IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.13446

# Optimizing Viola-Jones for Advanced Face Detection:A Comprehensive Study

**U SHIV KUMAR[1], SK.MOHAMMAD KAIF [2], K SIVA NAGASAI AJAY[3], SMRITILEKHA DAS[4], T GNANASRI ADILAKSHMI[5]**

Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Green fields, Vaddeswaram,A.P[1-5]

**Abstract**: In recent years, significant progress has been made towards human-like machine comprehension. However, it's crucial to be aware that teaching a computer to think like a person is a very challenging undertaking. Automation and how we interact with computers are becoming more fun as computer vision technology advances. The Viola Jones approach, which makes it easier to recognise persons in photos, is discussed in this research. No matter the illumination, we have the computer set up to perform this task automatically. In experiments, they studied the Viola-Jones Cascade Object Detector. In order to identify particular facial features, this detector employs a variety of filters and traits.

**Keywords:** SURF, Feature Detection, Face Recognition, Face Detection, Viola-Jones algorithm.

## I. INTRODUCTION

In our interactions with one another, faces play a huge role. Today, we utilise faces to identify people, whether it's for security purposes, to apprehend criminals, or for commercial purposes. In each of these situations, we start by locating the face.

We'll discuss various approaches to finding faces in the sections that follow. We'll go into further detail about the Viola-Jones approach and go through utilising neural networks to detect faces. Part V will describe the new approach we're advocating, and Part VI will present the findings of our studies.

## II. LITERATURE REVIEW

The Facial Action Coding System (FACS) was a generally accepted technique that was first proposed by Ekman and Friesen [7] in 1977 . By identifying the facial muscles that create these expressions, FACS is utilised to comprehend face expressions. To aid in the analysis of face expressions, it divides facial behaviours into 46 Action Units (AUs) It becomes a little more complex when it comes to the "line-based caricature of facial expression for the line edge map (LEM) descriptor" phase. In essence, it is drawing a streamlined image of a facial expression using lines to reflect how the face appears when exhibiting an emotion.

By calculating something known as the "line segment Hausdorff distance," they may determine how comparable this line-based representation of expressions is. This distance reveals the degree of separation between two sets of lines. In order to better understand how effectively the system recognises facial expressions, the researchers compare the line-based representation of expressions to the line-based description of test faces. The finding that overall, girls had a detection rate 7.8% higher than men was the most intriguing.Tensor-based techniques were utilised by Lajevardi and Wu to represent colour pictures. Using this method, they were able to identify face expressions with an accuracy of 68.8% even when the images had varying degrees of clarity and colour. Specifically designed neural networks with just one hidden layer were employed by Ma and Khorasani [14] to reduce the size of the photos. They discovered that employing a block size of 12 and a certain quantity of concealed units produced an accuracy rate that was above 93.75%.

The MPEG-4 standard was utilised by researchers to define how various facial features move. In order to develop a system for identifying facial emotions, Aleksic and Katsaggelos [2] concentrated on certain characteristics like the brows and outer lips. Based on the data they had gathered, they calculated the likelihood of various facial expressions using a mathematical model known as a hidden Markov model. They were successful in accurately identifying up to 93.66% of the test phrases using this method.Based on the data they had gathered, they calculated the likelihood of various facial expressions using a mathematical model known as a hidden Markov model. They were successful in accurately identifying up to 93.66% of the test phrases using this method. In order to make it simpler to identify faces in hazy or

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁐ Peer-reviewed & Refereed journal ⁐ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.13446

poor-quality images, Huang and He developed a method. They achieved this by utilising a method known as canonical correlation analysis (CCA) to look for comparable patterns in both high-quality and low-quality photos. They employed a technique known as radial basis functions (RBF), which performs somewhat similarly to a "nearest neighbour" strategy, to increase the accuracy of the detection process.

