iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** I **Month of publication:** January 2018

DOI: http://doi.org/10.22214/ijraset.2018.1159

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor :6.887
Volume 6 Issue I, January 2018- Available at www.ijraset.com

The Discrete Image Search uses Feature Extraction and Classification for Content Based Image Retrieval

Harpreet Kaur [1], Er.Sukhpreet Kaur [2]
[1, 2] *Department of Computer Science, Engineering, SUSCET, Tangori*

Abstract*: Content Based image retrieval develops representations of unique characteristics that are automatically extracted from the images themselves. The recent CBIR systems allow for query by the illustration, a method wherein an image is chosen by the customer/user as the query. The system features/ unique properties extracted of the query image, identified the database for images with same features, and exhibits relevant images to the user in-order of same to the query image. In this research, content adds among other features, unique properties such a shape, text, color and spatial data.Various CBIR systems have been developed that compare , analyze and retrieve images based-on one or more features. In content base image retrieval concept some issues occur like shape , texture and the rotation, etc. We resolve the main issues with the help of proposed techniques. In this thesis the problem of image retrieval will be solved using an inverse filter to resolve the blurriness of the image and combination of PCA and BPNN algorithm to extracted features and classify with BPNN algorithm the features based in searching of the queue image by category wise. Firstly feature extraction will be done using BPNN. Then, at first, the neural system is trained based on the features of imageries in the database. The image features considered here are an average value, min value & max. Value. The training is accepted out using BPNN algorithm. This trained when presented with a query image recovers and displays the pictures which are relevant and similar to a query from the database. The results show a considerable improvement in terms of error rates and accuracy of image retrieval. The whole simulation has been taking place in MATLAB environment.

Keywords*: Content based image retrieval, query searching, Feature Extraction (PCA) and classification methods(NN) .

I. INTRODUCTION

Recent advances in science and technology has increased the use of image data in diverse areas such has excitement, workmanship exhibitions, education, fashion design, industry, medicine and so forth. Explosion of World Wide Web (WWW) in the last decade has seen an enormous increase in the usage of digital images and the ease of access these randomly stored images in secluded databases [1]. That is why it is critical to execute an algorithm that helps a lot in searching. The text-based approach is also employed in picture recovery. In content based methodology, the pictures are by hand annotated by text description and then these descriptors are used by database management system to perform image retrieval. The concept of image retrieval from large database has been introduced by Florence. Earlier automatic retrieval was not possible so lots of labor was required. Faced with such limitations, CBIR system has been introduced.

Advancement in the image acquisition and storage lads in the development of the large database for images. It is very imperative to develop a scheme which can manage a large number of images [2]. For this the most common method is CBIR systems (content based image retrieval). These systems try to search the similar image from large database by specifying the query picture in the database. The main benefit is for automatic fitting of the image.

The main concept of the CBIR system is to get the image of similar features from the database [3]. This section will also represent how text based image retrieval is different from content based image retrieval.

Table 1. Difference between CBIR and TBIR

CBIR	TBIR
The Content Based Image Retrieval (CBIR) technique uses text embedded on images to search and retrieve advanced pictures	In text based image retrieval images are searched out using strings

Examples include in which similar images can be searched by specifying the query image.	Examples are texture images that cannot be portrayed by content. Picture contains data that can get it firstly by humans then by computers.
The main goal of CBIR is efficiency during image indexing as well as recovery, in this way lessening the requirement for human negotiation in the indexing procedure.	Systems based on categorizing images in semantic classes like "cat" as a subclass of "animal" avoid this issue, yet confront the similar scaling problems

CBIR systems can be classified into 3 levels based on the complexity [4]:

A. *Level -1 :* Searching for images based on the features of images like color, shape or location of image regions. There is no need of knowing to KDD. The features used for searching objective as well as being direct obtainable. If completely satisfactory performance is required the use of level 1 retrieval is used for trademark registration. In practice, however, most of the existing systems for general-purpose CBIR operate on this level
B. *Level-2:* Level 2 uses the derived attributes for the searching of images from database. There are two features from which differentiation can be made: a) Repossession of objects of a given type and b) Repossession of individual objects. This level of searching of images needs outside information. Most of the searches for newspaper image libraries fall into level 2
C. *Level-3:* Uses abstract attributes for retrieval. This means that high level features are extracted from images. Also this level can usefully be divided into two: a) Repossession of names, events or types of activity b) Repossession of pictures with emotional or religious [5].

Table 2. Basic CBIR systems

CBIR system	Features	Category
VisualSEEk	Matching by color features	Academic
Photobook	Matching by texture features	Academic
Multimedia Analysis and Retrieval Systems (MARS)	Get relevance feedback	Academic [13]
NeTra	Image segmentation before searching	Academic
QBIC	Texture and shape features	Commercial
Virage	Global, local features and similarity measure	Commercial [6]

In this research work, CBIR fights with connecting the semantic gap, defined as the separation between the high-level difficulty of CBIR and human perception and the low-level implementation structures and techniques. The problem of this research work is to extract the features of the image sets using PCA algorithm and train the classification on the basis of the features extracted and also to test the image on the source of the features in the archive and the features extracted of the image to be tested[17].

