

Article

Neural Network-Based Automated Assessment of Fatigue Damage in Mechanical Structures

Hassan Alqahtani ^{1,2} and Asok Ray ^{1,3,*}

¹ Department of Mechanical Engineering, Pennsylvania State University, University Park, PA 16802, USA; hha107@psu.edu

² Department of Mechanical Engineering, Taibah University, Medina 42353, Saudi Arabia

³ Department of Mathematics, Pennsylvania State University, University Park, PA 16802, USA

* Correspondence: axr2@psu.edu

Received: 24 November 2020; Accepted: 9 December 2020; Published: 16 December 2020



Abstract: This paper proposes a methodology for automated assessment of fatigue damage, which has been tested and validated with polycrystalline-alloy (A ℓ 7075-T6) specimens on an experimental apparatus. Based on an ensemble of time series of ultrasonic test (UT) data, the proposed procedure is found to be capable of detecting fatigue-damage (at an early stage) in mechanical structures, which is followed by online evaluation of the associated risk. The underlying concept is built upon two neural network (NN)-based models, where the first NN model identifies the feature of the UT data belonging to one of the two classes: undamaged structure and damaged structure, and the second NN model further classifies an identified damaged structure into three classes: low-risk, medium-risk, and high-risk. The input information to the second NN model is the crack tip opening displacement (CTOD), which is computed by the first NN model via linear regression from an ensemble of optical data, acquired from the experiments. Both NN models have been trained by using scaled conjugate gradient algorithms. The results show that the first NN model classifies the energy of UT signals with (up to) 98.5% accuracy, and that the accuracy of the second NN model is 94.6%.

Keywords: fatigue damage; crack tip opening displacement; detection and classification; linear regression; neural network

1. Introduction

Structural integrity of large-scale systems (e.g., aircraft, large-scale transport vehicles, and power plants) deteriorates over time due to damage in their mechanical components. According to Farrar and Wonden [1], the damage is the cumulative effect of changes that are initiated in a system and potentially degrades the system reliability; these phenomena may lead to both anticipated and unanticipated failures unless appropriate timely actions are taken. Damage in mechanical structures may evolve either at a micro-scale level from inherent local defects (e.g., voids and inclusions) in materials, or at a macro-scale level from global defects (e.g., corrosion and existing cracks). The fatigue damage, which belongs to a special class of structural damage, is caused by fluctuating stresses that can be well below the respective yield points. It is well known [2–4] that ~90% of the structural failures occur due to fatigue damage. In general, failures due to fatigue damage evolve through the following three stages.

- *Defect generation:* Defects may exist in the structural materials from which the machinery parts are manufactured, or defects may occur during the manufacturing process itself.
- *Damage evolution:* Fatigue damage in the structures of machinery components is an evolutionary process during the course of its operation.

- *Crack propagation*: As the damage in machinery components evolves, their cumulative effects may give rise to cracks that could propagate and directly affect the machinery operation, possibly leading to a complete failure and shutdown.

This paper focuses on assessment of the evolving fatigue damage in machinery structures, which plays a critical role in planning and management of machinery operation, via detection and characterization of the damage. Furthermore, the risk of health degradation of machinery components is evaluated from the assessed damage along with the estimation of a time schedule for (possibly condition-based) maintenance. Thus, damage assessment is considered to be one of the fundamental tasks that are directly related to the safety of both plant operating personnel and equipment. In general, the task of damage assessment is applicable to a variety of industrial equipment (e.g., boilers, turbines, and pumps), transport vehicles (e.g., vehicles, ships, and airplanes), and civil engineering infrastructures (e.g., buildings and bridges) [5–7].

Identification of occurrence, location, and severity of the damage in mechanical structures is considered to be one of the difficult tasks in damage assessment. For example, structural health monitoring (SHM) is used widely in civil infrastructures such as bridges to identify and evaluate the damage, evolving over time. Another example is condition monitoring (CM) that is used in rotating machinery for a similar purpose as SHM, but the impact of the environment on CM is limited as compared to that of SHM [8–11].

In many industries, the expertise of (human) inspectors is essential for damage assessment (e.g., making a decision on when a critical structure should be taken out of service for maintenance), which is expensive and error-prone. Therefore, various automated methods of fatigue-damage assessment have emerged for the dual purpose of damage detection and assessment (i.e., quantification). To this end, the work reported in this paper develops an automated method of fatigue-damage assessment to assist the maintenance team in making timely and feasible decisions.

Two models have been developed on the principle of neural networks, hereafter abbreviated as NN. These NN-based models, reported in the paper, have been developed, tested, and validated on laboratory-scale experimental apparatuses, where the most effective failure factor is the applied fatigue loads on the test specimens; and other less significant factors (e.g., environmental effects) are not addressed in this paper [12–14]. The first model is built upon NN analysis of an ensemble of ultrasonic testing (UT) signals, which is used to identify the onset of fatigue damage in the test specimens. The second NN model is developed based on optical images from a confocal microscope and a digital microscope, which measure the surface topography and the crack tip opening displacement (CTOD) to define three levels (i.e., low, medium, and high) of damage risk. The data base for CTOD is augmented as a function of the measurements of crack length from a digital microscope. In summary, the reported work focuses on real-time assessment of fatigue damage and the associated risk analysis for structural materials, made of polycrystalline alloys.

Contributions of the paper: The major contributions of this paper are delineated below.

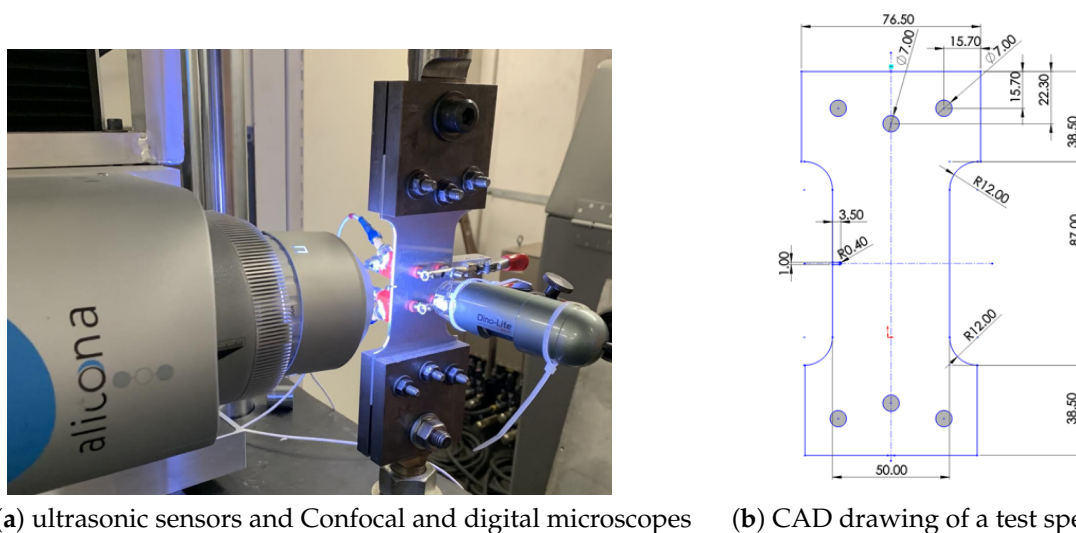
1. *Extraction of relevant features*: The extracted features are used for detection of structural damage and classification of damage risk.
2. *Development of an analytical model relating the crack length with the CTOD*: The information on the estimated CTOD data (which are derived from the measured data of crack length) is used as input data for NN models.
3. *Construction of a neural-network-based pattern classifier*: The first stage of the neural network classifies the features of ultrasonic data into the categories of undamaged structure and damaged structure in the first stage of the proposed method, while the second stage classifies the information on CTOD into low-risk, medium-risk, and high-risk.

