



# Sistem Mobile Deteksi Gangguan Kejiwaan Berbasis Suara Menggunakan Metode Deep Convolutional Neural Network

Kristiawan Nugroho<sup>1,\*</sup>, Alek Jusran<sup>1</sup>, Linda Kartika Sari<sup>2</sup>, Muhamat Nofiyanto<sup>2</sup>, Suprihartini<sup>3</sup>

<sup>1</sup> Fakultas Teknologi Informasi dan Industri, Teknologi Informasi, Universitas STIKUBANK, Semarang, Indonesia

<sup>2</sup> RSJD Dr. Amino Gondohutomo, Semarang, Indonesia

<sup>3</sup> RS UNIMUS, Semarang, Indonesia

Email: <sup>1,\*</sup>[kristiawan@edu.unisbank.ac.id](mailto:kristiawan@edu.unisbank.ac.id), <sup>2</sup>[alexjusran@gmail.com](mailto:alexjusran@gmail.com), <sup>2</sup>[kartikasari.linda@gmail.com](mailto:kartikasari.linda@gmail.com), <sup>2</sup>[muhamatnur@gmail.com](mailto:muhamatnur@gmail.com), <sup>3</sup>[suprihartini8@gmail.com](mailto:suprihartini8@gmail.com)

(\* : coresponding author)

**Abstrak**-Gangguan kejiwaan merupakan masalah kesehatan global yang sering kali tidak terdeteksi secara dini, sehingga membutuhkan pendekatan inovatif dalam deteksinya. Penelitian ini bertujuan untuk mengembangkan sistem mobile yang mampu mendeteksi gangguan kejiwaan berbasis suara menggunakan teknologi Deep Convolutional Neural Network. Metode yang digunakan dalam penelitian ini adalah pengumpulan data suara dari individu yang mengalami gejala gangguan kejiwaan, diikuti dengan ekstraksi fitur suara dan penerapan model Deep Convolutional Neural Network untuk klasifikasi gangguan tersebut. Sistem ini diuji dengan menggunakan dataset suara yang telah diolah, yang mencakup berbagai jenis gangguan kejiwaan, termasuk depresi dan kecemasan. Hasil penelitian menunjukkan bahwa model Deep Convolutional Neural Network mampu mencapai akurasi deteksi yang tinggi, dengan kemampuan mengenali gangguan kejiwaan berdasarkan ciri-ciri suara yang spesifik. Temuan ini membuka peluang baru untuk deteksi gangguan kejiwaan secara lebih cepat dan efisien menggunakan perangkat mobile, yang dapat diakses oleh masyarakat luas. Penelitian ini juga menunjukkan potensi besar teknologi deep learning dalam bidang kesehatan mental, khususnya dalam pencegahan dan diagnosis gangguan kejiwaan.

**Kata Kunci:** Gangguan; Kejiwaan; Sistem; Deteksi; Deep Learning.

**Abstract**-Mental disorders are a global health problem that often goes undetected early, requiring innovative approaches to their detection. This study aims to develop a mobile system capable of detecting mental disorders based on voice using Deep Convolutional Neural Network technology. The method used in this study is the collection of voice data from individuals experiencing symptoms of mental disorders, followed by voice feature extraction and the application of a Deep Convolutional Neural Network model for the classification of these disorders. The system was tested using a processed voice dataset, which includes various types of mental disorders, including depression and anxiety. The results showed that the Deep Convolutional Neural Network model was able to achieve high detection accuracy, with the ability to recognize mental disorders based on specific voice characteristics. This finding opens new opportunities for faster and more efficient detection of mental disorders using mobile devices, which are accessible to the wider community. This study also demonstrates the great potential of deep learning technology in the field of mental health, particularly in the prevention and diagnosis of mental disorders.

**Keywords:** Disorders; Mental Health; System; Detection; Deep Learning.

## 1. PENDAHULUAN

Perkembangan teknologi informasi (TI) dalam beberapa dekade terakhir telah membawa perubahan signifikan dalam berbagai sektor kehidupan manusia, salah satunya adalah dalam dunia medis. Dalam era digital saat ini, teknologi informasi telah berkembang dengan pesat, memperkenalkan berbagai inovasi yang mampu meningkatkan kualitas pelayanan kesehatan. Pemanfaatan teknologi dalam dunia medis mencakup berbagai bidang, mulai dari diagnosis, pengobatan, hingga manajemen rumah sakit. Salah satu perkembangan yang sangat menarik adalah penerapan kecerdasan buatan (AI), yang memungkinkan sistem untuk meniru kemampuan kognitif manusia[1] dalam menganalisis data besar, mengidentifikasi pola, dan memberikan rekomendasi atau keputusan yang akurat. Salah satu cabang kecerdasan buatan yang sedang berkembang pesat adalah deep learning, yang digunakan dalam berbagai aplikasi medis, mulai dari pengenalan gambar medis hingga analisis suara.

Kecerdasan buatan telah membuka berbagai kemungkinan baru dalam dunia medis, dengan aplikasinya yang sangat luas dalam berbagai aspek. Salah satu contoh penerapan AI dalam medis adalah penggunaan algoritma deep learning dalam pemrosesan citra medis untuk diagnosis penyakit[2]. Namun, tidak hanya terbatas pada gambar medis, kecerdasan buatan juga mulai digunakan dalam analisis suara, yang menjadi fokus utama dalam penelitian ini. Suara manusia, terutama yang berkaitan dengan gangguan kejiwaan, menyimpan berbagai informasi penting yang dapat digunakan untuk mendeteksi kondisi kesehatan mental pasien. Teknologi ini sangat berguna, mengingat gangguan kejiwaan sering kali sulit dideteksi secara klinis pada tahap awal dan membutuhkan deteksi yang lebih cepat agar penanganannya dapat segera dilakukan. Penggunaan teknologi suara berbasis AI menawarkan cara yang lebih praktis dan efektif dalam mendiagnosis gangguan mental.

