Research Article

Classification of Earthquake-Induced Damage for *R/C* Slab Column Frames Using Multiclass SVM and Its Combination with MLP Neural Network

Ali Kia and Serhan Sensoy

Department of Civil Engineering, Eastern Mediterranean University, Gazimagusa, Mersin 10, Turkey

Correspondence should be addressed to Ali Kia; ali_kia_cien@yahoo.com

Nonlinear time history analysis (NTHA) is an important engineering method in order to evaluate the seismic vulnerability of buildings under earthquake loads. However, it is time consuming and requires complex calculations and a high memory machine. In this study, two networks were used for damage classification: multiclass support vector machine (M-SVM) and combination of multilayer perceptron neural network with M-SVM (MM-SVM). In order to collect data, three frames of *R/C* slab column frame buildings with wide beams in slab were considered. For NTHA, twenty different ground motion records were selected and scaled to ten different levels of peak ground acceleration (PGA). Thus, 600 obtained data from the numerical simulations were applied to M-SVM and MM-SVM in order to predict the global damage classification of samples based on park and Ang damage index. Amongst the four different kernel tricks, the Gaussian function was determined as an efficient kernel trick using the maximum total accuracy method of test data. By comparing the obtained results from M-SVM and MM-SVM, the total classification accuracy of MM-SVM is more than M-SVM and it is accurate and reliable for global damage classification of *R/C* slab column frames. Furthermore, the proposed combined model is able to classify the classes with low members.

1. Introduction and Background

Artificial neural networks (ANNs) are one of the popular computational models applied widely throughout the sciences. Also they are specially used in many fields of civil engineering like materials strength prediction, thermographic inspection of electrical installations within buildings, traffic management and transportation systems, forecast water pressure in pipes, and so forth. Generally ANNs are used to solve complex problems by considering effective indices and establish a good relationship between the input and output parameters. Moreover, these networks can be applied in classification problems. The first classification algorithm was presented by Fisher [1]. In this algorithm, minimizing the classification error of train data was evaluated as an optimization criterion. This method has been used in many classification algorithms, yet there are some problems encountered mainly the generalization properties of the classifiers, which are not directly involved in the cost function. Also for doing the training process, determining the structure of the network was not easy. As an example, to determine the optimum number of neurons in the hidden layers of the multilayer perceptron (MLP) neural networks or the number of Gaussian functions in radial basis function (RBF) neural networks are difficult and time consuming. Cortes and Vapnik [2] introduced a new learning statistical theory which leads to presenting the support vector machines (SVMs). The significant features of these networks are their ability to minimize the classification errors, maximize the geometric margins between classes, design the classifiers with maximum generalization, and automatically determine the architecture of network for classifiers and modeling the nonlinear separator functions using nonlinear cores.

In recent years, several different neural networks such as SVM have been applied in different branches of civil engineering. In a tunnel construction, an intelligent controlling system was presented by Jun et al. [3]. This system needed to recognize the geophysical parameters to find the optimum

solution of problems. Therefore, a nonlinear optimization technique was employed using the least square support vector machine (LSSVM). The results showed that this method is timesaving and easy to use in local optimal problems. Mingheng et al. [4] employed several different models of traffic flow using SVM to find the best intelligent traffic control tool. They obtained that amongst the three proposed models, the SVM with the historical pattern data for the target road section model has the best performance. Vafaei et al. [5] applied MLP neural network to identify the real-time seismic damage for concrete shear walls. It was observed that the neural network could detect the amount of imposed damage with high accuracy. Two different neural networks, the adaptive neuro-fuzzy inference system (ANFIS) and the three-layered artificial neural network (TL-ANN) model, were used to estimate the earthquake load reduction factor for industrial structures by Ceylan et al. [6]. They showed that the ANFIS model was more successfully than the TL-ANN model. Xie et al. [7] investigated the amount of voids inside the concrete using SVM. The grid-algorithm and the genetic-algorithm were used to determine the kernel function and network parameters. The obtained results presented that the SVM exhibits a promising performance for identification of voids inside the reinforced concrete. In addition, ANNs were used in conjunction with each other. Köroğlu et al. [8] applied MLP neural network in two models: single MLP and combined MLP with itself (CMLP) for estimation of the flexural capacity of the quadrilateral FRP-confined R/C columns. They obtained the model of CMLP having lower prediction error than the single MLP model. In order to classify the cardiac arrhythmias, Castillo et al. [9] considered a hybrid intelligent system which consists of the fuzzy K-nearest neighbors with the MLP and a very high classification rate was obtained. To predict the short-term wind power generation, combination of genetic algorithm (GA) and orthogonal least squares (OLS) algorithm with RBF neural network was proposed by Chang [10]. The test results indicated that the proposed model is reliable with the sufficient performance.

The main aim of this research is to classify the vulnerability of R/C frames taken from slab column frame buildings built in Famagusta, Cyprus, by using ANNs. The distinguishing characteristics of this building type are the rectangular columns and the slabs supported on columns with wide beams in slab. Since the nonlinear time history analysis (NTHA) is time-consuming and imposes a significant computational burden, two networks include M-SVM and MM-SVM were proposed as good alternatives and efficient networks. Several common kernel tricks were tested to determine the best kernel function and applied to these machines for learning process. Therefore, using the M-SVM and MM-SVM, the classes of global damage for the similar frames can be predicted easily and earthquake-induced damage can be prevented by retrofit plan.

2. Methodology

2.1. Nonlinear Time History Analysis (NTHA).
Nonlinear time history analysis (NTHA) is one of the most common numerical analysis methods for evaluation of building's behavior under seismic loads. It can be applied to a system with single or multifreedom degrees. In this method, several ground motion records are selected based on the soil condition of a specific zone, distance from fault line, the time duration of earthquakes, amplitude and frequency content, and so forth. Also the dynamic responses in the incremental state can be determined by scaling the records based on the different levels of ground motion parameter such as peak ground acceleration (PGA), peak ground velocity (PGV), or peak ground displacement (PGD), which is called incremental nonlinear time history analysis (INTHA).

2.2. Sample Frames and Material Properties.
One of the existing building types in Cyprus is slab-column frame buildings with wide beams and rectangular columns. In this study, three R/C slab-column frames with the 4, 6, and 8 levels were selected from this type of buildings which is representative of the midrise frames in Famagusta, Cyprus. These buildings were designed according to 1975 version of the Turkish seismic design code [11]. The compressive strength of concrete, yield, and ultimate strength of steel were specified based on the previous researches equal to 15 MPa, 220 MPa, and 300 MPa, respectively [12]. The properties of soil type IV (D) were considered for this zone. Based on the detailed information mentioned in the building documents, the rectangular columns with aspect ratio between 2 and 3 (height/width ratio of cross section area) were used. Additionally, the beams heights were equal to slab thickness (around 15 cm) and were used as connection elements between columns. Figure 1 depicts the plan views for the 4-, 6-, and 8-story frames.

2.3. Ground Motions.
A significant step for doing the NTHA is selecting a set of ground motion records. In this study, due to the lack and uncertainty of ground motion records for the Famagusta region, twenty records were selected carefully which their average had the most correlation with Turkish design code [13]. This set of records was chosen based on the D-type site properties, the strike-slip fault mechanism, and the distance less than 100 km from the fault line [14]. The records were taken from the Berkeley database site [15] and then multiplied by factor of 2.3 in order to fit with the Turkish design spectrum using root mean square error reduction technique. The characteristics of these ground motion records are tabulated in Table 1. Additionally, the Turkish design spectrum [13], the mean, and the response spectrums of these scaled records are shown in Figure 2.

