COLOR IMAGE IDENTIFICATION BASED ON 2-D POWER SPECTRUM BASED ON NEURAL NETWORK

M.Sc. Hassan J. Hassan*

Received on: 30 / 4/2006
Accepted on: 16 / 1/2007

Abstract

Image identification plays a great role in industrial, remote sensing, and military applications. It is concerned with the generation of a signature to the image.

This work proposes a dynamic program (use Neural Network) to identify the color image depending on the distribution of the monochrome colors (red, green, and blue) in the same image to make image signature accordingly, which is represented by a values named power spectrum. The first step is to analyze the three-band monochrome image (color image) to Red, Green and Blue image, then deal with each image as a grey scale one which is represented as a 2-D matrix. The second step is to make Fourier Transform to each grey scale image in order to extract the implicit information in that image. The calculations of 2-D Power Spectrum for each image have been done to construct the final feature vector for each one. Finally, in the third step, and in order to handle problems of large input dimensions, a multilayer perceptron Neural Network has been used with two hidden layers. The input of the Neural Network structure is the final feature vectors which are obtained from the previous step. All programs are written using MATLAP VER. 6.5 programming language.

Keywords: **Image Recognition, Neural Network, Transformation Technique, Texture Classification.**

الخلاصة

إن تحليل نسيج الصورة لغرض التعرف عليها وتشخيصها يلعب دورا كبيرا في مجالات شتى, منها الصناعة والاستشعار من بعد إضافة إلى التطبيقات العسكرية.هذا البحث يقترح برنامج ديناميكي يعتمد على شبكة الخلايا العصبية (Neural Network) في تشخيص الصور الملونة, معتمدا على توزيع اللون الاحادي (الاحمر, الاخضر, والازرق) في تلك الصورة لغرض تكوين البصمة الخاصة التي تميزها والمتمثلة بحساب طيف الطاقة (power spectrum).**الخطوة الأولى** تبدأ بتحليل الصورة الملونة والمتكونة من الألوان الأحادية الأساسية (الاحمر, الاخضر, والازرق) والتعامل مع كل صورة أحادية ناتجة كصورة رمادية (grey scale image) والتي يتم تمثيلها بمصفوفة ذات بعدين.**الخطوة الثانية** يتم من خلالها تحويل الصورة الأحادية باستخدام محول فورير(Fourier transformer) لغرض استخلاص المعلومات الضمنية في تلك الصورة وحساب طيف الطاقة(power spectrum) , وبالمحصلة يتم تكوين متجة الخواص (feature vector) لكل صورة احادية.**الخطوة الثالثة والاخيرة** ولغرض التمكن من التغلب على عمليات التطبيق ذات المداخل الكثيرة, تم استخدام شبكة الخلايا العصبية (Neural Network) ذات الطبقتين المتواريتين (two hidden layers) . أما المدخلات لهذة الشبكة فانها تتمثل بمتجة يتكون من مجموعة من القيم تعكس كيفية توزيع الألوان الأساسية في كل صورة والتي تم تكوينها من الخطوة الثانية.

*Control and System Department University of Technology Baghdad-Iraq

1-Introduction

Image identification, digital signal processing, computer vision, and image classification are some of the fastest growing areas of computer science. The applications of the above areas in industry, medicine, the military, geology, biology, and other fields, coupled with the progress in computer hardware suggests the continual growth of these areas at ever increasing rates. The concern of most researchers in these areas has been to improve the quality of the images and/or to compute some information contained in them with the aim of understanding their content [1].

In image processing, the input is an image, and the output is also an image. On the other hand, in computer vision, the input is an image and the output consists of information about the image itself that may be used by the other process. Thus, the output will not be an image any more [2].

Identification systems often begin with some kind of preprocessing to remove noise and redundancy in the measurements through gray level modification, sharpening and smoothing, thereby ensuring an effective and efficient image description. Image representation and description involves the study of image characteristic measurements, which could be either numeric or non-numeric. The decision-making phase deals with the design of the identification system and selection of similarity measures for matching. It is quite clear that image-processing techniques are always the first stage of computer vision; moreover, sometimes image processing algorithms are used as feature extraction algorithms in the first step of the image identification procedure [1].

Several techniques for digital signal processing have been proposed for image identification. These can be generally classified into two main branches namely transformation and filtering. In fact great attention was given to the transformation [3].

The transformation is a process that translates an object from a given domain to another in order to have some important implicit information, which can be used for its recognition. In general, the transformation process is performed by convoluting the given signals with their bases functions. These bases functions are usually orthogonal which make the transform inevitable. This means that the original functions in the original domain can be back extracted without loss of information.

One of the conventional transformations is the Fourier transform which usually transforms the signal from its time domain to the frequency domain. In this work Fourier Transform has Pbeen used to extract implicit information in order to construct the image feature vector by calculating the 2-D Power Spectrum to the image.

2-Color Image Representation

Color images can be modeled as three-band monochrome image data, where each band of data corresponds to a different color. The actual information stored in the digital image data is the brightness information in each spectral band. When the image displayed, the corresponding brightness information is displayed on the screen by picture elements that emit light energy corresponding that particular color. Typical color images are represented as red, green and blue, or RGB images.

Using the 8-bit monochrome standard as a model, the corresponding color image would have 24 bits/pixel – 8 bit for each of the three color bands (red, green, and blue). In figure (1) we see a representation of a typical RGB color image with a single pixel's red, green, and blue values as a color pixels vector [4].

