Comparative and Evaluation of Anomaly Recognition by Employing Statistic Techniques on Humanoid Robot

Nuratiqa Natrah Mansor[1], Muhammad Herman Jamaluddin[2], Ahmad Zaki Shukor[3], Muhammad Sufyan Basri[4]

Faculty of Electrical Engineering, Universiti Teknikal Malaysia Melaka, Durian Tunggal, Melaka, Malaysia[1, 2, 3]

MAP2U Sdn Bhd, PT9951, Jalan BBN 1/3K, Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia[4]

Abstract—This paper presents the study to differentiate between normal and anomaly conditions detected by humanoid robots using comparative statistics. The study has been conducted in robotic software as a platform to examine the scenario and evaluate between the anomalies and normal behaviour in different conditions. This study employed a machine vision technique to run an image segmentation process and carry out semi-supervised object training within a controlled environment. The robot is trained by differentiating the measurement size of the target object, its location, and the object's visibility within three different frames. The effect is measured by extracting the positive predictive value (PPV) value, mean and standard deviation value from the captured image using statistical techniques in machine vision. The results showed that the mean value decreased by around 50% from the normal scenario when an anomaly occurred. Aside from that, the standard deviation values were more than twofold compared to the common scenario, especially after the object's size grew. In contrast, the deviation value is remarkably small when the target is situated in the middle of adjacent frames, compared to the value when the entire shape is positioned in the frame. Simultaneously, the mean values from the processed image produced a minor difference.

Keywords—Anomaly detection; humanoid robot; vision system; statistical analysis; robot recognition

I. INTRODUCTION

Over time, the application and comprehension of machine vision and robotics research have become increasingly valuable and helpful, especially for humans. Among the numerous studies, one that has piqued researchers' interest is applying picture segmentation approaches for spotting objects and people in various settings and orientations. Hence, it is scientifically referred to as anomaly detection. Anomaly is also defined as a low likelihood or probability of occurrence in a given setting. To put it simply, anomalies occur when the item's state deviates from the standard or customary order [1,2]. Certain anomaly detection studies generally implemented computational intelligence knowledge, including fuzzy logic-based outlier identification and neural networks.

Moreover, this detection algorithm employs machine learning and statistical detection methods, as the subject's previous behaviour or trends need to be learned and interpreted [3,4]. It may also be utilized for practical purposes in product quality control, enhancing national security, and military purposes. Anomaly detection might also be used for mobile robots and unmanned aerial vehicles (UAV). These robots can detect anomalous behaviour and sceneries in various environments. Among them, the humanoid robot is a type of mobile robot. The robot can be programmed to use picture segmentation and processing techniques to determine the output pixel value and identify anomalous behaviour.

In anomaly detection, there are several issues that need to be considered, such as objects' features and illumination circumstances [5], dense 3D shape models, image resolution, objects in different frames, and many other things [5-7]. Although some of the methodologies and procedures worked well under specific conditions or requirements, the accuracy for distinguishing anomalous behaviour has significant degradation due to different issues. Anomaly detection is perhaps one of the promising and fascinating fields to study.

Among the issues, the issue of recognizing static objects when located in different positions is addressed in this study. Apparently, the issue occurs when there is confusion regarding the identical frames and variable object sizes. Besides that, the issue also arises when an object is placed at the boundary or when the object is located in a different position or corner in the frame. The object may not appear identical to its predecessor as it holds the corresponding pixel value in every scenario. Mean values and deviation of pixels value also vary between photos. Hence, this research focuses on utilizing humanoid robots by integrating machine vision studies to detect any irregularities in their view and classify them using statistical techniques. This project was created using the CoppeliaSim and Python (x,y) programming to demonstrate the method and simulation of how the work functions.

The next section will delve into the broad idea of machine vision with the theories of robotics vision and humanoid vision as part of the component with the block diagram of necessary processes in the system, including techniques for learning in anomaly conditions. Then in the methodology part, the procedures and methodologies performed on each experiment are presented in detail. In the following section, all recorded data and information observed from every single experiment are tallied and illustrated in graph form. Next, Section V will discuss and evaluate the outcome after the simulation of each experiment to support the hypothesis and proposed method. Meanwhile, Section VI summarizes the key

findings and conveys realistic recommendations for improvements in future works.

II. Literature Review

A. Machine Vision

Visual exploration in unknown settings is a common and crucial job for intelligent automation in the present day. One of the most compelling aims for machine vision applications is for robots to be able to search for identifiable things in an unknown environment; even though some robots have a deficiency in visual skills because of unable to compete with human ability, particularly in brains and eyesight [8]. Regardless, it may still be explored and improved regularly with various concepts and theories, especially with the continuation of innovative technology.

Machine Vision is a "simple" processing system that captures visual information by processing, analyzing and measuring the image attributes using a combination of hardware and software. This is used to extract the information utilized in decision-making in order to generate the required result. The application of vision has been developed into a software platform to make people unnecessary spending an extended time developing algorithms for image processing [9].

Furthermore, the platform is user-friendly and allows for a simple extraction and modification of image information. Some of the examples are HALCON from MVTec Software, IMAQ from National Instruments, Visilog from Norpix, and PatMAX from Cognex Corporation [10]. It may sometimes surpass humans in specialized tasks. As shown in Fig. 1, the practical strategy for recognizing objects in vision systems mainly uses local characteristics as a primary tool to match scene material to models of recognized items.

B. Robotics Vision

Robots are growing increasingly capable in line with the decade's developments because they can execute a growing number of activities that would be too risky or difficult for humans. Flexibility has been required as job complexity has risen. The need for adaptability has grown in tandem with the complexity of modern jobs. The creation of humanoid robots gives way to the establishment of societies of robots that collaborate with humans as humanoid robot technology advances [9,10]. The evolution of robots goes from single and tiny robots to the establishment of humanoid and gigantic robots that can work together with humans.