II. LITERATURE SURVEY

Abdolraheem Khader et al., (2017) [7] studied in CBIR get the similarity results and retrieve images based on one type of feature which are color, texture or shape. In this study authors proposed a fusion based retrieval model for merging results taken from color and texture image features based different fusion methods. After implementing our proposed retrieval model on Wang image dataset which widely used in CBIR, the results show that the CombMEAN fusion approach has the best and high precision value and outperformed both individual color and texture retrieval model in both top10 and top20 retrieved images. Apurva Sharma [8] (2014) clarified a capable calculation rely on upon SURF, color histogram, SVM and NN. Proposed a strategy for figure Matching relied on SURF Algorithm utilizing SVM Classifier, NN nourish forward and color histogram. CBIR was a testing work which

recovers the comparable figures from the extensive database. Numerous CBIR procedures have been proposed before yet they were sufficiently bad and can be briefly messed with so the work was not satisfied. CBIR alone with Surf and SVM Method couldn't give better results. Consequently utilize CBIR with Surf, SVM, NN and color histogram giving improved results. Dr.K.VELMURUGAN et al., [9] (2014) exhibited a paper in procedures IJARCSSE named A Survey of Content-Based Image Retrieval Systems utilizing SIFT. In this paper creator portrays the expanding disagreeability of CBIR and SIFT innovation which was ended up being the most proficient strategy among another strategies grew by the scientist. Ms. A. N. Ganar. et al., [10] (2014) displayed a paper in procedures IJAREEIE named Duplicate Retrieval by Using Color, Texture & Shape Features. This paper portrays about the consolidating of the color, shape, and surface data in CBIR to accomplish effective recovery results. Certain strategies are consolidated like careful calculation to ascertain highlight vector for shape and co-event lattice to figure highlight vector for composition. Henceforth color, surface, shape intertwined elements are strong, then color shape composition highlight based picture recovery strategy. Suraya Abu Bakar et al., [11] (2013) this paper introduces an option approach for Content Based Image Retrieval (CBIR) utilizing Scale Invariant Feature Transform (SIFT) calculation for parallel and dim scale pictures. The inspiration to utilize SIFT calculation for CBIR was because of the way that SIFT was invariant to scale, revolution and interpretation and in addition part of the way invariant to relative bending and light changes. Future work will take a gander at more effective parameterization of highlight descriptors and option routines for better recovery performance. A.Anandh et al., 2016 [12] achieve an efficient result; this paper a proposed a technique for the generation of image content descriptor with three features viz., Color auto-Correlogram, Gabor Wavelet and Wavelet Transform. Color Auto-Correlogram Feature is associated with color information of an image which is derived from the RGB color space of an image. The Gabor Wavelet Feature is has the texture information to extract the textural features associated with the image and the Wavelet Transform Feature is linked with shape information in the extraction of edges in an image. The feature extraction process is accomplished based on the input query image from the IDB and the features are stored in a feature dataset. The Manhattan distance is applied on the user given query image and feature vector computed from database images for measuring similarity. Finally, the proposed technique retrieves the meaningful image from the image database which satisfies the user expectation. The performance of the retrieval system has been analysed by the performance measures Precision and Recall.

III.PROPOSE METHODOLOGY

CBIR is a large number of new techniques and systems are emerging in content-based image retrieval in the last decades, the gap between low-level visual features and high-level semantic understanding of images, which is also known as the semantic gap problem, the gap between the object in the world and the information in a computational description derived from a recording of that scene is the block to further improvement of the presentation of a content-based image retrieval system. Therefore, in order to solve the problems and improve CBIR performance, image annotation, Region-Based Image Retrieval approaches and relevance feedback have been receiving more attention in the recent years. Most research efforts have focused on expanding the number or breadth of visual features available for use in CBIR systems. As mentioned, color, texture, shape, and spectral methods have been developed and combined to varying degrees with different levels of success. Content-Based Image Retrieval has been used in several applications, such as medicine, fingerprint identification, biodiversity information systems, digital libraries, crime prevention, and historical research, among others. In particular, histograms as a representation of color.

Most research efforts have focused on expanding the number or breadth of visual features available for use in CBIR systems. As mentioned, color, texture, shape, and spectral methods have been developed and combined to varying degrees with different levels of success. Content-Based Image Retrieval has been used in several applications, such as medicine, fingerprint identification, biodiversity information systems, digital libraries, crime prevention, and historical research, among others. In particular, histograms as a representation of color.

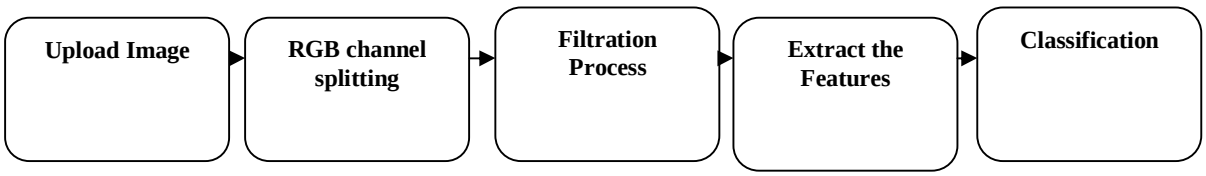


Fig 1. Proposed Flow Chart

Step 1: First, we download the dataset form the uci machine repository site. Define the categories of the CBIR images. Upload the image form the database. Calculate the RGB component in the original image. To convert the original image to gray-scale conversion cause of reducing the pixel size of the original image.
Step 2: Check the noise attack in the gray scale image. If noise attack is present in the gray-scale image, we implement the filtration approach (Bilateral and Inverse) for remove the noise attack in the original image.
Step 3: Implement the Principle component analysis algorithm in the image. Because, to find the uniqueness of the features of the filtration image. Feature Extraction algorithm to find the Eigen values and Eigen vector features is called extracted features.
Step 4: The proposed work, we implement the back propagation neural network for classification. It will work in two phases;

