EAI Endorsed Transactions
on Internet of Things
Research Article EAI.EU

Revitalizing Image Retrieval: AI Enhancement and Metaheuristic Algorithm Adaptation

Kumaravel Pichaimani[1,*], S.Thabasu Kannan[2]

[1]Periyar University, Salem, India
[2]Pannai College of Engineeering and Technology, Sivagangai, India

Abstract

INTRODUCTION: In recent years, development of digital technology has led to number of images, which can be stored in digital format. However, searching and retrieval of images in large image DB (Database) is a mammoth task. Therefore, different image retrieval techniques have been used for retrieving the suitable images, which includes retrieval of images using keywords or annotations, however, these methods are considered to be time consuming and leads to imprecise outcome.
OBJECTIVES: Therefore Effective and precise retrieval of suitable images from huge DB can thrived by utilizing CBIR (Content Based Image Retrieval) system. However, incorporation of CBIR in most existing studies resulted in low accuracy for IR. So, proposed model incorporates Modified ResNet50 (M-ResNet50) and VGG 16 model for feature extraction in order to extract the best features as M-ResNet50 utilizes extra dense layers which aids in better feature extraction process.
METHODS: After feature extraction, the features are fused using PCA and fed to Modified PSO (M-PSO) model for obtaining optimized features since M-PSO is fast and aids in selecting optimal features after processing from insignificant features that primarily set the preferred number of necessary features.
RESULTS: Moreover, M-PSO require less parameters to tune instead of a huge number of parameters by incorporating K parameters of KNN algorithm in order to find the nearest images to Query Images (QI), thereby making the model appropriate for IR process with better similarity score.
CONCLUSION: The proposed model utilizes 8 different sun images at different intervals for IR process. Finally, the proposed model is evaluated by using several metrics such as accuracy, precision, recall and F1 score, besides the proposed model is compared with various existing models in order to evaluate the efficiency of the proposed model.

Keywords: Content Based Image Retrieval, Image Retrieval, Resnet50, Particle Swarm Optimization, VGG-16, Feature Extraction, Feature Fusion

Received on 04 03 2024, accepted on 15 11 2024, published on 17 01 2025

doi: 10.4108/eetiot.5293

1. Introduction

CBIR is also termed as QBIC (Query by Image Content) [1], is a method of retrieving the images from the DB. In CBIR, a user states a QI and fetches the images in the DB similar to the QI. Retrieval of images were categorized by either local or global characteristics, which heavily depends on visual details. The properties of the images consist of form, texture and color of the images. In CBIR color is considered one of the commonly employed visual function. In addition to color, texture and form of images, descriptors are also considered to be an appropriate factor for Image Retrieval (IR). Therefore, to find the most similar

*Corresponding Author Email: kumaravelpichaimani@gmail.com

images, CBIR compares the content of the input image to DB images. CBIR is considered to one of the significant areas for image processing and computer vision [2]. This could be used in various fields such as health, medicine and prevention of crime and so on. CBIR model usually comprises of 2 important steps, which includes feature extraction [3] and measuring the similitude of indicators. These two steps are vitally carried out in CBIR system. Therefore, various existing studies have incorporated CBIR system for efficient image retrieval.

CBIR process employed in the suggested paper includes pre-processing of the images followed by applying PCA for feature extraction and reduction of dimensionality. Then the images which are extracted were trained by VGG16 network as it aids in generating the training dataset. Followed by the process of generating the training data, binary codes were generated for all images and query has been applied for displaying the relevant images. Finally, the model is assessed by using different metrics such as recall and precision for evaluating the efficiency of the suggested model [4]. Similarly, the image retrieval could be performed by employing neural networks as Deep convolution has the capability to produce efficient outcome for image retrieval system. Therefore, recommended paper has employed softmax function in DL. In addition, image retrieval system held the index of the visual features on main memory which makes it better for efficient image retrieval system. Further, recommended paper has employed VGG16, a pre-trained model along with cosine similarity as the score to detect the similarity score. Finally, the performance of the model has been assessed for generating the outcome [5].

Likewise, the amalgamation of CNN and Sparse representation has been used for increasing the retrieval speed of images as well as the accuracy of the image retrieval process. This process has been performed by employing hand-crafted feature extraction process along with the process of classification for obtaining the output. Further, the process of image retrieval has been assessed by using different metrics like recall and precision [6]. Correspondingly, recommended paper focused on providing better way for retrieving images as the existing methods were considered to be quite inefficient for CBIR. Hence, meta-heuristic inspired ABC (Artificial Bee Colony Optimization) has been considered to be for finding optimal solutions. Further, CBIR system has incorporated GLCM (Gray Level Occurrence Matrix) with ABC algorithm with the aim to improve the accuracy rate of IR system. The process of the suggested model started with feature extraction process, which was based on GLCM from which the image that has been extracted were fed to ABC algorithm as input. Finally, depending upon the feature of the input, ABC algorithm aided in classifying or retrieving the images [7].

CBIR usually comprises of retrieving the most visually related and relevant images to a given query images from a DB of images. However, the output of the existing approaches were considered to be inefficient due to its ability to deliver better similarity score for IR system. Hence, suggested study has employed Pigeon optimization algorithm as feature selection technique for extracting the appropriate images for the model. Process involves by giving the query images for extracting the images, further the training images are fetched from the dataset and then equated with the query images by employing ANN classifier for retrieving the suitable images. Moreover, the model is assessed by employing accuracy, precision, recall and F1- score values [8]. In recent years, various studies has experienced multi-extraction process for more effective IR process. Likewise, suggested paper has focused on multi-extraction process by employing multi-class PSO for feature selection as PSO has the ability to select the best features required for IR. Further, different classifiers were used for classification of features by employing SVM, DT and KNN as these classifiers were known to perform faster than the existing classifiers. After the process of classification, the model is assessed by employing different metrics like precision, F-score, accuracy and recall [9].

Though existing studies delivered reasonable outcome for image retrieval process, it still lagged in delivering better similarity score which is required for effective image retrieval. Therefore, proposed model employs Modified ResNet50 and VGG16 for feature extraction as modified ResNet50 uses extra layers for extracting the best features. After feature extraction, these features are fused by using PCA as PCA assist in reducing the dimensionality of the features thereby making the model effective for retrieving better images. Further, Modified PSO (M-PSO) is used for obtaining optimized features since M-PSO is fast and aids in selecting optimal features after processing from insignificant features that primarily set the preferred number of necessary features. Besides, M-PSO require less parameters to tune rather than huge number of parameters hence incorporation of K parameters of KNN algorithm is used with the aim of find the nearest images to query images, thereby making the model appropriate for IR process with better similarity score Finally, the proposed model is assessed by incorporating various evaluation metrics such as accuracy, precision, F1 score and recall. Major objectives of the paper includes,

- To employ Modified ResNet50 and VGG16 for feature extractions in order to attain best features for image retrieval
- To fuse the extracted features by using PCA (Principal Components Analysis) for feature fusion.
- To find the most relevant from all fused features by using Modified PSO with K parameters of KNN algorithm.
- To assess the efficiency of the model by incorporating performance metrics such as accuracy, precision, recall and F1 score.

