ISSN 2515-8260 Volume 07, Issue 09, 2020

WEIGHT BASED MULTI-FEATURE FUSION (WMFF) VIA IMPROVED ARTIFICIAL BEE COLONY (IABC) AND RE-RANKING WITH CLICK-BASED SIMILARITY FOR WEB IMAGES

Sree Rajeswari.M, *B.E,M.E.,(Ph.D),Assistant Professor, Department of Information Technology Dr.N.G.P.Institute of Technology,Coimbatore-641048.*
Dr.M.Rajalakshmi , *B.E,M.E.,Ph.D,Associate Professor, Department of Information Technology Coimbatore Institute of Technology,coimbatore-641014*

ABSTRACT: In image search re-ranking, a major problem restricting the image retrieval development is a intent gap, which is a gap between user's real intent and query/demand representation, besides well-known semantic gap. In the past, for achieving effective web image retrieval, classifier space or feature space is explored at a time by researchers. Visual information and images initial ranks with single feature are only considered in conventional re-ranking techniques for measuring typicality and similarity in web image retrieval, while overlooking click-through data influence. For image retrieval, various image features aggregation shows its effectiveness in recent days. But, uplifting the best features impact for a specific query image presents a major challenge in computer vision problem. In this paper, Weight based Multi-Feature Fusion (WMFF) is fused by Improved Artificial Bee Colony (IABC) for presenting a re-ranking algorithm to retrieve web image. Based on web query, features are assigned with weights, where different weights are received by different queries in ranked list. IABC algorithm used to compute weights is a data-driven algorithm and it does not require any learning. At last, in a web, color and texture features are fused using fusion and these features are extracted with respective modalities. A Semi-supervised Consensus Clustering re-ranking with click-based similarity and typicality procedure termed as SCCCST is used in re-ranking technique. Convolutional Neural Network (CNN) classifier with Multiple Kernel Learning (CNN-MKL) is used here for performing click-based similarity. Its operation is depends on selection of click-based triplet's and a classifier is used for integrating multiple features into a unified similarity space. The web image search re-ranking performance is greatly enhanced using proposed technique.

Keywords: Click through data, re-ranking, Semi-supervised Consensus Clustering (SCC) , Convolutional Neural Network (CNN) classifier with Multiple Kernel Learning (CNN-MKL), Weight based Multi-Feature Fusion (WMFF), Improved Artificial Bee Colony (IABC), Texture Feature, Color Feature and Keyword-Based Search.

1. INTRODUCTION

In various fields, huge images can be accessed as well as managed easily using the rapid growth in technology and web environment. On web, comprehensive image collection retrieval is an important research and industrial issue. Characteristics of typical Content Based Image Retrieval (CBIR) systems and web image retrieval are different. Users can search images using keywords interface and/or via query using web image search systems. In general, there are text annotations in web images, which can be obtained from web pages, where images are stored.

Images text information are used in conventional web image retrieval systems as well as in text (keyword) retrieval systems. Texts and general image information like non-graph/graph, image format, image size are used in some systems and for relevance feedback, user interface input are provided by some other system. In various situations, for large-scale image collections, efficiency and effectiveness of text-based image search technique is proved. But, it has some drawbacks like incapability of associated text in describing image content.

For enhancing search performance, proposed an image search re-ranking, where initial ranking orders are adjusted via auxiliary knowledge leveraging or visual content mining. In recent days, it has gained the attention in industry and academia [1], [2]. For overcoming "semantic gap", which is a gap between high-level and low-level feature semantics, visual information is used by most of the available re-ranking techniques in a passive and unsupervised manner [3].

Useful visual information are mined further using multiple visual modalities [4], but, there will be a limited enhancement in performance. This is due to the fact that, "intent gap" which is a gap between user's real intent and users' query/demand representation is neglected in these re-ranking techniques. And also, label information is also not considered in this re-ranking paradigm and only limited enhancement in performance only achieved.

For rectifying this issue, proposed Relevance Feedback (RF) technique for acquiring user's search intention to further enhancing retrieval performance. Other than this, a user term feedback technique is proposed by Zhang et al. [7] for refining retrieved images. However, in Text Based Information Retrieval (TBIR), feedback is not effective as mentioned by them.

However, computation of sufficient as well as explicit feedback from user is very difficult since, reluctant feedbacks are not provided by the users to search engines very often. User issued queries and respective clicked images can be recorded in search engines. On specific query image pairs relevance, explicit user preference are not reflected by the clicked images with its respective queries.

In information retrieval area, click-through data is having wide use [8-9]. Beyond the this fact, image thumbnails are browse by users in image search before choosing images to click and decision to click, which are based on the image relevance. So, click through data is assumed as "implicit" user feedback with hypothesis that, highly clicked images are having relevance with specified query. As user search behaviours footprints, implicit relevance feedback of user, cannot only be provided by click-through data. It is available readily and can be accessed freely by search engines [10].

Assumptions are widely accepted and generally applied methods of most image search re-ranking techniques. In a ranking list, there will be less distance between visually similar images and high rank will be given to images with high relevance. In order to produce satisfying re-ranking results, image typicality which is an image relevance level and image similarity are the highly important factors [11-13].

Visual information and images initial ranks with single feature are only considered in conventional re-ranking techniques for measuring typicality and similarity in web image retrieval, while overlooking click-through data influence. A Semi-supervised Consensus Clustering re-ranking with click-based similarity and typicality (SCCCST) is entirely adopted in this proposed re-ranking technique for guiding image typicality and similarity learning.
For re-ranking, proposed a novel image similarity learning algorithm called Click-based Multi Weight Feature Fusion Similarity Learning (CMWFFSL) via Convolutional Neural Network with Multiple Kernel Learning (CNN-MKL). For multiple features, after learning similarity metric for grouping semantically and visually similar images as same clusters proposed a Semi-supervised Consensus Clustering (SCC). Cluster typicality and within-clusters image typicality in descending order are computed for computing final re-rank list.

2. LITERATURE REVIEW

Related works are categorized in two dimensions: search with click-through data and visual search re-ranking.
Tian et al [6] implemented a local-global discriminative dimension reduction algorithm, where discriminative information and local geometry of labelled images are transferred into entire image database for learning submanifold. Proposed active re-ranking scheme's effectiveness is demonstrated via experimentation on real web image search dataset and synthetic datasets. The proposed scheme includes local-global discriminative dimension reduction algorithm and active sample selection technique based on structural information.
Zhang et al [10] leveraged click-through data, which is "implicit" user feedback for understanding the query in an effective manner. In this work proposed a novel reranking algorithm, termed as click-based relevance feedback. User search intention are identified in this algorithm by emphasizing click-through data successfully. It also learns query-dependent fusion weights adaptively for multiple modalities by leveraging multiple kernel learning algorithms. Around 11.62% initial search results are enhanced by proposed re-ranking technique as shown in results of experimentation and for most kind of queries like top, middle and tail queries, various existing techniques are outperformed by this technique.
Liu and Mei [14] stated that globally optimum ranked list will not be produced if classification performance is not good enough. They formulated an optimization problem from re-ranking, where, if any two arbitrary documents in the list are ranked properly based on relevance, an optimum ranked list will be produced. For computing relevant relation of every pairwise re-ranking, introduced two pairwise re-ranking techniques called, Exclusion Pairwise reranking (EP-reranking) and Difference Pairwise reranking (DP-reranking). At last, final ranked list is recovered by exploring Round Robin criterion. Over text search baseline and other re-ranking techniques, consistent enhancement is shown using this technique.
Duan et al [15] implemented a Generalized Multi-Instance (GMI) setting for this application. The retrieval performance is enhanced using developed GMI-SVM classifier, where, labels from bag level are propagated to instance level. All the bags are ranked using proposed bag ranking technique based on defined bag ranking score.
Top ranked bags are utilized as pseudo positive training bags, where few irrelevant images which are not having any association with texture query are sampled randomly for computing pseudo negative training bags. With manually labelled training bags which are derived from relevance feedback, better performance can be achieved using proposed GMI-SVM as demonstrated in experimentation.

