EAI Endorsed Transactions
on AI and Robotics
Review Article **EAI.EU**

The Face Detection / Recognition, Perspective and Obstacles in Robotic: A Review

Nafiz Md Imtiaz Uddin[1], Ata Jahangir Moshayedi[1,*], Hong lan[1], Yang Shuxin[1]

[1]School of Information Engineering Jiangxi University of Science and Technology, No 86, Hongqi Ave, Ganzhou, Jiangxi,341000, China

Abstract

Facial recognition research is one of the different types of research in this world today. In recent years, facial recognition in robots has attracted increased study interest. Robotic platforms now utilize a variety of object detection methods, with face detection being a viable use. Face detection in robotics is a computer technique that recognizes human faces in digital pictures and is used in a range of applications. Different authors have performed their research in different ways on the use of detection systems. This paper aims to give future researchers a better idea of using facial recognition systems in robotics. In this study, we reviewed research by various authors over recent years to facilitate future facial recognition research. In addition, scholars have addressed the topics, how they have done so, and the specifics of their approaches are described. This paper reviewed an overview of hardware implementation and software implementation by various authors. It can automatically focus cameras or count the number of people who have entered a location. Commercial applications of the method include displaying tailored advertisements in response to a recognized face along with the algorithms, functions and architectures used in facial recognition and giving the opinions of various authors mentioned. The comparative analysis of facial recognition and its architecture system is highlighted.

Keywords: Face Detection and Recognition, Computer Vision, Detection Tracking, Human Robot, Service robot.

Received on 02 November 2022, accepted on 11 November 2022, published on 15 December 2022

doi: 10.4108/airo.v1i1.2836

Introduction

Machine vision as the broad umbral moved the robotic world into various applications. Face detection and recognition as one of the branches applied in the robotic arena. To follow this task the data collection phase uses a series of sensors or external interfaces to obtain images. At this stage, image processing techniques such as noise removal algorithms can be used to adjust the input**s** [1]. Face detection has advanced from simple computer vision approaches to more complex artificial neural networks (ANN) and associated technologies, leading to ongoing performance increases. Face tracking, face analysis, and facial recognition are just a few of the primary applications it currently serves as the foundation for. Sequential activities in the application are significantly impacted by face detection [2]. Face detection is the foundation for many other applications, including face tracking, face analysis, and face recognition. Face analysis uses facial detection to instruct the algorithms used to determine age, gender, and emotions based on facial expressions. Furthermore, face detection is necessary for facial recognition for the algorithms to choose which components of an image (or video) to use in order to create the face prints that are then compared with previously recorded (A training data system can also be used) face prints in order to establish if there is a match [3]. The most recent research study is focused on many elements of face detection and its usage in the robotic area. Face detection has lately been increasingly in demand in the robotic world. Face detection is the initial stage in many surveillance systems and automated face recognition systems. It has been extensively used in several sectors. A corner-verifying-based face detection technique is described. First, skin colour recognition swiftly filters out most of the backdrop

* Corresponding author. Email: ajm@jxust.edu.cn

region [4]. The necessity for intelligent systems to automatically comprehend and examine data has increased due to the exponential growth of video and picture databases since doing so manually is becoming more difficult. If there is a focus on a single domain, the faces of humans are among the most distinct items that may be identified in the photos. Face recognition is getting more difficult due to the growing number of apps that employ it [5]. In the detection system, image analysis and stone detection of the robotic system was performed within the pipe, which is a revolutionary way of robotic image detection [6]. Face detection serves as a crucial component in facial imaging applications like facial analysis and detection, and it offers consumers many benefits, including:

1. Enhanced Safety:

Face detection makes surveillance more effective and aids in finding terrorists and criminals. Since passwords and other sensitive information can't be stolen or altered, personal security is also increased [7].

2. Effortless Integration:

Technology for facial recognition and face detection is simple to incorporate, and the majority of available solutions are appropriate for use with most security programs.

3. Automatically Identification:

When detection was done manually in the past, it was often wrong and inefficient. Automated surveillance aims to identify persons on a watch list and is a significant application of interest [8].

The remainder of this paper is organized as follows: The second section introduces object detection; the third section is the face detection segment; the fourth section presents the face detection methods, strategies, and algorithms; the fifth section discusses the face detection methods and robotic applications; the sixth section contains the results and discussion, and the last section is the paper conclusion.

Section II: Object detection

The process of finding instances of objects of a certain class inside an image is known as object detection. These methods include pattern matching, colour recognition, feature recognition, and matching, which are discussed below.

1. Pattern matching:

Pattern matching is a relatively simple algorithm for matching image fragments to images. The way it works is that by moving the received image pattern and calculating the errors, the difference between the two image pixels is the location where these errors have the lowest value. It is possible to improve on this method and scan the image with a scaled pattern to find an object even if its size changes. Another possibility is to select a possible location of the object and search around that location to find it. As shown in Figure 1 illustrates the pyramid diagram that is produced by employing this technique and repeating this procedure [9].

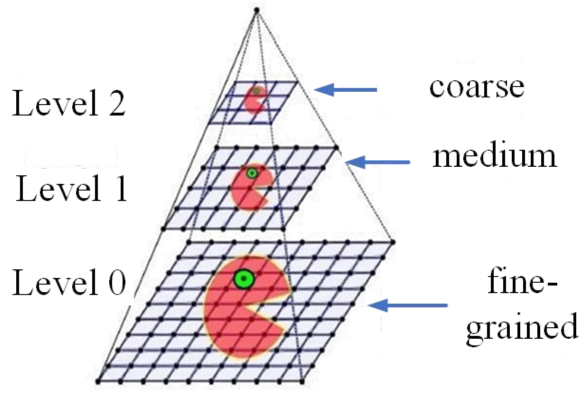

Figure 1. Image Pyramid Design [9]

As it is shown in Figure 1 This approach compares the sample picture with the provided image from the lowest resolution to the highest resolution.

2. Color recognition:

Colour tracking is one technique for tracking things. The image's pixels are examined in this manner until the required colour is discovered. The primary picture, which is the result of the filtering algorithm, is a binary image in which the desired colours are displayed as white and the undesirable colours are displayed as black. [10]. Among the disadvantages of the mentioned method, it can point out the sensitivity to changes in light. In this way, the detected colour changes with the change of the ambient light, and therefore, in practice, it is necessary to consider the tolerance for colour detection. In addition, if there is another colour in the environment with the same hue as the object, this will cause incorrect output from the robot, and therefore, it will cause problems in locating. The results of the technique, which can simultaneously identify several objects and distinguish the orange colour from other hues, are displayed in Figure 2 after it has been experimentally implemented using the OpenCV library [11].

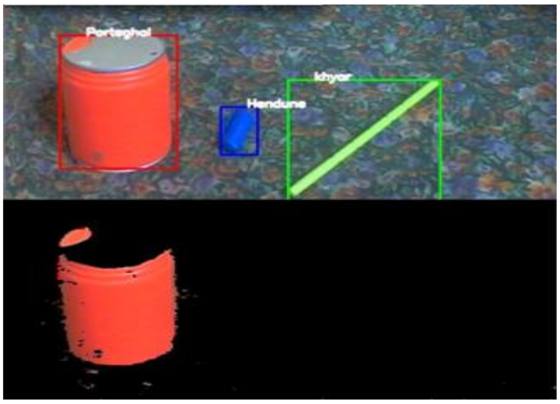

Figure 2. The colour detection algorithm for object detection**.**

3. Feature recognition and matching:

In image processing, a "feature" defines something that is noticeable, something by which information can be obtained from the image. For example, key points (corners), which are defined by the location of neighbouring points, are one of these features. Other examples are edges and lines.

