



DenseNet121 sebagai Feature Extractor pada Denoising Autoencoder untuk Deteksi Anomali Unsupervised Citra X-Ray Dada

Raihan Muhammar Zikra, Febi Yanto, Benny Sukma Negara^{*}, Surya Agustian, Reski Mai Candra

Fakultas Sains dan Teknologi, Program Studi Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim, Riau, Indonesia

Email: ¹12250114265@students.uin-suska.ac.id, ²febiyanto@uin-suska.ac.id, ^{3*}bsnegara@uin-suska.ac.id, ⁴surya.agustian@uin-suska.ac.id, ⁵reski.candra@uin-suska.ac.id

Email Penulis Korespondensi: bsnegara@uin-suska.ac.id

Abstrak—Deteksi anomali pada citra X-ray dada menjadi tantangan dalam pencitraan medis karena pendekatan supervised learning membutuhkan data berlabel dalam jumlah besar yang sulit diperoleh dan mahal untuk dianotasi. Penelitian ini mengusulkan sistem deteksi anomali unsupervised learning yang mengintegrasikan DenseNet121 pretrained sebagai ekstraktor fitur dengan Denoising Autoencoder (DAE), sehingga pelatihan hanya memerlukan citra normal tanpa anotasi anomali. Model dilatih pada citra normal dan diuji pada citra COVID-19 serta pneumonia untuk mengevaluasi kemampuan deteksi anomali berdasarkan reconstruction error terhadap threshold optimal. Evaluasi dilakukan menggunakan dataset Covid19-Pneumonia-Normal Chest X-Ray Images yang terdiri atas 5.228 citra, dengan membandingkan performa DenseNet121 dan ResNet50 sebagai feature extractor pada tiga variasi latent dimension. Konfigurasi DenseNet121 dengan latent dimension 128 mencapai performa keseluruhan tertinggi pada mayoritas metrik, yaitu akurasi 91%, sensitivity 90,75%, Macro F1-Score 72%, dan validation loss (MSE) 0,0952 pada 3.608 citra uji, meskipun AUC (0,9526) dan specificity (87,36%) tidak selalu menjadi yang tertinggi di antara seluruh konfigurasi yang diuji. Hasil penelitian menunjukkan bahwa penggunaan DenseNet121 sebagai ekstraktor fitur mampu meningkatkan kemampuan DAE dalam membedakan citra paru-paru normal dan anomali, sehingga berpotensi menjadi alternatif skrining awal yang efisien pada kondisi keterbatasan data berlabel.

Kata Kunci: Deteksi Anomali; Denoising Autoencoder; DenseNet121; Citra X-Ray Dada; Unsupervised Learning; Reconstruction Error

Abstract—Anomaly detection in chest X-ray images remains a challenge in medical imaging, as supervised learning approaches require large amounts of labeled data that are difficult and costly to annotate. This study proposes an unsupervised learning anomaly detection system that integrates a pretrained DenseNet121 as a feature extractor with a Denoising Autoencoder (DAE), so that training requires only normal images without anomaly annotations. The model was trained on normal images and tested on COVID-19 and pneumonia images to evaluate its anomaly detection capability based on reconstruction error relative to an optimal threshold. Evaluation was conducted on the Covid19-Pneumonia-Normal Chest X-Ray Images dataset comprising 5,228 images, comparing the performance of DenseNet121 and ResNet50 as feature extractors across three latent dimension configurations. The DenseNet121 configuration with a latent dimension of 128 achieved the highest overall performance on most metrics, namely 91% accuracy, 90.75% sensitivity, 72% Macro F1-Score, and a validation loss (MSE) of 0.0952 on 3,608 test images, although its AUC (0.9526) and specificity (87.36%) were not consistently the highest among all tested configurations. These results demonstrate that using DenseNet121 as a feature extractor improves the DAE's ability to distinguish normal from anomalous lung images, suggesting its potential as an efficient preliminary screening approach under conditions of limited labeled data.

Keywords: Anomaly Detection; Denoising Autoencoder; DenseNet121; Chest X-Ray Images; Unsupervised Learning; Reconstruction Error

1. PENDAHULUAN

Citra X-ray dada (chest X-ray) merupakan salah satu modalitas pencitraan medis yang paling banyak digunakan dalam diagnosis berbagai penyakit paru, seperti pneumonia dan COVID-19, karena memiliki biaya pemeriksaan yang relatif rendah, proses akuisisi yang cepat, serta tersedia secara luas pada berbagai fasilitas pelayanan kesehatan (Scarpiniti et al., 2022; Siddiqi & Javaid, 2024). Seiring dengan perkembangan deep learning, citra X-ray juga dimanfaatkan sebagai masukan pada sistem Computer-Aided Diagnosis (CAD) untuk membantu proses identifikasi penyakit secara otomatis. Sebagai contoh, (Rajput et al., 2022) menerapkan deep learning berbasis fungsi aktivasi cogent pada citra mammogram untuk mendukung diagnosis dini kanker payudara dan memperoleh akurasi klasifikasi yang tinggi. Meskipun berbagai metode berbasis supervised learning telah menunjukkan performa yang tinggi, pendekatan tersebut masih bergantung pada ketersediaan data berlabel dalam jumlah besar. Pada praktik klinis, proses anotasi citra memerlukan keterlibatan tenaga ahli sehingga membutuhkan waktu, biaya, dan sumber daya yang tidak sedikit, sementara data patologi yang representatif sering kali terbatas (Nawaz et al., 2024). Kondisi ini mendorong pengembangan pendekatan unsupervised anomaly detection yang mampu mempelajari karakteristik citra normal dan mengidentifikasi penyimpangan tanpa memerlukan data anomali berlabel.

Berbagai penelitian menunjukkan bahwa Denoising Autoencoder (DAE) merupakan salah satu pendekatan yang efektif untuk deteksi anomali pada berbagai domain. Pada bidang industri, (Kim & Soo Kim, 2024) melaporkan bahwa Deep Denoising Autoencoder (DDAE) mampu meningkatkan nilai Area Under the Curve (AUC) sekitar 18% pada deteksi anomali berbasis sinyal suara. Sementara itu, (Farooq & Savaş, 2024) menunjukkan bahwa DAE berbasis CNN dua tahap mampu mereduksi noise pada citra digital dengan tetap mempertahankan struktur tepi dan detail penting. Masih pada domain industri, (Liu et al., 2025) mengusulkan pendekatan reconstruction-based yang memanfaatkan perbedaan bias frekuensi antara citra normal dan anomali untuk mengatasi trade-off antara fidelitas rekonstruksi citra normal dan kemampuan membedakan citra anomali pada inspeksi visual manufaktur. Penerapan DAE turut meluas ke



berbagai domain lain, antara lain keamanan siber untuk deteksi intrusi pada jaringan Internet of Things (Alrayes et al., 2024), deteksi anomali berbasis graf melalui kerangka anomaly-denoised autoencoder (He et al., 2024), pertanian presisi untuk mendeteksi objek berisiko di sekitar kendaraan pertanian otonom (Mujkic et al., 2022), sistem rekomendasi untuk mengidentifikasi natural noise pada data rating pengguna (Park et al., 2023), serta sistem tenaga listrik untuk identifikasi topologi jaringan distribusi ketika sebagian data pengukuran hilang (Raghuvamsi et al., 2023). Hasil-hasil tersebut menunjukkan bahwa DAE memiliki kemampuan yang baik dalam mempelajari representasi data normal serta menghasilkan model yang robust terhadap gangguan pada data masukan. Pada domain pencitraan medis, teknik denoising secara umum berperan penting dalam meningkatkan kualitas citra guna mendukung proses diagnosis (Patil & Bhosale, 2022). Secara lebih spesifik, (Kascenas et al., 2022) menerapkan DAE secara unsupervised untuk mendeteksi tumor pada citra MRI otak dan menunjukkan bahwa penambahan noise selama pelatihan dapat meningkatkan kemampuan generalisasi model. Selain itu, (Scarpiniti et al., 2022) mengembangkan Deep Denoising Autoencoder untuk mendeteksi area paru yang terindikasi COVID-19 secara unsupervised dan memperoleh performa yang baik meskipun menggunakan data berlabel dalam jumlah terbatas. Pada pencitraan ilmiah beresolusi tinggi, (Alvarado et al., 2023) mengembangkan DAE yang dilatih pada struktur hasil simulasi untuk mereduksi noise pada citra scanning transmission electron microscopy (STEM) kromatin tanpa menghilangkan detail struktural penting. Temuan-temuan tersebut mengindikasikan bahwa DAE memiliki potensi yang besar untuk diterapkan pada deteksi anomali citra medis.

Meskipun demikian, hasil telaah terhadap penelitian terdahulu menunjukkan beberapa keterbatasan. Sebagian besar penelitian masih menggunakan arsitektur encoder konvolusional standar atau bekerja langsung pada ruang piksel (pixel space), sehingga representasi fitur yang dipelajari belum memanfaatkan kemampuan feature extractor CNN pretrained. Selain itu, evaluasi umumnya hanya dilakukan pada satu jenis citra atau satu jenis anomali sehingga kemampuan model dalam mendeteksi berbagai patologi pada citra X-ray dada masih belum banyak dieksplorasi. Kondisi tersebut menunjukkan bahwa masih terdapat peluang untuk mengembangkan pendekatan DAE yang memanfaatkan representasi fitur yang lebih kaya guna meningkatkan performa deteksi anomali pada citra medis.

Berdasarkan telaah terhadap penelitian terdahulu, masih terdapat beberapa keterbatasan yang membuka peluang untuk pengembangan metode deteksi anomali pada citra X-ray dada berbasis unsupervised learning. Penelitian ini mengidentifikasi tiga research gap utama. (1) Sebagian besar penelitian Denoising Autoencoder pada citra medis masih menggunakan encoder konvolusional standar atau melakukan proses rekonstruksi langsung pada ruang piksel (pixel space), sehingga representasi fitur yang dipelajari belum sepenuhnya memanfaatkan kemampuan feature extractor CNN pretrained dalam menangkap karakteristik visual yang lebih diskriminatif. (2) Berdasarkan literatur yang ditelaah, pemanfaatan DenseNet121 sebagai fixed feature extractor sebelum proses rekonstruksi Denoising Autoencoder untuk deteksi anomali pada citra X-ray dada masih sangat terbatas. Padahal, karakteristik dense connections pada DenseNet121 memungkinkan pemanfaatan kembali informasi fitur antarlapis sehingga menghasilkan representasi fitur yang lebih kaya dan mendukung proses pembelajaran yang lebih efektif dibandingkan arsitektur CNN konvensional (Zhou et al., 2024). Meskipun kombinasi DenseNet dan Denoising Autoencoder telah dilaporkan pada tugas klasifikasi supervised kanker serviks (Yang et al., 2024), penerapannya dalam kerangka unsupervised anomaly detection untuk citra X-ray dada multi-patologi masih belum banyak dilaporkan. (3) Sebagian besar penelitian hanya mengevaluasi satu jenis patologi sehingga kemampuan model dalam mendeteksi lebih dari satu jenis anomali paru dalam satu kerangka unsupervised learning masih belum banyak dieksplorasi. Berdasarkan ketiga research gap tersebut, penelitian ini mengusulkan integrasi DenseNet121 sebagai feature extractor dengan Denoising Autoencoder untuk deteksi anomali citra X-ray dada serta mengevaluasi kemampuannya dalam mendeteksi dua kategori anomali, yaitu COVID-19 dan Pneumonia.

Berdasarkan research gap tersebut, penelitian ini mengusulkan suatu kerangka *unsupervised anomaly detection* yang mengintegrasikan DenseNet121 sebagai *feature extractor* dengan Denoising Autoencoder (DAE) untuk mendeteksi anomali pada citra X-ray dada. Pada pendekatan yang diusulkan, DenseNet121 digunakan untuk mengekstraksi representasi fitur berdimensi tinggi, sedangkan DAE mempelajari distribusi fitur citra normal melalui proses rekonstruksi. Selanjutnya, kinerja model dievaluasi berdasarkan kemampuan reconstruction error dalam membedakan citra normal dari citra anomali (COVID-19 dan Pneumonia), dengan penentuan nilai ambang menggunakan metode *Youden Index*.

Penelitian ini bertujuan untuk merancang dan mengevaluasi sistem deteksi anomali berbasis unsupervised learning yang mengintegrasikan DenseNet121 sebagai feature extractor dengan Denoising Autoencoder (DAE) pada citra X-ray dada. DenseNet121 digunakan untuk mengekstraksi representasi fitur dari citra masukan, sedangkan DAE mempelajari distribusi fitur citra normal melalui proses rekonstruksi. Kinerja model dievaluasi berdasarkan nilai Mean Squared Error (MSE) sebagai anomaly score, dengan penentuan nilai ambang (threshold) menggunakan kurva Receiver Operating Characteristic (ROC) dan Youden Index untuk membedakan citra normal dari citra anomali (COVID-19 dan Pneumonia). Pendekatan ini diharapkan dapat meningkatkan kualitas representasi fitur yang digunakan pada proses deteksi anomali tanpa bergantung pada data anomali berlabel selama proses pelatihan.

Kontribusi penelitian ini meliputi tiga aspek. (1) Penelitian ini mengusulkan integrasi DenseNet121 sebagai feature extractor dengan Denoising Autoencoder dalam kerangka deteksi anomali citra X-ray dada berbasis unsupervised learning. (2) Penelitian ini mengevaluasi kemampuan model dalam mendeteksi dua kategori anomali, yaitu COVID-19 dan Pneumonia, menggunakan reconstruction error sebagai dasar identifikasi anomali. (3) Penelitian



ini menerapkan penentuan threshold berbasis kurva ROC dan Youden Index sebagai pendekatan yang objektif dan dapat direproduksi untuk mendukung proses pengambilan keputusan pada sistem Computer-Aided Diagnosis (CAD).

2. METODOLOGI PENELITIAN

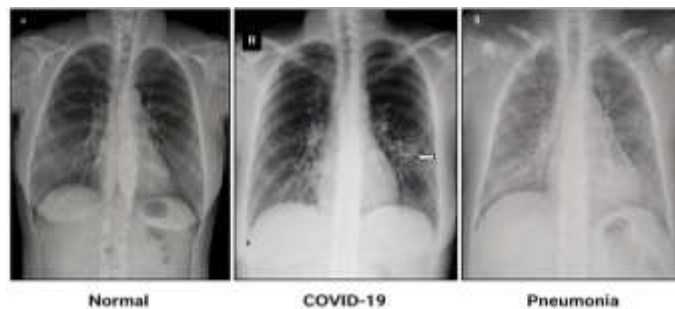
2.1 Dataset dan Pembagian Data

Dataset yang digunakan dalam penelitian ini adalah Covid19-Pneumonia-Normal Chest X-Ray Images yang dipublikasikan oleh (Sachin Kumar, 2022) melalui platform Mendeley Data (URL: <https://data.mendeley.com/datasets/dvntn9yhd2/1>). Dataset ini mencakup tiga kelas citra X-ray dada, yaitu Normal, COVID-19, dan Pneumonia. Sesuai dengan prinsip unsupervised anomaly detection, hanya citra kelas Normal yang digunakan dalam proses pelatihan model, sedangkan kelas COVID-19 dan Pneumonia digunakan secara eksklusif pada tahap pengujian sebagai representasi kelas anomali. Komposisi lengkap dataset disajikan pada Tabel 1.

Tabel 1. Komposisi Dataset Citra X-Ray Dada

No	Kategori Data	Jumlah Citra	Ukuran Citra	Format	Penggunaan dalam Penelitian
1	Normal	1.802	232×232 piksel	PNG	Pelatihan & Pengujian (uji Normal)
2	Pneumonia	1.800	232×232 piksel	PNG	Pengujian eksklusif (Anomali)
3	COVID-19	1.626	256×256 piksel	PNG	Pengujian eksklusif (Anomali)
Total		5.228	—	—	—

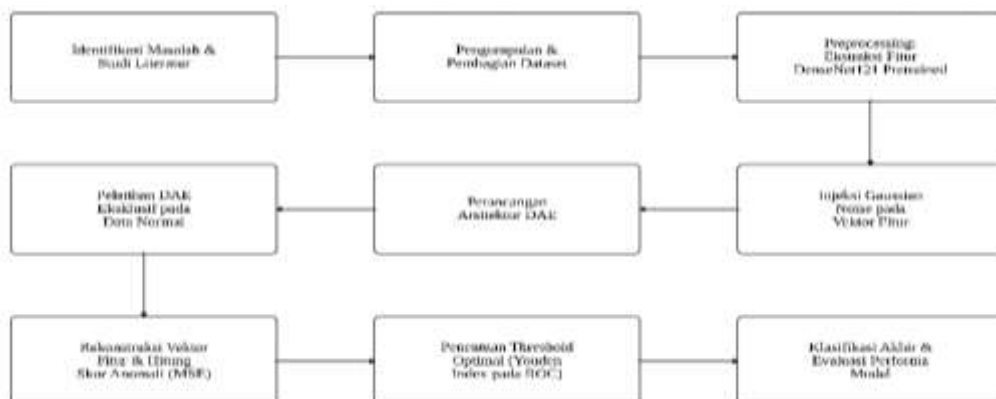
Data kelas Normal dibagi menggunakan random seed tetap (42) menjadi tiga subset: 80% sebagai data pelatihan (~1.442 citra), 10% sebagai data validasi (~180 citra), dan 10% sebagai data uji normal (182 citra). Total data uji yang digunakan dalam evaluasi akhir adalah 3.608 citra, terdiri dari 182 citra Normal dan 3.426 citra anomali (COVID-19: 1.626 dan Pneumonia: 1.800). Rasio kelas pada data uji adalah 1:18,8 (Normal:Anomali), mencerminkan kondisi nyata di mana kasus normal lebih jarang menjadi target skrining dibandingkan kasus yang perlu diverifikasi. Contoh citra dari masing-masing kelas ditampilkan pada Gambar 1.



Gambar 1. Contoh Citra X-Ray Paru-Paru: (a) Normal, (b) COVID-19, (c) Pneumonia

2.2 Tahapan Penelitian

Penelitian ini dilaksanakan melalui sembilan tahapan sistematis: (1) identifikasi masalah dan studi literatur; (2) pengumpulan dan pembagian dataset; (3) preprocessing berbasis ekstraksi fitur DenseNet121 pretrained; (4) injeksi Gaussian noise pada vektor fitur; (5) perancangan arsitektur DAE; (6) pelatihan DAE eksklusif pada data Normal; (7) rekonstruksi vektor fitur dan perhitungan skor anomali (MSE); serta (8) penentuan threshold optimal menggunakan Youden Index pada kurva ROC; (9) diikuti klasifikasi akhir dan evaluasi performa. Alur sistem secara keseluruhan diilustrasikan pada Gambar 2



Gambar 2. Tahapan Penelitian



2.3 Preprocessing: Ekstraksi Fitur Berbasis DenseNet121

Tahap preprocessing merupakan komponen penting dalam penelitian ini karena berfungsi menghasilkan representasi fitur yang lebih informatif sebelum proses rekonstruksi oleh Denoising Autoencoder (DAE). Setiap citra X-ray terlebih dahulu diubah ukurannya menjadi 224×224 piksel dengan tiga kanal warna (RGB) agar sesuai dengan spesifikasi masukan DenseNet121. Selanjutnya, DenseNet121 dengan bobot awal (pretrained weights) dari ImageNet digunakan sebagai feature extractor untuk menghasilkan representasi fitur dari setiap citra. Lapisan klasifikasi akhir dihilangkan dan digantikan dengan lapisan Global Average Pooling sehingga setiap citra direpresentasikan sebagai vektor fitur berdimensi 1.024. Seluruh parameter DenseNet121 dibekukan (frozen atau non-trainable) sehingga model berfungsi sebagai fixed feature extractor selama proses pelatihan DAE. Konfigurasi lengkap ditunjukkan pada Tabel 2.

Tabel 2. Konfigurasi Preprocessing Ekstraksi Fitur CNN Pretrained

Parameter	DenseNet121	ResNet50 (Pembanding)
Arsitektur	DenseNet121	ResNet50
Dataset Pretrain	ImageNet (1.000 kelas)	ImageNet (1.000 kelas)
Ukuran Input Citra	$224 \times 224 \times 3$ (RGB)	$224 \times 224 \times 3$ (RGB)
Lapisan Output yang Digunakan	Global Average Pooling	Global Average Pooling
Dimensi Vektor Fitur Output	1.024 dimensi	2.048 dimensi
Status Bobot	Frozen / Non-trainable	Frozen / Non-trainable
Mekanisme Utama	Dense Connections (aliran fitur ke semua layer berikutnya)	Residual Connections (skip connections)

Penggunaan bobot pretrained dari ImageNet didasarkan pada konsep transfer learning, yaitu memanfaatkan representasi fitur visual umum yang telah dipelajari dari dataset berskala besar. Meskipun ImageNet terdiri atas citra alami yang berbeda dengan citra medis X-ray, penelitian terdahulu menunjukkan bahwa lapisan awal CNN umumnya mampu mengekstraksi fitur dasar, seperti tepi, tekstur, dan pola bentuk, yang masih dapat ditransfer ke berbagai domain pencitraan. Pada penelitian ini, DenseNet121 tidak dilakukan fine-tuning, melainkan digunakan sebagai fixed feature extractor agar proses pembelajaran difokuskan pada Denoising Autoencoder dalam mempelajari distribusi fitur citra normal. Pendekatan ini juga dipilih untuk mengurangi kompleksitas pelatihan dan meminimalkan risiko overfitting akibat keterbatasan jumlah data pelatihan. Potensi perbedaan domain (domain shift) antara ImageNet dan citra X-ray diakui sebagai salah satu keterbatasan pendekatan ini, sehingga efektivitas representasi fitur dievaluasi secara empiris melalui perbandingan dengan ResNet50 sebagai feature extractor pembanding.

Keunggulan arsitektur DenseNet121 terletak pada mekanisme dense connections yang menghubungkan setiap lapisan dengan seluruh lapisan berikutnya melalui aliran informasi secara feed-forward. Mekanisme ini memungkinkan setiap lapisan memanfaatkan kembali (feature reuse) representasi fitur yang telah dipelajari pada lapisan sebelumnya sehingga menghasilkan representasi yang lebih kaya dan mendukung aliran gradien yang lebih baik. Melalui proses tersebut, setiap citra X-ray yang semula direpresentasikan sebagai data piksel berdimensi $224 \times 224 \times 3$ (150.528 fitur) ditransformasikan menjadi vektor fitur semantik berdimensi 1.024 yang selanjutnya digunakan sebagai masukan bagi Denoising Autoencoder.

2.4 Arsitektur Denoising Autoencoder (DAE)

DAE dalam penelitian ini beroperasi pada feature space berdimensi 1.024 yang dihasilkan DenseNet121 (atau 2.048 untuk ResNet50). Encoder terdiri atas tiga lapisan Dense berurutan (512, 256, dan 128 neuron), masing-masing dilengkapi dengan aktivasi ReLU dan Batch Normalization, kemudian diikuti oleh lapisan latent yang dimensinya divariasikan sebagai hyperparameter percobaan (32, 64, atau 128). Decoder memiliki struktur simetris yang mengembalikan representasi ke dimensi fitur asli (1.024 untuk DenseNet121 dan 2.048 untuk ResNet50). Injeksi Gaussian noise ($\sigma = 0,1$) ditambahkan pada vektor fitur masukan selama pelatihan, sedangkan target rekonstruksi tetap menggunakan vektor fitur asli tanpa gangguan sehingga DAE dapat mempelajari distribusi fitur normal yang lebih robust. Tabel 3. Pemilihan latent dimension sebesar 32, 64, dan 128 didasarkan pada pertimbangan kapasitas representasi ruang laten. Dalam autoencoder, ukuran latent space menentukan tingkat kompresi informasi dari data masukan. Dimensi yang terlalu kecil berpotensi menghilangkan informasi penting sehingga meningkatkan kesalahan rekonstruksi, sedangkan dimensi yang terlalu besar dapat mempertahankan informasi yang tidak relevan dan mengurangi sensitivitas model dalam membedakan data normal dan anomali. Oleh karena itu, penelitian ini mengevaluasi tiga tingkat kapasitas ruang laten, yaitu rendah (32), sedang (64), dan tinggi (128), untuk menganalisis pengaruhnya terhadap kualitas rekonstruksi dan performa deteksi anomali. Rincian arsitektur DAE disajikan pada Tabel 3.

Tabel 3. Arsitektur Denoising Autoencoder (DAE)

Komponen	Lapisan / Konfigurasi	Keterangan
Input	Vektor Fitur [1.024 / 2.048]-dim	Output DenseNet121 / ResNet50 + Gaussian Noise $\sigma=0,1$
Encoder	Dense(512) + BatchNorm	Aktivasi ReLU
	Dense(256) + BatchNorm	Aktivasi ReLU



Komponen	Lapisan / Konfigurasi	Keterangan
Latent Space	Dense(128) + BatchNorm	Aktivasi ReLU
	Dense(32 / 64 / 128)	Variasi hyperparameter eksperimen
Decoder	Dense(128)	Aktivasi ReLU
	Dense(256)	Aktivasi ReLU
	Dense(512)	Aktivasi ReLU
Output	Dense(1.024 / 2.048)	Rekonstruksi vektor fitur; tanpa aktivasi

2.5 Strategi Pelatihan Model

Model DAE dilatih secara eksklusif menggunakan data fitur kelas Normal sesuai dengan prinsip unsupervised anomaly detection, dengan konfigurasi hyperparameter sebagaimana dirangkum pada Tabel 3. Strategi pelatihan mencakup tiga mekanisme adaptif, yaitu: (1) EarlyStopping berbasis validation loss dengan patience 7 epoch dan restore_best_weights=True untuk menghentikan pelatihan ketika model tidak lagi mengalami peningkatan performa sehingga dapat mengurangi risiko overfitting; (2) ModelCheckpoint untuk menyimpan bobot model dengan nilai validation loss terbaik secara otomatis; serta (3) ReduceLROnPlateau yang menurunkan learning rate sebesar faktor 0,5 apabila validation loss tidak membaik selama tiga epoch berturut-turut sehingga proses optimasi dapat keluar dari kondisi stagnasi dan mencapai konvergensi yang lebih stabil. Konfigurasi tersebut diterapkan secara identik pada seluruh skenario eksperimen sehingga perbedaan performa yang diperoleh terutama dipengaruhi oleh variasi feature extractor dan latent dimension, bukan oleh perbedaan strategi pelatihan. Pemilihan strategi pelatihan ini didasarkan pada kebutuhan untuk memperoleh proses optimasi yang stabil sekaligus meningkatkan kemampuan generalisasi model. Penggunaan EarlyStopping dan ModelCheckpoint membantu mencegah pembelajaran berlebih (overfitting) pada data pelatihan, sedangkan ReduceLROnPlateau memungkinkan penyesuaian learning rate secara adaptif ketika proses pembelajaran mulai mengalami stagnasi. Dengan demikian, konfigurasi tersebut mendukung proses pelatihan Denoising Autoencoder agar menghasilkan representasi fitur yang lebih robust terhadap data normal sebelum digunakan pada tahap deteksi anomali.

2.6 Arsitektur Denoising Autoencoder (DAE)

Setelah pelatihan, skor anomali untuk setiap citra dihitung sebagai nilai MSE antara vektor fitur asli dan vektor fitur hasil rekonstruksi dari model. Nilai ambang batas (threshold) optimal ditentukan menggunakan Youden Index (J) pada kurva ROC, mengikuti pendekatan estimasi cutpoint optimal yang telah divalidasi secara luas dalam literatur statistik (Thiele & Hirschfeld, 2021), yaitu nilai τ yang memaksimalkan $J = \text{argmax}_{\tau} (\text{TPR}(\tau) - \text{FPR}(\tau))$, di mana TPR adalah True Positive Rate (sensitivitas) dan FPR adalah False Positive Rate ($1 - \text{spesifisitas}$). Citra dengan nilai $\text{MSE} \geq \tau$ diklasifikasikan sebagai Anomali; sebaliknya diklasifikasikan sebagai Normal. Metrik evaluasi yang digunakan mencakup: AUC (Area Under the ROC Curve), akurasi, sensitivity (recall kelas Anomali), specificity (recall kelas Normal), Macro F1-Score, presisi, dan confusion matrix. Nilai AUC menjadi metrik utama karena tidak bergantung pada threshold tertentu dan merepresentasikan kemampuan pemisahan kelas model secara keseluruhan, meskipun pada kondisi kelas yang tidak seimbang metrik pelengkap seperti Matthews Correlation Coefficient (MCC) turut direkomendasikan untuk melengkapi evaluasi (Chicco & Jurman, 2023), sehingga penelitian ini turut melaporkan precision, recall, dan F1-Score sebagai metrik pendamping.

3. HASIL DAN PEMBAHASAN

3.1 Ringkasan Perbandingan Seluruh Skenario Eksperimen

Penelitian ini menghasilkan enam skenario eksperimen dari kombinasi dua feature extractor (DenseNet121 dan ResNet50) dengan tiga variasi latent dimension (32, 64, dan 128). Seluruh model dilatih menggunakan data citra Normal secara eksklusif (± 1.442 citra) dan dievaluasi pada test set yang terdiri dari 3.608 citra (182 Normal dan 3.426 Anomali). Ringkasan perbandingan performa seluruh skenario disajikan pada Tabel 4.

Tabel 4. Ringkasan Perbandingan Hasil Seluruh Skenario Eksperimen

Feature Extractor	Latent	Epoch	Val Loss	AUC	Sensitivity	Specificity	F1 Macro	Accuracy
DenseNet121	32	57	0,1003	0,9543	89,61%	87,91%	70%	90%
DenseNet121	64	70	0,1029	0,9569	88,97%	92,31%	70%	89%
DenseNet121	128	56	0,0952	0,9526	90,75%	87,36%	72%	91%
ResNet50	32	43	0,1263	0,9335	76,42%	95,05%	58%	77%
ResNet50	64	47	0,1175	0,9488	81,90%	94,51%	63%	83%
ResNet50	128	50	0,1089	0,9529	84,88%	91,21%	65%	85%

Berdasarkan Tabel 4, seluruh skenario berbasis DenseNet121 secara konsisten menunjukkan performa yang lebih unggul dibandingkan skenario ResNet50 pada hampir seluruh metrik evaluasi. Konfigurasi terbaik secara keseluruhan diperoleh pada kombinasi DenseNet121 dengan latent dimension 128 (ditandai dengan tulisan tebal), yang menghasilkan validation loss terendah (0,0952), akurasi tertinggi (91%), sensitivity tertinggi (90,75%), dan Macro F1-



Score tertinggi (72%). Meskipun nilai AUC tertinggi secara absolut dimiliki oleh konfigurasi DenseNet121 latent 64 (0,9569), selisihnya dengan latent 128 hanya 0,43 poin persentase, sementara konfigurasi latent 128 lebih unggul pada sensitivity (metrik paling kritis dalam skrining medis), akurasi, dan Macro F1-Score. Performa lengkap kelima skenario lainnya (DenseNet121 latent 32 dan 64; ResNet50 latent 32, 64, dan 128) telah terangkum pada Tabel 4 di atas; subbagian berikut berfokus pada pembahasan rinci konfigurasi terbaik serta perbandingan agregat antar feature extractor.

3.2 Analisis Konfigurasi Terbaik: DenseNet121 dengan Latent Dimension 128

Konfigurasi terbaik secara keseluruhan diperoleh pada DenseNet121 dengan latent dimension 128, mencapai best epoch pada epoch ke-56 dari maksimum 100 epoch dengan validation loss terbaik sebesar 0,0952 terendah di antara seluruh enam skenario (lihat Tabel 5). Sebagai pembanding, konfigurasi DenseNet121 latent 32 mencapai best epoch ke-57 dengan validation loss 0,1003, sementara latent 64 mencapai best epoch ke-70 dengan validation loss 0,1029. Tabel 6 menyajikan detail seluruh metrik evaluasi konfigurasi terbaik, termasuk distribusi rata-rata reconstruction error per kelas.

Tabel 5. Hasil Evaluasi DenseNet121, Latent Dimension 128

Metrik Evaluasi	Nilai
Best Validation Loss (MSE)	0,0952 (terendah seluruh eksperimen)
Reconstruction Error Mean (keseluruhan)	0,0928
Reconstruction Error Std	0,0268
Mean Error — Kelas Normal	0,093
Mean Error — Kelas COVID-19	0,3637 (rasio 3,91× terhadap Normal)
Mean Error — Kelas Pneumonia	0,1712 (rasio 1,84× terhadap Normal)
AUC	0,9526
Threshold Optimal (Youden Index)	0,1155
Accuracy	91% (tertinggi seluruh eksperimen)
Sensitivity (Recall Anomali)	90,75% (tertinggi seluruh eksperimen)
Specificity (Recall Normal)	87,36%
Macro F1-Score	72% (tertinggi seluruh eksperimen)
Best Epoch / Stopped Epoch	56 / 63

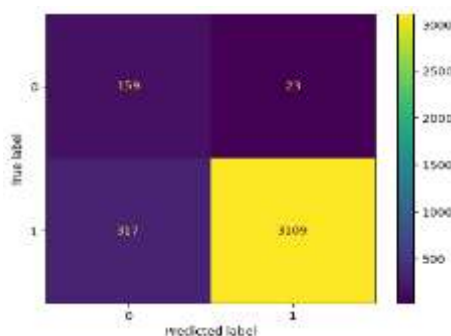
Tabel 6. Classification Report — DenseNet121, Latent Dimension 128

Kelas	Precision	Recall	F1-Score	Support
Normal (0)	33%	87%	48%	182
Anomali (1)	99%	91%	95%	3.426
Accuracy	—	—	91%	3.608
Macro Avg	66%	89%	72%	3.608
Weighted Avg	96%	91%	92%	3.608

Tabel 7. Confusion Matrix — DenseNet121, Latent Dimension 128

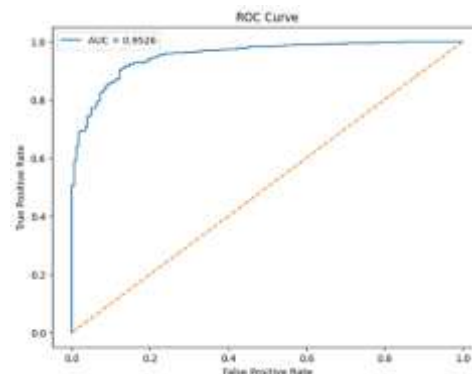
Aktual \ Prediksi	Normal (Pred.)	Anomali (Pred.)	Total
Normal (182)	159 (TN)	23 (FP)	182
Anomali (3.426)	317 (FN)	3.109 (TP)	3.426

Dari Tabel 6, model mencapai F1-Score kelas Anomali sebesar 95% dengan presisi 99% dan recall 91%, mengindikasikan bahwa model sangat andal dalam mendeteksi kasus anomali. Dari Tabel 7, dari 3.426 citra Anomali sebanyak 3.109 (TP) berhasil dideteksi dengan benar, sementara hanya 317 (FN) yang gagal terdeteksi. Dari 182 citra Normal, 159 (TN) teridentifikasi benar dan hanya 23 (FP) yang keliru diklasifikasikan sebagai Anomali. Visualisasi hasil konfigurasi terbaik ini disajikan pada Gambar 3, 4, dan 5.



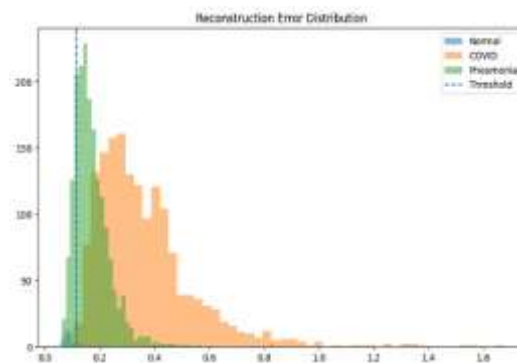
Gambar 3. Confusion Matrix — DenseNet121, Latent Dimension 128

Pada Gambar 3, confusion matrix menunjukkan hasil klasifikasi pada 3.608 citra uji, terdiri atas 182 citra normal dan 3.426 citra anomali. Model berhasil mengidentifikasi 3.109 citra anomali dengan benar (True Positive) dan 159 citra normal dengan benar (True Negative), menghasilkan sensitivity sebesar 90,75% dan specificity sebesar 87,36%.



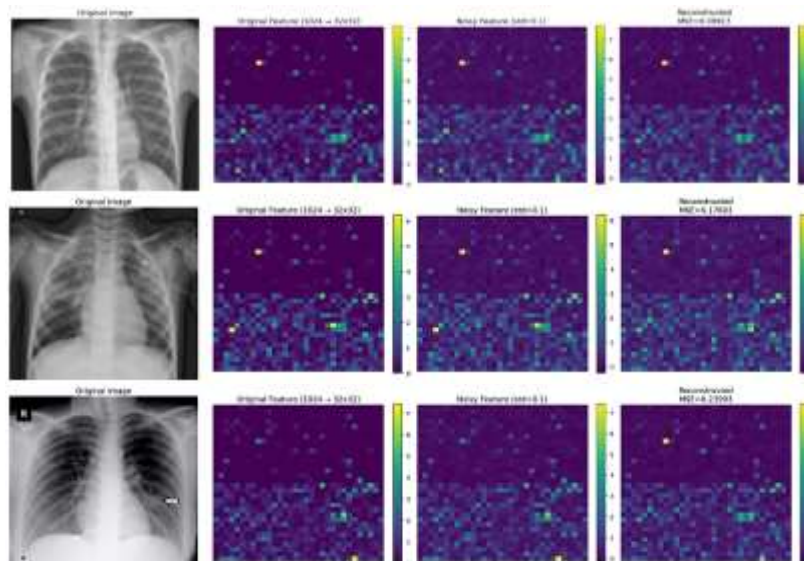
Gambar 4. Kurva ROC — DenseNet121, Latent Dimension 128 (AUC = 0,9526)

Pada Gambar 4, kurva Receiver Operating Characteristic (ROC) memperlihatkan kemampuan model dalam membedakan citra normal dan anomali pada berbagai nilai threshold. Nilai Area Under Curve (AUC) sebesar 0,9526 menunjukkan kemampuan diskriminasi yang sangat baik, dengan tingkat pemisahan kelas mencapai lebih dari 95%.



Gambar 5. Distribusi Reconstruction Error per Kelas — DenseNet121, Latent Dimension 128

Pada Gambar 5, distribusi reconstruction error menunjukkan perbedaan yang jelas antara kelas Normal, Pneumonia, dan COVID-19. Kelas Normal memiliki nilai error yang cenderung rendah dan terkonsentrasi di bawah threshold, sedangkan kelas Pneumonia dan COVID-19 menghasilkan error yang lebih tinggi. Temuan ini mengindikasikan bahwa model berhasil mempelajari distribusi data normal dan mampu mendeteksi penyimpangan sebagai anomali berdasarkan nilai reconstruction error.



Gambar 6. Visualisasi proses rekonstruksi fitur menggunakan DenseNet121-DAE pada sampel (a) Normal, (b) Pneumonia, dan (c) COVID-19.



Untuk memberikan gambaran mengenai cara kerja sistem yang diusulkan, Gambar 6 menunjukkan skenario proses deteksi anomali pada citra X-ray dada menggunakan kombinasi DenseNet121 dan Denoising Autoencoder (DAE). Pada tahap awal, citra masukan diproses menggunakan DenseNet121 pretrained dengan bobot yang dibekukan (frozen weights) untuk mengekstraksi representasi fitur berdimensi 1.024 yang mengandung informasi penting mengenai struktur paru-paru. Vektor fitur yang dihasilkan kemudian diberikan Gaussian Noise ($\sigma = 0,1$) sebagai mekanisme denoising sebelum dimasukkan ke dalam DAE. Selanjutnya, DAE melakukan proses rekonstruksi untuk menghasilkan kembali representasi fitur yang mendekati fitur asli. Perbedaan antara fitur asli dan fitur hasil rekonstruksi dihitung menggunakan Mean Squared Error (MSE) sebagai nilai reconstruction error yang berfungsi sebagai skor anomali. Nilai tersebut kemudian dibandingkan dengan threshold yang telah ditentukan untuk menghasilkan keputusan klasifikasi. Apabila nilai reconstruction error berada di bawah threshold, citra dikategorikan sebagai Normal, sedangkan nilai yang melebihi threshold dikategorikan sebagai Anomali. Mekanisme ini memungkinkan sistem mendeteksi penyimpangan terhadap distribusi data normal yang telah dipelajari selama proses pelatihan tanpa memerlukan data anomali pada tahap pembelajaran model.

3.3 Ringkasan Hasil ResNet50

Berbeda dengan DenseNet121, ketiga skenario ResNet50 (Tabel 5) menunjukkan pola yang konsisten: specificity yang sangat tinggi namun disertai sensitivity yang jauh lebih rendah. Pada latent 32, ResNet50 mencatat specificity tertinggi di antara seluruh eksperimen (95,05%) namun sensitivity terendah (76,42%), setara dengan 808 citra anomali (dari 3.426) yang gagal terdeteksi sebagai false negative. Peningkatan latent dimension dari 32 ke 128 pada ResNet50 secara konsisten menurunkan validation loss ($0,1263 \rightarrow 0,1089$) dan meningkatkan sensitivity ($76,42\% \rightarrow 84,88\%$), mengindikasikan bahwa ResNet50 membutuhkan kapasitas latent space yang lebih besar untuk menghasilkan representasi yang memadai. Meskipun demikian, bahkan konfigurasi terbaik ResNet50 (latent 128, AUC 0,9529, akurasi 85%) masih berada di bawah performa DenseNet121 latent 128 pada metrik akurasi dan sensitivity ($84,88\%$ vs $90,75\%$). Perbandingan kuantitatif rata-rata kedua feature extractor disajikan pada subbagian berikut.

3.4 Perbandingan Komprehensif DenseNet121 vs ResNet50

Tabel 8. Perbandingan Rata-Rata Performa DenseNet121 vs ResNet50 (Rata-Rata 3 Latent Dim)

Metrik (Rata-rata 3 Latent Dim)	DenseNet121	ResNet50	Selisih (DenseNet121 – ResNet50)
Validation Loss (MSE)	0,0995	0,1175	-0,0181 ↓ (DenseNet121 lebih baik)
AUC	0,9546	0,9451	0,0095
Accuracy	90,00%	81,70%	8,30%
Sensitivity	89,77%	81,07%	8,71%
Specificity	89,19%	93,59%	-4,40% (ResNet50 lebih tinggi)
Macro F1-Score	70,70%	62,00%	8,70%
False Negative Rate	10,23%	18,93%	-8,7% (DenseNet121 lebih baik)

Tabel 8 mengkonfirmasi secara kuantitatif superioritas DenseNet121 atas ResNet50 pada hampir seluruh dimensi performa. Keunggulan sensitivity rata-rata 8,71 poin persentase memiliki implikasi klinis yang sangat signifikan: dalam konteks skrining medis, setiap poin persentase sensitivity yang lebih rendah berarti lebih banyak kasus COVID-19 atau Pneumonia yang tidak terdeteksi dan tidak mendapat penanganan segera. Sebagai contoh konkret, pada konfigurasi latent 32, ResNet50 menghasilkan 808 false negative (kasus anomali yang tidak terdeteksi) dibandingkan hanya 356 dari DenseNet121 sebuah selisih 452 kasus dari 3.426 citra anomali yang diuji. Satu-satunya dimensi di mana ResNet50 lebih unggul adalah specificity dengan selisih rata-rata 4,40%, yang berarti ResNet50 lebih jarang mengklasifikasikan citra Normal sebagai Anomali; namun hal ini terjadi dengan konsekuensi false negative yang jauh lebih banyak, sebuah trade-off yang tidak menguntungkan dalam konteks skrining medis.

Perlu ditegaskan bahwa perbandingan pada Tabel 8 bertujuan mengevaluasi konsistensi arah performa (bukan signifikansi statistik formal), mengingat fokus utama penelitian ini adalah pada mekanisme klasifikasi berbasis selisih reconstruction error (MSE) terhadap threshold, bukan pada perbandingan competitive benchmarking antar arsitektur. Perbedaan performa yang konsisten pada hampir seluruh metrik (accuracy, sensitivity, F1-Score) tetap menjadi indikasi kuat keunggulan DenseNet121 sebagai feature extractor dalam pipeline yang diusulkan.

3.5 Analisis Reconstruction Error per Kelas

Tabel 9. Statistik Reconstruction Error per Kelas — Konfigurasi Terbaik (DenseNet121, Latent 128)

Statistik	Normal	COVID-19	Pneumonia
Mean Reconstruction Error	0,093	0,3637	0,1712
Threshold Optimal (Youden Index)	0,1155	0,1155	0,1155
Rasio Error terhadap Normal	1,00×	3,91×	1,84×
Selisih Mean vs Normal	—	0,2707	0,0782
Klasifikasi pada Threshold $\tau=0,1155$	Normal (di bawah τ)	Anomali (di atas τ)	Anomali (di atas τ)



Tabel 9 mengungkap pola yang sangat informatif mengenai kemampuan diskriminatif sistem. Terdapat perbedaan yang sangat signifikan antara mean reconstruction error kelas Normal (0,0930) dengan kelas COVID-19 (0,3637), dengan rasio mencapai 3,91 kali lipat. Gap yang sangat besar ini mengkonfirmasi bahwa model DAE yang dilatih pada fitur DenseNet121 kelas Normal berhasil mempelajari manifold distribusi fitur normal dengan sangat baik sehingga fitur COVID-19 yang memiliki distribusi sangat berbeda secara patologis menghasilkan reconstruction error yang jauh lebih tinggi sebagai sinyal anomali yang kuat.

Reconstruction error kelas Pneumonia (0,1712) yang berada di antara Normal (0,0930) dan COVID-19 (0,3637) konsisten dengan karakteristik visual penyakit Pneumonia. COVID-19 umumnya menunjukkan pola ground-glass opacity bilateral yang sangat difus dan khas sehingga fiturnya sangat jauh dari distribusi Normal, sementara Pneumonia memiliki morfologi yang lebih bervariasi dan beberapa kasusnya menunjukkan kemiripan parsial dengan pola citra normal (Kascenas et al., 2022). Pola ini konsisten dengan prinsip kerja sistem: reconstruction error yang dihasilkan DAE merepresentasikan seberapa jauh fitur suatu citra menyimpang dari manifold distribusi fitur Normal yang telah dipelajari selama pelatihan. Semakin besar penyimpangan pola fitur suatu citra terhadap distribusi Normal, semakin tinggi nilai MSE yang dihasilkan, sehingga selisih MSE terhadap threshold optimal inilah yang menjadi dasar keputusan klasifikasi Normal atau Anomali pada sistem ini.

3.6 Analisis Pengaruh Variasi Latent Dimension

Tabel 10. Perbandingan Performa DenseNet121 pada Variasi Latent Dimension (Data Lengkap)

Metrik	Latent 32	Latent 64	Latent 128
Validation Loss (MSE)	0,1003	0,1029	0,0952
AUC	0,9543	0,9569	0,9526
Accuracy	90%	89%	91%
Sensitivity	89,61%	88,97%	90,75%
Specificity	87,91%	92,31%	87,36%
Macro F1-Score	70%	70%	72%
True Positive (TP)	3.070	3.048	3.109
False Negative (FN)	356	378	317
True Negative (TN)	160	168	159
False Positive (FP)	22	14	23
Threshold Optimal	0,124	0,1293	0,1155
Best Epoch	57	70	56

Dari Tabel 10, beberapa pola penting dapat diidentifikasi. Pertama, peningkatan latent dimension dari 32 ke 128 secara umum menurunkan validation loss (0,1003 $\rightarrow$ 0,0952), mengindikasikan kapasitas representasi yang lebih besar memungkinkan DAE merekonstruksi distribusi fitur Normal dengan lebih akurat. Namun, terdapat pola non-monoton pada latent 64 yang menghasilkan AUC tertinggi (0,9569) dan specificity tertinggi (92,31%) namun validation loss lebih tinggi dari latent 32, mengindikasikan bahwa hubungan antara kapasitas latent space dengan performa model tidak selalu monoton pada rentang konfigurasi yang diuji dalam penelitian ini (32, 64, dan 128); dengan hanya tiga titik data, observasi ini bersifat lokal dan belum dapat digeneralisasi sebagai sifat umum hubungan tersebut (Nawaz et al., 2024).

Kedua, trade-off sensitivity-specificity yang teramati antara konfigurasi latent 64 dan 128 mencerminkan perbedaan dalam penetapan threshold optimal: latent 64 menghasilkan threshold lebih tinggi (0,1293 vs 0,1155) sehingga lebih sedikit FP (14 vs 23) namun lebih banyak FN (378 vs 317). Dalam konteks skrining medis di mana meminimalkan false negative adalah prioritas utama untuk keselamatan pasien, konfigurasi latent 128 yang menghasilkan hanya 317 FN (vs 356 pada latent 32 dan 378 pada latent 64) menjadikannya kandidat paling sesuai untuk implementasi klinis berdasarkan data pengujian ini. Ketiga, latent 128 juga menghasilkan lebih banyak TP (3.109) dibandingkan kedua konfigurasi lainnya, mengindikasikan deteksi anomali yang lebih komprehensif.

3.7 Pembahasan

Hasil penelitian ini secara empiris mengkonfirmasi hipotesis utama bahwa integrasi CNN pretrained sebagai feature extractor sebelum DAE secara signifikan meningkatkan performa deteksi anomali pada citra X-ray dada. Keunggulan DenseNet121 atas ResNet50 pada hampir seluruh metrik evaluasi berkorelasi langsung dengan perbedaan arsitekturalnya: mekanisme dense connections pada DenseNet121 menghasilkan representasi fitur yang lebih kaya dari setiap citra X-ray, sehingga DAE dapat mempelajari distribusi fitur normal dengan lebih konsisten dan menghasilkan reconstruction error yang lebih diskriminatif sebagai indikator anomali (Nawaz et al., 2024). Hal ini sejalan dengan temuan (Kascenas et al., 2022) yang menunjukkan bahwa kualitas representasi fitur sangat berpengaruh terhadap kemampuan autoencoder dalam membedakan pola normal dan anomali pada citra medis.

Dibandingkan dengan pendekatan DAE pada pixel space seperti yang diterapkan (Scarpiniti et al., 2022), pipeline DenseNet121-DAE yang diusulkan menghasilkan rasio pemisahan reconstruction error sebesar $3,91\times$ antara COVID-19 dan Normal. Sebagai perbandingan konseptual, pendekatan yang beroperasi langsung pada piksel memiliki keterbatasan dalam menghasilkan pemisahan yang setajam ini karena perbedaan piksel antar kelas dipengaruhi oleh variasi pencahayaan, artefak, dan noise yang tidak relevan secara semantis. Representasi fitur 1.024 dimensi dari



DenseNet121 yang memanfaatkan pengetahuan dari 1,2 juta gambar ImageNet secara inheren lebih robust terhadap variasi tersebut (Farooq & Savaş, 2024).

Nilai AUC 0,9526 yang dicapai konfigurasi terbaik berada dalam kisaran yang dilaporkan oleh (Nawaz et al., 2024) untuk ensemble autoencoder (0,90–0,98) pada berbagai jenis data biomedis, namun dengan keunggulan arsitektur single model yang secara struktural lebih sederhana dibandingkan pendekatan ensemble, yang berpotensi lebih mudah diimplementasikan dalam praktik klinis. Pengukuran efisiensi komputasi secara kuantitatif (waktu inferensi, jumlah parameter) belum dilakukan pada penelitian ini dan menjadi salah satu arah evaluasi lanjutan. Nilai sensitivity 90,75% yang diperoleh berada pada level yang kompetitif, meskipun perbandingan langsung dengan nilai yang dilaporkan (Kim & Soo Kim, 2024) pada domain suara (~72% sebelum peningkatan DDAE) perlu dimaknai secara hati-hati mengingat perbedaan karakteristik data dan tugas deteksi anomali antar domain (audio vs. citra medis); dengan catatan tersebut, hasil ini tetap mengindikasikan bahwa representasi fitur visual dari DenseNet121 memberikan fondasi yang baik untuk sistem deteksi anomali yang diusulkan.

Mengenai keterbatasan penelitian: presisi kelas Normal yang relatif rendah (33%) pada konfigurasi terbaik perlu dipahami dalam konteks ketidakseimbangan kelas yang sangat signifikan pada data uji (rasio 1:18,8 antara Normal dan Anomali). Dengan hanya 182 citra Normal berbanding 3.426 citra Anomali, bahkan sejumlah kecil false positive (23 citra) secara matematis menekan presisi kelas Normal secara dramatis. Ini merupakan karakteristik umum pada skenario unsupervised anomaly detection dan bukan indikasi kelemahan fundamental sistem (Torabi et al., 2023). Keterbatasan kedua adalah evaluasi yang dilakukan pada satu dataset publik, sehingga generalisasi pada distribusi dataset berbeda dari institusi lain masih memerlukan validasi lebih lanjut. Ketiga, sistem belum memiliki kemampuan lokalisasi area anomali (anomaly localization) yang akan sangat meningkatkan nilai klinis sistem sebagai alat bantu diagnosis. Hasil yang diperoleh menunjukkan bahwa masih terdapat ruang untuk penyempurnaan sistem pada penelitian berikutnya. Dari sisi ekstraksi fitur, arsitektur seperti EfficientNet, Vision Transformer (ViT), maupun varian DenseNet yang lebih dalam dapat dieksplorasi untuk memperoleh representasi fitur yang lebih optimal. Dari sisi interpretabilitas, penerapan pendekatan seperti Grad-CAM atau attention visualization berpotensi membantu menjelaskan area citra yang berkontribusi terhadap tingginya nilai reconstruction error. Sementara itu, pengujian pada dataset yang lebih heterogen dari berbagai institusi juga diperlukan untuk mengevaluasi konsistensi performa model pada distribusi data yang berbeda, termasuk kemungkinan penerapan mekanisme threshold adaptation guna mempertahankan kinerja sistem ketika terjadi perubahan karakteristik data dari waktu ke waktu.

4. KESIMPULAN

Temuan penelitian ini menjawab rumusan masalah secara langsung: integrasi DenseNet121 pretrained ImageNet sebagai feature extractor sebelum Denoising Autoencoder (DAE) terbukti meningkatkan kemampuan deteksi anomali unsupervised pada citra X-ray dada dibandingkan penggunaan feature extractor ResNet50, tanpa memerlukan data berlabel anomali selama pelatihan. Yang menarik, pemilihan arsitektur feature extractor justru memberikan pengaruh jauh lebih dominan terhadap performa akhir dibandingkan penyesuaian latent dimension, sehingga kualitas representasi fitur, bukan penyetelan hyperparameter DAE, menjadi faktor penentu utama keberhasilan deteksi anomali berbasis reconstruction error pada domain citra medis. Temuan ini membawa implikasi praktis yang cukup relevan: pendekatan unsupervised semacam ini menawarkan alternatif deteksi anomali yang tidak bergantung pada ketersediaan data berlabel dalam jumlah besar, sehingga berpotensi diterapkan pada fasilitas kesehatan dengan keterbatasan sumber daya anotasi. Perbedaan pola reconstruction error antar kelas (Normal, Pneumonia, COVID-19) turut mengindikasikan potensi sistem dalam mengenali variasi tingkat abnormalitas citra, meskipun validasi lebih lanjut terhadap label keparahan klinis tetap diperlukan sebelum temuan ini dapat mendukung aplikasi triase. Keterbatasan utama penelitian ini terletak pada belum diujinya baseline DAE konvensional pada ruang piksel sebagai pembanding, sehingga besarnya kontribusi feature extractor pretrained terhadap performa akhir belum dapat diisolasi secara pasti. Penelitian selanjutnya disarankan menambahkan baseline tersebut, mengeksplorasi arsitektur feature extractor alternatif seperti EfficientNet atau Vision Transformer, serta menguji generalisabilitas sistem pada dataset multi-institusi dengan variasi populasi dan perangkat pencitraan yang lebih luas.

REFERENCES

- Alrayes, F. S., Zakariah, M., Amin, S. U., Iqbal Khan, Z., & Helal, M. (2024). Intrusion Detection in IoT Systems Using Denoising Autoencoder. *IEEE Access*, 12, 122401–122425. <https://doi.org/10.1109/ACCESS.2024.3451726>
- Alvarado, W., Agrawal, V., Li, W. S., Dravid, V. P., Backman, V., de Pablo, J. J., & Ferguson, A. L. (2023). Denoising Autoencoder Trained on Simulation-Derived Structures for Noise Reduction in Chromatin Scanning Transmission Electron Microscopy. *ACS Central Science*, 9(6), 1200–1212. <https://doi.org/10.1021/acscentsci.3c00178>
- Chicco, D., & Jurman, G. (2023). The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Mining*, 16(1), 4. <https://doi.org/10.1186/s13040-023-00322-4>
- FAROOQ, Y., & SAVAŞ, S. (2024). Noise Removal from the Image Using Convolutional Neural Networks-Based Denoising Auto Encoder. *Journal of Emerging Computer Technologies*, 3(1), 21–28. <https://doi.org/10.57020/ject.1390428>



- He, J., Xu, Q., Jiang, Y., Wang, Z., & Huang, Q. (2024). ADA-GAD: Anomaly-Denoised Autoencoders for Graph Anomaly Detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(8), 8481–8489. <https://doi.org/10.1609/aaai.v38i8.28691>
- Kascenas, A., Pugeault, N., & O’Neil, A. Q. (2022). Denoising Autoencoders for Unsupervised Anomaly Detection in Brain MRI. *Proceedings of Machine Learning Research*, 172, 653–664.
- Kim, S.-M., & Soo Kim, Y. (2024). Enhancing Sound-Based Anomaly Detection Using Deep Denoising Autoencoder. *IEEE Access*, 12, 84323–84332. <https://doi.org/10.1109/ACCESS.2024.3414435>
- Liu, T., Li, B., Du, X., Jiang, B., Geng, L., Wang, F., & Zhao, Z. (2025). Simple and effective Frequency-aware Image Restoration for industrial visual anomaly detection. *Advanced Engineering Informatics*, 64, 103064. <https://doi.org/10.1016/j.aei.2024.103064>
- Mujkic, E., Philipsen, M. P., Moeslund, T. B., Christiansen, M. P., & Ravn, O. (2022). Anomaly Detection for Agricultural Vehicles Using Autoencoders. *Sensors*, 22(10), 3608. <https://doi.org/10.3390/s22103608>
- Nawaz, A., Khan, S. S., & Ahmad, A. (2024). Ensemble of Autoencoders for Anomaly Detection in Biomedical Data: A Narrative Review. *IEEE Access*, 12, 17273–17289. <https://doi.org/10.1109/ACCESS.2024.3360691>
- Park, H., Jeong, J., Oh, K.-W., & Kim, H. (2023). Autoencoder-Based Recommender System Exploiting Natural Noise Removal. *IEEE Access*, 11, 30609–30618. <https://doi.org/10.1109/ACCESS.2023.3262026>
- Patil, R., & Bhosale, S. (2022). Medical Image Denoising Techniques: A Review. *International Journal on Engineering, Science and Technology*, 4(1), 21–33. <https://doi.org/10.46328/ijonest.76>
- Raghuvamsi, Y., Teeparthi, K., & Kosana, V. (2023). Denoising autoencoder based topology identification in distribution systems with missing measurements. *International Journal of Electrical Power & Energy Systems*, 154, 109464. <https://doi.org/10.1016/j.ijepes.2023.109464>
- Rajput, G., Agrawal, S., Biyani, K., & Vishvakarma, S. K. (2022). Early breast cancer diagnosis using cogent activation function-based deep learning implementation on screened mammograms. *International Journal of Imaging Systems and Technology*, 32(4), 1101–1118. <https://doi.org/10.1002/ima.22701>
- Sachin Kumar. (2022). *Covid19-Pneumonia-Normal Chest X-Ray Images*. Mendeley Data. <https://doi.org/10.17632/dvntn9yhd2.1>
- Scarpiniti, M., Sarv Ahrabi, S., Baccarelli, E., Piazza, L., & Momenzadeh, A. (2022). A novel unsupervised approach based on the hidden features of Deep Denoising Autoencoders for COVID-19 disease detection. *Expert Systems with Applications*, 192, 116366. <https://doi.org/10.1016/j.eswa.2021.116366>
- Siddiqi, R., & Javaid, S. (2024). Deep Learning for Pneumonia Detection in Chest X-ray Images: A Comprehensive Survey. *Journal of Imaging*, 10(8), 176. <https://doi.org/10.3390/jimaging10080176>
- Thiele, C., & Hirschfeld, G. (2021). cutpointnr : Improved Estimation and Validation of Optimal Cutpoints in R. *Journal of Statistical Software*, 98(11). <https://doi.org/10.18637/jss.v098.i11>
- Torabi, H., Mirtaheri, S. L., & Greco, S. (2023). Practical autoencoder based anomaly detection by using vector reconstruction error. *Cybersecurity*, 6(1), 1. <https://doi.org/10.1186/s42400-022-00134-9>
- Yang, H., Aydi, W., Innab, N., Ghoneim, M. E., & Ferrara, M. (2024). Classification of cervical cancer using Dense CapsNet with Seg-UNet and denoising autoencoders. *Scientific Reports*, 14(1), 31764. <https://doi.org/10.1038/s41598-024-82489-2>
- Zhou, S., He, Z., Chen, X., & Chang, W. (2024). An Anomaly Detection Method for UAV Based on Wavelet Decomposition and Stacked Denoising Autoencoder. *Aerospace*, 11(5), 393. <https://doi.org/10.3390/aerospace11050393>