

Improving YOLO Performance with Advanced Data Augmentation for Soccer Object Detection

Rahayuning Febriyanti Puspita¹, Muhammad Naufal^{2*}, Farrikh Al Zami³

Teknik Informatika, Universitas Dian Nuswantoro, Semarang, Indonesia

111202214583@mhs.dinus.ac.id¹, m.naufal@dsn.dinus.ac.id², alzami@dsn.dinus.ac.id³

Article Info

Article history:

Received 2025-09-18

Revised 2025-11-12

Accepted 2025-11-15

Keyword:

Football,

Data augmentation,

Object detection,

YOLOv8m.

ABSTRACT

This study developed an object detection system for soccer games using the YOLOv8m algorithm with four main classes: player, goalkeeper, referee, and ball. The dataset, consisting of 372 annotated images, exhibited class imbalance, with significantly fewer ball instances compared to players. The basic YOLOv8m architecture was used without internal modifications, but adjustments were made to the output layer and fine-tuning of the pre-trained weights to adapt to the new dataset. Two models were compared: one without and one with advanced augmentation techniques (*mosaic*, *mixup*, *cutmix*). The experimental results showed an increase in *mAP@50* from 74.9% to 81.4% in the augmented model, with a statistically significant difference ($p < 0.01$). However, model performance still decreased under extreme conditions such as high occlusion, rapid movement, and uneven lighting. The combination of data augmentation, output layer adaptation, and fine-tuning proved effective in improving object detection accuracy and provided the basis for the development of a real-time artificial intelligence-based soccer match analysis system.



This is an open access article under the [CC-BY-SA](#) license.

I. PENDAHULUAN

Permainan sepak bola modern menuntut kecepatan dan ketepatan dalam pengambilan keputusan, baik dari sisi taktik pelatih, analisis pertandingan, maupun dalam mendukung kewenangan wasit di lapangan. Dalam kondisi permainan yang kompleks dan dinamis, pendekatan observasi konvensional secara manual memiliki keterbatasan signifikan, terutama dalam hal kecepatan dan konsistensi pemantauan berbagai elemen penting seperti pemain, bola, penjaga gawang, dan wasit. Oleh karena itu, diperlukan sistem berbasis pengolahan citra yang mampu mengenali objek secara otomatis dengan ketepatan tinggi dan waktu pemrosesan yang singkat[1].

Salah satu pendekatan yang berkembang pesat dalam pengenalan objek visual adalah metode deteksi berbasis kerangka kerja satu tahap seperti YOLO (You Only Look Once), yang sejak diperkenalkan pada 2015 telah mengalami berbagai pembaruan struktural [1]. Hal ini didukung oleh kajian sistematis yang menelusuri perkembangan YOLO dari

versi pertama hingga versi terkini, yang menunjukkan perbaikan dalam arsitektur internalnya serta kecocokan dengan berbagai perangkat keras ringan [2], [10]. Dalam proses deteksi, efisiensi pemrosesan, dan peningkatan akurasi pada objek berukuran kecil [2]. Pembaruan ini memungkinkan pengaplikasian teknologi ini dalam situasi yang menuntut hasil waktu nyata, termasuk dalam pertandingan olahraga.

Dalam penelitian terbaru, diterapkan pendekatan deteksi otomatis untuk mengenali bola dan pemain pada pertandingan sepak bola dengan posisi kamera dari jarak jauh (long-shot), yang selama ini menjadi kendala bagi metode pengenalan visual lainnya. Studi tersebut menunjukkan bahwa penerapan YOLOv8 memberikan hasil yang lebih unggul dibandingkan metode terdahulu seperti FootAndBall, khususnya dalam mendekripsi objek-objek kecil di lingkungan visual yang padat. Hal ini sejalan dengan temuan Chen et al. (2024) yang menunjukkan efektivitas YOLOv8n untuk mendekripsi objek kecil pada citra UAV [3], serta penelitian Wang et al. (2024) yang

mengusulkan peningkatan deteksi objek kecil melalui mekanisme *Mixed Local Channel Attention* dalam YOLOv8 untuk citra penginderaan jauh [4].

Peningkatan hasil deteksi dapat diperoleh melalui penambahan variasi pada data pelatihan (augmentasi data), yang secara empiris terbukti mampu memperbaiki nilai ketepatan (precision) dan tingkat keberhasilan sistem dalam mengenali objek (recall). Beberapa penelitian menunjukkan bahwa dengan teknik augmentasi yang sesuai, seperti pemotongan acak, pencermian, rotasi, dan perubahan pencahayaan, dapat meningkatkan nilai mAP (mean Average Precision) secara signifikan [5], [6].

Prototype ini tidak terbatas pada pendekslsian pemain, wasit dan bola saja, namun diperluas dengan melakukan perhitungan pada objek yang terdeteksi, sehingga memberikan peluang bagi pengembangan sistem pendukung keputusan otomatis dalam pertandingan [7]. Selain itu, telah dilakukan pemisahan citra operator kamera dari siaran pertandingan menggunakan sistem deteksi yang sama, dan hasilnya menunjukkan keberhasilan dalam segmentasi objek yang tidak relevan secara kontekstual, sehingga dapat meningkatkan kualitas visualisasi pertandingan secara keseluruhan [8].

Dalam sebuah studi komparatif yang membandingkan versi YOLOv7, YOLOv8, dan YOLOv9, ditemukan bahwa versi kedelapan memiliki keunggulan dari sisi akurasi dan efisiensi waktu pelatihan, serta lebih andal dalam mendekripsi bola dan pemain pada berbagai posisi kamera [9]. Hal ini didukung oleh kajian sistematis yang menelusuri perkembangan YOLO dari versi pertama hingga versi terkini, yang menunjukkan perbaikan dalam arsitektur internalnya serta kecocokan dengan berbagai perangkat keras ringan [2], [10].

Beberapa pengembangan juga telah dilakukan untuk lebih mengoptimalkan kinerja YOLOv8 dalam mendekripsi objek berukuran kecil, dengan menambahkan modul pemrosesan khusus dan teknik penyesuaian lanjutan, seperti modul peliputan spasial adaptif dan pendekatan multi-skala [11]. Dalam penerapan untuk objek bergerak cepat, seperti pelacakan bola yang dilontarkan, sistem ini menunjukkan kemampuan mendekripsi lintasan secara tepat dengan kecepatan pemrosesan yang tetap terjaga [12].

Sementara itu, di lingkungan visual yang tidak biasa seperti video 360 derajat, penyesuaian bentuk augmentasi juga telah diterapkan untuk meningkatkan kemampuan deteksi pada bidang pandang melingkar. Hasil penelitian menunjukkan bahwa pendekatan ini memberikan peningkatan yang signifikan dibandingkan dengan augmentasi konvensional [13]. Teknik serupa juga telah digunakan dalam pengamatan citra pertanian, di mana lingkungan lapangan terbuka dan variasi pencahayaan menuntut ketangguhan sistem deteksi yang tinggi [14].

Kajian terbaru terhadap YOLOv8 dan versi penerusnya, seperti YOLOv10, menekankan bahwa meskipun versi yang lebih baru memiliki kompleksitas struktur yang lebih tinggi,

YOLOv8 masih menjadi pilihan yang seimbang antara akurasi, kecepatan, dan kemudahan penerapan di berbagai perangkat [15]. Oleh karena itu, penting untuk melakukan pengujian lanjutan pada sistem ini dalam konteks nyata pertandingan sepak bola.