The application of vision is becoming a critical element in robotics and machine automation. Hence, the studies have been implemented in robotics advancement, known as robotic vision. In standard function, robot vision is utilized for part identification and navigation. In advanced function, the vision enables to lead and guide the robots or other automated machines to recognize, think and perform the task, resembling a human [11,12]. By controlling the special machine vision software, it can make the part identification and replace or complement manual inspections, navigation and measurements with digital cameras and imaging processes. Then the controller will automate and coordinate a necessary

process to perform the task. The general process can be seen in Fig. 2.

C. Humanoid Vision

Ideally, the vision has been applied to many types of robotics and machines, not being left behind in humanoid robots. With the intelligent vision system used in humanoids, the robot can interpret images in terms of colour, size, shape, and position of the object on the scene. In addition, both sights of a humanoid robot can recognize various objects with different traits. The image will undergo pre-processing and image segmentation stage before continuing with object matching and tracking [13, 14]. An intelligent system is also able to perform a high-level recognition task. Some of the essential elements in the automated analysis and extraction of the required information are shown in Fig. 3.
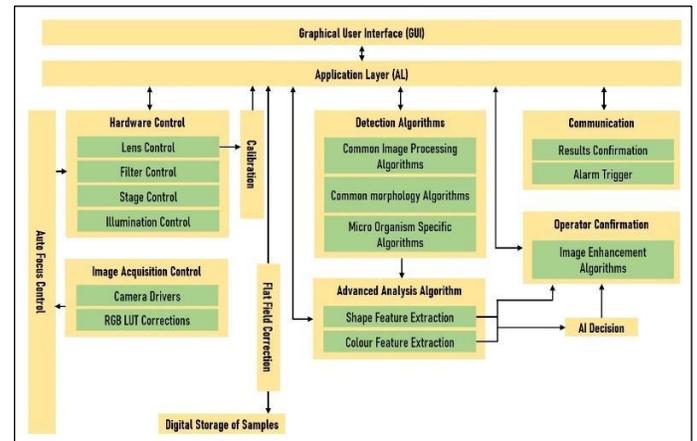

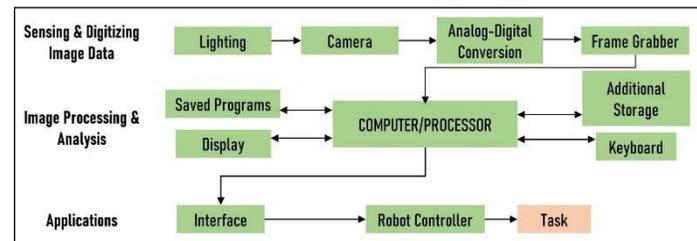Fig. 1. Machine vision software architecture.

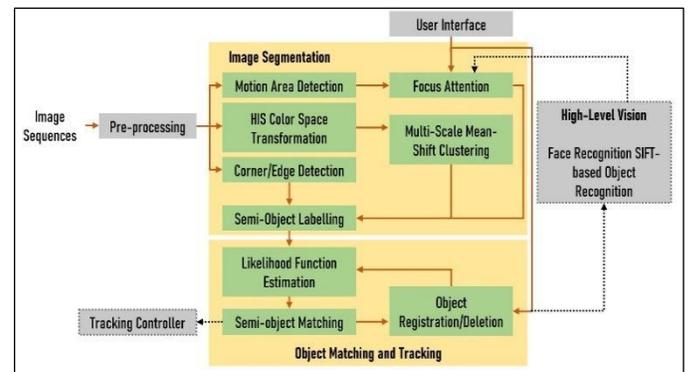Fig. 2. Machine vision process with three important components.

Fig. 3. Structure of the robotic vision system.

D. Methodology Used

Anomaly detection can occur in different surroundings and circumstances, depending on what abnormal behaviour are set. Therefore, anomaly detection methods are classified according to their input type, output type, learning method and the type of target anomaly. There are three types of learning methods in anomaly detection, which are supervised, semi-supervised and unsupervised [9,10].

For supervised learning, data sets labelled with normal or anomalous are used for training the model. While for semi-supervised, the model is trained with only normal data. The data points that do not match normal behaviour are considered anomalous during the detection. Differ for unsupervised learning case, no training data required. The input data is expected to contain a high ratio of normal to anomalous points.

Alternatively, the model will consider frequent anomalous behaviour as a common condition [15]. As mentioned earlier, the main objective of this research is to detect an anomaly behavior of non-motion objects via a humanoid robot. Besides that, the anomaly behavior in different frames is also observed using the single camera attached to the robot. The experiment is handled in a roomy space in which all the variables are controlled, fixed and not moving.

Thus, the method highlighted earlier is reasonable and much more apposite to be carried out in this project for anomaly behaviour as it can detect objects even though in a different position of a frame. First, this method is tested on the same base background and light intensity with different object sizes. After that, the outcome from each experiment and observation is turned into the percentage of accuracy, precision, and limitations. Finally, the data is discussed in the analysis and result section.

III. METHODOLOGY

Generally, this study aims to let the humanoid robot learn to detect and identify objects in anomaly behavior. In order to observe and evaluate the performance of object identification and anomalies cases, a series of experiments is conducted to prove the object's behaviour in different cases. Primarily, specific aspects and parameters need to be determined and controlled to maintain the accuracy of the experiments. As an example, lighting and frame background. Originally, the methodology to conduct the experiment was broken down into two stages. The first one is the system design phase which discusses about the software and type of simulation used to run the experiment. The other stage explains how the methodology is implemented, where the details on how experiments are handled to get reliable results.

A. System Designation

The experimental platform to work for this project is Python (x,y) for Windows and CoppeliaSim software. This study is developed and runs through the simulation process by these two softwares. The coding code for the programming is developed in Python. At the same time, the environment scene and robot used for the experiment are built in a robotic simulator, CoppeliaSim. These two softwares are linked to each other and act as terminals to show the overall simulation and experiment process.

In overall, this project undergoes four different types of experiments. The experiments are semi-structured scenes, calculating the mean and the standard deviation values based on different sizes of the object, determining the location and position of the object from the robot and lastly, collecting the mean and standard deviation values based on the different positions of an object in one frame.

