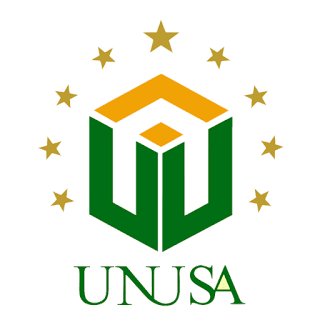
****

**TUGAS AKHIR**

**PERBANDINGAN MODEL ARSITEKTUR DENSENET201 DAN XCEPTION DALAM KLASIFIKASI CITRA X-RAY PARU-PARU**

**ABDUR ROCHMAN**

**3130021021**

**DOSEN PEMBIMBING:**

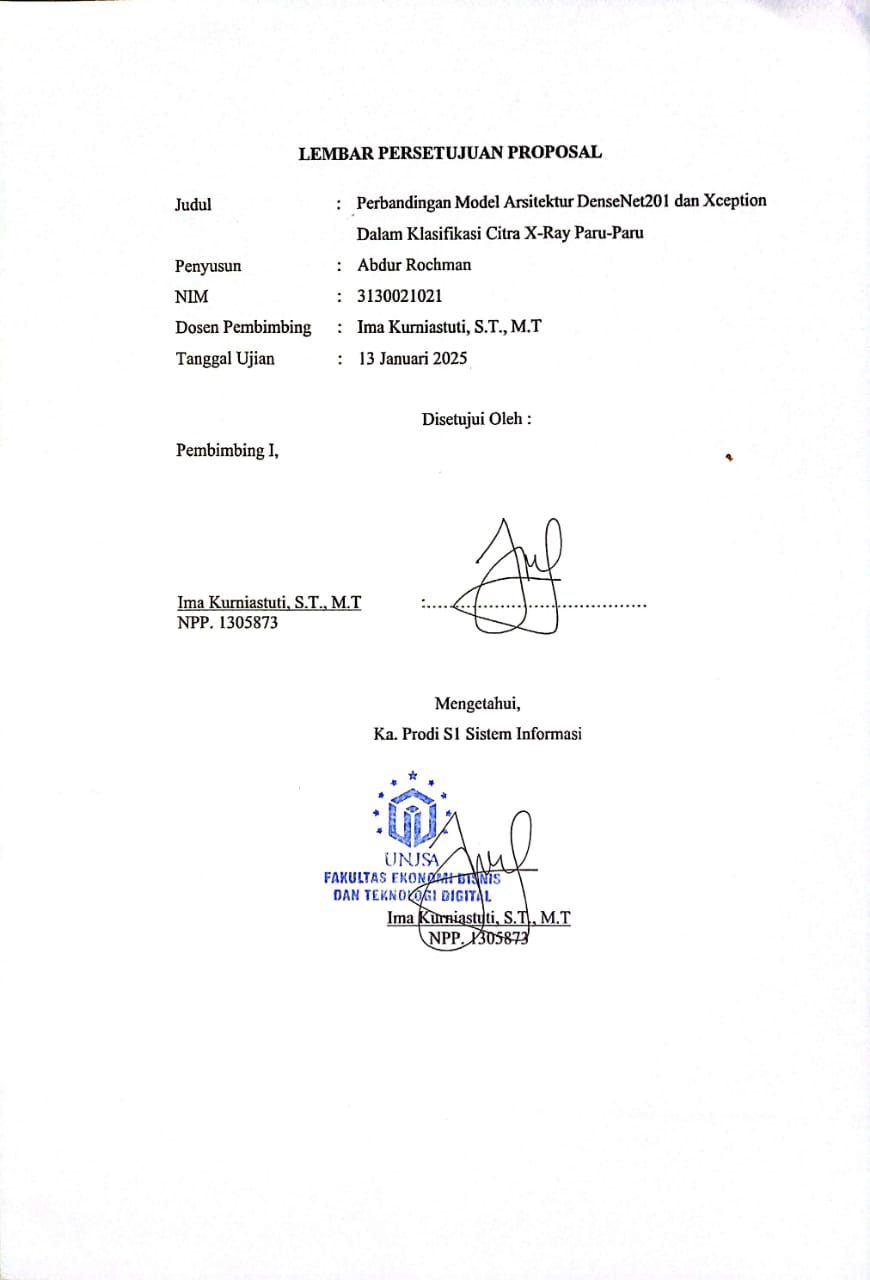
**IMA KURNIASTUTI, S.T., M.T.**

**PROGRAM STUDI S1 SISTEM INFORMASI**

**FAKULTAS EKONOMI BISNIS DAN TEKNOLOGI DIGITAL**

**UNIVERSITAS NAHDLATUL ULAMA SURABAYA**

**2025**

****

***PERBANDINGAN MODEL ARSITEKTUR DENSENET201 DAN XCEPTION DALAM KLASIFIKASI CITRA X-RAY PARU-PARU***

**Nama Mahasiswa : Abdur Rochman**

**Nim : 3130021021**

**Prodi : Sistem Informasi**

**Pembimbing : Imar Kurniatiastuti, S.T., M.T.**

# ABSTRAK

*Pneumonia adalah kondisi ketika alveoli paru-paru terisi cairan atau nanah, menyebabkan penderitanya mengalami kesulitan bernapas. Tuberculosis (TBC) merupakan penyakit menular akibat infeksi bakteri Mycobacterium tuberculosis yang menyerang paru-paru dan merusak jaringannya. Kedua penyakit ini dapat dideteksi menggunakan citra X-Ray paru-paru, yaitu gambar radiologi yang menunjukkan kondisi paru pasien. Dalam penelitian ini, dilakukan klasifikasi citra Chest X-Ray menggunakan metode deep learning, khususnya Convolutional Neural Network (CNN). Deep learning merupakan cabang dari machine learning yang menggunakan jaringan saraf tiruan untuk mengekstraksi pola data secara otomatis. Dua model CNN yang digunakan adalah DenseNet201 dan Xception, yang terbukti efektif dalam klasifikasi citra medis karena kemampuannya mengatasi masalah vanishing gradient dan meningkatkan akurasi. Tahapan penelitian terdiri dari persiapan (pengumpulan data, preprocessing, augmentasi), implementasi (pembuatan dan pelatihan model), serta evaluasi (perbandingan performa kedua model). Diharapkan model ini dapat membantu meningkatkan akurasi dalam mendiagnosis pneumonia dan tuberculosis.*

*Kata Kunci – Pneumonia, Tuberculosis, X-Ray, Deep Learning, DenseNet201, Xception*

# KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala berkat, bimbingan, pertolongan dan perlindungannya selama proses penyusunan sehingga dapat menyelesaikan Tugas Akhir yang berjudul “**Perbandingan Model Arsitektur DenseNet201 dan Xception Dalam Klasifikasi Citra X-Ray Paru-Paru**” dengan baik dan tepat pada waktunya.

Penulisan Tugas Akhir ini disusun sebagai syarat dalam menyelesaikan Pendidikan S1 Sistem Informasi di Fakultas Ekonomi Bisnis dan Teknologi Digital Universitas Nahdlatul Ulama Surabaya. Selama penyusunan dan penyelesaian Tugas Akhir ini banyak hal yang dilalui oleh Penulis, dalam prosesnya Tugas Akhir ini tidak akan dapat terselesaikan tanpa bantuan dan dorongan dari berbagai pihak.

Oleh karena itu, penulis ingin mengucapkan terima kasih, kepada:

1. Bapak Prof. Dr. Ir. Achmad Jazidie, M.Eng., selaku Rektor Universitas Nahdlatul Ulama Surabaya.
2. Bapak Dr. Ubaidillah Zuhdi, S.T., M.Eng., M.SM, selaku Dekan Fakultas Ekonomi Bisnis dan Teknologi Digital.
3. Ibu Ima Kurniastuti, S.T., M.T., selaku Kepala Program Studi S1 Sistem Informasi Fakultas Ekonomi Bisnis dan Teknologi Digital Universitas Nahdaltul Ulama Surabaya.
4. Ibu Ima Kurniastuti, S.T., M.T., selaku dosen pembimbing selama perkuliahan dan penyusunan tugas akhir yang selalu memberikan arahan, bimbingan, serta nasehat kepada penulis selama proses pelatihan.
5. Para Dosen penguji Tugas Akhir, yaitu Bapak Firman Yudianto, S.T., M.T., dan Ibu Rizqi Putri Nourma Budiarti, S.T., M.T., yang telah memberikan masukan dan koreksi yang sangat bermanfaat dalam penyusunan Tugas Akhir.
6. Untuk kedua Orang Tua dan Keluarga Saya yang senantiasa mendoakan dan mendukung penulis selama masa perkuliahan.
7. Untuk teman-teman S1 Sistem Inforasi Angkatan 2021, terutama para sahabat di grub “Punten Skripsi”(Galeh Ariya Irwana, Nanda Dwi Cahyo Wibowo, Rio Rizki, Dina Ayu) yang saling memberikan dukungan dan motivasi untuk segera menyelesaikan Tugas Akhir ini.

Penulis menyadari bahwasannya penulisan Tugas Akhir ini jauh dari sempurna, penulis memohon maaf atas keterbatasan ilmu yang dimiliki , untuk itu saran dan kritik yang membangun sangat penulis harapkan. Penulis berharap semoga Tugas Akhir ini dapat bermanfaat bagi para pembaca.

|  |  |  |
| --- | --- | --- |
|  |  | Surabaya, 12 Juni 2025 |
|  |  | Penulis |
|  |  | (Abdur Rochman) |
|  |  |  |

# DAFTAR ISI

# DAFTAR GAMBAR

# DAFTAR TABEL

# DAFTAR LAMPIRAN

# BAB 1 PENDAHULUAN

## **1.1 Latar Belakang**

Pneumonia adalah sebuah kondisi dimana alveoli penderita dipenuhi nanah dan cairan sehingga penderita mengalami kesakitan saat bernafas dan oksigen yang masuk kedalam paru paru kurang (Rindrasari, 2018). Pneumonia termasuk penyakit yang menyebabkan kematian tertinggi didunia khususnya pada anak anak (Elvionita et al., 2023). Berdasarkan data *World Health Organization* (WHO) pada 2017, terdapat 25.481 kasus kematian anak anak karena infeksi pernafasan akut atau 17 persen dari seluruh kematian anak anak. Indonesia berada diperingkat ke-7 didunia sebagai negara yang memiliki kasus pneumonia tertinggi (Pranita & Sumartiningtyas, 2020). Selain pneumonia, penyakit paru-paru lainnya seperti *tuberculosis* (TBC) juga menjadi masalah kesehatan global yang serius. Menurut laporan data WHO pada tahun 2017 tercatat diwilayah asia tenggara 62% kasus penderita TBC (world health organizations, 2017). Indonesia menjadi salah satu negara yang memiliki kasus terbanyak dalam menghadapi penyakit TBC, tercatat pada tahun 2017 sebesar 446.732 kasus yang terjangkit TBC. TBC adalah suatu penyakit menular yang disebabkan oleh bakteri Myobacterium tuberculosis. Bakteri ini menyerang paru-paru yang mengakibatkan kerusakan jaringan paru dan gangguan pernafasan (Kosarkar et al., 2022).

Peran dokter spesialis paru tentu sangat dibutuhkan. Namun jumlah dokter spesialis paru di Indonesia masih terbatas. Berdasarkan data Perhimpunan Dokter Paru Indonesia (PDPI) pada 2020, tercatat berjumlah 1.206 orang. Jumlah ini belum seimbang dengan jumlah penduduk di Indonesia yang saat ini berjumlah sekitar 250 juta jiwa. Idealnya Indonesia memiliki minimal 2.500 dokter spesialis paru, dengan kata lain saat ini kekurangan 1.294 dokter paru (Manafe, 2020). Diagnosis yang dilakukan oleh dokter spesialis paru sering kali bersifat subjektif. Oleh karena itu dibutuhkanlah sistem yang mampu memberikan diagnosis objektif

dan akurat dalam meningkatkan kualitas pelayanan kesehatan, khsusnya dalam menangani penyakit pneumonia dan TBC.

Identifikasi pneumonia biasanya melibatkan test klinis, analisis riwayat medis, dan test pecintraan seperti citra *chest x-ray* atau *computed tomography* (CT) (Huy & Lin, 2023). Berdasarkan wawancara dengan dokter spesialis paru, citra *X-ray* adalah salah satu yang dapat memberikan gambaran kondisi paru-paru, mendeteksi adanya infeksi, peradangan, atau akumulasi cairan yang menjadi ciri khas pneumonia dan TBC. Hasil dari citra medis *X-ray* akan didiagnosis oleh dokter spesialis untuk dilakukan tindakan medis selanjutnya. Diagnosis pneumonia yang akurat menjadi faktor krusial untuk memastikan pengobatan yang efektif, sehingga dapat meningkatkan prognosis pasien (Soewu et al., 2022). Dalam konteks penelitian ini, identifikasi pneumonia juga menjadi dasar pengembangan sistem prediksi berbasis deep learning yang diharapkan dapat mempercepat proses diagnosis dan meningkatkan akurasi deteksi penyakit paru-paru secara otomatis. Hal ini menunjukkan potensi besar teknologi kecerdasan buatan dalam mendukung praktik medis modern.

Dengan memanfaatkan kecerdasan buatan, berbagai penelitian telah dilakukan untuk membantu dalam meningkatkan analisis penyakit pneumonia dan TBC. Salah satu upaya ialah dengan mendeteksi pasien yang terdampak pneumonia dan TBC dengan cepat dan akurat menggunakan metode *deep learning*. *Deep learning* adalah cabang dari *machine learning* yang berfokus pada pengembangan algoritma berbasis jaringan saraf tiruan yang dirancang untuk belajar dan mengekstrasi pola atau representasi data secara otomatis, mirip dengan cara otak manusia bekerja (Heaton, 2018). Metode *deep learning* yang cukup efektif digunakan dalam pengolahan citra adalah *Convolutional Neural Network* (CNN). CNN mampu melakukan ekstraksi fitur dari citra secara lebih mendetail, menyimpan dan menjadikan acuan hasil hasil pembelajaran yang telah dilakukan dengan adanya parameter *weight* dan bias di dalam arsitekturnya (Iswari, 2021). Dalam penelitian ini, akan digunakan model arsitektur CNN dalam mengklasifikasikan pneumonia pada citra *X-Ray* paru-paru. Arsitektur CNN yang digunakan adalah *DenseNet201* dan *Xception*. *DenseNet201* memiliki kelebihan *dense connectivity*, setiap lapisan terhubung langsung dengan lapisan lainnya, sehingga informasi dapat tersalurkan dengan optimal dan ukuran keluaran tetap konsisten diseluruh lapisan. Sedangkan *Xception* memiliki kelebihan *depthwise seperable convolutions*, yang menggantikan konvolusi standar dengan dua tahap yaitu *pointwise convolution* yang kemudian diikuti oleh *deptwhise convolution*, sehingga dapat menghemat jumlah parameter dan meningkatkan efisiensi komputasi sekaligus mempertahankan akurasi.

Salah satu metode yang paling banyak digunakan untuk klasifikasi penyakit adalah *Convutional Neural Network* (CNN). Penelitian dengan menggunakan CNN antara lain penelitian oleh (Kong & Cheng, 2022) “ *Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion*” didapatkan model yang diusulkan mencapai akurasi rata-rata 98.0% untuk klasifikasi biner dan 97.3% untuk klasifikasi tiga kategori dimana menggunakan arsitektur *DenseNet* dan *VGG16*. Kemudia pada penelitian (Upasana et al., 2022) “*An Attention-based Pneumothorax Classification using Modified Xception Model*” mengenai mendeteki otomatis citra *X-ray* pneumotoraks, menghasilkan kesimpulan bahwa model *Xception* memiliki nilai akurasi yang baik dengan memiliki akurasi pelatihan 99.18% dan *validation accuracy* 87.53%.

Banyak pendekatan yang telah dikembangkan dan diterapkan untuk mendeteksi pneumonia, namun masih ada ruang untuk peningkatan kerja. Salah satu usulan untuk meningkatkan akurasi dan efisiensi dalam klasifikasi citra *X-Ray* adalah dengan mengimplementasikan metode *DenseNet201* dan *Xception*. Kedua model ini telah terbukti efektif dalam bidang klasifikasi citra medis karena kemampuannya dalam menangani vanishing gradient dan meningkatkan kemampuan generalisasi. Diharapkan penelitian ini dapat membantu dokter dalam menteksi awal pneumonia.

## **1.2 Perumusan Masalah**

Adapun rumusan masalah yang dapat diangkat dari penelitian ini yaitu: Bagaimana hasil akurasi, presision, sensitivity, dan Fscore dari klasifikasi citra *chest x-ray* paru-paru ?

## **1.3 Batasan Masalah**

Batasan masalah yang digunakan pada penelitian yaitu:

1. Data klasifikasi jenis citra *x-ray* yang digunakan antara lain : *x-ray* normal, *x-ray* pneumonia, dan *x-ray* tbc
2. Penelitian berfokus pada tingkat akurasi, presision, sensitivity, dan Fscore klasifikasi citra *x-ray* normal, *x-ray* pneumonia, dan *x-ray* tbc
3. Model dari arsitektur cnn yang digunakan adalah DenseNet201 dan Xception

## **1.4 Tujuan Penelitian**

Dari rumusan masalah yang dilakukan, maka tujuan dari penelitian ini yaitu : Untuk mengetahui hasil akurasi, presision, sensitivity, dan Fscore dari klasifikasi citra *x-ray* paru-paru

## **1.5 Manfaat Penelitian**

Manfaat dari penelitian ini yaitu sebagai berikut:

1. Bagi Masyarakat

Hasil dari penelitian tersebut dapat digunakan untuk meningkatkan akurasi dan efisiensi dalam mendeteksi pneumonia melalui analisis citra *chest X-ray*, sehingga dapat mendukung tenaga medis dalam mendiagnosis penyakit lebih cepat dan tepat. Dengan pemanfaatan model dari *deep learning*, sistem dapat membantu mengurangi risiko kesalahan diagnosis serta mempercepat proses pengambilan keputusan. Terutama ditengah keterbatasan jumlah dokter spesialis paru yang sedikit. Hal ini diharapkan dapat berkontribusi pada penanganan yang lebih dini dan efektif, sehingga meningkatkan peluang kesembuhan pasien dan secara keseluruhan mendukung peningkatan kualitas layanan kesehatan Masyarakat.

1. Bagi Pembaca

Hasil dari penelitian ini diharapkan dapat menjadi rujukan atau refrensi bagi mahasiswa Sistem Informasi yang akan melakukan penelitian lanjutan dengan topik *deep learning*.

1. Bagi Penulis

Manfaat yang didapatkan oleh Penulis yakni berupa menambah wawasan pada bidang *deep learning* dan oleh data citra gambar. Selain itu juga pengalaman saat proses pengerjaan tugas akhir dimulai dari proses pengambilan data hingga menulis laporannya.

## **1.6 Relevansi**

Relevansi pengerjaan penelitian ini berkaitan dengan beberapa mata kuliah seperti:

1. Sistem Cerdas : dalam mata kuliah Sistem Cerdas, konsep dan penerapan algoritma *deep learning* seperti *DenseNet201* dan *Xception* menjadi dasar dalam mengembangkan model untuk mendeteksi pneumonia.
2. Statitiska : dalam mata kuliah Statistika berkontribusi pada analisis data, seperti pengolahan statistik deskriptif dan inferensial untuk mengevaluasi performa model yang dikembangkan.
3. Data Science : dalam mata kuliah Data Science menyediakan landasan dalam pengelolaan dan pemrosesan dataset citra *chest X-ray*.
4. Visualisasi Informasi : Visualisasi Informasi berperan dalam menyajikan hasil penelitian secara informatif melalui grafik atau diagram yang memudahkan menangkap sebuah informasi oleh pengguna, termasuk tenaga medis.

# BAB 2 TINJAU PUSTAKA

## **2.1 Studi Literatur**

Tabel 2. 1 Tabel Studi Literatur

| No | Judul | Pembahasan | Keterkaitan | Perbedaan |
| --- | --- | --- | --- | --- |
| 1 | *An Alzheimer’s disease classification model using transfer learning Densenet with embedded healthcare decision support system* (Saleh et al., 2023). | Menggabungkan beberapa jenis model *DenseNet* yaitu *DenseNet* 121,169, dan 201. Menggunakan transfer learning untuk meningkatkan performa klasifikasi pada dataset *Magnetic Resonance Imaging* (MRI) yang lebih kecil, serta penggunaan augmentasi data untuk meningkatkan performa model. | Mengunakan *transfer learning*, menggunakan Teknik augmentasi data, dan menggunakan data citra medis. | Data yang digunakan MRI, Penggabungan 3 Jenis *DenseNet* 121,169, dan 201, dan menggunakan *Area Under the Curve* untuk evaluasi model. |
| 2 | *Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion* (Kong & Cheng, 2022). | Menambahkan *global attention machine block* dan *category attention block* untuk mengekstraksi fitur-fitur dalam. Jaringan residual (*ResNet*) digunakan untuk segmentasi informasi gambar. Hasil dari pelatihan model menunjukkan bahwa model yang diusulkan memiliki akurasi yang tinggi. | Penelitian sama-sama klasifikasi penyakit citra *X-Ray* paru-paru. Menggunakan teknologi *deep learning* yaitu *DenseNet* sebagai salah satu arsitektur model dalam citra *X-Ray*. Untuk mengevaluasi performa model menggunakan akurasi sebagai indikator utama keberhasilan. | Penelitian COVID-19 mencakup deteksi beberapa penyakit (COVID-19, pneumonia, normal). Penelitian COVID-19 menggabungkan fitur dari dua arsitektur untuk meningkatkan akurasi. Penelitian COVID-19 memiliki pendekatan yang lebih kompleks dengan penambahan mekanisme attention untuk mengoptimalkan performa model.  Penelitian |
| 3 | *FemurTumorNet: Bone tumor classification in the proximal femur using DenseNet model based on radiographs* (Pan et al., 2023). | Dengan menggunakan Model *DenseNet* dapat mengungguli akurasi diagnosis manusia dalam hal sensitivitas, spesitifitas, akurasi dan skorf1. Hasil dari penelitian menunjukkan bahwa model *DenseNet* memiliki tingkat akurasi yang tinggi. | Objek yang digunakan adalah citra medis, model yang digunakan sama yaitu *DenseNet* untuk klasifikasi citra medis. | Penyakit yang diklasifikasikan dan bagian tubuh yang menjadi objek penelitian berbeda seperti muskuloskeletal. Penelitian *FemurTumorNet* hanya fokus pada DenseNet. |
| 4 | *Optimization and fine-tuning of DenseNet model for classification of COVID-19 cases in medical imaging* (Chauhan et al., 2021). | Dengan menggunakan metode transfer learning, untuk meningkatkan performa model dengan cara bobot yang telah data dilatih sebelumnya. Teknik *fine-tuning* termasuk *early-stopping* digunakan guna untuk meningkatkan akurasi model. Hasil menunjukkan bahwa *optimizer adam* dengan fungsi kerugian *Cross Entropy* dan *scheduler StepLR* memberikan hasil akurasi yang tertinggi. | penelitian menggunakan citra medis *X-Ray* dada. Menggunakan *DenseNet* sebagai model utama dalam *deep learning*. Memanfaatkan transfer learning untuk mengatasi keterbatasan dataset medis. Menggunakan evaluasi akurasi model untuk menentukan keberhasilan klasifikasi. | Penelitian *DenseNet* lebih luas cakupannya seperti (COVID-19, PNEUMONIA dan NORMAL). Hanya fokus pada *DenseNet* tanpa memperbandingkan. *Fine-tuning* menjadi bagian untuk meningkatkan performa model *deep learning*. |
| 5 | *AI diagnostics in bone oncology for predicting bone metastasis in lung cancer patients using DenseNet-264 deep learning model and radiomics* (Zeng et al., 2024). | Menggunakan Fitur radiomik diekstraksi dari gambar CT dan seleksi fitur dilakukan menggunakan Minimum *Redundancy Maximum Relevance* (mRMR) dan *Least Absolute Shrinkage and Selection Operator* (LASSO). Model DenseNet-264 menunjukkan kinerja yang unggul. | Menggunakan model *DenseNet*. Menggunakan Transfer learning. Evaluasi model pada menggunakan *accuracy* untuk mengukur performa klasifikasi. | Object yang digunakan kanker metastasis tulang. Penelitian *Bone Oncology* menggunakan data CT dengan fitur radiomic. Model yang digunakan *DenseNet-264*. Lebih menonjolkan *evaluasi Under the Curve* (AUC) untuk mengukur performa. |
| 6 | *An in-depth analysis of Convolutional Neural Network architectures with transfer learning for skin disease diagnosis* (Sadik et al., 2023). | Menggunakan metode *transfer learning* untuk menemukan lebih banyak fitur. Mengevaluasi kinerja pendekatan yang diusulkan dengan beberapa arsitektur CNN popular diantara nya, *ResNet50*, *InceptioV3*, *Inception-ResNet* dan *DenseNet*. | Menggunakan data citra medis. Menggunakan *transfer learning* sebagai metode untuk meningkatkan kinerja model. Menggunakan model *DenseNet* dan *Xception*. Menggunakan augmentasi data. Mengevaluasi kinerja arsitektur model CNN untuk menentukan model yang paling efektif. | Menggunakan citra dermatoskopi. Penelitian kulit mencakup evaluasi lebih banyak arsitektur CNN diantara nya *ResNet50*, *InceptioV3*, *Inception-ResNet* dan *DenseNet*. |
| 7 | *An Attention-based Pneumothorax Classification using Modified Xception Model* (Upasana et al., 2022). | Penggunaan radiografi dada digunakan sebagai alat diagnosis yang signifikan untuk mendeteksi gangguan paru-paru. Menggabungkan jaringan *Xception* dengan attention module untuk meningkatkan akurasi deteksi pneumotoraks pada gambar *X-ray* dada. Model yang diusulkan memperoleh akurasi yang tinggi dan Area AUC yang sangat baik. | Menggunakan citra *X-Ray* dada. Menggunakan arsitektur model *Xception*. Mmenggunakan *transfer learning*. Menggunakan metrik evaluasi untuk mengukur tingkat keberhasilan klasifikasi model. | Berfokus pada pneumotoraks. Memoodifikasi arsitektur seperti menambahkan *attention module*. Penelitian Pneumotoraks berfokus pada pengembangan model. Menggunakan metrik AUC sebagai evaluasi kinerja model. |
| 8 | *Prediction of fetal brain gestational age using multihead attention with Xception* (Hasan et al., 2024). | Penggunaan model *Xception* yang telah dilatih sebelumnya dan mekanisme *multihead attention* (MHA) digunakan untuk memprediksi usia gestasi dari gambar MRI otak janin. Penelitian ini menunjukkan bahwa model yang diusulkan dapat membantu klinisi dalam memprediksi usia gestasi dengan akurasi tinggi. | Menggunakan arsitektur model CNN yaitu *Xception*. Menggunakan *transfer learning*. Menggunakan metrik evaluasi untuk mengukur tingkat keberhasilan klasifikasi model. | Menggunakan dataset MRI otak janin. Memodifikasi model *Xception* dengan MHA. Menggunakan *Mean Absolute Error* (MAE) untuk mengevaluasi hasil regresi. |
| 9 | *Boosted dipper throated optimization algorithm-based Xception neural network for skin cancer diagnosis: An optimal approach* (Tang & Rashid Sheykhahmad, 2024). | Algoritma *Boosted Dipper Throated* (BDTO) adalah teknik optimasi yang terinspirasi dari perilaku burung *Dipper Throated* yang dapat menentukan parameter dan bobot optimal untuk arsitketur CNN *Xception*. Sedangkan dataset International *Skin Imaging Collaboration* (ISIC) digunakan untuk meningkatkan kualitas dan keragaman gambar, serta menggunakan augmentasi data. Metode ini menunjukkan kinerja yang lebih baik dibandingkan dengan pendekatan kontemporer lainnya. | Menggunakan arsitektur model CNN *Xception*. Menggunakan augmentasi data. | Menggunakan dataset ISIC yang terdiri dari citra kulit. Menggunakan optimasi model *Xception* dengan algoritma BDTO. |
| 10 | *Concatenated Xception-ResNet50 — A novel hybrid approach for accurate skin cancer prediction* (Panthakkan et al., 2022). | Menggunakan pendekatan hibrida antara dua model *Xception* dan *ResNet50* model dilatih dan diuji menggunakan teknik *sliding window* untuk memastikan kinerja yang konsisten dan menghindari *overfitting*. Model yang diusulkan mencapai akurasi yang tinggi. | Menggunakan arsitektur model CNN yaitu *Xception*. Menerapkan penghindaran *overvitting* dan augmentasi data. | Jenis penyakit yang diklasifikasikan berbeda kanker kulit. Menggunakan pendekatan hibrida antara dua model *Xception* dan *ResNet50*. |

## **2.2 Dasar Teori**

### 2.2.1 Pneumonia

Pneumonia adalah salah satu jenis penyakit paru paru yang disebabkan oleh bakteri,virus,jamur atau parasit. Pneumonia memiliki sebuah kondisi dimana alveoli penderita dipenuhi nanah dan cairan sehingga penderita mengalami kesakitan saat bernafas dan oksigen yang masuk kedalam paru paru kurang (Rindrasari, 2018). Pasien yang mengidap pneumonia biasanya mengalami gejala berupa sesak napas, batuk berdahak, demam ataupun menggigil. Pneumonia biasanya ditularkan dengan berbagai cara antara lain melalui batuk dan bersin (Sartiwi et al., n.d).

