JOURNAL OF CRITICAL REVIEWS

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

ANALYSIS ON CONTENT BASED IMAGE RETRIEVAL TO IMPROVE THE PERFORMANCE AND ITS NOVEL APPROACH

Somala Ramakishore[1], Dr.P. Karpagavalli[2]

[1]Research Scholar, Dept. of Electronics and Communication Engineering, Sri Satya Sai University of Technology & Medical Sciences, Sehore, Bhopal-Indore Road, Madhya Pradesh, India.
[2]Research Guide, Dept. of Electronics and Communication Engineering, Sri Satya Sai University of Technology & Medical Sciences, Sehore, Bhopal Indore Road, Madhya Pradesh, India.

Received: 21.04.2020 Revised: 22.05.2020 Accepted: 19.06.2020

ABSTRACT: The issue of content based image retrieval in dynamic environment is tended to. It isn't doable for frameworks that examine images continuously where the images are put away or included a progressing premise. - In the realm of digital image handling since the requirement for content-based image retrieval has expanded with increase size and volume of digital images. In this paper we execute the successful structure which is utilized to recover most comparable images from enormous images database for the image gave by the user. We proposed methodology, an image present by a lot of areas, while correlation of images are representing, each image speak to by a histogram, subsequently the estimation of the district correspondence change into a histogram coordinating issue. Also, by utilizing and image separation ideas, the distinction between images acquired. Test results show that the proposed histogram image coordinating execution is adequate.

KEYWORDS: Dynamic, Digital, Image, Retrieval, Comparable.

 DOI: http://dx.doi.org/10.31838/jcr.07.14.178

I. INTRODUCTION

Content based image retrieval (CBIR), otherwise called inquiry by image content (QBIC) and content based visual data retrieval (CBVIR) is the use of PC vision strategies to the image retrieval issue. That is, the issue of scanning for digital images in huge database. Content based image retrieval (CBIR) was proposed in the 1990's content based image retrieval (CBIR) is a significant and noticeable region in image preparing because of its assorted application in entomb sight and sound, organizations image documents and media image chronicles.

Two fundamental parts in visual-based are feature extraction and feature reduction with use of certain likeness measurement. In CBIR frameworks, image content incorporates the color, texture, and shape. They are extracted for resulting utilizes during ordering and retrieval process. These features are viewed as low level and are generally displayed in numerical structures and doled out in feature vectors.

In any case, most of users are in every case more inspired by explicit districts as opposed to the whole image. Therefore, the greater part of the current CBIR frameworks are district based where the features are extracted distinctly from the Regions of Interest (ROI). Portrayal of the images at local level is distinguished to be nearer to human discernment framework.

The fundamental need of any image retrieval model is to search and arrange the images that are in a visual semantic relationship with the query given by the user. Most of the search engines on the Internet retrieve the images on the basis of text-based approaches that require captions as input. The user submits a query by entering some text or keywords that are matched with the keywords that are placed in the archive. The output is generated on the basis of matching in keywords, and this process can retrieve the images that are not relevant.

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

II. LITERATURE REVIEW

Zin et al (2018) Content-based Image Retrieval (CBIR) helps radiologist to distinguish comparable medical images in reviewing past cases during finding. Albeit a few algorithms have been acquainted with extract the content of the medical images, the procedure is as yet a test because of the idea of the feature itself where a large portion of them are extracted in low level structure. Notwithstanding the dimensionality reduction issue brought about by the low-level features, current features are likewise lacking to pass on the semantic importance of the images.

This paper surveys the ongoing works in CBIR that endeavors to diminish the semantic gap in extracting the features from medical images, unequivocally for mammogram images. Approaches, for example, the utilization of importance input, ontology just as AI algorithms are abridged and examined.

Prof. Manish Shriwas, Payal Deotale, Ankita Naik, Komal Wankhade, Arti Marbade (2019) In the world of digital image handling since the requirement for content-based image retrieval has expanded with increase size and volume of digital images. In this paper we actualize the compelling structure which is utilized to recover most comparable images from enormous images database for the image gave by the user.

We proposed methodology, an image present by a lot of locales, while examination of images are posing, each image speak to by a histogram, subsequently the estimation of the district correspondence change into a histogram coordinating issue. Moreover, by utilizing and image separation ideas, the contrast between images got. Trial results show that the proposed histogram image coordinating execution is satisfactory.

