International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 10 | Oct 2018 www.irjet.net p-ISSN: 2395-0072

A Review on Various Techniques for Face Detection

Nikita Raut[1], Varun Dhuldhoya[2]

[1,2]*Student, Dept. of Computer Engineering, KJ Somaiya College of Engineering, Maharashtra, India*

Abstract - *Face detection has been an intriguing area in the field of computer vision and machine learning, especially during the past several years. Face detection is the process of automatically locating human faces in digital images or videos. Faces are one of the most prevailing and very specific objects in visual media which can be tracked with the help of face detection algorithms. In spite of the inherent complexity of environment (illumination, pose, color coding, occlusion etc.), we have seen the development of some intelligent and accurate face detection algorithms. In this paper we have made an attempt to study the various methods namely Viola Jones, Genetic Algorithm, CNN, SVM, Hough transform with convolution neural network, MMX feature extraction, Min max with embedding for face detection.*

Key Words: **face detection, viola and jones, soft computing, artificial intelligence, neural networks, machine learning, accuracy**

1. INTRODUCTION

The subject of face detection and recognition is as old as computer vision. Researches on automatic face detection began from 1960s, but over the past decade, this groundbreaking technology has not just become viable, it has become widespread. Although other methods of identification such as fingerprints, or iris scans can be more accurate, face detection and recognition remains a major focus of research because of its non-invasive nature. Face detection is primarily the specific case of object detection. Object detection deals with detecting instances of objects such as people, automobiles, buildings or faces in an image or a video.

Face detection refers to identifying people's faces within digital images or videos. It is the first basic step to all the facial analysis methods like face recognition, face modeling, face tracking, face verification etc. First, the face has to be identified by the computer then only the further evaluation can be done.

Face detection has a wide range of applications. It is being used in photography as a way to help cameras autofocus on people's faces. Another application is smile recognition, which helps people take pictures at the perfect moment in order to capture smiles.

One of the most significant applications of face detection is a facial recognition. A facial recognition system matches an individual's face instantly against a database of photographs in order to establish identity. Facial recognition tools are used to secure phones and apps.

Face detection is being used by some companies in order to spot people walk by a certain area. It can also be used to predict age, gender and other factors. Biometric techniques like identifying fingerprints are now being replaced by a more efficient biometric technique- face detection.

Face recognition is one of the most relevant applications of image analysis. It's a true challenge to build an automated system which equals human ability to recognize faces. One aspect that has slowed down progress in this area is the lack of base lining studies whose goal is not just the development of complete systems but the analysis of how the different pieces of a system contribute to its success. The goal of this study then is to perform a systematic comparison of the different techniques. In upcoming sections, we discuss some recent approaches like Viola Jones, Genetic Algorithm, CNN, SVM, Hough transform with convolution neural network, MMX feature extraction, Min max with embedding which perform accurate skin cancer detection with the help of various algorithms and discuss the advantages and disadvantages of the different techniques used.

2. LITERATURE REVIEW

A. Viola Jones

The Viola Jones Object Detection framework [9] was proposed by Paul Viola and Michael Jones in 2001, where it was used for detecting objects. Ever since, it has also been used for face detection which is a sub-domain of object detection. In [1], Vikram and Dr.S.Padmavathi present a system based on Viola-Jones Cascade Object Detector, for detecting faces in an image and locating the facial features like eyes, nose and mouth in it. The advantage of this system is that in addition to detecting faces it also detects the facial features, which is otherwise a difficult task. In this algorithm the distance between the two eyes is pre-defined, thus when any image is given as an input this distance will be compared with the actual distance and therefore the eyes will be detected. It detects nose and mouth in a similar way.

The Viola - Jones consists of 3 techniques for the facial parts detection: 1) The Haar like features for the feature extraction. The presence of the features in any of the input image is determined by the Haar features 2) Ada boost machine-learning method for detecting the face. 3) Cascade classifier used to combine many of the features efficiently. The Haar Cascade classifier is the main part of the face detection. The Haar features have been allowed to convolve and scan through the images from the top left corner and it ends the detection of face processing at the right bottom corner.

