Received: July 26, 2022. Revised: September 22, 2022.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Enhanced Content-Based Image Retrieval with Trio-Deep Feature Extractors with Multi-Similarity Function

G. Krishna Raju[1]* **P. Padmanabham[2]** **A. Govardhan[3]**

[1]*Media & Communication Arts Department, Matrusri PG College, Hyderabad, India*
[2]*Bharat Engineering College, Hyderabad, India*
[3]*Computer Science and Engineering Department, Rector JNT University Hyderabad, India*
* Corresponding author's Email: Krishnashu1818@gmail.com

Abstract: Introduction of the World Wide Web (WWW) and progressions in the field of multimedia and computer technology have enlarged the amount of image collections and databases including art galleries, digital libraries, and medical imageries that extend to millions of images. From these large-scale datasets, the retrieval procedure of images is carried out by classical approaches like chi square distance, colour histogram and text-based image retrieval, which however requires more time for getting the desired images. Thus, there is a need of proposing an effective retrieval system for images by handling these large numbers of images. For this reason, a new Content-Based Image Retrieval (CBIR) system has been implemented in this paper for retrieving the desired images of users from the collected images. Initially, different kinds of images like medical images, texture images, and environmental images etc., are collected from the standard image database. The deep feature extraction using Visual Geometry Group network (VGG-16), Inception v3, and Xception are used to get the features of all the images in the database. The parameters in these three deep learning techniques are tuned with an enhanced optimization algorithm of Modified Bypass-Rider Optimization Algorithm (MB-ROA). The features from VGG16, Inception and Xception are used for optimal weighted fused feature selection using enhanced ROA under training phase. During testing with query images, it undergoes deep feature extraction with the same set of techniques and, then fed into the optimal weighted fused feature selection to get the features of query image. Then, the multi-similarity function considering the cosine, Euclidean and Jaccard similarity are checked between the optimal features of database images and query image. The database images with the minimum similarity with the query image are retrieved from the database. The overall precision and recall to retrieve the top ten images related to the provided query image from the Corel dataset are 77.25% and 76.89%, respectively, while 70.15% and 70.26%, respectively from the VisTex dataset. The experimental findings show that the recommended model is effective at retrieving images from the database.

Keywords: Content based image retrieval, Trio-deep feature extractors with multi-similarity function, VGG16, Inception, Xception, Cosine, Euclidean, Jaccard similarity, Modified bypass-rider optimization algorithm.

1. Introduction

CBIR models demonstrate the procedure of getting images appropriate to a query image from a huge number of visual contents [1]. It has to search the feature representations of image for retrieving a ranked collection of images regarding similarity to the query illustration [2, 3]. A major problem correlated with CBIR is to get the noteworthy information from raw data for eliminating the semantic gap [4]. It shows the variation among the higher-level representations of the images and low-level concepts [5]. The existing research works aim to work on primitive features for showing the image information regarding shape, texture and color. The most of existing works have been described the way of exploring the semantic rich image representations [6]. Some of the works combined bag-of-words models with local descriptors like Scale Invariant Feature Transform (SIFT), Vector of Locally Aggregated Descriptors (VLAD) and the fisher vector descriptors [7].

The two major factors like similarity measurement and feature representation are taken for the design of the CBIR program [7]. Various techniques have been implemented in CBIR research owing to differentiation among the low-resolution image pixels attained by high level human sensing and machines [8]. This issue is taken as the basic problem for intelligence due to the high-level inspection that helps to train and build intelligent machines as humans to perform real-world tasks [9]. The traditional approaches focus on discovering the corresponding pictures by taking the image features, in which some works build mathematical models for calculating the similarity of feature among images [10]. A specific model extracts feature factors including partial, texture, or color regions [11]. Though, the image feature factors do not broadly extracted in the mathematical models. For this reason, Local Binary Pattern (LBP) algorithm is used that relies upon image texture however, it shows bad performance in terms of color matching [12]. In the field of CBIR, various research works have aimed to address the retrieval outcomes, which can completely state high level semantic theories concerning to the feature description with low-level image content. Though, the results of image retrieval methods frequently drop under the expectations of users.

CBIR techniques often use machine learning techniques for solving the aforementioned issues and to reach the appropriate results. In recent decades, various machine learning approaches have observed more advancement [13]. They focus on demonstrating the high-quality data extraction with the aid of deep implementation approaches, which has designed with various offline parameters. Deep learning models learn features by discovering deep design frameworks at different levels of creativity from data. It uses complex functions to directly map input data to output without relying on human-made features [14]. The deep learning approaches have gained tremendous success in recent days specifically for retrieval tasks for images. Although, the various researches have been done with help of deep learning techniques in various image classification and recognition tasks in computer vision, there is not adequate focus on prioritizing the CBIR applications [15]. Deep learning techniques like Convolution Neural Network (CNN) model is trained enhancing the efficacy of feature learning by saving time and for improving the limitations of classical mathematical models [16]. The image retrieval tasks suffer from reducing the semantic gap among similarity analysis for human visual semantics [17]. It is performed by considering the semantically similar patterns among the similar group, and manually deleted irrelevant and noisy information in images. These challenges on CBIR and progressions on deep learning encourage us to implement a new CBIR model.

The highlights of the designed CBIR model are listed here.

- To implement an enhanced CBIR scheme with trio-deep feature extractors for enhancing the efficiency of retrieved images.
- To get the deep features from both gathered input images and also query images through utilizing the deep learning techniques like VGG16, Inception V3 and Xception networks. Rather than using general feature extractors, the deep learning techniques get more unique and high-level features. This trio-deep feature extraction helps in maximization of CBIR performance.
- To propose a new MB-ROA technique for selecting the optimal and fused set of image features under training phase from the gathered deep features and also for choosing the optimal weighted fused feature selection to get the features of query image. This optimal deep feature extraction increases the performance of CBIR in terms of standard performance metrics.
- To retrieve the CBIR images by considering the query images with the help of multi-similarity function concerning with cosine, Euclidean and Jaccard similarity measures.
- To analyze the efficacy of the designed CBIR model among the retrieved and query images with standard evaluation measures by comparing with traditional approaches.

The “remaining sections of the designed model” are given here. Section 2 studies the existing research works. Section 3 suggests the CBIR model using trio-based deep feature extraction. Section 4 proposes a trio-based “deep feature extraction” with optimal weighted fused feature selection. Section 5 recommends the development of novel multi-similarity function for enhanced CBIR model with meta-heuristic algorithm. Section 6 “evaluates the results. Section 7 gives conclusion” to this paper.

2. Existing research works

2.1 Related works

In 2017, Liu et al. [18] have implemented an efficient image retrieval approach through integrating low-level features from “Dot-Diffused Block Truncation Coding (DDBTC)” and the high-level features from CNN model. The high-level features like human perception were efficiently captured by CNN whereas low-level features such as color and texture were built by vector quantization-indexed histogram captured from maximum, minimum quantizers and DDBTC bitmap. The fusion of the CNN and DDBTC features used to improve the overall retrieval rate by similarity reweighting, dimension reduction, and two-layer codebook. Various datasets were examined in this model by determining two measures. As noticed in the evaluation, the designed approaches have accomplished noteworthy efficiency regarding the retrieval rate by comparing with the traditional approaches with either high- or low- level features.

In 2019, Swati *et al.* [19] have suggested a deep VGG19-CNN-derived feature extraction scheme and used closed-form metric learning for measuring the similarity among the database and query images. Moreover, a block-wise fine-tuning scheme and the transfer learning were utilized for enhancing the efficiency of retrieval. The experiments were done on CE-MRI dataset. This dataset has included various classes of brain tumors which were gathered from 233 patients. They have not utilized any handcrafted features, and so, it does not need higher pre-processing. This model has evaluated on fivefold cross-validation and shown the higher average precision compared to traditional methods.