Popova and Neshov (2013) endeavored to improve the retrieval adequacy and exactness of image search in tremendous databases. Their works prescribed two arrangements of various feature blends that perform well for medical image classifications, IRMA. The best three features utilized in the examination were Color Layout, Edge Histogram and DCT Coefficients, which were altogether consolidated for a higher image re-unimportant positioning than others that were based on singular features. The outcomes restored a 14.49% improvement of the retrieval MAP (50.68% for joined features Set). Later on, the examination is set to explore new features and apply proper loads for the specific features utilized in the mix.

In taking care of boisterous image in Relevance Feedback CBIR frameworks, Huang et al., (2013) proposed a two-advance methodology. The initial step was to apply a loud disposal algorithm. The subsequent advance was to embrace Fuzzy Support Vector Machine (FSVM) to re-rank the images. Boisterous Smoothing Relevance Feedback (NS-RF) was accounted for to yield better retrieval execution when contrasted with the Support Vector Machine-Noisy-Smoothing-Relevance Feedback (SVM-NS-RF).

The avocation was that NS-RF can additionally deal with the boisterous images by considering diverse applicable images with various importance probabilities in the image re-positioning technique, though SVM-NS-RF treated the held pertinent image similarly. What's more, the exploration noticed that the impact of boisterous disposal algorithm in NS-RF is delicate to the slack factor in the proposed equation. Subsequently, it is proposed to direct further examination on the incentive so as to improve the uproarious end algorithm later on.

Da Silva et al., (2011) led three tests utilizing mammograms taken from the University of South Carolina and University of Sao Paulo. This strategy of feature determination upgraded comparability search precision and brought down generously information dimensionality, which therefore improved access methods effectiveness and the CBIR framework.

For future work, it is suggested that another nearby search into GA and the collaboration between channel based methods and the GA wrapper-based method in CBIR frameworks can be acquainted with improve the productivity of the proposed method. Other than those, all literary data on patients' clinical history can be fused into the search mechanism.

III. PROPOSED WORK METHODOLOGY

The flow diagram of the proposed work is given in Figure 1 which delineates all the processing steps utilized in this work as following:

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

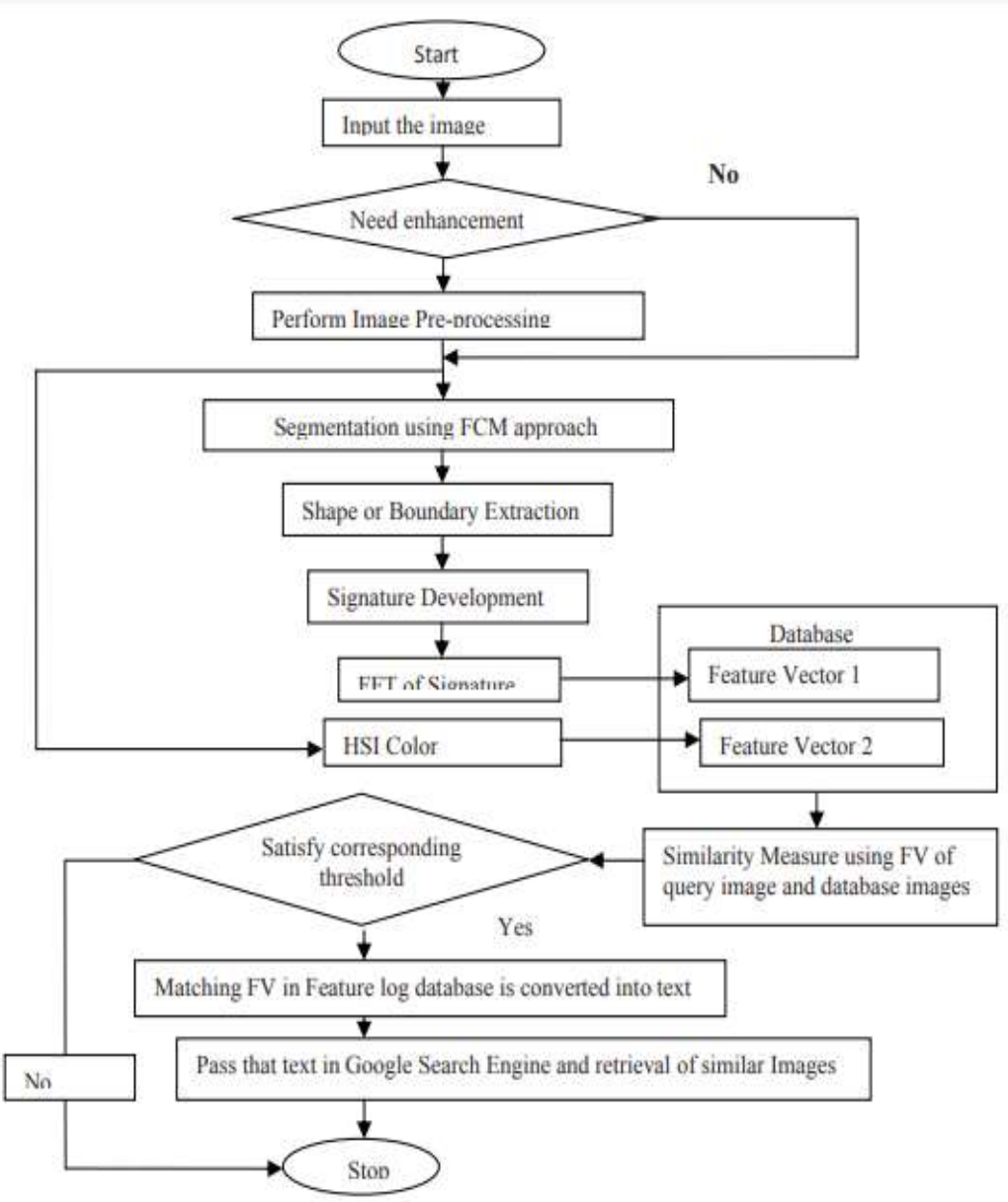


Figure 1: Flow Diagram of proposed work

- **Extracting Visual Content**

Visual content of images involves color, texture and shape. For compelling retrieval of comparative ages visual content of images structure best feature vector for correlation. In this proposed work shape and color is utilized to speak to visual content of the image.

1. Shape Feature

In the proposed work Fuzzy C-Mean clustering algorithm is utilized for division. After effective division limit of the extracted item is changed over into signature whose Fast Fourier Transform is determined. FFT give the cluster vector relating to the quantity of districts acquired after division and is put away in database as first feature vector. Shape is also considered as an important low-level feature as it is helpful in identification of real-world shapes and objects. Zhang and Lu presented a comprehensive review of the application of shape features in the domain of image retrieval and image representation. Region-based and contour-based are the main classifications of shape features. Figure 4 presents a basic overview of the classification of shape features. Trademark-based image retrieval is one of the specific domains where shape features are used for image representation.

