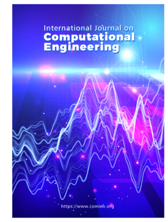
International Journal on Computational Engineering

Journal Homepage www.comien.org/index.php/comien



Comparative Analysis of Multi-pose Face Detection Between Yolo and Haar Cascade Classifier

Muhammad Fiqih Irfiansyah [a], Anang Kukuh Adisusilo [a,*]

[a] *Department of Informatics, Faculty of Engineering, Universitas Wijaya Kusuma Surabaya, Surabaya, Indonesia*

*Corresponding author: *anang65@uwks.ac.id*

Abstract— **Face detection is a biometric technology used to identify individuals based on their facial features. However, this technology faces challenges when detecting faces from various angles, which can affect the accuracy and speed of detection. To address this issue, algorithms such as You Only Look Once (YOLO) and the Haar Cascade Classifier have been used for object detection. YOLO is a real-time object detection algorithm, while the Haar Cascade Classifier is a simpler method that uses Haar-like features to detect objects. Several previous studies have tested both algorithms for object detection such as vehicles and crowd counting, with results showing that YOLO offers higher accuracy. This study aims to analyze the performance of YOLO and the Haar Cascade Classifier in detecting faces from various angles. The test results show that YOLO can consistently detect faces with 100% accuracy in all conditions. Meanwhile, the Haar Cascade Classifier also shows high accuracy, but experiences a significant drop at extreme angles of -90°. When the face is smiling in normal lighting, its accuracy is 36.96% for testing on images and 42.81% for testing on videos. Although the Haar Cascade Classifier has faster detection times, YOLO still excels in detection accuracy and consistency. Therefore, the algorithm selection can be tailored to the system's needs—whether prioritizing processing speed or detection accuracy.**

Keywords— **Face detection; YOLO; Haar cascade classifier; object detection; accuracy; detection speed.**

Manuscript received 15 Apr. 2025; revised 22 Jul. 2025; accepted 13 Sep. 2025. Date of publication 30 Sep. 2025.

I. Introduction

Face detection or face recognition is a biometric technology used to identify or verify a person's identity based on their facial features. A face detection system uses a camera to capture an image of a person's face, then matches it with stored data [1]. Face recognition is part of the advancement of artificial intelligence (AI) and machine learning, which makes it increasingly accurate [2].

However, this technology faces challenges in detecting faces from various angles, which can affect detection accuracy and speed. To overcome this issue, algorithms such as You Only Look Once (YOLO) and Haar Cascade Classifier have been implemented for object detection. YOLO is an algorithm capable of detecting objects in real-time, while Haar Cascade Classifier is a simpler method that uses Haar features for detection.

According to tests conducted by previous researchers, the YOLO method achieved the highest accuracy rate in motor vehicle detection at 74%, while the Haar Cascade method obtained an accuracy rate of 41% [3]. In a similar study, the results of visitor count detection using the YOLO method showed a success rate of 93.3%, while the Haar Cascade Classifier method achieved a success rate of 85.7%. Based on both test results, it can be concluded that YOLO achieved the highest level of accuracy [4].

Several recent studies further reinforce the advantages of YOLO in face detection tasks. YOLO5Face, a customized variant of YOLO-v5, achieves state-of-the-art performance on the WIDER Face dataset by optimizing the backbone architecture and landmark detection head for multi-pose scenarios [5]. More broadly, a comprehensive review tracing the evolution of YOLO from v1 to v8 highlights YOLO-v8's robust architecture and real-time detection capabilities, making it particularly effective for face and object detection applications [6]. Meanwhile, Advanced Face Detection with YOLOv8 demonstrates successful real-time face detection integration into AI modules, maintaining high accuracy even under challenging conditions such as occlusion and variable lighting [7], [8].

On the other hand, enhancements to the classic Haar Cascade algorithm continue to be explored. An improved version of Haar Cascade for gate pass security achieved a remarkable 98.39% detection accuracy under complex scenes, thanks to preprocessing steps like RGB conversion and logical filtering [9]. Additionally, comparative analysis in a driver drowsiness detection context shows that YOLO-based detectors provide more detailed regions of interest (ROI) than Haar Cascade, albeit with higher computational cost [10]. Such trade-offs underscore the need to evaluate both detection accuracy and processing efficiency when comparing these algorithms [11], [12].

Therefore, this study aims to analyze the performance of the YOLO and Haar Cascade Classifier algorithms in detecting faces from various different viewpoints. The benefit of this research is to provide a technical understanding of the most optimal face detection algorithms for different conditions. In addition, the results of this study can serve as a foundation for developing more accurate and efficient face detection systems, especially for applications in the security field that require real-time face detection.

II. Material and Method

The research was conducted through a structured sequence of stages to ensure an effective and timely process. The main stages carried out are as follows:

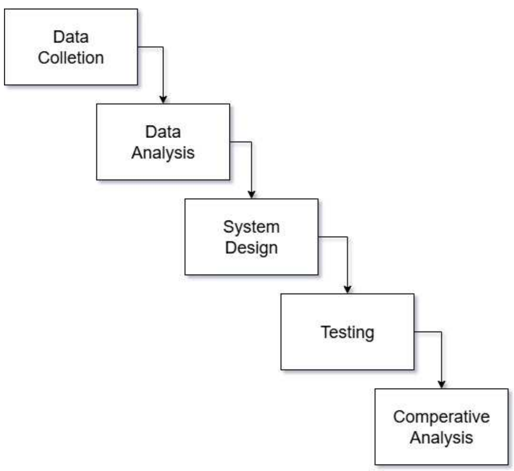


Fig. 1 Research Flow

A. Data Collection

The data were collected in the form of facial images with three variations: frontal, left side, and right side. The images were captured using a 720p camera and stored in .jpg format. The results of the data collection can be seen in Table I.

TABLE I
DATA COLLETION

Viewpoint	Resolution	Format	Number of Images
Frontal	720x1280	.jpg	100
Left Side	720x1280	.jpg	94
Right Side	720x1280	.jpg	49

As shown in Table 1, the dataset consists of 100 images for the frontal face viewpoint, 94 images for the left side, and 49 images for the right side.

B. Data Analysis

The purpose of this data analysis is to ensure that the data used in the face detection process meets the quality required by the YOLO and Haar Cascade Classifier algorithms.

