ISSN No:-2456-2165

Building an Efficient System to Catch Miscreants using Various Face Detection Techniques

Jeet A. Kansagara
Vishwakarma Institute of Information Technology, Pune - 411048
Academic year:4th year
Course level: Bachelor
Course name: Bachelor of Technology in Computer Engineering
Course semester:7th Semester

Supervisor Details: Yogesh Sharma
Affiliation: Vishwakarma Institute of Information Technology, Pune - 411048
Post: Professor, Department of Computer Engineering

Abstract:- In the year 2020 India has experienced an average of 80 murders and 77 rape cases daily. Most of the people don't even pay keen attention to it as it goes to around 29,000 murders and 28,105 rape cases in a year. This count is just a drop in ocean if we consider the vast population of 1.3 billion people. But this is just the count of murders and rapes if we take into consideration the other criminal acts almost 60 to 70 % cases go unregistered. According to NCRB [National Crime Records Bureau] there was rise of 28% in registered cases. By every passing day the crime rate is increasing, so the number of miscreants. It's quite tough for the law enforcement agencies to catch those miscreants and put them behind the bars. As in today's world the technology is playing an important role. Each and every thing revolves around technology. This paper focuses on using this technology specially various face detection techniques to build an efficient system to help the law enforcement agencies to catch the miscreants.

***Keywords*:-** face detection, neural networks, statistical approach.

I. INTRODUCTION

Yes, the face detection techniques are used in some of the high end law enforcement agencies for high level miscreants' detection such as terrorist. But the system they use is very complex and cost to run and implement it need tons of money. That's the reason why it's not used or been able to implement at the local level. Face detection an effortless task for humans is a complex problem for computer. But memory an effortless task to computer is limited to human capabilities. Face detection is computer vision problem which plays important part in face recognition, facial expression recognition, head pose estimation and human computer interaction. In layman's term face detection is computer technology that determines the location and size of face from a digital image. This paper focuses on building a robust system with the use of various face detection algorithms and comparing them on various terms to build and efficient, accurate system and running it on the CCTV's installed on streets by government to find the miscreants and help the respective law enforcement agencies to catch them and conclude this paper with several promising directions for future research and advancement of the system proposed.

In this paper, structured classification of the related literature is been provided which can also be used for not only building a system for catching and finding miscreants but for the person who is interested in learning the face detection techniques and to find which technique is best for particular need..

The following are some of the contributions of this paper:

Different algorithms which use various statistical and neural network approaches are been discussed, this algorithms can be implemented according to need since there is huge evolution in hardware.

Systematic discrepancies are shown between algorithms for each single method. A comprehensive comparison between all the methods is presented. A list of research challenges in face detection. The base of the system to build on the above comparison which suits best for catching the miscreants at local level. Further modifications which can be implemented in the system according to daily technological advancements.

II. MATERIALS AND METHODS

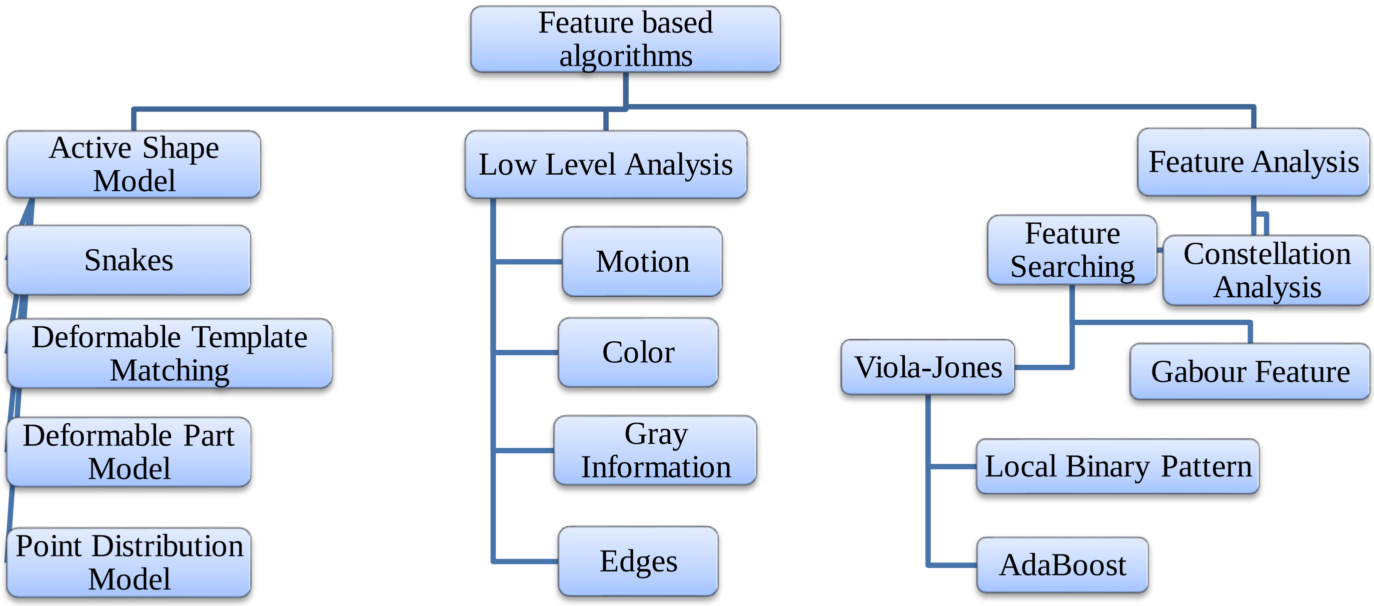


Fig. 1: Feature-based algorithms are divided into 3 categories active shape model, low level analysis, feature analysis which are further divided into sub divisions as shown in figure

III. FEATURE BASED ALGORITHMS

As mentioned earlier feature based algorithms focuses on the features in the image to find the face. It is further more divided into three parts, shown in the figure 1. The first category in feature based algorithms active shape model is an iterative process which deals with non-rigid and highly complex objects by deforming it to fit in the given sample. The second category low level analysis image pixel information is used and segmented and mostly focuses on the components of the face like eyebrows, eyes, nose, lips and many more. In third category i.e., feature analysis it's focus is on the facial geometry.

A. Active Shape Model

In ASM when system detects or views the image, it tries to link various facial features which may include nose, eyes, lips, mouth, ears, and so on. The moment when it comes close to this facial features. A mask is prepared by taking into the account of the points of the features and it's been considered as a map. The mask which is prepared can also be changed manually. Like when the system picks a shape of the face, there is the room given to user by system to adjust the shape. The greater number of images it is been trained the more accurate shapes can be formed by system without any human intervention. The Active shape models are further divided into four categories which are snakes, deformable part model (DPM), deformable template model (DTM), and point distribution model (PDM).

a) Snakes

Snakes model was first proposed by Kass et al. in 1987 [1]. The main purpose of the snakes is to locate the head boundaries [2]and, in that way, it identifies the face from the image. As the name suggests the snake's model is based on the movement of the snakes which is to expand and to shrink. First the snake is placed near the head region which is done on approximation, then it uses the qualities of snakes and gradually deforms itself into the shape of the face. Whether or not it's a good segmentation it can be matched with the energy function E(C) which is denoted by:

$$E(C) = E_{internal}(C) + E_{external}(C),$$

Here $E_{internal}(C)$ and $E_{external}(C)$ represent internal and external energy functions respectively. Where $E_{internal}(C)$ highly depends on the curve and its shape, on the other hand $E_{external}(C)$ depends on the edges of the image. The main purpose of the snake which was placed initially is to minimize the above energy function as much as possible in iterative way.

In this iterative process of minimization, the shrinking and expanding of snake is done to find the internal energy. Whereas, the curve which is formed is made to fit according to the image edges which are in close proximity to maintain the state of equilibrium. This iterative process requires high computational power. To counter this and to make the model faster, new greedy algorithms were used [4].

Snakes model are self-adapting to find there minimal energy state. By using Gaussian smoothing in the energy function the snakes can be modelled to the image scaling. In the real time snakes' model can work more efficiently as the energy function works as a noise filter, the ambiguities or other noise does not cause much of the problem to the snakes. Snakes model is much easy to manipulate.

There are some limitations of the snakes as well, if we want to extract the non-convex features the snake's model is not the best option. The criteria used for energy minimization is directly proportional to the accuracy, the higher or tighter the criteria higher the accuracy, but it is also directly proportional to computational time and space. Snakes consumes lot of space, to counter this Viterbi algorithm is used which exchanges the space for time. This implies longer time to process and find the face.

b) Deformable Template Matching (DTM)

In 1992, Yuille et al. [5] took the snakes model further by using eyes as the global feature for better finding of face from the image. Deformable template matching as the name tells it actively deforms the preset boundaries to fit according to shape of the face. Other facial features like eyes, nose, lips, mouth, ears and eyebrows plays a vital part in detecting the face along with the outer edge of the face. The problem with the traditional template based models is that it only works for rigid shape, which can be solved by DTM by adjusting itself in different shapes of face. DTM is highly recommended for the non-rigid face shapes.

DTM further works with two different approaches of forming deformable shape of face to fit, which is done by preset shapes. The shapes are made with two approaches i. polygonal templates (PT) and ii. hierarchical template (HT). As shown in figure 2 polygon template uses number of triangles and a face is made by deforming the triangles to form the face [6]. Whereas as shown in figure 3 hierarchical template uses the formation of binary tree [7].

