Mobile Based Visual Search

Pallavi Nagap
B.E Computer Department, Mumbai University1
pallavi.nagapp121@gmail.com

Abstract**—**Visual search has been an interesting topic for researchers since more than a decade. Content based image retrieval has been researched broadly. But there are many challenges which are faced for mobile based image retrieval. These problems have been addressed in this project. Nowadays mobile devices and cell phones are used at a large scale since they are very user friendly.

The HAAR classifier is used to carry out the face detection. The query image is captured using the mobile camera is used as the input to the system. Feature database is created by extracting LBP features from the images in the training set. The feature extraction of the query image is carried out using LBP. The features of the query image are then compared with the features of the training images. The system efficiently detects the face within the query image irrespective of the orientation of the face. The test results have shown the accuracy of the system. Thus keeping in mind the benefits of the algorithm the system uses HAAR Classifier for face detection, LBP (Linear Binary Pattern) for feature extraction and SVM classifier for face recognition.

Keywords ***— local binary pattern, image search , visual search***

*****

I. INTRODUCTION

Content based image retrieval has been broadly researched and used. But with the increase in use of smart phones there is need of mobile based applications. With the rapid development of mobile computing, biometric identification technologies including face recognition on smart phones become increasingly attractive. Face recognition and detection can be carried on mobile devices having camera. This feature can further be used for various reasons such as authorization, security etc. Mobile phones with front camera makes this more interesting. There are two main challenges of face recognition for mobile computing: one is the problem of the variations of facial images acquired in unconstrained environment, the other is the computing limitation of mobile phone platform. Keeping these problems in mind mobile based visual search is developed.

II. ISSUES OF MOBILE BASED VISUAL SEARCH

There are many issues which have to be faced while visual search on mobile device. The display screen is small as compared to the desktop. The transmission cost is also more. The resolution is low. The storage memory as well as computation power is also low. All these factors have to be considered while developing a mobile based image retrieval application. But in spite of these difficulties mobile based visual search is more beneficial as compared to desktop based visual search.

III. ISSUES WITH FACE RECOGNITION

1. Human face is not unique or rigid object; hence the variations in face appearance can make face recognition difficult.
2. Varying face appearance of same person like with or without glasses, facial hair, cosmetics etc.
3. Some other problems like illumination, pose, scale, and imaging parameters likeresolution, focus, noise etc.
4. Ageing is also a problem while recognizing a face.

IV. RELATED WORK

QawlanAkariman ,AgungNugrohoJati and AstriNovianty have proposed an effective technique for android based face recognition using Local Binary Pattern (LBP). Keeping in mind the simple and fast computation of LBP technique, LBP is used for face recognition. LBP will take the facial feature by dividing it into several small areas. [1] They have carried out comparison of various techniques and considered their advantages and disadvantages. In this paper various face recognition methods such as PCA (Peripheral Component Analysis), LDA (Linear Discriminant Analysis), LVQ(Learning Vector Quantization) have been compared The system has been tested using two types of self-made database and Yale face database.

Kamal Nasrollahi and Thomas B. Moeslund has proposed a system which deals with many of the challenges in a robust and real-time approach. This paper has presented a novel approach to face recognition using Haar-like features extracted from integral face images. The extracted features are then used for learning the complicated space of human faces in the Haar-like feature domain using a probabilistic classifier. The proposed system has been tested using public databases and the obtained results show that it indeed outperforms the state-of the-art face recognition systems. Several experimental results have been considered for testing the system against different imaging degradation including noise corruption, occlusion, headpose and down-sampling.

Integral images introduced by Viola and Jones for rapid Haar-like features extraction for the purpose of object detection by AdaBoost classifier in [2] more than a decade ago. AdaBoost with Haar-like features is very efficient for object detection, but it does not result in good performance for object recognition. Keeping this in mind HAAR-like features have been used along with PNN classifier for face recognition. Cross validation is carried out using different databases such as Faces94 Database, ORL Database and UMIST Database.

LBP (Local Binary Pattern) is simple to implement as compared to PCA and LDA. Pre-processing such as cropping, rotating and histogram equalization enhances the face recognition. [4]LBP performs better in feature extraction than in face recognition, need of multi classclassifier[4]. Drawbacks of CBIR(Content Based Image Retrieval) , need of Android based face recognition application[1]. HAAR features carries out robust real time face detection [7].

V. FLOWCHART OF THE SYSTEM

The query image is given as an input to the system. The face detection is carried out at the second stage and then these features are used for face recognition.

