A Modified Optimal Clustering Technique for Image Categorization and Summarization

Dr. Goutam Sarker
Senior Member IEEE,
Associate Professor
Dept. of CSE,
NIT Durgapur.

Saswati Das
M. Tech Student,
Dept. of CSE,
NIT Durgapur.

Antara Pal,
M. Tech Student
Dept. of CSE,
NIT Durgapur.

ABSTRACT

Image Categorization and Summarization has become a promising task with the swift development of social networks and image sharing sites. Image categorization is the process of categorizing images into groups based on image pixel similarity. On the other hand Image summarization is the process of selecting a small set of representative image from a collection of similar images. In our present work we propose hierarchical categorization of images followed by the summarization of each categories. By hierarchical we mean, we have further categorized a resulted category (which has scope for further categorization). Categorization of images is achieved by using Optimal Clustering algorithm and summaries are generated using cluster centroids. Furthermore we have imposed compression on the summary set to yield a concise summary set. User-based evaluation demonstrates the effectiveness of the proposed work.

Keywords

Image Categorization, Image Summarization, OCA, Hierarchical OCA, Centroid, Representative Image.

1. INTRODUCTION

Since the growth of social networks and image sharing sites is on its peak, it has become difficult for users to find what they are interested in from a large amount of images in the Internet. To obtain images of a particular category with high degree of similarity may seem difficult. Image categorization achieves the task of getting similar images grouped into a category. The term image categorization refers to the labeling of images into one of a number of predefined categories. Similarly image summarization is also a key Data Mining technique to get a concise and brief representation of a large amount of similar images. The objective of Image Summarization is to select a few images of a large scale image collection to represent the image collection. The images selected in the process of summarization are called Representative images. Various multimedia applications can benefit from image summarization. In this paper we present a framework for Hierarchical Categorization followed by Summarization. Further Categorization of formed categories is called Hierarchical Categorization. We have used OCA for that very purpose. Clustering is imperative to both categorization as well as summarization. Clustering is the method of formation of groups of data from an input data set, so that elements of the same group are similar with certain similarity measure and elements in different groups are dissimilar with the identical measure. Clustering plays the most important role of generalization or induction for learning. This generalization is the process of extracting the salient features of various groups of a particular data set. After the Hierarchical Categorization, the input Data Set is divided into their Categories. These categories are then subjected to the Summarization process. The output of summarization provides us with the representative images of each category. The summary is then compressed to reduce the size of images to make the summarization result more reasonable.

Recently many approaches have been proposed for image summarization. Some are based on annotating images for semantic based image search [5,6]. The summarization of images taking the corresponding tags into account has also provided interesting results. This kind of summarization has been given the name of Hybrid Summarization [7].

The remaining content of this paper is organized as follows: In section 2 we have explained the theory of the entire system, the sequence of functions it performs and brief description of the same. In the next section we present the overview of the framework. The features and the algorithms have been illustrated in this very section. In section 4 experimental result and results are reported. In section 5 we present the salient features of user-system interaction. Finally section 6 gives the conclusion of the entire work.

2. THEORY OF OPERATION

The Image Categorization and Summarization system broadly focuses on performing two tasks:

i. Categorization of given input dataset.
ii. Summarization of resulted categories.

A brief presentation of the sequence of functions performed by the system and description of the same has been shown by the following flowchart.

