Genetic Algorithm and ANN for Estimation of SPIV of Micro Beams

M. Heidari *

Department of Mechanical Engineering,
Aligudarz Branch, Islamic Azad University, Aligudarz, Iran
E-mail: m.heidari@iau-aligudarz.ac.ir
*Corresponding author

Abstract: In this paper, the static pull-in instability (SPIV) of beam-type micro-electromechanical systems is theoretically investigated. Herein, modified strain gradient theory in conjunction with Euler–Bernoulli beam theory have been used for mathematical modeling of the size dependent instability of the micro beams. Considering the mid-plane stretching as the source of the nonlinearity in the beam behavior, a nonlinear size-dependent Euler-Bernoulli beam model is used based on a modified couple stress theory, capable of capturing the size effect. Two common beam-type systems including double-clamped and clamped-free cantilever have been investigated. By selecting a range of geometric parameters such as beam lengths, width, thickness, gaps and size effect, we identify the static pull-in instability voltage. Back propagation artificial neural network (ANN) with three functions have been used for modelling the static pull-in instability voltage of micro beam. Effect of the size dependency on the pull-in performance has been discussed for both micro-structures. The network has four inputs of length, width, gap and the ratio of height to scale parameter of beam as the independent process variables, and the output is static pull-in voltage of microbeam. The number of nodes in the hidden layer, learning ratio and momentum term are optimized using genetic algorithms (GAs). Numerical data, employed for training the network and capabilities of the model in predicting the pull-in instability behaviour has been verified. The output obtained from neural network model is compared with numerical results, and the amount of relative error has been calculated. Based on this verification error, it is shown that the back propagation neural network has the average error of 6.36% in predicting pull-in voltage of cantilever micro-beam. Resultant low relative error value of the ANN model indicates the usability of the BPN in this area.

Keywords: Artificial neural networks, Euler-Bernoulli, Genetic algorithms, Nonlinear micro-beam, Modified couple stress theory, Static pull-in instability

Biographical notes: M. Heidari received his MSc in Mechanical Engineering from University of Shahid Chamran, in 2003. He is currently Assistant Professor at the Department of Mechanical Engineering, Aligudarz Branch, Islamic Azad University, Aligudarz, Iran.

1   INTRODUCTION

Emerging revolution of nanotechnology gives the opportunity to develop high performance precise ultra-small systems for engineering applications. Recently, micro/nano electromechanical systems (MEMS/NEMS) has found enormous applications in many science branches e.g. engineering, chemistry, optic, magnetic, electronics, etc. A beam-type MEMS/NEMS constructed from two conductive electrodes which one electrode is movable and the other one is fix (ground). Applying voltage between the electrodes leads to deformation of the movable electrode toward fixed electrode. When electrostatic force exceeds the elastic resistance of the beam, the instability occurs and movable electrode suddenly adheres to the ground. Instability characteristics of MEMS in micro-scales has been investigated by previous researchers during previous decade [1].

Parametric excitation occurs in a wide range of mechanics, due to time dependent excitations, especially periodic ones; some examples are columns made of nonlinear elastic material, beams with a harmonically variable length, parametrically excited pendulums and so forth. Investigating stability analysis on parametrically excited MEM systems is of great importance. In 1995 Gasparini et al. [2] studied on the transition between the stability and instability of a cantilevered beam exposed to a partially follower load. Applying voltage difference between an electrode and ground causes the electrode to deflect towards the ground. At a critical voltage, which is known as pull-in voltage, the electrode becomes unstable and pulls-in onto the substrate. The pull-in behavior of MEMS actuators has been studied for over two decades without considering the casimir force [3-5].

Osterberg et al. [3], [4] investigated the pull-in parameters of the beam-type and circular MEMS actuators using the distributed parameter models. Sadeghian et al. [5] applied the generalized differential quadrature method to investigate the pull-in phenomena of micro-switches. A comprehensive literature review on investigating MEMS actuators can be found in Ref. [6]. Further information about modeling pull-in instability of MEMS has been presented in Ref. [7], [8]. 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 [9], Mindlin [10] and Toupin [11] 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 [12]. Employing the equilibrium equation of moments of couples beside the classical equilibrium equations of forces and moments of forces, a modified couple stress theory was introduced by Yang, Chong, Lam, and Tong [13], 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. [14], and Asghari et al. [15], respectively. Rong et al. [16] present an analytical method for pull-in analysis of clamped–clamped multilayer beam.

This paper investigates the pull-in instability of micro-beams with a curved ground electrode under action of electric field force within the framework of von-Karman nonlinearity and the Euler–Bernoulli beam theory. The static pull-in voltage instability of clamped-clamped and cantilever micro-beam are obtained by using MAPLE commercial software. The effects of geometric parameters such as beam lengths, width, thickness, gaps and size effect are discussed in detail through a numerical study.