1.1. Paper Organization

Section II deals with conventional methods done on a similar domain with a diverse methods, as shown in Further, Section III represents the methodology executed in the projected system. The results and outcomes accomplished by the projected method are shown in Section IV. Finally, the conclusion and future work of the projected system is shown in Section V.

2. Literature Review

Various existing studies with different algorithms have been reviewed in the subsequent section**.**

Due to the development of CBIR technology, the efficiency of retrieval for image retrieval process is getting higher and higher. Based on features of the colour, shape and other aspects of the data images, efficiency of the image retrieval can be identified. However, retrieval of single image is considered to be inadequate for meeting certain requirement of people. Therefore suggested study has employed an image retrieval algorithm which was based on amalgamation of shape and colour features. The further, color feature of the images has been calculated using cumulative histogram approach and shape features has been calculated by using 7 Hu invariant methods. Moreover, the similarity of the images have been identified by using Euclidean distances and shape and colour features were combined with color and shape features. Concurrently, precision and recall metrics were used for measuring the efficacy of the proposed model for image retrieval [10]. The existing CBIR systems have confined efficiency since they extract the only limited set of features. Therefore, recommended paper has focused on extracting vast and significant features from image database and storage of the features in form of feature vector by employing meta-heuristic algorithms such as GA with simulating annealing for similarity evaluation. The suggested CBIR method was constructed based on RGB color along with neutroscopic clustering algorithm and extraction of color features were performed by using canny edge method, further gray level co-occurrence matrix was used for extracting texture features. From the experimental outcome, it was identified that, amalgamation of the existing techniques aided in aggregating the performance of IR, moreover, the efficiency of the model was assessed by using precision and recall [11].

It has been revealed that, retrieval performance of CBIR is vitally dependent on similarity measurements along with feature representation. Therefore, the suggested study has employed a capable algorithm for solving these issues by incorporating DBN (Deep Belief Network) for extraction of features and classification of the extracted features. The experimental outcome of the suggested study delivered positive performance of the model. Feature extraction approach utilized in the study has deliver better and reliable performance than the existing methods as DBN generated huge dataset for learning features and provided an effective content extraction [12]. Likewise, CBIR has been used in the recommended paper for extracting the feature vector from an image in order to efficiently retrieve the content based images. Therefore, recommended paper has employed 2 different sorts of image feature descriptor which includes ORB (Oriented Fast and Rotated Brief) and SIFT (Scale Invariant Feature Transform) extraction methods. SIFT has been used for analyzing the images which was based on different scale and orientation and ORB was used as fast key points. Further, K- means clustering has been employed in order to obtain the clusters. Additionally, the dimensionality reduction was carried out by using locality preserving projection for reducing the dimensions of the image feature vector. The model was performed by using various datasets and the results were evaluated by using precision [13].

Different researches have been performed for retrieval of images, however, the sematic gap between he extracted features and the human idea reduces the meticulousness of image retrieval. Therefore, the existing study employed CNN for extracting deep and high-level features from the images. Further, retrieval system is modelled by using optimization problem such as GA and PSO algorithm. However, the existing optimization algorithms was considered to be slow and inefficient. Therefore, M-GOA (Modified -Grasshopper Optimization Algorithm) has been used for solving the difficult optimization issues. M-GOA had the potential to retrieve similar images effectively, in spite of total search in DB. Therefore, from the experimental outcome, it was identified that amalgamation of CNN-MGOA delivered better accuracy for image retrieval than prevailing methods [14]. On contrary, existing paper has adopted Siamese network, in which the pair of images have been used as inputs and a specific model was learnt to make images belonging to the same class and had similar features by utilizing contrastive loss function and weight sharing function. Further, the recommended model has adapted CNN + hashing methods with Siamese network in order to learn features and performed hash coding in 1 network. Further, the efficiency of existing model had the potential to retrieve the similar images quickly than the conventional hash methods and typical DL methods [15].

Similarly, the suggested study has employed GA (Genetic Algorithm) and SVM (Support Vector Machine) for image retrieval in multiclass set-up. Initially, the features were extracted by using Haar, bi-orthogonal and daubechies wavelets and the features were refined by using GA. Once the features were refined, SVM was used for training the multiclass setup. Moreover, similarity measurement function was identified by using L2 norm, as the L2 norm was employed in between the query and retrieved images alongside the query image from the image repository. From the experimental outcome, it was identified that class imbalance issue was addressed efficaciously in CBIR [16].

The CBIR plays a vital role in different domains as the volume of images present in the DB upsurges enormously. Therefore, it is considered to be quite difficult for comparing the query image feature with every other images in the dataset during the process of retrieval. Further, computational complexity and issues associated to search space degrades the performance of the model. Therefore, the recommended model employed different search space reduction methods for classifying the collection of images into a subset of related images. This was achieved by employing hybrid KMFO (k Means Moth Flame Optimization Algorithm). This method aided in enhancing the performance of K-means classifier by allocating the ideal count of clusters and cluster centroids by employing number of flames and the values of flames which have been attained in MFO. The feature vectors used in the suggested model were HSV color histograms, GLCM and other methods as feature vectors. The suggested hybrid KMFO model was tested in Corel1k dataset and delivered competent outcome with compared with existing retrieval methods [17].

CBIR using low level features like texture and color has been demonstrated by employing BPSO (Bio-geography particle swarm optimization) method. This method relied on content utilized in the image along with classification of the images. The usage of the existing algorithms ensured in effective image retrieval and resulted in highly applicable to the content of an image query. Features such as color of images and filters were used for extracting the feature. Further, the efficacy of the model was identified by using precision and recall [18]. Similarly, the suggested paper has employed DCNN model called as MaxNet for CBIR. Moreover, MaxNet model aided in overcoming the over-fitting issues by employing dropout layer after each inception block and just before the output layer. Further, the similarity index was computed for retrieving the images which were identical to the query images [19]. Likewise, SAR (Deep Search and Rescue) based CBIR has been used for efficient retrieval of appropriate images. DNN-SAR used in the model was pre-processed initially. Then multiple feature extraction has been performed. After feature extraction, the feature fusion has taken place, then the clustering of the features were performed by using SFO algorithm (Sunflower Optimization). Finally from the experimental outcome, it was revealed that relevant images have been retrieved by using SNN-SAR optimization algorithm [20].

Existing paper has employed VGG-19 model for robust image retrieval for CBIR instead of employing hand-crafted feature extraction techniques. Further, ELM (Extreme Learning Machine) has been incorporated for parameter tuning and GA has been employed or reducing the computational expenses by passing the extracted features [21]. Suggested model has used DL based CBIR model for image encryption, termed as IEDL-SCBIR methods for securely retrieving the images. CSO optimization technique was used for optimization. Further, VGG was used for feature extraction with Euclidean distance based similarity evaluation [22]. Likewise, the suggested study has implemented pre-trained CNN model for predicting the class of input query image and aids in reducing the search space of the images. Further, semantic sorting method was used for final image retrieval and the efficiency of the model was assessed by using different performance metrics [23].

