INTERNATIONAL JOURNAL FOR INNOVATIVE RESEARCH IN MULTIDISCIPLINARY FIELD ISSN: 2455-0620 Volume - 7, Issue - 7, July – 2021
Monthly, Peer-Reviewed, Refereed, Indexed Journal with IC Value: 86.87 Impact Factor: 6.719
Received Date: 01/07/2021 Acceptance Date: 16/07/2021 Publication Date: 31/07/2021

A New strategy for Automatic Face Annotation

[1]Twisha Patel, [2]Kajol Patel
[1, 2]Assistant Professor
[1, 2]Computer Engineering Department, Parul Institute of Engineering and Technology – Diploma Studies,
Parul University, Vadodara, India.
Email – [1]twisha.patel12775@gmail.com [2]kajol.patel12633@paruluniversity.ac.in

Abstract: *Exact identification and naming of face has been significant in various regions of computer science. A tremendous number of photos are shared over the social passages, a portion of these photos are marked with appropriate names yet immense quantities of them are unlabelled. So, it ends up being hard to understand the name of the individual from the image if any sporadic individual sees it. We have proposed a new automatic face annotation approach that consequently distinguish and annotate the face with higher accuracy. The experimental results show that our proposed algorithm gives improved precision and recall value. Especially it reduces false positive image retrieval.*

Key Words: *automatic image annotation; automatic face annotation; feature extraction; Convolutional Neural Network; Gabor Wavelet Transform; Classification; Support Vector Machine; Similarity measurement.*

1. INTRODUCTION:

Automatic face annotation is method to distinguish human faces from a photograph and dispense their relating human names [3, 4, 5, 6, 7, 12, 22, 23]. Face annotation is closely related to the face detection and face recognition which are used in computer vision and image processing [3, 4, 5, 6, 7]. Face annotation is as an extended face recognition problem in which face classification models are trained from a collection of well labeled facial images using supervised or semi-supervised machine learning techniques [3, 4, 26]. A broad part of photographs shared over the web are human facial pictures. A portion of these facial pictures are named appropriately with names, yet most of them are unlabeled. Along these ways, this has given the motivation for assessment of automatic face annotation techniques to overcome this issue. There are some limitations of model-based strategies of face annotation. It is expensive and time- consuming to collect large amount of human facial images which are well labeled. To overcome these restrictions, search based face annotation is used which aim to tackle the auto face annotation task by exploiting content-based image retrieval technique in mining number of facial images on the web. The main purpose of search-based face annotation is to assign correct name labels to a given query image [7]. Face annotation is useful in many applications such as identity verification (passports, driving licenses, employee IDs), criminal justice systems (forensics), wild landmark face annotation, and online photo album management [3, 6].

We propose an automatic face annotation algorithm. This algorithm annotates the facial image with a label automatically. Our proposed algorithm will reduce false positive facial image retrieval. In light of our thorough study on existing feature extraction techniques we have chosen convolutional neural network with Gabor wavelet transform. For extracting the local feature, we have used Gabor wavelet transform and for global feature extraction we have used convolutional neural network**.** Similarly, based on our study on existing classifiers, we have chosen Support Vector Machine (SVM) for automatic face annotation. SVM classifier gives good performance in terms of accuracy and gives correct classification. So, we have used SVM classifier in our proposed system. In this paper, segment I portrays the introduction part of automatic face annotation. Segment II is about the proposed automatic face annotation algorithm. Segment III portrays the performance analysis. Area IV portray the conclusion.

2. AUTOMATIC FACE ANNOTATION ALGORITHM

Proposed automatic face annotation algorithm contains these steps: i) Feature extraction, ii) Classification, iii) Similarity measurement, iv) Retrieval, and v) Annotation.

Input: ‘q’ given as query image by user. Output: Similar facial images with annotation.

- Feature vector of query image = Fq
- Database images d= {d_1, d_2,…,d_n}
- Feature Vector of database Fd= {F_1, F_2,….,F_n}

Steps for the algorithm:

- For database image d do:
  - Apply Gabor Wavelet Transform and CNN to extract the features of database images respectively.

