Improving Object Detection Quality in Football Through Super-Resolution Techniques

Karolina Seweryn[1[0000−0003−0617−7301]], Gabriel Chęć[2], Szymon Łukasik[1,2[0000−0001−6716−610X]], and Anna Wróblewska[3[0000−0002−3407−7570]]

[1] NASK - National Research Institute, Warsaw, Poland
[2] AGH University of Krakow, Krakow, Poland
[3] Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland
{karolina.seweryn}@nask.pl

Abstract. This research examines the effectiveness of super-resolution techniques in improving object detection accuracy in football. Given the sport's fast pace and the need for precise tracking of players and the ball, super-resolution can offer significant improvements. The study applies super-resolution techniques to SoccerNet football videos and evaluates their impact on Faster R-CNN detection accuracy. Findings reveal a significant boost in object detection accuracy following the application of super-resolution preprocessing. Enhancing object detection by integrating super-resolution techniques provides substantial advantages, particularly in low-resolution settings, with a 12% rise in mean Average Precision (mAP) at an IoU (Intersection over Union) range of 0.50:0.95 for 320x240 pixel images when the resolution is quadrupled using RLFN. As the image dimensions grow, the extent of improvement becomes less pronounced; however, a consistent enhancement in detection quality remains clear. Moreover, the implications of these results for real-time sports analytics, player tracking, and the overall viewing experience are discussed.

Keywords: object detection · super-resolution · low-quality videos

1 Introduction

The tracking of players and the ball in football matches is a critical aspect of sports analytics, significantly impacting the tactical elements of the game. Player tracking data provides valuable insights into individual and team performance, enabling coaches to analyze movements, formations, and strategies more effectively [5, 17]. For instance, tracking data can be used to assess player fitness, monitor fatigue, and reduce the risk of injury by understanding players' physical demands during a match. Furthermore, ball-tracking technology plays a key role in enhancing the spectator experience, contributing to more accurate decisions through technologies like goal-line technology and video assistant referee (VAR) [32].

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

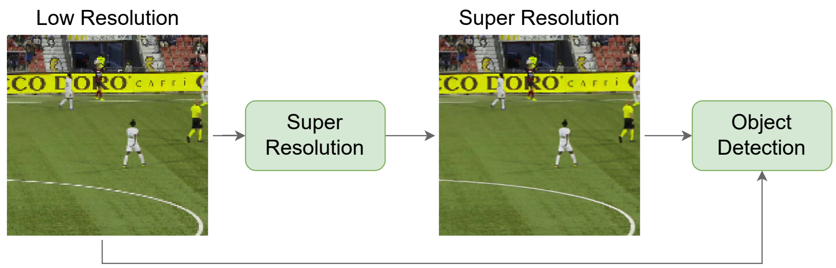

Fig. 1. An image is enhanced by a super-resolution model and fed to an object detector.

Despite the wealth of research and annual publications on tracking football players and the ball [17–19, 21, 22], the precise tracking of small, fast-moving objects like a football ball remains a challenge. These advanced tracking solutions frequently depend on the availability of high-definition video input to ensure appropriate results. However, a significant portion of football clubs, ranging from grassroots clubs to professional organizations, struggle with the provision of such high-quality data. This situation is not solely a consequence of the prohibitive costs associated with state-of-the-art recording equipment. Still, it is also exacerbated by the prevalent use of drones and portable devices that, while offering versatility and convenience, often compromise on video quality. Additionally, the widespread practice of compressing video content for distribution on platforms like YouTube or other streaming services further degrades the fidelity of footage available for analytical purposes, presenting substantial hurdles to the deployment and effectiveness of sophisticated tracking models in real-world settings. This disparity in technological access not only affects the competitiveness of smaller clubs but also limits the overall development of sports analytics by constraining data diversity and research opportunities.

While high-quality video costs may decrease in the future, alternative solutions remain valuable. This research explores the potential of super-resolution techniques to enhance object detection in football, reducing dependence on high-resolution input data. By integrating super-resolution with modern detection models, teams can lower costs, minimize data storage requirements, and eliminate the need for expensive equipment. These methods not only enhance detection performance on low-resolution footage but also improve accuracy even when high-resolution images are available, making super-resolution a valuable preprocessing step.