2.1. Problem Identification

From the assessment of the above-existing works, core concerns are emphasized as explored below,

- Precision and Recall obtained by the model for image retrieval is considerably less in the existing model [10].
- Existing method is considered to be ineffectual due to extraction of limited number of features [11].

3. Proposed methodology

The proposed study emphasizes on CBIR using DL and meta-heuristic algorithms for retrieving appropriate images and calculating similarity score of the images which have been retrieved. In existing methods are prone to errors and considered to be an expensive process for training the data. Therefore, proposed model has employed Modified ResNet50 model for feature extraction and Modified PSO model for finding the most relevant features from all extracted features as Modified PSO algorithm is fast and it possess the robustness to control the parameters, which aids in determining high correlated images by incorporating k parameters of KNN algorithm for finding the nearest images to query images. Figure.1 shows the mechanism of the proposed model.

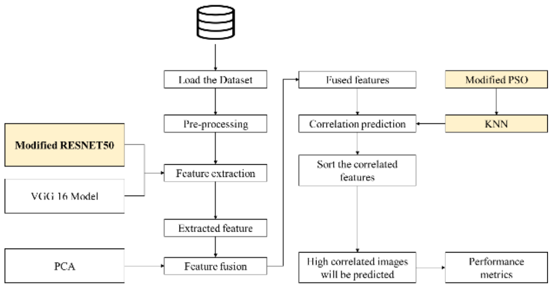


Figure 1. Overall mechanism of the Proposed Framework

Figure.1 shows the overall mechanism of the proposed model. Initially, the dataset is loaded, which consist of Gamma images, Infrared images, Magnetic images, microwave images, radio images, UV images, visible images and X-Ray images. After the process of pre-processing, feature extraction is carried out by employing Modified ResNet50 and VGG-16 model. The modified ResNet50 model incorporates additional newly added dense layers. This incorporation of various dense layers helps in extracting the best features required for the model. Similarly, VGG-16 model is used along with Modified ResNet50 model for effective feature extraction. After feature extraction, feature fusion is carried out by using PCA in order to fuse the features extracted from the images that are more discriminative than the input images and eliminates the redundant information. Further, Modified PSO is implemented for correlation prediction that are associated to each input images. The correlation prediction is accomplished through modified PSO, as the optimal similarity score can be obtained by using this algorithm. Moreover, M-PSO require less parameters to tune, rather than huge number of parameters by incorporating K parameters of KNN algorithm with the aim to find the nearest images to query images, thereby making the model appropriate for IR process with better similarity score. Further, attainment of nearest similar score aids in retrieving topmost high similar images. Finally, the performance of the proposed framework is assessed by incorporating different evaluation metrics which includes accuracy, F1 score, recall and precision.

3.1. Feature extraction – Modified ResNet50

Feature extraction is the process of reducing the number of features in the dataset by producing new features form the prevailing ones. Further, the technique of extracting the features is helpful specially when the dataset is huge and need to eliminate the number of resources without losing any significant and relevant information. However, the existing feature extraction techniques require lot of resources in order to detect the most relevant features from a huge dataset. Therefore, effective feature extraction techniques should be implemented for obtaining the best features. Figure shows the process of the feature extraction in the proposed model by using VGG-16 Model and Modified ResNet50 model.

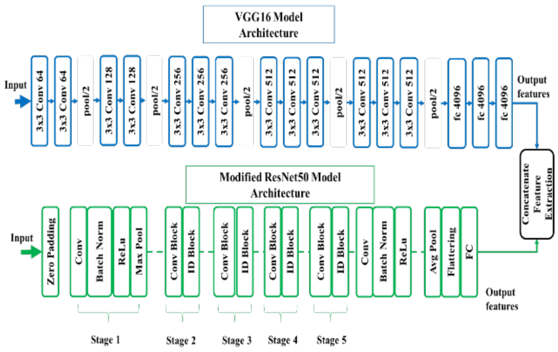


Figure 2. Feature Extraction

The figure shows the modified ResNet50 for extracting the features using extra layers for best feature extraction. In stage 1, different layers and process are used such as convolutional layer, batch normalization, ReLU and Max pooling layer. The Convolution layer (CL) is a layer which comprises of various convolution units for feature extraction. In the next layer, batch normalization is applied as this layer permits each layer of the network to do learning more self-reliantly. Batch normalization aids in normalizing the output of the preceding layers. Further, overfitting of the model can be avoided by incorporating batch normalization. After batch normalization, ReLU activation function has been utilized for resolving the issues associated to vanishing gradient. Further, ReLU helps in preventing the exponential growth in computation which is required to operate the NN. In stage 2, 3, 4 and 5, 2 different block are utilized, which includes convolutional block and identity block. Identity block is considered when there is no alteration in input and output dimensions, correspondingly, the conv block is similar to identity block but it consist of CL, in short cut path in order change the dimension, so that dimension of the input and output matches. The stage 2 conv block consist of 3 layers. Similarly, stage 3, 4 and 5 conv block comprises of 3 layers. Incorporation of extra layers in Modified ResNet50 model aids in extracting best features for feature extraction process. The architecture conclude with an average pooling layer, flattering, and a fully connected layer. The output from this network also feeds into the concatenate feature extraction block, which merges the feature representations from both models, potential enhancing robustness by leveraging their respective strengths. This dual-network architecture is suitable for tasks requiring both deep hierarchical features and residual connections to capture intricate patterns.

The modification to the ResNet50 are specifically enhance the feature extraction capability particularly or tasks involving large dataset. The significant change is to eliminate the down sampling in the last two residual blocks, which preserve the spatial resolution and allows for the retention of fine grained details critical for accurate features extraction. Additionally, the incorporation of atrous convolution increase the receptive field without compromising spatial resolution, allowing for effective multi scale contextual information capture. The modified ResNet50 produces multi-level feature maps that leverage both low level and high level feature, improving performance. The original final layer FC_1000, FC1000_Softmax and FC classification are replaced with FC_3, FC3_softmax and class output to align outputs by enhancing classification accuracy. The feature fusion head emphasizing significant features during fusion, while transfer learning employs pre trained weights for task adaptation. High level features are extracted from the average pooling layer and the architecture retains depth through detailed layer operation and hyper parameter tuning. These modifications reduce redundancy and enhance the discriminative power of extracted features, improving overall model performance compared to standard ResNet50.

Similarly, VGG16 model is used for extracting the features for feature extraction. The main objective of employing VGG16 model is to pass the image representation through stack of CL having ReLU activation function with the total of 16 layers. The implementation of VGG16 model aids in improvising the process of extraction as well as the process of localization. Eventually, the large scale image features extraction with DL are dealt and concatenated with VGG16 and Modified ResNet50 model. Since the features are learned layer by later, the loss of data can be resolved easily.