2.4. Damage Identification Method.
Several different criterions such as ductility ratio [16], interstory drift [17], flexural damage ratio [18], and maximum permanent drift [19, 20], were introduced in order to evaluate the damage level of buildings. These criteria calculated the amount of structural vulnerability only based on a proper theoretical background. Later, mathematical models of damage that have been determined based on vulnerability of buildings under the earthquakes were defined as the functions of structural

TABLE 1: Characteristics of the twenty natural records.

| Name	Event	Year	Time effective (s)	Dominant frequency (Hz)	PGA (g)	PGA/PGV (g/cm·s^{-1})
TH1	Park field	1966	27.80	0.708	0.059	0.010
TH2	Park field	1966	06.99	0.781	0.476	0.006
TH3	Imperial Valley-06	1979	12.82	0.269	0.171	0.004
TH4	Imperial Valley-06	1979	23.32	0.220	0.078	0.006
TH5	Victoria-Mexico	1980	10.64	4.297	0.101	0.013
TH6	Victoria-Mexico	1980	15.37	1.489	0.150	0.006
TH7	Westmorland	1981	08.40	3.101	0.171	0.029
TH8	Westmorland	1981	18.50	1.196	0.155	0.006
TH9	Morgan Hill	1984	35.98	0.867	0.032	0.006
TH10	Superstition Hills-B	1987	16.86	0.830	0.211	0.007
TH11	Superstition Hills-B	1987	28.75	0.488	0.207	0.006
TH12	Superstition Hills-B	1987	16.05	1.538	0.358	0.008
TH13	Landers	1992	36.32	1.172	0.136	0.012
TH14	Landers	1992	17.62	0.720	0.245	0.005
TH15	Kobe-Japan	1995	24.52	1.575	0.070	0.016
TH16	Kocaeli-Turkey	1999	15.34	0.269	0.268	0.004
TH17	Kocaeli-Turkey	1999	09.39	1.062	0.242	0.008
TH18	Kocaeli-Turkey	1999	14.99	1.929	0.152	0.007
TH19	Duzce-Turkey	1999	19.22	0.732	0.042	0.005
TH20	Duzce-Turkey	1999	16.09	3.284	0.114	0.010
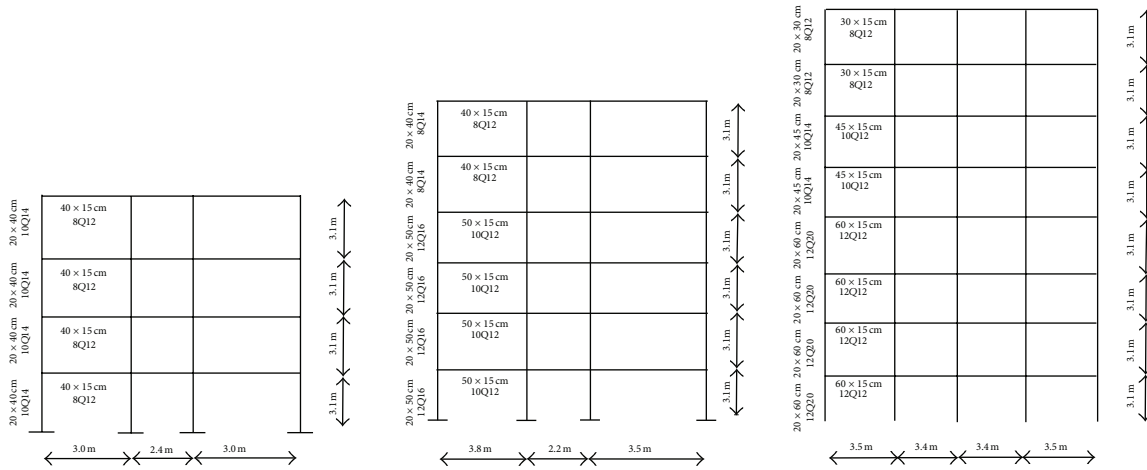
FIGURE 1: Plan view of four-, six-, and eight-story frames.

members' ductility, strength of materials, distance from the fault line, effective time duration of the earthquake, and so forth. Therefore it leads to improved damage function according to theoretical calculations and practical tests. In this study, Park et al. damage index was selected among the improved indices and is defined as the linear combination of the maximum displacement and the dissipated energy [21]. This index is expressed in the following equation:

$$DI = \frac{\delta_m}{\delta_u} + \frac{\beta}{\delta_u \cdot P_y} \int dE_h, \tag{1}$$
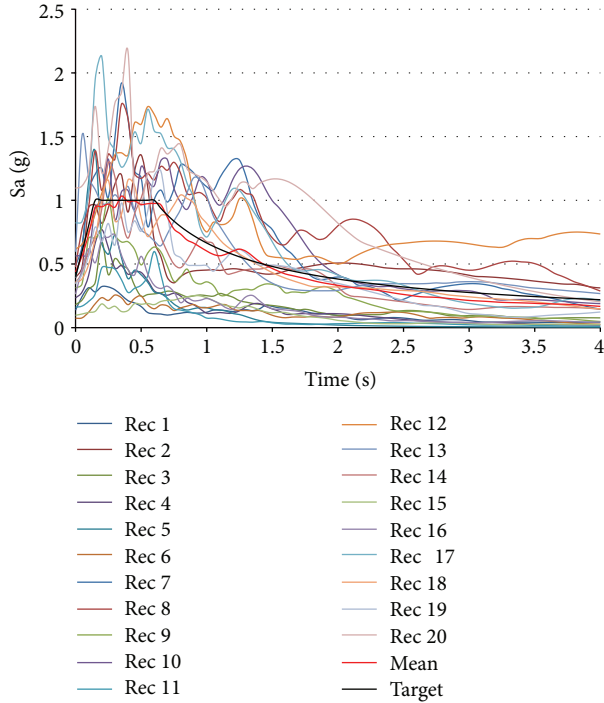
FIGURE 2: The response spectrums of individual scaled records, mean and Turkish design spectrum (Target).

TABLE 2: The classification of global damage based on the Park and Ang investigation.

Class number	State of structure	Amount of damage
Class 1	Repairable (economic)	DI ≤ 0.4
Class 2	Beyond repair (not economic)	0.4 < DI < 0.1
Class 3	Loss of building	DI ≥ 1

where δ_m and δ_u are the maximum deformation and ultimate deformation of element under monotonic loading, respectively; P_y is the yield strength of the structure, which can be calculated by nonlinear dynamic analysis; $\int dE_h$ is the hysteretic energy absorbed by the structural element during the response history; and β is a constant parameter. Therefore, based on this index, three classes of damages are used which is shown in Table 2.

3. Support Vector Machine

SVM has been introduced for the classification and pattern recognition problems by Cortes and Vapnik [2]. It is a relatively new learning algorithm used for binary classification problems. The main difference between SVM and the other algorithms is that the SVM minimizes the operational risk as an objective function instead of minimizing the classification error. The original pattern classification of this machine is to classify the linear input data using the perfect hyperplane into two classes with the largest margin in between classes. For nonlinear input data, a nonlinear mapping is used to transfer the input data from the primal space to the higher

dimensional feature space and lead to finding the proper hyperplane. Furthermore, SVMs have also been extended to solve multiclass problems.

3.1. Linear SVM. In this section, a simple introduction of the linear SVM is presented [22]. Consider a train sample data include $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, where each sample has the inputs ($x_i \in R^d$), and one class label ($y_i \in \{+1, -1\}$) which is shown in Figure 3.

