Design of Image Processing-Based Painting Museum Security Monitoring Design Using the YOLOv5 Method

An Netta Irene Winedar[1], Rachmad Saptono[2*], Amalia Eka Rakhmania[3]

[1,2] Digital Telecommunication Network Study Program, Department of Electrical Engineering, State Polytechnic of Malang, Malang City, 65141, Indonesia

[3] Telecommunication Engineering Study Program, Department of Electrical Engineering, State Polytechnic of Malang, 65141, Indonesia

[1]annettairene17@gmail.com, [2]rachmad.saptono@polinema.ac.id, [3]amaliaeka.rakhmania@polinema.ac.id

Abstract— **Paintings are valuable art assets in museums and are highly vulnerable to physical damage caused by direct human contact. Conventional surveillance systems such as CCTV rely heavily on manual monitoring and often fail to provide rapid responses to violations. This study proposes an image processing based museum security monitoring system using the YOLOv5 method to automatically detect hand movements approaching or entering the painting area. The system is implemented as a prototype using a Raspberry Pi 4, a webcam as visual input, and integrated outputs in the form of a web based monitoring interface, Telegram notifications, and audio warnings. System performance is evaluated by testing various camera distances and viewing angles to determine optimal placement. Experimental results show that the best detection performance is achieved at a distance of 1 meter with a camera angle of 45 degrees, producing an average detection accuracy of 67.2 to 68.8 percent. The average object detection processing time is 0.11 milliseconds, indicating that the system is capable of real time operation despite hardware limitations. The proposed system demonstrates the feasibility of automated museum security monitoring using lightweight object detection models on embedded platforms.**

Keywords—Image Processing, Museum Security, Object Detection, YOLOv5, Raspberry Pi

I. INTRODUCTION

Paintings hold the highest status in global auction classifications and are considered highly valuable commodities that occupy a special position for collectors. This significance arises because paintings are handmade creations that reflect unique artistic skills, emotion, and individuality, making them limited and enduring works of art whose value consistently increases over time [1]–[3]. Due to their high artistic and economic value, museums require reliable security systems to protect paintings from physical damage. Although security guards are commonly assigned to monitor museum galleries, the number of artworks usually exceeds the number of guards available, making it difficult to ensure continuous supervision of every painting [4]–[6].Therefore,comprehensive monitoring is needed to minimize potential threats and prevent damage.

To support this need, museums often install surveillance systems such as CCTV cameras. These cameras help observe visitor movements and detect inappropriate behavior around artworks [7–10]. Guards stationed in monitoring rooms can relay information to field personnel when violations occur. One form of harmful behavior frequently encountered is visitors touching paintings directly. Human fingertips contain oils, sweat, and acids that can transfer onto surfaces and potentially degrade the paint layer or protective coating over time [11]–[15]. Because of this risk, direct contact with paintings is strictly prohibited in most galleries.

However, relying solely on human monitoring is often inefficient. Communication delays between monitoring staff and field guards can slow response time, reducing the effectiveness of protective measures. This limitation highlights the need for automated supervision systems. Image-processing-based security solutions can analyze video captures in real time to detect violations more accurately and respond faster than manual surveillance alone. One promising approach involves detecting human hands near or inside the bounding box of a painting. Such detection enables instant identification of prohibited actions so that a warning or alert can be generated immediately.

The YOLOv5 method is highly suitable for this purpose because of its fast object-detection capability, high accuracy, and compatibility with devices that have limited computing resources, such as the Raspberry Pi .These advantages make YOLOv5 an effective model for real-time monitoring applications, particularly in environments requiring continuous observation like museums.

Based on these considerations, a system titled "Design for Museum Security Monitoring for Works of Art, Paintings Based on Image Processing Using the YOLOv5 Method" was developed. The system integrates several technologies to

*Corresponding author

enhance museum safety. A Raspberry Pi 4 functions as the primary controller, connecting the webcam, YOLOv5 detection process, website interface, and Telegram notification service. The webcam captures real-time visual data, which YOLOv5 processes to detect hands approaching the painting. The website serves as a monitoring dashboard for museum staff, displaying violation alerts and relevant information regarding system status. Meanwhile, Telegram notifications ensure that security personnel receive immediate

II. METHOD

A. System Block Diagram

The system block diagram will describe the input, process and output parts. The following is a block diagram of the system from the Design and Construction of Museum Security Monitoring for Works of Art, Paintings Based on Image Processing Using the YOLOv5 Method. Figure 1.

