Jurnal Teknik Informatika (JUTIF)
Vol. 5, No. 3, June 2024, pp. 811-820

DOI: https://doi.org/10.52436/1.jutif.2024.5.3.2015
p-ISSN: 2723-3863
e-ISSN: 2723-3871

VISUAL ENTITY OBJECT DETECTION SYSTEM IN SOCCER MATCHES BASED ON VARIOUS YOLO ARCHITECTURE

Mochamad Althaf Pramasetya Perkasa[1], R. Reza El Akbar[2], Muhammad Al Husaini*[3], Randi Rizal[4]

[1,2,3]Informatics, Engineering Faculty, Universitas Siliwangi, Indonesia
[4]Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Malaysia
Email: [1]207006080@student.unsil.ac.id, [2]reza@unsil.ac.id, [3]alhusaini@unsil.ac.id

(Article received: May 27, 2024; Revision: May 13, 2024; published: June 03, 2024)

Abstract

In this study, a performance comparison between the YOLOv7, YOLOv8, and YOLOv9 models in identifying objects in soccer matches is conducted. Parameter adjustments based on GPU storage capacity were also evaluated. The results show that YOLOv8 performs better, with higher precision, recall, and F1-score values, especially in the "Ball" class, and an overall accuracy (mAP@0.5) of 87.4%. YOLOv9 also performs similarly to YOLOv8, but YOLOv8's higher mAP@0.5 value shows its superiority in detecting objects with varying degrees of confidence. Both models show significant improvement compared to YOLOv7 in overall object detection performance. Therefore, based on these results, YOLOv8 can be considered as the model that is close to the best performance in detecting objects in the dataset used. This study not only provides insights into the performance and characteristics of the YOLOv7, YOLOv8, and YOLOv9 models in the context of object detection in soccer matches but also results in a dataset ready for additional analysis or for training deep learning models.

Keywords: *Deep Learning, Parameter, Soccer, YOLO.*

1. INTRODUCTION

Computer science has shown its potential in the sports industry in recent years. Comparison with photos shows that videos provide more information about changes in the situation over time. Object detection in sports videos requires more storage capacity and computing power than object detection in photos. However, this may be necessary in some situations where more than four objects are detected in a single frame [1]. Sports analytics plays a very important role in efforts to improve player performance, which is why the application of computer vision and virtual reality-based technologies is becoming increasingly necessary. This technology is used to perform accurate posture correction in various sports [2]. Computer vision-based evaluation systems on computers provide the necessary support in the decision-making process for training in sports. Application of object detection for sports analysis [3].

Soccer holds the title as the most popular sport worldwide, with over 40% of the respondents in the survey showing significant interest by answering "interested" or "very interested" in the sport [4]. As soccer is such a popular sport around the world, attention continues to be paid to the skills of the players. The ever-increasing popularity of soccer creates additional pressure for players to perform well on the pitch [5]. The expectation to get better results in every match makes the responsibility for coaches enormous. They are required to design the best strategy that will be reflected in the course of the match as well as the final result. However, it is known that there are various factors that can affect the performance of the players, both individually and as a whole [6].The implementation of intelligent algorithms in the field of data science has also become a common practice in the world of sports [7]. The rapid growth of sports video data on various internet platforms poses significant challenges in handling this information scientifically in the current era. Although recent years have witnessed significant progress in object detection and action detection research through deep learning, there has been little achievement in sports video detection [8].

Traditional methods use sensors to detect and record the athlete's key positions. After the raw data is analyzed using a deep learning-based approach, recommendations for training are then provided based on the results of the analysis [9] The addition of further sensors will increase costs, which may negatively impact the athlete's performance. Even in intense competition, the use of sensors can help identify weaknesses and strengths. It is difficult for a coach to remember and analyze every movement and action of every player after a match, in order to utilize that knowledge to guide players and prevent potential mistakes in the future. As a result, the job of a performance analyst, also known as a notation analyst, is to take on the responsibility of documenting the entire event, gathering information about the players' activities, their movements, as well as the timing of those activities. The information would then be comprehensively presented to the coach[10], as precise and effective identification of both players and the ball stands as a crucial aspect in any system aimed at automating the analysis of soccer match videos[11]. However, recent technologies in computer vision have replaced the traditional method of performance analysis with sports video analysis. The main function of this analysis technique is to present such information comprehensively to the coach [12]. Applying a deep learning approach is the key to achieving the goal of object detection in soccer sports videos [13]. Unlike previous machine learning methods that require manual generation of features to be extracted from the input, deep artificial neural networks are capable of learning and extracting features directly from the input.

The proposed model uses a pre-prepared YOLO-based deep learning network. This deep learning has been customized for object detection in sports videos [14]. YOLO (You Only Look Once) is a very simple approach, where a single convolutional network predicts bounding boxes as well as class probabilities for those boxes simultaneously. YOLO is trained on the full image and instantly optimizes the detection performance. This unified model has several advantages over traditional methods of object detection. The speed of YOLO is impressive. Since this approach views detection as a regression problem, no complex pipeline is required. Running the neural network on a new image at test time allows users to predict detection without the need for complicated batch processing. The basic network can run at 45 frames per second without batch processing on the Titan X GPU, and the faster version can even reach more than 150 FPS. It can thus process streaming video in real-time with less than 25 milliseconds of latency [15].

Research applying the YOLO algorithm to sports matches was conducted by Patel and Kamdar [14], which aims to detect objects from hockey matches, specifically the players and balls used. The main focus of this research was to improve the overall average object detection accuracy of the model used. The total dataset used reached 1119 annotated images, with a total of 4 classes. The first experiment was conducted with 100 epoch iterations, resulting in an average accuracy of 88.9%. The second experiment, with 200 epoch iterations, achieved an accuracy score of 91.2%, while the third experiment used 300 epoch iterations, resulting in an accuracy score of 91.3%.

