



Dual Phase CBIR Model using Hybrid Feature Extraction and Manhattan Distance Measure

Yashaswini Doddamane Kenchappa^{1*} Kwadaki Karibasappa²

¹Department of Information science and engineering, DonBosco Institute of Technology, India

²Department of Electrical and Computer Engineering,
Dayananda Sagar Academy of Technology and Management, India

* Corresponding author's Email: yashaswinidk5@gmail.com

Abstract: Recently, the multimedia complexity is noticeably increased due to the growth of internet and multimedia technologies and it has become one of the prime research topics in multimedia content retrieval. In this research, a dual phase model is proposed to enhance the performance of content based image retrieval. Initially, normalization technique is employed to improve the visibility level of the images that are collected from Wang and COREL datasets. Next, color moments and Alex-Net convolutional neural network are used to extract the feature values from the normalized image. By combining low and high level features, the semantic space between the extracted feature values is less that helps in retrieving the relevant content images significantly. Finally, the Manhattan distance measure is applied to measure the distance between the query image and an image dataset to retrieve the relevant images. The experimental result showed that the proposed dual phase model achieved better performance in content based image retrieval by means of recall, precision and f-score. From the comparative analysis, the proposed dual phase model showed a maximum of 0.43 precision value and 0.06 recall value improvement in content based image retrieval compared to the existing works like Bi-layer system, spatial and frequency domain features with Euclidean distance, and local directional pattern with color histograms.

Keywords: Alex-Net convolutional neural network, Color moments, Content based image retrieval, Image normalization, Manhattan distance measure.

1. Introduction

In recent decades, various photography devices like mobile phones, web camera, and digital camera generates huge image datasets, where searching, saving, managing and organizing of digital datasets has become an indispensable procedure [1, 2]. In order to handle large dataset, Content Based Image Retrieval (CBIR) procedure is developed to retrieve the relevant image from the dataset based on the similarity between the query image and an image dataset [3, 4]. Currently, several local, global and deep learning features are developed for CBIR by considering different mid-level representations and different visual properties like texture, color and shape [5, 6]. The core issue in CBIR is to estimate the relevance of a query image with respect to the dataset

images [7]. Most of the existing approaches mainly concentrated on matching the similar images by leveraging multi-visual feature. The visual space cannot characterize the relevance from a semantic perspective, due to the semantic space between higher level and lower level concepts [8-10]. In order to address the aforementioned concern, a dual phase model is proposed in this paper to improve the performance of CBIR.

At first, the input images are collected from Wang and COREL datasets and then image normalization technique is applied to improve the visibility level of the collected images by altering the range of image pixels. Next, hybrid feature extraction is accomplished by using color moments and Alex-net Convolutional Neural Network (CNN) feature to extract the feature vectors from the normalized images. In comparison to the feature selection

methodologies, hybrid feature extraction is less exposed to overfitting issue and perform good retrieval accuracy. Meanwhile, the semantic gap between the extracted feature vectors is less by combining the lower and higher-level feature vectors that results in better retrieval accuracy. After feature extraction, Manhattan distance measure is used to retrieve the relevant content images. In the experimental section, the developed dual phase model performance is analysed in light of recall, precision, f-score.

This research paper is prepared as follows: In Section 2, a few recent research papers on the topic "CBIR" is reviewed. In Section 3, the proposed dual phase model is explained briefly with mathematical expressions. The experimental analysis of the proposed dual phase model is indicated in the Section 4. The conclusion of the present research work is given in the Section 5.

2. Literature survey

P. Chhabra, N.K. Garg, and M. Kumar [11] used two feature descriptors; Scale Invariant Feature Transform (SIFT) and Oriented Fast and Rotated BRIEF (ORB) to extract the features from the images for effectively retrieving the relevant content based images. After extracting the features, k-means clustering and locality preserving projection approaches were utilized to decrease the dimensions of the extracted features. In retrieval phase, the optimal features were stored in the dataset and then matched with testing features for CBIR. Finally, trained classifiers like random forest, multilayer perceptron and decision tree utilized to predict the similarity between whole dataset features and testing data features for retrieving the relevant images. In this literature, the developed model performance was tested on COREL and Wang datasets and the simulation outcome showed that the developed model achieved better performance in CBIR by means of retrieval time, precision, and root mean square error. To retrieve desired information from the bulky dataset is still a major issue in this literature study. R. Bibi, Z. Mehmood, R.M. Yousaf, T. Saba, M. Sardaraz, and A. Rehman [12] used sparse complementary features to extract the feature vectors from the images. Next, fuzzy c-means clustering algorithm and locality-preserving projection were employed for optimal feature selection. The obtained optimal feature vectors were fed to soft label support vector machine for image classification. In this literature, corel-1500, holidays and corel-1000 datasets are used to validate the performance of the developed model in terms of retrieval accuracy, recall

and precision. The developed model has sensitivity problems in case of complex deformations like change in viewpoint and rotation.

