CONTENT-BASED IMAGE RETRIEVAL: UNLEASHING THE POTENTIAL OF A NEW FEATURE DESCRIPTOR

Ramza Fida[1*], Syed Muhammad Adnan[2], Muhammad Zain Uddin Umar[3], Wakeel Ahmad[4]

1. MS Computer Science (Scholar), Department of Computer Science, University of Engineering and Technology (UET), Taxila, Punjab, Pakistan.
2. Assistant Professor, Department of Computer Science, University of Engineering and Technology (UET), Taxila, Punjab, Pakistan.
3. MS Cyber Security, Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad, Capital Terrotory, Pakistan.
4. Lecturer, Department of Computer Science, University of Engineering and Technology (UET), Taxila, Punjab, Pakistan.

* Corresponding Author: Ramza Fida

ABSTRACT

As digital images continue to proliferate, content-based image retrieval represents an increasingly formidable task. There is a huge production of digital images in every domain of life. So, every field gains benefits from CBIR. Extensive research has been done to improve CBIR as many minute details of an image such as texture are not properly retrieved. In many images, all the features are not extracted because an image's directional and multiscale information and color features are missing. In this research, we proposed a transform domain system to effectively retrieve an image's texture and basic content. In this method, we first used the wavelet transform to obtain directional and multiscale information about an image and then fused it with tri-directional pattern to gain confusion matrix of 15 local patterns. Lastly, we combined the patterns to retrieve similar images. We used Corel 5k dataset on natural images. Our research turns out to give 62% precision rate of retrieved images. The research outcomes confirm the method's efficiency in the retrieval of images.

Keywords: Image retrieval, CBIR, digital images, Corel 5k dataset, multiscale information.

I. INTRODUCTION

With the rapid growth of technology, many new techniques are introduced for image acquisition, and information is stored in digital images. All the information obtained from different strands of living is captured in digital formats. In medical imaging, doctors are facilitated with the use of major databases like CT scans, MRIs, and X-Ray images. Similarly, in many different departments, information is stored in images, e.g., in the crime investigation department, pictures are used to maintain the evidence record. Over the past decade, there has been significant interest and development in content-based image retrieval. Many researchers have done much work on image retrieval in depth and breadth. Kato [1] introduced a pioneering approach that utilizes image color and shape to achieve automated image retrieval. The concept of CBIR originated from this early work.

Before the emergence of CBIR, image retrieval predominantly relied on text-based methods, with captions playing a crucial role in the process. The reason to produce irrelevant results was due to the disparity between human visual perception and manual interpretation. Due to these limitations, image interpretation and human perception were not fruitful.CBIR was introduced as a more robust solution to fill the gaps in text-based retrieval of images. CBIR involves a two stage process i.e. feature extraction and query matching. The goal of CBIR is to locate images in the database that have the highest visual resemblance to a provided query image. The beginning of the CBIR process involves a query image as input. Afterward, the query of the image's visual features is compared with those of database images, and images with similar content are determined based on the similarity of their feature vector. The main objective of CBIR aims to find and arrange images that bear a strong resemblance to the user's query in terms of visual content.

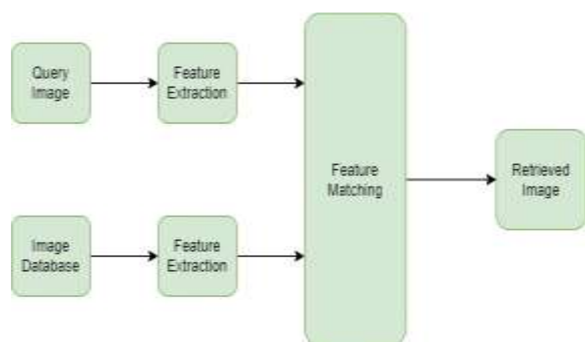

Figure 1: Basic CBIR model

II. RELATED WORKS

CBIR consists of different stages, some of which are mandatory, and some are optional. At the onset of the image retrieval process, users are required to provide quimagesmage. The processes applied to the query image will be applied in the same order to all other images[2]. In earlier times, Haralic proposed GLCM as a method for image classification [3]. Co-occurrence pairs of pixels are used by GLCM to extract features. For the extraction of texture features, Zhang et al. [4]used Pthe rewitt operator for detecting and extracting edge features from images.

