IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 22 May 2023, accepted 18 June 2023, date of publication 26 June 2023, date of current version 1 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3289759

RESEARCH ARTICLE

Reverse Image Search for Collage: A Novel Local Feature-Based Framework

MUHAMMAD ZUBAIR[1], MUHAMMAD AFFAN ALIM[1,2], IMRAN NASEEM[3,4], MUHAMMAD MANSOOR ALAM[2,5,6], AND MAZLIHAM MOHD SU'UD[2]

[1]College of Computing and Information Sciences, Karachi Institute of Economics and Technology (KIET), Karachi, Sindh 75190, Pakistan
[2]Faculty of Computing and Informatics, Multimedia University (MMU), Cyberjaya 63100, Malaysia
[3]College of Engineering, Karachi Institute of Economics and Technology (KIET), Karachi 75190, Pakistan
[4]School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Crawley, WA 6009, Australia
[5]Faculty of Computing, Riphah International University, Islamabad 46000, Pakistan
[6]Malaysian Institute of Information Technology, University of Kuala Lumpur, Kuala Lumpur 50250, Malaysia

Corresponding author: Mazliham Mohd Su'ud (mazliham@mmu.edu.my)

ABSTRACT Collage, a popular form of visual-content summarization technique is commonly used by internet users and digital artists. Social media usage is a rising trend that significantly affects the increasing demand for collages. The primary source of collage generation is social media, but other sources also generate it. Searching for a required query image in this corpus is a crucial demand and also valuable. The query image can be retrieved using Reverse Image Search (RIS), either in its exact form or with a small variation. Well-known search engines like Google and Yandex have this functionality, but their method has not been made public. In this research, we propose a consolidated framework for reverse image searching for the problem of collage. Essentially, the local features of collage images are extracted by using SIFT, SURF, and ORB algorithms. These features undergo the localization of the region of interest (ROI) process which handles by binning technique. We propose to use the Manhattan distance to calculate the similarity. The proposed model is extensively evaluated on standard databases and is shown to always have good results using SIFT algorithm. The proposed model is entirely generic and attains 90.96%, accuracy using the SIFT algorithm. The proposed approach is also evaluated on flip and scale variant college and achieves a result of 83% and 78% respectively, using SIFT algorithm.

INDEX TERMS Collage, collage detection, computer vision, machine learning, reverse image search (RIS).

I. INTRODUCTION

The term "collage" is well-known among internet users, especially those who use social media. It's frequently used by the internet community and digital artists for creative ways to condense a growing amount of visual data into a single view. Collage, on the other hand, can be used to depict pictures, music, text, and other art media in a single piece, but it is most frequently used to refer to visual arts. There is no fixed pattern for blending images in a collage. In this study, we have concentrated on visual collages. The organization of internal images in the collage, as seen in Figure (1), necessitates a diverse set of skills. However, while the internal image arrangement is important, other aspects such as rotation, zoom, crop, structure, and color modification are equally important when constructing collages.

Reverse Image Search (RIS) is a feature of the Content-based Information Retrieval (CBIR) system. It works similarly to a search engine that takes an image as input instead of meta-data or text as a query. Google image search tool, currently, is the most popular and widely used CBIR system [1]. A difference in the searching process between a regular search engine and CBIR is shown in Figure (2).

Essentially, Google has the highest number of index pages, thus its image-searching capability is robust compared to others [2]. The CBIR works as a system through which a user can search for relevant images in the available database. The output of the CBIR can be categorized into two types (i) exact duplicates or slightly variant copies of the query images, and

The associate editor coordinating the review of this manuscript and approving it for publication was Mingbo Zhao.

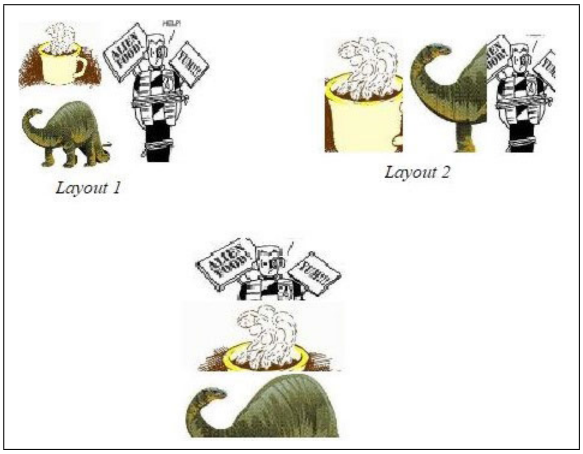


FIGURE 1. Example of different layouts for 3-image collage.

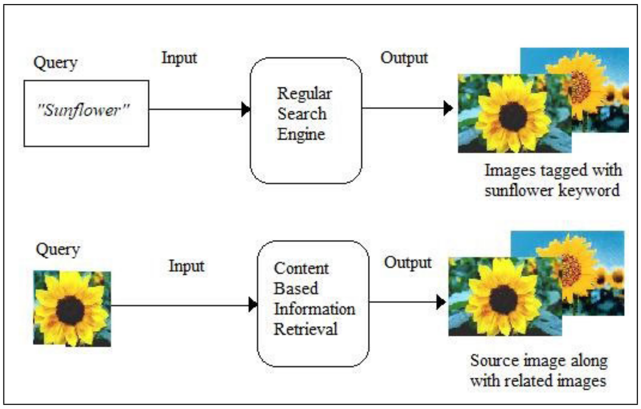


FIGURE 2. Example of difference between a regular search engine and CBIR.

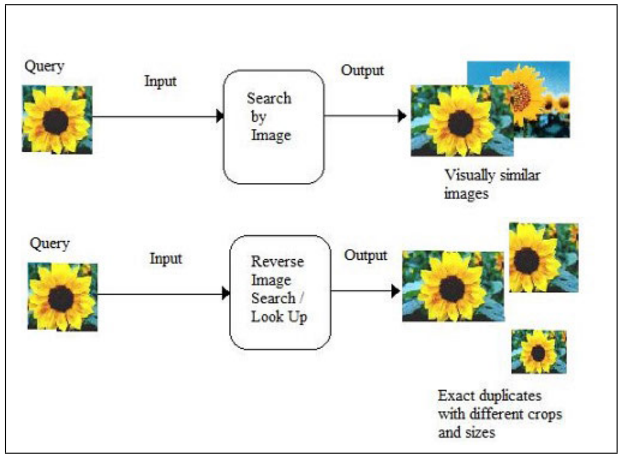


FIGURE 3. Example of outputs of search by image and reverse image search/look up.

