Human-centric Computing
and Information Sciences

RESEARCH

An effective image retrieval based on optimized genetic algorithm utilized a novel SVM-based convolutional neural network classifier

Mudhafar Jalil Jassim Ghrabat[1*†] , Guangzhi Ma[1*†], Ismail Yaqub Maolood[1], Shayem Saleh Alresheedi[1] and Zaid Ameen Abduljabbar[2]

*Correspondence: mudhafarjalil@hust.edu.cn; maguangzhi@hust.edu.cn
†Mudhafar Jalil Jassim Ghrabat and Guangzhi Ma contributed equally to this work
[1] School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China
Full list of author information is available at the end of the article

Abstract

Image retrieval is the process of retrieving images from a database. Certain algorithms have been used for traditional image retrieval. However, such retrieval involves certain limitations, such as manual image annotation, ineffective feature extraction, inability capability to handle complex queries, increased time required, and production of less accurate results. To overcome these issues, an effective image retrieval method is proposed in this study. This work intends to effectively retrieve images using a best feature extraction process. In the preprocessing of this study, a Gaussian filtering technique is used to remove the unwanted data present in the dataset. After preprocessing, feature extraction is applied to extract features, such as texture and color. Here, the texture feature is categorized as a gray level cooccurrence matrix, whereas the novel statistical and color features are considered image intensity-based color features. These features are clustered by k-means clustering for label formation. A modified genetic algorithm is used to optimize the features, and these features are classified using a novel SVM-based convolutional neural network (NSVMBCNN). Then, the performance is evaluated in terms of sensitivity, specificity, precision, recall, retrieval and recognition rate. The proposed feature extraction and modified genetic algorithm-based optimization technique outperforms existing techniques in experiments, with four different datasets used to test the proposed model. The performance of the proposed method is also better than those of the existing (RVM) regression vector machine, DSCOP, as well as the local directional order pattern (LDOP) and color co-occurrence feature + bit pattern feature (CCF + BPF) methods, in terms of the precision, recall, accuracy, sensitivity and specificity of the NSVMBCNN.

Keywords: Image retrieval, Pre-processing, Image feature, Gaussian filter, K-means clustering, Label information, CNN

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.*   (2019) 9:31

Page 2 of 29

Introduction

Recently, digital images have played a vital role in the depiction and dissemination of pictorial information. Therefore, huge sets of databases have been created and employed in various of applications, such as geographic information systems, identification of criminal activities, and multimedia encyclopedias. In several domains, such as biomedical and satellite imaging, digital images are suitable media for the storage and description of the temporal, spatial, physical and spectral components of the information. Image retrieval is mostly used in real time satellites for processing geotagged landmark images [1]. Useful information may be lost, resulting in miscategorized images. Consequently, image retrieval is useful in various fields. The retrieval process is mostly applied to computer vision recognition, pattern recognition, object visualization, and texture recognition [2]. Extracting similar images for a given query is a major problem for the image retrieval process that is known as mismatch error. Many real time applications have certain major impacts on this research, for instance, guiding mobile tourists and information inquiry services [3]. These applications include digital library managing, petroleum exploration, mineral identification, natural resource management, crop analysis, and medical imaging. For effective retrieval, feature extraction is performed. Picture retrieval can be determined using texture, color, and other features.

Feature extraction is the method in which the key points and the surrounding regions containing the raw pixel values are transformed into fewer domains that are single values [4]. Certain examples of feature extraction are the calculations of means, line thicknesses, standard deviations and distance between two pixels [5]. While applying several operators, such as measuring the size and angle of a noticed corner, there is a chance of extracting different features from similar regions. However, there are no requirements for all of the features extracted from a given area. Therefore, the unrelated features obtained are eliminated, which is called feature selection.

Next, in the feature selection approach, only a subset of the extracted features is selected. Finally, the image descriptor is a model where key point detection and feature extraction are performed together using computer vision and pattern recognition [6]. Henceforth, the input of the descriptor is specified as the raw pixel values of the descriptor, and the feature vector is an output value.

Depending on the features extracted and the training instances, effective classification is done through training the model. Mostly, domain experts only perform feature extraction and detection due to the high complexity. Thus, to automate these tasks, image descriptors have emerged, and the design part of this descriptor requires domain-expert interference. Furthermore, a huge number of training samples are usually required for the machine learning algorithms for better performance.

The key objectives of this work are focus distinguish query images from database images. Where we consider retrieving images from large databases by the use of the GLCM, NSF and IIBCF, which differ in regard to texture, color and other attributes. There are methods that try to increase the retrieval rate and reduce the retrieval time.

The processes in our work involving shape and color are as follows. We present an NSVMBCNN classifier that extracts features for the retrieval of images similar to the provided query. Moreover, we propose the GLCM and NSF are adopted for texture, and

IIBCF is used for color extraction. In this study, the k-means clustering technique is used for label formation. Finally, we measure the performance of the retrieval process.

To deal with significantly image retrieval data we present three main contributions in this work. Firstly, we propose an NSVMBCNN classifier and modified genetic algorithm have been utilized for image retrieval. The accuracy of the texture feature (GLCM, NSF) can be improved to the best results. The accuracy and the retrieval rate of the challenge are lower, the proposal method overcome. Secondly, we proposed using A NSVMBCNN classifier is implemented to overcome limitations in the previous methods to come out with efficient featured extraction method and classification process. Ultimately, in our work the analysis results show our proposed method is more efficient.

In the proposed system, three different types of features, including color features (IIBCF) and texture features, are integrated to develop the hybrid feature extraction mechanism. These features are more useful for obtaining clear information about the image. Then, the k-means algorithm is utilized to identify the groups or classes based on the extracted features, which is used to train the classifier. Finally, the NSVMBCNN classification technique is implemented to classify the labels with increased efficiency and accuracy. The remaining section of the paper is organized as follows: "Related works" section presents a brief review of the existing research work related to image retrieval in image processing. "Proposed work" section presents a detailed description of the proposed work. "Clustering technique" section illustrates a performance analysis and comparative analysis of proposed work. "The novel SVM based convolutional neural networkclassifier (NSVMBCNN)" section involves a short discussion of the conclusions of the proposed work and future work.

Related works

This section presents selected existing works related to the image retrieval process [7] presented a novel image descriptor approach for retrieving several image scenes. In this approach, groups of pixels were considered together for accomplishing a high level semantics. The proposed method is used to identify equal, similarity among images. Proposed CIBR method comprises of three sections in it namely, dataset partitioning, feature extracted, and training classifier. Features are extracted by scale-space extrema detection, key point localization, orientation assignment, key point descriptor. A significant advantage of this approach was that highly discriminative features were produced for image content labeling, and four datasets were used for the evaluations.

Al-sahaf in [6] elucidated an innovative approach emphasizing the key components of a GP-criptor and evaluating of the progress of the program. This new approach needed direct contact with the raw pixel values, so there was no requirement for providing pre-defined/extracted features with human interventions. The proposed method generates a feature vector by accepting input as an image. Texture images rotation variation is managed in the proposed method by using a new terminal set, function set.

Image descriptors developed by using two labeled instances for a class. Furthermore, for this GP-criptorri, the manual combination of the key points for designing the rotation in-variant was not required. Fewer training instances were required for the development of the descriptor, which was useful for applications containing limited labeled data. Thus, the training cost was reduced due to the operating capabilities.

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.*   (2019) 9:31

Page 4 of 29

Ciocca et al. [8] elucidated a new approach for predicting complications over textured images using a genetic programming framework. Nonlinear combinations were allowed in this approach, where the related image feature interactions in the complexity perceptions could be clarified. Different GP candidate solutions were statistically analyzed; thus, the roughness memorability, number of regions, and chroma measurements were evaluated in the gray level image, which was dominant over other traditional methods. In [9] presented a challenging task regarding the application of satellite images through retrieval of useful images from the database. A content-based image retrieval technique was used for searching the specific images in the database, which was related to the search query. The textural features were extracted using a statistical approach. Ultimately, a region-based comparison was made with a Bayesian classifier that classified images via a probabilistic approach.

Related works on feature extractions

In [10], local features composed of spatial domains were evaluated to demonstrate the implications of an comparing an image to a catalogue of an image. For overall feature extraction, the segmentation procedure was mandatory. The feature extraction is divided into two kinds: local features, global features. colour, shape, texture is local features utilized for detecting objects. Global features are employed for object classification. A new CBIR procedure was introduced to fuse the color and texture features. A color histogram (CH) was used to mine the color information. Texture features were mined by the use of the discrete wavelet transform (DWT) and edge histogram descriptor (EHD). A major limitation of this work was that this method did not use machine learning methodologies such as ANN .

Devi [11] clarified an innovative method for face recognition using an SVM classifier and a neural network. The SVM was used as a recognizer for the query image from the images returned by the retrieval process. Thus, the CBIR approach was more efficient than other conventional methods in terms of recognition rate and retrieval time.

Lu et al. [12] elucidated a novel face recognition approach by fusing shape and texture features. The shape feature representation, texture feature representation, fast faced retrieval by course to the find method techniques employed in the proposed method. The texture information of faces was extracted using modified Google Net. Next, these two features were fused and balanced with principal component analysis (PCA). To increase the efficiency, a coarse-to-fine search mechanism was employed for finding effectively similar objects. A scalability, the performance of facial attributes are improved. Wang [13] presented the challenges faced by content-based image retrieval (CBIR). A (CBIR) saves image database in the index file for joining it with the original image. Image descriptor is represented in vector format. The shape and the texture features were efficiently combined to improve the image retrieval rate. From the experimental results, significant improvements were achieved on well-known databases for the proposed method compared with other traditional methods.

Liu et al. [14] endorsed an image feature depiction method that was termed as a color difference histogram (CDH), and this method was utilized to describe image features for image retrieval. This method provided good discrimination power for color, texture, and shape features and spatial layout. And performance of the proposed method is analyzed

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.*   (2019) 9:31

Page 5 of 29

by precision, recall. Two features are merged to obtain good results. The authors in [15] proposed an efficient method for retrieval of images of faces. For this process, singular values and potential-field representation were employed. Image is divided for representing and applies rotation invariant property, shift invariant property, scale-invariant property.

For the feature extraction process, properties of singular values were used to develop a compact global feature for face-image representation. The use of singular values as rotation–shift–scale–invariant global features produced reasonable retrieval results. General images such as texture, trademark requires attention, improvement for image representation. Kumar et al. [16] surveyed the retrieval of 2D images with multiple dimensions and multi-modality images from sources including a diverse collection of medical data. A (CBIR) is image method applying visual features like color, texture, shape for retrieving images. Then, these features were used to reduce the two limitations of sensory gaps and semantic gaps. The Euclidean distance was used to measure the features in terms of vector and distance metrics. The (CBIR) structure was utilized practically in the health care domain.

Guo et al. [17] used error diffusion block truncation coding features in content-based image retrieval. An extension of the EDBTC image retrieval system was used to index video by considering the video as a sequence of images. Then CHF, Bit pattern histogram identifies match among query and targeted image. This methodology achieved high accuracy, and the features were added into EDBT indexing. A disadvantage of this work was high complexity. An Image index for CIBR, Image compression was performed effectively by the proposed method.

Related works on existing methods

The Dubey [18], represented a methodology of constructing a descriptor a locally by using a large scale of the neighbourhood that can be done by an order which is local directional out of other intensity values at various scales in a specific direction. The LDOP (local directional order pattern)—is evaluated by determining the connection between the central pixel and order index that are locally directional.