B. Stereo segmentation

Interest points that are matching in the left picture are divided into groups based on their location and discrepancy. Then, a greedy clustering approach is applied, starting with a single interest point. After that, if the x-position, y-position, and disparity are within the 5-pixel threshold of any neighbouring cluster member, it is supplemented with new ones. The segmentation usually occurs at a comparable depth but in a different spatial position or at a nearby spatial location but at a different depth. Fig. 4(a) and 4(b) showed the pattern of the process.

To enhance the segmentation step, feature locations are divided into on-object and outside-object groups (white vs other colours). According to the above image in Fig. 4(a) and 4(b), every bit of outside-object locations and any image pixels outside the bounding region are considered outside-object pixels. This region is emphasized in black colour after the segmentation process. Following that, the location for the on-object is highlighted in grey colour. To indicate that the particular shade is an anomaly, an image's pixel value is generated in histogram form so that it is easier to compare between the normal and abnormal or determine the different cases for every anomaly.

C. Anomaly Criterion

While learning the object samples, two sorts of anomalies can be identified; For starters, the event patterns that never have been seen before and employed as variable for classification. The flow is shown in Fig. 5 below. The anomaly scenario is defined as samples that deviate significantly (in terms of distance) from other clusters in the model [16]. Secondly, the model contains exemplars of other sorts of anomalies. It does, however, have a minimum frequency this time. As a result, it might also be characterized as a false positive. Both forms of abnormalities may be discovered by taking distance and frequency into account to classify.
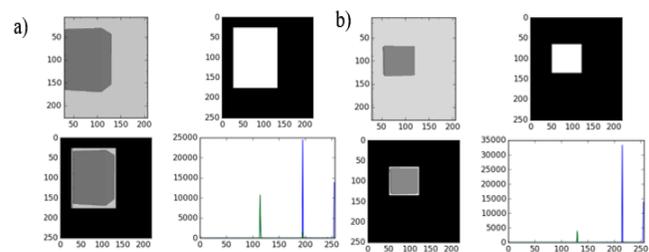

Fig. 4. Image segmentation, where the feature for matching exists, is clustered and plotted into the histogram.

According to equation (1), an intra-class distance Dist (C_i) in each cluster is defined as the greatest distance between any cluster member and the cluster centre.

$$Dist\ (C_i) = jmax\|f\left(e_j^{C_i}\right) - f(C_i)\|_2 \qquad (1)$$

Thus, as long as the abnormal grid event is detected, the frame triggered the condition as abnormal, either in significant or small deviations [16]. The deviation is calculated by using a simple formula. In this study, the deviation is extracted using the plug-in Open CV and Numpy Software in Python (x,y).
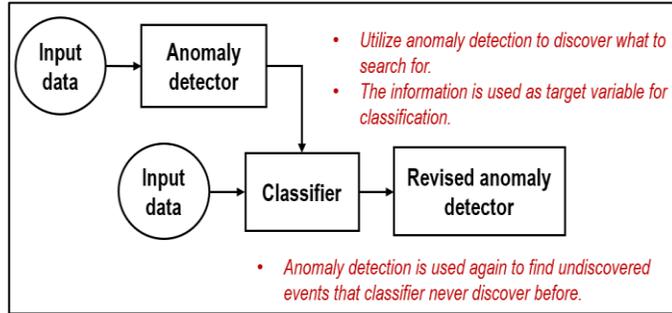

Fig. 5.   Example of flow process of anomaly behavior for classification.

D. Research Procedure

1) Experiment A: Semi-supervised Scene: This experiment investigates the object's behaviour and the pixel number when it is located at a different position in the same black background frame. The object is fixed in a triangular shape and red colour. The output results were positive predictive value (PPV), which is technically the pixel data extracted from the image captured. In addition, intra class-distance is calculated manually to find the percentage of anomaly that can happen. Finally, the value is compared with the smallest distance value to identify true-positive and false-positive anomalies.

2) Experiment B: influence on the object size: This experiment measures the performance and effect of the variant size of an object on the system. The object that has been used is a cuboid with the same colour as the previous experiment. The frame background is fixed to a grey colour to differentiate it from black. Black is the default colour for the surroundings, as seen on the vision sensor recorded in the CoppeliaSim. There are three 50 cm-width and 80 cm-high walls labelled as Frame 1, Frame 2 and Frame 3. Frame 1 is set at the origin at 180˚. All frames have the same surface, colour and size.
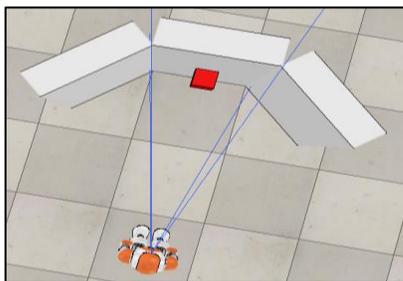

Fig. 6.   The top view of the experimental setup is placed ahead of the NAO robot.

Meanwhile, Frame 2 is located at -40˚ and deviates from the horizontal axis. Meanwhile, Frame 3 has deviated as much as 40˚ from the horizontal axis. The setup can be seen in Fig. 6 above. The parameter for the object's size is set from 0.4 cm x 0.4 cm up to 45 cm x 45 cm. This dimension is for Frame 1. However, for Frame 2 and Frame 3, the object size added an extra size of 50 cm x 50 cm. This is because both frames need a larger cuboid surface to cover up all areas of the wall frame.

3) Experiment C: detection of object's location: This experiment is conducted to locate the anomaly by calculating the distance and angle between the robot and the wall screen. The experiment scene setup is still the same, as shown in Fig. 6. In this experiment, the cuboid's maximum size was not identical to the cuboid in Experiment B. The maximum size of the cuboid for Frame 1 and Frame 3 is 20 cm x 20 cm, while for Frame 2 is 25 cm x 25 cm, depending on all cuboid sides that are possible to be fitted in a frame. Using the formula of Pythagoras' Theorem and Sine Rule, the location of the cuboid in each frame can be ascertained. The experiment is continued with the same procedure in each particular frame.

4) Experiment D: impact on various placement: From the result obtained in experiment B, only one size of a cuboid is selected as a fixed parameter in each frame. The cuboid size is picked based on the output mean and standard deviation difference between normal and abnormal conditions. For this time, the middle range of difference is selected. The cuboid is placed independently from 0 cm to 50 cm along the wall. Then, the effect of object position on the system is measured from the data attained in the mean and standard deviation of the pixel value. The experiment is continued with the same procedure in each particular frame. The difference between the normal and anomaly condition is observed based on the data outcomes.

