Genetic Design of Biologically Inspired Receptive Fields for Neural Pattern Recognition

Claudio A. Perez, *Member, IEEE*, Cristian A. Salinas, Pablo A. Estévez, *Member, IEEE*, and Patricia M. Valenzuela

***Abstract*—This paper proposes a new method to design, through simulated evolution, biologically inspired receptive fields in feed forward neural networks (NNs). The method is intended to enhance pattern recognition performance by creating new neural architectures specifically tuned for a particular pattern recognition problem. It is proposed a combined neural architecture composed of two networks in cascade: a feature extraction network (FEN) followed by a neural classifier. The FEN is composed of several layers with receptive fields constructed by an additive superposition of excitatory and inhibitory fields. A genetic algorithm (GA) is used to select the receptive fields parameters to improve the classification performance. The parameters are the receptive field size, orientation, and bias as well as the number of different receptive fields in each layer. Based on a random initial population where each individual represents a different neural architecture, the GA creates new enhanced individuals. The method is applied to the problems of handwritten digit classification and face recognition. In both problems, results show strong dependency between the NN classification performance and the receptive fields architecture. The GA selected parameters of the receptive fields that produced improvements in the classification performance on the test set up to 90.8% for the problem of handwritten digit classification and up to 84.2% for the face recognition problem. On the same test sets, results were compared advantageously to standard feed forward multilayer perceptron (MLP) NNs where receptive fields are not explicitly defined. The MLP reached a maximum classification performance of 84.9% and 77.5% in both problems, respectively.**

***Index Terms*—Face recognition, genetic algorithms (GAs), handwritten digit classification, neural pattern recognition, receptive fields.**

I. Introduction

A. Biological Receptive Fields

THE MAMMALIAN visual cortex has evolved over millions of years to perform pattern recognition tasks with remarkable precision [40]. In the visual system, neuroscientists have identified a hierarchical structure proceeding along multiple parallel pathways through a number of anatomically defined layers. There is physiological evidence that neurons respond to more complex patterns as the flow of information proceeds from the retina to the inferotemporal cortex in the visual pathways [2], [7], [41].

A receptive field is the region of the sensor where an adequate stimulus elicits a response [8]. The receptive fields of retinal ganglion cells in mammals are organized in center/surround configurations [28]. The receptive fields from retinal and lateral geniculate nucleus (LGN) neurons have circular symmetry and they respond almost equally to all stimulus orientations [50]. Hubel and Wiesel built a comprehensive picture of the basic functional architecture of the visual cortex [27]. They defined "simple cells" as cells where it is possible to map the excitatory and inhibitory regions of the receptive field by monitoring the cell's response to a spot of light. The receptive fields of simple cells were implemented by overlapping the receptive fields of center/surround cells from LGN [25], [26]. Simple cells at the visual cortex have oriented receptive fields, and hence they respond to stimuli in some orientations better than others being excellent at detecting the presence of simple visual features such as lines and edges of a particular orientation [50]. Orientation selective receptive neurons are found throughout layers 2 and 3 of the primary visual cortex and are relatively rare in the primary inputs within layer $4C$ [41], [50].

The receptive fields are also local in the two-dimensional (2-D) spatial frequency domain. The spatial contrast-sensitivity functions of cortical neurons were measured and resulted to be narrower than those of retinal ganglion cells [9]. This localization in 2-D space means that cells respond to a small band of radial spatial frequencies and to a small band of orientations [51]. The spatial frequency response characteristic of a cortical simple cell can conveniently be described in terms of Gabor functions [29].

"Complex cells" receive excitatory inputs from neighboring simple cells receptive fields of similar preferred orientation. Whenever a line or edge stimulus of the correct orientation falls within the receptive field of one of the simple cells, the complex cell is activated [26]. The receptive field properties become progressively more sophisticated as the flow of information proceeds from the retina to the inferotemporal cortex [50]. Understanding of cortical visual areas is still in early phase of scientific study and as the cortex is explored it is expected to find neurons with new receptive field properties. It will be necessary to characterize these receptive fields adequately to understand their computational role in vision.

The spatial receptive fields of simple cells in mammalian striate cortex have been described well physiologically and can be characterized as being localized, oriented in space, and bandpass in the frequency domain [39]. It has been shown that one of the early vision tasks is to reduce the redundancy of the input signals [3], [5]. Also, theories about efficient coding provide in-

Manuscript received February 7, 2002; revised April 15, 2002. This work was supported in part by CONICYT-Chile through Project FONDECYT 1990621 and by the Department of Electrical Engineering, Universidad de Chile. This paper was recommended by Associate Editor H. Takagi.

C. A. Perez and P. A. Estévez are with the Department of Electrical Engineering, Universidad de Chile, Santiago, Chile (e-mail: clperez@cec.uchile.cl; pestevez@cec.uchile.cl).

C. A. Salinas was with the Department of Electrical Engineering, Universidad de Chile, Santiago, Chile. He is now with Case Western Reserve University, Cleveland, OH 44106 USA (e-mail: cas35@po.cwru.edu).

P. M. Valenzuela is with the Department of Pediatrics, Pontificia Universidad Católica de Chile, Santiago, Chile (e-mail: pvalenzu@med.puc.cl).

Digital Object Identifier 10.1109/TSMCB.2003.810441

sights about cortical image representation [39], [40]. It was also shown that by using a learning algorithm to find sparse linear codes for natural scenes, a complete family of localized, oriented, and bandpass receptive fields, similar to those found in the striate cortex, were developed [17], [39], [40].

B. Receptive Fields in Artificial Neural Networks (ANNs)

Fully connected feed forward neural networks (NNs) allow connections between each unit of one layer and all units in the previous layer without an explicit specification of the receptive field architectures as in the visual system. Several papers have described the use of biological models of vision to solve pattern recognition problems [2], [11], [20], [32], [46]. Successful applications have been performed in bandwidth compression, image quality assessment, and image enhancement [18]. Several algorithms for image processing such as edge detection based on gradient operators and multiresolution architectures are based on the visual system [32]. For example, a cortical column architecture was used in a multilayer network to perform pattern recognition in speech and visual recognition [2]. The application of the retinal "Mexican hat" filter yielded a high-pass image emphasizing contours and sharp changes in luminosity [11]. Furthermore, it was shown that the retinal filter is optimized to decorrelate the incoming luminosity signal based on its correlation function [3].

According to the biological evidence described in the previous section, multilayered NNs may be built to correspond to the successive cortical maps with learning rules which enable them to learn many types of input–output transformations [2]. The basic structure of this model has been captured in a neural model where each layer is composed of several ordered ensembles of units called "planes" and a connectivity among layers enabling the appearance of simple and complex receptive fields. Examples of this type of neural architecture have been used in [1], [12], [13], [38], and [51].

In particular, biologically inspired receptive fields have been used in the neocognitron model and its variations [13], [38], and in modeling the visual cortex retinotopical and orientation maps [12]. One of the advantages of NNs with receptive fields is that the number of weights is efficiently reduced in relation to fully connected architectures [38].

The neocognitron was trained to recognize 24 simple patterns regardless where they appeared on a 16×16 element visual field. It was assumed that complex C cells from one cortical patch led to a pair of simple (S) and complex (C) cell layers [10]. Other areas were also organized in pairs of S and C units. In the neocognitron, the weights corresponding to units in the S layers were adjusted by training but the basic architecture of the receptive fields, the number of planes per layers and the number of total layers remained fixed [13]–[16], [31]. However, an automatic method to design or adjust the receptive field architecture has not been found in the literature.

C. Evolutionary Neural Networks

Evolutionary NNs refer to a special class of ANNs in which evolution is an essential form of adaptation in addition to learning [54]. Evolution has been introduced in NNs in three main levels: 1) connection weights; 2) architectures; and 3) learning rules [54]. As we employ genetic algorithms (GAs) to evolve the architecture of the receptive fields, we focus only on architecture evolution.

Evolution of NN topological structures can be performed by a direct encoding scheme where each connection in the architecture is directly specified by its binary representation. Another alternative is to use an indirect encoding scheme, e.g., a set of rules (grammar) that can be applied to produce an NN architecture [34]. The indirect encoding scheme can produce more compact genotypical representation of the neural architecture [54].

GAs have been used to find near optimal training parameters on neocognitron models [33]. Also, the selectivity parameters of the neocognitron were optimized using GAs [55]. In other work, GAs were used to optimize the number of layers, number of hidden units, learning rate, momentum, number of epochs, and initial set of weights values in a perceptron-type architecture [45], [52]. In our literature review, we have not found any application of evolutionary algorithms to the specific design of receptive fields to solve pattern recognition problems as it is proposed in this paper.

D. Purpose of the Present Study

In this paper a new method is proposed for automatic design of biologically inspired receptive fields in ANNs. The method, based on a GA, builds up an adequate receptive field architecture searching for the appropriate dimensions, orientation, and bias of center-surround receptive fields to maximize the classification performance of the network. The proposed NN architecture is composed of two combined networks connected in cascade. The first one, a feature extraction network (FEN), is a neural architecture with biologically inspired connectivity and receptive fields such as the neocognitron model [13]. The second network is a fully connected MLP classifier [23], [30], [47]. A GA is used to select parameters of the FEN receptive fields to improve the classification performance of the combined networks. The method designs a near optimum receptive field architecture for a specific problem and therefore could be applied to several classification tasks. The method is expected to enhance pattern recognition performance by an appropriate design of the combined network architecture. In this paper, the method is applied to the problems of handwritten digit classification and face recognition. The results are compared to those of fully connected MLP where no explicit receptive fields are defined.

