Journal of Computer Science

Original Research Paper

Multimedia Image Retrieval Using Novel Modelling of Combined Binary Patterns with Reduced Dimensions

M. Geetha Yadav and SP. Chokkalingam

Department of Computer Science and Engineering, School of Computing, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

Article history
Received: 21-09-2024
Revised: 11-12-2024
Accepted: 10-01-2025

Corresponding Author:
M. Geetha Yadav
Department of Computer Science and Engineering, School of Computing, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India
Email: geethayadav22@gmail.com

Abstract: Advancements in computer vision have transformed image retrieval. However, many existing systems struggle to bridge the semantic gap between low-level features and high-level user expectations. This study proposed a new approach that combines texture and colour features to enhance retrieval accuracy and user satisfaction. The proposed method employs the HSV colour model to extract colour features, focusing on dominant features from Hue (H) and Saturation (S) components at different levels of histogram bins, while texture features are derived from the Value (V) component using the Relational Edge Patterns (REPs). These patterns, generated by analyzing binary relationships in all directions between centre pixels and their neighbours for every 5'5 matrix, effectively capture the texture properties of objects in images. After concatenating colour and texture features, Principal Component Analysis (PCA) is applied to reduce dimensionality and retain the most discriminative features. The effectiveness of the proposed method is evaluated on the Corel-10k and 102-flowers datasets, demonstrating superior precision and recall compared to existing methods. Experimental results highlight its ability to form semantically meaningful clusters, bridging the semantic gap and enabling efficient image retrieval.

Keywords: Image Retrieval, HSV Model, Relational Edge Patterns (REP), Texture Features, Corel-10k, 102-Flowers Dataset

Introduction

Image retrieval is a critical technology that enables efficient organization and access to vast collections of visual data, making it indispensable for various industrial applications. In e-commerce, image retrieval systems enhance user experience by allowing customers to search for visually similar products, such as clothing or furniture, using images instead of keywords. In healthcare, these systems facilitate quick and accurate retrieval of medical images, aiding in diagnostics and treatment planning. Additionally, in digital asset management, industries like media and entertainment rely on image retrieval to organize, index and access large-scale image libraries. With advancements in feature extraction and similarity computation, modern image retrieval methods are increasingly robust and scalable, making them suitable for addressing real-world challenges across diverse industries.

Searching for similar images in large repositories poses a significant challenge for current technological advancements. With the exponential growth in image usage across various fields, both online and offline image repositories are expanding rapidly. However, retrieving similar images based solely on keyword annotations has numerous limitations, resulting in low accuracy. In this study, we take up these challenges and recommend a novel approach to improve the effectiveness of searching for similar images. In order to address the limitations of keyword-based image retrieval, the field of Content-Based Image Retrieval (CBIR) has emerged, focusing on retrieving similar images based on the content of the query image. This involves extracting relevant features such as colour, texture and shape from the query image. The literature (Smeulders *et al*., 2000; Liu *et al*., 2007; Pourghassem and Ghassemian, 2008) provides detailed insights into these techniques. CBIR encounters various obstacles that require solutions, such as developing efficient methods to represent visual content, performing both manual and automatic image annotation, closing the semantic gap between low-level image features and high-level semantics and resolving issues related to image ranking. Two key aspects require attention to enhance the accuracy of image retrieval systems: Extracting efficient features from images and minimizing the semantic gap.

Low-level features, including colour, texture and shape, play a vital role in effectively representing the object information contained within an image. These features provide valuable details about the visual characteristics of objects, enabling accurate image retrieval. Enhancing the accuracy of image retrieval systems relies on the proper extraction and utilization of these low-level features. Moreover, reducing the semantic gap, which refers to the mismatch between low-level features and high-level semantics, is crucial. By bridging this gap, images can be represented in a more meaningful way, aligning the extracted features with the intended semantics of the images and improving retrieval accuracy (Sucharitha *et al*., 2023b).

Colour features are essential in effectively representing colour image information. Among various colour models, RGB, HSV and YCbCr are commonly used, with HSV closely resembling human perception of colour and exhibiting robustness against lighting variations. Consequently, HSV outperforms other colour models. Statistical metrics such as colour correlograms, colour histograms, moments, covariance matrix, etc., are employed for extracting colour characteristics (Pass and Zabih, 1999; Srivastava *et al*., 2015; Malviya and Ladhake, 2016). These techniques contribute to enhancing the representation of colour features for image retrieval purposes.

Various colour-based feature extraction techniques are utilized in image retrieval to capture important information related to image intensities and spatial relationships. The process of generating a feature vector through a colour histogram involves tallying the frequencies of each intensity within various colour channels (Varnish and Pal, 2015). On the other hand, a colour correlogram depicts the spatial correlation among image intensities (Huang *et al*., 1997). The Color Coherence Vector (Salmi and Boucheham, 2014) is a feature that combines the coherence and incoherence characteristics of pixel colours to construct a feature vector. It incorporates information about the relationship and consistency of colours within an image. These techniques play a significant role in extracting meaningful colour-based features for image retrieval purposes. In image analysis, several spatial feature extraction methods have been introduced to enhance the understanding of images. Among these methods, the Motif Cooccurrence Matrix (Jhanwar *et al*., 2004) stands out, as it generates a 3D matrix that represents the local statistics of the image. This matrix effectively captures patterns and relationships among motifs, enabling valuable insights into the spatial distribution of features within the image. Furthermore, in Avula *et al*. (2014), the mean pixel intensity is utilized as a technique to identify cancer regions in an image. These spatial feature extraction methods play a crucial role in uncovering important information and patterns within images, contributing to advanced image analysis techniques.

Texture is another important aspect of an image that helps differentiate it from others. Texture properties can distinguish images based on variations in local intensities related to surface characteristics such as coarseness, smoothness and regularity. Various mathematical transformation methods can be employed to extract texture features (He *et al*., 2009). Two notable transforms commonly used for image retrieval are Gabor filters (Idrissa and Acheroy, 2002) and wavelet packets (Laine and Fan, 1993; de Rivaz and Kingsbury, 2003) introduced a complex wavelet transform that addresses the computational complexity associated with Gabor wavelets when extracting texture features. Additionally, a modified curvelet transform for texture-based image retrieval was proposed by Gonde *et al*. (2013).

