Ball Detection for Boccia Game Analysis

Alexandre Calado, Vinícius Silva, Filomena Soares, *Member, IEEE*, Paulo Novais e Pedro Arezes*

Abstract— **The present article proposes the training, testing and comparison of two models for ball detection, taking into account its final implementation in a Boccia game analysis computer-vision algorithm, within the "iBoccia" framework. The goal is to have a versatile and flexible algorithm towards different game environments. The selected ball detectors were a Histogram-of-Oriented-Gradients feature based Support Vector Machine (HOG-SVM) and a Convolutional Neural Network (CNN) based on a less complex implementation of the You Only Look Once model (Tiny-YOLO). Both detectors were evaluated offline and in real-time. The subsequent results showed that their performance was similar in both evaluations, however, Tiny-YOLO outperformed HOG-SVM by a small margin in all the used metrics. In real-time, both detectors achieved an accuracy of approximately 90%. Despite the high accuracy values, the detector requires further improvement because a single non-detection can influence the computer-vision algorithm's output, making the system unreliable.**

I. Introduction

Physical Inactivity is currently the fourth leading risk factor for global mortality, as reported by the World Health Organization [1]. It is estimated to be responsible for 3.2 million deaths worldwide each year. Moreover, it is also associated with premature mortality and the occurrence of several disorders and illnesses, such as coronary heart and cerebrovascular diseases, type 2 diabetes, hypertension, obesity, and different types of cancer [2], [3].

According to Dumith *et al*. [4], physical inactivity is more prevalent in developed countries and the most affected age group is the elderly one. Taking into account the aforementioned remarks, along with the current increase in the number of older adults and the wide range of benefits from physical activity [5]–[8], it is paramount to design and implement innovative solutions to change the exercising habits of this age group.

In order to address this issue, the authors developed a framework called "iBoccia" [9]–[11] for monitoring and promoting physical activity amongst the elderly. Boccia is a simple precision ball sport which can be easily adapted to individuals with different types of motor or mental impairments. Within this framework, a computer-vision based system was implemented for Boccia game analysis. The current algorithm [12], [13], with the usage of a webcam, is capable of detecting balls, according to its shape and colour, calculate the distance between each of the red and blue balls to the jack (white ball) and compute the score of a Boccia game in real-time (for more information about the scoring rules, the reader should refer to the BISFed [Boccia International Sports Federation] International Boccia Rules [14]), which is displayed in a User Interface (UI) [15]. A previous survey [15] sustains that this approach may motivate the elderly to play and engage in physical activity more often.

The current implementation of the algorithm uses a colour segmentation technique for detecting the Boccia balls based on the usage of colour masks defined in the HSV (Hue Saturation and Value) colour model, erode and dilate operations and detection of contours, i.e. curves that join all continuous points along a boundary that have the same colour and intensity. Although this implementation had very satisfiable results when applied to videos [12], [16], it is necessary to adjust a set of parameters, such as the masks' colour limits, number of iterations for morphological operations and minimum/maximum radius of the object to be detected. In real-time situations it is complicated and time-consuming to manually adjust these values due to issues such as luminosity conditions and camera positioning, therefore it is necessary to implement an algorithm that can be more flexible and robust in different testing environments. Recently, a collaboration was made possible with Sporting Clube de Braga's Boccia team, which enables the possibility of using this system for the athletes' training. Thus, for system usage, it is essential that calibration is not time consuming and no parameters adjustments are required.

The first step of this task is to correctly detect the Boccia balls with no need of calibration, which is the main focus of the present paper. Regarding computer-vision in sports, several other approaches for ball detection can be found in the literature. Chakraborty & Meher [17] ... the position of a basketball in real-time. To this end, the moving object is detected by using the approximate median method of background subtraction, followed by morphological operations and edge detection by using a Canny Edge detector [18]. After the moving object is detected, size, shape and compactness filters are used for identifying ball candidates. Although the accuracy for ball detection was high (94.77%), this solution is not viable for Boccia, since this approach depends of the object's movement and Boccia balls are static for most of the game.

A different approach was developed by Tong *et al*. [19] regarding soccer ball detection. Since the colour of the soccer field is stable and uniform, it can be represented by a

A. Calado is with the Algoritmi Research Centre, University of Minho, Campus de Azurem, 4800-058 Guimarães, Portugal (e-mail: alexandreluiscalado@gmail.com).

V. Silva is with the Algoritmi Research Centre, University of Minho, 4800-058 Guimarães, Portugal (e-mail: a65312@alumni.uminho.pt).

F. Soares is with the Algoritmi Research Centre, Dept. of Industrial Electronics, University of Minho, Campus de Azurem, 4800-058 Guimarães, Portugal (corresponding author phone: +351253510180; fax: +351253510179; e-mail: fsoares@dei.uminho.pt).

P. Novais is with the Algoritmi Research Centre, Dept. of Informatics, University of Minho, Campus of Gualtar, 4710 - 057 Braga, Portugal (e-mail: pjon@di.uminho.pt).

P. Arezes is with the Algoritmi Research Centre, Dept. of Production Systems, University of Minho, Campus de Azurem, 4800-058 Guimarães, Portugal (parezes@dps.uminho.pt).

constant mean colour and used for field extraction. Thus, after the field is extracted, only the remaining objects need to be considered for detection. Afterwards, a coarse-to-fine method is used for the identification of a single ball object and its contour is analysed for describing its shape features. Finally, the optimal ball region is found by fusing colour and shape similarity. This method proved to be valid and fast in image sequences, however, it depends of the stability and uniformity of the area of play for field extraction.

Other approaches include the use of other object detection methods, such as machine learning or deep learning. A properly trained machine learning or deep learning model for object detection can offer independence from the background, ball movement and colour, making these two options very attractive for the task at hand. In the context of the RoboCup International Robot Soccer Competition, Menashe *et al*. [20] used different types of Deep Neural Networks (DNNs) and Support Vector Machines (SVMs) for the detection of the soccer ball. The classification results showed that a Convolutional Neural Network (CNN) achieved a 99.48% precision and 99.41% recall, while a Polynomial SVM achieved 97.20% precision and 88.10% recall for the same dataset containing a total of 34,684 labelled images. Another example of machine learning usage for ball detection can be observed in the work of Ji *et al*. [21], where different combinations of pre-processing methods (for extracting image features) and classifiers, including SVMs and Adaboost, were used for the real-time detection of a table tennis ball. Overall, the combination of Local Gradient Patterns (LGP) and Adaboost obtained the highest accuracy (more than 85%).

The output of the aforementioned works show that deep learning and machine learning can have very good performances while detecting objects such as balls. However, there are downsides for these types of approaches, such as a higher computation time, when compared to image segmentation methods. Regardless, detection speed is not as important for Boccia game analysis as for the analysis of sports such as basketball, soccer or table tennis, where the movement of the ball is faster and practically constant. On the other hand, during a match of Boccia, the balls remain static for a longer time. Besides, the score is only noted when all of the balls within the Boccia court are not moving.

The goal of the present work is to implement a deep learning and a machine learning approach for the detection of Boccia balls and compare their performance, both offline and in real-time. The final algorithm should allow the detection of the balls within the court, regardless of luminosity conditions, camera angle and colours, thus avoiding the adjustment of parameters and cumbersome calibration.

The present paper is divided into five sections. Section II describes the system architecture, the used deep learning and machine learning algorithms and the methodology used for assessing the performance of both classifiers in real-time and offline. Section III presents the obtained results of both classifiers, which are respectively discussed in section IV. Finally, section V addresses the final remarks and future work.

II. METHODS

This section presents the proposed system architecture, the implemented classifiers for the detection of Boccia balls as well as the methodology for assessing their performance.

A. System Architecture

The system used for testing in the present paper is composed of an RGB camera and a computer (Figure 14). The used camera is a full high definition RGB web camera (Hewllet-Packard HD 4310). This camera was used in order to detect the balls during a gameplay session of Boccia by using computer-vision algorithms. The used computer to test the system featured a NVIDIA GeForce 920M graphics card, an Intel Core i7-5500U Quad-Core Processor and 8GB of RAM.
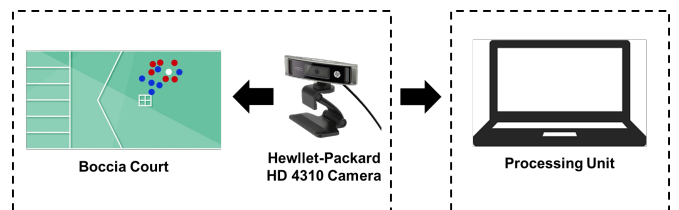

Figure 1: System Architecture

The used computer-vision algorithms were based on the DLIB Histogram-of-Oriented-Gradient (HOG) object detector [22] and You Only Look Once (YOLO) [23], which are described in the next two subsections.

B. HOG-SVM classifier

The DLIB library has a built-in trainable object detector that uses HOG features and a binary Support Vector Machine (SVM) classifier [22] to detect objects in scenes. Therefore, a model was trained in Python using the DLIB API. The model was trained by using the dataset described in subsection D with a 64 by 64 sliding window over 437 iterations with the following parameters: $C = 5$ and $\epsilon = 0.01$. C is a regularization parameter that allows a trade-off between the train set accuracy and capability of generalization. On the other hand, ϵ is used to define the SVM's safety margin. Both parameters affect the SVM complexity.

C. YOLO Classifier

Traditional detection systems apply the model to an image at multiple locations and scales. The high scoring image are considered detections. Conversely, the YOLO (You Only Look Once) convolutional neural network architecture proposed by Redmon *et al*. [24] consists in applying a single neural network to the full image, dividing the image into regions (a grid of 13 by 13 cells) and predicts bounding boxes and probabilities for each region. Then, YOLO outputs a confidence score that means the probability of a bounding box enclosing an object desired for detection. Additionally, for each bounding box it also predicts a class.

The confidence score and the class prediction are combined into a final score, informing the probability that a bounding box contains a specific object. The YOLO model used in the present work is the Tiny-YOLO, which it is smaller than the original YOLO model. The Tiny-YOLO model features 13 convolutional layers (3×3 and 1×1 convolutional layers) [23]. Finally, a confidence threshold

value is used to discard the bounding boxes with a final score value less than 30%.

This model was trained using the Darknet framework [25] with 8000 iterations. The network's input resolution size during the training was random, but always a multiple of 32. On the other hand, the input size image during test was of 416×416. The images were also flipped, augmenting the dataset. Moreover, the model was trained with a constant learning rate of 0.001.

D. Dataset

The dataset is comprised of 210 images with a total of 1416 annotations. The training set consisted of 140 images (with 932 annotations) of gameplay situations in the nursing home, laboratory, and in actual Boccia courts, which were kindly provided by the Sporting Clube of Braga. Additionally, it also contained images taken from official videos of the BISFed 2018 Liverpool World Boccia Championships [26], [27].

Regarding the test dataset, it was composed of 70 images (with 484 annotations) of Boccia game situations taken from 6 videos of BISFed regional and world competitions from 2017 and 2018 [28]–[33], as well as some images taken during game situations in laboratorial environment and in a Boccia court.

E. Evaluation Methodology

In order to evaluate the proposed system, two types of tests were conducted – offline and real-time evaluations. Concerning the offline evaluation, to quantify the performance of each classifier, the following metrics were used: recall, precision and average precision.
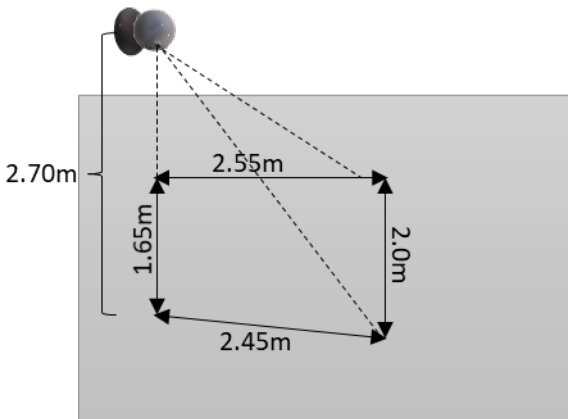

Figure 2: Representation of the area of the Boccia court (and its respective dimensions) covered by the camera's Field-of-View

As for the real-time evaluation, tests were conducted in a Boccia court. The camera was placed at 2.70 meters above the ground, directly above the court. Its Field-of-View (FOV) and correspondent area of the Boccia court covered is depicted in Figure 2.

A total of 12 Boccia games were simulated, 6 for each classifier. During each game, the number of ground-truths, true positives and false positives was registered for each new game situation, which was considered to change each time a Boccia ball entered the camera's FOV. The detection was considered a true positive if the ball stopped its movement inside the camera's field and was detected by the algorithm.

Additionally, if the ball entered the camera's FOV but stopped its movement outside of it, it was considered to be a true positive if the ball was successfully detected at least once during its movement. Finally, the performance of each model was assessed in terms of accuracy.

III. RESULTS

This section presents the results of both offline and real-time evaluations of the Tiny-YOLO and the HOG-SVM performances for the detection of Boccia balls, according to the methodology described in the previous section.

A. Offline Evaluation

As mentioned previously, the performance of both models for offline classification was evaluated with the use of three metrics: recall, precision and average precision. Average precision is a standard metric for object detection evaluation, and it can be computed as the area under the precision-recall curve [34]. The used Intersection of Union (IoU) threshold was of 0.5. Both models were used for detection on the same test set, which was composed by 70 images, as mentioned previously. The results of both Tiny-YOLO and HOG-SVM for offline evaluation can be observed in Table I.

TABLE I. RESULTS FROM OFFLINE EVALUATION

Classifier	Recall (%)	Precision (%)	Average Precision (%)
HOG-SVM	79.05	71.62	72.68
Tiny-YOLO	79.96	83.77	74.25

B. Real-Time Evaluation

Regarding the results of the real-time evaluation, a total of 61 game situations were considered while using the HOG-SVM detector and 58 while using the Tiny-YOLO detector. As mentioned in section II, the number of ground-truths, true positives and false positives was registered for each of the game situations. The total ground-truths, true positives, false positives, as well as accuracy, can be observed in Table II. Ground-truths corresponds to the total number of Boccia balls within the camera's FOV, thus accuracy was calculated as follows: *Accuracy = Total True Positives/Total Ground-Truths*

TABLE II. RESULTS FROM REAL-TIME EVALUATION

Classifier	Total Ground-Truths	Total True Positives	Total False Positives	Accuracy (%)
HOG-SVM	300	272	0	90.67
Tiny-YOLO	281	256	0	91.10

IV. DISCUSSION

Considering the offline evaluation, the results of both classifiers were very similar. The most noticeable difference can be observed in the precision values, where Tiny-YOLO achieved a considerably higher value than the HOG-SVM.
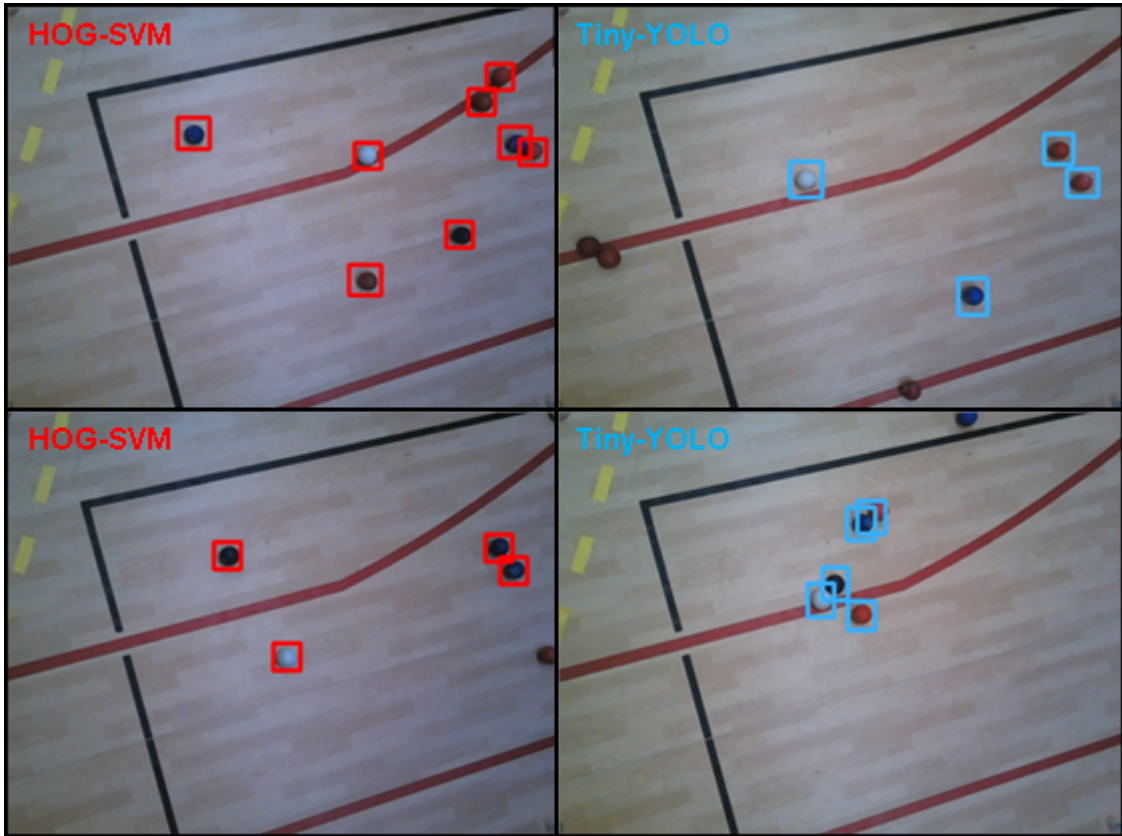
Figure 3: Example of game situations that occurred during testing and detections represented by the bounding boxes coloured red (for HOG-SVM) and light blue (Tiny-YOLO). The upper left and right figures depict situations when the ball stopped above a court line of the same colour for both HOG-SVM and Tiny-YOLO detectors, respectively. On the other hand, the lower left and right figures depict situations when the ball is close to margin of the area covered by the camera FOV for both HOG-SVM and Tiny-YOLO detectors, respectively.

This means that the Tiny-YOLO predictions were more accurate, i.e., less false positives were detected. On the other hand, the identical recall values mean that both classifiers had a similar performance while finding all the positives. However, the most complete metric to consider is the average precision, which is the average of the maximum precisions for different recall values, which can give a better overall idea of the classifiers' performance. Taking this into account, Tiny-YOLO had a slightly better performance on offline classification. However, this does not necessarily mean that it has the best real-time performance.

As in the offline scenario, the results from real-time evaluation were also similar between the two detectors. The accuracies were considerably high in both situations (> 90%) and no false positives were detected. The lack of false positives can be justified by the fact that all of the tests were performed in the same area of the court where the colour of the floor was homogeneous, with the exception of some lines belonging to other different courts, as it can be observed in Figure 3.

Despite the high accuracies, considering the final application of these detectors, an approximate value of 90% is not sufficiently good because a non-detection could affect the real game score computation. Thus, it is required that the accuracy approximates 100%.

According to the values obtained in both offline and real-time evaluations, Tiny-YOLO had a slightly better performance. However, some debilities were observed while testing the system in real-time. When a ball stopped its movement above a court line that matched its colour, the Tiny-YOLO could not detect them. The HOG-SVM also showed some difficulties in this type of situations, however, intermittent detections were observed. Examples of these situations are depicted in Figure 3. Both classifiers also failed to detect balls that stopped above the court area margins covered by the camera, as it can be also observed in Figure 3.

Regarding the performance in terms of computation time, a variable that can be decisive for the selection of a detector, the Tiny-YOLO detector outperformed by a small margin the HOG-SVM detector, achieving an average of approximately 7-8 frames per second (FPS) versus approximately 5-6 FPS, respectively. Regarding Tiny-YOLO, this corresponds to a frame being processed in an average time of 134 ms. On the other hand, HOG-SVM takes an average of 184 ms.

Overall, Tiny-YOLO provided the better performance, however, it can be further improved by training the model with a higher number of game situations, such as the ones where the ball is above court lines or when the totality of the ball is not visible.

V. FINAL REMARKS AND FUTURE WORK

The work described in this paper presented the implementation and testing of two different ball detectors, one based in a machine learning approach and the other based in a deep learning one. The two detectors were evaluated in both offline and real-time scenarios. The development of these detectors had the objective of making the current computer-vision algorithm from the iBoccia

framework more flexible by eliminating the need of resetting parameters each time the user wants to use the system in a different Boccia court or in different lighting conditions.

The selected classifiers for this task were an SVM based on Histogram-of-Oriented-Gradient features (HOG-SVM) and a CNN (Tiny-YOLO). The results of the classifiers were very similar in both offline and real-time scenarios, although Tiny-YOLO scored superior values in all of the metrics used for evaluation, including a higher FPS rate. However, despite the high accuracy values both classifiers achieved in real-time, both need to be further improved.

Future work includes the training of the Tiny-YOLO with a larger dataset, which should include situations when the ball is harder to detect, as in situations where the ball is occluded, above court lines that match its colour or when several balls are next to each other.

In order to increase the algorithm's speed, the code will be rewritten in C# programming language, which provides faster computation than Python.

A tracker can also be implemented to ensure that the algorithm does not lose the ball coordinates given by the detection bounding box, which may happen when detections are not continuous from frame to frame, e.g. when the ball stops above a court line of the same colour or when the ball is moving too fast. Regardless, its usage requires the fine tuning of its parameters to ensure its tracking quality and to avoid false positives.

Overall, it was interesting to observe that an SVM, which features a relatively simple implementation, can have a similar object detection performance to a more complex model, such as a CNN, even with a smaller dataset (images were flipped during the training of Tiny-YOLO).

After the detection process is optimized to fit the requirements of the game, it is also predicted to use a machine learning or deep learning approach for detecting colours and enabling the system to differentiate between the three different ball colours (red, blue and white) without setting parameters.

References

[1] World Health Organization, "Global recommendations on physical activity for health," *Geneva World Heal. Organ.*, p. 60, 2010.

[2] I. M. Lee *et al.*, "Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy," *Lancet*, vol. 380, no. 9838, pp. 219–229, 2012.

[3] J. A. Knight, "Physical Inactivity: Associated Diseases and Disorders," *Ann. Clin. Lab. Sci.*, vol. 42, no. 3, pp. 320–337, 2012.

[4] S. C. Dumith, P. C. Hallal, R. S. Reis, and H. W. K. Iii, "Worldwide prevalence of physical inactivity and its association with human development index in 76 countries," *Prev. Med. (Baltim).*, vol. 53, no. 1–2, pp. 24–28, 2011.

[5] I.-M. Lee, E. J. Shiroma, F. Lobelo, P. Puska, S. N. Blair, and P. T. Katzmarzyk, "Impact of Physical Inactivity on the World's Major Non-Communicable Diseases," *Lancet*, vol. 380, no. 9838, pp. 219–229, 2012.

[6] T. H. Westhoff *et al.*, "Too old to benefit from sports? The cardiovascular effects of exercise training in elderly subjects treated for isolated systolic hypertension," *Kidney Blood Press. Res.*, vol. 30, no. 4, pp. 240–247, 2007.

[7] J. M. Stessman, Jochanan; Hammerman-Rozenberg, Robert; Cohen, Aaron; Ein-Mor, Eliana; Jacobs, "Physical activity, function, and longevity among the very old..pdf," vol. 169, no. 16, pp. 1476–1483, 2009.

[8] F. Landi, M. Cesari, G. Onder, F. Lattanzio, E. M. Gravina, and R. Bernabei, "Physical activity and mortality in frail, community-living elderly patients.," *J. Gerontol. A. Biol. Sci. Med. Sci.*, vol. 59, no. 8, pp. 833–837, 2004.

[9] V. Silva *et al.*, "iBoccia: A Framework to Monitor the Boccia Gameplay in Elderly," *Lect. Notes Comput. Vis. Biomech.*, vol. 27, 2018.

[10] V. Silva *et al.*, "A wearable and non-wearable approach for gesture recognition – Initial results," *9Th Int. Congr. Ultra Mod. Telecommun. Control Syst.*, pp. 185–190, 2017.

[11] A. Calado, P. Leite, F. Soares, P. Novais, and P. Arezes, "Real-Time Gesture Classification for Monitoring Elderly Physical Activity Using a Wireless Wearable Device," *Ninth Int. Conf. Sens. Device Technol. Appl.*, pp. 164–168, 2018.

[12] P. Leite, A. Calado, and F. Soares, "Boccia Court Analisys for Real-time Scoring," *Proc. 15th Int. Conf. Informatics Control. Autom. Robot.*, vol. 2, pp. 511–516, 2018.

[13] A. Calado, P. Leite, F. Soares, P. Novais, and P. Arezes, "Boccia Court Analysis for Promoting Elderly Physical Activity," *Innov. Eng. Entrep. HELIX 2018. Lect. Notes Electr. Eng.*, vol. 505, 2018.

[14] BISFed, "BISFed International Boccia Rules (v.2)," 2017.

[15] A. Calado, P. Leite, F. Soares, P. Novais, and P. Arezes, "Design of a Framework to Promote Physical Activity for the Elderly," in *Human Systems Engineering and Design*, 2019, pp. 589–594.

[16] A. Calado, P. Leite, F. Soares, P. Novais, and P. Arezes, "Boccia Court Analysis for Promoting Elderly Physical Activity," in *Innovation, Engineering and Entrepreneurship*, 2019, pp. 158–164.

[17] B. Chakraborty and S. Meher, "Real-time position estimation and tracking of a basketball," *2012 IEEE Int. Conf. Signal Process. Comput. Control. ISPCC 2012*, 2012.

[18] L. Ding and A. Goshtasby, "On the Canny edge detector," *Pattern Recognit.*, vol. 34, no. 3, pp. 721–725, 2001.

[19] X. Tong, H. Lu, and Q. Liu, "An Effective and Fast Soccer Ball Detection and Tracking Method," *Proc. 17th Int. Conf. Pattern Recognition, 2004*, vol. 4, pp. 795–798, 2004.

[20] J. Menashe, J. Kelle, K. Genter, J. Hanna, and E. Liebman, "Fast and Precise Black and White Ball Detection for RoboCup Soccer," in *RoboCup-2017: Robot Soccer World Cup XXI*, 2017.

[21] Y. Ji, J. Zhang, Z. Shi, M. Liu, and J. Ren, "Research on real – time tracking of table tennis ball based on machine learning with low-speed camera," *Syst. Sci. Control Eng.*, vol. 6, no. 1, 2018.

[22] D. E. King, "Max-Margin Object Detection," 2015.

[23] J. Redmon and A. Farhadi, "YOLOv3 : An Incremental Improvement," 2018.

[24] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once : Unified , Real-Time Object Detection," 2015.

[25] J. Redmon, "Darknet: Open Source Neural Networks in C." [Online]. Available: http://pjreddie.com/darknet/. (accessed on November 2018)

[26] Paralympic Games, "Day 5 | 2018 BISFed Boccia World Championships," 2018. [Online]. Available: https://www.youtube.com/watch?v=m9sdtvlJft0&. (accessed on November 2018)

[27] Paralympic Games, "Day 6 | 2018 BISFed Boccia World Championships," 2018. [Online]. Available: https://www.youtube.com/watch?v=Rmx63CutfhY&.(accessed on November 2018)

[28] Paralympic Games, "Day 7 | 2018 BISFed Boccia World Championships," 2018. [Online]. Available: www.youtube.com/watch?v=5myD-vo0iXI&. (accessed on November 2018)

[29] Ck Stadium, "BISfed 2017 Bangkok World Open Thailand," 2017. [Online]. Available: https://www.youtube.com/watch?v=4avPcXDSBgM. (accessed on November 2018)

[30] Japan Boccia Association, "BISFed 2018 ISE Boccia Regional Open Day4," 2018. [Online]. Available: https://www.youtube.com/watch?v=9FaEwP15cM4. (accessed on

November 2018)

[31] 台灣地板滾球運動聯盟, "【LIVE】BISFed 2018 Chinese Taipei BOCCIA Regional Open (11/27~12/2) 【直播】2018亞太區地板滾球公開賽 1127~1202," 2018. [Online]. Available: https://www.youtube.com/watch?v=oKKI3putRWw. (accessed on November 2018)

[32] LaLigaSports, "Open Regional Boccia 2018: Viernes tarde," 2018. [Online]. Available: https://www.youtube.com/watch?v=QK3qXQNmcVU&. (accessed on November 2018)

[33] USA Boccia, "Kansas City Open PBC3 AUS vs JPN 2017 09 25," 2017. [Online]. Available: https://www.youtube.com/watch?v=FNFrJ_F1PLw&. (accessed on November 2018)

[34] P. Henderson and V. Ferrari, "End-to-end training of object class detectors for mean average precision," 2017.