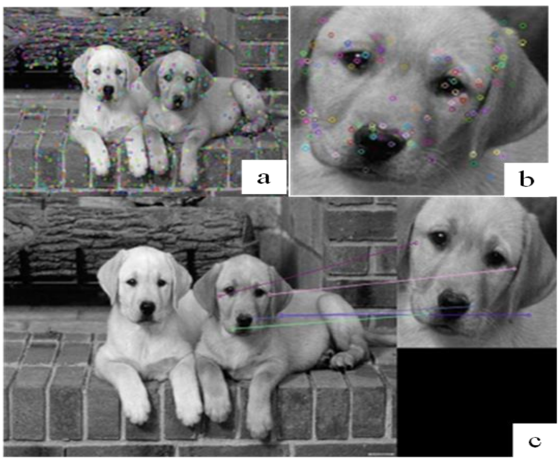

Figure 3: (a) and (b) show the features found in the image; (c) Correct matching of the characteristics of (a) and (b) [12].

Figure 3 shows the features of an image that were found by the feature detection algorithm. Initially, the required image can be considered, and a feature detection algorithm is applied to it. Then the algorithm tries to find the features it needs and matches them. This step is called feature description. At this stage, the algorithm tries to store the features that have been found in vector and compressed form. After that, a similar process is implemented on the second image to obtain the features of the second image. Finally, an algorithm is used to match the features of the two images. There are various methods for feature detection named fixed scale feature transformation (SIFTS), Robust accelerated SURF features, and Binary independent primary stable features (BRIEF) [12].

Section III: Face detection Segment

Three components make up the system's architecture: the database, the enrollment module (which scans and records a live thing's analogue or digital picture), and (An entity that handles compression, processing, and storage and also accounts for the comparison of the captured data with stored data), a module for Detection (this module interfaces with the application system). An architectural view is mentioned in Figure 4. A typical facial detection system includes user identification, a registration module, a Structural system, and a validation module. As shown, User identification records a face's analogue or digital picture. The collected sample is preprocessed and examined in the Registration module. For comparison in the future, this analyzed data is kept in the database (Structure of the system). The sample that was acquired is compressed and stored in the database. It should also have a retrieval property, which compares all the stored samples with the recently acquired sample and retrieves the matched sample for user verification and deciding if the stated match accuracy. A preprocessing system is another component of the verification module. The freshly acquired sample is preprocessed and contrasted in this validation module with the sample saved in the database. Depending on the database match, a decision is made.

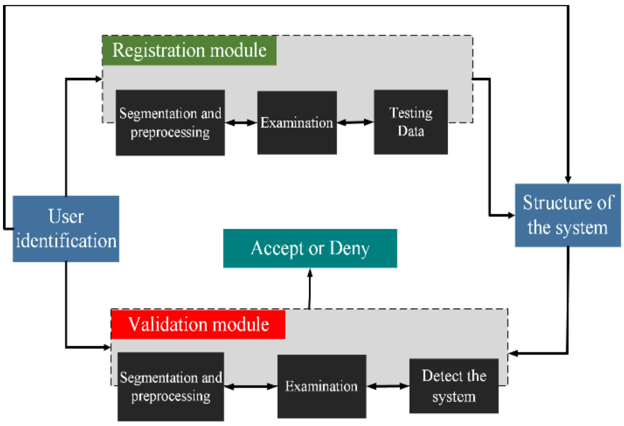

Figure 4. The general architecture of face detection system.

The sample is accordingly approved or disapproved. We may use an identity module instead of a Validation module. In this instance, the sample is compared to every other sample kept in the database. Simply put, it takes a picture as input, determines if a "face" is there, and determines its location on the image. During the detection and verification process, the system identifies the image with a reasonable degree of certainty and then determines if the probability is True or False. i.e., various classification approaches are used to compare the input vectors with the stored vectors in the database. One of the early approaches to characterizing the human face was the eigenface algorithm [13].

Section IV: Face detection methods Strategies and Algorithms

One of the study fields in computer vision and robotics that is expanding the fastest is face detection using vision, which is extensively employed in various human-related applications. However, it has been shown that normal lighting circumstances are the only ones where vision-based face detection is reliable. Both normal and harsh lighting conditions must be considered when creating an algorithm for face detection. One method is to transform photos of faces taken in diverse lighting situations into ones with an invariant look while keeping the texture and other face-specific traits [13]. The tracking module employs YOLO and F-RVO to extract detected items from subsequent frames and track humans. The Lightweight and Ground-Optimized Lidar Odometry and Mapping (LeGO-LOAM) technique, based on a 3D Lidar sensor, is used to map. Z. Chen et al. (2020) employed the Normal Distributions Transform (NDT) localization technique to locate the robot in the created map by comparing lidar data [14]. The important part is that Viola & Jones represented the face detection technological breakthrough. After much training, utilizing a cascade of "weak classifiers" with basic Haar features may provide stunning results. The face detection algorithm that uses this method the most is now. OpenCV comes with a straightforward implementation. A generic face model and potential instances of the item in the picture are compared using the Hausdorff distance. Also, it proved how well the detection algorithm worked on a database of 17000 photos from more than 2900 distinct subjects [15]. There are significant differences in the backdrop, lighting, and head size. The database's achieved detection rate is higher than 93% in their system [16]. Face detection data is necessary for the algorithms that determine which portions of an image or video are required to build a faceprint in a facial recognition system, which mathematically maps an individual's facial characteristics and saves the data as a faceprint. Once found, the new faceprint may be compared to previously collected faceprints to see whether there is a match. In general, considering techniques for face detection, different methods can be listed below:

1. The Viola-Jones method first locates the face on the grayscale image; then, it finds the location on the colored image. Although face real-time detection of faces was partially achieved, the accuracy of its detections is not very high. Since stiff things like chopsticks and cups cannot be handled by the block features in the Viola-Jones method, the face detection algorithm is more likely to identify fake faces when rigid objects are present in the face picture. This study proposes employing composite features based on the Viola-Jones method to address the problems above. They also suggest testing to show that this plan is effective [17].

2. The AdaBoost face detection algorithm detects faces together with skin colour, but certain non-faces also need to be found. After Harris corner detection, the corners of human faces are checked to minimize mistakes further. According to experimental findings, this technique can efficiently and correctly recognize faces via multi-feature fusion [17].

3. You only look once (YOLO): Target detection in a real-time context is appropriate for YOLO due to its quick detection speed. Its detection accuracy is higher, and its detection time is shorter than other target detection systems of a similar kind [18].

4. Haar cascade-like feature at a specific location in the detection window considers adjacent rectangular sectors, sums the pixel intensities in each, and then calculates the difference between these sums [19].

5. The Local binary patterns-based algorithm has employed a range of scale-spaces, cascade train techniques, and the hardware core of the system. The YOLOv2 and Haar-like commonly determine the face location in photos. They demonstrate how to program a humanoid robot to make face-to-face contact while anticipating its facial position [19].

6. Based on the **Kanade-Lucas-Tomasi (KLT) tracker and 2DPCA**, a system for face tracking and identification in a video series is suggested (Two-Dimensional Principal Component Analysis). All faces in the picture or video series are first detected using the Viola-Jones face detection technique before the KLT algorithm is used to track faces. KLT maintains long-term tracking when faces come in or out after being recognized in successive frames [20].

7. Convolutional neural network (CNN) cascade runs at various resolutions, swiftly discards background areas in the fast-low resolution stages, and carefully examines a limited number of demanding candidates in the final high-resolution stage.

The OpenCV classifier that uses Haar is much quicker. The video lags when switching to the MTCNN detector. Real-time operation is still possible, but the quality isn't as excellent. In every circumstance, Viola-Jones outperforms the KLT and has a higher detection rate [21].

A sophisticated object identification method utilized by CNNs, YOLO is often used in combination with other YOLO algorithms. YOLO v2.0's average accuracy is just roughly 16% greater than that of faster r-CNN [22].

Compared to Local Binary Patterns, Haar has higher accuracy. In testing, the Local Binary Pattern Histogram face recognition has a 93% accuracy rate with a confidence level of 2 to 5 [23]. Also, it is possible to identify persons and recognize faces, dynamic and static objects, environments, and custom items by using the Object Recognition and Classification (ORC) method. Depending on the extent and quality of the information, it may also be able to provide specifics like gender, age, or even mood. Thus, the ORC algorithm may be used in a security system or for autonomous navigation. For object detection, Robotech Vision employs its datasets [24].

