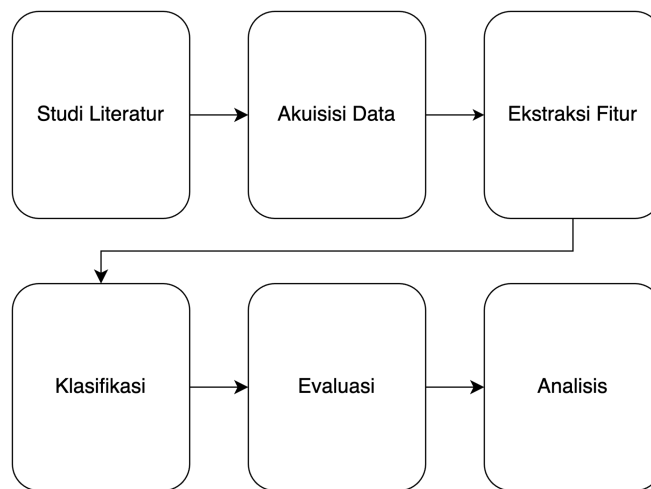


BAB III

ANALISIS DAN PERANCANGAN SISTEM

3.1 Metode Penelitian

Pada penelitian ini dilakukan dengan tahapan untuk meneliti dan analisis yang bertujuan untuk mengetahui masalah dan menjawab hasilnya, dengan ini maka alur metode nya yang digunakan terdiri dari studi literatur, akuisisi data, pengolahan data, ekstraksi fitur, klasifikasi, evaluasi, dan analisis. Untuk alur yang dilakukan dengan arah ada di Gambar 3.1, metode dan alur penelitian.



Gambar 3.1 Metode & Alur Penelitian

3.1.1 Studi Literatur

Penelitian ini berfokus pada deteksi dini gejala depresi melalui analisis sinyal audio dengan menggunakan teknologi machine learning. Studi literatur ini menyediakan konteks dan pemahaman mendalam tentang topik yang akan dibahas, mengacu pada penelitian yang sudah ada dan mengembangkannya dengan menggunakan metode analisis audio dan Convolutional Neural Network (CNN) untuk klasifikasi. Penulis sudah diskusi dengan dosen di program studi teknik komputer untuk penelitian ini. Sumber untuk studi literatur sudah mencakup keseluruhan untuk men analisis dan penulis sudah mempelajari bagaimana cara mengolah data dimana akan dilakukan untuk penelitian ini.

3.1.2 Akuisisi Data

Penelitian ini menggunakan kumpulan data yang tersedia untuk umum MODMA, yang merupakan singkatan dari Multi-modal Open Dataset for Mental-disorder Analysis. Kumpulan data ini dikumpulkan oleh Universitas Lanzhou di China. Pasien dengan gangguan depresi mayor (MDD) direkrut di antara pasien rawat inap dan rawat jalan dari Rumah Sakit Kedua Universitas Lanzhou, Gansu, Tiongkok, didiagnosis dan direkomendasikan oleh setidaknya satu psikiater klinis. *Normal control* (NC) direkrut melalui poster. Penelitian ini disetujui oleh Komite Etika Rumah Sakit Afiliasi Kedua Universitas Lanzhou, dan persetujuan tertulis diperoleh dari semua subjek sebelum percobaan dimulai. Semua pasien MDD menerima *Mini-International Neuropsychiatric Interview* (MINI) terstruktur yang memenuhi kriteria diagnostik untuk depresi berat dari Manual Diagnostik dan Statistik Gangguan Mental (DSM) berdasarkan DSM-IV. Psikiater menggunakan MMSE, sebuah kuesioner 30 poin, untuk memilih peserta yang mengalami depresi. Tujuan dari kuesioner ini adalah membantu psikiater menentukan sejauh mana gangguan kognitif pada pasien dengan penyakit otak yang menyebar melalui wawancara langsung. Selain itu, untuk cross-referencing, semua peserta diminta untuk menyelesaikan skala berikut.

- Patient Health Questionnaire (PHQ-9) adalah alat serbaguna yang terdiri dari sembilan pertanyaan, digunakan untuk mendiagnosis, menyaring, melacak, dan mengukur tingkat keparahan depresi. Kuesioner ini dipilih untuk menyelidiki hubungan antara fitur EEG dan tingkat depresi. Rata-rata skor PHQ-9 yang diperoleh adalah 9.6.
- Life Event Scale (LES) memiliki 48 pertanyaan yang mencakup peristiwa terkait keluarga, pekerjaan, dan dukungan sosial. Dampak dari setiap peristiwa dinilai berdasarkan frekuensi, durasi, dan intensitasnya. Kuesioner ini dipilih untuk cross-referencing. Rata-rata skor LES adalah -18.9.

