Deteksi Pelanggaran Sepak Bola Berbasis YOLOv11n dengan Dataset Roboflow

1st Raka Muhammad Sixmarch
Program Studi S1 Informatika
Direktorat Kampus Surabaya
Universitas Telkom
Surabaya, Indonesia
rakamsixmarch@student.telkomuniversity.ac.id

2nd Ahmad Wali Satria Bahari Johan, S.ST., M.Kom.
Program Studi S1 Informatika
Direktorat Kampus Surabaya
Universitas Telkom
Surabaya, Indonesia
ahmadsatria@telkomuniversity.ac.id

3rd Dr. Tanzilal Mustaqim, S.Kom., M.Kom.
Program Studi S1 Informatika
Direktorat Kampus Surabaya
Universitas Telkom
Surabaya, Indonesia
tanzilal@telkomuniversity.ac.id

Abstrak **--** Pelanggaran dalam pertandingan sepak bola sering terjadi secara cepat, dinamis, dan tumpang tindih antar pemain sehingga sulit diamati secara konsisten oleh wasit. Penelitian ini bertujuan mengimplementasikan dan mengevaluasi model YOLOv11n untuk mendeteksi pelanggaran sepak bola berbasis citra menggunakan dataset publik Roboflow. Dataset awal berjumlah 2.242 citra, kemudian setelah proses ekspor, augmentasi, dan penggabungan subset diperoleh 5.380 citra yang terdiri atas tiga kelas, yaitu foul, fouling_player, dan victim. Dataset dibagi menjadi 4.304 citra sebagai train pool dan 1.076 citra sebagai test set. Pelatihan dilakukan menggunakan skema 5-Fold Cross Validation selama 50 epoch pada setiap fold dengan bobot awal yolo11n.pt, ukuran citra 640 x 640, batch size 8, dan GPU NVIDIA A100. Evaluasi menggunakan precision, recall, mAP50, dan mAP50-95. Hasil rata-rata lima fold memperoleh precision 95,07%, recall 96,10%, mAP50 98,03%, dan mAP50-95 77,94%. Fold 1 menjadi model terbaik dengan mAP50-95 sebesar 78,16% dan menghasilkan mAP50-95 test set sebesar 77,70%. Hasil tersebut menunjukkan bahwa YOLOv11n mampu mendeteksi kelas pelanggaran secara konsisten sebagai dasar sistem pendukung analisis visual pertandingan.

Kata kunci-- deteksi pelanggaran sepak bola; YOLOv11n; Roboflow; computer vision; 5-Fold Cross Validation; mAP.

I. PENDAHULUAN

Sepak bola merupakan olahraga dengan intensitas gerak tinggi sehingga keputusan wasit sering diambil dalam kondisi visual yang cepat, padat, dan kompleks. Situasi seperti kontak fisik antar pemain, perubahan arah gerak, sudut kamera yang terbatas, serta occlusion dapat menyebabkan kejadian pelanggaran sulit dianalisis secara objektif. Teknologi Video Assistant Referee (VAR) telah membantu meningkatkan akurasi keputusan, tetapi proses peninjauan masih memerlukan interpretasi manusia dan dapat menambah durasi pertandingan [1].

Perkembangan computer vision membuka peluang untuk membantu analisis pelanggaran secara lebih otomatis. Berbagai penelitian telah menggunakan pendekatan visual untuk menganalisis kejadian dalam pertandingan, baik melalui estimasi pose pemain, klasifikasi aksi, maupun deteksi objek. Fang et al. [2] menggunakan estimasi pose untuk memprediksi foul dari video siaran, sedangkan Shao dan He [3] mengembangkan VAR-YOLOv8s untuk mendukung deteksi foul berbasis Internet of Things. Rabee et al. [4] membandingkan beberapa arsitektur deep learning untuk deteksi pelanggaran sepak bola otomatis, sehingga menunjukkan bahwa pendekatan berbasis pembelajaran mendalam relevan untuk mendukung analisis foul. Wang dan Li [5] mengembangkan metode pembagian fitur aksi foul berbasis salience image semantics. Temuan-temuan tersebut menunjukkan bahwa pendekatan visual dapat digunakan untuk membaca pola interaksi fisik antar pemain.