IV.   RESULTS AND DISCUSSION

A. Semi-Structured Scenes (Semi-Supervised)

For the early experiment, only one frame is used despite the different positions of the target. The reason for this is to observe the performance of the target object. In this case, a triangular object is placed on the screen. The black background colour is chosen to make the number of pixels easy to calculate, as this experiment has not yet applied the image segmentation process. The object has been used as the target and labelled by 'A'. Even though the object is in different positions on the frame, 'A' is successfully recognised and marked as normal behavior. The series of the captured image by the robot vision is shown in Fig. 7.

Normal scenes are classified whenever the object is still inside the frame range, not at the edge or outside the boundary. From the early hypothesis, a subject is marked as an anomaly when it is located at the frame's edge. The reason for this is that the frame's pixel value is counted differently than it is supposed to be. However, depending on the situation and output data, it can be categorized as a false positive or a true positive.
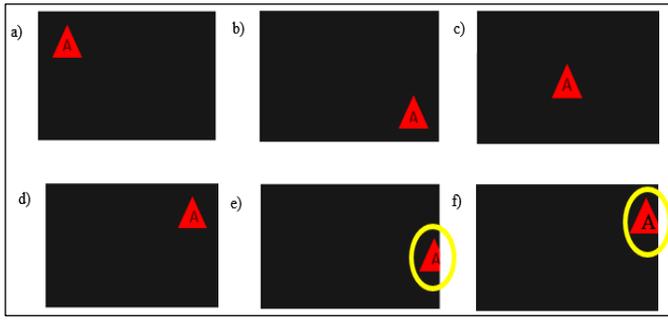
Fig. 7. A collection of object samples observed by the camera in a particular circumstance. Images (a)-(d) depict normal behaviour, whereas images (e)-(f) depict anomalous behaviour.

Precision or positive predictive value (PPV) is the ratio of the number of object pixels that are correctly recognized object pixels to the number of detected object pixels. When the ratio computation is near one, the performance is good. It is also known as the True Positive Rate (TPR), where the ratio of the number of object pixels correctly acknowledged to the number of object pixels in the ground truth. These quantitative analyses can be seen in Table I.

TABLE I.   PPV VALUE OF OBJECT SAMPLES IN DIFFERENT POSITIONS

Sample	1st	2nd	3rd	4th	5th	Average PPV Value
a	0.95	0.96	0.97	0.96	0.96	0.96
b	0.94	0.95	0.94	0.94	0.94	0.94
c	0.98	0.99	0.98	0.99	0.98	0.98
d	0.95	0.94	0.95	0.95	0.95	0.95
e	0.67	0.67	0.67	0.67	0.67	0.67
f	0.87	0.88	0.89	0.89	0.89	0.89

$$PPV = \frac{number\ of\ object\ pixels\ is\ correctly\ identified}{number\ of\ object\ pixels\ detected} \quad (2)$$

By calculating the mean, variance and standard deviation using the formula follows:

1) Mean, µ

$$\mu = \frac{\sum X}{N} \quad (3)$$

$$\mu = \frac{0.96+0.94+0.98+0.95+0.67+0.89}{6}$$

$$= 0.8983$$

2) Variance, σ^2

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} \ (x_i - \mu)^2 \quad (4)$$

$$\sigma^2 = \frac{1}{6}[(0.96 - 0.8983)^2 + (00.94 - 0.8983)^2 +$$
$$(0.98 - 0.8983)^2 + (0.95 - 0.8983)^2 +$$
$$(0.67 - 0.8983)^2 + (0.89 - 0.8983)^2$$

$$= 0.0112$$

3) Standard Deviation, σ

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \ (x_i - \mu)^2} \quad (5)$$

$$\sigma = \sqrt{0.0112} = 0.106$$

4) Intra - class distance, Dist (Ci)

$$Dist\ (C_i) = jmax\|f\left(e_j^{C_i}\right) - f(C_i)\|^2 \quad (6)$$

$$Dist(Ci) = \|0.67 - 0.945\|^2$$
$$= 0.0756 \approx 0.08 \quad (for\ case\ e)$$
$$Dist(Ci) = \|0.89 - 0.945\|^2$$
$$= 0.0030 \quad (for\ case\ f)$$

5) Smallest distance between f(ex) and f(Cx)

$$d(e_x, C_x) = 0.98 - 0.945$$
$$= 0.0035$$

The parameters that have been calculated above show that the value of standard deviation, σ ≈ 0.1 and intraclass distance is 0.08. Thus, it shows that the percentage of anomalies that can be occurred is very low. To prove the anomaly in Experiment 1, the calculated value from d(ex, Cx) is less than Dist (Cx), which shows that abnormality in Fig. 7(e) has happened. However, the value of Dist (Cx) and d(ex, Cx) in Fig. 7(f) is much alike. Therefore from Fig. 7(e) and Fig. 7(f), two out of six cases are abnormal. Fig. 7(f) can be labelled as a false positive as the PPV value is very near to the PPV reading of ordinary cases. Meanwhile, the image in Fig. 7(e) is quite distant from the mean value and other ordinary cases. Thus, Fig. 7(e) can be labelled as true-positive.

Aside from that, the shape in Fig. 7(e) showed a bigger proportion surpassing the boundary of the frame compared to the smaller proportion of the object in Fig. 7(f). The percentage of the whole object in Fig. 8(f) is about 90%, but in Fig. 7(e) is almost 60%. Data that does not match normal behavior in semi-structured scenes is considered anomalous during the detection.

