STM JOURNALS

Journal of
VLSI Design Tools & Technology

https://journals.stmjournals.com/jovdtt

ISSN: 2249-474X (Online)
ISSN: 2321-6492 (Print)
Volume 14, Issue 2, 2024
May–August
DOI (Journal): 10.37591/JoVDTT

Review JoVDTT

Surveillance Car Using ESP32 Cam Module by Using GAN Model

Raj Shaikh[1], Anushri Kulkarni[1,*], Suraj Shinde[1], Yashraj Deshpande[2]

Abstract

This study shows the Surveillance Car system which leverages the ESP32 Cam module and incorporates advanced image processing through a Generative Adversarial Network (GAN) model to redefine the landscape of mobile surveillance systems. The ESP32 Cam serves as the core hardware platform, offering compact design and wireless capabilities for real-time image capture and remote monitoring. The system's innovation lies in the integration of a GAN model for image processing, enhancing the system's ability to intelligently analyze captured data. This combination of cutting-edge hardware and artificial intelligence aims to provide a sophisticated, context-aware surveillance solution. The GAN model enables the identification of anomalies and specific objects within surveillance frames, enhancing adaptability to dynamic environments. With potential applications in security, law enforcement, and smart city initiatives, this study underscores the synergy between ESP32 Cam and GAN-based image processing to create an intelligent and adaptable surveillance system for contemporary needs.

Keywords: Surveillance, ESP32 CAM, GAN model, smart phone, camera, motor driver

INTRODUCTION

The Surveillance Car project utilizing the ESP32 Cam module and image processing with a Generative Adversarial Network (GAN) model represents a cutting-edge integration of hardware and artificial intelligence. The project aims to enhance traditional surveillance systems by incorporating advanced image processing capabilities. The ESP32 Cam module, known for its compact size and wireless connectivity features, serves as the core hardware platform.

It is equipped with a camera module that captures real-time images or video feeds. The ESP32 Cam's wireless capabilities enable seamless communication and data transfer, providing an efficient means for remote monitoring [1]. The innovative aspect of this project lies in the utilization of GAN for image processing. GANs, a class of deep learning models, are employed to generate and manipulate images, enabling the system to discern anomalies or identify objects of interest.

This application of GANs enhances the surveillance capabilities by providing more sophisticated and context-aware analysis of the captured images. The combination of ESP32 Cam and GAN-based image processing not only facilitates real-time monitoring but also opens up possibilities for intelligent surveillance, where the system can learn and adapt to different scenarios over time. This project represents a convergence of hardware and artificial intelligence to create a smart surveillance solution with broader applications in security and monitoring domains.

LITERATURE SURVEY

The paper by Yuanfu *et al*. can detect the presence of objects, functioning as an obstacle

***Author for Correspondence**
Anushri Kulkarni
E-mail: skncoe@sinhgad.edu

[1]Student, Department of Electronic & Telecommunication Engineering, Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India
[2]Assistant Professor, Department of Electronic & Telecommunication Engineering, Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India

Received Date: July 18, 2024
Accepted Date: July 25, 2024
Published Date: July 30, 2024

Citation: Raj Shaikh, Anushri Kulkarni, Suraj Shinde, Yashraj Deshpande. Surveillance Car using ESP32 Cam Module by using GAN Model. Journal of VLSI Design Tools & Technology. 2024; 14(2): 28–37p.

avoider, and providing real-time positioning information [2]. It is computationally controlled, enhancing its usability. The integration of Raspberry Pi software with the hardware was accomplished using the Python programming language. The authors highlight that this work has significant potential in the field of robotics, as various attributes can be added to enhance its precision and functionality.

Cho cardiography is a technique used to obtain real-time images of heart structures using ultrasound waves. It offers several advantages, including low operational cost, non-invasiveness, wide availability, and minimal discomfort to patients [3]. To assess heart functionality, it is necessary to obtain various cardiac parameters, such as the thickness of the heart wall, the enclosed area, and the variation of these shape attributes throughout the cardiac cycle. These shape attributes are acquired by identifying the endocardial (inner) and epicardial (outer) boundaries of the heart wall, which is achieved through echocardiographic image segmentation techniques [4, 5]. Image segmentation involves partitioning an image into separate regions that collectively form the original image. The image is divided into segments with similar properties, such as gray intensity levels, texture, and color. Over time, various automated and semi-automated methods have been developed to perform this task, providing more reliable and consistent results compared to manual segmentation, which is time-consuming and varies between individuals [1].

