Black Sea Journal of Artificial Intelligence
doi: 10.5281/zenodo.18213771

Open Access Journal
e-ISSN: 3108-4281

Research Article
Volume 1 - Issue 2: 35-39 / December 2025

ANALYSIS OF OPTIMIZED DATASETS FOR BASIC IMAGE PROCESSING ALGORITHMS

Duygu Sedef ÇALIŞKAN[1]*, Mehmet Serhat ODABAŞ[1], Recai OKTAŞ[2]

[1]Ondokuz Mayıs University, Institute of Graduate, Department of Smart Systems Engineering, Samsun, Türkiye
[2]Ondokuz Mayıs University, Faculty of Engineering, Department of Computer Engineering, Samsun, Türkiye

Abstract: In recent years, rapid advances in artificial intelligence and computer vision have significantly enhanced object detection systems, which are now widely used in fields such as autonomous driving, surveillance, and sports analytics. This study focuses on evaluating and comparing four state-of-the-art object detection architectures (YOLOv11, YOLOv12, Roboflow 3.0, and RF-DETR) to determine their effectiveness in real-time detection of basketball players. A publicly available dataset containing 170 annotated images from basketball game scenarios was obtained from the Roboflow platform. Each model was trained using identical hyperparameter configurations to ensure a fair comparison, and its performance was evaluated using mAP@50, Precision, and Recall metrics. The results demonstrate that RF-DETR achieved the highest overall accuracy (mAP@50 = 91.5%), while YOLOv11 showed the best balance between recall (84.3%) and precision (90.2%), making it ideal for real-time applications. These findings underscore the increasing capability of modern AI models to perform reliable object detection in complex and dynamic environments. As deep learning technologies continue to evolve, such comparative studies provide essential insights for selecting the most efficient architectures for real-world implementations.

Keywords: Image processing, Dataset, YOLO

***Corresponding author:** Ondokuz Mayıs University, Institute of Graduate, Department of Smart Systems Engineering, Samsun, Türkiye
E mail: sedefcalkn@gmail.com (D. S. ÇALIŞKAN)
Duygu Sedef ÇALIŞKAN https://orcid.org/0009-0005-7588-3498
Mehmet Serhat ODABAŞ https://orcid.org/0000-0002-1863-7566
Recai OKTAŞ https://orcid.org/0000-0003-3282-3549

Received: September 10, 2025
Accepted: December 05, 2025
Published: December 15, 2025

Cite as: Çalışkan, D. S., Odabaş, M. S., & Oktaş, R. (2025). Analysis of optimized datasets for basic image processing algorithms. *Black Sea Journal of Artificial Intelligence*, *1*(2), 35–39.

1. Introduction

Image processing, a fundamental branch of artificial intelligence, enables computers to interpret and analyze visual information from images and videos (Zhao et al., 2019; Sun et al., 2024). With the integration of deep learning, it has evolved into data-driven models that perform complex tasks such as detection, segmentation, and object tracking across diverse fields like healthcare, automation, and surveillance (Tao et al., 2022). Object detection is a crucial component of computer vision systems, enabling the automated identification and localization of objects within images or video frames (Jegham et al., 2024). It plays a vital role in various computer vision applications, including autonomous driving, surveillance, and sports analytics. This advanced technology serves as a foundation for a wide range of applications across multiple industries, including robotics, security, and healthcare. Therefore, object tracking is a crucial component of video analysis, as it enhances object detection by ensuring temporal consistency across consecutive frames (Trigka and Dritsas, 2025).

YOLO (You Only Look Once) is one of the leading and most popular object detection approaches for performing object detection. Unlike traditional two-stage methods, YOLO processes the entire image at once and simultaneously predicts class probabilities and bounding box coordinates, which significantly improves speed. Various techniques have been designed for crowd detection in videos and images (Bochkovskiy et al., 2020). Providing a balance between high accuracy and low latency (Esteva et al., 2017). In recent years, deep learning–based object detection models, such as YOLO and transformer-based architectures, like Detection Transformers (DETR), have significantly advanced detection accuracy and inference speed (Bochkovskiy et al., 2020; Carion et al., 2020; Simic and Gavrovska, 2025). However, detecting small or overlapping objects under varying lighting and background conditions remains a persistent challenge (Khalili and Smyth 2024; Xiaozheng et al., 2025).

