

The Deep learning Development for Real-Time Ball and Goal Detection of *Barelang-FC*

Susanto, Eko Rudiawan, Riska Analia, Daniel Sutopo P, Hendawan Soebakti

Mechatronic Engineering

Politeknik Negeri Batam

Batam, Indonesia

{susanto, ekorudiawan, riskaania, daniel, hendawan}@polibatam.ac.id

Abstract—Studies of object detection have recently attracted increased interest. One of the applications of object detection is robotics. This paper present the real-time object detection integrated to humanoid robot soccer. In order to enhance the vision to detect ball and goal, the You Only Look Once (YOLO) methods is used as deep-learning object detection method. The real-time experiments have been carried out in LINUX OS by using NVIDIA JETSON TX1 controller board. The experimental results show that the proposed method capable to detect and distinguish objects in the different lighting condition, with interference from other objects, also from the different angle of capturing an image.

Keywords—Artificial Intelligence, deep-learning, object detection, robotic vision, YOLO.

I. INTRODUCTION

The evolution of object detection was getting popular not only for image processing but also for robotic vision. In the robotic vision, applications of objects detection commonly used to recognize the objects around the robot while a robot is moving in its environment. It needs a lot of effort to embed the object detection system into robot, especially a robot which needs a good perceive of the object faster and accurate in real-time.

One of the applications of object detection in robotics is in robot soccer. In robot soccer, it is crucial to have a good, fast, adaptive and accurate system to recognize the objects. Many researchers have studied how to detect object effectively in robot soccer. In [1], Cheng, et.al proposed the object detection by using segmented object based on image depth. They used Kinect sensor to perform the 3D information of the image. This sensor enables the system to refresh the image up to 30 FPS. The 3D information from this sensor is used in omnidirectional camera to complete the task. In this research, they offer the method of object detection in four parts such as cloud filtering, background removing, cloud clustering, and object localization.

While in [2], Peter, et.al present the adaptive estimation for real-time object localization by using color-coded based on color-based pixel-wise subtraction. They utilize the adaptive estimation for real-time object localization to detect the position of ball and goal keeper. This method also applies a threshold to segment the color in Hue Saturation and Value (HSV) on recognizing the ball and goal keeper in the field.

For robot soccer, ball detection has done by combining the color and depth information which presented by Fitriana, et.al in [3]. The combined color and depth image information are used in omnidirectional camera to detect the 3D shape of the ball according to the color information of the object. By using this method, they claim it can perform the processing images faster and suitable for real-time application.

Hock, et.al in [4] suggested the object detection method by separating into two categories. These categories are color blob segmentation and grid-based segmentation. To detect the red ball in real-time, they used color blob segmentation combined with depth estimation through trigonometric properties. In order to detect the goal post, the grid-based segmentation pattern and recognition was applied to the system.

To detect and follow the spherical object without knowing the information of object color, it can be detected by color domain which contributes with latest background images through three steps such that detect the color of sphere object automatically, sphere object detection method, and using Kalman Filter to detect the position of sphere object. This method can be found in [5] which been done by Mulya, et.al.

The Hough transform also can be used as object detection instead of Kalman Filter which presented by Nevel, at.al [6]. They modify the object distance parameter to ensure the speed of processing object. The line process detection algorithm also point out in this research.

Bima et.al [8] used a Hough transform to detect real-time background images. This transform used to detect the color object and the position. The color-based subtraction and pixel-wise are used to detect the adaptive background. This method also can estimate the distance of goal from the robot body using triangle similarity. In [9], they used NAO humanoid robot camera to detect the object color and shape. They also used the color blob segmentation and depth color estimation based on trigonometry and used pattern grid-based segmentation to detect the goal position.

All these methods are having difficulty to detect an object with others if the object has the color the same with others. To overcome this problem, a method call YOLO (You Only Look Once) is used. This paper will perform object detection using this method and implemented on humanoid robot soccer called Barelang-FC. A prototype of Barelang-FC has been designed and constructed in our lab as shown in Fig.1.