The features extracted from Modified ResNet50 and VGG16 model are concatenated and fused together in feature fusion process by using PCA. Since feature extraction is considered as a significant part of CBIR, as the content of an image is summarized to a feature vector. Further, the images are compared together according to their features which includes color and texture and shape of images. Moreover, Feature fusion technique is employed, with the aim to fuse the features extracted from the images that are more discriminative than the input images and eliminates the redundant information. PCA is used for feature fusion since PCA assist in dimensionality reduction. Dimensionality reduction is considered to be one of the important aspects in image processing as reductions of dimension aids in mitigating overfitting of the model. PCA utilizes orthogonal transformation with the aim to detect the correlated variables and convert the variables into uncorrelated ones. In order to make the proposed model effective for image retrieval, the dimensions of the FF (fused features) is reduced by utilizing PCA model. Consequently, the most discriminant features are attained for the construction of the model. FF has n dimensions$FF = a_1, a_2, \ldots . a_n$, which requires to be decreased into $K \ll n$ dimensions. Fused features are obtained by employing the following steps,

The steps involves scaling of data, computation of covariance matrix, calculation of eigen value and eigen vector and selection of eigen value.

1) Data scaling is performed by using the equation 1,

$$a_j^i = \frac{a_j^i - \overline{a_j}}{\sigma_j} \tag{1}$$

2) Computation of co-variance matrix by using equation 2,

$$\Sigma = \frac{1}{m} \Sigma_i^m (a_i)(a_i)^T, \Sigma \in R^{nxn} \tag{2}$$

3) Calculation of eigen vector and eigen value by utilizing equation 3,

$$v^t \Sigma = \lambda\mu U = \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix}, v_i \in R^n \tag{3}$$

4) Selection of Eigenvalue

$$a_i^{new} = \begin{bmatrix} v_1^T & a^i \\ v_2^T & a^i \\ v_k^T & a^i \end{bmatrix} \in R^k \tag{4}$$

Once the features are fused by using PCA, the best features for similarity are determined by using Modified PSO algorithm by incorporating KNN algorithm for effective image retrieval.

Table 1. Parameters of PCA

Parameters	
PCA	
n_components	30
n_oversamples	10
tol	1

Table 1 shows the parameters employed for PCA in order to fuse the features obtained from feature extraction process.

3.2. Modified PSO – For Global Optimal Solutions and Better Similarity Detection

The fused features are optimized using Modified PSO for selecting the image feature description with incorporation of k parameter of KNN algorithm for suitable selection of features. The optimized features which are obtained using modified PSO is classified by employing KNN algorithm for selecting next suitable features for image retrieval. Figure.3 shows the flow of Modified-PSO.

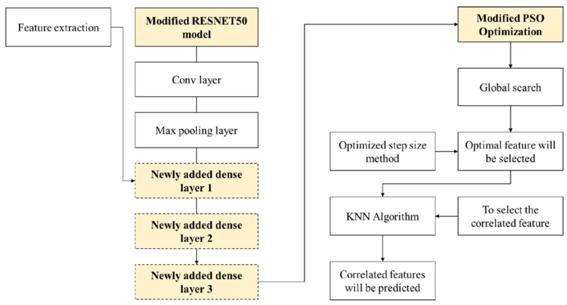


Figure 3. Modified PSO with KNN for Global Optimal Solutions and Better Similarity Detection

PSO algorithm is an intellectual way of solving complex problems by mimicking how creatures work together. PSO employs tiny agents which move around with the aim of detecting the best solution. PSO is motivated by cooperative conduct of natural organism such as birds which moves together in order to accomplish the goal. In PSO, process starts by allocating arbitrary position to every particle of the swarm. Further, the Fitness Value (FV) for each particle is assessed and then the position, individual best as well as the velocity for each particle is updated consequently. Moreover, at every single iteration, the particle which possess the best FV is recognized. Further, the fitness value and the position is saved as the $gbest$ (Global Best). Moreover, the fitness value of the updated particle is estimated and he individual best,$gbest$ position and velocity of each particle is updated for the 2nd time. The process continue till the user- defined termination value is attainted. Eventually, the best solution is obtained (highest fitness value).

Similarly, the modification done to the traditional PSO (Particle swarm optimization), specifically the modified PSO enhances the calculation of each particle's next position to improve convergence and exploration capabilities. The key aspect includes refined particle positioning that enhances convergence, dynamic adjustments based on previous states, and improved encoding and representation of solution areas and exploitation of known solution. The primary goal of incorporating KNN parameters into modified PSO is to optimize the K value for the K nearest neighbor (KNN) algorithm, ensuring that the selected K value improves classification performance. This modification allows for more effective features selection by leveraging the strength of both PSO and KNN, thereby enhancing the search for global optimal solution compared to existing PSO variant. Overall, these improvements aims to increase the efficiency and effectiveness of PSO in parameter optimization tasks, particularly in applications such as image retrieval and fault diagnosis for bearing vibrations.

In PSO, every particle in swarm denotes a solution to the problem and is distinct with its position and velocity. In D-dimensional search space, the position of its particle is denoted by D-dimensional vector$\mathrm{pos_i} = (\mathrm{pos_{il}}, \dots \mathrm{pos_{id}}, \dots . \mathrm{pos_{iD}})$. The velocity can be represented by$\mathrm{vec_{id}} = (\mathrm{vec_{i1}}, \dots . \mathrm{vec_{id}}, \dots . \mathrm{vec_{iD}})$. The best position is formerly noticed by the ith particle as $\mathrm{p_i} = (\mathrm{p_{i1}}, .. \mathrm{p_{id}}, \dots . \mathrm{p_{iD}})$ and p_g is denoted as the index of the particle which looks back the best position of the swarm, further $p_g$will become as the best solution identified. And the velocity and position of the particle is identified by using 2 equations, in which inertia weight w is added by,

$$\mathrm{vec_{id}} = \mathrm{wgtvec_{id}} + c_1 r\, (\mathrm{p_{id}} - \mathrm{pos_{id}}) + c_2\, R\, (\mathrm{p_{gd}} - \mathrm{pos_{id}}) \quad (5)$$

$$\mathrm{pos_{id}} = \mathrm{pos_{id}} + \mathrm{vec_{id}} \quad (6)$$

In which, c_1 and c_2 denoted as positive constants and r and $R$is denoted as 2 arbitrary functions in the range [0,1]; $pos_i = (pos_{il}, \dots pos_{id}, \dots . pos_{iD})$ denotes the location of the ith particle. The parameter w in the equation 5 is the inertia weight which upsurges the entire performance of PSO. It has been reported that larger value of w offers advanced capability for global search, whereas on the other hand lower value of w indicates a developed capability for local research. Therefore, in order to accomplish the higher performance, value of inertia weight w is decreased. The linearly decreases value of inertia is accompanied by using the formula,

$$\mathrm{wgt} = \mathrm{wgt_{max}} - \mathrm{iter}.\frac{\mathrm{wgt_{max}} - \mathrm{wgt_{min}}}{\mathrm{iter_{max}}} \quad (7)$$

Where, $iter_{max}$ is defined as maximum of iteration in the process of evolution and wgt_{max} is denoted as the maximum value of inertia weight, wgt_{min} is denoted as minimum value of inertia weight and $iter$ is defined as the current value of iteration. However, the conventional PSO algorithm possess the capability of ill-using the global maximum, it cannot affirm to accomplish the maximum but frequently falls into local maxima. Therefore, in order to overcome the issues, the proposed algorithm has employed modified PSO algorithm where k parameter of KNN is used for selection features and aids in obtaining the similar images.