Ojala *et al*. (1996) introduced a powerful texture feature extractor known as the Local Binary Pattern (LBP), which exhibits excellent capabilities in extracting scale-invariant, rotation-invariant and texture features (Ojala *et al*., 1996). This method leverages the variations between the central pixel and its eight neighbouring pixels within a 3'3 pattern, utilizing a threshold to generate a binary pattern. In contrast, the Local Ternary Pattern (LTP), introduced in a subsequent work (Murala *et al*., 2012), extends this concept by employing an interval instead of a single threshold, resulting in the generation of two distinct binary patterns. In the context of image retrieval, researchers in Dubey *et al*. (2015; 2016) have investigated advanced adaptations of Local Binary Patterns (LBP) to achieve efficient texture feature extraction. Two specific variations, namely Local Bit Plane Decoded Pattern (LBDP) and Local Diagonal Extrema Patterns (LDEP), are examined. These techniques leverage Gabor filters and moments in combination with the extended LBP variants, resulting in an enhanced representation of texture features. By integrating these methods, superior performance is observed across a range of image retrieval applications. The integration of local and global features has been shown to yield significant performance enhancements in diverse applications. In Sucharitha and Senapati (2020), a method is presented that combines Local Directional Edge Binary Patterns (LDEBP) with Zernike moments to facilitate efficient medical image retrieval. By incorporating both texture and shape features of objects present in the images, this approach effectively improves precision and recall metrics. This combination of features enables more comprehensive and accurate retrieval of medical images.

In Nitin *et al*. (2023), the authors introduced a new approach named Mean-Variance-Median based LBP (MVM-LBP); the proposed approach incorporates additional statistical measures. The MVM-LBP descriptor utilizes the mean, variance and median values of the pixels within a 3'3 pixel window, resulting in a more powerful feature vector. By incorporating these statistics, the MVM-LBP descriptor enhances the overall feature representation. Karanwal and Diwakar (2021a-b) introduce a novel pattern known as Neighborhood and Center Difference Based LBP (NCDB-LBP). NCDB-LBP is computed for every 3x3 window within the image, resulting in the generation of four label codes: NCDBc for clockwise direction and NCDBac for anti-clockwise direction. By adopting this unique methodology, the NCDB-LBP pattern effectively captures local texture variations, thereby improving the system's retrieval capabilities. The authors also introduced another efficient binary pattern called OD-LBP for face recognition. In Sucharitha *et al*. (2023a; 2024), the authors introduced a new pattern RDEBP on the consideration of directional relation among the pixels. In this, the binary patterns are extracted with the directional relation among the neighbourhood pixels after evaluating their relation with a centre pixel. Four binary patterns are extracted, and the two-step security is discussed by transforming the watermark encrypted data to cloud security.

Local Feature Extractors

Local Binary Pattern (LBP)

de Rivaz and Kingsbury (2003) introduced the Local Binary Pattern (LBP) as a method for analyzing rotation-invariant texture features. This approach involves generating a binary pattern and its corresponding decimal number for each pixel in the image. To construct the binary pattern, a 3×3 matrix is selected, with the centre pixel being compared to its eight neighbouring pixels. The binary pattern is generated based on the sign of the differences between the centre pixel and its neighbours. This technique allows for the characterization of local texture variations in a rotation-invariant manner:

$$LBP_{n,t} = \sum_{i=0}^{n-1} Bp((I(P_k - P_c)) \quad (1)$$

where:

$$Bp(k) = \begin{cases} 0 \; if \; k < 0 \\ 1 \; if \; k \geq 0 \end{cases} \quad (2)$$

Here, n is the no. of the neighbourhood, and t is the radius of the neighbourhood.

Relational Edge Patterns (REPs)

The relational edge patterns represent a method to extract and encode spatial relationships between the centre pixel and its nearby neighbours in an image. The goal is to capture the local structure information to ensure the challenges of the scale-invariant and rotation-invariant of an image are met. A 5'5 matrix cantered on a pixel serves as the local analysis window, on which the edges are computed at two radii: R = 1(inner edges) and r = 2 (outer edges). The inner edges are calculated using the local differences between the centre pixel and its eight neighbours. In contrast, the outer edges are divided into two patterns to incorporate differences in $\pm 0^0$, $\pm 45^0$, $\pm 90^0$ and $\pm 135^0$ directions with pixels and their complementary values, as shown in Fig. (1b). Binary patterns are then constructed based on the incremental magnitude in clockwise and anti-clockwise directions. A binary digit “1” is assigned if a directional increment is opposite to the previous one, and '0' otherwise. These patterns are consolidated into REPs using mathematical formulations that capture local differences and directionality. By combining features from r = 1 and r = 2, REPs encode both local and global structures, offering a compact, rotation-invariant representation suitable for tasks like texture classification and image retrieval. The Eqs. (3-6) clearly express the procedure for the extraction of REPs.

The identification of variations in the continuity of texture patterns relies heavily on the edges present in the images. To create a feature vector database for all the images, we utilize histograms. Each pattern's histogram is computed and then concatenated to form the final feature vector for that specific image. This process is repeated for all images in the database, resulting in a comprehensive feature vector database:

$$RIE_1(I_C) = \sum_{n=1|r1}^{N=8} f\left(I_n - I_{(n+1)}\right) X \, 2^{n-1} \quad (3)$$

$$RIE_2(I_C) = \sum_{n=1|r1}^{N=8} f\left(I_n - I_{(1-n)}\right) X \, 2^{n-1} \quad (4)$$

$$ROE_1(I_C) = \sum_{n=1|r2}^{N=8} f(I_n - I_{(n+2)}) X 2^{n-1} \quad (5)$$

$$ROE_2(I_C) = \sum_{n=1|r2}^{N=8} f(I_n - I_{(2-n)}) X 2^{n-1} \quad (6)$$

$$ROE_3(I_C) = \sum_{n=2p+1|r2}^{N=8} f(I_n - I_{(n+2)}) X 2^{n-1} \; for \; p = 1,2,3,4,5,6,7 \quad (7)$$

$$ROE_4(I_C) = \sum_{n=2p+1|r2}^{N=8} f(I_n - I_{(2-n)}) X 2^{n-1} \; for \; p = 1,2,3,4,5,6,7 \quad (8)$$

$$f(x) = \begin{cases} 1 \; if \; f(x) \geq f(x+1) \\ 0 \; if \; f(x) < f(x+1) \end{cases} \quad (9)$$

Pseudo-Code

Take an image from the database

Procedure

Step 1: Initialize parameters

- For a given image, consider each 5×5 matrix 7 define centre pixel I_C
- Define Coordinates to observe local patterns
- Define. Inner and Outer radius matrices: r1 and r2

Step 2: Iterate over image pixels

For each pixel (ignoring boundary pixels)

- Set the iteration range:
  for ki=3: n-3 (rows)
  for li=3: n:3 (columns)

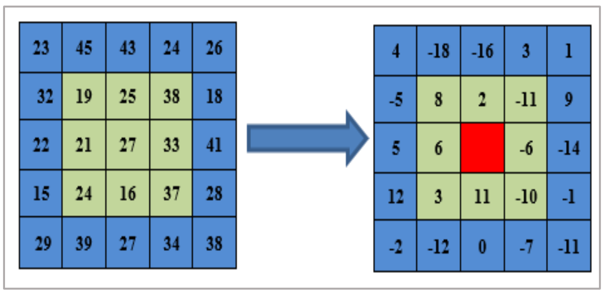
- For the RI and Ro matrix
- Identify and extract neighbouring pixels for both RI and Ro for every centre pixel I_C.