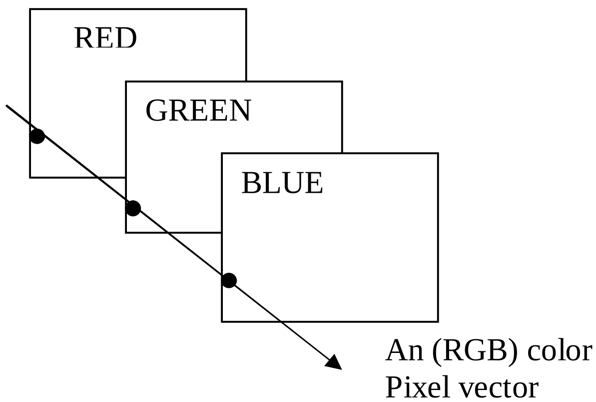


Figure (1): RGB color image with a single pixel's red, green, and blue color pixel vector.

3-Fourier Transform

The French mathematician Jean-Baptiste Fourier showed that any composite signal is a sum of a set of sine waves of different frequencies [5]. Fourier transform was developed by Fourier to explain the distribution of temperature and heat conduction. Since that time the Fourier transform has found numerous uses, including vibration analysis in mechanical engineering, circuit analysis in electrical engineering, and here in computer imaging. This transform allows for the decomposition of an image into a weighted sum of 2-D sinusoidal terms. Assuming an $m \times n$ image, the equation for the 2-D discrete Fourier transform is:

$$F(u,v) = \frac{1}{N}\sum_{r=0}^{N-1}\sum_{c=0}^{N-1} I(r,c)e^{-j2\pi\frac{(ur-vc)}{N}}$$

(1)

where

N: Image dimension (m x n).

I: Image element (pixel) coordinate

r: Row coordinate.

c: Column coordinate.

The base of the natural logarithmic function e is about 2.71828; j, the imaginary coordinate for a complex number, equals $\sqrt{-1}$. The basis functions are sinusoidal in nature, and can be seen by Euler's identity:

$$e^{jx} = \cos x + j\sin x$$

(2)

So we can also write the Fourier transform equation as:

$$F(u,v) = \frac{1}{N}\sum_{r=0}^{N-1}\sum_{r=0}^{N-1} I(r,c)[\cos(\frac{2\pi}{N}(ur+vc)) + j\sin(\frac{2\pi}{N}(ur+vc))] \quad \textbf{(3)}$$

In this case, $F(u,v)$ is also complex, with the real part corresponding to the cosine terms and imaginary parts corresponding to the sine terms. If we represent a spectral component by $F(u,v) = R(u,v) + jI(u,v)$, where $R(u,v)$ is the real part and $I(u,v)$ is the imaginary part, then we can find the magnitude and phase of a complex spectral component as:

$$\left|F(u,v\right| = \sqrt{[R(u,v)]^2 + [I(U,v)]^2}$$

(4)

$$\Phi(u,v) = \tan^{-1}\left[\frac{I(u,v)}{R(u,v)}\right]$$

(5)

The Fourier power spectrum can be found according to the following equation [4]:

$$P(u,v) = \left|F(u,v\right|^2 = [R(u,v)]^2 + I(u,v)]^2$$

(6)

where

$|F(u,v)|$: Magnitude of a complex spectral component, $\Phi(u,v)$:Phase of a complex spectral component, $p(u,v)$: power spectral component.

In this work, the power spectrum is used as a main color image feature.

4-Neural Network

A *neural network* is an interconnected network of simple processing element, e.g. scaling and filtering. The processing elements interact along paths of variable connection strengths which when suitably adapted can collectively produce complex overall desired behavior[6]. Figure (2) shows the neuron model which is described by:

$$y = f\left(\sum_{i=1}^{N} w_i x_i - v_t\right)$$

(7)

Where

x_i : Input signals.

w_i : Synaptic weights.

v_t : Threshold or bias.

$f(.)$: Activation function or squashing function or processing element.

y : Output signal of the neuron.

The use of threshold v_i is to provide a bias to the activation function $f(.)$.

5- Proposed Identification System

The proposed identification system consists of three phases:

1- Generation of reference image templates.

2- Neural Network training phase to the different images' features.

3- Identification phase.

5.1- Generation of reference image templates

This phase deals with problems of how to create reference templates from large number of images. For every identification system, there must be a reference of the images concerned within the system; it differs from one application to another. So these reference templates will consist of information about the images. Figure (3) shows the procedures of generating image templates which can be explained as bellow:

1- Input a color image to the generating algorithm.

2- Analyze the entered color image to three-band monochrome images (red, green, and blue).

3- Subdivide each band into **four** sub-images

4- Transform each sub-image using Fourier transform.

5- Calculate the power spectrum sum to each sub-image using equation (6).

6- Construct the feature vector for the image which is in a result consisted of **twelve** values, each one represents the summation of power spectrum for each sub-image (**four** values for each band).

7- Repeats steps from 1-6 for all images related with the identification system.

