

Revolusi Diagnostik Pneumonia: Tinjauan Sistematis Aplikasi *Deep Learning* pada Citra Radiografi Dada

 **Novalia Nurbaiti^{1*}**,  **Novia Fitri Azarah,**

^{1,2} Institut Teknologi dan Sains Nahdlatul Ulama Lampung, Indonesia

 novalianurbaiti@gmail.com^{*}, noviafitriazarah@gmail.com

Article Information:

Received April 18, 2025

Revised May 15, 2025

Accepted Juni 26, 2025

Keywords: Pneumonia,
Deep Learning,
Convolutional Neural
Network (CNN),
Chest Radiography,
Medical Diagnostics,
Systematic Literature Review.

Abstract

Pneumonia remains a leading cause of global morbidity and mortality, particularly in vulnerable populations such as children and the elderly. Early and accurate diagnosis is crucial, but is often hampered by the scarcity of trained radiologists and the inherent subjectivity in chest X-ray (CXR) image interpretation. In recent years, the rapid advancement of deep learning (DL), especially Convolutional Neural Networks (CNNs), has presented transformative solutions to these challenges. This systematic literature review (SLR) thoroughly analyzes twelve primary studies exploring the application of DL models for pneumonia detection from CXR images. Through a rigorous SLR methodology, these studies were identified, evaluated, and synthesized. Results indicate that DL models, whether built from scratch or based on transfer learning, are capable of achieving very high diagnostic accuracy, often surpassing or matching the performance of human radiologists. The discussion highlights key findings, addresses the formulated research questions, and identifies existing study limitations. The conclusion affirms the significant potential of DL in revolutionizing pneumonia diagnostics, and outlines future research directions to enhance the generalization, interpretability, and implementation of DL models in daily clinical practice.

PENDAHULUAN

Pneumonia merupakan masalah kesehatan global yang serius, menjadi penyebab utama morbiditas dan mortalitas, khususnya pada kelompok rentan seperti anak-anak dan lansia. Diagnosis yang cepat dan akurat sangatlah penting untuk penanganan dan pengobatan yang efektif. Meskipun berbagai teknik diagnostik tersedia, radiografi dada (CXR) adalah metode yang paling umum dan standar karena efektivitas biaya dan ketersediaannya yang luas, bahkan di fasilitas medis dengan sumber daya terbatas. Namun, interpretasi citra CXR untuk diagnosis pneumonia memiliki tantangan inheren. Prosesnya sangat bergantung pada keahlian dan pengalaman radiolog, yang dapat menyebabkan subjektivitas dan ketidaksepakatan antar-radiolog (Gabruseva et al., 2020; Szepesi & Szilágyi, 2022). Kelangkaan radiolog terlatih, terutama di negara berkembang dan daerah terpencil, semakin memperparah masalah ini, dengan lebih dari dua pertiga populasi dunia tidak memiliki akses yang memadai ke layanan radiolog (Omarov & Bazarkulova, 2023).

Dalam beberapa tahun terakhir, kemajuan pesat dalam *deep learning* (DL), khususnya *Convolutional Neural Networks* (CNNs), telah menawarkan solusi transformatif untuk tantangan ini. CNNs, sebagai arsitektur DL yang terinspirasi dari sistem visual biologis, telah menunjukkan kemampuan luar biasa dalam mengekstraksi fitur hierarkis secara otomatis dari citra, menjadikannya sangat cocok untuk tugas klasifikasi citra medis (Li et al., 2022). Sejumlah besar penelitian telah menginvestigasi penerapan CNN untuk mendeteksi pneumonia dari citra CXR, dengan hasil yang seringkali melampaui atau setara dengan kinerja radiolog manusia (Omarov & Bazarkulova, 2023; Sharma & Guleria, 2023; Abid et al., 2024). Tinjauan literatur sistematis (SLR) ini bertujuan untuk menganalisis secara mendalam aplikasi model DL, khususnya CNN, dalam deteksi pneumonia dari citra radiografi dada. Dengan menganalisis metodologi, hasil, dan temuan utama dari studi-studi primer yang relevan, kami berharap dapat memberikan gambaran komprehensif tentang kemajuan terkini di bidang ini, mengidentifikasi tren yang muncul, serta menyoroti tantangan dan peluang penelitian di masa depan.

