


A Visual Saliency-Based Approach for Content-Based Image Retrieval

Aamir Khan, Independent Researcher, India

Anand Singh Jalal, GLA University, Mathura, India

 <https://orcid.org/0000-0002-7469-6608>

ABSTRACT

During the past two decades an enormous amount of visual information has been generated; as a result, content-based image retrieval (CBIR) has received considerable attention. In CBIR the image is used as a query to find the most similar images. One of the biggest challenges in CBIR system is to fill up the “semantic gap,” which is the gap between low-level visual features and the high-level semantic concepts of an image. In this paper, the authors have proposed a saliency-based CBIR system that utilizes the semantic information of image and users search intention. In the proposed model, firstly a significant region is identified with the help of method structured matrix decomposition (SMD) using high-level priors that highlight the prominent area of the image. After that, a two-dimensional principal component analysis (2DPCA) is used as a feature, which is compact and effectively used for fast recognition. Experiment results are validated on different image dataset having an extensive collection of semantic classifications.

KEYWORDS

Image Retrieval, Query-Based Image Retrieval, Region of Interest

INTRODUCTION

It is an era of searching on the internet. Nowadays searching is done using a digital media known as Multimedia Information Retrieval (MIR). The biggest challenge in today's era is to find the solution for the searching of digital information. The need of as Multimedia Information Retrieval system is required to search the image/video effectively and efficiently from the vast corpus of data available on the internet with the help of search engine. Searching on the web is done either in text query or by multimedia as a query. MIR deals with multimedia information like image/video which is further divided into branches such as Content Based Image Retrieval (CBIR), Michael S Lew et al. (2006). In early years, Myron Flickner et al. (1995) developed a content based retrieval system known as QBIC (Query by Image Content). In CBIR the image is used as a query to find the matched image from the database of image corpus on the web. Features of images are used for the matching of the images which are visually similar i.e., shape or texture or color or with the combination of these features.

DOI: 10.4018/IJCINI.2021010101

This article, published as an Open Access article on November 6, 2020 in the gold Open Access journal, International Journal of Cognitive Informatics and Natural Intelligence (converted to gold Open Access January 1, 2021), is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

There are various features frequently used in CBIR systems: Color, textures, shape etc. The color feature represents the visual appearance of an object. It is also invariant to many geometric transformations. Color histograms, color moments, color coherent vector (CCV), and color correlogram are the descriptors used to represent the color feature. Image texture defines the spatial arrangement of color or intensities. Image texture can be described in term of smoothness, coarseness, roughness. Recently, SIFT feature is also used as a local feature for image representation by David G Lowe (2004). Timo Ojala et al. (1996) proposed texture histograms feature by combing simple texture features. Shape is also used as one of the global feature. There are different shapes descriptors such as circularity, moment invariant, aspect ratio, boundary segments, eccentricity, orientation and Fourier descriptors are used in CBIR systems. Apart from using single feature, researchers have exploited the combinations of different features such as color and texture used by researchers Xiang-Yang Wang et al. (2014), color and shape by researchers Theo Gevers and Arnold WM Smeulders (2000) and color, texture and shape by researchers Xiang-Yang Wang et al. (2011). Combing features have improved the efficiency and accuracy of retrieval systems. However, combining features create the feature vector heavy and retrieval system more complex. So there is a need of feature that is compact and as well as efficient.

The majority of the works in CBIR have focused on narrow image fields which usually represent images related to a particular domain such as satellite images; medical images Arnold WM Smeulders et al. (2000). In domain-specific images, the images are restricted around a small semantic concept. On the other hand, broad domains represent images having a large set of semantic concepts such as natural images on the web Arnold WM Smeulders et al. (2000). These images also contain significant lack of consistency within them. Developing a system that can deal with the broad image domain is much more challenging. In general, in an image retrieval system, user visual attention is only focused in specific regions also known as objects of interest.

Most of the CBIR frameworks assume each picture as a whole semantic unit and assemble primitive components Ritendra Datta et al. (2009). The global features are computed from the entire image. Therefore, the retrieval precision such systems are low Ying Liu et al. (2007). In whole matching approaches, the complete image is used for matching. In such approach, a lot of irrelevant regions are also included because the features of the entire image area are integrated into one or several global feature vectors (e.g., color histograms). Many researchers proposed region-based approaches, to overcome the issues related to the global features and to extract the semantic information, P Manipoonchelvi and K Muneeswaran (2015). A region-based image retrieval (RBIR) framework contains different modules such as segmentation, feature extraction and similarity matching. The segmentation module extracts the various homogeneous regions from an image. There are different segmentation methods such as pixel-based, boundary based and region-based methods used in literature. The feature extraction module extracts either low-level features such as color, shape, and texture or high-level semantic features. The high-level semantic features are computed from the low-level features. The prior knowledge also plays an important role in extracting high-level semantic features.

Cheng-Chieh Chiang et al. (2009) proposed a method which allows the user to specify regions of interest as a query. Based on the color-size histogram irrelevant images are filtered out. An Earth Mover's Distance (EMD) based similarity measure is used to rank and matching the resulting candidate images. P Manipoonchelvi and K Muneeswaran (2015) proposed visual attention based method for significant region extraction. A color layout descriptor (CLD) is used to represent the regions. In their method, first the image is segmented into homogeneous color regions. Then for each region in the image, they compute visual descriptors such as color contrast, proximity to the centre of the image, size and nearness to image's boundary for each region in the image. Finally, on the basis of the visual metrics, the significance of each region is measured and identifies the central regions.