This research will expand the scope of Patel and Kamdar [14] research dataset, by including datasets that cover various matches. Previously, the comparison was only based on the number of iterations. In this study, comparisons will be made using different versions of YOLO, namely YOLOV7, YOLOV8, and YOLOV9. The parameters used from each version of the YOLO model will be adjusted to the available GPU storage capacity. The purpose of this research is to evaluate and compare the performance of the YOLO model in identifying objects in soccer matches, and determine which version has the best performance.

2. RESEARCH METHOD

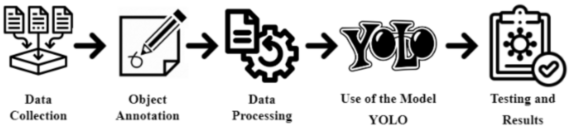

Fig. 1. Stages of YOLO Model Implementation

Fig. 1 shows the implementation stage of the model, where the development of the detection system using YOLO begins with the data collection stage of the Bundesliga 2022 video dataset. After the data is collected, the object annotation stage is carried out on each video frame to mark the location and type of objects such as players, referees, balls, and goals, using the Roboflow platform. The next step is data processing, which includes several stages, including data separation, data preprocessing, data augmentation, data normalization, and input preparation. After the data processing is complete, the YOLO model usage stage is continued. YOLOv7, YOLOv8, and YOLOv9 are used to develop object detection models with parameter adjustments. The next stage is training the object detection model using the labeled dataset, using the Python programming language and Google Colab to utilize large computing resources. After the model is trained, the detection results are tested and evaluated using validation datasets and videos.

3. RESULTS AND DISCUSSION

3.1. Data Collection

The dataset is taken from the Kaggle website, which provides match footage from the DFL-Bundesliga Data Shootout. The DFL is a data source that provides video footage of Bundesliga matches in 2022, which has been uploaded by the DFL through the Kaggle platform for use by researchers and practitioners in various analysis and model development.

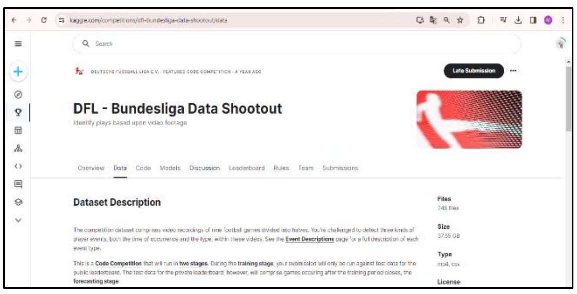

Fig. 2. Source of Dataset Used

Fig. 2 shows a snapshot explaining in detail about the dataset provided by DFL- Bundesliga Data Shootout via kaggle.

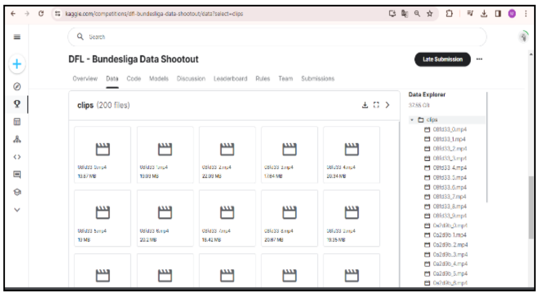

Fig. 3. Video Dataset

Fig. 3 shows the entire dataset in the clips folder provided by the DFL-Bundesliga Data Shootout. The contents of the folder are match footage from the Bundesliga, where each footage has an average duration of 30 seconds with a total file size of 200 videos, and the overall file size has a capacity of 3.7 GB.

3.2. Object Annotation

The object annotation stage using Roboflow for a dataset consisting of four classes, namely players, balls, goalkeepers, and referees, begins with the selection of the dataset to be annotated. After the dataset is selected, the next step is to choose an annotation tool that suits the needs, namely using the bounding box tool to mark objects in the image. The annotation process starts by marking each object by creating a bounding box around it, while providing a label corresponding to the class of the marked object. For example, each player, ball, goalkeeper, and referee will be labeled accordingly.


Fig. 4. Object Annotation Process

Figure. 4 displays the process of annotating images from the dataset using the Roboflow platform, which is an important step in data preparation for object detection model training. This process involves adding bounding boxes to identify and distinguish relevant objects, such as players, balls, goalkeepers, and referees.

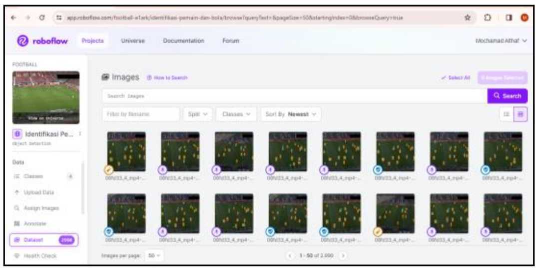

Fig. 5. Overall Dataset

Fig. 5 shows the captured annotation dataset after the annotation process. The figure shows the annotation results that have been used as a dataset to train the object detection model.

Table. 1. Metadata

Description	Value
Overall Image	2990
Number of Classes	4
Not Annotated	0
Training Data	2100 (70%)
Validation Data	588 (20%)
Test Data	299 (10%)
Dimension Size	1920×1080 pixels

Table. 1 is the metadata derived from the images that have gone through the annotation stage and then included in the dataset. This metadata provides an overview of the structure, distribution and important attributes of the dataset without requiring an individual review of each data sample. The information presented includes the total number of images in the dataset, the number of classes or labels present, and the number of images that have no annotations or labels. The table also provides details regarding the distribution of data for training, validation, and testing, along with the respective percentages.

3.3. Data Processing

The end result of this data processing is a dataset that is ready to be used for further analysis or training of deep learning models. This dataset consists of 2990 images that have been divided into three subsets, 70% for training data, 20% for validation data, and 10% for test data. The data has gone through preprocessing techniques to ensure cleanliness and readability before further analysis. Data augmentation has been applied to increase the variety and amount of training data, allowing the deep learning model to learn from diverse situations and improve its performance.