Pneumonia merupakan masalah kesehatan yang signifikan di Indonesia, berdasarkan data Riskesdas (2013) terjadi peningkatan prevalansi pneumonia pada semua umur, terutama pada balita yang cukup tinggi yaitu 4,5 per 100 balita. Hal ini dari 4,5 per 100 balita menderita penyakit pneumonia. Sementara itu, menurut laporan WHO pada tahun 2017 menunjukkan, 15% kematian anak dibawah umur 5 tahun disebabkan oleh pneumonia. Berdasarkan survey Balitbangkes 2016 jumlah anak yang menderita pneumonia di Indonesia diperkirakan melebihi dari 800.000 anak. Pada pasien dewasa, kasus pneumonia paling banyak terjadi pada rentan umur 56-65 tahun. Kondisi ini dipengaruhi oleh perubahan anatomi dan fisiologi akibat proses penuaan, seperti penurunan elesitisitas jaringan paru-paru, melemahnya sistem kekebalan tubuh (Farida et al., 2017). Faktor-faktor tersebut dapat menyebabkan penurunan kemampuan paru-paru untuk terkena infeksi paru-paru, sehingga lansia rentan terhadap penumonia.

Pneumonia juga dapat dipengaruhi oleh berbagai faktor lain seperti kebiasaan merokok, paparan polusi udara, serta kondisi lingkungan yang tidak higienis (Nurin, 2024). Selain itu seseorang dengan penyakit seperti diabetes, penyakit jantung, atau gangguan sistem imun lebih rentan terjangkit pneumonia (CDC, 2023). Pencegahan penyakit ini dapat dilakukan dengan menerapkan gaya hidup sehat seperti menjaga kebersihan tangan, menghindari kontak dengan orang yang terinfeksi, serta menjalani vaksinasi, seperti vaksin pneumokokus dan influenza . Langkah-langkah pencegahan ini sangat penting untuk mengurangi risiko terkena pneumonia, terutama bagi kelompok rentan seperti anak-anak, lansia, dan individu dengan penyakit kronis (Pittara, 2022).

### 2.2.2 *Tubercolosis*

*Tubercolosis* (TBC) adalah suatu penyakit menular yang disebabkan oleh bakteri *Myobacterium tuberculosis*. Bakteri ini biasanya menyerang paru-paru akan tetapi bakteri tuberkolosis dapat menyerang bagian tubuh mana saja seperti ginjal, tulang belakang, dan otak (Mathofani et al., 2020). Penyakit ini dapat menular melalui udara ketika seseorang yang terinfeksi batuk, bersin, atau bahkan berbicara. Gejala TBC aktif meliputi batuk yang berlangsung lebih dari tiga minggu, nyeri dada, atau dahak. Selain itu penderita tuberkolosis sering mengalami penurunan berat badan, kelelahan, demam, dan keringat malam. Deteksi infeksi TBC laten melibatkan penggunaan alat diagnosis seperti tes kulit tuberkulin atau uji pelepasan inferno-gamma. Sedangkan diagnosis tuberkolosis aktif bergantung pada citra x-ray dada dan kultur dahak (Kosarkar et al., 2022).

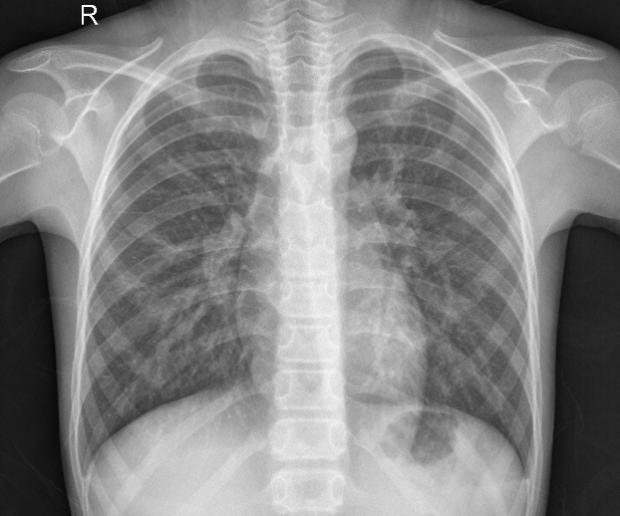
TBC menjadi salah satu penyakit yang terus menimbulkan ancanaman global yang signifikan. Menurut laporan data WHO pada tahun 2017 tercatat diwilayah asia tenggara 62% kasus penderita tuberkolosis (world health organizations, 2017). Indonesia menjadi salah satu negara yang memiliki kasus terbanyak dalam menghadapi penyakit tuberkolosis, tercatat pada tahun 2017 sebesar 446.732 kasus yang terjangkit tuberkolosis.

Kebiasaan merokok merupakan sebuah faktor pemicu yang juga meningkatkan resiko untuk terjangkit penyakit TBC. Resiko terjadinya tuberkolosis paru 17.500 kali lebih besar pada responden dengan kebiasaan merokok dibandingkan dengan responden yang tidak merokok. Selain merokok terdapat beberapa faktor lain yang dapat memicu resiko TBC. Kemiskinan, kepadatan hunian, dan lingkungan hidup yang tidak sehat juga dapat meningkatkan paparan bakteri TBC karena sanitasi yang buruk dan akses pelayanan kesehatan yang terbatas. HIV/AIDS juga merupakan sebuah faktor terjangkitnya TBC karena status imun menjadi melemah sehingga mudah untuk terkena penyakit lainnya (Banu et al., 2017).

### 2.2.2 Chest X-Ray

*Chest X-Ray* (CXR) adalah suatu proyeksi pecintraan medis yang menggunakan radiasi untuk menghasilkan gambar dari organ atau struktur didalam dada. Termasuk seperti paru-paru, jantung, tulang rusuk dan pembuluh darah. *Chest X-Ray* alat yang sangat umum digunakan untuk mendiagnosis berbagai kondisi seperti infeksi paru-paru, gagal jantung, kanker paru-paru dan masalah yang lainnya (*Chest X-Rays – The Heart Clinic*, 2015).

Citra paru-paru dapat dikategorikan normal apabila citra tersebut tidak ada bercak, teksturnya terlihat halus, serta ukuran paru-paru sesuai dengan anatomi yang normal. Sebaliknya, citra paru-paru dianggap abnormal apabila citra tersebut memiliki bercak, teksturnya tidak terlihat halus, serta perubahan bentuk atau ukuran paru-paru berubah. Perubahan tersebut dapat diketahui dari intensi piksel pada objek paru-paru (Wikanargo & Thenata, 2018).

  
Gambar 2. 1 Chest X-Ray Paru-Paru

### 2.2.3 Citra Digital

Citra digital adalah sebuah reprentasi dari suatu objek yang disimpan dalam suatu bentuk elektronik yang dapat diolah oleh komputer. Citra digital digambarkan sebagai fungsi(x,y) dengan x dan y merupakan koordinat pada sebuah bidang datar yang mempresentasikan kumpulan pixel dalam dua dimensi (Gonzalez & Woods, 2008). Secara umum citra digital dibagi menjadi tiga, yaitu citra biner, citra *grayscale* dan citra RGB. Citra biner hanya terdiri dari dua nilai, yaitu hitam dan putih, yang mewakili objek dan latar belakang. Citra *grayscale* terdiri dari nilai intensitas abu-abu yang menggambarkan perbedaan kecerahan, sedangkan citra RGB menggabungkan tiga komponen warna dasar, yaitu merah (Red), hijau (Green), dan biru (Blue), untuk membentuk warna yang lebih kompleks (Sulistiyani Ratna et al., 2016).

Terdapat banyak metode pengolahan gambar yang dapat digunakan, salah satunya adalah augmentasi gambar. Augmentasi merupakan sebuah metode untuk memperbanyak data gambar untuk training dengan membuat perubahan untuk memperoleh gambar training yang serupa tetapi berbeda. Memotong gambar dengan cara yang berbeda dapat memberikan posisi gambar yang berbeda, sehingga dapat mengurangi ketergantungan model pada posisi dimana objek muncul. Teknik augemntasi juga dapat memeberi penyusaian terhadap kecerahan, warna, serta faktor-faktor lain untuk mengurangi sensitivitas model terhadap warna. Melalukan flip ke kiri dan ke kanan pada gambar umumnya tidak mengubah kategori dari objek. Sehingga, metode ini adalah salah satu metode augmentasi gambar yang sering digunakan (Zhang et al., 2021).

### 2.2.4 Grayscale

*Grayscale* adalah format representasi citra yang menggunakan satu chanel untuk setiap piksel dengan nilai kecerahan yang bervariasi dari 0 hingga 255. Nilai ini menggambarkan intensitas Cahaya yaitu, 0 mewakili hitam (tanpa cahaya) dan 255 mewakili putih (intensitas cahaya maksimum) (Gonzalez & Woods, 2002). Dalam aplikasi pemrosesan citra *grayscale* sering digunakan karena informasi warna tidak selalu relevan, terutama pada jenis citra seperti *X-Ray*. Dimana analisis lebih berfokus pada pola intensitas daripada kombinasi warna.

Penggunaan citra *grayscale* memiliki peran penting dalam pengolahan citra medis seperti *X-Ray* paru-paru. Citra *X-Ray* umumnya sudah berupa *grayscale* secara *default*, karena tujuan utama dari citra tersebut adalah untuk menunjukkan detail struktur internal tubuh berdasarkan distribusi intensitas cahaya yang melewati jaringan tubuh. Dengan menggunakan format *grayscale*, analisis fitur seperti pola abnormal, perbedaan densitas jaringan, atau identifikasi anomali menjadi lebih mudah dilakukan (Ramdhan et al., 2014). Selain itu, penggunaan citra grayscale membantu menyederhanakan algoritma pemrosesan, termasuk metode peningkatan kontras seperti *Contrast Limited Adaptive Histogram Equalization* (CLAHE), karena metode ini hanya berfokus pada distribusi intensitas piksel tanpa memperhatikan informasi warna.

### 2.2.5 *Contrast Limited Adaptive Histogram Equalization* (CLAHE)

*Pre-processing* bertujuan untuk meningkatkan kualitas citra, sehingga memudahkan dan mempercepat kinerja sistem dalam mengenali citra *X-Ray* paru-paru. Proses ini sangat penting untuk memastikan bahwa citra yang digunakan oleh model deep learning memiliki kualitas yang optimal. Tanpa *pre-processing* yang tepat sistem dapat kesulitan untuk mengidentifikasi fitur-fitur penting dala citra. Salah satu metode *pre-processing* citra untuk meningkatkan kontras gambar adalah CLAHE.

CLAHE merupakan sebuah metode dengan memberikan nilai batas pada histogram. Nilai batas ini disebut dengan clip limit yang menyatakan batas maksimum tinggi suatu histogram. Dengan menggunakan CLAHE kontras citra dapat diperbaiki terutama pada area dengan kontras rendah. Metode ini sangat berguna untuk citra medis seperti *X-Ray* yang seringkali memiliki detail yang sulit terlihat karena kontras yang rendah. Metode ini bertugas membagi citra menjadi beberapa bagian kecil dan masing-masing bagian dihitung histogramnya secara terpisah. Proses ini dilakukan untuk menghindari efek noise yang berlebihan di area dengan kontras tinggi.

Proces CLAHE terdiri dari dua tahap, yaitu pada tahap pertama membagi citra kedalam beberapa bagian dengan ukuran yang sama dan merata. Pada tahap kedua menghitung nilai histogram masing-masing bagian sehingga mendapatkan nilai *clip limit* dari citra tersebut. Dengan begitu, area yang memiliki kontras rendah dapat ditingkatkan tanpa merusak detail yang ada pada area dengan kontras tinggi. Cara menghitung clip limit suatu histogram dapat didefinisikan dengan persamaan berikut (Koonsanit et al., 2017).

δ = (1 + (-1)) (2.1)

Keterangan:

δ = clip limit

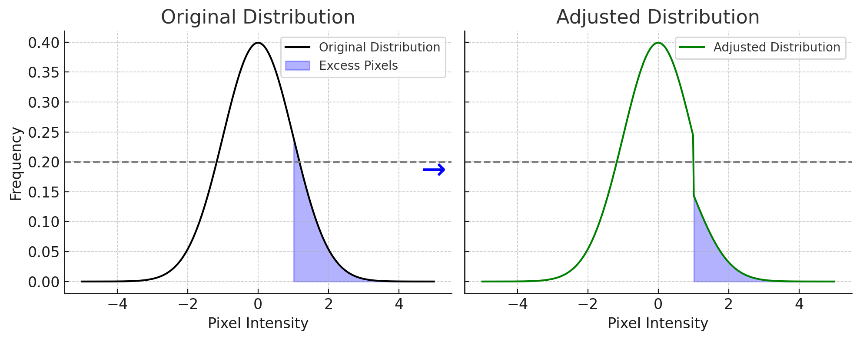
Q×R = luas citra dimensi citra

C = nilai komponen warna

a = clip factor (batas limit suatu histogram)

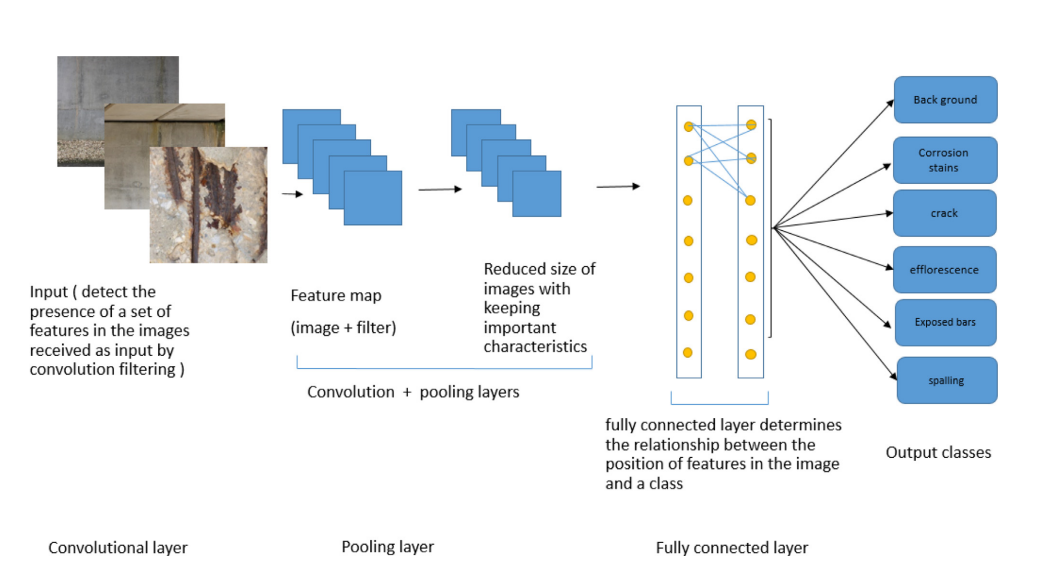
= nilai gradien maksimum

Histogram diatas nilai *clip limit* dianggap kelebihan (*exceess*) piksel yang akan didistribusikan kepada area sekitar di bawah *clip limit*, sehingga histogram merata (). Ilustrasi distribusi *excess* piksel dapat dilihat pada gambar 2.2.

  
Gambar 2. 2 Distribusi excess pixel pada histogram

### 2.2.6 *Convutional Neural Network* (CNN)

CNN adalah salah satu metode dari deeplearning yang digunakan dalam klasifikasi citra gambar. Pada dasarnya, CNN tidak jauh berbeda dengan neural network biasanya yang memiliki bobot, bias dan fungsi aktivasi. CNN mengatur neuron dalam bentuk tiga dimensi yaitu lebar, panjang dan tinggi. Secara umum, arsitektur CNN terdiri dari tiga *layer* utama yaitu *convutional layer*, *pooling layer* dan *fully connected layer (Sewak et al., 2018)*. Berikut merupakan ilustrasi arsitektur metode CNN yang ditunjukkan pada Gambar 2.3.

  
Gambar 2. 3 Arsitektur Convolutional Neural Network (CNN)

1. *Convolutional* *Layer*

Fungsi utama dari *convolution layer* adalah mengekstraksi fitur dari citra input. Proses konvolusi adalah mengalikan sebuah gambar dengan sebuah convolution kernel atau filter yang dinyatakan dalam bentuk matrik dengan ukuran yang biasanya lebih kecil dari ukuran gambar. *Convolutional layer* membutuhkan input yang kemudian mengaplikasikan *convotional kernel* dan memberikan sebuah hasil berupa feature map sebagai output yang diilustrasikan pada Gambar 2.4 (Zhang et al., 2021). Operasi konvolusi apabila input dan kernel berukuran dua dimensi ditulis pada persamaan (2.2).

(2.2)

dimana :

= *feature map* pada pixel ke-a,b pada citra ke*-*

= *bias* pada *feature map*

= bobot pada pada *convolutional kernel* ke-c,d pada citra ke-

*X* = input

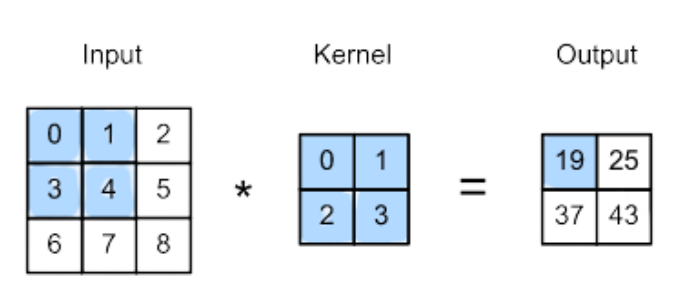
*a* = 1,2,…,*A*. A merupakan panjang pixel pada *feature map*

*b* = 1,2,…,*B*. B merupakan lebar pixel pada *feature map*

*c* = 1,2,…,*C*. C meruapakan panjang pixel pada *convolutional kernel*

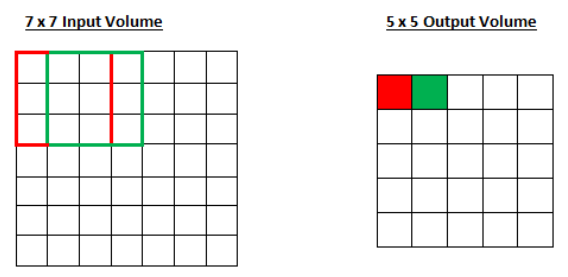
*d* = 1,2,…,*D*. D merupakan lebar pixel pada *convolutional kernel*

= 1,2,…,*L.* L merupakan citra yang digunakan

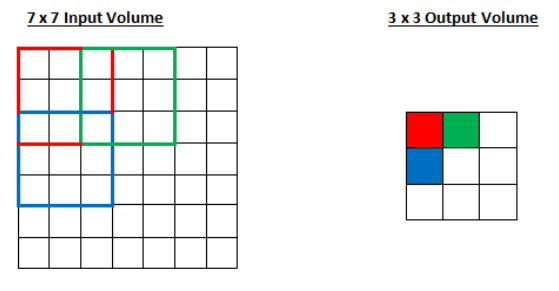
  
Gambar 2. 4 Ilustrasi Convolutional Layer

Dalam perhitungan output, kernel digeser untuk setiap pixel. Demi meningkatkan efisiensi komputasi, kernel digeser lebih dari satu pixel. Banyaknya pixel yang digeser dalam input disebut stride yang ditampilkan pada Gambar 2.5.

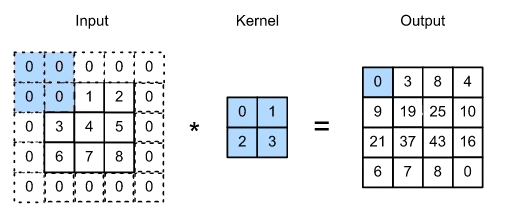
Stride 1



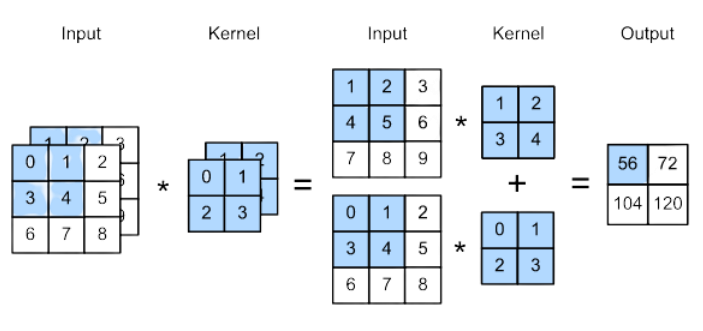
Stride 2

  
Gambar 2. 5 Perbedaan Antara Stride 1 dan Stride 2

Salah satu masalah yang terjadi akibat pengaplikasian *convutional layer* adalah kehilangan informasi *pixel* yang terletak pada bagian tepi gambar. Solusi yang mudah adalah menambah *pixel* tambahan dipinggiran gambar, sehingga memperbesar dimensi gambar. Biasanya, *pixel* tambahan bernilai 0 yang diilustrasikan pada Gambar 2.6.

  
Gambar 2. 6 Ilustrasi Padding Berukuran 1x1

Jika input data mengandung banyak *chanel*, maka perlu dibentuk *kernel* konvolusi yang memiliki *chanel* sebanyak *chanel* input sehingga dapat dilakukan korelasi silang. Apabila diasumsikan jika banyaknya *chanel* input adalah , maka kernel konvolusi harus sebanyak juga. Korelasi silang dihitung dengan menjumlahkan operasi konvolusi untuk setiap *chanel*, sehingga didapatkan output berdimensi dua yang diilustrasikan pada gambar 2.7.

  
Gambar 2. 7 Ilustrasi Padding Berukuran 1x1

Terlepas dari banyaknya input *chanel*, sejauh ini menghasilkan output dengan *chanel* tunggal. Akan tetapi, arsitektur CNN yang popular pada umumnya memiliki output dengan banyak *chanel*. Hal ini dilakukan untuk menyimpan lebih banyak informasi ketika *pooling* dilakukan.

1. Pooling Layer

*Pooling layer* berada setelah convolution layer dan tidak memiliki parameter. *Pooling layer* bersifat deteminstik sehingga fungsi yang bisa digunakan adalah maksimum dan mean untuk mengurangi input data (Zhang et al., 2021). Operasi ini masing-masing dinamakan *max pooling* dan *average pooling*. Fungsi *max pooling* disajikan pada persamaan berikut (Hafemann et al., 2017).

(2.3)

*FM* = *feature map*

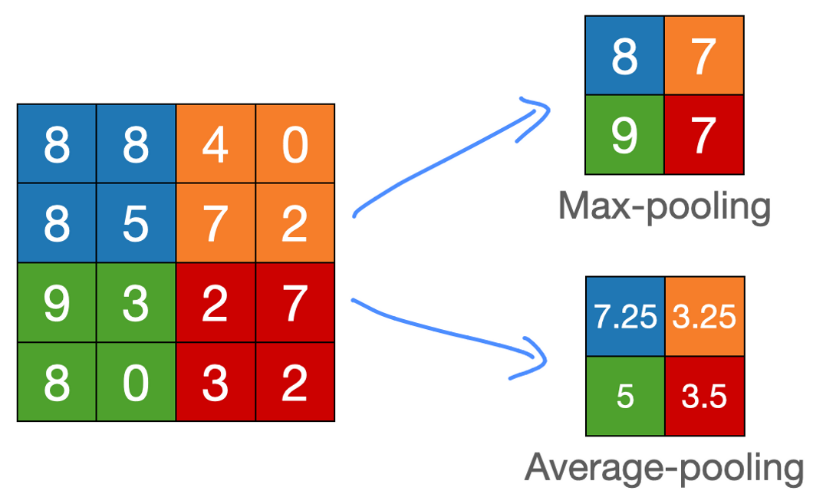
*Pool* = hasil *pooling layer*

*a* = 1,2,…,*A*. A merupakan panjang pixel

*b* = 1,2,…,*B*. B merupakan lebar pixel

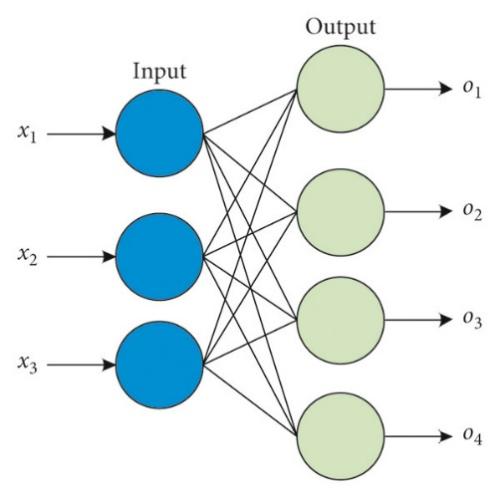
*c* = 1,2,…,*C*. C merupakan citra yang digunakan

Ilustrasi operasi *max pooling* dan *average pooling* yang dapat dilihat pada gambar. *Pooling layer* terdiri dari sebuah filter dengan ukuran dan *stride* tertentu yang akan secara bergantian bergeser pada seluruh area *feature map*.

  
Gambar 2. 8 Pooling Layer

1. *Fully Connected Layer*

*Fully connected layer* merupakan bagian terakhir dari arsitektur CNN yang digunakan untuk melakukan transformasi dimensi data agar dapat diklasifikasikan secara linear. Layer ini mengkoneksasikan semua neuron disatu layer dan kesemua neouron layer lainnya (Zhang et al., 2021). Citra digital yang memiliki dua atau tiga dimensi akan diubah menjadi suatu vector (data berdimensi satu) sebelum masuk ke *fully connected layer* karena input layer ini adalah suatu vektor. Fully connected layer memiliki fungsi aktivasi, *output layer* dan *loss function*. Berikut merupakan ilustrasi *fully connected layer* tanpa menggunakan hidden layer yang akan ditunjukkan pada gambar 2.9.

  
Gambar 2. 9 Ilustrasi Fully Connected Layer

Metode CNN menggunakan fungsi aktivasi pada *convolutional layer* sebelum *pooling layer* adalah fungsi aktivasi ReLu yang tercantum pada persamaa (2.5) Sedangkan fungsi aktivasi yang digunakan pada layer terakhir adalah fungsi aktivasi sigmoid yang tercantum pada persamaan (2.4), jika kelas klasifikasi yang digunakan adalah biner.

### 2.2.7 Fungsi Aktivasi Softmax

Fungsi aktivasi sigmoid merupakan fungsi aktivasi yang digunakan dalam perhitungan probabilitas untuk menentukan klasifikasi multi kelas dengan output kelas yang mempunyai nilai probabilitas yang paling tinggi (Agustina Purwitasari et al., 2022). Fungsi aktivasi softmax dapat mengurangi kemungkinan nilai error yang dihasilkan oleh fungsi objektif *cross-entropy (Rokhana et al., 2019)*. Perhitungan fungsi aktivasi softmax ditunjukkan pada persamaan (2.4) sebagai berikut.

(2.4)

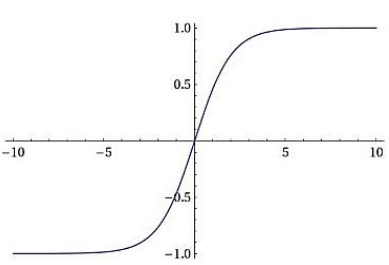
dimana :

= output setiap neuron output layer tanpa fungsi aktivasi

K = jumlah kelas dalam klasifikasi multikelas

p() = output setiap neuron *output layer* setelah dimasukkan dalam fungsi aktivasi

Berikut merupakan Ilustrasi fungsi sigmoid ditunjukkan pada gambar 2.10.

  
Gambar 2. 10 Distribusi Fungsi Sigmoid

### 2.2.8 Fungsi Aktivasi *Rectified Linear Unit* (ReLu)

Fungsi aktivasi ReLu merupakan fungsi aktivasi yang menghilangkan *vanishing gradient* yang cukup populer digunakan (Zhang et al., 2021). Fungsi ReLu digunakan sebagai non-linearitas telah menunjukkan kemungkinan pelatihan CNN tanpa memerlukan unsupervised pre-training. Perhitungan fungsi ReLu ditunjukkan pada persamaan berikut.

(2.5)

dimana :

*FM* = *feature map*

*a* = 1,2,…,*A*. A merupakan panjang *pixel*

*b* = 1,2,…,*B*. B merupakan lebar *pixel*

*l* = 1,2,…,*L*. L merupakan citra yang digunakan

### 2.2.6.3 Loss Function

*Loss function* adalah nilai error antara nilai aktual dengan nilai yang akan diprediksi. *Loss function* biasanya berupa nilai non-negatif dimana jika bernilai lebih kecil akan menghasilkan prediksi yang lebih akurat (Zhang et al., 2021). Terdapat banyak metode untuk mengukur nilai error pada *deep learning*. Pada klasifikasi yang memiliki tiga kelas perhitungan loss yang digunakan adalah *categorical-crossentropy*. Perhitungan *categorical-crossentropy* ditunjukkan pada persamaan berikut.

(2.6)

dimana :

N = jumlah sempel dalam batch

K= jumlah kelas

= nilai target berupa nilai 0 dan 1

=*predicted value* didapatkan dari persamaan (2.4)

### 2.2.9 Optimasi Parameter *Adaptive Moment Estimation* (ADAM)

Optimasi parameter digunakan untuk meminumkan nilai loss sehingga loss merupakan kunci dalam optimasi parameter bias dan bobot. Optimasi parameter yang digunakan dalam penelitian ini yaitu optimasi parameter ADAM. ADAM menggunakan gradient, lalu estimasi momen pertama dan kedua serta mengoreksi dengan *bias correction*. Penyelesaian gradient pada network ini menggunakan *chain rule* untuk mendapatkan *partial derivative (Bishop, 2006)*. Dalam mengoptimasi parameter dengan menggunakan Adam, hal pertama yang harus dilakukan adalah menghitung gradient dari loss function terhadap parameter yaitu bias dan pembobot, sehingga berikut persamaan yang digunakan.

(2.7)

Hal kedua setelah menghitung gradient dari loss function adalah menghitung Momentum term dan RMSProp term, sehingga berikut persamaan yang digunakan.