To address these limitations, this study compares four state-of-the-art architectures (YOLOv11, YOLOv12, Roboflow 3.0, and RF-DETR) using a unified dataset obtained from the Roboflow platform. The selected dataset consists of basketball-player images designed for real-time detection tasks. Each model was trained under identical conditions, and their performance was evaluated based on precision, recall, and mAP@50 metrics to determine the most effective architecture for accurate and real-time sports-oriented object detection.

2. Materials and Methods

2.1. Dataset

Historically, object detection datasets are created by collecting a large collection of images and sourcing annotators to label objects within a fixed set of classes (Ciaglia et al., 2022). The Roboflow platform facilitates this process by allowing users to upload, annotate, and share datasets with high accuracy. As noted in (Sakib et al., 2023), with Roboflow's assistance, all objects in the dataset were annotated with great detail and accuracy, ensuring that each object was accurately classified and its location was determined using bounding boxes.

In this study, an open-source dataset for detecting basketball players was obtained from the Roboflow platform. Each image contained multiple players in different poses and lighting conditions, and their positions were labeled with bounding boxes. Using Roboflow’s bounding box tool and the generated annotation files, the ground truth bounding boxes in each image were automatically constructed, ensuring consistent and highly accurate annotations (Yigit, 2025). The selected dataset was deemed more suitable for object detection models because it represents a dynamic and challenging environment characterized by multiple overlapping targets, rapid movement, and complex backgrounds, similar to a basketball game. This makes it suitable for evaluating YOLO and transformer-based architectures in terms of accuracy, robustness, and real-time performance.

The YOLO family approaches object detection as a single-stage regression problem, directly estimating the class probabilities p(ci) and location coordinates (x, y, w, h) of objects in an image (Ultralytics, 2024).

Overall detection confidence is defined as given in Equation 1:

$$Confidence = P(Object)\ x\ IoU_{pred}^{truth} \tag{1}$$

YOLOv11 and YOLOv12 are the latest versions of the Ultralytics series, which balance high accuracy with real-time performance using a CSPDarknet-based backbone and PANet/ConvNeXt-based feature fusion layers (Audu and Ndirmbula, 2024).

Roboflow 3.0 enhances the model's generalization performance by combining the optimization processes of the YOLO algorithm with advanced data preprocessing and data augmentation techniques.

RF-DETR (Roboflow-DETR) employs a transformer-based encoder-decoder architecture, capturing the contextual relationships between objects in the image through attention mechanisms (Redmon et al., 2016).

2.2. Dataset Properties

The dataset consists of a total of 170 annotated images, each depicting basketball game scenarios captured from YouTube videos at various angles and lighting conditions. It includes nine object classes, with “Player” being the primary class. The object categories are: Ball, Hoop, Period, Player, Referee, Shot Clock, Team Name, Team Points, and Time Remaining. The dataset segmentation was given in Figure 1.

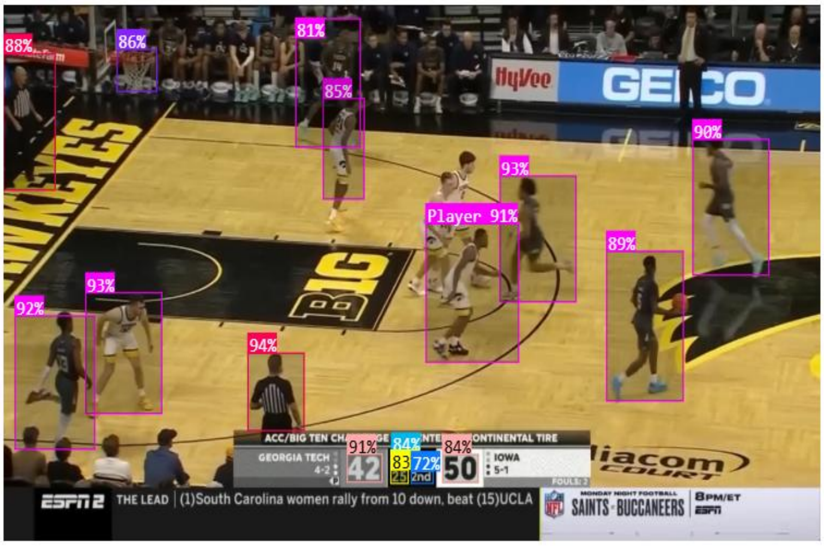


Figure 1. Dataset segmentation.