Rumah Sakit Jiwa Daerah Amino Gondo menjadi salah satu rumah sakit yang menyediakan layanan kesehatan bagi pasien dengan gangguan kejiwaan. Meskipun fasilitas dan layanan medis yang ada sudah cukup baik, masalah utama yang dihadapi oleh rumah sakit ini adalah terbatasnya jumlah tenaga medis yang terampil dalam mendiagnosis gangguan kejiwaan secara dini. Gangguan kejiwaan, seperti depresi, kecemasan, dan gangguan bipolar, sering kali tidak terlihat jelas dalam pemeriksaan fisik atau tes laboratorium, sehingga diagnosis sering tertunda hingga gejala menjadi parah. Selain itu, rumah sakit ini juga menghadapi tantangan dalam mencatat dan menganalisis data medis pasien secara efektif.



Pengumpulan data suara pasien yang dapat memberikan indikasi adanya gangguan kejiwaan seringkali tidak terkelola dengan baik karena keterbatasan waktu dan sumber daya yang ada. Oleh karena itu, diperlukan sebuah solusi yang lebih efisien untuk mendeteksi gangguan kejiwaan sejak dini.

Guna mengatasi masalah tersebut, diperlukan sistem yang dapat mendeteksi gangguan kejiwaan pada pasien dengan lebih cepat dan efisien. Salah satu solusi yang sangat potensial adalah pengembangan sistem deteksi gangguan kejiwaan berbasis suara yang dapat dioperasikan melalui perangkat mobile. Dengan menggunakan teknologi suara dan analisis berbasis deep learning[3], sistem ini akan memungkinkan rumah sakit untuk mendeteksi gangguan kejiwaan pada pasien secara lebih cepat dan akurat. Penggunaan aplikasi berbasis mobile juga akan memudahkan tenaga medis untuk melakukan pemeriksaan kapan saja dan di mana saja, tanpa terbatas pada waktu atau tempat. Selain itu, aplikasi mobile ini dapat membantu mengurangi beban kerja tenaga medis, yang sering kali menghadapi banyak pasien dalam waktu yang terbatas.

Penelitian ini bertujuan untuk mengembangkan sebuah aplikasi deteksi gangguan kejiwaan berbasis mobile yang memanfaatkan teknologi deep convolutional neural network (DCNN). Teknologi deep learning, khususnya DCNN, dipilih karena kemampuannya dalam memproses data besar dan kompleks[4], seperti suara manusia, untuk mengidentifikasi pola-pola yang tidak dapat dideteksi dengan mudah oleh manusia. Aplikasi ini akan bekerja dengan mengumpulkan data suara dari pasien, mengidentifikasi ciri-ciri suara yang terkait dengan gangguan kejiwaan, dan kemudian memberikan diagnosis awal berdasarkan analisis tersebut. Dengan menggunakan aplikasi ini, diharapkan tenaga medis dapat melakukan deteksi dini terhadap gangguan kejiwaan[5] pada pasien tanpa memerlukan peralatan medis yang kompleks dan mahal. Selain itu, aplikasi ini juga dapat membantu pasien untuk mendapatkan perawatan lebih cepat sebelum gangguan kejiwaan mereka berkembang menjadi lebih parah.

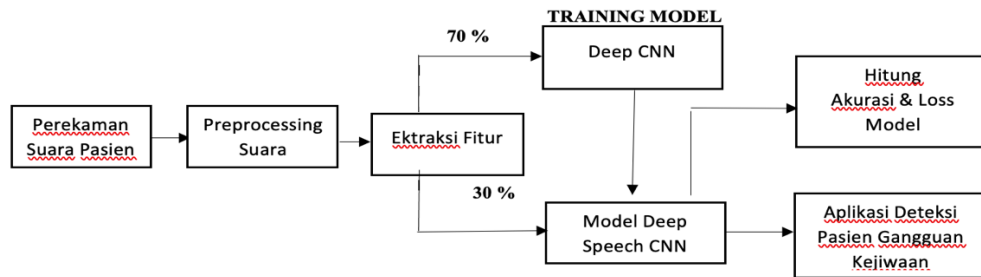
Beberapa penelitian mengenai peranan kecerdasan buatan dalam deteksi gangguan jiwa telah dilakukan antara lain oleh Wahyuni[6] dengan aplikasi pemantau pasien dengan gangguan jiwa maupun Sitanggang[7] mempergunakan algoritma LSTM (Long Short-Term Memory) menghasilkan kinerja model yang cukup menjanjikan. Selain itu beberapa penelitian lain juga dilakukan oleh Hidayat[8], Khairan[9] dan Saputra[10] dengan mempergunakan teknologi ChatBot sebagai sarana identifikasi awal pasien dengan gangguan kejiwaan. Sampai dengan saat ini penelitian berkembang dengan penerapan teknologi AI dari Machine Learning sampai dengan penerapan Deep Learning.

Hasil dari penelitian ini diharapkan dapat menghasilkan sebuah sistem yang efektif, efisien, dan mudah diakses oleh tenaga medis di Rumah Sakit Amino Gondo maupun rumah sakit lain yang menghadapi tantangan serupa. Sistem deteksi gangguan kejiwaan berbasis suara ini akan memperkenalkan metode baru dalam diagnosis gangguan mental yang lebih cepat dan praktis. Selain itu, aplikasi mobile ini juga dapat meningkatkan kualitas pelayanan kesehatan mental dengan memberikan akses lebih mudah kepada masyarakat untuk mendapatkan diagnosis awal. Dengan demikian, penelitian ini berkontribusi dalam menciptakan solusi yang inovatif untuk meningkatkan sistem kesehatan mental di Indonesia, sekaligus memanfaatkan kemajuan teknologi informasi dan kecerdasan buatan untuk meningkatkan deteksi dan penanganan gangguan kejiwaan pada pasien.

## 2. METODOLOGI PENELITIAN

Penelitian ini menghasilkan model deteksi pasien dengan gangguan kejiwaan berbasis suara menggunakan metode Convolutional Neural Network (CNN) dalam deep learning yang merupakan pendekatan yang populer dan robust dalam berbagai penelitian[11]. Kegiatan riset dimulai dengan pengumpulan data suara pasien yang mengalami gangguan kejiwaan. Data suara ini dikumpulkan melalui perekaman percakapan pasien di berbagai kondisi, seperti saat pasien berbicara dengan tenaga medis atau saat berbicara sendiri. Selanjutnya, data suara yang terkumpul akan melalui tahap preprocessing, di mana suara tersebut diubah menjadi fitur-fitur yang dapat dianalisis, seperti spektrogram atau fitur mel-frequency cepstral coefficients[12],[13],[14], yang dapat merepresentasikan karakteristik suara secara lebih detail.