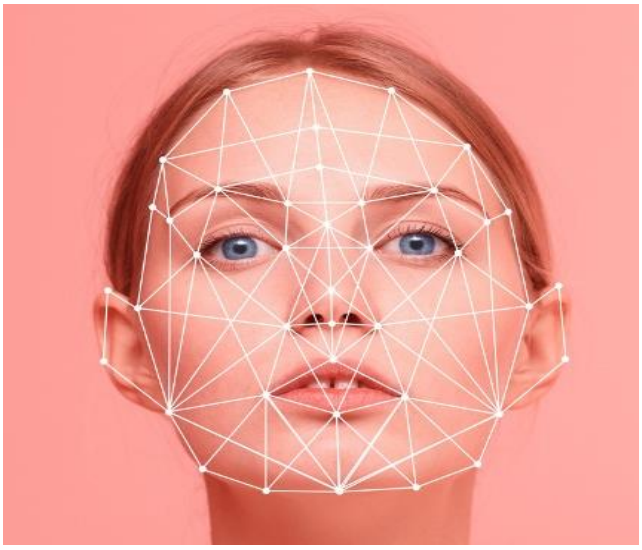


Fig. 2: Polygon template of the human face. Where the face is made by number of triangles which are deformed to form the face

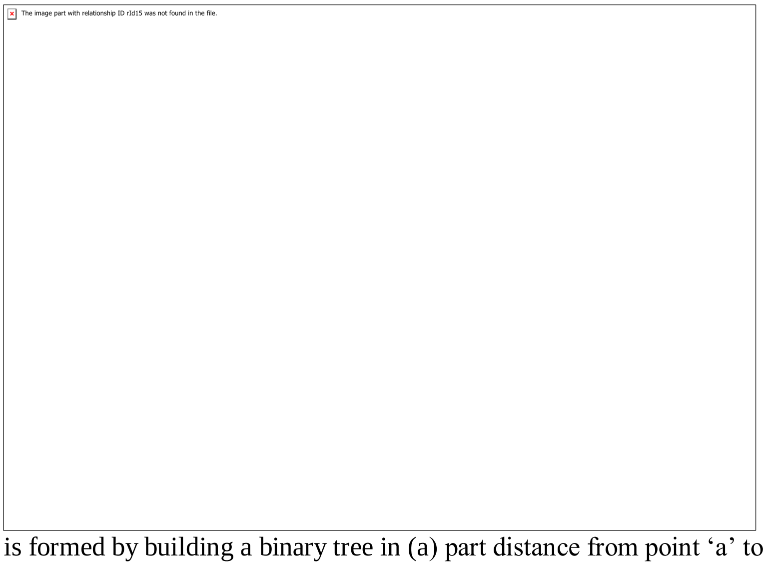


Fig. 3: Hierarchical template is formed by building a binary tree in (a) part distance from point 'a' to point 'b' is made by middle points in the deformation process in (b) part result of the deformation is converted to a hierarchical binary tree

As shown in figure 3 let us assume we want to find the curve from ‘a’ to ‘b’. Binary tree from ‘a’ to ‘b’ is been formed with ‘c’ as the midpoint. Two halves are recursively been formed with ‘c’ as the midpoint. The tree and its nodes are made in the same manner by selecting the midpoints. The local curvature is been described by sub trees, which provides us with local curvatures. By adding noise to the curvature, we can manipulate the structure which can be used to fit the global face(shape). In this deformation process the steepest gradient descent minimization of external energy is implicated.

$$E = E_v + E_e + E_p + E_i + E_{internal}\ .$$

Here, $E_v, E_e, E_p, E_i, E_{internal}$ are energy due to valley, edges, peak, image brightness and internal energy respectively.

DTM gives us better extraction because it mixes local and global information [5]. It can be used in real time as it can accommodate any given shape in the data. The energy term weights are difficult to interpolate. DTM requires lots of processing time as minimization process is a consecutive task. To add to drawbacks, DTM is very sensitive to starting position.

c) Deformable Part Model (DPM): In 1973,Fischler and Elschager et al. proposed the use of pictorial structure for face detection which is the base of DPM. The common use of DPM is to detect faces in comics and to detect human faces [8, 9]. A mask is made by modelling some discrete parts like eyes, nose, mouth, etc. separately which are rigid parts. It is a training based model. Geometric constraints are fixed between this parts like giving the distance between eyes, distance between nose and can be imaged as the springs attached to it which can be compressed or expanded to fit the given shape of the face.

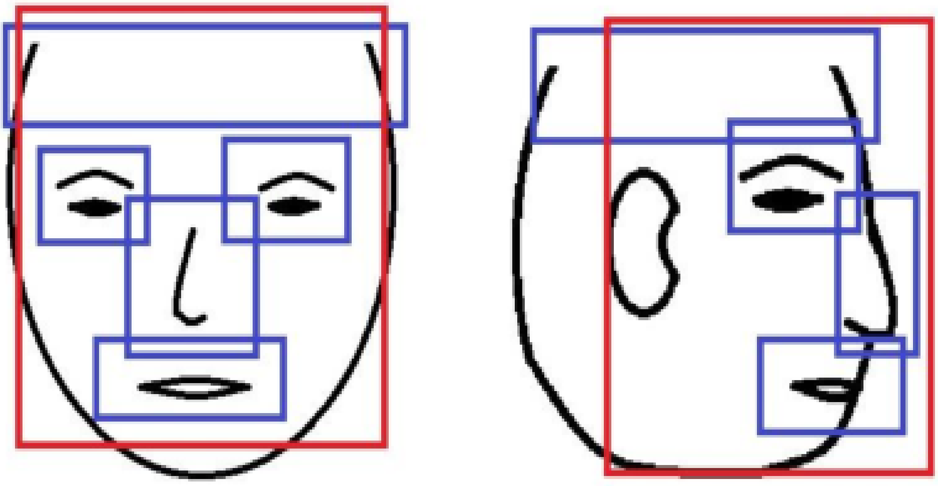


Fig. 4: The face mask is prepared, where red box is root and the blue box indicates the parts, whose distances can be changed accordingly to fit the given face which does not put too much pressure on the springs

The pictorial structure of DPM can be divided into two sections, root filter and part filter. Part filter includes the parts like eyes, nose, mouth which are fixed but the distance between them can be changed to fit the given face from the image.

In detecting faces with different shapes DPM performs well and it works very efficiently in the real time. As it uses spring analogy, it is easy to detect the face with various poses, various illumination and brightness of the image. Slowness or in technical terms bottleneck is the main challenge for the DPM, also DPM faces challenge in detection of new objects of face categories.

d) Point Distribution Model (PDM)
In 1992 Cootes et al. [11] invented the PDM, but the initial face PDM was devised by Lantis et al. [12] As the name describes in PDM points are used to describe the face. PDM heavily relies on the landmark points. A landmark point is the annotation of any image on any given shape on training set of images. The process of PDM begins by planting the landmark points on the shape of the training set of images. PDM is built with the global face shape which have the formation of eyes, nose, ears, mouth.

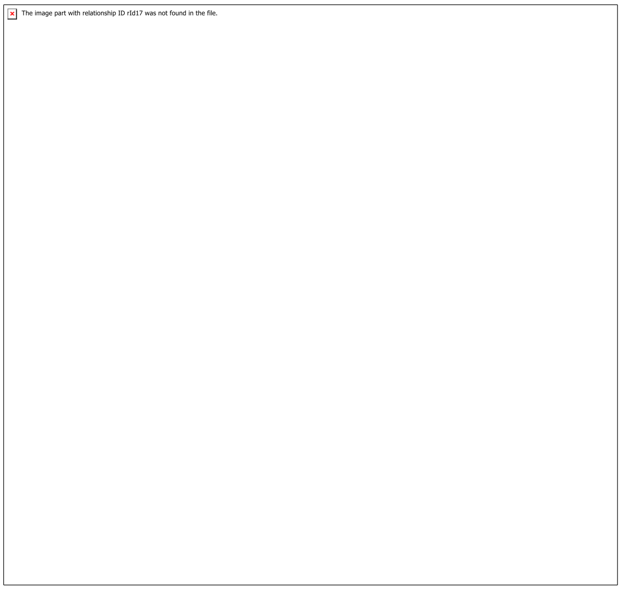


Fig. 5: A typical set of training images consisting various face elements which are mostly face shapes

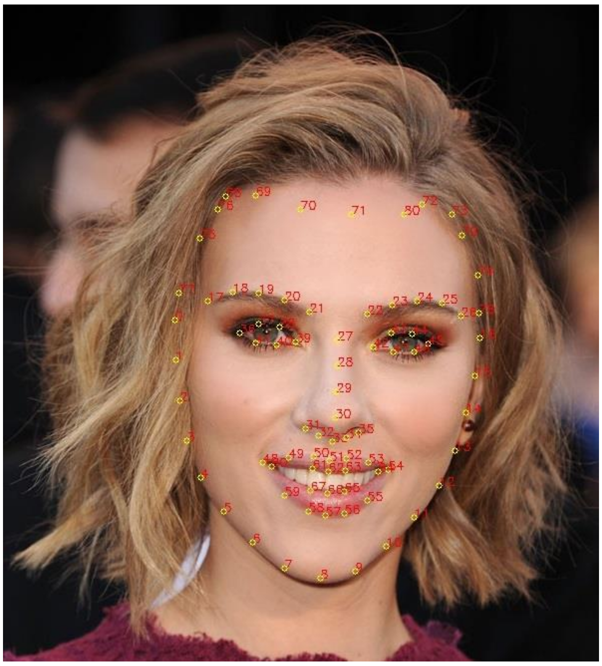


Fig. 6: Projection of points onto the face with the help of landmark points which were formed in the training images with the global face shapes

In the initial stage which is the training stage of PDM, various samples are taken where image holds number of points, and building the shape for each shape. Mean and covariance matrix is calculated by holding the points rigidly and in aligned fashion. When the PDM is fitted on the face the mean of the shape formed in the training stage is placed near the face. A search strategy named gray-scale search strategy is performed to deform the shape formed by points to fit the given face. In this process, the training set is the controlling the deformation according to the modelling of information.

The plus point of using PDM is it requires very less computational time for finding the face elements, as the elements are already founded when the global face is prepared in the training stage. Also, the occlusion is heavily reduced as the compactness of the global face formed with the points compensates for the occluded area [13]. The drawback of PDM is that the process of building a global face from the training set of images is inevitable and it is error prone. To add to the drawbacks the control points movement is restricted to only straight lines.

B. Low Level Analysis (LLA)

LLA works on taking the description out of image that are present in the image. The edges, color information and other things in the image are known as the descriptors. LLA descriptors can signify the edges of the face from the image and also different color variation. LLA is further divided into four parts motion, color, gray information and edges.

a) Motion

In motion face detection video and continuous image frames is required. Most of the valuable information can be fetched from the movement of objects. There are two main ways in detecting visual motion which are moving image contours and frame variance analysis. In frame variance analysis moving parts which include face are detected by gathering frame difference [14, 15]. Various face parts can also be detected in this manner [16, 17]. Commonly eye pairs are taken as the reference position while measuring face difference. Face contours gives us better result than frame variance.

Optical flow analysis is an advanced type of motion analysis. Apparent velocity of brightness is used in the optical flow analysis. The face motion is detected at the start, after that information is used to distinguish face. Motion analysis is very useful in real time as its robust and precise and also it uses less search space as it main focus is on movement. Motion analysis has a drawback of not detecting the eyes if the major axis is not perpendicular to eye center connecting the line [14]. If a face has beard or sunglasses then the motion analysis may not produce great results.

b) Color Information
In 2003 Kovac et al. [19], in 2010 Lik et al. [20], in 2013 Ban et al. [21], and many other have popularly used the skin color based face detection. As the color processing can be done faster face detection also become fast.

There are various color models used for face detection the ones which are commonly used are red, green, blue (RGB), hue, saturation, intensity (HSI), luminance, inphase, quadrature (YIQ). In RGB model, the basic three colors are mixed to build a color, there are some specific values of red, green, blue. After building the color the pixels which show maximum likeliness are deduced form the image. In HSI model has higher performance than other color models, it works by giving color clusters of face features a larger variance. HSI model can also detect various face elements like lips, eyes, eyebrows. In YIQ model RGB colors are bolstered. The conversion which is build shows difference between face and background which make face detection easy in natural environment.

