ICIC Express Letters
Part B: Applications
Volume **13**, Number **3**, March **2022**

ICIC International ©2022 ISSN 2185-2766
pp. 321–327

LOW-COST VISION-BASED FACE RECOGNITION USING ESP32-CAM FOR TRACKED ROBOT

Widodo Budiharto[1], Edy Irwansyah[1], Jarot Sembodo Suroso[2] and Alexander Agung Santoso Gunawan[3]

[1]Computer Science Department, School of Computer Science
[2]Information Systems Department, BINUS Graduate Program – Master of Information Systems Program
[3]Mathematics Department, School of Computer Science
Bina Nusantara University
Jl. K. H. Syahdan No. 9, Kemanggisan, Palmerah, Jakarta 11480, Indonesia
{ wbudiharto; eirwansyah; jsembodo; aagung }@binus.edu

Received April 2021; accepted July 2021

Abstract. *The main problem in computer vision for robot vision is to detect and recognize objects quickly and accurately. Face detection is needed by the tracked robot for understanding the environment, such as military robots for tracking and shooting the enemy. In this paper, we propose a prototype of a tracked robot with video streaming capability and the ability for face recognition using ESP32-CAM. We use a tracked robot as the simulation of a military robot. We propose an algorithm for face recognition controlling the robot using ESP32-CAM. The system can recognize an object from a video streaming run with 4-5 fps. The methods were explained, and experimental results were presented.*

Keywords: ESP32-CAM, Wi-Fi technology, Computer vision, Video streaming, Face detection

1. **Introduction.** Nowadays, most robotics systems use computer vision and sensors for general purposes, such as for surveillance and obstacle avoidance [1]. Remote-controlled robots for surveillance are also combined with video streaming for wireless operation such as developed by [2]. Robots for military purposes, in general, called an unmanned ground vehicle (UGV) is used to augment the soldier's capability. Many military robots were developed to maintain security and spies in conflict areas or borders based on cameras, firearms, and missiles. The study of the military tank robot system has been carried out, for example, [3,4].

Object detection and recognition are essential tasks in all applications of computer vision. There are many research projects considering the problem of recognizing different objects. Most of these projects work under certain conditions, where there is a finite number of objects and the environment is somewhat controlled. Contours are very well suited where a shape represents an object and they allow a certain level of tolerance since a value of how many objects must be similar can be added [5].

In developing a military robot, to produce autonomous systems, the system must be able to track targets/recognize objects based on computer vision. Robots are also expected to be able to recognize faces/objects that can be enemies to be conquered. Uncertainty is very common in tracking objects based on vision, so the application of probabilistic robotics in the development of intelligent robots is very important [6]. Robotic systems can take many forms, be stealthy or intentionally noisy, cloak themselves and deceive the enemy physically, electronically, and behaviorally. There are fundamental ethical implications in allowing full autonomy for these robots. Among the questions to be asked are as follows.

DOI: 10.24507/icicelb.13.03.321

- Will autonomous robots be able to follow established guidelines of the Laws of War and Rules of Engagement, as specified in the Geneva Conventions?
- Will robots know the difference between military and civilian personnel?
- Will they recognize a wounded soldier and refrain from shooting [7]?

Zim [8] analyzed the use of TinyML for neural network application using Xtensa LX6, the microprocessor inside ESP32. The experiment uses a different number of inputs neurons (9, 36, 144, and 576) with one and two hidden layers. From the experiment, the result shows that data transfer between different types of memory has less impact on the total run time. The advantage of this study is that this robot is controlled using Wi-Fi, to make the connection range wider than using Bluetooth, and use a low-cost controller. The major contribution of this research is to propose an alternative approach that has a cost lower than another model, to be used in the organization that has a limited budget to develop the robot itself. The result of our experiment shows the system was able to recognize an object from video streaming about 4-5 fps.

Wi-Fi is a wireless networking technology and stands for "wireless fidelity". Wi-Fi is a high Internet connection and was invented by NCR Corporation/AT&T in the Netherlands in 1991. This research is very important to produce models and methods of a military robot that can be controlled remotely. We present an introduction in Part 1, Part 2 as a concept of object detector, the proposed method in Part 3 and experimental results in Part 4 then conclusion in Part 5. Figure 1 shows a prototype of Wi-Fi-based tracked robot as proposed.

