Image classification using Hybrid MLP method

Ashish Mohan Yadav¹, B P S Sengar²

¹P.G. Scholar, ²Asst. Professor, ASCT Bhopal, M.P., India

Abstract—Content based image retrieval system is a research area which is growing very fast, here the query image’s visual contents are used to search similar or exact images from a very large scale database of images. An effective system is proposed in our method in which both the semantically and visually relevant features are used for retrieval of the related images. For an efficient retrieval of image, it is necessary for a classification algorithm to be much efficient that it can find and locate the image at a different like in a compressed level. Our system proposes a novel training algorithm for Multi-Layer Perceptron Network to classify compressed images more efficiently. Here a hybrid MLP method is proposed for retrieval several features and shortens the semantic gap between low-level visual feature and high-level perception.

Keywords: CBIR, MLP, SVM, Semantic gap, image classification, compressed image.

I. INTRODUCTION

Over a past few years, the growth in quantity of images used in the digital form is increased tremendously and with this also have raised many such demanding issues in data management and retrieval system of these digital data. But recently by some researchers Content Based Image Retrieval (CBIR) system is seen as an efficient solution to limit some of these issues efficiently. As we know that CBIR is a system for retrieving relevant images from a large digital image database by measuring similarities between the queried image and database images like when the picture of a ship is given, the system should retrieve similar images from the image database given. A feature extraction process which is automatic is very important and crucial also for image retrieval systems [1]. Feature extraction is achieved through extraction of features from images, the feature like color, texture, and shape. Now these extracted features of the images are then compared with the image which is a query image and database images. A similarity algorithm is used to calculate the similarity degree between the queried image and database images.

Now those database images which have matched features as the queried image (with the highest similarity measure) are then ranked and showed to the user. In this CBIR system we have two steps one for Feature Extraction and another for Image Matching (also called feature matching). In the Feature Extraction phase we extracts image features to a discerning extent like color, texture and shape features for every image available in the image database and are called feature vectors. In image matching phase we use the features of queried and database both images and make match between them to check features similar to those in database images. Image segmentation represents a digital image in simple multiple region forms to facilitate analysis. Objects, boundaries in an image are sited and also object and background is segmented using image segmentation [2] [3] [4] [5] [6] [7].

In the process of image segmentation we first segments color which is followed by edge detection process. A good quality of image segmentation required for good feature extraction as over segmentation leads to many details of the object and under segmentation groups many object into single region so level of segmentation is a deciding factor for success or failure of analysis. Image segmentation algorithm takes advantage of discontinuity and similarity of the grey levels of image. Some of the most commonly used segmentation techniques are histogram thresholding, clustering, vector based and fuzzy technique. In this study, a novel image classification algorithm hybrid MLP is proposed.

The rest of the study is organized as follows: Section 2 reviews fuzzy techniques used in CBIR available in literature, section 3 details about the proposed methods, section 4 gives the experimental details and results and section 5 concludes the study.
II. RELATED WORKS

A content based color image retrieval system based on Fast Compression Distance (FCD) concept was proposed by Cerra and Dateu [8]. A hybrid model of Artificial Neural Networks (ANN) using multiple linear regression models to get precise classification accuracy was proposed by Khashei, et al., [9]. The new method did not include total decompression. Histogram features were extracted from wavelet coefficients used for retrieval. Results prove improved retrieving accuracy than that of current algorithms. Khashman and Dimililer [10] trained a neural network to relate radiograph image contents to optimum image compression ratio. When trained, the NN chose ideal Haar wavelet compression ratio of x-ray images on being presented to the network. The model can be used for a 2 class and multi class problems. Computationally less complex, it is capable of use on large data sets. An image indexing/retrieval system suiting JPEG2000 compressed images was presented by Tang, et al., [11].

Packet header information was decoded for image retrieval and performed better than the pixel based Gabor filter and other wavelet based retrieval procedures.

Zargari, et al., [12] suggested a new compressed domain texture based visual information retrieval process. Experiments suggest that the proposed system efficiently compresses radiographs with high image quality. The trained NN correctly recognized optimum compression ratios for 25 training images as expected, yielding 100% training set recognition. Testing the trained NN using 23 images from Test Set 1 not presented to the network earlier, yielded a 95.65% recognition rate, where 22 out of 23 images with known optimum compression ratios were assigned correct ratio. A minimum accuracy level of 89% was accepted in this work. Using this accuracy, the NN yielded 95.65% correct recognition rate among optimum compression ratios. The proposed method’s successful implementation using NN was revealed by high recognition rates and minimal time costs when operating a trained NN. An image retrieval technique for JPEG images in the compressed domain was proposed by Zargari, et al., [13].