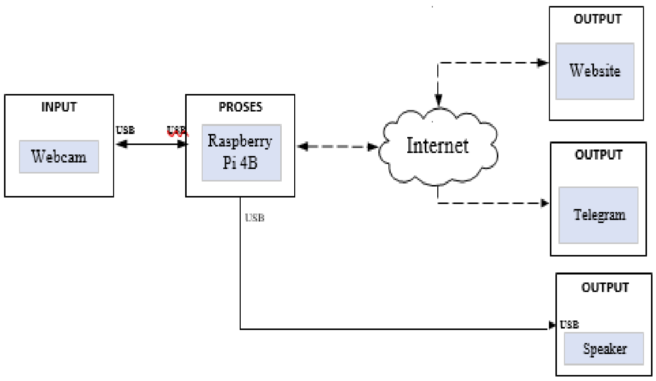


Figure 1 Block System Diagram

Based on the block diagram in Figure 1 above, there are three blocks, namely input block, process block and output block.

1. In the input block there is a webcam connected to the Raspberry Pi 4B using a USB cable. The webcam captures the captured image of the painting and the hand and the Raspberry Pi 4B will send the capture to the system.
2. The process block contains the Raspberry Pi 4B. Raspberry Pi 4B as a computer to process webcam captures using the YOLOv5 method to produce three outputs. To be able to access the system, the Raspberry Pi 4B is connected to the internet. Here the smartphone hotspot acts as internet for the Raspberry Pi 4B and Laptop so they can connectlocally. And smartphones are guaranteed to have a data quota so they can access websites online via laptop
3. The output block contains a speaker which is connected to the Raspberry Pi 4B using a USB cable and sends data via the website and Telegram. Websites and Telegram are accessed using the internet. A website notification will appear on the website page and a warning sound will appear when a hand is in the painting area. Then a message is sent via Telegram, the message contains the picture, time information and name of the painting.

There are three outputs from this system, namely websites, Telegram messages and speakers. The following is an explanation of the three outputs,

1. Website for monitoring security officers in the control room. If the detection condition is that there are hands around and/or inside the painting's boundary box, a notification will appear in the form of a pop-up on the website display and the detection results will be saved on the report page on the website.
2. Telegram messages as notifications to officers in the field or in the museum. This message is in the form of a display of the captured photo, a description of the time and the name of the painting.
3. Speakers to output sound appeals to visitors in the museum.

B. Tool Flow Diagram

In this research there is a flow diagram of how the tool works, as shown in Figure 2 below. Starting with initializing the Raspberry Pi or turning on the Raspberry Pi. Then the Raspberry Pi takes an image from the camera with the level of brightness and color sharpness obtained by the camera. Then, if the light captured by the camera is low so that the painting cannot be seen clearly, the website can adjust the brightness and sharpness of the colors in the image until the painting can be seen clearly. Then the Raspberry Pi detects the situation captured by the camera. If the detection condition is that there are hands around and/or inside the painting's bounding box, the Raspberry Pi sends a message to Telegram then displays a notification on the website then the speaker emits a warning sound. And if the detection condition is that there are no hands, the Raspberry Pi does not produce any output.

