The International Journal of Advanced Manufacturing Technology (2019) 104:3899–3913
https://doi.org/10.1007/s00170-019-04108-z

ORIGINAL ARTICLE

# Modeling of microscope images for early detection of fatigue cracks in structural materials

Najah F. Ghalyan[1] · Ibrahim F. Ghalyan[2] · Asok Ray[3]

Received: 8 May 2019 / Accepted: 7 July 2019 / Published online: 24 July 2019


**Abstract**
From the perspectives of health monitoring and life extension of structural materials, this paper addresses the problem of early detection of fatigue cracks in metallic materials (e.g., polycrystalline alloys). To this end, optical images have been collected from an ensemble of test specimens to construct computationally efficient models of crack evolution; these images are segmented into two major categories. The first category comprises images of (structurally) healthy specimens, while the second category contains images of specimens with cracks, including those in early stages of crack evolution. Based on this information, algorithms for early detection of crack formation are formulated in the setting of image classification, where the bag-of-words (BoW) technique has been used to develop models of the sensed images from a microscope, resulting in computationally efficient crack detection algorithms. To evaluate the performance of these crack detection algorithms, experiments have been conducted on a special-purpose fatigue testing apparatus, equipped with a computer-controlled and computer-instrumented confocal microscope system. The results of experimentation with multiple test specimens show excellent crack detection capabilities when the proposed BoW-based feature extraction is combined with quadratic support vector machine (QSVM) for pattern classification. Comparative evaluation with other classification tools establishes superiority of the proposed BoW/QSVM technique.

**Keywords** Bag-of-words · Crack detection · Image classification

## 1 Introduction

Fatigue crack damage is one of the most common sources of failures in mechanical structures, which could have severe consequences if not adequately addressed in a timely manner. Consequently, fatigue crack damage has attracted much attention of research and industrial institutions (e.g., [1, 20, 27, 29, 42, 43] and the references therein). Initial defects (e.g., dislocations, voids, inclusions and slip bands) exist in the microstructure of critical components even before a machinery is put into service. In general, the evolution of fatigue damage is critically dependent on these initial defects, from which cracks start to nucleate and eventually merge together, generating larger cracks that could potentially lead to a structural failure [45]. These microstructural initial defects are usually randomly distributed and produce uncertainties in the crack initiation and propagation processes even under known deterministic loading conditions. Therefore, evolution of fatigue crack in mechanical structures is treated as a stochastic process.

Early detection of fatigue crack can be critical for Pareto optimization of two competing requirements of: machinery performance and service life. An example in aircraft applications [41] is flight scheduling with life-extending control that could potentially reduce the structural damage with no significant loss of performance. Another example in electric power generation systems [26] is life-extending control of

✉ Asok Ray
axr2@psu.edu

Najah F Ghalyan
nfg103@psu.edu

Ibrahim F Ghalyan
ibrahim.ghalyan@nyu.edu

1 Department of Mechanical Engineering, The Pennsylvania State University, University Park, PA 16802, USA

2 Department of Mechanical and Aerospace Engineering, Tandon School of Engineering, New York University, Six Metrotech Center, Brooklyn, NY 11201, USA

3 Department of Mechanical Engineering and Department of Mathematics, The Pennsylvania State University, University Park, PA 16802, USA

steam turbines, for which grid disturbances and fluctuating output power make plant components undergo transient and sustained stresses, which may require frequent inspection and maintenance [8, 30]. Usually such systems have a number of critical components that characterize the effective service life. With reliable modeling of fatigue crack evolution, a life-extending control system could be designed to mitigate detrimental effects of fluctuating stresses in critical components while maintaining the performance at an acceptable level [9, 41]. Moreover, a real-time reliable estimation of the current state of fatigue crack could be used to optimize mission planning and condition-based maintenance, and facilitate identification of parts to be replaced/fixed based on their estimated damage. This could significantly reduce maintenance cost and enhance system availability.

Many of the predictive models of fatigue crack evolution, reported in literature (e.g., [7, 23, 33, 44]) strongly depend on the (unknown) initial defects in the structural materials, which give rise to uncertainties. As explained earlier, these uncertainties may lead to a significant difference in the expected remaining service lifer as predicted by the models. Moreover, randomly varying changes in the service conditions (e.g., temperature, humidity, and loading dynamics) may also significantly affect the cumulative fatigue damage and diminish the capability of model-based methods to reliably predict the remaining service life of the underlying mechanical system.

The above discussion evinces the need for development of data-driven methods that would model the fatigue damage largely based on sensor measurements and the prior knowledge of failure incidents. In this regard, various measurement devices, such as acoustic emission [19], ultrasonic [11, 20], and eddy current [52] sensors, have been widely used to detect the fatigue cracks. Machining marks and scratches, possibly due to surface finishing and cutting, are common defects from which fatigue cracks may initiate and propagate [17]. Therefore, there has been a considerable interest in detecting surface cracks using images, which can be used for probabilistic modeling of the fatigue crack damage [2, 10, 12, 15, 34, 36, 38, 39, 50]. While manual inspection of images is possible for simple cases, it typically depends on the specialist's knowledge and experience [36]. Therefore, automated inspection is generally considered to be more feasible and desirable, especially if there are many images to investigate or if online image classification is used to detect surface cracks upon their occurrence.

In the above context, there have been different techniques for automated crack detection using images. In [10, 34, 38, 50], convolution neural network (CNN) has been used to detect cracks in civil infrastructures (e.g., concrete and steel) using surface images. Although CNN typically yields excellent performance in image classification [28], one of the main issues in CNN is that it is data intensive. Without feeding the network with a large number and variety of images for training, CNN is expected to suffer from the overfitting problem [48]. Cubero-Fernandez et al. [12] reported a decision tree heuristic algorithm to classify images for 88% successful detection of cracks and 80% successful identification of the crack type while the false alarm rates were within a permissible bound. Zhong et al. [39] developed an improved percolation model for robust detection of cracks in noise-contaminated images. They used the characteristic of brightness and the feature of crack length to remove the noisy regions and detect cracks with reduced iterations and computation time.

The current paper investigates the performance of a natural language–based machine learning technique, called bag-of-words (BoW) [49, 51], which is flexible with a variety of options to use different classification techniques. The BoW method has been used in this paper for learning and detecting the behavior of crack evolution from (noise-contaminated) surface images from a microscope, in conjunction with the analytical tool of support vector machines (SVM) in the pattern classification stage [37]. The proposed BoW algorithm has been experimentally validated on a laboratory apparatus that is built upon an MTS 831.10 Elastometer fatigue test system, equipped with a computer-controlled and computer-instrumented Alicona confocal microscope system. Polycrystalline alloy specimens, subjected to externally applied cyclic loading, are tested on this apparatus, where images are taken from the microscope at the notch surface of the test specimen over many load cycles until the specimen breaks. The results of experimentation (i.e., pattern classification of noise-contaminated surface images from the microscope for crack damage estimation) have been used to develop decision models for crack detection. For the purpose of comparative performance evaluation, 21 different classification techniques have been tested within the BoW framework; among these classifiers, the quadratic support vector machine (QSVM) showed the best performance in terms of the classification accuracy and the area under the curve (AUC) of *receiver operating characteristics*[1] (ROC) for crack detection in the images.

**Contributions of the paper** The major contributions of the current paper are succinctly summarized below:

– *Nondestructive test and evaluation*: A statistical-analysis-based method of nondestructive test and evaluation is developed and validated with experimental data. The

[1]The receiver operating characteristic (ROC) curve for a binary classifier is a plot of the *True Positive Rate* (TPR) versus *False Positive Rate* (FPR) achieved by the classifier as the decision threshold is varied. Area under this curve (AUC) is a measure of the classifier performance; higher AUC scores are better, and the maximum achievable AUC score is 1 [37].

