

## BAB III

### METODE PENELITIAN

Pada bab ini membahas mengenai *software* yang digunakan serta perancangan sistem analisis penyakit pneumonia menggunakan citra x-ray yang berisi diagram alir pengerjaan, diagram alir analisis performansi VGG-16, diagram alir *preprocessing*, diagram alir skenario pengujian, performansi sistem, dan evaluasi *confusion matrix* sebagai berikut

#### 3.1 *Software* yang digunakan

##### 3.1.1 Google Colabs

Google colab merupakan salah satu pengembangan kode dengan basis pemrograman python. Google colab dibuat oleh Google Research yang diciptakan untuk memungkinkan pengguna dapat menulis dan menjalankan bahasa pemrograman python pada browser dan google colabs menyediakan akses gratis ke sumber daya komputasi seperti GPU dan TPU. Pengguna google colab dapat menyimpan pemrograman pada google drive serta dapat melakukan kolaborasi dengan pengguna lainnya.



**Gambar 3.1 Google colaboratory.**

Gambar 3.1 merupakan logo *Google Colaboratory*, *Google Collaboratory* atau Google Colab merupakan salah satu tools yang berbasis cloud dan bebas biaya. Google Colab dibuat menggunakan *environment* jupyter dan mendukung banyak *library* yang diperlukan dalam pengembangan *Artificial Intelligence* (AI), google colab merupakan jupyter notebook yang dapat dijalankan secara online dan gratis. Pengguna tidak perlu menginstal *software* tambahan karena google colab telah dilengkapi dengan *tools* yang dibutuhkan untuk menulis dan menjalankan kode Python

### 3.1.2 Python

Python merupakan bahasa pemrograman yang dapat dibuat dan bekerja pada berbagai *platform*. Perbedaan python dengan bahasa pemrograman yang lain terletak pada penekanan pembacaan kode agar mudah memahami sintaks. Bahasa pemrograman python merupakan bahasa pemrograman yang mendukung paradigma *object oriented programming* atau *scripting* [23].



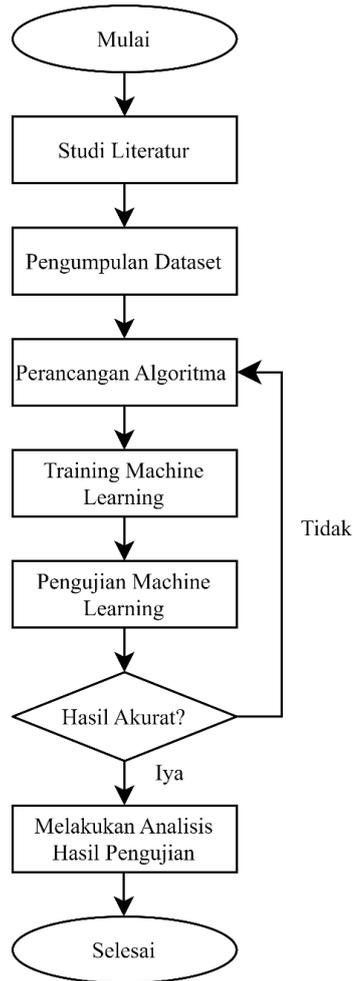
**Gambar 3.2 Python**[23]

Gambar 3.2 merupakan logo dari bahasa pemrograman python. Python akan diproses saat runtime oleh interpreter yang berarti pengguna tidak perlu mengkompilasi program sebelum dijalankan [24]. Bahasa pemrograman python digunakan untuk menggantikan bahasa pemrograman PERL sebagai aplikasi basis data GGI dikarenakan sintaksnya yang lebih mudah dipelajari dan pemrograman python berorientasi pada objek [25].

## 3.2 ALUR PENELITIAN

### 3.2.1 Diagram Alir Pengerjaan

Langkah-langkah dalam pengerjaan metode ini dijelaskan dalam diagram alir pada gambar 3.3. Tahapan pertama yang dilakukan adalah studi literatur untuk mendapatkan informasi yang mendukung pembuatan aplikasi ini, tahapan kedua yaitu melakukan pengumpulan dataset dan pengumpulan dataset berupa citra xray paru-paru normal dan pneumonia pada kaggle, tahapan ketiga yaitu melakukan perancangan algoritma image processing, algoritma yang digunakan adalah VGG-16 dengan 13 layer konvolusi dan 3 *fully connected layer* dan menggunakan 1000 citra x-ray pneumonia dan 1000 xray paru-paru normal sebagai data *train*. Tahapan selanjutnya, dalam melakukan tahapan *testing* digunakan 249 citra x-ray pneumonia dan 230 citra x-ray normal. Tahap keempat yaitu melakukan *training machine learning* menggunakan data x-ray yang dimiliki dan outputan berupa model. Pada tahapan kelima melakukan pengujian *machine learning* dengan melakukan perancangan *confusion matrix* untuk mengetahui akurasi seperti pada sub bab 3.2.4.

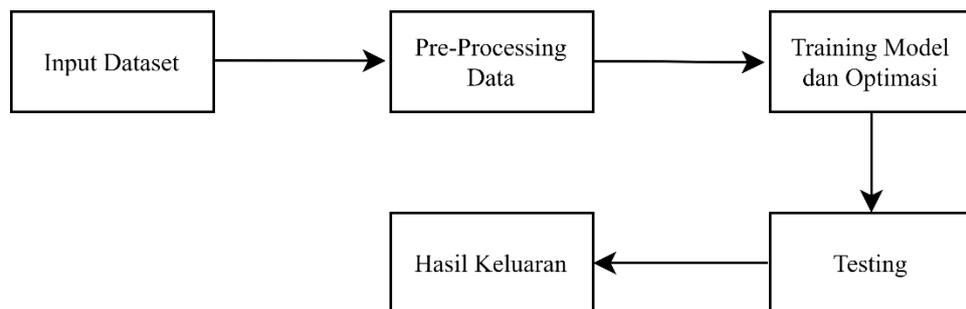


**Gambar 3.3 Diagram alir pengerjaan sistem**

Selanjutnya, setelah proses pengujian *machine learning* jika telah didapatkan hasil yang baik maka menuju tahapan analisis hasil pendeteksian, dan jika hasil akurasi belum sesuai maka kembali ke tahapan ketiga yaitu perancangan algoritma

### 3.2.2 Blok Diagram Sistem Analisa Performansi VGG-16

Pada gambar 3.4 dibawah ini menjelaskan mengenai alur sistem yang dibuat.

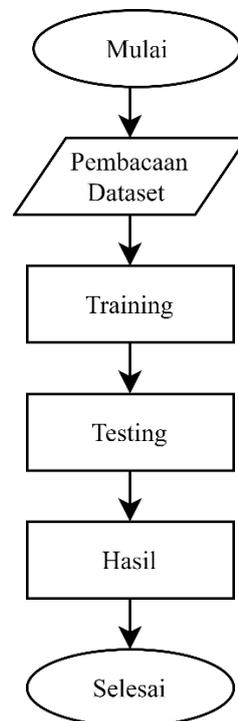


**Gambar 3.4 Blok diagram sistem analisa perdomansi vgg-16**

Sistem yang akan dirancang menggunakan input data untuk memasukan dataset yang digunakan dan berupa citra x-ray. Setelah data diinput terlebih dahulu dilakukan proses pre-processing seperti pada sub bab 3.2.3 dan dilanjutkan proses *training* model dan optimasi algoritma. Pada tahapan tersebut dilakukan *training* menggunakan algoritma VGG-16 dan optimasi dengan *loss functio*. Tahapan selanjutnya adalah melakukan testing menggunakan citra x-ray dan akan didapatkan hasil keluaran berupa hasil pendeteksian citra x-ray. Setelah mendapatkan hasil keluaran maka dapat dilakukan penghitungan *confusion matrix* untuk mengetahui tingkat hasil yang didapatkan.

### 3.2.3 Preprocessing

Dalam pengolahan citra terdapat proses preprocessing yang digunakan untuk optimalisasi data, tahapan *preprocessing* dapat dilihat pada gambar 3.5.