Then, the best features among the 1,60,000 features have been found using Adaboost technique. And finally the cascade method is used, that will pass the case if the image is of a human face, otherwise it will be considered as a non-face region.

B. Skin Color Based Detection

Hakam et al. [5] propose a method to identify human skin and faces from colored images where the human skin colored pixels are detected via a matrix that contains a collection of possible skin color values which represents the skin color of most people. They have used the circularity feature to distinguish human faces from other objects with similar color.

One of the most important information in identifying a human face is the color. However, the color is not enough to detect and localize human faces in images since other parts of the body often have the similar color. To solve this problem, they have used the geometry of the face since it differs from other parts of the body.

They created a database containing all the pixels that represent the colors of the human skin. The bigger the database, the better the accuracy but slower the computational speed. The algorithm will compare all the pixels in the database with each pixel in the possessing image, and if the pixel in the processing image is close enough to one of the pixels in the database image, then that pixel will be labelled as 1, meaning it is a skin color, if not the pixel will be labelled as 0 i.e. non-skin pixel.

The produced image goes into a set of modifications. The first modification is removing the small areas which are present away from the skin in the image and filling the holes present on the skin. Another process employed is to create a morphological structuring to remove the dots and the non-uniformity in the image. . If the background pixels have values similar to the skin color, then the background and the faces will be as one object. In order to solve this problem edge detection was used, where they transformed the input image into gray level format and then the edges in the image were found using sobel and canny detectors. Once we get the objects, circularity is used to find how circular the object is, so if the circularity value is above the threshold, a face has been detected.

The main drawback of this method is the creation of the reference matrix database. Another problem is the possibility of errors. However, number of errors can be reduced by considering more features to distinguish the face form other similar skin colored objects.

C. Detection using Min-Max Features Enhanced with Locally Linear Embedding

In their paper Rahmat et al. [12] have improved the classification accuracy in their previous work by integrating a technique called Locally Linear Embedding (LLE).

Essentially, the MMX feature extraction method equally divides an image into three horizontal, and three vertical face regions. Then, the sum of horizontal and vertical face regions was calculated and combined to form the final MMX features. After MMX feature extraction LLE transformation is done. LLE resolves a low dimensional representation that approximates the high dimension features this method is very efficient and under certain circumstances it is possible to derive an embedding only from the geometric situation without looking at scale, distance or connectivity between interval data.

First parameter is the nearest neighbors (K) to support the global geometry of the data. Second parameter dmax defines how much data has to be retained by LLE. Next a set of weights is computed that can reconstruct each data point. The embedding vector Y, is computed with the aforementioned weights .After this classification is performed using multilayer perceptron .Number of hidden layers is varied using trial and error to get the best results .Both hidden and output layers are use tangent sigmoid activation function. The training algorithm used was Scale conjugate Gradient as it shows excellent results in pattern classification.

Two MLPS were taken one for LLE and one for non LLE data. For non LLE data All the hidden units tested produced good to excellent accuracy (above 80%), with some variation from 81.8% to 95.2% For LLE data all hidden units tested produced excellent accuracy with minimum as 94.4% and maximum as 96.4%

Based on the observations, it was found that MMX was already a powerful feature representation algorithm even before applying LLE features and after applying LLE features the representation of MMX features had improved. This is because the LLE algorithm can represent the features in such a way that it maximizes the separability between face and non-face cases.

D. Convolution Neural Networks

Chenghao Zheng et al. [6] propose a real time approach for face detection by utilizing a single end-to-end deep neural network with multi-scale feature maps, multi-scale prior aspect ratios as well as confidence rectification. This system has been proposed to overcome the major shortcoming of the existing CNN-based face detectors such as need of high computational cost and slow calculation. They have proposed a face detection algorithm based on a single end-to-end CNN model with similarities to (single shot multibox detector) SSD. This model utilizes the three major advantages of SDD 1) Multiscale feature maps which guarantee the scale invariance to some extent 2) Convolution detectors, which predict the objects 3) Default boxes and aspect ratios. For confidence rectification, regression method is utilized instead of SoftMax function – a function used by the SSD, because it is easier to determine a face if the box is big enough with a higher resolution. In addition to this, human intuition is used where the easily detected faces fall

in class 1, whereas difficult faces fall in class 0.5. Finally, in order to improve the speed of the network, they have reduced the network depth and the number of convolution maps.