In M-PSO, the solutions are encoded as binary vectors and the position of the ith particle is represented by following vector,

$$\mathrm{x} = (\mathrm{p_i^1}, \mathrm{p_i^2}, \cdots, \mathrm{p_i^D}), \;\; \mathrm{i} = 1,2, \dots, \mathrm{E}, \quad (8)$$

In which, $p_i^d = (d = 1,2,3 \dots D)$ represents the location of the particle i in the d^{th} dimension and E represents the count of the particles. Further, in M-PSO, next position of each particle is identified by using,

$$\mathrm{IF}(\mathrm{vec}_i^d\ (t + 1)) = |\tanh(\mathrm{vec}_i^d\ (t + 1))] \tag{9}$$

Here, the parameter k of KNN model is used for M-PSO and $(\mathrm{vec}_i^d\ (t + 1)$ is used as transform function in order to fit the evolution of particles.

$$p_i^d(t+1) = \frac{p_i^{\sim d}(t)\mathrm{rand} < \mathrm{IF}\left(\mathrm{vec}_i^d(t+1)\right)}{p_i^d\ (t)\mathrm{rand} \geq \mathrm{IF}\left(\mathrm{vec}_i^d(t+1)\right)} \tag{10}$$

$p_i^{\sim d}(t)$ is denoted as complement of $p_i^d(t)$, and $rand$ represents the random values from the range 0~1.

The value of the parameter K is selected by randomly initializing the position of E particles, then the position of the particle is checked. Further, the fitness value of each particle is assessed. Further, based on the outcome of the fitness evaluation and then update the P_{best} (position best) and g_{best} (global best) position. Further, velocity and position of the particle is selected, finally the optimized parameter k of KNN can be obtained in order select the correlated feature by using,

$$w_s^k(a, c_j) = \sum_{y \in kNN(a,T):\ \Phi(b)=c_j} s(a,b) \tag{11}$$

Where $kNN(a,T)$ of the labelled document belongs to the training set and it is ordered with respect to up-surging the values of the distance function. The label assigned to the object a by the classifier is the class $c_j \in C$ which increases the sum of similarity between a and the documents labelled c_j. Further, the similarity between the images are computed using $s = 1 - d$.

$$V_{doc}\left(\widehat{\Phi_s^k}, I_i\right) = 1 - \frac{\arg \max\limits_{c_j \in C\{\widehat{\Phi_s^k}, (I_i)\}} w^k(I_i, c_j)}{\arg \max\limits_{c_j \in C} w^k(I_i, c_j)} \tag{12}$$

The KNN classification is obtained by using the function $\widehat{\Phi_s^k}(I_i, \mathrm{I})$ utilized in equation 12. The classification confidence has been used to decide whether the predicted label possess higher probability to be corrected. Where the value of V denotes the high confidence. Therefore, this process aids in speeding up the process of classification process thereby resulting in effective image retrieval process.

4. Results and Discussion

The proposed design has been executed with Python. The first section discusses description of the dataset. The second section describes the performance metrics. The third section presents experimental outcome of the proposed model, fourth section describes the performance exploration of the proposed model. Finally, the Fifth section presents comparative analysis to determine the efficacy of the proposed approach over conventional methods.

4.1. Dataset Description

The dataset consists of 8 different images which includes Gamma, infrared, magnetic images, microwave, Radio, UV, visible and X-Ray images. The images have been collected through diverse types of telescopes in the region of kodaikanal. Particularly, certain images are taken from WARM telescope. These telescope has been mounted at kodaikanal observatory which comprises of doublet lens of 20.06 cm through which 8 different images have been collected with different size of files. Table 2 shows the specifications of the telescope used to capture the sun images at different intervals and table depicts the dataset images and file size of the images incorporated in the proposed study.

Table 2. Specification of telescope

Objective Lens	Doublet Lens Of 15 Cm
Light-Feeding System	Two Mirror-Coelostat System
Image Size Of The Sun	1.25 Cm
Size Of Pixel	7.5 µm
Optical Filters	A) Central Wavelength: 4305 A B) Central-Wavelength: 3933 A

Table 3. Images and file size incorporated in the dataset

Images	File size
Gamma images	64 images
Infrared images	52 images
Magnetic images	51 images
Microwave images	56 images
Radio images	23 images
UV images	46 images
Visible images	51 images
x-ray	45 images

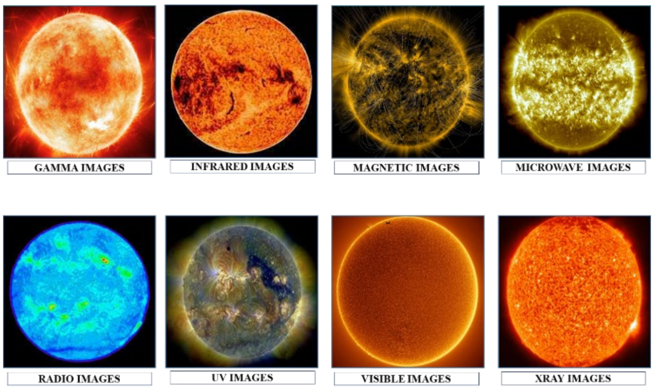


Figure 4. sample dataset images

Figure 4. shows the sample dataset images used in the proposed study, which includes gamma images, infrared images, magnetic images, microwave images, radio images, UV images, visible images and XRay images.

4.2. Performance Metrics

Performance metrics are employed to evaluate the performance of the proposed model by using metrics like accuracy, recall, precision, and F1 score.

Accuracy

Accuracy is defined as the measure of total accurate classification. The accuracy range is calculated with the following equation 13,

$$\mathrm{Acc} = \frac{\mathrm{TRN+TRP}}{\mathrm{TRN+FLN+TRP+FLP}} \tag{13}$$

Where TRN is True Negative, TRP is True positive, FLN is False Negative, and FLP is False Positive.

Recall

Recall is indicated as the solitary of the production metric, which evaluates the total of accurate positive groups created out of all the positive classes. Recall is estimated using subsequent equation 14,

$$\mathrm{Rc} = \frac{\mathrm{TRP}}{\mathrm{FLN+TRP}} \tag{14}$$

In equation 12, FLN refers to False-negative.

Precision

Precision is calculated by measuring the precise classification count. It is measured via improper classification. The precision is calculated with the following equation 1,

$$\mathrm{precision} = \frac{\mathrm{TRP}}{\mathrm{FLP+TRP}} \tag{15}$$

Where TRP is True Positive, FLP is False Positive.