The objective of this paper is to establish neural network model for estimating the pull-in instability voltage of cantilever beams. More specifically, back propagation neural network is used to construct the pull-in instability voltage. Effective parameters influencing pull-in voltage and their levels for training were selected through preliminary calculations carried out on instability pull-in voltage of micro-beam. Network trained by the same numerical data are then verified by some numerical calculations different from those used in training phase, and the best model was selected based on the criterion of having the least average values of verification errors. To the authors' best knowledge, no previous studies which cover all these issues are available. To the authors' best knowledge, no previous studies which cover all these issues are available.

2   FUNDAMENTAL OF MODIFIED COUPLE STRESS THEORY

In the modified couple stress theory, the strain energy density $\bar{u}$ for a linear elastic isotropic material in infinitesimal deformation is written as [17]:

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

Where:

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

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

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

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

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 [18]:

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

For an Euler–Bernoulli beam, the displacement field can be expressed as:

$$u_x = u(x,t) - z\frac{\partial w(x,t)}{\partial x}, \ u_y = 0, \ u_z = w(x,t) \qquad (7)$$

Where u is the axial displacement of the centroid of sections, and w denotes the lateral deflection of the beam. The parameter $\partial w/\partial x$ stands for the angle of rotation (about the y-axis) of the beam cross-sections. Assuming the above displacement field, after deformation, the cross sections remain plane and always perpendicular to the center line, without any change in their shapes. It is noted that parameter z represents the distance of a point on the section with respect the axis parallel to y-direction passing through the centroid.

3   CONSTITUTIVE NONLINEAR EQUATION OF BEAM TYPE MEMS

In this section, the governing equation and corresponding classical and non-classical boundary conditions of a nonlinear microbeam modeled on the basis of the couple stress theory are derived. The coordinate system and loading of an Euler–Bernoulli beam have been depicted in Fig. 1. In this figure, *F(x,t)* and *G(x,t)* refer to the intensity of the transverse distributed force and the axial body force, respectively, both as force per unit length.
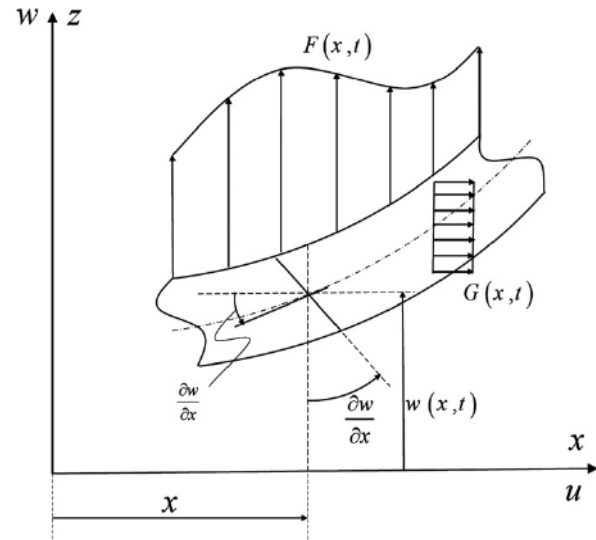

Fig. 1   An Euler–Bernoulli, loading and coordinate system

By assuming small slopes in the beam after deformation, the axial strain, i.e. the ratio of the elongation of a material line element initially in the axial direction to its initial length, can be approximately expressed by the von-Karman strain as:

$$\varepsilon_{xx} = \frac{\partial u_x}{\partial x} + \frac{1}{2}(\frac{\partial w}{\partial x})^2 = \frac{\partial u}{\partial x} - z\frac{\partial^2 w}{\partial x^2} + \frac{1}{2}(\frac{\partial w}{\partial x})^2 \qquad (8)$$

It is noted that finite deflection w is permissible and only it is needed that the slopes be very small. Hereafter, we use Eq. (8) for the axial strain, instead of the infinitesimal definition presented in Eq. (3). Substitution of Eqs. (7) and (8) into (3)-(5) yields the non-zero components. Also, combination of Eqs. (6) and (7) gives [19]:

$$\theta_y = -\frac{\partial w}{\partial x}, \theta_x = \theta_z = 0 \qquad (9)$$

Substitution of Eq. (9) into (5) yields the following expression for the only non-zero components of the symmetric curvature tensor:

$$\chi_{xy} = \chi_{yx} = -\frac{1}{2}\frac{\partial^2 w}{\partial x^2} \tag{10}$$

It is assumed that the components of strains, rotations and their gradients are sufficiently small. By neglecting the Poisson's effect, substitution of Eq. (8) into Eq. (2) gives the following expressions for the main components of the symmetric part of the stress tensor in terms of the kinematic parameters:

$$\sigma_{xx} = E\varepsilon_{xx} = E(\frac{\partial u}{\partial x} - z\frac{\partial^2 w}{\partial x^2} + \frac{1}{2}(\frac{\partial w}{\partial x})^2),$$

other $\sigma_{ij} = 0$ \tag{11}

Where E denotes the elastic modulus. In order to write the non-zero components of the deviatoric part of the couple stress tensor in terms of the kinematic parameters, one can substitute Eq. (10) into Eq. (4) to get:

$$m_{xy} = -\mu l^2 \frac{\partial^2 w}{\partial x^2} \tag{12}$$

Where μ and l are shear modulus and the material length scale parameter, respectively. To obtain the governing equations, the kinetic energy of the beam T, the beam strain energy due to bending and the change of the stretch with respect to the initial configuration U_{bs}, and the increase in the stored energy with respect to the initial configuration due to the existence of initially axial load U_{is} and finally the total potential energy $U = U_{bs} + U_{is}$ are considered as follows:

$$T = \frac{1}{2}\int_0^L \int_A \rho\{(\frac{\partial u}{\partial t} - z\frac{\partial^2 w}{\partial t \partial x})^2 + (\frac{\partial w}{\partial t})^2\}dA\,dx \tag{13a}$$

$$U = \frac{1}{2}\int_0^L \begin{Bmatrix} \{EI(\frac{\partial^2 w}{\partial x^2})^2 + EA(\frac{\partial u}{\partial x} + \frac{1}{2}(\frac{\partial w}{\partial x})^2)^2 \\ + N_0[2\frac{\partial u}{\partial x} + (\frac{\partial w}{\partial x})^2] + \frac{\mu A l^2}{2}(\frac{\partial^2 w}{\partial x^2})^2\} \end{Bmatrix}dx \tag{13b}$$

Where N_0, I and ρ are the axial load, area moment of inertia of section about y– axis and the mass density,

respectively. The work done by the external loads acting on the beam is also expressed as (13c):

$$\delta W = \int_0^L F(x,t)\delta w\,dx + \int_0^L G(x,t)\,\delta u\,dx + (\hat{N}\,\delta u)\Big|_{x=0}^{x=L} + (\hat{V}\,\delta w)\Big|_{x=0}^{x=L}$$

$$+ (\hat{M}\,\delta(\frac{\partial w}{\partial x}))\Big|_{x=0}^{x=L} + (\hat{P}^h\,\delta(\frac{\partial u}{\partial x} + \frac{1}{2}(\frac{\partial w}{\partial x})^2)\Big|_{x=0}^{x=L}$$

$$+ (\hat{Q}^h\,\delta(\frac{\partial^2 w}{\partial x^2}))\Big|_{x=0}^{x=L}$$

Where $\hat{N}$ and $\hat{V}$ represent the resultant axial and transverse forces in a section caused by the classical stress components acting on the section. The resultant axial and transverse forces are work conjugate to u and w, respectively. Also, $\hat{P}^h$ and $\hat{Q}^h$ are the higher-order resultants in a section, caused by higher-order stresses acting on the section. These two higher-order resultants are work conjugate to $\varepsilon_{xx} = \partial u / \partial x + 1/2(\partial w / \partial x)^2$ and $\partial^2 w / \partial x^2$, respectively. The parameter $\hat{M}$ is the resultant moment in a section caused by the classical and higher-order stress components. Now, the Hamilton principle can be applied to determine the governing equation:

$$\int_{t_1}^{t_2}(\delta T - \delta U + \delta W)dt = 0 \tag{14}$$

Where δ denotes the variation symbol. By applying Eqs. (13) and (14), the governing equilibrium micro beam is derived as:

$$S\frac{\partial^4 w}{\partial x^4} - N\frac{\partial^2 w}{\partial x^2} + \rho A\frac{\partial^2 w}{\partial t^2} = F(x,t) \tag{15}$$

Where:

$$N = N_0 + \frac{EA}{2L}\int_0^L(\frac{\partial w}{\partial x})^2 dx \tag{16}$$

$$S = EI + \mu A l^2 \tag{17}$$

If in Eq. (15), *N=0*, then the model of beam is called the linear equation (linear model) without the effect of geometric nonlinearity. The cross sectional area and length of beam are *A* and *L* respectively.

3.1 MODELING EXTERNAL FORCES

In Eq. (15), the external force per unit length of the beam, $f_{external}$, is the summation of the electrostatic Coulomb force per unit length (f_{elec}) and the intermolecular van der Waals (vdW) force per unit length of the nanostructures (f_{vdW}). In this work, effect of vdW force on the size-dependent pull-in instability of micro-beams is not investigated. Considering fringing field correctness effect for narrow beam model, the electrostatic Coulomb attraction in Eq. (15) is written as the following [20]:

$$F_{elec}(x,t) = \frac{\varepsilon_0 B V^2}{2(g-w)^2}[1+0.65\frac{(g-w)}{B}] \qquad (18)$$

Where $\varepsilon_0 = 8.854 \times 10^{-12} C^2 N^{-1} m^{-2}$ is the permittivity of vacuum, V is the applied voltage, g is the initial gap between the movable and the ground electrode and B is width of beam. For clamped-clamped beam, the boundary conditions at the ends are:

$$w(0)=0, \quad \frac{dw(0)}{dx}=0; \quad w(L)=0, \quad \frac{dw(L)}{dx}=0 \qquad (19)$$