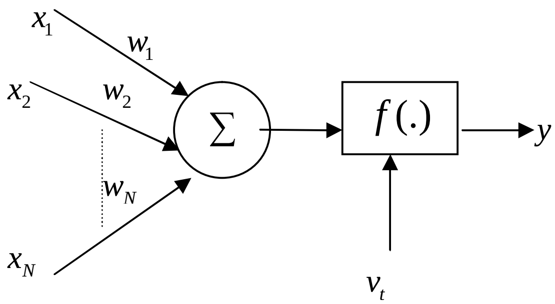
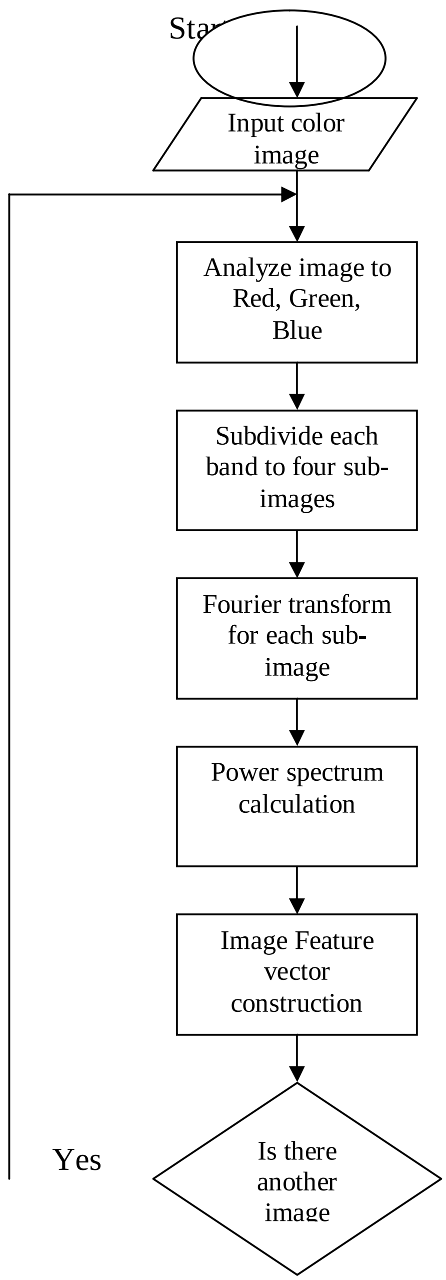


Figure (2): The Artificial Neuron Model



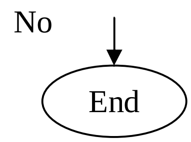


Figure (3): Flowchart of generating reference image templates

-4-

<u>5.1.1 Demonstrative example</u>

In order to show how to create a reference template for each image, we should follow the previous procedures mentioned in section (5.1):

1- Enter a color image as shown in figure (4-a).

2- Analyze the entered image to three-band monochrome images (red, green, and blue) as shown in figure (4-b, c, d).

3- Subdivide each band into four sub-images as shown in figure (4-e, f, g, h, i, j, k, l, m, n, o, p).

4- Transform each sub-image using Fourier transformation.

5- Calculate the power spectrum sum to each sub-image using equation (6).The results of calculation were: (4.3e+5, 2.02e+5, 1.4e+5, 1.8e+5, 4.68.5, 2.14e+5, 2.9e+5, 9.6e+4, 7.8e+5, 1.76e+4, 5.001e+3, 6.2e+3) respectively.

6- The construction of the reference template will be as follow: (e:\matlab\image 1 feature vector, 4.3e+5, 2.02e+5, 1.4e+5, 1.8e+5, 4.68.5, 2.14e+5, 2.9e+5, 9.6e+4, 7.8e+5, 1.76e+4, 5.001e+3, 6.2e+3).

7- Save the image structure which will be used in the training phase of neural network.

Figure (5,6 ,and7) shows the distribution of each color band in its sub-bands. From these figures, we can easily distinguish the difference in the color distribution to the same image. That is why we use the power spectrum as a main feature in the identification system.

5.2 Neural Network training phase

The proposed neural network classifier for image identification is a feed forward multi- layer net with an input layer, two hidden layers and an output layer, trained by the **Back propagation** training algorithm. After many trials it has been shown that the best architecture for best convergence is with 12 nodes in the input layer which correspond to the number of feature vector used in classification, 20 nodes in the first hidden layer, 40 nodes in the second hidden layer, and 100 nodes in the output layer which correspond to the classes of images in the application system working with. The network structure is shown in figure (8).

In the **training phase**, the net starts off with a random set of weights and a training pattern is presented at the input layer. The outputs of the network are then evaluated and compared to the expected output vector for this particular pattern. The resulting errors are fed- back from the output layer and the weights are adjusted. The procedure is repeated for all the patterns in the training set and each time the weights are adjusted. The training set is presented iteratively to the net in an attempt to minimize the mean square error. The aim is that, for a low system error, the output vector representing the classification of a particular input pattern should be equivalent to the target output vector for this pattern indicated by ground data. For example, if a pixel belongs to class 5 (image number five), the output from node 5 in the output layer should be high (+1) and the outputs from the other nodes should be low (-1). The network continues training with the input patterns from the reference set until the mean square error value between the value of the desired and target vectors reaches the error goal $E = 10^{-7}$, at that point training will stop, and the weights and all trained network variables are saved in a file to be used in a testing phase. Figure (9) shows a flowchart for the Neural Network training phase.

It's a recognized fact that the Learning rate and momentum rate was chosen automatically by the program after the operation of initialization at the beginning of training.

5.3 The identification phase

Figure (10) shows a block diagram for the identification process using Neural Network. The perceptron elements of the input layer of such a network receive the direct input feature values from the pattern to be classified. The output layer has one perceptron element per data class and should compute a high output signal at the appropriate element representing the correct class for this pixel. In practice this will only happen if the internal weight values attached to the connections have been properly learnt.

Figure (11) shows the training graph which explains the steps to reach the target vector in 340 iterations.

6- Conclusions

The following points are concluded from the test of the proposed system:-

1- Extracted features which are depended on

the distribution of monochrome colors in each color image are sufficient to construct the final signature color image.

2- The Neural Network gives significantly higher rate of successful classification for a given data set.

3- Neural Network can handle problems of large dimensions efficiently, the using of Neural Network in classifying 100 color images with 1200 values (12 values for each image) gives a result with percentage of successful rate equals to 100%.

