A Seminar Report

On

FACE DETECTION TECHNOLOGY

Submitted in partial fulfilment of award of

B.Tech. (IT-A) V Semester III Year

BY

DEEPALI

2016736

Department of Computer Science

Faculty of Mathematics & Computing

Banasthali Vidyapith

Rajasthan

(July – December 2022)

ABSTRACT

With the marvelous increase in video and image database there is an incredible need of automatic understanding and examination of information by the intelligent systems as manually it is getting to be plainly distant. Face plays a major role in social intercourse for conveying identity and feelings of a person. Human Beings have not tremendous ability to identify different faces than machines. So, automatic Face Detection System plays an important role in Face Recognition, Facial Expression Recognition, Head-Pose Estimation, Human-Computer Interaction etc. Face Detection is a computer technology that determines the location and size of a human face in a digital image. Face Detection has been a standout amongst topics in the computer vision literature. Here knowledge about Face Recognition, Face Detection along with their differences, method and algorithm of Face Detection are also given along with challenges, real life applications, advantages and disadvantages.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION …………………………………………………………………..4

1.1: WHAT IS FACE RECOGNITION? ………………………………………………4

1.2: FACE DETECTION TECHNOLOGY …………………………………………4-5

1.3: FACE DETECTION VERSUS FACE RECOGNITION …………………………5

CHAPTER 2: HOW FACE DETECTION WORKS? ………………………………………...6

CHAPTER 3: METHODS OF FACE DETECTION …………………………………………7

CHAPTER 4: THE VIOLA JONES ALGORITHM …………………………………………8

4.1: WORKING ……………………………………………………………..…8

4.1.1: DETECTION ……………………………………………………8-10

4.1.2: TRAINING ……………………………………………….....10-12

CHAPTER 5: APPLICATIONS ……………………………………………………..…...13-14

CHAPTER 6: CHALLENGES ………………………………………………………………….15

CHAPTER 7: PROS AND CONS …………………………………………………………16

CHAPTER 8: CONCLUSION ……………………………………………………………..17

REFERENCES ……………………………………………………………………………18

CHAPTER 1

INTRODUCTION

1.1: WHAT IS FACE RECOGNITION?

Face Recognition is a way of identifying or confirming an individual’s identity using their face. Facial Recognition Systems can be used to identify people in photos, videos, or in real-time.
Facial Recognition is a category of biometric security. Other forms of biometric software include Voice Recognition, Fingerprint Recognition, and Eye Retina or Iris Recognition. The technology is mostly used for security and law enforcement, though there is increasing interest in other areas of use.

The various steps involved in Face Recognition System are –

Step 1: Face Detection

Step 2: Face Analysis

Step 3: Converting the image to data

Step 4: Finding a match

1.2: FACE DETECTION TECHNOLOGY

Face Detection also called Facial Detection is an Artificial Intelligence (AI) based computer technology used to find and identify human faces in digital images. Face Detection Technology can be applied to various fields including security, biometrics, law enforcement, entertainment and personal safety to provide surveillance and tracking of people in real time.
It now plays an important role as the first step in many key applications including Face Tracking, Face Analysis and Face Recognition. Face Detection has a significant effect on how sequential operations will perform in the application.
In Face Analysis, Face Detection helps identify which parts of an image or video should be focused on to determine age, gender and emotions using facial expressions. In a Facial Recognition System which maps an individual's facial features mathematically and stores the data as a face print, Face Detection data is required for the algorithms that discern which parts of an image or video are needed to generate a

face print. Once identified, the new face print can be compared with stored face prints to determine if there is a match.

1.3: FACE DETECTION VERSUS FACE RECOGNITION

Face Detection simply means that the Face Detection System can identify that there is a human face present in an image or video – it cannot identify that person.

Face Detection is a component of Facial Recognition Systems – the first stage of Facial Recognition is detecting the presence of a human face in the first place. Face Detection can also be used in cameras to help with auto-focus on digital cameras and phones, where a small box appears around the faces of people detected within the image, allowing the camera to prioritize focus on those faces. Facial Recognition thus is able to recognize and identify the detected face based on a match stored in a database.