For cantilever beam, the boundary conditions at the ends are:

$$w(0)=0, \quad \frac{dw(0)}{dx}=0; \quad \frac{d^2w(L)}{dx^2}=0, \quad \frac{d^3w(L)}{dx^3}=0 \quad (20)$$

In the static case, we have $\frac{\partial}{\partial \tau}=0$ and $\frac{\partial}{\partial x}=\frac{d}{dx}$. Hence, Eq. (15) is reduced to:

$$(EI + \mu A l^2)\frac{d^4w}{dx^4} - [N_0 + \frac{EA}{2L}\int_0^L (\frac{dw}{dx})^2 dx]\frac{d^2w}{dx^2}$$
$$= \frac{\varepsilon_0 B V^2}{2(g-w)^2}[1+0.65\frac{(g-w)}{B}] \qquad (21)$$

A uniform micro-beam has a rectangular cross section with height h and width B, subjected to a given electrostatic force per unit length. Let us consider the following dimensionless parameters:

$$\alpha = \frac{AL^2}{2I}, \quad \beta = \frac{\varepsilon_0 B V^2 L^4}{2g^3 EI}, \quad \gamma = 0.65\frac{g}{B},$$
$$\delta = \frac{\mu A l^2}{EI}, \quad \tilde{w}=\frac{w}{g}, \quad \tilde{x}=\frac{x}{L}, \quad \Gamma=\frac{N_0 L^2}{EI} \qquad (22)$$

In the above equations, the non-dimensional parameter, δ is defined as the size effect parameter. Also, β is non-dimensional voltage parameter. The normalized nonlinear governing equation of motion of the beam can be written as [21]:

$$(1+\delta)\frac{d^4\tilde{w}}{d\tilde{x}^4} - \{\Gamma + \alpha \int_0^1 (\frac{d\tilde{w}}{d\tilde{x}})^2 d\tilde{x}\}\frac{d^2\tilde{w}}{d\tilde{x}^2}$$
$$= \frac{\beta}{(1-\tilde{w})^2} + \frac{\gamma\beta}{(1-\tilde{w})} \qquad (23)$$

4 ANNS

Artificial NNs are non-linear mapping systems with a structure loosely based on principles observed in biological nervous systems. In greatly simplified terms as can be seen from Fig. 2(a), a typical real neuron has a branching dendritic tree that collects signals from many other neurons in a limited area; a cell body that integrates collected signals and generates a response signal (as well as manages metabolic functions); and a long branching axon that distributes the response through contacts with dendritic trees of many other neurons.

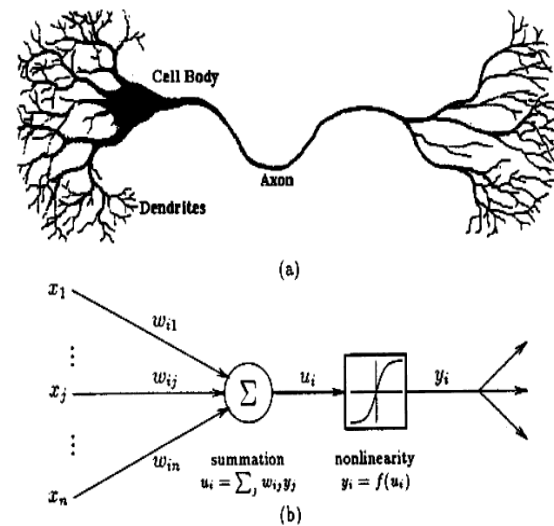**Fig. 2** (a): A biological nervous systems and (b): an artificial neuron model

The response of each neuron is a relatively simple non-linear function of its inputs and is largely determined by the strengths of the connections from its inputs. In spite of the relative simplicity of the individual units, systems containing many neurons can generate complex and interesting behaviors [22]. An ANN

shown in Fig. 3 is very loosely based on these ideas. In the most general terms, a NN consists of large number of simple processors linked by weighted connections. By analogy, the processing nodes may be called neurons. Each node output depends only on information that is locally available at the node, either stored internally or arriving via the weighted connections. Each unit receives inputs from many other nodes and transmits its output to yet other nodes.
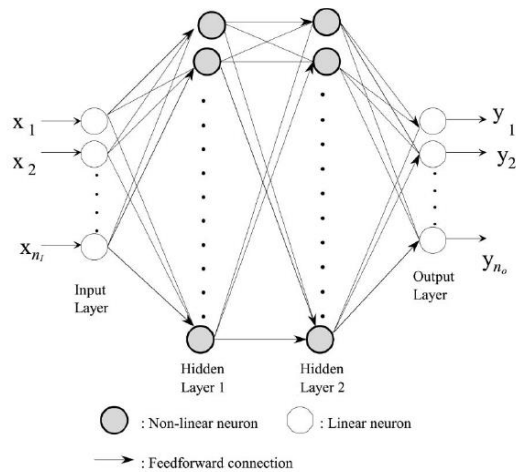