The method is intended to adapt the receptive field basic architecture of an ANN to create new models specifically tuned for a particular pattern recognition problem. Modeling the biological computations performed in the visual system is beyond the scope of this paper. Preliminary results of the proposed method applied to find the receptive fields x and y dimensions in a fixed neural architecture of two planes per layer were presented in [42] and [44].

II. METHODS

A. Combined Network Architecture

Functionally, the NN architecture can be decomposed into two networks in cascade as illustrated in Fig. 1. The first one,

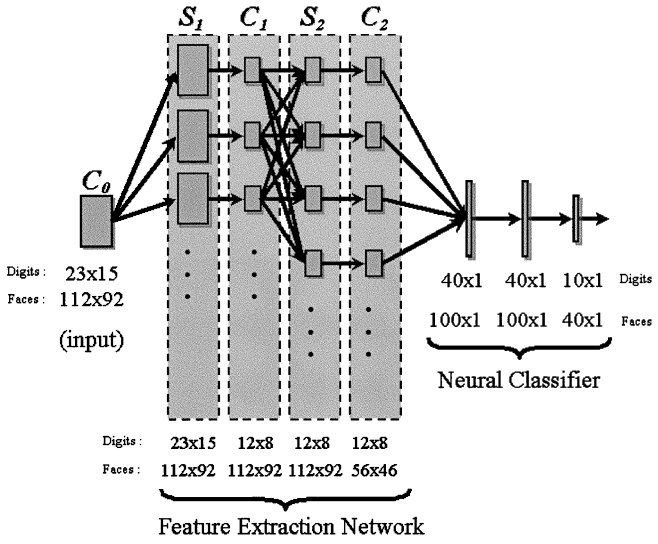

Fig. 1. Combined neural architecture composed of two networks in cascade: a feature extraction network (FEN) followed by an MLP classifier. The FEN is convolutional-type network composed of four layers with variable number of planes per layer and a threshold function. A scale reduction is performed between layers S_1 and C_1 for the handwritten classification problem and between S_2 and C_2 for the face recognition problem. Dimensions of each plane are specified for each problem (digits or faces). The neural classifier is an MLP composed of two hidden layers of 40 or 100 hidden units each one for the handwritten digit or the face recognition problem, respectively.

the FEN, is a hierarchical convolutional-type network [13], [38]. The feature extraction process is performed through successive filtering of the input image by four hidden layers connected in cascade. The second network is a neural classifier with an MLP architecture.

1) The Feature Extraction Network (FEN): The FEN is a neural architecture with biologically inspired connectivity and receptive fields such as the neocognitron model. The FEN is organized in four alternated S- and C-type layers following the notation used by Fukushima [13]. As shown in Fig. 1, the model consists of layers S_1, C_1, S_2, and C_2, where each layer is composed of several planes. As in the neocognitron, a C-type layer contains units with receptive fields over units belonging to a single plane of the previous layer, i.e., a unit of a C-type layer is only connected to units within the same plane of the previous layer. An S-type layer contains units connected to other units in different planes of the previous layer. The number of planes in a C-type layer is the same as the number of planes in the previous S-type layer. Each unit within a plane shares the same receptive field parameters (size, rotation and bias) with the other units in the same plane. Therefore, the number of different receptive fields depends on the total number of planes present in all layers. Unlike the neocognitron model, in our model the number of planes in the S-type layers is variable and it is selected by simulated evolution. As in the neocognitron model, a scale reduction is located between layer S_1 and C_1 and is performed by subsampling all planes in layer C_1. The scale reduction decreases the number of weights required in the network.

2) Receptive Field Spatial Filtering: The input image C_0 in Fig. 1 is propagated through the layers of the FEN and in each plane discrete convolution takes place between the image and the receptive field of that plane. Equation (1) describes the output of units in a C-type plane and (2) the output of the units of an S-type plane as follows:

$$C_l^k(i,j) = \sum_{m=1}^{Sx_l^k} \sum_{n=1}^{Sy_l^k} S_l^k(m,n)\, hc_l^k(i-m, j-n) \tag{1}$$

$$S_l^k(i,j) = \sum_{r=1}^{R_l} \sum_{m=1}^{Cx_{l-1}^r} \sum_{n=1}^{Cy_{l-1}^r} C_{l-1}^r(m,n)\, hs_l^k(i-m, j-n) \tag{2}$$

where $C_l^k(i,j)$ is the output of the unit in position (i,j) in the kth plane of the lth layer C, $S_l^k(i,j)$ is the output of the unit in position (i,j) in the kth plane of the lth layer S, hc_l^k, and hs_l^k are the receptive fields associated to each plane, respectively. R_l is the number of planes in layer C prior to layer S_l. Sx_l^k and Sy_l^k are the x and y dimensions, respectively, of the kth plane in the lth layer S. Cx_{l-1}^r and Cy_{l-1}^r are the dimensions x and y, respectively, of the rth plane in the $(l-1)$th layer C. The index l represents layers 1 and 2 of S-type or C-type layers. The scale reduction performed in the plane C_1 for the case of the handwritten digit classification problem and in C_2 in the face recognition problem is represented in

$$C_1^k(i,j) = \sum_{m=1}^{Sx_1^k} \sum_{n=1}^{Sy_1^k} S_1^k(m,n) \cdot hc_1^k(2i-m, 2j-n). \tag{3}$$

The scale reduction has an effect on the computational time required to convolve the receptive fields and the images. From this point of view, it is preferred to choose the scale reduction in C_1 as it was done in the handwritten digit problem. Nevertheless, human face classification depends on spatial information in the range 8–13 cycles/face [37] and therefore higher spatial resolution is required. By performing the scale reduction in C_2, the feature extraction is performed at higher spatial resolution in the face recognition problem. Besides, this feature extraction is feasible within reasonable computational time since the face database has only 160 training patterns. The output of every unit in the FEN goes through a threshold function [14] given by

$$\Phi(x) = \begin{cases} x, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0. \end{cases} \tag{4}$$

3) Receptive Field Construction: Even though the FEN architecture is based on the neocognitron model, the receptive fields are built in a complete different manner determining genetically their geometry, orientation, and bias. As illustrated in Fig. 2, each receptive field consists of an additive superposition of two separate fields: an excitatory and an inhibitory field, both rotated in a specific angle. All weights are equal within each excitatory or inhibitory field. The sum of all weights within each excitatory field is 1 and this sum is -1 for each inhibitory field. The influence of the inhibitory field over the complete receptive field may be attenuated by a parameter named bias B which takes values in the interval [0,1]. The bias adds low-pass filtering characteristics to the receptive field [53]. Thus, each receptive field is determined by six parameters: the x and y dimensions in pixels of the excitatory field, x_{exc} and y_{exc}, and of the inhibitory field, x_{inh} and y_{inh}; the orientation angle α; and the bias, B.

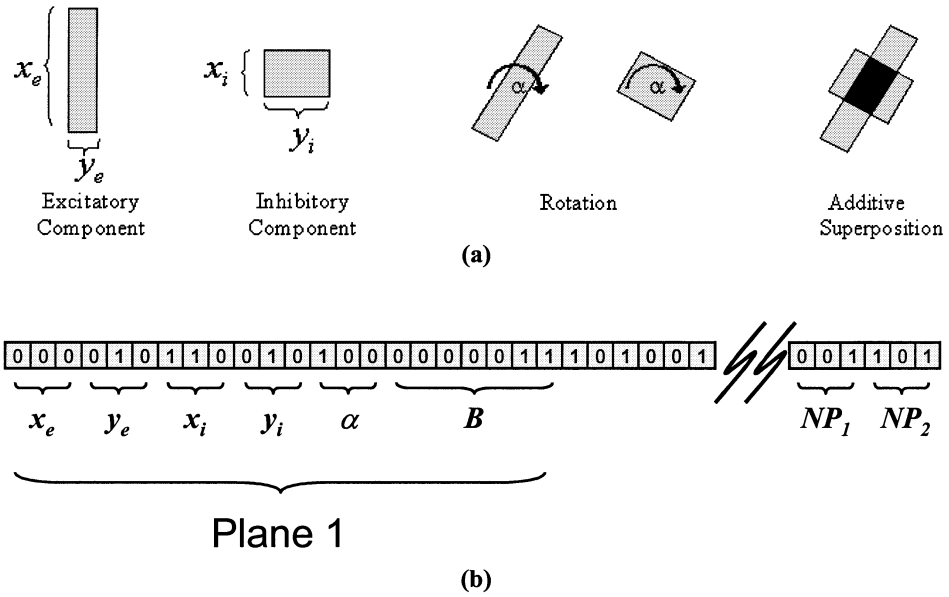

Fig. 2. Receptive field creation. (a) Determination of x and y dimensions of the excitatory and inhibitory components of the receptive field. Rotation in angle α. Additive superposition of the receptive field components including a bias B. (b) Coding of the neural architecture into a binary string. It is shown a coding example for the receptive field of one plane and the coding of the number of planes in layers 1 and 2, NP_1 and NP_2, respectively.