Based on experiment A that has been done, the pixel values are close to 1 despite the object's different position on a frame. The PPV value is significant when the object is at the edge of a frame. This value is categorised as an anomaly, whether false-positive or true-positive. False-positive is when the condition is an anomaly. Still, the output data is near the normal condition. At the same time, true-positive can be labelled when the condition is obviously an anomaly and slightly far from the deviation. Somehow, a false-positive can still be described as an anomaly. However, the percentage and accuracy are not much likeable and noticeable. For instance, the PPV value of Fig. 8(f) in the semi-structured experiment is 0.89. It has the lowest difference, with an overall mean value of 0.8983. Hence, the percentage of the object positively identified in Fig. 8(f) is almost 90%. Still, the image captured in Fig. 8(e) shows that the object has surpassed the frame edge at almost 40%. Based on the standard deviation calculation, σ the value is approximately 0.1, which is quite good for anomaly detection. It is because the range or frequency for an anomaly to happen is less and close to 0.
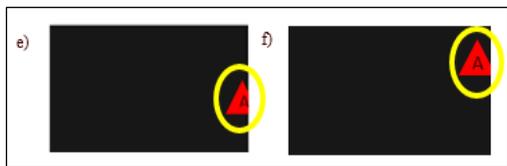
Fig. 8.   True-positive anomaly versus False-Positive anomaly.

B. Object Size

The main objective of this experiment is to observe the pattern of mean and standard deviation values of each image taken from the vision sensor humanoid robot. With the pattern, it can be deduced on what condition the frames can be labelled as abnormal or normal. First, the colour and shape of the object are fixed throughout the experiment to ensure the output result is clear and easy to analyse. The object chosen is a cuboid and red. To control the experiment and for comparison, the mean and standard deviation value for the normal frame is recorded first. The recorded values are tabulated in Table II below. A normal scene is when no object is placed in the corresponding frame area.

TABLE II.   MEAN AND STANDARD DEVIATION VALUES IN A NORMAL BEHAVIOR

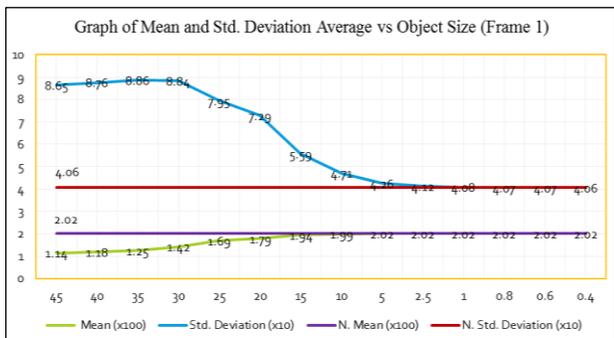| Frame Number	Data	Value
Frame 1 (Center)	Mean (x10^2)	2.02x10^2
	Std. Deviation (x10^1)	4.06x10^1
Frame 2 (Left)	Mean (x10^2)	2.20x10^2
	Std. Deviation (x10^1)	3.68x10^1
Frame 3 (Right)	Mean (x10^2)	2.06x10^2
	Std. Deviation (x10^1)	3.94x10^1

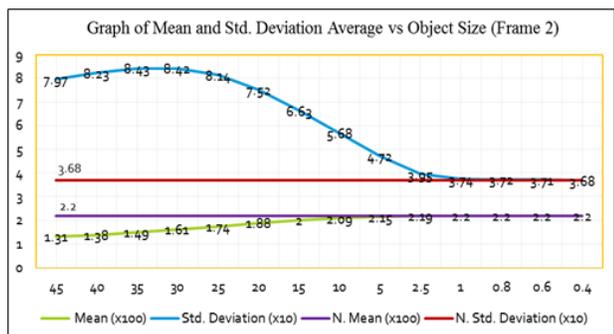Fig. 9.   Graph for mean and std. deviation average versus object size (Frame 1).

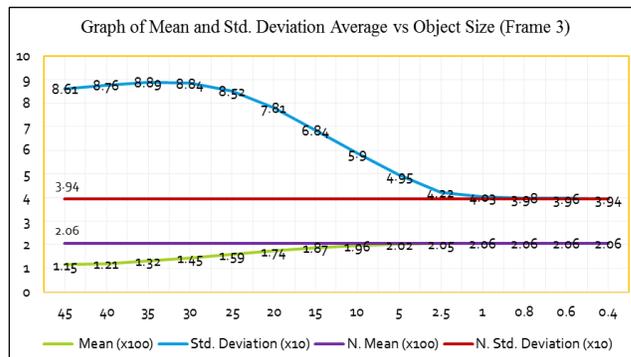Fig. 10. Graph for mean and std. deviation average versus object size (frame 2).

Fig. 11. Graph for mean and std. deviation average versus object size (Frame 3).

In Experiment B, only the object's size is manipulated. Other than that, the colour, shape and position are controlled during the experiment. The output data are plotted in the graph version, as shown in Fig. 9 to Fig. 11. The graphs show that the anomaly is identified when the cuboid size gets bigger than the typical size. The differences between the normal mean value and normal standard deviation value are very apparent too, with the biggest difference is recorded in Frame 3. The size for the cuboid reaches up to 50 cm x 50 cm and covers the whole frame wall. Moreover, anomalies also can be identified on the smaller object as long as the standard deviation value passes the threshold value. From the plotted graph above, cuboid size 1 cm x 1 cm is the minimum size detected as an anomaly in Frame 1. For Frame 2 and Frame 3, the smallest size is 0.8 cm x 0.8 cm. It is because when the object is smaller, the pixel value for the coloured object has no differences from the normal condition. Technically, it is considered normal behaviour, although the object is inside the frame. Therefore, it is known as a false-positive anomaly. More simply, pixel levels in the output image are too small and do not pass the normal frame value to be labelled as abnormal.

As for the bigger size, the pixel values for the coloured object are different from the plain frame as pixel intensity for the red colour is much higher than the grey one. This is also known as a true-positive anomaly.

C. Location of Object

In this experiment, the shape and colour of the cuboid are set not to change. However, the size is still different from the smallest to the biggest, which makes it possible to fit the cuboid size and calculate its position. The main objective of this experiment is to observe the location and position of the red cuboid image in a particular frame. With the pattern, it can be deduced which size of cuboid and position is suitable for humanoid robots to detect the anomalies. As stated earlier, the colour and shape of the object are fixed throughout the experiment to ensure the output result is clear and easy to analyze. For example, Frame 1 is still at the centre at 0˚ while Frame 2 is +40˚ and Frame 3 is -40˚.

