

**PENERAPAN METODE CNN-LSTM DALAM KLASIFIKASI
PENYAKIT PARU BERDASARKAN CITRA X-RAY**

SKRIPSI

FRANDA CHRISTIANO SIAHAAN

201402089



**PROGRAM STUDI S1 TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA
MEDAN
2024**

**PENERAPAN METODE CNN-LSTM DALAM KLASIFIKASI
PENYAKIT PARU BERDASARKAN CITRA X-RAY**

SKRIPSI

Diajukan untuk melengkapi tugas dan memenuhi syarat memperoleh ijazah Sarjana
Teknologi Informasi

FRANDA CHRISTIANO SIAHAAN

201402089



**PROGRAM STUDI S1 TEKNOLOGI INFORMASI
FAKULTAS ILMU KOMPUTER DAN TEKNOLOGI INFORMASI
UNIVERSITAS SUMATERA UTARA**

MEDAN

2024

PERSETUJUAN

Judul : PENERAPAN METODE CNN-LSTM DALAM
KLASIFIKASI PENYAIT PARU
BERDASARKAN CITRA X-RAY

Kategori : Skripsi

Nama Mahasiswa : Franda Christiano Siahaan

Nomor Induk Mahasiswa : 201402089

Program Studi : Sarjana (S-1) Teknologi Informasi

Fakultas : Ilmu Komputer dan Teknologi Informasi
Universitas Sumatera Utara

Medan, 14 Juni 2024

Komisi Pembimbing:

Pembimbing 2,

Dr. Mohammad Andri Budiman S.T.,
M.Comp.Sc., M.E.M.
NIP. 197510082008011011

Pembimbing 1,

Annisa Fadhilah Pulungan
S.Kom., M.Kom
NIP. 199308092020012001

Diketahui/disetujui oleh
Program Studi S-1 Teknologi Informasi
Ketua,

Deddy Wandi S.T., M.Kom.
NIP. 197908312009121002

PERNYATAAN

PENERAPAN METODE CNN-LSTM DALAM KLASIFIKASI PENYAKIT PARU BERDASARKAN CITRA X-RAY

Saya mengakui bahwa skripsi ini adalah hasil karya saya sendiri, kecuali beberapa kutipan dan ringkasan yang masing-masing telah disebutkan sumbernya.

Medan, 25 Juni 2024



Franda Christiano Siahaan

201402089

UCAPAN TERIMA KASIH

Segala puji dan syukur penulis panjatkan ke hadirat Tuhan yang Maha Esa yang atas kasih karunia-Nya kepada penulis sehingga mampu merampungkan tugas akhir ini yang menjadi salah satu ketentuan kelulusan penulis dan mendapat gelar Sarjana Komputer pada program studi S1 Teknologi Informasi Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara.

Penulis dalam penyusunan tugas akhir ini sadar bahwa pencapaian ini tidak akan tergapai tanpa adanya masukan dari banyak pihak. Pada oportunitas ini penulis beritikad menuturkan terima kasih yang sebesar-besarnya kepada:

1. Keluarga penulis yaitu ibu Tiurma Siallagan S.H., abang Perystito Siahaan S.Si. serta keluarga dan saudara-saudari yang lain yang telah mendukung, mendoakan dan memotivasi penulis untuk menyelesaikan skripsi.
2. Ibu Dr. Maya Silvi Lydia, B.Sc., M.Sc. selaku Dekan Fasilkom-TI Universitas Sumatera Utara.
3. Bapak Dedy Arisandi S.Kom., M.Kom. selaku Ketua Program Studi S1 Teknologi Informasi Universitas Sumatera Utara.
4. Bapak Ivan Jaya S.Si., M.Kom. selaku Sekretaris Program Studi S1 Teknologi Informasi Universitas Sumatera Utara.
5. Ibu Annisa Fadhillah Pulungan S.Kom., M.Kom. sebagai dosen pembimbing I dan Bapak Dr. Mohammad Andri Budiman S.T., M.Comp.Sc., M.E.M. sebagai dosen pembimbing II penulis yang telah membagikan ilmu, kritik, saran dan motivasi kepada penulis.
6. Dosen penguji I dan dosen penguji II yang memberikan masukan beserta saran dalam menyempurnakan skripsi ini.
7. Seluruh dosen program studi S1 Teknologi Informasi atas ilmu dan pengalaman kepada penulis.
8. Seluruh staff dan pegawai Fakultas Ilmu Komputer dan Teknologi Informasi Universitas Sumatera Utara atas bantuannya kepada penulis dalam mengurus semua hal tentang administrasi selama masa perkuliahan penulis.
9. Teman-teman penulis dari awal, pertengahan hingga akhir masa perkuliahan yaitu Hizkia Winter, Reinhart Domingo, Raihan Alifya, Muhammad Ridwan

Rizki, Kelvin Natanael, Bertrand Toby, Muhammad Iqbal Manalu yang telah berjuang bersama penulis dalam menyelesaikan perkuliahan.

10. Teman-teman dekat SMA penulis dari “KKF” yang memberikan semangat untuk merampungkan tugas akhir ini.
11. Teman-teman dekat SMP penulis yang juga memberikan semangat untuk merampungkan tugas akhir ini.
12. Teman seperjuangan TI USU angkatan 2020 yang bersama penulis melewati masa-masa perkuliahan.
13. Beberapa senior yang memberikan ilmu dan masukan kepada penulis yaitu bang Ammar, bang Saddam Zikri Dalimunthe dan kak Fenni Sarumaha yang bersedia menjawab pertanyaan penulis selama mengerjakan tugas akhir.
14. Seluruh kerabat dan teman-teman lainnya yang tidak bisa disebutkan namanya.

Penulis menyadari bahwa masih terdapat kekurangan pada tugas akhir ini, penulis berharap akan ulasan dan saran yang dapat membantu menyempurnakan tugas akhir ini.

Medan, Maret 2024

Penulis

PENERAPAN METODE CNN-LSTM DALAM KLASIFIKASI PENYAKIT PARU BERDASARKAN CITRA X-RAY

ABSTRAK

Tuberculosis, pneumonia dan kanker paru adalah penyakit paru yang sering diderita oleh masyarakat di Indonesia. Kemudian pada awal tahun 2020 muncul penyakit baru yang masuk ke Indonesia yaitu COVID-19. Gejala antara penyakit pneumonia dengan COVID-19 memiliki kemiripan, sehingga terkadang sulit dibedakan tanpa pemeriksaan medis. Tuberculosis dan kanker paru memiliki suatu ciri yang sama yaitu adanya nodul pada paru. Nodul paru merupakan sebuah opasitas yang berdiameter kurang dari tiga sentimeter dan apabila lebih dari tiga sentimeter maka harus dipertimbangkan apakah memiliki tanda-tanda kanker paru. Diagnosis penyakit paru dapat dilakukan melalui pemeriksaan X-Ray oleh dokter akan tetapi tidak dapat dipungkiri bahwa kesalahan diagnosis mungkin saja dapat terjadi oleh karena sesuatu hal. Untuk mengurangi kesalahan diagnosis tersebut, diperlukan suatu sistem yang dapat mengklasifikasi keempat penyakit tersebut berdasarkan citra X-Ray paru. CNN-LSTM adalah metode yang diterapkan pada penelitian ini. Tahapan yang dilakukan sebelum proses klasifikasi yaitu *grayscale*, *pre-processing* dan *image augmentation*, selanjutnya akan mengekstrak fitur yang ada pada citra oleh CNN dan kemudian diklasifikasi menggunakan LSTM. Data yang digunakan berjumlah 1250 yang terdiri dari 875 data *training*, 250 data *validation* dan 125 data *testing*. Berdasarkan pengujian yang dilakukan, sistem mampu mengklasifikasi penyakit pada paru dengan akurasi 93,6%.

Kata Kunci: Penyakit Paru, CNN-LSTM, Klasifikasi Gambar, Pengolahan Citra.

IMPLEMENTATION OF CNN-LSTM METHOD IN LUNG DISEASE CLASSIFICATION BASED ON X-RAY IMAGE

ABSTRACT

Tuberculosis, pneumonia and lung cancer are lung diseases that are often suffered by people in Indonesia. Then in early 2020, a new disease that entered Indonesia, namely COVID-19 emerged. The symptoms between pneumonia and COVID-19 are similar so it is sometimes difficult to distinguish without a medical examination. Tuberculosis and lung cancer share a common feature, namely the presence of nodules in the lungs. A lung nodule is an opacity that is less than three centimeters in diameter and if it is more than three centimeters then it should be considered whether it has signs of lung cancer. Diagnosis of lung disease can be done through X-Ray examination by a doctor but it cannot be denied that misdiagnosis may occur for some reason. To reduce these misdiagnoses, a system that can classify diseases from X-Ray images is a necessity. The method applied in this research is CNN-LSTM. The stages conducted before the classification process are gray scaling, pre-processing and image augmentation. After those process, the features in the image are extracted using CNN and then classified using LSTM. The data used amounted to 1250 consisting of 875 training data, 250 validation data and 125 testing data. Based on the tests conducted, the system is able to classify those four lung diseases with 93.6% accuracy.

Keywords: Lung Diseases, CNN-LSTM, Image Classification, Image Processing.

DAFTAR ISI

PERSETUJUAN	iii
PERNYATAAN	iv
UCAPAN TERIMA KASIH	v
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
Daftar Tabel	xii
Daftar Gambar	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Manfaat Penelitian	3
1.6 Metodologi Penelitian	4
1.6.1 Studi Literatur	4
1.6.2 Analisis Permasalahan	4
1.6.3 Perancangan Sistem	4
1.6.4 Implementasi	4
1.6.5 Pengujian Sistem	4
1.6.6 Penyusunan Laporan	4
1.7 Sistematika Penulisan	4
BAB II LANDASAN TEORI	6
2.1 Pneumonia	6

2.2 Tuberkulosis	6
2.3 COVID-19	7
2.4 <i>Lung Nodules</i>	7
2.5 Citra Digital	8
2.5.1 <i>Citra Grayscale</i>	8
2.6 Pemrosesan Citra Digital	9
2.7 Augmentasi Data	9
2.8 <i>Convolutional Neural Network (CNN)</i>	10
2.9 <i>Long Short-Term Memory (LSTM)</i>	11
2.10 CNN-LSTM	12
2.11 <i>Confusion Matrix</i>	13
2.11.1 <i>Precision</i>	14
2.11.2 <i>Recall</i>	14
2.11.3 <i>F1-Score</i>	14
2.12 <i>Confidence Score</i>	14
2.13 <i>Labelling</i>	15
2.14 <i>Cropping</i>	15
2.15 <i>Resizing</i>	15
2.16 Penelitian Terdahulu	15
2.17 Perbedaan Penelitian	18
2.18 Dataset yang digunakan	18
BAB III ANALISIS DAN PERANCANGAN SISTEM	20
3.1 Data yang digunakan	20
3.2 Analisis sistem	20
3.2.1 <i>Image Acquisition</i>	22
3.2.2 <i>Pre-Processing</i>	22
3.2.3 <i>Image Classification</i>	27

3.2.4 <i>Learned Model</i>	32
3.2.5 <i>Proses Training</i>	32
3.2.6 <i>Proses Validation</i>	35
3.2.7 <i>Proses Testing</i>	35
3.2.8 <i>Model Evaluation</i>	36
3.3 Perancangan Sistem	36
3.3.1 Desain Antarmuka	36
BAB IV Implementasi dan Pengujian	39
4.1 Implementasi Sistem	39
4.1.1 Perangkat Keras dan Perangkat Lunak	39
4.2 Implementasi Data	39
4.2.1 Implementasi Augmentasi Data	41
4.3 Implementasi Desain Sistem	42
4.3.1 Tampilan Utama	42
4.4 Prosedur Operasional	44
4.5 <i>Training model</i>	44
4.6 Pengujian Sistem	46
BAB V Kesimpulan dan Saran	52
5.1 Kesimpulan	52
5.2 Saran	52
Daftar Pustaka	53
Lampiran	57

Daftar Tabel

Tabel 2.1 Penelitian Terdahulu	16
Tabel 3.1 Pembagian data Training, Validation dan Testing.	22
Tabel 4.1 Pembagian data setelah proses augmentasi	42
Tabel 4.2 Hasil training data dengan berbagai kombinasi hyperparameter	45
Tabel 4.3 Hasil Pengujian Sistem	47
Tabel 4.4 Confusion Matrix	48
Tabel 4.5 Nilai TP, FP dan FN	48
Tabel 4.6 Nilai Precision, Recall dan F1-Score	49
Tabel 4.7 Hasil klasifikasi setelah implementasi CLAHE	50

Daftar Gambar

Gambar 2.1 Pneumonia (Mooney, 2018)	6
Gambar 2.2 Tuberkulosis (Rahman, et al. 2021)	7
Gambar 2.3 COVID-19 (JTIPTJ, 2021)	7
Gambar 2.4 Lung Nodules (Raddar, 2020)	8
Gambar 2.5 Citra Grayscale oleh https://unsplash.com/@hulkiokantabak	9
Gambar 2.6 Arsitektur CNN (Alom, et al. 2019)	11
Gambar 2.7 Arsitektur LSTM (Islam, et al. 2020)	12
Gambar 2.8 Arsitektur CNN-LSTM (Islam, <i>et al.</i> 2020)	13
Gambar 3.1 Citra Gambar Paru-Paru Tuberculosis (Rahman, et al. 2021)	20
Gambar 3.2 Arsitektur Umum	21
Gambar 3.3 Labelling	23
Gambar 3.4 Citra paru-paru sebelum dan setelah cropping	23
Gambar 3.5 Proses grayscaling	24
Gambar 3.6 Sebelum dan setelah Horizontal flip	24
Gambar 3.7 Sebelum dan setelah rotasi 30 derajat	25
Gambar 3.8 Sebelum dan setelah shear sebesar 0.1	25
Gambar 3.9 Sebelum dan setelah height shift range sebesar 0.1	25
Gambar 3.10 Sebelum dan setelah width whift range sebesar 0.1	25
Gambar 3.11 Sebelum dan setelah zoom sebesar 0.25	26
Gambar 3.12 Kombinasi Arsitektur CNN-LSTM	27
Gambar 3.13 Proses konvolusi	28
Gambar 3.14 Posisi perkalian matriks citra terhadap matriks filter	29
Gambar 3.15 Perkalian matriks citra dengan matriks kernel	30
Gambar 3.16 Feature map	31
Gambar 3.17 Feature map setelah ReLU	31
Gambar 3.18 Max Pooling dengan stride 2 dan kernel 2x2	32
Gambar 3.19 Penerapan LSTM layer	32
Gambar 3.20 Rancangan halaman pilih gambar	36
Gambar 3.21 Rancangan halaman hasil klasifikasi gambar	37
Gambar 3.22 Activity diagram	38
Gambar 4.1 Data Citra Pneumonia	40
Gambar 4.2 Data Citra Tuberculosis	40

Gambar 4.3 Data Citra Paru-Paru Normal	41
Gambar 4.4 Data Citra COVID-19	41
Gambar 4.5 Data Citra Lung Nodules	41
Gambar 4.6 Tampilan Utama sebelum memilih gambar.	43
Gambar 4.7 Tampilan Utama setelah mengklik tombol classify.	43
Gambar 4.8 Tampilan Utama setelah proses klasifikasi.	44
Gambar 4.9 Akurasi Training dan Validation pada model	45
Gambar 4.10 Loss Training dan Validation pada model	46

BAB I

PENDAHULUAN

1.1 Latar Belakang

Paru-paru adalah organ penting dalam sistem pernafasan yang berperan sebagai tempat terjadinya pertukaran antara oksigen dan karbondioksida. Karena kualitas udara yang tercemar dan lingkungan yang tidak sehat, paru-paru menjadi salah satu organ respirasi dasar yang diserang oleh penyakit (Rahmadewi & Kurnia 2016). Contoh penyakit paru yaitu pneumonia, Tuberculosis, COVID-19 dan *lung nodules*. Indonesia menduduki posisi kedua setelah India terkait penyakit tuberculosis dengan 969 ribu jumlah kasus dengan banyak kematian per tahun adalah 93 ribu yang setara dengan 11 kematian per jam (Rokom, 2022).

Salah satu metode dalam mendiagnosis penyakit pada paru-paru adalah dengan menggunakan citra medis. Salah satu peran penting dari citra medis adalah mengidentifikasi suatu penyakit. Banyak cara yang dapat digunakan untuk mengidentifikasi penyakit melalui citra medis, salah satunya adalah X-Ray. X-Ray atau dapat disebut radiografi bekerja dengan mengirimkan radiasi ke seluruh tubuh. Area dengan kadar kalsium yang tinggi seperti tulang dan gigi menghalangi radiasi dan membuat mereka tampak putih pada gambar sedangkan jaringan lunak memungkinkan radiasi melewatinya yang membuat mereka akan tampak abu-abu atau hitam pada gambar. X-Ray adalah metode pencitraan tercepat dan paling mudah untuk diakses yang membuat X-Ray menjadi metode yang paling umum digunakan untuk mengidentifikasi penyakit pada paru-paru. Pemeriksaan X-Ray membutuhkan dokter maupun medis yang ahli dalam membaca citra X-Ray tersebut dan terkadang membutuhkan waktu yang lama dan subyektif yang dapat menyebabkan terjadi kesalahan identifikasi oleh dokter maupun medis tersebut. Terdapat pula masalah keterbatasan ahli radiologi di berbagai sistem layanan kesehatan terutama di negara berkembang dan daerah terpencil. COVID-19 memiliki kemungkinan pola patologi yang mirip dengan Pneumonia sehingga sulit untuk membedakan X-Ray COVID-19 dengan Pneumonia (Hariri & Avşar 2023).