In the two-dimensional space, the discriminator is a line in the middle of the margin between the classes. Thus, for N-dimension space, the discriminator is a hyperplane. Suppose the distance between the each separate data and the discriminator is equal to 1, the two support hyperplanes are considered parallel to the discriminator, and the classifier function can be obtained as follows (see Figure 3):

$$
\begin{aligned}
w^T \cdot x_i + b \geq 1, \quad &\text{if } y_i = 1 \; i = 1, 2, \ldots, n \\
w^T \cdot x_i + b \leq -1, \quad &\text{if } y_i = -1 \; i = 1, 2, \ldots, n.
\end{aligned}
\tag{2}
$$

For a unique separator, the maximum margin between classes is needed. Thus, if the distance between the support hyperplanes is equal to M, using (2), the optimum margin (M) is given by

$$
M = \frac{(|b + 1| - |b - 1|)}{\|w\|} = \frac{2}{\|w\|}.
\tag{3}
$$

After calculating the maximum margin, the target function is defined as follows:

$$
\begin{aligned}
\text{Maximize } (M) &= \text{Maximize} \frac{2}{\|w\|} \\
&= \text{Minimize} \|w\| = \text{Minimize} \frac{1}{2} \|w\|^2 \\
&= \text{Minimize} \frac{1}{2} w^T \cdot w.
\end{aligned}
\tag{4}
$$

$$
\begin{aligned}
\text{Subject to (s.t.):} \quad & w^T \cdot x_i + b \geq 1, \quad \text{if } y_i = 1 \\
& \qquad\qquad i = 1, 2, \ldots, n \\
& w^T \cdot x_i + b \leq -1, \quad \text{if } y_i = -1 \\
& \qquad\qquad i = 1, 2, \ldots, n.
\end{aligned}
\tag{5}
$$

Since the probability of being the separated data in nature is very low and more datasets are inseparable; therefore, the discriminator (hyperplane) is also determined based on minimum number of errors. As a result, those members belonging to another class are penalized based on the distance from the boundary of its own class (δ) (see Figure 3). This strategy is represented as a model of soft margin SVM. For
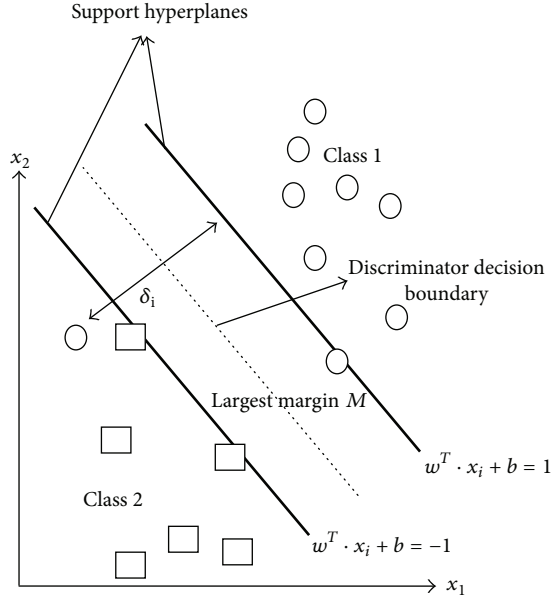
FIGURE 3: A sample of linear soft margin SVM.

this reason, nonnegative variables (δ_i) are defined and called as slack variable s.t. $\delta_i \geq 0$. Thus, (5) is changed as follows:

$$\text{Minimize} \quad \frac{1}{2} w^T \cdot w$$
$$\text{s.t.:} \quad w^T \cdot x_i + b \geq 1, \quad \text{if } y_i = 1$$
$$i = 1, 2, \ldots, n$$
$$w^T \cdot x_i + b \leq -1, \quad \text{if } y_i = -1 \quad (6)$$
$$i = 1, 2, \ldots, n$$
$$\delta_i \geq 0.$$

By multiplying both sides of first s.t. of (6) by y, the primal problem becomes

$$\text{Minimize} \quad \frac{1}{2} w^T \cdot w$$
$$\text{s.t.:} \quad y_i \left(w^T \cdot x_i + b \right) \geq 1 - \delta_i \quad (7)$$
$$i = 1, 2, \ldots, n, \ \delta_i \geq 0,$$

thus

$$L_p = \frac{1}{2} w^T \cdot w - \sum_{i=1}^{n} \left[y_i \left(w^T \cdot x_i + b \right) - 1 + \delta_i \right] \quad (8)$$
$$i = 1, 2, \ldots, n.$$

The primal problem is a quadratic program but it cannot be solved easily because it does not just depend on the parameters which are related to input vectors. Therefore, this equation changes from the primal form to dual form by using

the Lagrange method. The Lagrange factors (α_i, μ_i) must be nonnegative real coefficients and (8) becomes

$$L_p = \frac{1}{2} w^T \cdot w + C \sum_{i=1}^{n} \delta_i$$
$$- \sum_{i=1}^{n} \alpha_i \left[y_i \left(w^T \cdot x_i + b \right) - 1 + \delta_i \right]$$
$$- \sum_{i=1}^{n} \mu_i \cdot \delta_i, \quad \alpha_i, \mu_i \geq 0, \quad (9)$$

where C is penalty factor. In this case, L_p is a saddle point. Thus, at this point, the minimum value should be taken with respect to the parameters w, b, and δ and the maximum value should be taken with respect to the Lagrange multipliers (α_i, μ_i). This can be done by taking the partial derivative with respect to w, b, and δ in order to change the primal problem to a maximum problem as follows:

$$\frac{\partial L_p}{\partial w} = 0 \longrightarrow w - \sum_{i=1}^{n} \alpha_i \cdot y_i \cdot x_i = 0, \quad (10)$$
$$\frac{\partial L_p}{\partial b} = 0 \longrightarrow \sum_{i=1}^{n} \alpha_i \cdot y_i = 0, \quad (11)$$
$$\frac{\partial L_p}{\partial \delta} = 0 \longrightarrow C - \alpha_i - \mu_i = 0. \quad (12)$$

By substituting (10) and (11) into (9), the dual problem is obtained as follows:

$$L_D = -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \cdot \alpha_j \cdot y_i \cdot y_j \cdot x_i^T \cdot x_j + \sum_{i=1}^{n} \alpha_i. \quad (13)$$

Also based on (12), the box constrains are defined as

$$\begin{matrix} \alpha_i \geq 0 \\ \mu_i \geq 0 \\ \alpha_i + \mu_i = c \end{matrix} \longrightarrow \begin{matrix} 0 \leq \alpha_i \leq C \\ 0 \leq \mu_i \leq C. \end{matrix} \quad (14)$$

In addition, by considering h_{ij} and f as the following definition:

$$h_{ij} = y_i \cdot y_j \cdot x_i^T \cdot x_j,$$
$$f = \begin{bmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{bmatrix} \quad (15)$$

and substituting h_{ij} and f into (13), the dual formulation becomes

$$L_D = -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \cdot \alpha_j \cdot h_{ij} + \sum_{i=1}^{n} \alpha_i, \quad \alpha_i \geq 0, \quad (16)$$
$$L_D = -\frac{1}{2} \alpha^T \cdot H \cdot \alpha - f^T \cdot \alpha,$$

where H and α are defined as

$$H = [h_{ij}] = \begin{bmatrix} h_{11} & \cdots & h_{1n} \\ \vdots & \ddots & \vdots \\ h_{n1} & \cdots & h_{nn} \end{bmatrix}, \quad H \in R^{n \cdot n},$$

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}. \tag{17}$$

Therefore, the target function is expressed as follows:

$$\text{Minimize} \quad \frac{1}{2}\alpha^T \cdot H \cdot \alpha + f^T \cdot \alpha$$

$$\text{s.t. :} \qquad w^T \cdot x_i + b \geq 1, \qquad \text{if } y_i = 1$$

$$i = 1, 2, \ldots, n \tag{18}$$

$$w^T \cdot x_i + b \leq -1, \qquad \text{if } y_i = -1$$

$$i = 1, 2, \ldots, n,$$

$$\sum_{i=1}^{n} \alpha_i \cdot y_i = 0 \quad 0 \leq \alpha_i \leq C. \tag{19}$$

The quadratic programming problem (see (19)) can be solved easily by using the quadprog function in the Matlab software and the values of α_i are calculated. Then by substituting α_i values into (10), the values of w are obtained. Also for calculating the bias term, the Karush-Kuhn-Tucker (KKT) conditions [23] are necessary and sufficient for the optimization problems. Therefore, these conditions should be established in optimum point (see (9)). The bias value is calculated as

$$\text{K.K.T} \longrightarrow \alpha_i \left[y_i \left(w^T \cdot x_i + b \right) - 1 + \delta_i \right] = 0,$$

$$\mu_i \cdot \delta_i = (C - \alpha_i) \cdot \delta_i = 0. \tag{20}$$

Thus based on K.K.T conditions three cases occurred.