As color processing is faster, this models are much faster than other facial feature processing. But color information is very sensitive to the luminance change and different cameras gives us different color values. The algorithm is inaccurate for side viewed faces.

c) Gray Information
In grayscale images, the pixels of an image are representing a certain amount of light. In Layman's terms every pixel has intensity information, which is described as gray information. There are two basic colors black and white and there are many shades of gray [23]. Mostly the edges of face and other features are darker than the surrounding part. By exploiting this dichotomy various facial features and face can be deduced from background.

As gray information works on two dimensional whereas color information works on three dimensional. Therefore, gray processing requires less complex processing time. But gray information is less efficient.

d) Edge
In 1972 Sakai et al. [24] used edge information in face detection. As edge detection was earliest techniques in computer vision. A sharp change in image brightness is considered as edge. Sakai et al. effectively employed analyzing of line drawings to detect various face features then detecting the full face. On the works of Sakai et al. Craw et al. [25] were able to develop a human head outline detection method which uses hierarchical framework.

The edges in an image are matched with the pre model for the accurate detection. There are many different filters and operators used to detect edges in an image.

Sobel Operator: Sobel operator works by computing the approximation of the gradient of the image intensity function. It is most commonly used operator.

Marr-Hildreth edge operator: The Marr-Hildreth edge operator works by combining the image with Laplacian of Gaussian function. Then, zero crossing is detected in the filtered results to obtain the edges.

Steerable filter: Steerable filter has three steps of working, which are edge detection, filter orientation of edge detection and tracking neighboring edges.

Minimum scanning is required for the edge based face detection. It is robust and cost effective. But it does not give the desired output in the noisy images and also it cannot examine edges in all scales.

C. Feature Analysis (FA)
The facial features functions as the detectors, by observing the individual features the face can be located. In LLA sometimes background is detected as the face, which is solved by high analysis of features. To find the actual face structure geometrical face analysis is employed rigorously, which is obtained ambiguously by low level analysis. There are two ways in which we can use geometric shape information first is the positioning of face features by the relative position of these features, second is flexible face structures. Feature analysis is divided into two parts feature searching and constellation analysis.

a) Feature Searching
Feature searching is conventional technique. In this the major notable face features are searched first, then other features are found by keeping the major features as the reference which were found first. Eyes

pair, face axis, facial outline and body below head are some major features used as reference.

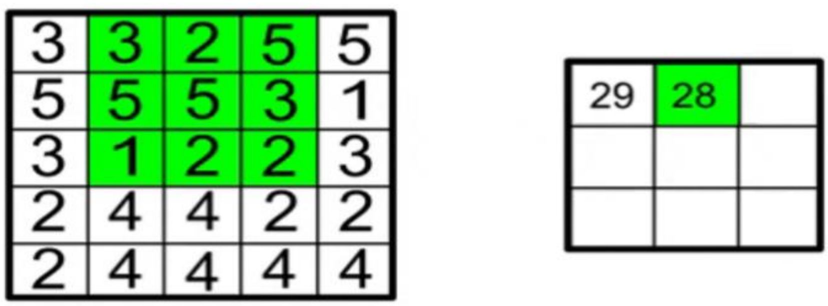
- **Face Outline:** De Silva et al. [30] presented the algorithm which starts by searching the forehead of the face. After finding the forehead it goes to find the eye pairs, which can be easily found as there is a sudden high variation of densities. The forehead and the eye pair are taken as the reference point and then the minor features are found. The algorithm which was presented showed 82% of accuracy. Facial outline-based algorithm cannot detect images with glass on the faces and hair on the forehead.
- **Eye Reference:**Jeng et al. [31] proposed to take the eye pair directly as a reference point in searching the face in the image. The algorithm which was proposed first searches for the possible eyes pair in images. The second step is similar to the outline, it searches for other facial features like mouth, nose, etc., keeping the eye pair as reference. This algorithm has 86% of rate in detecting of face.
- **Eye Movement:**Herpers et al. [32] proposed GAZE. The algorithm uses the eye movement in detecting the face. By using multi orientation Gaussian filter a rough representation is made. This rough representation is then used to locate the most important feature which is highly notable. Next the notability of the drawn out area is plunged and next location is drawn out in next iteration and the remaining features are deduced in the next iterations. With the 3 iteration Herpers et al. detected the moving eyes with 98% accuracy

Feature searching is further divided into two parts Viola-Jones algorithm and Gabor feature.

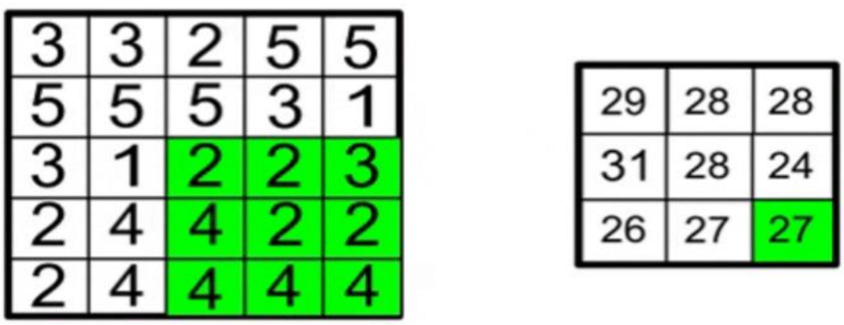
- **Viola-Jones Algorithm**

In 2001, Viola and jones came up with an object detection framework [33, 34]. The algorithm solved many challenges in real time face detection as it is faster and has high detection accuracy, and the computational complexity is also less. The algorithm works in two steps training and detection.

In detection part the image is converted to the gray scale and box search is performed throughout the image. After finding it in grayscale it finds location in colored image. Haar-like feature are used to search the face in the grayscale image. All the human faces have same features, the same is exploited in Haar-like features by making three types of Haar features for face, named edges, line and four sided features [35]. With the help of this value of each feature is calculated. An integral image is made out of them and compared quickly. This integral image is what makes the algorithm faster as it cuts the processing time by cutting the number of references of array. In the next stage boosting algorithm named Adaboost algorithm is used to select important and prominent features out of huge set to make the face detection efficient. In the last stage a cascading classification is assigned to throw out the non-face images in which prominent facial features are missing.



(a)



(b)

Fig. 7: Example of integral image: an 5x5 sized input image expressed with pixel values. Integral Image of the given image is been made by calculating the value form the original input image and its 3x3 pixels to one pixel in output image.

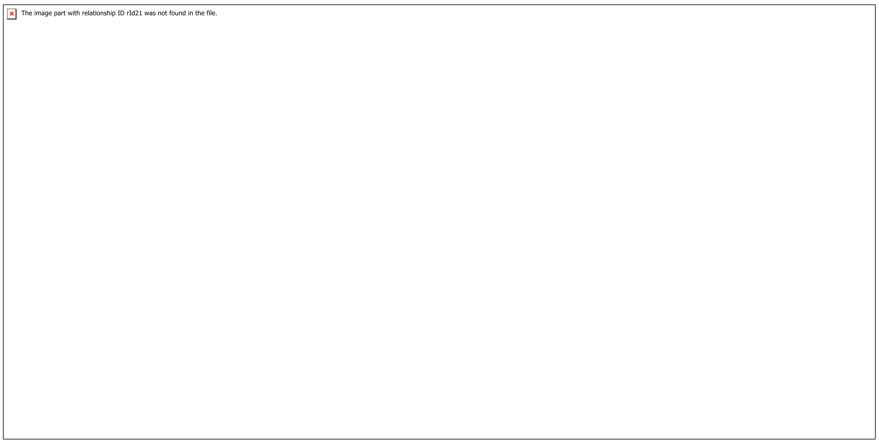


Fig. 8: Integral image array reference: (a) the sum within the dark shaped location is computed as 4+1−(2+3) (four array references); (b) the sum within the two dark shaped location is computed as 4+1−(2+3) and 8+5−(6+7), respectively, (eight array references); (c) the sum within the two adjacent dark shaped location is computed as 4+1−(2+3) +6+2−(4+5) and hence, 6+1−(3+5) (six array references)

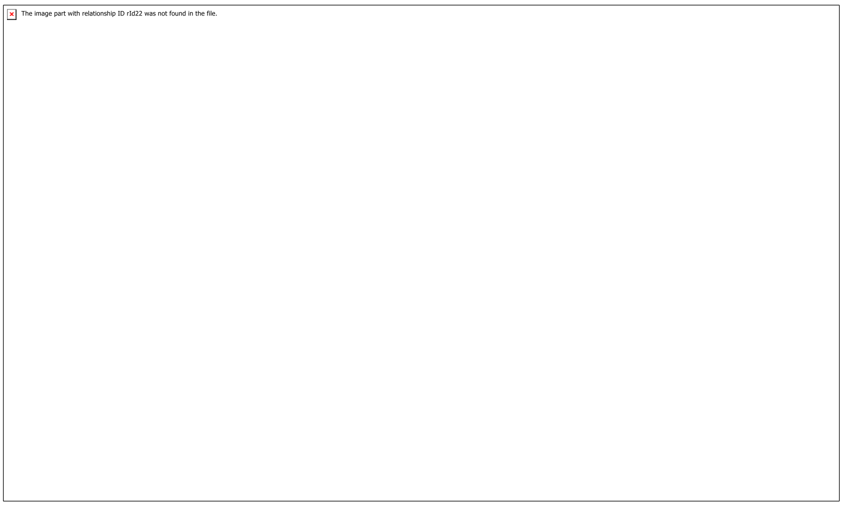


Fig. 9: Schematic diagram of detection cascade

The detection accuracy of the algorithm is very high at the time of release it was 95% a recent study by Jamal et al. reported it had accuracy of 97.88%. This algorithm is widely used as it has small computation time. Also, the false positive rate is very low. The Viola-Jones algorithm was 15% quicker than other algorithms at the time of release. Only frontal face can be detected through this algorithm. The training time is quite large for this algorithm and if we go forward with the limited number of classifiers the accuracy of result is decreased. The two further parts of Viola-Jones are Local Binary Pattern (LBP) and Adaboost.

- **Local Binary Pattern (LBP)**

Texture analysis is base of the LBP. In 1990 texture analysis model was proposed [37]. In 1994, Ojala et al. [38] described the LBP. It works robustly as texture descriptor and has a significant performance boost when working with the histogram of oriented object (HOG) [39]. There are many variants of the LBP, for instance, spatial temporal LBP (STLBP), ϵLBP, center symmetric local binary patterns (CS-LBP), spatial color binary patterns (SCBP), opponent color local binary pattern (OC-LBP), double local binary pattern (DLBP), uniform local binary pattern (ULBP), local SVD binary pattern (LSVD-BP), etc.

LBP works on nine pixels at an instance a 3X3 matrix and its main focus is central pixel. LBP starts by comparing the central pixel(cp) to its neighboring pixel(np) and it assigns 0 if np<cp and 1 if np>cp. In next step it converts eight binary np into one single byte which is known as LBP code or decimal number. This is done by multiplying the matrix component wise with an eight bit number representative matrix. The decimal number is then used in training process. Our interest is in edges, in the image the transition between 1 and 0 and vice versa is known as change in brightness. These changes are edge descriptors.