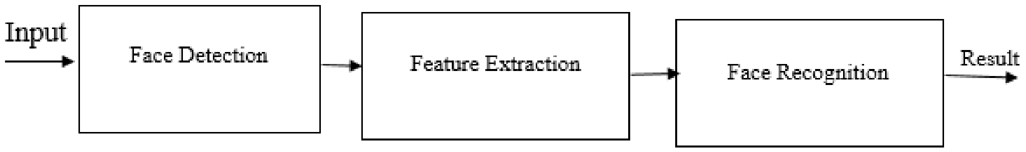


Figure 1: Block diagram of proposed system

VI. Flow chart

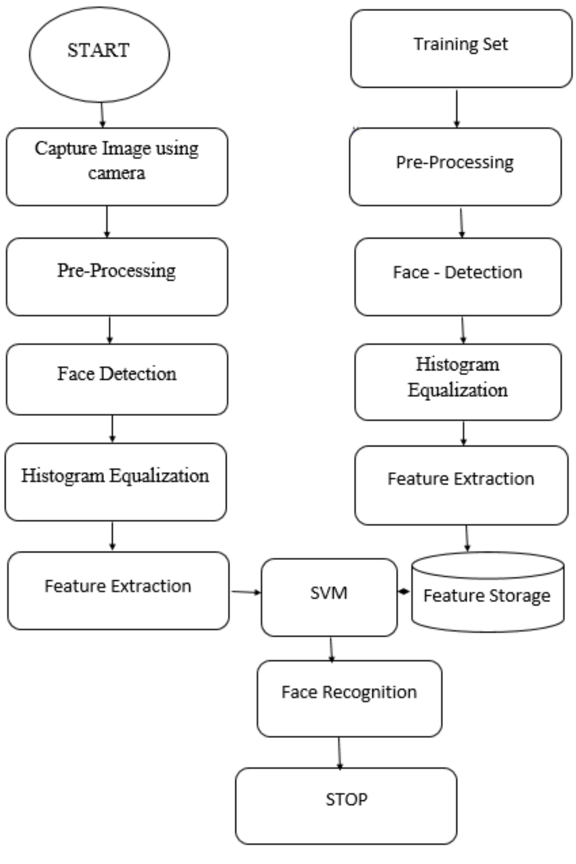


Figure 2: Flowchart of proposed system

Proposed approach includes following steps:

1) Part I : Dataset training

a) Collect images which are to be recognized

b) Collect at least 10 images of each individual in different facial expression

c) The more the number of images the more will be the accuracy of the system

d) Data training is the most important step in data recognition

2) Part II : Pre-processing

a) Rotate the image from right to left

b) Crop the image into standard size of 92x112

3) Part III : Face detection

a) Using Haar like feature the face detection is carried out.

b) Query image which is clicked with the camera is used to detect the face.

4) Part IV : Histogram equalization

a) Histogram equalization is carried .

5) Part V: Feature extraction

a) Features are extracted from the face detected image

b) Extracted features are stored in compared with the once stored in the database

c) The matching results are displayed.

d) Features are extracted using LBP

e) Histogram equalization is carried out

6) Part VI : Feature Storage

a) Extracted features are then stored in the database in csv file.

b) Trained database is then used for face recognition

7) Part VII : Face Recognition

a) Features extracted of the query image are matched with the features in the database

b) On successful matching the results are displayed.

The Fig: 2 shows the flowchart for mobile based visual search. The first stage is collecting the images of the facial images of which the features have to be extracted. This stage is called the training set. Labelling of the images are done. Second stage is pre-processing of the facialimages. Histogram equalisation is carried out to eliminate the illumination variation. The image is converted to grey scale image. Haar features are extracted from the images to detect the face. This is the third stage. In the fourth stage the LBP features are extracted. Histograms of these features are then used to find out the similar images. The histograms are then passed to the SVM classifier. This is done at the server's end. At the front end that is in the mobile application after the login process is done. The query image is captured with the camera. The query image is then sent to the server. The next process at the server side is face detection. After the face has been detected LBP features are extracted.

VII. Implementation details

This existing system with detail methodology is described in this chapter, along with this proposed methodology its working modules are also explained in details.

a. Existing system

In the existing system LBP has been used for feature extraction and face recognition. While
HAAR cascade has been used for face detection. The existing systems gives satisfactory results, but still the results need to be enhanced. But it has been observed that instead of using
LBP for feature extraction and face recognition both, using a classifier will be more beneficial wherein one-to-one as well as one-to-many comparison can be carried out. There are various limitations of mobile device as well.

b. Modules and Description

Different modules of the proposed methodology are being described as follows:

Training Set

AT&T dataset has been used to test the accuracy of the system. Facial images with different expressions and face orientations are been used. The AT&T facial database, sometimes also referred to as ORL (Olivetti Research Laboratory) Database of Faces, contains ten different images of each of 40 distinct subjects from which I have used 13 subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement). The images are in grey scale, with dimension of 92x112 pixels. The AT&T database was sorted into different folders and given unique name i.e. class labels. Along with AT&T database, self-made database of facial images has also been made. Facial images of 6 subjects were taken against plane background with different facial expressions. The system is tested with this database as well.