Ojala et al. [5], pioneered an approach that focuses on studying the association between pixels within the immediate neighborhood of an image. In the realm of feature descriptors, some have been suggested in the past to catto er various applications including LBP, LDP [6], LMEBP[7], , and DLEP [9], DBWP[10], LMeP[11], LTCoP [12].

Features of an image can be extracted by either analyzing it in the spatial domain or converting it to the frequency domain. Numerous CBIR systems have been evaluated in [13], considering their limitations and potential future directions. In [14], three processes for retrieval i.e. for surrounding images, minimum edge,s and integrated features. Wang et al. [15]proposed the DTCWTILTrP (Dual-Tree Complex Wavelet Transform Improved Local Tetra Pattern) method, a fusion of features extracted from both the spatial and transform domains is employed. In [16], wavelet transform is applied tobtainin multiresolution information, whilethe HOG technique is applied for feature extraction. Ahmad et al. [17] proposed a query-based image retrieval, which includes a user-friendly GUI for improved retrieval functionality. In [an 18], image decomposition process is proposed which involves using the Discrete Non-Separable Shearlet Transform (DNST), and the subbands obtained from it are used to compute the Local Non-Isotropic Pattern (LNIP). Additionally, statistical features are utilized to represent the images. To measure intensity changes, Weber's law is modified, and local multiple patterns are computed in [19]. In DLTCOP, four directional filters are used to compute the major directional responses[20]. Different radii are incorporated to further extend the pattern (DLTCoP) and enhance its capabilities [21]. Histograms are utilized to incorporate color information into both features, ensuring a comprehensive representation. Lungs X-Ray images are retrieved by leveraging transfer learning on a deep CNN work in [22]. For the efficient processing of massive image data, MapReduce is employed in six distinct modes [23]. In [24], authors proposed a system that includes two stages of retrieval, with different features used in each stage. The first stage involves coarse retrieval, while the second stage focuses on fine retrieval. In [25], research is conducted using distributed deep learning (DDL) using the AlexNet architecture. They introduced three modes of DDL: synchronous, asynchronous, and Hogwild. Images are retrieved using transfer learning in [26] using different classifiers such as VGG16, VGG19, Re,stNet and their variants. A novel hybrid CBIR technique is presented in [27] that combines global and local features, employing a two-layer filtering approach. IN [28], authors suggested a method for image retrieval that combines low-level features and deep features through an aggregation approach. Hassan et al. [29], proposed a color feature descriptor namely color octree quantization descriptor (COQD) to extract color strings. For feature extraction, low-dimentional feature descriptor is constructed using texture and color contents in [30]. Authors in [31] proposed an efficient image retrieval application that combines several descriptors from MPEG-7. Each image in the database is represented by six descrptors, and similarity between images is computed using specific distance measures to find the nearest image.

With the origination of new methodologies, images are acquired in a 3D plane. Many new feature descriptors are proposed to retrieve the image in a 3D plane. First of all, Volumetric LBP (VLBP) was proposed for dynamic texture recognition. Murala et al.[32] in 2015 used spherical directions to introduce new method of image retrieval of 3D volume. Gonde et al.[33], improved upon the existing methods of using spherical directions for feature extraction by incorporating wavelet transform for each of R G B channels in colored images. The aim of this research is to fuse the concept of a 3D local transform pattern with tri-directional pattern. We used the concept of the tri-directional pattern [34] to improvthe e confusion matrix and to generate 15 local patterns and combine them to retrieve the given image.

III. PROPOSED METHOD

After 3D-LTrP, we tried to solve the hurdles left behind in the previous research. They used local neighborhoods to retrieve the image whereas the retrieval system can be made more accurate by involving adjacent neighborhoods along with the center pixel in every symmetric spherical direction. In the proposed methodology, the problem is overcome by implementing the tri-directional pattern in every symmetric direction of the 3D plane. Moreover, the use of wavelet transform will maintain the multiscale resolution of the images. Figure 2 explains the directions of the patterns formed by the transformed images in R G B planes respectively. A novel descriptor is developed by utilizing a set of five symmetric directions.