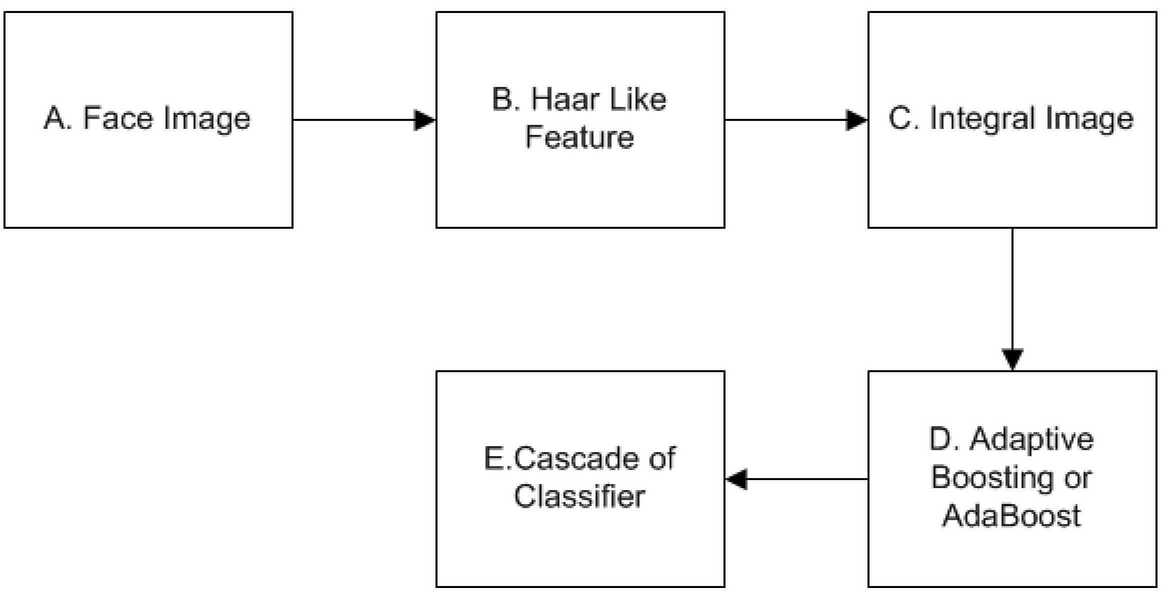
Fig. Face Detection Versus Face Recognition

CHAPTER 2

HOW FACE DETECTION WORKS?

Face Detection applications use algorithms and ML to find human faces within larger images, which often incorporate other non-face objects such as landscapes, buildings and other human body parts like feet or hands. Face Detection algorithms typically start by searching for human eyes, one of the easiest features to detect. The algorithm might then attempt to detect eyebrows, the mouth, nose, nostrils and the iris. Once the algorithm concludes that it has found a facial region, it applies additional tests to confirm that it has, in fact, detected a face.

To help ensure accuracy, the algorithms need to be trained on large **data** sets incorporating hundreds of thousands of positive and negative images. The training improves the algorithms' ability to determine whether there are faces in an image and where they are.



CHAPTER 3

METHODS OF FACE DETECTION

Yan, Kriegman, and Ahuja presented a classification for Face Detection methods. These methods divided into four categories. These categories are as follows –

1. Knowledge-Based:-

The knowledge-based method depends on the set of rules, and it is based on human knowledge to detect the faces. Example- A face must have a nose, eyes, and mouth within certain distances and positions with each other. The big problem with these methods is the difficulty in building an appropriate set of rules. There could be many false positive if the rules were too general or too detailed. This approach alone is insufficient and unable to find many faces in multiple images.

2. Feature-Based:-

The feature-based method is to locate faces by extracting structural features of the face. It is first trained as a classifier and then used to differentiate between facial and non-facial regions. The idea is to overcome the limits of our instinctive knowledge of faces. This approach divided into several steps and even photos with many faces they report a success rate of 94%.

3. Template Matching:-

Template matching method uses pre-defined or parameterized face templates to locate or detect the faces by the correlation between the templates and input images. Example- a human face can be divided into eyes, face contour, nose, and mouth. Also, a face model can be built by edges just by using edge detection method. This approach is simple to implement, but it is inadequate for Face Detection. However, deformable templates have been proposed to deal with these problems.

4. Appearance-Based:-

The appearance-based method depends on a set of delegate training face images to find out face models. The appearance-based approach is better than other ways of performance. In general appearance-based method rely on techniques from statistical analysis and machine learning to find the relevant characteristics of face images. This method also used in feature extraction for Face Recognition.