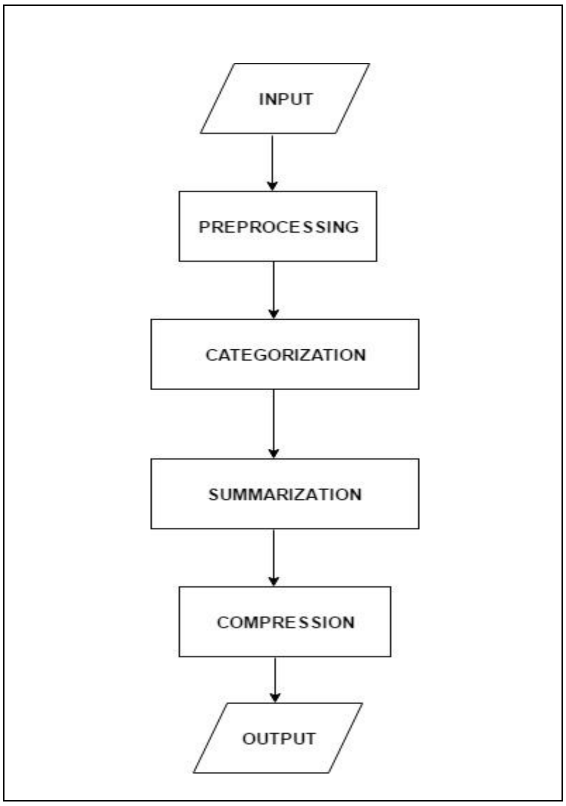

Fig 1: Schematic diagram of the system

(a) Image preprocessing : It is a common name for operation with image at the lowest level of abstraction. The aim of preprocessing is to improve the image data that eliminates unwanted distortions or enhances some features of image important for further processing. In our present work, the system resizes all images in the input dataset to a uniform dimension for further processing. The dimension chosen is 256 × 256.

(b) Image Categorization : Image Categorization does the work of grouping images based on their visual similarity. Images belonging to the same categories are highly similar whereas those belonging to different categories are highly dissimilar. OCA is used to perform grouping of similar images. We chose this algorithm since it provides us with single global and optimal solution. Here we do not require any prior data of the number of clusters. It can handle noisy and singleton clusters and hence can detect presence of any outliers. In this paper, we have proposed a novel approach of implementing hierarchical clustering using OCA to obtain even subgroups of a resulted group. A brief illustration of OCA is as follows.

Algorithm 1 : OCA

Input :
(i) Threshold
(ii) D: Dataset containing n objects.
Output : A set of clusters

Method :

Step1. Randomly select an object from D as the initial cluster. Assign the data point as the centroid of the cluster.
Step2. Repeat
Step3. Randomly choose an object not yet clustered
for each existing cluster
Step4. Calculate the distance between cluster mean and the data object.
end for
Step5. Find the minimum distance and the corresponding cluster (say C).
Step6. Compare minimum distance and threshold
If minimum distance ≤ Threshold
Step7. Assign data object to C.
Step8. Update the cluster centroid.

else
Step9. Make a new cluster containing the selected data object
Step10. Assign the data point as the centroid of the cluster.
end if
Until all data objects are clustered.

(c) Image Summarization : Image Summarization can be defined as selection of set of images that efficiently represents the visual content of a large number of similar images or images of same category. The ideal summary contains relatively very few images in order to represent a set of large image of that category. The clustering of similar images is always required for image summarization. The image chosen as summary is called the representative image. This representative image has to be obtained for each of the categories

to obtain the summary of the entire input data. After categorization of input images, the cluster of different categories are formed. The representative image of each cluster is chosen as the one which is closest to the center image of that cluster. A brief illustration of the summarization algorithm has been shown below:

Algorithm 2 : Summarization process

Input : Cluster set containing m clusters. **Output :** Set containing m representative images.

Method :
Step1. Calculate the centroid for each cluster.
for each cluster C
Step2. Find that image of C that is closest to the centroid of C.
Assign that image as the representative image of cluster C.
end for

(d) Image Compression : Image compression is applied on digital images, aimed to reduce their storage cost or transmission. In our present work we have applied compression technique on the resulted summarized image set to further reduce the cost of storage. For this purpose we have divided the image pixels into blocks of size n×n. Here value of n depends on by what factor we want to compress the image.(for e.g. If compression factor is 0.25,then n=4). Then we have applied mode function on each of the blocks to compress the image. The mode function chooses the most frequent pixel of the block and replaces it with the entire block. If no such frequent pixel is found, we find the average of value and chose that pixel to replace the block which is closest to average.

(e) Evaluation : Evaluation of categorization is necessary to access the accuracy of categorization. The evaluation measures are Accuracy, Precision, Recall and F-Score. All these measures can be calculated through confusion matrix. The confusion matrix is a very powerful tool for analyzing the correctness of categorization. Before discussing the above mentioned measures, let's suppose there are two categories of object viz. class1 and class2. For each object considered, we compare the category labeled by the system with the object's actual category.

