

ANALISIS CITRA FITUR WAJAH UNTUK DETEKSI DINI DEPRESI DENGAN TEKNIK HIBRIDA BERDASARKAN GABUNGAN FITUR *CONVOLUTIONAL NEURAL NETWORK* (CNN)

Hermawan Syahputra^{*1}, Syawal Gultom², Arnita³, Dedy Kiswanto⁴, Saatira Hilmi Batubara⁵,
Jeremia Manurung⁶

^{1,2,3,4,5,6} Universitas Negeri Medan, Kabupaten Deli Serdang

Email: ¹hsyahputra@unimed.ac.id, ²syawalgultom@unimed.ac.id, ³arnita@unimed.ac.id,
⁴dedykiswanto@unimed.ac.id, ⁵saatirahilmibatubara@gmail.com, ⁶jeremiamanurung@mail.ugm.ac.id,

^{*}Penulis Korespondensi

(Naskah masuk: 30 Oktober 2025, diterima untuk diterbitkan: 03 Maret 2026)

Abstrak

Depresi merupakan salah satu masalah kesehatan mental yang umum dan menjadi penyebab utama disabilitas di seluruh dunia. Deteksi dini sangat penting karena gejala awal sering tidak dikenali atau diabaikan. Penelitian ini mengusulkan sistem deteksi dini depresi berbasis kecerdasan buatan melalui analisis citra wajah menggunakan pendekatan hibrida Convolutional Neural Network (CNN). Data diperoleh dari rekaman wawancara klinis pasien dengan diagnosis depresi dan individu normal. Ekstraksi fitur dilakukan menggunakan tiga arsitektur CNN, yaitu VGG16, ResNet101V2, dan MobileNetV3, yang kemudian digabungkan untuk membentuk representasi fitur gabungan. Selanjutnya, algoritma Random Forest digunakan sebagai klasifikator. Hasil eksperimen menunjukkan bahwa kombinasi VGG16 dan MobileNetV3 memberikan performa terbaik dengan akurasi 94.55%, presisi 98.04%, dan spesifisitas 98.18%. Temuan ini menunjukkan bahwa pendekatan hibrida CNN dengan Random Forest berpotensi menjadi metode efektif untuk mendukung deteksi dini depresi berbasis citra wajah.

Kata kunci: *Depresi, Ekspresi Wajah, CNN, VGG16, ResNet101V2, MobileNetV3, Random Forest*

FACIAL FEATURE IMAGE ANALYSIS FOR EARLY DETECTION OF DEPRESSION USING A HYBRID TECHNIQUE BASED ON A COMBINATION OF CONVOLUTIONAL NEURAL NETWORK (CNN) FEATURES

Abstract

Depression is a major mental health disorder and one of the leading causes of disability worldwide. Early detection is crucial since initial symptoms are often overlooked or unrecognized. This study proposes an artificial intelligence-based early detection system for depression through facial image analysis using a hybrid Convolutional Neural Network (CNN) approach. Facial data were obtained from clinical interview recordings of patients diagnosed with depression and healthy individuals. Feature extraction was carried out using three CNN architectures VGG16, ResNet101V2, and MobileNetV3 whose features were then combined into a joint representation. A Random Forest algorithm was employed for classification. Experimental results indicate that the hybrid model combining VGG16 and MobileNetV3 achieved the best performance, with 94.55% accuracy, 98.04% precision, and 98.18% specificity. These findings demonstrate that the proposed CNN Random Forest hybrid approach is a promising method for supporting early depression detection based on facial features.

Keywords: *Depression, Facial Expression, CNN, VGG16, ResNet101V2, MobileNetV3, Random Forest*

1. PENDAHULUAN

Depresi merupakan salah satu gangguan kesehatan mental yang paling umum dan menjadi penyebab utama disabilitas di seluruh dunia. Organisasi Kesehatan Dunia (WHO, 2023) memperkirakan bahwa lebih dari 280 juta orang mengalami depresi, atau sekitar 3,8% dari populasi global, dengan prevalensi yang lebih tinggi pada perempuan dibandingkan laki-laki. Gangguan ini

ditandai oleh suasana hati yang menurun, kehilangan minat, gangguan tidur, perubahan nafsu makan, dan menurunnya konsentrasi (Abdurachman & Chahyati, 2025; Wardani & Tiastiningsih, 2023). Dalam kasus yang berat, depresi dapat menyebabkan keinginan untuk mengakhiri hidup. Kondisi tersebut bukan hanya berdampak pada kesejahteraan individu, tetapi juga menimbulkan beban sosial dan ekonomi yang signifikan bagi masyarakat (Humairah & Komalasari, 2024).

