Application of Artificial Neural Networks to Pattern Recognition

Işıl Celasun, M. Burçin Gökgöz, M. Kemal Cılız and Yorgo Istefanopulos
Department of Electrical Engineering, Boğaziçi University, Bebek 80815, Türkiye.

Abstract

In this paper, the recognition problem of fighter / bomber planes is investigated by using artificial neural networks as pattern recognition tools. The aim was to remove noise effects and to represent images in such forms that they are quite smaller in size and so that compatible with the neural network characteristics. There are lots of redundant information in images which should be discarded to increase the performance of neural networks. Five different models of aircrafts are used to train the artificial neural networks.

1. INTRODUCTION

Reference images are obtained by using a Vidicon camera that is connected to a PC through a digitizer circuit.The images are taken under bad (nonuniform) lighting conditions to simulate real cases. Grabbed images consist of 256*256 pixels where each has a value of intensity in the range [0,63]. Then 108*108 subimages that contain the picture of airplanes are extracted. Binary images are obtained by thresholding these gray-level images.

An aircraft which was initially represented by its digitized image is embedded with Gaussian noise having different signal to noise ratio values (6, 12, 18 and 24) and with non-Gaussian noise(Uniform noise in the interval [-9,+9]) Fig.2. Two noise removal techniques which are smoothing and cellular neural networks (CNNs) are applied to remove the noise in the above two cases and then they are compared in terms of performance. Images of airplanes with two different orientations and distances are taken in order to test whether the recognition method is rotational and scale invariant respectively.

Hence, artificial neural network models are used as a noise removal tool and as a classifier. These issues are discussed in sections 2 and 3.

2. NOISE REMOVING METHODS

Cellular neural networks are used as a noise removal technique in [4]. The main difference of CNNs is their nearest neighborhood connection structure as compared to a fully connected Hopfield net. Neurons are only connected to their nearest neighbors in the network matrix. Adjacent cells can interact directly with each other and cells not directly connected may also affect each other indirectly because of the propagation effects of the continuous time dynamics of the network. CNNs can be programmed to perform several tasks through their neighborhood connection weights which are also called cloning templates. A cloning template is a space invariant matrix that shows the connection weights of the cells in a neighborhood. Connection weights are twofold. A matrix A that holds the weights for output voltages of neighboor cells, which is also called the feedback operator and a matrix B that holds the weights for input voltages of neighbor cells, the control operator. The averaging operator which is used as the feedback operator, is chosen as the dynamic rule for noise removing CNN. The steady state of a cell C(i,j) depends on the average of those of its neighbor cells. This feedback operator has been used both for Gaussian and non-Gaussian noise embedded cases of images. Cellular neural network gives better performance for all the aircrafts in all positions when embedded in Gaussian noise with different variance values but approximately similar results are obtained with using smoothing and CNN when images are embedded in Uniform noise. To improve the performance obtained, some changes can be done on the feedback operator for non-Gaussian noise embedded images.

Table 1.
Table presenting the average number of wrong pixels for a 108*108 image after noise removing process using the two methods discussed above ("Smth." denotes smoothing).

	G6		G12		G18		G24	
	Smth.	CNN	Smth.	CNN	Smth.	CNN	Smth.	CNN
F15	153	118	154	121	157	124	162	130
F117	119	91	134	128	134	134	150	150
MIG27	302	224	303	226	306	226	306	226
MIRAGE	79	36	85	39	86	42	88	43
F16	178	129	181	139	189	145	194	153

3. CLASSIFICATION

Five types of airplanes are employed for classification which are F15, F16, MIG27, F117, MIRAGE. To test the recognition performance for the methods that will be discussed in the sequel, images of airplanes with two different orientations and distances are taken into account. The planes are rotated about 45 degree to change the orientation and taken two times far away to scale it. (Fig. 1(b) and (c)) Recognition process is performed for three forms of five airplanes (normal, rotated and scaled) with two type of noise(Uniform and Gaussian with four different variance values).

Recognition process is based on using pecstrum and multilayer perceptron (MLP). The first aim of using pecstrum is to provide translation, rotation and scale invariance properties and the second is to provide small input vectors to MLP. Pecstrum is calculated by applying morphological operations to the two-valued images. This approach is generally based on the analysis of two-valued images in terms of some predetermined geometric shape known as a structing element.