Yingxin Lou et al. [7], present a novel object detection network based on a pre-trained and multi-feature VGG-16 network. This model has been proposed as an improvement to faster R-CNN. According to [8], Faster R-CNN has certain drawbacks like (1) It only adopts Conv5 feature map which is not accurate enough for object detection and bounding (3) Faster R-CNN generates proposals using anchors which is coarse for enclosing the various objects. (4) Cannot enclose objects precisely. To overcome these shortcomings, they have proposed the Pt-Net model. Their model is first initialized by a pre-trained VGG-16 model and its own CNN, then the image is provided to the pre-trained CNNs, the outputs of the selected layers are aggregated into multi-feature maps and the proposals are generated on Conv5 feature map via particle filter, the proposals are then mapped to multi feature maps and the maps are cropped, finally, the face is detected using a novel overlap loss function.

E. Hough Transform with Convolution Neural Networks

In their paper, Oshin et al. [9], develop novel feature extractor method by the combination of Hough transform and CNN defining face features matrix dimension.Adaptive cascade DCNN represents a fine CNN model which is able to extract high quality features which enhances the predictions or initialization. When the initialization is obtained than deep convolutional neural network provide two level convolutional networks to rectify every landmark individually received from local receptive fields. In Deep CNN technique the major facial key points like eyes, nose tip, mouth (M) are detected using fine dot point while in Hough transformation the horizontal central line of higher threshold is measured for the detection purpose.

Although both ACDCNN and Hough transform - CNN is based on CNN and follow the cascade frame work. We show that the proposed system can achieve batter performance the initialization results and the final results of the ACDCNN

The score is calculated for each region of interest of detection and alignment, which matches the information represented by the spectrum. It can be used for color identification, inspection and other application, which can be used for comparison purpose. For image identification, set of references of image and their matching score obtained during the matching process. It is also called as the true positive rate of detection.

The faces are detected with more accuracy, in the point of mean error and lower failure rate shows that the applied method is comparatively more efficient than the earlier methods. The accuracy of the detection of faces increases ranging from 90% to 100% .and the accuracy of detection of alignment of faces is also increases.

The variation in the illumination, illustrates the robustness of the system. Therefore the future scope of the work is to develop a computer based fast method for the face detection and face recognition. This work can be extended for the study of other intra subject variations like expression, aging, occlusion, gender classification and other biometric devices.

F. Genetic Algorithm

According to [3], features extracted using neural network are not sufficient and have low accuracy. Face detection using neural networks suffers from problems while extending the classifier for newly added people and learning updated information about the existing people

In this paper, Mr. Deepak et al. uses genetic algorithm followed by Scale-invariant feature transform (or SIFT) algorithm to improve the accuracy by feature extraction. The main characteristics of the algorithm presented are the use of a combination of classification results and its incremental nature. The proposed algorithms consists of the following stages: 1) Taking the input image 2)extracting the features from image using genetic algorithm based SIFT algorithm 3)Detecting faces using feature vectors.

The first step of the genetic algorithm is selection. The fitness value for each pixel is calculated and a fitness function is used to select the child chromosomes. Then PCA-mutation was performed on the selected pixels. The mutation is necessary to insure high levels of diversity in the population. The next step is SIFT feature extraction. Then, the SIFT algorithm which is an algorithm in computer vision to detect and describe local features, will find the points whose surrounding patches are distinctive

G. Support Vector Machines

Sanjeev et al. [10] propose an improved face detection method for highly corrupted images. There are many unwanted elements present in image which are commonly known as noise that should be removed from an image for any further processing. Mean and Median filter are normally used to reduce noise present in an image and for preserving useful detail in the image. They have used Adaptive filtering to reduce noise and preserve useful information in an image as it is more selective than mean and median filtering.. Once the noise from an image has been removed the image is sent to the SVM trainer for identification of faces.