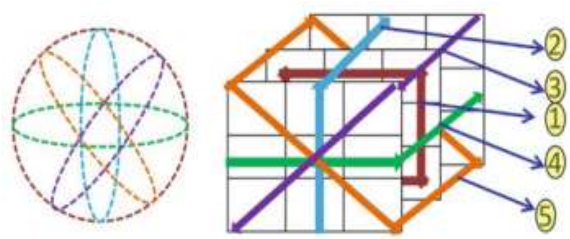

Figure 2: Spherical symmetric directions

By considering the specific value of every direction, a new descriptor value can be calculated by considering its relationship with the center pixel. Considering equation 1, the new relationship is formed by subtracting the center pixel value Ic(Gt) from every pixel as explained above. Subtracting the middle pixel from every pixel, the color pixel values of the image are calculated. To convert them to the binary value we use tri-directional code [27] to convert them to two binary values ($V_{LTriDP1}$, $V_{LTriDP2}$).

$$V_{P=8} = \begin{cases} v(I_0(G_t),(I_1(G_t),(I_2(G_t),(I_3(G_t),(I_4(G_t),(I_5(G_t),(I_6(G_t),(I_7(G_t) & \alpha = 1 \\ v(I_2(R_t), I_c(R_t), I_6(R_t), I_6(G_t), I_6(B_t), I_c(B_t), I_2(B_t), I_2(G_t) & \alpha = 2 \\ v(I_5(R_t), I_c(R_t), I_1(R_t), I_1(G_t), I_1(B_t), I_c(B_t), I_5(B_t), I_5(G_t) & \alpha = 3 \\ v(I_4(R_t), I_c(R_t), I_0(R_t), I_0(G_t), I_0(B_t), I_c(B_t), I_4(B_t), I_4(G_t) & \alpha = 4 \\ v(I_3(R_t), I_c(R_t), I_7(R_t), I_7(G_t), I_7(B_t), I_c(B_t), I_3(B_t), I_3(G_t) & \alpha = 5 \end{cases} \quad 1$$

The proposed feature descriptor is described in Equation 2.

$$equalDescriptor_{p,q,\alpha} = \sum_{q=0}^{p=1} 2^q * V_{(LTriDP_1, LTriDP_2, LTriDP_{mag})q} \quad 2$$

Once the local patterns have been identified, the construction of a histogram is employed to represent the entire image.

Histogram for this descriptor is described using Equation 3 and Equation 4.

$$H_s(q) = \frac{1}{N_1 \times N_2} \sum_{j=1}^{N_1} \sum_{k=1}^{N_2} f_4(Descriptor(j,k), q); \quad q \in [0, P\{P-1) + 2] \quad 2$$

$$f_4(x,y) = \begin{cases} 1 & x = y \\ 0 & x \neq y \end{cases} \quad 3$$

The size of the image, which is N_1XN_2, is considered.

IV. PROPOSED ALGORITHM

A. Feature Extraction

Figure 3 explains the working of the proposed method, while the steps of the proposed algorithm are outlined below.

ALGORITHM

$H_S(q)$	Histogram
α	Spherical symmetric directions
K	Decomposition levels in wavelet transform
Input	*color image*
Output	*Feature vector*

1. Load color image as input and separate R G B channels.
2. At K levels of decomposition, a wavelet transform is implemented on each of the channel separately.
3. For k= 1 to N
   a. Generate 3D volume to each R G B transform images.
   b. Compute the pattern in five symmetric directions.
   c. Convert each code into three binary codes ($LTriDP_1$, $LTriDP_2$, LTriDPmag)
   d. Construct a histogram for each binary pattern by using equation 3.
   e. Concatenate histogram
4. End of K
5. Concatenate all histograms at step 3 to obtain the feature vector.

B. Similarity Measurement

The process to select images from a given set of images that bear a resemblance to the query image is achieved using a similarity measurement method. By the algorithm outlined in section 4.1, features are gathered from the given database. The feature vector of the query image is compared with the feature vector of the image in the test database using the given equation 5 for evaluation

$$D(Q,I) = \sum_{i=1}^{N} \left| \frac{f_{I,i} - f_{Q,i}}{1 + f_{I,i} + f_{Q,i}} \right| \quad 4$$

where

Q Query image.
N length of feature vector;
I database image.
$F_{I,i}$ ith feature of I_{th} image in the database;
$F_{Q,i}$ ith feature of query image Q

A. Evaluation Measures

To evaluate the performance, we used two metrics .