In 2019, Liu *et al.* [20] have implemented an intelligent framework for offering secure CBIR framework to carry out the image retrieval task on the cloud without interacting the user's constant. They have extracted the deep features from the images through pre-trained deep CNN model called VGG-16. The lattice-based homomorphic strategy was used for concealing the information about the network. The recommended system has applied a "secure image similarity scoring protocol" that has facilitated the cloud servers for comparing two images without requiring deep features of any information. The experimental and comprehensive outcomes have demonstrated that the designed scheme was accurate and efficient.

In 2018, Tzelepi and Tefas [21] have adopted a CBIR scheme to learn most effective convolutional illustrations for producing most competent compact descriptors of images. They have enhanced the memory requirements and the retrieval efficiency based on accessible information. This designed approach has used three basic model retraining techniques: relevance feedback-based retraining, fully unsupervised retraining and retraining with relevance information. The efficiency of the designed retrieval model has outperformed others in evaluating the entire datasets.

In 2022, Reena and Ameera [22] have recommended a CBIR model with feature extraction using deep learning technique and feature reduction by a classical learning approach for retrieving the same images from a database for assisting initial or early lymphoma detection. A pre-trained network termed ResNet-101 was used for extracting the images features for distinguishing various classes of cells. This model has used data augmentation for solving the class imbalance issue through over-sampling the training data. Next, they have gathered deep learning features and further, the discriminant features were selected by dimensionality reduction techniques for performing the retrieval of images. The most eminent similarity measure called Euclidean distance was utilized for retrieving the equivalent images from the database. The performance analysis was done on gathered microscopic blood image dataset. The performance analysis has confirmed that the recommended model has verified superior performance regarding effectiveness and practicability.

In 2022, Zhang *et al.* [23] have implemented a new CBIR model through a Convolutional Siamese Neural Network (CSNN). Initially, the lesion patches were taken from the images for composing the NMTB and LC datasets and then, a patch-pair dataset was formed. Secondly, a CSNN was trained by using patch-pair dataset. Thirdly, a query was taken as a test patch. Further, the trained CSNN was utilized for computing the distance among the 20 patches in both datasets and formed query. The majority voting was used for performing the final prediction by taking the patches closest to the query. The designed CBIR-CSNN was trained and tested on a dataset with 719 patients and also on external dataset with 30 patients. The recommended CBIR-CSNN accurately discriminated LC from NMTB through Computerized Tomography (CT) images.

In 2021, Desai *et al.* [24] designed CBIR system to retrieve the desired user images from the gathered images. This framework adopted Support Vector Machine (SVM) and CNN to perform fast image retrieval task. The noteworthy features were extracted using CNN, and the SVM was applied for classification task. At last, the results showed the robustness of the recommended model over others.

In 2021, Thi *et al.* [25] have adopted a novel approach for medical image retrieval tasks through deep learning and salient regions. They have utilized two stages like an online task and offline task, where the local object features from the medical images are done by offline task whereas the CBIR in database was carried out by an online task. Initially, by considering the intensity, texture and shape features were gathered as local object features from the medical images through deep learning used in saliency of decomposition. In the second stage, CBIR with online task was carried out. Most similar images were retrieved by this system along with user given query images as an input through determining the similarity evaluation. The experimentation was done through measures like recall and precision metrics.

2.2 Problem definition

CBIR is considered as the significant problem in the field of computer vision. Recently, different works are developed with the help of labelled images. But, it is cost efficient, less feasible and time consuming while applying it in various applications. Therefore, many CBIR models are implemented for rectifying these challenges, and some of them are depicted in Table 1. CNN [18] ensures the high retrieval rate and minimum matching time of the features. But, it does not produce competitive efficiency while estimating with the textural datasets. Deep CNN VGG19 [19] is highly efficient in computation and memory owing to its lower dimensionality. However, the performance of the model is completely dependent on distance metric learning and feature representation. VGG16 [20] requires less time for searching even when increasing the image count. Yet, it does not evaluate the communication delay and considered only the time cost of the entire server. Deep CNN [21] can be used for improving the retrieval efficiency even for the huge-scale of databases. But, it needs to enhance the image retrieval efficiency with more compact deep features, and it does run on the GPUs. ResNet-101 [22] can also be applicable in the practical mage retrieval applications as it generates high robustness. Yet, this model suffers from higher complexity and slower process in implementation. CBIR-CSNN [23] requires less computational time and provides more robust performance. Still, it requires larger database for training and processing the neural network. CNN and SVM [24] reduce the time consumption of the retrieving the image results. But, it needs large dataset to process and train the neural network. CNN [25] solves the problems related to low-quality of medical images according to the denoising coefficients. But, it does not deeply analyze the context and the correlation in the image among the object features and so, it attains less accuracy in image retrieval. Therefore, a new CBIR model needs to be developed with an improved strategy.

3. Content-based image retrieval model using trio-based deep feature extraction

3.1 Proposed content-based image retrieval model

CBIR is a new research field suggested for addressing the challenges in existing studies as it is performed by considering the visual evaluation of contents by considering the query images. In CBIR model, the problem of achieving similar

Table 1. Benefits and issues of existing CBIR models

Author	Methodology	Features	Challenges
Liu et al. [18]	CNN	It ensures the high retrieval rate and minimum matching time of features.	It does not produce competitive efficiency while estimating with textural datasets.
Swati et al. [19]	Deep CNN VGG19	It is highly efficient in computation and memory owing to its lower dimensionality.	The efficacy of the model is highly based on the distance metric learning and feature representation.
Liu et al. [20]	VGG16	It requires less time for searching even when increasing the image count.	It does not evaluate the communication delay and considered only server time.
Tzelepi and Tefas [21]	Deep CNN	It can be used for improving the retrieval efficiency even for the huge-scale of databases.	It needs to enhance the image retrieval efficiency with more compact deep features and it does run on the GPUs.
Reena and Ameera [22]	ResNet-101	It can also be applicable in the practical image retrieval applications as it generates high robustness.	This model suffers from higher complexity and slower process in implementation.
Zhang et al. [23]	CBIR-CSNN	Provides more robust performance and requires less computational time.	It requires larger database for training and processing the neural network.
Desai et al. [24]	CNN and SVM	It reduces the time consumption of the retrieving the image results.	It is very expensive for training owing to the usage of complicated data.
Thi et al. [25]	CNN	It solves the problems related to lower quality of images according to the denoising coefficients.	It does not deeply analyze context and correlation in an image among the object features and so, it attains less accuracy in image retrieval.

images with superior accuracy is now challenging owing to the large amount of image databases. The traditional searching of images is performed with the help of text queries. Thus, data annotations are accessed by evaluating the query by text, which requires further analysis. On the contrary, the image query-based retrieval system needs to retrieve the most similar images with precise annotations for all the images. Moreover, it is complicated to identify the annotations for applying on a specific image since it requires to reach higher accuracy of the search. CBIR system is more effective as the query by image processes by themselves as a more helpful option. It is observed that the CBIR model is utilized for analyzing the queries with the help of image processing techniques from different research works. It helps in searching the suitable images from the database, however, it is considered as the challenging problem due to the process of large scale images. The basic requirement of any image retrieval approach is to arrange and search the images by considering the "visual semantic relationship with the user query". Some of the search engines on the Internet achieves the images through text-based approaches, which assumes the input as text. Moreover, they are however unfeasible for applying the concept of manual labelling to conventional huge scale image archives includes millions of images. The automated image annotation techniques are dependent on how precisely detects the images regarding factors like shape, spatial layout, texture, edges, and color-related information. The CBIR outputs images that are matching with contents of query images by finding the visual similarity or closeness. Moreover, this proposed CBIR model uses MB-ROA which is a modification of ROA [26], focuses on retrieving the images related to query image with the help of multi-similarity function, where the diagrammatic illustration of the designed CBIR is given in Fig. 1.