Y.D. Mistry [13] used fuzzy histogram based descriptors, segmentation based fractal texture analysis, Gabor wavelet texture features, binary Gabor pattern, and edge histogram descriptor for extracting texture and color feature vectors from the images. Next, Laplacian score was utilized to reduce the dimension of the extracted feature vectors. Finally, the relevant images were retrieved using Euclidean and city block distance measures. The extensive experiment performed on Wang dataset and the simulation outcome showed that the developed model attained better performance in CBIR in light of precision. The semantic gap between the extracted features is high in this study that may lead to poor retrieval performance. A. Shakarami, and H. Tarrah, [14] used local binary pattern, Alexnet convolutional neural network, and histogram of oriented descriptors for extracting the feature vectors from the images that were collected from Corel-1000, Oliva and Torralba (OT), and Caltech-101 datasets. Next, principle component analysis was used to reduce the dimension of the extracted feature vectors. This Simulation result showed that the developed model achieved a good performance in CBIR by means of accuracy and mean average precision. The undertaken feature descriptors consider only certain features like texture for feature extraction which may ineffective in retrieving the relevant images from the large datasets.

S. Singh, and S. Batra [15] developed a bi-layer CBIR system that majorly comprises of two phases. In the initial phase, feature vectors were extracted from the images by means of texture, shape, and color of the images. The second phase consists of two layers; initially the images were matched with query image and the relevant images were retrieved based on texture and shape feature space. Again, the retrieved images were matched with query image and the relevant content images were retrieved based on color and shape feature space. Extensive experiments showed that the developed bi-layer CBIR system outperformed the existing systems in light of precision, f-score, and recall. The developed system significantly overcomes the dimensionality issue, but it is sensitivity to inter class and intra class varieties. J. B. Naik, G. B. Kande, and C. Srinivasarao [16] introduced a new CBIR system based on a non-linear similarity matching measure and an informative pattern descriptor. In the developed system, similarity measure uses the similarity degree in logarithmic scale and then the similar images were retrieved based on the logarithmic value. The

developed CBIR system needs manual intervention to match the query and dataset images, which is considered as a major challenge in this literature study.

Y. Mistry, D. T. Ingole, and M. D. Ingole, [17] developed a hybrid CBIR system by using numerous distance measures. In this literature study, frequency domain feature like stationary wavelet transform and Gabor wavelet transform, and spatial domain feature like color moment, color auto-correlogram and HSV histogram feature were used for feature extraction. After feature extraction, various distance measures; Euclidean, Minkowski, City block and Mahalanobis were utilized to retrieve the relevant content based images. In the experimental phase, developed hybrid CBIR system attained better performance by means of precision value. However, semantic similarity and domain variance are still challenging concerns in the developed hybrid system. J. X. Zhou, X. D. Liu, T. W. Xu, J. H. Gan, and W. Q. Liu, [18] developed a new color image retrieval system using Local Directional Pattern (LDP) and color histograms. At first, feature vectors were extracted from the images by utilizing color histograms and LDP. Then, an efficient fusion process was developed, which includes a similarity metric and feature normalization. The simulation outcome showed that the developed system attained effective performance on two benchmark datasets; Wang and Corel 10000 datasets in light of recall and precision. The developed system retrieved relevant images based on the contents like color, shape and texture which is ineffective in retrieving the relevant images from large datasets.

In order to highlight the aforementioned issues, a new dual phase model is proposed in this research paper.

3. Methodology

In recent decades, an enormous number of image datasets are developed due to the rapid growth of multimedia technologies and internet. An effective retrieval and search technique is required to retrieve the relevant content images from the datasets that is helpful in many applications like three dimensional model retrieval, near-duplicate video detection, medical diagnosis, etc. [19-20]. The traditional image retrieval process is performed by characterizing image with a text annotation and retrieve the relevant images by searching the keywords [21]. The traditional image retrieval process is ambiguous and laborious, due to the diversity of the image contents and rapid increase in the number of images [22]. To highlight the aforementioned issue, a dual phase

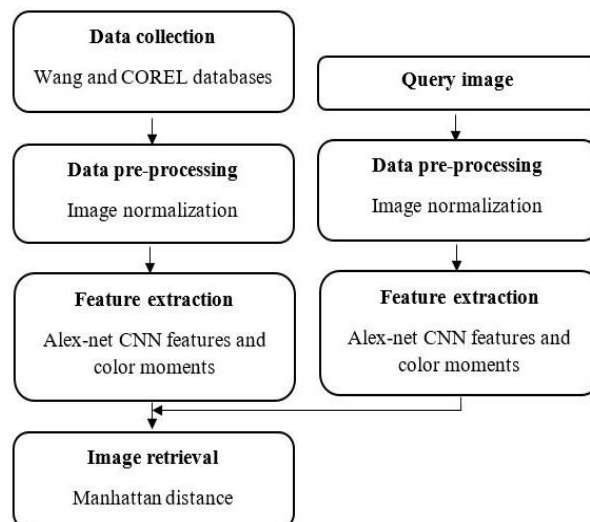


Figure. 1 Flow diagram of proposed dual phase model

model is proposed in this research to improve the image retrieval performance. In the proposed dual phase model, the image normalization technique is utilized to improve the visibility level of the images that are collected from Wang and COREL datasets. Next, feature extraction is accomplished using Alex-net CNN features and color moments and the extracted feature vectors are fed to Manhattan distance measure for an effective image retrieval. The work flow of the proposed dual phase model is graphically presented as Fig. 1.

