The influence of data augmentation on the performance of YOLOv8 in object detection in basketball games

Kyana Marckx

January 25, 2026

Abstract

TODO

Keywords – data augmentation, YOLOv8, object detection, basketball

1 Introduction

TODO

2 Literature

2.1 Object detection and the YOLO family

2.1.1 Object detection

Object detection is a fundamental task within computer vision that aims to both localize and classify objects in images or video frames [4]. Unlike image classification, object detection requires models to predict bounding boxes in addition to class labels, which makes it a more complex problem. Recent advances in deep learning have led to significant improvements in accuracy and robustness, enabling object detection models to be deployed in real-time applications.

2.1.2 The YOLO family

The *You Only Look Once (YOLO)* family of architectures is among the most widely used approaches for real-time object detection. Since the introduction of the original YOLO model by Redmon et al. [9], successive versions have focused on improving detection accuracy, inference speed and architectural stability. A key advantage of YOLO-based models is their single-stage detection pipeline, which allows them to process images efficiently while maintaining competitive performance.

YOLOv8, introduced by Ultralytics in 2023 [2], represents a substantial update compared to earlier versions. It features an improved backbone, anchor-free detection heads and updated loss functions, leading to improved performance, particularly for small and fast-moving objects. Due to its high inference speed, YOLOv8 is well suited for real-time applications such as surveillance systems, autonomous systems and sports analytics.

Recent comparative studies further support the effectiveness of YOLOv8. Research comparing YOLOv5, YOLOv8 and YOLOv9 in combat sport scenarios showed that YOLOv8 offers a strong balance between detection accuracy and computational efficiency across object detection, tracking and human action recognition tasks [8]. These findings reinforce the suitability of YOLOv8 for dynamic sports environments.

At the time of this research (January 2026), YOLOv12 represents the most recent officially released version, with further versions already announced. Nevertheless, YOLOv8 remains a stable and well-supported model, making it a suitable choice for systematic experimental evaluation.

2.2 Data augmentation

A recurring challenge in deep learning-based computer vision systems is the limited availability of diverse and representative training data. Data augmentation addresses this limitation by artificially increasing dataset diversity through transformations applied to existing images. These techniques have been shown to reduce overfitting and improve the generalization capabilities of deep learning models [3].

2.2.1 Augmented data versus synthetic data

Although both augmented data and synthetic data aim to increase dataset size and diversity, they differ fundamentally in their approach. Augmented data is created by applying relatively simple transformations to real images, such as rotation, cropping or brightness adjustments. In contrast, synthetic data is fully generated using techniques such as generative models or 3D simulations. Synthetic data enables the creation of controlled and rare scenarios that may be difficult to capture in real-world settings and is particularly valuable in domains with privacy or data availability constraints, such as medical imaging [5].

2.2.2 Types of data augmentation

Data augmentation techniques can be categorized based on the data modality to which they are applied; most commonly audio, text and image data [1]. Since this research focuses on visual data from basketball broadcasts, image-based augmentation techniques are of primary relevance.

Image augmentation methods can be grouped into several categories:

- **Geometric transformations** such as *crop, flip, scale, rotate, translate* and *stretch* improve robustness to camera viewpoint changes and object orientation.
- **Color space transformations** including adjustments to *brightness, contrast* and *saturation* simulate varying lighting conditions commonly encountered in broadcast footage.
- **Kernel-based filters** such as *blur* or *sharpen* help models become more robust to *motion blur* and focus variations.
- **Random erasing** removes random regions of an image to simulate partial occlusions.
- **Image mixing techniques** combine multiple images to increase contextual diversity.

These forms of augmentation are particularly relevant for sports analytics, where fast movement and changing lighting conditions frequently occur.

2.3 Object detection in sports and basketball analytics

Object detection has become an essential component of modern sports analytics, enabling automated player detection, referee identification, ball tracking and the extraction of game statistics. Due to their real-time processing capabilities, YOLO-based models are frequently adopted in sports-related computer vision systems.