Step 3: Calculate clockwise and Anti-clockwise directional relations as

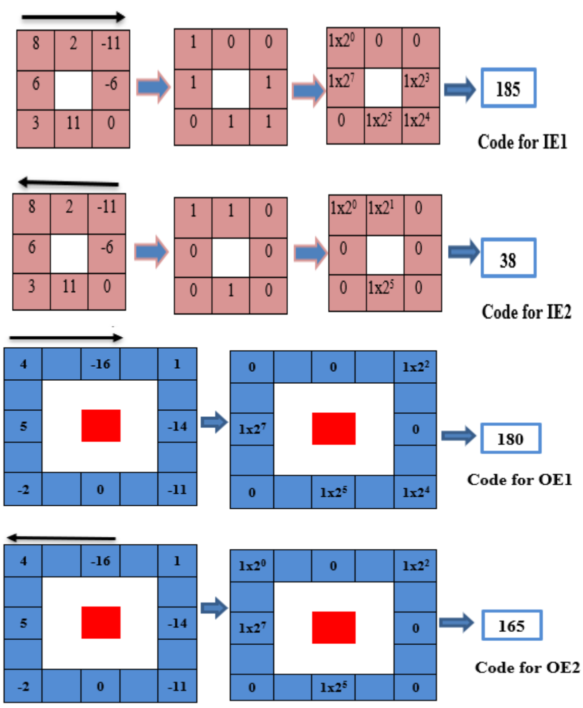
R_{IE1}, R_{IE2}, R_{OE1}, R_{OE2}, R_{OE3}, and R_{OE4} are shown in Fig. (1).

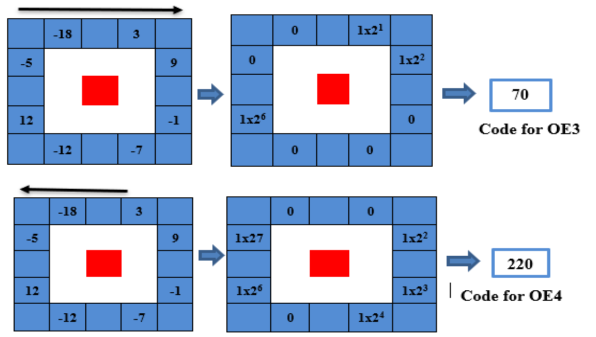
Step 4: Construct the Binary vectors BP_{IE1}, BP_{OE1} BP_{IE2} BP_{OE2}, BP_{OE3}, BP_{OE4}

Step 5: To construct a feature vector, calculate the histograms for each binary vector.



(a)





(b)

Fig. 1:(a) A 5×5 matrix from image and represented r = 1 and r = 2. (b) Procedure to calculate relative directional edge patterns

Principal Component Analysis (PCA)

PCA is one of the powerful statistical machine learning models used to reduce the dimensionality of data while retaining its most significant features. It converts a set of information of conceivably correlated variables into a set of values of completely uncorrelated variables termed principal components via an orthogonal transformation. The PCA is explained by considering a collective of n-dimensional vectors $I = \{[I_1, \quad I2, I_3, \ldots, I_n]^T\}$ whose dissemination is centred at origin $E(I) = 0$, where E is the anticipation operator. The covariance given in between every set of the variable is $I_{ij} = E\{(I_i - I_i')(I_j - I_j')\} = E\{I_i I_j\}$, a nxn covariance matrix can be repositioned with I_{ij} parameters:

$$I_x = E\{(I - I')(I - I')^T\} = E\{II^T\}$$

If $det(I_x) \neq 0$, then eigen vector decomposition can be applied to decompose I_x into a product of 3 matrices:

$$I_x = A\Lambda A^{-1}$$

where, $\Lambda = diag(\lambda_1, \lambda_2, \ldots \lambda_n)$ are known as eigen values and $A = [a_1, a_2, ., a_n]^T$ are the respective eigen vectors and $A^T A = I$ because A is the orthogonal matrix. As a result, the columns of A create a new orthogonal basis, which is a linear transformation of the original basis. The feature distributions can be whitened using the eigenvector decomposition by projecting the fundamental feature vector I on the eigenvector basis to attain the coordinates I, which is equal to the rotated feature basis. While PCA reduces dimensionality and retains significant features, the transformation does not guarantee that the resulting features have unit variance. Prewhitening can be applied to further refine the feature representation. Prewhitening ensures that the transformed data is uncorrelated and normalized, enhancing the interpretability and statistical properties of the features.

This is achieved by rescaling the coordinates of the PCA-transformed vector I using the eigenvalues λ_i. The whitened feature vector I_w is computed as:

$$I_w = A\Lambda^{-1/2}A^T I$$

where, $\Lambda^{-1/2}$ is a diagonal matrix with elements $1/\sqrt{\lambda_i}$.

This dimensionality reduction simplifies data representation, eliminates redundancy, and enhances interpretability, making PCA a key tool in applications such as feature extraction, image processing, and visualization. When combined with techniques like Relational Edge Patterns (REP), which encode pixel-level directional relations and capture local and global spatial features, PCA can further refine data by emphasizing statistically significant patterns and reducing noise, resulting in robust and efficient feature extraction for tasks such as texture classification, image retrieval and cross-modal learning as shown in Fig. (2).