Organization of the paper: The paper is organized into five sections. Section 2 describes the laboratory apparatus that serves as the data generator for validation of the reported method of fatigue-damage

assessment. Section 3 illustrates the methodology adopted in this paper, including an overview of linear regression and neural networks; this section consists of the following three subsections: the concept of curve fitting by linear regression; an overview of neural networks; and online damage assessment in mechanical structures. Section 4 discusses the results generated from the proposed method. Section 5 summarizes and concludes the paper along with recommendations for future research.

2. Description of the Experimental Apparatus

This section describes the experimental apparatus, as shown in Figure 1a, which is built upon a computer-instrumented and computer-controlled fatigue testing machine (Manufacturer: MTS Systems Corporation, Berlin, NJ, USA), equipped with ultrasonic testing probes (Manufacturer: Olympus, Tokyo, Japan), a confocal microscope (Manufacturer: Alicona Imaging GmbH | Dr.-Auner-Strasse 21a | 8074 Raaba/Graz, Austria), and a digital microscope (Manufacturer: QUESTAR®, New Hope, Pennsylvania, USA). The main objective here is to acquire ensembles of fatigue test data for evaluation of the damage state of the structure under consideration (e.g., test specimens in the experiments). In general, under medium-cycle to high-cycle fatigue loading of (ductile-alloy) machinery structures, a good part of the service life is consumed before reaching the crack onset stage. Therefore, the knowledge on the onset of fatigue cracks is necessary to reduce the probability of unanticipated failures as well as to maintain the machinery performance, which enhances both reliability and availability of machinery operation at a mitigated maintenance cost.



(a) ultrasonic sensors and Confocal and digital microscopes (b) CAD drawing of a test specimen.

Figure 1. The fatigue testing apparatus and ancillaries; in (b), all dimensions are in mm.

Results from 18 typical experiments have been reported to help formulating a strategy of structural health management in terms of the fatigue-damage properties of polycrystalline alloys. Figure 1b shows the CAD drawing of a notched test specimen (made of Al7075-T6 alloy), where each specimen is 3 mm thick with a 50 mm wide gauge section and the flanges with three holes on both ends are 76.5 mm wide to fix the specimens on the test apparatus by using custom-made grips. Table 1 presents the mechanical properties of Al7075-T6 alloy.

Table 1. The mechanical properties of Al7075-T6 alloy.

Mechanical Properties	The Value
Ultimate Tensile Strength (MPa)	572
Tensile Yield Strength (MPa)	503
Modulus of Elasticity (GPa)	71.7
Fatigue Strength (MPa)	159

2.1. Ultrasonic Testing

In the ultrasonic testing (UT) probes (see Figure 1), high frequency acoustic pulses (i.e., 15 MHz ultrasonic waves) are injected into each specimen by a piezo-electric transducer, called the transmitter, and are received by another piezo-electric transducer called the receiver, which is located on the other side of the transmitter. The strength of the signal is measured after it has propagated through the material. The strength of the signal at the receiver is influenced by the material features (e.g., grain boundaries, voids, and inclusions) that exist on the path of the propagated signals. While the effects of the pre-existing flaws such as voids, inclusions, or grain boundaries on the signal strength are assumed to be very gradually evolving and stable over the crack onset stage of the structure, the strength of the signal decreases dramatically once the crack propagation starts through the material because significant parts of the signals are reflected back and thus do not reach the receiver.

2.2. Optical Metrology Device

The optical metrology device (Infinite-Focus Alicona) (see Figure 1) provides 3D surface images. In the Focus-Variation system of Alicona, the topographical (colored) information is created from variations in the focus, where the small depth of the focus in an optical system is combined with vertical scanning. The vertical resolution of the Infinite-Focus system can be as low as 20 nm. The size of the generated image using Alicona is 0.4 mm by 0.4 mm, and each image has 4,161,600 pixels. Thus, the Alicona optical metrology device has the ability to detect very small cracks that are significantly less than 0.25 mm; these crack lengths are considered in this paper to belong to the crack onset regime. In the experiments, Alicona images have been taken (approximately) synchronously with ultrasonic testing (UT) data in order to provide a ground truth for the results of analysis from UT signals. Since the Alicona metrology also provides information on surface topography, they have been used to measure both the surface average roughness (S_α , the arithmetical mean height of a surface) and crack tip opening displacement (CTOD) [15,16].

2.3. Digital Microscope

Measurements from the Quester digital microscope (QDM) (see Figure 1), were taken synchronously with the measurements of average roughness (S_α) and crack tip opening displacement (CTOD) to provide the corresponding crack length α . The image resolution of the QDM is 640×480 pixels and the images are taken with 10-200X variable magnification.

3. Methodology of Damage Analysis

This section briefly introduces the methodologies, adopted in this paper for analysis of the experimental data (see Section 2), including overviews of linear regression, neural networks, and online damage assessment.

3.1. Curve Fitting by Linear Regression

Curve fitting by linear least squares is a method of identifying the model that delivers the best fit to the particular curves of the available (e.g., experimental) data set, where the error of the model is the least in some sense. The least-square method is the simplest and most widely used statistical technique for minimizing the error of the model (i.e., the summed square of residuals). The residuals (r_i) are defined as the differences between the observed responses (y_i) and estimated results, (\hat{y}_i). [17,18].

$$r_i = y_i - \hat{y}_i \quad \text{for } i = 1, 2, \dots, n. \quad (1)$$

The residual vector r is the length of the n -array of r_i values. The summed square of residuals is estimated as follows:

$$S = \sum_{i=1}^n r_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

For linear least squares, the summed square of residuals in Equation (2) is estimated as:

$$S = \sum_{i=1}^n (y_i - (a + bx_i))^2 \quad (3)$$

where the parameters a and b need to be estimated, such that S is minimized. Hence, S in Equation (3) is differentiated with respect to each parameter a and b , and the results must be identically equal zero at an extremal point.

$$\frac{\partial S}{\partial a} = -2 \sum_{i=1}^n (y_i - (a + bx_i)) = 0 \quad (4)$$

$$\frac{\partial S}{\partial b} = -2 \sum_{i=1}^n x_i (y_i - (a + bx_i)) = 0 \quad (5)$$

The parameters a and b are obtained by simultaneously solving Equations (6) and (7) as:

$$b = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad (6)$$

$$a = \frac{1}{n} (\sum y_i - b \sum x_i) \quad (7)$$

3.2. Neural Network

Neural network (NN) is a computational method that attempts to mimic the logic of a human brain. The neural network, in its simplest form, is composed of a set of nodes and a set of connections that link the neurons layerwise. Hence, an NN tends to imitate the most vital mechanism of the brain, which is the neural association. In essence, an NN works by building connections between nodes, and different types of connections create different types of NN. The feed-forward neural network is considered to be one of the most common types of NN.