3.1 Data collection and pre-processing

In this research work, Wang and COREL datasets are used to analyse the performance of the proposed dual phase model. Wang dataset consists of 1000 images for 10 classes such as buildings, dinosaurs, mountain, African people, elephant, bus, beach, horse, flower, and food in JPEG format with the image pixel size of 256×384 and 384×256 [23]. Wang dataset is extensively utilized to investigate the performance of CBIR methods, because it includes a variety of class information and limited dataset size. The sample images of Wang dataset are graphically represented in Fig. 2. In addition, COREL dataset consists of 6410 images for 40 classes that includes dog, elephant, door, indoor, iceberg, mountain, woman, flower, train, eagle, waterfall, cat, etc. The

Table 1. Statistics of COREL and wang datasets

Dataset	Image size	Number of classes	Total number of images
COREL	80×120 and 120×80	40	6410
Wang	384×256 and 256×384	10	1000



Figure. 2 Sample images of wang dataset



Figure. 3 Sample images of COREL dataset

sample images of COREL dataset are graphically indicated in Fig. 3. The statistics of COREL and Wang datasets are denoted in Table 1.

After collecting the images from Wang and COREL datasets, data pre-processing is carried out by linear normalization technique [24] and it enhances the visibility level of the collected images by altering the range of image pixel intensity values. The formula of linear normalization technique is mathematically represented in Eq. (1).

$$I_n = (I - Min) \times \frac{newMax - newMin}{Max - Min} + newMin \quad (1)$$

Where, I_n is specified as normalized images, I is represented as collected images, and $Max - Min$ is



Figure. 4 Normalized images of wang dataset



Figure. 5 Normalized images of COREL dataset

stated as minimum and maximum range of pixel value that ranges between 0 to 255. The sample normalized images of Wang and COREL datasets are denoted in the Fig. 4 and 5.

3.2 Hybrid feature extraction

After image pre-processing, feature extraction is accomplished using Alex-net CNN features [25] and color moments [26] to extract the feature vectors from the images. At first, pre-trained Alex-net model with Image Net dataset weights are used for feature extraction. Here, the feature sets are extracted with Image Net dataset learned filters of the CNN architecture. In pre-trained CNN model, 1st group layers include lower level feature sets that contain information about the image edges. In addition, the fully connected layer of 4096 neurons are added to the convolutional layers to collect the feature vectors from the convolutional layers. Next, color moments (mean, standard deviation and skewness) are utilized to extract the features f from the normalized images. Usually, the color information is distributed in the lower end moments; mean μ , standard deviation σ and skewness s that are mathematically defined in the Eqs. (2), (3), and (4).

$$\mu_{r,j} = \frac{1}{N} \sum_{i=1}^N p_{i,j} \quad j = H, S, V \quad (2)$$

$$\sigma_{r,j} = \left(\frac{1}{N} \sum_{i=1}^N (p_{i,j} - \mu_{r,j})^2 \right)^{1/2} \quad j = H, S, V \quad (3)$$

$$s_{r,j} = \left(\frac{1}{N} \sum_{i=1}^N (p_{i,j} - \sigma_{r,j})^3 \right)^{1/3} \quad j = H, S, V \quad (4)$$

Where, H, S and V are represented as hue, saturation and value, N is indicated as number of pixels, r is represented as number of images, and p_{ij} is specified as probability of j^{th} color channel at i^{th} location in the normalized image. Then, the extracted features f are given as the input to Manhattan distance measure for retrieving the relevant content based images.

3.3 Distance measure

After extracting the feature vectors from the query x_f and dataset images y_f , Manhattan distance measure is used to retrieve the relevant content images [27]. Manhattan distance measure is also named as Manhattan length or L_1 distance measure that calculates the distance between two data points in a grid like path manner. In this distance measure, diagonal distance is calculated by using the Pythagorean Theorem. The Manhattan distance between the query x_f and dataset images y_f is

