Journal of Artificial Intelligence and Data Mining (JAIDM), Vol. 9, No. 2, 2021, 245-257.

Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research Paper

Relevance Feedback-based Image Retrieval using Particle Swarm Optimization

Fatemeh Jafarinejad[1*] and Rezvaneh Farzbood[2]

1. Faculty of Computer Engineering, Shahrood University of Technology, Shahrood, Iran.
2. Faculty of Computer Engineering, Shahid Beheshti University, Tehran, Iran.

Article Info

Article History:
Received 10 October 2019
Revised 01 April 2020
Accepted 24 November 2020

DOI:10.22044/jadm.2020.9014.2037

Keywords:
Image Retrieval, Swarm Optimization Algorithm, Relevance Feedback, Image Classification.

**Corresponding author: jafarinejad@shahroodut.ac.ir (F. Jafarinejad).*

Abstract

Image retrieval is a basic task in many content-based image systems. Achieving a high precision, while maintaining the computation time, is very important in relevance feedback-based image retrieval systems. This paper establishes an analogy between this and the task of image classification. Therefore, in the image retrieval problem, we will obtain an optimized decision surface that separates the dataset images into two categories of relevant/irrelevant images corresponding to the query image. This problem is viewed and solved as an optimization problem using the particle optimization algorithm. Although the particle swarm optimization (PSO) algorithm is widely used in the field of image retrieval, no one uses it for a direct feature weighting. The information extracted from the user feedbacks will guide particles in order to find the optimal weights of various features of images (color-, shape- or texture-based features). Fusion of these very non-homogenous features require a feature weighting algorithm that will take place by the help of the PSO algorithm. Accordingly, an innovative fitness function is proposed to evaluate each particle's position. The experimental results on the Wang dataset and Corel-10k indicate that the average precision of the proposed method is higher than the other semi-automatic and automatic approaches. Moreover, the proposed method suggests a reduction in the computational complexity in comparison with the other PSO-based image retrieval methods.

1. Introduction

In a content-based image retrieval (CBIR) system, digital images are searched and retrieved from a massive dataset. Today, this approach has found its application in search engines [1], web-based image retrieval systems [2, 3], medical imaging recovery [4, 5], creation of an image ontology [6, 7], and so on.

The image retrieval methods can be divided into three general categories of metadata-based image retrieval, automatic CBIR, and semi-automatic image retrieval using relevance feedback (RF).

In the metadata-based image retrieval methods [8, 9], tags, keywords or text descriptions are used in order to retrieve text-based images [9]. Combination of topic-related features, derived from the latent Dirichlet allocation model, and spatial features, computed by a spatial location model, [8] can be mentioned as another kind of features used in this perspective. Although these methods are highly accurate in special-purpose applications with accurate labeled datasets, they cannot be used in public applications due to the timing of the labeling task.

Completely automatic CBIR methods (e.g. [10-12]) use machine learning approaches to retrieve the related images. This approach retrieves images with the help of image classification. In [11, 13], some neural network-based architectures are proposed for content-based image retrieval. Mohammed and Abdelhalim [11] have used an optimized pulse-coupled neural network to extract the visual features of the image, and have applied the K-nearest neighbors algorithm for classification and matching. In [13], wavelet packets and Eigen values of Gabor filters have been used as efficient features in this regard. A bag-of-visual-words (BoVW) model has been addressed in [14]. This method uses the visual words integration of the local intensity order pattern (LIOP) feature and the local binary pattern variance (LBPV) feature in order to reduce the issue of the semantic gap and enhance the performance of CBIR. Priyanka [15] has proposed a convolutional neural network (CNN)-based architecture for image retrieval. He has proposed a micro-structure pattern extraction technique in order to isolate the object from the background in an image. Thereafter, the similarity-based approach based on distance metrics is used for feature mapping. It has a good result on automatic CBIR. Rao *et al.* [16] have used three features called dynamic dominant color (DDC), motif co-occurrence matrix (MCM), and difference between pixels of scan pattern (DBPSP) to retrieve the images. Automatic image retrieval efforts have relatively ignored the two distinct characteristics of the CBIR systems: 1) The gap between high level concepts and low-level features; 2) Subjectivity of human perception of visual content [17]. Due to scrimmaging with semantic and meaning of images, the accuracy of these completely automatic methods is overshadowed.

Among these two approaches, semi-automatic relevance feedback-based methods[6, 18] are some iterative methods using the user feedback in each run to improve the retrieval procedure. The support vector machines (SVMs) have been used in [19, 20] for semi-automatic image retrieval. Cai-Hong *et al*. [21] have used a combination of SVM and the particle swarm optimization (PSO) algorithm for this task. SVM has been used to classify the dataset images into some categories in the first run. The feedbacks received from the user and PSO make some corrections on the original SVM-based decision surfaces. The PSO algorithm has been used for retrieving relevant images in [22]. At the first run of the retrieval system, distance of dataset images and query image is computed. Using this criterion, the PSO initialization takes place. Receiving the user feedbacks in each run, the weights of various features change using just the information from the relevant/irrelevant images and separately from the PSO parameters. Finally, the PSO parameters will be updated, and a feature re-weighting phase takes place, according to which, new relevant images will be shown to the user. Image representation for CBIR is in [23], which is based on complementary visual words. In this paper, integration of speeded up robust features (SURF) and co-occurrence histograms of oriented gradients (CoHOG) have been used as the local and feature descriptor, respectively.

Other metaheuristic algorithms [24], genetic algorithm [25, 26], reinforcement learning [27], and deep learning [28, 29] are some other approaches used in this regard to illustrate how to use the user feedback to retrieve relevant images. This approach represents a retrieval framework that exploits a hybrid feature space that is built by integrating the low-level image features and high-level semantic terms through rounds of RF and performs similarity-based retrieval to support semi-automatic image interpretation [30]. This will cause a final accurate result for critical tasks, for example in medical image retrieval, and bone tumor radiograph [30]. Reduction in the semantic gap, attaining more accuracy, reduction in computation complexity, and hence in the execution time, are some other advantages of this methodology [28].

The learning-based and relevance feedback-based image retrieval methods use information about the content of image, not the information in the tags and captions describing the image, to find the related images. The features used in this regard are divided into three general categories: color-, shape- and texture-related features [31]. The features associated with the color of the image generally include the histogram of different color channels of a color model, e.g. RGB model. The features associated with the shape of the image are generally obtained by the first, second or third moment of the image or by the edge detection methods. The texture-based features are generally divided into four categories [32] of statistical-based, transform based, model-based, and structural-driven transfers. In [32] , the authors have provided a comparison among all of these methods and their uses in other research works.

Generally, the metadata-based and automatic image retrieval approaches are very precise. However, in addition to data images, they require additional information about images (such as text labels or image classification information). Providing this information will cause the data production process to be a very time-consuming task. In addition, the metadata-based approach is just applicable to images with text labels. The RF-based image retrieval approach is a general approach that can be applied to a set of images without any further information about them. In this approach, the lack of quantitative data (tags or categories) will be compensated by using the user's feedback, which also covers the different and time-varying users' need. Bridging between the user and the retrieval system, this approach leads to a much improved retrieval performance by updating a query and similarity measures according to the user's preference.

