

Pengaruh *Progressive Resizing* pada EfficientNetB4 untuk Klasifikasi Pneumonia Bakterial dan Viral

Benedict Michael Pepper¹, Elroi Yonatan Raharjo², Windra Swastika³

^{1,2,3}Program Studi Teknik Informatika, Universitas Ma Chung, Villa Puncak Tidar Blok N No.1,
Kabupaten Malang, Jawa Timur, Indonesia, 65151

Correspondence: Benedict Michael Pepper (312310007@student.machung.ac.id)

Received: 27 05 2026 – Revised: 29 07 2026 - Accepted: 04 08 2026 - Published: 05 08 2026

Abstrak. Pneumonia merupakan penyebab utama kematian akibat infeksi pada anak di bawah usia lima tahun, menyumbang sekitar 14% dari seluruh kematian anak menurut WHO. Diferensiasi antara pneumonia bakterial dan viral berdasarkan citra X-ray dada merupakan tantangan klinis yang signifikan karena keduanya menunjukkan pola radiologis yang tumpang tindih, namun memerlukan penanganan terapeutik yang berbeda, yaitu antibiotik untuk bakteri dan penanganan suportif untuk virus. Penelitian ini membandingkan dua strategi pelatihan berbasis transfer learning pada arsitektur EfficientNetB4 untuk klasifikasi biner pneumonia bakterial versus viral: (1) strategi *Fixed Size* yang melatih model langsung pada resolusi 384x384 piksel selama 32 *epoch*, dan (2) strategi *Progressive Resizing* yang melatih melalui tiga tahap resolusi meningkat (224, 256, dan 384 piksel) selama total 41 *epoch*. Kedua model menggunakan arsitektur identik berupa *Concat Pooling head* (GlobalAveragePooling2D + GlobalMaxPooling2D), *Focal Loss*, augmentasi Albumentations dengan MixUp, serta optimizer AdamW. Menggunakan dataset Kermany et al. dengan 3.883 citra pelatihan-validasi dan set uji terpisah sebesar 390 citra (242 bakterial, 148 viral), dengan evaluasi menggunakan Test-Time Augmentation (TTA) 2-langkah deterministik, model *Fixed Size* mencapai akurasi 86,92% dan ROC AUC 0,9198, mengungguli model *Progressive Resizing* yang mencapai akurasi 78,46% dan ROC AUC 0,8011. Visualisasi Grad-CAM mengonfirmasi bahwa Model *Fixed Size* menghasilkan peta aktivasi yang lebih terlokalisasi secara spasial pada parenkima paru yang patologis. Temuan ini menunjukkan bahwa pelatihan langsung pada resolusi target lebih efektif untuk tugas klasifikasi citra medis yang bergantung pada fitur resolusi tinggi.

Kata kunci: Pneumonia, EfficientNetB4, *Progressive Resizing*, *Transfer Learning*

PENDAHULUAN

Pneumonia bertanggung jawab atas sekitar 14% dari seluruh kematian anak di bawah usia lima tahun menurut WHO (World Health Organization, 2022). Pembedaan antara etiologi bakterial dan viral menjadi prioritas klinis karena pneumonia bakterial memerlukan terapi antibiotik segera, sedangkan pneumonia viral dikelola secara suportif. Pemberian antibiotik yang tidak tepat pada kasus viral berkontribusi terhadap krisis resistensi antimikroba global (World Health Organization, 2021).

Pencitraan X-ray dada tetap menjadi modalitas diagnostik lini pertama karena ketersediaan dan biayanya yang terjangkau. Namun, diferensiasi radiologis antara pneumonia bakterial dan viral memiliki tingkat kesulitan yang tinggi bahkan bagi radiolog berpengalaman, karena kedua entitas ini sering kali menunjukkan pola konsolidasi dan infiltrat yang tumpang tindih. Tantangan ini telah mendorong upaya otomatisasi berbasis deep learning (Rajpurkar et al., 2017).

Kemajuan dalam transfer learning memungkinkan model yang dilatih pada dataset skala besar seperti ImageNet untuk diadaptasikan secara efisien pada tugas citra medis dengan dataset terbatas (Yosinski et al., 2014). Arsitektur EfficientNet (Tan & Le, 2019) melalui prinsip compound scaling telah menunjukkan performa unggul pada berbagai benchmark klasifikasi citra. Sebagai contoh penerapannya pada citra medis, Devasia et al. (2023) berhasil menggunakannya untuk deteksi tuberkulosis pada citra dada.

Meskipun beberapa penelitian telah menerapkan CNN pada klasifikasi pneumonia, studi yang secara khusus membandingkan strategi pelatihan untuk diferensiasi bakterial-viral pada EfficientNetB4 masih terbatas. Penelitian ini bertujuan: (1) menerapkan dan melatih dua model EfficientNetB4 dengan strategi pelatihan berbeda namun arsitektur identik; (2) membandingkan performa kuantitatif menggunakan metrik akurasi, presisi, *recall*, *F1-score*, dan ROC AUC; serta (3) menganalisis peta aktivasi Grad-CAM untuk memvalidasi relevansi klinis fitur yang dipelajari.

