Estimation of Natural Frequency of Micro Beams Using BPNN

Mohammad Heidari

Abstract**—**This paper presents an artificial neural network for estimation of natural frequency of micro beams. The dynamic problems of Euler-Bernoulli beams have been solved analytically on the basis of modified couple stress theory. In this study, back propagation neural network (BPNN) has been used for determining the natural frequency of micro beams. The BPN method is found to be the most accurate and quick. The networks have seven inputs, and the output is natural frequency of micro-beams. The output obtained from neural network models are compared with analytical results, and the amounts of relative errors have been calculated. The best results were obtained by the BPN by Levenberg-Marquardt algorithm training with 14 neurons in the one hidden layer. Training was continued until the mean squared error is less than 1e-5. Desired error value was achieved in the BPN and, the BPN was tested with both data used and not used for training. By training of this network, it is possible to estimate the natural frequency of beams at any conditions. It is shown that the BP neural network has the average errors of 2.46% in predicting natural frequency of micro beams.

Keywords**—** Natural Frequency, Micro-Beams, Neural network.

I. Introduction

MICRO-ELECTROMECHANICAL systems (MEMS) are widely being used in today's technology. So investigating the problems referring to MEMS, owns a great importance. Micro-beams are one of the major structures used widely in MEMS. Micro scale beams are widely used in microstructure devices and systems such as sensors [1-4] and actuators [5, 6], in which thickness of beams is typically on the order of microns and sub-microns. The size dependence of deformation behavior in micro scale beams had been experimentally observed in metals [7-9], polymers [10-12] and polysilicon [13]. The classical continuum mechanics theories are not capable of prediction and explanation of the size-dependent behaviors which occur in micron- and sub-micron-scale structures. However, some non-classical continuum theories such as higher-order gradient theories and the couple stress theory have been developed such that they are acceptably able to interpret the size-dependencies. In 1960s some researchers such as Koiter [14], Mindlin [15] and Toupin [16] introduced the couple stress elasticity theory as a non-classic theory capable to predict the size effects with appearance of two higher-order material constants in the corresponding constitutive equations. In this theory, beside the classical stress components acting on elements of materials, the couple stress components, as higher-order stresses, are also available which tend to rotate the elements. Utilizing the couple stress theory, some researchers investigated the size effects in some problems [17]. Employing the equilibrium equation of moments of couples beside the classical equilibrium equations of forces and moments of forces, a modified couple stress theory introduced by Yang, Chong, Lam, and Tong [18], with one higher-order material constant in the constitutive equations. Recently, size-dependent nonlinear Euler–Bernoulli and Timoshenko beams modeled on the basis of the modified couple stress theory have been developed by Xia et al. [19], and Asghari et al. [20], respectively. Park and Gao [21] were one of the first researchers to develop Euler–Bernoulli beam theories based on modified couple stress theory. Kong et al. [22] investigated the size-dependent natural frequency of Euler- Bernoulli beams; this study showed that for rectangular simply supported beams natural frequencies obtained by modified couple stress theory is 2.6 times greater than classical ones. Note that this result is valid when material length scale parameter and beam thickness are in the same order; however, when the ratio of thickness to higher-order material constant is 10, the calculated results by two different theories converge. The size effect of Bernoulli–Euler beam model has also been studied by Tsepoura et al. [23] and Papargyri-Beskou et al. [24]. Bernoulli–Euler flexural beam are dynamically analysed by analytic means on the basis of the simple theory of gradient elasticity due to Aifantis [25]. Free and forced flexural vibrations of simple beams are studied and the results show that natural frequencies for all natural modes are size dependence. Recently, the resonant frequency of a micro beam has been analysed theoretically using couple stress theory and the results show that the resonant frequency is size dependence [26]. The aim of this study is to estimation of natural frequency of Bernoulli–Euler micro beams using neural network models. The material is assumed to obey the modified couple stress theory [27]. The paper is structured as follows. The basic equations of modified couple stress theory are reviewed in Section 2. Then the displacement field typical for a Bernoulli–Euler beam, the governing equation, initial conditions and boundary conditions for the Bernoulli–Euler beams are established in Section 3. Section 4 recalls the artificial neural network. Section 5 proposes network development. Simulation results and discussion of the problem are given in Section 6. The paper gives a conclusion in Section 7.

Mohammad Heidari is with Abadan Branch, Islamic Azad University, Abadan, Iran (+989131835779; moh104337@yahoo.com).

