Content Based Image Retrieval using Multi-level 3D Color Texture and Low Level Color Features with Neural Network Based Classification System

Priyesh Tiwari
Department of Electronics and Communication Engineering
Manipal University Jaipur
Jaipur, India
priyesht81@gmail.com

Kulwant Singh
Department of Electronics and Communication Engineering
Manipal University Jaipur
Jaipur, India
kulwant.singh@jaipur.manipal.edu

Shivendra Nath Sharan
Department of Electronics and Communication Engineering
NIIT University Nimrana
Neemrana, India.
snathsharan@yahoo.com

Suraj Kamya
Sr. Data Scientist
Lagozon Technologies Pvt. Ltd.
Noida,India
kamyasuraj@yahoo.com

Abstract—Content based image retrieval (CBIR), is an application of real-world computer vision domain where from a query image, similar images are searched from the database. The research presented in this paper aims to find out best features and classification model for optimum results for CBIR system. Five different set of feature combinations in two different color domains (i.e., RGB & HSV) are compared and evaluated using Neural Network Classifier, where best results obtained are 88.2% in terms of classifier accuracy. Color moments feature used comprises of: Mean, Standard Deviation, Kurtosis and Skewness. Histogram features is calculated via 10 probability bins. Wang-1k dataset is used to evaluate the CBIR system performance for image retrieval. Research concludes that integrated multi-level 3D color-texture feature yields most accurate results and also performs better in comparison to individually computed color and texture features.

Keywords—CBIR, Color Feature, Texture Feature, Multi-level 3D color texture feature, Neural Network classifier, Wang

I. INTRODUCTION

Technological advances have made enormous data streaming a reality from all aspects of our life. As per the Google search statistics it performs more than 3.5 billion searches per day. As per the Business Today, only WhatsApp users shares 55 billion texts, 4.5 billion photos, 1 billion videos per day. So, if we analyses each technological aspect than images transfer and stores are way bigger than we think and searching right image from these databases can be a herculean task. Because every image is not annotated or labelled as per the content of the image. To find right image from the database big companies like Google and Microsoft have their own search engines for query image retrieval based on content (color, shape and/or texture etc.) [1].

A 3-year child can recognize different concepts and correctly identify various objects based on their physical properties [2]. Semantic gap is the difference between human cognition and mathematically driven feature metrices. To reduce the gap multiple features can be combined to design better algorithms [3].

Shape, color, texture and salient points in an image are most used low-level features in CBIR. Image retrieval majorly targets on similar image search based on multidimension features of an image [4].

As the resultant image and query image usually have semantic gap and it's difficult to eliminate, because multidimensional approach is required to solve such problem. A new approach is proposed to utilize the benefits of low-level color features along with advantage of multidimensional combined color and texture features. For evaluation and comparison purpose CNN, convolutional neural network classifier is used.

FIG. 1. SOME EXAMPLE IMAGES OF BEACHES, BUS AND PEOPLE CATEGORY FROM WANG-1K DATABASE.

Novelty of the present work is in feature extraction where color and texture both features are combined in 3D space and evaluated as a single feature. These combined features are evaluated and compared in both RGB and HSV domain along with other low-level features. For performance evaluation and benchmarking Wang 1-K database is used. Database contains images of 10 categories i.e. beaches, buildings, buses, dinosaurs, elephant, flowers, food, horses, mountains and people. Each category consists of 100 images which in total make 1,000 images. Figure 1. shows some of the examples from the database. All dataset images are of dimension 120x80 [2] [5].

Following this section, novel works proposed by other researchers are explained in section II. Proposed methodology is represented in section III which discuss block diagram, workflow and computational logic in detail. Section IV represents results and comparison for proposed work. Finally,

section V concludes and discuss further area of improvement for presented research work.

II. BACKGROUND WORK

For reducing semantic gap and increasing the accuracy of search results, researchers have proposed several different features and classification algorithms. Usually to extract the low-level features in spatial domain, it is supported that dividing the image into small blocks is a reliable and precise way of feature extraction [6] [7]. Some relevant studies can be found in [21] and [22].

