ENHANCING SPORTS PERFORMANCE ANALYSIS: AN AI APPROACH FOR BASKETBALL AND VOLLEYBALL

Ana PETRANIS[1*], Mihail VINOGRADSCHII[2]

[1]*Software Engineering and Automatics, Technical University of Moldova, Group IS-231M, Faculty of Computers, Informatics and Microelectronics, Chișinău, Republic of Moldova*
[2]*ML Engineer, Chișinău, Republic of Moldova*

*Corresponding author: Ana Petranis, ana.petranis@isa.utm.md

Tutor/coordinator: **Corina BEȘLIU**, university lecturer, *Technical University of Moldova*

Abstract. *Artificial intelligence is increasingly being used in all areas of human activity, from the familiar text editor to the cutting-edge satellite that has just entered Earth's orbit. Modern team sports also need to implement AI to analyze the results of matches, in order to identify the strengths and weaknesses of each player, as well as to develop a right strategy for subsequent games against a specific opponent. Our study investigates the process of data collection and processing for sports analytics using basketball and volleyball games as examples. Data for analysis was sourced from the official FIBA YouTube channel "FIBA - The Basketball Channel" and the Baller TV replay library, which archives matches from various youth sports. The obtained videos were segmented into individual frames. A portion of these frames were manually labeled to create training and validation datasets. The remaining frames formed the unlabeled test dataset, crucial for evaluating the accuracy of the YOLOv8 model chosen as the foundation for this study. Our focus was on identifying players through jersey number recognition, detecting the ball and its location on the court, classifying game situations, and processing the score and timer using OCR technology. The fine-tuned YOLOv8 achieved an accuracy of 93% based on the mAP50-95 metric, which evaluates the overlap between predicted and actual bounding boxes.*

Keywords: *computer vision, machine learning, object detection, YOLOv8*

Introduction

The AI industry is rapidly and confidently entering every domain of human activity. Nowadays, we have drones that can detect plant illnesses and determine when a crop is ready for harvest, along with autonomous greenhouses. Millions of people are benefiting from generative AI in their daily job and education. Additionally, modern cars are equipped with autopilots that rely on AI technology.

The challenge in sports performance analysis lies in extracting insights from vast data volumes together with diverse factors affecting athletic outcomes. Our goal is to use AI-driven technology to transform player evaluation and enhancement in basketball and volleyball. Knowing the experience of the prior studies, we embrace complex sporting conditions, utilizing advanced AI models like You Only Look Once (YOLOv8) for precise real-time analysis. Our empirical data supports these claims, showcasing improved performance analysis. Our approach modifies AI models for various sports contexts, guaranteeing accurate analysis in a range of specific circumstances. By using sophisticated algorithms and careful data annotations, we create a comprehensive framework that may be easily implemented in any business related to the sport analytics field.

Data Gathering

Any data-driven analysis starts with the collection of data. This data usually comes from match recordings that could be found on resources like *Baller TV* and *FIBA - The Basketball Channel*. It is essential to collect data with a variety of features in order to guarantee reliable and broadly applicable AI models. This covers changes in scene lighting, video quality, camera angles, and stabilisation.

An effective performance analysis depends on the AI model being able to learn a richer collection of features, which is made possible by diverse data. For example, if a model is trained just on static shots, it may miss information relating to player positioning and movement relative to the court. For this reason, data should contain video from a range of camera perspectives, including baseline cameras, standard broadcast views (Fig. 1), and even fisheye lenses (Fig. 2), which provide a broader field of view.

Figure 1. Standard broadcast camera view (Basketball)

Figure 2. Camera with fisheye lens (Volleyball)

Recordings from different locations with different lighting configurations and court layouts should be taken into consideration to further enhance the data collection. Video from both indoor and outdoor courts may be included in it to show the impact of various playing conditions. Additionally, a range of video quality levels, from high-definition broadcasts to lower-resolution recordings, should be gathered into the dataset. Finally, in terms of lighting, the data should contain footage shot in a variety of lighting scenarios, ranging from intensely lit training facilities to brightly lit professional venues.

Once the diverse data is collected, the video recordings are segmented into individual frames. Then, particular interesting game moments are found and labeled for additional examination. These labeled game moments serve as the foundation for extracting relevant performance statistics, such as the number of successful passes leading to a shot, shooting accuracy, player speed, and point tallies from scoring sequences.

Data Labeling

Data labeling is a critical and time-consuming step in AI research, especially for tasks involving object recognition and analysis. Data labeling in the context of sports performance analysis is the process of carefully locating and annotating objects of interest within video frames. LabelMe, a popular annotation tool (Fig. 3), can be used to create bounding boxes around specific objects in each frame [1]. In fact, these bounding boxes specify the object's size and location within the image. Examples of objects commonly labeled in basketball and volleyball data include the ball, players identified by jersey numbers or body features, jerseys themselves, and the score bug.