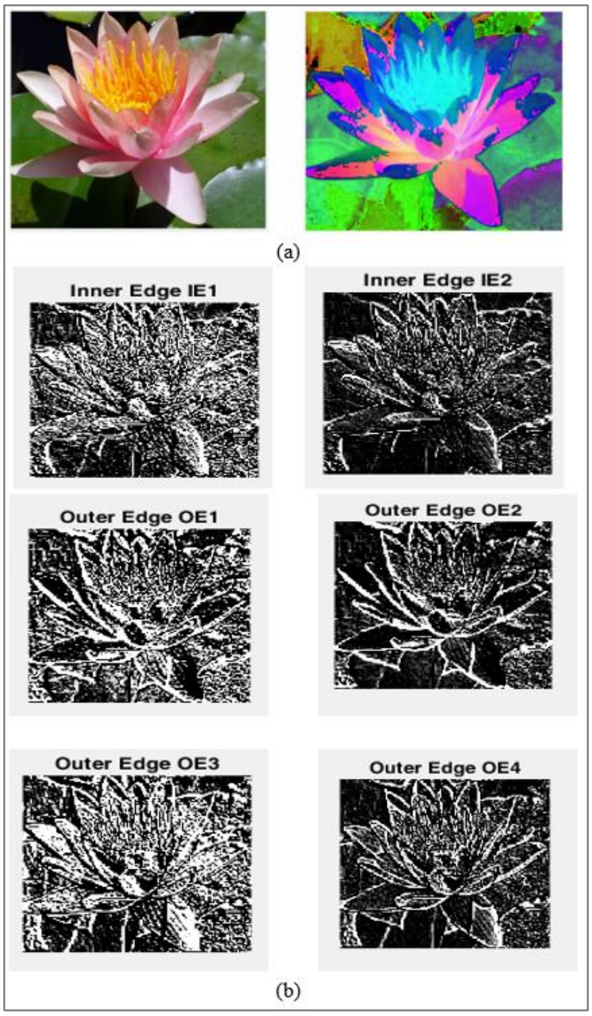


Fig. 2: (a) Original image and HSV image (b) All six relational edge patterns

Similarity Measurement

The image retrieval system heavily relies on the similarity metric, which governs the degree of similarity between the query image and the images within the database. By utilizing the computed similarity metric values, the system retrieves and ranks similar images from the database.

This method utilizes two distinct similarity metrics, namely the Euclidean distance and d1 distance, to gauge the similarity between the feature vectors of the query image and the images stored in the database. The Euclidean distance calculates the straight-line distance between the feature vectors, while the d1 distance computes the absolute difference between corresponding elements of the feature vectors. By utilizing these similarity metrics, the retrieval system can effectively identify and rank images based on their similarity to the query image. If the feature vector of the query image is $F_Q = \left(F_{q1}, F_{q2}, F_{q3} \dots\dots, F_{qn}\right)$ and the database images feature vectors are $F_{dB} = (F_{dB1}, F_{dB2}, F_{dB3}, \dots\dots, F_{dBn})$. The mathematical expressions for Euclidean and d1 distances are given in Eqs. (10-11):

$$Euclidian\ Distance\ D_E = \sqrt{\sum_{i=1}^{n}(F_{qi} - F_{DBi})^2} \quad (10)$$

$$Distance\ d_1 = \sum_{i=1}^{n} \left| \frac{F_{dB}(i) - F_{qj}(i)}{1 + F_{dB}(i) + F_{qj}(i)} \right| \quad (11)$$

where, F_Q is the query image features extracted using the proposed method, and F_{dB} is the feature vector database for all images of the database, $F_{dB1,}$ F_{dB2}, *etc.*, are feature vectors of individual images.

Evaluation Matrices

The evaluation metrics precision and recall are used to validate any retrieval systems due to their extensive applicability and efficiency in evaluating the retrieval performance. Precision measures the ratio of relevant images retrieved out of the total retrieved images, while recall measures the ratio of relevant images retrieved out of all relevant images in the dataset. Together, these metrics provide a comprehensive view of the system's ability to retrieve relevant results while minimizing irrelevant ones.

In the context of image retrieval, where the semantic gap poses significant challenges, precision and recall are particularly valuable as they emphasize the system's capacity to retrieve semantically meaningful results. High precision ensures that retrieved images align with user expectations, while high recall guarantees that relevant images are not overlooked. By evaluating the system using these metrics on benchmark datasets such as Corel-10 k and 102-flowers, we demonstrate the effectiveness of the proposed feature representation in addressing the semantic gap and achieving robust performance under varying conditions.

Rank-based evaluation methods are commonly used to assess image retrieval systems. In IR systems, prioritizing the top ranks for a given query image is essential. Several performance evaluation metrics are employed to evaluate the retrieval performance of such systems. Average Precision Rate (APR) calculates the average precision across all query images, considering the relevance of the retrieved images and their positions in the ranking list. It provides an overall assessment of the system's retrieval accuracy. Average Recall Rate (ARP) measures the average recall across all query images. It considers the ability of the system to retrieve relevant images, regardless of their position in the ranking. The precision-recall curve plots the precision against the recall values at different retrieval depths. It helps in analyzing the trade-off between precision and recall and provides a comprehensive evaluation of system performance. F-Measure is a metric that combines precision and recall into a single value, taking into account both metrics simultaneously. It provides a balanced assessment of the system's performance. By utilizing these evaluation metrics, the performance of medical image retrieval systems can be effectively assessed, enabling comparisons, optimizations and improvements in retrieval accuracy.

The specified evaluation metrics for a database are calculated using Eqs. (12-16):

$$Precision\ P = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ images\ retrieved} \tag{12}$$

$$Recall\ R = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ relevant\ images\ from\ the\ database} \tag{13}$$

$$ARP = \frac{1}{D_N} \sum_{i=1}^{D_N} P_i \tag{14}$$

$$ARR = \frac{1}{D_N} \sum_{i=1}^{D_N} R_i \tag{15}$$

$$F_{Measure} = \frac{2XARPXARR}{ARP+ARR} \tag{16}$$

According to the placements of the images, the precision and recall values are computed, and a precision-recall curve is created using these data. The precisions for all query images will be averaged to determine ARP.