1) Data Inspection: All collected images were manually inspected to ensure the face was clearly centered and correctly angled. A total of 243 images were collected and all of them were declared suitable for use in the labeling and training process. The labeled face image files are saved in .xml format. Examples of the collected facial images and their labeling process are shown in Fig. 2 and Fig. 3.

Fig. 2 Collected Facial Images

Fig. 3 Labeling Process

2) Preprocessing Data: The images were resized to 640×640 resolution to meet YOLO's input requirements. Additionally, the images were converted to grayscale specifically for the Haar Cascade Classifier to improve processing speed during detection.

Fig. 4 Resized and Grayscale Face Images

3) Dataset Requirements Analysis: The dataset was reviewed to ensure it met the required criteria in terms of quantity, variety, and suitability. This dataset includes three variations of facial angles: frontal with 100 images, left side with 94 images, and right side with 49 images. Although the distribution was not completely balanced, the amount of data was deemed sufficient to test both algorithms, covering a wide range of viewpoints to support the system's detection capabilities.

4) Dataset Preparation for Training: The dataset was prepared according to the training needs of YOLO and Haar Cascade Classifier. For Haar Cascade, images were divided into positive (with faces) and negative (without faces) data. Positive samples were used to generate a .vec file containing face coordinates, while negative samples helped reduce false detections. In contrast, YOLO used the original images and annotations directly, without requiring data separation or .vec file generation.

C. System Design

The face detection process in this research uses three types of input media: images, video recordings, and real-time. The system workflow can be seen in Fig. 5.

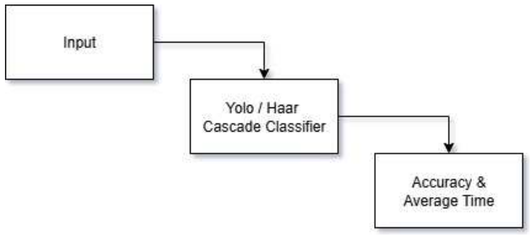


Fig. 5 System Workflow

As shown in Fig. 5, the system workflow begins with real-time input in the form of images, videos, or cameras. This input is then processed using one of two algorithms: YOLO or the Haar Cascade Classifier. From this process, the system produces a detection accuracy rate and average processing time.

YOLO detects objects in a single process by dividing the image into grids. Each grid predicts whether there is an object and its location. The following are the main formulas in YOLO:

$$\text{Confidence} = \text{P(Object)} \times \text{IOU}_{\text{pred}}^{\text{truth}} \tag{1}$$

Based on the formulas above, the following explanations are provided:

- P(Object): The probability or likelihood of an object being present in the grid, indicating whether a face is detected or not.
- $\text{IOU}_{\text{pred}}^{\text{truth}}$: The Intersection Over Union between the predicted bounding box (pred) and the labeled ground truth (truth).
- YOLO predicts face position based on the midpoint of the bounding box (x, y) width and height of the bounding box (w, h) Object class (face) with probability.

The Haar Cascade Classifier detects faces based on light and dark patterns on the face, using features known as Haar features. To accelerate the process, an integral image is applied, which allows the total pixel values to be computed very quickly.

Haar feature formula:

$$\text{Fitur} = \sum(\text{Bright Area}) - \sum(\text{Dark Area}) \tag{2}$$

Integral image formula:

$$ii(x,y) = \sum_{x^1=0}^{x} \sum_{y^1=0}^{y} i(x^1, y^1) \tag{3}$$

Based on the formulas above, the following explanations are provided:

- ii(x,y): The value of the integral image at coordinate (x, y).
- $i(x^1,y^1)$: The grayscale intensity (light-dark value) of the original pixel.
- The resulting features are used by the classifier at each stage to determine whether an image region is likely to contain a face or not.

Face detection testing using image input was conducted to analyze the performance of the YOLO and Haar Cascade Classifier algorithms. The process begins with model initialization, then each image is processed for face detection, and the results are stored to calculate accuracy and average processing time. The following is a flowchart of the image testing process, shown in Fig. 6.

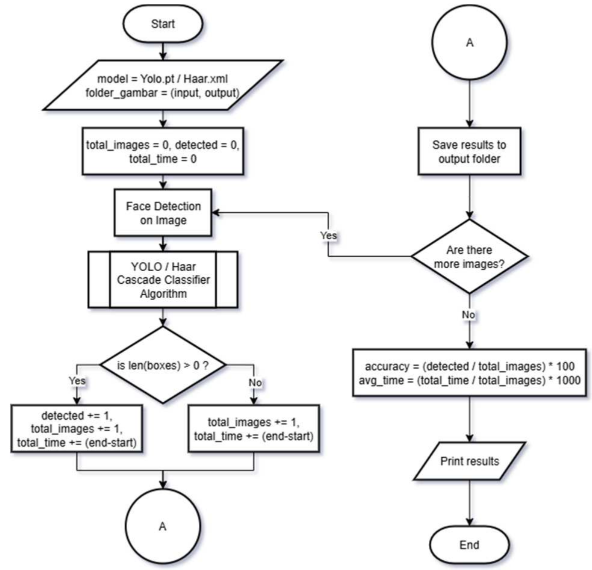


Fig. 6 Flowchart of Image Testing

As illustrated in the flowchart, the process starts with initializing the YOLO/Haar model and the input-output folder. Each image is analyzed using the face detection algorithm, and the system determines whether a face is detected. The number of images, detections, and processing times are recorded. If there are still more images, the process continues until all images are tested. At the end, accuracy is calculated as the ratio of detected images to total images, while the average processing time is computed by dividing the total time by the number of images.

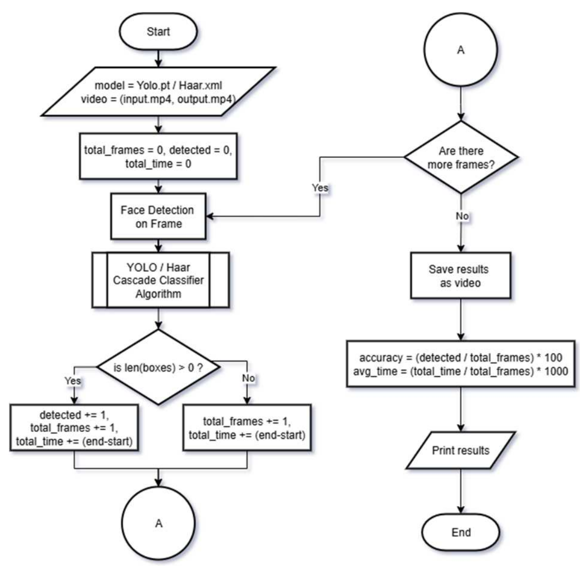


Fig. 7 Flowchart of Video Testing

Face detection testing on video input was performed to analyze the algorithm's performance when working with sequential moving frames. Each video frame was analyzed for face detection, and the results were saved as an output video for further evaluation. The following is a flowchart of the video testing process, shown in Fig. 7.