3.2.1. Feed-Forward Neural Network

The sequence of processes in a feed-forward neural network (FFNN) is simple and unidirectional, where the sequence starts from input nodes and ends at the output nodes. While the complexity of FFNN may vary from a simple architecture to more complex architectures, the simplest architecture of FFNN is a single-layer neural network as seen in Figure 2a that is composed of an input layer (which is not counted as an NN layer) and an output layer. The second type of neural network architecture is a Shallow or Vanilla multi-layer neural network as seen in Figure 2b that consists of an input layer, one (or a few) hidden layer(s), and an output layer. If there are several hidden layers, the neural network architecture is called a deep neural network [19] as seen in Figure 2c. The rationale for defining the architecture in terms of hidden layers is that these layers are not accessible from outside of the neural network.

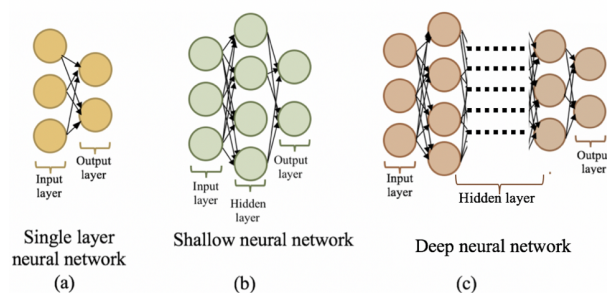


Figure 2. Three different types of neural network architecture.

Figure 3 illustrates the basic operation of an artificial neuron, where the input data (x_1, x_2, x_3) are multiplied by weights (w_{11}, w_{12}, w_{13}), respectively, and added with a bias b before leaving the node. By transforming the above result by a (nonlinear) activation function ϕ , then the output y is:

$$a = w * x + b = w_{11}x_1 + w_{12}x_2 + w_{13}x_3 + b \quad (8)$$

$$y = \phi(a) = \phi(wx + b) \quad (9)$$

One of the popular activation functions that are used in neural networks is the (smooth and nonlinear) sigmoid function that generates an analogue output that is limited between 0 and 1, as shown in Figure 4. The sigmoid function is described by:

$$y = \frac{1}{1 + e^a} \quad (10)$$

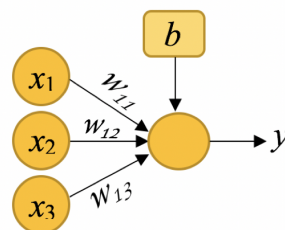


Figure 3. An artificial neuron.

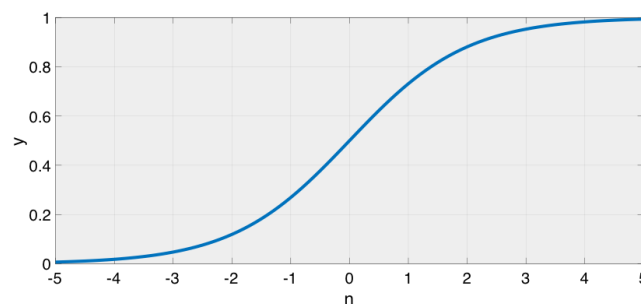


Figure 4. The sigmoid function.

Another function that is used in neural networks is the softmax function, which is usually applied at the last layer in the NN architecture. This layer is used to convert the output of the hidden layer into normalized class probabilities. The softmax function is defined as:

$$f_j(z) = \frac{e^{z_j}}{\sum_k e^{z_k}} \quad (11)$$

In essence, the softmax layer delivers the probabilities of each output class, where it takes in a vector with real values and produces a vector with elements between zero and one that sum to one.

The information in neural networks is stored in terms of weights. These weights are adjusted during the training of the neural network based on the error that is the difference between the output of the neural networks and the correct output [20–23].

3.2.2. Back Propagation

In the training/learning phase, the connection weights (w_{ij}) are adjusted to improve the performance of the NN model. From the perspectives of NN, an epoch refers to one cycle of the training/learning phase. At each epoch, a set of input data are passed through the NN architecture

producing outputs, and these outputs are compared with the target set of the output. Based on the computed error, the back-propagation algorithm is applied, where the measured error is passed in the reverse direction of the architecture, from the output layer to the input layer, to re-adjusted weights. This procedure is repeated continuously for the next epochs until the desired error is admissible. The mean squared error of the NN is expressed as:

$$J_{NN} = \frac{1}{N} \sum_{t=1}^N \sum_{e=1}^O (T_{NN}(t, e) - y_{NN}(t, e))^2 \quad (12)$$

where

N : Number of trained data (input, output).

e : Index identifying the output.

t : Index identifying training data.

O : Number of the NN outputs.

T : The NN target values.

3.2.3. Gradient Descent (GD)

The gradient descent method (also called steepest descent) is used to adjust the weights in the direction of the performance function that declines most rapidly (e.g., the most negative of the possible gradients). Equation (13) presents the adjustment of one of the network weights by using the GD algorithm:

$$w_{n+1} = w_n - 2\alpha J_n E_n \quad (13)$$

where

n : Number of iteration.

J : Jacobian matrix of J_{ANN} .

E : The computed error between ANN outputs and the target.

Although the GD algorithm is a well-known optimization method, it has the following four main disadvantages [24,25]:

- The learning rate is low.
- The direction is not perfectly scaled (i.e., convergence depends on the scale of the problem).
- The local minimal point could be missed.
- The results are sensitive to exogenous noise.

3.2.4. Conjugate Gradient

One of the methods that are established to improve the performance of the above GD algorithm belongs to the class of conjugate gradient (CG) algorithms. The CG algorithms are executed by searching in the steepest descent direction on the first iteration. Subsequently, a line search is implemented to find the optimal point to reach along the current search direction. Then, the next search direction is established such that it is conjugate to prior search directions. Generally, determining the new search direction requires a trade-off between the new steepest descent direction and the previous search direction as explained below:

$$p_0 = -g_0, \text{ where } g = J^T E \quad (14)$$

$$w_{n+1} = w_n + \alpha_n p_n \quad (15)$$

$$p_n = -g_n + \beta_n p_{n-1} \quad (16)$$

where the parameter (β_n) at each iteration is calculated to force the successive directions to be conjugate.

Different types of CG algorithms are distinguished by the method in which the parameter β_n is calculated. Most of the conjugate gradient algorithms involve a line search at each iteration. The line search technique is computationally expensive because the network response to all training data are processed various times for each search. Moller [26] established a method that overcomes the time-consuming line search. This method is known as the scaled conjugate gradient algorithm (SCG). The basic concept of this method is to combine the model-trust region approach, where the maximum distance is selected first, followed by the direction, with the CG approach [26–28].

In this paper, the SCG algorithm is used to set ANN weights, and Equations (17) and (18) present iterative computations of the parameter β_n and the direction of the new search:

$$\beta_n = \frac{|\mathbf{g}_{n+1}|^2 - \mathbf{g}_{n+1}^T \mathbf{g}_n}{\mathbf{g}_n^T \mathbf{g}_n} \quad (17)$$

$$\mathbf{p}_{n+1} = -\mathbf{g}_{n+1} + \beta_n \mathbf{p}_n \quad (18)$$

3.3. Online Damage Assessment in Mechanical Structures

This subsection proposes a neural network-based (NN)-based method for online damage assessment in mechanical structures. The proposed method consists of two cascaded NN models. The first NN model assesses the structural integrity of the system, which is comprised of two classes: the undamaged structure and the damaged structure. The second NN model evaluates the risk level of the damaged structure, which is further identified to belong to one of the following three classes: low-risk, medium-risk, and high-risk. Figure 5 presents the classification hierarchy of the proposed damage assessment method.

