International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 09 | Sep -2016 www.irjet.net p-ISSN: 2395-0072

# Content Based Image Retrieval and Classification Using Image Features and Deep Neural Network

**Sandeep Kaur[1], Gaganpreet Kaur[2]**

[1]M.Tech Student, Dept. of CSE, SGGSW University, Fatehgarh Sahib, Punjab, India

[2]Assitant Professor, Dept. of CSE, SGGSW University, Fatehgarh Sahib, Punjab, India

**Abstract-** *With the development of the Internet, and the availability of image capturing devices such as image scanners, the size of digital image collection is rising quickly. A well-organized image searching, browsing and retrieval tools are essential for users from different domains, including remote sensing, crime prevention, publishing, medicine, architecture, etc. Therefore, several general purpose image retrieval systems have been exploited. In CBIR, images are indexed by their visual content. Content based image retrieval consists of two parts: feature extraction/indexing and retrieval part. The techniques which are used to extract features of an image are called feature extraction techniques. The choice of features plays a significant role in image retrieval. Some of the features used are color, texture and shape. Combination of these features provides better performance than single feature. Here we have proposed a combination of statistical features and deep neural network which has improved the performance of the system which is shown in the results.*

***Keywords-*** CBIR, feature extraction, computer vision, retrieval, classification

## 1. INTRODCUTION

Image retrieval on the basis of image features, textures and color has become one of the most researched areas in the field of computer vision. The major utilization of all the techniques used to retrieve images based on content relies on how well the features are being extracted. With advances in feature extraction methods, the field is getting more and more sophisticated. Given an image database, we are interested in finding relevant images for a given query image. The "relevant" images are visually and semantically similar to the query image. Traditionally, retrieval was done by utilizing meta-data associated with the image. This meta-data includes image name, textual description associated with the image, time at which the picture was taken etc. But in a large image database, manually assigning labels/tags for each image is impossible. Further the meta-data associated with the image may be noisy, leading to incorrect retrieval. This motivates the use of content based retrieval systems.

Content based retrieval systems utilize "image content" to find similarity between the images. The "image content" refers to color, textures objects, shapes etc., present in the image. For the retrieval system to be scalable, the main challenge is to find a representation for the image which

1. Is robust to geometric transformations
2. Consumes less memory per image
3. Requires less computation for computing similarity between two images

The first property is required for robust image retrieval. The second property is necessary for storing the image representation of the entire database in RAM of the computer, to avoid costly disc accesses. The third property is necessary for online retrieval. Whenever a new image is added to the database we must able to quickly compute the distance of the new image with other images in the database. Recently proposed method called, Vector of Locally Aggregated Descriptors(VLAD) satisfies these properties and has been shown to be suitable for large image databases. In this thesis, we improve upon the existing Color statistical feature extraction approach by using DEEP NEURAL NETWORK color algorithm before aggregation and using feature fusion for fusing complementary information in images. Further, we propose an approach which uses features obtained from a pre-trained convolutional neural network for constructing compact image representation. In order to make use of the vast amount of data on the World Wide Web, efficient and effective techniques to retrieve the information based on its content are required to be developed. Among the various media types, images are one of the most important sources of information. It is not only the most

widely used media type besides text, but also one of the most widely used bases for representing and retrieving videos and other multimedia information. Despite the extensive research effort, the retrieval techniques used in CBIR systems lag behind the corresponding techniques in today's best techniques available.
At the early stage of CBIR, research primarily focused on exploring various feature representations, hoping to find a "best" representation for each feature. It is interesting to look back toward the beginning and see which of the original ideas have blossomed, which haven't, and which were made obsolete by the changing landscape of computing. Many programs and tools have been developed to formulate and execute queries based on the visual or audio content and to help browsing large multimedia repositories. Answers to many questions with respect to speed, semantic descriptors or objective image interpretations are still unanswered. The work presented here gives a brief review of the research carried out in this field and also the proposed techniques that will help improve the system.