(ii) visually similar images based on the contents of the query images, which are called reverse image search or reverse image look-up and search by image, respectively [3], [4]. The difference between the above two methods is shown in Figure (3).

The CBIR process consists of two phases (i) Feature extraction, and (ii) finding a match. In the process of feature extraction, it extracts the low-level information of images such as color, texture, hues, and shapes by using different algorithms and presents it in a highly descriptive way called features of an image. Essentially, the feature extraction process has two types (i) global and (ii) local feature extraction methods. The global feature extraction refers to the feature which represents the whole image as a single unit [5], whereas the local feature extraction uses information on the local patches of the image to get segmented region [5]. Within the last two decades extensive work has been done on feature extraction techniques and state-of-the-art algorithms have been proposed [1]. The most widely used feature extraction methods in the field of pattern recognition are (HOG) [6], Local Binary Pattern (LBP) [7], [8], Scale Invariant Feature Transform (SIFT) [8], speeded-up robust features (SURF) [10], [11], oriented Fast and Rotated Brief (ORB) [12]. We have proposed to use SIFT, SURF, and ORB to extract the features from the collage and used these features for further refinement. To the best of our knowledge, the issue of reverse image searching in collage (RISiC) is not currently being addressed in any proposed works. In this research, it is for the first time that the reverse image searching for collage is proposed. The RISiC method is offered by contemporary search engines like Google, Yandex, and Bing, but the technique and algorithms are not made available to the public. There is a dire need for a generalized method for RISiC which should be available publicly. To fill this gap, we have developed a framework based on the local feature. Although, in prior studies, some similar work has been done. The author has compared the performance of SIFT, SURF, BRIEF, and ORB in the context of 2D object recognition in [13]. In [14], the strength of CNN for class membership prediction was underlined by the authors, who also developed a method that uses it to retrieve images using a modified distance function in the wavelet feature space. Under the umbrella of the proposed method, the author combined the low-level and deep features of an image representation. In [15], graph-based reasoning attention pooling with curriculum design (GRAP-CD) is proposed for content-based image retrieval (CBIR). In addition to examining relationships between salient regions, GRAP-CD may gradually train the network to produce improved local minima.

In the current research we propose a fully automatic framework for reverse image searching in collage, Figure (9) depicts the workflow of the proposed method. we have proposed to use SIFT, SURF, and ORB for feature extraction based on brute force matcher (BFM). the resulting pair of matches is dispersed across the collage. Many of these pairs of matches are selected from the false regions. To find the region of interest (ROI), all false regions are discarded by using the proposed method Median of Mode (MoM). An updated ROI is extracted by aligning the center point of the query image with the Median with respect to the height and width of the query image as shown in Figure (4).

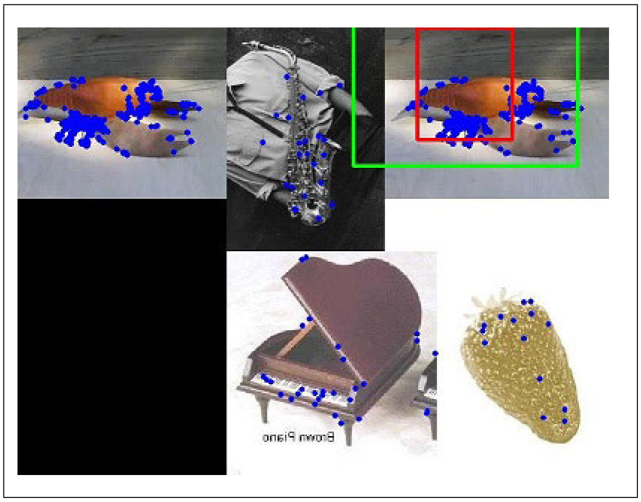


FIGURE 4. Updated bounding box of ROI.

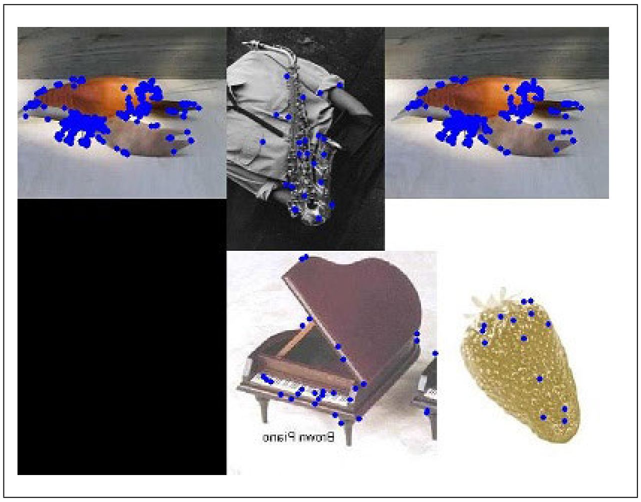


FIGURE 5. Matching feature points of collage to query.

A complete flow of the proposed model is shown in Figure (9). Finally, to validate the updated ROI, a Euclidean distance is used on the color ratio set. Those will be the best match where the updated ROI is had minimum distance.

The rest of the paper is organized as follows: The previous related works are presented in section II followed by the proposed framework for the collage reverse image search is discussed in section III. Detailed experimental work and its results are discussed in section IV, and finally, the paper is concluded in section V.

II. LITERATURE REVIEW

Content-based image retrieval (CBIR) is a technology that takes an image as input often referred to as a query image, and after processing, it returns output related to the query image. Essentially, the approach utilizes the image's actual content rather than its meta-data, such as image keywords or descriptions. The Reverse Image is a type of (CBIR) that is typically used in a search engine and aims to find a similar, identical image on the web according to the query image. Nevertheless, the output of the search images is cropped, rotated, illumination changes, etc.