In this paper, we will address the issue of RF-based image retrieval. The novelty of this work is to use the PSO algorithm for computation of suitable feature weights (degree of importance of each feature). In other words, using the PSO algorithm, we will look for a suitable relevant/irrelevant separator decision hyper plane as the scenario occurs in the classification problem. This will cause a direct PSO-based feature weighting and retrieval, which will decrease the computational complexity of image retrieval in comparison with other PSO-based CBIR (e.g. [22]), while preserving the precision. Accordingly, we propose a new fitness function for particles. This causes our method to have a higher precision than the other semi-automatic CBIR approaches on the Wang dataset [33]. The automatic CBIR methods (that have additional information related to image categorization) use the image classification paradigm. In other words, in those methods, the issue of image retrieval is changed to a multi-class classification problem. However, in the RF-based approach, we are deprived of image categorization information. In fact, each user's feedback will provide information about the binary related/unrelated classes in the subset of dataset images shown to the user.

Some other research works (e.g. [22]) have used PSO for feature re-weighting after a computationally complex phase for manually weighting features after each user feedback. Although the proposed method has a less computational complexity than the method in [22, 24], the results of testing this method on the Wang dataset [33] indicate the comparative precision of this method versus these methods on this dataset. In more detail, the proposed algorithm is more accurate than [24] in retrieving images of all categories of this dataset. Moreover, except in one category, it is more accurate or has the same precision as [۲۲]. This indicates the usefulness of the proposed method in the semi-automatic content-based image retrieval based on the user feedback. A detailed description of the background idea and implementation of the proposed approach is given in Section 3.

The remainder of this paper is structured as what follows. Section 2 introduces the basic concepts used in this paper. Section 3 presents the proposed procedure of image retrieval. In this section, we will introduce the features and fitness functions used. In Section 4, an example of recovered images is provided, followed by evaluation of the proposed method in comparison with some analogous methods and some automatic image retrieval methods. Finally, we conclude the paper.

2. Particle Swarm Optimization

The swarm intelligence-based techniques [34-36] use divide and conquer methodology to overcome the complexity of solving large and difficult problems. PSO, originated from the analysis of the behavior of birds catching food [37], is a stochastic local-search optimization technique. In this method, a population of individuals (swarms) search for an optimal solution in the work-space of problem. Image retrieval [21], structural design optimization of vehicle components [38], and multi-objective optimization of vehicle crashworthiness [39] are some of the extensive applications of the PSO algorithm.

In this algorithm, two key concepts, i.e. position and velocity, are used. The concept of position corresponds to the characteristics of each solution. On the other hand, the rate of change of the current position of a particle to find the successor position is recognized by the velocity. At initialization, the position and velocity of each particle are set randomly. Movements of particles and actually exploration of state-space by particles take place in the form of some mathematical formulas represented in Equations 1 and 2. In these formulas, the next position of a particle is calculated by a combination of the best position they have experienced so far and the global best position experienced by the entire population.

$$V_i(t+1) = \omega V_i(t) + \eta_1 r(p_i - X_i(t)) + \eta_2 r(g - X_i(t)) \quad (1)$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \quad (2)$$

where $V_i(t)$ and $X_i(t)$ represent the velocity and position of the ith particle in the tth iteration, respectively, ω is the inertial factor that is reduced from one to zero as training continues, η_1 and η_2 are the cognitive learning rate and the social learning rate, respectively, p_i and g represent the particle's best position and globally best position, respectively, and r is a random number generated in each iteration. The velocity update (Equation 1) depends on three different terms: inertia term, cognitive term, and social term. The inertia term is related to the tendency to maintain the previous velocity. The cognitive term determines the relative influence of memory of individuals in maintaining their best experience. The social term corresponds to the tendency to use the best experience of the population. The value of each component of velocity vector $\boldsymbol{v}$ can be clamped to the range of $[V_{min}, V_{max}]$ to reduce the likelihood of the particles leaving the search space. The overall procedure of the PSO technique is illustrated in algorithm 1.

Algorithm 1- Particle Swarm Optimization Algorithm:

Output: best solution obtained in a search space
For each particle i
    X_i and V_i are randomly initialized
    p_i originally sets to X_i
    $e_i = ft(X_i)$, fitness of ith particle with position X_i is calculated
    $g = argmin_{i=1}^{|partilces|} e_i$
For each time step t
    For each particle i
        Update particle's velocity:
            Generate random value r
            Compute V_i using equation 1.
        Update particle's position according to equation 2.
        Evaluate fitness of this position as $ft(X_i)$
        Update particle's best position if necessary
Update globally best position if necessary.

2.1. PSO in Image Retrieval

In order to better understand the common use of the PSO algorithm in image retrieval as well as a better comparison of the proposed method with previous ones, we will explain one of the PSO-based retrieval procedures addressed in [20] in this section. The whole procedure is schematically expressed in Figure 1.

According to Figure 1, after receiving a query image, the retrieval system computes its distance with all dataset images (according to some features) and shows the closest images to the user to get his/her RF. The distance metric is a weighted some of distances of each feature sets, addressed in (3).

$$Dist(\mathrm{Im}_{req}, \mathrm{Im}_j) = \sum_{f \in feature\_set} WMSE^f(\mathrm{Im}_{req}^f, \mathrm{Im}_j^f) \quad (3)$$

$$WMSE^f(x,y) = \frac{1}{S}\sum_{s=1}^{S}(x_s, y_s)^2 w^{k,s} \quad (4)$$

where $Dist$ is a function of how distant two images are from each other. It is computed as the sum of the weighted mean squared error of features of the two images as (4). S is the number of total features of the feature category and $w^{k,S}$ is the weight of importance of feature set S in the kth iteration. This weight is uniform in the first iteration, and will be updated after each RF as (5).

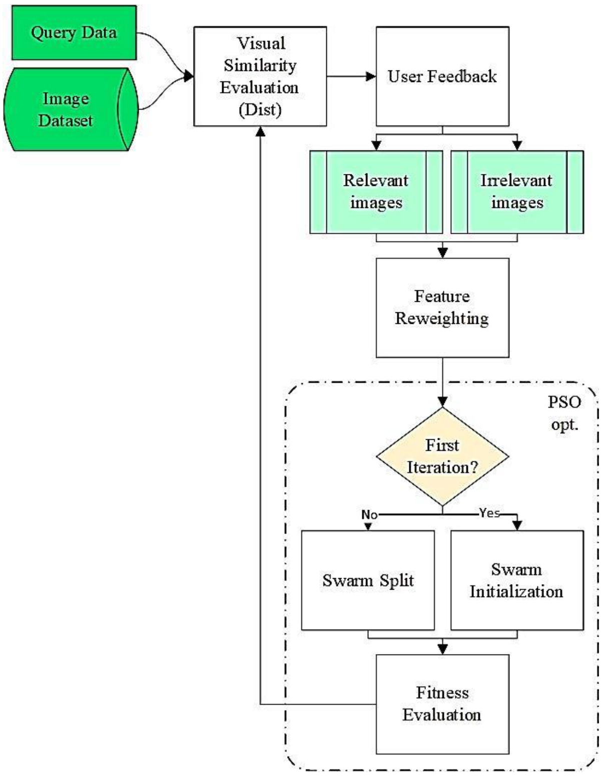

Figure 1. A flowchart of the retrieval procedure in [20].