Materials and Methods

Frame Work of Proposed Methods

The proposed method and its framework work are shown in Fig. (3). It has been explained in the sequence of the following steps:

1. In the initial stage, both the database and query images undergo a preprocessing step to ensure uniform size (dimensions) to 256'256 pixels to ensure consistency and facilitate feature extraction. To enhance image quality, a Gaussian filter is applied to remove any noise present in the images

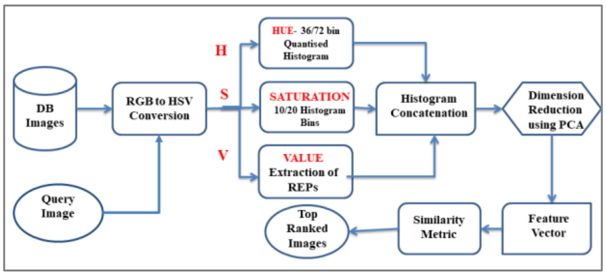


Fig. 3: Structure of proposed approach

2. After preprocessing, RGB images are converted to the HSV colour model, which separates chromatic content (Hue and Saturation) from intensity (Value). The HSV model was chosen for its robustness in light variations and its alignment with human visual perception. This separation allows for the independent extraction of colour and texture features
3. For individual histograms of H and S components, by choosing the appropriate bins of 36-72, colour features are extracted, and colour feature vectors are created
4. The proposed texture feature is applied to the value component of the HSV model and constructed as a feature vector with the help of a histogram
5. The database contains a feature vector for each image, obtained by merging the colour and texture features of every image into a unified feature vector.
6. The PCA is applied to the feature vector database to reduce the size and create new dimensions for the feature vector database
7. A similarity metric is employed on the database to identify and measure the similarity between images

Results and Discussion

The significance of the proposed algorithm has been verified on two challenging datasets, Corel-10k and 17 flower datasets. The next sections provide a comprehensive explanation of the specifics of each database and its findings.

Specifically, REP patterns capture texture information by encoding the relationships between centre pixels and their neighbours, offering a more comprehensive representation compared to methods like NCDB-LBP, which rely on predefined patterns. Similarly, the use of the HSV colour model allows for the extraction of dominant and perceptually meaningful colour features. PCA further improves performance by reducing dimensionality, retaining only the most discriminative features, which mitigates overfitting and enhances computational efficiency during retrieval. This combination ensures better separation of clusters in the feature space, as validated by comparative analysis of feature distributions and retrieval performance. We have also included additional comparisons with NCDB-LBP and MVM-LBP to demonstrate the superior precision and recall of our method, clarifying the mechanisms underlying its improved performance and strengthening the scientific contribution of our work.

Performance Analysis On Corel-10k

Corel-10k has 10,000 images in 100 categories, and each category has 100 similar images. Figure (4) shows the proposed algorithm performance for an arbitrary query image, as well as its performance with respect to the top pertinent retrieved images from n = 10 to 100, as shown in Table (1). The comparative analysis of the proposed framework with other state-of-the-art techniques is shown in Fig. (5), from the Figs. (4-5). The proposed framework has shown a significant improvement in the retrieval of similar images.

Performance Analysis on 102 Flowers Dataset

The Oxford Flowers 102 dataset (102 Flowers Dataset, n.d.) comprises 102 different flower categories that are commonly found in the United Kingdom as shown in Fig. (6). Each category consists of 80-256 similar images, and a total of 9549 images of flowers are available in the dataset.

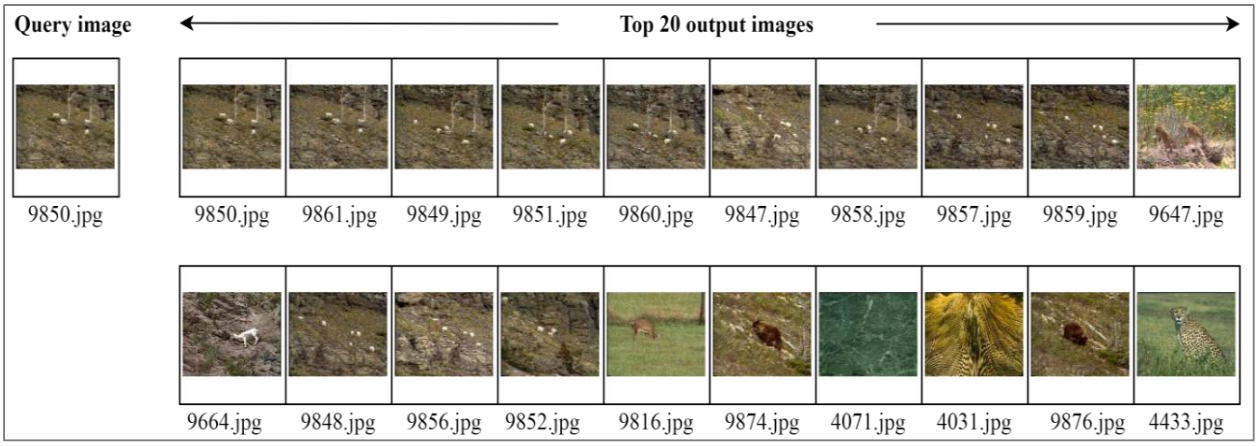


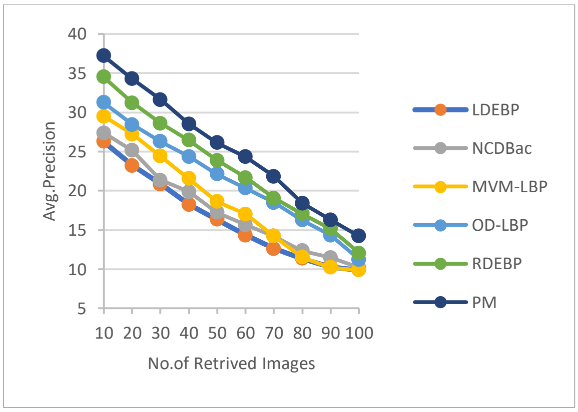
Fig. 4: Retrieved results for n = 10 for a random query image from corel-10k

Table 1: Average Precision and recall values on Corel-10K

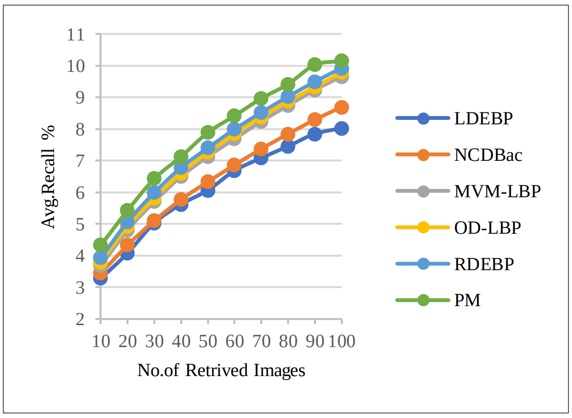
Corel 10k dataset						
Precision						
n	LDEBP	NCDBac	MVM-LBP	OD-LBP	RDEBP	PM
10	26.26	27.39	29.51	31.25	34.51	37.21
20	23.21	25.15	27.23	28.46	31.23	34.29
30	20.84	21.36	24.39	26.26	28.62	31.62
40	18.25	19.89	21.56	24.36	26.46	28.51
50	16.34	17.23	18.63	22.13	23.87	26.12
60	14.32	15.65	17.02	20.31	21.69	24.32
70	12.62	14.21	14.23	18.51	19.06	21.84
80	11.34	12.32	11.54	16.23	17.06	18.4
90	10.28	11.48	10.21	14.32	15.21	16.29
100	9.88	10.21	9.88	11.23	12.05	14.22
Recall						
n	LDEBP	NCDBac	MVM-LBP	OD-LBP	RDEBP	PM
10	3.28	3.44	3.69	3.8	3.93	4.33
20	4.08	4.33	4.81	4.92	5.07	5.43
30	5.03	5.1	5.71	5.82	5.99	6.43
40	5.62	5.77	6.50	6.61	6.78	7.13
50	6.06	6.33	7.13	7.26	7.40	7.89
60	6.68	6.86	7.69	7.85	7.99	8.42
70	7.08	7.37	8.24	8.39	8.52	8.97
80	7.45	7.84	8.74	8.88	9.02	9.42
90	7.84	8.29	9.23	9.33	9.49	10.03
100	8.02	8.69	9.66	9.79	9.92	10.15