ISSN No:-2456-2165

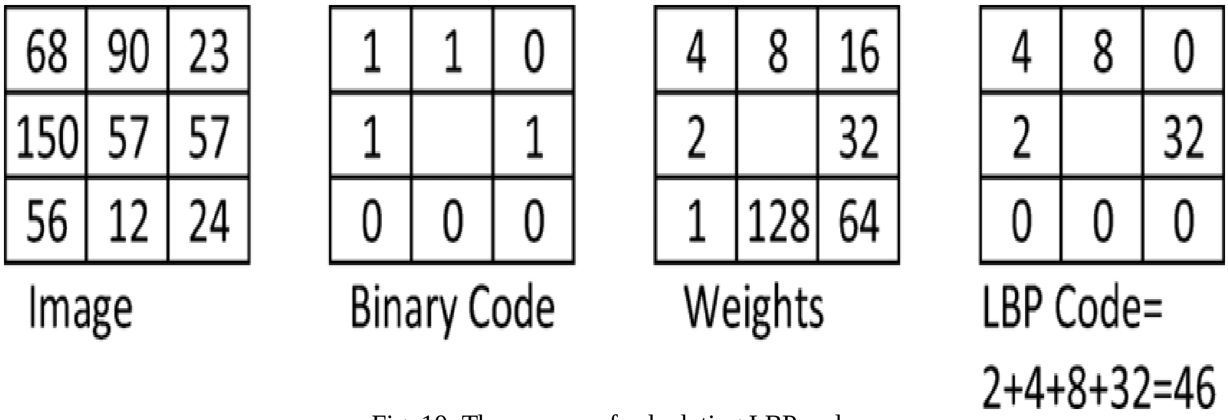


Fig. 10: The process of calculating LBP code

Computational simplicity and fast performance are the major plus points of the LBP. LBP can also detect a moving object by subtracting the background, and low false detection rate. The accuracy is same in offline and real time detection.

- AdaBoost

“Adaptive Boosting”, which is the commonly known as AdaBoost. In 1996, Freund and Schapire came up with this pragmatic boosting algorithm. The main focus of the AdaBoost is on classification and regression problem. The main function of the algorithm is to combine weak classifiers and make a strong classifier. More weight is added to the hard to define classifiers and less on already sorted out classifiers. This is how algorithm develops a better functioning classifier.

The degree of precision of AdaBoost is very high. The time for training is quite huge and it is also very sensitive to the noisy background images and null rejection is not supported.

- **Gabor Feature**

The Gabor filter is named after Dennis Gabor which is extensively used for the edge detection [41]. The filter is linear filter which is used for the texture analysis of the image [42]. As we know images have sharp change in contrast which is known as edges. This edges holds the prominent information and can help us locate the presence of face. Fourier transform is prominent in change analysis but does not work efficiently in dealing with sharp changes. Fourier transforms are sine waves oscillating in infinite time and space. But for detecting face we need something which works in finite time and space. Here is where Gabor wavelet is used which is rapidly plunging oscillation with mean zero [47]. For detection of the edges initially the wavelet is placed at random position if an edge is not found it is positioned at next random position [43].

With the help of Gabor feature impressive results were reported by dynamic link architecture (DLA), elastic bunch graph matching (EBGM), Gabor Fisher classifier (GFC), and AdaBoosted Gabor Fisher classifier (AGFC). The Gabor feature analysis has time complexity and image quality challenges.

- **Constellation Analysis**

As constellation is cluster of similar things. In constellation analysis, the facial feature group is formed and then the face is searched in the image [44, 45]. There is no rigidity in the algorithm. It is easy for the algorithm to detect face in the image which has lots of noisy backgrounds. Many algorithms failed to find the face in the complex background. But the constellation analysis was able to solve this problem.

Many face constellations have been proposed by many scientists. Here three of the most widely used are been discussed. Which is statistical theory by Brul et al. [46], probabilistic shape model by Yow and Cipolla et al. and graph matching. Statistical shape theory algorithm works efficiently in the rotation, translation and scale to a certain magnitude. It works smoothly with the features that are missing. Also, it has accuracy of 84%. Probabilistic shape model marks plunge in the detection of invalid features from noisy image and also have a 92% accuracy. It can handle minor variations in scale and orientation. Graph matching performs face detection in an automatic system and has higher detection accuracy.

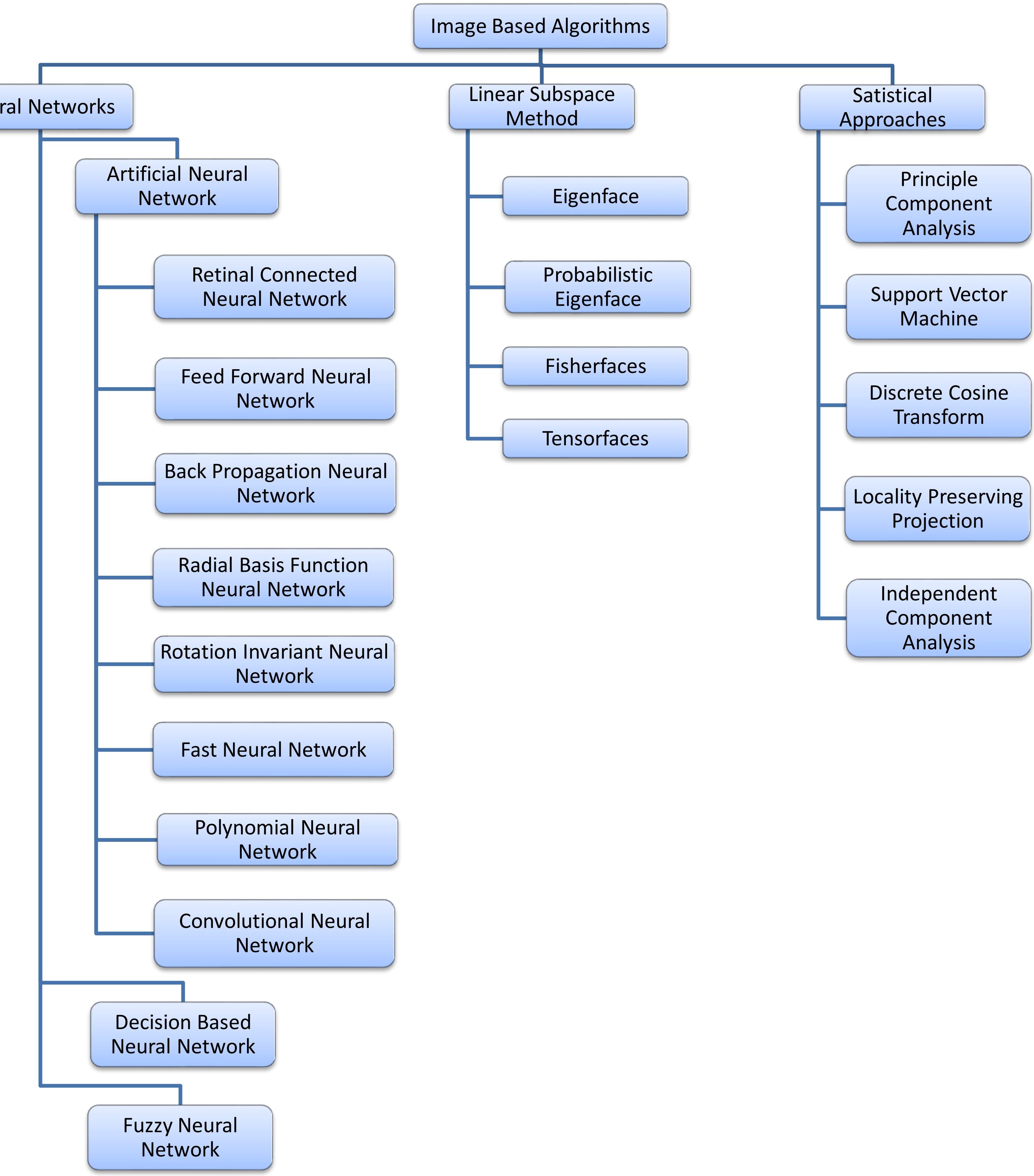


Fig. 11: Image based approaches for face detection which are divided into three types which are neural networks, linear subspace methods, statistical approaches this are further divided into many parts

ISSN No:-2456-2165

IV. IMAGE BASED ALGORITHMS

Face detection on a highly complex backgrounds gave a way for most of the image based approaches. Window based scanning is used for most of the image based approaches. Pixel by pixel classification is done on the window to find face and non-face. There is only small change in terms of the scanning window, step size, iteration number and sub-sampling rate which produces a more efficient approach. Image based approaches are non-conventional and recent techniques which have emerged in the face detection and are divided into three parts neural networks, linear subspace methods and statistical approaches.

A. Neural Network

The inspiration behind the neural network algorithms is human brain's neural network. Neural network take the data and they train themselves to recognize the pattern which in our case is face pattern. The networks then predict the output for the new set of similar faces. Neural networks are then further divided into three main parts artificial neural network (ANN), decision based neural network (DBNN) and fuzzy neural network (FNN).

a) Artificial Neural Network (ANN)
ANN is based on collection of the nodes like the human brain. The connected nodes are called artificial neurons. The learning of patterns in the data gives ANN to produce better results with more data. There is many ANN available but mostly used ANN are discussed below.

a. Retinal Connected Neural Network (RCNN)
In 1998, Rowley, Baluja and Kanade et al. [49] proposed a neural network based on the retinal connection of the human eyes. The ANN which was proposed was named retinal connected neural network (RCNN). RCNN takes a small part of the image to analyze whether the part contains the face or not. RCNN applies the filter on the image. The filter which is placed is based on the neural network. Temporary arbitrator is used to merge the output to a single node. The input image is searched, and different size of frame is applied to search for the face. The output node, with the help of an arbitrator, eliminates the overlapping features and combining face features gathering from the filtering.

The RCNN has some amount of the false positives as the output. The procedure is complex in terms of implementing and encounter frontal faces looking at the camera.

b. Feed Forward Neural Network (FFNN)
The feed forward neural network (FFNN) is also the simplest ANN and is a multi-layer perception. In 1958, Frank Posenblatt et al.[50] the neural network was upheld from perceptron's. Perceptron's are the methodologies of the brain to store and organize information. The face feature information is moved to the output node from the input node where it is done by hidden layers. In the training process the hidden layers assigns the weights to the face features. In the next stage which is detection stage the weights which are assigned are compared to report a result on the given image. FFNN handles the very large tasks, and accuracy is directly proportional to training sample.

