Review and Comparison of Face Detection Algorithms

Kirti Dang
CSE Department, ASET
Amity University
Noida, India
dangkirti@gmail.com

Shanu Sharma
CSE Department, ASET
Amity University
Noida, India
Shanu.sharma16@gmail.com

***Abstract*— With the tremendous increase in video and image database there is a great need of automatic understanding and examination of data by the intelligent systems as manually it is becoming out of reach. Narrowing it down to one specific domain, one of the most specific objects that can be traced in the images are people i.e. faces. Face detection is becoming a challenge by its increasing use in number of applications. It is the first step for face recognition, face analysis and detection of other features of face. In this paper, various face detection algorithms are discussed and analyzed like Viola-Jones, SMQT features & SNOW Classifier, Neural Network-Based Face Detection and Support Vector Machine-Based face detection. All these face detection methods are compared based on the precision and recall value calculated using a DetEval Software which deals with précised values of the bounding boxes around the faces to give accurate results.**

Keywords*— *Face detection, Viola-Jones face detector, SMQT Features and SNOW Classifier, Support Vector Machines-Based face detection, Neural Network-Based Face Detection, Precision, Recall.

I. INTRODUCTION

The pace in development of technology is because the computers are becoming intelligent. These intelligent computers lead to computer-human interaction which is a new era helping in exploring various fields[1]. Face detection i.e. the sub domain of object detection is one of the field based on computer-human interaction. Object Detection is a process that deal with detecting instances of objects from a particular class (such as people, cars, buildings or faces)in an image or video. There are numerous number of applications of object detection in this upcoming era like face detection, pedestrian detection etc. Face detection is the first basic step to all the facial analysis methods like face recognition, face alignment, face modeling, face verification and facial tracking etc.[2] Firstly, the face is detected by the computer then only the further analysis can be done. The face detection procedure is performed on an arbitrary image which determines the faces in an image by differentiating the patterns formed the faces from the other patterns and returns the dimensions of each face[3]. This can be achieved if face modeling and face segmentation is done accurately. This also requires considering various aspects of facial appearance like pose, resolution, focus, noise, color[4], shadow and occlusion[5] etc. The general process of face recognition is shown in Fig 1. In the past few decades the face detection is being in research very much as it is a trivial task for computers. Various algorithms are proposed for this task. In this paper four basic algorithms are discussed[6] which are used for object detection:1. Viola-Jones 2.SMQT Features and SNOW Classifier 3.Neural Network-Based Face Detection 4. Support Vector Machines-Based face detection.

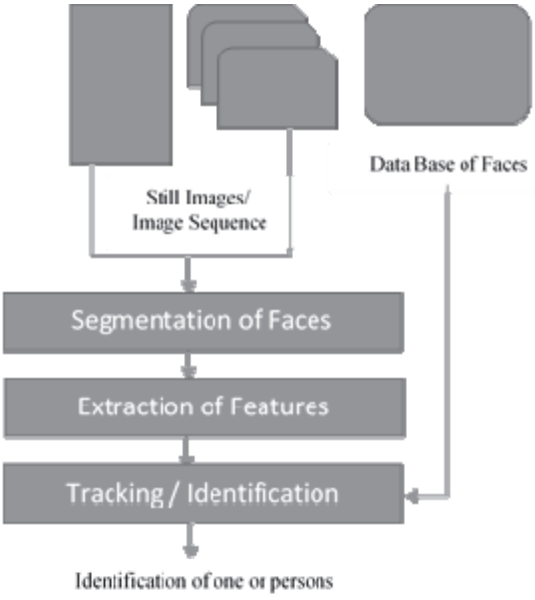


Figure 1. Process of face detection

After the detection of face using these algorithms, all the algorithms are compared on the basis of precision and recall

The values of precision and recall is calculated using DetEval Software[7] which takes two different sets of XML files as input. The first XML consist of ground truth rectangle dimensions value of the and the other one has the dimensions of detected rectangles (bounding boxes) of face locations. These two XMLs are compared and performance measures are calculated. Traditionally for the evaluation of object detection algorithms, a single detection file and its corresponding ground truth file the recall and precision, can be calculated defined as follows:

$$R_{OB} = \frac{\text{No. of correctly detected rectangles}}{\text{No. of rectangles in the database}}$$

$$P_{OB} = \frac{\text{No. of correctly detected rectangles}}{\text{Total No. of detected rectangles}}$$

Precision and recall are the measures used to measure the relevance of the algorithm. Precision also known as positive predictive value which states the fraction of retrieved instances relevant and gives information about the false alarms. Recall gives the information about the number of objects detected or relevant instances retrieved. For the perfect system the more it is close to 1 more accurate the results of the algorithm[8].