Each image was annotated with bounding boxes that precisely delineate object boundaries. The dataset was divided into 67% training, 19% validation, and 14% testing subsets to ensure balanced model evaluation. All images were standardized to a resolution of 640 × 640 pixels during preprocessing to maintain consistency across all trained models.

2.3. Data Augmentation

To improve the generalization capability of the models, various data augmentation techniques were applied through the Roboflow platform. These include horizontal and vertical flipping, brightness and contrast adjustment, rotation, and Gaussian blur. Such augmentations simulate real-world variations, helping the models to become more robust against scale, pose, and lighting changes commonly encountered in sports environments.

2.4. Model Architecture and Training Settings

In this study, four different object detection architectures (YOLOv11, YOLOv12, Roboflow 3.0, and RF-DETR) were trained and evaluated under identical experimental conditions. Each model was selected to represent a different approach to object detection: the YOLO family for real-time convolutional architectures, and RF-DETR for transformer-based detection (Chandana and Ramachandra, 2022; Zhang et al., 2023). Models were trained using the Roboflow platform, which provides an integrated training pipeline with GPU acceleration. Each training session lasted approximately 40–30 minutes, depending on the architecture and model complexity. The same dataset version, preprocessing pipeline, and augmentation techniques were applied across all models to ensure a fair comparison.

2.5. Hyperparameters

All models (YOLOv11, YOLOv12, Roboflow 3.0, and RF-DETR) were trained using the same hyperparameter configuration to ensure a fair and unbiased comparison (Table 1). This uniform setup ensures that any observed performance differences are due solely to the architectural design of each model, not to differences in training conditions.

Table 1. Training hyperparameters and configuration

Image size	640 × 640 pixels
Batch size	16
Epochs	50
Optimizer	Adam optimizer with default learning-rate scheduling
Checkpoint	MS COCO pretrained weights were used to initialize all models

3. Results and Discussion

3.1. Evaluation Metrics

All four models were evaluated on the same test subset of the dataset to ensure a fair comparison. Model performance was assessed using three key metrics: mAP@50, Precision, and Recall. The mAP@50 metric measures the overall detection accuracy by calculating how precisely each model identifies and localizes objects. Precision indicates how many of the detected objects were correct, while Recall measures how many of the actual objects in the images were successfully detected. Using the same dataset and training configuration ensured that performance differences resulted only from variations in the model architectures.

The results indicate that RF-DETR achieved the highest overall detection accuracy, with a mAP@50 score of 91.5%, followed by YOLOv11 (87.3%), YOLOv12 (87.0%), and Roboflow 3.0 (86.2%). In terms of precision, YOLOv11 recorded the best performance (90.2%), closely followed by Roboflow 3.0 (90.1%) and RF-DETR (90.9%), demonstrating that all models maintained a low rate of false detections. Regarding recall, YOLOv11 again outperformed the others (84.3%), suggesting that it detected a greater proportion of true objects compared to the other architectures. Figure 2 shows the output of the proposed object detection model on a test frame from the basketball dataset.

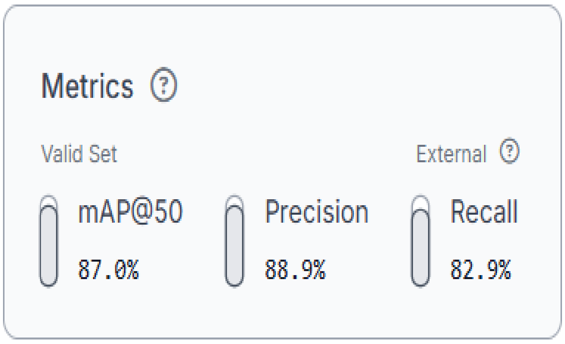
The bounding boxes demonstrate the model's ability to correctly detect multiple object classes, such as Player, Ball, and Hoop, with high confidence levels (e.g., 90%, 83%, and 98%). This qualitative example visually confirms the quantitative findings presented in Table 2, showing that the model can accurately identify and localize objects in complex, real-world scenes.