- Average retrieval precision (ARP)
- Average retrieval rate (ARR)

ARP and ARR are defined as

$$ARP = \frac{1}{N_q} \sum_{k=1}^{N_q} P(Q_k) \quad 5$$

$$ARR = \frac{1}{N_q} \sum_{k=1}^{N_q} R(Q_k) \quad 6$$

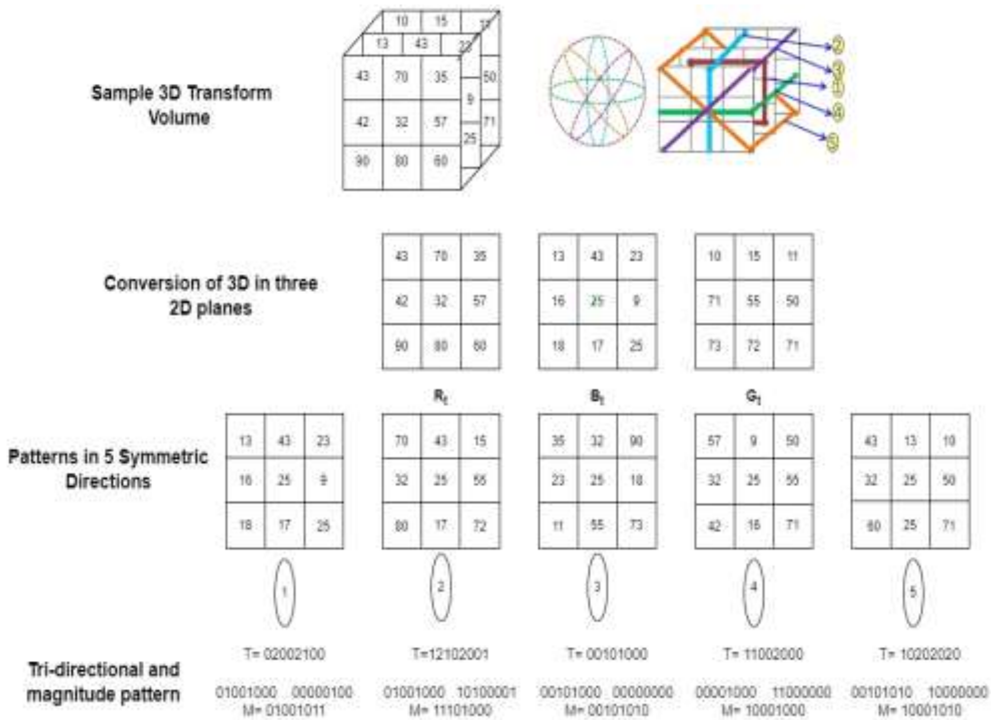

Figure 3: Proposed methodology

V. RESULTS

To assess the effectiveness of our approach, we conducted experiments using the COREL dataset. We evaluated the performance of our descriptor using precision and recall metrics. All the experiments were applied on an image base to retrieve data.

A. Dataset

COREL 5k image database comprises 5000 images categorized into 10 different classes, each of which contains 100 images from various groups. Fig 4 provides examples of some of the images in the COREL dataset. Researchers in this field often utilize these classes as standard benchmarks for evaluating and comparing various methods and techniques.