2 Related Work

Initial approaches to extracting tracking data relied on views from multiple cameras [6, 7, 16, 18] for precise player localization. While effective, these approaches came with significant drawbacks, including high installation and maintenance

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

costs, as well as complex technical challenges such as camera calibration and object re-identification [7]. These limitations made widespread adoption difficult, particularly for teams with limited resources.

More and more latest models find the coordinates of players and the ball based on just one camera [8, 9]. Still, locating small objects like the ball remains a challenge due to occlusion, false detection, lighting, and frame rate variations [10]. Also, interactions between players, and players and the ball can cause complex problems. Additionally, solutions for detecting players and a ball often rely on high-resolution input videos or zoom-in views and their trajectory imputations, i.e. inferring their locations based on adjacent frames in which their detections were feasible [11]. Tracking in a football match can be classified as a multi-object tracking task, and it remains challenging due to factors like abrupt object appearance differences and even severe object shadings and obscurations [23]. The latest solutions enhance neural architectures to capture small objects, e.g. adding focal loss to YOLOv7 [24], or using neural radiance fields [25].

A similar issue of tracking football players in low-quality videos has been addressed in the works [22] and [20]. In both cases, the researchers focus on adapting systems to challenging visual conditions typical of football broadcasts. The authors of [22] adapt advanced multi-object tracking systems for use with low-quality videos. At the same time, the other research concentrates on detecting and tracking small, less visible players without the need for manual data annotation. All those approaches highlight the increasing capability of sports analytics to handle visual challenges in football.

In recent years, super-resolution approaches have been significantly improved [12,13]. Currently, a bunch of research copes with applying and adapting super-resolution techniques to many detection tasks, such as satellite imagery [26], underwater object detection [28], and occluded small commodities [27]. In [34], the authors explored the effects of using super-resolution as a preprocessing step in object classification. They demonstrated that this approach significantly improves the detection of small objects. However, the analysis of performance in this work is somewhat limited, as it focuses solely on the number of true positives and the mean probability returned by the model. In sports analytics, only one study has explored super-resolution for ball detection, focusing on tennis [29]. Instead of integrating super-resolution into the model, it was used as a preprocessing step to enhance frame resolution before labeling, improving annotation accuracy. This approach markedly diverges from the use of super-resolution techniques within model architectures, as it was applied to refine the input data before annotation.

3 Experiments

Data Our study uses the SoccerNet dataset [3], consisting of 12 full football matches recorded from the main camera in 1080p (1920×1080). It includes 100 clips, each 30 seconds long, spanning various games and seasons to ensure a

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

diverse evaluation of object detection methods. The dataset is split into 42,750 training frames and 36,750 test frames, with no match overlap. We followed this division, using the training set for model training and the test set for evaluation. The dataset classifies on-field elements into 8 categories (player team left, player team right, goalkeeper team left, goalkeeper team right, main referee, side referee, staff, and ball) for object detection. For simplicity, our experiments grouped them into two: ball and person.

For super-resolution model training, we used the SoccerNet train set and UHDSR8K [14], which includes 2,100 images in 4K resolution. This dataset's high resolution offers a different range of challenges and opportunities for super-resolution model training. The UHDSR8K dataset is part of a study that aimed to benchmark single image super-resolution (SISR) methods on Ultra-High-Definition (UHD) images, including both 4K and 8K resolutions. It was used to evaluate the performance of SISR methods under various settings, contributing to the development of baseline models for super-resolution. In our experiments, we specifically utilized the 4K images from this dataset.

Super-resolution Choosing the super-resolution network architecture for image reconstruction affects its quality, speed, and efficiency. The architecture selection in this study was based on the NTIRE 2022 Challenge on Efficient super-resolution results [1]. The challenge emphasized designing networks capable of efficiently converting single images to higher resolutions, with a focus on metrics like runtime, parameter count, floating-point operations, and memory usage while maintaining or exceeding a Peak Signal-to-Noise Ratio (PSNR) [30] of 29 on the DIV2K dataset, where PSNR measures the quality of image reconstruction. The winning architecture, Residual Local Feature Network (RLFN) [4] was chosen for this study due to its compact size, fast learning, and effective high-resolution reconstruction capabilities. RLFN combines convolutional layers with attention mechanisms to extract local features and maintain global pixel relationships, crucial for detailed and accurate high-resolution image reconstruction.

