Comparative Analysis of Multi-layer Perceptron and Radial Basis Function for Contents Based Image Retrieval

Monis Ahmed Thakur, Syed Sajjad Hussain, Kamran Raza, Manzoor Hashmani
Faculty of Engineering, Sciences and Technology, IQRA University, Karachi, Pakistan
{(monisthakur89, shrizvi, kraza, mhashmani)@iqra.edu.pk}

Abstract. In the last century, most of the information was text based. But with the rapid growth in the field of computer network and low cost permanent storage media, the shape of information has become more interactive. Despite the significance of multimedia based files major problem is its retrieval as the search engine searches the text associated with the multimedia files, instead the contents of information. Therefore, retrieving a desired result is a major problem. Research trend shows that soft computing tools can play a significant role in intelligent image retrieval therefore in this paper we present a comparative analysis of two of the variants of artificial neural network (ANN) that is multilayer perceptron (MLP) and radial based function (RBF) for intelligent image retrieval with the objective of probability of correct detection.

Index Terms — Contents based image retrieval, multilayer perceptron (MLP), radial based function (RBF).

I. INTRODUCTION

With the rapid enhancement of human computer interaction (HCI) now days we are mostly depending on visual information and its contents. To access and archive multimedia information the new technique for these information is required. While only textual searches on metadata are allowed in most of databases [1]. The actual content may not be retrieved in applications like multimedia libraries, law enforcement agencies, geographical information systems, art collection, medical image management, education, document archives and websites because most of the text is irrelevant with content’s information. This also requires huge human effort to search the relevant content. This is the reason leading to the need of optimized and intelligent retrieval technique. In this paper, comparative analysis of Multilayer perceptron and Radial Based Function (RBF) for training and testing is proposed. Initially for both techniques BP and RBF all the training images are converted into feature vector (FV) containing color and texture feature, subsequently, FV of each image is recorded in feature vector table (FVT) to use for training and adjusting the weights of neural network. After off-line training of same data set of both technique separately, the weights of neural network testing data set are applied on both the techniques separately to the adjusted weight of ANN in order to retrieve most similar images.

Introduction and objectives of the paper are discussed in session 1 and the rest of paper organization is as follows: in session 2 related works is discussed, session 3 and 4 presents an overview of the Back Propagation Algorithm and Radial Based Function Algorithm. In session 5 discusses simulation results and in session 6 comparative analysis are presented and in session 7 conclusion of paper is discussed.

II. RELATED WORK

In 2011, B.Ramamurthy et al. use artificial neural network (ANN) approach to retrieved medical images using their shape feature. Using ANN specify that with the help of these techniques the image retrieval (IR) process of medical images has enhanced and there are a lot of significant performance [2]. In 2010, MOCANU et al. use machine learning tools to bridge the gap between low level feature and high level semantic and on the basis of precision, recall, execution time and compared time different methods [3]. In 2010, Olkiewicz et al. projected an approach in which they examine all the possibilities of the usage of an artificial neural network through which images can be labeled with the help of emotional keywords which are based on visual features and on used emotion filter which are based on the process of similar image retrieval [4].

In 2010, S. Kulkarni et al. projected fuzzy logic for evaluation of resemblance between input images and the images in database. The negative aspect of this research is that on conceptual or semantic level the low level image feature are too often restricted to describe the image [5]. In 2010, Hartvedt et al. projected a technique which uses textual data and visual data in the form of context.
In 2008, H. A. Ahmed et al; projected a unique approach for matching graph which resembled with human the thinking process [11]. In 2007, R. C. Joshi et al; projected an approach which uses genetic algorithm for computing the region based image similarity. The result of this technique was encouraging but there were some drop backs of this research as well, the low level feature like color and texture of the image are generally very less and ineffective for effective and efficient image retrieval of unconstained images. The problems of greedy algorithms are being overcome by the usage of many individual populations and evolutionary operators but still we consider that genetic algorithm can be used for feature matching [12].