Meskipun demikian, sebagian penelitian sebelumnya belum secara langsung memfokuskan lokalisasi objek berbasis bounding box pada kelas-kelas yang berkaitan dengan pelanggaran, seperti foul, fouling_player, dan victim. Selain itu, penerapan model ringan generasi terbaru masih perlu dikaji agar sistem dapat diarahkan pada pemrosesan yang lebih efisien. Penelitian ini mengisi ruang tersebut dengan menerapkan YOLOv11n sebagai model deteksi objek yang ringan, menggunakan dataset Roboflow yang telah dianotasi, dan mengevaluasi performa model melalui skema 5-Fold Cross Validation serta test set terpisah.

Tujuan penelitian ini adalah menerapkan YOLOv11n untuk mendeteksi pelanggaran sepak bola berbasis citra dan menganalisis hasil evaluasinya berdasarkan precision, recall, mAP50, dan mAP50-95. Sistem yang dikembangkan tidak dimaksudkan menggantikan kewenangan wasit, tetapi menjadi dasar decision support yang mampu menampilkan hasil deteksi visual secara cepat, konsisten, dan mudah dibaca.

II. KAJIAN TEORI

A. Visi Komputer

Visi komputer merupakan cabang kecerdasan buatan yang memungkinkan komputer memproses, menganalisis, dan menafsirkan informasi visual dari citra atau video. Dalam konteks pertandingan sepak bola, visi komputer digunakan untuk membaca pola gerak pemain, lokasi objek, serta interaksi fisik yang berpotensi berkaitan dengan pelanggaran. Teknologi ini relevan karena pertandingan sepak bola memiliki kondisi visual dinamis, seperti perubahan pencahayaan, variasi sudut kamera, objek yang saling tumpang tindih, dan pergerakan pemain yang cepat [6]

Pada penelitian ini, visi komputer digunakan untuk mengubah citra pertandingan menjadi informasi berbentuk bounding box, label kelas, dan confidence score. Hasil tersebut dapat membantu pengguna memahami posisi kejadian pelanggaran, pemain pelaku, dan pemain korban secara visual. Dengan demikian, luaran model tidak hanya berupa keputusan ada atau tidak adanya pelanggaran, tetapi juga lokasi objek yang menjadi dasar pembacaan kejadian.

B. Deep Learning dan Deteksi Objek

Deep learning merupakan pendekatan pembelajaran mesin yang menggunakan jaringan saraf berlapis untuk mempelajari representasi fitur dari data. Pada citra, fitur yang dipelajari dapat berupa tepi, tekstur, bentuk tubuh, hingga pola interaksi antar objek. Deteksi objek merupakan tugas yang menggabungkan klasifikasi dan lokalisasi, yaitu

menentukan jenis objek sekaligus posisi objek dalam citra melalui bounding box.

Metode deteksi objek modern terbagi menjadi two-stage detector dan one-stage detector. Two-stage detector melakukan proses region proposal terlebih dahulu sebelum klasifikasi, sedangkan one-stage detector melakukan prediksi kelas dan bounding box secara langsung. YOLO termasuk one-stage detector sehingga sesuai untuk kebutuhan analisis pertandingan yang membutuhkan pemrosesan cepat [7]

C. YOLOv11n

You Only Look Once versi 11 nano (YOLOv11n) merupakan varian ringan dari keluarga YOLOv11. Varian nano dipilih karena ukuran model lebih kecil dan kebutuhan komputasi lebih rendah dibandingkan varian yang lebih besar, sehingga lebih sesuai untuk pengembangan sistem deteksi yang diarahkan pada efisiensi. Dalam penelitian ini, model menggunakan bobot awal yolo11n.pt yang telah melalui pretraining pada dataset COCO. Pretraining tersebut memberikan pengetahuan awal terhadap objek umum, seperti manusia dan objek lapangan, sebelum model di-fine-tuning pada dataset pelanggaran sepak bola [8], [9]

Pemilihan YOLOv11n didasarkan pada kebutuhan untuk mendeteksi tiga kelas yang memiliki hubungan spasial erat dalam satu citra, yaitu foul, fouling_player, dan victim. Ketiga kelas tersebut sering muncul berdekatan atau saling menutupi sehingga model perlu memiliki kemampuan lokalisasi yang cukup baik. Penggunaan varian nano juga mendukung arah pengembangan sistem yang lebih ringan untuk integrasi antarmuka pengguna.