Wang et al. [19] modeled conditional random field (CRF) by using words that are semantically related, in which each and every vertex represent the final decision. This CRF used RVM to classify local evidence. Three major sections in the proposed architecture were binary image classifier, normalized Google distance, CRF model. Group of words for images was obtained by the binary classifier. The distance between the two words is extracted by NGD. CRF integrate ontologies, refine images. This paper also developed an algorithm to learn the weighted attributes of the CRF model.

Guo et al. [17] developed an image retrieval methodology which used ODBTC features to construct image attributes such as BPF and CCF. They are easily extracted from two ODBTC bitmap and quantizers with the involvement of codebooks visually. ODTBC is simple for describing images in CBIR system. Similar Images are analyzed by relative distance. Edge, shape, content of image are characterized by bit pattern feature.

Zhao et al. [20] represented seep semantic ranking for multi-image retrieval. In this paper deep convolution network is merged with hash functions for performing hash codes mapping. Handcrafted features are difficult to represent in the proposed method.

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.*   (2019) 9:31

Page 6 of 29

The proposed method was done using three ways deep hash functions, semantic ranking supervision, surrogate loss optimization. An input image is transformed and provided to convolution layers, fully connected layers in deep hash functions. Problems in SVM was resolved by semantic ranking supervision.

Lin et al. [21] proposed an efficient deep learning framework for creating binary codes by using (CNN) to retrieve large scale images. The idea was, in the presence of data labels, hidden layers are employed to represent the latent concept for the dominant class labels. The hash codes and images are learned in a pointwise manner instead of pairwise in other supervised methods of image retrieval. With simple alterations in deep CNN, this method showed 1% to 30% improved retrieval precision on MNIST and CIFR-10 data sets.

From the above review of the existing approaches, certain major issues in image retrieval was observed, such as high computational complexity, inaccurate feature extraction and classification, increased computational time, high cost, semantic gaps, reliability, and ineffective retrieval of images.

Previous research has examined the image retrieval process using different algorithms; however, certain limitations are present in the traditional image retrieval process, such as ineffective feature extraction, manual image annotation, and lower accuracy. To overcome these issues, a preprocessing technique with a median filter is used to remove the noise for better accuracy and reliability. The retrieval of images is effective due to using a trained classifier for the classification of selected features to retrieve the relevant data.

Proposed work

This section discusses the implementation of the image retrieval and classification technique to classify images that are present in a database. Initially, the images are taken as input from a dataset for preprocessing. Subsequently, the unwanted data are removed by using Gaussian filtering, and features are extracted to obtain the best color and texture features. The features obtained are optimized using modified genetic programming and are classified by using the NSVMBCNN.

An overall proposed flow is shown in Fig. 1. Initially, an input image is taken from the dataset, and preprocessing is done using Gaussian filtering, which is very efficient for reducing contrast. Next, the color and texture features are extracted using IIBCF, GLCM and novel statistical feature techniques for acquiring effective features. Then, the k-means clustering technique is applied for label formation because of the efficiency of this technique.

A modified genetic programming method is utilized for optimization. The key point to make those algorithms are working, a genetic algorithm is used for obtaining optimized features from extracted color, shape and texture features. Here the used objective function is estimating the correlation for each extracted image features. The optimized features are processed by segregating as train features and test features and also the labeled results which are obtained from k-means clustering are considered as train and test label. These inputs are processed in NSUM-CNN classifier for retrieving the relevant images. Entire database images are labeled based on extracted features and this labeling result is used for classifying the image. Functioning's of pre-proposed

Fig. 1 Overall proposed work-flow
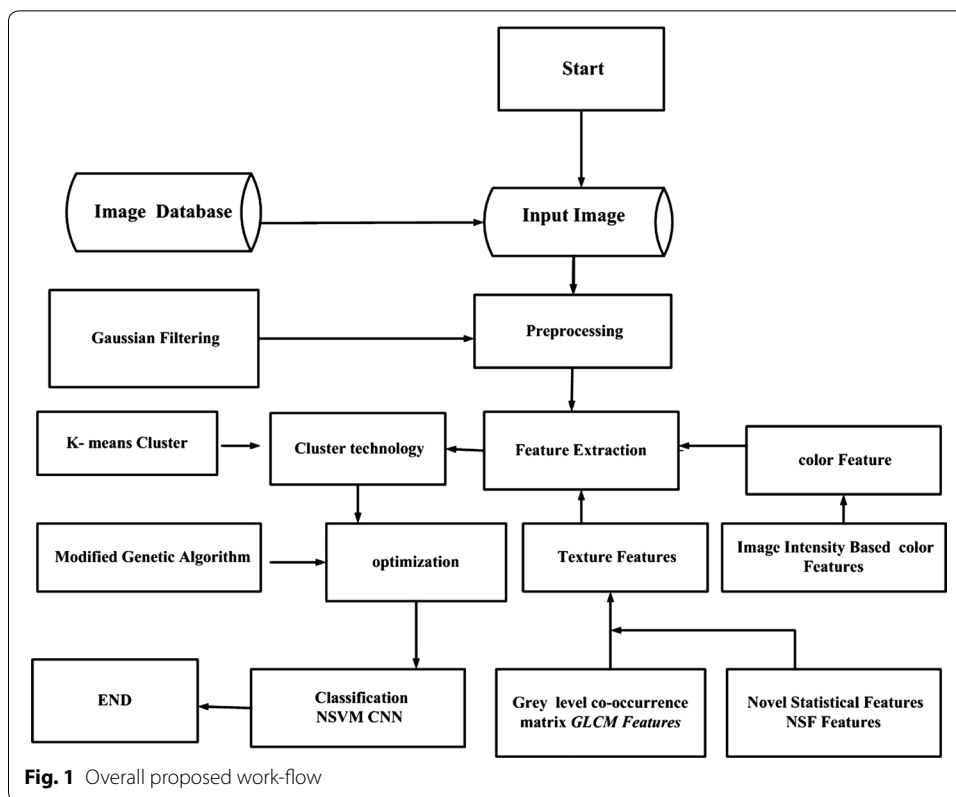
Pre-processing

Preprocessing is an initial step in image processing or retrieval [22]. From the database, the input images are extracted and then preprocessed to increase the efficiency. At this stage, the image features are enhanced using a Gaussian filter, which is used to remove the noise and blur present in the images. The preprocessing technique removes the image noise, and the image dimensions are also reduced. Additionally, this filter is helpful for edge detection. The Gaussian filter function is expressed as

$$G(x,y) = \frac{1}{2\pi(\sigma)^2} * \exp\frac{(x^2) + (y^2)}{2(\sigma)^2} \tag{1}$$

where σ is the standard deviation of the distribution, x is the distance from the origin along the horizontal axis, and y is the distance from the origin along the vertical axis. An advantage of our preprocessing is that the Gaussian filter makes the system much more efficient in terms of reducing contrast and blurring edges.

Figure 2a is an input image taken from the European 1M dataset. Similarly, Fig. 2b is from the Corel 1K dataset, Fig. 2c is from the LFW dataset and Fig. 2d is from the Flickr dataset. The above images are preferred as the input images for the four datasets. With these input images, we perform preprocessing.

Figure 3 shows the preprocessed images obtained from the input images. The European 1M, Corel 1K, LFW and Flickr 27 dataset preprocessed images of the given input are shown in Fig. 3a–d.

Fig. 2 Input images

Fig. 3 Pre-processed image
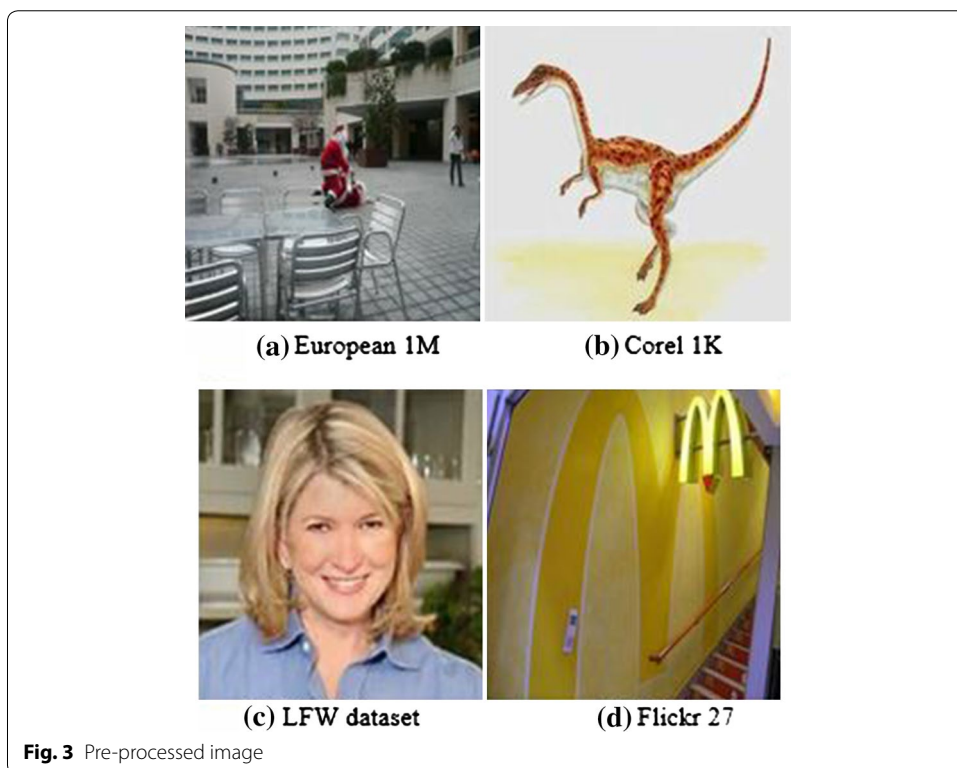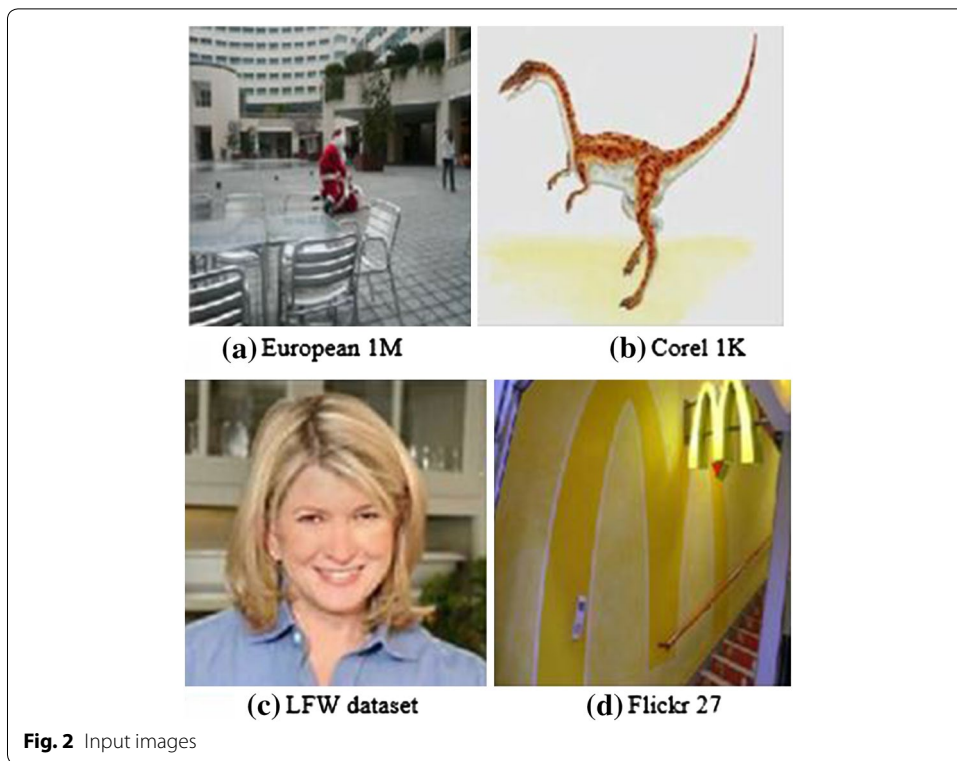
Feature extraction

Feature extraction is the process in which the features are extracted from the input data and transferred into a set of features, splitting the accurate data from the massive set of databases and retrieving the appropriate and revealing features (also known as dimensionality reduction). In this stage, the image features (texture and color) are extracted [23].