The flowchart shows that the system begins by reading the input video file and initializing the calculation parameters. Each frame is processed using the YOLO or Haar Cascade algorithm for face detection. The number of frames, detections, and processing times are recorded. If there are still more frames, the system continues until the entire video has been processed. Once completed, the detection results are saved into an output video file. Accuracy is calculated as the ratio of detected frames to total frames, while the average processing time is obtained by dividing the total time by the number of frames.

Real-time face detection testing is conducted using a live camera as input. This test aims to evaluate the speed and accuracy of the algorithms under real-world conditions. The process is carried out over a specific duration, and the results are saved as a video recording of real-time face detection. The following is a flowchart of the real-time testing process, shown in Fig. 8.

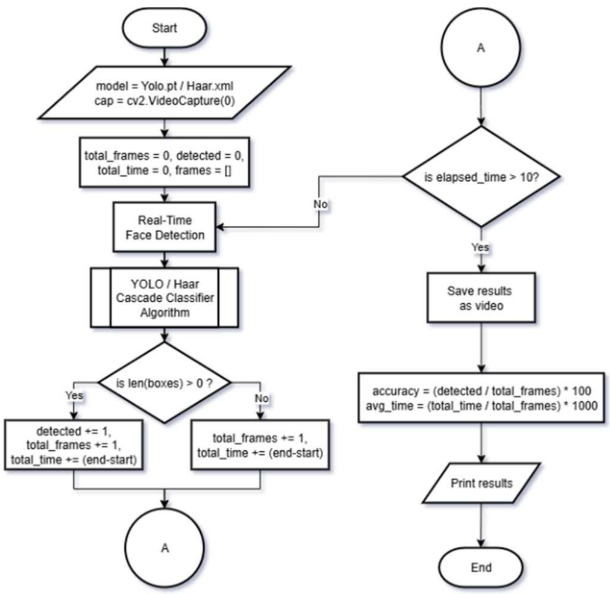


Fig. 8 Flowchart of Real-Time Testing

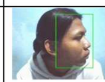
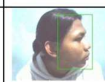
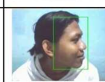
Based on the flowchart, the system begins by accessing the camera to capture live input. Each incoming frame is processed by the YOLO or Haar Cascade algorithm to detect faces. Detections, frame counts, and processing times are recorded. The process continues until the specified duration is reached (e.g., more than 10 seconds). Once the time limit is exceeded, the recorded results with detected faces are saved as an output video. Accuracy is calculated as the ratio of detected frames to total frames, and the average processing time is computed by dividing the total time by the number of frames.

III. Result and Discussion

A. Testing on Images

The system was tested using facial images with various viewing angles, ranging from -90° to 90°. Each angle was analyzed under bright and normal lighting conditions. Two types of facial expressions were also considered, namely neutral and smiling. The results were analyzed based on detection accuracy and processing time across all conditions. Detailed results are presented in Tables II and III, which provide a comparison of the detection results.

TABLE II
Testing Results of YOLO Algorithm on Images

Angle	Example Visual	Lighting	Expression	Total Images	Detected	Accuracy (%)	Time (ms)
-90°		Normal	Neutral	54	54	100%	30.86 ms
-90°		Bright	Neutral	54	54	100%	32.01 ms
-90°		Normal	Smile	46	46	100%	31.92 ms
-90°		Bright	Smile	46	46	100%	33.29 ms
-75°		Normal	Neutral	50	50	100%	34.57 ms
-75°		Bright	Neutral	50	50	100%	31.35 ms
-75°		Normal	Smile	54	54	100%	30.41 ms
-75°		Bright	Smile	54	54	100%	31.07 ms
-60°		Normal	Neutral	52	52	100%	29.92 ms
-60°		Bright	Neutral	52	52	100%	30.12 ms
-60°		Normal	Smile	55	55	100%	34.45 ms
-60°		Bright	Smile	55	55	100%	30.47 ms
-45°		Normal	Neutral	55	55	100%	31.76 ms
-45°		Bright	Neutral	55	55	100%	30.46 ms
-45°		Normal	Smile	55	55	100%	29.82 ms
-45°		Bright	Smile	55	55	100%	32.05 ms
-30°		Normal	Neutral	53	53	100%	36.02 ms
-30°		Bright	Neutral	53	53	100%	29.94 ms

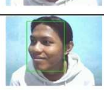
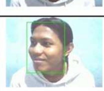
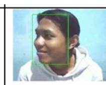
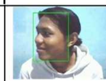
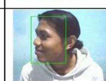
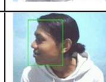
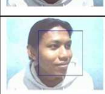
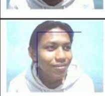
-30°		Normal	Smile	50	50	100%	32.18 ms
-30°		Bright	Smile	49	49	100%	33.68 ms
-15°		Normal	Neutral	55	55	100%	30.15 ms
-15°		Bright	Neutral	55	55	100%	32.77 ms
-15°		Normal	Smile	47	47	100%	31.24 ms
-15°		Bright	Smile	47	47	100%	34.52 ms
0°		Normal	Neutral	50	50	100%	33.09 ms
0°		Bright	Neutral	50	50	100%	36.09 ms
0°		Normal	Smile	51	51	100%	36.37 ms
0°		Bright	Smile	51	51	100%	32.38 ms
15°		Normal	Neutral	52	52	100%	30.77 ms
15°		Bright	Neutral	52	52	100%	31.32 ms
15°		Normal	Smile	53	53	100%	32.09 ms
15°		Bright	Smile	53	53	100%	35.24 ms
30°		Normal	Neutral	54	54	100%	34.65 ms
30°		Bright	Neutral	54	54	100%	30.86 ms
30°		Normal	Smile	50	50	100%	31.14 ms
30°		Bright	Smile	50	50	100%	32.98 ms
45°		Normal	Neutral	52	52	100%	31.24 ms
45°		Bright	Neutral	52	52	100%	33.89 ms
45°		Normal	Smile	50	50	100%	31.15 ms
45°		Bright	Smile	50	50	100%	35.81 ms
60°		Normal	Neutral	51	51	100%	33.39 ms
60°		Bright	Neutral	51	51	100%	33.68 ms
60°		Normal	Smile	52	52	100%	32.20 ms
60°		Bright	Smile	52	52	100%	32.03 ms
75°		Normal	Neutral	52	52	100%	29.83 ms
75°		Bright	Neutral	52	52	100%	34.30 ms
75°		Normal	Smile	50	50	100%	31.61 ms
75°		Bright	Smile	50	50	100%	32.66 ms
90°		Normal	Neutral	55	55	100%	32.29 ms
90°		Bright	Neutral	55	55	100%	35.56 ms
90°		Normal	Smile	54	54	100%	30.80 ms
90°		Bright	Smile	54	54	100%	31.47 ms

