



Pebri Andhi Herry Pratama, ST

Profil

Saya adalah sarjana fresh graduate lulusan Teknik Informatika, Universitas Palangkaraya pada tahun 2021. Saya memiliki skill dan pengalaman di bidang IT seperti aplikasi perkantoran, pemrograman Web, dan Machine Learning.

Kontak

Telepon/WA:
0812-5673-3006

Email:
nyahuidjabar@gmail.com

Github:
<https://github.com/NDYP>

Alamat:
Jl. Hentak No.20, Tewah, Gunung Mas

Pengalaman

Pelatihan IBM - Machine Learning

Fresh Graduate Academy Digi Talent Scholarship | 31 Mei - 23 Juli 2021

- Menganalisis dan membuat visualisasi data menggunakan Python
- Membuat model Machine Learning untuk klasifikasi data teks, dan data digital (gambar)

Pendidikan

Sarjana Teknik

Teknik Informatika, Universitas Palangkaraya | 2016-2021

- IPK 3.30
- Judul Skripsi : "Deteksi COVID-19 Berdasarkan Hasil Rontgen Dada Menggunakan Python"

Keahlian

Apikasi Perkantoran Bahasa Pemrograman Framework

- | | | |
|------------------------|----------------|--------------------|
| • Microsoft Excel | • HTML dan CSS | • Bootstrap |
| • Microsoft Word | • PHP | • Codeigniter |
| • Microsoft Powerpoint | • Python | • Keras/Tensorflow |

Kemampuan

- Data Analisis dan Visualisasi Data (Machine Learning)
- Desain dan Pemrograman Website
- Desain dan implementasi Database
- Dapat bekerja di dalam lingkungan tim, maupun secara individual
- Kemampuan mengikuti instruksi dengan baik

Bahasa

Indonesia
Fasih

Dayak
Fasih

Inggris (US)
Percakapan



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN UNIVERSITAS PALANGKA RAYA

Terkreditasi B, SK BAN-PT No. : 1032/SK/BAN-PT/Akred/PT/VI/2016, tanggal, 17 Juni 2016

Nomor : 552012021000153

Dengan ini menyatakan bahwa :

PEBRI ANDHI HERRY PRATAMA

Nomor Induk Mahasiswa : DBC116039

Nomor Induk Kependudukan : 6210031102980001

Lahir di Tewah tanggal 11 Februari 1998

Telah menyelesaikan dengan baik dan memenuhi segala syarat pendidikan SARJANA (S1) Pada Program Studi :

TEKNIK INFORMATIKA pada FAKULTAS TEKNIK

Oleh karena itu kepada diberikan Ijazah dengan Gelar

SARJANA TEKNIK (S.T)

Beserta segala hak dan kewajiban yang melekat pada gelar tersebut.

Di YUDISIUM pada tanggal : 31 Maret 2021

Dekan,

Ir. Waluyo Nuwantoro, M.T.

NIP. 19651119 199302 1 001



Diberikan di Palangka Raya

Pada tanggal : 22 April 2021

Rektor,

Dr. Andrie Elia, SE., M.Si

NIP. 19590812 198701 1 001

KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS PALANGKA RAYA
FAKULTAS TEKNIK

TRANSKRIP

Nomor Registrasi : 021/FT-TI/S1/IV/2021

NAMA	PEBRI ANDHI HERRY PRATAMA	JUR/PROG. STUDI	: Teknik Informatika
Nomor Induk Mahasiswa	: DBC 116 039	JENJANG	: Strata 1 (S1)
Selesai Tanggal	: 31 Maret 2021	AKREDITASI	: B

NO	Mata Kuliah	Kode MK	SKS	NILAI		SKSN	Keterangan
				Angka	Huruf		
I. Mata Kuliah Pengembangan Kepribadian							
1	Pendidikan Agama	DBPK 1011	3	4	A	12	
2	Pendidikan Pancasila dan Kewarganegaraan	DBPK 1061	3	3	B	9	
3	Bahasa Indonesia	DBPK 1081	3	4	A	12	
4	Bahasa Inggris I	DBPK 1051	2	4	A	8	
5	Bahasa Inggris II	DBPK 1092	2	2.5	C+	5	
II. Mata Kuliah Keilmuan dan Keterampilan							
1	Fisika	DBKK 1021	2	2	C	4	
2	Pengantar Teknologi Komputer & Informasi	DBKK 1041	3	4	A	12	
3	Kalkulus I	DBKK 1051	2	2.5	C+	5	
4	Kalkulus II	DBKK 1062	2	2.5	C+	5	
5	Logika Matematika	DBKK 1081	3	4	A	12	
6	Algoritma & Pemrograman I	DBKK 1091	3	3.5	B+	10.5	
7	Algoritma & Pemrograman II	DBKK 1102	3	3	B	9	
8	Algoritma & Pemrograman III	DBKK 1111	3	3	B	9	
9	Sistem Digital	DBKK 1151	3	3	B	9	
10	Statistika I	DBKK 1212	3	3	B	9	
11	Statistika II	DBKK 1221	3	4	A	12	
12	Sistem Berkas	DBKK 1261	3	4	A	12	
13	Pengantar Arsitektur Komputer	DBKK 1272	3	2.5	C+	7.5	
14	Metode Numerik	DBKK 1271	3	3	B	9	
15	Aljabar Vektor & Matriks	DBKK 1162	3	2.5	C+	7.5	
16	Struktur Data	DBKK 1172	3	4	A	12	
17	Matematika Diskrit	DBKK 1191	3	4	A	12	
18	Metodologi Penelitian	DBKK 1232	2	2.5	C+	5	