Setelah itu, dataset yang telah diproses akan dibagi menjadi data pelatihan dan data pengujian untuk membangun model CNN. Model CNN akan dilatih untuk mengenali pola-pola tertentu[15] dalam fitur suara yang berkaitan dengan gangguan kejiwaan, seperti perubahan pola bicara, kelancaran, dan intonasi suara. Proses pelatihan ini melibatkan optimisasi parameter dalam arsitektur CNN untuk memaksimalkan akurasi deteksi. Setelah model terlatih, tahap evaluasi dilakukan menggunakan data pengujian untuk mengukur performa model, seperti akurasi, presisi, recall, dan F1-score. Hasil dari model ini diharapkan dapat memberikan alat bantu diagnosis yang efektif dan efisien dalam mendeteksi gangguan kejiwaan melalui analisis suara pasien, serta memberikan kontribusi pada perkembangan teknologi di bidang kesehatan mental berbasis AI. Tahapan selengkapnya penelitian yang dilakukan divisualisasikan pada gambar 1 sebagai berikut:



**Gambar 1.** Model deteksi pasien dengan gangguan kejiwaan

### 2.1 Perekaman Suara

Proses pengumpulan data dengan cara perekaman suara pasien dengan gangguan kejiwaan di ruang IGD dilakukan untuk mendapatkan informasi yang akurat terkait dengan kondisi mental pasien. Perekaman ini dimulai dengan persiapan alat perekam suara yang memiliki kualitas tinggi untuk memastikan kejelasan suara yang terekam. Selanjutnya, tenaga medis atau peneliti akan mendekati pasien dengan penuh perhatian dan kehati-hatian, mengingat kondisi pasien yang mungkin cemas atau tidak kooperatif. Proses perekaman dilakukan dengan cara berbicara atau berinteraksi dengan pasien, untuk memancing respon verbal yang dapat menggambarkan kondisi mental mereka, seperti delusi, halusinasi, atau gangguan bicara lainnya. Selama proses ini, pengamatan terhadap ekspresi wajah dan bahasa tubuh pasien juga penting untuk memperoleh gambaran menyeluruh tentang keadaan psikologis mereka. Data suara yang terkumpul kemudian dianalisis untuk melihat pola-pola tertentu yang dapat mengindikasikan kondisi gangguan kejiwaan. Perekaman ini tidak hanya bertujuan untuk mengumpulkan data, tetapi juga untuk memberikan pemahaman yang lebih dalam tentang dampak gangguan kejiwaan terhadap komunikasi verbal pasien di ruang IGD.

### 2.2 Preprocessing Data

Proses preprocessing merupakan langkah awal pemrosesan data dan penting dalam penelitian[15], proses preprocessing suara dimulai dengan pengambilan sampel suara setiap detik menggunakan Adobe Audition. Pada tahap ini, suara yang direkam dibagi menjadi potongan-potongan kecil, masing-masing berdurasi satu detik, untuk mempermudah analisis dan pengolahan selanjutnya. Setelah pengambilan sampel suara, langkah berikutnya adalah membersihkan noise atau gangguan suara yang tidak diinginkan. Adobe Audition menyediakan berbagai alat untuk menghilangkan noise, seperti fitur *Noise Reduction* dan *Click Removal*, yang membantu mengurangi suara latar belakang atau suara interferensi yang bisa mengganggu kualitas data. Proses pembersihan ini sangat penting untuk memastikan bahwa suara yang digunakan dalam analisis adalah suara yang jelas dan bebas dari gangguan. Setelah proses pembersihan noise selesai, hasil sampel suara tersebut siap untuk diekstraksi fitur, yang merupakan langkah penting dalam analisis suara[16], seperti pengenalan pola atau pemrosesan lebih lanjut. Dengan proses preprocessing yang cermat, kualitas data suara yang digunakan untuk analisis akan jauh lebih baik dan akurat.

### 2.3 Ekstraksi Fitur

Ekstraksi fitur suara adalah proses mengubah sinyal suara mentah menjadi representasi numerik yang lebih ringkas dan informatif, sehingga lebih mudah dianalisis dan diproses oleh sistem komputer. Fitur yang diekstraksi ini mencakup karakteristik penting dari suara, seperti frekuensi, amplitudo, dan tempo, yang digunakan dalam berbagai aplikasi seperti pengenalan suara, identifikasi pembicara, analisis emosi, dan pengolahan bahasa alami. Dengan ekstraksi fitur, sistem dapat mengenali pola dalam suara, memisahkan informasi relevan dari gangguan, dan meningkatkan akurasi dalam tugas-tugas seperti klasifikasi atau deteksi. Ekstraksi fitur suara menggunakan Mel-Frequency Cepstral Coefficients (MFCC) merupakan salah satu teknik utama dalam pengolahan sinyal suara, terutama dalam bidang pengenalan suara dan pemrosesan bahasa alami.

MFCC bekerja dengan cara mengubah sinyal suara yang berbentuk gelombang menjadi representasi numerik yang lebih mudah dianalisis, dengan menekankan pada fitur-fitur penting yang ada dalam spektrum frekuensi suara. Proses ekstraksi dimulai dengan merekam sinyal suara dan membaginya menjadi frame-frame kecil, biasanya dalam rentang waktu yang sangat pendek. Setiap frame kemudian diubah menjadi spektrum frekuensi menggunakan Transformasi Fourier. Selanjutnya, spektrum ini dipetakan ke dalam skala Mel, yang lebih sensitif terhadap frekuensi yang dapat didengar oleh telinga manusia. Setelah itu, dilakukan proses diskretisasi dengan menghitung koefisien cepstral untuk masing-masing frame. Hasil akhirnya adalah serangkaian koefisien yang menggambarkan karakteristik suara pada tiap frame, yang dapat digunakan untuk berbagai aplikasi seperti pengenalan suara, analisis emosi, atau identifikasi pembicara. MFCC sangat populer karena kemampuannya untuk menangkap informasi frekuensi penting dalam suara manusia, meskipun proses ini dapat dipengaruhi oleh kualitas rekaman dan noise yang ada dalam sinyal.