4- Because we use different images with different features, so we may expect that the curve of training shown in figure (11) may has an overshoot through its running to reach the error goal $E = 10^{-7}$. So if we used another 100 images, figure (11) would differ and it may has an overshoot or not at another place of figure, depending on the difference among the features of images entered to the neural network structure.

7-References

[1] Tuber, J., “Digital Image Processing”, *Prentice Hall,* 1989.

[2] Riedmiller, M., Braun, H., “A direct adaptive method for faster backpropagation learning: the RPROP algorithm”, Proc. Of the IEEE Int. Conf. On Neural Networks, 1993.

[3] Liew, A.W.C., Law, N.F., and Nguyen, D.T., “Direct Reconstruction method for Wavelet Transform extreme representation”, *IEEE proceeding-Vis. Image Signal Processing*, vol. 144, no. 4, pp. 193-198, April 1997.

[4] Gonzales, R.C., Woods, R.E., “Digital Image Processing”, *Addison-Wesley puplishing Company* 1992.

[5] Behrouz A. Forouzan, “Data Communication and Networking” DeAnza College, 2003.

[6] R. Lippmann, “An Introduction to Computing with Neural Network,” IEEE ASSP Mag., pp 4-22, April 1987.

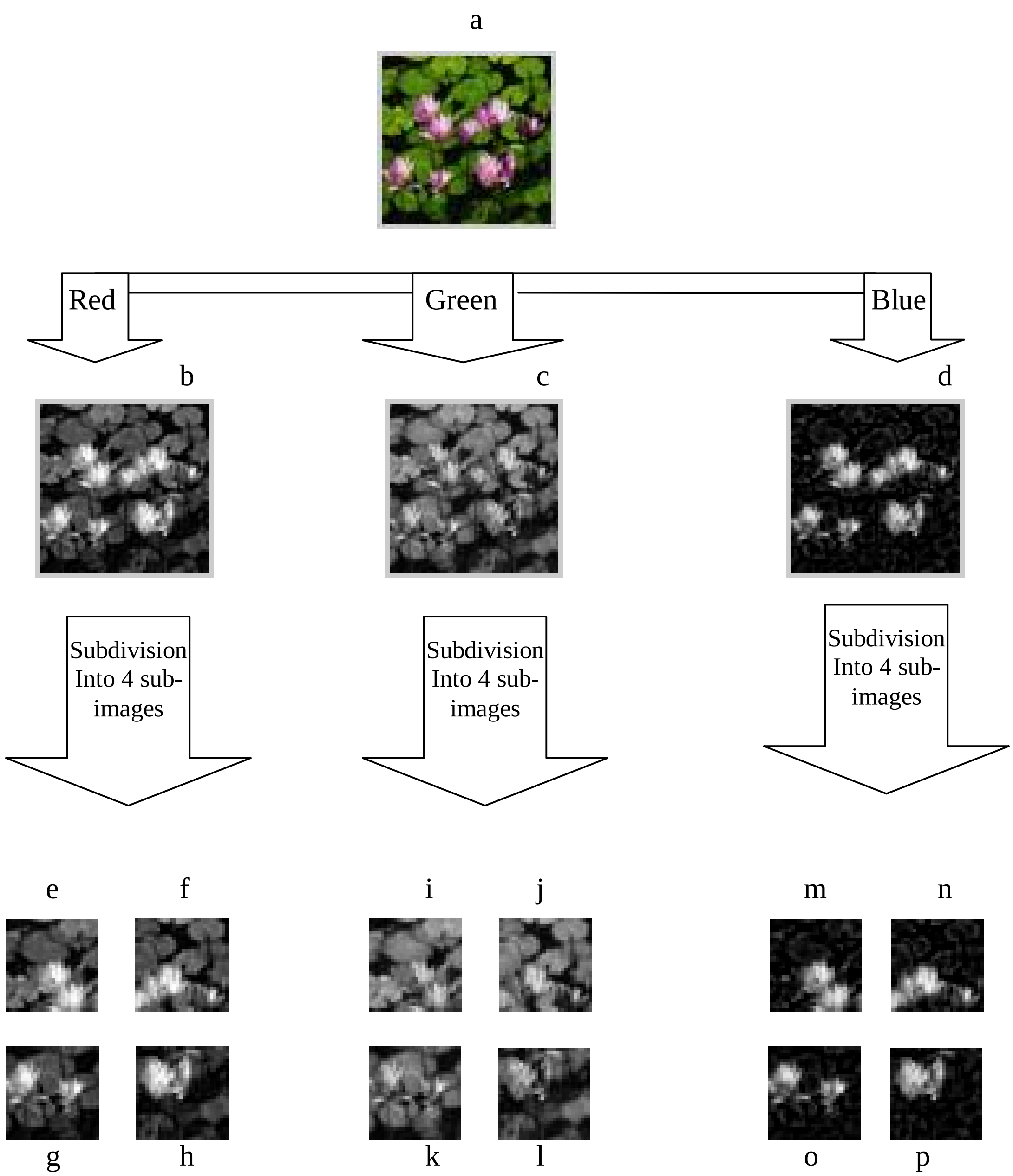


Figure (4): a-Main color image.
b-Red band with its four sub-bands (e, f, g, h).
c- Green band with its four sub-bands (i, j, k, l).
d- Blue band with its four sub-bands (m, n, o, p).

