International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2016): 79.57 | Impact Factor (2015): 6.391

CBIR-MSVM: Content-based Image Retrieval using Multi-Labelled Support Vector Machines

Justiner Joseph[1], Xuewen Ding[2]

Tianjin University of Technology and Education, Department of Electronics Engineering, 1310,
Dagu South Road, Hexi District, Tianjin, P.R. China

Abstract: ***Content Based Image Retrieval (CBIR) technique is the emergent application to extract the appropriate query based images. But, the query based extraction is the one of complicated task for reducing the classification accuracy. To overcome these issue, proposed the CBIR based Multi-labelled Support Vector Machine classifier is used to enhance the classification outcomes. The preprocessing stage is processed into two main forms such as image resizing and the image filtering. In this framework, Gaussian Filtering technique is performed to remove the unwanted features and filter the relevant content based features. Then, three feature extraction process are as color, shape, and the texture feature are extracted based on the Color Histogram, REGIONPROPS, and the Grey Level Co-occurrence Matrix (GLCM). The Color Histogram technique is utilized to remove the unwanted RGB based structures from the results of preprocessed image and applying REGIONPROPS shape feature to extract the specific area and the perimeter based shapes. Then, performing GLCM texture based feature extraction to extract the statistical related features. Among the extracted features, the similarity computation process is accomplished to classify the content based images. Finally, MSVM classifier is processed to classify the content based pictures. The presentation result of the proposed framework is predicted with the help of parameters such as precision, specificity, recall, sensitivity, and the classification accuracy. Hence, the proposed research work is superior to the other existing techniques.***

Keywords: Content Based Image Retrieval, Multi-labelled, Histogram, Texture feature, Shape, Color

1. Introduction

Content based image retrieval (CBIR) [1, 2] framework have been termed as the query by image content (QBIC) technique and its generally used to depict the way toward recovering the preferred pictures from the substantial gathering on the premise of features, (for example, shading, surface and shape) [3] that can be consequently removed from the pictures themselves. The features utilized for recovery can be either primitive or semantic, however the extraction procedure must be prevalently programmed. The CBIR system can be classified into two forms such as text query and the image query based extraction. There are some techniques of image retrieval are as: Relevant Feedback (Human Interaction), Semantic-based image recovery, Texture Feature Extraction, Extraction of Color, Web Based Image Retrieval, Low Level Image Features, and the Wavelet-Based CBIR.

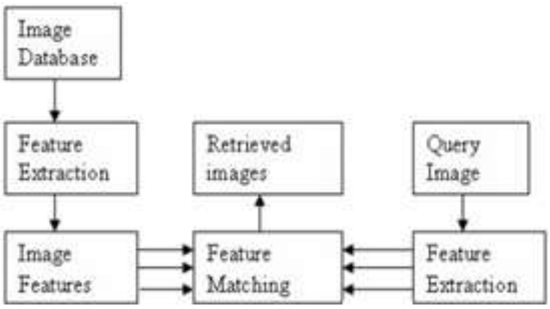

Figure 1: CBIR Work Flow

Image processing covers a much wider field, including image enhancement, compression, transmission, and interpretation. While there are grey areas (such as object recognition by feature analysis), the distinction between mainstream image analysis and CBIR is usually fairly clear-cut. Fig. 1 represents the overall CBIR work flow. The two major important tasks performed in the CBIR system to extract the relevant image are as follows: feature extraction, and the similarity measurement technique. Therefore, the various feature extraction process and the corresponding drawbacks are as follows. The multiple support vector machines [4] is used to ensemble the content based images. This MSVM technique effectively classified the query based images that mainly depends on the feature extraction process. The Daubechies wavelet transformation is used for extracting the feature vectors of images. The classification accuracy of this suggests system is the major problem. The color histogram procedure and the spatial orientation tree (SOT) [5] is used to calculate the vector point of an images.
The input image is subdivided into number of blocks then the wavelet coefficient of an each low pass vector point are built by an SOT. Subsequently, color histogram features are composed fromeach sub-block. After that, construct the vocabulary tree of an each vector point in an image. The Ordered-Dither Block Truncation Coding (ODBTC) [6] feature technique is employed to descripts the relevant image content. In the plan, the Bit Pattern Features (BPF) and the Color Co-occurrence Features (CCF) are computed by utilizing the bitmap picture and the shading quantizers in an ODBTC. Furthermore, it can assess the comparability level. The precision rate is the major problem in this truncation Coding feature technique.