### 2.4 Deep CNN

Metode Deep Convolutional Neural Networks (Deep CNN) pada pengenalan suara merupakan pendekatan canggih dalam bidang pengolahan sinyal suara yang memanfaatkan kekuatan jaringan saraf tiruan untuk mengenali dan memproses



suara. Deep CNN bekerja dengan cara menganalisis fitur suara secara otomatis melalui serangkaian lapisan konvolusi yang bertujuan untuk mengekstraksi pola-pola yang relevan dari data suara. Dalam pengenalan suara, proses ini dimulai dengan konversi sinyal suara menjadi spektrum frekuensi, seperti spektrogram, yang merepresentasikan perubahan energi suara sepanjang waktu. Jaringan CNN kemudian dilatih untuk mengenali pola-pola ini, baik itu terkait dengan kata-kata tertentu, suara pembicara, atau jenis suara lainnya. Keunggulan utama dari Deep CNN dalam pengenalan suara adalah kemampuannya untuk mengotomatisasi ekstraksi fitur tanpa memerlukan intervensi manual[17].

Dengan adanya banyak lapisan konvolusi yang dapat menangkap berbagai tingkat representasi dari data suara, model CNN dapat mengenali pola yang sangat kompleks, bahkan yang tidak mudah dikenali oleh manusia. Selain itu, teknik ini juga dapat mengatasi variasi dalam kualitas rekaman suara, aksen, atau gangguan eksternal dengan lebih baik dibandingkan metode tradisional. Deep CNN telah terbukti efektif dalam berbagai aplikasi pengenalan suara, seperti sistem asisten virtual, perangkat pengenalan suara, dan analisis emosi suara, karena kemampuannya dalam memberikan akurasi tinggi dalam pengenalan suara yang bervariasi dan kompleks.

### 2.5. Pengujian Model

Pengujian model menggunakan matriks akurasi merupakan langkah penting dalam mengevaluasi kinerja suatu model machine learning atau deep learning[18]. Matriks akurasi mengukur sejauh mana model dapat memprediksi dengan benar berdasarkan data yang diuji. Dalam konteks pengujian model, akurasi dihitung dengan membandingkan jumlah prediksi yang benar dengan jumlah total prediksi yang dilakukan. Formula akurasi biasanya dihitung sebagai rasio antara jumlah prediksi yang benar dengan jumlah total data yang diuji, yaitu:

$$Akurasi = \frac{\text{Jumlah Prediksi Benar}}{\text{Jumlah Total Data}} \quad (1)$$

Meskipun akurasi merupakan metrik yang sederhana dan mudah dipahami, ia tidak selalu memberikan gambaran lengkap tentang performa model, terutama pada data yang tidak seimbang. Misalnya, jika model hanya memprediksi kelas mayoritas pada dataset yang sangat tidak seimbang, akurasi tinggi dapat tercapai meskipun model tidak efektif dalam mengenali kelas minoritas. Oleh karena itu, selain akurasi, sering kali digunakan juga metrik lain seperti presisi, recall, F1-score, dan confusion matrix untuk mendapatkan gambaran yang lebih mendalam tentang kekuatan dan kelemahan model dalam berbagai aspek klasifikasi. Pengujian dengan matriks akurasi sangat berguna untuk menilai kemampuan dasar model, namun evaluasi lebih lanjut diperlukan untuk memastikan model bekerja dengan baik pada berbagai jenis data.

### 2.5 Rancang Bangun Aplikasi

Perancangan aplikasi deteksi pasien dengan gangguan kejiwaan ini dikembangkan menggunakan bahasa pemrograman Python dengan framework Streamlit sebagai antarmuka utama, serta diimplementasikan agar dapat diakses melalui perangkat Android. Streamlit dipilih karena kemampuannya dalam membangun aplikasi berbasis web yang interaktif[19], ringan, dan responsif, sehingga memudahkan proses visualisasi data dan hasil analisis secara real-time. Aplikasi ini dirancang untuk mengolah data masukan pasien, seperti parameter suara, emosi, atau indikator perilaku tertentu (sesuai model yang diterapkan), kemudian memprosesnya menggunakan algoritma analisis atau kecerdasan buatan yang telah dilatih sebelumnya.

Arsitektur sistem aplikasi terdiri dari beberapa komponen utama, yaitu antarmuka pengguna (user interface), modul pemrosesan data, modul analisis dan klasifikasi, serta modul penyajian hasil. Pengguna dapat mengakses aplikasi melalui perangkat Android dengan membuka layanan Streamlit yang telah dikonfigurasi agar kompatibel dengan tampilan mobile. Data yang dimasukkan akan diproses oleh backend Python untuk menghasilkan informasi berupa indikasi awal kondisi kejiwaan pasien, yang kemudian disajikan dalam bentuk visual dan narasi yang mudah dipahami oleh pengguna maupun tenaga medis.

## 3. ANALISA DAN PEMBAHASAN

Hasil penelitian mengenai perancangan model deteksi bagi pasien dengan gangguan kejiwaan diperoleh melalui beberapa tahapan utama yang saling terintegrasi. Tahapan tersebut diawali dengan proses perekaman suara pasien, pengolahan suara, pembentukan sampai dengan evaluasi model yang bisa dijelaskan sebagai berikut :

### 3.1 Perekaman Data Suara

Perekaman data suara pasien dengan gangguan kejiwaan yang dilaksanakan di Instalasi Gawat Darurat (IGD) RSJD Amino Gondo memberikan kontribusi penting terhadap kualitas dan validitas data penelitian. Lingkungan IGD yang bersifat dinamis dan merepresentasikan kondisi awal pasien saat datang berobat memungkinkan diperolehnya sampel suara yang mencerminkan keadaan emosional dan psikologis pasien secara nyata. Data suara yang direkam pada fase ini menunjukkan variasi karakteristik akustik, seperti intonasi, tekanan suara, dan kestabilan artikulasi, yang relevan untuk dianalisis sebagai indikator awal gangguan kejiwaan.