In each run, after receiving an RF, the shown images are categorized into two distinct classes of relevant and irrelevant images. According to these classes, the weight of importance of each feature is computed as (5).

$$w^{k,f} = \frac{\delta^{k,f}}{\delta_{REK}^{k,f}} \quad (5)$$

$$\delta^{k,f} = 1 - \frac{\phi_{CIrr}^{k,f}}{\phi_{Irr}^{k}} \quad (6)$$

where Φ_{Irr}^{k} is the number of irrelevant images at the kth iteration, while $\Phi_{CIrr}^{k,f}$ is the number of images in the set of irrelevant images that have the feature f within the range associated to the corresponding feature in the set of relevant images, and $\delta_{REL}^{k,f}$ is the standard deviation of the fth feature in the set of relevant images at the kth iteration. Thereafter, the PSO algorithm will be used more and more to retrieve the images that are close to relevant images and far from irrelevant ones according to fitness function (7).

$$fittness(p_n) = \frac{1}{N_{rel}^{k}}\sum_{r=1}^{N_{rel}} Dist(p_n^k, \mathrm{Im}_r^k) + \frac{1}{\frac{1}{N_{irr}^{k}}\sum_{irr=1}^{N_{irr}^{k}} Dist(p_n^k, \mathrm{Im}_{ir}^k)} \quad (7)$$

where the particles correspond to some images of dataset, and the fitness function produces lower values when the particle is close to the relevant set. N_{rel}^{k} (N_{irr}^{k}) is the number of relevant (irrelevant) images at the kth iteration, and Im_{r}^{k} (Im_{ir}^{k}) are the images of the relevant (irrelevant) set.

Thereafter, the 16 best particles (evaluated with fitness function 7) are shown to the user to receive his/her RF. The operators of this image retrieval procedure are very costly. In the next section, we will propose a PSO-based retrieval system with more simple operators (more simple fitness function that computes the weight of features and removes costly computation of feature weights as (5, 6) as well as finding the related images with this fitness function without the need for a further search to retrieve them with (7)).

3. Proposed Method

The problem of image retrieval can be considered as a sort of image classification problem. In this case, we want to categorize the images of a dataset into two related/unrelated categories (according to the input query image). Of course, lack of training samples leads to a major difference between these two categories, and it is the main problem in image retrieval. In image classification, we are looking for a separator decision hyperplane that separates the input into some categories. The equation of this hyperplane can ultimately be considered as the degree of importance of each one of the n different features of this (n + 1)-dimensional solution space. Moreover, some problems do not have a linearly separable solution (something that is highly probable in the classification in high-dimensional spaces, e.g. images). The assuming images of a specified category are located in a convex hull; several decision surfaces, instead of one, is required to distinguish between the two classes.

In the image retrieval problem, we will look for a weight vector that converts the difference of image properties to weighted overall difference between the relevant/irrelevant images and the query image. Hence, the problem becomes as the issue of finding a decision plane in classification problem. Of course, there are two problems: the non-linearly separable classes and the lack of training samples. In fact, the feedback received from the user plays the role of supervisor providing a small training data. Using this data, we find the weight vector that divides the images displayed to the user into two categories well. We will give the task of finding this surface to the particles of the PSO algorithm. Furthermore, given that, in general, the related/unrelated images are not linearly separable, we may need more than one weight vector. These weight vectors will be created separately in each run of the algorithm. First, we obtain a vector randomly. Then we use the feedback of the user to improve it to provide a weight vector, according to which, most of the images shown to the user are categorized correctly. In the next section, the idea will be put into practice.

3.1. Proposed Procedure

According to the idea addressed in Section 3.1, we proposed a procedure for retrieving the relevant images in a semi-automatic content-based image retrieval system using RF.

The general procedure of the proposed image retrieval system is shown in Figure 2. The system extracts features of all database images offline. These feature vectors form a feature repository given to the system at the computation time. The user initially sends a request image to retrieve similar images. These offline extracted features will be used to determine the similarity of dataset images with the requested image.

Latif *et al*. [24] and Sleit *et al*. [31] have categorized the features used in CBIR into three general categories of color-, shape- and texture-related features. Five sets of features are used in this paper, covering all of these categories as:

(a) RGB color histogram (a kind of color feature), which is addressed in Equation (8).

$$hist(\mathrm{Im}, c, b) = \frac{N_k}{N} \tag{8}$$

where $hist$ is the normalized histogram of channel c of image im in RGB color model, N is the total number of pixels (64*64), and N_k is the number of pixels with values in kth bin, b_k of histogram. The total number of features of this feature set is 3 * 256.

(b) median, variance, and skewness of histogram of the image in HSV color space, (c) variance and skewness of gray scale image, and (d) gradient magnitude computed using Sobel operator are used as some kinds of shape features. The basic operation of these features is addressed in Equations (9)-(12).

$$\mu = \frac{\sum_{i=1}^{N} \mathrm{Im}_i}{N} \tag{9}$$

$$\sigma^2 = \frac{E[(\mathrm{Im} - \mu)^2]}{N} \tag{10}$$

$$s = \frac{E[(\text{Im} - \mu)^3]}{\sigma^3} \tag{11}$$

$$G = \sqrt{G_X^2 + G_y^2)}$$
$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}, \tag{12}$$
$$G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

where $\mu, \sigma^2,$ and s are the median, variance, and skewness of the image Im, respectively. G is the gradient magnitude. The feature-set (b) containing median, variance, and skewness of the histogram of H, V, S, and Y channels have 12 features, totally. The feature-set (c) containing variance and skewness of gray scale image have 2 * 64 dimensions. Finally, the dimension of the feature-set (d) is 3844.

(e) Entropy filter with 9-by-9 neighborhood around a corresponding pixel (representing the texture), which has the same dimension as the main image (i.e. 64 * 64).

Therefore, the total number of features are 8848 features, which are categorize into 5 feature sets, each of which has a unique importance weight. In each run of the algorithm, the sum of these weights will be normalized to 1. Moreover, it can be seen that each feature set has a different dimension that can make them non-homogenous in computing their importance weight. Hence, we finally divide each value of each feature set by its number of features. Utilization of these non-homogenous features requires their fusion via some intelligent algorithms, e.g. PSO. Accordingly, the distance between each feature in the database images and the requested image are calculated. In order to obtain the final distance of each DB image and query image, the weighted average of these feature distances will be used. Hence, obtaining a suitable weight vector is very important in computation of the overall distances.

In the first run, the weight of all features is the same (each feature set has the weight of 1/5). According to this uniformly distributed weight, more similar images are obtained and shown to the user.

Getting the user feedback on whether each of the extracted images are relevant or not, the algorithm enters another phase in which the weight of the features will be changed using the particle optimization method. Thereafter, according to the newly obtained weight, which is the globally best position of particles, new relevant images are computed and shown to the user.

As shown in the flowchart of Figure 2, the algorithm enters the optimization stage at each step after receiving the user feedback. The position of each particle is a vector that shows the importance degree or weight of different features. At the beginning, the weights of each particle are obtained randomly and will eventually be normalized. Proposing a suitable objective function, the particle group moves towards new images in the search space. Table 1 describes the setting used for hyper-parameters of swarm optimization.