F1 score

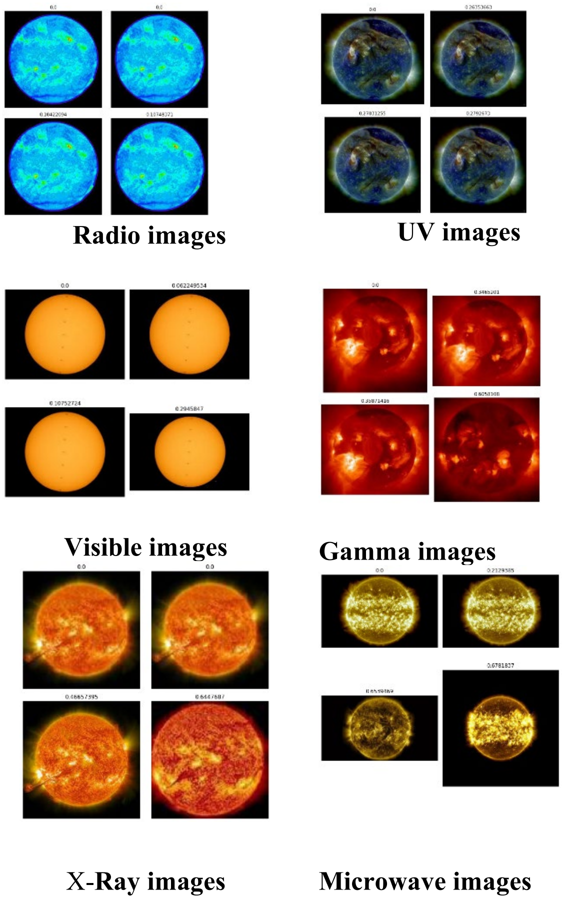
F1 score is denoted as the weighted harmonic-mean value of precision and recall; the F1 score is estimated with the following equation 16,

$$\mathrm{F1-score} = 2 \times \frac{\mathrm{Rc \times Pc}}{\mathrm{Rc+Pc}} \tag{16}$$

Where P is denoted as precision and R is denoted as recall.

4.3. Experimental Results

Outcome that has been attained by the implementing the proposed system is depicted in the subsequent section. Figure 3 shows the similarity score of different images such as radio, UV, visible, gamma, X-ray, infrared, magnetic and images that has obtained between the DB images and query images.



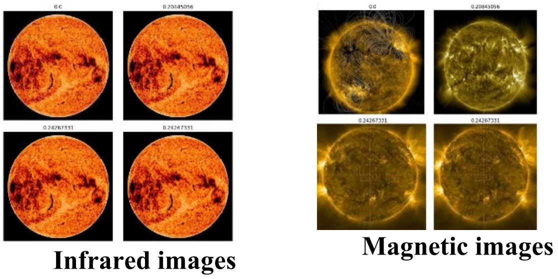


Figure 5. Similarity score of the images

Figure 5 shows the similarity score obtained by the dataset images. From the figure, it can be identified that, Proposed model aided in delivering better similarity score, which leads to effective image retrieval process.

4.4. Performance Analysis

Performance of the proposed model is analyzed using numerous metrics, which includes F1-score, Recall and precision and accuracy. Likewise, confusion matrix is also used for identifying the performance of the model. A Confusion matrix (CM) envisages and encapsulates the performance of a classification algorithm. CM depicts how many predictions are correct and incorrect per class. Figure 6 shows the confusion matrix of the proposed model.

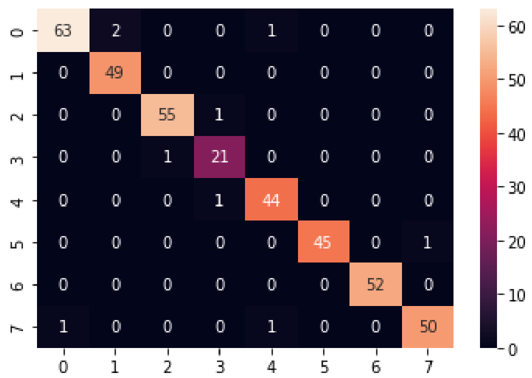


Figure 6. Confusion matrix

Table 4 shows the precision, accuracy, F1 score and recall obtained by the proposed model for the sun chromospheric dataset. Accuracy obtained was 0.98, precision obtained was 0.98, similarly, recall and F1 score obtained by the proposed model is 0.98 and 098. Figure 7 shows the graphical representation of the table.

Table 4. Performance analysis

	Accuracy	Precision	Recall	F1 score
Sun chromospheric dataset	0.98	0.98	0.98	0.98

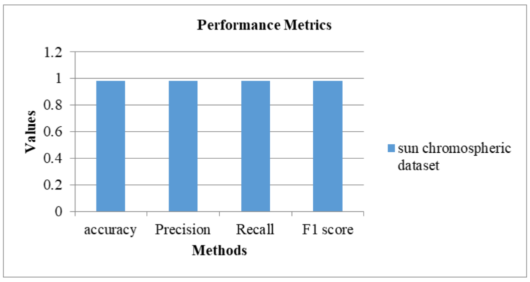


Figure 7. Performance of the proposed model with and without feature extraction

Table 5 and Figure 8 depicts the performance of the proposed model with feature extraction and without feature extraction. Modified ResNet50 has exposed 0.95142 as accuracy with feature extraction and VGG16 has exposed 0.9459 as accuracy with extraction. Similarly, Modified ResNet50 has exposed 0.9214 as accuracy without feature extraction and VGG16 has exposed 0.904 as accuracy without extraction.

Table 5. Proposed model with and without feature extraction

	With extraction	Without extraction
Modified ResNet50	0.95142	0.9214
VGG_16	0.9459	0.904
Feature concatenation	0.96366	0.9227

Feature concatenation rate with feature extraction has been exposed to be high than without extraction. This has confirmed that, accuracy rate enhances with the consideration of feature extraction process. Due to this, the proposed system has been able to explore high accuracy rate in CBIR.

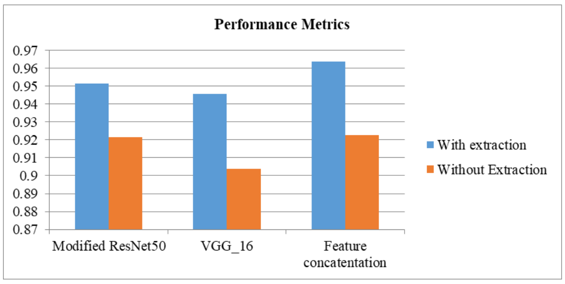


Figure 8. Performance of the proposed model with and without feature extraction

Classification of proposed model is depicted in the table-6. Where the precision, recall, F1 score and support of the proposed model is tabulated. Precision obtained by the model is 0.98, similarly, recall obtained by the proposed model is 0.98, F1 score obtained by the proposed mode is 0.98 and finally support value acquired by the proposed model is 388.