1.1 Objectives

The chief contribution of this research work is as follows:

- To retrieve the similar images from the database.
- To abstract the features among the color features, texture features and the shape features that depends on the extraction process is implemented.
- To increase the retrieval processing accuracy.

1.2 Organization

The rest of the areas of this paper is sorted out as takes after: Segment 2 audits a portion of the existing works identified with the CBIR. Segment 3 gives the definite depiction of the general proposed framework structure. Segment 4 exhibits the execution consequences of the proposed framework. At last, this paper is concluding in the segment 5.

2. Literature Survey

This section illustrates the different traditional image retrieval techniques and its drawbacks. *Poursistani, et al*. [7] offered the new, operative appearance indexing and the retrieval techniques. In a JPEG compressed images, extract the relevant features based on the codebook generated by using a K-means clustering procedure and the Vector Quantization (VQ) techniques. The extracted features based on the JPEG components of the histogram from a DCT coefficients. The efficacy was the major drawbacks in this offered system.*Wang, et al. [8]* presented a novel hashing approach, order preserving hashing (OPH), and the approximate nearest neighbor (ANN) search algorithm. In between the original and hamming space they estimate the similarity orders and maximized the alignment was the main role of learning hash functions. The Sigmoid relaxation, active set, and the stochastic gradient descent were employed to solve the problems of multiple binary-class classification and corrected the misalignment rate.

It does not guarantee to produce a good nearest neighbor. *Subrahmanyam, et al.* [9] offered the modified color motif co-occurrence matrix (MCMCM) to recover the substance based pictures. The red, green, and blue shading planes between the connection were gathered in this plan. The offered strategy incorporates the MCMCM and the (DBPSP) features with indistinguishable heaviness in difference to the framework that absorbs the matrix features. Therefore, the DBPSP with the color histogram and additionally required a k-mean highlights to improve the weights of a features. *Vipparthi and Nagar* [10] recommended the color directional local quinary patterns (CDLQP) for the extraction of color texture feature. The neighborhood directional edge information with grey level differences among the binary and ternary pixels were employed to extract the texture features and then retrieval the content based images. *Guo and Prasetyo* [11] suggested the Halftoning based block truncating coding feature extraction for the reclamation of an query result.

The ODBTC compacted informations stream are used to extract the significant feature by an index representation. Except the color quantizes and the bitmap image features were removed based on the dual processes such as Bit Pattern Feature (BPF) and the Color Co-occurrence Feature (CCF). *Guo, et al*. [12] presented an Error Diffusion Block Truncation Coding (EDBTC) for retrieving the content based images. In this technique, the feature descriptors are generated by utilizing theVector Quantization (VQ). In between target and the query image to evaluate the similarity based on the two different feature extraction techniques were introduced such as CHF and the BHF. The VQ-indexed bitmap image and the VQ-indexed color quantizer were employed to compute the CHF and BHF. After predicted the distance by utilized the integration method of CHF and BHF which shown the similarity. *Li, et al.* [13] suggested a novel Robust Structured Subspace Learning (RSSL) algorithm for representing the data.

The image identifying, extraction of feature, and correlating the extracted features were all considered in this suggested learning algorithm. The formulation methods include the feature extraction to remove the unwanted pixel information. Finally, optimization algorithm was performed to improve the classification effectiveness of the suggested algorithm. Hence, the multi-subspace setting utilized the sparse representation which has not been sufficiently answered or explored.