CHAPTER 4

THE VIOLA JONES ALGORITHM

It was developed in 2001 by Paul Viola and Michael Jones, the Viola-Jones algorithm is an object-recognition framework that allows the detection of image features in real-time. Despite being an outdated framework, Viola-Jones is quite powerful and its application has proven to be exceptionally notable in real-time face detection.

4.1: WORKING

There are 2 stages in the Viola-Jones Algorithm:

1) Training
2) Detection

4.1.1: TRAINING

a) TRAINING CLASSIFIERS

It is training the machine to identify the features. It involves feeding information, and subsequently training it to learn from the information to predict. So ultimately, the algorithm is setting a minimum threshold to determine whether something can be classified as a feature or not.

The algorithm shrinks the image to 24 x 24 and looks for the trained features within the image. It needs a lot of facial image data to be able to see features in the different and varying forms. That's why there is need to supply lots of facial image data to the algorithm so that it can be trained. Viola and Jones fed their algorithm 4,960 images (each manually labeled). For some images, the mirror image of a particular image can also be fed, which would be brand new information for a computer.

There is also a need to supply the algorithm non-facial images so it can differentiate between the two classes. Viola and Jones supplied their algorithm 9,544 non-facial images. Within these, some images may look similar to features in a face, but the algorithm will understand which features are more likely to be on a face and which features would obviously not be on a face.

b) ADAPTIVE BOOSTING (ADABOOST)

The algorithm learns from the images we supply it and is able to determine the false positives and true negatives in the data, allowing it to be more accurate. We would get a highly accurate model once we have looked at all possible positions and combinations of those features. Training can be super extensive because of all the different possibilities and combinations to check for every single frame or image.

For example – Consider an equation for the features that determines the success rate (as seen in the image), with f1, f2 and f3 as the features and a1, a2, a3 as the respective weights of the features. Each of the features is known as a **weak classifier.** The left side of the equation F(x) is called a **strong classifier.** Since one weak classifier may not be as good, a strong classifier is got by combining two or three weak classifiers. By continuously adding, it gets stronger and stronger. This is called an **ensemble.** Here the most important features are required, but the question is how to find the most important or the 'best' features? That's where Adaptive Boosting comes into play.

For example – Consider 10 pictures, 5 facial images and 5 non-facial.
The model gives us 3 out of 5 true positives and 2 out of 5 true negatives. It predicts correctly for these images but there are also some errors: 3 false positives and 2 false negatives. So it didn't find the feature on these 2 images they are actually faces, whereas it found the feature in 3 non-facial images.
In the next step, adaptive boosting uses another feature, the one to best complement the current strongest feature. So it doesn't look for the second-best feature, but one that complements the current best feature. So it increases the importance of the images that it got wrong as false negatives, and finds the next best feature that would fit these images, in a way, increasing the weight of these images on the overall algorithm. So, as new features are added, it comes down to one image at the end that would be given a higher weight. Once the algorithm is optimized and is able to calculate all positives and negatives correctly, move on to the next step: cascading.

c) CASCADING

Cascading is another sort of "hack" to boost the speed and accuracy of our model. Start by taking a sub-window and within this sub-window, take the most important or best feature and check if it is present in the image within the sub-window. If it is not in the sub-window, then the sub-window is discarded. If it is present, look at the second feature in the sub-window. If it isn't present, then simply reject the sub-window. Go on for the number of features that are there, and reject the sub-windows without the feature. Evaluations may take split seconds but since it has to be done for each feature, it could take a lot of time. Cascading speeds up this process a lot, and the machine is able to deliver results much faster.

4.1.2: DETECTION

Viola-Jones was designed for frontal faces, so it is able to detect frontal the best rather than faces looking sideways, upwards or downwards. Before detecting a face, the image is converted into grayscale, since it is easier to work with and there's lesser data to process. The Viola-Jones algorithm first detects the face on the grayscale image and then finds the location on the colored image.

Viola-Jones outlines and searches for a face within the box. It is essentially searching for the haar-like features, explained later. The box moves a step to the right after going through every tile in the picture. In this case, a large box size and large steps are taken for demonstration, but in general, the box size and step size can vary according to need.