A. Training Phase and (ii) Testing Phase.

In training Phase, we check the uniqueness of the extracted features and saved in the training phase. In testing Phase, identifies the category of the content base image retrieval. Evaluate the performance parameters i.e false acceptance rate, false rejection rate, precision, recall and compare the existing parameters (Precision and Recall).

IV. DATA SET DESCRIPTION AND RESULT ANALYSIS

In this section, the datasets used in object classification are presented. The datasets can be divided in four groups: Facial images, car images, flower images and sunset images.

Fig 2. Sunset Images [14]

Fig 3. Rose Images

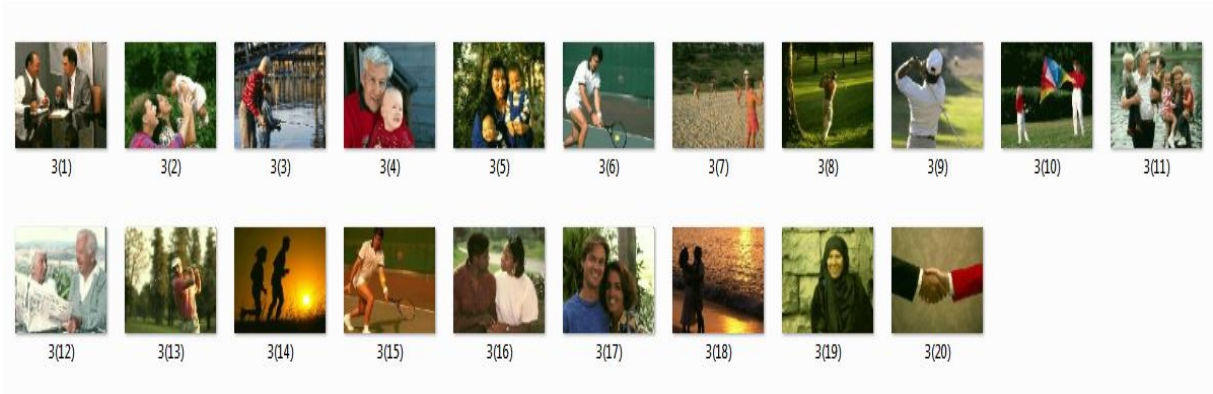


Fig 4. Facial Images

Fig 5. Car Images[15]

Fig 6. Image uploaded

The figure shows that the upload the category one image which is facial image, Category image two is car, category three is sunset and category four is flower. the gray scale image means reduce the original pixel matrix using rgb2gray scale command used . To reduce the image size convert the 2D transformation. The edge detection,image means generate the regions based on maximum, minimum and average values. The canny edge detector is an edge detection that uses a multistage algorithm to identify a world wide range of edges in images. Noise check in the grayscale image. We have found the salt and pepper and Gaussian noise. An effective noise reduces and optimize method of this kind of noise is an inverse filter or a morphological filter. An inverse filter is a non-linear digital filtering method, given used to remove noise. Some noise feature extracted is a kind of preprocessing phase to enhance the consequences of later processing. In feature extraction technique is used to recover most revealing terms from amount of matrix. This study used Component Analysis technique to calculate and study the Eigenvector and values to find the feature values and then to direct individual data with its principle components / Eigen Vector.

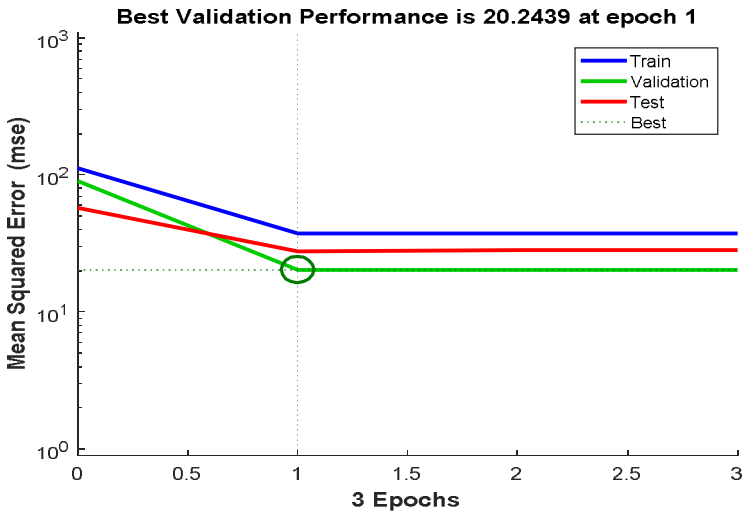


Fig 7. Performance Analysis

The above figure defined the architecture based on hidden neurons. In back propagation neural network data shows in three form input,output and hidden layers. Initialize the dataset define that based on 100 iterations, but the input process in 3 iterations , time consumed in 2 seconds , performance is 37.4 , gradient value is 9.12 and mutation 1.0 , validation checks 2 process out of 6 . A best validation process based on epoch1 .3 epoch data process, but based on mean square error means training error and the test error

sum is equal to MSE. The training state evaluation based on gradient means check the line of slope based on best validation performance. Mutation means how many data changed. The validation check mean data divide into phases, i.e, modules.