3. Proposed Method

This section deliberates the thorough explanation of the proposed CBIR system using Multi-label Support Vector Machine (MSVM) classifier. The main intention of this work is to extract the relevant image from the large database. The overall process of the proposed workflow is shows in Fig 2, which includes the following stages,

- Image Acquisition
- Preprocessing
- Feature Extraction
- Classification

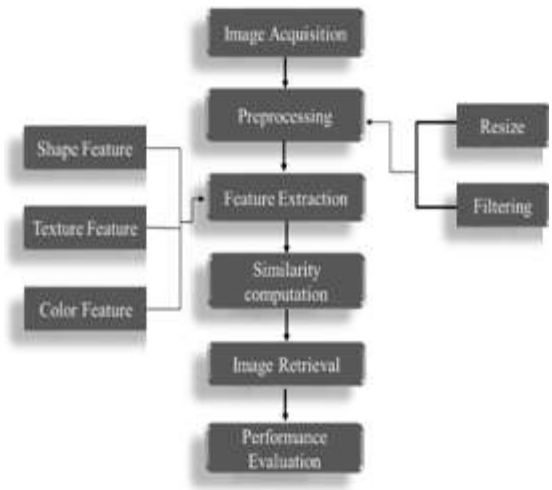

Figure 2: Proposed System Work Flow

3.1 Image Acquisition

Image acquisition is one of the initial process in image processing technique which defined as the progression of retrieving an image from the large dataset. A picture is an arrangement of rectangular exhibit of qualities (pixels). Every pixel information symbolizes the property of an image that estimate within a particular format. There are absolutely 138 pictures are taken for testing the performance. The possessions of an image which includes numerous layers such as red, green and blue filters (three values) that shows the image into limited brightness level as well as brightness. The qualities are typically spoken to by an 8 piece whole number, generous a scope of 256 levels of brilliance. Fig. 3 represents the case diagram.

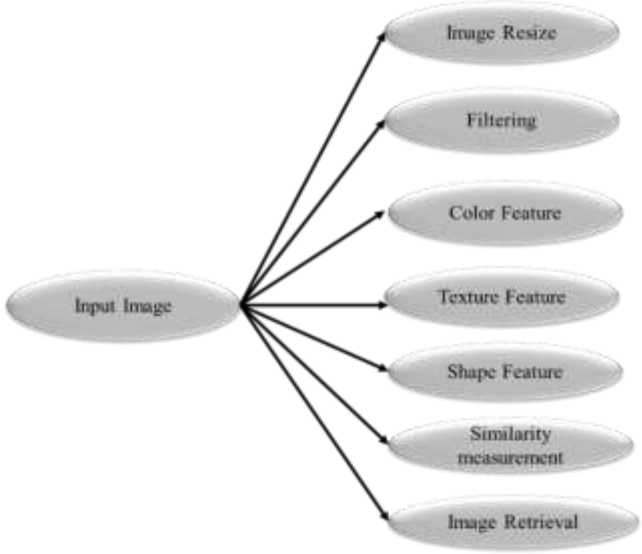

Figure 3: Proposed Case Diagram

3.2 Preprocessing

After image acquisition process, pre-processing stage is implemented to extract the relevant features based on the user query image. In this pre-processing stage, there are two different process is performed to remove the unwanted features.

- Image resizing
- Image Filtering

Primarily, image resizing is the process of reducing the input image size without loss of pixel information. If the image is reduced or enlarged, then the image looks like a “soft” or “jagged”. In order to protect the edge information of the input image is protected with the help of Gaussian filter. Secondarily, the resized image is given to the input of Gaussian filtering technique to filter the unwanted pixels and extract the content based pixel. The GF is used to remove the noisy pixel and blur the image. The mathematical derivation of GF as,

$$GF(x) = \frac{1}{\sqrt{2\pi\rho^2}} e^{-\frac{x^2}{2\rho^2}} \quad (1)$$

Where, ρ represents the standard deviation of an image. The filtered image is given to the feature extraction process.