Sistematika pembahasan dalam makalah ini diawali dengan pemetaan masalah dan perumusan hipotesis secara mendalam pada bagian berikutnya. Berlandaskan pada batasan masalah tersebut, keseluruhan desain penelitian, metodologi eksperimen terkontrol, hasil evaluasi kuantitatif metrik, hingga analisis interpretabilitas model menggunakan Grad-CAM akan dijabarkan secara berturut-turut pada bab-bab selanjutnya.

MASALAH

Diagnosis konvensional pneumonia mengandalkan interpretasi citra radiografi dada. Membedakan pneumonia bakterial dari viral berdasarkan citra X-ray merupakan tugas yang menantang bahkan bagi ahli radiologi berpengalaman akibat tumpang tindih fitur visual. Kesalahan diagnosis dapat menyebabkan penggunaan antibiotik yang tidak tepat, yang berkontribusi pada resistensi antimikroba, atau penundaan terapi pada kasus bakteri yang dapat berakibat fatal.

Pada level keilmuan, pertanyaan penelitian yang belum terjawab adalah: manakah antara strategi Fixed Size dan *Progressive Resizing* yang menghasilkan performa klasifikasi lebih tinggi pada arsitektur EfficientNetB4 yang identik? Hipotesis penelitian ini adalah bahwa strategi Fixed Size akan menghasilkan akurasi yang lebih tinggi karena alokasi waktu pelatihan yang lebih lama pada resolusi target 384×384 memungkinkan model mengoptimalkan fitur pada skala yang relevan secara langsung.

Aspek teknis yang perlu diperhatikan meliputi: (1) EfficientNetB4 dioptimalkan untuk resolusi sekitar 380×380 , sehingga pelatihan pada resolusi sub-optimal dalam *Progressive Resizing* berpotensi menghambat pembelajaran; (2) fitur diagnostik X-ray seperti batas konsolidasi lobar dan pola infiltrat interstisial bersifat resolusi-kritis; (3) pemilihan fungsi *loss* dan penanganan ketidakseimbangan kelas secara simultan dapat memengaruhi performa yang diamati.

METODE PELAKSANAAN

Desain Penelitian

Penelitian ini menggunakan desain studi eksperimental komparatif terkontrol. Variabel bebas adalah strategi pelatihan (Fixed Size vs. *Progressive Resizing*), sedangkan variabel terikat adalah metrik performa klasifikasi. Semua variabel lain dikendalikan secara identik: arsitektur model, fungsi *loss*, optimizer, pipeline augmentasi, dan dataset.

Dataset

Dataset yang digunakan adalah Chest X-Ray Images (Pneumonia) oleh Kermay et al. (2018), tersedia di Kaggle, mencakup citra X-ray dada dari pasien pediatri berusia 1–5 tahun yang divalidasi oleh minimal dua dokter spesialis. Hanya subset bakterial dan viral dari folder train dan val Kaggle yang digabungkan. Distribusi pool pelatihan-validasi: 2.538 bakterial, 1.345 viral, total 3.883 citra (rasio 1,89:1). Set uji menggunakan folder test

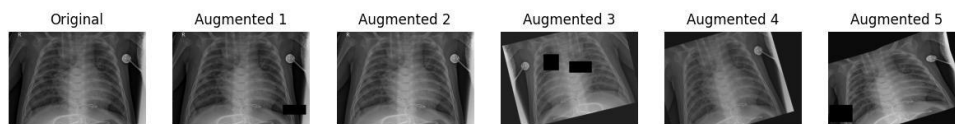
Kaggle secara terpisah (N=390; 242 bakterial, 148 viral) dan tidak pernah dilibatkan dalam pelatihan.

Tabel 1. Distribusi Dataset (Aktual dari Log Eksekusi)

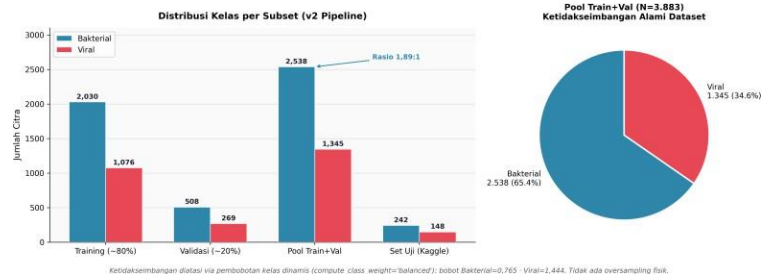
Subset	Bakterial	Viral	Total
Training (80%)	±2.030	±1.076	3.106
Validasi (20%)	±508	±269	777
Uji (terpisah)	242	148	390
Pool Train+Val	2.538	1.345	3.883

Partisi 80/20 diterapkan pada pool train+val melalui pengacakan terstratifikasi (random_state=42). Ketidakseimbangan kelas diatasi menggunakan pembobotan kelas dinamis via `compute_class_weight('balanced')` scikit-learn, menghasilkan bobot: bakterial=0,7650, viral=1,4435.

