlation is (± 1) when two features are linearly dependent, while it is 0 when the features are uncorrelated. The correlation is defined in the following equation:

$$r_{(a,\delta)} = \frac{\sum_{i=1}^{n}[(a_i - \bar{a})(b_i - \bar{b})]}{\sqrt{\sum_{i=1}^{n}(a_i - \bar{a})^2 \sum_{i=1}^{n}(b_i - \bar{b})^2}}, \tag{4}$$

where $\bar{a}$ and $\bar{b}$ are the average value of feature a and the average value of feature b, respectively.

$$\bar{a} = \frac{1}{n}\sum_{i=1}^{n}(a_i) \quad \bar{b} = \frac{1}{n}\sum_{i=1}^{n}(b_i). \tag{5}$$

The principle concept of using the correlation method in feature selection is that features with high correlation (close to $+1$) are strongly correlated and hence have almost the same influence on the dependent variable. Therefore, we can eliminate one of the two features when two features are correlated.[37,38]

3.3 | The ANN

The ANN presents a set of nodes that are associated in a syntactic way that is similar to the association of neurons in the human brain; thus, ANN follows the system of a human brain. ANN has different forms of node connections, and each form creates a new type of ANN. The feed-forward neural is commonly used in ANN applications. The mechanism process of the feed-forward neural starts from input nodes and ends at the output nodes; hence, the direction of the sequence in this type is a forward direction. This paper presents an ANN structure called the Shallow or Vanilla multilayer neural network (NN). The Shallow ANN architecture contains an input layer, one hidden layer, and an output layer, as shown in Figure 6.

The computational process of each node is shown in Figure 7. The input data (x_1, x_2, x_3) are multiplied by weights (w_1, w_2, w_3) and added by a bias b. Practically, when the input data have valuable information, a higher weight is applied and vice versa. Thus, weight has played an important role in the ANN computational process.
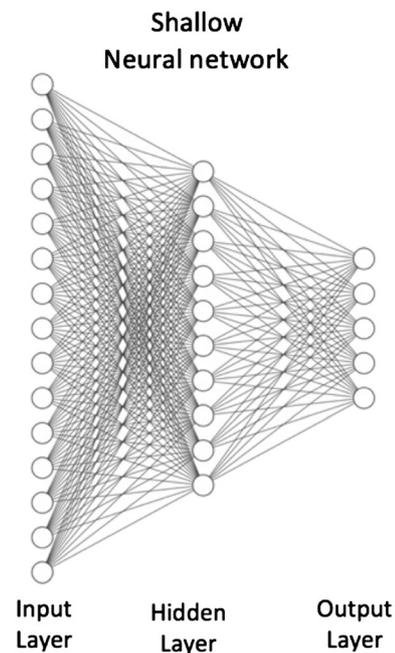
FIGURE 6 The Shallow neural network architecture
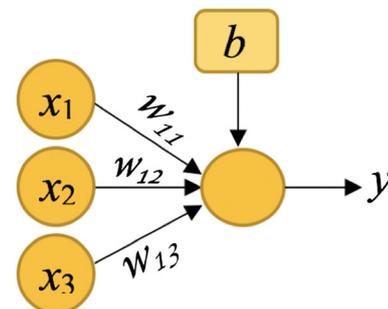
FIGURE 7 Typal structure of an artificial neuron [Colour figure can be viewed at wileyonlinelibrary.com]

The following equation expresses the output of each node:

$$\mathrm{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad w = [w_1 w_2 w_3],$$

$$a = w * x + b = w_1 x_1 + w_2 x_2 + w_3 x_3 + b, \qquad (6)$$

$$y = \phi(a) = \phi(\mathrm{w}x + b). \qquad (7)$$

Here, ϕ is the activation function. The applied activation function in this paper is known as the Sigmoid function, and it allows a NN model to learn complex problems by finding nonlinear relationships between data features. The Sigmoid function, as shown in Figure 8, generates an analogue output that is limited between 0 and 1. The Sigmoid function is defined by

