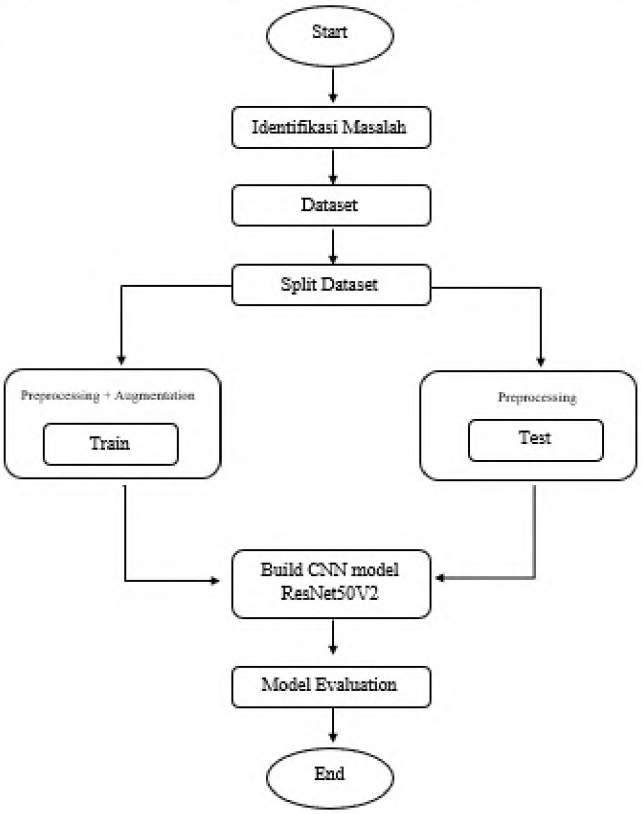
BAB III
METODE PENELITIAN

Pada penelitian ini topik yang dibahas adalah pemrosesan gambar untuk dapat digunakan dalam melakukan klasfikasi penyakit pneumonia dari dataset yang tersedia dengan memanfaatkan arsitektur model ResNet50V2 sebagai base model. Pada Gambar dijabarkan bahwa terdapat beberapa tahapan yang digunakan pada penelitian ini agar mampu mendapatkan hasil yang diinginkan. Proses pertama yang dilakukan adalah identifikasi masalah yang akan dibahas pada penelitian ini yaitu klasifikasi citra. Selanjutnya tahap pengumpulan dataset yang dilakukan pembagian ke dalam dua kelas yaitu *train* dan *test* dimana akan dilakukan proses *preprocessing* dan *data augmentation*. Proses selanjutnya yaitu membangun model dengan menggunkana *pretrained* model yang dipilih. Langkah terakhir dalam penelitian ini adalah *Model Evaluation* yang digunakan untuk mengetahui bagaimana performa dari model yang telah dirancang pada penelitian ini terhadap data yang telah disediakan untuk melakukan klasifikasi penyakit pneuimonia.



Gambar 3. 1 Flowchart Perancangan Penelitian

3.1 Identifikasi Masalah

Berdasarkan beberapa jurnal pendukung, dikatakan bahwa jumlah tenaga medis untuk dapat melakukan diagnosis penyakit pneumonia saat ini sangat terbatas, dan biaya yang diperlukan cukup besar. Berdasarkan masalah tersebut, peneliti ingin membuat suatu sistem yang dapat melakukan klasifikasi penyakit pneumonia berdasarkan citra x-ray sehingga mampu membatu tenaga medis dalam mendiagnosis penyakit pneumonia. Dalam jurnal rujukan utama telah dibuat sebuah model untuk dapat melakukan klasifikasi pneumonia dengan membandingkan empat model pre-*train*ed dan satu arsitektur model buatan sendiri yang mendapatkan hasil akurasi sebesar 90%. Dari hasil dari penelitian sebelumnya perlu untuk melakukan analisa sehingga mampu menaikkan angka akurasi menjadi lebih baik.

3.2 Lingkungan Pengembangan

Proses pembangunan dan pelatihan model pada penelitian ini menggunakan Google Colab dengan memanfaatkan *framework deep learning* pada Tensorflow dan Keras. Untuk perangkat keras yang digunakan pada penelitian ini menggunkaan laptop dengan processor Intel Core i5 dengan memori sebesar 12 GB.

3.3 Dataset

Dataset yang digunakan pada penelitian ini adalah data gambar x-ray dada dengan nama Chest X-Ray Images (Pneumonia) yang bersumber dari Guangzhou Women and Children's Medical Center, Guangzhou. Dataset pada penelitian ini dapat diakses pada https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia. Total gambar yang ada dalam dataset berjumlah 5856 gambar yang terdiri dari dua kategori yaitu normal dan pneumonia, untuk kategori normal terdiri dari 1583 gambar x-ray dada dan untuk kategori pneumonia terdiri dari 4273 gambar x-ray dada. Untuk sampel citra yang digunakan pada penelitian ini dapat dilihat pada Gambar 3.2 yang merupakan contoh sampel citra x-ray dada normal dan pada Gambar 3.3 merupakan contoh sampel citra x-ray dada pneumonia.