This paper implements a new CBIR model by trio-based deep feature extraction with multi-similarity function. The designed model consists of various phases like (i) Data collection, (ii) Deep feature extraction, (iii) Optimal weighted fused feature selection, (iv) multi-similarity computation, and (v) Retrieval of images. Initially, the query images and database images are gathered from two different datasets including various sectors like animal, beaches, etc. The gathered images including both query images and database images are further fed to the deep feature extraction process, where the deep features from the collected images are done through three different techniques like VGG16, Inception V3 and Xception networks. The gathered

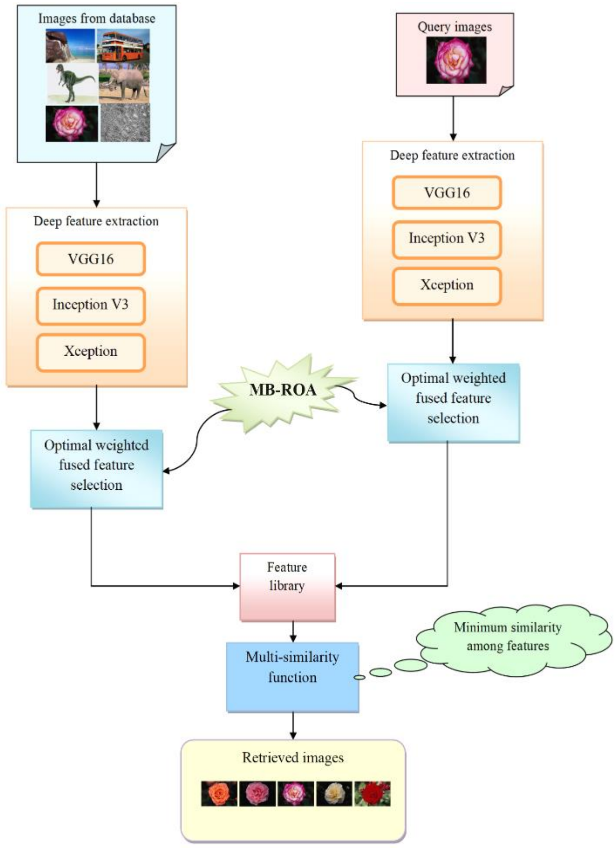


Figure. 1 Content-based image retrieval model using trio-based deep features with multi-similarity function

deep features are fused and fed to the newly developed MB-ROA technique, which results in selecting the optimal features of both query images and database images. Further, the optimal features of query images and database images are multiplied with two optimal weights, where the weights are optimized with the help of MB-ROA algorithm. Finally, the optimal weighted fused features are attained and forwarded to the feature library during training process. Then, while performing the testing process, when query image is given, the optimal features of query image should be matched with the feature library for retrieving the images. The multi-similarity function is computed among the features of query image with database feature library regarding metrics like cosine, Euclidean and Jaccard similarity functions. At last, the minimum similarity among the database images with the query image is retrieved from the database. The superior retrieval outcomes are attained by evaluating the features of both database images and query images.

3.2 Description of CBIR dataset

The proposed CBIR model gathers the input and query images from two standard datasets for experimentation, which has been discussed here.

Received: July 26, 2022. Revised: September 22, 2022.

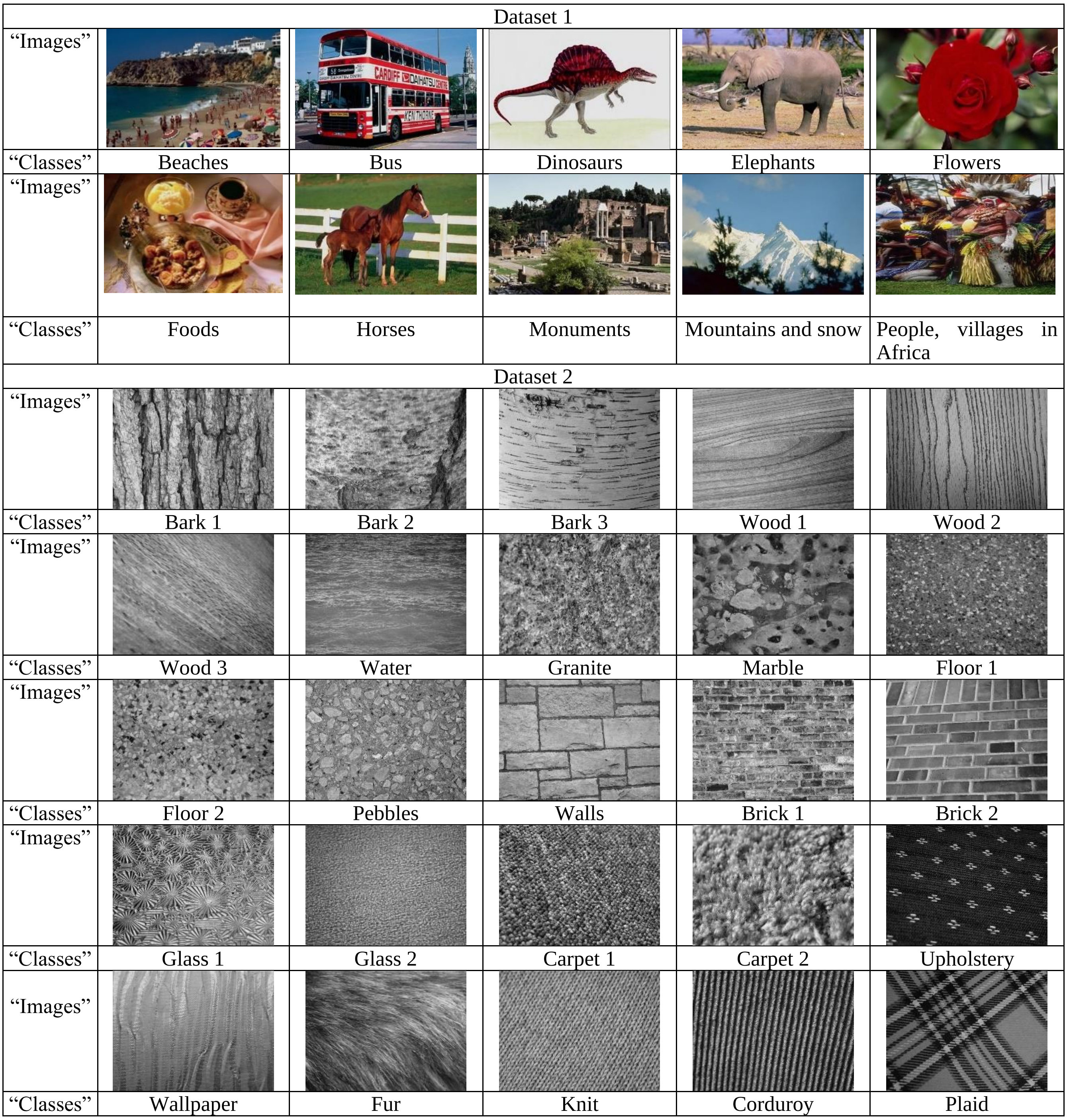


Figure. 2 Sample images gathered from two datasets for experimenting the content-based image retrieval model

Dataset-1 Corel Images: It is collected from "https://www.kaggle.com/datasets/elkamel/corel-images: Access date: 25-05-2022". It consists of 10 concept groups of images composed each by 100 images. For every concept group, the images are classified into 90 images for training and 10 images for test. The 10 groups in the dataset are animal, beaches, monuments, mountains and snow, foods, flowers, elephants, dinosaurs, bus, people and villages in Africa.