ISSN: 2321-8169
34 – 45

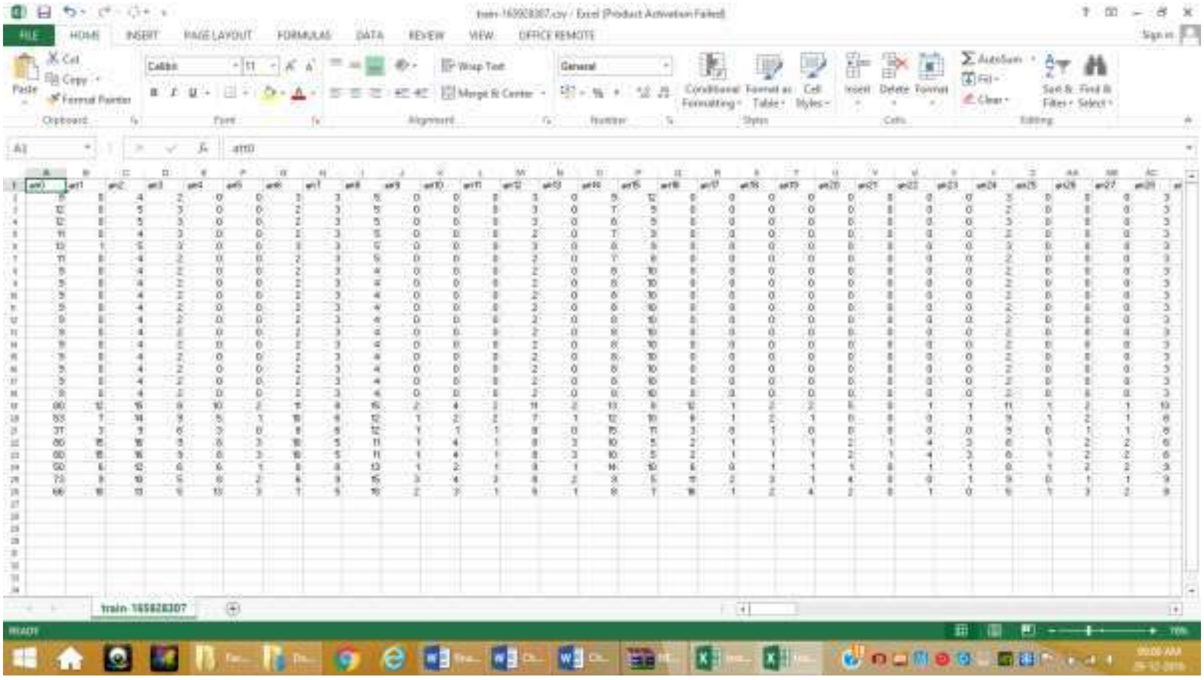


Figure 3 : The features stored for the training set.

While preparing the training set the features are extracted and stored in a CSV file. The above figure shows the attributes of the training images. There are 256 attributes of each image.

Pre-Processing

Since the image is captured real time there are few pre-processing steps done. The captured image is rotated by the system. So once it is uploaded on to the server the image which is initially rotated right is then rotated to the left. The image is cropped to a standard size of 92x112.

Face detection using HAAR-Like Features

The input query image undergoes pre – processing. The input image when captured gets rotated automatically. So the image is the rotated back and the stored. The captured image is converted into a diminished size and into grey scale image. The Haar like features are then extracted to detect the face. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window which sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. For example, if we have an image database with human faces. It is observed that among all faces the region of the eyes is darker than the region of the cheeks.

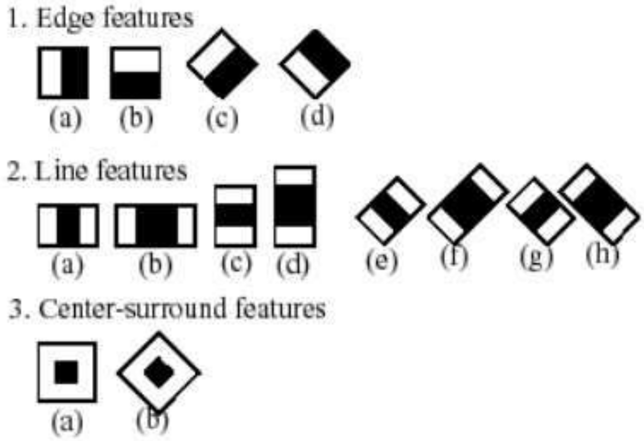


Figure 4 :Haar features[6]

Hence a common Haar feature is detected for face detection which is a set of two adjacent rectangles that lie above the eye and the cheek region. The position of these rectangles is defined relative to a detection window that acts like a bounding box to the target that is the face. The key advantage of a Haar-like feature over most other features is its calculation speed.

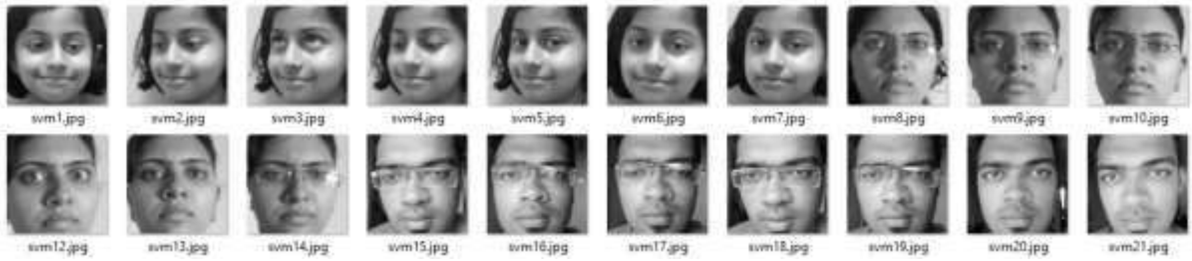
Histogram equalization:

Histogram equalization is a method in image processing for contrast adjustment using the image's histogram. This method usually increases the global contrast of many images, especially when the usable data of the image is represented by close contrast values. Through this adjustment, the intensities can be better distributed on the

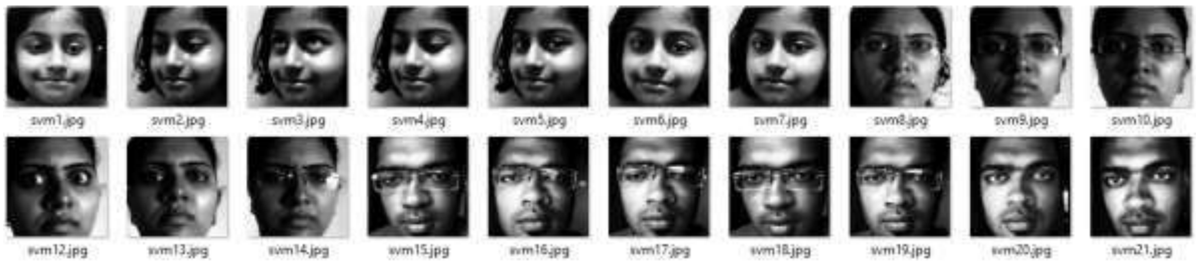
ISSN: 2321-8169
34 – 45

histogram. This allows for areas of lower local contrast to gain a higher contrast. Histogram equalization accomplishes this by effectively spreading out the most frequent intensity values.

The method is useful in images with backgrounds and foregrounds that are both bright or both dark. A key advantage of the method is that it is a fairly straightforward technique and an invertible operator. So in theory, if the histogram equalization function is known, then the original histogram can be recovered. The calculation is not computationally intensive. In order to equalize the feature histogram created from the features extracted by LBP. Since the training database is not uniform in nature histogram equalization is carried out.



a



b

Figure 5: a) Self-made images before applying histogram equalization b) Self-made images after applying histogram equalization.

a

b

Figure 6: a) AT&T images before applying histogram equalization b) AT&T images after applying histogram equalization.

