IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 3, 2021, accepted March 13, 2021, date of publication April 7, 2021, date of current version April 20, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3071581

Symmetric Image Contents Analysis and Retrieval Using Decimation, Pattern Analysis, Orientation, and Features Fusion

KHAWAJA TEHSEEN AHMED[1], SUMAIRA ASLAM[1], HUMAIRA AFZAL[1], SAJID IQBAL[1], ARIF MEHMOOD[2], AND GYU SANG CHOI[3]

[1]Department of Information Technology, Bahauddin Zakariya University, Multan 60800, Pakistan
[2]Department of CS and IT, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan
[3]Department of Information and Communication Engineering, Yeungnam University, Gyeongsan 38541, South Korea

Corresponding author: Gyu Sang Choi (castchoi@ynu.ac.kr)

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1A2C1006159) and MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2020-2016-0-00313) supervised by the IITP(Institute for Information & communications Technology Promotion).

ABSTRACT Image retrieval procedures locate interest points and consider the features as the visual property of an image in computer vision. These primitive features define visual attributes as local or global features for content based image retrieval. The visual attributes of an image, include spatial information, shape, texture, object and color, describe the image category. The present research contributes feature detector by performing non-max suppression after detecting edges and corners based on corner score and pixel derivation-based shapes on intensity-based interest points. Thereafter, interest point description applied on interest point features set by applying symmetric sampling to cascade matching produced by validating dense distributed receptive fields after estimating perifoveal receptive fields. Spatial color-based features vector are fused with retinal and color-based feature vector extracted after applying L2 normalization on spatially arranged color image. Dimensions are reduced by applying PCA on massive feature vectors produced after symmetric sampling and transmitted to bag-of-word in fused form for indexing and retrieval of images. Extensive experiments are performed on well-known benchmarks corel-1000, core-10000, caltech-101, image net, alot, coil, ftvl, 102-flowers and 17-flowers. In order to measure the competitiveness, we designed a comparison of proposed method with seven descriptors and detectors RGBLBP, LBP, surf, sift, DoG, HoG and MSER. The proposed method reports remarkable AP, AR, ARP, ARR, P&R, mAP and mAR rates in many categories of image datasets.

INDEX TERMS Interest point detection, image extraction, principal component coefficients, image descriptor, sliding window.

I. INTRODUCTION

Computer vision [1] is used to automatically extract, examine and understand useful information of query image and also retrieved images. Normally computer vision algorithms cover a numerous applications such as robots [2] and unmanned vehicle's navigation [3], video encoding on these unmanned vehicles [4], visual slam [5], video tracking [6], image retrieval [7], texture classification [8], object recognition [9], object categorization [10], image registration [11], face detection [12], face recognition [13] and also video shot retrieval [14]. According to computer vision, the features can be considered as the visual property of any image. Image processing is managed with different steps, and process of extracting features is major step for image processing, that works by extracting the local features as well as the global features of an image. Global features [15] refer to shape, spatial layout, texture and color where local features [16] of any image refer to interest point, region and spatial feature. Global features are not appropriate for partial matching and also arbitrary viewpoint consists of same object. To overcome this problem interest point detectors [17] are introduced that

The associate editor coordinating the review of this manuscript and approving it for publication was Wenming Cao.

extract local features of an image. Procedure of extracting the local features of an image based on the two steps process. First step is used to extract information of an image in form of interest point or key factors. Detectors are used for this feature extraction purpose. In second step neighborhoods of each interest point are further processed by descriptor.

Process of detecting the interest point is a dominant tool used infield of computer vision. Mostly three types of detectors are used for the detection of interest point, edge detectors [18], blob-based detectors [19] and corner-based detectors [20]. Edge detectors are essential part of computer vision system. An edge in any digital image may be the variation in texture, shade, color and light absorption, and all these changes in results may be used to determine the size, depth, surface properties and orientation of an image. To analyze the digital image, edge detection filters the irrelevant information for the selection of edge points. Complete process of continuous edge detection is time consuming and difficult because of noise that corrupt an image. Edge detectors cover the areas like military, robotics, pattern recognition system, meteorology, geography and medicine. Blob detectors are also crucial for computer vision systems. Blob-detectors aim to detect the region of different properties in digital image like color or brightness as compared to surrounding regions. Blob-detection based on two methods local extreme method and differential method. Local extreme calculates the local minima and local maxima of function, where differential method works on the basis of derivative values with respect to position. Motivators for the blob-detectors are such as: providing complementary information about any region that is not easily possible from corner detector and edge detectors, to obtain the interested region in object tracking and object recognition, peak detection in histogram, stereo matching, texture segmentation and ridge detection. Blob-detection can be computed by using the determinant of the hessian operator, the determinant of the hessian, the Laplacian of Gaussian, the hybrid Laplacian, the difference of Gaussian approach, and affine-adapted differential blob-detectors, Spatio-temporal blob-detector, the scale space blobs, grey level blob trees, grey level blob and maximally stable extremal regions (MSER) methods.

Corner based detector is an extreme curvature of the 2d image for its sharp curves of the edges and brightness change. These curvature points are widely used in virtual science reconstruction, image registration, camera calibration, motion estimation and mostly in computer vision task. Prominent purposes to use these corners are stereo matching, object tracking, visual odometer and optical flow calculation. Hence, it is imperative to efficiently implement these corner detection algorithms for the embedded systems which maintain power, resource constraints and stringent performance. Normally there are two corner-based detector methods in computer vision: based on gray scale image method and second one is image edges method. Image edge method is always required to encode the edges of an image, which depends upon the segmentations of image and extraction of edges from these segmentations. All this process is difficult to calculate because with a little bit change in segmentation complete operation will lead to a failure. Because of this reason edge method is not widely used for corner detection method. Second method is gray scale image, used for corner detection by calculating gradient and curvature. It avoids all the disadvantages faced in image edge method. Prominent operators of this gray scale image method are susan operator, harris operator, moravec operator and so on. In our proposed method we use Harris corner detection method[21] for the extraction of local features present in an image by analyzing principal intensity in neighborhood of a point. Second moment matrix is used for this complete process. Harris is not only used to detect corner but also used to detect noise and texture pattern.

In computer vision after completing the first step of key factors or interest point detection by using interest point detectors it is required to further process the neighborhoods of these detected key points. Computer vision system uses descriptors for this purpose. Description of neighborhoods of each key point is a challenge in local feature extraction. For this purpose, many descriptors are used but still it is a problem to select ideal pairs of neighborhoods and match this pair. Histogram is an admired descriptor of computer vision and is used to illustrate scale invariant feature transform (sift) key points [22]. Sift is used for the extraction of features of image and store these features in k-mean cluster. Sift and its variations are widely used in remote sensing for feature matching. Sift is computationally heavy and mathematically complicated as it is based on histogram that required to compute the gradient value of each pixel because of this its computational cost is also high. And also, it is highly affected by noise. The feature vectors extracted by surf [23] are almost indistinguishable to that of sift. It covers key points by grid and also further divides these grids into sub-grids. Histogram is created using these grids.

To overcome these problems we use fast retina keypoints (freak) descriptor [24] that is highly inspired by retina, part of human visual system. This method works in two steps, in first step sampling grid is used for the comparison of pixel intensities and then near the central point of pixel we generate the retinal level sampling grid by computing the differentiation of higher intensity points. Each intensity point is required to be less sensitive to noise in a smooth way.

II. RELATED WORK

There is a large collection of images collected from various sources which need to store and index in such a way that they can easily accessible. To access the relevant results from selected dataset, a content base image retrieval (cbir) structure is required. Cbir extracts only relevant features by using various techniques which are used for feature selection. These techniques simplify the image retrieval process and increase the retrieval ratio of relevant features. A model was proposed in [21] for the purpose of cbir and divided into three parts named feature discrimination feature extraction, and

feature selection. Gabor filter and 3d color histogram are used for feature extraction, genetic algorithm is used for the selection of features, and their deeply discrimination. Cbir systems returns similar images on the basis of query image from the large image sources. In [25],neural network system is used to improve the capabilities of cbir system. In order to retrieve the accurate image, k-nearest neighbors around the query image introduce a scheme of partial supervised learning. For efficient retrieval of images in cbir many new techniques are presented. Motif co-occurrence matrix (mcm) [26] is based on the idea of transformed images. In this technique complete image is transformed into grids of 2×2 pixels and both grids are replaced by the idea of scan which improves the retrieval performance. A technique for face recognition, represented by color based patterns is also introduced by [27]. This method discriminates both texture and color of any face image for feature recognition purpose. The method uses three databases for face recognition: color ferret, xm2vtsdb and cmu-pie. As similar to the feature detection, pedestrian detection is also used in content base image retrieval. For efficient performance of pedestrian detection [28] histogram of oriented gradient (hog) is used that works mostly in two steps. In first step sliding hog [29]window is used to scan the image and second step reduces the dimensions of features with principal component analysis. The algorithm is used to consider the image search or classification based on the fisher vectors or bag of words. In this way, regular patches of images are extracted in a dense way according to several scales.

In [30] some alternative choices are used to extract regular patches in a dense way such as edges, super pixels and zernike filters are used as feature detectors. It works better when extracted patches are around the edge rather than regions. Detection of 2d features are also supported in cbir, these features depends upon efficiency and invariance properties. Some 2d features detectors are binary robust invariant scalable keypoints, features of selected maximally stable external regions,, oriented binary robust independent elementary features, speeded up robust features, accelerated segment test, scale invariant feature transform and fast retina keypoints. Normally these feature detection methods can be tested by using long traffic scenes. Random sample consensus and brute-force matcher[31] are used to check how these methods are robust and to remove inconsistent matches ransack is used. Cbir system helps to extract the similar image of query image by comparing the local features and global features of query image with the images of database. Were, it is a technical job to find the semantic difference between the features of low- and high-level features of an image. One way to overcome this problem, to some extent, is to inspect combinations of mid-level relevant features. In order to extract the relevant images effectively bag-of-features model was used by [32]. Two combinations are defined hog and lbp [33], sift descriptor and also two integration models are suggested for this purpose including image-based integration and patch-based integration. The method shows better results for ambiguous objects and also for those images having noisy background.

With the advancement of technology, multimedia complexity is perceptibly increased and its retrieval process based on the content of multimedia. As we know, for the image retrieval CBIR system is used that extract the images from enormous stored databases in response of demanded image. In some scenario CBIR system restrict the extraction of extra image features that may confine competences of that system. To overcome the issue of that specific scenario [34] contributes to extract the important and robust features in response of QI from the huge databases and create a repository to store the extracted features in form of vectors like texture, shape and color. Metaheuristic algorithm was selected for the extraction of feature vectors from both QI image and databases. When an image is decided as a QI, distance metrics is selected to search the relevant images from database. Algorithm of neutrsophic clustering is selected for the extraction of RGB color, co-occurrence matrix of gray level for texture and method of canny edge for shapes and colors. For efficiency evaluation of designed method they calculated the value of precision-recall. Image classification based on texture has a significant role in various applications like object tracking, image segmentation, analysis of medical image, system of visual inspection etc. Many descriptors are designed for the extraction of texture such as LBP is designed for the description of local contrast and spatial structure of an image, LQP is introduced in advancement of LBP to improve the accuracy of classification but have some disadvantages like its input parameters are static and don't providing important binary patterns [35] Designed an improved version of LQP, known as improved LQP that divided the code of local quinary to four patterns based on local pattern, consists of local features. Evaluation of proposed approach based on data sets of Brodatz and Outex. An approach [36] for texture classification is introduced in which RTV analysis of features extraction is presented. Normally fractal geometry is used for the extraction of texture features but this method provides less accurate results for texture classification. They designed a mixed approach of co-occurrence matrix and RTV that calculate the first dimension results of extracted features by using threshold method and after that entropy results are calculated for co-occurrence matrix are calculated by RTV method. Outex and Broadatz were selected to calculate the comparative results. For creating relationship of centre pixels with its neighbor pixels on the basis of pixel direction [37] introduced a novel design in CBIR that consists of Local Hexadeca Patterns (LHdP), Local Octa Patterns (LOtP) and Direction Encoded Local Binary Pattern (DELBP) and calculate the derivative results of a pixel in its diagonal, horizontal and vertical direction. LBP method is used as feature vector and For calculating the derivative results every neighbor of a referenced pixel is considered, in result 256 directions are calculated for every referenced pixel, results are compared with different databases. In [38] presented a technique by the combination of color and texture information in transform and spatial domain jointly. The logic is used to extract the global texture, color information and local texture in two

frequency and spatial domains. Gaussian filter is selected to filter the image, extract the statistical features after calculating co-occurrence results of features in different directions. Quantized histogram is also applied to extract the global color texture in spatial domain and Gabor filter is applied to extract the features of local texture in frequency domain. improvement is done [39] using combined method of different feature extraction techniques by rotation invariant uniform local binary patterns (RULBP), color auto correlogram and local energy. MRMD filtering technique is used to extract the high frequency RULB and local energy features of gray image. Autocorrelogram is selected to extract the two dimensional features of a color space and Mahalanobis for similarity calculations of these features. In [40] designed an approach using two different types of texture descriptors to extract the features of an image from the database in response of QI. In first step, QI is examined to extract it color features and its texture features are extracted using two texture descriptors such as local binary and predefined patterns. After that distance criteria is measured to count the similarity of these features. As similar to feature detection some methods are also used for feature description. Lbp is one of them which are used for classification and feature description image processing. It is essential to combine the lbp selected from all the images of database for the description of color images. Whereas, when this combination is used, dimensions of pattern concatenated with lbp have an efficient increment. To overcome this dimensionality in [41] two schemas based on the decoder-based and adder are introduced. Color texture image databases: usptex, colored brodatz, mit-vistex, corel-1k, and 12 benchmarks natural scenes are used for experimental session. Content base image retrieval (cbir) system also supports scale and rotation invariant for texture classification. Gabor filter approach is used in [42] to collapse the responses of filter to orientation and scale of the texture. The process starts with two-dimensional matrix of features which is rearranged to calculate the homogeneous texture of the images. The sift invariance property is used to propose scale and rotation invariant features of the image. For this purpose brodatz texture album is used for experiment session. In annotation of medical images described a method such as surf (speeded up robust feature) descriptor [43]. By using this method most of the information gained by images, act as a backbone of content base image retrieval, especially for medical databases. The proposed methods are used for detection, extraction and description of image's features and match the features of required query image with the database images. Through this process false matching features are eliminated.

III. METHODOLOGY

First step involves the conversion of colored query image into black and white image by the function shown in equation 1.

$$\mathrm{I} = rgb2gray \tag{1}$$

This function is used to convert color images to grey scale images. The reason of using this grey scale image is to decrease the complexity level of an image like contrast, shades, shapes, edges, shadow, texture and so on with exploring colors. Mostly grey scale image is more valuable than its color image since, grey image is easier to deal with for the reason of its single color. Other reason for using this conversion is that colors are managed with human, but test images or query images are managed with machine or a system that is more difficult. Feature vectors are created after conversion of colored image into grey scale image by using process of interest point detector and descriptor algorithms as shown in figure 1.