Table 1. Values of hyper-parameters of PSO algorithm.

Hyper-parameter	Value
Population size	30
Particle size	5
$\boldsymbol{V_{max}}$	1
$\boldsymbol{\omega}$	0.7
$\boldsymbol{\eta_1}$	2
$\boldsymbol{\eta_2}$	2

The calculation of the best position of each particle and the globally best general position uses the fitness function of Equation 13, in which $ft(X_i)$ is referred to the fitness of the ith particle with position X_i, N_{rel}^k (N_{ir}^k) being the number of relevant (irrelevant) images in the kth iteration of RF-based retrieval system (a maximum of 20 iteration is considered), and Im_{req} (Im_{rel}, Im_{ir}) being the requested image (rlth relevant image/irth irrelevant image).

$$ft(X_i) = \frac{1}{N_{rel}^k} \sum_{r=1}^{N_{rel}^k} Dist(\text{Im}_q, \text{Im}_r) - \frac{1}{N_{irr}^k} \sum_{ir=1}^{N_{irr}^k} Dist_i(\text{Im}_{req} - \text{Im}_{ir}) \tag{13}$$

$Dis_i t$ is a measure of how distant the features of an image are from the feature in the requested image according to measures of particle i. We use the weighted mean absolute error for this criterion, which is addressed in Equation 14.

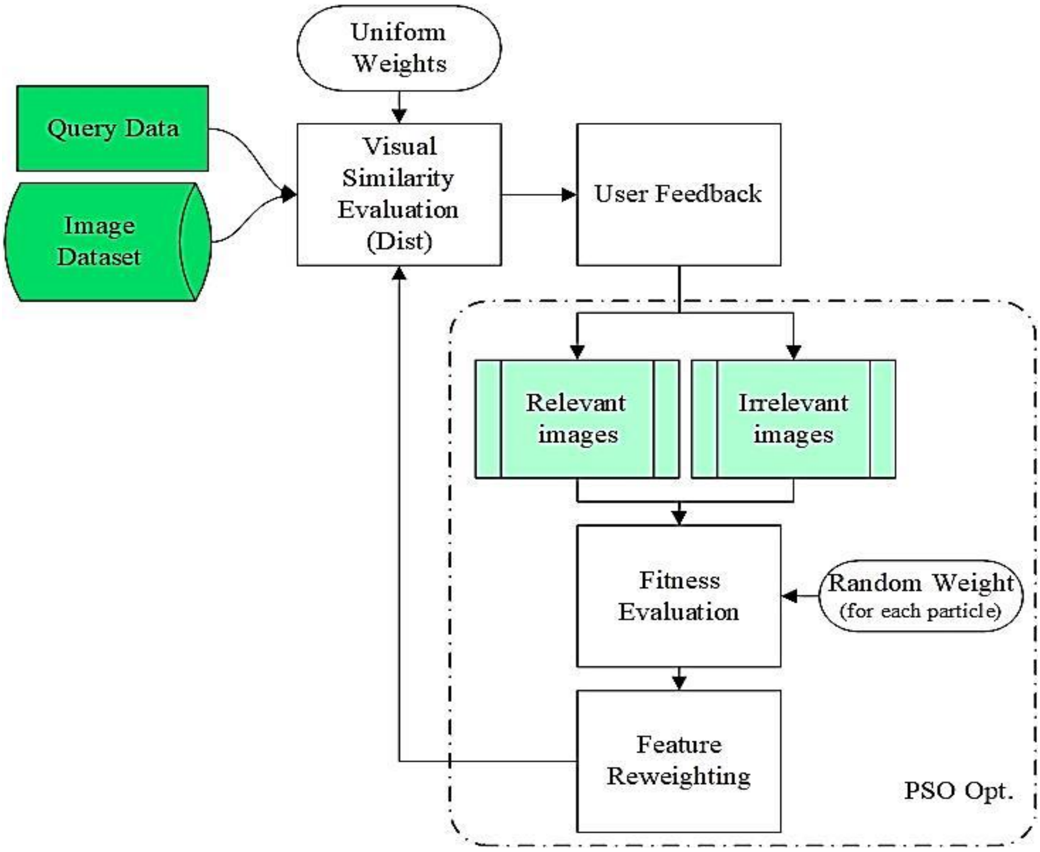

Figure 2. A flowchart of the overall proposed procedure.

$$Dist_i(\mathrm{Im}_{req}, \mathrm{Im}_j) = \sum_{f=1}^{F} |\mathrm{Im}_{req}^f - \mathrm{Im}_j^f| . X_i^f \qquad (14)$$

where the term F illustrates the number of feature sets, Im_{req}^f (Im_j^f) explains the value of the fth feature of Im_{req} (Im_j), and X_i is the position of the ith particle that shows the weight of the fth feature in the kth iteration, $w^{k,f}$.

In other words, similarity measurement is done implicitly in the fitness function of Equation 13. Actually, positions of particles of PSO in a 5D space shows the weight of importance of each one of the 5 feature sets. These weights are used to compute the weighted mean absolute error or distance of query image from other relevant/irrelevant images in Equation 14. According to Equation 13, a better particle with better weights creates a smaller distance between the relevant images and a larger distance in the irrelevant images.

Choosing this fitness function, in addition to the simple operators used in it, is also some kind of innovation of this work that will decrease the computation complexity. This function imprecisely looks for images that have the maximum distance from the set of unrelated images and the minimum distance from the set of related images. However, in some references, e.g. [22], a precise computation of the relevant/irrelevant images complicates the fitness function computations.

The global best position of particles of each run is calculated from Equation 15. Using this value, images with the lowest total weighted distance will be announced as relevant.

$$g^k = \arg\min_{i=1}^{|particles|} \{ft(p_i)\} \qquad (15)$$

Thereafter, the relevant images are shown to the user, and by obtaining the user feedback, all of this procedure will iterate. A termination condition for these iterations is to retrieve a preset number of relevant images (16 images), all of which satisfy the user or to reach a specified number of iterations (20 iterations).

4. Results and Evaluation

In order to evaluate the proposed procedure, we compare the proposed semi-automatic procedure with two other semi-automatic retrieval systems [22-24] and some automatic retrieval systems [13, 14, 16] that published their result on the Wang dataset [33] and separately for different categories of it. This dataset is a well-known one that selects 10 very discriminative classes of the Corel data set non-automatically. It contains diverse images of 10 classes with 100 images in each category, namely African people, beaches, buildings, buses, dinosaurs, elephants, flowers, horses, mountains, and food. These images are JPEG with a resolution of 384 × 256. As these datasets are well-classified, it is possible to quantitatively evaluate and compare the performance. In order to homogenize the size of the feature vectors in the images, at the processing time, all images were resized to 64 × 64. Moreover, in order to homogenize the effects of different properties, normalization was performed on each kind of feature so that the range of values of all the characteristics is the same. In the following, we will begin by providing an example of retrieval of a sample image. Then we compare the proposed procedure with some other image retrieval systems in terms of the accuracy and speed of image recovery.

4.1. An Example of Retrieval Procedure

In order to retrieve the related images, an image is firstly selected by the user as a requested image. The algorithm will retrieve similar images from the dataset. Figure 3 shows the first run of the retrieval procedure that explicitly restores similar images according to the properties of the database images and a uniform weight vector.