Reference [6] provided a comparative overview of the most famous image analysis and extraction systems. Systems are compared and evaluated based on how much similar results they can produce. A relative analysis of 9 systems i.e. TinEye, YANDEX, AKIWI, FIRE, IOSB, VIRAL, LIRE, PASTEC and MIFILE. 50% of the examined systems requires image query with annotations such as text description or tags selection from users, whereas Yandex and TinEye were tested, only based on image query to consume shortest amount of time. It's also suggested that to save time and improvement in accuracy, relevance feedback is relatively useful.

Reference [8] proposed a hybrid features combining low level and frequency domain features. Proposed system was evaluated using different distance metrics i.e. City block, Minkowski, Euclidean and Mahalanobis. Wang 1-K database is used to evaluate the system and an accuracy of 87.5 was achieved. The results achieved outperforms similar researches [9], [10], [11] and a comparison table is also provided.

Reference [12] proposed a CBIR framework using frequency domain features DWT based on local tera pattern. Corel1-K, Corel 5-Kand Corel 10-K databases are tested to evaluate the accuracy of the system. Artificial Neural Network, ANN classifier is used for query matching and searching results which also solves the shortcomings of minimum distance classifier. It is suggested that local wavelet pattern is more robust and accurate for the task of image retrieval. Combined feature improves the ARR, Average retrieval rate.

Reference [7] combined the features namely color and texture. To extract the content from the image an overview of MPEG-7 edge histogram descriptor is also provided. For classification SVM, Support Vector Machine algorithm is used. Edge histogram is obtained in 5 bins after dividing the database images into 4x4 sub-images. 5 bins used are horizontal, vertical, 45-degree, 135-degree and non-directional. To evaluate the algorithm IMAGEVARY 1-K dataset is used and maximum precision achieved is 75%.

Reference [13] focused on reducing the semantic gap by using RF, relevance feedback and low-level visual features. BDEE, biased discriminative Euclidean embedding technique is proposed to evaluate these low-level visual features. Corel dataset is used to evaluate the algorithm and they achieved a significant increase in accuracy by using semi supervised BDEE. Proposed algorithm is compared against other novel algorithms i.e., biased discriminant analysis (BDA), directed BDA (DBDA), marginal biased analysis (MBA), and support vector machine (SVM) based relevance feedback algorithms.

III. PROPOSED SYSTEM

Working of a CBIR system depends upon algorithm and type of features being extracted [14]. Proposed system provides a new method for feature extraction, which is multilevel 3D color-texture feature. Effectiveness of combined feature is much greater than individual feature algorithms. Different combination of features in RGB and HSV domains are evaluated and compared for optimized results. All algorithms and classifiers are evaluated in MATLAB 2020a.

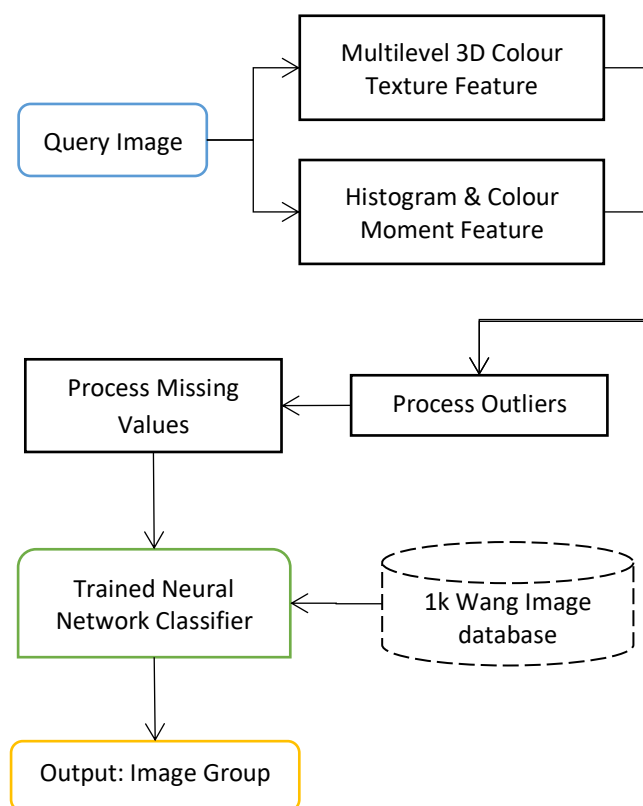A functional block diagram of the system is represented in figure 2.