Table 3.False Acceptance Rate and False Rejection Rate (Proposed work)

No of Images	FAR	FRR
Car	0.0054	0.00125
Sunset	0.0061	0.0013
Facial	0.00973	0.00145
Flower	0.0082	0.0016

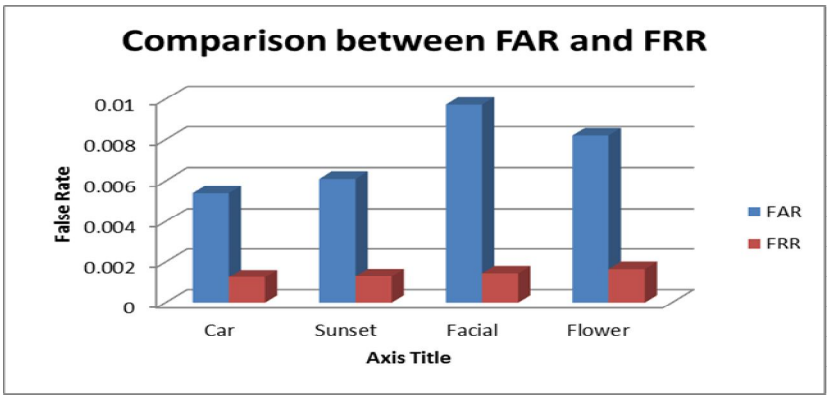


Fig 8. FAR and FRR

The above figure defined that the false acceptance rate means how much wrong data acceptable in the original form. The false rejection rate means how much wrong data reject able in the original data set.

Table 4. Precision and Accuracy in Proposed work

No of Images	Accuracy	Precision	Recall
Sunset	90	0.91	0.92
Car	91	0.93	0.93
Facial	93	0.93	0.94
Flower	93.5	0.95	0.96

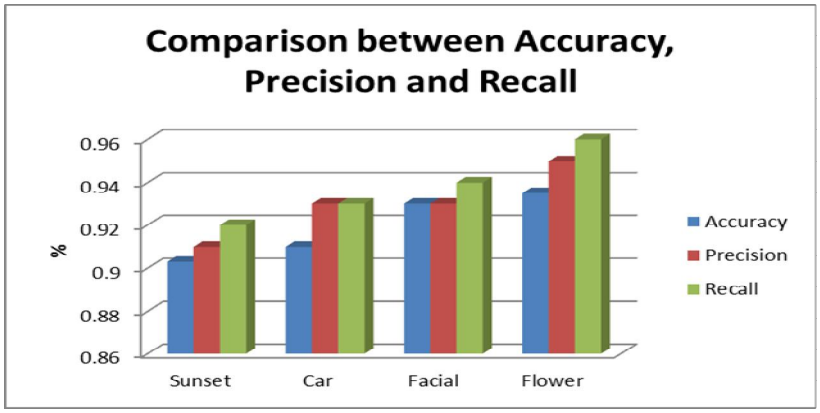


Fig 9. Performance parameter accuracy and precision, recall

The above figure defined that the accuracy means accurate the system based on BPNN and False rejection Rate and False Rejection rate is less then improve the performance the image of the dataset. It represents that the precision that the nearest of two or more measurements to each other.

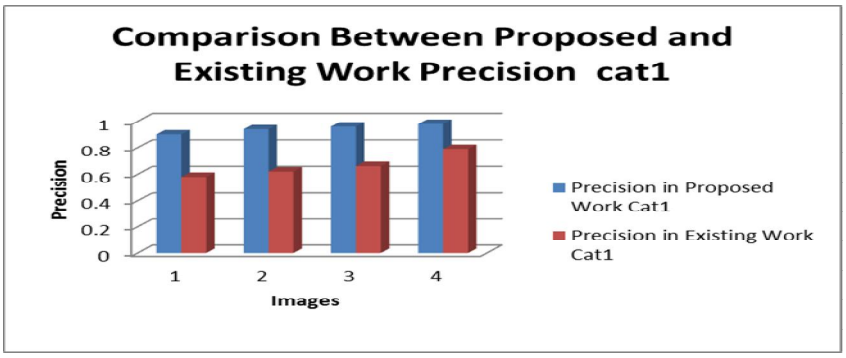


Fig 10. Comparison between proposed and existing work (precision in cat 1)