Feature extraction using LBP

In order to carry out face recognition the features have to be extracted from the query image.

The first step in constructing the LBP texture descriptor is to convert the image to grayscale.

For each pixel in the grayscale image, we select a neighbourhood of size 9 surrounding the centre pixel.

The proposed implementation of LBP is different from that of the original LBP function. The image is first divided into 3X3 matrix. The bin is initialised to zero. Then in the clockwise manner the first pixel is compared to the centre pixel and so on till the 9th pixel. If the first pixel is greater than or equal to the centre pixel then the bin size is incremented by 1. Thus for all the other pixels if they are greater than or equal to the centre pixel they are incremented by the power of 2 i.e 1 for 0th pixel, 2 for 1st pixel, 4 for 2nd pixel, 8 for the 5th pixel, 16 for 8th pixel, 32 for the 7th pixel , 64 for the 6th pixel and 128 for the 3rd pixel. The bin value is incremented by these values only if the pixel is greater than or equal to the centre pixel. These bin values are then converted to a histogram. The max value of the histogram is found out. Then the following formula is applied.

histogram[i] / max * 127.0

The value obtained after applying this formula is then rounded of. These values are then stored in a CSV file.

Advantages of LBP

1. LBP is tolerant against monotonic illumination changes.
2. LBP has simple computation and easy to understand and implement.

Face recognition using SVM

The SVM is first trained with the features extracted from the known images which are to be stored in the database. The features extracted by LBP which are then given to LBP Support vector machines (SVMs) are formulated to solve a classical two class pattern recognition problem.

To train the SVM the problem of difference in space is formulated, which explicitly captures the dissimilarities between two facial images. In difference space, focus is on the following two classes: the dissimilarities between images of the same individual, and dissimilarities between images of different people. These two classes are the input to a SVM algorithm. For face recognition, the decision surface is re-interpret to produce a similarity metric between two facial images.

The proposed system uses linear SVM. We are given a training dataset of n points of the form. as given in Eq (6.1)

$$(x_{1,}y_1) \dots \dots \dots \dots (x_{n,}y_n) \dots\dots\dots\dots\dots\dots\dots\dots\dots \text{ Eq (6.1)}$$

Where the y_i are either 1 or -1, each indicating the class to which the point x_i belongs. Each is a p-dimensional real vector. We want to find the "maximum-margin hyperplane" that divides the group of points x_i for which $y_i = 1$ from the group of points for which $y_i = -1$, which is defined so that the distance between the hyperplane and the nearest point x_i from either group is maximized.

Any hyperplane can be written as the set of points x satisfying the given in Eq (6.2)

$$w^*x - b = 0 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \text{ Eq (6.2)}$$

Where w is the (not necessarily normalized) normal vector to the hyperplane. The parameter $\frac{b}{||w||}$ determines the offset of the hyperplane from the origin along the normal vector w.

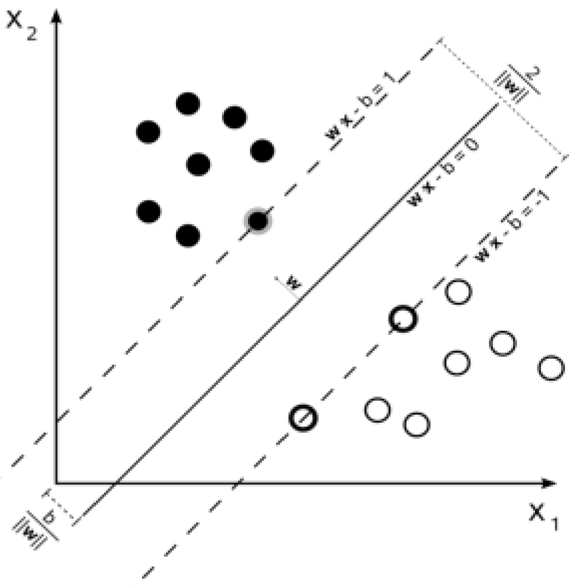


Figure 7: Support Vector Machine Classification.