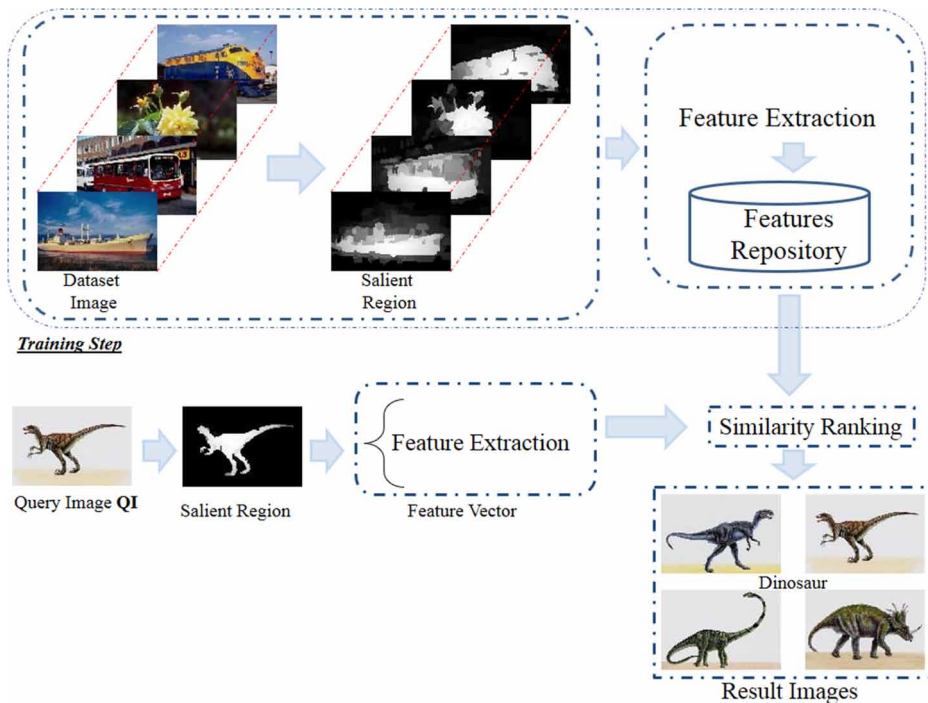
In a region-based approach, image is divided into different regions and visual features are computed on these regions. Region based feature have shown higher retrieval precision as compared to

global features. However, the success of region based approaches largely depends on the segmentation technique. In recent year saliency-based approaches have shown better performance in image segmentation by Alex Papushoy and Adrian G Bors (2015), HouwenPeng et al. (2017). Humans are capable to automatically detect regions of interest in their surroundings using visual search which is the ability to locate an object through attention by Jeremy M Wolfe (1998) consciously. A multi-level visual attention model is proposed by the author X Wang and H Duan (2017) for the task of saliency detection, this method is especially designed for the atomic nuclei which are linked to the visual attention mechanism. Author R. Manke et al. (2018) and M Singh et al. (2019) uses high features such as color, contrast and semantic prior to identify the salient region. Visual attention helps the brain for high-level cognitive processing by extracting specific visual information, i.e., salient details by researcher Simone Frintrop et al. (2010). Inspired by this cognitive approach, the proposed system computes salient region which represents the visual focus of attention. Therefore, in this paper, authors have proposed a saliency-based CBIR system having low complexity. The proposed image retrieval approach utilizes the semantic information of image and users search intention. In addition to this, in the proposed method, authors have used a 2DPCA Jian Yang et al. (2004) as a feature, which is compact and can be effectively used for fast image search.

Proposed Methodology

The block diagram of the proposed framework is elaborated in Figure 1. In the proposed framework, first, a Region of Interest (ROI) is identified using a salient based approach, HouwenPeng et al. (2017) to localize the feature vector. The second step is feature extraction, in which features are extracted using two-dimensional principal component analysis (2DPCA), Jian Yang et al. (2004). Finally, a Euclidian Distance method is used to measure the similarity between the query image and database images. In Figure 1, the first block is for training, from left to right, Original image set, Saliency Map

Figure 1. Proposed framework for content based image retrieval



Generation, and feature extraction, below the training block from left to right Query image, Salient region of query image, feature extraction, Similarity Ranking.

Region of Interest Identification

In the proposed approach, an existing saliency detection method known as Structured Matrix Decomposition (SMD) Houwen Peng et al. (2017) is used to detect the region of interest in the image. In this method the low-level features are integrated with high-level priors to find the salient region. The low-level salient object detection consists of four steps: image abstraction, index tree construction, matrix decomposition and saliency assignment HouwenPeng et al. (2017). In the image abstraction step, an input image is partitioned into compact and perceptually homogeneous patches. Each patch P_i is represented by a feature vector f_i , and all these feature vectors form the feature matrix as $F = [f_1, f_2, \dots, f_N]$. After that an index tree T is constructed to encode structure information via hierarchical segmentation.

The high-level priors are computed using the fusion of three types of priors, i.e. location, color and background priors, to generate a high-level prior map. A Gaussian distribution based on the distances of the pixels from the image centre is used to compute the location prior. The color prior is computed on the fact that human eye sensitivity to red and yellow color. The background prior calculates the probabilities of image regions connected to image boundaries. By integrating the low level prior with high level priors in the SMD process it works better for the complex background image, therefore also useful to retrieve the complex backgrounds images.

Feature Extraction

Feature extraction is the most crucial step in the content based image retrieval system as it directly affects the retrieval rate. Apart from using multiple features, in the proposed approach, a single feature i.e., a 2DPCA has been exploited for image retrieval task. Unlike multiple feature this single feature advantage to form efficient feature vector. Due to its stability and performance, PCA is considered as one of the standard approach to extract feature. However, PCA is failed to extract all local variance. The training data is explicitly used to provide this information. To overcome this issue, a variant of PCA known as 2D Principal Component Analysis (2DPCA) is proposed by Jian Yang et al. (2004). An image in 2PCA is represented as matrices, which helps to preserve the local information and extract more significant features.

Let A is an n -dimensional unitary column vector. The idea of 2DPCA is to project a given image X , a $m \times n$ matrix, onto A to get an m -dimensional vector Y by the linear transformation:

$$Y = XA \quad (1)$$

Authors call Y the projected feature vector of X . The procedure of 2DPCA to obtain the A vector can be characterized as follows.

Suppose there are n number of classes in R^{mn} space and there is a sample set including N images, $\{X_1, X_2, \dots, X_N\}$, where $X_i \in R^{mn}$; each sample belongs to a class j , $j \in \{1, 2, \dots, c\}$. Total scatter matrix can be defined as:

$$S_T = \frac{1}{N} \sum_{k=1}^N (X_k - \bar{X})(X_k - \bar{X})^T \quad (2)$$

where N denotes the number of training samples and $\bar{X} = \frac{1}{N} \sum_{k=1}^N (X_k)$ is the mean of all training samples.