All approaches that use the reverse image search (RIS) are based on Content-based image retrieval (CBIR) which has been aforementioned in section I. Each proposed method has some advantages and disadvantages. In [16], the author has proposed a color-based approach that utilizes the composite and stacked color histogram. The classical concept of Euclidean distance, for instance, has been reported to achieve improved results. In [17] the process of image classification used a model by SURF and texture features. A performance comparison between the local features algorithm SIFT, SURF, BRIEF, and ORB for RIS has been studied in [13]. The performance has also been measured on a variety of orientations such as rotation, scale, etc. In [18] local feature algorithms viz SURF, BRISK, and ORB have been used for feature extraction followed by the brute force method used for classification. In [19], the author has proposed an MSTD-based methodology that consists of color, edge orientation, and texture quantized values, the color factor contributes huge drift in the performance for feature description of visual contents. Another similar research has been used in [20], multiple color encoding i.e. RGB & YCbCr, shape, and texture were used for extraction of basic features followed by the genetic algorithm proposed for the feature matching. The concept of the dominant color-based approach for the MPEG-7 dataset was introduced in [21]. A segmentation-based approach has been proposed in [5], as it requires extensive training time, which is a significant performance disadvantage. The SIFT and ORB are used commonly for the collection of local features for reverse image search (RIS), In [22] scale-invariant feature transform (SIFT) and oriented Fast and rotated BRIEF (ORB) have been used to extract features vector from image and successfully retrieve content-based images. Deep learning is becoming more and more popular across all industries due to its effective performance. In [23] a deep learning-based reverse image search system has been proposed. A variety of deep learning-based content-based image retrieval techniques have been discussed in a survey paper [24]. In contrast to CBIR, a novel technique bi-layer content-based image retrieval (BiCBIR) system has been proposed in [25].

The aforementioned strategies and methods for reverse image search (RIS) have all been developed for a single image on the canvas. It is still a research gap in that no method or model exists for reverse image searching in collage (RISiC). However, some systems have been developed for the creation of collages of different shapes, sizes, and forms. A method for automatically creating eye-catching collages that is driven by visual perception is presented in [26]. The author concentrated on how to create a collage layout that not only makes it simple to access the overall image's theme but also complies with human visual perception. Grid collages of small image collections are common and helpful in many applications, including graphic design, internet photo

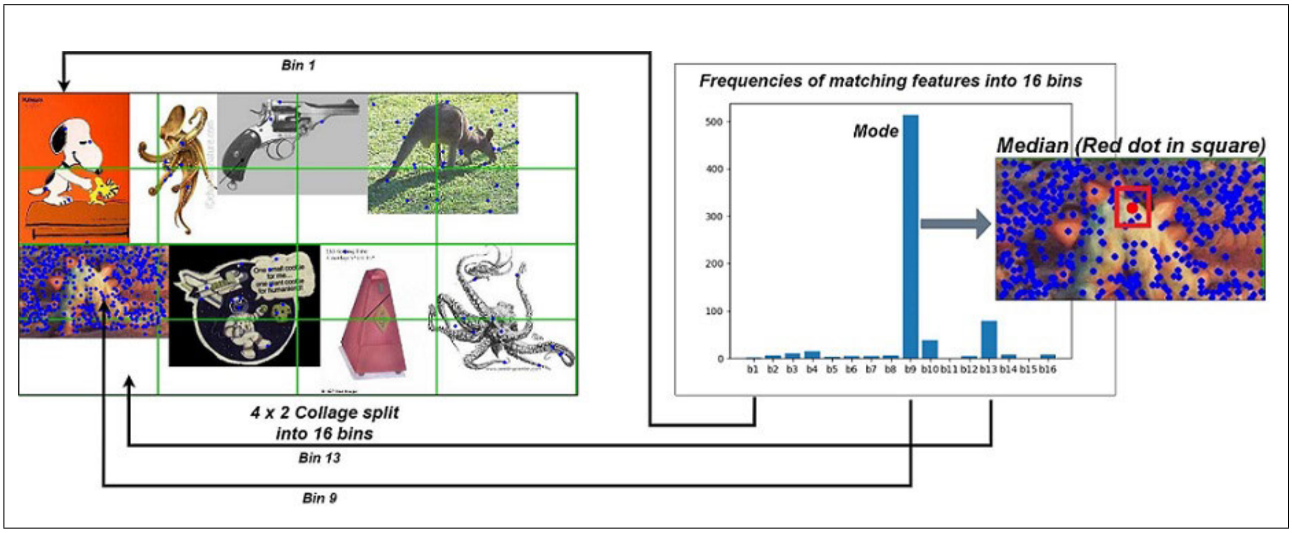


FIGURE 6. Median of Mode approach for the dense region and center orientation.

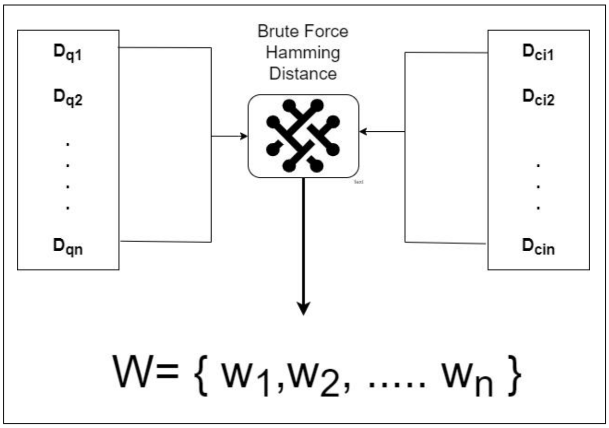


FIGURE 7. Hamming distance using Brute Force.

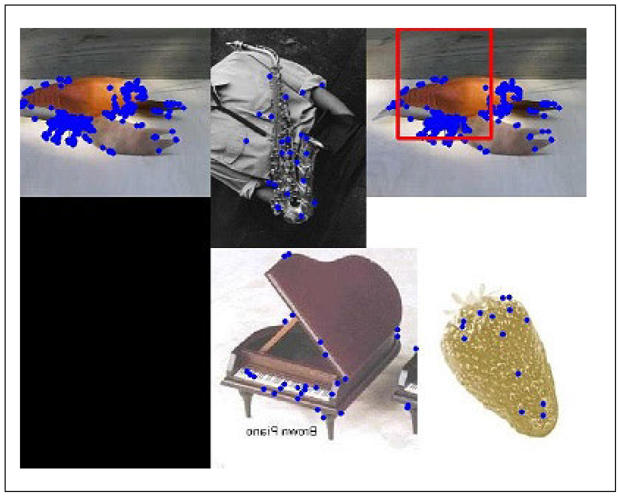


FIGURE 8. Bounding Box for ROI localization.

sharing, and managing personal albums. A Grid-style Layout collage generator model is proposed in [27].

In [28], the author has presented a content-based method for representing image collections in the collage that automatically creates a succinct and informative visual summary based on design principles and cognitive psychology. A tree shape visualization collage generation has been discussed in [29]. In [30], the author has targeted the user semantics information for a domain of sketches. The video frames have been converted into interactive collages based on conversation bubbles and comic art, proposed in [31]. The comic layout is more popular than other layouts, another comic-based collage generation has been proposed in [32]. In [33] a method has been proposed for collage generation based on size and dimension.