In 2004, Y. Wang proposed a structure for automatic metadata generation which is based on fuzzy K-NN classification [13]. In 2004, R. Krishnapuram et al; presented how fuzzy logic set theory can used effectively and efficiently, they described an image retrieval system [14]. In 2002, J. Han projected a representation of a new color histogram which is called FCH. The main focus is based on the color resemblance of each pixel’s color which is related to all the histogram bins through the fuzzy set function [15]. In 2001, H. K. LEE presented ANN approach which is specifically Radial based function (RBF) on the color and wavelet coefficient [16].

III. BACK PROPAGATION ALGORITHM

Two segments in training of model involved in the back propagation algorithm are (forward propagation and backward propagation): a forward propagation part trail by a backward propagation part. In the first part, from the layers of artificial neural network the input signals were sent. In each layer, for the transformation of the incoming signals, a nonlinear activation function is used; those are neurons also known as computational units. Through link weights every neuron of layer one is linked with all the neuron in the following layer. By that method, output of the first part is calculated. In the second part which is back propagation, by comparing the output which we get from forward propagation with the desired output the error is calculated. By using some adaptive optimization rule the weights of every link is updated through this error. The means square error (MSE) is the most general objective function for adaptive optimization rule which can be defined as:

\[
MSE\ (n) = \sum_{j=1}^{N} (d_{j}\ (n) - y_{j}\ (n))^{2} = \sum_{j=1}^{N} e_{j}\ (n)
\]  

(1)

In the above equation \(e_{j}(n)\) is the error between the output of the \(j^{th}\) neuron(\(y_{j}(n)\)), \((d_{j}(n))\) is the corresponding desired response and for number of output nodes there is a \(N\). For the customized back propagation algorithm, the weight linked with the \(j^{th}\) neuron of the \((l-1)^{th}\) layer to the \(j^{th}\) neuron of the \(l^{th}\) layer is updated as follows [17]:

\[
\begin{align*}
\begin{array}{c}
\text{New weights} \\
\text{Old weights} \\
\text{Local gradient}
\end{array}
\end{align*}
\begin{align*}
\begin{array}{c}
\omega_{ji}\ (n+1) = \omega_{ji}\ (n) + \mu\ (n) \delta_{j}\ (n) y_{j}\ (l-1)(n) \\
\delta_{j}\ (n) \subseteq \left\{ \begin{array}{ll}
(\delta_{j}\ (l)(n))
\end{array} \right\} & \text{output layer},
\end{array}
\begin{array}{c}
\delta_{j}\ (l)(n) \subseteq \left\{ \begin{array}{ll}
\delta_{j}\ (l)(n) w_{ji}\ (l-1)(n) \\
\phi_{j}\ (l)(n) y_{j}\ (l)(n)
\end{array} \right\} & \text{hidden layer}.
\end{array}
\end{align*}
\]

(2)

where learning rate is \(\mu\) and \(\delta_{j}\ (n)\) is the local gradient associated with the \(j^{th}\) neuron for the \(l^{th}\) layer and can be calculate according to the following:

\[
\delta_{j}\ (n) = \left\{ \begin{array}{ll}
\left( d_{j}\ (n) - y_{j}\ (l)(n) \right) \phi_{j}\ (l)(n) w_{ji}\ (l-1)(n),
\delta_{j}\ (l+1)(n) w_{ji}\ (l)(n) \\
\phi_{j}\ (l)(n) y_{j}\ (l)(n) \sum_{k=1}^{N} \delta_{k}\ (l+1)(n) w_{jk}\ (l)(n),
\end{array} \right.
\]

(3)

IV. RADIAL BASED FUNCTION ALGORITHM

In the training of radial based function model Gaussian function is used. The layers on radial based function are the input layer in which the size of training model, the hidden layer there are neurons at least to the number of classes of input layer and the output layer which gives the result of the network and it is linear. In the first part, the input signals were sent from the layers of artificial neural network. Then, for the transformation of the incoming signals a nonlinear activation function is used, those neurons are also called computational units. After getting the centered value the random weights and some
bias valves are added. With this method, output is calculated. In the second step, by comparing the output which we get from first part with the desired output the error is calculated. By using some adaptive optimization rule the weights of every link is updated through this error. The means square error (MSE) is the most general objective function for adaptive optimization rule which can be defined as:

$$MSE(n) = \sum_{j=1}^{N} (d_j(n) - y_j(n))^2 = \sum_{j=1}^{N} e_j^2(n)$$

(4)