- Combine the above extracted features to obtain resultant feature vector Fd of database images d.
- Apply SVM train to the feature vector fd and train the database images into its belonging class.

- For query image q do:
  - Apply Gabor Wavelet Transform and CNN to extract the features of query image respectively.
  - Combine the above extracted features to obtain resultant feature vector Fq of query image q.
  - Apply SVM classifier on feature vector Fq to obtain class of query image.
- Apply similarity measure between Fd and Fq to discover the correspondence between the test picture and database picture class.
- Retrieval of topmost N similar images with smallest distance.
- Annotate the facial images.

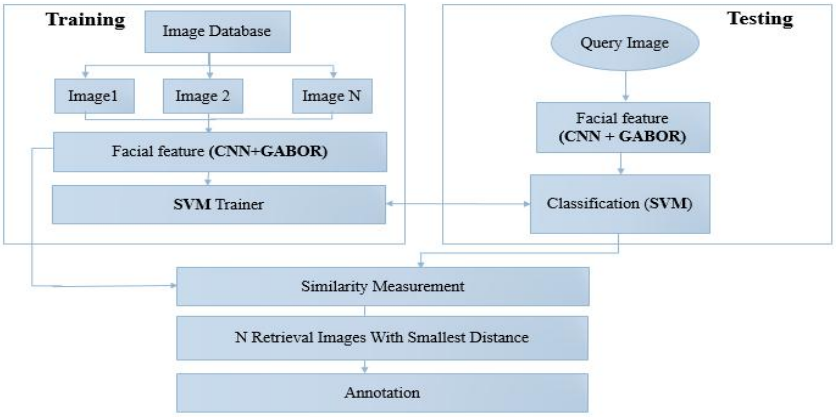

Figure 1. Proposed Flow of Automatic Face Annotation

2.1. Facial Feature Extraction:

Feature extraction is the technique of extracts various features like eyes, nose, mouth, etc. With feature extraction corresponding feature vector is generated for each feature. Accuracy of the feature dependent on the choice of feature extraction method. Hence, to pick a fitting feature extraction technique, we have examined a several feature extraction techniques which are SURF, SIFT, Convolutional neural network (CNN), Gabor Wavelet Transform, Eigenfaces, LBP and Haar. Local feature which are usually geometric and focuses on pixel intensities and Local features gives texture information. In local feature extraction, local features like nose, eyes, lips, etc. are segmented. Global features are used for whole analysis of facial features. In global feature extraction, whole face taken as the input. Convolutional Neural Network (CNN) extract features globally and Gabor Wavelet Transform (GWT) extract features locally. From the comparative study of feature extraction techniques, Convolutional Neural Network (CNN) and Gabor Wavelet Transform (GWT) are more beneficial than other techniques. CNN reduce the memory footprint and improves the performance. It also handles facial images with various poses, facial expressions and illumination. GWT concentrates on important components of faces such as mouth, eyes, nose, etc. It considers all key points of the face. It also invariance to illumination, rotation, scale, transform and have optimal localization property in frequency and spatial domain. Hence, we are going to combine CNN and GWT in terms of good accuracy and result.

2.2. Classification using SVM Classifier:

After feature vector is generated, next step is classification take place. We are using supervised learning approach therefore we need to train the images. Classification consists of training and testing of images. We have surveyed various classifiers which are Support Vector Machine (SVM), Artificial Neural Network (ANN) and K Nearest Neighbor (KNN). Based on our study on existing classifiers SVM is more beneficial than another classifier. SVM is basically deals with pattern classification problem. It is used to classify linear and non-linear classification patterns. Number of misclassifications is less in SVM. It attains optimal class boundaries by finding the maximum distance between classes. Moreover, SVM is robust, faster and required less memory. SVM is less complex than ANN. ANN is more time consuming as compare to SVM. Hence, we are going to use SVM classifier for image classification. Therefore, SVM is frequently used classifier in image processing [16] [18].

2.3. Similarity Measurement:

Apply similarity measure between feature vector of database images (Fd) and feature vector of query images (Fq) to find the similarity between the test image and database image class. We have studied different distance

which are Euclidean Distance, city block distance (Manhattan Distance) and minkowski Distance [19]. Based on our study on different distances, city block distance (Manhattan distance) attains higher precision value compared to the Euclidian distance and minkowski distance. So, we are going to use city block distance for similarity measurement in our proposed system.