Meskipun YOLOv8m dikenal memiliki akurasi tinggi dan efisiensi yang baik dalam berbagai aplikasi deteksi objek, penerapannya pada konteks pertandingan sepak bola masih menghadapi sejumlah kendala praktis. Variasi kondisi pencahayaan antara siang dan malam, pergerakan cepat dari bola maupun pemain, serta tumpang tindih objek (*occlusion*) sering kali menyebabkan hasil deteksi menjadi tidak stabil. Dalam situasi tersebut, objek kecil seperti bola berpotensi tidak teridentifikasi dengan tepat, atau pemain yang saling berdekatan terdeteksi sebagai satu entitas. Permasalahan ini menandakan bahwa keterbatasan model tidak hanya bersumber dari arsitektur internal, tetapi juga dari kompleksitas lingkungan visual pertandingan. Oleh karena itu, penelitian ini berfokus pada peningkatan kemampuan generalisasi model melalui penerapan teknik augmentasi data dan proses *fine-tuning* agar YOLOv8m mampu beradaptasi dengan kondisi dinamis dan pencahayaan yang bervariasi di lapangan sepak bola.

Penelitian ini bertujuan untuk meningkatkan kinerja model deteksi objek YOLOv8m dalam konteks pertandingan sepak bola melalui penerapan dan evaluasi beberapa teknik augmentasi data, seperti *mosaic*, *mixup*, dan *cutmix*. Fokus utama penelitian ini tidak hanya untuk mengukur peningkatan performa secara keseluruhan, tetapi juga mengeksplorasi efektivitas relatif dari kombinasi teknik augmentasi tersebut terhadap akurasi deteksi, terutama pada objek kecil dan kondisi visual yang kompleks.

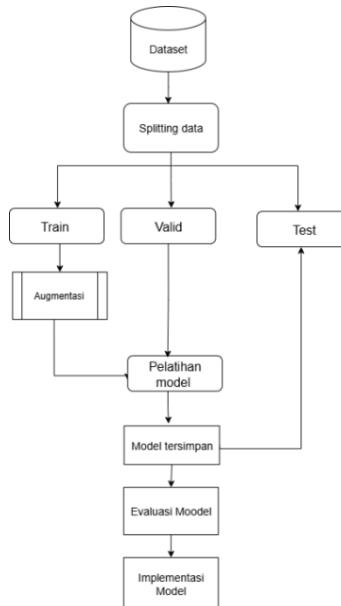
II. METODE

Penelitian ini dilaksanakan melalui tahapan rinci agar proses berjalan teratur dan hasil yang diperoleh dapat dipertanggungjawabkan. Alur penelitian ditunjukkan pada Gambar 1. Gambar 1 menunjukkan alur penelitian yang dilakukan dalam pengembangan sistem deteksi objek permainan sepak bola. Dataset yang tersedia terbagi dalam tiga bagian, yaitu data latih, validasi, dan pengujian, dengan penerapan teknik augmentasi pada data latih. Selanjutnya, proses pelatihan menggunakan YOLOv8m menghasilkan bobot terbaik (*best.pt*) yang kemudian divalidasi dan diuji menggunakan dataset terpisah. Hasil evaluasi model akhirnya diimplementasikan ke dalam sistem deteksi sederhana berbasis web.

A. Dataset Sepakbola

Tahap awal penelitian ini memanfaatkan dataset sepak bola yang bersumber dari platform open-source *Roboflow* [25]. Dataset ini terdiri dari 372 citra yang telah dianotasi secara manual ke dalam empat kelas objek utama, yaitu *player*, *goalkeeper*, *referee*, dan *ball*. Dataset tersebut

berasal dari rekaman pertandingan sepak bola yang diambil dari sumber siaran televisi dan dokumentasi lapangan terbuka, mencakup pertandingan lokal maupun internasional. Citra-citra tersebut mencerminkan variasi kondisi visual yang beragam, meliputi perbedaan pencahayaan antara siang dan malam, sudut pandang kamera yang berbeda (depan, samping, dan atas), serta tingkat kepadatan pemain yang bervariasi di setiap frame.



Gambar 1. Flowchart alur penelitian

Dataset dibagi menjadi 298 gambar (sekitar 80%) untuk pelatihan, 49 gambar (sekitar 13%) untuk validasi, dan 25 gambar (sekitar 7%) untuk pengujian menggunakan pendekatan *stratified train-validation-test split*, sehingga proporsi setiap kelas tetap seimbang di seluruh subset. Pendekatan ini memastikan evaluasi model dilakukan secara objektif terhadap data yang tidak digunakan selama proses pelatihan. Skema *k-fold cross-validation* tidak diterapkan karena keterbatasan ukuran dataset; namun, pelatihan dilakukan dengan lima nilai *random seed* berbeda untuk menjaga konsistensi hasil dan mengurangi potensi bias.

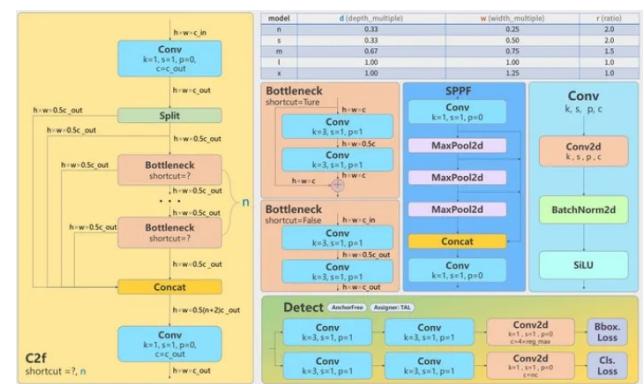
B. Augmentasi Data

Setelah data diperoleh, dilakukan proses augmentasi pada data latih untuk meningkatkan keragaman dan jumlah sampel, sehingga model menjadi lebih adaptif terhadap variasi kondisi pencahayaan, posisi, dan skala objek. Beberapa teknik augmentasi yang digunakan antara lain: penyesuaian intensitas warna ($hsv_v = 0,25$) untuk mensimulasikan perubahan pencahayaan, pergeseran posisi objek ($translate = 0,2$) dan perubahan skala ($scale = 0,25$) untuk memperluas variasi spasial, serta pembalikan horizontal ($flplr = 0,25$) untuk menambah variasi orientasi objek. Selain itu, digunakan pula kombinasi citra berbasis mosaic (0,25), *mixup* (0,25), dan *cutmix* (0,25) guna

memperkaya representasi fitur antar sampel dan meningkatkan kemampuan generalisasi model terhadap data baru.

C. YOLOv8m

YOLOv8m merupakan salah satu varian dari algoritma deteksi objek You Only Look Once generasi kedelapan yang dikembangkan oleh Ultralytics [16]. Huruf "m" pada YOLOv8m menunjukkan varian medium, yaitu versi yang dirancang untuk memberikan keseimbangan antara kecepatan inferensi dan tingkat akurasi [10]. Dengan demikian, YOLOv8m dipilih dalam penelitian ini karena mampu menghadirkan performa yang cukup stabil tanpa menuntut perangkat keras yang terlalu mahal.