The new method is for spatially predicted I-frames in H.264 video coding standard. I-Frame coding uses various prediction modes to spatially predict pixels of a block from upper or left adjacent pixels. A block’s selected prediction mode indicates how the block’s pixels are related to neighbouring parts. A suggestion was that histogram of prediction modes be used as a compressed I-frames texture descriptor. As the method is based on independent I-Frame coded pictures, it can be used for H.264 coded videos analysis or I-frame based coded images image retrieval like advanced image coding. Simulation indicates the superior performance and lower computational load compared to a Gabor filter based efficient realization of pixel domain texture retrieval method. Also, it is robust to variations in image/coding parameters. Le Hoang Thai, Tran Son Hai, Nguyen Thanh Thuy et al. proposed a method in which convey mutually two areas in which are Artificial Neural Network (ANN) and Support Vector Machine (SVM) applying for image classification. Firstly, separate the image into many sub-images based on the features of images. Each sub-image is classified into the responsive class by an ANN. Lastly, SVM has been accumulated all the categorize result of ANN [14].

A CBIR system based on a multi-scale geometric analysis (MGA) tool, called ripplet transform type-I (RT) have been presented by Chowdhury et al.[15]. Laplacian transform of the sharpened grey-scale image is statistically quantized into colour histogram bins in Malik and Baharudin [16].

Fig. 1 Block diagram for CBIR System
III. FEATURE EXTRACTION

Feature extraction phase is a very important step for image classification used in CBIR. With this step what system is do that all the relevant or irrelevant features of all the images available in database are extracted and after that by these features classification of image is performed. Actually, by feature extraction we map an image from image space to feature space. And with feature space we measure the similarity of an input space extracted from input image with the help of kernel function.

In digital image, basically there are many features like colour, shape, text, size and dimension etc. which can be used for feature extraction but for extraction of these features which are more relevant is not an easy task in fact a difficult task. The output given by this step is in the form of vector [17] [18] [19].

Basically image features can be divided into 2 parts.
1. Visual Features
2. Semantic Features

Feature which are extracted by human vision are called visual feature.
They are further divided into-
1. General Features
2. Domain specific

General features are those which can be used for searching like colour, shape, texture and feature which are used for particular domain and have knowledge about them [20].

For example, we are searching for face of a girl which belongs to human category, so here domain is human. Another one is we are searching for elephant which belongs to animal category. These features are domain specific [21].

Some features are semantic these are the features those have some meaningful information about the image and those are very difficult to extract. The categories came into this feature are mean value, RGB value, Histogram value, Standard deviation and entropy but again these features are not so easy to find [22].

So to find this the set of features with the help of input data is called feature extraction [23].

Below diagram shows the process of image classification and feature extraction [24].

IV. PROPOSED METHODOLOGY

A. Classifiers
1) Naïve Bayes Classifier

The Naïve Bayes classifier works on a simple, intuitive concept. It is also seen that Naïve Bayes outperforms many comparatively complex algorithms, making use of variables in data samples, by observing them individually and independently [25, 26].

Naïve Bayes is a model which works on supervised learning classifier. Bayes rule for supervised learning is represented for an unknown target function f: X→Y, or equivalently P(Y|X). where Y is actually a boolean-valued random variable, and X is a vector containing ‘n’ boolean attributes.
That is \( X = (X_1, X_2, \ldots, X_n) \), where \( X_i \) is the Boolean random variable denoting the \( i \)th attribute of \( X \). Applying Bayes rule, \( P(Y = y_i | X) \) can be represented as

\[
P(Y = y_i | X = x_k) = \frac{P(X = x_k | Y = y_i) P(Y = y_i)}{\sum_j P(X = x_k | Y = y_j) P(Y = y_j)}
\]

where \( y_m \) denotes the \( m \)th possible value for \( Y \), \( x_k \) denotes the \( k \)th possible vector value for \( X \), and where denominator summation is over all legal values of the random variable \( Y \). A method to learn \( P(Y | X) \) is using training data to estimate \( P(X | Y) \) and \( P(Y) \) which are then used together with Bayes rule to determine \( P(Y | X = x_k) \) for any new instance \( x_k \).

2) Support vector machine (SVM)

Support vector machine (SVM) is a linear machine constructing a hyperplane as a decision surface [27]. It is based on structural risk minimization method; the error rate here is calculated by sum of error rate of the random variable \( Y \). A method to learn \( P(Y | X) \) is using training data to estimate \( P(X | Y) \) and \( P(Y) \) which are then used together with Bayes rule to determine \( P(Y | X = x_k) \) for any new instance \( x_k \).

The following equation will calculate implementation error function:

\[
\frac{1}{2} w^T w + C \sum_{i=1}^{N} \xi_i + C \sum_{i=1}^{N} \xi_i
\]

where \( C \) is capacity constant, \( w \) vector of coefficients, \( b \) a constant and \( \xi_i \) parameters for handling non separable data (inputs). The index \( i \) labels \( N \) training cases.

3) Multilayer perceptron (MLP)

Multilayer perceptron (MLP) is technique which is very popular for supervised learning network consisting of layer used for input, one or more hidden layers and a layer used for output. Connections between different layers are made to connect every node from a layer to next layer’s neurons. During training, weights are adjusted for each connection. The nodes of output layer are used to produce outputs. Feature vector \( x \) is given as a input at input layer with a output at output layer representing a discriminator between its class and other classes. Training examples, in training, are fed and the predicted outputs computed. The output and target output are compared and measured error is propagated back through network and weights adjusted [29, 30].

The training set of size \( m \) is represented as \( T_m = \{(x_1, y_1), \ldots, (x_m, y_m)\} \) where \( x_i \in \mathbb{R}^a \) are the input vectors of dimension \( a \) and \( y_i \in \mathbb{R}^b \) are output vectors of dimension \( b \) and \( \mathbb{R} \) represents a real numbers set. Let \( f_w \) represent the function with weight \( w \) for the neural network. Supervised learning adjusts weights so that:

\[
f_w(x_i) = y_i; \forall (x_i, y_i) \in T_M
\]
V. PROPOSED APPROACH

Though Levenberg-Marquardt method is considered efficient, computing large Jacobians needs a large memory. The large matrices needed to be inverted for computation, results in bigger computation time. Hence, to reduce computation cost, the following changes are introduced in Levenberg-Marquardt method. Performance index to be optimized in Levenberg-Marquardt algorithm is given as where \( w \) refers to all network weights. \( d_{ic} \) is desired/required value of \( i \)th output and \( ct \)h pattern. \( o_{ic} \) is actual value. The following performance index is introduced in Levenberg-Marquardt method. This leads to major reduction in matrix size, thereby reducing computation cost.

The proposed system is formalised by following steps:

**Algorithm Steps:**
1. Input : Image dataset, Query image
2. Verify image dataset.
3. IF available = false Then
   Error Msg “ Insufficient image in database”
ELSE
Features extract of all images stored in database
\{Color feature: color histogram, color moments………\} now stored it in \( M \)- dimensional feature vector \( DBFi \)
ENDIF
4. Query Image : Extract color feature of query image
5. Stored it in separate variable \( QF \)
\{Used RGB color model to construct a feature vector\}
6. Color co-occurrence matrixes of RGB generate. \( PRGB = [PRGB_{ij}] \)
7. Retrieved pixel information from previous step
8. Sum average value of row by column \( \{i, j \text{ are row and columns respectively}\} \)
9. Apply distance formula

\[
Dist = \sum_{i=1}^{P} w_i \times |f_{DBFi} - f_{QFi}|
\]
\( \{p\text{-dimension, } w_i\text{-weight of } f^{th} \text{ feature}\} \)
\( f_{DBFi} \text{ = } i \text{th feature of DB image} \)
\( f_{QFi} \text{ = } i \text{th feature of query image} \)
\( \text{\\Apply proposed MLP} \)
10. The classification accuracy for retrieval of uncompressed images and various compressed images is evaluated using the proposed MLP-NN.
11. Initialize weights and parameter \( \mu \) (\( \mu=0.01 \) is appropriate).
12. Compute sum of squared errors over inputs \( F(w) \).
13. Obtain increment of weights \( \Delta w \) by