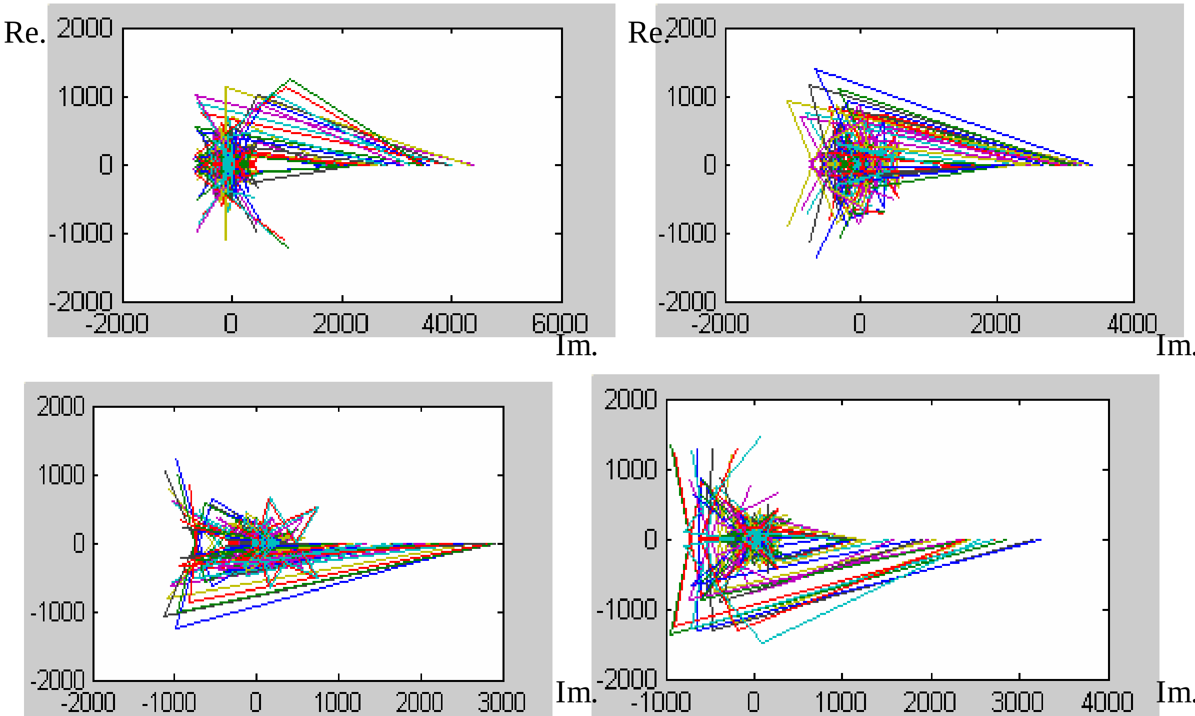


Figure (5): Red color distribution in four red sub-bands

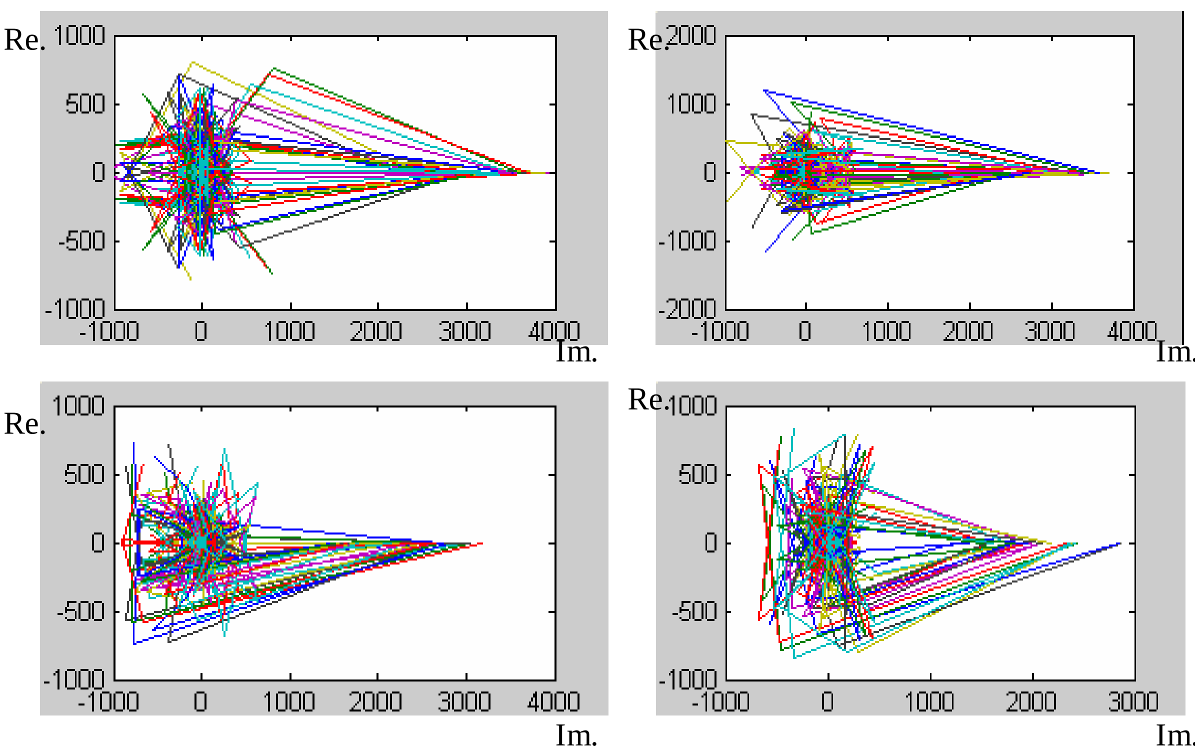


Figure (6): Green color distribution in four green sub-bands