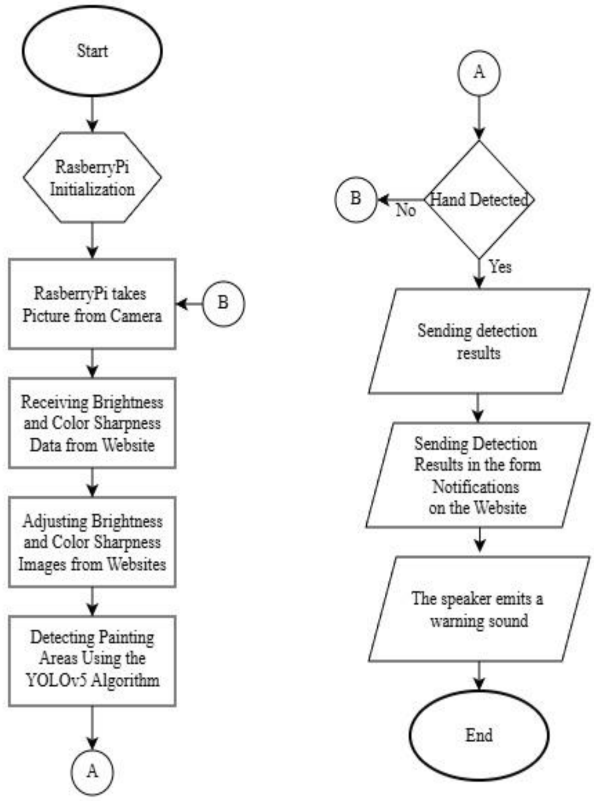
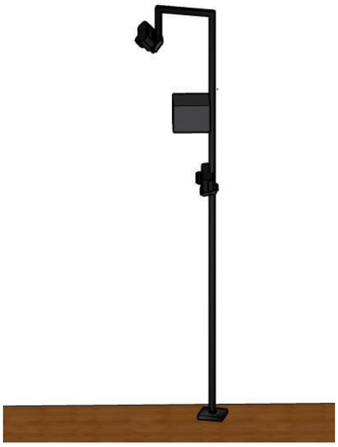


Figure 2 Tool Flow Diagram

C. Mechanical Design

In tool design, there is also a need for a mechanical design so that the tool design can be well shaped and easy to use. The following is a mechanical view in Figure 3

Figure 3. Tool Design



The following is an explanation of the mechanical design of the tool:

1. Raspberry Pi 4B is installed in the middle of the tripod using a holder.
2. The webcam is connected to the Raspberry Pi 4 using a USB cable and placed on the top end of the tripod
3. The speaker is attached using a double tip to the tripod.

D. Website Design

In designing this website, flow diagrams and website design will be explained. The following will be explained as below.

1). Website User Flow Diagram

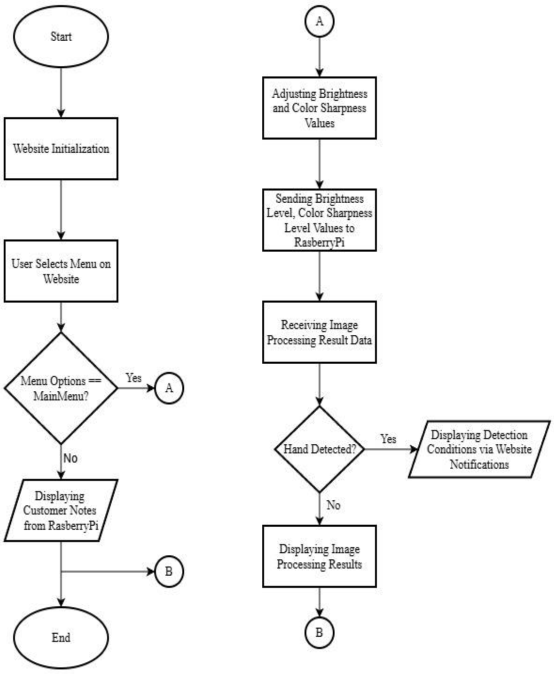


Figure 4 Website Flow Diagram

Figure 4 explains the flow diagram of the website system thatwill be used by museum security officers. The features on the website include: homepage and reports. The website is accessed without logging in. When the website is successfully accessed, the user can then select the menu on the website. If the user selects Main Menu, he will be directed to the main page which contains brightness level settings and color sharpness, frame rate and detection conditions. Then, if there is a hand in the painting area, the website will receive the image processing data and then display the detection conditions in the form of a notification on the website. Next,in the report menu section, the website will display a record page for violations that have been previously arrested.

2). Website Appearance

In this section, we will explain the design appearance of the museum security monitoring website. The website was created using the open-source framework Streamlit. The following is a view of the website design:

a. Website homepage display

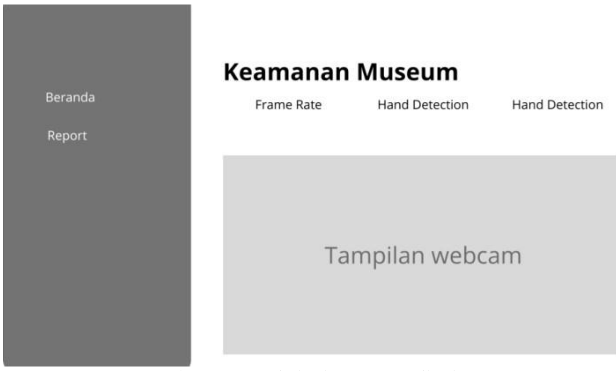


Figure 5. Website homepage display

b. Report page display

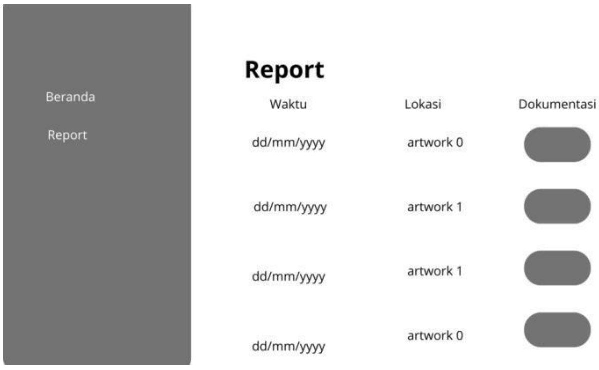


Figure 6 Report page display

On the website there is no log in menu and there are two menu displays, namely Home and Report. Each menu has its own uses which can be explained as follows:

1. Home page. On the home page there is a real-time webcam display as in Figure 5 above.
2. Report page. This report page serves to view records of violations (hands in the painting area) that have occurred

previously as in Figure 6 above.

III. RESULTS AND DISCUSSION

In this sub-chapter, the results that have been designed will be explained, including tools, websites and Telegram. The following are the results of the design of tools and websites used in the painting art museum security monitoring system.

A. Tool Results

The results of the tool that has been made based on the mechanical design are shown in the Figure 7, 8, 9.

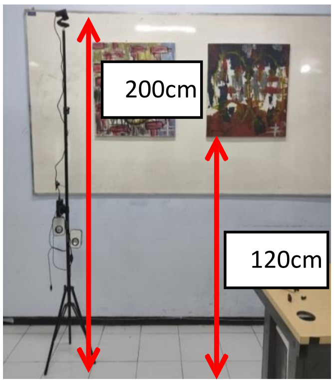


Figure 7 Front view

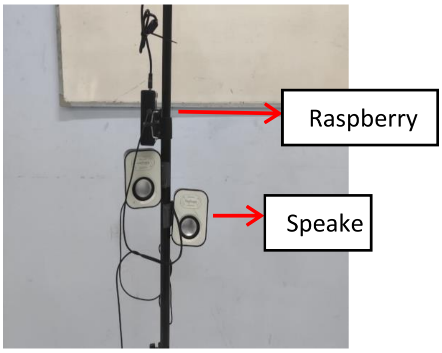


Figure 8 Speakers on the Mechanic

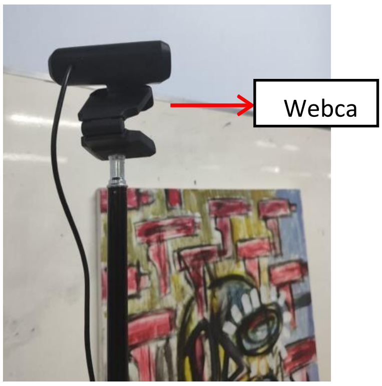


Figure 9 Camera on Mechanics

Painting placement specifications:

1. Two paintings are attached to the wall 120cm above thefloor level.
2. The distance between painting 1 and painting 2 is not tooclose, so a distance of 10cm is taken.

B. Optimal Camera Distance and Location

This test was carried out to measure and see the accuracy of this system. Tests on museum security were taken from several positions and distances. This test was carried out based on three distances, namely 1 meter, 2 meters and 3 meters. And five positions, including angles 0˚, 45˚, 90, 135˚, and 180˚. To determine the optimal distance and location, you can look at the violation detection results in the form of large and small percentage values. The best value will be taken as the optimal distance and location. The following can be described as this test was carried out.

1. Position the camera in relation to the painting

The position of the camera in this painting is taken from the midpoint between painting 1 and painting 2 as the center point. Then from the center point, a meter line is measured and drawn according to the distance to be tested. This research is tested from a distance of 1 meter, 2 meters and 3 meters so that it will produce the optimal camera placement point. The following is an illustration of the location of the camera in the painting.

