Harnessing Viola-Jones for Effective Real-Time Crowd Monitoring Based on Image Processing Techniques

T. Vasudeva Reddy[1], Santhosh Kumar[2], V Sreelatha Reddy[3], T. L. Kayathri[4], Raja Suresh[5] and D. Srikar[1]

[1]*Dept. of ECE. B V Raju Institute of Technology, Narsapur, Medak (dist), Telangana, India*
[2]*Dept of ECE, BVRIT HYDERABAD College of Engineering for Women, Hyderabad, Telangana, India*
[3]*EIE dept., CVR College of Engineering Ibrahimpatnam, Hyderabad, India*
[4]*Department of ECE, M. Kumarasamy College of Engineering, Thalavapalayam, Karur, T,N, India*
[5]*Dept. of ECE, Sri Venkateswara College of Engineering, Tirupati, Andhra Pradesh, India*

Keywords: Public Safety, Event Logistics, Urban Planning, Emergency Response, Student Safety, and Educational Technology.

Abstract: A ground breaking approach to classroom management has been developed, harnessing the power of advanced image processing techniques to monitor student activity and gauge cro wd density in real-time. By leveraging sophisticated algorithms, this innovative system empowers educators to respond swiftly and effectively to shifting classroom conditions. This technology has far-reaching implications that extend beyond the classroom walls. Its applications in public safety, event logistics, and urban planning can significantly enhance the overall quality of life. For instance, in public safety, this technology can facilitate rapid emergency response and minimize the risk of accidents. In event logistics, it can optimize crowd flow and density, ensuring a more enjoyable and secure experience for attendees In urban planning, it can inform data-driven decisions, enabling city officials to design more efficient and sustainable public spaces. By providing actionable insights into crowd behaviour, this system can have a profound impact on various aspects of society. In educational settings, it can improve student safety, enhance the learning experience, and foster a more productive and inclusive classroom environment.

1 INTRODUCTION

One of the pivotal techniques in video surveillance is people counting, a task that often encounters several challenges in crowded settings. These challenges include low resolution, occlusions, fluctuations in lighting, diverse imaging angles, and background clutter, making it a complex task. People counting serves as an essential function within intelligent video surveillance systems, offering substantial utility and commercial value across various venues such as banks, train stations, shopping centers, and educational institutions. The complexities of accurately counting individuals in densely populated surveillance zones are amplified by these environmental and technical factors (Aish, Zaqoot, et al. , 2023). During the ongoing pandemic, the ability to count and analyse the distribution of people within a camera's view can play a crucial role in mitigating the spread of COVID-19. Various techniques such as segmentation, pixel counting, and feature extraction are employed to detect crowds, though challenges arise, such as background elimination, when camera setups are widespread.

Our focus is primarily on crowd counting which involves distinguishing the crowd from background disturbances and quantifying the number of individuals within the crowd. While substantial research has been conducted in human recognition, it tends to be most effective in less crowded environments. Traditional CCTV systems often suffer from low-resolution issues, prompting the introduction of cost- effective high-resolution cameras. However, these solutions face their own set of challenges, including maintaining high image quality and managing the increased processing load from multiple cameras operating simultaneously. Additionally, in densely populated areas, the high

Reddy, T. V., Kumar, S., Reddy, V. S., Kayathri, T. L., Suresh, R. and Srikar, D.
Harnessing Viola-Jones for Effective Real-Time Crowd Monitoring Based on Image Processing Techniques.
DOI: 10.5220/0013617000004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 3*, pages 368-376
ISBN: 978-989-758-763-4

crowd density reduces the number of pixels available per person, complicating the use of standard individual detection methods. (Ma, Siyuan, et al. , 2023).

Face detection involves locating one or more human faces within images or videos and is crucial for applications in biometric authentication, security, surveillance, and media indexing. While humans can easily recognize faces, this task presents significant challenges for computers. It is treated as major issues in the computer vision due to considerable variations within the same class brought about by differences in facial features, lighting conditions, and expressions. This research utilizes the MATLAB r2020b nt to develop a face detection using the Viola-Jones algorithm. The model having an ability to detect faces in real-time and has become a standard in identifying frontal faces within images (Ramasamy, Praba, et al. , 2020), (Meivel, Indira Devi, et al. , 2021). The objective of this paper is to delve into the Viola-Jones algorithm's methodology for counting heads. Initially designed for rapid image detection, the algorithm employs a two-stage process: training and detection, leveraging Haar- like features. This study further examines the use of Integral Images, Training Classifiers, Adaptive Boosting, and Cascading techniques, which collectively enhance the accuracy and efficiency of the algorithm.