An end-to-end algorithm called CrowdMove is utilized to move the robot toward the desired location after gathering all the data required for motion planning. The robot will begin to play a recorded verbal instruction to warn people to maintain an appropriate social distance if it senses its distance from the masses is less than 5 meters while approaching. Face detection may also automatically focus cameras or track how many people are in space. The technology may also be used in marketing; for instance, it can show specialized adverts when a particular face is detected [25].

Section V: Face detection methods and robotic application.

In recent years, the use of face-detection robot systems has increased significantly. It is one of the types of biometric techniques that refers to the automatic detection of faces by computerized systems by taking a look at faces. Robot systems are currently being used in a number of civilian contexts, including agriculture, 3D mapping, rescue operations, and even face identification technologies. The use of object detection systems of all these robotic platforms is increasing day by day. The face detection system first supports the face detection system in the vision robot. The aim notion of face tracking may be used to subsequently apply to smart equipment such as robots in order to pursue moving face targets and look for lost face targets [26]. In social interaction, a person's face is crucial for communicating their identity and sentiments. Humans don't have a very good capacity to distinguish between distinct faces compared to robots. Therefore, face detection, facial expression recognition, head posture estimation, human-computer interaction, etc., rely heavily on automated face detection systems [27].

The mathematical model that demonstrates the kinematic equations pertaining to differential drive robots was reviewed by A. J. Moshayedi et al. (2022). The robot must constantly know the translation and rotation matrix for autonomous navigation. The Instantaneous Centre of Curvature is a location on the common axis of the two driving wheels that the robot must revolve around as its speed changes [28].

A. Face Recognition Technique in Robotics

When utilizing a camera to identify faces, a face recognition robot uses an image processing technique. It is a technique for recognizing or confirming an individual's identification using their face—one of the most significant computer vision applications with significant commercial appeal. Deep learning-based algorithms have recently dramatically improved facial recognition technology [28] [29].

This model can be named a Model-Based Technique, holistic technique (appearance-based), Support Vector Device (SVM), and Artificial Neural Network (ANN), which are described as follows.

1. Model-Based Technique

Model-based face recognition techniques build a face model that can capture facial variations or highlights based on geometric connections between people, also known as geometry-based, template-based, or template-based techniques. There are two types of model-based techniques:3D statistical model of the form and texture of the human face is known as the **3D Morphable Technique (3DMM).** Instead of only modelling the look of 3D faces, it can also represent their essential qualities, such as form and skin texture. Widely used for generic face representation, image analysis, and image synthesis, 3DMM gathers data via a set of well-supervised 2D and 3D face scans and creates a mapping between a low-dimensional parameter space and a high-dimensional space of textured 3D models. And the second one is **Elastic Bunch Graph Matching (EBGM)** which locates a collection of landmark characteristics before matching the common structural similarities and features with other photos in the database to identify people. The same kind of network is used to represent each instance, from which a structure with nodes for local textures and variations is built. Only items with a shared structure—like faces in a frontal pose—and those with a common set of landmarks may be subject to EBGM [29].

2. Holistic Technique (appearance-based)

The appearance-based approaches, among the most effective and well-researched methods for face identification, employ pixel intensity values, which directly correlate to the brightness of light produced by the object and certain rays in space [30].

This method also contains four categories named Independent Component Analysis (ICA), Kernel Principal Component Analysis (KPCA), Linear Discriminant Analysis (LDA), and Principal Component Analysis (PCA) which are described as follows.

The **Independent Component Analysis (ICA)**: ICA offers a representation based on statistically independent variables, whereas PCA aims to represent the inputs based on uncorrelated variables. ICA, a generative model for the observed multivariate data, identifies underlying elements that govern collections of measurements, random variables, or signals. The primary use of ICA is to address the issue of blind signal separation. For the **Kernel Principal Component Analysis (KPCA**), a non-linear technique for identifying crucial characteristics or landmarks to take advantage of the complex spatial structure, the pictures are first translated from image space into a higher-dimensional feature space using a kernel function. The last one is **Linear Discriminant Analysis (LDA)**. Also known as Fisherface, LDA is a dimensionality-reduction method based on appearance. It minimizes a dataset's number of variables (or dimensions) while preserving most class-specific discriminating data. Employing more than one prepared image for a single lesson uses a regulating learning method. The **Principal Component Analysis (PCA)** is a feature extraction approach that keeps a face's most useful and expressive characteristics while discarding the "least important" variables. It attempts to build a computer model that best reflects the face by minimizing the number of variables and bringing out key variables (eigenfeatures like eyes, noses, lips, cheeks, etc.) that best reflect the face [31] [32].

3. Support Vector Device (SVM)

One of the best methods for solving classification issues is Support Vector Machines (SVM). Inferentially, given a collection of points from two classes, an SVM determines the hyperplane that minimizes the distance from either class to the hyperplane while separating the greatest percentage of points from the same class on the same side. This hyperplane, known as the Optimal Separating Hyperplane (OSH), reduces the possibility of misclassifying both the test set's hidden instances and those in the training set [33] [34].

4. Artificial Neural Network (ANN)

Different ANN designs and models have been employed in recent years for face identification and recognition because humans can mimic how neurons in the human brain function. Some of the most effective examples of face recognition using ANN are: Convolutional Neural Networks (CNN), Evolutionary Optimization of Neural Networks, Multilayer Perceptron (MLP), Back Propagation Neural Networks (BPNN), Gabor Wavelet Faces with ANN, Skin Color and BPNN, etc. [35].

Automatic machine face recognition offers various benefits over human face recognition. Since a computer never gets tired, it can keep more face pictures in a gallery of well-known people's faces, remembering more faces and identifying each one more quickly than a person. This has made applications for security, surveillance, medical care, and other things possible.

Concerning the mentioned method, various authors have shown different approaches to the robotic field for face detection and recognition task. As the reviewed paper shown, *M. Turk et al. (1991)* produced a computer system to follow and locate a subject's head and then identify the subject by comparing the face's features to those of well-known people. They built a six-level Gaussian pyramid for each picture based on the results of their experiment, resulting in a resolution drop from 512 X 512 pixels to 16 x 16 pixels. Their approach produced an average classification accuracy of 96% when accounting for illumination variance, 85% when accounting for orientation variation, and 64% when accounting for size variation [36]. In a picture, *S.-H. Jeon et al. (2016)* proposed reconstructing a frontal face from a non-frontal face. This study aims to demonstrate that the presented approach for the face recognition system is resistant to changes in head attitude. Additionally, they contrast the suggested strategy with a straightforward approach that rotates in-plane utilizing eye locations. They use SIFT and Support Vector Machine as the foundation of their gender classifier. According to their trial, the new strategy is more accurate than the existing one (77% vs 74%). When utilizing an Intel i7 3.5GHz CPU, the processing time is less than 300ms [37]. *Muhtadin et al. (2016)* report research results using interactive communication between humans and robots. In this procedure, the operator's voice acts as the system input. The system will start to pick up on nearby noises as soon as it picks up a human face. The system will track the discovered human face if it determines that there are noises nearby. The system that resulted could recognize faces constantly at angles between 0° and 30°, identify human voice with an accuracy of 80.45%, and respond to human speech with a correct response of 47.58% [38]. In another proposal research work, A. S. Yasaswini et al. (2017) described a system consisting of three modules: robot movement, facial expression detection, and face detection. The H-operation bridge in the hardware portion is based on the L293 Motor driver.