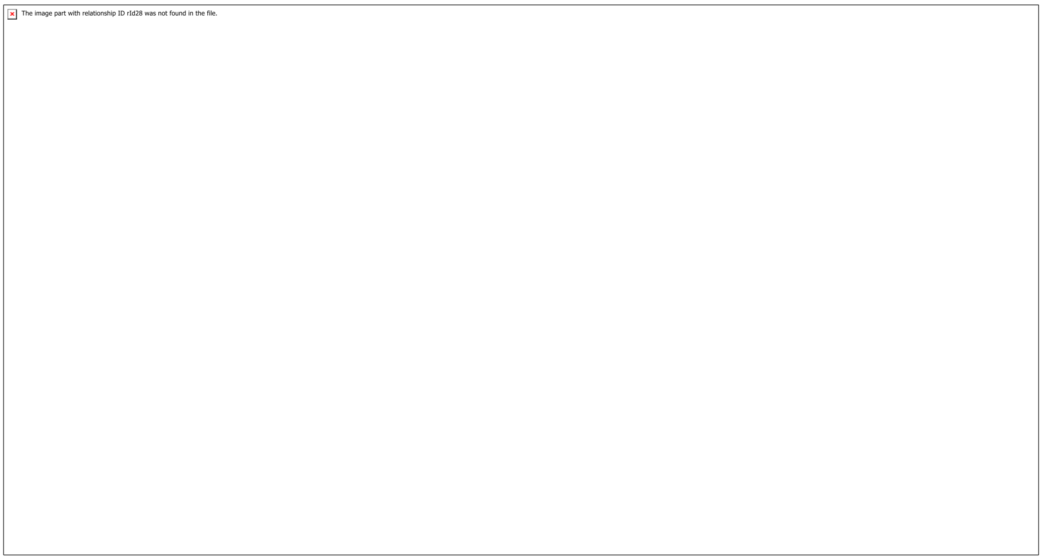


Fig. 12: Schematic diagram of three layer FFNN for face detection.

c. Back Propagation Neural Network (BPNN)
In 1963 Steinbunch et al. [51] proposed a learning matrix. In 1975, Seppo et al. developed the updated version of BPNN this version is also called reverse mode automatic differentiation (RMAD). BPNN gained attention when Rumelhart, Hinton and Williams published there paper in 1986 [52].

The base of BPNN is "learning by example". It calculates the error back to the input node from the output to make the adjustment in the weights in the hidden layer to increase the accuracy of the results. In the training stage many number of face features are used as the input. A larger amount of the weights is assigned to every features and the comparison is made with the input node with the errors. The error rate is checked and if it is higher the weight is decreased in the next attempt and comparison is made with the input node again. This makes sure that minimal error reporting weights are generated and employed in the next stage which is detection for the new images which are given. BPNN is easy to implement and program. It is also fast. BPNN gets stuck into local minima.

d. Radial Basis Function Neural Network (RBFNN)
In 1988, Broomhead and Lowe published radial basis function neural network (RBFNN) [53]. RBFNN share some similarities with BPNN. RBFNN also has input, hidden and output layers. But in RBFNN there is only one hidden layer in RBFNN and it remains to only one hidden layer which is named as feature vector. Gaussian potential function is used in RBFNN for mapping and neuron activating.

Computations in RBFNN are easy [54]. The first two stages of the training algorithm are used to train the network. In RBFNN approximation property are possessed. RBFNN has easy design and it does not affect much to input noise, it has online learning ability and has a good generalization, it has flexible control system. The inadequate number of neurons results in failure of the network and large number results in overlapping.

e. Rotation Invariant Neural Network (RINN)
In 1997, Rowley, Baluja and Kanade proposed rotation invariant neural network (RINN) [55]. Conventional algorithms are restricted to find in the frontal face only this problem is solved with RINN which detects the face at any angle of rotation. At the start, a network name router network holds every input network to find its orientation. After this the network works on preparing the window to detect the one or more detector networks. Higher classification performance is displayed in RINN [56]. Only small number of features can be learned by RINN and it performs very well only with a small number of training sets.

f. Fast Neural Network (FNN)
In 2002, Hazem El-Bakry presented a neural network which reduces the computation time of the neural network [57]. FNN is very fast in computing and detecting faces in the image. The image is divided into sub images in FNN after that sub images are searched for the faces. In FNN the challenge of sub image centering and normalization in the Fourier space is solved. Parallel processing is implemented in the network which makes the neural networks fast. But FNN is computationally expensive.

g. Polynomial Neural Network (PNN)
In 2003, Haung et al. presented the polynomial neural network (PNN) based face detection technique [58]. PNN was proposed by Ivakhnenko in 1971 [59]. The algorithm is also named as group methods of data handling (GMDH). GMDH has two inputs and one output where the output is quadratic combinations of two inputs.

A frame which can be slide over the image is introduced to detect the face. The detector labels the frames which have the face into it. Test image is then divided into variable scales to examine the numerous face shapes. But this dividing process is in short rescaling the input image into standard frame. Images are then arbitrated to avoid the overlapping due to re scaling and multiple faces in the detection region. Pixel intensities are fixed to compose a feature vector of 368 measurements, lighting conditions are ratified with an optimal plane, which causes minimal error. PNN can handle the image with the complex and dense background, the algorithm has high detection rate and low false positive rate. The algorithm has a problem in detecting faces in an image with large number of faces in it.

h. Convolutional Neural Network (CNN)
There are many ambiguities in who first proposed the CNN. CNN share some similarity to FFNN [60]. CNN has convolutional and some other layers along with hidden layers, an input is convoluted then passed to the next layer. The training stage trains the network and the best weighted values or filters are saved for detection. The network is trained with the usual back propagation gradient descent procedure. In the next stage which is detection stage the filters are scanned over the image to find patterns where patterns can be edges, shapes or the colors. CNN is very good at feature extracting also it has higher accuracy and high computational efficiency. CNN requires big dataset to train and is slow and holds high computational costs.

b) Decision Based Neural Network (DBNN)
In 1995, Kung et al. presented face detection algorithm based on the decision based neural network (DBNN) [62]. It can detect face in both still and moving images in still image static process is used and a temporal strategy is used for moving images(video). In the presented algorithm training stage, the pattern of the face is annealed to make the eye plane horizontal and produce structure where distance between the eyes is constant. A 16 x 16 pixels from images containing face or non-face is assembled by a Sobel edge map. This Sobel edge map is then used as the input to DBNN. When the whole image is considered only the founded sub image which contains the face is located as the location of the face. Effectiveness of DBNN in computation performance and time is high. The hierarchical structure of the DBNN provides a good understanding of structural richness. The processing speed and recognition accuracy is quite high. The detection rate is only higher only when the facial orientation is between -15 and 15 degrees.

c) Fuzzy Neural Network (FNN)
Fuzzy Neural Network (FNN) is the combination of the human like reasoning style of a fuzzy system with learning and connection establishing structure of the neural network which is known as neuro fuzzy hybridization which is known as fuzzy neural network (FNN). Many scientists used FNN for face detection among them some major and popular paper where are Rhee et al. in 2001 [63], Bhattacharjee et al. in 2009, Petrosino et al. in 2010, Pankaj et al. in 2011 and Chandrashekar et al. in 2017 [64].

Fuzzy membership degrees allocate the preprocessed 20 x 20 frame, whether or not it contains the face or not. This fuzzy degrees are then used as the input to the network. The training of the network is done by the fuzzy degrees along with error back propagation. After the training is done an evaluation is run over the network, which defines the degree of which a given window contains a face or not. If a frame is labeled to have a face, post processing is then carried out. FNN is highly accurate than the other neural networks. FNN requires fewer hidden neurons and can handle noisy backgrounds. But in FNN system user is required to describe the linguistic rules and it does not learn by example as in other networks.

B. Linear Subspace

Linear subspace is a vector space which is subset of the larger vector space. In layman’s terms smaller portion of a frame is called subspace. The linear subspace is further classified into four parts eigenfaces, probabilistic eigenspaces, fisher faces and tensor faces.

a) Eigen faces
Pentland and Turk [15, 66] first implemented the use of eigen faces in the face detection which was first proposed by Kirby and Sirovich [65]. Face detection is the process of finding a face in the given image. The system in which training of model is involved they are mostly train with the images which have high noise in them. Noise in image is created by changing in the pose, rotation of the face, changing the lighting conditions and many other minor factors. Even the images are highly noisy every image has its pattern and specially the images with the face always have the pattern. The pattern is created by the facial elements like nose, eyes, lips, eyebrows, etc. This elements or facial features are called as eigen faces. The eigen faces mostly are obtained by the principal component analysis (PCA). With the help of PCA, an eigen face corresponding to the elements are been prepared from the training data set.

When we combine many eigen faces in there right proportion the original face can be recreated. Here each eigen face represent a single facial feature. Due to the noise,it is not possible all the time that we can get all the eigen faces. But if the feature is there in the image, the proportion or share of that feature in the sum of the eigen faces is higher. The summation of all the weighted eigen faces gives us the full face from the image. The share of the feature is dependent on the weight assigned to it. In this way the face is detected and extracted from the given image with the help of weighted eigen faces and its summation.

Eigen faces method is very sensitive to the scale of the image if there is a minute change in the scaling the whole result may vary. As most of the other models the training and the learning time for the model is very time consuming. Eigen faces are very efficient when the face in the image is bigger or large in the dimensions. But we can do the data compression in the eigen faces by using low dimensional subspace representation and the knowledge of the geometry of the face is not required while using the eigen face model.

b) Probabilistic Eigenspaces
When there is constrained environment the eigen faces model showed great results in the face detection. But the eigen faces model works only on faces with rigid features. The model which was proposed by Pentland and Moghaddam solves this issue of rigid faces by making use of probabilistic similarity measure which is based on parametric estimate of probability density which is named as the probabilistic eigen spaces [67]. The model which was proposed has the higher capacity to handle the occlusion [68]. The process includes the probability distribution (PD) of the reconstruction error of each classes which were employed in the model to detect the face and the distribution of the class members of the eigen spaces is take into account in finding the face in this model.

c) Fisher faces
In 1997, Kriegman, Hespanha, Belhumeur came up with the idea to use the fisher face in the face detection [69]. To find the proper date representation is the major challenge in the face detection the

solution of this challenge involves principal component analysis by finding eigen faces. We can say that the sub space of the image which has most of the face can be considered as the eigen faces. But eigen faces cannot clearly define the similarity of the face subspace. To solve this anomality the subspaces which has the same classes at one spot and the subspaces which are very dissimilar are required. This process of achieving the defined tasks is known as discriminant analysis (DA). There are many flavors of the DA but the widely used and known is linear discriminant analysis (LDA). When the LDA is used to find the subspace in the given image which has the face this is called as Fisher face. The use of Fisher faces in the face detection helps a lot when the images where face detection is required have huge variation in the expressions of the face and the illumination. Not many models can detect the face with the glasses on the eyes but in this model the error rate of finding the face with the glasses is quite low than the eigen faces. The computation time for the fisher faces is also quite small in comparison to the eigen faces but fisher faces is highly dependent on the input data.

d) Tensor faces

In 2002, Terxopoulos and Vasilescu presented the idea of the tensor faces [70, 71]. A tensor is the generalization of the matrix in a multi-dimensional basis which makes tensor faces a multi linear approach. There are many dominating factors in the image like structure, the illumination and the view point. Multi linear algebra always solves this type of problem along with the band of images. This mathematical method uses the higher dimensional tensor to represent the image ensemble. When finding a face with the decomposition of the image an extension of singular value decomposition (SVD) known as the N-mode SVD is used. This N-mode SVD gives us the tensor faces from the given images. Many computer vision problems can be solved if tensor faces are employed as the unified frame work. Also, tensor faces perform good than the eigen faces.