Di Indonesia, prevalensi depresi juga menunjukkan angka yang mengkhawatirkan. Berdasarkan Survei Kesehatan Indonesia tahun 2023, sekitar 1,4% penduduk berusia di atas 15 tahun mengalami depresi, namun hanya 12,7% yang mendapatkan penanganan medis. Artinya, lebih dari 80% penderita tidak menjalani pengobatan (Ahdiat, 2024). Rendahnya angka penanganan ini mengindikasikan adanya kesenjangan besar dalam deteksi dini dan layanan kesehatan mental. Gejala depresi sering kali tidak disadari atau diabaikan, sehingga penderita baru mendapatkan bantuan ketika kondisinya sudah parah. Oleh karena itu, dibutuhkan pendekatan diagnostik yang lebih objektif, cepat, dan mudah diterapkan.

Analisis ekspresi wajah menawarkan potensi besar untuk deteksi dini depresi. Secara klinis, tenaga medis berpengalaman dapat mengidentifikasi tanda-tanda depresi melalui perubahan ekspresi wajah, seperti penurunan intensitas senyuman, tekanan bibir yang lebih kuat, atau gerakan alis tertentu (Bubulac et al., 2025; Kraft-Feil et al., 2023). Selain itu, Ekspresi wajah dan perilaku seseorang dapat menjadi indikator penting dalam menilai adanya gangguan depresi (Pratamaatmadja & Prasetyo, 2024). Penelitian (Pampouchidou et al., 2020; Lestari, 2026) menunjukkan bahwa analisis visual berbasis citra wajah dapat menjadi ukuran objektif status psikoemosional dan berfungsi sebagai alat bantu diagnosis depresi yang efisien.

Perkembangan teknologi kecerdasan buatan (AI) dan pembelajaran mendalam (*deep learning*) telah membuka peluang baru untuk menganalisis data visual secara otomatis (Purnama et al., 2025; Maxnun & Sadih, 2025). Convolutional Neural Network (CNN) merupakan salah satu arsitektur *deep learning* yang paling efektif dalam mengekstraksi fitur dari citra wajah. CNN mampu mengenali pola nonlinier yang kompleks dan menyesuaikan bobot pembelajaran melalui proses *backpropagation* (Chinthia et al., 2025; Mulyani et al., 2024). Beberapa arsitektur CNN yang umum digunakan antara lain VGG16, ResNet101V2, dan MobileNetV3, yang masing-masing memiliki keunggulan tersendiri dalam mendeteksi fitur visual. Namun, penggunaan satu arsitektur CNN tunggal sering kali belum mampu menghasilkan performa optimal, terutama ketika data pelatihan terbatas atau ekspresi wajah bervariasi secara signifikan (Reksomulyo, 2025; Akbar, Saifullah, & Prapcoyo, 2024).

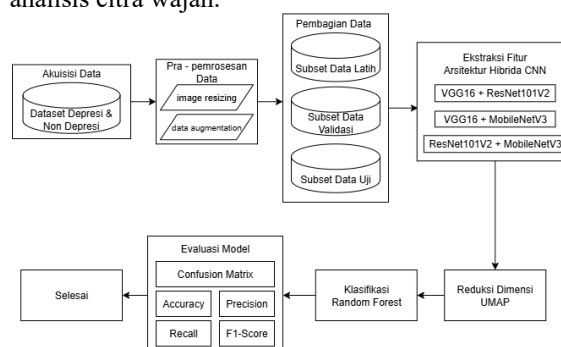
Untuk mengatasi keterbatasan tersebut, pendekatan *hybrid deep learning* dikembangkan dengan menggabungkan beberapa model CNN untuk memperkaya representasi fitur. Pendekatan ini telah terbukti meningkatkan akurasi pada berbagai tugas klasifikasi citra (Awaji, A., Khan, M. U., & Kim, 2023; Karim et al., 2025). Dalam konteks deteksi depresi, penggabungan fitur dari beberapa model CNN memungkinkan sistem untuk menangkap variasi ekspresi yang lebih halus dan kompleks

(Syaiqalloh, 2025; Fattah, Rosid, & Aji, 2026). Penelitian ini menerapkan gabungan fitur dari arsitektur VGG16, ResNet101V2, dan MobileNetV3, kemudian melakukan reduksi dimensi menggunakan algoritma *Uniform Manifold Approximations and Projection* (UMAP) untuk menghasilkan representasi fitur yang kompak dan bermakna. Proses klasifikasi selanjutnya dilakukan dengan algoritma *Random Forest* (RF), yang dikenal mampu menangani data berdimensi tinggi dan mengurangi risiko *overfitting*.

Melalui integrasi antara CNN dan RF, penelitian ini bertujuan mengembangkan sistem deteksi dini depresi berbasis citra wajah yang mampu memberikan hasil klasifikasi yang akurat, efisien, dan objektif. Sistem ini diharapkan dapat mendukung proses skrining awal pasien dalam praktik klinis serta menjadi langkah awal menuju pengembangan teknologi diagnostik cerdas di bidang kesehatan mental.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimental berbasis *deep learning* dengan tujuan mengembangkan sistem deteksi dini depresi melalui analisis citra wajah.