Example of Albumentations



Gambar 1. Contoh citra X-ray dada dari dataset Kermay. Kiri: pneumonia bakterial. Kanan: pneumonia viral.



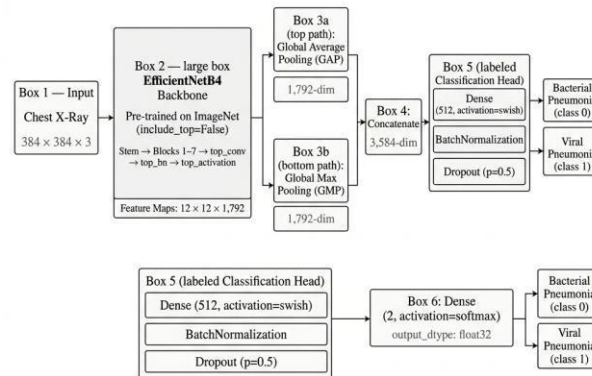
Gambar 2. Distribusi kelas pada pool pelatihan-validasi (N=3.883 citra) dan seluruh subset. Ketidakseimbangan 1,89:1 diatasi via pembobotan kelas dinamis, bukan oversampling fisik.

Arsitektur Model

Kedua model menggunakan arsitektur identik: *backbone* EfficientNetB4 (Tan & Le, 2019) dengan bobot pra-latih ImageNet dan *classification head* kustom berbasis *Concat Pooling*. *Concat Pooling* menggabungkan GlobalAveragePooling2D dan GlobalMaxPooling2D secara paralel (3.584-dim), dilanjutkan Dense(512, aktivasi Swish; Ramachandran et al., 2018), BatchNormalization, Dropout(0,5), dan Dense(2, Softmax). Prinsip compound scaling EfficientNetB4 (Alzubaidi et al., 2021; LeCun et al., 1998)

dengan koefisien $\phi=4$ menghasilkan resolusi input optimal 380×380 piksel dengan 17,67 juta parameter backbone.

EfficientNetB4 with Concatenated Pooling Head



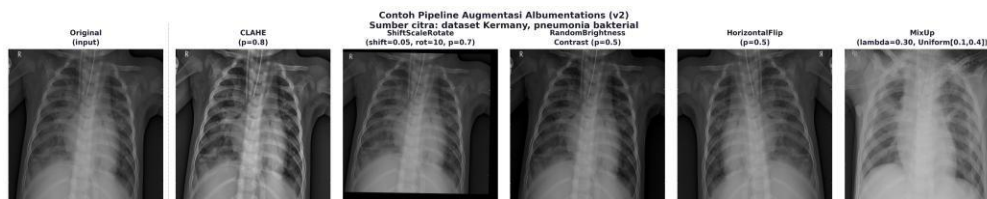
Gambar 3. Diagram arsitektur model: backbone EfficientNetB4 dengan Concat Pooling head yang menggabungkan GlobalAveragePooling2D dan GlobalMaxPooling2D secara paralel (3.584-dim).

Tabel 2. Spesifikasi Arsitektur Model

Komponen	Spesifikasi
Backbone	EfficientNetB4 (ImageNet pretrained)
Total parameter	18.594.913 (70,93 MB)
Concat Pooling	GlobalAverage + GlobalMax (3.584-dim)
Dense	512 neuron, aktivasi Swish
BatchNormalization	Diterapkan
Dropout	Rate = 0,5
Dense (output)	2 neuron, Softmax
Parameter head	921.090

Pipeline Augmentasi Data

Augmentasi diterapkan menggunakan library Albumentations (Buslaev et al., 2020) dengan transformasi yang mencerminkan variasi klinis realistis pada citra X-ray. MixUp (Zhang et al., 2018) diterapkan dalam varian bounded uniform mixing ($\lambda \sim \text{Uniform}[0,1; 0,4]$) yang mempertahankan keterbacaan fitur diagnostik. CoarseDropout tidak digunakan pada implementasi ini karena berpotensi menghilangkan region anatomis penting.



MixUp diterapkan per-batch pada GPU (lambda ~ Uniform(0.1, 0.4)); kolom kanan menunjukkan pencampuran dengan sampel viral (lambda=0.30). CoarseDropout tidak digunakan di v2.

Gambar 4. Contoh pipeline augmentasi Albumentations (v2). Dari kiri ke kanan: Original → CLAHE → ShiftScaleRotate → RandomBrightnessContrast → HorizontalFlip → MixUp.