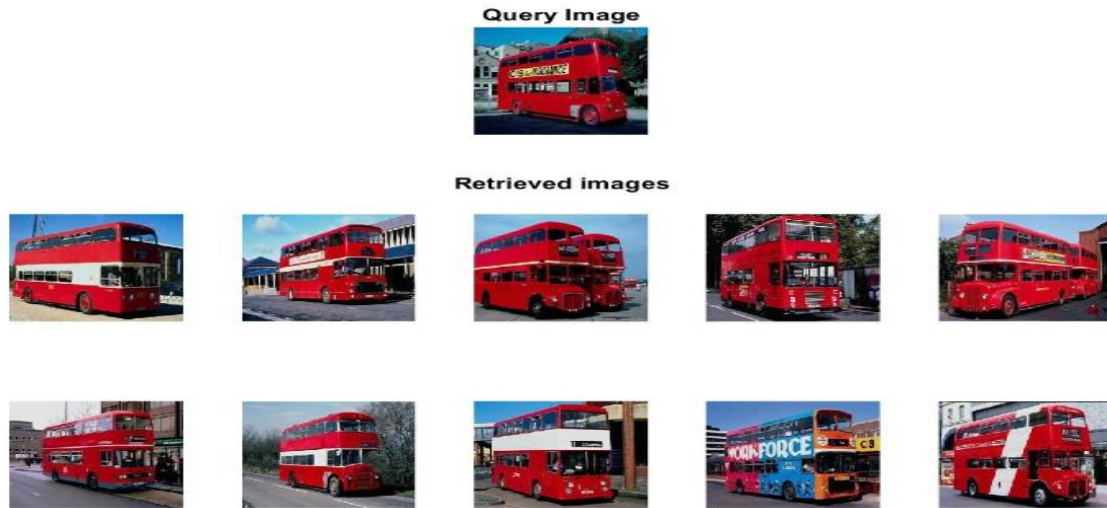


Figure. 6 Sample retrieval output image of wang dataset



Figure. 7 Sample retrieval output image of COREL dataset

mathematically represented in Eq. (5).

$$\text{Manhattan distance } (x_f, y_f) = |x_{f1} - y_{f1}| + |x_{f2} - y_{f2}| \quad (5)$$

If the data points have n dimensions like $x = (x_{f1}, x_{f2} \dots, x_{fn})$ and $y = (y_{f1}, y_{f2} \dots, y_{fn})$, and the Eq. (5) is updated as shown in Eq. (6).

$$\text{Manhattan distance } (x_f, y_f) = |x_{f1} - y_{f1}| + |x_{f2} - y_{f2}| + \dots + |x_{fn} - y_{fn}| = \sum_{i=1}^n |x_{fi} - y_{fi}| \quad (6)$$

Where, $i = 1, 2 \dots n$. The sample retrieval output image of Wang and CORLE datasets are given in the Fig. 6 and 7.

4. Experimental results

In this research article, the proposed dual phase model is simulated by using MATLAB (2018a) tool with the system requirements; Operating System: windows 10, RAM: 16GB, and Processor: Intel core i7. Here, the proposed dual phase model performance is compared with the benchmark models like bi-layer CBIR system [15], spatial and frequency domain features with Euclidean distance [17], and LDP with color histogram [18] on Wang and COREL datasets to validate its efficiency. In this scenario, the proposed dual phase model performance is evaluated using precision, f-score, and recall. Precision is defined as the number of relevant images retrieved by a search divided by the total number of images retrieved. Meanwhile, recall is defined as the number of images retrieved by a search divided by the total number of existing relevant images. F-score is

Table 2. Performance analysis of the proposed dual phase model with different distance measures on wang dataset

Classes	Hybrid feature with Euclidian			Hybrid feature with Cosine			Hybrid feature with Manhattan		
	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score
African people	0.79	0.09	0.03	0.82	0.06	0.13	0.89	0.1	0.14
Beach	0.77	0.02	0.16	0.76	0.06	0.15	0.84	0.11	0.19
Buildings	0.72	0.02	0.02	0.78	0.08	0.06	0.82	0.1	0.08
Bus	0.94	0.02	0.18	0.97	0.08	0.15	1	0.12	0.22
Dinosaurs	0.98	0.06	0.17	0.98	0.11	0.15	1	0.13	0.22
Elephant	0.91	0.03	0.17	0.95	0.09	0.15	0.97	0.12	0.22
Flowers	1	0.08	0.2	0.96	0.07	0.15	1	0.13	0.22
Food	0.88	0.1	0.02	0.83	0.06	0.05	0.88	0.11	0.1
Horse	0.95	0.08	0.2	0.99	0.11	0.17	1	0.12	0.22
Mountains	0.91	0.08	0.18	0.9	0.04	0.12	0.91	0.11	0.2
Average	0.91	0.03	0.07	0.91	0.04	0.11	0.93	0.12	0.18

Table 3. Performance analysis of the proposed dual phase model with individual and hybrid feature extraction on wang dataset

Classes	Color moment with Manhattan			Alex-net CNN with Manhattan			Hybrid feature with Manhattan		
	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score
African people	0.5	0.06	0.01	0.76	0.1	0.14	0.89	0.1	0.14
Beach	0.3	0.04	0.07	0.83	0.11	0.2	0.84	0.11	0.19
Buildings	0.5	0.06	0.01	0.82	0.1	0.11	0.82	0.1	0.08
Bus	0.25	0.03	0.06	1	0.12	0.22	1	0.12	0.22
Dinosaurs	0.45	0.01	0.01	0.99	0.13	0.22	1	0.13	0.22
Elephant	0.75	0.06	0.01	0.97	0.12	0.22	0.97	0.12	0.22
Flowers	0.6	0.12	0.21	1	0.12	0.22	1	0.13	0.22
Food	0.8	0.1	0.08	0.9	0.11	0.08	0.88	0.11	0.1
Horse	0.75	0.12	0.21	0.99	0.12	0.22	1	0.12	0.22
Mountains	0.3	0.04	0.07	0.9	0.11	0.2	0.91	0.11	0.2
Average	0.52	0.07	0.07	0.92	0.11	0.18	0.93	0.12	0.18