Many problems in image processing, computer graphics, and computer vision can be viewed as translating an input image into a corresponding output image. Like how a concept can be expressed in either English or French, a scene can be rendered as an RGB image, a gradient field, an edge map, a semantic label map, etc. We define automatic image-to-image translation as the task of converting one possible representation of a scene into another, given sufficient training data. Traditionally, each of these tasks has been addressed with separate, specialized machinery, despite the commonality that the setting always involves predicting pixels from pixels. Gong et al aims to develop a unified framework to address all these problems within a common context. Optical satellite remote sensing images are used in various applications, and to better utilize these images and improve their resolution, researchers have focused on super resolution reconstruction (SRR) methods [2]. These methods convert low-resolution (LR) images into high-resolution (HR) images, and the results of SRR methods from LR images are referred to as super resolution (SR) images. The SRR task involves an ill-posed problem, requiring methods to fully exploit prior knowledge. Some methods use prior knowledge by establishing degradation models, while others are based on sparse representation and learn a dictionary for transforming between LR patches and HR patches [6]. However, the limited learning ability of these methods restricts their performance. To address this, Bernard *et al*. introduced a convolutional neural network (CNN) with high learning ability into SRR tasks, proposing the SRCNN [7]. Despite its innovative approach, this method uses pixel loss to optimize the network, resulting in overly smooth images as pixel loss does not consider perceptual quality.

SYSTEM ARCHITECTURE

The surveillance car's architecture centers on the ESP32-CAM board, which controls image capture and processing through its integrated camera module. Wi-Fi connectivity enables communication with other devices and cloud services. Motion detection triggers actions like image capture, while a GAN module enhances image quality or aids in object recognition. Optional features include object detection, data storage, user interfaces, power management, real-time communication, sensor integration, security measures, remote control, and mobility considerations.

OPERATIONAL MODES

Enhanced Image Quality Mode

In this mode, the surveillance car focuses on capturing high-quality images and videos using the ESP32-CAM.

Upon capturing images, the GAN module processes them to enhance image quality, improving clarity, sharpness, and contrast [5, 8, 9]. Enhanced images are then used for object recognition or detection, where the system identifies specific objects or persons of interest.

ISSN: 2249-474X (Online), ISSN: 2321-6492 (Print)

The surveillance car can store these enhanced images locally or upload them to a cloud server for further analysis or archival purposes. This mode is particularly useful for scenarios where detailed image analysis is required, such as identifying license plates or facial recognition.

Real-time Threat Detection Mode

In this mode, the surveillance car prioritizes real-time threat detection and response. The ESP32-CAM continuously captures live video feeds and processes them in real-time for motion detection. When motion is detected, the GAN module quickly analyzes the frames to identify potential threats or suspicious activities [10, 11].

Upon detection of a threat, the surveillance car triggers immediate actions such as sending alerts to designated recipients, activating sirens or lights, or recording video footage. Enhanced Image Quality Mode system is shown in Figure 1.