Dari hasil analisis awal, proses perekaman di IGD juga memperlihatkan tantangan teknis, seperti gangguan kebisingan lingkungan dan keterbatasan durasi interaksi dengan pasien. Namun demikian, keterlibatan tenaga medis dalam proses perekaman membantu memastikan bahwa data yang diperoleh tetap layak dianalisis dan sesuai dengan kondisi klinis pasien. Pembahasan ini menunjukkan bahwa perekaman data suara di IGD RSJD Amino Gondo tidak hanya feasibel secara operasional, tetapi juga strategis dalam menghasilkan data autentik yang mendukung pengembangan model deteksi dini gangguan kejiwaan. Temuan ini memperkuat bahwa pemanfaatan data suara dari fase kegawatdaruratan dapat meningkatkan sensitivitas sistem dalam mengenali pola-pola awal gangguan mental, sehingga berpotensi menjadi dasar pengambilan keputusan klinis yang lebih cepat dan tepat.

### 3.2 Preprocessing Suara

Preprocessing suara merupakan tahap penting dalam sistem deteksi pasien dengan gangguan kejiwaan karena kualitas sinyal suara sangat memengaruhi akurasi hasil analisis. Pada tahap ini, sinyal suara pasien yang diperoleh dari proses perekaman terlebih dahulu melalui teknik sampling suara untuk menyeragamkan laju sampel (sampling rate), sehingga seluruh data audio memiliki format dan resolusi waktu yang konsisten. Proses sampling ini bertujuan untuk memastikan bahwa informasi frekuensi penting dalam suara[20] pasien dapat direpresentasikan secara optimal serta mengurangi beban komputasi pada tahap pemrosesan selanjutnya.

Setelah proses sampling, dilakukan reduksi noise untuk menghilangkan gangguan suara latar yang tidak relevan, seperti kebisingan lingkungan, gema, atau interferensi perangkat perekam. Teknik reduksi noise diterapkan dengan memanfaatkan metode penyangkutan sinyal, seperti filtering frekuensi dan noise suppression, sehingga sinyal utama yang merepresentasikan karakteristik suara pasien menjadi lebih dominan. Hasil dari tahap preprocessing ini adalah sinyal suara yang lebih bersih, stabil, dan representatif, sehingga mampu meningkatkan kinerja algoritma analisis maupun model

### 3.3 Ekstraksi Fitur

Ekstraksi fitur suara merupakan tahapan penting dalam sistem deteksi gangguan kejiwaan berbasis suara, karena kualitas fitur yang dihasilkan sangat menentukan kinerja model klasifikasi. Pada penelitian ini, ekstraksi fitur suara pasien dengan gangguan kejiwaan dilakukan menggunakan pendekatan Mel-Frequency Cepstral Coefficients (MFCC), yang terbukti efektif dalam merepresentasikan karakteristik spektral sinyal suara manusia. MFCC bekerja dengan meniru mekanisme pendengaran manusia melalui pemetaan frekuensi linier ke skala mel, sehingga mampu menangkap pola artikulasi, intonasi, dan emosi yang terkandung dalam suara pasien. Proses ekstraksi MFCC diawali dengan tahap pra-pemrosesan sinyal suara, meliputi normalisasi, framing, dan windowing untuk mengurangi distorsi sinyal.

Selanjutnya, sinyal suara ditransformasikan ke domain frekuensi menggunakan Fast Fourier Transform (FFT), kemudian dilewatkan melalui filter bank mel untuk memperoleh energi pada setiap pita frekuensi. Hasil log-energi filter tersebut diproses menggunakan Discrete Cosine Transform (DCT) guna menghasilkan koefisien MFCC yang bersifat ringkas dan diskriminatif. Koefisien MFCC inilah yang digunakan sebagai representasi fitur suara pasien dan menjadi masukan utama bagi algoritma pembelajaran mesin atau deep learning dalam mengidentifikasi indikasi gangguan kejiwaan secara dini.

### 3.4 Metode Deep CNN

Metode Deep Convolutional Neural Network (Deep CNN) digunakan dalam penelitian ini sebagai pendekatan utama untuk mendeteksi gangguan kejiwaan pada pasien berbasis analisis suara. Deep CNN dipilih karena kemampuannya yang unggul dalam mengekstraksi fitur kompleks dan hierarkis dari data sinyal, termasuk sinyal suara yang memiliki karakteristik nonlinier dan dinamis. Pada tahap awal, suara pasien direkam dan diproses melalui tahapan prapemrosesan, seperti penghilangan noise, normalisasi sinyal, serta segmentasi suara untuk memastikan kualitas data yang optimal. Selanjutnya, sinyal suara diubah ke dalam bentuk representasi fitur, seperti Mel-Frequency Cepstral Coefficients (MFCC), spektrogram, atau mel-spektrogram, yang kemudian dijadikan sebagai input ke dalam arsitektur Deep CNN.

Arsitektur jaringan terdiri dari beberapa lapisan konvolusi, pooling, dan fully connected layer yang berfungsi untuk menangkap pola-pola penting pada spektrum suara yang berkaitan dengan kondisi psikologis pasien, seperti perubahan intonasi, kecepatan bicara, dan energi suara. Proses pelatihan dilakukan menggunakan dataset suara pasien dengan label kondisi kejiwaan tertentu, sehingga model dapat mempelajari perbedaan karakteristik suara antara pasien dengan gangguan kejiwaan dan individu sehat. Melalui mekanisme backpropagation dan optimasi parameter, model Deep CNN mampu meningkatkan akurasi klasifikasi secara signifikan. Dengan pendekatan ini, sistem deteksi yang dihasilkan bersifat objektif, non-invasif, dan memiliki potensi besar untuk digunakan sebagai alat bantu skrining dini gangguan kejiwaan dalam mendukung pengambilan keputusan klinis oleh tenaga kesehatan. Pembelajaran mempergunakan model Deep CNN untuk melakukan training data suara pasien dengan gangguan kejiwaan. Hasil pengukuran kinerja model Deep CNN dapat disampaikan seperti pada tabel 1 berikut:

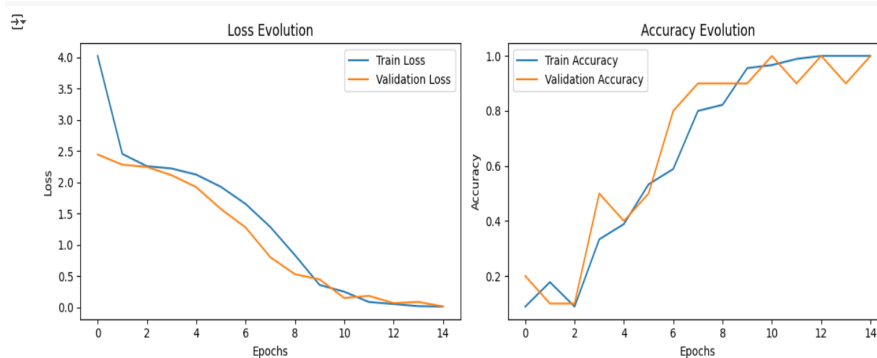
**Tabel 1.** Pengukuran akurasi dan loss model

Epoch	Akurasi	Loss
1	0.0851	3.7020
2	0.1705	2.4946

Epoch	Akurasi	Loss
3	0.0766	2.2445
4	0.2976	2.2222
5	0.4210	2.1398
6	0.4373	1.9781
7	0.6055	1.6522
8	0.7867	1.3455
9	0.8583	0.8675
10	0.9595	0.3658

Pada tabel 1 diperoleh hasil pengukuran kinerja model Deep CNN dari epoch 1 sampai dengan 10 diperoleh hasil bahwa model memiliki kinerja yang semakin membaik dengan angka maksimal akurasi dicapai pada epoch 10 yaitu sebesar 95.95% dan loss model sebesar 0.3636, hal ini menunjukkan model memiliki kinerja yang baik dalam mendeteksi suara pasien dengan gangguan kejiwaan. Berdasarkan hasil pelatihan model Deep Convolutional Neural Network (Deep CNN) selama 10 epoch, diperoleh tren peningkatan kinerja model yang sangat signifikan, baik dari sisi akurasi maupun nilai loss. Pada tahap awal pelatihan, khususnya pada epoch 1 hingga epoch 3, model menunjukkan akurasi yang masih sangat rendah, yaitu berkisar antara 7,66% hingga 17,05%, dengan nilai loss yang relatif tinggi mencapai 3,7020 pada epoch pertama. Kondisi ini mengindikasikan bahwa model masih berada pada fase awal pembelajaran dan belum mampu mengekstraksi pola serta karakteristik data secara optimal. Memasuki epoch ke-4 hingga epoch ke-6, terlihat adanya peningkatan akurasi yang cukup konsisten, dari 29,76% menjadi 43,73%, disertai dengan penurunan nilai loss hingga 1,9781.

Hal ini menunjukkan bahwa proses pembelajaran mulai stabil dan model mulai memahami representasi fitur yang relevan dari data masukan. Peningkatan performa yang paling signifikan terjadi pada epoch ke-7 hingga epoch ke-10, di mana akurasi meningkat secara tajam dari 60,55% menjadi 95,95%, sementara nilai loss menurun drastis dari 1,6522 menjadi 0,3658. Penurunan loss yang konsisten seiring dengan peningkatan akurasi menandakan bahwa model mengalami konvergensi yang baik tanpa indikasi ketidakstabilan selama proses pelatihan. Hasil ini mengonfirmasi kemampuan Deep CNN dalam mempelajari fitur-fitur kompleks dan nonlinier secara efektif, sehingga mampu menghasilkan prediksi yang sangat akurat. Secara keseluruhan, kombinasi peningkatan akurasi yang progresif dan penurunan loss yang stabil menunjukkan bahwa arsitektur Deep CNN yang diusulkan memiliki kinerja yang sangat baik dan layak digunakan sebagai model klasifikasi, khususnya dalam mendukung sistem deteksi dini berbasis kecerdasan buatan. Performa model selengkapnyanya divisualisasikan pada gambar 2 sebagai berikut :



**Gambar 2.** Grafik Kinerja Model Deep CNN

Pada gambar 2 terlihat grafik kinerja akurasi menunjukkan yang semakin membaik dengan tingkat akurasi maksimal mendekati 100%, Hal ini menunjukkan model memiliki kinerja yang unggul dalam pengenalan suara pasien dengan gangguan kejiwaan. Berdasarkan grafik evolusi nilai loss dan akurasi selama proses pelatihan model Deep Convolutional Neural Network (Deep CNN), terlihat bahwa model menunjukkan kinerja pembelajaran yang sangat baik dan stabil baik pada data pelatihan (training) maupun data validasi. Pada grafik *loss evolution*, nilai loss training pada epoch awal masih relatif tinggi, yaitu berada di atas 4,0, sedangkan loss validasi berada di kisaran 2,4. Kondisi ini menunjukkan bahwa pada tahap awal pelatihan, model belum mampu mempelajari pola data secara optimal dan masih mengalami kesalahan prediksi yang cukup besar. Seiring bertambahnya jumlah epoch, baik loss training maupun loss validasi mengalami penurunan yang konsisten dan signifikan. Penurunan ini mengindikasikan bahwa model secara bertahap mampu menyesuaikan bobot dan parameter internalnya untuk meminimalkan kesalahan prediksi. Pada epoch menengah hingga akhir, nilai loss training dan validasi mendekati nol, serta menunjukkan jarak yang sangat kecil satu sama lain, yang menandakan proses konvergensi berjalan dengan baik dan stabil.

Selanjutnya, grafik *accuracy evolution* memperlihatkan tren peningkatan akurasi yang sangat jelas baik pada data training maupun validation. Pada epoch awal, akurasi training masih berada di bawah 20%, sedangkan akurasi validasi menunjukkan fluktuasi ringan akibat model yang masih berada pada fase awal pembelajaran. Namun, mulai epoch ke-4

hingga epoch ke-7, terjadi peningkatan akurasi yang cukup tajam, di mana akurasi training dan validasi sama-sama meningkat hingga di atas 70%. Hal ini menunjukkan bahwa model mulai berhasil mengekstraksi fitur-fitur penting dari data masukan dan mampu melakukan klasifikasi dengan lebih akurat. Pada epoch selanjutnya, akurasi training mencapai nilai mendekati 100%, sementara akurasi validasi juga berada pada rentang yang sangat tinggi, meskipun terdapat sedikit fluktuasi kecil. Fluktuasi ini masih dapat ditoleransi dan tidak menunjukkan indikasi overfitting yang signifikan, karena perbedaan antara akurasi training dan validasi relatif kecil.