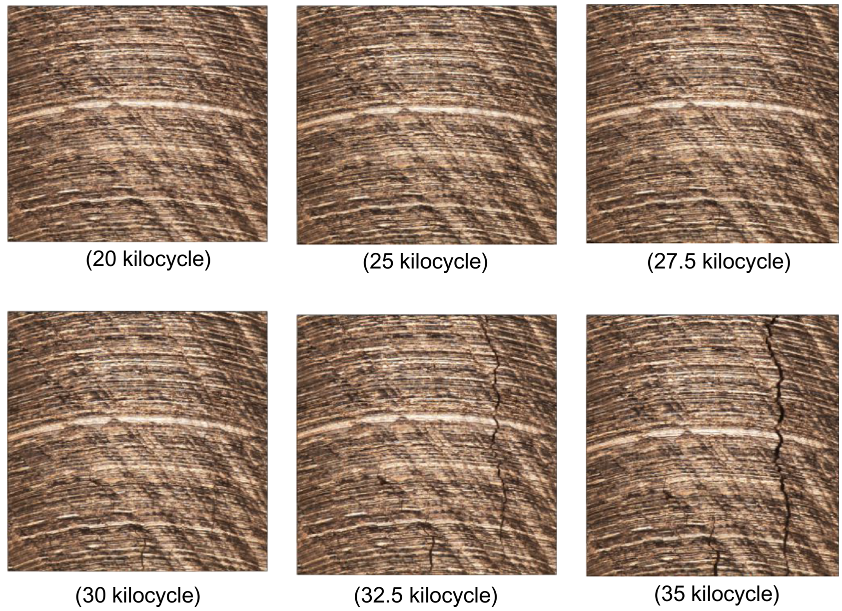


**Fig. 1** Images of fatigue crack evolution at increasing load cycles

underlying algorithms are built upon the concepts of bag-of-words (BoW) and quadratic support vector machine (QSVM) to train data-driven models that can be used for early detection of evolving fatigue cracks from surface images.

– *Validation with microscope images*: The results, predicted by the proposed BoW/QVSM method, have been validated with the data collected from an experimental apparatus.

**Organization of the paper** The paper is organized in five sections, including the present one, and two appendices. Section 2 presents a description of the main challenges in detecting cracks from the noise-contaminated microscope images and how such detection could be useful for early detection and estimation of evolving fatigue damage Section 3 provides some preliminary materials from machine learning theory as required for understanding the proposed algorithm for crack detection using microscope images. Section 4 briefly introduces the theory of bag-of-words (BoW) and its implementation for damage estimation, based on the crack detection from the images. Section 5 presents experimental validation of the proposed method. Section 6 summarizes and concludes the paper along with recommendations for future research. Appendix 1 briefly outlines the relative merits and demerits of various classification methods that are used in the performance evaluation of the proposed BoW-based method. Appendix 2 succinctly describes the theory of support vector machines (SVM) that serve as pattern classifiers in the proposed method.

## 2 Problem description

Fatigue crack evolution is a stochastic process that can be estimated using surface images at the locations where the structure is likely to start developing cracks, such as riveted joints, sharp corner and notches. The following paragraph explains the challenges of crack detection from such images and how this detection could be useful for estimation of fatigue crack evolution at an early stage.

The images in Fig. 1 were taken from a typical specimen under an external cyclic load of 3 kN amplitude applied on the specimen for the entire experiment at an average frequency of 50 Hz. For each image, the corresponding

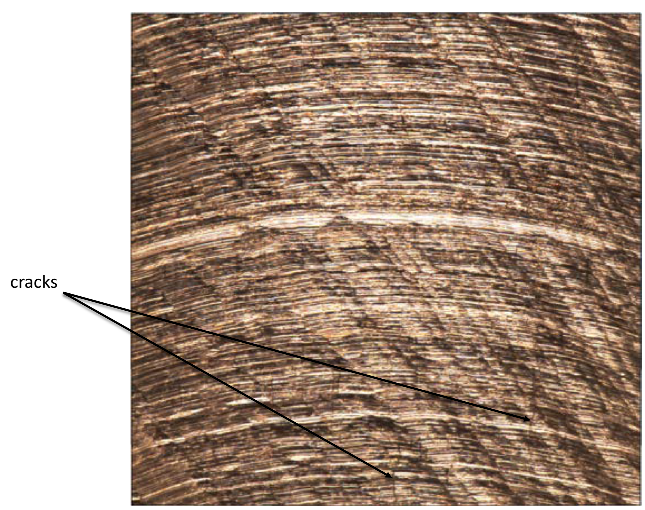


**Fig. 2** Early fatigue cracks at 27.5 kilocycles

number of load cycles is shown under the image. As shown in the enlarged image in Fig. 2, visible cracks start appearing around $\sim$ 27.5 kilocycles, where two cracks are barely visible in the image. As seen in the image in Fig. 2, cracks are not the only dark lines in the image and other dark lines could be due to material surface finishing and scratches. Distinguishing cracks from such types of dark lines in an automated manner is a challenge in this regard. Moreover, cracks may not occur in similar shapes as they propagate in a random fashion, yielding very different crack shapes and sizes. This makes the objective of *learning* the statistics of cracks from the image even more challenging; on the other hand, cracks could be due to an earlier phase of the fatigue damage process. For example, the specimen whose sample images are shown in Fig. 1 broke after 45 kilocycles. However, the cracks shown in Fig. 2 are corresponding to 27.5 kilocycles, which is an early stage of the specimen life, and the detection of such early fatigue cracks would be useful for maintenance scheduling and mission planning in order to mitigate the probability of an unanticipated failure.

Next, a nondestructive evaluation method is proposed, which makes use of a computer vision method, called bag-of-words (BoW) [49, 51], to detect fatigue cracks from surface images. It is reiterated that early detection of fatigue cracks, predicted by this method, is potentially applicable to life-extending control [9, 26]. However, before delving into details of the BoW-based fatigue crack detection, the following section summarizes preliminary notions and concepts of machine learning and computer vision that will be needed in developing the proposed detection technique.

## 3 Preliminaries

Classification and clustering are two commonly used machine learning techniques for interpreting data sets and developing models that are capable of predicting pertinent information. In this setting, classification is considered to be a supervised learning process, where input data and their corresponding labels are used to develop models that are capable of predicting labels, as accurately as possible, when these models are provided with new data. In contrast, clustering is considered to be an unsupervised learning process, where models are developed to interpret unlabeled input data and understand the hidden relationships among the elements of the given data set without regard to their labels. In the following subsections, both topics of classification and clustering are summarized, with a focus on image classification, for realization of BoW-based fatigue crack detection.

### 3.1 Classification

There are several techniques (e.g., support vector machines (SVM) and Bayesian methods) [37] available for classification of feature vectors. Given a data set $\{(x_1, y_1), \ldots, (x_n, y_n)\}$, where $x_i$ is the input vector and $y_i$ is the label of the input vector, $(i = 1, \ldots, n)$, let $L$ be the total number of classes, or categories, to which the labels $y_i$ belong. Then, classification can be defined as the process of employing the given data set to develop $L$ models that can realize the following mapping:

$$y_j(x_i) = \begin{cases} 1, & \text{if } x_i \in \text{class } j, \text{ where } j = 1, \cdots, L \\ 0, & \text{otherwise,} \end{cases} \tag{1}$$

The relationship in Eq. (1) is the general setting of classification for inputs that might necessitate pre-processing, especially for high-dimensional input vectors. For example, in image classification, inputs are images and if they are used without pre-processing, then the classification process would have inputs as pixels of given images that are high-dimensional, possibly rendering the task of image classification unrealizable. Thus, for successful realization of an image classification task, reasonably low-dimensional feature vectors are extracted out of a given set of images. This is accomplished by feature extraction that converts each image into a feature vector of lower dimension [37]. Feature extraction of images can be realized by myriad of techniques, such as speeded up robust features (SURF) [4] and scale-invariant feature transform (SIFT) [31]. The extracted feature vectors are used to develop models for the classification process.

### 3.2 Clustering

The main objective of clustering is to partition a given data set $\{x_1, \ldots, x_n\}$ into a number of subsets, where the elements of each subset share a common attribute, such as a threshold for a distance metric to the center of the subset, and possibly other attributes. Such a partitioning process is called clustering and it is performed without the need for labeling of the given data set; that is why it is called an unsupervised learning process. Multiple techniques can be employed for clustering of a data set; examples include Gaussian mixtures model [25], hierarchial clustering [16], and fuzzy c-means [24]. However, one of the most efficient clustering techniques is $K$-means, in which the given data set is grouped into $K$ clusters with $K \leq n$. The $K$-means clustering is achieved by minimizing the following objective functional:

$$J = \sum_{i=1}^{K} \sum_{x \in C_i} d(c_i, x)^2 \tag{2}$$

where $C_i$ is the $i$th cluster, $d(x, y)$ is the Euclidean distance between the data points $x$ and $y$, and $c_i$ is the centroid of the cluster $C_i$ obtained as:

$$c_i \triangleq \frac{1}{|C_i|} \sum_{x \in C_i} x, \tag{3}$$

with $|C_i|$ being the number of points in the $i$th cluster $C_i$. Minimization of the objective functional in Eq. (2) using $K$-means is implemented as follows:

- Initialize a set of $K$ centroids $\{c_1, \ldots, c_K\}$.
- Assign each $x_i$ to the cluster with the nearest centroid. Use Eq. (3) to update the values of centroids.
- Implement the cluster assignment and centroid steps iteratively until there are no changes in the cluster assignments and values of centroids.