Fig. 3   A layered feed-forward artificial NN

By itself, a single processing element is not very powerful, it generates a scalar output with a single numerical value, which is a simple non-linear function of its inputs. The power of the system emerges from the combination of many units in an appropriate way. A network is specialized to implement different functions by varying the connection topology and the values of the connecting weights. Complex functions can be implemented by connecting units together with appropriate weights. In fact, it has been shown that a sufficiently large network with an appropriate structure and property chosen weights can approximate with arbitrary accuracy any function satisfying certain broad constraints. Usually, the processing units have responses like (see Fig. 2(b)):

$$y = f(\sum_i u_i) \tag{24}$$

Where, u_i are the output signals of hidden layer to output layer, $f(u_i)$ is a simple non-linear function such as the sigmoid, or logistic function. This unit computes a weighted linear combination of its inputs and passes this through the non-linearity to produce a scalar output. In general, it is a bounded non-decreasing non-linear function; the logistic function is a common choice. This model is, of course, a drastically simplified approximation of real nervous systems. The

intent is to capture the major characteristics important in the information processing functions of real networks without varying too much about physical constraints imposed by biology. The impressive advantages of NNs are the capability of solving highly non-linear and complex problems and the efficiency of processing imprecise and noisy data. 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 are used in test phase [23].

5   BACK PROPAGATION NEURAL NETWORK

Back propagation neural network (BPN), developed by Rumelhart [24], 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 interconnectivity of the weights is 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 [25].

6   GENETIC ALGORITHMS

Genetic algorithms have become popular following Holland's work [26]. GAs consisting of continuous and binary forms are designed to efficiently search huge, non- linear, discrete and poorly understood search spaces, where expert knowledge is scarce or difficult to model and traditional optimization techniques has failed [27]. Typically, a simple GA consists of three operations: (1) parent selection, (2) crossover, and (3) mutation. In this research, Roulette wheel selection scheme has been applied among the selection operators [28]. Two-point crossover was used for each chromosome of the chromosome-pair having a 50% chance of selection, the two parents selected for

crossover exchange information lying between two randomly generated points within the binary string. In addition, that most of the successful applications of GAs [29] is greatly dependent on finding a suitable method for encoding the chromosome, the creation of a fitness function to rank the performance of a specific chromosome is also of paramount importance for the success of the training process. The genetic algorithm rates its own performance around that of the fitness function; consequently, if the fitness function does not adequately take account of the desired performance features, the genetic algorithm is unable to meet the deserved requirements of the user. The proposed chromosome includes five genes for learning ratio, two genes for momentum term and five genes pertain to the number of neurons of hidden layer. To sum up, GA was used to optimally search learning ratio, momentum term and the number of neurons in hidden layer. Tables 1 and 2 explore, respectively, the parameters and variables of the GA exploited in the research.

Table 1 The GA parameters

Number of generation	100
Population	30
Chromosome length	12
Selection operator	Roulette
Fitness normalization	Rank
Elitism	1
Crossover	Pc = 0.8, two-point, uniform
Mutation	Pm = 0.1, Gaussian, mean = 0.0, std = 1.0

Table 2 The GA variables

Variable name	Range	Optimized value
Number of the hidden-layer neurons	4-15	8
Learning ratio	0.1-1	0.85
Momentum term	0-1	0.90

7   RESULTS AND DISCUSSION

When the applied voltage between the two electrodes increases beyond a critical value, the electric field force cannot be balanced by the elastic restoring force of the movable electrode and the system collapses onto the ground electrode. The voltage and deflection at this state are known as the pull-in voltage and pull-in deflection, which are of utmost importance in the design of MEMS devices. The pull-in voltage of cantilever and fixed-fixed beams is an important variable for analysis and design of micro-switches and other micro-devices. Typically, the pull-in voltage is a function of geometry variable such as length, width, and thickness of the beam and the gap between the beam and ground plane. To study the instability of the

Nano-actuator, Eq. (23) is solved numerically and simulated. To highlight the differences between linear and nonlinear geometry model results of Euler-Bernoulli microbeam, we first compare the pull-in voltage for a fixed-fixed and cantilever beams with a length of $100\,\mu m$, a width of $50\,\mu m$, a thickness of $1\,\mu m$ and two gap lengths.

For a small gap length of $0.5\,\mu m$, we observe that linear and nonlinear geometry model give identical results. However, for a large gap length of $2\,\mu m$, we observe that pull-in voltage for fixed-fixed beam is significantly different. The difference in the pull-in voltage is even larger when a gap length of $4.5\,\mu m$ is considered. It is evident that pull-in voltage of fixed-fixed beam is larger than fixed-free beam. The gap lengths used vary from 5 to $30\,\mu m$. For gaps smaller than $15\,\mu m$ and lengths larger than $350\,\mu m$, we observe that the pull-in voltage obtained with linear and nonlinear geometry model are very close.