TINJAUAN PUSTAKA

Pneumonia, sebagai infeksi pernapasan akut yang menyerang paru-paru, terus menjadi ancaman serius bagi kesehatan masyarakat global. Diperkirakan 5–7 juta anak-anak meninggal akibat pneumonia setiap tahunnya, dan sekitar 7% populasi dunia terpengaruh oleh penyakit ini, dengan lebih dari 4 juta kematian per tahun secara global akibat penyakit pernapasan yang berhubungan dengan polusi udara, termasuk pneumonia (Sourab & Kabir, 2022; Omarov & Bazarkulova, 2023; Abid et al., 2024). Deteksi dini dan diagnosis yang akurat merupakan prasyarat esensial untuk manajemen dan pengobatan pneumonia yang efektif.

Meskipun berbagai teknik diagnostik tersedia, seperti X-ray dada, MRI, dan CT scan, radiografi dada (CXR) adalah metode yang paling umum dan standar yang digunakan di seluruh dunia. Ini karena CXR menawarkan keseimbangan biaya-efektivitas dan ketersediaan luas, bahkan di fasilitas medis dengan sumber daya terbatas (Szepesi & Szilágyi, 2022). Namun, diagnosis pneumonia dari citra CXR memiliki tantangan inheren. Proses interpretasi sangat bergantung pada keahlian dan pengalaman radiolog, yang dapat menimbulkan variasi subjektif dan ketidaksepakatan antar-radiolog (Gabruseva et al., 2020; Szepesi & Szilágyi, 2022). Lebih lanjut, kelangkaan radiolog terlatih di banyak negara berkembang dan daerah terpencil memperparah masalah ini, di mana lebih dari dua pertiga populasi dunia tidak memiliki akses yang memadai ke layanan radiolog (Omarov & Bazarkulova, 2023).

Convolutional Neural Networks (CNNs), sebagai arsitektur DL yang terinspirasi dari sistem visual biologis, telah menunjukkan kemampuan luar biasa dalam secara otomatis mengekstraksi fitur hierarkis dari citra, menjadikannya sangat cocok untuk tugas klasifikasi citra medis (Li et al., 2022). Sejumlah besar penelitian telah menginvestigasi penerapan CNN untuk mendeteksi pneumonia dari citra CXR, dengan hasil yang seringkali melampaui atau setara dengan kinerja radiolog manusia (Omarov & Bazarkulova, 2023; Sharma & Guleria, 2023; Abid et al., 2024).

Szepesi dan Szilágyi (2022) memperkenalkan CNN dengan *dropout* pada lapisan konvolusional, mencapai akurasi 97.2%. Stephen et al. (2019) merancang model CNN dari awal untuk klasifikasi pneumonia dengan akurasi validasi 93.73%, menunjukkan efektivitasnya bahkan dengan *dataset* terbatas melalui augmentasi data. Omarov dan Bazarkulova (2023) mengusulkan model sekuensial berbasis CNN dengan akurasi 92%. Urey et al. (2019) mencapai 78.73% akurasi dengan arsitektur CNN dan Residual Network yang lebih efisien (hanya 3 lapisan konvolusional) setelah *preprocessing* citra ekstensif. Li et al. (2022) membangun model CNN gaya VGG yang mencapai 97% akurasi. Abid et al. (2024) mengusulkan CNN tujuh lapis yang mencakup lapisan konvolusional, normalisasi *batch*, *max-pooling dropout*, dan lapisan *dense*.

Banyak studi juga memanfaatkan arsitektur CNN pra-terlatih melalui pendekatan *transfer learning*. Model-model ini seringkali digunakan sebagai ekstraktor fitur, diikuti oleh lapisan *dense* atau pengklasifikasi *machine learning* tradisional. Qaimkhani et al. (2022) menemukan VGG19 memberikan akurasi tertinggi (98.27%) di antara ANN, CNN kustom, dan VGG19. Sharma dan

Guleria (2023) menunjukkan VGG16 dengan *Neural Networks* (NN) mencapai 95.4% akurasi. Yang dan Jin (<https://www.google.com/search?q=2022>) menggunakan DenseNet-121 pra-terlatih dengan akurasi sekitar 70.5% untuk klasifikasi 8 penyakit toraks. Vetrithangam et al. (2023) mengusulkan model *deep learning* ResNet152v2 yang dimodifikasi, mencapai akurasi 99.77%.