## 4 Fatigue crack detection using the bag-of-words technique

The proposed methodology for fatigue crack detection is formulated as an image classification problem, where the underlying model is triggered as soon as a crack is identified in the image. The following three subsections present details of the detection procedure, starting with a brief description of the bag-of-words (BoW) technique [49, 51].

### 4.1 Bag-of-words technique

The BoW technique is embodied by three main stages: feature extraction, quantization, and classification, as described below [18].

- **Feature extraction:** The first stage of BoW is feature extraction which consists of two main steps.
  1. In the first step, the points of interest in the images are identified by using detection tools such as scale-invariant feature transform (SIFT) (see [3, 46] for further details on detection schemes).
  2. The second step estimates the descriptor in the vicinity of the points of interest, detected in the first step. Multiple schemes can be employed for estimation of descriptors, such as histogram of oriented gradients (HOG) and wavelets (see, for example, [13, 35, 46]). Therefore, the input to this stage is an image, and the output of this stage is a set of descriptors which represent the features extracted from the image.
- **Quantization:** The second stage of BoW is the quantization process that also consists of two steps.
  1. The clustering process that groups the descriptors into subsets that are called visual words. Any clustering technique can be employed to implement this step; examples are $k$-means, fuzzy c-means, and Gaussian mixture model (see, for example, [6, 37] for further details on clustering techniques). Due to its simplicity and efficient performance, $K$-means, explained in Section 3, is frequently employed for clustering in the BoW technique.
  2. Construction of the histogram of the visual words. This is done by computing the frequency of occurrence of each visual word by counting how many descriptors in each cluster. Thus, the quantization process generates a histogram of the visual words for each image. Although the histograms can be used directly in the classification process, their normalization might be preferable to make the frequency of occurrence of words within a certain range. The main purpose of such normalization is to avoid high numbers in the subsequent computations in classification stage.
- **Classification:** The last stage of BoW is the classification process by which a decision rule is trained for image classification. The inputs to the classification stage are the histograms that are generated earlier; and the outputs are the class labels of the images. Many classification schemes can be employed to implement this stage; examples are decision tree classifier (DTC), linear discriminant classifier (LDC), and support vector machine (SVM) (see, for example, [6, 37] for further details on classification techniques). Appendix 1 presents a brief description of the performance of a set of classifiers used in this paper. Due to its efficiency, quadratic support vector machine (QSVM) is chosen among these classifiers for accomplishing the classification stage. Details on SVM are presented in Appendix 2.

### 4.2 BoW-based detection

The problem of fatigue crack detection is formulated in this paper as a binary classification, where the nominal (healthy) case is denoted as Class 0 and the damage case is Class 1. Thus, two models are required to be developed that can generate, as precisely as possible, the following binary classification rule:

$$y(\mathcal{I}) = \begin{cases} 0, & \text{if } \mathcal{I} \in \text{Class } 0 \\ 1, & \text{if } \mathcal{I} \in \text{Class } 1. \end{cases} \tag{4}$$

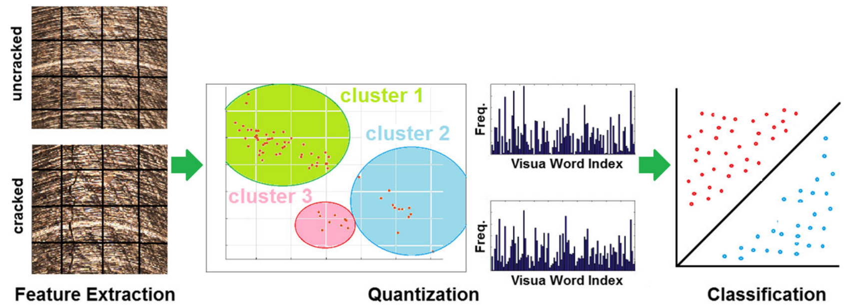


**Fig. 3** Concept of bag-of-words (BoW) for detection of fatigue crack

where $y(\mathcal{I})$ is the output of the classification algorithm, and $\mathcal{I}$ is the sensed image.

The BoW modeling technique, explained above, is employed to realize Eq. (4), and Fig. 3 provides a conceptual presentation of the modeling sequence. Multiple specimens are considered and different cyclic loads are applied to the specimens until they break. During this process, microscope images at the notch surface of the specimens are captured and grouped into nominal and damaged sets. Then, the BoW concept is employed to develop models for the nominal and damaged phases for crack detection. This procedure would yield detection of a crack by examining the image, and further action is required to measure the detected crack. Once a crack is detected, the next step is the measurement step from the image of the confocal microscope. Alternatively, one could consider $M + 1$ hypotheses, instead of the binary hypotheses, where the first model corresponds to the nominal (undamaged) case and the remaining $M$ models are corresponding to $M$ levels of the normalized crack length [40]. In this case, the BoW algorithm would yield both the detection of the crack existence and, based on positive detection, a quantized estimate of the normalized crack length. Further details are given in Section 5, where the results are presented for experiments conducted on multiple specimens.

### 4.3 Usage of BoW fatigue damage model in real-life applications

The BoW algorithm has been developed and validated in the laboratoy environment, where a confocal microscope has been used for generating the images to detect the cracks; however, in real-life applications, usage of such a confocal microscope may not be feasible due to its cost and fragility. Therefore, an appropriate sensor (e.g., ultrasonic sensor [20]) may be deployed to obtain measurements that could be used as inputs to a data-driven model for online monitoring of fatigue damage, instead of using images from a microscope. Alternatively, using the information on applied load (e.g., forces and moments), a finite element model (FEM) could estimate the normalized crack length $\hat{c}_t$ via the following equations [40]:

$$\hat{a}_t - \hat{a}_{t-\delta t} = h\left(\Delta K_t^{eff}\right)\delta t \quad \text{with } h(0) = 0 \text{ for } t \geq t_0 \text{ and given } \hat{a}_{t_0} > 0 \tag{5}$$

$$\Delta K_t^{eff} = \Delta S_t^e \sqrt{\pi \hat{a}_{t-\delta t}}\, F(\hat{a}_t) \tag{6}$$

$$\Delta S_t^e = \left[S_t^{\max} - \max(S_t^0, S_{t-\delta t}^{\min})\right] U(S_t^{\max} - S_{t-\delta t}^0) \tag{7}$$

$$\hat{c}_t = \frac{\hat{a}_t}{w} \tag{8}$$

where $t$ is the current time upon completion of a stress cycle; $t_0$ is the initial time; $\hat{a}_t$ is the estimated mean value of crack length; $h(.)$ is a non-negative measurable function that is dependent on the material and geometry of the stressed component; $F$ is a dimensionless correlation factor that is a function of the geometric configuration (e.g., thickness, width, and the crack type in the stressed component) and the crack length; $w$ is a normalizing factor that depends on the component geometry; $\Delta S_t^e$ is the effective stress range during the cycle completed at time $t$ with the corresponding crack opening stress $S_t^0$, maximum stress $S_t^{\max}$, and minimum stress $S_t^{\min}$; and $U(.)$ is the unit step function defined as:

$$U(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geq 0 \end{cases} \tag{9}$$

Regardless of whether an inexpensive sensor-based model or an FEM model is used, the resulting damage monitoring is likely to be inaccurate and therefore needs to be calibrated by a more reliable source such as the image-based BoW model developed in this paper. From this perspective, the BoW fatigue damage model can be used to routinely train/calibrate other (inexpensive and less reliable) models to ensure the accuracy of damage estimation.

As an example, Fig. 4 illustrates how the proposed BoW model could be used to calibrate an ultrasonic sensor model, where both (confocal microscope) images and ultrasonic data are synchronously collected for the same specimen, and the ultrasonic sensor model's parameters are calibrated such that the error of damage estimation by the sensor model relative to the damage estimated by the BoW model is significantly reduced.

