



Analisis Kinerja Random Forest Dalam Deteksi Gejala Alergi Rongga Mulut Berbasis Warna Gusi

Juli Hartati Gea^{*}, Agustinus Rudatyo Himamunanto, Haeny Budiati

Fakultas Fisika dan Komputer, Program Studi Informatika, Universitas Kristen Immanuel, Yogyakarta, Indonesia

Email: ^{1,*}julihartatigea@gmail.com, ²rudatyo@gmail.com, ³heny@ukrim.ac.id

Email Penulis Korespondensi: julihartatigea@gmail.com

Abstrak—Deteksi dini alergi di rongga mulut masih menjadi tantangan karena penilaian visual bersifat subjektif dan keterbatasan fasilitas diagnostik. Penelitian ini menawarkan pendekatan baru dengan algoritma Random Forest untuk mengklasifikasikan tingkat keparahan gejala alergi berdasarkan analisis warna gusi pada citra digital. Sebanyak 2.742 gambar gusi dikategorikan secara klinis menggunakan skor Modified Gingival Index (MGI) menjadi kondisi ringan, sedang, dan berat. *Preprocessing* mencakup konversi ruang warna HSV, segmentasi adaptif menggunakan threshold warna merah pada kanal hue 0–10 dan 160–180, saturation > 50, dan value > 40. Fitur statistik yang meliputi rata-rata, standar deviasi, kemiringan, kurtosis, entropi diekstrak dan dinormalisasi menggunakan Z-Score. Enam kombinasi parameter diuji dengan rasio data latih-ujii sebesar 80:20. Konfigurasi optimal dengan *n_estimators*=80, *max_depth*=9, *min_samples_leaf*=2 mencapai akurasi 95,81%. Kinerja tertinggi tercapai pada kelas ringan dengan precision dan recall sebesar 98,91%, serta stabil pada kelas sedang yaitu 93,80% dan berat sebesar 94,74% dengan selisih hanya 0,94%. Evaluasi *Cross validation* menunjukkan stabilitas model yang sangat baik, dengan akurasi rata-rata 95,30% dan standar deviasi 0,67%, menandakan konsistensi performa antar subset data. Analisis *importance* fitur menunjukkan dominasi kanal *hue* dan *saturation*, khususnya fitur kurtosis dan rerata kejenuhan warna. Hasil penelitian ini membuktikan bahwa sistem deteksi alergi berdasarkan warna gusi dengan *Random Forest* sangat akurat dan efektif sebagai skrining non-invasif di wilayah dengan sumber daya terbatas yang berpotensi meningkatkan akses skrining awal di fasilitas kesehatan primer.

Kata Kunci: Deteksi Alergi; Warna Gusi; HSV; Random Forest; Machine Learning

Abstract—Early detection of allergies in the oral cavity remains challenging due to the subjective nature of visual assessment and limited access to diagnostic facilities. This study proposes a novel approach using the Random Forest algorithm to classify the severity of allergic symptoms based on gum color analysis from digital images. A total of 2,742 gum images were clinically categorized using the Modified Gingival Index (MGI) into mild, moderate, and severe conditions. Preprocessing included conversion to HSV color space and adaptive segmentation using red thresholds on the hue channel (0–10 and 160–180), saturation > 50, and value > 40. Statistical features, including mean, standard deviation, skewness, kurtosis, and entropy, were extracted and normalized using Z-Score. Six parameter combinations were tested with an 80:20 train-test split. The optimal configuration with *n_estimators*=80, *max_depth*=9, and *min_samples_leaf*=2 achieved an accuracy of 95.81%. The highest performance was achieved in the mild class with precision and recall of 98.91%, and stable results in the moderate (93.80%) and severe (94.74%) classes, with only a 0.94% difference. Cross-validation evaluation demonstrated excellent model stability, with an average accuracy of 95.30% and a standard deviation of 0.67%, indicating consistent performance across data subsets. Feature importance analysis showed the dominance of the hue and saturation channels, particularly kurtosis and mean saturation. This study demonstrates that a Random Forest-based allergy detection system using gum color is highly accurate and effective as a non-invasive screening tool in dental and oral health, especially in resource-limited settings, with the potential to improve early screening access in primary healthcare facilities.

Keywords: Allergy Detection; Gum Color; HSV; Random Forest; Machine Learning

1. PENDAHULUAN

Deteksi penyakit alergi di rongga mulut merupakan tantangan dalam kesehatan masyarakat yang prevalensinya terus meningkat secara global. Salah satu jenis manifestasi alergi di rongga mulut yaitu stomatitis alergika berupa peradangan pada mukosa mulut yang disebabkan oleh paparan alergen seperti makanan, logam ortodontik, obat-obatan, bahan restorasi gigi, dan bahan akrilik [1]. Data epidemiologi menunjukkan prevalensi *oral allergy syndrome* (OAS) berkisar antara 5–26% pada populasi dengan riwayat alergi. Prevalensi gingivitis terus meningkat secara global, mencapai 70–75% populasi [2]. Selain itu *World Allergy Organization* (WAO) melaporkan prevalensi reaksi hipersensitivitas atau *oral allergy*, secara umum sekitar 22% populasi di dunia, dengan variasi 2% hingga 83% tergantung wilayah dan kelompok.

Gingiva atau gusi merupakan jaringan lunak pada rongga mulut yang membungkus leher dan menutupi tulang alveolar [3]. Warna gingiva normal umumnya merah muda pucat, namun dapat mengalami perubahan menjadi merah terang, ungu kemerahan, atau bahkan kehitaman akibat peradangan, iritasi, atau reaksi alergi [4]. Perubahan warna ini terjadi karena vasodilatasi pembuluh darah dan infiltrasi sel inflamasi sebagai respons terhadap alergen, yang mengakibatkan peningkatan aliran darah dan akumulasi mediator inflamasi di jaringan gingiva.

Beberapa penelitian sebelumnya menunjukkan bahwa perubahan warna gingiva dapat diinduksi oleh faktor eksternal seperti nikotin dan alergen. Sebuah penelitian yang menunjukkan adanya hubungan antara pigmentasi gingiva dan kebiasaan merokok. Studi tentang pigmentasi gusi pada perokok menunjukkan peningkatan warna gelap pada area mandibular dan maxillar, yang mengindikasikan bahwa warna gusi merupakan indikator klinis penting terhadap kondisi patologis, termasuk reaksi alergi [4].

Meskipun berbagai penelitian telah mengeksplorasi perubahan warna gusi dalam konteks penyakit periodontal dan kebiasaan merokok, belum ada penelitian yang secara spesifik menganalisis korelasi antara warna gusi dengan manifestasi alergi sistemik. Hal ini menciptakan gap penelitian yang signifikan, mengingat stomatitis alergika sering menjadi



manifestasi awal dari reaksi hipersensitivitas [1]. Deteksi dini melalui analisis warna gusi dapat menjadi *screening tool* non-invasif yang *cost-effective*, terutama di daerah dengan keterbatasan akses ke spesialis alergi-imunologi [5].

Deteksi dini manifestasi alergi pada gingiva sangat penting untuk mencegah komplikasi sistemik yang lebih serius. Namun, pemeriksaan klinis terhadap warna gusi masih sangat subjektif, dipengaruhi oleh kondisi pencahayaan dan pengalaman klinisi. Masalah utama dalam mendeteksi alergi pada gingiva tidak hanya terletak pada kesulitan identifikasi perubahan warna secara objektif, tetapi juga pada kesulitan mendiagnosis alergi, yang memerlukan pengujian khusus seperti *skin prick test* atau test IgE spesifik. Variabilitas interpretasi visual antar praktisi, perbedaan kondisi pencahayaan, dan keterbatasan infrastruktur diagnostik di daerah terpencil memperburuk situasi ini. Pendekatan berbasis *machine learning* dapat menjadi solusi untuk mengatasi limitasi tersebut sambil memberikan standarisasi objektif dalam proses diagnosis. *Random forest* adalah salah satu algoritma *machine learning* yang meningkatkan akurasi prediksi yang kuat dengan menggabungkan keunggulan dari beberapa pohon keputusan [6]. Metode ini akan membentuk sejumlah pohon, dimana satu pohon akan memberikan satu unit suara [7]. Karena menghasilkan banyak pohon keputusan, algoritma ini disebut sebagai "*forest*" atau hutan [8].

Penelitian lain telah menunjukkan efektivitas algoritma *Random Forest* dalam klasifikasi penyakit gigi dan mulut. Penelitian ini membandingkan performa tiga algoritma berbasis *tree Decision Tree*, *Random Forest*, dan *XGBoost* dalam mendeteksi penyakit infeksi gigi dan mulut seperti gingivitis, pulpitis, periodontitis, dan abses odontogenik. Dalam penelitian tersebut, *Random Forest* berhasil mencapai akurasi sebesar 91,7% tanpa SMOTE, dan meningkat hingga 93,1% setelah penanganan ketidakseimbangan kelas menggunakan teknik SMOTE (*Synthetic Minority Oversampling Technique*) [8]. Hasil ini menunjukkan bahwa *Random Forest* memiliki kemampuan yang sangat baik dalam mengklasifikasikan kondisi gigi dan mulut berdasarkan data klinis gejala sekaligus menjadi landasan untuk mengembangkan aplikasi serupa dalam mendeteksi alergi berbasis perubahan warna gusi.

Tujuan penelitian ini adalah mengembangkan sistem klasifikasi berbasis algoritma *Random Forest* untuk mendeteksi dan mengklasifikasikan tingkat keparahan gejala alergi pada gusi manusia berdasarkan analisis warna dalam ruang HSV (*Hue, Saturation, Value*). Selain itu, penelitian ini bertujuan mengevaluasi performa model *Random Forest* dengan mengukur akurasi, presisi, dan *recall* pada berbagai kombinasi parameter serta rasio pembagian data latih dan uji. Penelitian juga menganalisis kontribusi masing-masing fitur warna *hue*, *saturation*, dan *value* terhadap hasil prediksi menggunakan teknik *feature importance*. Untuk memastikan konsistensi sistem, model diuji stabilitasnya melalui metode *cross-validation 10-fold*. Dengan pendekatan ini, diharapkan sistem deteksi alergi berbasis warna gusi dapat memberikan hasil yang cepat, akurat, dan objektif, sehingga mendukung upaya deteksi dini yang efektif di bidang kesehatan rongga mulut.

2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *Random Forest* untuk mendeteksi gejala alergi berdasarkan perubahan warna gusi manusia. Data warna gusi akan dianalisis dan diolah menggunakan teknik *image processing* untuk kemudian dijadikan input pada model *Random Forest* guna menentukan apakah terjadi gejala alergi atau tidak. Algoritma ini dipilih karena kemampuannya dalam menangani dataset dengan dimensi tinggi dan menghasilkan prediksi yang akurat.

Penelitian ini diawali dengan proses pengumpulan data, di mana gambar gusi diperoleh dari dua sumber dataset yaitu *Mendeley Data* untuk kelas ringan dan *Kaggle* untuk kelas sedang dan berat. Setiap gambar diklasifikasi berdasarkan kondisi visual gusi menjadi tiga kelas: ringan, sedang, dan berat. Tahap selanjutnya adalah *preprocessing* data yang meliputi *resize* gambar, *white balancing*, konversi ke ruang warna HSV, segmentasi area gusi berdasarkan warna merah, serta *filtering* untuk mengurangi *noise*. Setelah *preprocessing* selesai, dilakukan ekstraksi fitur warna dari area gusi yang telah disegmentasi. Selanjutnya, dilakukan normalisasi data menggunakan metode *Z-Score normalization* agar seluruh fitur memiliki skala yang sebanding dan distribusi yang stabil. Tahapan berikutnya adalah klasifikasi menggunakan algoritma *Random Forest*, yang bertujuan untuk membedakan kondisi gusi berdasarkan pola warna yang terdeteksi. Setelah model dikembangkan, dilakukan evaluasi metode menggunakan metrik klasifikasi seperti akurasi, presisi, *recall*, dan *F1-score* terhadap data uji, guna mengetahui performa model dalam mendeteksi gejala alergi. Selain itu, digunakan juga *cross-validation 10-fold* untuk menguji konsistensi dan stabilitas model terhadap berbagai subset data. Tahap akhir dari proses ini adalah analisis *feature importance*, yang dilakukan untuk mengukur kontribusi masing-masing fitur HSV terhadap prediksi model. Tahapan penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data

Dataset yang digunakan dalam penelitian ini merupakan data sekunder yang dikombinasikan dari dua sumber utama untuk membentuk tiga kelas kondisi gusi yaitu ringan, sedang, dan berat. Kelas ringan diperoleh dari dataset dengan anotasi



Modified Gingival Index (MGI) yang tersedia di *Mendeley Data* (DOI: <https://doi.org/10.17632/3253gj88rr.1>). Dataset ini berisi citra intraoral dengan skor MGI 0 hingga 4. Untuk penelitian ini, gambar dengan skor MGI 0 dan 1 dipilih karena secara klinis mencerminkan kondisi gusi sehat hingga ringan yaitu kemerahan ringan tanpa pembengkakan [9]. Sementara itu, kelas Sedang dan Berat diperoleh dari dataset *Oral Diseases* yang tersedia di *Kaggle* (<https://www.kaggle.com/datasets/salmansajid05/oral-diseases>), yang mencakup citra rongga mulut dengan berbagai penyakit oral, termasuk gingivitis. Citra dari dataset ini kemudian diklasifikasikan ulang berdasarkan kriteria objektif yang telah ditetapkan [10]. Klasifikasi dilakukan berdasarkan penilaian visual serta referensi skor MGI, yaitu untuk skor MGI 0-1 dikategorikan sebagai kelas ringan, mencakup kemerahan ringan tanpa pembengkakan. Skor MGI 2-3 dikategorikan sebagai kelas sedang, mencakup kemerahan lokal hingga menyeluruh dan pembengkakan ringan, dan skor MGI 4 dikategorikan sebagai kelas berat, ditandai dengan inflamasi parah, pembengkakan jelas, dan perubahan tekstur jaringan gusi.