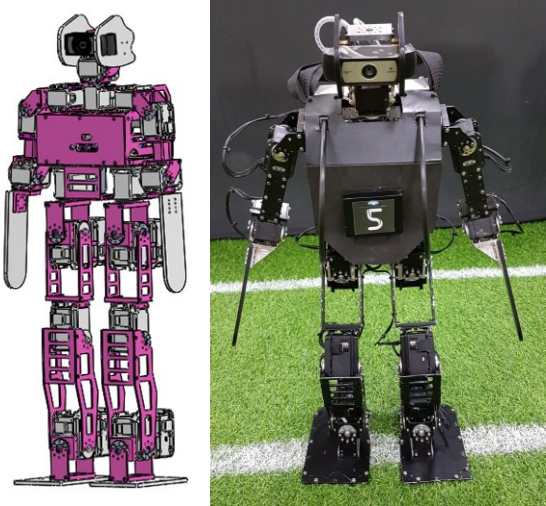


Fig. 1 Prototype of Barelang-FC humanoid robot soccer.

YOLO is a new approach to object detection introduced by Redmon, et.al [7]. This algorithm detected the object to classify and localize an object. A single neural network employs to this system to predict the bounding boxes and class probabilities from full image in one evaluation. This method can be performing effectively up to 45 FPS in real-time object detection. It makes YOLO has some benefits such as fast than other object detection method. YOLO enable to predict an object. This is done by using region proposal-based techniques and learns to generalize representations of objects.

The rest of this paper is organized as follows: Section II will describe about proposed method which been used as object detection approach. In section III, the controller design was explained to implement the proposed method to the robot in real-time. The results of experiments are presented in section IV. And finally, in Section V contains the conclusion and future work from the results obtained.

II. OBJECT DETECTION

Fast and accurate object detection method is needed to accomplish robot soccer tasks for example to detect the ball and goal. YOLO is a new approached which faster than others detection systems. The processing image with YOLO is simple and straightforward. As illustrated in Fig.2, at the beginning this system will resize the input image to 320 x 320. After the image is already resized, then it runs the convolution network on the image to predict the bounding boxes and class probabilities on the image simultaneously, train on full image and directly to optimize the detection performance. Then do thresholds the resulting detections by the model's confidence. This work reframes the object detection into a single regression problem and straight from the pixels of an object to bounding box coordinates and class probabilities. By this description, YOLO can determine the objects and location of these objects in the full image frame [7]. YOLO system is tended to unify the separate components of object detection becomes a single neural network. It makes bounding box of entire image feature should be considered.

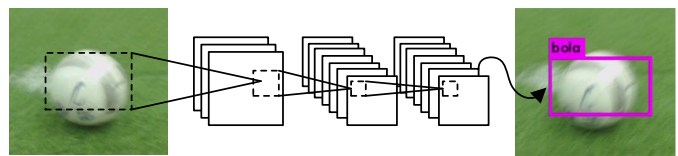


Fig. 2 YOLO recognition system.

This method divides the input image into $s \times s$ grid. This grid cell is responsible for detecting the object when the center of the object falls into this cell. The aim of each grid is to predict B bounding boxes and confidence score of the boxes. The confidence scores respect to the how confidence and accurate the model which the box contains an object to predicts. The confidence factor can be expressed as (1). If there is no object exists in that cell, then the confidence score defines as zero. The confidence score should be equal to intersection over unit (IOU) between predict box and ground truth.

$$\Pr(\text{Object}) * IOU_{predict}^{truth} \quad (1)$$

In this method, it consists of 5 predictions such that x, y, w, h , and confidence. The (x, y) coordinates describe the center of the bounding box of the grid cell, while w and h represent the prediction of width and height relative to the whole image.