Gambar 1 Alur Penelitian

Dapat dilihat pada Gambar 1 sistem ini dirancang menggunakan kombinasi arsitektur *Convolutional Neural Network* (CNN) dan algoritma *Random Forest* (RF) dalam satu pipeline hibrida. Penelitian dilakukan pada tahun 2025 di Rumah Sakit Jiwa Prof. Dr. Muhammad Ildrem, Provinsi Sumatera Utara.

Tahapan penelitian meliputi:

1. Pra-pemrosesan data, mencakup ekstraksi frame dari video, *resizing* gambar ke 224×224 piksel, normalisasi, dan *augmentasi* untuk meningkatkan variasi data serta mencegah *overfitting*.
2. Pembagian dataset menjadi 80% data latih, 10% validasi, dan 10% uji (total 1008 citra, masing-masing 504 per kelas).
3. Ekstraksi fitur menggunakan model CNN hibrida, yaitu kombinasi VGG16 +

ResNet101V2, VGG16 + MobileNetV3, dan ResNet101V2 + MobileNetV3. Masing-masing CNN berfungsi sebagai *feature extractor* untuk menghasilkan representasi mendalam dari ekspresi wajah.

4. Reduksi dimensi menggunakan UMAP (Uniform Manifold Approximation and Projection) untuk menyaring fitur penting dan mengurangi kompleksitas komputasi.
5. Klasifikasi dengan algoritma Random Forest, yang memanfaatkan hasil ekstraksi dan reduksi fitur untuk membedakan antara citra wajah depresi dan non-depresi.
6. Evaluasi model dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, *specificity*, dan AUC dari kurva ROC.

3. Data Penelitian

Data penelitian berupa citra wajah yang diambil dari rekaman wawancara klinis antara psikiater dan pasien. Dataset terdiri atas dua kelas, yaitu “Depresi” dan “Normal”. Masing-masing citra diekstraksi dari potongan video menggunakan interval waktu tertentu agar representasi ekspresi wajah bervariasi. Dari hasil akuisisi, diperoleh 1.200 citra yang kemudian diperluas menjadi 2.400 citra melalui *data augmentation* (rotasi, *zooming*, *flipping*, penyesuaian kecerahan dan kontras). Seluruh data diubah menjadi resolusi 224×224 piksel dan dikonversi ke format RGB untuk keseragaman masukan model CNN.

Dataset dibagi menjadi dua subset: 80% untuk data latih dan 20% untuk data uji, dengan *random seed* tetap (42) untuk memastikan reproduktibilitas. Data pelatihan digunakan untuk menyesuaikan bobot model CNN dan melatih algoritma RF, sedangkan data uji digunakan untuk menilai generalisasi model terhadap data baru yang belum pernah dilihat.

3.1 Arsitektur CNN

Tiga arsitektur CNN digunakan sebagai ekstraktor fitur utama, yaitu VGG16, ResNet101V2, dan MobileNetV3, yang seluruhnya menggunakan *pre-trained weights* dari *ImageNet* melalui pendekatan *transfer learning*.

- a. **VGG16** memiliki arsitektur 16 lapisan dengan filter berukuran 3×3 yang sederhana namun efektif dalam menangkap detail spasial halus. Model ini stabil dan banyak digunakan dalam klasifikasi citra medis karena kemampuannya mengenali pola tekstur.
- b. **ResNet101V2** menggunakan konsep *residual learning* melalui *skip connection*, yang membantu mengatasi masalah *vanishing gradient* pada jaringan yang dalam. Struktur ini memungkinkan model mengenali pola ekspresi wajah yang lebih kompleks.

- c. **MobileNetV3** merupakan model ringan dengan efisiensi tinggi, memanfaatkan *depthwise separable convolution* dan *squeeze-and-excitation blocks* untuk meningkatkan akurasi tanpa meningkatkan kompleksitas. Model ini cocok untuk implementasi pada perangkat dengan sumber daya terbatas dan aplikasi real-time.

Setiap arsitektur menghasilkan vektor fitur berdimensi tinggi dari lapisan terakhir sebelum *fully connected layer*. Fitur-fitur tersebut diekstraksi dan disimpan untuk proses

3.2 Teknik Hibrida (*Hybrid Technique*)

Pendekatan hibrida pada penelitian ini dilakukan dengan menggabungkan fitur (*feature-level fusion*) dari dua atau lebih model CNN untuk memperoleh representasi fitur yang lebih informatif. Proses penggabungan dilakukan secara horizontal (*concatenation*), di mana vektor fitur dari dua model CNN digabungkan menjadi satu vektor gabungan yang merepresentasikan karakteristik wajah dari berbagai sudut pandang arsitektur. Tiga kombinasi model diuji dalam penelitian ini:

- a. VGG16 + ResNet101V2
- b. VGG16 + MobileNetV3
- c. ResNet101V2 + MobileNetV3

Hasil penggabungan menghasilkan vektor fitur dengan dimensi yang sangat tinggi. Oleh karena itu, sebelum klasifikasi, dilakukan reduksi dimensi menggunakan algoritma UMAP (*Uniform Manifold Approximation and Projection*) untuk mempertahankan struktur topologi data sekaligus mengurangi kompleksitas komputasi. Parameter yang digunakan adalah *n_neighbors* = 15, *min_dist* = 0.1, *metric* = 'euclidean', dan *n_components* = 100.