- Pittsburgh Sleep Quality Index (PSQI) memiliki tujuh komponen untuk mendeteksi masalah tidur, terdiri dari 19 item laporan diri. Indikator ini dipilih untuk menyelidiki hubungan langsung antara karakteristik tidur dan depresi EEG. Rata-rata skor PSQI adalah 7.5.
- Generalized Anxiety Disorder Scale-7 (GAD-7) hanya terdiri dari tujuh pertanyaan laporan diri untuk tujuan menyaring dan mengukur tingkat keparahan gangguan kecemasan umum. Kuesioner ini digunakan untuk cross-referencing kecemasan dan kesedihan. Hasilnya menunjukkan rata-rata skor GAD-7 sebesar 7.
- Rata-rata skor CTQ-SF yang teridentifikasi adalah 45.4.
- Rata-rata skor SSRS adalah 37.9.

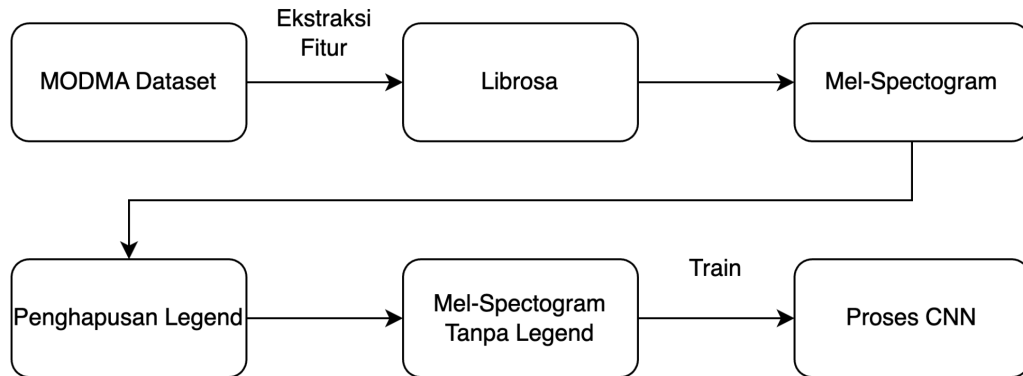
Dari beberapa kuesioner yang diberikan kepada peserta, kuesioner yang digunakan untuk mengkategorikan tingkat depresi adalah kuesioner PHQ-9. Rata-rata skor PHQ-9 untuk semua pasien yang mengalami depresi adalah 9.6 atau dibulatkan menjadi 10.

3.1.3 Ekstraksi Fitur

Dalam penelitian ini, ekstraksi fitur dari data audio melibatkan penggunaan Librosa, sebuah pustaka Python yang digunakan untuk analisis dan pemrosesan sinyal audio. Librosa digunakan untuk mengidentifikasi dan mengekstrak berbagai fitur audio, termasuk Mel-spectrogram. Mel-spectrogram dihasilkan dengan mengubah sinyal audio menjadi representasi visual dari spektrum frekuensi seiring waktu, yang sangat berguna untuk analisis menggunakan Convolutional Neural Network (CNN).

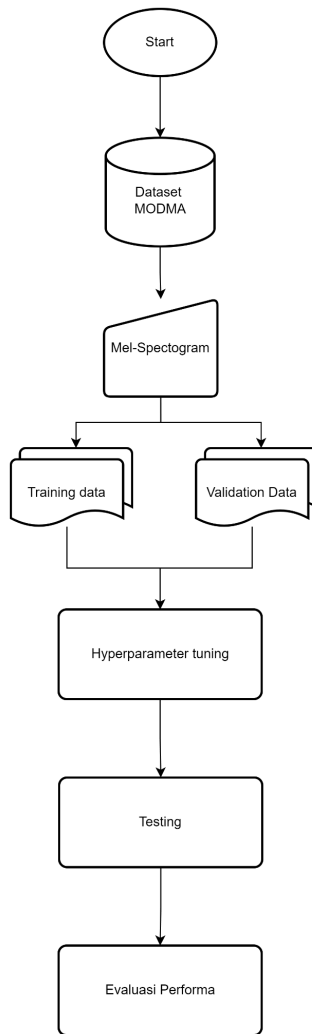
Fitur-fitur ini kemudian diolah untuk mengidentifikasi pola-pola yang mungkin mengindikasikan keberadaan depresi pada subjek. Proses ini memanfaatkan kemampuan Librosa dalam menangani kompleksitas data suara dan menyediakan dataset fitur yang baik untuk dianalisis. Dengan mengubah audio menjadi Mel-spectrogram lalu diproses kembali untuk menghilangkan anotasi dan juga legend yang ada digambar, sehingga hanya gambar

mel-spectrogram saja yang diproses, kita dapat memanfaatkan CNN yang dirancang untuk memproses data dalam bentuk gambar, sehingga meningkatkan akurasi dalam klasifikasi depresi.