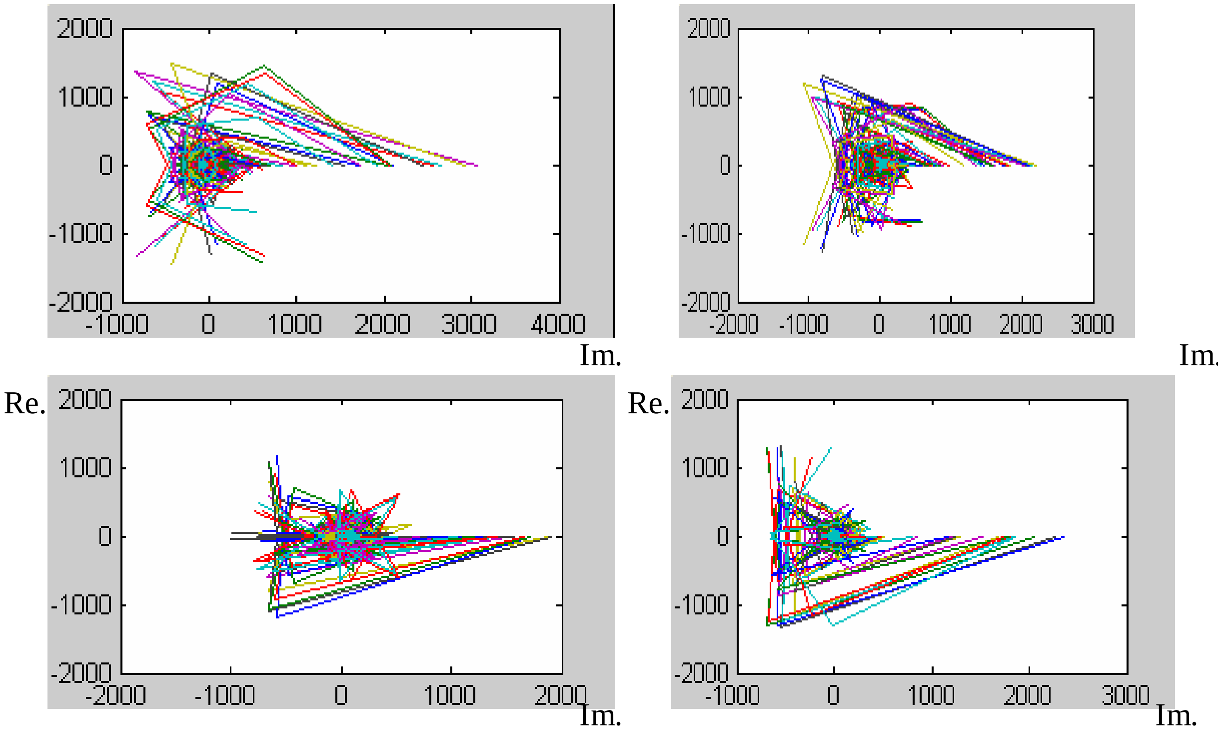


Figure (7): Blue color distribution in four blue sub-bands

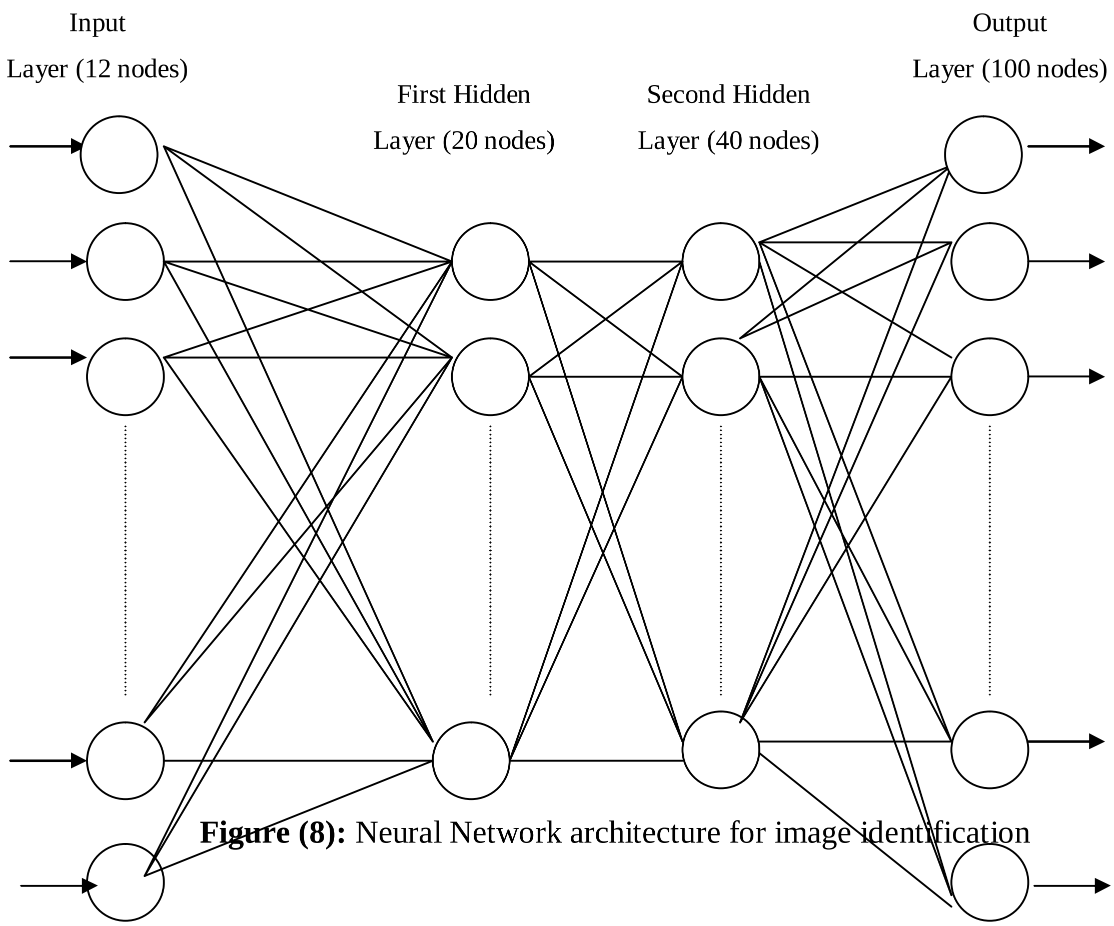


Figure (8): Neural Network