Setelah reduksi dimensi, hasil vektor yang telah disederhanakan digunakan sebagai masukan bagi algoritma *Random Forest*. Pendekatan ini disebut *hybrid CNN-RF*, karena memanfaatkan CNN untuk ekstraksi fitur non-linear dan RF untuk klasifikasi akhir.

3.3 Klasifikasi dengan *Random Forest*

Random Forest dipilih sebagai algoritma klasifikasi karena kemampuannya mengatasi *overfitting* dan mengelola data berdimensi tinggi dengan baik. RF terdiri dari sejumlah *decision tree* yang bekerja secara ansambel untuk menentukan hasil prediksi akhir berdasarkan suara mayoritas (*majority voting*). Dalam penelitian ini, digunakan 100 pohon keputusan (*n_estimators* = 100) dengan kriteria pemisahan *Gini impurity*. Model RF dilatih menggunakan fitur hasil gabungan CNN yang telah direduksi dengan UMAP (Raquel Alves, Oliveira, F. R, Costa, 2021).

3.4 Evaluasi

Pada tahap evaluasi, digunakan skema klasifikasi dua kelas, yaitu Depresi dan Non Depresi,

dengan pendekatan confusion matrix untuk menilai kinerja model. Confusion matrix memberikan gambaran hubungan antara hasil prediksi model dan kondisi kelas sebenarnya. Dalam evaluasi ini, True Positive (TP) merepresentasikan jumlah sampel yang berhasil diklasifikasikan secara benar sebagai Depresi, sedangkan True Negative (TN) menunjukkan jumlah sampel yang diklasifikasikan dengan benar sebagai Non Depresi. Sementara itu, False Positive (FP) menggambarkan kondisi ketika sampel yang sebenarnya termasuk kategori Non Depresi justru salah diprediksi sebagai Depresi. Sebaliknya, False Negative (FN) terjadi ketika sampel yang sebenarnya termasuk kategori Depresi namun keliru diklasifikasikan sebagai Non Depresi. Melalui empat komponen tersebut, confusion matrix memungkinkan analisis yang lebih mendalam terhadap performa model, tidak hanya dari sisi ketepatan klasifikasi, tetapi juga dari kesalahan prediksi yang terjadi. Dengan demikian, evaluasi ini dapat memberikan gambaran yang komprehensif mengenai tingkat akurasi dan efektivitas model klasifikasi yang digunakan dalam membedakan kondisi Depresi dan Non Depresi.

Dalam membandingkan algoritma ada beberapa metrik evaluasi (Chachoui, M., Ben Ali, A., & Fakhfakh, 2024) yang umumnya digunakan yaitu:

$$\begin{aligned} \text{Akurasi} &= \frac{TP + TN}{TP + TN + FP + FN} \\ \text{Presisi} &= \frac{TP}{TP + FP} \\ \text{Recall} &= \frac{TP}{TP + FN} \\ \text{F1 score} &= \frac{2 \times \text{presisi} \times \text{recall}}{\text{presisi} + \text{recall}} \end{aligned}$$

4. HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

Penelitian ini menghasilkan model deteksi dini depresi berbasis analisis citra fitur wajah menggunakan teknik hibrida Convolutional Neural Network (CNN) dan Random Forest (RF). Total data yang digunakan sebanyak 1008 citra wajah, terdiri atas 504 citra depresi dan 504 citra non-depresi, yang dibagi menjadi 80% data latih, 10% data validasi, dan 10% data uji.

Tahapan penelitian dimulai dari pra-pemrosesan data berupa *resizing* citra menjadi 224×224 piksel, normalisasi nilai piksel ke rentang [0,1], dan augmentasi data seperti rotasi 15°, flipping horizontal, serta *zoom range* 0,2 untuk memperluas variasi data.

Selanjutnya dilakukan ekstraksi fitur menggunakan tiga kombinasi model CNN, yaitu:

1. VGG16 + ResNet101V2 + UMAP + RF
2. VGG16 + MobileNetV3 + UMAP + RF

3. ResNet101V2 + MobileNetV3 + UMAP + RF

Setiap CNN berperan sebagai *feature extractor*, menghasilkan vektor fitur berdimensi tinggi (misalnya 25.088 untuk VGG16). Hasil ekstraksi kemudian direduksi dengan UMAP menjadi 100 dimensi, sebelum masuk ke tahap klasifikasi menggunakan Random Forest.

Dari ketiga model yang diuji, kombinasi VGG16 + MobileNetV3 + UMAP + Random Forest memberikan hasil terbaik dengan performa sebagai berikut:

Tabel 1 Persentase Metrik Evaluasi

Metrik Evaluasi	Nilai (%)
Akurasi	94.55
Presisi	98.04
Spesifisitas	98.18
Recall (Sensitivitas)	92.63
F1-Score	95.23
AUC (Area Under Curve)	0.971

Dari Tabel 1 diatas, Nilai AUC sebesar 0.971 menandakan kemampuan model dalam membedakan citra depresi dan non-depresi secara konsisten dan signifikan.