3.3 Feature Extraction

Feature Extraction is the procedure of removing some features and extract one or more features based on the measurements of some quantifiable property of a preprocessed image. It can be categorized into three forms of features extraction such as

- Color Features
- Shape Features
- Texture Features

3.3.1 Color Features

First of all, the color features extraction is one of the most important process in image processing. The RGB (red, green, blue) based outlines are very much useful in the color histogram to analyze the image in an order. A color histogram CH is performed with the input of an Gaussian filter result image and defined in the vector form as,

$$CH = \{CH[1], CH[2], \dots . CH[i], \dots CH[N]\} \quad (2)$$

Where i represents the color in the color histogram, $CH[i]$ represents the number of pixels in color i in an image, and the N represents the number of bins in the color histogram. To compare the different size images normalized color histogram is used as,

$$CH'[i] = \frac{CH[i]}{ij} \quad (3)$$

Here i, j represents the total quantity of pixel in an image. The color histogram. Each image have derive the color histogram is known as feature vector. It removes the size of an image. The extracted R, G, B color image as given to the shape feature extraction process.

3.3.2 Shape Features

Secondarily, the shape based feature extraction process is developed for measuring the similitude between the shapes represented by their user query. Shape content description is hard to characterize on the grounds that measuring the closeness between shapes is troublesome. In this manner, two stages are fundamental fit as a fiddle based picture recovery, such as feature withdrawal and the similarity evaluation among the features. The RegionProps is a shape feature process to measure the properties of an image regions. The various shape measurements such as 'Area', 'Euler Number', ‘Diameter’, 'Orientation', 'Extent', 'Perimeter', 'Centroid', 'Extrema', 'SubarrayIdx', etc. In this measurements, the area and the perimeter measure are used to derive the shape features.

The Pixel Value Measurements requires the grayscale image as an input such as 'Mean Intensity', 'Max Intensity', 'Weighted Centroid', 'Min Intensity', and the 'Pixel Values'. The section is characterized as the genuine quantity of white pixels in the picked object province. The protest area was ascertained by tallying the quantity of white or '1' pixels inside the object limit. Therefore, the mathematical derivation of area is defined as,

$$radii = diameters/2 \quad (4)$$

$$Diameters = 2 * radii$$

The perimeter of object is defined as the summation of the separation between each connecting pair of pixels around the object border (B). The mathematical formulation of perimeter is derived as,

$$delta = diff(B).\hat{}2; \quad (5)$$

$$perimeter = sum(sqrt(sum(delta, 2))); \quad (6)$$

Therefore, the area and perimeter have been derived. After, implement the texture feature extraction process.

3.3.3 Texture Features

Texture feature is another important property of an image which has been utilized to characterize and perceive the items and also a part of discovering similitudes between the pictures. A GLCM technique is employed to portray the surface features and it processes the scalar estimations of an image. It alludes to the direct relationship among the present picture pixel-i and the reference pixel-j which is available in the extricated shape feature image. A GLCM is defined by first identifying a displacement vector $d = (dx, dy)$ and counting all pairs of pixels isolated by d having gray levels between i and the j. A co-occurrence matrix is a twodimensional cluster in which both the rows and the columns represent an arrangement of conceivable image esteems. The scalar features which are extracted by utilizing the factual measures, for example, contrast, cluster prominence, correlation, cluster shade, dissimilarity, entropy, homogeneity, energy, maximum probability, auto correlation, sum of squares, and the inverse different Moment. Let G be the number of gray-values in the image, then the dimension of the co-occurrence matrix $CM(i,j)$ will be N × N.