Case 1. None support vectors if ($\alpha_i = 0$)

$$\alpha_i = 0 \longrightarrow \begin{cases} \mu_i = C \longrightarrow \delta_i = 0 \\ y_i \left(w^T \cdot x_i + b \right) - 1 \geq 0. \end{cases} \tag{21}$$

Case 2. Outliers if ($\alpha_i = C$)

$$\alpha_i = C \longrightarrow \begin{cases} \mu_i = 0 \longrightarrow \delta_i \geq 0 \\ y_i \left(w^T \cdot x_i + b \right) - 1 + \delta_i = 0. \end{cases} \tag{22}$$

Case 3. Support vectors if ($0 < \alpha_i < C$)

$$0 < \alpha_i < C \longrightarrow \begin{cases} 0 < \mu_i < C \longrightarrow \delta_i = 0 \\ y_i \left(w^T \cdot x_i + b \right) - 1 = 0. \end{cases} \tag{23}$$

In Case 3, each x_i corresponding to α_i are support vector machines. Thus by multiplying both sides of first s.t. of (19) by y as follows:

$$w^T \cdot x_i + b = y_i, \tag{24}$$

the amount of the bias term can be obtained as follows:

$$S = \{ i \mid 0 < \alpha_i < C \}, \quad i \in S,$$

$$b = \frac{1}{|S|} \sum_{i=1}^{S} \left(y_i - w^T \cdot x_i \right), \tag{25}$$

and also using (10), w becomes

$$w = \sum_{i=1}^{n} \alpha_i \cdot y_i \cdot x_i. \tag{26}$$

Finally, by having the amounts of w and b, the optimal hyperplane decision function can be expressed as follows:

$$y_i = \text{Sign} \left(w^T \cdot x_i + b \right). \tag{27}$$

3.2. Nonlinear SVM. For nonlinearly data, the selection of optimal hyperplane for separation of data is difficult. For this case, Cortes and Vapnik [2] used the Hilbert-Schmidt theory [24] in order to transform the d-dimensional input vector x into (usually higher) an N-dimensional feature vector by using an N-dimensional vector function $\emptyset$:

$$\emptyset : R^d \longrightarrow R^N,$$

$$\emptyset : x \longrightarrow z, \tag{28}$$

$$z = \emptyset(x).$$

Therefore based on the SVM algorithm, the discriminator equation can be applied into z space instead of x space as follows:

$$w^T \cdot x + b = 0 \longrightarrow w^T \cdot z + b = 0$$

$$\longrightarrow w^T \cdot \emptyset(x) + b = 0. \tag{29}$$

And according to the properties of soft margin classifier method, the dual problem is obtained as follows:

$$L_D = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \cdot \alpha_j \cdot y_i \cdot y_j \cdot \emptyset\left(x_i^T\right) \cdot \emptyset\left(x_j\right) - \sum_{i=1}^{n} \alpha_i, \tag{30}$$

$$\alpha_i \geq 0,$$

by substituting $K(x_i, x_j)$ instead of $\emptyset(x_i^T) \cdot \emptyset(x_j)$, the dual formulation becomes

$$L_D = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \cdot \alpha_j \cdot y_i \cdot y_j \cdot K\left(x_i, x_j\right) - \sum_{i=1}^{n} \alpha_i, \tag{31}$$

$$\alpha_i \geq 0,$$

TABLE 3: Properties of kernel functions.

Kernel function	Expression	Comment
Linear	$K\left(x_i, x_j\right) = r + \alpha \cdot x_i^T \cdot x_j$	$\alpha > 0$
Polynomial	$K\left(x_i, x_j\right) = \left(r + \alpha \cdot x_i^T \cdot x_j\right)^p$	$\alpha > 0$
Gaussian	$K\left(x_i, x_j\right) = \exp\left(-\dfrac{1}{2\sigma^2}\left\|x_i - x_j\right\|^2\right)$	$\sigma \neq 0$
Sigmoid	$K\left(x_i, x_j\right) = \tanh\left(\beta_0 + \beta_1 \cdot x_i^T \cdot x_j\right)$	$\beta_1 > 0$

x_i, x_j are input vectors and $r, \alpha, \sigma, \beta_0,$ and β_1 are kernel parameters.

where $K(x_i, x_j)$ is kernel trick (nonlinear function) and is applied to change the linear discriminator model into nonlinear form. In this study, four common kernel tricks were applied in SVM in order to find the best kernel function including linear kernel function, polynomial kernel function, Gaussian kernel function, and sigmoid kernel function as presented in Table 3.

Therefore, the optimal hyperplane decision function is expressed as follows:

$$y = \text{sign}\left(\sum_{i=1}^{n} \alpha_i \cdot y_i \cdot K\left(x_i, x\right) + b\right). \tag{32}$$

3.3. Multiclass SVM (M-SVM). The basic theory of SVM is designing the discriminator (hyperplane) with maximum margin between the two classes, while most of classification problems are in the multiclass models [25]. For N classes' model, Cortes and Vapnik [2] presented a strategy to compare one class with the remaining classes and this leads to generating the N classifiers. Therefore, this method needs the solution of the N quadratic programming optimization problems. This strategy can be named as "one-versus-rest" and was used in this research.

4. Data Generation, Selection of Input Parameters

For classification of the imposed global damage of these sample frames, M-SVM and MM-SVM were employed. In order to do so, twenty suitable ground motion records are selected and scaled to ten levels of PGA, then each scaled record applied to each frames and the amount of damage were obtained using IDARC-2D software [26]. After carrying out the NTHA, 600 data were generated and divided into three classes based on the Park and Ang damage definition (see Table 2) including 153, 62, and 385 data for class number 1, class number 2, and class number 3, respectively.

For generation of each input data, the appropriate parameters should be selected that are able to describe the properties of *R/C* frames and ground motion characteristic. In this investigation, seven and four parameters were chosen for structural and ground motion properties, respectively. The structural parameters were selected based solely on geometry and dimension of sample frames and without any relation to the engineering analysis such as first mode period, ductility and energy absorption of structural elements, and top displacement. In addition, for ground motion parameters, the main and available characteristics of records were selected. The ranges and definition of parameters are shown in Table 4.

The scaling of the data set is very important for training and also testing process of network. Thus, before presenting the data to the network, it is advised to normalize them. Therefore, linear normalization method was used to change the input parameters range between zero and one.

5. Models Used for Classification

In this study, two models of neural networks consisting of M-SVM and MM-SVM were applied. In order to find the best kernel function for training process, the total accuracy prediction scores of the test data were calculated. Also, the kernel parameters (p, r, α, σ, β_0, and β_1) and penalty factor (c) should be determined to reach the maximum margin between classes and the minimum classification error between real and predicted data. The amounts of p, r, α, β_0 and β_1 were obtained using trial and error. Also for the two remaining parameters (c and σ), the grid-search method was considered and the best values were selected automatically using Libsvm-3.17 [27] in the Matlab software. The results showed that 143, 149, 155, and 152 class labels from the total of 180 test data class labels were correctly predicted for linear, polynomial (5 degree), Gaussian and sigmoid functions, respectively. Therefore the Gaussian function was chosen as the best kernel trick function.

5.1. M-SVM Model. In the M-SVM, the set of normalized data which include 600 input data and each data containing eleven elements were shuffled and then applied to this machine that 70% and 30% of the total data were used for training and testing process, respectively. Figure 4 shows the comparison of actual classes and predicted classes of the imposed global damage for train data, test data, and all data of M-SVM. In this Figure, the hollow circles and stars indicated the actual classes and predicted classes, respectively, which if the classification be correctly done, then the hollow circles and stars will be overlapped together. The obtained results showed that the M-SVM was predicted the classes number 1 and number 3 with high precision. But for class number 2, this performance was very low because the M-SVM was not able to determine the proper margins based on feature of input data.

Also for evaluation of the obtained results from classified data, the confusion matrix is used and is defined as an error matrix or a contingency table to determine the performance

TABLE 4: The range of structural and ground motion parameters used.