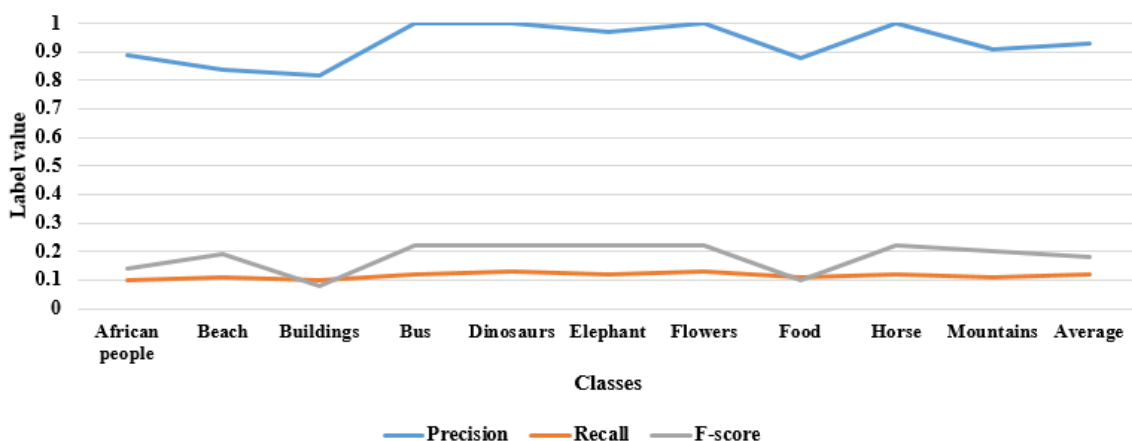


Figure. 8 Graphical analysis of the proposed dual phase model on Wang dataset

determined as a measure of test’s accuracy and it is the harmonic mean of the recall and precision. The mathematical expressions of precision, recall and f-score are defined in the Eq. (7-9).

$$Precision = \frac{|I(Ret) \cap I(Rel)|}{|I(Ret)|} \tag{7}$$

$$Recall = \frac{|I(Ret) \cap I(Rel)|}{|I(Rel)|} \tag{8}$$

Table 4. Performance analysis of the proposed dual phase model with different distance measures on COREL dataset

Classes	Hybrid feature with Euclidian			Hybrid feature with Cosine			Hybrid feature with Manhattan		
	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score
Car	0.98	0.07	0.22	0.98	0.13	0.21	1	0.13	0.22
Dog	0.88	0.1	0.09	0.95	0.1	0.1	0.95	0.11	0.1
Door	0.9	0.12	0.18	0.94	0.11	0.21	0.96	0.12	0.21
Eagle	0.96	0.09	0.19	0.99	0.12	0.21	1	0.13	0.22
Elephant	0.63	0.08	0.06	0.63	0.08	0.08	0.64	0.08	0.08
Flower	0.92	0.11	0.18	0.98	0.11	0.21	0.98	0.12	0.22
Indoor	0.88	0.05	0.15	0.95	0.12	0.2	0.95	0.12	0.21
Mountain	0.99	0.11	0.15	0.97	0.12	0.21	0.99	0.12	0.22
Train	0.88	0.04	0.19	0.94	0.1	0.2	0.95	0.12	0.21
Cat	0.96	0.07	0.18	0.99	0.11	0.21	1	0.12	0.22
Average	0.92	0.11	0.18	0.93	0.1	0.18	0.94	0.12	0.19

Table 5. Performance analysis of the proposed dual phase model with individual and hybrid feature extraction on COREL dataset

Classes	Color moment with Manhattan			Alex-net CNN with Manhattan			Hybrid feature with Manhattan		
	Precision	Recall	F-score	Precision	Recall	F-score	Precision	Recall	F-score
Car	0.51	0.06	0.09	1	0.13	0.22	1	0.13	0.22
Dog	0.35	0.12	0.16	0.93	0.12	0.21	0.95	0.11	0.1
Door	0.55	0.14	0.02	0.97	0.12	0.22	0.96	0.12	0.21
Eagle	0.29	0.07	0.09	0.99	0.12	0.22	1	0.13	0.22
Elephant	0.54	0.03	0.05	0.6	0.08	0.22	0.64	0.08	0.08
Flower	0.79	0.15	0.02	0.97	0.12	0.21	0.98	0.12	0.22
Indoor	0.66	0.2	0.23	0.93	0.12	0.21	0.95	0.12	0.21
Mountain	0.87	0.17	0.1	0.96	0.12	0.21	0.99	0.12	0.22
Train	0.83	0.14	0.26	0.94	0.12	0.21	0.95	0.12	0.21
Cat	0.31	0.05	0.13	1	0.12	0.22	1	0.12	0.22
Average	0.57	0.11	0.11	0.93	0.12	0.22	0.94	0.12	0.19

$$F - score = \frac{2 \times precision \times recall}{Precision + recall} \tag{9}$$

Where, $I(Rel)$ is denoted as relevant images, and $I(Ret)$ is represented as retrieved images.