However, for large gaps and for short beams, we observe that the difference in the pull-in voltage obtained with linear and nonlinear geometry model is not negligible. When the gap increases, the error in pull-in voltage with linear model increase significantly. Furthermore, contrary to the case of cantilever beams, the thickness has a significant effect on the error in pull-in voltages. Another observation is that the length of the beam has little effect on the error in pull-in voltage. This observation is also different from the case of cantilever beams. From the results, it is clear that the linear model is generally not valid for the fixed-fixed beams case, except when the gap is very small, such as the $0.5\,\mu m$ case. The results represent that the size effect increases the pull-in voltage of the Nano-actuators.

8   MODELING OF SPIV OF CANTILEVER BEAM USING BACK PROPAGATION ANN

Modeling of pull-in instability of micro-beam with BP neural network is composed of two stages: training and testing of the networks with numerical data. The training data consisted of values for beam length (L), gap (g), width of beam (b) and (h/l), and the corresponding static pull-in instability voltage (V_{PI}). Total 120 such data sets were used, of which 110 were selected randomly and used for training purposes whilst the remaining 10 data sets were presented to the trained networks as new application data for verification (testing) purposes. Thus, the networks were evaluated using data that had not been used for training. Training/Testing pattern vectors are formed, each formed with an input condition vector, and the

corresponding target vector. The network has four inputs of beam length (*L*), gap (*g*), width of beam (*b*) and (*h/l*) ratio and one output of static pull-in voltage (V_{PI}). Table 3 shows 10 numerical data sets which have been used for verifying or testing network capabilities in modeling the process.

Therefore, the general network structure is supposed to be 4-n-1, which implies 4 neurons in the input layer, n neurons in the hidden layer, and 1 neuron in the output layer. Then, by varying the number of hidden neurons, different network configurations are trained, and their performances are checked. For training problem, equal learning rate and momentum constant of $\eta = \alpha = 0.9$ were used [26]. Also, error stopping criterion was set at *E=0.01*, which means training epochs continued until the mean square error fell beneath this value. The size of hidden layer(s) is one of the most important considerations when solving actual problems using multi-layer feed-forward network. However, it has been shown that BP neural network with one hidden layer can uniformly approximate any continuous function to any desired degree of accuracy given an adequate number of neurons in the hidden layer and the correct interconnection weights [25]. Therefore, one hidden layer was adopted for the BP model.

Table 3 Beam geometry and pull-in voltage for verification analysis

Test No.	$L(\mu m)$	$b(\mu m)$	h/l	$g(\mu m)$	V_{PI} (volt)
1	75	0.5	4	0.5	0.179
2	100	5	6	1	2.44
3	125	10	8	1.5	7.31
4	150	20	10	2	16.82
5	175	25	12	2.5	26.78
6	200	30	14	3	40.27
7	225	35	16	3.5	53.84
8	250	40	18	4	68.01
9	275	45	20	4.5	84.53
10	300	50	22	5	103.62

To determine the number of neurons in the hidden layer, a procedure of trail and error approach needs to be done. As such, attempts have been made to study the network performance with a different number of hidden neurons. Hence, a number of candidate networks are constructed, each of trained separately, and the "best" network were selected based on the accuracy of the predictions in the testing phase. It should be noted that if the number of hidden neurons is too large, the ANN might be over-trained giving spurious values in the testing phase. If too few neurons are selected, the function mapping might not be accomplished due to under-training. Three functions, namely newelm, newff and newcf [30] have been used for creating BP networks. Then, by varying the number of hidden

neurons, different network configurations are trained, and their performances are checked. The results are shown in Table 4. Both the required iteration numbers and mapping performances were examined for these networks. As the error criterion for all networks was the same, their performances are comparable. As a result, from Table 4, the best network structure of BP model is picked to have 8 neurons in the hidden layer with the average verification errors of 6.36% in predicting V_{PI} by newelm function. Tables 5, 6 and 7 show the comparison of calculated and predicted values for static pull-in voltage in verification cases. After 1884 epochs, the MSE between the desired and actual outputs becomes less than 0.01. At the beginning of the training, the output from the network is far from the target value. However, the output slowly and gradually converges to the target value with more epochs and the network learns the input/output relation of the training samples.

Table 4 The effects of different number of hidden neurons on the BP network performance

No. of hidden neurons	Epoch	Average error in V_{PI} (%) with newelm function	Average error in V_{PI} (%) with newcf function	Average error in V_{PI} (%) with newff function
4	18914	12.31	10.27	12.30
5	4970	14.38	18.38	20.19
6	1783	8.19	11.65	12.75
7	3984	9.72	9.39	11.17
8	1884	6.36	8.28	10.14
9	2770	13.39	11.86	19.98
10	2683	11.67	16.40	15.48

Table 5 Comparison of V_{PI} calculated and predicted by the BP neural network model with newelm function

Test No.	V_{PI} (volt)		
	Calculated	BP model (newelm)	Error (%)
1	0.179	0.190	6.56
2	2.44	2.61	7.28
3	7.31	7.32	0.16
4	16.82	19.21	14.24
5	26.78	28.22	5.39
6	40.27	42.17	4.74
7	53.84	57.13	6.12
8	68.01	71.60	5.29
9	84.53	89.39	5.76
10	103.62	112.03	8.12

