International Journal of Advanced Science and Technology
Vol. 29, No. 9s, (2020), pp. 5094-5105

Automated Visual Tracking and Live Data Analysis in Badminton

Mona Teja K[1*] and Prabakaran N[2]

[1,2] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore-632014, India.
[1]monateja9999@gmail.com and [2]prabakaran.n@vit.ac.in

Abstract

In the modern sports era, competition is all that matters. In order to compete in a sport, several strategies and analysis are required. Strategies against selected opponents can yield better results. But currently there is no such tool which gives the statistics without manual procedure. The model proposed will automate several processes and yields several visualisations along with the statistical data. The visual representation is a key factor for understanding the data. It works automatically on the live feed video or video stored in the database and generates the visual analysis. The code is computationally faster and the visualizations are simple. Hence, it can even be used in the live matches to aid coaches and players instantly. The model is built by using simple deep learning modules followed by basic computer vision operations and techniques. Data obtained is very simple to understand unlike spreadsheets, so even amateurs can make use of the model.

Keywords: *Automation, Player Tracking, Computer Vision, Data Visualization, Deep Learning, Image Processing, Heat Maps, Live Analysis.*

1. Introduction

In the modern era of sports, analysis has become a vital part of performance of any player or a team. Strategies and visual analysis of opponent's technique is aiding the players to improve their individual performances and identify the loop holes in the current strategies. But the problem is that in the analysis the spread sheets is making things complicated when it comes to player or coach's point of view. Even the technology used for analysis and graphics involves lot of manual procedure. The equipment is highly costly and complex to understand. Thereby automation helps us in reducing the human effort and complexity in the analysis. In order to make the model light and work faster we need to use basic and standard algorithms.

In the model that is proposed OpenCV and Deep learning is being used. The model will be able to provide us the results on a live feed or already broadcasted video. We also use a technique called homography in order to scale the things that we are viewing. As the perspectives of view changes the analysis might go wrong so we need to properly scale down the things. In this regard homography helps us in providing the standard or ideal view of the field and the players. The Idea is to provide the coaches and the players a visual analysis by which they will actually relate it to either the field of the game or player. Few of such visualizations are heatmaps, region analysis etc. Apart from these visualizations the model offers the statistics of how much each individual player is moving and to which areas he is being forced to move.

There are several methods to do player tracking but currently all those applications are quite expensive and are licensed. Only professional operators will be using that tools to get the desired output. But there is no such tool that amateurs can make use of. These football turfs and amateur leagues are also becoming popular. People want to post their plays on the social media. As the technology is quite expensive analysis is being limited to only professional games. Hence, there is a great requirement for a tool with which even amateurs can operate and get the statistical data. Model is very light and easy to understand. The visualizations obtained by the model is useful to make new strategies or understand the opponent's strategies.

2. Related Work

In the recent trends of research and development few approaches include the usage of belief propagation. In this method the information obtained in one view is used as an advantage while computing other view and there is redundant view. This method of identification works fine from some views and produce redundant outputs in case there are any occlusions or noise based on the likelihood scores [1]. In the general broadcast they use high-resolution cameras which is not cost effective. It includes lot of expenditure due to which the usage of some other algorithms such as K-Shortest Paths. By using this algorithm, the analysis can be even done by capturing through low resolution cameras with a customized Tripod-height. This method gives a great amount of accuracy but there is a possibility for failure in corner cases. Several prototypes are being made to detect and track players in multiple camera discontinuous sequences and efforts are being made to reduce the complexity and computation time [2].

There is a thesis stating procedure for multiple player detection using basic computer vision techniques. It follows a sequence of operations in order to detect the players. The first step is finding out the objects that are not stationary in the consecutive set of frames. This is achieved by well-known method i.e. background subtraction. Major portion of the image which includes the pitch or ground remains stationary. As background i.e. the pitch is stationary, foreground elements i.e. players are obtained using background subtraction. After the data is obtained from all the cameras in different views, they co-relate them using the relationship between them using a technique called Homography [3]. As there are many players present on the pitch at the same time, their paths may intersect during which identification algorithm mistakes one player for another and this phenomenon in some studies is termed as convex global optimization problem [4]. Whereas there are few models where data is fetched and processed in static multiple cameras and finally after all the processing is done the data from the different individual static camera is integrated [5]. As we have discussed earlier regarding background subtraction there are few models where they don't use background subtraction instead substitute it by tracking continuously by using a geometric model and homography analysis. Later once the elements are identified the tracking is done using local tracking modules [6].

In fact, multiple players tracking requires locating and identifying each individual player and labelling them accordingly. This is easy when the required targets are isolated. But this is not the case every time so when the isolated targets interact it's very difficult to label them accurately. This phenomenon is termed as Bayesian network inference problem [7]. Few research teams have used the basic histogram analysis in terms of Hue and Saturation for ground or pitch identification and segmentation. In case of shadow, weather and brightness issues they have incremental model to solve those issues. And this is followed by a player detection using a linear model that uses decision tree approach [8].

Now that they have detected players from the foreground that is obtained using histogram analysis, they wanted to identify the players and the Id. The principle of connected components is applied to detect the individual components which is later followed by Skeleton Pruning and reverse Euclidean transform [9]. So, by this analysis we can come to a brief conclusion that several researches are being carried out on several sports for analysing strategies, recovering scenes, automation in indexing and use artificial intelligence for obtaining statistics and analysis. There is a huge demand in the market for automated tracking system using imaging data [10].

A. Existing Gap

In the modern sports statistics of the players and opposition players play a major role in terms of the performance of the individual/team. But it is quite challenging to manually analyse each and every move. Moreover, it's further complicated when it comes to the team sports as there are multiple players, camera views and actions involved at the same time. Now a days, broadcasters perform several operations using manual annotations and thereby record the statistics.

This procedure requires many people to manually work on it. They are getting the analysis after certain amount of time post-match which does not make any sense for the current game performance. As the statistics are manually calculated there is a chance of getting marginal errors which makes data redundant and not useful. Even after obtaining the data from manual annotations, the data in the spreadsheets are quite complex for the coaches or amateurs to understand.

Due the heavy computational requirements the cost of equipment is unimaginable. In this case only the professionals and people who are able to afford can get the statistics. Even after spending huge amount of money if the data is in spreadsheets and co-ordinates, it's quite complex to analyse the data. Moreover, there are no proper sets for visualizing the statistical data. Some generic challenges that we face are colour-identification, player-motion, noise-occurrence, shuffling-views, discontinuity in frames.

3. Proposed Methodology

The prior task is to identify the players followed by identifying the individual players. The model that we are trying to develop needs to automatically adapt itself for the discontinuity in the frames and should be able to identify the players. For this we use basic OpenCV. If we want to analyse the statistics of a live match then we need to make use of parallel processing for better efficiency. The model must be light enough and automated to the maximum extent.

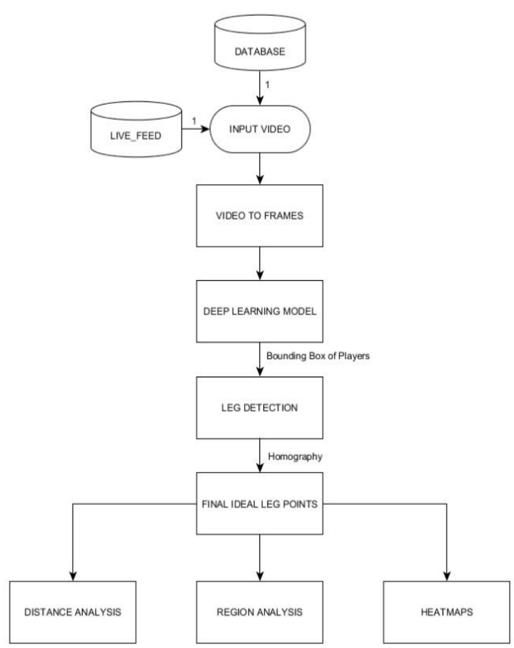

Fig 1. Architecture Diagram

ISSN: 2005-4238 IJAST

4. Implementation

Before performing any kind of analysis on any game, one must remember that the view is 2 Dimensional. Hence there might many occlusions among the players. Hence, we need to select an attribute that is unique for each player so that tracking becomes easier.

A. Training

The basic step in our procedure to identify the players. For player identification we need to train a deep learning model. In building any deep learning model the prior step is to build the data. In order to build the data, I have considered several images that are extracted from already broadcasted video using ffmpeg. ffmpeg is used to convert video to frames and frames to video respectively. In the live matches we get approximately 25-30 FPS but we consider 5 to 6 frames per second as the computational cost increases. Now, after retrieving frames from the broadcasted video, we annotate the frames and mark the extreme co-ordinates of player in the top court and player in the bottom court.

Figure 2. Training Images and respective XML's

Let's have a glance of the format of the XML's that are used as input for training. The below figure depicts the raw format of the XML that is used for the training purposes.


```xml
<annotation>
    <filename>/1_08.xml</filename>
    <size>
        <width>1280</width>
        <height>720</height>
        <depth>3</depth>
    </size>
    <segmented>0</segmented>
    <object>
        <name>0</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <difficult>0</difficult>
        <bndbox>
            <xmin>645</xmin>
            <ymin>205</ymin>
            <xmax>669</xmax>
            <ymax>287</ymax>
        </bndbox>
    </object>
    <object>
        <name>0</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <difficult>0</difficult>
        <bndbox>
            <xmin>681</xmin>
            <ymin>325</ymin>
            <xmax>728</xmax>
            <ymax>440</ymax>
        </bndbox>
    </object>
</annotation>
```


Figure 3. Sample XML

Now we pass these files and obtain for training and get MobileNet SSD, a caffe model as output. In our live analysis, we pass the normal images into the model and get the image with bounding boxes surrounding the players.

Coming to the game of badminton leg points are the unique attributes that never get occluded as players of different teams are in different regions of the field. This is the reason why identifying leg points is a key factor for tracking the players. Refer Link given in annexure for visual understanding.

B.Leg Detection

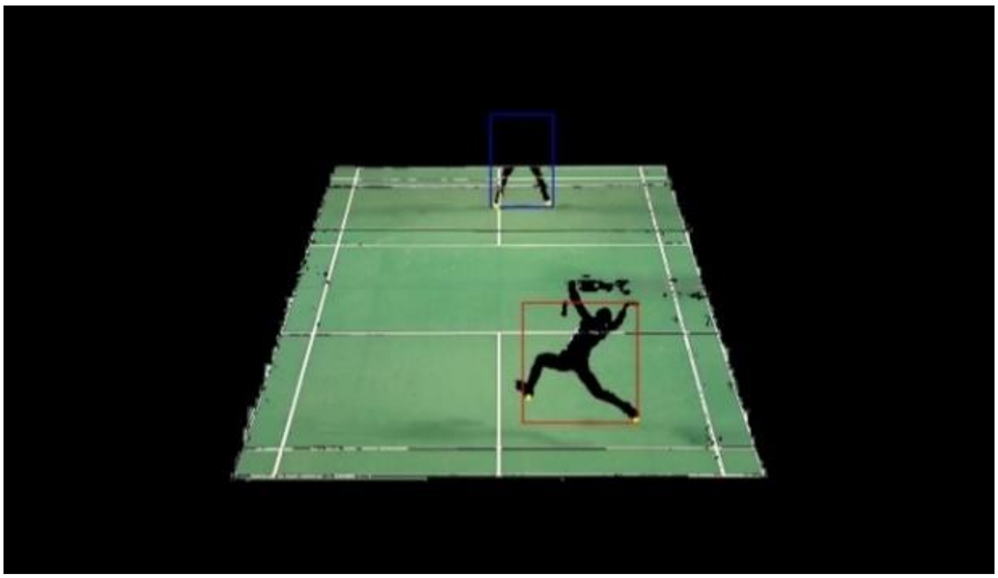

Figure 4. Leg Detection

First, we need to mask the field and lines for achieving greater accuracy. The dimensions of the court differ from stadium to stadium. Hence, we mark four extreme points of the field to get field mask and lines for getting analysis and statistics. This is just to ensure the higher accuracy

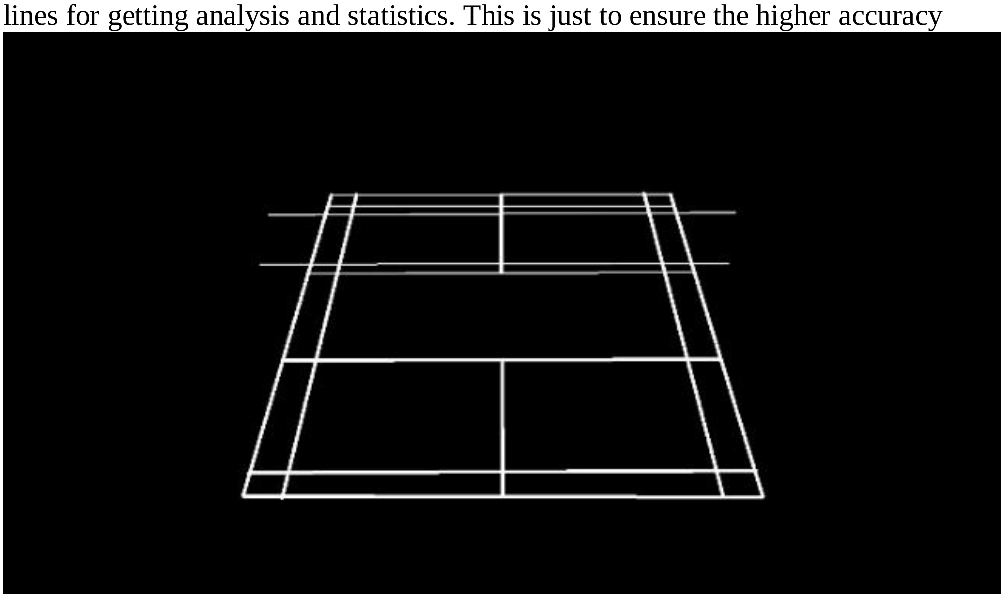

Figure 5. Line Mask

Now, we convert the input video (broadcast video or live feed) into frames. Now after this conversion we perform our operations on each frame. We will pass all frames through a model trained for detection of players on the court using MobileNet architecture. Now from this model we get an image and a bounding box around each player. From here we perform operations on the bounding boxes to locate the legs of the players. It is followed by homography to scale the top and bottom player ideally.

C.Player Tracking

Now we create two NumPy arrays of half court shape to individually analyse each player. Following this we will find the centroid of the legs and update the respective NumPy array from centroid (x, y) location as the centroid moves over the frames. Now we calculate the Euclidean distance for previous centroid which depicts the distance moved by the player from previous frame to the current frame. We create two more arrays of same shape initialized to zeros as the centroid keeps proceeding to a location (x, y) we need to increment the value at the Xth row and Yth column of our 2D array. This ends up giving us the frequency matrix of player reaching out to location (x, y) of the court during the course of the game.

Figure 6. Player Tracking

D.Homography

Now we use homography to scale the players and generate their statistics such as distance and heatmaps after scaling. In the following the ideal pitch displayed

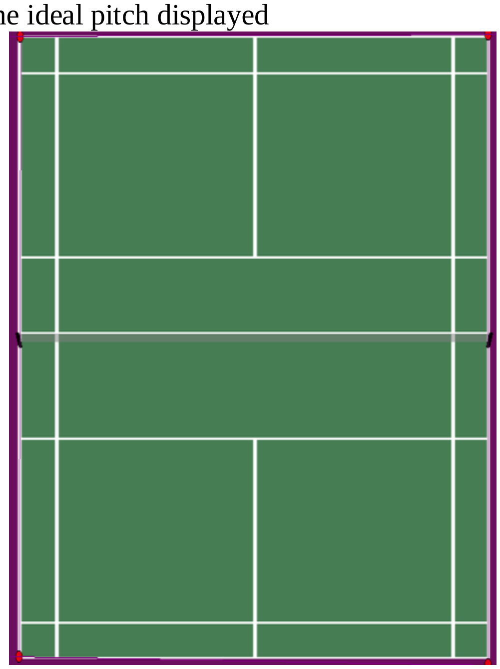

Figure 7. Homography Visualization

E. Heatmaps

As we obtain the frequency matrices of each player after homography. We create a customized heatmap by selections the range of the colours that need to be transitioned from low to high. Low Transition colour indicates that player has visited that particular part of the court a smaller number of times. Same its vice-versa for the high transition colour. We divide our frequency ranges into bins and for the top fraction of the bins (i.e. 5% for example) is considered and respectively for each transition colour accordingly heatmaps are plotted.

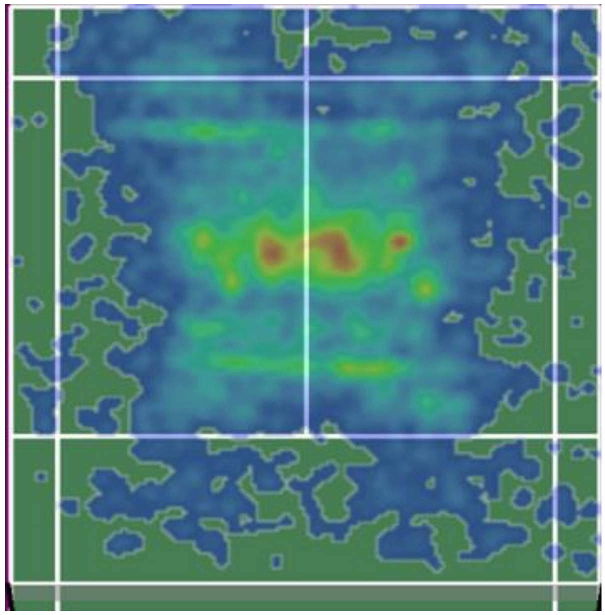

Figure 8. Heat Map

But, we are broadcasting we create a extra channel called alpha channel apart from RGB channels. We consider the shape of alpha channel to same as any of the three channels. Now we operate on alpha channel and generate heatmaps on the alpha final. Finally, this alpha channel is overlaid with some bias on to the live broadcast feed with some opacity or graphics. And the below figures depict the alpha channel.

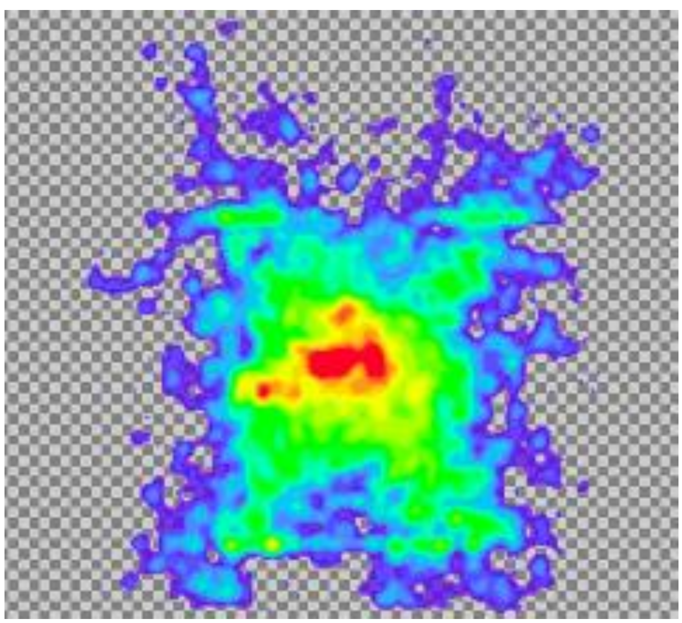

Figure 9. Alpha Channel

F.Distance Analysis

We calculate the distance metrics by the Euclidean distances between the current player centroids in this frame and centroids of respective players in the previous frame and store the output in this respective format. Time_Stamp, Player_ID, Rally_Distance, Total_Distance, Player_ID, Rally_Distance, Total_Distance, Rally_Start_Frame, Rally_End_Frame.

```
07:02:37.29 1 25.81 25.81 2 28.72 28.72 117 189
07:02:47.31 1 12.53 38.34 2 10.28 39.0 250 289
07:03:00.52 1 23.56 61.91 2 26.04 65.04 353 413
07:03:07.56 1 13.02 74.92 2 15.79 80.83 486 514
07:03:20.61 1 28.58 104.01 2 23.71 106.28 592 649
07:03:37.57 1 27.78 131.79 2 29.64 135.92 724 805
07:03:59.61 1 41.25 181.22 2 48.59 200.85 912 1011
07:04:08.58 1 19.61 204.93 2 15.62 220.04 1122 1159
07:04:13.66 1 8.33 213.26 2 8.94 228.98 1284 1301
07:04:17.35 1 6.23 219.49 2 9.44 238.42 1367 1383
07:04:30.94 1 23.34 242.83 2 21.92 260.34 1443 1506
07:04:38.02 1 10.87 255.07 2 12.1 277.58 1715 1742
07:04:48.31 1 20.85 275.93 2 19.6 297.18 1816 1865
07:05:08.81 1 33.53 309.46 2 32.52 329.69 1974 2066
```

Figure 10. Distance Sample

Cumulative distance is the total distance and we break the rally distance once the rally is done and again initialized to zero before the next rally. We can also plot bar plots for comparison of distance run by the players in each point.

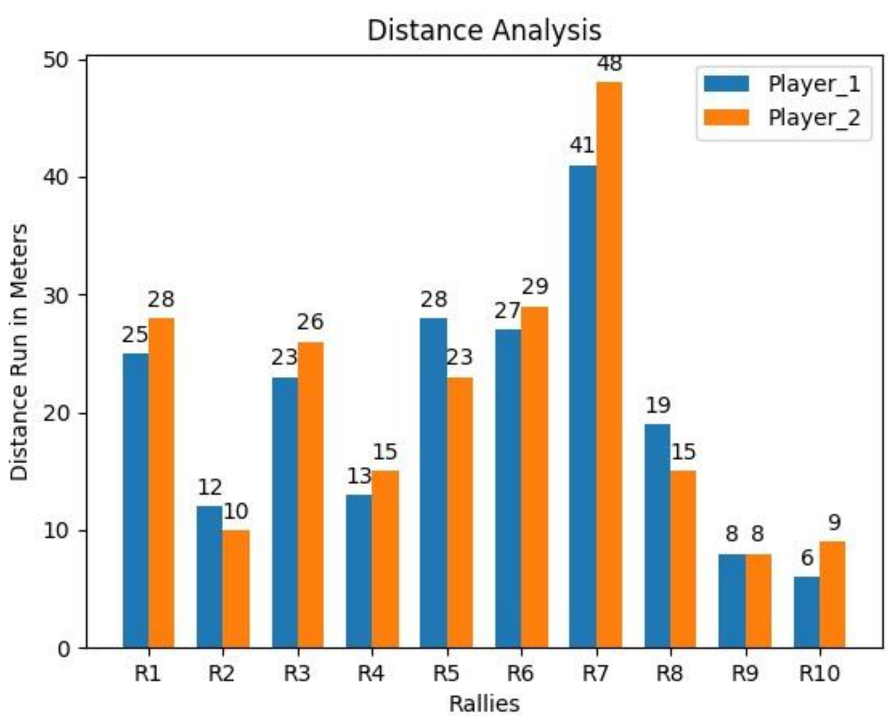

Figure 11. Distance Analysis

We will use an ideal pitch image to map the original image into ideal pitch where the statistics and player data is completely scaled and can be compared. We can also use Plotly to make the bar charts more interactive and increase the interactivity of the graphs.

G.Region Analysis

We can segment the ground into n regions and we can compare the frequencies of the players in that particular region. We can colour the region with respect to the player jersey whose frequency is greater. Thereby we can infer that in two ways i.e. particular player is dominant than another player in that respective regions. If we don't consider the difference then we can infer that particular player is being forced to that region of the court compared to the other regions.

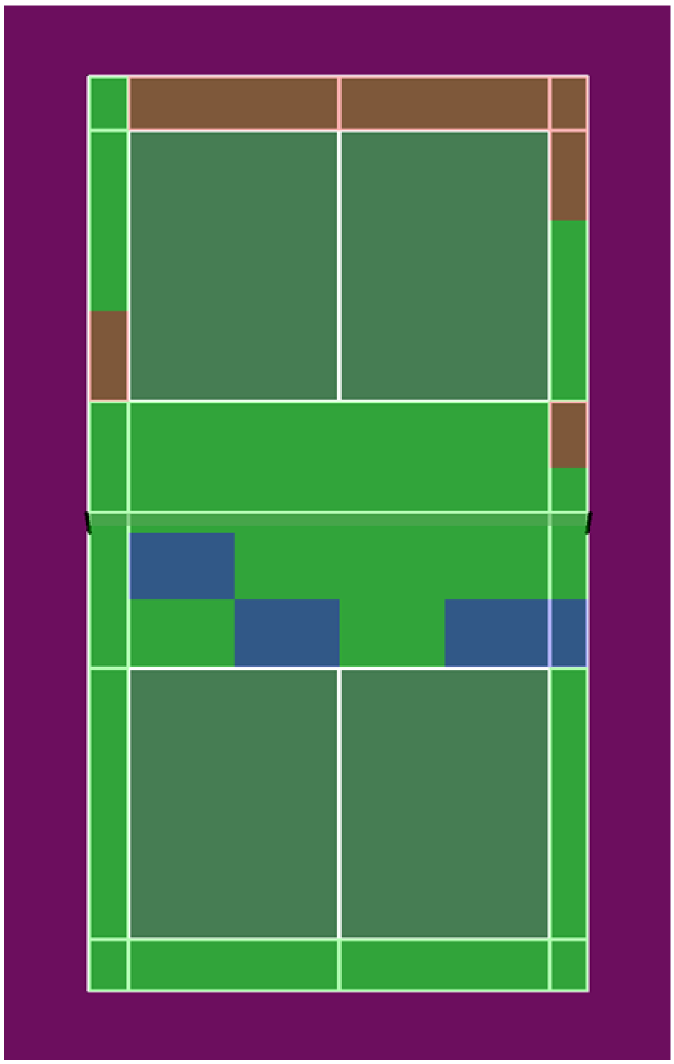

Figure 12. Region Analysis

Apart from the regional analysis currently in the broadcast the broadcasters are manually keeping track of specific attributes. For example, in badminton Fore-Hand Winners and Errors, Back-Hand Winners and Errors etc. These are recorded by the logger manually. Hence there might be errors in computing as they have to track the movements and shots of the players almost parallelly which is very difficult. Advancements are being made to implement shot prediction which automatically predicts the shot and computes in the logger. It reduces the major amount of work done by the logger. Enhancement of Bar plots in distance metrics and several other things can be expected as the outcome in future.

H.Experimental Setup

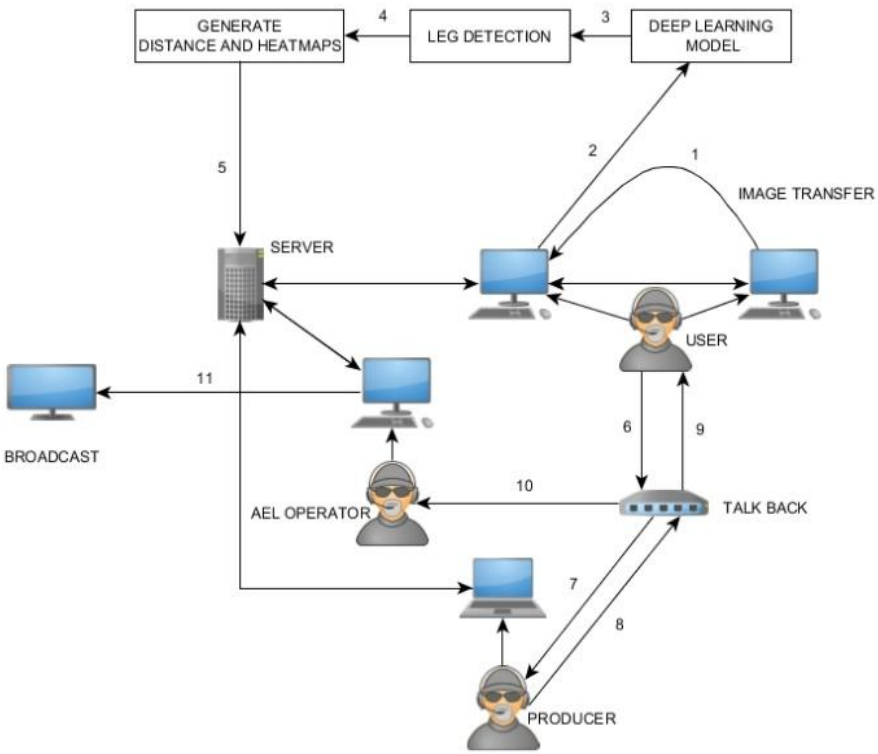

Figure 13. Experimental Setup

5. Proposed Cultural Swarm based Task Scheduling Technique

The experimental setup clearly depicts the flow and how the data is analysed and broadcasted in a detail manner. Coming to the analysis part you can see the heat maps of two different players which clearly shows that one player's center is concentrated in the center whereas the other player's center is completely scattered. This infer that one player is stay strong in the center region and using his/her reach whereas the other player is made to run all over the court. The same difference can even be clarified by observing the distance travelled by the players.

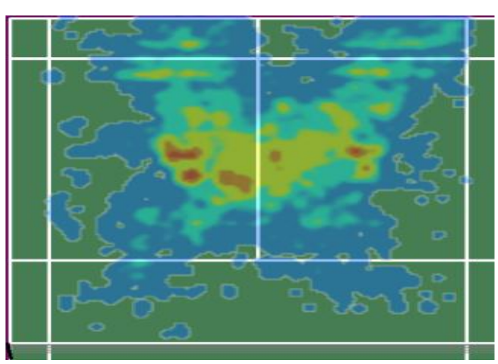

Figure 14. Comparison of Heatmaps

The experimental setup clearly depicts the flow and how the data is analysed and broadcasted in a detail manner. Coming to the analysis part you can see the heat maps of two different players which clearly shows that one player's center is concentrated in the center whereas the other player's center is completely scattered. This infer that one player is stay strong in the center region and using his/her reach whereas the other player is made to run all over the court. The same difference can even be clarified by observing the distance travelled by the players.

Note: As the game progress the players change the courts after each set of the match. At that moment we need to switch the NumPy arrays from top to bottom in order to sustain the player data of previous sets.
By making using of regional analysis we can clearly depict the area where an individual player is being forced to move. If a player is inconsistent in move to specific regions of the court, we can make use of the analysis to make that player work more on those regions. Distance analysis also plays a major role in identifying the flexibility and wear and tear of the player. If a specific player's distance statistics is always lower than the opponent it indicates that he/she needs to change their area of placement, so that opponent is forced to move towards their weaker regions.

6. Conclusion

The visualizations and the instant result can improve the player strategies and aid them in improving their performance. It also helps to trace out the flaws in their opponent's play. Detailed analysis, visualizations and challenges that are faced are explained. Player tracking is achieved at a better accuracy. Region specific analysis helps to change their game play and force their opponents to their weaker regions. Most of the implementations that are carried out in the current model is being automated and moreover achieving them at greater speed and accuracy.

References

[1] Shah-Mansouri, Hamed, Vincent WS Wong, and Robert Schober. "Joint optimal pricing and task scheduling in mobile cloud computing systems." IEEE Transactions on Wireless Communications 16.8 (2017): 5218-5232.
[2] Du, W., Hayet, J. B., Piater, J., & Verly, J. (2006, May). Collaborative multi-camera tracking of athletes in team sports. In Workshop on computer-vision based analysis in sport environments (CVBASE) (Vol. 5439, pp. 2-13).
[3] Nishikawa, Y., Sato, H., & Ozawa, J. (2018, January). Multiple sports player tracking system based on graph optimization using low-cost cameras. In 2018 IEEE International Conference on Consumer Electronics (ICCE) (pp. 1-4). IEEE.
[4] Mondal, D. C. (2014). Multi camera soccer player tracking (Doctoral dissertation).
[5] Shitrit, H. B., Berclaz, J., Fleuret, F., & Fua, P. (2011, November). Tracking multiple people under global appearance constraints. In 2011 International Conference on Computer Vision (pp. 137-144). IEEE.
[6] Xu, M., Orwell, J., Lowey, L., & Thirde, D. (2005). Architecture and algorithms for tracking football players with multiple cameras. IEE Proceedings-Vision, Image and Signal Processing, 152(2), 232-241.
[7] Hayet, J. B., Mathes, T., Czyz, J., Piater, J., Verly, J., & Macq, B. (2005, September). A modular multi-camera framework for team sports tracking. In IEEE Conference on Advanced Video and Signal Based Surveillance, 2005. (pp. 493-498). IEEE.
[8] Nillius, P., Sullivan, J., & Carlsson, S. (2006, June). Multi-target tracking-linking identities using bayesian network inference. In 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06) (Vol. 2, pp. 2187-2194). IEEE.

[9] Duh, D. J., Chang, S. Y., Chen, S. Y., & Kan, C. C. (2013). Automatic broadcast soccer video analysis, player detection, and tracking based on color histogram. In Intelligent Technologies and Engineering Systems (pp. 123-130). Springer, New York, NY.

[10] Vinoth Kumar V, Karthikeyan T, Praveen Sundar P V, Magesh G, Balajee J.M. (2020). A Quantum Approach in LiFi Security using Quantum Key Distribution. International Journal of Advanced Science and Technology, 29(6s), 2345-2354.

[11] wase, S., & Saito, H. (2004, August). Parallel tracking of all soccer players by integrating detected positions in multiple view images. In Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004. (Vol. 4, pp. 751-754). IEEE.