Bersambung ke Halaman Berikut

Lanjutan dari Halaman Muka

NAMA : PEBRI ANDHI HERRY PRATAMA Nomor Induk Mahasiswa : DBC 116 039

NO	Mata Kuliah	Kode MK	SKS	NILAI		SKSN	Keterangan
				Angka	Huruf		
III. Mata Kuliah Keahlian Berkarya							
1	Pemrograman Berorientasi Objek	DBKB 1012	3	2	C	6	Judul Skripsi : Deteksi Covid-19 Berdasarkan Hasil Rontgen Dada (Chest X-Ray) Menggunakan Python
2	Interaksi Manusia dan Komputer	DBKB 1172	3	3.5	B+	10.5	
3	Basis Data 1	DBKB 1192	3	3	B	9	
4	Basis Data 2	DBKB 1201	3	4	A	12	
5	Sistem Informasi	DBKB 1212	3	3.5	B+	10.5	
6	Internet & Intranet	DBKB 1222	3	4	A	12	
7	Keamanan Jaringan	DBKB 1221	3	2.5	C+	7.5	
8	Sistem Operasi	DBKB 1042	3	3	B	9	
9	Organisasi Komputer	DBKB 1031	3	3	B	9	
10	Komunikasi Data	DBKB 1052	3	3.5	B+	10.5	
11	Jaringan Komputer	DBKB 1061	3	2.5	C+	7.5	
12	Rekayasa Perangkat Lunak	DBKB 1091	3	4	A	12	
13	Teori Bahasa dan Otomata	DBKB 1101	3	2.5	C+	7.5	
14	Kecerdasan Buatan	DBKB 1132	3	3.5	B+	10.5	
15	Program Profesional	DBKB 1142	3	3.5	B+	10.5	
16	Grafika Komputer	DBKB 1161	3	2.5	C+	7.5	
17	Manajemen Proyek	DBKB 1232	3	3.5	B+	10.5	
18	Metode Pengembangan Perangkat Lunak	DBKB 1242	3	4	A	12	
19	Multimedia	DBKB 1252	3	3	B	9	
IV. Mata Kuliah Perilaku Berkarya							
1	Etika dan Profesi	DBPB 1012	2	4	A	8	
V. Mata Kuliah Berkehidupan Bermasyarakat							
1	Skripsi / Tugas Akhir	DBBB 1030	6	4	A	24	
2	Ilmu Sosial Budaya Dasar	DBBB 1042	3	4	A	12	
3	Kerja Praktek	DBBB 1010	2	4	A	8	
4	Kuliah Kerja Nyata	DBBB 1020	4	4	A	16	
VI. Mata Kuliah Pilihan							
1	Sistem Penunjang Keputusan	DBKP 1010	3	2	C	6	
2	Analisa Desain Sistem Informasi	DBKP 1040	3	4	A	12	
3	Pemodelan dan Simulasi Data	DBKP 1110	3	2	C	6	
4	Sistem Cerdas	DBKP 1120	3	3	B	9	
JUMLAH SKS			149			491.5	Predikat Yudisium : Sangat Memuaskan
Jumlah SKS Yang Dicapai (SKS) =				149		Jumlah SKS Kali Nilai (SKSN) =	491.5

Palangka Raya, 22 April 2021

Dekan

Ir. WALUYO NUSWANTORO, M.T.
NIP.196511191993021001





T.
Basaruddi
n

Digitally signed by
T. Basaruddin
ou= BAN-PT
Date: 2021.06.23
08:09:19 WIB

SERTIFIKAT AKREDITASI

Badan Akreditasi Nasional Perguruan Tinggi
berdasarkan Keputusan BAN-PT No. 557/SK/BAN-PT/Ak-PPJ/PT/VI/2021, menyatakan bahwa

Universitas Palangka Raya, Kota Palangka Raya

memenuhi syarat peringkat

Akreditasi B

*Sertifikat akreditasi perguruan tinggi ini berlaku
sejak tanggal 18 - Juni - 2021 sampai dengan 18 Juni 2026*



Jakarta, 22 - Juni - 2021

A handwritten signature in blue ink that appears to read "Basaruddin".

Prof. T. Basaruddin
Direktur Dewan Eksekutif



BAN-PT


Basaruddin

Digitally signed by
T. Basaruddin
ou= BAN-PT, o=
Kemristekdikti
Date: 2019.05.22 08:13:
08 WIB

Sertifikat Akreditasi

Badan Akreditasi Nasional Perguruan Tinggi

berdasarkan Keputusan BAN-PT No. 1627/SK/BAN-PT/Akred/S/V/2019, menyatakan bahwa
Program Studi **Teknik Informatika**, Pada Program Sarjana **Universitas Palangka Raya, Kota Palangka Raya**
Terakreditasi dengan peringkat

Terakreditasi B

*Sertifikat akreditasi ini berlaku 5 (lima) tahun
sejak tanggal 21 - Mei - 2019 sampai dengan 21 - Mei - 2024*



Jakarta, 21 - Mei - 2019

Prof. T. Basaruddin
Direktur Dewan Eksekutif



FGA Fresh Graduate Academy

SERTIFIKAT PELATIHAN

0182180221-4/DTS/BLSDM.KOMINFO/07/2021

Badan Penelitian dan Pengembangan Sumber Daya Manusia
Kementerian Komunikasi dan Informatika Republik Indonesia,
menyatakan bahwa :

Pebri Andhi Herry Pratama

Telah Mengikuti Pelatihan **IBM – Machine Learning**
Program Fresh Graduate Academy Digital Talent Scholarship
Sebagai Peserta mulai tanggal 31 Mei 2021 – 23 Juli 2021
yang meliputi 160 Jam Pelatihan

Jakarta, 25 Juli 2021

Kepala Badan Penelitian dan Pengembangan Sumber Daya Manusia Kementerian
Komunikasi dan Informatika



Dr. Hary Budiarto, M.Kom.