Table 6 Comparison of V_{PI} calculated and predicted by the BP neural network model with newff function

Test No.	V_{PI} (volt)		
	Calculated	BP model (newff)	Error (%)
1	0.179	0.191	7.12
2	2.44	2.52	3.39
3	7.31	7.54	3.16
4	16.82	18.29	8.75
5	26.78	28.16	5.19
6	40.27	46.03	14.31
7	53.84	57.49	6.79
8	68.01	78.39	15.27
9	84.53	86.71	2.59
10	103.62	120.53	16.32

Table 7 Comparison of V_{PI} calculated and predicted by the BP neural network model with newcf function

Test No.	V_{PI} (volt)		
	Calculated	BP model (newcf)	Error (%)
1	0.179	0.193	8.29
2	2.44	2.59	6.28
3	7.31	7.70	5.39
4	16.82	19.21	14.24
5	26.78	28.76	7.41
6	40.27	43.99	9.25
7	53.84	57.96	7.67
8	68.01	82.14	20.78
9	84.53	92.51	9.45
10	103.62	116.74	12.67

CONCLUSIONS

The primary contributions of the paper are summarized as follows.

1. The BP neural network is capable of constructing model using only numerical data, describing the static pull-in instability behavior.
2. 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.
3. For cantilever beams, length has a significant effect on the error in pull-in voltages, while for fixed-fixed beams, the length has little effect on the error. On the other hand, for fixed-fixed beams, thickness has significant effect on the error in pull-in voltage, while for cantilever beams it has little effect.
4. The static pull-in instability voltage of clamped–clamped and cantilever beam are compared. For both clamped–clamped and cantilever beams, the pull-in

voltage in nonlinear geometry beam model is bigger than linear model.
5. For both fixed-fixed and cantilever beams by increasing of gap length, the pull-in voltage is significantly increased.
6. For both fixed-fixed and cantilever beams by increasing of thickness of beams, the pull-in voltage is significantly increased.
7. For both fixed-fixed and cantilever beams by increasing of length of beams, the pull-in voltage is significantly decreased.
8. By using modified couple stress theory, it is found that the dimensionless pull-in voltage of MEMS increases linearly due to the size effect. This emphasizes the importance of size effect consideration in design and analysis of MEMS.
9. When the ratio of *h/l* increases, the pull-in voltage predicted by modified couple stress theory and ANN is constant approximately.

The conclusion above indicates that the geometry of beam has significant influences on the electro-static characteristics of micro-beams that can be designed to tailor for the desired performance in different MEMS applications.

REFERENCES

[1] Tadi Beni, Y., Karimipour, I., and Abadyan, M., "Modeling the Instability of Electrostatic Nano-Bridges and Nano-Cantilevers Using Modified Strain Gradient Theory", Applied Mathematical Modelling, Vol. 39, No.9, 2015, pp. 2633-2648.

[2] Gasparini, A. M., Saetta, A. V., and Vitaliani, R. V., "On the Stability and Instability Regions of Non-Conservative Continuous System under Partially Follower Forces", Computer Methods in Applied Mechanics and Engineering, Vol. 124, No. (1-2), 1995, pp. 63-78.

[3] Osterberg, P. M., Senturia, S. D., "M-TEST: a Test Chip for MEMS Material Property Measurements Using Electrostatically Actuated Test Structures", Journal of Microelectromechanical Systems, Vol. 6, No. 2, 1997, pp. 107-118.

[4] Osterberg, P. M., Gupta, R. K., Gilbert, J. R., and Senturia, S. D., "Quantitative Models for the Measurement of Residual Stress, Poisson Ratio and Young's Modulus Using Electrostatic Pull-in of Beams and Diaphragms", Proceedings of the Solid- State Sensor and Actuator Workshop, Hilton Head, SC, 1994.

[5] Sadeghian, H., Rezazadeh, G., Osterberg, P., "Application of the Generalized Differential Quadrature Method to the Study of Pull-in Phenomena of MEMS Switches", Journal of Microelectromechanical Systems, Vol. 16, No. 6, 2007, pp. 1334-1340.

[6] Salekdeh, Y. A., Koochi, A., Beni, Y. T., and Abadyan, M., "Modeling Effect of Three Nano-Scale Physical Phenomena on Instability Voltage of Multi-Layer MEMS/NEMS: Material Size Dependency, Van Der Waals Force and Non-Classic Support Conditions", Trends in Applied Sciences Research, Vol. 7, No. 1, 2012, pp. 1-17.

[7] Batra, R. C., Porfiri, M., Spinello, D., "Review of Modeling Electrostatically Actuated Microelectromechanical Systems", Smart Materials and Structures, Vol. 16, No. 6, 2007, R23-R31.

[8] Lin, W. H., Zhao, Y. P., "Pull-in Instability of Micro-Switch Actuators: Model Review", International Journal of Nonlinear Sciences and Numerical Simulation, Vol. 9, No. 2, 2008, pp.175-184.