It includes images that exhibit various factors such as large-scale variations, diverse poses, and variations in lighting conditions. Additionally, some categories within the dataset showcase significant variations within the category itself, while others are very similar to each other.

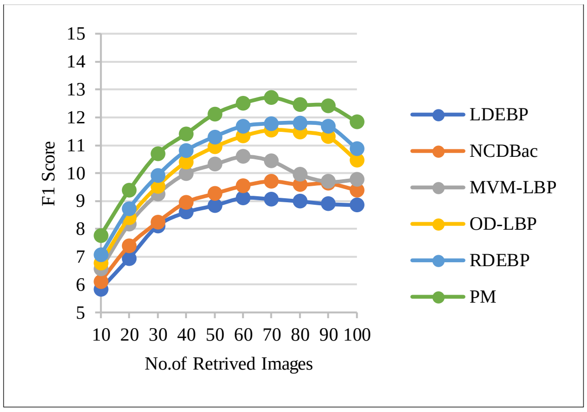
The retrieval performance for the top 20 images of the proposed framework for a typical query image is shown in Fig. (7) and in Table (2) the retrieval performance in terms of precision and recall for the top 10-80 images along with proposed method many other existing methods has shown. Figure (8) completely represents the significance of the proposed framework in terms of the major parameters of the image retrieval systems.



(a)



(b)



(c)

Fig. 5: Comparative analysis of the proposed framework on Corel-10k with respect to (a) Avg. precision vs No. Of received images (b) Avg. Recall vs No. Of retrieved images (c) F1 score

Fig. 6: Sample images from 102 Flowers Dataset

Computational Complexity

The computational complexity of the proposed method is implemented on an Intel Core i7 with 16GB RAM, 4GB NVIDIA GPU and MATLAB 2021a and is analyzed in terms of feature extraction time and the length of the feature vector.

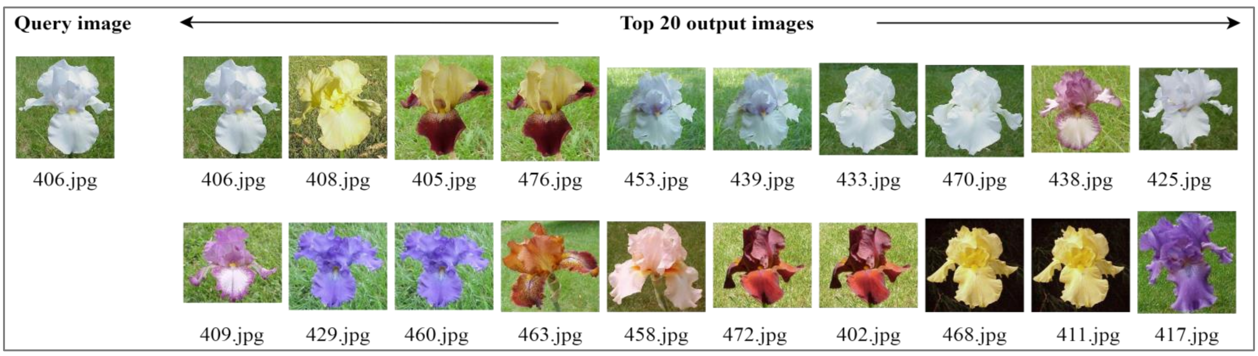


Fig. 7: Retrieved results for n = 10 for a random query image from 17 flower dataset

Table 2: Average Precision and Recall for top n matches from 102-Flower dataset

Flowers Dataset						
Precision---						
n	LDEBP	NCDBac	MVM-LBP	OD-LBP	RDEBP	PM
10	22.65	25.29	26.18	27.37	29.42	30.88
20	16.65	18.89	20.46	20.93	24.16	25.21
30	14.45	16.26	17.42	17.99	20.36	21.11
40	13.02	14.78	15.84	16.37	18.47	18.97
50	12.13	13.79	14.23	15.41	16.98	17.56
60	11.54	13.1	13.84	14.60	15.19	15.98
70	11.01	12.53	12.94	13.89	14.21	15.12
80	10.62	12.06	12.31	13.35	12.38	14.46
Recall---						
n	LDEBP	NCDBac	MVM-LBP	OD-LBP	RDEBP	PM
10	2.83	3.16	3.25	3.42	3.78	3.86
20	4.16	4.72	4.98	5.23	5.85	5.84
30	5.42	6.1	6.51	6.74	7.05	7.55
40	6.51	7.39	7.86	8.18	8.26	9.09
50	7.58	8.62	8.98	9.63	10.16	10.59
60	8.65	9.82	10.01	10.95	11.35	11.99
70	9.63	10.96	11.08	12.16	12.85	13.23
80	10.62	12.06	12.14	13.35	13.98	14.46

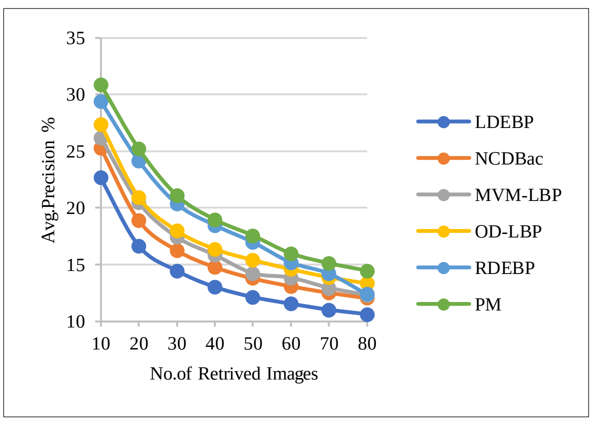
Table 3: Computational complexity comparison between PM to existing methods in terms of feature extraction and retrieval time(s)