Gambar 2. Struktur YOLOv8

Secara arsitektural, YOLOv8m menggunakan backbone yang telah disempurnakan dengan mekanisme konvolusi dan struktur blok C2f untuk mengekstraksi fitur lebih mendalam, serta dilengkapi dengan Path Aggregation Network (PAN) guna menggabungkan informasi lintas skala. Selain itu, pada tahap deteksi, YOLOv8m memanfaatkan *anchor-free detection head* [13], yang menjadikan proses prediksi kotak deteksi lebih fleksibel dan mampu mengenali objek dengan berbagai ukuran. Karakteristik ini menjadikan YOLOv8m lebih andal dalam mendeteksi objek kecil, seperti bola dalam pertandingan sepak bola, ataupun objek dengan bentuk kompleks seperti gestur wasit di tengah kerumunan pemain.

Keunggulan lain dari YOLOv8m adalah kompatibilitasnya dengan berbagai teknik augmentasi data modern, termasuk MixUp, CutMix, dan Mosaic, yang terbukti meningkatkan keragaman data latih [6]. Kombinasi arsitektur efisien dan dukungan terhadap augmentasi lanjutan inilah yang mendasari pemilihan YOLOv8m sebagai model utama dalam penelitian ini, dengan harapan dapat menghasilkan deteksi real-time yang cepat sekaligus akurat pada situasi pertandingan sepak bola yang dinamis.

Dalam penelitian ini, digunakan arsitektur dasar YOLOv8m tanpa modifikasi struktur internal seperti backbone atau detection head. Namun, dilakukan penyesuaian pada lapisan output agar sesuai dengan empat kelas objek yang digunakan dalam dataset sepak bola, yaitu

player, goalkeeper, referee, dan ball. Selain itu, proses pelatihan dilakukan dengan pendekatan *fine-tuning* terhadap *weight* pra-latih (pre-trained weights) YOLOv8m yang telah disediakan oleh Ultralytics. Pendekatan ini memungkinkan model untuk mempertahankan kemampuan deteksi umum dari model dasar sambil beradaptasi dengan karakteristik visual yang spesifik pada domain pertandingan sepak bola. Dengan demikian, penelitian ini tetap memanfaatkan kekuatan arsitektur dasar YOLOv8m, tetapi menambahkan kontribusi berupa adaptasi layer output dan *fine-tuning* untuk domain baru.

D. Pelatihan Model

Tahap berikutnya adalah pelatihan model menggunakan arsitektur YOLOv8m. Penelitian ini membandingkan dua pendekatan utama, yaitu pelatihan dengan konfigurasi YOLOv8m tanpa augmentasi tambahan serta pelatihan dengan penerapan augmentasi data sebagaimana dijelaskan pada tahap sebelumnya. Masing-masing pendekatan dijalankan sebanyak lima kali dengan menggunakan seed acak yang berbeda. Penggunaan lima seed ini bertujuan untuk mengevaluasi konsistensi performa model [9], sekaligus mengurangi bias yang mungkin muncul jika hanya dilakukan satu kali pelatihan. Strategi ini diharapkan dapat memberikan gambaran yang lebih akurat mengenai performa model secara keseluruhan.

TABEL I
KONFIGURASI HYPERPARAMETER PELATIHAN

Parameter	Nilai
Epoch	200
Batch size	32
Learning Rate (lr)	0,005
Image Size	640 x 640
Optimizer	AdamW
Data Source	data.yaml
Weight	YOLOv8m.pt

Konfigurasi parameter pelatihan yang digunakan ditunjukkan pada Tabel I, dengan pengaturan utama meliputi *epoch* sebanyak 200, *batch size* 32, *learning rate* 0.005, ukuran citra 640×640 piksel, dan *optimizer* AdamW. Seluruh proses pelatihan dijalankan di platform Google Colaboratory dengan dukungan GPU NVIDIA Tesla T4 dan lingkungan CUDA 12.1. Sistem secara otomatis menyimpan *best weights* berdasarkan performa terbaik pada data validasi, dan bobot tersebut digunakan pada tahap evaluasi akhir.

Setiap sesi pelatihan memerlukan waktu sekitar 45–60 menit untuk menyelesaikan 200 *epoch* dengan konfigurasi di atas. Hasil pengujian menunjukkan bahwa model hasil *fine-tuning* mampu melakukan inferensi pada citra tunggal dengan waktu rata-rata 25–30 milidetik per frame, yang

menunjukkan potensi penerapan sistem dalam analisis pertandingan sepak bola secara *real-time*. Pemilihan arsitektur YOLOv8m dilakukan karena memiliki keseimbangan yang baik antara jumlah parameter, kebutuhan komputasi, dan tingkat akurasi, sehingga sesuai untuk aplikasi berbasis waktu nyata.

E. Pengujian Model pada Data Test

Pengujian model dilakukan menggunakan data uji yang telah dipisahkan sejak awal proses eksperimen. Evaluasi dilakukan sebanyak lima kali sesuai dengan variasi seed yang digunakan dalam penelitian, sehingga diperoleh gambaran performa model yang lebih menyeluruh. Pada tahap ini, kinerja model diukur dengan menggunakan metrik evaluasi utama, yaitu *precision*, *recall*, serta *mean Average Precision (mAP)*. Hasil pengujian pada data uji memberikan informasi mengenai efektivitas model dalam mendeteksi objek target secara konsisten pada berbagai kondisi citra, serta menjadi dasar pembandingan antara model dengan augmentasi dan tanpa augmentasi.

F. Evaluasi Model

Pada tahap evaluasi model, yang dilakukan adalah membandingkan secara ilmiah apakah terdapat perbedaan performa yang signifikan antara YOLOv8m tanpa augmentasi dan YOLOv8m dengan augmentasi, yang dilakukan dengan analisis statistik menggunakan uji t-test [17]. Uji ini diterapkan terhadap hasil mAP dari kedua pendekatan, yang masing-masing diperoleh dari lima kali percobaan dengan seed yang berbeda menggunakan data uji. Dengan cara ini, hasil penelitian dapat memberikan bukti yang lebih kuat apakah peningkatan performa akibat augmentasi bersifat signifikan atau sekedar variasi acak yang tidak bermakna secara statistik.

G. Implementasi Model

Tahap akhir dari penelitian ini adalah implementasi model dalam bentuk sistem berbasis web. Implementasi dilakukan menggunakan framework Gradio untuk menyediakan antarmuka interaktif yang terhubung langsung dengan model pembelajaran mesin. Melalui aplikasi ini, pengguna dapat mengunggah citra uji dan langsung memperoleh hasil deteksi objek dari model yang telah dilatih. Dengan adanya implementasi ini, penelitian tidak hanya menghasilkan model secara teoritis, tetapi juga menghadirkan solusi praktis yang dapat dimanfaatkan oleh pengguna secara lebih luas.