Colour features

Color cooccurrence features (CCFs) A huge quantity of information regarding the image content is presented in the color distribution of the pixels in the image. The distribution of image color provides the attributes of that image with the support of the CCFs. This type of matrix is used to compute the probability of the occurrence of a pixel together with nearby pixels, allowing for the evaluation of the color features of the desired color. Color features are extracted based on the intensity of the image, which makes the system more effective. Thus, the color feature extraction in our proposed work is effective due to the intensity-based color feature retrieval.

Algorithm 1: IIBCF - Image Intensity Based Color Features:

Input: Filt_Color_Image
Output: Color_Feature
Procedure:
step 1: RGB to gray conversation for future processes. GrayImage = rgb2gray(Filt_Color_Image);
step 2: After the gray conversation, we take the color features for the gray image using the NIBCF.
Output_Image = IIBCF (GrayImage,6,2,1);
where IIBCF is an image intensity-based color feature
step 3: We calculate the IIBCF values using the following steps.
Input_Img = round(GrayImage); [r c] = size(Input_Img); Mini_intensity = min(min(Input_Img));
Maxi_intensity = max(max(Input_Img));
Out_I = zeros(Maxi_intensity-Mini_intensity+1);
Then, we determine the values of Dir_x and Dir_y.
Dir_x = Dir_x*Dis;
Dir_y = Dir_y*Dis;
Then, we determine the values of Intensity1 and Intensity2.
Based on the intensity values, we calculate the Output_Image.
The values of Intensity1 and Intensity2 depend on the minimum intensity and maximum intensity.
step 4 After the IIBCF and Output_Image calculations, we estimate the Color_Features for the input image.
M= Output_Image
for k1 = 1:4:size(M,1)-3
idxr = 1+fix(k1/4);
for k2 = 1:4:size(M,2)-3
idxc = 1+fix(k2/4);
MF = M(k1:k1+3,k2:k2+3);
A(idxr,idxc) = mean(MF(:));
end
end
IIBCF_Feature = abs(A(1:28)); The IIBCF_Feature values are the novel features for the color image.
End

Algorithm 1 depicts the image intensity-based the color features. Initially, the RGB is converted into a gray image. Then, in the second step, the color features are determined. Subsequently, the IIBCF values are calculated based on the intensity values. Thus, ultimately, the color features are extracted from the input image, for which the IIBCF feature values are the novel features for the color image [24].

Texture feature

Texture features

Texture retrieval for feature extraction requires the following: gray level cooccurrence matrix (GLCM), and novel statistical features (NSF).

Gray level cooccurrence matrix GLCM

The GLCM is a statistical method of examining texture that states the spatial relationship of pixels, a gray-level cooccurrence matrix (GLCM), also known as the gray-level spatial dependence matrix, is a statistical method of examining texture that determines the spatial relationships of pixels. The GLCM operations demonstrate the texture of an image by evaluating how often pairs of pixel with exact values and specified spatial relationships occur in an image, producing a GLCM and extracting statistical measures from this matrix [25].

A GLCM includes pixel position information, which assumes the values of a gray-level matrix. This parameter is a two-dimensional array whose columns and rows depict the set of probable values of that image. Twelve statistical features, namely, autocorrelation, inverse different moment, entropy, energy, homogeneity, sum of squares, dissimilarity, correlation, cluster prominence, maximum probability, inverse different moment and contrast, are extracted from the cooccurrence matrix. A GLCM Pd [i, j] is initially defined by the identification of a displacement vector d = (dx, dy) and the addition of all sets of pixels separated by d that are have gray levels i and j.

Information regarding the positions of the pixels is retrieved by the GLCMA GLCM is utilized in this work to acquire more information about an image and provides an interpretation of the co-occurrence matrix. An advantage of the cooccurrence matrix calculations is that the cooccurring pairs of pixels can be spatially related in several orientations related to distance and angular spatial relationships, considering the associations between two pixels at a time. Finally, the mixture of gray levels and the respective positions are determined [26]

Novel statistical features (NSF)

The statistical features are defined as the correlations between adjacent pixels. The features are extracted by computing the mean, median, standard deviation, skewness, kurtosis, coefficient of variance, covariance, correlation, and entropy. These computations give a set of feature vectors for an image. An NSF is an efficient method for acquiring statistical information.

Shape feature, here the Gaussian filter pre-processed the input image which in turn extract the shape feature. The shape is an important attribute for identifying and recognizing objects. This feature can be extracted by using two techniques namely, (1) contour based technique—which calculates are the shape from image boundary, (2) region based technique—extracts the whole region of the image.

Clustering technique

In this clustering approach, k-means clustering is used. The k-means clustering algorithm is commonly used. Clustering techniques are used to label images in the database regarding color, size and shape. The entire database images are labeled based on

extracted features and this labelling result is used for classifying the image. An optimization is performed on the extracted features using a modified genetic algorithm to acquire a better image retrieval process. The k-means clustering technique is used for classification processes in which an activation update is performed. Finally, the resulting images are retrieved [27].

k-means clustering

Algorithm 2: K-Means Clustering

Input: Initialized group of image pixels
Output: Labeled Clusters
step 1: Compute the K centroids for each originally initialized cluster. K is the number of clusters primarily assigned according to equation (3).
step 2: For each pixel, compute the Euclidean distances between the pixel and all the centroids according to equation (4).
step 3: According to the Euclidean distance, assign each pixel to the nearest cluster with the newly computed centroids calculated based on the estimation of the distance between the centroids and pixels .
step 4: Compute the new centroid with all the pixels assigned to the same cluster.
step 5: Iterate Steps 1-4 until the maximum number of iterations is reached or all the resulting clusters remain unchanged.
step 6: The image pixels processed are ultimately labeled with the last assigned cluster centroids.
End

A specified set of image pixels is segregated into predefined K groups, as in [28]; thus, the pixels similar to each other are in the same group. Primarily, the K-centroid is assessed for a given group of pixels, and the distance of a pixel to the centroid of its group is assumed to be shorter than those of the pixel to the centroids of other groups. To find the most suitable centroid for each pixel, the Euclidean distances between the pixel and each centroid are computed and compared; then, the pixel is assigned to the group with the nearest centroid.

After all the pixels are computed and assigned, the new centroids are computed according to the pixels in the same group. The new centroids may be computed or estimated several times until the maximum number of iterations is reached or the centroids remain unchanged. The final centroids are used as the labels of the pixels in the corresponding centroid group, and later, the labels are used to train the naïve Bayes classifier. Thus, the k-means results generally become increasingly lower as the iterative computations progress [28].

$$Kmeans_{result} = \sum_{j=1}^{k} \sum_{i=1}^{n} \|(x_i^j - c_j)\|^2 \tag{2}$$

where $\|(x_i^j - c_j)\|^2$ denotes the distance between pixels *xij* to the group to the group center, n is the total number of data points obtained in the group, and K is the total number of groups or clusters. The centroid calculation for K clusters is $k = 1, 2, 3, ..., K$

$$C_k = lim_{k \to 1tok} \frac{\sum_{i=1}^{m} Feat_i}{k} \tag{3}$$

Euclidean distance calculation:

$$Ed_i = lim_{i=1:m1} argmax_{k \to 1tok} = \sqrt{(Feat_i - C_k)^2} \tag{4}$$

In this work k means clustering approach has been chosen due to its simple, fast and efficient clustering procedure

The novel SVM based convolutional neural network classifier (NSVMBCNN)

A NSVMBCNN classifier is utilized in our work to acquire the benefits of both SVM and CNN approaches. The classification approach is better in which every image features are acquired from the train feature matrix to get the class values. With this novel classifier, the precision and recall value achieves higher as it specified below in the metrics. The optimized features are processed by segregating as train features and test features and also the labeled results which are obtained from k-means clustering are considered as train and test label. These inputs are processed in NSVMB-CNN classifier for retrieving the relevant images.

Algorithm 3: NSVMBCNN- Novel SVM Based Convolutional Neural Network Classifier

---

Input: Test_Feat, Train_Feat and Label
Output: : Output_Class.
step 1: : Take every feature of the image from Test_Feat, obtain the size of the Train_Feat
  matrix, and obtain the class values.
step 1.1: [m,n]= size
(Test_Feat)
iTrainrfin=[];
step 2: : for I = 1:m
step 2.1: : Train_t= Test_Feat(i,:)
step 2.2: : Unique_Class= unique(Label);
N=length(Unique_Class);.
step 3: : Calculate SVM Struct.
step 3.1: if (N > 1)
iTrain_r = 1;
classes = 0;
step 3.2: Cond = max(Class_Label)-min(Class_Label); c1=(Label == Unique_Class(iTrain_r));
newClass = c1;
svmStruct=
step3.3:svmtrain(Train_Feat,newClass,'kernel_function','rbf');
step 4 : Estimate Train_Feat based on SVM Struct.
step4.1:for X=1:size(Train_Feat,1) .
Train1(X,:)=svmStruct.ScaleData.scaleFactor.*Train_Feat (X,:);
end
step 4.2: for X=1:size(Train_t,1) Trainst1(X,:) =
  svmStruct.ScaleData.scaleFactor.*Train_t(X,:);
end
step 5:: Train the convolutional neural network (CNN).
step 5.1: Train_layers = Convolution_Layer and Sub_Sampling_Layer
step 5.2: Get the batch size of the test feature
size([Test(:) ; Test(:)],1)
CNN_res = test_y(cnn_out);
End

Algorithm 3 shows the NSVMBCNN. Every image feature is taken from Test_Feat, the size of the Train_Feat matrix is obtained, and the class values are acquired. Next, SVM Struct is calculated. Based on SVM struct, the training features are estimated. Finally, the convolutional neural network is trained.

Optimization using modified genetic algorithm

In this study genetic algorithm is used for obtaining optimized features from extracted colour, shape and texture features. Here the used objective function is estimating the correlation for each extracted image features. There are several key components that are needed to define the genetic programming, including terminals, functions, fitness

functions, reproductions, crossovers, and mutations. To find an optimal solution for assigning reasonable weights to the classes based on features such as color, texture and shape, a genetic algorithm is employed in this proposed approach. The iteration process starts with the set of training images combined with the acknowledged decisions.
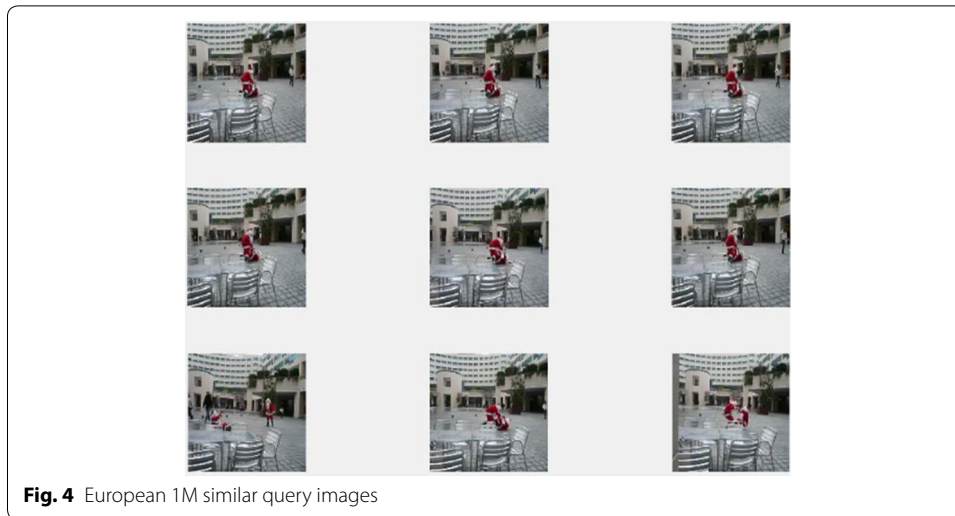
Initially, the GP activates on a very large population of random combination functions. According to the information from the training images, the combination functions are evaluated. The genetic transformation is used for processing until the stopping criteria is reached. The process generally creates and estimates the next generation population in an iterative manner. By developing a population with various generations, the GP pursues improved combination functions. By applying genetic transformations, for example, mutations, crossovers and reproductions, the population individuals are modified.