San Pedro et al [16] reviewed the visual aesthetic quality influence in search results as a complementary attribute of relevance. Introduced a new ranking parameter based on considered aesthetics, which is used for enhancing top ranks quality with the availability of huge relevant results. Proposed two techniques for aesthetic rating inference: On is visual content based, another one is user comments analysis based for detecting opinions about images quality.

Wang et al [17] explored the multiple modalities in graph-based learning technique for introducing web image search re-ranking technique. Single modality or multiple modalities are implemented by integrating long feature vector for integrating relevance scores, modalities weights and distance metrics effectively and every modality is scaled into a unified technique. Robustness of proposed technique when compared with every individual modality is shown using an experimental results and better performance is exhibited using this re-ranking technique.

Umesh [18] implemented an image retrieval system based on semantic for retrieving set of relevant images for specified query image from web. Visual dictionary is generated using Dense SIFT feature extraction technique and Global color space model via Self Organizing Maps (SOM) clustering algorithm. Distance between input image and images set in image database are measure using histogram intersection technique for retrieving similar images. For demonstrating proposed system's effectiveness, over a collection of 1000 generic Web images, experimentation is carried out.

Cai et al [19] exploited semantic attributes for image search re-ranking. For all pre-set attributes, according to the classifier, every image is represented by an attribute feature with response from all classifiers. For exploring two information sources simultaneously, proposed a Visual-attribute joint hypergraph learning technique. Relationship between images are modelled by constructing hypergraph. The joint hypergraph learning technique's effectiveness is demonstrated using experimental results.

Ji et al [20] introduced a ranking function termed as hypersphere-based rank (H-Rank), feature extraction algorithm termed as hypersphere-based relevance preserving projection (HRPP). In specific, an original high-dimensional feature space is transformed into an intrinsically low-dimensional hypersphere space using anspectral embedding algorithm called HRPP via preservation of manifold structure and relevance relationship among images. The HRPP technique with reversed K-Nearest Neighbor (KNN) is termed as one-click-based HRPP (OC-HRPP). Effectiveness of proposed algorithm is shown by the extensive experimentation results on three huge-real world dataset. This technique requires a labelling of only one relevant image.

Yang et al [21] explored click-through data, which is used as a user searching behavior's footprint, which is used for understanding the query effectively and to provide recurrent patterns identification basis, which can be used in re-ranking. Retrieval is done using this Click-boosting multi-modality graph-based reranking.

Similar images which are not clicked are located by leveraging clicked images in this algorithm and they are re-ranked in a multi-modality graph-based learning technique. On a real-world image dataset, which is collected from a commercial search engine with click-through data, this work reports encouraging results.

Yang et al [22] introduced a novel re-positioning methodology, named spectral clustering re-positioning with click-based likeness and regularity. In the first place, to gain proficiency with a fitting closeness estimation, click-based multi-include similitude learning calculation is proposed which conducts metric learning dependent on click-based triplets choice, and coordinates various highlights into a bound together comparability space through numerous piece learning. At that

ISSN 2515-8260 Volume 07, Issue 09, 2020

point, in view of the educated snap based picture comparability measure, lead ghostly bunching to assemble outwardly and semantically comparative pictures into same groups, and get the last re-rank rundown by figuring click-based bunches averageness and inside groups click-based picture commonality in plummeting request.

Despite the fact that visual substance of pictures has been mined through different ways, Euclidean separation or cosine separation is for the most part used to gauge picture closeness in the above methodologies. The significance of picture likeness estimation is regularly neglected, also, for numerous visual modalities. Past that, client cooperation is constantly considered to explain the plan hole; in any case, clients regularly decline to give such input. Therefore, it is an intense issue to defeat plan hole without client support during the inquiry procedure.

3. PROPOSED METHODOLOGY

An effective as well as less time consumed web image retrieval process is focused in this work. In the first stage, colors are extracted from given keyword text using YUV and RGB techniques. Gray level Co-occurrence Matrix (GLCM) is used for extracting features. Then, Weight based Multi-Feature Fusion (WMFF) is used for fusing the extracted color and texture features. In WMFF algorithm, theImproved Artificial Bee Colony (IABC) is used for creating features weight.

Based on web query, weights are assigned for features, where different weights are assigned to ranked list of various queries. For re-ranking, proposed a novel image similarity learning algorithm called Click-based Multi Weight Feature Fusion Similarity Learning (CMWFFSL) via Convolutional Neural Network with Multiple Kernel Learning (CNN-MKL). For multiple features, after learning similarity metric for grouping semantically and visually similar images as same clusters proposed a Semi-supervised Consensus Clustering (SCC). Cluster typicality and within-clusters image typicality in descending order are computed for computing final re-rank list.

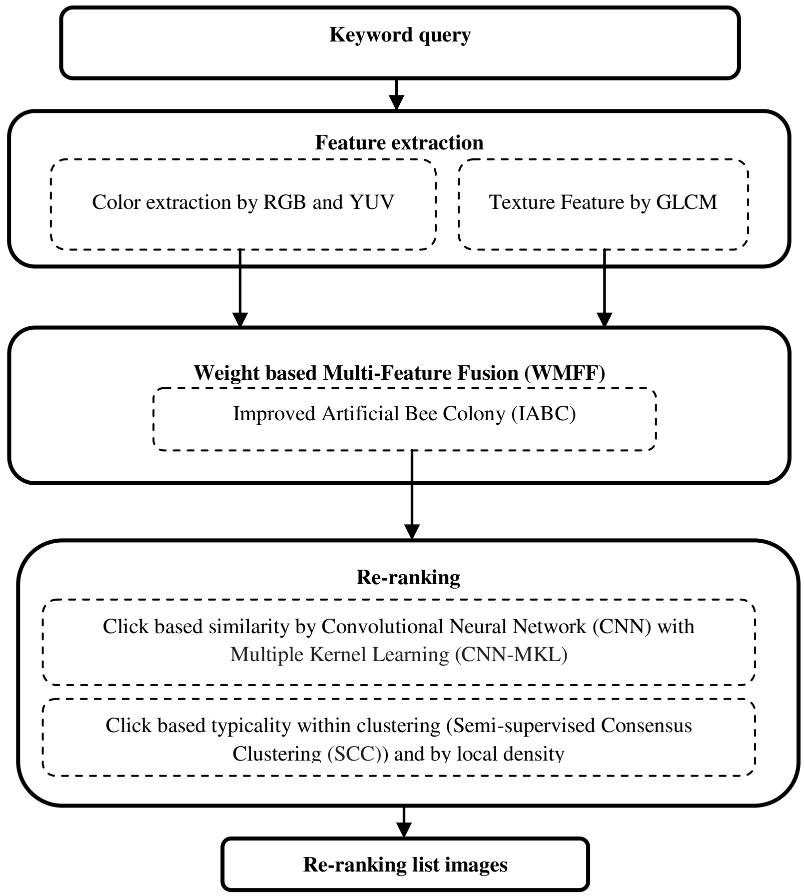

FIGURE 1.OVERVIEW OF THE PROPOSED ARCHITECTURE

A new direction has emerged forimproving the search performance and termed as image search Re-ranking. In this technique, text based searching function outcomes are reordered using visual information. From initial search function outcomes, this work considers image feature's visual characteristics, which follows the ranking function construction and for ranking, functions are used for reordering the images conclusively.

Image's visual similarity are included for performing image re-ranking and for getting better results from image and also for obtaining multiplicity of image result. In image search process, re-ranking objective is to upgrade text based search's impact which requires image content extraction, which is a visual. Re-ranking also considers visual contents of images. For confining user's goal with least human intervention, significant image search technique is proposed for searching related images via necessitation of single click by user over images, which are initially searched.

3.1.Feature extraction

In feature extraction, shape, texture and color are the three major feature descriptors applied. Here, texture and color extraction are mainly focused. Following section presents, brief discussion about it. After the specification query, this entire process takes place. Keyword type query is specified in this work.

In any content-based image retrieval approach, visual feature extraction plays a major role. In general, features includes both visual features like texture, shape, color and text based features like keywords and annotations. Further, in visual feature range, features are classified as low-level and high-level features.

3.1.1. Color feature extraction by using RGB and YUV

The RGB is most prevalent color space, which represents Red-Green-Blue. These colors are assumed as a additive primary colors of light and are included in this space. In visible spectrum, any existing colors can be created by combining these three colors in different intensity level. Space if formed according to device and perceptually, it is non-uniform. When observed by a human eye, colors might not be seen as closer as indicated by this, even they are comparatively close in RGB space.