$$y = \frac{1}{1 + e^a}. \qquad (8)$$

Adjusting ANN structure weights is a vital process, where these weights are adjusted based on the value of the error between the ANN model outputs and desired outputs. This technique is defined as the back-propagation technique. In this technique, a set of input data is passed through the ANN structure generating outputs. Based on the measured error, the back-propagation technique is applied in the reverse path of ANN architecture (from the output layer to the input layer (the reverse path of ANN architecture) to re-adjusted the weights. This cycle is defined as an epoch, and the number of epochs is determined by achieving the best model performance (the minimum error). Generally, the model performance improves gradually as the model error gradually decreases. The mean squared error of the ANN model is defined by
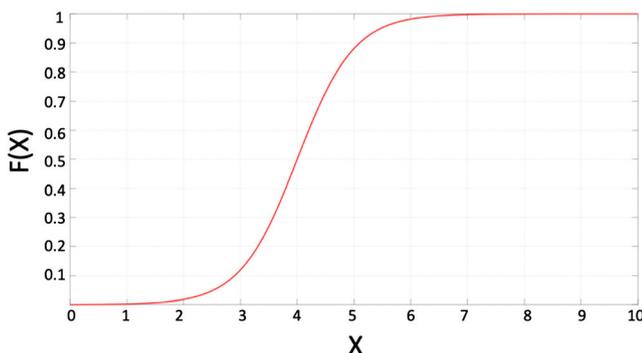
$$J_{ANN} = \frac{1}{N} \sum_{t=1}^{N} \sum_{e=1}^{O} \left(T_{ANN}(t,e) - y_{ANN}(t,e) \right)^2. \qquad (9)$$

N: Number of trained data (input, output).
e: Index denotes the number of the output.
t: Index denotes the number of trained data.
O: Number of the ANN outputs.
T: The ANN target values.

In this paper, the scaled conjugate gradient (SCG) back-propagation algorithm was applied to re-adjust weights and biases of the ANN by using the Deep Learning ToolboxTM of MATLABTM 2020a.

Due to the unavailability of a reliable crack-initiation model for developing a physic-informed ML algorithm, the authors applied the pattern-recognition method to construct a data-driven NN model. In pattern recognition, the set of experimental data has been classified into the undamaged class and the damaged class, where a pattern is a pair of variables $\{a,b\}$, where a is a vector that presents all features of the surface topography and b is the corresponding label of the feature vector. Along this line, a classification methodology is used to define the pattern allocation criteria for fatigue-damage detection. In this paper, the concept of cross-entropy (CM)[39] is used to evaluate the performance of the ANN model whose output is a probability value between zero and one, where cross-entropy loss (CEL) decreases as the predicted probability converges to its actual value. In this way, a perfect model would have a CM of zero.[40] The CEL is defined by the following equation:

$$CEL = -\frac{1}{n} \sum_{i=1}^{n} [y_i \times \log(h_\theta(x_i)) + (1 - y_i) \times \log(1 - h_\theta(x_i))], \qquad (10)$$

where
n: number of training samples.
x_i: input for training sample n.
y_i: target label for training sample n.
h_θ: model with NN weights.

The input data are randomly distributed each time a machine learning model is trained. This is done by initializing a random number generator (RNG) with an arbitrarily selected seed value before each training.

Figure 9 shows a schematic diagram of the NN architecture that is constructed with 10 Sigmoid hidden

FIGURE 8 The Sigmoid function [Colour figure can be viewed at wileyonlinelibrary.com]

neurons and two Softmax output neurons. The splitting ratio of the input data, which is measurements of the surface roughness, is 70% for training, 15% for validation, and 15% for testing.[41–43]

4 | RESULTS OF EXPERIMENTAL VALIDATION

This section presents and discusses the experimental results to identify the best surface textures, based on the performance of the ANN model. In this paper, 208 distinct areas in the image of the Alicona apparatus were investigated to build a model for fatigue-damage detection. Each measured area has six features; these features are $S_a, S_q, S_v, S_p, S_{z10}$, and S_z. Before using the ANN model, the data and features were split into two classes. The first class is defined as the undamaged surface, and it presents all data before the occurrence of the crack (93 measured

areas). The second class is called the damaged surface, which is characterized by the surface cracks in the observed images (115 measured areas). Figure 10A,B shows the measurements for undamaged surface and damaged surface classes, respectively.