A. INTERESTPOINT DETECTION

Process of detecting the interest point is performed for a system to compare two images on the basis of similarity as represented by figure 2. Corner detection or interest point detection both words can be used interchangeably. Interest point can be defined as a clear, well founded and has a well-defined position in an image. Normally a corner is an intersection of two edges and is detected by the variation in features of an image. Harris corner detector is based on the mathematical model used to extract the points having large intensity values in pattern recognition and classification model[44]. All this is done in a simple way by discerning a window of a patch that can generate a variety of variations after moving around [45]. This algorithm is ascertained by calculating the gradient value of each pixel.

If gradient values of both directions are considerable, then this pixel will be considered as a corner. After considerations of the corner score's difference according to direction, a mathematical relation is used to express a relation between the patches of window.

The steps of harris corner detector are as follows[46]:

First step is to calculate the both derivations of an image also the q.

$$\tilde{\mathfrak{K}}_p \leftrightarrow \frac{\beta\tilde{\mathfrak{K}}}{\beta p}\tilde{\mathfrak{K}}_q \leftrightarrow \frac{\beta\tilde{\mathfrak{K}}}{\beta q}\tilde{\mathfrak{K}}_p\tilde{\mathfrak{K}}_q \leftrightarrow \frac{\beta\tilde{\mathfrak{K}}}{\beta p}\frac{\beta I}{\beta q}$$

Second step of this process is to calculate the product of these derivatives for every pixel.

$$\tilde{\mathfrak{K}}_p^2 = \tilde{\mathfrak{K}}_p * \tilde{\mathfrak{K}}_p\tilde{\mathfrak{K}}^2 = \tilde{\mathfrak{K}}_q * \tilde{\mathfrak{K}}_q\tilde{\mathfrak{K}}_p\tilde{\mathfrak{K}}_q = \tilde{\mathfrak{K}}_p * \tilde{\mathfrak{K}}_q$$

Third step is the calculation of the covariance matrix l

$$L = \sum_{p,q} N(p,q)\begin{bmatrix} \tilde{\mathfrak{K}}_p^2 & \tilde{\mathfrak{K}}_p\tilde{\mathfrak{K}}_q \\ \tilde{\mathfrak{K}}_p\tilde{\mathfrak{K}}_q & \tilde{\mathfrak{K}}_q^2 \end{bmatrix}$$

This covariance matrix l is calculated by using a window function: $N(p, q)$

That is:

$$F(m,r) = \sum pqN(p,q)[\tilde{\mathfrak{K}}(p+m, q+r) - \tilde{\mathfrak{K}}(p,q)]^2 \tag{2}$$

The following equation can be explained as

- F presents the difference value between original and moved window.

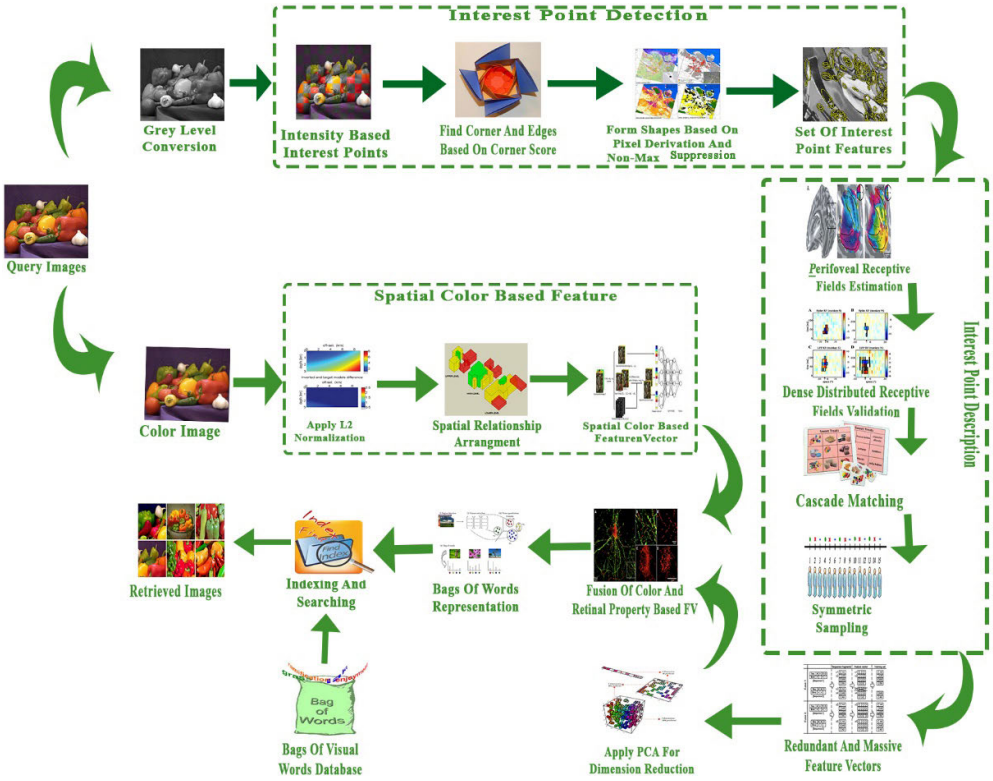


FIGURE 1. The proposed method.

FIGURE 2. The interest point detection.

- M is the movement of window directed in p direction.
- R is the movement of window directed in q direction.
- N(p, q) presents window placement toward the position (p, q).
- $\Re$ presents the image's intensity for position (p, q).
- $\Re$ (p+m, q+r) is a presentation of intensity of windows that are moved.
- $\Re$ (p, q) is the original intensity.

Simplified form of above equation is

$$F\,(m, r) = [mr]\,L \begin{bmatrix} m \\ r \end{bmatrix} \tag{3}$$

At the present point l is second moment matrix.

This step calculates the λ eigen value: if the value of λ is near about 0 then can be considered as an edge rather than the corner. To find out the corner value of λ must be greater than 0.

Hence to detect a corner in an image it is essential to compute all these steps.

B. INTEREST POINT DESCRIPTION

A binary descriptor for key point's comparison across images is considered in the proposed method. The proposed descriptor computes a flow of binary strings on the basis of comparison test of pairs of pixels intensities around the key points. Complete working of selected descriptor consists of inspiration by the human retina and its system. In first step, sampling grid is used for the comparison of pixel intensities which is circular. Furthermore, grid of retinal sampling is also generated by computing the differentiation of higher intensity points around the center. Each intensity point is required to be less sensitive to noise in a smooth way.

As similar to the brisk, local gradients are added for the orientation of key points while the proposed descriptor refers to those pairs which have symmetric fields around the center of receptive patch rather than the long pairs. To match the retinal model, kernels work effectively for sample points and Gaussian kernels calculate the standard deviation of all sampling point. Gaussian kernel according to size is changed

by using log-polar retinal pattern for improved performance. Normally accessible fields are overlapped for better performance. Now consider the interested fields that are $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ with concentration K_i in a way like:

$$\mathfrak{K}_X > \mathfrak{K}_Y, \quad \mathfrak{K}_Y > \mathfrak{K}_Z, \ \mathfrak{K}_X > \mathfrak{K}_Z$$

where $\mathfrak{K}_X > \mathfrak{K}_Z$ is affected for overlapped fields.

C. APPLY PCA FOR DIMENSION REDUCTION

Principal component analysis (pca) works on the basis of statistical techniques and belongs to linear transform. The method is considered as a dominant tool for pattern recognition and data analysis which is frequently used in image and signal processing as a method for data dimension reduction and data compression as well.

Pca for signal processing is a way to transformation of vector x, which is a complete set of s given input vectors containing the same length W formed in the vectors of s-dimensions $x = [x1, x2, \ldots .xs]^n$ into a Y vector according to

$$y = d(x - t_x) \tag{4}$$

The above equation (4) declares that for one input there are W values for each of the vector x. The vector t_x belongs to input variables and consists of mean values of all these input variables.

Image color reduction algorithms are lossy at some points, whereas, their results of these algorithms are still satisfactory for some working applications.

D. APPLY L2 NORMALIZATION

An algorithm is basic rule to create the segments of an image on the basis of information, gathered from the object and scenes of an image. Normally, the process of segmentation on the basis of object, scene and the label process of each pixel consequent to these objects especially in colored image are considered difficult. To overcome this difficulty, l2 normalization is selected in our work. Usually, the reflection of crossing point in blurred object can be described by the function of physical model, explained in following equation (5).

$$Q(b, \lambda) = \begin{cases} R_a(b) E_a(\lambda) \, Metal \\ R_a(b) E_a(\lambda) + R_O(b) E_h(\lambda) \, K.F. \end{cases} \tag{5}$$

where q is the presentation of photometric angles and for the measurement of wavelength λ is used.the symbol a denotes the reflection surface and o is the presentation reflection of body. *K.f.* Indicates dielectric object in an inhomogeneous way. Reflected surface a and the body reflection of k.f. Of inhomogeneous object can be connected just in one direction. $R_a(b)$ helps to determine the sensor values illumined for the objects that are dielectric. Value of this reflection for illumined objects is greater than zero and for the blurred object this value is zero. To count the effect of spectral power on illuminated surface in an image, this formula utilizes the n sensors with f spectral sensitivities; calculated value can be achieved by equation 6 and 7.

$$T_i(x, y) = \int_\lambda \mathcal{F}_i(\lambda) \mathcal{L}(\lambda) R(O) E(\lambda) j\lambda \tag{6}$$

$$= R(O) \int_\lambda \mathcal{F}_i(\lambda) \mathcal{L}(\lambda) E(\lambda) j\lambda \tag{7}$$

where o is theoccupation to consider the scene geometry in location (x,y).normally l1 normalization is considered for the coordinates(x,y) of color space.

IV. EXPERIMENTATION

A. DATASETS

Experiments are done on a variety of standard to achieve the high accuracy level of the designed method. Results obtained from the designed method are used to make comparison with predefined methods as: surf [47], sift [48], dog [49], hog [50], rgblbp [51], mser [52] and lbp [53]. The results are normally affected by features of image such as size, location, quality, color and cluttering. Experiments are performed on nine different databases including corel-1000[54], caltech-101 [55], image net [56], corel-10k [57], coil [58], alot [59], ftvl, 17-flowers [60]and 102-flowers [61]. Sample images from each category of each dataset are demonstrated below in figure 2.

1) COREL 1000

The corel-1000 dataset is generally selected to organize, retrieve and index the images for the classification and categorization task. Corel database contains the multiple categories of an image including beach, mountain, elephant, flowers, building, dinosaur, horse, bus, africa and food. This widely used dataset consists of 10 groups with each contain the 100seleted images of 348 × 256 pixels. Sample image selected from each category is shown in figure 5.

2) COREL 10,000 DATASET

Corel-10k dataset consists of 100 categories each with 100 images covering 10,000 (100 × 100) images. Size of each image is 128 × 84 in jpeg format. 15 categories are counted for experimentations including hospital, hospital, shining stars, text, texture, sun set, flags, flowers, planet, food, ketch, human texture, cars, trees, butterfly, and animals. Figure 5 shows the sample images of selected categories for corel 10k in which 15 images are presented, one image from each category. Arrangement of these categories are butterfly, flowers, flags, sun set, shining stars, text, ketch, texture, trees, hospital, food, planet, animals, human texture and cars.

3) CALTECH 101

Caltech-101 is considered as an ordinary dataset, used for image classification, recognition and categorization. The total numbers of images in the dataset are 9146. Fifteen categories were selected including chandeliers, airplane, tortoise, brain, ketch, bonsai, leopard, bikes, buddha, wristwatch, butterfly, things, face easy, face and ewer. Figure 5 presents 15 images

of caltech-101, which are selected from a total of 102 categories by considering the shape, pattern, color and texture of images. Each image is selected as a sample image which presents its category, named airplane, bonsai, chandeliers, things, ketch, leopard, wristwatch, bikes, tortoise, face easy, ewer, butterfly, face, brain and buddha.

4) ALOT

Alot dataset is a large collection of 250 rough textures of color images which are recorded for scientific purposes. Each class contains the 100 images and pixel size of each image is 384 × 235. Total number of images in alot dataset is 2500. 10 categories were selected for experimentation, downloaded from the alot repository including fruit-sprinkles, rope, red-coal, orange-parts, toy-marble, coins, corn, stones, ice-thick-layer and mandarin-pee. Figure 3 shows 10 images of alot dataset, selected for experimentation. We have selected 10 categories and one image from each category is presented in figure 5 as fruit-sprinkles, rope, red-coal, orange-parts, toy-marbles, coins, corn, stones, ice-thick-layer and mandarin-pee.

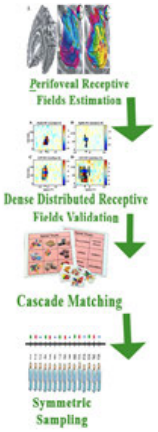


FIGURE 3. The interest point description.

5) COIL

Dataset of columbia object image library (coil-100) contains 100 categories. Each category has 72 images. Total images of coil are (100 × 72) 7,200. For experiments 15 categories were selected, named tomato, car, pink cup, cat, white cup, truck, statue, stick, soft edri, rolaids, mud pot, frog, herb-box, fancy feast and jug. Figure 5 is a presentation of 15 categories of columbia object image library (coil-100) database arranged as tomato, pink cup, car, stick, herb-box, white cup, statue, soft-edri, rolaids, jug, cat, truck, frog, mud pot and fancy feast.

6) FTVL

Ftvl dataset contains 15 categories with different sizes. Total images of ftvl dataset are 2612 with size 1024 × 768. The selected categories of ftvl dataset are aata_potato, watermelon, asterix_potato, taiti_lime, cashew, spanish_pear, granny_smith_apple, diamond_peach, fuji_apple, plum,

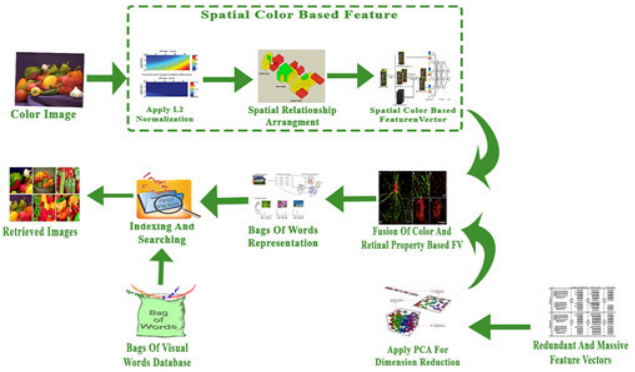


FIGURE 4. The L2 normalization and apllying PCA with BoW generation.

orange, onion, honneydew_melon, nectarine, nectarine and kiwi.

7) 17-FLOWERS

17-flowers dataset contains the 17 categories, each category consist of 80 images of (17 × 80) pixels. Selected categories of 17-flowers are iris, crocus, cowslip, colts foot, bluebell, tiger lily, lily valley, daisy, buttercup, dandelion, wind flower, snowdrop, tulip, pansy, sun flower, daffodils, and fritillary. In this dataset related images are retrieved on the basis of color.