Table 6. Classification for proposed model

	Value of Precision	**Value of Recall**	**F1-Score**	**Support**
0	0.98	0.95	0.97	66
1	0.96	1	0.98	49
2	0.98	0.98	0.98	56
3	0.91	0.95	0.93	22
4	0.96	0.98	0.97	45
5	1	0.98	0.99	46
6	1.00	1	1	52
7	0.98	0.96	0.97	52
Proposed method	**0.98**	**0.98**	**0.98**	**388**

Various aspects have been considered by the proposed model in order to analyze the performance of the proposed model. Confusion matrix has been considered since correct classification predicted is more than the misclassification, likewise, accuracy, recall, precision, F1-score has delivered better outcome. Similarly, with feature extraction and without feature extraction has been considered for analyzing the performance of the proposed model, in which, with feature extraction delivered better results than without extraction. Though, proposed model delivered better outcome, the model is compared with other prevailing methods with the aim of determining the efficacy of the proposed framework.

4.5. Comparative Analysis

Comparative analysis is used to compare the proposed model with the exiting model for analyzing the performance of the proposed model for IR. Table 7 shows the accuracy and precision obtained by existing and proposed model for image retrieval. Accuracy obtained by the existing model is 87.33% whereas the accuracy obtained by the proposed model is 98%. Similarly, precision obtained by the existing model is 86.3%, whereas proposed model obtained precision of 98%.

Table 7. Comparative Analysis of Proposed And Existing Model [24]

Methods	**Accuracy**	**Precision**
Existing (COREL)	87.33%	86.36%
Proposed Model	98	98

Figure 9 shows the comparative analysis of the existing and proposed model for image retrieval process. From the figure it can be identified that proposed model delivered better accuracy and precision than the existing model.

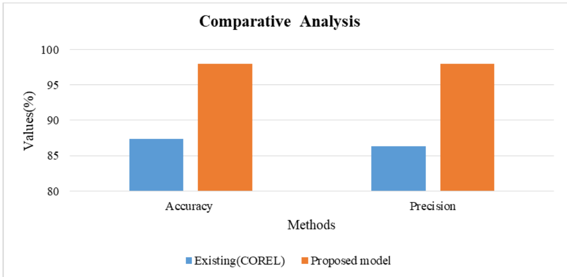


Figure 9. Comparative Analysis of Proposed And Existing Model [24]

Table 8 displays the comparative analysis of the existing and proposed model for Corel-1K dataset. Various models such as MTSD (Multi Trend Structure Descriptor), LTCoP (Local Ternary Co-Occurrence Pattern), LMeP (Local Mesh Pattern), LQEP (Local Quantized Extrema Pattern), DLTerQEP (Directional Local Quantized Ternary Extrema Pattern), existing model were used for comparing with the proposed model. Accuracy obtained by MTSD is 83.23%, similarly accuracy obtained by LTCoP, LMeP, LQEP, DLTerQEP and the existing model is 81.21%, 83.24%, 85.45%, 87.23% and 87.78%. Figure.10. shows the comparative analysis of the proposed and existing model.

Table 8. Comparative Analysis of Existing and Proposed model [25]

	MTSD	LTCoP	LMeP	LQEP	DLTerQEP	Existing	Proposed
Dataset: Corel-1k	83.23	81.21	83.24	85.45	87.23	87.78	98

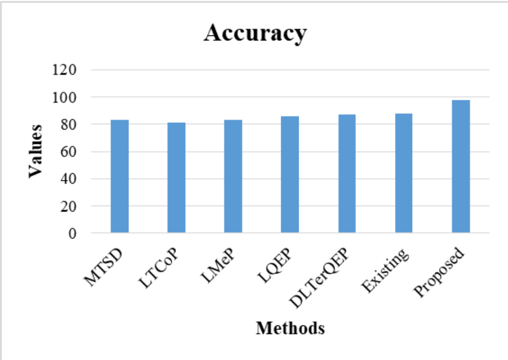


Figure 10. Comparative analysis of the proposed method

However, the proposed model attained accuracy of 98%. From the experimental outcome, it was identified that the proposed model delivered enhanced outcome than the prevailing models. Table 9 shows the dimensions and feature extraction time of the prevailing and proposed model. Dimensions of the prevailing and proposed model obtained is 219 and 224. Likewise, feature extraction time obtained by the existing model is 5.026, whereas the feature extraction time obtained by the proposed model is 2.11.

Table 9. Dimensions and Feature Extraction Time of Existing and Proposed Model [24]

	Dimensions	Feature extraction time
Existing model	219	5.026
Proposed Model	224	2.11

Table 10 shows the dimensions obtained by the prevailing approaches and the proposed method for the images. The methods used are LTCoP with dimensions of 512 similarly, the dimensions obtained by the LMeP, LQEP, DLTerQEP, MTST were 768, 4096, 137. Similarly, the dimension obtained by the existing method is 685 and 145. However, the proposed model obtained dimensions of 224.

Table 10. Analysis with regard to dimensions [25]

METHODS	DIMENSIONS
DLTerQEP	2 x 4096
LTCoP	512
LQEP	4096
MTSD	137
LMeP	768
Existing	685
	145
Proposed model	**224**

From the experimental outcome, it has been identified that, proposed model has delivered better results than the existing models for image retrieval. This effective outcome for retrieving the similar images is delivered by employing modified ResNet50 and VGG16 model for feature extraction, as modified ResNet50 uses extra dense layers which makes the model efficient for extracting the best features along VGG16 model. Further, the features which are extracted by feature extraction process is fused using PCA as PCA aids in reducing the dimensionality and helps in fusing the features which are required for effective image retrieval process. Besides, correlated features for retrieving the images is obtained by using Modified PSO algorithm. The modified PSO algorithm utilizes k parameters of KNN algorithm for selection of best features which assist in selecting the correlated features for retrieving the suitable images.

Table 11. Comparison of Genetic Algorithm

Proposed Framework	With GA
Modified ResNet50	0.9162
VGG_16	0.9
Feature concatenation	0.9112

Table 11 indicates that the proposed framework outperforms the Genetic Algorithm in terms of feature extraction effectiveness, with values of 0.9162 for the modified ResNet50, compared to 0.9 for VGG_16 and 0.9112 for feature concatenation. This demonstrates that the proposed method not only provides better results but also highlights the advantages of using modified approach to optimize the feature extraction in image retrieval tasks. By including these comparisons, the effectiveness of the proposed method can be contextualized within the broader landscape of existing techniques, thereby strengthening its contribution to the field and addressing the reviewer's concerns regarding the depth of analysis.

Table 12. Computational Efficiency of the Proposed Framework

Proposed Framework	Training Time (Seconds)
Modified ResNet50 + MPSO	45
VGG_16 + MPSO	67.5
Feature concatenation + MPSO	56

The result presented in table 12 highlights the computational efficiency of the proposed framework which are alternatively checked instead of modified PSO, demonstrating that the Modified ResNet50 combined with Modified PSO achieves a training time of just 45 seconds. In comparison, VGG_16 with MPSO takes 67.5 seconds, and feature concatenation with MPSO requires 56 seconds. This indicates that the proposed methods not only maintain the high performance but also significantly reduce the training time, suggesting a more efficient utilization of computational resource despite the added complexity from dense layers benchmark provides valuable insights into the training time the paper would benefit from additional metrics such as inference speed and resource consumption to offer a comprehensive analysis of computational efficiency.