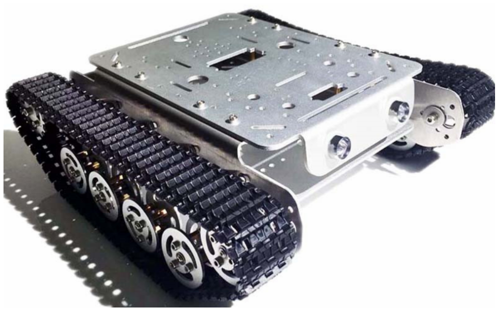

FIGURE 1. A prototype of the tracked robot using ESP32-CAM

2. Related Works.

2.1. Military robot using wireless technology. Different kinds of robots are specifically employed for doing special tasks in military applications. In military services, there are some areas in which some of the tasks involve greater risk and danger, and therefore, those tasks must be performed without military personnel, solely by the robots. For example, the Guardium is a part of a new category of military robot/UGV. Israel is the first country in the world to use these robots to replace soldiers on missions like border patrols. The Guardium is based on a Tomcar dune-buggy-like vehicle and is equipped with a range of sensors, cameras, and weapons. It can be driven by a soldier sitting in a command center mile away or receive a predesignated route for its patrol, making it completely autonomous [9].

A radio frequency controlled robotic vehicle is designed using a robotic vehicle that is interfaced with a radio frequency remote control. RF transmitter is used by a control panel or controlling person and RF receiver is connected to the robotic vehicle that is to be controlled remotely. Radiofrequency remote control works over an adequate range (up to 200 meters) by facilitating with a proper antenna with an RF power below 1 Watt. Lin et al. [10] studied the use of the second-order terminal sliding mode controller (SOTSMC), which has the function to control the formation of unmanned aerial vehicles (UAV). Based on the experiment, the result of using SOTSMC performs better than using conventional SMC controller.

Prabhu and Hebbal [11] researched unarmed robot for defense and security using Arduino Uno. The robot can be controlled using apps on the smartphone. The connection between the robot and the smartphone uses Bluetooth. It can be used to know exactly how many people are hidden in a room, and based on that action can be taken by the armed forces. Ghute et al. [12] researched about military surveillance robot using Raspberry Pi3. This robot is equipped with Bluetooth connectivity, several sensors, a camera, and a Wi-Fi network, to stream the video captured by the robot through another device.

2.2. **Wi-Fi technology for robot.** A robot in a military context is a powered machine that (1) senses, (2) thinks (in a deliberative, non-mechanical sense), and (3) acts. There are several possible AI applications for military robots. Replacing frozen software with systems that do not need to be refreshed periodically creates a broad potential for creating more nimble systems, possibly at a lower cost. Again, AI could be used in training systems [2].

Wi-Fi is a wireless networking technology, based on the IEEE 802.11 standards, which are commonly used for local area networking of devices and Internet access. Wi-Fi uses multiple parts of the IEEE 802 protocol family and is designed to seamlessly interwork with its wired sibling Ethernet. Compatible devices can network through a wireless access point to each other as well as to wired devices and the Internet.

The ESP32-CAM is a full-featured microcontroller that also has an integrated video camera and microSD card socket. It is inexpensive and easy to use and is perfect for IoT devices requiring a camera with advanced functions like image tracking and recognition. The ESP32-CAM is based upon the ESP32-S module, so it shares the same specifications. It has the following features:

- 802.11b/g/n Wi-Fi
- Bluetooth 4.2 with BLE
- UART, SPI, I2C, and PWM interfaces
- Clock speed up to 160 MHz
- Computing power up to 600 DMIPS
- 520 KB SRAM plus 4 MB PSRAM
- Support Wi-Fi Image Upload
- Multiple Sleep modes
- Firmware Over the Air (FOTA) upgrades are possible
- 9 GPIO ports
- Built-in Flash LED

Figure 2 below shows the connection between the power supply and FTDI programmer for ESP32-CAM.