Additive primaries are there in phosphor luminescence in a monitor and through coefficients (α, β, γ), all colors are parameterized as $C = \alpha R+\beta G+\gamma B$. This coefficents value starts from zero and it indicated luminescence presence to one representing output as full phosphor. Volume of cube is loaded by color coordinates available in parameterization with black as vertices, blue, green and red as three primaries, white, yellow, magenta and cyan as secondary fusions.

3.1.2. Texture extraction based on Gray level co-occurrence matrix

Textures are analyzed using a statistical approach called Gray-level co-occurrence matrix (GLCM) [23]. Spatial relationship between pixels are reflected using this textures. The GLCM functions are used for describing the texture of an image by computing how certain values are possessed by frequent pixel pairs and in a definite spatial relationship, appear in image, producing a GLCM, followed by arithmetical measures extracted from matrix.

A gray-level co-occurrence matrix (GLCM) is generated in MATLAB using a gray co matrix function. It computes how often a pixel with intensity I appears in specific spatial relationship with a pixel with intensity j. Spatial relationship is described as interest pixel, pixels lying in its direct right.

In consequential GLCM, every available element(i, j) is merely a aggregation of pixels count with value i, which is indicated in spatial relationship with a pixel with value j present in input image. Shape, color, Texture, Entropy, Energy, Intensity, correlation, Homogeneity and contrast features are extracted in this work. Similarity measurement is done using these features.

3.2.Weight based Multi-Feature Fusion (WMFF)

In image search re-ranking, when compared with single feature, multi-feature fusion shows its effectiveness in mining relevant recurrent visual patterns of multiple features [24]. Various aspects of image content are represented using multiple features, where same image's semantics are shared potentially. Direct computation of weights are difficult due to different scaling of similarity scores from various features.

For computing weight $we_{m,}$ an Improved Artificial Bee Colony (IABC) is described. Query-specific weights are computed using this for measuring feature's importance in specific query. For m extracted features, there exist a m Mahalanobis distance.

For a specified query q, where $q \in Q$(Qis a queries set), which represents image instance as x_i, where $x_i \in \mathbb{R}^d$(d is a image's feature dimension) and $i = 1,2,\dots,n$(n is a images count to be re-ranked). Image $x_i$can be expressed as x_i^p according to every feature p ($p = 1,2,\dots,m$). For every feature p, data is mapped into a Reproducing Kernel Hilbert Space (RKHS),$\mathcal{H}$ using feature map ϕwith corresponding kernel function. Here $\mathrm{dis}(x_i, x_j)$ is used for combining multi-feature's similarity score.

In this work, IABC algorithm is followed for generating every feature vector's weight values. Honey bees foraging process is simulated in IABC algorithm [25]. There are three classes of bees in ABC algorithm, namely, scout, onlooker and employed bees. For a user specified query image, features information are shared by three types of bees. Set of images with their food sources feature vectors with each other. According to similarity, food sources are computed by employed bees as features weight values via distance function and in live, food sources features are delivered to onlooker bees.

According to the received feature vector, employed bees are followed by onlooker bees for exploring weight values. In food source so called weight value search process, after a continuous trail, food source is not enhanced beyond the pre-set value, it is abandoned by employed bees and becomes a scout bees which searches food source randomly. The term "limit" defines the pre-set value of feature p's weight value for elimination. In ABC algorithm, it is an important controlling parameter. Following lists the ABC algorithm flow.

Step 1: Using the below mentioned expression, randomly created N weight values,

$$\mathrm{we}_{ij} = \mathrm{we}_j^{min} + \mathrm{rand}[0,1](\mathrm{we}_j^{max} - \mathrm{we}_j^{min}) \quad (1)$$

Where, artificial bee is represented as we_{ij}, which is at the i^{th}weight value's j^{th} position for feature 'p'; $i = 1, 2, \ldots, N$; $j = \mathrm{rand}[1, 2, \ldots, n]$;weight values count is given by N, dimension os search space is given by n, feature's j^{th}weight value's upper limit is given by we_j^{min} and lower limit is given by we_j^{max}.

Step 2:A new weight source$V_i = [\mathrm{we}_{i1}, \mathrm{we}_{i2}, \ldots, V_{ij}, \ldots, \mathrm{we}_{in}]$ are generated by employed bees using following expression, with current weight value position is $\mathrm{WE}_i = [\mathrm{we}_{i1}, \ldots, \mathrm{we}_{in}]$. $\mathrm{WE}_i = [\mathrm{we}_{i1}, \ldots, V_{ij}, \mathrm{we}_{in}]$in following equation (2)

$$v_{ij} = \mathrm{we}_{ij} + \theta_{ij}(\mathrm{we}_{ij} - \mathrm{we}_{kj}) \quad (2)$$

Where, new source position computed by i^{th}employed bee is represented as v_{ij}, $i = 1, 2, \ldots, N, j = \mathrm{rand}[1, 2, \ldots, n], k \in \{1, 2, \ldots, N\}$, and $k \neq i, \theta_{ij}$is a random number chosen from[−1, 1].

Step 3: According to food sources quality, employed bees are followed by every onlooker bee. The pr_i represents food source quality which is expressed as,

$$\mathrm{pr}_i = \frac{\mathrm{ft}_i}{\sum_{i=1}^{N} \mathrm{ft}_i} \quad (3)$$

Where, □thsolution's fitness value is represented as ft_i. For the problem of minimization, solution's fitness value is expressed as,

$$\mathrm{ft}_i = \begin{cases} \frac{1}{1+f(\mathrm{WE}_i)} & \text{if } f(\mathrm{WE}_i) > 0 \\ 1 + |f(\mathrm{WE}_i) & \text{if } f(\mathrm{WE}_i) \leq 0 \end{cases} \quad (4)$$

Where, at WE_i, objective function value is represented as $f(\mathrm{WE}_i)$.

Step 4: Abandoned the weight values, if it is not getting better in a predefined limit value. Then scouts bees are formed from employed bees. Using expression (1), randomly generated a new weight value. Onlooker bees search technique is modified in basic ABC algorithm, which is an enhancement done in this phase.

In ABC, algorithm's effectiveness is reduced due to new food source production. This is because, new weight value production is not utilizing the existing information. Inspired by Differential Evolution (DE), an Improved ABC algorithm(IABC) is proposed by Zhu and Kwong [26], where global best food source information(from images set, feature vector p's weight value) is utilized. Following expresses the IABC search technique,

$$v_{ij} = \mathrm{we}_{best,j} + \theta_{ij}\left(\mathrm{we}_{ra_1,j} - \mathrm{we}_{ra_2,j}\right) \quad (5)$$

Where, in all weights, best weight value is represented as we_{best}. Mutually exclusive integers are represented using the indices ra_1 and ra_2, they are randomly selected from {1, 2, . . . ,}, which is different from base index i. In order to enhance the exploitation further, further modifications are done in the search technique of IABC. New technique used for computing new weight value as food source is expressed as,

$$V_{ij} = we_{best,j} + \theta_{ij}\left(we_{best,j} - we_{ra_2,j}\right)(6)$$

In comparison with expression (5), for a feature vector 'p', new weight is produced around best weight value in expression (6) in addition to the usage of best weight value called food source's information for guiding search trajectory. This leads to better enhancement in algorithm's exploitation ability. From feature vector 'p', weight values are generated using distance value among two images. In a dataset, query is assigned with a query-specific weight.

Higher weight is generated for less distance between feature vectors, else it is assumed as non-query image. However, non-query images are also applied with this. As non-query images are excluded in evaluation, there is no requirement to adjust its fusion weights. After computing weights, applied the feature fusion vector. Integration of colour and texture feature vector with respective weight values are used for doing the same.

3.3.Click-Based Multi-Feature Similarity

A novel image similarity learning algorithm, which integrates multiple kernel learning and click-based triplets detection with Convolutional Neural Network (CNN) learning is proposed and it is termed as Click-based Multi Weight Feature Fusion Similarity Learning (CMWFFSL). This idea is inspired from similarity information collection in triadic form or relative comparisons [27] and integration of heterogeneous data as single and unified similarity space [28], using Multiple Kernel Partial Order Embedding (MKPOE).