Figure 4. Examples from COREL datasets

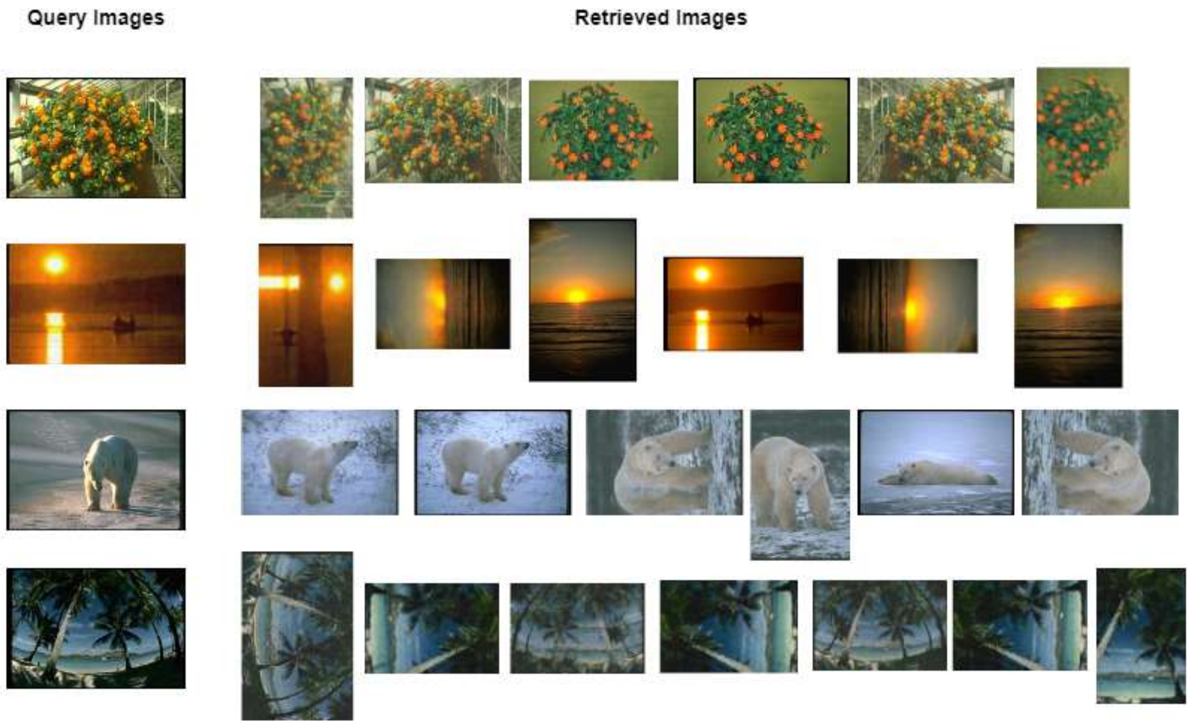

Figure 5 Retrieved images from COREL 5k Dataset

B. Results

To assess the performance of our approach, we utilized similarity measures to identify the most relevant images. Precision and recall values were calculated by considering the top twenty retrieved images, as well as randomly selected query images. We ran our experiments five times for each image class and calculated the average results. The precision and recall values of different CBIR descriptors were compared and presented in Table 1. The comparison demonstrated the high efficiency of our new descriptor for image retrieval. Precision and recall of 60 different set of images using different descriptors is presented in Figure 6 and Figure 7 respectively.

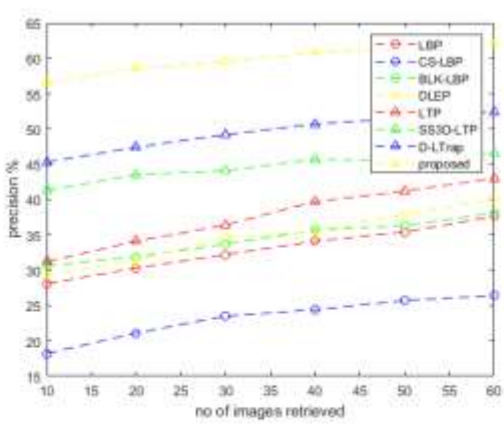

Figure 6 Precision for the number of retrieved images

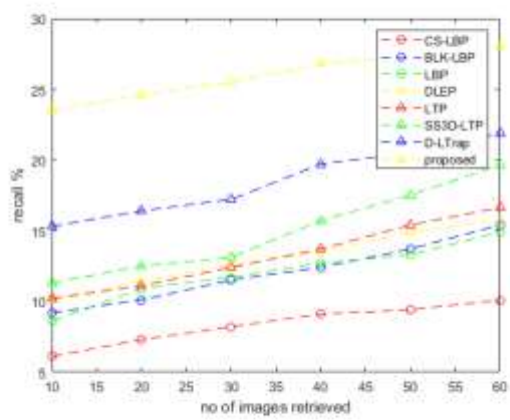

Figure 7 Recall for number of retrieved images

Table 1: Comparative analysis of the different techniques

Method	Performance	
	Precision (%)	**Recall (%)**
CS_LBP	26.40	10.10
BLK_LBP	38.10	15.30
LBP	37.60	14.90
DLEP	40.00	15.70
LTP	42.95	16.62
SS-3D-LTP	46.25	19.64
3d-LTraP	52.46	21.90
Proposed work	**62.00**	**28.00**