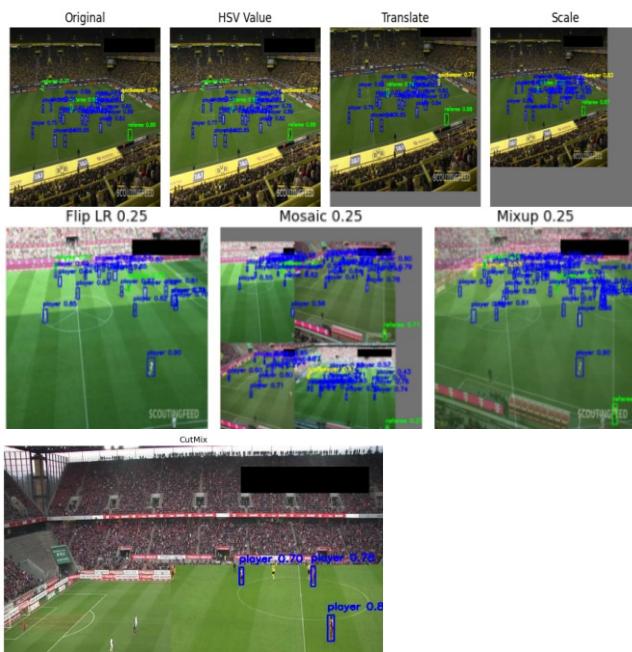
III. HASIL DAN PEMBAHASAN

Setelah melalui tahapan pengumpulan dataset dari Roboflow, anotasi objek, penerapan teknik augmentasi, serta pembagian data menjadi train, validasi, dan test set, langkah berikutnya adalah menyajikan hasil serta pembahasan dari eksperimen yang telah dilakukan. Hasil yang diperoleh dari

pelatihan model YOLOv8m, baik versi tanpa augmentasi maupun dengan augmentasi, kemudian dievaluasi melalui metrik kinerja utama dan diuji secara statistik untuk mengetahui signifikansi perbedaannya. Selain itu, implementasi model ke dalam sistem berbasis web juga dilakukan guna menguji sejauh mana sistem yang dibangun dapat diterapkan secara praktis.

A. Hasil Augmentasi Data

Tahap augmentasi data yang diimplementasikan secara langsung di dalam fungsi pelatihan “*train_and_evaluate*” menghasilkan variasi citra baru dengan keragaman visual yang lebih tinggi dibandingkan dataset asli. Setiap citra latih memiliki peluang untuk dimodifikasi melalui perubahan warna, intensitas pencahayaan, pergeseran posisi objek, perubahan skala, hingga penggabungan beberapa citra ke dalam satu bingkai. Dengan adanya variasi ini, model memperoleh representasi data yang lebih beragam sehingga mampu meningkatkan kemampuan generalisasi pada saat diuji dengan data yang belum pernah dilihat sebelumnya. Sample dari hasil augmentasi data dapat dilihat pada gambar 3 berikut.



Gambar 3. Visualisasi augmentasi data

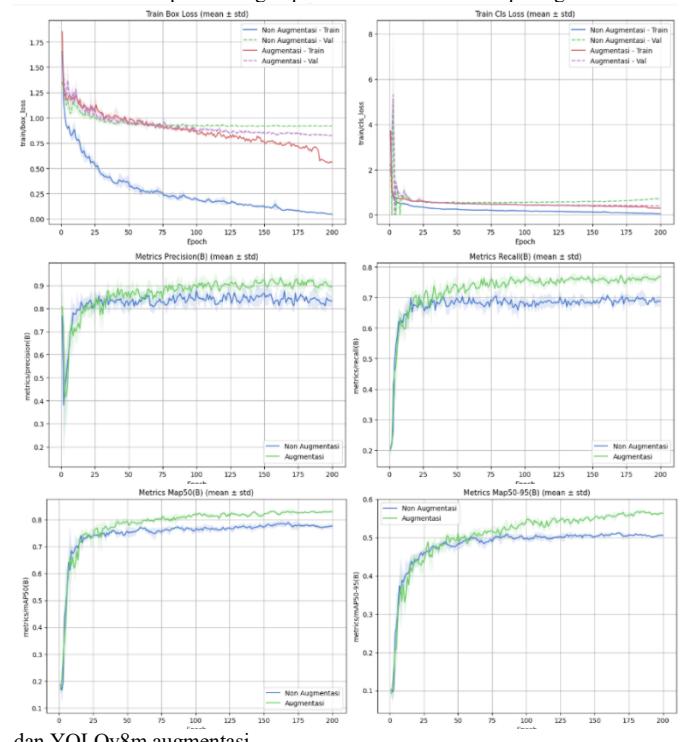
Visualisasi hasil augmentasi memperlihatkan bahwa objek dapat muncul dalam kondisi yang jauh berbeda dari citra aslinya. Sebagai contoh, sebuah citra pemain bola yang awalnya berada di sisi tengah gambar dapat digeser ke tepi kanan, diberi perubahan intensitas warna, atau bahkan digabung dengan citra lain menggunakan teknik *mosaic* maupun *cutmix*. Dengan demikian, augmentasi ini membantu memperkaya distribusi data latih sehingga model tidak hanya mengandalkan pola visual statis, tetapi juga belajar untuk

mengenali objek dalam situasi yang lebih realistik dan bervariasi. Strategi ini sangat penting karena kondisi nyata di lapangan pertandingan sepak bola sering kali melibatkan pencahayaan yang tidak merata, sudut kamera yang berubah-ubah, dan keberadaan banyak objek dalam satu bingkai.

B. Hasil Pelatihan Model

Proses pelatihan model dilakukan dengan dua pendekatan berbeda, yaitu YOLOv8m tanpa augmentasi dan YOLOv8m dengan augmentasi. Masing-masing pendekatan dijalankan sebanyak lima kali menggunakan seed acak yang telah ditentukan, yaitu 42, 123, 999, 7, dan 2023. Tujuan penggunaan seed yang berbeda adalah untuk menguji kestabilan hasil serta mengurangi bias yang mungkin muncul akibat inisialisasi parameter secara acak. Hasil pelatihan perbandingan performa model dari 5 seed tersebut bisa dilihat dari Gambar 4 berikut.

Gambar 4. Grafik perbandingan performa YOLOV8m tanpa augmentasi



dan YOLOv8m augmentasi

Berdasarkan grafik perbandingan, terlihat bahwa model tanpa augmentasi mengalami penurunan *training box loss* yang lebih cepat dan stabil dibandingkan dengan model augmentasi. Garis *loss* pada model tanpa augmentasi menurun tajam hingga mendekati nol, sedangkan augmentasi cenderung konvergen lebih lambat. Akan tetapi, pada grafik metrik evaluasi seperti *precision*, *recall*, serta *map*, model dengan augmentasi konsisten menunjukkan nilai yang lebih tinggi daripada model tanpa augmentasi. Hal ini menunjukkan bahwa meskipun *loss* pelatihan lebih kecil pada model tanpa augmentasi, kemampuan generalisasi

lebih baik dimiliki oleh model dengan augmentasi. Hal ini juga didukung dengan hasil nilai akhir rata-rata dan standar deviasi yang bisa dilihat di Tabel II berikut.

TABEL II
RINGKASAN NILAI AKHIR (MEAN ± STD)

Metrik	Tanpa Augmentasi	Augmentasi
Train/box loss	0.04 ± 0.00	0.56 ± 0.01
Val/box loss	0.92 ± 0.00	0.82 ± 0.00
Train/cls loss	0.04 ± 0.00	0.29 ± 0.00
Val/cls loss	0.71 ± 0.02	0.39 ± 0.00
Metrics/precision(B)	0.83 ± 0.03	0.89 ± 0.03
Metrics/recall(B)	0.69 ± 0.02	0.77 ± 0.01
Metrics/mAP@50(B)	0.78 ± 0.01	0.81 ± 0.01
Metrics/mAP@50-95(B)	0.51 ± 0.00	0.56 ± 0.00

Dari hasil pada Tabel II tersebut, terlihat bahwa pada *train/box_loss*, model tanpa augmentasi mencapai nilai akhir 0.04 ± 0.00 , jauh lebih rendah dibandingkan dengan model augmentasi (0.56 ± 0.01). Namun, pada metrik performa utama, model augmentasi lebih unggul. *Precision* model augmentasi mencapai 0.89 ± 0.00 , dibandingkan dengan tanpa augmentasi yang hanya 0.83 ± 0.00 . Demikian pula, nilai *recall* (0.77 ± 0.00 dibandingkan dengan 0.69 ± 0.02) serta mAP@50 (0.78 ± 0.00 dibandingkan dengan 0.73 ± 0.00) dan mAP@50-95 (0.56 ± 0.00 dibandingkan dengan 0.51 ± 0.00) juga menunjukkan keunggulan pada model dengan augmentasi.

