Computational Materials Science 125 (2016) 136–145

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Prediction of fatigue stress concentration factor using extreme learning machine

Baoxian Wang [a,b], Weigang Zhao [a,b,*], Yanliang Du [a,b], Guangyuan Zhang [a], Yong Yang [b]

[a] Structure Health Monitoring and Control Institute, Shijiazhuang Tiedao University, Shijiazhuang 050043, China
[b] Key Laboratory for Health Monitoring and Control of Large Structures of Hebei Province, Shijiazhuang 050043, China

ARTICLE INFO

Article history:
Received 7 July 2016
Received in revised form 23 August 2016
Accepted 25 August 2016

Keywords:
Fatigue stress concentration factor
Material fatigue resistance
Extreme learning machine
Prediction

ABSTRACT

Fatigue stress concentration factor (FSCF) plays a vital role in studying the limitation of material fatigue resistance. Theoretically, FSCF not only reflects the level of fatigue stress concentration but also indicates the notch sensitive degree. In this work, a novel and efficient numerical model is presented for predicting FSCF, which exploits an emergent learning technique, i.e., Extreme Learning Machine (ELM). Specifically, we adopt seven parameters (i.e., tensile strength, yield strength, fatigue strength, theoretical stress concentration factor, notch root radius, samples size and notch fatigue limit) as the inputs, and the corresponding FSCF value is used as the output. With the randomly generated hidden neuron parameters, the ELM-based predictor can be fast trained. Besides, a pairwise metric constraint is introduced in the presented model, which can elevate the forecasting accuracy. A series of cross validation experiments demonstrate that the proposed FSCF predictor performs favorably against the existing empirical formulas and other learning based methods.

1. Introduction

Fatigue property is a considerable issue for material designing [1], and fatigue is also one of the chief culprits which result in the part failure of some equipments (such as large-scale bridge, astronautics devices, conveying machinery, etc.). The statistical results reveal that most of mechanical fracture accidents are caused by fatigue failure. Therefore, fatigue analysis plays an important role, especially in designing the equipment structure under alternating stress. As the limitation of material fatigue resistance, fatigue stress concentration factor (FSCF) not only manifests the degree of fatigue stress concentration but also reflects the notch sensitive level [2].

So far, most of the existing estimators for FSCF are some empirical formulas, which are based on the experimental research. In addition, these empirical methods often pay more attention to some specific conditions (such as material types, material strength limit, notch radius, etc.).

Neuber [3] considers that the notch fatigue limit is under the control of notch root radius, and thus, the FSCF can be computed as follows

$$k = 1 + \frac{K_t - 1}{1 + \sqrt{\alpha/\rho}} \tag{1}$$

here k is the estimated FSCF, K_t is the theoretical stress concentration factor, ρ is the size of notch root radius, and α is the Neuber length parameter.

Similarly, Peterson [4] holds the opinion that if the material fatigue failure happens, the corresponding surface stress must be larger than its fatigue limit. Hence, the FSCF value of material can be estimated as follows

$$k = 1 + \frac{K_t - 1}{1 + \beta/\rho} \tag{2}$$

here β is the Peterson length parameter.

In [5], the relationship between the FSCF value and the strain concentration factor has been further studied. According to their research, the materials can be divided into high-notch sensitive type and low-notch sensitive one. Therefore, the resultant FSCF predictor will have a piecewise expression.

Now it can be seen that the empirical formulas mentioned above are generally based on some priori hypotheses. However, in principle, the FSCF value is affected by the comprehensive factors of materials [2]. These empirical formulas may only explain the distribution of FSCF value in some cases, so they can not always perform well in the predicting task of FSCF.

* Corresponding author at: Structure Health Monitoring and Control Institute, Shijiazhuang Tiedao University, Shijiazhuang 050043, China.
E-mail address: zhaoweig2002@126.com (W. Zhao).

http://dx.doi.org/10.1016/j.commatsci.2016.08.035
0927-0256/

Recently, there emerge some learning based predictors [6–12]. Compared with the empirical formulas, the learning based predictors can fully utilize the hidden relationships between the feature representations of material, and finally find what is the best-fitting output with a certain amount of training examples.

Hou et al. [6] apply the support vector machine (SVM) into the prediction of FSCF. This SVM based predictor achieves better performance than Neuber [3] and Peterson [4] formulas, which validates the potentiality of machine learning methods. However, the training of SVM needs to solve a quadratic programming (QP) problem, which may result in high computation.

During the recent decade, artificial neural networks (ANN) has been widely used in the material design field [7,12,13]. In [7], Restrepo et al. advocated a back propagation neural network model to predict the grain boundary energy. Similarly, an implementation of ANN model in the fracture energy prediction for polymer nanocomposites has been presented in [8]. However, it is well known that the parameters tuning of ANN is a complex and time-consuming process, which may hinder its practical use in the material designing.

In this paper, we attempt to propose a new and efficient predicting model for FSCF, which takes advantage of the excellent learning capability of an emergent learning technique, i.e., Extreme Learning Machine (ELM). The ELM proposed by Huang et al. [14,15] is a novel machine learning framework. Compared with traditional learning methods (ANN or SVM), the hidden node parameters of ELM are randomly generated and it only computes the output weights among hidden layers and the output layer. Experiments demonstrate that ELM not only has a much faster learning speed but also obtains a better generalization in regression or classification applications [15]. Inspired by the success of ELM in related fields (e.g., fault diagnosis [16], feature learning [17], discharge forecasting [18], etc.), we apply it into the FSCF prediction, and try to achieve better forecasting in terms of accuracy and efficiency. The main contributions are summarized as follows.