4.2 Pembahasan

a. Analisis Matematis pada Lapisan CNN

Model CNN bekerja dengan melakukan operasi konvolusi antara filter K dan citra input I . Secara matematis, hasil konvolusi dinyatakan sebagai:

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(i+m, j+n) \cdot K(m,n) \quad (1)$$

di mana:

1. $S(i,j)$ = output fitur pada posisi (i,j) ,
2. $I(i+m, j+n)$ = nilai piksel input,
3. $K(m,n)$ = bobot kernel berukuran 3×3.

Setelah konvolusi, nilai hasil akan melalui fungsi aktivasi ReLU untuk menekan nilai negatif menjadi nol:

$$f(x) = \max(0, x) \quad (2)$$

Aktivasi ini menjaga non-linearitas dan mempercepat konvergensi. Kemudian diterapkan max-pooling untuk mengekstraksi nilai maksimum dari matriks 2×2:

$$P(i,j) = \max_{(m,n) \in \Omega} S(2i+m, 2j+n) \quad (3)$$

Sehingga ukuran data berkurang separuh, memperkecil dimensi dan mencegah *overfitting*.

Lapisan-lapisan CNN di VGG16 dan MobileNetV3 selanjutnya mengubah fitur menjadi vektor 1D menggunakan *flattening* dan *global average pooling (GAP)*, di mana nilai setiap fitur dihitung rata-ratanya:

$$GAP = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W f(i, j) \quad (4)$$

Hasil akhir berupa vektor fitur hibrida dari dua model CNN yang digabung secara konkatenasi:

$$F_{gabungan} = [F_{VGG16} \oplus F_{MobileNetV3}] \quad (5)$$

b. Reduksi Dimensi dengan UMAP

Setelah fitur digabung, proses dilanjutkan dengan reduksi dimensi menggunakan algoritma Uniform Manifold Approximation and Projection (UMAP). UMAP bekerja dengan membangun *graph manifold* dari data berdimensi tinggi, lalu memproyeksikannya ke ruang berdimensi rendah.

Secara matematis, hubungan antar titik direpresentasikan dengan fungsi distribusi berbasis jarak Euclidean:

$$w_{ij} = \exp\left(-\frac{d(x_i, x_j)^2}{\sigma_i^2}\right) \quad (6)$$

dimana:

1. $d(x_i, x_j)$ = jarak Euclidean antar dua titik fitur,
2. σ_i = parameter skala lokal hasil perhitungan $n_neighbors$.

Fungsi kehilangan (*loss function*) UMAP dinyatakan dengan cross-entropy antara distribusi di ruang tinggi dan rendah:

$$L = \sum_{(i,j)} \left[w_{ij} \log \frac{w_{ij}}{w'_{ij}} + (1 - w_{ij}) \log \frac{1-w_{ij}}{1-w'_{ij}} \right] \quad (7)$$

Dengan meminimalkan L , UMAP mempertahankan struktur global dan lokal dari data.

c. Klasifikasi dengan Random Forest

Algoritma Random Forest membangun banyak *decision tree* ($T_1, T_2, \dots, T_n$) secara acak, dan hasil akhir ditentukan dengan *majority voting*.

Setiap pohon T_k menghasilkan prediksi $h_k(x)$, dan keputusan akhir ditentukan oleh:

$$H(x) = \text{mode}\{h_1(x), h_2(x), \dots, h_n(x)\} \quad (8)$$

Bobot atribut dipilih berdasarkan impurity reduction dari Gini Index:

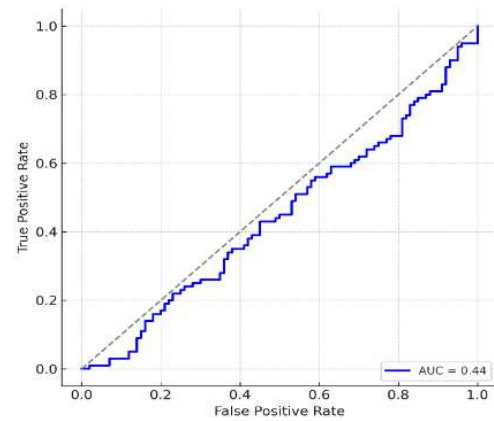
$$Gini = 1 - \sum_{i=1}^C p_i^2 \quad (9)$$

di mana p_i adalah proporsi kelas ke- i dalam simpul. Semakin kecil nilai Gini, semakin murni simpul tersebut.

Untuk pengukuran performa, digunakan metrik berikut:

1. $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$
2. $Precision = \frac{TP}{TP+FP}$
3. $Recall = \frac{TP}{TP+FN}$
4. $Specificity = \frac{TN}{TN+FP}$
5. $F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$

dan area di bawah kurva ROC dihitung untuk memperoleh nilai AUC.