In 2DPCA algorithm, the optimal projection vector A_{opt} satisfies with:

$$A_{opt} = \operatorname{argmax} |A^T S_T A| = [A_1, A_2, \dots, A_m] \quad (3)$$

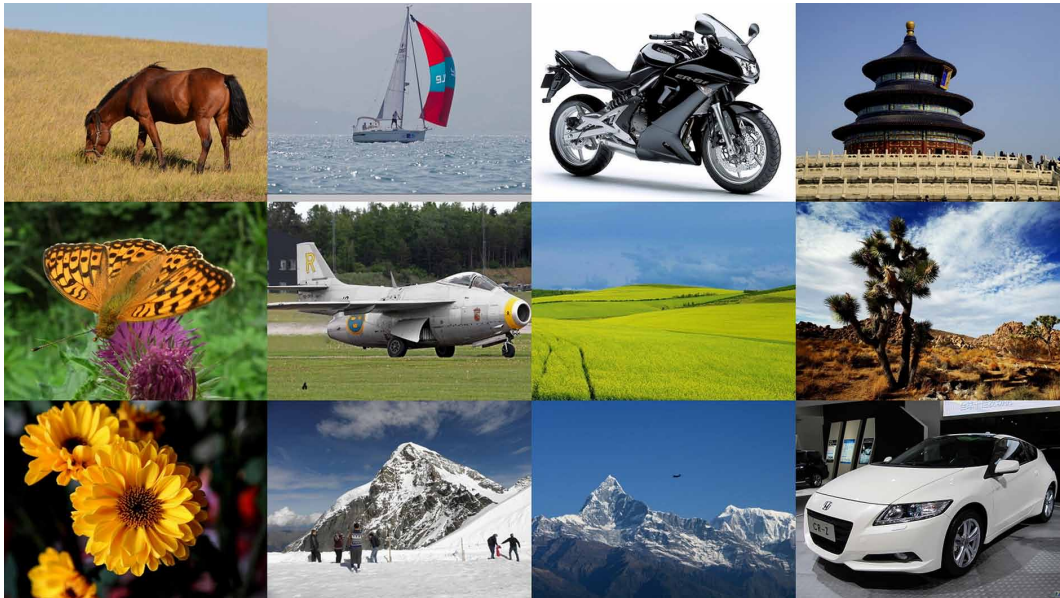
where $\{A_i | i = 1, 2, \dots, m\}$ are m orthogonal eigenvectors ST of corresponding to the largest m eigen values, respectively.

The recognition process is applied on the weight vectors of input test image to the weight vectors of the trained images. First researchers are finding the distance between the input images to the weight vectors of the trained image. By calculating the distance between these images researchers get some weight value. Hence, by setting the threshold researchers get some value and the minimum value showing accurate results as image identification.

Figure 2. Algorithm of proposed image retrieval model

Algorithm 1: Proposed Image Retrieval Algorithm
Input : Query image I_Q Output : Retrieved images $I_r = I_{r1}, I_{r2}, I_{rn}$
1. Segment the query image I_Q into Region of interest using SMD 2. Extract the Salient Region I_{SR} 3. Perform Normalization on an n-dimensional grayscale image $I : \{X \subseteq R^n\} \longrightarrow \{newMin, newMax\}$ With intensity values in the range (Min,Max), into a new image $I_N : X R^{nnewMin, newMax}$ With intensity values in the range (newMin, newMax). $\overline{I_S R^N} = (I - Min) \frac{newMax - newMin}{Max - Min} + newMin$
4. Compute feature vector (Z) using 2DPCA as follows: <ul style="list-style-type: none"> • For the image ISR calculates the mean, image differencing • $Y_i = x_i - \overline{I_S R^N}$ • Where $\overline{I_S R^N} = \frac{1}{n} \sum_{i=1}^n I_N$, Where n is the numbers of ROI of training image, and $\overline{I_S R^N}$ is the average value of the images • Calculate covariance Matrix for Y_i and Eigen vector is computed using the covariance matrix.
5. Compute Similarity Ranking using Euclidian Distance $\ p\ = Z_i - A ^2$ Where i=1, 2, 3, 4, 5, 6 ... n number of image and A is the feature vector of the Test image.
6. Image Retrieval(FV_{cLD}, FV_{cD})

Figure 3. Sample images of GHIM-10K dataset



Experimental Results and Analysis

To validate the accuracy and performance of the proposed framework extensive experiments are performed on different standard datasets. The Proposed system is implemented on MATLAB 2014a, having hardware configuration of Intel Core i5 2.90 GHZ CPU with 16 GB of RAM. Coral-10K (Guang-Hai Liu. (2017)), GHIM-10K (Guang-Hai Liu. (2017)), Olivia-2688 (Aude Oliva and Antonio Torralba (2001)) datasets are used to validate the proposed system, in many QBIR system these are used as a benchmark.

DATASETS

Dataset 1 (GHIM-10K)

The dataset GHIM-10K (Guang-Hai Liu. (2017)) consists images of twenty categories. In this dataset, there are wide varieties of natural images such as fireworks, insects, horses, bikes etc. There are total 10,000 images in this dataset. The size of the images is either 300 x 400 or 400 x 300. Each class has 500 images. For the experimental evaluation Aircraft, Boat, Bike, Butterfly, Car, Flower, Hen, Horse, Ship categories of image are considered as shown in Figure 3.

Dataset 2 (Corel-10K)

Corel-10K dataset (Guang-Hai Liu. (2017)) is the second dataset that researchers have used for evaluation. The images have been divided into hundred different categories. Each image is of size either 256 x 384 or 384 x 256. There are 100 images in each category. The number of images in this dataset is 10,000. Bus, Cat, Dinosaur, Dog, Doll, Flower, Mountain, Ship, Train image categories are considered for the experimental evaluation as shown in Figure 4.