Object detection We used Faster-RCNN [2] with ResNet50 backbone for object detection, chosen for its pre-trained availability in PyTorch, widespread adoption, and robust performance across diverse datasets. Faster R-CNN works by integrating two key components: a Region Proposal Network (RPN) and a Fast R-CNN detector. The RPN first scans the input image with a sliding window approach to generate object proposals, identifying regions that most likely contain an object. These proposals are then passed to the Fast R-CNN detector, which extracts features using a convolutional neural network, classifies the objects within these regions, and refines their bounding boxes for precise localization.

Experimental Setup In our experiments using the SoccerNet dataset, images from successive frames were degraded to lower resolutions by reducing their original dimensions of 1920x1080 by factors of 2, 3, 4, and 6. For UHDSR8K images with an original size of 3840x1920, similar reductions were applied.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

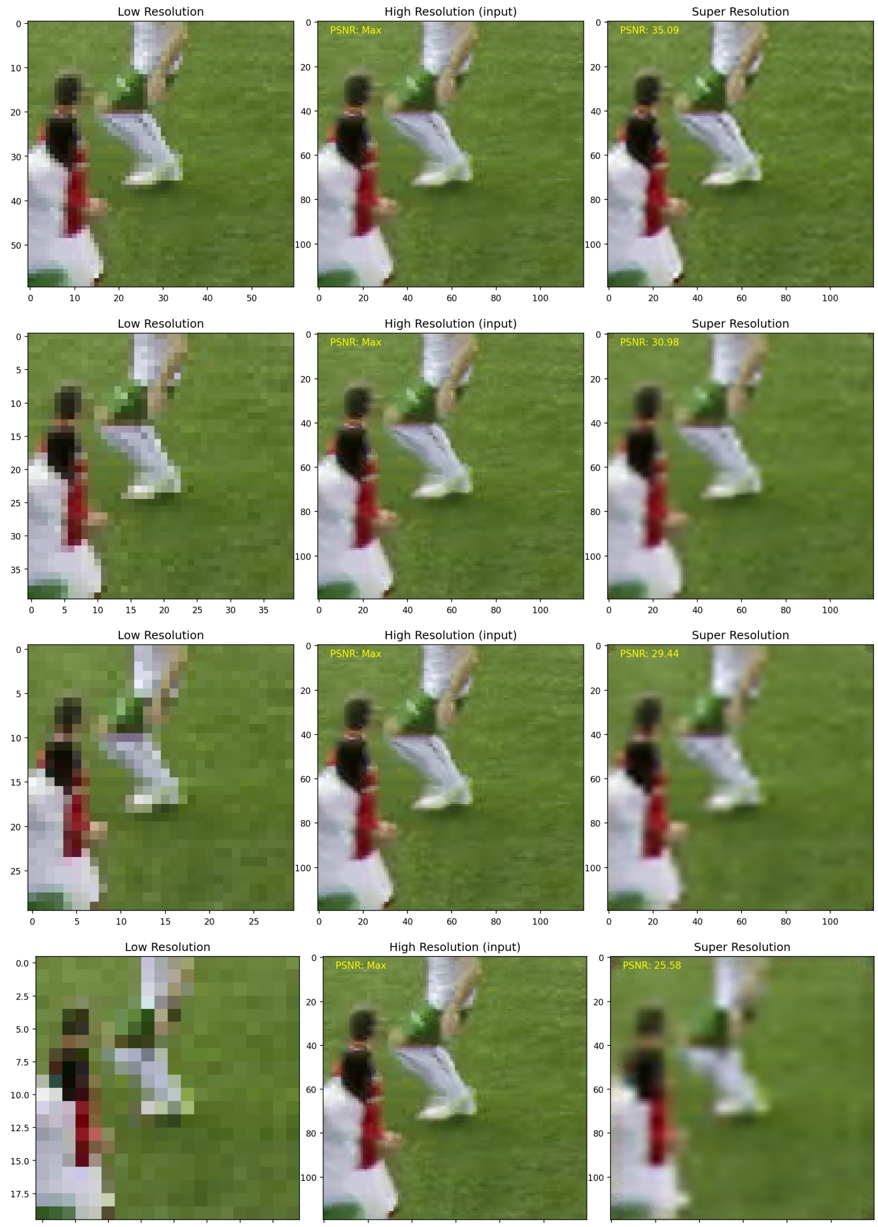

Fig. 2. Sequential visualizations of super-resolution enhancements showcasing progressive upscaling factors. Rows of images illustrate the enhancements achieved through super-resolution RLFN techniques for scales of x2, x3, x4, and x6, respectively. The middle column displays the original image (ground truth image), the left column shows the low-resolution version, and the right column presents the enhanced result after applying super-resolution to the low-resolution image.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

We tested the object detection performance of Faster R-CNN on these low-resolution images and images enhanced by the super-resolution network RLFN, trained on both SoccerNet and UHDSR8K datasets (see Figure 1). Experiments were conducted using NVIDIA A100-40GB GPUs. Training of super-resolution models and object detection was performed with a batch size of 32. The super-resolution model was trained with a learning rate ranging from 1e-5 to 0.01 using the Adam optimizer [33].

4 Results