The L293D dual H-bridge motor driver consists of two H-bridges and two clockwise counters lock-wise-controllable DC motors that may be interfaced. In their paper, they took the 640x320 RGB-coloured images for this study using a 2MP Microsoft VX-3000 web camera. This system's main benefit is that it is comfortable for those with physical limitations and doesn't need bodily movement [39]. *M. D. Putro et al. (2018)* suggest real-time human face tracking for application in human-robot interaction. The proposed hardware was Intel Core I5-6600 CPU @ 3.30 GHz, 8 GB RAM. The proposed technique can preserve a face's ability to remain in the centre of the frame area and follow a single face. While using accessories and under harsh situations, this system will always track and maintain the face in the centre of the picture. According to their testing findings, a speed of 28.32 frames per second was the maximum speed [40]. *T.M.W. Vithanawasam et al. (2018)* recognized face and upper-body emotions with the RGB-D (Kinect) sensor for service robots because it offers robust upper-body detection. 1000 photos were used in the dataset's creation, with each image representing a different emotion. Initially, 20,000 pictures were shot of the subject as they spontaneously expressed their feelings. While the subject communicated their feelings, the coordinates of each joint were taken. 33 characteristics were added to a dataset comprising 6400 samples. The 11 joints in the upper body's 33 characteristics corresponded to their X, Y, and Z coordinates. The sensor records 30 frames per second at a frame rate of 30 samples per second. According to the findings of facial emotion detection, the percentages of anger, fear, boredom, and neutrality were 83.3%, 79.5%, 82.1%, and 84.2%, respectively [41].

Y. Okafuji et al. (2019) proposed a contact method for robots to interact face-to-face that may reduce delays. They used four techniques to track the movement and positioning of people: optical flow, YOLOv2 for face and person recognition, and Haar-like face detection [42].

W.-y. LU et al. (2019) suggest using composite features based on the Viola-Jones algorithm to enhance the block characteristics of stiff objects and demonstrate the viability of this approach via tests. From the FDDB data collection, 100 random photos were chosen for the experiment's face detection. Additionally, the number of faces that were genuinely included in the 100 photos and the number of missing and erroneous detections. Based on this, a total of 10 experiments were carried out, using a total of 1000 images as experimental data. By comparing the composite feature approach based on the Viola-Jones algorithm with the original Viola-Jones algorithm in testing, the textual strategy used in the face recognition process is demonstrated to be accurate [43].

P. B. Nithin et al. (2019) demonstrated the efficacy of well-known face recognition techniques on an offline computer. Recommended in hardware was Intel Core I7-930 CPU @ 3.01 GHz, 8GB RAM. The hardware system's primary purpose is to locate a person's face in the actual world and track its location. According to their findings, compared to existing feature extractor techniques, the deep neural network-based face detection algorithm has higher accuracy (97.5%). Their paper observed that LBPH underperformed in low illumination situations [44].

P. Chakraborty et al. (2019) developed a human-robot interaction system by estimating the visual focus of the human attention level. The requisite real-time picture frames from the robotic interfaces and sensors were taken from their research. After extracting the eye and head from the picture frames, the gaze was estimated using the eyeball's movement. The assumption was that the person had no attention if he was always seeing something. The suggested ML approaches' validation and test accuracy were close to 99.24% and 99.43%, respectively [45].

J. Heredia et al. (2021) proposed an adaptive multimodal emotion detection architecture for social robots. By considering various analytical techniques, they want to enhance previous research findings while preserving the findings' relevance to social robots and using those robots' built-in sensors. They also demonstrated how, in the context of social robots, our system is adaptable and versatile at every level [46]. The objective of moon discovery by A. J. Moshayedi et al. (2021) on the fixed dataset is to identify the moon as the primary target in the picture. The robotic platform has a novel method for examining various lunar images. The form identification method is assisted in finding the object and revealing the moon phases by the SAZ algorithm [47]. Using Nvidia Jetson NX and the GuartBot, S. Manzoor et al. (2022) demonstrated a two-stage face mask recognition system tailored for real-time inference. Based on the model's performance in terms of frames per second, execution time, and pictures per second, their research performed a number of studies. The CNN model for classification-based face mask recognition outperforms MobileNet-v2, Xception, and InceptionNet-v3 in their experiment, achieving 94.5%, 95.9%, and 94.28% accuracy on training, validation, and testing datasets, respectively [48].The authors propose the vision task for the robot platforms for face recognition can be into four categories and structures face Detection in a Robotic Arm Based System, Vehicle System, Humanoid Robotic System and Other Service Robotic Platforms, which are described as follows: D. Banerjee et al. (2018) developed a robotic arm (Figure 6 a) by using automated facial recognition software testing. A key component of this test automation scenario is the robotic arm. They used the software WINCAPS 3 to program the robotic arm. A robotic arm was used to construct an automated face recognition test suite as part of their discussion of the computer vision use case for facial recognition. The face detection rate is still high at 20°–30° face yaw angles, reaching around 91%, but the true positive rate falls to 69% [49]. In the other research, H. Zhang et al. (2019) proposed a six-degree of freedom tracking method (Figure 6 b) that attaches a face detector with a camera to the robot's wrist to get real-time information on face depth and attitude. The AUBO i5 robot, a computer, and a monocular CMOS camera module make up most of the proposed six-DOF visual servo face tracking system's hardware. The core robot functions utilized to move the robot are encapsulated in the AUBO SDK. The computer is equipped with an 8-core Intel(R) Core (TM) i5-83000H processor clocked at 2.30 GHz, an NVIDIA GeForce GTX 1050 graphics card, and 8192MB of RAM. It is a Lenovo Legion Y7000. Processing images and data on face position and posture are done using the Open Face toolkit 2.2.0. To conduct matrix operations, utilize the Eigen 3.3.9 package. The camera utilized in this work has a frame rate of 29±2 FPS [50]. The second approach is to use Face Detection with Vehicle System. M. Karahan et al. (2020) presented a camera on the Quadrotor UAV as part of a human face identification and tracking system (Figure 6 c). The Tello Edu Drone's camera is used in this research to take pictures and capture video while in the air. The Tello Edu Drone delivers pictures and videos to the computer while it is wirelessly attached. The MATLAB system examines photos and videos and finds human faces. An outline of a rectangle is drawn around the human face when this algorithm locates it. A face-tracking algorithm called Kanade-Lucas-Tomasi (KLT) is used. The KLT method locates feature locations that can be reliably monitored. According to their findings, the tracking algorithm could recognize facial characteristics and follow the facial motions of the face, while the face detection algorithm correctly recognizes faces [51]. The approach presented by C. Iaboni et al. (2021) used event cameras (Figure 6 d) to recognize and track several mobile ground robots in real-time. The software program in C++ and Python took care of data collecting, detection and tracking, and real-time data display. In studies involving up to three robots, the experimental findings demonstrated 100% detection and tracking fidelity (as high as 93% for four robots) in the presence of event camera noise and robot stops. There was a gradual decline in the quality of detection and tracking as the lighting conditions changed [52]. Face Detection in Humanoid Robotic Systems as the third target reported in research works.Using design strategies like variable image scaling and simultaneous processing of many classifiers without integral picture creation, S. S. Lee et al. (2016) created a fully pipelined architecture (Figure 6 f) that was implemented on the FPGA platform. The suggested test bed is built on the Xilinx FPGA platform and a human-like robot. According to their paper, the design enables real-time Face Detection (FD) processing for a VGA video at 30 frames per second [53]. *D.S. Pamungkas et al. (2018*) developed a head robot with a camera (Figure 6 e) that uses a local binary pattern histogram approach to identify faces using the face identification algorithm Haar Cascade. The major hardware they suggested was the HD Webcam C310, servo motors, and Arduino Uno. And the suggested system leverages libraries from OpenCV. Each participant in the exam has their data collected 20 times from different angles and stances. The samples are then kept in a folder marked with the individual's unique number. The number is used to identify the person's facial sample. According to their test findings, people are then positioned between 50 and 150 cm from the camera to capture their complete faces [54]. *B. Cilmi et al. (2018)* developed a humanoid robot that is proficient at accurately and dependably recognizing, tracking, and viewing a human face (Figure 6 g). This project aims to construct the required humanoid robot using the Kanade-Lucas-Tomasi feature tracker and the Viola-Jones face detection framework [55]. For robot-human intercommunication, *S. Sharmin et al. (2021)* created a duplex eye contact system that fulfils both requirements. They created a robotic head (as a platform) on which to perform tests on human-robot communication, using this platform to put our theoretical idea of duplex eye contact into practice (Figure 6 h). This study suggests a conceptual model of a duplex eye contact mechanism considering two cases: human initiative and robot initiative. The outcomes of the robotic framework in a scenario using the suggested duplex eye contact mechanism are reported. The findings reveal that, for the human initiative case and the robot initiative example, the suggested strategy made eye contact with 92% and 86% accuracy [56].