4. Conclusion

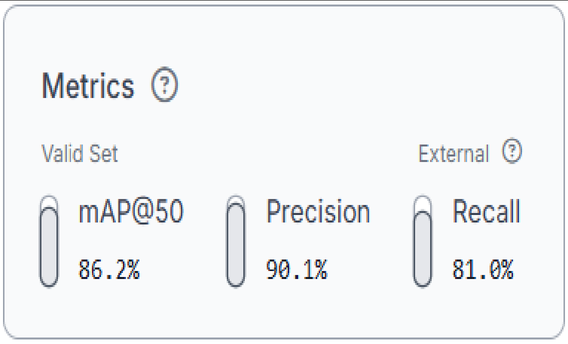
This study presented a comparative evaluation of four modern object detection architectures (YOLOv11, YOLOv12, Roboflow 3.0, and RF-DETR) trained in the same experimental environment to assess their effectiveness in real-time basketball player detection. The results showed that RF-DETR achieved the highest detection accuracy thanks to its attention-based transformative structure, while YOLOv11 offered the most balanced performance in terms of precision, recall, and inference speed, making it a strong candidate for real-time sports-focused applications. These findings highlight the growing maturity of contemporary AI models in handling visually complex scenarios involving rapid movements, occlusions, and varying lighting conditions.

Figure 2. Qualitative detection results on basketball frames.

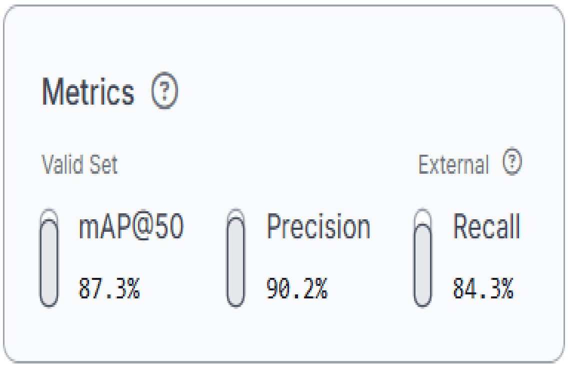
Table 2. Evaluation metrics



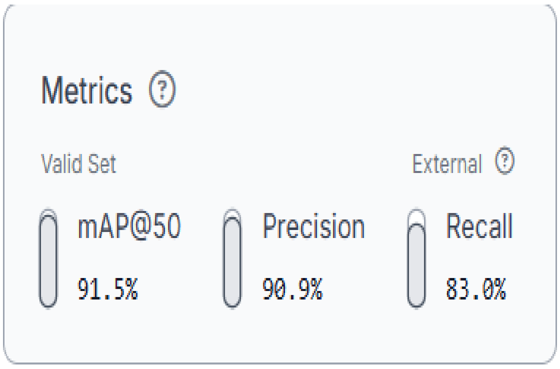
YOLOv12 Object Detection (Fast)



Roboflow 3.0 Object Detection (Fast)



YOLOv11 Object Detection (Fast)



RF-DETR (Base)

Despite these advancements, challenges remain, particularly in detecting small or overlapping targets and ensuring stability in dynamic environments. Future work could expand the dataset, incorporate multiple camera perspectives, evaluate lightweight architectures on edge devices, or explore hybrid CNN-transformer models to further improve accuracy and computational efficiency. Moreover, integrating advanced tracking algorithms could enable complete detection and tracking pipelines supporting next-generation sports analytics and automated broadcast systems. Furthermore, embedding temporal attention mechanisms could improve frame-to-frame consistency in fast-paced sequences, while domain adaptation strategies could help maintain robustness across different arenas and recording conditions. Ultimately, extending these approaches to multimodal frameworks that combine video, sensor, or contextual data could lead to more holistic and intelligent decision-making systems for real-world applications.

Author Contributions

The percentages of the authors' contributions are presented below. All authors reviewed and approved the final version of the manuscript.