Originally developed by Paul Viola et al; in 2001, the Viola-Jones algorithm is a prominent machine learning framework for object and face detection known for its use of a boosted cascade of simple features. It integrates Ada Boost and Haar feature-based cascaded classifiers for effective object discrimination. Haar features, which are small rectangular patterns of image intensities, help distinguish between different objects. AdaBoost, a machine learning method, enhances classifier effectiveness by converting weak classifiers into strong ones.

During the initial phase, the algorithm trains multiple weak classifiers on a large dataset containing both positive and negative examples. These classifiers, simple decision trees, use Haar features to recognize specific patterns in the images (Garg, Hamarneh, et al. , 2020), (Rajeshkumar, Samsudeen, et al. , 2020) Subsequently, AdaBoost amalgamates these weak classifiers into a robust classifier. In the detection phase, the algorithm scans the image with a window of varying sizes, applying the strong classifier to each sub- window. If a sub- window is identified as likely containing an object, it is passed on to the next level of the cascade. This cascade, composed of increasingly complex and computationally demanding classifiers, is designed to quickly dismiss non-object sub- windows while advancing those likely containing objects. (Zhou, Wei, et al. , 2020). Due to its speed and accuracy, the Viola-Jones algorithm has gained popularity for facial recognition. It is utilized in numerous applications, such as mobile devices, security systems, and digital cameras.

2 LITERATURE SURVEY

In the referenced study (Lienhart, Kuranov, et al. , 2003), Viola and colleagues developed a quick detection system using an improved cascade of straight forward classifiers. In this research, two significant enhancements to their approach and provide a detailed empirical analysis of these improvements. First, incorporate a new set of Haar-like features. Innovative features not only retain simplicity of the original qualities found in (Rajeshkumar, Samsudeen, et al. , 2020) but are also easily computable, significantly enhancing detection capabilities. Utilizing these rotating features, our prototype face detector achieves an average reduction in the false alarm rate by 10% at the same hit rate.

Secondly, conducting an exhaustive evaluation of the detection efficiency and computational demands of various boosting methods, specifically Discrete, Real, and Gentle Ada Boost, alongside different weak classifiers. Our findings indicate that Gentle Ada Boost demonstrates superior performance over Discrete Ada Boost, particularly when utilizing small CART trees as the foundational classifiers. This analysis offers valuable insights into optimizing detection frameworks for better accuracy and efficiency in real-world applications.

In reference (Timo, Matti, et al. , 2002), the authors introduced a method for classifying grayscale and rotation-invariant textures through a simple yet effective multi resolution approach that utilizes binary patterns and nonparametric discernment between sample and prototype. This technique builds on the concept that certain local binary patterns, known as "uniform," play a crucial role in defining local image textures, and that their histogram distribution serves as a powerful feature for texture analysis.

To enhance this framework, developed a widespread operator for gray-scale and rotation invariance that facilitates the integration of multiple operators for multi resolution analysis. This advancement allows the identification of "uniform" patterns across any quantization of angular space and at whichever level of three-dimensional resolution.

Furthermore, by its very nature, this operator remains invariant to any monotonic grayscale transformations, making our proposed method exceptionally robust against variations in grayscale. This resilience enhances the utility and applicability of the technique in diverse and dynamic imaging conditions. In reference (Kruppa, Santana, et al. , 2003), the authors highlight that face detection plays a critical role in initiating tracking algorithms within visual surveillance systems. Recent findings in psychophysics suggest the importance of incorporating the local context of a face, such as head contours and the torso, for effective detection. The detector developed in this study leverages this concept of local context, which enhances its robustness beyond the capabilities of traditional face detection methods. This enhanced robustness makes it particularly appealing for use in surveillance applications. The study referenced in (Marco, Oscar, et al. , 2007) explores a face detection system that surpasses traditional approaches usually applied to still images. This system is uniquely designed to leverage the temporal coherence found within video streams, creating a more dependable detection framework. By utilizing cue combination, it achieves multiple and real-time detection capabilities. For each detected face, the system constructs a feature-based model and tracks it across successive frames using varied model information. The research focuses specifically on video streams, where the advantage of incorporating temporal coherence can be fully realized. In reference (Marco, Oscar, et al. , 2007), a notable aspect of advanced driver assistance systems (adas) in contemporary vehicles is traffic sign recognition (tsr), also referred to as road sign recognition (rsr). To meet the critical demands to achieve real- time performance and resource efficiency, we propose a highly optimized hardware implementation for traffic sign recognition (tsr). The tsr process is divided into two main phases: detection and recognition. During the detection phase, the normalized rgb colour transform is utilized along with single-pass connected component labelling (ccl) to effectively identify potential traffic signs. Most German traffic signs are assumed to be red or blue and typically take the form of circles, triangles, or rectangles. Our enhancement to single- pass ccl eliminates "merge-stack" operations by recording related regional relationships during the scanning phase and updating the labels in the iteration phase, thereby streamlining the process..