Table. 2. Data Processing Results

Description	Before Data Processing	After Data Processing
Overall Image	2990	7173
Number of Classes	4	4
Not Annotated	0	0
Training Data	2100 (70%)	6276 (87%)
Validation Data	588 (20%)	598 (13%)
Test Data	299 (10%)	299 (4%)
Dimension Size	1920×1080	640×640

Table. 2 is the result of processing the dataset. It can be seen that the total number of images increased from 2990 to 7173 after data processing. The percentage of training data increased from 70% to 87%, while the percentage of validation data and test data decreased after data processing. The change of image dimension size from 1920×1080 to 640×640 is also followed by the change of image dimension size from 1920×1080 to 640×640 after data processing.

3.4. Parameter Testing

The model was trained using data from 2022 Bundesliga matches to recognize objects at football matches in various match situations. Each YOLO model will undergo a different training process depending on the architecture and parameters used.

Table. 3. Parameter Comparison

Parameter	YOLOv7	YOLOv8	YOLOv9
Batch Size	14	16	8
Image Size	800 pixels	1080 pixels	640 pixels
Number of Iterations	30	30	30

Table. 3 gives an overview of how these parameters differ between the three versions of the YOLO model. For example, YOLOv8 uses a larger batch size (16) compared to YOLOv7 (14) and YOLOv9 (8). In addition, YOLOv8 also uses a larger image size (1080 pixels), while YOLOv9 uses a smaller image size (640 pixels). Despite the differences in parameters, these three versions use the same number of iterations (30), which indicates that the training process is performed with a similar number of iterations to compare the performance of models with different parameters.

3.5. Model Evaluation

This evaluation is done using performance metrics such as precision, recall, and f1-score. Helping to ensure that the model can accurately detect and identify patterns in the data that have not been seen before, as well as ensuring that the model is reliable in detection.

3.5.1. Presicion

The graphs in Fig. 6, 7, and 8 show how the precision of the YOLOv7, YOLOv8, and YOLOv9 models evolve over time. Precision is calculated using validation data. At a confidence point of 0.5, the precision reaches more than 0.8. This shows that the model is able to predict well at that confidence level.

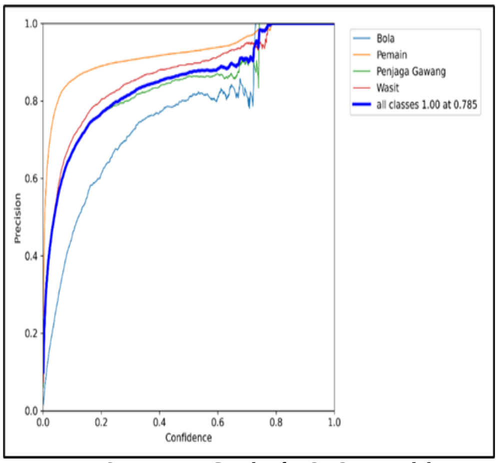

Fig. 6. Precision Graph of YOLOv7 Model

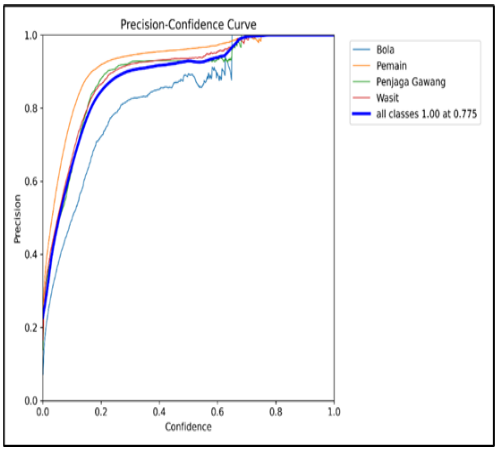

Fig. 7. Precision Graph of YOLOv8 Model

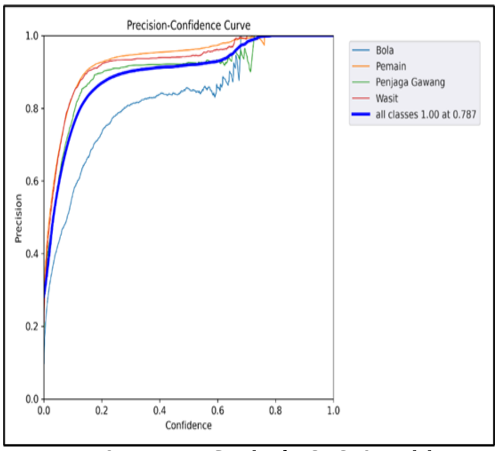

Fig. 8. Precision Graph of YOLOv9 Model

3.5.2. Recall

Recall measures the model's effectiveness in identifying all true positive instances. As the confidence threshold increases, the recall curve typically declines, indicating that as confidence rises, the model may overlook some positive instances. The validation result graphs (refer to Fig. 9, 10, and 11) for the YOLOv7, YOLOv8, and YOLOv9 models demonstrate that, at a confidence level of 0.8, the recall value exceeds 0.8. This implies that all three models consistently generate numerous predictions with high confidence levels.

