A NEURO-FUZZY APPROACH TO CONTENT BASED IMAGE RETRIEVAL

N.Srikrishna*, K.Vindhya* and P.Satyanarayana**
* National Institute of Technology, Warangal-506 004, AP, India
** National Remote Sensing Agency, Department of Space, Hyderabad-37
E-mail: srisrikrishna@rediffmail.com, satyanarayana_p@nrsa.gov.in
Phone No.: 040-2-3884014, Fax No.: 040-2-3884013

ABSTRACT
In this paper, an efficient Content Based Image Retrieval method making use of a Fuzzy-Neural hybrid system is presented, in which both fuzzy logic techniques and neural networks are utilized separately to establish two decoupled subsystems which perform their own tasks in serving different functions in the combined system. A Feed Forward Back Propagation Neural Network (FNN) is adopted for Image Classification. The various input features used in the fuzzy inference system and the neural network include Image Intensity, Colour, texture and Image Morphological features.

I. Introduction
Remotely sensed data moved from being analog to digital. Now a time has come for it to become online. This calls for organizing the Remote sensing data in such a way that the user can access relevant portions of it. Content-based Image retrieval (CBIR) is a paradigm, which addresses this issue.

Today a very large archive of Remote sensing data with varying characteristics exists. The user would like to access only those portions, which are of interest to him. Deriving and storing information using traditional methods is infeasible because of the sheer effort involved and very little of it will be utilized. Efficient image storage, indexing and retrieval systems are needed in order to make this vast quantity of data useful. Hence the relevance of CBIR for remotely sensed data arises for proper Data Resource Management and Planning.

Typically, the Content of an image can be characterized by a variety of visual properties known as Features. It is common to compare images by colour, texture, and shape, although these entail different levels of computational complexity. Recognition and Classification methodologies for CBIR can be categorized. Though there is some overlap, three such procedures of Pattern Recognition [1] are:
1. **Statistical Pattern Recognition**, wherein features are algorithmically extracted and template matching techniques are used.

2. **Syntactic (Structural) Pattern Recognition**, where primary components of patterns are extracted and the relations between them are defined with decision trees or graphs [2].

3. **Artificial Neural Network Approach** is a ‘Black Box Model’, which involves testing the applicability of various learning methods and algorithms of ANN for Visual Image Recognition, making use of different ANN Models and Paradigms and their Discriminating abilities to facilitate CBIR.

4. **Fuzzy Logic Approach**. Fuzzy systems are *structured numerical estimators*. They start from highly formalized insights about the structure of categories found in the real world and then articulate fuzzy IF-THEN rules as a kind of expert knowledge.

   An appropriate colour space, a colour Quantization scheme, a histogram representation, a function approximator (for mapping inputs with outputs) and a similarity metric are the main ingredients required for the design of a retrieval system. The present system attempts at utilizing the maximum number of details of an image as provided by its histogram.

   Both spatial and temporal relations can be used to select image sequences that match a query. The distance between a query image and an image from the database can be a weighted sum of the differences between pairs of the corresponding query image features and database image features.

   One can create many distance measures from an image basing on the set of features and a scoring mechanism. A combination of colour and texture distances will prove better than either alone.

   An image is typically characterized by intrinsic attributes of its content ranging from low-level features such as colour and texture, to more complex, relatively higher-level features such as shape.

   Structuring and visualizing digital images based on their content similarities, however, is not as mature as its text-based counterpart. Currently, many CBIR systems have higher-level feature extraction capabilities, but much remains to be done.

   An **Automated Pattern Recognition System** such as CBIR system is an operational system that minimally contains [2].
1. An Input subsystem that accepts sample pattern vectors and
2. A decision maker subsystem that decides the classes to which an input pattern vector belongs.

