

LAPORAN AKHIR
MATA KULIAH DEEP LEARNING (B)
“KLASIFIKASI PNEUMONIA MENGGUNAKAN MODEL
EFFICIENTNETB0”



Nama Anggota:

- | | |
|-------------------------------|-------------------------|
| 1. Krisnawan | : 22083010068 (Ketua) |
| 2. Talitha Adilla Fujisai P.P | : 22083010079 (Anggota) |
| 3. Cahya Eka Melati | : 22083010090 (Anggota) |
| 4. Wahyu Putra Pratama | : 22083010092 (Anggota) |
| 5. Hanin Anindya Putri | : 22083010086 (Anggota) |

Dosen Pengampu :

Amri Muhaimin, S.Stat., M.Stat., M.S. (NIP. 19950723202406100)

Program Studi Sains Data

Fakultas Ilmu Komputer

Universitas Pembangunan Nasional “Veteran” Jawa Timur

2025

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi digital telah mendorong kemajuan signifikan di berbagai sektor, termasuk bidang kesehatan. Salah satu inovasi yang banyak dimanfaatkan adalah kecerdasan buatan (*artificial intelligence*), khususnya teknik *deep learning*. Teknologi ini memiliki kemampuan untuk mempelajari pola dari data dalam jumlah besar, sehingga sangat efektif digunakan dalam berbagai aplikasi, terutama analisis citra medis. Dalam dunia radiologi, pemanfaatan *deep learning* semakin relevan karena proses interpretasi gambar medis seperti foto rontgen (X-ray) karena dapat meningkatkan efisiensi dan akurasi diagnosis penyakit paru-paru seperti pneumonia [1]. Penyakit ini merupakan infeksi pada jaringan paru-paru yang menimbulkan peradangan dan sering kali menyebabkan gangguan pernapasan serius yang disebabkan oleh berbagai patogen, termasuk virus, bakteri, dan jamur [2]. Pneumonia masih menjadi satu diantar penyebab utama angka kesakitan dan kematian di banyak negara, terutama pada kelompok rentan seperti anak-anak, lanjut usia, dan pasien dengan imunitas rendah. Deteksi dini sangat penting untuk mencegah perburukan kondisi pasien, tetapi interpretasi citra X-ray secara manual memerlukan tenaga ahli yang berpengalaman dan dapat dipengaruhi oleh beban kerja, kelelahan, serta subjektivitas penilaian yang dibutuhkan [1].

Melihat berbagai tantangan tersebut, pengembangan sistem pendukung diagnosis berbasis *deep learning* menjadi alternatif yang menjanjikan. Salah satu arsitektur yang terbukti efisien untuk klasifikasi citra adalah *Efficient Net*. Model ini dirancang menggunakan pendekatan *compound scaling* sehingga dapat mencapai akurasi tinggi dengan penggunaan sumber daya komputasi yang lebih efisien [3]. Dalam konteks klasifikasi pneumonia, *Efficient Net* dinilai mampu mengenali pola halus pada citra X-ray secara lebih stabil dibandingkan arsitektur konvolusional lainnya [4]. Penelitian ini memanfaatkan arsitektur *Efficient Net* dengan pendekatan *transfer learning* melalui dua tahap pelatihan, yaitu *training head* dan *fine-tuning*. Tahap *training head* dilakukan untuk melatih lapisan klasifikasi tanpa mengubah bobot lapisan dasar, sedangkan tahap *fine-tuning* membuka kembali beberapa lapisan terakhir agar model dapat menyesuaikan diri dengan pola pneumonia yang lebih spesifik. Proses pelatihan juga didukung mekanisme *callbacks* seperti *ModelCheckpoint*, *ReduceLROnPlateau*, dan *EarlyStopping* untuk meningkatkan stabilitas, efisiensi, dan mencegah overfitting. Melalui pendekatan ini, penelitian diharapkan mampu menghasilkan model klasifikasi pneumonia yang akurat, cepat, dan dapat berfungsi sebagai alat bantu dalam proses diagnosis klinis.

1.2 Rumusan Masalah

1. Bagaimana kinerja model EfficientNetB0 dalam mengklasifikasikan citra X-ray dada antara kondisi paru-paru normal dan kondisi yang terindikasi pneumonia?

2. Bagaimana cara untuk mengatasi *imbalanced* data pada dataset citra penyakit pneumonia?
3. Bagaimana pengaruh dua tahapan pelatihan *training head* dan *fine-tuning* terhadap peningkatan akurasi serta kemampuan model dalam mengenali pola-pola pneumonia pada citra X-ray medis?

1.3 Tujuan

1. Menilai kemampuan model EfficientNetB0 dalam melakukan klasifikasi citra X-ray dada antara kondisi paru-paru normal dan yang menunjukkan indikasi pneumonia.
2. Menggunakan teknik *class weighting* pada kelas minoritas di dataset citra penyakit pneumonia
3. Mengidentifikasi sejauh mana tahapan pelatihan *training head* dan *fine-tuning* berkontribusi terhadap peningkatan akurasi serta ketepatan model dalam mengenali pola-pola pneumonia pada citra medis.

BAB II

TINJAUAN PUSTAKA

2.1 Pneumonia

Pneumonia adalah penyakit infeksi pada paru-paru yang menyebabkan peradangan pada jaringan alveoli, yaitu kantung udara tempat pertukaran oksigen terjadi [5]. Pada kondisi normal, alveoli berisi udara, namun pada pneumonia bagian ini dapat terisi cairan, lendir, atau nanah sehingga mengganggu proses pernapasan dan menurunkan suplai oksigen ke tubuh. Penyakit ini umumnya disebabkan oleh bakteri *Streptococcus pneumoniae*, tetapi juga dapat dipicu oleh virus, jamur, maupun mikroorganisme lainnya. Secara klinis, pneumonia ditandai dengan gejala seperti batuk, demam, sesak nafas, nyeri dada, hingga penurunan saturasi oksigen. Tingkat keparahannya bervariasi, mulai dari ringan hingga mengancam jiwa, terutama pada anak-anak, lansia, dan individu dengan sistem kekebalan tubuh yang lemah. Dalam konteks radiologi, pneumonia umumnya terlihat pada citra X-ray melalui adanya opasitas, infiltrasi, atau konsolidasi pada area paru. Pola-pola ini sering kali memiliki karakteristik halus dan bervariasi sehingga memerlukan ketelitian dalam interpretasi. Oleh karena itu, penggunaan model *deep learning* seperti *Efficient Net* menjadi relevan untuk membantu proses deteksi dini dan klasifikasi pneumonia secara lebih cepat, akurat, dan konsisten.