II. Preliminaries

In the modified couple stress theory, the strain energy density $\bar{u}$ for a linear elastic isotropic material in infinitesimal deformation is written as [27]:

$$\bar{u} = \frac{1}{2}\left(\sigma_{ij}\varepsilon_{ij} + m_{ij}\chi_{ij}\right) \; i,j = 1,2,3 \quad (1)$$

Where

$$\sigma_{ij} = \lambda\varepsilon_{mm}\delta_{ij} + 2\mu\varepsilon_{ij} \quad (2)$$

$$\varepsilon_{ij} = \frac{1}{2}((\nabla u)_{ij} + (\nabla u)_{ij}^{T}) \quad (3)$$

$$m_{ij} = 2l^2\mu\chi_{ij} \quad (4)$$

$$\chi_{ij} = \frac{1}{2}((\nabla\theta)_{ij} + (\nabla\theta)_{ij}^{T}) \quad (5)$$

In which σ_{ij}, ε_{ij}, $m_{ij}$and χ_{ij} denote the components of the symmetric part of stress tensor σ, the strain tensor ε, the deviatoric part of the couple stress tensor *m* and the symmetric part of the curvature tensor χ, respectively. Also, *u* and θ are the displacement vector and the rotation vector. The two Lame constants and the material length scale parameter are represented by λ, μ and *l*, respectively. The Lame constants are written in terms of the Young's modulus *E* and the Poisson's ratio ν as $\lambda = \nu E/(1+\nu)(1-2\nu)$ and $\mu = E/2(1+\nu)$.

The components of the infinitesimal rotation vector θ_i are related to the components of the displacement vector field u_i as [28]:

$$\theta_i = \frac{1}{2}(curl(u))_i \quad (6)$$

III. Governing Equation of Motion

The axes for plane beam analysis are established, as shown in figure 1. The total length of the beam is L. According to the basic hypotheses of Bernoulli–Euler beams and the one-dimensional beam theory, the displacement field can be written as [29]:

$$u = -z\psi(x,t) \quad v = 0 \quad w = w(x,t) \quad (7)$$

Where u, v, w are the x-, y- and z- components of the displacement vector, and ψ(x) is the rotation angle of the centroidal axis of the beam given approximately by

$$\psi(x) \approx \frac{\partial w(x,t)}{\partial x} \quad (8)$$

for small deformation considered here.

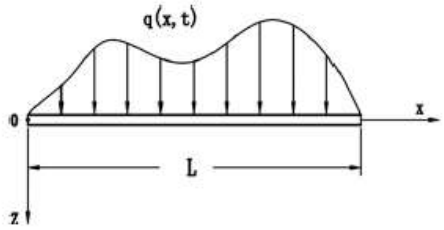

Fig. 1 Coordinate system for Euler –Bernoulli beam

The strain energy of the beams is:

$$U = \frac{1}{2}\int_0^L (EI + \mu Al^2)\left(\frac{\partial^2 w}{\partial x^2}\right)^2 \, dx \quad (9)$$

Where *I* is the usual second moment of cross-sectional area, A is the cross-sectional area of the beam, μ is the shear modulus and *E* is Young's modulus of the beam material. The additional parameter *l* is regarded as a material property characterizing the effect of couple stress. The work done by the external forces in the form of transverse loading *q(x, t)*, as shown in figure 1, reads

$$V = \int_0^L q(x,t)w(x,t)dx \quad (10)$$

And the kinetic energy can be written as

$$T = \frac{1}{2}\int_0^L \rho(x)A(x)\left(\frac{\partial w(x,t)}{\partial t}\right)^2 dx \quad (11)$$

Where *ρ(x)* is the density of the beam material. The dynamic governing equation of this beam as well as all possible boundary conditions can be determined with the aid of the Hamilton's principle [24]

$$\delta\{\int_{t_1}^{t_2}(T - V - U)dt\} = 0 \quad (12)$$

In view of the Eq. (12), the dynamic governing equation of the beam in terms of w(x, t) is given by

$$\rho A\ddot{w} + (EI + \mu Al^2)w^{(4)} = q(x,t) \quad (13)$$

Where

$$w^{(4)} = \frac{\partial^4 w}{\partial x^4}, \quad \ddot{w} = \frac{\partial^2 w}{\partial t^2}$$

For a simply supported (S.S) beam, in which *L* is the length of beam, the boundary conditions of this problem are

$$w(0,t) = 0 \quad \frac{\partial^2 w(0,t)}{\partial x^2} = 0 \quad (14)$$

$$w(L,t) = 0 \quad \frac{\partial^2 w(L,t)}{\partial x^2} = 0 \quad (15)$$

Then

$$\omega = \sqrt{\frac{EI+\mu Al^2}{\rho AL^4}}\,(n\pi)^2 \quad (16)$$

Where *ω* is the natural frequency. By letting *l = 0*, the new model reduce to the classical beam model.

