

An Automated Player Detection and Tracking in Basketball Game

P. K. Santhosh^{1,*} and B. Kaarthick²

Abstract: Vision-based player recognition is critical in sports applications. Accuracy, efficiency, and Low memory utilization is alluring for ongoing errands, for example, astute communicates and occasion classification. We developed an algorithm that tracks the movements of different players from a video of a basketball game. With their position tracked, we then proceed to map the position of these players onto an image of a basketball court. The purpose of tracking player is to provide the maximum amount of information to basketball coaches and organizations, so that they can better design mechanisms of defence and attack. Overall, our model has a high degree of identification and tracking of the players in the court. We directed investigations on soccer, basketball, ice hockey and pedestrian datasets. The trial comes about an exhibit that our technique can precisely recognize players under testing conditions. Contrasted and CNNs that are adjusted from general question identification systems, for example, Faster-RCNN, our approach accomplishes cutting edge exactness on three sorts of recreations (basketball, soccer and ice hockey) with 1000×fewer parameters. The all-inclusive statement of our technique is additionally shown on a standard passer-by recognition dataset in which our strategy accomplishes aggressive execution contrasted and cutting-edge methods.

Keywords: Player detection, basketball game, player tracking, court detection, color classification, mapping, pedestrian detection, heat map.

1 Introduction

Automated player detection and tracking in team-sport games is of growing importance [Petilla, Yap and Zheng et al. (2018)]. As the profits from sports are increasing substantially, teams are heavily in-vested more in gathering statistics on their athletes. Certain statistics, such as distance run during a match, can provide information on player's health [Barris and Button (2008)]. Moreover, real-time detection of players can be valuable in identifying the opponent's formation and strategy, and might give some insight on the likelihood of a certain play be successful. This can lead to better strategies [Lefevre, Bombardier and Charpentier et al. (2018)]. One of the principle challenges in sports science is a target examination of player execution. While an individual player's physical capacities can be promptly tried in the research center, a group's execution must be seen amid a real diversion. This procedure may incorporate propelled investigation strategies, for example, video recording and measurable examination, however it by and

¹ Anna University, Chennai, Tamilnadu, 600025, India.

² Coimbatore Institute of Engineering and Technology, Coimbatore, Tamilnadu, 641109, India.

* Corresponding Author: P. K. Santhosh. Email: santhoshpkphd@gmail.com.

by depends on perception and manual explanation by sports specialists, with the potential danger of winding up excessively subjective. Also, manual comment is a tedious and dreary undertaking, generally constrained either to scholarly research or to the modest number of groups that can manage the cost of an adequate number of qualified specialists [Le, Lefevre, Bombardier et al. (2018)].

In addition, a few analysts have discovered that even games specialists regularly can't watch and review every one of the points of interest that can demonstrate critical for the right translation of the outcomes. This is the explanation behind an expanding volume of research worried about the programmed or self-loader acknowledgment and investigation of human conduct in sports. A definitive objective of such research is to create techniques for the programmed understanding and examination of a group's execution, which would show a succinct synopsis of the group's and the players' qualities, feeble nesses, and slip-ups. The principle focal point of this article is the test of watching a ball game and translating the movement of the group on the court [Thomas, Gade, Moeslund et al. (2017)]. The nature of the group has two imperative segments. The first part includes the abilities of the players, and is communicated as their specialized information. The second part is communicated as the general group strategies. All together for a group to be effective, it needs people with amazing specialized aptitudes. By and by, these people must have the capacity to act together as a gathering-an errand that requires great coordination between the individual players and must be accomplished with a considerable measure of preparing [Kamble, Keskar, Bhurchandi et al. (2017)]. Following this test, the focal point of this article is on an investigation of composed movement in group activities, specifically basketball [Shukla and Danganwala (2017)].

It is generally acknowledged that the best players in basketball game have the capacity to respond diversely in comparable circumstances. On the group level, the circumstance is comparable great groups can rapidly change their strategies if necessary. Such conduct keeps the restricting groups from setting up a decent resistance, and keeps the play intriguing, as the two groups need to constantly adjust to the circumstance on the court. This presents a specific level of unpredictability and irregularity to a group's execution and makes the outline of a completely programmed examination framework, which would perceive, comprehend, and grade each conceivable circumstance on the court, to a great degree troublesome. In any case, because of the idea of a game's standards and thorough player preparing, the movement of the players over the court is not completely arbitrary, and it is sensible to expect that it is conceivable to separate some normal highlights of the group, particularly while considering that organized group play is typically rehearsed ahead of time [Kagalagomb and Sunanda (2017)]. This exploration paper can assemble data in regards to the situation of the players in the court, also, as information identified with the style of play of each group. Such information could be crucial to the winning of matches, if well analyzed by the team's coach [Liu, Yan and Liu (2017)]. More generally, our research could be adapted to the tracking of players of any sports match.

The remainder of this paper is structured as follows. In the rest of this section, a short overview of the related work presented in Section 2. The proposed methodology is presented in Section 3. The analysis and results are presented in Section 4. The discussion

and Conclusions are presented in Section 5 and 6 respectively.

2 Related work

A great deal of work concerning mechanized movement displaying an investigation has been introduced for bushel ball gamed lately. For instance, Johnson and Hogg [Johnson and Hogg (1996); Johnson and Hogg (2002)] introduced two methodologies for displaying variable, non-straight conduct. In their first approach an aggressive learning, a neural system was utilized on flow vectors from picture groupings of walkers [Johnson and Hogg (1996)], and in the other, a probabilistic movement demonstrate was acquired utilizing a Gaussian blend show, speaking to the framework state changes of a person on foot movement [Johnson and Hogg (2002)]. The more refined, direction based, multi-specialist, activity acknowledgment, and examination approaches include the conservative portrayal and displaying of activities and connections and their coherent and fleeting relations [Rao, Yilmaz and Shah (2002); Li and Woodham (2005); Intille and Bobick (2001); Hogeng and Nevatia (2001), Jug, Pers, Dezman et al. (2003)]. Rao et al. [Rao, Yilmaz and Shah (2002)] introduced a way to deal with see invariant activity acknowledgment that is equipped for clarifying an activity as far as significant activity units called dynamic moments and interims. Li et al. [Li and Woodham (2005)] displayed a framework for speaking to and thinking about chosen ball plays in view of direction information, increased with area specific learning, for example, forward/in reverse skating, puck ownership, and so forth. Intille et al. [Intille and Bobick (2001)] fabricated models of basketball plays utilizing conviction systems and fleeting diagrams.

A comparative approach was utilized by Jug et al. [Jug, Pers, Dezman et al. (2003)] to survey group execution in ball offense. The fundamental commitment of the last two methodologies is the portrayal of multi-specialist movement and acknowledgment from uproarious direction information. This is finished by partitioning the multi-operator action into individual outwardly grounded, objective based natives that are probabilistically coordinated with the low-arrange transient and coherent connections. Nonetheless, there are two fundamental issues with such an approach. The first one is the requirement for an exact transient division of the examined directions. The second issue is the trouble of building transient and coherent connections, particularly due to the wide range of parameters that should be defined physically. In this way, such an approach isn't especially appropriate for situations when either a huge amount of information must be examined or a wide range of conduct models are utilized as a part of the examination. Other game related research has concentrated on the substance based ordering of video film, learning movement models in baseball [Jug, Pers, Dezman et al. (2003)]. The long haul point of the exploration depicted above is to furnish competitors with the criticism they have to enhance their kinematic abilities. It can likewise be utilized to create and screen the right execution of certain pre-defined mechanical activities (e.g., ball taking care of) or to take in the right execution of more unpredictable group exercises.

In our work, we address the issue of a computerized examination of b-ball, with the point of defeating the depicted issues. We picked our approach in view of master sports learning. Like with the strategy utilized as a part of games explore, we play out a two-advance investigation process, where the amusement is first portioned by the periods of

play (offense, protection, time out), and afterward every one of the sections is broke down in detail. In order to achieve our end goal of a two dimensional image with player positioning, we made use of a five step algorithm, each of which will be further expanded:

- 1) Court Detection-find lines of the court;
- 2) Individual Detection-detect individuals standing on the court;
- 3) Color Classification-Separate these individuals into two teams;
- 4) Player Tracking-Keep positions information frame by frame;
- 5) Mapping-translate onto a court

The data for this research work consisted of multiple YouTube videos which were then cropped in order for us to do our analysis. We mainly selected videos in which we were able to see all major lines of the court in order to accurately perform the homography.

3 Methodology

The block diagram of the proposed methodology is illustrates in Fig. 1. The subsequent subsections are discussed in each stage of the algorithm in detail.

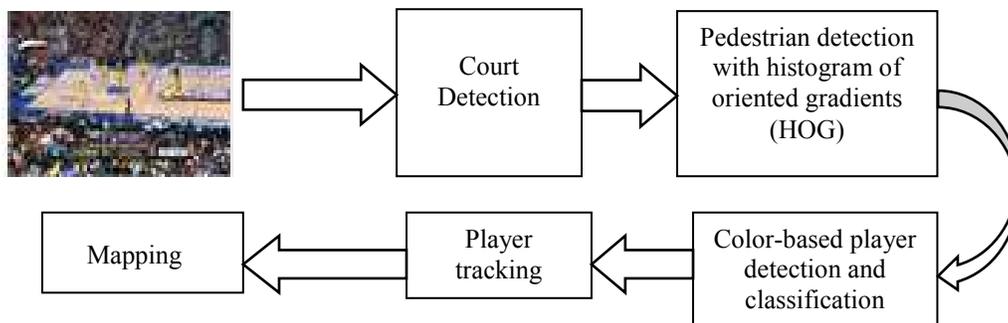


Figure 1: Block diagram of algorithm for player detection and tracking

3.1 Court detection

The video frames that we obtained from YouTube were initially converted from the BGR to the HSV (hue, saturation and value) color model. We then focused on the H-plane in order to create a binary model of the system. Then, we proceeded to perform erosion and dilation of the image in order to get rid of artifacts that were not related to the court. Subsequently, we made use of the Canny edge detector to detect the lines in our system. Finally, we performed the Hough transform in order to detect the straight lines in the system. This process is illustrated by Fig. 2.

3.2 Pedestrian detection by histogram of oriented gradients

The next stage is pedestrian detection through Histogram of Oriented Gradients (HOG). HOG essentially builds histograms of the gradient orientations in localized portions of an image, which can be used to identify objects in a image. While it is difficult to establish definite characteristics for these histograms in order to detect a certain object, machine learning classifiers, such as support vector machine (SVM), can be used to identify a

desired object in an image based on a training data set. For pedestrian detection no single feature has been shown to outperform HOG. However, the performance can be improved by using additional features to provide complementary information [Dollar, Wojek and Schiele et al. (2012)].

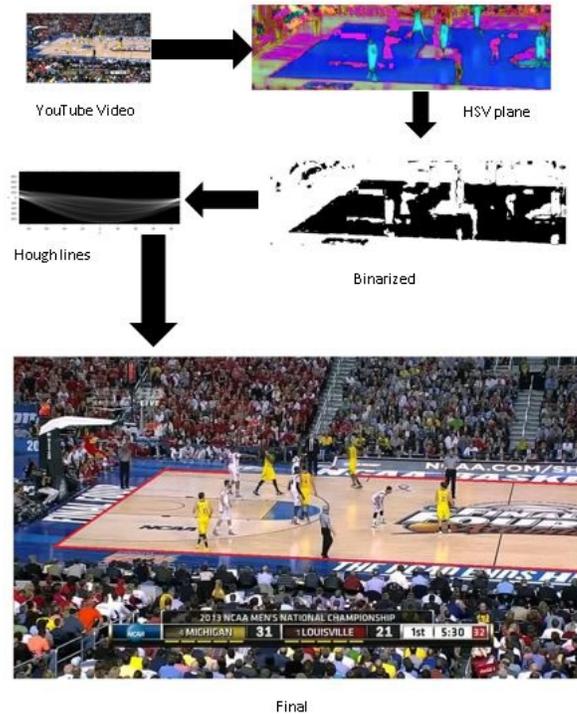


Figure 2: This schematic represents the work performed in order to detect the edge lines of the system

For this research work we used the HOG detector from OpenCV. This was mainly motivated by the fact that OpenCV has already a default data set for pedestrian detections, and that the HOG feature calculation and SVM were already efficiently implemented. We used the “Daimler” dataset for pedestrian detection. This detector is trained using window size of 48 by 96 pixels. Thus, the HOG detector expects pedestrians to be of at least that size.

Fig. 3 illustrates an example of pedestrian detection of basketball players using HOG descriptors and SVM classifier. The inset in that figure illustrates the HOG descriptors of one of the players as well as an inverse, which corresponds to a reconstructed image from the HOG descriptors. This image shows that HOG descriptors carry significant information of the detection. In the sample frame all players were detected, but there were some false positives, and two players were detected by the same box because they were too close together.

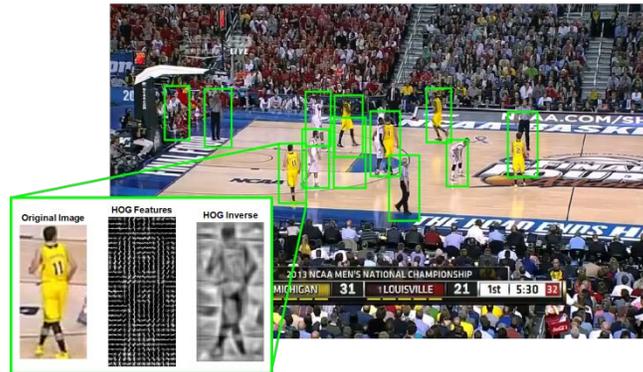


Figure 3: Example of basket player's detection using openCV pedestrian detection with HOG

The HOG features with SVM have a miss rate of about 70% for pedestrian detection [Dollar, Wojek, Schiele et al. (2012)]. However for this particular application, the accuracy of the HOG detectors is expected to be smaller, since the players can fall, jump, or crunch to get the ball, and consequently will not be detected by the HOG detector. Thus, it is necessary a redundancy system to detect the players when the HOG detector fails. To this end, we built a color-based detector and classifier. This color-based detector detects the players based on their jersey's colors. The purpose of this detector is two-folded. Firstly, it is responsible for classifying the players according to their team as well as ruling out other "pedestrians" that might be detected by the HOG detector such as referees, coaches, audience members, for example. Secondly, the color detector should identify players within a HOG box. Quite often some players will be too close together or partially obstructed by other players. In these situations the HOG detector might return a single detection (box) for all those players.

3.3 Color-based detection and classification

The color-based detector performs player's detection within a HOG box, which is a region of the original image classified as a pedestrian by the HOG detector. Color-based detection could also be used in a larger image (not necessary in a HOG box only), for example detecting players in the entire court as done in previous projects [Lucas and Kanade (2009); Jug, Pers and Dezman]. However, the HOG detector greatly improves the performance of this color detector, since the HOG boxes limit the scope to a couple of players. Thus other objects that might have the same color of the player's jerseys, such as details in the floor typically found in basketball courts, will not be detected as often as if an entire frame were used.

The color detector performs detection by using thresholds in the HSV space. The choice of the HSV space as oppose to RGB was motivated by the fact that the HSV enables higher discrimination between changes in color rather than saturation and brightness. For instance, a RGB-color-based initially implemented would constantly obtain false positives given by the reflections on floor. Given a set of images containing players from both teams, referees, members of the audience, etc. The histograms for all coordinates H,

S, and V are calculated and thresholds are calculated using Otsu's method. Depending on the color of the two teams more than one threshold might be necessary to allow distinction between other elements that might appear in the image (e.g., yellow and floor). Once these thresholds are known we can derive logical expressions for color detection. For instance, yellow corresponds to (60°, 100%, and 100%) in the HSV space, whereas white corresponds to (0°, 0%, and 100%). Thus, we can distinguish between yellow and white by requiring the hue and saturation coordinates to be higher than their respective thresholds for yellow, and below a certain threshold for white.

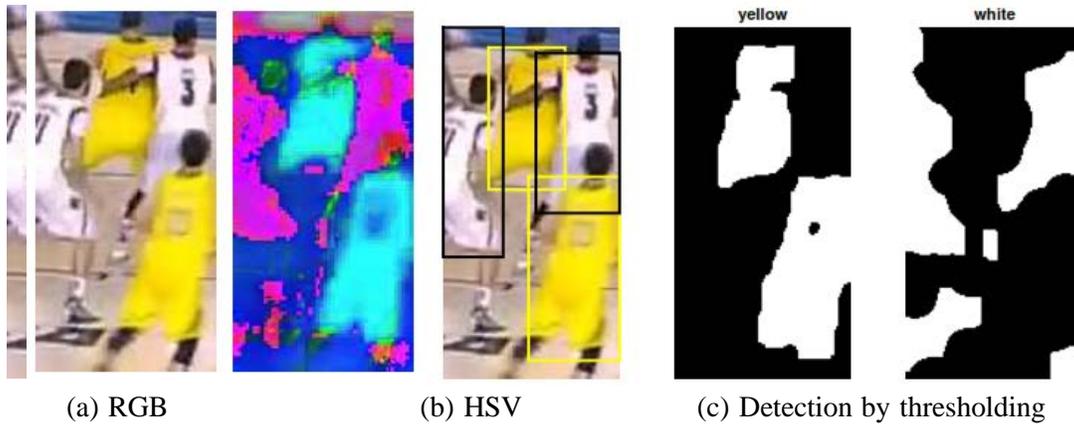


Figure 4: Illustration of color-based classifier. (a) represents the original image in RGB, (b) is the original image converted to HSV, (c) is the image after binarization through thresholding. Closing and dilation using circle of radius 10, and 15×5 rectangle is performed in the resulting binary image. The algorithm selects only boxes that meet certain criteria in size and extent. Thus, smaller boxes are ignored and false positives are avoided. As we can see in (c) all four players were classified correctly

Fig. 4 illustrates the color detection process for a white vs. yellow detection. Once the thresholds are determined from a training data set, the color detection can be performed fairly fast since it basically comprises of comparisons and logical operations. To avoid some false positives, some additional criteria is enforced. For instance, boxes that correspond to more less than 5% of the image area are disregarded. Moreover, valid boxes are expected to have height higher than width since basketball players are tall and they are standing. The color detection also plays an important role as a backup of the HOG detector. As the game proceeds some players might become partially obstructed by other players, or they might simply not be detected by the HOG detector in a certain frame. In these different scenarios, the color-based detector will be called to find the missing player by performing detection in small neighborhood of the corresponding box from the previous frame. This works fairly well because we used videos of rate of 24 fps, thus the player is expected to be in the surroundings of where he was in the previous frame. This different condition in adding players and dropping players in the detection is done by the tracking algorithm discussed in the next subsection.

Fig. 5 shows two sample frames after color detection. Note that Fig. 5(a) corresponds to the same frame of Fig. 4, and the false positives were eliminated and the ambiguity of

two players being detected by the same box was eliminated.



(a) Michigan vs. Louisville

(b) Michigan vs. Syracuse

Figure 5: Example of player's detection using HOG pedestrian detection and color-based classification in two different games.

3.4 Tracking

Once detected, the next goal is to establish a frame by frame positioning of the individual players in order to understand the play in total. Thus, a tracking algorithm kept track of the players' movements. This algorithm used the information from the previous frames for initial conditions on tracking. We dealt with the following scenarios.

3.4.1 Scenario 1: player detected in consecutive frames

In this scenario, a player was detected by the HOG detector and color classifier in back to back frames. Because the frames are taken at a rate of 24 fps, the position of player from one frame to the next are highly related. Thus these boxes that are detected would have high overlap. In this case, the position of this player is updated with the new position, and the new position is saved.

3.4.2 Scenario 2: neighborhood estimate

In this scenario, a player was not detected by the HoG detector, but can be found by a color detection in the neighborhood. With this frame-by-frame scenario, the HOG does not detect every player in every frame. Players can be in motion and blurred, crouching, or in one way be undetectable to the HOG detector in a particular frame. A player detected in the previous frame would then have no correlation to a player in the current frame. We search for this missing player within a 20 pixel bound around the location of the player in the previous frame. If the color detector can identify a player with the same jersey within this box, it matches this player to this new position. If multiple players are found within this box, it matches this player to the closest previous position. Thus lose a HOG detection for a frame does not result in the loss of the player' positioning.

3.4.3 Scenarios 3 and 4: adding and dropping

The next two scenarios, though quite different, have a similar solution:

Scenario 3: A player was not originally detected in the first frame, but was found at a

later time. In this scenario, a player was not detected in the 1st frame. The HoG detector was not able to originally find this player. This scenario will be referred to as an add. In the original frame, 9 players were detected on the court in Fig. 6(a). however, in a later frame, this 10th player was added in Fig. 6(b).

Scenario 4: A player who had been previously identified by the HoG detector and color classifier was dropped. This will happen when players of the same team merge together, and the color classifier is unable to distinguish them as separate players. This scenario will be referred to as a drop. In the first frame for example, one is able to distinguish five separate players on the yellow team in Fig. 6(a). However, because the players converge, only 3 yellow masses are distinguishable in Fig. 6(c). Therefore, information is lost.

These scenarios have a combined solution, a minimum distance correlator. On every single frame, the boxes and positions are stored. If a player is dropped, after a certain period of time, the HoG detector will detect him again. Because no box from the previous frame correlates to him, he will have had no established previous position. The algorithm checks to see if a player who had been dropped is relatively close to his position. The distance allowed is fixed distance multiplied by the number of frames since this player was lost. If this player was within the distance of any previously dropped players, he is then correlated back to this original player, combating the drop issue. However, if no previous player was within range of this current player, a new player is added to the tracking data, combating the add issue.

This tracking algorithm solving scenarios 3 and 4 is far from perfect. Tracking itself is highly dependent on the detection information. False positives from the bench or the court can lead to bad minimum distance guesses if a player is lost. With players merging on the court, solving the drop problem is especially difficult. Players' motion change rapidly and do get convoluted, and using a players' original motion is not reliable to match missing players to previous frames. The algorithm is also sensitive to camera jitter. The best solution to this add and drop problem would be to have multiple stable camera viewpoints, which is not available to us from a broadcast, but would be available within a professional environment such as a NBA team.



a) The first frame of the play (b) A player was added (c) 3 players merged

Figure 6: Different tracking scenarios that might occur throughout the game

3.5 Mapping via homography

The last step was related to the projection of each player's location in the top-down view of the court. By having the dimensions of the court, we are able to find a 3×3 homography matrix that is computed using an affine transform. Each player's

position is then multiplied by the homography matrix that projects them into the model court, as shown in Fig. 7.

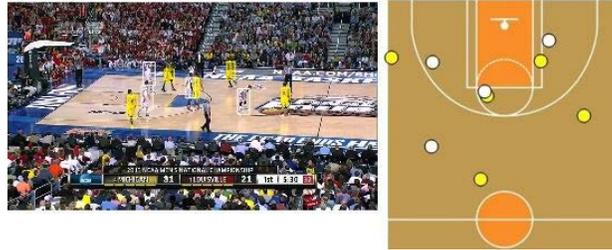


Figure 7: Comparison between the players detected and the projected image in the court. As you can see the players match the position in the top-down court model

4 Analysis and results

In this section we use the algorithm developed to analyze the positions of the players as the play progresses. As an example, Fig. 8. shows the heat map of the players positioning as the play progresses. The position of the team is compatible with the video since we can see that the white team remains on the defense throughout the play and the yellow team only crosses the white team's defense line to score a basket.

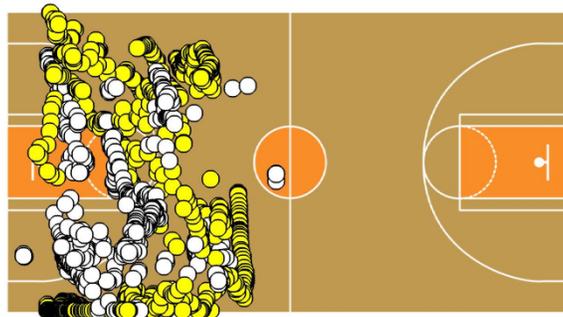


Figure 8: Heat map of the player's location between teams yellow and white

A similar heat map can be created for the second game between Michigan (represented by yellow) and Syracuse (represented by red), Fig. 9. We can see the distribution of players in the field is compatible to that seen in the video, which indicates that our player detection and homography are working in conjunction in order to provide us with useful data related to the players location in the court.

In order to further validate the results of our tracking let's break down the Syracuse and Michigan game into the two teams. Then, if we analyze only the positions of the players from Michigan we get the image compatible with Fig. 10. for the first 60 frames. In those, we can see that we are correctly labeling players 1 through 5 in the Michigan game. Furthermore, there movements seem to be compatible to those represented in the video. What is most impressive, however, is the ability of our player identification and tracking system to capture the movement of one single player that crosses the court, passing through many other players and is still correctly labeled and identified. This can

be perceived by Fig. 11.

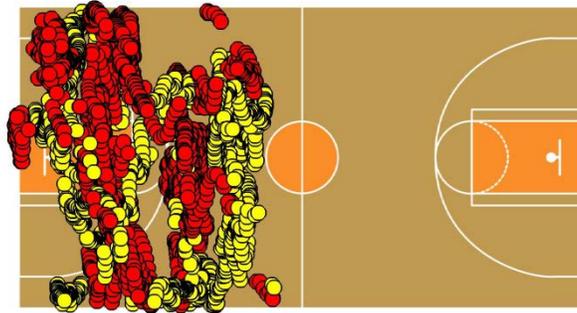


Figure 9: Heat map of the player's location between teams yellow and red

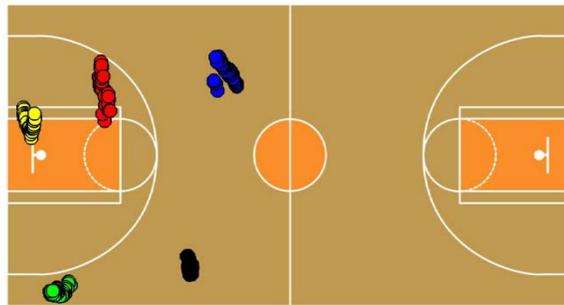


Figure 10: Positions of players from Michigan team. Each player is represented by a different color. The movements of the players seem to be compatible to that of the video

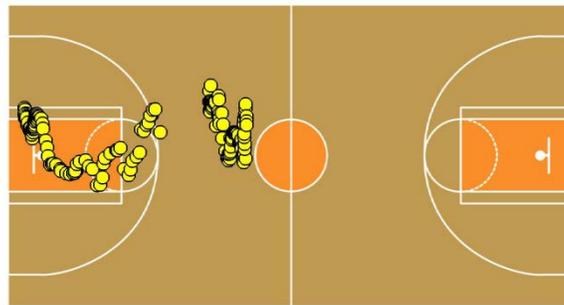


Figure 11: Position of Michigan player identified as player 1. As we can see the color segmentation and the tracking appear as a viable way to track this player across the court

5 Discussion

The comparison was done with the proposed approach with Integral Channel Features (ICF) [Dollar, Wojek and Schiele et al. (2010)], Faster Recurrent Convolution Neural Network (RCNN) [Ibrahim, Muralidharan and Deng et al. (2016)] and Single Shot Multibox Detector (SSD 512) [Liu, Anguelov and Erhan et al. (2016)]. Fig.12. shows the detection performance is measured by Area Under Curve (AUC) of ROC curve for Soccer, Ice Hockey dataset. From the Fig. 12 our method indicating higher generality of our approach.

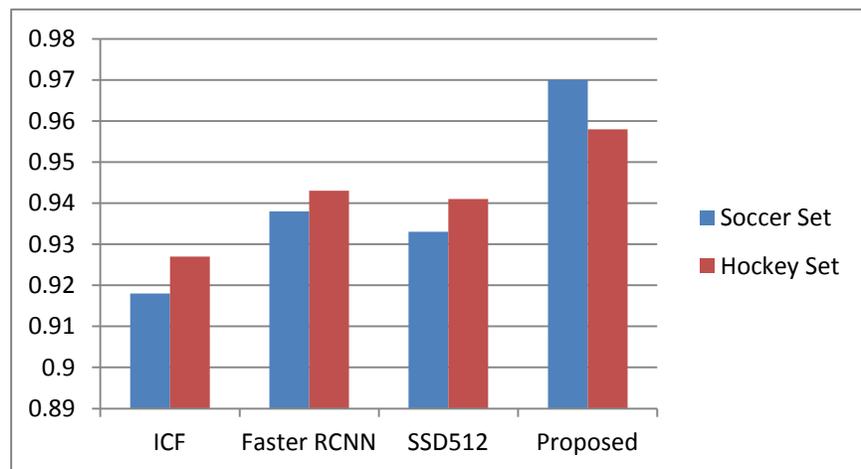


Figure 12: Comparisons of our method with other state of the art method

We also analyze the performance on players with different pixel resolutions (player height in pixels) in Fig. 13. Our method is more robust than other methods such as ICF [Dollar, Wojek, Schiele et al. (2010)], Faster RCNN [Ibrahim, Muralidharan and Deng et al. (2016)], SSD512 [Liu, Anguelov and Erhan et al. (2016)] when the pixel resolutions declines.

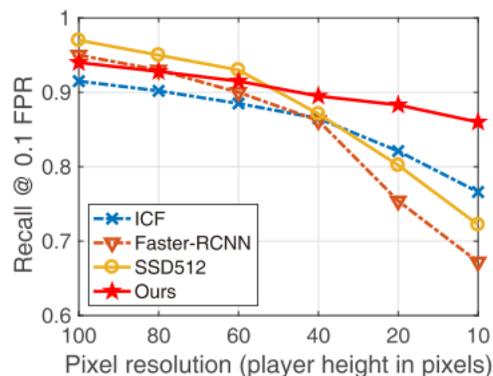


Figure 13: Recall with different pixel resolutions. The plot shows recall rate @ 0.1 false positive rate (FPR) as a function of players' pixel resolution on the soccer dataset

To further evaluate the generality of our method, we perform experiments on the ETH Zurich pedestrian dataset [Lucas and Kanade (2009)]. The performance of our method is compared with several state-of-the-art approaches: VJ [Viola and Jones (2004)], HOG [Dalal and Triggs (2005)], FPDW [Dollar, Wojek and Schiele (2012)], ICF [Dollar, Belongie and Perona (2010)], DPM [Felzenszwalb, Girshick and McAlester et al. (2010)] and ContDeepNet [Zeng, Ouyang and Wang (2013)]. As suggested in Dollar et al. [Dollar, Wojek and Schiele (2012)], we use miss rate vs. False-Positive-Per-Image (FPPI) curves to analyze the performance on pedestrian detection. A lower curve indicates better performance. As shown in Fig. 14, our method is better than ICF, FPDW, HOG and VJ. It is also competitive with the state-of-the-art ContDeepNet method which is specifically designed for pedestrian detection.

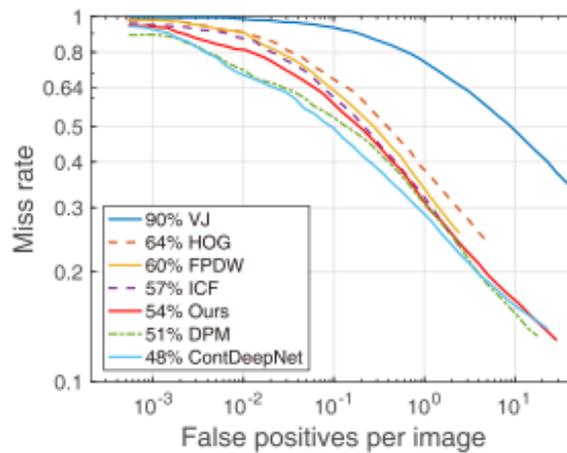


Figure 14: Comparison on the ETHZ pedestrian dataset Lucas et al. [Lucas and Kanade (2009)]. The percentages in the legend are log-average miss rates of these methods. Lower curves stand for better performance

The qualitative results of our method are shown in Fig. 15. It successfully detects players and pedestrians under various conditions. For example, it detects players/pedestrians with different aspect ratios on all the datasets, and accurately outputs team membership on soccer basketball datasets.



Figure 15: Qualitative results. The proposed method is able to effectively reject non-player samples stage by stage and output accurate bounding boxes for players and pedestrians. In this figure, NMS stands for non-maximum suppression. Different colors indicate different team membership

6 Conclusions and future work

We have developed an algorithm that accurately detects basketball players in a video and is able to accurately place them in a 2D top-view court. We have achieved such results making use of an association of the hog detector from OpenCV and color segmentation. Our method is extensively tested on soccer, basketball, ice hockey and pedestrian datasets. The experimental results suggest that the proposed approach is light, effective and robust compared with many state-of-the-art detection methods. Future work for this project would involve better adjusting the color segmentation thresholds to avoid artifact identification as well as improving the accuracy of our tracking system.

References

- Barris, S.; Button, C.** (2008): A review of vision-based motion analysis in sport. *Sports Medicine*, vol. 38, no. 12, pp. 1025-1043.
- Dalal, N.; Triggs, B.** (2005): Histograms of oriented gradients for human detection. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 886-893.
- Dollar, P.; Belongie, S.; Perona, P.** (2010): The fastest pedestrian detector in the west. *British Machine Vision Conference*, vol. 2, no. 3, pp. 1-11.
- Dollar, P.; Wojek, C.; Schiele, B.; Perona, P.** (2012): Pedestrian detection: an evaluation of the state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 4, pp. 743-761.
- Felzenszwalb, P. F.; Girshick, R. B.; McAlester, D.; Ramanan, D.** (2010): Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627-1645.
- Hongeng, R.; Nevatia, S.** (2001): Multi-agent event recognition. *Proceedings of IEEE International Conference on Computer Vision*, pp. 84-91.
- Ibrahim, M. S.; Muralidharan, S.; Deng, Z.; Vahdat, A.; Mori, G.** (2016): A hierarchical deep temporal model for group activity recognition. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1971-1980.
- Intille, S. S.; Bobick, A. F.** (2001): Recognizing planned, multi person action. *Computer Vision and Image Understanding*, vol. 18, no. 3, pp. 414-445.
- Johnson, H.; Hogg, D.** (1996): Learning the distribution of object trajectories for event recognition. *Image and Vision Computing*, vol. 14, no. 1, pp. 609-615.
- Johnson, H.; Hogg, D.** (2002): Representation and synthesis of behavior using Gaussian mixtures. *Image and Vision Computing*, vol. 20, no. 12, pp. 889-894.
- Jug, M.; Pers, J.; Dezman, B.; Kovacic, S.** (2003): Trajectory based assessment of coordinated human activity. *Proceedings of Third International Conference Computer Vision Systems*, pp. 534-543.
- Kagalagomb, C. G.; Sunanda, D.** (2017): Player tracking in sports video using optical flow analysis. *Proceedings of the International Conference on Data Engineering and Communication Technology*, pp. 733-739.
- Kamble, P. R.; Keskar, A. G.; Bhurchandi, K. M.** (2017): Ball tracking in sports: a

survey. *Artificial Intelligence Review*, vol. 48, no. 3, pp. 1-51.

Le, N. C.; Lefevre, F.; Bombardier, V.; Charpentier, P.; Krommenacker, N. et al. (2018): Automatic video editing: original tracking method applied to basketball players in video sequences. *Lecture Notes in Computer Science*, vol. 10884, pp. 117-126.

Lefevre, F.; Bombardier, V.; Charpentier, P.; Krommenacker, N.; Petat, B. (2018): Automatic camera selection in the context of basketball game. *Lecture Notes in Computer Science*, vol. 10884, pp. 72-79.

Li, F.; Woodham, R. J. (2005): Analysis of player actions in selected hockey game situations. *Proceedings of the Second Canadian Conference on Computer and Robot Vision*, pp. 152-159.

Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S. et al. (2016): SSD: Single shot multibox detector. *Proceedings of European Conference on Computer Vision*, pp. 21-27.

Liu, W.; Yan, C. C.; Liu, J. (2017): Deep learning based basketball video analysis for intelligent arena application. *Multimedia Tools and Applications*, vol. 76, no. 23, pp. 24983-25001.

Lucas, B.; Kanade, T. (2009): Moving obstacle detection in highly dynamic scenes. *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 56-63.

Parsons, S.; Rogers, J. (2014): Basketball player tracking and automated analysis. *Winter Project Report*. Stanford University, USA.

Petilla, C. A. B.; Yap, G. D. G.; Zheng, N. Y.; Yuson P. L. L.; Ilaio, J. P. (2018): Single player tracking in multiple sports videos. *Proceedings of Mechatronics and Machine Vision in Practice*, pp. 73-89.

Rao, C.; Yilmaz, A.; Shah, M. (2002): View-invariant representation and recognition of actions. *International Journal of Computer Vision*, vol. 50, no. 2, pp. 203-226.

Shukla, N.; Dangarwala, K. J. (2017): Survey on sportsperson positional marking system. *Proceeding of IEEE International Conference on Innovations in Information, Embedded and Communication Systems*, pp. 1-5.

Thomas, G.; Gade, R.; Moeslund, T. B.; Carr, P.; Hilton, A. (2017): Computer vision for sports: current applications and research topics. *Computer Vision and Image Understanding*, vol. 159, no. 1, pp. 3-18,

Viola, P.; Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137-154.

Wilson, M.; Giese, J. (2014): Basketball localization and location prediction. *Winter Project Report*. Stanford University, USA.

Zeng, X.; Ouyang, W.; Wang, X. (2013): Multi-stage contextual deep learning for pedestrian detection. *Proceedings of IEEE International Conference on Computer Vision*, pp. 121-128.