Table. 1 : Confusion Matrix

Predicted Class \ Actual Class	class1	class2
class1	a	b
class2	c	d

The confusion matrix shown for a binary classification problem can be easily made for multiple classes in a similar manner. The evaluation measures are calculated from the above confusion matrix as follows:

(a) Accuracy: It can be stated as the percentage of object that are correctly classified by the classifier.

$$\text{Accuracy} = \frac{a+d}{a+b+c+d}$$

Where a, b, c and d are defined in the above matrix.

(b) Precision: It is the probability of actually a file being in a class if they are predicted to be classified in the same class.

$$\text{precision} = \frac{a}{a+b}$$

(c) Recall: It is defined as the probability of a file as being in a class it actually belongs to that class.

$$\text{recall} = \frac{a}{a+c}$$

(d) F-score: Harmonic mean of precision and recall.

$$\text{F-score} = \frac{2*precision\ *recall}{precision\ +recall}$$

3. OVERVIEW AND ALGORITHM FOR PRESENT WORK

We have used the following features in the process of categorization and summarization:

1. *Input Image set:* It is the input to the system containing set of images belonging to different categories. In the present work we have taken images of forest, deserts, beach, sunset, and facial images of different persons as input.

2. *Distance:* It is the distance between two images. Since we have taken colour images, each (i, j) pixel of an image contains 3 components viz. red(r), green(g) and blue(b). The r, g, b component of (i, j) pixel in image(I) can b defined as I_{rij} ,I_{gij} and I_{bij} respectively.

Let d_{ij} be the distance between the (i, j) pixels of image a and b. Then the distance between image a and b, d_{ab} can b calculated by the following expressions.

$$d_{ij} = \sqrt{\left(a_{rij} - b_{rij}\right)^2 + \left(a_{gij} - b_{gij}\right)^2 + \left(a_{bij} - b_{bij}\right)^2}$$

$$d_{ab} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{d_{ij}}{n \times n}$$

3. *Threshold:* The threshold is a value that determines whether an image belongs to a cluster. It is used as a decision parameter for the clustering algorithm.

4. *Centroid:* It is the middle of a cluster. A centroid is used to measure the cluster location.
(i,j) pixel value of a centroid is the ratio of summation of (i, j) pixel value of each image to the total number of image in the cluster.

5. *Representative Image:* This is the image chosen as the one which is closest to the centroid of a cluster in terms of distance. Each cluster after the summarization process produces one representative image. The set of representative image is the summary of an input image set.

The overall work has been presented in the following diagram:

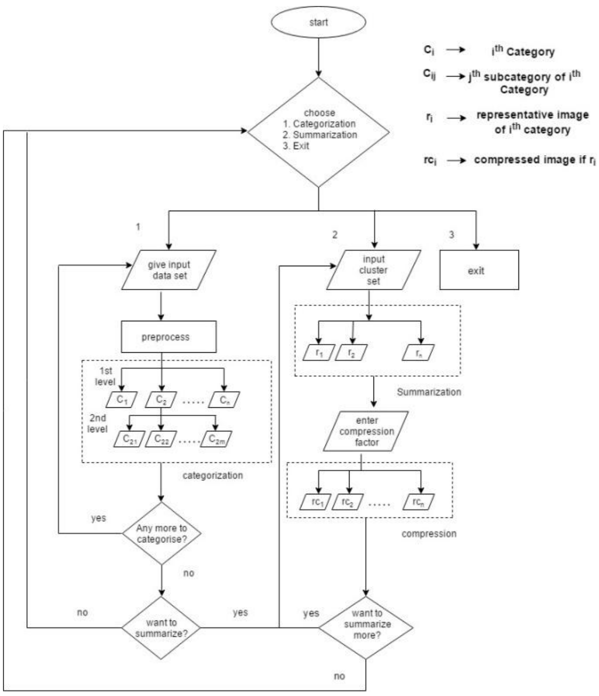

Fig 2: System Model

Algorithms for the proposed work has been illustrated below

Algorithm 3: Algorithm to determine the threshold

We have taken 5 categories (clusters) to analyze the intra cluster distance within each cluster. Let the intra cluster distance of i^{th} cluster ID_i and distance between image a and b be d_{ab}.