Tabel 3. Pipeline Augmentasi Albumentations

Transformasi	Parameter
CLAHE (Pizer et al., 1987)	clip_limit=2,0; tile_grid=(8,8); p=0,8
ShiftScaleRotate	shift=0,05; scale=0,10; rotate=10°; p=0,7
RandomBrightnessContrast	brightness=0,10; contrast=0,10; p=0,5
HorizontalFlip	p=0,5
MixUp (Zhang et al., 2018)	$\lambda \sim \text{Uniform}[0,1; 0,4]$ (GPU, per-batch)
Normalisasi	Pixel ÷ 255,0 (skala [0,1]); konsisten di seluruh split

Konfigurasi Pelatihan

Kedua model menggunakan konfigurasi umum sebagai berikut: Focal Loss (Lin et al., 2017) dengan $\gamma=2,0$, $\alpha=0,25$, dan label_smoothing=0,1 sebagai fungsi loss; optimizer AdamW (Loshchilov & Hutter, 2019) dengan LossScaleOptimizer; framework TensorFlow 2.19.0/Keras; GPU Tesla P100-PCIE-16GB; presisi numerik mixed precision (float16/float32); serta TTA 2-langkah deterministik (citra asli + flip horizontal). Raghu et al. (2019) dan Yosinski et al. (2014) memvalidasi relevansi fitur ImageNet untuk domain medis sebagai dasar penggunaan transfer learning.

Tabel 4. Hyperparameter Umum Kedua Model

Hyperparameter	Nilai
Fungsi loss	CategoricalFocalCrossentropy ($\gamma=2,0$; $\alpha=0,25$; label_smoothing=0,1)
Optimizer	AdamW dengan LossScaleOptimizer
LR head training	1×10^{-3}
LR fine-tuning (Fixed Size)	1×10^{-4} (ReduceLROnPlateau, factor=0,5, patience=3)
LR fine-tuning (Progressive)	CosineDecayRestarts (awal 1×10^{-5} ; t_mul=2,0; m_mul=0,9)
Framework	TensorFlow 2.19.0 / Keras
GPU	Tesla P100-PCIE-16GB
Presisi numerik	Mixed precision (float16/float32)
TTA: forward pass	2 (asli + flip horizontal)

Model 1 – *Progressive Resizing* (224→256→384). Strategi ini terinspirasi dari prinsip curriculum learning (Bengio et al., 2009) dan dipopulerkan oleh Howard dan Guggenberger (2020) dalam fastai. Pelatihan dibagi menjadi tiga stage resolusi berurutan. Bobot diwarisi antar-stage melalui transfer file .weights.h5; classifier head tidak diinisialisasi ulang. Seluruh fine-tuning dilakukan dengan backbone dibuka dari block6a ke atas.

Tabel 5. Konfigurasi *Progressive Resizing* per Stage (Aktual dari Log Eksekusi)

Parameter	Stage 1 (224×224)	Stage 2 (256×256)	Stage 3 (384×384)
Batch size	32	24	16
Head training epochs	4	diwarisi	diwarisi
Fine-tuning epochs	15	7	15
Total epochs (stage)	19	7	15
Fine-tune dari block	block6a+	block6a+	block6a+

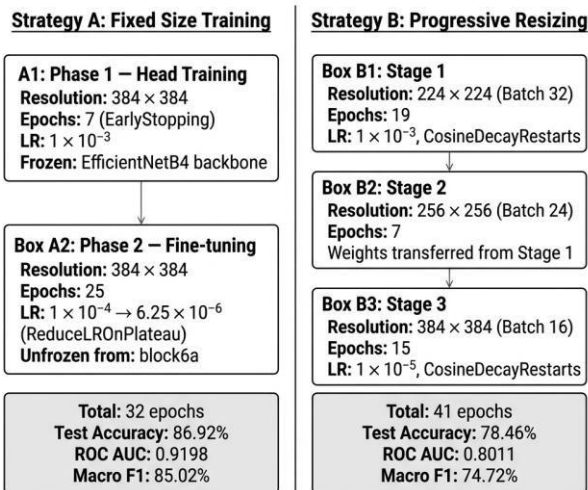
EarlyStopping patience (FT)	5	5	5
<i>Epoch</i> kumulatif	19	26	41

Model 2 – Fixed Size (384×384). Model Fixed Size melatih langsung pada resolusi target 384×384 menggunakan skema dua-fase dengan EarlyStopping berbasis val_auc. LR fine-tuning menggunakan ReduceLROnPlateau yang menurunkan LR dari 1×10^{-4} hingga $6,25 \times 10^{-6}$ secara adaptif.

Tabel 6. Konfigurasi Fixed Size (Aktual dari Log Eksekusi)

Parameter	Phase 1: Head Training	Phase 2: Fine-tuning
Resolusi	384×384	384×384
Batch size	16	16
<i>Epochs</i> (aktual)	7	25
Parameter terlatih	921.090	9.556.294
LR	1×10^{-3}	$1 \times 10^{-4} \rightarrow 6,25 \times 10^{-6}$
EarlyStopping patience	4 (monitor: val_auc)	8 (monitor: val_auc; restore_best_weights=True)
Total <i>epoch</i>	32 (7+25)	–

Comparison of Training Resolution Strategies



Gambar 5. Perbandingan skema dua strategi pelatihan. Kiri: Model 2 Fixed Size, dua fase pada resolusi tetap 384×384 (total 32 *epoch*). Kanan: Model 1 *Progressive Resizing*, tiga stage resolusi bertingkat (total 41 *epoch*).