III. METHODOLOGY

Collage is a mosaic type, where images are used to construct a visual art piece. Reverse image search (RIS) till now has only been applied to a single image and in the context of object detection. As per our knowledge, the RIS method has not been applied to collages. Although, for a single image, RIS uses global and local features. We proposed a unified framework based on local features for reverse image search in collages (RISiC). The sequence of modules in the proposed framework is (i) extraction of local features, (ii) localization of ROI, (iii) validation of ROI, and (iv) classification.

A. MODULE-1: EXTRACTION OF LOCAL FEATURE

1) COLLAGE FEATURE DATABASE-TRAINING PART

The feature database of each collage in the dataset is created in this phase. We have proposed to use the SIFT to extract the features (keypoints and descriptors) of each collage. Then color information has been acquired only for the extracted features. Essentially, a single feature M_i consists of three components, a keypoint K, a descriptor D, and a color channel set S shown in Equation (1). The elaboration of color ratio set S in Equation (2), and the collage feature set is described in

TABLE 1. Performance of SURF, SIFT, and ORB.

Algorithms	Max no. of features (collage / query)	Accuracy	Avg. Search time of single query image in sec.	Precision@3	Precision@5	Precision@10
SIFT	12645 / 1794	90.96	0.159	0.909	0.924	0.939
SURF	19081/1382	90.64	0.156	0.906	0.921	0.936
ORB	20590/8911	89.00	0.349	0.890	0.904	0.919

Equation (3).

$$M_{i(feature)} = [K_{(keypoint)}, D_{(descriptor)}, S_{(color\ ratio\ set)}] \tag{1}$$

$$S_{(RGB\ ratio\ set)} = [\frac{CC_0}{\sum_{i=0}^{2} CC_i}, \frac{CC_1}{\sum_{i=0}^{2} CC_i}, \frac{CC_2}{\sum_{i=0}^{2} CC_i}] \tag{2}$$

$$C_i = [M_0, M_1, M_2, \ldots, M_n] \tag{3}$$

Thus finally the dataset of collage is represented as Equation (4).

$$B = [C_0, C_1, C_2, \ldots, C_n] \tag{4}$$

2) QUERY FEATURE SET- (FOR CLASSIFICATION)

This section involves sending a non-collage (single image) for matching with the collage database. A specific method has been used for feature extraction of query images.

For the procedure of query image matching a set C_q is defined in Equation 5 from the query image is extracted using the same approach described in III-A1.

$$C_q = [M_{q0}, M_{q1}, M_{q2}, \ldots, M_{qn}] \tag{5}$$

A matching procedure in which a set of query image features $M_{qn} \in C_q$ is matched from each feature belonging to $C \in B$. M consists of a keypoint, Descriptor, and RGB ratio set, for both collage and query images. Essentially, descriptor $D \in M$ is used for the closest matches in our proposed methodology. We proposed to use a Hamming distance [35] for matching based on brute force as shown in Figure (7).

Hamming Distance-based on Brute Force
1 : for every D_q in C_q repeat
2 : for every C in B repeat
3 : for every D_c in C repeat
4 : $Ham(D_q, D_c)$

The resultants are $w_n \in W$, which is a pair of closest match between $D_{q_i} \in C_q$ *with* $D_{c_j} \in C_i$ shown in equation 6.

$$W = \{w_0, w_1, w_2, \ldots, w_n\}$$
$$where\ w_n\ is$$
$$w = \{(D_{qi}, D_{cj}) | i\ and\ j = |W|\} \tag{6}$$

Despite being the closest pairs, a considerable number of pairs are false matches that are dispersed across the collage as illustrated in Figure (5).

B. LOCALIZATION OF ROI

The false matches deceive the classification algorithm and lower the performance. It is a dire need to eliminate the false matches from set W. A frequency distribution method (FDM) as shown in Figure (6) with respect to 2-dimension (x,y) is proposed for localization of Region of Interest (ROI).

Essentially, a median of mode method is used after frequency distribution for finding the dense region from W as shown in Figure (6). For marking the boundary of the dense region, a collage is split into N^2 bins such that $F = \{b_1, b_2, b_3, \ldots b_n\}$ where $n \in N^2$as shown in Figure (6). Each $w_i \in W$ is assigned a bin number based on the occurrence in the respective bin as shown in Figure (6). A single bin is assigned to multiple $w_i s$, so a mode is considered to be applied for finding the dense region. A bounding box inside the selected bin is marked by using a rectangular approach defined in equations 7.

$$P_1 = (min(x), min(y))$$
$$P_2 = (max(x), max(y)) \tag{7}$$

where P_1 and P_2 are two points for rectangular boundary, which contains the x, and y of different $w_i s \in W$. The representation of the bounding box is shown in Figure (8).

Figure (8) illustrates the empirical finding that the bounding box in question lacks an adequate number of densely-packed points.

For refinement of bounding box, new coordinates of P'_1 and P'_2 is obtained by using Equation 8.

$$P'_1 = (O.x - Q.w/2, O.y - Q.h/2)$$
$$P'_2 = (O.x + Q.w/2, O.y + Q.h/2) \tag{8}$$

where O denotes the appropriate median point w'_i with regard to a location within the dense region., and $Q.w$ and $Q.h$ are the width and height of the query image, respectively. A refined bounding box is illustrated in Figure (4). All points in refined bounding is represented $W' = \{w'_1, w'_2, w'_3, \ldots, w'_n\}$ where $W' \in W$.

C. VALIDATION AND CLASSIFICATION

In this stage, we verify the collected points $w_i s$ that were selected by structure information. For this purpose of validation, a distance, based on color intensities is calculated between query and collage whose pairs exist in refined ROI. A Manhattan distance is used to calculate the similarity defined in equation 9.

