

## **BAB III**

### **METODOLOGI PENELITIAN**

Pada bab ini akan dijelaskan secara menyeluruh mengenai metodologi yang digunakan untuk melakukan penelitian pengenalan suara menggunakan metode *deep learning*.

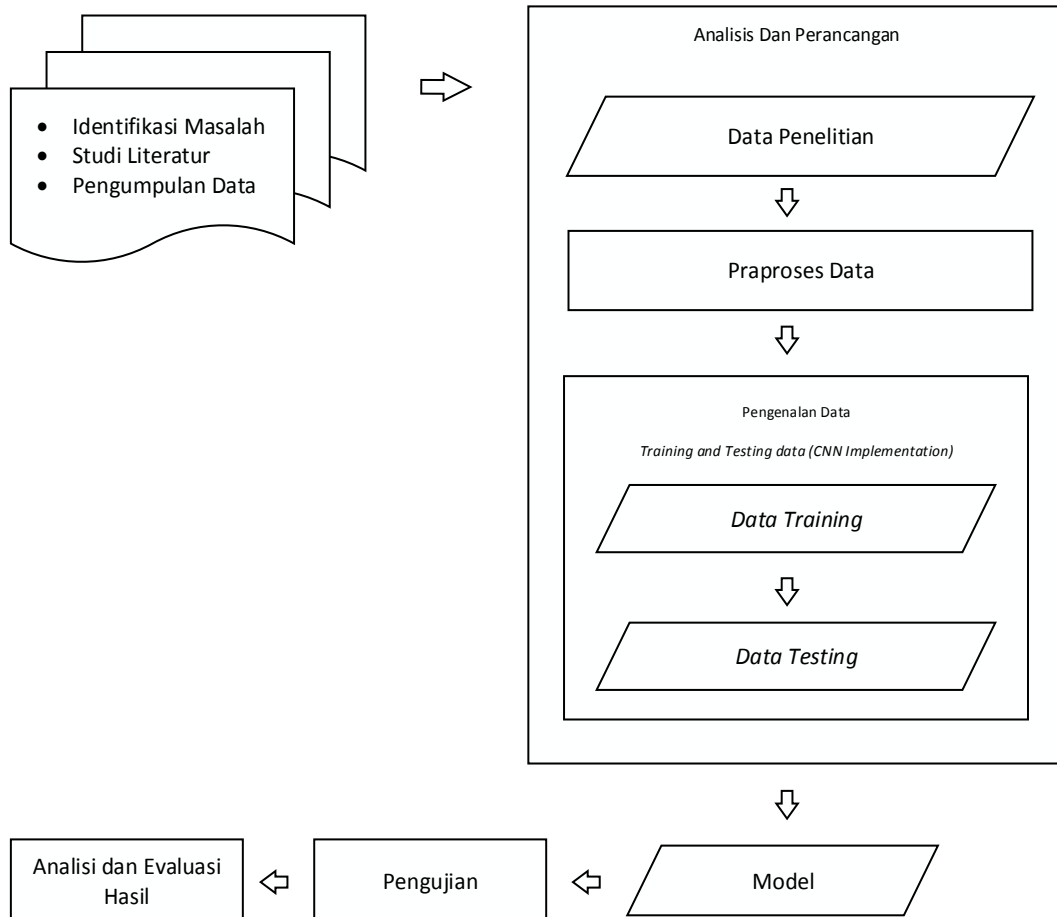
#### **3.1 Desain Penelitian**

Desain penelitian adalah tahapan atau gambaran yang akan dilakukan dalam penelitian. Desain penelitian ini dibuat untuk memberikan kemudahan dalam melakukan penelitian.