C. Statistical Approaches

Statistical approaches are the mostly widely studied area for the face detection than other face detection algorithms or models. There are many statistical approaches available but the ones which are widely used and know are principal component analysis (PCA), support vector machine (SVM), discrete cosine transform (DCT), locality preserving projection (LPP) and independent component analysis (ICA).

a) Principal Component Analysis (PCA)

In 1901, Karl Pearson [72] proposed the idea of principal component analysis (PCA), further modification and advancement were done by Harlod in 1933 [73]. PCA changes it name on the basis of the application or field it is working on like proper orthogonal decomposition (POD) in mechanical engineering, Kerhunen-Loeve transform (KLT) in signal processing, and many more. The very first implementation of the PCA in face detection was done by Pentland and Turk.

Lots of data compression is done into some captures of real data in PCA. Orthogonal transformation mathematical procedure is been used in PCA in conversion of the set of values of possibly correlated M variables to possibly uncorrelated set of K variables, this are known are the principal components. Gathering or collection of component is done from the training data set and only first few components are taken in consideration other components are rejected. The components which we get are also known as the eigen faces. The process of detection of face is done by projecting a test image on the subspace spanned with the eigen spaces. Where there is constrained environment present PCA shows the higher efficiency and also it is faster than other statistical approaches. The further advanced statistical PCA (SPCA) has very high face detection rate and also it has simple computation. But PCA is dependent on the linear assumption and scale variants.

b) Support Vector Machine (SVM)

In 1963, Lerner and Vapnik introduced the modification of the generalized portrait algorithm [74]. In 1964, Vapnik and Chervonenkis made improvements and updated the generalized portrait algorithm. In 1992, Vapnik, Guyon and Boser proposed the next update of the SVM which is now widely used SVM flavor. Many researcher and scientists made use of SVM in face detection [75].

As it is training based model in the training stage the extraction of facial features is done using PCA or Histogram of Oriented Grading (HOG) or any other feature extraction technique. Using this data, the SVM model is trained to identify the face and non-face from the image. SVM draws a hyper plane between this classes. In the next stage which is detection stage the frame which is received by the model is compared to the trained images to put it into their respective classes of face or non-face.

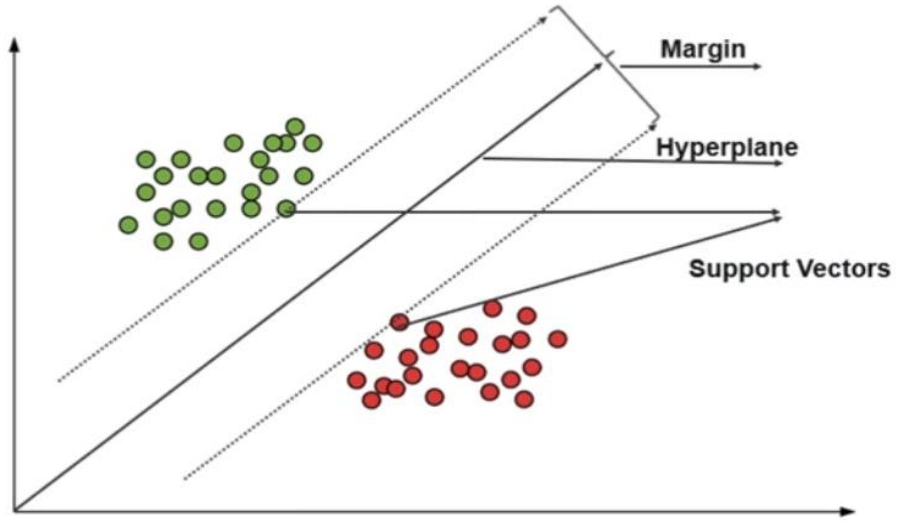


Fig. 13: Schematic representation of SVM. Here we may consider green dots as face data and red dots as non-face data. Hyperplane divides the both the classes with certain margin and the points which are near to the hyper plane are the support vectors

The risk of overfitting is very small in the SVM as the generalization is taking place in SVM models. When there is higher dimensional data present the SVM models are quite effective. Also, the SVM models are know to be memory efficient. When there is large dataset present the SVM models are not recommended and also the SVM models does not gives accurate results when the datasets have lots of noisy images in it.

c) Discrete Cosine Transform (DCT)
In 1974, Nasir Ahmed proposed the Discrete Cosine Transform (DCT) [76]. The main task of the DCT is to compress the image. In 2001, Ziad et al. made the use of DCT for face detection and recognition and many more researchers and scientists like Aman et al. in 2011 [77], Surya et al. in 2012 used DCT in face detection.

Here we have to manually enter the position or location of the eyes in the image. But this does not make the algorithm less worthy as this algorithm can be used with a localization system [77]. When the system gets the eye coordinates and the image it then performs the geometric and illumination normalization. Then the computation of the DCT of normalized face takes place. The subset of DCT coefficients which describe the faces is held as the feature vector. The subsets which are formed have the highest variance which are low to mid frequency. After that the final step to detect and recognize face this face vector is been compared with the database of the feature vectors of faces. Due to the normalization process DCT has higher detection rates and also it is computationally less expensive in comparison to the Karhunen-Loeve transform (KLT). DCT produces rich information of the face descriptors and gives us a simpler way to deal with the 3D facial distortions. To make some decisions it is required to ignore some high frequency components.

d) Locality Preserving Projection (LPP)
In 2004, Ha and Niyogi proposed locality preserving projection (LPP) [78]. It works as the substitute for PCA. The main purpose of the LPP is to store the locality structure which helps the LPP fast in recognizing the patterns as it explores the nearest patterns.

Similar to PCA in LPP subspace with a face is been searched which is mostly smaller in dimension than the actual image. The actual image is scaled and normalized to orientation. The normalization takes place in such a way that the two eyes are aligned into the same position. After this the cropping of image is done to 32 x 32 pixels. Then each image is then represented by the 1024 dimensional vector with 256 gray levels per pixel. He used the training set consisting of six images per individual. The use of training samples was done to learn projection and text images where projected on the reduced image space. LPP is highly recommended for practical applications as it is fast and preserves local structures and the error rate is quite less in comparison to PCA and LDA. The drawback of LPP is that it is sensitive to the noise and the outliers.

e) Independent Component Analysis (ICA)
In 1984, Bernard Ans and JeannyHerault proposed the initial idea of independent component analysis (ICA) [79]. Which gained popularity when Pierre Comon wrote and published his paper in 1994. Many scientists made the use of the ICA in the face detection some popular were Deniz et al. in 2001, Marian Barlett et al. in 2002, Zaid Aysseri in 2015 and many others.

In PCA we try to find the correlation by maximizing variation, in ICA we maximize independence of the features. ICA tries a linear transformation of the feature space to a new feature space in which the features are mutually independent, and the mutual information between the features of the original feature space and the new feature space is high as possible. ICA is lot better than PCA on many parameters. ICA is highly sensitive to high order data but in PCA only higher variance matters. The probabilistic model formed by ICA is far better than PCA and ICA model or algorithm is iterative in nature. But there are some problems in ICA like it cannot handle huge amount of data and there are display difficulties in ordering of the source vector.

V. CHALLENGES THAT ARE FACED DURING THE FACE DETECTION

There are some major challenges present in the face detection technology which have a significant impact on the accuracy and detection of face from the images or videos. The major challenges to tackle are face occlusion, variance in scale of images, variance in pose, complex and noisy background, very odd expression of the face, low resolution of images and most important too many faces in the image or video. Various algorithm tackles this challenges in their own ways to increase the accuracy and detection rate of the face from the given image. The algorithms which are available at present have variance in their performance and their own strength and weakness in face detection and recognition. Some algorithms are time consuming some have high computational costs whereas some have issue of over fitting, some algorithms require lots of space, some are very efficient computationally.

But in our case in which we want to build an efficient system to catch the miscreants which can be directly installed and be operational in already installed CCTV's all over the country. Keeping in mind the computational cost, space costs, hardware costs, and efficiency we need such an algorithm which is easy to implement which yields higher efficiency and gives us accurate results. At the current stage Fuzzy Neural Network (FNN) is the best option to work with as it is highly accurate, have low error rate, and high detection rate, also it is computationally less expensive than other models or algorithms and easy to program and implement on large scale and lots of scope for future modifications and updating the network. As Fuzzy neural network uses the fuzzy logic with the neural network, we need to code the linguistic rules this is the only thing which is holding the FNN backwards. But if we consider the advantages and our requirements and constraints in which this system will be implemented Fuzzy Neural Network is best option to move forward with.

VI. BASIC DESIGN AND WORKING OF THE SYSTEM TO CATCH THE MISCREANTS USING FACE DETECTION

Most of the hardware components is already installed like CCTV's cameras in most of the cities and places. The work which we have to do is that to build a Fuzzy Neural Network. After building the Network we can implement it on the system of CCTV's which are already in place and we need some extra hardware so that the network can run smoothly without any space problem. The face detection problem is solved by the neural network but the major concern is that how to recognize the face and match whose face it is. If we take India as example, it has a population of approximately 130 crores and every minute there are hundreds of people be captured in every CCTV camera and at present in most of cities there are almost two to three cameras in every street. So, we have a huge problem of recognizing the face as the database is quite large and till the time, the system recognizes the face the miscreant is long gone from that location.

In the developed countries like United States of America they have pyramid like analogy to store the data of the miscreants. Like the top 10 most wanted miscreants are at the top of the pyramid, then top 100 most wanted miscreants are below them, them top 1000 most wanted miscreants and the pyramid goes on to every citizen of the country and the citizens who are not involved in any criminal activities are at the lowest level of the pyramid. This are not the exact details but an analogy given to understand the current system for matching the data of miscreants. This kind of system takes a day or two to go to the bottom of the pyramid. Keeping in mind the conditions of India we may require a week or more to identify and we cannot lose that amount of time. We need a system which is more efficient in matching with the database.