FIG. 2. FUNCTIONAL BLOCK DIAGRAM

A. Color Domain

In terms of human interpretation HSV color domain is much efficient with respect to RGB [15]. In proposed system features are evaluated and compared in both color domains i.e. RGB and HSV. For color space conversion rgb2hsv standard function is used from MATLAB which uses hex-cone model [16].

B. Feature Extraction

Five different set of feature combinations are compared and evaluated to optimize system performance. Color moments feature comprises of: Mean, Standard Deviation, Kurtosis and Skewness. Histogram features is calculated via 10 probability bins.

1) Color Moments & Histogram in RGB domain
2) Color Moments & Histogram in HSV domain
3) Multi-Level 3D Color Texture & Color Moments in RGB domain

4) Multi-Level 3D Color Texture & Color Moments in HSV domain
5) Multi-Level 3D Color Texture, Color Moments & Histogram in HSV domain

Mean is defined in equation (1), where A denotes variable vector, formed up by N scalar observations

$$\mu = \frac{1}{N} \sum_{i=1}^{N} Ai \qquad (1)$$

Standard deviation is defined in equation (2)

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} |A_i - \mu|^2} \qquad (2)$$

Kurtosis is defined in equation (3), where in terms of x, μ is the mean, σ is the standard deviation and E(t) is expected value of t. Kurtosis value greater than 3 means distribution is more likely to have outliers and vice versa, whereas 0 means normal distribution.

$$K = \frac{E(x-\mu)^4}{\sigma^4} \qquad (3)$$

Skewness is used to evaluate asymmetry in data around the mean and represented by following equation (4). Perfect symmetric distribution results in 0, whereas +value signifies right shift and vice versa.

$$S = \frac{E(x-\mu)^3}{\sigma^3} \qquad (4)$$

To evaluate multilevel 3D color features CLCM (color level co-occurrence matrix) is used which is based upon GLCM (grey level co-occurrence matrix). Harlick [24] defined fourteen statistical features co-occurrence matrices for texture features and classification, but due to strong co-relation only four out of fourteen are used. These are defined as contrast, correlation, energy and homogeneity defined respectively by equation (5), (6), (7) and (8):

$$Contrast = \sum_{i,j} |i-j|^2 p(i,j) \qquad (5)$$

$$Correlation = \sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j} \qquad (6)$$

$$Energy = \sum_{i,j} p(i,j)^2 \qquad (7)$$

$$Homogeneity = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \qquad (8)$$

Above four equations define that how many times a grey pixel is corelated with its adjacent pixels. To embed the color information while evaluating these features adjacent pixels are used from different color planes with parameters distance, d = 1 or 0 and an angle Θ = 0 [17]. In this manner extracted features comprises 12 feature points which are RG0, RB0, BG0, RR, BB, GG, RG, RB, GR, GB, BR and BG, where R, G, B represents the Red, Green and Blue planes respectively. By default (d, Θ) pair are (1,0) otherwise d = 0 is mentioned after the plane. To compute the features complete single image of dimension 120x80 is used which results a 12x4

feature matrix. Rather than using full image a kernel of smaller size such as 3x3, 10x10 or MxN can also be used, where M (rows) and N (columns) should be smaller than actual size of image. But in this case computational complexity will be exponentially increased. Lower kernel size will lead to increased computational cost. Because in single image for each iteration a new 12x4 feature matrix will be generated.

In HSV plane again (d, Θ) pairs are similar and the notation format. 12 feature points for HSV planes are: HS0, HV0, SV0, HH, SS, VV, HS, HV, SH, SV, VH and VS. It also provides a 12x4 feature matrix for further evaluation.
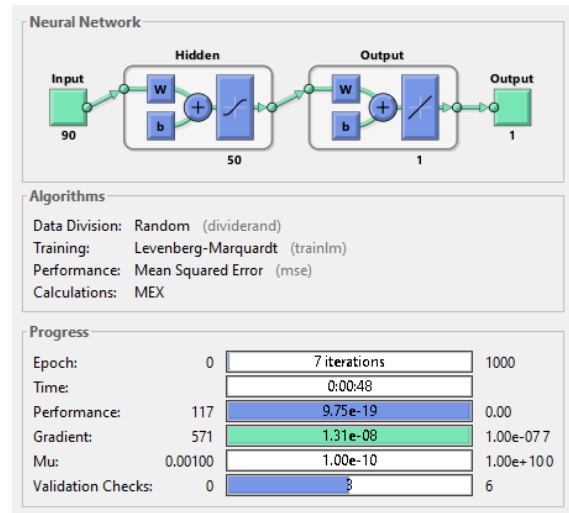
C. Neural Network Classifier

FIG. 3. CLASSIFIER STRUCTURAL DIAGRAM

Figure 3 represents structural diagram with details of neural network classifier used for feature system of Multi-Level 3D Color Texture, Color Moments & Histogram in HSV domain. The neural network, NN used for classification has 50 hidden layers and as an output it provides the group number to which it belongs. At input it shows 90, which basically represents the feature points. From multilevel 3D color texture feature we got a feature matrix of 12x4, which translates to 1x48 in vector format and reaming 1x42 feature vector points are obtained from color moments and histogram features. For training of network Levenberg-Marquardt algorithm is used which requires more memory but less time to train the network [20]. From Wang-1k database it is separated into 3 parts i.e. training, validation and testing which respectively comprises 70%, 15% and 15% randomly divided data respectively. Performance of network is evaluated using mean square error, mse. Figure 4. represents the error histogram from neural network which also indicates outliers present in the data. Wider spread represents more outliers and vice versa. Figure 5. represents flow diagram of proposed methodology.

Out of five feature sets, fifth one outperforms all other four feature-set. Due to this reason network diagram and related performance evaluation figures are given for Multi-Level 3D Color Texture, Color Moments & Histogram in HSV domain system.
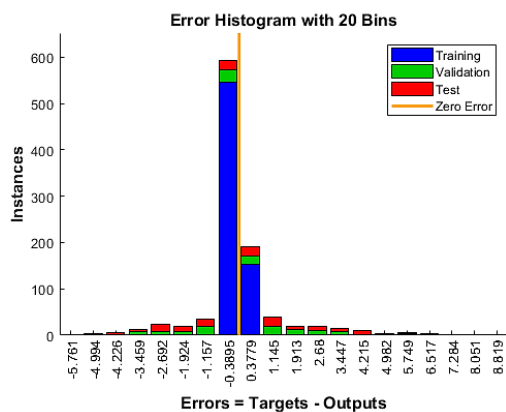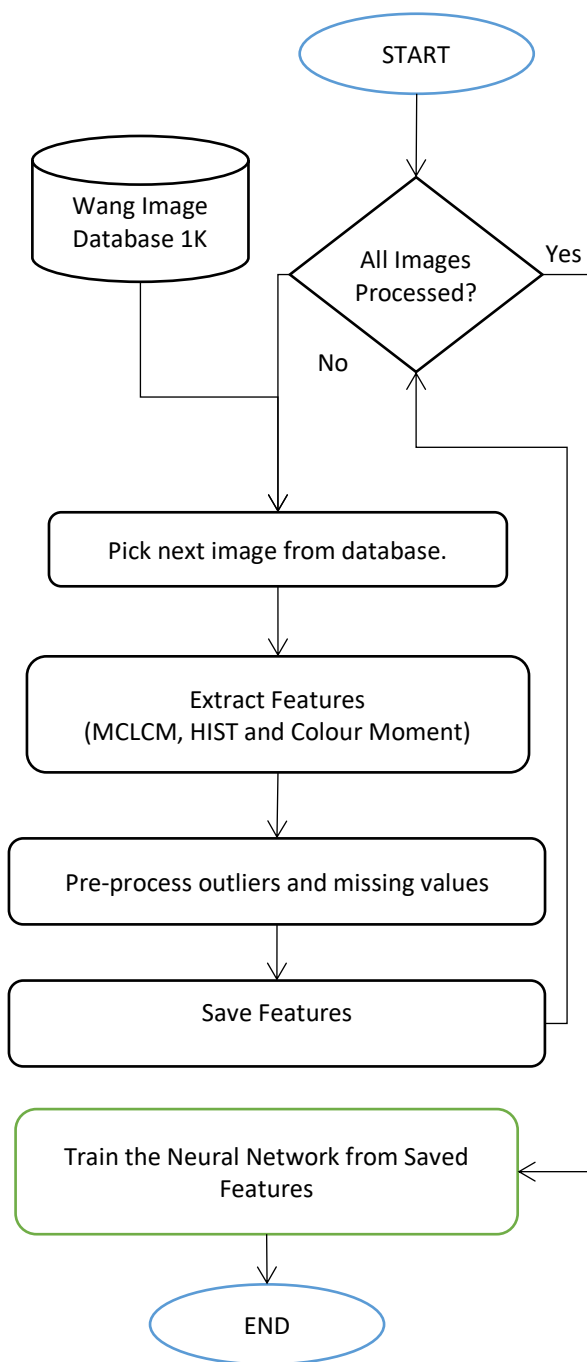
FIG. 4. ERROR HISTOGRAM