In this modified GP, from the elements of the terminal and function sets, individual programs are constructed. A tree based GP is used for characterizing the individual (i.e., image descriptor) progress through the GP-criptor, in which the terminal set provides the terminal nodes, i.e., leaves. From the function set, every nonterminal node is drawn. Furthermore, strongly typed GP (STGP) is employed for announcing the node restrictions. Each individual is a set of synthesized formulas that is used to extract the feature vectors.

Four nodes comprise the terminal set: (1) min (_x), (2) max (_x), (3) mean(_x), and (4) stdev (_x), which are functions that return the minimum, maximum, mean, and standard deviation values of the elements of a vector, respectively. These functions are preferred based on their order independence while extracting features. On the other hand, vector value scuffling does not affect the results of these functions. The rotation variants of the pixels is significant and should be handled properly. The vector of the integer value is acquired from the terminal nodes, and a single floating point value is returned. The terminal set is the main difference between the GP-criptor and GP-criptorri.

In GP-criptor, the original pixel values of the sliding window, which is randomly selected, are the leaf nodes of the evolved program; however, the leaf nodes in the GP-criptorri are the pixel values statistics of the sliding window. There are five nodes in the function set, i.e., code, and the four arithmetic operators $+$, $/$, $-$ and $*$. The arithmetic operators have the standard definitions, yet / is secure, thus, this operator returns zero when the denominator is zero to avoid the "division by zero" problem. Spaced out from code, two input arguments are taken by those functions, and a single output is returned.

Furthermore, the input and output are floating-point values; therefore, an input node can be taken from the output of one node. Several operators are used in this approach, including trigonometric functions and logical operators, which signifies another crucial advantage of the GP-criptorri: The descriptors designed by the researchers are not more flexible. The predefined numbers of the argument are taken from the code node, and binary code is returned. Due to type incongruity between the output of the code node and the output of the other function node, the code node does not appear as a child of the other node. The code node exists in the individual tree root, and separate individuals consume only one code node. By using 0 as a threshold, the node converts the input to binary. To construct the feature vector, the generated codes are used [6].

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.*   (2019) 9:31

Page 14 of 29

Fig. 4 European 1M similar query images

Algorithm 4: MGP: Modified Genetic Algorithm

Input: Initialised group of image pixels
Output: Labelled clusters
step 1: : t← 0 **step 2:**InitPopulation [P(t)]; Initialises the population.
step 3: EvalPopulation [P(t)]; Evaluates the population .
step 4: While not terminating, do.
step 4.1: P'(t) ← Variation [P(t)]; Creates new solutions .
step 4.2: EvalPopulation [P'(t)]; Evaluates the new solutions.
step 4.3: P(t+1) ← ApplyGeneticOperators [P'(t) U Q]; Next,generation pop. t ← t + 1;
end while.

The MGP is shown in Algorithm 4. The population is initialized, and then evaluations are made. After this process, a new solution has been created and evaluated. Finally, the genetic operators are applied to obtain effective results

Figure 4 shows similar query images from the European 1M dataset. Five images are selected as queries from each group. If the group contains less than five images, all of the images are used as a query. Figure 5 shows similar query images from the Corel 1K dataset. This research uses the commercially accessible Corel photograph gallery, which contains 1000 color images sized 384 * 256 pixels.

Figure 6 shows similar query images from the Labeled Faces in the Wild (LFW) dataset, which is a database of photographs of faces designed for studying the problem of unconstrained facial recognition. The Flickr logo dataset similar query images are shown in Fig. 7.

Figure 4 shows the similar query images of the European 1M, Corel-1k, LFW and Flickr 27 datasets. Figure 4 shows the similar query image for the European 1M dataset. Five images are selected as queries from each group. If the group contains less than five images, then all the images are used as queries.

Figure 5 presents the similar query image for the Corel-1k dataset. The commercially accessible Corel photograph gallery, which contains 1000 coloured images with $384 * 256$ pixels, used in this study. Figure 6 shows the similar query image for the LFW dataset, which is a database of face photographs designed for studying the

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.* (2019) 9:31

Page 15 of 29

Fig. 5 Corel 1K similar query images

Fig. 6 (LFW) dataset similar query images

Fig. 7 Flickr logo dataset similar query images
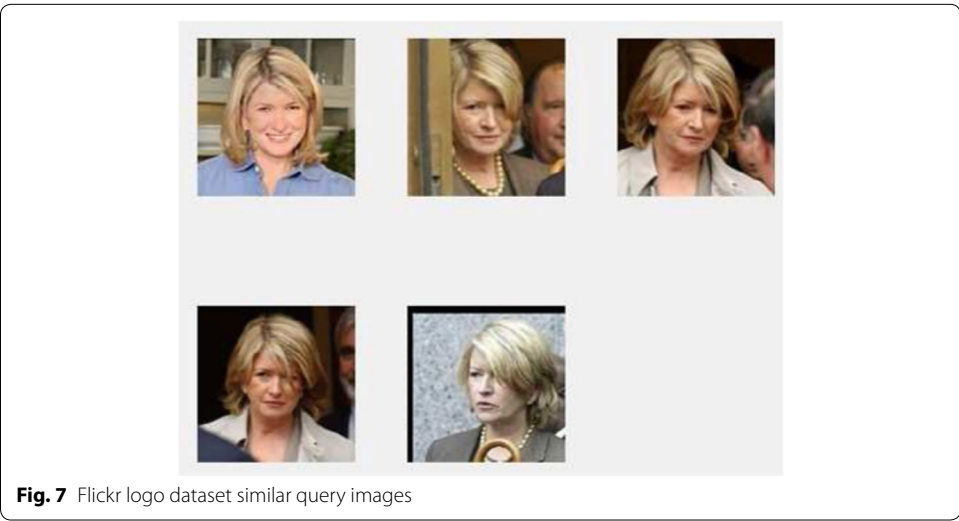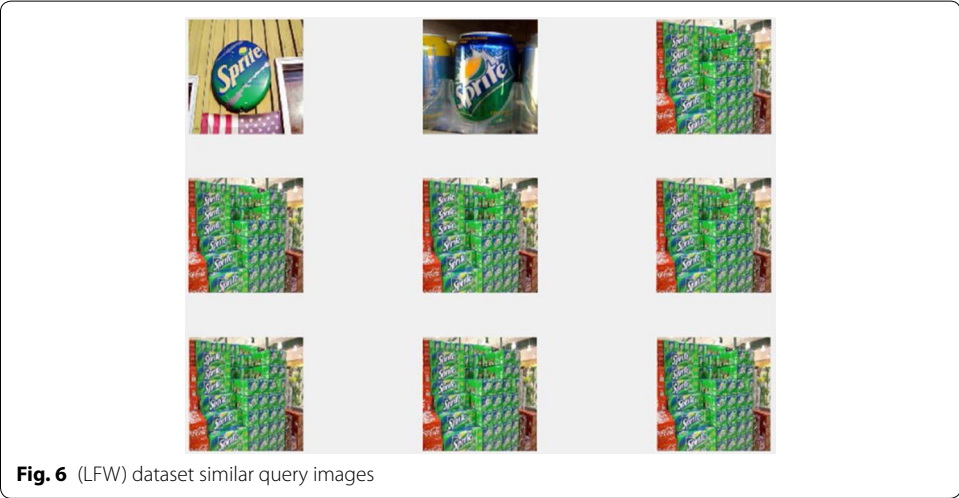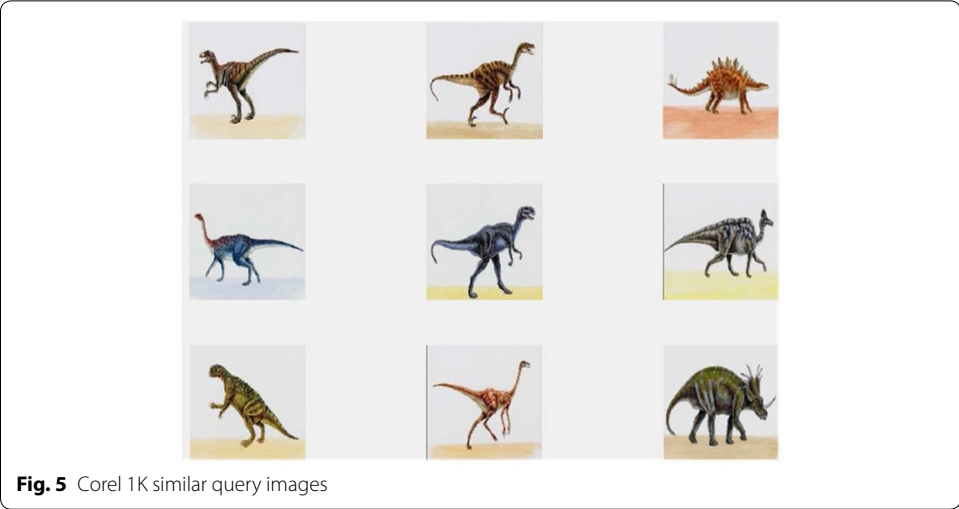
problem of unconstrained face recognition. Similar query images are predicted based on the distance, i.e., images with similar features that are nearby are selected. The Flickr logo dataset similar query images are shown in Fig. 7

Performance analysis

This section discusses the performance results of the proposed technique on the dataset. Comparative analysis of the proposed dataset is used to evaluate the proposed method versus an existing other methods. The metrics are sensitivity, specificity, accuracy, precision, recall, and retrieval time. Notably, the NSVMBCNN classification provides effective results.

Datasets

The European 1M, Corel 1K, Flickr 27, LFW datasets are used in this work to evaluate the performance measures. These datasets are individually discussed below to evaluate the efficiency of the proposed method. Table 2 shows the parametric analyses of the various datasets.

Performance measures

Accuracy: The accuracy is how close a measured value is to a standard or known value [29]. Accuracy is also associated with the weighted arithmetic mean of the precision and inverse precision (weighted by the bias) and the weighted arithmetic mean of the recall and inverse recall (weighted by the prevalence)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{5}$$

Precision [30]: The precision is a basic measure for evaluating the performance of classification techniques. In image retrieval, precision is the fraction of retrieved images that are relevant to the query images. Precision is calculated as follows:

$$Precision(X(I_i)) = \frac{A}{B}. \tag{6}$$

The average retrieval precision is calculated by the following equation,

$$ARP = \frac{1}{\mid Ds \mid} \sum_{i=1}^{\mid Ds \mid} X(I_i). \tag{7}$$

where Ds is the total number of correctly retrieved images.