If we want to minimize the time, we can do the parallel processing but the computational costs will rise up significantly. I propose that we can use the area wise bifurcation and use of binary search and randomizing it. Like for instance take three areas in consideration like Sangli, Kolhapur, Satara every region will have its own database of the miscreants which will be arranged in the sorted manner and the data which will be stored is the face converted to the values using the FNN. Why to store the values because we can then directly match the output of the FNN to the database. Which will save our time in converting the values to face then comparing. Why to keep the database area wise because then instead of searching in the national database first we will search the database of that region which will save us lots of time for searching which we would usually waste in unnecessary searching of whole data base.

Once the face is recognized the system then can send an alert to the concerned authorities which can proceed accordingly. In the system there will be a functionality in which the law enforcement agencies can add the miscreant to the database which will save the time of them sending the image to the Neural network manager or database administrator and then they adding to database. The flow of the basic system is given below:

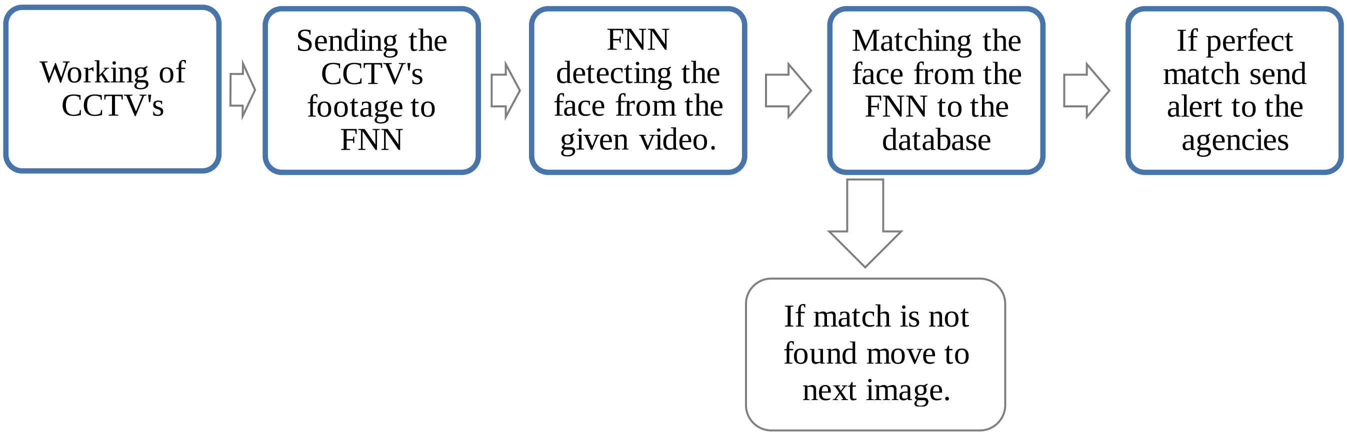


Fig. 14: Basic flow of the system

VII. RESULTS

With every day advancement in the technology there is always scope for modifications in every system present which makes them efficient and accurate with every modification. This system can also be modified and new features can be added to it such as Aadhaar card linked with the data set and or other government issued documents linked to the system and database which can help us keep track of the citizens also and if any mishap happens, we can easily identify the person and act accordingly. We can build the system in such a way that once a miscreant is been recognized by the system then it will automatically track the movements of the miscreants without any need of human intervention which will save us lots of man power. The linguistic rules set can be made more complex to give more accurate results. The system can be fine-tuned to save the computational costs and man power.

Also, newer algorithm or model can be implemented which may be environment friendly as every face detection algorithm requires lots of computation process lots of energy is used to implement and run this algorithms so newer algorithms can be developed or invented which may use less energy and perform better than the existing algorithms or models. A combination of the algorithms can be implemented which may yield the better results. And keeping in the mind the recent condition of Covid-19 face masks and shields are becoming the new normal so the system should also be able to tackle this challenge.

VIII. CONCLUSION

In this paper various face detection algorithms were discussed its drawbacks and advantages were also mentioned in detail. The major challenges in face detection technology are face occlusion, computational cost, accuracy, computational and processing time, and so on. Various algorithms and models tackle this challenges in there own way. Not every algorithm or model is suitable for every type of applications. For specific kind of applications or specific kind of data some algorithms work best with those data and some does not. So, while using the algorithm the detailed study of application of algorithm the data it will be working on is important according to that we must select the algorithm, This paper gives the details and working and its working condition of most of the face detection algorithms. The basic system design is also given which can be used for catching miscreant. The system not only can be used to catch miscreants but it can be used to take the attendance of the students hassle free without requiring any human power. Some of the face detection algorithms mentioned above are a century old. While using the algorithms we must keep in mind the process of algorithm how the algorithm is implemented and then use the algorithm. A detailed study of algorithm and application is important before using it.

- **Conflict of Interest:** Neither the author Jeet Kansagara nor the supervisor Yogesh Sharma have any conflict of interest in publishing this research paper. This research paper is written for the sole purpose on the advancement of the face detection techniques.

REFERENCES

[1.] Kass, M.; Witkin, A.; Terzopoulos, D. Snakes: Active contour models. Int. J. Comput. Vis. 1988, 1, 321–331. [Google Scholar]

[2.] Nikolaidis, A.; Pitas, I. Facial feature extraction and pose determination. Pattern Recognit. 2000, 33, 1783–1791. [Google Scholar]

[3.] Gum, S.R.; Nixon, M.S. Active Contours for Head Boundary Extraction by Global and Local Energy Minimisation. IEE Colloq. Image Process. Biomed. Meas. 1994, 6, 1. [Google Scholar]

[4.] Huang, C.L.; Chen, C.W. Human facial feature extraction for face interpretation and recognition. Pattern Recognit. 1992, 25, 1435–1444. [Google Scholar]

[5.] Yuille, A.L.; Hallinan, P.W.; Cohen, D.S. Feature extraction from faces using deformable templates. Int. J. Comput. Vision 1992, 8, 99–111. [Google Scholar]

[6.] Felzenszwalb, P. Representation and detection of deformable shapes. IEEE Trans. Pattern Anal. Mach. Intell. 2005, 27, 208–220. [Google Scholar]

[7.] Zhu, L.; Chen, Y.; Yuille, A. Learning a Hierarchical Deformable Template for Rapid Deformable Object Parsing. IEEE Trans. Pattern Anal. Mach. Intell. 2010, 32, 1029–1043. [Google Scholar]

[8.] Nishida, K.; Enami, N.; Ariki, Y. Detection of facial parts via deformable part model using part annotation. In Proceedings of the 2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), Hong Kong, China, 16–19 December 2015. [Google Scholar]

[9.] Yanagisawa, H.; Ishii, D.; Watanabe, H. Face detection for comic images with deformable part model. In Proceedings of the The Fourth IIEEJ International Workshop on Image Electronics and Visual Computing, Samui, Thailand, 7–10 October 2014. [Google Scholar]

[10.] Hasan, M. K., Ahsan, M. S., Abdullah-Al-Mamun, Newaz, S. H. S., & Lee, G. M. (2021). Human Face Detection Techniques: A Comprehensive Review and Future Research Directions. Electronics, 10(19), 2354. MDPI AG. Retrieved from http://dx.doi.org/10.3390/electronics10192354

[11.] Cootes, T.F.; Taylor, C.J. Active shape models—'Smart snakes'. In BMVC92; Springer: London, UK, 1992; pp. 266–275. [Google Scholar]

[12.] Lanitis, A.; Cootes, T.; Taylor, C. Automatic tracking, coding and reconstruction of human faces, using flexible appearance models. Electron. Lett. 1994, 30, 1587–1588. [Google Scholar]

[13.] Lanitis, A.; Hill, A.; Cootes, T.F.; Taylor, C. Locating Facial Features Using Genetic Algorithms. In Proceedings of the 27th International Conference on Digital Signal Processing, Limassol, Cyprus, 26–28 June 1995; pp. 520–525. [Google Scholar]

[14.] Van Beek, P.J.; Reinders, M.J.; Sankur, B.; van der Lubbe, J.C. Semantic segmentation of videophone image sequences. In Proceedings of the Visual Communications and Image Processing'92, International Society for Optics and Photonics, Boston, MA, USA, 1 November 1992; Volume 1818, pp. 1182–1193. [Google Scholar]

[15.] Turk, M.; Pentland, A. Eigenfaces for recognition. J. Cogn. Neurosci. 1991, 3, 71–86. [Google Scholar]

[16.] Luthon, F.; Lievin, M. Lip motion automatic detection. In Proceedings of the Scandinavian Conference on Image Analysis, Lappeenranta, Finland, 9–11 June 1997. [Google Scholar]

[17.] Crowley, J.L.; Berard, F. Multi-modal tracking of faces for video communications. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Juan, PR, USA, 17–19 June 1997; pp. 640–645. [Google Scholar]

[18.] McKenna, S.; Gong, S.; Liddell, H. Real-time tracking for an integrated face recognition system. In Proceedings of the Second European Workshop on Parallel Modelling of Neural Operators, Faro, Portugal, 7–9 November 1995; Volume 11. [Google Scholar]

[19.] Kovac, J.; Peer, P.; Solina, F. Human skin color clustering for face detection. In Proceedings of the IEEE Region 8 EUROCON 2003, Computer as a Tool, Ljubljana, Slovenia, 22–24 September 2003; Volume 2, pp. 144–148. [Google Scholar]

[20.] Liu, Q.; Peng, G.z. A robust skin color based face detection algorithm. In Proceedings of the 2010 2nd International Asia Conference on Informatics in Control, Automation and Robotics (CAR 2010), Wuhan, China, 6–7 March 2010; Volume 2, pp. 525–528. [Google Scholar]

[21.] Ban, Y.; Kim, S.K.; Kim, S.; Toh, K.A.; Lee, S. Face detection based on skin color likelihood. Pattern Recognit. 2014, 47, 1573–1585. [Google Scholar]

[22.] Zangana, H.M. A New Skin Color Based Face Detection Algorithm by Combining Three Color Model Algorithms. IOSR J. Comput. Eng. 2015, 17, 06–125. [Google Scholar]

[23.] Graf, H.P.; Cosatto, E.; Gibbon, D.; Kocheisen, M.; Petajan, E. Multi-modal system for locating heads and faces. In Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, Killington, VT, USA, 14–16 October 1996; pp. 88–93. [Google Scholar]

[24.] Sakai, T. Computer analysis and classification of photographs of human faces. In Proceedings of the First USA—Japan Computer Conference, Tokyo, Japan, 3–5 October 1972; pp. 55–62. [Google Scholar]

[25.] Craw, I.; Ellis, H.; Lishman, J.R. Automatic extraction of face-feature. Pattern Recognit. Lett. 1987, 5, 183–187. [Google Scholar]

[26.] Sikarwar, R.; Agrawal, A.; Kushwah, R.S. An Edge Based Efficient Method of Face Detection and Feature Extraction. In Proceedings of the 2015 Fifth International Conference on Communication Systems and Network Technologies, Gwalior, India, 4–6 April 2015; pp. 1147–1151. [Google Scholar]

[27.] Suzuki, Y.; Shibata, T. Multiple-clue face detection algorithm using edge-based feature vectors. In Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, QC, Canada, 17–21 May 2004; Volume 5. [Google Scholar]