Input: Set of cluster C = $\{C_1, C_2, \ldots, C_m\}$. **Output:** Threshold
Method: for each cluster C_i **Step1.** Set distance to zero. for j=1 to n for k= j+1 to n **Step2.** distance = distance + d_{ij} end for end for **Step3.** ID_i = distance/n , n=no of images in cluster C_i. end for **Step4.** Assign the average of intra-cluster distance as Threshold. Threshold= $\frac{\sum_{i=1}^{m} ID_i}{m}$

Algorithm 4. Optimal Clustering Algorithm

Input: (i) Dataset D, D = $\{D_1, D_2, \ldots \ldots . D_n\}$ (ii) Threshold. **Output:** (i) Cluster set C, C = $\{C_1, C_2, \ldots, C_m\}$ (ii) Set of centroids c = $\{c_1, c_2, \ldots, c_m\}$

Method:
Step1. Randomly select an image from D as the initial cluster.
Assign pixel values of the chosen image as the centroid
of the cluster.
for each image not yet clustered
Step2. Randomly choose an image D_i from D
Begin
for each existing cluster C_j, $1 \le j \le x$, where x is the number of cluster formed so far
Step3. Calculate distance d_j between D_i and C_j.
end for
Step4. Find the minimum distance d_{min}, d_{min} =min(d_j) and the corresponding cluster C_{min}.
Step5. Compare d_{min} and Threshold
if($d_{min} \le$ Threshold)
Step6. Assign image D_i to cluster C_{min}
Step7. Update the centroid c_{min}

of C_{min} .
else
Step8. Make a new cluster C_{x+1}
Step9. Assign the pixel values of D_i as the centroid c_{x+1} of cluster C_{x+1}
end if
end
end for

Algorithm 5: Hierarchical Optimal Clustering Algorithm

Input: Dataset D, D = { $D_1, D_2, \ldots\ldots\ldots D_n$ }.
Output: (i) Cluster set C, C = {$C_1, C_2, \ldots\ldots, C_m$} for 1st level OCA
(ii) Set of clusters F = {$F_1, F_2, \ldots\ldots, F_p$} for 2nd level OCA

Method:
Step1. Apply OCA on dataset D.
Step2. Randomly choose one facial image (I) from D.
for each cluster C_j
Step3. Find the distance d_j between I and centroid of cluster C_j
end for
Step4. Find the minimum distance d_{min} = min(d_j) and the corresponding cluster C_{min}
Step5. Apply OCA on C_{min}

Algorithm 6. Algorithm for summarization

We have passed the set of cenroids c = {$c_1, c_2, \ldots\ldots, c_m$} calculated in OCA as input to the summarization process.

Input: (i) Cluster set C, C = {$C_1, C_2, \ldots\ldots, C_m$}
(ii) Set of centroids c = {$c_1, c_2, \ldots\ldots, c_m$}
Output: Set of representative images R = {$r_1, r_2, \ldots\ldots, r_m$}

Method:
for each cluster C_j
for each image in C_j
Step1. Find the distance d_i between C_j and the i^{th} image
End for
Step2. Find the image with least distance value , d_{min} =min(d_i) where i is the number of image in c_j
Step3. Assign d_{min} as the representative image r_j of cluster C_j
End for

Algorithm 7. Algorithm for compression

Let the size of the image be n×n , i^{th} block be denoted as B_i and compression factor be a factor of n.