Secara matematis, metrik evaluasi tersebut dihitung dengan rumus berikut:

- *Precision* adalah metrik yang digunakan untuk mengukur ketepatan prediksi model, yaitu seberapa banyak deteksi yang benar dari seluruh deteksi yang dihasilkan [4].

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

dimana TP adalah *True Positive* dan FP adalah *False Positive*.

- *Recall* adalah metrik yang digunakan untuk mengukur sejauh mana model mampu menemukan semua objek yang ada pada data uji [4].

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

dimana FN adalah *False Negative*.

- *mAP* adalah metrik gabungan yang memperhitungkan *precision* dan *recall* pada berbagai ambang batas *IoU*(Intersection over Union) [5].

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N AP_i \quad (3)$$

dengan AP_i adalah *Average Precision* dari setiap kelas, dan N jumlah kelas.

Dengan demikian, hasil ini menegaskan bahwa penerapan augmentasi data memberikan peningkatan yang signifikan terhadap kinerja model, terutama pada aspek generalisasi dan konsistensi performa deteksi objek.

C. Hasil Pengujian Model Pada Data Test

Pengujian pada data uji dilakukan untuk mengevaluasi sejauh mana model yang telah dilatih mampu melakukan generalisasi terhadap data baru yang sebelumnya tidak pernah digunakan dalam proses pelatihan. Beberapa metrik utama seperti *precision*, *recall*, mAP@50, dan mAP@50-95 digunakan sebagai indikator efektivitas dan keandalan model dalam mendeteksi objek target. Hasil pengujian untuk kedua model ditampilkan pada Tabel III dan Tabel IV.

TABEL III
HASIL PENGUJIAN MODEL YOLOV8

Seed	Precision	Recall	mAP@50	mAP@50-95
42	0.904	0.663	0.754	0.524
123	0.751	0.686	0.739	0.51
999	0.88	0.657	0.731	0.51
7	0.766	0.69	0.777	0.531
2023	0.93	0.685	0.743	0.524
mean ± std	0.846 ± 0.08	0.676 ± 0.01	0.749 ± 0.01	0.520 ± 0.00

TABEL IV
HASIL PENGUJIAN MODEL AUGMENTASI+YOLOV8

Seed	Precision	Recall	mAP@50	mAP@50-95
42	0.838	0.773	0.818	0.576
123	0.878	0.777	0.804	0.577
999	0.848	0.774	0.817	0.578
7	0.892	0.776	0.823	0.584
2023	0.842	0.778	0.807	0.572
mean ± std	0.859 ± 0.02	0.776 ± 0.00	0.814 ± 0.00	0.577 ± 0.00

Berdasarkan hasil pengujian pada Tabel III dan Tabel IV, terlihat adanya perbedaan performa yang jelas antara model YOLOv8 tanpa augmentasi dan dengan augmentasi. Model tanpa augmentasi menunjukkan nilai rata-rata *precision* sebesar $0,846 \pm 0,082$, *recall* sebesar $0,676 \pm 0,015$, *mAP@50* sebesar $0,749 \pm 0,018$, serta *mAP@50-95* sebesar $0,520 \pm 0,009$. Hasil ini mengindikasikan bahwa meskipun model mampu menghasilkan presisi yang cukup tinggi, nilai recall yang relatif rendah menunjukkan masih terdapat beberapa objek yang tidak terdeteksi dengan baik. Sebaliknya, penerapan augmentasi pada proses pelatihan memberikan dampak positif terhadap kinerja model. Rata-

rata nilai *precision* meningkat menjadi $0,859 \pm 0,024$, *recall* naik menjadi $0,776 \pm 0,002$, *mAP@50* menjadi $0,814 \pm 0,008$, serta *mAP@50-95* mencapai $0,577 \pm 0,004$.

Analisis lebih lanjut menunjukkan bahwa teknik *mosaic* berperan penting dalam meningkatkan kemampuan model mendekripsi objek kecil seperti bola, karena memperluas variasi spasial dan latar belakang dalam data latih. Sementara itu, *mixup* membantu mengurangi *overfitting* dengan mengombinasikan dua citra secara acak, dan *cutmix* memperbaiki ketahanan model terhadap tumpang tindih objek (*occlusion*). Kombinasi dari ketiga teknik ini terbukti efektif dalam meningkatkan kemampuan generalisasi model terhadap kondisi visual yang kompleks pada pertandingan sepakbola.

Secara keseluruhan, hasil pengujian memperlihatkan bahwa penggunaan teknik augmentasi berperan penting dalam memperkuat performa deteksi YOLOv8m, baik secara kuantitatif maupun kualitatif. Untuk memahami lebih jauh sumber peningkatan tersebut, dilakukan analisis terhadap pola kesalahan deteksi yang muncul pada kedua model. Model tanpa augmentasi masih menunjukkan beberapa kesalahan khas, seperti bola yang tidak terdeteksi pada kondisi *motion blur*, kesalahan klasifikasi antara pemain dan penjaga gawang akibat kemiripan warna seragam, serta *multiple detection* pada objek yang sama. Setelah augmentasi diterapkan, frekuensi kesalahan tersebut berkurang secara signifikan, dan model mampu menghasilkan *bounding box* yang lebih presisi serta *confidence score* yang lebih stabil di atas 0,80.

Hasil ini mengindikasikan bahwa peningkatan performa numerik yang dicapai tidak hanya bersifat statistik, tetapi juga mencerminkan perbaikan nyata pada stabilitas dan keandalan deteksi di tingkat visual. Dengan demikian,

eksplorasi terhadap kombinasi teknik augmentasi terbukti memberikan dampak yang substansial terhadap peningkatan akurasi dan konsistensi deteksi objek dalam lingkungan pertandingan yang dinamis.

D. Analisis Statistik (Uji t-test)

Untuk memastikan apakah perbedaan performa antara YOLOv8m tanpa augmentasi dan YOLOv8m dengan augmentasi signifikan secara statistik, dilakukan uji *t-test* berpasangan pada nilai *mAP@50* dari data uji [16]. Secara matematis uji *t-test* berpasangan dapat dituliskan sebagai berikut:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \quad (4)$$

\bar{d} = rata-rata selisih dari pasangan data

s_d = standar deviasi dari selisih

n = jumlah pasangan sampel

Nilai *t* kemudian dibandingkan dengan distribusi *t* pada derajat kebebasan ($n - 1$) untuk menentukan apakah perbedaan yang dihasilkan signifikan [18]. Hasil perhitungan ditunjukkan pada Tabel IV.