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

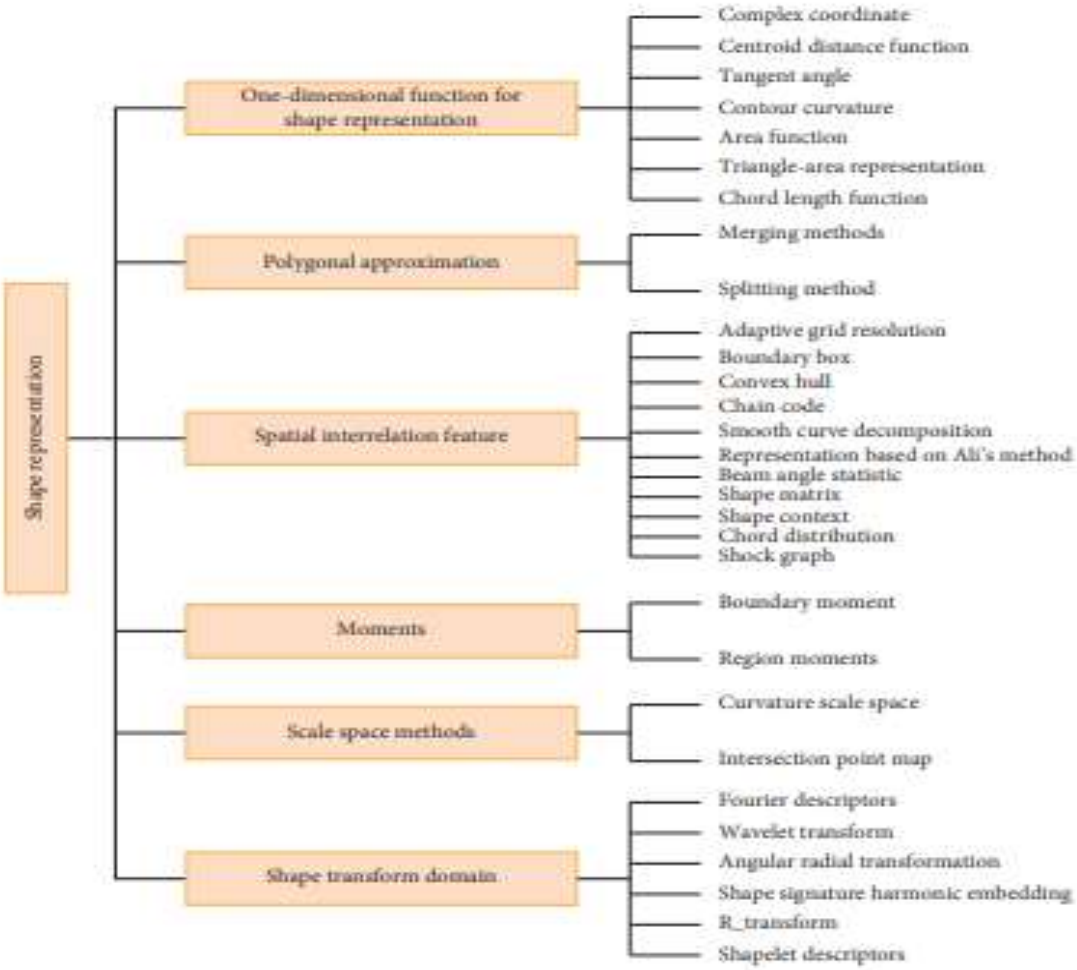


Figure 2: An Overview of Shape-based Feature Extraction Approaches

i) Segmentation

Step 1: For inquiry image, discover the clusters and its comparing focus utilizing FCM clustering approach.
Stage 2: Apply connected part technique so as to get connected areas or veil utilizing section cover and number of fragments as input. For connected segment methodology we utilize 4-connectivity.

ii) Signature Development

Step 1: For signature improvement we utilize the real image, region cover and number of regions acquired utilizing division as inputs.
Stage 2: Find centroid, limit, gap, mark of the region cover.
Stage 3: Find out the mark of every region by figuring separation between limit components and centroid of the region.
Stage 4: As this mark is variation to interpretation so we figure the shape number of every region and afterward mark of that region veil is created.
Stage 5: Generate the mark of the query image by consolidating the mark of the region veil.
Stage 6: After figuring mark of the image discover the FFT utilizing the formula characterized in Equation (1).
For length N input vector x, the DFT is a length N vector X, with elements

$$X(k) = \sum_{n=1}^{N} x(n)*\exp(-j*2*pi*(k-1)*(n-1)/N), \quad 1 <= k <= N \qquad (1)$$

X(k) is the array vector which is used to form first Feature Vector F_1.

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

2. Color Feature

Colour features are designed based on colour space. Colour representations that are commonly used in RBIR are RGB, LAB and LUV, HSV (HSL), YCrCb, and Hue-Min-MaxDifference (HMMD). All of these are designed to be closer to human perception. They are represented using colour-covariance matrix, colour histogram, colour correlogram, dominant colour descriptor, colour moments, colour coherent vector or colour co-occurrence matrix. Colour features are reported to be very stable and robust due to their resilience towards rotation, translation and changes during scaling [12]. However, most medical images are in grayscale except for colour photographs that are used for diagnosis in ophthalmology, pathology and dermatology. Researcher used elective colour features to retrieve similar skin lesions when comparing to an actual case to verify the diagnosis or confer with similar symptoms. The colour feature representations used are limited to rank matrix learning vector quantisation (LiRaMLVQ) and a Large Margin Nearest Neighbor (LMNN) approach. They performed a comparison on the retrieval rates achieved with extracted and original features for eight different standard colour spaces, and the research has achieved a significant result for each colour space

We use HSI color model where hue and saturation component of image form second feature vector F_2.