Pada penelitian ini, desain penelitian yang akan dilakukan meliputi langkah-langkah sebagai berikut :

1. Identifikasi masalah yaitu merumuskan masalah dan dengan cara apakah masalah tersebut akan diselesaikan.
2. Studi literatur yaitu pengumpulan informasi terkait dengan penelitian yang didapatkan melalui buku, data kuliah, dan internet.
3. Analisis dan perancangan, yaitu langkah menganalisa kebutuhan sistem secara intensif dan spesifik. Untuk memahami sifat sistem yang akan dibangun, pembuat sistem (analisis) harus memahami domain informasi untuk sistem, fungsi yang dibutuhkan, perilaku, kinerja dan antarmuka sistem.

Adapun desain penelitian yang dipergunakan dalam penelitian ini digambarkan dengan skema seperti pada Gambar 3.1 berikut.



**Gambar 3.1 Desain Peneliti**

Gambar 3.1 menunjukkan tahapan dalam menyelesaikan penelitian ini. Berikut merupakan penjelasan dari masing-masing tahapannya.

### 1.1.1 Identifikasi Masalah

Identifikasi masalah merupakan tahapan pertama pada suatu penelitian. Tahapan ini adalah merumuskan masalah dan menentukan bagaimana tahapan-tahapan selanjutnya dapat dilakukan dengan cara apakah masalah tersebut akan diselesaikan.

### 1.1.2 Studi Literatur

Studi literatur adalah dengan mencari referensi terkait dengan penelitian yang dilakukan. Pada penelitian ini yaitu mengenai suara, pengolahan suara, spektrogram, *deep learning*, kecerdasan buatan dan jaringan syaraf tiruan serta pembahasan tentang penelitian-penelitian sebelumnya yang berkaitan. Penjelasan literatur dan teori-teori yang mendukung penelitian ini dapat dilihat pada BAB II.

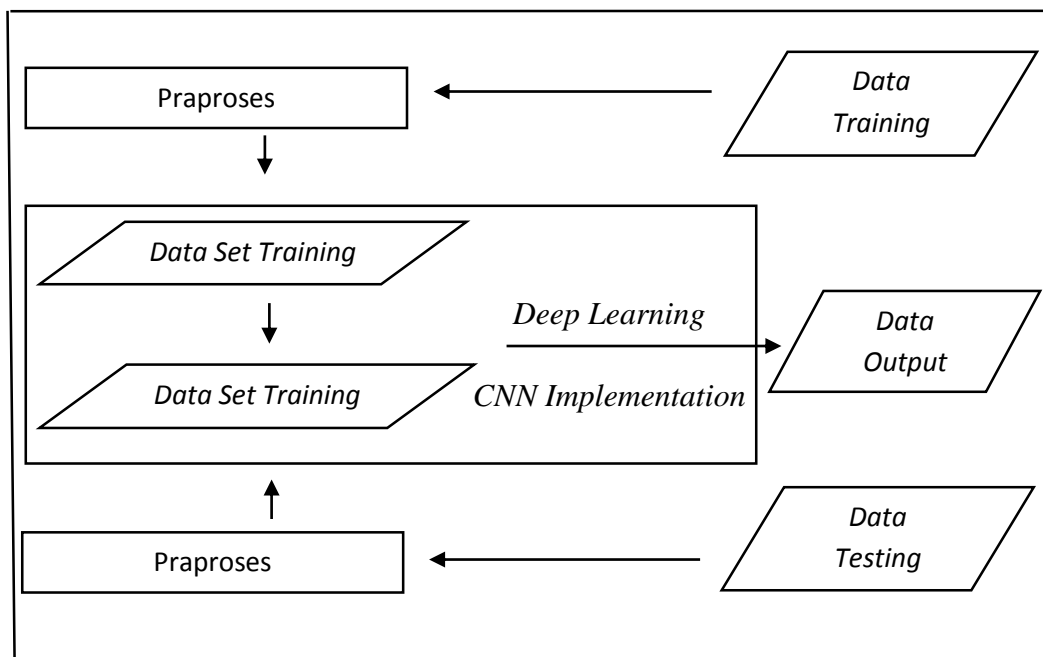
### 1.1.3 Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan dengan merekam suara menggunakan *smartphone* terhadap beberapa orang untuk dijadikan model pengenalan suara. Beberapa orang tersebut adalah 10 mahasiswa Universitas Pendidikan Indonesia, yaitu penulis dan bersama sembilan orang lainnya.