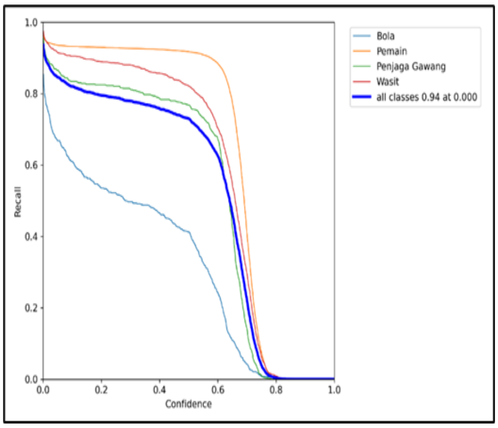

Fig. 9. Recall Graph of YOLOv7 Model

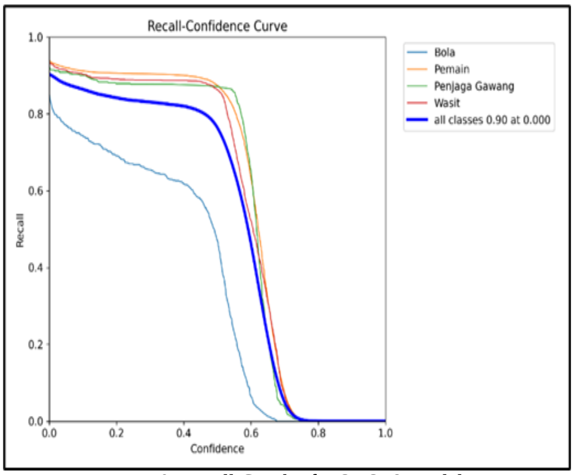

Fig. 10. Recall Graph of YOLOv8 Model

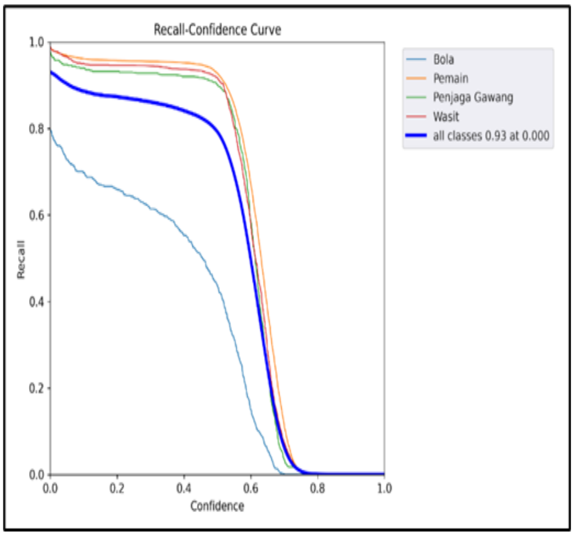

Fig. 11. Recall Graph of YOLOv9 Model

3.5.3. F1-Score

The F1 value is the harmonic mean of the precision and recall values calculated in the validation process, the higher the F1 value, the better the quality of the model prediction. The graphs in Fig. 12, 13, and 14 show the F1 value against the confidence value of the class, for each class has an F1 value above 0.8 at confidence points above 0.2.

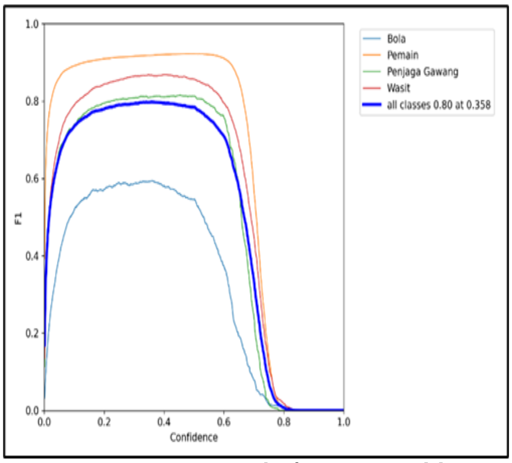

Fig. 12. F1-Score Graph of YOLOv7 Model

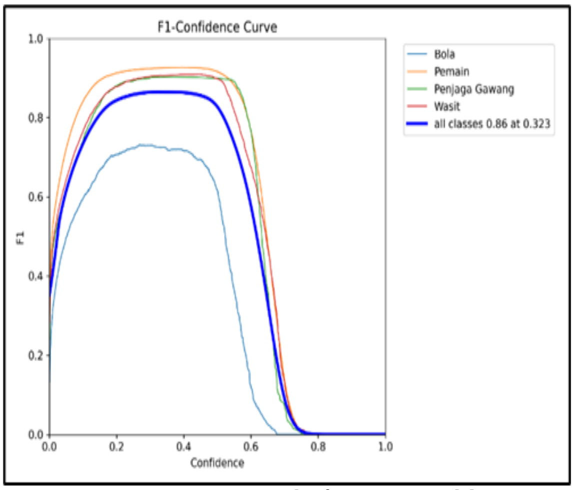

Fig. 13. F1-Score Graph of YOLOv8 Model

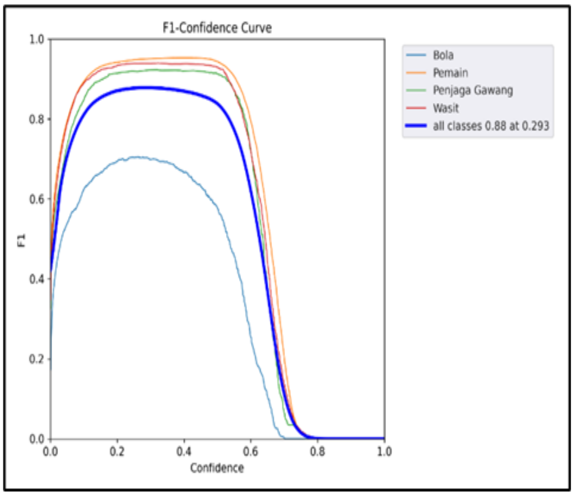

Fig. 14. F1-Score Graph of YOLOv9 Model

3.5.4. Confussion Matrix

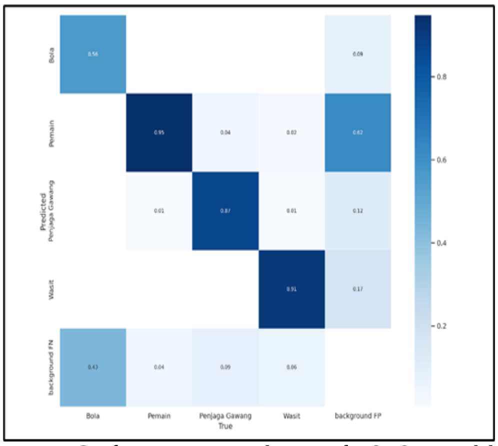
Fig. 15. Confusion Matrix Evaluation of YOLOv7 Model