**Gambar 3.5 Diagram alir *preprocessing***

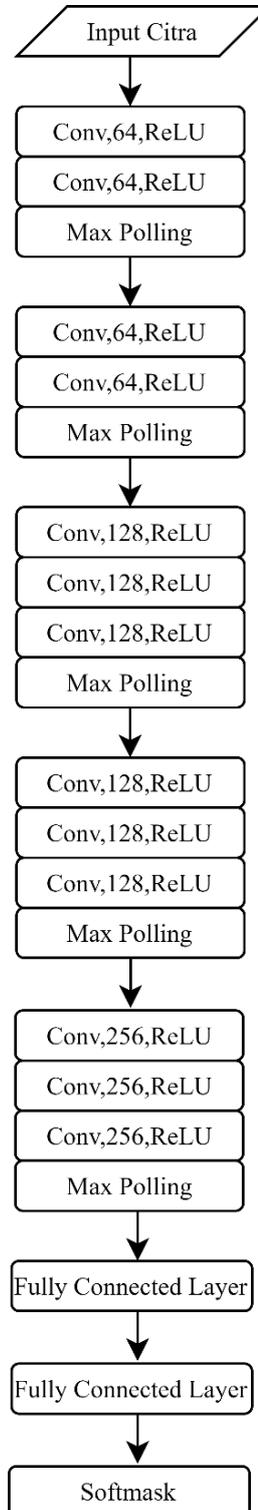
Berikut merupakan tahapan-tahapan pada *preprocessing* :

#### 1. Pembacaan Dataset

Proses pembacaan dataset adalah dengan melakukan penambahan dataset yang telah didapatkan kedalam sistem agar dapat dilanjutkan proses *training*. Pada proses pembacaan dataset ini dilakukan dengan membaca dataset yang akan digunakan dan telah diupload pada github untuk mempermudah dalam mengakses dataset tersebut.

## 2. Training

Pada pemrosesan data citra x-ray menggunakan algoritma VGG-16 memiliki beberapa tahapan yang dijelaskan pada diagram alir pada gambar 3.6. Secara umum VGG-16 memiliki enam garis besar.



**Gambar 3.6 Model pelatihan VGG-16.**

Tahapan pertama, *Input* citra merupakan tahapan melakukan penambahan data citra kedalam CNN. Kedua, *convolution layer* akan membentuk sebuah filter yang memiliki panjang dan tinggi (*pixels*). Tahapan ketiga, ReLu akan menerapkan fungsi aktivasi terhadap nilai dari hasil konvolusi, pada aktivasi di ReLu nilai *output* dari neuron dapat dinyatakan sebagai 0. Jika inputan nya bernilai negatif dan jika fungsi aktivasi bernilai positif, maka *output* dari neuron merupakan nilai input aktivasi itu sendiri. Keempat, tahap pada *max polling layer*, pada tahapan ini berfungsi untuk mengurangi dimensi dari *featur map* (*downsampling*) dan dapat mempercepat komputasi. Kelima, tahap *fully connected layer*. Pada tahap ini semua neuron aktivasi dari lapisan sebelumnya akan terhubung dengan lapisan selanjutnya sehingga bertujuan untuk mengolah data sehingga dapat diklasifikasikan. Tahapan terakhir melakukan tahapan aktivasi *softmax* yang berfungsi untuk menghitung probabilitas pada output, pada output layer akan diambil nilai dengan probabilitas tertinggi sebagai prediksi.

Diagram alir diatas merupakan alur dari arsitektur VGG-16, VGG-16 memiliki 16 layer yang terdiri dari 13 *convolution layer* dan 3 *fully connecteed*. Setelah ditambahkan data citra x-ray pada tahap pertama dilakukan *covolution* dengan 64 pixel dan fungsi aktivasi sebanyak dua kali lalu dilakukan *max polling layer*, juga *pooling layer*, selanjutnya pada tahap kedua dilakukan *convolution* dengan 128 pixel dan fungsi aktivasi sebanyak dua kali dan dilakukan *max pooling*, pada tahapan ketiga dilakukan pengolahan *convolution* dengan 256 pixel dan ditambahkan fungsi aktivasi sebanyak tiga kali lalu dilakukan *max pooling*, tahapan keempat dan kelima yaitu melakukan *convolution* 512 pixel dan fungsi aktivasi sebanyak tiga kali lalu dilakukan *max polling*, pada tahapan keenam, ketujuh, dan kedelapan dilakukan *fully connected layer*, dan tahapan terakhir yaitu dilakukan *softmax*.