2.4. Nearest-Neighbor Retrieval:

After similarity measurement, next step is nearest neighbor retrieval. Top N similar images are retrieved with smallest distance.

2.5. Face Annotation:

After the nearest-neighbor picture retrieved, next task is face annotation in which facial images are annotated with a label associated with retrieved nearest-neighbors. The label of these nearest neighbor images is shared with the query image and finally face annotation task is completed.

3. PERFORMANCE ANALYSIS:

We have conducted experiments in Matlab-2013. We have used the Indian Movie Face database (IMFDB). Indian Movie Face database (IMFDB) [24] is a large unconstrained face database consisting of 34512 images of 100 Indian actors collected from more than 100 videos. All the images are manually selected and cropped from the video frames resulting in a high degree of variability in terms of scale, pose, expression, illumination, age, resolution, occlusion, and makeup. Among 34512 images, we decided to use 500 images of Indian actors. We divided images into 5 classes. Each class contain 100 images. We take 50% images for training and 50% images for testing. Precision and recall are the most widely used measurement method to evaluate the retrieval accuracy. Recall measures the ability of the system to retrieve all images that are relevant, while precision measures the ability of the system to retrieve only images that are relevant. The performance of the classifier is measured in terms of accuracy. This term refers to the ability of the method to correctly predict the class of new unseen data. So, classification accuracy is calculated by determining the percentage of cases in which the test sets are correctly classified in their respective classes. A good classifier always has a high accuracy. The precision and recall and accuracy of the classifier can be calculated as follows [21].

$$\text{Precision} = \frac{TP}{TP+FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{Accuracy} = \frac{\text{True positive + true negative}}{\text{Toal number of images}} \quad (3)$$

TP= No. of retrieved images which are like to the query.
FP= No. of retrieved images unlike to the query.
FN= No. of images in the database which are alike to query but not retrieved.
Where true positive is the quantity of right predictions when an instance is certain, true negative is the quantity of right predictions when an instance is negative.

4. EXPERIMENTAL ANALYSIS AND RESULTS:

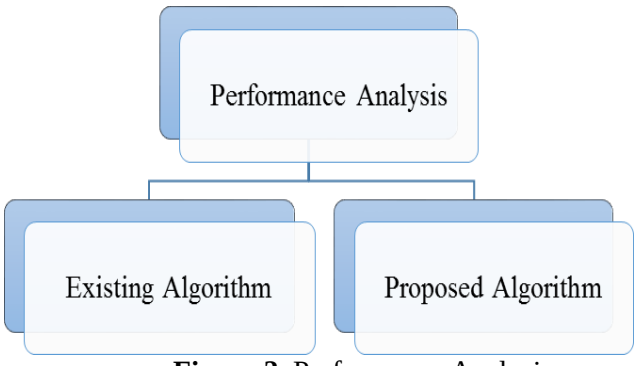

Figure 2. Performance Analysis

- **Existing Algorithm:**

We have compared our proposed algorithm with [7]. For face annotation approach, in [7] feature extraction is

finished using SURF detection technique and labels are given to the picture using these closest neighbors.

Figure 3. Feature Extraction using SURF

Figure 4. Retrieval and Annotation results of existing algorithm

Figure 5. Retrieval and Annotation results of existing algorithm

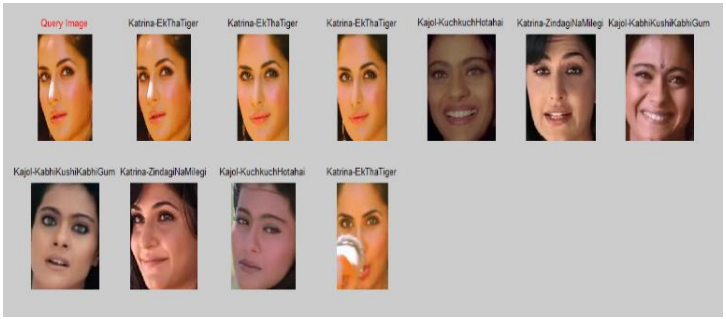

Figure 6. Retrieval and Annotation results of existing algorithm

Performance Analysis of Existing Algorithm:

The outcomes are determined dependent on top 10 retrieved images from Indian Movie Face database (IMFDB) for query image. The Euclidean distance is utilized for similarity measure between query image and database images. The estimations of precision and recall parameters for every 5 class are determined and appeared in the table.