5. Conclusion

CBIR is a process of retrieving the most visually similar images to a given query images from DB of images. However, the existing methods used for retrieving the images are considered to be slow and not precise for retrieving the images. Therefore, proposed model employed 8 different images of sun, namely which includes Gamma, infrared, magnetic images, microwave, Radio, UV, visible and X-Ray images for retrieving similar images. Effective image retrieval was performed by employing modified ResNet50 and VGG16 for extracting best features since, M-ResNet50 model utilized extra dense layers which aided in extracting suitable features along with VGG16 model. The concatenated features which were obtained using feature extraction process is fused by using PCA, as PCA has been used for dimensionality reduction, thereby aided in selecting the optimal features for IR. Moreover, features which were required for IR was obtained by using M-PSO. Furthermore, selection of next appropriate features was accomplished by utilizing k parameters of KNN algorithm, thereby making the model effective for retrieving the appropriate images. Finally, efficiency of the model was assessed by using different metrics like accuracy, recall, precision and f1 score, in which the values obtained by the proposed model for accuracy, recall, precision and F1 score was 98%, 98%, 98% and 98%. Further, the proposed model was compared with the existing models for assessing the efficacy of the proposed model. In future, different meta-heuristic algorithms could be used for enhancing the accuracy of image retrieval process.

References

[1] S. C. Yenigalla, S. Rao, and P. S. J. E. P. Ngangbam, "Implementation of Content-Based Image Retrieval Using Artificial Neural Networks," vol. 34, no. 1, p. 25, 2023.
[2] V.-H. J. M. T. Vu and Applications, "Content-based image retrieval with fuzzy clustering for feature vector normalization," pp. 1-21, 2023.
[3] M. Hadid, Q. M. Hussein, Z. Al-Qaysi, M. Ahmed, M. M. J. I. J. F. C. S. Salih, and Mathematics, "An Overview of Content-Based Image Retrieval Methods And Techniques," vol. 4, no. 3, pp. 66-78, 2023.
[4] R. Sharma, B. Rai, S. J. W. J. o. C. Sharma, and M. Science, "Classification on Unsupervised Deep Hashing With Pseudo Labels Using Support Vector Machine for Scalable Image Retrieval," vol. 2, no. 2, pp. 40-55, 2023.
[5] H. J. a. p. a. Tanioka, "A Fast Content-Based Image Retrieval Method Using Deep Visual Features," 2019.
[6] A. Sezavar, H. Farsi, S. J. M. T. Mohamadzadeh, and Applications, "Content-based image retrieval by combining convolutional neural networks and sparse representation," vol. 78, pp. 20895-20912, 2019.
[7] A. J. C. I. Banharnsakun, "Artificial bee colony algorithm for content-based image retrieval," vol. 36, no. 1, pp. 351-367, 2020.
[8] M. Buvana, K. Muthumayil, and T. J. A. o. t. R. S. f. C. B. Jayasankar, "Content-based image retrieval based on hybrid feature extraction and feature selection technique pigeon inspired based optimization," pp. 424-443, 2021.
[9] M. Garg, G. J. N. C. Dhiman, and Applications, "A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants," vol. 33, pp. 1311-1328, 2021.
[10] X. Zenggang, T. Zhiwen, C. Xiaowen, Z. Xue-min, Z. Kaibin, and Y. J. J. o. s. p. s. Conghuan, "Research on image retrieval algorithm based on combination of color and shape features," vol. 93, pp. 139-146, 2021.
[11] M. K. J. A. J. f. S. Alsmadi and Engineering, "Content-based image retrieval using color, shape and texture descriptors and features," vol. 45, no. 4, pp. 3317-3330, 2020.
[12] R. R. Saritha, V. Paul, and P. G. J. C. C. Kumar, "Content based image retrieval using deep learning process," vol. 22, pp. 4187-4200, 2019.

[13] P. Chhabra, N. K. Garg, M. J. N. C. Kumar, and Applications, "Content-based image retrieval system using ORB and SIFT features," vol. 32, pp. 2725-2733, 2020.
[14] A. Sezavar, H. Farsi, and S. J. I. J. o. E. Mohamadzadeh, "A modified grasshopper optimization algorithm combined with cnn for content based image retrieval," vol. 32, no. 7, pp. 924-930, 2019.
[15] Y. Cai, Y. Li, C. Qiu, J. Ma, and X. J. I. a. Gao, "Medical image retrieval based on convolutional neural network and supervised hashing," vol. 7, pp. 51877-51885, 2019.
[16] U. A. Khan, A. Javed, R. J. M. T. Ashraf, and Applications, "An effective hybrid framework for content based image retrieval (CBIR)," vol. 80, pp. 26911-26937, 2021.
[17] A. Joseph, E. S. Rex, S. Christopher, and J. J. A. J. o. G. Jose, "Content-based image retrieval using hybrid k-means moth flame optimization algorithm," vol. 14, no. 8, p. 687, 2021.
[18] D. Dubey and G. S. J. W. P. C. Tomar, "Interactive Biogeography Particle Swarm Optimization for Content Based Image Retrieval," vol. 107, no. 2, pp. 907-921, 2019.
[19] S. Hussain, M. A. Zia, and W. J. E. S. w. A. Arshad, "Additive deep feature optimization for semantic image retrieval," vol. 170, p. 114545, 2021.
[20] N. Keisham and A. J. S. C. Neelima, "Efficient content-based image retrieval using deep search and rescue algorithm," vol. 26, no. 4, pp. 1597-1616, 2022.
[21] R. Bibi, Z. Mehmood, A. Munshi, R. M. Yousaf, and S. S. J. P. o. Ahmed, "Deep features optimization based on a transfer learning, genetic algorithm, and extreme learning machine for robust content-based image retrieval," vol. 17, no. 10, p. e0274764, 2022.
[22] M. Elsharkawy, A. N. J. J. o. C. Al Masri, and I. Management, "A novel image encryption with deep learning model for secure content based image retrieval," no. 2, pp. 54-4-64, 2021.
[23] J. Pradhan, A. K. Pal, M. S. Obaidat, and S. H. Islam, "A post dynamic clustering approach for classification-based image retrieval," 2020: IEEE, pp. 1-7.
[24] R. Raja, S. Kumar, and M. R. Mahmood, "Color object detection based image retrieval using ROI segmentation with multi-feature method," *Wireless Personal Communications,* vol. 112, no. 1, pp. 169-192, 2020.
[25] S. Sathiamoorthy and M. Natarajan, "An efficient content based image retrieval using enhanced multi-trend structure descriptor," *SN Applied Sciences,* vol. 2, no. 2, pp. 1-20, 2020.