2.2 Klasifikasi Citra Medis

Klasifikasi citra medis merupakan satu diantara aplikasi penting dalam bidang kecerdasan buatan yang bertujuan untuk mengidentifikasi kondisi atau penyakit tertentu berdasarkan citra diagnostik, seperti X-Ray, CT-Scan, atau MRI [6]. Proses klasifikasi ini berfokus pada pengelompokan citra ke dalam kategori tertentu, misalnya citra normal dan citra yang menunjukkan indikasi penyakit. Klasifikasi citra medis memiliki beberapa tantangan antara lain adalah variasi kualitas citra, perbedaan anatomi antar pasien, keterbatasan jumlah data berlabel, serta keberadaan *noise* pada citra. Interpretasi hasil citra sangat bergantung pada keahlian dokter spesialis, yang tidak jarang menimbulkan variasi dalam diagnosis dan potensi kesalahan [7]. Untuk mengatasi tantangan tersebut, berbagai pendekatan komputasional telah dikembangkan, mulai dari metode komputasi tradisional berbasis ekstraksi fitur manual hingga pendekatan modern berbasis *deep learning* yang mampu mempelajari pola secara otomatis dari data dalam jumlah besar.

2.3 Deep Learning

Deep learning merupakan salah satu cabang dari pembelajaran mesin yang memanfaatkan model jaringan saraf tiruan (*artificial neural networks*) dengan banyak lapisan (*layers*) untuk menganalisis data [8]. Secara teoritis, *deep learning* berfokus pada pemodelan struktur dan pola kompleks dalam data melalui hierarki representasi. Setiap lapisan dalam jaringan saraf mengolah input secara bertahap untuk menghasilkan *output* yang lebih abstrak, yang memungkinkan sistem untuk belajar secara otomatis tanpa memerlukan fitur yang

ditentukan secara manual. Dalam dunia medis, teknologi ini telah diterapkan secara luas, terutama dalam pengolahan citra medis guna mendeteksi dan mendiagnosis penyakit secara otomatis dan lebih akurat [9]. Penerapan *deep learning* tidak hanya meningkatkan akurasi dalam mendeteksi penyakit, tetapi juga memungkinkan identifikasi penyakit secara lebih dini daripada metode manual konvensional [8].

2.4 Transfer Learning

Transfer learning adalah teknik dalam *deep learning* yang memungkinkan model yang telah dilatih pada satu tugas untuk digunakan kembali dalam tugas yang berbeda namun serupa. Secara teoritis, *transfer learning* memanfaatkan pengetahuan yang telah dipelajari oleh model dari data dan pola yang ada pada tugas pertama, dan kemudian mentransfer pengetahuan tersebut ke tugas lain yang memiliki kesamaan dalam hal fitur atau struktur data [10]. Pendekatan ini sangat berguna ketika data untuk tugas baru terbatas atau sulit diperoleh, karena model dapat memanfaatkan fitur-fitur yang telah dipelajari sebelumnya untuk mempercepat proses pelatihan dan meningkatkan performa [8]. Dalam *transfer learning*, model yang telah dilatih pada *source task* biasanya akan menggunakan bobot dan struktur jaringan yang sama pada *target task*, meskipun beberapa lapisan terakhir pada jaringan dapat disesuaikan ulang untuk lebih sesuai dengan tugas baru. *Transfer learning* banyak digunakan dalam aplikasi seperti pengenalan gambar, pemrosesan bahasa alami, dan analisis suara, di mana model dapat memanfaatkan pelatihan yang telah dilakukan pada dataset besar dan kemudian diadaptasi untuk tugas yang lebih spesifik dengan data yang lebih kecil. Dengan cara ini, *transfer learning* mengurangi kebutuhan akan pelatihan model dari awal, yang menghemat waktu dan sumber daya komputasi [10].

2.5 Model EfficientNet

EfficientNet merupakan keluarga arsitektur *Convolutional Neural Network* (CNN) yang dikembangkan untuk mencapai keseimbangan optimal antara akurasi, efisiensi komputasi, dan jumlah parameter. Model ini dikembangkan sebagai solusi atas keterbatasan CNN generasi sebelumnya, yaitu peningkatan kedalaman jaringan (*depth*), lebar filter (*width*), atau resolusi input (*resolution*) sering dilakukan secara terpisah sehingga membuat model menjadi berat, tidak stabil, dan kurang efisien. EfficientNet memiliki 8 model yang dimulai dari B0 hingga B7, makin tinggi nomor model maka parameternya juga bertambah dan akurasi juga terus meningkat [11]. Tan dan Le dari *Google Brain* pada tahun 2019 memperkenalkan metode *compound scaling* untuk mengatasi masalah tersebut. Pendekatan ini menskalakan *depth*, *width*, dan *resolution* secara proporsional sehingga struktur jaringan tetap seimbang dan stabil ketika diperbesar [12]. Arsitektur dasar EfficientNet-B0 sendiri diperoleh melalui *Neural Architecture Search* (NAS), yang secara otomatis mengoptimalkan konfigurasi lapisan berdasarkan kriteria efisiensi dan performa. Secara struktural, EfficientNet terdiri dari beberapa blok berulang yang terdiri dari *Depthwise Separable Convolutions*, disertai dengan *Swish Non Linearity* yang berfungsi sebagai fungsi aktivasi [13]. Kombinasi ini membantu model menangkap pola visual yang kompleks

secara efisien. Model EfficientNet-B0 tetap sangat efektif dalam upaya klasifikasi gambar walaupun jumlah parameternya tergolong sedikit [14]. Teknik *Depthwise Separable Convolutions* memperkenalkan pendekatan baru dengan membagi proses konvolusi standar menjadi dua tahap, yaitu *depthwise convolution* dan *pointwise convolution*. Dalam tahap *depthwise*, saluran *input* individual diproses secara terpisah menggunakan filter yang hanya berisi satu saluran. Selanjutnya, pada tahap *pointwise*, konvolusi 1x1 digunakan untuk menggabungkan saluran - saluran ini, menghasilkan *output* akhir [14].

Keunggulan utama EfficientNet terletak pada rasio akurasi terhadap komputasi yang lebih tinggi dibandingkan arsitektur populer seperti ResNet atau Inception. Dengan parameter yang lebih sedikit namun akurasi lebih tinggi, model ini sangat sesuai untuk aplikasi medis yang membutuhkan analisis cepat dan presisi. Dalam klasifikasi pneumonia pada citra X-ray, EfficientNet efektif dalam mengenali pola radiologis halus seperti infiltrasi dan opasitas. Melalui transfer learning, bobot pra-latih dari ImageNet dapat disesuaikan dengan domain medis melalui training head dan fine-tuning, sehingga menghasilkan performa yang stabil dan akurat.

BAB III

METODOLOGI

3.1 Sumber Data

Penelitian ini menggunakan data sekunder Chest X-Ray Images (Pneumonia) yang bersumber dari Kaggle. Link dataset: [Chest X-Ray Images \(Pneumonia\)](#).

3.1.1 Deskripsi Dataset

Dataset pada penelitian ini terdiri dari citra X-ray dada yang di kelompokkan menjadi dua kelas:

1. Normal

Citra yang menunjukkan kondisi paru-paru sehat atau bersih tanpa area opasifikasi abnormal.

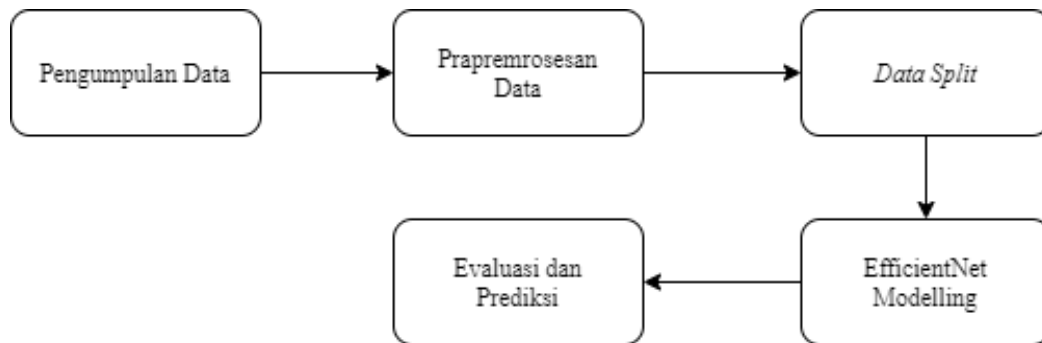
2. Pneumonia

Citra yang menunjukkan kondisi pneumonia. Pneumonia dengan bakteri yang biasanya terdapat konsolidasi lobar fokal, sedangkan pneumonia virus bermanifestasi dengan pola interstisial yang lebih difus di kedua paru-paru.

Dataset *Chest X-Ray Images* (Pneumonia) memiliki total 5.863 citra, yang dibagi menjadi 3 bagian, yaitu:

1. *Train* dengan jumlah citra normal 1.341 dan pneumonia 3.875.
2. *Test* dengan jumlah citra normal 234 dan pneumonia 380.
3. *Validation* dengan jumlah citra normal 8 dan pneumonia 8.

3.2 Flowchart



Gambar 3.1 Diagram Alur Kerja

3.3 Metodologi

3.3.1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder berupa citra X-Ray dada yang diperoleh dari dataset *Chest X-Ray Images* (Pneumonia) yang tersedia pada *platform* Kaggle. Dataset ini dipilih karena telah banyak digunakan dalam penelitian sebelumnya serta menyediakan anotasi kelas yang jelas, yaitu kelas *Normal* dan *Pneumonia*. Total citra yang

Model EfficientNet yang digunakan dalam penelitian ini yaitu EfficientNet-B0, yang dipilih untuk membandingkan performa arsitektur dengan kompleksitas rendah dan tinggi. Model memanfaatkan bobot awal (*pre-trained weights*) dari dataset ImageNet untuk mempercepat proses pembelajaran fitur dasar citra.

Pada tahap *fine-tuning*, sebanyak 30 lapisan terakhir dari masing-masing arsitektur EfficientNet dibuka (*unfrozen*) dan dilatih ulang, sementara lapisan lainnya dipertahankan untuk menjaga fitur umum yang telah dipelajari sebelumnya. Lapisan klasifikasi akhir disesuaikan dengan jumlah kelas pada penelitian ini, yaitu dua kelas (*Normal* dan *Pneumonia*). Proses pelatihan dilakukan secara iteratif dalam beberapa *epoch* menggunakan data latih, dengan data validasi digunakan untuk memantau potensi *overfitting* serta kestabilan kinerja model.

Model EfficientNet yang digunakan memanfaatkan bobot awal (*pre-trained weights*) dari dataset ImageNet untuk mempercepat proses pembelajaran fitur dasar citra. Lapisan awal (*feature extractor*) dipertahankan, sementara lapisan akhir disesuaikan dengan jumlah kelas pada penelitian ini, yaitu dua kelas. Proses pelatihan dilakukan secara iteratif dalam beberapa *epoch* menggunakan data latih, dengan data validasi digunakan untuk memantau potensi *overfitting*.

3.3.4. Evaluasi Model

Evaluasi kinerja model dilakukan menggunakan data uji yang tidak dilibatkan dalam proses pelatihan. Beberapa metrik evaluasi digunakan untuk memberikan gambaran performa model secara menyeluruh, meliputi *accuracy*, *precision*, *recall*, dan *F1-score*. Selain itu, *confusion matrix* digunakan untuk menganalisis kemampuan model dalam mengklasifikasikan masing-masing kelas.

Untuk mengevaluasi kemampuan diskriminatif model secara lebih mendalam, penelitian ini juga menggunakan *Receiver Operating Characteristic (ROC) curve* dan nilai *Area Under the Curve (AUC)*. Metrik AUC-ROC dipilih karena mampu memberikan evaluasi yang lebih stabil terhadap performa model pada dataset dengan distribusi kelas yang tidak seimbang.

Penggunaan beberapa metrik evaluasi ini bertujuan untuk menghindari bias penilaian kinerja model, terutama karena distribusi kelas pada dataset tidak sepenuhnya seimbang. Hasil evaluasi selanjutnya dianalisis untuk menilai efektivitas model EfficientNet dalam melakukan klasifikasi citra X-Ray dada menjadi kelas *Normal* dan *Pneumonia*.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Dataset dan *Splitting Data*

Dataset yang digunakan dalam penelitian ini berupa citra Chest X-Ray yang terdiri dari dua kelas, yaitu normal dan pneumonia. Dataset ini digunakan untuk membangun model klasifikasi berbasis *deep learning* dengan tujuan mendeteksi penyakit pneumonia secara otomatis berdasarkan citra rontgen dada. Setelah melalui tahap pembersihan data, jumlah total citra yang digunakan dalam penelitian ini adalah 5.856 citra. Dataset kemudian dibagi menjadi tiga bagian, yaitu data latih, data validasi, dan data uji dengan proporsi 80% : 10% : 10%. Pembagian dataset dilakukan menggunakan teknik *stratified split* untuk menjaga proporsi masing-masing kelas tetap seimbang pada setiap subset data.

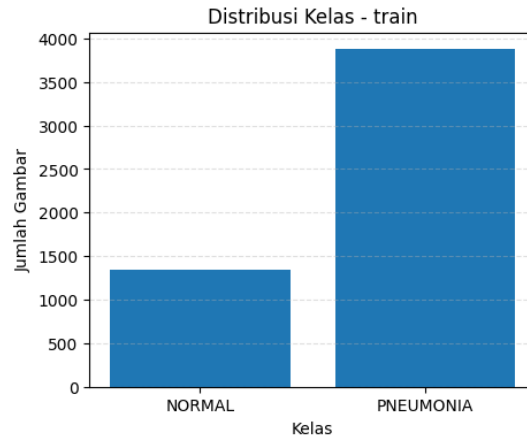
Tabel 4.1 Distirbusi Pembagian Dataset Pneumonia

Subset	Normal	Pneumonia	Total
<i>Training</i>	1.341	3.875	4.684
<i>Validation</i>	158	428	586
<i>Testing</i>	158	428	586

Berdasarkan Tabel tersebut, dapat dilihat bahwa jumlah citra pada kelas pneumonia lebih banyak dibandingkan kelas normal, sehingga dataset bersifat tidak seimbang. Kondisi ini menjadi perhatian penting dalam proses pelatihan model agar tidak terjadi bias terhadap kelas mayoritas.

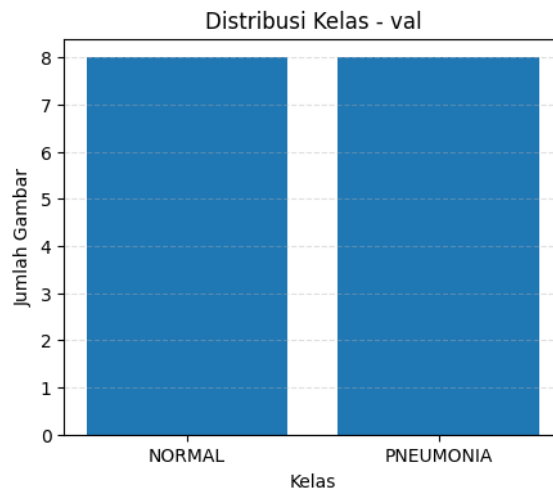
4.2 Distribusi Kelas Dataset

Distribusi kelas pada dataset divisualisasikan menggunakan grafik batang untuk menunjukkan proporsi jumlah citra pada masing-masing kelas, yaitu normal dan pneumonia, pada setiap subset data. Visualisasi ini bertujuan untuk memberikan gambaran awal mengenai keseimbangan data yang digunakan dalam proses pelatihan dan evaluasi model.



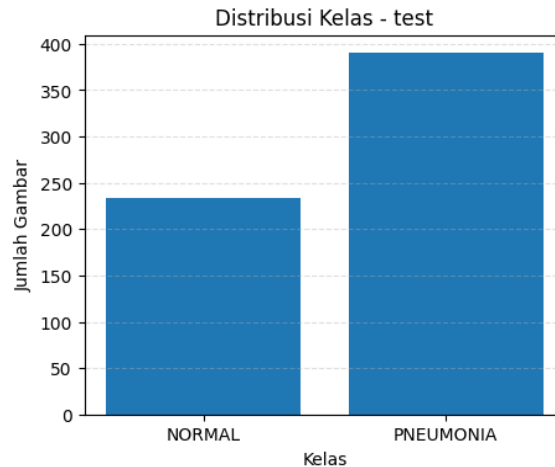
Gambar 4.1 Distribusi *Data Training*

Berdasarkan visualisasi *data training*, terlihat bahwa jumlah citra kelas pneumonia jauh lebih banyak dibandingkan kelas normal. Kondisi ini menunjukkan bahwa data latih bersifat tidak seimbang sehingga berpotensi menyebabkan model lebih cenderung mempelajari pola dari kelas mayoritas apabila tidak ditangani dengan tepat.



Gambar 4.2 Distribusi *Data Validation*

Pada *data validation*, proporsi antara kelas normal dan pneumonia tetap mengikuti distribusi *data training*. Hal ini menunjukkan bahwa proses pembagian data dilakukan secara *stratified*, sehingga distribusi kelas pada data validasi tetap merepresentasikan kondisi data latih dan dapat digunakan secara optimal untuk memantau performa model selama proses pelatihan.



Gambar 4.3 Distribusi *Data Testing*

Distribusi kelas pada data testing juga menunjukkan dominasi kelas pneumonia dibandingkan kelas normal. Keseragaman distribusi antara *data training*, *validation*, dan *testing* menandakan bahwa pembagian dataset telah dilakukan secara konsisten dan representatif, sehingga hasil evaluasi model dapat mencerminkan performa yang sebenarnya.

Ketidakseimbangan jumlah data antar kelas pada ketiga subset tersebut berpotensi mempengaruhi kinerja model, khususnya dalam mendeteksi kelas minoritas. Oleh karena itu, pada proses pelatihan model diterapkan strategi *class weighting*, yaitu pemberian bobot yang lebih besar pada kelas normal. Hal ini bertujuan agar model dapat belajar secara lebih seimbang dan meningkatkan kemampuan dalam mengklasifikasikan kedua kelas.

4.3 *Preprocessing*

Preprocessing merupakan tahap persiapan data sebelum digunakan untuk melatih model. Dalam proyek klasifikasi pneumonia ini, *preprocessing* dimulai dengan mengorganisir dataset ke dalam struktur folder yang sesuai. Dataset awal yang terdiri dari folder NORMAL dan PNEUMONIA dibagi menjadi tiga subset, yaitu *train*, *validation*, dan *test* dengan proporsi 80/10/10. Proses ini menggunakan *stratified splitting* untuk memastikan distribusi kelas yang seimbang di setiap subset. Dari total 5.856 *file* gambar, terdapat 1.583 gambar normal dan 4.273 gambar pneumonia yang kemudian didistribusikan ke folder-folder tersebut. Setelah pembagian data, dilakukan *handling* untuk menangani *imbalance class* menggunakan *class weight* yang dihitung dengan metode *balanced*. *Class weight* ini memberikan bobot lebih tinggi pada kelas minoritas (NORMAL) untuk mengkompensasi ketidakseimbangan jumlah data. Tahap *preprocessing* juga mencakup konversi gambar dari *grayscale* (1 *channel*) menjadi RGB (3 *channel*) karena model yang akan digunakan memerlukan *input 3 channel*. Gambar-gambar kemudian dilakukan *resize* menjadi ukuran standar 600x600 *pixel* dan dilakukan normalisasi nilai *pixel*. *Batch size* dikurangi menjadi 4 untuk mencegah *error Out of Memory* saat *training*. Semua konfigurasi ini disiapkan dalam variabel-variabel yang akan digunakan pada tahap

training model, termasuk pengaturan *Seed* untuk *reproducibility* dan penggunaan *Autotune* untuk optimasi performa.

4.4 Arsitektur Model EfficientNet

Model yang digunakan dalam penelitian ini adalah EfficientNet-B0 dengan pendekatan *transfer learning*. Model memanfaatkan bobot awal hasil *pre-training* pada dataset ImageNet sehingga mampu mengekstraksi fitur visual secara lebih optimal pada citra Chest X-Ray. Arsitektur model terdiri dari *backbone* EfficientNet-B0 yang diikuti oleh lapisan *Global Average Pooling*, *Dropout*, dan satu lapisan *Dense* dengan fungsi aktivasi sigmoid untuk melakukan klasifikasi biner antara kelas normal dan pneumonia. Struktur ini dirancang untuk menghasilkan representasi fitur yang ringkas sekaligus mengurangi risiko *overfitting*.

Layer (type)	Output Shape	Param #
input_layer_5 (InputLayer)	(None, 224, 224, 3)	0
efficientnetb0 (Functional)	(None, 7, 7, 1280)	4,049,571
global_average_pooling2d_2 (GlobalAveragePooling2D)	(None, 1280)	0
dropout_2 (Dropout)	(None, 1280)	0
dense_3 (Dense)	(None, 1)	1,281

Total params: 4,050,852 (15.45 MB)
Trainable params: 1,281 (5.00 KB)
Non-trainable params: 4,049,571 (15.45 MB)

Gambar 4.4 Model EfficientNetB0

Proses pelatihan dilakukan kedalam dua tahap, yaitu pada tahap awal, *backbone* EfficientNet-B0 dibekukan sehingga hanya lapisan *classifier* yang dilatih. Kondisi ini ditunjukkan dengan jumlah parameter *trainable* sebanyak 1.281 dan parameter *non-trainable* sebanyak 4.049.571. Selanjutnya, dilakukan tahap *fine-tuning* dengan membuka beberapa lapisan terakhir *backbone* untuk menyesuaikan fitur model dengan karakteristik citra Chest X-Ray. Model dilatih menggunakan fungsi *loss* Binary Crossentropy dan *optimizer* Adam, yang umum digunakan pada permasalahan klasifikasi biner. Pendekatan ini memungkinkan proses pelatihan berjalan secara stabil dan menghasilkan performa model yang optimal.

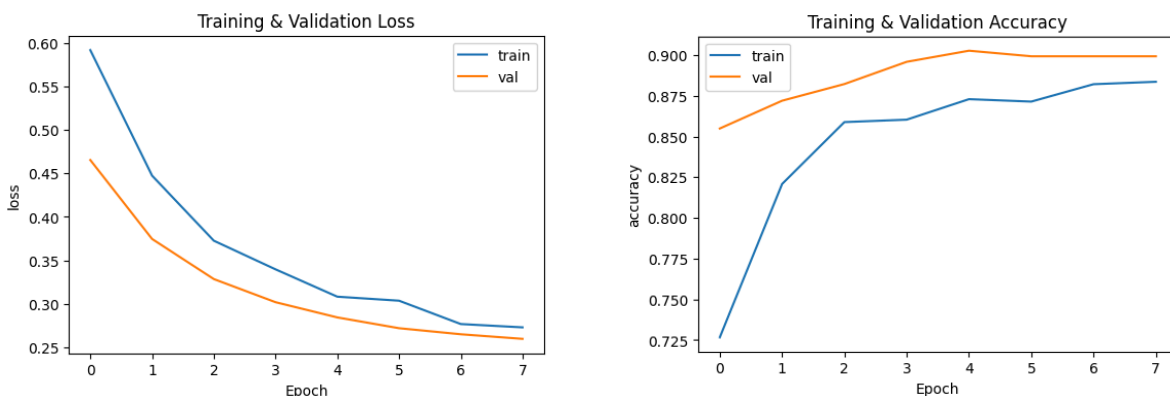
4.5 Fine-Tuning Model

Dalam proyek ini, akan digunakan model *pre-trained* EfficientNetB0 yang sudah dilatih pada dataset besar seperti ImageNet. Proses *fine tuning* dimulai dengan *head training*, yaitu hanya *layer* klasifikasi yang dilatih dengan *base model* yang sepenuhnya di *freeze* (*training* = *False*). Pada tahap ini, model dilatih menggunakan *callbacks* yang terdiri dari *ModelCheckpoint* untuk menyimpan bobot terbaik berdasarkan *validation loss* dengan parameter ‘*save_best_only=True*’ dan ‘*verbose=1*’, *training* dilakukan selama jumlah *epoch* yang

ditentukan oleh 'EPOCHS_HEAD' dengan menggunakan *class_weight* untuk menangani *imbalance*. Tahap kedua adalah *fine tuning* 30 *layer* terakhir dari *base model* yang di *unfreeze* untuk *training* ulang dengan *learning rate* yang jauh lebih rendah ($1e-5$) untuk adaptasi *fine-grained* terhadap dataset X-ray paru-paru. Proses *unfreeze* dilakukan dengan mengatur 'layer.trainable = False' untuk semua *layer* hingga *index* tertentu, kemudian 'layer.trainable = True' untuk *layers* setelahnya. Model kemudian dilakukan *compile* ulang dengan *learning rate* yang lebih kecil dan dilatih kembali selama 'EPOCHS_FINE' *epoch*. Strategi *two-stage training* ini memungkinkan model untuk pertama kali belajar klasifikasi dasar dengan cepat, kemudian melakukan penyesuaian halus pada *feature extractor* untuk karakteristik spesifik citra X-ray Pneumonia.

4.6 Pelatihan Model

Proses pelatihan model dilakukan menggunakan konfigurasi *batch size* 16 dan *image size* 224x224 *pixel*. *Training* model menggunakan *learning rate* $1e-4$ untuk melatih *layer* klasifikasi *custom*, diikuti dengan *learning rate* yang lebih rendah yaitu $1e-5$ untuk melatih 30 *layer* terakhir dari *base model*. Hasil pelatihan model menunjukkan proses *training* yang berjalan selama 8 *epoch* dengan performa dengan performa yang terus meningkat. Detail lebih lanjut divisualisasikan dalam plot *loss* dan *accuracy* di bawah ini



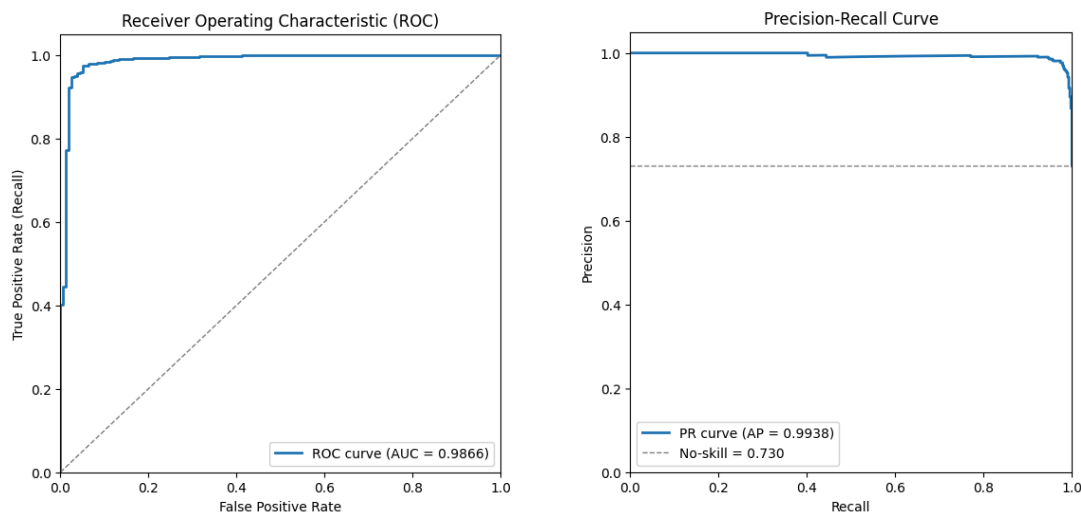
Gambar 4.5 Plot Loss dan Accuracy EfficientNetB0

Gambar plot *loss* dan *accuracy* mencapai *training accuracy* 93,79%, AUC 98,80%, *loss* 0,1541, *precision* 98,89% dan 97,89%, dan *recall* 90,78%. Model berhasil menyimpan *checkpoint* terbaik berdasarkan *improvement* pada *validation loss* di berbagai *epoch* dengan *file output* 'best_effb0.h5'. *Callback* ReduceLROnPlateau tidak terpicu karena *validation loss* terus mengalami perbaikan, menunjukkan bahwa *learning rate* awal sudah optimal. Tren metrik menunjukkan konsistem pada *accuracy* dari *epoch* 1 (82,95%) hingga *epoch* 8 (93,79%) pada *training set*, dan dari 87,45% hingga 92,66% pada *validation set*. Nilai AUC yang tinggi (>95%) di semua *epoch* mengindikasikan bahwa model memiliki kemampuan diskriminasi yang sangat baik antara kelas NORMAL dan PNEUMONIA. Perbedaan kecil antara *training* dan *validation metrics* menunjukkan bahwa model tidak mengalami *overfitting* yang signifikan,

mengindikasikan bahwa regularisasi dengan *dropout* dan strategi *fine tuning* bekerja dengan efektif.

4.7 Evaluasi Model

Evaluasi model dilakukan dengan memuat model terbaik yang telah disimpan selama proses pelatihan dan mengujinya pada 586 citra data uji yang tidak dilibatkan dalam proses training. Evaluasi ini bertujuan untuk menilai kemampuan generalisasi model EfficientNetB0 dalam mengklasifikasikan citra X-ray dada ke dalam kelas Normal dan Pneumonia.



Gambar 4.6 Plot ROC dan AUC Score

Hasil evaluasi menunjukkan bahwa model memiliki kemampuan diskriminasi yang sangat baik dengan nilai ROC-AUC sebesar 0,9866 (98,66%), yang mengindikasikan bahwa model mampu membedakan kedua kelas secara efektif. Kurva ROC yang dihasilkan menunjukkan karakteristik yang mendekati kondisi ideal, yaitu True Positive Rate yang tinggi dengan False Positive Rate yang rendah. Selain itu, evaluasi menggunakan kurva Precision, dan Recall menghasilkan Average Precision sebesar 99,38%. Hal ini menunjukkan bahwa model mampu mempertahankan tingkat precision yang tinggi pada berbagai nilai recall, khususnya pada kondisi dataset yang tidak seimbang.

Berdasarkan evaluasi ROC dan AUC score, model EfficientNetB0 menunjukkan performa yang stabil dan baik. Nilai ROC-AUC dan Average Precision yang tinggi mengindikasikan bahwa model tidak hanya akurat secara keseluruhan, tetapi juga memiliki kemampuan generalisasi yang baik dalam membedakan citra paru-paru normal dan pneumonia. Analisis performa secara lebih rinci pada masing-masing kelas, termasuk precision, recall, F1-score, serta pola kesalahan prediksi, dibahas lebih lanjut pada bagian Classification Report dan Confusion Matrix.

4.7.1 Classification Report

Classification report menunjukkan bahwa model memiliki kinerja yang baik pada kedua kelas, yaitu Normal dan Pneumonia. Evaluasi ini mencakup metrik *precision*, *recall*, *F1-score*, dan *accuracy* untuk masing-masing kelas, yang selanjutnya disajikan secara lebih rinci melalui visualisasi *classification report* pada gambar di bawah ini.

```

=== Classification report (test) ===

```

	precision	recall	f1-score	support
NORMAL	0.93	0.94	0.93	158
PNEUMONIA	0.98	0.97	0.98	428
accuracy			0.96	586
macro avg	0.95	0.96	0.95	586
weighted avg	0.96	0.96	0.96	586

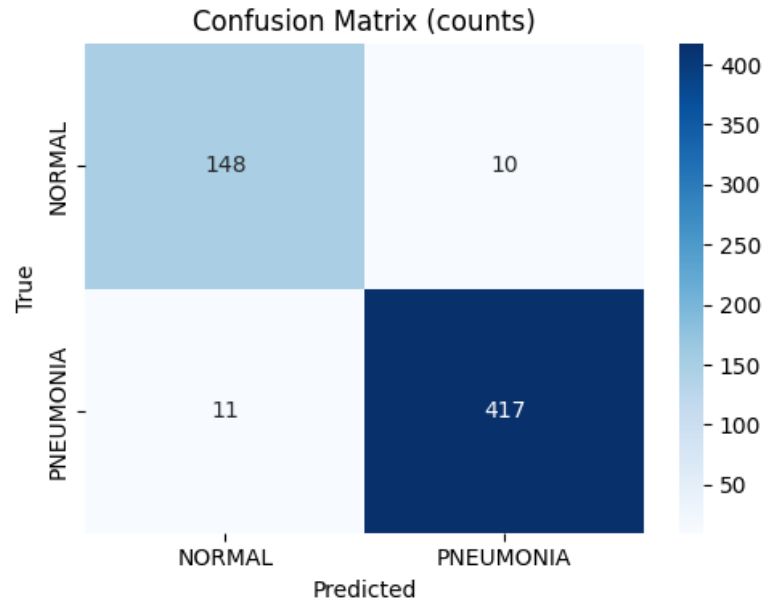
Gambar 4.7 Evaluasi *Classification Report*

Classification report menunjukkan bahwa model EfficientNetB0 memiliki kinerja yang sangat baik dan seimbang pada kedua kelas, yaitu Normal dan Pneumonia. Untuk kelas Normal, model mencapai *precision* 93%, *recall* 94%, dan *F1-score* 93%, yang menandakan kemampuan model dalam mengenali citra paru-paru normal secara akurat. Sementara itu, pada kelas Pneumonia, model menunjukkan performa yang lebih tinggi dengan *precision* 98%, *recall* 97%, dan *F1-score* 98%, yang sangat penting dalam konteks medis untuk meminimalkan kasus pneumonia yang tidak terdeteksi.

Secara keseluruhan, model memperoleh akurasi sebesar 96% pada data uji. Nilai *macro average* dan *weighted average* yang hampir sama *precision*, *recall*, dan *F1-score* sekitar 95-96% menunjukkan bahwa meskipun dataset bersifat tidak seimbang, model tetap mampu mempertahankan performa yang konsisten tanpa bias berlebihan terhadap kelas mayoritas. Hasil ini menunjukkan bahwa penerapan *class weighting* dan *stratified splitting* efektif dalam mengatasi permasalahan ketidakseimbangan data.

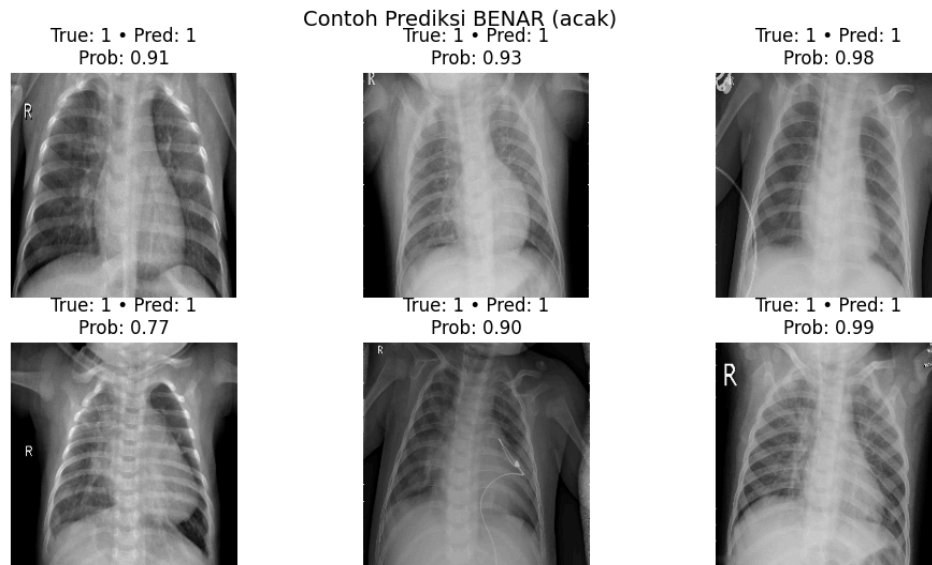
4.7.2 *Confusion Matrix*

Confusion matrix digunakan untuk menampilkan distribusi hasil prediksi model, yang terdiri atas *True Negative* (TN), *False Negative* (FN), *True Positive* (TP), dan *False Positive* (FP) pada setiap label kelas. Melalui visualisasi ini, dapat dianalisis kesalahan prediksi model, khususnya pada dua kelas penyakit yang mengalami salah klasifikasi, sehingga memberikan gambaran performa model dalam membedakan masing-masing label.



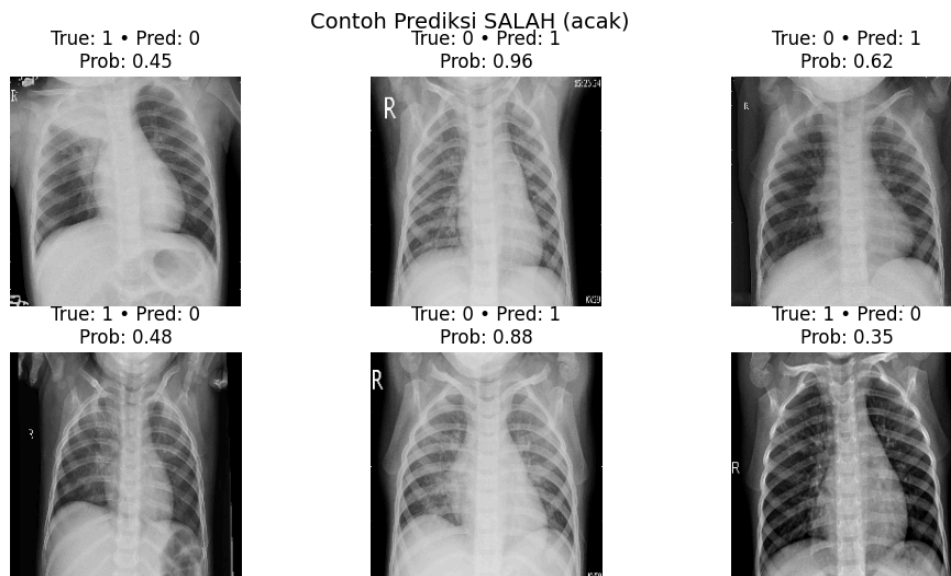
Gambar 4.8 Evaluasi *Confusion Matrix*

Berdasarkan hasil evaluasi *confusion matrix*, model berhasil mengklasifikasikan 148 citra berlabel normal dan 417 pneumonia dengan benar. Sementara itu, terdapat 10 citra label normal yang salah diprediksi sebagai pneumonia dan 11 citra pneumonia yang salah diprediksi sebagai normal. Secara keseluruhan, model hanya menghasilkan 21 kesalahan dari 586 data uji, atau sekitar 3,6% error rate, yang menunjukkan tingkat akurasi klasifikasi yang tinggi. Nilai *recall* per kelas juga menunjukkan performa yang konsisten, yaitu 94% untuk kelas normal dan 97% untuk kelas pneumonia. Hasil evaluasi menandakan bahwa sebagian besar kasus dari kedua kelas berhasil dikenali dengan benar. Dominasi nilai pada diagonal *confusion matrix* mengindikasikan bahwa model memiliki kemampuan klasifikasi yang baik, dengan jumlah kesalahan klasifikasi yang relatif kecil, sehingga model dapat diandalkan dalam membedakan citra paru-paru normal dan pneumonia. Beberapa evaluasi lanjutan juga dilakukan dengan menampilkan label aktual dan label hasil prediksi pada sejumlah citra X-ray secara acak, disertai dengan nilai probabilitas prediksi yang dihasilkan oleh model. Visualisasi ini bertujuan untuk memberikan gambaran kualitatif mengenai tingkat keyakinan model serta pola kesalahan yang masih terjadi.



Gambar 4.9 Citra Prediksi Benar oleh Model

Citra yang diprediksi benar menunjukkan bahwa model EfficientNetB0 mampu mengenali pola radiologis pneumonia secara konsisten dengan tingkat keyakinan yang tinggi. Pada sebagian besar contoh prediksi benar, nilai probabilitas berada pada rentang 0,77 hingga 0,99, yang mengindikasikan bahwa model memberikan keputusan klasifikasi dengan *confidence* yang kuat. Hal ini mencerminkan kemampuan model dalam mengekstraksi fitur penting seperti opasitas, infiltrasi, dan konsolidasi paru yang menjadi ciri khas pneumonia, sehingga prediksi yang dihasilkan selaras dengan label aktual pada data uji.