- **Similarity Measure**

Similarity measures are used to identify the similarity between query image and database images. Thus, a good similarity measure would need to be employed in order to retrieve the most similar and relevant images. Many similarity measures have been introduced by researchers to improve the effectiveness and efficiency in CBIR. The measurement required for the similarity comparison is determined by the selected feature vectors. The extraction of the multi-dimensional points' features from the images allows the calculation of the distances between the multi-dimensional points. Popular metrics that are utilised to gauge the distance between two points in multi-dimensional space are Euclidean distance, Weighted Euclidean distance, Manhattan distance, Cross Correlation distance, Minimum Mean distance rule and Statistical distance.

Another similarity measure is performed using graph-based approach. Graph representations have received more attention recently because it is an elective tool to represent relational information. Proposed a new method to determine the similarity between histological images through graph-theoretic description and matching in retrieving histological images from larger databases. In contrast to the above methods that directly measure the similarity in terms of image information alone, Classifier-based similarity measures do not directly measure similarity in terms of image information as compared to the method mentioned earlier. It uses the classification of a query image based on a fixed set of predetermined labels to assess similarity.

Step 1: In the wake of computing FVs of query image we will think about the FVs of database images which are as of now processed and put away in database. Presently, contrast between the FVs determined and a threshold (th) is chosen. Matches threshold is changed over into content of comparing image and afterward content is passed in browser for searching in dynamic environment.

Similarity measure can be discovered utilizing separation measure between FVs of query images and FVs of database images utilizing the formula demonstrated as follows:

$$\text{Difference} = (1/N) \sum(Fn\text{-}F'n)^2 \quad n=1,2 \quad \textbf{(2)}$$

Where F= FVs of query image and

F'= FVs of database images stored in log files (extracted sequentially)
N=total number of elements in an array vector of FV.
Step 2: On the off chance that contrast between feature vector matches to a particular threshold, object get perceived and changed over into content. This content contains the name of the coordinating image, got through information base.
Stage 3: Pass that content in browser for searching in Google or Yahoo search motor.

IV. RESULTS AND DISCUSSION

We have taken 100 images in the database and calculated their feature vectors and tested the proposed method with 20 test images.

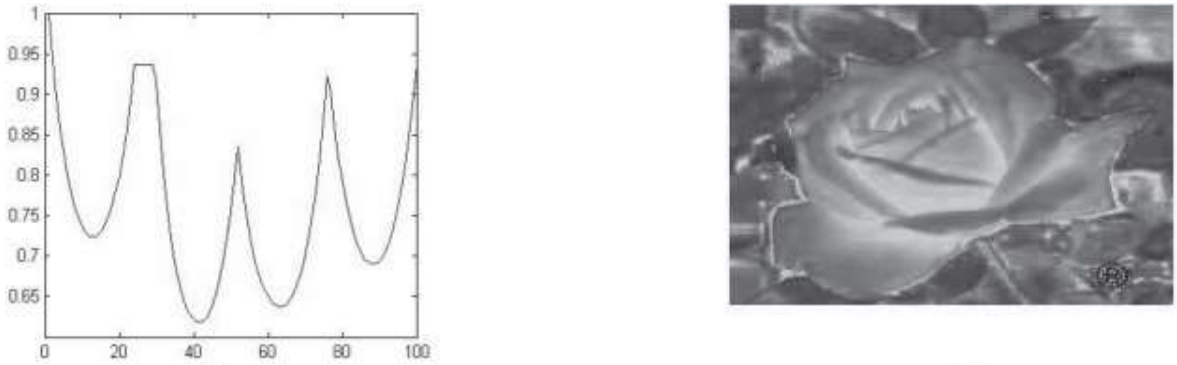
Step 1: Input the query image shown in figure 3.
Step 2: The resultant image after segmentation for given input image is shown in figure 3(b):

ISSN- 2394-5125 VOL 7, ISSUE 15, 2020

(a) (b)

Figure 3: (a) Input Yellow Flower Image, (b) Segmented Image Using FCM Approach

Step 3: First Feature Vector i.e. Signature is developed for image shown in figure 4(a).
Step 4: Second Feature Vector is HSI color processing of input image, HSI image is shown in figure 4(b).