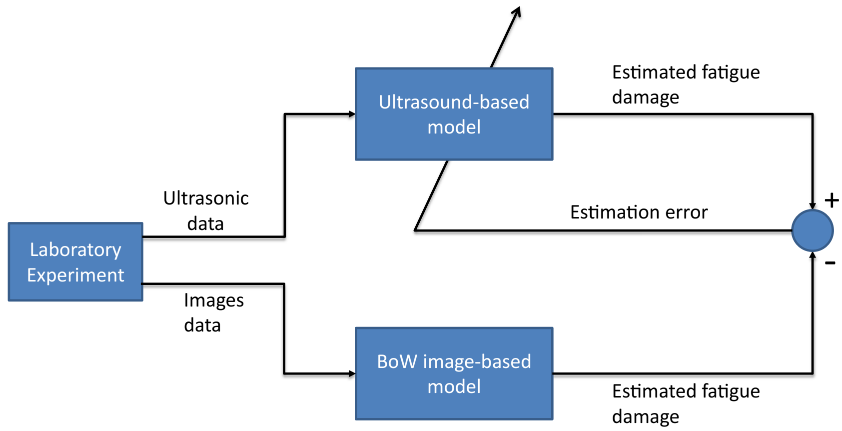


**Fig. 4** Calibration of an ultrasonic sensor using BoW

## 5 Results and discussion for model validation

This section presents the results of the BoW model, developed in Section 4, for validation by using the fatigue data on a laboratory apparatus.

### 5.1 Description of the experimental apparatus

Figure 5 shows the experimental apparatus that is built upon an MTS 831.1 Elastometer fatigue testing machine, which can be used to apply external load to test specimens with the desired cyclic load properties: amplitude, frequency, and the shape of the force function; it is also capable of applying random loading. The other major component of the apparatus is a computer-equipped confocal microscope, Alicona R25 sensor system, which has a high resolution and magnification power and an excellent capability to take images from different angles and convert the image to RGB numerical data. The microscope is mounted to an Aerotech Motion stage, which is a high-precision moving frame that can precisely move the microscope to the desired position relative to the specimen notch surface.

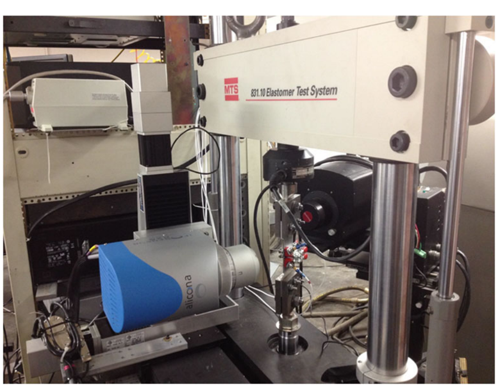

**Fig. 5** The fatigue test machine equipped with a traveling confocal microscope

The images with no cracks are considered to belong to the nominal model (Class 0), while images with possible cracks belong to the damaged model (Class 1). The cross-validation method [6] has been used to train and test the classification performance, where the BoW model is trained at several time epochs based on labeled specimens. Then, this trained model is used to classify the other unknown images and make a decision for each image whether it belongs to Class 0 (nominal) or Class 1 (damaged). Then, the classification error is computed for each classifier being used to find the classification performance.

### 5.2 Model validation with experimental data

As explained in Section 4, the quantization stage in the BoW technique results in assigning a histogram of the visual words to each image and therefore each image is represented as a point in the feature space of visual words. Consequently, good feature extraction and quantization would enhance the quality of discrimination between nominal and damaged images in the feature space of visual words. To demonstrate this fact, the feature extraction and quantization stages of the BoW are implemented on two different images of the same specimen: one with a crack and one without a crack. The speeded up robust features (SURF) technique, which is a speeded up version of the scale-invariant feature transform (SIFT) [3], is adopted in the feature extraction stage, and the *k*-means is employed in the quantization stage. To choose the number of clusters, we tested several choices on a validation subset of the image ensemble. Among these choices, we pick the number of clusters = 1200 which showed the best detection performance. The systematic selection process of the optimal number of clusters in this context is in fact an open problem which will be discussed in Section 6. Figure 6 presents the results, which show that different histograms are obtained for nominal and damaged phases. However, just the differences in histograms between the images of

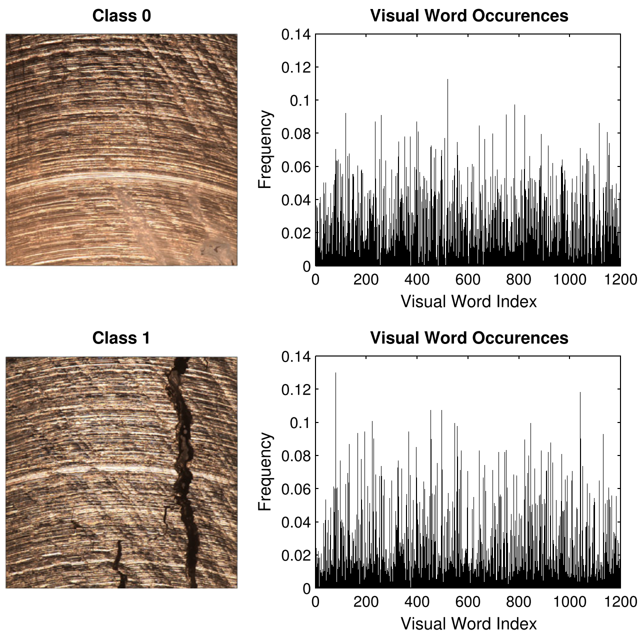


**Fig. 6** Histograms for nominal (Class 0) and damaged (Class 1) cases

nominal and damaged phases may not be sufficient for reliable classification. The key point here is to identify common patterns (i.e., signature) for the damaged images despite diversity in the shapes and size of cracks, with different textures of the images, and other patterns for the undamaged images which may also largely differ in their textures. The performance in this regard would depend on the three stages (namely, feature extraction, quantization, and classification) of the BoW technique as outlined in Section 4. It is also noted that different classifiers may result in different classification performance.

### 5.3 Performance evaluation of the BoW-based detection

This subsection evaluates the performance of the BoW-based detection, using different methods in the classification stage, for classifying images of specimens without and with cracks. The evaluation is made based on 174 images from a microscope, some of which contains cracks, ranging from very tiny cracks to big ones, and the others do not contain any cracks. These images belong to different specimens and corresponds to different regions in the notch part of the specimen, where cracks are most likely to initiate. Figure 1 shows some of the images for a sample specimen at different load cycles.

The SURF technique [3] is used in the feature extraction phase, which generates 1,887,436 descriptors from all the 174 images. The quantization phase utilizes $k$-means clustering with 1200 clusters. Each of these clusters represents a visual word and each image generates a histogram of the visual words, where the frequency of each visual word is given by the number of descriptors that belong to the cluster representing this visual word.

**Table 1** Classification performance of the BoW model with different classifiers

| No. | Classifier description | Performance (%) |
|---|---|---|
| 1 | Decision Trees | 85.7 |
| 2 | Linear Discriminant | 82.4 |
| 3 | Quadratic Discriminant | 89.0 |
| 4 | Logistic Regression | 63.7 |
| 5 | Linear SVM | 83.5 |
| 6 | Quadratic SVM | *94.5* |
| 7 | Cubic SVM | 93.4 |
| 8 | Fine Gaussian SVM | 80.2 |
| 9 | Medium Gaussian SVM | 86.8 |
| 10 | Coarse Gaussian SVM | 80.2 |
| 11 | Fine KNN | 86.8 |
| 12 | Medium KNN | 84.6 |
| 13 | Coarse KNN | 80.2 |
| 14 | Cosine KNN | 85.7 |
| 15 | Cubic KNN | 85.7 |
| 16 | Weighted KNN | 87.9 |
| 17 | Ensemble Boosted Trees | 80.2 |
| 18 | Ensemble Bagged Trees | 90.1 |
| 19 | Ensemble Subspace Discriminant | 93.4 |
| 20 | Ensemble Subspace KNN | 85.7 |
| 21 | Ensemble RUSBoosted Trees | 87.9 |

The three important observations from the above evaluation process are as follows: (i) a large number of feature vectors (e.g., 1,887,436 descriptors) for a relatively small (only 174) number of test images; (ii) different textures of images from the same group (i.e., belonging to either damaged or undamaged) of test images; and (iii) various different shapes and sizes of the emerging cracks to be learned. A large number of feature vectors is required for learning the highly random cracks with the different image textures.

We use 21 different classifiers in the classification stage of the BoW method. Table 1 lists the classification performance of each one of these classifiers, and a detailed discussion of these classifiers performance is given in Appendix 1. As seen in Table 1, quadratic support vector machine (QSVM), cubic support vector machine (CSVM), and ensemble subspace discriminant (ESD) classifiers yield better performance compared with the other classifiers, while QSVM is found to have the best classification performance of 94.5%.