Gambar 4.10 Citra Prediksi Salah oleh Model

Citra yang diprediksi salah umumnya memiliki nilai probabilitas yang lebih rendah hingga sedang, yaitu berkisar antara 0,35 hingga 0,62, yang menunjukkan adanya ketidakpastian dalam proses klasifikasi. Kesalahan ini dapat disebabkan oleh karakteristik citra yang bersifat ambigu, seperti manifestasi pneumonia yang ringan atau atipikal, serta adanya artefak atau struktur anatomi normal yang menyerupai pola pneumonia. Temuan ini menunjukkan bahwa kesalahan prediksi lebih sering terjadi pada kasus borderline, sehingga hasil prediksi model pada kondisi tersebut masih memerlukan verifikasi tambahan dan tidak disarankan untuk digunakan sebagai dasar keputusan klinis tunggal.

BAB V

Kesimpulan

Penelitian ini mengembangkan model klasifikasi pneumonia berbasis citra X-ray dada menggunakan arsitektur EfficientNetB0 dengan pendekatan transfer learning. Melalui dua tahapan pelatihan, yaitu *training head* dan *fine-tuning*, serta penerapan strategi *class weighting* untuk menangani ketidakseimbangan data, model mampu mencapai kinerja yang sangat baik. Hasil evaluasi pada data uji menunjukkan bahwa model memperoleh akurasi sebesar 96%, dengan nilai ROC-AUC mencapai 98,66%, yang menandakan kemampuan diskriminatif yang sangat kuat dalam membedakan citra paru-paru normal dan pneumonia. Selain itu, model mencatat *precision* sebesar 98% dan *recall* sebesar 97% pada kelas Pneumonia, yang sangat penting dalam konteks medis untuk meminimalkan risiko kesalahan diagnosis.

Berdasarkan analisis *classification report* dan *confusion matrix*, model menunjukkan performa yang seimbang pada kedua kelas, dengan *F1-score* sebesar 93% untuk kelas Normal dan 98% untuk kelas Pneumonia, serta tingkat kesalahan prediksi yang rendah, yaitu sekitar 3,6% dari total data uji. Hasil ini menunjukkan bahwa pelatihan dua tahap dan penanganan data tidak seimbang telah berjalan secara efektif. Namun, penelitian ini masih memiliki keterbatasan, khususnya pada penggunaan dataset publik dengan variasi karakteristik citra yang terbatas. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset klinis yang lebih beragam, mengintegrasikan teknik augmentasi lanjutan, serta mengeksplorasi arsitektur atau *attention mechanism* tambahan untuk meningkatkan robustnes dan generalisasi model pada penerapan nyata.

REFERENSI

- [1] 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>
- [2] Ibrahim, A.U., Ozsoz, M., Serte, S. *et al.* Pneumonia Classification Using Deep Learning from Chest X-ray Images During COVID-19. *Cogn Comput* 16, 1589–1601 (2024). <https://doi.org/10.1007/s12559-020-09787-5>
- [3] M. Tan and Q. V. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, Proc. Mach. Learn. Res., vol. 97, pp. 6105–6114, 2019.
- [4] M. Ucan, B. Kaya, O. Aygun, M. Kaya, and R. Alhajj, “Comparison of EfficientNet CNN models for multi-label chest X-ray disease diagnosis,” *PeerJ Comput. Sci.*, vol. 11, p. e2968, Jul. 2025, doi: 10.7717/peerj-cs.2968.
- [5] R. Aprilia, F. . Faisal, Irwandi, Suharni, and Efriza, “Tinjauan Literatur: Faktor Risiko dan Epidemiologi Pneumonia pada Balita”, *Scientificj*, vol. 3, no. 3, pp. 166–173, May 2024.
- [6] N. T. Ujianto, Gunawan, H. Fadilah, A. P. Fanti, A. D. Saputra, and I. G. Ramadhan, “Penerapan Algoritma K-Nearest Neighbors (KKN) untuk Klasifikasi Citra Medis,” *IT-Explore: Jurnal Penerapan Teknologi Informasi dan Komunikasi*, vol. 2, no. 02, pp. 33–43, 2023.
- [7] R. Sari, A. E. Minarno, and Y. Azhar, “Implementasi Jaringan CNN-LSTM untuk Deteksi Citra X-Ray Dada COVID-19,” *Jurnal Repositor*, vol. 4, no. 4, 2022.
- [8] M. Munadi, “Penerapan Algoritma Deep Learning untuk Deteksi Dini Penyakit dari Citra Medis,” *Journal of Computer Science and Information Technology*, vol. 1, no. 2, pp. 53–59, Aug. 2025, doi: 10.70716/jocsit.v1i2.259.
- [9] S. H. Abdullah, R. Magdalena, and R. Y. N. Fu’adah, “Klasifikasi diabetic retinopathy berbasis pengolahan citra fundus dan deep learning,” *Journal of Electrical and System Control Engineering*, vol. 5, no. 2, pp. 84–90, 2022.
- [10] L. Nurlaela, Y. Suhanda, A. Sopian, C. S. Dewi, and R. Syahrial, “PENGEMBANGAN FRAMEWORK DATA MINING BERBASIS DEEP NEURAL NETWORK DENGAN EKSPLORASI TEKNIK TRANSFER LEARNING UNTUK PREDIKSI DAN KLASIFIKASI DATA,” *Jurnal Rekayasa Informasi Swadharma (JRIS)*, vol. 05, no. 01, pp. 132–141, Jan. 2025.
- [11] U. Atila, M. Ucar, K. Akyol, dan E. UCar, “Plant Leaf Disease Classification Using EfficientNet Deep Learning Model,” *Ecol Inform*, vol. 61, Mar. 2021.
- [12] A. C. Milano, “KLASIFIKASI PENYAKIT DAUN PADI MENGGUNAKAN MODEL DEEP LEARNING EFFICIENTNET-B6,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 1, Jan. 2024, doi: 10.23960/jitet.v12i1.3855.
- [13] A. Howard *et al.*, “Searching for MobileNetV3,” in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, Oct. 2019, pp. 1314–1324. doi: 10.1109/ICCV.2019.00140.
- [14] A. N. Fajrina, Z. H. Pradana, S. I. Purnama, and S. Romadhona, “Penerapan Arsitektur EfficientNet-B0 Pada Klasifikasi Leukimia Tipe Acute Lymphoblastik Leukimia,” *Jurnal Riset Rekayasa Elektro*, vol. 6, no. 1, p. 59, June 2024, doi: 10.30595/jrre.v6i1.22090.