\[
\Delta w = (J^T J + \mu I)^{-1} J^T e
\]
where \( J \) is Jacobian matrix, \( \mu \) is learning rate
14. Recompute sum of squared errors \( F(w) \)
Using \( w + \Delta w \) as the trial \( w \), and judge
IF trial \( F(w) < F(w) \) in step 2 THEN
\( w = w + \Delta w \)
\( \mu = \mu \cdot \beta \) (\( \beta = 0.1 \))
Go back to step 2
ELSE
go back to step 4
ENDIF
15. Sort similarity distances in ascending order and retrieve first \( N \) images

VI. EXPERIMENTAL RESULT

This section presents the simulation results implemented by the proposed method SVMID3 and the simulation is done using the best known simulator MATLAB 2009a and some reputed image dataset.

A. Image Data Set

The coral image data set is very famous image data set for research purpose of image classification and retrieval, in this experimental data set they used 180 images which contain total 18 classes and each class have fixed 10 images from out of coral dataset of thousands images.

B. Performance Evaluation of CBIR

The level of the kind of accuracy achieved by a system is self-sufficient to establish its performance. And if the outcome shown by the system is satisfactory and promising, it can be used as a standard in future research works. In CBIR, precision-recall is the most widely used measurement method to evaluate the retrieval accuracy[31].

**Precision** is defined as the ratio of the number of retrieved relevant images to the total number of retrieved images [32].

\[
\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}}
\] [32]
Recall is defined as the ratio of the number of relevant images retrieved to the total number of relevant images in the whole database [33].

\[
\text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in database}} \quad [33]
\]

Error Rate is defined as the ratio of the number of non-relevant retrieved images to the total number of images in the whole database. [34]

\[
\text{Error rate} = \frac{\text{Number of non-relevant images retrieved}}{\text{Total number of images retrieved}} \quad [34]
\]

C. Result analysis of classified images

Here in the figure 5.2.1 and figure 5.2.2 compare result with class 12 dinosaur images, class 12 is defined for dinosaur images in our database, where the accuracy measurement of SVM and proposed method are 70% and 100% respectively.

Similarly I have tested both method in eighteen different classes, and next section summarized the all classes result. The result analysis of classified images based on compared two methods with retrieved images. The table 5.2.1 and table 5.2.2 shows that the detailed analysis of the SVM and SVMID3 methods respectively.

<table>
<thead>
<tr>
<th>Image Class</th>
<th>SVM</th>
<th>Proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>7</td>
<td>80</td>
<td>100</td>
</tr>
</tbody>
</table>
Table II Overall Accuracy

<table>
<thead>
<tr>
<th>Overall Accuracy</th>
<th>SVM</th>
<th>Proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>78.34</td>
<td>91.12</td>
</tr>
</tbody>
</table>

Here table 5.2.3 shows that the individual comparative class wise accuracy of both method, where class taken randomly for retrieving the image classes. After that the figure 5.2.3 shows graph analysis of the individual classes accuracy of the both method, where proposed method shows the better result up to maximum instant when it compared with SVM method.

D. Accuracy chart for all these classes

Here table 5.2.4 shows the overall accuracy of the both methods where the result is examined in the bases of 18 classes of the images. And finally produced the output where SVM method has overall 71.12 percentage and SVMID3 has overall 90.56 percentage accuracies. After that in the figure 5.2.4 its shows that the comparative graph analysis of the both method with the help of 3D bar graph.

Table-II Overall Accuracy

<table>
<thead>
<tr>
<th>Overall Accuracy</th>
<th>SVM</th>
<th>Proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>78.34</td>
<td>91.12</td>
</tr>
</tbody>
</table>

VII. CONCLUSION

Here our focus with this proposed work is to give a comparative study of neural network and support vector machine. For the classification of the images we use support vector machines as a classifier and after applying this classification process for the features extracted from the images in the step of feature extraction. Its main function is to find a hyper plane with maximum margin in a high dimensional feature space. Here all the conditions such as nonlinear support vector machines, optimal separating hyper planes, and linearly non-separable case are discussed upon which support vector machines work and a kernel based learning method which is used for the mapping purpose is also discussed. So with our proposed method we get much better performance than the other traditional methods and get optimal results.

References