Figure 2 presents the results of the RLFN model applied to a SoccerNet test sample. The middle column illustrates the original input image, which was resized by factors of 2, 3, 4, and 6 (left), followed by processing through the super-resolution network (right). With a small - double reduction, the network performs well in reconstructing the original image. A PSNR value over 35 indicates a high quality of reconstruction. However, the smaller the dimensions of the low-resolution image, the more challenging it becomes to replicate the input. In the last case - a sixfold reduction - the PSNR drops to 25.58. This example demonstrates that the network is capable of fairly accurately reproducing objects despite having highly blurred pixels as input.

Table 1. Impact of the input image shape on detection performance of Faster RCNN.

Image Shape	**mAP**		**mean IoU**	
	IoU=0.5:0.95	**IoU=0.5**	$\tau_{IoU} = 0.5$	$\tau_{IoU} = 0.9$
(320×240)	24.3	46.7	78.7	92.7
(640×480)	27.6	52.3	79.6	92.8
(720×576)	27.5	51.8	79.6	92.8
(1280×720)	29.5	56.3	79.9	92.9
(1920×1080)	29.5	56.0	79.9	92.9

Table 1 presents the results of Faster R-CNN model tested on SoccerNet with various image shapes. Increasing the input size improves mAP, with a sixfold increase (from 320×240 to 1920×1080) boosting mAP by over 21% at IoU 0.50:0.95 and around 20% at mAP@IoU=0.50. The average IoU remained consistent across various thresholds. These results confirm that higher image resolution enhances detection quality.

After training super-resolution models, we evaluated their performance on SoccerNet using PSNR and MSE metrics (Table 2). The analysis reveals that larger reductions in image size (denoted as xN) before input into the super-resolution model correlate with increased challenges in image reconstruction. Notably, the RLFN model trained on the SoccerNet dataset achieved higher PSNR values on its test set, underscoring the advantage of training on images

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

Table 2. Quality of super-resolution models on SoccerNet and UHDSR8K. 'x[N]' indicates that the image's dimensions were downscaled by a factor of N before processing by the super-resolution network. ↑ - the higher the better, ↓ - the lower the better.

Trained on UHDSR8K			**Trained on SoccerNet**	
SR	**PSNR ↑**	**MSE ↓**	**PSNR ↑**	**MSE ↓**
RLFN x2	33.15	36.47	34.63	26.44
RLFN x3	29.73	79.43	30.54	66.76
RLFN x4	27.83	119.82	29.02	92.82
RLFN x6	25.48	204.47	26.45	166.41

Table 3. Detection performance (mAP) of Faster R-CNN model for various super-resolution scales and training datasets.