In specific, users implicit feedback are represented using image click-through data and for a specified query, clicked (unclicked) images are represented as pseudo-labeled relevant (irrelevant) images. Thus, click-based triplets are leveraged with one unclicked and two clicked images. Unclicked image is used as a side information for driving metric learning algorithm. With one unclicked image x_k, and two clicked image $x_i^0$and x_j^0, similarity side-information is represented as,

"$x_i^0$is more similar to x_j^0than to x_k." (7)

In addition to the usage of click-through data as images semantic guidance for measuring images in distinct visual aspects, multiple visual modalities like texture and colour in CMWFFS used here. These are integrated as a unified similarity space using multiple kernel embedding for learning Mahalanobis distances. Between two images $x_i$and x_j, Mahalanobis distance is measured as,

$$dis_A\left(x_i, x_j\right) = \left|\left|x_i, x_j\right|\right|_A = \sqrt{\left(x_i - x_j\right)^T A\left(x_i - x_j\right)}(8)$$

Where, a positive semi-definite matrix is represented as A, and $A \succcurlyeq 0$. For m extracted features, there will be m different Mahalanobis distance. Data is mapped into a Reproducing Kernel Hilbert Space (RKHS),$\mathcal{H}$ for every feature p using a feature map ϕ having respective kernel function like linear, RBF etc, as

$$ke\left(x_i, x_j\right) = \langle\phi(x_i), \phi\left(x_j\right)\rangle(9)$$

For a feature p, inner product symmetric matrix $A^p \in \mathbb{R}^{n\times n}$expressed in (8) is indicated as,

$$A^p = ke^{p^T}W^pK^p(10)$$

Where, for feature p, kernel symmetric matrix is represented as ke^p, required unsolved matrix is represented as W^p and $W^p \succcurlyeq 0$. Target may be converted into a learning of Mahalanobis distance metric $\Phi^T W\Phi$over $\mathcal{H}$[28], after kernel function based original data mapping. Distance between image $x_i$and $x_j$by combining various visual modalities through multiple kernel embedding as,

$$\mathrm{dis}(x_i, x_j) = \sum_{p=1}^{m} \left(ke_i^p - ke_j^p\right) W^p \left(ke_i^p - ke_j^p\right) (11)$$

Where, kernel matrix's i^th column is represented as ke_i. Core concept of multiple kernel learning is an integration of numerous kernel functions into a single function. For a specified kernel functions $ke_1(x_i, x_j), ke_2(x_i, x_j), \ldots, ke_M(x_i, x_j)$,, kernel's linear grouping can be expressed as,

$$ke(x_i, x_j) = \sum_{m=1}^{M} \mathrm{dis}_m ke_m(x_i, x_j) (12)$$

Based on click count difference among two clicked image, click-based triplets are selected for detecting click-based triplets [24]. Then, as mentioned below, clicked –clicked image pairs set are defined.

$$\mathcal{J} \triangleq \{(i,j) | 0 \leq c_i - c_j \leq \delta\} (13)$$

Where, threshold used for controlling clicked-clicked image pair selection is represented as δ and it is greater than 0. Then, click-based triplets set can be defined as,

$$S \triangleq \{(i,j,k) | (c_i - c_j) \in \mathcal{J}, c_k = 0\} (14)$$

For clicked-clicked pairs, a new function is defined using a supplementary regularization and a classifier with multiple kernel embedding. There a fully connected layer, pooled layer and convolutional layer in Convolutional Neural Network (CNN) [29]. The major part of CNN is convolutional layer. From feature maps or input images, features are extracted using this layer functions. There are multiple convolution kernels in every convolutional layer, which can be used for computing multiple feature maps. The convolution layer is computed as,

$$x_b^l = f\left(\min_{W^p} \sum_{a \in p_b} x_a^{l-1} * \mathrm{tr}\left(W^a ke_{ab}^l\right) + ot_b^l\right) + \sum_{(i,j) \in \mathcal{J}} \mathrm{dis}(x_i, x_j) + \sum_{(i,j,k) \in S} \lambda_{ijk} (15)$$

$$\lambda_{ijk} = 1/\exp\left(\frac{c_i - c_j}{2\gamma^2}\right) (16)$$

Where, previous layer output's characteristic map is represented as x_a^{l-1}, b^th convolution layer's a^th channel output is represented as x_b^l and activation function is represented as $f(\cdot)$. Input feature maps subset is represented as p_b, convolution kernel is given by ke_{ab}^l and is offset is represented as ot_b^l, adapting penalty parameter is represented as λ, for query q, in $\mathcal{J}$, among all pairs, average difference value of clicks is given by γ and it is expressed as,

$$\gamma = \frac{\sum_{(i,j) \in \mathcal{J}} |c_i - c_j|}{|\mathcal{J}|} (17)$$

Where, set J's size is represented as $|J|$ and (i≠j) (i, j = 1, 2, ..., n). Between two convolutional layers, pooling layer is sandwiched in general. Feature map dimention reduction is a major functions of this layer and feature's scale invariance are maintained to some extent. Convolution and pooling process are similar, a sliding window similar to a filter is involved in this with simple computations. In an image area, average value is used as area's pooled value in mean pooling.

Image background are well preserved using this approach. Image area's, maximum value is taken as a area's pooled value in max pooling and image texture's are preserved well. Multiple image maps computed by passing the image through various convolution layers are integrated using fully connected layer and for subsequent image classification, high-layer semantic features are obtained by pooling layers.

Projected Sub-Gradient Descent (PSD) based solution is proposed. For single feature, with respect to Frobenius inner products, distance is computed as follows,

$$\mathrm{dis}(x_i, x_j) = (ke_i - ke_j)^T W (ke_i - ke_j) = (e_i - e_j)^T keWke(e_i - e_j) (18)$$

ISSN 2515-8260 Volume 07, Issue 09, 2020

$$= \mathrm{tr}\left(\mathrm{keWke}\left(\mathrm{e_i} - \mathrm{e_j}\right)\left(\mathrm{e_i} - \mathrm{e_j}\right)^{\mathrm{T}}\right)(19)$$
$$= \mathrm{tr}\left(\mathrm{WkeE_{ij}ke}\right) = \langle \mathrm{W}, \mathrm{keE_{ij}ke}\rangle_{\mathrm{F}}(20)$$

Where, i^{th} standard basis vector is given by e_i, so that $e_i\, ke^T e_i$,and$E_{i,j} = (e_i - e_j)(e_i - e_j)^T$. Summarization of distance among two images $x_i$and xre expressed as,

$$\mathrm{dis}\left(\mathrm{x_i}, \mathrm{x_j}\right) = \sum_{p=1}^{m} \mathrm{dis}^{p}\left(\mathrm{x_i}, \mathrm{x_j}\right) = \sum_{p=1}^{m} \mathrm{tr}\left(\mathrm{ke^p W^p ke^p E_{ij}^p}\right) = \sum_{p=1}^{m} \langle \mathrm{W^p}, \mathrm{ke^p E_{ij}^p ke^p}\rangle_{\mathrm{F}}(21)$$

Over m distinct n× n matrices W^p, a Semi-Definite Program (SDP) is used for formulating optimization after multiple kernel embedding. Every W^p is constrained as diagonal matrix by redefining optimization problem for dealing with huge datasets. Diagonal value's non-negativity is similar to positive definiteness, if $W^p$are all diagonal, it produces matrix's eigenvalues. This permits the replacement of constraints$W^p \succcurlyeq 0$ with linear constraints$W_{ii}^p \succcurlyeq 0$ and a linear program is formulated by simplifying optimization rather than SDP [31]. Model flexibility is reduced b this model, but it results in a highly effective optimization procedure requirement [28].

3.4.Learning Click-Based Typicality

The 'm' different $W^p(p = 1, 2, \dots, m)$, are obtained after learning click-based multi-feature similarity procedure and image similarity is measured using an expression (18). For detecting relevant recurrent patterns according to learnt similarity metrics, Semi-supervised Consensus Clustering (SCC) for grouping semantically as well as visually similar images into same clusters. From data, pairwise-affinity matrix is constructed easily according to learnt image similarity metrics and relevant operation is performed directly.