Input: (i) Set of representative images R = {$r_1, r_2, \ldots\ldots, r_m$}
(ii) Compression Factor(cf)
Output: Compressed set of R, RC = { $rc_1, rc_2, \ldots\ldots, rc_m$ }

Method:
Step1. Compute number of blocks, no. Of blocks = $\frac{n \times n}{cf \times cf}$
for each r_i
for each B_j in r_i
Step2. If there exist a most frequent pixel
Step3. Replace the pixel with the B_j
else
Step4. Find the average of pixels in B_j
Step5. Calculate the distance of each pixel in B_j from average
Step6. Replace the block with the pixel value having the least distance from average.
end if
end for
end for

4. Experiment Result and Performance Evaluation

The hierarchical image categorization and summarization system is evaluated using an Intel(R) Xeon(R) CPU E5506 @2.13GHz having 12GB RAM and Windows 7 Ultimate 64-bit OS. The image files of different categories like forest, sunset, desert, beach etc has been collected from internet. We have used benchmark data(http://cycl.mit.edu/database.htm) as input to the system. We have also taken benchmark facial data of different person along with the above mentioned categories to be used for next level classification.

We have implemented 2 levels of categorization viz. 1st and 2nd . In 1st level categorization the input set is an image set containing images of different category such as forest, sunset, desert, beach and facial images of different persons. In 2nd level categorization the facial image group formed as result of former categorization is then broken down to subcategories with facial images of particular person. The result for 1st level categorization and 2nd level categorization in a tabular form is shown below:

Table 2. 1st Label Categorization Result

Category	Size	Correct	Wrong	Misclassifications	Time (sec)	Acuuracy(%)
Sunset	18	17	1	1	207.97	99.47
Beach	67	67	0			
Desert	16	16	0			
Face	33	33	0			
Forest	55	55	0			

Abbreviation:
Size: Size of Category
Correct: No. of Images Correctly Categorized
Wrong: No of Image Wrongly Categorized
Misclassifications : Total No of Misclassifications
Time: Time for categorization

Table 3. 2nd Label Categorization Result

Category	Size	Correct	Wrong	Misclassifications	Time (sec)	Acuuracy(%)
Person1	9	9	0	0	35.32	100
Person2	10	10	0			
Person3	8	8	0			
Person4	6	6	0			

Abbreviation:
Size: Size of Category
Correct: No. of Images Correctly Categorized
Wrong: No of Image Wrongly Categorized
Misclassifications : Total No of Misclassifications
Time: Time for categorization

The performance evaluation of Categorization(1st Level) in the form of fuzzy confusion matrix with all the performance measure has been presented below:

Table 4. Confusion Matrix

Predicted Class

	Sunset	Beach	Desert	Face	Forest
Sunset	17	1	0	0	0
Beach	0	67	0	0	0
Desert	0	0	16	0	0
Face	0	0	0	33	0
Forest	0	0	0	0	55

Accuracy = 99.47%

category	Precision(%)	Recall(%)	f-score(%)
Sunset	100	94.4	97.14
Beach	98.5	100	99.25
Desert	100	100	100
Face	100	100	100
Forest	100	100	100

After Categorization the set of categories are then subjected to the summarization process. For each category in the set a representative image is produced.

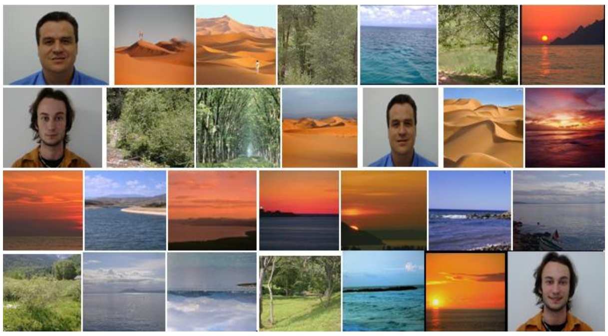

Fig 3. Portion of Input Data

Fig 4. A 5 image summary of input set of 5 categories

5. SALIENT FEATURES OF USER SYSTEM INTERACTION

```
>> menu
MAIN MENU
1. CATEGORIZATION
2. SUMMARIZATION
3.EXIT
Enter a choice: 1
Input the image set for categorization: ds
Elapsed time is 207.970289 seconds.
the cluster having facial images is cluster4
```

```
Elapsed time is 35.319389 seconds.
Any more image set for categorization? y/n : n
Do you want to summarize? y/n :y
Enter the cluster set you want to summarize: result
Elapsed time is 21.695808 seconds.
enter the compression factor:0.5
Elapsed time is 4.985268 seconds.
Any more set to compress? y/n :n
Do you want to summarize? y/n :y
Enter the cluster set you want to summarize:
result_face
Elapsed time is 3.692402 seconds.
enter the compression factor:0.25
Elapsed time is 2.063974 seconds.
Any more set to compress? y/n :n
Any more set to summarize? y/n :n
MAIN MENU
1. CATEGORIZATION
2. SUMMARIZATION
3.EXIT
Enter a choice: 3
>>
```