Input Shape	**Train Dataset**	**Output Shape**		**mAP**		
				@IoU=0.50:0.95	**@IoU=0.50**	**@IoU=0.75**
320x240	-	-	-	24.3	**46.7**	22.7
	SoccerNet	640x480	x2	26.0	39.8	31.6
		960x720	x3	26.9	40.8	31.9
		1280x960	x4	**27.3**	41.4	**32.2**
		1920x1440	x6	26.6	40.8	31.2
	UHDSR8K	640x480	x2	24.7	40.2	27.3
		960x720	x3	24.9	39.7	28.8
		1280x960	x4	25.3	40.3	29.1
		1920x1440	x6	27.2	41.3	31.9
640x480	-	-	-	27.6	**52.3**	26.4
	SoccerNet	1280x960	x2	27.8	40.1	30.6
		1920x1440	x3	28.1	40.3	31.1
		2560x1920	x4	**28.5**	40.5	31.6
		3840x2880	x6	28.3	40.4	**32.1**
	UHDSR8K	1280x960	x2	28.3	40.9	31.7
		1920x1440	x3	28.0	40.1	31.6
		2560x1920	x4	**28.5**	40.3	**32.2**
		3840x2880	x6	28.4	40.0	**32.2**
1280x720	-	-	-	29.5	**56.3**	27.5
	SoccerNet	2560x1440	x2	30.1	42.0	34.3
		3840x2160	x3	30.1	41.8	33.5
		5120x2880	x4	30.3	42.4	34.0
		7680x4320	x6	29.8	41.8	32.8
	UHDSR8K	2560x1440	x2	30.4	42.8	34.3
		3840x2160	x3	30.3	42.8	34.3
		5120x2880	x4	**30.6**	42.6	**34.5**
		7680x4320	x6	29.6	41.8	32.2

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

akin to the target dataset for optimal results. A PSNR over 30 is regarded as good quality. Thus, the RLFN models with x2 and x3 reductions on SoccerNet, and the x2 reduction on UHDSR8K, met the criterion of good quality.

Table 4. Mean IoU results of Faster R-CNN model for various super-resolution scales and training datasets, evaluated at $\tau_{IoU} = 0.5$ and $\tau_{IoU} = 0.9$. τ_{IoU} denotes the IoU threshold that the model uses to determine if an object is predicted correctly.

Input Shape	**Train Dataset**	**Output Shape**	**Upscale Factor**	**mean IoU** $\tau_{IoU} = 0.5$	**mean IoU** $\tau_{IoU} = 0.9$
320x240	-	-	-	78.7	92.7
	SoccerNet	640x480	x2	81.1	91.9
		960x720	x3	81.6	92.6
		1280x960	x4	81.8	92.7
		1920x1440	x6	81.1	**93.1**
	UHDSR8K	640x480	x2	80.2	92.2
		960x720	x3	80.4	91.7
		1280x960	x4	80.8	91.8
		1920x1440	x6	**82.1**	92.5
640x480	-	-	-	79.6	92.8
	SoccerNet	1280x960	x2	83.7	92.9
		1920x1440	x3	83.9	92.7
		2560x1920	x4	84.0	**93.0**
		3840x2880	x6	83.8	92.8
	UHDSR8K	1280x960	x2	83.7	92.9
		1920x1440	x3	84.1	92.8
		2560x1920	x4	84.2	**93.0**
		3840x2880	x6	**84.3**	92.8
1280x720	-	-	-	79.9	92.9
	SoccerNet	2560x1440	x2	84.5	92.9
		3840x2160	x3	84.2	92.9
		5120x2880	x4	84.2	92.9
		7680x4320	x6	84.2	93.0
	UHDSR8K	2560x1440	x2	84.4	92.8
		3840x2160	x3	84.3	93.0
		5120x2880	x4	**84.8**	92.9
		7680x4320	x6	84.3	**93.1**

Table 3 and Table 4 present our findings on the impact of super-resolution techniques on object detection. We evaluated an object detection model in its original form, without any additional training on the the dataset. Applying super-resolution to 320x240 images leads to an approximate 12% improvement in the mAP @IoU=0.50:0.95 metric, while the mAP@IoU=0.50 metric shows a decline. A similar trend is observed with 640x480 images, where the mAP @IoU=0.50:0.95 metric increased from 27.6 to 28.5 and 1280x720, where we observed an increase from 29.5 to 30.6. Super resolution can reduce performance

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

at lower IoU thresholds (e.g., 0.50), indicating a trade-off between detection sensitivity and localization accuracy. The mAP@IoU=0.75 results further confirm that super-resolution primarily benefits higher-precision localization, especially for very low-resolution inputs (e.g., 320×240), but also brings noticeable improvements at higher resolutions—for instance, from 26.4 to 32.1 and from 27.5 to 34.3. Super-resolution also enhanced average IoU values across evaluated thresholds, improving bounding box precision. However, increasing image size does not always yield better detection performance. A sixfold upscaling did not show a clear advantage over a fourfold increase, possibly due to distortions introduced by the super-resolution network. Nonetheless, super-resolution consistently outperformed the baseline without image enhancement.