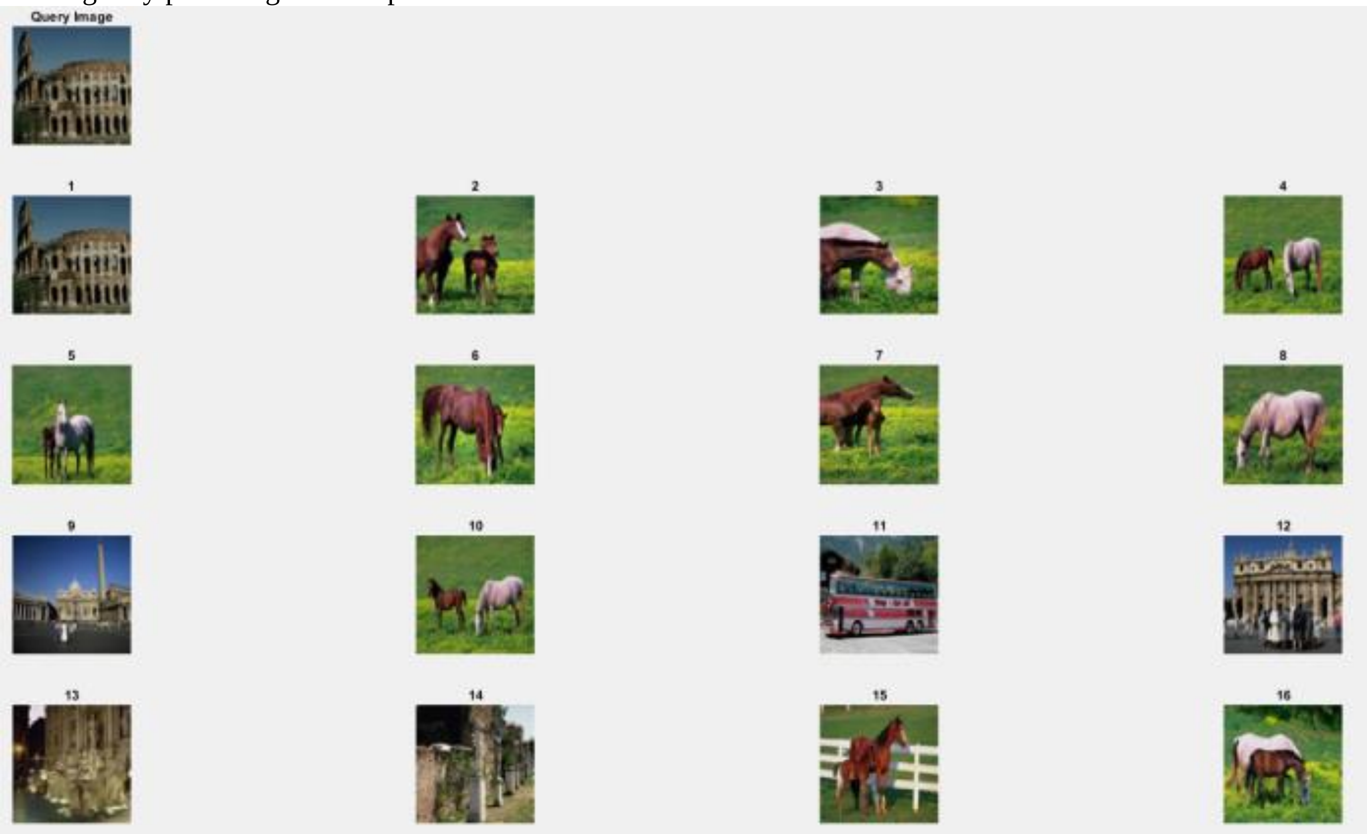

Figure 3. Retrieved images in the first run of the algorithm before using the relevance feedback. The query image is shown in the upper-left block of the image. The 16 retrieved images are shown in the underlying blocks.

After displaying the top-16 images that have minimum distances from the requested image, the related and unrelated images will be obtained from the user feedback. Thereafter, the particle optimization algorithm will update the weight of the features. This weight update will be such as to minimize the distance between the relevant images maximize the distance between irrelevant images from the requested image using the information about the relatedness of the shown images provided by the user feedback. By keeping the related images of the previous phase, the particles will find other candidate relevant images that will further be displayed to the user in order to get feedback.

Applying the particle optimization algorithm on Figures 3 shows uncertainly of the new related images that are highly dependent on the calculated weight. The procedure will continue to iterate until the user announces that all images are related or iteration 20 times. In the case of Figure 3, the related retrieved images after 11 repetitions are shown in Figure 4.