4.1 Quantitative investigation on wang dataset

In this section, the proposed dual phase model performance is evaluated on the Wang dataset by means of precision, recall, and f-score. The Wang dataset comprises of 1000 images for 10 classes; buildings, dinosaurs, mountains, African people, elephant, bus, beach, horse, flower, and food with the pixel size of 384×256 and 256×384 . In Table 2, the performance analysis carried out with different distance measures (Euclidean, cosine, and Manhattan) in combination with hybrid feature extraction. By analyzing Table 2, the combination; hybrid feature with Manhattan achieved better CBIR performance compared to other two combinations by means of precision, recall, and f-score. By analysing

the average result of all classes, hybrid feature with Manhattan distance measure showed a maximum of 0.02 value improvement in precision. Additionally, the combination; hybrid feature with Manhattan distance showed maximum of 0.09 and 0.11 value improvement in recall, and f-score related to other combinations; hybrid feature with cosine and hybrid feature with Euclidean. The graphical comparison of the proposed dual phase model; hybrid feature with Manhattan on Wang dataset is denoted in Fig. 8.

In Table 3, the performance analysis is carried out by individual and hybrid feature extraction in combination with Manhattan distance measure. As same as Table 2, the hybrid feature extraction with Manhattan distance achieved better performance in CBIR by means of precision, recall, and f-score. By analysing Table 3, the hybrid feature with Manhattan showed a maximum of 0.41, 0.05, and 0.11 value improvement in precision, recall, and f-score compared to individual feature extraction with Manhattan distance measure. In this research paper, semantic gap between the extracted feature vectors is

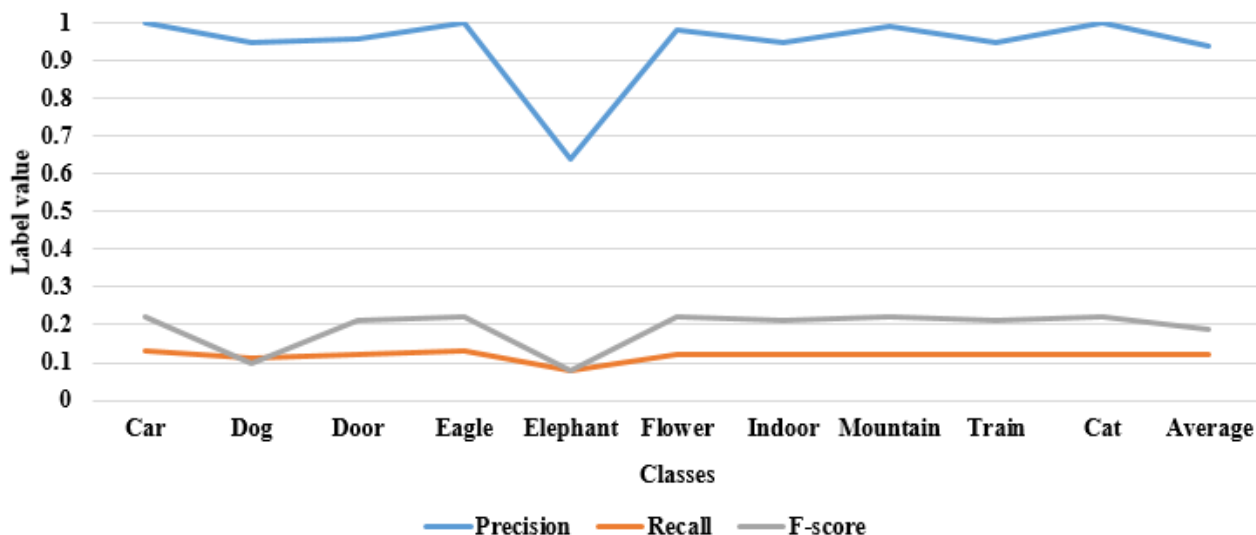


Figure. 9 Graphical analysis of the proposed dual phase model on COREL dataset

very low by utilizing both higher and lower level features that helps to achieve effective performance in CBIR.

4.2 Quantitative investigation on COREL dataset

In this section, the dual phase model performance is evaluated on COREL dataset by means of precision, recall and f-score. The COREL dataset contains 6410 images for 40 classes with the size of 80×120 and 120×80 . By analyzing the average result of all classes, the combination; hybrid feature with Manhattan distance achieved better CBIR performance compared to other two combinations. In Table 4, the proposed hybrid feature with Manhattan almost showed a maximum of 0.02, 0.02, and 0.01 value improvement in precision, recall and f-score values compared to other combinations. Related to other distance measures, Manhattan distance is very robust to outliers that retrieves the relevant content images effectively.

In Table 5, the performance analysis is carried out by individual and hybrid feature extraction in combination with Manhattan distance on COREL dataset. As similar to Table 3, the proposed dual phase model (hybrid feature with Manhattan distance) achieved better performance in CBIR related to individual features with Manhattan distance. Graphical comparison of the proposed dual phase model on COREL dataset is represented in Fig. 9.