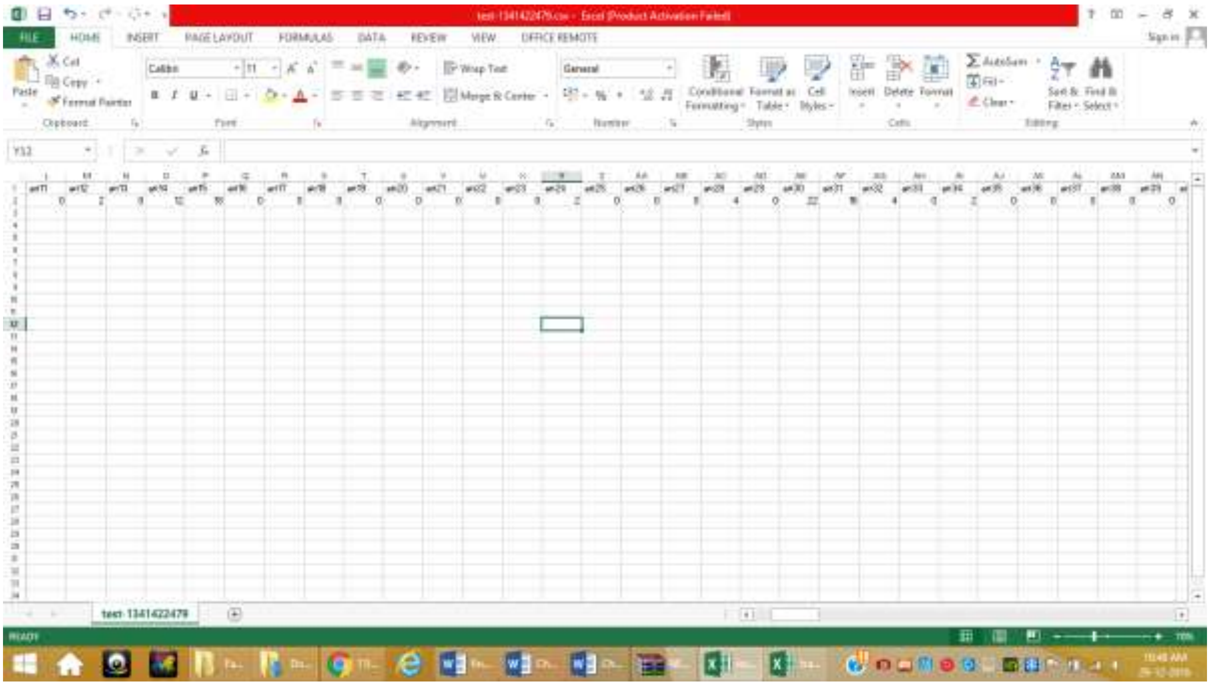


Figure 8 : Feature extracted from query image.

Advantages of SVM classifier:-

1. By introducing the kernel, SVMs gain flexibility in the choice of the form of the threshold separating solvent from insolvent companies, which needs not be linear and even needs not have the same functional form for all data, since its function is non-parametric and operates locally. As a consequence they can work with financial ratios, which show a non-monotone relation to the score and to the probability of default, or which are non-linearly dependent, and this without needing any specific work on each non-monotone variable.

2. Since the kernel implicitly contains a non-linear transformation, no assumptions about the functional form of the transformation, which makes data linearly separable, is necessary. The transformation occurs implicitly on a robust theoretical basis and human expertise judgement beforehand is not needed.

Figure 9 : Processing the captured image

After the image is captured using the front camera, you have to press the save button. The captured image is sent to the server where it is taken as the query image. This query image is then further processed.

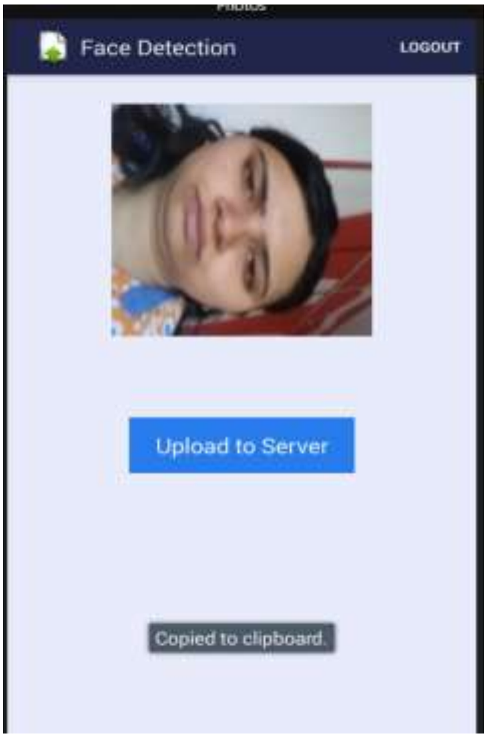


Figure10 :Upload the query image to server

After the query image is captured it has to be uploaded on to the server.

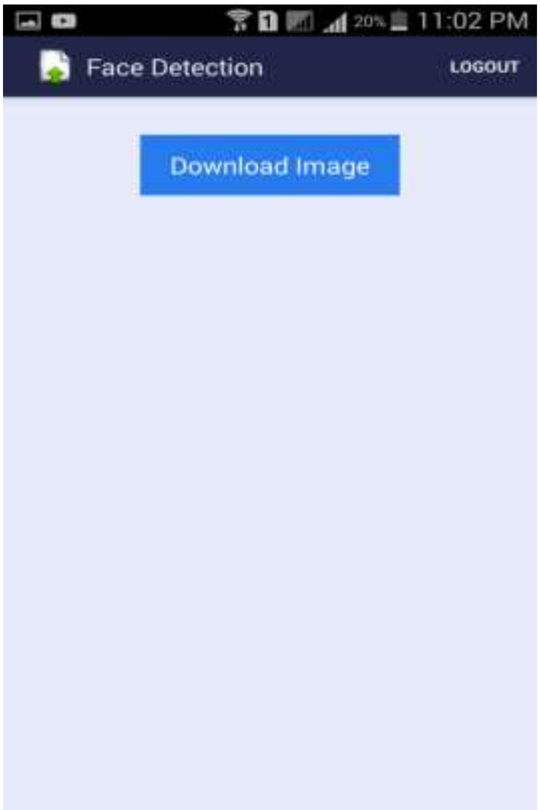


Figure11 :Download detected facial image.

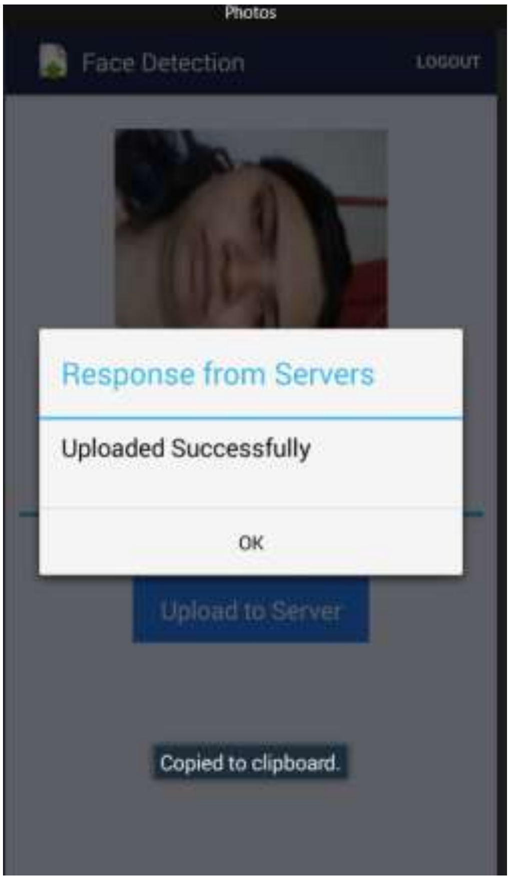


Figure12 : Image uploaded successfully to the server.