Dataset 2 VisTex database: It is collected from Link : http://vismod.media.mit.edu/pub/VisTex/". It includes various classes, which include bark with 3 scenes, wood with 3 scenes, water, granite, marble, floor with 2 scenes, pebbles, wall, brick with 2 scenes, glass with 2 scenes, carpet with 2 scenes, upholstery, fur, knit, corduroy, and plaid.

Finally, the gathered database images are termed as H_x, where $x = 1,2,3,\cdots,X$ and the total number of gathered images is known as X and the gathered

Received: July 26, 2022. Revised: September 22, 2022.

query images are termed as Q_x, where $x = 1,2,3,\cdots,X$. Few collected query images are shown in Fig. 2.

4. Trio-based deep feature extraction with optimal weighted fused feature selection

4.1 Trio-based deep feature extraction

In the designed CBIR model, a new trio-based deep feature extraction is used for gathering the most significant features for enhancing the accuracy of CBIR model. For this purpose, the VGG-16, Inception v3, and Xception networks are used for gathering deep features. These networks get the features from the gathered database images H_x, and also from the query images are termed as Q_x.

VGG-16: It is most efficient network used for extracting the deep features of the images. Moreover, the existing works report high accuracy through the VGG-16 [27]. Generally, VGG-16 is useful in precise feature extractor. The VGG-16 considers the input images as database images and query images with fixed size of 3×224×224. These input images are forwarded to a stack of different convolutional layers of various receptive fields. The stride rate for pooling and convolutional layers is similar throughout the VGG-16 network, where the 2×2 with stride 2 is included in pooling layer and 3×3 with stride 1 is presented in convolutional layer. VGG-16 has two different convolutional layers with 64 and 128 filters, correspondingly. Before every convolution process, the border pixels are included for preserving the sizes of feature maps similar to the input. Finally, VGG-16 has three fully connected layers with 4096 neurons and the features are compressed into 1000 dimensions through final fully connected layer. The features extracted from VGG-16 for both database images and query images are correspondingly known as F_S^{VGG-16} and V_S^{VGG-16}.

Inception v3: Inception v3 is one of the networks of CNNs, which is especially utilized for extracting the features from the both query images and database images [28]. It has the benefit of factoring convolutions into various branches successively operating on space and channels. The Inception-V3 model is an updated version of the Inception-V1 model. It utilizes various numbers of techniques for optimizing the network. The major concept of the Inception framework is to alter small kernels with the larger kernel for learning the multi-scale representations and reduced the total number of constraints and computational complexity [29]. Finally, the features extracted from Inception v3 for both database images and query images are correspondingly known as F_S^{incept} and V_S^{incept}.

Xception: It is an extended version of Inception architecture. It is "a linear stack of depthwise separable convolution layers with residual connections" [30]. These layers help in reducing the memory requirements and computational cost. Xception consists of 36 convolutional layers divided into 14 modules, in which "all include linear residual connections, except for the first and last modules". The learning of space-wise and channel-wise features is performed by separating the separable convolution in Xception. The representational bottlenecks and vanishing gradients are solved by adopting the residual connection through creating a "shortcut in the sequential network". A summation operation is used for getting the final outputs through the shortcut connection. At last, the features extracted from Xception for both database images and query images are correspondingly known as F_S^{xcept} and V_S^{xcept}. These gathered "deep features" are fed to the optimal weighted fused feature selection stage.

4.2 Optimal weighted fused feature selection

The recommended CBIR model introduces a novel optimal weighted fused "feature selection process", where the optimal features are selected with the help of MB-ROA technique. Here, from the gathered set of VGG-16 features F_S^{VGG-16} and V_S^{VGG-16}, Inception v3 features F_S^{incept} and V_S^{incept} and Xception features F_S^{xcept} and V_S^{xcept}, 10 features are selected for each feature set through MB-ROA technique. This optimal feature selection helps in choosing the most significant features with the help of heuristic strategy. It aids in reducing the training time, enhancing the accuracy, eliminating the irrelevant or redundant features and minimizing the overfitting. The optimally chosen features of database images using MB-ROA technique are known as $F_S^{VGG-16*}$, $F_S^{incept*}$ and F_S^{xcept*}, and the optimal features of query images using MB-ROA technique are termed as $V_S^{VGG-16*}$, $V_S^{incept*}$ and V_S^{xcept*} that are forwarded to the weighted feature selection process, where the optimal weights known as W_1 and W_2 are multiplied with the optimally selected features for getting the optimal weighted fused features. Here, the MB-ROA technique helps in optimizing the weights W_1 and W_2 to get the optimal weighted fused feature selection process. This procedure for database images is derived in following equations.

$$WF_1 = (F_s^{VGG-16*} \times W_1) + \left(F_s^{incepr*} \times (1 - W_1)\right) \quad (1)$$

$$F_s^* = (W_2 \times WF_1) + \left((1 - W_2) \times F_s^{xcept*}\right) \quad (2)$$

Finally, the selected optimal weighted fused features of database images are known as F_s^*, in which the selected optimal weighted fused features of database images are fed to the feature library. Similarly, optimal weighted fused feature selection process of query image is derived here.

$$WS_1 = (V_s^{VGG-16*} \times Wg_1) + \left(V_s^{incepr*} \times (1 - Wg_1)\right) \quad (3)$$

$$V_s^* = (Wg_2 \times WS_1) + \left((1 - Wg_2) \times V_s^{xcept*}\right) \quad (4)$$

Then, the selected optimal weighted fused features of query image are termed as V_s^*, which is attained during the testing process while getting the query image. These selected optimal weighted fused features of query image during the testing process are further subjected to the feature library to retrieve the output images regarding the query image by comparing with those features by finding the similarity. Here, the range of the selected optimal weighted fused features F_s^* and V_s^* are correspondingly assigned in the bounding limit of [0, 10], and the weights W_1 , W_2 , Wg_1 and Wg_2 are correspondingly fixed among [0.01, 0.99]. The objective of minimizing the variance among optimal weighted fused features of database images is given in Eq. (5).

$$of^{database} = \underset{\{F_s^{VGG-16*}, F_s^{incept*}, F_s^{xcept*}, W_1, W_2\}}{argmin} \left(\frac{1}{var(F_s^*)}\right) \quad (5)$$

The objective of minimizing the variance among optimal weighted fused features of query image is given in Eq. (6).

$$of^{query} = \underset{\{V_s^{VGG-16*}, V_s^{incept*}, V_s^{xcept*}, Wg_1, Wg_2\}}{argmin} \left(\frac{1}{var(V_s^*)}\right) \quad (6)$$

The illustration of optimal weighted fused feature selection stage is known in Fig. 3.

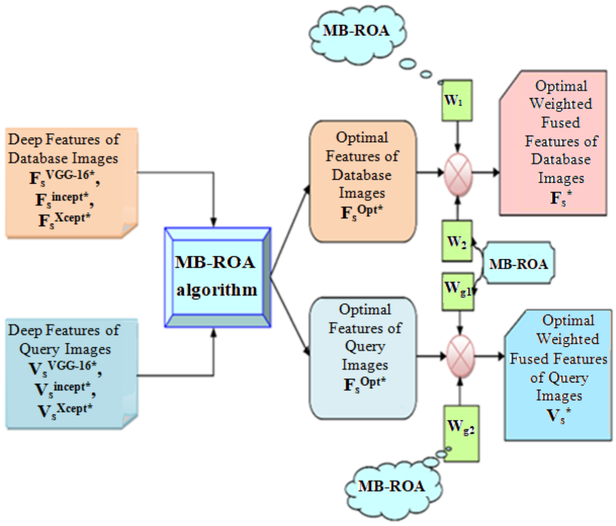


Figure. 3 Optimal weighted fused feature selection stage using MB-ROA technique

5. Development of novel multi-similarity function for enhanced content-based image retrieval model with meta-heuristic algorithm

5.1 Proposed MB-ROA

In this designed CBIR model, a new MB-ROA technique is implemented for selecting the optimal features of both database images and query image and optimizing the weights used for selecting the optimal weighted fused feature selection. This MB-ROA-derived feature selection helps in maximizing the performance of the retrieval rate of images. This new MB-ROA method is implemented for getting the superior performance rate of retrieved images. It is implemented by inspiring the concept of ROA.