This helps increase accuracy by searching a database for faces that are similar to those in a low-quality input image. They evaluated the effectiveness of their technique using several face datasets, and they were able to successfully detect 93.0% of faces in the FERET database, 95.0% of faces in the ORL database, and a fair percentage of faces in the UMIST database.Turk and Pentland's Eigenface approach, one of the most widely used methods for identifying faces, was originally presented.. This technique for identifying facial expressions was developed by Murthy and Jadon, who then put it to the test on the Cohn-Kanade (CK) [9]. Using the projected gradient approach, Zhi, Flierl, Ruan, and Kleijn created a method known as "graph-preserving sparse non-negative matrix factorization"(GSNMF).The Japanese female facial expression (JAFFE) database and the Facial Expression database are two examples of facial expression datasets from which significant elements may be extracted using a technique known as the Facial Action Coding System (FACS). In essence, it is a method for extracting and using relevant information from massive collections of face expressions. In terms of identifying various face problems, they had good success. They were able to identify covered eyes with 93.3% accuracy, covered noses with 94.0% accuracy, covered mouths with 90.1% accuracy, and fast, unexpected facial expressions with 96.6% accuracy. Mase and Pentland developed a method for tracking the motion of face muscles using a method known as "dense optical flow." By integrating it with a face model that employs a unique technique for continuously estimating and enhancing the tracking, they were able to make it even better. They were able to reach a 98.0% accuracy rate because to these enhancements. Simply put, they created a technique to precisely measure facial muscle movements in a person.

The "enhanced ratio template algorithm" was the technique McOwan employed to identify faces from the front. The researchers employed a "multichannel gradient model" to track face movements. They put their detection system to the test using the "support vector machine" (SVM) classifier, and they discovered that it had an 81.82% detection rate. Simply said, McOwan employed a variety of methods to recognise frontal visages and track facial motions, and the system they developed successfully detected them 81.82% of the time. In their study, they analysed facial expressions using a technique known as "elastic graph matching (EGM)". They also applied a novel technique known as "2-class kernel discriminate analysis" to raise the efficiency of their facial emotion recognition system.With a technique based on Gabor features and utilising elastic graph matching, they were able to detect targets with an accuracy of 90.5%. With an improved detection accuracy of 91.8%, the technique based on normalised morphological characteristics and elastic graph matching outperformed the others. In the beginning, their system could only detect emotions in images taken in standard visible light. They widened it, though, to incorporate both visible and infrared pictures. This implies that their system can now recognise facial expressions of emotion in additional contexts.They also discovered that the mathematical tools known as Gabor wavelets were useful since they are effective at identifying particular regions in a picture and portraying things in a way that consumes less data. The Enhanced Fisher linear discriminant Model (EFM), developed by Liu and Wechsler, is a face identification technique based on Gabor wavelets. To boost the functionality of their system, they included more Gabor features. They used 200 distinct datasets or scenarios to evaluate their strategy. Using a technique developed by Zhang and Tjondronegoro [11], pictures are divided into smaller sections (called patches), from which characteristics are extracted using Gabor wavelets. Utilising assessments of their similarity or difference, they compared these input picture patches with patches from previously trained images. They employed a classification strategy that combined four distinct mathematical methods known as "kernels" with support vector machines to decide which matches to consider. In essence, they created a system that can identify things in photos by disassembling them into their component parts and comparing those parts with ones it has already seen.Two databases were used to evaluate their techniques. They attained an accurate detection rate of 92.93% for the JAFFE database and a rate of 94.8% for the CK database.They presented two novel methodologies in their research [21]. The first one concentrated on directly identifying dynamic facial expressions, whereas the second one relied on identifying certain facial action components. They used a specific technique to sort the outcomes.

## III. PROPOSED ALGORITHM

Paul Viola and Michael Jones [12] invented the Viola-Jones algorithm in 2001, which is a ground-breaking technique for identifying numerous face traits. Real-time face and trait recognition is made possible by this breakthrough.The Matlab toolbox, which is a component of the computer vision toolkit, was used to create this technique. To do this, it employs a method known as a "cascade object detector." You can use a Region of Interest (ROI), change the Scale Factor, select Minimum and Maximum Sizes, establish a Merge Threshold, or apply a Classification Model in this system. The cascade classification model must first be trained before it can be used.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⩳ Peer-reviewed & Refereed journal ⩳ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.13446

You can use a Region of Interest (ROI), change the Scale Factor, select Minimum and Maximum Sizes, establish a Merge Threshold, or apply a Classification Model in this system.[2] To find items of varying sizes, we utilise the "Scale Factor" option with a value that is only a little bit higher than 1.0001. The "Min Size" and "Max Size" options determine how finely the detection resolution scales.