As the last group, Service Robot Platforms, *G. -S. J. Hsu et al. (2016)* developed a camera robot (Figure 6 i) that includes a multiview face detector to capture faces in various poses. A battery, an Intel Core i5-3230M 2.6GHz processor, 4GB of RAM, and the Toshiba PORTEGE R930 were all recommended features. Also linked to the platform through the Toshiba PORTEGE R930 are two cameras: a Microsoft Kinect RGB-D camera and a Logitech C920 webcam. 15 million pixels and 1080p resolution are available on the Logitech webcam for shooting images. Furthermore, they mentioned that the standard separation for a single face is 0.8 meters [57].

W. Jiang et al. (2017) proposed deep convolutional neural networks for home service robots based on facial detection and recognition software (Figure 6 j). From their paper, the face detection network is pre-trained over the challenging WIDER dataset, which contains 159,424 training faces. An 85.7% true positive rate (with 50 false positives in the test set) may still be attained with as little as 100 ideas. As a compromise between accuracy and performance, we employ 100 ideas in the final system, and their unoptimized code operates at around 5 frames per second. When employing the fixed centre crop, the experiments show a classification accuracy of 976.1% ± 0.23 [58]. Combining Single-Shot Detection, FaceNet, and Kernelized Correlation Filter, *C. -L. Hwang et al. (2020)* created a deep learning technique (SSD-FN-KCF). The proposed hardware was three dc servomotors, motion control, one laptop, RGB-D camera (Figure 6 k). The image-based adaptive finite-time hierarchical constraint control achieves the necessary posture for seeking or tracking (a particular) person depending on the image processing results. In their investigation, the resolution of 320 x 240 allowed for a human detection rate of 98.6% and a frame rate above 30 fps, regardless of the person and camera's distance (8m) or view angle equal to 180 [59].

Section VI: Result and Discussion

The information stage is the most important part of a machine vision system. At this stage, the information collected is transformed into a structure that is more useful to the target system. From this stage, information that is not important to the system can be removed, and information that is important to the system can be emphasized. A good representation of the data at this stage makes the task much easier at later stages, and consequently, a lot of time is spent on the design of the machine vision system for the correct definition of the appropriate representation. At the decision-making step, a choice should be made regarding the output of the system using the methods that are accessible.

These phases necessitate the use of machine learning algorithms and pattern recognition for difficult tasks like object recognition, face recognition, or picture alignment. The Object detection step are shown in Figure 5.

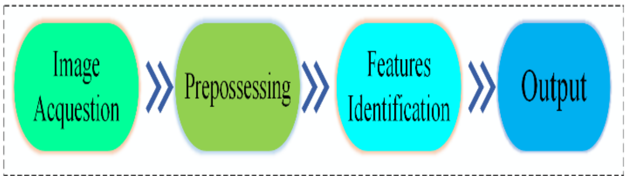

Figure 5. The process of detecting objects from the input image.

According to Figure 5, after the image acquisition section the images are noised, the background is removed, or any action that speeds up the processing in finding the desired area is indicated perform in the pre-processing step. After that, the image features are extracted and specified, and then the output can be displayed [10-13] [60]. In general, a methodical split of the image is the first step in resolving the issue of how identifying items in a picture or video. An area can first be identified by capturing an item, after which the tool will apply an algorithm to the input. Prepossessing then starts in accordance with its requirements. Following that, the computer would provide a variety of item suggestions based on the settings. The next phases of detection involve categorizing objects using models, applying features, and returning the classes and places within the framework of ultimately approved proposals. The next phases of detection involve categorizing objects using models, applying features, and returning the classes and places within the framework of ultimately approved proposals. Through such processing, the desired output is obtained. In a surveillance system, video cameras are placed in fixed locations that send their images to a specific point, while in flying robots, the camera is mobile, has no fixed location, and has six degrees of freedom. This makes it difficult to detect, locate, and track objects, which may be caused by one of the following factors: Loss of information due to the conversion of 3D images into 2D; Presence of noise in the image; Presence of non-rigid objects in the image; The presence of partial or total obstacles to seeing the object; Objects with complex shapes; Changing the light of the image; Need for simultaneous processing [1] [60].

A computer method known as face detection makes it feasible to identify faces in digital photographs. In order to improve human-robot interaction in service robots, people recognition is a crucial component. Various methods for detecting faces exist, but they presuppose parameters impractical for service robots, such as a picture with a centred face and regulated lighting [61].

One of the most popular uses of computer vision is face detection. It is a crucial issue in pattern recognition and computer vision. Several facial feature detection techniques have been developed in the previous ten years. Recently, convolutional neural networks (CNN) and deep learning have successfully powered extremely accurate face detection algorithms.

The viola-Jones algorithm is an object recognition framework that allows the detection of human faces. Viola-Jones algorithm is robust, powerful, and faster despite being outdated. The KLT algorithm tracks a set of feature points across the video frames. Once the detection locates the face, the next step is identifying feature points that can be reliably tracked. To distinguish between positive (i.e., photographs with a face) and negative (i.e., images without a face), face detection methods are used (i.e. images without a look). The algorithms must be trained on enormous datasets, including thousands of face photos and non-facial images, to be able to do this task properly. Another approach uses object recognition, bounding box prediction, feature extraction, and sparse feature matching as part of its tracking mechanism. With the aid of the motion modelling technique F-RVO, they employed YOLO to identify pedestrians and updated the traces of pedestrians by matching sparse characteristics. Table 1 shows the Face detection methods' merit and demerit in robotics.

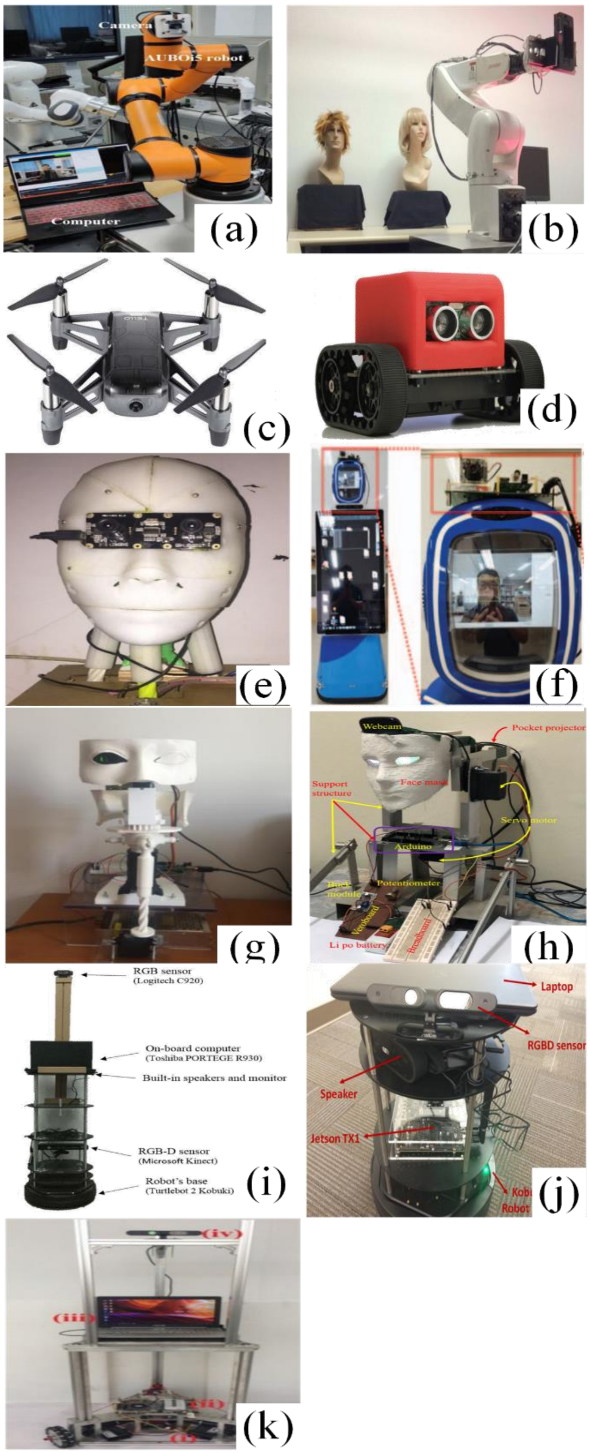

Figure 6. Robotic Arm-Based in Face Detection. (a) Six-DOF Face Tracking System [49], (b) Face Recognition Testing Automation [48], Vehicle System in Face Detection. (c) Tello Edu Quadrotor UAV (quadrotor DJI Ryze) [50] (d) Ground Robot [51]. Humanoid Robot Interaction in Face Detection. (e) Robot head [53], (f) Based on Xilinx Virtex-6 FPGA [52], (g) Humanoid robot neck [54], (h) Robotic head prototype with hardware configurations