Gambar 3. 2 Citra X-Ray Normal

Gambar 3. 3 Citra X-Ray Pneumonia

3.4 Split Dataset

Pada penelitian ini dataset yang digunakan sama dengan penelitian terdahulu [13][14]. Dataset terdiri dari 5856 citra x-ray akan dibagi ke dalam tiga bagian, yaitu data *train* dengan jumlah citra x-ray sebanyak 5216, data *test* dengan jumlah 624 citra x-ray, dan data validation dengan jumlah 16 citra x-ray. Namun, merujuk pada penelitian [13] dataset dibagi kedalam dua bagian, terdapat proses penggabungan antara data *train* dan data validation sehingga jumlah pembagian data yang digunakan pada penelitian ini sebanyak 5232 citra x-ray untuk data *train*, dan 624 citra x-ray untuk data *test*. Data *train* adalah kumpulan data yang nantiya akan digunakan untuk melatih model agar dapat mengenali penyakit pneumonia dari citra x-ray. Sedangkan untuk data *test* adalah kumpulan data yang digunakan untuk proses optimasi pada saat pelatihan model.

3.5 Data Preprocessing dan Augmentation

Dalam penelitian ini terdapat proses pre-processing dengan melakukan perubahan ukuran gambar menjadi 224 x 224 pixel. Serta terdapat proses augmentasi data yang berguna untuk menambah variasi dari data gambar yang tersedia, data augmentasi merupakan salah satu teknik yang digunakan untuk mengatasi overfitting pada saat pelatihan model [31]. Pada penelitian ini dilakukan beberapa teknik augmentasi yang akan dijabarkan pada Tabel 3.1

Tabel 3. 1 Augmentasi yang akan dilakukan pada data

No	Proses	Keterangan
1	rescale	Parameter untuk mengubah nilai citra
2	zoom_range	Parameter untuk perubahan perbesaran citra
3	horizontal_flip	Parameter untuk melakukan proses pembalikan citra secara horizontal
4	vertical_flip	Parameter untuk melakukan proses pembalikan citra secara vertikal
5	brightnes_range	Parameter untuk melakukan proses perubahan tingkat kecerahan dari citra
6	rotation_range	Parameter untuk menerapkan rotasi citra

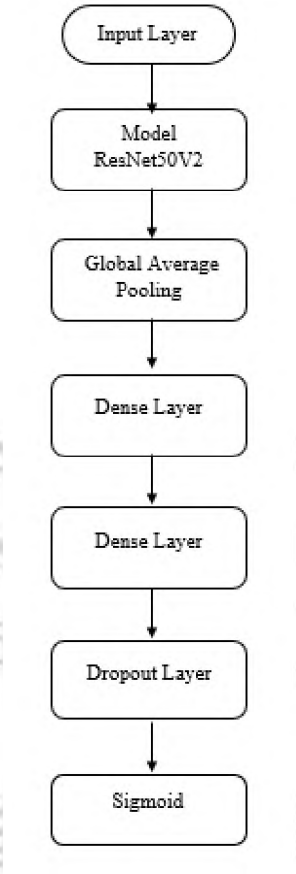
Pada proses augmentasi ini menggunakan ImageDataGenerator dari Keras, dimana pada saat proses *training* akan dilakukan proses augmentasi gambar pada setiap *epochs*. Proses augmentasi tesebut akan menghasilkan 5232 data citra baru yang berbeda – beda pada setiap *epochs* yang dilakukan. Hal tersebut dapat mengurangi penggunaan memory pada saat *train*ing dan dapat meningkatkan variasi dari data sehingga mampu mengurangi terjadinya *overfitting* dan menambah hasil akurasi dari model latih.

3.6 Build CNN Model ResNet50V2

Pada penelitian ini dilakukan pembuatan model *Convolution Neural Network* dengan memanfaatkan *pre-trained* model ResNet50V2. Pembuatan model pada penelitian ini yaitu *input layer* dengan ukuran yang di sarankan sebesar 224x224 pixel, dilanjutkan dengan arsitektur model dari ResNet50V2 dengan bobot *'imagenet'* yang digunakan sebagai *base model.* Layer selajutnya terdapat layer *Global Average Pooling* yang bertujuan untuk dapat mengurangi terjadinya overfitting saat melakukan pelatihan model. Untuk *Fully Connected Layer* terdapat *dense layer* serta *Dropout Layer*. Untuk *output layer* pada penelitian ini menggunakan 1 neuron dengan aktivasi *sigmoid*. Pelatiham model dilakukan dengan *epochs* sebanyak 100 yang dibagi kedalam dua tahap.

a. Skenario Pengujian Pertama pada arsitektur model menggunakan optimizer adam
b. Skenario Pengujian Pertama pada arsitektur model menggunakan optimizer adadelta

Setelah pembuatan arsitektur dilakukan proses pelatihan model dengan Berikut adalah model arsitektur ResNet50V2 yang diusulkan pada penelitian ini yang dapat dilihat pada Gambar 3.4.



Gambar 3. 4 Rancangan Arsitektur Model

3.7 Model Evaluation

Model yang telah dirancang dan telah melalui proses pelatihan model akan dilakukan proses evaluasi. Pada penelitian ini hasil evaluasi digunakan untuk melakuakn pengujian performa model yang telah melalui proses pelatihan dengan menggunakan grafik akurasi dan *loss*, *Classification Report*, dan *Confusion Matrix*. Informasi yang didapatkan pada classification report dapat digunakan untuk mengetahu hasil *accuracy, precision, recall, dan f1-score*. Pengujian dataset pada *Confusion Matrix* dapat memberikan informasi seberapa banyak data yang dapat diklasifikasikan dengan benar dan salah pada matrix uji coba.