Table 1 presents outcomes of the ANN model for all features individually. These outcomes were discussed in detail in following reference.[44]

It is well known that adding more features is an easy way to boost the accuracy of the ANN model. In our previous study,[44] each input data represents only one feature, and the best feature that provides the best model performance was S_z, while S_a provides the lowest ANN model performance for the fatigue-damage prediction. This is because S_a takes the average of the heights of the texture across a measured surface, and the size of the

FIGURE 9 Pattern-recognition ANN applied in this investigation [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 The results of ANN for all measured surface topography parameters.

	Cross-entropy loss (error)	ANN model performance %
S_a	0.69	68.3
S_p	0.55	71.6
S_v	0.52	68.8
S_q	0.55	69.2
S_z	0.44	87.0
S_{z10}	0.49	68.8

(A) $S_a = 431.8\ \mu m,\ S_q = 431.9\ \mu m,\ S_p = 442.6\ \mu m,\ S_v = -418.9\ \mu m,$
$S_z = 23.66\ \mu m,\ S_{10z} = 863.8\ \mu m$

FIGURE 10 Examples of analyzed surfaces with investigated features, six surface topography parameters correspondingly, as follows: (A) undamaged surface and (B) damaged surface [Colour figure can be viewed at wileyonlinelibrary.com]
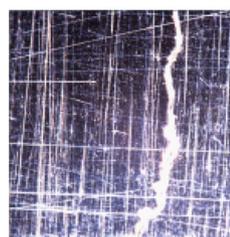
(B) $S_a = 531.7\ \mu m,\ S_q = 513.8\ \mu m,\ S_p = 526.7\ \mu m,\ S_v = -496.1\ \mu m,$
$S_z = 30.6\ \mu m,\ S_{10z} = 1024.1\ \mu m.$

microcracks according to the measured area ($0.4mm \times 0.4mm$) is almost negligible. Hence, the S_a fails to predict microcracks successfully. On the other hand, S_z accounts for the maximum valley S_v and the maximum peak height S_p, and Mínguez-Martínez et al. found that the S_z is very sensitive to surface defects and surface scratches.[45] The appearance of microcracks on the surface is similar to surface scratches. Therefore, S_z provides the best ANN model performance among other surface topography parameters. In this paper, each data represents six features to improve the ANN model performance, where the performance of all features will be compared with S_z model performance, This model was defined as *all feature model (AFM)*. In addition, the feature selection using the Pearson correlation method has been utilized to improve the performance of AFM, and this model is defined as best feature model (BFM). Table 2 presents the correlation matrix for all features.

Figure 11 shows the confusion matrix of the *AFM* model. The confusion matrix is a table that is applied to illustrate the performance of a classifier model. Columns of the confusion matrix refer to the target class (true class), while rows of the confusion matrix refer to the output of the classifier (the predicted class). The diagonal cells of the confusion matrix show samples that are correctly classified, while the off-diagonal cells of the confusion matrix show samples that are incorrectly classified. The bottom of the confusion matrix is known as the recall of the classifier, where it shows the percentages of all the samples belonging to each class that is classified correctly and incorrectly. The column on the far right of the confusion matrix is known as the precision of the classifier, and it illustrates the percentages of all samples predicted to belong to each class that is classified correctly and incorrectly. The cell in the bottom right of the confusion matrix presents the overall accuracy of the classifier within a specified interval of statistical confidence,[46] which is chosen as 95% in this paper.

It follows from Table 2 that S_z and S_p are strongly correlated (93%); hence, S_p is removed because S_z is significantly affected by crack detection, where the ANN performance of S_z is 87.0%, while the ANN performance of S_p is 71.9%. Also ,S_v and S_q are strongly correlated (71%); but the ANN model performance for both features is almost close. In the present case, S_v is chosen over S_q because S_v shows the maximum valley depth of the investigated surface and when a crack propagates on the free surface of the material; it is seen that S_v is significantly affected by crack propagation. The best features (i.e., surface topography parameters) are S_a, S_v, S_{z10}, and S_z. Figure 11A,B shows the classifier accuracy of all features (i.e., both AFM and BFM) model at performance accuracy of 90.9% and 94.2%, respectively.