3 FACE DETECTION USING VIOLA JONES ALGORITHM

The widely celebrated application of the Viola-Jones algorithm is in face detection. The fundamental stages indicates here are:

Data Preparation: This initial stage involves assembling a dataset consists of both positive and negative images. Positive images contain faces, while negative images do not. These collections are crucial for training the algorithm.

Feature Extraction: The second stage involves using Haar features to extract relevant data from the images by calculating the difference between the sum of pixel intensities in white and black areas, which are represented by rectangular patches.

Training: During this phase, the Ada Boost algorithm is employed to train a series of weak classifiers. Each classifier is specifically trained to recognize a certain Haar feature.

Detection: In the detection stage, the algorithm slides a window of various sizes over a test image, applying a cascade of classifiers to each window. This cascade determines whether each window contains a face or not, with faces being indicated by a bounding box.

Post-processing: Finally, the algorithm applies post- processing techniques to refine the detection results, reducing false positives and enhancing the accuracy of the bounding boxes. The Viola-Jones algorithm is favoured in numerous applications due to its speed and accuracy in face detection. It has been effectively implemented in devices and systems such as digital cameras, facial recognition software, and security systems, showcasing its practical utility in real- world scenarios.

Data Collection: The data is gathered from secondary resources such as scholarly articles, books, and digital platforms. Each source was meticulously chosen for its relevance and reliability concerning the subject matter. Scholarly publications and texts pertaining to the Viola-Jones algorithm and image processing formed the cornerstone of the research. Digital resources including blogs, forums, and other online platforms provided additional insights, enriching the primary data acquired from academic texts.

Figure 1: Edge Features

Data Analysis: The analysis of data in this study employed a qualitative approach. This involved a thorough examination of the collected sources to extract pertinent details. This information was then organized into thematic categories, which guided the development of the research methodology. Qualitative research emphasizes the interpretation of social phenomena to grasp their deeper meaning and complexity. It is particularly effective in addressing the qualitative dimensions of human experiences that are beyond numerical measurement. In this study, the qualitative method facilitated the detection of patterns, linkages, and themes within the data. Organizing the data into themes provided structured framework that enhanced the analysis, highlighting similarities and variances within the data. This structured approach not only clarified the understanding of the topic but also ensured that the devised methodology was deeply rooted in the empirical data and aligned with the research objectives.

4 METHODOLOGY

Detection Process: The Viola-Jones algorithm is engineered primarily to election of frontal faces, experiencing limitations when identifying faces angled to the side, upward, or downward. Initially, the image is transformed into grayscale, a format that simplifies processing due to its reduced data requirements. The algorithm first pinpoints a face within these grayscale images before locating the same face in the collared image. It delineates a rectangular box that commences its search from the top-right corner, moving leftward. This process involves scanning for Haar-like features, which will be elaborated on later in this discussion. The search process incrementally shifts the right boundary of the rectangle with each step across the image.

4.1 Haar-like Features

Named after Alfred Haar, a Hungarian mathematician known for his work on Haar wavelets, these features are crucial for the algorithm's function and characterized by adjacent rectangular regions at a specific position within a detection window, contrasting in pixel intensities, which the algorithm uses to distinguish different facial features. Viola and Jones categorized these into three main types: a. Edge Features, b. Line Features, c. Four- Rectangle Features, d. Integral Images, e. Training Classifiers, f. Adaptive Boosting, g. Cascading.