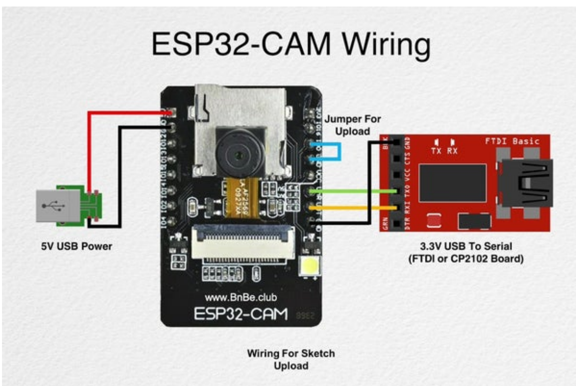

FIGURE 2. ESP32-CAM power supply and programmer [13]

3. **Proposed Method.** Autonomy in a robot is the capacity to operate in the real-world environment without any form of external control, once the machine is activated and at least in some areas of operation, for extended periods. In this study, we propose a tracked robot as a basis for the military robot as shown in Figure 3. This robot model can also be controlled remotely (teleoperated). Motor drivers and tank wheels and a strong and adequate mechanical model are also needed to be able to move on the battlefield [6]. The previous research has produced a tank robot model but is controlled by Android apps and Bluetooth which has certainly lower security than Wi-Fi [14].

The authors have conducted prior research for object tracking using a method of color-based object detection and Kalman filters that are adequate for tracking objects. Programming a robot with a good algorithm and supported by artificial intelligence is an important aspect today [15]. The camera system on the robot will be used in this research to obtain the vision of a target. We use high voltage 24 V and current 5 A for the DC motor and drivers of the tank able to move with enough power [14]. The USB Wi-Fi will be used for accepting commands from laptop under the browser application and Raspberry controller as shown in Figure 3. A single camera was connected to the Raspberry for processing video streaming. Figure 3 shows the architecture of the robot with ultrasonic sensors and the driver motor with high current 5 A.

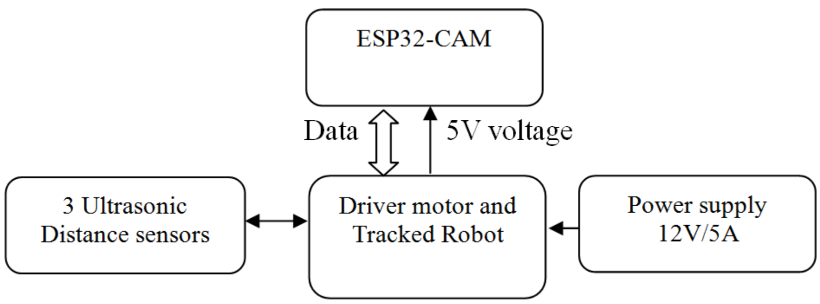

Figure 3. The architecture of a tracked robot with object recognition capability

We develop a program based on Arduino and a library of computer vision in C++. First, the program will make a connection between the server and the robot using Wi-Fi [16]. The program will find a face in front of the camera. The algorithm for video streaming, face recognition, and controlling the robot is shown in Algorithm 1.

Algorithm 1. Video streaming, face recognition and controlling a robot using Wi-Fi

```
declare variables
const ssid = ""; //Wi-Fi Name SSID
const password = ""; //Wi-Fi Password
function face_recognition(dl_matrix3du_t *image_matrix, box_array_t *net_boxes)
function startCameraServer()
function ObstacleAvoidance()
while (WiFi.status() == CONNECTED)
    startCameraServer()
    Serial.print("Camera Ready! Use http://")
    Serial.print(WiFi.localIP())
    Serial.println("to connect")
    ObstacleAvoidance()
    face_recognition() //auto scanning left/right
      if (matchFace == false) then
      begin
```

```
            move robot forward
            print "Enemy not found"
        endif
        if (matchFace == true) then
        begin
            print "Enemy detected"
        endif
end

definition face_recognition(dl_matrix3du_t *image_matrix, box_array_t *net_boxes)
begin
matched_id = recognize_face(&id_list, aligned_face);
if (matched_id >= 0)
begin
    printf ("Match Face ID: %u", matched_id);
    printf (image_matrix, FACE_COLOR_GREEN, "Enemy detected %u", matched_id);
    matchFace = true;
endif
else
    begin
        print "No Match Found"
        matched_id = −1;
    end
endif
end
```