Figure 13:Recognised face with name label

ISSN: 2321-8169
34 – 45

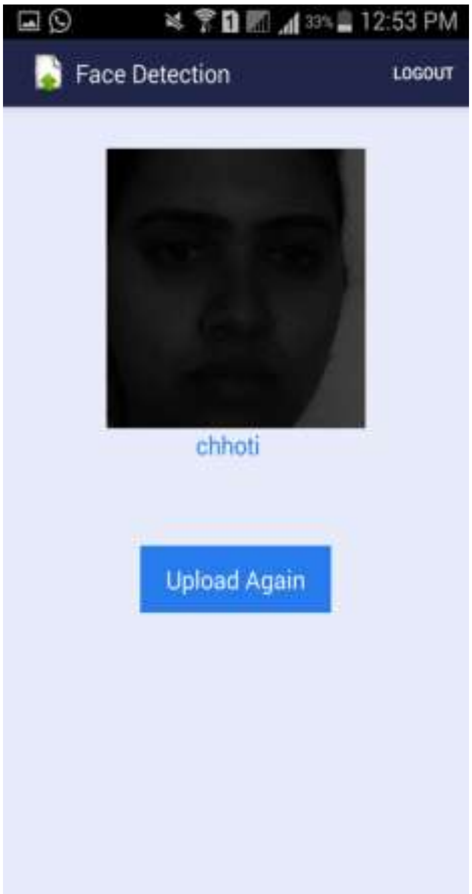


Figure14 : Output in case of low brightness. This gives wrong prediction.

Figure 15 : Missing trained images or processing error then null output

Figure 16 : Query images with different facial expressions and brightness levels.

VIII. Result and Analysis

The MOBILE BASED VISUAL SEARCH is used for face recognition on mobile device. The system is giving the output as expected. The face detection and recognition is being carried out efficiently. The system was tested using Weka software. The csv file of the training set of self-made database was given as input. 10 fold cross validation was carried out to check the accuracy. The validation shows 88% accuracy.The system was tested using Weka software. The csv file of the training set of AT&T database was given as input. 10 fold cross validation was carried out to check the accuracy. The validation showed 90.91 % accuracy.The system was tested using Weka software. The csv file of the training set of AT&T database was given as input. 66% split validation was carried out to check the accuracy. The validation showed 83.33% accuracy. The system was tested using Weka software. The csv file of the training set of self-made database was given as input. 66% split validation was carried out to check the accuracy. The validation showed 87.5% accuracy.

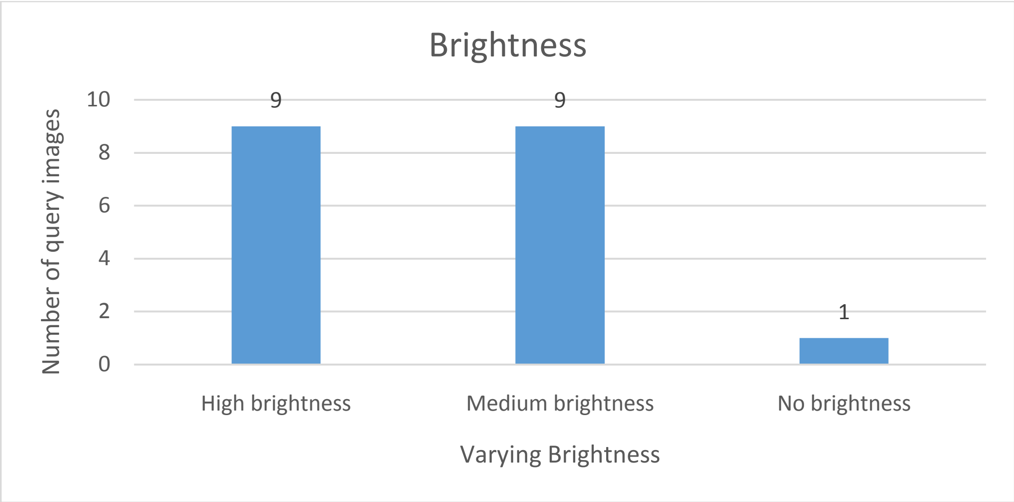


Figure 17 Graphical representation of analysis on AT&T database on the basis of brightness

The system was tested with 10 query images with varying level of brightness. The above graph shows the statistics of the testing.