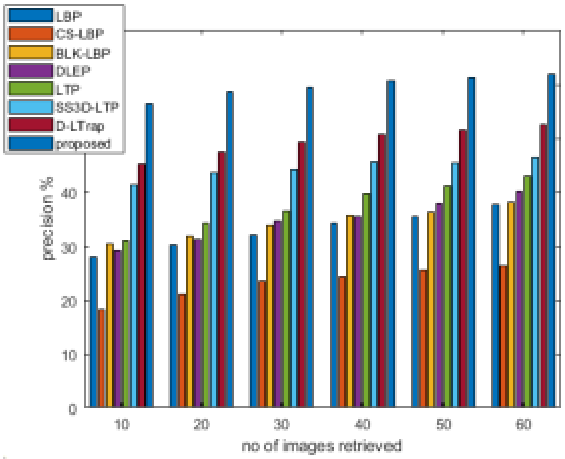

Figure 8 Precision rate of retrieved images using a bar graph

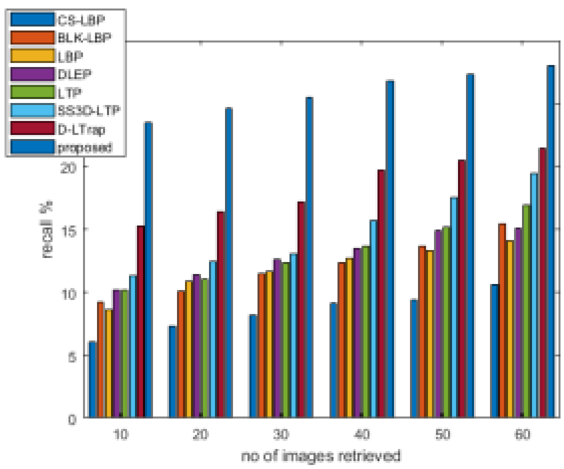

Figure 9 Recall rate of different images using a bar graph

VI. CONCLUSION

The primary aim of our research is to enhance the effectiveness of content-based image retrieval, which is widely recognized and challenging in the field of computer vision. For this, we introduced a novel descriptor that combines VLBP and LTriDP descriptors. The novel descriptor is employed to extract texture features from the entire image in the transform domain, considering each color plane separately. By incorporating information from the joined neighborhood, the LTriDP component of the descriptor improves the accuracy of retrieved images. The results of our experiments show that this research surpasses other methods in both precision and recall. The retrieval precision of images based on shape and color is demonstrated to be 62% by the proposed system.

REFERENCES

1. Kato, T. *Database architecture for content-based image retrieval*. in *image storage and retrieval systems*. 1992. SPIE.
2. Kherfi, M.L., D. Ziou, and A. Bernardi, *Image retrieval from the world wide web: Issues, techniques, and systems.* ACM Computing Surveys (Csur), 2004. 36(1): p. 35-67.
3. Haralick, R.M., *Statistical and structural approaches to texture.* Proceedings of the IEEE, 1979. 67(5): p. 786-804.
4. Zhang, J., G.-l. Li, and S.-w. He. *Texture-based image retrieval by edge detection matching GLCM*. in *2008 10th IEEE International Conference on High-Performance Computing and Communications*. 2008. IEEE.
5. Ojala, T., M. Pietikäinen, and D. Harwood, *A comparative study of texture measures with classification based on featured distributions.* Pattern recognition, 1996. 29(1): p. 51-59.
6. Zhang, B., et al., *Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor.* IEEE Transactions on image processing, 2009. 19(2): p. 533-544.
7. Subrahmanyam, M., R. Maheshwari, and R. Balasubramanian, *Local maximum edge binary patterns: a new descriptor for image retrieval and object tracking.* Signal Processing, 2012. 92(6): p. 1467-1479.
8. Murala, S., R. Maheshwari, and R. Balasubramanian, *Local tetra patterns: a new feature descriptor for content-based image retrieval.* IEEE Transactions on image processing, 2012. 21(5): p. 2874-2886.
9. Murala, S., R. Maheshwari, and R. Balasubramanian, *Directional local extrema patterns: a new descriptor for content based image retrieval.* International journal of multimedia information retrieval, 2012. 1(3): p. 191-203.
10. Murala, S., R. Maheshwari, and R. Balasubramanian, *Directional binary wavelet patterns for biomedical image indexing and retrieval.* Journal of Medical Systems, 2012. 36: p. 2865-2879.
11. Murala, S. and Q.J. Wu, *Local mesh patterns versus local binary patterns: biomedical image indexing and retrieval.* IEEE Journal of Biomedical and health informatics, 2013. 18(3): p. 929-938.
12. Murala, S. and Q.J. Wu, *Local ternary co-occurrence patterns: a new feature descriptor for MRI and CT image retrieval.* Neurocomputing, 2013. 119: p. 399-412.
13. Rout, N.K., M. Atulkar, and M.K. Ahirwal, *A review on content-based image retrieval system: Present trends and future challenges.* International Journal of Computational Vision and Robotics, 2021. 11(5): p. 461-485.