$$D_{C_i} = \frac{\sum_i |S_{q_i} - S_{c_i}|}{|W| + 1} \tag{9}$$

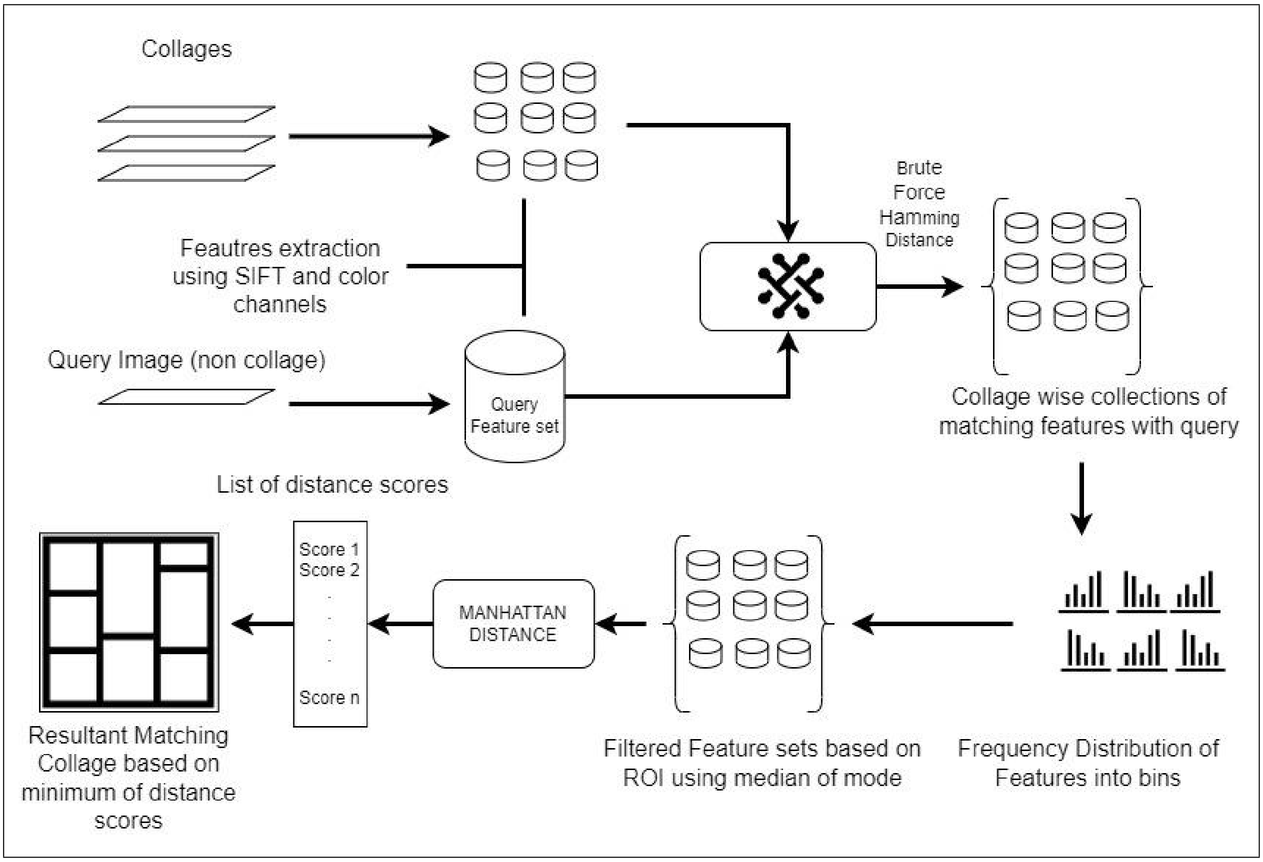


FIGURE 9. Flow diagram of proposed model.

where S_{q_i} and S_{c_i} are S sets of query and collage respectively.

Algorithm - Searching Closest Matches

for i = 1 to $|B|$; where B is a set of collage's feature
  W[i] = brute force match(C_q, C_i)
  F = freq-dist(W[i], N^2); total number of N×N cells
  b = max(F)
  P_1 = min(b.x), min(b.y) P_2 = max(b.x), max(b.y)
  R = bounding box(P_1,P_2)
  O = median(R)
  P'_1 = (O.x - Qw/2, O.y - Qh/2)
P'_2 = (O.x + Qw/2, O.y + Qh/2)
  R_N = bounding-box(P'_1, P'_2)
  W_S = W.subset(R_N) W_S includes all the pairs w_i which are located in the bonding box R_N
  distances[i] = Manhattan(C_i.S,C_q.S, W_S)
  Match = B[index(min(distances))]
[end for]

IV. EXPERIMENTS AND RESULTS

A. PREPARATION OF DATASET

To evaluate the performance of the proposed method extensive experiments were conducted on a benchmark dataset CalTech101 [34]. The dataset contains 9164 images with 101 different categories viz elephant, binoculars, mountains, buildings, cat, etc. Each category has a different number of images with variations of rotation, scaling, color, distortion, etc shown in Figure (10).

FIGURE 10. Variant sample images of the same category from dataset.

A total of 372 random distinct images are selected from the pool of 9164 for the construction of 30 collages with different dimensions and layouts as shown in Figure (11).

A detailed description of layouts is described in Table 3. During the testing process, 310 query images were selected

TABLE 2. Performance of SURF, SIFT, and ORB on flip orientations.

Algorithms	Acc. Vertical Flip	Acc. Horizontal Flip	Acc. Both Flips	Vertical Flip Precision			Horizontal Flip Precision			Both Flips Precision		
				3	5	10	3	5	10	3	5	10
SIFT	40.60	51.61	83.22	0.406	0.412	0.419	0.516	0.524	0.533	0.832	0.845	0.859
SURF	52.56	51.93	79.92	0.525	0.534	0.543	0.519	0.527	0.536	0.799	0.812	0.825
ORB	13.87	11.11	68.88	0.139	0.140	0.143	0.111	0.112	0.114	0.688	0.700	0.712

TABLE 3. Collage set layouts.

Layout	Count	Sizes
2 x 2	6	(600 x 540) (564 x 600)(511 x 600)(600 x 540)(564 x 600) (511 x 600)(600 x 457)(600 x 567)(534 x 600)
4 x 2	6	(1115 x 600)(9062 x 600)(1200 x 569) (1045 x 600)(1094 x 600)(1092 x 536)
3 x 3	6	(861 x 799)(891 x 900)(900 x 848) (888 x 900)(872 x 900)(900 x 851)
4 x 4	6	(1230 x 1200)(1412 x 1262)(1151 x 1200) (1320 x 1220)(1161 x 1162)(1161 x 1200)
5 x 5	6	(1720 x 1575)(587 x 1477)(1500 x 1435) (1640 x 1535)(1497 x 1500)(1493 x 1453)

FIGURE 11. A typical diagram of different layouts of collage.

from 372 images that are used for the construction of the collage.

B. RESULTS AND DISCUSSION

The experiments were conducted on 30 collages having different dimensions and layouts and query images, it is a single image for searching, for this research, 310 query images were used. In this study, it is proposed to subdivide the images into bins. A strong integration is developed of local feature algorithm with a localized image outperform.