Algorithm	Size of the feature vector	Corel-10K		102-Flowers Dataset	
		Feature ext. time(s)	Retrieval time(s)	Feature ext. time(s)	Retrieval time(s)
LDEBP	512	16.97	7.37	21.45	4.89
NCDBac	1024	34.54	10.36	35.25	8.23
MVM-LBP	768	21.45	8.24	27.72	6.58
OD-LBP	768	20.74	7.85	26.65	7.13
RDEBP	512	15.89	8.65	18.24	8.12
PM	768	22.46	7.49	17.49	5.02

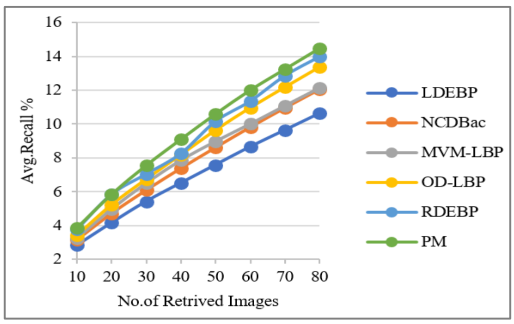
For feature extraction and retrieval, Table (3) compares the proposed method with existing methods on two datasets. The feature extraction time.

Comparison between PCA and t-SNE

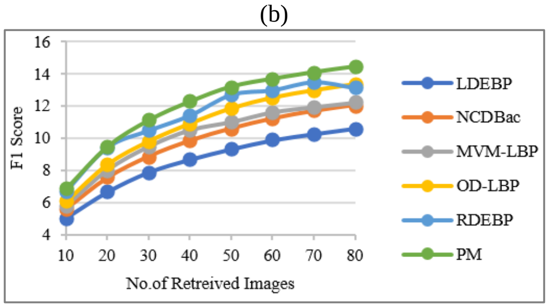
PCA has shown its significance over t-SNE (t-Distributed Stochastic Neighbour Embedding) in terms of scalability, computational efficiency, and ability to handle high-dimensional datasets like Corel-10k and 102-flowers.



(a)



(b)



(c)

Fig. 8: Comparative analysis of the proposed framework on 102 flowers dataset with respect to the (a) Avg. Precision vs No. Of retrieved images (b) Ang. Recall vs No. Of retrieved images (c) F1 score

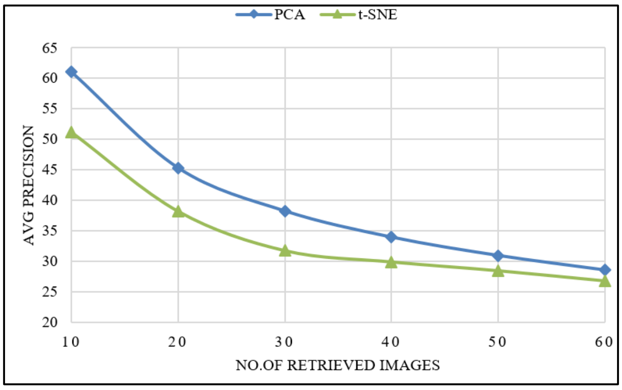


Fig. 9: Comparative analysis of PCA over t-SNE over precision on the 102-flowers dataset

The retrieval performance of these two methods over average precision is given in Fig. (9). PCA's linear transformations make its results interpretable, enabling easier analysis of feature representations. It effectively reduces dimensionality while decorating features, creating a compact and robust feature space suitable for retrieval tasks.

Conclusion

This research paper focuses on enhancing colour image retrieval efficiency through the utilization of colour and texture features, along with Principal Component Analysis (PCA). A novel descriptor called REP is introduced, which captures directional, unidirectional and relative directional relationships among pixels within a 5x5 matrix up to a scale of 2. This results in the generation of six binary patterns. The colour information is extracted from the images using the HSV colour model, and both the colour and texture features are concatenated to form the feature vector. PCA is applied to the feature vector database to reduce the dimensionality of the feature vector. Experimental evaluations conducted on the 102-Flowers and Corel-10k databases demonstrate significant improvements in terms of precision and recall. These findings highlight the effectiveness of the proposed method in enhancing colour image retrieval performance.

Future work will focus on extending the proposed method to handle large-scale datasets and cross-modal retrieval tasks, leveraging datasets like Fashion_DB, COCO and Flickr30k that integrate text and image data to achieve a more comprehensive retrieval framework. Additionally, the scalability and robustness of the method will be explored under varying environmental conditions, such as illumination changes and occlusions. Further optimization of the feature extraction and dimensionality reduction processes will be investigated to improve computational efficiency for real-time applications. Finally, the use of deep learning models to complement or replace handcrafted features like REP will be explored to adapt the methodology to evolving trends in computer vision and retrieval tasks.

Acknowledgment

All authors would like to thank the Dean, School of computing, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

M. Geetha Yadav: Defining the research concept, analysis, implementation, and interpretation of results.

SP. Chokkalingam: Collected related literature, analysis, and interpretation of results.

Ethics

This article is original and unpublished. Corresponding authors confirm that all other authors have read and agree that the manuscript does not involve ethical issues.

References

102 Flowers Dataset. (n.d.). https://www.robots.ox.ac.uk/~vgg/data/flowers/102

Avula, M., Lakkakula, N. P., & Raja, M. P. (2014). Bone Cancer Detection from MRI Scan Imagery Using Mean Pixel Intensity. *2014 8th Asia Modelling Symposium*. 2014 8th Asia Modelling Symposium (AMS), Taipei, Taiwan. https://doi.org/10.1109/ams.2014.36

de Rivaz, P., & Kingsbury, N. (2003). Complex Wavelet Features for Fast Texture Image Retrieval. *Proceedings 1999 International Conference on Image Processing (Cat. 99CH36348)*. 6th International Conference on Image Processing (ICIP'99), Kobe, Japan. https://doi.org/10.1109/icip.1999.821576

Dubey, S. R., Singh, S. K., & Singh, R. K. (2015). Local Diagonal Extrema Pattern: A New and Efficient Feature Descriptor for CT Image Retrieval. *IEEE Signal Processing Letters*, *22*(9), 1215–1219. https://doi.org/10.1109/lsp.2015.2392623