For a cantilever beam or fix-free (F.F), in which *L* is the length of beam, the boundary conditions of this problem are [29]

$$w(0,t) = 0 \quad \frac{\partial w(0,t)}{\partial x} = 0 \quad (17)$$

$$\frac{\partial^2 w(L,t)}{\partial x^2} = 0 \quad \frac{\partial^3 w(L,t)}{\partial x^3} = 0 \quad (18)$$

Then yields

$$\omega = (s_n^2)\sqrt{\frac{EI+\mu Al^2}{\rho A}} \quad (19)$$

Where

$$s_n L = 1.875, 4.694, 7.855, 10.996, 14.137, \ldots$$
$$n = (1, 2, 3, 4, 5, \ldots) \quad (20)$$

IV. Neural Networks

Neural networks are artificial intelligence algorithms for cognitive tasks, such as learning and optimisation. Neural networks are of interest because of their ability to learn, to make decisions, and to draw conclusions from examples without knowledge of the underlying rules. The motivation for neural networks came from attempts to simulate the processes of the human brain, and so to enhance the capabilities of computers. Neural networks are applicable in those situations in which a relationship between the predictor variables (independents, inputs) and predicted variables (dependents, outputs) exists, even when that relationship is very complex and not easy to articulate in the usual terms of "correlations" or "differences between groups". Commonly neural networks are adjusted, or trained, so that a particular input leads to a specific target output. Such a situation is shown in figure 2, where, the network is adjusted, based on a comparison of the output and the target, until the network output matches the

target. Typically many such input/target pairs are used, in this supervised learning, to train a network [30].

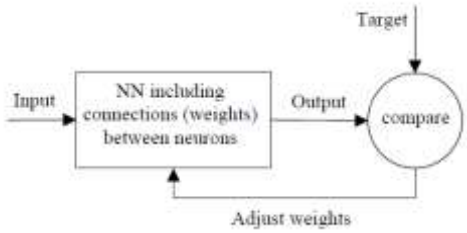

Fig. 2 Processing of NN

A neural network is a non-linear system consisting of a large number of highly interconnected processing units, nodes or artificial neurons. Each input signal is multiplied by the associated weight value and summed at a neuron. The result is put through an activation function to generate a level of activity for the neuron. This activity is the output of the neuron. When the weight value at each link and the correction pattern are determined, the neural network is trained. The process is accomplished by learning from the training set and by applying certain learning rule. The trained network can be used to generalize for those inputs that are not included in the training set. The main advantages of neural networks are as follows:

1. They can learn and generalize by being trained from a series of examples, without knowledge of underlying rules, to produce meaningful solutions to problems.

2. Data presented for training neural networks can be theoretical data, experimental data, empirical data or a combination of these [31].

A typical neural network is arranged in three (or more) layers of nodes as figure 3 shows:

1. An input layer (receiving the input values),

2. An output layer (presenting the output values),

3. One or more hidden layers between the input and output layers.

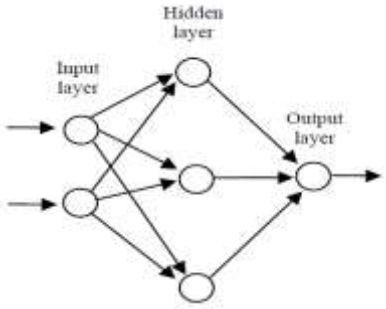

Fig. 3 A typical neural network

Mainly, there are three types of training condition for NNs; namely supervised training, graded training and self-organization training. Supervised training, which is adopted in this study, can be applied as:

(1) First, the dataset of the system, including input and output values, is established;

(2) The dataset is normalized according to the algorithm;

(3) Then, the algorithm is run;

(4) Finally, the desired output values corresponding to the input used in test phase.

V. Back Propagation Neural Network

Back propagation neural network (BPN), developed by Rumelhart [32], is the most prevalent of the supervised learning models of ANN. BPN used the gradient steepest descent method to correct the weight of the interconnectivity neuron. BPN easily solved the interaction of processing elements by adding hidden layers. In the learning process BPN, the interconnective the weights are adjusted using an error convergence technique to obtain a desired output for a given input. In general, the error at the output layer in the BPN model propagates backward to the input layer through the hidden layer in the network to obtain the final desired output. The gradient descent method is utilized to calculate the weight of the network and adjusts the weight of interconnectives to minimize the output error. The formulas used in this algorithm are as follows:

1) Hidden layer calculation results:

$$net_i = \sum x_i w_i \quad (21)$$

$$y_i = f(net_i) \quad (22)$$

Where x_i and $w_i$are input data and weights of the input data, respectively. The *f* is activation function, and y_i is result obtained from hidden layer. Output layer calculation results:

$$net_k = \sum y_i w_{jk} \quad (23)$$

$$o_k = f(net_k) \quad (24)$$

Where w_{jk} are weights of output layer, and o_k is result obtained from output layer. Activation functions used in layers are logsig, tansig and linear:

$$f(net_i) = \frac{1}{1+e^{-net_i}} \text{ (logsig)} \quad (25)$$

$$f(net_i) = \frac{1-e^{-net_i}}{1+e^{-net_i}} \text{ (tansig)} \quad (26)$$

$$f(net_i) = net_i \text{ (linear)} \quad (27)$$

Errors made at the end of one cycle:

$$e_k = (t_k - o_k)o_k(1 - o_k) \quad (28)$$

$$e_i = y_i(1 - y_i)\sum e_k w_{ij} \quad (29)$$

Where t_k is result expected from output layer, $e_k$is error occurred at output layer, and e_i is error occurred at hidden layer. Weights can be changed using these calculated error values according to Eqs. (30) and (31).