1) EXPERIMENT-1: SELECTION OF BINS AND PERFORMANCE

For local feature extraction, three state-of-the-art algorithms were considered i.e. ORB, SIFT, and SURF. A wrapper-based approach was adapted to obtain the optimal number of bins for all three algorithms. The number of bins is calculated by N^2 where $N$is $2, 3, 4, \ldots$. The experiment was conducted on different values of N, SIFT, and SURF outperforms on N=4, and ORB produces the high result on N=6 as shown in Figure (12).

At the time of feature extraction, SIFT collected a less amount of features of 12645/1794 from collage and query, respectively, and the other two algorithms SURF and ORB obtained 19081/1382 and 20590/8911, receptively. In this experiment, the proposed approach using SIFT obtained a high accuracy of 90.96% utilizing a 36 number of bins. This shows that the proposed model by using SIFT outdid the proposed approach using SURF and ORB by a margin of 0.32% and 1.96%, respectively.

Precision@k is a widely-used assessment measure in the fields of information retrieval and machine learning for evaluating the performance of a classification model. This metric calculates the accuracy of predicting positive instances by the system or model in the top k predictions, where k is a specified number. For the purpose of validating the experiment, multiple values of k for precision@k (3, 5, and 10) were tested in this experiment. At k=10, SIFT demonstrated the highest precision of 93.9%, which is slightly superior to SURF and superior to the ORB by 1.5% and 3%, respectively. A detailed result of experiments are shown in Table 1

2) EXPERIMENT-2: PERFORMANCE ON COLLAGE'S FLIP

In this experiment, collages were flipped in three different orientations viz horizontal, vertical, and both horizontal & vertical. Some flipped collages are shown in Figure (13).

Three sub-experiments were performed individually on three orientations, detailed results are shown in Table 2. In the case of vertical and horizontal flips, the proposed method using SURF outperformed and achieved 52.56% and 51.93% accuracies respectively. In contrast to SURF, the SIFT of proposed methods obtained 48.60% and 51.61 % whereas ORB achieved 13.87% and 11.11% respectively. In their current configuration, none of the three methods are well-suited for handling vertical and horizontal flips. The precision@k was evaluated for three different values of k (3, 5, and 10) for all three types of flips. Through experimentation, it was observed that SURF demonstrated slightly superior outcomes in vertical flip scenarios across all three k values, as compared to SIFT and ORB. During the horizontal flip, both SIFT and SURF achieved identical results up to two decimal places, indicating a significantly

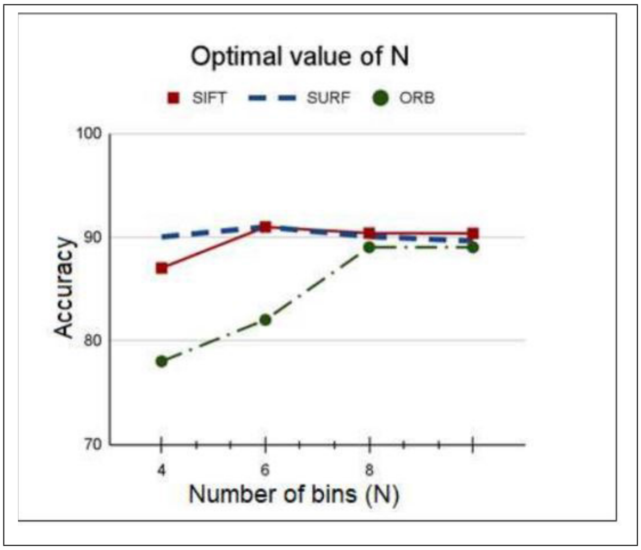


FIGURE 12. Efficiency vs bins relationship.

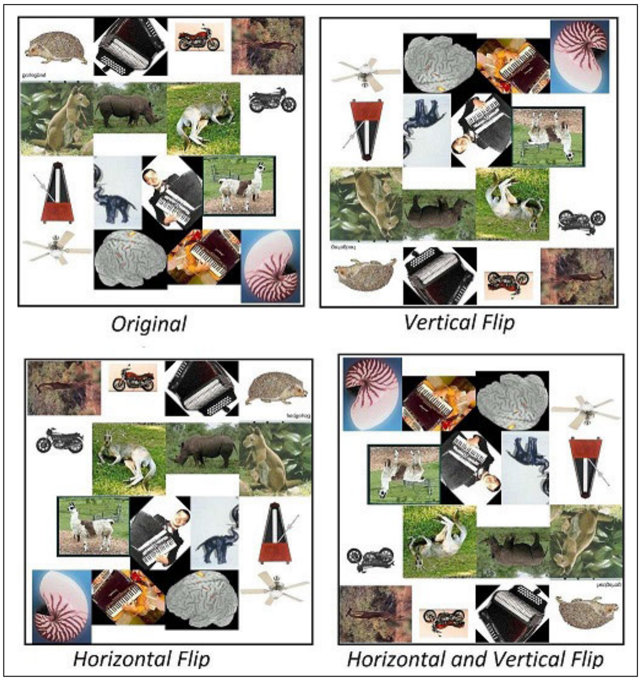


FIGURE 13. A typical image of have three different flips.

better performance compared to ORB. However, overall, the outcomes are not encouraging with the present setup. The experiment of both flips, where the highest accuracy of 85.9% by utilizing k=10 was gained by the proposed method using SIFT algorithm with a substantial margin of 3.4% and 14.7% by using SURF and ORB. Precision@k was computed, and SIFT outperformed both SURF and ORB, achieving a precision of 83.2% with k=3. This outcome was better than SURF's 79.9% and ORB's 68.80%.

3) EXPERIMENT-3: PERFORMANCE ON COLLAGE'S SCALING

Scaling was applied to collages in this experiment. Scaling was done in three different ways: 125%, 150%, and 50%.

TABLE 4. Performance of SURF, SIFT, and ORB on scale variance.

Algo.	Acc. 125% scale	Acc. 150% scale	Acc. 50% scale
SIFT	78.00	78.00	61.00
SURF	78.00	76.00	58.00
ORB	39.00	35.00	12.90

On these scales, the proposed methodology's scale in variance is tested using the respective optimal value of N for all algorithms. In the case of 125% and 150%, SIFT leads SURF by a margin of 2%, in the latter, both have an accuracy of 78% in the former. When performance is scaled down to 50%, the performance in this situation among scale-up is modest, yet SIFT still outperforms the others with a margin of 3%. The results are illustrated in Table 4

V. CONCLUSION

We have extracted the local feature from the collage, these features are used for the localization of the region of interest (ROI) through binning technique. The advantage of localization of ROI is that the proposed framework becomes a generic framework for any local features and works at a very low computational cost. It was proposed to use the median of the mod (MOM) for the ROI's localization. The advantage of using it is that the proposed framework does not reliant on the dimensions of the collage.