[55] Service Robotic Platforms. (i) Photographer Robot "Fotor" [56], (j) Based on Jetson TX1 [57], (k) Omnidirectional Mobile Robot [58].

Table 1. The face detection methods Via robotics application

Ref	Aim and Method	Advantage	Disadvantage	Application (platform/task)	Accuracy
[2]	Face Detection, Tracking, and Recognition with AI	Computer vision-assisted detection and tracking of masks worn by individuals	The eigenface approach was unable to identify the mask-wearers.	The mobile platform is used for the face detection system	With reasonable CPU frame rates of 5–10 FPS and 50 FPS, respectively
[3]	A humanoid robot with facial expression control	They are very comfortable for those with physical limitations and do not call for any physical movement.	The robot will be controlled by a single customer, who will provide all commands (wrong identification).	Conduct a variety of tasks in human civilization	-
[34]	Execution of Human-Robot Interaction	Using external memory to lower hardware complexity	DDR3 is used as external memory	Universal Asynchronous Receiver/Transmitter (UART) interface is supported for sending the Face Detection system	The processing speed of 30 frames per second.
[41]	The RGB-D captures facial and upper-body expressions	When facing the Kinect sensor while sitting, standing, or moving, all emotions were properly caught in both ways.	slight variation in emotion for some cases	service robots	According to training data, Emotion of anger83.3%, fear79.5%, bored 82.1 %, neutral was 84.2%
[43]	Feature application of the Viola-Jones algorithm	Utilized a lower missed detection rate and erroneous detection rate than the VJ algorithm.	The issue with the Viola-Jones method is that stiff objects reduce the recognition rate.	-	This provides a solid framework for guaranteeing face recognition's correctness.
[48]	Viola Jones, KLT algorithms, CNN, Face Detection algorithms	It can detect single or multiple persons.	The static pictures produced by the Tello are shaky, soft, and monotonous in	Unmanned Aerial Vehicles (UAV) Autonomous Detection System	-
[49]	Home Service Robot end-to-end process	To make a unified network, an efficient training technique is employed.	In complex situations, the algorithm overlooked certain faces.	End-to-end deep neural networks	Used the fixed centre crop, accuracy was 97.61%±0.23
[50]	Real-Time Face Tracking Human Robot	Both the servo motor and the vision algorithm operate dependably.	Not all parts could be drilled or bonded, which reduced the system's modularity.	Using an analogue input moves the robotic neck platform	90% detection for the AFW dataset and a frame rate of over10 frames per second with GPU acceleration
[52]	Using event cameras to detect and track moving indoor ground robots	Without any training, they were able to execute reliable detection and tracking.	The study will be expanded to include the detection of several quadrotors.	Pencil balancing, industrial robotics, and a dual copter platform for 1D attitude control	The testing results demonstrated 100% fidelity in detection and tracking.

As table 1 shown, the implemented method on the robot for face detection is compared in terms of pros, cons and accuracy. Here Viola-Jones algorithm is most used in face detection, where real-time face tracking of humanoid robotics is more advantageous. Additionally, an accuracy technique, event cameras have shown 100% fidelity in identifying and following moving interior ground robots. The incapability of Eigenface to recognize people wearing masks was its main defect. The methodology of Face Detection & Recognition in Robotics is mentioned in Figure 7.

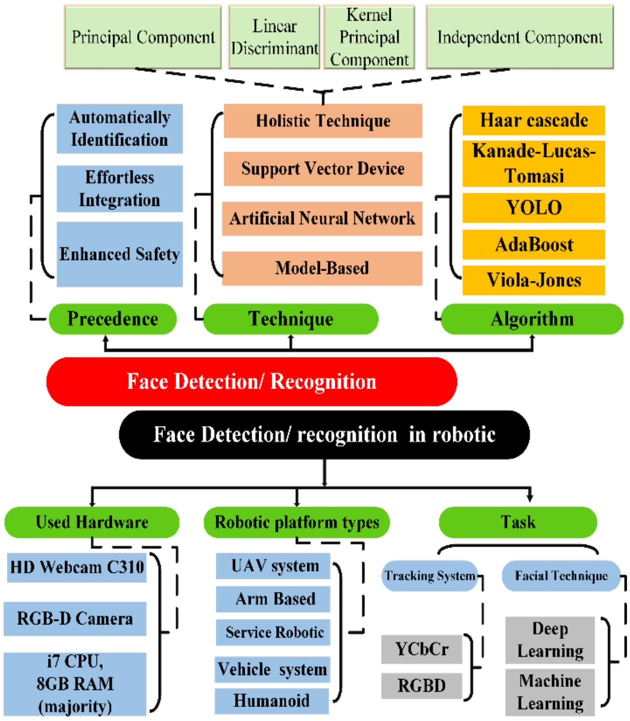

Figure 7. Methodology of Face Detection & Recognition in Robotic

As shown in Figure 7, face detection and recognition in general and in robotics in terms of algorithm, approach, and precedence are summarized. Many benefits are available through this system. The use of facial recognition techniques is significantly more effective than biometric and iris recognition, and these systems exist originally investigated in security systems to recognize and compare human faces. The algorithms concentrate on finding frontal human faces. It is comparable to image detection, which matches a picture of a person bit by bit. Robotics need computer vision as a fundamental component. Moreover, facial recognition and detection have improved because of advancements made by artificial intelligence technologies.

Section VII: CONCLUSION

Detection is known to find where and how many faces are present in an image or sight. To do this, one can search for faces at every angle, shape and position. A computer program called face detection uses artificial intelligence to discover and recognize human faces in digital images and videos. It may be thought of as a specific instance of object-class detection, where the objective is to identify all the objects that are a part of a particular class and determine their positions and sizes. The introductory purpose of incorporating computer vision into artificial intelligence is to develop a model of visual perception that can see a situation, carry out a job without human assistance, and make appropriate judgments. The whole procedure includes gathering datasets, processing, analysing, and comprehending digital pictures to use them in a practical setting. The majority of the components, as listed in this paper, along with the application, method, and hardware equipment that the author utilized, show that, as the hardware implementation part, most of the research has been done using an i7 CPU and 8 GB of RAM in this case. Based on all of our assessments, it's determined that Viola-Jones is the most commonly used algorithm for face identification because it is used to identify and detect things like this human face. In future work, it planned to compare and implement the various face detection method on a platform and compare them via the same application.

REFERENCES

[1] Rout, R. Kumar. "A survey on object detection and tracking algorithms." PhD diss., 2013.

[2] S. Ç. Tekkök, M. E. Söyünmez, B. Bostancı and P. O. Ekim, "Face Detection, Tracking and Recognition with Artificial Intelligence," 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2021, pp. 1-5, DOI: 10.1109/HORA52670.2021.9461356.

[3] A. J. Moshayedi, A. S. Khan, S. Yang and S. M. Zanjani, "Personal Image Classifier Based Handy Pipe Defect Recognizer (HPD): Design and Test," 2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP), 2022, pp. 1721-1728, DOI: 10.1109/ICSP54964.2022.9778676.