Evaluasi Model

Evaluasi pada set uji (N=390; 242 bakterial : 148 viral) menggunakan prediksi TTA 2-langkah deterministik. Metrik yang dihitung: akurasi, presisi, *recall* (sensitivitas), *F1-score*, ROC AUC, dan *confusion matrix*. Interpretabilitas model dievaluasi menggunakan Grad-CAM (Selvaraju et al., 2017) pada lapisan aktivasi terakhir EfficientNetB4 (top_activation). Perbedaan performa antar model merupakan estimasi titik dari satu partisi

tetap tanpa uji signifikansi statistik; pembaca disarankan mempertimbangkan hal ini saat menginterpretasi selisih metrik.

HASIL DAN PEMBAHASAN

Perbandingan Performa Kuantitatif

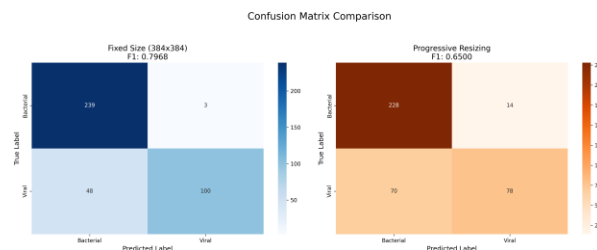
Tabel 7 merangkum seluruh metrik evaluasi kedua model pada set uji (N=390; 242 bakterial : 148 viral) menggunakan prediksi TTA 2-langkah. Model Fixed Size (Model 2) secara konsisten mengungguli Model *Progressive Resizing* (Model 1) di semua metrik evaluasi.

Tabel 7. Perbandingan Metrik Evaluasi Kedua Model (Set Uji, N=390; dengan TTA)

Metrik	Model 1 (Prog. Resizing)	Model 2 (Fixed Size)	Δ (Fixed-Prog.)
Akurasi	78,46%	86,92%	+8,46%
ROC AUC	0,8011	0,9198	+0,1187
Presisi (Bakterial)	76,51%	83,28%	+6,77%
Presisi (Viral)	84,78%	97,09%	+12,31%
Recall (Bakterial)	94,21%	98,76%	+4,55%
Recall (Viral)	52,70%	67,57%	+14,87%
F1-Score (Bakterial)	84,44%	90,36%	+5,92%
F1-Score (Viral)	65,00%	79,68%	+14,68%
Macro F1	74,72%	85,02%	+10,30%

Matriks Konfusi

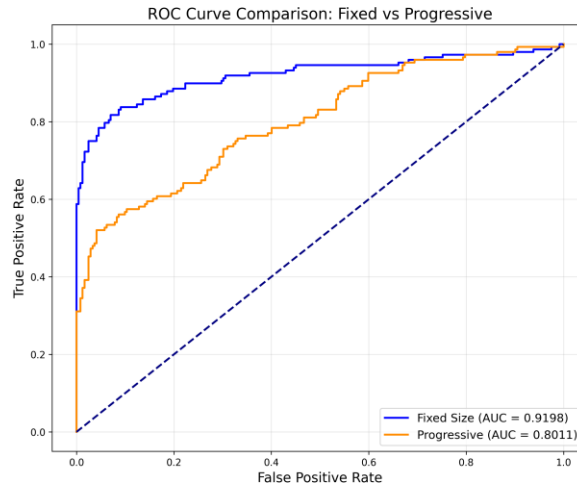
Model 2 (*Fixed Size*) menunjukkan perbaikan di kedua kelas: hanya 3 dari 242 bakterial salah diklasifikasikan sebagai viral (1,2% FNR bakterial) dan 48 dari 148 viral salah diklasifikasikan sebagai bakterial (32,4% FNR viral). Model 1 (*Progressive Resizing*) menunjukkan asimetri yang lebih besar: 14 dari 242 bakterial salah sebagai viral (5,8% FNR bakterial) dan 70 dari 148 viral salah sebagai bakterial (47,3% FNR viral). Kedua model mengalami kesulitan lebih besar pada kelas viral, namun Model 2 lebih baik dalam mendeteksi kedua kelas secara konsisten.



Gambar 6. *Confusion matrix* pada set uji (N=390, dengan TTA). Kiri: Model 2 (*Fixed Size*). Kanan: Model 1 (*Progressive Resizing*).

Kurva ROC

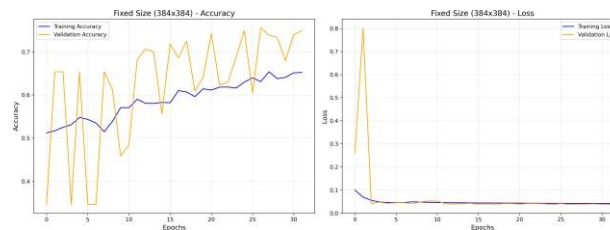
Model 2 (*Fixed Size*, AUC=0,9198) menunjukkan kurva ROC yang secara konsisten berada di atas Model 1 (AUC=0,8011), terutama pada rentang *False Positive Rate* rendah (0,0–0,2) yang paling relevan secara klinis.



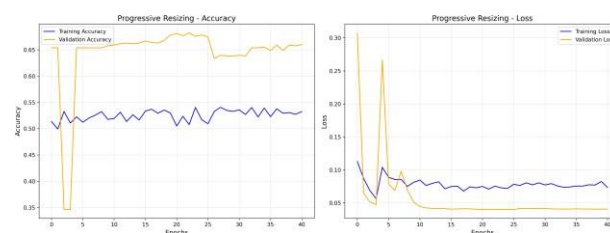
Gambar 7. Kurva ROC pada set uji (N=390). Biru: Model 2 *Fixed Size* (AUC=0,9198). Oranye: Model 1 *Progressive Resizing* (AUC=0,8011).

Kurva Pelatihan

Model 1 (*Progressive Resizing*, 41 *epoch* total) menunjukkan lonjakan loss pada pergantian stage (*epoch* 19 dan 26), yang merupakan artefak normal dari weight transfer antar-resolusi. Nilai val_auc terbaik yang dicapai selama pelatihan adalah 0,6989 (Stage 1, *epoch* 12). Model 2 (*Fixed Size*, 32 *epoch* total) menyelesaikan pelatihan berdasarkan batas maksimum *epoch* dengan *epoch* terbaik pada *epoch* ke-18 (val_auc=0,8041); bobot terbaik dipulihkan via `restore_best_weights=True`. `ReduceLROnPlateau` menurunkan LR secara bertahap dari 1×10^{-4} ke $6,25 \times 10^{-6}$.



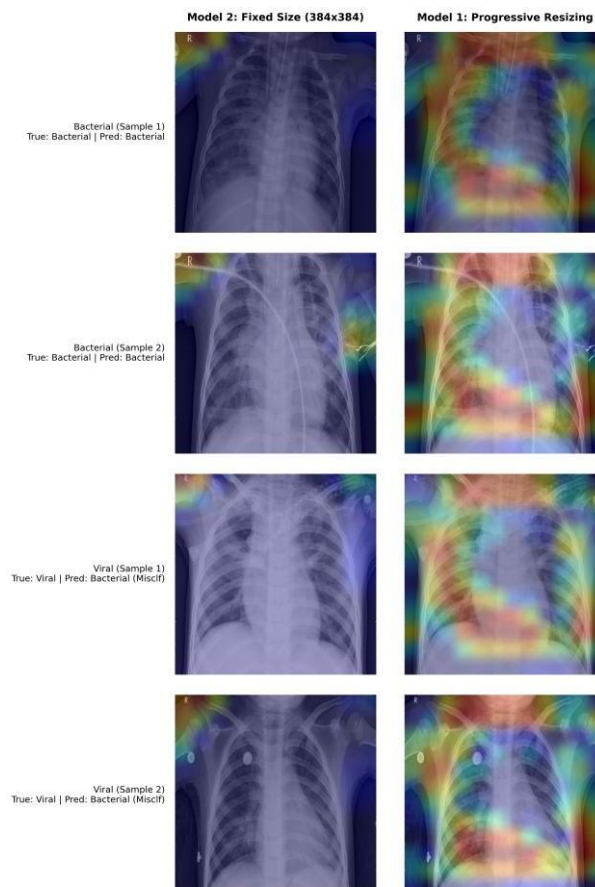
Gambar 8. Kurva akurasi dan *loss* pelatihan: Model 2 *Fixed Size* (32 *epoch* total). Train=biru, Validasi=oranye.



Gambar 9. Kurva akurasi dan *loss* pelatihan: Model 1 *Progressive Resizing* (41 *epoch* kumulatif; tiga stage).

Visualisasi Grad-CAM

Grad-CAM (Selvaraju et al., 2017) pada lapisan *top_activation* EfficientNetB4 ($12 \times 12 \times 1.792$) mengonfirmasi bahwa kedua model mengaktifkan region anatomis konsisten dengan lokasi lesi paru. Model *Fixed Size* menunjukkan peta aktivasi yang lebih terlokalisasi secara spasial pada parenkima paru yang patologis. Pada kasus viral yang salah diklasifikasikan (misklasifikasi sebagai bakterial pada kedua model), Grad-CAM menunjukkan bahwa aktivasi merespons pola bilateral yang menyebar, karakteristik infiltrat interstisial, namun kedua model gagal menemukan fitur diskriminatif yang cukup kuat.



Gambar 10. Perbandingan visualisasi Grad-CAM pada lapisan *top_activation* EfficientNetB4 ($12 \times 12 \times 1.792$). Model *Fixed Size* (kiri) menunjukkan aktivasi lebih terlokalisasi pada region patologis.