Table II presents the face detection test results using the YOLO algorithm based on a combination of variations in viewing angles, lighting conditions, and facial expressions. The viewing angles ranged from -90° to 90°, with two lighting conditions (normal and bright) and two facial expressions (neutral and smiling). Each combination was tested with a number of images ranging from 46 to 55, indicating a fairly even distribution of data across conditions, which provides a consistent basis for comparison across all scenarios. Test results show that the YOLO algorithm consistently detects faces with 100% accuracy across all test conditions, including varying lighting and expressions. Average detection times ranged from 29 ms to 37 ms, which is considered fast and efficient.

TABLE III
TESTING RESULTS OF HAAR CASCADE CLASSIFIER ON IMAGES

Angle	Example Visual	Lighting	Expression	Total Images	Detected	Accuracy (%)	Time (ms)
-90°		Normal	Neutral	54	54	92.59%	16.70 ms
-90°		Bright	Neutral	54	54	100%	21.04 ms
-90°		Normal	Smile	46	17	36.96%	15.56 ms
-90°		Bright	Smile	46	46	100%	24.45 ms
-75°		Normal	Neutral	50	50	100%	15.91 ms
-75°		Bright	Neutral	50	50	100%	18.40 ms
-75°		Normal	Smile	54	54	100%	18.79 ms
-75°		Bright	Smile	54	54	100%	24.39 ms
-60°		Normal	Neutral	52	52	100%	21.24 ms
-60°		Bright	Neutral	52	52	100%	25.29 ms
-60°		Normal	Smile	55	55	100%	21.65 ms
-60°		Bright	Smile	55	55	100%	25.84 ms
-45°		Normal	Neutral	55	55	100%	19.44 ms
-45°		Bright	Neutral	55	55	100%	33.99 ms
-45°		Normal	Smile	55	55	100%	16.58 ms
-45°		Bright	Smile	55	55	100%	25.80 ms
-30°		Normal	Neutral	53	53	100%	20.95 ms
-30°		Bright	Neutral	53	53	100%	23.37 ms
-30°		Normal	Smile	50	50	100%	21.12 ms
-30°		Bright	Smile	49	49	100%	24.88 ms
-15°		Normal	Neutral	54	54	100%	15.78 ms
-15°		Bright	Neutral	54	54	100%	29.43 ms
-15°		Normal	Smile	47	47	100%	21.64 ms
-15°		Bright	Smile	47	47	100%	29.30 ms
0°		Bright	Neutral	50	50	100%	26.44 ms
0°		Normal	Smile	51	51	100%	21.57 ms
0°		Bright	Smile	51	51	100%	26.07 ms
15°		Normal	Neutral	52	52	100%	18.42 ms
15°		Bright	Neutral	52	52	100%	29.75 ms
15°		Normal	Smile	53	53	100%	24.01 ms
15°		Bright	Smile	53	53	100%	26.86 ms
30°		Normal	Neutral	54	54	100%	30.43 ms
30°		Bright	Neutral	54	54	100%	30.25 ms
30°		Normal	Smile	49	49	100%	25.98 ms
30°		Bright	Smile	50	50	100%	22.46 ms
45°		Normal	Neutral	52	52	100%	17.98 ms
45°		Bright	Neutral	52	52	100%	18.47 ms
45°		Normal	Smile	50	50	100%	15.46 ms
45°		Bright	Smile	50	50	100%	16.33 ms
60°		Normal	Neutral	51	51	100%	17.80 ms
60°		Bright	Neutral	51	51	100%	25.81 ms

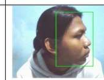
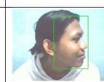
60°		Normal	Smile	52	52	100%	13.73 ms
60°		Bright	Smile	52	52	100%	15.00 ms
75°		Normal	Neutral	52	52	100%	16.71 ms
75°		Bright	Neutral	52	52	100%	20.74 ms
75°		Normal	Smile	50	50	100%	10.90 ms
75°		Bright	Smile	50	50	100%	11.93 ms
90°		Normal	Neutral	55	55	100%	16.27 ms
90°		Bright	Neutral	55	55	100%	19.47 ms
90°		Normal	Smile	54	54	100%	11.12 ms
90°		Bright	Smile	54	54	100%	10.01 ms

Table III presents the face detection test results using the Haar Cascade Classifier algorithm based on variations in viewing angles, lighting conditions (normal and bright), and facial expressions (neutral and smiling). The number of images tested in each combination ranged from 46 to 55. The results show that the Haar Cascade Classifier generally achieved high accuracy in most conditions, particularly at viewing angles from -75° to 90°. However, a notable decrease occurred at the -90° angle under normal lighting with a smiling expression, where accuracy dropped to 36.96%. The average detection time per image was relatively fast, ranging from 10.01 ms to 33.99 ms, demonstrating its strength in efficiency despite accuracy limitations under certain conditions.

B. *Testing on Video*

The system was tested using recorded face videos, where each frame containing a face was detected. The recorded videos included viewing angles ranging from -90° to 90°, under both normal and bright lighting conditions. Two types of facial expressions were also considered, namely neutral and smiling. The results were analyzed based on detection accuracy and processing time across all tested conditions. Detailed results for each algorithm are presented in Tables IV and V.