The three primary methods used by the Viola-Jones approach are as follows:

1.In order to build an integral image, feature extraction is done using Haar-like features.

2.A machine-learning approach known as Ada Boost,which creates powerful classifiers using boosting techniques since the classifiers are fairly complicated, is used to recognise faces.

3.The Cascade classifier, which produces a classifier with many filters, is the third strategy. It effectively mixes various characteristics.

We are able to distinguish between non-facial and facial regions with effectiveness when we employ cascading powerful classifiers. The nose, upper chest, lips, eyes, and pupils are just a few of the facial traits that this object detector can reliably recognise. This study explores the many feature vectors that may be employed with the method to identify faces[4]. The cascade object detection framework manages the face detection. The benefit of the suggested approach is that it can recognise faces in any lighting situation. The primary component of this procedure, which makes use of classifier using a Haar cascade. These Haar features can reveal whether a feature is present in the provided picture. Every function returns a single value that is determined by the difference in the sum of the sum of the pixels in the black region and the pixels in the white region. We think of Haar features as a rectangular zone for quick facial feature identification. The following figure lists some Haar filters that can be used to identify faces:

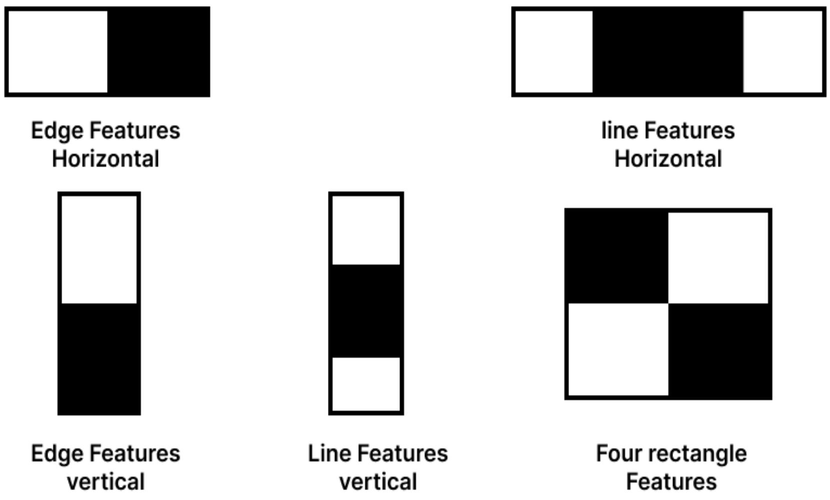


Fig:2.1

it is calculated as in (1)*S(x) = Sum of black region / Sum of White region (1)*

When we enter an image into the system, the first thing we do is begin at the top left corner and scan the entire thing. As we proceed from the top left to the bottom right, we're looking for any indications of face characteristics. To locate faces in the picture, we continually use this scanning procedure.

We employ a technique known as an integral image to quickly determine rectangular characteristics. Using just four numbers from a rectangle's four corners, this technique enables us to[5] calculate the total of the pixels in any given region. The value at a certain location (a, b) in the integral picture is determined by adding up all the pixels above and to the left of that location.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⅔ Peer-reviewed & Refereed journal ⅔ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.13446

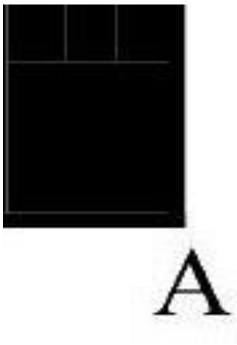


Fig:2.2

The Haar-like characteristics may be seen in this illustration. "f0(x) = A" is written in an area that looks to be dark. Let's now examine a different filter in the graphic.