Recall [30]: The recall measures the prediction ability of models and is mainly used to select the instance of a certain class from a dataset. The recall is also called the sensitivity, which is calculated as follows:

$$Recall(T(I_i)) = A/(Total\ number\ of\ relevant\ images\ in\ database) \tag{8}$$

Average retrieval rate is calculated as

$$ARP = \frac{1}{\mid Ds \mid} \sum_{i=1}^{\mid Ds \mid} T(I_i). \tag{9}$$

Table 1 Performance analysis of the SVM, CNN algorithms and the proposed NSVMBCNN

Parameters	Proposed NSVMBCNN	Algorithm SVM	Algorithm CNN
Tp	5	5	5
Tn	52	62	44
Fp	13	3	21
Fn	0	0	0
Sensitivity	80	68.5714	70
Specificity	99.4615	97.5824	97.6923
Precision	80.3651	52.4956	65.6593
Recall	82.8451	68.5714	70
Accuracy	98.8429	95.5102	95.7143

Table 2 Parameter analysis for the four datasets

Parameter	European 1M	Corel-1k	LFW	Flickr 27
Sensitivity (%)	90	98	99	89
Specificity (%)	98	98	98	99
Precision (%)	89	98	99	83.78
Recall (%)	98.7	98	100	99.69
Accuracy (%)	96.77	96	97.14	92.72

where, A—number of relevant images retrieved, B—total number of images retrieved, ARP—average retrieval precision, TP—true postive, TN—true negative, FP—false postive, FN—false negative, and DB—database.

Table 1, shows that provide the evidence of our proposed method, we have used four different data set such as European 1M dataset, Corel-1k Dataset, LFW Dataset and Flickr Logos 27. However, we tested the SVM, and CNN with proposed methods on datasets based on performance analysis for each of TP, TN FP, FN, accuracy, precision, recall. Our proposed method better results than existing other methods.

Table 2 and Fig. 8 portray the analysis of the performance measures in terms of precision, recall, accuracy, sensitivity and specificity. The accuracy is one of the significant measures; here, the algorithm acquires higher accuracy on the LFW dataset than on the other datasets. Similarly, the Corel 1K dataset achieves better values for all the parametric measures. The ranges of sensitivity four datasets performance measures are analysis to identify the efficiencies of the algorithm on the datasets.

European 1M dataset

The European Cities 1M dataset consists of 909,940 geo-tagged images from 22 European cities, scraped from Flickr using geographic queries covering a window of each city center. A subset of 1081 images from Barcelona is annotated into 35 groups depicting the same scene; 17 of the groups are landmark scenes, and 18 are nonlandmark scenes. Since not all the scenes are landmarks, the annotation cannot rely on tags but rather relies on a combination of visual query expansion and manual clean-up. Five images are selected as queries from each group.

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.*　　(2019) 9:31

Page 18 of 29

Fig. 8 Performance measures accuracy, precision, recall, sensitivity and specificity analyses

Fig. 9 Precision, recall, and F-measure analyses for landmark images in the European dataset
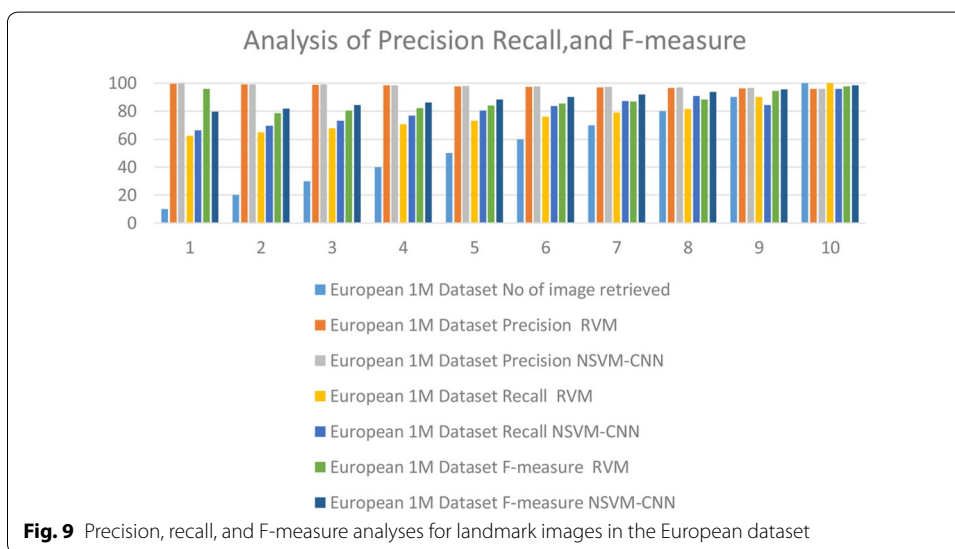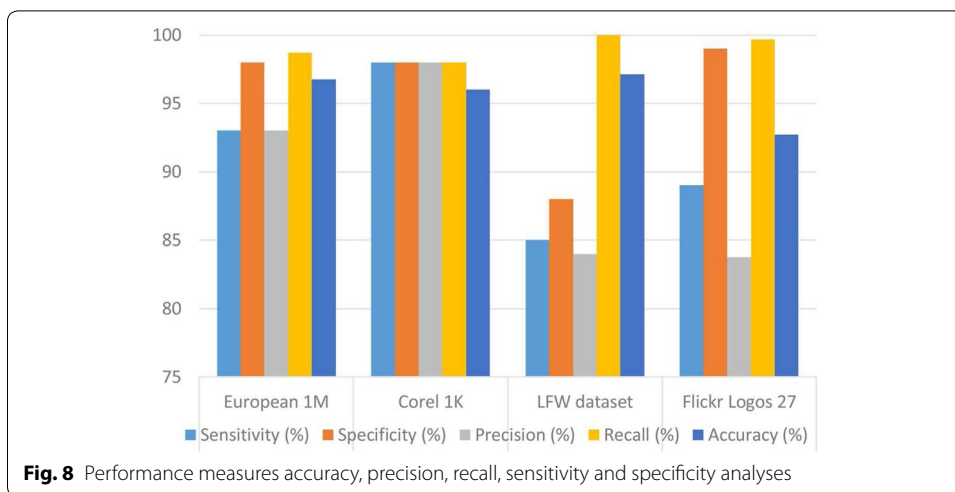
If the group contains less than five images, all of them are used as a query. The NSVMBCNN learning model achieves better accuracy, precision and recall rate than the RVM owing to the accurate clustering model, which is shown in Fig. 9. The categorization of separate sets of features through image intensity-based color feature approaches that are employed in earlier clustering enables the accurate processing of labels with respect to these images. Therefore, the classification performance of the entire system escalates accordingly. The devised NSVMBCNN approach shows significantly enhanced performance compared with the existing RVM [19] classifier[1]

Figure 9, it depicts that the RVM method is compared with our proposed method to evaluate the precision, recall and F-measure. In this approach, European Cities 1M dataset acquires 96%, 95.9% and 97.3% which is comparatively much better than the RVM approach.

---

[1] http://image.ntua.gr/iva/datasets/ec1m/.

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.* (2019) 9:31

Page 19 of 29

Table 3 Comparative analysis of precision, recall and F-measure for existing RVM and proposed NSVM-CNN

No of image retrieved	European 1M dataset					
	Precision		Recall		F-measure	
	RVM	NSVM-CNN	RVM	NSVM-CNN	RVM	NSVM-CNN
10	99.5	99.8	62.34	66.21	95.78638	79.60675
20	99.1	99.25	65.02	69.67	78.52159	81.87009
30	98.65	98.98	67.84	73.23	80.39421	84.17984
40	98.23	98.54	70.56	76.72	82.12701	86.2717
50	97.8	98.13	73.31	80.23	83.80244	88.28179
60	97.38	97.72	76.06	83.74	85.40963	90.19148
70	96.95	97.32	78.81	87.25	86.9439	92.01029
80	96.53	96.91	81.56	90.76	88.41582	93.73423
90	96.1	96.51	90	84.31	94.27	95.37685
100	95.91	96.023	100	95.91	97.73	98.35769
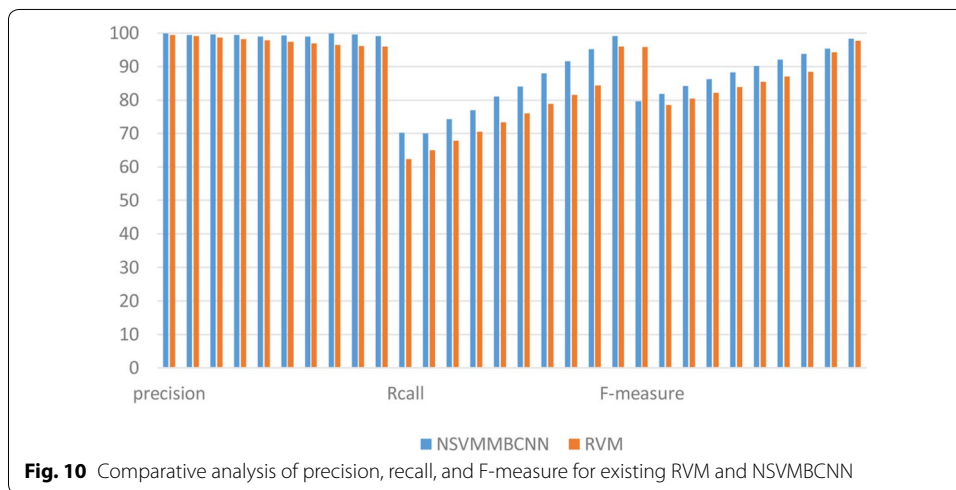
Fig. 10 Comparative analysis of precision, recall, and F-measure for existing RVM and NSVMBCNN

Table 3 and Fig. 10 show comparative analyses of the precision for the existing and proposed methods. Precision comparisons are made for different numbers of retrieved images. At a retrieval image rate of 10, the existing approach attains a precision of 99.5, which is 0.5% lower than that of the proposed approach. For the image retrieval rate of 20, the RVM achieves a precision of only 99.1, which is comparatively lower than that of the NSVMBCNN. Ultimately, for the 100 image retrieval rate, NSVMBCNN achieves s a precision of 98.023.

Table 3 and Fig. 10 show comparative analyses of the recall of the RVM and the proposed method. The table and graph show that for the consecutive image retrieval rate, the precision is high for our proposed NSVMBCNN when compared with the RVM. At an image retrieval rate of 100, a precision of 95.91 is achieved, which is almost 4% less than the proposed NSVMBCNN precision of 99.73. At last 100 image retrieval rate of F-measure, the proposed NSVM-CNN for the existing method is higher than the RVM method.