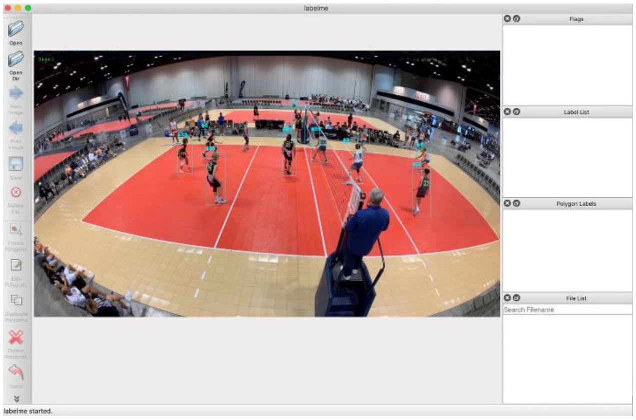

Figure 3. LabelMe interface

An additional technique employed for data labeling is the use of homography masks (Fig. 4). In essence, these masks are specialized overlays that precisely delineate the region of the playing court in the video frame. The AI model needs this data in order to distinguish between activities that take place on and off the court [2].

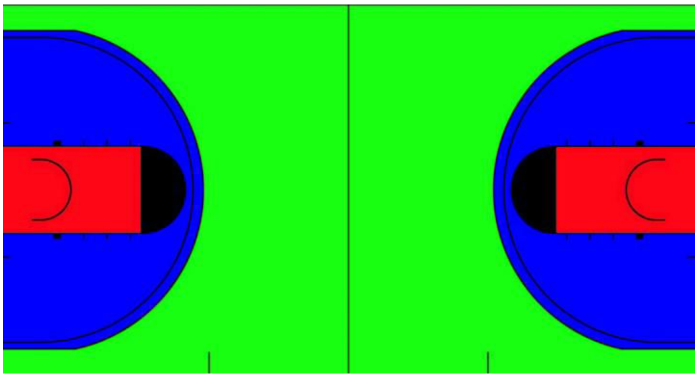

Figure 4. Basketball court homography mask

Through the combined use of bounding boxes and homography masks, three distinct datasets are typically created: training, validation and test. Training dataset consists of labeled video frames used to train the initial AI model. The labels define the location and type of each object within the frame. A different labeled dataset called the validation dataset is used to assess how well the model performed during training. This allows for adjustments to the model and prevents overfitting on the training data. After the training phase is finished, an unlabeled test dataset is used to evaluate the accuracy of the model. This dataset provides a realistic measure of how well the model performs on unseen data [3].

Model Training

Any AI model's ability to function depends on the hardware configuration that is chosen with care. We selected an efficient configuration with an Intel Core i7-13700K processor and a powerful NVIDIA GeForce RTX 4070 Ti graphics card for our experimental AI analyzer. On such a combination of hardware components, any model will perform fast and efficiently.

You Only Look Once (YOLOv8) [4] is a cutting-edge deep learning model designed especially for real-time object detection. Instead of using multiple neural networks for detection, YOLOv8 uses a single neural network to analyze the entire image at once. The image is divided into a grid of regions by this network, which then forecasts bounding boxes and probabilities for each region, indicating the kind and quantity of objects that are present there. YOLO's simplified methodology results in remarkable processing speeds, which makes it perfect for a range of real-time applications.

The model training procedure is led by a carefully written Python script. Importing the pre-defined YOLOv8 model architecture is the first step in the script's setup. Next, crucial training parameters should be set and the version of YOLOv8 being used should be specified. Data is one of these parameters. The number of epochs (the number of times the entire training dataset will be passed through the model for learning), the batch size (the number of samples used in one epoch to train a neural network), the number of workers in pool (the number of parallel worker processes to leverage for data loading and preprocessing, accelerating the training process), and the augmentation (enhances the size and quality of machine learning training datasets) are all important details about the training data.

Figure 5 shows a training batch example. This batch consists of video frames where the ball has been manually labeled with an ID of 0 representing its object class. Following 500 training epochs, the YOLOv8 model learns to predict the ball's location within each frame.

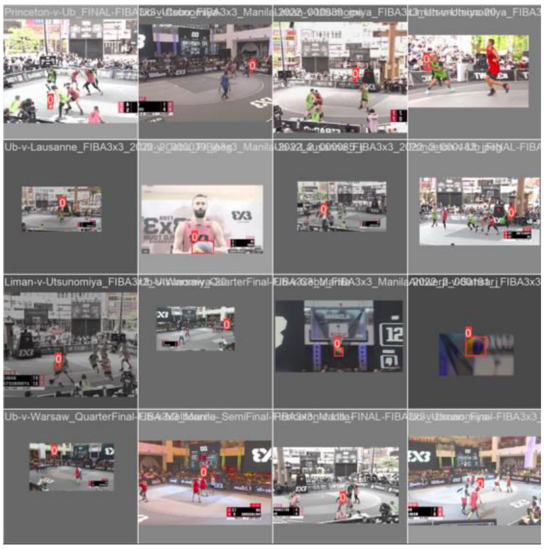

Figure 5. Training batch

Figure 6 shows the graphical representation of this prediction. We can easily evaluate the model's accuracy in locating the ball's bounding box thanks to the visual format.