TABLE IV
TESTING RESULTS OF YOLO ON VIDEOS

Angle	Example Visual	Lighting	Expression	Total Frames	Detected	Accuracy (%)	Time (ms)
-90°		Normal	Neutral	324	324	100%	28.32 ms
-90°		Bright	Neutral	324	324	100%	26.85 ms
-90°		Normal	Smile	278	278	100%	29.23 ms
-90°		Bright	Smile	278	278	100%	25.01 ms
-75°		Normal	Neutral	300	300	100%	25.23 ms
-75°		Normal	Neutral	300	300	100%	25.23 ms
-75°		Bright	Neutral	300	300	100%	26.93 ms
-75°		Normal	Smile	323	323	100%	31.41 ms
-75°		Bright	Smile	323	323	100%	26.04 ms
-60°		Normal	Neutral	314	314	100%	34.28 ms
-60°		Bright	Neutral	314	314	100%	32.42 ms
-60°		Normal	Smile	327	327	100%	28.05 ms
-60°		Bright	Smile	327	327	100%	30.63 ms
-45°		Normal	Neutral	327	327	100%	35.44 ms
-45°		Bright	Neutral	327	327	100%	24.88 ms
-45°		Normal	Smile	332	332	100%	29.64 ms
-45°		Bright	Smile	332	332	100%	33.19 ms
-30°		Normal	Neutral	318	318	100%	33.14 ms
-30°		Bright	Neutral	318	318	100%	24.18 ms
-30°		Normal	Smile	302	302	100%	27.54 ms

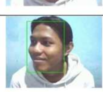
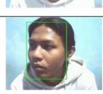
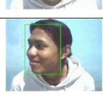
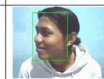
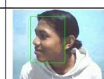
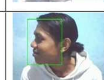
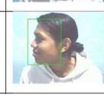
-30°		Bright	Smile	302	302	100%	28.36 ms
-15°		Normal	Neutral	326	326	100%	29.68 ms
-15°		Bright	Neutral	326	326	100%	31.60 ms
-15°		Normal	Smile	284	284	100%	26.08 ms
-15°		Bright	Smile	284	284	100%	26.42 ms
0°		Normal	Neutral	300	300	100%	28.79 ms
0°		Bright	Neutral	300	300	100%	27.09 ms
0°		Normal	Smile	308	308	100%	27.85 ms
0°		Bright	Smile	308	308	100%	29.82 ms
15°		Normal	Neutral	309	309	100%	30.15 ms
15°		Bright	Neutral	309	309	100%	29.78 ms
15°		Normal	Smile	315	315	100%	29.97 ms
15°		Bright	Smile	315	315	100%	27.11 ms
30°		Normal	Neutral	322	322	100%	29.72 ms
30°		Bright	Neutral	322	322	100%	34.53 ms
30°		Normal	Smile	296	286	100%	28.32 ms
30°		Bright	Smile	296	296	100%	26.14 ms
45°		Normal	Neutral	311	311	100%	33.34 ms
45°		Bright	Neutral	311	311	100%	30.94 ms
45°		Normal	Smile	299	299	100%	28.96 ms
45°		Bright	Smile	299	299	100%	30.04 ms
60°		Normal	Neutral	303	303	100%	32.65 ms
60°		Bright	Neutral	303	303	100%	33.23 ms
60°		Normal	Smile	309	309	100%	29.08 ms
60°		Bright	Smile	309	309	100%	33.63 ms
75°		Normal	Neutral	314	314	100%	31.15 ms
75°		Bright	Neutral	314	314	100%	27.02 ms
75°		Normal	Smile	299	299	100%	32.27 ms
75°		Bright	Smile	299	299	100%	31.56 ms
90°		Normal	Neutral	327	327	100%	27.04 ms
90°		Bright	Neutral	327	327	100%	25.44 ms
90°		Normal	Smile	325	325	100%	34.08 ms
90°		Bright	Smile	325	325	100%	34.51 ms

Table IV presents the results of testing the YOLO algorithm for face detection in video recordings across various combinations of viewing angles, lighting conditions (normal and bright), and facial expressions (neutral and smiling). Each video contained a fixed combination of angle, lighting, and expression, and the system was tested to detect faces in every frame of the video. The results show that YOLO consistently detected faces with 100% accuracy across all viewing angles from -90° to 90°, under both normal and bright lighting, and with both neutral and smiling expressions. The average detection time per frame ranged from 24.88 ms to 35.44 ms. These findings indicate that the YOLO algorithm delivers highly stable and efficient face detection performance under a wide range of simulated video conditions.

TABLE V
TESTING RESULTS OF HAAR CASCADE CLASSIFIER ON VIDEOS

Angle	Example Visual	Lighting	Expression	Total Frames	Detected	Accuracy (%)	Time (ms)
-90°		Normal	Neutral	324	294	90.74%	10.42 ms
-90°		Bright	Neutral	324	324	100%	13.50 ms

-90°	Normal	Smile	278	119	42.81%	12.26 ms
-90°	Bright	Smile	278	278	100%	15.13 ms
-75°	Normal	Neutral	300	300	100%	11.62 ms
-75°	Bright	Neutral	300	300	100%	14.15 ms
-75°	Normal	Smile	323	323	100%	14.66 ms
-75°	Bright	Smile	323	323	100%	17.53 ms
-60°	Normal	Neutral	314	314	100%	10.31 ms
-60°	Bright	Neutral	314	314	100%	14.31 ms
-60°	Normal	Smile	327	327	100%	16.50 ms
-60°	Bright	Smile	327	327	100%	15.77 ms
-45°	Normal	Neutral	327	327	100%	10.53 ms
-45°	Bright	Neutral	327	327	100%	15.78 ms
-45°	Normal	Smile	332	332	100%	17.48 ms
-45°	Bright	Smile	332	332	100%	16.53 ms
-30°	Normal	Neutral	318	318	100%	16.43 ms
-30°	Bright	Neutral	318	318	100%	24.57 ms
-30°	Normal	Smile	302	302	100%	23.32 ms
-30°	Bright	Smile	302	302	100%	21.72 ms
-15°	Normal	Neutral	326	326	100%	12.85 ms
-15°	Bright	Neutral	326	326	100%	19.21 ms
-15°	Normal	Smile	284	284	100%	21.30 ms