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.*   (2019) 9:31

Page 20 of 29

Table 4 Comparative analysis of precision, recall and F-measure for existing DSCOP and proposed NSVM-CNN

No of image retrieved	Corel 1K dataset					
	Precision		Recall		F-measure	
	DSCOP	NSVM-CNN	DSCOP	NSVM-CNN	DSCOP	NSVM-CNN
10	80	99.52	9	93.45	16.17978	96.38953
20	75	99.46	15	93.97	25	96.63709
30	72	99.32	21	94.49	32.51613	96.84482
40	69	99.23	28	95.01	39.83505	97.07416
50	65	99.13	32	95.53	42.8866	97.29671
60	63	99.03	38	96.05	47.40594	97.51724
70	60	98.93	42	96.57	49.41176	97.73576
80	59	98.83	47	97.09	52.32075	97.95227
90	57	98.73	50	97.61	53.27103	98.16681
100	55	98	55	98	55	98

Fig. 11 Precision, recall, and F-measure analyses for landmark Images in the Corel 1K dataset
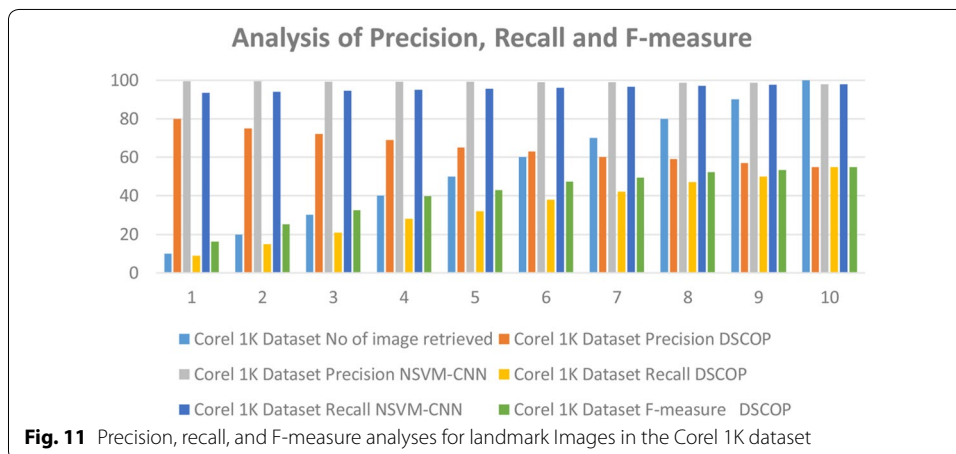
Corel-1k dataset

In the implementation of this method, the Corel 1K dataset is used. This research uses the commercially accessible Corel photograph gallery, which contains 1000 color images with 384 * 256 pixels. All the images in the database are grouped into 10 classes, where each class involves 100 images with several semantic groups, such as 'beach', 'building', 'people', 'elephant', 'dinosaur', 'flower', 'mountain', 'food', 'horse' and 'bus'. Pictures belonging to the same class or semantic group are considered similar images[2]

Figure 11 shows that the NSVMBCNN learning model achieves a better accuracy, precision and recall rate than the DSCop [31] owing to the accurate clustering model. Therefore, the classification performance of the entire system escalates accordingly. The devised NSVMBCNN approach shows significantly enhanced performance compared with the existing DSCoP classifier with a stipulated enhancement of 45 and 40%
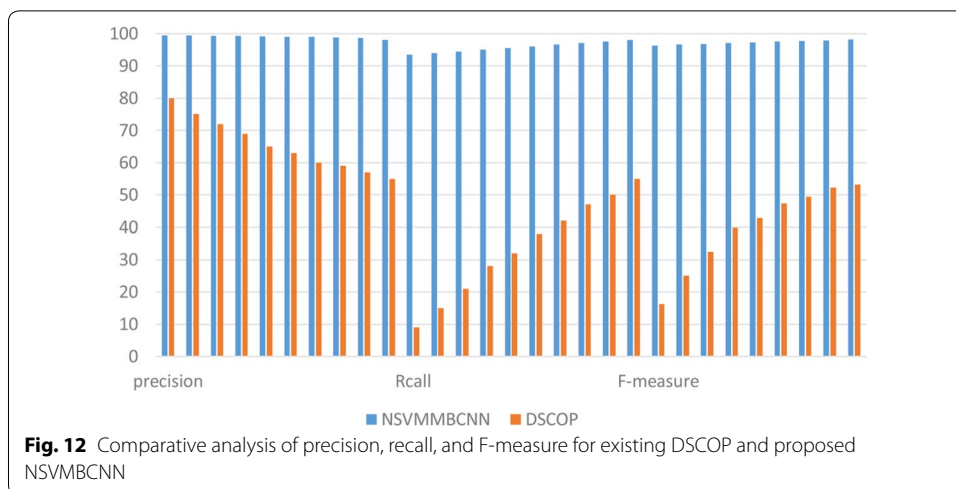
---

[2] https://blog.csdn.net/garfielder007/article/details/51483759.

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.* (2019) 9:31

Page 21 of 29

Fig. 12 Comparative analysis of precision, recall, and F-measure for existing DSCOP and proposed NSVMBCNN

Table 5 Comparative analysis of precision

Visual features	HSV	GLCM	SIFT	CNN	NSVMBCNN
Precision	43	39	53	79	89.143

for the precision and recall, respectively. Figure 11 depicts that the DSCOP method is compared with our proposed method to evaluate the precision, recall and F-measure. In this approach, Flickr dataset acquires 98%, 98% and 98% which is comparatively much better than the LDOP approach

Table 4 and Fig. 12 show comparative analyses of the precision for the existing DSCOP and proposed NSVMBCNN. An existing DSCOP method is considered for analysis of the precision value with our proposed method. When the proposed NSVMBCNN is compared with the existing DSCOP regarding the Corel 1K dataset, the results achieved by the DSCOP show greater deviations. For an image retrieval rate of 20, the precision rate of the DSCOP is 75, which is less than the 99.46 of the proposed NSVMBCNN. Finally, for an image retrieval rate of 100, the proposed method yields 98, providing better precision.

Table 4 and Fig. 12 show comparative analyses of the recall for the existing DSCOP [29] and proposed NSVMBCNN. An existing DSCOP method is considered for analysis of the recall value with our proposed method. When the proposed NSVMBCNN is compared with the existing DSCOP regarding the Corel 1K dataset, the results achieved by the DSCOP show greater deviations. For an image retrieval rate of 20, the recall rate of the DSCOP is 15, which is less than the 93.97 of the proposed NSVMBCNN. Finally, for an image retrieval rate of 100, the proposed method yields 98, providing better recall. Image retrieval rate of F-measure, the proposed NSVM-CNN from these results, it is recognized that our proposed work does well than the existing approach.

Table 5 [32] show the comparative analysis of the proposed and existing system for the Corel 1k dataset in which the proposed technique provides 89.143 precision rate.

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.* (2019) 9:31

Page 22 of 29

Table 6 Comparative analysis of recall

Class/method	WATH approach (SVM based)	Spatial level-2 approach	Spatial LGH approach	GMM spatiogram	Color fusion	Spatial BoF	NSVMBCNN
Recall	*17.37*	*17.25*	16.84	16.11	15.66	13	95.7

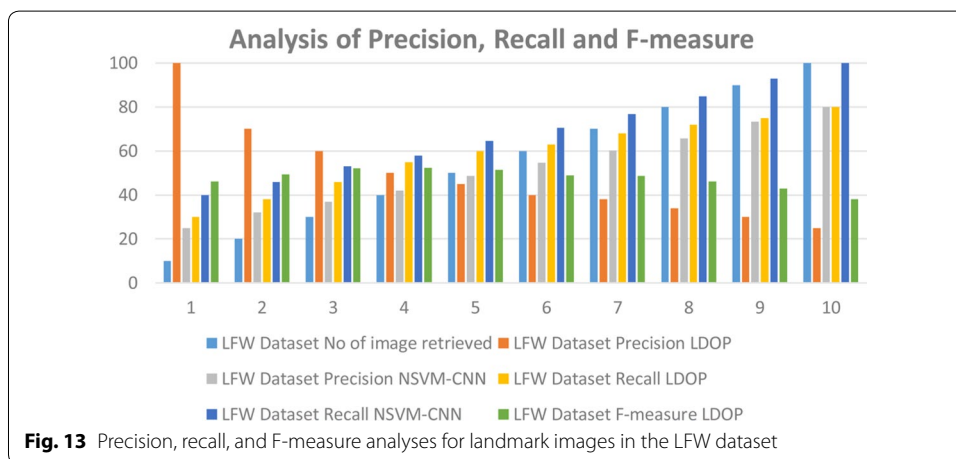The italic values indicate the recall values (recall > 17) from previous methods [30]

Fig. 13 Precision, recall, and F-measure analyses for landmark images in the LFW dataset

Table 6 [33] shows the comparative analysis of the existing and proposed technique in terms of recall. From this it is shown that the proposed technique provides better result than the existing technique

LFW dataset

Labeled Faces in the Wild is a database of face photographs designed for studying the problem of unconstrained face recognition. The dataset contains more than 13,000 images of faces collected from the web. Each face has been labeled with the name of the person pictured. A total of 1680 of the people pictured have two or more distinct photos in the dataset. The only constraint on these faces is that the faces were detected by the Viola–Jones face detector[3].

Figure 13 shows the LDOP method compared with our proposed method regarding the precision and recall measures. Our approach achieves values of 80 and 100% in the LFW dataset for 100 images retrieved, which is comparatively much better than those of the LDOP approach [18]. Figure 13 depicts that the LDOP method is compared with our proposed method to evaluate the precision recall and F-measure. In this approach, Flickr dataset acquires 80%, 100% and 88.8% which is comparatively much better than the LDOP approach.

Table 7 and Fig. 14 shows the precision values for the existing and proposed methods. At the initial image retrieval rate of 10, the LDOP has better precision than the proposed method. However, for the largest image retrieval rate of 100, the precision is

---

[3] http://vis-www.cs.umass.edu/lfw/.

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.* (2019) 9:31

Page 23 of 29

Table 7 Comparative analysis of precision, recall and F-measure for existing LDOP and proposed NSVM-CNN

No of image retrieved	LFW dataset					
	Precision		Recall		F-measure	
	LDOP	NSVM-CNN	LDOP	NSVM-CNN	LDOP	NSVM-CNN
10	100	25	30	40	46.15385	30.76923
20	70	32	38	46	49.25926	37.74359
30	60	37	46	53	52.07547	43.57778
40	50	42	55	58	52.38095	48.72
50	45	48.78	60	64.5	51.42857	55.54926
60	40	54.6	63	70.6	48.93204	61.57764
70	38	60.2	68	76.7	48.75472	67.45566
80	34	65.8	72	84.8	46.18868	74.10146
90	30	73.4	75	92.9	42.85714	82.00673
100	25	80	80	100	38.09524	88.88889

Fig. 14 Comparative analysis of precision, recall, and F-measure for existing LDOP and proposed NSVMBCNN
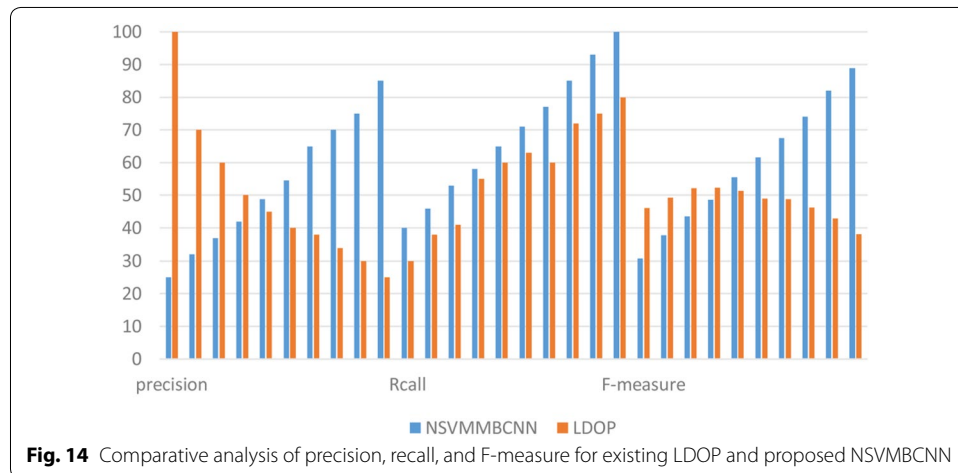
high for NSVMBCNN and is much better than that of the LDOP, which yields only 25% precision.