Momentum term :

(2.8)

Dimana :

= koefisien untuk momentum

RMSProp term :

(2.9)

Dimana :

= koefisien untuk estimasi kedua

Hal ketiga setelah menghitung Momentum term dan RMSProp term adalah menghitung Bias correction, sehingga berikut persamaan yang digunakan

Bias correction :

(2.10)

Langkah terakhir ialah menghitung parameter update, sehingga berikut persamaan yang digunakan.

Dimana :

= iterasi saat ini dan koefisien momentum

= iterasi saat ini dan koefisien untuk estimasi kedua

Parameter update :

(2.11)

Dimana :

= Learning rate,

= Stabilizer untuk menghindari pembagian dari nol.

### 2.2.10 *Transfer Learning*

*Transfer Learning* adalah salah satu metode dari *deep learning* dengan memanfaatkan model yang sudah dilatih terhadap suatu dataset besar untuk menyelesaikan permasalahan lain. Teknik ini digunakan untuk mempercepat proses *training* dan meningkatkan kualitas model. Dengan menggunakan *pre-trained* model dapat menghemat sumber daya komputasi dan waktu yang diperlukan untuk melatih model dari awal (Patterson & Gibson, 2017). Contoh *pre-trained* model adalah *ResNet, VGGNet, InceptionV3, DenseNet, EfficentNet* dan *MobilNet*. Model-model ini telah dilatih pada dataset skala besar dan kompleks sehingga memiliki kemampuan yang baik untuk berbagai tugas pengenalan pola dan klasifikasi. Setiap model memiliki arsitektur yang unik dan kelebihan masing-masing yang dapat dipilih berdasarkan kebutuhan spesifik dari permasalahan yang dihadapi.

Salah satu dataset yang sering digunakan untuk melatih *pre-trained* model adalah *ImageNet*. Dataset ini berisi sekitar 1,2 juta gambar dengan 1.000 kelas yang berbeda, mencakup berbagai objek dan scene (Iswari, 2021). Dengan dilatih pada dataset yang besar, *pre-trained* model seperti *DenseNet* dan *Xception* mampu menangkap fitur-fitur yang sangat beragam. Sehingga dapat diterapkan pada berbagai aplikasi dari pengenalan objek hingga segmentasi gambar. Penggunaan *pre-trained* model memungkinkan transfer pengetahuan dari satu domain ke domain lain. Meningkatkan efisiensi dan efektivitas dalam pengembangan model deep learning.

Selain efisiensi dan kemampuan generalisasi yang baik, *Transfer Learning* juga memungkinkan pengembangan model dengan dataset yang relatif kecil. Ketika dataset yang tersedia terbatas pelatihan model dari awal bisa menjadi tantangan besar (Hosna et al., 2022). Dalam situasi ini, *pre-trained* model dapat bertindak sebagai dasar yang kuat, di mana model hanya perlu disesuaikan atau di *fine-tune* pada dataset spesifik yang lebih kecil.

1. Ekstraksi Fitur

Ekstraksi fitur adalah proses penting dalam pemrosesan citra yang bertujuan untuk mengambil informasi signifikan atau pola yang merepresentasikan karakteristik utama dari sebuah citra. Proses ini bertujuan untuk mereduksi dimensi data mentah dengan tetap mempertahankan informasi penting yang dapat digunakan untuk analisis atau klasifikasi. Dalam *deep learning* ekstraksi fitur dilakukan secara otomatis oleh lapisan-lapisan dalam model CNN (Huang et al., 2016). Salah satu metode yang umum digunakan dalam *transfer learning* adalah memanfaatkan model *pre-trained*, seperti *DenseNet*, *ResNet*, atau *Xception*, *VGGNet* yang sudah dilatih pada dataset besar seperti *ImageNet*. Model ini berfungsi sebagai base model untuk melakukan ekstraksi fitur dari data baru.

Pada *transfer learning*, lapisan-lapisan dalam model *pre-trained* sering kali dibekukan dengan mengatur properti *False*. Hal ini bertujuan untuk menjaga parameter yang telah dipelajari dari dataset awal sehingga model dapat berfungsi sebagai ekstraktor fitur. Fitur-fitur yang telah diekstraksi kemudian dapat digunakan untuk *custom layers* yang dirancang khusus untuk tugas tertentu, seperti klasifikasi *binary* atau multi-kelas. Ekstraksi fitur otomatis ini memungkinkan, memanfaatkan keunggulan model pre-trained tanpa perlu melatih ulang seluruh model, sehingga menghemat waktu dan sumber daya komputasi (He et al., 2015).

1. *Task Specifics Layers*

*Task specific layers* adalah lapisan yang ditambahkan ke model *deep learning* yang sudah dilatih sebelumnya untuk menyesuaikan model tersebut dengan tugas atau masalah tertentu (Rahman et al., 2025). Dalam konteks transfer learning model pre-trained seperti *DenseNet*, *VGGNet*, atau *ResNet* digunakan untuk melakukan ekstraksi fitur dari data baru. Fitur-fitur yang telah dipelajari dari dataset besar seperti *ImageNet* digunakan untuk memahami pola-pola dasar dalam data, seperti tepi, tekstur, atau bentuk. Namun, untuk menyelesaikan tugas tertentu, seperti klasifikasi citra atau deteksi objek, model memerlukan lapisan tambahan yang dirancang khusus untuk tugas tersebut.

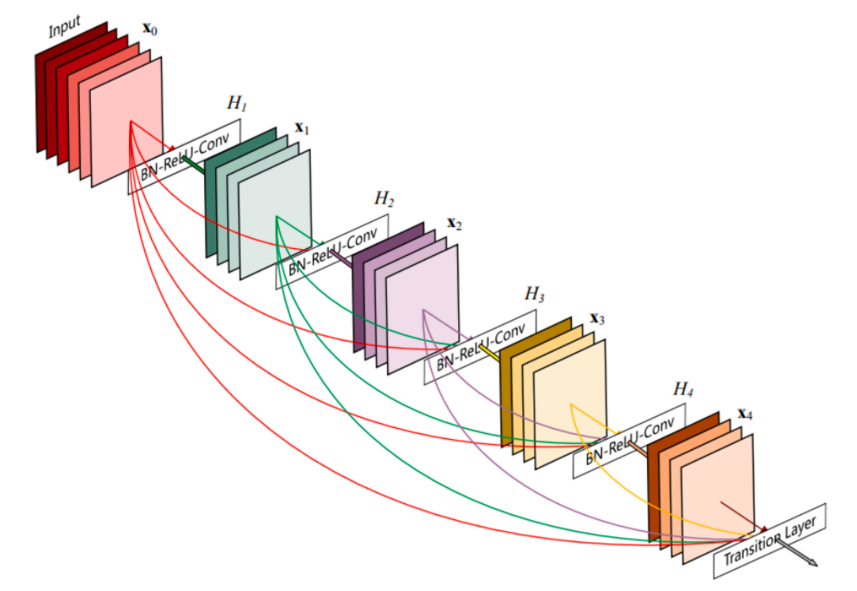
Lapisan yang ditambahkan ini sering disebut sebagai *task specific layers*, yang bertujuan untuk mengambil fitur yang telah diekstraksi oleh model *pre-trained* dan memanfaatkan informasi tersebut untuk memecahkan masalah spesifik. Keuntungan utama dari penggunaan *task specific layers* adalah memungkinkan model untuk memanfaatkan data yang telah dipelajari dari data besar tanpa memerlukan pelatihan ulang pada seluruh model, sehingga menghemat waktu dan sumber daya komputasi (Rahman et al., 2025).

### 2.2.11 Model *DenseNet201*

*DenseNet201* adalah salah satu metode *deep learning* dari CNN yang terdiri dari 201 lapisan, yang diperkenalkan oleh Gao Huang dkk, pada tahun 2017 (Virnodkar et al., 2022). Jaringan ini dikenal karena menggunakan *dense connections* yang terdiri dari *dense blocks* dan *transition layers*. *Dense blocks* yang terdiri dari *Bottleneck Layers*, berfungsi sebagai komponen utama dalam jaringan model. Mode koneksi padat didalam *dense blocks* memungkinkan setiap lapisan terhubung langsung dengan lapisan lainnya, sehingga informasi dapat tersalurkan dengan optimal dan ukuran keluaran tetap konsisten diseluruh lapisan. *DenseNet* mengatur jumlah *channels* melalui *bottleneck layers*, *transition layers* dan *growth layer* (Huang et al., 2016). Pendeketan ini tidak hanyak dapat mengurangi jumlah parameter dan mengatasi masalah *overfitting* akan tetapi juga dapat secara signifikan menurunkan beban komputasi. Secara keseluruhan arsitektur DenseNet201 yang digunakan dapat dilihat pada Gambar 2.11.

Salah satu keunggulan utama dari *DenseNet201* adalah kemampuannya dalam memitigasi masalah vanishing gradient yang sering menjadi tantangan dalam *deep learning*. Dengan menghubungkan setiap lapisan langsung ke setiap lapisan lainnya dalam *dense blocks DenseNet* memastikan bahwa gradien dapat mengalir lebih mudah melalui jaringan selama proses pelatihan. Ini berarti bahwa informasi penting tidak hilang saat melalui banyak lapisan dan pelatihan dapat berlangsung lebih efisien dan efektif (Kong & Cheng, 2022).

*DenseNet201* juga memberikan penawaran seperti efisiensi parameter yang tinggi. Karena *dense connections* memungkinkan reuse fitur dari lapisan sebelumnya, sehingga jumlah parameter yang diperlukan jauh lebih sedikit dibandingkan dengan arsitektur jaringan konvolusi tradisional (Devi et al., 2024). Ini tidak hanya mengurangi kebutuhan akan sumber daya komputasi yang besar tetapi juga membantu dalam menghindari *overfitting*, terutama ketika bekerja dengan dataset yang lebih kecil.

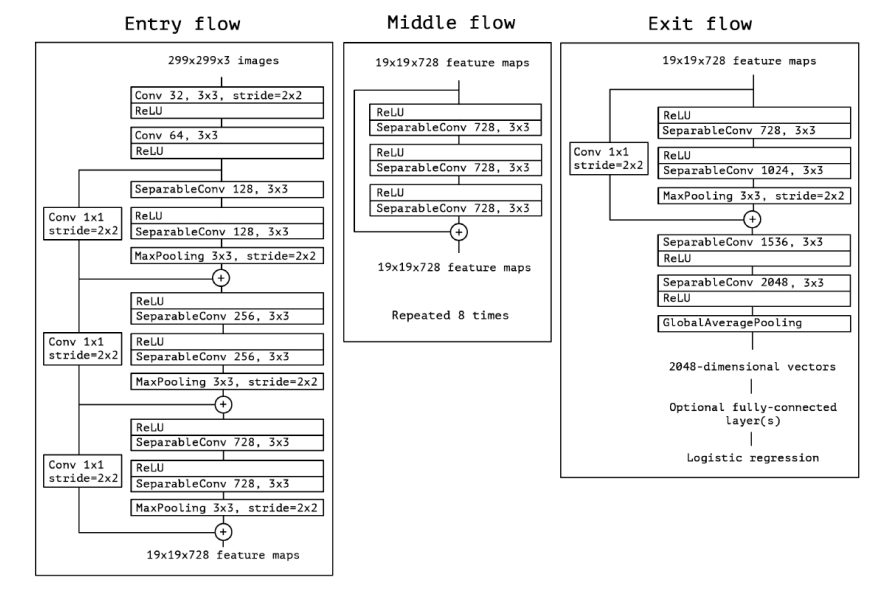
  
Gambar 2. 11Arsitektur DenseNet201

### 2.2.12 Model *Xception*

*Xception* adalah arsitektur model dari CNN *open source* yang diteliti dan dikembangkan oleh Google. Nama *Xception* sendiri merupakan singkatan dari *Extreme version of Inception* yang berarti Versi yang lebih ekstrem dari Inception atau bisa diartikan versi yang lebih intens dari model yang sebelumnya dibuat oleh google. Model ini menggabungkan beberapa inovasi dari Inception dan mengembangkan pendekatan baru untuk meningkatkan kinerjanya. *Xception* dirancang untuk memberikan fleksibilitas dan performa tinggi dalam berbagai tugas klasifikasi gambar, sehingga yang menjadikannya salah satu model populer dalam bidang klasifikasi gambar (Chollet, 2017).

*Xception* memiliki lapisan konvolusi yang menggabungkan *pointwise convolution* yang kemudian diikuti oleh *deptwhise convolution*. Kombinasi ini memungkinkan *Xception* untuk menghemat jumlah parameter dan meningkatkan efisiensi komputasi sekaligus mempertahankan akurasi. Dari kombinasi tersebut terbukti memberikan kinerja yang lebih baik dalam melakukan klasifikasi gambar pada berbagai dataset besar seperti *ImageNet* (Chollet, 2017).

Arsitektur *Xception* sendiri memiliki lapisan yang terdiri dari 36 lapisan *konvolusi* yang menjadi dasar jaringan ekstraksi fitur. Ke-36 lapisan *konvolusi* tersebut disusun kedalam 14 modul, dimana semuanya memiliki koneksi *residual linear* disekitarnya kecuali untuk modul pertama dan terakhir. Selain itu *Xception* sendiri memiliki arsitektur yang memudahkan dalam memodifikasi pada parameter-parameter yang dimilikinya (Carnagie et al., 2022). Spesifik lengkap tentang jaringan *Xception* dapat dilihat Gambar 2.12.

  
Gambar 2. 12 Arsitektur Xception

### 2.2.13 Evaluasi Ketetapan Klasifikasi

Evaluasi model klasifikasi digunakan untuk menentukan model terbaik dengan cara Melihat kemungkinan kesalahan klasifikasi yang dilakukan oleh suatu model. Model klasifikasi akan menghasilkan nilai *output* dalam bentuk dikrit maupun kontinu. Nilai *diskrit* akan memprediksi label kelas dari testing, sedangkan *kontinu* akan mempresentasikan estimasi dari Probabilitas kelas prediksi (Tharwat, 2018). Dalam mengukur ketepatan klasifikasi, perlu diketahui jumlah data pada setiap kelas prediksi dan kelas aktual yang terdiri TP (*True Possitive*), TN (*True Negative*), FP (*False Positive*) dan FN (*False Negative*). True Positive (TP) adalah jumlah kasus di mana model berhasil memprediksi dengan benar bahwa data termasuk ke dalam kelas positif ketika label sebenarnya adalah positif dan model juga memprediksi positif. True Negative (TN) adalah jumlah kasus di mana model berhasil memprediksi dengan benar bahwa data termasuk ke dalam kelas negatif ketika label sebenarnya negatif dan model juga memprediksi negatif. False Positive (FP) atau dikenal sebagai kesalahan tipe I terjadi ketika model salah memprediksi data sebagai kelas positif padahal label sebenarnya adalah negatif. Sedangkan False Negative (FN) atau kesalahan tipe II terjadi ketika model salah memprediksi data sebagai kelas negative, label sebenarnya adalah positif (Choi et al., 2024). Berkut merupakan *confusion matrix* yang memuat keempat nilai tersebut.

Tabel 2. 2 Confusion Matrix

|  |  |  |
| --- | --- | --- |
| Kelas Aktual | Kelas Prediksi | |
| Positif | Negatif |
| Positif | TP | FN |
| Negatif | FP | TN |

Kinerja klasifikasi dapat diukur dengan menggunakan *accuracy*, *precision*, *recall*, dan *Fscore*. *Accuracy* adalah banyak pengamatan yang terklasifikasi secara tepat. *Precision* adalah banyak pengamatan yang tepat terprediksi positif dari keseluruhan dengan hasil prediksi positif. *Recall* adalah banyaknya pengamatan yang tepat diklasifikasikan sesuai kategorinya. *Fscore* didapatkan dari nilai kombinasi antara *precision* dan *recall*. Dalam mengidentifikasi penyakit pada bidang medis diperlukan diagnosis yang tepat sebanyak mungkin sehingga hal yang perlu diperhatikan adalah proporsi penyakit yang teridentifikasi benar pada semua kasus penyakit. Oleh karena itu pada bidang medis sebaiknya menggunakan *sensitivity (Chen, 2019)*. Perhitungan *accuracy*, *precision*, *recall* dan *Fscore* dapat dilakukan dengan menggunakan persamaan.

(2.12)

(2.13)

(2.15)

(2.16)

### 2.2.14 *Python*

*Python* adalah bahasa pemrograman computer serbaguna yang sering digunakan untuk membangun situs website, software atau aplikasi, mengotomasikan tugas dan melakukan analisis data. Bahasa pemrograman ini termasuk bahasa tujuan umum. Fleksibilitasnya memungkinkan pengguna untuk menciptakan berbagai jenis program dimulai dari pengolahan data, pengembangan kecerdasan buatan, hingga mengotomasikan dalam bidang industri. Python juga memiliki sintaks yang sederhana dan mudah dipahami sehingga sangat cocok bagi pemula yang baru mulai belajar pemrograman (Sharma et al., 2020).

Popularitas *python* terus meningkat berkat sifatnya yang mudah dipahami dan ekosistem yang luas. Berdasarkan survei pengembangan *Stack Overflow* tahun 2022, python menempati peringkat keempat sebagai bahasa pemrograman terpopuler. Hampir 50% responden mengaku menggunakan *python* untuk menyelesaikan tugas-tugas harian mereka. Dalam bidang pengembangan perangkat lunak *python* banyak digunakan banyak digunakan untuk pengolahan data, pembelajaran mesin dan pengembangan antarmuka *Application Programming Interface* (API). Bahkan python menjadi bahasa pemrograman dasar untuk *framework* seperti *Django* dan *Flask* yang memudahkan pengembangan aplikasi berbasis web (Marchand, 2022).

Nama *python* sendiri memiliki latar belakang yang unik. Ketika Guido van Rossum menciptakan bahasa ini pada akhir 1980-an ia terinspirasi dari acara komedi Inggris *Monty Python’s Flying Circus* yang ia tonton saat itu. Menurutnya nama tersebut terdengar singkat,menarik, dan sedikit misterius. Hal ini mencerminkan filosofi *python* sebagai bahasa pemrograman yang sederhana dan mudah dipahami namun kuat (Rossum, 1996).

### 2.2.15 *TensorFlow*

*TensorFlow* adalah sebuah *framework* *open source* yang dirancang untuk komputasi dan *machine learning*. Dikembangkan oleh tim Google Brain pada tahun 2015. *TensorFlow* menawarkan fleksibilitas tinggi dalam membangun model *machine learning* dan *deep learning*. *Framework* ini mendukung operasi berbagai operasi matematika kompleks yang digunakan dalam pembuatan modeling seperti regresi, klasifikasi, pengenalan pola. *TensorFlow* memiliki arsitektur yang memungkinkan komputasi lintas platform dari perangkat *mobile* hingga *kluster* komputasi dasar menjadikannya pilihan populer dikalangan.

Salah satu fitur utama *TensorFlow* adalah kemampuannya untuk memanfaatkan grafik aliran data. Grafik ini memberitahu perhitungan dalam bentuk *nodes* dan *edges*, dimana *nodes* operasi matematika dan *edges* adalah data *tensor* yang mengalir diantara *nodes*. Pendekatan ini memungkinkan optimalisasi efisien, terutama untuk tugas-tugas yang memerlukan pemrosesan data besar secara paraler. *TensorFlow* mendukung berbagai bahasa pemrograman yang dapat memudahkan integrasi dengan sistem yang sudah ada seperti *Python*, *C++*, dan *javascript* (Abadi et al., 2015).

*Framework* ini terus berkembang dengan adanya fitur seperti *TensorFlow Lite* untuk pengembangan aplikasi pada perangkat *mobile* seperti *TensorFlow.js* untuk implementasi pada aplikasi berbasis web, *TensorFlow Extended* (TFX) untuk manajemen alur kerja *machine learning* secara end-to-end. Dengan ekosistem yang luas dan komunitatif *TensorFlow* salah satu *framework* yang digunakan dalam pengembangan model kecerdasan buatan. Keunggulannya dalam skalabilitas dan fleksbilitas menjadikannya alat yang sangat efektif untuk menangani permasalahan data yang kompleks (*TensorFlow*, n.d.).

### 2.2.16 Google Colab

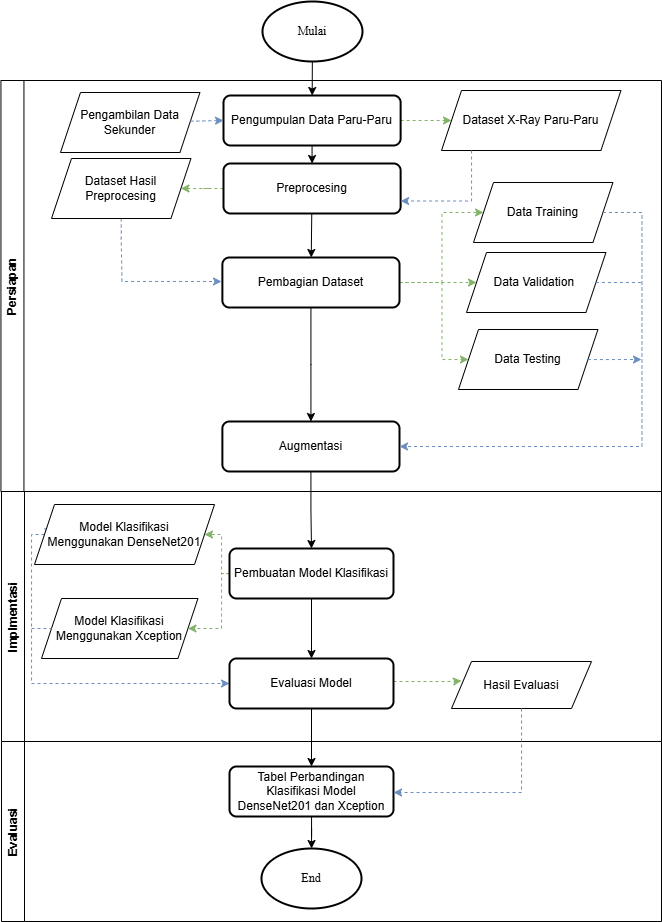
Google Colab adalah sebuah platform pemrograman berbasis cloud yang dikembangkan oleh oleh Google. Google Colab memberikan layanan menjalankan kode Python langsung dibrowser tanpa memerlukan pengaturan lingkungan lokal. Dengan integrasi google drive, google colab mempermudah kolaborasi antar pengguna yang memungkinkan berbagai dan mengedit notebook secara bersamaan. Colab mendukung berbagai aplikasi termasuk data, visualisasi, pengembangan model *deep learning* dan *machine learning* yang menjadikannya populer dikalangan peneliti dan praktis data (Carneiro et al., 2018).

Colab menawarkan akses gratis ke sumber daya komputasi yang kuat seperti GPU dan TPU, yang dapat digunakan untuk mempercepat pelatihan model dalam *deep learning* dan *machine learning*. Google Colab juga menyediakan akses mudah ke berbagai library *python* populer seperti *TensorFlow*, *NumPy*, dan *Matplotlib* tanpa memerlukan instalasi manual. Google Colab juga dapat mengintegrasikan dengan layanan cloud lain seperti google *cloud storage* sehingga memperluas kemampuannya untuk mengelola data besar (Scholar & Bagane, 2024).

# BAB 3 METODOLOGI PENELITIAN

## **3.1 Tahapan Penelitian**

Pada tahapan penelitian ini, akan dijelaskan mengenai alur penelitian yang digunakan dalam pengerjaan tugas akhir. Tahapan penelitian ini memberi pedoman berupa alur penelitian yang dilakukan selama penelitian berlangsung.

  
Gambar 3. 1 Flowchart Tahapan Penelitian

### 3.1.1 Persiapan

1. Pengumpulan Data Paru-Paru

Pada proses ini merupakan tahapan yang digunakan untuk pengambilan data yang dibutuhkan dalam pengembangan model klasifikasi gambar. Dataset yang digunakan mencakup data citra *X-Ray* pneumonia, *X-Ray* TBC, dan *X-Ray* normal. Populasi penelitian ini berasal dari data sekunder. Untuk data sekunder diperoleh pada platform kaggle dan data dari rumah sakit. Data berasal dari platform kaggle berjumlah 7225 dataset citra x-ray dan data dari rumah sakit sejumlah 300 citra x-ray dengan rincian 100 data citra x-ray paru-paru normal, 100 data citra x-ray paru-paru pneumonia, dan 100 data citra x-ray paru-paru tuberculosis.

1. *Preprocessing*

Tahap *preprocessing* data terhadap data citra merupakan tahap awal yang penting untuk dilakukan sebelum melakukan proses pengenalan citra. Tahap ini bertujuan untuk mempersiapkan data citra agar dapat diolah dengan baik oleh model pengenalan citra.

1. Pengurangan data citra

Pengurangan data citra merupakan suatu proses dari preprocessing data yang bertujuan untuk menyederhanakan dataset tanpa mengurangi representasi atau kualitas informasi yang relavan. Proses ini dilakukan untuk memastikan dalam pengembangan model lebih terfokus dan efisien, baik dari segi waktu pemrosesan maupun kebutuhan sumber daya komputasi. Pengurangan data citra dilakukan secara manual dengan mengurangi jumlah data menjadi 1000 data untuk setiap label.

2. Menerapkan CLAHE pada data citra

CLAHE merupakan langkah yang digunakan meningkatkan kontras gambar sehingga fitur penting dalam citra menjadi lebih terlihat dan dapat diolah dengan baik oleh model. CLAHE bekerja dengan menyesuaikan histogram lokal pada setiap gambar, menghindari peningkatan kontras yang secara berlebihan pada data (Koonsanit et al., 2017).

1. Pembagian Dataset

Dalam penelitian ini, terdapat 2000 data citra *X-Ray* yang dibagi menjadi tiga kelompok yaitu, data latih, data uji, dan data validasi. Data tersebut diambil secara acak dari keseluruhan dataset citra *X-Ray*. Sebanyak 1400 data *X-Ray* digunakan sebagai data latih untuk membangun sistem. Sebanyak 300 data validasi digunakan selama pelatihan untuk memonitor kinerja model. Sebanyak 300 data digunakan untuk data uji untuk mengevaluasi kinerja akhir model setelah pelatihan selesai.

1. Augmentasi

Augmentasi data pada citra dilakukan untuk meningkatkan variasi dataset tanpa menambah jumlah gambar asli secara langsung. Augmentasi citra dilakukan menggunakan berbagai metode. Augmentasi citra yang tepat akan meningkatkan performa model pengenalan citra.

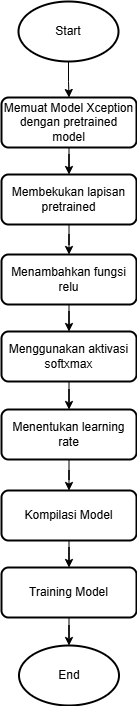
### 3.1.2 Implementasi

1. Pembuatan Model

Dalam tahap pembuatan model ini, model yang digunakan adalah *DenseNet201* dan *Xception*. Berikut adalah tahapan pembuatan model *DenseNet201* dan *Xception*.

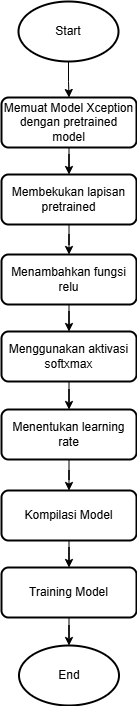
1. Mode*l DenseNet201*

Model *DenseNet201* adalah yang terdiri dari 201 lapisan, yang mempunyai *dense connections* yang terdiri dari *dense blocks* dan *transition layers*. Setiap lapisan terhubung langsung dengan lapisan lainnya, sehingga informasi dapat tersalurkan dengan optimal dan ukuran keluaran tetap konsisten diseluruh lapisan (Huang et al., 2016). Tahapan pembuatan model *DenseNet201* antara lain model menggunakan *ImageNet* dengan memanfaatkan pretrained untuk membantu model dalam mengklasifikasikan, menambahkan *freze* model yang memiliki fungsi agar lapisan model yang sudah dilatih sebelumnya tidak akan diperbarui selama pelatihan untuk mencegah overfitting, menambahkan fungsi relu, menggunakan aktivasi sigmoid, menentukan learning rate, menambahkan kompilasi model yang digunakan untuk optimalisasi model seperti optimasi adam, melakukan pelatihan atau training model (Cobilla et al., 2023).

  
Gambar 3. 2 Alur Pembuatan Model DenseNet201

1. Model *Xception*

Model *Xception* adalah model memiliki lapisan konvolusi yang menggabungkan *pointwise convolution* yang kemudian diikuti oleh *deptwhise convolution*, yang memiliki kelebihan menghemat jumlah parameter dan meningkatkan efisiensi komputasi sekaligus mempertahankan akurasi (Chollet, 2017). Terdapat beberapa tahap untuk pembuatan model *Xception* antara lain, model menggunakan *ImageNet* dengan memanfaatkan pretrained untuk membantu model dalam mengklasifikasikan, karena model tidak perlu lagi untuk mempelajari fitur-fitur dasar dari awal yang memerlukan dataset banyak. Menambahkan *freze* model yang memiliki fungsi agar lapisan Model yang sudah dilatih sebelumnya tidak akan diperbarui selama pelatihan, hal ini dilakukan untuk mencegah *overfitting*. Menambahkan fungsi relu. Menggunakan aktivasi sigmoid. Menentukan learning rate. Menambahkan kompilasi model yang digunakan untuk optimalisasi model seperti optimasi adam. Melakukan pelatihan atau training model (Cobilla et al., 2023).

  
Gambar 3. 3 Alur Pembuatan Model Xception

1. Evaluasi Model

Kinerja klasifikasi Model DenseNet201 dan Model Xception diukur dengan menggunakan *accuracy*, *precision*, *sensitivity*, dan *Fscore* yang diambil dari dataset testing. Untuk mengetahui nilai dari *accuracy*, *precision*, *sensitivity*, dan *Fscore* dapat diketahui dari nilai TN, TP, FN, dan FP yang diperoleh dari *confusion matrix* (hart et al., 2024). Perhitungan *accuracy*, *precision*, *recall* dan *Fscore* dapat dilakukan dengan menggunakan persamaan berikut.