[9] Koiter, W. T., "Couple-Stresses in the Theory of Elasticity: I and II.", Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen Series B, 1964, pp. 6717-6744.

[10] Mindlin, R. D., Tiersten, H. F., "Effects of Couple-Stresses in Linear Elasticity", Archive for Rational Mechanics and Analysis, Vol. 11, No. 1, 1962, pp. 415-448.

[11] Toupin, R. A., "Elastic Materials with Couple-Stresses, Archive for Rational Mechanics and Analysis", Vol. 11, No. 1, 1962, pp. 385-414.

[12] Anthoine, A., "Effect of Couple-Stresses on the Elastic Bending of Beams", International Journal of Solids and Structures, Vol. 37, No. 7, 2000, pp. 1003-1018.

[13] Yang, F., Chong, A. C. M., Lam, D. C. C., and Tong, P., "Couple Stress Based Strain Gradient Theory for Elasticity", International Journal of Solids and Structures, Vol. 39, No. 10, 2002, pp. 2731-2743.

[14] Xia, W., Wang, L., and Yin, L., "Nonlinear Non-Classical Microscale Beams: Static Bending, Post Buckling and Free Vibration", International Journal of Engineering Science, Vol. 48, No. 12, 2010, pp. 2044-2053.

[15] Asghari, M., Ahmadian, M. T., Kahrobaiyan, M. H., and Rahaeifard, M., "On the Size-Dependent Behavior of Functionally Graded Micro-Beams", Materials and Design, Vol. 31, No. 5, 2010, pp. 2324-2329.

[16] Rong, H., Huang, Q. A., Nie, M., and Li, W., "An Analytical Model for Pull-in Voltage of Clamped–Clamped Multilayer Beams", Sensors and Actuators A: Physical, Vol. 116, No. 1, 2004, pp. 15-21.

[17] Yang, F., Chong, A. C. M., Lam, D. C. C., and Tong, P., "Couple Stress Based Strain Gradient Theory for Elasticity", International Journal of Solids and Structures, Vol. 39, No. 10, 2002, pp. 2731-2743.

[18] Shengli, K., Shenjie, Z., Zhifeng, N., and Kai, W., "The Size-Dependent Natural Frequency of Bernoulli–Euler Micro-Beams", Journal of Engineering Science, Vol. 46, No. 5, 2008, pp. 427-437.

[19] Ma, H. M., Gao, X. L., and Reddy, J. N., "A Microstructure-Dependent Timoshenko Beam Model Based on a Modified Couple Stress Theory", Journal of the Mechanics and Physics of Solids, Vol. 56, No. 12, 2008, pp. 3379-3391.

[20] Tadi Beni, Y., Koochi, A., and Abadyan, M., "Theoretical Study of the Effect of Casimir Force, Elastic Boundary Conditions and Size Dependency on the Pull-In Instability of Beam-Type NEMS", Physica E: Low-dimensional Systems and Nanostructures, Vol. 43, No. 4, 2011, pp.979-988.

[21] Zhao, J., Zhou, S., Wanga, B., and Wang, X., "Nonlinear Microbeam Model Based on Strain Gradient Theory", Applied Mathematical Modelling, Vol. 36, No. 6, 2012, pp. 2674-2686.

[22] Freeman, J. A., Skapura, D. M., "Neural Networks: Algorithms, Applications, and Programming Techniques", Addision-Wesley, 1992.

[23] Gao, D., Kinouchi, Y., Ito, K., and Zhao, Z., "Neural Networks for Event Extraction from Time Series: A Back Propagation Algorithm Approach, Future Generation Computer Systems", Vol. 21, No. 7, 2005, pp. 1096-1105.

[24] Rumelhart, D. E., Hinton, G. E., and Williams, R. J., "Learning Representations by Back Propagating Error", Nature, Vol. 323, 1986, pp. 533-536.

[25] Zhang, H., Wei, W., and Mingchen, Y., "Boundedness and Convergence of Batch Back-Propagation Algorithm with Penalty for Feedforward Neural Networks", Neurocomputing, Vol. 89, 2012, pp. 141-146.

[26] Holland, J. H., "Adaption in Natural and Artificial Systems", Ann Arbor: University of Michigan Press, 1975.

[27] He, Y., Guo, D., and Chu, F., "Using Genetic Algorithms and Finite Element Methods to Detect Shaft Crack for Rotor-Bearing System", Mathematics and Computers in Simulation, Vol. 57, No. 1-2, 2001, pp. 95-108.

[28] Wong, M. L. D., Nandi, A. K., "Automatic Digital Modulation Recognition Using Artificial Neural Network and Genetic Algorithm", Signal Processing, Vol. 84, No. 2, 2004, pp. 351-365.

[29] Tang, K. S., Man, K. F., Kwong, S., and He, O., "Genetic Algorithms and Their Applications", IEEE Signal Processing Magazine, Vol. 13, No. 6, 1996, pp. 22-37.

[30] Demuth, H., Beale M., Matlab Neural Networks Toolbox, User's Guide, The Math Works, Inc., http://www.mathworks.com, 2001.