The co-occurrence matrix texture feature is reduced the feature dimensionality and lists the following some features as,

$$Cont = \sum_i \sum_j (i-j)^2 CM(i,j)^n \quad (7)$$

$$Homegenity = \sum_i \sum_j \frac{CM(i,j)}{1+|i-j|} \quad (8)$$

$$Engy = \sum_i \sum_j CM(i,j)^2 \quad (9)$$

$$Entro = -\sum_i \sum_j CM(i,j) \log CM(i,j) \quad (10)$$

$$Max\ Prob = \max_{i,j} CM(i,j) \quad (11)$$

$$Cluster\ prominence = \sum_{i=0}^{CM-1} \sum_{j=0}^{CM-1} \{i+j-\mu_x - \mu y4 \times CM(i,j) \quad (12)$$

$$Inertia = \sum_i \sum_j (i,j)^2 CM(i,j) \quad (13)$$

$$Correlation = \frac{\sum_i \sum_j (i,j) CM(i,j) - \sigma_i \sigma_j}{\mu_i \mu_j} \quad (14)$$

$$Dissimilarity = \sum_{i,j=1}^{CM} CM_{ij} |i-j| \quad (15)$$

$$Diff\ moment = \sum_i \sum_j \frac{1}{1+(i-j)^2} CM(i,j) \quad (16)$$

Where,

$$\sigma_i = \sum_i i \sum_j CM(i,j) \quad (17)$$

$$\sigma_j = \sum_j j \sum_i CM(i,j) \quad (18)$$

$$\mu_i = \sum_i (i-\sigma_i)^2 \sum_j CM(i,j) \quad (19)$$

$$\mu_j = \sum_i (j-\sigma_j)^2 \sum_j CM(i,j) \quad (20)$$

The extracted features is characterized into the matrix form which includes the number of rows and columns that referred as the gray levels features. The features obtained from the GLCM technique are finally concatenated to extract the content based features.

3.4 Similarity Computation based Classification

The similarity computation is the main task of the image recovery as well as classification process for which calculating the similarity distance for the adjacent neighbor estimation between the two images. The similarity distance estimation is mainly depending on the performance of the classification results. If the system completes the estimation of similarity distance and it produce a conventional group of images based on the increasing order of the similarity distance score. In the IR process, the images are searched by their general image description model representation. The model is based on content based similarity retrieval. The model language is a sophisticated window-based graphical interface. The user interface supports the visual expression of a query and allows query refinement and manipulation of the results. The model supports retrieval by attribute values, color, shape, texture, pictorial example similarity and spatial constrains. The experimental results of the MSVM technique is shows in the figure 4.

Figure 4 (a): Input image

Figure 4(b): Gaussian filtered

Figure 4 (c): Classified results

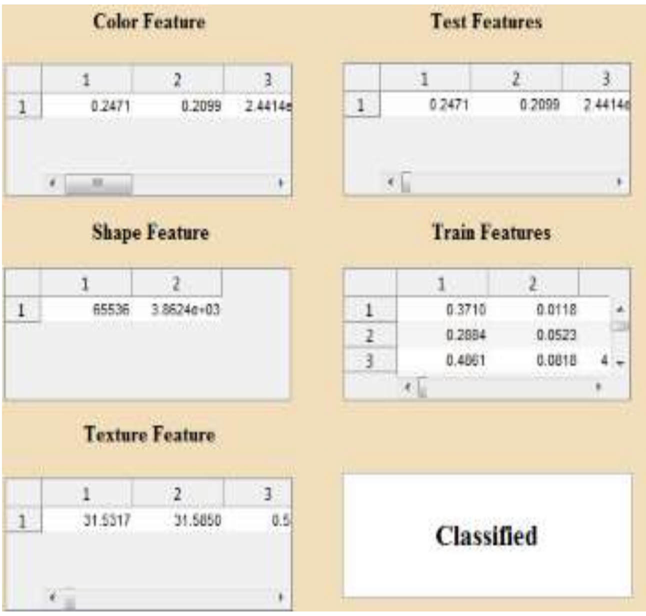

Figure 4(d): Extracted features lists

Figure 4: (a) input image, (b) Gaussian filtered image, (c) classified results, and (d) extracted features lists

The image retrieval process is performed based on the multi-labeled SVMs (MSVM) which utilized the multi hyper planes to separate the training data based on the eight classes of grade level. The MSVM are used for classification and for reducing the dimensionality of the feature vectors effectively. There are 138 images are considered as the testing cases.