(3.1)

(3.2)

(3.3)

(3.3)

Akurasi ini memiliki tingkat nilai diagnosa yaitu :

1. Akurasi antara 0.90 dan 1.00 diklasifikasikan sebagai berikut “excellent classification” (klasifikasi sangat baik).
2. Akurasi antara 0.80 dan 0.90 diklasifikasikan sebagai berikut “good classification” (klasifikasi baik).
3. Akurasi antara 0.70 dan 0.80 diklasifikasikan sebagai berikut “fair classification” (klasifikasi cukup).
4. Akurasi antara 0.60 dan 0.70 diklasifikasikan sebagai berikut “poor classification” (klasifikasi kurang baik).
5. Akurasi antara 0.50 dan 0.60 diklasifikasikan sebagai berikut “failure” (gagal).

### Evaluasi

Pada tahap evaluasi dilakukan dengna membuat tabel perbandingan dari hasil klasifikasi model *DenseNet201* dan *Xception*. Kinerja kedua model arsitektur CNN ini dibandingkan berdasarkan akurasi, presisi, recall, dan F1-Score. Metrik-metrik tersebut dihitung menggunakan confusion matrix yang diperoleh dari hasil klasifikasi terbaik masing-masing model.

## **3.2 Jadwal Kegiatan**

Tabel 3. 1 Tabel Jadwal Kegiatan

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| No | KEGIATAN | Bulan I | | | | Bulan II | | | | Bulan III | | | | Bulan IV | | | |
| I | II | III | IV | I | II | III | IV | I | II | III | IV | I | II | III | IV |
| 1 | Pengumpulan Data Pneumonia |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | Preprocessing Data |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 | Pembagian dataset |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | Permodelan Model Klasifikasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 | Evaluasi Model |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 6 | Perbandingan Model |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

# BAB 4 HASIL DAN PEMBAHASAN

## **4.1 Pengumpulan Data Paru-Paru**

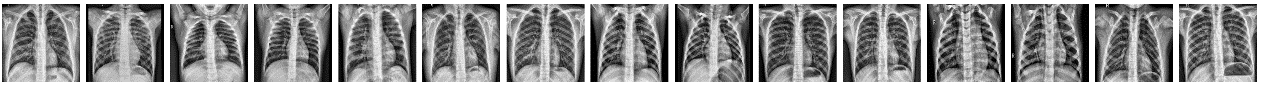
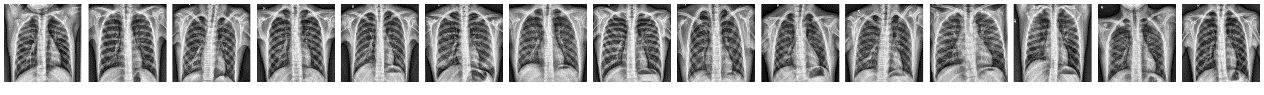
Pada proses ini merupakan tahapan yang digunakan untuk mengumpulkan data yang dibutuhkan dalam pengembangan model klasifikasi gambar. Dataset yang digunakan mencakup data citra *X-Ray* normal, *X-Ray* pneumonia, dan *X-Ray* tbc. Populasi penelitian ini berasal dari data sekunder. Untuk data sekunder diperoleh pada platform kaggle dan data dari rumah sakit. Dataset dari kaggle diperoleh dari 3 sumber. Data untuk kategori pneumonia dan normal diperoleh dari <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>, sedangkan data untuk tuberculosis diperoleh dari dua sumber yaitu <https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset> dan <https://www.kaggle.com/datasets/raddar/tuberculosis-chest-xrays-shenzhen>. Total dataset sekunder yang diperoleh dari platform kaggle dan rumah sakit berjumlah 7218 yang terbagi menjadi 3 kategori yaitu normal, pneumonia, dan tbc. Dataset yang diperoleh pada platform kaggle berjumlah 7225 dataset citra x-ray. Sementara itu data yang berasal dari rumah sakit berjumlah 100 citra x-ray. Adapun rincian data sekunder kaggle dan rumah sakit dapat dilihat pada tabel.

Tabel 4. 1 Pembagian Data Sekunder

|  |  |  |
| --- | --- | --- |
| Sumber | Citra X-Ray | Jumlah |
| Kaggle | Normal | 1583 |
| Pneumonia | 4273 |
| Tbc | 1362 |
| Rumah Sakit | Normal |  |
| Pneumonia | 60 |
| Tbc | 40 |
| Total | 7318 | |

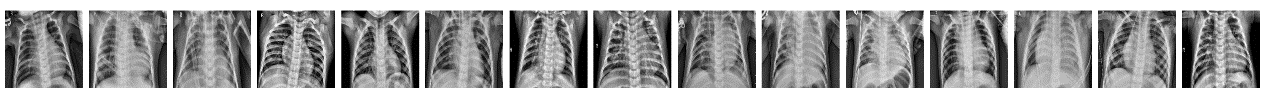
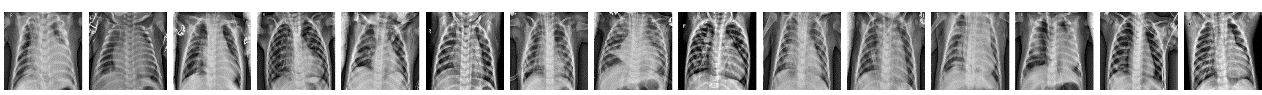
Berdasarkan dataset diatas, dapat dikelompokkan menjadi empat kategori kelas yaitu:

1. Normal, dengan jumlah data sebanyak 1583 citra x-ray. Adapun kategori kelas normal dapat dilihat seperti pada gambar.



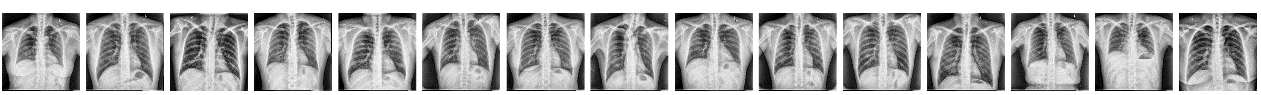
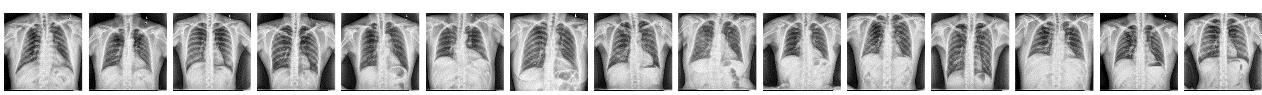
Gambar 4. 1 Dataset Kategori Normal

1. Pneumonia, dengan jumlah data sebanyak 4333 citra x-ray. Adapun kategori kelas pneumonia dapat dilihat seperti pada gambar.



Gambar 4. 2 Dataset Kategori Pneumonia

1. Tuberculosis, dengan jumlah data sebanyak 1402 citra x-ray. Adapun kategori kelas normal dapat dilihat seperti pada gambar.



Gambar 4. 3 Dataset Kategori Tuberculosis

## **4.2 Preprocesing**

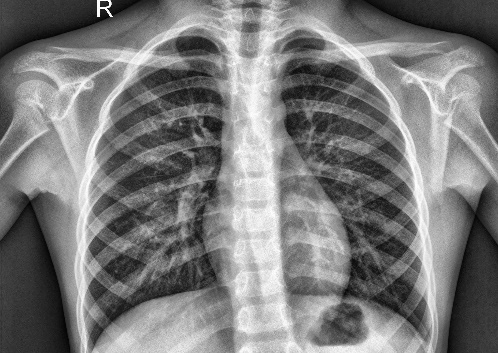
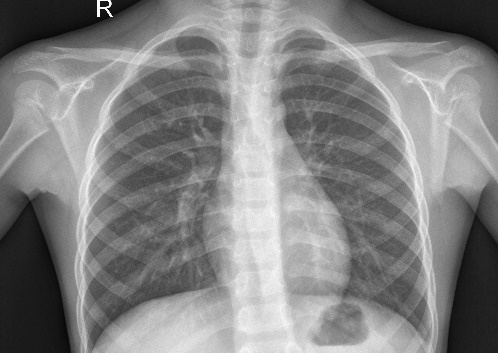
### 4.2.1 Pengurangan Data Citra

Setelah data telah terkumpul dilakukan pengurangan data citra digunakan untuk menyesuaikan kebutuhan penelitian. Data yang awalnya diperoleh dari Kaggle dan rumah sakit berjumlah 7318 dataset citra x-ray yang terbagi dalam 3 kategori yaitu, Normal, Pneumonia, dan Tuberculosis. Proses pengurangan data dilakukan untuk memastikan distribusi yang lebih seimbang serta mengurangi beban komputasi dalam proses pelatihan model. Dengan jumlah yang lebih sedikit dapat dilakukan lebih efisien tanpa mengorbankan representasi dari masing-masing kategori.

Proses pengurangan data dilakukan secara manual dengan mempertimbangkan kesimbangan jumlah data dalam setiap kategori. Pemilihan citra yang dipertahankan atau dihapus dilakukan berdasarkan beberapa kriteria, seperti kualitas gambar, kejelasan struktur paru-paru, dan kesesuaian label dengan kondisi yang ditampilkan pada citra. Setelah dilakukan pengurangan, dataset yang digunakan dalam penelitian ini terdiri dari 3.000 citra X-Ray, dengan pembagian yang tetap mempertahankan keseimbangan antara kategori Normal, Pneumonia, dan Tuberculosis. Jumlah ini dipilih berdasarkan pertimbangan efisiensi dalam proses pelatihan model serta keterbatasan sumber daya komputasi.

### 4.2.2 Peningkatkan Kontras Metode Clahe

Metode Clahe digunakan untuk meningkatkan nilai kontras citra terutama pada area yang memiliki tongkat kontras yang rendah. Sehingga detail dalam citra menjadi lebih jelas dan mudah diinterpresentasikan. Dalam penerapan metode clahe, digunakan nilai clip limit sebesar 2 dan untuk ukuran tile grid sebesar 8x8. Sehingga dengan parameter ini, CLAHE dapat mampu meningkatkan kualitas visual citra secara optimal tanpa mengorbankan informasi penting yang terakndung didalamnya.



Gambar 4. 4 Citra X-Ray Sebelum (kiri) dan Sesudah CLAHE (kanan)

## **4.3 Pembagian Dataset**

Pembagian dataset merupakan langkah penting untuk memastikan bahwa model deteksi objek dapat dilatih, divalidasi, dan diuji dengan cara yang terstruktur dan representasif. Pembagian dataset menjadi 3 diantaranya data training, data validation, data testing. Data training digunakan untuk melatih model dengan mengidentifikasi pola, karakteristik, dan objek yang relevan untuk dideteksi. Selama pelatihan, model mengoptimalkan bobot dan biasnya melalui beberapa epoch berulang kali.

Data validation digunakan untuk mengevaluasi performa model selama pelatihan untuk menghindari overfitting. Setiap epoch, model diuji pada data validasi untuk memberikan gambaran kinerja pada data baru. Hasil ini digunakan untuk menyempurnakan hyperparameter dan dapat menyebabkan penghentian pelatihan lebih awal atau early stopping jika diperlukan.

Data testing digunakan untuk mengevaluasi performa akhir model setelah data selesai. Data ini hanya digunakan sekali dan tidak berpatisipasi dalam proses pelatihan atau validasi, sehingga memberikan evaluasi yang akurat terhadap model yang sudah dilatih.

Dalam penelitian ini pembagian data yang digunakan meliputi

* 80% data training, 10% data validation, 10% data testing
* 70% data training, 20% data validation, 10% data testing
* 70% data training, 15% data validation, 15% data testing
* 60% data training, 30% data validation, 10% data testing

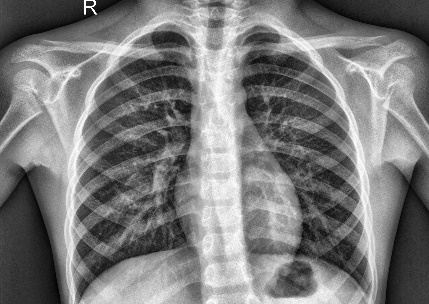
Tabel 4. 2 Pembagian Data Training, Validation, dan Testing

|  |  |  |  |
| --- | --- | --- | --- |
| Pembagian Data | Data Training | Data Validation | Data Testing |
| 80%:10%:10% | 2400 | 300 | 300 |
| 70%:20%:10% | 2100 | 600 | 300 |
| 70%:15%:15% | 2100 | 450 | 450 |
| 60%:30:10% | 1800 | 900 | 300 |

Semakin banyak data yang digunakan untuk training maka semakin baik model yang dilatih. Pembagian ini juga memastikan ada cukup data yang digunakan untuk validasi serta pengujian untuk mengevaluasi kinerja model.

## 4.4 Augmentasi Data

Augmentasi data adalah teknik yang digunakan untuk meningkatkan keberagaman data pelatihan dengan menerapkan transformasi seperti rotasi, pergeseran, flipping, dan zooming pada citra tanpa mengubah label aslinya. Teknik ini bertujuan untuk meningkatkan kemampuan generalisasi model dalam mengenali pola pada data baru. Berbagai teknik augmentasi dilakukan pada penelitian ini seperti rescale dengan menormalisasikan nilai piksel ke rentang 0,1, rotasi acak hingga 40 derajat, pergeseran horizontal dan vertikal hingga 20%, transformasi shear dan zoom hingga 20%, dan flip horizontal untuk meningkatkan variasi data. Data ini kemudian dilakukan pengubahan ukuran gambar untuk memastikan bahwa semua gambar dalam dataset memiliki ukuran yang sama sebelum diproses oleh model. Dalam penelitian ini gambar dirubah keukuran 512x512 piksel dengan batch size sebesar 32. Augmentasi dilakukan secara real-time (on-the-fly). Variasi citra dihasilkan secara dinamis selama proses pelatihan berlangsung, sehingga tidak menambah jumlah data secara fisik dalam direktori dataset. Setiap citra diproses menjadi berbagai bentuk baru pada setiap epoch, memungkinkan model menerima data yang lebih bervariasi di setiap iterasi dan meningkatkan kemampuan generalisasi.

Gambar 4. 5 Mengubah Ukuran Gambar Menjadi 512x512

## **4.5 Pembuatan Model Klasifikasi**

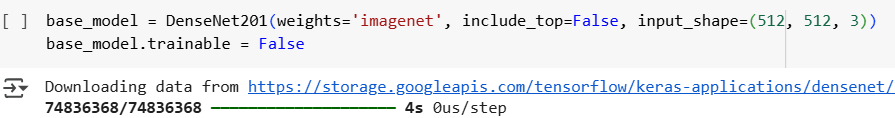
Pada tahap ini dilakukan pembuatan model untuk melakukan klasifikasi citra X-ray paru-paru kedalam tiga kelas, yaitu normal,pneumonia, dan tuberculosis. Model yang digunakan dalam penelitian ini adalah DesneNet201 dan Xception. Kedua model tersebut merupakan salah satu arsitektur model *convutional neural network* (CNN) yang terbukti efektif dalam dalam klasifikasi citra. Proses pembuatan model mencangkup pemanggilan arsitektur model, penyesuaian struktur model untuk klasifikasi tiga kelas, kompilasi model, dan pelatihan menggunakan data yang telah dibagi dalam tiga skema berbeda yang akan ditampilkan dalam bentuk tabel.

Tabel 4. 3 Pembagian Data Training, Validation, Testing

|  |  |  |  |
| --- | --- | --- | --- |
| Pembagian Data | Data Training | Data Validation | Data Testing |
| 80%:10%:10% | 2400 | 300 | 300 |
| 70%:20%:10% | 2100 | 600 | 300 |
| 70%:15%:15% | 2100 | 450 | 450 |
| 60%:30:10% | 1800 | 900 | 300 |

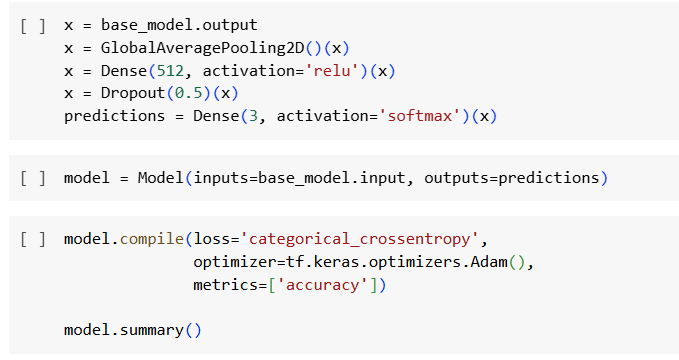
Model dilatih menggunakan data latih dan divalidasi menggunakan data validasi untuk memantau performa selama proses pelatihan. Teknik early stopping digunakan agar proses pelatihan dapat berhenti secara otomatis jika tidak ada peningkatan performa dalam beberapa epoch, sehingga menghindari overfitting. Hasil pelatihan model kemudian dianalisis berdasarkan metrik akurasi dan loss, baik pada data latih maupun validasi.

### 4.5.1 Pembuatan Model DenseNet201



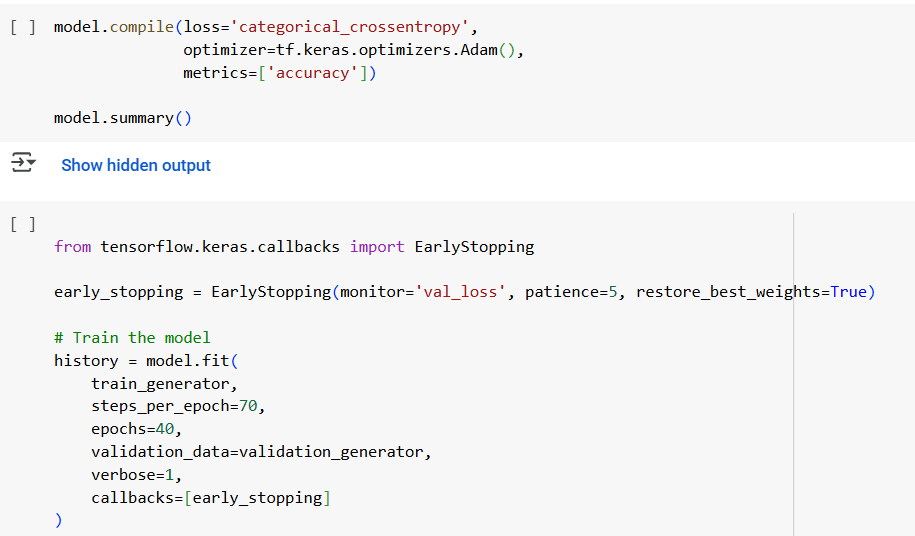
Gambar 4. 6 Pemanggilan Model DesneNet201

Dalam penelitian ini, digunakan arsitektur Dense Convolutional Network (DenseNet201) sebagai model dasar untuk klasifikasi citra x-ray paru-paru. Pada Gambar 4.5 ditunjukkan bahwa model ini diimplementasikan melalui pustaka *keras.applications*, dengan bobot awal yang diambil dari hasil pelatihan sebelumnya menggunakan dataset ImageNet, sehingga model tidak perlu dilatih dari awal. Selain itu, parameter *include\_top* disetel ke nilai false untuk menghilangkan bagian akhir dari arsitektur DenseNet201 yang berfungsi sebagai fully connected layer bawaan model. Hal ini dilakukan agar peneliti dapat menambahkan layer klasifikasi yang disesuaikan dengan jumlah kelas target dalam penelitian, yaitu tiga kelas: normal, pneumonia, dan tuberculosis. Ukuran input citra yang digunakan adalah 512 × 512 piksel dengan tiga saluran warna, yaitu *red*, *green*, dan *blue* (RGB), yang sesuai dengan format standar citra warna. Untuk menghindari pelatihan ulang terhadap seluruh parameter model dasar, nilai atribut *trainable* pada objek model dasar diatur menjadi *false*. Dengan konfigurasi ini, seluruh layer pada DenseNet201 tidak akan diperbarui selama proses pelatihan, sehingga model hanya berfungsi sebagai ekstraktor fitur.



Gambar 4. 7 Struktur Model Transfer Learning DenseNet201

Pada Gambar 4.6, di atas output dari DenseNet201, ditambahkan beberapa layer klasifikasi lanjutan. Proses ini diawali dengan penerapan layer *Global Average Pooling 2D* untuk meratakan fitur hasil ekstraksi. Selanjutnya, digunakan satu layer *Dense* dengan 512 neuron dan fungsi aktivasi *Rectified Linear Unit* (ReLU), yang kemudian diikuti oleh layer *dropout* dengan nilai 0,5 sebagai strategi untuk mengurangi risiko *overfitting*. Sebagai layer output, digunakan layer *dense* dengan tiga neuron dan fungsi aktivasi *softmax* untuk mengklasifikasikan citra ke dalam tiga kategori: normal, pneumonia, dan tuberculosis.



Gambar 4. 8 Proses Kompilasi dan Pelatihan Model

Pada Gambar 4.7, model dikompilasi menggunakan fungsi loss categorical\_crossentropy, algoritma optimasi Adam, dan metrik evaluasi berupa *accuracy*. Untuk mencegah *overfitting* dan pelatihan berlebihan, diterapkan teknik *Early Stopping* yang memantau nilai *validation loss* dan menghentikan proses pelatihan apabila tidak terjadi peningkatan dalam lima *epoch* berturut-turut. Parameter restore\_best\_weights=True memastikan bahwa bobot model akan dikembalikan ke kondisi terbaik sepanjang proses pelatihan. Proses pelatihan dilakukan dengan memanfaatkan fungsi model.fit() dengan konfigurasi *steps per epoch* sebesar 70, jumlah maksimum *epoch* sebanyak 40, dan data validasi diambil dari parameter validation\_generator sebanyak 5 langkah setiap *epoch*. Dengan konfigurasi ini, model diharapkan mampu melakukan pembelajaran secara optimal terhadap data latih, divalidasi secara berkala pada setiap epoch, serta meminimalkan risiko *overfitting* melalui penerapan *Early Stopping*.

#### Pembagian data 80%:10%:10%

Pada tabel ditampilkan hasil pelatihan pada pembagian 80% untuk data *training*, 10% untuk data *validation*, 10% untuk data testing dengan pelatihan selama 40 *epoch*.

Tabel 4. 4 Eksperimen Pelatihan Menggunakan 40 Epoch Pada Pembagian 80:10:10%

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 80%:10%:10% | | | | |
| Epoch | Akurasi | Loss | Akurasi Validasi | Loss Validasi |
| 1 | 0.6989 | 0.7037 | 0.8900 | 0.2909 |
| 2 | 0.9068 | 0.2421 | 0.9176 | 0.2496 |
| 3 | 0.9104 | 0.2450 | 0.8967 | 0.2422 |
| 4 | 0.9193 | 0.2633 | 0.7800 | 0.4844 |
| 5 | 0.9141 | 0.2215 | 0.8767 | 0.2987 |
| 6 | 0.9334 | 0.1791 | 0.8767 | 0.2793 |
| 7 | 0.9182 | 0.2121 | 0.8400 | 0.3782 |
| 8 | 0.9199 | 0.2043 | 0.8800 | 0.3155 |

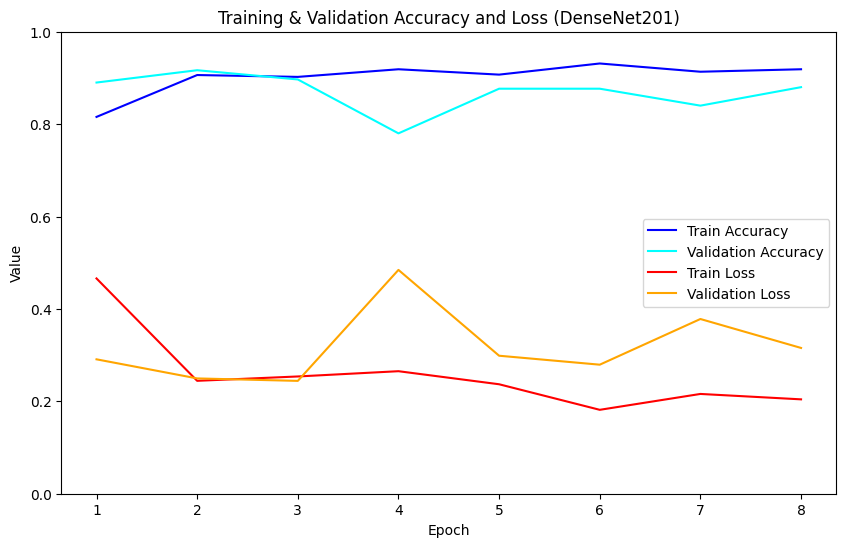
Pelatihan pertama menggunakan skema pembagian data 80%:10:10% seperti yang disajikan pada tabel . Kita dapat melihat bahwa *epoch* pertama, model mencapai Akurasi sebesar 0.6989 dan Loss 0.7037 yang artinya model sudah mampu mengenali pola dasar dalam data dengan cukup baik, meskipun masih terdapat kesalahan prediksi yang cukup tinggi, yang sebagaimana ditunjukkan nilai loss 0.7037. Akan tetapi, untuk Akurasi Validasi sebesar 0.8900 dan Loss Validasi sebesar 0.2909 menunjukkan bahwa model justru berkinerja lebih baik pada data validasi dibandingkan data latih pada epoch pertama. Hal ini menunjukkan bahwa model belum sepenuhnya belajar pola-pola dari data latih, namun sudah mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat.

Pada *epoch* ke kedua, model mengalami peningkatan dengan akurasi 0.9068 dan Loss 0.2421. Hal ini menandakan bahwa model telah mempelajari pola dengan yang sangat baik. Model mulai mampu melakukan klasifikasi dengan tingkat kesalahan yang lebih rendah dibandingkan dengan *epoch* pertama. Selain itu, nilai akurasi validasi sebesar 0.9176 dan loss validasi 0.2496 menunjukkan bahwa performa model pada data validasi sangat sejalan dengan performa pada data pelatihan.

Namun, pada *epoch* ketiga, mengalami penurunan dengan nilai akurasi 0.8879 dan loss 0.2837. Penurunan ini menunjukkan bahwa meskipun model masih memiliki kemampuan klasifikasi yang baik, terdapat fluktuasi dalam proses pelatihan yang mungkin disebabkan oleh beberapa faktor seperti salah satunya awal *overfitting*. Meskipun demikian, nilai akurasi validasi tetap stabil diangkat 0.9062 dan nilai loss validasi 0.2800, yang menunjukkan bahwa model masih mampu melakukan generalisasi dengan baik terhadap data validasi.

Selanjutnya, model mengalami peningkatan yang stabil dan konsisten dengan puncak performa pada *epoch* ke empat dengan nilai akurasi 0.9062 dan nilai loss 0.2506. Namun, nilai akurasi validasi justru menurun menjadi 0.8125 dan nilai loss validasi 0.4268. Hal ini menunjukkan adanya kemungkinan awal *overfitting*, di mana model mulai terlalu menyesuaikan diri dengan data pelatihan sehingga kinerjanya pada data validasi menurun.

Setelah itu, performa model tetap stabil dan tidak memberikan kenaikan secara signifikan hingga pada *epoch* 16, meskipun sebelumnya telah direncanakan berlangsung hingga *epoch* 40. Hal ini disebabkan oleh *early stopping* yang menandakan tidak adanya peningkatan yang jelas atau signifikan, sehingga pelatihan dihentikan lebih awal untuk mencegah *overfitting* dan menghemat sumber daya komputasi.



Gambar 4. 9 Grafik Akurasi dan Loss Pada Pelatihan dan Validasi

Grafik pada gambar menunjukkan performa model selama proses pelatihan, yang mencakup akurasi dan loss pada data pelatihan serta validasi. Pelatihan dilakukan dengan pembagian data sebesar 80% untuk training, 10% untuk validasi, dan 10% untuk testing. Pada gambar sebelah kiri sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai akurasi. Sedangkan gambar sebelah kanan sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai loss.

Pada gambar sebelah kiri, grafik yang menunjukkan perkembangan akurasi. Garis berwarna oranye menunjukkan nilai akurasi pada data validasi, sedangkan garis berwarna biru menunjukkan akurasi pada data pelatihan. Pada epoch ke-2, terlihat bahwa akurasi pelatihan mencapai sekitar 0.90 dan akurasi validasi juga tinggi, menunjukkan bahwa model mampu mengenali pola dengan cukup baik pada tahap awal. Kemudian pada epoch ke-4, terjadi peningkatan yang signifikan pada akurasi pelatihan hingga mendekati 0.93 dan akurasi validasi pun turut meningkat, yang menandakan bahwa model sedang berada dalam fase belajar yang efektif. Namun, seiring berjalannya epoch, akurasi validasi yang ditunjukkan oleh garis oranye menjadi fluktuatif dan menurun drastis di akhir pelatihan, sementara akurasi pelatihan terus meningkat. Pola ini menunjukkan adanya indikasi overfitting, di mana model terlalu menyesuaikan diri terhadap data pelatihan dan kehilangan kemampuan generalisasi terhadap data baru.

Sementara itu, pada gambar sebelah kanan, grafik yang menunjukkan perkembangan loss. Garis berwarna oranye mewakili nilai loss pada data validasi, sedangkan garis berwarna biru menunjukkan loss pada data pelatihan. Pada epoch ke-2, loss pelatihan turun tajam hingga sekitar 0.26, dan loss validasi pun menurun menjadi sekitar 0.29, menunjukkan efisiensi model dalam meminimalkan kesalahan pada tahap awal pelatihan. Pada epoch ke-4, nilai loss pelatihan kembali menurun, dan meskipun loss validasi mengalami sedikit peningkatan, keduanya masih berada pada nilai yang relatif rendah. Namun, setelah epoch-epoch berikutnya, loss validasi menunjukkan pola yang tidak stabil dan bahkan meningkat tajam di akhir pelatihan, sementara loss pelatihan terus menurun. Hal ini memperkuat indikasi bahwa model mengalami overfitting, yaitu terlalu fokus pada data pelatihan hingga kinerjanya memburuk pada data validasi.