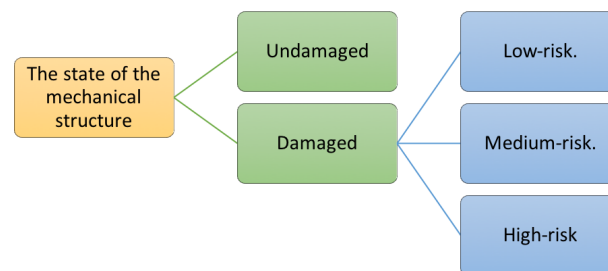


Figure 5. The classification hierarchy.

3.3.1. Feature Extraction for Machine Learning

In this paper, the task of feature extraction [29] provides pertinent information (i.e., UT signal attenuation) for detection and classification of fatigue crack damage. The rationale is based on the fact that, for the UT signal, attenuation is a consequence of partial reflection due to the damage evolution (e.g., crack onset). For example, the maximum signal strength implies the undamaged structure (e.g., the fresh test specimen), while the minimum signal strength denotes the damaged structure at the end of its service life (e.g., fully cracked specimen). In the reported work, UT data have been generated synchronously with the images obtained from the digital microscope and confocal microscope (see Figure 1) to provide ground truth for UT signal attenuation; and the signal energy is computed as:

$$E_\alpha = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (19)$$

As shown in Figure 6 beyond the vertical red dashed line, the signal energy is significantly attenuated as a consequence of crack evolution. This phenomenon is chosen to be the feature that is used for the classification for the first NN model.

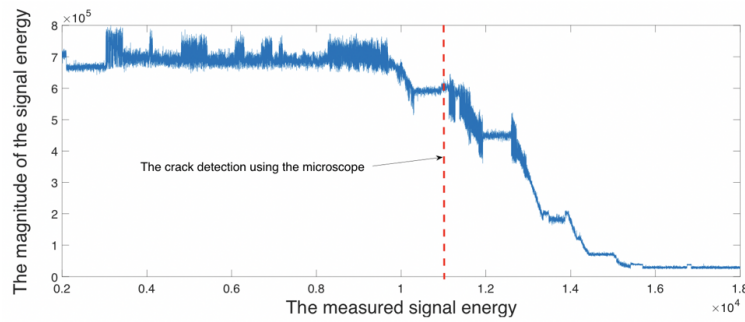


Figure 6. The UT signals energy for one experiment, where the x -axis presents 19,029 UT energy measurements during each experiment; and the y -axis shows the value of measured signal energy.

3.3.2. Classification of Damage Patterns

In this paper, the concept of pattern recognition [29,30] has been used for damage detection and classification based on experimental data. In NN, a pattern is a pair of variables $\{\phi, \lambda\}$, where ϕ is a feature vector constructed from measurement data and λ is the corresponding label of the feature vector. Along this line, in the reported work, classification procedures have been used for defining pattern allocation criteria for damage assessment. The following two NN models have been constructed for damage detection and risk identification, respectively:

1. *The first NN model for detection of structural damage:* The state of the stressed structure (e.g., a specimen) is classified based on the energy, \mathcal{E}_{UT} , of UT signals. As seen in Figure 7, before fatigue crack onset (i.e., to the left of the vertical solid line), \mathcal{E}_{UT} is labeled to belong to an undamaged class while, after the fatigue crack onset (i.e., to the right of the vertical solid line), \mathcal{E}_{UT} is labeled to belong to a damaged class.

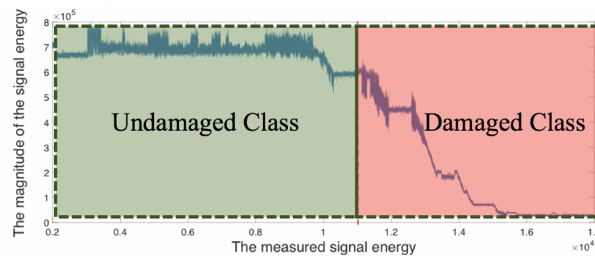


Figure 7. The The classification of the first stage, the damage initiation.

2. *The second NN model for classification of damage risk:* The risk of the structural damage (in the damaged class identified by the first NN model) is evaluated based on two criteria. The first criterion is the roughness average (S_α) of the investigated area, and the second criterion is the critical crack length α_{cr} . As seen in Figure 1, a confocal microscope is used to measure the crack tip opening displacement (CTOD) and the roughness average (S_α), while the digital microscope is used to measure the crack length α . The observations from the confocal microscope and the digital microscope are synchronized such that measurements can be taken simultaneously when the fatigue testing machine is in operation. Following these experimental observations, the damage risk is identified in the following three classes:

- (a) *High-risk damage class:* In this class, the critical crack length (CCL) method is applied to determine the damage, in which all measured data after exceeding CCL (α_{cr}) are considered to be of high risk. The critical crack length is computed as:

$$\alpha_{cr} = \frac{K_{IC}^2}{\pi \sigma_{max}^2 Y_\alpha^2} \quad (20)$$

where K_{IC} is the fracture toughness (for the Al7075-T6 alloy, $K_{IC} = 20.0 \text{ MPa}\cdot\text{m}^{0.5}$); α_{cr} is the critical crack length (CCL); σ_{max} is the maximum applied stress; and the dimensionless parameter Y_α is computed from the crack length α and specimen width w of the test specimen. For edge cracks in tension, Zahavi et al. [31] introduced the following formula to calculate Y_α .

$$Y_\alpha = 1.12 - 0.231\left(\frac{\alpha}{w}\right) + 10.55\left(\frac{\alpha}{w}\right)^2 - 21.72\left(\frac{\alpha}{w}\right)^3 + 30.39\left(\frac{\alpha}{w}\right)^4 \quad (21)$$

The CCL in all tested specimens is $\alpha_{cr} \approx 9.6 \text{ mm}$, as seen in Figure 8, where all measurements were taken after reaching CCL and are labeled to belong to the high-risk class.

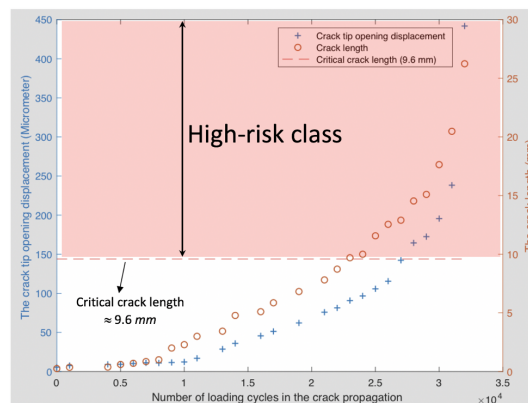


Figure 8. Determining the high-risk region using the crack length and CTOD.