TABEL IV
HASIL UJI T-TEST

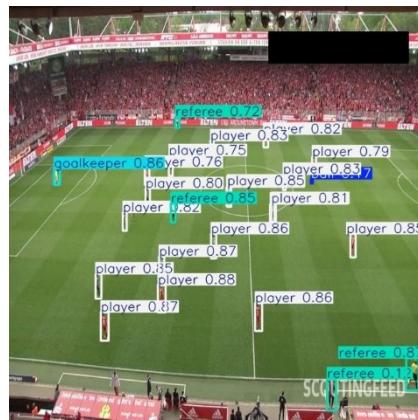
Data	Model	Mean	Variance	t Stat	df	P (T≤ t) two-tail	α(0.05)	Keterangan
Test	Tanpa Augmentasi	0.749	0.0003172	-7.447	8	7.28E-05	0.05	Sangat Signifikan
	Augmentasi	0.814	6.37E-05					

Berdasarkan hasil uji *t-test* pada data uji, nilai rata-rata *mAP@50* model tanpa augmentasi adalah 0,749 sedangkan model dengan augmentasi mencapai 0,814. Nilai *t-stat* sebesar $-7,447$ dengan *p-value* (two-tail) sebesar $7,28 \times 10^{-5}$ ($p < 0,01$) menunjukkan bahwa terdapat perbedaan yang sangat signifikan antara kedua metode pada tingkat signifikansi 1%. Hal ini membuktikan bahwa penggunaan augmentasi tidak hanya meningkatkan performa rata-rata, tetapi juga memberikan perbedaan yang nyata secara statistik dalam kemampuan generalisasi model pada data uji. Dengan demikian, augmentasi berperan penting dalam meningkatkan keandalan deteksi serta mengurangi risiko kesalahan model pada kasus nyata. Selain signifikansi statistik yang telah dibuktikan melalui uji *t-test*, peningkatan performa sebesar

6,5% (dari 74,9% menjadi 81,4%) juga memiliki arti praktis yang penting. Dalam konteks pelacakan waktu nyata pada pertandingan sepak bola, peningkatan ini berpengaruh terhadap kestabilan hasil deteksi dan kecepatan respon sistem dalam mengenali objek yang bergerak cepat. Model dengan augmentasi terbukti lebih konsisten dalam mempertahankan deteksi terhadap pemain, bola, dan wasit meskipun terjadi perubahan sudut kamera atau pencahayaan. Dengan demikian, peningkatan performa tersebut tidak hanya signifikan secara statistik, tetapi juga memberikan manfaat nyata bagi penerapan sistem deteksi objek secara *real-time* di lingkungan pertandingan sebenarnya.

Meskipun hasil uji *t* menunjukkan adanya peningkatan performa yang signifikan secara statistik pada model dengan

augmentasi data, pengamatan lebih lanjut terhadap hasil deteksi mengungkapkan bahwa performa model masih mengalami penurunan pada kondisi ekstrem tertentu. Secara khusus, pada citra dengan tingkat *occlusion* tinggi—misalnya ketika beberapa pemain saling menutupi atau bola tertutup sebagian oleh tubuh pemain—nilai *confidence score* deteksi cenderung menurun dan menghasilkan *false negative*. Hal serupa juga terjadi pada situasi di mana objek bergerak sangat cepat (*motion blur*) atau pencahayaan tidak merata, yang menyebabkan model kesulitan mempertahankan konsistensi deteksi antar frame. Meskipun teknik augmentasi seperti *mosaic* dan *cutmix* telah membantu memperluas variasi data latih, model tetap menunjukkan keterbatasan dalam mengenali objek dengan deformasi atau tumpang tindih ekstrem. Hasil ini menunjukkan bahwa meskipun



Gambar 7. Visualisasi prediksi YOLOv8m original

Pada model yang dilatih tanpa augmentasi (Gambar 7), deteksi terhadap objek utama seperti *player*, *referee*, dan *goalkeeper* umumnya berhasil, namun masih ditemukan beberapa kekurangan. Objek berukuran kecil seperti bola sering kali tidak terdeteksi, terutama saat tertutup sebagian oleh pemain lain (*occlusion*) atau berada di area dengan pencahayaan tidak merata. Selain itu, beberapa hasil prediksi menunjukkan nilai *confidence* yang rendah dan bounding box yang tidak sepenuhnya sejajar dengan bentuk objek sebenarnya.

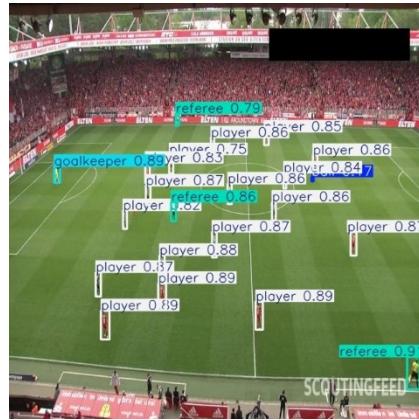
Sebaliknya, model dengan augmentasi (Gambar 8) memperlihatkan peningkatan yang signifikan secara visual. Deteksi bola menjadi lebih stabil bahkan ketika posisinya sebagian tertutup atau berada di area gelap. Objek *referee* dan *goalkeeper* juga dapat dikenali dengan lebih tepat, didukung oleh *bounding box* yang lebih proporsional dan nilai *confidence* yang merata di atas 0,80. Peningkatan ini menunjukkan bahwa penerapan teknik augmentasi seperti *mosaic*, *cutmix*, dan *mixup* mampu memperluas keragaman data pelatihan, sehingga model menjadi lebih adaptif terhadap variasi kondisi visual di lapangan.

Hasil ini memperkuat analisis kuantitatif sebelumnya, di mana model dengan augmentasi menunjukkan peningkatan *mAP@50* sebesar 6,5% dibandingkan model tanpa

peningkatan performa akibat augmentasi bersifat signifikan secara statistik, masih terdapat ruang untuk pengembangan model lebih lanjut, seperti penerapan pendekatan berbasis *temporal feature aggregation* atau *motion-aware learning* untuk meningkatkan ketahanan model terhadap kondisi visual yang kompleks dalam pertandingan sepak bola nyata.

E. Visualisasi Hasil Prediksi

Visualisasi hasil prediksi pada data uji memberikan gambaran kualitatif mengenai perbedaan performa antara model YOLOv8m tanpa augmentasi dan YOLOv8m dengan augmentasi yang dapat dilihat pada Gambar 7 dan Gambar 8.



Gambar 8. Visualisasi prediksi YOLOv8m augmentasi

augmentasi. Oleh karena itu, perbaikan performa tidak hanya terbukti secara numerik, tetapi juga terlihat secara kualitatif melalui hasil deteksi yang lebih stabil, presisi, dan relevan terhadap kondisi nyata pertandingan sepak bola.

Walaupun penerapan teknik augmentasi data terbukti meningkatkan akurasi dan stabilitas hasil deteksi, strategi ini juga memiliki konsekuensi dari sisi efisiensi komputasi. Berdasarkan hasil pengamatan selama proses pelatihan, waktu training pada model dengan augmentasi meningkat sekitar 20–25% dibandingkan model tanpa augmentasi. Kenaikan ini terjadi karena proses pembentukan citra sintetis seperti *mosaic*, *mixup*, dan *cutmix* memerlukan tahapan komputasi tambahan di setiap iterasi pelatihan. Meskipun demikian, teknik augmentasi yang digunakan tidak menambah kompleksitas arsitektur YOLOv8m, sebab seluruh proses dilakukan pada tahap praproses data sebelum masuk ke jaringan utama. Dengan kata lain, metode ini hanya berdampak pada durasi pelatihan, bukan pada jumlah parameter model maupun kecepatan inferensi. Peningkatan waktu pelatihan tersebut dinilai sepadan dengan kenaikan performa deteksi yang signifikan, sehingga pendekatan berbasis augmentasi tetap efisien dan relevan untuk pengembangan sistem deteksi real-time.