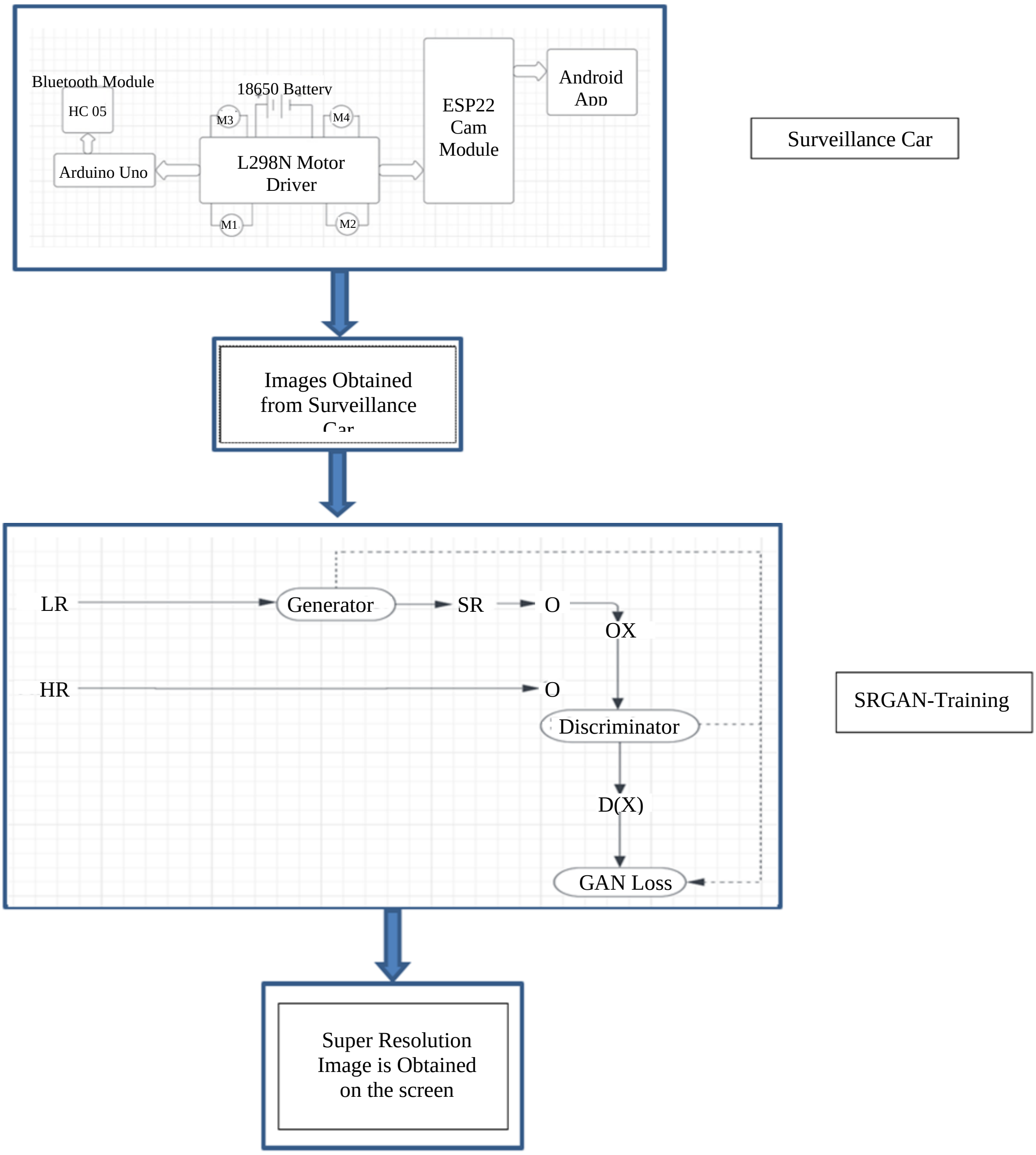

Figure 1. Block Diagram of the system.

Generative Adversarial Networks (GANs) gained significant attention in the field of computer vision for their ability to generate realistic images. Researchers and developers started leveraging GANs for tasks such as image synthesis, super-resolution, style transfer, and image-to-image translation [12]. GANs became a valuable tool for enhancing visual data in various applications, including surveillance and security.

Camera System

This includes one or more cameras strategically placed on the car to capture video footage from different angles. These cameras are the primary sensors for surveillance and capture visuals of the surrounding environment.

Sensor Interface

This module interfaces with various sensors beyond cameras, such as LIDAR, radar, ultrasonic sensors, and GPS. These sensors provide additional data about the car's surroundings, including distances, object detection, and location information.

Image/Video Processing

The captured video streams from the cameras are sent to this module. Here, image and video processing algorithms are applied, including object detection, tracking, and classification. Advanced techniques like deep learning may be employed for enhanced accuracy.