View angle 0˚

The first position is taken at an angle of 0˚. The following is a depiction of the location of the camera at an angle of 0˚ to the painting, as shown in Figure 10.

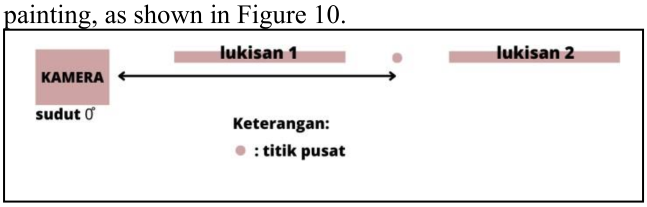


Figure 10. Depiction of angle 0

45˚ angle view

The second position taken is at an angle of 45˚. The following is a depiction of the location of the camera at a 45° angle to the painting, as shown in Figure 11

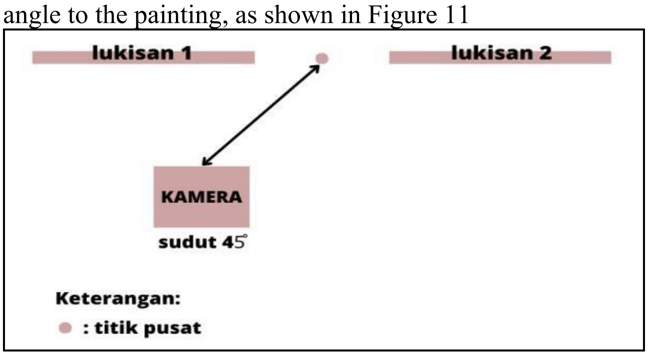


Figure 11. Depiction of a 45 angle

90° angle view

The third position taken is at an angle of 90°. The following is a depiction of the location of the camera at a 90° angle to the painting.

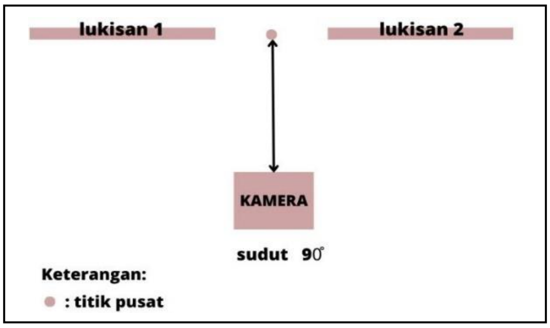


Figure 12. Depiction of a 90° angle

135˚ angle view

The fourth position taken is at an angle of 135˚. The following is a depiction of the location of the camera at an angle of 135° to the painting, as shown in Figure 13.

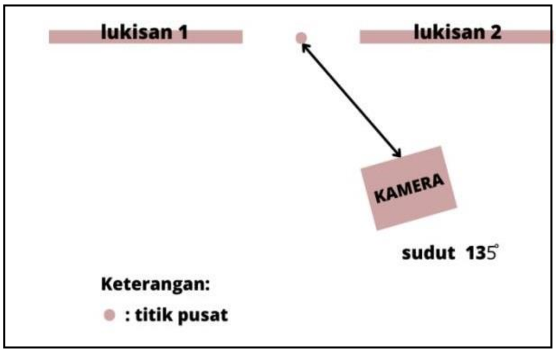


Figure 13 Depiction of the 135 angle

180° angle view

The fifth position taken is at an angle of 180°. The following is a depiction of the location of the camera at a 180° angle to the painting, as shown in Figure 14

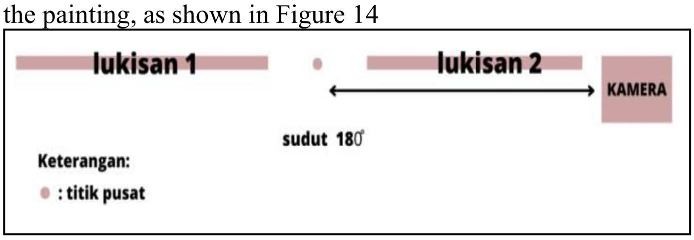


Figure 14. Depiction of a 180° angle

1) Optimal Camera Distance and Location test results

The tests that have been carried out by taking a distance of 1 meter, 2 meters and 3 meters and taking five positions each, namely 0˚, 45˚, 90˚, 135˚, and 180˚, produce percentages as in Table 4.1 and Table 4.2 below. Testing produces accuracy from the website and Telegram for each painting. This is due to the limitations of the Raspberry Pi 4B GPU in sending data so that the accuracy results obtained can be different.