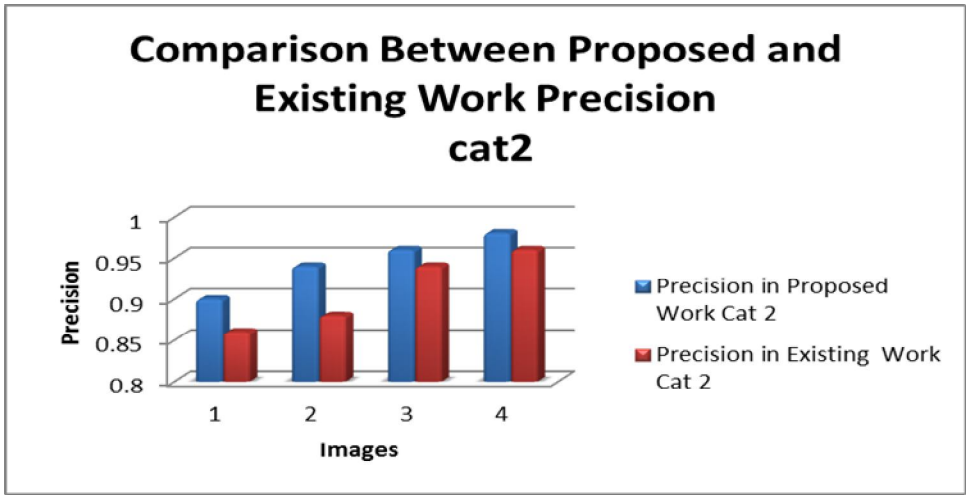


Fig 11. Comparison Between proposed and existing work (Precision in Cat 2)

The figure shows that the comparison between the proposed and existing work in precision performance in category -2 in flower images. The performance parameters in precision value are proposed and existing work value is 0.98 and 0.96.

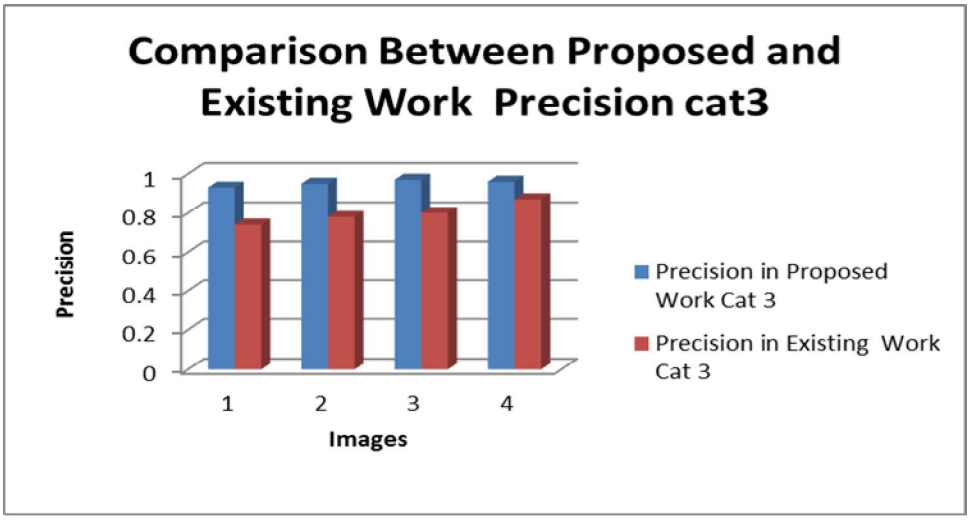


Fig 12. Comparison Between Proposed and Existing work in cat -3 precision

The figure shows that the comparison between the proposed and existing work in precision performance in category -3 in facial images. The performance parameters in precision value are proposed and existing work value is 0.96 and 0.87.

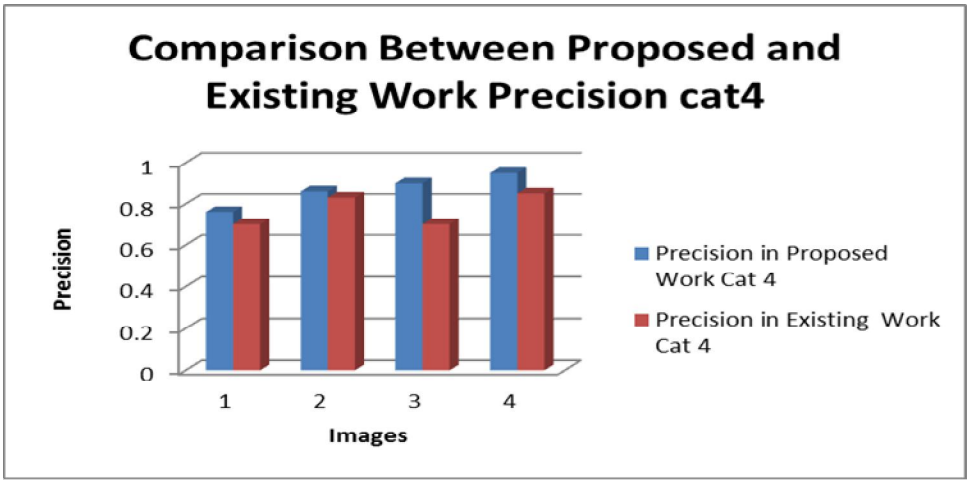


Fig 12. Comparison Between Proposed and Existing work in cat -4 precision

The figure shows that the comparison between the proposed and existing work in precision performance in category -4 in car images. The performance parameters in precision value are proposed and existing work value is 0.95 and 0.85.