8) 102-FLOWERS

The dataset of 102-flowers contains 102 categories of 8189 images of flowers. Each category consists of variant length 40 to 258 images. List of all categories is passion flower, daisy, water lily, Barberton, cyclamen, anthodium, water cress, lotus, frangipani, hibiscus, wall flower, clematis, rose, poinsettia and petunia. The images have large pose, scale and light variations. Figure 2 shows images for 102-flowers dataset which is the improved version of 17-flowers dataset. The images are selected for their scale, large pose and light variation. The arrangement of these images is as follows: passion flower, water lily, cyclamen, water cress, frangipani, wall flower, rose, petunia, poinsettia, clematis, hibiscus, lotus, anthodium, thorn apple, Barberton and daisy.

9) IMAGE NET SYNSET

The image net synset is a database consists of extensive images, used to retrieve index, annotate multimedia data and organize these image according to query image. Each image of word net describes multiple concepts or words phrase or word is called a synset. Word net consists of more than almost100,000 synset, almost dictated by nouns of (80,000+) concepts.

B. COMPARISON OF AVERAGE RETRIEVAL RECALL (ARR), AVERAGE RETRIEVAL PRECISION (ARP) AND PRECESION AND RECALL (P&R) RATES

Figure 6 highlights the graphical representation of the comparisons performed for dataset of 17-flowers. The designed

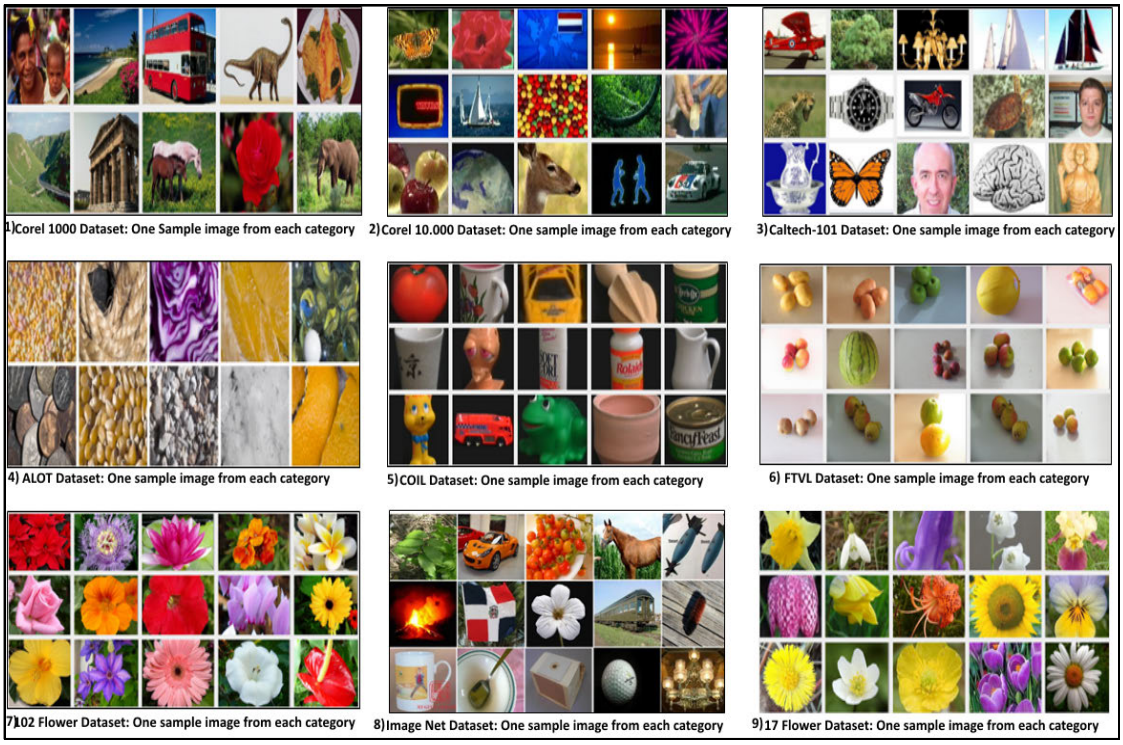


FIGURE 5. Sample images of all datasets.

method is compared with different standardized methods. Figure 6(a) presents graph of ARP for 17 different categories of flowers. The designed method shows ARP rate from 0.750 to 0.970. The presented method outperforms in all categories as compared to other methods. Figure 6(b) illustrates the ARR rate of the designed method. The introduced method shows marvelous results in most of the categories with rate of average recall from 0.133 to 0.102. Figure 6(c) is the presentation of precision and recall (PR) graph. PR graph presents average based precision through y-axis and average based recall x-axis. The proposed method reports remarkable PR rates.

Figure 7(a) is the representation of FTVL dataset, average retrieval precision (ARP) graph shows the comparison of calculated results of the precision. ARP rate of designed method shows significant results from 0.933 to 1. Figure 7(b) illustrates the ARR results to make a comparison among different methods. The designed method shows mostly tremendous rate with average recall from 0.100 to 0.108. Figure 7(c) is a presentation of precision and recall (PR) graph with considerable results. Image net results for average retrieval precision (ARP) are shown in figure 7(d), graph shows comparison results of the designed method with trending methods for image net dataset. ARP rate is from 0.320 to 0.405. The proposed method outperforms, comparative to other methods. Figure 7(e) illustrates the average retrieval recall (ARR) rate to compare different methods with the proposed method. The designed method frequently shows tremendous results in most of the categories with average recall rate from 0.258 to 0.315. Figure 7(f) is a presentation of PR graph for image net dataset. The proposed method reports improved results for PR.

Graph of ARP is presented in figure 8(a) for the coil dataset which compares the results of proposed method with defined methods. Arp graph rate is from about 0.890 to 0.830. Proposed method out performs to comparative methods. Figure 8(b) illustrates the average retrieval recall (ARR) rate to compare different methods with presented method. Proposed method shows tremendous results in most categories with average recall rate from 0.112 to 0.122. Figure 8(c) is a presentation of precision recall (PR) graph. Precision recall graphs basically are single value metrics which are calculated on the basis of features extracted by images. Order of extracted features to compare the images also matters. All features are indexed in a ranked sequence to form a database of features. By computing recall and precision for every feature in the indexed database, one can present a graphical presentation of precision recall graph.

Figure 8(d) presents average retrieval precision (ARP) graph for the caltech-101 that rate is from about 0.660 to 0.489. Proposed method better results as compared to all comparisons method. Figure 8(e) illustrates the average retrieval recall (ARR) rate to compare different methods with proposed method. Proposed shows average recall rate from 0.152 to 0.224. Figure 8(f) is a presentation of precision recall (PR) graph.

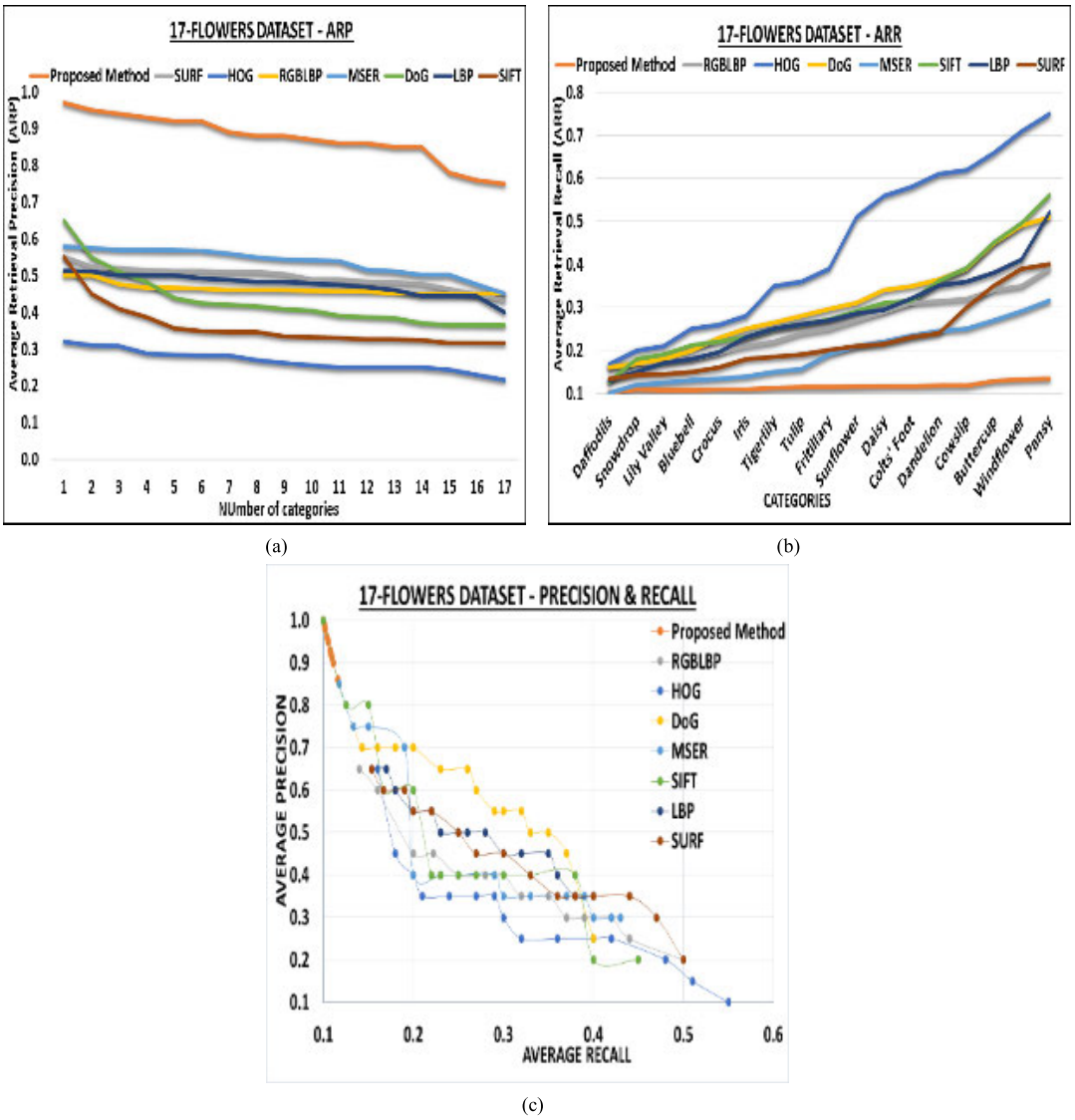


FIGURE 6. ARP, ARR and PR ratio for the 17-flowers database.

Alot dataset's results are shown from figure 8(g) that shows the comparison of proposed method with defined methods for average retrieval precision (ARP). Arp graph rate is from about 0.951 to 0.915. Proposed method out performs to comparative methods. Figure 8(h) illustrates the average retrieval recall (ARR) rate to compare different methods with proposed method. Proposed method presents marvelous results in almost all categories with average recall rate from 0.104 to 0.110. Figure 8(i) presents the calculated results for the precision and recall graph of Alot dataset.

Figure 9(a) presents average retrieval precision (ARP) graph that presents the comparative results of designed system with defined methods for dataset of corel-1000. Results of ARP graph are from 0.709 to 0.850. Proposed method outperforms to comparative methods. Figure 9(b) illustrates the average retrieval recall (ARR) rate to compare different methods with proposed method. Proposed shows high results in most categories with average recall rate from 0.118 to 0.143. Figure 9(c) is a presentation of precision recall (PR) graph. Precision recall graph basically are single value metrics which are calculated on the basis of features extracted by images. Order of extracted features to compare the images also matters. All features are indexed in a ranked sequence to form a database of features. By computing recall and precision for every feature in the indexed database, one can present a graphical presentation of precision-recall graph. Proposed method show better precision recall rate for Corel 1000 dataset.

Figure 10(a) presents average retrieval precision (ARP) graph of Corel 10,000 that compare the calculated results of our method with defined methods. Arp graph rate is

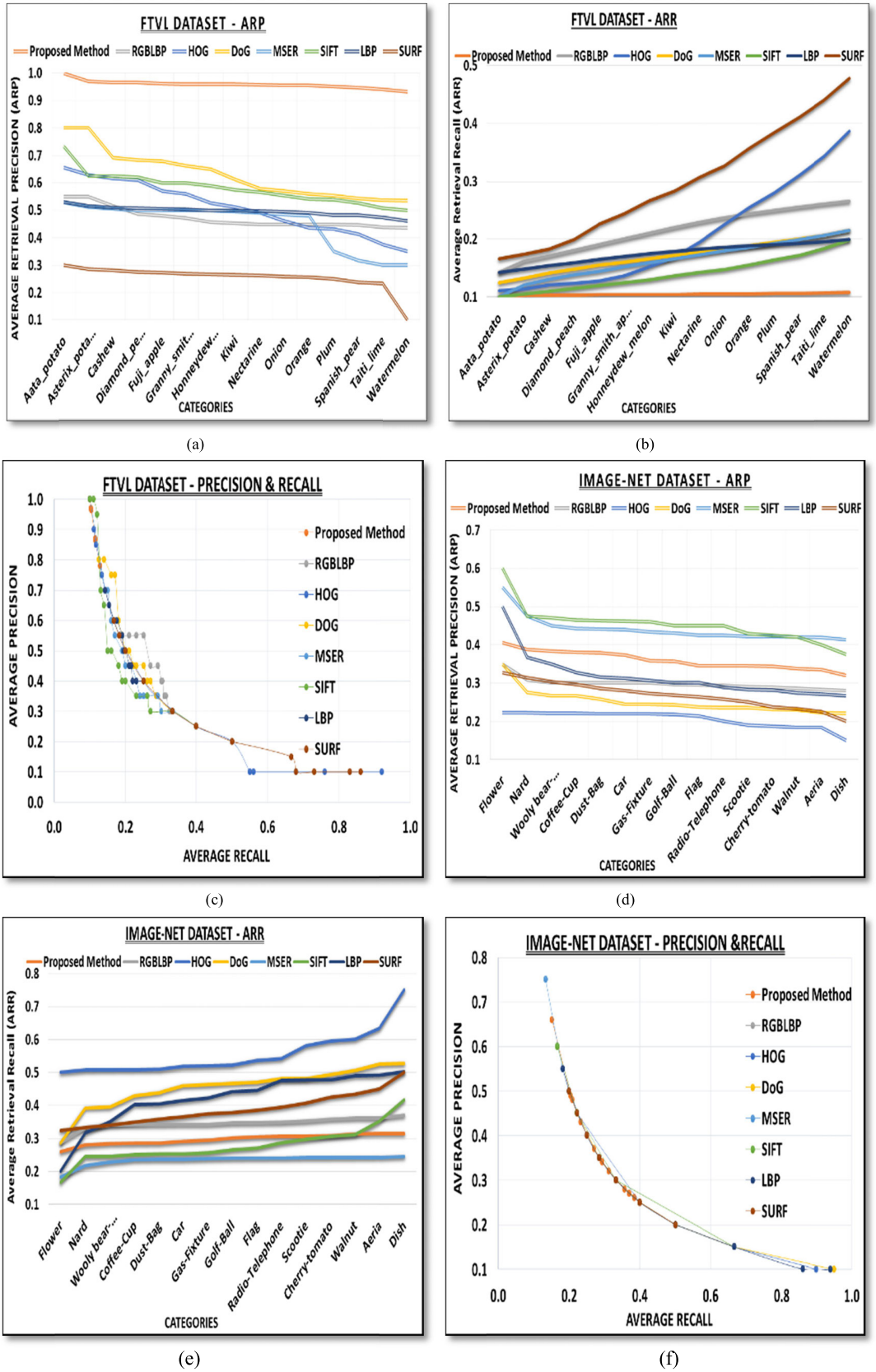


FIGURE 7. ARP, ARR and PR ratio for the FTVL and Imagenet datasets.