In the context of collage, a generic framework is proposed for the reverse image search. We have proposed to use the SIFT to extract the local features of the collage. A single feature essentially consists of three parts: a key point, a descriptor, and a color channel set. In the raw feature collection, multiple erroneous matches exist in the feature set. We have proposed to develop the ROI using the binning technique for collecting the relevant local features. By using the median of mode (MoM) approach, the additional dense region of informative characteristics is recovered. The Manhattan distance is then used to calculate the degree of similarity between the query image and the collage dataset. The SIFT algorithm has shown better results compared to other feature selection techniques i.e. SURF and ORB. Three different orientations are considered for experiments such that vertical, horizontal, and mixed vertical-horizontal flips. Extensive experiments on standard databases demonstrate the efficacy of the proposed framework. The proposed research offers a structured methodology for reverse image search which is shown to be critical for collage searching.

To further enhance the performance on rotation and scale variance, more research is required. In the future, we intend to work on a more intricate structure-based collage in order to expand the proposed framework and make it suitable for a broad scope.

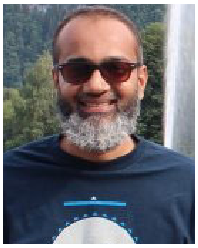
REFERENCES

[1] X. Li, J. Yang, and J. Ma, "Recent developments of content-based image retrieval (CBIR)," *Neurocomputing*, vol. 452, pp. 675–689, Sep. 2021.

[2] Y. Bitirim, S. Bitirim, D. C. Ertugrul, and O. Toygar, "An evaluation of reverse image search performance of Google," in *Proc. IEEE 44th Annu. Comput., Softw., Appl. Conf. (COMPSAC)*, Jul. 2020, pp. 1368–1372.
[3] P. Nieuwenhuysen, "Image search process in the web using image copy," *J. Multimedia Process. Technol.*, vol. 9, no. 4, pp. 124–133, Dec. 2018.
[4] M. Reilly and S. Thompson, "Reverse image lookup: Assessing digital library users and reuses," *J. Web Librarianship*, vol. 11, no. 1, pp. 56–68, Jan. 2017.
[5] R. Bibi, Z. Mehmood, R. M. Yousaf, T. Saba, M. Sardaraz, and A. Rehman, "Query-by-visual-search: Multimodal framework for content-based image retrieval," *J. Ambient Intell. Humanized Comput.*, vol. 11, no. 11, pp. 5629–5648, Nov. 2020.
[6] T. Surasak, I. Takahiro, C.-H. Cheng, C.-E. Wang, and P.-Y. Sheng, "Histogram of oriented gradients for human detection in video," in *Proc. 5th Int. Conf. Bus. Ind. Res. (ICBIR)*, May 2018, pp. 172–176.
[7] C. Song and S. Ji, "Face recognition method based on Siamese networks under non-restricted conditions," *IEEE Access*, vol. 10, pp. 40432–40444, 2022.
[8] G. Lowe, "Sift-the scale invariant feature transform," *Int. J.*, vol. 2, nos. 91–110, p. 2, Mar. 2004.
[9] L. Han, J. Wang, Y. Zhang, X. Sun, and X. Wu, "Research on adaptive ORB-SURF image matching algorithm based on fusion of edge features," *IEEE Access*, vol. 10, pp. 109488–109497, 2022.
[10] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571.
[11] E. Karami, S. Prasad, and M. Shehata, "Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images," 2017, *arXiv:1710.02726*.
[12] P. S. Hiremath and J. Pujari, "Content based image retrieval using color, texture and shape features," in *Proc. IEEE 15th Int. Conf. Adv. Comput. Commun. (ADCOM)*. Dec. 2007, pp. 780–784.
[13] E. M. Martey, H. Lei, X. Li, and O. Appiah, "Effective image representation using double colour histograms for content-based image retrieval," *Informatica*, vol. 45, no. 7, pp. 97–105, Dec. 2021.
[14] D. Srivastava, B. Rajitha, and S. Agarwal, "An efficient image classification using bag-of-words based on SURF and texture features," in *Proc. 14th IEEE India Council Int. Conf. (INDICON)*, Dec. 2017, pp. 1–6.
[15] A. Jakubovic and J. Velagic, "Image feature matching and object detection using brute-force matchers," in *Proc. Int. Symp. ELMAR*, Sep. 2018, pp. 83–86.
[16] S. Sathiamoorthy and M. Natarajan, "An efficient content based image retrieval using enhanced multi-trend structure descriptor," *Social Netw. Appl. Sci.*, vol. 2, no. 2, pp. 1–20, Feb. 2020.
[17] M. K. Alsmadi, "Content-based image retrieval using color, shape and texture descriptors and features," *Arabian J. Sci. Eng.*, vol. 45, no. 4, pp. 3317–3330, Apr. 2020.
[18] S. Somnugpong and K. Khiewwan, "Content-based image retrieval using a combination of color correlograms and edge direction histogram," in *Proc. 13th Int. Joint Conf. Comput. Sci. Softw. Eng. (JCSSE)*, Jul. 2016, pp. 1–5.
[19] P. Chhabra, N. K. Garg, and M. Kumar, "Content-based image retrieval system using ORB and SIFT features," *Neural Comput. Appl.*, vol. 32, no. 7, pp. 2725–2733, Apr. 2020.
[20] A. Kiran, S. A. Qureshi, A. Khan, S. Mahmood, M. Idrees, A. Saeed, M. Assam, M. R. A. Refaai, and A. Mohamed, "Reverse image search using deep unsupervised generative learning and deep convolutional neural network," *Appl. Sci.*, vol. 12, no. 10, p. 4943, May 2022.
[21] S. R. Dubey, "A decade survey of content based image retrieval using deep learning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 5, pp. 2687–2704, May 2022.
[22] S. Singh and S. Batra, "An efficient bi-layer content based image retrieval system," *Multimedia Tools Appl.*, vol. 79, nos. 25–26, pp. 17731–17759, Jul. 2020.
[23] Y. Cao, R. W. H. Lau, and A. B. Chan, "Look over here: Attention-directing composition of Manga elements," *ACM Trans. Graph. (TOG)*, vol. 33, no. 4, pp. 1–11, 2014.
[24] L. Tan, Y. Song, S. Liu, and L. Xie, "ImageHive: Interactive content-aware image summarization," *IEEE Comput. Graph. Appl.*, vol. 32, no. 1, pp. 46–55, Jan. 2012.
[25] X. Pan, F. Tang, W. Dong, C. Ma, Y. Meng, F. Huang, T.-Y. Lee, and C. Xu, "Content-based visual summarization for image collections," *IEEE Trans. Vis. Comput. Graphics*, vol. 27, no. 4, pp. 2298–2312, Apr. 2021.
[26] Y. Liang, X. Wang, S.-H. Zhang, S.-M. Hu, and S. Liu, "PhotoRecomposer: Interactive photo recomposition by cropping," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 10, pp. 2728–2742, Oct. 2018.
[27] L. Liu, H. Zhang, G. Jing, Y. Guo, Z. Chen, and W. Wang, "Correlation-preserving photo collage," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 6, pp. 1956–1968, Jun. 2018.
[28] Z. Yu, L. Lu, Y. Guo, R. Fan, M. Liu, and W. Wang, "Content-aware photo collage using circle packing," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 2, pp. 182–195, Feb. 2014.
[29] X. Han, C. Zhang, W. Lin, M. Xu, B. Sheng, and T. Mei, "Tree-based visualization and optimization for image collection," *IEEE Trans. Cybern.*, vol. 46, no. 6, pp. 1286–1300, Jun. 2016.
[30] X. Zheng, X. Qiao, Y. Cao, and R. W. H. Lau, "Content-aware generative modeling of graphic design layouts," *ACM Trans. Graph.*, vol. 38, no. 4, pp. 1–15, Aug. 2019.
[31] P. M. Chutel and A. Sakhare, "Evaluation of compact composite descriptor based reverse image search," in *Proc. Int. Conf. Commun. Signal Process.*, Apr. 2014, pp. 1430–1434.
[32] Y. Gan, Y. Zhang, Z. Sun, and H. Zhang, "Qualitative photo collage by quartet analysis and active learning," *Comput. Graph.*, vol. 88, pp. 35–44, May 2020.
[33] Y. Song, F. Tang, W. Dong, and C. Xu, "Non-dominated sorting based multi-page photo collage," *Comput. Vis. Media*, vol. 8, no. 2, pp. 199–212, Jun. 2022.
[34] K. T. Ahmed, S. Aslam, H. Afzal, S. Iqbal, A. Mehmood, and G. S. Choi, "Symmetric image contents analysis and retrieval using decimation, pattern analysis, orientation, and features fusion," *IEEE Access*, vol. 9, pp. 57215–57242, 2021.
[35] A. Bookstein, V. A. Kulyukin, and T. Raita, "Generalized Hamming distance," *Inf. Retr.*, vol. 5, pp. 353–375, 2002.