Deep learning yang merupakan *subset* dari machine learning memiliki dampak yang besar di beberapa bidang keilmuan selama beberapa tahun terakhir tidak terkecuali bidang kesehatan. Hal ini dibuktikan dengan teknologi *Deep Learning* mampu mengungguli metode lain dalam pemrosesan dan analisis citra (Egger, *et al.* 2022). Beberapa penelitian yang pernah dilakukan untuk mendeteksi penyakit menggunakan Deep Learning seperti: Pneumothorax (Malhotra, *et al.* 2022), COVID-19 (Wang, *et al.* 2023), Pneumonia (Maysanjaya, 2020), Tuberculosis (Oloko-Oba & Viriri, 2022), kanker paru (Wang, 2022), tumor otak (Zeineldin, *et al.* 2023), kanker payudara (Nasser & Yusof, 2023) dan kanker kulit (Mehr & Ameri, 2022).

Pada penelitian untuk mendeteksi Pneumonia, Tuberculosis dan COVID-19 yang dilakukan oleh Septhyan *et al.* (2022) diperoleh akurasi 95%. Penelitian lainnya oleh Diar *et al.* (2022) yang membandingkan empat kelas yaitu normal, viral pneumonia, COVID-19 dan tuberculosis dengan menggunakan preprocessing CLAHE memperoleh hasil akurasi sebesar 94,687 %. Ada juga penelitian untuk mendeteksi Pneumonia menggunakan Autoencoder dan four-layer convolutional network (FCNN) memperoleh akurasi sebesar 98% (El-Shafai, *et al.* 2022) dan pada penelitian Huang & Liao pada tahun 2022 yang mengklasifikasi lima kelas penyakit pada paru diperoleh akurasi sebesar 98.71% untuk Stacking Ensemble dan 99.21 untuk ECA-EfficientNetV2.

Berdasarkan penelitian sebelumnya dapat disimpulkan bahwa metode CNN mampu mengenali citra X-Ray paru. Untuk mendapatkan hasil yang lebih maksimal, pada penelitian ini penulis menambah jaringan LSTM didalam arsitektur CNN yang digunakan. CNN digunakan untuk *feature extraction* sedangkan LSTM digunakan untuk mengklasifikasi penyakit. Jaringan LSTM masih dapat menyimpan pengetahuan yang sudah dipelajari sebelumnya didalam memori internalnya. Pada *fully connected networks*, setiap layer terhubung dengan layer sebelumnya sedangkan *node* di tiap layer tidak terhubung dan hanya memproses satu input. Dengan menerapkan LSTM, *node-node* akan terhubung melalui sebuah graf berarah sepanjang urutan waktu tertentu yang diperlakukan sebagai input dengan urutan tertentu (Islam, *et al.* 2020).

Oleh karena itu, sistem yang dikembangkan diharapkan dapat mengklasifikasi penyakit paru-paru yang terdiri dari empat penyakit yaitu Pneumonia, *Coronavirus Disease-19* (COVID-19), Tuberkulosis dan *Lung Nodules*.

1.2 Rumusan Masalah

Pembacaan citra medis biasanya dilakukan oleh ahli radiologi dan dokter, namun metode ini bisa saja menghasilkan kesalahan diagnosis. Untuk memberikan diagnosis yang keakuratannya mendekati dokter dan ahli, diperlukan suatu pendekatan yang dapat membantu atau menjadi opsi lain dalam mendeteksi penyakit berdasarkan citra X-Ray.

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah menerapkan metode CNN-LSTM untuk mengklasifikasi penyakit paru yang datanya berupa citra X-Ray.

1.4 Batasan Masalah

Supaya penelitian ini terarah dan tercapainya tujuan penelitian yang diharapkan maka diperlukan batasan masalah. Batasan masalah penelitian ini yaitu:

1. Penyakit paru yang diklasifikasi hanya empat yaitu Pneumonia, Tuberkulosis, COVID-19 dan *Lung Nodules* dengan ekstensi file .jpg, .png dan .jpeg yang diambil dari Kaggle.
2. Hanya dapat mengidentifikasi satu jenis penyakit pada satu citra
3. Tidak dapat menentukan letak posisi keberadaan 3indaka pada paru yang menentukan jenis penyakit yang menyerang paru.

1.5 Manfaat Penelitian

Berikut adalah manfaat yang diberikan dari penelitian ini yaitu:

1. Menjadi opsi lain bagi medis untuk membantu dalam mempercepat proses mendiagnosa penyakit pada paru.
2. Mengurangi terjadinya kesalahan diagnosis penyakit paru.

1.6 Metodologi Penelitian

Metodologi yang diterapkan penelitian ini diuraikan sebagai berikut.

1.6.1 Studi Literatur

Penulis mengumpulkan beberapa referensi yang diambil dari buku, jurnal, skripsi, artikel dan sumber lainnya yang terkait dengan penelitian.

1.6.2 Analisis Permasalahan

Pada tahap ini penulis akan memahami konsep dari metode CNN-LSTM dan bagaimana konsep penggabungan kedua metode tersebut untuk mengklasifikasi penyakit paru.

1.6.3 Perancangan Sistem

Penulis merancang arsitektur umum yang digunakan pada sistem.

1.6.4 Implementasi

Pada tahap implementasi, penulis akan melakukan pengimplementasian metode CNN-LSTM pada sistem berbasis web yang dibangun.

1.6.5 Pengujian Sistem

Setelah sistem berhasil dibangun, selanjutnya akan dilakukan pengujian sistem dari metode CNN-LSTM dalam melakukan klasifikasi penyakit paru.

1.6.6 Penyusunan Laporan

Pada tahap ini penulis mendokumentasikan hasil penelitian yang dilakukan.

1.7 Sistematika Penulisan

Sistematika penulisan pada skripsi ini terdiri atas lima bab sebagai berikut:

BAB 1: PENDAHULUAN

Bab pertama memuat latar belakang dari penelitian klasifikasi empat penyakit paru berdasarkan citra X-Ray, rumusan masalah yang menjadi alasan dari penelitian ini, tujuan penelitian klasifikasi penyakit paru berdasarkan citra X-Ray, batasan masalah terkait penggunaan citra X-ray yang digunakan, manfaat penelitian, metodologi penelitian yang berisi studi literatur, analisis permasalahan, perancangan sistem, implementasi, pengujian sistem dan penyusunan laporan.

BAB 2: LANDASAN TEORI

Bab kedua memuat teori-teori yang terkait dengan penelitian. Beberapa teori yang dimuat diantaranya adalah teori terkait Pneumonia, Tuberkulosis, COVID-19, *Lung Nodules* dan arsitektur CNN-LSTM.

BAB 3: ANALISIS DAN PERANCANGAN

Bab ketiga berisi perancangan arsitektur umum dari CNN-LSTM yang digunakan disertai dengan *pseudocode*, analisis data citra X-Ray yang digunakan dan perancangan sistem web.

BAB 4: IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab keempat memuat penjelasan tentang pelaksanaan yang dilakukan seperti jumlah data setelah augmentasi, tampilan sistem web yang dibangun dan hasil pengujian menggunakan sistem.

BAB 5: KESIMPULAN DAN SARAN

Bab kelima berisi kesimpulan dari penelitian klasifikasi penyakit paru berdasarkan citra X-Ray yang dilakukan dan saran untuk penelitian selanjutnya.

BAB II

LANDASAN TEORI

2.1 Pneumonia

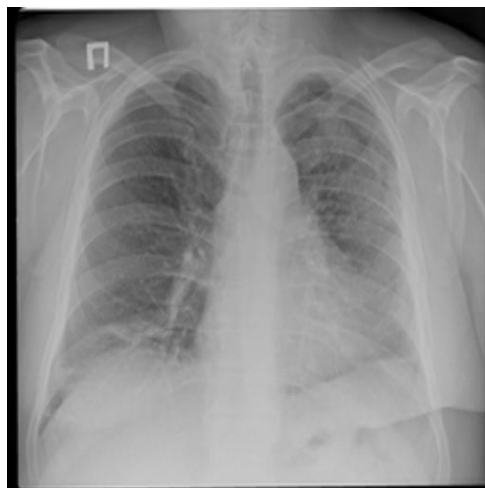
Pneumonia merupakan infeksi paru akut yang menginfeksi salah satu atau kedua paru-paru. Infeksi dari penyakit ini menyebabkan kantung udara atau alveoli pada paru-paru berisi cairan atau nanah. Bakteri, virus atau jamur merupakan sumber dari pneumonia. Gambar pneumonia ditampilkan oleh Gambar 2.1.



Gambar 2.1 Pneumonia (Mooney, 2018)

2.2 Tuberkulosis

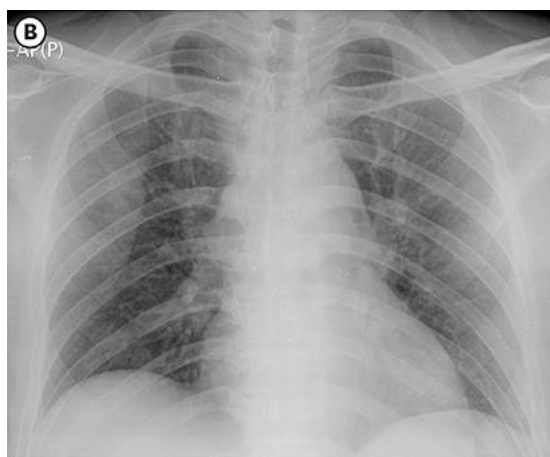
Tuberkulosis merupakan penyakit menular yang berasal dari bakteri *Mycobacteria Tuberculosis* yang kemudian ditularkan ke manusia melalui saluran pernafasan dan paling sering menyerang paru-paru. Gambar tuberkulosis ditampilkan oleh Gambar 2.2.



Gambar 2.2 Tuberkulosis (Rahman, *et al.* 2021)

2.3 COVID-19

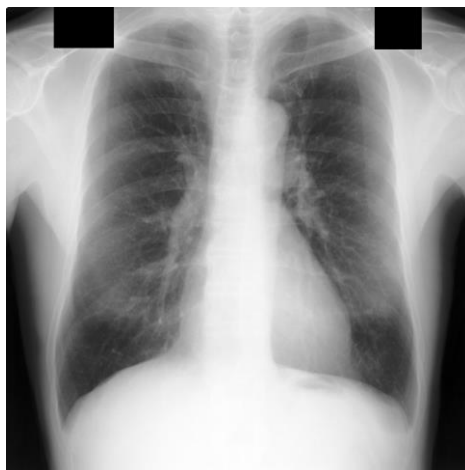
COVID-19 adalah penyakit infeksi paru yang berasal dari virus menular coronavirus 2 (SARS-CoV-2). Gambar COVID-19 ditampilkan oleh Gambar 2.3.



Gambar 2.3 COVID-19 (JTIPTJ, 2021)

2.4 Lung Nodules

Lung nodules merupakan pertumbuhan abnormal yang terbentuk di salah satu atau kedua paru-paru. *Lung nodules* yang bersifat jinak biasanya berukuran hingga 3 cm sedangkan jika lebih maka harus dipertimbangkan sebagai indikasi awal kanker paru-paru (Loverdos *et al.* 2019). Gambar Lung Nodules ditampilkan oleh Gambar 2.4.



Gambar 2.4 Lung Nodules (Raddar, 2020)

2.5 Citra Digital

Citra digital adalah proyeksi dari tampilan tiga dimensi ke bidang proyeksi dua dimensi dengan x dan y adalah letak posisi piksel gambar dan berisi nilai intensitas. Citra dapat disebut citra digital apabila nilai x , y dan intensitas citra bernilai diskrit (Tyagi, 2018). Citra digital secara umum dapat diartikan sebagai matriks representasi dari gambar dua dimensi yang menerapkan elemen sel poin berupa nilai angka terbatas yang disebut piksel. Setiap piksel direpresentasikan dalam bentuk nilai numerik. Pada citra *grayscale*, satu nilai 8 bit sudah cukup dalam menunjukkan intensitas piksel (antara 0 dan 255); citra berwarna memiliki tiga nilai yang menunjukkan besaran warna merah atau R, hijau atau G, dan biru atau B yang tersimpan. Jika gambar hanya terdiri dari dua intensitas, citra tersebut merupakan citra biner.

2.5.1 Citra *Grayscale*

Grayscale merupakan citra berwarna abu-abu yang didapat dari value yang berada di antara hitam (0) dan putih (255). Contoh citra *grayscale* ditampilkan oleh Gambar 2.5.



Gambar 2.5 Citra *Grayscale* oleh <https://unsplash.com/@hulkiokantabak>

2.5.1.1 Pengubahan citra RGB ke citra *grayscale*

Salah satu algoritma untuk mengubah citra berwarna ke citra *grayscale* adalah *Luminance* (Kanan & Cottrel, 2012). Teknik *Luminance* dilakukan dengan mengalikan setiap pixel citra dengan bobot warna R, G dan B yang kemudian dijumlahkan.

$$\gamma \leftarrow 0.2989R + 0.5870G + 0.1140B \quad (2.1)$$

2.6 Pemrosesan Citra Digital

Pemrosesan citra digital adalah penggunaan komputer digital untuk melakukan suatu urutan atau serangkaian tindakan pemrosesan pada suatu gambar.

2.7 Augmentasi Data

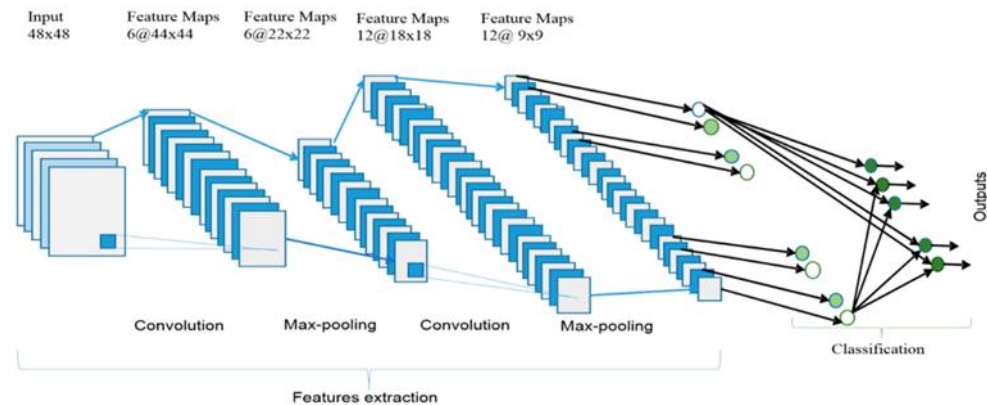
Salah satu cara untuk menghindari *overfitting* adalah dengan melatih model dengan data yang banyak. Hal tersebut dapat dicapai dengan augmentasi data. Serangkaian teknik yang dilakukan dalam augmentasi data dapat meningkatkan ukuran, kualitas dan keragaman dataset dengan membuat data baru dari data yang sudah ada sehingga dapat meningkatkan kinerja model *Deep Learning*.

Contoh augmentasi data berasal dari transformasi sederhana seperti pembalikan horizontal, augmentasi ruang warna, dan pemotongan acak. Transformasi ini mengkodekan banyak ketidakseragaman yang ada didalam data di mana hal ini memberikan tantangan pada tugas pengenalan gambar. Terdapat beberapa metode augmentasi seperti transformasi geometris, transformasi ruang warna, filter kernel,

pencampuran gambar, penghapusan acak, augmentasi ruang fitur, *adversarial training*, augmentasi berbasis GAN, *neural style transfer*, dan *meta-learning schemes*. Terdapat juga berbagai teknik augmentasi data yang berbasis manipulasi gambar dasar, yaitu: *geometric transformations*, *flipping*, *color space*, *rotation*, *scaling*, *shear*, *translation*, *cropping*, *color space transformations* dan *noise injection*.

2.8 Convolutional Neural Network (CNN)

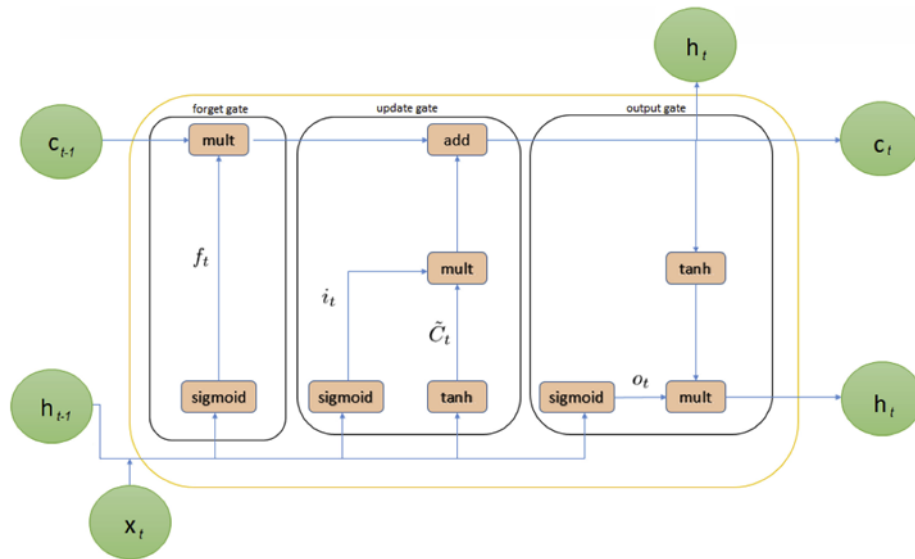
Convolutional Neural Network (CNN) adalah algoritma *deep learning* yang berkemampuan tinggi di berbagai tugas visual. CNN dapat digunakan untuk menganalisis, mengidentifikasi, mensegmentasi dan mengklasifikasikan gambar. Metode CNN bekerja dengan menyederhanakan gambar untuk analisis yang lebih baik. Desain arsitektur yang terdapat di CNN terdiri dari empat *layer* yaitu *layer* konvolusi, *pooling layer*, *nonlinearity layer* dan *fully connected layer* (Bezdan & Džakula 2019). *Layer* konvolusi berfungsi untuk menyaring gambar di mana bagian ini akan mengekstrak fitur penting yang terdapat pada gambar yang bertujuan untuk meningkatkan fitur. *Pooling layer* berfungsi untuk menurunkan atau mengurangi ukuran sampel gambar dari sampel fitur yang diekstraksi. Hal ini membuat pemrosesan menjadi lebih cepat, karena parameter berkurang. Dalam CNN, fungsi aktivasi yang umum digunakan adalah *Rectified Linear Unit*. Fungsi aktivasi digunakan untuk membentuk non-linearitas pada model. Hasil keluaran dari lapisan terakhir dari *layer* konvolusi akan menjadi input untuk *fully connected layer* atau disebut juga lapisan klasifikasi. Lapisan klasifikasi akan menggunakan jaringan syaraf tiruan feed-forward karena memiliki kinerja yang lebih baik (Hinton, *et al.* 2006); (Nair & Hinton, 2010). Kemudian setelah melewati lapisan yang terhubung penuh, output diprediksi ke dalam kelas-kelas. Detail arsitektur CNN ditampilkan oleh Gambar 2.6.