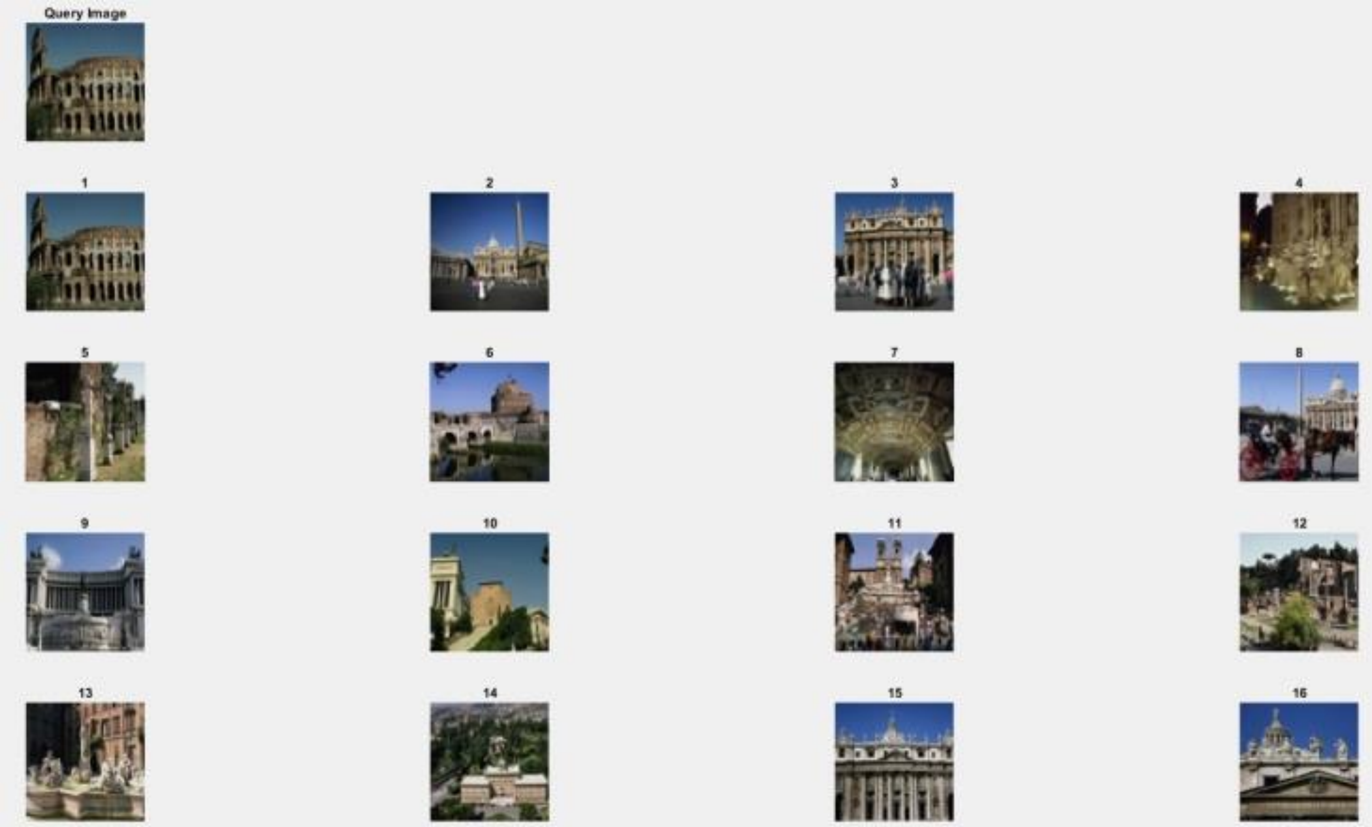

Figure 4. Final retrieved images for the query image after 11 iterations of the proposed algorithm. The query image and retrieved images are shown in the upper left block and the other bellowing blocks, respectively.

4.2. Performance Comparison

The results of comparing the performance of the proposed method with some other automatic or semi-automatic image retrieval systems on Wang dataset [33] are shown in Figures 5 and 6 and Tables 2 and 3. The comparisons are obtained from the number of iterations required to reach the 16 related images, the response time of each procedure, and the precision of each method, which are separately reported for 10 classes of images in the Wang dataset.

Figure 5 compares the proposed method with three other relevance feedback-based methods [22-24] in term of the precision of the system (maximum = 20 iterations).

As shown in Figure 5, precision of the image retrieval by the proposed algorithm is better for some classes, e.g. beach, buildings, elephants, horses, and mountains than the method of [22] , and it is completely better than the method of [24] and [23].

In order to further investigate the effect of direct feature weighting against feature reweighting in PSO-based image retrieval paradigm, we made another comparison with the algorithm of [22] in terms of the computational time; the results obtained are shown in Figure 7.

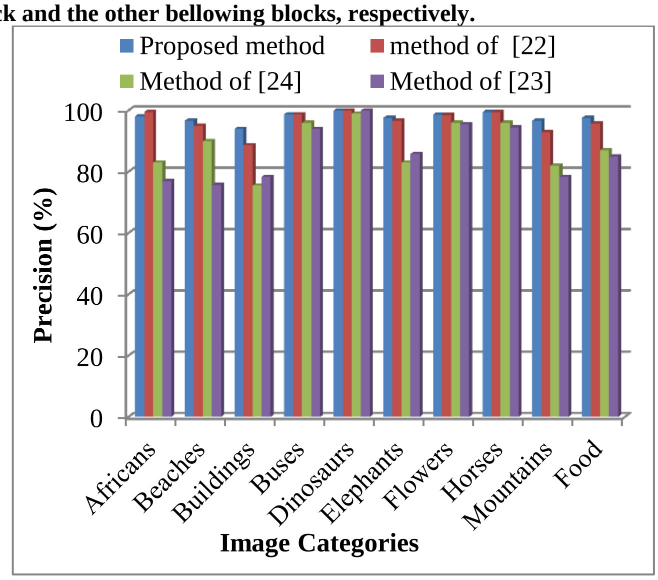

Figure 5. Comparison between the proposed method and methods of [22, 24] in term of precision of the retrieval system for retrieving each one of the 10 different image categories of the Wang dataset.

As shown in this figure, for most of the classes, the elapsed time for calculations in the proposed method is fewer. This reduction in the elapsed time is obtained due to the reduced computational complexity in the evaluation function used for weight assessment. For example, in order to retrieve the building images, the precision and timing of the proposed algorithm are better than the method [22].

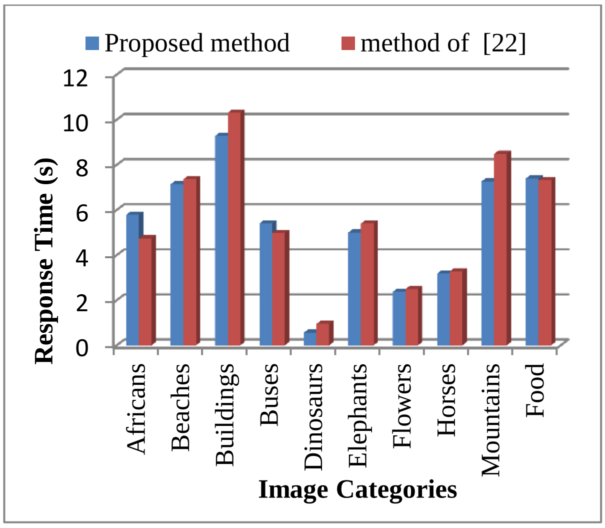

Figure 6. Comparison between the proposed method and method of [22] in terms of the response time for retrieving different image categories.

Since the semi-automatic image retrieval task is based on user feedback, and the algorithm iterates and improves its evaluation criteria until the user satisfies, the number of necessary iterations can be considered as another important criterion; as more iteration shows less accuracy. This criterion is addressed in Table 2 alongside the two other criteria of precision and response time.

Comparison of the proposed method with the method of the two automatic retrieval systems [13, 16] is reported in Table 3. Comparison of the proposed method against [13, 16] shows the strength of the proposed method versus these methods. On the other hand, when the dataset provides more visual information of its images, as in [14], the results of automatic CBIR will be better than the other CBIR approaches. The results of the proposed method are compared with those of another article on the famous Corel data set, which can be seen in Table 4. The second article was an automatic method. As shown in this table, this method could have a better precision and recall compared to the second article. Table 5 generally compares the proposed method with the other methods.

In case of application of deep learning in CBIR, we review some papers in Section 1. Priyanka [15] reports application of its deep learning method on the Wang dataset. The average precision of its algorithm is 93.457, which is a good result for an automatic CBIR approach. However, this precision is still less than the precision of the proposed method and method of [15] (with precisions of 97.761 and 96.525, respectively). Comparison of the semi-automatic and automatic CBIR methods shows that semi-automatic approaches are most reliable for high accurate tasks.

Moreover, the results of the proposed method are compared with another automatic image retrieval method [40] on the Corel-10K data set, which can be seen in Table 4. As shown in this table, the proposed method has a better precision and recall compared to [40]. Table 5 summarizes all the results.

Table 2. Comparing the proposed algorithm and the methods of [22, 24] in terms of precision, response time, and average number of iterations to retrieve relevant images.