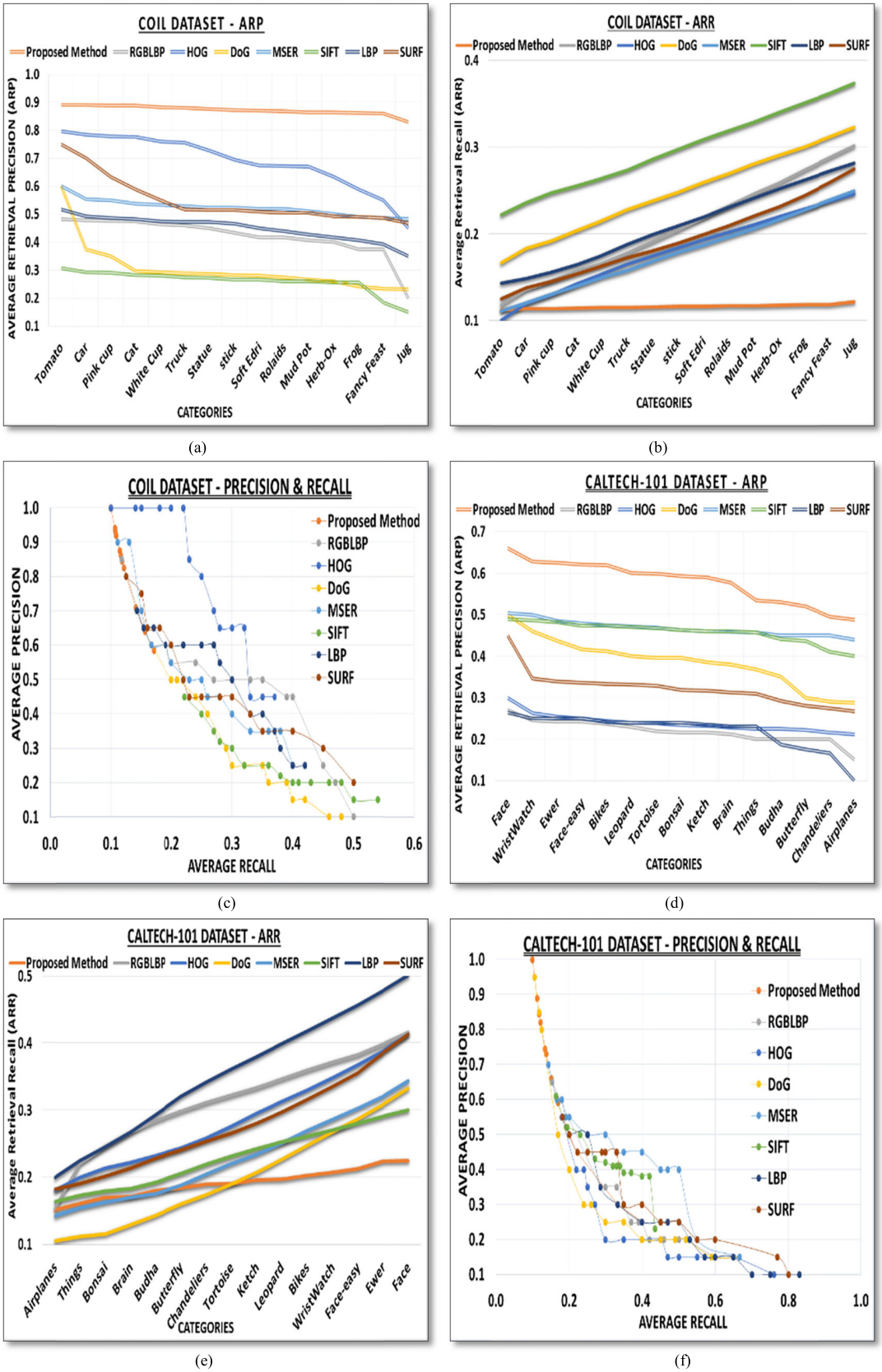
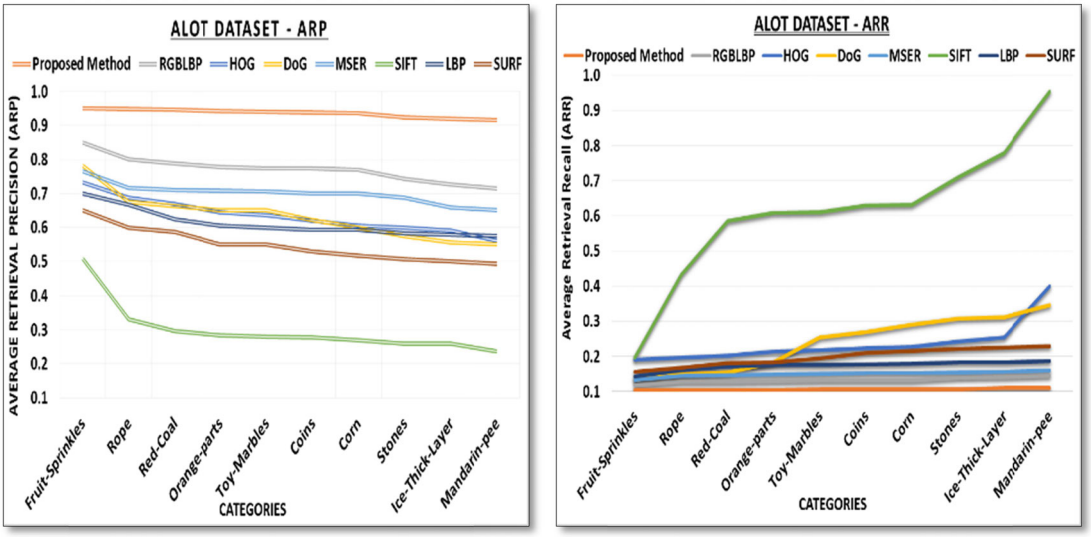
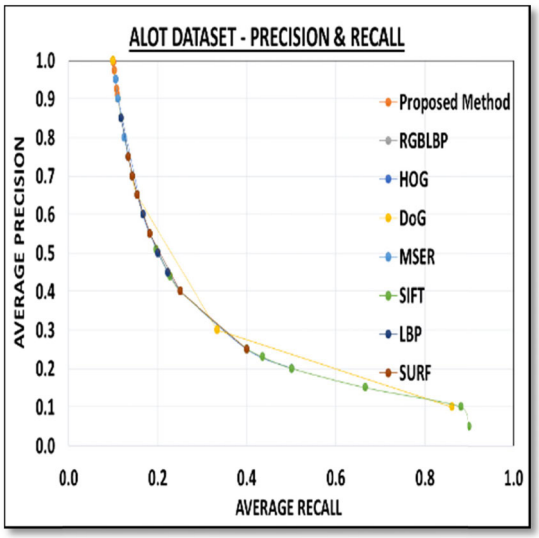


FIGURE 8. ARP, ARR and PR ratio for the coil, Caltech-101 and alot datasets.



(g) (h)



(i)

FIGURE 8. *(Continued.)* **ARP, ARR and PR ratio for the coil, Caltech-101 and alot datasets.**

from about 0.495 to 0.630. Proposed method outperforms to comparative methods. Figure 10(b) illustrates the average retrieval recall (ARR) rate to compare different methods with proposed method. Proposed shows tremendous results in most categories with average recall rate from 0.165 to 0.202. Precision recall (PR) graph is shown in figure 10(c). 102-flowers dataset's results shown from figure 10(d). Average retrieval precision (ARP) graph rate is from about 0.940 to 1. Proposed method outperform to comparative methods. Figure 10(e) illustrates the average retrieval recall (ARR) rate to compare different methods with presented method. Proposed shows marvelous results in most categories with average recall rate from 0.1 to 0.107. Figure 10(f) is a presentation of precision recall (PR) graph of 102 flowers dataset.

C. AVERAGE PRECISION (AP), AVERAGE RECALL (AR) FOR THE INTRODUCED METHOD VS. TRENDING METHODS

Figure 11 presents a graphical representation of derived results for AP rates of the designed method and comparative methods. The calculated rates present that our introduced system give valuable outcome in the majority of categories of 102-flowers and FTVL database. Figure 11(a) shows the comparative values of AP calculated by new method and different existing methods for the 102-flowersdataset, 11(b) shows the AP rates for the 17 -flowers, 11(c) shows the AP for the FTVL and 11(d) shows the results for the image net dataset.

Figure 12 presents a graphical representation of ar. Derived results of the proposed method are calculated to compare it with other existing method. The graphical representation

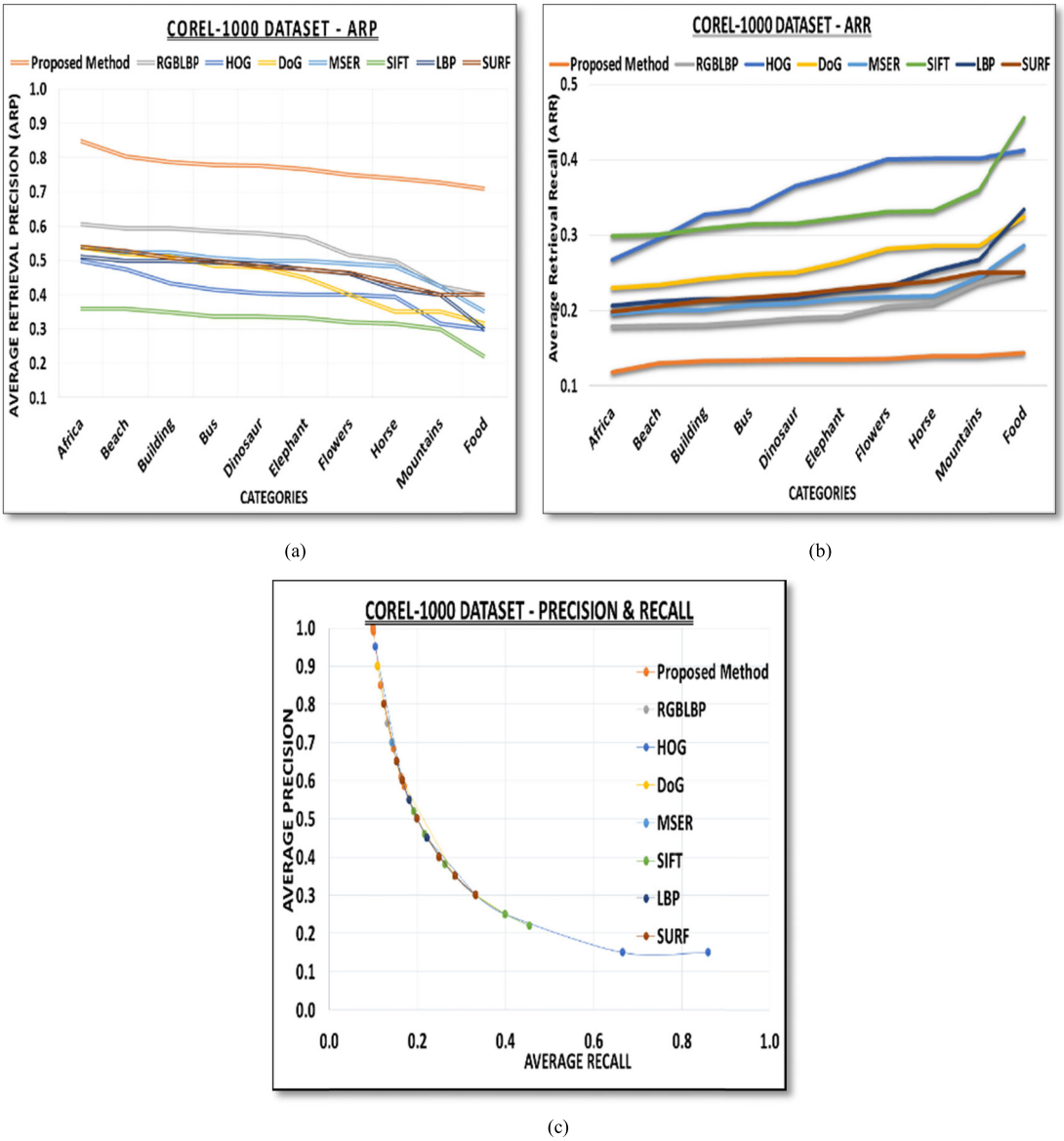


FIGURE 9. ARP, ARRand PR ratio for the Corel-1000 dataset.

shows that the introduced method gives outstanding performance in different categories of 102-flowers and FTVL database. Figure 12(a) shows results of ap achieved by new method and existing methods for the 102-flowersdataset, 12(b) shows the ap rates for the 17 -flowers, 12(c) presents the AP for the FTVL and 12(d) shows the results for the image net dataset.

Figure 13 presents the results of average precision (AP) counted from proposed method and also other defined method. The results presents that the introduced system shows valuable outcomes in various categories of coil, corel 1000 database. Figure 13(a) show the comparative values of AP of designed method and existing methods for the coil dataset, 13(b) shows the AP rates for the caltech-101, 13(c) shows the AP for the alot, 13(d) shows the results for the corel-1000 and 13(e) shows the results for the corel 10,000 dataset. Proposed method gives outstanding performance in various categories like horse, africa, flowers, dinosaur, dinosaur, and food. Designed system achieves AP value that range starts from 0.585 and end at 1.00 for selected dataset of corel 1k. Predefined system also shows improved outcome in a number of categories like hog show outstanding performance for bus and rgblbp for 'building' category. For the corel 10,000 dataset proposed method show better performance in different categories like text, sunset, food, trees and animals. Comparative methods are also good in some categories like difference of gaussian (DoG) performs well for butterfly, cars, ketch and planets. Dog normally is used