$$w_{jk} = w_{jk} + \alpha e_k y_i + \beta \Delta w_{jk} \quad (30)$$

$$w_{ij} = w_{ij} + \alpha e_i x_i + \beta \Delta w_{ij} \quad (31)$$

Where w_{ij} are weights of output layer. Δw_{jk} and Δw_{ij} are correction made in weights at the previous calculation. Parameter α is learning ratio, and β is momentum term, that is used to adjust weights. In this paper, α=0.65 and β=0.9, are used. Square error, occurred in one cycle, can be found by Eq. (32).

$$e = \sum 0.5|t_k - o_k|^2 \quad (32)$$

The completion of training the BPN, relative error (RE) for each data and mean relative error (MRE) for all data are calculated according to Eqs. (33) and (34), respectively.

$$RE = \left(\frac{100(t_k - o_k)}{t_k}\right) \quad (33)$$

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{100(t_k - o_k)}{t_k}\right) \quad (34)$$

Where *n* is the number of data [33].

VI. Network Development

A. Input and output data

The cross section area of beam, length of beam, shear modulus, moment of inertia, density of beam material, young modulus and scale parameter are the input of the network. Finally, the outputs of net are the natural parameter of the micro beam.

B. Network configuration

The nodes at the input and output layer are determined by the number of predictor and predicted variables. In this research there are 7 nodes in the input layers due to the number of input variables, and 1 node in the output layer, for similar reasons. There are no rules given to determine the exact number of hidden layers and the number of nodes in hidden layers. A large number of hidden-layer nodes will lead to an over-fit at intermediate points, which can slow down the operation of NN. On the other hand an accurate output may not be achieved if too few hidden layer nodes are included in the neural network. The results show that the best configuration of the network is achieved by one hidden layer. The number of nodes in the input layer, in the hidden layer and in the output layer is chosen to 7-14-1, respectively. The activation function in the input and the hidden layers is sigmoid function and linear function in the output layer.

C. Pre-processing the data

For a proper working of the neural network a pre-processing of the input and output data is performed. The input values are normalized between -1 and 1, since the activation function is a sigmoid function in the input layer. Normalization is made by the following function:

$$x_{norm} = 2.\frac{x - x_{min}}{x_{max} - x_{min}} - 1 \quad (35)$$

The output values are normalized between 0 and 1 and a linear function in the output layer [34].

D. Training of the network

Once a network is structured for a particular application, that network is ready to be trained. To start this process the initial weights are chosen randomly. During the training the weights are iteratively adjusted to minimize the network performance function. As performance function the mean square error–the average squared error between the network output and the target output is applied. For the training of the network the Matlab Neural Network Toolbox is used [35]. The Levenberg-Marquardt algorithm is chosen to perform the training with the default values suggested in. In this work, for training is used of three functions, *newelm*, *newff* and *newcf*. The stopping criteria are adjusted, that the mean square error should be less then 10^{-5} and the number of epochs (iterations) should be less then 5000. The BPN learning process involves a forward propagation pass calculating the outputs using the inputs, weights and neuron transfer functions, as well as a back propagation pass correcting the weights using the error between the predicted and target values. The major advantage of the BPN model is its ability to learn from examples without requiring principal knowledge of domain problems. In addition, it is very effective in dealing with large amounts of data. The structure of the BPN model can easily be constructed according to the domain problem and the availability of data attributes.

VII. Results

In this study, the back propagation learning algorithm is used in a feed forward, single hidden layer network. A variable transfer function is used as the activation function for both the hidden layer and the output layer. Many back propagation training algorithms were repeatedly applied until satisfactory training was achieved. The names of training algorithms are shown in table 1. The activation function for the hidden layer and the output layer that is used are shown in table 2.

TABLE I
THE VARIABLE TRAINING METHODS [35]

Acronym	Description
LM	Levenberg-Marquardt
BFG	BFGS Quasi-Newton
RP	Resilient Back propagation
SCG	Scaled Conjugate Gradient
CGB	Conjugate Gradient with Powell/Beale Restarts
CGF	Fletcher-Powell Conjugate Gradient
CGP	Polak-Ribiére Conjugate Gradient
OSS	One Step Secant
GDX	Variable Learning Rate Back propagation

TABLE II
THE VARIABLE ACTIVATION FUNCTIONS IN THE LAYERS

No	The activation function	
	Hidden layer	Output layer
1	logsig	logsig
2	logsig	tansig
3	logsig	purelin
4	tansig	tansig
5	tansig	logsig
6	tansig	purelin

The best combination for all methods that is used in this paper is logsig for hidden layer and purelin for output layer. In the hidden layer, a number of neurons from 5 to 20 are used. The data set available for natural frequency included 60 data patterns. From these, 40 data patterns were used for training the network, and the remaining 20 patterns were randomly selected and used as the test data set. The regression value (R^2) of the output variable values for the test data set for various neurons in hidden layer is shown in table 3. It should be noted that these data were completely unknown to the network. The closer this value is to unity the better is the prediction accuracy. The best (R^2) value obtained is 0.9999, and it is obtained from the *LM* algorithm by 14 neurons in hidden layer.