[28.] Froba, B.; Kublbeck, C. Robust face detection at video frame rate based on edge orientation features. In Proceedings of the Fifth IEEE International

Conference on Automatic Face Gesture Recognition, Washington, DC, USA, 21 May 2002; pp. 342–347. [Google Scholar]

[29.] Suzuki, Y.; Shibata, T. An edge-based face detection algorithm robust against illumination, focus, and scale variations. In Proceedings of the 2004 12th European Signal Processing Conference, Vienna, Austria, 6–10 September 2004; pp. 2279–2282. [Google Scholar]

[30.] De Silva, L.; Aizawa, K.; Hatori, M. Detection and Tracking of Facial Features by Using Edge Pixel Counting and Deformable Circular Template Matching. IEICE Trans. Inf. Syst. 1995, 78, 1195–1207. [Google Scholar]

[31.] Jeng, S.H.; Liao, H.Y.M.; Han, C.C.; Chern, M.Y.; Liu, Y.T. Facial feature detection using geometrical face model: An efficient approach. Pattern Recognit. 1998, 31, 273–282. [Google Scholar]

[32.] Herpers, R.; Kattner, H.; Rodax, H.; Sommer, G. GAZE: An attentive processing strategy to detect and analyze the prominent facial regions. In Proceedings of the International Workshop on Automatic Face and Gesture Recognition, Zurich, Switzerland, 26–28 June 1995; pp. 214–220. [Google Scholar]

[33.] Viola, P.; Jones, M. Robust Real-time Object Detection. Int. J. Comput. Vis. 2001, 4, 34–47. [Google Scholar]

[34.] Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), Kauai, HI, USA, 8–14 December 2001; Volume 1, p. I. [Google Scholar]

[35.] Li, J.; Zhang, Y. Learning surf cascade for fast and accurate object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; pp. 3468–3475. [Google Scholar]

[36.] Viola, P.; Jones, M.J. Robust real-time face detection. Int. J. Comput. Vision 2004, 57, 137–154. [Google Scholar]

[37.] He, D.C.; Wang, L. Texture unit, texture spectrum, and texture analysis. IEEE Trans. Geosci. Remote Sens. 1990, 28, 509–512. [Google Scholar]

[38.] Ojala, T.; Pietikainen, M.; Harwood, D. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In Proceedings of the 12th International Conference on Pattern Recognition, Jerusalem, Israel, 9–13 October 1994; Volume 1, pp. 582–585. [Google Scholar]

[39.] Wang, X.; Han, T.X.; Yan, S. An HOG-LBP human detector with partial occlusion handling. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 29 September–2 October 2009; pp. 32–39. [Google Scholar]

[40.] Freund, Y.; Schapire, R.E. Experiments with a new boosting algorithm. ICML 1996, 96, 148–156. [Google Scholar]

[41.] Gabor, D. Theory of communication. Part 1: The analysis of information. J. Inst. Electr.-Eng.-Part III Radio Commun. Eng. 1946, 93, 429–441. [Google Scholar]

[42.] Sharif, M.; Khalid, A.; Raza, M.; Mohsin, S. Face Recognition using Gabor Filters. J. Appl. Comput. Sci. Math. 2011, 5, 53–57. [Google Scholar]

[43.] Rahman, M.T.; Bhuiyan, M.A. Face recognition using gabor filters. In Proceedings of the 2008 11th International Conference on Computer and Information Technology, Khulna, Bangladesh, 24–27 December 2008; pp. 510–515. [Google Scholar]

[44.] Burl, M.C.; Perona, P. Recognition of planar object classes. In Proceedings of the CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 18–20 June 1996; pp. 223–230. [Google Scholar]

[45.] Huang, W.; Mariani, R. Face detection and precise eyes location. In Proceedings of the 15th International Conference on Pattern Recognition (ICPR-2000), Barcelona, Spain, 3–7 September 2000; Volume 4, pp. 722–727. [Google Scholar]

[46.] Burl, M.C.; Leung, T.K.; Perona, P. Face localization via shape statistics. In Proceedings of the Internatational Workshop on Automatic Face and Gesture Recognition. Citeseer, Zurich, Switzerland, 26–28 June 1995; pp. 154–159. [Google Scholar]

[47.] Bhuiyan, A.A.; Liu, C.H. On face recognition using gabor filters. World Acad. Sci. Eng. Technol. 2007, 28, 195–200. [Google Scholar]

[48.] Yow, K.C.; Cipolla, R. Feature-based human face detection. Image Vision Comput. 1997, 15, 713–735. [Google Scholar]

[49.] Rowley, H.A.; Baluja, S.; Kanade, T. Neural network-based face detection. IEEE Trans. Pattern Anal. Mach. Intell. 1998, 20, 23–38. [Google Scholar]

[50.] Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. Psychol. Rev. 1958, 65, 386. [Google Scholar]

[51.] Steinbuch, K.; Piske, U.A. Learning matrices and their applications. IEEE Trans. Electron. Comput. 1963, EC-12, 846–862. [Google Scholar]

[52.] Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. Nature 1986, 323, 533–536. [Google Scholar]

[53.] Broomhead, D.S.; Lowe, D. Multivariable Functional Interpolation and Adaptive Networks. Complex Syst. 1988, 2, 321–355. [Google Scholar]

[54.] Orr, M.J. Introduction to Radial Basis Function Networks; Centre for Cognitive Science University of Edinburgh: Edinburgh, UK, 1996. [Google Scholar]

[55.] Rowley, H.A.; Baluja, S.; Kanade, T. Rotation invariant neural network-based face detection. In Proceedings of the 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No. 98CB36231), Santa Barbara, CA, USA, 25 June 1998; pp. 38–44. [Google Scholar]

[56.] Marcos, D.; Volpi, M.; Tuia, D. Learning rotation invariant convolutional filters for texture classification. In Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 2012–2017. [Google Scholar]

[57.] El-Bakry, H.M. Face detection using neural networks and image decomposition. In Proceedings of the 2002 International Joint Conference on Neural Networks,

IJCNN'02 (Cat. No. 02CH37290), Honolulu, HI, USA, 12–17 May 2002; Volume 1, pp. 1045–1050. [Google Scholar]
[58.] Huang, L.L.; Shimizu, A.; Hagihara, Y.; Kobatake, H. Face detection from cluttered images using a polynomial neural network. Neurocomputing 2003, 51, 197–211. [Google Scholar]
[59.] Ivakhnenko, A.G. The group method of data of handling; a rival of the method of stochastic approximation. Sov. Autom. Control 1968, 13, 43–55. [Google Scholar]
[60.] Le Cun, Y.; Jackel, L.D.; Boser, B.; Denker, J.S.; Graf, H.P.; Guyon, I.; Henderson, D.; Howard, R.E.; Hubbard, W. Handwritten digit recognition: Applications of neural network chips and automatic learning. IEEE Commun. Mag. 1989, 27, 41–46. [Google Scholar]
[61.] Matsugu, M.; Mori, K.; Mitari, Y.; Kaneda, Y. Subject independent facial expression recognition with robust face detection using a convolutional neural network. Neural Networks 2003, 16, 555–559. [Google Scholar]
[62.] Kung, S.Y.; Lin, S.H.; Fang, M. A neural network approach to face/palm recognition. In Proceedings of the 1995 IEEE Workshop on Neural Networks for Signal Processing, Cambridge, MA, USA, 31 August–2 September 1995; pp. 323–332. [Google Scholar]
[63.] Rhee, F.C.H.; Lee, C. Region based fuzzy neural networks for face detection. In Proceedings of the Joint 9th IFSA World Congress and 20th NAFIPS International Conference (Cat. No. 01TH8569), Vancouver, BC, Canada, 25–28 July 2001; Volume 2, pp. 1156–1160. [Google Scholar]
[64.] Chandrasekhar, T. Face recognition using fuzzy neural network. Int. J. Future Revolut. Comput. Sci. Commun. Eng. 2017, 3, 101–105. [Google Scholar]
[65.] Sirovich, L.; Kirby, M. Low-dimensional procedure for the characterization of human faces. Josa A 1987, 4, 519–524. [Google Scholar]
[66.] Turk, M.; Pentland, A. Face recognition using eigenfaces. In Proceedings of the 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Maui, HI, USA, 3–6 June 1991; pp. 586–587. [Google Scholar]
[67.] Moghaddam, B.; Pentland, A. Probabilistic visual learning for object detection. In Proceedings of the IEEE International Conference on Computer Vision, Cambridge, MA, USA, 20–23 June 1995; pp. 786–793. [Google Scholar]
[68.] Midgley, J. Probabilistic Eigenspace Object Recognition in the Presence of Occlusion; National Library of Canada: Ottawa, ON, Canada, 2001. [Google Scholar]
[69.] Belhumeur, P.N.; Hespanha, J.P.; Kriegman, D.J. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. 1997, 19, 711–720. [Google Scholar]
[70.] Vasilescu, M.A.O.; Terzopoulos, D. Multilinear subspace analysis of image ensembles. In Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Madison, WI, USA, 18–20 June 2003; Volume 2. [Google Scholar]
[71.] Vasilescu, M.A.O.; Terzopoulos, D. Multilinear Analysis of Image Ensembles: Tensorfaces; European Conference on Computer Vision; Springer: Copenhagen, Denmark, 2002; pp. 447–460. [Google Scholar]
[72.] Pearson, K. LIII. On lines and planes of closest fit to systems of points in space. Lond. Edinb. Dublin Philos. Mag. J. Sci. 1901, 2, 559–572. [Google Scholar]
[73.] Hotelling, H. Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 1933, 24, 417. [Google Scholar]
[74.] Vapnik, V. Pattern recognition using generalized portrait method. Autom. Remote Control 1963, 24, 774–780. [Google Scholar]
[75.] Guo, G.; Li, S.Z.; Chan, K.L. Support vector machines for face recognition. Image Vision Comput. 2001, 19, 631–638. [Google Scholar]
[76.] Ahmed, N.; Natarajan, T.; Rao, K.R. Discrete cosine transform. IEEE Trans. Comput. 1974, 100, 90–93. [Google Scholar]
[77.] Chadha, A.R.; Vaidya, P.P.; Roja, M.M. Face recognition using discrete cosine transform for global and local features. In Proceedings of the 2011 International Conference On Recent Advancements in Electrical, Electronics And Control Engineering, Sivakasi, India, 15–17 December 2011; pp. 502–505. [Google Scholar]
[78.] He, X.; Niyogi, P. Locality preserving projections. Adv. Neural Inf. Process. Syst. 2004, 16, 153–160. [Google Scholar]
[79.] Hérault, J.; Ans, B. Réseau de neurones à synapses modifiables: Décodage de messages sensoriels composites par apprentissage non supervisé et permanent. C. R. Séances L'Académie Sci. Série Sci. Vie 1984, 299, 525–528. [Google Scholar]