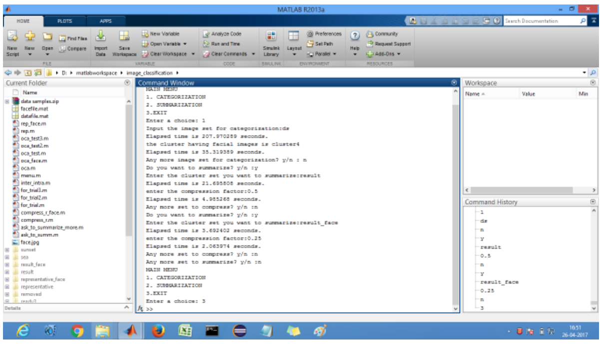

Fig 5. Snapshot of User System Interaction

6. CONCLUSION

In the present work we have performed Hierarchical Categorization of images followed by the Summarization of categories formed. Input is an image set containing images of various categories collected from internet. Categorization is achieved through Optical Clustering Algorithm (OCA). Summarization does the task to choosing a representative image of each category by computing distance between images and their respective Cluster Centroid. The one which is closest to centroid is selected as the representative for that category. Further we have performed compression on the resulted summaries using the mode function.

The summaries can be further improved by taking other parameters into account such as tags and salient (uncommon) features etc. Such kind of summarization is called Hybrid Summarization. In future these aspects can be added to our system to further improvise the quality and relevancy of summaries.

The time for categorization and summarization is affordable. Accuracy is also acceptable. The result of the system is affective on a real internet collection of images.

7. REFERENCE

[1] Hongliang Yu, Zhi-Hong Deng, Yunlun Yang and Tao Xiong. 2014. A Joint Optimization Model for Image Summarization Based on Image Content and Tags. Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence.

[2] Radhakrishna Achanta, Appu Shaji, Pascal Fua, Sabine Susstrunk. 2009. Image Summaries using Database Saliency. The 2nd ACM SIGGRAPH Conference and Exhibition in Asia (SIGGRAPH Asia), Yokohama, Japan, December 16-19, 2009 - ACM SIGGRAPH ASIA 2009.

[3] I. Simon, N. Snavely, S.M. Seitz, Scene Summarization for online image collections, in: Proceedings of the IEEE 11th International Conference on Computer Vision, 2007, pp. 1–8.

[4] Hao Xu, Jingdong Wang, Xian-Sheng Hua, Shipeng Li. 2011. Hybrid Image Summarization. Proceeding MM '11 Proceedings of the 19th ACM international conference on Multimedia Pages 1217-1220.

[5] J. Tang, R. Hong, S. Yan, T.S. Chua, G.J. Qi, R. Jain, 2011. Image, Annotation by knn-sparse graph-based label propagation over noisily-tagged web images, ACM Trans. Intel. Syst. Technol. 2 (2).

[6] J. Tang, H. Li, G.J. Qi, T.S. Chua, 2010. Image annotation by graph-based inference with integrated multiple/single instance representations, IEEE Trans. Multi- media 12 (2).

[7] Minxian Li, Chunxia Zhao , Jinhui Tang, 7 November 2013, Hybrid image summarization by hypergraph partition. Neurocomputing, Volume 119, Pages 41–48.

[8] Sarker, G.(2000),A Learning Expert System for Image Recognition, Journal of The Institution of Engineers (I), Computer Engineering Division.,Vol. 81, 6-15.

[9] Sarker, G.(2002),SELECTWIN – A New K-WTA Optimal Learning Technique for Pattern Classification and Recognition, Journal of The Institution of Engineers(I), Computer Engineering Division.,Vol. 83, 16 – 21.