- (b) *Medium-risk damage class:* As the crack length reaches a certain level, the roughness average (S_a) of the investigated area changes dramatically and the crack growth process tends to become unstable; and the effect of the crack on the surface topography becomes noticeable. This change is often called as the S_{α_alert} . In all experiments, the S_{α_alert} occurred between the crack onset and CCL. Therefore, all observations of the damage after the S_{α_alert} and before CCL are classified to belong to the medium-risk class.
- (c) *Low-risk damage class:* The damage observations before CCL are considered to be non-risky. However, this consideration might be ambiguous because the damage risk of observations just before CCL is not similar to that at the damage initiation. Therefore, identifying the damage risk of non-risky damage is very necessary. In the reported work, the roughness average S_α has been used to quantify the low-risk damage because the effects of tiny cracks on the external surface are, in fact, negligible. In addition, as shown in Figure 9, readings of S_a before the crack onset are stable, and they are also still stable after the crack onset.

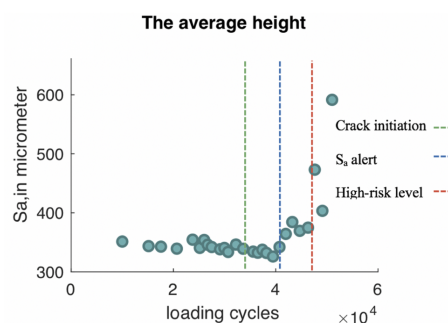


Figure 9. The average surface roughness, S_a , measurements of one experiment.

Figure 10a–d present the surface topographies of typical (non-damaged), (low-risk), (medium-risk), and (high-risk) damaged specimens, respectively.

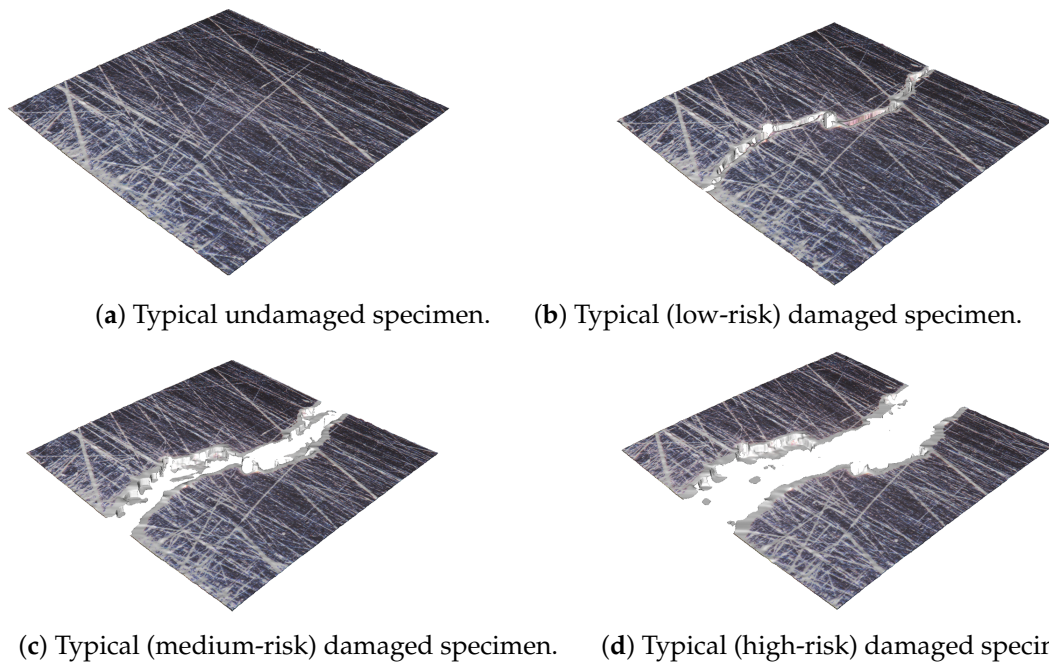


Figure 10. Illustration of fatigue-damage evolution in a typical test specimen.

3.3.3. Data Preparation and Neural Networks Procedure

This subsection provides the necessary steps for building an NN model with a shallow architecture, as delineated in Figure 11.

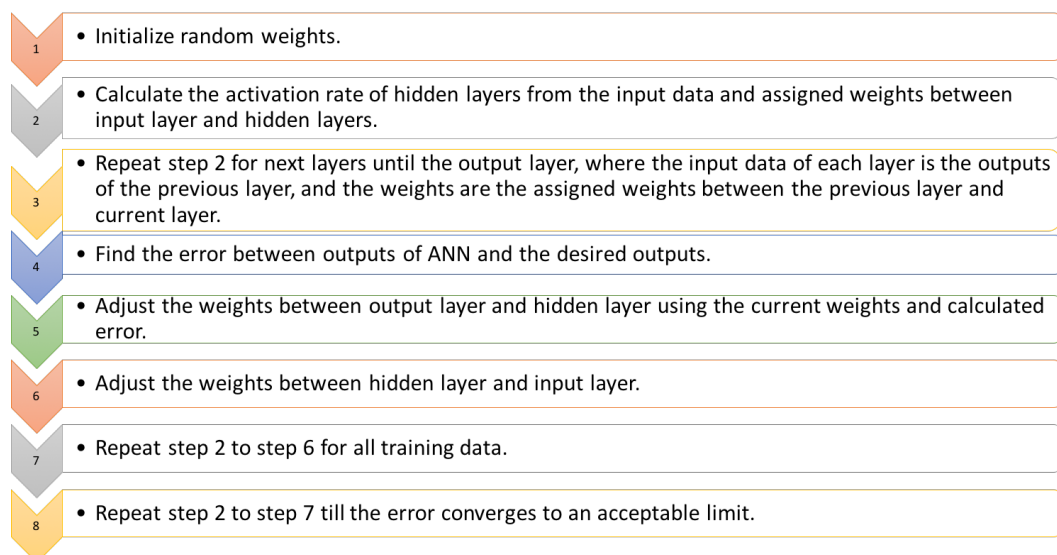


Figure 11. The neural network analysis procedure.

In the first NN model, the energy of UT signals is normalized because the response of the signal energy for all tested specimens is of similar texture, but the values of the energy may differ from one to another. The normalization is called a z-score, and it transforms UT signal energy to have a mean of zero and a standard deviation of one. The UT signal energy is normalized as:

$$z_n = \frac{x_n - \bar{x}}{\sigma} \quad (22)$$

where \bar{x} is the mean and σ is the standard deviation.

In the second NN model, a linear regression model is created relating the relationship between CTOD and CCL to fit the experimental data. The rationale for creating a linear regression model in this paper is as follows:

1. Taking CTOD measurements is experimentally expensive, especially at the high-risk level.
2. To ensure that the NN model is effective, the training data set must satisfy two conditions: (i) Every group of data must be normalized so that each group pattern is represented in the data set; (ii) statistical deviations must be effectively represented within each class, where both NN models must contain the total range of (noise-corrupted) data [32].

4. Results and Discussion

This section presents and discusses the experimental results for validation of the proposed linear regression and (neural network (NN) models. The first part of this section presents the results of linear regression, while the second part covers the results of NN models.

4.1. Linear Regression Model

As stated in the previous section, before applying the tools of neural network, the amount of the data must be considered, such that every group must contain enough data for training, validating, and testing the neural network model. A selection of the percentage of splitting the dataset into train, validation, and test set depends on the size of the dataset. A common splitting ratio, which is used in this study, is 70% of the data for a train set, 15% for validation set, and 15% for the test set.