II. Scope of the Present Work

Of the aforesaid methods of Pattern Recognition for CBIR, the present study makes use of the following:

1. Fuzzy Logic Approach (FLA).
2. Artificial Neural Network Approach (ANN).

Since fuzzy systems do not have much learning capability, it is difficult for a human operator to tune rules and membership functions from training data set. Also, because the internal layers of neural networks are always opaque to the user, the mapping rules in the network are not visible and are difficult to understand; furthermore, the convergence of learning is usually very slow and not guaranteed. Thus, a promising approach for reaping the benefits of both fuzzy systems and neural networks (and solving their respective problems) is to merge or fuse them into an integrated system. This fusion of two different technologies can be realized in three directions, resulting in systems with different characteristics:


The system architecture adopted in the present work is: Fuzzy-Neural Hybrid Schema, in which both fuzzy logic techniques and neural networks are utilized separately to establish two decoupled subsystems which perform their own tasks in serving different functions in the combined system. The architecture of fuzzy-neural hybrid systems is usually application-oriented. Making use of their individual strengths, fuzzy logic and neural network subsystems complement each other efficiently and effectively to achieve a common goal.

The present work essentially consists of two stages, the first one being a Fuzzy system and the later one, an Artificial Neural Network system. The initial stage accounts for the uncertainty arising due to the colours of an image. The final stage comprises of classification of images based on the outputs obtained from the trained neural network.
III. Methodology

In the present approach, the *Pixel level of input data Abstraction* is resorted to, as this has the maximum number of applications for pattern recognition and Image Processing [5].

1. Experimental Configuration

As in Fig. 1, the procedure adopted here for CBIR using ANN consists of an initial Image Feature Extraction and Data Preparation. This process usually requires an apriori uniform dimensioning of the image attributes such as the image resizing, image data type conversion, to facilitate computerized image analysis and other image enhancement processes.

The next phase comprises of a Fuzzy-Inference System (FIS), which assumes fuzziness among the color details of the image as depicted by its color histogram. The color details of the image are scaled in this stage.

This is followed by the preparation of the ANN Sample Input vectors in a way to match the input specifications of the neural network. The Input Subsystem that accepts sample pattern vectors, processes the Input Pattern vectors to enable further analysis.

A decision maker subsystem decides the class to which an input pattern vector belongs and is followed by an Image Retrieval Schema.

2. Data Preparation
i. Selected Image Data Set.

The image data set consists of 427 JPEG colour images with 24-bit depth. The size of the images is 128x128 pixels. 375 images are considered for representing the training image set and the rest 52 images are taken as the Query images.

ii. Tools for feature extraction.

To obtain the colour Histogram features of the images, a program in JAVA programming language that uses - ‘PIXEL GRABBER’ Class, JAVA.AWT.IMAGE’ and ‘GETIMAGE’ method is used.

To obtain the other image features as mentioned in the section II-2 above, the Image Processing Toolbox (Version 3.1-R12.1) of MATLAB software is used.

3. Image Feature Extraction

The following details of the images are considered as inputs to the present system:

i. Image Colour Details such as:

a. **Colour Quantization** scheme involving color indexing of the images.

b. **Correlation** between the Red, Green and Blue band histograms.

ii. Image Morphological features (obtained by **Edge Detection**) such as:

iii. Image Texture details such as:

a. **Second moment** of gray-level histogram.

b. **Image Entropy**.

iv. Histogram of **Discrete Cosine Transform** coefficients.

v. In addition to the above the following derived features of the image are also used:

a) **Standard deviation** of the Quantized color index values.

b) **Contrast** of Texture.

c) **Edge density** of the binary images

d) **Compactness/ Roundness** of the binary images

**Colour Quantization Scheme**

Reducing the number of colours in an image involves quantization. Here, for all the images the Minimum Variance Quantization is obtained. Minimum variance quantization works by associating pixels into groups based on the variance between their pixel values. Also the minimum variance has been set to obtain an indexed image with 16 levels [4] of
colours. Thus, by making use of this scheme, for each image, 48 (3x16) index values have been obtained.

**Correlation**

The Correlation among the 3 color bands (R, G, and B) of each image is calculated. Thus 3 values for each image are obtained.

**Edge Detection Schedule**

Each image is *Edge detected* for all the 16 bin ranges of the intensity histogram. Thus 16 edge detected *Binary Images* are obtained for each image. Now for each binary image, *Euler Number* and *Foreground Image Areas* are calculated.