Fresh Graduate Academy Digital Talent Scholarship 2021

IBM – Machine Learning sejumlah 160 JP

- Kelas 1 – Memulai Pemrograman dengan Python (Dicoding)**
- Kelas 2 – Data Analysis with Python (IBM SkillsBuild)**
- Kelas 3 – Belajar Dasar Visualisasi Data (Dicoding)**
- Kelas 4 – Data Visualization with Python (IBM SkillsBuild)**
- Kelas 5 – Elements of AI (University of Helsinki)**
- Kelas 6 – Belajar Machine Learning untuk Pemula (Dicoding)**
- Kelas 7 – Belajar Pengembangan Machine Learning (Dicoding)**



Deteksi COVID-19 Berdasarkan Hasil Rontgen Dada (*Chest Xray*) Menggunakan Python

Pebri Andhi Herry Pratama¹⁾, Rony Teguh²⁾, Abertun Sagit Sahay³⁾, Valencia Wilentine⁴⁾

¹⁾²⁾³⁾ Teknik Informatika, Fakultas Teknik, Universitas Palangkaraya
Kampus Tanjung Nyaho Jl. Yos Sudarso Palangka Raya 73112

⁴⁾ Laboratorium Kesehatan dan Kalibrasi Provinsi Kalimantan Tengah
Palangka Raya
¹⁾ nyahuidjabar@gmail.com 1

Abstrak

Rontgen Dada (*Chest X-Ray*) memiliki fungsi penting dalam tiga bidang kesehatan yaitu diagnosis, perawatan dan pemeriksaan ulang. Studi dari Cina menyatakan *Chest radiograph* (X-Ray) dan *Chest computed tomography* (CT) scans dapat membantu mendiagnosa COVID-19. Oleh hal tersebut maka *Chest radiographs* (X-Ray) dan *Chest computed tomography* (CT) scans adalah metode yang tepat untuk mendeteksi infeksi paru-paru akibat dari COVID-19. Berdasarkan hal tersebut penulis mencoba membuat sebuah model untuk klasifikasi citra digital hasil rontgen dada (*Chest X-Ray*) dengan label kelas Normal, *Pneumonia*, *Tuberculosis* (TBC), dan COVID-19. Melalui model yang dihasilkan tersebut akan dibandingkan model yang terbaik untuk digunakan.

Metode yang digunakan untuk membuat model tersebut adalah melalui pelatihan (*training*) dan pengujian (*testing*) dataset menggunakan model arsitektur *Convolutional Neural Network* (CNN) yaitu VGG19, ResNet50, dan InceptionV3. Jumlah citra yang digunakan sebanyak 1.000 Gambar *Chest X-Ray*. Dataset tersebut dibagi menjadi data latih (*train*) dan validasi (*validation*) dalam beberapa rasio 20% : 80%, 50% : 50%, dan 80% : 20%. Sedangkan data yang digunakan sebagai data uji (*test*) diambil 10% dari data latih (*train*) dan 4 gambar *chest x-ray* sebagai dataset prediksi.

Berdasarkan hasil dari penelitian yang telah dilakukan, didapatkan model terbaik yaitu VGG19 pada iterasi (*epoch*) 41 dari 61 epoch dengan rasio 20% : 80% menjadi model terbaik untuk penelitian ini. Dengan menghasilkan akurasi pelatihan mencapai 0,9444 atau 94,44% (dalam persentase) dan nilai Loss 0,1084. Pada pengujian menghasilkan *Confussion Matrix* dengan akurasi sebesar 0.96 dan kehilangan 0.0939 Lalu untuk pengujian prediksi data baru menghasilkan akurasi dengan nilai 98.97% untuk label Normal, 99.16% untuk COVID-19, 99.56% untuk *Pneumonia* dan 99.79% untuk *Tuberculosis* (TBC).

Kata kunci: Deep Learning, Convolutional Neural Network, Python, Chest X-Ray, COVID-19.

Abstract

Chest X-ray have an important function in the three areas of healthcare, namely diagnosis, treatment and re-examination. Studies from China suggest Chest Radiographs (X-Ray) and Chest Computed Comography (CT) scans can help diagnose COVID-19. Therefore, chest x-rays (x-rays) and chest computed tomography (CT) are appropriate methods for lung infections due to COVID-19. Based on this, the authors tried to make a model for the classification of digital images of Chest X-Ray results with the labels of Normal, Pneumonia, Tuberculosis (TBC), and COVID-19. Through the resulting model, the best model to use will be compared.

The method used to create this model is through training and testing the dataset using the Convolutional Neural Network (CNN) architectural model, namely VGG19, ResNet50, and InceptionV3. The number of images used is 1,000 Chest X-Ray images. The dataset is divided into training and validation datasets in several ratios of 20% : 80%, 50%: 50%, and 80%: 20%.

While testing uses 10% from train dataset chest x-ray images as a confusion matrix dataset and 4 chest x-ray images as a prediction dataset.

From the results of the research that has been done, the best model is VGG19 at 41 of 61 epoch and a ratio of 20% : 80%. Where the VGG19 model produces 94.44% for accuracy and 0,1084 loss value for training. Whereas at the testing stage with a configuration matrix, 95% accuracy value was obtained. Then for testing the new data prediction produces the best accuracy with 98.97% accuracy for the Normal label, 99.16% for COVID-19, 99.56% for Pneumonia, and 99.79% for Tuberculosis (TBC).

Keywords: Deep Learning, Convolutional Neural Network, Python, Chest X-Ray, COVID-19.

1. PENDAHULUAN

Metode *Chest Radiograph* atau *Chest X-Ray* telah lama digunakan untuk mendeteksi anatomi tubuh manusia. Pemeriksaan *Chest X-Ray* adalah salah satu pemeriksaan pencitraan medis yang paling sering digunakan karena lebih terjangkau. Pembacaan foto *Chest X-Ray* memiliki keterbatasan, yaitu dalam menentukan hasilnya masih menggunakan metode manual dan sulit terdeteksinya penyakit, sehingga memerlukan waktu lama sebelum tenaga medis atau dokter mendiagnosa penyakit yang diderita oleh pasien. Salah satu metode untuk mengatasi masalah tersebut adalah dengan mengklasifikasikan citra *Chest X-Ray* ke dalam kelas atau label tertentu dengan menggunakan *Deep Learning*. Data yang berupa citra *Chest X-Ray*, merupakan data gambar dua dimensi yang dapat diproses menggunakan *deep learning* dengan metode *Convolutional Neural Network* (CNN). CNN digunakan untuk mengklasifikasi data yang terlabel dengan menggunakan metode *Supervised Learning*, yang mana cara kerja cari supervised learning adalah terdapat data yang dilatih dan terdapat variabel yang ditargetkan sehingga tujuan dari metode ini adalah mengklasifikasikan suatu data ke dalam data yang sudah ada. Kemampuan CNN di klaim sebagai model terbaik untuk memecahkan permasalahan *object detection* dan *object recognition* pada data citra sebuah gambar hasil rontgen dada.

Penelitian mengenai *object recognition* telah dilakukan oleh beberapa peneliti. Salah satunya penelitian yang berjudul penelitian Implementasi Deep Learning Menggunakan Convolutional Neural Network (CNN) Untuk Klasifikasi Alat Tulis. Pada penelitian tersebut bertujuan untuk pengenalan objek dengan kelas *ballpoint*, penghapus, dan penggaris. Data *sample* yang digunakan sebanyak 300 gambar merupakan hasil *crawling* dari *google image*. Akurasi yang dihasilkan model untuk data train adalah 100% dan untuk data test sebesar 95%. Kemudian dilakukan percobaan untuk 3 data baru, didapatkan 1 gambar masuk kategori yang salah [1].

Berdasarkan permasalahan tersebut penulis ingin menerapkan penelitian yang telah dilakukan peneliti sebelumnya menggunakan metode yang sama dengan kasus yang berbeda. Penulis mengusulkan sebuah model yang dapat membaca hasil rontgen dada (*Chest X-Ray*) untuk diagnosis COVID-19 dengan menggunakan tiga model Convolutional Neural Network (CNN), yaitu model VGG19, ResNet50, dan InceptionV3. Dari ketiga model CNN tersebut akan dibandingkan model yang cocok digunakan untuk diagnosa beberapa label atau kelas dataset yaitu normal, COVID-19, *Tuberculosis*, dan *Pneumonia* dari hasil rontgen dada. Lalu model yang menghasilkan akurasi terbaik akan diimplementasikan ke sebuah aplikasi berbasis Web, sehingga dapat digunakan secara umum. Sehingga teknologi *deep learning* dapat diterapkan dalam bidang kesehatan dan menjadi alat bantu seorang radiologis ataupun dokter dalam hal diagnosis hasil rontgen dada (*Chest X-Ray*).

2. TINJAUAN PUSTAKA

Penelitian Pertama dengan jurnal penelitian yang berjudul “Implementasi Sistem Pakar Menggunakan Metode Certainty Factor Untuk Mendiagnosa Dini Corona Virus Disease (COVID-19)”. Dari 152 data pasien yang di input pada penelitian ini mendapatkan hasil 114 ODP dengan rata-rata nilai CF 91,38% , 36 PDP dengan rata-rata nilai CF 98,25% dan 2 NON

dengan rata-rata nilai CF 40%. CF dengan percobaan perhitungan sistem yaitu data yang mewakili pasien mendapatkan nilai CF 0.998848 atau 99.88% menjadi PDP [2].

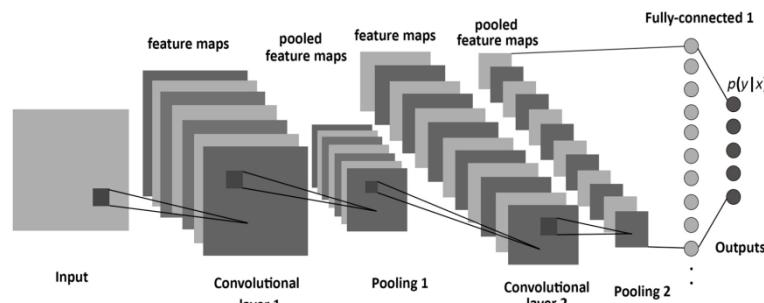
Penelitian kedua dengan jurnal penelitian yang berjudul Implementasi Metode Forward Chaining Untuk Diagnosa Penyakit COVID-19 di Rsud Berkah Pandeglang Banten. Berdasarkan hasil penelitian ini metode forward chaining sangat cocok untuk penelitian diagnosis Virus COVID 19 dikarenakan algoritma forward chaining menggunakan algoritma runut maju yang setiap langkah mendiagnosis gejala dan langkah selanjutnya yaitu membuat system aplikasi agar penelitian ini dapat berkembang menjadi sistem pakar deteksi Virus COVID 19 dengan metode forward chaining kedepannya [3].

Penelitian Ketiga dengan penelitian yang berjudul penelitian Implementasi Deep Learning Menggunakan Convolutional Neural Network Untuk Klasifikasi Alat Tulis. Akurasi yang dihasilkan model untuk data train adalah 100% dan untuk data test sebesar 95%. Kemudian dilakukan percobaan untuk 3 data baru, didapatkan 1 gambar masuk kategori yang salah [1].

3. METODE PENELITIAN

3.1 Convolutional Neural Network (CNN)

Metode yang digunakan dalam penelitian adalah melalui pelatihan (*training*) dan pengujian (*testing*). CNN pada dasarnya adalah susunan banyak layer yang terdiri dari *convolution layer*, *pooling layer* yang disebut *feature layer*, dan *fully connected layer* yang disebut *classification layer*. Model arsitektur *Convolutional Neural Network* (CNN) yang digunakan pada penlitian yaitu VGG19, ResNet50, dan InceptionV3 yang telah dimodifikasi. Model-model yang digunakan tersebut dirancang untuk digunakan untuk klasifikasi ribuan label. Modifikasi dilakukan pada *classification layer*, karena jumlah label yang digunakan pada penelitian hanya empat kelas yaitu normal, *tuberculosis*(TBC), *pneumonia*, dan COVID-19. Setelah model melalui tahap pelatihan dan menghasilkan model untuk klasifikasi citra rontgen dada ke dalam 4 kelas. Tahap selanjutnya adalah megudi model tersebut dengan prediksi data baru menggunakan model yang menghasilkan nilai akurasi dan kehilangan paling optimal.