TABLE III
THE (R^2) VALUES FOR NATURAL FREQUENCY OF MICRO- BEAMS WITH VARIOUS NEURONS IN THE HIDDEN LAYER

Number of hidden neuron	Acronym of training method								
	LM	BFG	RP	SCG	CGB	CGF	CGP	OSS	GDX
5	0.9977	0.9888	0.9819	0.9957	0.9738	0.9918	0.9905	0.9913	0.9891
6	0.9981	0.9676	0.992	0.9909	0.9942	0.9864	0.9846	0.9969	0.9916
8	0.9999	0.9981	0.9942	0.9578	0.9986	0.9978	0.9681	0.9963	0.9945
10	0.9985	0.9998	0.9967	0.9996	0.9996	0.9995	0.9595	0.9890	0.9856
12	0.9967	0.9928	0.9994	0.9998	0.9998	0.9698	0.9898	0.9898	0.9881
14	0.9999	0.9899	0.9992	0.9699	0.9997	0.9698	0.9699	0.9797	0.9950
16	0.9996	0.9699	0.9994	0.9699	0.9699	0.9599	0.9899	0.9479	0.9983
18	0.9944	0.9916	0.9996	0.9797	0.9694	0.9792	0.9799	0.9688	0.9782
20	0.9989	0.9599	0.9995	0.9899	0.9899	0.9899	0.9799	0.9896	0.9870

In tables 4, 5 and 6 the results of training the network using nine different training algorithms by 14 neurons in the hidden layer and logsig-purlin activation function are summarized.

TABLE IV
THE RESULTS OF THE VARIABLE TRAINING METHODS IN THE BPN WITH NEWELM FUNCTION

Acronym	Epoch in goal	Error Goal	Train time (s)	Test time (s)
LM	1190	met	22.724082	0.054526
BFG	1528	met	47.482451	0.046583
RP	3598	Not met	57.677318	0.046466
SCG	1902	Not met	88.328176	0.044047
CGB	3210	Not met	108.102075	0.052230
CGF	1495	Not met	98.894782	0.046125
CGP	3190	Not met	99.665335	0.046852
OSS	49000	Not met	44.071065	0.046677
GDX	2800	Not met	78.243551	0.046046

TABLE V
THE RESULTS OF THE VARIABLE TRAINING METHODS IN THE BPN WITH NEWCF FUNCTION

Acronym	Epoch in goal	Error Goal	Train time (s)	Test time (s)
LM	1925	met	28.726182	0.052241
BFG	2458	met	65.481454	0.041678
RP	2079	Not met	69.16278	0.043146
SCG	2180	Not met	43.19076	0.052047
CGB	2932	Not met	89.15905	0.065130
CGF	3415	Not met	128.15282	0.046985
CGP	3780	met	55.150335	0.034152
OSS	2900	met	53.011265	0.048677
GDX	2960	met	41.120551	0.058946

TABLE VI
THE RESULTS OF THE VARIABLE TRAINING METHODS IN THE BPN WITH NEWFF FUNCTION

Acronym	Epoch in goal	Error Goal	Train time (s)	Test time (s)
LM	2293	met	31.445082	0.054789
BFG	2700	met	67.402451	0.056673
RP	3090	met	78.277318	0.056906
SCG	3940	Not met	92.749761	0.065237
CGB	3912	Not met	88.48905	0.062230
CGF	2450	met	76.69642	0.053565
CGP	3162	Not met	109.23335	0.056907
OSS	3040	Not met	99.41895	0.066677
GDX	2306	met	59.446511	0.056044

The fastest algorithm for this problem is the LM. On the average, it is over two times faster than the next fastest algorithm. This is the type of problem for which the LM algorithm is best suited. In tables 7, 8 and 9 a comparison between the ratios of actual natural frequency ($\frac{\omega}{\omega_0}$) and predicted with the artificial neural network for the LM method are presented. Where ω_0 is the natural frequency in the classical beam model. As can be seen, the error with newelm function is very small.