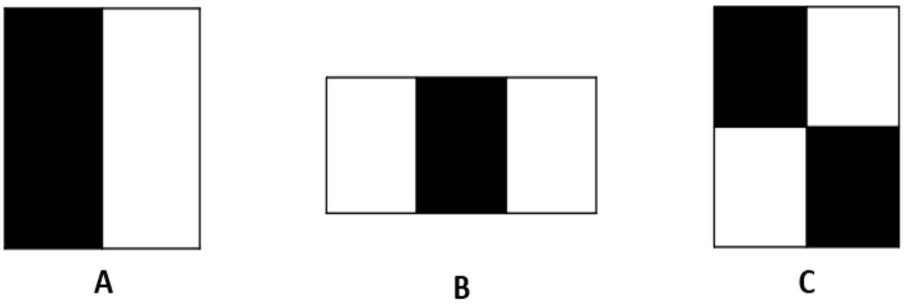


Fig:2.3

A 24x24 window is used as the first evaluation space for image characteristics in the Viola-Jones method. There are possibly 160,000 distinct characteristics that may be calculated inside this time, making it impossible to manage manually. So, they effectively choose the most pertinent characteristics from this huge pool using a machine learning approach known as Ada Boost.

By integrating these characteristics using a linear combination, Ada Boost produces a powerful classifier. In plainer language, it determines which elements are crucial and then cleverly blends them to aid in accurate facial recognition.

$$F(x) = \alpha 1F1(x) + \alpha 2F(x) + \dots\dots\dots \quad (2)$$

In other words, if a picture successfully completes all the processes, it may be identified as a human face. However, it is not regarded as a human face if it fails at any one of these processes. This encapsulates the algorithm's operation.

## IV. EXPERIMENTAL RESULTS

The team ran studies utilising various collections of portrait images of people. The "Bao database," one of the datasets they used, has images of faces in various lighting and positions. Their programme has a 92% accuracy rate for accurately identifying faces in the database. Additionally, they tested their method on several databases, including the "Yale Face database" and the "AR-Face database." These databases mostly include plain, unchanging pictures of faces.

They utilised these datasets by choosing photographs with different characteristics for their trials in order to test their algorithm. They were able to properly train their facial recognition system as a result. They can use specialised strategies based on setting a limit to detect the many components of a person's face in their algorithm and delete any regions that do not belong there. These tests' outcomes demonstrate how well their system recognises face characteristics.

 