14. Nesakumari, G.R., *Image retrieval system based on multi feature extraction and its performance assessment.* International Journal of Information Technology, 2021: p. 1-13.
15. Wang, H., et al., *Texture image retrieval based on the fusion of local and global features.* Multimedia Tools and Applications, 2022. 81(10): p. 14081-14104.
16. Dixit, U.D. and M. Shirdhonkar, *Language-based document image retrieval for Trilingual System.* International Journal of Information Technology, 2020. 12(4): p. 1217-1226.
17. Ahmad, K., et al., *An efficient image retrieval tool: query-based image management system.* International Journal of Information Technology, 2020. 12: p. 103-111.
18. Wang, X., et al., *Texture image retrieval using DNST domain local neighborhood intensity pattern.* Multimedia Tools and Applications, 2022. 81(20): p. 29525-29554.
19. Yang, W., X. Zhang, and J. Li, *A local multiple patterns feature descriptor for face recognition.* Neurocomputing, 2020. 373: p. 109-122.
20. Singhal, A., M. Agarwal, and R.B. Pachori, *Directional local ternary co-occurrence pattern for natural image retrieval.* Multimedia Tools and Applications, 2021. 80: p. 15901-15920.
21. Agarwal, M. and A. Singhal, *Directional local co-occurrence patterns based on Haar-like filters.* Multimedia Tools and Applications, 2022: p. 1-15.
22. Agrawal, S., et al., *Content-based medical image retrieval system for lung diseases using deep CNNs.* International Journal of Information Technology, 2022: p. 1-9.
23. Kanaparthi, S.K. and U. Raju, *Content based image retrieval on big image data using local and global features.* International Journal of Information Technology, 2022. 14(1): p. 49-68.
24. Shaheen, F. and R. Raibagkar, *Efficient Content-Based Image Retrieval System with Two-Tier Hybrid Frameworks.* Applied Computer Systems, 2022. 27(2): p. 166-182.
25. Raju, U., et al. *Distributed Deep Learning for Content-Based Image Retrieval*. in *Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021*. 2023. Springer.
26. Chughtai, I.T., et al., *Content-based image retrieval via transfer learning.* Journal of Intelligent & Fuzzy Systems, 2023(Preprint): p. 1-26.
27. Salih, F.A.A. and A.A. Abdulla, *Two-layer content-based image retrieval technique for improving effectiveness.* Multimedia Tools and Applications, 2023: p. 1-22.
28. Lu, Z., et al., *Image retrieval using dual-weighted deep feature descriptor.* International Journal of Machine Learning and Cybernetics, 2023: p. 1-11.
29. Zekkouri, H., B. Aksasse, and M. Ouanan. *Content-Based Image Retrieval Using Octree Quantization Algorithm*. in *International Conference on Advanced Intelligent Systems for Sustainable Development*. 2022. Springer.
30. Varish, N., et al., *Color Image Retrieval Method Using Low Dimensional Salient Visual Feature Descriptors for IoT Applications.* Computational Intelligence and Neuroscience, 2023. 2023.
31. Aziz, E., H. Silkan, and A. Boulezhar. *Combining Descriptors for Efficient Retrieval in Databases Images*. in *International Conference on Advanced Intelligent Systems for Sustainable Development*. 2022. Springer.
32. Murala, S. and Q.J. Wu, *Spherical symmetric 3D local ternary patterns for natural, texture and biomedical image indexing and retrieval.* Neurocomputing, 2015. **149**: p. 1502-1514.
33. Gonde, A.B., et al. *3D local transform patterns: a new feature descriptor for image retrieval*. in *Proceedings of International Conference on Computer Vision and Image Processing: CVIP 2016, Volume 2*. 2017. Springer.
34. Verma, M. and B. Raman, *Local tri-directional patterns: A new texture feature descriptor for image retrieval.* Digital Signal Processing, 2016. **51**: p. 62-72.