Secara keseluruhan, keselarasan antara penurunan loss dan peningkatan akurasi pada data training dan validation menunjukkan bahwa arsitektur Deep CNN yang digunakan memiliki kemampuan generalisasi yang sangat baik. Model tidak hanya mampu belajar secara efektif dari data pelatihan, tetapi juga mempertahankan kinerja yang tinggi ketika diuji pada data validasi. Dengan demikian, hasil ini mengonfirmasi bahwa model yang diusulkan layak digunakan sebagai sistem klasifikasi yang andal, khususnya dalam konteks pengembangan sistem deteksi gangguan kejiwaan berbasis suara yang membutuhkan akurasi tinggi dan stabilitas model.

## 4. IMPLEMENTASI

### 4.1 Aplikasi Deteksi Pasien Gangguan Kejiwaan

Model deteksi pasien dengan gangguan kejiwaan kemudian dikembangkan menjadi aplikasi deteksi gangguan kejiwaan yang bisa dijalankan pada perangkat berbasis Android. Hasil pengembangan aplikasi tersebut terlihat pada gambar gambar sebagai berikut :



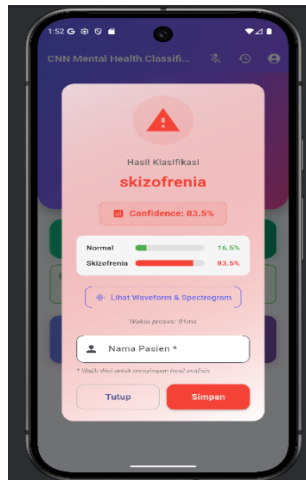
Gambar 3. Halaman Login Aplikasi

Aplikasi deteksi pasien dengan gangguan kejiwaan dibangun menggunakan streamlit dan bahasa Python memiliki tampilan login awal dimana pemakai harus memasukkan user name dan password. Jika belum memilikinya maka pemakai bisa melakukan registrasi awal untuk mendapatkan akun user name dan password. Tampilan login aplikasi bisa dilihat tampilannya pada gambar 2. Setelah memasukkan user name dan password maka pemakai akan dibawa ke halaman dashboad untuk melakukan deteksi suara seperti ditunjukkan pada gambar 4.



Gambar 4. Halaman perekaman dan upload suara

Pada gambar 4 terlihat bahwa halaman dashboard deteksi suara bisa dipilih dengan 2 cara yaitu dengan melakukan perekaman suara kita secara langsung dengan terlebih dahulu mengklik tombol start recording, kemudian diikuti dengan berbicara langsung selama beberapa detik maka aplikasi akan melakukan perekaman otomatis. Cara ke 2 bisa dilakukan dengan melakukan upload file suara yang sudah kita persiapkan sebelumnya. Kedua cara tersebut akan membawa kita ke halaman hasil deteksi suara seperti terlihat pada gambar 5.



Gambar 5. Hasil deteksi suara

Berdasarkan hasil deteksi suara menggunakan aplikasi akan didapatkan hasil analisis model dimana akan ditampilkan hasilnya dalam 2 kelas yaitu kelas hasil deteksi normal dan hasil deteksi skizofrenia (gangguan kejiwaan).

## 5. KESIMPULAN

Penelitian ini berhasil mengembangkan sebuah sistem deteksi gangguan kejiwaan berbasis suara yang memanfaatkan teknologi Deep Convolutional Neural Network (DCNN) dan dapat dioperasikan secara optimal melalui perangkat mobile. Sistem yang dikembangkan dirancang untuk menganalisis karakteristik suara pasien, seperti intonasi, frekuensi, energi, dan pola temporal, yang kemudian diproses menggunakan model deep learning guna mengidentifikasi indikasi gangguan kejiwaan, khususnya depresi dan kecemasan. Hasil pengujian menunjukkan bahwa sistem mampu mencapai tingkat akurasi yang sangat tinggi, yaitu sebesar 95,95%, yang mengindikasikan bahwa pendekatan berbasis DCNN memiliki performa yang andal dan konsisten dalam melakukan klasifikasi kondisi kesehatan mental berdasarkan data suara. Penerapan metode deep learning, terutama DCNN, dalam analisis suara memberikan kontribusi signifikan terhadap peningkatan kemampuan deteksi dini gangguan kejiwaan yang selama ini sering sulit diidentifikasi melalui pemeriksaan klinis konvensional. Banyak gangguan kejiwaan tidak menunjukkan gejala fisik yang jelas pada tahap awal, sehingga pendekatan berbasis suara menjadi solusi alternatif yang objektif dan non-invasif. Selain itu, integrasi sistem ini ke dalam perangkat mobile memungkinkan tenaga medis melakukan proses deteksi secara fleksibel di berbagai lokasi, baik di fasilitas kesehatan maupun di luar rumah sakit, tanpa memerlukan peralatan khusus yang kompleks. Keberadaan aplikasi ini juga berpotensi membantu mengurangi beban kerja rumah sakit dan tenaga kesehatan, karena proses skrining awal dapat dilakukan dengan lebih cepat dan efisien. Dengan demikian, sumber daya medis dapat difokuskan pada penanganan pasien yang membutuhkan intervensi lanjutan. Lebih jauh lagi, sistem deteksi gangguan kejiwaan berbasis suara ini memiliki peluang besar untuk diimplementasikan di berbagai rumah sakit atau pusat layanan kesehatan lain yang menghadapi keterbatasan tenaga ahli kesehatan mental. Pemanfaatan perangkat mobile yang mudah diakses menjadikan teknologi ini relevan untuk diterapkan secara luas, khususnya di daerah dengan keterbatasan fasilitas kesehatan. Secara keseluruhan, hasil penelitian ini membuka peluang pengembangan lebih lanjut dalam bidang kesehatan mental digital, baik untuk keperluan pencegahan, pemantauan, maupun diagnosis awal gangguan kejiwaan. Dengan adanya sistem ini, diharapkan kualitas pelayanan kesehatan mental di Indonesia dapat meningkat secara signifikan, memberikan akses yang lebih merata bagi masyarakat, serta mendukung upaya deteksi dan penanganan gangguan kejiwaan sejak dini guna mengurangi dampak sosial, ekonomi, dan psikologis yang ditimbulkan.