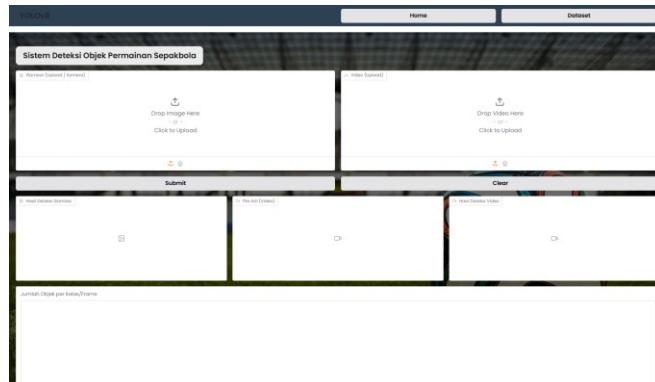
F. Implementasi Hasil

Sebagai tahap akhir penelitian, model YOLOv8m yang telah dilatih kemudian diimplementasikan dalam bentuk aplikasi berbasis web menggunakan framework Gradio. Aplikasi ini dirancang sederhana namun fungsional, sehingga dapat diakses dengan mudah melalui antarmuka web. Terdapat tiga halaman utama yang disediakan, yaitu *homepage*, *detection page*, dan *dataset page*.



Gambar 9. Tampilan homepage

Pada *homepage*, sistem menampilkan informasi umum mengenai aplikasi deteksi objek permainan sepakbola beserta menu navigasi.

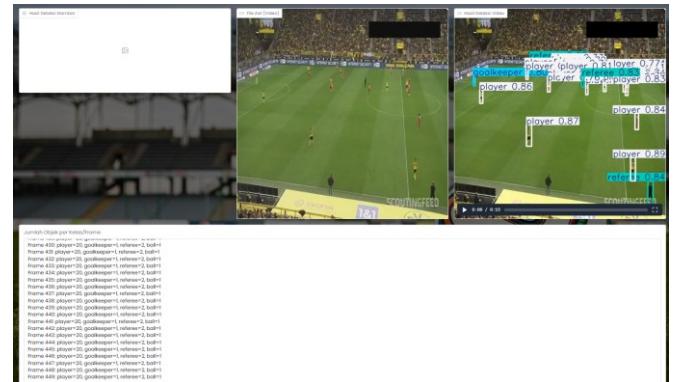


Gambar 10. Tampilan detection page

Pada *detection page*, pengguna dapat langsung melakukan uji coba deteksi objek dengan berbagai cara, seperti mengunggah gambar (*jpg*, *png*), video (*mp4*), atau mengambil gambar secara langsung melalui webcam.

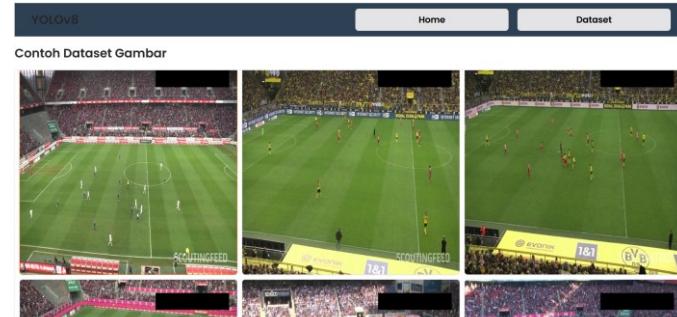


Gambar 11. Tampilan proses deteksi gambar



Gambar 12. Tampilan proses deteksi video

Hasil deteksi gambar maupun video akan ditampilkan secara visual dalam bentuk *bounding box* lengkap dengan label dan *confidence score*, serta informasi jumlah objek yang terdeteksi pada setiap kelas maupun per frame untuk data video. Tampilan ini memungkinkan pengguna tidak hanya melihat sisi visualisasi, tetapi juga memahami distribusi objek yang terdeteksi pada data input.



Gambar 13. Tampilan dataset page

Sementara itu, pada *dataset page*, sistem menyajikan contoh-contoh gambar yang digunakan dalam pelatihan model. Halaman ini berfungsi sebagai dokumentasi agar pengguna memahami jenis data yang menjadi dasar pelatihan model. Dengan struktur ini, implementasi berbasis Gradio memberikan pengalaman interaktif, memudahkan pengujian model secara langsung, sekaligus memperlihatkan

potensi penerapannya dalam sistem pendukung analisis pertandingan sepak bola secara *real-time*.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan teknik augmentasi data memberikan dampak yang signifikan terhadap peningkatan performa model YOLOv8m dalam mendekripsi objek pada permainan sepak bola. Meskipun model tanpa augmentasi mencapai nilai *training loss* yang lebih rendah, hasil validasi menunjukkan bahwa model tersebut memiliki kemampuan generalisasi yang lebih terbatas. Sebaliknya, model dengan augmentasi memperlihatkan konsistensi yang lebih baik, ditunjukkan oleh nilai *validation loss* yang lebih stabil serta peningkatan pada metrik evaluasi seperti *precision*, *recall*, *mAP@50*, dan *mAP@50-95*.

Selain penggunaan teknik augmentasi, penelitian ini juga melakukan penyesuaian pada lapisan output YOLOv8m agar sesuai dengan empat kelas objek yang digunakan, yaitu *player*, *goalkeeper*, *referee*, dan *ball*. Proses pelatihan dilakukan dengan pendekatan *fine-tuning* terhadap *pre-trained weights* YOLOv8m yang disediakan oleh Ultralytics, sehingga model dapat beradaptasi dengan karakteristik visual pada domain sepak bola tanpa perlu mengubah struktur arsitektur dasarnya. Pendekatan ini memungkinkan pemanfaatan efisiensi model dasar sambil meningkatkan akurasi deteksi melalui adaptasi kontekstual terhadap dataset baru.

Meskipun hasil pengujian menunjukkan peningkatan performa yang signifikan secara statistik, analisis lebih lanjut mengindikasikan bahwa model masih mengalami penurunan performa pada kondisi ekstrem, seperti saat terjadi *occlusion* tinggi, objek bergerak cepat (*motion blur*), atau pencahayaan yang tidak merata. Situasi ini menyebabkan penurunan *confidence score* dan peningkatan kemungkinan *false negative*, terutama pada kelas bola dan wasit. Oleh karena itu, arah penelitian lanjutan akan difokuskan pada pengembangan pendekatan berbasis *temporal feature aggregation* dan *motion-aware learning* untuk meningkatkan ketahanan model terhadap variasi kondisi visual yang kompleks pada pertandingan sepak bola nyata.

Secara keseluruhan, penelitian ini menegaskan bahwa kombinasi antara augmentasi data lanjutan, adaptasi lapisan output, dan strategi *fine-tuning* yang tepat dapat meningkatkan kemampuan generalisasi YOLOv8m dalam domain olahraga. Pendekatan ini dapat menjadi dasar pengembangan sistem deteksi objek real-time yang lebih akurat dan tangguh di masa depan, serta membuka peluang integrasi dengan sistem analisis taktik atau asisten wasit berbasis kecerdasan buatan.