D. 5-Fold Cross Validation

K-Fold Cross Validation adalah skema evaluasi yang membagi data latih menjadi beberapa bagian, kemudian setiap bagian digunakan secara bergantian sebagai data validasi. Penelitian ini menggunakan lima fold sehingga train pool dibagi menjadi lima subset. Pada setiap iterasi, satu subset berperan sebagai validation set dan empat subset lainnya menjadi training set. Hasil dari setiap fold kemudian dihitung rata-rata dan standar deviasinya untuk membaca kestabilan performa model [10]

Skema ini dipilih karena evaluasi dengan satu kali pembagian data dapat menghasilkan estimasi performa yang bergantung pada distribusi data tertentu. Dengan 5-Fold Cross Validation, model diuji pada variasi data validasi yang lebih beragam. Fold terbaik dipilih berdasarkan mAP50-95 karena metrik tersebut lebih ketat dalam menilai kualitas lokalisasi bounding box.

E. Metrik Evaluasi Deteksi Objek

Evaluasi model dilakukan menggunakan precision, recall, Average Precision (AP), mAP50, dan mAP50-95. Precision mengukur ketepatan prediksi positif, sedangkan recall mengukur kemampuan model menemukan objek yang seharusnya terdeteksi. AP menggambarkan luas area di bawah kurva precision-recall untuk satu kelas. mAP50 menghitung rata-rata AP pada ambang Intersection over Union (IoU) sebesar 0,50, sedangkan mAP50-95 menghitung rata-rata AP pada ambang IoU 0,50 hingga 0,95 dengan kenaikan 0,05 [11], [12]

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

$$AP = \int_0^1 p\,(r)\,dr \quad (3)$$

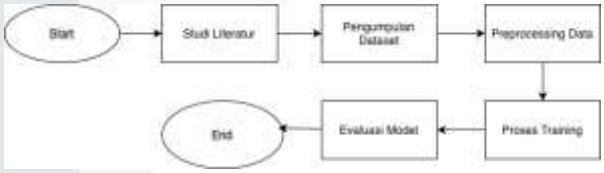
$$\text{mAP50} = \frac{1}{K}\sum_{k=1}^{K} A\,P_k^{0.50} \quad (4)$$

$$\text{mAP50-95} = \frac{1}{|T|K}\sum_{t\in T}\sum_{k=1}^{K} A\,P_k^{t} \quad (5)$$

TP adalah true positive, FP adalah false positive, FN adalah false negative, K adalah jumlah kelas, k adalah indeks kelas, dan T adalah himpunan ambang IoU dari 0,50 sampai 0,95 dengan kenaikan 0,05. Precision yang tinggi menunjukkan prediksi model cenderung benar, sedangkan recall yang tinggi menunjukkan objek yang seharusnya terdeteksi jarang terlewat. mAP50-95 digunakan sebagai metrik utama pemilihan fold terbaik karena lebih ketat dibandingkan mAP50.

III. METODE

Penelitian ini menggunakan pendekatan eksperimen berbasis deteksi objek. Tahapan penelitian dimulai dari studi literatur, pengumpulan dataset, preprocessing, pembagian data, pelatihan YOLOv11n, evaluasi menggunakan 5-Fold Cross Validation, pemilihan model terbaik, pengujian pada test set, dan implementasi hasil ke rancangan antarmuka web. Alur umum penelitian ditunjukkan pada Gambar 1.



GAMBAR 1
(ALUR UMUM PENELITIAN)

A. Dataset dan Kelas Deteksi

Dataset yang digunakan berasal dari Roboflow dalam format anotasi YOLO. Data awal berjumlah 2.242 raw image, kemudian setelah proses ekspor, augmentasi, resize, dan penggabungan subset diperoleh 5.380 citra. Dataset terdiri atas tiga kelas, yaitu foul sebagai kejadian pelanggaran, fouling_player sebagai pemain pelaku pelanggaran, dan victim sebagai pemain yang menjadi korban pelanggaran. Contoh anotasi bounding box disajikan pada Gambar 2.