IV. RESULT & DISCUSSION

In terms of classification accuracy, best results obtained are 88.2%. To obtain these results, color domain used is HSV and features used are 3D MCLCM, Color Moment and Histogram. On using these features total features points becomes 90 which are fed to Neural Network Classifier as inputs and shown in Figure 3. Regression results and performance of all phases, training, validation and testing are respectively 0.998, 0.622 and 0.620, as presented in Figure 6.
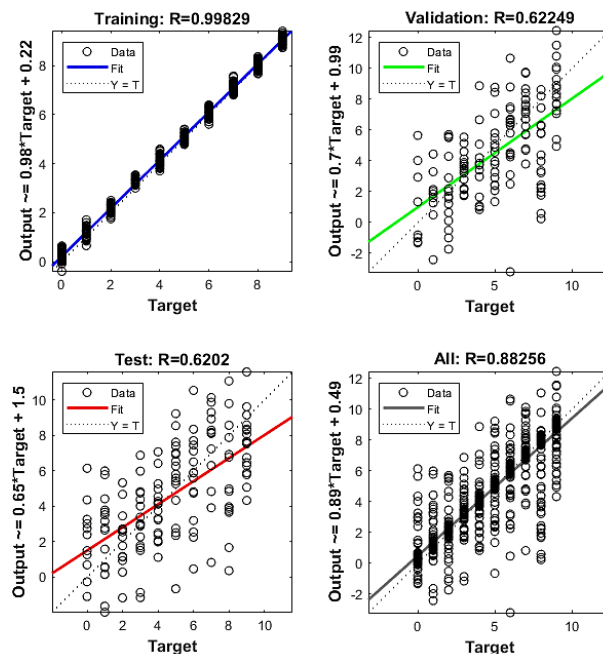

FIG. 6. REGRESSION ANALYSIS

TABLE I COMPARATIVE ANALYSIS OF FIVE FEATURE SETS

Sr	Domain	Features Type	Feature Points	Layers	Regression
1	RGB	Color Moment & Histogram	42	30	0.75
2	HSV	Color Moment & Histogram	42	30	0.79
3	RGB	3D MCLCM and Color Moment	60	20	0.75
4	HSV	3D MCLCM and Color Moment	60	30	0.83
5	HSV	3D MCLCM, Color Moment & Histogram	90	50	0.88

In similar manner, Table-I represents the comparisons of all feature combinations in both (RGB/HSV) color domains and number of feature points which are input to NN Classifier respectively. Neural network classifier parameters are exactly same for all five feature types except number of layers. Number of layers are found empirically for best regression fit in each case. Figure 7 represents feature and layer data of Table-I in bar chart format for easier analysis and comparison of layers in feature points. Figure-8 represents line graph showing accuracy of all five Neural Network Classifier's in term of regression from Table-I.
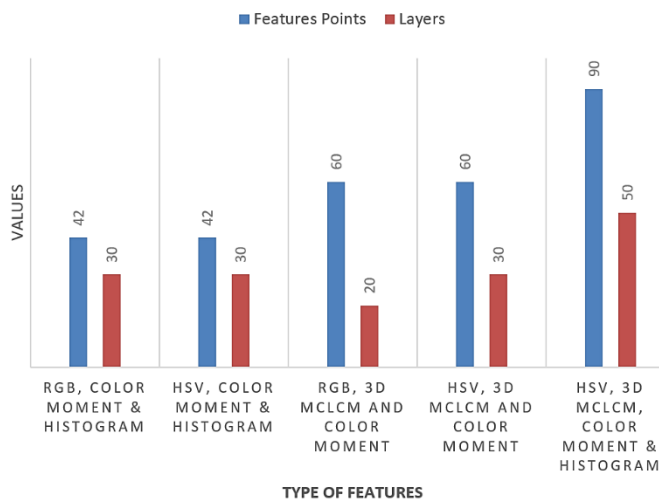