In the case of limited data available for training, a regression model that represents typical measurement data can be used, instead of the actual measured data. As seen in Figure 12, a relationship between the crack length (shown as the x -axis) and the crack tip opening displacement (CTOD) (shown as the y -axis) is built by using a linear regression approach from the experimental data, where the dots represent the actual measured data of an experiment, and the fitted straight line represents the linear regression model.

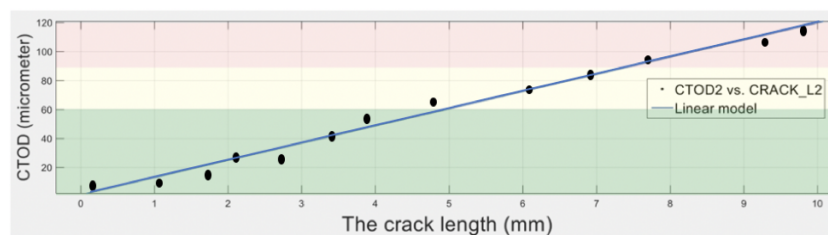


Figure 12. The linear regression model of the crack length and the crack tip opening displacement.

In Figure 12, the green area presents the low-risk damage class, while the yellow and red areas present the medium-risk and high-risk classes, respectively. Due to the limited amount of available experimental data, it is difficult to have an accurate NN model using the actual measured data. For example, since the medium-risk class has only three measured data, and if one uses the splitting ratio (70% train, 20% test, and 20% validation), either the testing set or validation set will be empty.

Therefore, this problem has been alleviated by incorporating a linear regression model that estimates the CTOD. The linear regression model of Figure 12 is constructed as:

$$CTOD = 11.88 a + 1.186 \quad (23)$$

where $CTOD$ is the crack tip opening displacement in micrometers, and a is the crack length in mm . Furthermore, the size of the input data of the neural network is now increased from 13 measured points to 986 estimated points by using the linear regression model.

4.2. Results Generated from Neural Network Models

This subsection presents the results generated from the two NN models, where the pattern recognition part is handled by the *neural network* toolbox of MATLAB, which consists of a two-layer feed-forward network, with ten sigmoid hidden neurons (two softmax output neurons for the first model and three softmax output neurons for the second model). The amount of the input data for the first NN model is 19,029 samples, and 986 samples are the inputs for the second NN model. Table 2 illustrates the splitting of data between the two NN models.

Table 2. Splitting of input data for training, validation, and testing of the NN models

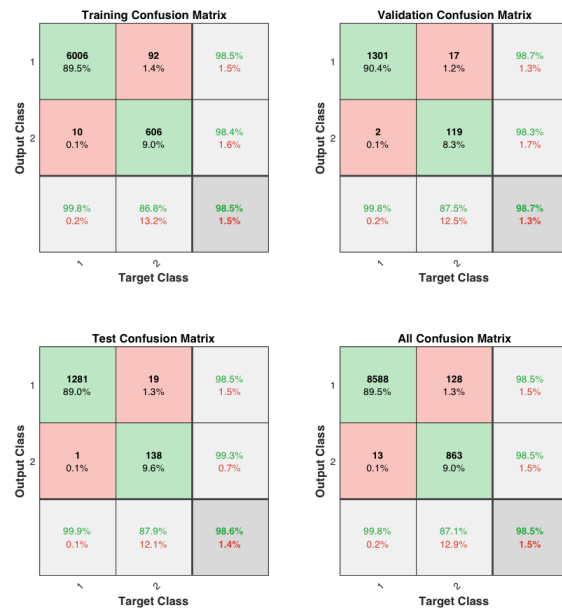
	The Input Data (Total Samples)	The Training Data (70% Samples)	The Validation Data (15% Samples)	The Testing Data (15% Samples)
First, NN model	19,029	13,321	2854	2854
Second NN model	986	690	148	148

The plates in Figure 13a,b show the confusion matrices of the first NN model and the second NN model, respectively, for training, testing, and overall validation. The outputs of the NN models appear to be very accurate; and the high percentages of correct responses are indicated in the green squares, while the low percentages of incorrect responses appear in the red squares. The lower right gray squares show the overall accuracy/error.



(a) Confusion matrices of the first NN model.

Figure 13. Cont.

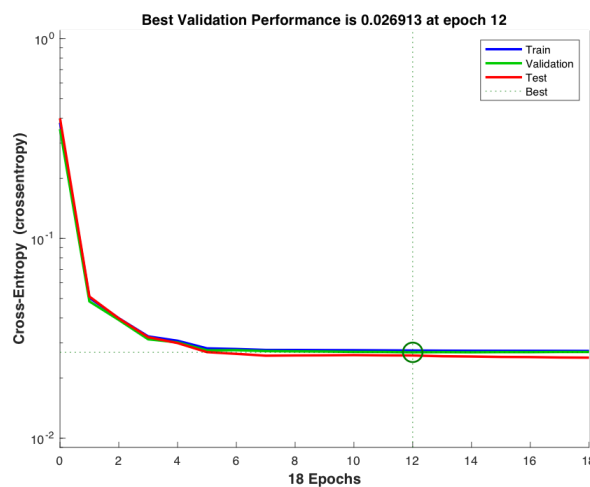


(b) Confusion matrices of the second NN model.

Figure 13. Performance of the first and second neural network models for all experiments.

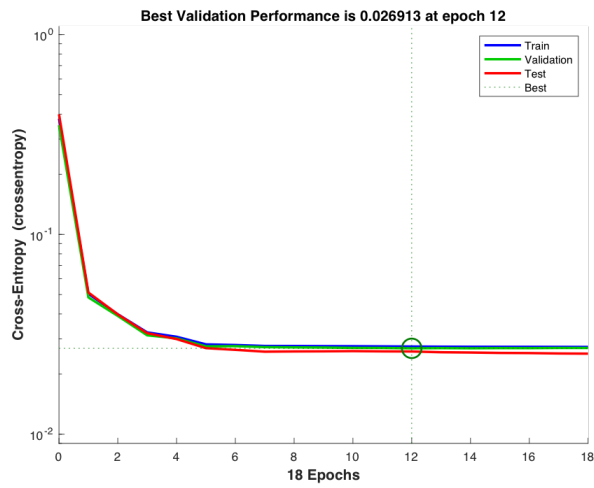
The first NN model, which makes the (binary) detection of undamaged and damaged classes, achieves an accuracy of 98.5% for the training data set, 98.7% for the validation data set, and 98.6% for the testing data set; and the first NN model achieves an overall accuracy of 98.5%. The second NN model, which makes the (trinary) classification of damaged class into high-risk, medium-risk, and low-risk, achieves an accuracy of 95% for the training data set, 94.8% for the validation data set, and 92.8% for the testing data set; and the overall accuracy of the second NN model is equal to 94.6%.

Figure 14a,b show that the best validation performance for the first NN model is achieved at the epoch 12, while that for the second NN model is achieved at the epoch 45. Figure 15a,b illustrate the respective error histograms of the first NN and the second NN models, respectively, for training, validation, and testing. It is noted that the data fitting errors are distributed within a region that is close to the zero-error point.



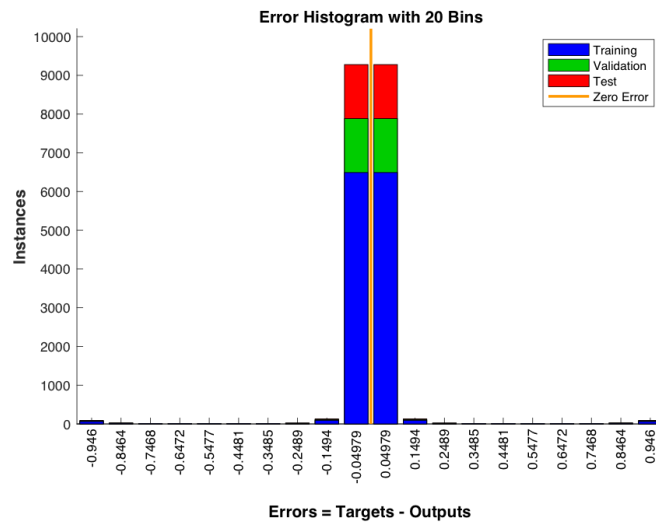
(a) Best validation performance of the first NN model.

Figure 14. Cont.

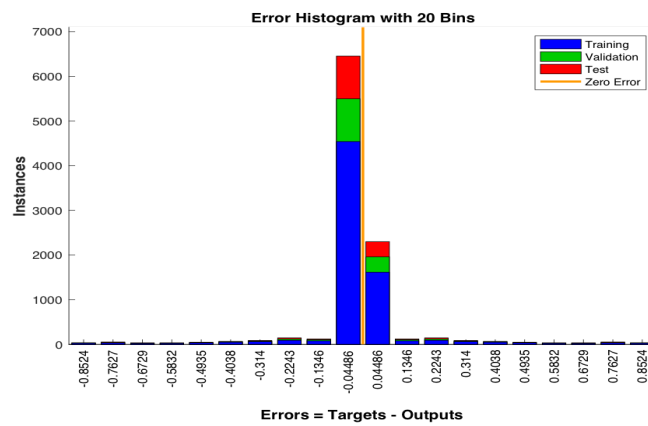


(b) Best validation performance of the second NN model.

Figure 14. Best validation performance for all experiments.



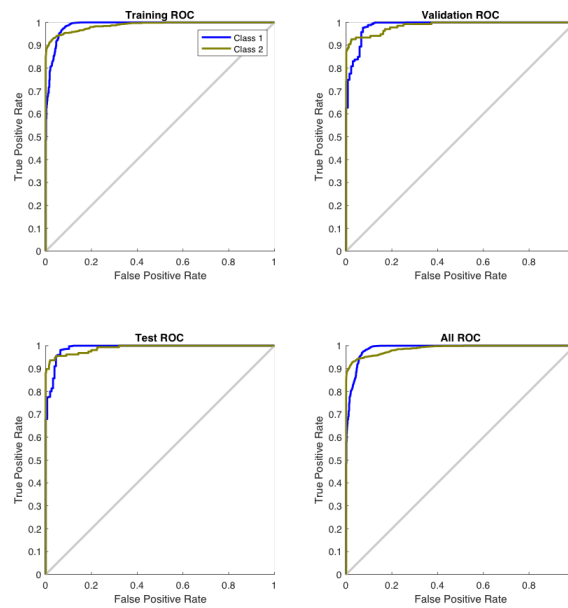
(a) Error histogram of the first NN model.



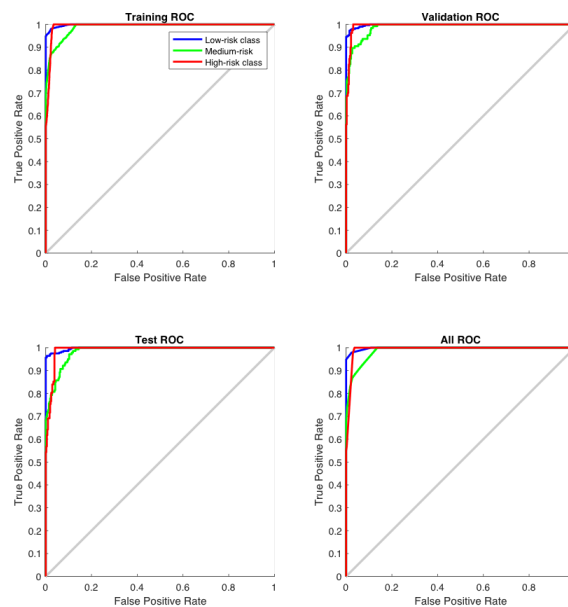
(b) Error histogram of the second NN model.

Figure 15. Error histogram for all experiments.

Figure 16a,b show ROC curves (The Receiver Operating Characteristic (ROC) curves [33] have been used in this paper to assess the performance of both NN-based models for damage detection and classification.) for the first NN model and second NN model, respectively, where the colored lines in each figure represent the ROC curves. The ROC curves represent the relationship between the true positive rate of detection (also called sensitivity) on the y -axis versus the false positive rate of detection (also called specificity) on the x -axis as the threshold is varied. The optimal model (100% sensitivity and 0% specificity) is achieved in the upper-left hand corner. As seen in Figure 16a,b, the performance is excellent for both of the NN models.



(a) ROC plots of the first NN model.



(b) ROC plots of the second NN model.

Figure 16. Receiver operating characteristics (ROC) for all experiments.

5. Summary, Conclusions, and Future Work

This paper has reported the development and experimental validation of a methodology for fatigue-damage assessment in mechanical structures of machinery. The goal is to increase the structural reliability of operating machinery and to avoid grave consequences of the damaged structure, such as endangered safety of plant operating personnel and equipment as well as environmental protection, especially in hazardous engineering applications. For example, real-time assessment of fatigue damage should be applied on petrochemical reactors, where the non-destructive testing (NDT) sensors are embedded in critical locations to provide an updated evaluation for the status of the reactor. Once the fatigue damage is monitored and assessed, human experts should make the final decision on plant operation and maintenance.

In this work, two neural network (NN)-based models have been constructed on the experimental data, acquired from a computer-instrumented and computer-controlled fatigue testing apparatus, equipped with ultrasonic test probes, a confocal microscope, and a digital microscope. It is noted that delicate instruments like optical microscopes are to be used in the laboratory environments only to calibrate the algorithms that make use of (relatively) inexpensive and rugged ultrasonic probes for field applications.

Decision making on damage assessment relies on the outputs of the afore-mentioned two NN models, where the the first model classifies the signal energy of ultrasonic test (UT) data into two classes, namely undamaged class and damaged class; and the second NN model classifies the crack tip opening displacement (CTOD) data into high-risk, medium-risk, and low-risk. Furthermore, a linear regression model has been constructed to augment the amount of the data by estimating the CTOD as inputs to the NN models.

The results of this investigation show that the classification accuracy of the first NN model reaches 98.5%, and that of the second NN model reaches 94.6%. Therefore, the proposed methodology of fatigue-damage assessment systems has the potential of successfully detecting the fatigue damage onset in mechanical structures of machinery and evaluating the associated risk of operation.

While there are many areas of theoretical and experimental research to improve the proposed classification method so that it can be gainfully applied to real-life problems, the following topics are suggested for future research:

- *Development of a detailed simulation model of the operating machinery under consideration to generate ample data for training and testing of deep neural networks:* This model must be experimentally validated at a few discrete operating points so that the interpolated (and also extrapolated) points of operation are reasonably correct representations of the real-life situations.
- *Usage of more advanced tools of neural networks (e.g., convolutional neural networks (CNN) [34] and recurrent neural networks (RNN) [35]) for fatigue-damage assessment:* Such advancements are expected to yield wider ranges of applications at the expense of requiring extensive training data.

