

Lung Disease Detection Using Gradient-Weighted Class Activation Mapping (Grad-CAM)

Wan Sofiyah¹, Benny Sukma Negara², Muhammad Irsyad³, Iwan Iskandar⁴, Febi Yanto⁵
^{1,2,3}Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim, Pekanbaru, 28293, Indonesia

Informasi Artikel

Diterima : 30 Mei 2025
Revisi : 6 Juni 2025
Publikasi : 20 Juni 2025

Kata Kunci:

XAI
Covid-19
Pneumonia
Chest X-Ray
Grad-CAM

ABSTRAK

Deteksi dini terhadap penyakit pernapasan seperti *Coronavirus Disease-19* (Covid-19) dan Pneumonia sangat penting untuk mempercepat penanganan dan mencegah komplikasi yang lebih serius. Penelitian ini mengusulkan metode klasifikasi citra Chest X-ray (CXR) menggunakan Convolutional Neural Network (CNN) untuk membedakan Covid-19, Pneumonia, dan paru-paru normal. Pelatihan model melibatkan eksplorasi berbagai kombinasi *hyperparameter* untuk menemukan konfigurasi optimal. Hasil terbaik dicapai pada kombinasi *learning rate* 0.001, 50 *epoch*, dan *batch size* 32, dengan akurasi 96.33%. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, dan *confusion matrix*. Penelitian ini menggunakan *Gradient-Weighted Class Activation Mapping* (Grad-CAM) sebagai alat interpretasi transparan dari keputusan model. Kontribusi utama dari penelitian ini adalah penerapan Grad-CAM pada klasifikasi *multi-class* CXR untuk meningkatkan interpretabilitas model dalam diagnosis penyakit paru.

ABSTRACT

Early detection of respiratory diseases such as Coronavirus Disease-19 (Covid-19) and Pneumonia is crucial for accelerating treatment and preventing more serious complications. This study proposes a method for classifying Chest X-ray (CXR) images using a Convolutional Neural Network (CNN) to distinguish between Covid-19, Pneumonia, and normal lungs. Model training involved exploring various hyperparameter combinations to find the optimal configuration. The best results were achieved with a learning rate of 0.001, 50 epochs, and a batch size of 32, yielding an accuracy of 96.33%. Evaluation was conducted using accuracy, precision, recall, F1-score, and confusion matrix metrics. This study uses Gradient-Weighted Class Activation Mapping (Grad-CAM) as a transparent interpretation tool for model decisions. The main contribution of this study is the application of Grad-CAM in multi-class CXR classification to enhance model interpretability in lung disease diagnosis.

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



*Penulis Koresponden

Email: bsnegara@uin-suska.ac.id

Cara sitasi IEEE::

W. Sofiyah, B.S. Negara, M. Irsyad, I. Iskandar, F. Yanto, "Lung Disease Detection Using Gradient-Weighted Class Activation Mapping (Grad-CAM)," *Journal of Artificial Intelligence and Software Engineering (J-AISE)*, vol. 5, no. 2, pp. 720-730, Juni 2025. doi: 10.30811/jaise.v5i2.7041

1. PENDAHULUAN

Kesehatan paru merupakan salah satu aspek vital dalam menjaga fungsi sistem pernapasan yang optimal, mengingat paru-paru berperan langsung dalam pertukaran oksigen dan karbon dioksida yang esensial bagi kehidupan. Permasalahan pada paru, seperti infeksi, inflamasi, atau gangguan kronis, dapat menurunkan kualitas hidup dan bahkan berujung pada kematian bila tidak ditangani secara tepat [1]. Selama ini, proses identifikasi penyakit paru umumnya dilakukan secara manual melalui pemeriksaan klinis, radiologis, dan laboratorium, yang meskipun efektif, memiliki keterbatasan dalam hal kecepatan, akurasi, dan ketergantungan pada keahlian tenaga medis [2]. Dalam konteks ini peran teknologi, khususnya dalam bidang kecerdasan buatan dan analisis citra medis, mulai memainkan peran penting dalam mendukung deteksi dini dan diagnosis penyakit paru secara lebih efisien, akurat, dan terstandarisasi.

Penyakit paru dapat terlihat dari perubahan struktur dan pola radiologis pada organ paru, salah satunya melalui pemeriksaan CXR yang merupakan metode pencitraan dasar namun krusial dalam evaluasi kondisi paru [3]. Hal tersebut menjadi penting dalam upaya deteksi dini penyakit paru, karena perubahan yang teridentifikasi pada tahap awal dapat membantu mencegah progresivitas penyakit dan komplikasi serius. Terdapat jenis penyakit paru yang dapat dikenali melalui temuan khas pada CXR, seperti pneumonia yang menunjukkan *consolidation* yaitu berupa bercak putih tebal yang umumnya berisi cairan, nanah, dan menunjukkan adanya infeksi berat [4]. Lain halnya Covid-19 yang memperlihatkan *Ground-Glass Opacity* (GGO) yaitu bercak putih yang tidak terlalu tebal yang menunjukkan adanya inflamasi, infeksi dan cedera ringan, di zona perifer dan bawah paru [5]. Dengan melakukan deteksi dini terhadap penyakit paru melalui analisis CXR, dokter dapat segera mengambil langkah penanganan yang tepat, mengoptimalkan pengobatan, dan meningkatkan peluang kesembuhan serta keselamatan pasien.

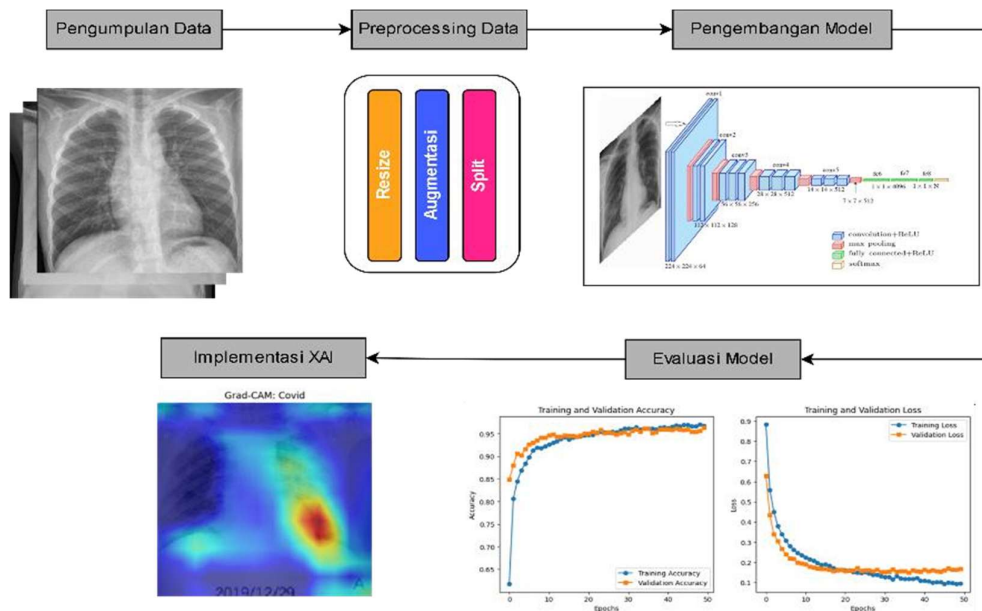
Penelitian ini menemukan bahwa kemiripan fitur visual antara citra Covid-19 dan Pneumonia menjadi tantangan tersendiri dalam proses klasifikasi, karena keduanya menunjukkan karakteristik radiologis yang saling tumpang tindih. Pada hakikatnya Pneumonia adalah infeksi yang disebabkan oleh virus, bakteri, atau jamur. Mirip dengan Covid-19, Pneumonia rentan mengancam jiwa kelompok usia di bawah 2 tahun dan orang berusia di atas 65 tahun. Pneumonia merupakan penyakit menular seperti halnya Covid-19 dengan beberapa gejala yang mirip. Namun, Covid-19 telah terbukti lebih fatal daripada Pneumonia, karena dapat menyebabkan *Acute Respiratory Distress Syndrome* (ARDS) [6].

Karena belum tersedia metode diagnosis yang cerdas, saat ini dibutuhkan pendekatan yang mampu mengidentifikasi pasien yang diduga terinfeksi Covid-19 secara cepat dan akurat. Meskipun metode deteksi yang ada saat ini cukup andal, prosesnya masih memerlukan waktu yang relatif lama. Saat ini, terdapat tiga jenis metode utama yang digunakan untuk mendeteksi Covid-19, yaitu *Reverse Transcription Polymerase Chain Reaction* (RT-PCR), *Computed Tomography* (CT) *Scan*, dan CXR. Di antara ketiganya, RT-PCR merupakan salah satu teknik yang memakan waktu, sedangkan CXR dapat mendeteksi peradangan pada paru-paru beserta posisi, bentuk, dan ukurannya, dan CT-Scan merupakan cara yang lebih efektif untuk mendiagnosis Covid-19 dan Pneumonia karena dapat memberikan gambaran rinci tentang kantung udara, namun teknik ini memiliki suatu kekurangan yaitu memerlukan biaya yang cukup mahal.

Oleh karena itu, dalam penelitian ini, kami memilih CXR sebagai media utama untuk klasifikasi karena bersifat lebih praktis, cepat, dan ekonomis. Selain itu, CXR mampu memberikan informasi visual yang cukup untuk mendeteksi indikasi adanya infeksi atau kelainan pada paru-paru, termasuk yang disebabkan oleh Covid-19 maupun Pneumonia. Dengan dukungan teknologi deep learning, citra CXR dapat dimanfaatkan secara optimal untuk membantu proses diagnosis secara otomatis dan akurat. Pendekatan ini diharapkan dapat menjadi solusi pendukung dalam sistem deteksi dini penyakit paru yang efisien dan mudah diimplementasikan.

2. METODE

Untuk mencapai tujuan penelitian, digunakan pendekatan metodologis yang sistematis yang mencakup pengumpulan data, *preprocessing* data, pemodelan menggunakan arsitektur model, evaluasi model, serta implementasi XAI dengan Grad-CAM [7].



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data

Dataset diperlukan untuk membangun model komputasi yang efektif sehingga dapat akurat dalam melakukan deteksi medis. Dataset yang digunakan merupakan data CXR dengan resolusi asli 256x256 piksel dan 232x232 piksel yang tersedia secara publik [8]. Dataset dibagi menjadi tiga kelas yaitu Covid-19, Normal (tanpa kelainan paru), dan Pneumonia dengan total jumlah 3.000 data training dengan 1.000 data untuk masing-masing kelas. Selanjutnya, kami menggunakan 30 data CXR tambahan yang belum pernah dilatih untuk dijadikan sebagai data uji, dengan masing-masing 10 data perkategori kelas yang sama.



Gambar 2. Dataset CXR pada kategori (a) Covid-19 (b) Normal (c) Pneumonia

2.2 Pre-processing Data

Data CXR yang digunakan melalui proses *preprocessing* terlebih dahulu untuk menyeragamkan ukuran citra, menormalkan piksel, dan augmentasi data. Adapun tahapan yang dilakukan dalam *preprocessing*, diantaranya:

2.2.1 Resize

Proses ini dilakukan dengan mengubah ukuran citra asli menjadi 224×224 piksel untuk memenuhi persyaratan input model VGG16. Ukuran ini dipilih karena merupakan standar input pada arsitektur yang digunakan, yang memungkinkan fitur spasial tetap terjaga tanpa membebani proses komputasi.

2.2.2 *Split Data*

Dataset dibagi menjadi dua bagian utama menggunakan fungsi `train_test_split`, dengan rasio 80% untuk latih dan 20% untuk validasi. Untuk memastikan distribusi kelas tetap seimbang, pembagian dilakukan secara stratifikasi berdasarkan label asli. Dengan demikian, model dapat dilatih secara optimal dan diuji pada data yang benar-benar baru untuk mengukur generalisasi.

2.2.3 *Augmentasi data*

Augmentasi data adalah teknik untuk meningkatkan keragaman dan jumlah data yang digunakan untuk melatih model *machine learning* (ML) guna mengurangi *overfitting* dan meningkatkan generalisasi model. Teknik augmentasi yang digunakan yaitu rotasi acak hingga 15° untuk menambahkan pemahaman model terhadap data dan dengan pengisian area kosong menggunakan nilai piksel terdekat (`fill_mode="nearest"`). Proses ini diterapkan secara *real-time* selama pelatihan menggunakan modul `ImageDataGenerator`.

2.3 **Pengembangan Model**

2.3.1 *Convolutional Neural Network (CNN)*

Convolutional Neural Networks (CNN) merupakan arsitektur *deep learning* yang efektif untuk analisis data visual melalui ekstraksi dan pembelajaran fitur secara hierarkis. CNN terdiri atas *convolution layer*, *pooling*, *fully connected*, dan *non-linear activation* yang meniru mekanisme kerja visual manusia. Pada *convolution layer*, digunakan filter untuk memindai fitur penting dari citra dan menghasilkan *feature map* sebagai representasi spasial dari informasi yang relevan [9]. Filter tersebut berperan dalam mengenali pola lokal seperti tepi, tekstur, dan fitur-fitur penting lainnya yang membantu dalam memahami struktur visual yang kompleks [10]

Secara matematis, proses konvolusi dilakukan dengan menggeser filter di atas gambar input, kemudian menghitung hasil perkalian *element-wised* dan menjumlahkannya [11]. Hasil dari proses ini adalah peta fitur yang menekankan karakteristik penting dari data input, sehingga model dapat berfokus pada informasi yang relevan dan menyaring data yang tidak signifikan. Dengan pendekatan bertingkat dan terstruktur ini, CNN menjadi alat yang sangat tepat untuk menangani tantangan dalam analisis citra medis, memberikan hasil yang akurat dan andal [12] Penelitian terkait menemukan potensi CNN dalam pengembangan model deteksi penyakit yang akurat dan efisien pada berbagai aplikasi termasuk medis menggunakan gambar *CT-scan* untuk memprediksi kanker paru-paru dengan penerapan CNN dan mencapai rata-rata akurasi yang cukup baik dengan 72,41% [13].

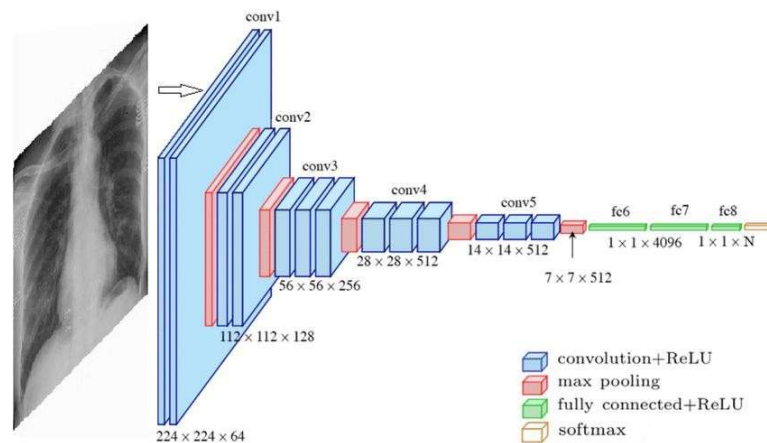
2.3.2 *Visual Geometry Group 16 (VGG16)*

Visual Geometry Group 16 (VGG16) merupakan salah satu arsitektur CNN populer yang terdiri dari 16 lapisan, yang dirancang khusus untuk keperluan klasifikasi citra. Model ini banyak digunakan dalam bidang pencitraan medis karena kemampuannya dalam mengekstrak fitur secara bertingkat, sehingga menghasilkan performa yang handal. Temuan sebelumnya mengklasifikasikan gambar *CT-scan* dengan VGG16 untuk mendeteksi tingkat penyakit kanker paru-paru dengan tiga kategori: jinak, normal, dan ganas. Hasilnya model dapat mengenali dan membedakan kondisi paru-paru dengan efektif [14]. Penelitian lain menguji berbagai arsitektur CNN untuk deteksi kanker paru-paru menggunakan *CT-scan* dan *PET scan*. Hasilnya VGG16 unggul karena mampu menganalisis jaringan kanker paru-paru dengan tepat dan meningkatkan akurasi diagnostic dibandingkan arsitektur lain [15].

VGG16 menggunakan lapisan konvolusi berfilter kecil (3×3) untuk mengekstraksi fitur bertingkat, mulai dari pola sederhana hingga struktur kompleks, sehingga memungkinkan representasi citra medis yang akurat dan efisien [16] Filter tersebut secara bertahap menangkap berbagai fitur, mulai dari pola sederhana seperti tepi hingga struktur yang lebih kompleks. Hal ini memungkinkan representasi citra medis yang efektif.

Setiap operasi konvolusi dilanjutkan dengan *activation function* ReLU yang memperkenalkan elemen non-linearitas, meningkatkan kemampuan model untuk mempelajari dan memodelkan hubungan kompleks dalam data [17]

Fitur yang diperoleh dari lapisan konvolusi kemudian diproses oleh *fully-connected layer* yang bertugas mengintegrasikan informasi untuk keperluan klasifikasi. Contohnya, dalam klasifikasi citra medis, VGG16 dapat mengkategorikan gambar menjadi kelas seperti normal, jinak, atau ganas [14]. Terakhir, lapisan *softmax* mengubah hasil akhir menjadi distribusi probabilitas, memberikan kemungkinan prediksi yang jelas untuk tiap kategori, sehingga memungkinkan prediksi menghasilkan akurasi yang baik. Struktur yang rapi dan sistematis ini menjadikan VGG16 berguna untuk analisis gambar medis, khususnya dalam mendeteksi dan mengklasifikasikan gangguan pada paru-paru. Kombinasi antara ekstraksi fitur secara bertahap dan kemampuan klasifikasi yang kuat menjamin hasil yang tepat dan dapat dipercaya, yang sangat penting dalam aplikasi medis yang kritis [18].



Gambar 3. Arsitektur VGG16 pada CNN

2.3.3 Konfigurasi Pelatihan Model

Penelitian ini menggunakan *transfer learning* dengan menginisialisasi model VGG16 dengan bobot yang telah dilatih sebelumnya dari dataset *ImageNet*, memanfaatkan kemampuan *feature extraction* yang *robust* yang dikembangkan pada kumpulan gambar besar dan beragam. Model ini kemudian di *fine-tuning* pada dataset CXR untuk menyesuaikan secara khusus proses klasifikasi kelainan paru. Selain itu, digunakan optimisasi menggunakan *Adaptive Moment Estimation* (ADAM) yaitu algoritma *optimizer* yang populer untuk CNN, yang dikenal dengan momentum adaptifnya [19]. Penelitian terbaru telah mengeksplorasi penerapan ADAM dengan arsitektur VGG16 di berbagai domain. Dalam klasifikasi tumor otak, VGG16 dengan ADAM mencapai akurasi 97,15% pada dataset uji, mengungguli *optimizer* lainnya [20].

Adapun dalam memilih kombinasi *hyperparameter* kami menggunakan teknik *Manual Random Search* digunakan dengan tujuan mendapatkan konfigurasi. Studi terkini membandingkan *random search* dan teknik lain bernama *Grid Search*. Teknik *grid search* menawarkan eksplorasi menyeluruh, tetapi memerlukan sumber daya komputasi yang besar. Di sisi lain, *random search* memberikan hasil yang lebih cepat melalui pemilihan acak, meskipun tidak mencakup seluruh kombinasi *hyperparameter* secara sistematis [21].

Table 1. *Hyperparameter* pelatihan model

<i>Hyperparameter</i>	Nilai
Rasio data	80:20
<i>Optimizer</i>	Adam
<i>Activation function</i>	ReLU
<i>Learning rate</i>	0.00001 – 0.1

<i>Epoch</i>	25, 50, 100
<i>Batch size</i>	8,32, 64

2.4 Evaluasi Model

Evaluasi performa model dilakukan dengan menggunakan *classification report* yang mencakup metrik-metrik utama seperti *precision*, *recall*, *f1-score*, dan *accuracy* untuk masing-masing kelas. Selain itu, *confusion matrix* disertakan guna memberikan representasi yang lebih detail terkait kemampuan model dalam mengklasifikasikan citra CXR ke dalam tiga kategori, yaitu Covid-19, Normal, dan Pneumonia. Perhitungan metrik-metrik evaluasi tersebut didasarkan pada rumus standar yang dijelaskan berikut ini:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Keterangan:

- TP = Jumlah data positif yang diklasifikasi benar.
- TN = Jumlah data negatif yang diklasifikasi benar.
- FP = Jumlah data positif yang diklasifikasi salah.
- FN = Jumlah data negatif yang diklasifikasi salah.

2.5 Implementasi XAI dengan Grad-CAM

Explainable Artificial Intelligence (XAI) dianggap perlu untuk menafsirkan bagaimana model mengambil suatu keputusan dengan transparan. XAI dikembangkan untuk mengklasifikasikan metode interpretasi citra berbasis *deep learning*, yang menekankan peranan penting penjelasan dalam pengambilan keputusan model [22]. Dalam klasifikasi citra CXR untuk Covid-19, Pneumonia, dan Tuberkulosis, oleh *deep learning* dengan metode XAI dapat mencapai akurasi pengujian 94.31% [23]. Grad-CAM telah muncul sebagai alat visualisasi XAI yang kuat untuk menginterpretasikan model CNN [24].

Grad-CAM bekerja dengan meng-*highlight* area fitur yang mempengaruhi keputusan model untuk menghasilkan *heatmap*. *Heatmap* yang dihasilkan kemudian ditumpangkan pada gambar CXR. Pendekatan ini memberikan representasi yang dapat ditafsirkan dari proses pengambilan keputusan model dengan lebih terbuka. Warna peta panas Grad-CAM menunjukkan pentingnya atau relevansi berbagai elemen gambar masukan terhadap keputusan model. Warna yang lebih hangat (misalnya, merah) menandakan pentingnya yang lebih besar, sementara warna yang lebih dingin (misalnya, biru) menunjukkan pentingnya yang lebih kecil. Visualisasi peta panas memungkinkan identifikasi area pada gambar yang paling berkontribusi terhadap keluaran model, sehingga membantu dalam memahami proses pengambilan keputusannya. [25].

Dalam diagnosis medis, *heatmap* yang dihasilkan Grad-CAM telah terbukti membantu para ahli radiologi dalam menafsirkan hasil klasifikasi berbasis AI [26]. Penelitian lain menyebutkan metode ini mengungguli metode lain seperti CAM dan Grad-CAM++ dalam melokalisasi area penting untuk prediksi CNN [27]. Penelitian ini menggunakan Grad-CAM untuk meningkatkan kemampuan interpretasi prediksi model, sehingga proses diagnostik menjadi lebih transparan dan dapat diandalkan.

Secara umum, langkah-langkah Grad-CAM meliputi:

1. Melakukan *forward pass* untuk memperoleh prediksi model.
2. Menghitung gradien dari skor kelas target terhadap keluaran dari *convolution layer* terakhir.
3. Menghitung rata-rata gradien global untuk masing-masing *feature maps*.

4. Melakukan kombinasi linier antara *feature map* dan bobot gradien untuk menghasilkan *class activation map*.
5. Melakukan ReLU pada *activation map* agar hanya mempertahankan fitur yang mendukung kelas target.

Berikut rumus perhitungan Grad-CAM:

$$L_{Grad-CAM}^c = ReLU \left(\sum_k \alpha \frac{c}{k} A^k \right) \quad (5)$$

Keterangan:

- $L_{Grad-CAM}^c$ = *Lokalisasi map* untuk kelas c
- $\alpha \frac{c}{k}$ = Bobot *feature map* ke- k
- A^k = *Activation map* lapisan k

3. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dari eksperimen yang telah dilakukan serta pembahasan terhadap temuan yang diperoleh. Fokus utama berada pada evaluasi performa model berdasarkan pengaturan *hyperparameter* yang bervariasi. Selain itu, dilakukan analisis terhadap pengaruh masing-masing konfigurasi terhadap akurasi model guna mengidentifikasi kombinasi yang memberikan hasil paling optimal.

3.1 Hasil Pelatihan Model

Bagian ini menyajikan sejumlah eksperimen pelatihan model. Skenario pengaturan menggunakan kombinasi *hyperparameter* yang beragam, diantaranya: *learning rate*, *epoch*, dan *batch size* yang menghasilkan nilai akurasi yang berbeda. Skenario disusun dengan *manual random search* dengan sebagian percobaan didasarkan pada penelitian yang sudah ada. Tujuannya untuk mengidentifikasi kombinasi parameter terbaik yang mampu menghasilkan kinerja model paling optimal pada dataset yang digunakan. Hasil skenario eksperimen tersebut disajikan pada Tabel 2. Tabel ini memperlihatkan performa model pada setiap kombinasi *hyperparameter* yang diuji, termasuk nilai akurasi yang diperoleh. Melalui penyajian ini, dapat dianalisis pola dan pengaruh masing-masing parameter terhadap kinerja model secara keseluruhan.

Tabel 2. Eksperimen Pelatihan Model

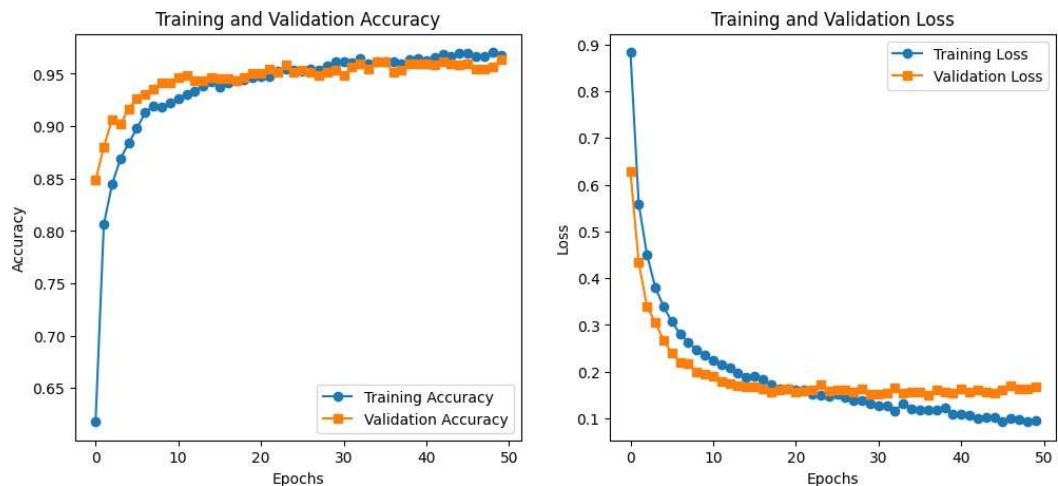
No.	Rasio	<i>Learning Rate</i>	<i>Epoch</i>	<i>Batch Size</i>	<i>Accuracy</i>
1	80:20	0.01	25	8	94.83%
2	80:20	0.001	25	8	95.67%
3	80:20	0.0001	25	8	91.00%
4	80:20	0.01	50	8	95.50%
5	80:20	0.001	50	8	95.33%
6	80:20	0.01	50	32	96.00%
7	80:20	0.001	50	32	96.33%
8	80:20	0.0001	50	32	92.33%
9	80:20	0.00001	50	32	84.00%
10	80:20	0.0001	50	64	92.50%
11	80:20	0.001	100	8	94.67%
12	80:20	0.0001	100	32	94.17%

Tabel 2. menyajikan hasil eksperimen pelatihan model dengan berbagai skenario, Seluruh eksperimen menggunakan rasio data latih dan validasi 80:20, *activation function* ReLU, dan *optimizer* Adam. Berdasarkan tabel, kombinasi *hyperparameter* dengan *learning rate* 0.001, *epoch* 50, dan *batch size* 32 menghasilkan akurasi tertinggi dengan 96.33%. Sebaliknya, penggunaan *learning rate* 0.00001 dan kombinasi *epoch* dan

batch size yang sama justru menunjukkan penurunan dengan akurasi hanya mencapai 84.00%. Hal ini mengindikasikan bahwa *learning rate* yang terlalu rendah menyebabkan proses pelatihan berjalan tidak efektif.

3.2 Evaluasi Kinerja Model

Untuk memantau efektivitas proses pelatihan model, dilakukan visualisasi terhadap perkembangan nilai akurasi dan *loss* pada data latih dan validasi di setiap *epoch*. Visualisasi ini bertujuan untuk mengidentifikasi potensi *underfitting* atau *overfitting* selama pelatihan berlangsung. Grafik berikut menunjukkan perubahan nilai akurasi dan *loss* sepanjang 50 *epoch* pelatihan.



Gambar 4. Grafik Akurasi dan Loss

Dapat dilihat bahwa nilai akurasi pada data pelatihan meningkat secara konsisten seiring bertambahnya *epoch*. Akurasi pada data validasi juga menunjukkan tren positif dan relatif stabil, yang menandakan bahwa model tidak mengalami *overfitting* serta mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat sebelumnya. Sementara itu, pada grafik *loss* menunjukkan perubahan nilai kurva mengalami penurunan yang konsisten, terutama pada fase awal pelatihan, yang mengindikasikan bahwa model berhasil meminimalkan kesalahan prediksi. Dari Gambar 4. di atas disimpulkan bahwa proses pelatihan berlangsung dengan baik dan model dapat dievaluasi lebih lanjut melalui evaluasi klasifikasi dan interpretasi visual.

3.3 Evaluasi Klasifikasi

Model dengan performa terbaik tersebut selanjutnya digunakan sebagai dasar untuk tahap evaluasi lanjutan, termasuk evaluasi analisis metrik pada *accuracy*, *presisi*, *recall*, dan *F1-score*.

Tabel 3. Hasil Klasifikasi Model

	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Covid-19	0.99	0.98	0.99
Normal	0.95	0.99	0.97
Pneumonia	0.98	0.95	0.96
Akurasi			0.97

Hasil evaluasi ditampilkan pada Tabel 3 menunjukkan bahwa model memiliki performa klasifikasi yang baik di semua kelas. Selain itu, dilakukan pula evaluasi *confusion matrix* untuk memahami lebih lanjut pola prediksi model. *Confusion matrix* dapat mengidentifikasi jumlah prediksi yang benar maupun salah pada setiap kelas.

Tabel 4. Confusion Matrix

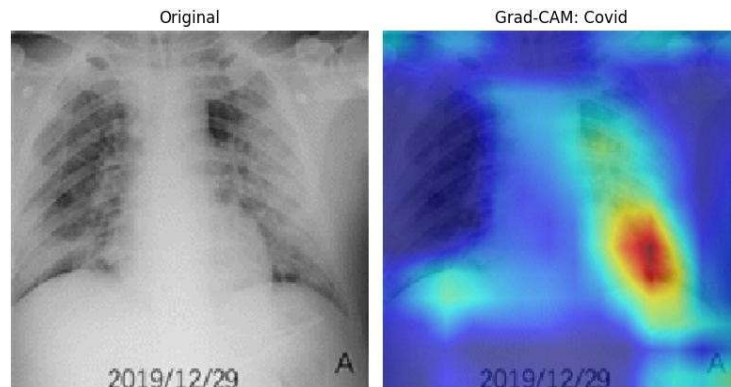
	Covid-19	Normal	Pneumonia
Covid-19	197	2	1
Normal	0	196	4
Pneumonia	2	5	193

Berdasarkan Tabel 4, dapat dilihat bahwa model berhasil mengklasifikasikan sebagian besar data dengan benar. Dari jumlah 200 sampel kelas Covid-19, sebanyak 197 terklasifikasi dengan benar, sementara sisanya keliru diklasifikasikan sebagai Normal dan Pneumonia. Hal serupa terjadi pada kelas Normal dan Pneumonia, di mana tingkat kesalahan prediksi juga relatif rendah. Hal ini menunjukkan bahwa model memiliki kemampuan generalisasi yang baik terhadap data uji dan mampu membedakan antar kelas secara efektif.

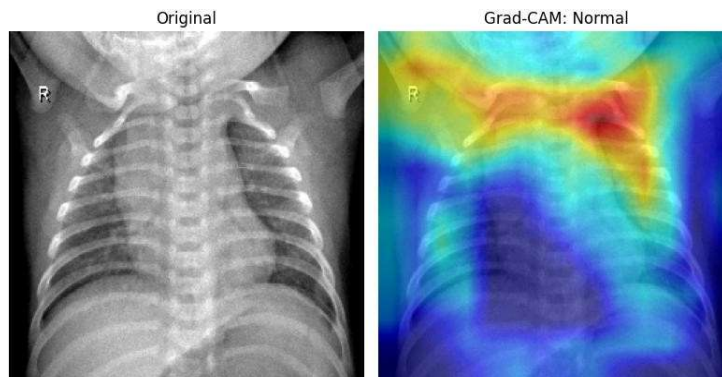
3.4 Visualisasi Grad-CAM

Penelitian ini juga melakukan interpretasi visual terhadap model yang telah diuji sebelumnya menggunakan metode Grad-CAM. Gambar yang dihasilkan memberikan pemahaman lebih mendalam mengenai fokus atensi model terhadap area-area penting pada citra CXR. Hal ini membantu mengevaluasi apakah model membuat keputusan berdasarkan fitur yang relevan secara medis atau tidak.

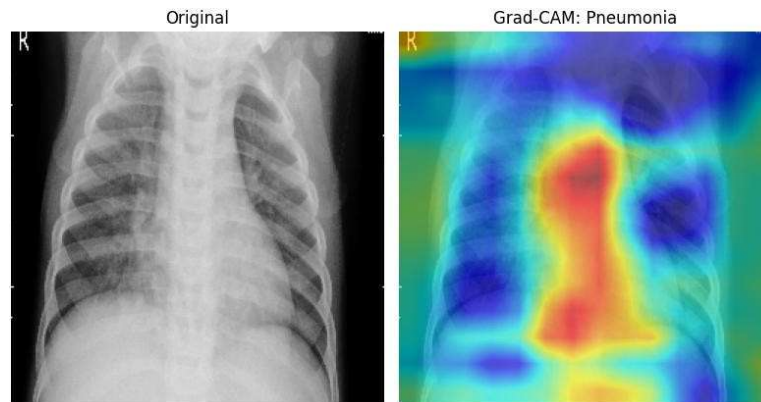
Berikut hasil interpretasi visual menggunakan Grad-CAM untuk setiap kategori:



Gambar 5. Visualisasi Grad-CAM pada pasien Covid-19



Gambar 6. Visualisasi paru-paru normal



Gambar 7. Visualisasi Grad-CAM pada pasien Pneumonia

Berdasarkan hasil di atas, Grad-CAM dapat menginterpretasikan keputusan model yang dikembangkan dengan memberikan gambaran visual mengenai area penting pada citra yang menjadi fokus perhatian model saat melakukan klasifikasi. Interpretasi visual ini membantu dalam memahami alasan di balik prediksi model serta meningkatkan kepercayaan terhadap performa model dalam mengidentifikasi kategori yang tepat.

4. KESIMPULAN

Penelitian ini mengusulkan model klasifikasi citra *Chest X-ray* berbasis *deep learning* yang menunjukkan kemampuan awal dalam membedakan antara kategori Covid-19, Normal, dan Pneumonia berdasarkan data yang tersedia. Proses pelatihan model dilakukan dengan eksplorasi kombinasi *hyperparameter*, dan konfigurasi terbaik diperoleh pada *learning rate* 0.001, 50 *epoch*, dan *batch size* 32. Model tersebut menghasilkan akurasi tinggi sebesar 97%, dengan nilai *precision*, *recall*, dan *f1-score* yang seimbang pada setiap kelas. Evaluasi performa melalui *confusion matrix* menunjukkan proporsi prediksi benar dan kesalahan pada masing-masing kelas, sementara visualisasi Grad-CAM memberikan interpretasi visual atas area yang menjadi fokus model dalam pengambilan keputusan. Hal ini memperkuat keyakinan terhadap kemampuan dan transparansi sistem yang dibangun. Secara keseluruhan, pendekatan ini menunjukkan bahwa klasifikasi berbasis CNN terhadap citra CXR dapat menjadi alat bantu diagnosis yang cepat dan efektif, serta berpotensi diterapkan dalam sistem pendukung keputusan medis. Penelitian ini belum melibatkan validasi dari pakar medis secara langsung. Oleh karena itu, disarankan pada penelitian selanjutnya untuk melakukan konfirmasi hasil prediksi model kepada tenaga ahli seperti pulmonolog atau radiolog guna memastikan akurasi dan relevansi klinis dari sistem yang dikembangkan.

REFERENSI

- [1] W. W. Timur, F. B. Sabiti, and N. L. Amajida, "HUBUNGAN POLA TERAPI DENGAN KUALITAS HIDUP PASIEN TB PARU DI BALAI KESEHATAN MASYARAKAT WILAYAH PATI," *Jurnal Farmasi & Sains Indonesia*, vol. 6, no. 2, pp. 126–130, Jan. 2024, doi: 10.52216/jfsi.vol6no2p126-130.
- [2] S. Prasad Koyyada and T. P. Singh, "An explainable artificial intelligence model for identifying local indicators and detecting lung disease from chest X-ray images," *Healthcare Analytics*, vol. 4, p. 100206, Dec. 2023, doi: 10.1016/j.health.2023.100206.
- [3] I. Benlala, F. Laurent, and G. Dournes, "Structural and functional changes in <sc>COPD</sc>: What we have learned from imaging," *Respirology*, vol. 26, no. 8, pp. 731–741, Aug. 2021, doi: 10.1111/resp.14047.
- [4] R. Mulyadi, R. Rahmawati, E. Arief, E. Syahril, and B. Natsir, "GAMBARAN RADIOLOGI FOTO TORAKS PADA PASIEN RAWAT INAP YANG TERDIAGNOSIS PNEUMONIA KOMUNITAS," *PREPOTIF : JURNAL KESEHATAN MASYARAKAT*, vol. 8, no. 2, pp. 2916–2928, Jul. 2024, doi: 10.31004/prepotif.v8i2.29570.
- [5] D. Cozzi *et al.*, "Ground-glass opacity (GGO): a review of the differential diagnosis in the era of COVID-19," *Jpn J Radiol*, vol. 39, no. 8, pp. 721–732, Aug. 2021, doi: 10.1007/s11604-021-01120-w.
- [6] M. Jafari-Oori *et al.*, "Acute Respiratory Distress Syndrome and COVID-19: A Scoping Review and Meta-analysis," 2021, pp. 211–228. doi: 10.1007/978-3-030-59261-5_18.
- [7] R. Klangbunrueng, P. Pookduang, W. Chansanam, and T. Lunrasri, "AI-Powered Lung Cancer Detection: Assessing VGG16 and CNN Architectures for CT Scan Image Classification," *Informatics*, vol. 12, no. 1, p. 18, Feb. 2025, doi: 10.3390/informatics12010018.
- [8] S. Kumar *et al.*, "<sc>LiteCovidNet</sc>: A lightweight deep neural network model for detection of <sc>COVID</sc> - 19 using X-ray images," *Int J Imaging Syst Technol*, vol. 32, no. 5, pp. 1464–1480, Sep. 2022, doi: 10.1002/ima.22770.
- [9] N. Nurkhasanah and M. Murinto, "Klasifikasi Penyakit Kulit Wajah Menggunakan Metode Convolutional Neural Network," *Sainteks*, vol. 18, no. 2, p. 183, Feb. 2022, doi: 10.30595/sainteks.v18i2.13188.

- [10] K. J. Aditya, I. Kanedi, and A. Sudarsono, "SEGMENTASI DETEKSI TEPI PADA CITRA DIGITAL PATAH TULANG ORANG DEWASA MENGGUNAKAN METODE SOBEL DAN METODE PREWITT," *Djtechno: Jurnal Teknologi Informatika*, vol. 3, no. 2, pp. 224–233, Dec. 2022, doi: 10.46576/djtechno.v3i2.2735.
- [11] L. Stanković and D. Mandić, "Convolutional Neural Networks Demystified: A Matched Filtering Perspective-Based Tutorial," *IEEE Trans Syst Man Cybern Syst*, vol. 53, no. 6, pp. 3614–3628, Jun. 2023, doi: 10.1109/TSMC.2022.3228597.
- [12] G. Kourounis, A. A. Elmahmudi, B. Thomson, J. Hunter, H. Ugail, and C. Wilson, "Computer image analysis with artificial intelligence: a practical introduction to convolutional neural networks for medical professionals," *Postgrad Med J*, vol. 99, no. 1178, pp. 1287–1294, Nov. 2023, doi: 10.1093/postmj/qgad095.
- [13] A. Bherje *et al.*, "Design of Deep Learning-based Approach to Predict Lung Cancer on CT Scan Images," in *2024 5th International Conference on Innovative Trends in Information Technology (ICITIIT)*, IEEE, Mar. 2024, pp. 1–5. doi: 10.1109/ICITIIT61487.2024.10580370.
- [14] H. Huang, M. Wang, Q. Ye, and Z. Zhou, "Diagnosis of Lung Cancer Based on CT Scans Using Convolutional Neural Networks," in *2022 International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACAI)*, IEEE, Aug. 2022, pp. 338–341. doi: 10.1109/ICDACAI57211.2022.00073.
- [15] C. Tejaswini, P. Nagabushanam, P. Rajasegaran, P. R. Johnson, and S. Radha, "CNN Architecture for Lung Cancer Detection," in *2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT)*, IEEE, Apr. 2022, pp. 346–350. doi: 10.1109/CSNT54456.2022.9787650.
- [16] S. A. E. ALBAKIA and R. A. Saputra, "Identifikasi Jenis Daun Tanaman Obat Menggunakan Metode Convolutional Neural Network (CNN) Dengan Model VGG16," *Jurnal Informatika Polinema*, vol. 9, no. 4, pp. 451–460, Aug. 2023, doi: 10.33795/jip.v9i4.1420.
- [17] D. R. Sarvamangala and R. V. Kulkarni, "Convolutional neural networks in medical image understanding: a survey," *Evol Intell*, vol. 15, no. 1, pp. 1–22, Mar. 2022, doi: 10.1007/s12065-020-00540-3.
- [18] S. Tammina, "Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images," *International Journal of Scientific and Research Publications (IJSRP)*, vol. 9, no. 10, p. p9420, Oct. 2019, doi: 10.29322/IJSRP.9.10.2019.p9420.
- [19] S. Chaudhury and T. Yamasaki, "Robustness of Adaptive Neural Network Optimization Under Training Noise," *IEEE Access*, vol. 9, pp. 37039–37053, 2021, doi: 10.1109/ACCESS.2021.3062990.
- [20] M. Sah and C. Direkoglu, "A survey of deep learning methods for multiple sclerosis identification using brain MRI images," *Neural Comput Appl*, vol. 34, no. 10, pp. 7349–7373, May 2022, doi: 10.1007/s00521-022-07099-3.
- [21] N. Subaşı, "Comprehensive Analysis of Grid and Randomized Search on Dataset Performance," *European Journal of Engineering and Applied Sciences*, vol. 7, no. 2, pp. 77–83, Dec. 2024, doi: 10.55581/ejeas.1581494.
- [22] B. H. M. van der Velden, H. J. Kuijff, K. G. A. Gilhuijs, and M. A. Viergever, "Explainable artificial intelligence (XAI) in deep learning-based medical image analysis," *Med Image Anal*, vol. 79, p. 102470, Jul. 2022, doi: 10.1016/j.media.2022.102470.
- [23] M. Bhandari, T. B. Shahi, B. Siku, and A. Neupane, "Explanatory classification of CXR images into COVID-19, Pneumonia and Tuberculosis using deep learning and XAI," *Comput Biol Med*, vol. 150, p. 106156, Nov. 2022, doi: 10.1016/j.compbiomed.2022.106156.
- [24] Y. Shen and X. Huang, "A Comparative Visualization Analysis of Neural Network Models Using Grad-CAM," *Science and Technology of Engineering, Chemistry and Environmental Protection*, vol. 1, no. 10, Dec. 2024, doi: 10.61173/yzp9wt79.
- [25] Md. Z. Hasan *et al.*, "Fast and Efficient Lung Abnormality Identification With Explainable AI: A Comprehensive Framework for Chest CT Scan and X-Ray Images," *IEEE Access*, vol. 12, pp. 31117–31135, 2024, doi: 10.1109/ACCESS.2024.3369900.
- [26] J.-C. Chien, J.-D. Lee, C.-S. Hu, and C.-T. Wu, "The Usefulness of Gradient-Weighted CAM in Assisting Medical Diagnoses," *Applied Sciences*, vol. 12, no. 15, p. 7748, Aug. 2022, doi: 10.3390/app12157748.
- [27] Y. Zhang, D. Hong, D. McClement, O. Oladosu, G. Pridham, and G. Slaney, "Grad-CAM helps interpret the deep learning models trained to classify multiple sclerosis types using clinical brain magnetic resonance imaging," *J Neurosci Methods*, vol. 353, p. 109098, Apr. 2021, doi: 10.1016/j.jneumeth.2021.109098.