Gambar 1. CNN Concept [4]

3.2 Populasi dan Sampel

Populasi dalam penelitian ini adalah citra hasil rontgen dada (*Chest X-Ray*). Sedangkan sampel yang digunakan dalam penelitian ini adalah dataset citra rontgen dada yaitu normal, *tuberculosis*(TBC), *pneumonia*, dan COVID-19.

Tabel 1. Definisi Operasional

No	Variabel	Definisi
1	Normal	Dataset citra rontgen dada paru-paru normal yang tidak terkena infeksi pneumonia, <i>Tuberculosis</i> (TBC) ataupun COVID-19.
2	<i>Tuberculosis</i> (TBC)	Dataset citra rontgen dada terinfeksi <i>Tuberculosis</i> (TBC).
3	<i>Pneumonia</i>	Dataset citra rontgen dada infeksi paru-paru.
4	COVID-19	Dataset citra rontgen dada terinfeksi positif COVID-19.

3.3 Jenis dan Sumber Data

Jenis data yang digunakan untuk pelatihan (*training*) dan pengujian (*testing*) pada penelitian ini adalah data sekunder. Data yang digunakan berjumlah 1.000 gambar *Chest Xray* yang digunakan untuk pelatihan (*training*), dengan pembagian rasio 20% : 80%, 50% : 50%, dan 80% : 20% untuk data uji (*train*) dan data validasi (*validation*). Sedangkan data uji (uji) yang digunakan untuk pengujian (*testing*) sebesar 10% data yang dibagi dari data uji (*train*).

Tabel 2. Sumber Dataset Pelatihan (*training*)

No	Variabel	Judul	Sumber (link)	Jumlah
1.	Normal	Chest X-Ray Images (pneumonia) [5]	https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia	250
2.	<i>Tuberculosis</i> (TBC)	Pulmonary Chest X-Ray Abnormalities [6]	https://www.kaggle.com/kmader/pulmonary-chest-xray-abnormalities?select=Montgomery	250
3.	<i>Pneumonia</i>	Chest X-Ray Images [5]	https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia	250
4.	<i>COVID-19</i>	COVID-19 Radiography Database [7]	https://www.kaggle.com/tawsifurrahman/covid19-radiography-database	219
		COVID-19 chest xray [8]	https://www.kaggle.com/bachrir/covid-chest-xray	31
Jumlah				1.000

3.4 Rancangan Dataset

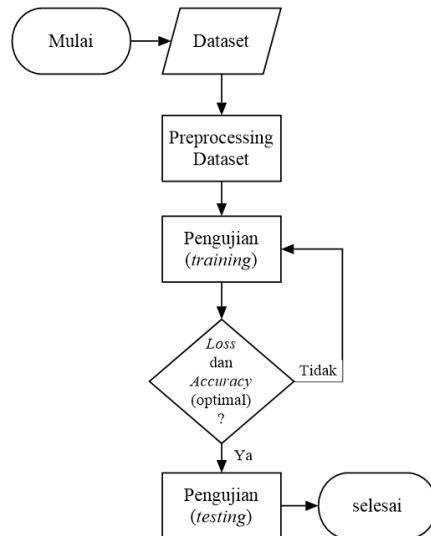
Dataset dibagi berdasarkan 4 label yang dapat dilihat pada Gambar 2 yaitu gambar hasil rontgen dada Normal, *Tuberculosis* (TBC), *Pneumonia*, dan positif COVID-19.



Gambar 2. Label Dataset (Normal, COVID-19, *Pneumonia*, dan *Tuberculosis*)

3.5 Rancangan Deep Learning

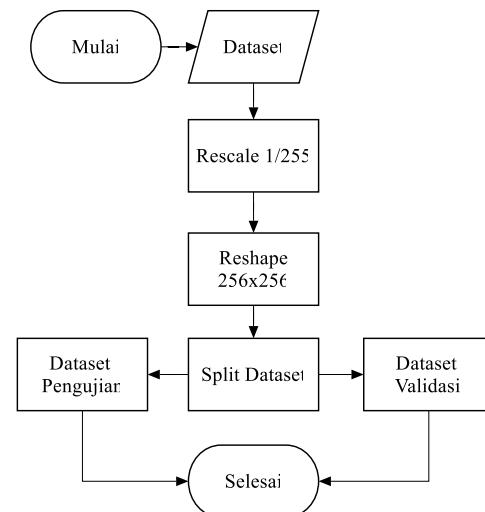
Rancangan *deep learning* terdiri dari beberapa tahapan, dimulai dari pengumpulan data citra Normal, *Tuberculosis* (TBC), *Pneumonia* dan COVID-19 yang akan digunakan untuk citra latih (*train*) dan citra uji (*test*). Lalu masuk ke tahap *preprocessing*, kemudian pelatihan menggunakan model-model *Convolutional Neural Network* (CNN), lalu terakhir data citra yang digunakan sebagai pengujian (*testing*) di klasifikasikan ke dalam label dataset yang telah dilatih.