## 2. RELATED WORK

**N. Varish and A. K. Pal [6] 2015** proposed that the content based image retrieval is the process of searching images from an image database based on the visual contents of the input query image. In this paper, the authors have presented a CBIR technique using color based feature. Since a color image consist of three basic color components i.e. red, green, blue. In this initially the three probability histograms are constructed for each color component and then divide them into several number of significant bins and from each bin several statistical values like standard deviation, skewness, kurtosis. The computed statistical values are used as extracted features of the image data. The processing cost of the presented CBIR technique is significantly low. The technique has been tested on standard image databases and satisfactory results have been achieved.

**S. Arivazhagan, L. Ganesan [17] 2013** proposed that wavelet is used to transform a picture from spatial into the frequency domain. The transform displays a superposition function of family of the basis of functions, these are called wavelets. The Wavelet transforms explore the information from signal at distinct scales crossing the signal via high pass filters and low pass. Wavelets provide a good energy compaction and multi solution capability. Wavelets are powerful with the respect to image color shifts. DWT compressed a signal into multiple Wavelet Functions and Basis Functions. The transform of a 2-dimensional picture is also a multi solution method that applies sub-sampling and recursive filtering.

**Cerra, Daniele, and M. Datcu [4] 2012** proposed a compression-based measure, the (FCD) Fast Compressing of Distance that associate the correctness of NCD by the decreased complexity of PRDC. In an initial offline step, the pictures are quantized in a specific color space or changed into the strings, after the changed to reserve textural information in process: representative, subsequently dictionaries are expressed from an each object. There are some similarities in between different images which are calculated by the comparing with an each couple of dictionaries.

**M.Singha, K. Hemachandran [16] 2011** described a process of getting specific images from a huge collection of DB on the basis of color, texture features. This method defines the Wavelet Based Color Histogram Image Retrieval technique. The color and texture features are expressed through color histogram and wavelet transformation and combination of some powerful features for translation and scaling of objects in a picture. It has demonstrated a fast retrieval method on a WANG picture database including 1000 general color images. Moreover, some other computational tips are effectively decreased with the usage of wavelet transformation.

**Murala, Subrahmanyam, R. P. Maheshwari, and R. Balasubramanian[17] 2012** proposed that texture analysis has become more useful and broadly used in pattern recognition and computer vision software applications causing its potential in exploring the prominent characteristics. The performance enhancement can be attained by calculating the thresholds while using genetic algorithm for CBIR application. It is a branch of texture optimization and analysis that has been attracted wide attention from the industries has used the discrete transform for texture classification.

**Singha, M. K. Hemachandran, and A. Paul[12] 2012** proposed that color feature is most important attribute of picture retrieval, because its fast and an easy computation. Color is also responsive feature that runs as a vital role in picture matching. The expressing of image feature from digital picture, that depend on the color and it display in digital images. The color histogram is used as color feature explanation for picture retrieval. The genuine idea, to get the CH for returns comes from Ballard and Swain, who realized the

energy to identify the color or an object. The CH has merits involving less complexity of the computation and an estimate invariance of rotation, translation, and scale, but it did not get the information of pixels to other images.

**Amanatiadis, A., et al. [3] 2011** proposed an evaluation of MPEG-7 size descriptors the effectiveness of Fourier descriptors and Zernike moments that was confirmed with experimental conclusion. In the scale space curvature descriptor performs the evaluated size and shape descriptors when it has compared with the Core Experiment of MPEG-7. In these spectral transforms and descriptors based moments, like FD and Zernike moments are proved to the good choices for normal shape applications. During the mentioned descriptors which are the most crucial shape descriptors, that haven't evaluated opposite to the each other. The comparison results and retrieval performance are discussed. Moreover, their complexity in terms of amount which is required retrieval computational cost and coefficient is present.