It is a novel optimization algorithm utilized for finding the target position by various components such as "bypass rider, follower, overtake, and attackers' parameters." The ROA is implemented by considering the riders proceeded to the target positions for winning the race. Each rider follows some strategies for reaching the target location. Initially, through the help of bypassing leading riders, the target position is reached by the bypass rider. The leading rider will be followed by the follower, and they won't disobey the leading rider. The over-taker follows their path for reaching the target position by considering the route of leading rider. Finally, the attackers reach the target location by following the maximum speed theory. Thus, all the riders utilize the predefined theory for reaching the target location.

As a modification to the existing ROA, the bypass rider reaches the final location by using the fitness assisted random number instead of randomly generated number in the range of [0, 1]. The bypass rider reaches the target by taking the leading path, where the group riders are initialized in Eq. (7).

$$Y_k = \{Y_k(j,l)\}, 1 \le j \le D; 1 \le l \le A \qquad (7)$$

In Eq. (7), the location of j^{th}rider at time $k$is mentioned as Y_k, the time instant is noted as k, the dimension of the optimization issue is given as A, and the number of riders is specified as D.The rider who is in the leading position is considered as the bypass rider, which is determined in Eq. (8).

$$Y_{k+1}^{BP} = \xi\left[Y_k(\beta,l) * \delta(l) + Y_k(\alpha,l) * [1 - \delta(l)]\right] \qquad (8)$$

Here, term δ indicates a random value with the size of $1\times A$ in the range of [0, 1], a number selected from 1 to D is denoted as α, a random number with $[1, D]$ is indicated as β and a random number with [0, 1] in the conventional ROA is mentioned as ξ whereas in the designed MB-ROA technique, it is updated based on the fitness solutions as per Eq. (9).

$$\xi = \frac{of_k}{of_k^{worst} * of_k^{mean}} \qquad (9)$$

In Eq. (9), the mean among fitness solutions is noted as of_k^{mean}, the worst fitness solution is indicated as of_k^{worst} and the recent fitness solution is denoted as of_k.

Secondly, the follower moves to the destination through following the bypass rider, which is more relevant to bypass rider and so, the follower position is updated as given in Eq. (10).

$$Y_{k+1}^{FP}(l,m) = Y^{BL}(BL,l) + \left[cos\left(E_{j,m}^k\right) * Y^{BL}(BL,l) * di_j^k\right] \qquad (10)$$

In Eq. (10), the distance to be travelled of the j^{th}rider is known as di_j^k, the steering angle in the m^{th} coordinate of the j^{th} rider is mentioned as $E_{j,m}^k$, the index of bypass rider is represented as BL, position of bypass rider is noted as Y^{BL}, and the coordinate selector is referred to as m. The overtaker rider follows the information collected based on the bypass rider along with their own position. The over-taker position is updated by considering three constraints like direction indicator, relative success rate, and coordinator selector. These factors play a major role due to the requirement of collecting the own information of over-taker and bypass rider. Thus, the position of over-taker is updated in Eq. (11).

$$Y_{k+1}^{OP}(l,m) = Y_k(l,m) + \left[Di_k^R(l) * Y^{BL}(BL,m)\right] \qquad (11)$$

In Eq. (11), the direction indicator of the j^{th} rider is known as $Di_k^R(l)$ and derived in Eq. (12), and the location in the m^{th} coordinate of thej^{th}rider is expressed as $Y_k(l,m)$, where the success rate of j^{th} rider is given as $GF_k^D(l)$.

$$Di_k^R(l) = \left[\frac{2}{1 - log\left(GF_k^D(l)\right)}\right] - 1 \qquad (12)$$

For reaching target location, the attacker utilizes maximum speed strategy, in which the attacker focuses on obtaining the location of leader and follower. Eq. (13) determines the location of the attacker.

$$Y_{k+1}^{AT}(j,l) = Y^{BL}(BL,l) + \left[cos\left(E_{j,m}^k\right) * Y^{BL}(BL,l) * di_j^k\right] \qquad (13)$$

The leader position is noted as $Y^{BL}(BL,l)$. The process is repeated until reaching the optimal solutions and meeting termination criteria. The flowchart of the designed MB-ROA technique is given in Algorithm 1.

Algorithm 1: Suggested MB-ROA technique

Assume the rider population and parameters like steering angle
Evaluate the fitness among solutions
Determine the random number $\xi$by Eq. (9)
While $k < K_{max}$
  For $j = 1 to D$
    "Update the bypass rider using" Eq. (8)
    "Update the follower rider using" Eq. (10)
    "Update the overtaker using" Eq. (11)
    "Update the attacker using" Eq. (13)
  End for
  Update the best solutions
Return best solutions
End while
Exit

5.2 Multi-similarity function for image retrieval

The selected optimal weighted fused features of query image V_s^* is compared with the feature library of database images F_s^* with multi-similarity function for getting the images based on the query image. In the research field of CBIR, similarity determination with standard measures is played a significant role, which helps in calculating the visual similarities among the images in a database and a query image. Hence, the retrieval outcome is not a single image but a number of images arranged through their similarities with the query image. Utilization of various similarity measures result in reducing the retrieval efficient in any image retrieval systems and thus, there is a need of finding the best distance measure for CBIR system. While getting the smaller value for the distance, the query images will be more equivalent to the database images. The determination of objective with three distance measures like cosine, Euclidean and Jaccard similarity metrics are used for measuring the similarity among the features of database images and query image.

Here, cosine similarity cs "measures the similarity between two vectors of an inner product space and it is measured by the cosine of the angle between two vectors and determines whether two vectors are pointing in roughly the same direction" as equated in Eq. (14).

$$cs(F_S^*, V_S^*) = \frac{F_S^* \cdot V_S^*}{\|F_S^*\| \cdot \|V_S^*\|} \tag{14}$$

Then, the Euclidean distance is described as "the distance between two feature vectors" as given in Eq. (15).

$$ds(F_S^*, V_S^*) = \sqrt{\sum (F_S^* \cdot V_S^*)^2} \tag{15}$$

Further, Jaccard similarity is described as "the size of the intersection divided by the size of the union of two lists" and also it is a common statistical metric that determines "the similarity between two feature sets based on their intersection" and is formulated in Eq. (16).

$$Jac(F_S^*, V_S^*) = \frac{|F_S^* \cap V_S^*|}{|F_S^* \cup V_S^*|} \tag{16}$$

Thus, this multi-similarity function among the database images and query image result in better retrieval images. Finally, the suitable images related to query image are achieved with the help of trio-based deep feature extraction and multi-similarity determination.

6. Results and discussions

6.1 Experimental setup

The proposed multi-similarity function-based CBIR model was developed in Python. Here, the performance of the proposed model was compared over the conventional models like VGG-16 [27], Inception v3 [28], and Xception [30] and heuristic approaches like PSO [31], GWO [32], DHOA [33], and ROA [26] in terms of F1-score, Precision and Recall. The evaluation was carried out on "two datasets" and analyzed the performance of retrieved images with various performance measures.