Data Fusion

Data from multiple sensors, such as cameras, LIDAR, radar, and GPS, are combined and fused together to create a comprehensive situational awareness picture. This improves the overall accuracy of the surveillance system.

Communication Module

This component handles data communication between the surveillance car and external systems. It might include cellular, Wi-Fi, or other communication technologies to transmit real-time data to a central monitoring station or other relevant parties.

Control System

The control system manages the movement of the car based on the inputs from the various sensors. It ensures that the car operates safely and efficiently while performing its surveillance tasks.

Object Detection and Tracking

This module processes the video data to identify and track objects of interest, such as pedestrians, vehicles, or other potential points of concern.

Storage Unit

Recorded video footage, along with processed data, can be stored in this unit for later analysis or evidence purposes. Depending on the application, this could be a local storage device or a cloud-based system.

Central Monitoring Station

The surveillance car's data and video feed can be transmitted to a central monitoring station where human operators can review the data in real-time. Alerts and notifications can also be generated if certain predefined events or anomalies are detected.

COMPONENT DETAILS

Arduino uno

The Arduino Uno is a popular microcontroller board based on the ATmega328P microcontroller. It is part of the Arduino family of open-source hardware platforms that are designed for easy prototyping and development of interactive projects. Arduino UNO is shown in Figure 2. Here are some key details about the Arduino Uno:

Microcontroller

The Arduino Uno is powered by the ATmega328P microcontroller, which is an 8-bit AVR (Advanced Virtual RISC) microcontroller.

Clock Speed

The ATmega328P on the Arduino Uno runs at 16 MHz.

Digital I/O Pins

The board has 14 digital input/output pins (of which 6 can be used as PWM outputs).

Analog Input Pins

There are six analog input pins.

Voltage Regulator

The board has a built-in voltage regulator, allowing it to be powered with an external DC power supply (7–12 V) or through USB.

USB Connection

The Arduino Uno can be connected to a computer via a USB cable, which also provides power to the board.

Power Supply

The board can be powered using a USB cable, an external power supply, or a battery.

Memory

The ATmega328P has 32 kb of flash memory for storing the user's program, 2 kb of SRAM, and 1 kb of EEPROM.

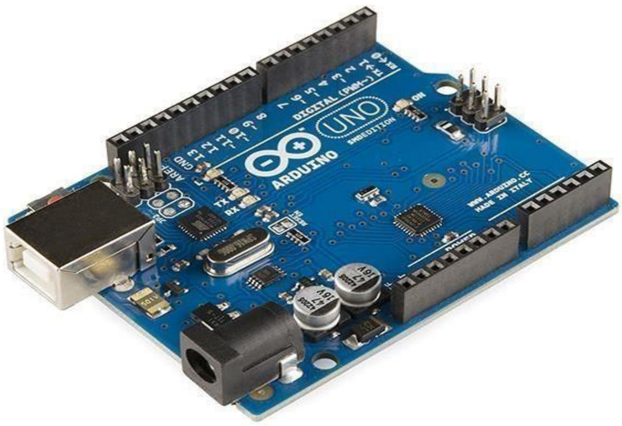

Figure 2. Arduino UNO.

Communication

The Arduino Uno has a USB connection for serial communication with a computer. It also has a UART (Universal Asynchronous Receiver-Transmitter) for serial communication with other devices.

Operating Voltage

The operating voltage of the Arduino Uno is 5 V.

Esp32 Camera Module

The ESP32 Camera module is a versatile piece of hardware that integrates the ESP32 microcontroller with a camera sensor, enabling it to capture and process images. Camera Module ESP32 is shown in Figure 3. Here are the key components and details of the ESP32 Camera module:

Figure 3. ESP32 Camera Module.

ESP32 Microcontroller

- *Processor:* The ESP32 module is powered by a dual-core Ten silica LX6 microcontroller.
- Clock Speed: Typically runs at 160 or 240 MHz.
- *Memory:* Includes both Flash memory for program storage and SRAM for data storage.

Camera Sensor

- *Type:* The module includes an OV2640 camera sensor, which is a popular choice for embedded camera applications.
- *Resolution:* The OV2640 supports a maximum resolution of 1600×1200 pixels.
- Output Format: Provides JPEG, YUV, RGB, and other image formats.