The vector features are categorized into $C1, C2, C3, C4, C5, C6, C7$, and the $C8$ respectively based on their hyper-planes. The 8 categories is shows under the following classifications as,

C1 as elephant images

$C1 = \{'\text{Name}'\ '\text{Place}'\ '\text{Length}'\ '\text{Lifespan}'\ '\text{Mass}'\ '\text{Gestation Period}'\ '\text{Height}'\};$

C2 as Rose flower images

$C2 = \{'\text{Name}', '\text{Kingdom}', '\text{Division}', '\text{Class}', '\text{Order}', '\text{Family}', '\text{Subfamily}', '\text{Genus}'\};$

C3 as vehicle images

$C3 = \{'\text{Name}', '\text{Length}', '\text{Width}', '\text{Height}', '\text{Mass}'\};$

C4 as dinosaur images

$C4 = \{'\text{Name}', '\text{Scientific Name}', '\text{Higher Classification}', '\text{Phylum}', '\text{Subclass}'\};$

C5 as horse images

$C5$ ={'Name', 'Scientific Name', 'Lifespan', 'Speed', 'Mass', 'Gestation period'};

C6 as food item images

$C6$ ={'Name'};

C7 as male images

$C7$ ={'Personal_Name', 'Place' 'Age' 'Academic' 'School' 'Year' 'Occupation' 'Associate Professor' 'Profession'};

C8 as the female images

$C8$ ={'Personal_Name', 'Place' 'Age' 'Academic' 'School' 'Year' 'Occupation'};

These are the classification process performance follows under the CBIR techniques. Figure 5 represents the overall process of Image Retrieval diagram.

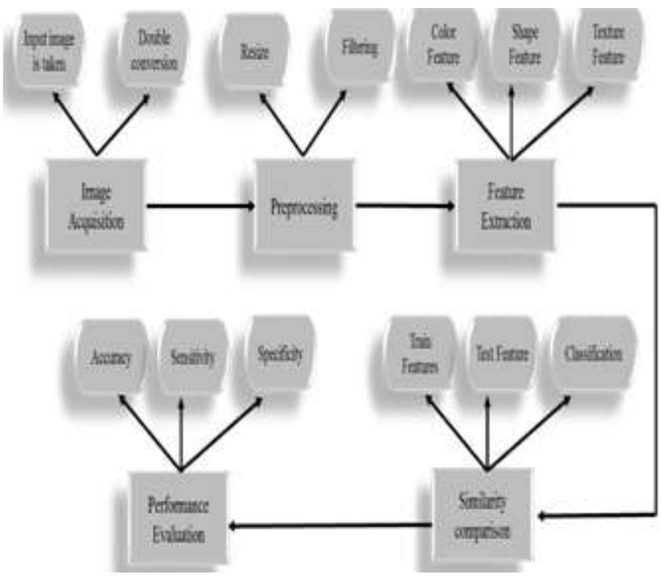

Figure 5: Image Retrieval diagram

4. Performance Analysis

This division offers the results of the projected CBIR system using MSVM classifier. The results are analyzed and evaluated in terms of,

- Accuracy
- Sensitivity
- Specificity
- Precision
- Recall
- F-measure
- Geometric-mean

4.1 Accuracy, Sensitivity, Specificity

Accuracy is well-defined as the nearness of the standard fixed significance and it estimate the efficiency of the proposed system.

$$Acc = \frac{(TP+TN)}{(FP+TN)+(TP+FN)} \quad (21)$$

Sensitivity of the proposed MSVM classifier is defined as the process of calculating the correct classification results.

$$Sen = \frac{TP}{TP+FN} \quad (22)$$

Specificity of the proposed MSVM classifier is defined as the process of calculating the wrongly classified results rejects ratio.

$$Spec = \frac{TN}{TN+FP} \quad (23)$$

Where, TP characterizes the correctly identification values, TN characterizes the correctly identified wrong values, FP characterizes the false positive values, and the FN represents the false negative values.