4.3 Comparative investigation

The comparative investigation of the proposed and the existing systems is given in Tables 6 and 7. S. Singh, and S. Batra, [15] introduced a bi-layer CBIR

system that majorly contains two steps. Based on shape, texture and color, feature vectors were extracted from the images which were collected from Wang dataset. In the next step, the collected images were matched with query image and the relevant images were retrieved on the basis of shape and texture feature space. In addition, the retrieved content images were matched once again with query image and the relevant images were retrieved on the basis of shape and color feature space. Experimental outcome showed that the developed Bi-layer system achieved average precision value of 0.92. Y. Mistry, D. T. Ingole, and M. D. Ingole, [17] used frequency domain features like stationary wavelet transform

Table 6. Comparative investigation of the proposed and existing models in light of precision

Wang dataset			
Classes	Bi-layer system [15]	Spatial and frequency domain features with Euclidean distance [17]	Dual phase model
African people	0.95	0.81	0.89
Beach	0.75	0.92	0.84
Buildings	0.85	0.79	0.82
Bus	1	0.93	1
Dinosaurs	1	0.99	1
Elephant	0.9	0.79	0.97
Flowers	1	0.99	1
Food	0.95	0.80	0.88
Horse	1	0.85	1
Mountains	0.8	0.88	0.91
Average	0.92	0.875	0.93

Table 7. Comparative investigation of the proposed and existing models in light of precision and recall

Wang dataset		
Method	Precision	Recall
LDP with color histogram [18]	0.5	0.06
Dual phase model	0.93	0.12

and Gabor wavelet transform, and spatial domain features like color moment, color auto-correlogram and HSV histogram features for feature extraction. Then, the extracted feature vectors were fed to four distance measures such as Euclidean, city block, Minkowski, and Mahalanobis in order to retrieve the relevant content based images. In that, spatial and frequency domain features with Euclidean distance achieved better average precision value of 0.875 on Wang dataset. Related to these existing works, the proposed dual phase model achieved better average precision value of 0.93. In this research paper, the proposed dual phase model combined low and high level features, so the semantic gap between the extracted features is less that helps in achieving better retrieval performance and also to overcome the traditional issues like semantic similarity, partial similarity, and domain variance.

J. X. Zhou, X. D. Liu, T. W. Xu, J. H. Gan, and W.Q. Liu, [18] developed an image retrieval system on the basis of LDP and color histograms. In this literature, an effective fusion process was developed that includes similarity metrics (Euclidean and city-block) and feature normalization. The experimental outcome showed that the developed system achieved average precision value of 0.5 and the average recall value of 0.06 on Wang dataset. Hence, the proposed dual phase model achieved average precision value of 0.93 and average recall value of 0.12 on Wang dataset, which are better related to the existing work. In the proposed work, color moments and Alex-Net convolutional neural network represents the spatial correlation of texture and color orientations and also has better discrimination power with color, texture and shape features that overcomes the problems like inter and intra class varieties in the collected dataset.

5. Conclusion

In this research paper, a dual phase model is proposed to enhance the performance of CBIR. The proposed model majorly includes two phases; hybrid feature extraction and a distance measure to retrieve the relevant content images from the large datasets. The color moment and Alex-net CNN features are applied for extracting the feature vectors and then Manhattan distance measure is used to identify the distance between the query and dataset images for relevant image retrieval. The experimental results

showed that the proposed dual phase model attained better CBIR performance compared to the existing works like Bi-layer system, spatial and frequency domain features with Euclidean distance, and LDP with color histogram. On Wang dataset, the proposed dual phase model showed maximum of 0.43 precision value and 0.06 recall value improvement in CBIR related to the existing works. In the future work, a new deep learning concept is included in the proposed dual phase model to further improve the performance of CBIR.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st author. The supervision and project administration, have been done by 2nd author.

References

- [1] M. K. Alsmadi, "Content-Based Image Retrieval Using Color, Shape and Texture Descriptors and Features", *Arabian Journal for Science and Engineering*, Vol. 45, pp. 3317–3330, 2020.
- [2] M. Agarwal and R. P. Maheshwari, "Multichannel Local Ternary Co-occurrence Pattern for Content-Based Image Retrieval", *Transactions of Electrical Engineering*, Vol. 44, No. 1, pp. 495-504, 2020.
- [3] M. Tzelepi and A. Tefas, "Deep convolutional learning for content based image retrieval", *Neurocomputing*, Vol. 275, pp. 2467-2478, 2018.
- [4] K. T. Ahmed, S. Ummesafi, and A. Iqbal, "Content based image retrieval using image features information fusion", *Information Fusion*, Vol. 51, pp. 76-99, 2019.
- [5] R. S. Bressan, P. H. Bugatti, and P. T. Saito, "Breast cancer diagnosis through active learning in content-based image retrieval", *Neurocomputing*, Vol. 357, pp. 1-10, 2019.
- [6] C. Cui, P. Lin, X. Nie, Y. Yin, and Q. Zhu, "Hybrid textual-visual relevance learning for content-based image retrieval", *Journal of Visual Communication and Image Representation*, Vol. 48, pp. 367-374, 2017.
- [7] D. C. G. Pedronette and R. D. S. Torres, "Unsupervised rank diffusion for content-based