Although Table 2 suggests that super-resolution models trained on the SoccerNet dataset have superior quality, the relation to their detection efficacy remains ambiguous. The detector's performance was similar regardless of the training dataset. When upscaling 320x240 images by six times, the UHDSR8K-trained model achieved a higher mAP (27.2 vs. 26.6 for SoccerNet). Interestingly, even for 1280x720 images quality improved, but the most significant detection enhancement occurred at 320x240, with a 12.3% mAP@IoU=0.50:0.95 increase, compared to 3.3% for medium and 3.7% for large images.

In Table 5, the impact of classification of each class is considered for the original image with dimensions of 320x240. Notably, the table reveals that the application of super-resolution enhances the detection of players (person class). However, the detection of smaller objects, such as a ball, continues to pose challenges; in the original image, the mean Average Precision (mAP) for the ball was already low at 11. After applying super-resolution, this metric declined further to 0. Detecting small objects remains a significant hurdle, as discussed in [35], and unfortunately, the proposed method did not yield improvements in the detection quality of small objects.

Table 5. Detection performance of Faster R-CNN model for various super-resolution scales trained on SoccerNet for input of shape 320x240.

		mAP @IoU=0.50		**mAP @IoU=0.50:0.95**	
Output Shape	**Scale**	**Person**	**Ball**	**Person**	**Ball**
-	-	45.1	**3.5**	82.3	**11.2**
(640, 480)	x2	52.0	0.0	79.7	0.0
(960, 720)	x3	53.9	0.0	81.7	0.0
(1280, 960)	x4	**54.6**	0.0	**82.8**	0.0
(1920, 1440)	x6	53.1	0.0	81.6	0.0

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

5 Discussions and Conclusions

Our study highlights the crucial role of advanced super-resolution (SR) techniques in improving object detection in football. Integrating SR methods with object detection models resulted in 12% increase in mAP for low-quality images. We also highlights the importance of selecting a suitable dataset for training SR models. Our analysis revealed that training the SR algorithm with football-specific data slightly enhanced the quality of super-resolution on SoccerNet, as compared to SR models that were trained on generic data.

The practical implications of these enhancements are far-reaching. For instance, in automated video analysis for coaching and tactical evaluation, higher detection accuracy enables more detailed and nuanced analysis of players' movements and actions. Similarly, in the context of automated officiating and player tracking, improved object detection can contribute to more reliable and fair decision-making processes. Furthermore, these improvements can lead to cost reductions as there is less need to invest in high-quality cameras; the super-resolution technology compensates by enhancing the quality of the images captured by more standard equipment.

Although this study focuses solely on object detection with super-resolution in the context of football, its findings can be extended to other sports and even entirely different domains. For example, they can enhance facial recognition, license plate reading, and suspicious activity detection in low-resolution CCTV footage, thereby improving public safety and investigative capabilities. Similarly, in manufacturing, particularly in quality control, super-resolved images can help identify defects in products or machinery even when using standard cameras.

The limitation of our approach is the lack of real-time processing capability, as the SR preprocessing introduces computational overhead that may be unsuitable for latency-sensitive applications such as live broadcasts; however, it remains valuable in offline scenarios, which are common in professional sports environments where post-game analysis is typically conducted.

Looking ahead, our research opens several avenues for future exploration. One promising direction is the investigation of real-time SR and object detection algorithms that can operate efficiently in a live broadcast environment. Another area of interest is the exploration of domain-specific SR techniques that are optimized for varying weather conditions and lighting environments typical of football matches. In our future work, we also plan to explore various super-resolution techniques and object detection models. We aim to evaluate the effectiveness of cutting-edge methods in enhancing image quality and accurately identifying objects. Our goal is to find optimal solutions for diverse applications. What is more, we would like to analyse how fine-tuning object detectors and using sliding windows [31] affects detection quality.

