AIAA JOURNAL
Vol. 41, No. 4, April 2003

Neural Networks for Inverse Problems in Damage Identification and Optical Imaging

Yong Y. Kim* and Rakesh K. Kapania†
Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061

Artificial neural networks (ANNs) are employed in solving inverse problems in damage detection in structures, as well as detection, using optical imaging, of an anomaly in a light-diffusive media, such as a human tissue. Both of these problems, namely, identifying the damage parameters in a damaged structure and identifying the representative properties in a tissue, require solving highly complex inverse problems. The neural networks (NNs) for both problems are similar, and a method found suitable for solving one type of problem can be applied for solving the other type of problem. In the damage identification problem, the natural frequencies of a damaged beam model obtained from analytical and numerical methods were used to identify damage parameters by employing feedforward backpropagation, and also radial basis NNs. In the optical imaging problem, the tissue under investigation was illuminated by a number of near-infrared light sources placed around the circumference of the tissue. Both the location and the size of the anomaly were identified by studying the influence of the anomaly on the light intensity received at the boundary of the tissue. The near-infrared light measurements are assumed to be available at a number of light detector positions, also along the circumference of the tissue. NNs were used to determine the location and the size of the anomaly in a tissue. The direct problem for the case of optical imaging was solved using the finite element method to generate the training and testing sets for NNs.

Introduction

THE transfer of technology developed in one area to another area is very desirable. A soft computing method such as artificial neural networks (ANNs), which is finding applications in almost all branches of science and engineering, can be used in different engineering fields without changing much of the basic methodology. Here, we used neural networks on two problems and gain benefits from this experience. The capability of solving inverse problems using neural networks (NNs) has been studied for two different kinds of problems in our work, employing numerical simulations. One is a damage identification problem, and the other is an optical imaging problem. The objective in the damage identification problem is to identify damage parameters in a structure by measuring variations in the dynamic behavior of damaged and perfect systems. The objective in the optical imaging problem is to identify an anomaly in a light-diffusive media, such as a human tissue, by studying the effect of the anomaly on the propagation of near-infrared light passing through the tissue. Both problems have a common feature in that an attempt was made to find an anomaly in an otherwise perfect system using limited data such as natural frequencies and diffused measurements only at the boundary. Therefore, there exists the possibility of using similar approaches, such as an application of ANNs, to solve both types of inverse problems.

The ANNs are imitations of real neurons in the human brain in their structure, data processing and restoring method, and learning ability. In ANNs and real neurons, knowledge is acquired by the network through a learning process, and interneuron connection strengths known as synaptic weights are used to store the knowledge. This organization of neurons is what enables the brain to perform certain tasks, like pattern recognition and motor control, much faster than the fastest digital computer available even today. NNs can be considered universal approximators for arbitrary functions that make probabilistic assumptions about data. The learning process can be considered a method for finding parameter values that appear to be closest to the available data. NNs, as an inverse problem solving algorithm, have advantages in that they are capable of representing nonlinear systems, are fault tolerant, and are capable of providing instantaneous responses once properly trained.

The most used NN structures are feedforward backpropagation NNs as shown in Fig. 1. The output t of the network is a transferred sum of weighted inputs p, with added bias using the sigmoid or a linear function. The relation of input and output is shown in Eq. (1), where w is weight and f is a transfer function[1]:

$$t_i = f\left(\sum_{j=1}^{n} w_{i,j} \cdot p_i + b_i\right) \tag{1}$$

Feedforward networks with one or more hidden layers of sigmoid neurons and a linear output layer are often used. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output. This structure of NN can approximate any function with a finite number of discontinuities, given that an adequate number of neurons are present in the hidden layers.

Another structure of NN that is frequently used is radial-basis function NN, shown in Fig. 2. A radial-basis function NN has one input layer, only one hidden layer, and one output layer. The relation between the input p_1 and the output t_1 is shown in Eq. (2) where b is bias, f_r is radial basis function, and $\| \cdot \|$ indicates the dot product of two vectors[1]:

$$t_i = f_r(\|\boldsymbol{w} \cdot \boldsymbol{p}\| b_i) \tag{2}$$

The outputs of the first-layer neurons, each of which represents a basis function, are determined by the distance between the network input and the center of the basis function. A radial-basis function NN uses a linear output layer, and a radial-basis function NN can be constructed very quickly by adding as many radial-basis neurons as there are in the input layer.

There have been many efforts to use NNs in studying damage identification problems. Kudva et al.[2] used backpropagation NNs in identifying holes of various diameters in a plate stiffened with bays. NNs were used to identify the location and size of holes based on the strain-gauge data. They demonstrated that an NN can be used to determine damage location and size in a typical structure. Wu et al.[3] used a back-propagation NN in identifying damage in a three-story building. The NN simulated the level of damage in each

Received 6 March 2002; revision received 5 November 2002; accepted for publication 15 November 2002.

*Graduate Assistant, Department of Aerospace and Ocean Engineering.
†Professor, Department of Aerospace and Ocean Engineering. Associate Fellow AIAA.

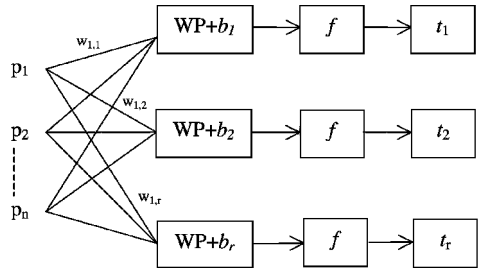


Fig. 1 Feedforward NN (p = inputs, W = weights, b = bias, f = transfer function, n = number of inputs, r = number of neurons, and t = outputs).

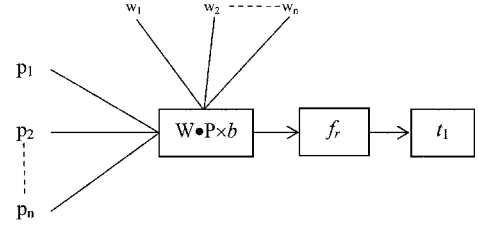


Fig. 2 Radial-basis NN (p = inputs, W = weights, f_r = radial-basis function, and t = outputs).

of the members, based on the Fourier transform of the acceleration data. They were able to predict damages on the first and the third floor but not on the second floor.

Leath and Zimmerman[4] used multilayer backpropagation NNs in identifying damage in a four-element cantilever beam based on the first two bending frequencies. They also developed a training algorithm that can create an NN that fits the data with a minimal number of neurons. The damage was modeled by reducing Young's modulus by up to 95%. They fixed the number of hidden nodes to be one less than the number of data points and also found that exactly fitting the data can cause problems in most cases. The damage was identified within an error of 35%. Spillman et al.[5] used feedforward NN in identifying damage in a steel bridge using an accelerometer and a fiber optic modal sensor. Damage was modeled by loosening or removing a bolted plate in the bridge. The NN used the height and frequency of the first two modal peaks and the impact intensity and location as inputs. They were able to diagnose damage correctly in the proportion of 58% of the testing cases. Worden and Iomlinson[6] used a back-propagation NN in a framework structure. Damage was modeled by removing one member from the framework. The NN identified the map from static strain data to a measure of the damage. They used three hidden layers and used a finite element method (FEM) to generate training data. An experimental model of the same geometry was used to test the trained NN. The trained NN was able to identify the location of the damage most of the time.

Kirkegaard and Rytter[7] used backpropagation NN in a steel lattice mast. Damage was simulated by bolted joints of reduced thickness. The NN simulated the degree of damage in stiffness based on the first five natural frequencies. They also used FEM to generate the training data, and the training data had a 0–100% reduction in diagonal cross-sectional area. The trained NN was able to locate the damage, which has a magnitude of more than 50%. Manning[8] used backpropagation NN in a truss structure. Damage was modeled by a change in one of the member's cross section. The NN simulated the member's cross-sectional area based on a pole-zero location extracted from the frequency response functions between the actuator and the two piezoceramic sensors on the same member and the stiffness information of the member. The networks were tested on three examples and could predict member area within 10% for most of the members. Povich and Lim[9] used a backpropgation NN in a planar truss with struts. Damage was modeled by removing struts from the structure. The NN simulated damage in each member based on the Fourier transform of the acceleration time history. The NN had 384 inputs and 60 outputs, which correspond to each strut. The NN was trained with 61 examples and was able to correctly identify a missing strut in 21 out of a total of 38 testing cases.

Tsou and Shen[10] used a backpropagation NN in three-degree-of-freedom (DOF) and eight-DOF spring–mass systems. Damage was modeled by reducing one of the spring constants by an amount that varied from 10 to 80%. The NN simulated the change in the spring stiffness based on the changes in the natural frequencies for the case of a three-DOF system and that using the residual modal force in an eight-DOF system. They employed a two-step approach in identifying the damage using NNs. The first NN identified which spring is damaged. Based on the output of the first NN, an appropriate NN is chosen to quantify the damage in the damaged spring. They were able to identify the extent of the damage with an accuracy of 5% in the interpolation case, but the error went up to 30% if the NN had to use extrapolation. Rhim and Lee[11] used backpropagation NN in identifying delamination in a composite cantilevered beam modeled using the FEM. An autoregressive system identification was performed on the transfer function of the beam from the force input to the displacement. Next, the NN was used to simulate an experimental damage scale, based on the denominator of the transfer function. The network was trained with 10 training patterns and was able to identify the damage in three test examples correctly.

Barai and Pandey[12] used backpropagation NN to identify the changes in the stiffness in different elements in the truss structures using nodal time histories. They used a finite element simulation of the truss structure representing a bridge with a moving point force to simulate a vehicle at constant velocity. They were able to predict the change in stiffness within an accuracy of 4%. Ceravolo and De Stefano[13] used a backpropagation NN in the truss structure by an FEM. Damage was modeled by removing elements from the structure. The NN simulated the location of damage using the first 10 natural frequencies. They found that an NN with two hidden layers performs better than an NN with only one hidden layer. Schwarz et al.[14] used a backpropagation NN in spring–mass systems. They used the NN to identify the changes in the spring constants that simulate damage based on the changes in the natural frequencies.

Ganguli and Chopra[15] used a comprehensive physics-based model of a helicopter rotor in forward flight. The rotor model employed finite elements in both space and time. They used two NNs. The first was to classify the type of damage, and the second was to characterize the level of damage from noise-contaminated simulated vibration and blade response test data. A backpropagation NN with one hidden layer was used with adaptive learning. They were able to identify single and multiple faults by moisture absorption, damaged lag damper, and a damaged pitch control system using an NN trained on simulated ideal and noise-contaminated data. They found out that it is important to train an NN using noise-contaminated response data for accurate estimation. The NN, trained on noisy data, gave almost zero error for noise levels less than 10%.

Marwala[16] used a committee of NNs with three types of information, namely, the frequency-response function, modal properties, and wavelet data, to identify four types of faults in a cylindrical shell. The committee of three networks used together resulted in a lower mean square error than the average mean square error given by any individual approach. He also found that the effectiveness of the method was enhanced when experimental data were used because the committee approach assumes that the three types of data used in the three approaches are uncorrelated. Palakal et al.[17] used a backpropagation NN with wavelet transform to predict material loss based on images of corroded regions on the structure. They used wavelet transforms using the Harr filter for extracting parameters for segmenting the image. They obtained a large set of data required for the NN training by perturbing images. The trained NN was able to predict the material loss quite well.

There have been a number of researchers who tried to use NNs types other than the feedforward backpropagation NN. Szewczyk and Hajela[18] used a counterpropagation NN, which is similar to building an adaptive lookup table from the given data, to find damage in a truss structure. The advantage of this NN is that the architecture is selected by the data. The trained NN identified the stiffness of the members based on the static deformation under a given load within an error of 30%. Klenke and Paez[19] used a probabilistic NN and a probabilistic pattern classifier in identifying the existence of damage. They used a Gaussian kernel-density estimator to approximate

the probability density function to be used in their scheme. For a piece of new data, an estimated likelihood was computed for membership into two classes, damaged and undamaged. The greater of the two likelihoods was taken as the guess for class membership. Damage was modeled by changes in spring constants. In all of the test cases, damage was clearly identified.

Luo and Hanagud[20] proposed a dynamic learning rate steepest descent method in detection of delaminations and impact damage using modal analysis and an FEM. Real-time health monitoring techniques based on the modal analysis or finite element analysis are limited by time-consuming signal processing. Their method, based on the simple steepest descent method, improved the learning convergence speed significantly, so that they were able to show that an NN can be used in real-time signal processing with modal data and finite element analysis. Doyle and Fernando[21] used a fast Fourier transform in the preprocessing stage and backpropagation and a learning vector quantization NN in detecting impact damage in a composite material using an optical fiber vibration sensor system. They achieved a lower error with fewer iterations using a learning vector quantization NN than with a backpropagation NN.

Li et al.[22] used an adaptive resonance theory NN embedded with fuzzy classifiers. The adaptive resonance theory based on NN is made up of two layers of cells, an input feature layer and an encoded category representation layer. When the pattern of an input matches one of the stored examples, only the matching cell resonates with the input and outputs a value that is close to one. The trained NN was able to correctly identify the condition of significant wear on the tool before the rapid wear stage, based on the optical scattering image of a surface. Wu et al.[23] used wavelet-based NNs using a data-fusion method for the damage detection of anisotropic composite materials. They used a piezoelectric ceramic patch as an actuator and another as a sensor to obtain output acceleration signals. Using wavelet packets, they decomposed signals into eight different wave bands and were able to obtain different features in decomposed wave bands for a healthy structure and for structures with various levels of damage. The obtained features for different cases of damage were used to train an adaptive NN to recognize the characteristics of different damage. The trained NN was able to distinguish between different levels of damage.

In our damage identification problem, a diffusively damaged Euler–Bernoulli beam model, described by Cerri and Vestroni,[24] was utilized to evaluate the use of NNs for identification purposes. In Cerri and Vestroni's work, a procedure based on modal equation was used to solve the inverse problem. Differences in natural frequencies of damaged beam and the perfect beam were utilized in determining three damage parameters. The process used by Cerri and Vestroni in finding the parameters was graphical and was difficult to implement for some cases, for example, ill-conditioned cases. On the other hand, NNs are able to map this three-dimensional parameter space automatically and give the values of damage parameters, given the natural frequencies of the damaged beam, in an instantaneous manner. Our work differs from previous work using NNs in that we are identifying the extent of the damage in addition to the location and the magnitude of the damage. In our case, the inputs for the NNs are the natural frequencies of a damaged beam and the targets to be identified using an NN are the location, the extent, and the magnitude of the damage. For this purpose, the most suitable structure of NNs and training methods were studied. The experience in utilizing NNs in this damage identification problem was found to be very valuable in solving the optical imaging problem and vice versa.

Optical imaging refers to the methodology of using light in a narrow wavelength band in the near-infrared (~700–1000-nm) spectrum to transilluminate a light-diffusive media such as a human tissue and to use the resulting measurements of intensity on the tissue boundary to reconstruct a map of the internal optical properties. The main advantage of optical imaging is its capability of safely and portably measuring tissue functions to detect nonfunctioning cells, such as cancerous cells. It is known that to be able to observe the functioning of cells, a continuous and noninvasive imaging method is required. Moreover, optical imaging has advantages over x-ray, computer tomographic (CT) scans, and positron emission tomography because of its portability and lower equipment cost.[25] The main idea of optical imaging is that light passes through the body in small amounts, carrying with it characteristics about tissues through which it has passed. Based on these characteristics, optical images can be obtained by solving the inverse problem of light propagation. However, diffusive and scattering characteristics of near-infrared light when propagating in a tissue leads to a highly nonlinear inverse problem whose solution requires large amounts of computational time, even for relatively coarse measurements, if conventional methods are used.[26–33]

The application of NNs may help reduce this computational time, thus, making the use of optical imaging as a method of solving the underlying inverse problem a more viable approach. The process of solving this inverse problem consists of two parts. The first part entails solving the direct problem of a light-diffusion equation to predict propagation of photons in a tissue. The second part, based on the solutions of the direct problem and the information observed in the detectors, entails obtaining an optical image by solving the inverse problem.

There also have been a number of researchers who have used NNs in various medical imaging problems, but none have used NNs in optical medical imaging. Chiu and Yau[34] used NNs for the reconstruction of x-ray CT images of a time-varying object that has motion artifacts. They introduced an imaging reconstruction method based on the previous knowledge of the projections and developed a novel training algorithm. Nakao et al.[35] used backpropagation algorithm to construct a CT image using four projections and compared the performance of this algorithm with that of the algebraic reconstruction technique (ART). They found that their algorithm is more effective than the ART. Kiliç and Korürek[36] used a finite element based NN algorithm that calculates conductivity distribution from the electrical measurements on the medium boundary. Aizenberg et al.[37] used a special kind of cellular NN based on multiple-valued threshold logic in the complex plane to detect edges and enhance resolution in a CT image.

Among the existing approaches to solving the direct problem of photon transport in tissue, the fastest method at present is perhaps the FEM for solving the photon-diffusion equation. In our research, we have used a FEM program and have integrated this program, with NNs available in the Neural Network Toolbox in MATLAB®, to identify the presence of an anomaly using measurements of near-infrared light that has propagated through a human tissue and has gone through a high degree of both scattering and absorption. As for the optical imaging, the inverse problem has been solved using traditional techniques such as a gradient-based optimization scheme and the perturbation approach.[26–29] These conventional methods are slow and may get trapped in a local minima. We have used NNs in this problem to be able to obtain both a set of outputs (location and size of an anomaly) to the given set of inputs (near-infrared light measured at a number of detectors and the light emitted by a number of sources) and to simultaneously achieve a higher resolution in an instantaneous manner.

Damage Identification

As shown in Fig. 3, the damage zone of the beam is represented as being a distance X away from the midspan, having a length L_D, a density ρ, area A, and a zone of reduced stiffness of EI^D in a beam that has initial bending stiffness EI and an overall span of L. These quantities are nondimensionalized with respect to the total beam dimensions to form three nondimensionalized damage parameters[24]: location $x = X/(L/2)$, damage length $b = L^D/L$, and damage magnitude $\beta = (EI^U - EI^D)/EI^U$.

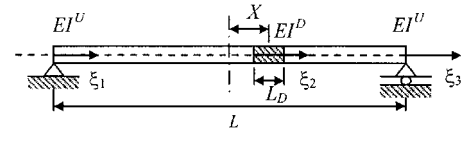


Fig. 3 Damaged beam model with three damage parameters (X = location, L^D = extension of damage, and EI^D = magnitude of damage).

The governing equation of a vibrating Euler–Bernoulli beam is given by

$$EI\frac{\partial^4 v(x,t)}{\partial x^4} + \rho A\frac{\partial^2 v(x,t)}{\partial t^2} = 0 \tag{3}$$

where v is a vertical displacement and t is time.

When the damaged beam with three zones, ξ_1, ξ_2, and ξ_3, is segmented, each segment can be expressed as Eq. (4), where $\lambda_i^4 = \omega^2[\rho A/(EI)_i]L^4$, $\xi_i = x_i/L$. When the solution in the form shown in Eq. (5) is utilized and the boundary conditions at the beam ends and at each of the segment boundaries are satisfied, the characteristic equation for the problem is obtained.[23] The solution of the direct problem was obtained by numerically solving this characteristic equation for each beam segment:

$$\frac{\mathrm{d}^4 V_i}{\mathrm{d}\xi_i^4} - \lambda_i^4 V_i = 0, \qquad i = 1, 2, 3 \tag{4}$$

$$V_i(\xi_i) = A_{i1}\sinh\lambda_i\xi_i + A_{i2}\cosh\lambda_i\xi_i + A_{i3}\sin\lambda_i\xi_i + A_{i4}\cos\lambda_i\xi_i \tag{5}$$

MATLAB was used to solve the characteristic equation. To develop an NN to identify damage parameters, Neural Network Toolbox in MATLAB has been used. An NN was trained for the first three to five natural frequencies, which themselves are obtained by solving the direct problem as inputs and the damage parameters as the training targets. The trained NN was able to provide damage parameters as the output for a new set of inputs (natural frequencies). These cases, where the NNs have to identify one, two, and three damage parameters, have been studied separately.

In the first case, only one damage parameter was identified from the three known natural frequencies. A total of three inputs, the three natural frequencies, and, as output, the damage location, were considered. In the first step, the characteristic equation was solved to get natural frequencies for various values of possible locations of damage. These natural frequencies, coupled with the target, the location of the damage, were used to train an NN.

The artificial NN used in this analysis was a two-layer feedforward NN with two hidden layers. Transfer functions found to be most efficient for the NN were the sigmoid transfer function for the two hidden layers and a linear transfer function for the output layer. The most efficient training method in this case was a resilient backpropagation method. This method eliminates the harmful effects of the magnitude of partial derivatives required in the steepest descent method and enables a faster training of a multilayer NN that uses sigmoid transfer functions.[38] Figure 4 shows the simulation errors, which are the errors between the desired value and the output of the trained NN, in the case when the NN was trained with 40 data points. In this case, the errors were relatively small compared to the case when the NN was trained only with 20 data points.

Next, the two-parameter case was studied. Here, the diffused damage is represented by a concentrated damage in the form of a reduced beam bending rigidity. Figure 5 shows an example of training and testing points. The training points are shown as squares and the testing points, which lie in the midway between training points, are shown as diamonds.

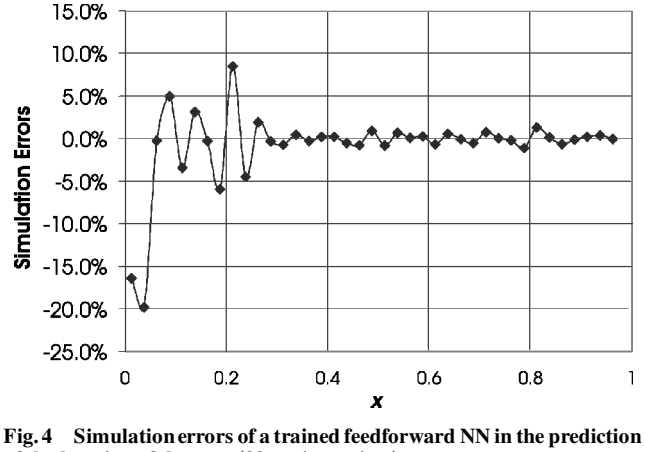


Fig. 4 Simulation errors of a trained feedforward NN in the prediction of the location of damage (39 testing points).

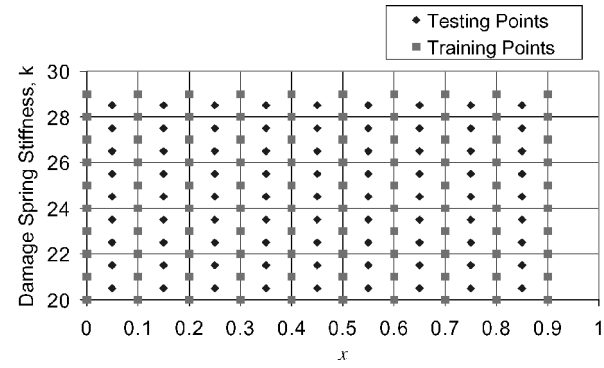


Fig. 5 Distribution of testing and training points for an NN to identify two parameters.

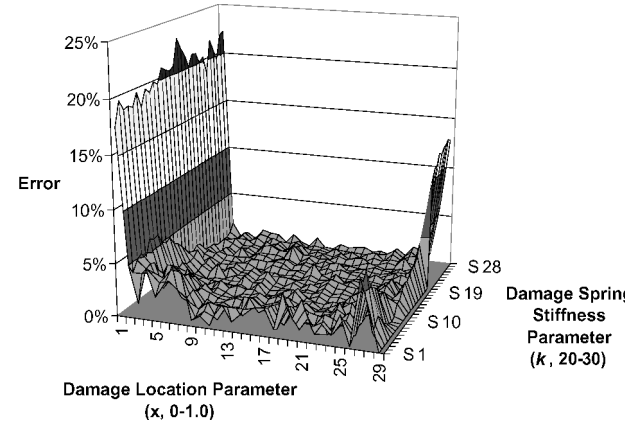


Fig. 6 Simulation errors of a trained NN at testing points in identifying two damage parameters when the first three natural frequencies of the damaged beam are given (841 testing data points, $x = 29$ and $k = 29$).

Selecting the location x and the stiffness k values for 30 points each resulted in 900 training data points. To train NNs for 900 points, a radial-basis NN was used instead of a multilayer feedforward NN because a radial-basis NN is better suited for the case in which a large number of testing data points are available.

A radial-basis NN has one input layer, one hidden layer, and one output layer, and has a radial-basis function as the transfer function. The difference between a feedforward NN and a radial-basis NN is that, in the case of a feedforward NN, an NN with a random parameter is created, and the NN so created is trained with the training data set. However, in the case of a radial-basis NN, the radial-basis NN is created with a training data set that has near-zero error vectors. When a multilayer feedforward NN was tried, it took too much time to train that NN. Therefore, the radial-basis NN is better suited for a case with a large number of training data points. The "newrb" function in MATLAB creates a radial-basis NN by adding neurons to the network until the sum-squared error falls below an error goal, or a maximum number of neurons have been reached. Figure 6 shows the case when the NN, which is trained with 900 data sets, was tested using the "test" data points.

Moreover, an NN was trained to identify three parameters. The training was done for 20 locations of x, eight values of span size b between 0.05 and 0.2, and eight different values of the damage-magnitude parameter β between 0.1 and 0.5. In this case, a multilayer feedforward NN was used with resilient backpropagation training. The trained NN was tested for a new set of data that lies between the training points. The trained NN showed only a relative error of 1% for the testing points, as shown in Fig. 7.

The effect of measurement error in the frequency inputs was studied. The noise was simulated by introducing errors in the

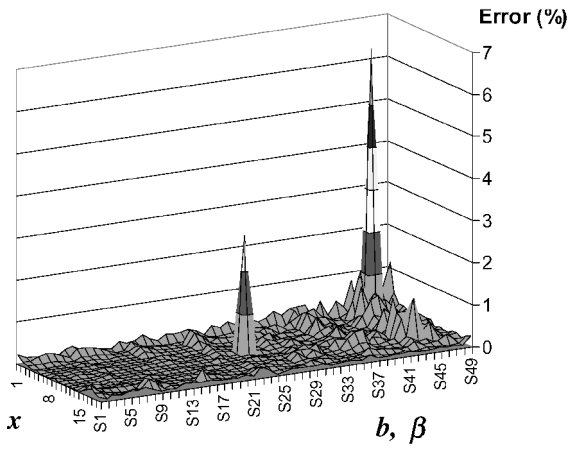


Fig. 7 Simulation errors of a trained NN at testing points in identifying three parameters (931 testing points, $x = 19$, $b = 7$, and $\beta = 7$).

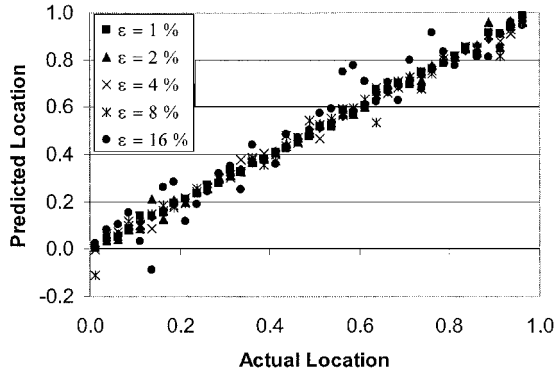


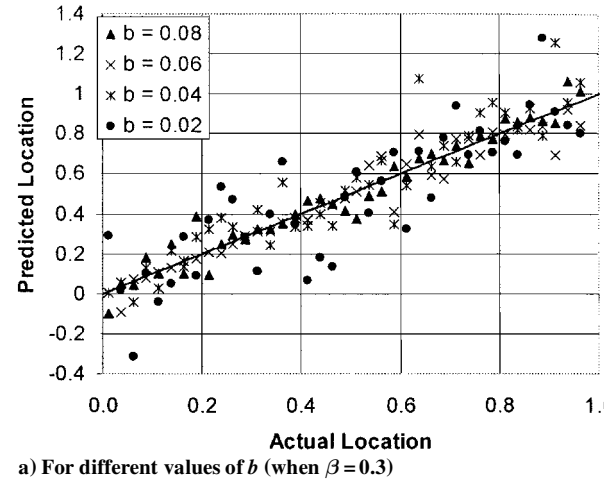
Fig. 8 Prediction of the location of damage for different degrees of error in the measurement of the difference of frequency between damaged and undamaged beam ($b = 0.05$ and $\beta = 0.3$).

measurement of the difference between the natural frequencies of damaged and undamaged beam for the first five modes, and these error-infested data were used as inputs for the trained NNs to predict the location of damage. Figure 8 shows the predicted location of damage vs the actual location for different degrees of errors in the measurements of the difference between the natural frequency of the damaged beam and that of the undamaged beam. The straight line with a slope of unity shows the points with no errors in predicting the damage location. From Fig. 8, we can see that the error in the measurement should be less than 8% for NNs to predict the location of damage correctly.

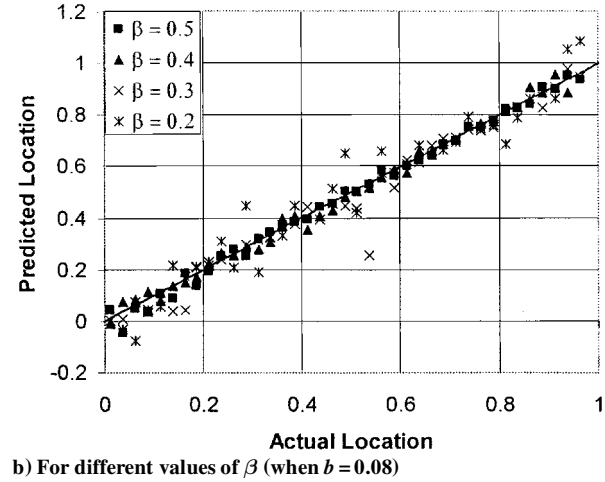
Next, we investigated how well the damage in a beam can be predicted using error-infested frequency data for different values of the span and the degree of damage. Figure 9a shows the prediction of the location of damage for different values of the span of damage b. Figure 9b shows the prediction of the location for different values of the degree of damage β. The straight lines with a slope of unity also show the points with no errors in predicting the damage location. From Figs. 8 and 9, we can see that a trained NN can predict the location of the damage even though errors are present in the input. However, the span of the damage should be larger than 8% of the whole length of the beam, and the reduction in the bending stiffness should be larger than 20% of the bending stiffness value of an undamaged beam for the present NN-based approach to identifying the damage.

Optical Imaging Problem

The first step in an optical imaging problem is solving the direct problem, which is, for the given location and strength of a light source, calculating the intensity of light or photon time of flight measurements at detector locations. This can be done using an FEM that solves that photon-diffusion equation. The photon-diffusion equation is known to be a very good approximation of photon transport



a) For different values of b (when $\beta = 0.3$)



b) For different values of β (when $b = 0.08$)

Fig. 9 Prediction of the location of damage for different values of the span and the degree of the damage ($\varepsilon = 8\%$).

in human tissue.[30] A diagram depicting the direct problem solving is shown in Fig. 10.

The photon-diffusion equation, which is the approximate solution of the photon transport equation, is shown hereafter. This is valid for the scattering dominant case with an isotropic source condition[28]:

$$\frac{1}{c_n}\frac{\partial}{\partial t}\Phi(\boldsymbol{r}, t) - \nabla \cdot [D(\boldsymbol{r})\nabla\Phi(\boldsymbol{r}, t)] + \mu_a(\boldsymbol{r})\Phi(\boldsymbol{r}, t) = S(\boldsymbol{r}, t) \quad (6)$$

where $\Phi(\boldsymbol{r}, t)$ is the photon intensity at position $\boldsymbol{r}$ and time t, $D(\boldsymbol{r})$ is the diffusion constant, $F(t)$ is the diffusion intensity, $S(\boldsymbol{r}, t)$ is the source strength, $\mu_a(\boldsymbol{r})$ is the absorption coefficient, and c_n is the speed of light in the medium. The solution of the light-diffusion equation can be obtained by the FEM using a Galerkin approach, as follows (see Ref. 30):

$$[\boldsymbol{K}(\kappa) + \boldsymbol{C}(\mu_a, c)]\boldsymbol{\Phi} + \boldsymbol{B}\frac{\partial \boldsymbol{\Phi}}{\partial t} = \boldsymbol{Q}$$

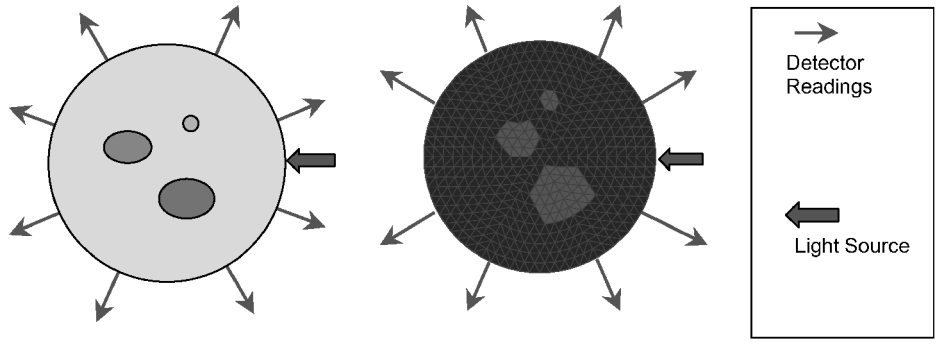
$$K_{ij} = \int_\Omega \kappa(\boldsymbol{r})\nabla\psi_j(\boldsymbol{r}) \cdot \nabla\psi_i(\boldsymbol{r})\, d\Omega$$

$$C_{ij} = \int_\Omega \mu_a(\boldsymbol{r})c\psi_j(\boldsymbol{r})\psi_i(\boldsymbol{r})\, d\Omega$$

$$B_{ij} = \int_\Omega \psi_j(\boldsymbol{r})\psi_i(\boldsymbol{r})\, d\Omega, \qquad Q_j(t) = \int_\Omega \psi_j(\boldsymbol{r})q_0(\boldsymbol{r}, t)\, d\Omega$$

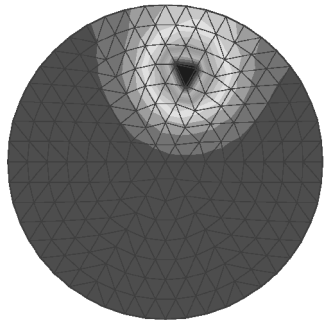
$$\boldsymbol{\Phi} = [\Phi_1(t), \Phi_2(t), \ldots, \Phi_D(t)]^T, \qquad \kappa(\boldsymbol{r}) = \frac{c}{3[\mu_a(\boldsymbol{r}) + \mu_s'(\boldsymbol{r})]} \quad (7)$$

where $\boldsymbol{K}$, $\boldsymbol{C}$, and $\boldsymbol{B}$ are the system matrices, $\boldsymbol{Q}$ represents source values, ψ_i are the element interpolation functions, κ is the diffusion

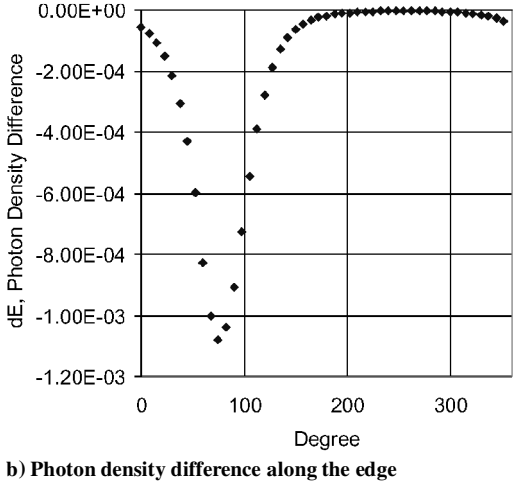


Actual experimental reading **Calculation using FEM**

Fig. 10 Solution of the direct problem.



a) Photon density difference contour



b) Photon density difference along the edge

Fig. 11 Photon density difference with and without the presence of an absorption anomaly (optical properties of tissue: medium, $\mu_a = 0.025\ \text{mm}^{-1}$ and $\mu_s = 2.0\ \text{mm}^{-1}$; and anomaly, $\mu_a = 0.05\ \text{mm}^{-1}$).

coefficient, Ω is the tissue domain, μ_a is the absorption coefficient, μ_s is the reduced scattering coefficient, and c is the speed of light.

For the time-independent case, expression (7) is reduced as follows[30]:

$$[\boldsymbol{K}(\kappa) + \boldsymbol{C}(\mu_a, c)]\boldsymbol{\Phi} = \boldsymbol{Q} \tag{8}$$

To be able to solve the inverse problem using an NN, a large enough number of training cases should be generated to train an NN to solve the inverse problem correctly.

The information used as an input to the NN is the light intensity distribution along the boundary of a tissue. Figure 11 shows the clear differences in the light intensity distribution inside a tissue with and without the presence of an anomaly.

Optical Imaging Results

Various methods and complexities were utilized in using NNs in identifying an anomaly in a circular-shaped tissue. A simple problem that only identifies one parameter was solved first, and the complexity was increased gradually to the case where the number of outputs is same as the number of elements in an FEM. Throughout the solution, a multilayer feedforward NN with two sigmoid hidden layers and linear output layers, as well as a resilient backpropagation training scheme, were used. Here, the tissue is assumed to have a large degree of scattering coefficient, which is the valid assumption for which to use a photon-diffusion equation to approximate photon transport in tissue. The anomaly is assumed to have different scattering coefficient and should have a larger absorption coefficient. The exact values of coefficients are different from problem to problem and are noted in the figure captions.

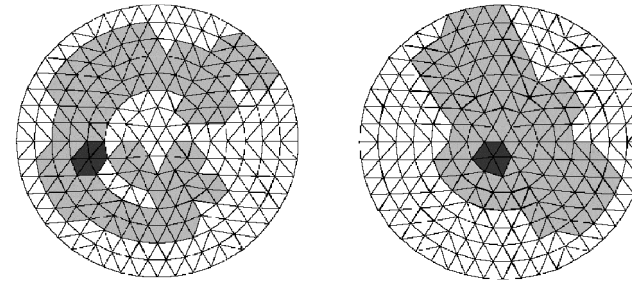


Fig. 12 Examples of an anomaly in a nonhomogeneous medium with different shapes of inside composition [tissue properties: white area: medium, $\mu_a = 0.01\ \text{mm}^{-1}$ and $\mu_s' = 1.0\ \text{mm}^{-1}$; gray area: parenchyma, $\mu_a = 0.03\ \text{mm}^{-1}$, and $\mu_s' = 0.5\ \text{mm}^{-1}$; and dark area: anomaly (tumor), $\mu_a = 0.5\ \text{mm}^{-1}$ and $\mu_s' = 3.0\ \text{mm}^{-1}$].

Preliminary results were obtained using an optical finite element program available at the Biomedical Optic Research Group's homepage, URL:http://www.medphys.ucl.ac.uk/research/borg/index.htm. However, we subsequently developed our own finite element code to solve the direct problems. Because utilizing an existing finite element code had limitations with regard to the code's integration with the MATLAB Neural Network Toolbox, such as in generating a large enough number of training cases, we have developed our own finite element code for solving the direct problem in two dimensions as part of solving the inverse problem. The code was developed in the MATLAB environment for easy and efficient integration with the MATLAB Neural Network Toolbox. The finite elements used in our code are linear triangular elements to accommodate round and irregular shapes. In evaluating element matrices, exact integral formulas were used.[39]

In human tissue, the rough shape of the cross section of a given tissue would be similar from person to person; however, there would be a variety of variations in the cross sections size and detailed shape. Thus, we wanted to find out whether a trained NN would be able to discern the existence of an anomaly in a nonhomogeneous medium with known absorption and scattering properties, but having an unknown distribution of its constituents such as fat and parenchyma. Figure 12 provides examples of a nonhomogeneous medium with anomalies. The white-colored medium is the fat, and the light-gray-colored area is the parenchyma. The dark-colored

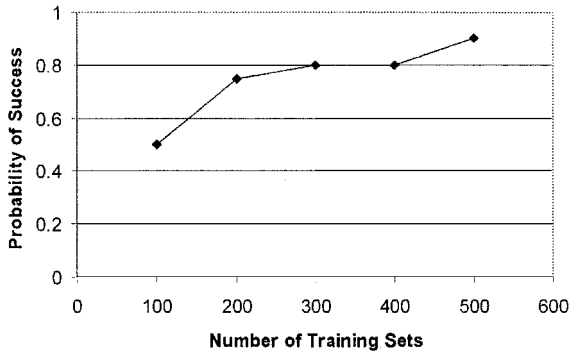


Fig. 13 Probability of successfully predicting the existence of an anomaly in a nonhomogeneous medium by a trained NN, for different numbers of training sets, which were randomly generated.

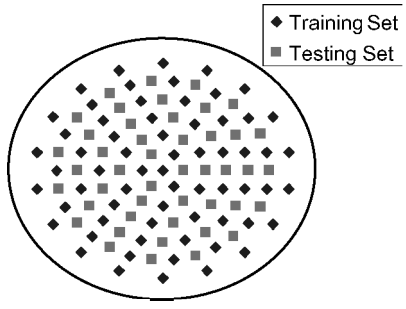


Fig. 14 Distribution of training and testing set for an NN that predicts the location and the size of an anomaly (optical properties of tissue: medium, $\mu_a = 0.025\ \text{mm}^{-1}$ and $\mu_s' = 2.0\ \text{mm}^{-1}$; and anomaly, $\mu_a = 0.05\ \text{mm}^{-1}$).

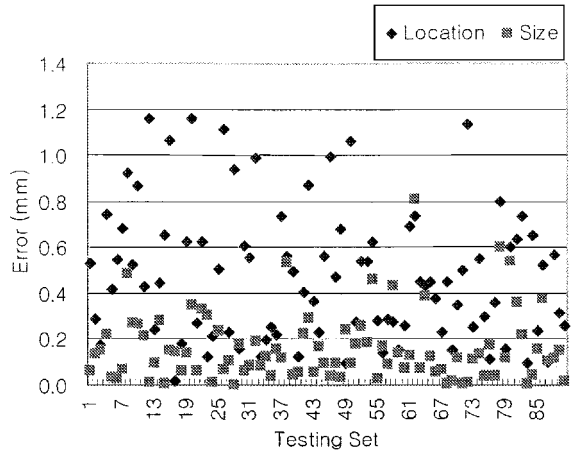


Fig. 15 Simulation error of an NN in predicting the location and the size of an anomaly in 384 elements finite element model.

region is the anomaly. The NN was trained with randomly selected parenchyma locations and shapes and both with and without anomalies. After training, the probability of a correct prediction was found to be only 50% with an NN trained with 100 training sets. However, this probability increased to 70% for an NN trained with 200 samples and to 90% for an NN trained with 500 samples. Figure 13 shows the increase in the probability of correctly predicting the presence of an anomaly as the training set size increases. From this, we can see that it is possible for a trained NN to tell whether a region having an abnormal property exists in a tissue of known absorption and scattering properties but with an unknown distribution of fat and parenchyma.

In the next case, an NN was trained to identify three anomaly parameters, namely, the location and the size of an anomaly. They are two coordinate values of the center location of the anomaly and the radius of the anomaly. The training was done based on training sets distributed evenly over the cross section. Figure 14 shows the locations of testing and training points. The training succeeded in converging to a simulation error of 10^{-5}. Figure 15 shows that the simulation error displaying the deviation in location prediction was mostly less than 1 mm in a sample with the overall diameter of 50 mm. From these results, we can expect that higher accuracy can be achieved by using a higher number of training sets and a finer mesh.

Next, we trained an NN in a different way. We made each output neuron represent the absorption coefficient of each element in the FEM, as shown in Fig. 16. In this way, drawing contours based on the values of the simulation output of neurons corresponding to each finite element can reveal the intrinsic tomography of the section.

We used two finite element meshes having different resolutions, for solving the direct and inverse problems, respectively, for restudying the aforementioned case. Recall that, in that case, each output

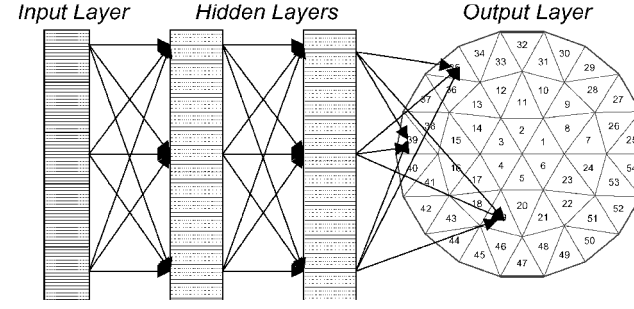
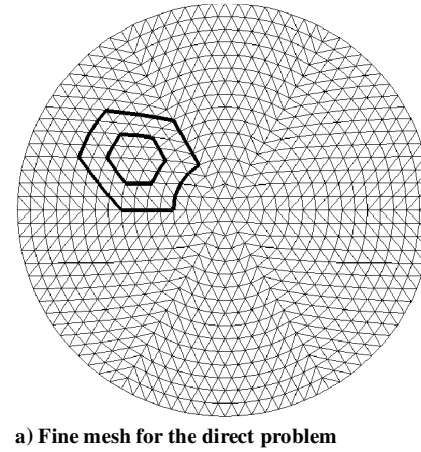
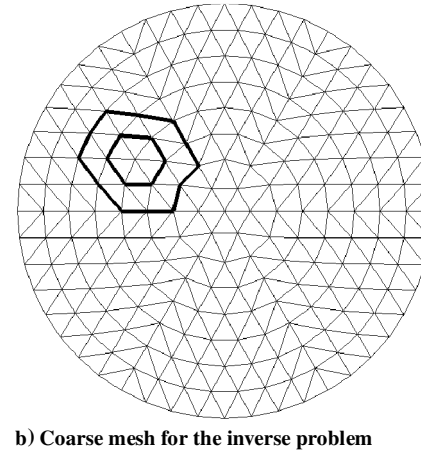


Fig. 16 Structure of multilayer feedforward NNs to identify the property of each element; each output neuron corresponds to the parameters of each element.



a) Fine mesh for the direct problem



b) Coarse mesh for the inverse problem

Fig. 17 Fine mesh for direct problem solving and coarse mesh for inverse problem solving, showing two sizes of anomalies, corresponding to each other's mesh used in the training (fine mesh, 1536 elements; coarse mesh, 384 elements; optical properties of tissue: medium, $\mu_a = 0.025\ \text{mm}^{-1}$ and $\mu_s = 2.0\ \text{mm}^{-1}$; and anomaly, $\mu_a = 0.5\ \text{mm}^{-1}$).

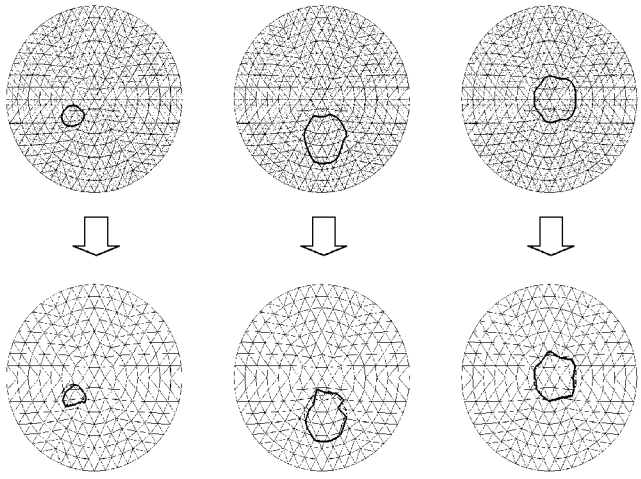


Fig. 18 Actual target and simulated output of trained NNs displaying the edges of an anomaly on fine and coarse finite element meshes in testing set: actual target and simulation output.

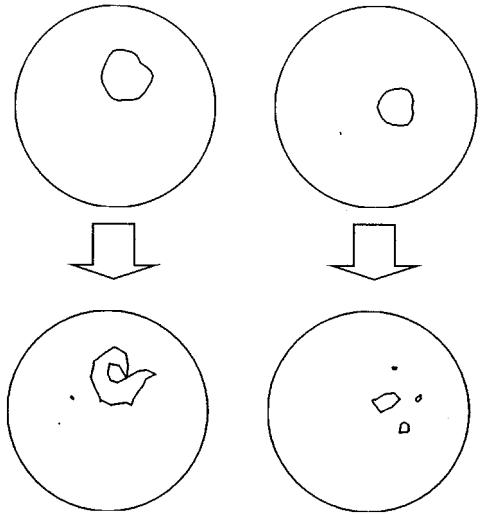


Fig. 19 Worst cases in the simulation results using different meshes for direct and inverse problems: actual target and simulation output.

neuron represents the absorption coefficient of an element in the finite element mesh. In Fig. 17, the two meshes are shown for the case with the presence of an anomaly in the tissue. The fine mesh, with the anomaly present, is used in solving the direct problem to simulate various detector readings. At the same time, the absorption coefficients of each element in the corresponding coarse mesh with the same shape of anomaly are used as the training targets for the NN's output layer. The training was done for every possible location in the coarse mesh, with two different sizes of the anomaly, as shown in Fig. 18. In this way, we have more accurate results for the direct problem and a smaller mapping space for the NN. At this point, the training took 3000 epochs. This is much shorter than the case when only a coarse mesh was used for obtaining both the data for simulating the measurements and for training the NN. When only a coarse mesh was used, it took more than 6000 epochs for training.

The trained NN was tested with new data obtained using the fine mesh and with an anomaly whose edges do not coincide with those of the anomaly that can be represented in a coarse mesh. In four out of five testing cases, the NN revealed very similar anomaly shapes. Figure 18 shows the close prediction given by a trained NN. Although Fig. 19 shows less accurate results, we can still observe that the NN gives good information about the anomaly. In Fig. 18, both the actual target in the testing set and the simulation result by the trained NNs to inputs corresponding to the actual target are shown on the fine and coarse meshes. Here, we can see that, although the trained NN has only seen the anomaly whose edges coincide with a coarse finite element mesh, the trained NN can still successfully predict the shape of an anomaly whose edges do not coincide with the edge of the coarse finite element mesh.

Conclusions

The first step in solving the inverse problem of damage detection and optical imaging is finding the best information that will most efficiently identify the parameters that we are seeking. In the damage identification problem, natural frequencies of a beam provided sufficient information. In the optical imaging problem, the intensity of light observed at various locations was utilized to both locate and determine the size of an anomaly.

The second step lies in finding the best way to employ NNs to solve the specific inverse problem. In both problems, feedforward backpropagation NNs and a resilient backpropagation training method have been found to be most efficient. Also, it was found that using radial-basis NNs in solving some aspects of the problem can have benefits. Causing output neurons to have a parameter value for each finite element, as was done in the optical imaging case, can have applications for the damage identification problem as well. Also, the damage identification problem in two dimensions would need a step similar to that in the optical imaging problem. In both of the problems, for NNs to predict accurate values, the generation of adequate amounts of training sets, as well as long learning processes, are needed.

When NNs are applied for solving inverse problems related to optical imaging, the most promising feature that the NNs provide seems to be their ability to classify the existence of an anomaly in a tissue in a manner that inspires confidence. Likewise, the NNs provide information related to both the geometry and the optical properties of a constituent in a tissue. This would also be the case for the damage identification problem wherein a well-trained NN can provide both the location and the extent of damage.

This work provides an example of applying a technology (identifying damage using NNs), developed in the field of structural engineering, to a problem in a totally different field, as well as applying the experience gained in the second field back to the structural engineering field. Eventually, we will see the benefits of solving inverse problems in both of the fields.

For the damage identification problem, it was seen that the present approach works extremely well if the error in the differences between the damaged and undamaged frequency is less than 16%, the size of the damage is more than 6–8%, and the reduction in the stiffness is 20% or more.

Acknowledgments

The authors wish to thank the Carilion Biomedical Institute, Roanoke, Virginia, and the Optical Sciences and Engineering Research Center at Virginia Polytechnic Institute and State University, Blacksburg, Virginia, for their valuable support.

References

[1]Haykin, S. S., *Neural Network: Fundamental Foundation*, 2nd ed., Prentice–Hall, Upper Saddle River, NJ, 1998, pp. 256–317.

[2]Kudva, J. N., Munir, N., and Tan, P. W., "Damage Detection in Smart Structures Using Neural Networks and Finite-Element Analyses," *Smart Materials and Structures*, Vol. 1, No. 2, 1992, pp. 108–112.

[3]Wu, X. J., Ghaboussi, J., and Garrett, J. H., "Use of Neural Networks in Detection of Structural Damage," *Computers and Structures*, Vol. 42, No. 4, 1992, pp. 649–659.

[4]Leath, W. J., and Zimmerman, D. C., "Analysis of Neural Network Supervised Training with Application to Structural Damage Detection," *Proceedings of the 9th Virginia Polytechnic Institute and State University Symposium on Dynamics and Control of Large Structures*, Virginia Polytechnic Inst. and State Univ., Blacksburg, VA, 1993, pp. 583–594.

[5]Spillman, W., Huston, D., Fuhr, P., and Lord, J., "Neural Network Damage Detection in a Bridge Element," *Proceedings of SPIE Vol. 1918, Smart Structures and Materials 1993: Smart Sensing, Processing, and Instrumentation*, edited by R. O. Claus, Society of Photo-Optical Instrumentation Engineers—The International Society for Optical Engineering, Bellingham, WA, 1993, pp. 288–295.

[6]Worden, K., and Tomlinson, G. R., "Damage Location and Quantification Using Neural Networks," *Engineering Integrity Assessment*, edited by J. M. Edwards, J. Kerr, and P. Stanley, Chameleon Press, London, 1993, pp. 11–33.