In the above equation $e_j(n)$ is the error between the output of the $j^{th}$ neuron $(y_j(n))$, $(d_j(n))$ is the corresponding desired response and for number of output nodes there is a $N$. For the recursive radial based algorithm Gaussian function is used where the weight linked with the $i^{th}$ neuron is updated as follows:

$$w_i(n) = w_i(n) + \mu \frac{1}{n} \frac{e(n)}{\text{Error}} y_i(n)$$

(5)

where learning rate is $\mu$, $e$ is error between actual output and the desired output and $y_i(n)$ is the calculated output of the single image and can be calculate according to

$$f(x) = \sum w_j h_j(x)$$ (Output layer)

$$h_j(x) = \exp(-{(x - c_j)^2}/r_j^2)$$ (Hidden layer)

(6)

(7)

Where $c_j$ is center of a region and $r_j$ is width of the receptive field. Figure 1 describes the network diagram of artificial neural network radial based function:

![Figure 1: Structure of Back propagation](image)

V. EXPERIMENTAL RESULTS

In this section the study of the retrieval precision of IR system and on MATLAB is discussed. The content based image retrieval prototype model was implemented. Initially images were read that will have to be trained from image dataset. The image dataset used have 360 images of four classes which are bus, horse, flower and scenery, in which each class has 90 images different from each other. Then, each image is converted into matrix form. Each image matrix has $L \times B \times 3$ dimension where $L$ is the horizontal resolution, $B$ is vertical resolution and 3 in the RGB plane, and there size is 256 x 256 x 3. We split all plain (RGB) and then find mean and standard deviation of all plain and normalize them. Then entropy of each plane is calculated for texture feature and concatenates it with color features. Each FV contain color feature and texture feature. Next the feature vector of each image in an array of matrices which we called FVT was stored. Then, back propagation algorithm, (MLP) and Gaussian algorithm, (used by RBF) was applied on these inputs for training of the images and update the weights.

Then for testing, the image retrieval prototype model 35 testing images were read from the testing database which are different from the training images and all 35 images are of same four classes of training images. Then, again the same procedure is performed to get FV of each image which contains color and texture feature. After that, the back propagation algorithm is applied on these newly calculated input and updated weight which were obtained from MLP and test the content based image retrieval prototype model of MLP which were used. Similarly, the Gaussian algorithm was applied on these newly calculated input and updated weight which were obtained from RBF and test the content based image retrieval prototype model of RBF which we used.

A. Result of MLP:

Training:

To apply the back propagation algorithm on training image, array of matrix at hidden layer with 40 neurons was provided. Subsequently, weights of neuron are adjusted by the back propagation algorithm. Figure 2 shows a graphical representation of the desired output with blue color lines and in Figure 3 is the graphical representation with red lines of training output.
The above Figure 2 shows the actual output which is predicted. The x-axis is the output values which are given to the four classes of the dataset in which 0 represent bus, 1 represent the flower, 2 represent the horse and 3 represent the scenery. While y-axis represents the total number of images in the training dataset. In Figure 3 there is an output graph of trained images values that were obtained after the training of images using MLP technique. In the below Figure 3, the x-axis is the output values which were obtained for the four classes of the dataset. Where, y-axis represents the total number of images that were trained.

The Figure 2 and 3 shows that the actual output is very closed to the desired output of the training image. Then, the sum of error between desired output and the trained output from NN is calculated for 100 epochs which is shown in Figure 4.

The above Figure 4 shows the sum of error of iterations, where x-axis represents the sum of error and y-axis represents the number of iterations. It also shows that the sum of error is decreasing and tends towards almost 0 that is why the desired output and the trained output results are very much similar. The time of each epoch for training is presented in Figure 5 where x-axis represents the time and y-axis represents the number of iterations.