Sebagai arah pengembangan berikutnya, penelitian ini dapat diperluas ke tahap pelacakan multi-objek (*Multi-Object Tracking/MOT*) untuk memantau pergerakan setiap pemain dan bola secara berkesinambungan pada setiap frame

video. Integrasi antara deteksi dan pelacakan ini akan memungkinkan analisis temporal terhadap dinamika pertandingan serta interaksi antar pemain. Selain itu, pengembangan menuju sistem pengenalan aksi (*action recognition*) juga menjadi langkah logis untuk mengenali aktivitas spesifik seperti umpan, tendangan, atau pelanggaran yang relevan bagi analisis performa dan pengambilan keputusan di lapangan.

Selain arah penelitian tersebut, sistem deteksi ini juga berpotensi dikembangkan dengan menambahkan kelas objek baru yang relevan dalam konteks pertandingan sepak bola, seperti garis lapangan, gawang, kartu wasit, atau bola cadangan. Penambahan kelas-kelas ini akan memperluas cakupan analisis, memungkinkan sistem tidak hanya mengenali entitas manusia dan bola, tetapi juga elemen kontekstual yang penting untuk analisis posisi, pelanggaran, dan situasi permainan. Dengan demikian, model deteksi dapat menjadi fondasi untuk sistem analisis taktik yang lebih komprehensif dan mendukung penerapan kecerdasan buatan dalam manajemen pertandingan sepak bola secara menyeluruh.

Lebih lanjut, sistem yang dikembangkan ini dirancang agar dapat diintegrasikan dengan perangkat real-time, seperti kamera stadion, drone pengawas, atau sistem VAR (*Video Assistant Referee*). Aliran video dari perangkat tersebut dapat diproses secara langsung melalui *edge computing* atau server lokal yang menjalankan model YOLOv8m, sehingga hasil deteksi objek dapat ditampilkan secara instan pada *dashboard* analisis. Dengan demikian, sistem ini berpotensi untuk mendukung pengambilan keputusan di lapangan secara cepat dan objektif dalam konteks pertandingan nyata.

DAFTAR PUSTAKA

- [1] X. Li, et al., "Automatic object detection for behavioural research using YOLOv8," *Behavior Research Methods*, vol. 56, pp. 2453–2472, 2024.
- [2] F. Zhang, et al., "Small Object Detection with YOLOv8 Algorithm Enhanced by MobileViTv3 and Wise-IoU," in *Proc. ICCPR*, 2023, pp. 35–40.
- [3] Y. Chen, et al., "Small Object Detection in UAV Images Based on YOLOv8n," *Int. J. Computational Intelligence Systems*, vol. 17, no. 1, 2024.
- [4] M. Naufal, H. A. Azies, and R. M. Brilianto, "Enhanced Brain Tumor Classification Through Gamma Correction in Deep Learning," *Sistemasi: Jurnal Sistem Informasi*, vol. 13, no. 6, pp. 2348–2358, Oct. 2024.
- [5] A. Ilyana, Nurdin, and Maryana, "Real-Time Detection of Coffee Cherry Ripeness Using YOLOv11," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 4, pp. 1170–1178, Aug. 2025
- [6] YOLOv8," *J. Computational Theoretical and Applied Mechanics*, vol. 1, no. 1, 2024.
- [7] J. Liu, et al., "Improved YOLOv8 for Small Object Detection," in *Proc. ICCN/IANet*, 2024, pp. 65–72. doi: 10.1145/3670105.3670150.
- [8] H. Wang, et al., "Detection of Small Object based on Improved-YOLOv8," *Frontiers in Computing and Intelligent Systems*, vol. 3, no. 2, pp. 115–128, 2024.

- [9] M. Khan, et al., "An improved YOLOv8 model for prohibited item detection with deformable convolution and dynamic head," *J. Real-Time Image Process.*, 2025. doi:[10.1145/3394171.3413828](https://doi.org/10.1145/3394171.3413828)
- [10] L. Zhou, et al., "A Small-Object Detection Model Based on Improved YOLOv8s for UAV Image Scenarios," *Remote Sensing*, vol. 16, no. 13, p. 2465, 2024.
- [11] R. Pramasyeta, et al., "Visual Entity Object Detection System in Soccer Matches Based on Various YOLO Architecture," *Jurnal Teknik Informatika (JUTIF)*, vol. 5, no. 3, 2024.
- [12] L. Yang, et al., "SFG-YOLOv8: efficient and lightweight small-feature gesture keypoint detector," *J. King Saud Univ. – Comput. Inf. Sci.*, 2025.
- [13] T. Ahmed, et al., "Object Detection and Tracking for Football Data Analytics," in *Proc. IACIDS EAI*, 2024.
- [14] Y. Li, et al., "SOD-YOLO: Small-Object-Detection Algorithm Based on Improved YOLOv8 for UAV Images," *Remote Sensing*, vol. 16, no. 16, p. 3057, 2024.
- [15] H. Ryu, et al., "YOLOv8 with Post-Processing for Small Object Detection Enhancement," *Applied Sciences*, vol. 15, no. 13, p. 7275, 2025.
- [16] Muljono, S. A. Wulandari, H. A. Azies, M. Naufal, W. A. Prasetyanto, and F. A. Zahra, "Breaking Boundaries in Diagnosis: Non-Invasive Anemia Detection Empowered by AI," *IEEE Access*, vol. 12, pp. 133578–133588, Jan. 2024, doi: 10.1109/ACCESS.2024.3353788.
- [17] A. Khalili and B. Smyth, "SOD-YOLOv8—Enhancing YOLOv8 for Small Object Detection in Aerial Imagery and Traffic Scenes," *Sensors*, vol. 24, no. 19, p. 6209, 2024.
- [18] S. Postupaev, et al., "Real-Time Camera Operator Segmentation with YOLOv8 in Football Video Broadcasts," *AI*, vol. 5, no. 2, pp. 842–872, 2024.
- [19] A. S. Fan, *YOLOv8-Lite: A Lightweight Object Detection Model for Real-Time Autonomous Driving Systems*, TETAI, vol. 1, no. 1, pp. 1–16, Apr. 2024. doi: 10.62762/TETAI.2024.894227
- [20] G. Corder and D. Foreman, *Nonparametric Statistics: A Step-by-Step Approach*, 2nd ed. Hoboken, NJ: Wiley, 2014.
- [21] A. Pramasyeta et al., *Research on Small Object Detection Algorithm Based on YOLOv8*, in *Proc. CSAI'24*, Beijing, Dec. 2024, pp. (paper number) doi :[10.1145/3709026.3709027](https://doi.org/10.1145/3709026.3709027)
- [22] Y. Zheng et al., *HTFD-YOLO: Small Target Detection in Drone Aerial Photography Based on YOLOv8s*, *J. Supercomputing*, vol. 81, art. no. 545, Feb. 2025. doi: [10.1007/s11227-025-07067-3](https://doi.org/10.1007/s11227-025-07067-3)
- [23] X. Fan et al., *Improvement of YOLOv8 for Vehicle Small Object Detection Research*, in *Proc. AIAHPC'24*, Zhuhai, Jul. 2024.doi: [10.1145/3690931.3690946](https://doi.org/10.1145/3690931.3690946)
- [24] Anonymous (2025), *Enhanced YOLOv8 for Small-Object Detection in Multiscale UAV Imagery: Innovations in Detection Accuracy and Efficiency*, *Digital Signal Processing*, vol. 158, art. no. 104964, Mar. 2025. doi: 10.1016/j.dsp.2024.104964
- [25] Roboflow, Soccer Dataset for Object Detection (*Referee, Goalkeeper, Player, Ball*).[Online].Available: <https://universe.roboflow.com/roboflow-jyuqo/football-players-detection-3zvbc>