Gambar 2.6 Arsitektur CNN (Alom, *et al.* 2019)

2.9 Long Short-Term Memory (LSTM)

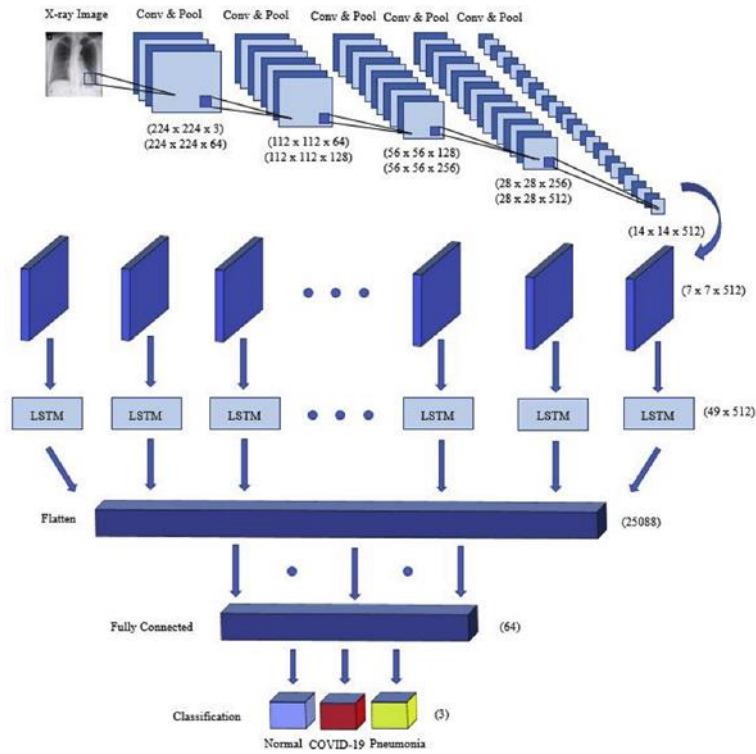
LSTM merupakan pengembangan dari RNN (*Recurrent Neural Network*) untuk mengatasi kekurangan dalam hal menyimpan informasi untuk waktu yang lama dan mengatasi masalah *vanishing* dan *exploding gradient* dengan menerapkan *memory block*. Tujuan dari pengembangan terhadap RNN ini adalah bahwa informasi yang telah lama diproses tersebut memiliki kemungkinan untuk dibutuhkan lagi di suatu kondisi tertentu pada saat proses pengumpulan informasi lebih lanjut. LSTM mewujudkannya dengan mampu menyimpan dan menghubungkan informasi yang telah diperoleh. LSTM terdiri dari tiga *gate* yang berada dalam satu *cell state*, yaitu: *forget*, *input* dan *output gate*. *Forget gate* berperan untuk menentukan *cell state* mana yang akan dipertahankan dari komputasi sebelumnya. *Input gate* berperan dalam mengontrol berapa banyak data input baru yang harus disimpan didalam *cell state*. *Output gate* berperan untuk menyempurnakan memori yang terpilih, memperbarui informasi dan memberikan output potongan informasi menggunakan fungsi *sigmoid* dan *tanh*. Detail arsitektur LSTM ditampilkan oleh Gambar 2.7.



Gambar 2.7 Arsitektur LSTM (Islam, *et al.* 2020)

2.10 CNN-LSTM

Dalam gabungan arsitektur CNN-LSTM, CNN berfungsi untuk mengekstrak fitur-fitur kompleks dari sebuah citra sedangkan LSTM berfungsi sebagai penentu kelas yang menggantikan *fully connected layer*. Gabungan dari arsitektur CNN dengan LSTM dapat mengolah informasi spasial dan temporal dari sebuah gambar yang berguna untuk klasifikasi (Fachrel, *et al.* 2023). Layer LSTM ditempatkan setelah *layer convolutional* terakhir untuk mendapatkan nilai *output* dari layer *convolutional* yang kemudian akan menjadi nilai *input* pada LSTM. Detail arsitektur CNN-LSTM ditampilkan oleh Gambar 2.8.



Gambar 2.8 Arsitektur CNN-LSTM (Islam, *et al.* 2020)

2.11 Confusion Matrix

Salah satu cara yang dipakai untuk melihat kemampuan model klasifikasi dalam membuat prediksi adalah dengan menggunakan *Confusion matrix*. Dari *confusion matrix* dapat dihitung skor untuk akurasi, presisi, *recall*, dan F1-score. Skor ini digunakan untuk menganalisis kinerja model. Nilai prediksi dan fakta aktual disajikan dalam *confusion matrix* dengan ukuran $n \times n$ di mana n adalah jumlah kelas (Fachrel, *et al.* 2023).

Confusion matrix merepresentasikan jumlah dari nilai yang diprediksi dan nilai aktual. TP atau *True Positive* adalah banyaknya data aktual positif dengan prediksi positif. TN atau *True Negative* menunjukkan banyak data aktual negatif dengan prediksi negatif. FP atau *False Positive* adalah banyak data aktual negatif dengan prediksi positif dan FN atau *False Negative* adalah banyaknya data aktual positif dengan prediksi negatif. Salah satu metrik yang digunakan dalam mengklasifikasi adalah akurasi. Keakuratan suatu model dapat dihitung dengan menggunakan Rumus 2.2 (Powers, 2011)

$$\text{Akurasi} = \frac{\text{Jumlah data dengan prediksi benar}}{\text{Jumlah keseluruhan data}} \quad (2.2)$$

2.11.1 Precision

Precision merupakan nilai yang menunjukkan seberapa akurat model untuk memprediksi nilai positif. Rumus mencari nilai *precision* ditampilkan oleh Persamaan 2.3 (Powers, 2011).

$$Precision = \frac{TP}{TP + FP} \quad (2.3)$$

2.11.2 Recall

Recall atau juga dikenal sebagai *Sensitivity* digunakan untuk mengukur kemampuan suatu model untuk memprediksi hasil yang positif dari seluruh data yang bernilai positif. Rumus mencari nilai *recall* ditampilkan oleh Persamaan 2.4 (Powers, 2011).

$$Recall = \frac{TP}{TP + FN} \quad (2.4)$$

2.11.3 F1-Score

F1-score merupakan metrik yang menggunakan nilai presisi dan nilai recall dari pengklasifikasi. *F1-score* adalah metrik lain yang umum digunakan dalam pengukuran klasifikasi. *F1-score* dihitung dengan menggunakan rata-rata *weighted harmonic* antara *precision* dan *recall*. Untuk klasifikasi data yang bernilai positif, hal ini membantu untuk memahami hubungan *tradeoff* antara ketepatan dan jumlah. Rumus untuk mencari nilai *F1-Score* terdapat pada Persamaan 2.5 (Powers, 2011 & 2015).

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2.5)$$

2.12 Confidence Score

Confidence score adalah metrik yang dihitung sebagai standar evaluasi. *Confidence score* menunjukkan probabilitas gambar terdeteksi dengan benar oleh algoritma yang digunakan yang disajikan dalam bentuk persentase.

2.13 Labelling

Labelling merupakan pemberian identitas atau nama pada setiap citra gambar. *Labelling* diperlukan sebagai pemberi informasi dari citra yang akan dipelajari sehingga model deep learning dapat memahami hubungan antara citra gambar yang dipelajari dengan labelnya.

2.14 Cropping

Cropping merupakan teknik memotong gambar untuk mendapatkan fokus yang lebih ke bagian gambar yang diinginkan.

2.15 Resizing

Resizing digunakan untuk menyesuaikan ukuran citra gambar sesuai dengan yang diinginkan pada input model. Tujuan lain dari melakukan *resizing* adalah untuk mengurangi beban komputasi yang diakibatkan oleh nilai *pixel* citra gambar yang terlalu besar.

2.16 Penelitian Terdahulu

Sudah banyak penelitian terhadap pendeteksian maupun pengklasifikasian penyakit berdasarkan citra X-Ray. Penelitian yang dilakukan oleh (Ahmed, M.S et al, 2023) untuk mengklasifikasi empat kelas (COVID-19, Pneumonia, Tuberculosis, Normal) menggunakan CNN mendapatkan rerata akurasi sebesar 98.72%. Penelitian yang dilakukan oleh Elshennawy N. M & Ibrahim D. M pada tahun 2020 untuk mendeteksi pneumonia dengan menggunakan empat model CNN berbeda yaitu vanilla CNN, dua pre-trained model (ResNet152V2 MobileNetV2) dan CNN-LSTM memperoleh hasil akurasi 92.19%, 99.12%, 96.48% dan 91.80% secara berurutan. Penelitian Fachrel J *et al* pada tahun 2023 menghasilkan f1-score sebesar 98.87% dalam mendeteksi Pneumonia pada paru-paru dengan menggunakan metode CNN-LSTM dengan tambahan penggunaan *10-Fold Cross-Validation*.

Tabel 2.1 Penelitian Terdahulu

No	Penulis	Judul	Tahun	Keterangan
1	Elshennawy, N.M, et al	Deep-Pneumonia Framework Using Deep Learning Models Based on Chest X-Ray Images	2020	Penelitian ini mendeteksi Pneumonia pada paru-paru dengan menggunakan empat model CNN berbeda yaitu vanilla CNN, dua pre-trained model yaitu ResNet152V2 MobileNetV2 dan CNN-LSTM. Diperoleh hasil untuk CNN, dua pre-trained model yaitu ResNet152V2, MobileNetV2 dan CNN-LSTM mendapatkan akurasi sebesar 92.19%, 99.12%, 96.48% dan 91.80%.
2	Islam M.Z., et al	A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images	2020	Penelitian ini mengklasifikasi Pneumonia dan COVID-19 pada paru-paru menggunakan CNN-LSTM. Penelitian ini memperoleh hasil F1-score sebesar 99.4%.
3	Win, K. Y., et al	Ensemble Deep Learning for the Detection of COVID-19 in Unbalanced	2021	Penelitian ini menggunakan Ensemble Learning dengan menggunakan sebelas transfer learning model. Penelitian ini mendapatkan hasil akurasi tertinggi yaitu 99.23% untuk

		Chest X-ray Dataset		mendeteksi Pneumonia dan COVID-19 menggunakan ensemble lima transfer learning model yaitu XceptionNet, MobileNetV2, DensetNet201, InceptionResNetV2 dan NasNetMobile dengan tambahan augmentasi data.
4	Mabrouk, A., et al	Pneumonia Detection on Chest X-ray Images Using Ensemble of Deep Convolutional Neural Networks	2022	Penelitian ini menggunakan Ensemble Learning dengan menggunakan tiga algoritma yaitu MobileNet, DenseNet dan Vision Tranformer. Diperoleh hasil akurasi sebesar 93.91%.
5	Ahmed, M. S., et al	Joint Diagnosis of Pneumonia, COVID-19, and Tuberculosis from Chest X-ray Images: A Deep Learning Approach	2023	Penelitian ini mengklasifikasi 4 kondisi paru-paru yaitu Pneumonia, COVID-19, Normal dan Tuberculosis menggunakan CNN dengan memperoleh hasil akurasi validasi sebesar 98.72%
6	Fachrel, J., et al	Enhancing an Imbalanced Lung Disease X-ray Image Classification	2023	Penelitian ini mendeteksi Pneumonia pada paru-paru dengan menggunakan metode CNN-LSTM dengan tambahan penggunaan 10-Fold Cross-Validation.

		with the CNN-LSTM Model		Penelitian ini memperoleh hasil F1-score sebesar 98.87%.
7	Huang, M. L. & Liao, Y. C.	Stacking Ensemble and ECA-EfficientNetV2 Convolutional Neural Networks on Classification of Multiple Chest Diseases Including COVID-19	2023	Salah satu metode penelitian yang dilakukan adalah <i>stacking-ensemble</i> enam EfficientNetV2 untuk mengklasifikasi penyakit paru yaitu tuberculosis, bacterial pneumonia, COVID-19 dan viral pneumonia dengan hasil akurasi sebesar 98.71(± 0.26)

2.17 Perbedaan Penelitian

Perbedaan penelitian penulis dengan penelitian Islam M.Z. *et al* dan Fachrel, J. *et al* adalah penelitian yang dilakukan oleh kedua peneliti tersebut menggunakan CNN-LSTM untuk mendeteksi Pneumonia dan COVID-19 dan Fachrel, J. *et al* sedangkan penelitian ini menambah dua penyakit lain yaitu Tuberculosis, dan lung nodules.

2.18 Dataset yang digunakan

Dataset yang diambil dari Kaggle dan digunakan pada penelitian ini adalah sebagai berikut:

1. JTIPTJ, (*Chest X-Ray Covid-19 dan Tuberculosis*)
<https://www.kaggle.com/datasets/jtiptj/chest-xray-pneumoniacovid19tuberculosis>
2. Paul Mooney, (*Chest X-Ray Images Pneumonia*)
<https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>

3. TAWSIFUR RAHMAN, Dr. Muhammad Chowdhury dan Amith Khandakar (*Tuberculosis*)
<https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset>
4. RADDAR (Nodules in Chest X-rays (JSRT))
<https://www.kaggle.com/datasets/raddar/nodules-in-chest-xrays-jsrt>
5. NATIONAL INSTITUTES OF HEALTH CHEST X-RAY DATASET and Chris Crawford (NIH Chest X-rays) <https://www.kaggle.com/datasets/nih-chest-xrays/data>

BAB III

ANALISIS DAN PERANCANGAN SISTEM

3.1 Data yang digunakan

Penelitian ini menggunakan dataset yang diperoleh dari Kaggle yang adalah sebuah platform yang membebaskan semua orang untuk mendapatkan dataset yang ingin digunakan untuk membangun model *Artificial Intelligence* (AI). Ekstensi citra X-Ray paru-paru yang digunakan adalah .jpg dan .jpeg yang terdiri dari lima kelas, yaitu Tuberculosis, Pneumonia, *Lung Nodules*, COVID-19 dan Normal dengan total jumlah data adalah 1250 di mana jumlah data di setiap kelas adalah sama yaitu 250. Contoh data citra X-Ray paru dapat ditampilkan oleh Gambar 3.1



Gambar 3.1 Citra Gambar Paru-Paru Tuberculosis (Rahman, *et al.* 2021)

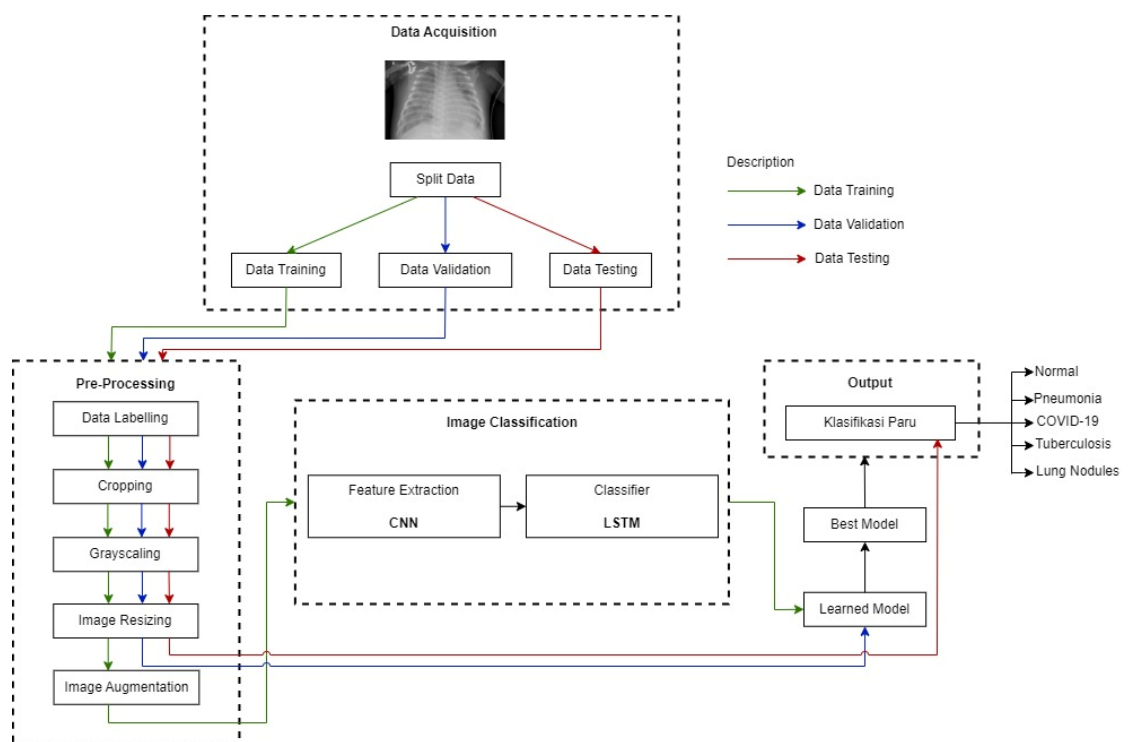
3.2 Analisis sistem

Penelitian ini terdiri dari beberapa tahapan yang diperlukan untuk mengklasifikasi penyakit pada paru, tahapan tersebut adalah *image acquisition*, *pre-processing* dan *classification*. *Output* yang dihasilkan dari tahapan-tahapan tersebut adalah berupa model *deep learning* yang digunakan untuk menganalisis citra paru-paru. Tahap pertama yaitu *image acquisition* merupakan tahapan untuk mendapatkan citra gambar yang akan digunakan. Data gambar paru-paru yang terdiri dari lima kelas

yang diperoleh dari Kaggle terdiri dari Tuberkulosis, Pneumonia, *Lung Nodules*, COVID-19 dan Normal.