It follows from Figure 11A that the AFM model accuracy is 90.9%, which is better than that of the S_z model by 3.9%. Figure 11B shows the BFM accuracy, which is 94.2%. The BFM accuracy is better than both the AFM accuracy by 3.1%, and that of the S_z model by 7.2%. The improvement of the model accuracy as compared to that
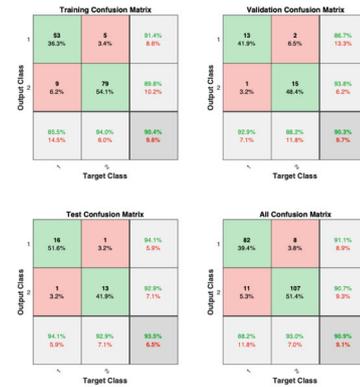
(A) Confusion matrices of all feature model (AFM) : ~ 90.9% performance accuracy.
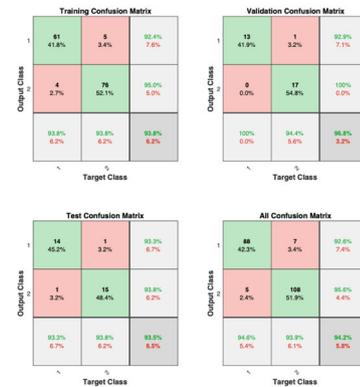
(B) Confusion matrices of best feature model (BFM): ~ 94.2% performance accuracy .

FIGURE 11 The model accuracy of AFM and BFM [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 The correlation matrix for all measured surface topography parameters.

	S_a	S_q	S_p	S_v	S_{z10}	S_z
S_a	1.00	0.00	0.54	0.70	0.00	0.00
S_q	0.00	1.00	0.55	0.71	0.00	0.00
S_p	0.54	0.55	1.00	0.00	0.01	0.93
S_v	0.70	0.71	0.00	1.00	0.01	0.61
S_{z10}	0.00	0.00	0.01	0.01	1.00	0.07
S_z	0.00	0.00	0.93	0.61	0.07	1.00
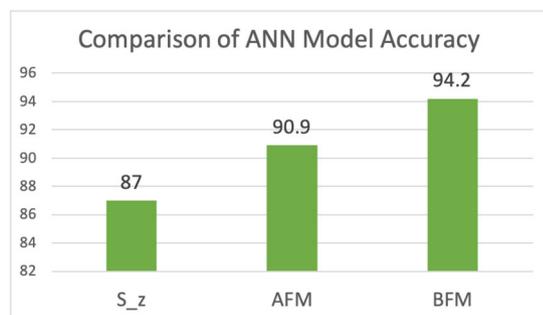
(A) Accuracy comparison of AFM, BFM, and S_z models.
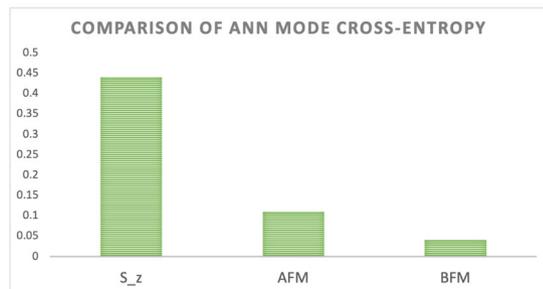
(B) Cross-entropy comparison of AFM, BFM, and S_z models.

FIGURE 12 Performance comparison of AFM and BFM [Colour figure can be viewed at wileyonlinelibrary.com]

of the S_z model is 5.5% for the AFM, and 8.3% for the BFM, as shown in Figure 12A.

Figure 12B shows the comparison of CEL among AFM, BFM, and S_z models. As compared to the CEL of the S_z model, that of the AFM is reduced by 75%, and it is significantly reduced by almost 91% for the BFM.

Generally, accuracy is very useful to validate the performance of a model when the samples are uniformly distributed (50% undamaged samples and 50% damaged samples); but in the present case, the distribution of the samples is slightly imbalanced (e.g., 44% undamaged samples and 56% damaged samples). In case of imbalanced data, the receiver operating characteristics (ROC) is used, because the ROC describes the relationship of the false positive rate of detection (specificity) on the x-axis with the true positive rate of detection (sensitivity) on the y-axis across different cut-off thresholds.[47] To confirm the outputs of the model accuracy of AFM and BFM, here, the ROC is applied to evaluate the performance of both AFM and BFM for damage detection and classification.