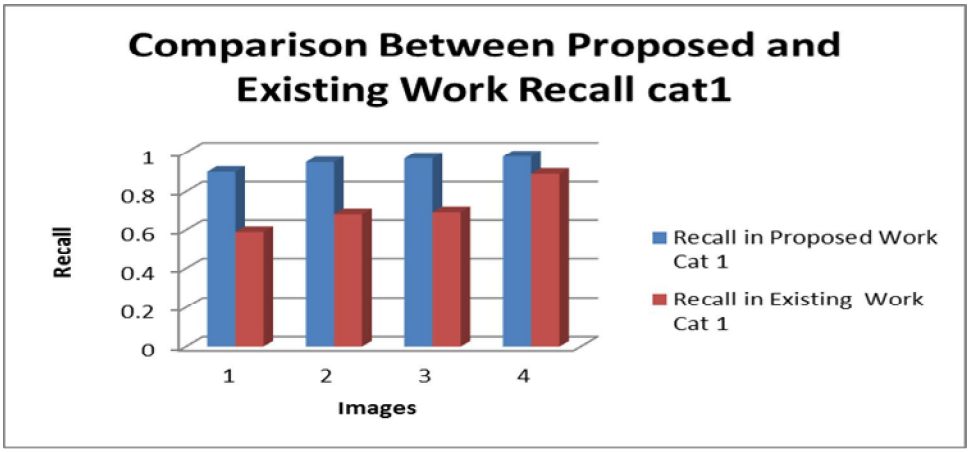


Fig 13. Comparison Between Proposed and Existing work in cat -1 Recall

The figure shows that the comparison between the proposed and existing work in recall performance in category -1 in sunset images. The performance parameters in recall value are proposed and existing work value is 0.98 and 0.89.

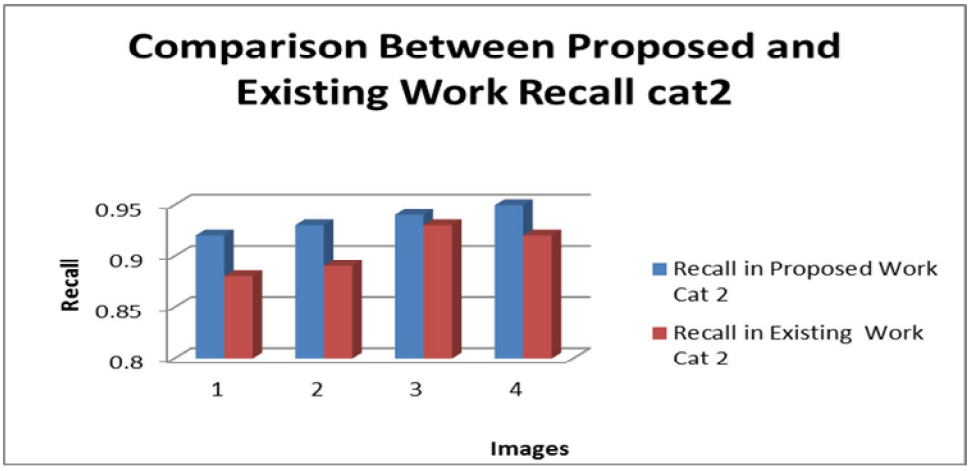


Fig 14. Comparison Between Proposed and Existing Work in cat 2 Recall

The figure shows that the comparison between proposed and existing work in recall performance in category -2 in flower images. The performance parameters in recall value are proposed and existing work value is 0.95 and 092.

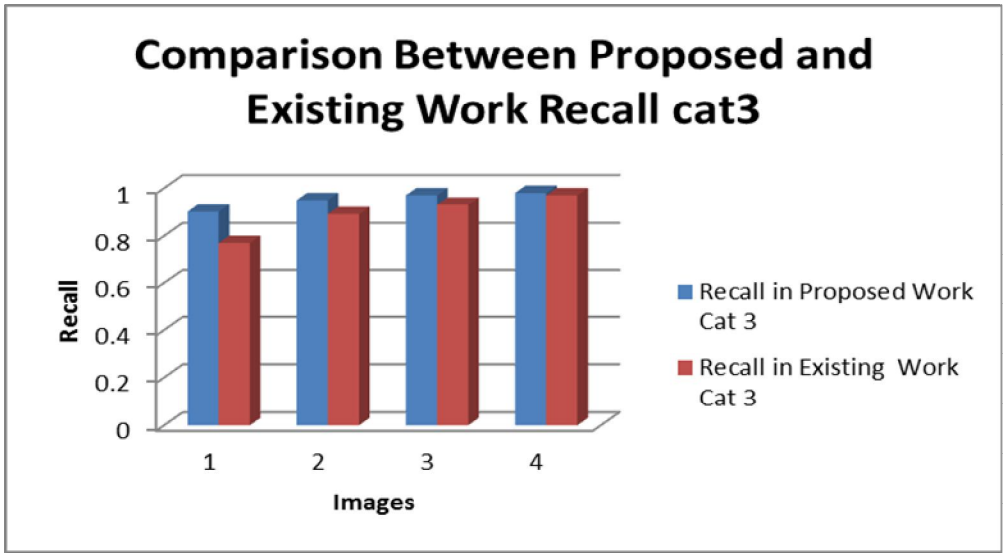


Fig 15. Comparison Between Proposed and Existing Work in cat 3 Recall

The figure shows that the comparison between proposed and existing work in recall performance in category -3 in facial images. The performance parameters in recall value are proposed and existing work value is 0.95 and 0.85.