Dataset 3 (Olivia-2688)

Olivia-2688 dataset (Aude Oliva and Antonio Torralba (2001)) is another dataset that researchers have used for experimentation. The dataset have eight different categories of images, namely, Coast,

Figure 4. Some of the sample images from the coral-10K dataset



Forest, Highway, Inside city, Mountain, Open country, Street, Tall building as shown in Figure 5. Each image is of size 256 x 256. Total number of images in this dataset is 2688.

RESULT AND DISCUSSION

Qualitative Analysis

The saliency region extraction based technique used in this paper effectively suppresses the background regions and retains only the foreground region. This mechanism considers both the high level and low level features, with 2DPCA requires less number of parameters. The proposed approach is aimed to improve the efficiency and accuracy of the content- based imaging applications especially CBIR. The experimental results illustrate that the proposed approach improves accuracy and effectiveness

Figure 5. Sample images of Olivia-2688 dataset

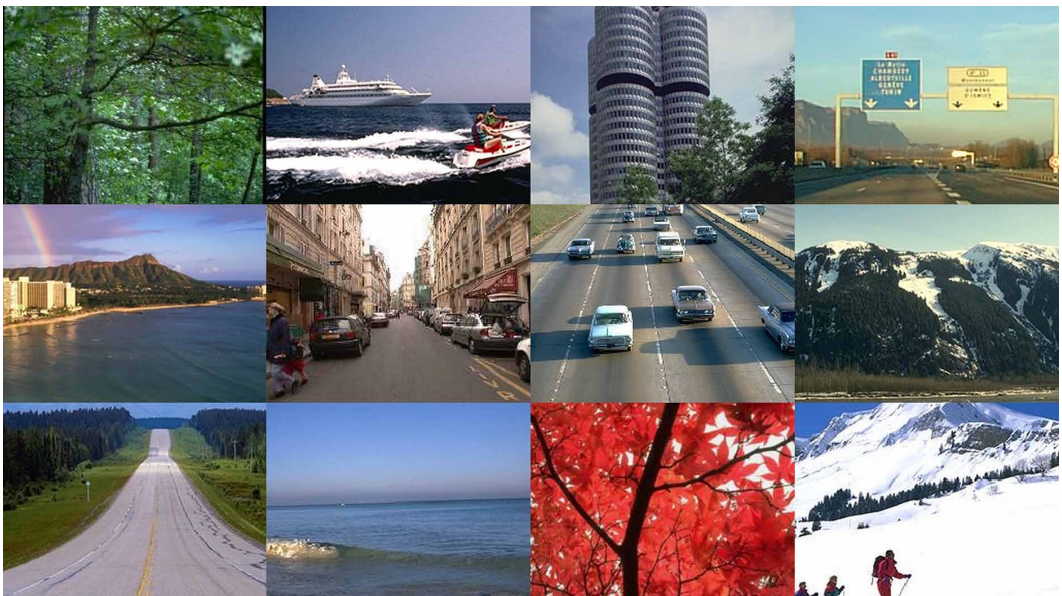
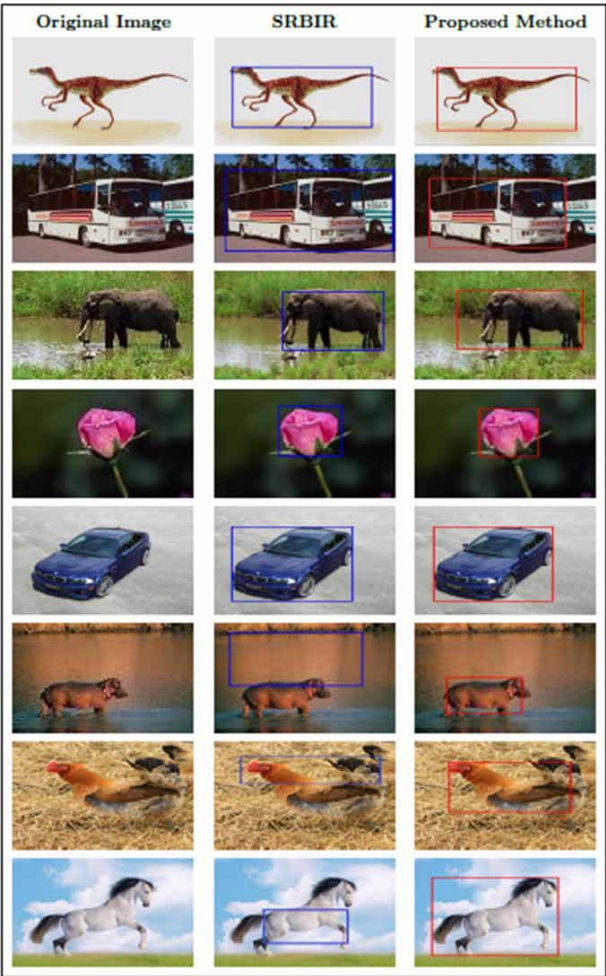


Figure 6. Region of Interest (ROI) extraction using the proposed model and SRBIR



compared to the existing systems with GSD, Megha Agarwal and RP Maheshwari (2012) and SRBIR, P Manipoonchelvi and K Muneeswaran (2015). In the figure 5 some sample results of salient region output of the proposed method is compared SRBIR, P Manipoonchelvi and K Muneeswaran (2015). When the background in simple both methods are able to identify the significant region (i-v), however, when the background of the image is complex (vi-vii) existing approach is lagging and the proposed model outperform.

One of the key expectations from a CBIR system is to fill the semantic gap. It means, the low-level features imply the query image's semantic content. To achieve the same, authors assume that each image contains semantically meaningful regions. Out of these semantically meaningful regions there is one region known as salient region, which reflect the image's semantic content. Figure 6 demonstrate that the results meet the users' satisfaction even in the complicated background.

Quantitative Analysis

The quantitative performance of the proposed approach has been evaluated on the basis of Accuracy, F-measure, Precision and Recall. All of the four measuring parameters are used to compute the relevancy of the retrieved images.