Figure 13A,B shows ROC curves for the AFM and BFM, respectively; the colored lines in each figure show the respective ROC curves. The optimal model (i.e., 100% sensitivity and 0% specificity) is obtained in the upper-left hand corner. As seen in Figure 13A,B, the ROC performance of BFM surpasses that of AFM, because the ROC curves of BFM are more close to reaching the upper-left-hand corner than those of AFM.
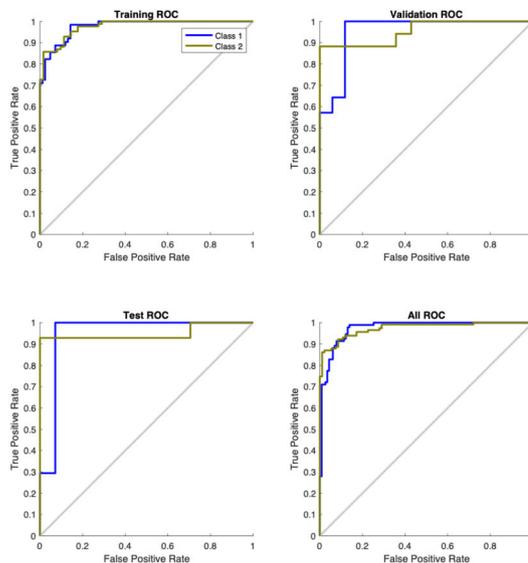
(A) ROC plots of the AFM NN model.
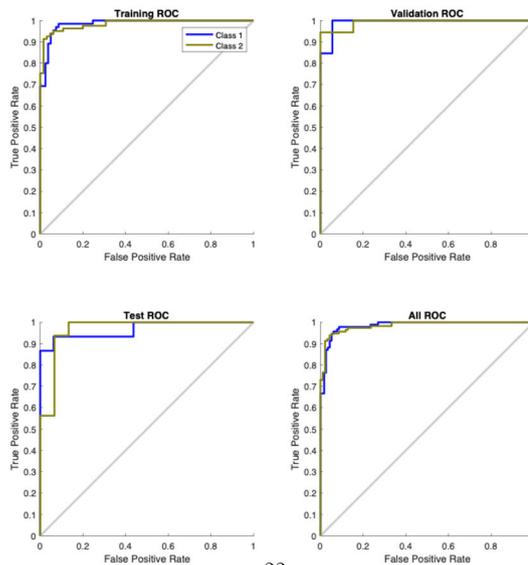
(B) ROC plots of the BFM NN model.

FIGURE 13 Receiver operating characteristics (ROC) for all experiments [Colour figure can be viewed at wileyonlinelibrary.com]

5 | SUMMARY, CONCLUSIONS, AND FUTURE WORK

This paper has developed an automated ANN model for fatigue-damage detection using the experimental data of surface topography, where the measurements were obtained from a computer-instrumented and computer-controlled fatigue testing apparatus, equipped with a confocal microscope (Alicona). Alicona has been used to measure the following surface textures: the arithmetical

mean height S_a, the root-mean-square height S_q, the maximum peak height S_p, the maximum valley depth S_v, the 10-point height S_{z10}, and the maximum height S_z.

The following three main conclusions can be drawn from this work:

- The surface textures have been tested to predict the status of fatigue damage using ANN-based models. The single feature S_z yields the best performance where the classification accuracy is 87%.
- The feature selection using the Pearson correlation method has been utilized to select the best features that provide the best model performance. The ANN outputs of the BFM were compared with those of both the AFM and the S_z model. The classification accuracy of the BFM reaches 94.2%, and that of the AFM reaches 90.9%. In essence, BFM improves the performance of fatigue-damage detection by 7.2%.
- The findings reported here shed new light on building an automated system for fatigue-damage prediction systems using the surface topography parameters data and ANN models.

Further research is needed to obtain a deeper understanding of the relationships between the surface topography parameters and fatigue damage in mechanical structures, such that the ANN models can be gainfully applied to real-life problems. In this context, the following topics are suggested for future research:

- *Investigation on different types of materials*: Similar studies should be conducted on other materials, supported by a comparison of the results with those of the proposed ANN models.
- *Usage of alternative surface topography tools for fatigue-damage assessment*: The proposed methodology could be tested in various industrial applications to detect fatigue damage using different tools of surface topography measurements, tool that is widely used in industries.
- *fatigue-damage risk assessment*: The degree of severity of the fatigue damage could be assessed by using surface topography measurements; a feasible means to achieve this goal is determination of surface topography that could provide clear indications for the risk assessment.