-15°	Bright	Smile	284	284	100%	16.11 ms
0°	Normal	Neutral	300	300	100%	14.26 ms
0°	Bright	Neutral	300	300	100%	18.14 ms
0°	Normal	Smile	308	308	100%	19.34 ms
0°	Bright	Smile	308	308	100%	21.96 ms
15°	Normal	Neutral	309	309	100%	14.68 ms
15°	Bright	Neutral	309	309	100%	15.98 ms
15°	Normal	Smile	315	315	100%	17.02 ms
15°	Bright	Smile	315	315	100%	17.52 ms
30°	Normal	Neutral	322	322	100%	21.41 ms
30°	Bright	Neutral	322	322	100%	24.47 ms
30°	Normal	Smile	296	296	100%	20.94 ms
30°	Bright	Smile	296	296	100%	13.07 ms
45°	Normal	Neutral	311	311	100%	19.33 ms
45°	Bright	Neutral	311	311	100%	22.79 ms
45°	Normal	Smile	299	299	100%	18.34 ms
45°	Bright	Smile	299	299	100%	16.85 ms
60°	Normal	Neutral	303	303	100%	21.95 ms
60°	Bright	Neutral	303	303	100%	18.18 ms
60°	Normal	Smile	309	309	100%	16.86 ms
60°	Bright	Smile	309	309	100%	16.03 ms

75°		Normal	Neutral	314	314	100%	21.61 ms
75°		Bright	Neutral	314	314	100%	22.45 ms
75°		Normal	Smile	299	299	100%	12.56 ms
75°		Bright	Smile	299	299	100%	13.39 ms
90°		Normal	Neutral	327	327	100%	17.42 ms
90°		Bright	Neutral	327	327	100%	19.78 ms
90°		Normal	Smile	325	325	100%	10.30 ms
90°		Bright	Smile	325	325	100%	10.41 ms

Table V presents the test results of the Haar Cascade Classifier in detecting faces from video footage under various viewing angles, lighting conditions (normal and bright), and facial expressions (neutral and smiling). Each video contains a fixed combination of these variables, and detection is performed on each frame. The results show varying performance depending on the conditions. At an angle of -90°, the accuracy is 90.74% under normal lighting with no expression and 42.81% with a smiling expression. At angles from -75° to 90°, the accuracy reaches 100% across all lighting and expression variations. The average detection time per frame ranges from 10.30 ms to 24.57 ms, indicating fast and efficient processing, although accuracy decreases under extreme angles and certain expressions.

C. *Testing in Real-Time*

The system was tested for real-time face detection over a duration of 10 seconds as the face gradually moved from -90° to 90°. Testing was conducted under two lighting conditions (normal and bright) and with two facial expressions (neutral and smiling). The goal was to assess the algorithm's ability and consistency in detecting faces during gradual changes in viewpoint, while also considering the effects of lighting and expressions. Detailed results are presented in Table VI.

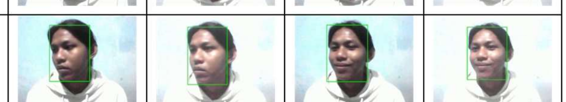
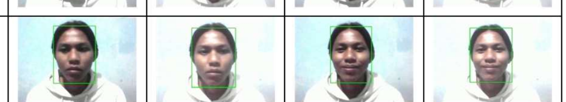
Based on Table VI, the results of real-time face detection testing over a 10-second duration using the YOLO and Haar Cascade Classifier algorithms are presented under different lighting conditions (normal and bright) and facial expressions (neutral and smiling). The total number of frames varied for each algorithm, depending on how many frames were successfully processed during testing.

YOLO consistently achieved 100% detection accuracy across all tested conditions, with an average detection time of 37.72 ms to 39.44 ms, demonstrating stable and efficient real-time performance. In comparison, the Haar Cascade Classifier reached 91.41% and 90.53% accuracy under normal lighting for neutral and smiling expressions, but achieved 100% under bright lighting. Its detection speed was faster, averaging 11.72–14.80 ms per frame, although with lower accuracy in certain conditions. Visual examples of the real-time tests for both algorithms are shown in Tables VII and VIII.

TABLE VI
TESTING RESULTS OF YOLO AND HAAR CASCADE IN REAL-TIME

Algorithm	Lighting	Expression	Total Frames	Detected	Accuracy (%)	Time (ms)
YOLO	Normal	Neutral	157	157	100%	37.72 ms
Haar Cascade Classifier	Normal	Neutral	256	234	91.41%	12.52 ms
YOLO	Normal	Smile	149	149	100%	38.84 ms
Haar Cascade Classifier	Normal	Smile	243	220	90.53%	13.27 ms
YOLO	Bright	Neutral	146	146	100%	39.44 ms
Haar Cascade Classifier	Bright	Neutral	268	268	100%	11.72 ms
YOLO	Bright	Smile	154	154	100%	38.02 ms
Haar Cascade Classifier	Bright	Smile	274	274	100%	14.80 ms

TABLE VII
VISUAL EXAMPLES OF YOLO IN REAL-TIME

Angle	Normal Lighting Neutral Expression, Bright Lighting Neutral Expression, Normal Lighting Smiling Expression, Bright Lighting Smiling Expression
-90°	
-75°	
-60°	
-45°	
-45°	
-30°	
-15°	
0°	
15°	
30°	
45°	

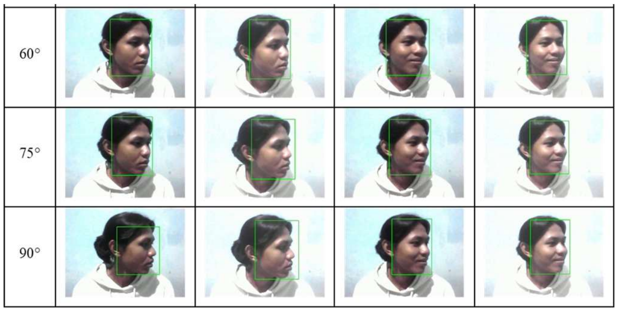
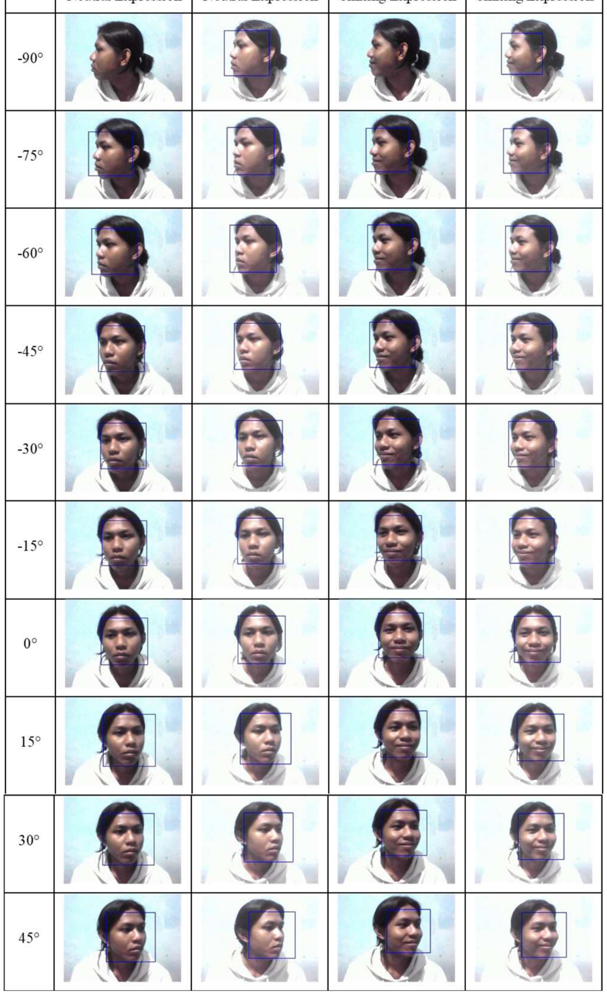