In spectral clustering, prior knowledge is used in SCC. From domain knowledge, pairwise constraints are used in this. Between two semantically as well visually similar images, pairwise constraints are represented as cannot-links (in different classes) and as must-links (in the same class). For every mustlink(i, j) pair, $\mathrm{sim_{ij}} = \mathrm{sim_{ji}} = 1$ is assigned and for every cannot-link (i, j) pair, $\mathrm{sim_{ij}} = \mathrm{sim_{ji}} = 0$ is assigned. In the graph, connection between two semantically as well as visually similar images with high similar features is made, if SSC is used to cluster images in feature vector via t-nearest neighbour graph representation.

Between image pair, similarity is changed to 0 for cannot-link constraints and in graph, edge between images pair is broken. This study applies only must-link constraints. Algorithm 1 describes the Semi-supervised spectral clustering algorithm details. For a specified images $I_1, \dots, I_n$, created l pairwise must-link constraints. Similarity function expressed in (22) is used for obtaining similarity matrix 'S'.

$$sim_{ij} = \exp\left(-\frac{\left||x_i - x_j|\right|^2}{2\sigma^2}\right)(22)$$

where, scaling parameter is represented as σ, which is used to measure similarity between two images. Sparse matrix is formed by modifying S, where, for every image in S, only t nearest neighbours are maintained. Then, in S, applied the l pairwise constraints. Normalized spectral clustering algorithm is followed in steps 5 to 10 [32]. Algorithm 2 describes the semi-supervised consensus clustering algorithm. As in [33], for a specified n×$m$web image with n samples and m *features*, generated a n×$q$image subspace ($q < d$) as,

$$q = q_{min} + [\alpha(q_{max} - q_{min})](23)$$

Where, a uniform random varibale is represented as α, which lies between 0 to 1, subspace's lower bound is represented as q_{min} and upper bound is represented as q_{max}. 0.75d is said as

q_{min} and 0.85d is said as q_{max}. Assume a cluster ensemble as $\pi = \pi_1, \ldots, \pi_u$ with u clustering solutions. On every subspace dataset, applied this clustering for computing clustering results (u). Fixed clusters count u is used and one clustering solution is represented by every $\pi_i = u_{i1}, \ldots, u_{ik}$. Between clusters and images, according to crisp associations, generated the Cluster-Association Matrix 'AM*'*, it has $m*u$ clusters and n images.

If $I_i$belongs to a cluster u_j, $\mathrm{AM}(I_i, u_j) = 1, i = 1, \ldots, n;\ j = 1, \ldots, m$, otherwise $\mathrm{AM}\left(I_i, u_j\right) = 0$. In $R\mathrm{AM}(I_i, u_j)$, new association values are estimated for generating a Refined Cluster-Association matrix 'RAM', if $\mathrm{AM}(I_i, u_j) = 1$. Between $u_j$and other clusters to which $I_i$probably belongs to them each other, similarity function is represented as $RAM(I_i, u_j)$. From a clusters weighted graph, similarity between any clusters can be obtained in cluster ensemble. At last, for computing final clustering solution, on RAM, applied the spectral clustering.

Algorithm 1: Semi supervised spectral clustering

Input :n number of images $I_1, \ldots I_n$, clusters count (u) and pairwise constraints count 'l'

Output: Visually and semantically similar images are grouped into u clusters

1. From images,l must-link constraints are generated.
2. Similarity matrix S is constructed, where $sim_{ij} \geq 0$ represents similarity between images $x_i \& x_j$
3. Using t-nearest neighbor graph, S is modified as sparse matrix.
4. The l pairwise constraintsis applied on S $sim_{ij} = sim_{ji} = 1$
5. Normalized laplacianmatrix $L = IM - D^{-\frac{1}{2}} S D^{-\frac{1}{2}}$ is computed. Diagonal matrix defines degree matrix D with degrees $d_1, .. d_n$ on diagonal $d_i = \sum_{j=1}^{m} sim_{ij}$
6. First u eigenvectors $ev_1, \ldots ev_k$ of L are computed
7. The eigen vector $EV \in \mathbb{R}^{n*u}$ to be matrix with vectors $ev_1, .. ev_k$ as columns
8. Form matrix $T \in \mathbb{R}^{n*u}$ from EV by normalizing rows to norm $1 t_{ij} = \frac{ev_{ij}}{\left(\sum_u ev_{iu}^2\right)^{1/2}}$
9. For i=1,..n let $y_i \in \mathbb{R}^u$ be a vector with respect to i^{th} row of T
10. Images cluster y_i with Semi-supervised consensus clustering (SCC)

Algorithm 2: Semi-supervised consensus clustering (SCC)

Input: Specify an image $n \times m$ $I_1, \ldots, I_n$with n images and m features. Cluster count u, pairwise constraints count l, size of ensemble *'e'*, folds count h in cross-validation are set.

Output: $I_1, \ldots, I_n$are grouped into u clusters

1. In every run, images having feature vectors are split into h fold. In every fold, run steps 2-5.
2. From other $h - 1$ fold images, must-link's $l$pairwise constraints are generated.
3. A cluster ensemble$\prod = \pi_1, \ldots, \pi_u$ is generated with uclustering solutions, $\pi_i = u_{i1}, u_{i2}, \ldots u_{ie}$.
   3.1.The k subspace images $B_i, i = 1, \ldots, k, B_i \in \mathbb{R}^{n*q}, q < d$ are generated
   3.2.On $B_1, \ldots, B_m$, algorithm 1 in steps 2-10 is applied with the fixed clusters count$u$, and π_i is computed
   3.3.In cluster ensemble $\prod$, π_i is stored.
4. From $\prod$, cluster-association matrix RAM is generated.
5. On RAM, spectral clustering is applied and images are clustered into u clusters.

Following sections presents click-based local typicality and click-based cluster typicality in a separate manner.

1) Cluster Typicality: The relative cluster similarity sim(u) and initial cluster confidence confd (u) are used for deciding cluster typicality. Where, clusters are represented using u. With respect to specified query, in this cluster, overall sample relevance are measured using initial cluster confidence. Combination of click score and initial score are used for defining initial cluster confidence as,

$$\mathrm{confd}(u) = \sum_{x_i \in u} 1 - \frac{r(x_i)}{n} + \frac{c_i}{\mathrm{sum}(c)} \quad (24)$$

Where, images count to be re-ranked is represented as n, x's initial ranked order is represented as r (x), in initial ranked list, from all clicked images, click counts summation is represented as sum(c). On the other side, image representativeness to a specified query topic is reflected using typicality. There will be similarity between images in the same cluster, if they have good typicality. So, ratio between intra-cluster and inter-cluster similarity defines relative cluster similarity, which is given by,

$$\mathrm{sim}(u) = \frac{\mathrm{avg}_{x_i, x_j \in u \wedge i \neq j} \mathrm{dis}(x_i, x_j)}{\mathrm{avg}_{x_i \in u \wedge x_k \notin u} \mathrm{dis}(x_i, x_k)} \quad (25)$$

Within cluster u, between images, average similarity is represented using the first term in the above mentioned expression, in the same cluster, between any image, average similarity is represented using the second term. According to above two measures, defined the cluster typicalityCT(u) as a linear weighted combination.

$$CT(u) = \gamma * conf(u) + (1 - \gamma) * sim(u) \quad (26)$$

2) Local Typicality: Local density ld(x), image initial confidence con(x) are used for computing local typicality. In cluster u, image x's local density $ld(x|u)$ is computed using Kernel Density Estimation (KDE).

$$ld\left(x_j^i \middle| u^i\right) = \frac{1}{|u_i|} \sum_{j=1}^{|u^i|} F\left(x^i - x_j^i\right) \quad (27)$$

Where, kernel function which satisfies $\int F(x)dx = 1$ and F(x) > 0 is represented as F(x). Exponential kernel function F(x)=exp(||−x|| /2δ) is adopted here. On different metrics, there will be a discussion about results and analysis in the following section and comparison is made with available works.

4. RESULTS AND DISCUSSION

Experimental settings are introduced in this section with results. In order to validate the proposed re-ranking techniques effectiveness, overall performance is specified through multiple features integration, usage of various distance metrics with comparison. The MATLAB is used for implementing available and proposed techniques.