Based on the optimal location and distance test table above, you can see the average value of each test. Starting from the optimal distance and location test table in painting 1, the percentage value is 67.2% - 68.8%. For the average presentation value of the optimal distance and location in painting 2, the percentage average value is 63.6% - 62.9%. It can be seen that the values for optimal location and distance in painting 2 are more than sufficient. The optimal location and distance for painting 1 is higher so the bounding box detection reading is better, as shown in Table I.

TABLE I
PAINTING TEST RESULTS 1

No	Gap (meter)	Location	Accuracy Result (*Website*)	Accuracy Result (Telegram)
1		0˚	95%	85%
2		45˚	82%	95%
3	1	90˚	91%	91%
4		135˚	81%	88%
5		180˚	0	0
6		0˚	87%	88%
7		45˚	86%	87%
8	2	90˚	81%	92%
9		135˚	93%	84%
10		180˚	0	0
11		0˚	71%	71%
12		45˚	70%	77%
13	3	90˚	81%	81%
14		135˚	94%	94%
15		180˚	0%	0%
	Average		67.2%	68.8%

In the calculation results, the average presentation value of theoptimal distance and location for placing the camera on the two paintings is at a distance of 1 meter with an angle of 45°.

C. Duration of YOLOv5 detecting hand objects.

This test was carried out by measuring the duration of YOLOv5 in detecting images. When data processing reaches the exit of the bounding box, the system will emit a sound lasting 2 seconds. The detection results of the violation object will be seen in milliseconds. The duration of detection is calculated as follow , as shown in Table II.

TABLE II
DETECTION DURATION TEST RESULTS ON PAINTINGS

No	Gap (Meter)	Location	Approach Duration
1		0°	0,15 ms
2		45°	0,20 ms
3	1	90°	0,16 ms
4		135°	0,12 ms
5		180°	0 ms
6		0°	0,12 ms
7		45°	0,15 ms
8	2	90°	0,14 ms
9		135°	0,12 ms
10		180°	0,12 ms
11		0°	0 ms
12		45°	0,15 ms
13	3	90°	0,15 ms
14		135°	0,16 ms
15		180°	0 ms
Average			0,11 ms

Based on the results of testing the duration of painting 1 and painting 2, the average image processing until the bounding box appears, the accuracy reaches 0.11 ms and the duration of the sound after the accuracy appears is 0.11 ms. For detection, it is quite fast using a Raspberry Pi 4B. However, if you use a computer for computing, the detection duration can be below the average detection using Raspberry Pi 4B computing or in other words faster than using Raspberry Pi 4B. This is because the Raspberry Pi 4B does not have a GPU and YOLOv5 itself has a GPU, therefore there is quite a long delay but object detection is quite fast.

D. Testig the optimal location of the Access Point relative to the camera

In this test, signal strength was measured using the WiFi Analyzer application on a smartphone. This test aims to find out the optimal location of the internet or access point at what angle in the room. Here the test uses a smartphone hotspot as an internet tool. Testing is carried out by placing the Access Point or hotspot at a distance of 1 meter and at 5 angles, including 0˚, 45˚, 90˚, 135˚, and 180˚. The following is an illustration of testing the optimal location of the Access Point relative to the camera , as shown in Figure 15.

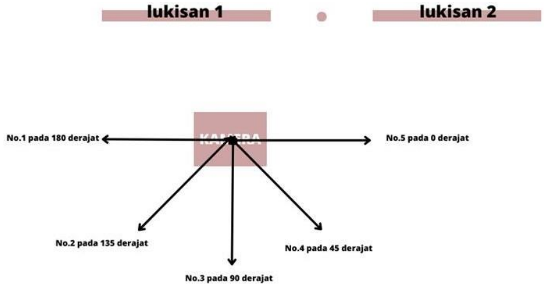


Figure 15. Depiction of testing the optimal location of the Access Point relative to the camera

Based on the depiction in the image above, the signal strength and speed are shown in Table III.