- image retrieval”, *Neurocomputing*, Vol. 260, pp. 478-489, 2017.
- [8] S. Fadaei, R. Amirfattahi, and M. R. Ahmadzadeh, “Local derivative radial patterns: a new texture descriptor for content-based image retrieval”, *Signal Processing*, Vol. 137, pp. 274-286, 2017.
- [9] Kumar, S. Dyer, J. Kim, C. Li, P. H. Leong, M. Fulham, and D. Feng, “Adapting content-based image retrieval techniques for the semantic annotation of medical images”, *Computerized Medical Imaging and Graphics*, Vol. 49, pp. 37-45, 2016.
- [10] D. Feng, J. Yang, and C. Liu, “An efficient indexing method for content-based image retrieval”, *Neurocomputing*, Vol. 106, pp. 103-114, 2013.
- [11] P. Chhabra, N. K. Garg, and M. Kumar, “Content-based image retrieval system using ORB and SIFT features”, *Neural Computing and Applications*, Vol. 32, No. 7, pp. 2725-2733, 2020.
- [12] R. Bibi, Z. Mehmood, R. M. Yousaf, T. Saba, M. Sardaraz, and A. Rehman, “Query-by-visual-search: multimodal framework for content-based image retrieval”, *Journal of Ambient Intelligence and Humanized Computing*, Vol. 11, pp. 5629–5648, 2020.
- [13] Y. D. Mistry, “Textural and color descriptor fusion for efficient content-based image retrieval algorithm”, *Iran Journal of Computer Science*, Vol. 3, No. 3, pp. 169-83, 2020.
- [14] Shakarami and H. Tarrach, “An efficient image descriptor for image classification and CBIR”, *Optik*, Vol. 214, pp. 164833, 2020.
- [15] S. Singh and S. Batra, “An efficient bi-layer content based image retrieval system”, *Multimedia Tools and Applications*, Vol. 79, pp. 17731–17759, 2020.
- [16] J. B. Naik, G. B. Kande, and C. Srinivasarao, “Logarithmic distance measure with improved local vector pattern for content-based image retrieval”, *The Imaging Science Journal*, Vol. 66, No. 4, pp. 239-253, 2018.
- [17] Y. Mistry, D. T. Ingole, and M. D. Ingole, “Content based image retrieval using hybrid features and various distance metric”, *Journal of Electrical Systems and Information Technology*, Vol. 5, No. 3, pp. 874-888, 2018.
- [18] J. X. Zhou, X. D. Liu, T. W. Xu, J. H. Gan, and W. Q. Liu, “A new fusion approach for content based image retrieval with color histogram and local directional pattern”, *International Journal of Machine Learning and Cybernetics*, Vol. 9, No. 4, pp. 677-689, 2018.
- [19] H. Sun, Z. Zhao, Q. Tian, X. Jin, L. Zhang, and B. Li, “Content-based image retrieval using color vector angle difference histogram”, *Journal of the Chinese Institute of Engineers*, Vol. 40, No. 3, pp. 246-256, 2017.
- [20] K. F. Hwang, Y. F. Chang, and K. Y. Chung, “A content-based image retrieval system based on object-moment feature”, *The Imaging Science Journal*, Vol. 62, No. 8, pp. 415-422, 2014.
- [21] M. K. Alsmadi, “An efficient similarity measure for content based image retrieval using memetic algorithm”, *Egyptian Journal of Basic and Applied Sciences*, Vol. 4, No. 2, pp. 112-122, 2017.
- [22] O. A. Adegbola, D. O. Aborisade, S. I. Popoola, O. A. Amole, and A. A. Atayero, “Modified one-class support vector machine for content-based image retrieval with relevance feedback”, *Cogent Engineering*, Vol. 5, No. 1, pp. 1541702, 2018.
- [23] D. Giveki, “A new content based image retrieval model based on wavelet transform”, *Journal of Computer and Communications*, Vol. 3, No. 3, pp. 66, 2015.
- [24] K. M. Koo and E. Y. Cha, “Image recognition performance enhancements using image normalization”, *Human-centric Computing and Information Sciences*, Vol. 7, pp. 33, 2017.
- [25] A. Barbhuiya, R. K. Karsh, and R. Jain, “CNN based feature extraction and classification for sign language”, *Multimedia Tools and Applications*, 2020.
- [26] J. Jing, Q. Li, P. Li, and L. Zhang, “A new method of printed fabric image retrieval based on color moments and gist feature description”, *Textile Research Journal*, Vol. 86, No. 11, pp. 1137-1150, 2016.
- [27] L. Greche, M. Jazouli, N. Es-Sbai, A. Majda, and A. Zarghili, “Comparison between Euclidean and Manhattan distance measure for facial expressions classification”, In: *Proc. of International Conf. on Wireless Technologies, Embedded and Intelligent Systems (WITS)*, pp. 1-4, 2017.
- [28] Dataset links:
Wang: <https://www.kaggle.com/rdmisal/wang-dataset-keras/data>
COREL: <https://sites.google.com/site/dctresearch/Home/content-based-image-retrieval>