The best combination (BoW/QSVM) has been used to compute the receiver operating characteristic (ROC) curve [37] for crack detection in the images. The results are shown

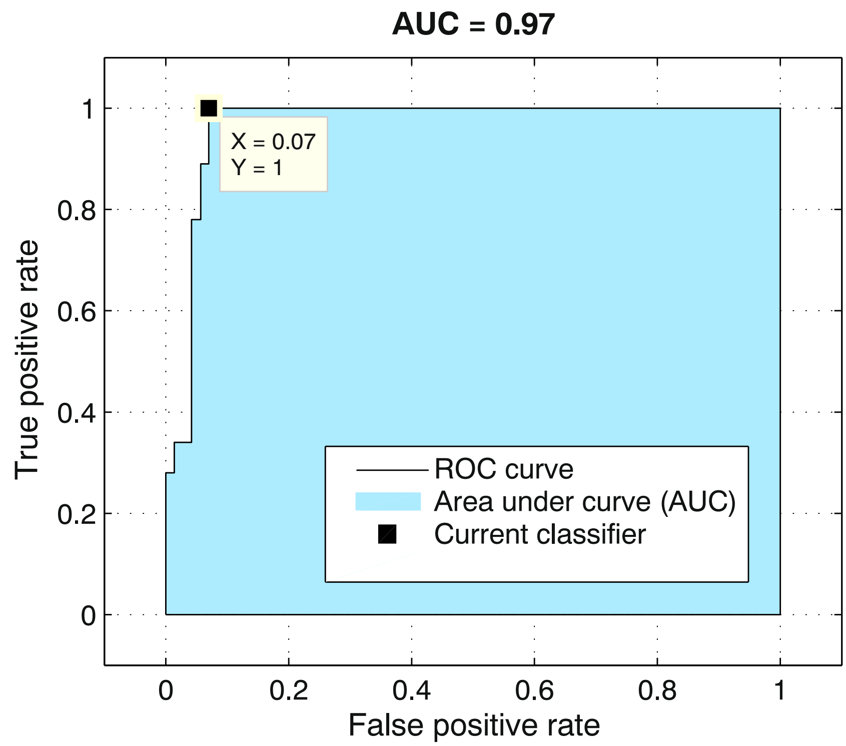


**Fig. 7** ROC curve of BoW/QSVM for detecting the nominal (Class 0) model

in Figs. 7 and 8 for nominal (Class 0) (i.e., no crack) and cracked (Class 1) (i.e., with cracks) cases, respectively. For both cases, the BoW method shows the area under the curve (AUC) to be 0.97. This is an excellent performance in view of the fact that the maximum value of AUC is 1.0.

The images used in this paper have been taken from different specimen samples, at different locations within the notch area for each specimen, and under different lighting conditions. Figure 9 shows samples of the images used

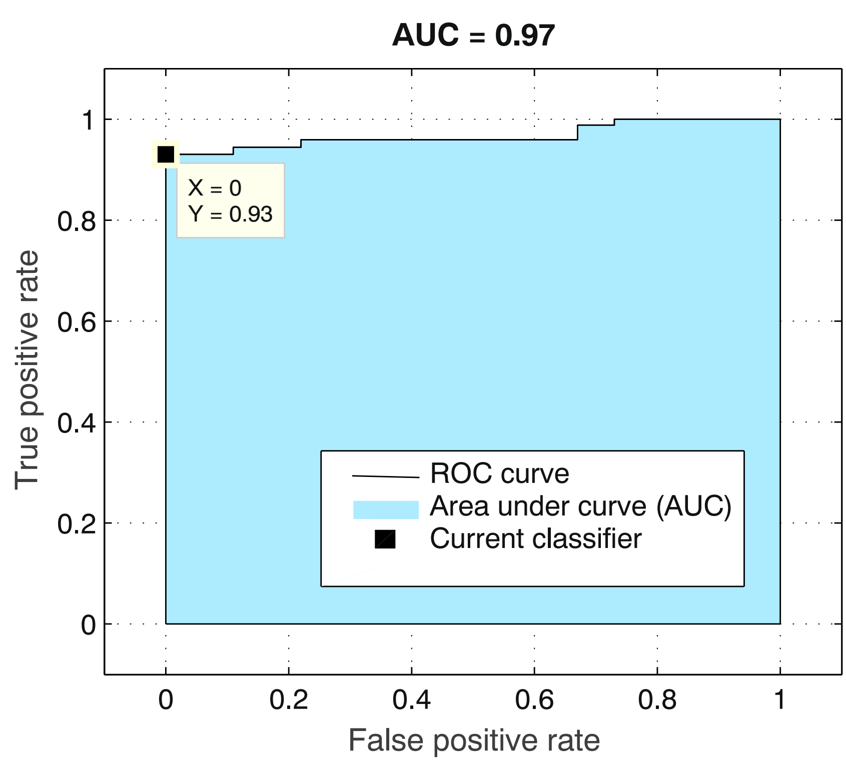


**Fig. 8** ROC curve of BoW/QSVM for detecting the damaged (Class 1) model

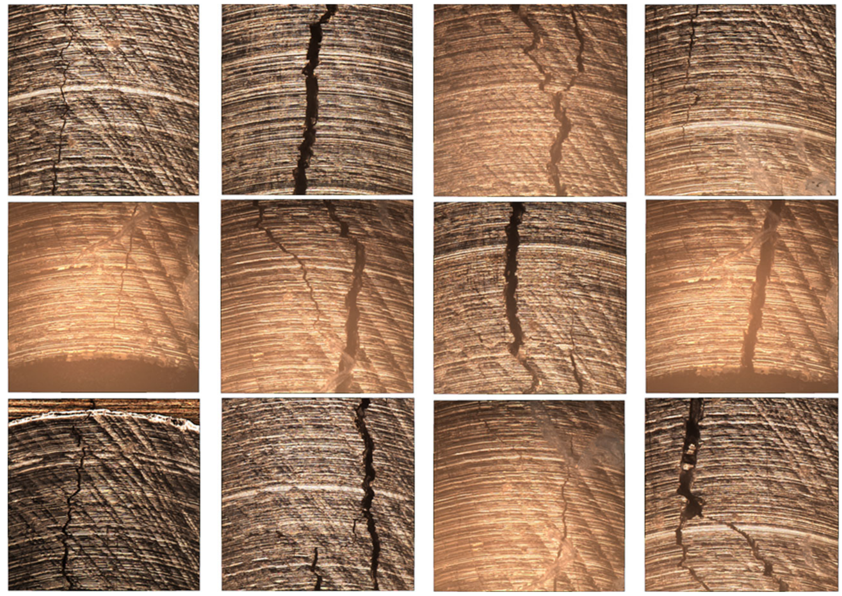

**Fig. 9** Variation of sample images in illumination, object surface property, and cracks' shape and size

in this work, where the images in the individual twelve plates are subjected to different levels of illumination, object surface property, and shape and size of cracks. The classification accuracy of $\sim 94.5\%$ and the ROC detection performance of $\sim 0.97$ achieved by the proposed BoW-based feature extraction algorithm combined with a QSVM demonstrate robustness of the proposed technique to these environmental changes.

## 6 Summary and conclusions

This paper has developed and validated a data-driven method of detecting fatigue crack evolution in polycrystalline alloys. The proposed method has industrial applications toward automated inspection of machinery components. The potential applications include power plants, chemical plants, and transportation vehicles.

Relying on the images captured by a confocal microscope under a variety of environmental conditions (e.g., different illumination, object surface property, and shape and size of cracks), the algorithm of crack detection is formulated as a binary classification problem that is composed of two classes: (i) nominal (Class 0) (i.e., no cracks) corresponding to the healthy phase of the structural material and (ii) damaged (Class 1) corresponding to images with cracks. The bag-of-words (BoW) technique is proposed for classification of images with the QSVM algorithm. The classification performance has been computed to be $\sim 94.5\%$ reflecting an excellent performance for early detection of fatigue crack damage by using the BoW/QVSM algorithm. The performance of the BoW algorithm is comparatively evaluated by employing several classification schemes and superiority of the BoW/QSVM is established.