Figures 15, 16, and 17 display the Confusion Matrix of the models trained using YOLOv7, YOLOv8, and YOLOv9. The Confusion Matrix provides a visualization of how well the model predicts the class correctly or incorrectly for each class. The results of the validation process show that the most difficult class to predict is the ball class, with a rate below 50%, because the ball has a small shape and white color that is similar to many other objects on the field, such as white player shoes, penalty box points, and other objects that have elements of white color. In addition, sometimes other classes are not successfully predicted and are considered as background. Nonetheless, overall the model gave satisfactory results for most of the classes.

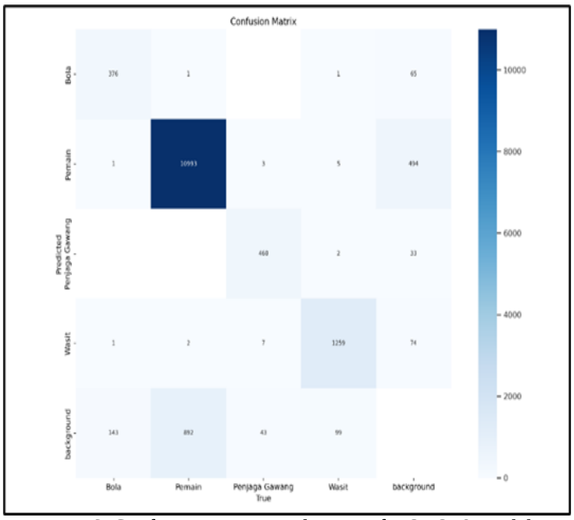

Fig. 16. Confusion Matrix Evaluation of YOLOv8 Model

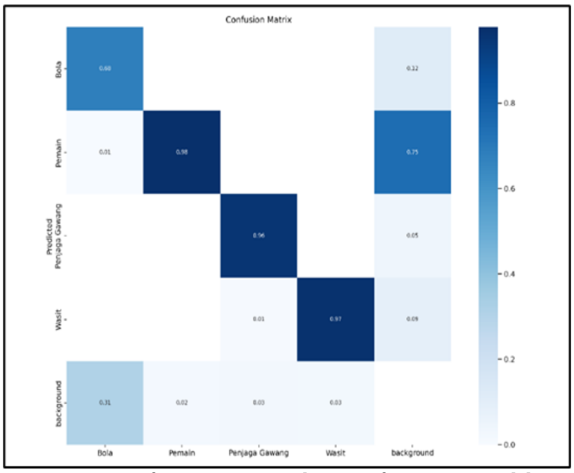

Fig. 17. Confusion Matrix Evaluation of YOLOv9 Model

3.6. Model Evaluation Results

The model evaluation results include F1-Score, mAP (mean Average Precision), recall, and precision values, which help in assessing the overall quality of the model in detecting objects in soccer matches using YOLOv7, YOLOv8, and YOLOv9. The results of the model evaluation for the three models are shown in Table. 4, 5, and 6.

Table. 4. YOLOv7 Model Evaluation Results

Class Name	Precision	Recall	F1-Score	mAP@0.5	Overall Accuracy (mAP@0.5)
Ball	0.758	0.488	0.594	0.523	**77.1%**
Player	0.912	0.925	0.918	0.916	
Goalkeeper	0.825	0.797	0.810	0.778	
Referee	0.87	0.865	0.868	0.867	

Table. 5. YOLOv8 Model Evaluation Results

Class Name	Precision	Recall	F1-Score	mAP@0.5	Overall Accuracy (mAP@0.5)
Ball	0.834	0.677	0.745	0.716	**87.4%**
Player	0.951	0.942	0.948	0.948	
Goalkeeper	0.929	0.92	0.923	0.902	
Referee	0.936	0.934	0.935	0.931	

Table. 6. YOLOv9 Model Evaluation Results

Class Name	Precision	Recall	F1-Score	mAP@0.5	Overall Accuracy (mAP@0.5)
Ball	0.805	0.626	0.704	0.638	**86%**
Player	0.946	0.946	0.946	0.958	
Goalkeeper	0.915	0.915	0.915	0.901	
Referee	0.933	0.945	0.933	0.942	

3.6.1. Evaluation Results of the YOLOv7 Model

The evaluation results of the YOLOv7 model can be seen in Table. 4, which shows the variation of values between different object classes. Overall, the high precision, recall, and F1-score values indicate that the model tends to be good at identifying and predicting objects for the "Player" and "Referee" classes, with the highest F1-score value in the "Player" class. However, for the "Ball" and "Goalkeeper" classes, the precision, recall, and F1-score values tend to be lower, indicating challenges in object detection and prediction for these classes. Nonetheless, the overall mAP@0.5 value of the model remains relatively good at predicting, reaching 77.1%, indicating that the model can generally perform object detection well.

3.6.2. Evaluation Results of the YOLOv8 Model

The evaluation results in Table. 5 of the YOLOv8 model illustrate the variation in performance in detecting objects for each class. Overall, high precision, recall, and F1-Score values indicate good quality predictions for the "Player" and "Referee" classes, with the highest F1-Score value found in the "Player" class. However, for the "Ball" and "Goalkeeper" classes, the precision, recall, and F1-Score values tend to be lower, indicating challenges in detecting and predicting objects for these classes. Even so, the overall mAP@0.5 value of the model remains relatively high, reaching 87.4%, indicating the model's ability to detect objects well.