We also acknowledge potential ethical concerns. Enhanced tracking may raise privacy issues, particularly outside sports contexts. Additionally, detection models can exhibit biases, such as varying accuracy across skin tones or lighting. Future work should include fairness testing and privacy impact assessments to ensure responsible deployment.

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

In conclusion, the integration of advanced super-resolution methods into object detection frameworks presents a significant advancement in the analysis of football imagery. By addressing the challenges posed by low-resolution images, this approach enhances the accuracy of object detection, thereby offering valuable insights for both tactical analysis and the development of automated officiating systems.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Acknowledgments

A.W. was supported by OMINO grant (no 101086321) funded by the European Union under the Horizon Europe and by the Polish Ministry of Education and Science within the International Projects Co-Financed program. (However, the views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the European Research Executive Agency can be held responsible for them.)

References

1. Li, Y., Zhang, Y., Timofte, R., Van Gool, L., Yu, L., Li, Y., Li, X., Jiang, T., Wu, Q., Han, M., et al.: NTIRE 2023 challenge on efficient super-resolution: Methods and results. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1922–1960 (2023)
2. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE transactions on pattern analysis and machine intelligence, 39(6), 1137-1149 (2016)
3. Deliege, A., Cioppa, A., Giancola, S., Seikavandi, M.J., Dueholm, J.V., Nasrollahi, K., Ghanem, B., Moeslund, T.B., Van Droogenbroeck, M.: Soccernet-v2: A dataset and benchmarks for holistic understanding of broadcast soccer videos. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4508–4519 (2021)
4. Kong, F., Li, M., Liu, S., Liu, D., He, J., Bai, Y., Chen, F., Fu, L.: Residual local feature network for efficient super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 766–776 (2022)
5. Torres-Ronda, L., Beanland, E., Whitehead, S., Sweeting, A., Clubb, J.: Tracking systems in team sports: a narrative review of applications of the data and sport specific analysis. Sports Medicine-Open **8**(1), 15 (2022)
6. Olagoke, A.S., Ibrahim, H., Teoh, S.S.: Literature survey on multi-camera system and its application. IEEE Access **8**, 172892–172922 (2020)
7. Wang, X.: Intelligent multi-camera video surveillance: A review. Pattern Recognition Letters **34**(1), 3–19 (2013)
8. Breitenstein, M.D., Reichlin, F., Leibe, B., Koller-Meier, E., Van Gool, L.: Online multiperson tracking-by-detection from a single, uncalibrated camera. IEEE Transactions on Pattern Analysis and Machine Intelligence **33**(9), 1820–1833 (2010)