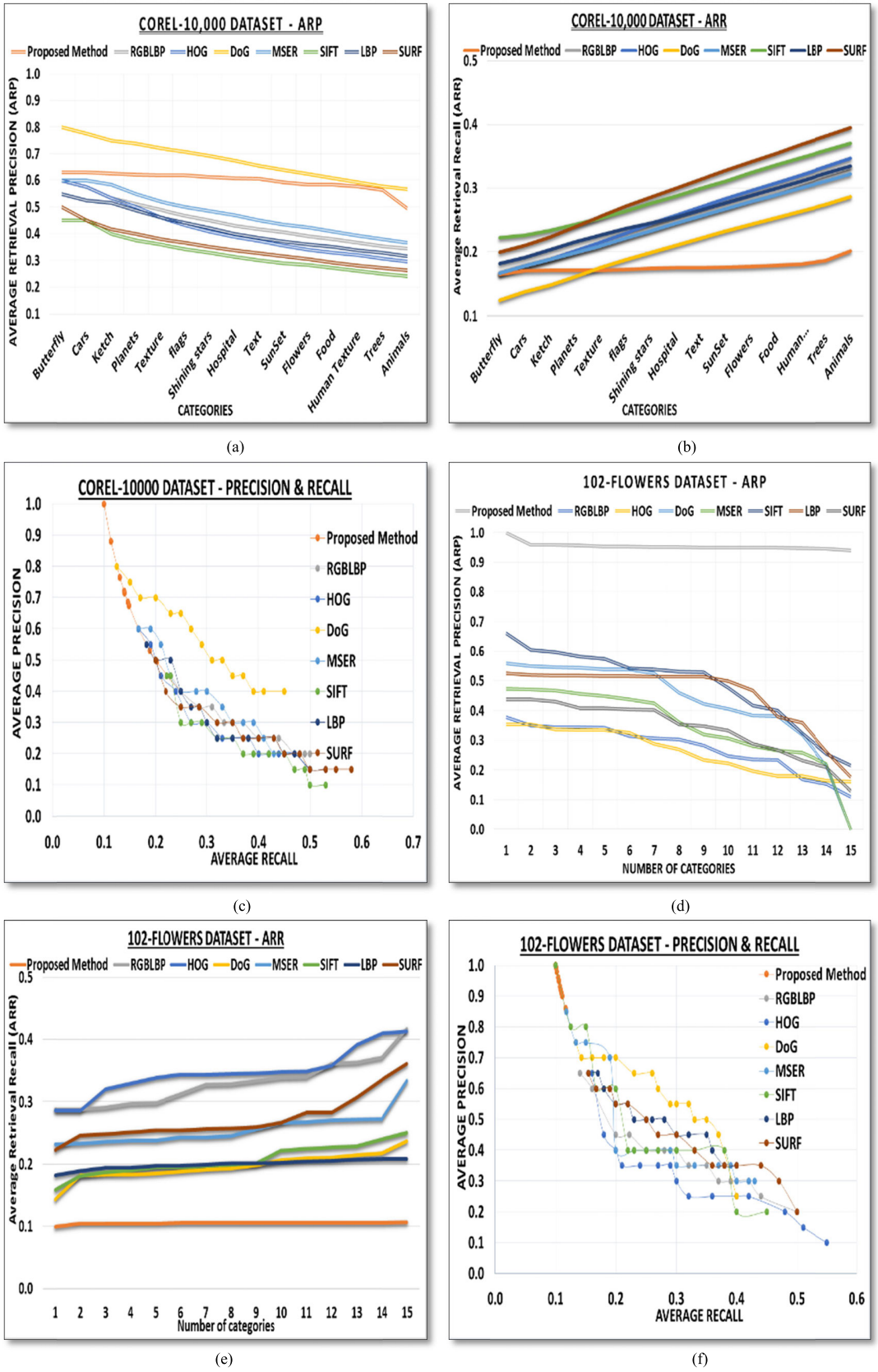


FIGURE 10. ARP, ARR and PR rates for the Corel-10000 and 102–Flowers dataset.

FIGURE 11. Average precision rates for the 102-Flowers 17 -Flowers, FTVL and Image net dataset.

for edge detection and also consider as a feature enhancement algorithm. Average precision rate that our proposed method calculates is approximate from 0.40 to 1.00.

Figure 14 presents a graphical representation of AR results, derived from the proposed method and made a comparison of these values with other introduced system. The obtained results present significant performance of introduced method in different categories of coil, corel 1000 database. Figure 14(a) show the comparative values of average recall of new method and existing methods for the coil dataset, 14(b) shows the AR rates for thecaltech-101, 14(c) shows the AR for the alot, 14(d) shows the results for the corel-1000 and 14(e) shows the results for the corel 10,000 dataset. Average recall value achieved from corel 1000 dataset starts from 0.100 and ends at 0.171 retrieved results by proposed method. The result is a presentation to show better recall rates of proposed method in different categories of corel 1000 database. AR extracted for corel 10,000 dataset in 14(f) is about from 0.100 to 0.250.

D. COMPARISON OF MEAN AVERAGE PRECISION (mAP) RATES

As mentioned above AP can be defined in image processing as the average of all queries and present this average as a single score. Suppose we are searching for an image of any category like flowers and for this purpose we provide an image of flower like lily as a required query image. In return of this selected query image in return we get a set of images in

TABLE 1. Corel-1000 average precision.

Corel-1000 Average Precision								
Categories	Proposed Method	MSER	LBP	RGBLBP	HOG	SURF	DoG	SIFT
Africa	0.85	0.35	0.30	0.40	0.30	0.40	0.35	0.22
Beach	0.65	0.50	0.50	0.45	0.50	0.40	0.35	0.38
Building	0.69	0.60	0.45	0.70	0.15	0.50	0.25	0.40
Bus	0.65	0.55	0.60	0.45	0.95	0.60	0.65	0.35
Dinosaur	1.00	0.70	0.65	0.90	0.60	0.80	0.80	0.25
Elephant	0.61	0.45	0.45	0.50	0.10	0.35	0.30	0.30
Flowers	0.99	0.40	0.55	0.70	0.30	0.65	0.70	0.46
Horse	1.00	0.65	0.60	0.75	0.25	0.30	0.90	0.52
Mountains	0.59	0.30	0.35	0.50	0.45	0.35	0.30	0.35
Food	0.85	0.40	0.30	0.60	0.45	0.30	0.60	0.25

TABLE 2. Caltech-101 average precision.

Caltech-101 Average Precision								
Categories	Proposed Method	RGBLBP	HOG	DoG	MSER	SIFT	LBP	SURF
Airplanes	0.66	0.15	0.30	0.50	0.50	0.41	0.10	0.45
Things	0.59	0.25	0.15	0.20	0.40	0.39	0.25	0.10
Bonsai	0.52	0.25	0.20	0.20	0.45	0.61	0.15	0.45
Brain	0.60	0.20	0.40	0.25	0.45	0.55	0.25	0.25
Budha	0.30	0.15	0.15	0.30	0.40	0.41	0.50	0.15
Butterfly	0.30	0.20	0.20	0.95	0.55	0.52	0.25	0.20
Chandeliers	0.45	0.20	0.15	0.30	0.45	0.52	0.15	0.45
Tortoise	0.75	0.35	0.15	0.25	0.50	0.38	0.20	0.50
Ketch	0.61	0.20	0.35	0.80	0.55	0.42	0.30	0.30
Leopard	1.00	0.35	0.20	0.85	0.60	0.39	0.50	0.25
Bikes	0.82	0.30	0.55	0.20	0.70	0.50	0.10	0.55
WristWatch	0.85	0.35	0.20	0.15	0.15	0.43	0.15	0.30
Face-easy	0.73	0.20	0.10	0.20	0.40	0.41	0.10	0.45
Ewer	0.25	0.25	0.15	0.40	0.60	0.23	0.35	0.45
Face	0.89	0.65	0.40	0.15	0.20	0.38	0.25	0.20

a ranked way that ordered from more similar to less similar. Sometimes it may be possible that not even a single image is similar. So, at this point we calculate the precision of every correct retrieved image and suppose we mark it as '1' and non-relevant images are marked as '0'. Then take average of this precision. In the same way map is an extended

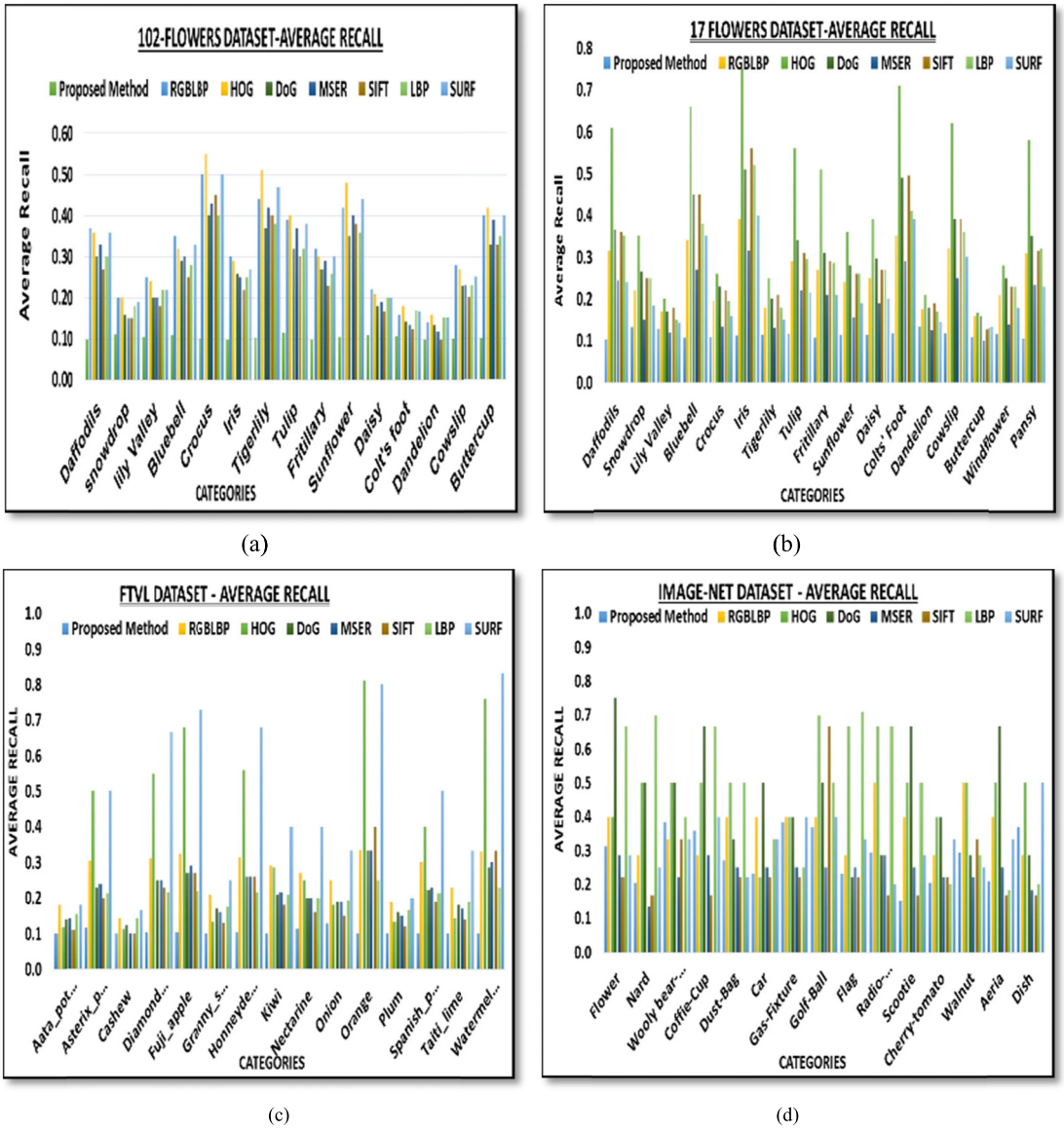


FIGURE 12. Average recall rates for the 102-Flowers 17 Flowers, FTVL and Image net dataset.

version of AP. In this process we take mean of all AP values for multiple query images and repeat the process of AP in same way.

Figure 15(a) shows mAP and 15(b) shows mAR rate for 102–flowers dataset that is extended version of 17–flowers dataset (a) shows that the mean map results of designed method are dominant than existing method that is about 0.956 15(b) shows the comparative results of mean average recall proposed method with different descriptors and has significantly better results of mean average recall (mar) almost around 0.105 as compared to the predefined methods. 15(c) shows that the mean average precision (map) rate calculated for 17–flowers dataset has significant improvement with respect to proportional method. Our introduced method gives the map rate almost around 0.918 for dataset of 17-flowers 15(d) shows the comparison of mean average recall and proves that proposed method has significantly better mAR results around 0.110 than the existing methods.

Mean average precision and recall calculated for ftvl dataset, 16(a) presents rate of mean average precision (mAP) of proposed method and compare it with different methods like surf, HoG, mser, lbp and DoG. Proposed method show mAP results around 0.961 that is remarkable than existing methods 16(b) presents the calculated rate of mean average recall and proposed method has better mean average recall (mAR) results around 0.105. Figure 16(c) shows mean average precision and 16(d) shows mean average recall rate for image net dataset (c) shows that the mean average

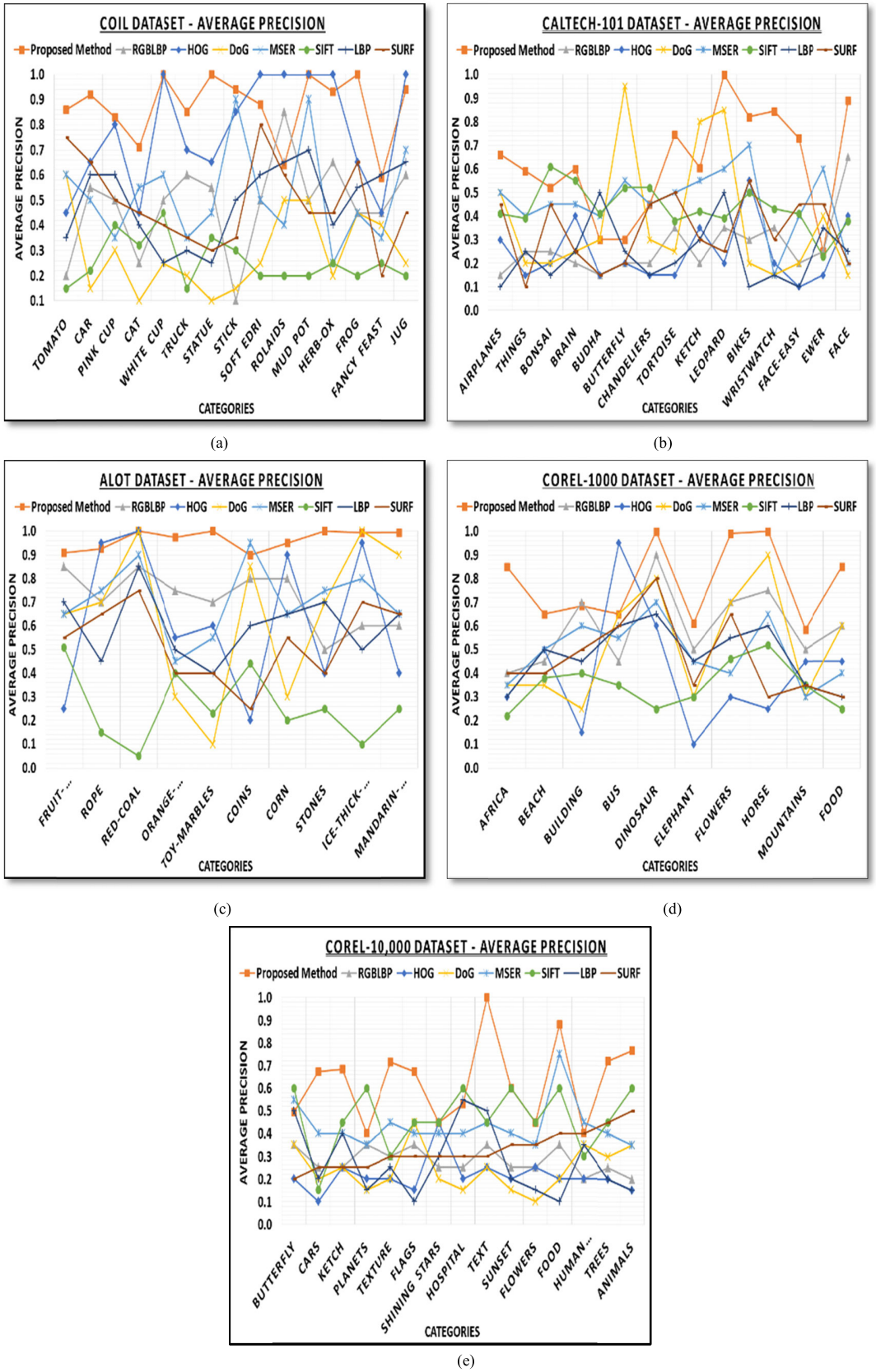


FIGURE 13. Average precision rates for the COIL, Caltech-101, alot, corel 1k and corel 10k dataset.