Despite the excellent performance of the BoW/QSVM method, there are several areas of theoretical and experimental research that must be conducted before the proposed BoW technique can be optimally applied to real-life problems. To this end, the authors suggest the following topics for future research:

– *Calibration of inexpensive and environmentally rugged sensors with the proposed BoW/QSVM algorithm*: The usage of a confocal microscope in real-life applications may not be feasible due to its cost and fragility. Therefore, an appropriate sensor (e.g., ultrasonic sensor [20]) may be deployed for online monitoring of fatigue damage, instead of using images from a microscope; alternatively, the information generated from a finite element model (FEM) could be used to estimate the crack length [40]. The BoW model developed in this paper could be used in the laboratory setting to calibrate/enhance such an inexpensive damage detector.
– *Estimation of the number of visual words*: This information resulting in optimal fatigue crack damage detection performance is an open problem and might result in a significant enhancement of the Bow/QSVM algorithm.
– *Multiclass classification*: The BoW method is used in this paper to do binary classification of images and detect whether a crack exists in the image or not. However, to estimate the fatigue damage level, an estimate of the normalized crack $c_t$ is needed. Once a crack is detected in the image, the confocal microscope utilized in this paper can be used to measure the crack. Alternatively, one can partition the space

of the normalized crack lengths into $M$ regions, with each region representing a crack length level. Then, $M + 1$ classes are considered, where the first class corresponds to the nominal (undamaged) case and the remaining $M$ classes are corresponding to the $M$ levels of the normalized crack length. In this case, a BoW algorithm can be developed to learn $M + 1$ statistical models for the nominal (undamaged) case and for each crack level, and images are classified into $M + 1$ classes, instead of two only. In this scenario the BoW algorithm would yield both the detection of the crack existence and estimation of the fatigue damage level in an automated fashion, without the need for the crack length measurement by the confocal microscope. However, this would impose more challenge to the problem, where instead of learning common patterns for the nominal images and other common patterns for the damaged images, the developed algorithm needs to learn patterns for each crack length level despite the random nature of the crack shapes. This should be a good topic for future experimental study.

**Funding information** The work reported in this paper has been supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) under Grant Nos. FA9550-15-1-0400 and FA9550-18-1-0135 in the area of dynamic data-driven application systems (DDDAS). Any opinions, findings, and conclusions in this paper are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

## Appendix 1. Comparison of classification methods

The various classification methods, listed in Table 1, which yield relatively low performance, generally have specific shortcomings (see [6, 21, 37, 47] for more details), possibly because they rely on assumptions that may not be valid under the encountered conditions. These issues are discussed below.

1. While the *Decision Trees* method is known to suffer from lack of robustness, where small changes in the input data can cause large deviations in the tree structure, the complexity of this method may also lead to overfitting problems.
2. The *Discriminant Analysis* technique may become very sensitive to outliers and also rely on assumptions that may not be valid in the specific application. For example, each discriminant function has approximately the same variance over different classes of images, which is unlikely to be fulfilled, because different images from the same group (nominal or damaged) may have very different textures and crack shapes and sizes, where the corresponding feature vectors are likely to be largely different.
3. The *Logistic Regression* method has a major disadvantage that it requires independent inputs. For example, if multiple images are taken at the same location (e.g., at the notch of a specimen) under different load cycles, then the resulting images could have very similar texture, which would generate highly correlated feature vectors. Logistic Regression is also known for overfitting problems.
4. The *K nearest neighbor (KNN)* classifier has a fundamental flaw that it does not rely on a learning procedure to enable capturing patterns in the data efficiently. Given a new input, *KNN* would find the closest $K$ neighbors to the new input from the training data, and assign it the most common label among the $K$ closest neighboring points. The problem at hand requires the classifier to be capable of *learning* existence of cracks, despite their different shapes and sizes, and to identify their patterns in the data that would uniquely distinguish them from other types of dark lines in the images (e.g., scratches and material surface finishing lines).
5. The *ensemble subspace discriminant (ESD)* classifier has two main advantages [5, 22]: (i) it combines the models produced by several discriminant learners into an ensemble that performs better than the original learners, and (ii) only part of the features are used for training each one of the original learners by randomly sampling from the original features. By doing so, the individual learners will tend not to over-focus on features that appear highly predictive in the training set, but not for points outside that set. This could be very useful in situations where the number of features is much larger than the number of training points; this is the situation in our case where we have 1,887,436 descriptor (feature) with only 174 total images.
6. The *support vector machines (SVM)* serve as classifiers and rely on convex optimization for which local optimality is global as well. Linear SVM happen to be the simplest SVM configuration (which is optimal for linearly separable data) and is unlikely to be suitable for images with different texture and having different cracks lengths and shapes that would normally give rise to largely nonlinearly separable data. Therefore, a significant improvement is expected if a nonlinear SVM is used instead of a linear SVM. As explained in Appendix 2, among many types of nonlinear SVM, the quadratic SVM (QSVM) and cubic SVM (CSVM) have the key feature that, while they are capable of handling nonlinearly separable data, they are less prone to overfitting compared with higher order kernels. It is noted that QSVM is less susceptible to overfitting than CSVM, as explained in Appendix 2.

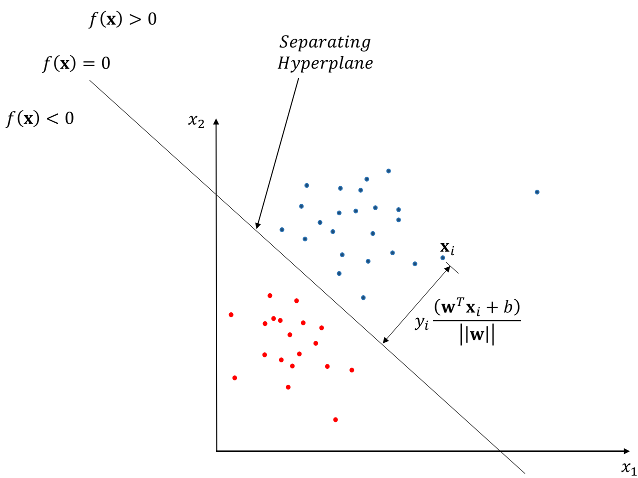


**Fig. 10** Binary classification for linearly separable data

## Appendix 2. Support vector machines

In classification problems, a classifier is constructed upon a *discriminant function* $f(\mathbf{x})$ that assigns each input vector $\mathbf{x}$ to one of the (a priori defined) $K$ classes. To put this in perspective, consider a two-dimensional *linearly separable* data set $\mathcal{D} \triangleq \{\mathbf{x}_i, y_i\}$, for which each input vector $\mathbf{x}_i$ belongs to one of the two classes $\mathcal{C}_0$ or $\mathcal{C}_1$, where each $\mathbf{x}_i$ is labeled by a *class index* $y_i \in \{-1, 1\}$ such that:

$$y_i = \begin{cases} -1, & \mathbf{x}_i \in \mathcal{C}_0 \\ 1, & \mathbf{x}_i \in \mathcal{C}_1 \end{cases} \tag{10}$$

This is called the *binary classification* problem, where a linear discriminant function is trained as:

$$f(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b \tag{11}$$

by finding a *weight vector* $\mathbf{w}$ and a *bias* $b$ such that, for all $\mathbf{x}_i \in \mathcal{D}$:

$$\begin{aligned} f(\mathbf{x}_i) < 0\,, \ \mathbf{x}_i \in \mathcal{C}_0 \\ f(\mathbf{x}_i) > 0\,, \ \mathbf{x}_i \in \mathcal{C}_1 \end{aligned} \tag{12}$$

Then, for any new input $\mathbf{x}$, the corresponding class index $y$ is estimated as:

$$\hat{y} = \operatorname{sign}(\mathbf{w}^T\mathbf{x} + b) \tag{13}$$

where sign(.) is the signum function defined as:

$$\operatorname{sign}(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0 \end{cases}$$

Figure 10 illustrates the idea of training a classifier on a linearly separable data, where a hyperplane separating the two classes is given by the straight decision line $f(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b = 0$.