II. Face DetectioN Algorithms

A. *Viola Jones Face detection Algorithm*

The Viola Jones Object Detection framework [9]was proposed by Paul Viola and Michael Jones in 2001 which was the first framework to give competitive object detection rates. It can be used for detecting objects in real time but it is mainly applied to face detection application. The detection rate of this framework is quiet high (true positive rate) and very low false-positive rate which makes the algorithm as robust and it also processes the images quickly. It is main goal is face detection not recognition that is to distinguish faces from non-faces which is as the first step recognition. Four main steps followed in this algorithm are:

- Haar Feature Selection: As all human face have similar properties like eyes region is darker than the nose bridge region etc. These properties are compared using Haar feature also known as digital image features based upon Haar basis functions.
- Creating an Integral Image: An integral image is formed by computing the rectangles adjacent to the rectangle present at (x,y) into a single image representation which helps in speeding the procedure.
- Adaboost Training: In this Adaboost learning algorithm is used to build the classifier to be trained. This algorithm helps in finding small critical visual features from large set of potential features.
- Cascading Classifiers: Process of combining the classifiers which quickly discards the background windows so that more computation can be performed on face-like regions.

B. *Successive Mean Quantization Transform(SMQT) Features and Sparse Network of Winnows (SNOW) Classifier Method*

This method is divided into of two phases. In the first phase(face luminance), pixel information of an image is gathered which can be later used in the second phase. The second phase (detection phase),for the feature extraction in object detection local SMQT features, as these features were able to cope up with illumination. For better results and to speed up the computation of the standard SNOW classifier , split up SNOW classifier is used which trains a single classifier and divided arbitrarily into several weak classifiers in cascade. This new approach for face detection is computationally faster and has high detection rates[10].

C. *Neural Network-Based Face Detection*

Neural Network inspired by human brain composed of simple artificial neurons also known a perceptrons are connected to each other in multiple layer. Each perceptron consist of mathematical function either a summation function(sum up all the inputs and pass) or threshold function(limit the input).This is a self learning network which is trained and not explicitly programmed[11]. In the case of face detection, neural network system examines each and every window (considerably small in size) to determine whether it consist of face or not. It reduces the computational task as it doesn't require to train with non-face images[12]. This process is divided into two steps. In the first step, region of the image (20*20 pixel size) is fed as input to filter made up of the neural network. The output of this filter lies between [-1, 1] depicting the absence or presence of face. The filter is applied to all the regions of the image for detection of faces. The second step is to overcome the false detections found in the first step and to increase the efficiency for better results. This is possible if all the overlapping detections of the single neural network are fused together.

D. *Support Vector Machines:*

A supervised learning model constructs a maximum marginal hyper plane or set of hyper planes in a high- or infinite-dimensional space which clusters the data into two classes and this hyper plane separate the one class from another[13]. The margin is defined as the distance between the hyper plane and the closest data points. Support vectors are the data points that lie on the boundary of the margin of the hyper plane. When this method implied on the object detection then the positive samples are assumed to be the object classes amongst the classes and the negative samples are the non-object classes. For manually generated training examples labeled as object and non object, quadratic programming is used to compute the hyper plane from infinite number of hyper planes present.

For non linear classification using SVM , kernel trick is applied to the features extracted from the input. It transforms the non linearly separable data into a high dimensional space where it is likely to be separable. These kernels can be either polynomial kernels like Gaussian kernel and two perceptron kernel or radial kernel like sigmoid function. The main task is

to find the appropriate kernel for the given problem which is not easy. Every time the kernel is selected it is tested for the classification problem and it may not yield to good results as produced by the sample set.