TABLE III
OPTIMAL ACCESS POINT LOCATION TEST RESULTS

No	Gap (Meter)	Locations	Sinyal Strength (dBm)	Speeds (Mbps)
1		0°		192 Mbps
2		45°	- 59 dBm	192 Mbps
3	1	90°	- 36 dBm	192 Mbps
4		135°	- 51 dBm	192 Mbps
5		180°	- 52 dBm	192 Mbps

In the test results for the location of the access point, the most optimal point was taken, namely a distance of 1 meter with a 90° angle to the device because the signal strength has very low noise. The speed obtained from several locations taken is192 Mbps.

E. Website and Telegram system functionality

This functionality test was carried out in three conditions, namely the condition that the hand was detected, objects other than the hand, and the hand was not detected. Table IV are the detection results in several conditions:

TABLE IV
CONDITION : THERE ARE HANDS

Condition	Website	Telegram	Speaker Condition	Reports
Hand Detection	Succeed Detection	Succeed Detection	ON	Get into trouble

In Figure 16, it can be seen that there is a hand in the painting..

Figure 16. There is a hand at a distance of 1 meter from the painting to the camera.

Below in Figure 17 you can see the detection of a hand in the painting at a distance of 2 meters from the painting to the camera and in the area of the painting that has been detected by YOLOv5.

Figure 17. There is a hand in the painting at a distance of 2 meters from the camera.

Below in Figure 18 you can see the detection if there is ahand in the painting with a distance of 2 meters between the painting and the camera at 180° and in the area of the painting that has been detected by YOLOv5.

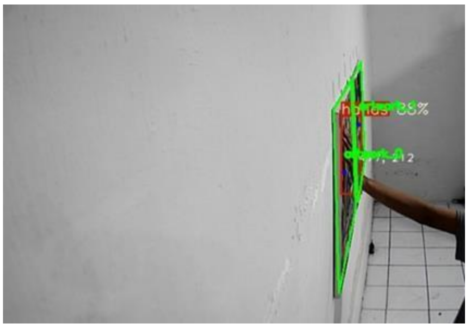


Figure 18 There are hands in the painting at a distance of 2 meters from the camera.

If a hand is detected in the painting area, the system will display a warning notification via the website and a message to Telegram which will then be followed by a warning sound on the speaker. Violations will be recorded on the report pageas in Figure 19 below.

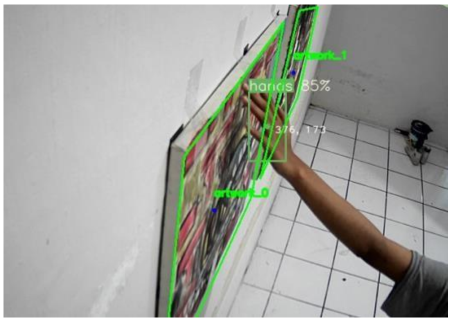


Figure 19. Entry Detection in Reports

Reports will be recorded sequentially from the earliest to the latest detection. In Table V, logs can also be deleted so that there is not too much data stored.

TABLE V
CONDITION OF OBJECTS OTHER THAN HANDS

Condition	Website	Telegram	Speaker Condition	Reports
Things other than hands	No Detection	No Detection	OFF	Didn"t enter

Figure 20. Objects other than Hands

Figure 20 and Table VI are the detection results if the condition of an object other than a hand, for example a person is passing by, the system does not read the object. Then the system does not provide output.

TABLE VI
CONDITION NO HANDS

Condition	Website	Telegram	Speaker Condition	Reports
Things other than hands	No Detection	No Detection	OFF	Didn"t enter

The following are the detection results if there are no hands in the painting area. This condition does not produce any output.

F. Testing the Effect of Light on Object Detection Results

This test is carried out by detecting in a room with varying light intensity. The variation value of light intensity is measured using a luxmeter measuring instrument. Detection is carried out by placing the camera on the painting 1 meter away with an inclination of 45°. The following are the results of testing the effect of light on object detection results:

TABLE VII
RESULTS OF TESTING THE INFLUENCE OF LIGHT ON OBJECT DETECTION RESULTS

No	Brightness Value (LUX)	Detection Results
1	0	No Detection
2	30	Succeed Detection
3	13	Succeed Detection
4	16	No Detection

Based on the results of the table on the influence of light on detection results in this system, it can be seen that the LUX brightness values are detected and not detected. Starting froma table testing the effect of light on painting and hand detection results. Where the test was carried out 4 times and it was detected 2 times it was not detected and 2 times it was detected. So in testing the influence of light on the detection results, this is very influential. If the color brightness level is too low, it will be difficult for the YOLOv5 algorithm to be carried out, otherwise if the brightness level is too high, it will be difficult for YOLOv5 to read objects. However, in theresults of this test there were false positives, where at a brightness value of 13 lux it was successfully detected and at a brightness value of 16 lux it was not detected.