**J. Yue, Z. Li, L. Liu, and Z. Fu [14] 2011** proposed the common feature of low-level including reflecting texture, color, salient and shape points in picture. Due to the effectiveness, robustness, low storage merit advantages and implementation simplicity. Color contains the most effective and useful feature or all CBIR systems take colors. CIE or HSV Lab and LUV gaps are used to display color combination of RGB space. Normally, the division of color was displayed by color formed and histograms in the images' of feature vectors.

**R. Datta, J. Li, and J. Z. Wang [18] 2008** proposed the featured shape with images, reliability segmentation was critical that the shape approximates are hugely meaningless. Even then the normal problem of these segmentation in context of human being perception is large from being resolved, some interesting newest directions, most crucial segmentation on the Normalized Cuts criteria. It based on the spectral clustering, that has been expressed to texture picture segmentation with using cues of texture and contour differences.

## 3. PROPOSED METHODOLOGY

### 3.1 DEEP NEURAL NETWORK

In behavioural sciences, also as in most biological sciences, applied mathematics analyses victimization ancient algorithms don't invariably result in a satisfactory answer, notably in classification analysis. Current classification strategies suppose constant quantity or non-parametric variable analyses: discriminant analysis, cluster analyses, etc.

These strategies square measure typically rather inefficient once the information square measure nonlinearly distributed, even once variable transformation. Therefore, here we tend to propose a classification technique, supported the principles of deep neural networks. Throughout the eighties, the utilization of the NN developed explosively within the areas of word recognition

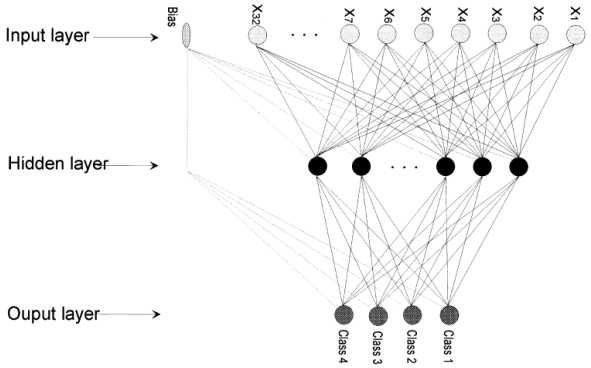



**Fig 1.** Structure of a neural network used in the experiments

Waibel et al., 1989; Lefebvre et al., 1990), and character or image recognition (Belliustin et al., 1991; physicist et al., 1990; Tirakis et al., 1990; Omatu et al., 1990; Fukushima and Wake, 1990; Iwata et al., 1990). Yet, solely some applications were associated with the biological sciences. In hydrobiology models, compared to multiple correlations, NN clearly improved prediction performance

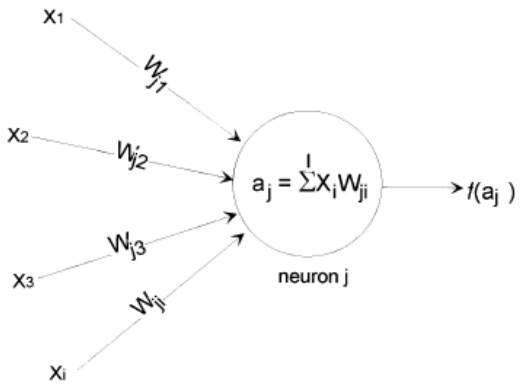



**Fig 2.** Detail of one neuron

For classification functions, NN are used for the analysis of macromolecule structure, the classification of seaweeds (Smits et al., 1992), and therefore the recognition of impulsive noises in marine mammals (Nicolas et al., 1989). During this paper, NN square measure accustomed discriminate the vocalizations of 4 male Damadama, deer (Mammalia, Cer6idae), throughout the rutting amount.

**Algorithm 1**

Step 1: Image is obtained from training dataset.

Step 2: Calculate the R, G B value from image

Step 3: Calculate the histogram and probability histogram value from image

Step 4: Divide the probability histogram into 10 bins and for each bin calculate the standard deviation, skewness and kurtosis. Also calculate the color mean features as:

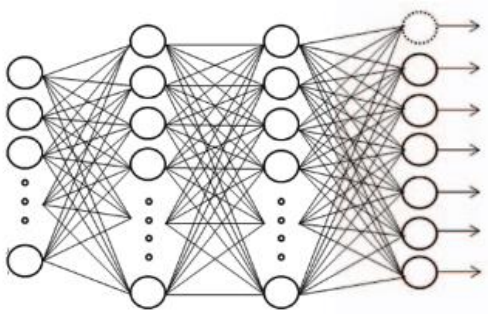
1. Ravg $= \frac{\sum_{i=1}^{h} \sum_{j=1}^{w} R(i,j)}{h*w}$
2. Gavg $= \frac{\sum_{i=1}^{h} \sum_{j=1}^{w} G(i,j)}{h*w}$
3. Bavg $= \frac{\sum_{i=1}^{h} \sum_{j=1}^{w} B(i,j)}{h*w}$

Step 5: Features of image is extracted and save it in training file.

Step 6: If this image is the last image, then preprocess the training file and train the classifier, otherwise go to step 1.

Step 7: Use Deep Neural Network for classification of images and training the neurons:

1. The architecture for Deep neural Network can be shown as:



Deep neural network architecture

2. The activation function of hidden units is the logistic sigmoid

$$g(z) = \frac{1}{1+exp(-z)}.$$

3. In the function, z=Wx, where x is the input vector and W is the weight parameter.
4. The output is given by:

$$h_W(x_i) = \begin{bmatrix} p(y_i = 1|\, x_i; W) \\ p(y_i = 2|\, x_i; W) \\ \vdots \\ p(y_i = m|\, x_i; W) \end{bmatrix} = \frac{1}{\sum_{j=1}^{m} e^{W_j^T x_i}} \begin{bmatrix} e^{W_1^T x_i} \\ e^{W_2^T x_i} \\ \vdots \\ e^{W_m^T x_i} \end{bmatrix}$$

**Algorithm 2**

Step 1: Now, image is obtained from testing data set.

Step 2: Extract the feature of image.

Step 3: Apply the features as input to the trained classifier, using the equation

$$h_W(x_i) = \begin{bmatrix} p(y_i = 1|\, x_i; W) \\ p(y_i = 2|\, x_i; W) \\ \vdots \\ p(y_i = m|\, x_i; W) \end{bmatrix} = \frac{1}{\sum_{j=1}^{m} e^{W_j^T x_i}} \begin{bmatrix} e^{W_1^T x_i} \\ e^{W_2^T x_i} \\ \vdots \\ e^{W_m^T x_i} \end{bmatrix}$$

Step 4: Find the class to which the image belongs using the output of the classifier

Step 5: Pick random images from the selected category.

## 4. RESULTS


**Fig 3**: Retrieved images of peoples using DEEP NEURAL NETWORK features based on Query Image.The primary image retrieved is same because the question image. This shows the effectiveness of the method which can then be quantified in terms of Precision and Recall and the results can be mentioned in the form of graphs and a comparison result or graph can be displayed.


**Fig 4**: Retrieved Images of bus of the query image. The primary image retrieved is same because the question image. This shows the effectiveness of the method which can then be quantified in terms of Precision and Recall and the results can be mentioned in the form of graphs and a comparison result or graph can be displayed.


**Fig 5:** Results for retrieved image of buildings using DEEP NEURAL NETWORK feature. The primary image retrieved is same because the question image. This shows the effectiveness of the method which can then be quantified in terms of Precision and Recall and the results can be mentioned in the form of graphs and a comparison result or graph can be displayed.

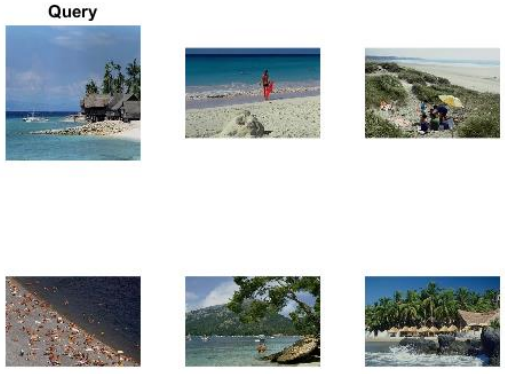



**Fig 6**- Results for beach image query using DEEP NEURAL NETWORK Feature. The primary image retrieved is same because the question image. This shows the effectiveness of the method which can then be quantified in terms of Precision and Recall and the results can be mentioned in the form of graphs and a comparison result or graph can be displayed.


**Fig. 7-** Confusion Matrix for the classification

The diagonal cells show the number of cases that were correctly classified, and the off-diagnal cells show the misclassified cases. The blue cell in the bottom right shows the total percent of correctly classified cases(in green) and the total percent of misqualified cases(in red). The results show very good recognition.

Table 1: Precision and Recall for Proposed work for five categories

| | Base Method[7] | Proposed Method |
|---|---|---|
| **People** | 67 | 91 |
| **Beaches** | 51 | 71 |
| **Buildings** | 54 | 70 |
| **Buses** | 79 | 92 |
| **Dinosaurs** | 99 | 100 |

**Table 2:**Comparison between Base and Proposed method

| | Precision | Recall |
|---|---|---|
| **People** | 91 | 9.1 |
| **Beaches** | 71 | 7.1 |
| **Buildings** | 70 | 70 |
| **Buses** | 92 | 9.2 |
| **Dinosaurs** | 100 | 10 |

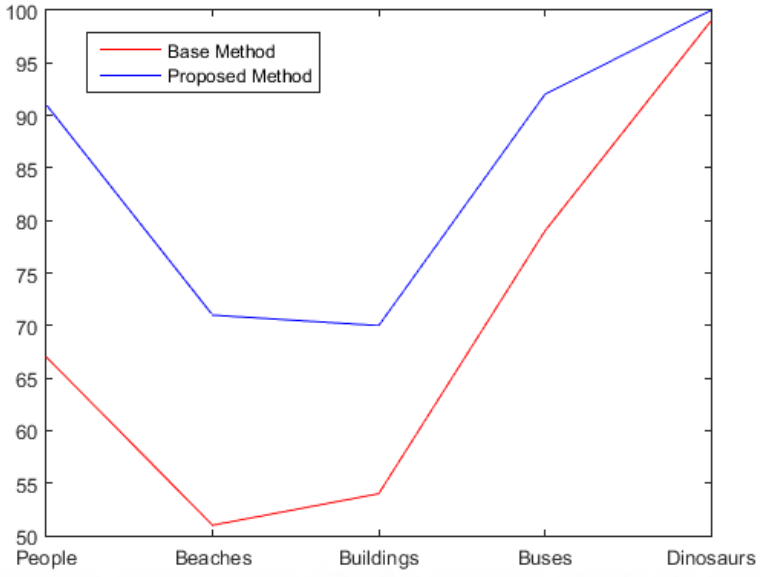



**Fig. 8-** Graph of results

Graph shows the results for the comparison of base and proposed method by taking the values from table 2.