Teknik *preprocessing* dan augmentasi data merupakan aspek krusial yang digunakan di hampir semua studi untuk memperkaya *dataset* dan mencegah *overfitting*. Teknik umum meliputi rotasi, *zooming*, *shifting* (lebar/tinggi), dan *horizontal flipping* (Omarov & Bazarkulova, 2023; Stephen et al., 2019; Sourab & Kabir, 2022). Gabruseva et al. (2020) bahkan bereksperimen dengan berbagai tingkat augmentasi (ringan, berat, rotasi kustom) untuk optimalisasi. Pra-pemrosesan citra seperti standarisasi ukuran citra (misalnya, 224x224 atau 150x150 piksel), normalisasi nilai piksel (seringkali membagi dengan 255), dan konversi format (misalnya, RGB ke *grayscale*) adalah langkah-langkah umum (Sourab & Kabir, 2022; Stephen et al., 2019). Urey et al. (2019) secara spesifik menyelidiki dampak peningkatan kontras, perluasan ruang warna, dan pencerahan buatan pada citra untuk memperjelas fitur. Vetrithangam et al. (2023) juga menekankan langkah-langkah *preprocessing* yang diperlukan untuk meningkatkan kualitas *dataset*.

Model-model DL secara konsisten menunjukkan akurasi yang sangat tinggi. Sourab dan Kabir (2022) melaporkan akurasi *hybrid* CNN-RF tertinggi 99.52%. Szepesi dan Szilágyi (2022) melaporkan akurasi 97.2%. Li et al. (2022) mencapai akurasi 97%. Sharma dan Guleria () menunjukkan VGG16+NN mencapai 95.4% pada *dataset* yang lebih besar. Vetrithangam et al. (2023) mencapai akurasi 99.77% dengan model ResNet152v2 yang dimodifikasi. Abid et al. (2024) melaporkan akurasi hingga 96%. Selain akurasi, metrik seperti *precision*, *recall* (sensitivitas), F1-score, dan *Area Under the Curve* (AUC) juga secara luas digunakan.

Recall sering ditekankan karena kepentingannya dalam meminimalkan *false negative* di diagnostik medis. Omarov dan Bazarkulova (2023) mencapai *recall* 96% untuk pneumonia. Beberapa studi juga menunjukkan keunggulan model hibrida. Sourab dan Kabir (2022) menunjukkan keunggulan model hibrida CNN sebagai ekstraktor fitur yang dikombinasikan dengan pengklasifikasi *machine learning* tradisional seperti *Random Forest* (RF), *Support Vector Machine* (SVM), dan *K-Nearest Neighbor* (KNN). Hibrida CNN-RF secara konsisten memberikan kinerja terbaik. Beberapa studi memperluas deteksi pneumonia ke klasifikasi multikelas, termasuk identifikasi COVID-19 bersama pneumonia dan normal (Sharma & Guleria, 2023), atau klasifikasi 8 hingga 14 penyakit toraks berbeda (Yang & Jin, 2022; Gabruseva et al., 2020).

Tinjauan literatur sistematis (SLR) ini bertujuan untuk mengidentifikasi, menganalisis, dan mensintesis bukti ilmiah mengenai aplikasi model *deep learning*, khususnya CNN, untuk deteksi pneumonia dari citra radiografi dada. Dengan menganalisis metodologi, hasil, dan temuan utama dari studi-studi primer yang relevan, kami berharap dapat memberikan gambaran komprehensif tentang kemajuan terkini di bidang ini, mengidentifikasi tren yang muncul, serta menyoroti tantangan dan peluang penelitian di masa depan.

METODOLOGI

Metodologi tinjauan literatur sistematis (SLR) ini dirancang untuk memastikan proses identifikasi, seleksi, dan sintesis bukti ilmiah yang transparan, replikabel, dan komprehensif terkait aplikasi *deep learning* (DL) dalam deteksi pneumonia dari citra radiografi dada (CXR). Pendekatan ini mengikuti prinsip-prinsip yang umumnya diatur dalam pedoman pelaporan tinjauan sistematis.

2.1 Rumusan Pertanyaan Penelitian (RQ)

Tinjauan ini dipandu oleh tiga pertanyaan penelitian (RQ) yang spesifik dan terstruktur, bertujuan untuk menginvestigasi peran DL dalam diagnostik pneumonia:

- a) **RQ1:** Bagaimana model *deep learning*, khususnya *Convolutional Neural Networks* (CNN), diaplikasikan untuk deteksi pneumonia dari citra radiografi dada?
- b) **RQ2:** Seberapa efektif model *deep learning* dalam mendiagnosis pneumonia dari citra

radiografi dada, dibandingkan dengan metode konvensional atau model *machine learning* lainnya?