With different metrics, available Click-Based Multifeature Similarity Learning (CMSL), Support Vector Machine (SVM) based web image retrieval, Multiple Support Vector Machine with Kernel Learning (MSVM-KL) and proposed CMWFFSL algorithm are compared. Clustering results of proposed SCCCST and available techniques like feature extraction and re-ranking using click based similarity and typicality (FE-ReRCST), SCCST-SIFT (Scale-Invariant Feature Transform), Spectral Clustering re-ranking with Click-based Similarity and Typicality (SCCST) are also compared.

Dataset have to add

4.1.Performance measures

For comparison of performances, metrics like accuracy, NDCG, mean Average Precision (mAP), Average Precision (AP) are used in this research work.

ISSN 2515-8260 Volume 07, Issue 09, 2020

4.1.1. Normalized Discounted Cumulative Gain (NDCG)

Annotators labels every query-image pair in above mentioned dataset carefully on a scale 0 to 1: 0–"irrelevant," and 1–"relevant." For measuring the performance Normalized Discounted Cumulative Gain (NDCG) is adopted, which is having wide use in information retrieval with two or more relevance levels. For a specified ranked list, at depth *l,* NDCG is expressed as,

$$NDCG@l = Z_l \sum_{i=1}^{l} \frac{2^{r^i}-1}{\log(1+i)} \quad (28)$$

Where, i^{th} image's relevance score is represented as r^i and normalization constant is represented as Z_l for guaranteeing a perfect NDCG ranking when *l*is equal to 1.

4.1.2. Average Precision (AP)

The retrieval results of single query's retrieval quality is evaluated using a metric called Average Precision (AP). Both recall and precision values are considered in this AP. Fraction of relevant retrieved (top k) images represents precision and fraction of retrieved relevant images (top k returned results) represents recall.

Increase in recall or increase in retrieved image count decreases the retrieval system's precision in general. From the rank positions, precision values are averaged using AP, where a relevant image is retrieved, which is expressed in (29).

The mAP is adopted usually for summarizing the retrieval quality over multiple query images, where average precision is averaged over entire queries.

$$AP = \frac{\sum_{k=1}^{n} P(k).rel(k)}{R} \qquad (29)$$

Where, for current query image, relevant results count are represented as R, top k retrieval result's precision are represented as p(k), binary indicator function us represented as rel(k), this value will be 1 for match between current query image and k^{th} retrieved result, else it will be zero, total retrieved results count is represented as n.

4.1.3. mean Average Precision (mAP)

For every query, average precision score's means defines t mean Average Precision (mAP) for a queries set.

4.1.4. Accuracy

Among total number of cases, ratio between true positive results and false negative results defines accuracy.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \times 100 \qquad (30)$$

True Positive is represented as TP, True Negative is represented as TN, False Positive is represented as FP, False Negative is represented as FN.

4.2.Results comparison

Results of different metrics discussed above are compared based on various image retrieval techniques.

TABLE 1. ASSESSMENT OF VARIOUS TECHNIQUES ON IMAGE RETRIEVAL WITH RESPECT TO NDCG

Measure	Retrieval methods-NDCG			
NDCG@l	**SCCST-SIFT**	**SCCST**	**FE-ReRCST**	**SCCCST**
NDCG@10	0.753	0.768	0.801	0.823
NDCG@20	0.762	0.773	0.823	0.847
NDCG@30	0.775	0.795	0.835	0.862
NDCG@40	0.785	0.801	0.854	0.871
NDCG@50	0.796	0.814	0.869	0.887

NDCG@60	0.802	0.825	0.878	0.909
NDCG@70	0.814	0.838	0.889	0.915
NDCG@80	0.826	0.847	0.908	0.926
NDCG@90	0.837	0.862	0.918	0.939
NDCG@100	0.841	0.872	0.924	0.957

TABLE 2. VARIOUS IMAGE RETRIEVAL TECHNIQUES WITH RESPECT TO AVERAGE PRECISION (AP)

Number of images	**Retrieval methods -Average Precision (AP)**			
	CMSL	**SVM**	**MSVM-KL**	**CMWFFSL**
20	0.8256	0.8452	0.8654	0.8871
40	0.8315	0.8524	0.8715	0.8945
60	0.8418	0.8647	0.8879	0.9042
80	0.8571	0.8749	0.8928	0.9261
100	0.8712	0.8892	0.9128	0.9318

TABLE 3. VARIOUS IMAGE RETRIEVAL TECHNIQUES WITH RESPECT TO MEAN AVERAGE PRECISION (mAP)

Number of images	**Retrieval methods - mean Average Precision (mAP)**			
	CMSL	**SVM**	**MSVM-KL**	**CMWFFSL**
20	0.8158	0.8318	0.8541	0.8692
40	0.8252	0.8563	0.8693	0.8925
60	0.8389	0.8692	0.8852	0.9125
80	0.8463	0.8789	0.8928	0.9281
100	0.8625	0.8812	0.9076	0.9363

TABLE 4 .VARIOUS IMAGE RETRIEVAL TECHNIQUES WITH RESPECT TO ACCURACY

NUMBER OF IMAGES	**RETRIEVAL METHODS – ACCURACY (%)**			
	CMSL	**SVM**	**MSVM-KL**	**CMWFFSL**
20	80.22	82.52	86.52	90.19
40	81.78	83.89	87.37	91.45
60	82.63	84.54	88.48	92.73
80	84.56	85.69	89.66	93.86
100	85.74	87.45	90.78	94.75

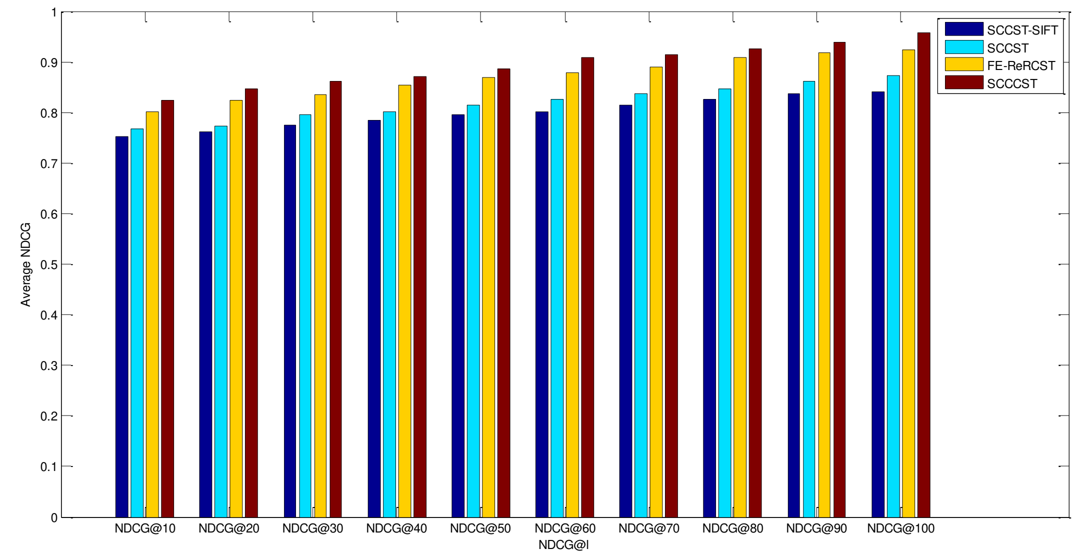

FIGURE 2. COMPARISON OF RE-RANKING APPROACHES IN TERMS OF NDCG

Figure 2 shows the overall performance of different reranking approaches on dataset. On the whole, the proposed SCCCST outperforms other methods, and the improvements are consistent and stable at different depths of NDCG. Using proposed SCCCST re-ranking approach the NDCG values are improved significantly. To be specific, the NDCG@100 produces of 0.957, whereas other methods such as SCCST-SIFT, SCCST and FE-ReRCST gives only 0.841, 0.872 and 0.924 on the entire dataset (See Table 1).