#### Pembagian data 70%:20%:10%

Pada tabel ditampilkan hasil pelatihan pada pembagian 70% untuk data *training*, 20% untuk data *validation*, 10% untuk data testing dengan pelatihan selama 40 *epoch*.

Tabel 4. 5 Eksperimen Pelatihan Menggunakan 40 Epoch Pada Pembagian 70:20:10%

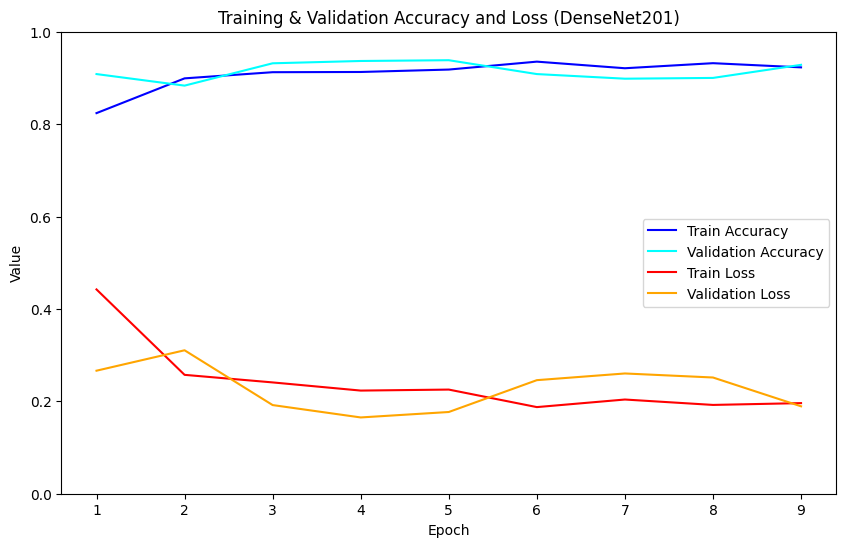
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 70%:20%:10% | | | | |
| Epoch | Akurasi | Loss | Akurasi Validasi | Loss Validasi |
| 1 | 0.7216 | 0.6458 | 0.9083 | 0.2663 |
| 2 | 0.8905 | 0.2836 | 0.8833 | 0.3105 |
| 3 | 0.9035 | 0.2634 | 0.9317 | 0.1919 |
| 4 | 0.9137 | 0.2387 | 0.9367 | 0.1651 |
| 5 | 0.9207 | 0.2248 | 0.9383 | 0.1769 |
| 6 | 0.9341 | 0.1916 | 0.9083 | 0.2458 |
| 7 | 0.9134 | 0.2180 | 0.8983 | 0.2603 |
| 8 | 0.9319 | 0.1898 | 0.9000 | 0.2516 |
| 9 | 0.9211 | 0.1982 | 0.9283 | 0.1892 |

Pelatihan pertama menggunakan skema pembagian data 70%:20:10% seperti yang disajikan pada tabel . Kita dapat melihat bahwa *epoch* pertama, model mencapai Akurasi sebesar 0.7290 dan Loss 0.6190 yang artinya model sudah mampu mengenali pola dasar dalam data dengan baik, meskipun masih terdapat kesalahan prediksi yang cukup tinggi, yang sebagaimana ditunjukkan nilai loss 0.6190. Akan tetapi, untuk Akurasi Validasi sebesar 0.9125 dan Loss Validasi sebesar 0.2331 menunjukkan bahwa model justru berkinerja lebih baik pada data validasi dibandingkan data latih pada epoch pertama. Hal ini menunjukkan bahwa model belum sepenuhnya belajar pola-pola dari data latih, namun sudah mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat.

Pada *epoch* ke kedua, model mengalami peningkatan dengan akurasi 0.9007 dan Loss 0.2628. Hal ini menandakan bahwa model telah mempelajari pola dengan yang sangat baik. Model mulai mampu melakukan klasifikasi dengan tingkat kesalahan yang lebih rendah dibandingkan dengan *epoch* pertama. Namun, nilai akurasi validasi justru menurun menjadi 0.8000 dan loss validasi 0.4944. Penurunan ini mengindikasikan adanya fluktuasi dalam proses pelatihan, yang mungkin disebabkan oleh faktor-faktor seperti awal mula terjadinya *overfitting*, di mana model mulai terlalu menyesuaikan diri terhadap data pelatihan.

Selanjutnya, model mengalami peningkatan yang stabil dan konsisten dengan puncak performa pada *epoch* ke tiga dengan nilai akurasi 0.9196 dan nilai loss 0.2365. Namun, nilai akurasi validasi menurun menjadi 0.7188 dan nilai loss validasi 0.7299. Hal ini menunjukkan adanya kemungkinan awal *overfitting*, di mana model mulai terlalu menyesuaikan diri dengan data pelatihan sehingga kinerjanya pada data validasi menurun.

Setelah itu, performa model tetap stabil dan tidak memberikan kenaikan secara signifikan hingga pada *epoch* 10, meskipun sebelumnya telah direncanakan berlangsung hingga *epoch* 40. Hal ini disebabkan oleh *early stopping* yang menandakan tidak adanya peningkatan yang jelas atau signifikan, sehingga pelatihan dihentikan lebih awal untuk mencegah *overfitting* dan menghemat sumber daya komputasi.



Gambar 4. 10 Grafik Akurasi dan Loss Pada Pelatihan dan Validasi

Grafik pada gambar menunjukkan performa model selama proses pelatihan, yang mencakup akurasi dan loss pada data pelatihan serta validasi. Pelatihan dilakukan dengan pembagian data sebesar 70% untuk training, 20% untuk validasi, dan 10% untuk testing. Pada gambar sebelah kiri sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai akurasi. Sedangkan gambar sebelah kanan sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai loss.

Pada gambar kiri, grafik yang menunjukkan perkembangan akurasi. Garis berwarna biru menunjukkan akurasi data pelatihan, yang meningkat secara konsisten dari awal hingga akhir epoch, dan cenderung stabil di atas 0.92 pada epoch-4 hingga epoch-10. Sedangkan garis berwarna oranye menunjukkan akurasi data validasi, yang mengalami fluktuasi signifikan. Pada epoch ke-2, akurasi validasi turun drastis hingga sekitar 0.72, menunjukkan performa buruk pada data validasi. Namun, pada epoch ke-4, akurasi validasi mengalami peningkatan tajam hingga hampir menyamai akurasi pelatihan, sebelum kembali fluktuatif pada epoch-epoch berikutnya. Fluktuasi ini menandakan bahwa model kemungkinan mengalami overfitting, yaitu terlalu cocok terhadap data pelatihan namun kurang mampu melakukan generalisasi terhadap data baru.

Pada gambar kanan, grafik yang menunjukkan perkembangan loss. Garis berwarna biru menunjukkan nilai loss pada data pelatihan, yang menurun secara konsisten dan stabil pada kisaran rendah mulai dari epoch ke-3 hingga ke-10. Hal ini menunjukkan bahwa model semakin baik dalam mempelajari pola dari data pelatihan. Sebaliknya, garis berwarna oranye menunjukkan nilai loss pada data validasi, yang sangat fluktuatif. Pada epoch ke-2, nilai loss validasi meningkat tajam mendekati 0.75, menunjukkan bahwa model gagal memprediksi data validasi dengan baik saat itu. Namun, pada epoch ke-4, loss validasi turun drastis, mencerminkan adanya perbaikan performa terhadap data validasi, sebelum kembali mengalami fluktuasi pada epoch berikutnya. Pola ini menunjukkan bahwa meskipun model cukup baik dalam pelatihan, ia masih kesulitan dalam melakukan generalisasi yang stabil terhadap data validasi.

#### Pembagian data 70%:15%:15%

Pada tabel ditampilkan hasil pelatihan pada pembagian 70% untuk data *training*, 15% untuk data *validation*, 15% untuk data testing dengan pelatihan selama 40 *epoch*.

Tabel 4. 6 Eksperimen Pelatihan Menggunakan 40 Epoch Pada Pembagian 70:15:15%

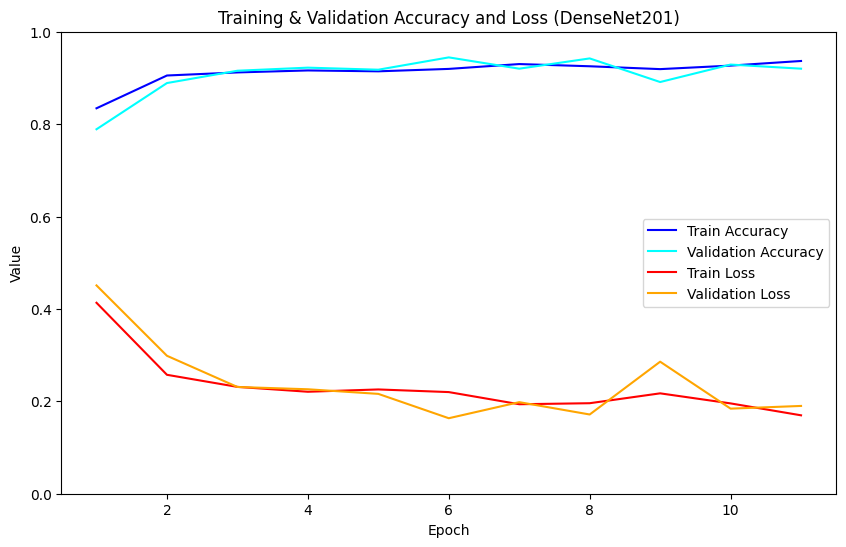
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 70%:15%:15% | | | | |
| Epoch | Akurasi | Loss | Akurasi Validasi | Loss Validasi |
| 1 | 0.7317 | 0.5913 | 0.7889 | 0.4509 |
| 2 | 0.9114 | 0.2520 | 0.8889 | 0.2986 |
| 3 | 0.9083 | 0.2360 | 0.9156 | 0.2311 |
| 4 | 0.9114 | 0.2247 | 0.9222 | 0.2261 |
| 5 | 0.9190 | 0.2187 | 0.9178 | 0.2161 |
| 6 | 0.9264 | 0.2130 | 0.9444 | 0.1634 |
| 7 | 0.9428 | 0.1783 | 0.9200 | 0.1982 |
| 8 | 0.9286 | 0.1863 | 0.9422 | 0.1715 |
| 9 | 0.9166 | 0.2165 | 0.8911 | 0.2858 |
| 10 | 0.9267 | 0.1947 | 0.9289 | 0.1842 |
| 11 | 0.9461 | 0.1559 | 0.9200 | 0.1900 |

Pelatihan pertama menggunakan skema pembagian data 70%:15:15% seperti yang disajikan pada tabel . Kita dapat melihat bahwa *epoch* pertama, model mencapai Akurasi sebesar 0.7580 dan Loss 0.5843 yang artinya model sudah mampu mengenali pola dasar dalam data dengan baik, meskipun masih terdapat kesalahan prediksi yang cukup tinggi, yang sebagaimana ditunjukkan nilai loss 0.5843. Akan tetapi, untuk Akurasi Validasi sebesar 0.7312 dan Loss Validasi sebesar 0.6647 menunjukkan bahwa model mulai mampu mengenali pola dalam data, meskipun performa pada data validasi sedikit lebih rendah dibandingkan data pelatihan.

Pada *epoch* ke kedua, model mengalami peningkatan dengan akurasi 0.9082 dan Loss 0.2628. Hal ini menandakan bahwa model telah mempelajari pola dengan yang sangat baik. Model mulai mampu melakukan klasifikasi dengan tingkat kesalahan yang lebih rendah dibandingkan dengan *epoch* pertama. Selain itu, nilai akurasi validasi sebesar 0.9187 dan loss validasi 0.2543 menunjukkan bahwa performa model pada data validasi sangat sejalan dengan performa pada data pelatihan.

Selanjutnya, model mengalami peningkatan yang stabil dan konsisten dengan puncak performa pada *epoch* ke empat dengan nilai akurasi 0.9196 dan nilai loss 0.2147. Akurasi validasi pun menunjukkan peningkatan yang konsisten, dengan nilai sebesar 0.9563 dan loss validasi sebesar 0.1784. Hal ini menunjukkan bahwa model tidak hanya belajar dengan baik pada data pelatihan, tetapi juga mampu melakukan generalisasi yang sangat baik terhadap data validasi.

Setelah itu, performa model tetap stabil dan tidak memberikan kenaikan secara signifikan hingga pada *epoch* 10, meskipun sebelumnya telah direncanakan berlangsung hingga *epoch* 40. Hal ini disebabkan oleh *early stopping* yang menandakan tidak adanya peningkatan yang jelas atau signifikan, sehingga pelatihan dihentikan lebih awal untuk mencegah *overfitting* dan menghemat sumber daya komputasi.



Gambar 4. 11 Grafik Akurasi dan Loss Pada Pelatihan dan Validasi

Grafik pada gambar menunjukkan performa model selama proses pelatihan, yang mencakup akurasi dan loss pada data pelatihan serta validasi. Pelatihan dilakukan dengan pembagian data sebesar 70% untuk training, 15% untuk validasi, dan 15% untuk testing. Pada gambar sebelah kiri sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai akurasi. Sedangkan gambar sebelah kanan sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai loss.

Pada gambar sebelah kiri, grafik yang menunjukkan perkembangan akurasi. Garis berwarna oranye menunjukkan nilai akurasi pada data validasi, sedangkan garis berwarna biru menunjukkan akurasi pada data pelatihan. Pada epoch ke-2, akurasi validasi meningkat tajam dan menyamai akurasi pelatihan di angka sekitar 0.91, menunjukkan bahwa model mulai mampu mengenali pola pada data validasi. Kemudian pada epoch ke-4, akurasi validasi mengalami peningkatan signifikan hingga mencapai lebih dari 0.95, lebih tinggi dari akurasi pelatihan, yang mengindikasikan model memiliki generalisasi yang sangat baik pada titik tersebut. Setelah itu, meskipun terjadi sedikit fluktuasi, akurasi validasi tetap berada pada kisaran tinggi di atas 0.90.

Sementara itu, pada gambar sebelah kanan, grafik yang menunjukkan perkembangan loss. Garis berwarna oranye mewakili nilai loss pada data validasi, sedangkan garis berwarna biru menunjukkan loss pada data pelatihan. Pada epoch ke-2, loss validasi menurun drastis dari sebelumnya, mendekati nilai loss pelatihan, menunjukkan peningkatan performa generalisasi model. Pada epoch ke-4, loss validasi kembali menurun dan mencapai titik terendahnya, mencerminkan kemampuan model dalam memprediksi data validasi dengan baik. Meskipun setelahnya terdapat sedikit fluktuasi pada loss validasi, tren keseluruhan tetap stabil dan rendah. Loss pelatihan sendiri terus mengalami penurunan yang konsisten dari awal hingga akhir epoch, menandakan proses pembelajaran model berjalan baik.

#### Pembagian data 60%:30%:10%

Pada tabel ditampilkan hasil pelatihan pada pembagian 60% untuk data *training*, 30% untuk data *validation*, 10% untuk data testing dengan pelatihan selama 40 *epoch*.

Tabel 4. 7 Eksperimen Pelatihan Menggunakan 40 Epoch Pada Pembagian 60:30:10

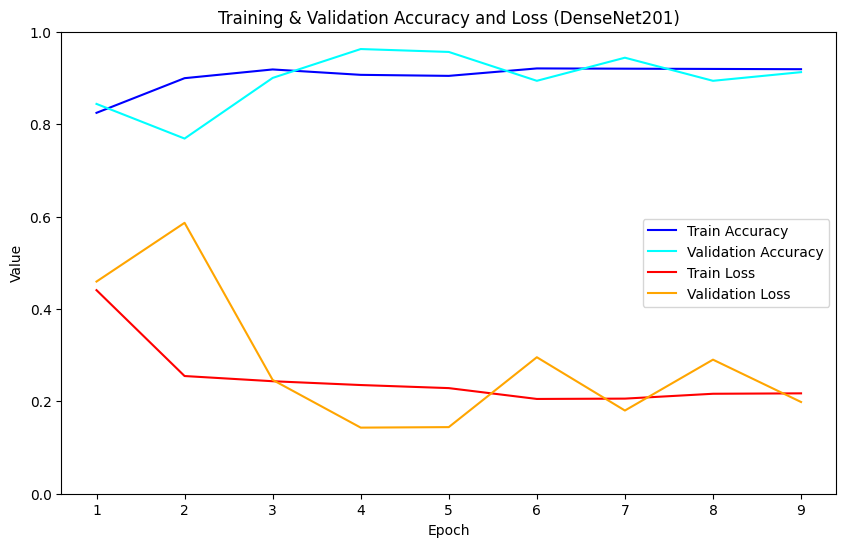
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 60%:30%:10% | | | | |
| Epoch | Akurasi | Loss | Akurasi Validasi | Loss Validasi |
| 1 | 0.7424 | 0.6214 | 0.8438 | 0.4592 |
| 2 | 0.8873 | 0.2885 | 0.7688 | 0.5864 |
| 3 | 0.9157 | 0.2571 | 0.9000 | 0.2464 |
| 4 | 0.9082 | 0.2342 | 0.9625 | 0.1431 |
| 5 | 0.9071 | 0.2244 | 0.9563 | 0.1442 |
| 6 | 0.9206 | 0.2076 | 0.8938 | 0.2954 |
| 7 | 0.9239 | 0.1960 | 0.9438 | 0.1801 |
| 8 | 0.9193 | 0.2044 | 0.8938 | 0.2902 |
| 9 | 0.9240 | 0.2099 | 0.9125 | 0.1985 |

Pelatihan pertama menggunakan skema pembagian data 60%:30:10% seperti yang disajikan pada tabel . Kita dapat melihat bahwa *epoch* pertama, model mencapai Akurasi sebesar 0.7112 dan Loss 0.6818 yang artinya model sudah mampu mengenali pola dasar dalam data dengan cukup baik, meskipun masih terdapat kesalahan prediksi yang cukup tinggi, yang sebagaimana ditunjukkan nilai loss 0.6818. Akan tetapi, untuk Akurasi Validasi sebesar 0.9125 dan Loss Validasi sebesar 0.2503 menunjukkan bahwa model justru berkinerja lebih baik pada data validasi dibandingkan data latih pada epoch pertama. Hal ini menunjukkan bahwa model belum sepenuhnya belajar pola-pola dari data latih, namun sudah mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat.

Pada *epoch* ke kedua, model mengalami peningkatan dengan akurasi 0.9001 dan Loss 0.2784. Hal ini menandakan bahwa model telah mempelajari pola dengan yang sangat baik. Model mulai mampu melakukan klasifikasi dengan tingkat kesalahan yang lebih rendah dibandingkan dengan *epoch* pertama. Selain itu, nilai akurasi validasi sebesar 0.9187 dan loss validasi 0.2259 menunjukkan bahwa performa model pada data validasi sangat sejalan dengan performa pada data pelatihan.

Selanjutnya, model mengalami peningkatan yang stabil dan konsisten dengan puncak performa pada *epoch* ke tiga dengan nilai akurasi 0.9111 dan nilai loss 0.2434. Namun, nilai akurasi validasi menurun menjadi 0.8875 dan nilai loss validasi 0.2476. Hal ini menunjukkan adanya kemungkinan awal *overfitting*, di mana model mulai terlalu menyesuaikan diri dengan data pelatihan sehingga kinerjanya pada data validasi menurun.

Setelah itu, performa model tetap stabil dan tidak memberikan kenaikan secara signifikan hingga pada *epoch* 16, meskipun sebelumnya telah direncanakan berlangsung hingga *epoch* 40. Hal ini disebabkan oleh *early stopping* yang menandakan tidak adanya peningkatan yang jelas atau signifikan, sehingga pelatihan dihentikan lebih awal untuk mencegah *overfitting* dan menghemat sumber daya komputasi.



Tabel 4. 8 Grafik Akurasi dan Loss Pada Pelatihan dan Validasi

Grafik pada gambar menunjukkan performa model selama proses pelatihan, yang mencakup akurasi dan loss pada data pelatihan serta validasi. Pelatihan dilakukan dengan pembagian data sebesar 60% untuk training, 30% untuk validasi, dan 10% untuk testing. Pada gambar sebelah kiri sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai akurasi. Sedangkan gambar sebelah kanan sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai loss.

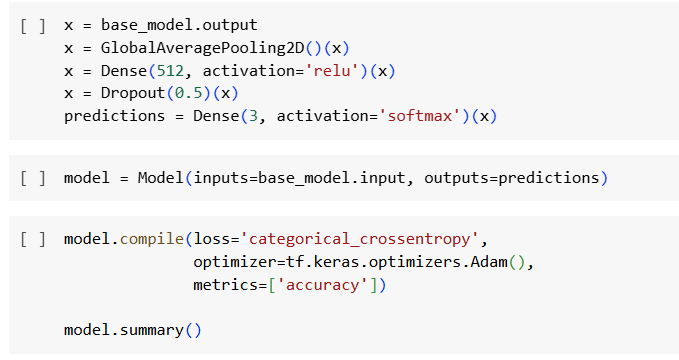
Pada gambar sebelah kiri, grafik menunjukkan perkembangan akurasi selama proses pelatihan model. Garis berwarna oranye menunjukkan nilai akurasi pada data validasi, sedangkan garis berwarna biru menunjukkan akurasi pada data pelatihan. Pada epoch ke-2, akurasi validasi mengalami sedikit penurunan dibandingkan epoch sebelumnya, namun kemudian mengalami peningkatan pada epoch ke-4, yang menunjukkan adanya perbaikan performa sementara pada data validasi. Meskipun demikian, akurasi validasi secara keseluruhan terlihat berfluktuasi dan tidak stabil, mengindikasikan bahwa model mungkin mengalami kesulitan dalam melakukan generalisasi terhadap data baru. Di sisi lain, akurasi pelatihan meningkat secara konsisten dan stabil di atas 92% setelah beberapa epoch, menunjukkan bahwa model mampu belajar dengan baik dari data pelatihan.

Sementara itu, pada gambar sebelah kanan, grafik menunjukkan perkembangan loss selama proses pelatihan. Garis berwarna oranye mewakili nilai loss pada data validasi, sedangkan garis berwarna biru menunjukkan loss pada data pelatihan. Pada epoch ke-2, loss validasi berada di sekitar 0.26 dan kemudian mengalami penurunan pada epoch ke-4 menjadi sekitar 0.20, menandakan adanya peningkatan performa. Namun, nilai loss validasi secara umum tampak tidak stabil dan mengalami lonjakan yang signifikan pada epoch ke-13, yang menunjukkan indikasi overfitting. Sebaliknya, loss pada data pelatihan mengalami penurunan yang konsisten seiring bertambahnya epoch, menandakan bahwa model mampu mengoptimalkan performanya terhadap data pelatihan dengan baik.

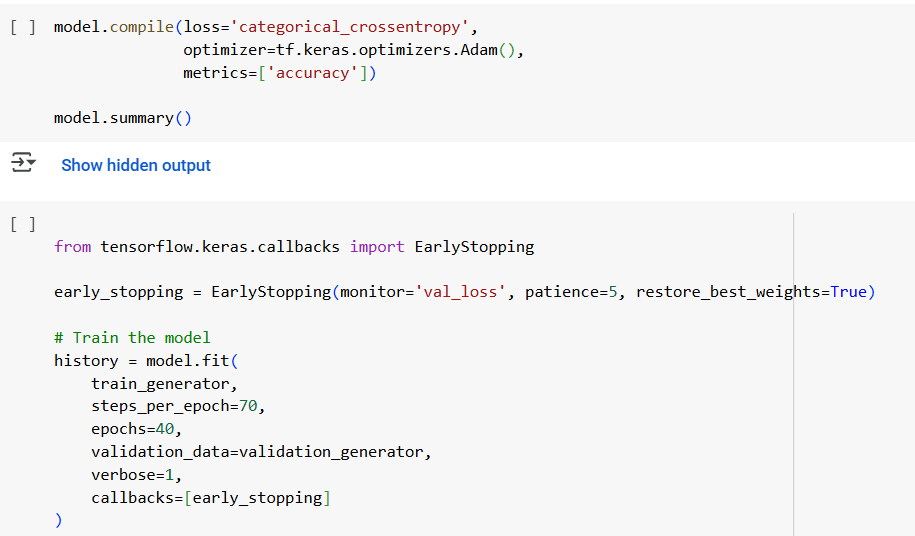
### 4.5.2 Pembuatan Model Xception



Dalam penelitian ini, digunakan arsitektur xception sebagai model dasar untuk klasifikasi citra x-ray paru-paru. Pada Gambar 4.5 ditunjukkan bahwa model ini diimplementasikan melalui pustaka *keras.applications*, dengan bobot awal yang diambil dari hasil pelatihan sebelumnya menggunakan dataset ImageNet, sehingga model tidak perlu dilatih dari awal. Selain itu, parameter *include\_top* disetel ke nilai false untuk menghilangkan bagian akhir dari arsitektur Xception yang berfungsi sebagai fully connected layer bawaan model. Hal ini dilakukan agar peneliti dapat menambahkan layer klasifikasi yang disesuaikan dengan jumlah kelas target dalam penelitian, yaitu tiga kelas: normal, pneumonia, dan tuberculosis. Ukuran input citra yang digunakan adalah 512 × 512 piksel dengan tiga saluran warna, yaitu *red*, *green*, dan *blue* (RGB), yang sesuai dengan format standar citra warna. Untuk menghindari pelatihan ulang terhadap seluruh parameter model dasar, nilai atribut *trainable* pada objek model dasar diatur menjadi *false*. Dengan konfigurasi ini, seluruh layer pada Xception tidak akan diperbarui selama proses pelatihan, sehingga model hanya berfungsi sebagai ekstraktor fitur.



Pada Gambar 4.6, di atas output dari xception, ditambahkan beberapa layer klasifikasi lanjutan. Proses ini diawali dengan penerapan layer *Global Average Pooling 2D* untuk meratakan fitur hasil ekstraksi. Selanjutnya, digunakan satu layer *Dense* dengan 512 neuron dan fungsi aktivasi *Rectified Linear Unit* (ReLU), yang kemudian diikuti oleh layer *dropout* dengan nilai 0,5 sebagai strategi untuk mengurangi risiko *overfitting*. Sebagai layer output, digunakan layer *dense* dengan tiga neuron dan fungsi aktivasi *softmax* untuk mengklasifikasikan citra ke dalam tiga kategori: normal, pneumonia, dan tuberculosis.



Pada Gambar 4.7, model dikompilasi menggunakan fungsi loss categorical\_crossentropy, algoritma optimasi Adam, dan metrik evaluasi berupa accuracy. Untuk mencegah *overfitting* dan pelatihan berlebihan, diterapkan teknik *Early Stopping* yang memantau nilai *validation loss* dan menghentikan proses pelatihan apabila tidak terjadi peningkatan dalam lima *epoch* berturut-turut. Parameter restore\_best\_weights=True memastikan bahwa bobot model akan dikembalikan ke kondisi terbaik sepanjang proses pelatihan. Proses pelatihan dilakukan dengan memanfaatkan fungsi model.fit() dengan konfigurasi *steps per epoch* sebesar 70, jumlah maksimum *epoch* sebanyak 40, dan data validasi diambil dari parameter validation\_generator sebanyak 5 langkah setiap *epoch*. Dengan konfigurasi ini, model diharapkan mampu melakukan pembelajaran secara optimal terhadap data latih, divalidasi secara berkala pada setiap epoch, serta meminimalkan risiko *overfitting* melalui penerapan *Early Stopping*.

#### Pembagian data 80%:10%:10%

Pada tabel ditampilkan hasil pelatihan pada pembagian 80% untuk data *training*, 10% untuk data *validation*, 10% untuk data testing dengan pelatihan selama 40 *epoch*.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 80%:10%:10% | | | | |
| Epoch | Akurasi | Loss | Akurasi Validasi | Loss Validasi |
| 1 | 0.7605 | 0.5412 | 0.8600 | 0.3546 |
| 2 | 0.9075 | 0.2970 | 0.8867 | 0.3103 |
| 3 | 0.8965 | 0.2612 | 0.8133 | 0.4479 |
| 4 | 0.8656 | 0.3406 | 0.8100 | 0.4285 |
| 5 | 0.9194 | 0.2145 | 0.8733 | 0.3070 |
| 6 | 0.8994 | 0.2349 | 0.9233 | 0.2168 |
| 7 | 0.9068 | 0.2352 | 0.8567 | 0.3720 |
| 8 | 0.9147 | 0.2215 | 0.7967 | 0.4840 |
| 9 | 0.9182 | 0.2320 | 0.8767 | 0.3167 |
| 10 | 0.9497 | 0.1663 | 0.8400 | 0.3567 |
| 11 | 0.9227 | 0.1920 | 0.8133 | 0.4325 |

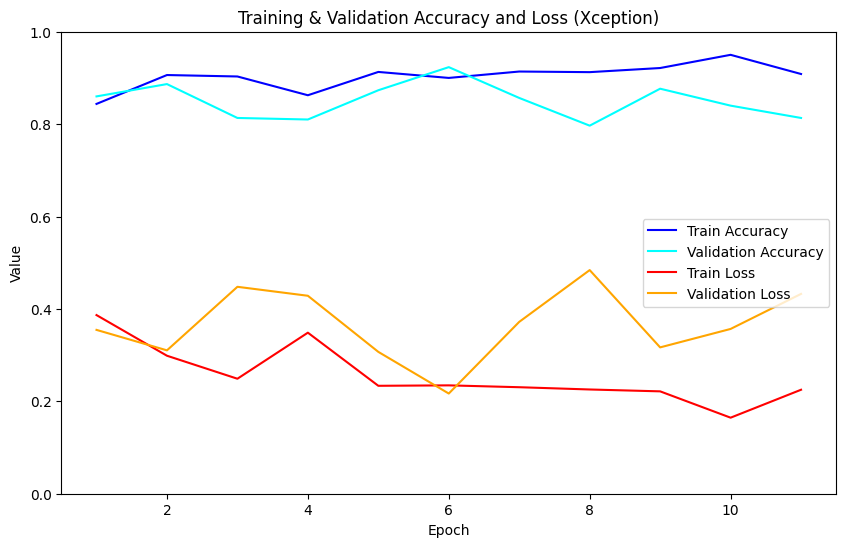
Pelatihan pertama menggunakan skema pembagian data 80%:10:10% seperti yang disajikan pada tabel . Kita dapat melihat bahwa *epoch* pertama, model mencapai Akurasi sebesar 0.7237 dan Loss 0.6127 yang artinya model sudah mampu mengenali pola dasar dalam data dengan cukup baik, meskipun masih terdapat kesalahan prediksi yang cukup tinggi, yang sebagaimana ditunjukkan nilai loss 0.6127. Akan tetapi, untuk Akurasi Validasi sebesar 0.8562 dan Loss Validasi sebesar 0.3070 menunjukkan bahwa model justru berkinerja lebih baik pada data validasi dibandingkan data latih pada epoch pertama. Hal ini menunjukkan bahwa model belum sepenuhnya belajar pola-pola dari data latih, namun sudah mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat.

Pada *epoch* ke kedua, model mengalami peningkatan dengan akurasi 0.8253 dan Loss 0.3643. Hal ini menandakan bahwa model telah mempelajari pola dengan baik. Model mulai mampu melakukan klasifikasi dengan tingkat kesalahan yang lebih rendah dibandingkan dengan *epoch* pertama. Namun, nilai akurasi validasi justru menurun menjadi 0.8062 dan loss validasi 0.4528. Penurunan ini mengindikasikan adanya fluktuasi dalam proses pelatihan, yang mungkin disebabkan oleh faktor-faktor seperti awal mula terjadinya *overfitting*, di mana model mulai terlalu menyesuaikan diri terhadap data pelatihan.

Pada *epoch* ketiga, model mengalami peningkatan dengan akurasi 0.8869 dan loss 0.2779. Peningkatan ini menunjukkan bahwa model semakin mampu mengenali pola dalam data pelatihan secara lebih akurat, dengan tingkat kesalahan yang semakin kecil. Selain itu, nilai akurasi validasi juga meningkat dibandingkan dengan epoch ke dua, yaitu mencapai 0.9062 dan nilai loss validasi 0.2436, yang menunjukkan bahwa model masih mampu melakukan generalisasi dengan baik terhadap data validasi. Peningkatan ini mengindikasikan bahwa model tidak hanya belajar dengan baik pada data pelatihan, tetapi juga menunjukkan generalisasi yang lebih baik terhadap data yang belum pernah dilihat sebelumnya.

Selanjutnya, model mengalami peningkatan yang stabil dan konsisten dengan puncak performa pada *epoch* ke empat dengan nilai akurasi 0.9056 dan nilai loss 0.2159. Namun, nilai akurasi validasi justru menurun menjadi 0.7375 dan nilai loss validasi 0.6250. Hal ini menunjukkan adanya kemungkinan awal *overfitting*, di mana model mulai terlalu menyesuaikan diri dengan data pelatihan sehingga kinerjanya pada data validasi menurun.

Setelah itu, performa model tetap stabil dan tidak memberikan kenaikan secara signifikan hingga pada *epoch* 8, meskipun sebelumnya telah direncanakan berlangsung hingga *epoch* 40. Hal ini disebabkan oleh *early stopping* yang menandakan tidak adanya peningkatan yang jelas atau signifikan, sehingga pelatihan dihentikan lebih awal untuk mencegah *overfitting* dan menghemat sumber daya komputasi.



Grafik pada gambar menunjukkan performa model selama proses pelatihan, yang mencakup akurasi dan loss pada data pelatihan serta validasi. Pelatihan dilakukan dengan pembagian data sebesar 80% untuk training, 10% untuk validasi, dan 10% untuk testing. Pada gambar sebelah kiri sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai akurasi. Sedangkan gambar sebelah kanan sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai loss.

Pada gambar kiri, garis berwarna biru menunjukkan akurasi pada data pelatihan, sementara garis berwarna oranye menunjukkan akurasi pada data validasi. Terlihat bahwa akurasi pelatihan meningkat secara konsisten dari epoch ke-0 hingga epoch ke-5, bahkan mencapai lebih dari 0.93 pada epoch ke-5. Ini menandakan bahwa model berhasil mempelajari pola dari data pelatihan dengan baik. Namun, akurasi validasi yang ditunjukkan oleh garis oranye bersifat fluktuatif, naik-turun tajam dari satu epoch ke epoch berikutnya. Sebagai contoh, pada epoch ke-2 akurasi validasi sempat tinggi hamper mendekati pelatihan, tetapi langsung turun drastis pada epoch ke-3, lalu naik lagi pada epoch ke-5, dan turun kembali pada epoch ke-6. Fluktuasi ini menunjukkan bahwa model mengalami ketidakstabilan dalam generalisasi, kemungkinan karena jumlah epoch yang masih sedikit atau model terlalu sensitif terhadap data validasi.

Pada gambar kanan, garis berwarna biru menunjukkan nilai loss pada data pelatihan, sementara garis oranye menunjukkan loss pada data validasi. Loss pelatihan menunjukkan tren penurunan secara konsisten hingga mencapai nilai mendekati 0.22 pada epoch ke-5, yang menandakan bahwa kesalahan model terhadap data pelatihan terus berkurang. Namun, nilai loss validasi menunjukkan pola yang sangat tidak stabil. Contohnya, loss validasi turun hingga mendekati loss pelatihan pada epoch ke-2, tetapi meningkat tajam pada epoch ke-3 dan tetap fluktuatif hingga akhir pelatihan. Ini memperkuat dugaan bahwa model mengalami overfitting atau tidak cukup stabil pada data validasi, meskipun performa pada data pelatihan terlihat baik.

#### Pembagian data 70%:20%:10%

Pada tabel ditampilkan hasil pelatihan pada pembagian 70% untuk data *training*, 20% untuk data *validation*, 10% untuk data testing dengan pelatihan selama 40 *epoch*.

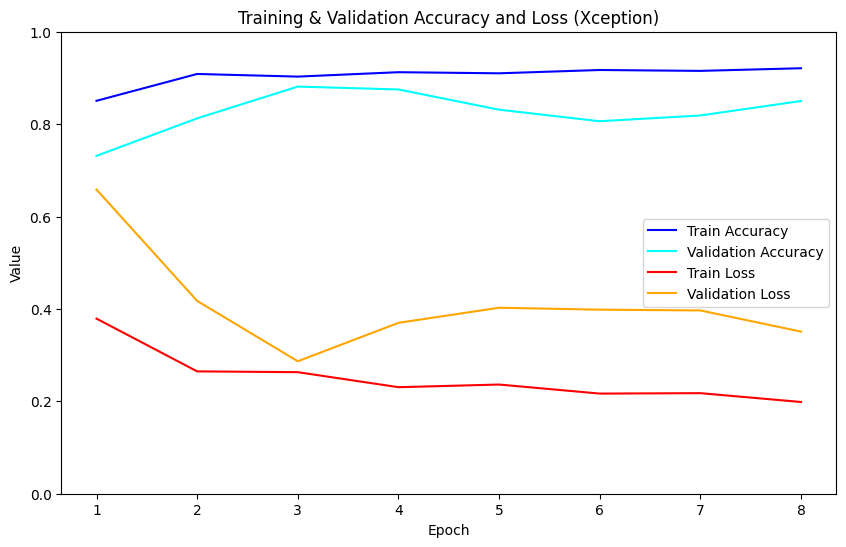
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 70%:20%:10% | | | | |
| Epoch | Akurasi | Loss | Akurasi Validasi | Loss Validasi |
| 1 | 0.7830 | 0.5258 | 0.7312 | 0.6584 |
| 2 | 0.9034 | 0.2728 | 0.8125 | 0.4174 |
| 3 | 0.9076 | 0.2564 | 0.8813 | 0.2867 |
| 4 | 0.9096 | 0.2337 | 0.8750 | 0.3699 |
| 5 | 0.9103 | 0.2339 | 0.8313 | 0.4026 |
| 6 | 0.9182 | 0.2080 | 0.8062 | 0.3984 |
| 7 | 0.9071 | 0.2348 | 0.8188 | 0.3967 |
| 8 | 0.9173 | 0.2096 | 0.8500 | 0.3509 |

Pelatihan pertama menggunakan skema pembagian data 70%:20:10% seperti yang disajikan pada tabel . Kita dapat melihat bahwa *epoch* pertama, model mencapai Akurasi sebesar 0.7201 dan Loss 0.6162 yang artinya model sudah mampu mengenali pola dasar dalam data dengan cukup baik, meskipun masih terdapat kesalahan prediksi yang cukup tinggi, yang sebagaimana ditunjukkan nilai loss 0.6162. Akan tetapi, untuk Akurasi Validasi sebesar 0.8562 dan Loss Validasi sebesar 0.3608 menunjukkan bahwa model justru berkinerja lebih baik pada data validasi dibandingkan data latih pada epoch pertama. Hal ini menunjukkan bahwa model belum sepenuhnya belajar pola-pola dari data latih, namun sudah mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat.

Pada *epoch* ke kedua, model mengalami peningkatan dengan akurasi 0.8986 dan Loss 0.2816. Hal ini menandakan bahwa model telah mempelajari pola dengan yang sangat baik. Model mulai mampu melakukan klasifikasi dengan tingkat kesalahan yang lebih rendah dibandingkan dengan *epoch* pertama. Namun, nilai akurasi validasi justru menurun menjadi 0.8250 dan loss validasi 0.4196. Penurunan ini mengindikasikan adanya fluktuasi dalam proses pelatihan, yang mungkin disebabkan oleh faktor-faktor seperti awal mula terjadinya *overfitting*, di mana model mulai terlalu menyesuaikan diri terhadap data pelatihan.

Selanjutnya, model mengalami peningkatan yang stabil dan konsisten dengan puncak performa pada *epoch* ke tiga dengan nilai akurasi 0.9110 dan nilai loss 0.2536. Akurasi validasi menunjukkan peningkatan dengan nilai sebesar 0.8438 dan loss validasi sebesar 0.3651. Hal ini menunjukkan bahwa model tidak hanya belajar dengan baik pada data pelatihan, tetapi juga mampu melakukan generalisasi yang sangat baik terhadap data validasi.

Setelah itu, performa model tetap stabil dan tidak memberikan kenaikan secara signifikan hingga pada *epoch* 13, meskipun sebelumnya telah direncanakan berlangsung hingga *epoch* 40. Hal ini disebabkan oleh *early stopping* yang menandakan tidak adanya peningkatan yang jelas atau signifikan, sehingga pelatihan dihentikan lebih awal untuk mencegah *overfitting* dan menghemat sumber daya komputasi.



Grafik pada gambar menunjukkan performa model selama proses pelatihan, yang mencakup akurasi dan loss pada data pelatihan serta validasi. Pelatihan dilakukan dengan pembagian data sebesar 70% untuk training, 20% untuk validasi, dan 10% untuk testing. Pada gambar sebelah kiri sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai akurasi. Sedangkan gambar sebelah kanan sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai loss.

Pada gambar sebelah kiri, grafik yang menunjukkan perkembangan akurasi. Garis berwarna oranye menunjukkan nilai akurasi pada data validasi, sedangkan garis berwarna biru menunjukkan akurasi pada data pelatihan. Akurasi pada data pelatihan meningkat secara konsisten dari awal hingga mencapai lebih dari 0.92, menunjukkan bahwa model mampu belajar dengan baik dari data pelatihan. Sebaliknya, akurasi pada data validasi terlihat fluktuatif dan tidak stabil. Sebagai contoh, pada epoch ke-2 akurasi validasi menurun ablei, kemudian pada epoch ke-4 mengalami peningkatan yang cukup signifikan. Namun setelah epoch ke-7, akurasi validasi ablei menurun dan tetap rendah, meskipun akurasi pelatihan terus meningkat. Hal ini mengindikasikan adanya kemungkinan overfitting, yaitu model terlalu menyesuaikan diri dengan data pelatihan dan kurang mampu melakukan generalisasi pada data baru.

Sementara itu, pada gambar sebelah kanan, grafik yang menunjukkan perkembangan loss. Garis berwarna oranye mewakili nilai loss pada data validasi, sedangkan garis berwarna biru menunjukkan loss pada data pelatihan. Loss pada data pelatihan terus menurun secara stabil sepanjang proses pelatihan, menandakan bahwa model berhasil meminimalkan kesalahan terhadap data pelatihan. Namun, loss pada data validasi menunjukkan pola yang tidak stabil. Misalnya, pada epoch ke-2, loss validasi meningkat tajam, lalu pada epoch ke-4 mengalami penurunan yang mencerminkan perbaikan performa. Meski demikian, secara umum, nilai loss validasi tetap mengalami fluktuasi yang cukup besar, bahkan naik pada beberapa epoch, meskipun loss pelatihan terus menurun. Ini memperkuat indikasi overfitting yang juga terlihat pada grafik akurasi.

#### Pembagian data 70%:15%:15%

Pada able ditampilkan hasil pelatihan pada pembagian 70% untuk data *training*, 15% untuk data *validation*, 15% untuk data testing dengan pelatihan selama 40 *epoch*.

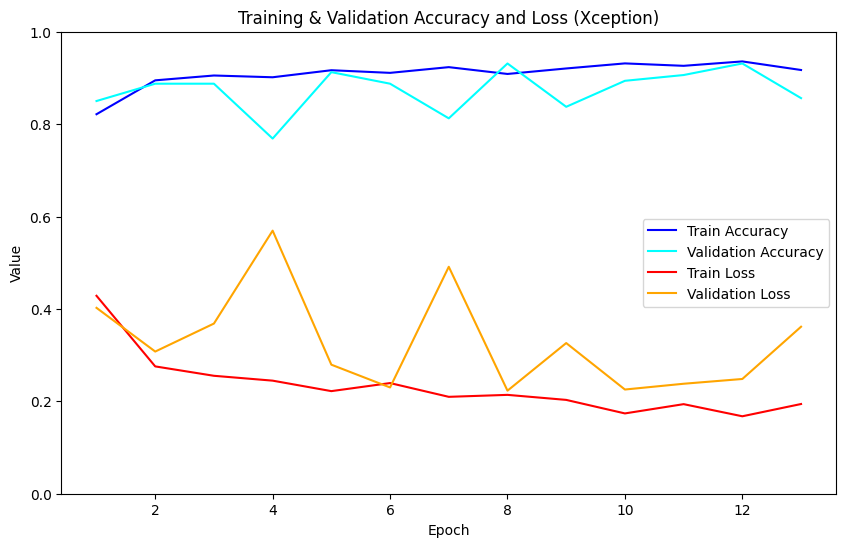
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 70%:15%:15% | | | | |
| Epoch | Akurasi | Loss | Akurasi Validasi | Loss Validasi |
| 1 | 0.7313 | 0.5809 | 0.8500 | 0.4023 |
| 2 | 0.8916 | 0.2745 | 0.8875 | 0.3077 |
| 3 | 0.8986 | 0.2536 | 0.8875 | 0.3683 |
| 4 | 0.8985 | 0.2496 | 0.7688 | 0.5694 |
| 5 | 0.9115 | 0.2313 | 0.9125 | 0.2793 |
| 6 | 0.9142 | 0.2246 | 0.8875 | 0.2300 |
| 7 | 0.9265 | 0.2138 | 0.8125 | 0.4912 |
| 8 | 0.9027 | 0.2271 | 0.9312 | 0.2230 |
| 9 | 0.9268 | 0.1911 | 0.8375 | 0.3261 |
| 10 | 0.9347 | 0.1709 | 0.8938 | 0.2255 |
| 11 | 0.9232 | 0.2086 | 0.9062 | 0.2380 |
| 12 | 0.9379 | 0.1647 | 0.9312 | 0.2484 |
| 13 | 0.9199 | 0.1852 | 0.8562 | 0.3615 |

Pelatihan pertama menggunakan skema pembagian data 70%:15:15% seperti yang disajikan pada table . Kita dapat melihat bahwa *epoch* pertama, model mencapai Akurasi sebesar 0.7347 dan Loss 0.6118 yang artinya model sudah mampu mengenali pola dasar dalam data dengan cukup baik, meskipun masih terdapat kesalahan prediksi yang cukup tinggi, yang sebagaimana ditunjukkan nilai loss 0.6118. Akan tetapi, untuk Akurasi Validasi sebesar 0.8062 dan Loss Validasi sebesar 0.4929 menunjukkan bahwa model justru berkinerja lebih baik pada data validasi dibandingkan data latih pada epoch pertama. Hal ini menunjukkan bahwa model belum sepenuhnya belajar pola-pola dari data latih, namun sudah mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat.

Pada *epoch* ke kedua, model mengalami peningkatan dengan akurasi 0.8866 dan Loss 0.2997. Hal ini menandakan bahwa model telah mempelajari pola dengan yang sangat baik. Model mulai mampu melakukan klasifikasi dengan tingkat kesalahan yang lebih rendah dibandingkan dengan *epoch* pertama. Namun, nilai akurasi validasi justru menurun menjadi 0.7688 dan loss validasi 0.4620. Penurunan ini mengindikasikan adanya fluktuasi dalam proses pelatihan, yang mungkin disebabkan oleh beberapa faktor seperti awal mula terjadinya *overfitting*, di mana model mulai terlalu menyesuaikan diri terhadap data pelatihan.

Selanjutnya, model mengalami peningkatan yang stabil dan konsisten pada *epoch* ke tiga dengan nilai akurasi 0.8990 dan nilai loss 0.2583. Akurasi validasi menunjukkan peningkatan dengan nilai sebesar 0.9187 dan loss validasi sebesar 0.2380. Hal ini menunjukkan bahwa model tidak hanya belajar dengan baik pada data pelatihan, tetapi juga mampu melakukan generalisasi yang sangat baik terhadap data validasi.

Setelah itu, performa model tetap stabil dan tidak memberikan kenaikan secara signifikan hingga pada *epoch* 8, meskipun sebelumnya telah direncanakan berlangsung hingga *epoch* 40. Hal ini disebabkan oleh *early stopping* yang menandakan tidak adanya peningkatan yang jelas atau signifikan, sehingga pelatihan dihentikan lebih awal untuk mencegah *overfitting* dan menghemat sumber daya komputasi.



Grafik pada gambar menunjukkan performa model selama proses pelatihan, yang mencakup akurasi dan loss pada data pelatihan serta validasi. Pelatihan dilakukan dengan pembagian data sebesar 70% untuk training, 15% untuk validasi, dan 15% untuk testing. Pada gambar sebelah kiri sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai akurasi. Sedangkan gambar sebelah kanan sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai loss.

Pada gambar sebelah kiri, grafik yang menunjukkan perkembangan akurasi. Garis berwarna oranye menunjukkan nilai akurasi pada data validasi, sedangkan garis berwarna biru menunjukkan akurasi pada data pelatihan. Akurasi pelatihan mengalami peningkatan yang stabil dari sekitar 0.84 hingga mencapai 0.92. Di sisi lain, akurasi validasi mengalami fluktuasi yang cukup signifikan, misalnya meningkat tajam pada epoch ke-2, menurun ablei pada epoch ke-4, kemudian ablei meningkat hingga mendekati nilai akurasi pelatihan. Pola ini mengindikasikan adanya potensi overfitting pada epoch tertentu, namun model berhasil menyesuaikan diri ablei pada epoch selanjutnya.

Sementara itu, pada gambar sebelah kanan, grafik yang menunjukkan perkembangan loss. Garis berwarna oranye mewakili nilai loss pada data validasi, sedangkan garis berwarna biru menunjukkan loss pada data pelatihan. Loss pada data pelatihan menunjukkan tren menurun yang konsisten, dari sekitar 0.42 ke 0.21, yang mengindikasikan bahwa model semakin baik dalam mempelajari data pelatihan. Namun, loss pada data validasi menunjukkan pola yang lebih fluktuatif. Terjadi lonjakan pada epoch ke-4, yang ablei memperkuat indikasi overfitting. Meskipun demikian, nilai loss validasi ablei menurun setelahnya, menunjukkan bahwa model mampu memperbaiki kemampuannya dalam mengenali data baru.

#### Pembagian data 60%:30%:10

Pada table ditampilkan hasil pelatihan pada pembagian 60% untuk data *training*, 30% untuk data *validation*, 10% untuk data testing dengan pelatihan selama 40 *epoch*.

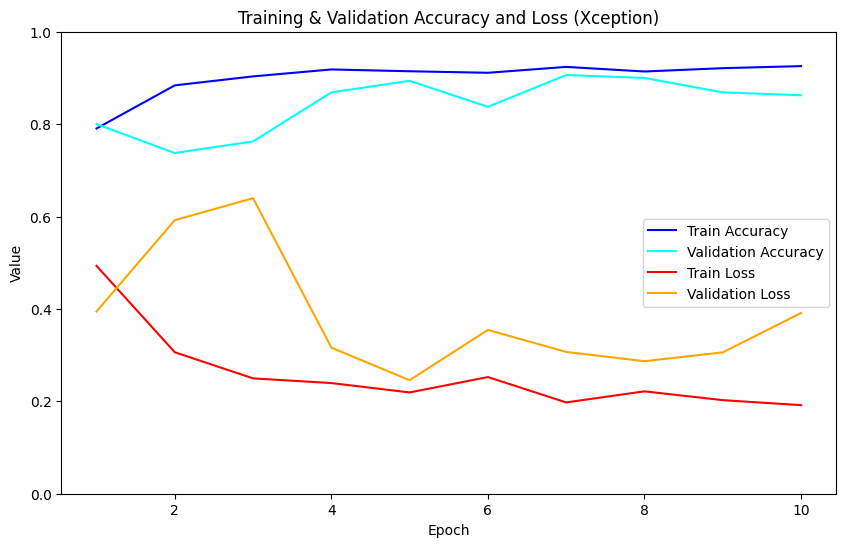
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 60%:30%:10% | | | | |
| Epoch | Akurasi | Loss | Akurasi Validasi | Loss Validasi |
| 1 | 0.6923 | 0.6888 | 0.8000 | 0.3944 |
| 2 | 0.8812 | 0.3138 | 0.7375 | 0.5921 |
| 3 | 0.9102 | 0.2394 | 0.7625 | 0.6397 |
| 4 | 0.9197 | 0.2529 | 0.8687 | 0.3162 |
| 5 | 0.9107 | 0.2290 | 0.8938 | 0.2457 |
| 6 | 0.9139 | 0.2505 | 0.8375 | 0.3545 |
| 7 | 0.9258 | 0.1916 | 0.9062 | 0.3069 |
| 8 | 0.9215 | 0.2161 | 0.9000 | 0.2868 |
| 9 | 0.9152 | 0.2138 | 0.8687 | 0.3060 |
| 10 | 0.9247 | 0.1962 | 0.8625 | 0.3912 |

Pelatihan pertama menggunakan skema pembagian data 60%:30:10% seperti yang disajikan pada table . Kita dapat melihat bahwa *epoch* pertama, model mencapai Akurasi sebesar 0.7478 dan Loss 0.5807 yang artinya model sudah mampu mengenali pola dasar dalam data dengan cukup baik, meskipun masih terdapat kesalahan prediksi yang cukup tinggi, yang sebagaimana ditunjukkan nilai loss 0.5807. Akan tetapi, untuk Akurasi Validasi sebesar 0.8250 dan Loss Validasi sebesar 0.4484 menunjukkan bahwa model justru berkinerja lebih baik pada data validasi dibandingkan data latih pada epoch pertama. Hal ini menunjukkan bahwa model belum sepenuhnya belajar pola-pola dari data latih, namun sudah mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat.

Pada *epoch* ke kedua, model mengalami peningkatan dengan akurasi 0.8900 dan Loss 0.2954. Hal ini menandakan bahwa model telah mempelajari pola dengan yang sangat baik. Model mulai mampu melakukan klasifikasi dengan tingkat kesalahan yang lebih rendah dibandingkan dengan *epoch* pertama. Namun, nilai akurasi validasi justru menurun menjadi 0.7688 dan loss validasi 0.6282. Penurunan ini mengindikasikan adanya fluktuasi dalam proses pelatihan, yang mungkin disebabkan oleh beberapa faktor seperti awal mula terjadinya *overfitting*, di mana model mulai terlalu menyesuaikan diri terhadap data pelatihan.

Selanjutnya, model mengalami peningkatan yang stabil dan konsisten pada *epoch* ke tiga dengan nilai akurasi 0.9088 dan nilai loss 0.2386. Akurasi validasi menunjukkan peningkatan dengan nilai sebesar 0.8938 dan loss validasi sebesar 0.2886. Hal ini menunjukkan bahwa model tidak hanya belajar dengan baik pada data pelatihan, tetapi juga mampu melakukan generalisasi yang sangat baik terhadap data validasi.

Setelah itu, performa model tetap stabil dan tidak memberikan kenaikan secara signifikan hingga pada *epoch* 11, meskipun sebelumnya telah direncanakan berlangsung hingga *epoch* 40. Hal ini disebabkan oleh *early stopping* yang menandakan tidak adanya peningkatan yang jelas atau signifikan, sehingga pelatihan dihentikan lebih awal untuk mencegah *overfitting* dan menghemat sumber daya komputasi.



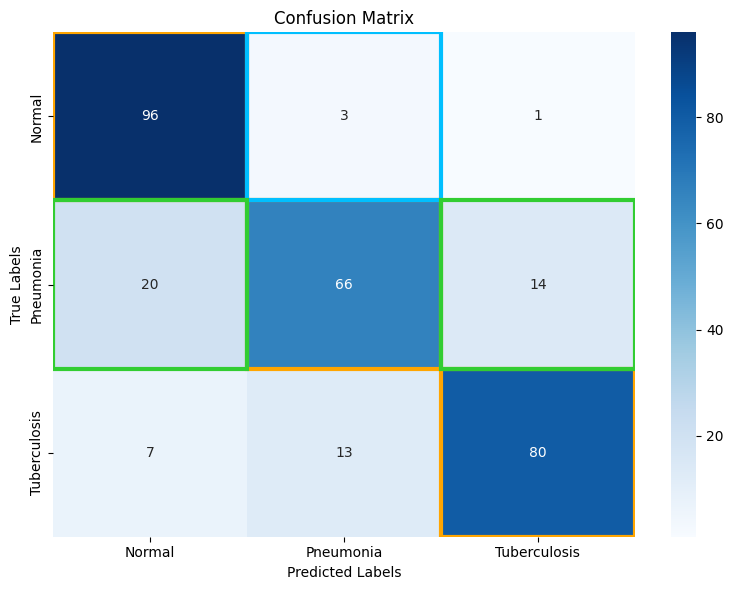
Grafik pada gambar menunjukkan performa model selama proses pelatihan, yang mencakup akurasi dan loss pada data pelatihan serta validasi. Pelatihan dilakukan dengan pembagian data sebesar 60% untuk training, 30% untuk validasi, dan 10% untuk testing. Pada gambar sebelah kiri sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai akurasi. Sedangkan gambar sebelah kanan sumbu x menunjukkan jumlah *epoch* pelatihan yang telah dilakukan, sumbu y menunjukkan nilai loss.

Pada gambar sebelah kiri, grafik menunjukkan perkembangan akurasi. Garis berwarna oranye menunjukkan nilai akurasi pada data validasi, sedangkan garis berwarna biru menunjukkan akurasi pada data pelatihan. Terlihat bahwa akurasi pelatihan meningkat secara konsisten dari sekitar 0.82 hingga mencapai lebih dari 0.93 pada akhir epoch. Sebaliknya, akurasi validasi tampak fluktuatif — menurun tajam pada beberapa epoch awal, lalu meningkat pesat pada epoch ke-6 dan mencapai nilai yang sebanding dengan akurasi pelatihan di epoch terakhir. Pola ini menunjukkan bahwa model sempat mengalami overfitting, tetapi kemudian mampu meningkatkan generalisasi terhadap data validasi.

Sementara itu, pada gambar sebelah kanan, grafik menunjukkan perkembangan loss. Garis berwarna oranye mewakili nilai loss pada data validasi, sedangkan garis berwarna biru menunjukkan loss pada data pelatihan. Loss pelatihan mengalami penurunan yang stabil dari sekitar 0.45 menjadi di bawah 0.2, mengindikasikan bahwa model semakin baik dalam mempelajari pola dari data pelatihan. Namun, loss validasi menunjukkan variasi yang besar, dengan lonjakan tinggi pada epoch ke-1 dan ke-4, serta fluktuasi pada epoch-epoch selanjutnya. Meskipun demikian, pada epoch terakhir, nilai loss validasi kembali menurun, menunjukkan adanya perbaikan dalam performa generalisasi model terhadap data yang tidak dilatih.

## **4.6 Evaluasi Model**

Evaluasi model DenseNet dan Xception yang telah dilatih pada penelitian ini menggunakan *confusion matrix*. Evaluasi dilakukan untuk mengetahui kinerja model yang telah dilatih dapat mengenali dan mendeteksi citra. *Confusion matrix* merupakan sebuah tabel yang menampilkan jumlah prediksi yang benar dan salah yang dibuat oleh model pada setiap kelas citra, yaitu normal, pneumonia tuberculosis.