Table 7 and Fig. 14 describe the recall values of the existing and proposed approaches. The existing LDOP value is more or less equal compared with that of our proposed approach. However, when the image retrieval rate increases, the recall value of the NSVMBCNN increases. Image retrieval rate of F-measure, an existing LDOP value is more or less equal when compared to our proposed approach. Yet, if the image retrieval rate increases the recall value of NSVM-CNN is also increased.

Flickr Logos 27

In the server uploading process, the proposed search engine is loaded with Flickr (for image) and geospatial (for web content) datasets. To show the efficiency of the proposed method, the system is validated with a set of performance metrics such as average time taken to retrieve the results, precision, recall, F-measure and average similarity. Figure 16 shows the CCF + BPF method compared with our proposed method regarding

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.* (2019) 9:31

Page 24 of 29

Fig. 15 Precision, recall, and F-measure analyses for landmark images in the Flickr Logos 27 dataset

Fig. 16 Comparative analysis of precision, recall and F-measure for existing CCF + BPF and proposed NSVMBCNN
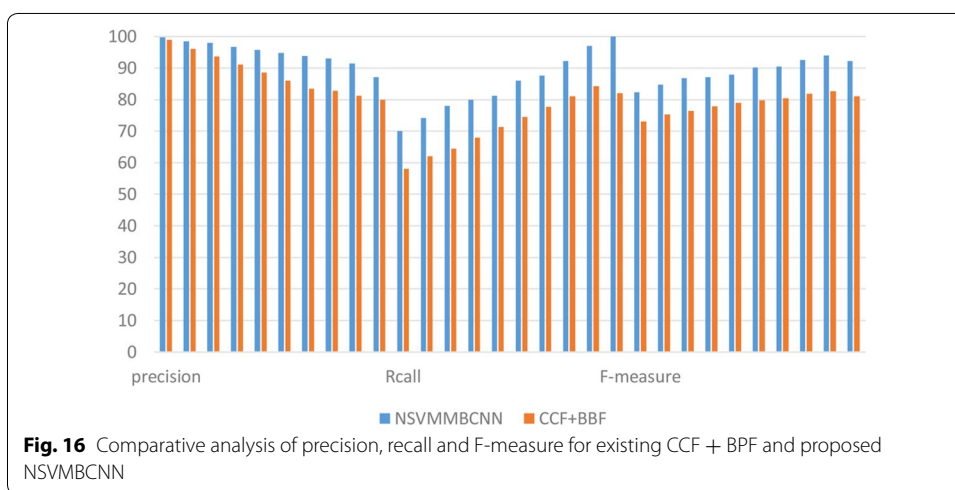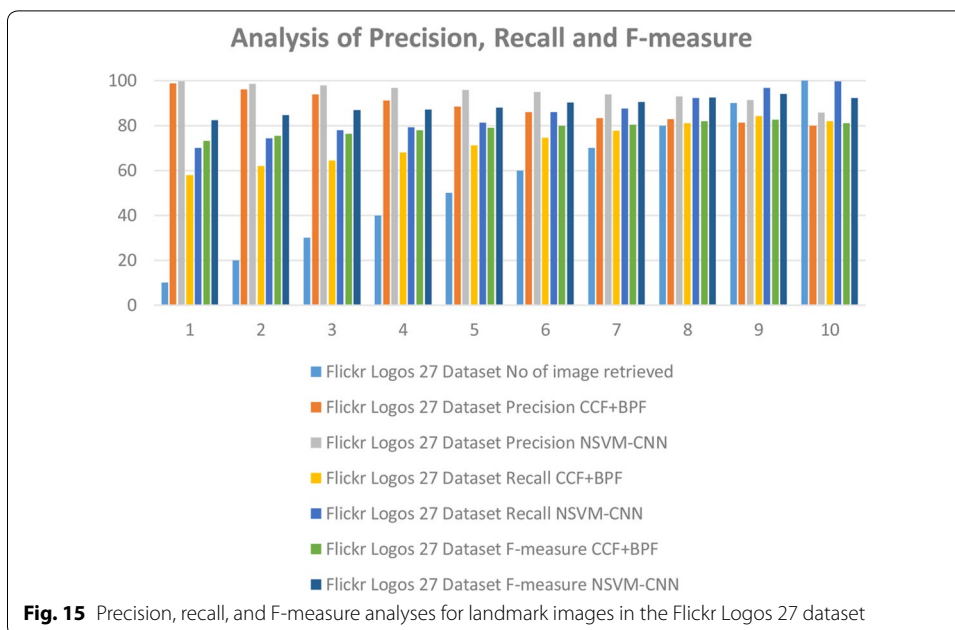
the precision and recall measures. Our approach achieves values of to 99.69% in the Flickr dataset for 100 images retrieved, which is comparatively much better than the values achieved by the CCF + BPF [17] approach.[4]

Figure 15 depicts the comparison of the CCF + BPF method with our proposed method to evaluate the precision and recall measures. It depicts that the CCF + BPF method is compared with our proposed method to evaluate the precision, recall and F-measure. In this approach, Flickr dataset acquires 85.78%, 99.69% and 92.2% which is comparatively much better than the CCF + BPF approach.

---

[4] http://image.ntua.gr/iva/datasets/flickr_logos.

Table 8 Comparative analysis of precision, recall and F-measure for existing CCF + BPF and proposed NSVM-CNN

No of image retrieved	Flickr Logos 27 dataset					
	Precision		Recall		F-measure	
	CCF + BPF	NSVM-CNN	CCF + BPF	NSVM-CNN	CCF + BPF	NSVM-CNN
1–4, 6–7, 10	98.87	99.78	58	70	73.11098	82.27824
20	96.11	98.54	62	74.25	75.37562	84.68771
30	93.73	97.97	64.5	77.98	76.41516	86.83945
40	91.09	96.73	68	79.25	77.86938	87.12186
50	88.52	95.82	71.25	81.24	78.95162	87.92971
60	85.95	94.85	74.5	85.88	79.81645	90.1424
70	83.38	93.81	77.75	87.53	80.46664	90.56126
80	82.81	92.99	81	92.18	81.895	92.58323
90	81.24	91.37	84.25	96.83	82.71763	94.0208
100	80	85.78	82	99.69	80.98765	92.21338
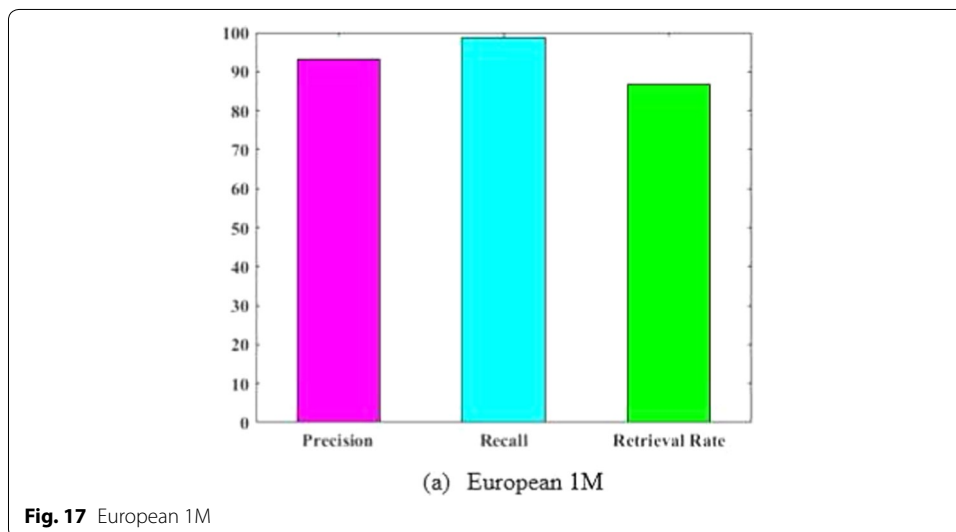
Fig. 17 European 1M

Table 8 and Fig. 16 depict comparative analyses of the precision for the existing CCF + BPF and proposed NSVMBCNN. For several numbers of retrieved images, precision comparisons are made. At a retrieval image rate of 10, the existing approach attains 98.87 precision, which is 99.78% less than the proposed approach. For the image retrieval rate of 20, the CCF + BPF achieves a precision of only 96.11, which is comparatively lower than that of the NSVMBCNN. For the 100 image retrieval rate, the NSVMBCNN achieves a precision of 88.78, which is better than that of the existing method.

Table 8 and Fig. 16 portray the comparative analysis of the recall for the CCF + BPF and proposed method. A table and graph show that for consecutive image retrieval rates, the recall is high for our proposed NSVMBCNN compared with that of CCF + BPF. At an image retrieval rate of 100, 82% recall is achieved, which is almost 17.69% lower than that obtained by the proposed NSVMBCNN, which produced a recall of 99.69. The simulation output supports these results. Image retrieval rate of F-measure. An existing

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.*   (2019) 9:31

Page 26 of 29

Fig. 18 Corel 1K

Fig. 19 LFW dataset
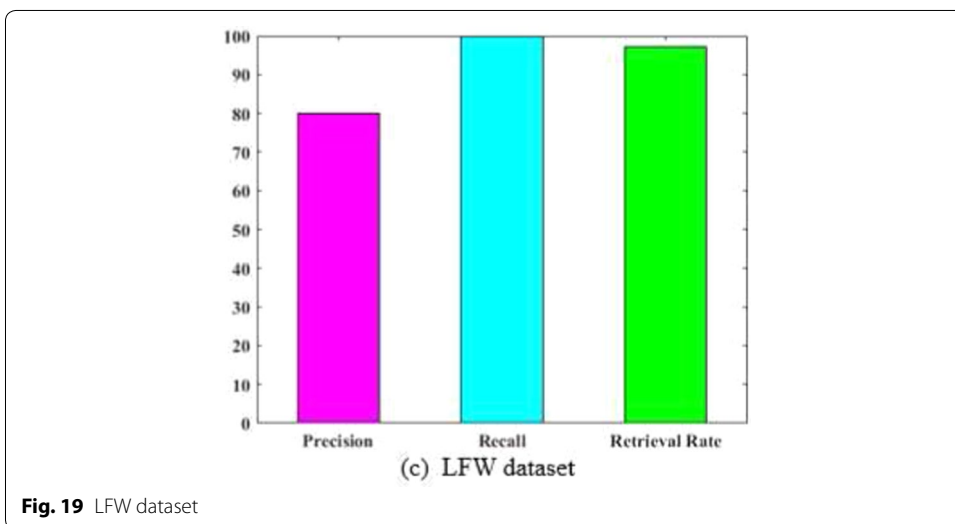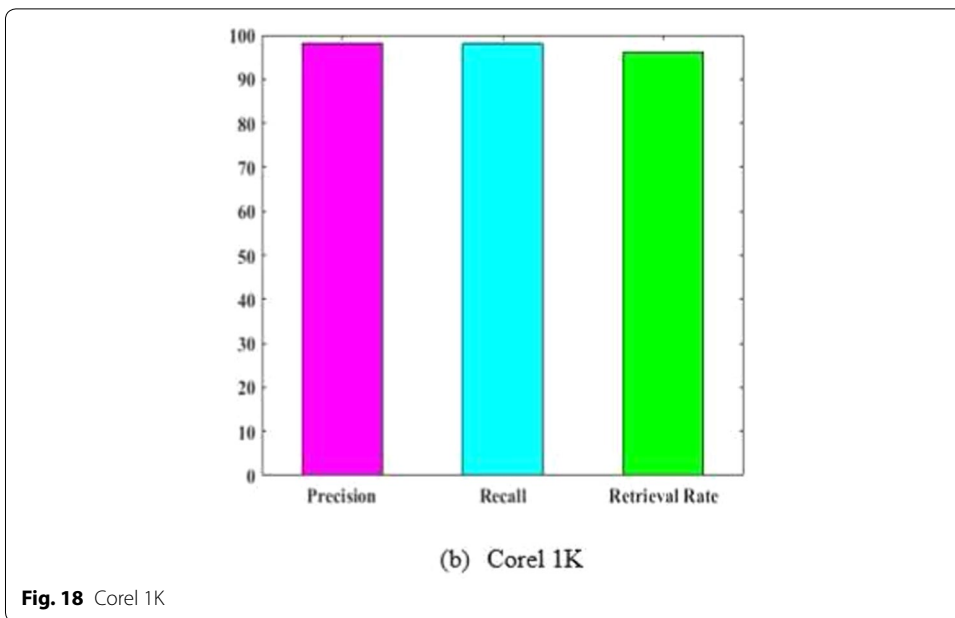
CCF + BBF value is more or less equal when compared to our proposed approach. Yet, if the image retrieval rate increases the recall value of NSVM-CNN is also increased.