**Euler Number**

The Euler number is a measure of the topology of an image. It is defined as the total number of objects (connected components) in the image minus the number of holes in those objects. Euler Numbers for both *4-Connected* and *8-Connected Pixel Neighborhoods* have been calculated. Thus for each image 16 values for *4-Connected Pixel Neighborhood* and 16 values for *8-Connected Pixel Neighborhood* are obtained.

**Foreground Image Area**

The area is a measure of the size of the foreground of the image. Roughly speaking, the area is the number of on pixels in the image. For this feature, 16 values are obtained for each image.

**Second Moment** of gray-level histogram (the *Variance*) is a measure of gray - level contrast that can be used to establish descriptors of relative smoothness.

**Entropy**

The Entropy of a source is defined as the average information generated by the source. It is a measure of randomness or unpredictability.

**Discrete Cosine coefficients.**

In the present work a histogram of DCT coefficients is created which is subsequently fed to the neural network. The advantage of cosine transform is that it has excellent energy compaction for highly correlated data.

**Contrast of Texture**

Contrast is the natural measure of the degree of spread of matrix values. It is essentially the moment of inertia of the matrix around its main diagonal.

**Edge density**

This is a measure of coarseness of random texture and represents the density of edge pixels. For the obtained edge map, the edge
density is measured by the average number of edge pixels per unit area.

**Compactness / Roundness**

Compactness of the binary images is defined as \((\text{Perimeter})^2/\text{Area}\). This is a dimensionless quantity and thus is insensitive to uniform scale changes. This is minimal for a disc-shaped region.

4. **Image Classification**

The entire training data set has been classified apriori into classes as shown in Table 1.

5. **Fuzzy Logic Model Adopted**

The second stage of the present system comprises of a Fuzzy-Inference System (FIS), which assumes fuzziness among the color details of an image as depicted by its color histogram. The FIS observes the relatively dominant colors within an image and results in an output in the form of weightages to the colors. This helps in elevating the significance of the dominant colors within the image, which otherwise tend to get diminished during the process of function approximation using the neural network, that is used to classify the data. Now the initial color statistics of the image are scaled accordingly, by multiplying with the weights obtained apriori.

The FIS assumes fuzzy values for the correlation between the intensity histograms and R, G and B color histograms. This in turn also assumes the correlation among the color histograms to be fuzzy. The FIS is made use of, to obtain the weightages for each color within a given image. These weightages are used in conjunction with the color quantized index values to serve as inputs to the Neural Network for image retrieval. This is done by multiplying the originally quantized color index values with their respective weightages. Hence, these scaled (weighted) values of the color indices are supplied to the neural network along with the other details of the image.

**Fuzzy Inference System**

- Number of Inputs = 9
- Number of Outputs = 3
- Number of Fuzzy rules = 24
- FIS architecture: Mamdani
- Defuzzification Method: LOM (Largest of Mean)

6. **Neural Network Adopted**

In the present work an Off-line learning method using the Feed Forward Back Propagation Network (FFN) is adopted for training the network. Back-Propagation is a...
generalization of the Widrow-Hoff learning rule to multiple-layer networks and nonlinear differentiable transfer functions. Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors, or classify input vectors in an appropriate way as defined by the user. Networks with biases, a sigmoid layer, and a linear output layer are capable of approximating any function with a finite number of discontinuities. The term back-propagation refers to the manner in which the gradient is computed for nonlinear multilayer networks.

Properly trained back-propagation networks tend to give reasonable answers when presented with inputs that they are never trained with. Typically, the query is classified based upon its similarity with the output values of the trained neural network. This generalization property makes it possible to train a network on a representative set of input/target pairs and get good results without training the network on all possible input/output pairs [6]. Table 2 gives the details of the FFN used and its parameters.

With the parameters mentioned in Table 2, the FFN is trained till the Training Goal is met. The sections to follow depict the results.

7. Procedure Adopted for Testing the Neural Network.
   i. The various new features of all the images are extracted as mentioned above.
   ii. These features are arranged in a way so as to be compatible with the Input pattern of the ANN model.
   iii. The FNN is trained with the set of input values obtained above, until an acceptable Goal is met with.
   iv. Now the trained FNN is simulated with a series of Query images.
   v. The images, which are efficiently classified, are further investigated to retrieve the Best matching Images in the order of similarity.