FIG. 5. FLOW CHART

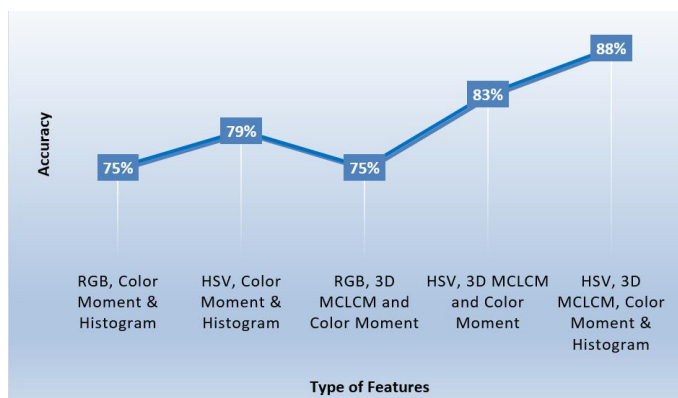FIG. 7. COMPARATIVE SUMMARY FOR FEATURES AND LAYERS
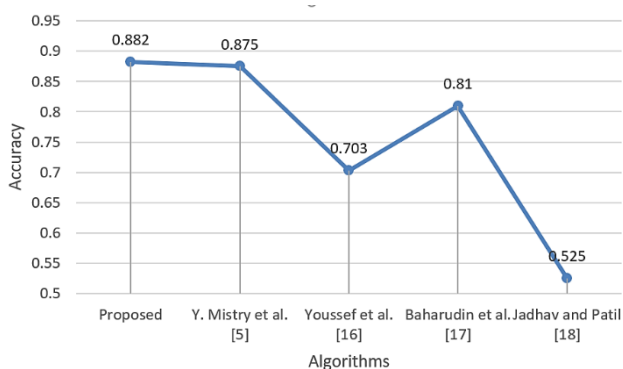

FIG. 8. RESULT ANALYSIS OF CLASSIFIERS

FIG. 9. RESULT ANALYSIS WITH OTHER NOVEL WORKS

In Figure-9 result of proposed method is compared with other similar novel works and found that proposed system outperforms all other methods in terms of accuracy.

V CONCLUSION

A new algorithm is proposed to evaluate the color and texture feature in combined form. So as benefits of proposed algorithm, it not also reduces the computational cost even also improves the accuracy of the CBIR system. Further, to evaluate and optimize the results new algorithm is tested with different features set and color domains. Proposed work is also compared with other novel works proposed by different authors and found that proposed algorithm outperforms all other in terms of accuracy. Limitation and further area of improvisation can be use of shape features in combined format so that all color, texture and shape can be calculated in one go with further lowering down the computational cost and accuracy can be enhanced.

VI Acknowledgement

References

[1] Zakhayu Rian, Viny Christanti, Janson Hendryli "Content-Based Image Retrieval using Convolutional Neural Networks" 2019 IEEE International Conference on Signals and Systems (ICSigSys).

[2] Jia Li, J.Z. Wang "Automatic Linguistic Indexing of Pictures by a statistical modeling approach" IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 25 , Issue: 9 , Sept. 2003).

[3] Luiz Gustavo S. Real, Renato Bueno, Marcela X. Ribeiro, "Evaluating Boundary Conditions and Hierarchical Visualization in CBIR", IEEE 2019 23rd International Conference Information Visualisation (IV).

[4] JunYue, ZhenboLi, LuLiu, ZetianFu, "Content-based image retrieval using color and texture fused features", Mathematical and Computer Modelling Volume 54, Issues 3–4, August 2011, Pages 1121-1127

[5] J.Z. Wang, Jia Li, G. Wiederhold "SIMPLIcity: semantics-sensitive integrated matching for picture libraries" IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 23 , Issue: 9 , Sep 2001).