MUHAMMAD ZUBAIR received the M.S. degree in computer science from the Karachi Institute of Economics and Technology (KIET), Karachi, in 2020, where he is currently pursuing the Ph.D. degree. His research interests include information retrieval systems, CBIR, computer vision, video and image processing, and machine learning.

MUHAMMAD AFFAN ALIM received the B.Sc. (Hons.) and Ph.D. degrees in mathematics from the University of Karachi, in 1996 and 1999, respectively, the M.S. degree in computer science from Mohammad Ali Jinnah University, in 2007, and the Ph.D. degree in computer science from the Karachi Institute of Economics and Technology (KIET), in 2018. His research interests include pattern classification, machine learning, and data science with a special emphasis on biometrics and bioinformatics applications.



IMRAN NASEEM received the B.E. degree in electrical engineering from the NED University of Engineering and Technology, Pakistan, in 2002, the M.S. degree in electrical engineering from the King Fahd University of Petroleum and Minerals (KFUPM), Saudi Arabia, in 2005, and the Ph.D. degree from The University of Western Australia, in 2010. He was a Postdoctoral Researcher with the Institute for Multi-Sensor Processing and Content Analysis, Curtin University of Technology, Australia. In 2011, he joined the College of Engineering, Karachi Institute of Economics and Technology (KIET), Pakistan, in 2011, as an Associate Professor. He is also an Adjunct Research Fellow with the School of Electrical, Electronic and Computer Engineering, The University of Western Australia. He has authored several publications in top journals and conferences, including IEEE Transactions on Pattern Analysis and Machine Intelligence and IEEE International Conference on Image Processing. His research interests include pattern classification and machine learning with a special emphasis on biometrics and bioinformatics applications. His benchmark work on face recognition has received more than 180 citations in less than four years. He is a Reviewer of IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Image Processing, and IEEE Signal Processing Letters.

MUHAMMAD MANSOOR ALAM received the M.S. degree in system engineering and the M.Sc. degree in computer science from France, U.K., and Malaysia, and the Ph.D. degree in computer engineering and the Ph.D. degree in electrical and electronics engineering. He was the Associate Dean of the College of Computer Science and Information Systems (CCSIS) and the HoD of the Department of Mathematics, Statistics, and Computer Science Institute of Business Management (IoBM), Pakistan. He is currently a Professor of computer science. He is also with Riphah International University, Islamabad. He is enjoying 20 years of research and teaching experience in Canada, England, France, Malaysia, Saudi Arabia, and Bahrain. He has the honor to work as an Online Laureate (Facilitator) for the MSIS Program run by Colorado State University, USA, and Saudi Electronic University, Saudi Arabia. He has also established research collaborations with Universiti Kuala Lumpur (UniKL) and Universiti Malaysia Pahang (UMP). He is currently an Adjunct Professor with UniKL and supervising 12 Ph.D. students. He has done postdoctoral research in Malaysia in machine learning approaches for efficient prediction and decision-making. He has authored more than 150 research articles which are published in well-reputed journals of high impact factor, Springer Link book chapters, Scopus-indexed journals, and IEEE conferences. The Universite de LaRochelle awarded him Très Honorable (Hons.) Ph.D. degree due to his research impact during the Ph.D. degree.

MAZLIHAM MOHD SU'UD received the Diploma degree in science, the bachelor's and master's degrees in electronics, electrotechnics, and automation, and the master's degree in electronics from the University of Montpellier II, France, the master's degree in electrical and electronics engineering from the University of Montpellier, in 1993, and the Ph.D. degree in computational intelligence and decision and the Ph.D. degree in computer engineering from Université de La Rochelle, in 2007. Since 2013, he has been the President/the CEO of Universiti Kuala Lumpur, Malaysia. He has vast experience in publishing in high-quality international scientific journals and conference proceedings. He has numerous years of experience in the industrial and academic fields.