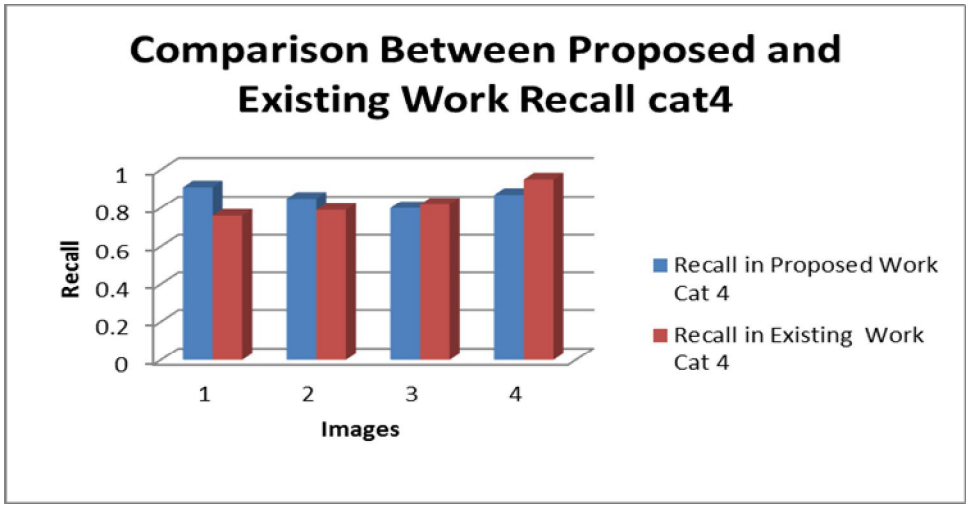


Fig 16. Comparison Between Proposed and Exising work in calt -4 Recall

The figure shows that the comparison between proposed and existing work in recall performance in category -4 in car images. The performance parameters in recall value are proposed and existing work value is 0.87 and 0.95.

V. CONCLUSION AND FUTURE SCOPE

Motivated by the limitations of such an approach, recent research focus in CBIR has moved to “interactive systems” and “human in the loop” that involves a human as part of the retrieval process. In this thesis the problem of image retrieval will be solved using a combination of Inverse, PCA and BPNN. Firstly feature extraction will be done using PCA. Then, at first, the neural network is trained based on the features of images in the database. In PCA algorithm used to extract the feature set. This algorithm is an initialization the feature extraction using eigen values and eigen vectors.The dramatic rise in the sizes of image databases has steered the development of effective and efficient retrieval systems. The growth of these systems started with retrieving images using textual suggestions, but later introduced image retrieval based on content. This came to be known as Content Based Image

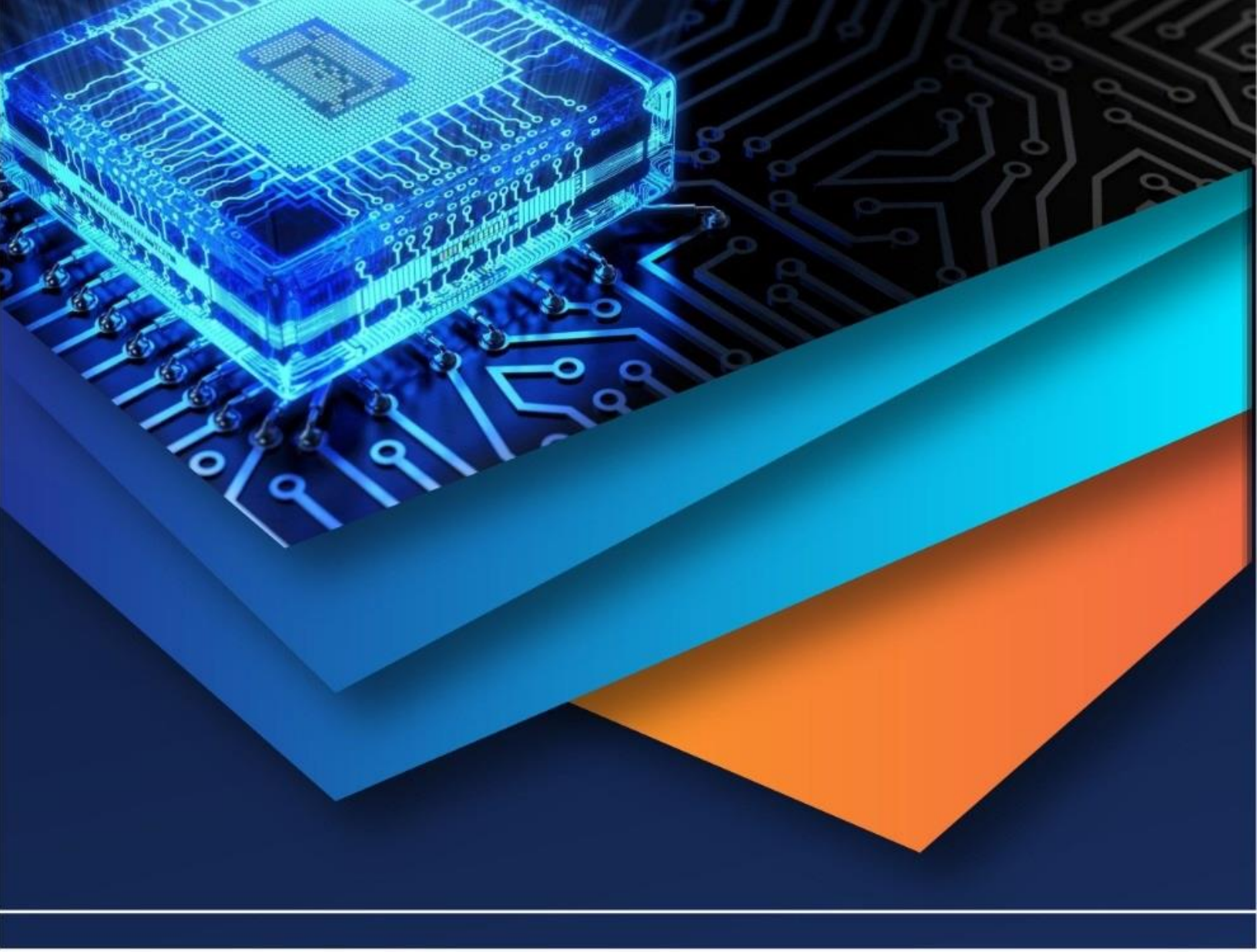
Retrieval or CBIR. Systems using CBIR retrieve images based on visual features such as texture, color and shape, as opposed to depending on image descriptions or textual indexing.