SD Card Slot

- *Storage Expansion:* Some ESP32 Camera modules include an SD card slot for additional storage capacity.

Antenna

- *Built-in Antenna:* Some ESP32 Camera modules come with a built-in PCB antenna for Wi-Fi and Bluetooth

L298N Motor driver

The L298N is a popular dual H-bridge motor driver integrated circuit (IC) used for controlling DC motors and stepper motors as shown in Figure 4. Here are the key details about the L298N motor driver:

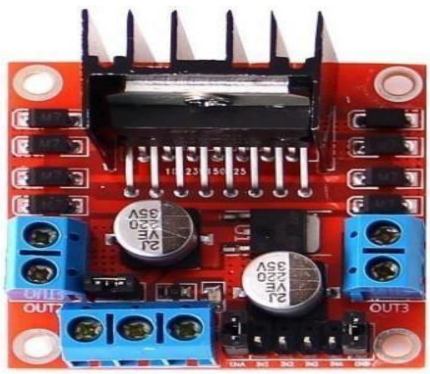

Figure 4. L298N Motor driver.

H-Bridge Configuration

The L298N is designed as a dual H-bridge driver, allowing it to control two motors independently.

Motor Types

It can be used to control both DC motors and bipolar stepper motors.

Maximum Voltage

The L298N can typically handle motor supply voltages in the range of 4.8 to 46 V.

Maximum Current

The maximum current per channel (per motor) is around 2 A, but this can vary based on the specific variant and implementation.

HC05 Bluetooth module

The HC-05 is a popular Bluetooth module that allows for wireless communication between electronic devices. Here are the key details about the HC-05 Bluetooth module, shown in Figure 5:

Figure 5. HC05 Bluetooth module.

Bluetooth Version

The HC-05 module is based on Bluetooth 2.0+EDR (Enhanced Data Rate) standard.

Operating Mode

The HC-05 can operate in either Master or Slave mode. In Master mode, it can initiate a connection to other Bluetooth devices, while in Slave mode, it can accept incoming connections. Power Supply (lithium-ion cell) shown in Figure 6.

Figure 6. Power Supply (lithium-ion cell).

IMPLEMENTATION DETAILS

Implementing a surveillance car using ESP32-CAM and a GAN module involves several steps, incl uding hardware setup, software development, and integration of various components. Here is a detailed implementation guide:

Hardware Setup

Assemble the surveillance car chassis, ensuring it can accommodate the ESP32-CAM board, motors, wheels, and any additional components. Connect the ESP32-CAM board to the chassis and ensure its securely mounted.

Software Development

Set up the Arduino IDE for ESP32 development and install the necessary libraries for ESP32-CAM and motor control. Develop firmware for the ESP32-CAM board to control the camera, capture images/video, and perform basic image processing tasks such as motion detection.

Integration

Integrate the firmware for the ESP32-CAM, motor control, and GAN module into a cohesive system. Ensure proper communication between the components, including sending image data from the ESP32-CAM to the GAN module for processing and receiving processed images back. Implement error handling and recovery mechanisms to address communication failures or other issues.

APPLICATIONS

Industrial Surveillance

In industrial settings such as warehouses or manufacturing facilities, the surveillance car can monitor for safety hazards, unauthorized access, or equipment malfunctions. It can navigate through aisles and inspect remote areas, providing real-time surveillance.

Smart Agriculture

In agriculture, the surveillance car can monitor crop fields for signs of pests, diseases, or irrigation issues. It can capture images of plants and analyze them using the GAN module to identify areas requiring attention.

Public Safety

Deployed in public areas, the surveillance car can assist law enforcement in monitoring crowds, identifying individuals of interest, and responding to emergencies. It can enhance security in parks, parking lots, and other public spaces.