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.13446

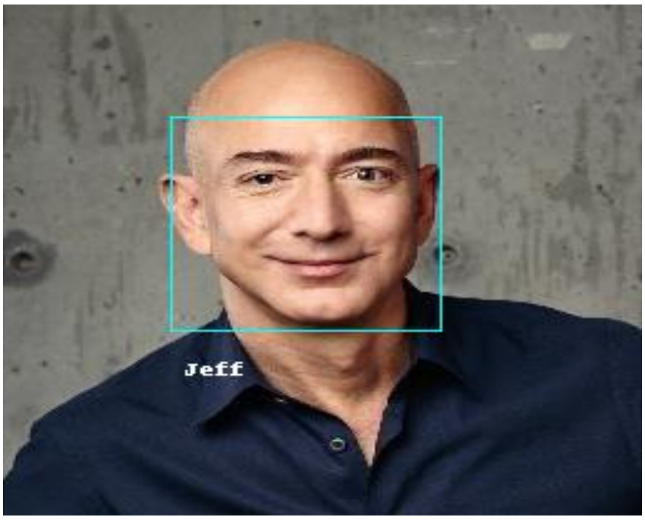


Fig:3.1

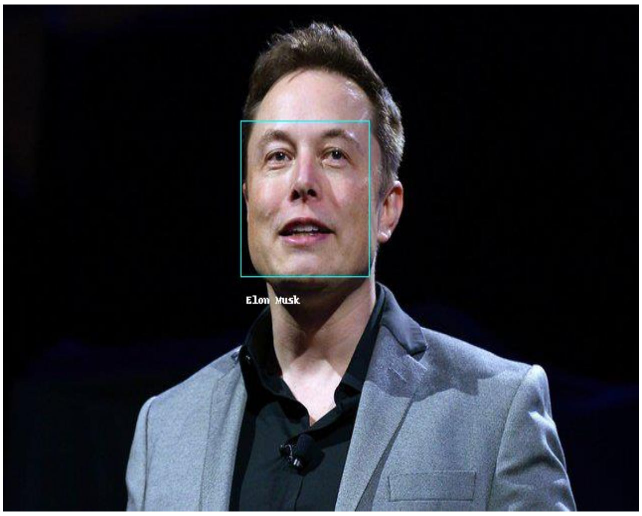

Fig:3.2

## V. CONCLUSION

This essay offers an improvement on the Viola Jones technique for identifying faces in images.We can identify faces with greater accuracy using this new method around 90% of the time. The good news is that it takes the same amount of time as the previous method to achieve this. Therefore, our novel technology locates faces more effectively without taxing the computer.

## REFERENCES

[1] Ms. Aswathy.R ,IOSR Journal of Computer Engineering (IOSR-JCE) e- ISSN: 2278-0661, p- ISSN: 2278-8727Volume 11, Issue 1 (May. - Jun. 2013), PP 61-64.
[2] P. S. Aleksic and A. K. Katsaggelos, “Automatic Facial Expression Recognition Using Facial Animation Parameters and Multistream HMMs”, IEEE Transactions on Information Forensics and Security, vol. 1, no. 1, pp. 3-11, March 2006.
[3]. TaSenthil Kumar and Saivenkateswaran, Sb, “Evaluation of video analytics for face detection and recognition”, International Journal of Applied Engineering Research, vol. 10, pp. 24003-24016, 2015.
[4]. Fernandez, Ma Christina D., et al. "Simultaneous face detection and recognition using Viola-Jones Algorithm and Artificial Neural Networks for identity verification." Region 10 Symposium, 2014 IEEE. IEEE, 2014.

DOI: 10.17148/IJARCCE.2024.13446

[5]. Arulkumar, C. V., and P. Vivekanandan. "Multi-feature based automatic face identification on kernel eigen spaces (KES) underunstable lighting conditions." Advanced Computing and communication Systems, 2015 International Conference on. IEEE, 2015.

[6]. Hsu, Rein-Lien, Mohamed Abdel-Mottaleb, and Anil K. Jain. "Face detection in color images." Pattern Analysis and Machine Intelligence, IEEE Transactions on 24.5 (2002): 696-706.

[7] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, "Classifying facial actions," IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 10, pp. 974–989, Oct. 1999.

[8]. Agrawal, Samiksha, and Pallavi Khatri. "Facial expression detection techniques: based on Viola and Jones algorithm and principal component analysis." Advanced Computing & Communication Technologies (ACCT), 2015 Fifth International Conference on. IEEE, 2015.

[9].G. Mehta ,S. Vatta ''An Introduction to a Face Recognition System using PCA, FLDA and Artificial Neural Networks'', IJARCSSE Volume 3, Issue 5, May 2013.

[10]. R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, Face Detection in Color Images, IEEE Transaction on Pattern Analysis and Machine Learning, 24 (5) (2002) 696-706.

[11]. C. Zhang and Z. Zhang, A Survey of Recent Advances in face Detection, Technical Report, MSRTR-2010-66, Microsoft Research, Microsoft Corporation, June 2010.

[12]. Viola, Paul, and Michael J. Jones. "Robust real-time face detection."International journal of computer vision 57.2 (2004): 137-154.

[13]. Yow, Kin Choong, and Roberto Cipolla. "Feature-based human face detection." Image and vision computing 15.9 (1997): 713-735

[14] L. Ma and K. Khorasani, "Facial Expression Recognition Using Constructive Feedforward Neural Networks", IEEE Transactions on Systems, Man and Cybernetics—Part B: Cybernetics, vol. 34, no. 3, pp.1588-1595, June 2004.

 