## 5. CONCLUSION

Content based image retrieval is a challenging method of capturing relevant images from a large storage space. Although this area has been explored for decades, no technique has achieved the accuracy of human visual perception in distinguishing images. Whatever the size and content of the image database is, a human being can easily recognize images of same category. Overall the performance of content based image retrieval depends on features, feature extraction techniques, similarity measures and the size of database. Several feature extraction techniques have been developed to the task of image retrieval. Further, it is proved that by combining different features, the performance can be increased. We have performed performance evaluation of the proposed method using and Deep Neural Network classifier with COREL database for determining the classification rate .It is observed that the proposed is giving desired results. Further, it is observed that in some cases there will be irrelevant images with the result of query image in some cases these irrelevant images are totally different from query image on basis of color and shape. Still, this is not the required image and hence there is a scope of improvement in the existing algorithm future work consists of using some other color space or improved texture extraction technique.

## REFERENCES

[1] Alaa, E., Hakim, A., Csift, A. F.,*"A sift descriptor with color invariant characteristics",* In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 1978–1983, 2006.

[2] Krizhevsky, A., Sutskever, I., and Hinton, G., *"Imagenet classification with deep convolutional neural networks",* In Advances in Neural Information Processing Systems 25, pp.1106–1114, 2012.

[3] Amanatiadis, A., *"Evaluation of shape descriptors for shape-based image retrieval",* Image Processing, IET 5.5, vol. 3, issue 3, pp. 493-499, 2011.

[4] Cerra., Daniele., Datcu, M., *"A fast compression-based similarity measure with applications to content-based image retrieval" ,* Journal of Visual Communication and Image Representation vol. 23 no.2, pp. 293-302, 2012.

[5] Lowe, D. G, *"Distinctive image features from scale-invariant keypoints",* International journal of computer vision, vol. 60, no. 2, pp. 91–110, 2004.

[6] Varish, N., Pal, A. K., *"Content based image retrieval using statistical features of color histograms"*, in 3rd international conference on signal processing, communication and networking, pp. 1-6, 2015.

[7] Khutwad, H. K., Prof. Vaidya, R. J. D., *"Content based image retrieval",* in International Journal of Image Processing and Vision Sciences, Vol. 2, pp. 19-24, 2013.

[8] Farsi, H., Mohamadzadeh, S., *"Colour and texture feature-based image retrieval by using Hadamard matrix in discrete wavelet transform",* in IET Image Processing, Vol. 7, Issue 3, pp. 212-218, 2013.

[10] Egou, H. E., Douze, M., Schmid, C., *"Hamming embedding and weak geometric consistency for large scale image search",* In Computer Vision, ECCV, pp. 304–317, 2008.

[11] Egou, H. E., Douze, M., Schmid, C., Perez, P., *"Aggregating local descriptors into a compact image representation",* In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3304–3311, IEEE, 2010.

[12] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., Imagenet, L. F. F., "*A large-scale hierarchical image database",* In IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255, IEEE, 2009.

[13] Delhumeau, J., Gosselin, P. H., Egou, H. E., Perez, P., *"Revisiting the valid image representation",* In Proceedings of the 21st ACM international conference on Multimedia, pp. 653–656, ACM, 2013.

[14] Yue, J., Li, Z., Liu, L., Fu, Z., *"Content-based image retrieval using color and texture fused features",* In Science Direct Journal of Mathematical and Computer Modeling, pp. 1121-1127, 2011 .

[15] Liu., Hai, G., Yang, J. Y., "*Content-based image retrieval using color difference histogram",* Pattern Recognition, vol. 46, no.1, pp. 188-198, 2013.

[16] Singha, M., Hemachandran, K., Paul, A., *"Content-based image retrieval using the combination of the fast wavelet transformation and the color histogram",* In IET Image Processing , Volume 6, pp. 1221-1226, 2011.

[17] Murala.., Subrahmanyam.., Maheshwari, R. P., Balasubramanian, R., *"Local tetra patterns: a new feature descriptor for content-based image retrieval",* Image Processing, IEEE Transactions on vol. 21, no. 5,pp. 2874-2886, 2012.

[18] Datta, R., Li, J., Wang, J. Z., *"Content-Based Image Retrieval - Approaches and Trends of the New Age",* In ACM Computing Surveys, Vol. 40, pp.253-262, 2008.