GAMBAR 2
(CONTOH ANOTASI BOUNDING BOX PADA CITRA PELANGGARAN)

TABEL I
(RINGKASAN DATASET)

Komponen	Jumlah / Keterangan
Raw image awal	2.242 citra
Total data eksperimen	5.380 citra
Train pool	4.304 citra
Test set	1.076 citra
Kelas	foul, fouling_player, victim

B. Preprocessing Data

Preprocessing dilakukan untuk menyesuaikan dataset dengan kebutuhan pelatihan YOLO. Setiap citra dan label disusun dalam struktur folder sesuai format YOLO, kemudian data diatur ke dalam file konfigurasi data.yaml. Proses resize dilakukan ke ukuran 640 x 640 piksel agar sesuai dengan konfigurasi input model. Selain itu, augmentasi digunakan untuk meningkatkan variasi data dan membantu model lebih kuat terhadap perubahan visual pada citra pertandingan.

Augmentasi yang digunakan mencakup transformasi visual seperti horizontal flip, shear, perubahan brightness, blur, median blur, grayscale, serta CLAHE. Augmentasi ini penting karena citra pertandingan sepak bola dapat memiliki pencahayaan, sudut pandang, dan kualitas visual yang berbeda-beda. Contoh hasil augmentasi dataset disajikan pada Gambar 3.



GAMBAR 3
(CONTOH AUGMENTASI DATASET ROBOFLOW)

C. Konfigurasi Pelatihan

Pelatihan model dilakukan dengan YOLOv11n menggunakan bobot awal yolo11n.pt. Proses training dijalankan di Google Colab dengan GPU NVIDIA A100 dan framework Ultralytics YOLO versi 8.3.224. Setiap fold dilatih selama 50 epoch dengan ukuran citra 640 x 640 piksel dan batch size 8. Konfigurasi pelatihan diringkas pada Tabel II.

TABEL II
(PARAMETER PELATIHAN YOLOV11N)

Parameter	Nilai
Model	YOLOv11n
Pre-trained weight	yolo11n.pt
Epoch	50 setiap fold
Image size	640 x 640
Batch size	8
Workers	2
Patience	50
Device	GPU NVIDIA A100
Framework	Ultralytics 8.3.224
Skema validasi	5-Fold Cross Validation

D. Skema Evaluasi

Dataset dibagi menjadi train pool dan test set. Train pool digunakan untuk 5-Fold Cross Validation, sedangkan test set dipisahkan sejak awal dan hanya digunakan untuk evaluasi akhir model terbaik. Pada setiap fold, model divalidasi menggunakan bagian data yang berbeda. Hasil setiap fold dibandingkan untuk menentukan model terbaik, kemudian hasil rata-rata digunakan untuk menilai kestabilan performa secara keseluruhan.

Pemilihan model terbaik didasarkan pada mAP50-95 tertinggi karena metrik ini menilai performa deteksi pada rentang IoU yang lebih luas. Setelah model terbaik dipilih, model tersebut diuji pada test set yang belum pernah digunakan pada tahap training dan validasi. Dengan skema ini, evaluasi akhir diharapkan lebih objektif dan tidak bias terhadap data validasi.

IV. HASIL DAN PEMBAHASAN

A. Hasil Rata-rata Evaluasi K-Fold

Hasil evaluasi menunjukkan bahwa model YOLOv11n memiliki performa yang stabil pada lima pembagian data validasi. Rata-rata precision sebesar 95,07% menunjukkan bahwa mayoritas prediksi model sesuai dengan anotasi ground truth. Recall sebesar 96,10% menunjukkan bahwa model mampu menemukan sebagian besar objek pelanggaran yang seharusnya terdeteksi. Nilai mAP50 sebesar 98,03% menandakan deteksi sangat baik pada ambang IoU 0,50, sedangkan mAP50-95 sebesar 77,94% menunjukkan bahwa performa model tetap cukup baik ketika kualitas lokalisasi dievaluasi pada ambang yang lebih ketat.