The other parameter which should predict is C . C is a conditional class probability which can be expressed as (2). Our system has two classes which named as "ball" and "goal" classes. Each grid cell containing the object will have the conditional probabilities. This method only predicts one set of class probabilities per grid cell or the number of boxes B .

$$\Pr(\text{Class}_i | \text{Object}) \quad (2)$$

To obtain the class-specific confidence scores for each box, the conditional class probabilities should be multiplied by the individual box confidence prediction which can be seen in (3). As for the tensor prediction, the formulation can be encoded as (4). For evaluate the proposed method to detect the object in Barelang-FC system, we determined the parameters as: $S=5$, $B=2$ and $C=2$.

$$\Pr(\text{Class}_i | \text{Object}) * \Pr(\text{Object}) * IOU_{predict}^{truth} = \Pr(\text{Class}_i) * IOU_{predict}^{truth} \quad (3)$$

$$S \times S \times (B * 5 + C) \text{ tensor} \quad (4)$$

Fig. 3 described the system architecture of YOLO detection system. Based on [7], YOLO system was implemented as a convolutional neural network and evaluated on the PASCAL VOC detection dataset. This object detection system consists of two main layers such as initial convolutional layers and fully connected layers. The initial convolutional layers will extract the features from image meanwhile the fully connected layers will predict the output probabilities and coordinates. The network architecture is inspired by the GoogLeNet model for image classification with 9 convolutional layers followed by 2 fully connected layers. The YOLO architecture used 1x1 reduction layers followed by 3x3 convolutional layers. From the expression (4), we have acquired the final output of the network is the $5 \times 5 \times 12$ tensor of predictions.

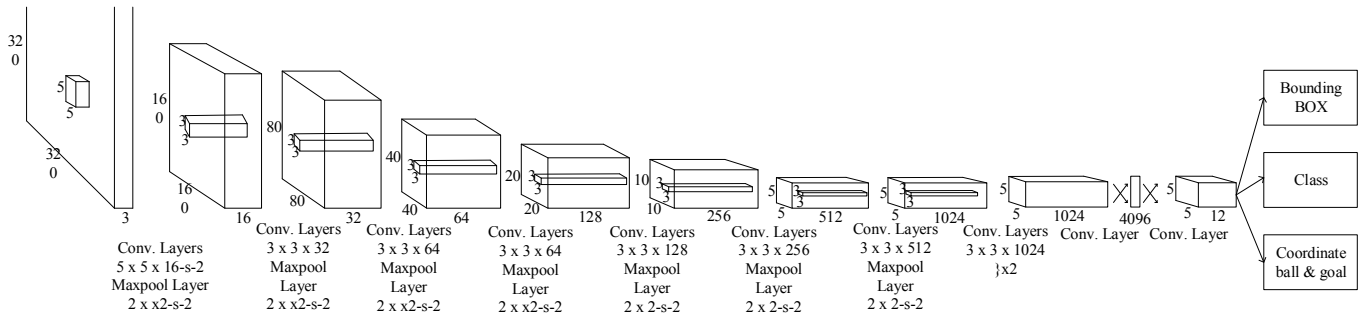


Fig. 3 The system architecture of object detection.

The output prediction illustrated as bounding BOX which is named as class name and position of the object detected. For training process we have collected 4000 images each class and 8000 images in total for both classes. To increase accurateness of detection, we have recorded the video of the robot during the robot is moving then convert and store it into an image file. Each condition happening during robot in operation is sorted. A blur image is added, enlarge, and also decrease of the image to rise up the confidence value of accuracy detection. Training process obtain accuracy over 60%. This process is spent two days training with computer specification are i7, RAM 16GB and double GTX1080 of GPU.