6.2 Performance measures

The evaluation metrics utilized for experimenting the effectiveness of the retrieved images are illustrated here.

6.2.1. Precision:

It is "the ratio of positive observations that are predicted exactly to the total number of observations that are positively predicted".

$$Pr = \frac{cr^{pos}}{cr^{pos} + sa^{pos}} \tag{17}$$

Here, terms cr^{pos}, cr^{neg}, sa^{pos}, sa^{neg} refer to the "true positives, true negatives, false positives, and false negatives," respectively.

6.2.2. Recall:

"the number of true positives, which are recognized exactly".

$$recall = \frac{2cr^{pos}}{cr^{pos} + sa^{neg}} \tag{18}$$

6.2.3. F1-score:

It is the harmonic mean between precision and recall. It is used as a statistical measure to rate performance".

$$F1score = \frac{2cr^{pos}}{2cr^{pos} + sa^{pos} + sa^{neg}} \tag{19}$$

6.3 Retrieved images with query image

Some of the retrieved images with query image are given in Fig. 4, in which the proposed MB-ROA achieves most relevant images based on the query images whereas the existing algorithms tries to match the query image but it fails.

6.4 Performance analysis over heuristic algorithms

When retrieving the number of top similar images associated to the query image is goes on increasing, the outcomes of optimization techniques such as PSO [31], GWO [32], DHOA [33], and ROA [26] are practically comparable. Furthermore, the high computing cost and precision of these

Received: July 26, 2022. Revised: September 22, 2022.

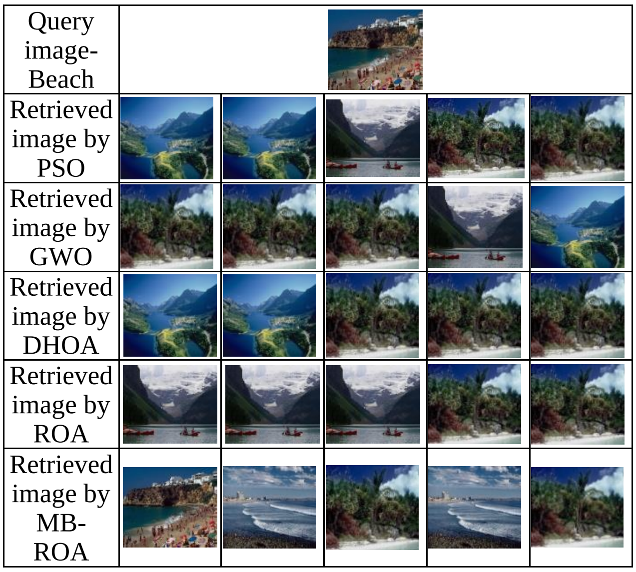


Figure. 4 Some retrieved images with query image over various heuristic algorithms on corel images dataset

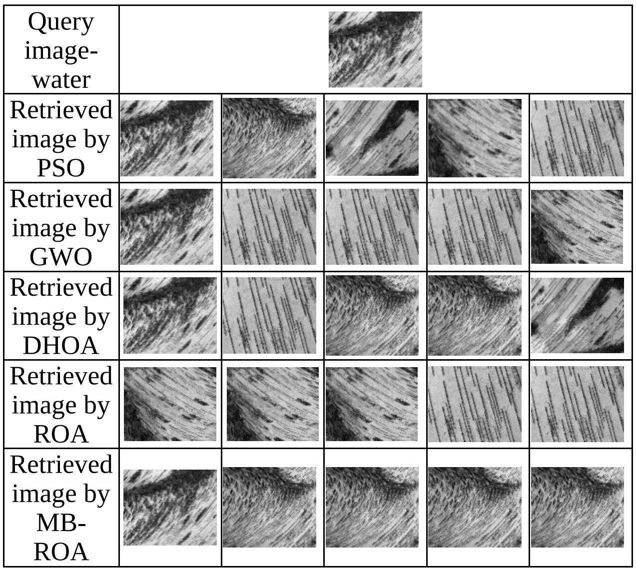


Figure. 5 Some of the retrieved images with query image over various heuristic algorithms on VisTex database

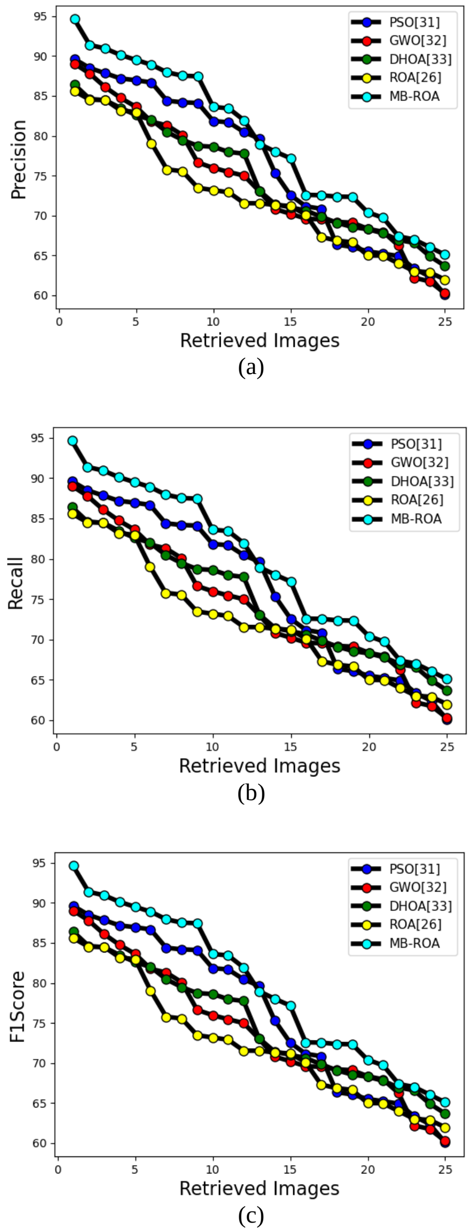


Figure. 6 Evaluation on the implemented enhanced CBIR with trio-deep feature extractors with multi-similarity function over optimization approaches "for dataset 2 in terms of (a) Precision, (b) Recall, and (c) F1-score"

optimization techniques remain a challenging task. Among these meta-heuristic optimization techniques, ROA achieves global optimal solutions with a limited number of rider groups. As a result, its computational cost has been decreased. However, in order to boost the convergence rate even further, ROA is modified and renamed as MB-ROA. The MB-ROA bypass most rider groups that are not required to deliver the optimal solution. As a result, MB-ROA outperformed the competition.

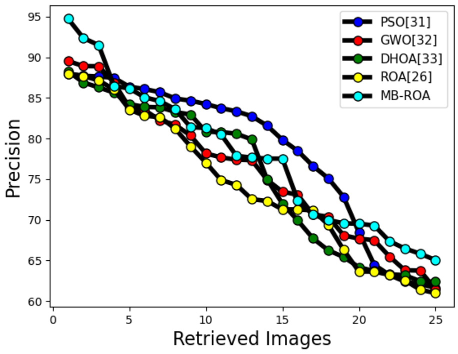
The investigation of the effectiveness of the designed CBIR mode over other optimization algorithms for both datasets is depicted in Fig. 6 and Fig. 7 respectively. The considered measures are evaluated regarding various retrieved images, which demonstrate the superior efficiency using MB-ROA.