	D.S.Ç.	M.S.O.	R.O.
C	40	30	30
D	40	30	30
S	40	30	30
DCP	40	30	30
DAI	40	30	30
L	40	30	30
W	40	30	30
CR	40	30	30
SR	40	30	30
PM	40	30	30
FA	40	30	30

C=Concept, D= design, S= supervision, DCP= data collection and/or processing, DAI= data analysis and/or interpretation, L= literature search, W= writing, CR= critical review, SR= submission and revision, PM= project management, FA= funding acquisition.

Conflict of Interest

The authors declared that there is no conflict of interest.

Ethical Consideration

Ethics committee approval was not required for this study because of there was no study on animals or humans.

References

Audu, A. I., & Ndirmbula, S. M. (2024). Machine learning based real-time detection of ripe and unripe sorghum. *Nigerian Journal of Engineering Science and Technology Research*, *10*(2), 124–137.

Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). *YOLOv4: Optimal speed and accuracy of object detection*. arXiv. https://doi.org/10.48550/arXiv.2004.10934

Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-end object detection with transformers. In A. Vedaldi, H. Bischof, T. Brox, & J. M. Frahm (Eds.), *Computer Vision – ECCV 2020* (pp. 213–229). Springer. https://doi.org/10.1007/978-3-030-58452-8_13

Chandana, R. K., & Ramachandra, A. C. (2022). *Real time object detection system with YOLO and CNN models: A review*. arXiv. https://doi.org/10.48550/arXiv.2208.00773

Ciaglia, F., Zuppichini, F. S., Guerrie, P., McQuade, M., & Solawetz, J. (2022). *Roboflow 100: A rich, multi-domain object detection benchmark*. arXiv. https://doi.org/10.48550/arXiv.2211.13523

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Jegham, N., Koh, C. Y., Abdelatti, M., & Hendawi, A. (2024). *YOLO evolution: A comprehensive benchmark and architectural review of YOLOv12, YOLO11, and their previous versions*. arXiv. https://doi.org/10.48550/arXiv.2411.00201

Khalili, B., & Smyth, A. W. (2024). *SOD-YOLOv8—Enhancing YOLOv8 for small object detection in traffic scenes*. arXiv. https://doi.org/10.48550/arXiv.2408.04786

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. https://doi.org/10.1109/CVPR.2016.91

Sakib, K. S., Abir, T. I., Kareena, S. S., & Alam, R. (2023). *Autonomous object detection dataset: A study on Bangladeshi roadways* [Undergraduate thesis, Brac University]. Brac University Repository.

Simic, N., & Gavrovska, A. (2025). Comparative analysis of YOLOv11 and YOLOv12 for AI-powered aerial people detection. *12th International Conference on Electrical, Electronic and Computing Engineering (IcETRAN)*, 1–4.

Sun, Y., Sun, Z., & Chen, W. (2024). The evolution of object detection methods. *Engineering Applications of Artificial Intelligence*, *133*, 108458. https://doi.org/10.1016/j.engappai.2024.108458

Tao, X., Gong, X., Zhang, X., Yan, S., & Adak, C. (2022). Deep learning for unsupervised anomaly localization in industrial images: A survey. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1–21. https://doi.org/10.1109/TIM.2022.3196436

Trigka, M., & Dritsas, E. (2025). A comprehensive survey of machine learning techniques and models for object detection. *Sensors*, *25*(1), 214. https://doi.org/10.3390/s25010214

Ultralytics. (2024). *YOLOv12 official release documentation*. https://docs.ultralytics.com (Erişim tarihi: 3 Eylül 2025).

Xiaozheng, Z., Zhongjun, Y., & Huaici, Z. (2025). DCS-YOLOv8: A lightweight context-aware network for small object detection in UAV remote sensing imagery. *Remote Sensing*, *17*(17), 2989. https://doi.org/10.3390/rs17172989

Yigit, G. (2025). Crowd detection: Leveraging YOLO for human recognition. *Turkish Journal of Engineering*, *9*(3), 571–577. https://doi.org/10.31127/tuje.1627839

Zhang, J., Zhang, H., Liu, B., Qu, G., Wang, F., Zhang, H., & Shi, X. (2023). Small object intelligent detection method based on adaptive recursive feature pyramid. *Heliyon*, *9*(7), e17730. https://doi.org/10.1016/j.heliyon.2023.e17730

Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems*, *30*(11), 3212–3232. https://doi.org/10.1109/TNNLS.2018.2876865