It is observed from this paper that the efficiency of using the nose structure as an identifying feature for face detection improves the success rate.

The drawback of this paper lies in the fact of its dependence on visibility of faces after the filter procedure. Future scope lies in improving the identification ratio, either by combining several feature extraction algorithms] or combining several classifiers or combining the above two methods.

H. Detection using K-means clustering and Local Binary Patterns (LBP) algorithm

Wisam et al. [11] propose a method where the targeted images were collected, and K-means clustering was applied to separate the colors that form these images. According to this method, a face should be separated from its background. In the proposed method, the targeted images were collected, and K-means clustering was applied to separate the colors that form these images. According to this method, a face should be separated from its background.

The LBP algorithm is applied to the faces in the image and then the LBP feature is divided into 6 groups 3 positive and 3 negative.

LBP is the process of returning image information in such a way that will make it a unique binary pattern.

1) The image is divided into blocks, and the procedure is applied separately for each block. 2) Each pixel should be compared with eight neighbors starting from the top left in a clockwise or counterclockwise direction. 3) The difference between center points and surrounding pixels is compared. If the value of a neighbor pixel is lower than that of a center pixel, then it should take a value of "0;" otherwise, the value should be "1." 4) The histogram for each block is computed, which will result in a feature vector. 5) The histogram is normalized (optional). 6) The histogram of all blocks is concatenated, which will provide a unique binary description for the entire image. 7) To classify information, a machine-learning algorithm can be used.

The LBP algorithm is applied to the faces in the image and then the LBP feature is divided into 6 groups 3 positive and 3 negative .On the basis of the distribution for these points, the classifying element is able to separate the face from a non-face area.

The first procedure of the algorithm involves segmenting an image to separate the face from the scene by using the K-means clustering code and reducing the time consumed for face detection.

As stated previously, the main concept of face detection is to identify the differences between the LBP for the face sample and that for the block taken from the image. The similarity between the two feature vectors is then calculated, and the face and non-face are segregated. The final result of LBP on an image can be represented by a vector, which contains a histogram for every block.

The main concept of the proposed algorithm is to classify these feature vectors to estimate the object it describes, whether or not it is a face or non-face object. Results show the efficiency of the algorithm in detecting faces with a high rate of accuracy. Detecting faces in many poses, such as turned and upside-down faces, is our main contribution to the literature.

Technique Used	Advantages	Disadvantages
Viola-Jones	In addition to detecting faces it also detects facial features	[19]Detector is most effective only on frontal images of face. It fails to detect face in an image with Partial features. This condition exists when face is slightly rotated.
Skin color based	Simplicity	Creation of reference matrix dataset, possibility of errors in many cases
Min-Max Features Enhanced with Locally Linear Embedding		
Convolution Neural Network	Improved Accuracy	Time consuming for a huge dataset
Hough Transform with convolution neural network		
Genetic Algorithm	Solves the problem of extending the classifier for newly added people and learning updated information about the existing people	Poor results if poor fitness function chosen
SVM		
K- means Clustering and Local Binary Patterns		

3. CONCLUSIONS

In this study, we have examined various techniques that give high accuracy for face detection. Face detection requires various stages like pre-processing, segmentation and feature extraction. This survey focuses on different strategies like Viola Jones, Genetic Algorithm, CNN, svm, Hough transform with convolution neural network, MMX feature extraction, Min max with locally linear embedding. As per the review, each algorithm is found to have its advantages and disadvantages. However, amongst the analysed algorithm, the Min Max with locally linear embedding algorithm has the least amount of disadvantages and thus, outweighs other algorithms like Viola-Jones, K-mean clustering and neural networks

REFERENCES

[1] Vikram, K., and S. Padmavathi. "Facial parts detection using Viola Jones algorithm." Advanced Computing and Communication Systems (ICACCS), 2017 4th International Conference on. IEEE, 2017..