The main objective of this paper is to retrieve the images from database in a fast and an efficient manner using modified Back Propagation Neural Network (BPNN). This trained when presented with a query image retrieves and displays the images which are relevant and similar to a query from the database. The results show a considerable improvement in terms of FRR, FAR and accuracy of image retrieval.

Future scope includes implementing the CBIR system considering more low-level image descriptors and highly efficient deep learning neural network, that may possibly verify to be quite fast as well as precise one. This work can be extended by integrating with Fuzzy C-means clustering algorithm for better efficiency.

REFERENCES

[1] Long, F. H., H. J. Zhang and D. D. Feng, Fundamentals of content-based image retrieval, in Multimedia Information Retrieval and Management - Technological Fundamentals and Applications, Springer-Verlag, pp.1-26, New York, 2003.
[2] Musoko, V, Kolínová, M., Procházka, A, Image Classification Using Competitive Neural Networks, 12th Scientific Conference MATLAB2004, Humusoft, 2004.
[3] Kato, K., Database architecture for content–based Image retrieval, 377 volume 1662 of SPIE Proceedings, pp. 112–123, San Jose, 1992.
[4] Müller, H., User interaction and evaluation in content-based visual Information retrieval, PhD thesis, University of Geneva, Geneva, 2002.
[5] Grubinger, M., Analysis and evaluation of visual information systems performance, 2007.
[6] Eakins, J.P. and Graham, M. E., Content–based image retrieval, JISC TechnologyApplication Program, University of Northumbria, Newcastle upon Tyne, 2000.
[7] Alhassan, Abdolraheem Khader, and Ali Ahmed Alfaki. "Color and texture fusion-based method for content-based Image Retrieval." In Communication, Control, Computing and Electronics Engineering (ICCCCEE), 2017 International Conference on, pp. 1-6. IEEE, 2017.
[8] ApurvaSharma" content based image retrieval with surf ,colour histogram, SVM and neural network "International Journal of Latest Scientific Research and Technology 1(2), July - 2014, pp. 31-36
[9] Dr.K.VELMURUGAN" A Survey of Content-Based Image Retrieval Systems using sift" ijarcsse Volume 4, Issue 1, January 2014
[10] Ms. A. N. Ganar" Image Retrieval by Using Colour, Texture and Shape Features" ijareeie Vol. 3, Issue 4, April 2014.
[11] Suraya Abu Bakar" Content-Based Image Retrieval using SIFT for Binary and grayscale Images" IEEE 2013.
[12] Anandh, A., K. Mala, and S. Suganya. "Content based image retrieval system based on semantic information using color, texture and shape features." In Computing Technologies and Intelligent Data Engineering (ICCTIDE), International Conference on, pp. 1-8. IEEE, 2016.
[13] S. Liapis and G. Tziritas, "Color and texture image retrieval using chromaticity histograms and wavelet frames," IEEE Trans. Multimedia, vol. 6, pp.676-683, Oct. 2004.
[14] S. Antani, R. Kasturi, and R. Jain, "A survey of the use of pattern recognition methods for abstraction, indexing and retrieval of images and video," Pattern Recognition, vol. 35, pp. 945–965, Apr. 2002.
[15] Sung BAE ET. Al, "Human Oriented Image Retrieval Syatem Using Genetic Algorithm ", ACM, Vol. 32, pp. 452-458, May 2002.
[16] Smeulders, A. W., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22(12), 1349-1380.
[17] X. S. Zhou and T. S. Huang, "Relevance feedback in content-based image retrieval: Some recent advances," Information of Science, vol. 148, pp. 129–137, Dec. 2002.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY