Journal of Vocational Education and Information Technology
Vol. 7, No. 2, Desember 2026, 619 - 629
E-ISSN: 2722-5305
Terakreditasi **SINTA Peringkat 4**
Surat Keputusan Direktur Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Nomor: 230/E/KPT/2024

Analisis Pengaruh Data Augmentation terhadap Kinerja YOLOv8 dalam Deteksi Fraktur Tulang pada Citra X-Ray

Rahma Allysa Abelya[1], Faisal Muttaqin[2,*], Yisti Vita Via[3]

[1,2,3] Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional "Veteran" Jawa Timur; Jl. Rungkut Madya No.1, Gunung Anyar, Surabaya; e-mail: [1]22081010004@student.upnjatim.ac.id , [2]faisalmuttaqin.if@upnjatim.ac.id , [3]yistivia.if@upnjatim.ac.id
* Korespondensi: e-mail: faisalmuttaqin.if@upnjatim.ac.id

Diterima: 11-08--2026; Review: 09-09-2026; Disetujui: 13-10-2026

Abstrak: Fraktur tulang merupakan kondisi terputusnya kontinuitas struktur tulang yang umumnya diidentifikasi melalui citra X-Ray, namun proses identifikasi manual rentan terhadap kesalahan akibat variasi pola fraktur yang kompleks. Penelitian ini bertujuan menganalisis pengaruh penerapan data augmentation terhadap kinerja model YOLOv8 dalam mendeteksi fraktur tulang pada citra X-Ray menggunakan dataset FracAtlas sebanyak 1.434 citra (717 fraktur dan 717 non-fraktur) yang dibagi menjadi data latih, validasi, dan uji dengan rasio 70:15:15. Penelitian menggunakan lima skenario pelatihan, yaitu S1 tanpa augmentasi (baseline), S2 Horizontal Flip, S3 Rotation, S4 Brightness, dan S5 kombinasi ketiga teknik. Setiap skenario dilatih menggunakan model YOLOv8n dengan konfigurasi epoch, ukuran citra, dan batch size yang sama, kemudian dievaluasi menggunakan metrik Precision, Recall, mAP@0.5, dan mAP@0.5:0.95. Hasil pengujian menunjukkan bahwa skenario baseline (S1) memperoleh Precision tertinggi sebesar 0,7047, sedangkan skenario augmentasi tunggal (S2, S3, S4) justru menurunkan nilai Precision dibandingkan baseline. Skenario S5 yang menggabungkan Horizontal Flip, Rotation, dan Brightness memberikan performa terbaik secara keseluruhan dengan mAP@0.5 sebesar 0,3829 dan mAP@0.5:0.95 sebesar 0,1449, meningkat masing-masing sebesar 0,0542 dan 0,0191 dibandingkan baseline. Hasil ini menunjukkan bahwa kombinasi beberapa teknik augmentasi lebih efektif dalam meningkatkan performa model pada data uji (testing set) dibandingkan penerapan augmentasi tunggal, meskipun nilai Recall pada seluruh skenario masih tergolong rendah (0,29-0,35). Penelitian ini diharapkan dapat menjadi acuan dalam pemilihan strategi augmentasi data yang tepat untuk meningkatkan performa model deteksi objek pada citra medis.

Kata kunci: YOLOv8, deteksi fraktur, data augmentation, FracAtlas, citra X-Ray.

Abstract: *A bone fracture is a discontinuity in bone structure that is commonly identified through X-ray images, although manual identification remains prone to error due to the complex variation of fracture patterns. This study analyzes the effect of data augmentation on the performance of the YOLOv8 model in detecting bone fractures in X-ray images using the FracAtlas dataset of 1,434 images (717 fracture and 717 non-fracture) split into training, validation, and testing data with a ratio of 70:15:15. Five training scenarios were used, namely S1 without augmentation (baseline), S2 Horizontal Flip, S3 Rotation, S4 Brightness, and S5 a combination of the three techniques. Each scenario was trained using the YOLOv8n model with the same epoch, image size, and batch size configuration, then evaluated using Precision, Recall, mAP@0.5, and mAP@0.5:0.95 metrics. The test results show that the baseline scenario (S1) obtained the highest Precision at 0.7047, while the single-augmentation scenarios (S2, S3, S4) reduced Precision compared to the baseline. Scenario S5, which combined Horizontal Flip, Rotation, and Brightness, achieved the best overall performance with an mAP@0.5 of 0.3829 and mAP@0.5:0.95 of 0.1449, an increase of 0.0542 and 0.0191 respectively over the baseline. These results indicate that combining several augmentation techniques is more effective at improving the model's model's performance on the testing set than applying a single augmentation, although the Recall value across all scenarios remained relatively low (0.29-0.35).*

Copyright@2026. LEMBAGA PENGEMBANGAN DAN INOVASI UNDHARI

This study is expected to serve as a reference for selecting appropriate data augmentation strategies to improve object detection model performance on medical images.

Keywords: *YOLOv8, fracture detection, data augmentation, FracAtlas, X-ray image.*

1. Pendahuluan

Fraktur tulang merupakan kondisi terputusnya kontinuitas struktur tulang yang dapat terjadi akibat trauma, tekanan berlebih, maupun kondisi patologis tertentu [1], [2]. Pemeriksaan radiografi atau citra X-Ray menjadi salah satu metode yang umum digunakan untuk mengidentifikasi adanya fraktur karena mampu memberikan gambaran struktur tulang pada bagian tubuh yang diperiksa [3]. Namun, identifikasi lokasi fraktur pada citra X-Ray dapat menjadi tantangan karena karakteristik fraktur yang beragam serta adanya pola fraktur dengan perubahan visual yang relatif kecil [4]. Kondisi tersebut mendorong pengembangan teknologi berbasis kecerdasan buatan untuk membantu proses analisis citra medis secara otomatis.Perkembangan *deep learning* telah memberikan kontribusi dalam pengolahan citra medis karena kemampuannya dalam mempelajari dan mengekstraksi karakteristik citra secara otomatis [5]. Pendekatan *deep learning* telah digunakan untuk menganalisis citra X-Ray dan mengidentifikasi fraktur tulang. Salah satu pendekatan yang dapat digunakan adalah *object detection*, yaitu metode yang tidak hanya mengidentifikasi keberadaan objek, tetapi juga menentukan lokasi objek melalui *bounding box* [2]. Pendekatan tersebut relevan dalam deteksi fraktur karena sistem diharapkan mampu menunjukkan lokasi area fraktur pada citra X-Ray.

Salah satu algoritma *object detection* yang berkembang adalah *You Only Look Once* (YOLO). YOLO melakukan proses deteksi objek dalam satu tahap sehingga mampu melakukan klasifikasi dan lokalisasi objek secara langsung [6], [7]. YOLOv8 merupakan salah satu pengembangan algoritma YOLO yang telah diterapkan pada berbagai tugas deteksi objek, termasuk deteksi fraktur pada citra radiografi [6], [8]. Meskipun demikian, performa model *deep learning* tidak hanya dipengaruhi oleh arsitektur yang digunakan, tetapi juga oleh kualitas dan keragaman data pelatihan. Pada bidang pencitraan medis, ketersediaan dataset beranotasi umumnya masih terbatas karena proses anotasi membutuhkan waktu yang lama serta melibatkan tenaga ahli [9]. Keterbatasan jumlah dan variasi data tersebut dapat menyebabkan model mengalami *overfitting* dan menurunkan kemampuan generalisasi ketika dihadapkan pada citra baru [10]. Salah satu pendekatan yang banyak digunakan untuk mengatasi permasalahan tersebut adalah data augmentation, yaitu teknik yang menghasilkan variasi data pelatihan melalui transformasi tertentu tanpa mengubah informasi utama pada objek [11], [12]. Penerapan data augmentation diharapkan mampu meningkatkan keberagaman data sehingga model dapat mempelajari representasi fitur yang lebih robust. Berbagai teknik data augmentation, seperti *Horizontal Flip*, *Rotation*, dan penyesuaian *Brightness*, telah banyak diterapkan pada penelitian berbasis *computer vision* [12]. Namun demikian, efektivitas setiap teknik augmentasi dapat berbeda bergantung pada karakteristik dataset dan objek yang dideteksi. Pada citra X-Ray muskuloskeletal, transformasi tertentu berpotensi meningkatkan kemampuan model dalam mengenali pola fraktur, sedangkan transformasi lainnya justru dapat mengurangi informasi penting yang dibutuhkan selama proses pembelajaran [10], [13]. Namun penelitian yang membahas pengaruh data augmentasi pada citra medis umumnya belum secara khusus membandingkan efektivitas penerapan augmentasi tunggal dan kombinasi augmentasi untuk kasus deteksi fraktur tulang pada citra x-ray menggunakan algoritma YOLOv8, sehingga sejauh mana masing-masing pedekatan augmentasi tersebut memngaruhi performa deteksi fraktur pada kasus ini belum banyak diketahui secara pasti. Oleh karena itu, penelitian ini diperlukan untuk menganalisis secara lebih mendalam pengaruh penerapan augmentasi secara tunggal maupun kombinasi terhadap performa model YOLOv8 dalam mendeteksi fraktur tulang pada citra x-ray.
Berdasarkan permasalahan tersebut, penelitian ini bertujuan menganalisis pengaruh penerapan berbagai teknik data augmentation terhadap kinerja model YOLOv8 dalam mendeteksi fraktur tulang pada citra X-Ray menggunakan dataset FracAtlas. Penelitian menggunakan lima skenario pelatihan yang terdiri atas satu skenario tanpa augmentasi sebagai *baseline* serta empat skenario dengan teknik augmentasi yang berbeda, yaitu *Horizontal Flip*, *Rotation*, *Brightness*, dan kombinasi ketiganya. Seluruh skenario dilatih menggunakan konfigurasi pelatihan yang sama sehingga perbedaan performa yang diperoleh dapat dianalisis berdasarkan teknik augmentasi yang diterapkan. Kinerja setiap model kemudian dievaluasi menggunakan metrik Precision,

Vol.7, No. 2, Desember 2026, 619 - 629

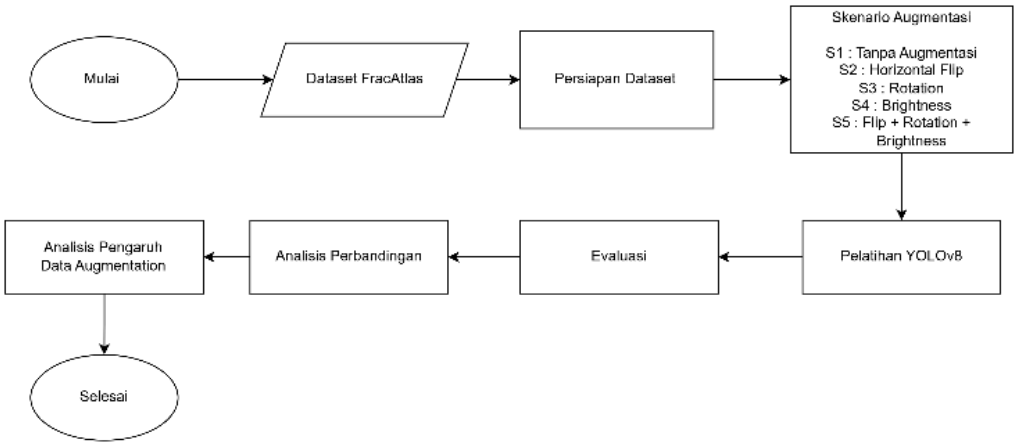
Recall, mAP@0.5, dan mAP@0.5:0.95 untuk mengidentifikasi teknik augmentasi yang memberikan performa terbaik dalam mendeteksi fraktur tulang pada dataset FracAtlas.

2. Metode Penelitian

Pada penelitian ini dilakukan serangkaian tahapan yang meliputi persiapan dataset, pembagian data, pelatihan model YOLOv8, serta evaluasi performa model. Tahapan penelitian diawali dengan pengumpulan dataset FracAtlas, kemudian dilakukan penyesuaian format anotasi ke format YOLO. Selanjutnya dataset dibagi menjadi data latih, validasi, dan pengujian sebelum digunakan dalam proses pelatihan model. Setelah proses pelatihan selesai, model dievaluasi menggunakan metrik *Precision*, *Recall*, *mean Average Precision* (mAP@0.5), dan mAP@0.5:0.95 untuk mengetahui kemampuan model dalam mendeteksi lokasi fraktur pada citra X-Ray.

2.1 Alur Penelitian

Penelitian ini dilakukan melalui beberapa tahapan yang disusun secara sistematis, mulai dari pengumpulan dataset hingga evaluasi model. Tahapan tersebut meliputi persiapan dataset, pelatihan model YOLOv8, serta evaluasi performa model dalam mendeteksi lokasi fraktur pada citra X-Ray. Alur penelitian secara keseluruhan ditunjukkan pada Gambar 1.

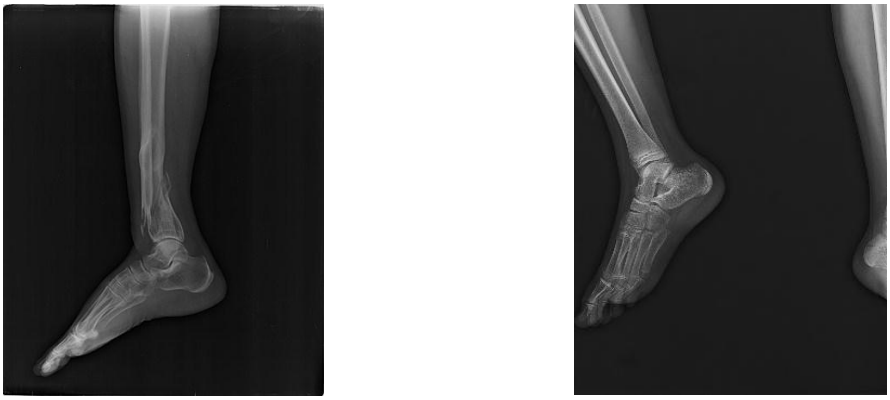


Gambar 1. Alur Penelitian

Berdasarkan Gambar 1, penelitian diawali dengan pengumpulan dataset FracAtlas yang kemudian dipersiapkan melalui proses seleksi data, konversi anotasi dari format COCO ke format YOLO, dan pembagian dataset menjadi data latih, validasi, serta pengujian. Dataset yang telah dipersiapkan selanjutnya digunakan pada proses pelatihan model YOLOv8 untuk menghasilkan model deteksi fraktur. Model yang telah dilatih kemudian dievaluasi menggunakan metrik *Precision*, *Recall*, *mean Average Precision* pada ambang *Intersection over Union* (IoU) 0,5 (mAP@0.5), dan mAP@0.5:0.95. Tahap terakhir adalah visualisasi hasil deteksi untuk menunjukkan kemampuan model dalam mengidentifikasi lokasi fraktur pada citra X-Ray.

2.2 Dataset

Dataset yang digunakan dalam penelitian ini adalah FracAtlas yang diperoleh dari platform Kaggle, yaitu dataset citra radiografi muskuloskeletal yang dirancang untuk mendukung penelitian deteksi fraktur. Dataset ini menyediakan citra X-Ray yang telah dilengkapi dengan anotasi *bounding box* pada area fraktur sehingga dapat digunakan untuk tugas *object detection*. FracAtlas mencakup berbagai bagian anatomi tubuh dengan variasi karakteristik fraktur, sehingga sesuai digunakan sebagai dataset untuk pengembangan dan evaluasi model deteksi fraktur.



(a) Citra dengan Fraktur (b) Citra tanpa Fraktur

Gambar 2. Contoh Dataset FracAtlas

Berdasarkan Gambar 2, dataset FracAtlas terdiri atas citra fraktur yang dilengkapi anotasi *bounding box* untuk menunjukkan lokasi fraktur serta citra non-fraktur yang tidak memiliki anotasi. Variasi lokasi anatomi, ukuran, dan bentuk fraktur pada dataset ini memberikan tantangan tersendiri dalam proses pelatihan model deteksi objek. Pada penelitian ini digunakan sebanyak 1.434 citra X-Ray yang terdiri atas 717 citra fraktur dan 717 citra non-fraktur. Pemilihan jumlah citra yang seimbang antara kedua kelas dilakukan untuk mengurangi potensi ketidakseimbangan data (*class imbalance*) selama proses pelatihan model. Selanjutnya, dataset dibagi menjadi tiga bagian, yaitu data latih (*training*), data validasi (*validation*), dan data pengujian (*testing*) dengan rasio 70:15:15. Rincian pembagian dataset ditunjukkan pada Tabel 1.

Tabel 1. Pembagian Dataset

Dataset	Fraktur	Non-Fraktur	Total
Training	501	502	1003
Validation	108	107	215
Testing	108	108	216
Total	717	717	1434

2.3 Augmentasi Data

Augmentasi data merupakan teknik prapemrosesan yang digunakan untuk meningkatkan variasi data latih dengan menghasilkan citra baru melalui transformasi tertentu tanpa mengubah informasi utama pada objek [12]. Pada penelitian ini, augmentasi diterapkan hanya pada data latih (*training set*), sedangkan data validasi dan data pengujian tidak diberikan augmentasi agar proses evaluasi dilakukan menggunakan data asli. Penelitian ini menggunakan tiga jenis augmentasi, yaitu Horizontal Flip, Rotation, dan Brightness. *Horizontal Flip* digunakan untuk menghasilkan citra hasil pencerminan secara horizontal sehingga model dapat mempelajari variasi orientasi objek. *Rotation* diterapkan untuk mensimulasikan perubahan sudut pengambilan citra, sedangkan *Brightness* digunakan untuk mensimulasikan variasi intensitas pencahayaan yang mungkin terjadi pada proses akuisisi citra X-Ray [10], [14].



Gambar 3. Contoh Hasil Penerapan Data Augmentation

Berdasarkan Gambar 3, setiap teknik *data augmentation* menghasilkan perubahan karakteristik visual yang berbeda pada citra X-Ray. Skenario S1 menggunakan citra asli tanpa augmentasi

sebagai *baseline*. Skenario S2 menghasilkan citra hasil pencerminan secara horizontal, S3 menghasilkan citra dengan rotasi sesuai parameter yang ditentukan, S4 menghasilkan perubahan intensitas kecerahan citra, sedangkan S5 merupakan kombinasi dari ketiga teknik augmentasi tersebut. Seluruh transformasi dilakukan tanpa menghilangkan informasi penting pada objek sehingga tetap sesuai digunakan sebagai data pelatihan model deteksi fraktur. Untuk menganalisis pengaruh augmentasi terhadap kinerja model YOLOv8, penelitian ini menyusun lima skenario pelatihan yang terdiri atas satu skenario tanpa augmentasi dan empat skenario dengan kombinasi augmentasi yang berbeda. Rincian skenario augmentasi ditunjukkan pada Tabel 2.

Tabel 2. Skenario Augmentasi Data

Skenario	Horizontal Flip	Rotation	Brightness
S1	–	–	–
S2	✓	–	–
S3	–	✓	–
S4	–	–	✓
S5	✓	✓	✓

Setiap teknik augmentasi diterapkan menggunakan parameter yang telah ditentukan berdasarkan konfigurasi pelatihan pada framework Ultralytics YOLOv8. Nilai parameter yang digunakan pada penelitian ini ditunjukkan pada Tabel 3.

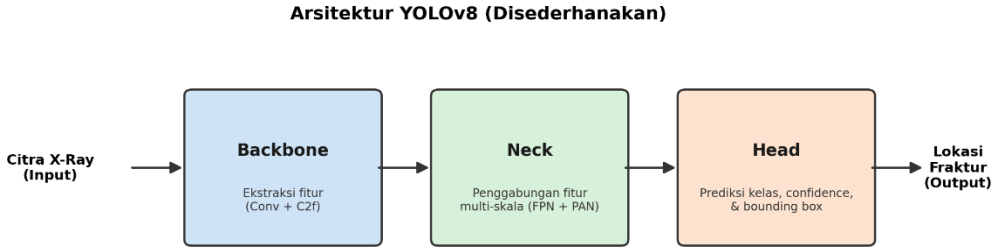
Tabel 3. Parameter Augmentasi

Parameter	Nilai
Horizontal Flip (fliplr)	0,5
Rotation (degrees)	10°
Brightness (hsv_v)	0,4

Berdasarkan skenario pada Tabel 2, S1 digunakan sebagai model pembanding tanpa augmentasi (*baseline*). Skenario S2, S3, dan S4 masing-masing menerapkan satu jenis augmentasi untuk mengamati pengaruh setiap teknik secara individual terhadap performa model. Selanjutnya, skenario S5 mengombinasikan ketiga teknik augmentasi untuk mengevaluasi pengaruh penerapan beberapa augmentasi secara bersamaan terhadap kemampuan YOLOv8 dalam mendeteksi lokasi fraktur pada citra X-Ray.

2.4 YOLOv8

YOLOv8 (*You Only Look Once version 8*) merupakan algoritma *one-stage object detection* yang dikembangkan oleh Ultralytics untuk mendeteksi objek secara cepat dan akurat. Berbeda dengan pendekatan *two-stage detector* yang memisahkan proses *region proposal* dan klasifikasi objek, YOLOv8 melakukan ekstraksi fitur, klasifikasi objek, serta prediksi *bounding box* dalam satu kali proses inferensi [6]. Pendekatan ini memungkinkan model menghasilkan waktu komputasi yang lebih efisien dengan tetap mempertahankan performa deteksi yang tinggi. Pada penelitian ini digunakan varian YOLOv8n (Nano) sebagai model deteksi fraktur tulang. Varian ini dipilih karena memiliki jumlah parameter yang lebih sedikit dibandingkan varian YOLOv8 lainnya, sehingga membutuhkan sumber daya komputasi yang lebih rendah dengan tetap memberikan performa deteksi yang baik. Selain itu, penggunaan YOLOv8n memungkinkan proses pelatihan dilakukan secara efisien pada setiap skenario *data augmentation* yang diterapkan [7], [14]. Arsitektur YOLOv8 secara umum terdiri atas tiga komponen utama, yaitu Backbone, Neck, dan Head [7], [15], sebagaimana ditunjukkan pada Gambar 4. Backbone berfungsi mengekstraksi fitur dari citra masukan menggunakan serangkaian lapisan konvolusi serta modul C2f untuk menghasilkan representasi fitur yang lebih kaya. Selanjutnya, Neck menggabungkan fitur dari berbagai tingkat resolusi melalui mekanisme Feature Pyramid Network (FPN) dan Path Aggregation Network (PAN) sehingga informasi objek pada berbagai skala dapat dipertahankan. Terakhir, Head menghasilkan prediksi berupa kelas objek, nilai *confidence score*, dan koordinat *bounding box* pada beberapa skala deteksi sehingga model mampu mengenali objek berukuran kecil, sedang, maupun besar [6], [7], [15].



Gambar 4. Arsitektur YOLOv8 (Disederhanakan)
Sumber: Disederhanakan oleh penulis, diadaptasi dari [6]

Berdasarkan Gambar 4, proses deteksi diawali dengan citra masukan yang diproses oleh Backbone untuk mengekstraksi fitur pada beberapa tingkat kedalaman. Fitur tersebut kemudian diteruskan ke Neck melalui proses *upsampling* dan *concatenation* sehingga informasi dari berbagai skala dapat digabungkan sebelum dikirim ke Head. Selanjutnya, Head menghasilkan prediksi akhir berupa kelas objek, tingkat kepercayaan (*confidence score*), dan koordinat *bounding box* yang digunakan untuk menentukan lokasi fraktur pada citra X-Ray [6]. Pada penelitian ini, model YOLOv8n dilatih menggunakan dataset FracAtlas pada lima skenario *data augmentation* yang berbeda. Seluruh skenario menggunakan konfigurasi pelatihan yang sama sehingga perbedaan performa model dapat dianalisis berdasarkan teknik augmentasi yang diterapkan. Model terbaik (*best.pt*) dari setiap skenario kemudian dievaluasi menggunakan metrik Precision, Recall, mAP@0.5, dan mAP@0.5:0.95 untuk mengetahui pengaruh *data augmentation* terhadap kinerja YOLOv8 dalam mendeteksi fraktur tulang pada citra X-Ray.

2.5 Pelatihan Model

Proses pelatihan model dilakukan menggunakan Ultralytics YOLOv8 dengan bahasa pemrograman Python pada platform Google Colaboratory yang memanfaatkan GPU NVIDIA A100-SXM4-40GB untuk mempercepat komputasi. Pelatihan dilakukan secara terpisah pada setiap skenario augmentasi yang telah dijelaskan pada subbab sebelumnya. Seluruh skenario menggunakan konfigurasi pelatihan yang sama sehingga perbedaan performa model hanya dipengaruhi oleh teknik augmentasi yang diterapkan.

Parameter pelatihan yang digunakan dalam penelitian ini ditunjukkan pada Tabel 4.

Tabel 4. Parameter Pelatihan Model

Parameter	Nilai
Model	YOLOv8n
Framework	Ultralytics YOLOv8
Bahasa Pemrograman	Python 3.12
Epoch	100
Ukuran Citra (Image Size)	640 × 640 piksel
Batch Size	16
Optimizer	AdamW (auto, momentum 0,9)
Learning Rate	0,002
Perangkat Komputasi	GPU NVIDIA A100-SXM4-40GB

Selama proses pelatihan, model mempelajari karakteristik citra X-Ray dan lokasi fraktur berdasarkan anotasi *bounding box* pada data latih. Pada setiap *epoch*, model divalidasi menggunakan data validasi untuk memantau proses pembelajaran dan memperoleh model dengan kemampuan generalisasi yang baik. Setelah seluruh proses pelatihan selesai, model terbaik (*best.pt*) dari masing-masing skenario digunakan pada tahap evaluasi menggunakan data uji.

2.6 Evaluasi Model

Evaluasi model dilakukan untuk mengukur kemampuan YOLOv8 dalam mendeteksi lokasi fraktur pada citra X-Ray. Proses evaluasi menggunakan data uji (*testing set*) yang tidak digunakan selama proses pelatihan maupun validasi sehingga hasil evaluasi dapat menggambarkan kemampuan generalisasi model terhadap data yang belum pernah dilihat sebelumnya.

Setiap model hasil pelatihan pada lima skenario augmentasi dievaluasi menggunakan metrik *Precision*, *Recall*, *mean Average Precision* pada ambang *Intersection over Union* (IoU) 0,5 (*mAP@0.5*), dan *mean Average Precision* pada rentang IoU 0,5 hingga 0,95 (*mAP@0.5:0.95*). Keempat metrik tersebut dipilih karena merupakan metrik yang umum digunakan untuk mengevaluasi performa model *object detection*. *Precision* menunjukkan kemampuan model dalam menghasilkan prediksi yang benar dari seluruh objek yang terdeteksi, sedangkan *Recall* mengukur kemampuan model dalam mendeteksi seluruh objek fraktur yang terdapat pada data uji. Nilai *mAP@0.5* digunakan untuk mengukur akurasi deteksi pada ambang IoU sebesar 0,5, sementara *mAP@0.5:0.95* memberikan evaluasi yang lebih komprehensif dengan menghitung rata-rata nilai *Average Precision* pada beberapa ambang IoU mulai dari 0,5 hingga 0,95. Hasil evaluasi dari setiap skenario augmentasi kemudian dibandingkan untuk menganalisis pengaruh penerapan teknik augmentasi terhadap performa model YOLOv8 dalam mendeteksi lokasi fraktur pada citra X-Ray.

3. Hasil dan Pembahasan

Bab ini menyajikan hasil pelatihan dan evaluasi model YOLOv8 pada setiap skenario augmentasi data. Selanjutnya, hasil tersebut dianalisis untuk mengetahui pengaruh augmentasi terhadap performa model dalam mendeteksi fraktur pada citra X-Ray.