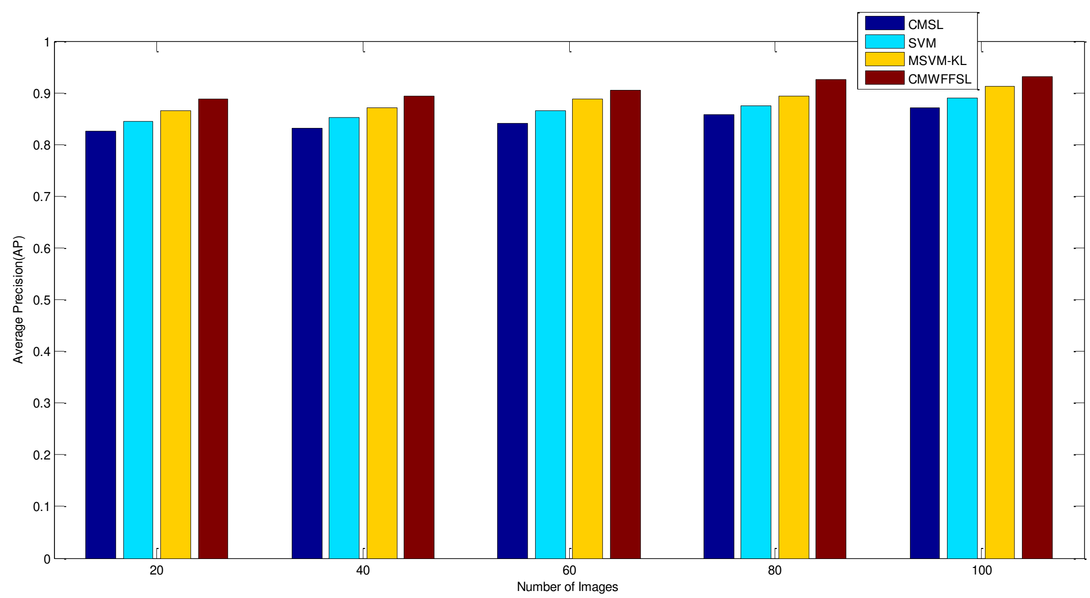

FIGURE 3. COMPARISON OF RETRIEVAL METHODS IN TERMS OF AVERAGE PRECISION (AP)

The AP performance comparison between different image retrieval techniques are shown in figure 3. It indicates that, existing methods are outperformed by proposed CMWFFSL and for various images counts, consistent and stable performance is exhibited using this technique. Around 0.9318 AP value is attained in proposed CMWFFSL, whereas, 0.8712 is attained in CMSL, 0.8892 is attained in SVM and 0.9128 (as shown in table 2) is attained in MSVM-KL for

100 image set as shown in that figure. When compared with existing techniques, high AP value can be attained using proposed CMWFFSL technique as depicted in implemented and validated results.

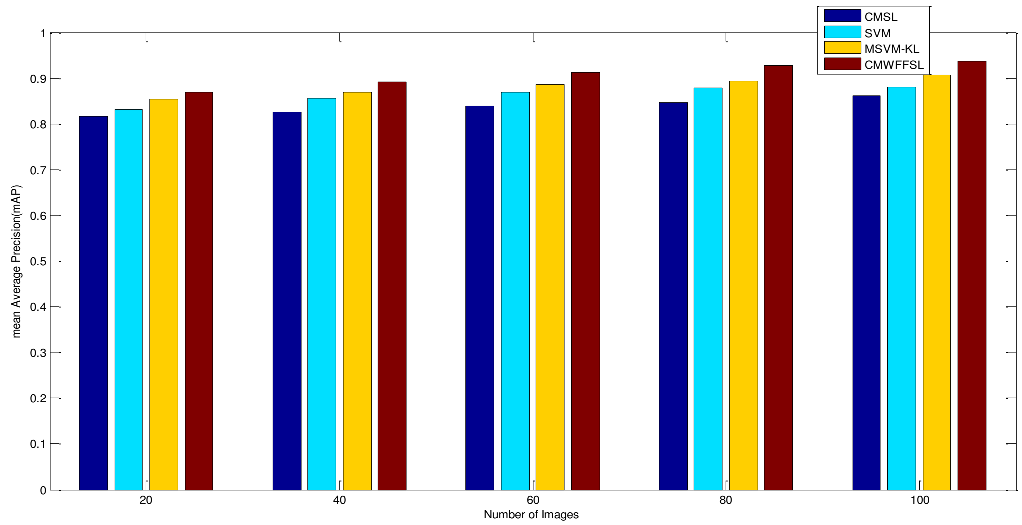

FIGURE 4. COMPARISON OF RETRIEVAL METHODS IN TERMS OF MEAN AVERAGE PRECISION (mAP)

The mAP performance comparison between different image retrieval techniques are shown in figure 4. It indicates that, existing methods are outperformed by proposed CMWFFSL and for various images counts, consistent and stable performance is exhibited using this technique. Around 0.9363 mAP value is attained in proposed CMWFFSL, whereas, 0.8625 is attained in CMSL, 0.8812 is attained in SVM and 0.9076(as shown in table 3) is attained in MSVM-KL for 100 image set as shown in that figure. When compared with existing techniques, high mAP value can be attained using proposed CMWFFSL technique as depicted in implemented and validated results.

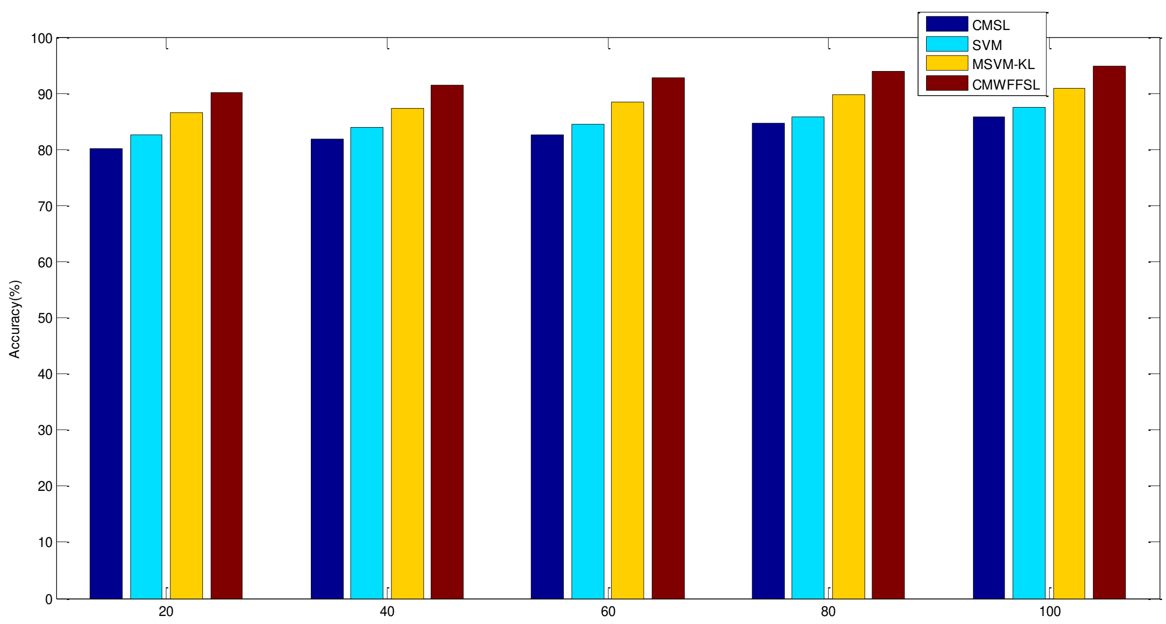

FIGURE 5. COMPARISON OF RETRIEVAL METHODS IN TERMS OF ACCURACY

Accuracy comparison between different image retrieval techniques are shown in figure 5. Around 94.75% of accuracy can be attained in proposed CMWFFSL, whereas, 85.74% is attained in CMSL, 87.45% is attained in SVM and 90.78% is attained in MSVM-KL for 100 image set as shown in that figure. When compared with existing techniques, high accuracy can be attained using proposed CMWFFSL technique as depicted in implemented and validated results.