Different classes of Wang dataset		Africans	Beaches	Buildings	Buses	Dinosaurs	Elephants	Flowers	Horses	Mountains	Food	Average
Precision	Proposed method	98.12	**96.69**	**94**	98.75	100	97.62	**98.62**	99.5	**96.69**	**97.62**	**97.761**
	method of [22]	**99.5**	95	88.62	98.75	100	96.69	98.5	99.5	92.94	95.75	96.525
	Method of [24]	83	90	75.5	96	99	83	96	96	82	87	88.75
	Method of [23]	77.0	75.7	78.3	94	100	85.8	95.5	94.5	78.35	85	86.4
Time	Proposed method	5.7873	**7.1397**	**9.2792**	5.4020	**0.5616**	**4.9975**	**2.3658**	**3.1777**	**7.2614**	7.3933	**5.33**
	method of [22]	4.7391	7.3613	10.3095	4.9783	0.9500	5.3977	2.4869	3.2733	8.4842	7.3217	5.53
Avg. num. of iterations	Proposed method	9.9900	12.8600	**17.0900**	9.9500	**1.0700**	**9.2300**	4.3900	5.8900	**13.9000**	14.1900	9.856
	method of [22]	8.3100	12.8500	18.1700	8.7500	1.7200	9.4300	4.3800	5.6900	14.6300	12.5600	**9.649**

Table 3. Comparing precision of the proposed method with two automatic retrieval systems.

Classes	Africans	Beaches	Buildings	Buses	Dinosaurs	Elephants	Flowers	Horses	Mountains	Food	Average
Proposed	**98.12**	**96.69**	**94**	**98.75**	**100**	**97.62**	**98.62**	**99.5**	**96.69**	**97.62**	**97.761**
Method of [13]	65	60	62	85	93	65	94	77	73	81	75
Method of [16]	56	54	61	89	98	58	90	78	51	68	70
Method of [14]	81	79	91	76	99.7	96.8	94.5	84.1	**97.4**	91.5	89.6

Table 4. Comparing the precision and recall of the proposed method with [40] on Corel-10k.

Method	Precision	Recall
Proposed method	64.46	10.31
Method of [40]	56.88	6.83

Table 5. Summarization of comparing precision of the proposed method with other methods on different datasets.

Data set	[13]	[14]	[16]	[22[]	[23]	[24]	[40]	proposed
Wang	75	89.6	70	96.5	86.4	88.75	--	**97.8**
Corel-10	--	--	--	--	--	--	56.88	**64.46**

5. Conclusion

In this work, we considered the issue of the content-based image retrieval using the relevance feedback. The features used to evaluate the similarity of the requested image with retrieved images included color histogram, first-to-third-color moments, image edges, and 2D wavelet transform, representing the texture. Fusion of these non-homogenous features requires a feature weighting algorithm. For a proper weighting, particle optimization algorithm was used, which could be referred to as article innovation. Information was extracted from the user feedback guides particles in weighting features to minimize (maximize) the distance between the user selected relevant (irrelevant) images and the query image.

The results obtained indicate a less computational complexity of the proposed method with a comparable accuracy to the other PSO-based methods. Moreover, the proposed algorithm has a higher average precision than the other RF-based CBIR methods. Comparison with automatic CBIR approaches including deep learning methods shows that the approaches are most reliable for high accurate tasks.

References

[1] K. Guo, R. Zhang, Z. Zhou, Y. Tang, and L. Kuang, "Combined Retrieval: A Convenient and Precise Approach for Internet Image Retrieval," *Information Sciences,* vol. 358-359, pp. 151-163, 2016/09/01/ 2016, doi: https://doi.org/10.1016/j.ins.2016.04.001.

[2] A. S. Hsu W1, Long LR, Neve L, Thoma GR., "SPIRS: a Web-Based Image Retrieval System for Large Biomedical Databases," *International Journal of Biomedical Informatics,* Vol. 78, pp. 13-24, 2009.

[3] C. Huang, H. Xu, L. Xie, J. Zhu, C. Xu, and Y. Tang, "Large-Scale Semantic Web Image Retrieval Using Bimodal Deep Learning Techniques," *Information Sciences,* Vol. 430-431, pp. 331-348, 2018/03/01/ 2018, doi: https://doi.org/10.1016/j.ins.2017.11.043.

[4] R. Lan, H. Wang, S. Zhong, Z. Liu, and X. Luo, "An Integrated Scattering Feature with Application to Medical Image Retrieval," *Computers & Electrical Engineering,* Vol. 69, pp. 669-675, 2018/07/01/ 2018, doi: https://doi.org/10.1016/j.compeleceng.2018.01.027.

[5] Yogesh Kumar, Ashutosh Aggarwal, and S. Tiwari, "An Efficient and Robust Approach for Biomedical Image Retrieval using Zernike Moments," *Biomedical Signal Processing and Control,* Vol. 39, pp. 459-473, 2018.

[6] A. khodaskar and S. Ladhake, "New-Fangled Alignment of Ontologies for Content Based Semantic Image Retrieval," *Procedia Computer Science,* Vol. 48, pp. 298-303, 2015/01/01/ 2015, doi: https://doi.org/10.1016/j.procs.2015.04.185.

[7] M. K. Sharma and T. J. Siddiqui, "An Ontology Based Framework for Retrieval of Museum Artifacts," *Procedia Computer Science,* Vol. 84, pp. 169-176, 2016/01/01/ 2016, doi: https://doi.org/10.1016/j.procs.2016.04.083.

[8] V. K. G. P. Shamna, K.A. Abdul Nazeer,, "Content based Medical Image Retrieval using Topic and Location Model," *Journal of Biomedical Informatics,* Vol. 91, 2019,.

[9] S. Zhang, Q. Tian, G. Hua, Q. Huang, and W. Gao, "ObjectPatchNet: Towards scalable and semantic image annotation and retrieval," *Computer Vision and Image Understanding,* Vol. 118, pp. 16-29, 2014/01/01/ 2014, doi: https://doi.org/10.1016/j.cviu.2013.03.008.

[10] R. W. Daiguo Deng, Hefeng Wu, Huayong He, Qi Li, Xiaonan Luo, "Learning Deep Similarity Models with Focus Ranking for Fabric Image Retrieval," *Image and Vision computing,* Vol. 70, pp. 11-20, 2018.

[11] M. M. Mohammed, A. Badr, and M. B. Abdelhalim, "Image Classification and Retrieval using Optimized Pulse-Coupled Neural Network," *Expert Systems with Applications,* Vol. 42, No. 11, pp. 4927-4936, 2015/07/01/ 2015, doi: https://doi.org/10.1016/j.eswa.2015.02.019.

[12] M. Tzelepi and A. Tefas, "Deep convolutional learning for Content Based Image Retrieval," *Neurocomputing,* Vol. 275, pp. 2467-2478, 2018/01/31/ 2018, doi: https://doi.org/10.1016/j.neucom.2017.11.022.

[13] A. Irtaza, A. Jaffar, E. Aleisa, and T.-S. Choi, "Embedding Neural Networks For Semantic Association in Content Based Image Retrieval," *Multimedia Tools and Applications,* Vol. 72, 09/01 2014, doi: 10.1007/s11042-013-1489-6.

[14] A. Sarwar, Z. Mehmood, T. Saba, K. A. Qazi, A. Adnan, and H. Jamal, "A Novel Method for Content-Based Image Retrieval to Improve the Effectiveness of the Bag-Of-Words Model Using a Support Sector Machine," *Journal of Information Science,* Vol. 45,

No. 1, pp. 117-135, 2019/02/01 2018, doi: 10.1177/0165551518782825.

[15] S. Priyanka, "Microstructure Pattern Extraction Based Image Retrieval," *Multimedia Tools and Applications*, Vol. 79, No. 3, pp. 2263-2283, 2020/01/01 2020, doi: 10.1007/s11042-019-08113-y.