We can access the data of the camera using Wi-Fi. The SSID and password are used by the controller for connecting with Wi-Fi. For capturing and streaming the video, we use MJPG Streamer, which can send real-time video cameras using Wi-Fi. For using MJPG Streamer, we must connect to the same Wi-Fi connection, and must run both software in the robot and our gadget. For this research, we use a small camera, that can work with low energy.

4. **Experimental Results.** We developed a program for computer vision using a laptop, OpenCV, and Python, and Wi-Fi for controlling the robot. Figure 4 shows the result of video streaming using ESP32-CAM.

Based on the experiment, video streaming is very good. This is because of the high-speed connection provided by the server. The robot used a single camera, but it has a limitation that is only able to measure an object that has a 4-meter maximum distance. The ability of ESP32-CAM as a controller for the military robot is very excellent. The weakness of the system is that it cannot track the moving object that moves very fast. Figure 5 shows the result of face recognition using our system with about 4-5 fps.

5. **Conclusions.** In this paper, we propose a model of a tracked robot using a low-cost controller and computer vision that can be used for face recognition. The ability of video streaming and Wi-Fi technology for the robot is very good. The system can recognize an object from a video streaming run with 4-5 fps. The tracked robot is an important tool for combat. Future warfare will involve operators and machines, not soldiers shooting at each other on the battlefield. For future work, we will propose a method for shooting a target with more precision and accuracy and using the real gun.