(a) (b)

Figure 4: (a) Signature of segmented Image, (b) HSI Component of Input Image

Step 5: First feature vector after applying FFT(y1) and Second feature vector (h2) for input image is:
y1=[8.6664 0.3758 0.2005 0.1440 0.1247 0.1220 0.1247 0.1440 0.2005 0.3758];
h2=0.7364;

For compelling Retrieval we need to process the image and store their comparing Feature Vectors in the log file. Query Image is processed and Feature vector is gotten. On the off chance that they got Feature Vector coordinates the Log File; at that point comparing query image is changed over into text and went to a search engine which is demonstrated as follows.

Stage 6: Resultant images after execution of proposed algorithm based on feature vector correlation:

(a) (b)

Figure 5: (a) Resultant Search Result for Query Image ” Yellow Rose” ; (b) Similar Yellow Flower having same Result

V. CONCLUSION

In this proposed work, FCM approach is utilized. The fuzzy C-means algorithm (FCM) sums up the hard c-means algorithm to enable a point to halfway have a place with different clusters. Content based image retrieval is a method to recover increasingly applicable images. Recover comparative images just is a standing issue in digital image processing. The presentation of CBIR framework is improved by presenting pertinence criticism procedures in the framework. We propose a feature extraction approach, called "Fast Fourier Transform", to extract the invariant array vector from signature acquired after segmentation of the image. The inferred array vector is later used to frame a first feature vector alongside the HSI segment of color image, which will be utilized as second feature vector. Consequences of the investigations on the rose flower database as indicated are the proof that the new algorithm can perform far superior to customary methods of Content based Image Retrieval.

VI. REFERENCES

[1] Zin, "Content-Based Image Retrieval in Medical Domain: A Review", *Journal of Physics: Conf. Series* 1019 (2018) 012044, 2018.

[2] Prof. Manish Shriwas, Payal Deotale, Ankita Naik, Komal Wankhade, Arti Marbade, "A Novel Approach For Content based Image Retrieval by using Histogram", *International Research Journal of Engineering and Technology (IRJET)*, 2019.

[3] Popova, A.A., & Neshov, N N. (2013). Combining Features Evaluation Approach in Content-Based Image Search for Medical Applications.

[4] Huang, W., Zhang, P., & Wan, M. (2013). A novel similarity learning method via relative comparison for content-based medical image retrieval. *Journal of digital imaging*, 26(5), 850- 865.

[5] Da Silva, S.F., Ribeiro, M.X., Neto, J.D.E.B., Traina-Jr, C., & Traina, A.J. (2011). Improving the ranking quality of medical image retrieval using a genetic feature selection method. *Decision Support Systems*, 51(4), 810-820.

[6] WangXing-yuan , ChenZhi-feng, YunJiao-jiao, An effective method for color image retrieval based on texture, Computer Standards & Interfaces Volume 34 (2012) Pages 31–35.

[7] Nikita Sethand Sonika Jindal, A Review on content based Image retrieval, *International journal of computers and Technology* (January 2017).

[8] Eduardo Valle, Matthieu Cord, Sylvie Philipp-Foliguet, and David Gorisse, Indexing Personal Image Collections: A Flexible, Scalable Solution, *IEEE Transactions on Consumer Electronics*, Vol. 56, No. 3, 2010, pp. 1167-1175.

[9] Hao Xia, Pengcheng Wu, Steven C.H. Hoi & Rong Jin, Boosting Multi- Kernel Locality-Sensitive Hashing for Scalable Image Retrieval, *ACM on SIGIR'*12, 2012.

[10] X. Li,W. Hu, Z. Zhang, & H.Wang, Heat Kernel based local binary pattern for face presentation, *IEEE Signal Process. Lett*, vol. 17, no. 3, pp. 308–311, Mar. 2010.

[11] Prabu, S., V. Balamurugan, and K. Vengatesan. "Design of cognitive image filters for suppression of noise level in medical images." *Measurement* 141 (2019): 296-301

[12] K. Vengatesan, A. Kumar, R. Naik and D.K. Verma, "Anomaly Based Novel Intrusion Detection System for Network Traffic Reduction," 2018 2nd *International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2018 2nd International Conference on, Palladam, India*, 2018, pp. 688-690.

[13] Kumar, V.D.A., Kumar, V.D.A., Malathi, S, Vengatesan. K, Ramakrishnan. M. "Facial Recognition System for Suspect Identification Using a Surveillance Camera". *Pattern Recognition and Image Analysis*, July 2018, Volume 28, Issue 3, pp 410–420.