Gambar 3. Flowchart Desgin Deep Learning

3.6 Rancangan Image Processing

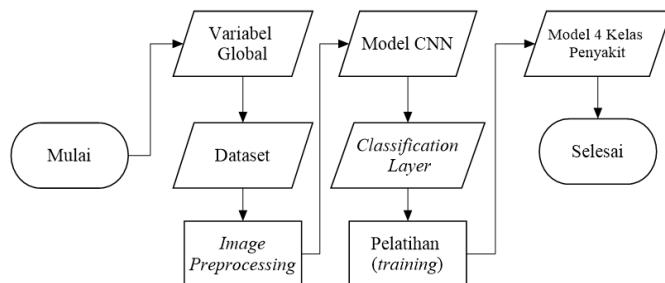
Perancangan *preprocessing* merupakan langkah yang wajib dilakukan sebelum masuk perancangan *Convolutional Neural Network* (CNN). *Preprocessing* berfungsi untuk menyiapkan citra agar dapat digunakan pada proses selanjutnya. Hal-hal yang dilakukan dalam perancangan *preprocessing* yang pertama yaitu dilakukan proses *thresholding* yang merupakan proses untuk menjadikan nilai citra menjadi 0-1 dari 0-255. Kemudian proses *resize* menjadi ukuran 256x256 untuk model VGG19, 244x224 untuk ResNet50, dan 299x299 untuk InceptionV3 seperti pada gambar 3.6 yang berguna untuk normalisasi citra agar seluruh dalam dataset sama. Setelah itu dilakukan proses *split* untuk membagi data dalam beberapa rasio (20% : 80% ; 50% : 50% ; 80% : 80%) yang telah dikumpulkan menjadi dua, yaitu data uji (*train*) dan data validasi (*validation*) yang digunakan untuk proses pengujian (*training*). Kemudian dari data uji (*train*) diambil sebesar 10% untuk digunakan sebagai data uji (*test*) pada proses pengujian (*testing*).



Gambar 4. Design flowchart preprocessing dataset

3.7 Rancangan Pelatihan (*training*)

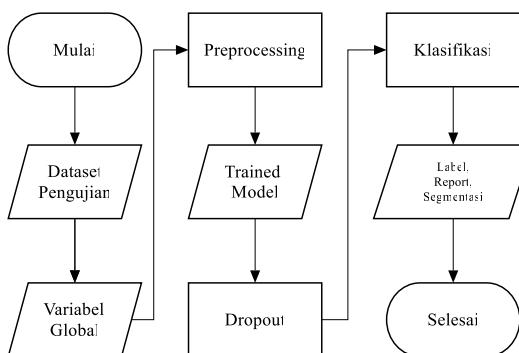
Dataset latih (*train*) dan data validasi (*validation*) yang telah melewati *preprocessing* dapat dilanjutkan untuk proses pelatihan (*training*). Proses *training* ini bertujuan untuk melatih algoritma *Convolutional Neural Network* (CNN) dalam mempelajari pola setiap gambar pada dataset yang telah melewati *image processing* dan membentuk sebuah model baru. Rancangan *training* yang diajukan untuk pada penelitian ini terdiri dari beberapa tahapan seperti pada Gambar 5.



Gambar 5. Flowchart Proses Pelatihan (*training*)

3.8 Rancangan Pengujian (*testing*)

Proses ini mencocokkan antara data masukkan dengan data yang telah melalui proses *training*. Setelah gambar *Chest X-Ray* telah melewati *preprocessing* Tahap selanjutnya adalah memilih file *training* yang telah disimpan untuk digunakan sebagai klasifikasi. Setelah proses klasifikasi dijalankan, maka akan ditampilkan Gambar *Chest X-Ray* beserta dengan hasil klasifikasi berupa label dan persentase akurasi.



Gambar 6. Flowchart Pengujian (*Testing*)

4. PEMBAHASAN

4.1 *Preprocessing*

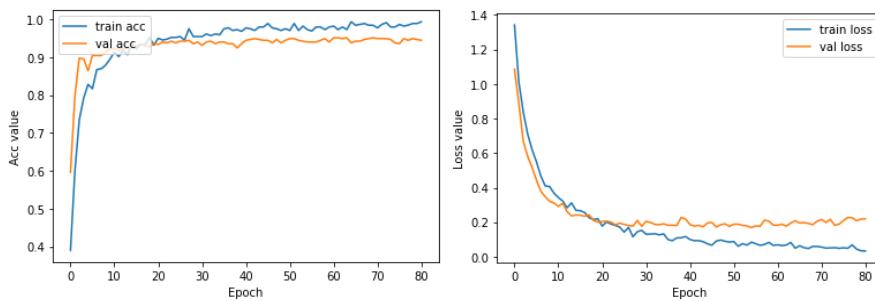
Pada tahap *peporcessing* ini, ukuran keseluruhan dataset dinormalisasi ukuran 256x256 untuk model VGG19, 244x224 untuk ResNet50, dan 299x299 untuk InceptionV3. Lalu dataset dilakukan proses *rescale* yang merupakan proses untuk menjadikan nilai citra menjadi 0-1 dari 0-255 yang gunanya mempermudah dalam proses input *training* dan *testing*. Setelah itu dilakukan proses *split* untuk membagi data dalam beberapa rasio (20% : 80% ; 50% : 50% ; 80% : 80%) yang telah dikumpulkan menjadi dua, yaitu data uji (*train*) dan data validasi (*validation*) yang digunakan untuk proses pengujian (*training*). Kemudian dari data uji (*train*) diambil sebesar 10% untuk digunakan sebagai data uji (*test*) pada proses pengujian (*testing*).

4.2 Implementasi Model

Setelah melakukan *preprocessing* pada dataset, tahap selanjutnya adalah menentukan model yang digunakan untuk proses *training*. Pada penelitian ini menggunakan beberapa model dari CNN, yaitu VGG19, ResNet50, dan InceptionV3. Model tersebut digunakan sebagai *feature layer*, sedangkan *classification layer* di modifikasi menyesuaikan jumlah label yang digunakan. Dikarenakan jumlah kelas pada imagenet adalah 1.000 maka ini tidak cocok dengan jumlah kelas yang kita miliki yaitu hanya 4. Maka pada penelitian menggunakan beberapa *layer* sebagai model output.

4.3 Pelatihan (*training*)

Setelah membuat model selanjutnya adalah melakukan proses *training*, dimana proses *training* ini merupakan proses dimana *machine learning* bekerja sehingga algoritma yang sudah kita definisikan dapat mengingat pola dari masing masing kelas pada data yang kita *train*. Pada proses *training* penelitian ini menggunakan proses pengulangan atau iterasi (*epoch*) *training* yang terhenti otomatis saat akurasi sudah tidak dapat naik dari akurasi dari beberapa *epoch* sebelumnya.



Gambar 7. Grafik Akuraasi (Accuracy) dan Kehilangan (Loss) VGG19 Rasio 50%:50%

Gambar 7 menunjukkan grafik pergerakan nilai akurasi (*acc/accuracy*) dan nilai *loss* dari hasil pelatihan (*training*) dan validasi (*validation*). Berdasarkan gambar tersebut, garis berwarna biru menunjukkan pergerakan nilai akurasi dan kehilangan untuk data pelatihan, sedangkan garis berwarna orange menunjukkan pergerakan nilai akurasi dan kehilangan untuk data validasi.

Pada penelitian ini saat proses pelatihan, setiap terjadi peningkatan akurasi maka pelatihan dilanjutkan, sedangkan jika tidak ada peningkatan akurasi dalam beberapa *epoch* maka pelatihan akan otomatis terhenti. Gambar 4.4 pelatihan terhenti pada *epoch* 81, karena tidak ada pengingkatan akurasi dari *epoch* 61. Sehingga akurasi tertinggi dihasilkan pada *epoch* 61. Begitu juga dengan model lainnya jika proses pelatihan tidak terdapat peningkatan pada beberapa *epoch* maka proses pelatihan akan terhenti otomatis.

Berdasarkan pelatihan model-model yang telah dijalankan, didapatkan variasi waktu untuk setiap model dengan iterasi (*epoch*) dan rasio dataset yang berbeda.

Tabel 4. Waktu Pelatihan

Rasio	Model	VGG19	ResNet50	InceptionV3
20%	Epoch	49	21	33
	Waktu	00:25:08	0:10:09	00:11:53
50%	Epoch	81	21	27
	Waktu	00:39:05	0:10:09	00:14:48
80%	Epoch	71	27	35
	Waktu	00:34:27	0:15:54	0:21:17

4.4 Validasi

Untuk mengetahui akurasi (*accuracy*) dan kehilangan (*loss*) rata-rata pada setiap model maka dilakukan evaluasi model menggunakan dataset *validation*. Tabel 4 adalah Tabel evaluasi hasil *training* setiap model yang telah dilakukan berulang-ulang untuk mendapatkan nilai akurasi (*accuracy*) dan kehilangan (*loss*) yang optimal untuk digunakan pada proses pengujian (*testing*).

Tabel 5. Tabel Evaluasi Hasil *Training*

Rasio	Model	VGG19	ResNet50	InceptionV3
	<i>Epoch</i>	41	1	13
20%	<i>Accuracy</i>	0,9444	0,25	0,6222
	<i>Loss</i>	0,1084	4,72	1,4805
50%	<i>Epoch</i>	61	1	7
	<i>Accuracy</i>	0.9335	0.2695	0,6328
80%	<i>Loss</i>	0.2354	3,1717	1,0025
	<i>Epoch</i>	51	7	15
	<i>Accuracy</i>	0.9335	0,2695	0,6523
	<i>Loss</i>	0.2329	3,7394	1,1784

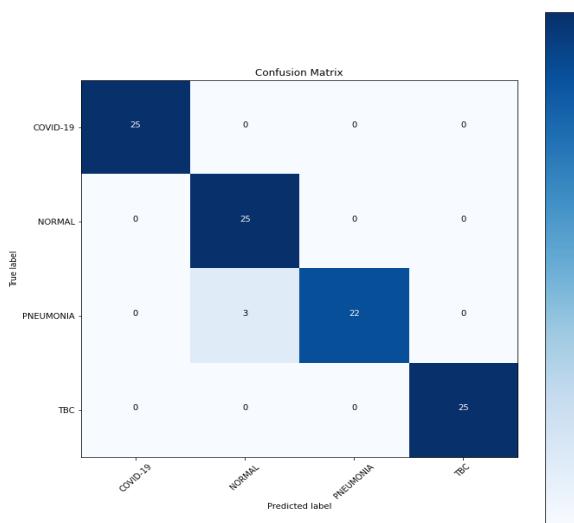
Berdasarkan evaluasi model tersebut didapatkan model optimal masing-masing model yaitu VGG19, ResNet50, dan InceptionV3. Model VGG19 menghasilkan nilai akurasi 0,9444 dan nilai Loss 0,1084 pada iterasi (*epoch*) terhenti pada 61 *epoch* karena tidak ada peningkatan dari 41 *epoch* dengan rasio 20% : 80%, ResNet50 menghasilkan akurasi optimal pada iterasi (*epoch*) terhenti pada 21 *epoch* karena tidak ada peningkatan dari 1 *epoch* dengan rasio 50% : 50% yaitu dengan nilai akurasi 0.2695 dan nilai Loss 3,1717. InceptionV3 menghasilkan akurasi optimal pada iterasi (*epoch*) terhenti pada 35 *epoch* karena tidak ada peningkatan dari 15 *epoch* dengan rasio 80% : 20% yaitu dengan nilai akurasi 0.6523 dan nilai Loss 1,1784.

4.5 Pengujian (*testing*)

Algoritma *Convolutional Neural Network* (CNN) membutuhkan proses *training* dan *testing*. Proses *testing* bertujuan menguji sebuah model yang dibentuk pada saat proses *training*. Pada pengujian (*testing*) ini dilakukan dengan menggunakan model terbaik saat proses pelatihan (*training*) yaitu VGG19 dengan nilai akurasi tertinggi berhenti di 41 iterasi (*epoch*) pada rasio 20% : 80%, ResNet50 dengan nilai akurasi tertinggi berhenti di 1 iterasi (*epoch*) pada rasio 50% : 50%, dan untuk model InceptionV3 dengan nilai akurasi tertinggi berhenti di 15 iterasi (*epoch*) pada rasio 80% : 20%.