[4] K. Dang and S. Sharma, "Review and comparison of face detection algorithms," 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, 2017, pp. 629-633, DOI: 10.1109/CONFLUENCE.2017.7943228.

[5] Q. Zhao and S. Zhang, "A face detection method based on corner verifying," 2011 International Conference on Computer Science and Service System (CSSS), 2011, pp. 2854-2857, DOI: 10.1109/CSSS.2011.5974784.

[6] A. J. Moshayedi, & Fard, Saeed & Liao, Liefa & Eftekhari, S. Ali. (2019). Design and Development of Pipe Inspection Robot Meant for Resizable Pipe Lines. International Journal of Robotics and Control. 2. 25. 10.5430/ijrc.v2n1p25.

[7] J. J. Patoliya and M. M. Desai, "Face detection based ATM security system using embedded Linux platform," 2017 2nd International Conference for Convergence in Technology (I2CT), 2017, pp. 74-78, DOI: 10.1109/I2CT.2017.8226097.

[8] P. Apoorva., H. C. Impana., S. L. Siri., M. R. Varshitha. and B. Ramesh., "Automated Criminal Detection by Face Recognition using Open Computer Vision Classifiers," 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), 2019, pp. 775-778, DOI: 10.1109/ICCMC.2019.8819850.

[9] R. Barták and A. Vykovský, "Any Object Tracking and Following by a Flying Drone," 2015 Fourteenth Mexican International Conference on Artificial Intelligence (MICAI), 2015, pp. 35-41, DOI: 10.1109/MICAI.2015.12.

[10] J. Y. Choi, S. G. Kim. "Collaborative Tracking Control of UAV-UGV." World Academy of Science, Engineering and Technology, International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering 6 (2012): 2487-2490.

[11] X. Farhodov, O. -H. Kwon, K. W. Kang, S. -H. Lee and K. -R. Kwon, "Faster RCNN Detection Based OpenCV CSRT Tracker Using Drone Data," 2019 International Conference on Information Science and Communications Technologies (ICISCT), 2019, pp. 1-3, doi: 10.1109/ICISCT47635.2019.9012043.

[12] C. Cruz, L. E. Sucar and E. F. Morales, "Real-time face recognition for human-robot interaction," 2008 8th IEEE International Conference on Automatic Face & Gesture Recognition, 2008, pp. 1-6, DOI: 10.1109/AFGR.2008.4813386.

[13] X. Wang and X. Tang, "Face Photo-Sketch Synthesis and Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 11, pp. 1955-1967, Nov. 2009, doi: 10.1109/TPAMI.2008.222.

[14] Z. Chen et al., "Autonomous Social Distancing in Urban Environments Using a Quadruped Robot," in IEEE Access, vol. 9, pp. 8392-8403, 2021, DOI: 10.1109/ACCESS.2021.3049426.

[15] Fröba, B., Külbeck, C. (2001). Real-Time Face Detection Using Edge-Orientation Matching. In: Bigun, J., Smeraldi, F. (eds) Audio- and Video-Based Biometric Person Authentication. AVBPA 2001. Lecture Notes in Computer Science, vol 2091. Springer, Berlin, Heidelberg. https://DOI.org/10.1007/3-540-45344-X_12

[16] Jesorsky, O., Kirchberg, K.J., Frischholz, R.W. (2001). Robust Face Detection Using the Hausdorff Distance. In: Bigun, J., Smeraldi, F. (eds) Audio- and Video-Based Biometric Person Authentication. AVBPA 2001. Lecture Notes in Computer Science, vol 2091. Springer, Berlin, Heidelberg. https://DOI.org/10.1007/3-540-45344-X_14

[17] D. Peleshko and K. Soroka, "Research of usage of Haar-like features and AdaBoost algorithm in Viola-Jones method of object detection," 2013 12th International Conference on the Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), 2013, pp. 284-286.

[18] W. Yang and Z. Jiachun, "Real-time face detection based on YOLO," 2018 1st IEEE International Conference on Knowledge Innovation and Invention (ICKII), 2018, pp. 221-224, DOI: 10.1109/ICKII.2018.8569109.

[19] Y. Okafuji, J. Baba and J. Nakanishi, "Face-to-Face Contact Method for Humanoid Robots Using Face Position

Prediction," 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2019, pp. 666-666, DOI: 10.1109/HRI.2019.8673175.

[20] N. H. Barnouti, M. H. N. Al-Mayyahi and S. S. M. Al-Dabbagh, "Real-Time Face Tracking and Recognition System Using Kanade-Lucas-Tomasi and Two-Dimensional Principal Component Analysis," 2018 International Conference on Advanced Science and Engineering (ICOASE), 2018, pp. 24-29, DOI: 10.1109/ICOASE.2018.8548818.

[21] Aashish, Kamath & A., Vijayalakshmi. (2017). Comparison of Viola-Jones And Kanade-Lucas-Tomasi Face Detection Algorithms. Oriental journal of computer science and technology. 10. 151-159. 10.13005/ojcst/10.01.20.

[22] N. A and K. Jaisharma, "A Deep Learning Based Approach for Detection of Face Mask Wearing using YOLO V3-tiny Over CNN with Improved Accuracy," 2022 International Conference on Business Analytics for Technology and Security (ICBATS), 2022, pp. 1-5, doi: 10.1109/ICBATS54253.2022.9758925.

[23] G. Xu, A. Sohail Khan, A. J. Moshayedi, X. Zhang, and Y. Shuxin, "The Object Detection, Perspective and Obstacles In Robotic: A Review: ", EAI Endorsed Trans AI Robotics, vol. 1, no. 1, p. e13, Oct. 2022.

[24] V. Mohan, P. Shanmugapriya and Y. Venkataramani, "Object Recognition using image descriptors," 2008 International Conference on Computing, Communication and Networking, 2008, pp. 1-4, doi: 10.1109/ICCCNET.2008.4787717.

[25] R. C. Luo, C. T. Liao and Y. J. Chen, "Robot - human face tracking and recognition using relative affine structure," 2008 IEEE Workshop on Advanced robotics and Its Social Impacts, 2008, pp. 1-6, doi: 10.1109/ARSO.2008.4653585.

[26] Y. Nishina, J. K. Tan, H. S. Kim and S. Ishikawa, "Development of an autonomous robot for face tracking," 2007 International Conference on Control, Automation and Systems, 2007, pp. 1178-1181, doi: 10.1109/ICCAS.2007.4406512.

[27] Kumar, A., Kaur, A. & Kumar, M. Face detection techniques: a review. Artif Intell Rev 52, 927–948 (2019). https://DOI.org/10.1007/s10462-018-9650-2

[28] A. J. Moshayedi, A. Shuvam Roy, S. K. . Sambo, Y. . Zhong, and L. Liao, " Review On: The Service Robot Mathematical Model ", EAI Endorsed Trans AI Robotics, vol. 1, no. 1, p. e8, Feb. 2022.

[29] T. Gerig et al., "Morphable Face Models - An Open Framework," 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), 2018, pp. 75-82, doi: 10.1109/FG.2018.00021.

[30] G. M. Zafaruddin and H. S. Fadewar, "Face recognition: A holistic approach review," 2014 International Conference on Contemporary Computing and Informatics (IC3I), 2014, pp. 175-178, doi: 10.1109/IC3I.2014.7019610.

[31] K. -C. Kwak and W. Pedrycz, "Face Recognition Using an Enhanced Independent Component Analysis Approach," in IEEE Transactions on Neural Networks, vol. 18, no. 2, pp. 530-541, March 2007, doi: 10.1109/TNN.2006.885436.

[32] R. K. Sadykhov and V. A. Samokhval, "Face Identification Algorithm Based on Synthetic Linear Descriptors," 2005 IEEE Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 2005, pp. 593-595, doi: 10.1109/IDAACS.2005.283053.