Tahap selanjutnya adalah *image pre-processing*. Tahap ini terdiri dari beberapa proses seperti *labelling*, *resizing* dan *augmentation*. *Labelling* bertujuan untuk memberikan identitas citra paru-paru yang digunakan. *Resizing* merupakan proses untuk memperkecil ukuran citra dengan mengurangi *pixel* citra gambar tersebut.

Tahap terakhir yaitu tahap klasifikasi, merupakan proses untuk memberikan kelas pada citra gambar di mana *Convolutional Networks* dari CNN digunakan mendapatkan fitur dari citra dan LSTM untuk menentukan kelas citra. *Output* yang dihasilkan setelah tahap klasifikasi adalah citra gambar paru-paru yang teridentifikasi memiliki salah satu dari keempat penyakit atau tidak teridentifikasi memiliki penyakit paru. Tahapan keseluruhan dapat dilihat pada gambar arsitektur umum pada Gambar 3.2.



Gambar 3.2 Arsitektur Umum

Berikut ini merupakan penjelasan dari proses arsitektur umum diatas.

3.2.1 Image Acquisition

Tahap ini merupakan tahap untuk mengumpulkan data citra X-Ray paru-paru. Citra yang diambil adalah citra paru-paru normal, Pneumonia, COVID-19, Tuberkulosis dan *Lung Nodules*. Data diperoleh dari dataset yang berasal dari Kaggle. Citra yang digunakan pada penelitian ini memiliki format file .jpg dan .png yang setelah itu dibagi kedalam tiga bagian yakni data *training*, data *validation* dan data *testing* dengan porsi bagian 70% : 20% : 10% (875 data untuk data *training*, 250 data untuk data *validation* dan 125 data untuk data *testing*).

Tabel 3.1 Pembagian data Training, Validation dan Testing.

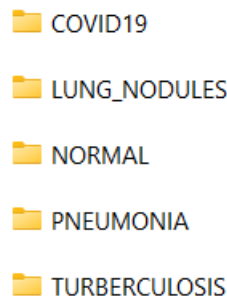
No.	Data	Data Training	Data Validation	Data Testing	Jumlah
1	Pneumonia	175	50	25	250
2	Tuberculosis	175	50	25	250
3	Lung Nodules	175	50	25	250
4	COVID-19	175	50	25	250
5	Normal	175	50	25	250
Total Data		875	250	125	1250

3.2.2 Pre-Processing

Tahap ini bertujuan untuk menaikkan kualitas citra supaya lebih mudah dikenali oleh sistem. *Pre-processing* yang diterapkan meliputi *labelling*, *cropping*, *grayscale*, *resizing* dan *image augmentation*.

a) Data Labelling

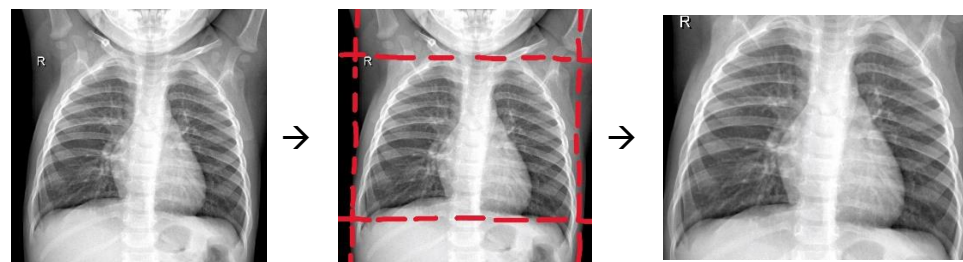
Label data yang diterapkan adalah Pneumonia, Tuberculosis, Lung Nodules, COVID-19 dan Normal.



Gambar 3.3 *Labelling*

b) *Cropping*

Cropping yang diterapkan pada citra di penelitian ini dimaksudkan untuk menghilangkan bagian citra yang tidak perlu agar pada saat feature extraction bagian citra tersebut tidak menjadi *feature* yang dapat mempengaruhi hasil klasifikasi. Contoh citra gambar sebelum dan sesudah *cropping* ditampilkan oleh Gambar 3.4.



Gambar 3.4 Citra paru-paru sebelum dan setelah cropping

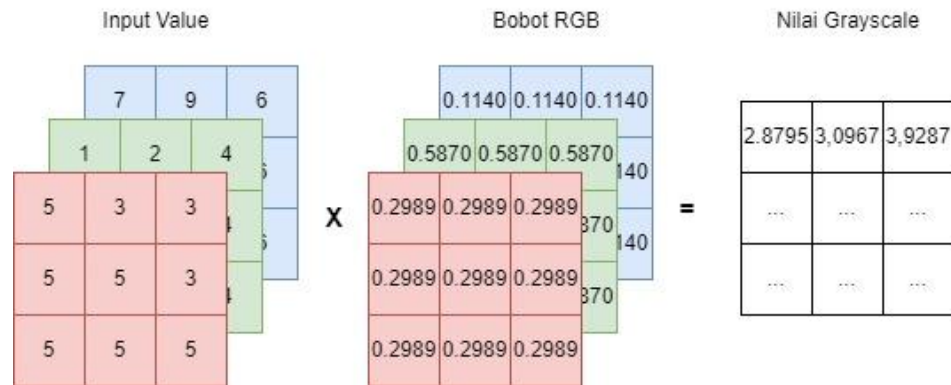
Dapat dilihat pada Gambar 3.4 bahwa *cropping* yang dilakukan untuk mendapatkan citra yang hanya pada bagian paru-paru sedangkan organ lain yang tidak diperlukan dihilangkan.

c) *Image Resizing*

Ukuran citra yang digunakan pada penelitian ini adalah 224 x 224 *pixel*. Dikarenakan ukuran citra yang didapat lebih besar dari 224 x 224 *pixel* maka diperlukan proses *resizing*.

d) *Grayscale*

Grayscale dilakukan karena tidak semua data yang digunakan pada penelitian ini adalah citra *grayscale*. Proses grayscale dapat dilihat pada Gambar 3.5.

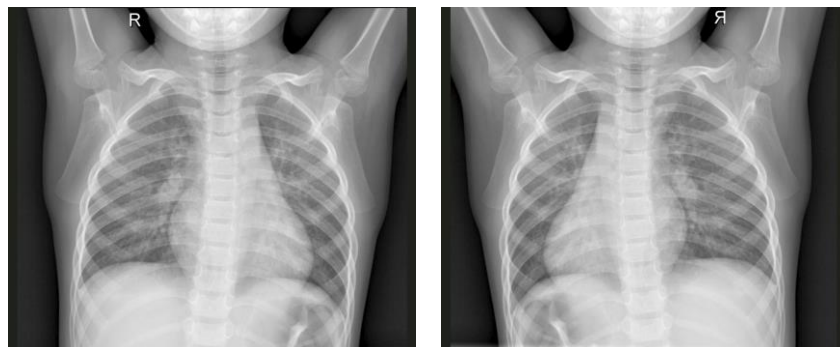


Gambar 3.5 Proses grayscale

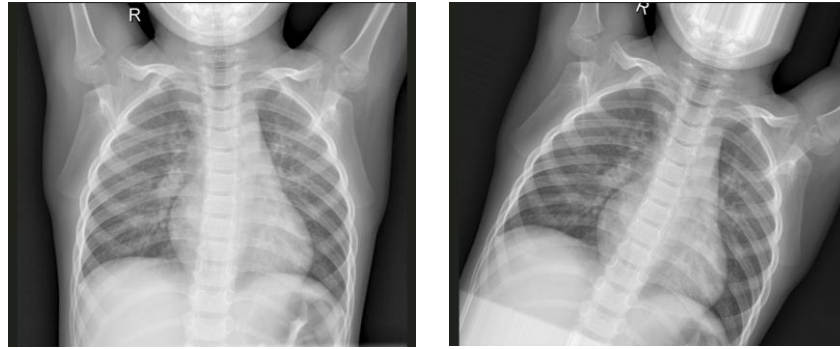
Perhitungan nilai grayscale dilakukan dengan mengalikan *value* piksel yang ada pada citra dengan bobot RGB secara *dot product* dan kemudian dijumlahkan.

e) *Penerapan Image Augmentation*

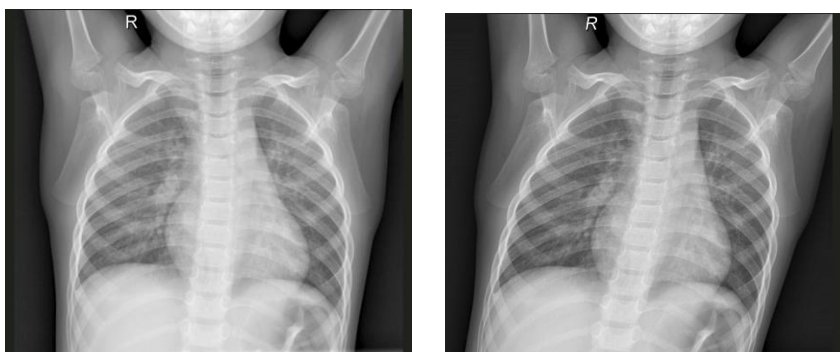
Augmentasi citra yang diaplikasikan pada penelitian ini adalah *flip*, *rotate*, *height shift range*, *width shift range*, *shear range* dan *zoom range*. *Flip* yang digunakan adalah horizontal *flip* (membalik gambar secara horizontal), *rotate* (memutar gambar) sebesar 30 derajat, *height shift range* (menggeser gambar secara vertical dengan rentang tertentu), *width shift range* (menggeser gambar secara horizontal dengan rentang tertentu) dan *shear range* (memiringkan gambar) yang dilakukan masing-masing sebesar 0.1 serta *zoom range* (memperbesar gambar) sebesar 25%.



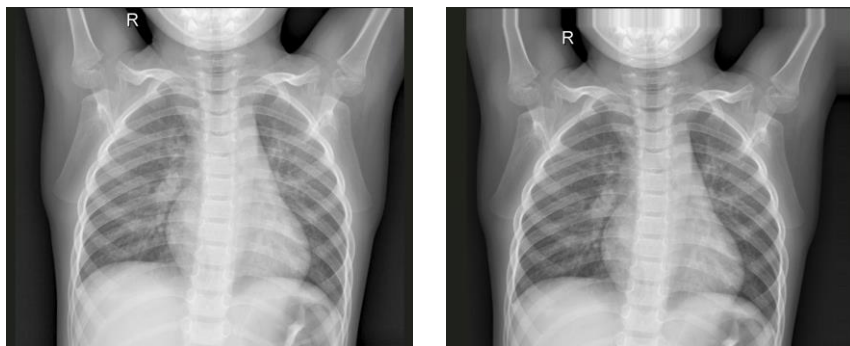
Gambar 3.6 Sebelum dan setelah Horizontal flip



Gambar 3.7 Sebelum dan setelah rotasi 30 derajat



Gambar 3.8 Sebelum dan setelah shear sebesar 0.1



Gambar 3.9 Sebelum dan setelah height shift range sebesar 0.1



Gambar 3.10 Sebelum dan setelah width whift range sebesar 0.1



Gambar 3.11 Sebelum dan setelah zoom sebesar 0.25

Berikut merupakan pseudocode augmentasi data yang dilakukan.

```
#buka file gambar
inputs ← Image.open('nama_file')
shape ← ('shape dari file, contoh: "1,244,244,3" ')

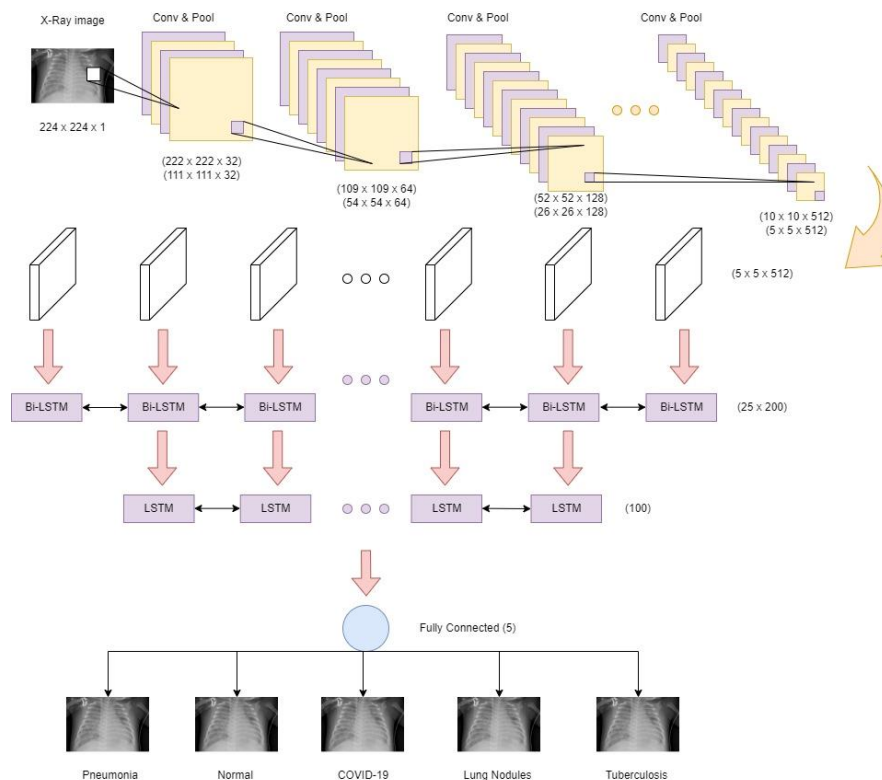
#buat variabel sementara untuk menampung value dari shape file
temp ← np.empty(shape=(shape[0], shape[1], shape[2], shape[3]))

for I in range(shape[0]):
    #ubah gambar ke bentuk array
    array_inputs = tf.keras.preprocessing.image.img_to_array(inputs[i])
    #masukkan augmentasi
    augmentasi = tf.keras.preprocessing.image."nama augmentasi".
    • random_zoom(array_inputs, 0.25)
    • random_rotation(array_inputs, 0.3)
    • random_shear(array_inputs, 0.1)
    • random_shift(array_inputs, (0.1, 0.1))
    • random_flip('horizontal')
    #masukkan ke variabel temp
    temp[i]= sheared

for i in range(shape[0]):
    # ubah kembali ke bentuk gambar dan kemudian ditampilkan
    tf.keras.preprocessing.image.array_to_img(temp[i]).show()
```


3.2.3 Image Classification

Proses klasifikasi pada penelitian ini menggunakan gabungan arsitektur CNN dengan LSTM di mana CNN akan mengekstrak fitur-fitur kompleks dari gambar dan LSTM digunakan sebagai pengklasifikasi. Detail arsitektur CNN-LSTM yang diterapkan penelitian ini terdapat pada Gambar 3.12.

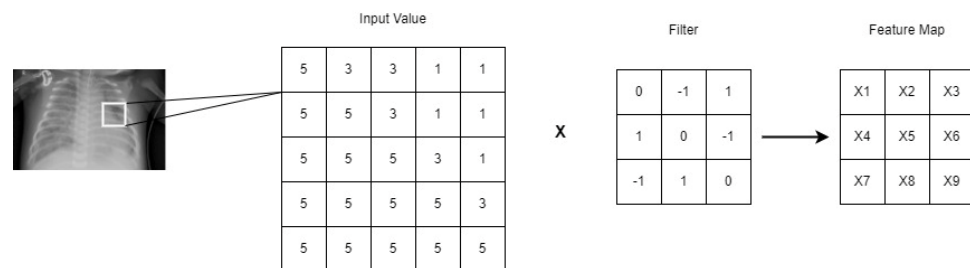


Gambar 3.12 Kombinasi Arsitektur CNN-LSTM

Layer yang digunakan berjumlah 13 layer: lima *convolution layer*, lima *pooling layer*, dua *bidirectional LSTM layer* dan satu *output layer*. Lima *convolutional layer* dikombinasikan dengan *pooling layer*. Ukuran *kernel* yang digunakan adalah 3x3 dengan menggunakan fungsi aktivasi ReLU. *Max pooling layer* dengan ukuran *kernel* 2x2 digunakan untuk mengurangi ukuran dari input image. Setelah proses *convolutional* terakhir maka *output shape* terakhir akan di ubah menjadi dua dimensi untuk bisa diproses di *bidirectional LSTM layer*. LSTM menerima nilai *output* dari *convolution layer* terakhir untuk proses klasifikasi.

a) *Convolutional Layer*

Convolutional layer digunakan untuk *feature extraction*. *Feature extraction* terhadap citra dilakukan dengan mengonvolusi matriks *N-dimensional* dari input citra dengan kernel konvolusi atau filter yang berupa *N-dimensional* matriks. Konvolusi dilakukan dengan melakukan perkalian *dot product* matriks citra dengan matriks *filter* yang kemudian nilai dari hasil perkalian tersebut dijumlahkan untuk menghasilkan *output* berupa matriks baru atau *feature map*.



Gambar 3.13 Proses konvolusi

Perhitungan konvolusi yang sesuai dengan yang ditampilkan oleh Gambar 3.13 dilakukan dengan mengalikan secara *dot product* dari matriks 3x3 dari filter terhadap matriks 5x5 dari citra dengan pergeseran atau *stride* sebesar 1x1. Contoh perhitungan dapat dilihat pada Gambar 3.14.

5	3	3	1	1
5	5	3	1	1
5	5	5	3	1
5	5	5	5	3
5	5	5	5	5

5	3	3	1	1
5	5	3	1	1
5	5	5	3	1
5	5	5	5	3
5	5	5	5	5

5	3	3	1	1
5	5	3	1	1
5	5	5	3	1
5	5	5	5	3
5	5	5	5	5

5	3	3	1	1
5	5	3	1	1
5	5	5	3	1
5	5	5	5	3
5	5	5	5	5

5	3	3	1	1
5	5	3	1	1
5	5	5	3	1
5	5	5	5	3
5	5	5	5	5

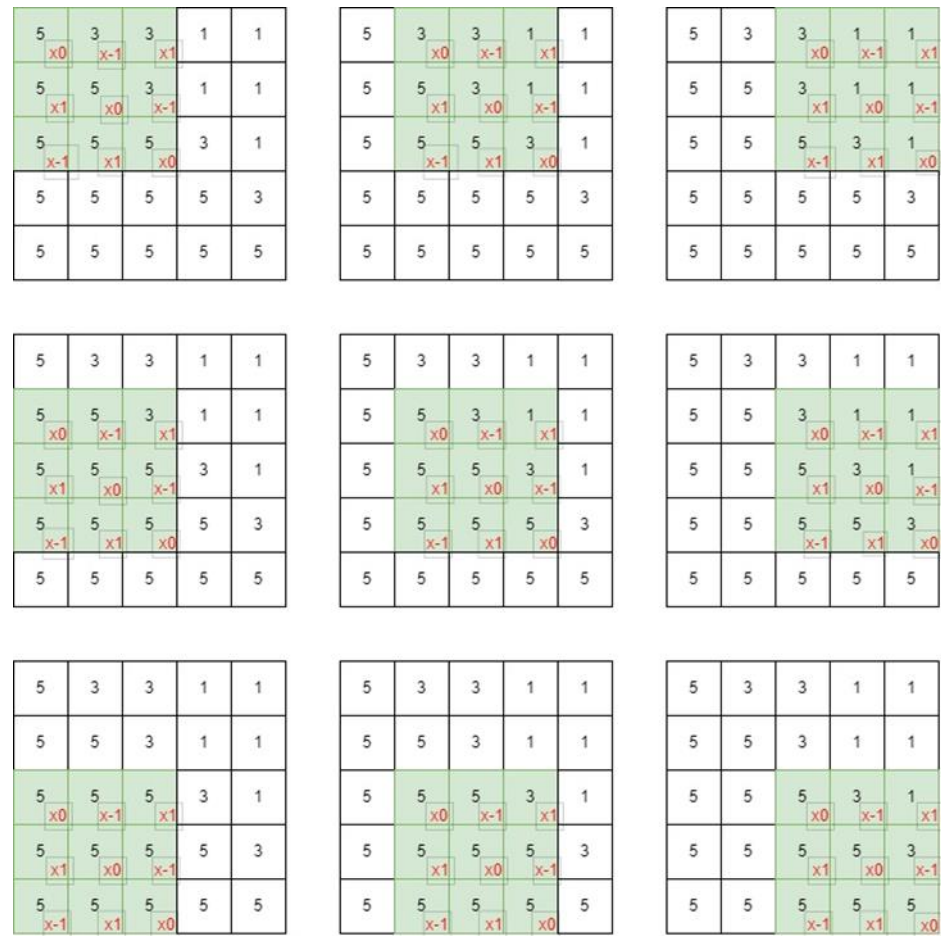
5	3	3	1	1
5	5	3	1	1
5	5	5	3	1
5	5	5	5	3
5	5	5	5	5

5	3	3	1	1
5	5	3	1	1
5	5	5	3	1
5	5	5	5	3
5	5	5	5	5

5	3	3	1	1
5	5	3	1	1
5	5	5	3	1
5	5	5	5	3
5	5	5	5	5

5	3	3	1	1
5	5	3	1	1
5	5	5	3	1
5	5	5	5	3
5	5	5	5	5

Gambar 3.14 Posisi perkalian matriks citra terhadap matriks filter



Gambar 3.15 Perkalian matriks citra dengan matriks kernel

Gambar 3.15 memperlihatkan perkalian *dot product* untuk menghasilkan *feature map*. Perhitungan *convolutional layer* pada penelitian ini dapat dilihat pada perhitungan dibawah:

$$X1 = (5 \times 0) + (3 \times -1) + (3 \times 1) + (5 \times 1) + (5 \times 0) + (3 \times -1) + (5 \times -1) + (5 \times 1) + (5 \times 0) = 0 + -3 + 3 + 5 + 0 + -3 + -5 + 5 + 0 = 2$$

...

$$X9 = (5 \times 0) + (3 \times -1) + (1 \times 1) + (5 \times 1) + (5 \times 0) + (3 \times -1) + (5 \times -1) + (5 \times 1) + (5 \times 0) = 0 + -3 + 1 + 5 + 0 + -3 + -5 + 5 + 0 = 0$$

Berdasarkan perhitungan diatas didapat *feature map* yang ditampilkan oleh Gambar 3.16.

2	2	0
-2	0	4
0	-2	0

Gambar 3.16 *Feature map*

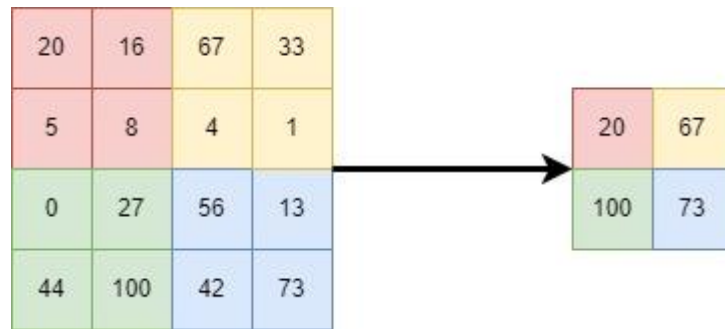
Fungsi aktivasi ReLU akan membuat matriks hanya akan berisi nilai positif di mana apabila terdapat nilai negatif maka akan diubah menjadi 0, sehingga matriks *feature map* akan berubah menjadi seperti yang ditampilkan oleh Gambar 3.17.

2	2	0
0	0	4
0	0	0

Gambar 3.17 *Feature map setelah ReLU*

b) Max Pooling

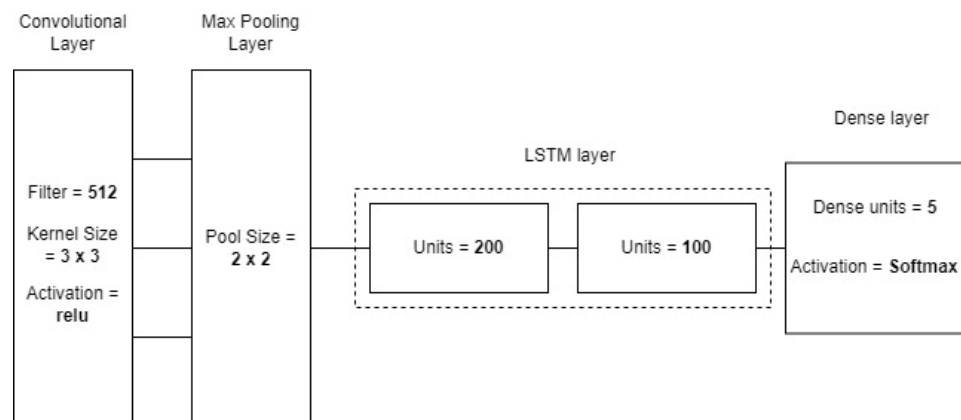
Pooling layer berfungsi untuk mendapatkan ringkasan statistik dari input nilai citra dan kemudian menurunkan dimensi spasial dari *feature map* dengan harapan bahwa tidak ada informasi esensial yang hilang. *Max pooling layer* mengambil nilai maksimal di antara semua nilai dari bagian matriks $N \times N$. Kombinasi *Max pooling* yang paling sering digunakan adalah *stride* 2 dengan ukuran kernel 2. Bentuk tersebut akan membagi *feature map* secara spasial ke dalam kotak dari blok persegi atau kubik dengan sisi 2 dan mengambil nilai maksimal dari blok-blok tersebut untuk setiap fitur input. Proses *max pooling* dapat dilihat pada Gambar 3.18.



Gambar 3.18 Max Pooling dengan *stride* 2 dan kernel 2x2

c) LSTM layer

LSTM layer akan menerima nilai *output* dari *convolutional layer* untuk menjadi nilai *input* pada LSTM dan kemudian proses diteruskan ke *dense layer* untuk proses klasifikasi. Penerapan LSTM layer dapat dilihat pada Gambar 3.19.



Gambar 3.19 Penerapan LSTM layer

3.2.4 Learned Model

Proses klasifikasi akan menghasilkan *Learned model* yang berisi pengetahuan yang didapat dari melatih dataset. *Learned model* berisi bobot yang didapat selama proses pelatihan data. *Learned model* terbaik akan dipilih sebagai *best model* yang digunakan pada sistem.

3.2.5 Proses Training

Proses pelatihan data adalah tahapan bagi algoritma *deep learning* untuk mempelajari dataset yang diberikan agar dapat menghasilkan model yang dapat

melakukan tugas yang diberikan, sebagai contoh dalam penelitian ini CNN-LSTM diberikan tugas untuk dapat mengenali dan membedakan gambar paru yang tidak terkena penyakit dan yang terkena empat jenis penyakit yang berbeda. Berikut merupakan kode untuk melakukan augmentasi data.

```

Import ImageDataGenerator
# Augmentasi data training
set train_datagen to ImageDataGenerator using arguments
    1. rescale  $\leftarrow$  1./255,
    2. horizontal_flip  $\leftarrow$  True,
    3. rotation_range  $\leftarrow$  30,
    4. width_shift_range  $\leftarrow$  0.1,
    5. height_shift_range  $\leftarrow$  0.1,
    6. shear_range  $\leftarrow$  0.1,
    7. zoom_range  $\leftarrow$  0.25.
# Data validation
Set validation_datagen to ImageDataGenerator using arguments
    1. rescale  $\leftarrow$  1./255
# Data test
Set test_datagen to ImageDataGenerator using arguments
    1. rescale  $\leftarrow$  1./255

# Konfigurasi batch size, size gambar, grayscale
Set {train_generator, validation_generator, test_generator} to {train_datagen,
validation_datagen, test_datagen} calls flow_from_directory using
arguments
    1. batch_size  $\leftarrow$  15,
    2. target_size  $\leftarrow$  (224,224),
    3. color_mode  $\leftarrow$  'grayscale'

```

Berikut merupakan kode untuk training data.

```

Import tensorflow library as tf
Import Adam from tensorflow.keras.optimizers
# Membuat model
Set model to tf calls keras calls models calls sequential contains
    1. tf calls keras calls layers calls Conv2D set
        1. filters ← 32,
        2. kernel_size ← (3,3),
        3. activation ← 'relu',
        4. input_shape ← (224, 224, 1)
    2. tf calls keras calls layers calls MaxPooling2D set pool_size ← (2, 2)
    3. tf calls keras calls layers calls Conv2D set
        1. filters ← 32,
        2. kernel_size ← (3,3),
        3. activation ← 'relu'
    4. tf calls keras calls layers calls MaxPooling2D set pool_size ← (2,2)
    5. tf calls keras calls layers calls Conv2D set
        1. filters ← 128,
        2. kernel_size ← (3,3),
        3. activation ← 'relu'
    6. tf calls keras calls layers calls MaxPooling2D set pool_size ← (2,2),
    7. tf calls keras calls layers calls Conv2D set
        1. filters ← 256,
        2. kernel_size ← (3,3),
        3. activation ← 'relu'
    8. tf calls keras calls layers calls MaxPooling2D set pool_size ← (2,2),
    9. tf calls keras calls layers calls Conv2D set
        1. filters ← 512,
        2. kernel_size ← (3,3),
        3. activation ← 'relu'
    10. tf calls keras calls layers calls MaxPooling2D set pool_size ←
        (2,2),
    11. tf calls keras calls layers calls Reshape set to (25, 512)

```



```

12. tf calls keras calls layers calls Bidirectional get tf calls keras calls
    layers calls LSTM set
    1. units ← 200,
    2. return sequences ← True
13. tf calls keras calls layers calls Bidirectional get tf calls keras calls
    layers calls LSTM set units ← 100
14. tf calls keras calls layers calls Dense set
    1. Dense units ← 5,
    2. activation = 'softmax'

# Compile model
Set model calls compile using arguments
    1. loss ← categorical_crossentropy,
    2. optimizer ← Adam set learning_rate ← 0.00015,
    3. metrics ← 'accuracy'

# Latih model
Set history to model calls fit using arguments
    1. train_generator,
    2. epochs ← 250,
    3. validation_data ← validation_generator

```

3.2.6 Proses *Validation*

Data validasi ditujukan sebagai indikator pembanding apakah model yang dilatih mengalami *overfitting* maupun *underfitting*.

3.2.7 Proses *Testing*

Proses pengetesan pada model merupakan proses untuk menguji apakah model yang dihasilkan mampu mempelajari pola informasi yang ada pada data training. Proses ini menunjukkan seberapa tinggi kemampuan model dalam mengklasifikasi citra gambar yang diinput. Kemampuan model dapat dilihat dari nilai akurasi yang dihasilkan pada proses data testing.

3.2.8 Model Evaluation

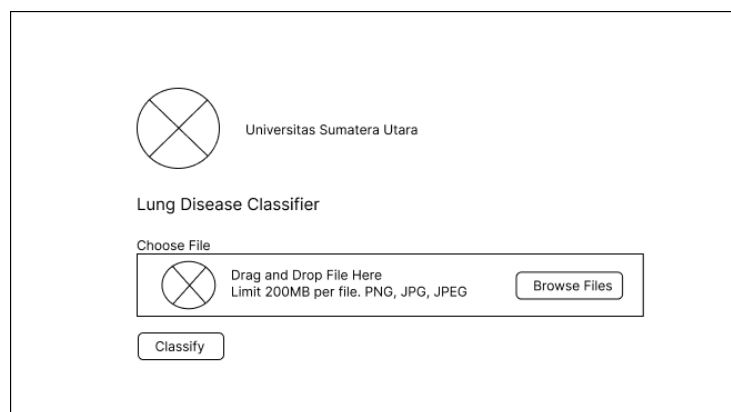
Evaluasi terhadap model diperlukan untuk melihat kinerja model dalam mempelajari pola informasi dari data training. Peneliti dalam hal ini menggunakan confusion matrix sebagai metrik evaluasi sekaligus untuk melihat *precision*, *recall* dan *F1-score* dari model.

3.3 Perancangan Sistem

Tahap selanjutnya adalah pembuatan sistem yang akan digunakan oleh pengguna untuk mendeteksi jenis penyakit yang terdapat pada paru-paru. Sistem berbasis web adalah sistem yang dikembangkan dalam penelitian ini. Input yang diterima pada sistem web ini berupa gambar dengan format PNG, JPG dan JPEG dengan limit maksimal 200MB sedangkan *output* yang diterima oleh pengguna berupa gambar, label kelas penyakitnya serta nilai *confidence*-nya.

3.3.1 Desain Antarmuka

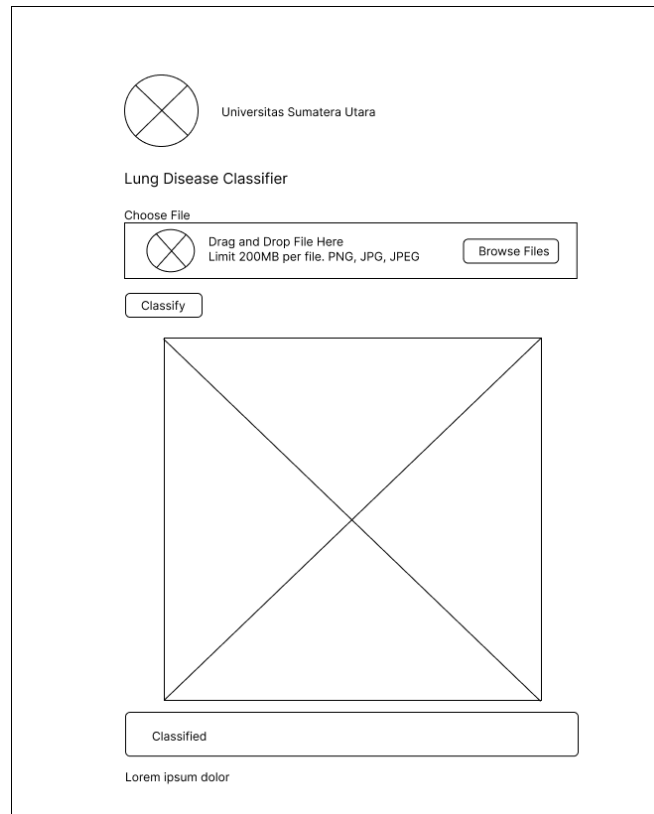
Rancangan antarmuka sistem ditampilkan oleh Gambar 3.20 dan Gambar 3.21 beserta dengan diagram aktivitas sistem aplikasi yang terdapat pada gambar 3.22.



Gambar 3.20 Rancangan halaman pilih gambar

Pada Gambar 3.20 terdapat logo dari Universitas Sumatera Utara dan juga tulisan Universitas Sumatera Utara disampingnya. Di bawah logo terdapat tulisan “Lung Disease Classifier” yang merupakan judul dari aplikasi ini. Di bawah judul terdapat *box* yang berisi fitur untuk memasukkan gambar untuk di

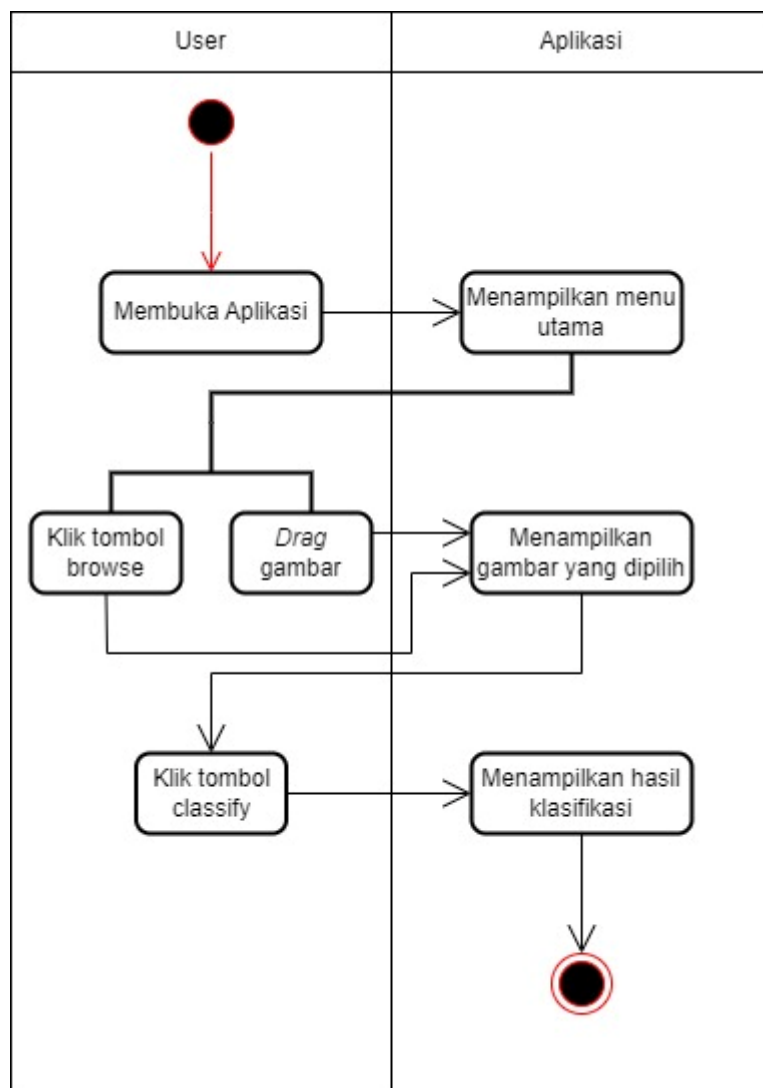
klasifikasi yang dapat dilakukan dengan mengklik tombol ‘browse files’ atau dengan *mendrag* gambar kedalam box. Di bawah *box* tersebut terdapat tombol ‘classify’ untuk memulai proses klasifikasi.



The wireframe shows a web interface for a 'Lung Disease Classifier'. At the top left is a circular logo with a cross inside, followed by the text 'Universitas Sumatera Utara'. Below this is the title 'Lung Disease Classifier'. A 'Choose File' section contains a drag-and-drop area with a circular icon and the text 'Drag and Drop File Here' and 'Limit 200MB per file. PNG, JPG, JPEG'. To the right of this area is a 'Browse Files' button. Below the file selection area is a 'Classify' button. Underneath the button is a large square box with a diagonal cross, intended for displaying the image. At the bottom, there is a 'Classified' label and a text area containing the placeholder text 'Lorem ipsum dolor'.

Gambar 3.21 Rancangan halaman hasil klasifikasi gambar

Setelah gambar berhasil dipilih maka gambar akan muncul di bawah tombol ‘classify’ tersebut (dapat dilihat pada Gambar 3.21) dan apabila tombol ‘classify’ diklik maka akan muncul tulisan ‘classified’ beserta hasil klasifikasinya di bawah tulisan ‘classified’ tersebut (pada gambar 3.21 diwakilkan oleh tulisan *lorem ipsum dolor*).



Gambar 3.22 Activity diagram

BAB IV

Implementasi dan Pengujian

4.1 Implementasi Sistem

Implementasi sistem klasifikasi citra paru-paru menggunakan metode CNN-LSTM adalah sebagai berikut.

4.1.1 Perangkat Keras dan Perangkat Lunak

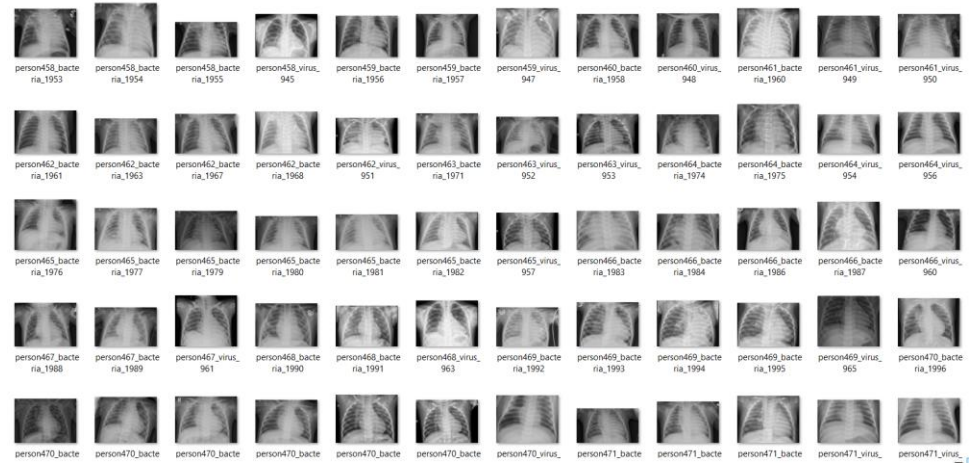
Berikut detail perangkat keras dan perangkat lunak yang dipakai untuk mengembangkan sistem ini.

1. Laptop Lenovo Legion 5 15IMH05
2. Processor Intel^R CoreTM i7-10750H 2.60GHz
3. GPU Nvidia GeForce GTX 1650Ti 4GB
4. Nvidia CUDA 11.2
5. Storage SSD 512GB
6. RAM 16GB
7. Sistem operasi Windows 11 Home Single Language
8. Tensorflow 2.9.0
9. Python 3.9.2
10. Visual Studio Code
11. Streamlit 1.31.1

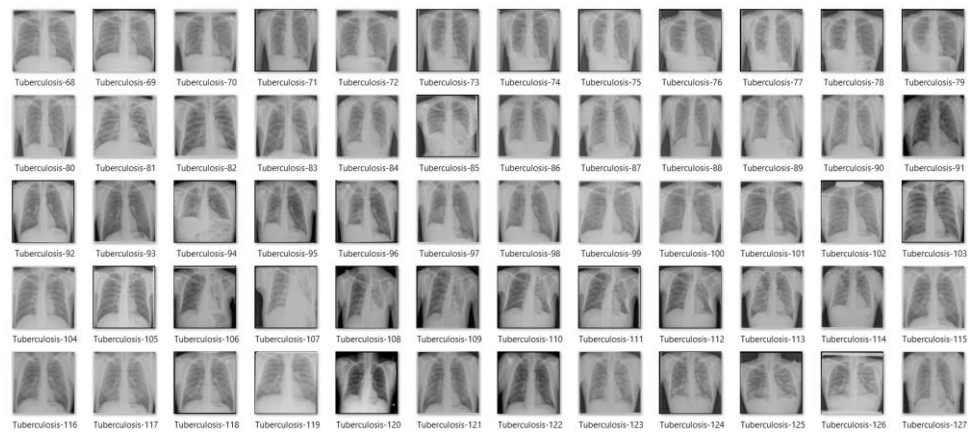
4.2 Implementasi Data

Data yang diimplementasi merupakan data dari Kaggle yang sudah dalam bentuk dataset. Dataset citra X-Ray paru-paru tersebut terdiri dari lima kelas yakni Pneumonia, Tuberculosis, Normal, COVID-19 dan *Lung Nodules*. Gambar yang digunakan berekstensi .jpg, .jpeg dan .png dengan total keseluruhan adalah sebanyak 1250 data dengan masing-masing kelas memiliki

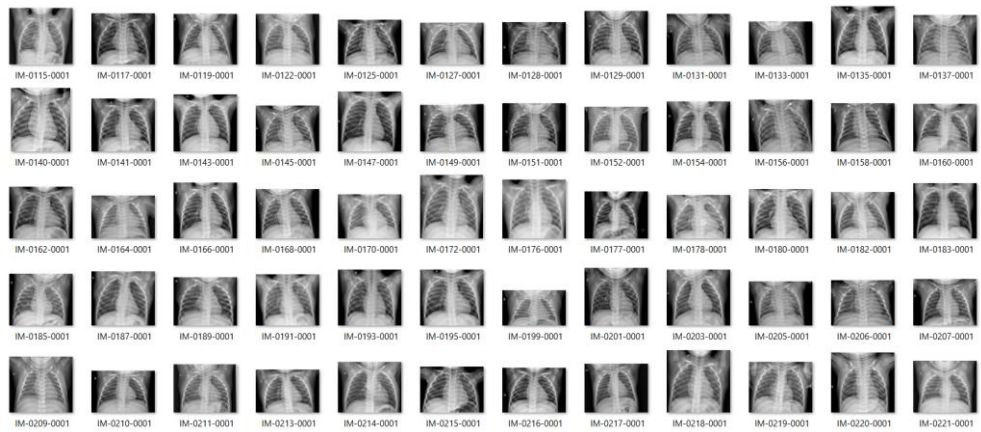
jumlah 250 data. Contoh data pada setiap kelas ditampilkan oleh Gambar 4.1, Gambar 4.2, Gambar 4.3, Gambar 4.4 dan Gambar 4.5.



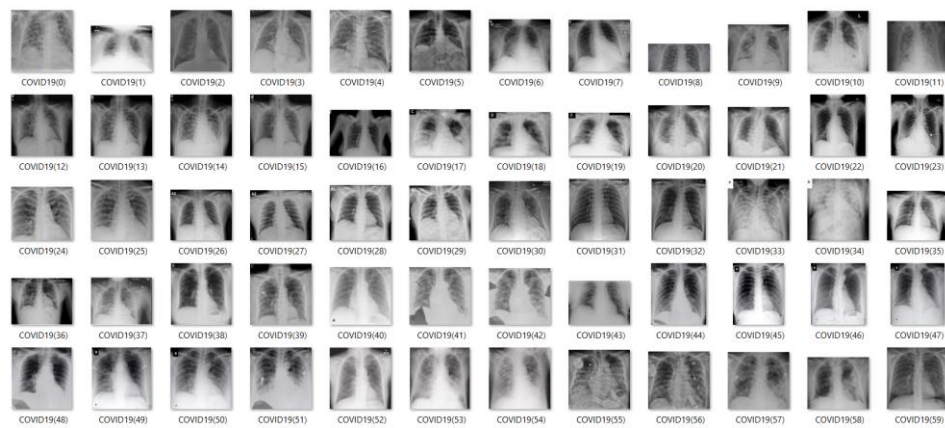
Gambar 4.1 Data Citra Pneumonia



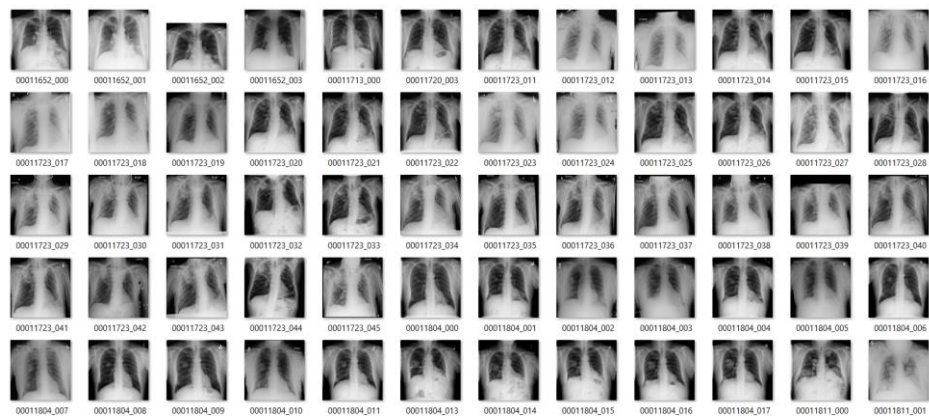
Gambar 4.2 Data Citra Tuberculosis



Gambar 4.3 Data Citra Paru-Paru Normal



Gambar 4.4 Data Citra COVID-19



Gambar 4.5 Data Citra *Lung Nodules*

4.2.1 Implementasi Augmentasi Data

Pada tahap ini data citra paru-paru yang sudah dikelompokkan sesuai dengan kelasnya akan diaugmentasi sebanyak enam kali. Data yang diaugmentasi hanya

data yang termasuk kedalam data training. Sesuai dengan data yang ditampilkan oleh Tabel 3.1, ada sebanyak 875 data training yang dibagi dalam 5 kelas yang akan diaugmentasi dan jika dikalikan dengan banyaknya proses augmentasi maka jumlah data training akan bertambah menjadi 5250. Perubahan dataset setelah diaugmentasi ditampilkan oleh Tabel 4.1.

Tabel 4.1 Pembagian data setelah proses augmentasi

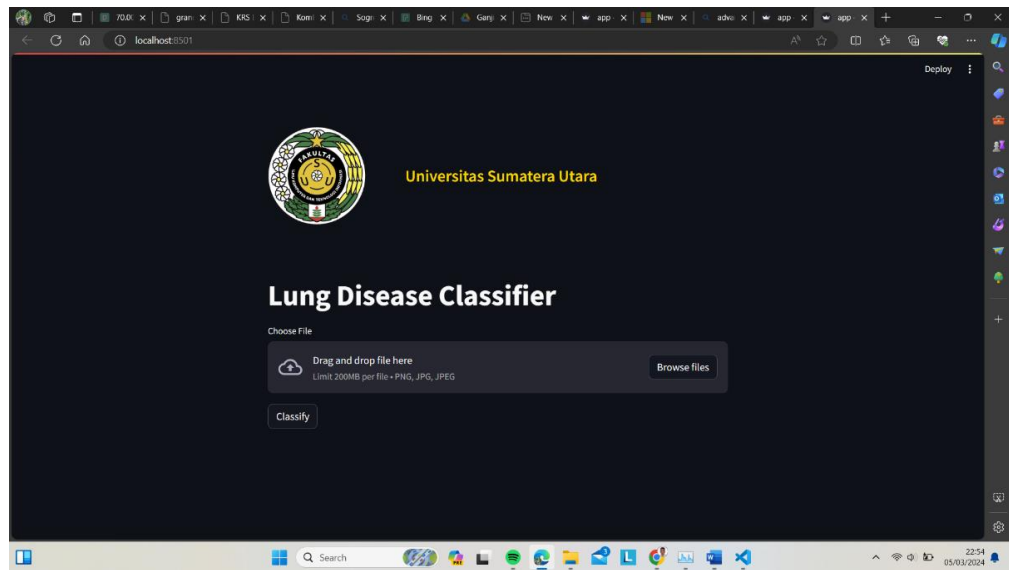
No.	Data	Data Training	Data Validation	Data Testing	Jumlah
1	Pneumonia	1050	50	25	250
2	Tuberculosis	1050	50	25	250
3	Lung Nodules	1050	50	25	250
4	COVID-19	1050	50	25	250
5	Normal	1050	50	25	250
Total Data		5250	250	125	1250

4.3 Implementasi Desain Sistem

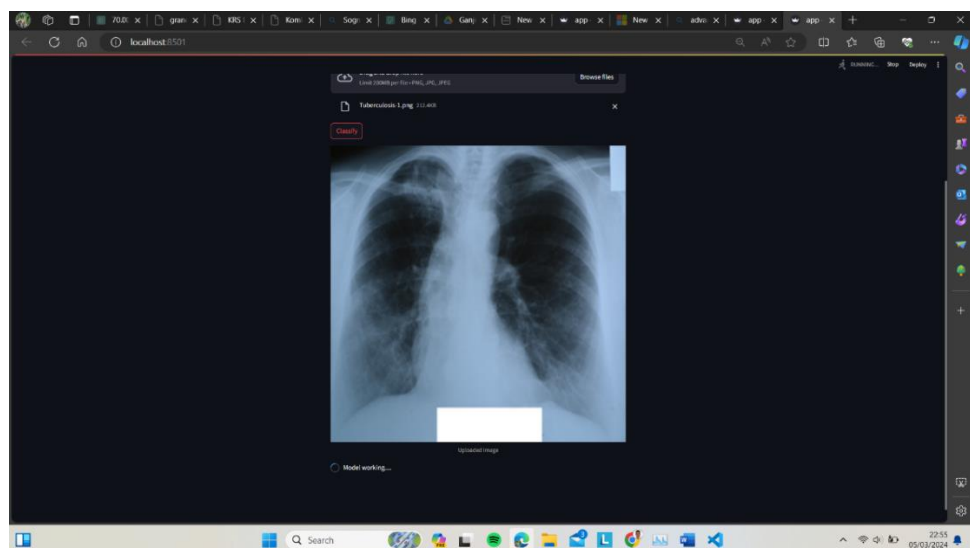
Tampilan antarmuka sistem memiliki satu halaman yaitu halaman utama.