ORCID
Hassan Alqahtani https://orcid.org/0000-0002-5164-3335

REFERENCES

1. Anderson TL, Anderson TL. *Fracture Mechanics: Fundamentals and Applications*: CRC Press; 2005.
2. Campbell FC. *Elements of Metallurgy and Engineering Alloys*: ASM International; 2008.
3. Chan KS. Roles of microstructure in fatigue crack initiation. *Int J Fatigue*. 2010;32(9):1428-1447.
4. Ranganathan N, Aldroe H, Lacroix F, Chalon F, Leroy R, Tougui A. Fatigue crack initiation at a notch. *Int J Fatigue*. 2011;33(3):492-499.
5. Saggar M, Nasr A, Bouraoui C. Initiation life prediction method for defective materials. *Advances in Materials, Mechanics and Manufacturing*: Springer; 2020:17-25.
6. Huang Z, Wagner D, Bathias C, Paris PC. Subsurface crack initiation and propagation mechanisms in gigacycle fatigue. *Acta Materialia*. 2010;58(18):6046-6054.
7. Huang Y, Li S, Xiao G, Chen B, He Y, Song K. Research on the fatigue failure behavior of 1Cr17Ni2 blades ground by abrasive belt with passivation treatment. *Eng Failure Anal*. 2021;129:105670.
8. Podulka P. The effect of surface topography feature size density and distribution on the results of a data processing and parameters calculation with a comparison of regular methods. *Materials*. 2021;14(15):4077.
9. Singh K, Sadeghi F, Correns M, Blass T. A microstructure based approach to model effects of surface roughness on tensile fatigue. *Int J Fatigue*. 2019;129:105229.
10. Zhao B, Song J, Xie L, Hu Z, Chen J. Surface roughness effect on fatigue strength of aluminum alloy using revised stress field intensity approach. *Sci Rep*. 2021;11(1):1-17.
11. Taylor D, Clancy OM. The fatigue performance of machined surfaces. *Fatigue Fract Eng Mater Struct*. 1991;14(2-3):329-336.
12. Itoga H, Tokaji K, Nakajima M, Ko H-N. Effect of surface roughness on step-wise S–N characteristics in high strength steel. *Int J Fatigue*. 2003;25(5):379-385.
13. Zuluaga-Ramírez P, Arconada A, Frövel M, Belenguer T, Salazar F. Optical sensing of the fatigue damage state of CFRP under realistic aeronautical load sequences. *Sensors*. 2015;15(3):5710-5721.
14. Ås SK, Skallerud B, Tveiten BW. Surface roughness characterization for fatigue life predictions using finite element analysis. *Int J Fatigue*. 2008;30(12):2200-2209.
15. El May M, Saintier N, Palin-Luc T, Devos O. Non-local high cycle fatigue strength criterion for metallic materials with corrosion defects. *Fatigue Fract Eng Mater Struct*. 2015;38(9):1017-1025.
16. Cha Y-J, Choi W, Büyüköztürk O. Deep learning-based crack damage detection using convolutional neural networks. *Comput-Aided Civil Infrastruct Eng*. 2017;32(5):361-378.

17. Drury CG, Watson J. Good practices in visual inspection. Human factors in aviation maintenance-phase nine, progress report, FAA/Human Factors in Aviation Maintenance. http://hfskyway.faa.gov; 2002.

18. Ghalyan NF, Ghalyan IF, Ray A. Modeling of microscope images for early detection of fatigue cracks in structural materials. *Int J Adv Manufact Technol*. 2019;104(9-12):3899-3913.

19. Li S, Zhao X. Convolutional neural networks-based crack detection for real concrete surface. In: Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2018, Vol. 10598. International Society for Optics and Photonics; 2018:105983V.

20. Mainblades. VISUAL INSPECTIONS aircraft visual inspection. https://mainblades.com/aircraft-visual-inspection/

21. Gope D, Gope PC, Thakur A, Yadav A. Application of artificial neural network for predicting crack growth direction in multiple cracks geometry. *Appl Soft Comput*. 2015;30:514-528.

22. Rafiq MY, Bugmann G, Easterbrook DJ. Neural network design for engineering applications. *Comput Struct*. 2001;79(17):1541-1552.