CONCLUSION

The Surveillance Car project leveraging the ESP32 Cam module and image processing with a Generative Adversarial Network (GAN) model represents a significant stride in the realm of intelligent surveillance systems. The integration of the ESP32 Cam provides a robust hardware foundation, offering real-time image capture and wireless connectivity for seamless remote monitoring. The incorporation of the GAN model enhances the project's analytical capabilities, enabling the identification of anomalies and specific objects within captured frames. Through this amalgamation of cutting-edge hardware and artificial intelligence, the Surveillance Car showcases adaptability and context-awareness in monitoring dynamic environments. The project holds immense potential for applications in security, law enforcement, and smart city initiatives, providing a sophisticated, intelligent, and adaptable solution to contemporary surveillance needs. As technology continues to advance, this project serves as a testament to the innovative possibilities arising from the synergy between hardware platforms like the ESP32 Cam and state-of-the-art image processing techniques using GAN models. The insights gained from this project not only contribute to the field of surveillance technology but also underscore the continual evolution and refinement of integrated systems for the benefit of society.

Acknowledgements

We would like to express our sincere gratitude to everyone who has contributed to the successful completion of this project. Firstly, we would like to thank our project guide Mrs. Anushri Kulkarni, for providing us with valuable guidance and support throughout the project. Her insightful comments and constructive feedback were instrumental in shaping our project and achieving its goals. We would also like to express our gratitude to the faculty and staff of the Department of Electronics and Telecommunication, Smt. Kashibai Navale College of Engineering, for providing us with the necessary resources and facilities to carry out this project.

Furthermore, we would like to extend our thanks to our friends and colleagues who have provided us with their valuable assistance and support during the course of this project. Their encouragement and help have been invaluable to us in completing this project successfully. We acknowledge with gratitude all those who have contributed to this project in one way or another, and we sincerely appreciate their efforts and support.

REFERENCES

1. Chen Wei, Wenjing Wang, Wenhan Yang, Jiaying Liu. Deep retinex decomposition for lowlight enhancement. Computer Vision and Pattern Recognition. arXiv:1808.04560 [cs.CV]. 2018 Aug.
2. Yuanfu Gong, Puyun Liao. Enlighten-GAN for Super Resolution Reconstruction in Mid-Resolution Remote Sensing Images. Remote Sens. 2021; 37(22): 586–595.
3. Kin Gwn Lore, Adedotun Akintayo, Soumik Sarkar. LLNet: A deep autoencoder approach to natural low-light image enhancement. Pattern Recognition Volume 61, January 2017, Pages 650–662
4. Pallavi Kulkarni, Deepa Madathil. A review on echocardiographic image speckle reduction filters. Biomed Res. 2018; 29(12): 2582–2589.
5. Mishra A, Dutta PK, Ghosh MK. A GA based approach for boundary detection of left ventricle with echocardiographic image sequences. Image Vis Comput. 2003; 21(11): 967–976.
6. Ifan Jiang, Xinyu Gong. EnlightenGAN: Deep Light Enhancement without Paired Supervision. *IEEE Trans* Image Process. 2021; 30: 2340–2349.
7. Bernard O, Bosch JG, *et al.* Standardized Evaluation System for Left Ventricular Segmentation Algorithms in 3D Echocardiography. IEEE Trans Med Imag. 2016; 35(4): 967–977.
8. Xiaojie Huang, Lin Ben A, Compas Colin B, Sinusas Albert J, Staib Lawrence H, Duncan James S. Segmentation of Left Ventricles from Echocardiographic Sequences via Sparse Appearance Representation. IEEE *Workshop on* Mathematical Methods in Biomedical Image Analysis. 2012; 305–312.
9. Swetava Ganguli, Pedro Garzon, Noa Glaser. GeoGAN: A Conditional GAN with Reconstruction and Style Loss to Generate Standard Layer of Maps from Satellite Images. arXiv:1902.05611 [cs.CV]. *2019 Apr*.
10. Phillip Isola, Jun-Yan Zhu, et al. Image-to-Image Translation with Conditional Adversarial Networks. Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 5967–5976
11. Devabalan P. Satellite Image Processing on a Grid Based Computing Environment. Int J Comput Sci Mob Comput. 2014 Mar; 3(3): 1039–1044.

12. Pallavi Kulkarni, Deepa Madathil. Fully automatic segmentation of LV from echocardiography images and calculation of ejection fraction using deep learning. Int J Biomed Eng Technol. 2020; 40(3): 241–261.