Several studies have demonstrated the effectiveness of YOLO architectures in sport contexts. In football analytics, systems combining YOLOv5 with DeepSORT have achieved strong performance in player detection and tracking tasks [10]. In basketball-specific research, YOLOv8 has been successfully applied to detect players, the ball and court elements, demonstrating its suitability for complex basketball scenes [6].

Despite these advances, object detection in sports remains a challenging task. Sports footage is characterized by rapid movement, frequent camera angle changes, varying illumination, and significant occlusions between players. Ball detection poses an additional challenge due to the small size and high velocity of the ball, as highlighted in recent reviews of ball detection techniques in sports [7]. These factors can negatively affect detection performance, particularly when models are trained on limited or insufficiently diverse datasets.

2.4 Research gap and motivation

The existing literature clearly shows that YOLO-based object detection models achieve strong performance in real-time and sports-related applications. Additionally, data augmentation is widely recognized as an effective strategy for improving model generalization and robustness. However, the specific impact of different image augmentation strategies on YOLOv8 performance in basketball contexts has not been studied in depth.

Most existing research focuses on model architecture comparisons or tracking frameworks, while data augmentation is often treated as a secondary component. This research addresses this gap by systematically evaluating how different categories of image augmentation influence YOLOv8 detection performance for basketball-specific objects, including players, referees, the ball, the hoop and scoreboard elements. The goal is to provide practical insights into building more robust and reliable real-time basketball detection systems.

3 Research questions

Based on the preceding literature review and problem definition, this study focuses on analyzing the impact of different image data augmentation techniques on the performance of the YOLOv8 model for object detection in basketball game footage.

3.1 Main research question

What is the effect of different image data augmentation techniques on the detection performance of YOLOv8 when identifying objects such players, the basketball, the shot clock and other relevant elements in basketball games?

3.2 Sub-research questions

- To what extent do different image augmentation techniques (such as geometric transformations, color space adjustments and kernel-based filters) improve the detection performance of YOLOv8 compared to a model trained on a non-augmented dataset?
- Which augmentation techniques have the greatest impact on the detection of small and fast-moving objects such as the basketball?
- Are certain object classes (e.g. players, ball, shot clock) more sensitive to specific augmentation techniques than others?

3.3 Hypotheses

- A YOLOv8 model trained using data augmentation achieves a significantly higher mean Average Precision (mAP) compared to a model trained without data augmentation.
- Models trained with a combination of multiple augmentation techniques outperform models trained with individual augmentation categories. Among individual augmentation types, kernel-based filters are expected to have the strongest positive impact on detection performance.
- The effectiveness of data augmentation varies across object classes, with smaller and faster-moving objects such as the basketball benefiting more from augmentations such as blur, rotation and brightness adjustments.

4 Methodology

4.1 Overview

This study investigates the impact of different image data augmentation strategies on the performance of a YOLOv8-based object detection system for basketball game footage. The methodology is structured as a sequence of experimental steps, each implemented in a separate notebook. The workflow starts with dataset preparation and augmentation, followed by model training, evaluation and quantitative comparison of results across multiple experimental settings.

All experiments are conducted using the same base model architecture and training configuration to ensure fair and reproducible comparisons.

4.2 Dataset and preprocessing

The experiments are based on a publicly available basketball object detection dataset[1] containing annotated images of players, referees, the ball, the hoop and other game elements (period, shot clock, team name, team points and time remaining). Before training, a preprocessing pipeline was applied uniformly across the original dataset since it serves as the baseline for all augmentations. This pipeline includes automatic image orientation correction and resizing to a fixed resolution of 640x640 pixels with white padding to preserve the aspect ratio of the original image. No augmentations were applied at this stage to the original dataset.

The dataset is split into training (120 images, 70%), validation (25 images, 15%) and test (25 images, 15%) sets using a consistent split across all experiments to ensure comparability of evaluation results.

[1] https://app.roboflow.com/kiejaana/basketball-players-fy4c2-dqb1n/1

4.3 Creation of augmented datasets

Three augmented datasets were created by applying specific categories of image augmentation techniques to the original dataset. Each augmented dataset focuses on a single augmentation category to allow for isolated evaluation of its effect. The considered augmentation groups include geometric transformations, color space adjustments and kernel-based filters.