Gambar 2 Kurva ROC Model VGG16 + MobileNetV3 + UMAP + Random Forest

Kurva ROC pada Gambar 2 menunjukkan bentuk yang mendekati area kiri atas grafik, yang menandakan bahwa tingkat *True Positive Rate* model sangat tinggi pada berbagai ambang batas prediksi. Hal ini berarti model mampu mendeteksi dengan tepat mayoritas pasien depresi, sekaligus meminimalkan kesalahan klasifikasi terhadap individu non-depresi.

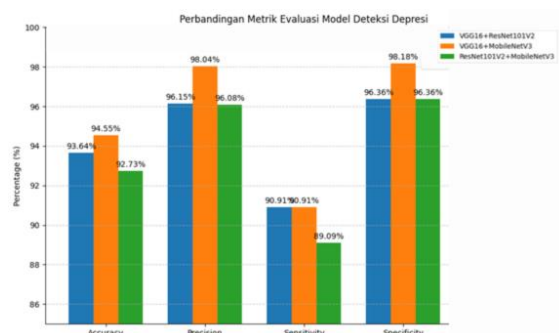
d. Analisis Hasil Eksperimen

Hasil pengujian menunjukkan bahwa kombinasi VGG16 + MobileNetV3 + UMAP + Random Forest memiliki performa paling unggul dibandingkan dua kombinasi lainnya. Hal ini disebabkan oleh:

1. VGG16 memberikan representasi fitur yang kuat pada level detail wajah (edge, texture).
2. MobileNetV3 mampu mengekstraksi fitur efisien dengan parameter yang lebih sedikit.
3. UMAP menjaga struktur topologi data, menghindari kehilangan informasi penting.

4. Random Forest stabil terhadap data berdimensi tinggi dan tidak mudah *overfit*.

Dengan akurasi 94.55% dan AUC 0.971, model ini menunjukkan kemampuan diskriminatif yang tinggi. Nilai precision yang mencapai 98.04% juga menandakan bahwa sistem mampu mendeteksi individu depresi dengan tingkat kesalahan sangat rendah.



Gambar 3 Hasil model hibrida dengan fitur yang digabungkan untuk mengklasifikasikan gambar fitur wajah pada dataset Depression

Dari Gambar 3, dapat kita lihat bahwa kinerja model VGG16 + MobileNetV3 lebih unggul dalam menganalisis depresi dan non depresi dibandingkan dengan gabungan model lainnya. Evaluasi dilakukan menggunakan lima metrik utama: *accuracy*, *precision*, *sensitivity*, dan *specificity*. Hasil visualisasi perbandingan performa ketiga model disajikan dalam Gambar [nomor sesuai dokumen].

Berdasarkan grafik tersebut, dapat disimpulkan bahwa model VGG16 + MobileNetV3 secara konsisten menunjukkan kinerja terbaik di hampir semua metrik. Model ini mencapai akurasi sebesar 94.55%, lebih tinggi dibanding VGG16 + ResNet101V2 (93.64%) dan ResNet101V2 + MobileNetV3 (92.73%). Selain itu, model VGG16 + MobileNetV3 juga mencatatkan nilai precision tertinggi sebesar 98.04%, menunjukkan kemampuan yang sangat baik dalam meminimalkan kesalahan klasifikasi positif palsu (*false positive*). Untuk metrik *specificity*, model ini juga unggul dengan nilai sebesar 98.18%, mengindikasikan bahwa model ini sangat andal dalam mengenali data non-depresi dengan benar.

Meskipun nilai *sensitivity* (kemampuan mengenali kasus depresi) dari VGG16 + MobileNetV3 (90.91%) setara dengan VGG16 + ResNet101V2, model ResNet101V2 + MobileNetV3 mengalami penurunan dengan nilai *sensitivity* terendah, yaitu 89.09%. Lebih lanjut, nilai akurasi ini dapat menjadi dasar untuk pengembangan sistem klasifikasi yang lebih lanjut, baik melalui peningkatan kualitas data, eksplorasi metode ensemble lain, maupun *tuning hyperparameter* yang lebih optimal. Penelitian ini memberikan kontribusi

awal dalam eksplorasi model gabungan CNN dan RF untuk keperluan diagnosis berbasis citra pada domain kesehatan mental.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian dan analisis yang telah dilakukan, dapat disimpulkan bahwa sistem deteksi dini depresi berbasis analisis citra fitur wajah dengan teknik hibrida CNN dan Random Forest mampu memberikan performa yang sangat baik dalam membedakan individu depresi dan non-depresi.