Table VII illustrates visual examples of real-time face detection using the YOLO algorithm across various combinations of facial angles, lighting conditions (normal and bright), and facial expressions (neutral and smiling). The visual results demonstrate that YOLO successfully detected faces with high accuracy under all tested conditions, including extreme angles such as -90° and 90°, under both lighting variations, and with both neutral and smiling expressions. This consistent detection performance highlights YOLO's robustness in handling diverse viewpoints and lighting conditions, which aligns with the quantitative results showing 100% accuracy across all tests.

TABLE VIII
VISUAL EXAMPLES OF HAAR CASCADE CLASSIFIER IN REAL-TIME

Angle	Normal Lighting Neutral Expression	Bright Lighting Neutral Expression	Normal Lighting Smiling Expression	Bright Lighting Smiling Expression



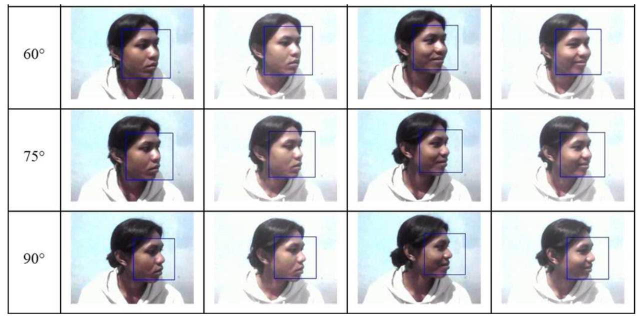


Table VIII presents visual examples of real-time face detection results using the Haar Cascade Classifier across different viewing angles, lighting conditions (normal and bright), and facial expressions (neutral and smiling). The visual results indicate that the Haar Cascade Classifier was able to detect faces fairly well at several viewing angles and under bright lighting, producing consistent bounding boxes in most of these conditions. However, its performance significantly declined at the extreme -90° angle for both neutral and smiling expressions, particularly under normal lighting, where the bounding box failed to appear. This limitation highlights the sensitivity of the Haar Cascade Classifier to extreme pose variations, especially when combined with less favorable lighting conditions.

D. *Comparitive Analysis*

Testing was conducted on three types of data, such as images, video recordings, and real-time, in order to obtain a comprehensive overview of the performance of the YOLO and Haar Cascade Classifier algorithms in face detection. Images were tested under varying angles, lighting, and expressions, while video frames were analyzed under consistent conditions. Real-time testing lasted 10 seconds with facial movements from -90° to 90°, assessing detection consistency under dynamic conditions.

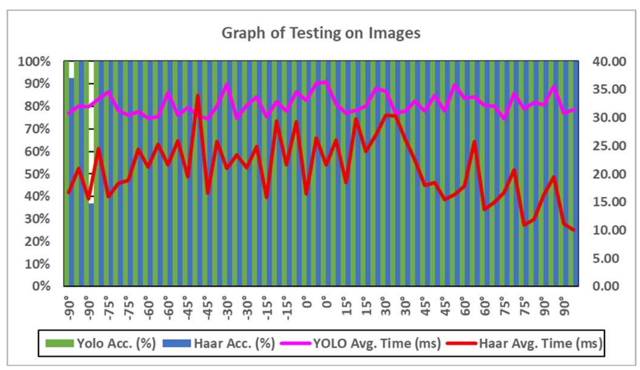


Fig. 9 Graph of Testing on Images

Based on the results shown in Fig. 9, testing on image data indicates that the YOLO algorithm achieved 100% accuracy across all viewing angles from –90° to 90°, under both normal and bright lighting conditions, and for facial expressions with neutral and smiling faces. The average detection time for YOLO ranged from 29 ms to 36 ms.

Meanwhile, the Haar Cascade Classifier showed decreased accuracy under certain conditions, particularly at –90° with normal lighting and a neutral expression, reaching 92.59%, and at –90° with normal lighting and a smiling expression, where accuracy dropped to 36.96%. Nevertheless, for most other angles and conditions, the Haar Cascade Classifier was able to achieve 100% accuracy, with a faster average detection time compared to YOLO, ranging from 15 ms to 26 ms.

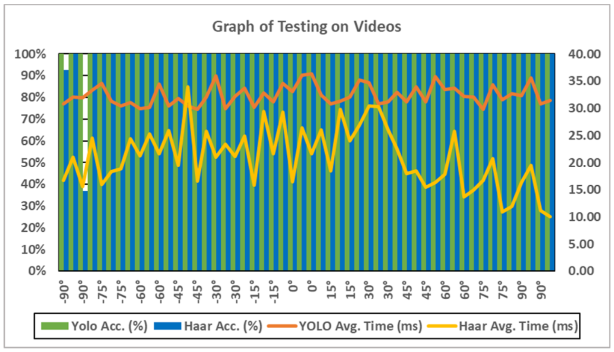


Fig. 10 Graph of Testing on Videos

Based on the results shown in Fig. 10, testing on video data indicates that the YOLO algorithm consistently detected faces with 100% accuracy across all viewing angles from –90° to 90°, under all lighting conditions (normal and bright) and facial expressions (neutral and smiling). The average detection time per frame ranged from 25 ms to 35 ms.