Care is taken to avoid duplication of original images across datasets. Each augmented dataset contains the original images alongside their augmented counterparts, ensuring that the total dataset size increases while preserving the original data distribution.

4.4 YOLOv8 training on baseline dataset

To establish a reference point, a baseline YOLOv8[2] model was trained using the original, non-augmented dataset. The lightweight YOLOv8n variant was selected to prioritize real-time performance while maintaining sufficient detection accuracy.

The baseline model is trained using fixed hyperparameters, including image size, batch size, number of epochs and random seed. During baseline training, built-in data augmentation within YOLOv8 is disabled to isolate the effect of external augmentation strategies introduced later in the workflow.

The trained baseline model is evaluated on the held-out test set, and standard object detection metrics such as mean Average Precision (mAP50 and mAP50-95), precision and recall are recorded.

4.5 YOLOv8 training on augmented datasets

Separate YOLOv8 models are trained on each augmented dataset using the same training configuration as the baseline experiment. This ensures that any observed performance differences can be attributed to the applied augmentation strategy rather than changes in model architecture or hyperparameters.

For each augmented dataset, the trained model is evaluated on the same test set used for the baseline model. Evaluation metrics are collected consistently across all experiments.

4.6 Creation of combined dataset

In addition to training on individual augmentation types, a combined dataset is constructed by merging the different augmented datasets into a single training set. Duplicate images are removed to prevent overrepresentation of original samples. This combined dataset aims to capture the complementary benefits of multiple augmentation strategies within a single training process.

4.7 YOLOv8 training on combined dataset

Another YOLOv8 model is trained on this combined dataset using the same training configuration and evaluated on the test set.

4.8 Evaluation metrics

Model performance is assessed using standard object detection metrics, including mean Average Precision at an Intersection over Union (IoU) threshold of 0.5 (mAP50), mean Average Precision averaged over thresholds from 0.5 to 0.95 (mAP50-95), precision and recall. In addition to accuracy-based metrics, inference-related timing metrics are recorded to evaluate the suitability of each model for real-time deployment.

All metrics are computed using the YOLOv8 validation framework to ensure consistency across experiments.

4.9 Results aggregation and comparison

To facilitate quantitative comparison, evaluation results from all experiments are exported to CSV files and subsequently concatenated into a single CSV file. This aggregated results file enables direct comparison across baseline, individual augmentation strategies and the combined dataset.

Visualizations are generated to compare performance metrics and inference times across experiments, allowing trends and trade-offs to be clearly identified.

[2] https://docs.ultralytics.com/models/yolov8/

5 Results

TODO

6 Conclusion

TODO

7 Discussion

TODO

References

[1] A. A. Awan. A complete guide to data augmentation, Dec. 2024.

[2] G. Jocher, A. Chaurasia, and J. Qiu. Ultralytics yolov8, 2023.

[3] E. Kavlakoglu and J. Murel. What is data augmentation? — ibm, May 2024.

[4] E. Kavlakoglu and J. Murel. What is Object Detection? | IBM, Jan. 2024.

[5] D. Keskin. Synthetic data and data augmentation, Nov. 2023.

[6] Z. Liang, J. Wang, T. Huang, Z. Sang, and J. Zhang. Basketball detection based on yolov8. *PLOS ONE*, 20(8):e0326964, Aug. 2025.

[7] C. Moreira, L. Ferreira, and P. Coelho. A comprehensive review of ball detection techniques in sports. *PeerJ Computer Science*, 11:e3079, 08 2025.

[8] E. Quinn and N. Corcoran. Comparative analysis of yolov5/v8/v9 for object detection, tracking, and human action recognition in combat sports. *International Conference on AI Research*, 4:364–373, 12 2024.

[9] J. Redmon, S. K. Divvala, R. B. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. *CoRR*, abs/1506.02640, 2015.

[10] B. Wang. Football sports video tracking and detection technology based on yolov5 and deepsort. *Discover Applied Sciences*, 7(6):563, May 2025.