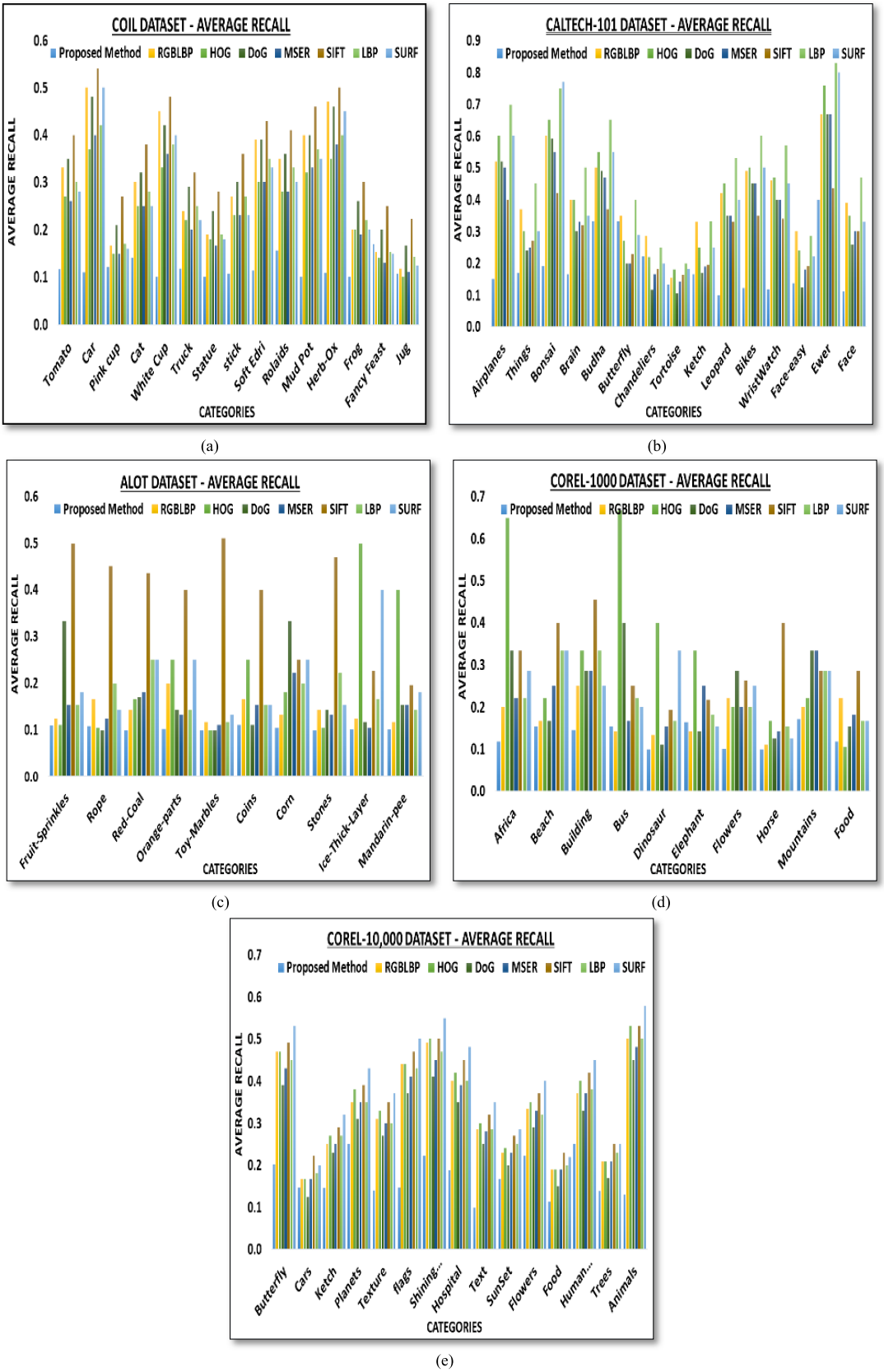


FIGURE 14. AP rates for the coil, Caltech-101, a lot, corel 1k and corel 10k dataset.

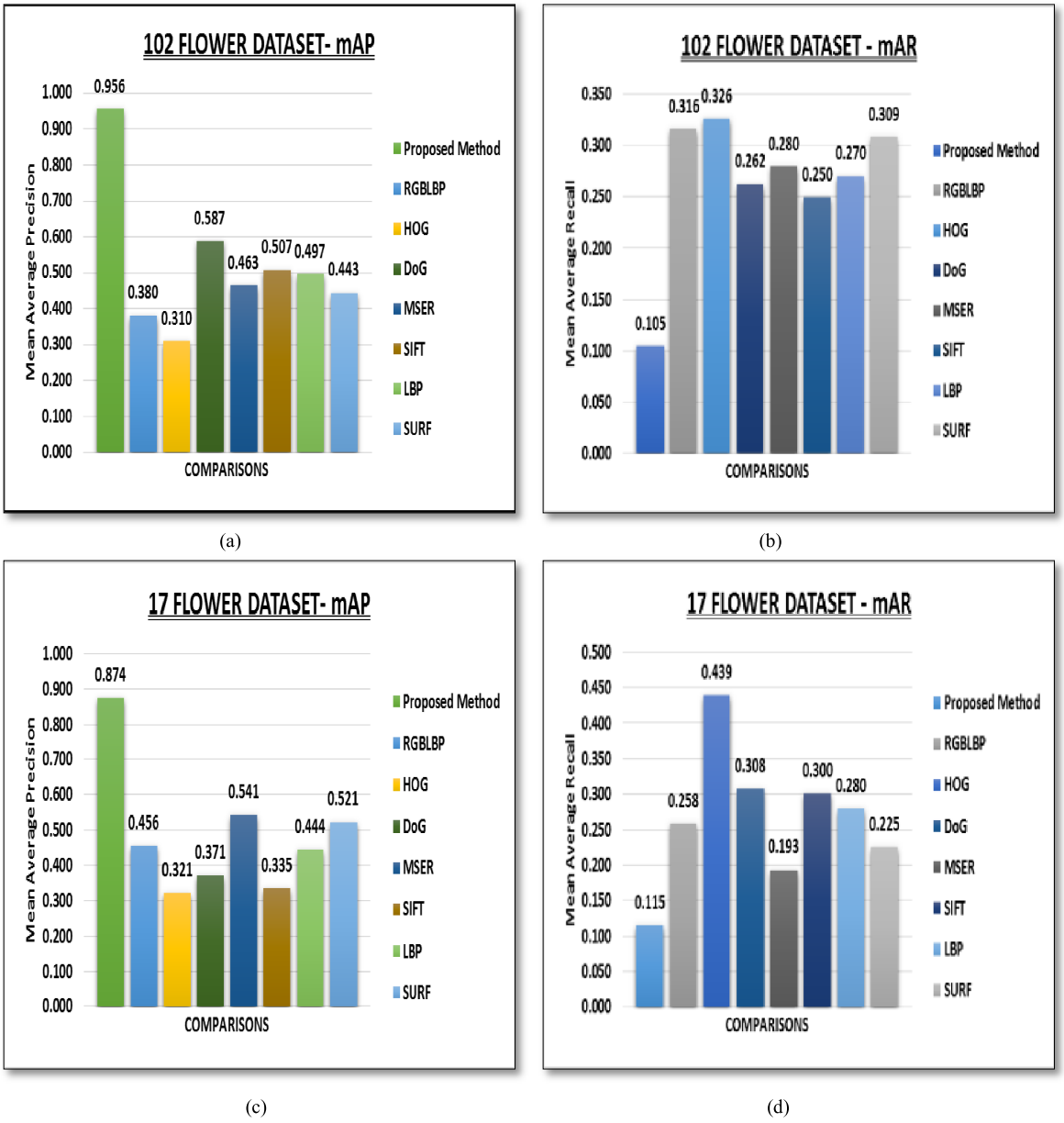


FIGURE 15. mAP and mAR ratio for the 17–flowers dataset and 102–flowers dataset.

precision (mAP) results for image net of designed method is higher than existing method that is about 0.956 (d) shows the comparative results of mar of other descriptor with introduced method and has remarkable mean average recall (mAR) results around 0.105.

Figure 17(a) shows mean average precision and 17(b) shows recall rate for coil dataset. 17(a) is a presentation of mean average precision (mAP) rate and shows the comparatively high value about 0.873 of proposed method. 17(b) is a presentation of mean average recall and present better result around 0.118 for proposed method. Figure 17(c) shows mean average precision and figure 17(d) mean average recall rate for difficult dataset caltech-101 and compare the calculated value of proposed method with different descriptors. 17(c) shows that the mean average precision (mAP) results of proposed method are marvelous according to existing method that is about 0.620. 17(d) shows the comparison of mean average recall and has higher mAR results almost0.191 than predefined methods.

Figure 18 presents of Map and mAR rate. 18(a) presents calculated mean average precision (map) for alot dataset and give higher calculated result of 0.965 than other comparative methods. A figure 18(b) presents mean average recall (mar) and has better results of mAR about 0.104 than

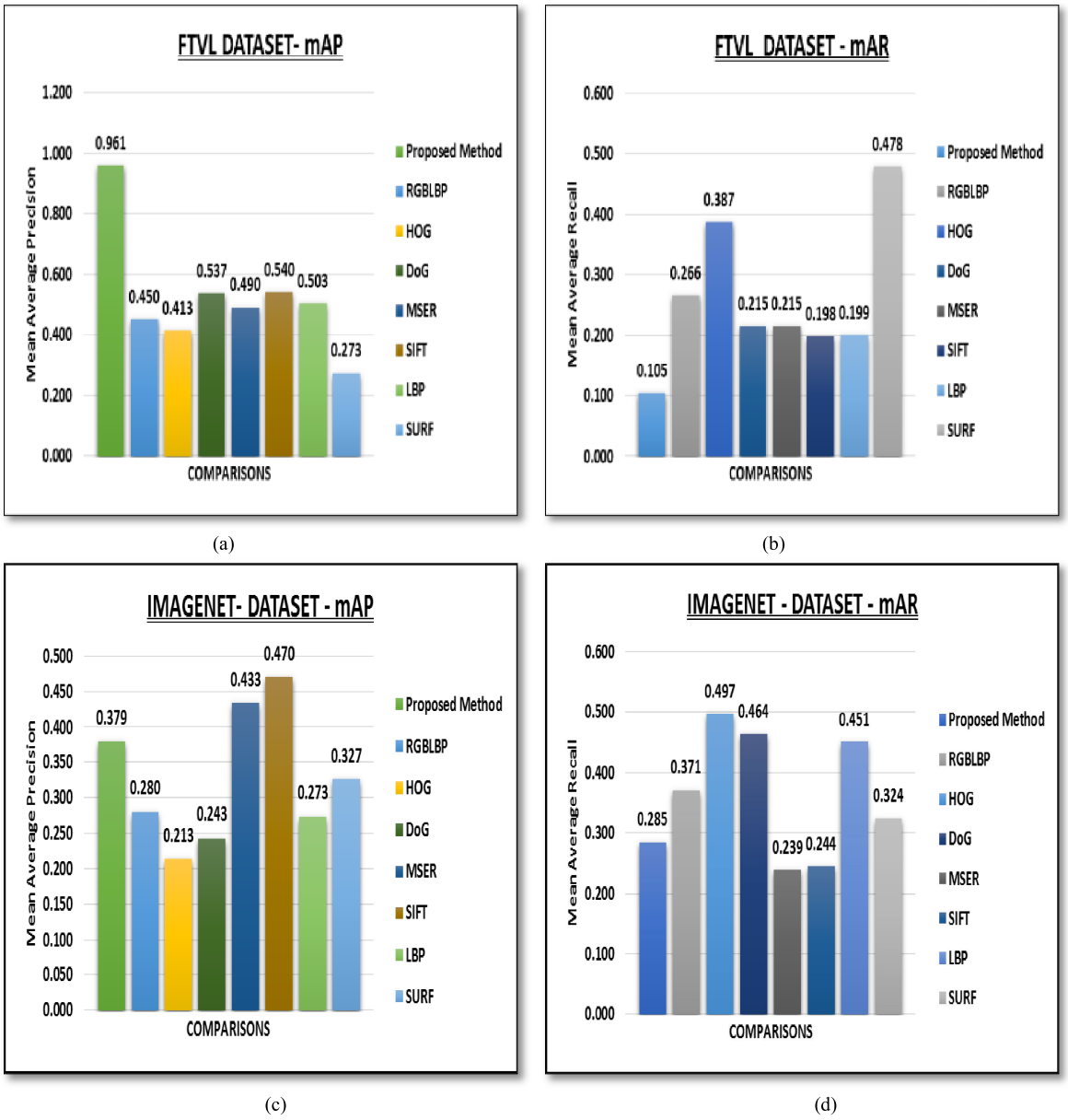


FIGURE 16. mAP and mAR ratio for the ftvl categories and image net dataset.

the comparative methods. Figure 18(c) and (d) shows mAR and Map results for corel-1000 of introduced method compared with predefined methods. 18(c) shows that the map results of designed method is remarkable than existing method that is about 0.787. 18(d) shows the comparison of mean average recall and has mAR results almost 0.132 than the competitive methods.

Figure 19(a) presents mAP rate for corel 10k dataset. Compare the calculated mean average precision with the calculated values of predefined methods and defines proposed method shows 0.629 values that is significantly better than other methods. 19(b) graphical presentation of mAR of corel 10k dataset and calculate the values around 0.171 than performs well.