architecture for image identification

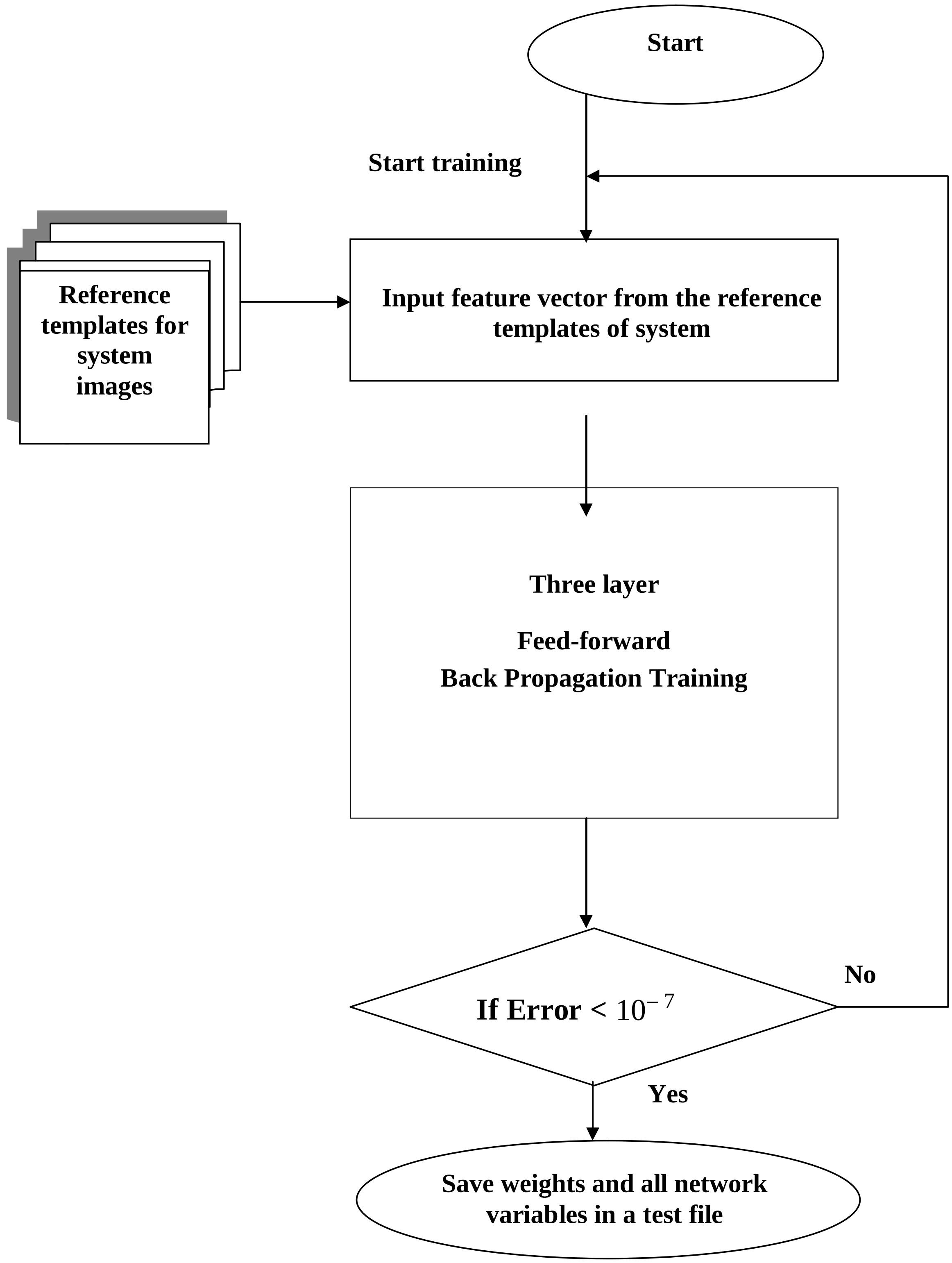


Figure (9): Flowchart for the Neural Network training phase

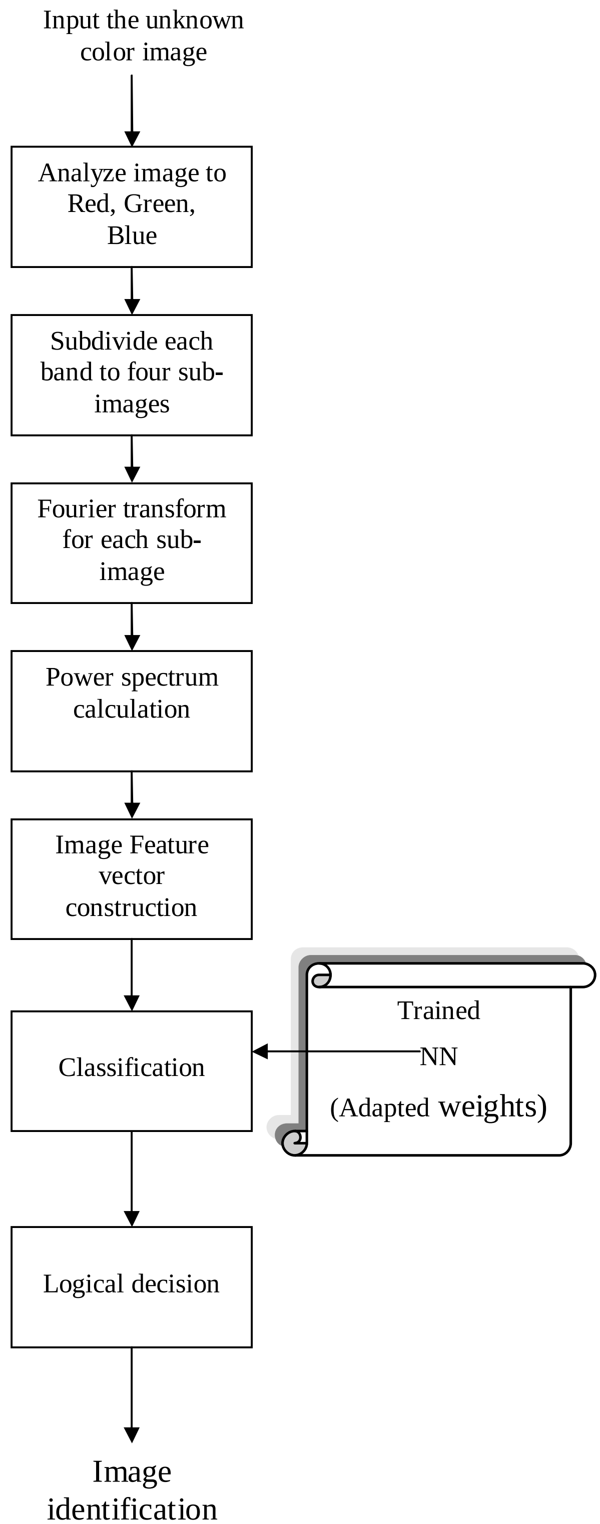