TABLE I. RESULT OF EXSTING ALGORITHM

Classes	**Precision**	**Recall**
Aamir Khan	0.32	0.0747
Amitabh Bachchan	0.38	0.076
Kajol	0.38	0.076
Kareena Kapoor	0.22	0.0712
Katrina kaif	0.42	0.168
Average	**0.344**	**0.09318**

- **Proposed Algorithm**

The major steps of proposed approach are:

- Facial feature extraction
- Classification
- Similarity measurement
- Nearest-Neighbor retrieval and Face annotation

Facial Feature Extraction using Gabor Wavelet Transform (GWT)

For local feature extraction we have used Gabor wavelet transform. In our proposed work, we have use 5 different scales and 6 orientations so we obtained total 60 features per image.

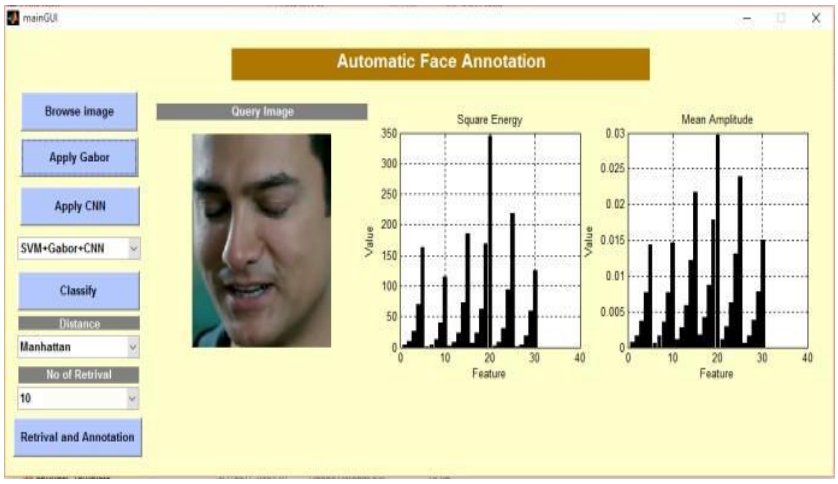

Figure 7. Feature Extraction using GWT

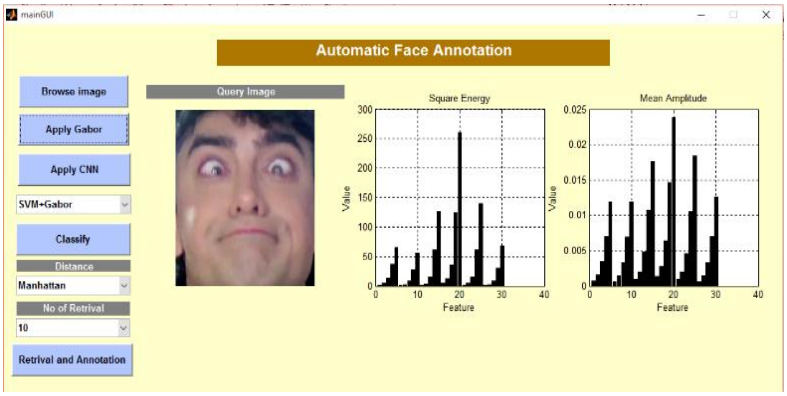

Figure 8. Feature Extraction using GWT

Facial Feature Extraction using Convolutional Neural Network (CNN)

For global feature extraction we have used convolutional neural network. We obtained total 2160 features per image.

.

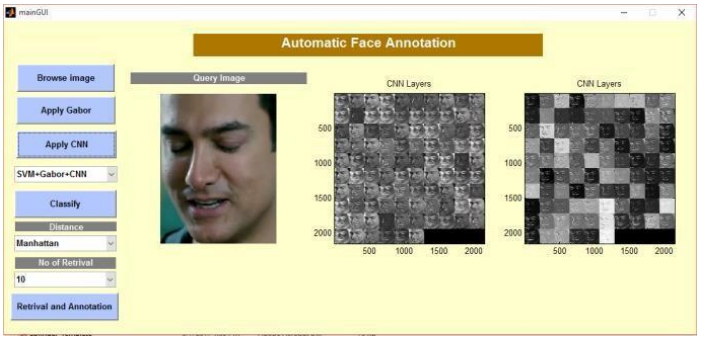

Figure 9. Feature Extraction using CNN

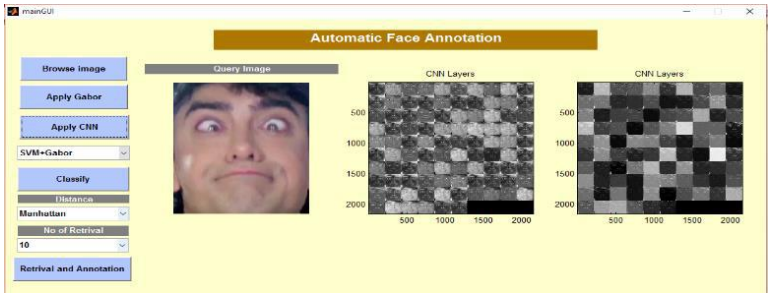

Figure 10. Feature Extraction using CNN

Classification using SVM Classifier

SVM classifier gives great execution as far as accuracy and gives correct classification. Thus, we have utilized SVM classifier in our proposed framework. The exhibition of the designed classifier is estimated as far as accuracy using confusion matrix.

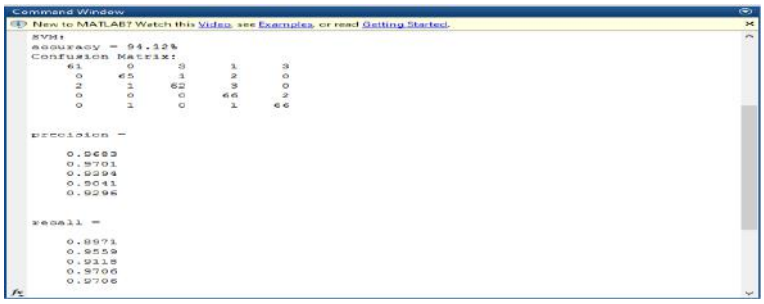

Figure 11. Accuracy, Precision, Recall and Confusion matrix of SVM Classifier

From Fig. 7 we can conclude that Accuracy of classifier is
94.12 %. We have tested every class query images and most of the time classifier gives correct result which proves the efficiency of our proposed approach. Class 1 defines the class of Aamir khan. Figure 8 shows the classification result.

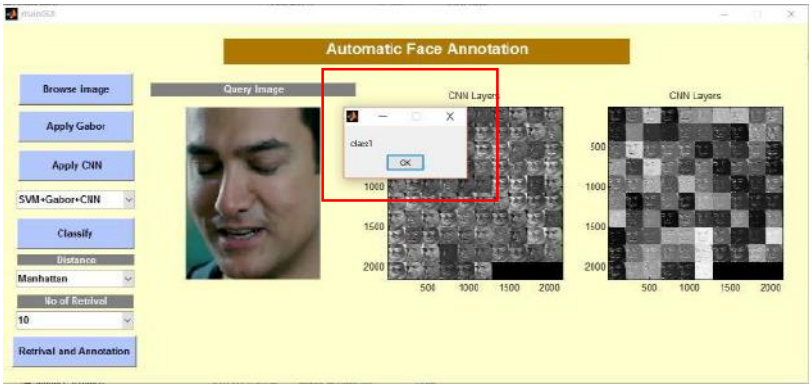

Figure 12. Classification result

We have analyzed different distance which are Euclidean Distance, Manhattan Distance and minkowski Distance. The results are calculated based on top 10 retrieved images from Indian Movie Face database (IMFDB) for query image. The values of precision and recall parameters for each 5 class are calculated for Euclidean distance, city block distance (Manhattan distance), and minkowski distance. From which city block distance (Manhattan distance) achieves higher precision value compared to the Euclidian distance and minkowski distance. So, we are going to use city block distance for similarity measurement in our proposed system

Figure 13. Retrieval and Annotation results of proposed

Figure 14: Retrieval and Annotation results of proposed algorithm

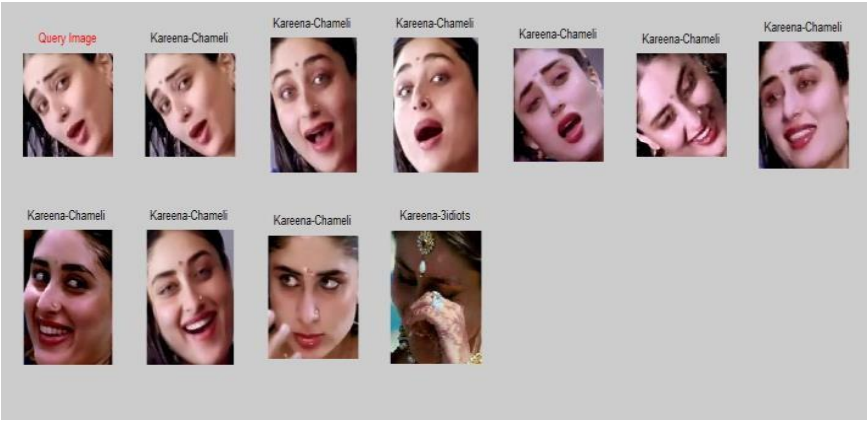

Figure 15: Retrieval and Annotation results of proposed algorithm

Performance Analysis of Existing Algorithm:

The outcomes are determined dependent on top 10 retrieved images from Indian Movie Face database (IMFDB) for query image. The Euclidean distance is utilized for similarity measure between query image and database images. The estimations of precision and recall parameters for every 5 class are determined and appeared in the table.

Performance Analysis of Proposed Algorithm:

The results are calculated based on top 10 retrieved images from specific class which is obtained by SVM for query image. The values of precision and recall parameters for each 5 class are calculated for city block distance (Manhattan distance) shown in the table 2.

TABLE II. RESULT OF PROPOSED ALGORITHM

Classes	**Precision**	**Recall**
Aamir Khan	0.84	0.202
Amitabh Bachchan	1	0.2
Kajol	0.9	0.18
Kareena Kapoor	0.88	0.29
Katrina kaif	0.84	0.336
Average	**0.892**	**0.2416**

Table 2 sows the result of Gabor wavelet transform and convolutional neural network for retrieving the relevant images from specific class which is obtain by SVM classifier. The value of precision and recall parameter for each 5 class are calculated by using city block distance.

- **Comparison of Existing and Proposed Work:**
  We have compared existing and proposed work in terms of precision and recall.

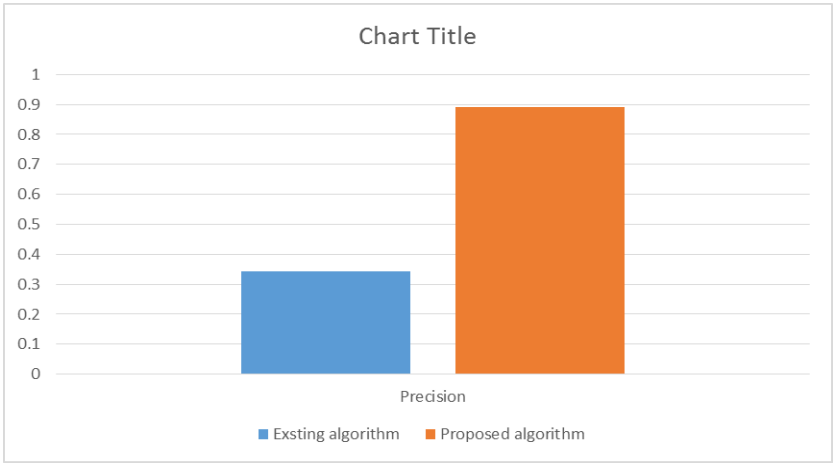

Figure 16: Comparison of precision for existing and proposed work

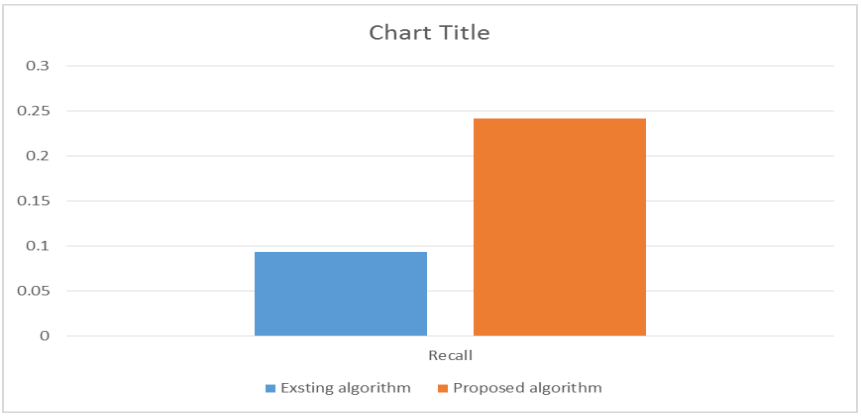

Figure 17: Comparison of recall for existing and proposed work

Result shows that the proposed algorithm gives the precision value 0.892 and recall value 0.2416. Existing algorithm gives the precision value 0.344 and recall value 0.09318. From the results, our proposed algorithm archives higher precision and recall value and reduced the false positive retrieval of facial images by using SVM classifier because query image is only compared with particular class get by the SVM and similarity measurement is only performed amid the query image and the image in the particular class get by the classification.

5. SUMMERY:

In this work, we have proposed an automatic face annotation approach that automatically label the facial images with their human names. To evaluate the performance of proposed approach, we have identified MATLAB-2013 as an efficient implementation tool and Indian Movie Face database (IMFDB. In our proposed system, SVM Classifier gives good accuracy using feature extraction techniques which are CNN and GWT. We have evaluated the performance of SVM classifier using confusion matrix. We have evaluated the retrieval results using precision and recall parameter. By comparing our proposed algorithm with existing algorithm, it gives the good result and reduced the false positive facial images.

REFERENCES:

Proceedings Papers:

1. D. Zhang, Md. M. Islam and G. Lu, "A review on automatic image annotation techniques", Pattern Recognition, vol. 45, no. 1, pp. 346–362, January 2012.
2. Kalaivani and S. Chitrakal, "Challenges & approaches in multi-label image annotation," Fourth International Conference on Computing, Communications and Networking Technologies, Tiruchengode, 2013, pp. 1-8
3. D. Wang, S. C. H. Hoi, Y. He and J. Zhu, "Mining Weakly Labeled Web Facial Images for Search-Based Face Annotation," IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 1, pp. 166-179, January 2014.
4. D. Wang, S. C. H. Hoi, Y. He, J. Zhu, T. Mei and J. Luo, "Retrieval-ased Face Annotation by Weak Label Regularized Local Coordinate Coding," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 36, no. 3, pp. 550-563, March 2014.
5. X. Zhang, L. Zhang, X. J. Wang and H. Y. Shum, "Finding Celebrities in Billions of Web Images," IEEE Transactions on Multimedia, vol. 14, no. 4, pp. 995-1007, August 2012.
6. C. Nandhagopal and K. Priyanka, "A comparative analysis of face annotation schemes for web facial images in social networks," IEEE International Conference on Electrical, Computer and Communication Technologies, Coimbatore, 2015, pp. 1-5.
7. E. Paul and A S Ajeena Beegom, "Mining images for image annotation using SURF detection technique," IEEE International Conference on Control Communication & Computing India, Trivandrum, 2015, pp. 724-728.
8. C. Yan, C. Lang, T. Wang, X. Du and C. Zhang, 2014, "Age estimation based on convolutional neural network" Pacific Rim Conference on Multimedia, springer, Malaysia, 2014, pp. 211- 220.
9. J. K. Kamarainen, "Gabor features in image analysis," IEEE International Conference on Image Processing Theory, Tools and Applications, Istanbul, 2012, pp. 13-14.
10. S. Murala, A. B. Gonde and R. P. Maheshwari, "Color and Texture Features for Image Indexing and Retrieval," IEEE International Advance Computing Conference, Patiala, 2009, pp. 1411-1416.
11. V. Purandare and K. T. Talele, "Efficient heterogeneous face recognition using Scale Invariant Feature Transform," IEEE International Conference on Circuits, Systems, Communication and Information Technology Applications, Mumbai, 2014, pp. 305-310.
12. N. Jamil and S. A. Sa'dan, "Automated face annotation for personal photo management," IEEE International Conference on Computational Science and Technology, Kota Kinabalu, 2014, pp. 1-5.
13. V. P. Kshirsagar, M. R. Baviskar and M. E. Gaikwad, "Face recognition using Eigenfaces," 3rd International Conference on Computer Research and Development, Shanghai, 2011, pp. 302- 306.
14. K. Meena and A. Suruliandi, "Local binary patterns and its variants for face recognition," IEEE International Conference on Recent Trends in Information Technology, Tamil Nadu, 2011, pp. 782-786.
15. S. M. Lajevardi and Z. M. Hussain, "Automatic facial expression recognition: Feature extraction and selection," Signal, Image Video Process., vol. 6, no. 1, pp. 159–169, March 2012.
16. C. Jin, J. Liu and J. Guo "A Hybrid Model Based on Mutual Information and Support Vector Machine for Automatic Image Annotation," Computer Science On-line Conference, Springer, 2015, pp.29-38.
17. T. R. Kalansuriya and A. T. Dharmaratne, "Facial image classification based on age and gender," IEEE International Conference on Advances in ICT for Emerging Regions (ICTer), Colombo, 2013, pp. 44-50.
18. M. Oujaoura, B. Minaoui and M. Fakir, "A semantic approach for automatic image annotation," IEEE International Conference on Intelligent Systems: Theories and Applications (SITA), Rabat, 2013, pp. 1-8.
19. F. Malik and B. Baharudin, "Analysis of distance metrics in content-based image retrieval using statistical quantized histogram texture features in the DCT domain", Journal of King Saud University - Computer and Information Sciences, vol. 25, no. 2, pp. 207-218, July 2013.
20. ElAlami, M. Esmel. "A new matching strategy for content-based image retrieval system." Applied Soft

Computing, vol. 14, 407- 418, October 2013

21. S. Setty, M Husain, P. Beham, J. Gudavalli, M. Kandasamy, R. Vaddi, V. Hemadri, J.C. Karure, R. Raju, B. Rajan, V. Kumar and C. V. Jawahar, "Indian Movie Face Database: A benchmark for face recognition under wide variations," IEEE National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), Jodhpur, 2013, pp. 1-5.
22. Kasthuri and A. Suruliandi, "A survey on face annotation techniques," IEEE 4th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, 2017, pp. 1-9.
23. B.N. Tandel, U. Desai, “Various Face Annotation Techniques: Survey” Intelligent Communication Technologies and Virtual Mobile Networks. ICICV 2019. Lecture Notes on Data Engineering and Communications Technologies, vol 33. Springer, Cham, 2019, pp. 94-102.
24. F. Kong, “Facial expression recognition method based on deep convolutional neural network combined with improved LBP features”, Pers Ubiquit Comput 23, Springer, 2019, pp. 531–539.
25. S. Khan, M. H. Javed, E. Ahmed, S. A. A. Shah and S. U. Ali, "Facial Recognition using Convolutional Neural Networks and Implementation on Smart Glasses," IEEE International Conference on Information Science and Communication Technology (ICISCT), Karachi, Pakistan, 2019, pp. 1-6.
26. V. A. Sindagi and V. Patel, "DAFE-FD: Density Aware Feature Enrichment for Face Detection," 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa Village, HI, USA, 2019, pp. 2185-2195.

Book:

1. MATLAB and Statistics Toolbox Release 2013a, The MathWorks, Inc., Natick, Massachusetts, United States.