Figure 17 portrays the European 1M dataset simulation output, in which the precision is 93, recall is 98.3 and retrieval rate (accuracy) is 86.77.

Figure 18 portrays the Corel 1K dataset simulation output, in which the precision is 98, recall is 98 and retrieval rate (accuracy) is 96.

Figure 19 portrays the LFW dataset simulation output, in which the precision is 80, recall is 100 and retrieval rate (accuracy) is 97.14.

Figure 20 portrays the Flickr 27 dataset simulation output, in which the precision is 75.78, recall is 99.69 and retrieval rate (accuracy) is 70.72.

Fig. 20 Flickr 27 dataset
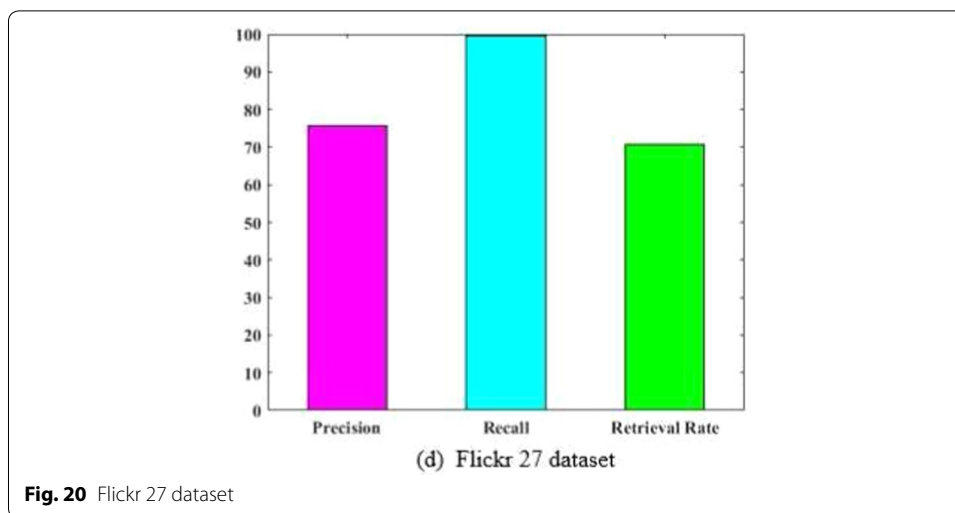
Figures 17, 18, 19 and 20 shows the performance analysis of the European 1M Corel 1K, LFW and Flickr 27 datasets in terms of precision, recall and retrieval rate.

Figure 17 presents the precision, recall and retrieval rate for the European 1M dataset, with values of 93.0, 98.7 and 86.7. Figure 18 shows the precision, recall and retrieval rate for the Corel 1K dataset, with values of 98, 98 and 96%. Better performance is achieved for the LFW dataset, especially the retrieval rate, which is shown in Fig. 19. Finally, in Fig. 20, the Flickr 27 dataset precision, recall, and retrieval rate performance metrics are presented, and the precision and recall values are better than retrieval rate.

The results are better for the proposed approach than for those of the existing RVM, DSCOP, LDOP and CCF + BPF algorithms. The best results for the proposed approach can be obtained by using a modified genetic algorithm with CNN-based classification.

Conclusion

In this paper, the image retrieval process is a major task. Existing method LDOP, DSCOP, CCCF, BPF are harder for selecting kernel function. In order to resolve this problem, the proposed method is employed by using NSVM-CNN. Proposed method comprises of data preprocessing, feature extraction, feature optimization. Preprocessing was applied by Gaussian filtering for eliminating improper data. Features texture, color are extracted depending upon image intensity. The texture feature is considered as GLCM, novel statistical and the color feature is extracted as image intensity based color features. These features are grouped by k-means clustering to form label. Finally, the genetic algorithm proposed to optimize the extracted features and then NSVM-CNN trained and retrieve the relevant images. The proposed method is preferred because of its improved accuracy, performance. The precision, recall and retrieval rate is increased when compared with the other existing techniques. This technique offers the best and accurate results. The empirical results that are represented here make the proposed approach to help in the medical field. i.e. in diagnose sections which completely rely on image retrieval. Currently, in this paper, the proposed approach consists of a range of 5000 datasets. Furthermore, in future, this proposed approach may be utilized in the Big Data Analytics field which will employ more than 2L datasets.

Author details
[1] School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China.
[2] Computer Science Department, College of Education for Pure Sciences, University of Basrah, Basra 61004, Iraq.

References
1. Zhou XS, Huang TS (2003) Relevance feedback in image retrieval: a comprehensive review. Multimedia Syst 8(6):536–544
2. Gudivada VN, Raghavan VV (1995) Content based image retrieval systems. Computer 28(9):18–22
3. Do MN, Vetterli M (2002) Wavelet-based texture retrieval using generalized gaussian density and Kullback–Leibler distance. IEEE Trans Image Process 11(LCAV–ARTICLE–2002–001):146–158
4. Hira ZM, Gillies DF (2015) A review of feature selection and feature extraction methods applied on microarray data. Adv Bioinform. https://doi.org/10.1155/2015/198363
5. Yue-Hei Ng J, Yang F, Davis LS (2015) Exploiting local features from deep networks for image retrieval. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 53–61
6. Al-Sahaf H, Al-Sahaf A, Xue B, Johnston M, Zhang M (2017) Automatically evolving rotation-invariant texture image descriptors by genetic programming. IEEE Trans Evol Comput 21(1):83–101
7. Giveki D, Soltanshahi MA, Montazer GA (2017) A new image feature descriptor for content based image retrieval using scale invariant feature transform and local derivative pattern. Optik 131:242–254
8. Ciocca G, Corchs S, Gasparini F (2016) Genetic programming approach to evaluate complexity of texture images. J Electron Imaging 25(6):061408
9. Sanu SG, Tamase PS (2017) Satellite image mining using content based image retrieval. Int J Eng Sci 13928
10. Nazir A, Ashraf R, Hamdani T, Ali N (2018) Content based image retrieval system by using hsv color histogram, discrete wavelet transform and edge histogram descriptor. In: 2018 international conference on computing, mathematics and engineering technologies (iCoMET), IEEE, pp 1–6
11. Devi NS, Hemachandran K (2017) Content based feature combination method for face image retrieval using neural network and SVM classifier for face recognition. Indian J Sci Technol 10(24):1–11
12. Lu Z, Yang J, Liu Q (2017) Face image retrieval based on shape and texture feature fusion. Comput Vis Media 3(4):359–368
13. Wang X-Y, Liang L-L, Li Y-W, Yang H-Y (2017) Image retrieval based on exponent moments descriptor and localized angular phase histogram. Multimedia Tools Appl 76(6):7633–7659
14. Liu G-H, Yang J-Y (2013) Content-based image retrieval using color difference histogram. Pattern Recogn 46(1):188–198
15. Jian M, Lam K-M (2014) Face-image retrieval based on singular values and potential-field representation. Signal Process 100:9–15
16. Kumar A, Kim J, Cai W, Fulham M, Feng D (2013) Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data. J Digit Imaging 26(6):1025–1039

Ghrabat *et al. Hum. Cent. Comput. Inf. Sci.*   (2019) 9:31

Page 29 of 29

17. Guo J-M, Prasetyo H, Chen J-H (2015) Content-based image retrieval using error diffusion block truncation coding features. IEEE Trans Circuits Syst Video Technol 25(3):466–481
18. Dubey SR, Mukherjee S (2018) Ldop: local directional order pattern for robust face retrieval. arXiv preprint arXiv :1803.07441
19. Wang Y, Gong S (2007) Refining image annotation using contextual relations between words. In: Proceedings of the 6th ACM international conference on image and video retrieval, pp 425–432, Citeseer
20. Zhao F, Huang Y, Wang L, Tan T (2015) Deep semantic ranking based hashing for multi-label image retrieval. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1556–1564
21. Lin K, Yang H-F, Hsiao J-H, Chen C-S (2015) Deep learning of binary hash codes for fast image retrieval. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 27–35
22. Kim H, Lee C, Lee J, Kim J, Yu T, Chung G, Kim J (2019) An explicit numerical algorithm for surface reconstruction from unorganized points using Gaussian filter. J Korean Soc Ind Appl Math 23(1):31–38
23. Zhou J, Liu X, Liu W, Gan J (2019) Image retrieval based on effective feature extraction and diffusion process. Multimedia Tools Appl 78(5):6163–6190
24. Chun YD, Kim NC, Jang IH (2008) Content-based image retrieval using multiresolution color and texture features. IEEE Trans Multimedia 10(6):1073–1084
25. Bhunia AK, Bhattacharyya A, Banerjee P, Roy PP, Murala S (2018) A novel feature descriptor for image retrieval by combining modified color histogram and diagonally symmetric co-occurrence texture pattern. arXiv preprint arXiv :1801.00879
26. Varish N, Pal AK (2018) A novel image retrieval scheme using gray level co-occurrence matrix descriptors of discrete cosine transform based residual image. Appl Intell 48(9):2930–2953
27. Yousuf M, Mehmood Z, Habib HA, Mahmood T, Saba T, Rehman A, Rashid M (2018) A novel technique based on visual words fusion analysis of sparse features for effective content-based image retrieval. Math Probl Eng. https:// doi.org/10.1155/2018/2134395
28. Wagstaff K, Cardie C, Rogers S, Schrödl S et al (2001) Constrained k-means clustering with background knowledge. ICML 1:577–584
29. Doğantekin A, Özyurt F, Avcı E, Koç M (2019) A novel approach for liver image classification: PH-C-ELM. Measurement 137:332–338
30. Alkhawlani M, Elmogy M, Elbakry H (2015) Content-based image retrieval using local features descriptors and bag-of-visual words. Int J Adv Comput Sci Appl 6(9):212–219
31. Bhunia AK, Bhattacharyya A, Banerjee P, Roy PP, Murala S (1801) A novel feature descriptor for image retrieval by combining modified color histogram and diagonally symmetric co-occurrence texture pattern. Pattern Anal Appl. https://doi.org/10.1007/s10044-019-00827-x
32. Huang, W-q, Wu Q (2017) Image retrieval algorithm based on convolutional neural network. In: Current trends in computer science and mechanical automation, Sciendo Migration, Vol 1, pp 304–314
33. Mehmood Z, Mahmood T, Javid MA (2018) Content-based image retrieval and semantic automatic image annotation based on the weighted average of triangular histograms using support vector machine. Appl Intell 48(1):166–181