4.3.1 Tampilan Utama

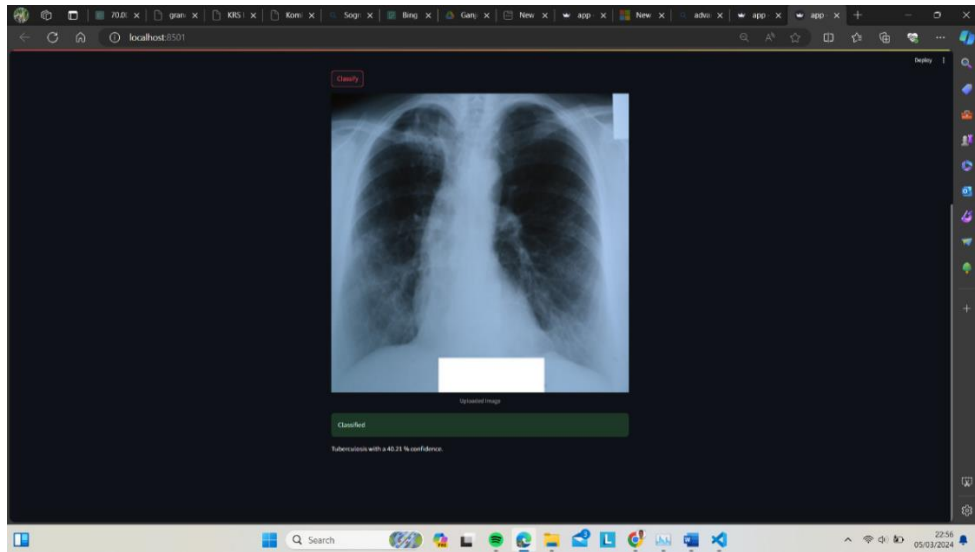
Tampilan utama berisi halaman untuk melakukan klasifikasi. Cara yang dapat dilakukan untuk mengklasifikasi citra adalah dengan men-*drag* file gambar ataupun dengan mengklik tombol “Browse files” dan setelah tidak ada error yang muncul maka selanjutnya adalah dengan mengklik tombol “classify”. Hasil klasifikasi akan muncul dibawah gambar X-Ray yang dipilih sebelumnya. Tampilan Utama sistem ditampilkan oleh Gambar 4.6, Gambar 4.7 dan Gambar 4.8.



Gambar 4.6 Tampilan Utama sebelum memilih gambar.



Gambar 4.7 Tampilan Utama setelah mengklik tombol classify.



Gambar 4.8 Tampilan Utama setelah proses klasifikasi.

4.4 Prosedur Operasional

Untuk menjalankan aplikasi ini pengguna terlebih dahulu menginstall Streamlit kedalam *environment* kemudian untuk menjalankan aplikasinya dilakukan dengan menjalankan perintah “`streamlit run app.py`” pada terminal. Setelah me-*run* streamlit maka streamlit akan terbuka langsung di *default browser* yang ditetapkan pada *device*.

Pada aplikasi ini terdapat tombol *browse files* untuk membuka *file explore* dan memilih gambar yang akan di klasifikasi atau dapat juga dilakukan dengan *drag and drop* di tempat yang disediakan, kemudian pengguna mengklik tombol *classify*. Setelah proses selesai maka sistem akan mengeluarkan output hasil klasifikasi beserta dengan nilai *confidence* nya.

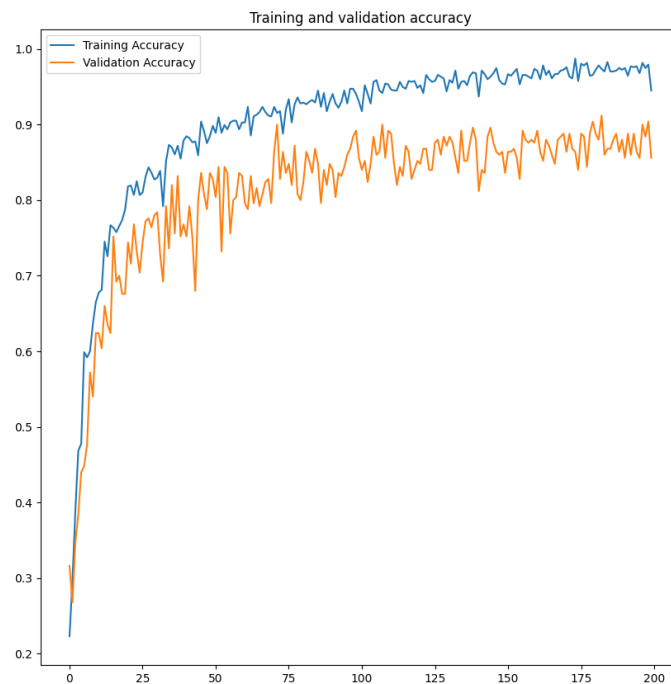
4.5 Training model

Sebelum melakukan pelatihan pada model, terdapat beberapa parameter yang dapat mempengaruhi atau mengontrol proses pembelajaran pada model, parameter ini disebut *Hyperparameter*. *Hyperparameter* tersebut dapat diubah-ubah untuk mendapatkan nilai yang terbaik. Hyperparameter yang diubah pada penelitian ini adalah *batch size*, *learning rate* dan jumlah filter di *Convolution* pertama. *Hyperparameter-hyperparameter* yang diterapkan ditampilkan oleh Tabel 4.2.

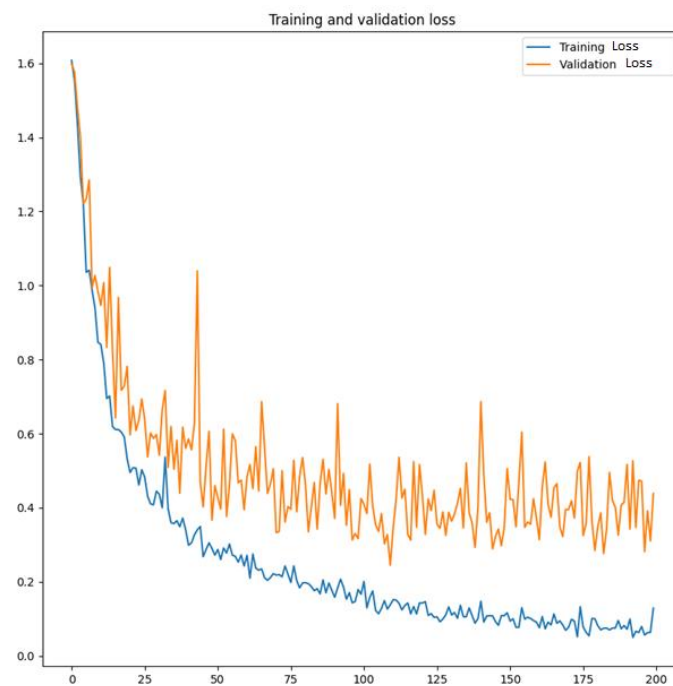
Tabel 4.2 Hasil training data dengan berbagai kombinasi hyperparameter

Percobaan	Batch Size	Learning Rate	Filter <i>Convolution</i> Pertama	Test Accuracy (-0.02 error)
Percobaan 1	10	0.0001	32	0.94 (0.92)
Percobaan 2	10	0.00015	32	0.91 (0.89)
Percobaan 3	10	0.0001	64	0.91 (0.89)
Percobaan 4	10	0.00015	64	0.89 (0.87)
Percobaan 5	10	0.00025	64	0.91 (0.89)
Percobaan 6	15	0.00015	64	0.95 (0.93)
Percobaan 7	15	0.00025	64	0.93 (0.91)
Percobaan 8	25	0.0001	32	0.88 (0.86)

Setelah melakukan pelatihan menggunakan hyperparameter yang sesuai dengan yang ditunjukkan oleh tabel 4.2, didapatkan model dengan hasil terbaik yaitu pada percobaan keenam yang menggunakan 15 batch, 0.00015 learning rate dan 64 filter pada convolution pertama dengan *epoch* setiap percobaan adalah 200. Hasil akurasi dan loss dari percobaan 6 ditampilkan oleh gambar 4.9 dan 4.10.

**Gambar 4.9** Akurasi *Training* dan *Validation* pada model

Berdasarkan grafik yang ditampilkan oleh Gambar 4.9, terlihat bahwa akurasi data latih (ditunjukkan dengan garis warna biru) dan data validasi (ditunjukkan dengan garis warna jingga) meningkat di setiap *epoch* nya, namun peningkatan akurasi pada data validasi tidak terlihat signifikan dibandingkan dengan yang terlihat pada data latih. Penambahan jumlah *epoch* memiliki kemungkinan untuk mendapatkan nilai akurasi yang lebih tinggi.



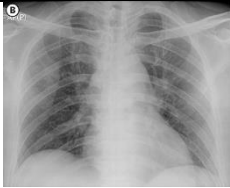


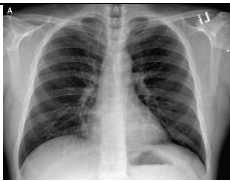


Gambar 4.10 Loss *Training* dan *Validation* pada model

Berdasarkan grafik loss yang ditunjukkan oleh Gambar 4.10, nilai *loss* dari data latih terus melandai turun sedangkan pada data validasi pada *epoch* pertama sampai *epoch* 100 melandai turun namun pada *epoch* diatas 100 nilai *loss* cenderung stagnan hingga *epoch* ke 200.

4.6 Pengujian Sistem

Ada sebanyak 125 data *testing* yang diuji oleh sistem. Sistem dapat mengidentifikasi citra paru-paru kedalam lima kelas klasifikasi. Hasil pengujian sistem ditampilkan oleh pada Tabel 4.3.

Tabel 4.3 Hasil Pengujian Sistem

No.	Citra	Aktual	Sistem	Status
1		COVID-19	COVID-19	Benar
2		COVID-19	COVID-19	Benar
3		COVID-19	COVID-19	Benar
4		COVID-19	Lung Nodules	Salah
...				
124		Tuberculosis	Tuberculosis	Benar
125		Tuberculosis	Tuberculosis	Benar

Berdasarkan hasil pengujian dari 125 data yang dites terdapat sebanyak delapan data yang hasil klasifikasinya meleset dari label sebenarnya.

Tabel 4.4 Confusion Matrix

	COVID-19	Lung Nodules	Normal	Pneumonia	Tuberculosis
COVID-19	24	1	0	0	0
Lung Nodules	0	24	0	0	1
Normal	1	0	22	2	0
Pneumonia	0	0	0	24	1
Tuberculosis	2	0	0	0	23

Pada hasil yang ditampilkan oleh tabel *confusion matrix*, kesalahan klasifikasi citra paling banyak terdapat pada kelas Normal. Faktor yang menjadi penyebab terjadinya kesalahan klasifikasi tersebut adalah pada citra data normal terdapat kemiripan terhadap citra gambar Pneumonia yang menyebabkan sistem sulit membedakan citra tersebut. Nilai TP, FP dan FN penelitian ini ditampilkan oleh Tabel 4.5.

Tabel 4.5 Nilai TP, FP dan FN

	TP	FP	FN
COVID-19	24	3	1
Lung Nodules	24	1	1
Normal	22	0	3
Pneumonia	24	2	1
Tuberculosis	23	2	2

Berdasarkan data yang ditampilkan oleh tabel 4.5, diperoleh skor precision, recall dan *F1-score* yang ditampilkan oleh tabel 4.6.

Tabel 4.6 Nilai Precision, Recall dan F1-Score

	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
COVID-19	0.89	0.96	0.92
Lung Nodules	0.96	0.96	0.96
Normal	1	0.88	0.94
Pneumonia	0.92	0.96	0.94
Tuberculosis	0.92	0.92	0.92
Average	0.938	0.936	0.936

Berdasarkan hasil pengujian yang dilakukan menggunakan sistem, akurasi klasifikasi penyakit paru dengan menggunakan metode CNN-LSTM dapat dihitung dengan rumus berikut:















$$\text{Akurasi} = \frac{24+24+22+24+23}{125} * 100\%$$



$$\text{Akurasi} = \frac{117}{125} * 100\%$$

$$\text{Akurasi} = 93,6 \%$$

Berdasarkan hasil pengujian yang didapat, tingkat akurasi yang didapat pada penelitian klasifikasi penyakit paru menggunakan algoritma CNN-LSTM adalah sebesar 93,6 %. Terjadinya kesalahan klasifikasi yang dihasilkan oleh sistem dipengaruhi oleh kontras citra. Hal ini dibuktikan oleh penulis dengan mengetes sistem dengan data citra yang sudah diimplementasikan dengan CLAHE yang merupakan salah satu teknik untuk memodifikasi kontras. Tabel 4.7 menunjukkan hasil setelah dan sebelum implementasi CLAHE dengan *clip_limit* = 1.5 terhadap citra yang hasil klasifikasinya salah.

Tabel 4.7 Hasil klasifikasi setelah implementasi CLAHE

No .	Citra sebelum CLAHE	Citra setelah CLAHE	Aktual	Sistem	Perbandingan dengan tes sebelumnya
1			COVID-19	Lung Nodules	Tetap
2			Lung Nodules	Lung Nodules	Berubah
3			Normal	Pneumonia	Tetap
4			Normal	COVID-19	Tetap
5			Normal	Normal	Berubah
6			Pneumonia	Pneumonia	Berubah
7			Tuberculosis	COVID-19	Tetap

8			Tuberculosis	COVID-19	Tetap
---	---	---	--------------	----------	-------

Berdasarkan hasil yang ditunjukkan oleh Tabel 4.7 terdapat 3 gambar yang hasil klasifikasinya berubah dan benar sehingga dapat disimpulkan bahwa modifikasi kontras dengan nilai tertentu mempengaruhi hasil klasifikasi sistem.

BAB V

Kesimpulan dan Saran

5.1 Kesimpulan

Kesimpulan dari penelitian dan pengujian sistem klasifikasi penyakit paru yang menerapkan metode CNN-LSTM adalah:

1. Metode CNN-LSTM mampu mengklasifikasi empat penyakit paru yaitu Pneumonia, Lung Nodules, COVID-19 dan Tuberculosis dengan akurasi sebesar 93,6%.
2. Kegagalan sistem dalam mengklasifikasi kelima kondisi paru dipengaruhi oleh kontras citra.

5.2 Saran

Berikut saran yang dapat dipertimbangkan untuk penelitian selanjutnya:

1. Membuat agar sistem dapat memberitahukan tingkat keparahan dari penyakit yang diklasifikasi.
2. Menggunakan metode yang lain yang dapat menghasilkan akurasi yang lebih baik.
3. Menambahkan metode yang dapat memengaruhi kontras dalam melakukan klasifikasi citra penyakit paru berdasarkan citra X-Ray.

Daftar Pustaka

- Ahmed, M.S., Rahman, A., AlGhamdi, F., AlDakheel, S., Hakami, H., AlJumah, A., AlIbrahim, Z., Youldash, M., Alam Khan, M.A., & Basheer Ahmed, M.I. 2023. Joint Diagnosis of Pneumonia, COVID-19, and Tuberculosis from Chest X-ray Images: A Deep Learning Approach. *Diagnostics*. 13, 2562. <https://doi.org/10.3390/diagnostics13152562>
- Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Hasan, M., Van Essen, B.C., Awwal, A.A.S., & Asari, V.K. 2019. A State-of-the-Art Survey on Deep Learning Theory and Architectures. *Electronics*. 8, 292. <https://doi.org/10.3390/electronics8030292>
- Adigun, R., Singh, R.. Tuberculosis. [Updated 2023 Jul 11]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK441916/>
- Bezdan T. & Džakula N. B. 2019. Convolutional Neural Network Layers and Architectures. International Scientific Conference on Information Technology and Data Related Research. pp. 445–451. <https://doi.org/10.15308/Sinteza-2019-445-451>.
- Bloom, B.R., Atun, R., Cohen, T., Dye, C., Fraser, H., Gomez, G.B., Knight G, Murray M, Nardell E, Rubin E, Salomon J, Vassall A, Volchenkov G, White R, Wilson D, & Yadav P. 2017. Tuberculosis. In: Holmes KK, Bertozzi S, Bloom BR, Jha P, editors. *Major Infectious Diseases*. 3rd ed. Washington (DC): The International Bank for Reconstruction and Development / The World Bank; 2017 Nov 3. Chapter 11. PMID: 30212088.
- Cascella, M., Rajnik, M., Aleem, A., et al. Features, Evaluation, and Treatment of Coronavirus (COVID-19) [Updated 2023 Aug 18]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK554776/>
- CT Scan Versus MRI Versus X-Ray: What Type of Imaging Do I Need?. 2023. www.hopkinsmedicine.org (Diakses tanggal 16 September 2023)
- Diar, R.M., Fu'Adah, R. Y. N. & Usman K. 2022. Klasifikasi Penyakit Paru-Paru Berbasis Pengolahan Citra X Ray Menggunakan Convolutional Neural Network (Classification of The Lung Diseases Based On X Ray Image Processing Using Convolutional Neural Network). *e-Proceeding of Engineering* 9(2): 476.
- Egger, J., Gsaxner, C., Pepe, A., Pomykala, L.K., Jonske, F., Kurz, M., Li, J., & Kleesiek J. 2022. Medical deep learning—A sistematic meta-review. *Computer Methods and Programs in Biomedicine*. Volume 221. 106874. ISSN 0169-2607. <https://doi.org/10.1016/j.cmpb.2022.106874>.
- El-Shafai, W., El-Nabi, S. A., El-Rabaie, E.-S. M., Ali, A. M., Soliman, N. F., Algarni, A. D., & Abd El-Samie, F. E. 2022. Efficient deep-learning-based autoencoder denoising approach for medical image diagnosis. *Computers, Materials & Continua*, vol. 70, no. 3, pp. 6107–6125, 2022
- Elshennawy, N.M., Ibrahim, D.M. 2020. Deep-Pneumonia Framework Using Deep Learning Models Based on Chest X-Ray Images. *Diagnostics*. 10. 649. <https://doi.org/10.3390/diagnostics10090649>
- Haddad M, Sharma S. 2023. Physiology, Lung. StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing.