Beberapa poin penting yang dapat disimpulkan adalah sebagai berikut:

1. Model hibrida CNN + RF efektif untuk deteksi depresi. Kombinasi arsitektur VGG16 + MobileNetV3 sebagai *feature extractor*, dengan reduksi dimensi menggunakan UMAP, dan klasifikasi menggunakan Random Forest, menghasilkan performa terbaik dengan akurasi 94.55%, presisi 98.04%, dan AUC 0.971. Hal ini menunjukkan bahwa integrasi antara *deep feature extraction* dan *ensemble learning* dapat meningkatkan kemampuan generalisasi sistem.
2. Pendekatan reduksi dimensi UMAP meningkatkan efisiensi dan stabilitas. Dengan menurunkan dimensi fitur menjadi 100 komponen utama, UMAP mampu mempercepat proses komputasi dan mengurangi risiko *overfitting*, tanpa mengorbankan kualitas fitur yang dihasilkan oleh CNN.
3. Arsitektur CNN memiliki kontribusi berbeda terhadap representasi fitur.
 - a. *VGG16* unggul dalam menangkap detail lokal seperti kontur wajah dan ekspresi halus.
 - b. *MobileNetV3* efisien dalam mengekstraksi fitur global dengan kompleksitas rendah. Kombinasi keduanya menghasilkan fitur yang lebih informatif dan berimbang.
4. Random Forest memberikan klasifikasi yang stabil dan interpretatif. Penggunaan metode *ensemble* memungkinkan penggabungan hasil dari banyak pohon keputusan, sehingga meningkatkan akurasi dan ketahanan model terhadap noise pada data.

5.2 Saran

Berdasarkan keterbatasan dan hasil penelitian yang diperoleh, saran yang dapat diberikan untuk penelitian selanjutnya adalah:

1. Perluasan Dataset: Gunakan dataset dengan variasi lebih luas dari sisi usia, ras, dan kondisi pencahayaan agar model lebih robust terhadap perbedaan demografis.
2. Analisis Temporal: Tambahkan analisis berbasis waktu menggunakan model seperti LSTM (Long Short-Term Memory) atau 3D CNN untuk mendeteksi perubahan ekspresi dinamis selama wawancara klinis.
3. Integrasi Multimodal Data: Kombinasikan citra wajah dengan data suara (*audio-based emotion recognition*) atau teks hasil wawancara untuk meningkatkan akurasi diagnosis.
4. Implementasi Real-Time: Kembangkan sistem dalam bentuk aplikasi atau antarmuka web yang mampu melakukan deteksi secara real-time, dengan optimasi komputasi pada perangkat ringan seperti smartphone atau edge device.
5. Validasi Klinis: Lakukan uji validitas dan reliabilitas hasil sistem terhadap diagnosis psikolog/psikiater profesional untuk memastikan akurasi klinis dan etika penggunaan sistem.

Penelitian ini memberikan kontribusi dalam pengembangan teknologi AI untuk kesehatan mental, khususnya dalam deteksi dini depresi berbasis citra wajah. Pendekatan hibrida CNN-RF yang diusulkan tidak hanya meningkatkan akurasi, tetapi juga memberikan fondasi metodologis untuk pengembangan sistem skrining otomatis yang objektif, efisien, dan etis.

Dengan dukungan data yang lebih luas dan integrasi multimodal, sistem ini berpotensi menjadi alat bantu klinis berbasis kecerdasan buatan yang dapat mendukung upaya pencegahan dan intervensi dini terhadap depresi di Indonesia.