[10] Sarker, G.(2010),A Probabilistic Framework of Face Detection , International Journal of Computer, Information Technology and Engineering (IJCITAE),4(1), 19-25.

[11] Sarker, G.(2011),A Multilayer Network for Face Detection and Localization, International Journal of Computer, Information Technology and Engineering (IJCITAE), 5(2), 35-39.

[12] Sarker, G.(2012),A Back Propagation Network for Face Identification and Localization ,International Journal of Computer, Information Technology and Engineering (IJCITAE),6(1), 1-7.

[13] Sarker, G.(2012),An Unsupervised Learning Network for Face Identification and Localization ,International Journal of Computer, Information Technology and Engineering (IJCITAE),6(2), 83-89.

[14] Sarker, G., Roy, K. (2013) ,A Modified RBF Network With Optimal Clustering For Face Identification and Localization ,International Journal of Advanced Computational Engineering and Networking, ISSN: 2320-2106.,1(3), 30 -35.

[15] Sarker, G., Roy, K.(2013),An RBF Network with Optimal Clustering for Face Identification ,Engineering Science International Research Journal (ISSN) – 2300 – 4338 ,1(1),ISBB: 978-93-81583-90-6, 70-74.

[16] Sarker, G.(2013),An Optimal Back Propagation Network for Face Identification and Localization ,International Journal of Computers and Applications (IJCA),ACTA Press, Canada.,35(2).,DOI 10.2316 / Journal .202.2013.2.202 – 3388.

[17] Sarker, G., Sharma. S.(2014),A Heuristic Based RBFN for Location and Rotation Invariant Clear and Occluded Face Identification ,International Journal of Computer

Information Technology and Engineering (IJCITAE), Serials Publications ,8(2),109-118.
[18] Sarker, G.(2014),A Competitive Learning Network for Face Detection and Localization, International Journal of Computer Information Technology and Engineering (IJCITAE), Serials Publications, 8(2),119-123.
[19] Sarker, G.(2002),A Semantic Concept Model Approach for Pattern Classification and recognition ,28th Annual Convention and Exhibition IEEE – ACE 2002.,December 20-21 2002 , Science City ,Kolkata, 271 – 274.
[20] Sarker, G.(2005),A Heuristic Based Hybrid Clustering for Natural Classification in Data Mining ,20th Indian Engineering Congress, organized by The Institution of Engineers (India), December 15-18, 2005, Kolkata, INDIA, paper no. 4.
[21] Sarker, G.(2011),A Back propagation Network for Face Identification and Localization ,2011 International Conference on Recent Trends in Information Systems (ReTIS–2011) held in Dec. 21-23, Kolkata, DOI: 10.1109/ReTIS.2011.6146834, pp 24-29.
[22] Sarker, G.(2012), 2012. An Unsupervised Learning Network for Face Identification and Localization ,2012 International Conference on Communications, Devices and Intelligent Systems (CODIS) Dec. 28 and 29, 2012, Kolkata, DOI: 10.1109/CODIS.2012.6422282, pp 652-655.
[23] Sarker, G., Roy, K.(2013),An RBF Network with Optimal Clustering for Face Identification, International Conference on Information & Engineering Science – 2013(ICIES -2013), Feb. 21-23 2013, organized by IMRF, Vijayawada, Andhra Pradesh, pp – 70-74.
[24] Sarker, G., Roy, K.(2013), 2013. A Modified RBF Network with Optimal Clustering for Face Identification and Localization, International Conference on Information & Engineering Science – 2013(ICIES -2013), Feb. 21-23 2013, organized by IMRF, Vijayawada, Andhra Pradesh pp 32-37.
[25] K. Roy, Sarker G. (2013), 2013. A Location Invariant Face Identification and Localization with Modified RBF Network ,International Conference on Computer and Systems ICCS-2013, 21-22 September, 2013, pp – 23-28, Bardhaman.
[26] Sarker, G.,Sharma, S. (2014), 2014. A Heuristic Based RBFN for Location and Rotation Invariant Clear and Occluded Face Identification ,International Conference on Advances in Computer Engineering and Applications, ICACEA – 2014, with IJCA), pp – 30-36.