IV. Case Studies

The Post-training Regression Analysis gives the following results.

\[ m = 1.0000; b = 9.6195 \times 10^{-5}; r = 1.0000 \]

The value of the Regression Coefficient (r) indicates that the training is complete.

The results of the images retrieved are as shown in Fig.2

Results of Simulation
V. Model Efficiency Evaluation

Precision and Recall are the parameters commonly used to evaluate the efficiency of the information retrieval systems as well as image retrieval systems. Recall indicates the proportion of relevant images in the database, which have been retrieved when answering a query. Precision, by the other hand, is the proportion of the retrieved images that are relevant for the query.

The values for Precision are calculated for each case. As can be seen above, higher the Classification Efficiency, higher is the Precision value and hence the Model is of high Efficiency.

VI. Conclusions

The results show that the fuzzy rules framed, based upon the assumptions regarding the correlation of the color histogram with the intensity histogram and correlation among color histograms, are effective to a fair degree of model efficiency.

Also the method adopted for bringing about a relation between the color histograms and the quantized color index values, which is also assumed to be fuzzy in nature, is a feature worth adopting to consider image color dominance.

It is also observed that the histogram of Discrete Cosine coefficients rather than conventional intensity histogram can be a better measure of an image detail.

At a basic level, details such as entropy, contrast, edge density, compactness are found to represent texture of an image that can be used in the neural network in conjunction with the other attributes of shape, color and intensity.

The values of Precision and model efficiency strengthen the need for correct sampling of the Training images. The number of representative Images in the training sample ought to be comprehensive so as to avoid possibility of any misclassification.
Hence, it is concluded that a CBIR system using all the aforesaid features of images would also be efficient.

As future research it is planned to include other visual features of the images and different combinations thereof, to develop a Robust CBIR system.

VII. Acknowledgements

The authors are highly indebted and grateful to Padmasri Dr. B.L.Deekshatulu, former Director, NRSA, Hyderabad and Shri V.N.Kameswara Rao, Professor, Dept. of Civil Engineering, National Inst. of Technology, Warangal for their continued guidance and motivation throughout this study.

VIII. References:


[2] Carl G.Looney, “Pattern Recognition using Neural Networks-Theory and algorithms for Engineers and Scientists”, Oxford University press, 0.4 pp 8,


IX. Tables and Figures

<table>
<thead>
<tr>
<th>Table 1. Image Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landscape</td>
</tr>
<tr>
<td>Sky</td>
</tr>
<tr>
<td>Trees</td>
</tr>
<tr>
<td>RSItaly</td>
</tr>
<tr>
<td>RSShore</td>
</tr>
<tr>
<td>RSTsangpo</td>
</tr>
<tr>
<td>RSCity</td>
</tr>
<tr>
<td>Rsgreen</td>
</tr>
<tr>
<td>RSRoads</td>
</tr>
</tbody>
</table>
Table 2. Details of the Feed-Forward Neural Network

<table>
<thead>
<tr>
<th>Sl.No.</th>
<th>Feature</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>BackPropagation Algorithm</td>
<td>Levenberg-Marquardt (trainlm)</td>
</tr>
<tr>
<td>2</td>
<td>Output Layer Transfer Function</td>
<td>Logarithmic sigmoid transfer function (logsig)</td>
</tr>
<tr>
<td>3</td>
<td>Number of Epochs</td>
<td>300</td>
</tr>
<tr>
<td>4</td>
<td>Goal</td>
<td>1e-5</td>
</tr>
<tr>
<td>5</td>
<td>Number of Hidden layers</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Number of Neurons in the Hidden Layer</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>Number of Neurons in the Output Layer</td>
<td>9</td>
</tr>
<tr>
<td>8</td>
<td>Number of Neurons in the Input Layer</td>
<td>108</td>
</tr>
</tbody>
</table>

Retrieved images

![Query](image-url)
Fig. 2. Retrieved Images (In decreasing order of Similarity)