Author Contributions: conceptualization, H.A. and A.R.; methodology, H.A. and A.R.; software, H.A.; validation, H.A. and A.R.; formal analysis, H.A. and A.R.; investigation, H.A. and A.R.; resources, A.R. and H.A.; data curation, H.A.; writ draft preparation, H.A. and A.R.; writ and editing, A.R.; visualization, H.A. and A.R.; supervision, A.R.; project administration, A.R. and H.A.; funding acquisition, A.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The first author gratefully acknowledges the financial support of the Saudi Arabian Cultural Mission (SACM). The work reported in this paper has been supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) under Grant No. FA9550-15-1-0400. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Farrar, C.R.; Worden, K. *Structural Health Monitoring: A Machine Learning Perspective*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
2. Suresh, S. *Fatigue of Materials*; Cambridge University Press: Cambridge, UK, 1998.
3. Campbell, F.C. *Elements of Metallurgy and Engineering Alloys*; ASM International: Almere, The Netherlands, 2008.
4. Lesiuk, G.; Szata, M.; Rozumek, D.; Marciniak, Z.; Correia, J.; DeJesus, A. Energy response of S355 and 41Cr4 steel during fatigue crack growth process. *J. Strain Anal. Eng. Des.* **2018**, *53*, 663–675.
5. Meruane, V.; Mahu, J. Real-time structural damage assessment using artificial neural networks and antiresonant frequencies. *Shock Vib.* **2014**, *2014*, 653279.
6. Meruane, V.; Heylen, W. An hybrid real genetic algorithm to detect structural damage using modal properties. *Mech. Syst. Signal Process.* **2011**, *25*, 1559–1573.
7. Meruane, V.; Heylen, W. Damage detection with parallel genetic algorithms and operational modes. *Struct. Health Monit.* **2010**, *9*, 481–496.
8. Dervilis, N. A Machine Learning Approach to Structural Health Monitoring with a View towards Wind Turbines. Ph.D. Thesis, University of Sheffield, Sheffield, UK, 2013.
9. Güemes, A.; Fernandez-Lopez, A.; Pozo, A.R.; Sierra-Pérez, J. Structural Health Monitoring for Advanced Composite Structures: A Review. *J. Compos. Sci.* **2020**, *4*, 13.
10. Dong, T.; Haftka, R.T.; Kim, N.H. Advantages of Condition-Based Maintenance over Scheduled Maintenance using Structural Health Monitoring System. In *Reliability and Maintenance—An Overview of Cases*; IntechOpen: London, UK, 2019.
11. Friedmann, H.; Kraemer, P. Vibration-based condition monitoring, structural health monitoring, population monitoring—Approach to a definition of the different concepts by means of practical examples from the field of wind energy. In Proceedings of the 8th European Workshop on Structural Health Monitoring (EWSHM 2016), Bilbao, Spain, 5–8 July 2016.
12. Gonzalez-Perez, C.; Valdes-Gonzalez, J. Identification of structural damage in a vehicular bridge using artificial neural networks. *Struct. Health Monit.* **2011**, *10*, 33–48.
13. Arangio, S.; Beck, J. Bayesian neural networks for bridge integrity assessment. *Struct. Control. Health Monit.* **2012**, *19*, 3–21.
14. Strączkiewicz, M.; Barszcz, T. Application of artificial neural network for damage detection in planetary gearbox of wind turbine. *Shock Vib.* **2016**, *2016*, doi:10.1155/2016/4086324.
15. Danzl, R.; Helml, F.; Scherer, S. Comparison of roughness measurements between a contact stylus instrument and an optical measurement device based on a colour focus sensor. In *Technical Proceedings of the 2006 NSTI Nanotechnology Conference and Trade Show*; CRC Press: Boca Raton, FL, USA, 2006; pp. 284–287.
16. Danzl, R.; Helml, F.; Scherer, S. Focus variation—A robust technology for high resolution optical 3D surface metrology. *Stroj.-Vestn.-J. Mech. Eng.* **2011**, *57*, 245–256.
17. Abazid, M.; Alkoud, D. *A Least-Squares Approach to Prediction The Future Sales of Pharmacy*; Blue Eyes Intelligence Engineering & Sciences Publication Pvt. Ltd.: Bhopal, India, 2018.
18. Mallick, P.; Nayak, C.; Behera, T. Least square based parabolic regression to determine calibration constant in measurement. In Proceedings of the 2010 5th International Conference on Industrial and Information Systems, Mangalore, India, 29 July–1 August 2010; pp. 334–338.
19. Wang, Z.; Yan, W.; Oates, T. Time series classification from scratch with deep neural networks: A strong baseline. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 1578–1585.
20. Coolen, A.C. A beginner's guide to the mathematics of neural networks. In *Concepts for Neural Networks*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 13–70.
21. Zakaria, M.; Al-Shebany, M.; Sarhan, S. Artificial neural network: A brief overview. *Int. J. Eng. Res. Appl.* **2014**, *4*, 7–12.
22. Islam, M.; Chen, G.; Jin, S. An Overview of Neural Network. *Am. J. Neural Netw. Appl.* **2019**, *5*, 7–11.
23. Zayegh, A.; Al Bassam, N. Neural network principles and applications. In *Digital Systems*; IntechOpen: London, UK, 2018.
24. Wang, X. Method of steepest descent and its applications. *IEEE Microw. Wirel. Compon. Lett.* **2008**, *12*, 24–26.

25. Ruder, S. An overview of gradient descent optimization algorithms. *arXiv* **2016**, arXiv:1609.04747.
26. Møller, M.F. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Netw.* **1993**, *6*, 525–533.
27. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006.
28. Dkhichi, F.; Oukarfi, B. Neural Network Training By Gradient Descent Algorithms: Application on the Solar Cell. *Int. J. Innov. Res. Sci. Eng. Technol.* **2014**, *3*, 15696–15702.
29. Ghorpade, S.; Ghorpade, J.; Mantri, S. Pattern recognition using neural networks. *Int. J. Comput. Sci. Inf. Technol. (IJCSIT)* **2010**, *2*, 92.
30. Ghotkar, V.N.; Indurkar, P.; Chilke, B. The Pattern Recognition using Spiking Neural Network and Neural Network. *Int. J.* **2015**, *5*.
31. Zahavi, E.; Torbilo, V.; Press, S. *Fatigue Design: Life Expectancy of Machine Parts*; CRC Press: Boca Raton, FL, USA, 1996.
32. Thirumalainambi, R.; Bardina, J. Training data requirement for a neural network to predict aerodynamic coefficients. In Proceedings of the Independent Component Analyses, Wavelets, and Neural Networks. International Society for Optics and Photonics, Orlando, FL, USA, 1 April 2003; Volume 5102, pp. 92–103.
33. Poor, H. *An Introduction to Signal Detection and Estimation*; Spriger: New York, NY, USA, 1994; p. 308.
34. Alqahtani, H.; Bharadwaj, S.; Ray, A. Classification of fatigue crack damage in polycrystalline alloy structures using convolutional neural networks. *Eng. Fail. Anal.* **2021**, *119*, doi:10.1016/j.engfailanal.2020.104908.
35. Husken, M.; Stagge, P. Recurrent neural networks for time series classification. *Neurocomputing* **2003**, *50*, 223–235.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).