4.5.1 Confusion Matrix

Confusion Matrix adalah matriks N x N yang digunakan untuk mengevaluasi kinerja model klasifikasi, di mana N adalah jumlah kelas target. *Confusion Matrix* membandingkan nilai target aktual dengan yang diprediksi oleh model *machine learning*. Ditampilkan dengan tabel yang menyatakan jumlah data uji yang benar diklasifikasikan dan jumlah data uji yang salah diklasifikasikan [9].



Gambar 8. Cofusion Matrix

Pada penelitian ini penulis menggunakan perintah *Classification report* pada python yang akan menampilkan nilai *Accuracy* dan *Loss* dari pengujian yang telah dilakukan. Berikut merupakan *classification report* Model VGG19 dengan iterasi (*epoch*) 41 pada rasio 20% : 80%, ResNet50 dengan iterasi (*epoch*) 1 pada rasio 50%: 50%, dan untuk model InceptionV3 dengan iterasi (*epoch*) 15 pada rasio 80% : 20%.

Tabel 6. *Classification Report*

Model	Accuracy	Loss
VGG19 (Rasio 20% : 80%, epoch 41)	0.96	0.0939
ResNet50 (Rasio 50%:50% , epoch 1)	0.25	1.6600
InceptionV3 (Rasio 50%:50% , epoch 15)	0.63	1.1077

Dari Tabel 7 dapat dilihat bahwa pengujian data menggunakan model VGG19 pada rasio 20% : 80% dengan iterasi (*epoch*) 41 menghasilkan akurasi sebesar 0.96 atau 96% dan Kehilangan 0,0939. Model ResNet50 pada rasio 50% : 50% dengan iterasi (*epoch*) 1 menghasilkan akurasi sebesar 0.25 atau 25% dan Kehilangan 1,660. Sementara itu model InceptionV3 pada rasio 80% : 20% dengan iterasi (*epoch*) 15 menghasilkan akurasi sebesar 0.63 atau 96% dan Kehilangan 1,1077.

4.5.2 Prediksi

Proses ini bertujuan untuk mengetahui label data baru yang belum diketahui berdasarkan pola gambar dengan menampilkan segmentasi dan persentasi akurasinya.pada penelitian ini dilakukan prediksi menggunakan 4 data baru menghasilkan *output* seperti pada Tabel 7.

Tabel 7. Hasil Prediksi VGG19

Label	Normal	COVID-19	Pneumonia	TBC
Akurasi (%)	98.97%	99.16%	99.56%	99,79%
Output				

Tabel 7 merupakan hasil prediksi dari model VGG19 dengan iterasi (*epoch*) 41 pada rasio 20% : 80%. Berdasarkan tabel tersebut menghasilkan prediksi seluruh label dengan benar. Dimana label normal diprediksi ke normal dengan akurasi 98,97%, label COVID-19 diprediksi ke COVID-19 dengan akurasi 99,16%, label *Pneumonia* diprediksi ke *Pneumonia* dengan akurasi 99,56%, sedangkan label TBC diprediksi ke TBC dengan akurasi 99,79.

5. KESIMPULAN

Berdasarkan hasil dari penelitian yang telah dilakukan dapat disimpulkan bahwa model VGG19 pada iterasi (*epoch*) 41 dari 61 epoch dengan rasio 20% : 80% menjadi model terbaik untuk penelitian ini. Dengan menghasilkan akurasi pelatihan mencapai 0,9444 atau 94,44% (dalam persentase) dan nilai Loss 0.1084. Pada pengujian menghasilkan *Confussion Matrix* dengan sebesar 0.96 atau 96% dan Kehilangan 0,0939, sedangkan prediksi menghasilkan prediksi keseluruhan label dengan benar dengan label NORMAL dengan nilai akurasi tertinggi yaitu 99,97%.

DAFTAR PUSTAKA

- [1] J. Pujoseno, “IMPLEMENTASI DEEP LEARNING MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK UNTUK KLASIFIKASI ALAT TULIS,” 2018.
- [2] M. F. Suryana, F. Fauziah, and R. T. K. Sari, “Implementasi Sistem Pakar Menggunakan Metode Certainty Factor Untuk Mendiagnosa Dini Corona Virus Disease (COVID-19),” *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 559, 2020, doi: 10.30865/mib.v4i3.2132.
- [3] R. Rizky, M. Ridwan, and Z. Hakim, “Implementasi Metode Forward Chaining Untuk Diagnosa Penyakit Covid 19 Di Rsud Berkah Pandeglang Banten,” *J. Teknol. Inf.*, vol. 4, no. 1, pp. 1–4, 2020.
- [4] V. Nigam, “Understanding Neural Networks. From neuron to RNN, CNN, and Deep Learning,” *medium.com*, 2018. <https://medium.com/analytics-vidhya/understanding-neural-networks-from-neuron-to-rnn-cnn-and-deep-learning-cd88e90e0a90>.
- [5] P. Mooney, “Chest X-Ray Images (Pneumonia),” 2018. <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>.
- [6] K. S. Mader, “Pulmonary Chest X-Ray Abnormalities,” 2018. <https://www.kaggle.com/kmader/pulmonary-chest-xray-abnormalities?select=Montgomery>.
- [7] T. Rahman, M. Chowdhury, and A. Khandakar, “COVID-19 Radiography Database,” 2020. <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>.
- [8] Bachir, “COVID-19 chest xray,” 2020. <https://www.kaggle.com/bachrr/covid-chest-xray>.
- [9] A. Bhandari, “Everything you Should Know about Confusion Matrix for Machine Learning,” *Analytics Vidhya*, 2020. <https://www.analyticsvidhya.com/blog/2020/04/confusion-matrix-machine-learning/>.
- [10] R. Teguh, A. S. Sahay, and F. F. Adjii, “Pemodelan Penyebaran Infeksi Covid-19 Di Kalimantan, 2020,” *J. Teknol. Inf. J. Keilmuan dan Apl. Bid. Tek. Inform.*, vol. 14, no. 2, pp. 171–178, 2020, doi: 10.47111/jti.v14i2.1229.