5. CONCLUSION AND FUTURE WORK

Issues related to leveraging click-through data for reducing image search intent gap are studied in this paper. Separate extraction of colour and texture from given keyword text is performed in the initial stage. Weight based Multi-Feature Fusion (WMFF) is used for fusing the extracted color and texture features. In WMFF algorithm, Improved Artificial Bee Colony (IABC) is used for creating feature weights. Based on web query, weights are assigned for features, where different weights are assigned to ranked list of various queries.

An Improved ABC algorithm (IABC) is proposed by Differential Evolution (DE), where global best food sources information are used. From image set, feature vector ‘p’s weight value is used. A Semi-supervised Consensus Clustering re-ranking with click-based similarity and typicality (SCCCST) is entirely adopted in this proposed re-ranking technique for guiding image typicality and similarity learning.

For re-ranking, proposed a novel image similarity learning algorithm called Click-based Multi Weight Feature Fusion Similarity Learning (CMWFFSL) via Convolutional Neural Network with Multiple Kernel Learning (CNN-MKL). For multiple features, after learning similarity metric for grouping semantically and visually similar images as same clusters proposed a Semi-supervised Consensus Clustering (SCC). Cluster typicality and within-clusters image typicality in descending order are computed for computing final re-rank list.

With respect to metrics like accuracy, Normalized Discounted Cumulative Gain (NDCG), mean Average Precision (mAP), Average Precision (AP) are used for comparing proposed SCCCST with various available re-ranking techniques. Superiority and availability of proposed SCCCST is demonstrated using extensive results. Moreover, on feature-weighting, its effectiveness is also exploited in computer vision problems like image segmentation, object detection, person re identification.

REFERENCES

1. Mei T., Y. Rui, S. Li, and Q. Tian, “Multimedia search reranking: A literature survey,” ACM Comput. Surv., vol. 46, no. 3, p. 38, 2014.
2. Hsu W. H., L. S. Kennedy, and S.-F. Chang, “Reranking methods for visual search,” IEEE Multimedia, vol. 14, no. 3, pp. 14–22, 2007.
3. Hsu W. H., L. S. Kennedy, and S.-F. Chang, “Video search reranking through random walk over document-level context graph,” in Proc. 15th ACM Int. Conf. Multimedia, 2007, pp. 971–980.
4. Yao, T., Ngo, C.W. and Mei, T., 2013. Circular reranking for visual search. IEEE Transactions on Image Processing, 22(4), pp.1644-1655.
5. Cui J., F. Wen, and X. Tang, “Real time google and live image search re-ranking,” in Proc. 16th ACM Int. Conf. Multimedia, 2008, pp. 729–732.
6. Tian, X., Tao, D., Hua, X.S. and Wu, X., 2010. Active reranking for web image search. IEEE Transactions on Image Processing, 19(3), pp.805-820.

7. Zhang C., J. Y. Chai, and R. Jin, “User term feedback in interactive text-based image retrieval,” in Proc. 28th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2005, pp. 51–58.
8. DupretG. and C. Liao, “A model to estimate intrinsic document relevance from the clickthrough logs of a Web search engine,” in Proc.3rd ACM Int. Conf. Web Search Data Mining, 2010, pp. 181–190.
9. DupretG. E. and B. Piwowarski, “A user browsing model to predict search engine click data from past observations,” in Proc. 31st Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., 2008, pp. 331–338.
10. Zhang Y., X. Yang, and T. Mei, “Image search reranking with querydependent click-based relevance feedback,” IEEE Trans. Image Process., vol. 23, no. 10, pp. 4448–4459, 2014.
11. Tang J., R. Hong, S. Yan, T.-S. Chua, G.-J. Qi, and R. Jain, “Image annotation by kNN-sparse graph-based label propagation over noisily tagged Web images,” ACM Trans. Intell. Syst. Technol., vol. 2, no. 2, p. 14, 2011.
12. Tang J., Z. Li, M. Wang, and R. Zhao, “Neighborhood discriminant hashing for large-scale image retrieval,” IEEE Trans. Image Process., vol. 24, no. 9, pp. 2827–2840, 2015.
13. Liu Y., T. Mei, M. Wang, X.Wu, and X.-S. Hua, “Typicality-based visual search reranking,” IEEE Trans. Circuits Syst. Video Technol., vol. 20, no. 5, pp. 749–755, 2010.
14. Liu Y. and T. Mei, “Optimizing visual search reranking via pairwise learning,” IEEE Trans. Multimedia, vol. 13, no. 2, pp. 280–291, 2011.
15. Duan L., W. Li, I. W. H. Tsang, and D. Xu, “Improving Web image search by bag-based reranking,” IEEE Trans. Image Process., vol. 20, no. 11, pp. 3280–3290, 2011.
16. San Pedro J., T. Yeh, and N. Oliver, “Leveraging user comments for aesthetic aware image search reranking,” in Proc. ACM 21st Int. Conf.World Wide Web, 2012, pp. 439–448.
17. Wang M., H. Li, D. Tao, K. Lu, and X. Wu, “Multimodal graph-based reranking for Web image search,” IEEE Trans. Image Process., vol. 21, no. 11, pp. 4649–4661, 2012.
18. Umesh, K.K., 2013. Semantic Based Image Retrieval System for Web Images. In Advances in Computing and Information Technology (pp. 491-499). Springer, Berlin, Heidelberg.
19. Cai, J., Zha, Z.J., Wang, M., Zhang, S. and Tian, Q., 2015. An attribute-assisted reranking model for Web image search. IEEE transactions on image processing, 24(1), pp.261-272.
20. Ji Z., Y. Pang, and X. Li, “Relevance preserving projection and ranking for Web image search reranking,” IEEE Trans. Image Process., vol. 24, no. 11, pp. 4137–4147, 2015.
21. Yang X., Y. Zhang, T. Yao, C.-W. Ngo, and T. Mei, “Click-boosting multi-modality graph-based reranking for image search,” Multimedia Syst., vol. 21, no. 2, pp. 217–227, 2015.
22. Yang, X., Mei, T., Zhang, Y., Liu, J. and Satoh, S.I., 2016. Web image search re-ranking with click-based similarity and typicality. IEEE transactions on image processing, 25(10), pp.4617-4630.
23. deSiqueira, F.R., Schwartz, W.R. and Pedrini, H., 2013. Multi-scale gray level co-occurrence matrices for texture description. Neurocomputing, 120, pp.336-345.
24. Yang X., T. Mei, and Y. Zhang, “Rescue tail queries: Learning to image search re-rank via click-wise multimodal fusion,” in Proc. ACM Int. Conf. Multimedia, 2014, pp. 537–546.
25. KarabogaD. and B. Basturk, “On the performance of artificial bee colony (ABC) algorithm,” Applied Soft Computing, vol. 8, no. 1, pp. 687–697, 2008.
26. Zhu G. and S. Kwong, “Gbest-guided artificial bee colony algorithm for numerical function optimization,” AppliedMathematics and Computation, vol. 217, no. 7, pp. 3166–3173, 2010.

27. Schultz M. and T. Joachims, “Learning a distance metric from relative comparisons,” in Proc. Adv. Neural Inf. Process. Syst. (NIPS), 2004,p. 41.
28. McFee B. and G. Lanckriet, “Learning multi-modal similarity,” J. Mach. Learn. Res., vol. 12, pp. 491–523, 2011
29. Wen, L., Li, X., Gao, L. and Zhang, Y., 2017. A new convolutional neural network-based data-driven fault diagnosis method. IEEE Transactions on Industrial Electronics, 65(7), pp.5990-5998.
30. Long, Y., Gong, Y., Xiao, Z. and Liu, Q., 2017. Accurate object localization in remote sensing images based on convolutional neural networks. IEEE Transactions on Geoscience and Remote Sensing, 55(5), pp.2486-2498.
31. Gu B., V. S. Sheng, K. Y. Tay, W. Romano, and S. Li, “Incremental support vector learning for ordinal regression,” IEEE Trans. Neural Netw. Learn. Syst., vol. 26, no. 7, pp. 1403–1416, 2015.
32. Luxburg UV: A tutorial on spectral clustering, statistics and computing. ACM ComputSurv 2007, 17(4):395–416.
33. Lam-on N, Boongoen T, Garett S: LCE: a link-based cluster ensemble method for improved gene expression data analysis. Bioinformatics 2010, 26(12):1513–1519.