Faktor Penyebab Keunggulan *Fixed Size*

Keunggulan *Fixed Size* dapat dikaitkan dengan empat faktor utama. Pertama, alokasi epoch lebih efektif: Model *Fixed Size* menghabiskan seluruh 32 *epoch* pada resolusi 384×384 , sedangkan *Progressive Resizing* hanya menghabiskan 15 *epoch* terakhir (*Stage* 3) pada resolusi target dari 41 *epoch* total. Kedua, perbedaan LR fine-tuning: *Fixed Size* menggunakan LR awal 1×10^{-4} ($10 \times$ lebih tinggi dari CosineDecayRestarts 1×10^{-5} pada *Progressive*), memungkinkan adaptasi bobot lebih cepat. Ketiga, sensitivitas resolusi EfficientNetB4: arsitektur dioptimalkan untuk 380×380 piksel, sehingga pelatihan pada resolusi 224 dan 256 piksel mengakibatkan adaptasi fitur yang tidak optimal. Keempat, karakteristik diagnostik X-ray bergantung pada resolusi tinggi: fitur kunci seperti batas konsolidasi lobar versus pola infiltrat interstisial halus hanya dapat dibedakan optimal pada resolusi tinggi, menantang asumsi curriculum learning (Bengio et al., 2009).

Perbandingan dengan Penelitian Sebelumnya

Tabel 8. Tabel 8. Perbandingan dengan Penelitian Sebelumnya pada Dataset Kermamy

Penelitian	Model	Tugas	Akurasi
Penelitian ini (<i>Fixed Size</i>)	EfficientNetB4	B vs. V	86,92%
Penelitian ini (Prog. Resizing)	EfficientNetB4	B vs. V	78,46%
Mardianto et al. (2024)	CNN Kustom	B vs. V	92,00%
Abubakar et al. (2021)	NASNet+SVM	B/N/V	92,90%
Huang et al. (2025)	ResNet50V2-SE	4-kelas	98,18%

Perbandingan langsung antar-studi memiliki keterbatasan nyata karena protokol augmentasi, penyeimbangan kelas, dan partisi set uji berbeda antar-studi. Kontribusi penelitian ini adalah pada investigasi sistematis pengaruh strategi pelatihan untuk tugas diferensiasi bakterial-viral. Model *Fixed Size* berpotensi diterapkan sebagai sistem pendukung keputusan diagnostik (CAD) awal, meski validasi eksternal dan perbandingan langsung dengan performa radiolog manusia tetap diperlukan.

Keterbatasan Penelitian

Keterbatasan penelitian ini meliputi: (1) dataset tunggal dari populasi pediatri spesifik, sehingga generalisasi belum divalidasi; (2) arsitektur tunggal EfficientNetB4, tanpa studi perbandingan ResNet, ViT, atau ConvNeXt; (3) tidak ada ablasi Focal Loss vs. standard cross-entropy atau ablasi pembobotan kelas dinamis secara terpisah; (4) tidak ada evaluasi pada dataset eksternal; (5) set uji tidak seimbang (242:148) merepresentasikan distribusi alami namun memengaruhi interpretasi metrik weighted average; (6) ketidaksetaraan total *epoch* (41 vs. 32) membatasi perbandingan efisiensi pelatihan secara

langsung; dan (7) selisih performa merupakan estimasi titik dari satu partisi, sehingga uji signifikansi statistik tidak dihitung.

KESIMPULAN

Penelitian ini membandingkan dua strategi pelatihan, yaitu *Fixed Size* (384×384) dan *Progressive Resizing* (224→256→384), pada arsitektur EfficientNetB4 yang identik untuk klasifikasi biner pneumonia bakterial versus viral pada citra X-ray dada. Menggunakan dataset Kermany et al. dengan 3.883 citra pool pelatihan-validasi dan set uji terpisah 390 citra, strategi *Fixed Size* terbukti lebih efektif dengan akurasi 86,92% dan ROC AUC 0,9198, mengungguli *Progressive Resizing* (78,46%, AUC 0,8011) dengan selisih +8,46% akurasi dan +0,1187 AUC. Perbedaan performa dikaitkan dengan: (1) alokasi *epoch* yang lebih efektif pada resolusi target, (2) LR fine-tuning 10× lebih tinggi pada *Fixed Size*, dan (3) kesesuaian langsung antara resolusi pelatihan dengan resolusi optimal EfficientNetB4. Grad-CAM mengonfirmasi aktivasi region paru yang relevan secara anatomis pada kedua model, dengan Model *Fixed Size* menunjukkan lokalisasi spasial yang lebih presisi. Penelitian lanjutan yang mendesak meliputi validasi pada dataset eksternal, studi ablasi komponen individual (*Focal Loss*, TTA, strategi unfreezing), dan investigasi arsitektur berbasis perhatian seperti Vision Transformer (ViT).