TABLE VII
COMPARISON BETWEEN ACTUAL NATURAL FREQUENCY AND THE BPN MODEL (WITH NEWELM FUNCTION)

No.	($\frac{\omega}{\omega_0}$) Ratio of natural frequency (S.S)			($\frac{\omega}{\omega_0}$) Ratio of natural frequency (F.F)		
	Actual	Predicted	%Error	Actual	Predicted	% Error
1	1	0.98	1.37	2.82	2.78	1.21
2	1.52	1.48	2.35	1.09	1.03	5.1
3	3.62	3.57	1.15	5.67	5.51	2.67
4	1.13	1.10	2.36	3.42	3.30	3.41
5	4.47	4.20	5.93	5.17	5.03	2.64
6	5.26	4.96	5.57	2.39	2.36	1.12
7	3.48	3.46	0.33	1.28	1.27	0.54
8	5.50	5.38	2.12	2.06	2.03	1.34
9	2.59	2.52	2.33	3.11	2.95	5.12
10	2.17	2.14	1.18	5.38	5.29	1.52
Average Error	2.46%					

TABLE VIII
COMPARISON BETWEEN ACTUAL NATURAL FREQUENCY AND THE BPN MODEL (WITH NEWCF FUNCTION)

No.	($\frac{\omega}{\omega_0}$) Ratio of natural frequency (S.S)			($\frac{\omega}{\omega_0}$) Ratio of natural frequency (F.F)		
	Actual	Predicted	%Error	Actual	Predicted	% Error
1	1	0.87	12.56	2.82	2.67	5.12
2	1.52	1.43	5.78	1.09	1.00	8.22
3	3.62	3.25	10.19	5.67	5.13	9.45
4	1.13	3.35	7.27	3.42	3.32	2.66
5	4.47	4.10	8.27	5.17	4.97	3.73
6	5.26	5.00	4.82	2.39	2.14	10.12
7	3.48	3.33	4.16	1.28	1.16	9.19
8	5.50	5.16	6.09	2.06	1.91	6.8
9	2.59	2.42	6.47	3.11	2.92	6.08
10	2.17	2.02	4.49	5.38	4.90	8.75
Average Error	7.01%					

TABLE VIIII
COMPARISON BETWEEN NATURAL FREQUENCY AND THE BPN MODEL (WITH NEWFF FUNCTION)

No.	($\frac{\omega}{\omega_0}$) Ratio of natural frequency (S.S)			($\frac{\omega}{\omega_0}$) Ratio of natural frequency (F.F)		
	Actual	Predicted	%Error	Actual	Predicted	% Error
1	1	0.89	10.98	2.82	2.52	10.42
2	1.52	1.36	9.89	1.09	0.95	12.05
3	3.62	3.28	9.25	5.67	5.12	9.57
4	1.13	1.00	11.16	3.42	3.08	9.96
5	4.47	4.05	9.25	5.17	4.73	8.46
6	5.26	4.83	8.11	2.39	2.14	10.44
7	3.48	3.19	8.08	1.28	1.17	8.51
8	5.50	4.91	10.67	2.06	1.82	11.38
9	2.59	2.29	11.31	3.11	2.77	10.69
10	2.17	1.89	12.46	5.38	4.86	9.61
Average Error	10.11%					

Figures 4 and 5 show the compression between size effect on the natural frequencies with respect to various sizes for simply supported and cantilever micro-beams by two methods respectively.

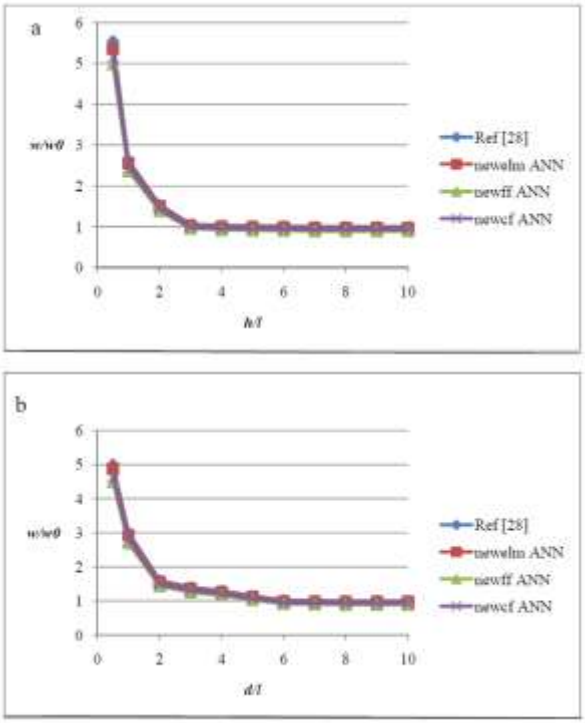

Fig. 4 Size effect on the natural frequencies with respect to various sizes for simply supported beam (a) rectangular beam with thickness *h* and (b) circular beam with diameter *d* by three ANN models.