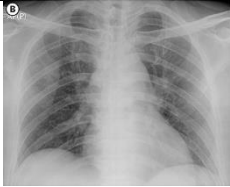



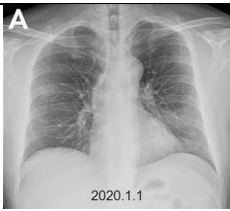


- Hansell, D. M., Bankier, A. A., MacMahon, H., McLoud, T. C., Müller, N. L., & Remy, J. 2008. Fleischner Society: glossary of terms for thoracic imaging. *Radiology*, 246(3), 697–722. <https://doi.org/10.1148/radiol.2462070712>
- Hariri, M., & Avşar, E. 2023. COVID-19 and pneumonia diagnosis from chest X-ray images using convolutional neural networks. *Network modeling and analysis in health informatics and bioinformatics*, 12(1), 17. <https://doi.org/10.1007/s13721-023-00413-6>
- Hinton, G. E., Osindero, S., & Teh, Y. W. 2006. A fast learning algorithm for deep belief nets. *Neural computation*, 18(7), 1527–1554. <https://doi.org/10.1162/neco.2006.18.7.1527>
- Huang, M. L., & Liao, Y. C. 2022. A lightweight CNN-based network on COVID-19 detection using X-ray and CT images. *Computers in biology and medicine*, 146, 105604. <https://doi.org/10.1016/j.compbiomed.2022.105604>
- Huang, M. L., & Liao, Y. C. 2023. Stacking Ensemble and ECA-EfficientNetV2 Convolutional Neural Networks on Classification of Multiple Chest Diseases Including COVID-19. *Academic radiology*, 30(9), 1915–1935. <https://doi.org/10.1016/j.acra.2022.11.027>
- Hunt, B. R. 1983. Digital Image Processing, *Advances in Electronics and Electron Physics*, Volume 60, 1983, Pages 161–221, [https://doi.org/10.1016/S0065-2539\(08\)60890-2](https://doi.org/10.1016/S0065-2539(08)60890-2)
- Indolia, S., Goswami, A. K., Mishra, S. P., & Asopa, P. 2018. Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. In *Procedia Computer Science* (Vol. 132, pp. 679–688). Elsevier B.V. <https://doi.org/10.1016/j.procs.2018.05.069>
- Indonesia Raih Rekor Capaian Deteksi TBC Tertinggi di Tahun 2022. 2023. ayosehat.kemkes.go.id. (Diakses tanggal 18 September 2023)
- Islam, M. Z., Islam, M. M., & Asraf, A. 2020. A Combined Deep CNN-LSTM Network for the Detection of Novel Coronavirus (COVID-19) Using X-ray Images. *Informatics in Medicine Unlocked*, Volume 20, 2020, 100412. ISSN 2352-9148, <https://doi.org/10.1016/j.imu.2020.100412>.
- Kanan C, Cottrell GW. 2012. Color-to-Grayscale: Does the Method Matter in Image Recognition ?. *PLoS ONE* 7(1): e29740. doi:10.1371/journal.pone.0029740
- Kulkarni, A., Chong, D., & Batarseh, F.A. 2020. *Data Democracy*. Editor(s): Feras A. Batarseh, Ruixin Yang. ISBN 9780128183663. Academic Press.
- Kundu R, Das R, Geem ZW, Han G-T, Sarkar R. 2021. Pneumonia detection in chest X-ray images using an ensemble of deep learning models. *PLoS ONE* 16(9): e0256630. <https://doi.org/10.1371/journal.pone.0256630>
- Leblanc, M., Tibshirani, R., 1996. Combining estimates in regression and classification. *Journal of the American Statistical Association* 91, 1641–1650. doi:10.1080/01621459.1996.10476733
- Li, P., Tang, H., Yu, J., & Song, W. 2020. LSTM and multiple CNNs based event image classification. *Multimedia Tools and Applications*. doi:10.1007/s11042-020-10165-4
- Loverdos, K., Fotiadis, A., Kontogianni, C., Iliopoulou, M., & Gaga, M. 2019. Lung nodules: A comprehensive review on current approach and management. *Annals of thoracic medicine*, 14(4), 226–238. https://doi.org/10.4103/atm.ATM_110_19

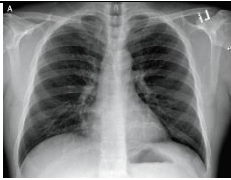


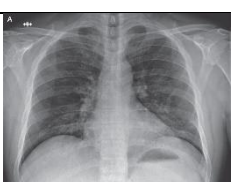
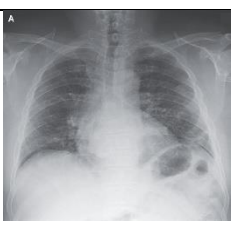
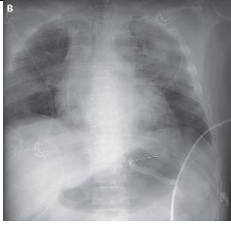


- Mabrouk, A., Redondo, R. P. D., Dahou, A., Abd Elaziz, M. & Kayed, M. 2022. Pneumonia Detection on Chest X-ray Images Using Ensemble of Deep Convolutional Neural Networks. *Applied Sciences*. 12(13):6448. <https://doi.org/10.3390/app12136448>
- Malhotra, P., Gupta, S., Koundal, D., Zaguia, A., Kaur, M., & Lee, H. N. 2022. Deep Learning-Based Computer-Aided Pneumothorax Detection Using Chest X-ray Images. *Sensors (Basel, Switzerland)*, 22(6), 2278. <https://doi.org/10.3390/s22062278>
- Mathiasen, V. D., Andersen, P. H., Johansen, I. S., Lillebaek, T., & Wejse, C. 2020. Clinical features of tuberculous lymphadenitis in a low-incidence country. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases*, 98, 366–371. <https://doi.org/10.1016/j.ijid.2020.07.011>
- Mbuh, T. P., Ane-Anyangwe, I., Adeline, W., Thumamo Pokam, B. D., Meriki, H. D., & Mbacham, W. F. 2019. Bacteriologically confirmed extra pulmonary tuberculosis and treatment outcome of patients consulted and treated under program conditions in the littoral region of Cameroon. *BMC pulmonary medicine*, 19(1), 17. <https://doi.org/10.1186/s12890-018-0770-x>
- Mehr, R. A. & Ameri, A. 2022. Skin Cancer Detection Based on Deep Learning. *Journal of biomedical physics & engineering*, 12(6), 559–568. <https://doi.org/10.31661/jbpe.v0i0.2207-1517>
- Maysanjaya, I.M.D. 2020. Classification of Pneumonia Based on Lung X-rays Images using Convolutional Neural Network. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 9(2), 190-195. <https://doi.org/10.22146/jnteti.v9i2.66>
- Nair, V. & Hinton, G.E. 2010. Rectified Linear Units Improve Restricted Boltzmann Machines. *Proceedings of the 27th International Conference on Machine Learning*, Haifa, 21 June 2010, 807-814.
- Nasser, M., & Yusof, U. K. 2023. Deep Learning Based Methods for Breast Cancer Diagnosis: A Sistematic Review and Future Direction. *Diagnostics (Basel, Switzerland)*, 13(1), 161. <https://doi.org/10.3390/diagnostics13010161>
- Oloko-Oba, M., & Viriri, S. 2020. Diagnosing Tuberculosis Using Deep Convolutional Neural Network. *Image and Signal Processing: 9th International Conference, ICISP 2020, Marrakesh, Morocco, June 4–6, 2020, Proceedings*, 12119, 151–161. https://doi.org/10.1007/978-3-030-51935-3_16
- Powers, D.M. 2011. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *ArXiv*, abs/2010.16061.
- Powers, D.M. 2015. What the F-measure doesn't measure: Features, Flaws, Fallacies and Fixes. *ArXiv*, abs/1503.06410.
- Pulmonary Nodules. my.clevelandclinic.org. (Diakses tanggal 19 November 2023)
- Pralambang, S.D., & Sona, S. 2021. Faktor Risiko Kejadian Tuberkulosis di Indonesia. *Jurnal FKM UI* 2021.
- Rahmadewi, R., & Kurnia, R. 2016. Klasifikasi Penyakit Paru Berdasarkan Citra Rontgen dengan Metoda Segmentasi Sobel. *Jurnal Nasional Teknik Elektro*, 5(1), 7–12. <https://doi.org/10.25077/jnte.v5n1.174.2016>
- Rokom.2022. Tahun ini, Kemenkes Rencanakan Skrining TBC Besar-besaran. sehatnegeriku.kemkes.go.id, 22 Maret 2022 (diakses 20 Oktober 2023).
- Septyhan, S., Magdalena, R. & Pratiwi, N. K. C. 2022. Deep Learning Untuk Deteksi Covid-19, Pneumonia, Dan Tuberculosis Pada Citra Rontgen Dada





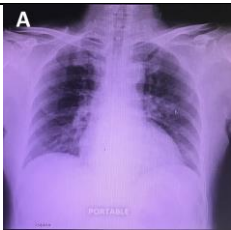


- Menggunakan CNN Dengan Arsitektur Alexnet. *e-Proceeding of Engineering* 8(6): 2869. ISSN: 2355-9365.
- Sharma, A., Ahmad Farouk, I., & Lal, S. K. 2021. COVID-19: A Review on the Novel Coronavirus Disease Evolution, Transmission, Detection, Control and Prevention. *Viruses*, 13(2), 202. <https://doi.org/10.3390/v13020202>.
- Shorten, C. & Khoshgoftaar, T.M. 2019. A survey on Image Data Augmentation for Deep Learning. *J Big Data* 6. 60. <https://doi.org/10.1186/s40537-019-0197-0>
- Siddiqui F, Vaqar S, Siddiqui AH. Lung Cancer. [Updated 2023 May 8]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK482357/>
- Teuwen, J., & Moriakov, N. 2020. Convolutional neural networks. In Handbook of medical image computing and computer assisted intervention (pp. 481-501). Academic Press.
- Tyagi, V. 2018. Understanding Digital Image Processing (1st ed.). CRC Press. <https://doi.org/10.1201/9781315123905>
- Wang, L. 2022. Deep Learning Techniques to Diagnose Lung Cancer. *Cancers (Basel)*. 2022 Nov 13;14(22):5569. doi: 10.3390/cancers14225569. PMID: 36428662; PMCID: PMC9688236.
- Wang, D., Mo, J., Zhou, G., Xu, L. & Liu, Y. 2020. An efficient mixture of deep and machine learning models for COVID-19 diagnosis in chest X-ray images. *PLoS One*. 2020 Nov 17;15(11):e0242535. doi: 10.1371/journal.pone.0242535. PMID: 33201919; PMCID: PMC7671547.
- Win, K.Y. Maneerat, N. Sreng, S. Hamamoto, K. 2021. Ensemble Deep Learning for the Detection of COVID-19 in Unbalanced Chest X-ray Dataset. *Applied Sciences*. 11(22):10528. <https://doi.org/10.3390/app112210528>
- Xing, D., Wang, Y., Sun, P. *et al.* 2023. A CNN-LSTM-att hybrid model for classification and evaluation of growth status under drought and heat stress in chinese fir (*Cunninghamia lanceolata*). *Plant Methods*. 19. 66. <https://doi.org/10.1186/s13007-023-01044-8>
- Zeineldin, R.A., Karar, M.E., Elshaer, Z. *et al.* 2022. Explainability of deep neural networks for MRI analysis of brain tumors. *Int J CARS* 17, 1673–1683. <https://doi.org/10.1007/s11548-022-02619-x>.






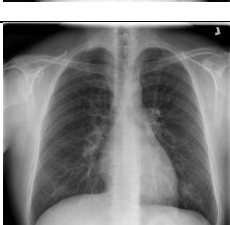
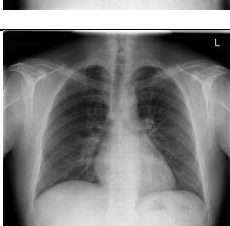
Lampiran








Tabel hasil keseluruhan tes








No.	Citra	Aktual	Sistem	Status
1		COVID-19	COVID-19	Benar
2		COVID-19	COVID-19	Benar
3		COVID-19	COVID-19	Benar
4		COVID-19	COVID-19	Benar
5		COVID-19	COVID-19	Benar
6		COVID-19	COVID-19	Benar
7		COVID-19	COVID-19	Benar








8		COVID-19	Lung Nodules	Salah
9		COVID-19	COVID-19	Benar
10		COVID-19	COVID-19	Benar
11		COVID-19	COVID-19	Benar
12		COVID-19	COVID-19	Benar
13		COVID-19	COVID-19	Benar
14		COVID-19	COVID-19	Benar
15		COVID-19	COVID-19	Benar










16		COVID-19	COVID-19	Benar
17		COVID-19	COVID-19	Benar
18		COVID-19	COVID-19	Benar
19		COVID-19	COVID-19	Benar
20		COVID-19	COVID-19	Benar
21		COVID-19	COVID-19	Benar
22		COVID-19	COVID-19	Benar











23		COVID-19	COVID-19	Benar
24		COVID-19	COVID-19	Benar
25		COVID-19	COVID-19	Benar
26		Lung Nodules	Tuberculosis	Salah
27		Lung Nodules	Lung Nodules	Benar
28		Lung Nodules	Lung Nodules	Benar
29		Lung Nodules	Lung Nodules	Benar











30		Lung Nodules	Lung Nodules	Benar
31		Lung Nodules	Lung Nodules	Benar
32		Lung Nodules	Lung Nodules	Benar
33		Lung Nodules	Lung Nodules	Benar
34		Lung Nodules	Lung Nodules	Benar
35		Lung Nodules	Lung Nodules	Benar
36		Lung Nodules	Lung Nodules	Benar











37		Lung Nodules	Lung Nodules	Benar
38		Lung Nodules	Lung Nodules	Benar
39		Lung Nodules	Lung Nodules	Benar
40		Lung Nodules	Lung Nodules	Benar
41		Lung Nodules	Lung Nodules	Benar
42		Lung Nodules	Lung Nodules	Benar
43		Lung Nodules	Lung Nodules	Benar











44		Lung Nodules	Lung Nodules	Benar
45		Lung Nodules	Lung Nodules	Benar
46		Lung Nodules	Lung Nodules	Benar
47		Lung Nodules	Lung Nodules	Benar
48		Lung Nodules	Lung Nodules	Benar
49		Lung Nodules	Lung Nodules	Benar
50		Lung Nodules	Lung Nodules	Benar








51		Normal	Normal	Benar
52		Normal	Normal	Benar
53		Normal	Normal	Benar
54		Normal	Normal	Benar
55		Normal	Normal	Benar
56		Normal	Normal	Benar
57		Normal	Normal	Benar
58		Normal	Normal	Benar
59		Normal	Normal	Benar








60		Normal	Normal	Benar
61		Normal	Normal	Benar
62		Normal	Normal	Benar
63		Normal	Normal	Benar
64		Normal	Normal	Benar
65		Normal	Pneumonia	Salah
66		Normal	Normal	Benar
67		Normal	Normal	Benar
68		Normal	Normal	Benar
69		Normal	Normal	Benar








70		Normal	Normal	Benar
71		Normal	Normal	Benar
72		Normal	Normal	Benar
73		Normal	COVID-19	Salah
74		Normal	Pneumonia	Salah
75		Normal	Normal	Benar
76		Pneumonia	Pneumonia	Benar
77		Pneumonia	Pneumonia	Benar
78		Pneumonia	Pneumonia	Benar
79		Pneumonia	Pneumonia	Benar



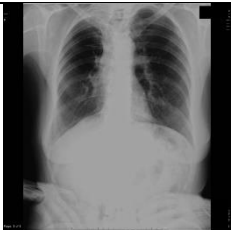

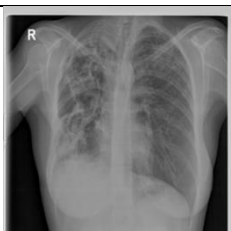
80		Pneumonia	Pneumonia	Benar
81		Pneumonia	Pneumonia	Benar
82		Pneumonia	Pneumonia	Benar
83		Pneumonia	Pneumonia	Benar
84		Pneumonia	Pneumonia	Benar
85		Pneumonia	Pneumonia	Benar
86		Pneumonia	Pneumonia	Benar
87		Pneumonia	Pneumonia	Benar
88		Pneumonia	Pneumonia	Benar
89		Pneumonia	Pneumonia	Benar

90		Pneumonia	Tuberculosis	Salah
91		Pneumonia	Pneumonia	Benar
92		Pneumonia	Pneumonia	Benar
93		Pneumonia	Pneumonia	Benar
94		Pneumonia	Pneumonia	Benar
95		Pneumonia	Pneumonia	Benar
96		Pneumonia	Pneumonia	Benar
97		Pneumonia	Pneumonia	Benar
98		Pneumonia	Pneumonia	Benar
99		Pneumonia	Pneumonia	Benar

100		Pneumonia	Pneumonia	Benar
101		Tuberculosis	Tuberculosis	Benar
102		Tuberculosis	Tuberculosis	Benar
103		Tuberculosis	Tuberculosis	Benar
104		Tuberculosis	Tuberculosis	Benar
105		Tuberculosis	Tuberculosis	Benar
106		Tuberculosis	Tuberculosis	Benar

107		Tuberculosis	COVID-19	Salah
108		Tuberculosis	COVID-19	Salah
109		Tuberculosis	Tuberculosis	Benar
110		Tuberculosis	Tuberculosis	Benar
111		Tuberculosis	Tuberculosis	Benar
112		Tuberculosis	Tuberculosis	Benar
113		Tuberculosis	Tuberculosis	Benar

114		Tuberculosis	Tuberculosis	Benar
115		Tuberculosis	Tuberculosis	Benar
116		Tuberculosis	Tuberculosis	Benar
117		Tuberculosis	Tuberculosis	Benar
118		Tuberculosis	Tuberculosis	Benar
119		Tuberculosis	Tuberculosis	Benar
120		Tuberculosis	Tuberculosis	Benar

121		Tuberculosis	Tuberculosis	Benar
122		Tuberculosis	Tuberculosis	Benar
123		Tuberculosis	Tuberculosis	Benar
124		Tuberculosis	Tuberculosis	Benar
125		Tuberculosis	Tuberculosis	Benar