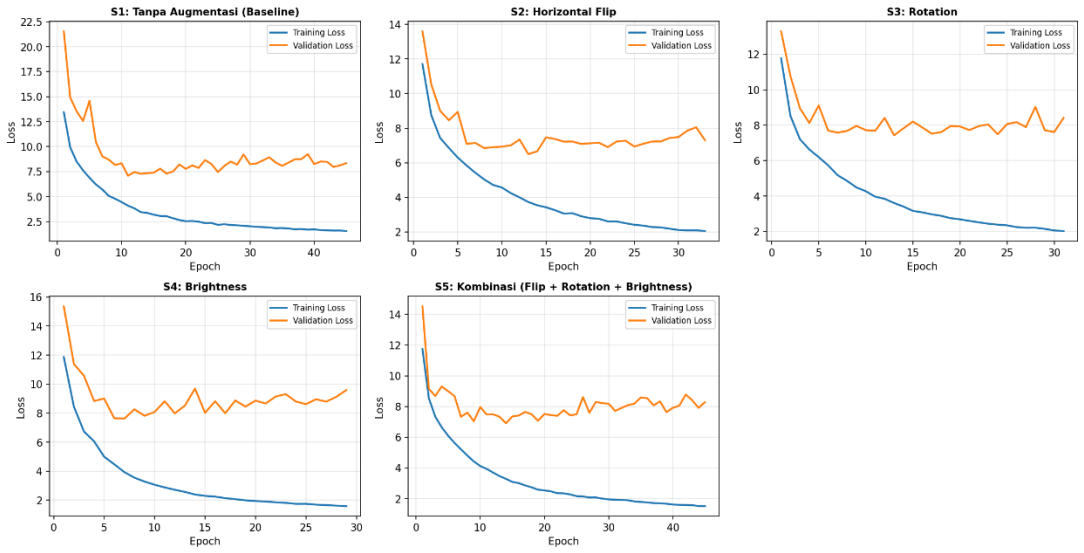
3.1 Hasil Pelatihan Model

Pelatihan model dilakukan pada lima skenario augmentasi data menggunakan konfigurasi pelatihan yang sama sebagaimana dijelaskan pada Bab 2. Setiap skenario menghasilkan satu model terbaik (*best.pt*) yang selanjutnya digunakan pada tahap evaluasi menggunakan data uji. Ringkasan hasil pelatihan masing-masing skenario ditunjukkan pada Tabel 5.

Tabel 5. Ringkasan Hasil Pelatihan Model

Skenario	Status Pelatihan	Model Terbaik	Jumlah Epoch
S1	Selesai (early stopping)	best.pt (epoch ke-25)	45
S2	Selesai (early stopping)	best.pt (epoch ke-13)	33
S3	Selesai (early stopping)	best.pt (epoch ke-11)	31
S4	Selesai (early stopping)	best.pt (epoch ke-9)	29
S5	Selesai (early stopping)	best.pt (epoch ke-25)	45

Perkembangan proses pelatihan pada setiap skenario ditunjukkan melalui grafik *training loss* dan *validation loss* sebagaimana ditampilkan pada Gambar 5.



Gambar 5. Grafik Training Loss dan Validation Loss per Skenario Augmentasi

Berdasarkan hasil pelatihan, model terbaik dari setiap skenario selanjutnya digunakan pada tahap evaluasi untuk membandingkan pengaruh teknik augmentasi terhadap performa YOLOv8 dalam mendeteksi fraktur pada citra X-Ray.

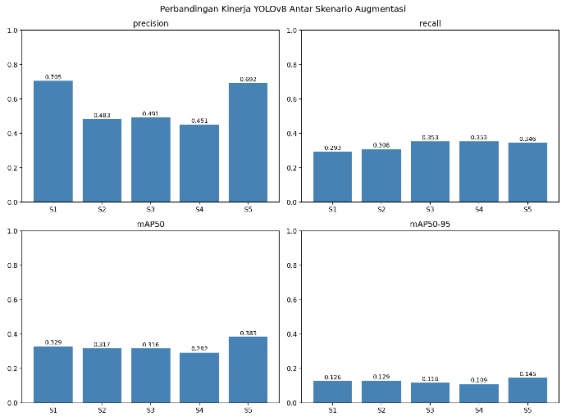
3.2 Hasil Evaluasi Model

Evaluasi model dilakukan menggunakan data uji (*testing set*) untuk mengukur kemampuan YOLOv8 dalam mendeteksi lokasi fraktur pada citra X-Ray. Setiap model yang dihasilkan dari lima skenario augmentasi dievaluasi menggunakan metrik *Precision*, *Recall*, *mean Average Precision* pada ambang *Intersection over Union* (IoU) 0,5 (*mAP@0.5*), dan *mean Average Precision* pada rentang IoU 0,5 hingga 0,95 (*mAP@0.5:0.95*). Hasil evaluasi masing-masing skenario ditunjukkan pada Tabel 6.

Tabel 6. Hasil Evaluasi Model

Skenario	Precision	Recall	mAP@0.5	mAP@0.5:0.95
S1	0.704727	0.293233	0.328690	0.125839
S2	0.483046	0.308271	0.316593	0.128715
S3	0.490569	0.353383	0.315915	0.116194
S4	0.451314	0.353383	0.291722	0.109290
S5	0.691572	0.345865	0.382864	0.144924

Untuk mempermudah perbandingan performa antar skenario augmentasi, hasil evaluasi juga disajikan dalam bentuk grafik sebagaimana ditunjukkan pada Gambar 6.



Gambar 6. Perbandingan Hasil Evaluasi Model

Berdasarkan hasil evaluasi pada Tabel 6 dan Gambar 6, setiap skenario augmentasi akan dibandingkan untuk mengetahui pengaruhnya terhadap performa YOLOv8 dalam mendeteksi fraktur pada citra X-Ray. Analisis lebih lanjut mengenai pengaruh masing-masing teknik augmentasi terhadap nilai *Precision*, *Recall*, *mAP@0.5*, dan *mAP@0.5:0.95* disajikan pada subbab berikutnya.