For measuring the accuracy of the retrieval system Precision is used. It is the ratio of retrieved images that are relevant to the query image:

$$Precision = \frac{tp}{tp + fp}$$

For measuring the robustness of the retrieval system Recall is used. It is defined as the ratio of relevant images in the database that are retrieved in response to a query image:

$$Recall = \frac{tp}{tp + fn}$$

Other related measures used in classification include true negative rate and accuracy:

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn}$$

- tp (True Positive) is considered as correctly identified i.e. actual cats that were correctly classified as cats;
- fp (False Positive) incorrectly identified i.e. dogs that were incorrectly labeled as cats;
- fn (False Negative) incorrectly rejected i.e. cats that were incorrectly marked as dogs;
- tn (True Negative) correctly rejected i.e. all the remaining animals, correctly classified as non-cats.

If and only if the query image and matched image both lies in the same semantic group, then the result is considered as relevant image.

An F-measure measure is the unified performance measure:

$$F = 2 * \frac{precision * recall}{precision + recall}$$

Table 1 illustrates the average precision, recall, accuracy and F-measure of the proposed with the existing CBIR system on coral-10K Dataset on the categories Bus, Cat, Dinosaur, Dog, Doll, Flower, Mountain, Ship and Train. It is depicted that proposed system achieved the highest average, precision, recall and F-measure.

Figure 7 depicts the parametric comparison of different approaches on images categories of Coral 10K dataset listed in the table 1. Result is compared with the existing approach SRBIR and GSD. It is demonstrated from the graph that proposed model shows better results even in case of complex background.

Table 2 compare the average retrieval precision, recall and f-measure of the proposed system with the existing CBIR system. The experiment is performed on GHIM-10K Dataset on the categories Aircraft, Boat, Bike, Butterfly, Car, Flower, Hen, Horse and Ship. The results illustrate that the proposed approach performs better than the state-of-art methods.

In Figure 8 a graphical comparison of existing approach SRBIR and GSD is demonstrated with the proposed model on GHIM 10K dataset.

Table 1. Comparison of precision, recall, accuracy and F-measure obtained by the proposed method with existing content-based image retrieval systems GSD, Megha Agarwal and RP Maheshwari (2012), and SRBIR, P Manipooncheli and K Muneeswaran (2015) for Coral 10K Dataset

	Precision			Recall			Accuracy			F-measure		
	SRBIR	GSD	Proposed	SRBIR	GSD	Proposed	SRBIR	GSD	Proposed	SRBIR	GSD	Proposed
Bus	76.2	78.9	85.0	80.0	75.0	85.0	95.5	95.5	96.8	78.0	76.9	85.0
Cat	70.0	75.0	80.0	70.0	75.0	80.0	94.1	95.0	95.7	70.0	75.0	80.0
Dinosaur	71.4	66.7	83.3	75.0	60.0	75.0	94.6	93.3	95.8	73.2	63.2	78.9
Dog	63.6	69.6	73.9	70.0	80.0	85.0	93.3	94.6	95.2	66.7	74.4	79.1
Doll	82.4	80.0	88.9	70.0	80.0	80.0	95.5	95.7	96.8	75.7	80.0	84.2
Flower	73.3	71.4	77.8	55.0	50.0	70.0	93.7	93.3	94.8	62.9	58.8	73.7
Mountain	61.9	52.4	61.9	65.0	55.0	65.0	92.5	91.3	92.3	60.0	53.7	63.4
Ship	76.2	76.2	85.7	80.0	80.0	90.0	95.5	95.5	97.3	78.0	78.0	87.8
Train	54.5	59.1	61.9	62.0	64.0	65.0	91.3	92.5	92.3	55.8	61.9	63.4