[33] S. Nasr, K. Bouallegue, M. Shoaib and H. Mekki, "Face recognition system using bag of features and multi-class SVM for robot applications," 2017 International Conference on Control, Automation and Diagnosis (ICCAD), 2017, pp. 263-268, doi: 10.1109/CADIAG.2017.8075668.

[34] S. Tayeb, A. Mahmoudi, F. Regragui and M. M. Himmi, "Efficient detection of P300 using Kernel PCA and support vector machine," 2014 Second World Conference on Complex Systems (WCCS), 2014, pp. 17-22, DOI: 10.1109/ICoCS.2014.7060953.

[35] S. A. Nazeer, N. Omar, K. F. Jumari and M. Khalid, "Face Detecting Using Artificial Neural Network Approach," First Asia International Conference on Modelling & Simulation (AMS'07), 2007, pp. 394-399, doi: 10.1109/AMS.2007.38.

[36] M. Turk, A. Pentland; Eigenfaces for Recognition. J Cogn Neurosci 1991; 3 (1): 71–86. doi: https://doi.org/10.1162/jocn.1991.3.1.71

[37] S. -H. Jeon, H. -S. Yoon and J. -H. Kim, "Frontal face reconstruction with symmetric constraints," 2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), 2016, pp. 854-855, DOI: 10.1109/URAI.2016.7733994.

[38] Muhtadin, A. R. Muttaqin and S. Sumpeno, "Implementation of face detection and recognition of Indonesian language in communication between humans and robots," 2016 International Conference on Information & Communication Technology and Systems (ICTS), 2016, pp. 53-57, DOI: 10.1109/ICTS.2016.7910272

[39] S. Yasaswini, B. Akshitha, R. S. Suchitra and M. R. Rao, "Facial expression controlled humanoid robot," 2017 International Conference on Trends in Electronics and Informatics (ICEI), 2017, pp. 754-757, DOI: 10.1109/ICOEI.2017.8300804.

[40] M. D. Putro and K. -H. Jo, "Real-time Face Tracking for Human-Robot Interaction," 2018 International Conference on Information and Communication Technology Robotics (ICT-ROBOT), 2018, pp. 1-4, DOI: 10.1109/ICT-ROBOT.2018.8549902.

[41] T. M. W. Vithanawasam and B. G. D. A. Madhusanka, "Dynamic Face and Upper-Body Emotion Recognition for Service Robots," 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS), 2018, pp. 428-432, doi: 10.1109/ICIS.2018.8466505.

[42] Y. Okafuji, J. Baba and J. Nakanishi, "Face-to-Face Contact Method for Humanoid Robots Using Face Position Prediction," 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 2019, pp. 666-666, DOI: 10.1109/HRI.2019.8673175.

[43] W. -y. LU and M. YANG, "Face Detection Based on Viola-Jones Algorithm Applying Composite Features," 2019 International Conference on Robots & Intelligent System (ICRIS), 2019, pp. 82-85, DOI: 10.1109/ICRIS.2019.00029.

[44] P. B. Nithin, A. Francis, A. J. Chemmanam, B. A. Jose and J. Mathew, "Face Tracking Robot testbed for Performance Assessment of Machine Learning Techniques," 2019 7th International Conference on Smart Computing & Communications (ICSCC), 2019, pp. 1-5, DOI: 10.1109/ICSCC.2019.8843628.

[45] P. Chakraborty, S. Ahmed, M. A. Yousuf, A. Azad, S. A. Alyami and M. A. Moni, "A Human-Robot Interaction System Calculating Visual Focus of Human's Attention Level," in IEEE Access, vol. 9, pp. 93409-93421, 2021, DOI: 10.1109/ACCESS.2021.3091642.

[46] J. Heredia et al., "Adaptive Multimodal Emotion Detection Architecture for Social Robots," in IEEE Access, vol. 10, pp. 20727-20744, 2022, DOI: 10.1109/ACCESS.2022.3149214.

[47] Moshayedi, A. J., Chen, Z., Liao, L., & Li, S. (2022). Sunfa Ata Zuyan machine learning models for moon phase detection: algorithm, prototype and performance comparison. TELKOMNIKA Telecommunication Computing Electronics and Control, 20(1), 129. https://DOI.org/10.12928/telkomnika.v20i1.22338
[48] S. Manzoor et al., "Edge Deployment Framework of GuardBot for Optimized Face Mask Recognition With Real-Time Inference Using Deep Learning," in IEEE Access, vol. 10, pp. 77898-77921, 2022, DOI: 10.1109/ACCESS.2022.3190538.
[49] D. Banerjee and K. Yu, "Robotic Arm-Based Face Recognition Software Test Automation," in IEEE Access, vol. 6, pp. 37858-37868, 2018, DOI: 10.1109/ACCESS.2018.2854754.
[50] H. Zhang, B. Fan, X. Zhang, H. Zhan and X. Li, "A Six-degree-of-freedom Face Tracking Method for Non-contact Physiological Detection Robot," 2021 IEEE International Conference on Mechatronics and Automation (ICMA), 2021, pp. 950-955, DOI: 10.1109/ICMA52036.2021.9512790
[51] M. Karahan, H. Kurt and C. Kasnakoglu, "Autonomous Face Detection and Tracking Using Quadrotor UAV," 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2020, pp. 1-4, DOI: 10.1109/ISMSIT50672.2020.9254469.
[52] C. Iaboni, H. Patel, D. Lobo, J. -W. Choi and P. Abichandani, "Event Camera-Based Real-Time Detection and Tracking of Indoor Ground Robots," in IEEE Access, vol. 9, pp. 166588-166602, 2021, DOI: 10.1109/ACCESS.2021.3133533.
[53] S. -S. Lee, S. -J. Jang, J. Kim and B. Choi, "A hardware architecture of face detection for human-robot interaction and its implementation," 2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia), 2016, pp. 1-2, DOI: 10.1109/ICCE-Asia.2016.7804752.
[54] D. S. Pamungkas and S. Al-Aidid, "Detection, Recognition, and Tracking Face Using 2 DoF Robot with Haar LBP Histogram," 2018 International Conference on Applied Engineering (ICAE), 2018, pp. 1-5, DOI: 10.1109/INCAE.2018.8579409.
[55] B. Cilmi and M. Mercimek, "Design and Implementation of Real-Time Face Tracking Humanoid Robot," 2018 6th International Conference on Control Engineering & Information Technology (CEIT), 2018, pp. 1-6, DOI: 10.1109/CEIT.2018.8751757.
[56] S. Sharmin, M. M. Hoque, S. M. R. Islam, M. F. Kader and I. H. Sarker, "Development of Duplex Eye Contact Framework for Human-Robot Inter Communication," in IEEE Access, vol. 9, pp. 54435-54456, 2021, DOI: 10.1109/ACCESS.2021.3071129.
[57] G. -S. J. Hsu and J. -W. Huang, "A photographer robot with multiview face detector," 2016 IEEE International Conference on Industrial Technology (ICIT), 2016, pp. 2152-2156, DOI: 10.1109/ICIT.2016.7475103.
[58] W. Jiang and W. Wang, "Face detection and recognition for home service robots with end-to-end deep neural networks," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017, pp. 2232-2236, DOI: 10.1109/ICASSP.2017.7952553.11
[59] C. -L. Hwang, D. -S. Wang, F. -C. Weng and S. -L. Lai, "Interactions Between Specific Human and Omnidirectional Mobile Robot Using Deep Learning Approach: SSD-FN-KCF," in IEEE Access, vol. 8, pp. 41186-41200, 2020, DOI: 10.1109/ACCESS.2020.2976712.
[60] M. Hu, Q. Zhang and Z. Wang, "Application of rough sets to image pre-processing for Face detection," 2008 International Conference on Information and Automation, 2008, pp. 545-548, doi: 10.1109/ICINFA.2008.4608060.
[61] A. J. Moshayedi, M. Gheibollahi, and L. Liao, "The quadrotor dynamic modeling and study of meta-heuristic algorithms performance on optimization of PID controller index to control angles and tracking the route," IAES Int. J. Robot. Autom., vol. 9, no. 4, p. 256, 2020,