Gambar 3.2 Alur Audio dan mel-spectrogram untuk proses Klasifikasi

3.1.4 Klasifikasi



Gambar Klasifikasi 3.3 Struktur Klasifikasi

Penulis menggunakan Convolutional Neural Network (CNN) digunakan untuk mengolah Mel-Spectrogram dari data audio karena CNN telah banyak digunakan dalam berbagai penelitian terkait analisis data visual dan audio. Alasan utama penggunaan CNN adalah kemampuannya yang unggul dalam mengenali pola kompleks dalam data, yang membuatnya menjadi pilihan populer di antara peneliti untuk tugas-tugas klasifikasi yang berkaitan dengan pengolahan sinyal dan image processing. Penggunaan ini didasarkan pada keberhasilan yang telah terdokumentasi dalam literatur [9], menunjukkan efektivitas CNN dalam membedakan karakteristik yang halus dalam data yang dapat mengindikasikan keberadaan kondisi seperti depresi.

Penelitian ini menggunakan dua pendekatan utama dalam implementasi CNN, yaitu CNN non-transfer learning dan CNN transfer learning. CNN non-transfer learning melibatkan pelatihan model dari awal menggunakan data yang tersedia, sedangkan CNN transfer learning memanfaatkan model yang telah dilatih sebelumnya pada dataset besar dan umum, kemudian diadaptasi dengan data spesifik yang kita miliki.

Alasan dari penggunaan dua model ini adalah untuk membandingkan mana model yang lebih baik terutama untuk data yang tidak terlalu besar jumlahnya. Model transfer learning, seperti InceptionV3 (InceptionV3), sering kali menunjukkan kinerja yang lebih baik pada dataset yang lebih kecil karena model tersebut telah memiliki pengetahuan awal yang luas dari pelatihan sebelumnya. Di sisi lain, menggunakan ShuffleNet sebagai model non-transfer learning memungkinkan untuk menguji sejauh mana model yang dibangun dari awal dapat mengenali pola dalam data audio yang terbatas.

Dengan membandingkan kedua pendekatan ini, peneliti dapat menentukan metode mana yang lebih efektif dalam mendeteksi depresi dari data audio dengan jumlah yang terbatas, serta memahami kelebihan dan kekurangan dari masing-masing pendekatan dalam konteks ini. Hasil perbandingan ini diharapkan dapat memberikan wawasan yang lebih mendalam tentang strategi yang paling efektif untuk mengoptimalkan deteksi otomatis kondisi mental melalui analisis suara.

3.1.5 Evaluasi

Evaluasi model CNN dengan matriks evaluasi dan *confusion matrix* untuk analisis yang lebih mendalam. Metrik seperti akurasi, presisi, recall, dan *F1-score* digunakan untuk menilai kinerja model dalam membedakan antara subjek depresi dan kontrol.

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN}$$

- TP (True Positives): Jumlah prediksi positif yang benar.
- TN (True Negatives): Jumlah prediksi negatif yang benar.
- FP (False Positives): Jumlah prediksi positif yang salah.
- FN (False Negatives): Jumlah prediksi negatif yang salah.

$$\text{Presisi} = \frac{TP}{TP+FP}, \text{ Recall} = \frac{TP}{TP+FN}, \text{ F1-Score} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}}$$

Confusion matrix membantu mengidentifikasi jumlah *true positives*, *false positives*, *true negatives*, dan *false negatives*, memberikan wawasan tentang jenis kesalahan yang dibuat oleh model. Pendekatan ini memungkinkan pemahaman yang lebih baik tentang efektivitas dan area peningkatan potensial untuk model CNN dalam deteksi dini depresi.

3.1.6 Analisis

Analisis dari evaluasi model CNN dalam deteksi dini depresi melalui audio menunjukkan bahwa dengan menggunakan matriks evaluasi dan *confusion matrix*, kita dapat memperoleh pemahaman yang komprehensif tentang kinerja model. Metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score* memberikan gambaran umum tentang efektivitas model, sementara *confusion matrix* memberikan gambaran spesifik tentang kekuatan dan kelemahan model dalam mengklasifikasikan kasus depresi.

Analisis ini bertujuan untuk membandingkan CNN non-transfer learning dan CNN transfer learning. Dengan mengevaluasi kedua pendekatan ini, kita dapat menentukan mana yang lebih efektif dalam menangani dataset yang terbatas. CNN non-transfer learning dilatih dari awal menggunakan Mel-spectrogram, sedangkan CNN transfer learning memanfaatkan model yang telah dilatih sebelumnya seperti InceptionV3 (InceptionV3)

Hasil analisis akan menunjukkan perbedaan kinerja kedua pendekatan ini, membantu kita memahami kelebihan dan kekurangan masing-masing dalam deteksi depresi berdasarkan data audio.