Here, Fig. 12 below shows the accomplished segmentation process to locate only the inside region of the object (dark grey colour) and eliminate the rest of the outer region (black colour).
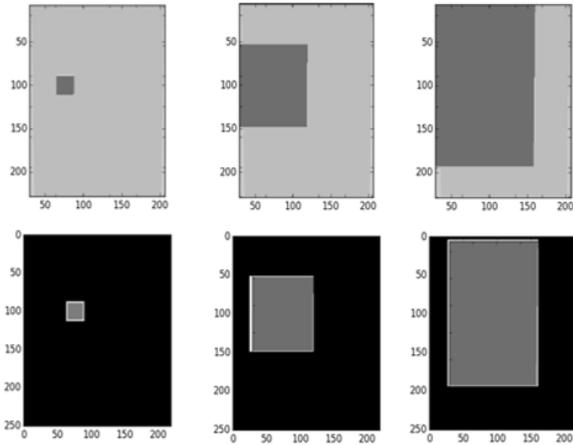
Fig. 12. Images captured by vision sensor for 5 cm x 5 cm, 20 cm x 20 cm and 35 cm x 35cm surface area.

Based on the calculation from each frame, it proves that the object's location from the humanoid robot is about 8˚31' on the left side of the frame wall. The value is calculated based on Pythagoras' Theorem and Sine Rule.
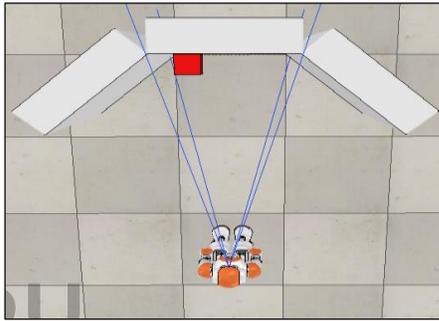

Fig. 13. Setting from the upper view between the wall frame and the robot to measure the angle.

The wall frame is 0.5 m in length. Starting from the centre until the left edge of the frame is 0.25 m. Then, the cuboid is set to be 0.1 m from the left side of the frame edge. So, the value of a = 0.15 m. Value b, the distance between the robot and the frame, is fixed at 1 m. To calculate c and x:

$$c = \sqrt{b^2 + a^2} \qquad (7)$$

$$c = \sqrt{1^2 + 0.15^2}$$

$$c = 1.01\text{m} \qquad \text{and,}$$

$$\frac{\sin X}{0.15m} = \frac{\sin\sin 90}{1.01m} \qquad (8)$$

$$1.01\text{m x } \sin X = 0.15\text{m x } \sin 90$$

$$X = 8\,˚31'$$

So, the cuboid is located 8˚31' between the centre and left of the robot. The outline can be seen in Fig. 13. Different situation with experiment C, the robot needs to locate the actual size and location of the anomalies on the frame wall after the anomaly behavior can be distinguished. Similar to Experiment B, the cuboid's colour, shape and position are fixed during the experiment. However, in this experiment, the

robot could spot the cuboid size location up to 20 cm x 20 cm for Frame 1 and Frame 3 and 25 cm x 25 cm for Frame 2. When the object is bigger than the limit size, the centre and edges of the object are not fully fitted in a particular frame. Therefore, the centre and edges of the cuboid are located differently and over the limit; even the standard deviation value differs from the normal behavior. This condition can be considered a true-negative anomaly. To make the robot detect bigger object than the limits, the size of the wall need to be bigger and extended. Using Pythagoras Theorem and Sine Rule, the actual location, right side and distance of the abnormal can be calculated. With this ability, anomaly detection can be improved as the robot can locate the object's location as long as it is still within the frame boundaries.

D. Position of Object

The main objective of this experiment is to observe the pattern of mean and standard deviation values in a particular frame despite the position not being fixed while the size is constant. With the pattern, it can be assumed that the best position is suitable for a humanoid robot to detect anomalies. From the results, the surface area of a 10 cm x 10 cm cuboid is chosen as the best measurement to get the median value. It also fits the normal frame and the other two slanted frames, regardless of various positions. As stated earlier, for making a comparison, the mean and standard deviation value for the normal frame is recorded first.
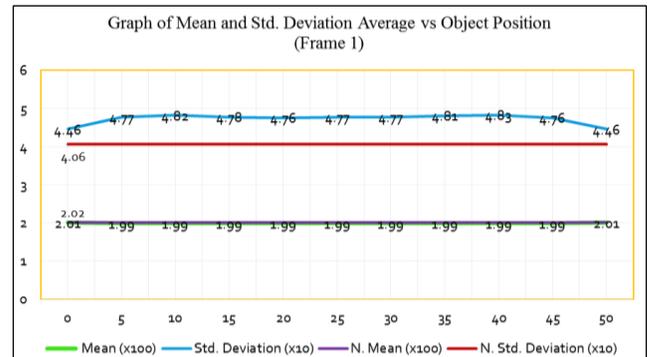

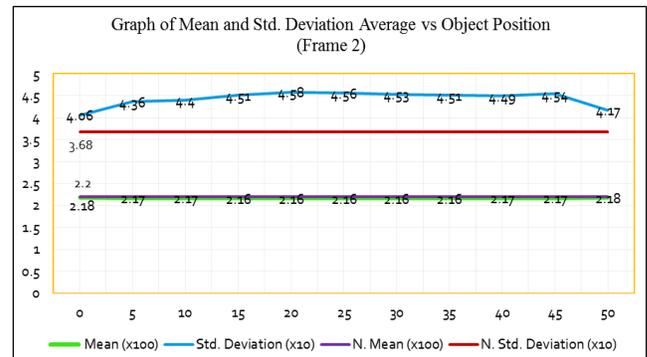Fig. 14. Graph of mean and std. deviation average vs object position (frame 1).