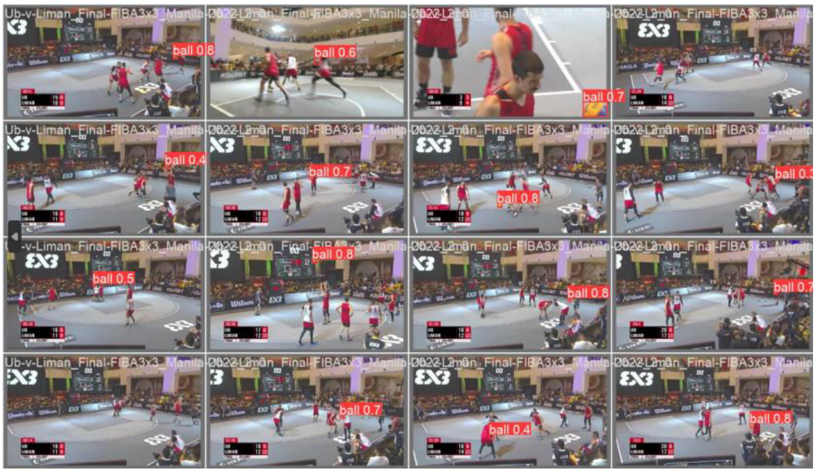

Figure 6. YOLOv8 predictions

Additionally, AI models can be trained to detect and locate score bugs within a video frame. Convolutional Neural Networks (CNNs) are highly effective in recognizing particular patterns in images. A CNN, in our case YOLOv8, trained on labeled video frames can identify distinctive visual characteristics of score bugs. It is able to identify the areas that contain these on-screen graphics by examining the pixel data of every frame. A well-labeled dataset is essential for training any object detection model, including score bug recognition. The location of various data points, such as the clock and the score, must be precisely labeled on every frame that contains a score bug during the training process. By labeling the various elements within the score bug, the model is able to distinguish between them more precisely, leading to the accurate extraction of the desired information.

Optical Character Recognition (OCR) techniques are used once the CNN has successfully located the score bug. Converting visual text into a format that is readable by machines is the speciality of OCR technology. When used in this way, OCR enables us to retrieve important game information from the score bug, including team names, scores, and remaining game time (Fig. 7) [4-5].

Figure 7. Score bug

Results

Our research produced a unified method for re-identifying players throughout different video sequences, detecting jerseys/players, and classifying teams. This strategy takes advantage of team sports' consistency. While the number of players on the court differs (ten in basketball, twelve in volleyball), identifying players relies on unique visual cues like their jersey numbers and team colors.

However, sports do have specific visual elements that require model adjustments. For example, the way the ball looks in different sports like basketball, volleyball and football can differ dramatically. In such cases, the YOLOv8 model requires re-training on brand-new labeled video frames containing the specific ball type. Similar to this, because courts vary in size, zones, markings, and fixed objects (baskets, nets, gates, etc.), court homography masks should be modified for each sport.

Training the model with a large enough dataset can achieve an overall prediction power of approximately 91.8%, even with these sport-specific variations. Out of all the components, the ball detection had the highest accuracy. Based on the mAP50-95 metric, which takes into account the overlap between the ball's predicted and actual bounding boxes, the refined YOLOv8 model achieved an accuracy of 93%. The single ball on the court eliminates the need for difficult re-identification tasks like those needed for players and becomes the reason for high accuracy results.

Conclusions

In conclusion, our study has effectively examined the problem of improving sports performance analysis by utilizing AI techniques. We have offered a solid answer to the challenges involved in evaluating volleyball and basketball performance through rigorous validation, sophisticated AI algorithms, and careful data collection. Our results validate the success of our methodology in obtaining significant insights from various match settings, consequently providing coaches, experts, and players with invaluable resources to enhance their on-court performance.

References

[1] Wang J., Quan H., Li L., *Automated Return Type Annotation for Python*. SSRN Electronic Journal, 2022. doi: 10.2139/ssrn.4088424

[2] Jiang K., Li J., Liu Z., Dong C., *Court Detection Using Masked Perspective Fields Network*. Proceedings of IEEE Pacific Rim International Symposium on Dependable Computing, PRDC, 2023. doi: 10.1109/PRDC59308.2023.00055

[3] Gong Y.Liu G.Xue Y.Li R.Meng L., *A survey on dataset quality in machine learning*. Information and Software Technology, 2023. doi: 10.1016/j.infsof.2023.107268

[4] Talib M., Al-Noori A. H., Suad J., *YOLOv8-CAB: Improved YOLOv8 for Real-time Object Detection*. Karbala International Journal of Modern Science, 2024. doi: 10.33640/2405-609X.3339

[5] Guntuboina C., Porwal A., Jain P., Shingrakhia H., *Deep Learning Based Automated Sports Video Summarization using YOLO*. Electronic Letters on Computer Vision and Image Analysis, 2021. doi: 10.5565/REV/ELCVIA.1286