Figure 7. Graph comparison of PM with GSD and SRBIR on Coral 10 dataset

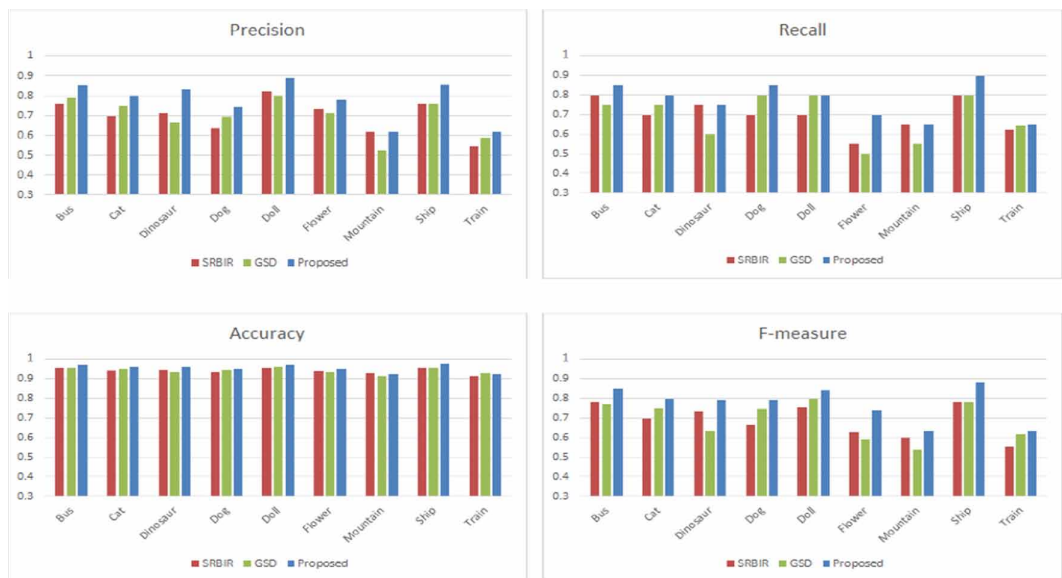


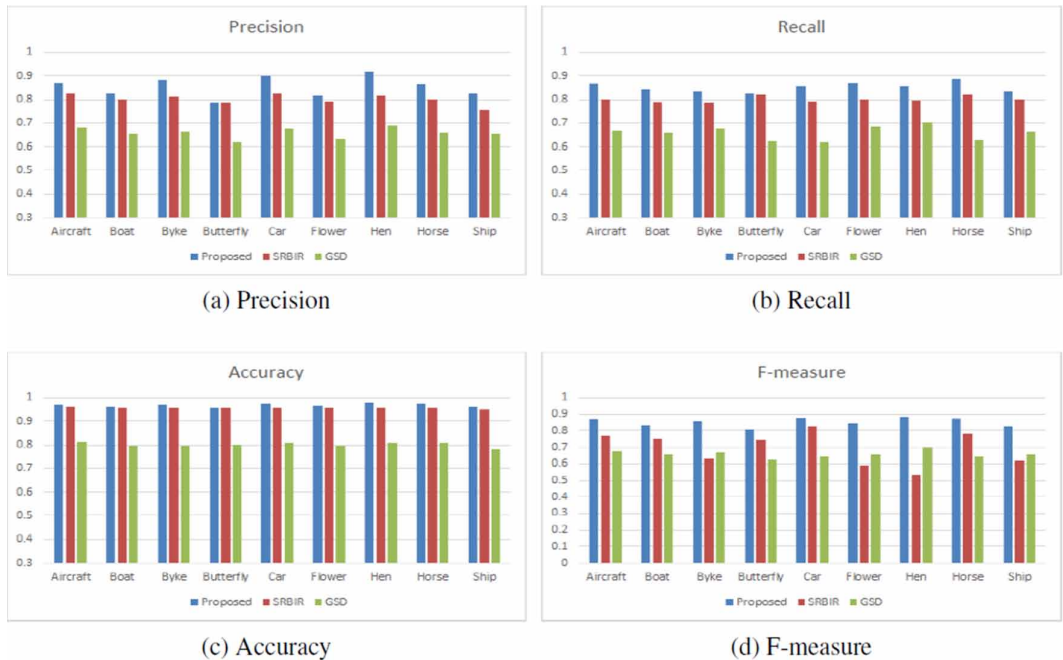
Table 3 illustrates the average precision, recall, accuracy and F-measure of the proposed with the state-of-art CBIR system on Olivia-2688 Dataset on the categories Coast, Forest, Highway, inside city, Mountain, Open country, Street and Tall building. The results indicate that the proposed system obtained better performance in term of these measures.

Figure 9 depicts the parametric comparison of different approaches on images categories of Olivia-2688 dataset listed in the table 3 In the Olivia-2688 dataset images are complex and visually less different like images of category inside city, street, tall building and highway are very less different in visual appearance. Result is compared with the existing approach SRBIR and GSD and it is analyzed that proposed model shows better results.

Table 2. Comparison of precision, recall, accuracy and F-measure obtained by the proposed method (PM) with existing content-based image retrieval systems GSD, Megha Agarwal and RP Maheshwari (2012), and SRBIR, P Manipoonchelvi and K Muneeswaran (2015) for GHIM 10K Dataset

	Precision			Recall			Accuracy			F-measure		
	SRBIR	GSD	Proposed	SRBIR	GSD	Proposed	SRBIR	GSD	Proposed	SRBIR	GSD	Proposed
Aircraft	82.5	68.2	86.9	80.0	66.8	86.6	96.1	81.2	97.2	76.9	67.5	86.8
Boat	80.1	65.3	82.2	78.8	65.8	84.2	95.7	79.4	96.4	75.0	65.5	83.2
Bike	81.2	66.3	88.2	78.4	67.6	83.4	95.8	79.8	97.0	63.2	66.9	85.7
Butterfly	78.4	61.9	78.6	82.0	62.4	82.2	95.7	80.2	95.7	74.4	62.2	80.4
Car	82.4	67.7	90.1	79.4	62.0	85.4	96.0	80.9	97.4	82.9	64.7	87.7
Flower	79.2	63.3	81.5	80.2	68.6	87.2	95.7	79.7	96.5	58.8	65.8	84.3
Hen	81.7	69.1	91.8	79.6	70.2	85.4	95.9	80.6	97.6	53.7	69.6	88.5
Horse	79.9	66.0	86.4	82.0	63.2	88.6	95.9	80.9	97.2	78.0	64.6	87.5
Ship	75.8	65.5	82.2	80.2	66.2	83.4	95.2	77.9	96.3	61.9	65.9	82.8

Figure 8. Graph comparison of PM with GSD and SRBIR on GHIM10k dataset

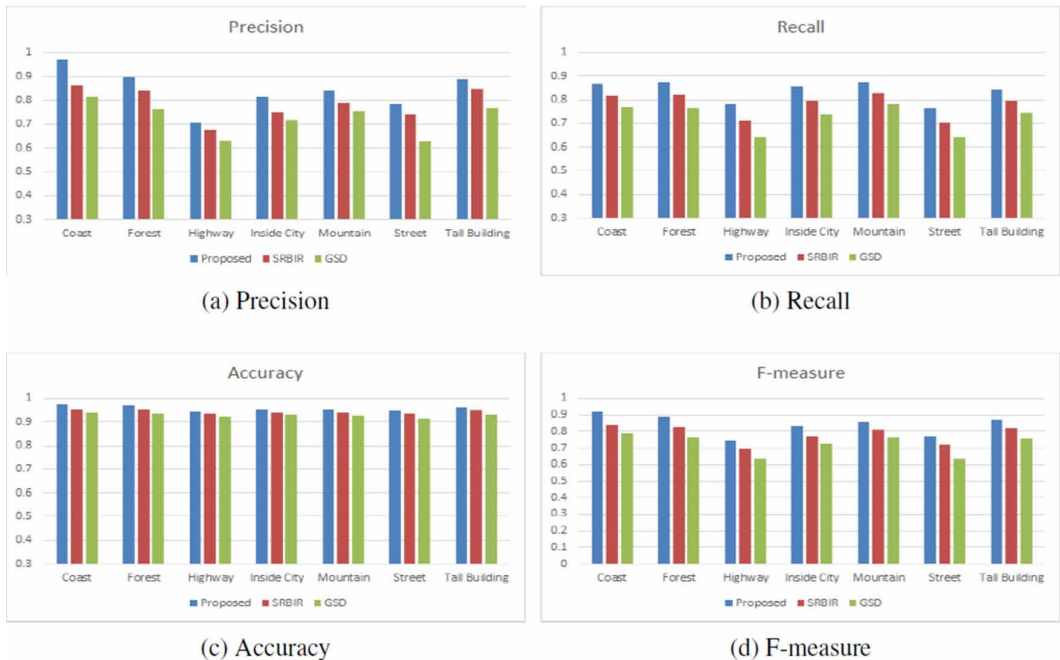


The proposed methodology presented in the paper is able to extract the foreground region effectively in the case of complex background image; the background of the image is suppressed successfully. In the technique after identifying the region of interest single feature is taken into considered. The main aim of the proposed technique is to enhance the accuracy and efficiency of the Query by Image Content (QBIC) application with using least feature. The experimental result is calculated on the parameters Precision, Recall, Accuracy and F-measure. It is found that the result of the proposed QBIR system is equivalent to the retrieval using multiple features.