23. Hasni H, Alavi AH, Jiao P, Lajnef N. Detection of fatigue cracking in steel bridge girders: a support vector machine approach. *Arch Civil Mech Eng*. 2017;17(3):609-622.

24. Wen L, Li X, Gao L, Zhang Y. A new convolutional neural network-based data-driven fault diagnosis method. *IEEE Trans Ind Electron*. 2017;65(7):5990-5998.

25. Mohanty JR, Mahanta TK, Mohanty A, Thatoi DN. Prediction of constant amplitude fatigue crack growth life of 2024 T3 Al alloy with R-ratio effect by GP. *Appl Soft Comput*. 2015;26:428-434.

26. Wang H, Song W, Zio E, Kudreyko A, Zhang Y. Remaining useful life prediction for lithium-ion batteries using fractional brownian motion and fruit-fly optimization algorithm. *Measurement*. 2020;161:107904.

27. D'Angelo G, Rampone S. Feature extraction and soft computing methods for aerospace structure defect classification. *Measurement*. 2016;85:192-209.

28. Dworakowski Z, Dragan K, Stepinski T. Artificial neural network ensembles for fatigue damage detection in aircraft. *J Intell Mater Syst Struct*. 2017;28(7):851-861.

29. Hu J, Deng S. Data-driven fatigue damage monitoring and early warning model for bearings. *Wireless Commun Mob Comput*. 2022;2022:7611670.

30. Alqahtani H, Ray A. Fatigue damage detection and risk assessment via neural network modeling of ultrasonic signals. *Fatigue Fract Eng Mater Struct*. 2022;45(6):1587-1604.

31. Danzl R, Helmli F, Scherer S. Comparison of roughness measurements between a contact stylus instrument and an optical measurement device based on a colour focus sensor. In: Proceedings of the Nanotechnology Conference. Nanotech; 2006:284-287.

32. Danzl R, Helmli F, Scherer S. Focus variation—a robust technology for high resolution optical 3d surface metrology. *Strojniški vestnik-J Mech Eng*. 2011;57(3):245-256.

33. Leach R. *Optical Measurement of Surface Topography*, Vol. 8: Springer; 2011.

34. Leach R. *Characterisation of Areal Surface Texture*: Springer; 2013.

35. Liu H, Dougherty ER, Dy JG, et al. Evolving feature selection. *IEEE Intell Syst*. 2005;20(6):64-76.

36. Peres FAP, Fogliatto FS. Variable selection methods in multivariate statistical process control: a systematic literature review. *Comput Ind Eng*. 2018;115:603-619.

37. Gopika N, ME AMK. Correlation based feature selection algorithm for machine learning. In: 2018 3rd International Conference on Communication and Electronics Systems (ICCES). IEEE; 2018:692-695.

38. Saidi R, Bouaguel W, Essoussi N. Hybrid feature selection method based on the genetic algorithm and Pearson correlation coefficient. *Machine Learning Paradigms: Theory and Application*: Springer; 2019:3-24.

39. Cover TM, Thomas JA. *Elements of Information Theory*. 2nd ed. John Wiley & Sons; 2006.

40. Ho Y, Wookey S. The real-world-weight cross-entropy loss function: modeling the costs of mislabeling. *IEEE Access*. 2019; 8:4806-4813.

41. Coolen ACC. A beginner's guide to the mathematics of neural networks. *Concepts for Neural Networks*: Springer; 1998:13-70.

42. Islam M, Chen G, Jin S. An overview of neural network. *Am J Neural Netw Appl*. 2019;5(1):7-11.

43. Zakaria M, Al-Shebany M, Sarhan S. Artificial neural network: a brief overview. *Int J Eng Res Appl*. 2014;4:7-12.

44. Alqahtani H, Ray A. Detection of fatigue damage via neural network analysis of surface topography measurements. *Current Perspectives and New Directions in Mechanics, Modelling and Design of Structural Systems*: CRC Press; 2022:618-623.

45. Mínguez-Martínez A, Maresca P, Caja J, Oliva JV. Results of a surface roughness comparison between stylus instruments and confocal microscopes. *Materials*. 2022;15(16):5495.

46. Brunk HD. *Introduction to Mathematical Statistics*. 3rd ed.: Xerox; 1975.

47. Poor HV. *An Introduction to Signal Detection and Estimation*. 2nd ed.: Springer-Verlag; 1994.