TABEL III
(RATA-RATA HASIL EVALUASI 5-FOLD)

Metrik	Rata-rata (%)	Std. Deviasi
Precision	95,07	0,16
Recall	96,10	0,39
mAP50	98,03	0,36
mAP50-95	77,94	0,19

Standar deviasi yang relatif kecil menunjukkan bahwa variasi performa antar fold tidak terlalu besar. Nilai standar deviasi precision sebesar 0,16 dan mAP50-95 sebesar 0,19 mengindikasikan bahwa model tidak hanya baik pada satu fold, tetapi cukup konsisten pada beberapa pembagian data. Hal ini penting karena dataset sepak bola memiliki variasi visual yang tinggi, terutama pada posisi pemain dan bentuk interaksi fisik.

B. Peringkat Fold dan Model Terbaik

Perbandingan antar fold menunjukkan bahwa Fold 1 memperoleh mAP50-95 tertinggi sebesar 78,16% sehingga dipilih sebagai model terbaik. Fold 4 berada pada peringkat kedua dengan mAP50-95 sebesar 78,07%, diikuti Fold 2, Fold 5, dan Fold 3. Walaupun Fold 3 memiliki precision dan recall tertinggi, nilai mAP50-95-nya lebih rendah dibandingkan Fold 1. Dengan demikian, Fold 1 dipilih karena memberikan kualitas lokalisasi paling baik pada evaluasi yang lebih ketat.

TABEL IV
(PERINGKAT HASIL EVALUASI FOLD)

Rank	Fold	Prec.	Rec.	mAP50	mAP50-95
1	Fold 1	95,16	95,92	97,40	78,16
2	Fold 4	95,00	95,96	98,32	78,07
3	Fold 2	95,01	96,13	98,04	77,95
4	Fold 5	94,89	95,75	98,17	77,83
5	Fold 3	95,31	96,75	98,20	77,68

C. Evaluasi Test Set

Model terbaik dari Fold 1 kemudian diuji pada 1.076 citra test set. Hasil pengujian memperoleh precision 95,60%, recall 95,90%, mAP50 98,00%, dan mAP50-95 77,70%. Nilai tersebut mendekati hasil validasi K-Fold, sehingga model dapat mempertahankan performa pada data yang tidak digunakan dalam pelatihan. Hal ini

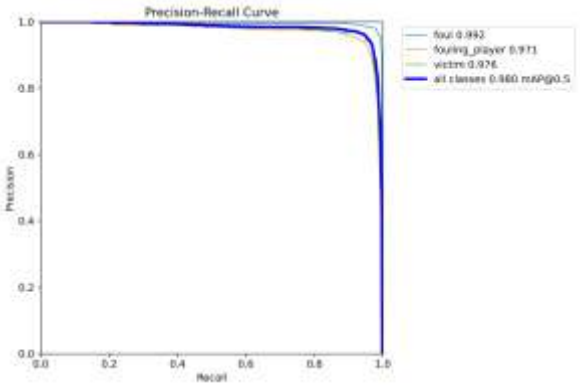
mengindikasikan bahwa YOLOv11n mampu mempelajari pola visual kelas foul, fouling_player, dan victim secara cukup general.

TABEL V
(HASIL EVALUASI TEST SET MODEL TERBAIK)

Metrik	Nilai (%)
Precision	95,60
Recall	95,90
mAP50	98,00
mAP50-95	77,70

D. Analisis Precision-Recall Curve

Precision-Recall Curve model terbaik menunjukkan hubungan antara ketepatan prediksi dan kemampuan model menemukan objek pada berbagai nilai confidence. Kurva yang berada dekat area kanan atas menunjukkan bahwa model mampu mempertahankan precision tinggi ketika recall meningkat. Kondisi ini memperlihatkan bahwa model tidak hanya akurat dalam menghasilkan prediksi, tetapi juga mampu menemukan objek pelanggaran dengan tingkat kelalaian yang rendah. Hasil kurva Precision-Recall model terbaik ditunjukkan pada Gambar 4.

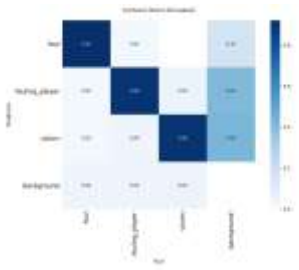


GAMBAR 4
(PRECISION-RECALL CURVE MODEL TERBAIK)

E. Analisis Confusion Matrix

Confusion matrix ternormalisasi digunakan untuk membaca distribusi prediksi model terbaik terhadap setiap kelas seperti yang ditujukan pada gambar 5. Secara umum, prediksi paling banyak berada pada diagonal utama sehingga model mampu membedakan kelas foul, fouling_player, dan victim. Pada matriks model terbaik, nilai diagonal untuk kelas foul, fouling_player, dan victim berada pada tingkat yang relatif baik. Hal tersebut menunjukkan bahwa model dapat mengenali kejadian pelanggaran, pemain pelaku, dan pemain korban secara cukup terpisah meskipun ketiga objek sering muncul dalam posisi berdekatan.