It can be used in many applications like detecting pedestrians or face detection. Haar wavelets applied on the positive and negative training examples to extract the features which helps in discriminating the classes[14].

III. COMPARATIVE EVALUATION

For the comparison of the above specified face detection algorithms, precision and recall is calculated using the DetEval Software. DetEval works on xml files which input the measurements of the bounding box of the detected object and the ground truth value. These two xmls are compared for the more accurate evaluation of the detection algorithms. As per the requirement, the ground truth value of the sample image (Fig 2) assumed is calculated using an api named ' Betafaceapi'[15] which gives précised dimensions of the bounding boxes(Fig 3). After obtaining the ground truth value with adequate dimensions, all the face detection algorithms are run in MATLAB R2016b and the dimensions are saved in a matrix for the second xml for evaluation.

Figure 2. Sample Image

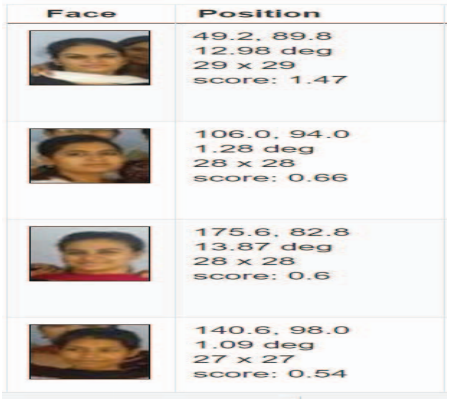
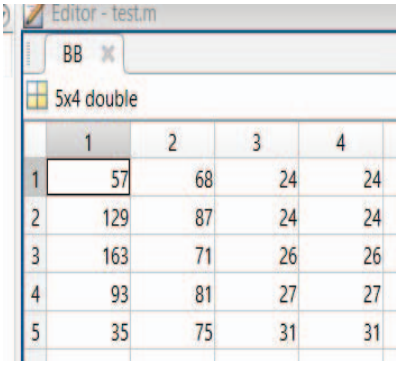


Figure 3. Ground Truth Value

The implementation of Viola jones in MATLAB 2016b was done using vision.CascadeObjectDetector System present in computer vision system toolbox. This cascade object detector is pre-trained using Viola-Jones detection algorithm to detect the face and other type of objects. The dimensions of the bounding boxes are retrieved (Fig 5) and substituted in DetEval xml file (Fig 6).This gives the values of precision and recall as:

Precision=0.27312 Recall=0.27312

Figure 4. Face detection using Viola Jones



	1	2	3	4
1	57	68	24	24
2	129	87	24	24
3	163	71	26	26
4	93	81	27	27
5	35	75	31	31

Figure 5. Dimensions of bounding boxes using Viola Jones

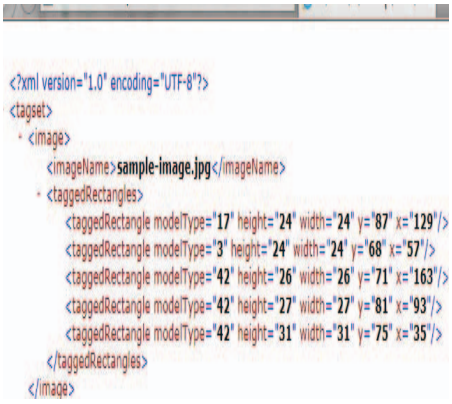


Figure 6. Values subsitututed in xml

Similarly, code for SMQ, SVM and Neural Network for facial detection is implemented in MATLAB 2016b and the substituted in DetEval for values of precision.

For SMQT Snow features are used and a single classifier is trained using the face and non face images and the cascades are formed by dividing the classifier. The results are shown in Fig. 7. The calculated values are:

Precision=0.26792 Recall=0. 26792

Figure 7. _SMQT Features and SNOW Classifier Method

In Neural network for face detection a network is trained using training set of images of faces and then the test image is put as an input to the neural network gives result values as:

Precision=0.0339450 Recall=.037582

The results of neural network based face detection is shown in Fig. 8.

Figure 8. Neural Network method

In SVM, the SVM is created for detection of faces, trained using face and non-face images and gabor filters are applied for which the values are calculated as follows and results are shown in Figure 9.