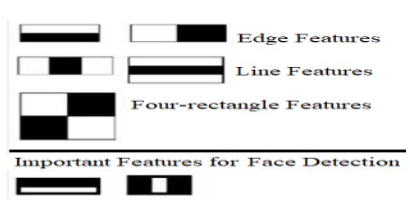
With smaller steps, a number of boxes detect face-like features (Haar-like features) and the data of all of those boxes put together helps the algorithm determine where the face is.

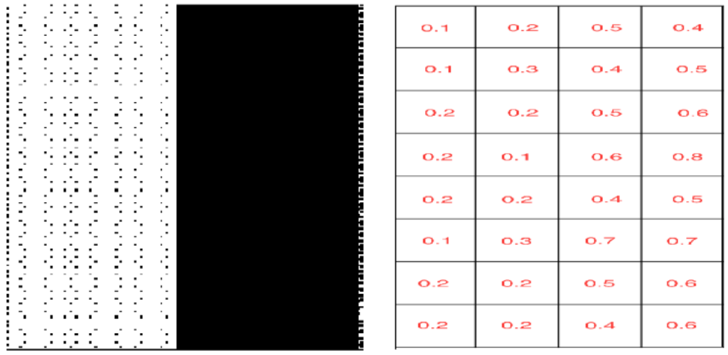
a) HAAR-LIKE FEATURES

Haar-like features are named after Alfred Haar, a Hungarian mathematician in the 19th century who developed the concept of Haar wavelets (kind of like the ancestor of haar-like features). The features below show a box with a light side and a dark side, which is how the machine determines what the feature is. Sometimes one side will be lighter than the other, as in an edge of an eyebrow. Sometimes the middle portion may be shinier than the surrounding boxes, which can be interpreted as a nose.

There are 3 types of Haar-like features that Viola and Jones identified in their research:

1) Edge features
2) Line-features
3) Four-sided features





These features help the machine understand what the image is. Imagine what the edge of a table would look like on a black and white image. One side will be lighter than the other, creating that edge like black and white feature as shown in the picture above.

In the two important features for Face Detection, the horizontal and the vertical features describe what eyebrows and the nose, respectively, look like to the machine. Additionally, when the images are inspected, each feature has a value of its own. It's quite easy to calculate: Subtract White area from the Black area. For example, look at the image below.

Imagine the Haar-like feature was converted into a grid. Each square represents a pixel. Let us consider a 4 x 8 grid, but in reality, there would be many more pixels and thus a much larger grid for a certain feature. The numbers in the boxes represent the darkness of the features. The higher it is, the darker the pixel. Thus, the numbers are higher on the right side than on the left side. Now if you add up the numbers on the two left-sided (white) columns, and subtract it from the sum of the right-sided (black) columns, it will give the value of the particular feature.

So in this case, the value of the feature is →
(0.5 + 0.4 + 0.5 +0.6 + 0.4 + 0.7 + 0.5 + 0.4 + 0.4 + 0.5 + 0.6 + 0.8 + 0.5 + 0.7 + 0.6 + 0.6) - (0.1 + 0.1 + 0.2 + 0. 2 + 0.2 + 0.1 + 0.2 + 0.2 + 0.2 + 0.3 + 0.2 + 0.1 + 0.2 + 0.3 + 0.2 + 0.2)

B-W=8.7-3

= 5.7

b) INTEGRAL IMAGE

So in the last section, we calculated the value of the feature. In reality, these calculations can be very intensive since the number of pixels would be much greater within a large feature.

The integral image plays its part by allowing to perform these intensive calculations quickly to understand whether a feature of a number of features fit the criteria.

Now imagine the one highlighted in red is the grid for a certain feature whose value is needed to be calculated. Normally, just the need is to add up the boxes, but since that can be computationally intensive, an integral image is created.

To calculate the value of a single box in the integral image, take the sum of all the boxes to its left. The image below shows an example –

8	5	6	4	4	2	3	4	4	3
7	5	6	5	5	6	6	5	5	4
5	5	7	8	9	7	7	6	5	3
2	4	3	3	4	5	6	6	5	4
2	3	4	5	5	6	6	7	5	5

Regular Image

8	13	19	23	27	29	32	36	40	43
15	25	37	46	55	63	72	81	90	97
20	35	54	71	79	94	110	125	139	149
22	41	63	83	95	115	137	158	177	191
24	46	72	97	114	140	168	202	226	245

Integral Image

The green box in the integral image is calculated as the sum of the highlighted area in the regular image. Do this for every box, a sequence will be obtained going through the grid and it may look something like the image below.