Kesalahan prediksi masih muncul pada bagian background, terutama ketika objek pemain mengalami tumpang tindih, perubahan pose tubuh, atau berada pada area visual yang kurang jelas. Kondisi tersebut wajar pada data pertandingan sepak bola karena kontak fisik sering terjadi dalam ruang yang sempit dan dinamis. Dengan demikian, peningkatan dataset primer dengan variasi kamera dan kualitas video yang lebih beragam dapat menjadi langkah pengembangan berikutnya.



GAMBAR 5
(CONFUSION MATRIX TERNORMALISASI MODEL TERBAIK)

Perbedaan antara mAP50 dan mAP50-95 juga memberikan informasi penting. Nilai mAP50 yang tinggi menunjukkan bahwa model mampu menemukan objek pada kriteria IoU 0,50, sedangkan nilai mAP50-95 yang lebih rendah menunjukkan bahwa akurasi lokalisasi akan semakin menurun ketika ambang IoU diperketat. Dengan kata lain, model sudah baik dalam mendeteksi objek, tetapi peningkatan kualitas bounding box masih dapat dilakukan melalui penambahan data, perbaikan anotasi, dan pengujian pada video pertandingan yang lebih bervariasi.

F. Implementasi Antarmuka

Model terbaik diintegrasikan ke rancangan Web UI sebagai sarana visualisasi hasil deteksi. Laman input digunakan untuk mengunggah citra atau video pertandingan dan mengatur confidence threshold. Pengaturan threshold diperlukan karena nilai confidence memengaruhi jumlah deteksi yang ditampilkan. Threshold yang terlalu rendah dapat meningkatkan false positive, sedangkan threshold yang terlalu tinggi dapat membuat objek pelanggaran tertentu tidak terdeteksi. Adapun web UI input ditujukan pada gambar 6.



GAMBAR 6
(TAMPILAN INPUT WEB UI)

Laman output menampilkan bounding box, label kelas, confidence score, tabel hasil deteksi, serta ringkasan jumlah objek pada setiap kelas. Tampilan ini membantu pengguna membaca hasil model secara terstruktur karena hasil visual dan hasil numerik ditampilkan dalam satu halaman. Dengan demikian, pengguna dapat melihat posisi objek pada citra sekaligus mengecek kelas dan tingkat keyakinan deteksi. Adapun web ui utput ditujukan pada gambar 7.



GAMBAR 7
(TAMPILAN OUTPUT WEB UI HASIL DETEKSI)

Dalam konteks penggunaannya, sistem ini berperan sebagai decision support. Artinya, sistem memberikan informasi visual yang dapat membantu analisis, tetapi keputusan akhir tetap berada pada manusia atau perangkat regulasi pertandingan. Pendekatan ini penting agar hasil deteksi model tidak diposisikan sebagai pengganti aturan sepak bola, melainkan sebagai alat bantu untuk membaca kejadian secara lebih objektif dan terstruktur.

G. Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan. Dataset yang digunakan merupakan dataset publik dari Roboflow, bukan dataset primer yang dikumpulkan secara langsung dari perekaman pertandingan. Oleh karena itu, variasi kamera, kualitas video, jarak pengambilan gambar, dan konteks pertandingan mengikuti karakteristik dataset yang tersedia. Selain itu, kondisi non-foul tidak ditambahkan sebagai kelas anotasi tersendiri, sehingga keluaran non-foul hanya dimaknai sebagai kondisi ketika model tidak menghasilkan deteksi terhadap tiga kelas utama pada ambang confidence tertentu.

Keterbatasan lain terletak pada fokus deteksi berbasis citra. Model dapat mengenali objek pada frame atau citra, tetapi belum menganalisis urutan waktu kejadian secara eksplisit. Pada pertandingan nyata, keputusan pelanggaran sering membutuhkan konteks sebelum dan sesudah kontak fisik terjadi. Oleh karena itu, penelitian selanjutnya dapat mengombinasikan deteksi objek dengan analisis temporal video agar sistem mampu membaca pola kejadian secara lebih lengkap.