Dubey, S. R., Singh, S. K., & Singh, R. K. (2016). Local Bit-Plane Decoded Pattern: A Novel Feature Descriptor for Biomedical Image Retrieval. *IEEE Journal of Biomedical and Health Informatics*, *20*(4), 1139–1147. https://doi.org/10.1109/jbhi.2015.2437396

Gonde, A. B., Maheshwari, R. P., & Balasubramanian, R. (2013). Modified Curvelet Transform with Vocabulary Tree for Content Based Image Retrieval. *Digital Signal Processing*, *23*(1), 142–150. https://doi.org/10.1016/j.dsp.2012.04.019

Huang, J., Kumar, S. R., & Mitra, M. (1997). Combining Supervised Learning with Color Correlograms for Content-Based Image Retrieval. *Proceedings of the Fifth ACM International Conference on Multimedia - MULTIMEDIA '97*. the fifth ACM international conference, Seattle, Washington, United States. https://doi.org/10.1145/266180.266383

He, Z., You, X., & Yuan, Y. (2009). Texture Image Retrieval Based on Non-Tensor Product Wavelet Filter Banks. *Signal Processing*, *89*(8), 1501–1510. https://doi.org/10.1016/j.sigpro.2009.01.021

Idrissa, M., & Acheroy, M. (2002). Texture Classification Using Gabor Filters. *Pattern Recognition Letters*, *23*(9), 1095–1102. https://doi.org/10.1016/s0167-8655(02)00056-9

Jhanwar, N., Chaudhuri, S., Seetharaman, G., & Zavidovique, B. (2004). Content Based Image Retrieval Using Motif Cooccurrence Matrix. *Image and Vision Computing*, *22*(14), 1211–1220. https://doi.org/10.1016/j.imavis.2004.03.026

Karanwal, S., & Diwakar, M. (2021a). Neighborhood and Center Difference-Based-LBP for Face Recognition. *Pattern Analysis and Applications*, *24*(2), 741–761. https://doi.org/10.1007/s10044-020-00948-8

Karanwal, S., & Diwakar, M. (2021b). OD-LBP: Orthogonal Difference-Local Binary Pattern for Face Recognition. *Digital Signal Processing*, *110*, 102948. https://doi.org/10.1016/j.dsp.2020.102948

Laine, A., & Fan, J. (1993). Texture Classification by Wavelet Packet Signatures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*(11), 1186–1191. https://doi.org/10.1109/34.244679

Liu, Y., Zhang, D., Lu, G., & Ma, W.-Y. (2007). A Survey of Content-Based Image Retrieval with High-Level Semantics. *Pattern Recognition*, *40*(1), 262–282. https://doi.org/10.1016/j.patcog.2006.04.045

Malviya, A. V., & Ladhake, S. A. (2016). Pixel Based Image Forensic Technique for Copy-move Forgery Detection Using Auto Color Correlogram. *Procedia Computer Science*, *79*, 383–390. https://doi.org/10.1016/j.procs.2016.03.050

Murala, S., Maheshwari, R. P., & Balasubramanian, R. (2012). Local Tetra Patterns: A New Feature Descriptor for Content-Based Image Retrieval. *IEEE Transactions on Image Processing*, *21*(5), 2874–2886. https://doi.org/10.1109/tip.2012.2188809

Nitin, A., Sucharitha, G., & Sharma, S. C. (2023). MVM-LBP: Mean−Variance−Median Based LBP for Face Recognition. *International Journal of Information Technology*, *15*(3), 1231–1242. https://doi.org/10.1007/s41870-023-01204-y

Ojala, T., Pietikäinen, M., & Harwood, D. (1996). A Comparative Study of Texture Measures with Classification Based on Featured Distributions. *Pattern Recognition*, *29*(1), 51–59. https://doi.org/10.1016/0031-3203(95)00067-4

Pourghassem, H., & Ghassemian, H. (2008). Content-Based Medical Image Classification Using a New Hierarchical Merging Scheme. *Computerized Medical Imaging and Graphics*, *32*(8), 651–661. https://doi.org/10.1016/j.compmedimag.2008.07.006

Pass, G., & Zabih, R. (1999). Comparing Images Using Joint Histograms. *Multimedia Systems*, *7*(3), 234–240. https://doi.org/10.1007/s005300050125

Salmi, M., & Boucheham, B. (2014). Content Based Image Retrieval Based on Cell Color Coherence Vector (Cell-CCV). *2014 4th International Symposium ISKO-Maghreb: Concepts and Tools for Knowledge Management (ISKO-Maghreb)*. 2014 4th International Symposium ISKO-Maghreb: Concepts and Tools for knowledge Management (ISKO-Maghreb), Algiers, Algeria. https://doi.org/10.1109/isko-maghreb.2014.7033451

Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-Based Image Retrieval at the End of the Early Years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(12), 1349–1380. https://doi.org/10.1109/34.895972

Srivastava, D., Wadhvani, R., & Gyanchandani, M. (2015). A Review: Color Feature Extraction Methods for Content Based Image Retrieval. *International Journal of Computational Engineering & Management*, *18*(3), 9–13.

Sucharitha, G., & Senapati, R. K. (2020). Biomedical Image Retrieval by Using Local Directional Edge Binary Patterns and Zernike Moments. *Multimedia Tools and Applications*, *79*(3–4), 1847–1864. https://doi.org/10.1007/s11042-019-08215-7

Sucharitha, G., Arora, N., & Sharma, S. C. (2023a). Medical Image Retrieval Using a Novel Local Relative Directional Edge Pattern and Zernike Moments. *Multimedia Tools and Applications*, *82*(20), 31737–31757. https://doi.org/10.1007/s11042-023-14720-7

Sucharitha, G., Godavarthi, D., Ramesh, J. V. N., & Khan, M. I. (2024). Secure and Efficient Content-Based Image Retrieval Using Dominant Local Patterns and Watermark Encryption in Cloud Computing. *Cluster Computing*, *27*(9), 11873–11889. https://doi.org/10.1007/s10586-024-04635-9

Sucharitha, G., Kalyani, B. J. D., Chandra Sekhar, G., & Srividya, Ch. (2023b). Efficient Image Retrieval Technique with Local Edge Binary Pattern Using Combined Color and Texture Features. *Computational Intelligence for Engineering and Management Applications*, 261–276. https://doi.org/10.1007/978-981-19-8493-8_21

Varish, N., & Pal, A. K. (2015). Content Based Image Retrieval Using Statistical Features of Color Histogram. *2015 3rd* International Conference on Signal Processing, Communication and Networking (ICSCN), Chennai, India. https://doi.org/10.1109/icscn.2015.7219922