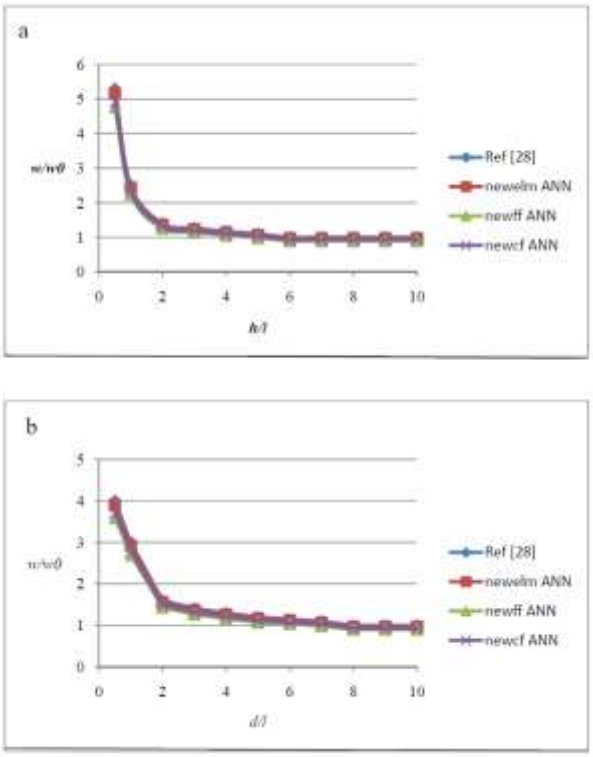

Fig. 5 Size effect on natural frequency with respect to various sizes for cantilever beam: (a) rectangular beam with thickness *h* and (b) circular beam with diameter *d* by three ANN models.

VIII. CONCLUSIONS

The present study shows that for the analyses of natural frequency of micro beams, the BPN is a suitable method. Date for developing the ANN model is obtained by analytical formula. Results from ANN model are compared with the results from the classical model. The best regression value for the simulation is 0.9999 with newelm function. The MRE value of the BPN model is 2.46%. The results show that newelm function is more accurate than newff and newcf functions. Also the Levenberg-Marquardt training is faster than other training methods. The BPN method also has the advantages of computational speed, low cost and ease of use by people with little technical experience. It is found that the natural frequencies of the beams predicted by the newly model are larger than that predicted by the classical beam model. These conclusions can be explained as follows: the intrinsic size dependence of materials decrease the static beam deflections and hence increase the stiffness leading to increased values of natural frequencies. The difference between the natural frequencies predicted by the two models is very significant when the characteristic sizes such as thickness and diameter are small, but is diminishing with the increase of the characteristic sizes.

REFERENCES

[1] J. Pei, F. Tian, T. Thundat, Glucose biosensor based on the micro cantilever. Analytical Chemistry, vol. 76, pp. 292–297, 2004.
[2] N.A. Hall, M. Okandan, F.L. Degertekin, Surface and bulk-silicon-micro- machined optical displacement sensor fabricated with the SwIFT-Lite process. Journal of Microelectromechanical Systems vol. 15 (4), pp. 770–776, 2006.
[3] W. Faris, A.H. Nayfeh, Mechanical response of a capacitive microsensor under thermal load. Communications in Nonlinear Science and Numerical Simulation vol. 12 (5), pp. 776–783, 2007.
[4] Y. Moser, M.A.M. Gijs, Miniaturized flexible temperature sensor. Journal of Microelectromechanical Systems vol. 16 (6), pp. 1349–1354, 2007.
[5] E.S. Hung, S.D. Senturia, Extending the travel range of analog-tuned elec- trostatic actuators. Journal of Microelectromechanical Systems vol. 8 (4), pp. 497–505, 1999.
[6] M.P. de Boer, D.L. Luck, W.R. Ashurst, R. Maboudian, A.D. Corwin, J.A. Walraven, J.M. Redmond, High-performance Surface-micromachined Inchworm Actuator. Journal of Microelectromechanical Systems, vol.13, pp. 63-74, 2004.
[7] W.D. Nix, Mechanical properties of thin films. Metallurgical and Materials Transactions vol. A 20, pp. 2217–2245, 1989.
[8] N.A. Fleck, G.M. Muller, M.F. Ashby, J.W. Hutchinson, Strain gradient plasticity: theory and experiment. Acta Metallurgica Et Materialia vol. 42, pp. 475–487, 1994.
[9] W.J. Poole, M.F. Ashby, N.A. Fleck, Micro-hardness of annealed and work- hardened copper polycrystals. Scripta Materialia, vol. 34 (4), pp. 559–564, 1996.
[10] D.C.C. Lam, A.C.M. Chong, Indentation model and strain gradient plasticity law for glassy polymers. Journal of Materials Research, vol. 14, pp. 3784–3788, 1999.
[11] D.C.C. Lam, F. Yang, A.C.M. Chong, J. Wang, P. Tong, Experiments and theory in strain gradient elasticity. Journal of Mechanics and Physics of Solids, vol. 51 (8), pp. 1477–1508, 2003.
[12] A.W. McFarland, J.S. Colton, Role of material microstructure in plate stiffness with relevance to micro cantilever sensors. Journal of Micromechanics and Microengineering vol. 15, pp.1060–1067, 2005.
[13] I. Chasiotis, W.G. Knauss, The mechanical strength of polysilicon films: Part 2. Size effects associated with elliptical and circular perforations. Journal of the Mechanics and Physics of Solids, vol. 51, pp.1551–1572, 2003.