3.3 Analisis Pengaruh Augmentasi Data

Berdasarkan hasil evaluasi pada Tabel 6 dan Gambar 6, dilakukan analisis terhadap pengaruh penerapan teknik augmentasi data terhadap performa model YOLOv8 dalam mendeteksi fraktur pada citra X-Ray. Analisis difokuskan pada perubahan nilai *Precision*, *Recall*, *mAP@0.5*, dan *mAP@0.5:0.95* pada setiap skenario augmentasi dibandingkan dengan model tanpa augmentasi (*baseline*). Hasil evaluasi pada Tabel 6 menunjukkan bahwa skenario S1 (tanpa augmentasi) memperoleh nilai Precision tertinggi, yaitu sebesar 0,7047. Penerapan augmentasi tunggal, yaitu S2 (Horizontal Flip), S3 (Rotation), dan S4 (Brightness), justru menurunkan nilai Precision masing-masing sebesar 0,2217, 0,2142, dan 0,2534 dibandingkan baseline. Kondisi ini menunjukkan bahwa penerapan satu jenis augmentasi secara tunggal meningkatkan variasi data latih sehingga model menghasilkan lebih banyak prediksi positif, termasuk prediksi yang kurang tepat (*false positive*), yang berdampak pada penurunan Precision. Meskipun demikian, ketiga skenario tersebut menunjukkan peningkatan Recall dibandingkan

Vol.7, No. 2, Desember 2026, 619 - 629

baseline, yaitu sebesar 0,0150 pada S2 serta 0,0602 pada S3 dan S4, yang mengindikasikan bahwa augmentasi tunggal membantu model mengenali lebih banyak objek fraktur meskipun disertai peningkatan kesalahan deteksi. Berbeda dengan skenario augmentasi tunggal, skenario S5 yang mengombinasikan Horizontal Flip, Rotation, dan Brightness menunjukkan hasil yang lebih seimbang. Nilai Precision pada S5 sebesar 0,6916 hanya mengalami penurunan tipis sebesar 0,0132 dibandingkan baseline, sementara Recall meningkat sebesar 0,0526. Skenario S5 juga memperoleh nilai mAP@0.5 dan mAP@0.5:0.95 tertinggi di antara seluruh skenario, yaitu masing-masing sebesar 0,3829 dan 0,1449, atau meningkat sebesar 0,0542 dan 0,0191 dibandingkan baseline. Hasil tersebut mengindikasikan bahwa kombinasi beberapa teknik augmentasi lebih efektif dibandingkan penerapan augmentasi tunggal, karena variasi orientasi objek dari Horizontal Flip dan Rotation yang digabungkan dengan variasi intensitas pencahayaan dari Brightness membantu model mempelajari representasi fitur fraktur yang lebih beragam tanpa mengorbankan Precision secara signifikan, sehingga model mampu menyeimbangkan antara ketepatan dan kelengkapan deteksi.

Temuan ini sejalan dengan penelitian sebelumnya yang menerapkan YOLOv8 pada citra X-Ray, yang menunjukkan bahwa variasi data latih berperan penting dalam meningkatkan kemampuan model mendeteksi fraktur pada citra radiografi. Meskipun demikian, nilai Recall pada seluruh skenario dalam penelitian ini masih tergolong rendah, yaitu berkisar antara 0,29 hingga 0,35, yang menunjukkan bahwa model masih melewatkan cukup banyak objek fraktur pada data uji. Kondisi ini kemungkinan dipengaruhi oleh jumlah data latih yang relatif terbatas serta karakteristik fraktur pada citra X-Ray yang memiliki perubahan visual relatif kecil, sehingga diperlukan strategi tambahan, seperti augmentasi yang lebih bervariasi atau penggunaan varian YOLOv8 dengan kapasitas lebih besar, untuk meningkatkan kemampuan model dalam mendeteksi seluruh area fraktur secara lebih menyeluruh.

3.4 Visualisasi Hasil Deteksi

Visualisasi hasil deteksi dilakukan untuk memberikan gambaran mengenai kemampuan model YOLOv8 dalam mengidentifikasi lokasi fraktur pada citra X-Ray. Hasil deteksi ditampilkan dalam bentuk *bounding box* pada area yang diprediksi sebagai fraktur. Visualisasi ini digunakan sebagai pendukung hasil evaluasi kuantitatif sehingga performa model dapat diamati secara langsung. Salah satu contoh hasil deteksi pada data uji menggunakan model terbaik skenario S5 (kombinasi Horizontal Flip, Rotation, dan Brightness) ditunjukkan pada Gambar 7. Contoh ini digunakan untuk memberikan gambaran mengenai karakteristik keluaran model, termasuk lokasi bounding box dan tingkat keyakinan (*confidence score*) yang dihasilkan pada proses deteksi.



Gambar 7. Visualisasi Hasil Deteksi Fraktur Menggunakan YOLOv8

Pada Gambar 7, model S5 menghasilkan satu prediksi *bounding box* dengan label *fracture* dan *confidence score* sebesar 0,27 pada citra tungkai bawah yang menunjukkan kondisi pasca pemasangan implan ortopedi. Nilai confidence yang rendah tersebut menunjukkan bahwa meskipun model berhasil mengarahkan prediksi pada area yang berpotensi mengandung fraktur, tingkat keyakinan model terhadap prediksi tersebut masih rendah dan berada di bawah ambang confidence yang umum digunakan (0,5), sehingga prediksi ini berpotensi tidak lolos apabila diterapkan threshold yang lebih ketat. Kondisi ini konsisten dengan nilai Precision dan Recall skenario S5 yang meskipun tertinggi di antara seluruh skenario augmentasi, yaitu masing-masing sebesar 0,6916 dan 0,3459, masih menunjukkan bahwa model belum sepenuhnya optimal dalam menghasilkan prediksi dengan keyakinan tinggi. Pada citra yang sama, kaki sebelah kiri yang secara visual memiliki karakteristik serupa tidak menghasilkan prediksi bounding box sama sekali. Karena tidak tersedia anotasi ground truth pada visualisasi ini, tidak dapat dipastikan apakah kondisi tersebut disebabkan oleh ketiadaan fraktur pada kaki tersebut atau merupakan kasus false negative akibat keterbatasan model. Contoh ini menggambarkan bahwa meskipun skenario S5 memberikan performa terbaik secara kuantitatif, model masih menghasilkan deteksi dengan tingkat keyakinan yang bervariasi, sehingga diperlukan evaluasi visual pada lebih banyak citra uji untuk memperoleh gambaran yang lebih menyeluruh mengenai kemampuan dan keterbatasan model.