[16] M. Rao, B. Rao, and D. Govardhan, "CTDCIRS: Content based Image Retrieval System based on Dominant Color and Texture Features," *International Journal of Computer Applications*, Vol. 18, 03/31 2011, doi: 10.5120/2285-2961.

[17] R. S. M. Premkumar, "Interactive Content Based Image Retrieval using Multiuser Feedback," *International Journal on Informatics Visualization*, Vol. 1, No. 4, pp. 165 - 169, 2017.

[18] N. M. A. J. Li, "Relevance Feedback in Content-Based Image Retrieval: A Survey," in *Handbook on Neural Information Processing*, Vol. 49. Berlin, Heidelberg: Springer, 2013, pp. 433-469.

[19] M. M. Rahman, S. K. Antani, and G. R. Thoma, "A Learning-Based Similarity Fusion and Filtering Approach for Biomedical Image Retrieval Using SVM Classification and Relevance Feedback," *IEEE Transactions on Information Technology in Biomedicine*, Vol. 15, No. 4, pp. 640-646, 2011, doi: 10.1109/TITB.2011.2151258.

[20] A. Marakakis, G. Siolas, N. Galatsanos, A. Likas, and A. Stafylopatis, "Relevance Feedback Approach for Image Retrieval Combining Support Vector Machines and Adapted Gaussian Mixture Models," *IET Image Processing*, Vol. 5, No. 6, pp. 531-540, 2011, doi: 10.1049/iet-ipr.2009.0402.

[21] M. Cai-Hong, D. Qin, and L. Shi-Bin, "A Hybrid PSO and Active Learning SVM Model for Relevance Feedback in the Content-Based Images Retrieval," in *2012 International Conference on Computer Science and Service System*, 11-13 Aug. 2012 2012, pp. 130-133, doi: 10.1109/CSSS.2012.40.

[22] M. Broilo and F. G. B. D. Natale, "A Stochastic Approach to Image Retrieval Using Relevance Feedback and Particle Swarm Optimization," *IEEE Transactions on Multimedia*, Vol. 12, No. 4, pp. 267-277, 2010, doi: 10.1109/TMM.2010.2046269.

[23] F. Baig *et al.*, "Boosting the Performance of the BoVW Model Using SURF–CoHOG-Based Sparse Features with Relevance Feedback for CBIR," *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, Vol. 44, No. 1, pp. 99-118, 2020/03/01 2020, doi: 10.1007/s40998-019-00237-z.

[24] A. Latif *et al.*, "Query-Sensitive Similarity Measure for Content-based Image Retrieval Using Meta-Heuristic Algorithm," *Journal of King Saud University - Computer and Information Sciences*, Vol. 30, No. 3, pp. 373-381, 2018/07/01/ 2018, doi: https://doi.org/10.1016/j.jksuci.2017.05.002.

[25] A. Ameer and K. S. Kumar, "Efficient Automatic Image Annotation Using Optimized Weighted Complementary Feature Fusion Using Genetic Algorithm," *Procedia Computer Science*, Vol. 58, pp. 731-739, 2015/01/01/ 2015, doi: https://doi.org/10.1016/j.procs.2015.08.094.

[26] M. Arevalillo-Herráez, F. J. Ferri, and S. Moreno-Picot, "Improving Distance Based Image Retrieval Using Non-Dominated Sorting Genetic Algorithm," *Pattern Recognition Letters*, Vol. 53, pp. 109-117, 2015/02/01/ 2015, doi: https://doi.org/10.1016/j.patrec.2014.05.008.

[27] Z. Shoaie and S. Jinni, "Semantic image retrieval using relevance feedback and reinforcement learning algorithm," in *2010 5th International Symposium On I/V Communications and Mobile Network*, 30 Sept.-2 Oct. 2010 2010, pp. 1-4, doi: 10.1109/ISVC.2010.5654900.

[28] L. Pinjarkar, M. Sharma, and S. Selot, "Deep CNN Combined With Relevance Feedback for Trademark Image Retrieval," *Journal of Intelligent Systems*, 09/15 2018, doi: 10.1515/jisys-2018-0083.

[29] X. Heng, W. Jun-yi, and M. Lei, "Relevance Feedback for Content-based Image Retrieval using Deep Learning," in *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, 2-4 June 2017 2017, pp. 629-633, doi: 10.1109/ICIVC.2017.7984632.

[30] C. K. L. Banerjee, A. E. Devorah, B. Do, D. L. Rubin, C. F. Beaulieu, "Relevance Feedback for Enhancing Content based Image Retrieval and Automatic Prediction of Semantic Image Features: Application to Bone Tumor Radiographs," *J Biomed Inform*,Vol. 84, pp. 123-135, 2018.

[31] A. Sleit, A. Abu Dalhoum, M. Qatawneh, M. Al-Sharief, R. a. Al-Jabaly, and O. Karajeh, "Image Clustering using Color, Texture and Shape Features," *KSII Transactions on Internet and Information Systems*, Nol. 5, pp. 211-227, 01/21 2011, doi: 10.3837/tiis.2011.01.012.

[32] A. A. Fahimeh Alaei, Umapada Pal, Michael Blumenstein, "A Comparative Study of Different Texture Features For Document Image Retrieval," *Expert Systems with Applications*, Vol. 121, pp. 97-114, 2019.

[33] J. Z. Wang, L. Jia, and G. Wiederhold, "SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23,No. 9, pp. 947-963, 2001, doi: 10.1109/34.955109.

[34] A. R. Y. Betül Sultan Yıldız, "Comparison of Grey Wolf, Whale, Water Cycle, Ant Lion and Sine-Cosine Algorithms For The Optimization of a Vehicle Engine Connecting Rod," *Materials Testing*, Vol. 60, No. 3, pp. 311-315, 2018.

[35] A. Yildiz, "A New Hybrid Artificial Bee Colony Algorithm for Robust Optimal Design and Manufacturing," *Applied Soft Computing*, Vol. 13, pp. 2906–2912, 05/01 2013, doi: 10.1016/j.asoc.2012.04.013.

[36] A. Yildiz, "A Novel Hybrid Immune Algorithm for Global Optimization in Design and Manufacturing," *Robotics and Computer-Integrated Manufacturing*, Vol. 25, pp. 261-270, 04/01 2009, doi: 10.1016/j.rcim.2007.08.002.

[37] X. Meng and Z. Pian, "Theoretical Basis for Intelligent Coordinated Control," in *Intelligent Coordinated Control of Complex Uncertain Systems for Power Distribution Network Reliability*, X. Meng and Z. Pian Eds. Oxford: Elsevier, 2016, Ch. 2, pp. 15-50.

[38] A. Yildiz and B. Yıldız, "The Harris hawks optimization algorithm, salp swarm algorithm, grasshopper optimization algorithm and dragonfly algorithm for structural design optimization of vehicle components," *Materialprufung*, Vol. 8, pp. 60-70, 08/03 2019, doi: 10.3139/120.111379.

[39] A. Yildiz and K. Solanki, *Multi-Objective Optimization Of Vehicle Crashworthiness Using Paticle Swarm Optimization Approach*. 2012.

[40] Z. Wei, G. Liu, "Image Retrieval Using the Intensity Variation Descriptor," *Mathematical Problems in Engineering*, 2020.