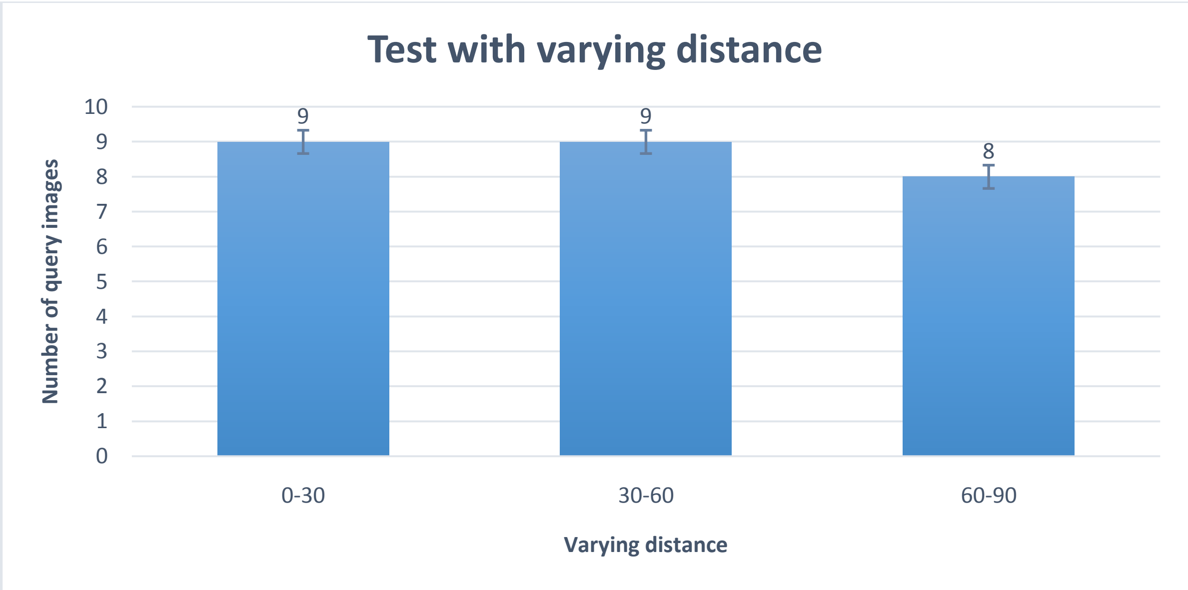


Figure 18: Graphical representation of analysis on the basis of distance between front camera and person.

The system was tested with capturing facial images with front camera at varying distance between the camera and the person. The above graph shows the statistics of the testing.

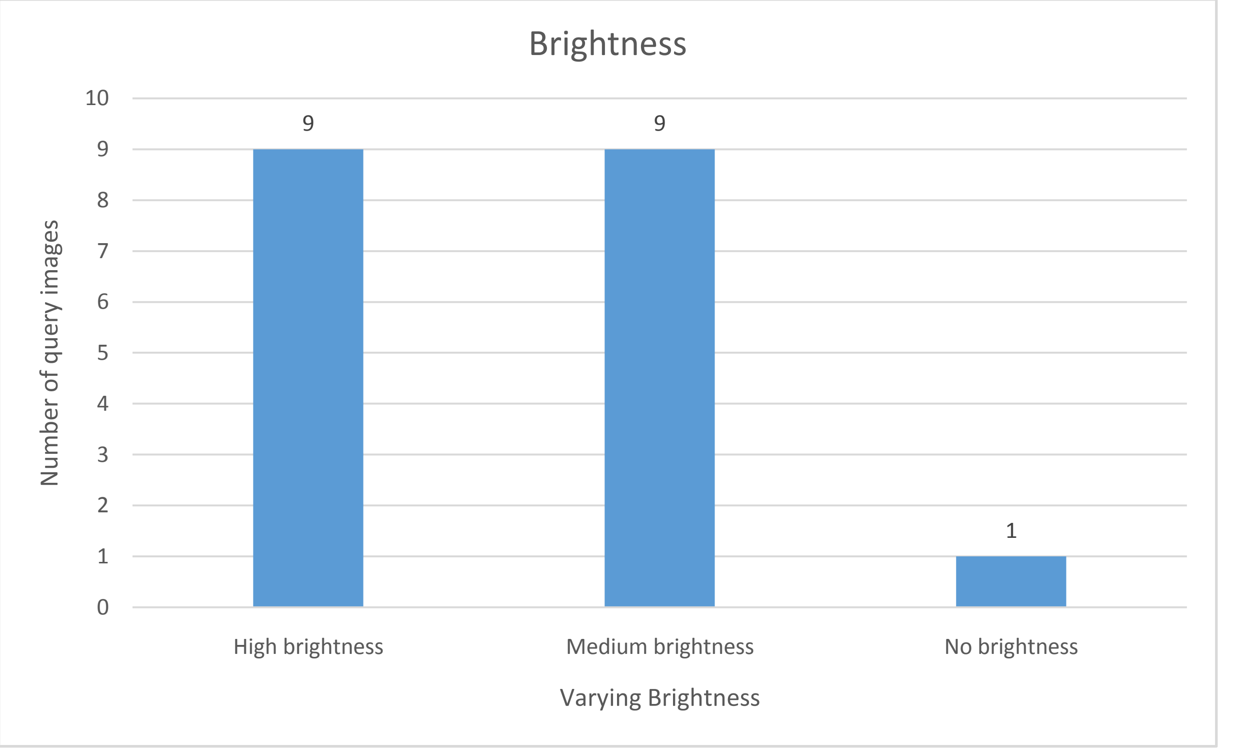


Figure 19: Graphical representation of analysis on the basis of different level of brightness

The system was tested with capturing query image in different brightness level. Ten images were tested.