(1) As far as we know, it is the first time that the ELM technique is utilized in the material designing, especially for the FSCF prediction. According to the ELM theories [14], ELM based predictor with random hidden parameters (with almost any nonlinear piecewise activation function) has the universal approximation capability. That is, unlike the ANN based FSCF predictors, where the hidden neuron parameters need to be iteratively fine-tuned, the proposed FSCF predicting method can provide a simple but effective estimating solution.
(2) A pairwise metric constraint is exploited in the ELM training. Specifically, if two kinds of materials have different FSCF values, they will exhibit differently in the ELM-based regression framework, and vice versa. Thus, the proposed predictor can utilize the mutual relationship within the two kinds of materials, and tends to achieve a better forecasting performance.
(3) The incremental updating equation of proposed FSCF predictor has been derived, which makes it very flexible to estimate the FSCF value of material with the new available training instances.

2. Preliminary knowledge

To facilitate the understanding of the proposed predicting model in the following sections, we briefly review the related theories/concepts of ELM. For a more detailed discussion and analysis, we refer the readers to [14,15,19–21]. It should be noted that the differences and relationships between ELM and other earlier works have been intensively analyzed in [22].

Huang et al. [14] originally proposed the ELM for generalized single hidden layer feed-forward neural networks (SLFNs), and recently extended it to the multi-layer case [23]. Suppose that SLFNs with L hidden nodes (see Fig. 1) are represented as

$$f_L(\mathbf{x}) = \sum_{j=1}^{L} G(\mathbf{w}_j, b_j, \mathbf{x})\beta_j = \sum_{j=1}^{L} h_j(\mathbf{x})\beta_j \tag{3}$$

where $\mathbf{w}_j$ is the input weight connecting the input layer to the j-th hidden node, and b_j is the bias of j-th hidden node; $G(\cdot)$ is the activation function; β_j is the output weight linking the j-th hidden node and the outputs; $h_j(\cdot)$ is the output vector of j-th hidden node.

Unlike the traditional understanding of neural networks, ELM theories [14] show that the hidden neurons need not to be adjusted. The corresponding implementation is random hidden neuron, and its parameters, i.e., $(\mathbf{w}, b)$ in the activation function $G(\mathbf{w}, b, \mathbf{x})$, are randomly generated based on a continuous probability distribution. In addition, Huang et al. further proved that ELM satisfies the universal approximation capability:

Theorem 1 (*Universal approximation capability [20]*). *Given any bounded nonconstant piecewise continuous function as the activation function, if the SLFNs can approximate any target function $f(\mathbf{x})$ via tuning the parameters of hidden neurons, the sequence $\{h_j(\mathbf{x})\}_{j=1}^{L}$ could be randomly generated based on any continuous sampling distribution, and with appropriate output weights, the $\lim_{L\to\infty}\left\|\sum_{j=1}^{L} h_j(\mathbf{x})\beta_j - f(\mathbf{x})\right\| = 0$ holds with probability one.*

The Eq. (3) can be rewritten as $f_L(\mathbf{x}) = \sum_{j=1}^{L} h_j(\mathbf{x})\beta_j = \mathbf{h}(\mathbf{x})\boldsymbol{\beta}$. Here, $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_L]^{\mathrm{T}}$ is the matrix of output weights, and $\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}), \ldots, h_L(\mathbf{x})]$ is the row vector representing the outputs of L hidden nodes with respect to the input $\mathbf{x}$. With the randomly generated hidden neuron parameters, $\mathbf{h}(\mathbf{x})$ is known to users. Thus, the ELM output function Eq. (3) becomes linear, and only the output weights $\boldsymbol{\beta}$ is unknown.

Given a training data set $\{\mathbf{X}, \mathbf{T}\} = \{\mathbf{x}^i, \mathbf{t}_i\}_{i=1}^{N}, \mathbf{x}^i \in \mathbb{R}^d$ is the i-th training data vector, and $\mathbf{t}_i \in \mathbb{R}^m$ represents the corresponding label. The above linear equations can be written in the matrix form

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \tag{4}$$

where $\mathbf{H}$ is the hidden layer output matrix (*randomized matrix*).

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}(\mathbf{x}^1) \\ \vdots \\ \mathbf{h}(\mathbf{x}^N) \end{bmatrix} = \begin{bmatrix} \mathbf{h}_1(\mathbf{x}^1) & \cdots & \mathbf{h}_L(\mathbf{x}^1) \\ \vdots & \ddots & \vdots \\ \mathbf{h}_1(\mathbf{x}^N) & \cdots & \mathbf{h}_L(\mathbf{x}^N) \end{bmatrix} \tag{5}$$

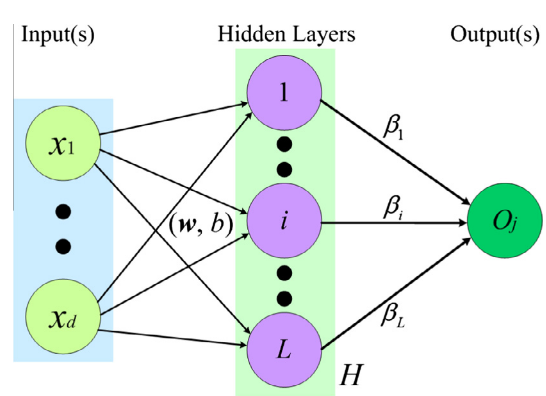

Fig. 1. Typical structure of an extreme learning machine framework.