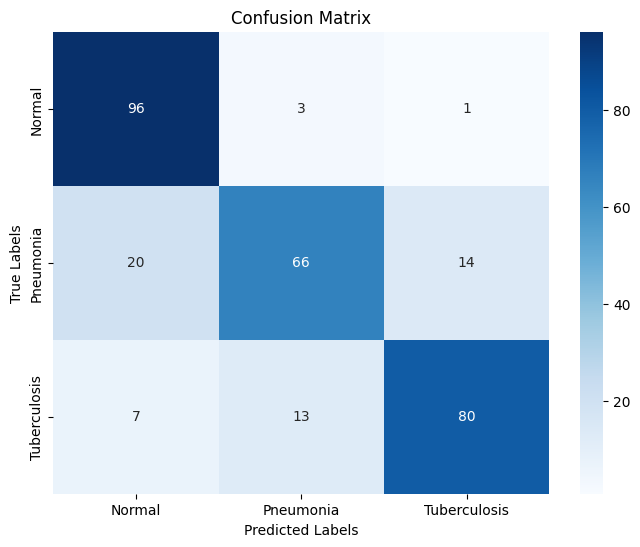


Pada gambar menampilkan hasil *confusion matrix* yang digunakan dalam mengevolusi performa model yang telah dilatih untuk mendeteksi tiga kelas normal, pneumonia, dan tuberculosis. Terdapat dua sumbu pada *confusion matrix*, yaitu sumbu x dan sumbu y. Sumbu y menunjukkan *True Label* (label sebenernya) dari data *testing*, sedangkan sumbu x menunjukan *Predicted Label* (label prediksi) yang dihasilkan oleh model. Terdapat pemetaan nilai *confusion matrix* berdasarkan garis berwarna kuning, biru, dan hijau dengan makna yang berbeda-beda. Nilai TP (True Positive) didefinisikan dengan warna kuning, Nilai FP (*False Positive*) didefinisikan dengan warna biru, FN (*False Negative*) didefinisikan dengan warna hijau.

### 4.6.1 Evaluasi Model DenseNet201

#### Pembagian data 80%:10%:10%

Pada gambar menampilkan hasil dari prediksi yang divisualisasikan menggunakan *confusion matrix* pada skema 80%:10%:10% selama 40 epoch. Visualisasi ini menggunakan dua parameter, yaitu *Actual* dan *Predicted*. *Actual* menunjukkan kelas sebenernya dari data citra (*Ground Truth*), sedangkan *Predicted* adalah kelas yang diprediksi oleh model. Selain digunakan untuk menghitung akurasi (*accuracy*), *confusion matrix* juga digunakan untuk menghitung presisi (*pression*), *recall*, dan F1 *Score* dengan melihat nilai TP (*True Positive*), FP (*False Positive*), dan FN (*False Negative*).



Pada Gambar diatas, dapat dilihat dari bahwa hasil pengujian menggunakan gambar dari data testing pada setiap kelas dengan total keseluruhan sebanyak 300 gambar, menunjukkan hasil yang memuaskan. Dari 300 gambar yang diuji, terdapat 18 kesalahan prediksi. Kesalahan prediksi ini dimungkinkan terjadi karena faktor-faktor kemiripan antar kelas, variasi data, dan kualitas gambar. Selanjutnya akan dilakukan perhitungan dari hasil *confusion matrix* untuk mengetahui performa kinerja model dengan nilai akurasi (*Accuracy*), presisi (*Precission*), *recall*, dan F1 *Score*.

Menghitung Akurasi (*Acuracy*):

Akurasi =

=

=

= 80.67%

Menghitung Presisi (*Precission*):

Presisi =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Positive | Presission |
| Normal | 96 | 27 |  |
| Pneumonia | 66 | 16 |  |
| Tuberculosis | 80 | 15 |  |
| Total | | | 2.4276 |

Macro Precission =

Macro Precission = = 80.92%

Menghitung Recall:

Recall =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Negative | Recall |
| Normal | 96 | 4 |  |
| Pneumonia | 66 | 34 |  |
| Tuberculosis | 80 | 20 |  |
| Total | | | 2.42 |

Macro Recall =

Macro Recall = = 80.67%

Menghitung F1 Score:

F1 - Score =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | *Precission* | *Recall* | F1-Score |
| Normal | 0.7805 | 0.96 |  |
| Pneumonia | 0.805 | 0.66 |  |
| Tuberculosis | 0.8421 | 0.80 |  |
| Total | | | 2.4076 |

Macro F1-Score =

Macro F1-Score = = 80.25%

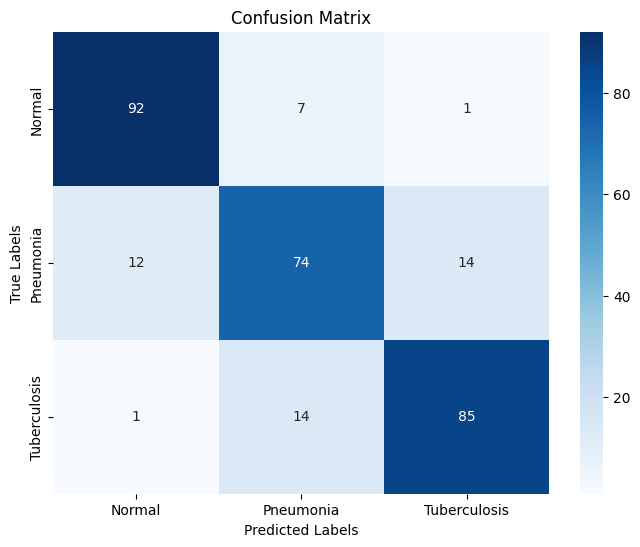
Setelah melakukan perhitungan dengan menggunakan total gambar dan menghasilkan nilai-nilai indikator seperti Akurasi, Presisi, *Recall*, dan F1 Score, akan ditunjukkan dalam ringkasan hasil evaluasi pada tabel. Tabel ini akan menunjukkan nilai masing- masing kelas untuk setiap indikator berdasarkan model dengan skema pembagian data 80%:10%:10% dan 40 epoch. Tabel ini memberikan gambaran kinerja model dalam memprediksi setiap kelas.

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | Precission | Recall | F1-Score |
| Normal | 0.7805 | 0.96 | 0.8611 |
| Pneumonia | 0.805 | 0.66 | 0.7259 |
| Tuberculosis | 0.8421 | 0.80 | 0.8206 |
| Macro Precission | 0.8092 | | |
| Macro Recall | 0.8067 | | |
| Accuracy | 0.8067 | | |
| Macro F1-Score | 0.8025 | | |

Hasil ringkasan menunjukkan bahwa akurasi yang diperoleh cukup tinggi yaitu sebesar 0.92 atau 92%. Selain itu, nilai yang sama pada presisi dan *recall* yaitu mencapai 0.92 atau 92% dan nilai F1 *Score* menacapai 0.9204 atau 92.04% Secara keseluruhan, model mampu mendeteksi dengan baik pada 3 jenis citra x-ray paru-paru yang telah dilatih.

#### Pembagian data 70%:20%:10%

Pada gambar menampilkan hasil dari prediksi yang divisualisasikan menggunakan *confusion matrix* pada skema 70%:20%:10% selama 40 epoch. Visualisasi ini menggunakan dua parameter, yaitu *Actual* dan *Predicted*. *Actual* menunjukkan kelas sebenernya dari data citra (*Ground Truth*), sedangkan *Predicted* adalah kelas yang diprediksi oleh model. Selain digunakan untuk menghitung akurasi (*accuracy*), *confusion matrix* juga digunakan untuk menghitung presisi (*pression*), *recall*, dan F1 *Score* dengan melihat nilai TP (*True Positive*), FP (*False Positive*), dan FN (*False Negative*).



Pada Gambar diatas, dapat dilihat dari bahwa hasil pengujian menggunakan gambar dari data testing pada setiap kelas dengan total keseluruhan sebanyak 300 gambar, menunjukkan hasil yang memuaskan. Dari 300 gambar yang diuji, terdapat 15 kesalahan prediksi. Kesalahan prediksi ini dimungkinkan terjadi karena faktor-faktor kemiripan antar kelas, variasi data, dan kualitas gambar. Selanjutnya akan dilakukan perhitungan dari hasil *confusion matrix* untuk mengetahui performa kinerja model dengan nilai akurasi (*Accuracy*), presisi (*Precission*), *recall*, dan F1 *Score*.

Menghitung Akurasi (*Acuracy*):

Akurasi =

=

=

= 83.67%

Menghitung Presisi (*Precission*):

Presisi =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Positive | Presission |
| Normal | 92 | 13 |  |
| Pneumonia | 74 | 21 |  |
| Tuberculosis | 85 | 15 |  |
| Total | | | 2.5049 |

Macro Precission =

Macro Precission = = 83.49%

Menghitung Recall:

Recall =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Negative | Recall |
| Normal | 92 | 8 |  |
| Pneumonia | 74 | 16 |  |
| Tuberculosis | 85 | 15 |  |
| Total | | | 2.51 |

Macro Recall =

Macro Recall = = 83.67%

Menghitung F1 Score:

F1 - Score =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | *Precission* | *Recall* | F1-Score |
| Normal | 0.8762 | 0.92 |  |
| Pneumonia | 0.7787 | 0.74 |  |
| Tuberculosis | 0.8500 | 0.85 |  |
| Total | | | 2.5064 |

Macro F1-Score =

Macro F1-Score = = 83.55%

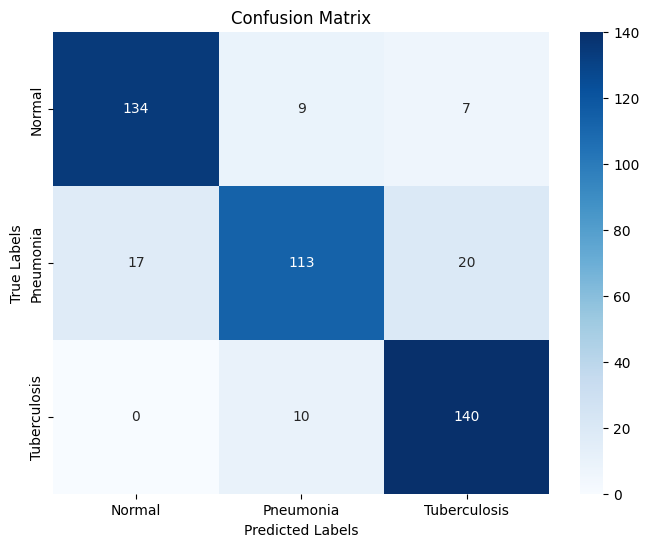
Setelah melakukan perhitungan dengan menggunakan total gambar dan menghasilkan nilai-nilai indikator seperti Akurasi, Presisi, *Recall*, dan F1 Score, akan ditunjukkan dalam ringkasan hasil evaluasi pada tabel. Tabel ini akan menunjukkan nilai masing- masing kelas untuk setiap indikator berdasarkan model dengan skema pembagian data 80%:10%:10% dan 40 epoch. Tabel ini memberikan gambaran kinerja model dalam memprediksi setiap kelas.

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | Precission | Recall | F1-Score |
| Normal | 0.8762 | 0.9200 | 0.8976 |
| Pneumonia | 0.7787 | 0.7400 | 0.7588 |
| Tuberculosis | 0.8500 | 0.8500 | 0.8500 |
| Macro Precission | 0.8349 | | |
| Macro Recall | 0.8367 | | |
| Accuracy | 0.8367 | | |
| Macro F1-Score | 0.8355 | | |

Hasil ringkasan menunjukkan bahwa akurasi yang diperoleh cukup tinggi yaitu sebesar 0.92 atau 92%. Selain itu, nilai yang sama pada presisi dan *recall* yaitu mencapai 0.92 atau 92% dan nilai F1 *Score* menacapai 0.9204 atau 92.04% Secara keseluruhan, model mampu mendeteksi dengan baik pada 3 jenis citra x-ray paru-paru yang telah dilatih.

#### Pembagian data 70%:15%:15%

Pada gambar menampilkan hasil dari prediksi yang divisualisasikan menggunakan *confusion matrix* pada skema 70%:15%:15% selama 40 epoch. Visualisasi ini menggunakan dua parameter, yaitu *Actual* dan *Predicted*. *Actual* menunjukkan kelas sebenernya dari data citra (*Ground Truth*), sedangkan *Predicted* adalah kelas yang diprediksi oleh model. Selain digunakan untuk menghitung akurasi (*accuracy*), *confusion matrix* juga digunakan untuk menghitung presisi (*pression*), *recall*, dan F1 *Score* dengan melihat nilai TP (*True Positive*), FP (*False Positive*), dan FN (*False Negative*).



Pada Gambar diatas, dapat dilihat dari bahwa hasil pengujian menggunakan gambar dari data testing pada setiap kelas dengan total keseluruhan sebanyak 450 gambar, menunjukkan hasil yang memuaskan. Dari 450 gambar yang diuji, terdapat 36 kesalahan prediksi. Kesalahan prediksi ini dimungkinkan terjadi karena faktor-faktor kemiripan antar kelas, variasi data, dan kualitas gambar. Selanjutnya akan dilakukan perhitungan dari hasil *confusion matrix* untuk mengetahui performa kinerja model dengan nilai akurasi (*Accuracy*), presisi (*Precission*), *recall*, dan F1 *Score*.

Menghitung Akurasi (*Acuracy*):

Akurasi =

=

=

= 86%

Menghitung Presisi (*Precission*):

Presisi =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Positive | Presission |
| Normal | 134 | 17 |  |
| Pneumonia | 113 | 19 |  |
| Tuberculosis | 140 | 27 |  |
| Total | | | 2.5818 |

Macro Precission =

Macro Precission = = 86.06%

Menghitung Recall:

Recall =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Negative | Recall |
| Normal | 134 | 16 |  |
| Pneumonia | 113 | 37 |  |
| Tuberculosis | 140 | 10 |  |
| Total | | | 2.5799 |

Macro Recall =

Macro Recall = = 86%

Menghitung F1 Score:

F1 - Score =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | *Precission* | *Recall* | F1-Score |
| Normal | 0.8874 | 0.8933 |  |
| Pneumonia | 0.8561 | 0.7533 |  |
| Tuberculosis | 0.8383 | 0.9333 |  |
| Total | | | 2.5747 |

Macro F1-Score =

F1-Score = = 85.82%

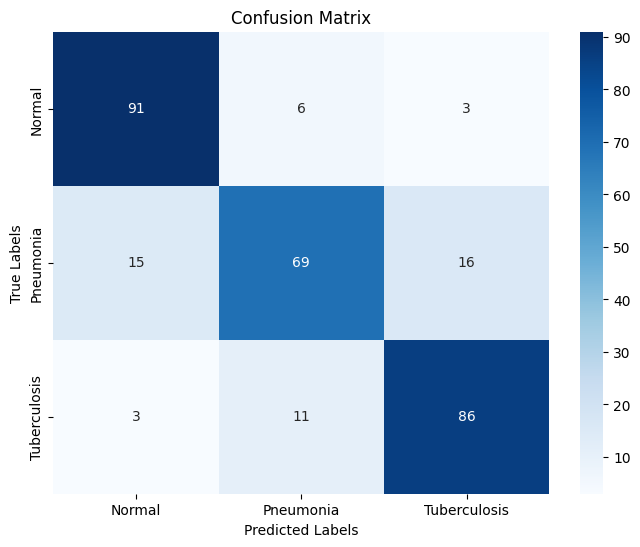
Setelah melakukan perhitungan dengan menggunakan total gambar dan menghasilkan nilai-nilai indikator seperti Akurasi, Presisi, *Recall*, dan F1 Score, akan ditunjukkan dalam ringkasan hasil evaluasi pada tabel. Tabel ini akan menunjukkan nilai masing- masing kelas untuk setiap indikator berdasarkan model dengan skema pembagian data 80%:10%:10% dan 40 epoch. Tabel ini memberikan gambaran kinerja model dalam memprediksi setiap kelas.

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | Precission | Recall | F1-Score |
| Normal | 0.8874 | 0.8933 | 0.8903 |
| Pneumonia | 0.8561 | 0.7533 | 0.8013 |
| Tuberculosis | 0.8383 | 0.9333 | 0.8831 |
| Macro Precission | 0.8606 | | |
| Macro Recall | 0.86 | | |
| Accuracy | 0.86 | | |
| Macro F1-Score | 0.8582 | | |

Hasil ringkasan menunjukkan bahwa akurasi yang diperoleh cukup tinggi yaitu sebesar 0.92 atau 92%. Selain itu, nilai yang sama pada presisi dan *recall* yaitu mencapai 0.92 atau 92% dan nilai F1 *Score* menacapai 0.9204 atau 92.04% Secara keseluruhan, model mampu mendeteksi dengan baik pada 3 jenis citra x-ray paru-paru yang telah dilatih.

#### Pembagian data 60%:30%:10%

Pada gambar menampilkan hasil dari prediksi yang divisualisasikan menggunakan *confusion matrix* pada skema 60%:30%:10% selama 40 epoch. Visualisasi ini menggunakan dua parameter, yaitu *Actual* dan *Predicted*. *Actual* menunjukkan kelas sebenernya dari data citra (*Ground Truth*), sedangkan *Predicted* adalah kelas yang diprediksi oleh model. Selain digunakan untuk menghitung akurasi (*accuracy*), *confusion matrix* juga digunakan untuk menghitung presisi (*pression*), *recall*, dan F1 *Score* dengan melihat nilai TP (*True Positive*), FP (*False Positive*), dan FN (*False Negative*).



Pada Gambar diatas, dapat dilihat dari bahwa hasil pengujian menggunakan gambar dari data testing pada setiap kelas dengan total keseluruhan sebanyak 300 gambar, menunjukkan hasil yang memuaskan. Dari 300 gambar yang diuji, terdapat 19 kesalahan prediksi. Kesalahan prediksi ini dimungkinkan terjadi karena faktor-faktor kemiripan antar kelas, variasi data, dan kualitas gambar. Selanjutnya akan dilakukan perhitungan dari hasil *confusion matrix* untuk mengetahui performa kinerja model dengan nilai akurasi (*Accuracy*), presisi (*Precission*), *recall*, dan F1 *Score*.

Menghitung Akurasi (*Acuracy*):

Akurasi =

=

=

= 82%

Menghitung Presisi (*Precission*):

Presisi =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Positive | Presission |
| Normal | 91 | 18 |  |
| Pneumonia | 69 | 17 |  |
| Tuberculosis | 86 | 19 |  |
| Total | | | 2.4568 |

Macro Precission =

Macro Precission = = 81.89%

Menghitung Recall:

Recall =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Negative | Recall |
| Normal | 91 | 9 |  |
| Pneumonia | 69 | 31 |  |
| Tuberculosis | 86 | 14 |  |
| Total | | | 2.46 |

Macro Recall =

Macro Recall = = 82%

Menghitung F1 Score:

F1 - Score =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | *Precission* | *Recall* | F1-Score |
| Normal | 0.8355 | 0.91 |  |
| Pneumonia | 0.8023 | 0.69 |  |
| Tuberculosis | 0.8190 | 0.86 |  |
| Total | | | 2.4516 |

Macro F1-Score =

Macro F1-Score = = 81.72%

Setelah melakukan perhitungan dengan menggunakan total gambar dan menghasilkan nilai-nilai indikator seperti Akurasi, Presisi, *Recall*, dan F1 Score, akan ditunjukkan dalam ringkasan hasil evaluasi pada tabel. Tabel ini akan menunjukkan nilai masing- masing kelas untuk setiap indikator berdasarkan model dengan skema pembagian data 80%:10%:10% dan 40 epoch. Tabel ini memberikan gambaran kinerja model dalam memprediksi setiap kelas.

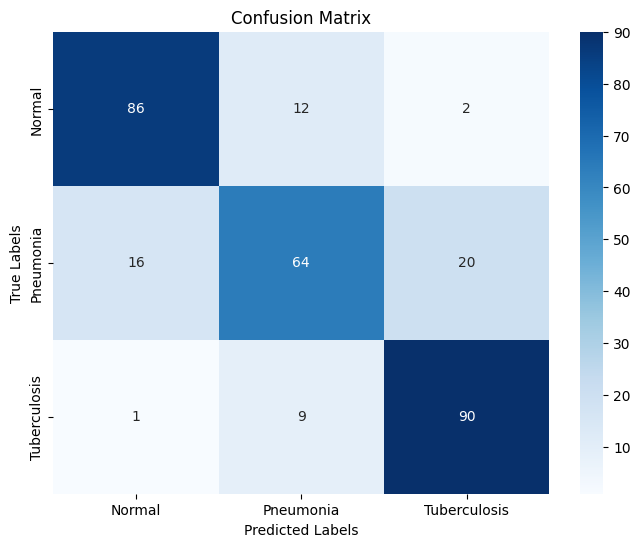
|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | Precission | Recall | F1-Score |
| Normal | 0.8355 | 0.91 | 0.8711 |
| Pneumonia | 0.8023 | 0.69 | 0.7416 |
| Tuberculosis | 0.8190 | 0.86 | 0.8389 |
| Macro Precission | 0.8189 | | |
| Macro Recall | 0.82 | | |
| Accuracy | 0.82 | | |
| Macro F1-Score | 0.8172 | | |

Hasil ringkasan menunjukkan bahwa akurasi yang diperoleh cukup tinggi yaitu sebesar 0.92 atau 92%. Selain itu, nilai yang sama pada presisi dan *recall* yaitu mencapai 0.92 atau 92% dan nilai F1 *Score* menacapai 0.9204 atau 92.04% Secara keseluruhan, model mampu mendeteksi dengan baik pada 3 jenis citra x-ray paru-paru yang telah dilatih.

### 4.6.2 Evaluasi Model Xception

#### Pembagian data 80%:10%:10%

Pada gambar menampilkan hasil dari prediksi yang divisualisasikan menggunakan *confusion matrix* pada skema 80%:10%:10% selama 40 epoch. Visualisasi ini menggunakan dua parameter, yaitu *Actual* dan *Predicted*. *Actual* menunjukkan kelas sebenernya dari data citra (*Ground Truth*), sedangkan *Predicted* adalah kelas yang diprediksi oleh model. Selain digunakan untuk menghitung akurasi (*accuracy*), *confusion matrix* juga digunakan untuk menghitung presisi (*pression*), *recall*, dan F1 *Score* dengan melihat nilai TP (*True Positive*), FP (*False Positive*), dan FN (*False Negative*).



Pada Gambar diatas, dapat dilihat dari bahwa hasil pengujian menggunakan gambar dari data testing pada setiap kelas dengan total keseluruhan sebanyak 300 gambar, menunjukkan hasil yang memuaskan. Dari 300 gambar yang diuji, terdapat 38 kesalahan prediksi. Kesalahan prediksi ini dimungkinkan terjadi karena faktor-faktor kemiripan antar kelas, variasi data, dan kualitas gambar. Selanjutnya akan dilakukan perhitungan dari hasil *confusion matrix* untuk mengetahui performa kinerja model dengan nilai akurasi (*Accuracy*), presisi (*Precission*), *recall*, dan F1 *Score*.

Menghitung Akurasi (*Acuracy*):

Akurasi =

=

=

= 80%

Menghitung Presisi (*Precission*):

Presisi =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Positive | Presission |
| Normal | 86 | 17 |  |
| Pneumonia | 64 | 21 |  |
| Tuberculosis | 90 | 22 |  |
| Total | | | 2.3915 |

Macro Precission =

Macro Precission = = 79.72%

Menghitung Recall:

Recall =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Negative | Recall |
| Normal | 86 | 14 |  |
| Pneumonia | 64 | 36 |  |
| Tuberculosis | 90 | 10 |  |
| Total | | | 2.4 |

Macro Recall =

Macro Recall = = 80%

Menghitung F1 Score:

F1 - Score =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | *Precission* | *Recall* | F1-Score |
| Normal | 0.8350 | 0.86 |  |
| Pneumonia | 0.7529 | 0.64 |  |
| Tuberculosis | 0.8036 | 0.90 |  |
| Total | | | 2.3893 |

Macro F1-Score =

Macro F1-Score = = 79.64%

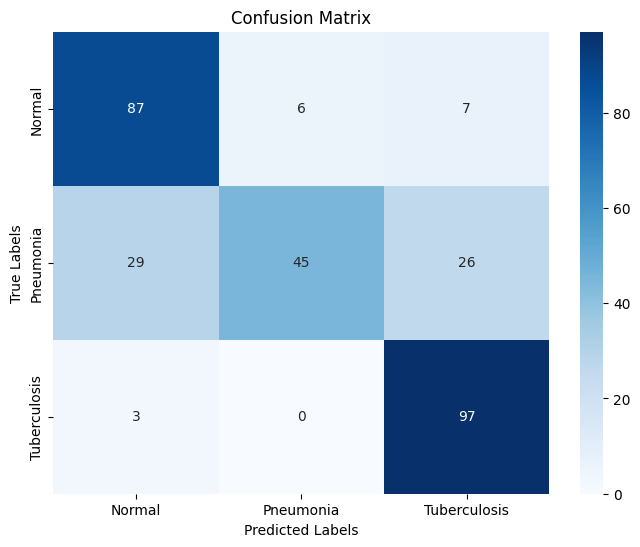
Setelah melakukan perhitungan dengan menggunakan total gambar dan menghasilkan nilai-nilai indikator seperti Akurasi, Presisi, *Recall*, dan F1 Score, akan ditunjukkan dalam ringkasan hasil evaluasi pada tabel. Tabel ini akan menunjukkan nilai masing- masing kelas untuk setiap indikator berdasarkan model dengan skema pembagian data 80%:10%:10% dan 40 epoch. Tabel ini memberikan gambaran kinerja model dalam memprediksi setiap kelas.

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | Precission | Recall | F1-Score |
| Normal | 0.8350 | 0.86 | 0.8474 |
| Pneumonia | 0.7529 | 0.64 | 0.6925 |
| Tuberculosis | 0.8036 | 0.90 | 0.8494 |
| Macro Precission | 0.7972 | | |
| Macro Recall | 0.80 | | |
| Accuracy | 0.80 | | |
| Macro F1-Score | 0.7964 | | |

Hasil ringkasan menunjukkan bahwa akurasi yang diperoleh cukup tinggi yaitu sebesar 0.92 atau 92%. Selain itu, nilai yang sama pada presisi dan *recall* yaitu mencapai 0.92 atau 92% dan nilai F1 *Score* menacapai 0.9204 atau 92.04% Secara keseluruhan, model mampu mendeteksi dengan baik pada 3 jenis citra x-ray paru-paru yang telah dilatih.

#### Pembagian data 70%:20%:10%

Pada gambar menampilkan hasil dari prediksi yang divisualisasikan menggunakan *confusion matrix* pada skema 70%:20%:10% selama 40 epoch. Visualisasi ini menggunakan dua parameter, yaitu *Actual* dan *Predicted*. *Actual* menunjukkan kelas sebenernya dari data citra (*Ground Truth*), sedangkan *Predicted* adalah kelas yang diprediksi oleh model. Selain digunakan untuk menghitung akurasi (*accuracy*), *confusion matrix* juga digunakan untuk menghitung presisi (*pression*), *recall*, dan F1 *Score* dengan melihat nilai TP (*True Positive*), FP (*False Positive*), dan FN (*False Negative*).



Pada Gambar diatas, dapat dilihat dari bahwa hasil pengujian menggunakan gambar dari data testing pada setiap kelas dengan total keseluruhan sebanyak 300 gambar, menunjukkan hasil yang memuaskan. Dari 300 gambar yang diuji, terdapat 37 kesalahan prediksi. Kesalahan prediksi ini dimungkinkan terjadi karena faktor-faktor kemiripan antar kelas, variasi data, dan kualitas gambar. Selanjutnya akan dilakukan perhitungan dari hasil *confusion matrix* untuk mengetahui performa kinerja model dengan nilai akurasi (*Accuracy*), presisi (*Precission*), *recall*, dan F1 *Score*.

Menghitung Akurasi (*Acuracy*):

Akurasi =

=

=

= 76.33%

Menghitung Presisi (*Precission*):

Presisi =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Positive | Presission |
| Normal | 87 | 32 |  |
| Pneumonia | 45 | 6 |  |
| Tuberculosis | 97 | 33 |  |
| Total | | | 2.3596 |

Macro Precission =

Macro Precission = = 78.65%

Menghitung Recall:

Recall =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Negative | Recall |
| Normal | 87 | 13 |  |
| Pneumonia | 45 | 55 |  |
| Tuberculosis | 97 | 97 |  |
| Total | | | 2.29 |

Macro Recall =

Macro Recall = = 76.33%

Menghitung F1 Score:

F1 - Score =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | *Precission* | *Recall* | F1-Score |
| Normal | 0.7310 | 0.87 |  |
| Pneumonia | 0.8824 | 0.45 |  |
| Tuberculosis | 0.7462 | 0.97 |  |
| Total | | | 2.2335 |

Macro F1-Score =

Macro F1-Score = = 74.45%

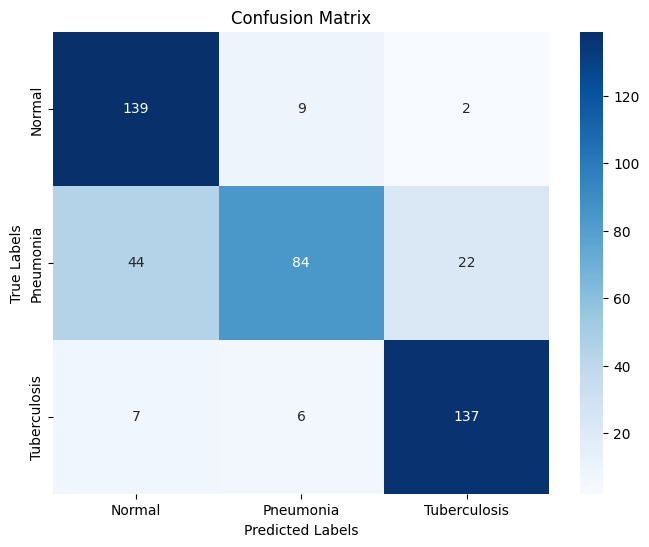
Setelah melakukan perhitungan dengan menggunakan total gambar dan menghasilkan nilai-nilai indikator seperti Akurasi, Presisi, *Recall*, dan F1 Score, akan ditunjukkan dalam ringkasan hasil evaluasi pada tabel. Tabel ini akan menunjukkan nilai masing- masing kelas untuk setiap indikator berdasarkan model dengan skema pembagian data 80%:10%:10% dan 40 epoch. Tabel ini memberikan gambaran kinerja model dalam memprediksi setiap kelas.