3.6.3. Evaluation Results of the YOLOv9 Model

The evaluation results in Table. 6 of the YOLOv9 model show the variation in performance in detecting objects for each class. Overall, good predictions are seen in the "Player" and "Referee" classes, with the highest F1-score value in the "Player" class. However, the "Ball" and "Goalkeeper" classes had lower predictions, indicating challenges in detecting objects for those classes. The overall model has a good ability to detect objects, with mAP@0.5 values reaching 86%.

3.7. Object Detection Results

Object detection results using video and the YOLOv7, YOLOv8, and YOLOv9 models involve a continuous object recognition process on each video frame. Each frame is individually analyzed by the model, which then identifies the objects contained in it. The detected objects are then marked with a bounding box and a corresponding label. This process continues for every frame in the video, allowing the model to recognize objects that move or change position. The following are the results of object detection on videos using the YOLOv7, YOLOv8, and YOLOv9 models:

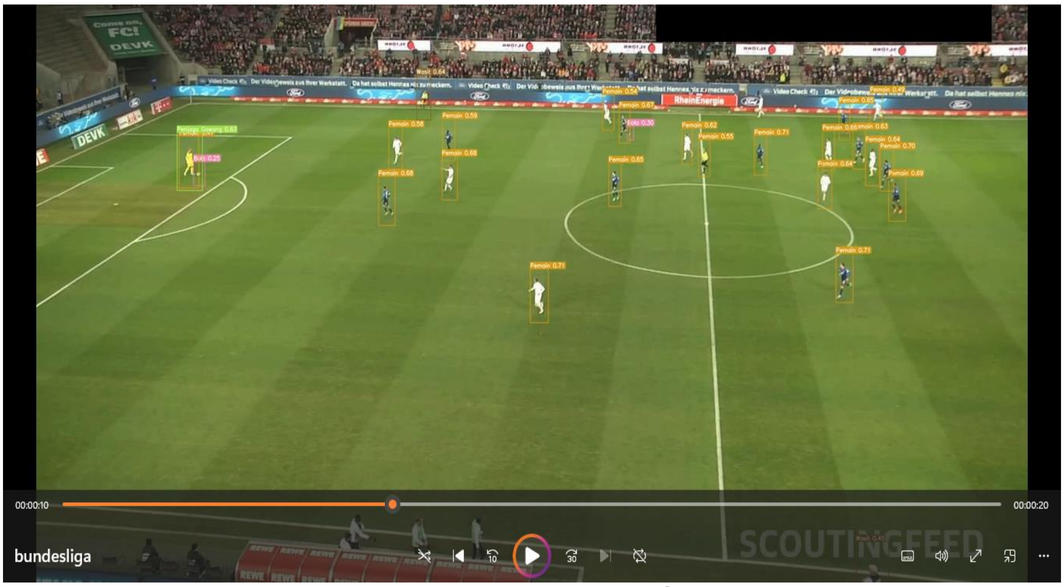

Fig. 18. YOLOv7 Detection Results

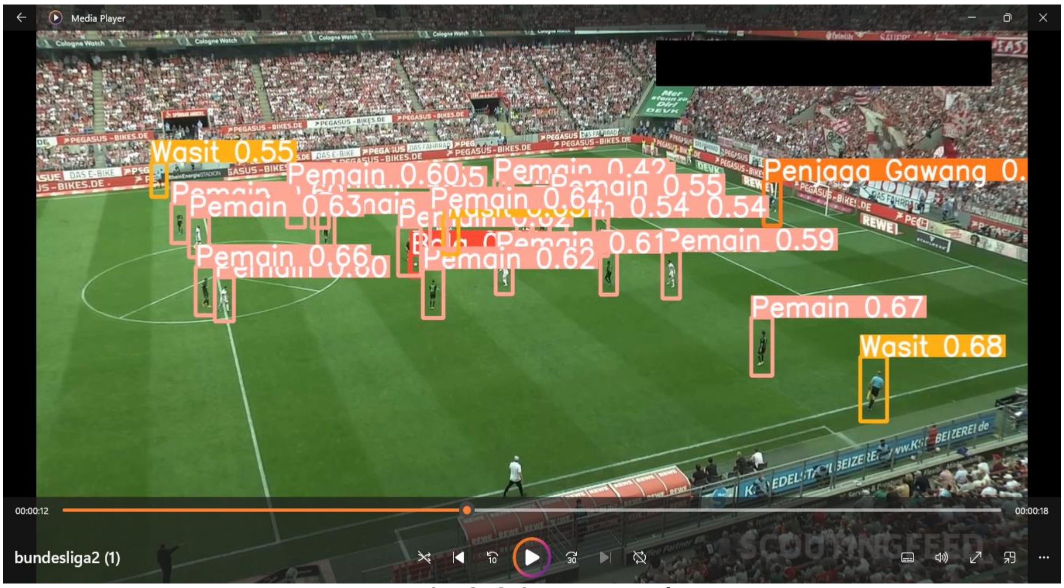

Fig. 19. YOLOv8 Detection Results

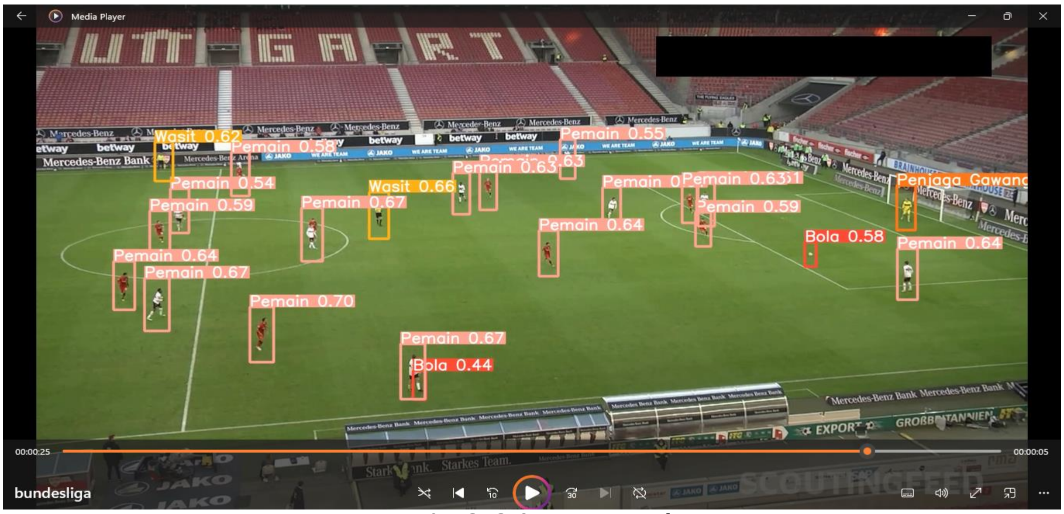

Fig. 20. YOLOv9 Detection Results

3.7.1. YOLOv7 Object Detection Result

The detection results using the YOLOv7 model on the video are shown in Fig. 18. Although the model is able to recognize most of the objects in the football match, there are still errors in prediction, especially in ball identification. For example, the model often predicts a small white object, such as a player's shoe or sock, as the ball. However, overall, the prediction performance is quite good, especially in recognizing objects other than the ball.

3.7.2. YOLOv8 Object Detection Result

The detection results using the YOLOv8 model on the video can be seen in Fig. 19. Overall, it performs better than YOLOv7 and YOLOv9 in predicting and identifying objects. Although the model succeeds in recognizing many objects with high accuracy, the prediction errors mainly occur in ball identification. But the prediction error is not as much as in the YOLOv7 and YOLOv9 models.

3.7.3. YOLOv9 Object Detection Result

The detection results using the YOLOv9 model on the video shown in Fig. 20 show that the model is close to the performance of YOLOv8 but still not optimal. There are still challenges especially in the ball class, similar to what happened with YOLOv7. Overall, YOLOv9 performs quite well in predicting and identifying objects.

3.8. Model Performance Comparison

The performance comparison of the YOLOv7, YOLOv8, and YOLOv9 models illustrates the detailed evaluation of each model in detecting objects in soccer matches. This evaluation includes an analysis of the detection performance for each object class as well as a comparison of the overall performance between models.

Table. 7. Model Performance Comparison

Model	Class Name	Presicion	Recall	F-1 Score	mAP @0.5	Overall Accuracy (mAP@0.5)
YOLOv7	Ball	0.758	0.488	0.594	0.523	**77.1%**