6.5 Investigation on deep learning techniques

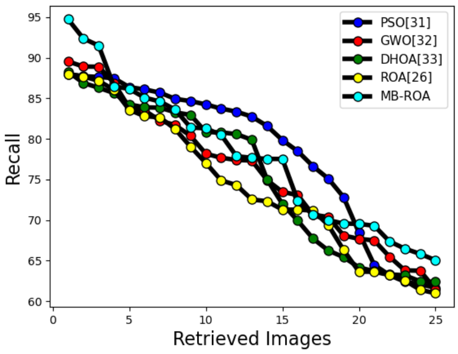
The evaluation on the recommended MB-ROA based CBIR model is examined over two datasets with various deep learning techniques as correspondingly shown in Fig. 8 and 9. The analysis on proposed MB-ROA approach gets maximum retrieval performance with trio-based deep feature extraction over deep learning techniques.

The works in [18-25] used deep learning models to extract the features. As a result, extra pre-processing is no longer required. However, the usage of a larger feature set limits the usefulness of

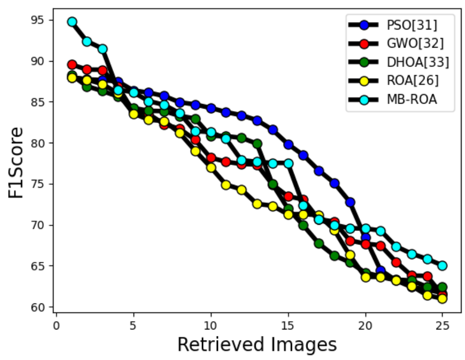
Received: July 26, 2022. Revised: September 22, 2022.



(a)

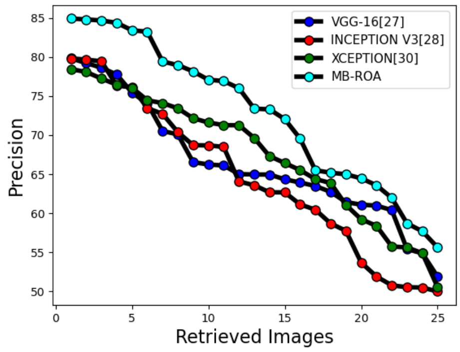


(b)

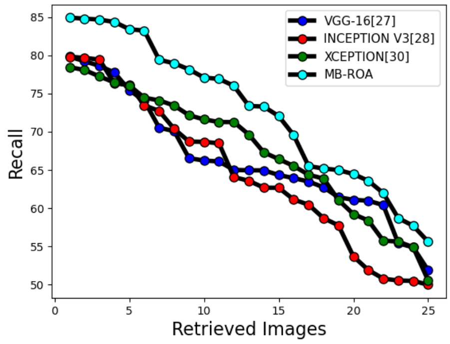


(c)

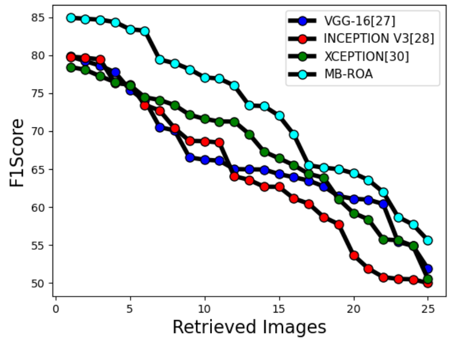
Figure. 7 Evaluation on the implemented enhanced CBIR with trio-deep feature extractors with multi-similarity function over optimization approaches “for dataset 2 in terms of (a) Precision, (b) Recall, and (c) F1-score”



(a)



(b)



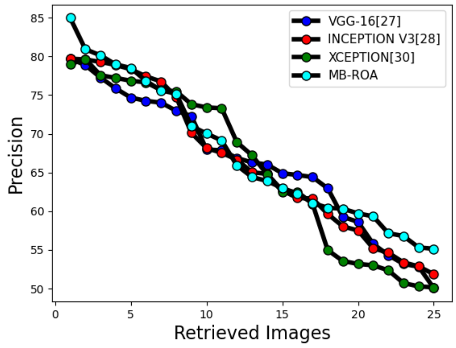
(c)

Figure. 8 Evaluation on the implemented enhanced CBIR with trio-deep feature extractors with multi-similarity function over deep learning approaches “for dataset 1 in terms of (a) Precision, (b) Recall, and (c) F1-score”

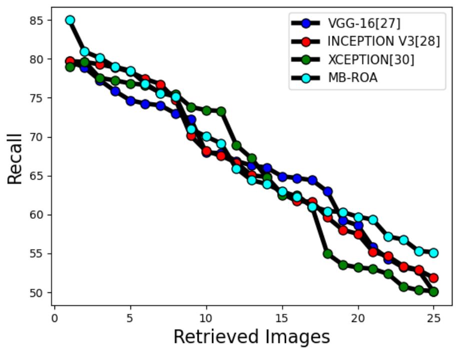
the works in [18-20, 23-25]. The works in [20-24] suffers due to complexity of extracted features and needs longer training time and a large database. These challenges can be addressed by adopting optimum feature set.

The VGG-16 [27] model improves accuracy mostly through increasing model depth. Inception [28] is capable of extracting multi-level features. Xception [30] substitutes regular Inception modules with depth-wise separable convolutions to further improve accuracy. As each of the above three models has its own strengths for improving retrieval performance, their feature sets are merged and produced a trio-based feature set to further enhance retrieval performance. However, as a result of this integration, the number of features in the trio-based feature set increased. This had decreased the efficiency and increased computational cost. To address this, MB-ROA optimization was done on trio-based features to get optimal features, resulting in lower computational costs and higher performance.

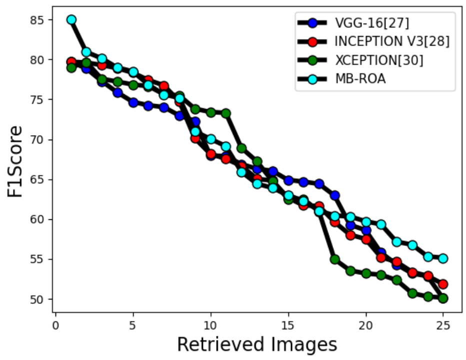
The recommended CBIR using MB-ROA in terms of precision is 5.5%, 4.8%, and 3.6% correspondingly superior to VGG-16, Inception v3



(a)



(b)



(c)

Figure. 9 Evaluation on the implemented enhanced CBIR with trio-deep feature extractors with multi-similarity function over deep learning approaches "for dataset 2 in terms of (a) Precision, (b) Recall, and (c) F1-score"

and Xception techniques while taking the number of retrieved images as 25 for dataset 1. Similarly, the maximum retrieval performance using the designed model is also evaluated for dataset 2 also.

7. Conclusion

Use A novel CBIR model was implemented in this paper for achieving better retrieval results. Here, different kinds of images from standard database were collected. These gathered images from the database were utilized in the deep feature extraction that was done using VGG-16, Inception v3, and Xception networks for extracting the features of all the images in the database including query image. Here, a new MB-ROA technique was implemented for optimizing the weight and selecting the features to implement optimal weighted fused feature selection. These extracted features of both database images and query image was stored into separate feature library. Then, the multi-similarity function was determined among the optimal features of query image and database images by considering the cosine, Euclidean and Jaccard similarity measures. The database images with the minimum similarity with the query image were attained from the database. Through the performance analysis, the efficiency of the trio-deep feature extraction using MB-ROA technique in terms of F1-score was 5%, 3.5%, and 4.9% correspondingly superior to VGG-16, Inception v3 and Xception techniques for VisTex database by "considering the number retrieved images" as 25. Finally, the analysis has demonstrated that the efficiency of the designed model on retrieving the images from the database.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data collection, 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 and 3rd author.