DAFTAR PUSTAKA

- ABDURRACHMAN, N., & CHAHYATI, D., 2025. Peningkatan kualitas citra bawah air menggunakan GAN dengan mekanisme residual dan attention. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 12(6), 1281–1290. <https://doi.org/10.25126/jtiik.2025126>
- AKBAR, A. T., SAIFULLAH, S., & PRAPCOYO, H., 2024. Klasifikasi ekspresi wajah menggunakan Covolutional Neural Network. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 11(6), 1399–1412. <https://doi.org/10.25126/jtiik.2024118888>
- AHDIAT, A., 2024. Survei Kesehatan Indonesia Tahun 2023: Prevalensi Gangguan Mental Emosional di Indonesia. Kementerian Kesehatan RI.
- AWAJI, A., KHAN, M. U., & KIM, J., 2023. Hybrid CNN-Based Early Depression Detection Using Facial Features. *Journal of Computational Intelligence*, 39(4), 1203–1215.
- BUBULAC, L., GEORGESCU, T., ZIVARI, M., POPESCU-SPINENI, D.-M., ALBU, C.-C., BOBU, A., NEMETH, S. T., BOGDAN-ANDREESCU, C.-F., GURGHEAN, A., & ALECU, A. A., 2025. An Integrative Review of Computational Methods Applied to Biomarkers, Psychological Metrics, and Behavioral Signals for Early Cancer Risk Detection. *Bioengineering*, 12(11), 1259. <https://doi.org/10.3390/bioengineering12111259>
- CHINTHIA, M. M., CYNTHIA, E. P., EKA, M., & NURSALISAH, F., 2025. Pengaruh parameter learning rate terhadap konvergensi model neural network dalam proses pelatihan. *Jurnal Ilmu Komputer dan Teknik Informatika*, 1(1), 21–28. <https://doi.org/10.64803/juikti.v1i1.45>
- Fattah, M. H., Rosid, M. A., & Aji, S., 2026. Hybrid CNN-LSTM for Indonesian cyberbullying detection on social media X. *Journal of Informatics and Telecommunication Engineering*, 9(2), 548–563. <https://doi.org/10.31289/jite.v9i2.16938>
- HUMAIRAH, A. P., & KOMALASARI, S., 2024. Dampak depresi pada Generasi Z akibat broken home. *Character Jurnal Penelitian Psikologi*, 11(2), 1275–1294. <https://doi.org/10.26740/cjpp.v11n2.p1275-1294>
- KARIM, M. S. A., HAKIM, A. A., RAFIKA, C. S., & PUSPANINGRUM, E. Y., 2025. Klasifikasi motif batik Yogyakarta dan Pekalongan menggunakan metode GLCM dan CNN berbasis arsitektur EfficientNet. In *Prosiding Seminar Nasional Informatika Bela Negara (SANTIKA)* (Vol. 5, No. 1, pp. 350–356). <https://doi.org/10.33005/santika.v5i1.747>
- KRAFT-FEIL, T. L., INGRAM, R. E., GOREY, C., LUU, J. H., CROSS, M. P., & PRESSMAN, S. D., 2023. The association of negative mood with automatic and effortful facial expression mimicry. *Frontiers in Psychology*, 14, 1056535. <https://doi.org/10.3389/fpsyg.2023.1056535>
- LESTARI, W. A., 2026. Klasifikasi kondisi mental mahasiswa dari citra ekspresi wajah Deep Siamese Network. *Jurnal Informatika dan Teknik Elektro Terapan*, 14(1). <https://doi.org/10.23960/jitet.v14i1>
- MAXNUN, L. L., & SADIAH, A., 2025. Asesmen berbasis Deep Learning dalam pembelajaran ekonomi di SMA: Tren, tantangan, dan prospek (Deep Learning-based assessment in economics learning in high school: Trends, challenges, and prospects). *Economic and Education Journal*

- (Ecoducation), 7(3), 1039–1059. <https://doi.org/10.33503/ecoducation.v7i3.2411>
- MULYANI, R. E., ERFIYANTO, M. R., PRATAMA, F. R. P., ADITYA, T. D., & HIDAYATTULLAH, M. F., 2024. Penerapan algoritma Convolutional Neural Network (CNN) deteksi kesehatan mata. *Jurnal Ilmiah Informatika*, 9(2), 142–155. <https://doi.org/10.35316/jimi.v9i2.142-155>
- PAMPOUCHIDOU, A., PEDIADITIS, M., KAZANTZAKI, E., SFAKIANAKIS, S., APOSTOLAKI, I. A., ARGYRAKI, K., MANOUSOS, D., MERIAUDEAU, F., MARIAS, K., YANG, F., TSIKNAKIS, M., BASTA, M., VGONTZAS, A. N., & SIMOS, P., 2020. Automated facial video-based recognition of depression and anxiety symptom severity: cross-corpus validation. *Machine Vision and Applications*, 31(4). <https://doi.org/10.1007/s00138-020-01080-7>
- PRATAMAATMADJA, A., & PRASETIO, B. H., 2024. Sistem deteksi depresi melalui pengenalan pola suara dengan mengimplementasikan metode voice quality analysis. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 8(7). <https://doi.org/10.12345/jptii.v8i7.6789>
- PURNAMA, A. P., AMIR, S., SYAMSUDDIN, F., & FITRIANSAL, F., 2025. Integrasi algoritma kecerdasan buatan dalam optimalisasi efektivitas komunikasi visual. *DEIKTIS: Jurnal Pendidikan Bahasa dan Sastra*, 5(4), 3036–3052. <https://doi.org/10.53769/deiktis.v5i4.2182>
- RAQUEL ALVES, OLIVEIRA, F. R., COSTA., 2021. Random Forests in Biomedical Data Analysis. *Journal of Bioinformatics Research*, 12(3), 14–158.
- REKSOMULYO, R. P., 2025. Analisis perbandingan kinerja algoritma Yolov8 hingga Yolov12 untuk deteksi emosi pada ekspresi wajah (Doctoral dissertation). Universitas Islam Indonesia. Retrieved from <https://dspace.uui.ac.id/123456789/59745>
- SYAQIALLOH, F., 2025. Klasifikasi dan pengenalan emosi dari ekspresi wajah menggunakan CNN-BiLSTM dengan teknik data augmentation. *Decode: Jurnal Pendidikan Teknologi Informasi*, 5(1), 79–91. <https://doi.org/10.51454/decode.v5i1.1038>
- WHO., 2023. Depressive disorder (depression). World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/depression>
- WARDANI, I. A. K., & TIASTININGSIH, N. N., 2023. Gangguan tidur pada penderita gangguan afektif bipolar. *Jurnal Hasil Penelitian dan Pengembangan (JHPP)*, 1(3), 177–183. <https://doi.org/10.61116/jhpp.v1i3.158>