Fig. 2(a) shows the three basic steps in the receptive field construction: selection of x_{exc}, y_{exc}, x_{inh}, and y_{inh} dimensions, rotation of the receptive field in angle α, and addition of the excitatory part, inhibitory part and bias of the receptive field. The analytical expression for the receptive field is given by

$$h(i,j) = h^{exc}_{x_{exc},y_{exc},\alpha}(i,j) + (1-B)h^{inh}_{x_{inh},y_{inh},\alpha}(i,j) \quad (5)$$

where $h^{exc}_{x_{exc},y_{exc},\alpha}$ is the excitatory component of the receptive field of dimensions x_{exc} by y_{exc} and orientation angle α, $h^{inh}_{x_{inh},y_{inh},\alpha}$ is the inhibitory component of dimensions x_{inh} by y_{inh} and orientation angle α. Equations (6) and (7) show the restrictions on the intensity for both components of the receptive field h^{exc} and h^{inh}, respectively

$$\sum_{i=1}^{x_{exc}} \sum_{j=1}^{y_{exc}} h^{exc}_{x_{exc},y_{exc},\alpha}(i,j) = 1 \qquad \forall x_{exc},\, y_{exc},\, \alpha \quad (6)$$

$$\sum_{i=1}^{x_{inh}} \sum_{j=1}^{y_{inh}} h^{inh}_{x_{inh},y_{inh},\alpha}(i,j) = -1 \qquad \forall x_{inh},\, y_{inh},\, \alpha. \quad (7)$$

The result of the application of the receptive fields over the input image is a nonlinear filtering process performed by each plane in the FEN. As parameters x_{exc}, y_{exc}, x_{inh}, y_{inh}, α, and B in (5) determine the spatial configuration of the receptive field, these parameters also determine the equivalent spatial filtering performed by the FEN.

4) Neural Classifier: As shown in Fig. 1, the second network in cascade, the neural classifier, is an MLP architecture [23] composed of two hidden layers and a number of outputs equal to the number of classes of the pattern recognition problem. For the handwritten digit classification the number of outputs is 10. In the case of the face recognition problem the number of outputs is 40 since there are 40 different subjects in the database. The maximum value among the outputs of the neural classifier determines the class for the input pattern

TABLE I
MLP NEURAL CLASSIFIER TRAINING PARAMETERS

Parameter	Value
Number of hidden layers	2
Units in the first hidden layer	40
Units in the second hidden layer	40
Range of initial random weights	[-0.5, 0.5]
Learning rate	0.01
Momentum	0.2
Training epochs	500

[30]. The neural classifier was trained by backpropagation with momentum for a fixed number of epochs [21], [47]. The chosen parameters for the MLP architecture have been determined in previous research for the same pattern recognition problems [43]–[45]. The number of units in each hidden layer was determined to be $N_h = 40$ for the handwritten digit problem and $N_h = 100$ for the face recognition problem. These numbers of hidden units were selected based on results of training several MLP models for different number of hidden units and selecting the smallest number with highest classification performance. The training parameters are summarized in Table I.

5) Databases: The method was applied to the handwritten digit classification and to the face recognition problems. In the first application, a database composed of 3674 handwritten digits from university students was partitioned into three sets. The partition separates randomly digits from different persons but leaves all digits from the same person in a single partition. Therefore, from the generalization point of view this form of partition represents one of the most difficult possible cases. The training set was composed of 1837 handwritten digits, the validation and test sets were composed of 918 and 919 digits, respectively. The handwritten database contains digits that are not normalized in size as shown in Fig. 3(a). Each handwritten digit is composed of 15×23 pixels, with two levels (binary) per pixel. This database is available in [22].

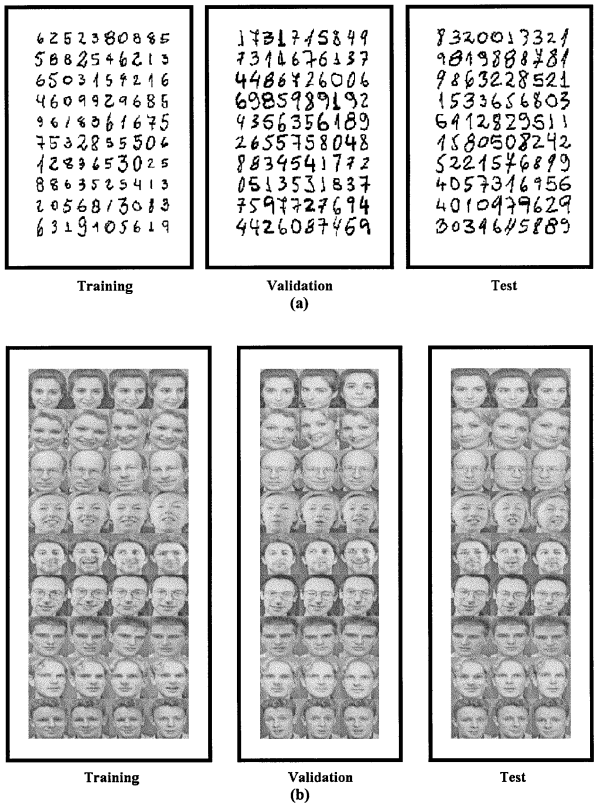

Fig. 3. Samples from the training, validation and test sets from the (a) handwritten digit and (b) face databases.

The second database is the face database from AT&T Laboratories Cambridge, where there are ten different images of each of 40 distinct subjects. The size of each image is 92×112 pixels, with 256 gray levels per pixel. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open/closed eyes, smiling/not smiling) and facial details (glasses/no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement) [49]. The face database was also partitioned randomly into three sets: 160 faces for training, 120 for validation and 120 for testing. This leaves for each individual in the database four faces for training, three for validation, and three for test. Some cases include individuals with and without glasses in different partition sets, which represent a difficult problem.

Samples for each partition on both databases are shown in Fig. 3. The training set is used to train each individual in the GA. The validation set is used to determine the fitness of each individual in the population and therefore to generate the next population according to the GA as it is explained in Section II-B. The generalization performance of each individual is measured on the test set [52].

B. Genetic Algorithm

The optimum size, orientation and bias of the receptive fields as well as the number of different receptive fields to maximize the classification performance of the network are not known. Given the problem high dimensionality and the lack of *a priori* knowledge, GAs emerge as an appropriate tool to find a suitable configuration for the receptive fields [42], [44].

TABLE II
FEN PARAMETERS ENCODED IN THE SPECIFIED NUMBER OF BITS. THE TABLE ALSO SHOWS THE MAXIMUM AND MINIMUM VALUE OF EACH PARAMETER AND THE MINIMUM STEP OF CHANGE BETWEEN THOSE VALUES

Parameter	Number of bits	Minimum Value	Maximum Value	Step
x_e	3	1	8	1
y_e	3	1	8	1
x_i	3	1	8	1
y_i	3	1	8	1
α (degrees)	3	0	157.5	22.5
Bias	7	0	1	0.0079
NP_1	3	1	8	1
NP_2	3	1	8	1

1) Coding: The proposed network architecture can be encoded into a binary string [24], [36]. Table II shows the FEN parameters encoded by the GA. The number of planes in layers S_1 and S_2 are defined by the parameters NP_1 and NP_2, respectively. According to the number of layers and number of planes per layer used in previous convolutional models [13], [38] the maximum number of planes per layer in S_1 and S_2 was set to 8, and the total number of layers in the network was 4. Therefore, there may be a total of 32 different receptive fields. Besides, each receptive field is defined by six parameters: excitatory and inhibitory x, y dimensions, rotation α, and bias B. Two global parameters were needed additionally to encode NP_1 and NP_2. An example of string coding for one plane and for the number of planes per layer is shown in Fig. 2(b). The total number of parameters in a single FEN is 194 $(6 \times 32 + 2)$. The number of bits assigned to each parameter defines the number of different values that each parameter may take, e.g., three bits take $2^3 = 8$ values. As shown in Table II each receptive field is encoded by 22 bits, therefore 704 bits are required to encode the 32 receptive fields. Additionally, six bits are required to encode the number of planes in S_1 and S_2. Therefore, the total number of bits required per architecture is 710.

2) Fitness: Each individual in a given population of the GA represents a specific network architecture that must be ranked according to its performance in the pattern recognition task. All architectures in the population are trained with the training set. During training, the classification performance in the validation set was also computed. Then, each individual was assigned a fitness that was equal to the maximum recognition rate reached by the combined network in the validation set.

The fitness measurement is a number in the interval [0,1] representing the range 0% to 100% recognition in the validation set. As the weights of the classifier are initialized randomly at the beginning of the training process, the fitness measurement varies from one starting point to another. From the computational point of view, the best choice is to select the minimum number of evaluations for each individual in the GA. As a tradeoff, three evaluations were chosen for each individual in the GA. This approach to reduce the computational cost in GAs has been employed previously [6]. In Section III, a discussion regarding the statistical significance of this choice is given. The final fitness assigned to the individual is the average of the three evaluations. If the same individual appears more than three times during the sim-

TABLE III
SUMMARY OF THE PARAMETERS FOR THE GA

No. of individuals per population (N)	20
No. of evaluation per individual	3
Crossover probability ($P_{crossover}$)	0.5
Mutation probability ($P_{mutation}$)	0.0001

ulation, it will inherit the fitness of the first three appearances. Therefore, the fitness of each individual and its three first evaluations must be stored. This helps to reduce computations at the expense of memory space.

3) Selection and Sampling: The genetic algorithm uses proportional selection [19], [36] to assign an individual of the current generation a probability P_i to be chosen as an individual of the next generation as follows:

$$P_i = \frac{f_i}{\sum_{k=1}^{N} f_k} \tag{8}$$

where f_i is the fitness of the individual i of the population and N is the number of individuals in the population. Therefore, the expected offspring of individual i for the next generation is defined by

$$N_i = N \cdot P_i. \tag{9}$$

However, N_i is a real number that must be converted to an integer number. This conversion is performed using the stochastic universal sampling method [36], [19], which assumes that the population is laid out in random order as in a pie graph. Each individual is assigned a space on the pie graph proportional to its fitness. Then, an outer roulette wheel is placed around the pie with N equally spaced pointers. A single spin of the roulette wheel picks all N members of the new population. The resulting sampling is optimal combining zero bias, minimum spread, and $O(n)$ time complexity [4], [19].

4) Crossover and Mutation: Crossover and mutation are performed on each population after the selection process. To avoid positional bias, uniform crossover was used [35], [48]. In uniform crossover, two parents are randomly selected and for each bit position, an information exchange is produced with a probability $P_{\text{crossover}}$. This means that each bit is inherited independently of the other bits present in the parent strings. The probability of bit interchange between two individuals must be selected according to the equilibrium between exploration and exploitation [36]. Each pair of parent strings generates a pair of siblings. The mutation operator [36] consists of random changes in the bits of the binary string. The probability of mutation P_{mutation} is normally small; however, it also depends on the equilibrium between exploration and exploitation. Table III shows the probabilities used in all our experiments.

5) Computational Cost: The proposed method is intensive in computations since an NN has to be trained for each different configuration of receptive field parameters. Even for a rather small handwritten digit training database of 1837 examples, the computational time for one individual of the population takes several minutes in a Pentium III-450 MHz computer under Linux operating system. Therefore, the number of individuals in the population was chosen to be relatively small (20), to avoid longer computational time due to population size [33], [52], [55]. Furthermore, it is expected to reach significantly higher classification results with a larger training set but simulation time would not be appropriate for our available computer resources (based on Pentium III-450 MHz computers). In the case of the face recognition problem, the training database considers only 160 training images, but the input weights are 10 304 for each 92×112 pixel image. The training time takes several minutes per individual, therefore, a population of 20 individuals was also chosen for this problem.

In both problems, part of the simulations results were obtained in a computer laboratory with a network of 20 PCs to compute in parallel the individuals of one population and one PC to control the GA. As the computational performance of PCs doubles roughly every two years, the proposed method would require in the future less extensive computer resources. Once the method is applied and the GA has determined the FEN architecture, the combined architecture, FEN + MLP classifier can be implemented in a simple PC to perform classification online.

C. Multilayer Perceptron Model (MLP)

A multilayer perceptron model (MLP) neural classifier without predefined receptive fields served as a reference to compare results with our method. An MLP was directly trained with the training set and results were measured on the validation and test sets for different number of hidden units. Because of the weights random initialization, ten simulations were performed for each different number of hidden units. In the case of the handwritten recognition problem, the input layer was composed of 345 units, and the output layer was of ten units, one for each digit ($0, \ldots, 9$). Consequently, the basic configuration of the two-hidden-layer network is $345 : N_h : N_h : 10$, where N_h represents the number of hidden units in each of the two hidden layers. In the case of the face recognition problem, there are 10 304 (92×112) input units and 40 output units (40 subjects in the database). The basic configuration of the two-hidden-layer MLP is $10\,304 : N_h : N_h : 40$. In both problems, the output of each unit of the MLP is mathematically expressed as

$$f(net_j) = \frac{1}{1 + e^{-net_j}}, \quad net_j = \sum_i w_{ji} o_i + bias_j \tag{10}$$

where net_j is the weighted sum of the outputs o_i of all the units in the preceding layer (inputs in the first layer), and f is a sigmoidal activation function used as threshold. The weights w_{ji} denote the connection strength between unit i of the preceding layer, and unit j of the current layer. Besides, each unit has a bias $bias_j$, allowing a shift in the relative position of the sigmoidal function along the net_j axis.

III. RESULTS

A. Multilayer Perceptron Neural Classifier

In this subsection, we show the results obtained with an MLP neural classifier alone in the problems of handwritten digit classification and face recognition. These results will serve as a reference to compare with the results of our proposed model.

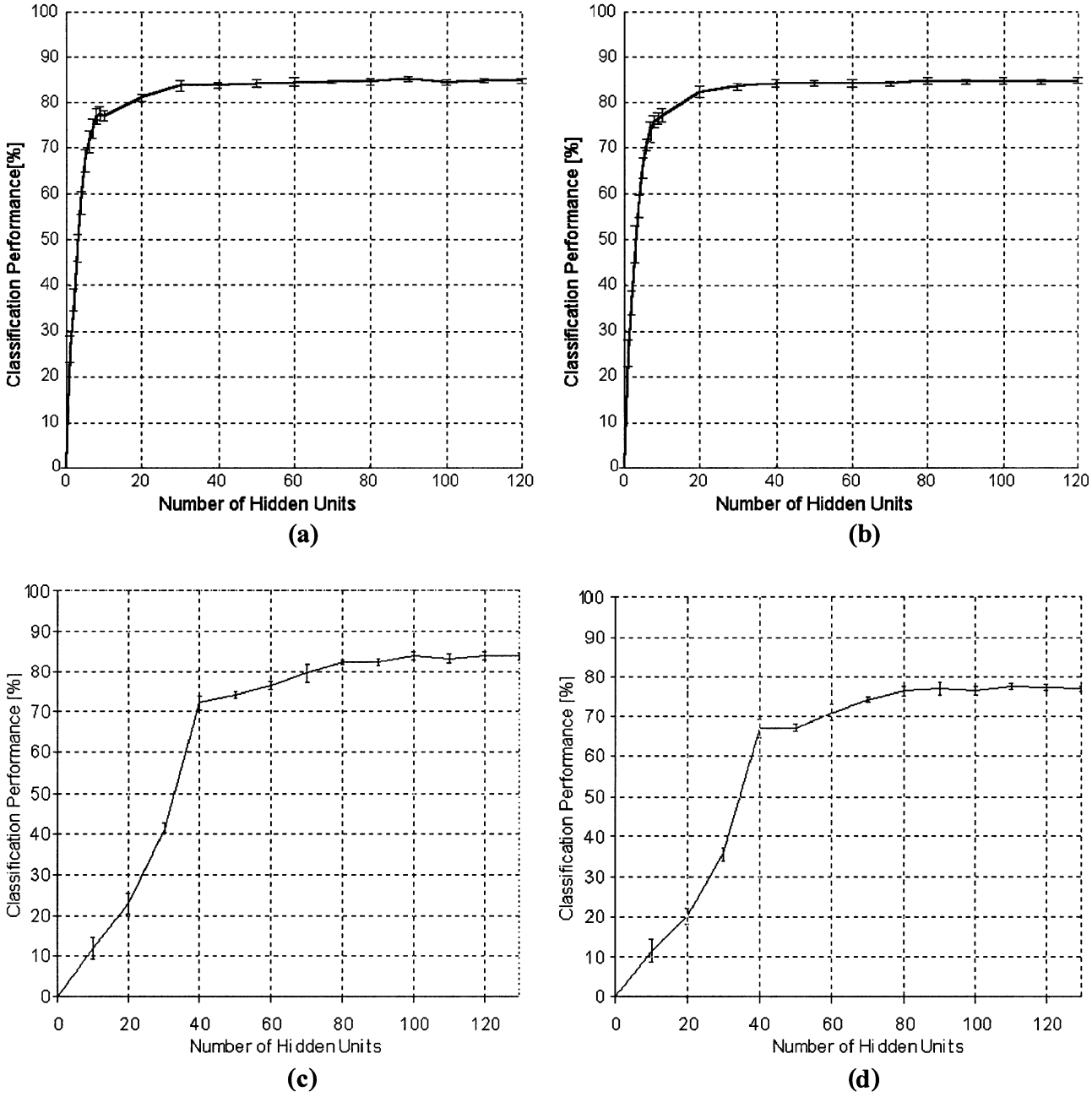

Fig. 4. Percentage of the (a) validation and (b) test patterns correctly classified by a two-hidden-layer MLP as a function of the number of hidden units, for the handwritten digit classification problem. Percentage of the (c) validation and (d) test patterns correctly classified by a two-hidden-layer MLP as a function of the number of hidden units, for the face recognition problem. Average and standard deviation for ten simulations with different weight initializations.

Fig. 4 shows the percentage of the validation and test patterns that are correctly classified by a two-hidden-layer MLP, as a function of the number of units in each hidden layer. As each network was trained for ten random starting weight sets, the average recognition rate and the standard deviation are shown in Fig. 4(a) and (b) for the validation and test set in the handwritten recognition problem, respectively. Fig. 4(c) and (d) shows the results for the validation and test sets for the face recognition problem. In all cases, it can be observed that the classification performance increases as a function of the number of hidden units for a small number of hidden units and then does not change significantly for a large number of hidden units. According to the ANOVA test, no statistically significant changes in classification performance were reached on the test set above 40 units in the hidden layers for the handwritten recognition problem and above 100 units in the face recognition problem. Among all number of hidden units, the best classification performance on the test set, computed as the average of ten simulations, for the handwritten digit problem reached 84.9% for an 80-hidden-unit MLP. The best recognition performance on the test set for the face recognition problem reached 77.5% for a 110-hidden-unit MLP.

B. Genetic Selection of FEN Architecture

1) Handwritten Digit Classification Problem:

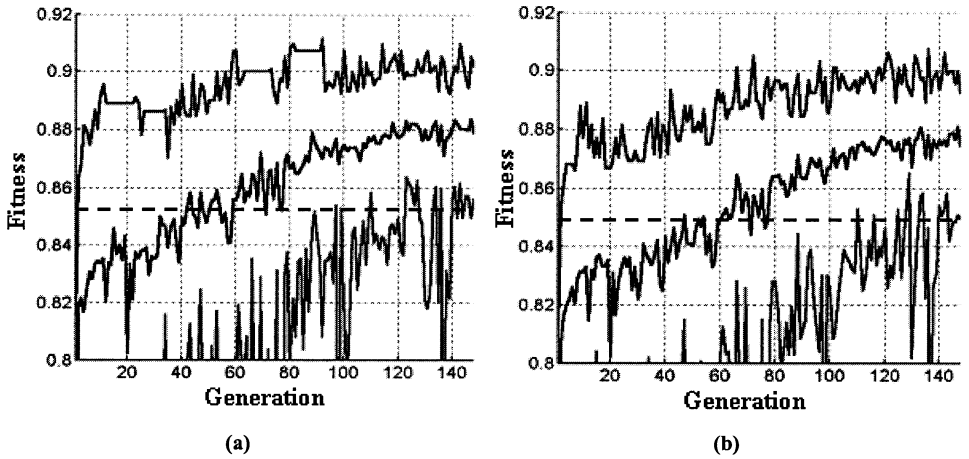

Fig. 5. Fitness for receptive field parameters and variable number of planes for 145 generations for the handwritten digit classification. Fitness of the best network (top), average fitness of the population (middle), and worst network (bottom) as a function of the generation number. (a) Validation set. (b) Test set.

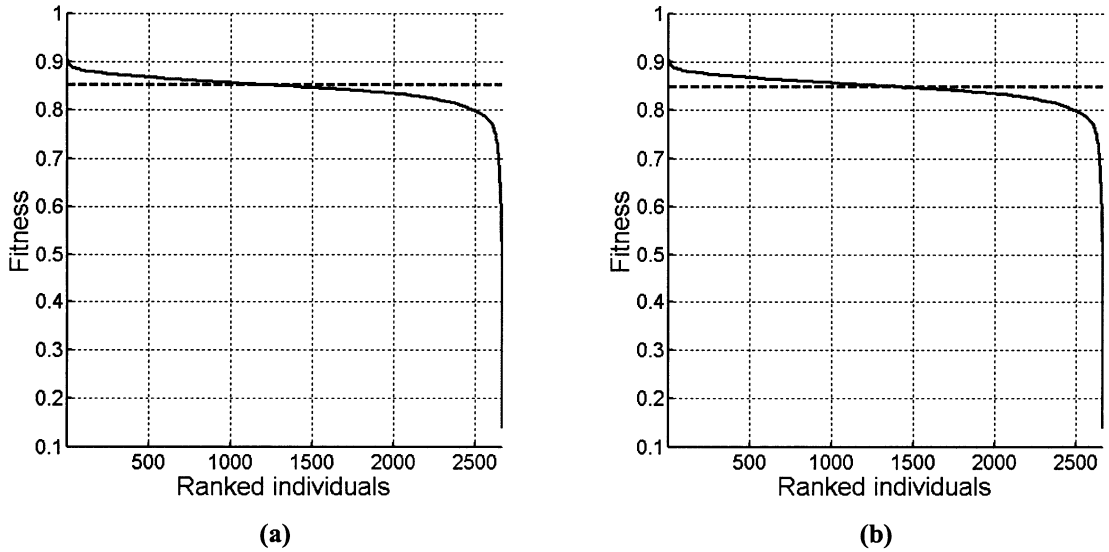

Fig. 6. Ranked classification performance by fitness including all GA individuals on the (a) validation set and (b) test set for the handwritten digit classification problem. The dashed line represents the best classification performance obtained by an MLP alone on the same database.

The results of the genetic selection of the FEN parameters for the problem of handwritten digit classification are presented in this subsection. A neural architecture with variable receptive field parameters and variable number of planes per layer was employed. Fig. 5 plots the best (top), average (middle), and worst (bottom) fitness for each population as a function of the generation number. Fig. 5(a) corresponds to the validation set and Fig. 5(b) to the test set. A total of 145 generations of 20 individuals each were computed. It can be observed that the classification performance of the system improved as the genetic algorithm selected appropriate parameters for the size, orientation and bias of the receptive fields as well as a proper number of planes per layer. The classification performance of the best individual on the validation set, as a function of the generation number, improved from 81.6% at the starting generation to 91.3%. The classification performance of the best individual on the test set, as a function of the generation number, improved from 81.3% to 90.8%. Therefore, the fitness of the best individual improved 9.7 and 9.5 percentage points in the validation and test sets, respectively. The horizontal dashed lines in Fig. 5(a) and (b) represent the best classification performance of 85.3% and 84.9%, reached by the best MLP architecture over the same validation and test sets, respectively.

Fig. 6(a) and (b) shows the classification performance of all individuals in the 145 generations ranked according to their fitness. These figures show that the difference between the classification performance of the best ranked individual and the worst one, changed between 91.3% and 14.1% (77.2 percentage points) in the 145 generations for the validation set and between 90.8% and 14.6% (76.2 percentage points) in the test set. In Fig. 6(a) and (b), the dashed line represents the best classification performance obtained by a two-hidden-layer MLP on the same validation and test sets. It can be observed that there are over 1200 individuals ranked with fitness above that of the two-hidden-layer MLP model. Therefore, a proper design of the receptive field architecture yields better classification performance than that of the MLP model alone. These results emphasize the significance of an appropriate receptive field design because a proper selection of receptive fields yields significantly better classification results than other choices.

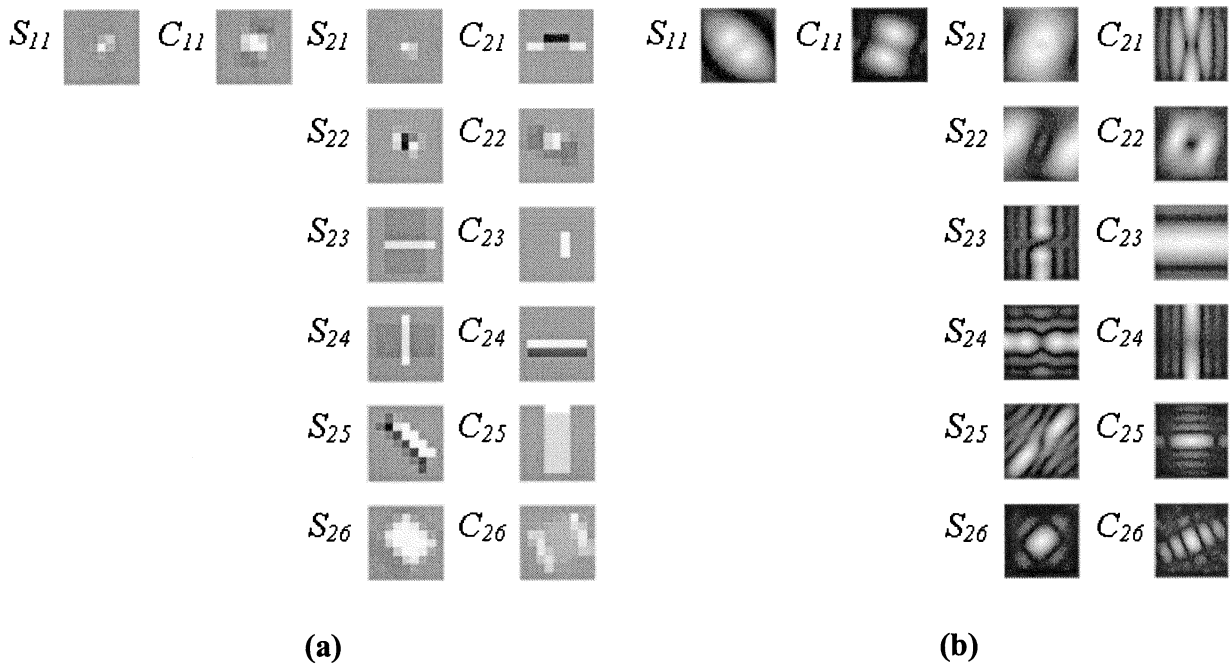

Fig. 7. (a) Architecture of the FEN for the best network in the problem of handwritten classification. Each column contains the number of different receptive fields in one layer: one receptive field in the first layer and five at the output layer. The gray levels represent the spatial configuration of the receptive fields. The background gray represents 0 and darker levels represent negative weights (inhibitory). (b) Spatial frequency filtering characteristics of each receptive field in the FEN (Gray scale: white is 1 and black is 0).

In the problem of handwritten digit classification, the GA selected as the best individual an FEN receptive field architecture composed of six planes of 12×8 units in the output layer and one plane in the input layer. The FEN applied to each input pattern can be interpreted as a nonlinear transformation from the 23×15 input image to produce six planes of 12×8 units. After the FEN, a neural classifier with 40 units in each of the two hidden layers requires a total of $6 \times 12 \times 8 \times 40 + 40 \times 40 + 40 \times 10 = 25\,040$ adjustable weights (biases are not counted). Without the FEN, a two-hidden-layer MLP architecture with 61 hidden units contains $15 \times 23 \times 61 + 61 \times 61 + 61 \times 10 = 25\,376$ adjustable weights which is comparable to the number of adjustable weights of the best network selected by the GA. Nevertheless, the MLP alone reaches only a maximum recognition rate of 84.9% with equivalent number of adjustable weights in the test set. Therefore, the FEN provides an improvement in classification performance of 5.9 percentage points relative to an MLP in the test set.

For the best individual in this simulation, Fig. 7(a) shows the layer type and the number of receptive fields in each layer. Fig. 7(a) also shows the receptive field spatial configurations in gray level. The white and black tones represent the maximum and minimum values, respectively, and the background gray level represents the value 0. Fig. 7(b) shows the spatial frequency filtering characteristics of each receptive field: white tone represents value 1 and black tone represents value 0. In Fig. 7(b), the reference for the spatial frequency coordinate system is at the center of each image, i.e., low frequencies are represented at the center and higher frequencies toward the edges.

Some of the receptive fields shown in Fig. 7(a) can be interpreted in terms of their main spatial filtering characteristics. Planes S_{11}, S_{21}, and S_{26} are mainly smoothing filters with different cutoff frequencies. The larger the excitatory part of the receptive field, the larger the smoothing effect on the input image. This effect can be observed in Fig. 7(b) where the spatial frequency characteristic of the respective receptive fields are represented. In Fig. 7(b), the spatial frequency response of plane S_{21} shows that only high frequencies at the edge of the graph will be filtered and therefore, a small smoothing effect will be produced by this plane. Nevertheless, the spatial frequency response of plane S_{26} shows that mainly low frequencies at the center will pass, thus with a strong low-pass spatial frequency result. Planes S_{23} and S_{24} are mainly line detectors in the horizontal and vertical directions, respectively. It also can be observed that filters such as S_{23} and S_{24} are orthogonal line detectors. The same for planes C_{23} and C_{24}. The plane S_{25} performs edge detection in 135°. The threshold function (4) applied after each plane generates additional nonlinear filtering effects. Interpreted from a filtering point of view, the FEN extracts different features from the original image that improve the classification performance of the combined network.

2) Face Classification Problem: The genetic selection of the receptive fields parameters for the face recognition problem are now presented. Fig. 8 shows the best (top) and average (bottom) fitness for each population as a function of the generation number. Fig. 8(a) is for the validation set and Fig. 8(b) for the test set. A total of 89 generations of 20 individuals each one were computed. The horizontal dashed line in Fig. 8(a), 84.2% and Fig. 8(b), 77.5%, represent the best classification results reached by the two-hidden-layer MLP alone on the validation and test sets, respectively. It can be observed that the GA selected several combination of the receptive field parameters (size, orientation, bias, and number of planes per layer) that

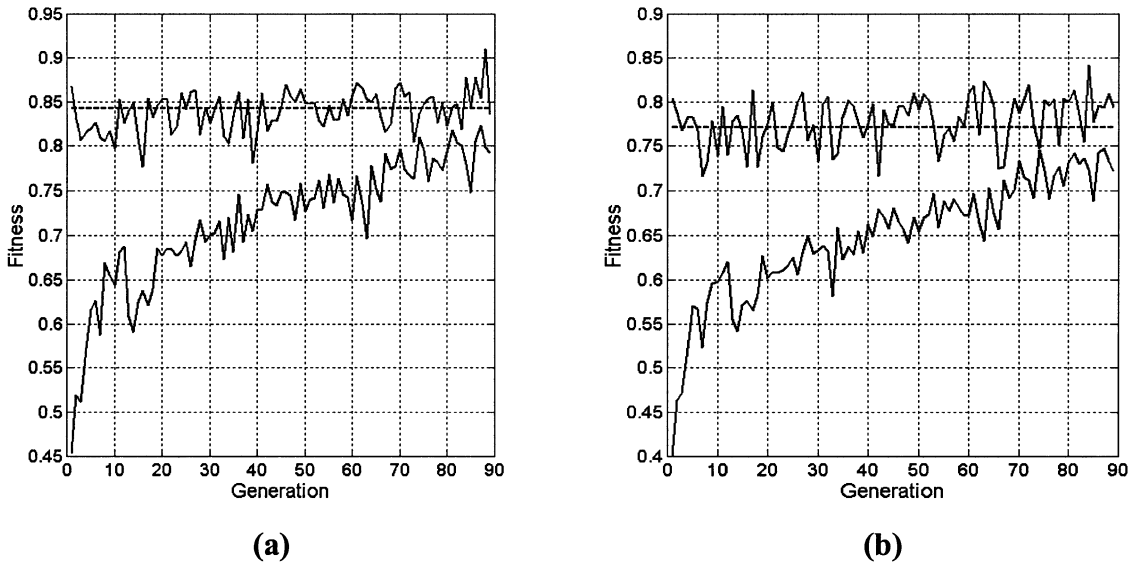

Fig. 8. Fitness for receptive field parameters and variable number of planes for 89 generations for the face recognition problem. Fitness of the best network (top) and average fitness of the population (bottom) as a function of the generation number. (a) Validation set. (b) Test set.

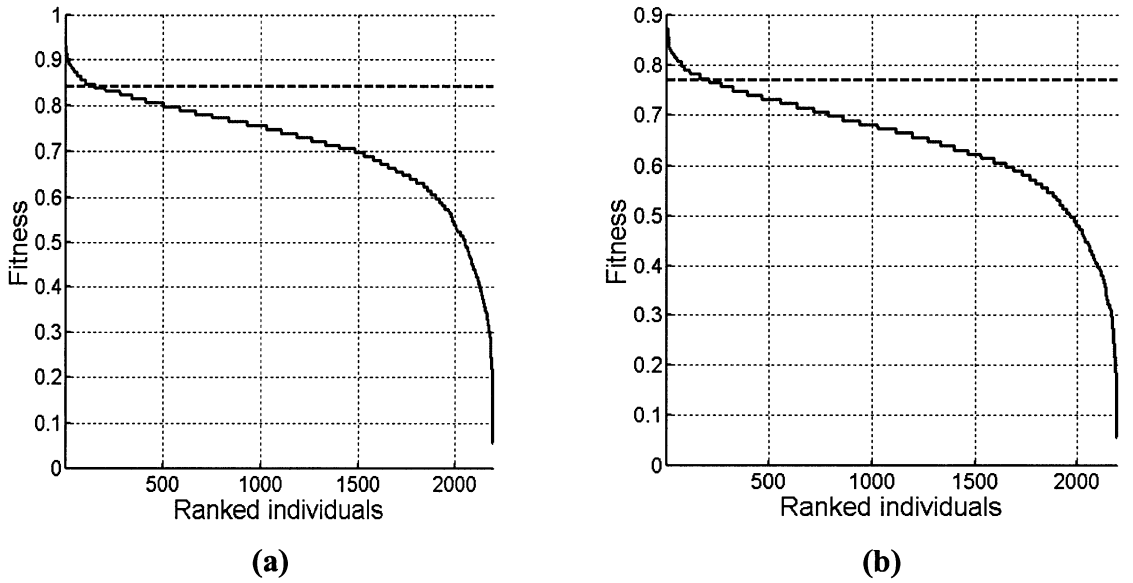

Fig. 9. Ranked classification performance by fitness including all individuals on the (a) validation set and (b) test set for the face recognition problem. The dashed line represents the best classification performance obtained by an MLP alone on the same database.

yielded better results than those reached by the MLP. The GA keeps a record of each individual and its fitness in every generation. Each individual in the GA encodes a receptive field architecture in a binary string. Therefore, as long as some of the individuals have reached classification performances above the MLP model, the receptive field architecture is known independently of the point where the GA is stopped. In fact, Fig. 9(a) and (b) shows the solutions ranked according to their fitness and it is found that several perform better than the MLP model. The GA found an individual with classification performance of 91.1% on the validation set when the best result at the starting generation was 86.6%. The best fitness found by the GA on the test set in the 89 generations was 84.2% and the best individual at the starting generation was 80.6%.

Fig. 9(a) and (b) shows the classification performance of the best individuals in the 89 generations ranked according to their fitness. These figures show that the difference between the classification performance for all ranked individuals ranged between 84.2% and 6% for the test set. It is also observed that there are several configurations of the FEN network that yield better results than the best solution reached by a two-hidden-layer MLP alone on the test set (77.5%). As in the handwritten recognition problem, these results show the significance of an appropriate receptive field design. In the face recognition problem, the receptive field spatial configuration for the best individual in the test set is shown in Fig. 10. The receptive fields are shown in gray levels in Fig. 10(a). White and black tones represent maximum and minimum values, respectively, and the background gray level represent value 0. Fig. 10(b) shows the spatial frequency characteristics of each receptive field (white represents value 1 and black tone represents value 0). The first two layers of the FEN are

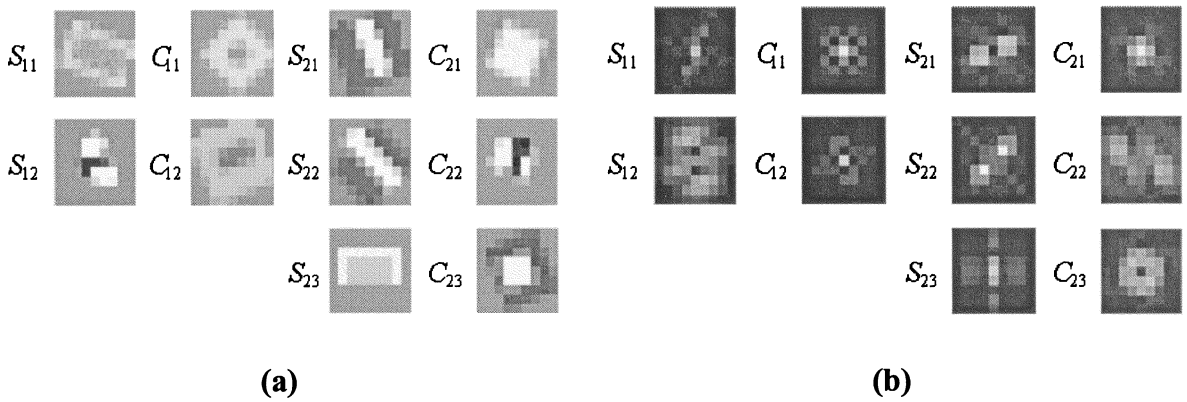

Fig. 10. (a) Architecture of the FEN for the best network in the face recognition problem. Each column contains the number of different receptive fields in one layer: two receptive fields in the first layer and three at the output layer. The gray levels represent the spatial configuration of the receptive fields. The background gray represents 0, darker levels represent negative weights (inhibitory) and lighter levels represent positive values (excitatory). (b) Spatial frequency filtering characteristics of each receptive field in the FEN (Gray scale: white is 1 and black is 0).

composed of two input S-type planes followed by two C-type planes. Then, the next two layers are composed of three S-type planes followed by three C-type planes. As in the case of the handwritten classification problem, the FEN can be interpreted as a nonlinear transformation from the 92×112 input image to a three plane image of 46×56 pixels. After the FEN an MLP neural classifier with 100 hidden units requires $3 \times 46 \times 56 \times 100 + 100 \times 100 + 100 \times 40 = 786\,800$ adjustable weights. A two-hidden-layer MLP with 75 hidden units has $92 \times 112 \times 75 + 75 \times 75 + 75 \times 40 = 781\,425$ adjustable weights, a number similar to that of the combined network selected by the GA. Nevertheless, the classification performance is significantly lower since the MLP reaches a maximum of 77.5% with 110 hidden units. In Fig. 10(a), the genetically evolved receptive fields, S_{21}, and S_{22} are selective in orientation and show excitatory center and inhibitory surround. Receptive fields S_{11}, C_{11}, and C_{12} have inhibitory center and excitatory surround and S_{11} and C_{12} are oriented in 45° and 135°, respectively. Receptive field C_{21} is a spatial low-pass filter with a large excitatory component and its spatial frequency response can be observed in C_{21} in Fig. 10(b).

The choice of three evaluations for each individual in the GA is a tradeoff between statistical significance and computational time employed. After ten simulations for different starting weight sets, the standard deviation for the MLP neural classifier is 2.5% in the handwritten digit classification problem and 2.3% in the face recognition problem, as shown in Fig. 4. These values correspond to 40 and 100 hidden units as used in cascade with the FEN in each problem, respectively. Considering these standard deviations, differences between recognition rates obtained with three evaluations are statistically significant with high level of confidence ($p < 0.05$) when the averages differ in approximately 3.6% in the range for averages around 80–91%. In the two problems treated in this paper, the difference between performance of each individual at the starting generation and the best fitness found by the GA is statistically significant. Furthermore, to evaluate the method it is used an independent test set that is different from the validation set used by the GA to evaluate the fitness of each individual.

Results of the handwritten digit classification and face recognition are summarized in Table IV for the test set. The recognition performance is shown for the fully connected MLP model alone and the combined network (FEN + MLP neural classifier). Improvements in classification performance by introducing the FEN network are 5.9 and 6.7 percentage points in each problem, respectively.

TABLE IV
RECOGNITION PERFORMANCE ON THE TEST SET FOR THE FULLY CONNECTED MLP MODEL AND THE COMBINED NETWORK, FEN + MLP CLASSIFIER MODEL

Problem	Fully Connected *MLP* Model [%]	*FEN+MLP* Classifier [%]
Handwritten Digit Recognition	84.9	90.8
Face Recognition	77.5	84.2

IV. CONCLUSIONS

This paper presented a new method for automatic design of biologically inspired receptive fields in feed forward NNs to enhance pattern recognition performance. It is proposed a neural architecture composed of two networks in cascade: an FEN followed by an MLP neural classifier. A GA was used to select the parameters of the FEN to improve the classification performance of the combined architecture. The GA optimized the size, orientation, and bias of biologically inspired receptive fields as well as the number of planes per layer in the FEN. The method was applied to the problems of handwritten digit classification and face recognition. Rather small databases were used in both problems to avoid excessive computational time, although results are expected to improve with larger databases.

Results of fully connected MLP NNs for different number of hidden units were presented as a reference to compare results of the proposed model. On the handwritten digit classification problem, results show that no further improvements were obtained by increasing the number of hidden units above 40. The best results for a two-hidden-layer MLP of 80 hidden units reached 85.3% on the validation set and 84.9% on the test set.

For the proposed combined model (FEN + MLP), after a genetic search through 145 generations, the selected architecture reached 90.8% of classification performance on the test set for the handwritten digit classification problem. This result is significantly better than the 84.9% obtained by fully connected

MLP model on the same test set. By ranking the classification performance of all the combined models with receptive fields selected by the GA, a strong dependency was found between the combined NN performance and the receptive field configuration (dimensions, orientation, bias and number of planes per layer). Differences between 90.8% and 14.6% of classification performance were found on the test set depending on the receptive field configuration of the FEN. As shown in Fig. 6, more than 1200 of the combined models performed better than the fully connected MLP model.

For the face recognition problem, the two-hidden-layer MLP model reached the best classification performance of 84.2% for the validation set and 77.5% for the test set. The results of the genetic selection of the FEN parameters reached a correct face recognition performance of 91.1% on the validation set and 84.2% on the test set. These results are significantly better than those reached by the fully connected MLP model on the same sets. Results of the ranked classification performance (see Fig. 9) on the face recognition problem after the genetic search for the FEN parameters show that the main improvements are due to the receptive field configuration. Results are compared among models with different receptive field configuration and to those of a fully connected NN where receptive fields are not explicitly defined.

It is possible to interpret the filtering effect of some of the receptive fields selected for the FEN. Some of them show spatial direction selectivity or spatial frequency selectivity. The FEN network can be interpreted as a nonlinear transformation over the input data producing a better separation among classes. It would be useful to study in the future the common characteristics of the receptive fields among the combined models that yielded best classification performance results.

This paper introduced a method for an automatic design of the receptive fields basic architecture and applied the method to improve the classification performance of two pattern recognition problems in 2-D. Modeling the biological architecture of the visual system is beyond the scope of this paper. Nevertheless, the proposed method makes use of the knowledge about the biological receptive field architecture to incorporate receptive fields into artificial neural models to create new architectures specifically tuned for a particular pattern recognition problem. Possible future developments include evolving more complex neural architectures to improve pattern recognition performance. The model should incorporate complex receptive fields as those found in the visual cortex to include curvature, corners and end-stopped line detectors. It would also be interesting to include the concept of modular specialization in the evolving model.

ACKNOWLEDGMENT

The authors would like to thank L. Medina for his help in the preparation of this manuscript.

REFERENCES

[1] F. W. Adams, H. T. Nguyen, R. Raghavan, and J. Slawny, "A parallel network for visual cognition," *IEEE Trans. Neural Networks*, vol. 3, pp. 906–922, Nov. 1992.

[2] F. Alexandre, F. Guyot, J. P. Haton, and Y. Burnod, "The cortical column: A new processing unit for multilayered networks," in *Neuro-Vision Systems*, M. Gupta and G. Knopf, Eds. Piscataway, NJ: IEEE Press, 1994, pp. 209–219.

[3] J. J. Atick and N. N. Redlich, "Toward a theory of early visual processing," *Neural Comput.*, vol. 2, pp. 308–320, 1992.

[4] J. E. Baker, "Reducing bias and inefficiency in the selection algorithm," in *Proc. 2nd Int. Conf. Genetic Algorithms and Their Applications*, J. Grefenstette, Ed. Mahwah, NJ: Lawrence Erlbaum, 1987, pp. 14–21.

[5] H. B. Barlow, "Possible principles underlying the transformation of sensory messages," in *Sensory Communications*, W. A. Rosenblith, Ed. Cambridge, MA: MIT Press, 1961.

[6] F. Z. Brill, D. E. Brown, and W. N. Martin, "Fast genetic selection of features for neural network classifiers," *IEEE Trans. Neural Networks*, vol. 3, pp. 324–328, Mar. 1992.

[7] P. Churchland and T. Sejnowski, "Thumbnail sketch of the mammalian visual system," in *The Computational Brain*, 4th ed. Cambridge, MA: MIT Press, 1996, pp. 148–157.

[8] ——, "Receptive fields: Size and center-surround organization," in *The Computational Brain*. Cambridge, MA: MIT Press, 1996, pp. 53–56.

[9] R. L. De Valois, D. G. Albrecht, and L. G. Thorell, "Spatial frequency selectivity of cells in macaque visual cortex," *Vis. Res.*, vol. 22, pp. 545–559, 1982.

[10] S. Deutsch and A. Deutsch, "Visual pattern recognition, neural networks, and household chores," in *Understanding the Nervous System*. Piscataway, NJ: IEEE Press, 1993, pp. 282–345.

[11] A. Dimitrov and J. Cowan, "Spatial decorrelation in orientation—Selective cortical cells," *Neural Comput.*, vol. 10, pp. 1779–1795, 1998.

[12] U. Ernst, K. Pawelzik, M. Tsodyks, and T. Sejnowski, "Relation between retinotopical and orientation maps in visual cortex," *Neural Comput.*, vol. 11, pp. 375–379, 1999.

[13] K. Fukushima, "Neocognitron: A hierarchical neural network capable of visual pattern recognition," *Neural Networks*, pp. 119–130, 1988.

[14] K. Fukushima, S. Miyake, and T. Ito, "Neocognitron: A neural network model for a mechanism of visual pattern recognition," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, pp. 826–834, Sept./Oct. 1983.

[15] K. Fukushima and S. Miyake, "Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position," *Pattern Recognit.*, vol. 16, no. 6, pp. 455–469, 1982.

[16] K. Fukushima, "Neural network model for a mechanism of pattern recognition unaffected by shift in position—Neocognitron," *Trans. IECE Jpn.*, vol. 62-A, no. 10, pp. 658–665, 1979.

[17] C. Fyfe and R. Baddeley, "Finding compact and sparse distributed representations of visual images," *Network*, vol. 6, pp. 333–344, 1995.

[18] D. J. Granrath, "The role of human visual models in image processing," *Proc. IEEE*, vol. 69, no. 5, pp. 555–561, 1981.

[19] J. Grefenstette, "Proportional selection and sampling algorithms," in *Evolutionary Computation 1: Basic Algorithms and Operators*, T. Bäck, D. B. Fogel, and T. Michalewics, Eds. Briston, U.K.: IOP, 2000, pp. 172–180.

[20] M. M. Gupta and G. K. Knopf, "Neuro-vision systems: A tutorial," in *Neuro-Vision Systems: Principles and Applications*, M. M. Gupta and G. K. Knopf, Eds. Piscataway, NJ: IEEE Press, 1994, pp. 1–34.

[21] J. M. Hannan and J. M. Bishop, "A comparison of fast training algorithms over two real problems," in *Proc. 5th Int. Conf. Artificial Neural Networks*. London, U.K.: IEE, 1997, conf. pub. 440, pp. 1–6.

[22] Handwritten database [Online]. Available: http://alcatraz.die.uchile.cl/handwritten_digit.htm

[23] S. Haykin, "Multilayer perceptrons," in *Neural Networks: A Comprehensive Foundation*, J. Griffin, Ed. Piscataway, NJ: IEEE Press, 1994, pp. 138–229.

[24] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: Univ. of Michigan Press, 1975, Second edition: MIT Press.

[25] D. Hubel, "The primary visual cortex," in *Eye, Brain and Vision*, ser. No. 22: Scientific American, 1995, pp. 59–91.

[26] D. H. Hubel, *The Visual Cortex of the Brain*: Scientific American, 1963, vol. 209, pp. 54–63.

[27] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Phys.*, vol. 160, no. 1, pp. 106–154, 1962.

[28] S. W. Kluffer, "Discharge patterns and functional organization of mammalian retina," *J. Neurophys.*, vol. 16, no. 1, pp. 37–68, 1953.

[29] R. E. Kronauer and Y. Y. Zeevi, "Reorganization and diversification of signals in vision," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, no. 1, pp. 91–101, 1985.

[30] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 4–22, Apr. 1987.

[31] D. R. Lovell, "The neocognitron as a system for handwritten character recognition: Limitations and improvements," Ph.D. dissertation, Univ. of Queensland, Queensland, Australia, July 1994.
[32] S. G. Mallat, "Multi frequency channel decompositions of images and wavelet models," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 2091–2110, Dec. 1989.
[33] T. Ming-Yeong, K. Li-Pheng, and S. Siang-Kok, "Application of genetic algorithms to optimize neocognitron network parameters," *Neural Network World*, no. 3, pp. 293–304, 1997.
[34] M. Mitchell, "Evolving neural networks," in *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press, 1996, pp. 65–83.
[35] ——, "Genetic algorithms: An overview," in *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press, 1996, pp. 1–31.
[36] ——, "Implementing a genetic algorithm," in *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press, 1996, pp. 155–178.
[37] R. Näsänen, "Spatial frequency bandwidth used in the recognition of facial images," *Vis. Res.*, vol. 39, pp. 3824–3833, 1999.
[38] C. Neubauer, "Evaluation of convolutional neural networks for visual recognition," *IEEE Trans. Neural Networks*, vol. 9, no. 4, pp. 685–696, 1998.
[39] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, pp. 607–609, 1996.
[40] ——, "Sparse coding with an overcomplete basis set: A strategy employed by V1?," *Vis. Res.*, vol. 37, pp. 3311–3325, 1997.
[41] L. A. Olzak and J. P. Thomas, "Neural recording in human pattern vision: Model and mechanisms," *Vis. Res.*, vol. 39, pp. 231–256, 1999.
[42] C. A. Perez, C. Salinas, and P. Estevez, "Designing biologically inspired receptive fields for neural pattern recognition technology," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Tucson, AZ, Oct. 7–10, 2001, pp. 58–63.
[43] C. A. Perez and C. Salinas, "Genetic neural architecture design based on visual receptive field organization for face recognition," in *Annu. Biomedical Engineering, Abst. Suppl., Biomedical Engineering Soc., 2001 Annu. Fall Meeting*, Durham, NC, Oct. 4–7, 2001, p. S-122.
[44] ——, "Biologically inspired receptive fields for computational vision," in *Proc. 1st Joint BMES/EMBS Conf.*, Atlanta, GA, Oct. 13–16, 1999, p. 924.
[45] C. A. Perez and C. A. Holzmann, "Improvements on handwritten digit recognition by genetic selection of neural network topology and by augmented training," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Orlando, FL, Oct. 12–15, 1997, pp. 1487–1491.
[46] C. A. Perez and G. Marin, "Handwritten digit recognition by a neural network with slope detection and multiple planes per layer architecture," in *Proc. 15th Annu. Int. Conf. IEEE/EMBS*, San Diego, CA, October 28–31, 1993, pp. 275–276.
[47] D. Rumelhart, G. Hinton, and R. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, D. Rumelhart and J. McClelland, Eds. Cambridge, MA: MIT Press, 1986, vol. 1, pp. 318–362.
[48] G. Syswerda, "Uniform crossover in genetic algorithms," in *Proc. 3rd Int. Conf. Genetic Algorithms*. San Mateo, CA: Morgan-Kaufmann, 1989, pp. 2–9.
[49] *The Database of Faces*. Cambridge, MA: AT&T Laboratories. [Online]. Available: http://www.uk.research.att.com/facedatabase.html.
[50] B. A. Wandell, "The cortical representation," in *Foundations of Vision*: Sinauer Assoc., 1995, pp. 153–193.
[51] A. B. Watson and A. J. Ahumada, "A hexagonal orthogonal-oriented pyramid as a model of image representation in visual cortex," *IEEE Trans. Biomed. Eng.*, vol. 36, no. 1, pp. 97–106, 1989.
[52] T. Whitfort, B. Choi, C. Metthews, and J. Mccullagh, "An evolutionary approach to the specification of high performing backpropagation neural networks," in *Proc. Int. Conf. Neural Information Processing and Intelligent Information Systems*, N. Kasabov, R. Kozma, R. O'Shea, G. Coghill, and T. Gedean, Eds. New York: Springer-Verlag, 1998, p. 4.
[53] M. J. Wright, "Training and testing of neural net window operators on spatiotemporal image sequences," in *Neural Networks for Vision, Speech, and Natural Language*, R. Linggard, D. J. Myers, and C. Nightingale, Eds., 1992, pp. 93–111.
[54] X. Yao, "Evolving artificial neural networks," *Proc. IEEE*, vol. 87, pp. 1423–1447, Sept. 1999.
[55] D. S. Yeung, Y. T. Cheng, H. S. Fong, and F. L. Chung, "Neocognitron based handwriting recognition system performance tuning using genetic algorithm," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, San Diego, CA, Oct. 11–14, 1998, pp. 4228–4233.

Claudio A. Perez (M'90) received the B.S. degree in electrical engineering, the P.E. degree in electrical engineering, and the M.S. degree in biomedical engineering, all from the Universidad de Chile, Santiago, in 1980 and 1985, respectively, and the Ph.D. degree in biomedical engineering from The Ohio State University (OSU), Columbus, in 1991.

Currently, he is an Associate Professor at the Department of Electrical Engineering, Universidad de Chile, where he is also the Graduate Studies Chairman. His research interests include computational vision, including biologically inspired models for pattern recognition and man–machine interfaces. He has been an Invited Speaker at OSU and is participating in a Fulbright Faculty Exchange between the University of California, Berkeley, and Univeridad de Chile.

Dr. Perez is a member of Sigma-Xi, BMES, the Pattern Recognition Society, and the OSU Alumni Association. He received a Fulbright Fellowship in 1986 to pursue his Ph.D. degree at OSU. In 1990, he was a Presidential Fellow and received a Graduate Student Alumni Research Award from OSU. In 1991, he received a Fellowship for Chilean Scientists from Fundacion Andes.

Cristian A. Salinas received the M.S. degree in biomedical engineering from the Universidad de Chile, Santiago, in 2001. His thesis was on pattern recognition using neural networks tuned by genetic algorithms. He is currently pursuing the Ph.D. degree in biomedical engineering at Case Western Reserve University, Cleveland, OH. His research interest include computational vision, neural networks to visual pattern recognition, genetics algorithms, and digital image processing.

Pablo A. Estévez (M'98) received the degree in electrical civil engineering from the Universidad de Chile, Santiago, and the Ph.D. degree from the University of Tokyo, Tokyo, Japan, in 1995.

Currently, he is an Assistant Professor in the Department of Electrical Engineering, Universidad de Chile. He has been an Invited Researcher at the Communication Science Laboratory, NTT-Kyoto, Japan, and at the Ecole Normale Superieure de Lyon, France. His research interests include the application of neural networks and genetic algorithms to pattern recognition and biomedical issues.

Dr. Estévez is a Member of the INNS and Forest Products Society. He received a Mombusho Fellowship to pursue his Ph.D. degree at the University of Tokyo.

Patricia M. Valenzuela received the M.D. degree and the M.S. degree in pediatrics from Pontificia Universidad Católica de Chile (PUC), Santiago, in 1983 and 1986, respectively, and the M.S. degree in preventive medicine from The Ohio State University (OSU), Columbus, in 1991.

From 1987 to 1991, she was a Visiting Scientist at the Department of Toxicology and Pharmacology at Children's Hospital, Columbus. Currently, she is an Assistant Professor in the Pediatrics Department at PUC. Her research interest include sensory systems, lead poisoning, and preventive medicine in children.

Dr. Valenzuela is a member of Phi Beta Delta, the Honor Society for International Scholars, the Preventive Medicine Alumni Society of OSU, and the Chilean Pediatrics Society.