E. RETRIEVE RESULTS

1) RETRIEVED RESULTS FOR THE DATASET OF COREL 1000 AND 10k

Figure20 presents 2 set of images retrieved on the basis of query image selected from corel 1000 dataset (a) is a presentation of 20 retrieved image results on the basis of query image dinosaur. (b) shows the results of retrieved image for elephant category used as a query image for the experiment of corel 1000 dataset. Figure 20(c)and (d) presents 2 set of images retrieved on the basis of query image selected from corel 10k dataset (c) is a presentation of 20 retrieved image results on the basis of query image flag query image that outperforms as compared to existing methods. (d) shows the results of retrieved images for car category of corel

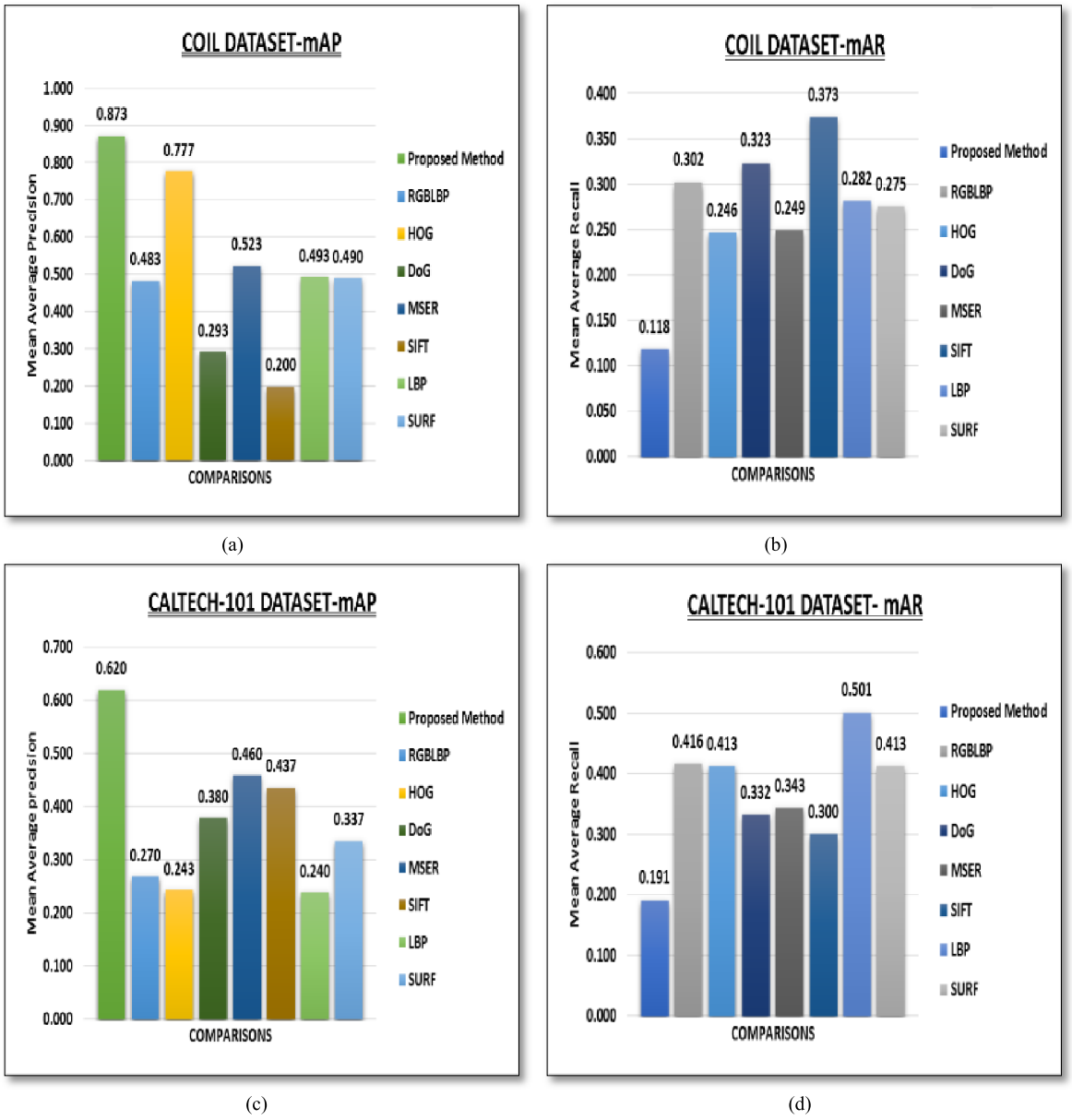


FIGURE 17. mAP and mAR ratio for the coil and caltech-101 dataset.

10k dataset. Proposed method retrieved 19 images when car is given as a query image.

2) RETRIEVED RESULTS FOR THE DATASET OF CALTECH-101 AND ALOT

Figure 21 presents images of dataset caltech-101 and figure 21(a) illustrates 20 retrieved image results on the basis of query image tortoise. Proposed method retrieved 17 images when tortoise is given as a query image that is better performance as compare with existing methods. 21(b) shows 20 images of leopard for experiment, show brilliant performance for leopard category of caltech-101 dataset. Figure 21(c) shows 20 images retrieved for rope category of alot dataset. These results provide more accuracy to the proposed method due to their brilliant and efficient result rate. 21(d) presents 20 images retrieved for fruit-sprinkles category of alot dataset. Selected category is important due to its texture format. Fruit sprinkles mostly used for decoration and in desserts based on textures like ice-cream, doughnuts or cupcakes.

3) RETRIEVED RESULTS FOR THE DATASET OF 102-FLOWERS AND COIL

Figure 22 presents 2 set of images retrieved on the basis of query image selected from 102–flowers dataset 22(a) and 22(b) is a presentation of 20 retrieved image

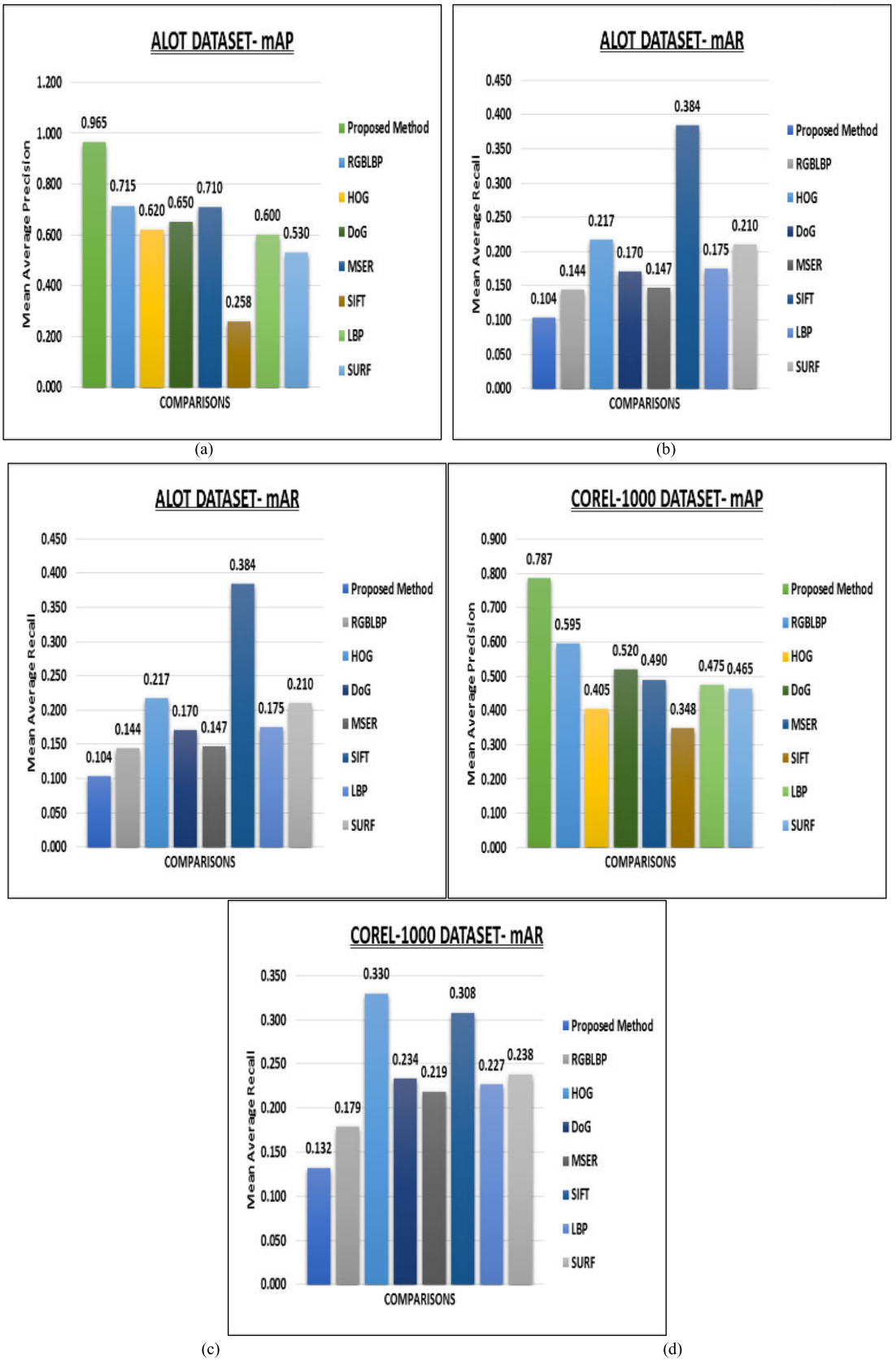
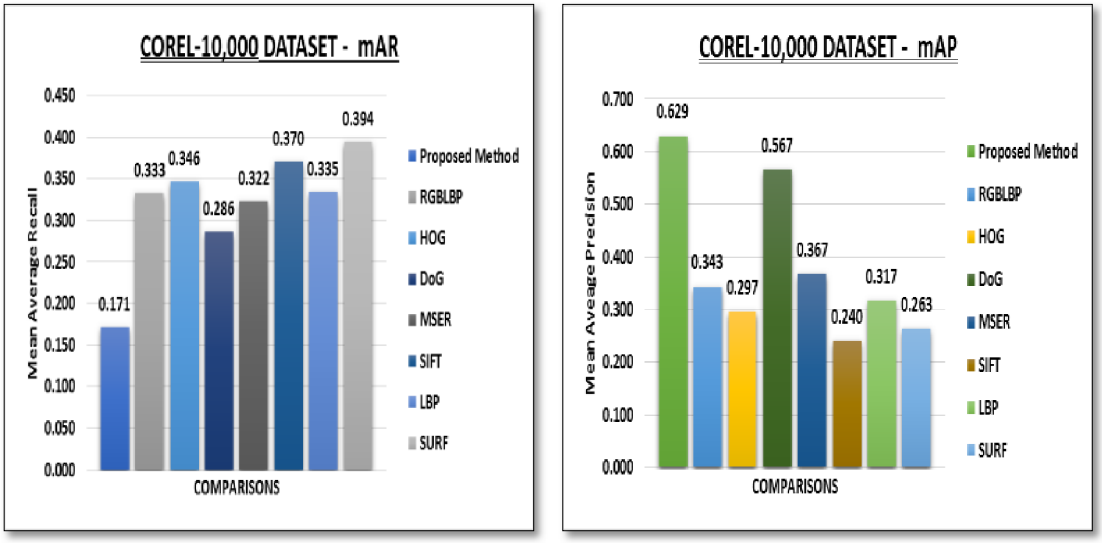
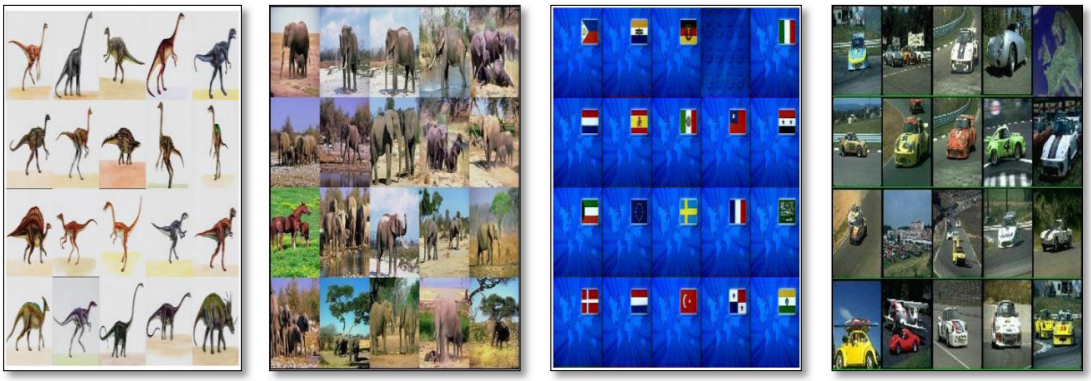


FIGURE 18. mAP and mAR ratio for the alot and corel-1000 dataset.



(a) (b)

FIGURE 19. mAP and mARfor corel-10,000 dataset.



(a) (b) (c) (d)

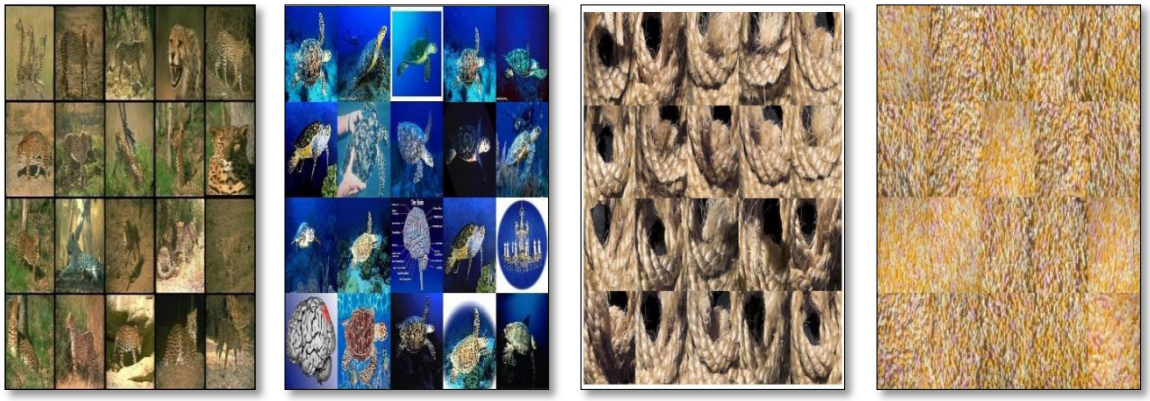
FIGURE 20. Retrieved results for the dataset of corel 1000 and corel 10k.

TABLE 3. mAP & mAR rate of selected datasets.

MAP & MAR RATE OF SELECTED DATASETS

Dataset	mAP	mAR
Corel 1000	0.787±0.4	0.132
Corel 10,000	0.629±0.3	0.171
Caltech-101	0.620±0.3	0.191
Image net	0.379±0.2	0.285
A lot	0.965±0.5	0.104
Coil	0.871±0.4	0.118
Ftvl fruits	0.961±0.5	0.105
17 -flowers	0.918±0.4	0.110
102 -flowers	0.956±0.5	0.105

results on the basis of query image. Figure shows brilliant performance for category of 102–flowers dataset. The proposed work is compared with state-of-the-art descriptors whose results are acceptable as benchmarks; the variants of these descriptors do not carry the potential milestones to be compared.



(a) (b) (c) (d)

FIGURE 21. Retrieved results for the dataset of caltech-101 and alot.

(a) (b) (c) (d)

FIGURE 22. Retrieved results for the dataset of 102-flowers.

(a) (b) (c) (d)

FIGURE 23. Retrieved results for the dataset of FTVL.

4) RETRIEVED RESULTS FOR THE DATASET OF FTVL AND 17-FLOWERS

Figure 23 (a) shows 20 images retrieved for cashew category of ftvl fruits dataset. These results provide more accuracy to the proposed method due to their brilliant and efficient result rate. 23(b) presents 20 images retrieved for granny smith apple category of ftvl fruits dataset. Figure 23 presents images from 17-flowers dataset 23(c) and 23(d) is a presentation of 20 retrieved image results on the basis of query image.

(a) (b)

FIGURE 24. Retrieved results for the dataset of the Imagenet.

TABLE 4. Feature length, feature extraction time, total computation time for all datasets.