Table 3. Comparison of precision, recall, accuracy and F-measure obtained by the proposed method (PM) with already existing content-based image retrieval systems GSD, Megha Agarwal and RP Maheshwari (2012), and SRBIR, P Manipoonchelvi and K Muneeswaran (2015) forOlivia-2688 Dataset

	Precision			Recall			Accuracy			F-measure		
	SRBIR	GSD	Proposed	SRBIR	GSD	Proposed	SRBIR	GSD	Proposed	SRBIR	GSD	Proposed
Coast	86.2	81.4	97.2	81.7	76.7	86.7	95.3	94.0	97.6	83.9	79.0	91.6
Forest	83.8	76.3	90.0	82.0	76.5	87.5	95.4	93.6	96.9	82.9	76.4	88.7
Highway	67.8	63.0	70.5	71.2	64.2	78.1	93.3	92.2	94.1	69.4	63.6	74.1
Inside City	74.9	71.8	81.7	79.5	73.7	85.4	94.0	93.0	95.6	77.2	72.8	83.5
Mountain	78.9	75.5	83.9	82.9	78.1	87.7	93.9	92.8	95.4	80.8	76.7	85.8
Street	74.0	62.8	78.2	70.2	64.0	76.4	93.5	91.3	94.6	72.1	63.4	77.3
Tall Building	84.7	77.0	88.8	79.5	74.4	84.6	94.9	93.1	96.1	82.0	75.7	86.6

Figure 9. Graph comparison of PM with GSD and SRBIR on Olivia-2688 dataset



Performance Evaluation With Different Features and Classifier

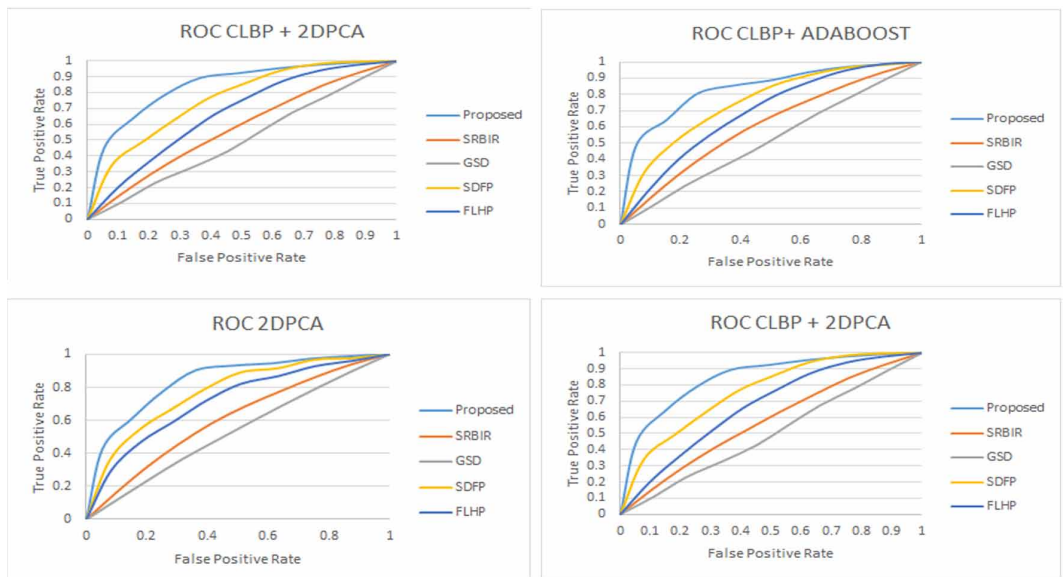
For a fair comparison, researchers have also compared the result of proposed method with different combination of features and classifier. On the bases of different parameters Precision, Recall, Accuracy and F-measure the result is compared in table 4 on Olivia-2688, Coral-10K and GHIM-10K datasets. The results illustrate that the proposed method with single feature shows better result.

From the experiment, it is analyzed that the proposed model out-perform. Figure 10 shows ROC comparison with the existing approach.

Table 4. Performance comparison of the proposed method other methods on performance measure precision, recall, accuracy and f-measure

	Precision			Recall			Accuracy			F-measure		
	Olivia-2688	Coral-10K	GHIM-10K	Olivia-2688	Coral-10K	GHIM-10K	Olivia-2688	Coral-10K	GHIM-10K	Olivia-2688	Coral-10K	GHIM-10K
CLBP + SVM	0.761	0.735	0.807	0.808	0.725	0.794	0.895	0.903	0.885	0.815	0.713	0.815
CLBP + AdaBoost	0.821	0.761	0.84	0.834	0.756	0.834	0.938	0.936	0.94	0.837	0.751	0.841
CLBP + 2DPCA	0.859	0.778	0.861	0.858	0.777	0.857	0.962	0.953	0.969	0.858	0.775	0.858
2DPCA	0.843	0.771	0.853	0.837	0.766	0.851	0.957	0.953	0.968	0.839	0.767	0.851

Figure 10. Comparison of ROC Curve obtained by the proposed method with existing content-based image retrieval systems GSD, Megha Agarwal and RP Maheshwari (2012), SRBIR, P Manipoonchelvi and K. Muneeswaran (2015), SDFP, Manke et al. (2018) and FLHP, M Singh et al. (2019) for GHIM 10K Dataset



CONCLUSION

In this paper, a saliency based efficient CBIR is proposed. Images having complex background are extracted from the image data corpus. ROI in complex background is extracted using well-known saliency approach SMD. For the fast recognition process, compact and efficient features are used known as 2DPCA. The experiment result of the process is compared on different benchmark datasets. Computed results are compared with the existing approach and found that the proposed approach in terms of retrieval with less features out performs. The proposed approach lags in the case where the ROI is unidentified or multiple different categories object present in the image. In the future work, multiple different categories objects can be identified and related images are extracted.