Precision=0.01392850 Recall=.00835708

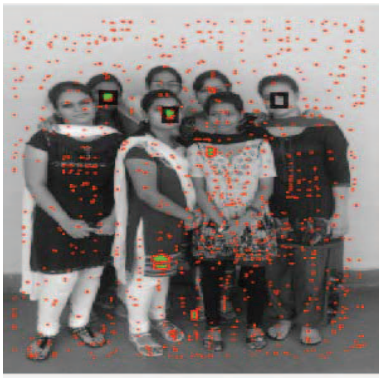


Figure 9. SVM Method

The comparative analysis of the explained face detection algorithms on the basis of precision and recall is given in Table 1.

TABLE I. COMPARATIVE ANALYSIS

S.No	Face Detection Algorthims	Precision	Recall
1	Viola-Jones face detector	0.27321	0.27321
2	SMQT Features and SNOW Classifier	0.26792	0.26792
3	Neural Network-Based Face Detection	0.339450	0.037582
4	Support Vector Machines-Based face detection	0.01392850	0.00835708

IV. CONCLUSION

The above experiment conduced it can be estimated on the basis of the values of precision and recall the highest value is of Viola-Jones followed by SMQT Features and SNOW Classifier Method then Neural Network for facial detection and at last is the Support vector Machine.So the best amongst all of these algorthims is Viola-Jones for the face detection. For future work, various other object algorithms can be compared using this procedure.

REFERENCES

[1] Cha Zhang and Zhengyou Zhang, —A Survey of Recent Advances in Face Detection: Microsoft Research Microsoft Corporation, June 2010

[2] Himani S. Parekh, Darshak G. Thakore , Udesang K. Jaliya ,A Survey on Object Detection and Tracking Methods International Journal of Innovative Research in Computer and Communication Engineering

[3] M.-H. Yang, D. J. Kriegman, and N. Ahuja. Detecting faces in images: A survey. IEEE Trans. on PAMI, 24(1):34–58, 2002.

[4] R. Hussin, M. Rizon Juhari, Ng Wei Kang, R.C.Ismail, A.Kamarudin, Digital Image Processing Techniques for Object Detection From Complex Background Image. International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012).

[5] Ragini Choudhury Verma, Cordelia Schmid, and Krystian Mikolajczyk, ―Face Detection and Tracking in a Video by Propagating Detection Probabilities‖, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 25, NO. 10, OCTOBER 2003.

[6] Ms. Varsha Gupta, Mr. Dipesh Sharma,A Study of Various Face Detection Methods. International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 5, May 2014.

[7] Christian Wolf and Jean-Michel Jolion. Object count/Area Graphs for the Evaluation of Object Detection and Segmentation Algorithms, In *International Journal on Document Analysis and Recognition* , 8(4):280-296, 2006.

[8] Akcay, H. G., Aksoy, S., July 2008. Automatic detection of geospatial objects using multiple hierarchical segmentations. IEEE Transactions on Geoscience and Remote Sensing 46 (7), 2097–2111.

[9] Viola, Paul and Michael J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. Volume: 1, pp.511–518.

[10] Ingvar Claesson. "Face Detection using Local SMQT Features and Split up Snow Classifier", 2007 IEEE International Conference on Acoustics Speech and Signal Processing-ICASSP 07, 04/2007

[11] H.A. Rowley, S. Baluja, and T. Kanade, ―Neural Networks Based Face Detection,‖ IEEE Trans. Pattern Analysis an Machine Intelligence, vol. 20, no. 1, pp. 22-38, Jan. 1998.

[12] K.K. Sung and T. Poggio, ―Example-Based Learning for View- Based Human Face Detection,‖ IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 39-51, Jan. 1998.

[13] BOSER, B., GUYON, I. M., AND VAPNIK, V. 1992. A training algorithm for optimal margin classifiers. In *ACM* Workshop on Conference on Computational Learning Theory (COLT). 142–152.

[14] Yilmaz, A., Javed, O., and Shah, M. 2006. Object tracking: A survey. *ACM Comput. Surv. 38*, 4, Article 13 (Dec. 2006), 45 pages.

[15] Oleksandr Kazakov, Munich, Germany-https://www.betafaceapi.com