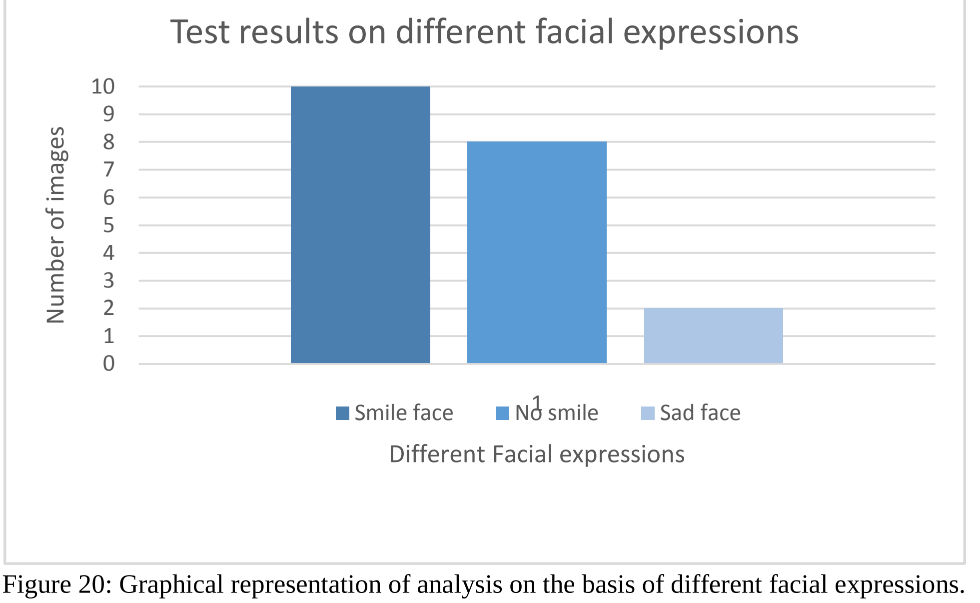


Figure 20: Graphical representation of analysis on the basis of different facial expressions.

The system was tested with query images with different facial expressions such as smiling face, no smile and sad face.

ISSN: 2321-8169
34 – 45

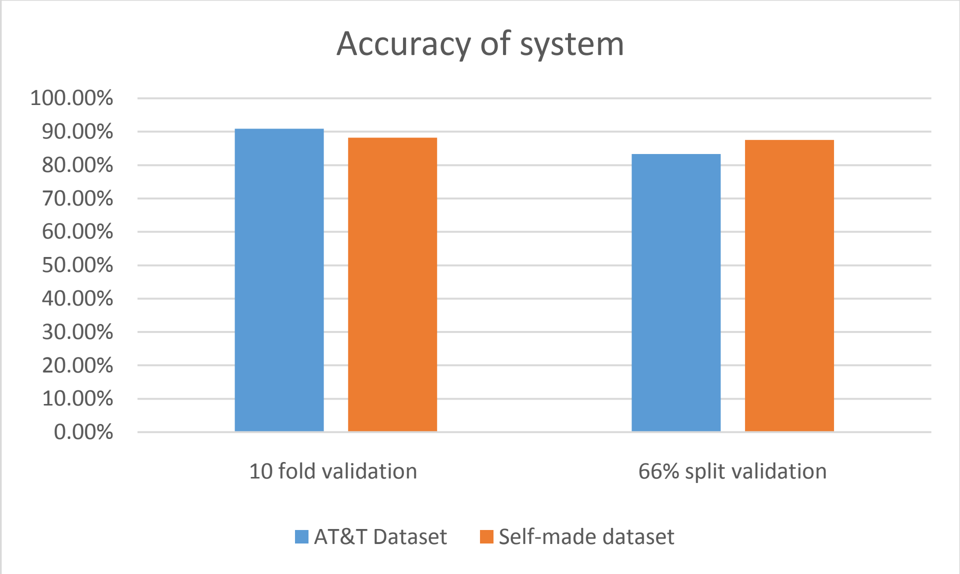


Figure 21: Graphical representation of analysis of accuracy.

The system was tested using Weka software. The csv file of the training set of self-made database was given as input. 10 fold cross validation was carried out to check the accuracy. The validation showed 88.24% accuracy. The system was tested using Weka software. The csv file of the training set of AT&T database was given as input. 10 fold cross validation was carried out to check the accuracy. The validation showed 90.91 % accuracy.

IX. Conclusion and Future Scope

Mobile based visual search is an important topic in recent research. Even though there is lots of development in this topic the scope for development still exists. The proposed system carries out face recognition over mobile device with accurate results to the maximum possibility. With constant change in the physical appearance of human body there is still some error rate possible.

The system can further be improvised by reducing the features extracted. With the growing age of a person the facial features of a person may change so the system can be improvised by updating the database on regular intervals.

References

[1] QawlanAkariman , AgungNugrohoJati , AstriNovianty," Face Recognition Based on the Android Device Using LBP Algorithm," in proceeding of 2015 International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC) 978-1-4799-8975-1/15/$31.00 ©2015 IEEE School of Electrical Engineering Telkom University Bandung, Indonesia

[2] P. Viola and M. Jones, "Rapid object detection using a boosted classifier of simple features," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2001, vol. 1, pp. 511–518.

[3] AbdenourHadid, MattiPietik¨ainen and TimoAhonen, "A Discriminative Feature Space for Detecting and Recognizing Faces," in Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04) 1063-6919/04 $20.00 © 2004 IEEE

[4] Divyarajsinh N. Parmar, Brijesh B. Mehta, " Face Recognition Methods & Applications" in the proceedings of Divyarajsinh N Parmar et al , Int.J.Computer Technology &Applications,Vol 4 (1),84-86 IJCTA | Jan-Feb 2013

[5] S. Suhas, A. Kurhe, Dr.P. Khanale, "Face Recognition Using Principal Component Analysis and Linear Discriminant Analysis on Holistic Approach in Facial Images Database", IOSR Journal of Engineering e-ISSN: 2250-3021, p-ISSN: 2278-8719, Vol. 2, Issue 12 (Dec. 2012), ||V4|| PP 15-23

[6] M. A. Turk and A. P. Pentland, "Face Recognition Using Eigenfaces", 1991.

[7] Jo Chang-yeon , "Face Detection using LBP features" CS 229 Final Project Report December 12, 2008

[8] Jiawei Hu, LiangruiPeng, Li Zheng, "XFace: A Face Recognition System for Android Mobile Phones" proposed in 2015 IEEE 3rd International Conference on Cyber-Physical Systems, Networks, and Applications.