REFERENCES

- Agarwal, M., & Maheshwari, R. P. (2012). Á trous gradient structure descriptor for content based image retrieval. *International Journal of Multimedia Information Retrieval*, 1(2), 129–138. doi:10.1007/s13735-012-0005-5
- Chiang, C. C., Hung, Y. P., Yang, H., & Lee, G. C. (2009). Region-based image retrieval using color-size features of watershed regions. *Journal of Visual Communication and Image Representation*, 20(3), 167–177. doi:10.1016/j.jvcir.2009.01.001
- Datta, R., Li, J., & Wang, J. Z. (2009). Exploiting the human-machine gap in image recognition for designing CAPTCHAs. *IEEE Transactions on Information Forensics and Security*, 4(3), 504–518. doi:10.1109/TIFS.2009.2022709
- Frintrop, S., Rome, E., & Christensen, H. I. (2010). Computational visual attention systems and their cognitive foundations: A survey. *ACM Transactions on Applied Perception*, 7(1), 6. doi:10.1145/1658349.1658355
- Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., & Steele, D. (1995). Query by image and video content: The QBIC system. *Computer*, 28(9), 23–32. doi:10.1109/2.410146
- Gevers, T., & Smeulders, A. W. (2000). Pictoseek: Combining color and shape invariant features for image retrieval. *IEEE Transactions on Image Processing*, 9(1), 102–119. doi:10.1109/83.817602 PMID:18255376
- Liu, G.-H. (2017). *Image Rank Matchine*. Retrieved from <http://www.ci.gxnu.edu.cn/cbir/>
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91–110. doi:10.1023/B:VISI.0000029664.99615.94
- Lew, M. S., Sebe, N., Djeraba, C., & Jain, R. (2006). Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing Communications and Applications*, 2(1), 1–19. doi:10.1145/1126004.1126005
- Liu, Y., Zhang, D., Lu, G., & Ma, W. Y. (2007). A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40(1), 262–282. doi:10.1016/j.patcog.2006.04.045
- Manipoonchelvi, P., & Muneeswaran, K. (2015). Significant region-based image retrieval. *Signal, Image and Video Processing*, 9(8), 1795–1804. doi:10.1007/s11760-014-0657-0
- Manke, R., & Jalal, A. S. (2018). A Method for Salient Region Detection by Fusing Prior Information. In *Proceedings of 3rd International Conference on Internet of Things and Connected Technologies (ICIOTCT)* (pp. 26-27). doi:10.2139/ssrn.3167304
- Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1), 51–59. doi:10.1016/0031-3203(95)00067-4
- Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision*, 42(3), 145–175. doi:10.1023/A:1011139631724
- Papushoy, A., & Bors, A. G. (2015). Image retrieval based on query by saliency content. *Digital Signal Processing*, 36, 156–173. doi:10.1016/j.dsp.2014.09.005
- Peng, H., Li, B., Ling, H., Hu, W., Xiong, W., & Maybank, S. J. (2017). Salient object detection via structured matrix decomposition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4), 818–832. doi:10.1109/TPAMI.2016.2562626 PMID:28113696
- Singh, M., Jalal, A. S. S., Manke, R., & Khan, A. (2019). A Visual Saliency Detection Approach by Fusing Low-Level Priors With High-Level Priors. *International Journal of Computer Vision and Image Processing*, 9(3), 23–37. doi:10.4018/IJCVIP.2019070102
- Smeulders, A. W., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 1349–1380. doi:10.1109/34.895972
- Wolfe, J. M. (1998). Visual search. *Attention*, 1, 13–73. PMID:9503918

- Wang, X. Y., Yu, Y. J., & Yang, H. Y. (2011). An effective image retrieval scheme using color, texture and shape features. *Computer Standards & Interfaces*, 33(1), 59–68. doi:10.1016/j.csi.2010.03.004
- Wang, X. Y., Zhang, B. B., & Yang, H. Y. (2014). Content-based image retrieval by integrating color and texture features. *Multimedia Tools and Applications*, 68(3), 545–569. doi:10.1007/s11042-012-1055-7
- Wang, X., & Duan, H. (2017). Hierarchical visual attention model for saliency detection inspired by avian visual pathways. *IEEE/CAA Journal of Automatica Sinica*, 6(2), 540–552.
- Yang, J., Zhang, D., Frangi, A. F., & Yang, J. Y. (2004). Two-dimensional PCA: A new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1), 131–137. doi:10.1109/TPAMI.2004.1261097 PMID:15382693

Mohd Aamir Khan received his B. Tech in Information Technology and M.Tech. in Computer Engineering and Applications. He is currently a Ph.D. Research Scholar at the Department of Computer Engineering and Applications at GLA University Mathura. His research interest includes Information Retrieval, Content-Based Image Retrieval, Face Matching, Saliency and video processing. He is Member of IEEE, IEEE WIE, IEEE Young Professional and IEEE-Computer Society. He has attended the number of short-term courses/workshops/seminars organized by various esteemed organizations such as IIT Roorkee, ISI Bangalore, etc. He has published research papers in International Journals /Conferences of repute and participated in various International/National conferences.

Anand Singh Jalal received the M.Tech degree in Computer Science from Devi Ahilya Vishwavidyalaya, Indore, India. He received the Ph.D. in the area of Computer Vision from Indian Institute of Information Technology (IIIT), Allahabad, India. He has 20 years of teaching and research experience and currently, he is working as a Professor & Head of Department, Department of Computer Engineering and Applications, GLA University, Mathura, India. His research interests include Image Processing, Computer Vision and Pattern Recognition.