There exists a well developed morphological algebra which is expressible as digital algorithms in terms of a few primitive morphological operations. Two of the fundamental operations are called Minkowski addition and Minkowski subtraction. Minkowski addition adds two regions on the analytical plane by adding all points of one region to those of the second. Minskowski subtraction finds points by first taking the symmetric region of the second operand about the origin, and then selecting those points if their addition with the symmetric region is on the first region. Erosion operation is defined as the subtraction of the symmetry of a region from another. If the second operand is already symmetrical around the origin, the erosion is equivalent to the Minskowski subtraction. Dilation operation is equivalent to the Minskowski addition. While erosion has an effect of shrinking the shapes, dilation expands them. Besides the fundamental operations, two other operations play a central role in image analysis. One of them is opening which is an erosion followed by a dilation. The other is closing which is a dilation followed by an erosion.

Pecstrum consists of a set of vectors where each gives the fractional change in the area of a shape when opening or closing is applied. The image is taken as the first operand of the opening operation. The second operand is a symmetrical region around the origin which is called the structing element. Pecstrum vector set is obtained by applying several opening operations on the original image while increasing the size of the structing element by a certain amount at each step. Two examples of opening sequences are illustrated in Fig.4 and Fig.5. Pecstrum is translational invariant since it is not related to the position but to the area and the form of the shapes. It is also rotational invariant provided the structing element is a circular one[4]. Pecstrum with above properties is not scale invariant. However, if every shape is prescaled to a standard area, pecstrum sets for the images of the same figure with different distances will be close to each other.

The pecstrum consists of at most 6 real values in our implementation (For 5 reference images; F15 : 5 , F117, :6, MIG27 : 3, MIRAGE : 5, F16 : 5). Binary form of pecstrum set is also presented to MLP network as well as pecstrum itself. In this case binary codes are concatenated to form an input vector to MLP. The length of such input vector is 36 instead of 6, since each element of pecstrum is converted into binary code up to 6 bit precision. In our implementation, these two options give the same result.

The MLP net consists of few layers where each contains a number of perceptrons.

Each perceptron in a layer has a connection with all perceptrons of the previous and the next layers. The network has an input layer of which perceptrons accept the input values to the network. The last layer of the network outputs the result of the operations. Each perceptron has a nonlinearity that maps the sum of incoming values to a single output value. Connections between perceptrons have weights which are multiplied by the output of the outcoming perceptron to calculate the input of the incoming perceptron. Weights are adjusted by the backpropagation algorithm while the network is trained.

A neural net consisting of 3 layers is used. The input layer has 36 nodes. Two hidden layers contain 45 and 30 nodes. The output layer has 5 nodes where each one is expected to output "1" while the others "0" for each airplane.

This method which is based on artificial neural network is compared with a conventional one which makes use of Fourier descriptors(FD) in shape representation. FDs in shape representations are also translational, rotational ans scale invariant. In this representaion a closed contour surrounding the shape is sampled by a certain number N of points. The axis and ordinate of each point constitute the real and imaginary part of a complex number respectively. Then a set of N complex Fourier coefficients are obtained and calculated by :

$$f_k = \frac{1}{N}\sum_{n=1}^{N} z_n \exp(-j2\pi kn/N)$$

where $k = -n_1, \ldots, n_2$ with $n_1+n_2+1 = N$.

These values are taken as N=128, $n_1=63$ and $n_2=64$ in the implementation. Comparison between two FDs F and G can be achieved by crosscorrelating them by the formula:

$$c = \max_n \left| \frac{1}{N} \sum_{k=-n_1}^{n_2} \frac{f_k * g_k}{\|f_k\| \, \|g_k\|} \exp(-j2\pi kn/N) \right|$$

or

$$c' = \max_n \left| \frac{1}{N} \sum_{k=-n_1}^{n_2} \frac{f_k * g'_k}{\|f_k\| \, \|g_k\|} \exp(-j2\pi kn/N) \right|$$

where g_k' is equal to $g_{n_2-n_1-k}$ then a distance term between them is d(F,G)=min(d , d') where $d^2 = 2[1-c]$ $d'^2 = 2[1-c']$ with d' being the distance term for the flipped form of the original shape. First five FDs for five different types of planes used in the application are found and stored then at the classification stage , a Fourier descriptor computed from the new image is compared by each of those five FDs. The minimum distance is taken for the classification scheme.

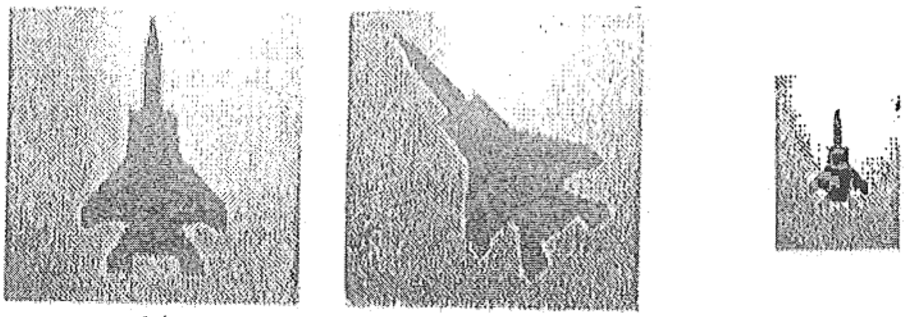
4.SIMULATION RESULTS

Simulations are done on a 286/16 PC with coprocessor and Turbo Pascal 6.0 is used as the programming language. Once the FDs of reference images are computed the recognition process of new coming noise removed image is completed in exactly 60 seconds. On the other side once the network is trained, the computation of pecstrum and recognition by the network takes 25 seconds for an image in average. The result of two classification methods are presented in Table 2. Fourier descriptor method best recognizes F15 but for example F16 is generally misrecognized as being confused with F15 in recognition. Pecstrum plus MLP recognizes F15 and F16, MIG27 almost in all cases. This method has the best performance at recognizing the similarly oriented planes with the reference planes and close performances in rotated orientation and the scaled cases

Table 2.

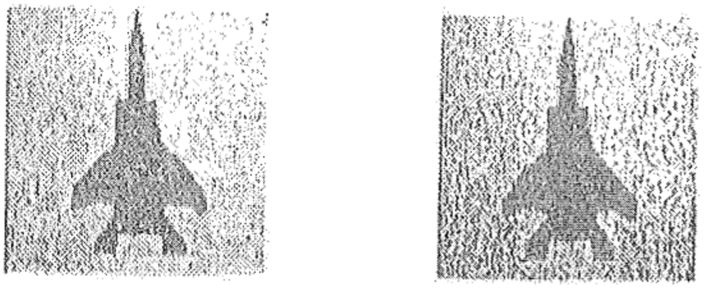
Comparison of the two methods in recognition of the airplanes ("Pec." denotes pecstrum and ■ , + denote the correct recognition for FD and Pec.+MLP respectively)

	Noise	F15		F117		MIG27		MIRAGE		F16	
		FD	**Pec.+ MLP**	**FD**	**Pec.+ MLP**	**FD**	**Pec.+ MLP**	**FD**	**Pec.+ MLP**	**FD**	**Pec.+ MLP**
Normal	G6	■	+	■	+	■	+	■	+		+
	G12	■	+	■	+	■	+	■			+
	G18	■	+	■		■	+	■	+		+
	G24	■	+	■		■	+	■	+		+
	U18		+				+		+		+
Rotated	G6	■	+		+		+	■			+
	G12	■	+				+	■			+
	G18	■	+				+	■			+
	G24	■	+					■			+
	U18		+				+		+		+
Scaled	G6		+		+		+				
	G12	■				■	+				+
	G18	■					+				+
	G24		+			■	+			■	+
Smaller	U18	■				■	+				+



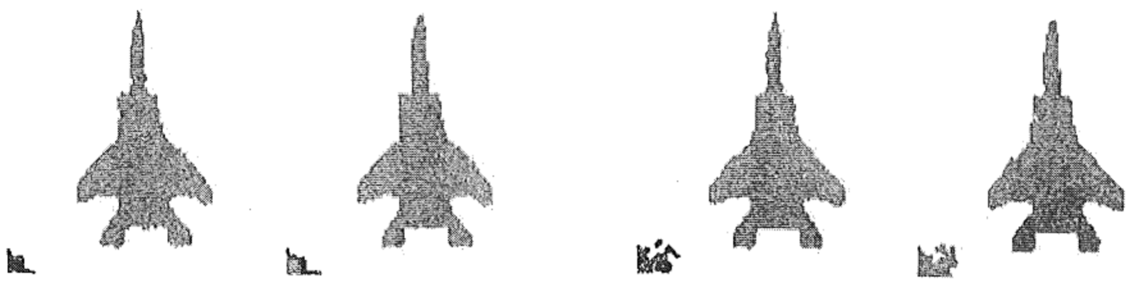
(a) (b) (c)

Figure 1. Gray level images of F15 (a) 108*108 reference image , (b) 45 degree rotated, (c) two times distant pose .



(a) (b)

Figure 2. Noisy images of F15 (a) Gaussian noise with variance 18 is added (b) Uniform noise is added.



S CNN S CNN

(a) (b)

Figure 3. Binary forms of noise removed version of F15 by smoothing (denoted by S) and using CNN of (a) Gaussian with variance 18 (denoted by G18) (b) Uniform noise (denoted by U) added image of F15

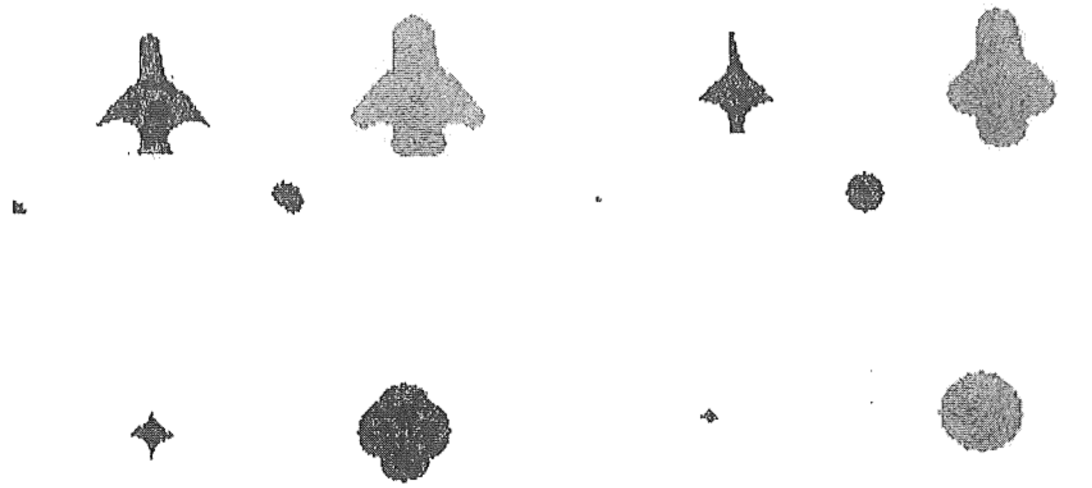


Figure 4. Opening sequence for F15.

Figure 5. Opening sequence for MIG27.

5. CONCLUSIONS

With using the morphological properties such as thickness ratio, relative area of some components of the shape, different pecstrums are available for all airplanes. So, in the recognition of objects with silhouettes like airplanes pecstrum provides good results. On the other hand, Fourier descriptor method uses the contour characteristics (frequency) of the shape. Since the contours obtained from the shapes of some planes are not much different from each other, 'FD' is not as good as 'pecstrum + MLP' in recognition performance for all cases discussed previously .

6. REFERENCES:

1. R.O. Duda, P.E. Hart, Pattern Classification and Scene Analysis, John Wiley & Sons, 1973.

2. Y. Istefanopulos, B. Sankur, I. Celasun, O. Toker, "Variational approach to pattern recognition using third order moments", International Signal Processing Workshop on Higher Order Statistics, France, July 1991.

3. Anil K. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, New Jersey, 1989.

4. L. Chua, Lin Yang, "Cellular neural networks: Applications" IEEE Trans. CAS, Vol. 35, No.10, pp.1273-1290

5. V. Anastassopoulos, A.N. Venetsanopoulos, " The classification properties of the pecstrum and its use for pattern identification", Circuits Systems Signal Processing, Vol. 10, No.3, pp. 293-326, 1991.

6. I. Pitas, A. Venetsanopoulos, " Morphological Shape Description", IEEE Trans. PAMI, Vol. 12, No. 1, pp.38-45, Jan. 1990.

7. T. Roska, et al., "Detecting simple motion using CNN's" Proc.IEEE First Int. Workshop of CNNA, pp. 127-138, Budapest, 1990.