4.2 Edge Features

Edge features are particularly effective in scenarios like eyebrow detection where there is a stark contrast between the dark pixels of the eyebrows and the lighter skin tones surrounding them. These features capture these abrupt changes in pixel intensity, aiding in the accurate delineation of facial features as showed in Figure1.

4.3 Line Features

Line features excel in identifying areas such as the lips, where the pixel intensity transitions from light to dark and back to light. These features are adept at capturing the nuanced changes in shading across the facial contour. For a visual example, see Figure2.

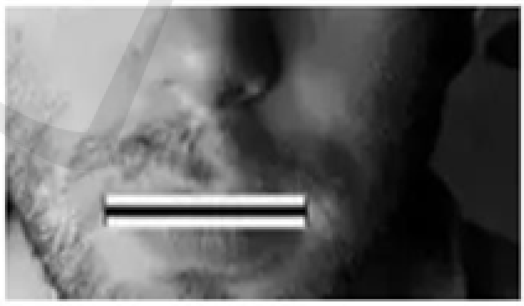

Figure 2: Line Features

4.4 Four Side Features

Each feature is assigned a value that reflects its effectiveness in aiding the machine's interpretation of the image. This value is determined by measuring the difference between sum of the pixels values of black and the white region. For a visual reference represented in Figure3, where the value is obtained by subtracting the white area from the black area.

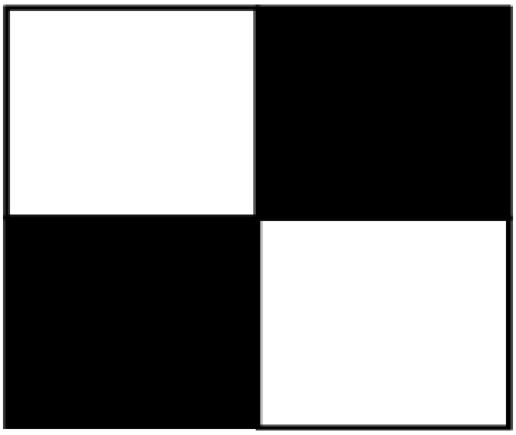

Figure 3: Four rectangular features

4.5 Integral Images

As the number of pixels increases or in a larger image, calculating the value of a feature becomes extremely difficult. To execute complex computations quickly and to determine to meet the requirements, the Integral Image idea is used as depicted in Figure4.

4.6 Training Classifiers

During the training stage of the Viola-Jones algorithm, the classified information is fed into the algorithm to learn from the data and make predictions. The algorithm establishes a minimal threshold to decide if anything qualifies as a feature or not. The image is reduced to 24x24 pixels during the training stage and is scanned for features.

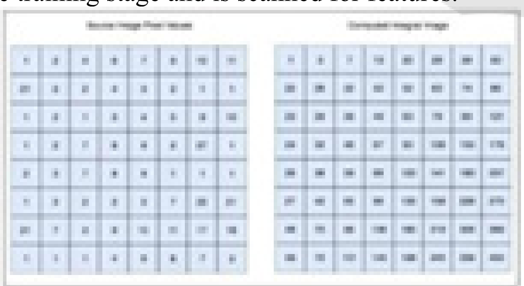

Figure 4: Working of integral image

4.7 Adaptive Boosting

To obtain an accurate model, all potential places and configurations of features are examined. Training can be computationally expensive as it takes a long time to examine every potential combination in each image. A strong classifier is referred to as F(x), and a single classifier is weak, but combining two or three weak classifiers produces strong classifier, which is called an ensemble.

4.8 Cascading

Cascading is an additional technique used to enhance the model's accuracy and speed. The sub-window is selected, from which the best characteristic is chosen to identify its presence in the image. If not, sub-window is rejected, and the next one is selected. If it is present, the second feature is examined, and if not, it is rejected, and the process continues. This process is sped up by cascading, and the machine produces results more quickly.

5 CASCADE OBJECT DETECTOR

It employs the Viola-Jones method to detect human faces, noses, eyes, mouths, and upper bodies. Viola-Jones algorithm processes a grayscale image, analyzing multiple smaller sub-regions to identify facial features by examining specific traits within each sub-region. Given that an image may contain faces of various sizes, the algorithm must inspect multiple locations and scales. Viola and Jones leveraged Haar-like features in their method to facilitate face detection. Furthermore, a custom classifier can be trained to operate with this system object using the Image Labeller tool. This allows for precise identification of upper body or facial features within an image.