17-Flowers			
No. of images	Feature length (words)	Features extraction time (sec.)	Total time (sec.)
Corel-1000	14	0.179	1.61
Corel-10000	12	0.169	1.51
Caltech-101	9	0.192	1.92
ALOT	9	0.155	1.68
COIL	7	0.112	1.34
FTVL	8	0.148	1.61
102-Flowers	8	0.175	1.89
ImageNet	16	0.199	2.25
17-Flowers	11	0.183	2.05

5) RETRIEVED RESULTS FOR THE DATASET OF THE IMAGE NET

Figure 24 displays images from image net dataset 24(a) is a presentation of 20 retrieved image results on the basis of query image. Proposed method retrieved 10 images when cherry tomato is given as a uery image that is better performance as compare with existing methods. Figure 24(b) shows 13 relevant images of bag query image for experiment out of 20 image.

V. CONCLUSION

The paper presents an image classification approach by extracting local and global features of images using well-known feature detector and descriptor. The proposed system extracts shape, spatial color features and objects for complex images. Average precision (AP), average rate (AR), ARP, ARR, P&R, map and mar rates are calculated to measure efficiency of the proposed method and evaluate with seven descriptors and detectors rgblbp, lbp, surf, sift, dog, hog and mser. Extensive experiments are performed on well-known benchmarks corel-1000, core-10000, caltech-101, image net, alot, coil, ftvl, 102-flowers and 17-flowers. The derived method reports remarkable results in many categories of image datasets. The derived method also shows considerable performance in diverse categories with complex background, texture and distinctive shape. An extension to the proposed work is to incorporate it with convolutional neural network (CNN). By embedding the vgg19, Resnet signatures with proposed algorithm will capable it to achieve the deep learning feature.

REFERENCES

[1] Y. Liu, J. Liu, and Y. Ke, "A detection and recognition system of pointer meters in substations based on computer vision," *Measurement*, vol. 152, Feb. 2020, Art. no. 107333.
[2] D. Scaramuzza, M. C. Achtelik, E. Kosmatopoulos, L. Doitsidis, F. Friedrich, A. Martinelli, M. W. Achtelik, M. Chli, S. Chatzichristofis, and L. Kneip, "Vision-controlled micro flying robots: From system design to autonomous navigation and mapping in GPS-denied environments," *IEEE Robot. Autom. Mag.*, vol. 21, no. 3, pp. 26–40, Sep. 2014.
[3] Y. Wang, S. Wang, and M. Tan, "Path generation of autonomous approach to a moving ship for unmanned vehicles," *IEEE Trans. Ind. Electron.*, vol. 62, no. 9, pp. 5619–5629, Sep. 2015.
[4] M. Bhaskaranand and J. D. Gibson, "Global motion assisted low complexity video encoding for UAV applications," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 1, pp. 139–150, Feb. 2015.
[5] A. Gil, O. M. Mozos, M. Ballesta, and O. Reinoso, "A comparative evaluation of interest point detectors and local descriptors for visual SLAM," *Mach. Vis. Appl.*, vol. 21, no. 6, pp. 905–920, Oct. 2010.
[6] L. Wang, T. Liu, G. Wang, K. L. Chan, and Q. Yang, "Video tracking using learned hierarchical features," *IEEE Trans. Image Process.*, vol. 24, no. 4, pp. 1424–1435, Apr. 2015.
[7] N. Neelima and E. S. Reddy, "An efficient multi object image retrieval system using multiple features and SVM," in *Advances in Signal Processing and Intelligent Recognition Systems*. New York, NY, USA: Springer, 2016, pp. 257–265.
[8] Y. Dong, D. Tao, X. Li, J. Ma, and J. Pu, "Texture classification and retrieval using shearlets and linear regression," *IEEE Trans. Cybern.*, vol. 45, no. 3, pp. 358–369, Mar. 2015.
[9] L. Bo, X. Ren, and D. Fox, "Unsupervised feature learning for RGB-D based object recognition," in *Experimental Robotics*. New York, NY, USA: Springer, 2013, pp. 387–402.
[10] A. Bergamo and L. Torresani, "Classemes and other classifier-based features for efficient object categorization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 10, pp. 1988–2001, Oct. 2014.
[11] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, Dec. 2015.
[12] R. Ranjan, V. M. Patel, and R. Chellappa, "HyperFace: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 1, pp. 121–135, Jan. 2019.
[13] P. Karczmarek, A. Kiersztyn, W. Pedrycz, and M. Dolecki, "An application of chain code-based local descriptor and its extension to face recognition," *Pattern Recognit.*, vol. 65, pp. 26–34, May 2017.
[14] M. Douze, J. Revaud, J. Verbeek, H. Jégou, and C. Schmid, "Circulant temporal encoding for video retrieval and temporal alignment," *Int. J. Comput. Vis.*, vol. 119, no. 3, pp. 291–306, Sep. 2016.
[15] J. M. Murphy, J. L. Moigne, and D. J. Harding, "Automatic image registration of multimodal remotely sensed data with global shearlet features," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1685–1704, Mar. 2016.

[16] C. Zhang, J. Sang, G. Zhu, and Q. Tian, "Bundled local features for image representation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 8, pp. 1719–1726, Aug. 2018.

[17] Y. Zhu, W. Chen, and G. Guo, "Evaluating spatiotemporal interest point features for depth-based action recognition," *Image Vis. Comput.*, vol. 32, no. 8, pp. 453–464, Aug. 2014.

[18] J. Canny, "A computational approach to edge detection," in *Readings in Computer Vision*, M. A. Fischler and O. Firschien, Eds. Los Altos, CA, USA: Morgan Kaufmann, 1987.

[19] A. Behnam, D. C. Wickramasinghe, M. A. A. Ghaffar, T. T. Vu, Y. H. Tang, and H. B. M. Isa, "Automated progress monitoring system for linear infrastructure projects using satellite remote sensing," *Autom. Construct.*, vol. 68, pp. 114–127, Aug. 2016.

[20] S. W. Teng, R. M. N. Sadat, and G. Lu, "Effective and efficient contour-based corner detectors," *Pattern Recognit.*, vol. 48, no. 7, pp. 2185–2197, Jul. 2015.

[21] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vis. Conf.*, 1988, pp. 10–5244.

[22] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Sep. 1999, pp. 1150–1157.

[23] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, 2008.

[24] A. Alahi, R. Ortiz, and P. Vandergheynst, "FREAK: Fast retina keypoint," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 510–517.

[25] A. Irtaza, M. A. Jaffar, E. Aleisa, and T.-S. Choi, "Embedding neural networks for semantic association in content based image retrieval," *Multimedia Tools Appl.*, vol. 72, no. 2, pp. 1911–1931, Sep. 2014.

[26] N. Jhanwar, S. Chaudhuri, G. Seetharaman, and B. Zavidovique, "Content based image retrieval using motif cooccurrence matrix," *Image Vis. Comput.*, vol. 22, no. 14, pp. 1211–1220, Dec. 2004.

[27] J. Y. Choi, K. N. Plataniotis, and Y. M. Ro, "Using colour local binary pattern features for face recognition," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 4541–4544.

[28] S. Pan, S. Sun, L. Yang, F. Duan, and A. Guan, "Content retrieval algorithm based on improved HOG," in *Proc. 3rd Int. Conf. Appl. Comput. Inf. Technol./2nd Int. Conf. Comput. Sci. Intell.*, Jul. 2015, pp. 438–441.

[29] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, pp. 886–893.

[30] A. Iscen, G. Tolias, P.-H. Gosselin, and H. Jegou, "A comparison of dense region detectors for image search and fine-grained classification," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2369–2381, Aug. 2015.

[31] Z. Song and R. Klette, "Robustness of point feature detection," in *Proc. Int. Conf. Comput. Anal. Images Patterns*, 2013, pp. 91–99.

[32] J. Yu, Z. Qin, T. Wan, and X. Zhang, "Feature integration analysis of bag-of-features model for image retrieval," *Neurocomputing*, vol. 120, pp. 355–364, Nov. 2013.

[33] T. Ojala, M. Pietikainen, and D. Harwood, "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions," in *Proc. 12th Int. Conf. Pattern Recognit.*, Oct. 1994, pp. 582–585.

[34] M. K. Alsmadi, "Content-based image retrieval using color, shape and texture descriptors and features," *Arabian J. Sci. Eng.*, vol. 45, no. 4, pp. 3317–3330, Apr. 2020.

[35] L. Armi and S. Fekri-Ershad, "Texture image classification based on improved local quinary patterns," *Multimedia Tools Appl.*, vol. 78, no. 14, pp. 18995–19018, Jul. 2019.

[36] F. Tajeri Pour, M. Rezaei, M. Saberi, and S. F. Ershad, "Texture classification approach based on combination of random threshold vector technique and co-occurrence matrixes," in *Proc. Int. Conf. Comput. Sci. Netw. Technol.*, Dec. 2011, pp. 2303–2306.

[37] U. S. N. Raju, K. S. Kumar, P. Haran, R. S. Boppana, and N. Kumar, "Content-based image retrieval using local texture features in distributed environment," *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 18, no. 1, Jan. 2020, Art. no. 1941001.

[38] N. T. Bani and S. Fekri-Ershad, "Content-based image retrieval based on combination of texture and colour information extracted in spatial and frequency domains," *Electron. Library*, vol. 37, no. 4, pp. 650–666, Aug. 2019.

[39] H.-H. Bu, N.-C. Kim, K.-W. Park, and S.-H. Kim, "Content-based image retrieval using combined texture and color features based on multi-resolution multi-direction filtering and color autocorrelogram," *J. Ambient Intell. Humanized Comput.*, vol. 1, pp. 1–9, Oct. 2019.

[40] N. Hor and S. Fekri-Ershad, "Image retrieval approach based on local texture information derived from predefined patterns and spatial domain information," 2019, *arXiv:1912.12978*. [Online]. Available: http://arxiv.org/abs/1912.12978

[41] S. R. Dubey, S. K. Singh, and R. K. Singh, "Multichannel decoded local binary patterns for content-based image retrieval," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 4018–4032, Sep. 2016.

[42] F. Riaz, A. Hassan, S. Rehman, and U. Qamar, "Texture classification using Rotation- and scale-invariant Gabor texture features," *IEEE Signal Process. Lett.*, vol. 20, no. 6, pp. 607–610, Jun. 2013.

[43] A. Jenitta and R. S. Ravindran, "Image retrieval based on local mesh vector co-occurrence pattern for medical diagnosis from MRI brain images," *J. Med. Syst.*, vol. 41, no. 10, pp. 1–10, Oct. 2017.

[44] A. Walha, A. Wali, and A. M. Alimi, "Video stabilization with moving object detecting and tracking for aerial video surveillance," *Multimedia Tools Appl.*, vol. 74, no. 17, pp. 6745–6767, Sep. 2015.

[45] W.-J. Wang, J.-W. Chang, S.-F. Haung, and R.-J. Wang, "Human posture recognition based on images captured by the Kinect sensor," *Int. J. Adv. Robot. Syst.*, vol. 13, no. 2, p. 54, Mar. 2016.

[46] P. Ram and S. Padmavathi, "Analysis of Harris corner detection for color images," in *Proc. Int. Conf. Signal Process., Commun., Power Embedded Syst. (SCOPES)*, Oct. 2016, pp. 405–410.

[47] A. Bundy, "Catalogue of artificial intelligence tools," in *Catalogue of Artificial Intelligence Tools*. New York, NY, USA: Springer, 1986, pp. 7–161.

[48] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Gray scale and rotation invariant texture classification with local binary patterns," in *Proc. Eur. Conf. Comput. Vis.*, 2000, pp. 404–420.

[49] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image Vis. Comput.*, vol. 22, no. 10, pp. 761–767, Sep. 2004.

[50] J. Li and J. Z. Wang, "Automatic linguistic indexing of pictures by a statistical modeling approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1075–1088, Sep. 2003.

[51] L. Fei-Fei, R. Fergus, and P. Perona, "Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories," in *Proc. Conf. Comput. Vis. Pattern Recognit. Workshop*, Jun./Jul. 2004, p. 178.

[52] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

[53] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLIcity: Semantics-sensitive integrated matching for picture libraries," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 9, pp. 947–963, Sep. 2001.

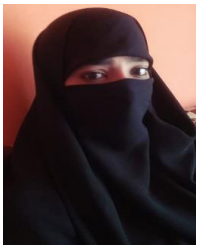
[54] S. A. Nene, S. K. Nayar, and H. Murase, "Columbia object image library (coil-100)," Tech. Rep., 1996.

[55] G. J. Burghouts and J.-M. Geusebroek, "Material-specific adaptation of color invariant features," *Pattern Recognit. Lett.*, vol. 30, no. 3, pp. 306–313, Feb. 2009.

[56] M.-E. Nilsback and A. Zisserman, "A visual vocabulary for flower classification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2006, pp. 1447–1454.

[57] M.-E. Nilsback and A. Zisserman, "Automated flower classification over a large number of classes," in *Proc. 6th Indian Conf. Comput. Vis., Graph. Image Process.*, Dec. 2008, pp. 722–729.

KHAWAJA TEHSEEN AHMED received the M.S. degree in computer science from Bahauddin Zakariya University, Multan, Pakistan, in 2010, and the Ph.D. degree in computer science from the University of Central Punjab (UCP), Lahore, Pakistan, in 2019. He is currently working as an Assistant Processor with the Department of Computer Science, Bahauddin Zakariya University. His research interests include computer vision, deep learning, pattern recognition, and machine learning.



SUMAIRA ASLAM received the M.S. degree from Bahuddin Zakriya University, Multan, Pakistan. Her research interests include computer vision, deep learning, pattern recognition, image enhancement, and machine learning.

HUMAIRA AFZAL received the M.Sc. degree (Hons.) in computer science from Bahauddin Zakariya University, Multan, Pakistan, in 1997, the M.Sc. degree in computer engineering from CASE, University of Engineering and Technology, Taxila, Pakistan, in 2010, and the Ph.D. degree in computer science from the School of Electrical Engineering and Computer Science, University of Bradford, U.K., in August 2014. She is currently working as an Assistant Professor with the Department of Computer Science, Bahauddin Zakariya University. Her research interests include MAC protocol design for cognitive radio networks, performance modeling, queuing theory, network security, software engineering, and sliding mode control.

SAJID IQBAL received the M.S. (CS) degree in computer science and the Ph.D. degree in computer science from Pakistan. He is currently working as an Assistant Professor with the Department of Computer Science, Bahauddin Zakariya University, Multan, Pakistan. His research interests include medical image processing and computer vision.

ARIF MEHMOOD received the Ph.D. degree from the Department of Information and Communication Engineering, Yeungnam University, South Korea, in November 2017. He is currently working as an Assistant Professor with the Department of Computer Science & IT, The Islamia University of Bahawalpur, Pakistan. His research interests include data mining, mainly working on AI and deep learning-based text mining, and data science management technologies.

GYU SANG CHOI received the Ph.D. degree from the Department of Computer Science and Engineering, Pennsylvania State University, University Park, PA, USA, in 2005. From 2006 to 2009, he was a Research Staff Member with the Samsung Advanced Institute of Technology (SAIT), Samsung Electronics. Since 2009, he has been a Faculty Member with the Department of Information and Communication, Yeungnam University, South Korea. His research interests include non-volatile memory and storage systems.