Type	Parameters	Definition	Range of parameters
Structural	H/B	The ratio of total building height to width	1.44–1.78
	L (m)	Frame load share from lateral bays	3.25–3.35
	IC (mm^4)	Total moment of inertia for first story columns	200000–106666.7
	IB (mm^4)	Total moment of inertia for first story beams	90000–50625
	N	Number of stories	4–8
	B	Number of bays	3-4
	M (m)	Maximum bay length	3–4.5
Ground motion	PGA (g)	PGA	0.1–1.0
	FC (Hz)	Frequency content	0.22–4.297
	E_T (sec)	Effective time	6.99–36.32
	A/V (g/cms^{-1})	Ration of PGA to PGV	0.004–0.029

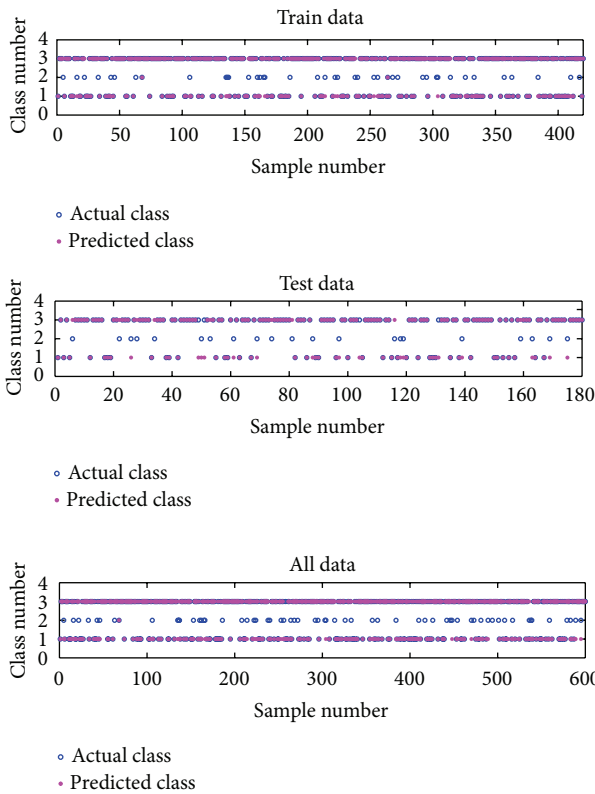
FIGURE 4: Comparison of the actual and predicted classes for train data, test data, and all data of M-SVM.

FIGURE 5: Sample of confusion matrix.

The confusion matrix for the train data, test data, and all data of M-SVM is given as follows.

Confusion matrix for the train data $= \begin{bmatrix} 112 & 0 & 0 \\ 20 & 2 & 19 \\ 0 & 0 & 267 \end{bmatrix}$, total accuracy = 90.71%, $C = 76.11$, $\sigma = 1.8885$,

Confusion matrix for the test data $= \begin{bmatrix} 41 & 0 & 0 \\ 12 & 0 & 9 \\ 4 & 0 & 114 \end{bmatrix}$, total accuracy = 86.11%, $C = 117.38$, $\sigma = 2.3453$,

Confusion matrix for the all data $= \begin{bmatrix} 153 & 0 & 0 \\ 35 & 1 & 26 \\ 4 & 0 & 381 \end{bmatrix}$, total accuracy = 90.71%, $C = 346.69$, $\sigma = 7.7198$.

Based on extracted confusion matrices, the amounts of SEN, SPC, PRE, ACC, Error, NPV, and PPV for each class and each set of data are presented in Table 6. The obtained results from these parameters showed that the best and worst classifications were done for class number 3 and class number 2 with ACC equal to 95% and 89.83%, respectively. Also, the low value of PRE for class number 2 represents inaccuracies in classification of this class.

5.2. MM-SVM Model. For generation of MM-SVM, a one-layer feed-forward MLP neural network [28, 29] was used at first level, then the output of this network was applied to M-SVM in form of input data. This combined artificial neural network was named MM-SVM and is shown in Figure 6.

of network. Each element of this matrix expresses the number of actual classes versus predicted classes. The structure of confusion matrix is shown in Figure 5.

Whereas TP is a true positive observation, TN is a true negative observation. FN is a false negative since observation is an actual negative (−) but the classifier label is positive (+) and FP is a false positive since observation is an actual positive (+), nonetheless, the classifier label is negative (−). For assessment of this matrix, some parameters can be used which are shown in Table 5.
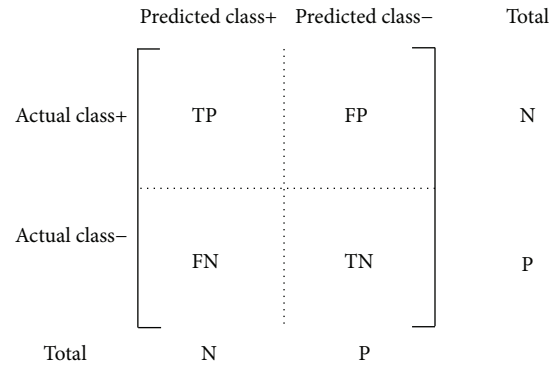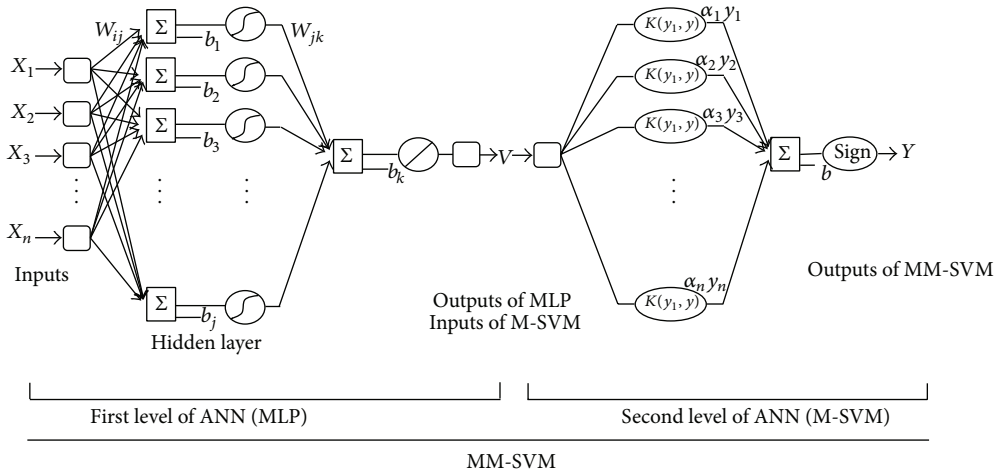
Figure 6: The architecture of combined MLP with M-SVM (MM-SVM).

Table 5: Properties of parameters used for evaluation of confusion matrix.

| Name	Definition	Function
Sensitivity or recall (SEN)	The rate of true positive	$\text{SEN} = \dfrac{\text{TP}}{\text{TP} + \text{FN}}$
Specificity (SPC)	The rate of true negative	$\text{SPC} = \dfrac{\text{TN}}{\text{FP} + \text{TN}}$
Precision (PRE)	The fraction of observations classified as positive that are actually positive	$\text{PRE} = \dfrac{\text{TP}}{\text{TP} + \text{FP}}$
Accuracy (ACC)	The proportion of the total number of predictions that are correct	$\text{ACC} = \dfrac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$
Error	The proportion of the total number of predictions that are incorrect	$\text{Error} = \dfrac{\text{FP} + \text{FN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$
NPV	The negative predictive values	$\text{NPV} = \dfrac{\text{TN}}{\text{TN} + \text{FN}}$
PPV	The positive predictive values	$\text{PPV} = \dfrac{\text{TP}}{\text{TP} + \text{FP}}$
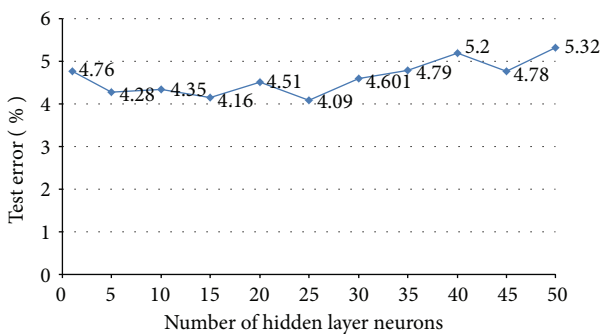
Figure 7: Number of hidden layer neurons versus test error.