- c) **RQ3:** Apa saja tantangan dan peluang dalam implementasi model *deep learning* untuk deteksi pneumonia dalam praktik klinis?

2.2 Protokol dan Strategi Pencarian

Pengembangan protokol pencarian dilakukan untuk mengidentifikasi studi primer yang relevan. Strategi pencarian dirumuskan menggunakan kombinasi kata kunci yang luas dan spesifik, meliputi: "*pneumonia detection*", "*chest X-ray*", "*deep learning*", "*convolutional neural network*" (CNN), "*medical imaging*", "*diagnosis*", dan "*AI in healthcare*". Kata kunci ini digabungkan menggunakan operator Boolean seperti "AND" dan "OR" untuk memperluas dan menyaring hasil. Pencarian dilakukan di beberapa basis data ilmiah terkemuka, termasuk PubMed, Scopus, Web of Science, dan IEEE Xplore. Meskipun tidak ada batasan tanggal publikasi yang ketat diberlakukan untuk menangkap evolusi bidang ini, prioritas diberikan pada studi-studi terbaru untuk merefleksikan kemajuan terkini dalam teknologi DL dan aplikasinya.

2.2 Kriteria Inklusi/Eksklusi

Studi-studi yang teridentifikasi dievaluasi berdasarkan kriteria inklusi dan eksklusi yang telah ditetapkan secara eksplisit:

2.2.1 Kriteria Inklusi:

- Artikel jurnal yang dipublikasikan dalam bahasa Inggris.
- Studi yang berfokus pada deteksi atau klasifikasi pneumonia menggunakan citra X-ray dada.
- Studi yang mengimplementasikan atau mengevaluasi model *deep learning* (terutama CNN) untuk tujuan deteksi pneumonia.
- Artikel yang telah melalui proses tinjauan sejawat (*peer-reviewed*).
- Studi yang melaporkan metrik kinerja kuantitatif yang jelas (misalnya, akurasi, presisi, *recall*, F1-score, Area Under the Curve/AUC).

2.2.2 Kriteria Eksklusi:

- Artikel yang tidak relevan secara langsung dengan topik inti tinjauan ini (misalnya, berfokus pada deteksi penyakit paru-paru lain selain pneumonia, atau menggunakan modalitas pencitraan selain X-ray dada sebagai sumber data utama).
- Jenis publikasi non-primer seperti tinjauan literatur, editorial, opini, abstrak konferensi tanpa naskah lengkap, atau bab buku (kecuali jika menyajikan data primer yang unik dan relevan).
- Studi yang hanya menggunakan metode *machine learning* tradisional tanpa komponen *deep learning*.

Data yang relevan diekstraksi dari setiap studi yang memenuhi kriteria inklusi menggunakan formulir ekstraksi data terstruktur yang telah dirancang sebelumnya. Informasi yang dikumpulkan meliputi: detail bibliografi studi (penulis, tahun publikasi, judul jurnal), jenis model DL yang diimplementasikan atau dievaluasi, arsitektur jaringan yang spesifik (misalnya, jumlah lapisan, jenis lapisan, penggunaan *transfer learning*), karakteristik *dataset* yang digunakan (ukuran, sumber, deskripsi populasi), metode *preprocessing* dan augmentasi data yang diterapkan, metrik kinerja diagnostik yang diperoleh, serta temuan-temuan kunci dan kesimpulan yang ditarik oleh penulis studi primer. Proses ini bertujuan untuk mengumpulkan informasi yang seragam dan relevan untuk sintesis data yang efektif.

HASIL DAN PEMBAHASAN

Berdasarkan proses seleksi yang ketat, dua belas studi primer yang relevan berhasil diidentifikasi dan dianalisis secara mendalam dalam SLR ini. Dua belas jurnal yang dianalisis ini dipublikasikan antara tahun 2019 dan 2024, menegaskan bahwa aplikasi DL untuk deteksi pneumonia

dari CXR merupakan bidang penelitian yang sangat aktif dan berkembang pesat. Mayoritas studi memanfaatkan *dataset* citra X-ray dada pediatrik dari Guangzhou Women and Children's Medical Center, yang seringkali tersedia melalui platform Kaggle atau Mendeley Data v2. Beberapa studi lain menggunakan *dataset* yang lebih besar seperti NIH ChestX-ray14 atau *dataset* gabungan (Yang & Jin, 2022; Gabruseva et al., 2020). Ukuran *dataset* bervariasi secara signifikan, dari beberapa ribu hingga lebih dari seratus ribu citra.