[2] Kavitha, J. C., and A. Suruliandi. "Texture and color feature extraction for classification of melanoma using SVM." Computing Technologies and Intelligent Data Engineering (ICCTIDE), International Conference on. IEEE, 2016.

[3] Mishra, Deepak, and Shaila Chugh. "Face detection using genetic based SIFT algorithm." Green Computing, Communication and Conservation of Energy (ICGCE), 2013 International Conference on. IEEE, 2013.

[4] Munteanu, Adrian, et al. "Wavelet-based lossless compression scheme with progressive transmission capability." International Journal of Imaging Systems and Technology 10.1 (1999): 76-85.

[5] Shehadeh, Hakam, Audai Al-khalaf, and Mahmood Al-khassaweneh. "Human face detection using skin color information." Electro/Information Technology (EIT), 2010 IEEE International Conference on. IEEE, 2010.

[6] Zheng, Chenghao, Menglong Yang, and Chengpeng Wang. "A Real-Time Face Detector Based on an End-to-End CNN." Computational Intelligence and Design (ISCID), 2017 10th International Symposium on. Vol. 1. IEEE, 2017.

[7] Lou, Yingxin, et al. "PT-NET: Improve object and face detection via a pre-trained CNN model." Signal and Information Processing (GlobalSIP), 2017 IEEE Global Conference on. IEEE, 2017.

[8] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems. 2015.

[9] Misra, Oshin, and Ajit Singh. "An approach to face detection and alignment using hough transformation with convolution neural network." Advances in Computing, Communication, & Automation (ICACCA)(Fall), International Conference on. IEEE, 2016.

[10] Kumar, Sanjeev, Asutosh Kar, and Mahesh Chandra. "SVM based adaptive Median filter design for face detection in noisy images." Signal Processing and Integrated Networks (SPIN), 2014 International Conference on. IEEE, 2014.

[11] Alobaidi, Wisam H., et al. "Face Detection Based on Probability of Amplitude Distribution of Local Binary Patterns Algorithm."

[12] Hidayat, Rahmat, et al. "Face Detection using Min-Max Features Enhanced with Locally Linear Embedding." (2018).

[13] Hamdy, Ahmed, Mohamed Elmahdy, and Maha Elsabrouty. "Face detection using PCA and skin-tone extraction for drowsy driver application." Information and Communications Technology, 2007. ICICT 2007. ITI 5th International Conference on. IEEE, 2007..

[14] Wong, Kwok-Wai, Kin-Man Lam, and Wan-Chi Siu. "An efficient algorithm for human face detection and facial feature extraction under different conditions." Pattern Recognition 34.10 (2001): 1993-2004.

[15] Sharifara, Ali, Mohd Shafry Mohd Rahim, and Yasaman Anisi. "A general review of human face detection including a study of neural networks and Haar feature-based cascade classifier in face detection." Biometrics and Security Technologies (ISBAST), 2014 International Symposium on. IEEE, 2014.

[16] Dang, Kirti, and Shanu Sharma. "Review and comparison of face detection algorithms." Cloud Computing, Data Science & Engineering-Confluence, 2017 7th International Conference on. IEEE, 2017.

[17] Hashemi, Seyyed Mohammad Reza, and Maryam Faridpour. "Evaluation of the algorithms of face identification." Knowledge-Based Engineering and Innovation (KBEI), 2015 2nd International Conference on. IEEE, 2015.

[18] Fitrianingsih, Elis, Endang Setyati, and Luqman Zaman. "Extraction of Eye and Mouth Features for Drowsiness Face Detection Using Neural Network." Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control 3.2 (2018): 103-112.

[19] Anitha, C., M. K. Venkatesha, and B. Suryanarayana Adiga. "High speed face detection and tracking." International Journal of Soft Computing and Artificial Intelligence, ISSN (2015): 84-90.