Call the object with arguments, as if it were a function It uses the Viola-Jones method to create an object detector. This method takes time to train but can quickly recognize faces in real-time.

The process has four key steps: a. choosing specific features b. making an integral image, Training with Ada Boost, c. Building classifier cascades

Figure 5: Input image for Cascaded Object Detection

Figure 6: Output image for Fig 5

Figures 5 and 6 illustrate that edge and line features help detect edges and lines, while four-sided features are used to find diagonal patterns.

Cascade Object Detection Algorithms: To develop an object recognition model, features are extracted from labeled images to capture the target object's characteristics. These features are then used to build a classification model, which is employed for robust object detection. Training image dimensions determine the smallest detectable object region, so the Min Size property should be set accordingly. Careful feature extraction, image size selection, and model parameter configuration are crucial for building a reliable classification model. By following this process, a model can be constructed for various object recognition tasks, enabling accurate detection and classification of objects in images.

Cascade of Classifiers: The detector employs a series of classifiers that are organized in a series to scan efficiently the image for the targeted object. As mentioned in Figure 7, each stage uses increasingly the binary classifiers that enabling the rapid dismissal of regions that lacks with the target object. If a region does not pass, then it is immediately discarded, by avoiding the necessity for more intensive analysis in later stages.

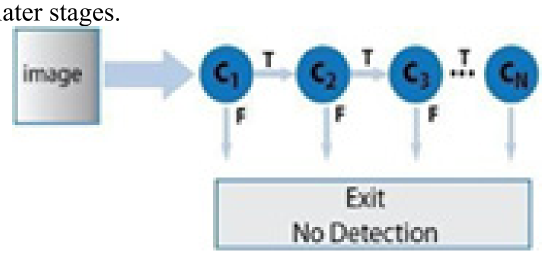

Figure7. Cascade of Classifiers

Multi-resolution Object Detection: Detector adjusts input image size to identify target objects, using a sliding window that matches training image dimensions. Scaling increments follow Scale Factor property.

Calculating search area size at each step Correlation among Man-size, Max-Size, and ScaleFactor: Object size and scale factor crucial for detection parameters. Min Size and Max Size define detectable object size range. Adjusting parameters when object size is known reduces computation time. Scale Factor impacts search window sizes.

Merge Detection Threshold: Search window detects objects at each scaling increment, producing multiple detections. Detections merged into single bounding box per object based on Merge Threshold property, ensuring accurate object detection.

rgb2gray: Converts RGB image or colour map to grayscale, preserving luminance information, discarding hue and saturation details. Syntax: I = rgb2gray (map).

Figure 8. Input image for Gray threshold

Figure 8: Output for the Figure8.

Insert Object Annotation It returns a true colour image marked with shapes and labels at specified positions. See Figure9. and Figure9 RGB = insert Object Annotation (I, shape, position, label) Counting number of people.

Figure 9: Input image for insert Object Annotation

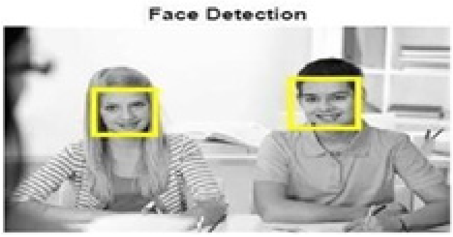

Figure 9: Output image for Figure9.1

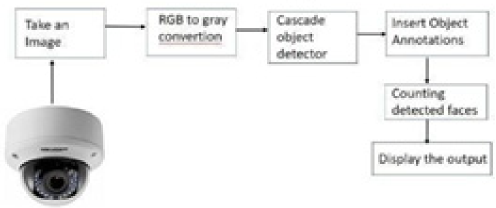

Figure 10: Block diagram

The counted number of people using the function num2str(n). It converts numbers to character array.

S = num2str(n)

The output format is determined by the magnitudes of the original values, making it effective for labelling and titling plots that contain numeric data.

6 IMPLEMENTATION

6.1 Mathematical Formulation

Let's denote the input image as I(x,y), where x and y are the spatial coordinates. The Haar wavelet feature extraction can be represented as:

$$H(x,y)=\sum i=0N-1hi\backslash*I(x+i,y) \quad (1)$$

where hi are the Haar wavelet coefficients and N is the number of coefficients. The AdaBoost training can be represented as:

$$F(x,y)=\sum i=0M-1fi\backslash*H(x+i,y) \quad (2)$$

where fi are the AdaBoost weights and M is the number of weak classifiers.