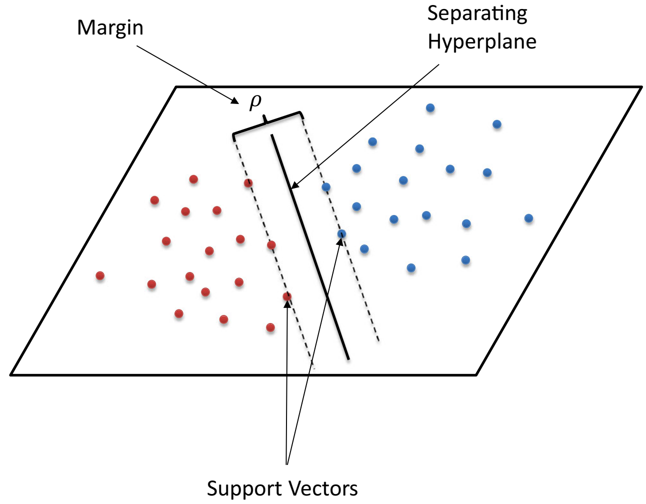


**Fig. 11** Linearly separable data

It can be shown that the distance $r(\mathbf{x}_i)$ from any point $\mathbf{x}_i$ to the decision line is given by:

$$r(\mathbf{x}_i) \triangleq y_i \frac{(\mathbf{w}^T\mathbf{x}_i + b)}{||\mathbf{w}||} \tag{14}$$

The data inputs with minimum distance to the separating hyperplane are called *Support Vectors*. The *Margin*, $\rho$, is the maximum width of the band that can be drawn separating the support vectors of the two classes. Figure 11 illustrates these two concepts. It follows from Eq. (14) that if one scales $\mathbf{w} \to c\mathbf{w}$ and $b \to cb$, the distance from any point to the decision line is unchanged. For convenience, we scale $\mathbf{w}$ and $b$ such that

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) = 1 \tag{15}$$

at the support vectors of each class. Hence, for all $(\mathbf{x}_i, y_i) \in \mathcal{D}$:

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 \tag{16}$$

and, from Eqs. (14) and (15), the margin is given by:

$$\rho = \frac{2}{||\mathbf{w}||} \tag{17}$$

The main idea of SVMs is to maximize the margin while zero classification error is maintained. The optimal separating hyperplane satisfying these two conditions is called the *Support Vector Machine* [47]. These two conditions can be formulated as:

- $\rho = \dfrac{2}{||\mathbf{w}||}$ is maximized, and
- For all $(\mathbf{x}_i, y_i) \in \mathcal{D}$, $y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1$

This can be converted to the minimization problem:

$$\left(\hat{\mathbf{w}}, \hat{b}\right) = \underset{\mathbf{w}, b}{\arg\min} \left[\frac{1}{2}||\mathbf{w}||^2\right] \tag{18}$$

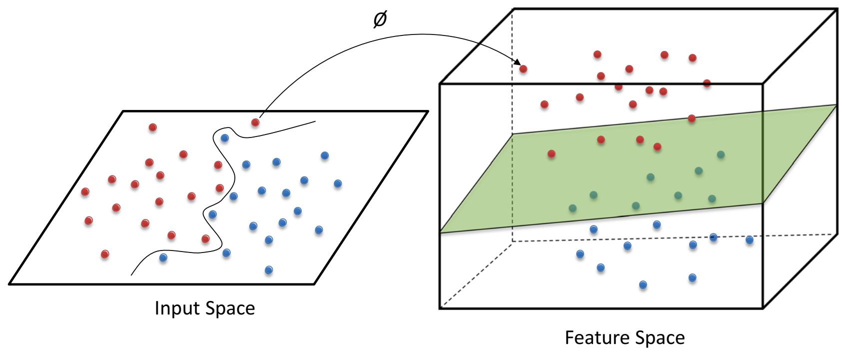


**Fig. 12** Mapping nonlinearly separable data into linearly separable feature vectors

subject to the constraint:

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1, \text{ for all } (\mathbf{x}_i, y_i) \in \mathcal{D} \tag{19}$$

It is worth pointing out that the cost function in Eq. (18) is convex and the constraints in Eq. (19) consists of linear functions. These two conditions together lead to a unique feature of SVMs that any local minimum is also global so that SVM optimal classifier is unique [47]. This constrained optimization problem can be solved using Lagrange multipliers technique [6]. Hence, the constraint (19) needs to be satisfied for all data inputs. Assuming that $N$ data inputs are available, $N$ Lagrange multipliers $\underline{\alpha} = [\alpha_1, \ldots, \alpha_N]$ are needed to obtain the Lagrangian function

$$L(\mathbf{w}, b, \underline{\alpha}) = \frac{1}{2}||\mathbf{w}||^2 - \sum_{i=1}^{N} \alpha_i\{y_i(\mathbf{w}^T\mathbf{x}_i + b) - 1\} \tag{20}$$

The solution [32] is given by:

$$\mathbf{w} = \sum_{i=1}^{N} \alpha_i y_i \mathbf{x}_i \tag{21}$$

and

$$b = y_k - \mathbf{w}^T\mathbf{x}_k \text{ for any } \mathbf{x}_k \text{ such that } \alpha_k \neq 0 \tag{22}$$

where the optimal Lagrange multipliers are computed by solving the Wolf-dual optimization problem [47]:

$$\underset{\underline{\alpha}}{\arg\max} \left( \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \right) \tag{23}$$

such that:

$$\alpha_i \geq 0, \quad i = 1, \ldots, N \tag{24}$$

$$\sum_{i=1}^{N} \alpha_i y_i = 0 \tag{25}$$

Once the optimal Lagrange multipliers have been computed, any new input $\mathbf{x}$ can then be classified by using the optimal SVM classifier:

$$f(\mathbf{x}) = \text{sign}\left( \sum_{i=1}^{N} \alpha_i y_i \mathbf{x}_i^T \mathbf{x} + b \right) \tag{26}$$

Interestingly, $\alpha_i$ is zero for all the data inputs except the support vectors. Since usually only few inputs are support vectors, this has two important advantages:

- The computation in Eqs. (21) and (26) is greatly reduced, and
- The complexity of the resulting classifier is characterized by the number of support vectors rather than the dimensionality of the feature space so that SVMs are generally less susceptible to problems of overfitting than other types of classifiers [14].

### A2.1 Nonlinear SVM

So far it has been assumed that the data set is linearly separable. However, this may not be the case in many real-life applications, where data are usually nonlinearly separable. SVM can efficiently handle such cases through mapping nonlinearly separable data into a higher-dimension linearly separable feature vectors for which linear decision hyperplane can be used for classification. The idea is illustrated in Fig. 12, where a two-dimensional nonlinearly separable data is mapped to a three-dimensional linearly separable feature vectors using a map $\phi$.

This is called the "the kernel trick," which is simply done by replacing $\mathbf{x}$ in the previous analysis by $\phi(\mathbf{x})$. Then, Eq. (26) is rewritten as:

$$f(\mathbf{x}) = \text{sign}\left( \sum_{i=1}^{N} \alpha_i y_i \mathcal{K}(\mathbf{x}_i, \mathbf{x}) + b \right) \tag{27}$$

where $\mathcal{K}(\mathbf{x}_i, \mathbf{x}) \triangleq \phi(\mathbf{x}_i)^T \phi(\mathbf{x})$ is called the *kernel function*. There are many kinds of kernels that have been used in this regard, among them are the polynomial kernel functions given by the general form [32]:

$$\mathcal{K}(\mathbf{x}_i, \mathbf{x}) \triangleq \left(1 + \mathbf{x}_i^T \mathbf{x}\right)^d \tag{28}$$

If $d = 1$, it is the linear kernel, which have been explained so far for linearly separable data, where the constant 1 would just change the threshold. Among the most commonly used kernel is the *Quadratic Kernel* given by Eq. (28) with $d = 2$.

A key feature of this type of kernel is that, while it is capable of handling nonlinearly separable data, it is less prone to overfitting problems compared with higher order polynomial kernels.