## REFERENCES

- [1] V. Albert, J. Mardame, F. P. Wardani, and S. Penabur, "Application of Artificial Intelligence Technology to Improve Responsiveness and Speed of Operations in Organizations Penerapan Teknologi Kecerdasan Buatan untuk Meningkatkan Responsivitas dan Kecepatan Operasional dalam Organisasi," *Pros. Semin. Nas. Ilmu Manajemen, Ekon. Keuang. dan Bisnis*, vol. 3, no. 2, pp. 145–164, 2024.
- [2] N. Nova, A. Mulyanti, C. Silmy, A. Putri, and L. Mulyani, "Systematic Review : Pemanfaatan Deep Learning untuk Diagnosis Penyakit Menggunakan MRI," *J. Penelit. Inov.*, vol. 5, no. 2, pp. 839–852, 2025.
- [3] H. Wijaya, "Teknologi Pengenalan Suara tentang Metode , Bahasa dan Tantangan : Systematic Literature Review," *J. Bin.*



- Digit. - Technol.*, vol. 7, no. 2, 2024, doi: 10.32877/bt.v7i2.1888.
- [4] Y. Safitri, M. Muhammad, and J. Samodro, "Strategi dan Efektivitas Deep Learning untuk Mitigasi Ancaman Keamanan Jaringan di Era IoT," *J. Comput. Sci. Informatics*, doi: 10.34304/scientific.v2i1.338.
- [5] N. Tahun, "Pendahuluan Masalah kesehatan jiwa merupakan salah satu masalah yang perlu mendapat perhatian dalam penanganannya . Hal ini dikarenakan jumlah kasus populasi ( Kementerian Kesehatan RI , 2018 ). Kabupaten Pena," *GEMAKES J. Pengabd. Kpd. Masy.*, vol. 5, pp. 245–255, 2025, doi: 10.36082/gemakes.v5i2.2171.
- [6] P. Kesehatan and M. Berbasis, "Pengembangan Aplikasi Berbasis Kecerdasan Buatan untuk Pemantauan dan Peningkatan Kesehatan Mental Berbasis Pengguna," *Indones. Res. J. Educ.*, vol. 5, pp. 2486–2491, 2025.
- [7] A. S. Sitanggang *et al.*, "Peran Kecerdasan Buatan terhadap Diagnosis dan Penanggulangan Masalah Kesehatan Mental Universitas Komputer Indonesia , Indonesia Menurut Federasi Kesehatan Mental Dunia ( World Federation for Mental Health ), kesehatan mental adalah ( 1 ) kondisi yang mem," *Merkurius J. Ris. Sist. Inf. dan Tek. Inform.*, vol. 3, no. September, 2025.
- [8] R. R. Hidayat and M. Fikry, "Chatbot Deteksi Awal Gangguan Kecemasan Menggunakan Dialogflow," *J. Teknol. TERPADU*, vol. 11, no. 2, 2023.
- [9] T. D. A. N. Prospek, "Chatbot Ai Dalam Identifikasi Awal Gangguan Kesehatan Mental Di Indonesia ;," *J. Empati*, vol. 13, pp. 498–508, 2024.
- [10] F. Sains, T. Universitas, And B. Darma, "Penerapan Chatbot Sebagai Alat Diagnosis Indikasi Penyakit Setres Dengan Menggunakan Metode Backward Chaining," *J. Sains Student Res.*, vol. 3, no. 2, pp. 441–447, 2025.
- [11] K. Azmi and S. Defit, "Implementasi Convolutional Neural Network ( CNN ) Untuk Klasifikasi Batik Tanah Liat Sumatera Barat," *J. Unitek*, vol. 16, no. 1, pp. 2580–2582, 2023.
- [12] F. Science and U. Diponegoro, "Comparison of Feature Extraction MFCC and LPC in Automatic Speech Recognition for Indonesian," *TELKOMNIKA*, vol. 15, no. 1, pp. 292–298, 2017, doi: 10.12928/TELKOMNIKA.v15i1.3605.
- [13] D. Cabrera, R. Medina, M. Cerrada, R. Sánchez, E. Estupiñan, and C. Li, "applied sciences Improved Mel Frequency Cepstral Coefficients for Compressors and Pumps Fault Diagnosis with Deep Learning Models," *Appl. Sci.*, 2024.
- [14] G. D. Saxena and S. Ali, "Extricate Features Utilizing Mel Frequency Cepstral Coefficient in Automatic Speech Recognition System," *I. J. Eng. Manuf.*, vol. 12, no. 6, pp. 14–21, 2022, doi: 10.5815/ijem.2022.06.02.
- [15] S. Ghandi and Y. R. Ramadhan, "Penerapan Metode Convolutional Neural Network ( Cnn ) Dalam Aplikasi Pendeteksi Penyakit Daun Tanaman Kentang Berbasis Android," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 5, pp. 8701–8708, 2024.
- [16] U. Machine and L. Algorithm, "Improved Feature Parameter Extraction from Speech Signals Using Machine Learning Algorithm," *sensors*, 2022.
- [17] M. Nurhidayanti, "Penerapan Deep Learning dalam Pengenalan Wajah untuk Sistem Keamanan," *J. Inform. Indones.*, vol. 1, no. 1, pp. 29–37, 2025.
- [18] S. Dewi, "Analisis Sentimen Aplikasi WhatsApp berdasarkan Ulasan di," *Arus J. Sains dan Teknol. ( AJST )*, vol. 2, no. 2, 2024.
- [19] O. V. Agesti *et al.*, "Perbandingan Algoritma Decision Tree Dan C4.5 Untuk Klasifikasi Kelayakan Calon Pendoror Darah Di Pmi Kota Semarang," *J. Inform. Teknol. dan Sains*, pp. 1149–1158, 2025.
- [20] C. Technologies *et al.*, "Analisis Penggunaan Frekuensi Sampling dan Filter Lowpass - FIR untuk Mencegah Aliasing," *J. Telecommun. Networks, Electron. Comput. Technol.*, vol. 5, pp. 36–48, 2025.