[6] Miroslav MARINOV, Irena VALOVA, Yordan KALMUKOV "Comparative Analysis of Content-Based Image Retrieval Systems" 2019 16th Conference on Electrical Machines, Drives and Power Systems (ELMA).

[7] Pradnya Vikhar, Pravin Karde "Improved CBIR system using Edge Histogram Descriptor (EHD) and Support Vector Machine (SVM)" IEEE 2016 International Conference on ICT in Business Industry & Government (ICTBIG).

[8] YogitaMistry, D.T.Ingole, M.D.Ingole "Content based image retrieval using hybrid features and various distance metric" Journal of Electrical Systems and Information Technology,Volume 5, Issue 3, December 2018, Pages 874-888.

[9] Youssef, S.M., Mesbah, S., Mahmoud, Y.M., 2012. An efficient content-based image retrieval system integrating wavelet-based image sub-blocks with dominant colors and texture analysis. In: Information Science and Digital Content Technology (ICIDT), 2012 8th International Conference on. Vol. 3. IEEE, pp. 518–523.

[10] Baharudin, B., et al., 2012. Effective content-based image retrieval: combination of quantized histogram texture features in the dct domain. In, Computer & Information Science (ICCIS), 2012 International Conference on. Vol. 1. IEEE, pp. 425–430.

[11] Jadhav, S.M., Patil, V., 2012. An effective content based image retrieval (cbir) system based on evolutionary programming (ep). In: Advanced Communication Control and Computing Technologies (ICACCCT), 2012 IEEE International Conference on. IEEE, pp. 310–315.

[12] Mayuri Sadafale, S. V. Bonde "Spatio-frequency local descriptor for content based image retrieval" 2017 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)

[13] Wei Bian, Dacheng Tao "Biased Discriminant Euclidean Embedding for Content-Based Image Retrieval" IEEE Transactions on Image Processing (Volume: 19 , Issue: 2 , Feb. 2010)

[14] Komal Juneja, Akhilesh Verma, Savita Goel, Swati Goel "A Survey on Recent Image Indexing and Retrieval Techniques for Low-Level Feature Extraction in CBIR Systems" 2015 IEEE International Conference on Computational Intelligence & Communication Technology.

[15] Muhsina Kaipravan, Rejiram R "A novel CBIR system based on combination of color moment and Gabor filter" IEEE 2016 International Conference on Data Mining and Advanced Computing (SAPIENCE).

[16] Smith, A. R. "Color Gamut Transform Pairs". SIGGRAPH 78 Conference Proceedings. 1978, pp. 12–19.

[17] Benco M, Hudec R, Kamencay P, et al. "An advanced approach to extraction of color texture features based on GLCM" Int J Adv Robot Syst 2014:1–8.

[18] Haralick RM, Shanmugam K. Dinstein Its'Hak textural features for image classification,system,man and cybernetics SMC IEEE Trans.1973;3(6):610-21

[19] Wei Bian, Dacheng Tao "Biased Discriminant Euclidean Embedding for Content-Based Image Retrieval" IEEE Transactions on Image Processing (Volume: 19 , Issue: 2 , Feb. 2010)

[20] MathWorks, (2020a). Deep Learning Toolbox: User's Guide (R2020a).

[21] Rafeek Mamdouh, Hazem M. El-Bakry, Alaa Riad, Nashaat El-Khamisy, Converting 2D-Medical Image Files "DICOM" into 3D-Models, based on Image Processing, and Analysing their Results with Python Programming, WSEAS Transactions on Computers, ISSN / E-ISSN: 1109-2750 / 2224-2872, Volume 19, 2020, Art. #2, pp. 10-20.

[22] Jose Augusto Cadena Moreano, Nora Bertha La Serna Palomino, Efficient Technique for Facial Image Recognition with Support Vector Machines in 2D Images with Cross-Validation in Matlab, WSEAS Transactions on Systems and Control, ISSN / E-ISSN: 1991-8763 / 2224-2856, Volume 15, 2020, Art. #18, pp. 175-183.