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | Precission | Recall | F1-Score |
| Normal | 0.7310 | 0.87 | 0.7946 |
| Pneumonia | 0.8824 | 0.45 | 0.5957 |
| Tuberculosis | 0.7462 | 0.97 | 0.8432 |
| Macro Precission | 0.7865 | | |
| Macro Recall | 0.7633 | | |
| Accuracy | 0.7633 | | |
| Macro F1-Score | 0.7445 | | |

Hasil ringkasan menunjukkan bahwa akurasi yang diperoleh cukup tinggi yaitu sebesar 0.92 atau 92%. Selain itu, nilai yang sama pada presisi dan *recall* yaitu mencapai 0.92 atau 92% dan nilai F1 *Score* menacapai 0.9204 atau 92.04% Secara keseluruhan, model mampu mendeteksi dengan baik pada 3 jenis citra x-ray paru-paru yang telah dilatih.

#### Pembagian data 70%:15%:15%

Pada gambar menampilkan hasil dari prediksi yang divisualisasikan menggunakan *confusion matrix* pada skema 70%:15%:15% selama 40 epoch. Visualisasi ini menggunakan dua parameter, yaitu *Actual* dan *Predicted*. *Actual* menunjukkan kelas sebenernya dari data citra (*Ground Truth*), sedangkan *Predicted* adalah kelas yang diprediksi oleh model. Selain digunakan untuk menghitung akurasi (*accuracy*), *confusion matrix* juga digunakan untuk menghitung presisi (*pression*), *recall*, dan F1 *Score* dengan melihat nilai TP (*True Positive*), FP (*False Positive*), dan FN (*False Negative*).



Pada Gambar diatas, dapat dilihat dari bahwa hasil pengujian menggunakan gambar dari data testing pada setiap kelas dengan total keseluruhan sebanyak 450 gambar, menunjukkan hasil yang memuaskan. Dari 450 gambar yang diuji, terdapat 50 kesalahan prediksi. Kesalahan prediksi ini dimungkinkan terjadi karena faktor-faktor kemiripan antar kelas, variasi data, dan kualitas gambar. Selanjutnya akan dilakukan perhitungan dari hasil *confusion matrix* untuk mengetahui performa kinerja model dengan nilai akurasi (*Accuracy*), presisi (*Precission*), *recall*, dan F1 *Score*.

Menghitung Akurasi (*Acuracy*):

Akurasi =

=

=

= 80%

Menghitung Presisi (*Precission*):

Presisi =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Positive | Presission |
| Normal | 139 | 51 |  |
| Pneumonia | 84 | 15 |  |
| Tuberculosis | 137 | 24 |  |
| Total | | | 2.432 |

Macro Precission =

Macro Precission = = 81.07%

Menghitung Recall:

Recall =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Negative | Recall |
| Normal | 139 | 11 |  |
| Pneumonia | 84 | 66 |  |
| Tuberculosis | 137 | 13 |  |
| Total | | | 2.4 |

Macro Recall =

Macro Recall = = 80%

Menghitung F1 Score:

F1 - Score =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | *Precission* | *Recall* | F1-Score |
| Normal | 0.7319 | 0.9267 |  |
| Pneumonia | 0.8485 | 0.56 |  |
| Tuberculosis | 0.8516 | 0.9133 |  |
| Total | | | 2.3731 |

Macro F1-Score =

Macro F1-Score = = 79.10%

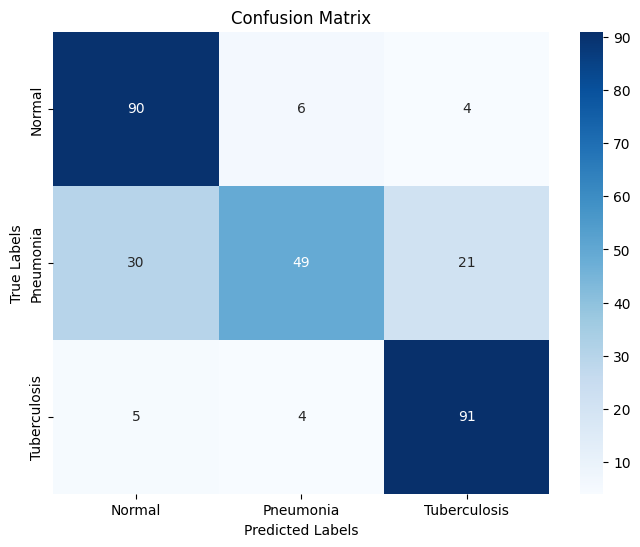
Setelah melakukan perhitungan dengan menggunakan total gambar dan menghasilkan nilai-nilai indikator seperti Akurasi, Presisi, *Recall*, dan F1 Score, akan ditunjukkan dalam ringkasan hasil evaluasi pada tabel. Tabel ini akan menunjukkan nilai masing- masing kelas untuk setiap indikator berdasarkan model dengan skema pembagian data 80%:10%:10% dan 40 epoch. Tabel ini memberikan gambaran kinerja model dalam memprediksi setiap kelas.

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | Precission | Recall | F1-Score |
| Normal | 0.7319 | 0.9267 | 0.8187 |
| Pneumonia | 0.8485 | 0.56 | 0.6731 |
| Tuberculosis | 0.8516 | 0.9133 | 0.8813 |
| Macro Precission | 0.8107 | | |
| Macro Recall | 0.80 | | |
| Accuracy | 0.80 | | |
| Macro F1-Score | 0.7910 | | |

Hasil ringkasan menunjukkan bahwa akurasi yang diperoleh cukup tinggi yaitu sebesar 0.92 atau 92%. Selain itu, nilai yang sama pada presisi dan *recall* yaitu mencapai 0.92 atau 92% dan nilai F1 *Score* menacapai 0.9204 atau 92.04% Secara keseluruhan, model mampu mendeteksi dengan baik pada 3 jenis citra x-ray paru-paru yang telah dilatih.

#### Pembagian data 60%:30%:10%

Pada gambar menampilkan hasil dari prediksi yang divisualisasikan menggunakan *confusion matrix* pada skema 60%:30%:10% selama 40 epoch. Visualisasi ini menggunakan dua parameter, yaitu *Actual* dan *Predicted*. *Actual* menunjukkan kelas sebenernya dari data citra (*Ground Truth*), sedangkan *Predicted* adalah kelas yang diprediksi oleh model. Selain digunakan untuk menghitung akurasi (*accuracy*), *confusion matrix* juga digunakan untuk menghitung presisi (*pression*), *recall*, dan F1 *Score* dengan melihat nilai TP (*True Positive*), FP (*False Positive*), dan FN (*False Negative*).



Pada Gambar diatas, dapat dilihat dari bahwa hasil pengujian menggunakan gambar dari data testing pada setiap kelas dengan total keseluruhan sebanyak 300 gambar, menunjukkan hasil yang memuaskan. Dari 300 gambar yang diuji, terdapat 32 kesalahan prediksi. Kesalahan prediksi ini dimungkinkan terjadi karena faktor-faktor kemiripan antar kelas, variasi data, dan kualitas gambar. Selanjutnya akan dilakukan perhitungan dari hasil *confusion matrix* untuk mengetahui performa kinerja model dengan nilai akurasi (*Accuracy*), presisi (*Precission*), *recall*, dan F1 *Score*.

Menghitung Akurasi (*Acuracy*):

Akurasi =

=

=

= 76.67%

Menghitung Presisi (*Precission*):

Presisi =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Positive | Presission |
| Normal | 90 | 35 |  |
| Pneumonia | 49 | 10 |  |
| Tuberculosis | 91 | 25 |  |
| Total | | | 2.335 |

Macro Precission =

Macro Precission = = 77.83%

Menghitung Recall:

Recall =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | True Positive | False Negative | Recall |
| Normal | 90 | 10 |  |
| Pneumonia | 49 | 51 |  |
| Tuberculosis | 91 | 9 |  |
| Total | | | 2.3 |

Macro Recall =

Macro Recall = = 76.67%

Menghitung F1 Score:

F1 - Score =

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | *Precission* | *Recall* | F1-Score |
| Normal | 0.72 | 0.90 |  |
| Pneumonia | 0.831 | 0.49 |  |
| Tuberculosis | 0.784 | 0.91 |  |
| Total | | | 2.258 |

Macro F1-Score =

Macro F1-Score = = 75.23%

Setelah melakukan perhitungan dengan menggunakan total gambar dan menghasilkan nilai-nilai indikator seperti Akurasi, Presisi, *Recall*, dan F1 Score, akan ditunjukkan dalam ringkasan hasil evaluasi pada tabel. Tabel ini akan menunjukkan nilai masing- masing kelas untuk setiap indikator berdasarkan model dengan skema pembagian data 80%:10%:10% dan 40 epoch. Tabel ini memberikan gambaran kinerja model dalam memprediksi setiap kelas.

|  |  |  |  |
| --- | --- | --- | --- |
| Kelas | Precission | Recall | F1-Score |
| Normal | 0.72 | 0.90 | 0.80 |
| Pneumonia | 0.8305 | 0.49 | 0.6158 |
| Tuberculosis | 0.7845 | 0.91 | 0.8412 |
| Macro Precission | 0.7783 | | |
| Macro Recall | 0.7667 | | |
| Accuracy | 0. 7667 | | |
| Macro F1-Score | 0. 7523 | | |

Hasil ringkasan menunjukkan bahwa akurasi yang diperoleh cukup tinggi yaitu sebesar 0.92 atau 92%. Selain itu, nilai yang sama pada presisi dan *recall* yaitu mencapai 0.92 atau 92% dan nilai F1 *Score* menacapai 0.9204 atau 92.04% Secara keseluruhan, model mampu mendeteksi dengan baik pada 3 jenis citra x-ray paru-paru yang telah dilatih.

## **4.7 Tabel Perbandingan Klasifikasi Model DenseNet201 dan Xception**

Tabel 4.44 berikut menyajikan perbandingan performa dari klasifikasi menggunakan data citra *x-ray* paru-paru antara model DenseNet201 dan Xception berdasarkan dari empat skenario pembagian data yang telah dilakukan. Evaluasi menggunakan metrik akurasi, presisi (macro), recall (macro), f1-score (macro).

Tabel 4. 9 Tabel Perbandingan Hasil Klasifikasi Model DenseNet201 dan Xception

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Skenario | Model | Akurasi | Presisi (Macro) | Recall (Macro) | F1-Score (Macro) |
| (80-10-10) | DenseNet201 | 0.8067 | 0.8092 | 0.8067 | 0.8025 |
| Xceptionn | 0.80 | 0.7972 | 0.80 | 0.7964 |
| (70-20-10) | DenseNet201 | 0.8367 | 0.8349 | 0.8367 | 0.8355 |
| Xception | 0.7633 | 0.7865 | 0.7633 | 0.7445 |
| (70-15-15) | DenseNet201 | 0.86 | 0.8606 | 0.86 | 0.8582 |
| Xception | 0.80 | 0.8107 | 0.80 | 0.7910 |
| (60-30-10) | DenseNet201 | 0.82 | 0.8189 | 0.82 | 0.8172 |
| Xception | 0.7667 | 0.7783 | 0.7667 | 0.7523 |

Berdasarkan hasil evaluasi pada tabel 4.44, dapat dilihat bahwa model DenseNet201 menunjukkan performa yang lebih baik dibandingkan dengan model Xception pada keempat skenario pembagian data yang digunakan. Evaluasi dilakukan menggunakan matrik akurasi, presisi makro, recall makro, dan f1-score makro, yang memberikan kualitas hasil dari klasifikasi model terhadap ketiga kelas yaitu, normal, pneumonia, tuberculosis.

Pada skenario 70-20-10, model DenseNet201 mencapai performa tertinggi dengan akurasi 95%, serta nilai presisi, recall, f1-score makro masing-masing sebesar 0.9499, 0.95, dan 0.9498. Hal ini menunjukkan bahwa model mampu mengklasifikasikan dengan sangat baik dan konsisten pada semua kelas. Sebaliknya, pada skenario 60-30-10, model xception mencapai akurasi terbaiknya, yaitu sebesar 89.33%, dengan nilai presisi, recall, f1-score makro masing masing sebesar 0.8989, 0.8933 dan 0.8929. Yang berarti arsitektur xception masih tergolong baik namun tidak sebanding dengan performa yang dicapai oleh DeneNet201.

Perbedaan performa antara kedua model tersebut menunjukkan bahwa arsitektur DenseNet201 lebih unggul dalam melakukan klasifikasi pada dataset citra x-ray paru-paru. Konektivitas pada antara layer-layer pada arsitektur DenseNet memungkinkan fitur dari layer awal tetap terhubung dan digunakan hingga layer akhir, sehingga menghasilkan representasi fitur yang lebih kompleks dan mendalam. Sementara itu, model Xception yang menggunakan arsitektur *depthwise seperable convolution* tetap menunjukkan performa yang baik, meskipun belum dapat mengungguli performa dari DenseNet201.

# BAB 5 KESIMPULAN DAN SARAN

## **5.1 Kesimpulan**

Berdasarkan hasil pelatihan dan pengujian yang dilakukan pada model DenseNet201 dan Xception dapat ditarik hasil kesimpulan sebagai berikut:

1. Penelitian ini membandingkan arsitektur model DenseNet201 dan Xception dalam mengklasifikasikan citra X-ray paru-paru normal, pneumonia, dan tuberculosis menggunakan 3000 gambar dari dataset sekunder. Peningkatan kontras dilakukan dengan metode CLAHE, sedangkan augmentasi data mencakup rescale, normalisasi piksel, rotasi, pergeseran, shear, zoom, dan flip horizontal secara real-time. Seluruh gambar diubah ke ukuran 512x512 piksel dengan batch size 32. Data dilatih dengan empat skema pembagian 80%:10%:10%, 70%:20%:10%, 70%:15%:15%, dan 60%:30%:10% selama 40 epoch dan 70 perulangan. Evaluasi performa dilakukan menggunakan *confusion matrix*, dan model dibandingkan berdasarkan akurasi, presisi, *recall*, dan *f1-score*.
2. Hasil evaluasi menggunakan *confusion matrix* menunjukkan bahwa model yang memiliki kinerja klasifikasi terbaik adalah model DenseNet201dengan pembagian data 70% untuk data *training*, 20 % untuk data *validation*, dan 10% untuk data *testing*. Nilai akurasi yang diperoleh sebesar 0.95 atau 95%, *macro precision* sebesar 0.9499 atau 94.99%, *macro* *recall* sebesar 0.95 atau 95%, dan *macro* *F1-Score* sebesar 0.9498 atau 94.98%.

## **5.2 Saran**

Salah satu faktor yang memengaruhi kinerja suatu model klasifikasi adalah jumlah data yang digunakan. Semakin banyak data yang tersedia, maka akurasi yang diperoleh cenderung lebih optimal. Namun, dalam penelitian ini, jumlah citra yang digunakan dibatasi hanya sebanyak 3.000 citra digital karena keterbatasan kapasitas penyimpanan. Oleh karena itu, peningkatan jumlah data diperlukan agar hasil akurasi dapat lebih ditingkatkan. Untuk mendukung hal tersebut, ke depannya dibutuhkan media penyimpanan yang lebih besar. Selain itu, untuk penelitian selanjutnya bisa ditambahkan lebih banyak variasi jenis citra x-ray paru-paru, dikarenakan pada penelitian ini masih menggunakan 3 jenis klasifikasi citra x-ray paru-paru, sementara jenis citra X-ray paru-paru yang tersedia sebenarnya masih sangat beragam.

# DAFTAR PUSTAKA

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., … Research, G. (2015). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. www.tensorflow.org.

Agustina Purwitasari, N., Soleh, M., Raya Puspiptek, J., & Tangerang Selatan, K. (2022). Implementasi Algoritma Artificial Neural Network Dalam Pembuatan Chatbot Menggunakan Pendekatan Natural Language Processing (Implementation Of Artificial Neural Network Algorithm In Chatbot Development Using Natural Language Processing Approach).

Banu, S., Sitepu, R., & Sulistiari, R. (2017). FAKTOR RISIKO KEJADIAN TB PARU DI PUSKESMAS HUTARAKYAT SIDIKALANG TAHUN 2017.

Bishop, C. M. . (2006). Pattern recognition and machine learning. Springer Science + Business Media.

Carnagie, J. O., Prabowo, A. R., Budiana, E. P., & Singgih, I. K. (2022). Essential Oil Plants Image Classification Using Xception Model. Procedia Computer Science, 204, 395–402. https://doi.org/10.1016/j.procs.2022.08.048

Carneiro, T., Da Nobrega, R. V. M., Nepomuceno, T., Bian, G. Bin, De Albuquerque, V. H. C., & Filho, P. P. R. (2018). Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications. IEEE Access, 6, 61677–61685. https://doi.org/10.1109/ACCESS.2018.2874767

CDC. (2023). Risk Factors for Pneumonia | Pneumonia | CDC. https://www.cdc.gov/pneumonia/risk-factors/index.html?form=MG0AV3

Chauhan, T., Palivela, H., & Tiwari, S. (2021). Optimization and fine-tuning of DenseNet model for classification of COVID-19 cases in medical imaging. International Journal of Information Management Data Insights, 1(2). https://doi.org/10.1016/j.jjimei.2021.100020

Chen, X. (2019). Image enhancement effect on the performance of convolutional neural networks. www.bth.se

Chest X-rays – The Heart Clinic. (2015, September 29). https://heartclinics.org/terminology/chest-x-rays/

Choi, Y., Hart, G., Soulliere, C., & Chen, M. (2024, May 31). TP / FP / FN / TN: Learn about true positives, false positives, false negatives, and true negatives to evaluate

ML model performance · Testing with Kolena. https://docs.kolena.com/metrics/tp-fp-fn-tn/?form=MG0AV3

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-January, 1800–1807. https://doi.org/10.1109/CVPR.2017.195

Cobilla, R., Carlo Dichoso, J., Minon, A. B., Kate Pascual, A., Abisado, M., Huyo-A, S. L., & Avelino Sampedro, G. (2023). Classification of the Type of Brain Tumor in MRI Using Xception Model. 2023 International Conference on Electronics, Information, and Communication, ICEIC 2023. https://doi.org/10.1109/ICEIC57457.2023.10049979

Devi, B. S., Chatrapati, K. S., & Sandhya, N. (2024). Enhanced Sugarcane Disease Detection Using DenseNet201 and DenseNet264 with Transfer Learning and Fine-tuning. Frontiers in Health Informatics, 13(3). www.healthinformaticsjournal.com

Elvionita, C., Sari, I. P., & Nuryastuti, T. (2023). Evaluation the Rationality of Clinical Outcomes of Antibiotic Use and Patterns of Bacterial Resistance to Antibiotics in Children with Pneumonia. Majalah Farmaseutik, 19(1). https://doi.org/10.22146/farmaseutik.v19i1.76103

Farida, Y., Trisna, A., & Nur, D. (2017). Study of Antibiotic Use on Pneumonia Patient in Surakarta Referral Hospital Studi Penggunaan Antibiotik Pada Pasien Pneumonia di Rumah Sakit Rujukan Daerah Surakarta. In Journal of Pharmaceutical Science and Clinical Research (Vol. 02).

Gonzalez, R. C. ., & Woods, R. E. . (2002). Digital image processing. Prentice Hall.

Gonzalez, R. C. ., & Woods, R. E. . (2008). Digital image processing. Prentice Hall.

Hafemann, L. G., Sabourin, R., & Oliveira, L. S. (2017). Learning features for offline handwritten signature verification using deep convolutional neural networks. Pattern Recognition, 70, 163–176. https://doi.org/https://doi.org/10.1016/j.patcog.2017.05.012

hart, gordon, Shen, B., Chen, M., & Choi, Y. (2024, December 3). Confusion Matrix: How to create and interpret confusion matrices to evaluate ML model performance · Testing with Kolena. https://docs.kolena.com/metrics/confusion-matrix/

Hasan, M. A., Haque, F., Roy, T., Islam, M., Nahiduzzaman, M., Hasan, M. M., Ahsan, M., & Haider, J. (2024). Prediction of fetal brain gestational age using multihead attention with Xception. Computers in Biology and Medicine, 182. https://doi.org/10.1016/j.compbiomed.2024.109155

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. http://arxiv.org/abs/1512.03385

Heaton, J. (2018). Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning. Genetic Programming and Evolvable Machines, 19(1–2), 305–307. https://doi.org/10.1007/s10710-017-9314-z

Hosna, A., Merry, E., Gyalmo, J., Alom, Z., Aung, Z., & Azim, M. A. (2022). Transfer learning: a friendly introduction. Journal of Big Data, 9(1). https://doi.org/10.1186/s40537-022-00652-w

Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2016). Densely Connected Convolutional Networks. http://arxiv.org/abs/1608.06993

Huy, V. T. Q., & Lin, C.-M. (2023). An Improved Densenet Deep Neural Network Model for Tuberculosis Detection Using Chest X-Ray Images. IEEE Access, 11, 42839–42849. https://doi.org/10.1109/ACCESS.2023.3270774

Iswari, R. A. (2021). KLASIFIKASI PNEUMONIA PADA GAMBAR X-RAY PARU-PARU MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK.

Kong, L., & Cheng, J. (2022). Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion. Biomedical Signal Processing and Control, 77. https://doi.org/10.1016/j.bspc.2022.103772

Koonsanit, K., Thongvigitmanee, S., Pongnapang, N., & Thajchayapong, P. (2017). Image enhancement on digital x-ray images using N-CLAHE. 2017 10th Biomedical Engineering International Conference (BMEiCON), 1–4. https://doi.org/10.1109/BMEiCON.2017.8229130

Kosarkar, N., Basuri, P., Karamore, P., Gawali, P., Badole, P., & Jumle, P. (2022). Disease Prediction using Machine Learning. 2022 10th International Conference on Emerging Trends in Engineering and Technology - Signal and Information Processing (ICETET-SIP-22), 1–4. https://doi.org/10.1109/ICETET-SIP-2254415.2022.9791739

Manafe, D. (2020). Covid-19 Masih Panjang, Indonesia Kekurangan 1.294 Dokter Paru. https://www.beritasatu.com/news/674573/covid19-masih-panjang-indonesia-kekurangan-1294-dokter-paru

Marchand, W. R. (2022). The Rising Popularity of Python - https://pythoncircle.com. https://pythoncircle.com/post/763/the-rising-popularity-of-python/?form=MG0AV3

Mathofani, P. E., Febriyanti, R., Studi, P., Masyarakat, K., & Abstrak, U. F. (2020). Faktor-Faktor Yang Berhubungan Dengan Kejadian Penyakit Tuberkulosis (TB) Paru di Wilayah Kerja Puskesmas Serang Kota Tahun 2019 The Factors Associated With The Incidence Of Pulmonary Tuberculosis In The Working Area Of Serang City Health Center 2019. In Jurnal Ilmiah Kesehatan Masyarakat (Vol. 12).

Nurin, F. (2024). 9 Faktor Risiko Pneumonia yang Perlu Anda Waspadai. https://hellosehat.com/pernapasan/pneumonia/faktor-risiko-pneumonia/?form=MG0AV3

Pan, C., Lian, L., Chen, J., & Huang, R. (2023). FemurTumorNet: Bone tumor classification in the proximal femur using DenseNet model based on radiographs. Journal of Bone Oncology, 42. https://doi.org/10.1016/j.jbo.2023.100504

Panthakkan, A., Anzar, S. M., Jamal, S., & Mansoor, W. (2022). Concatenated Xception-ResNet50 — A novel hybrid approach for accurate skin cancer prediction. Computers in Biology and Medicine, 150. https://doi.org/10.1016/j.compbiomed.2022.106170

Patterson, J., & Gibson, A. (2017). Deep Learning: A Practitioner’s Approach. O’Reilly Media. https://books.google.co.id/books?id=qrcuDwAAQBAJ

Pittara. (2022, May 10). Pencegahan Pneumonia - Alodokter. https://www.alodokter.com/pneumonia/pencegahan?form=MG0AV3

Pranita, E., & Sumartiningtyas, H. K. N. (2020). Indonesia Peringkat Ketujuh Kematian Balita akibat Pneumonia di Dunia, Apa Sebabnya? Halaman all - Kompas.com. https://www.kompas.com/sains/read/2020/11/09/173300723/indonesia-peringkat-ketujuh-kematian-balita-akibat-pneumonia-di-dunia-apa?page=all

Rahman, T., Baras, A. S., & Chellappa, R. (2025). Evaluation of a Task-Specific Self-Supervised Learning Framework in Digital Pathology Relative to Transfer Learning Approaches and Existing Foundation Models. Modern Pathology, 38(1). https://doi.org/10.1016/j.modpat.2024.100636

Ramdhan, A., Bustomi, M. A., & Faridawati. (2014). KLASIFIKASI CITRA RONTGEN PARU PARU DENGAN EKSTRAKSI FITUR HISTOGRAM DAN METODE NAIVE BAYES.

Rindrasari, R. (2018). KLASIFIKASI KELAS RISIKO PASIEN PNEUMONIA MENGGUNAKAN METODE SUPPORT VECTOR MACHINE-GENETIC ALGORITHM (SVM-GA) HYBRID.

Rokhana, R., Priambodo, J., Karlita, T., Made, I., Sunarya, G., Yuniarno, E. M., Purnama, K. E., & Purnomo, M. H. (2019). Convolutional Neural Network untuk Pendeteksian Patah Tulang Femur pada Citra Ultrasonik B-Mode. In JNTETI (Vol. 8, Issue 1).

Rossum, G. van. (1996). Foreword for “Programming Python” (1st ed.) | Python.org. https://www.python.org/doc/essays/foreword/

Sadik, R., Majumder, A., Biswas, A. A., Ahammad, B., & Rahman, M. M. (2023). An in-depth analysis of Convolutional Neural Network architectures with transfer learning for skin disease diagnosis. Healthcare Analytics, 3. https://doi.org/10.1016/j.health.2023.100143

Saleh, A. W., Gupta, G., Khan, S. B., Alkhaldi, N. A., & Verma, A. (2023). An Alzheimer’s disease classification model using transfer learning Densenet with embedded healthcare decision support system. Decision Analytics Journal, 9. https://doi.org/10.1016/j.dajour.2023.100348

Sartiwi, W., Rika Nofia, V., Komala Sari, I., Studi, P. S., & Syedza Saintika, S. (n.d.). Jurnal Abdimas Saintika LATIHAN BATUK EFEKTIF PADA PASIEN PNEUMONIA DI RSUD SAWAHLUNTO. https://jurnal.syedzasaintika.ac.id

Scholar, M. T., & Bagane, M. (2024). Google Colab: The Free Cloud Platform Powering Machine Learning. 12, 1.

Sewak, M., Karim, Md. R., & Pujari, P. (2018). Practical Convolutional Neural Networks: Implement advanced deep learning models using Python. Packt Publishing.

Sharma, A., Khan, F., Sharma, D., Gupta, S., & Student, F. Y. (2020). Python: The Programming Language of Future.

Soewu, T., Uday Kalyan, S. V, Rakhra, M., & Singh, D. (2022). Lung Cancer Detection using Image Processing. 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), 1206–1211. https://doi.org/10.1109/IC3I56241.2022.10072589

Sulistiyani Ratna, S., Setyawan Arianto, F., & Komarudin, M. (2016). PENGOLAHAN CITRA; DASAR DAN CONTOH PENERAPANNYA.

Tang, X., & Rashid Sheykhahmad, F. (2024). Boosted dipper throated optimization algorithm-based Xception neural network for skin cancer diagnosis: An optimal approach. Heliyon, 10(5). https://doi.org/10.1016/j.heliyon.2024.e26415

TensorFlow. (n.d.). Retrieved December 25, 2024, from https://www.tensorflow.org/

Tharwat, A. (2018). Classification assessment methods. Applied Computing and Informatics, 17(1), 168–192. https://doi.org/10.1016/j.aci.2018.08.003

Upasana, C., Tewari, A. S., & Singh, J. P. (2022). An Attention-based Pneumothorax Classification using Modified Xception Model. Procedia Computer Science, 218, 74–82. https://doi.org/10.1016/j.procs.2022.12.403

Virnodkar, S. S., Pachghare, V. K., Patil, V. C., & Jha, S. K. (2022). CaneSat dataset to leverage convolutional neural networks for sugarcane classification from Sentinel-2. Journal of King Saud University - Computer and Information Sciences, 34(6), 3343–3355. https://doi.org/10.1016/j.jksuci.2020.09.005

Wikanargo, M. A., & Thenata, A. P. (2018). IMAGE SEGMENTATION OF CHEST X-RAYS FOR ABNORMALITY PATTERN RECOGNATION IN LUNGS USING FUZZY C-MEANS METHOD. Jurnal Terapan Teknologi Informasi, 2(2), 101–111. https://doi.org/10.21460/jutei.2018.22.98

world health organizations. (2017). Global tuberculosis report 2017. World Health Organization.

Zeng, T., Chen, Y., Zhu, D., Huang, Y., Huang, Y., Chen, Y., Shi, J., Ding, B., & Huang, J. (2024). AI diagnostics in bone oncology for predicting bone metastasis in lung cancer patients using DenseNet-264 deep learning model and radiomics. Journal of Bone Oncology, 48. https://doi.org/10.1016/j.jbo.2024.100640

Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2021). Dive into Deep Learning — Dive into Deep Learning 1.0.3 documentation. https://d2l.ai/