Meanwhile, the Haar Cascade Classifier showed decreased accuracy under certain conditions. At –90° with normal lighting and a neutral expression, accuracy reached only 90.74%, and under smiling expressions with normal lighting, accuracy dropped to 42.81%. However, for all other tested conditions, the Haar Cascade Classifier achieved 100% accuracy. Haar Cascade also demonstrated faster detection times, ranging from 10 ms to 24 ms.

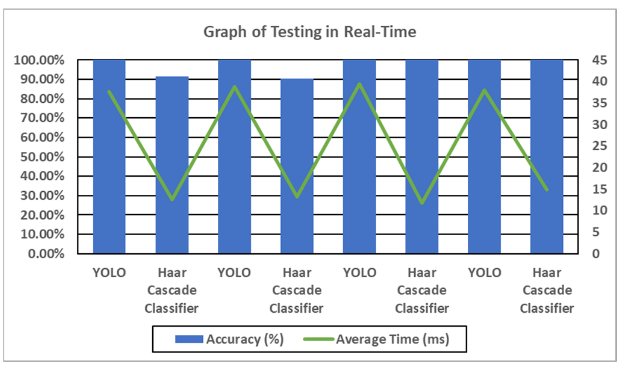


Fig. 11 Graph of Testing in Real-Time

Based on the results shown in Fig. 11, real-time testing indicates that the YOLO algorithm consistently detected faces with 100% accuracy across all test conditions, including normal and bright lighting, as well as neutral and smiling facial expressions. The average detection time per frame ranged from 37.72 ms to 39.44 ms.

Meanwhile, the Haar Cascade Classifier showed more variable accuracy, reaching 91.41% under normal lighting with a neutral expression and 90.53% under normal lighting with a smiling expression. However, under bright lighting, both neutral and smiling expressions achieved 100% accuracy. The average detection time for Haar Cascade was faster, ranging from 11.72 ms to 14.80 ms.

IV. Conclusion

Based on the conducted tests, the YOLO algorithm consistently achieved 100% accuracy in face detection across all viewing angles, lighting conditions (normal and bright), and facial expressions (neutral and smiling) for images, videos, and real-time streams. Meanwhile, the Haar Cascade Classifier exhibited decreased accuracy under specific conditions, particularly at extreme angles of –90°, with normal lighting and neutral or smiling expressions, while maintaining 100% accuracy under most other conditions. A comparison between the two algorithms indicates that YOLO outperforms Haar Cascade in terms of detection accuracy and consistency, whereas Haar Cascade demonstrates faster processing times. YOLO's detection time ranged from 30 to 39 ms, compared to 10 to 24 ms for Haar Cascade. Despite the slower processing, YOLO's superior accuracy makes it more suitable for applications where detection reliability is critical, whereas Haar Cascade may be preferred in systems prioritizing processing efficiency. For future research, this study can be extended by testing under a wider range of lighting conditions, diverse facial expressions, and larger datasets to obtain more representative results. Additionally, comparing YOLO and Haar Cascade with other recent face detection algorithms could provide a broader perspective on algorithm performance and relevance to modern system requirements.

References

[1] M. Hartomo, B. D., Hendrayudi, "Penerapan Computer Vision Untuk Absensi Wajah Berbasis Algoritma Cnn Pada Guru Smk Excellent 1 Tangerang," *J. Tek. Inform. Mahakarya JTIM*, vol. 4, no. 2, p. 70, 2021.

[2] L. Li, X. Mu, S. Li, and H. Peng, "A Review of Face Recognition Technology," *IEEE Access*, vol. 8, pp. 139110–139120, 2020, doi:10.1109/ACCESS.2020.3011028.

[3] Marzuman, Muharnis, Azizul, D. M. Rinaldi, and B. Prasetyo, "Perbandingan Metode Haar Cascade , YoloV3 , dan TinyYoloV3 Dalam Mendeteksi Kendaraan Bermotor Berbasis Video," *Proceeding Appl. Bus. Eng. Conf.*, no. November, pp. 17–19, 2022.

[4] Achmad Ubaidillah Ms, Achmad Fiqhi Ibadillah, and M. M. F. Nur, "Deteksi Jumlah Pengunjung Dan Penggunaan Masker Dengan Menggunakan Metode YOLO Dan Haar Cascade Classifier," *J. Zetroem*, vol. 5, no. 1, pp. 10–18, 2023, doi: 10.36526/ztr.v5i1.2564.

[5] D. Qi, W. Tan, Q. Yao, and J. Liu, "YOLO5Face: Why Reinventing a Face Detector," *Lect. Notes Comput. Sci.*, vol. 13805 LNCS, pp. 228–244, 2023, doi: 10.1007/978-3-031-25072-9_15.

[6] N. Zhao and Z. Qin, "A Survey and Test of the Face Detection Model Based on Yolov5," *Appl. Comput. Eng.*, vol. 2, no. 1, pp. 1026–1036, 2023, doi: 10.54254/2755-2721/2/20220620.

[7] M. Hussain, "YOLOv5, YOLOv8 and YOLOv10: The Go-To Detectors for Real-time Vision," pp. 1–12, 2024, [Online]. Available: http://arxiv.org/abs/2407.02988

[8] X. Ren, J. Tang, D. Yin, N. Chawla, and C. Huang, "A Survey of Large Language Models for Graphs," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 6616–6626, 2024, doi: 10.1145/3637528.3671460.

[9] C. A. Antipona, R. Magsino, R. Dioses, and K. Mata, "An Enhancement of Haar Cascade Algorithm Applied to Face Recognition for Gate Pass Security," *Ic-Itechs*, vol. 5, no. 1, pp. 1–9, 2024, doi: 10.32664/ic-itechs.v5i1.1500.

[10] M. N. Andrean *et al.*, "Comparing Haar Cascade and YOLOFACE for Region of Interest Classification in Drowsiness Detection," *J. Media Inform. Budidarma*, vol. 8, no. 1, p. 272, 2024, doi: 10.30865/mib.v8i1.7167.

[11] S. Minaee, P. Luo, Z. Lin, and K. Bowyer, "Going Deeper Into Face Detection: A Survey," pp. 1–17, 2021, [Online]. Available: http://arxiv.org/abs/2103.14983

[12] Y. Feng, S. Yu, H. Peng, Y. R. Li, and J. Zhang, "Detect Faces Efficiently: A Survey and Evaluations," *IEEE Trans. Biometrics, Behav. Identity Sci.*, vol. 4, no. 1, pp. 1–18, 2022, doi: 10.1109/TBIOM.2021.3120412.