References

[1] S. Sathiamoorthy and M. Natarajan, "Correction to: An efficient content based image retrieval using enhanced multi-trend structure descriptor", *SN Applied Sciences*, Vol. 2, No. 335, 2020.

[2] R. Zhang and Z. Zhang, "Effective Image Retrieval Based on Hidden Concept Discovery in Image Database", *IEEE Transactions on Image Processing*, Vol. 16, No. 2, pp. 562-572, 2007.

[3] H. Greenspan and A. T. Pinhas, "Medical Image Categorization and Retrieval for PA", *IEEE Transactions on Information Technology in Biomedicine*, Vol. 11, No. 2, pp. 190-202, 2007.

[4] C. S. Fuh, S. W. Cho, and K. Essig, "Hierarchical color image region segmentation

for content-based image retrieval system", *IEEE Transactions on Image Processing*, Vol. 9, No. 1, pp. 156-162, 2000.

[5] D. Xu, S. Yan, D. Tao, S. Lin, and H. Zhang, "Marginal Fisher Analysis and Its Variants for Human Gait Recognition and Content- Based Image Retrieval", *IEEE Transactions on Image Processing*, Vol. 16, No. 11, pp. 2811-2821, 2007.

[6] J. Kim, W. Cai, D. Feng, and H. Wu, "A New Way for Multidimensional Medical Data Management: Volume of Interest (VOI)-Based Retrieval of Medical Images with Visual and Functional Features", *IEEE Transactions on Information Technology in Biomedicine*, Vol. 10, No. 3, pp. 598-607, 2006.

[7] G. Paschos, I. Radev, and N. Prabakar, "Image content-based retrieval using chromaticity moments", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 15, No. 5, pp. 1069-1072, 2003.

[8] B. Zhu, M. Ramsey, and H. Chen, "Creating a large-scale content-based airphoto image digital library", *IEEE Transactions on Image Processing*, Vol. 9, No. 1, pp. 163-167, 2000.

[9] R. Rahmani, S. A. Goldman, H. Zhang, S. R. Cholleti, and J. E. Fritts, "Localized Content-Based Image Retrieval", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 11, pp. 1902-1912, 2008.

[10] J. Zhang and L. Ye, "Content Based Image Retrieval Using Unclean Positive Examples", *IEEE Transactions on Image Processing*, Vol. 18, No. 10, pp. 2370-2375, 2009.

[11] D. P. Huijsmans and N. Sebe, "How to complete performance graphs in content-based image retrieval: add generality and normalize scope", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 2, pp. 245-251, 2005.

[12] I. E. Naqa, Y. Yang, N. P. Galatsanos, R. M. Nishikawa, and M. N. Wernick, "A similarity learning approach to content-based image retrieval: application to digital mammography", *IEEE Transactions on Medical Imaging*, Vol. 23, No. 10, pp. 1233-1244, 2004.

[13] M. M. Rahman, P. Bhattacharya, and B. C. Desai, "A Framework for Medical Image Retrieval Using Machine Learning and Statistical Similarity Matching Techniques with Relevance Feedback", *IEEE Transactions on Information Technology in Biomedicine*, Vol. 11, No. 1, pp. 58-69, 2007.

[14] R. Mogili and G. Narsimha, "Detection of Cardiac Arrhythmia from ECG Using CNN and XGBoost", *International Journal of Intelligent Engineering and Systems*, Vol. 15, No. 2, pp. 414-425, 2022, doi: 10.22266/ijies2022.0430.38.

[15] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 12, pp. 1349-1380, 2000.

[16] W. Jiang, G. Er, Q. Dai, and J. Gu, "Similarity-based online feature selection in content-based image retrieval", *IEEE Transactions on Image Processing*, Vol. 15, No. 3, pp. 702-712, 2006.

[17] R. Krishnapuram, S. Medasani, S. H. Jung, Y. S. Choi, and R. Balasubramaniam, "Content-based image retrieval based on a fuzzy approach", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 16, No. 10, pp. 1185-1199, 2004.

[18] P. Liu, J. Guo, C. Wu, and D. Cai, "Fusion of Deep Learning and Compressed Domain Features for Content-Based Image Retrieval", *IEEE Transactions on Image Processing*, Vol. 26, No. 12, pp. 5706-5717, 2017.

[19] Z. N. K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, J. Lu, "Content-Based Brain Tumor Retrieval for MR Images Using Transfer Learning", *IEEE Access*, Vol. 7, pp. 17809-17822, 2019.

[20] F. Liu, Y. Wang, F. C. Wang, Y. Z. Zhang, and J. Lin, "Intelligent and Secure Content-Based Image Retrieval for Mobile Users", *IEEE Access*, Vol. 7, pp. 119209-119222, 2019.

[21] M. Tzelepi and A. Tefas, "Deep convolutional learning for Content Based Image Retrieval", *Neurocomputing*, Vol. 275, pp. 2467-2478, 2018.

[22] M. R. Reena and P. M. Ameer, "A content-based image retrieval system for the diagnosis of lymphoma using blood micrographs: An incorporation of deep learning with a traditional learning approach", *Computers in Biology and Medicine*, Vol. 145, No. 105463, 2022.

[23] K. Zhang, S. Qi, J. Cai, D. Zhao, T. Yu, Y. Yue, Y. Yao, and W. Qian, "Content-based image retrieval with a Convolutional Siamese Neural Network: Distinguishing lung cancer and tuberculosis in CT images", *Computers in Biology and Medicine*, Vol. 140, No. 105096, 2022.

[24] P. Desai, J. Pujari, C. Sujatha, A. Kamble, and A. Kambli, "Hybrid Approach for Content-Based Image Retrieval using VGG16 Layered Architecture and SVM: An Application of

Deep Learning”, *SN Computer Science*, Vol. 2, No. 170, 2021.

[25] V. T. H. Tuyet, N. T. Binh, N. K. Quoc, and A. Khare, “Content Based Medical Image Retrieval Based on Salient Regions Combined with Deep Learning”, *Mobile Networks and Applications*, Vol. 26, pp. 1300-1310, 2021.

[26] D. Binu and B. S. Kariyappa, “RideNN: A New Rider Optimization Algorithm-Based Neural Network for Fault Diagnosis in Analog Circuits”, *IEEE Transactions on Instrumentation and Measurement*, 2018.

[27] K. Tejasri and J. R. Singothu, “Secure Content Based Image Retrieval System Using CNN and VGG-16”, *JETIR*, Vol. 7, Issue. 11, 2020.

[28] V. Bhandi, K. A. S. Devi, “Image Retrieval Using Features from Pre-Trained Deep CNN”, *International Journal of Scientific & Technology Research*, 2020.

[29] B. Sreenivasulu, A. Pasala, G. Vasanth, “Adaptive Inception Based on Transfer Learning for Effective Visual Recognition”, *International Journal of Intelligent Engineering and Systems*, Vol. 13, No. 6, pp. 1-10, 2020, doi: 10.22266/ijies2020.1231.01.

[30] G. Pereira, S. Härtel, M. Solar, and M. Araya, “Content-Based Medical Image Retrieval and Intelligent Interactive Visual Browser for Medical Education, Research and Care”, *Diagnostics*, Vol. 11, 2021.

[31] M. E. H. Pedersen and A. J. Chipperfield, “Simplifying Particle Swarm Optimization”, *Applied Soft Computing*, Vol. 10, No. 2, pp. 618-628, 2010.

[32] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey Wolf Optimizer”, *Advances in Engineering Software*, Vol. 69, pp. 46-61, 2014.

[33] G. Brammya, S. Praveena, N. S. N. Preetha, R. Ramya, B. R. Rajakumar, and D. Binu, “Deer Hunting Optimization Algorithm: A New Nature-Inspired Meta-heuristic Paradigm”, 2019.