## References


1. Aeran A, Siriwardane S, Mikkelsen O, Langen I (2017) A new nonlinear fatigue damage model based only on s-n curve parameters. Int J Fatigue 103:327–341
2. Anwar SA, Abdullah MZ (2014) Micro-crack detection of multicrystalline solar cells featuring an improved anisotropic diffusion filter and image segmentation technique. EURASIP J Image Video Process 2014(1):15. https://doi.org/10.1186/1687-5281-2014-15
3. Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-up robust features (SURF). Comput Vis Image Underst 110(3):346–359. https://doi.org/10.1016/j.cviu.2007.09.014
4. Bay H, Tuytelaars T, Van Gool L (2006) Surf: Speeded up robust features. In: Leonardis A, Bischof H, Pinz A (eds) Computer vision – ECCV 2006. Springer, Berlin, pp 404–417
5. Bertoni A, Folgieri R, Valentini G (2005) Bio-molecular cancer prediction with random subspace ensembles of support vector machines. Neurocomputing
6. Bishop CM (2006) Pattern recognition and machine learning. Springer, New York
7. Bjerken C, Melin S (2003) A tool to model short crack fatigue growth using a discrete dislocation formulation. Int J Fatigue 25(6):559–566
8. Bovsunovskii A (2014) Asynchronous connection of a turbine generator to the mains as a factor of fatigue damage of steam turbine shafting. Strength Mater 46(6):810–819
9. Caplin J, Ray A, Joshi S (2001) Robust damage-mitigating control of aircraft for high performance and structural durability. IEEE Trans Aerosp Electron Syst 37(3):849–862
10. Cha YJ, Choi W, Büyüköztürk O (2017) Deep learning-based crack damage detection using convolutional neural networks. Comput-Aided Civ Infrastructu Eng 32(5):361–378. https://doi.org/10.1111/mice.12263
11. Cook D, Berthelot Y (2001) Detection of small surface-breaking fatigue cracks in steel using scattering of Rayleigh waves. NDT E Int 34(7):483–492
12. Cubero-Fernandez A, Rodriguez-Lozano FJ, Villatoro R, Olivares J, Palomares JM (2017) Efficient pavement crack detection and classification. EURASIP J Image Video Process 2017(1):39. https://doi.org/10.1186/s13640-017-0187-0
13. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 1, pp 886–893. https://doi.org/10.1109/CVPR.2005.177
14. Duda R, Hart P, Stork D (2012) Pattern classification, 2nd edn. Wiley, USA
15. Emanuel Aldea SLHM (2015) Robust crack detection for unmanned aerial vehicles inspection in an a-contrario decision framework. J Electron Imaging 24:24–24–16. https://doi.org/10.1117/1.JEI.24.6.061119
16. Fernández A, Gómez S (2008) Solving non-uniqueness in agglomerative hierarchical clustering using multidendrograms. J Classif 25(1):43–65. https://doi.org/10.1007/s00357-008-9004-x
17. Gaur V, Doquet V, Persent E, Mareau C, Roguet E, Kittel J (2016) Surface versus internal fatigue crack initiation in steel: influence of mean stress. Int J Fatigue 82:437–448
18. Ghalyan IFJ, Chacko SM, Kapila V (2018) Simultaneous robustness against random initialization and optimal order selection in bag-of-words modeling. Pattern Recogn Lett 116:135–142. https://doi.org/10.1016/j.patrec.2018.09.010
19. Grondal S, Delebarre C, Assaad J, Dupuis J, Reithler L (2002) Fatigue crack monitoring of riveted aluminium strap joints by lamb wave analysis and acoustic emission measurement techniques. NDT E Int 35(3):137–146
20. Gupta S, Ray A, Keller E (2007) Online fatigue damage monitoring by ultrasonic measurements: a symbolic dynamics approach. Int J Fatigue 29(6):1100–1114
21. Hastie T, Tibshirani R, Friedman J (2016) The elements of statistical learning: data mining, inference, and prediction, 2nd edn. Springer, Berlin
22. Ho T (1998) The random subspace method for constructing decision forests. IEEE Transactions on Pattern Analysis and Machine Intelligence
23. Ishihara S, McEvily A (2002) Analysis of short fatigue crack growth in cast aluminium alloys. Int J Fatigue 24(11):1169–1174
24. Jasim IF, Plapper PW (2013) T-S fuzzy contact state recognition for compliant motion robotic tasks using gravitational search-based clustering algorithm. In: 2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp 1–8. https://doi.org/10.1109/FUZZ-IEEE.2013.6622415
25. Jasim IF, Plapper PW (2014) Contact-state monitoring of force-guided robotic assembly tasks using expectation maximization-based gaussian mixtures models. Int J Adv Manuf Technol 73(5):623–633. https://doi.org/10.1007/s00170-014-5803-x
26. Kallappa P (2000) Ray: Fuzzy wide-range control of fossil power plants for life extension and robust performance. Automatica 36:69–82
27. Kwofie S, Rahbar N (2012) A fatigue driving stress approach to damage and life prediction under variable amplitude loading. Int J Damage Mechan 22:393–404
28. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444
29. Lee DG, Jang KC, Kuk JM, Kim IS (2005) Comparison of the fatigue life of f.f. shaft material according to various environmental temperatures. Int J Adv Manuf Technol 26(7):896–908. https://doi.org/10.1007/s00170-003-2047-6
30. Liu C, Jiang D, Chen J (2014) Coupled torsional vibration and fatigue damage of turbine generator due to grid disturbance. J Eng Gas Turb Power 136(6):062,501–1–062,501–9
31. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. Int J Comput Vis 60(2):91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94
32. Manning C, Raghavan P, Schütze H (2008) Introduction to information retrieval. Cambridge University Press, Cambridge
33. Meggiolaro M, Castro J (2004) Statistical evaluation of strain-life fatigue crack initiation prediction. Int J Fatigue 26(5):463–476
34. Modarres C, Astorga N, Droguett EL, Meruane V (2018) Convolutional neural networks for automated damage recognition and damage type identification. Struct Control Health Monitor 25(10):e2230. https://doi.org/10.1002/stc.2230.E2230STC-17-0222.R3
35. Mohan A, Papageorgiou C, Poggio T (2001) Example-based object detection in images by components. IEEE Trans Pattern Anal Mach Intell 23(4):349–361. https://doi.org/10.1109/34.917571
36. Mohan A, Poobal S (2018) Crack detection using image processing: A critical review and analysis. Alexandria Eng J 57(2):787 – 798. https://doi.org/10.1016/j.aej.2017.01.020. http://www.sciencedirect.com/science/article/pii/S1110016817300236

37. Murphy K (2012) Machine learning: a probabilistic perspective, 1st edn. The MIT Press, Cambridge
38. Ni F, Zhang J, Chen Z (2019) Pixel-level crack delineation in images with convolutional feature fusion. Struct Control Health Monitor 26(1):e2286. https://doi.org/10.1002/stc.2286.E2286STC-18-0168.R1
39. Qu Z, Lin LD, Guo Y, Wang N (2015) An improved algorithm for image crack detection based on percolation model. IEEJ Trans Electr Electron Eng 10(2):214–221. https://doi.org/10.1002/tee.22056
40. Ray A (2004) Stochastic measure of fatigue crack damage for health monitoring of ductile alloy structures. Struct Health Monit 3(3):245–263
41. Ray A, Caplin J (2000) Life extending control of aircraft: Trade-off between flight performance and structural durability. Aeronaut J 104(1039):397–408
42. Saggar M, Bouraoui C, Nasr A (2018) Fatigue reliability prediction of defective materials based on a useful equivalent wöhler curve. Int J Adv Manuf Technol 97(1):1011–1021. https://doi.org/10.1007/s00170-018-1975-0
43. Saggar M, Sallem H, Bouraoui C (2018) Fatigue life prediction under variable loading based on a new damage model devoted for defective material. Int J Adv Manuf Technol 95(1):431–443. https://doi.org/10.1007/s00170-017-1198-9
44. Slima KB, Penazzi L, Mabru C, Ronde-Oustau F (2013) Fatigue analysis-based numerical design of stamping tools made of cast iron. Int J Adv Manuf Technol 67(5):1643–1650. https://doi.org/10.1007/s00170-012-4597-y
45. Suresh S (2004) Fatigue of Materials, 2nd edn. Cambridge University Press, Cambridge
46. Szeliski R (2011) Computer vision: algorithms and applications. Springer, London
47. Theodoridis S, Koutroumbas K (2008) Pattern recognition, 4th edn. Academic Press, USA
48. Wu B, Liu Z, Yuan Z, Sun G, Wu C (2017) Reducing overfitting in deep convolutional neural networks using redundancy regularizer. In: Lintas A, Rovetta S, Verschure PF, Villa AE (eds) Artificial neural networks and machine learning – ICANN 2017. Springer International Publishing, Cham, pp 49–55
49. Wu L, Hoi SCH, Yu N (2010) Semantics-preserving bag-of-words models and applications. IEEE Trans Image Process 19(7):1908–1920. https://doi.org/10.1109/TIP.2010.2045169
50. Yokoyama S, Matsumoto T (2017) Development of an automatic detector of cracks in concrete using machine learning. Procedia Eng 171:1250–1255. https://doi.org/10.1016/j.proeng.2017.01.418. http://www.sciencedirect.com/science/article/pii/S1877705817304289. The 3rd International Conference on Sustainable Civil Engineering Structures and Construction Materials - Sustainable Structures for Future Generations
51. Zhang Y, Jin R, Zhou ZH (2010) Understanding bag-of-words model: a statistical framework. Int J Mach Learn Cybern 1(1):43–52. https://doi.org/10.1007/s13042-010-0001-0
52. Zilberstein V, Walrath K, Grundy D, Schlicker D, Goldfine N, Abramovici E, Yentzer T (2003) MWM eddy-current arrays for crack initiation and growth monitoring. Int J Fatigue 25(9-11):1147–1155