III. HARDWARE DESIGN

In this section, the embedded process from object detection results to humanoid robot soccer will describe. Fig. 4 shows the block diagram of the construction hardware of the robot soccer. Barelang-FC hardware construction consists of digital camera, computer, servo controller, and servo motors. The Digital Camera is used to capture the environment surrounding of the robot to detect and obtain the position of a ball and the goal. The C930 type was chosen as a digital camera for robot vision. The computer which mounted in the robot is used for controller board to process the proposed method, the NVIDIA JETSON TX1 is selected as a computer board in this system. Deep-learning computation needs a combination of Central Processing Unit (CPU) and Graphical Processing Unit (GPU) work together to speed up calculation during the process. NVIDIA JETSON TX1 has Quad ARM processor and 256 CUDA Cores, with its dimension is the same sizes as a credit card. It is descent processing unit to be applied in the Humanoid robot soccer. When the results of deep-learning process has been done in the controller boards, then all the coordinate and shape of object will be sent to the sub controller of the servo controller. This information is used to move the servo motors and let the robot knows the position of the ball and goal also itself.

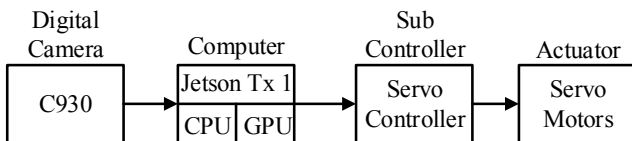


Fig. 4 Block diagram of hardware construction.

IV. EXPERIMENTS RESULTS

To determine the effectiveness of the proposed method, the YOLO system is embedded into Barelang-FC robot soccer in real-time. YOLO system is utilized to detect the ball and goal fast and accurate. To optimize the proposed method to detect the object accurately, the LOO (Leave One Out) cross validation is combined with proposed method to detect object in real-time application with 1000 data sample for each image.

In this section, the experimental results will present along with the figure explanation. The whole experiments are used white ball with different pattern which is been used in RoboCup 2017 competition. Literally, this proposed method also can detect a different variation of ball however, to detect the other variation of ball it is needed to be trained and it is possible to do it even need a lot of time to wait the training process to be completed. By using the proposed method, our robot can detect the ball and goal around 12cm to 4m depends on the image quality and the frame quantity of image from digital camera which is mounted on the robot

At first experiment, the moving robot tries to detect the ball in the area without a white line and set the pink box as a ball which written as “bola” and orange as a goal written as “gawang”.



Fig. 5 Ball detection in bright lighting condition.

As can see in Fig. 5, the system detected the balls with camera lighting in a bright condition so that the color of grass still green. While in Fig. 6, the system also can detect the ball with camera lighting in less bright condition so that the color of grass change into slightly blue. By doing these two experiments, the system still can detect the position and shape of ball. The distance between robot and ball in this experiment is around 20 cm.

The second robot task is to determine the ball near the white objects for example line or a pole. Fig.7 presents the YOLO results of detecting ball near the white line while in Fig. 8 the systems detect the ball near the white line and pole. The system can recognize the ball even though the angle and distance are varied. The distance from object to robot is about 60 cm. These results, also shown that the system can recognize the ball with different contrast produced by camera and determine the ball even in a long distance.



Fig. 6 Ball detection in less lighting condition.

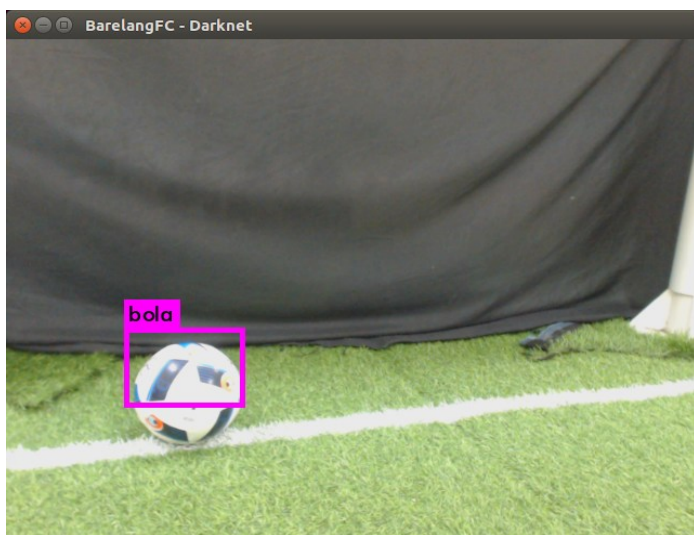


Fig. 7 Ball detection disturb by white line.

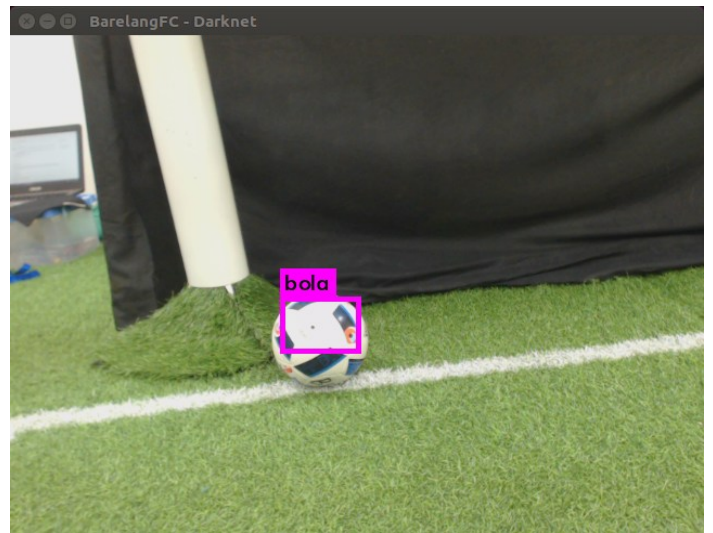


Fig. 8 Ball detection disturb by white line and pole of goal.

Furthermore, the experiment to detect the goal is also performed. The moving robot is trying to determine the goal with different contrast and angles. The goal detecting results is shown in Fig. 9 and Fig. 10. In this experiment, we cover the goal with black fabric to give a different color of the object. From this experimental result, it shows that the system can recognize the white goal stand.

The other experiment, we set the environment with different objects as seen in Fig. 11. This figure shows that the system can detect the intention of the object clearly and accurate even though the distance from the robot to the object was far enough. Furthermore, this system recognizes the ball and goal in “noisy” environment. We put some objects in the “noisy” environment such as black fabric to cover the white wall, some balls with different pattern, robots, and other stuff.

Next experiment was held in the more “noisy” condition. The black cover of goal was taken out and the noise from the white wall is added as a background. The robot is trying with the proposed method to recognize the ball and goal from some distance. The result can be seen in Fig. 12. The wall and goal have the same color also the camera in the robot can see the white line under the goal, with the distance from half of the field and crowded area. The proposed method enables to distinguish the ball and goal. Additionally, we also verify the proposed approach in the different angle, distance, and crowded object as shown in Fig.13. Fig.13 tells the result of object detection even in slight angle position. The system can detect the setting object without interfering with other objects accurately and sharply.

The proposed method enables to give the information of ball and goal coordinate as show as Fig. 14. This coordinate information is an important thing to robot for maneuvering in the field so that the robot will not place the ball in the wrong goal. From Fig. 14, we can notice that the proposed method of object detection evenly can do 20 FPS in order to detect the object as expected. The rate of FPS results of this method is depending on how many layers applied in the network.



Fig. 9 Goal detection in straight viewpoint from robot.

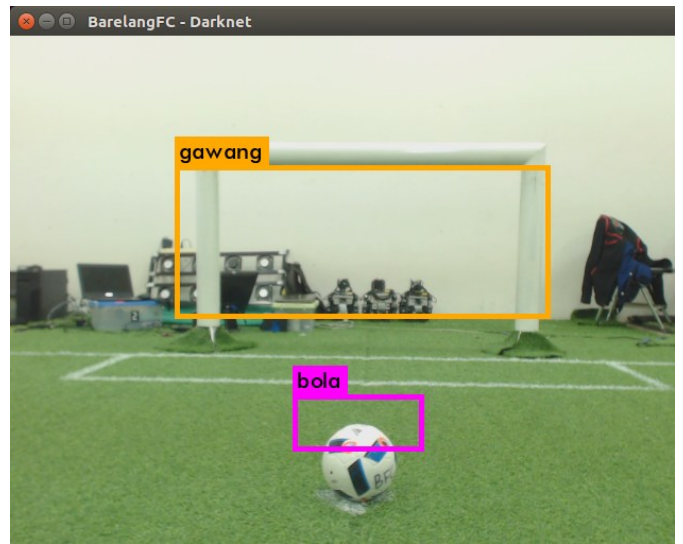


Fig. 12 Ball and goal detection disturb by white wall and some other objects.

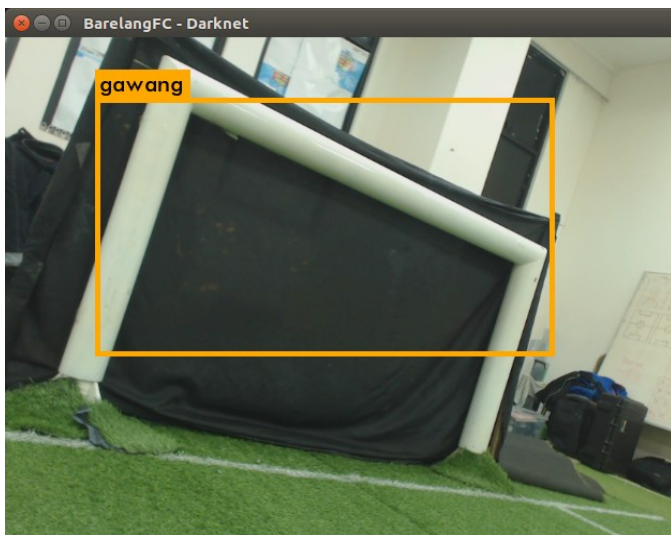


Fig. 10 Goal detection from several angle of robot.

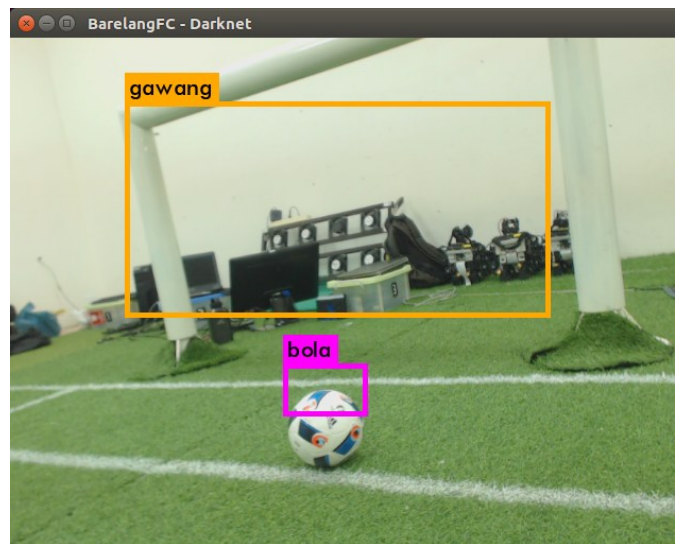


Fig. 13 Ball and goal detection disturb by white wall and some other objects from different angle.



Fig. 11 Ball and goal detection covered by black fabric.

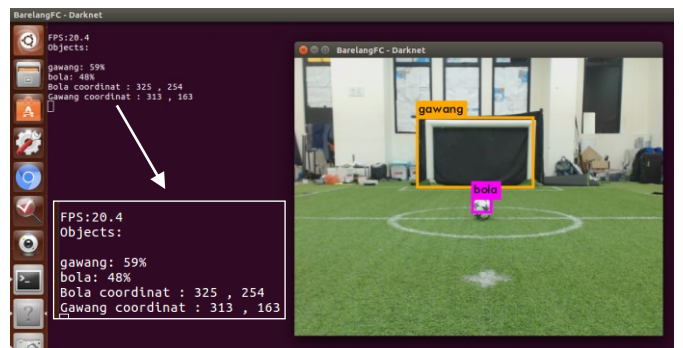


Fig. 14 System shows the coordinate of ball and goal position.

Moreover, the precision of an object detection is depending on how many class which been used the amount of the class is linear in comparison to the accurate result of object detection. For these experiments, two classes and 16 layers

consist of 9 convolution layers and 7 max-pool layers are used as described in Fig. 3.

V. CONCLUSION

This paper presents a real-time object detection unified to humanoid robot soccer to detect white ball and goal with different pattern and position. YOLO was selected as an object detection method and unified to real humanoid robot soccer in a real-time situation. From the experimental results it can be concluded that the proposed method of object detection can cope the task in a different situation. By using the proposed method, it also can show the coordinate of ball and goal which already detected. In future work, we will use the method to distinguish robot team member and opponent to get some new strategy approach while maneuvering in the field.

REFERENCES

- [1] Qiu Cheng, Shuijun Yu, Qinghua Yu, Junhao Xiao, "Real-time object segmentation for soccer robots based on depth images," IEEE-International Conference on Information and Automation (ICIA), pp. 1532-1537, Ningbo-China, August 2016.
- [2] Peter Chondro, Shanq-Jang Ruan, "An adaptive background estimation for real-time object localization on a color-coded environment," IEEE-International Conference on Advanced Computer Science and Information (ICACSIS), pp. 464-469, Malang-Indonesia, October 2016.
- [3] A.N. Fitriana, K. Mutijarsa, W. Adiprawita, "Color-based segmentation and feature detection for ball and goal post on mobile soccer robot game field," IEEE- International Conference on Information Technology Systems and Innovation (ICITSI), pp. 1-4, Bandung-Indonesia, October 2016.
- [4] Phua Seong Hock, S. Parasuraman, "Motion synchronization with predefined rhythms for humanoid robot," IEEE-Recent Advances in Intelligent Computational Systems (RAICS), pp. 294-299, Trivandrum-India, December 2015.
- [5] Anton Kurniawan Mulya, Fernando Ardilla, Dadet Pramadihanto, "Ball tracking and goal detection for middle size soccer robot using omnidirectional camera," IEEE- International Electronics Symposium (IES), pp. 432-437, Denpasar-Indonesia, September 2016.
- [6] António José Ribeiro Neves, Alina Trifan, Paulo Dias, José Luís Azevedo, "Detection of aerial balls in robotic soccer using a mixture of color and depth information," IEEE- International Conference on Autonomous Robot System and Competitions, pp. 227-232, Vila Real-Portugal, April 2015.
- [7] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unifield, Real-Time Object Detection," unpublished, 2016.
- [8] Bima Nugraha Sanusi, Widyawardana Adiprawita, Kusprasapta Mutijarsa, "Humanoid robosoccer goal detection using hough transform," IEEE- International Conference on Control, Electronics, Renewable Energy and Communications (ICCEREC), pp. 153-159, Bandung-Indonesia, August 2015.
- [9] Anže Troppan, Eduardo Guerreiro, Francesco Celiberti, Gonçalo Santos, Aamir Ahmad, Pedro U. Lima, "Unknown-color spherical object detection and tracking," IEEE- International Conference on Autonomous Robot Systems, pp. 1-4, Lisbon-Portugal, April 2013.