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Kermany, Zhang, dan Goldbaum atas penyediaan dataset Chest X-Ray Images (Pneumonia) melalui Mendeley Data, serta kepada Kaggle yang memfasilitasi akses terhadap data tersebut untuk komunitas riset. Terima kasih juga kepada seluruh pengembang library open-source yang digunakan dalam penelitian ini (TensorFlow, Albumentations, scikit-learn)

DAFTAR PUSTAKA

- Abubakar, M. M., Adamu, B. Z., & Abubakar, M. Z. (2021). Pneumonia classification using hybrid CNN architecture. Dalam Prosiding 2021 International Conference on Data Analytics for Business and Industry (ICDABI) (hlm. 520–522). IEEE. <https://doi.org/10.1109/ICDABI53623.2021.9655918>
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaria, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep

- learning: Concepts, CNN architectures, challenges, and future directions. *Journal of Big Data*, 8(1), 53. <https://doi.org/10.1186/s40537-021-00444-8>
- Bengio, Y., Louradour, J., Collobert, R., & Weston, J. (2009). Curriculum learning. Dalam *Prosiding 26th International Conference on Machine Learning (ICML)* (hlm. 41–48). ACM. <https://doi.org/10.1145/1553374.1553380>
- Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Papostolakis, A., Natekin, A., & Kalinin, A. A. (2020). Albumentations: Fast and flexible image augmentations. *Information*, 11(2), 125. <https://doi.org/10.3390/info11020125>
- Devasia, J., Goswami, H., Lakshminarayanan, S., Rajaram, M., & Adithan, S. (2023). Deep learning classification of active tuberculosis lung zones wise manifestations using chest X-rays: A multi label approach. *Scientific Reports*, 13, 887. <https://doi.org/10.1038/s41598-023-28079-0>
- Howard, J., & Gugger, S. (2020). Fastai: A layered API for deep learning. *Information*, 11(2), 108. <https://doi.org/10.3390/info11020108>
- Huang, K. A., Choudhary, H. K., Santiago, A., & Prakash, N. S. (2025). Squeeze-and-excitation enhanced convolutional neural networks for multi-class pneumonia classification on chest radiographs. *Cureus*, 17(12), e99649. <https://doi.org/10.7759/cureus.99649>
- Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., & Zhang, K. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5), 1122–1131.e9. <https://doi.org/10.1016/j.cell.2018.02.010>
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324. <https://doi.org/10.1109/5.726791>
- Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollar, P. (2017). Focal loss for dense object detection. Dalam *Prosiding IEEE/CVF International Conference on Computer Vision (ICCV)* (hlm. 2980–2988). IEEE. <https://doi.org/10.1109/ICCV.2017.324>
- Loshchilov, I., & Hutter, F. (2019). Decoupled weight decay regularization. Dalam *Prosiding International Conference on Learning Representations (ICLR)*. <https://doi.org/10.48550/arXiv.1711.05101>
- Mardianto, M. F. F., Yoani, A., Soewignjo, S., Putra, I. K. P. K. A., & Dewi, D. A. (2024). Classification of pneumonia from chest X-ray images using support vector machine and convolutional neural network. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 15(6). <https://doi.org/10.14569/IJACSA.2024.01506104>
- Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J. B., & Zuiderveld, K. (1987). Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*, 39(3), 355–368. [https://doi.org/10.1016/S0734-189X\(87\)80186-X](https://doi.org/10.1016/S0734-189X(87)80186-X)

- Raghu, M., Zhang, C., Kleinberg, J., & Bengio, S. (2019). Transfusion: Understanding transfer learning for medical imaging. *Advances in Neural Information Processing Systems (NeurIPS)*, 32. <https://doi.org/10.48550/arXiv.1902.07208>
- Rajpurkar, P., Irvin, J., Ball, R. L., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv:1711.05225*. <https://arxiv.org/abs/1711.05225>
- Ramachandran, P., Zoph, B., & Le, Q. V. (2018). Searching for activation functions. *Dalam Prosiding International Conference on Learning Representations (ICLR)*. <https://doi.org/10.48550/arXiv.1710.05941>
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Dalam Prosiding IEEE International Conference on Computer Vision (ICCV)* (hlm. 618–626). IEEE. <https://doi.org/10.1109/ICCV.2017.74>
- Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *Dalam Prosiding 36th International Conference on Machine Learning (ICML)*. <https://doi.org/10.48550/arXiv.1905.11946>
- World Health Organization. (2021). Antimicrobial resistance. <https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance>
- World Health Organization. (2022). Pneumonia in children. <https://www.who.int/news-room/fact-sheets/detail/pneumonia>
- Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems (NeurIPS)*, 27. <https://doi.org/10.48550/arXiv.1411.1792>
- Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. (2018). Mixup: Beyond empirical risk minimization. *Dalam Prosiding International Conference on Learning Representations (ICLR)*. <https://doi.org/10.48550/arXiv.1710.09412>

s



© 2026 by authors. Content on this article is licensed under a Creative Commons Attribution 4.0 International license. (<http://creativecommons.org/licenses/by/4.0/>).