The cascaded classifier can be represented as:

$$C(x,y)=\sum i=0L-1ci\backslash*F(x+i,y) \quad (3)$$

where ci are the cascaded classifier weights and L is the number of stages. ci are the cascaded classifier weights and L is the number of stages.

Application to Crowd Monitoring to apply the Viola- Jones algorithm to crowd monitoring, we can use the following approach:

6.1.1 Pre-processing

Apply pre-processing techniques such as background subtraction and thresholding to enhance the quality of the input image.

Background Subtraction:

Let I(x,y) be the input image and B(x,y) be the background image. The background subtracted image S(x,y) can be represented as

$$S(x,y)=\max(I(x,y),B(x,y))|I(x,y)-B(x,y)| \quad (4)$$

Thresholding

Let I(x,y) be the input image and T be the threshold value. The thresholded image T(x,y)

Noise Removal (Gaussian Filter)

Let I(x,y) be the input image and G(x,y) be the Gaussian filter. The filtered image F(x,y) can be represented as:

$$F(x,y)=\Sigma\, i=-NN\, \Sigma\, j=-NNG(i,j)\cdot I(x+i,y+j)/(2N+1)2 \quad (5)$$

Where N is the filter size and G(i,j) is the Gaussian filter kernel.

Edge Detection (Sobel Operator)

Let I(x,y) be the input image. The edge detected image E(x,y) can be represented as:

$$E(x,y)=(Gx2+Gy2)/2 \quad (6)$$

where Gx and Gy are the gradients in the x and y directions, respectively.

6.1.2 Feature Extraction

Use Haar wavelets to extract features from the pre- processed image. Let I(x,y) be the pre-rocessed image. The Haar wavelet feature extraction can be represented as:

$$H(x,y)=\Sigma\, i=0N-1hi\cdot I(x+i,y) \quad (7)$$

where hi are the Haar wavelet coefficients and N is the number of coefficients. Alternatively, you can also use the following equation:

$$H(x,y)=\sum i=0N-1\sum j=0N-1hi\cdot hj\cdot I(x+i,y+j) \quad (8)$$

This equation uses a 2D Haar wavelet transform to extract features from the image. hi and hj are the Haar wavelet coefficients, which are used to convolve with the image to extract features. Also, N is the number of coefficients, which determines the size of the feature vector. Adjust the value of N to change the number of features extracted from the image.

6.1.3 Classifier Training

Train an AdaBoost classifier on the extracted features to detect people in the crowd. Let H(x,y) be the extracted features from the pre-processed image. The AdaBoost classifier training can be represented as:

$$F(x,y)=\sum i=0M-1fi\cdot hi\cdot H(x+i,y) \qquad (9)$$

Where fi are the AdaBoost weights, hi are the Haar wavelet coefficients, and M is the number of weak classifiers. This equation uses a combination of Haar wavelet features and AdaBoost weights to train a strong classifier for people detection in the crowd. fi are the AdaBoost weights, which are used to combine the weak classifiers to form a strong classifier. M is the number of weak classifiers, which determines the complexity of the strong classifier. Adjust the value of M to change the accuracy and speed of the people detection algorithm.

6.1.4 Cascaded Classifier

Use the trained classifier in a cascaded framework to detect people in the crowd. Let F(x,y) be the strong classifier trained in the previous stage. The cascaded classifier can be represented as:

$$C(x,y)=\sum i=0L-1ci\cdot F(x+i,y) \qquad (10)$$

where ci are the cascaded classifier weights and L is the number of stages. This equation uses a product of strong classifiers to form a cascaded classifier, which improves the accuracy and speed of people detection in the crowd. ci are the cascaded classifier weights, which are used to combine the strong classifiers to form a cascaded classifier. L is the number of stages, which determines the complexity of the cascaded classifier. This proposal is implemented using MATLAB software Version r2020b.

Procedure to conversion: Take an image as input, Convert the input image into gray using “rgb2gray”,Cascade the gray image using “vision. Cascade Object Detector”, Insert the object annotations for cascaded image using “insert Object Annotations”. Count the people present in the image, refer Figure10. Display the count of people.

7 RESULTS

The Figure11.is taken as the input image.

Figure 11:Input Image

The Figure11 is converted to gray image using “rgb2gray”.