V. KESIMPULAN

Penelitian ini berhasil mengimplementasikan YOLOv11n untuk mendeteksi pelanggaran sepak bola berbasis citra pada tiga kelas utama, yaitu foul, fouling_player, dan victim. Dataset Roboflow dikonversi ke format YOLO, dibagi menjadi train pool dan test set, kemudian digunakan dalam pelatihan berbasis 5-Fold Cross Validation. Implementasi model menghasilkan bounding box, label kelas, dan confidence score yang dapat divisualisasikan melalui rancangan Web UI.

Hasil evaluasi menunjukkan bahwa rata-rata 5-Fold Cross Validation memperoleh precision 95,07%, recall 96,10%, mAP50 98,03%, dan mAP50-95 77,94%. Fold 1 menjadi model terbaik dengan mAP50-95 sebesar 78,16%. Evaluasi pada test set menghasilkan precision 95,60%, recall 95,90%, mAP50 98,00%, dan mAP50-95 77,70%. Hasil ini menunjukkan bahwa YOLOv11n mampu mempertahankan performa pada data uji yang tidak digunakan saat training maupun validasi.

Penelitian selanjutnya disarankan menggunakan dataset primer dari pertandingan nyata, memperbanyak variasi sudut kamera dan pencahayaan, serta menguji model pada video real-time dengan skenario pertandingan yang lebih kompleks. Selain itu, integrasi temporal berbasis video dapat dikembangkan agar sistem tidak hanya mendeteksi objek pada citra tunggal, tetapi juga memahami urutan kejadian pelanggaran secara lebih lengkap.

REFERENSI

[1] R. Maneiro *et al.*, “The effect of the video assistant referee (VAR) on referees’ decisions at FIFA Women’s World Cups,” 2022. [Online]. Available: https://longomatch.com/en/

[2] J. Fang, C. Yeung, and K. Fujii, “Foul prediction with estimated poses from soccer broadcast video,” Feb. 2024, [Online]. Available: http://arxiv.org/abs/2402.09650

[3] Y. Shao and Z. He, “VAR-YOLOv8s: IoT-based automatic foul detection in soccer matches,” *Alexandria Engineering Journal*, vol. 111, pp. 555–565, Jan. 2025, doi: 10.1016/j.aej.2024.10.031.

[4] A. Rabee, Z. Anwar, A. AbdelMoety, A. Abdelsallam, and M. Ali, “Comparative analysis of automated foul detection in football using deep learning architectures,” *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-96945-0.

[5] J. Wang and L. Li, “A method for feature division of Soccer Foul actions based on salience image semantics,” *PLoS One*, vol. 20, no. 6 JUNE, Jun. 2025, doi: 10.1371/journal.pone.0322889.

[6] F. Zheng, D. Z. Al-Hamid, P. H. J. Chong, C. Yang, and X. J. Li, “A Review of Computer Vision Technology for Football Videos,” May 01, 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/info16050355.

[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” 2016. [Online]. Available: http://pjreddie.com/yolo/

[8] R. Khanam and M. Hussain, “YOLOv11: An Overview of the Key Architectural Enhancements,” Oct. 2024, [Online]. Available: http://arxiv.org/abs/2410.17725

[9] T.-Y. Lin *et al.*, “Microsoft COCO: Common Objects in Context,” Feb. 2015, [Online]. Available: http://arxiv.org/abs/1405.0312

[10] S. Raschka, “Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning,” Nov. 2020, [Online]. Available: http://arxiv.org/abs/1811.12808

[11] Y. Wu, L. Geng, X. Guo, C. Wu, and G. Yu, “An Improved YOLOv11n Model Based on Wavelet Convolution for Object Detection in Soccer Scenes,” *Symmetry (Basel).*, vol. 17, no. 10, Oct. 2025, doi: 10.3390/sym17101612.

[12] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, “Generalized Intersection over Union: A Metric and A Loss for Bounding

Box Regression," Apr. 2019, [Online]. Available: http://arxiv.org/abs/1902.09630