	Player	0.912	0.925	0.918	0.916	
	Goalkeeper	0.825	0.797	0.810	0.778	
	Referee	0.87	0.865	0.868	0.867	
YOLOv8	Ball	0.834	0.677	0.745	0.716	**87.4%**
	Player	0.946	0.946	0.946	0.958	
	Goalkeeper	0.915	0.915	0.915	0.901	
	Referee	0.933	0.945	0.933	0.942	
YOLOv9	Ball	0.805	0.626	0.704	0.638	**86%**
	Player	0.946	0.946	0.946	0.958	
	Goalkeeper	0.915	0.915	0.915	0.901	
	Referee	0.933	0.945	0.933	0.942	

The comparison results of YOLOv7, YOLOv8, and YOLOv9 models in Table. 7 in detecting objects show that YOLOv8 performs better than the other two models. YOLOv8 has higher precision, recall, and F1-score values, especially in the “Ball” class. YOLOv8 has an overall accuracy (mAP@0.5) of 87.4%, which is higher than YOLOv9 (86.0%) and YOLOv7 (77.1%). YOLOv9 performs similarly to YOLOv8, but the higher mAP@0.5 value of YOLOv8 indicates that this model tends to excel at detecting objects with varying confidence levels. Both YOLOv8 and YOLOv9 show significant improvement over YOLOv7 in the overall object detection performance. Therefore, based on these results, YOLOv8 can be considered as the model that is close to the best performance in detecting objects in the dataset used.

4. DISCUSSION

In this study, YOLOv8 showed the best performance in identifying objects in soccer matches. YOLOv8 has higher precision, recall, and F1-Score values especially in the “Ball” class, with values of 83.4%, 67.7%, and 74.5%, respectively. In addition, YOLOv8 achieved an overall accuracy (mAP@0.5) of 87.4%, which is higher than YOLOv9 (86.0%) and YOLOv7 (77.1%).

Patel and Kamdar [14] applied the YOLO algorithm to sports matches, focusing on detecting objects from hockey matches, particularly players and balls. The primary objective of their study was to enhance the overall average object detection accuracy of the utilized model. They utilized a total dataset consisting of 1119 annotated images across 4 classes. The initial experiment involved 100 epoch iterations, yielding an average accuracy of 88.9%. Subsequently, the second experiment, comprising 200 epoch iterations, achieved an accuracy score of 91.2%, while the third experiment utilized 300 epoch iterations, resulting in an accuracy score of 91.3%.

A comparison between this study and the previous study conducted by Patel and Kamdar [14] reveals some significant differences, especially in terms of dataset usage and analysis coverage. The previous study by Patel and Kamdar only used a dataset from one match, while this study covers various matches with more diverse objects, including players, referees, and goalkeepers wearing different clothes. This diversity has an impact on the model's performance in detecting objects, as the model must be able to recognize different types of objects with diverse visual attributes. Therefore, even though the dataset used in this study is more diverse, the object detection accuracy may be lower compared to previous studies that only used datasets from a single match. This study adds to the understanding of model performance by utilizing diverse datasets, allowing measurement of whether diversity impacts detection accuracy.