Tabel 1. Rangkuman Data Statistik Utama

Referensi (Tahun)	Dataset Utama (Ukuran & Sumber)	Tipe Arsitektur DL Utama	Metrik Kinerja Utama (Akurasi/AUC)	Catatan Khusus Metode/Hasil
Vetrithangam et al. (2023)	1485 citra X-ray (RSNA Pneumonia Detection Challenge)	Modified ResNet152v2	Akurasi: 99.77% , Sensitivitas: 99.86%, Spesifisitas: 95.4%, Presisi: 99.86%	Model yang dimodifikasi untuk mengurangi kompleksitas komputasi dan waktu pelatihan; kinerja tinggi pada
Sourab & Kabir (2022)	5856 citra X-ray (Mendeley Data v2)	Hybrid CNN-RF (22 layers CNN + Random Forest)	Akurasi: 99.52% , AUC: 98.7%	Model hibrida CNN-RF menunjukkan kinerja terbaik dibandingkan CNN-KNN dan CNN-SVM.
Qaimkhani et al. (2022)		Dataset Kaggle (Guangzhou Women and Children's Medical Center)	VGG19 (Transfer Learning)	Akurasi: 98.27%
Szepesi & Szilágyi (2022)	5856 citra berlabel (Kaggle: Guangzhou Women and Children's Medical Center)	CNN Kustom dengan Dropout di Lapisan Konvolusional	Akurasi: 97.2% , Recall: 97.3%, Presisi: 97.4%, AUC: 0.982	Model dibangun dari awal, mengungguli VGG-19, InceptionV3, ResNet50 dalam akurasi dan efisiensi.
Li et al. (2022)	5863 citra JPEG (Guangzhou Women and Children's Medical Center)	VGG-style CNN Kustom	Akurasi: 97% , Presisi Pneumonia: 98%, Recall Pneumonia: 97%	Menggunakan Adamax optimizer
Abid et al. (2024)	5863 citra X-ray (Kaggle)	CNN Tujuh Lapis Kustom	Akurasi: Hingga 96%	Fokus pada peningkatan efisiensi dan akurasi diagnosis, mengatasi masalah kualitas citra dan interpretasi.
Sharma & Guleria (2023)	Dataset 1: 5856 CXR (Kaggle);	Dataset 2: 6436 CXR (Normal, Pneumonia, Covid-19 dari Kaggle/GitHub)	VGG16 + Neural Networks (NN)	Akurasi: 92.15% (DS1), 95.4% (DS2)
Omarov & Bazarkulova (2023)	5863 citra X-ray (Guangzhou Women's and	Model Sekuensial Sederhana	Akurasi: 92% , Recall: 96% , Presisi: 92% , F1-score: 94%	Dilatih dalam 26 menit selama 12

	Children's Medical Center via Kaggle)	(CNN 22 layers)		
Stephen et al. (2019)	5856 citra X-ray (Pediatrik)	CNN Kustom (Dibangun dari Awal)	Akurasi Validasi: 93.73% (untuk ukuran 200x200x3)	Menerapkan augmentasi data ekstensif untuk mengatasi
Gabruseva et al. (2020)	26684 citra X-ray (RSNA Pneumonia Detection Challenge)	SSD RetinaNet dengan SE-ResNext101 encoder	mAP: 0.24781 (Peringkat ke-2 dalam kompetisi)	Fokus pada deteksi wilayah
Urey et al. (2019)	Sekitar 4000 citra X-ray (Kaggle)	CNN dengan Residual Network (3 lapisan konvolusional)	Akurasi: 78.73%	Menggunakan langkah
Yang & Jin (2022)	112120 citra X-ray (NIH ChestX-ray14)	DenseNet Classifier 121-layer (Pra-terlatih)	Akurasi: ~70.5% (untuk klasifikasi 8 penyakit toraks)	Mampu melakukan klasifikasi dan

3.1 Analisis Tematik

Analisis tematik mengidentifikasi beberapa kategori utama dalam pendekatan dan temuan studi yang ditinjau:

3.1.1 Arsitektur CNN: Kustom vs. *Transfer Learning*

- Arsitektur Kustom (*From Scratch*):** Beberapa peneliti mengusulkan arsitektur CNN yang dibangun dari awal. Szepesi dan Szilágyi (2022) memperkenalkan CNN dengan *dropout* pada lapisan konvolusional, mencapai akurasi 97.2%. Stephen et al. (2019) merancang model CNN dari awal untuk klasifikasi pneumonia dengan akurasi validasi 93.73%, menunjukkan efektivitasnya bahkan dengan *dataset* terbatas melalui augmentasi data. Omarov dan Bazarkulova (2023) mengusulkan model sekuensial berbasis CNN dengan akurasi 92%. Urey et al. (2019) mencapai 78.73% akurasi dengan arsitektur CNN dan Residual Network yang lebih efisien (hanya 3 lapisan konvolusional) setelah *preprocessing* citra ekstensif. Li et al. (2022) membangun model CNN gaya VGG yang mencapai 97% akurasi. Abid et al. (2024) mengusulkan CNN tujuh lapis yang mencakup lapisan konvolusional, normalisasi *batch*, *max-pooling dropout*, dan lapisan *dense*.
- Pendekatan *Transfer Learning*:** Banyak studi memanfaatkan arsitektur CNN pra-terlatih seperti VGG19, ResNet50, InceptionV3, DenseNet-121, AlexNet, dan GoogLeNet. Model-model ini seringkali digunakan sebagai ekstraktor fitur, diikuti oleh lapisan *dense* atau pengklasifikasi *machine learning* tradisional. Qaimkhani et al (2022) menemukan VGG19 memberikan akurasi tertinggi (98.27%) di antara ANN, CNN kustom, dan VGG19. Sharma dan Guleria (2023) menunjukkan VGG16 dengan Neural Networks (NN) mencapai 95.4% akurasi. Yang dan Jin (2022) menggunakan DenseNet-121 pra-terlatih dengan akurasi sekitar 70.5% untuk klasifikasi 8 penyakit toraks. Vetrithangam et al. (2023) mengusulkan model *deep learning* ResNet152v2 yang dimodifikasi, mencapai akurasi 99.77%.

3.2 Metrik Kinerja dan Perbandingan:

- Akurasi Tinggi:** Model-model DL secara konsisten menunjukkan akurasi yang sangat tinggi. Sourab dan Kabir (2022) melaporkan akurasi *hybrid* CNN-RF tertinggi 99.52%. Szepesi dan Szilágyi (2022) melaporkan akurasi 97.2%. Li et al. (2022) mencapai akurasi 97%. Sharma dan Guleria (2023) menunjukkan VGG16+NN mencapai 95.4% pada *dataset* yang lebih besar.

Vetrithangam et al. (2023) mencapai akurasi 99.77% dengan model ResNet152v2 yang dimodifikasi. Abid et al. (2024) melaporkan akurasi hingga 96%.

- b) **Metrik Lain:** Selain akurasi, metrik seperti *precision*, *recall* (sensitivitas), F1-score, dan Area Under the Curve (AUC) juga secara luas digunakan. *Recall* sering ditekankan karena kepentingannya dalam meminimalkan *false negative* di diagnostik medis. Omarov dan Bazarkulova (2023) mencapai *recall* 96% untuk pneumonia.

Tabel 2. Analisis Kompetitor dengan akurasi, kekuatan, dan kelemahan

Referensi (Tahun)	Model yang Diusulkan/Dievaluasi	Akurasi (Tertinggi)	Kekuatan (Strengths)	Kelemahan (Weaknesses)
Vetrithangam et al. (2023)	Modified ResNet152v2	99.77%	Akurasi sangat tinggi , mengurangi kompleksitas komputasi dan waktu pelatihan , <i>highly interpretable model</i>	Keterbatasan pada <i>dataset</i> X-ray pediatrik (1485 gambar) , perlu validasi pada <i>dataset</i> yang lebih beragam.
Sourab & Kabir (2022)	Hybrid CNN-RF (22 layers CNN + Random Forest)	99.52%	Akurasi sangat tinggi , model hibrida unik , RF meningkatkan kinerja CNN secara signifikan	Tidak menunjukkan akurasi tinggi yang sama untuk semua <i>dataset</i> radiogram dada pneumonia yang ada
Qaimkhani et al. (2022)	VGG19 (Transfer Learning)	98.27%	Akurasi sangat tinggi , VGG19 mengungguli ANN dan CNN kustom , fitur yang dipelajari dari <i>large-scale dataset</i> bermanfaat.	Penggunaan <i>dataset</i> terbatas (5863 gambar) , perlu diuji pada data yang lebih besar dan kompleks di masa depan.
Szepesi & Szilágyi (2022)	Novel CNN with Dropout in Convolutional Part	97.2%	Akurasi tinggi , konvergensi lebih cepat , stabil , lebih efisien dibanding model <i>transfer learning</i> yang lebih besar.	<i>dataset</i> yang digunakan mungkin tidak representatif secara klinis.
Li et al. (2022)	VGG-style CNN Custom Model	97%	Akurasi tinggi , presisi tinggi untuk pneumonia (98%) , mampu membedakan pola unik pneumonia , mengatasi variasi kualitas gambar.	Tidak menyertakan lapisan <i>dropout</i> di kode awal , perluasan <i>dataset</i> dan uji klinis lanjutan diperlukan.
Abid et al. (2024)	Seven-layer CNN Custom Model	Hingga 96%	Akurasi tinggi , mengatasi keterbatasan model	Kurangnya cakupan untuk pasien dari berbagai latar

			<i>deep learning</i> tradisional dalam variabilitas penyakit dan interpretasi.	belakang dan fasilitas kesehatan , keterbatasan penggunaan modalitas pencitraan lain atau <i>dataset</i> klinis tambahan , masalah interpretasi hasil.
Sharma & Guleria (2023)	VGG16 + Neural Networks (NN)	95.4%	Akurasi tinggi pada dua <i>dataset</i> (termasuk multikelas dengan COVID-19) , mengungguli kombinasi VGG16 dengan SVM, KNN, RF, NB.	Penggunaan <i>dataset</i> terbatas, bisa ditingkatkan dengan augmentasi.
Omarov & Bazarkulova (2023)	Simple Sequential Model (CNN 22 layers)	92.3%	Mengurangi beban radiolog dan menghemat waktu , akurasi dan <i>recall</i> tinggi (96%) , model sederhana dengan daya komputasi rendah.	<i>Dataset</i> tidak seimbang , terbatas pada satu institusi , kurangnya <i>robustness</i> dalam pendekatan.
Stephen et al. (2019)	CNN Model Built from Scratch	93.73% (Validation)	Mengatasi masalah <i>overfitting</i> dengan augmentasi data , model sederhana.	Keterbatasan kedalaman data , akurasi validasi menurun sedikit dengan ukuran citra yang lebih besar.
Gabruseva et al. (2020)	SSD RetinaNet with SE-ResNext101 encoder	Peringkat ke-2 di RSNA Pneumonia Detection Challenge (0.24781 mAP)	Deteksi wilayah pneumonia (<i>localization</i>) , menggunakan <i>multi-task learning</i> , menangani <i>empty boxes</i> , mengatasi perbedaan proses pelabelan <i>train</i> dan <i>test</i> .	<i>mAP score</i> relatif rendah dibandingkan akurasi absolut , tidak melibatkan augmentasi <i>test-time</i> untuk trade-off akurasi-sumber daya
Urey et al. (2019)	CNN with Residual Network (3 conv layers)	78.73%	Menggunakan <i>preprocessing</i> citra yang unik (kontras, pencerahan, ruang warna) untuk memperjelas fitur , akurasi lebih tinggi dari CheXNet dengan lapisan lebih sedikit , mengurangi waktu komputasi.	Akurasi relatif lebih rendah dibandingkan model-model terbaru , hanya berfokus pada Pneumonia (bukan 14 penyakit).
Yang & Jin (2022)	DenseNet Classifier (121 layers, pre-trained)	~70.5% (untuk 8 penyakit toraks)	Mampu melakukan klasifikasi dan <i>localization</i>	Akurasi lebih rendah untuk deteksi pneumonia

(heatmap) untuk secara spesifik.
berbagai penyakit
toraks ,
transfer learning
menghemat sumber
daya komputasi.