Figure (10): Block diagram for the identification process using Neural Network

Error function

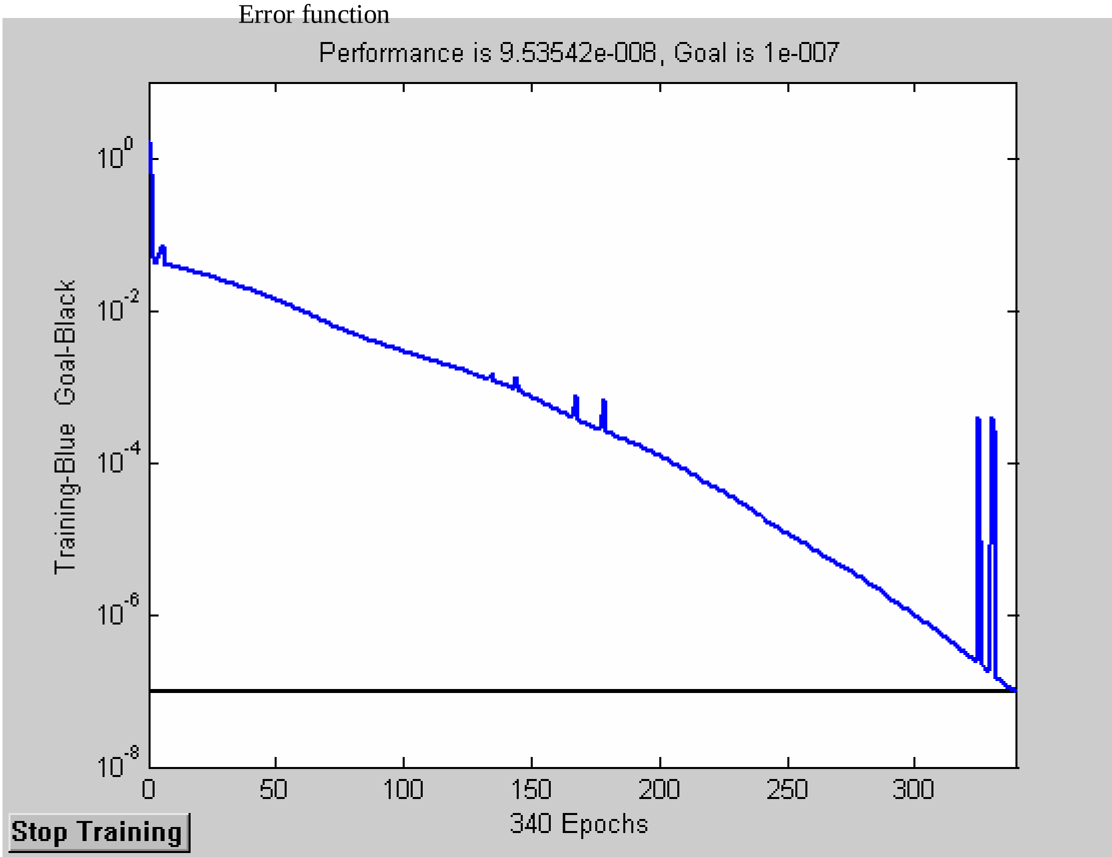


Figure (11): Training Steps to reach the target ($E = 10^{-7}$)