

Deteksi Edema Paru pada Chest X-ray Menggunakan YOLOv5n dan Grey Wolf Optimizer

Nasrudin Affandi Prasetyo¹, Cinantya Paramita², Amiq Fahmi³

^{1,2} Program Studi Teknik Informatika, Universitas Dian Nuswantoro, Indonesia

³ Fakultas Kedokteran, Universitas Dian Nuswantoro, Indonesia

¹11202214518@mhs.dinus.ac.id, ²cinantya.paramita@dsn.dinus.ac.id, ³amiq.fahmi@dsn.dinus.ac.id

Info Artikel

Riwayat Artikel:

Received 2026-03-11

Revised 2026-05-08

Accepted 2026-05-19

Abstract – Pulmonary edema is a lung disorder characterized by fluid accumulation in the alveolar and interstitial spaces, which disrupts gas exchange and reduces oxygen levels in the blood. Chest X-ray (CXr) imaging is commonly used for pulmonary edema assessment because it is fast and widely available; however, its interpretation still depends heavily on radiologist expertise and may lead to diagnostic variability, particularly in healthcare facilities with limited radiology resources. This study aims to develop an automated pulmonary edema detection system based on deep learning to support more consistent analysis of CXr images. The novelty of this study lies in the integration of Grey Wolf Optimizer (GWO) for hyperparameter optimization of the YOLOv5n model specifically for pulmonary edema detection in Chest X-ray images. The proposed method employs YOLOv5n as a lightweight object detection architecture because of its computational efficiency and suitability for resource-constrained environments. To improve detection performance and training stability, the hyperparameters of YOLOv5n are optimized using GWO. The model is trained and evaluated using annotated CXr images, and its performance is measured using precision, recall, mAP@0.5, and mAP@0.5–0.95. Experimental results show that the YOLOv5n + GWO model achieved a precision of 0.906, recall of 0.936, mAP@0.5 of 0.963, and mAP@0.5–0.95 of 0.737, indicating improved detection performance compared with the baseline YOLOv5n configuration. The proposed framework demonstrates potential as a decision-support tool for assisting medical personnel in early pulmonary edema screening through efficient and consistent analysis of CXr images.

Keywords: Chest X-ray, Deep Learning, Grey Wolf Optimizer, Pulmonary Edema, YOLOv5n.

Abstrak – Edema paru merupakan gangguan paru-paru yang ditandai dengan akumulasi cairan pada ruang alveolar dan interstitial, sehingga mengganggu pertukaran gas dan menurunkan kadar oksigen dalam darah. Pemeriksaan Chest X-ray (CXr) umum digunakan dalam penilaian edema paru karena cepat dan tersedia secara luas; namun, interpretasi citra CXr masih sangat bergantung pada keahlian radiolog sehingga dapat menimbulkan variasi diagnosis, terutama pada fasilitas kesehatan dengan keterbatasan sumber daya radiologi. Penelitian ini bertujuan untuk mengembangkan sistem deteksi edema paru otomatis berbasis deep learning guna mendukung analisis citra CXr yang lebih konsisten. Kebaruan penelitian ini terletak pada integrasi Grey Wolf Optimizer (GWO) untuk optimasi hyperparameter model YOLOv5n secara khusus pada deteksi edema paru berbasis citra Chest X-ray. Metode yang diusulkan menggunakan YOLOv5n sebagai arsitektur deteksi objek yang ringan karena efisien secara komputasi dan sesuai untuk lingkungan dengan sumber daya terbatas. Untuk meningkatkan performa deteksi dan stabilitas pelatihan, hyperparameter YOLOv5n dioptimasi menggunakan GWO. Model dilatih dan dievaluasi menggunakan citra CXr beranotasi, sedangkan performanya diukur menggunakan precision, recall, mAP@0.5, dan mAP@0.5–0.95. Hasil eksperimen menunjukkan bahwa model YOLOv5n + GWO memperoleh precision sebesar 0,906, recall sebesar 0,936, mAP@0.5 sebesar 0,963, dan mAP@0.5–0.95 sebesar 0,737, yang menunjukkan peningkatan performa deteksi dibandingkan konfigurasi baseline YOLOv5n. Kerangka kerja yang diusulkan menunjukkan potensi sebagai alat bantu keputusan bagi tenaga medis dalam skrining awal edema paru melalui analisis citra CXr yang efisien dan konsisten.

Kata Kunci: Chest X-ray, Deep Learning, Edema Paru, Grey Wolf Optimizer, YOLOv5n.

Corresponding Author:

Cinantya Paramita

Email:

cinantya.paramita@dsn.dinus.ac.id



This is an open access article under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license.

I. PENDAHULUAN

Pulmonary edema merupakan kondisi patologis pada paru-paru yang ditandai dengan akumulasi cairan pada ruang alveolar dan interstitial sehingga mengganggu proses pertukaran gas serta menurunkan kadar oksigen dalam darah [1]. Kondisi ini perlu mendapat perhatian karena dapat berkembang menjadi gangguan pernapasan akut apabila tidak dikenali sejak tahap awal. Oleh karena itu, identifikasi dini terhadap indikasi pulmonary edema menjadi penting dalam mendukung penanganan pasien secara cepat dan tepat. Salah satu pemeriksaan yang umum digunakan untuk mengetahui adanya indikasi pulmonary edema adalah Chest X-ray (CXr), karena citra CXr dapat memperlihatkan perubahan densitas jaringan paru akibat penumpukan cairan [1], [4]. Meskipun demikian, interpretasi citra CXr

membutuhkan keahlian tenaga medis, khususnya radiolog, sehingga hasil analisis masih dapat dipengaruhi oleh pengalaman, subjektivitas, dan ketelitian pemeriksa [6].

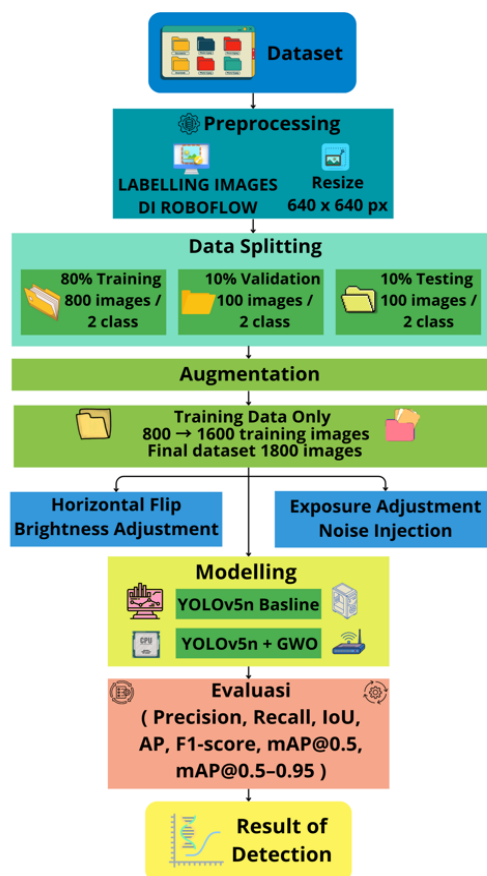
Permasalahan interpretasi citra CXR menjadi semakin kompleks pada fasilitas kesehatan dengan keterbatasan sumber daya. Tidak semua fasilitas kesehatan memiliki tenaga radiologi yang memadai, terutama pada wilayah terpencil atau fasilitas layanan kesehatan dengan infrastruktur terbatas. Keterbatasan tenaga medis dan fasilitas radiologi dapat menyebabkan proses pemeriksaan dan interpretasi citra menjadi lebih lambat, sehingga dibutuhkan teknologi pendukung yang mampu membantu analisis citra medis secara lebih cepat, objektif, dan konsisten. Dalam konteks tersebut, teknologi kecerdasan buatan dapat dimanfaatkan sebagai alat bantu untuk mendukung proses skrining awal pulmonary edema berbasis citra CXR, bukan untuk menggantikan peran tenaga medis, tetapi untuk membantu proses pengambilan keputusan klinis.

Perkembangan artificial intelligence, machine learning, dan deep learning telah membuka peluang besar dalam analisis citra medis secara otomatis. Berbagai penelitian telah menerapkan deep learning untuk deteksi dan analisis kelainan pada citra medis. Akbar et al. [1] menggunakan Siamese Convolutional Networks untuk memprediksi tingkat keparahan pulmonary edema pada citra Chest X-ray. Huynh et al. [2] mengembangkan model deep learning untuk penilaian radiografik pulmonary edema dengan dukungan biomarker klinis. Schulz et al. [26] menunjukkan bahwa deep learning dapat digunakan untuk memprediksi dan mengkuantifikasi pulmonary edema dari citra Chest X-ray. Selain itu, Choudhury et al. [12] membandingkan beberapa model CNN untuk klasifikasi tingkat keparahan pulmonary edema menggunakan citra Chest X-ray. Beberapa penelitian lain juga menerapkan YOLO untuk analisis citra medis, seperti deteksi lesi kecil pada Chest X-ray [9], segmentasi paru berbasis kombinasi SAM dan YOLO [14], abnormal target detection pada chest radiograph [21], serta strategi object detection untuk deteksi dan lokalisasi penyakit pada citra Chest X-ray [27]. Evaluasi pada penelitian-penelitian tersebut umumnya menggunakan metrik seperti precision, recall, AP, F1-score, dan mean Average Precision (mAP) untuk mengukur kualitas model deteksi atau klasifikasi [9], [12], [22], [27]. Namun, sebagian penelitian terdahulu lebih banyak berfokus pada klasifikasi, prediksi tingkat keparahan, segmentasi, atau deteksi kelainan CXR secara umum, sehingga kajian mengenai deteksi pulmonary edema berbasis bounding box menggunakan model YOLOv5n yang ringan masih relatif terbatas. Keterbatasan tersebut menunjukkan bahwa masih diperlukan penelitian yang tidak hanya berfokus pada klasifikasi atau prediksi tingkat keparahan, tetapi juga pada deteksi lokasi indikasi pulmonary edema secara langsung pada citra CXR. Pendekatan deteksi objek menjadi penting karena mampu memberikan informasi visual berupa bounding box yang dapat membantu interpretasi area target secara lebih jelas.

Berdasarkan kajian tersebut, penelitian ini mengusulkan penggunaan YOLOv5n sebagai model deteksi objek untuk mendeteksi pulmonary edema pada citra Chest X-ray. YOLOv5n dipilih karena merupakan varian YOLO yang ringan dan sesuai untuk pengembangan sistem deteksi dengan kebutuhan komputasi yang lebih efisien. Namun, performa model YOLO sangat dipengaruhi oleh konfigurasi hyperparameter seperti learning rate, momentum, weight decay, batch size, dan parameter pelatihan lainnya. Konfigurasi hyperparameter yang kurang optimal dapat menyebabkan performa deteksi menurun, proses pelatihan kurang stabil, atau model sulit mencapai hasil terbaik. Oleh karena itu, penelitian ini mengintegrasikan *Grey Wolf Optimizer* (GWO) sebagai metode optimasi hyperparameter untuk memperoleh konfigurasi parameter yang lebih optimal pada YOLOv5n [15], [16]. Research gap penelitian ini terletak pada masih terbatasnya kajian yang secara khusus mengintegrasikan YOLOv5n dan GWO untuk deteksi pulmonary edema pada citra Chest X-ray. Novelty penelitian ini adalah penerapan GWO untuk optimasi hyperparameter YOLOv5n pada kasus deteksi pulmonary edema berbasis CXR. Kualitas model dievaluasi menggunakan metrik precision, recall, IoU, AP, F1-score, mAP@0.5, dan mAP@0.5–0.95. Kontribusi utama penelitian ini meliputi: (1) pengembangan model deteksi pulmonary edema berbasis YOLOv5n pada citra Chest X-ray, (2) penerapan *Grey Wolf Optimizer* untuk optimasi hyperparameter YOLOv5n, dan (3) evaluasi performa model YOLOv5n baseline dan YOLOv5n + GWO menggunakan metrik evaluasi object detection.

II. METODE

Metode penelitian ini disusun untuk menjelaskan tahapan deteksi edema paru pada citra Chest X-ray menggunakan YOLOv5n dan *Grey Wolf Optimizer* (GWO). Tahapan penelitian meliputi pengumpulan dataset, preprocessing, pembagian data, augmentasi data, pemodelan menggunakan YOLOv5n baseline dan YOLOv5n + GWO, serta evaluasi performa model. Setiap tahapan dirancang secara sistematis agar proses pelatihan dan pengujian model dapat dilakukan secara terstruktur, konsisten, dan dapat direplikasi. Pendekatan ini digunakan untuk memastikan bahwa data yang digunakan telah melalui proses persiapan yang sesuai sebelum diterapkan pada model deteksi objek. Selain itu, perbandingan antara model baseline dan model hasil optimasi dilakukan untuk mengetahui pengaruh GWO terhadap peningkatan performa deteksi. Alur metode penelitian ditunjukkan pada Gambar 1.



Gambar 1. Flowchart metode penelitian

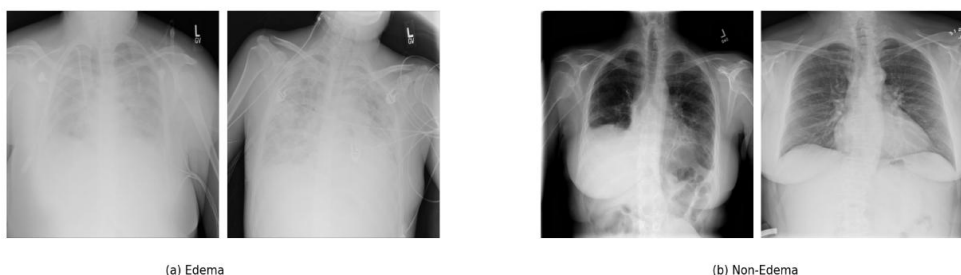
A. Dataset

Penelitian ini menggunakan dataset citra Chest X-ray (CXR) yang diperoleh dari platform Kaggle dengan judul *Pulmonary Edema Classified – By NIH (National Institutes of Health)*. Dataset terdiri atas 1.000 citra yang terbagi menjadi dua kelas, yaitu 500 citra edema dan 500 citra no-edema. Seluruh citra merupakan bagian dari NIH Chest X-ray Dataset yang telah diberi label berdasarkan laporan radiologi menggunakan pendekatan natural language processing.

Dataset asli tidak menyediakan anotasi lokasi edema, sehingga dilakukan pelabelan manual menggunakan Roboflow. Citra dengan label edema diberi anotasi bounding box pada area paru yang menunjukkan indikasi peningkatan densitas, sedangkan citra no-edema tidak diberi anotasi objek. Hasil pelabelan ini menghasilkan berkas koordinat bounding box yang digunakan sebagai data pelatihan model YOLOv5n.

Untuk menjaga kualitas anotasi, setiap bounding box diperiksa kembali secara visual berdasarkan kesesuaian posisi kotak terhadap area target. Anotasi yang terlalu melebar, tidak tepat, atau tidak mencakup area target diperbaiki sebelum digunakan dalam pelatihan. Pemeriksaan ulang ini dilakukan untuk meminimalkan kesalahan label karena kualitas anotasi berpengaruh terhadap kemampuan model dalam mendeteksi edema paru.

Citra Chest X-ray (CXR): Edema dan No Edema



(a) Edema

(b) Non-Edema

Gambar 2. Citra Chest X-ray (CXR) pada kelas edema dan no-edema

B. Preprocessing

Tahap preprocessing dilakukan untuk menyiapkan citra sebelum proses pelatihan model. Pada tahap ini, citra diperiksa orientasinya dan dilakukan auto-orientation untuk memastikan posisi anatomis gambar berada pada orientasi yang benar. Selanjutnya, seluruh citra diubah ukurannya menjadi 640×640 piksel sesuai dengan kebutuhan input YOLOv5n. Ukuran ini dipilih karena umum digunakan pada konfigurasi pelatihan YOLOv5 dan membantu menjaga konsistensi dimensi citra, mempercepat proses komputasi, serta mengurangi ketidaksesuaian ukuran data selama pelatihan. Tahap terakhir adalah normalisasi nilai piksel untuk mengurangi variasi intensitas yang tidak berkaitan langsung dengan pola visual edema paru, sehingga model dapat mempelajari fitur yang lebih relevan untuk proses deteksi.

C. Data Splitting dan Augmentation

Setelah melalui tahap preprocessing, dataset asli yang digunakan dalam penelitian ini terdiri dari 1.000 citra Chest X-ray. Dataset tersebut dibagi ke dalam tiga subset, yaitu training, validation, dan testing dengan komposisi 80% untuk training, 10% untuk validation, dan 10% untuk testing. Dengan demikian, sebanyak 800 citra digunakan sebagai data training, 100 citra digunakan sebagai data validation, dan 100 citra digunakan sebagai data testing. Pembagian ini bertujuan untuk memastikan bahwa data yang digunakan pada proses pelatihan, pemantauan performa, dan evaluasi akhir tidak saling tumpang tindih sehingga penilaian kinerja model dapat dilakukan secara objektif.

Data augmentation diterapkan hanya pada subset training untuk meningkatkan variasi data dan kemampuan generalisasi model. Teknik augmentasi yang digunakan meliputi horizontal flip, brightness adjustment, exposure adjustment, dan noise injection. Melalui proses ini, jumlah data training meningkat dari 800 citra menjadi 1.600 citra. Sementara itu, subset validation dan testing tidak mengalami proses augmentation dan masing-masing tetap berjumlah 100 citra. Hal ini dilakukan agar evaluasi model tetap menggunakan data asli yang tidak dimodifikasi. Dengan demikian, total dataset setelah proses augmentasi menjadi 1.800 citra.

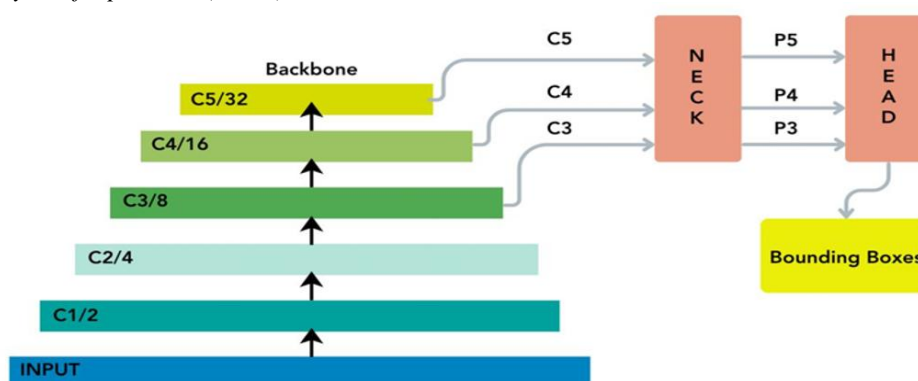
Subset training digunakan sebagai sumber pembelajaran utama model, sedangkan subset validation digunakan untuk memantau performa model selama proses pelatihan serta mendeteksi potensi overfitting. Adapun subset testing digunakan pada tahap akhir untuk mengevaluasi kemampuan generalisasi model terhadap data baru yang belum pernah digunakan sebelumnya.

D. Arsitektur YOLOv5n (Baseline)

Model baseline yang digunakan dalam penelitian ini adalah YOLOv5n, yaitu varian paling ringan dari keluarga *You Only Look Once* (YOLO). Model ini dipilih karena memiliki ukuran arsitektur yang kecil, kebutuhan komputasi yang rendah, serta kecepatan inferensi yang tinggi. Karakteristik tersebut menjadikan YOLOv5n sesuai untuk penelitian dengan dataset berukuran terbatas dan kebutuhan deteksi objek yang efisien pada citra Chest X-ray (CXR).

Secara umum, arsitektur YOLOv5n terdiri atas tiga komponen utama, yaitu backbone, neck, dan head. Bagian backbone berfungsi untuk mengekstraksi fitur penting dari citra masukan melalui lapisan konvolusi. Selanjutnya, bagian neck menggabungkan fitur dari berbagai skala untuk meningkatkan kemampuan model dalam mengenali objek dengan ukuran dan karakteristik visual yang beragam. Bagian head menghasilkan keluaran berupa koordinat bounding box, kelas objek, dan confidence score untuk setiap objek yang terdeteksi.

Pada tahap awal, YOLOv5n dilatih menggunakan konfigurasi hyperparameter standar sebagai baseline. Seluruh citra Chest X-ray yang telah melalui tahap preprocessing, dilengkapi anotasi bounding box, dan dibagi ke dalam subset training digunakan sebagai data pelatihan. Proses pelatihan dilakukan secara iteratif hingga nilai loss dan metrik evaluasi seperti precision, recall, mAP@0.5, dan mAP@0.5–0.95 menunjukkan kecenderungan konvergen. Hasil pelatihan baseline ini digunakan sebagai pembandingan terhadap model YOLOv5n yang dioptimasi menggunakan *Grey Wolf Optimizer* (GWO).



Gambar 3. Arsitektur YOLOv5n yang terdiri dari Backbone, Neck, dan Head

E. Modeling

Tahap modeling dilakukan dengan membangun dua skenario pemodelan, yaitu YOLOv5n baseline dan YOLOv5n dengan optimasi *Grey Wolf Optimizer* (GWO). Kedua skenario digunakan untuk membandingkan pengaruh optimasi hyperparameter terhadap performa deteksi pulmonary edema pada citra Chest X-ray (CXR).

Pada skenario pertama, model YOLOv5n dilatih menggunakan konfigurasi hyperparameter standar sebagai baseline. Seluruh citra yang telah melalui tahap preprocessing, pembagian dataset, dan anotasi bounding box digunakan dalam proses pelatihan dengan ukuran input 640×640 piksel. Hasil dari skenario baseline digunakan sebagai acuan pembandingan terhadap model yang telah dioptimasi.

Pada skenario kedua, dilakukan optimasi hyperparameter menggunakan algoritma GWO untuk mencari kombinasi parameter yang lebih optimal, seperti *learning rate*, momentum, dan weight decay. Parameter terbaik hasil optimasi kemudian digunakan untuk melatih kembali model YOLOv5n dengan prosedur pelatihan yang sama seperti skenario baseline.

Perbandingan hasil dari kedua skenario dilakukan menggunakan metrik evaluasi object detection, yaitu precision, recall, F1-score, mAP@0.5, dan mAP@0.5–0.95. Perbandingan ini bertujuan untuk mengetahui efektivitas optimasi GWO dalam meningkatkan performa deteksi pulmonary edema pada citra Chest X-ray.

F. Optimasi Hyperparameter menggunakan Grey Wolf Optimizer (GWO)

Pada skenario kedua, dilakukan optimasi hyperparameter menggunakan algoritma *Grey Wolf Optimizer* (GWO). Metode ini dipilih karena memiliki kemampuan pencarian solusi global melalui mekanisme exploration dan exploitation yang seimbang. Dalam penelitian ini, GWO digunakan untuk mengoptimalkan beberapa hyperparameter penting pada model YOLOv5n, yaitu learning rate, momentum, dan weight decay, yang berpengaruh terhadap stabilitas pelatihan dan kecepatan konvergensi model.

Proses optimasi dimulai dengan membentuk populasi awal yang merepresentasikan sejumlah kandidat kombinasi hyperparameter. Setiap kandidat dievaluasi menggunakan nilai fitness berdasarkan metrik mAP@0.5 pada data validation. Kandidat dengan nilai fitness terbaik digunakan sebagai acuan dalam pembaruan posisi kandidat lain pada setiap iterasi sesuai mekanisme pergerakan *Grey Wolf Optimizer*.

Setelah proses iterasi selesai, diperoleh kombinasi hyperparameter dengan nilai fitness tertinggi. Kombinasi parameter tersebut kemudian digunakan sebagai konfigurasi baru pada proses pelatihan YOLOv5n + GWO. Nilai hyperparameter terbaik hasil optimasi GWO ditampilkan pada Tabel I agar proses eksperimen dapat direplikasi.

TABEL 1
HYPERPARAMETER TERBAIK HASIL OPTIMASI GWO

No.	Hyperparameter	Nilai
1.	Learning rate	0,002936
2.	Momentum	0,600
3.	Weight decay	0,010
4.	Batch size	15
5.	Epoch	50
6.	Image size	640 × 640

G. Evaluasi Model

Evaluasi model dilakukan untuk menilai kemampuan sistem dalam mendeteksi pulmonary edema pada citra Chest X-ray (CXR) secara kuantitatif. Pada tahap ini, dua skenario pemodelan dibandingkan, yaitu YOLOv5n baseline dan YOLOv5n dengan konfigurasi hyperparameter hasil optimasi *Grey Wolf Optimizer* (GWO). Evaluasi dilakukan menggunakan data testing yang tidak dilibatkan selama proses pelatihan agar hasil yang diperoleh dapat merepresentasikan kemampuan generalisasi model terhadap data baru.

Penilaian kinerja model mengacu pada konsep confusion matrix dalam konteks object detection. Hasil deteksi ditentukan berdasarkan perbandingan antara bounding box prediksi dan bounding box ground truth menggunakan ambang batas Intersection over Union (IoU). Prediksi dikategorikan sebagai true positive (TP) apabila bounding box prediksi sesuai dengan ground truth, sedangkan false positive (FP) terjadi ketika model mendeteksi edema pada area yang tidak sesuai dengan ground truth. False negative (FN) menunjukkan area edema yang tidak berhasil dideteksi oleh model, sedangkan true negative (TN) menunjukkan citra no-edema yang tidak menghasilkan deteksi edema.

Metrik evaluasi yang digunakan meliputi precision, recall, F1-score, IoU, Average Precision (AP), mAP@0.5, dan mAP@0.5–0.95. Precision digunakan untuk mengukur ketepatan prediksi positif, recall digunakan untuk mengukur kemampuan model dalam mendeteksi seluruh objek edema, sedangkan F1-score digunakan untuk menilai keseimbangan antara precision dan recall. IoU digunakan untuk mengukur tingkat kesesuaian antara bounding box prediksi dan ground truth.

Precision digunakan untuk mengukur ketepatan model dalam menghasilkan prediksi positif. Nilai precision dihitung menggunakan persamaan berikut.

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

Recall digunakan untuk mengukur kemampuan model dalam mendeteksi seluruh area edema yang sebenarnya ada. Nilai recall dihitung menggunakan persamaan berikut.

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

F1-score digunakan untuk menilai keseimbangan antara precision dan recall. Nilai F1-score dihitung menggunakan persamaan berikut.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

Intersection over Union (IoU) digunakan untuk mengukur tingkat kesesuaian antara bounding box prediksi dan bounding box ground truth. Nilai IoU dihitung menggunakan persamaan berikut.

$$IoU = \frac{Intersection}{Union} \quad (4)$$

Average Precision (AP) digunakan untuk mengevaluasi performa deteksi berdasarkan luas area di bawah kurva precision-recall. Nilai AP dihitung menggunakan persamaan berikut.

$$AP = \int_0^1 Precision(Recall) d(Recall) \quad (5)$$

Mean Average Precision (mAP) diperoleh dari rata-rata nilai AP. Nilai mAP dihitung menggunakan persamaan berikut.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (6)$$

Dalam penelitian ini, mAP@0.5 dan mAP@0.5–0.95 digunakan sebagai indikator utama evaluasi. mAP@0.5 dihitung pada ambang IoU 0,5, sedangkan mAP@0.5–0.95 dihitung pada beberapa ambang IoU dari 0,5 hingga 0,95 untuk menilai konsistensi deteksi model secara lebih ketat. Perbandingan precision, recall, F1-score, dan mAP antara YOLOv5n baseline dan YOLOv5n + GWO digunakan untuk menilai pengaruh optimasi hyperparameter terhadap performa deteksi *pulmonary edema* pada citra CXR.

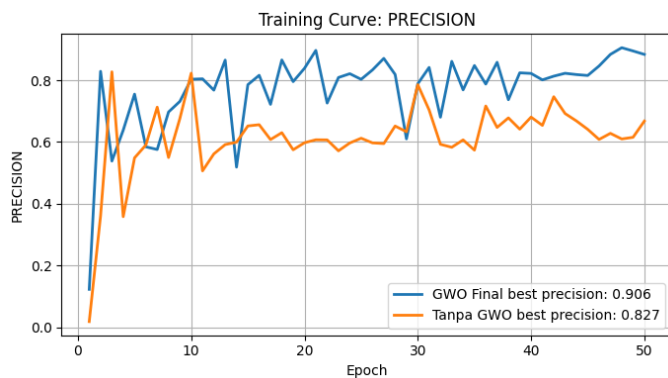
III. HASIL DAN PEMBAHASAN

A. Evaluasi Performa Model YOLOv5n dan YOLOv5n + GWO

Penelitian ini mengevaluasi kinerja YOLOv5n dan pengaruh optimasi hyperparameter menggunakan *Grey Wolf Optimizer* (GWO) dalam mendeteksi *pulmonary edema* pada citra Chest X-ray (CXR). Evaluasi dilakukan dengan membandingkan dua skenario, yaitu YOLOv5n baseline dan YOLOv5n + GWO. Dataset asli terdiri dari 1.000 citra dan dibagi dengan rasio 80% training, 10% validation, dan 10% testing. Dengan demikian, sebanyak 800 citra digunakan sebagai data training, 100 citra sebagai data validation, dan 100 citra sebagai data testing. Setelah proses augmentasi pada data training, jumlah data training meningkat menjadi 1.600 citra, sedangkan data validation dan testing tetap masing-masing 100 citra. Pelatihan dilakukan selama 50 epoch dengan batch size 15 dan resolusi citra 640×640 piksel menggunakan GPU T4 melalui Google Colab.

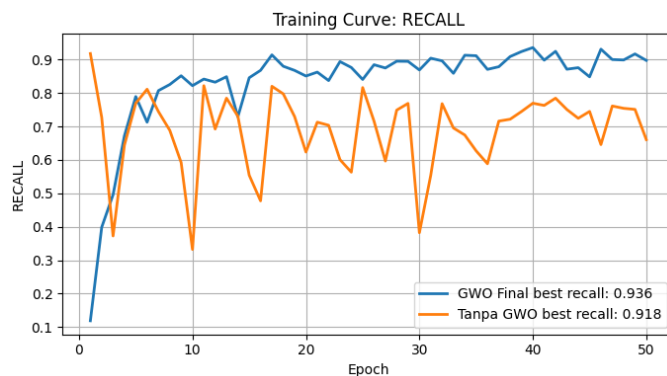
Sebagaimana ditunjukkan pada Gambar 4 dan Gambar 5, model YOLOv5n + GWO memperoleh precision sebesar 0,906 dan recall sebesar 0,936, lebih tinggi dibandingkan baseline dengan precision 0,827 dan recall 0,918. Peningkatan ini menunjukkan bahwa optimasi GWO membantu model menghasilkan prediksi positif yang lebih tepat dan meningkatkan kemampuan deteksi area edema pada citra CXR. Peningkatan performa juga terlihat pada metrik mAP. Berdasarkan Gambar 6 dan Gambar 7, model YOLOv5n + GWO memperoleh mAP@0.5 sebesar 0,963 dan mAP@0.5–0.95 sebesar 0,737, sedangkan baseline memperoleh mAP@0.5 sebesar 0,815 dan mAP@0.5–0.95 sebesar 0,444. Hasil ini menunjukkan bahwa optimasi GWO berkontribusi terhadap peningkatan kemampuan deteksi dan konsistensi lokalisasi bounding box.

Evaluasi akhir pada subset testing menunjukkan bahwa model YOLOv5n + GWO memperoleh hasil lebih baik pada metrik utama object detection, sebagaimana disajikan pada Tabel II. Gambar 8, Gambar 9, dan Gambar 10 menunjukkan kurva box loss, classification loss, dan distribution focal loss (DFL). Meskipun beberapa nilai loss baseline lebih kecil, model YOLOv5n + GWO tetap menunjukkan performa lebih baik pada metrik utama seperti precision, recall, F1-score, dan mAP.



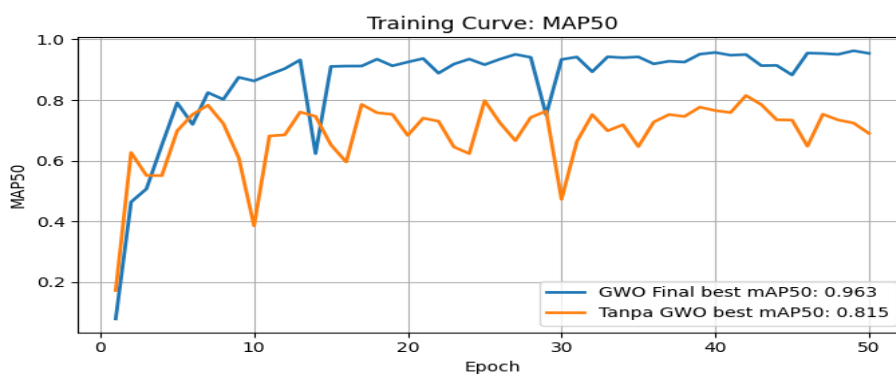
Gambar 4. Kurva Precision.

Gambar 4 menunjukkan perbandingan nilai precision antara model YOLOv5n baseline dan YOLOv5n + GWO selama proses pelatihan. Model YOLOv5n + GWO memperoleh nilai precision akhir sebesar 0,906, lebih tinggi dibandingkan model baseline sebesar 0,827. Peningkatan ini menunjukkan bahwa optimasi GWO membantu model menghasilkan prediksi positif yang lebih tepat, sehingga kesalahan deteksi berupa false positive dapat dikurangi.



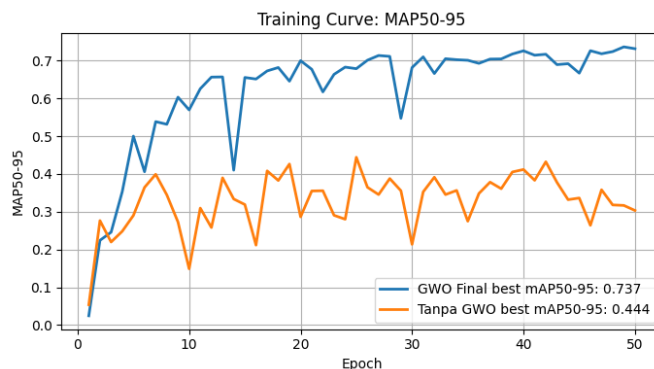
Gambar 5. Kurva Recall.

Gambar 5 menunjukkan perbandingan nilai recall antara model YOLOv5n baseline dan YOLOv5n + GWO selama proses pelatihan. Model YOLOv5n + GWO memperoleh nilai recall akhir sebesar 0,936, lebih tinggi dibandingkan model baseline sebesar 0,918. Peningkatan ini menunjukkan bahwa model hasil optimasi lebih mampu mengenali area edema yang sebenarnya ada pada citra CXR, sehingga potensi false negative dapat dikurangi.



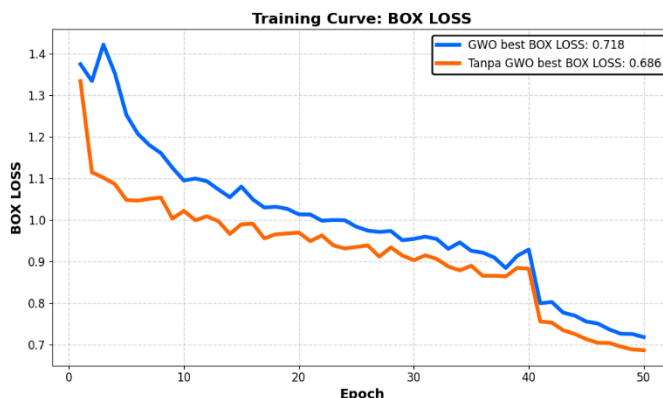
Gambar 6. Kurva mAP@0.5.

Gambar 6 menunjukkan perbandingan nilai mAP@0.5 antara model YOLOv5n baseline dan YOLOv5n + GWO selama proses pelatihan. Model YOLOv5n + GWO memperoleh nilai mAP@0.5 akhir sebesar 0,963, lebih tinggi dibandingkan model baseline sebesar 0,815. Hasil ini menunjukkan bahwa optimasi GWO berkontribusi terhadap peningkatan kemampuan deteksi objek pada ambang IoU 0,5.



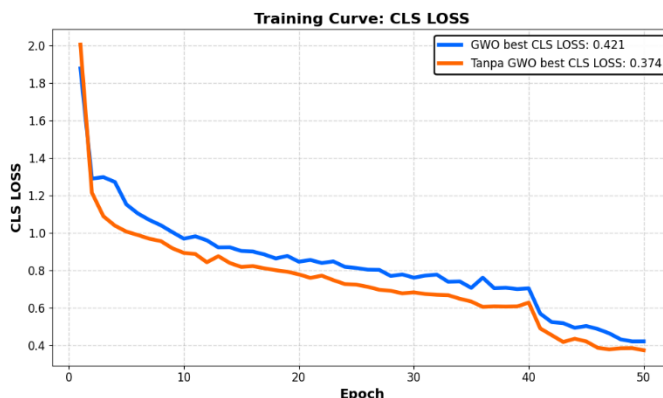
Gambar 7. Kurva mAP@0.5–0.95.

Gambar 7 menunjukkan perbandingan nilai mAP@0.5–0.95 antara model YOLOv5n baseline dan YOLOv5n + GWO selama proses pelatihan. Model YOLOv5n + GWO memperoleh nilai mAP@0.5–0.95 sebesar 0,737, lebih tinggi dibandingkan model baseline sebesar 0,444. Peningkatan ini menunjukkan bahwa optimasi GWO berkontribusi terhadap peningkatan konsistensi deteksi objek pada berbagai ambang Intersection over Union (IoU).



Gambar 8. Kurva Box Loss.

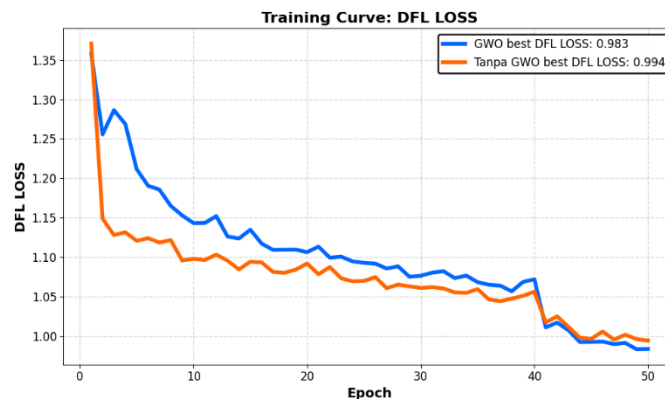
Gambar 8 menunjukkan perbandingan nilai box loss antara model YOLOv5n baseline dan YOLOv5n + GWO selama proses pelatihan. Kedua model menunjukkan penurunan box loss hingga akhir epoch, yang menandakan bahwa proses pelatihan berjalan stabil. Model baseline memperoleh nilai box loss akhir sebesar 0,686, sedangkan model YOLOv5n + GWO memperoleh nilai 0,718. Meskipun nilai box loss baseline lebih rendah, model YOLOv5n + GWO tetap menunjukkan performa lebih baik pada metrik utama deteksi objek seperti precision, recall, dan mAP.



Gambar 9. Kurva Classification Loss.

Gambar 9 menunjukkan perbandingan nilai classification loss antara model YOLOv5n baseline dan YOLOv5n + GWO selama proses pelatihan. Kedua model menunjukkan penurunan classification loss hingga akhir epoch, yang menunjukkan bahwa proses pembelajaran kelas berjalan stabil. Model baseline memperoleh nilai classification loss akhir sebesar 0,374, sedangkan model YOLOv5n + GWO memperoleh nilai 0,421. Meskipun nilai classification loss

baseline lebih rendah, model YOLOv5n + GWO tetap memperoleh performa lebih baik pada metrik utama deteksi objek seperti precision, recall, F1-score, dan mAP.



Gambar 10. Kurva Distribution Focal Loss.

Gambar 10 menunjukkan perbandingan nilai Distribution Focal Loss (DFL) antara model YOLOv5n baseline dan YOLOv5n + GWO selama proses pelatihan. Kedua model menunjukkan penurunan DFL loss hingga akhir epoch, yang menandakan bahwa proses pembelajaran lokalisasi bounding box berjalan stabil. Model YOLOv5n + GWO memperoleh nilai DFL loss akhir sebesar 0,983, sedikit lebih rendah dibandingkan model baseline sebesar 0,994. Hasil ini menunjukkan bahwa optimasi GWO berkontribusi terhadap peningkatan stabilitas lokalisasi bounding box, meskipun peningkatannya relatif kecil.

B. Perbandingan Performa Model YOLOv5n dan YOLOv5n + GWO

Evaluasi kinerja model dilakukan untuk membandingkan kemampuan YOLOv5n baseline dan YOLOv5n + GWO dalam mendeteksi pulmonary edema pada citra Chest X-ray (CXR). Evaluasi ini bertujuan untuk mengetahui pengaruh optimasi hyperparameter menggunakan *Grey Wolf Optimizer* terhadap kualitas deteksi model. Metrik yang digunakan meliputi precision, recall, F1-score, mAP@0.5, dan mAP@0.5–0.95. Precision digunakan untuk menilai ketepatan prediksi positif, recall digunakan untuk mengukur kemampuan model dalam menemukan area edema, sedangkan F1-score digunakan untuk melihat keseimbangan antara precision dan recall. Sementara itu, mAP@0.5 dan mAP@0.5–0.95 digunakan untuk menilai konsistensi lokalisasi bounding box pada berbagai ambang *Intersection over Union* (IoU).

TABEL 2
PERFORMA MODEL YOLOV5N DAN YOLOV5N + GWO

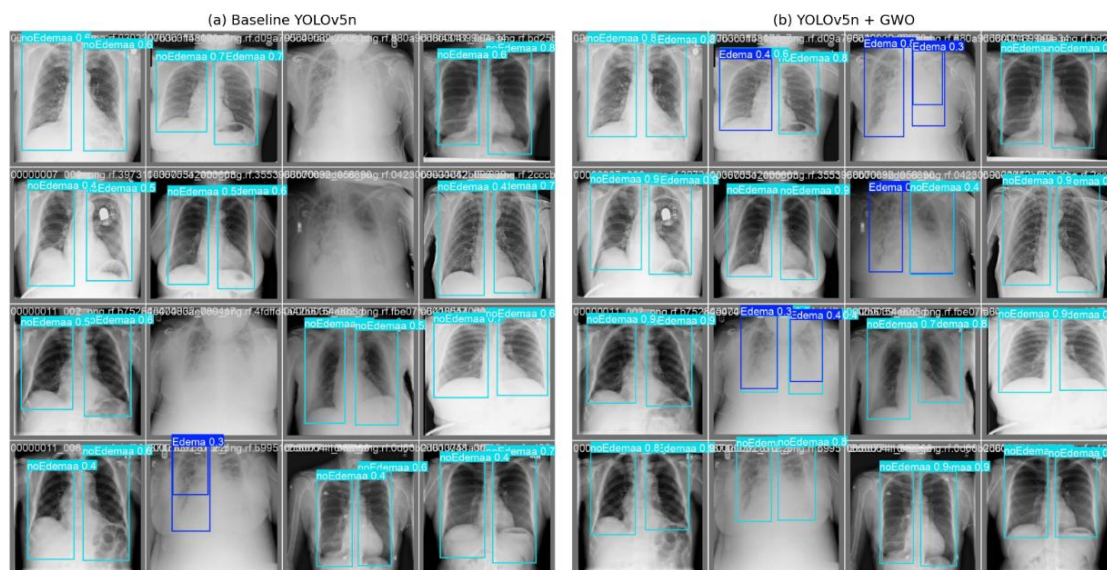
Model	Precision	Recall	mAP@0.5	mAP@0.5–0.95	F1-Score
YOLOv5n (Baseline)	0.827	0.918	0.815	0.444	0.872
YOLOv5n + GWO (Ours)	0.906	0.936	0.963	0.737	0.920

Berdasarkan Tabel II, model YOLOv5n + GWO menunjukkan peningkatan pada seluruh metrik utama dibandingkan model YOLOv5n baseline. Nilai precision meningkat dari 0,827 menjadi 0,906, recall meningkat dari 0,918 menjadi 0,936, F1-score meningkat dari 0,872 menjadi 0,920, mAP@0.5 meningkat dari 0,815 menjadi 0,963, dan mAP@0.5–0.95 meningkat dari 0,444 menjadi 0,737. Peningkatan ini menunjukkan bahwa optimasi hyperparameter menggunakan GWO berkontribusi terhadap ketepatan prediksi, kemampuan mendeteksi area edema, serta konsistensi lokalisasi bounding box pada citra CXR.

Secara keseluruhan, hasil evaluasi menunjukkan bahwa integrasi GWO pada YOLOv5n memberikan pengaruh positif terhadap performa deteksi pulmonary edema pada dataset yang digunakan. Namun, hasil ini tetap terbatas pada data dan skenario eksperimen dalam penelitian ini, sehingga validasi menggunakan dataset yang lebih besar dan lebih beragam masih diperlukan untuk menilai kemampuan generalisasi model.

C. Visualisasi Hasil Deteksi Model

Visualisasi hasil deteksi dilakukan untuk membandingkan keluaran model YOLOv5n baseline dan YOLOv5n + GWO pada citra Chest X-ray (CXR). Visualisasi ini bertujuan untuk menunjukkan kemampuan masing-masing model dalam mengidentifikasi area edema melalui bounding box pada data validasi. Perbandingan hasil deteksi kedua model ditampilkan pada Gambar 11.



Gambar 11. Hasil deteksi pulmonary edema pada citra Chest X-ray menggunakan (a) YOLOv5n baseline dan (b) YOLOv5n + GWO pada data validasi.

Berdasarkan Gambar 11(a), model YOLOv5n baseline mampu mendeteksi beberapa area yang diduga sebagai edema, tetapi masih terdapat prediksi dengan confidence score yang lebih rendah dan lokasi bounding box yang kurang tepat. Hal ini menunjukkan bahwa model baseline masih memiliki keterbatasan dalam mengenali pola visual edema pada citra CXR.

Sementara itu, Gambar 11(b) menunjukkan bahwa model YOLOv5n + GWO menghasilkan deteksi yang lebih konsisten pada contoh citra validasi yang ditampilkan. Model mampu menghasilkan bounding box yang lebih sesuai terhadap area target dan menunjukkan confidence score yang lebih baik pada beberapa hasil deteksi.

Secara umum, visualisasi hasil deteksi ini mendukung hasil evaluasi kuantitatif sebelumnya melalui metrik precision, recall, dan mAP. Namun, hasil visualisasi ini masih terbatas pada contoh data validasi, sehingga pengujian lebih lanjut menggunakan dataset yang lebih besar dan beragam diperlukan untuk memastikan kemampuan generalisasi model.

IV. SIMPULAN

Penelitian ini menganalisis pengaruh optimasi hyperparameter menggunakan *Grey Wolf Optimizer* (GWO) terhadap kinerja model YOLOv5n dalam mendeteksi pulmonary edema pada citra Chest X-ray (CXR). Hasil eksperimen menunjukkan bahwa model YOLOv5n + GWO memperoleh performa lebih baik dibandingkan model YOLOv5n baseline, dengan nilai precision sebesar 0,906, recall sebesar 0,936, F1-score sebesar 0,920, mAP@0.5 sebesar 0,963, dan mAP@0.5–0.95 sebesar 0,737. Hasil tersebut menunjukkan bahwa optimasi GWO berkontribusi terhadap peningkatan ketepatan prediksi, sensitivitas deteksi, dan konsistensi lokalisasi bounding box pada dataset yang digunakan. Meskipun demikian, pemanfaatan model masih terbatas pada dataset yang digunakan dalam penelitian ini dan belum divalidasi pada dataset multi-sumber atau skenario klinis nyata. Oleh karena itu, penelitian selanjutnya disarankan menggunakan dataset yang lebih besar dan beragam, melibatkan validasi anotasi oleh tenaga ahli, serta menguji model pada lingkungan klinis yang lebih representatif. Temuan ini menunjukkan bahwa pemilihan hyperparameter yang lebih optimal berperan penting dalam meningkatkan kinerja model deteksi objek, terutama pada citra CXR yang memiliki pola visual edema yang cenderung kompleks. Dengan demikian, kontribusi utama penelitian ini adalah menunjukkan bahwa integrasi GWO pada YOLOv5n dapat meningkatkan metrik utama deteksi pulmonary edema pada citra CXR, dengan batasan bahwa generalisasi model masih perlu dievaluasi lebih lanjut.

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada Program Studi Teknik Informatika, Universitas Dian Nuswantoro Semarang, atas dukungan dan fasilitas yang diberikan selama pelaksanaan penelitian ini.

DAFTAR PUSTAKA

- [1] M. N. Akbar, X. Wang, D. Erdogmus, and S. Dalal, "PENet: Continuous-Valued Pulmonary Edema Severity Prediction on Chest X-Ray Using Siamese Convolutional Networks," in Proc. 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Glasgow, Scotland, United Kingdom, 2022, pp. 1834–1838, doi: 10.1109/EMBC48229.2022.9871153.
- [2] J. Huynh, S. Masoudi, A. Noorbakhsh, A. Mahmoodi, S. Kligerman, A. Yen, K. Jacobs, L. Hahn, K. Hasenstab, M. Pazzani, and A. Hsiao, "Deep Learning Radiographic Assessment of Pulmonary Edema: Optimizing Clinical Performance, Training With Serum Biomarkers," *IEEE Access*, vol. 10, pp. 48577–48588, 2022, doi: 10.1109/ACCESS.2022.3172706.
- [3] M. Fetanat, M. Stevens, P. Jain, C. Hayward, E. Meijering, and N. H. Lovell, "Fully Elman Neural Network: A Novel Deep Recurrent Neural Network Optimized by an Improved Harris Hawks Algorithm for Classification of Pulmonary Arterial Wedge Pressure," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 5, pp. 1733–1744, 2022, doi: 10.1109/TBME.2021.3129459.
- [4] R. Rampriya, R. Santoshkumar Reddy, K. Lakshmi Prasad, H. B. Karthik Kumar, and R. Vimal Raja, "Diagnosis and Risk Analysis of Pulmonary Disorder in CXR Images Using Optimized Deep Learning Techniques," in Proc. International Conference on Inventive Research in Computing Applications (ICIRCA), 2025, pp. 1221–1226, doi: 10.1109/ICIRCA65293.2025.11089624.
- [5] A. M. Harale, V. K. Bairagi, E. Boonchieng, and M. R. Bachute, "Nodules Detection in Lungs CT Images Using Improved YOLOV5 and Classification of Types of Nodules by CNN-SVM," *IEEE Access*, vol. 12, pp. 140456–140471, 2024, doi: 10.1109/ACCESS.2024.3466292.
- [6] M. Suchithra, S. Guleria, and S. Patil, "NeXt-Ray: Next Gen X-Ray Diagnosis Using Deep Learning and GenAI," in Proc. International Conference on Innovative Trends in Information Technology (ICITIT), 2025, pp. 1–7, doi: 10.1109/ICITIT64777.2025.11041079.
- [7] R. T. Lucassen et al., "Deep Learning for Detection and Localization of B-Lines in Lung Ultrasound," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 9, pp. 4352–4361, 2023, doi: 10.1109/JBHI.2023.3282596.
- [8] V. Tidke et al., "Multi Disease Prediction Related to Pulmonary Area by Leveraging Deep Learning," in Proc. International Conference on Computational, Communication and Information Technology (ICCCIT), 2025, pp. 57–61, doi: 10.1109/ICCCIT62592.2025.10927913.
- [9] S. Hao, X. Li, W. Peng, Z. Fan, Z. Ji, and I. Ganchev, "YOLO-CXR: A Novel Detection Network for Locating Multiple Small Lesions in Chest X-Ray Images," *IEEE Access*, vol. 12, pp. 156003–156019, 2024, doi: 10.1109/ACCESS.2024.3482102.
- [10] W. Lu, Y. Xue, X. Chen, and H. He, "Enhancing Chest X-Ray Classification with Retrieval-Augmented OpenAI GPT Models," in Proc. International Conference on Natural Language Processing (ICNLP), 2025, pp. 394–398, doi: 10.1109/ICNLP65360.2025.11108514.
- [11] Q. Guan, Y. Huang, Y. Luo, P. Liu, M. Xu, and Y. Yang, "Discriminative Feature Learning for Thorax Disease Classification in Chest X-ray Images," *IEEE Transactions on Image Processing*, vol. 30, pp. 2476–2487, 2021, doi: 10.1109/TIP.2021.3052711.
- [12] M. S. Choudhury, R. Rahman, S. T. Bhuiyan, S. Wasi, and S. B. Alam, "Deep Learning for Accurate Classification of Pulmonary Edema Severity in Chest X-Rays: A Comparative Evaluation of CNN Models with Grad-CAM Feature Visualization," in Proc. 2025 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR), Bali, Indonesia, 2025, doi: 10.1109/ICWAPR66259.2025.11306863.
- [13] B. M. Wiley, B. Zhou, G. Pandompatam, J. Zhou, H. O. Kucuk, and X. Zhang, "Lung Ultrasound Surface Wave Elastography for Assessing Patients With Pulmonary Edema," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 11, pp. 3417–3423, 2021, doi: 10.1109/TBME.2021.3072891.
- [14] E. Khalili, B. Priego-Torres, A. León-Jiménez, and D. Sanchez-Morillo, "Automatic Lung Segmentation in Chest X-Ray Images Using SAM With Prompts From YOLO," *IEEE Access*, vol. 12, pp. 122805–122819, 2024, doi: 10.1109/ACCESS.2024.3454188.
- [15] Y. C. Kuyuu and N. Ozekmekci, "Grey Wolf Optimizer to the Hyperparameters Optimization of Convolutional Neural Network with Several Activation Functions," in Proc. 2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 2022, doi: 10.1109/ISMSIT56059.2022.9932838.
- [16] M. Safaldin, M. Otair, and L. Abualgah, "Improved Binary Gray Wolf Optimizer and SVM for Intrusion Detection System in Wireless Sensor Networks," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 2, pp. 1559–1576, 2021, doi: 10.1007/s12652-020-02228-z.
- [17] B. Zhao, L. Chang, and Z. Liu, "Fast-YOLO Network Model for X-Ray Image Detection of Pneumonia," *Electronics*, vol. 14, no. 5, Art. no. 903, 2025, doi: 10.3390/electronics14050903.
- [18] C. Paramita, C. Supriyanto, P. Šolić, C. Wada, and A. A. Dzaky, "Performance Evaluation of YOLOv8 Models for Multi-Class Skin Lesion Detection From Dermoscopic Images," in Proc. 2025 International Conference on Smart Computing, IoT and Machine Learning (SIML), 2025, doi: 10.1109/SIML65326.2025.11080819.
- [19] E. R. Subhiyakto et al., "Evaluation of Resampling Techniques in CNN-Based Heartbeat Classification," *Ingénierie des Systèmes d'Information*, vol. 29, no. 4, pp. 1323–1332, 2024, doi: 10.18280/isi.290408.
- [20] C. Paramita et al., "Comparative Analysis of YOLOv5 and YOLOv8 Cigarette Detection in Social Media Content," *Scientific Journal of Informatics*, vol. 11, no. 2, pp. 341–352, 2024, doi: 10.15294/sji.v11i2.2808.
- [21] W. Fan, X. Guo, L. Teng, and Y. Wu, "Research on Abnormal Target Detection Method in Chest Radiograph Based on YOLO v5 Algorithm," in Proc. 2021 IEEE International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology (CEI), Fuzhou, China, 2021, doi: 10.1109/CEI52496.2021.9574450.
- [22] M. G. Ragab, S. J. Abdulkadir, A. Muneer, A. Alqushaibi, E. H. Sumiea, R. Qureshi, S. M. Al-Selwi, and H. Alhussian, "A Comprehensive Systematic Review of YOLO for Medical Object Detection (2018 to 2023)," *IEEE Access*, vol. 12, pp. 57815–57836, 2024, doi: 10.1109/ACCESS.2024.3386826.
- [23] Y. Yuan, S. Yao, X. Ren, Y. Xiao, Z. Zhou, and Y. Lu, "Chest X-ray Disease Diagnosis Method Based on YOLO Model and Pseudo Color Conversion," in Proc. 2022 5th International Conference on Data Science and Information Technology (DSIT), Shanghai, China, 2022, doi: 10.1109/DSIT55514.2022.9943871.
- [24] J. Leka Shree, E. Joy, and M. S. Sri Ram, "Enhanced Radiological Anomaly Detection Using Optimized YOLO-NAS Model," in Proc. 2024 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE), Shivamogga, India, 2024, doi: 10.1109/AMATHE61652.2024.10582157.
- [25] W. Willy and A. Prabowo, "Penerapan Transfer Learning VGG-16 untuk Mendeteksi Penyakit Mata Manusia Berbasis Citra Fundus," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 11, no. 1, pp. 84–93, 2026, doi: 10.30591/jpit.v11i1.9291.
- [26] D. Schulz, S. Rasch, M. Heilmaier, R. Abbassi, A. Poszler, J. Ulrich, M. Steinhardt, G. A. Kaissis, R. M. Schmid, R. Braren, and T. Lahmer, "A Deep Learning Model Enables Accurate Prediction and Quantification of Pulmonary Edema From Chest X-Rays," *Critical Care*, vol. 27, Art. no. 201, 2023, doi: 10.1186/s13054-023-04426-5.
- [27] Y.-C. Cheng, Y.-C. Hung, G.-H. Huang, T.-B. Chen, N.-H. Lu, K.-Y. Liu, and K.-H. Lin, "Deep Learning-Based Object Detection Strategies for Disease Detection and Localization in Chest X-Ray Images," *Diagnostics*, vol. 14, no. 23, Art. no. 2636, 2024, doi: 10.3390/diagnostics14232636.