Adapun rekaman suara yang dilakukan adalah merekam suara ucapan berbahasa Indonesia setiap orang secara individu dari pengucapan 10 kata yang berupa kata satu sampai dengan sepuluh dan secara perkatanya dilakukan dengan 3 teknik pengucapan yaitu pelan, normal dan cepat. Untuk pelaksanaannya diatur pada kondisi lingkungan yang tenang supaya data rekaman suara memiliki hasil yang bagus. Jadi total data yang akan terkumpul adalah 300 data rekaman dari 10 orang (responden).

#### 4.1.1 Analisis dan Perancangan

Analisis dan perancangan adalah langkah untuk memahami sifat sistem yang akan dibangun dengan menganalisa kebutuhan sistem secara intensif dan spesifik. Pada tahapan ini pembuat sistem (analisis) dapat memahami fungsi yang dibutuhkan, perilaku, kinerja dan antarmuka dari sistem. Gambar 3.3 berikut ini menunjukkan rancangan alur sistem yang dibuat.



**Gambar 3.3 Analisis dan Perancangan**

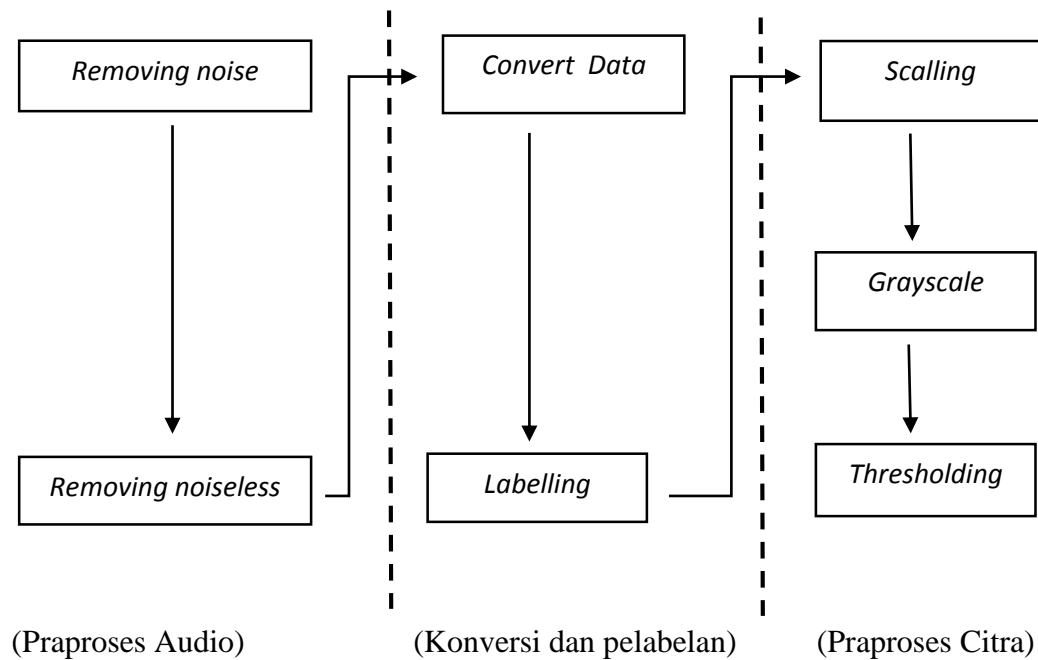
##### 1. Data *training* dan *testing*

Data penelitian yang telah dikumpulkan akan diatur menjadi data *training* dan data *testing*.

Data *training* merupakan data yang digunakan untuk membuat sistem mempelajari fitur-fitur yang ada pada data dan melakukan klasifikasi terhadap fitur-fitur tersebut. Sedangkan data *testing* merupakan data yang digunakan untuk menguji akurasi hasil klasifikasi fitur-fitur data *training* yang telah dipelajari oleh sistem.

## 2. Praproses

Praproses adalah proses menjadikan data lebih tepat digunakan. Pada penelitian ini, praproses dilakukan untuk mendapatkan data yang hanya berdurasi suara yang digunakan menjadi dalam format data gambar (citra). Alur praproses ini dibuat sebagai mana ditunjukkan pada Gambar 3.4 berikut:



**Gambar 3.4 Tahapan Praproses**

### A. Praproses Audio

Pada tahapan ini adalah akan dilakukan pemotongan durasi data rekaman, yaitu pemotongan dengan penghapusan durasi suara gangguan (*noise*) dan durasi suara diam (*noiseless*) yang merupakan bagian dari data yang ikut terekam tetapi tidak diperlukan. Hal ini dilakukan untuk mendapatkan data suara yang tepat.

### B. Konversi dan Pelabelan

Tahapan selanjutnya, data akan dikonversi dari bentuk *audio* menjadi bentuk gambar (citra).

Konversi tersebut yaitu dari data .mp3 menjadi .wav agar menjadikan data lebih detail, dan dari .wav menjadi *spectrogram image* untuk sebagai teknik yang dipergunakan dalam mengenali data suara. *Spectrogram image* merupakan bentuk dari grafik frekuensi (*spektrum*) yang di plot terhadap waktu dan amplitudo. Data tersebut selanjutnya diberi label untuk merepresentasikan perbedaan setiap data suara. Label ini juga sebagai indikator dalam pengujian yang dilakukan.

### C. Praproses Citra

Setelah itu, dalam bentuk *spectrogram image* akan dilakukan praproses citra dengan tujuan mendapat karakteristik data untuk proses pengenalan.

Praproses ini adalah *scaling*, yaitu dengan merubah seluruh ukuran data gambar menjadi 28x28 piksel (Deng, The MNIST Database of Handwritten, 2012). *Grayscale*, yaitu menjadikan citra berwarna (citra tiga dimensi), yaitu panjang, lebar dan warna dapat menjadi dua dimensi dengan cara membuang warna yang terdapat pada citra, sehingga membuat citra hanya memiliki warna hitam dan putih. Dan *thresholding*, yaitu memisahkan antara objek dan latar belakangnya yang terdapat pada citra. Hal ini dilakukan dengan mengubah citra ke dalam bentuk biner untuk menyederhanakan data agar mempercepat proses pada sistem dalam mengenali fitur pada data yang akan dideteksinya.

### 3. Implementasi *Deep learning*

Pada tahap ini, diimplementasikan model *Convolutional Neural Network* (CNN) menggunakan *Deep Learning Toolbox* untuk mempelajari data *input*, mengklasifikasikannya dan menghitung tingkat akurasi hasil klasifikasinya.. Tahapan pertama adalah menentukan jumlah *hidden layer* sesuai kebutuhan. Pada penelitian ini menggunakan empat *hidden layer* CNN, yaitu dua *layer* konvolusi dan dua *layer subsampling*. Berikut adalah penjelasan mengenai tahapan-tahapan ekstraksi ciri dan klasifikasi pada metode *deep learning* model *Convolution Neural Network*.

### A. *Convolution Layer*

Menghitung *convolution layer* menggunakan teknik konvolusi seperti yang dijelaskan pada bab II, yaitu menentukan jumlah *output map*. *Output map* adalah jumlah *neuron* di tiap *convolution layer*. Jumlah *output map* pada *layer* konvolusi yang pertama adalah 6 dan kedua 12. Selanjutnya menentukan jumlah kernel yang digunakan pada 2 *layer* konvolusi yaitu ukuran 5x5 piksel (Deng, The MNIST Database of Handwritten, 2012).

### B. *Subsampling Layer*

*Subsampling* merupakan pemotongan ukuran dari sebuah data citra. Proses ini dilakukan untuk memperoleh peningkatan invariasi posisi dari sebuah fitur. Pada penelitian ini *subsampling layer* menggunakan proses perhitungan *max pooling* dengan skala matriks 2x2. Dimana proses dari *max pooling* adalah membagi *output* yang dihasilkan dari proses *convolution layer* menjadi beberapa *grid* kecil.

### C. Jumlah *Epoch* dan *Batch*

Jumlah *epoch* yang dicobakan bervariasi untuk diketahui tingkat akurasi pada masing-masing *epoch*, percobaanya dimulai dari 1000 dan 10000. Sedangkan jumlah *batch* yang digunakan adalah sejumlah data *training*.

### D. Proses *Training* dan *Testing*

Sebelum melakukan klasifikasi, dilakukan *training* terlebih dahulu dengan menggunakan parameter hasil *training* pada CNN. Setelah itu barulah dilakukan proses *testing* untuk melihat hasil klasifikasi data citra suara yang dijadikan sebagai data *input*. Pada *Toolbox* yang digunakan, proses *training* menggunakan jaringan *backpropagation*, dan proses *testing* menggunakan jaringan *feed forward*. Nilai awal *kernel* pada proses *training* adalah *random*, lalu nilai *kernel* tersebut diperbaharui setiap melakukan *training* untuk menentukan pola klasifikasi. Sedangkan pada proses *testing* menggunakan nilai *kernel* yang telah melalui proses *training*.

#### **4.1.2 Model**

Model merupakan hasil yang didapat dari mempelajari fitur-fitur pada data untuk selanjutnya digunakan pada pengujian.

#### **4.1.3 Pengujian**

Pengujian adalah tahap uji coba untuk mengetahui apakah sistem atau model yang dibangun sudah bekerja dengan benar atau belum.

#### **4.1.4 Analisis dan Evaluasi**

Analisis dan Evaluasi adalah dilakukannya analisis dan evaluasi hasil untuk mengetahui kemampuan dan keakurasian serta evaluasi terhadap sistem yang telah dibangun.

### **3.2 Data Penelitian**

Data penelitian yang digunakan yaitu data rekaman suara yang direkam menggunakan *smartphone*. Data ini kemudian di praproses terlebih dahulu dan kemudian menjadi *spectrogram image* dalam bentuk *thresholding* dengan ukuran 28x28 pixel untuk dapat digunakan pada sistem pengenalan suara yang dibangun.

#### **3.1.1 Data Input**

Terdapat dua macam data *input* yang digunakan. Data pertama merupakan *file* variabel Matlab dengan format *MAT file* yang digunakan ketika pelatihan dan pengujian sistem. Sedangkan data *input* kedua adalah dengan format *image* yang digunakan pada sistem yang telah dibangun tampilan antarmukanya.

#### **3.1.2 Data Output**

Data *output* penelitian ini adalah klasifikasi data dari data *input*. Hasil klasifikasi tersebut ditampilkan dalam bentuk teks.



### 3.3 Spesifikasi *Hardware* dan *Software*

Untuk menunjang penelitian yang dilakukan ini, diperlukanlah spesifikasi *hardware* dan *software* untuk dapat mengimplementasikannya. Maka ditentukan bahwa *hardware* dan *software* yang digunakan adalah sebagai berikut:

1. *Smartphone* MI4C dengan aplikasi *record*.
2. Laptop dengan prosesor Intel Core i7 4500 2.80 GHz dan RAM 4 GB
3. Sistem operasi Microsoft Windows 10 64 bit
4. Aplikasi Matlab 2017a