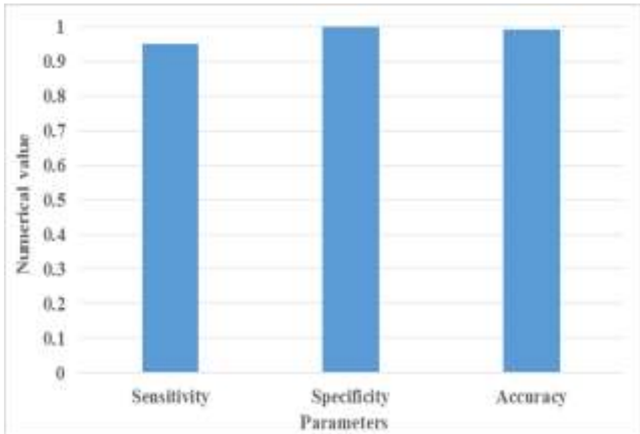

Figure 6: Accuracy, specificity, sensitivity

From the figure 6 shows that the proposed MSVM classifier experimental analysis results. The graphical structure values are illustrates: sensitivity as 0.9500, specificity as 1, and the accuracy as 0.992. Hence, the proposed system achieves greater performance.

4.2 Precision, Recall, F-measure, Geometric Mean

Precision is characterized as the division of recovered record that are applicable to the user inquiry. Precision considers all recovered reports, however it can likewise be assessed at a given cut-off rank, considering just the highest outcomes returned by the framework.

$$Pr = \left\{ \frac{\{relevant\ \ image\ \cap retreived\ \ document\ \}}{retr\ eived\ \ document} \right\} \quad (24)$$

Recall is characterized as the portion of the important reports that are effectively recovered.

$$Re = \left\{ \frac{\{relevant\ \ document\ \ \cap retreived\ \ document\ \}}{relevant\ \ document} \right\} \quad (25)$$

The f-measure is characterized as the weighted consonant mean of the exactness and the review of the test precision.

The geometric mean is defined as a sort of mean that characterizes the focal propensity or commonplace estimation of an arrangement of pixels by utilizing the result of their esteems.

$$Gmean = \sqrt{TPR * TNR} \quad (26)$$

Where, the G-mean is measured with the base of TPR and the TNR. TPR characterizes the correctly identified the positive result rate, and the TNR characterizes the correctly identified the wrong prediction result rate.