The total time for training is presented in Figure 6 where x-axis represents the time and y-axis represents the number of iterations.
Testing:

After giving the testing images with updated weight values which were obtained after trainings of images to content based image retrieval prototype model, the result is calculated and represented in Figure 7.

The above Figure 7 shows the results of tested images where x-axis shows the class of image retrieved against the input image and y-axis shows the total number of testing image provided for testing. The result is 62% of the proposed model for multi layer perceptron content based image retrieval system. This means, out of 100, 62 images were retrieved. In Figure 8 the probability of correct detections in each class of dataset which was used is represented. The x-axis of the Figure 8 shows the probability of retrieval and y-axis shows the classes of image dataset in which first bar represents bus, second bar represents flower, third bar represents the horse and the fourth bar represents the scenery.  

1.1 B. Result of RBF:

Training:

To apply the RBF technique on training image array of matrix at hidden layer with 4 neurons were provided. Each class of the training will be passed through single neuron. Then some bias values are provided to adjust the weights of neuron by the Gaussian algorithm. Figure 9 shows the desired output value and the training output value where the x-axis is the output values which we give to the four classes of the dataset in which 0 represents bus, 1 represents the flower, 2 represents the horse and 3 represents the scenery and y-axis represents the total number of images in training dataset. The graph with blue lines shows the actual output and the graph with red lines shows the training output values which were obtained after the training of images using RBF technique.
output and the trained output from neural network is calculated for 100 epochs which is illustrated in Figure 10.

![Figure 10: For 100 Epochs RBF Training Sum of Error](image)

The above Figure shows the sum of error of iterations, where x-axis represents the sum of error and y-axis represents the number of iterations. It also showing that the sum of error is reduced to some extant but after some iteration it is decreasing very slightly. The time of each epoch for training is presented in Figure 11 where x-axis represents the time and y-axis represents the number of iterations.

![Figure 11: Time of Each Epoch for RBF Training](image)

The total time for training is presented in Figure 12 where x-axis represents the time and y-axis represents the number of iterations.

![Figure 12: Total Time of RBF Training](image)

**Testing:**

After the training on neurons to test the content based image retrieval prototype model, testing images data set is giving to the proposed model with the updated weights to simulate the model, followed by calculated result shown in Figure 13.

![Figure 13: Testing of RBF, Blue Shows Desired Output and Red Shows Retrieved Output](image)

The results of tested images are shown in the Figure 13 above where x-axis shows the class of image retrieved against the input image and y-axis shows the total number of testing image provided for testing. Blue color in the figure shows the actual image output value and the red color shows the retrieved image value against the in input image. The result is approximately 72% of the proposed model for radial based function content based image retrieval system. This means, out of 100, 72 images were retrieved correctly. In Figure 14 the probability of correct detections in each class of dataset used is displayed. The x-axis of the figure shows the probability of retrieval and y-axis shows the classes of image dataset in which first bar represents bus, second bar represents flower, third bar...
represents the horse and the fourth bar represents the scenery.

![Figure 14: Probability of Retrieval of Each Class using RBF](image)

VI. COMPARATIVE ANALYSIS

Simulation result yields that RBF can increase the correct retrieval by 10% approximately, as 62% is obtained by MLP, and by using RBF we have 72% average probability of correct detection on the same dataset using same features. In the same figure the each iteration time for training in RBF and the total time taken by RBF for the training of images are also displayed. The probability of correct detection of each class of image separately in RBF is also calculated. Moreover, in the multilayer perceptron, there is an over fitting error where radial based function doesn’t have over fitting issue.

VII. CONCLUSION AND FUTURE WORK

In this paper, performance analysis of two of the variants of artificial neural network (ANN) is performed; that is multilayer perceptron (MLP) and radial based function (RBF) based on probability of correct detection of each class and over all probability of correct image retrieved. Simulation results prove that by using RBF the accuracy of the system is enhanced. By using RBF there is 72% average probability of correct detection. The RBF increases the correct retrieval by 10% approximately as 62% is reported by S.S. Hussain [18]. For future work it is suggested to bridge up the gap between low level feature and high level semantic that can be used to retrieve semantically similar images.

Reference: