Automatic Extraction of Heat Maps and Goal Instances of a Basketball Game Using Video Processing

Gerónimo Eberle(✉), Jimena Bourlot, César Martínez, and Enrique M. Albornoz

Research Institute for Signals, Systems and Computational Intelligence, sınc(i) UNL-CONICET, Ciudad Universitaria, Ruta Nac. No. 168, Km 472.4, (3000), Santa Fe, USA
contacto@geronimoeberle.com.ar,
{cmartinez,emalbornoz}@sinc.unl.edu.ar

Abstract. The use of image processing techniques and computer vision to obtain teams statistics in different sports, currently represents a new source of information very useful for season preparedness. In this work, we propose different methods to show the players position for a specific time, to perform the point counting, and to extract clips of goal situations in the match. In order to accomplish this, we use a combination of transfer learning using a pre-trained deep neural network with a database of basketball game excerpts, and video processing techniques. As a proof of concept, the method was applied to a basketball game of local teams, showing the feasibility of the proposed approach. AQ1

Keywords: Game statistics · Heat maps · Digital image processing · Deep neural network · Transfer learning

1 Introduction

Data extraction in the activity of team sports is a relatively new field. Through expensive systems and cameras, such as AutoStats in the NBA, professional management teams have the possibility to perform analysis of their matches, forms of play, and statistics that allow decision making for better performance. Besides, basketball coaches spend a lot of time in clipping and analyzing a match of the rival team for show to your players the tactics of the next game. To this end, it is important to provide a tool for automatically clip and extract highlights of the games. The motivation arises to reduce the prices of this kind of technology, so that small regional teams could have the possibility of analyzing their matches in a similar way -automatically- without the expense of time that represents doing it manually.

It is of interest to develop a method that, through software, allows to replace the expensive sensors that are used in the big leagues to track the players of a game, which are placed in different parts of the body, being somewhat annoying

P. Pesado and G. Gil (Eds.): CACIC 2021, CCIS 1584, pp. 1–11, 2022.
https://doi.org/10.1007/978-3-031-05903-2_7

for the athlete in his/her activity. Some examples are those that are located as girdles in certain areas such as the knee [17], chips that adhere to the athlete's body [16], or even inserted in mouthguards [12]. Microelectromechanical devices (MEMS sensors) can also be considered, which are located on the wrist of the players to allow the tracking of the players and monitor their performance [19] and [22]. Another aspect that can be addressed as a complement to the monitoring of the players, is that of the location of the ball on the field. There are different approaches today to carry out this task, such as multisensory radio frequency systems that combine player detection with ball tracking on the field [3]; radar-based systems for the detection and tracking of balls from different sports [15], or systems that use different image processing techniques to identify balls using as input the data obtained by Kinect cameras located in mobile robots [5].

The drawback found around these alternatives is that they are usually expensive and difficult to access for smaller regional teams. That is the reason why it would be useful to find an accessible way to use computer vision resources and statistical techniques to perform the game analysis of a team, such as heat maps, point counting, or the automatic extraction of clips of goal plays.

To carry out the above, it is necessary to identify the ball and track it, taking into account the moments when the ball enters the basket, which will be the goals of the match. With this action detected, it is direct to calculate the count of goals, as well as to determine start and end times in the whole video. Thus, by performing a task of recording player locations, it would be possible to generate a heat map of the team. This could allow any team to perform automatic game analysis in an accessible way, with the addition of using a non-invasive system since no sensor should be dressed during the activity.

Usually, television cameras do not perform a complete panning of the playing field, but different sectors are captured from a rotation of them. That is why to get maps of the game, it is required as a first stage the registration of the different frames of the video, and thus it is possible to generate a panoramic image that allows to visualize the complete court. Then, the result must be transformed into an aerial perpendicular view, where the visualization of the players' heat maps would be easier.

The detection of objects in images is a very incipient field, especially with the arrival of robust models of deep neural networks [2]. These models allow us to identify objects within images, which will be useful to find the players and the basketball thus obtain the data we are looking for.

This article extends the initial work presented in the CACIC conference [4] with the addition of a deep neural network for object detection, improved generation and display of heat maps over court diagrams, detection of important game events and automatic generation of short video clips for game analysis. The rest of the article is organized as follows. In Sect 2 the objectives of this work are described. Section 3 presents the details of the methods that form the proposed system. The experimental results are exposed in Sect. 4, and finally the conclusions are outlined in Sect. 5.

sinc(*i*) Research Institute for Signals, Systems and Computational Intelligence (sinc.unl.edu.ar)
G. Erbele, J. Bourlot, C. E. Martínez & E. M. Albornoz; "Automatic Extraction of Heat Maps and Goal Instances of a Basketball Game Using Video Processing"
Communications in Computer and Information Science - CACIC 2021, Vol. 1584, pp. 1-11, 2022.

2 Purpose and Objectives

The main objective of this work is to develop a system that allows the extraction of clips with relevant information for the analysis of a basketball team in a match. In addition, the system should work just using a TV footage video of the game.

According to technical managers and coaches of local teams (from Santa Fe province, Argentina), it is very useful to get the distribution of the players throughout the game and, at the same time, analyze the goal situations, not only in terms of the position of the team players but also the game mode of the opponents. In order to accomplish that, the identification of the ball and the players on the field is addressed. Furthermore, if the moment when the ball enters in the basket is detected, it will be possible to extract video clips of the more relevant moments which help for a post analysis.

Once the elements of interest have been identified, the desired outputs can be generated. In particular, the distribution of players throughout a match can be visualized on a heat map, and the points can be translated into a count along with a set of small videos showing the situations, for both teams. To obtain the heat maps, it will also be necessary to study and implement image registration methods to obtain an image of the entire court, in an upper and perpendicular view.

3 The Approach

The system was designed based on a series of successive processing blocks. Figure 1 shows a general scheme of the system.

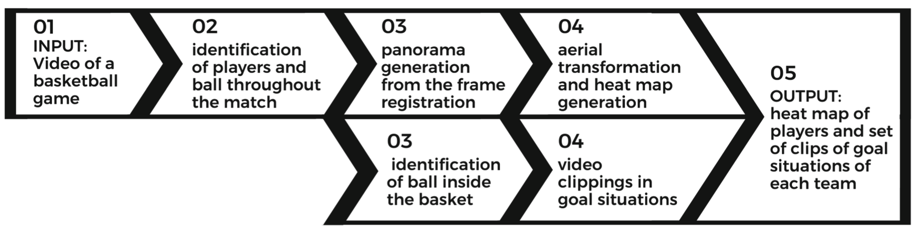

Fig. 1. Diagram of the proposed system.

The input is a TV footage video, captured from a single camera with a resolution of 1280×720 pixels, taken at 30 frames per second. The first step is carrying out the image registration of frames to build a panoramic image of the entire court. When registering, the transformation applied to each frame is calculated and also used to transfer the position of the players from the original frame to the panoramic view on the court at each time. For better visual interpretability, another transformation is performed that convert these results to an aerial view.

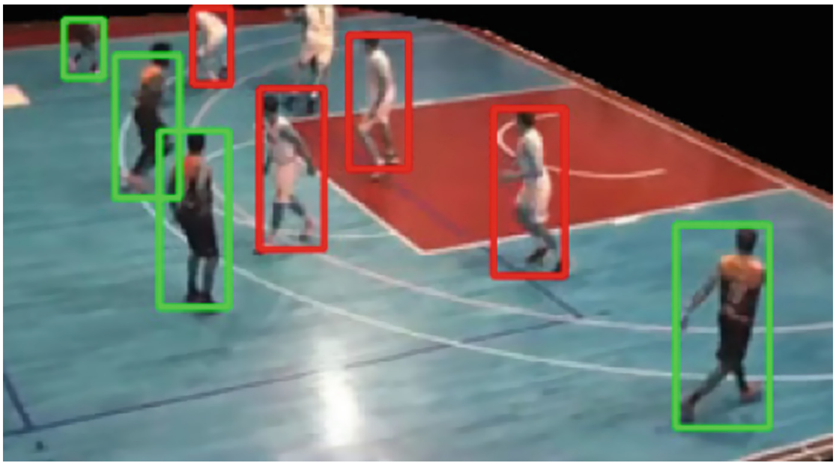

Fig. 2. Result of the identification of players (green and red boxes). (Color figure online)

Identification of Players and Ball Throughout the Match. At this stage, the YOLO object detection algorithm [2] is used. It is an algorithm based on deep neural networks trained to segment a wide variety of objects, including animals, people, among others. This network is one of the most used since through a single inference on the image, it can perform a very precise segmentation that makes it possible to implement both offline and in real time. This algorithm was trained with the Tensorflow library in Python language [1].

The publicly available YOLO network can not identify basket balls and points in the game, so these two desired outputs were trained by means of a transfer learning approach. The transfer learning takes a pre-trained neural network, leaving most of its parameters fixed and trains only a few layers for the new objects to be recognized. Compared to training a network from scratch, this process requires a substantially shorter training time, obtaining good results in terms of prediction accuracy and allowing a model to be trained with a relatively limited number of images.

The re-training was applied with a Pytorch implementation of the algorithm [11] using 942 images taken from different matches of basketball of the National Collegiate Athletic Association (NCAA). These images were labeled using the LabelImg software [18] and the results are two files: the original frame and a corresponding text file with the relative coordinates and size of the bounding boxes detected with the players and/or the point (ball inside basket situation).

The court is segmented to separate the players from the public, thus obtaining the players of any of the teams, and/or the referees. Finally, a post-processing is carried out that allows to identify which team each player belongs to, and also to discard the referees. This segmentation is carried out using different color

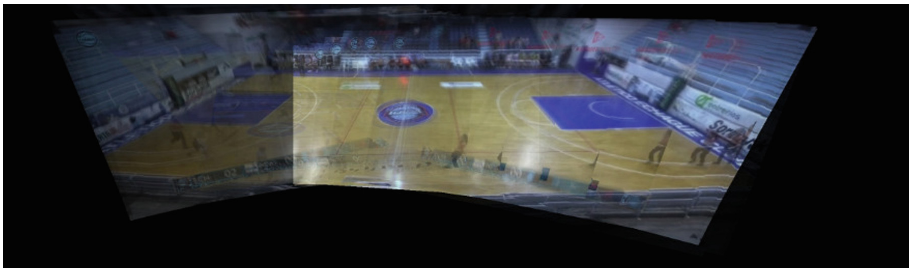

Fig. 3. Panoramic view generated by the image registration of the frames.

models [14]. An example of the segmentation can be seen in Fig. 2. Here, the identified players of each team are marked with boxes of different colors.

Panoramic Image Generation from Frame Registration. To generate the view of the court over which the results will be displayed, a panoramic view is first obtained from the input video [6]. To this end, an image registration algorithm was used with a step of identification of descriptors and key points detection by means of the ORB detector [9], and then a correlation search between those points using the BF Matcher [9]. After obtaining the relevant points between two successive frames, it is possible to calculate the tra3nsformation that transforms the frame from the original view to the panoramic view using a RANSAC linear estimation model [8]. All the methods were implemented in the OpenCV library [13]. An example of the final result of the registration of all the frames of the video can be seen in Fig. 3. The information calculated at this point will be relevant later to find the position of the players in this view.

Identification of Ball, Ball Inside the Basket and Video Clipping. After the transfer learning, the YOLO network is able to detect the ball through all the match, and identify when any team scores during the game, as in Fig. 4. In case of a'made basket' detection, the systems produces a short video-clip with the last 7 s and save this clip in a different registry. The algorithm detect all the frames in which the ball hit the basketball hoop, in case of missing the shot the frames detected are 2–5 approximately and in case of 'made basket' the frames count is higher than 5. In the first frame of the potential point, the algorithm takes the clip for the last 7 s and with the next frames detect if there was a score or not. An example of a moment of this situation is shown in Fig. 5.

Fig. 4. Identification of the ball.

Fig. 5. Identification of 'made basket' situation.

Aerial Transformation and Heat Map Formation. It is intended to show the heat map in an aerial view (upper and perpendicular to the court), since in this way it is possible to see the strong borders of each player, the main locations of the players, the general movement of a team, etc. To accomplish this, a transformation is made to the panoramic view of the court by placing the extreme points of the court in the panoramic view, and the destination points that respect the standard proportions of a basketball court. A perspective transformation matrix is obtained for this process [21]. The result is an aerial image of the court on which the heat maps of the match will be displayed, as can be seen in Fig. 6.

sinc(*i*) Research Institute for Signals, Systems and Computational Intelligence (sinc.unl.edu.ar)
G. Erbele, J. Bourlot, C. E. Martínez & E. M. Albornoz; "Automatic Extraction of Heat Maps and Goal Instances of a Basketball Game Using Video Processing"
Communications in Computer and Information Science - CACIC 2021, Vol. 1584, pp. 1-11, 2022.

Fig. 6. Aerial view of the court.

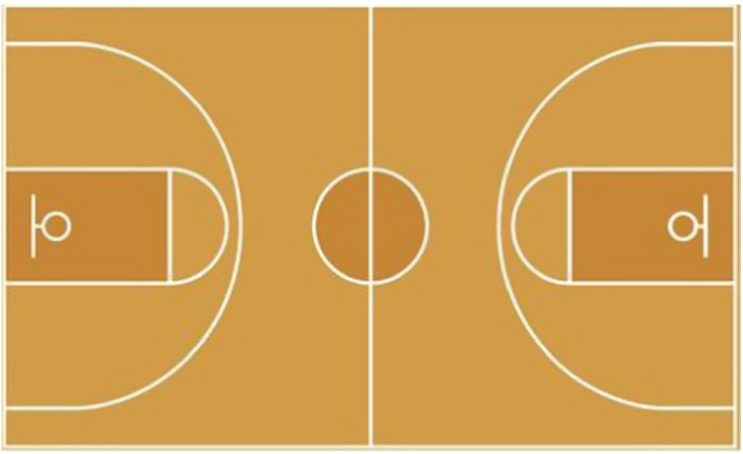

Fig. 7. Schematic representation of a basketball court.

The transformation is carried out on a panoramic image built from the average of a set of partial reconstructions obtained by the registration of different frames, so when performing the transformation to the upper view it becomes noticeable that the players were not completely eliminated from the court. While a good approximation to the top view of a basketball court is obtained, it still presents some small errors that could be improved with some non-linear noise reduction techniques (players who still appear partially).

In order to improve the detection of the edges of the court, it is an option to use its schematic representation, such as the one shown in Fig. 7. Here, a clearer visualization of the relevant information -the heat maps of the match- could be obtained, being more useful for a coach analyzing the game.

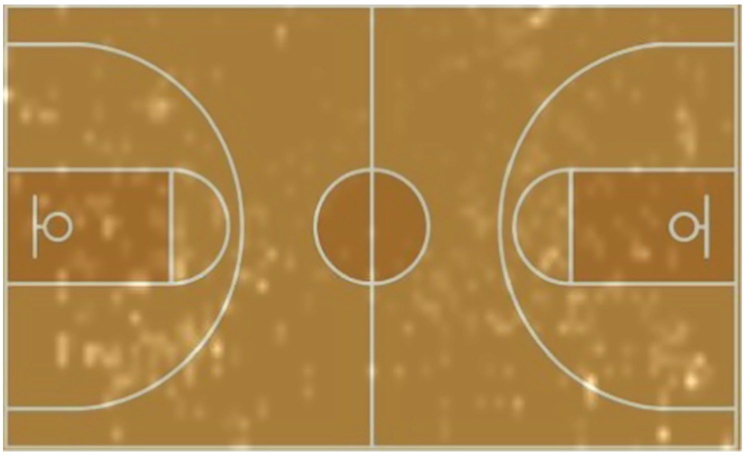

Fig. 8. Heat map for team 1 over the basketball court.

Once the coordinates that represent the location of players on the court have been obtained, the two transformations are carried out successively. The court is divided into non-overlapping square cells of 50 cm on each side [7]. With this discretization, a matrix is defined for each team, whose values will contain the number of players detected within the corresponding cells. Subsequently, with this information, the heat maps of each of the equipment are calculated using the seaborn library [20], used for the visualization of statistical data. It is open source, based on matplotlib [10] and implemented in Python, and has several models to build heat maps with different variants.

4 Experimental Results

The proposed approach was tested with a video of a basketball match recorded in 2021, between Atlético Echagüe Club and Colon S.F. Club, in the Torneo Nacional de Ascenso (Argentina).

The output images obtained after the analysis are showed in Figs. 8 and 9. These present the heat maps of each team over all the match, where each bright spot indicates presence of the team in the location and the brightest points those locations where the team has had more presence.

As can be seen, the heat map results were projected over the top view of the basketball court.

This first approach was designed and implemented to work with a typical television broad-cast video, captured with a single camera located in the upper-center of a lateral view and which only performs horizontal rotations. Respect to teams identification, it is performed using color segmentation of T-shirts and this process is parameterizable, allowing the adaptation to any teams. In addition, the basketball court detection is parameterizable, allowing its applicability to other courts.

sinc(i) Research Institute for Signals, Systems and Computational Intelligence (sinc.unl.edu.ar)
G. Erbele, J. Bourlot, C. E. Martínez & E. M. Albornoz; "Automatic Extraction of Heat Maps and Goal Instances of a Basketball Game Using Video Processing"
Communications in Computer and Information Science - CACIC 2021, Vol. 1584, pp. 1-11, 2022.

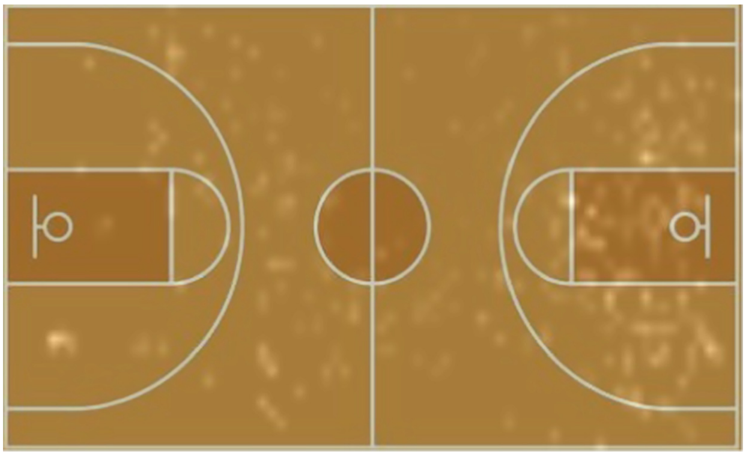

Fig. 9. Heat map for team 2 over the basketball court.

The other output provided by the system is the collection of video clips with 'goal' or 'no-goal' situations, the filenames specify each case. However, these have not information about the specific team that took the shot, then in some cases there are shots (clips) overlapped in 5 or more seconds.

Regarding the execution times, the detection of ball and basket is carried out in real time for 30FPS videos, given that the YOLO inference is very fast. The bottleneck of the performance is the creation of the heat maps, because the registration of frames only can run in real time for low resolution videos (smaller than 480p).

On the other hand, the system does not address the monitoring of individual players along the court, nor the detection failures due to occlusions. However, for the general statistics that was intended in this work, there is no significant difference.

In the registration step, a problem arises for fast game actions that took short time to go from one basket to the other. Here, the middle part of the court is poorly transformed because there is no enough frames. Nevertheless, for the general construction of the heat map, these instants could be discarded with no substantial drop in performance.

5 Conclusions

In this work, a method to obtain heat maps of the teams, and clips of goal situations, in a basketball match was presented. For the experiments, a typical television broad-cast video, captured with a single camera, was used.

Results are projected over a complete basketball court obtained through image transformations. Statistics and heat maps of players location during the match are presented as image augmentation. In addition, a set of video clips is generated resuming important moments of the match (such as the goal situation).

Future works include the improvement of the generation of the upper view of the court. In this sense, the first improvements could address to bug fixing, such as to enhance the court edges detection and to reduce noise that arises due to player shadows. Furthermore, in order to upgrade the statistics and the heat map, the ball possession and the player's position when taking a shot will be detected and analysed.

Acknowledgement. The authors would like to thank the **sınc**(i) institute and the UNL (with CAI+D 50620190100145LI), to Enzo Ferrante and Eric Priemer by the collaboration in the original conference work, and Ryan Werth for providing the corpus of images to train the network.

References

1. Abadi, M., et al.: {TensorFlow}: a system for {Large-Scale} machine learning. In: 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), pp. 265–283 (2016)
2. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: Yolov4: optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934 (2020)
3. Botwicz, M., Klembowski, W., Kulpa, K., Samczyński, P., Misiurewicz, J., Gromek, D.: The concept of an RF system for the detection and tracking of the ball and players in ball sports: an extended abstract for poster presentation. In: 2017 Signal Processing Symposium (SPSympo), pp. 1–3. IEEE (2017)
4. Bourlot, J., Eberle, G., Priemer, E., Ferrante, E., Martínez, C., Albornoz, E.: Generación de mapas de calor de un partido de básquetbol a partir del procesamiento de video. In: Mac Gaul, M. (ed.) XXVII Congreso Argentino de Ciencias de la Computación, CACIC, pp. 231–239 (2021)
5. Chen, B., Huang, Z., Yu, W., Xu, Y., Peng, J.: Object recognition and localization based on kinect camera in complex environment. In: 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 668–673. IEEE (2013)
6. Chen, K., Wang, M.: Image stitching algorithm research based on opencv. In: Proceedings of the 33rd Chinese Control Conference, pp. 7292–7297. IEEE (2014)
7. Cheshire, E., Halasz, C., Perin, J.K.: Player tracking and analysis of basketball plays. In: European Conference of Computer Vision (2013)
8. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM **24**(6), 381–395 (1981)
9. Howse, J., Minichino, J.: Learning OpenCV 4 Computer Vision with Python 3: get to Grips with Tools, Techniques, and Algorithms for Computer Vision and Machine Learning. Packt Publishing Ltd, Birmingham (2020)
10. Hunter, J.D.: Matplotlib: a 2d graphics environment. Comput. Sci. Eng. **9**(03), 90–95 (2007)
11. Jocher, G., Stoken, A., Borovec, J., Chaurasia, A., Changyu, L.: ultralytics/yolov5. Github Repository, YOLOv5 (2020)
12. Kim, J., et al.: Wearable salivary uric acid mouth guard biosensor with integrated wireless electronics. Biosens. Bioelectron. **74**, 1061–1068 (2015)
13. Laganière, R.: OpenCV 3 Computer Vision Application Programming Cookbook. Packt Publishing Ltd, Birmingham (2017)

14. Lu, W.L., Ting, J.A., Little, J.J., Murphy, K.P.: Learning to track and identify players from broadcast sports videos. IEEE Trans. Pattern Anal. Mach. Intell. **35**(7), 1704–1716 (2013)
15. Salski, B., Cuper, J., Kopyt, P., Samczynski, P.: Radar cross-section of sport balls in 0.8-40-GHZ range. IEEE Sens. J. **18**(18), 7467–7475 (2018). https://doi.org/10.1109/JSEN.2018.2862142
16. Sekine, Y., et al.: A fluorometric skin-interfaced microfluidic device and smartphone imaging module for in situ quantitative analysis of sweat chemistry. Lab on a Chip **18**(15), 2178–2186 (2018)
17. Stetter, B.J., Ringhof, S., Krafft, F.C., Sell, S., Stein, T.: Estimation of knee joint forces in sport movements using wearable sensors and machine learning. Sensors **19**(17), 3690 (2019)
18. Tzutalin, D.: Labelimg. GitHub Repository 6 (2015)
19. Wang, Y., Zhao, Y., Chan, R.H.M., Li, W.J.: Volleyball skill assessment using a single wearable micro inertial measurement unit at wrist. IEEE Access **6**, 13758–13765 (2018). https://doi.org/10.1109/ACCESS.2018.2792220
20. Waskom, M.L.: Seaborn: statistical data visualization. J. Open Source Softw. **6**(60), 3021 (2021)
21. Wen, P.C., Cheng, W.C., Wang, Y.S., Chu, H.K., Tang, N.C., Liao, H.Y.M.: Court reconstruction for camera calibration in broadcast basketball videos. IEEE Trans. Visual Comput. Graphics **22**(5), 1517–1526 (2015)
22. Zhang, X., Duan, H., Zhang, M., Zhao, Y., Sha, X., Yu, H.: Wrist mems sensor for movements recognition in ball games. In: 2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), pp. 1663–1667 (2019). https://doi.org/10.1109/CYBER46603.2019.9066506

Author Queries

Chapter 7

Query Refs.	Details Required	Author's response
AQ1	Per Springer style, both city and country name must be present in the affiliations. Accordingly, we have inserted the country name "USA" in affiliation. Please check and confirm if the inserted country name is correct. If not, please provide us with the correct country name.	