Let's look at the value taken earlier –

Just the need is to look at the 4 corners of the feature, and add the purples, subtract the greens. → **168–114 + 79–110 = 23**

So why do we use the integral image?

Because Haar-like features are actually rectangular, and the integral image process allows to find a feature within an image very easily and as already known the sum value of a particular square and to find the difference between two rectangles in the regular image, just the need is to subtract two squares in the integral image. So even if there is 1000 x 1000 pixels in the grid, the integral image method makes the calculations much less intensive and can save a lot of time for any facial detection model.

CHAPTER 5

APPLICATIONS

1. Gender Classification

Applications are built to detect gender information with face detection methods.

2. Crowd Surveillance

Face detection is used to detect and analyze crowds in frequented public or private areas.

3. Biometric Attendance

Facial recognition is used to detect the attendance of humans. It is often combined with biometric detection for access management.

4. Emotional Inference

Emotion recognition applications are still in the works; when they are fully developed, AI might be able to “read” nonverbal cues, gestures, body movements, and facial expressions to convey a person’s feelings.

5. Lip Reading

The detection, modeling, and tracking of lips during videos can be used to generate automatic subtitles. Such an application can be found on YouTube, where some videos have the option to turn on subtitles, even if the creator has not provided any.

6. Facial Feature Extraction

Facial features like nose, eyes, mouth, skin color and more can be extracted from images.

7. Facial Motion Capture

With applications such as Snapchat, people’s faces can be altered in real-time with fun filters. Facial detection makes this possible, as its algorithms tell the applications that there is a face that can be traced.

8. Facial Recognition

Facial Recognition adds increased security to nearly every global industry. It seeks to identify a person and then authenticate their identity – but for a person's face print to be analyzed via facial recognition, the facial area to be assessed is determined by face detection.

9. Photography

Facial recognition can be used to "tag" people's faces in photos across social media platforms, and facial detection forms the foundation of this application. Furthermore, facial detection technology can be used alongside tracking to focus on a person's face while the photo is being taken.

10. Marketing

Facial surveillance can help stores determine customers that have visited a few times and offer them perks or discounts – thus fostering increased customer loyalty.

CHAPTER 6

CHALLENGES

1. Unusual expressions

Human faces in an image may show unexpected or odd facial expressions.

2. Face occlusion

Faces may be partially hidden by objects such as glasses, scarves, hands, hairs, hats, and other objects, which impact the detection rate.

3. Illuminations

Some image parts may have very high or low illumination or shadows.

4. Complex background

A high number of objects in a scene reduces the accuracy and rate of detection.

5**. Many faces in one image**

An image with a high number of human faces is very challenging for an accurate detection rate.

6. Low resolution

Low-resolution images or image noise impacts the detection rate negatively.

7. Skin types

Detecting faces of different face colors is challenging for detection and requires a wider diversity of training images.

8. Distance

If the distance to the camera is too high, the object size (face size) may be too small.

9. Orientation

The face orientation and angle toward the camera impact the rate of face detection.

CHAPTER 7

PROS AND CONS

PROS

1. Better security
2. Easy to integrate
3. Automated identification
4. Helps find missing people
5. Protects businesses against theft
6. Improves medical treatment
7. Strengthens security measures
8. Makes shopping more efficient

CONS

1. Huge memory requirements
2. Detection can be vulnerable
3. Potential privacy issue
4. Imposes on personal freedom and rights
5. Misuse causing fraud and other crimes
6. Technology is still new and can be manipulated
7. Errors can implicate innocent people

CHAPTER 8

CONCLUSION

Thus, Face Detection is a system that gives its major contribution in many different fields including Facial Recognition Systems. The various methods of Face Detection can be implied along with the algorithm which can help in doing Face Analysis and make Face Recognition Systems. It has many advantages in many fields like education, marketing, gender classification and many more. Today, Face Detection is of paramount importance due to the wide range of benefits and applications they provide.

REFERENCES

https://towardsdatascience.com/the-intuition-behind-facial-detection-the-viola-jones-algorithm-29d9106b6999

https://www.researchgate.net/publication/326667118_Face_Detection_Techniques_A_Review

https://www.techtarget.com/searchenterpriseai/definition/face-detection#:~:text=The%20methods%20used%20in%20face,up%20with%20well%2Ddefined%20rules.

https://viso.ai/deep-learning/face-detection-overview/