[7]Kirkegaard, P., and Rytter, A., "Use of Neural Networks for Damage Assessment in a Steel Mast," *Proceedings of the 12th International Modal Analysis Conference*, Society of Photo-Optical Instrumentation Engineers—The International Society for Optical Engineering, Bellingham, WA, 1994, pp. 1128–1134.

[8]Manning, R., "Damage Detection in Adaptive Structures Using Neural Networks," *Proceedings of the 35th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, AIAA, Washington, DC, 1994, pp. 160–172.

[9]Povich, C., and Lim, T., "Artificial Neural Network Approach to Structural Damage Detection Using Frequency Response Functions," *Proceedings of 35th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, AIAA, Washington, DC, 1994, pp. 151–159.

[10]Tsou, P., and Shen, M. H. H., "Structural Damage Detection and Identification Using Neural Networks," *AIAA Journal*, Vol. 32, No. 1, 1994, pp. 176–183.

[11]Rhim, J., and Lee, S., "Neural Network Approach for Damage Detection and Identification of Structures," *Proceedings of 35th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, AIAA, Washington, DC, 1994, pp. 173–180.

[12]Barai, S. V., and Pandey, P. C., "Vibration Signature Analysis Using Artificial Neural Networks," *Journal of Computing in Civil Engineering*, Vol. 9, No. 4, 1995, pp. 259–265.

[13]Ceravolo, R., and De Stefano, A., "Damage Location in Structures Through a Connectivistic use of FEM Modal Analyses," *Modal Analysis: The International Journal of Analytical and Experimental Modal Analysis*, Vol. 10, No. 3, 1995, p. 176.

[14]Schwarz, B. J., McHargue, P. L., and Richardson, M. H., "Using SDM to Train Neural Networks for Solving Modal Sensitivity Problems," *Proceedings of the 14th International Modal Analysis Conference*, Society of Experimental Mechanics, Bethel, CT, 1996, pp. 1285–1291.

[15]Ganguli, R., and Chopra, I., "Helicopter Rotor System Fault Detection Using Physics-Based Model and Neural Networks," *AIAA Journal*, Vol. 36, No. 6, 1998, pp. 1078–1086.

[16]Marwala, T., "Damage Identification Using Committee of Neural Networks," *Journal of Engineering Mechanics*, Vol. 126, No. 1, 2000, pp. 43–50.

[17]Palakal, M. J., Pidaparti, R. M. V., and Rebbapragada, S., "Intelligent Computational Methods for Corrosion Damage Assessment," *AIAA Journal*, Vol. 39, No. 10, 2001, pp. 1936–1943.

[18]Szewczyk, P. Z., and Hajela, P., "Damage Detection in Structures Based on Feature-Sensitive Neural Networks," *ASCE Journal of Computing in Civil Engineering*, Vol. 8, No. 2, 1994, pp. 163–178.

[19]Klenke, S. E., and Paez, T. L., "Damage Identification with Probabilistic Neural Networks," *Proceedings of the 14th International Modal Analysis Conference*, Society of Experimental Mechanics, Bethel, CT, 1996, pp. 99–104.

[20]Luo, H., and Hanagud, S., "Dynamic Learning Rate Neural Network Training and Composite Structural Damage Detection," *AIAA Journal*, Vol. 35, No. 9, 1997, pp. 1522–1527.

[21]Doyle, C., and Fernando, G., "Detecting Impact Damage in a Composite Material with an Optical Fibre Vibration Sensor System," *Smart Materials and Structres*, Vol. 7, No. 4, 1998, pp. 543–549.

[22]Li, X. Q., Wong, Y. S., and Nee, A. Y. C., "Intelligent Tool Wear Identification Based on Optical Scattering Image and Hybrid Artificial Intelligence Techniques," *Journal of Engineering Manufacture*, Vol. 213, No. B2, 1999, pp. 191–196.

[23]Wu, Y. J., Shi, X. Z., and Zhuang, T. G., "Fusion of Wavelet Packets and Neural Network in Detection of Composites," *AIAA Journal*, Vol. 38, No. 6, 2000, pp. 1063–1069.

[24]Cerri, M. N., and Vestroni, F., "Detection of Damage in Beams Subjected to Diffused Cracking," *Journal of Sound and Vibration*, Vol. 234, No. 2, 2000, pp. 259–276.

[25]Müller, G. J., *Medical Optical Tomography: Functional Imaging and Monitoring*, SPIE Optical Engineering Press, Washington, DC, 1993, pp. 5–12.

[26]Arridge, S. R., "A Gradient-Based Optimization Scheme for Optical Tomography," *Optics Express*, Vol. 2, No. 6, 1997, pp. 213–226.

[27]Arridge, S. R., and Schweiger, M., "Image Reconstruction in Optical Tomography," *Philosophical Transactions: Biological Sciences*, Vol. 352, No. 1354, 1997, pp. 716–726.

[28]Barbour, R. L., Graber, H. L., Chang, J., Barbour, S. L., Koo, P. C., and Arronson, R., "MRI-Guided Optical Tomography: Prospects and Computation for a New Imaging Method," *IEEE Computational Science and Engineering*, Vol. 2, No. 4, 1993, pp. 63–77.

[29]Chang, J., Graber, H. L., Barbour, R. L., and Aronson, R., "Recovery of Optical Cross-Section Perturbations in Dense-Scattering Media by Transport-Theory-Based Imaging Operators and Steady-State Simulated Data," *Applied Optics*, Vol. 35, No. 20, 1996, pp. 3963–3978.

[30]Arridge, S. R., Schweiger, M., Hiraoka, M., and Delpy, D. T., "A Finite Element Approach for Modeling Photon Transport in Tissue," *Medical Physics*, Vol. 20, No. 2, 1993, pp. 299–309.

[31]Schweiger, M., and Arridge, S. R., "The Finite-Element Method for the Propagation of Light in Scattering Media: Frequency Domain Case," *Medical Physics*, Vol. 24, No. 6, 1997, pp. 895–902.

[32]Wang, L., and Jacques, S. L., "Monte Carlo Modeling of Light Transport in Multilayered Tissues in Standard C," Univ. of Texas M. D. Anderson Cancer Center, Houston, TX, 1992.

[33]Okada, E., Schweiger, M., Arridge, S. R., Firbank, M., Delpy, D. T., "Experimental Validation of Monte Carlo and Finite-Element Methods for the Estimation of the Optical Path Length in Inhomogeneous Tissue," *Applied Optics*, Vol. 35, No. 19, 1996, pp. 3362–3371.

[34]Chiu, Y. H., and Yau, S. F., "An Artificial Neural Network for Tomographic Reconstruction of Time Varying Object," *Proceedings of the International Symposium on Speech, Image Processing and Neural Networks*, Inst. of Electrical and Electronics Engineers, New York, 1994, pp. 409–412.

[35]Nakao, Z., Ali, F. E. A. F., and Chen, Y. W., "CT Image Reconstruction by Backpropagation," *Proceedings of the First International Conference on Knowledge-Based Intelligent Electronic Systems*, Inst. of Electrical and Electronics Engineers, New York, 1997, pp. 323–326.

[36]Kiliç, B., and Korürek, M., "A Finite Element Method Based Neural Network Technique for Image Reconstruction in Electrical Impedance Imaging," *Proceedings of the 2nd International Biomedical Engineering Days*, Inst. of Electrical and Electronics Engineers, New York, 1998, pp. 100–102.

[37]Aizenberg, I., Aizenberg, N., Hiltner, J., Moraga, C., and Meyer zu Bexten, E., "Cellular Neural Networks and Computational Intelligence in Medical Image Processing," *Image and Vision Computing*, Vol. 19, No. 4, 2001, pp. 177–183.

[38]Demuth, H., and Beale, M., *Neural Network Toolbox for use with MATLAB*, The MathWorks, Inc., Natick, MA, 2000, pp. 5.16, 5.17.

[39]Reddy, J. N., *An Introduction to The Finite Element Method*, 2nd ed., McGraw–Hill, New York, 1993, pp. 304–312.

J. P. Gore
Associate Editor