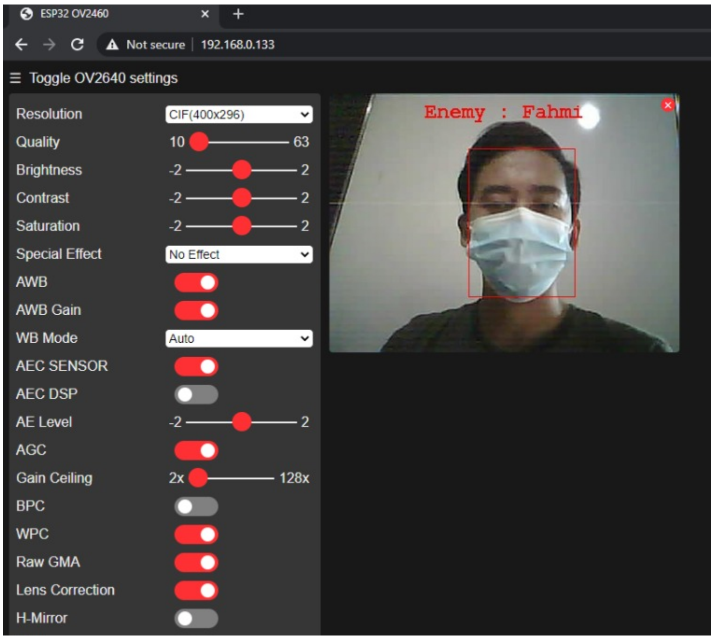

FIGURE 4. ESP32-CAM tools and displaying results from the camera

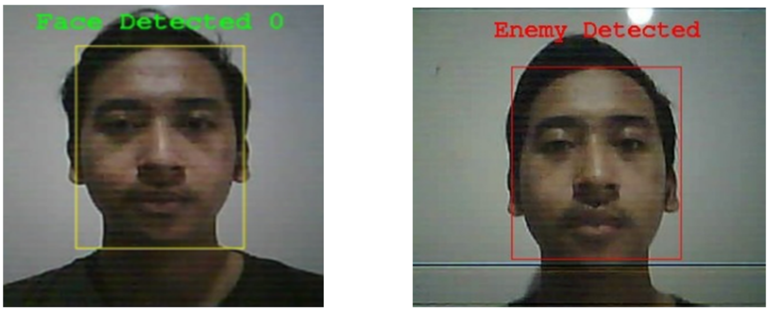

FIGURE 5. Face detection and recognition using ESP32-CAM

Acknowledgment. This work is supported by Directorate General of Research and Development Strengthening, Indonesian Ministry of Research, Technology, and Higher Education, as a part of Penelitian Dasar Terapan Unggulan Perguruan Tinggi Research Grant year 3 to Binus University titled "Pemodelan Sistem Kendali Robot Tempur dengan Teknologi Wireless dan Computer Vision" with contract no. 064/E4.1/AK.04.PT/2021 and contract date 12 July 2021, and inheritance contract no. 064/E4.1/AK.04.PT/2021, 3530/LL3/KR/2021 and contract date 12 July 2021.

REFERENCES

[1] W. Budiharto, Intelligent surveillance robot with obstacle avoidance capabilities using neural network, *Computational Intelligence and Neuroscience*, vol.2015, 2015.
[2] W. Budiharto, Design of tracked robot with remote control for surveillance, *International Conference on Advanced Mechatronic Systems*, Japan, 2014.
[3] J. H. Lee and W. S. Yoo, Velocity decision for unmanned ground vehicle considering road roughness, *Proc. of KSME 2011 Fall Annual Meeting*, 2011.
[4] W.-S. Yoo, D.-H. Cho, J.-H. Lee and K.-C. Yi, Maneuvering speed of an off-road autonomous vehicle, *Proceedings of MOVIC*, 2010.
[5] C. Astua, R. Barber, J. Crespo and A. Jardon, Object detection techniques applied on mobile robot semantic navigation, *Sensors*, vol.14, no.4, pp.6734-6757, 2014.
[6] S. Bhat and M. Meenakshi, Vision-based robotic system for military applications – Design and real-time validation, *2014 5th International Conference on Signal and Image Processing (ICSIP)*, pp.20-25, 2014.

[7] P. Lin, G. Bekey and K. Abney, *Autonomous Military Robotics: Risk, Ethics, and Design*, US Department of Navy, Office of Naval Research, California State Polytechnic University, San Luis Obispo, 2008.

[8] M. Z. H. Zim, *TinyML: Analysis of Xtensa LX6 Microprocessor for Neural Network Applications by ESP32 SoC*, DOI: 10.13140/RG.2.2.28602.11204/1, 2021.

[9] *Advanced Military Robot*, https://nypost.com/2017/01/29/why-israel-has-the-mosttechnologically-ad vanced-military-on-earth/, Accessed on 1 August 2020.

[10] H. Lin, K. Chen and R. Lin, Finite-time formation control of unmanned vehicles using nonlinear sliding mode control with disturbances, *International Journal of Innovative Computing, Information and Control*, vol.15, no.6, pp.2341-2353, 2019.

[11] B. A. Prabhu and S. Hebbal, Small unarmed robot for defense and security: A cost-effective approach using Arduino Uno, *2017 2nd International Conference on Emerging Computation and Information Technologies (ICECIT)*, pp.1-6, 2017.

[12] M. S. Ghute, K. P. Kamble and M. Korde, Design of military surveillance robot, *2018 1st International Conference on Secure Cyber Computing and Communication (ICSCCC)*, pp.270-272, 2018.

[13] *ESP32-CAM Board*, https://www.instructables.com/IP-Camera-With-Face-Detection-Using-the-ES P32-CAM-/, Accessed on 4 March 2021.

[14] W. Budiharto, A. A. S. Gunawan, E. Irwansyah and J. S. Suroso, Android-based wireless controller for military robot using Bluetooth technology, *2019 2nd World Symposium on Communication Engineering (WSCE)*, pp.215-219, 2019.

[15] A. Billard, S. Calinon, R. Dillmann and S. Schaal, Robot programming by demonstration, in *Handbook of Robotics*, B. Siciliano and O. Khatib (eds.), Secaucus, NJ, USA, Springer, 2008.

[16] *Object Recognition Using ESP32-CAM*, https://github.com/longpth/ESP32-CAM-ObjectDetection, Accessed on 10 February 2021.