Figure 12: Gray image for Figure11

The grayscale image in Figure 12 is analysed using the "vision. Cascade Object Detector," which applies the Viola- Jones algorithm for facial detection. To label the detected objects in the processed image, the "insert Object Annotation" function is utilized, as illustrated in Figure 13.

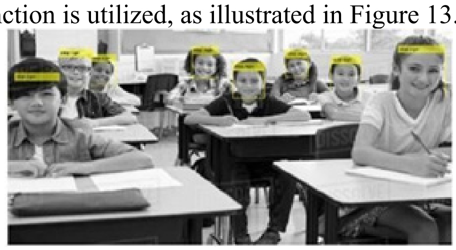

Figure 13: Cascaded image of the Figure12

Displayed the count of people. The output of the Figure11 shown in Figure14.

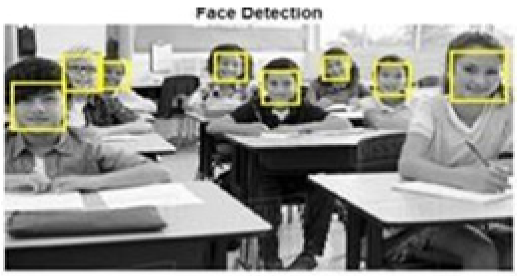

Figure14: Output image

8 CONCLUSION

The Viola-Jones algorithm is a powerful tool for detecting human faces in images, utilizing a combination of techniques such as Haar-like features, Integral Images, and Adaptive Boosting. This research sheds light on the algorithm's capabilities and limitations, paving the way for future exploration. Potential applications include crowd counting, surveillance, and healthcare, while integrating deep learning techniques could further enhance its accuracy and efficiency. By optimizing its performance and exploring new use cases, the Viola-Jones algorithm can become an even more valuable asset in various fields.

REFERENCES

Adnan M. Aish, Hossam Adel Zaqoot, Waqar Ahmed Sethar, Diana A. Aish; 2023. *Prediction of groundwater quality index in the Gaza coastal aquifer using supervised machine learning techniques*. Water Practice and Technology 2023.

Ma, Siyuan & Hu, Qintai & Zhao, Shuping & Wu, Wenyan & Wu, Jigang. 2023. *Multi-Scale Multi-Direction Binary Pattern Learning for Discriminant Palmprint Identification.* IEEE Transactions on Instrumentation and Measurement.

Ramasamy, Dhivya Praba, and Kavitha Kanagaraj. 2020. *"Object detection and tracking in video using deep learning techniques: A review." Artificial Intelligence Trends for Data Analytics Using Machine Learning and Deep Learning Approaches*

S. Meivel, K. Indira Devi, S. Uma Maheswari, J. Vijaya Menaka,, *2021 Real time data analysis of face mask detection and social distance measurement using Matlab*,Materials Today: Proceedings.

Garg S, Hamarneh G, Jongman A, Sereno JA, Wang Y. 2020. *ADFAC: Automatic detection of facial articulatory features. MethodsX.*.

Rajeshkumar, T., U. Samsudeen, S. Sangeetha, and U. Sudha Rani. 2020. *"Enhanced visual attendance system by face recognition USING K- nearest neighbor algorithm."* Journal of Advanced Research in Dynamical and Control Systems.

Zhou, Wei, Shengyu Gao, Ling Zhang, and Xin Lou. 2020.*"Histogram of oriented gradients feature extraction from raw bayer pattern images."* IEEE Transactions on Circuits and Systems II:

Lienhart R., Kuranov A., and V. Pisarevsky 2003.*"Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection."* Proceedings of the 25th DAGM Symposium on Pattern Recognition. Magdeburg, Germany,

Ojala Timo, Pietikäinen Matti, and Mäenpää Topi, 2002. *"Multiresolution Gray- Scale and Rotation Invariant Texture Classification with Local Binary Patterns".* IEEE Transactions on Pattern Analysis and Machine Intelligence,

Kruppa H., Castrillon-Santana M., and B. Schiele. 2003.*"Fast and Robust Face Finding via Local Context". Proceedings of the Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance,*

Castrillón Marco, Déniz Oscar, Guerra Cayetano, and Hernández Mario," ENCARA2: 2007. *Real-time detection of multiple facesat different resolutions in video streams"* Journal of Visual Communication and Image Representation,