[14] W.T. Koiter, Couple-stresses in the theory of elasticity: I and II. Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, Series B, pp. 6717-6744, 1964.
[15] R.D. Mindlin, H.F. Tiersten, Effects of couple-stresses in linear elasticity. Archive for Rational Mechanics and Analysis, vol. 11(1), pp. 415-448, 1962.
[16] R.A. Toupin, Elastic materials with couple-stresses. Archive for Rational Mechanics and Analysis, vol. 11(1), pp.385–414, 1962.
[17] A. Anthoine, 2000. Effect of couple-stresses on the elastic bending of beams. International Journal of Solids and Structures, vol. 37, pp. 1003-1018, 2000.
[18] F. Yang, A.C.M. Chong, D.C.C. Lam, P. Tong, Couple stress based strain gradient theory for elasticity. International Journal of Solids and Structures, vol. 39(10), pp. 2731-2743, 2002.
[19] W. Xia, L. Wang, L., Yin, Nonlinear non-classical microscale beams: Static bending, post buckling and free vibration. International Journal of Engineering Science, vol. 48, pp. 2044-2053, 2010.
[20] M. Asghari, M. Rahaeifard, M.H. Kahrobaiyan, M.T. Ahmadian, On the size-dependent behavior of functionally graded micro-beams. Materials and Design, vol. 32, pp. 1435-1443, 2011.
[21] S.K. Park, X.L. Gao, Bernoulli–Euler beam model based on a modified couple stress theory, J. Micromech. Microeng., vol. 16 (11), doi:10.1088/0960-1317/16/11/015, 2006.
[22] S.L. Kong, S.J. Zhou, Z.F. Nie, K. Wang, The size-dependent natural frequency of Bernoulli–Euler micro-beams. Int. J. Eng. Sci., vol. 46 (5) pp. 427–437, 2008.
[23] K. G. Tsepoura, S. Papargyri-Beskou, D. Polyzos, D. E. Beskos, Static and dynamic analysis of a gradient-elastic bar in tension. Archive of Applied Mechanics, vol. 72, pp. 483–497, 2002.
[24] D. E. Beskos, S. Papargyri-Beskou, D. Polyzos, K. G. Tsepoura, Bending and stability analysis of gradient elastic beams. International Journal of Solids and Structures, vol. 40 (2), pp. 385–400, 2003.
[25] S. Papargyri - Beskou, D. Polyzos, D.E. Beskos, Dynamic analysis of gradient elastic flexural beams. Structural Engineering and Mechanics, vol. 15 (6), pp. 705–716, 2003.
[26] X. Kang, Z.W. Xi, Size effect on the dynamic characteristic of a micro beam based on cosserat theory. Journal of Mechanical Strength, vol. 29 (1), pp.1–4, 2007.
[27] F. Yang, A.C.M. Chong, D.C.C. Lam, P. Tong, Couple stress based strain gradient theory for elasticity. International Journal of Solids and Structures, vol. 39 (10), pp. 2731–2743, 2002.
[28] K. Shengli, Z. Shenjie, N. Zhifeng, W. Kai, The size-dependent natural frequency of Bernoulli–Euler micro-beams. Journal of Engineering Science, vol. 46, pp.427-437, 2008.
[29] S. Kong, S. Zhou, Z. Nie, K. Wang, The size-dependent natural frequency of Bernoulli–Euler micro-beams. International Journal of Engineering Science, (46) 427–437, 2008.
[30] H. Shao, G. Zheng, Convergence analysis of a back-propagation algorithm with adaptive momentum. Neurocomputing, vol. 74(5), pp.749–752, 2011.
[31] D. Gao, Y. Kinouchi, K. Ito, Z. Zhao, Neural networks for event extraction from time series: a back propagation algorithm approach. Future Generation Computer Systems, vol. 21(7), pp.1096–1105, 2005.
[32] D. E. Rumelhart, G. E. Hinton, R. J. Williams, Learning representations by back-propagating errors, Nature, 323(6088) pp.533–536, 1986.
[33] M. Heidari, H. Homaei, Design a PID controller for suspension system by back propagation neural network, Journal of Engineering, vol. 2013, Article ID 421543, 9 pages, 2013. doi:10.1155/2013/421543.
[34] M. Heidari, H. Homaei, Estimation of Acceleration Amplitude of Vehicle by Back Propagation Neural Networks, Advances in Acoustics and Vibration, vol. 2013, Article ID 614025, 7 pages, doi:10.1155/2013/614025, 2013.
[35] H. Demuth and M. Beale, Matlab Neural Networks Toolbox, User's Guide, Copyright 1992–2001, The Math Works, Inc, http://www.mathworks.com.