5. CONCLUSION

Based on the results of the model performance comparison, it was found that YOLOv8 showed the best performance in identifying objects in soccer matches. YOLOv8 has higher precision, recall, and f1-score values especially in the “Soccer” class, with values of 83.4%, 67.7%, and 74.5% respectively. In addition, YOLOv8 achieved an overall accuracy (mAP@0.5) of 87.4%, which is higher than that of YOLOv9 (86.0%) and YOLOv7 (77.1%). Evaluating the effect of parameter adjustment on the object detection performance of each YOLO version, it was found that YOLOv8 showed improved performance by using a larger batch size (16) and higher image resolution (1080 pixels) compared to YOLOv7 and YOLOv9. This shows that parameter adjustment can affect the performance of object detection and proper parameter selection can improve the performance of the model.

REFERENCES

[1] Y. Wang and J. Doherty, “R. Van Dyck–Moving object tracking in video,” in *Proc. Conference on Information Sciences and Systems*, 2000.

[2] C. Zhu, R. Shao, X. Zhang, S. Gao, and B. Li, “Application of virtual reality based on computer vision in sports posture correction,” *Wirel Commun Mob Comput*, vol. 2022, pp. 1–15, 2022.

[3] M. Burić, M. Pobar, and M. Ivašić-Kos, “Object detection in sports videos,” in *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, IEEE, 2018, pp. 1034–1039.

[4] G. Lovett, “World football report 2018 2,” 2018.

[5] A. Ali, “Measuring soccer skill performance: a review,” *Scand J Med Sci Sports*, vol. 21, no. 2, pp. 170–183, 2011.

[6] C. Wright, C. Carling, C. Lawlor, and D. Collins, “Elite football player engagement with performance analysis,” *Int J Perform Anal Sport*, vol. 16, no. 3, pp. 1007–1032, 2016.

[7] F. R. Goes *et al.*, “Unlocking the potential of big data to support tactical performance analysis in professional soccer: A systematic review,” *Eur J Sport Sci*, vol. 21, no. 4, pp. 481–496, 2021.

[8] A. G. Prawiyogi, M. Hammet, and A. Williams, “Visualization Guides in the understanding of theoretical material in lectures,” *International Journal of Cyber and IT Service Management*, vol. 3, no. 1, pp. 54–60, 2023.

[9] P. Salvo, A. Pingitore, A. Barbini, and F. Di Francesco, “A wearable sweat rate sensor to monitor the athletes’ performance during training,” *Sci Sports*, vol. 33, no. 2, pp. e51–e58, 2018.

[10] M. Stein *et al.*, “Bring it to the pitch: Combining video and movement data to enhance team sport analysis,” *IEEE Trans Vis Comput Graph*, vol. 24, no. 1, pp. 13–22, 2017.

[11] J. Komorowski, G. Kurzejamski, and G. Sarwas, “Footandball: Integrated player and

ball detector," *arXiv preprint arXiv:1912.05445*, 2019.

[12] I. McKeown, K. Taylor-McKeown, C. Woods, and N. Ball, "Athletic ability assessment: a movement assessment protocol for athletes," *Int J Sports Phys Ther*, vol. 9, no. 7, p. 862, 2014.

[13] K. Rangasamy, M. A. As'ari, N. A. Rahmad, N. F. Ghazali, and S. Ismail, "Deep learning in sport video analysis: a review," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 4, pp. 1926–1933, 2020.

[14] S. H. Patel and D. Kamdar, "Object detection in hockey sport video via pretrained yolov3 based deep learning model.," *ICTACT Journal on Image & Video Processing*, vol. 13, no. 3, 2023.

[15] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[16] T. Wang and T. Li, "Deep learning-based football player detection in videos," *Comput Intell Neurosci*, vol. 2022, 2022.

[17] B. T. Naik and M. F. Hashmi, "Ball and player detection & tracking in soccer videos using improved yolov3 model," 2021.

[18] B. T. Naik, M. F. Hashmi, Z. W. Geem, and N. D. Bokde, "Deep player-track: player and referee tracking with jersey color recognition in soccer," *IEEE Access*, vol. 10, pp. 32494–32509, 2022.

[19] M. Husein, K. C. Ciptady, R. Suryadi, and M. Harahap, "Detecting and tracking player in football videos using two-stage mask r-cnn Approach," in *IAIC International Conference Series*, 2023, pp. 132–138.

[20] Hiemann, T. Kautz, T. Zottmann, and M. Hlawitschka, "Enhancement of speed and accuracy trade-off for sports ball detection in videos—finding fast moving, small objects in real time," *Sensors*, vol. 21, no. 9, p. 3214, 2021.

[21] M. Buric, M. Ivasic-Kos, and M. Pobar, "Player tracking in sports videos," in *2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)*, IEEE, 2019, pp. 334–340.

[22] K. Kausalya, "Object recognition in soccer sports videos," in *2022 Smart Technologies, Communication and Robotics (STCR)*, IEEE, 2022, pp. 1–5.

[23] Q. Chen and C. Poullis, "Tracking and identification of ice hockey players," in *International Conference on Computer Vision Systems*, Springer, 2023, pp. 3–16.

[24] Cioppa, A. Deliege, N. U. Huda, R. Gade, M. Van Droogenbroeck, and T. B. Moeslund, "Multimodal and multiview distillation for real-time player detection on a football field," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020, pp. 880–881.

[25] Y. Liu, L. G. Hafemann, M. Jamieson, and M. Javan, "Detecting and matching related objects with one proposal multiple predictions," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 4520–4527.

[26] J. A. Vicente-Martínez, M. Márquez-Olivera, A. García-Aliaga, and V. Hernández-Herrera, "Adaptation of yolov7 and yolov7 tiny for soccer-ball multi-detection with deep sort for tracking by semi-supervised system," *Sensors*, vol. 23, no. 21, p. 8693, 2023.