### 3. Testing

Proses *testing* merupakan proses klasifikasi menggunakan nilai dari hasil *training*. Pada proses ini dilakukan pendeteksian citra x-ray menggunakan

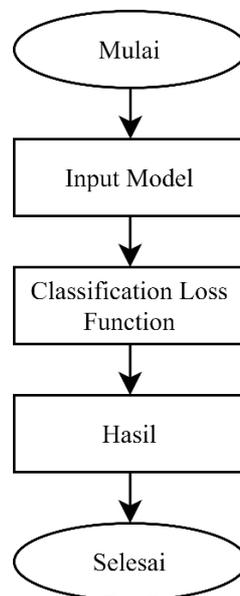
model yang telah dibuat. Proses ini menghasilkan akurasi dari klasifikasi yang dilakukan, data yang gagal diklasifikasikan.

#### 4. Hasil

Hasil yang didapatkan dari proses preprocessing adalah tingkat akurasi dari proses pengklasifikasian.

### 3.2.4 Skenario Pengujian

Pada penyusunan skripsi ini dilakukan skenario pengujian pada algoritma yang diterapkan menggunakan *loss function* dan *confusion matrix*. Pertama, *lost function* merupakan fungsi yang berguna untuk menghitung seberapa jauh model menyimpang dari prediksi yang benar, *loss function* juga menghitung seberapa akurat algoritma yang digunakan. Pada umumnya *lost function* digunakan pada tahapan optimalisasi dimana *loss function* membutuhkan model sebagai objek yang dihitung. Terdapat tiga jenis *lost function* yaitu *regression loss fungtion*, *classification loss function*, dan *ranking loss function*.



**Gambar 3.7 Diagram alir analisis performansi.**

Pada gambar 3.7 diatas merupakan alur analisis perfomansi yang dilakukan. Langkah pertama adalah menambahkan model yang akan dilakukan analisa, selanjutnya dilakukan analisa menggunakan *classification loss function* dan akan didapatkan hasil keluaran nilai dari perfomansi yang dianalisis.

### 3.2.5 Perfomansi Sistem

Tahapan perfomansi sistem merupakan proses yang dilakukan setelah melakukan proses *training* dan *testing* dataset yang telah dilakukan. Perfomansi

sistem memiliki fungsi untuk mengukur tingkat perfromansi dari klasifikasi terhadap model pelatihan. Terdapat lima indikator untuk menentukan perfromansi suatu sistem, yaitu tingkat akurasi, *loss*, presisi, *recall*, dan *f1-score*. Perhitungan perfromansi sistem akan dibantu menggunakan *confusion matrix*.

### 3.2.6 Evaluasi *Confusion Matrix*

Pada bagian ini akan dibahas mengenai metode evaluasi menggunakan *confusion matrix*. Setelah melatih model *CNN* pada dataset gambar *pneumonia*, dilakukan evaluasi kinerja model menggunakan data *test*. Evaluasi dilakukan dengan menghitung akurasi dan membangun *classification report*. *confusion matrix* memberikan gambaran tentang seberapa baik model mampu memprediksi kelas target. Untuk menentukan 5 indikator performansi sistem dengan dicari menggunakan *confusion matrix* yang memiliki nilai parameter *True Positif* (TP), *False Positif* (FP), *True Negatif* (TN), dan *False Negatif* (FN).

**Tabel 3.1 *Confusion matrix*.**

Aktual	Prediksi	
	Pneumonia	Normal
Pneumonia	TP	FN
Normal	FP	TN

Dari tabel 3.1 diatas akan dijelaskan seperti berikut :

1. *True Positive* (TP) adalah jumlah x-ray pneumonia diverivikasi benar x-ray pneumonia.
2. *True Negative* (TN) adalah jumlah x-ray normal diverivikasi benar x-ray normal.
3. *False Normal* (FN) adalah jumlah x-ray pneumonia diverivikasi benar x-ray normal.
4. *False Positive* (FP) adalah jumlah x-ray normal diverivikasi benar x-ray pneumonia.