Fig. 15. Graph of mean and std. deviation average vs object position (frame 2).
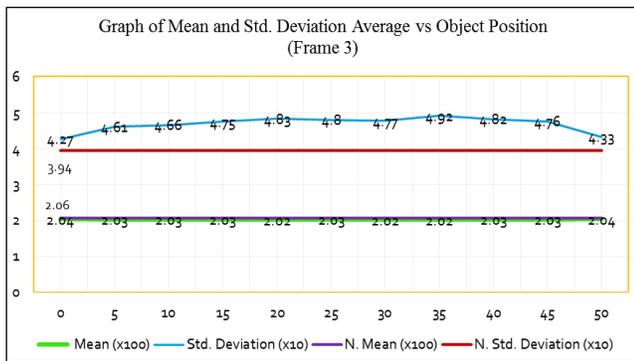
Fig. 16. Graph of mean and std. deviation average vs object position (frame 3).

In general, experiment D is quite similar to Experiment B. The changes in the scenario where the cuboid is placed at the halfway point but varies along different lengths on the frame. This experiment uses the output result in mean and standard deviation values to distinguish the anomaly pattern. A 10 cm x 10 cm cuboid was fixed during the experiment. The reason for using the mentioned size area is that it can be fitted on the surfaces of all kinds of frames. As mentioned in the prior subchapter, Frame 2 and Frame 3 were tilted (±40˚) from their original axis (±0˚). It also gives the best output image after being detected by robot vision. It has a median value with the highest and lowest mean and standard deviation pixels recorded from the previous experiment. The result shows that an anomaly can still be identified and evaluated, even though the standard deviation value has fluctuated in all frames. Because of the fixed size, the mean graph shown in Fig. 14 to Fig. 16 for each frame does not change critically. Their differences were at least 0.01 until 0.04 from the normal mean value.

V. DISCUSSION

With all the observations, the calculated mean, µ and standard deviation, σ for the initial and final point of the object in all three frames were recorded as the lowest. The biggest difference in σ value showed when the object moved nearer to the middle. It is because, at 0 cm and 50 cm, the cuboid is captured half of its actual size in frame boundaries. Thus the pixel intensity of the background value is the highest compared to the object colour value, making the standard deviation value drops significantly. The image pixel also shows the highest grey intensity when the object is located by the edge of any frame. All performance data in Experiments B, C and D are categorized as quantitative results.

Based on the graph trends corresponding to the tabulated data recorded in Experiment B, the σ curve ascends and achieves the highest peak when the object is 35 cm x 35 cm. In contrast, the mean line decreases slightly until its minimum value. It is because the pixel intensity for the red value surpasses the grey value, which gives great differences in deviation reading compared to the normal deviation data. However, it started to decline slightly after the highest peak as the pixel value for grey is untraceable inside the frame. It also happens when the object size is smaller. For example, when the measurement size is 1 cm x 1 cm, the line for standard deviation approaches the line for the normal value. The

changes happened in the mean line too. This can be explained as the pixel value for red being untraceable and considered a normal condition. This situation also applies in Frame 2 and Frame 3. Observing through each graph in Experiment D, the pattern for mean in anomaly and common condition shows little differences and is fixed horizontally. However, a different outlook is observed in the standard deviation pattern. The graph is up and down compared to the normal standard deviation line. The outcome shows that the standard deviation value is the best parameter to be observed when examining anomaly behavior and comparing them with the normal state.

Meanwhile, the camera on the humanoid robot followed the object no matter the target object's position in image segmentation. From the concept, the distance between the object on the frame and the NAO robot is always the same. Both images in the left and right eye are matched so that only the image of the target object is recognized, pinpointed and shaded in the bounding area. Interest points are grouped according to respective image position and disparity. Each of them carries a different pixel number.

However, the mean pixel number of a particular object is much alike, although it is located in various positions. To access each image's pixel value and calculate the mean and standard deviation, Open CV and Numpy plug-in library functions need to be explored. In this project, every image is set to produce only 3-bit pixel values. The pixels are red, green and blue intensity. The pixel value detected differed from what it was supposed to be whenever the object's position was between the edges of two frames, or the object surpassed the frame edge.

As a consequence, the segmented image produced different values of pixels. As stated earlier, anomaly behavior happens when there is unlike frequency happening at a time. Thus, when the mean, standard deviation, or PPV value produced differs from the normal state, it can be categorized as an anomaly, as stated by a study in [17, 18].

The proposition that can be made from Experiments A, B and D is that both µ and σ values are reliable. The difference between normal and anomaly is comparable and can be proved based on the physical data and recorded image. To validate, each experiment is repeated five times to get accurate results, and the average value is calculated and recorded in tables. However, minor errors happen when the head yaw of the humanoid robot is rotated 30˚ to the left and right during detecting and scanning the frames repeatedly. Sometimes, the robot turned less than 30˚ or more than 30˚ from what it was supposed to. The effect of this the captured image is not similar as expected because the frame and point of vision are visibly not parallel because of the few internal factors such as joint errors and kinematic errors [19]. Because of that, the simulation must stop and take a break for 5 to 6 seconds. Then, it resumed again to get a precise result and eliminate the insignificant error that might change the results, as proposed by [20].

VI. CONCLUSION AND RECOMMENDATIONS

In a nutshell, the main objectives of this project are successfully achieved using the semi-supervised approach to

learn the anomalies. PPV value, overall mean value, variance and standard deviation have been calculated to prove the anomaly statistically in that scenario. It also can be concluded that an object's positions greatly vary the standard deviation values. The location of the abnormal condition also has been proved during the experiment. It is proved by observing the maximum size of the object that can fit in the captured frame by ensuring all surfaces are within the range. Data obtained from Frame 2 differs from the data in Frame 3, although both slanted 40° from the horizontal. The setting for the frames and light tuning needs to be customized in the future. Besides that, a shadow effect is formed from different angles even though the light intensity for the surrounding has been set to default in the early setup. Furthermore, the calculation for mean and standard deviation pixel value can be made to access every bit image rather than only 3-bit colour to improve the study result. And lastly, the improvement could be focused on when there are ambient environmental changes in terms of illumination intensity, object colour and quantity of objects to be detected. The control system of the robot could also be improved by using closed loop estimation system [21] or by adaptive technique called disturbance observer [22].

References

[1] Zabalza, J., Fei, Z., Wong, C., Yan, Y., Mineo, C., Yang, E., Rodden, T., Mehnen, J., Pham, Q. C. and Ren, J., "Smart Sensing and Adaptive Reasoning for Enabling Industrial Robots with Interactive Human-Robot Capabilities in Dynamic Environments - A Case Study," Sensors, vol. 19, 2019. doi: 1354. 10.3390/s19061354.

[2] Pasquale, G., Mar, T., Ciliberto, C., Rosasco, L., and Natale, L., "Enabling Depth-Driven Visual Attention on the iCub Humanoid Robot: Instructions for Use and New Perspectives," Frontiers in Robotics and AI, vol. 3, 2015. doi: 10.3389/frobt.2016.00035.

[3] Leitner, J., Förster, A., and Schmidhuber, J., "Improving Robot Vision Models for Object Detection Through Interaction," Proceedings of the International Joint Conference on Neural Networks, 2014. doi: 10.1109/ICHR.2010.5686837.

[4] Flores, R., Karen, L., Trujillo-Romero, F. and Suleiman, W., "Object recognition modular system implementation in a service robotics context," pp. 1-6, 2017. doi: 10.1109/CONIELECOMP.2017.7891833.

[5] Laielli, M., Smith, J., Biamby, G., Darrell, T., and Hartmann, B., "LabelAR: A Spatial Guidance Interface for Fast Computer Vision Image Collection," pp. 987-998, 2019. doi: 10.1145/3332165.3347927.

[6] Pizzuto, G. and Cangelosi, A., "Exploring Deep Models for Comprehension of Deictic Gesture-Word Combinations in Cognitive Robotics," pp. 1-7, 2019. doi: 10.1109/IJCNN.2019.8852425.

[7] Khairy, D., Salem A, Areed, M. F., Mohamed A. A. and Rania A. A., "An Algorithm for Providing Adaptive Behavior to Humanoid Robot in Oral Assessment", International Journal of Advanced Computer Science and Applications (IJACSA), vol. 13, no. 9, 2022. doi: 10.14569/IJACSA.2022.01309119.

[8] Martínez, E., del Pobil, Angel P., "Vision for Robust Robot Manipulation," Sensors (Basel, Switzerland), vol. 19, 2019. doi:10.3390/s19071648.

[9] Sanchez, T., Caramiaux, B., Thiel, P., and Mackay, W., "Deep Learning Uncertainty in Machine Teaching," pp. 173-190, 2019. doi: 10.1145/3490099.3511117.

[10] L. Ma, M. Ghafarianzadeh, D. Coleman, N. Correll and G. Sibley, Simultaneous localization, mapping, and manipulation for unsupervised object discovery. In 2015 IEEE International Conference on Robotics and Automation (ICRA), (2015) 1344-1351. doi: 10.1109/ICRA.2015.7139365.

[11] Zhou, Z. and Yatani, K., "Gesture-aware Interactive Machine Teaching with In-situ Object Annotations," 2022, pp. 1-14. doi: 10.1145/3526113.3545648.

[12] Ari, D. and Alagoz, B. "A Review of Genetic Programming Popular Techniques, Fundamental Aspects, Software Tools and Applications. Sakarya University Journal of Science, 2021. doi: 10.16984/saufenbilder.793333.

[13] B. Nuño, C. Carlos and H. Álvaro, "Improving the detection of robot anomalies by handling data irregularities," Neurocomputing, vol. 459, 2021, pp. 419-431. doi:10.1016/j.neucom.2020.05.101.

[14] Gueye, T., Wang, Y., Mushtaq, R. T., Rehman, M., Ahmed, A., and Ali, H., "State of the art review on automatic sorting system for industrial robots using Internet of Robotic Things," 2022. doi: 10.21203/rs.3.rs-2329674/v1.

[15] Espinosa, O., Castañeda, L., and Martínez, F., "Minimalist Artificial Eye for Autonomous Robots and Path Planning, 2016, pp. 232-238. doi: 10.1007/978-3-319-24834-9_28.

[16] Giovanni, M., Carmine, R., and Rezia, M., "Vision and Locomotion Control Systems in a bioinspired Humanoid Robot," Proceedings of the Mediterranean Electrotechnical Conference – MELECON, 2015. doi:10.1109/MELCON.2014.6820564.

[17] Bucinskas, V., Dzedzickis, A., Sumanas, M., Sutinys, E., Petkevicius, S., Butkiene, J., Virzonis, D., and Morkvenaite-Vilkonciene, I., "Improving Industrial Robot Positioning Accuracy to the Microscale Using Machine Learning Method," Machines, vol. 10, no. 940, 2022. doi: 10.3390/machines10100940.

[18] Chunkai, Z., Chen, Y., Yin, A., and Wang, X., "Anomaly detection in ECG based on trend symbolic aggregate approximation," Mathematical Biosciences and Engineering, vol. 16, pp. 2154-2167., 2019. doi:10.3934/mbe.2019105.

[19] Kecskés, I., Odry, Á., and Odry, P., "Uncertainties in the Movement and Measurement of a Hexapod Robot". In: Awrejcewicz, J. (eds) Perspectives in Dynamical Systems I: Mechatronics and Life Sciences. Springer Proceedings in Mathematics & Statistics, vol . 362, 2019. doi: 10.1007/978-3-030-77306-9_12.

[20] Vocetka, M., Huňady, R., Hagara, M., Bobovský, Z., Kot, T., and Krys, V., "Influence of the Approach Direction on the Repeatability of an Industrial Robot," Applied Sciences, 2020. doi: 10. 8714. 10.3390/app10238714.

[21] Mansor, N. N., Jamaluddin, M. H., and Shukor, A. Z., "Adaptive Control Technique Effects on Single Link Bilateral Articulated Robot Arm," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 12, no. 7, pp. 512-520, 2021. doi: 10.14569/IJACSA.2021.0120759.

[22] Aivaliotis, P., Papalitsa, E., Michalos, G., and Makris, S., "Identification of dynamic robot's parameters using physics-based simulation models for improving accuracy," Procedia CIRP, vol. 96, pp. 254-259, 2021. doi: 10.1016/j.procir.2021.01.083.