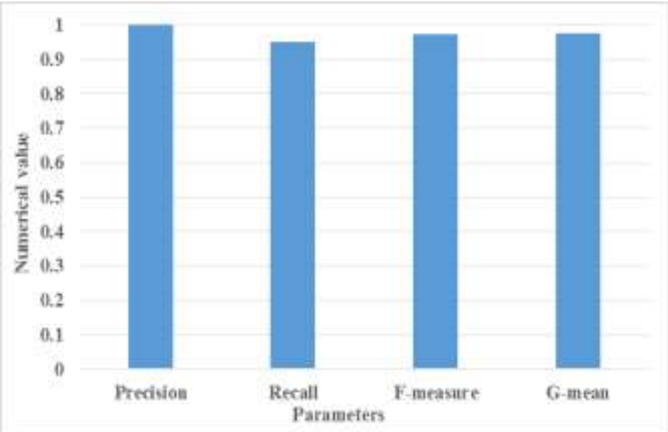

Figure 6: Different parameters

5. Conclusion

In this research proposed, the CBIR based MSVM classifier framework for retrieving the content based images. In this paper focuses to retrieve the similar images from the large dataset images. Initially, preprocessing step is processed to filter the unwanted image pixels and also the system includes three feature extraction modules, such as Color Histogram, Region Props based shape feature, and the GLCM based texture extraction. In this scheme, added the three extracted feature vectors and estimate the similarity based distance vectors. Finally, the adjacent content based images are retrieved successfully with the help of MSVM classifier. This technique experimental results are analyzed based on the parameters such as specificity, precision, sensitivity, f-measure, recall, g-mean, and the accuracy. The MSVM technique produces 99.2% accuracy rate. In future, the proposed system will be extended to limit the feature extraction process.

References

[1] P. M. Kulkarni and P. S. A. Bhalotra, "CONTENT BASED IMAGE RETRIEVAL REVIEW," 2014.

[2] D. S. Shete, M. Chavan, and K. Kolhapur, "Content Based Image Retrieval," InternationalJournal of Emerging Technology and Advanced Engineering, vol. 2, pp. 85-90, 2012.

[3] M. Meharban and S. Priya, "A Review on Image Retrieval Techniques," Bonfring International Journal of Advances in Image Processing, vol. 6, p. 7, 2016.

[4] E. Yildizer, A. M. Balci, M. Hassan, and R. Alhajj, "Efficient content-based image retrieval using multiple support vector machines ensemble," Expert Systems with Applications, vol. 39, pp. 2385-2396, 2012.

[5] M.Subrahmanyam, R. Maheshwari, and R. Balasubramanian, "Expert system design using wavelet and color vocabulary trees for image retrieval," Expert Systems with Applications, vol. 39, pp. 5104-5114, 2012.

[6] J.-M. Guo and H. Prasetyo, "Content-based image retrieval with ordered dither block truncation coding features," in Image Processing (ICIP), 2013 20th IEEE International Conference on, 2013, pp. 4006-4009.

[7] P. Poursistani, H. Nezamabadi-pour, R. A. Moghadam, and M. Saeed, "Image indexing and retrieval in JPEG compressed domain based on vector quantization," Mathematical and Computer Modelling, vol. 57, pp. 1005-1017, 2013.

[8] J. Wang, J. Wang, N. Yu, and S. Li, "Order preserving hashing for approximate nearest neighbor search," in Proceedings of the 21st ACM international conference on Multimedia, 2013, pp. 133-142.

[9] M. Subrahmanyam, Q. J. Wu, R. Maheshwari, and R. Balasubramanian, "Modified color motif co-occurrence matrix for image indexing and retrieval," Computers & Electrical Engineering, vol. 39, pp. 762-774, 2013.

[10] S. K. Vipparthi and S. K. Nagar, "Color directional local quinary patterns for content based indexing and retrieval," Human-Centric Computing and Information Sciences, vol. 4, p. 6, 2014.

[11] J.-M. Guo and H. Prasetyo, "Content-based image retrieval using features extracted from halftoning-based block truncation coding," IEEE Transactions on image processing, vol. 24, pp. 1010-1024, 2015.

[12] J.-M. Guo, H. Prasetyo, and J.-H. Chen, "Content-based image retrieval using error diffusion block truncation coding features," IEEE Transactions on Circuits and Systems for Video Technology, vol. 25, pp. 466-481, 2015.

[13] Z. Li, J. Liu, J. Tang, and H. Lu, "Robust structured subspace learning for data representation," IEEE transactions on pattern analysis and machine intelligence, vol. 37, pp. 2085-2098, 2015

Author Profile

Justiner Joseph received the Bachelor degree in Electronics and Communication Engineering from St. Joseph University in Tanzania in 2014. During November, 2014 - August,2015, She worked as Tutorial assistant at St. Joseph University in Tanzania under department of Physics. She is now studying M.S. degreein Signal and Information Processing Engineering at Tianjin University of Technology and Education.