Figure 21. Detection results for brightness value 16

In Figure 21 above is the detection result with a brightness value of 16 lux. This can be caused by the color of the hand being confused with the color of the painting so that the camera cannot capture the image properly. As well as cameras that have deficiencies in image sharpness.

IV. CONCLUSION

The results of this trial on detection for museum security showed that the level of accuracy in detection had an average of 67.2% - 68.8% in painting 1 and an average of 63.6% - 62.9% in painting 2. The difference in the average accuracy value was taken by distance. detection between the two paintings. The percentage value indicates the extent to which the model is confident that the identified object is within that bounding box. It was proven that the slopes of 0 and 180 at a distance of 3 meters could not be detected due to the limited sharpness of the image captured by the camera. Meanwhile, a detection distance of 1 meter to 2 meters can still be detected by the camera even though it takes quite a long time. Based on the results of the violation detection time, it can be observed that the museum security monitoring design system is more efficient because the average detection duration is 0.11ms. The duration results can be even faster if good computing is used for detection so that the FPS display does not run slowly on the system. Then, the website and Telegram functionality section shows that when there are hands in the painting area, the YOLOv5 algorithm will detect the violation object (hands in the painting area) and send three outputs, namely website notifications, Telegram messages, and sound notifications. Then, if there are other objects in the painting object, the YOLOv5 algorithm does not detect the presence of hands, then the three outputs do not work.

REFERENCES

[1] S. M. Metev and V. P. Veiko, *Laser Assisted Microtechnology*, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[2] J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.

[3] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, “A novel ultrathin elevated channel low-temperature poly-Si TFT,” *IEEE Electron Device Lett.*, vol. 20, pp. 569–571, Nov. 1999.

[4] J. Liang and J. Chen. “Joint Relay selection and Network coding for Error -Prone Two-Way Decode-and-Forward Relay Networks,” *IEEE Transactions. Commununications*, vol. 31, no. 3, pp. 476–488, July 2013.

[5] T. Vu, P. Duhamel, and M. Renzo, “On the Diversity of Network-Coded Cooperation with Decode-and-Forward Relay Selection”, *IEEE Trans. Wirel. Commun.*, vol. 14, no. 18, pp. 4369-4378, 2015.

[6] Y. Jiang, I. W.-H. Ho, Z. Sattar, Q. F. Zhou, and F. C. M. Lau, “Paired-relay-selection schemes for two-way relaying with network coding,” *IET Commun.*, vol. 9, no. 6, pp. 888–896, Apr. 2015

[7] M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, “High resolution fiber distributed measurements with coherent OFDR,” in *Proc. ECOC’00*, 2000, paper 11.3.4, p. 109.

[8] X. Liu, G. Wang, and B. Li, “A Physical-network-coding-based amplify and forward cooperation scheme,” *ICICS 2013 - Conf. Guid. 9th Int. Conf. Information, Commun. Signal Process.*, Nov. 2013, pp. 1–5.

[9] W. Yang and X. Zhao, “Robust Relay Selection and Power Allocation for OFDM-Based Cooperative Cognitive Radio Networks,” in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2016, pp. 1–6.

[10] R. E. Sorace, V. S. Reinhardt, and S. A. Vaughn, “High-speed digital-to-RF converter,” U.S. Patent 5 668 842, Sept. 16, 1997.

[11] *FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.

[12] “PDCA12-70 data sheet,” Opto Speed SA, Mezzovico, Switzerland.

[13] A. Karnik, “Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP,” M. Eng. thesis, Indian Institute of Science, Bangalore, India, Jan. 1999.

[14] J. Padhye, V. Firoiu, and D. Towsley, “A stochastic model of TCP Reno congestion avoidance and control,” Univ. of Massachusetts, Amherst, MA, CMPSCI Tech. Rep. 99-02, 1999.

[15] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std. 802.11, 1997