At the first level of MM-SVM, 600 input data applied to MLP neural network. In this network, each layer's nodes are connected to the next layer nodes with a specific weight similar to the synaptic weight in human neural networks. For training process, the Levenberg-Marquardt back propagation algorithm was employed to update the weights and bias terms of the MLP network. Therefore, using this network leads to changing the primal data space from the eleven dimensions to one dimension. The MLP network consists of eleven neurons in the input layer, optimum neurons in the hidden layer, and one neuron in the output layer. The number of neurons and the type of activation function in the hidden layer are very important parameters for the network training process. In this research, Gaussian and linear functions were used for the hidden layer and the output layer of the MLP network, respectively. In addition, the number of hidden layer neurons was determined based on the minimum test error, as in the following equation:

$$\text{Error}\,(\%) = \left(\frac{\sum_{i=1}^{k} |O_i - T_i|}{m \cdot n} \right) \cdot 100, \tag{33}$$

where O_i is the output of the neural network, T_i is the desired output, k is the number of testing or training data, m is the number of testing or training segments, and n is the number of neural network outputs for testing and training procedures [30]. The test error values for different number of neurons in hidden layer which were calculated based on the Gaussian function are shown in Figure 7.

TABLE 6: The SEN, SPC, PRE, ACC, Error, NPV, and PPV values for each class.

	Train data			Test data			All data		
	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
SEN (%)	84.85	100.0	93.36	71.93	00.00	92.68	79.69	100.0	93.61
SPC (%)	100.0	90.67	100.0	100.0	88.33	92.98	100.0	89.82	97.93
PRE (%)	100.0	04.88	100.0	100.0	00.00	96.61	100.0	01.61	98.96
ACC (%)	95.24	90.71	95.48	91.11	88.33	92.78	93.50	89.83	95.00
Error (%)	04.76	9.286	04.52	08.89	11.67	07.22	06.50	10.17	05.00
NPV (%)	93.51	100.0	87.58	88.49	100.0	85.48	91.28	100.0	87.91
PPV (%)	100.0	04.88	100.0	100.0	00.00	96.61	100.0	01.61	98.96

TABLE 7: The RMSE, MSE, R, μ, and σ values for each set of data.

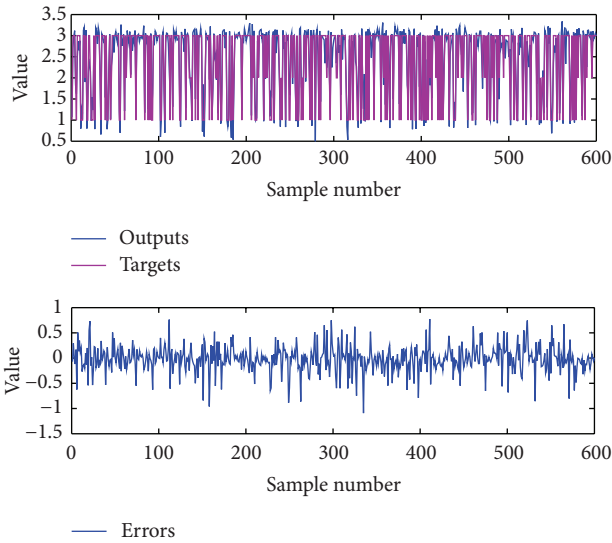|	Train data	Validation data	Test data	All data
RMSE	0.24341	0.30799	0.31044	0.26488
MSE	0.05925	0.09486	0.09637	0.07016
R	0.95944	0.93508	0.93819	0.95224
μ	−0.012685	−0.0013085	−0.056020	−0.017479
σ	0.24337	0.30971	0.30706	0.26452

FIGURE 8: Comparing the real and predicted values.

Therefore, the minimum test error was obtained equal to 4.09 for the 25 hidden layer neurons. In the MLP network, all the data was divided into three sets that include train data (70% of all data), validation data (15% of all data), and test data (15% of all data). For finding the best fit of the data set, the statistical models were used. For this reason, root mean square error (RMSE), mean square error (MSE), correlation coefficient (R), mean of error (μ), and standard deviation of error (σ) were considered [31]. In the training process, the network stopped after 16 iterations with MSE and gradient equal to 0.0413 and 0.0254, respectively. In addition, the best validation performance was 0.09486 at epoch 10. The values of RMSE, MSE, R, μ, and σ are presented in Table 7.

As seen in Table 7, the RMSE and MSE variables in testing cases are greater than in training and validating cases. Also the amount of Rs ranged between 0.93508 and 0.95944. Figure 8 compares the real output values and the predicted values of all data. The most variation of errors for all data sets happened between −0.5 and +0.5 which successfully represents the network damage prediction values with high accuracy. The regression and fit function for the train, validation, test, and all data are shown in Figure 9. The high value of R (around 0.95) indicates a good linear relationship between predicted values and actual values for the total response in the MLP model. Moreover, the histogram of error for all data is presented in Figure 10. The concentration of the bins error around the zero line with mean −0.017479 and standard deviation 0.26452 for all sets of data represents a good performance of this network.

At the second level of MM-SVM, the obtained outputs of the first level (MLP network) were applied to the M-SVM as input data. This set includes 600 input data and each data consists of only one element. Figure 11 shows the comparison of the actual classes and the predicted classes of the imposed global damage for the train data, test data, and all data of MM-SVM. For all classes, the results showed that the MM-SVM predicted the classification of global damage with high accuracy compared to M-SVM. Indeed, reduction in feature space of input data and creating high correlation between the input and output data by the MLP neural network lead to determine more precisely of margins for each class by the M-SVM.

For evaluation of the system's performance of MM-SVM network, the confusion matrix for the train data, test data, and all data is given as follows.

$$\text{Confusion matrix for the train data} = \begin{bmatrix} 105 & 7 & 0 \\ 0 & 36 & 5 \\ 0 & 4 & 263 \end{bmatrix},$$
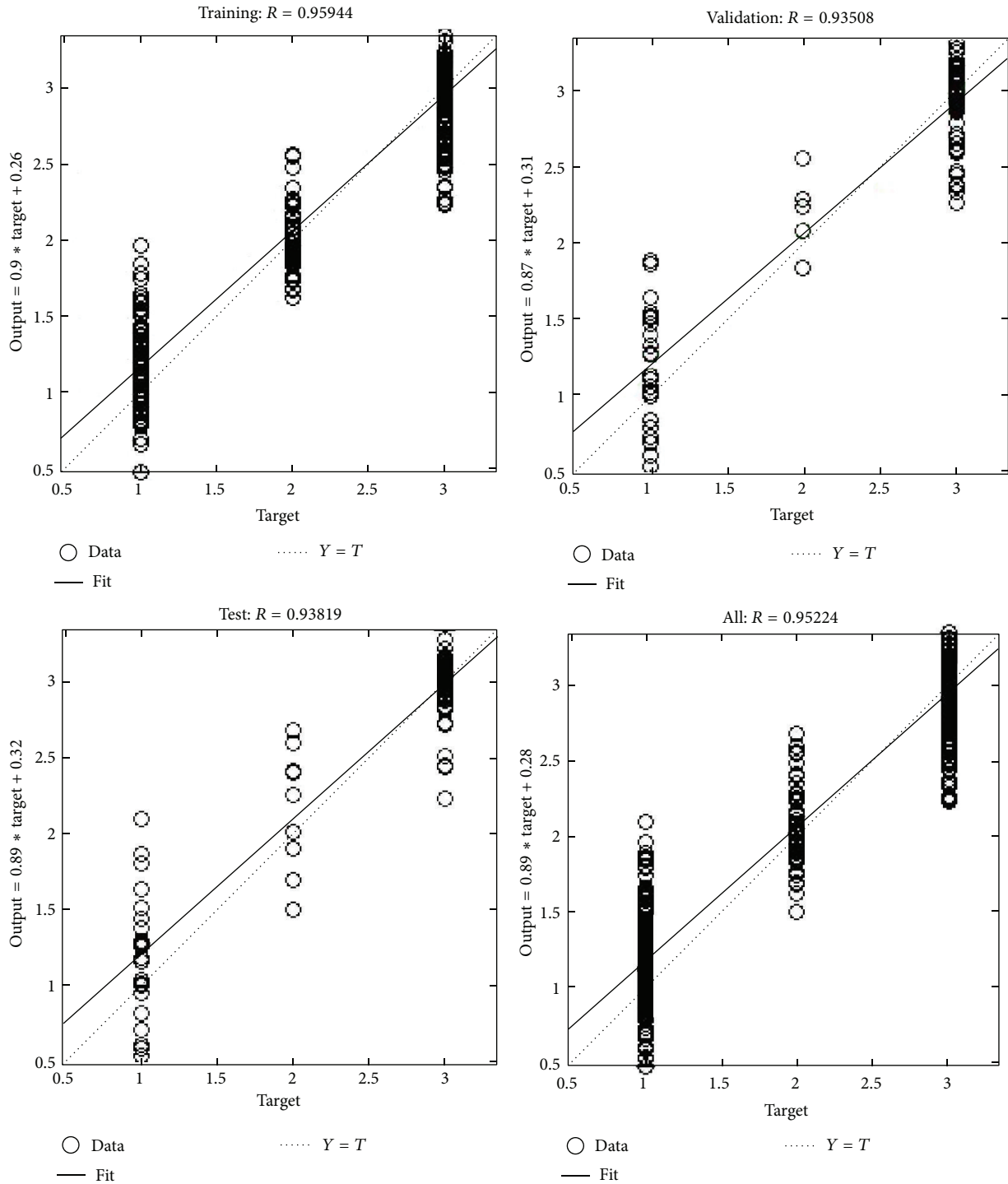$$\text{total Accuracy} = 96.19\%, C = 76.11, \sigma = 0.1928,$$

FIGURE 9: The regression and fit function for each set of data.

Confusion matrix for the test data = $\begin{bmatrix} 39 & 2 & 0 \\ 5 & 11 & 5 \\ 0 & 1 & 117 \end{bmatrix}$, total Accuracy = 92.78%, $C = 32$, $\sigma = 0.8781$,

Confusion matrix for the all data = $\begin{bmatrix} 144 & 9 & 0 \\ 7 & 44 & 11 \\ 0 & 5 & 380 \end{bmatrix}$, total Accuracy = 94.67%, $C = 181.0193$, $\sigma = 0.7071$.

Based on extracted confusion matrices from MM-SVM, the amounts of SEN, SPC, PRE, ACC, Error, NPV, and PPV

for each class and each set of data are presented in Table 8. The maximum and minimum of SPC were obtained equal to 98.66% and 93.98% for class number 2, respectively. Also for all data, the ACC-values for class number 1, class number 2, and class number 3 were extracted equal to 97.33, 94.67, and 97.33 respectively.

The above obtained results for all three classes showed that although the amount of data for each class is different,

Table 8: The SEN, SPC, PRE, ACC, Error, NPV, and PPV values for each class.

	Train data			Test data			All data		
	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
SEN (%)	100.0	76.60	98.13	88.64	78.57	95.90	95.36	75.86	97.19
SPC (%)	97.78	98.66	97.37	98.53	93.98	98.28	98.00	96.68	97.61
PRE (%)	93.75	87.80	98.50	95.12	52.38	99.15	94.18	70.97	98.71
ACC (%)	98.33	96.19	97.86	96.11	92.78	96.67	97.33	94.67	97.33
Error (%)	01.67	3.810	02.14	03.89	07.22	03.33	02.67	05.33	02.67
NPV (%)	100.0	97.10	96.73	96.40	98.11	91.94	98.43	97.40	94.89
PPV (%)	93.75	87.80	98.50	95.12	52.38	99.15	94.12	70.97	98.70
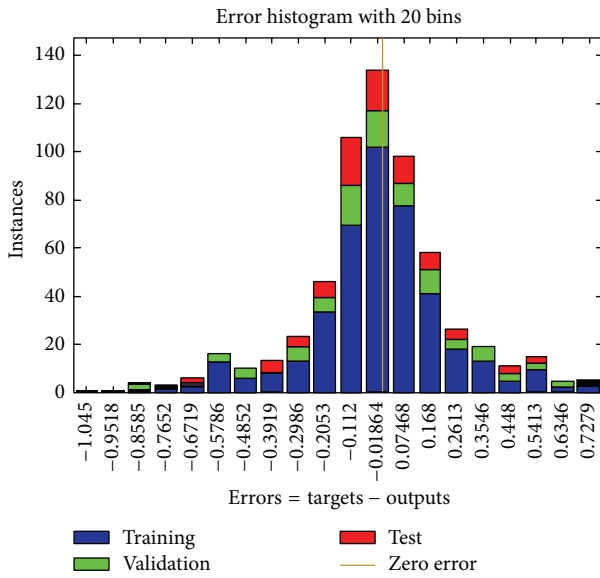
Figure 10: The error histogram for all data.

Figure 11: Comparison of the actual and predicted classes for train data, test data, and all data of MM-SVM.

the amount of these parameters are close together which indicated the MM-SVM model to be highly suited to classification of global damage for R/C slab column frames and provide reference for future seismic assessment of this frame's type.

6. Discussion

Since that the percentage of input data for class number, class number 2, and class number 3 was equal to 25.5%, 10.3%, and 64.2% of all data, respectively, the M-SVM model showed very weak performance for classification of class with minimum members. In fact, this model predicted only 2, 0, and 1 data from the set of 41, 21, and 62 considered data for the train, test and all data cases of class number 2, respectively, whereas the MM-SVM model was able to predict 36, 11, and 44 data for this class. Indeed, using the MLP model in first level of MM-SVM leads to reducing the dispersion and complication in feature space of input data and for this cause, in the second phase, the M-SVM was able to determine the margins for each class with high precision. Table 9 compares the ACC-value for M-SVM and MM-SVM models. These results showed that the MM-SVM
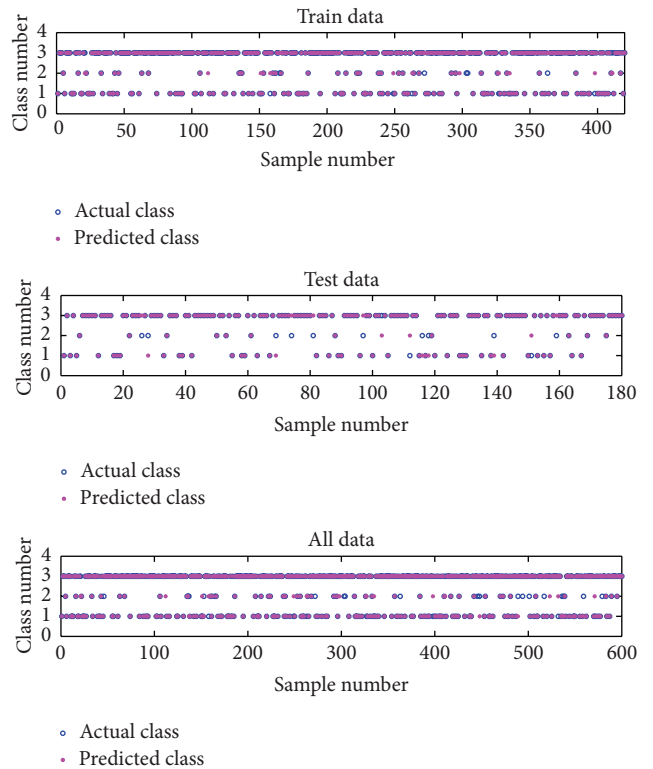
Table 9: Comparing the ACC-value for M-SVM and MM-SVM.

Data set	M-SVM (%)	MM-SVM (%)	Improvement (%)
Train data	90.71	96.19	5.48
Test data	86.11	92.78	6.67
All data	90.71	94.67	3.96

classifier improves significantly the performance in terms of recognition rate and error rate compared with M-SVM model for one classification task of global damages.

7. Conclusion

In this study, the classification of the imposed seismic damage under earthquake loads for R/C slab-column frames in the

Famagusta city was investigated. Based on the Park and Ang global damage classification, the three classes of damage are applied including repairable (economic), beyond repair (not economic), and loss of building (collapse). The NTHA was used to generate the 600 databases. Two ANNs, M-SVM, and MM-SVM were applied to establish the relationship between the structural and ground motion parameters (input data) and damage classification (output data). The following conclusions were obtained from this investigation.

In the M-SVM, to find the best kernel trick, four different kernel functions were applied including linear function, polynomial function (5 degree), Gaussian function, and sigmoid function and they evaluated using maximum accuracy of test data. The results showed that the Gaussian function had the maximum accuracy and it was employed as an efficient kernel trick function.

The number of hidden layer neurons for the first level of MM-SVM (MLP neural network) was selected based on the minimum test error, which was obtained equal to 25 neurons. Additionally, the value of R for train data, validation data, test data, and all data was calculated around 0.96, 0.94, 0.94, and 0.95, respectively.

Comparing the classification results of the M-SVM and MM-SVM showed that the total accuracy of MM-SVM is more than M-SVM. Also for class number 2 (class with the lowest member), the obtained values of PRE or PPV indicated that the MM-SVM predicted the label of this class with high efficiency towards the M-SVM. Thus, the MM-SVM is proven as an efficient and time saving way for classification of the imposed global damage under earthquake loads and it can be used for similar R/C slab-column frames solely by selecting the structural geometric and ground motion parameters. In addition, this method of damage classification can be used by the insurance companies because it is easy and fast.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] R. A. Fisher, "The use of multiple measures in taxonomic problems," *Annals of Eugenics*, vol. 7, pp. 179–188, 1936.

[2] C. Cortes and V. N. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[3] X. Jun, J. Annan, W. Zhiwu, and Q. Jingping, "A nonlinear optimization technique of tunnel construction based on DE and LSSVM," *Mathematical Problems in Engineering*, vol. 2013, Article ID 980154, 11 pages, 2013.

[4] Z. Mingheng, Z. Yaobao, H. Ganglong, and C. Gang, "Accurate multisteps traffic flow prediction based on SVM," *Mathematical Problems in Engineering*, vol. 2013, Article ID 418303, 8 pages, 2013.

[5] M. Vafaei, A. B. Adnan, and A. B. A. Rahman, "Real-time seismic damage detection of concrete shearwalls using artificial neural networks," *Journal of Earthquake Engineering*, vol. 17, no. 1, pp. 137–154, 2013.

[6] M. Ceylan, M. H. Arslan, R. Ceylan, M. Y. Kaltakci, and Y. Ozbay, "A new application area of ANN and ANFIS: determination of earthquake load reduction factor of prefabricated industrial buildings," *Civil Engineering and Environmental Systems*, vol. 27, no. 1, pp. 53–69, 2010.

[7] X. Xie, P. Li, H. Qin, L. Liu, and D. C. Nobes, "GPR identification of voids inside concrete based on the support vector machine algorithm," *Journal of Geophysics and Engineering*, vol. 10, no. 3, Article ID 034002, 2013.

[8] M. A. Köroğlu, M. Ceylan, M. H. Arslan, and A. Ilki, "Estimation of flexural capacity of quadrilateral FRP-confined RC columns using combined artificial neural network," *Engineering Structures*, vol. 42, pp. 23–32, 2012.

[9] O. Castillo, P. Melin, E. Ramírez, and J. Soria, "Hybrid intelligent system for cardiac arrhythmia classification with Fuzzy K-Nearest Neighbors and neural networks combined with a fuzzy system," *Expert Systems with Applications*, vol. 39, no. 3, pp. 2947–2955, 2012.

[10] W.-Y. Chang, "An RBF neural network combined with OLS algorithm and genetic algorithm for short-term wind power forecasting," *Journal of Applied Mathematics*, vol. 2013, Article ID 971389, 9 pages, 2013.

[11] Turkish Earthquake Code (TEC-1975), *Regulations on Structures Constructed in Disaster Regions*, Ministry of Public Works and Settlement, Ankara, Turkey, 1975.

[12] M. H. Arslan, "An evaluation of effective design parameters on earthquake performance of RC buildings using neural networks," *Engineering Structures*, vol. 32, no. 7, pp. 1888–1898, 2010.

[13] Turkish Earthquake Code(TEC-2007), *Regulations on Structures Constructedin Disaster Regions*, Ministry of Public Works and Settlement, Ankara, Turkey, 2007.

[14] Z. Cagnan and G. B. Tanircan, "Seismic hazard assessment for Cyprus," *Journal of Seismology*, vol. 14, no. 2, pp. 225–246, 2010.

[15] PEER, "Pacific earthquake engineering research center," 2013, http://peer.berkeley.edu.

[16] G. H. Powell and R. Allahabadi, "Seismic damage prediction by deterministic methods: concepts and procedures," *Earthquake Engineering & Structural Dynamics*, vol. 16, no. 5, pp. 719–734, 1988.

[17] M. Yoshimura, "Control of seismic drift demand for reinforced concrete buildings with weak first stories," *Earthquake Engineering and Engineering Seismology*, vol. 4, no. 1, pp. 27–35, 2003.

[18] C. Meyer and M. S. L. Roufaiel, "Analysis of damaged concrete frames for cyclic loads," *Earthquake Engineering & Structural Dynamics*, vol. 11, no. 2, pp. 207–228, 1983.

[19] S. Toussi and J. T. P. Yao, "Hysteresis identification of existing structures," *Journal of Engineering Mechanics*, vol. 109, no. 5, pp. 1189–1202, 1983.

[20] J. E. Stephens and J. T. P. Yao, "Damage assessment using response measurements," *Journal of Structural Engineering*, vol. 113, no. 4, pp. 787–801, 1987.

[21] Y. J. Park, A. H. S. Ang, and Y. K. Wen, "Seismic damage analysis of reinforced concrete buildings," *Journal of Structural Engineering*, vol. 111, no. 4, pp. 740–757, 1985.

[22] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[23] R. Fletcher, *Practical Methods of Optimization*, John Wiley & Sons, Chichester, UK, 1987.

[24] N. Heckman, *The Theory and Application of Penalized Least Squares Methods or Reproducing Kernel Hilbert Spaces Made Easy*, University of Birtish Columbia, Vancouver, Canada, 1997.

[25] K. Crammer and Y. Singer, "On the algorithmic implementation of multi-class SVMs," *Journal of Machine Learning*, vol. 2, no. 7, pp. 265–292, 2001.

[26] Y. J. Park, A. M. Reinhorn, and S. K. Kunnath, *IDARC-2D V7.0 Computer Program: Inelastic Damage Analysis of RC Building Structures*, State University of New York, 2010.

[27] C. C. Chang and C. J. Lin, "A library for support Vector machines," 2013, http://www.csie.ntu.edu.tw/~cjlin/libsvm/.

[28] S. S. Chiddarwar and N. R. Babu, "Comparison of RBF and MLP neural networks to solve inverse kinematic problem for 6R serial robot by a fusion approach," *Engineering Applications of Artificial Intelligence*, vol. 23, no. 7, pp. 1083–1092, 2010.

[29] I. Yilmaz and O. Kaynar, "Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5958–5966, 2011.

[30] Y. Özbay, R. Ceylan, and B. Karlik, "A fuzzy clustering neural network architecture for classification of ECG arrhythmias," *Computers in Biology and Medicine*, vol. 36, no. 4, pp. 376–388, 2006.

[31] N. R. Draper and H. Smith, *Applied Regression Analysis*, Wiley-Interscience, Hoboken, NJ, USA, 3rd edition, 1998.