KESIMPULAN

Tinjauan ini secara meyakinkan menunjukkan bahwa *deep learning*, khususnya arsitektur CNN, memegang kunci untuk merevolusi diagnosis pneumonia dari citra X-ray dada. Model-model ini telah menunjukkan kemampuan luar biasa dalam mencapai akurasi diagnostik yang tinggi, efisiensi operasional, dan potensi untuk mengatasi kesenjangan sumber daya manusia dalam perawatan kesehatan. Integrasi prinsip fisika dalam pengembangan teknologi pencitraan medis dan analisis citra telah menjadi fondasi yang kokoh bagi kemajuan ini.

Dengan mengatasi tantangan-tantangan ini melalui penelitian yang kolaboratif dan interdisipliner, kita dapat membuka jalan bagi implementasi *deep learning* yang lebih luas, yang pada akhirnya akan meningkatkan diagnosis pneumonia, menyelamatkan nyawa, dan memperkuat sistem perawatan kesehatan secara global.

DAFTAR PUSTAKA

- Abid, M. K., Qadir, M., Alam, M., & Khan, M. A. (2024). Using Deep Learning for Pneumonia from Chest X-Ray Images. *Technical Journal, University of Engineering and Technology (UET) Taxila, Pakistan*, 29(2), 69–75. <https://tj.uettaxila.edu.pk/index.php/technical-journal/article/view/2141/381>
- Gabruseva, T., Poplavskiy, D., & Kalinin, A. (2020). Deep Learning for Automatic Pneumonia Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 350–351. Doi: [10.1109/CVPRW50498.2020.00183](https://doi.org/10.1109/CVPRW50498.2020.00183)
- Li, S., Mo, Y., & Li, Z. (2022). Automated Pneumonia Detection in Chest X-Ray Images Using Deep Learning Model. *Innovations in Applied Engineering and Technology*, 1(1), 1–6. Doi: <https://doi.org/10.62836/iaet.vli1.002>
- Omarov, B., & Bazarkulova, I. (2023). A DEEP LEARNING MODEL FOR PNEUMONIA DETECTION FROM X-RAY IMAGES. *Scientific Journal of Astana IT University*, 14, 91–103. Doi: <https://doi.org/10.37943/14ASWX8861>
- Qaimkhani, F. M., Hussain, M. G., Shiren, Y., & Xingfang, J. (2022). Pneumonia Detection Using Deep Learning Methods. *International Journal of Scientific Advances*, 3(3), 489–493. Doi: [10.51542/ijscia.v3i3.32](https://doi.org/10.51542/ijscia.v3i3.32)
- Sharma, S., & Guleria, K. (2023). A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks. *Procedia Computer Science*, 218, 357–366. <https://doi.org/10.1016/j.procs.2023.01.018>
- Sourab, S. Y., & Kabir, M. A. (2022). A comparison of hybrid deep learning models for pneumonia diagnosis from chest radiograms. *Sensors International*, 3, 100167. Doi : [10.1016/j.sintl.2022.100167](https://doi.org/10.1016/j.sintl.2022.100167)
- Stephen, O., Sain, M., Maduh, U. J., & Jeong, D.-U. (2019). An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare. *Journal of Healthcare Engineering*, 2019, 4180949. <https://doi.org/10.1155/2019/4180949>
- Szepesi, P., & Szilágyi, L. (n.d.). Detection of pneumonia using convolutional neural networks and deep learning. *Biocybernetics and Biomedical Engineering*, 42(4), 1012–1022. <https://doi.org/10.1016/j.bbe.2022.08.001>
- Urey, D. Y., Saul, C. J., & Taktakoglu, C. D. (2019). Early Diagnosis of Pneumonia with Deep

- Learning. *Arxiv preprint arxiv:1904.00937v1*. <https://doi.org/10.48550/arxiv.1904.00937>
- Vetrihangam, D., Satve, P. P., Raja Kumar, J. R., Anitha, P., Vidhya, S., & Saini, A. K. (2023). PREDICTION OF PNEUMONIA DISEASE FROM X-RAY IMAGES USING A MODIFIED RESNET152V2 DEEP LEARNING MODEL. *Journal of Theoretical and Applied Information Technology*, 101(17), 6929–6942. <https://www.jatit.org/volumes/vol101no17/16vol101no17.pdf>
- Yang, Z., & Jin, F. (2022). *Pneumonia Detection on Chest X-Rays with Deep Learning*. Arxiv preprint arxiv:2201.07724. <https://doi.org/10.48550/arxiv.1711.05225>