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

9. Ma, C., Yang, F., Li, Y., Jia, H., Xie, X., Gao, W.: Deep trajectory post-processing and position projection for single & multiple camera multiple object tracking. International Journal of Computer Vision **129**, 3255–3278 (2021)
10. Akan, S., Varlı, S.: Use of deep learning in soccer videos analysis: survey. Multimedia Systems **29**(3), 897–915 (2023)
11. Kim, H., Choi, H.-J., Kim, C.J., Yoon, J., Ko, S.-K.: Ball trajectory inference from multi-agent sports contexts using set transformer and hierarchical bi-lstm. In: Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 4296–4307 (2023)
12. Chauhan, K., Patel, S.N., Kumhar, M., Bhatia, J., Tanwar, S., Davidson, I.E., Mazibuko, T.F., Sharma, R.: Deep learning-based single-image super-resolution: A comprehensive review. IEEE Access **11**, 21811–21830 (2023)
13. Lepcha, D.C., Goyal, B., Dogra, A., Goyal, V.: Image super-resolution: A comprehensive review, recent trends, challenges and applications. Information Fusion **91**, 230–260 (2023)
14. Zhang, K., Li, D., Luo, W., Ren, W., Stenger, B., Liu, W., Li, H., Yang, M.-H.: Benchmarking ultra-high-definition image super-resolution. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 14769–14778 (2021)
15. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10684–10695 (2022)
16. Zhang, R., Wu, L., Yang, Y., Wu, W., Chen, Y., Xu, M.: Multi-camera multi-player tracking with deep player identification in sports video. Pattern Recognition **102**, 107260 (2020)
17. Manafifard, M., Ebadi, H., Abrishami Moghaddam, H.: A survey on player tracking in soccer videos. Computer Vision and Image Understanding **159**, 19–46 (2017)
18. Komorowski, J., Kurzejamski, G.: Graph-based multi-camera soccer player tracker. In: 2022 International Joint Conference on Neural Networks (IJCNN), pp. 1–8 (2022) , 2022, pp. 1–8.
19. Naik, B.T., Hashmi, M.F., Geem, Z.W., Bokde, N.D.: DeepPlayer-track: player and referee tracking with jersey color recognition in soccer. IEEE Access **10**, 32494–32509 (2022)
20. Martins, E., Brito, J.H.: Soccer player tracking in low quality video. arXiv preprint arXiv:2105.10700 (2021)
21. Cioppa, A., Giancola, S., Deliege, A., Kang, L., Zhou, X., Cheng, Z., Ghanem, B., Van Droogenbroeck, M.: Soccernet-tracking: Multiple object tracking dataset and benchmark in soccer videos. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3491–3502 (2022)
22. Hurault, S., Ballester, C., Haro, G.: Self-supervised small soccer player detection and tracking. In: Proceedings of the 3rd International Workshop on Multimedia Content Analysis in Sports, pp. 9–18 (2020)
23. Luo, W., Xing, J., Milan, A., Zhang, X., Liu, W., Kim, T.-K.: Multiple object tracking: A literature review. Artificial Intelligence **293**, 103448 (2021)
24. Vicente-Martínez, J.A., Márquez-Olivera, M., García-Aliaga, A., Hernández-Herrera, V.: Adaptation of YOLOv7 and YOLOv7_tiny for soccer-ball multi-detection with DeepSORT for tracking by semi-supervised system. Sensors **23**(21), 8693 (2023)
25. Lewin, S., Vandegar, M., Hoyoux, T., Barnich, O., Louppe, G.: Dynamic NeRFs for soccer scenes. In: Proceedings of the 6th International Workshop on Multimedia Content Analysis in Sports, pp. 113–121 (2023)

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11

26. Wang, Y., Bashir, S.M.A., Khan, M., Ullah, Q., Wang, R., Song, Y., Guo, Z., Niu, Y.: Remote sensing image super-resolution and object detection: Benchmark and state of the art. Expert Systems with Applications **197**, 116793 (2022)
27. Dong, H., Xie, K., Xie, A., Wen, C., He, J., Zhang, W., Yi, D., Yang, S.: Detection of occluded small commodities based on feature enhancement under super-resolution. Sensors **23**(5), 2439 (2023)
28. Ji, X., Liu, G.-P., Cai, C.-T.: Collaborative framework for underwater object detection via joint image enhancement and super-resolution. Journal of Marine Science and Engineering **11**(9), 1733 (2023)
29. Hashmi, M.F., Naik, B.T., Keskar, A.G.: BDTA: events classification in table tennis sport using scaled-YOLOv4 framework. Journal of Intelligent & Fuzzy Systems **44**(6), 9671–9684 (2023)
30. Korhonen, J., You, J.: Peak signal-to-noise ratio revisited: Is simple beautiful? In: 2012 Fourth International Workshop on Quality of Multimedia Experience, pp. 37–38 (2012)
31. Akyon, F.C., Altinuc, S.O., Temizel, A.: Slicing aided hyper inference and fine-tuning for small object detection. In: 2022 IEEE International Conference on Image Processing (ICIP), pp. 966–970 (2022)
32. Held, J., Cioppa, A., Giancola, S., Hamdi, A., Ghanem, B., Van Droogenbroeck, M.: VARS: Video assistant referee system for automated soccer decision making from multiple views. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 5086–5097 (2023)
33. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
34. Na, B., Fox, G.C.: Object classifications by image super-resolution preprocessing for convolutional neural networks. Advances in Science, Technology and Engineering Systems Journal (ASTESJ) **5**(2), 476–483 (2020)
35. Cheng, G., Yuan, X., Yao, X., Yan, K., Zeng, Q., Xie, X., Han, J.: Towards large-scale small object detection: Survey and benchmarks. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(11), 13467–13488 (2023)

ICCS Camera Ready Version 2025
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-97626-1_11