4. Kesimpulan

Penelitian ini menganalisis pengaruh penerapan data augmentation terhadap kinerja model YOLOv8n dalam mendeteksi fraktur tulang pada citra X-Ray menggunakan dataset FracAtlas melalui lima skenario pelatihan, yaitu S1 (tanpa augmentasi), S2 (Horizontal Flip), S3 (Rotation), S4 (Brightness), dan S5 (kombinasi ketiga teknik). Hasil evaluasi menunjukkan bahwa penerapan augmentasi secara tunggal (S2, S3, dan S4) cenderung menurunkan nilai Precision dibandingkan baseline meskipun disertai peningkatan Recall, sedangkan skenario kombinasi S5 memberikan performa terbaik secara keseluruhan dengan mAP@0.5 sebesar 0,3829 dan mAP@0.5:0.95 sebesar 0,1449, meningkat masing-masing sebesar 0,0542 dan 0,0191 dibandingkan baseline tanpa mengorbankan Precision secara signifikan. Temuan ini menunjukkan bahwa kombinasi beberapa teknik augmentasi, yaitu variasi orientasi objek melalui Horizontal Flip dan Rotation yang digabungkan dengan variasi intensitas pencahayaan melalui Brightness, lebih efektif dalam meningkatkan performa model YOLOv8 pada data uji dibandingkan penerapan augmentasi tunggal. Meskipun demikian, nilai Recall pada seluruh skenario masih tergolong rendah, sehingga model masih berpotensi melewatkan sejumlah area fraktur pada citra uji. Untuk penelitian selanjutnya, disarankan untuk menguji varian YOLOv8 dengan kapasitas lebih besar (seperti YOLOv8s atau YOLOv8m), menambah jumlah dan variasi data latih, serta mengeksplorasi kombinasi teknik augmentasi lain guna meningkatkan Recall dan performa deteksi fraktur secara keseluruhan pada citra X-Ray.

Referensi

[1] A. Aqeil, R. Afriyanto, A. Haikal, B. Shamsul, and K. Adzhar, “Comparison of Deep Neural Network and Convolutional Neural Network Algorithms for Bone Fracture,” *IJATIS: Indonesian Journal of Applied Technology and Innovation Science*, vol. 3, no. 1, pp. 10–18, 2026, doi: 10.57152/IJATIS.v3i1.2271.

[2] J. Jung, J. Dai, B. Liu, and Q. Wu, “Artificial intelligence in fracture detection with different image modalities and data types,” *PLOS Digital Health*, vol. 3, no. 1, Jan. 2024, doi: 10.1371/journal.pdig.0000438.

[3] A. dkk. Hassan, “Fracture Detection In X-rays Using Custom Convolutional Neural Network (CNN) And Transfer Learning Models,” Sep. 2025, [Online]. Available: http://arxiv.org/abs/2509.06228

[4] A. Tahir, A. Saadia, K. Khan, A. Gul, A. Qahmash, and R. N. Akram, “Enhancing diagnosis: ensemble deep-learning model for fracture detection using X-ray images,” *Clin. Radiol.*, vol. 79, no. 11, pp. e1394–e1402, Nov. 2024, doi: 10.1016/j.crad.2024.08.006.

[5] N. Puspita Sari, “Analisis Performa Algoritma CNN dalam Klasifikasi Citra Medis Berbasis Deep Learning Analysis Of CNN Algorithm In Deep Learning-Based Medical Image Classification,” *Jurnal Komputer*, vol. 2, no. 2, pp. 87–92, Jan. 2024.

[6] Juan Terven et. al, “A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS,” vol. 5, pp. 1680–1716, Nov. 2023.
[7] E. Panja, H. Hendry, and C. Dewi, “YOLOv8 Analysis for Vehicle Classification Under Various Image Conditions,” *Scientific Journal of Informatics*, vol. 11, no. 1, pp. 127–138, Feb. 2024, doi: 10.15294/sji.v11i1.49038.
[8] Z. Bao, “The UAV Target Detection Algorithm Based on Improved YOLO V8,” *International Conference on Image Processing, Machine Learning and Pattern Recognition (IPMLP 2024), September 1315, 2024, Guangzhou, China*, vol. 1, Sep. 2024, doi: 10.1145/3700906.
[9] Ph. D. Mada Sanjaya W. S., *Deep Learning Citra Medis Berbasis Pemrograman Python*. 2023.
[10] J. Nalepa, M. Marcinkiewicz, and M. Kawulok, “Data Augmentation for Brain-Tumor Segmentation: A Review,” Dec. 11, 2019, *Frontiers Media S.A.* doi: 10.3389/fncom.2019.00083.
[11] E. Goceri, “Medical image data augmentation: techniques, comparisons and interpretations,” *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 12561–12605, Nov. 2023, doi: 10.1007/s10462-023-10453-z.
[12] A. Kebaili, J. Lapuyade-Lahorgue, and S. Ruan, “Deep Learning Approaches for Data Augmentation in Medical Imaging: A Review,” Apr. 01, 2023, *MDPI*. doi: 10.3390/jimaging9040081.
[13] I. Abedeen, Md. A. Rahman, F. Z. Prottyasha, T. Ahmed, T. M. Chowdhury, and S. Shatabda, “A Dataset for Fracture Classification, Localization and Segmentation of Musculoskeletal Radiographs,” *Sci. Data*, vol. 10, no. 1, p. 521, Aug. 2023, doi: 10.1038/s41597-023-02432-4.
[14] S. D. Mutia, R. Firliana, and A. Nugroho, “Optimasi Model Yolov8n Menggunakan Augmentasi Data Untuk Peningkatan Akurasi Sistem Dress-Code Surveillance,” *Bulletin of Information Technology (BIT)*, vol. 7, no. 2, pp. 222–231, 2026, doi: 10.47065/bit.v5i2.2760.
[15] J. Wang *et al.*, “Road defect detection based on improved YOLOv8s model,” *Sci. Rep.*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-67953-3.