Tabel 1. Distribusi Data

Kelas Kondisi	Jumlah Gambar	Sumber Dataset
Ringan	914	Mendeley Data (MGI 0-1)
Sedang	1.101	Kaggle - Oral Diseases (reklasifikasi)
Berat	1.248	Kaggle - Oral Diseases (reklasifikasi)
Total	3.263	Kombinasi kedua sumber

Pada Tabel 1 menunjukkan distribusi dataset yang tidak seimbang. Jumlah data awal yang terkumpul terdiri dari 914 gambar untuk kelas ringan, 1.248 gambar untuk kelas sedang, dan 1.101 gambar untuk kelas berat. Untuk menjaga keseimbangan kelas dan menghindari bias klasifikasi, dilakukan proses *undersampling* terhadap kelas sedang dan berat, sehingga masing-masing kelas memiliki jumlah sampel yang sama, yaitu 914 gambar. Total keseluruhan dataset yang digunakan untuk pelatihan dan pengujian model adalah 2.742 gambar, dengan distribusi seimbang sebanyak 914 gambar per kelas.

2.2 Preprocessing Data

Data yang telah dikumpulkan dilakukan tahap preprocessing untuk meningkatkan kualitas citra dan mempersiapkan data agar optimal saat proses pelatihan model. Proses ini mencakup beberapa tahapan penting sebagai berikut.

a. *Resize*

Resize adalah melakukan pengukuran dimensi gambar agar model memproses gambar dengan ukuran yang sama [10].

b. *White Balancing dan Histogram Equalization*

Proses *white balancing* dilakukan menggunakan pemisahan *channel* LAB, di mana *channel* L (*luminance*) digunakan untuk mengoreksi pencahayaan, sementara A dan B untuk mengoreksi warna. *White balancing* bertujuan menstabilkan warna secara adaptif, sehingga warna merah alami pada gusi dapat ditampilkan secara akurat. Kemudian *histogram equalization* diterapkan pada *channel luminance* (Y) untuk meningkatkan kontras area gusi.

c. Konversi ke Ruang Warna HSV

Citra dikonversi dari ruang warna RGB ke HSV (*Hue, Saturation, Value*). HSV dipilih karena lebih representatif untuk mendeteksi perubahan warna pada jaringan biologis seperti gusi.

d. Segmentasi Area Gusi

Dilakukan dengan pendekatan *thresholding* pada dua rentang warna merah dalam ruang HSV untuk memisahkan area gusi dari latar belakang. Segmentasi area gusi dilakukan dengan teknik *thresholding* berdasarkan dua rentang warna merah, yaitu:

1. *Hue*: 0–10 dan 160–180
2. *Saturation*: ≥ 50
3. *Value*: ≥ 50

Rentang ini dipilih berdasarkan spektrum warna merah alami pada jaringan gusi yang mengalami peradangan, dan telah ditentukan secara empiris melalui pengamatan visual terhadap sampel dataset.

e. *Filtering Noise*

Filtering Noise bertujuan untuk mengurangi noise tanpa menghilangkan detail penting. Dalam penelitian ini, digunakan median filter atau bilateral filter karena keduanya efektif untuk mempertahankan tepi objek dan kontur warna, yang sangat penting dalam analisis tekstur dan distribusi warna gusi.

f. Operasi Morfologi

Menggunakan teknik opening dan closing untuk memperhalus bentuk area gusi hasil segmentasi dan menghilangkan bagian kecil yang tidak relevan. Langkah ini penting agar fitur yang diekstraksi dari citra setelah segmentasi benar-benar berasal dari area gusi yang valid secara visual, tanpa gangguan bentuk yang disebabkan oleh *noise* atau ketidakaturan hasil segmentasi awal.

2.3 Ekstraksi Fitur

Feature extraction merupakan suatu teknik mengambil fitur ciri pada sebuah bentuk citra agar tahapan eksperimen berikutnya memiliki hasil yang maksimal. Beberapa fitur yang berbeda, termasuk ciri warna dan tekstur, dapat digunakan untuk ekstraksi fitur [11]. Setelah segmentasi, dilakukan ekstraksi fitur dari area gusi yang telah terisolasi. Fitur yang

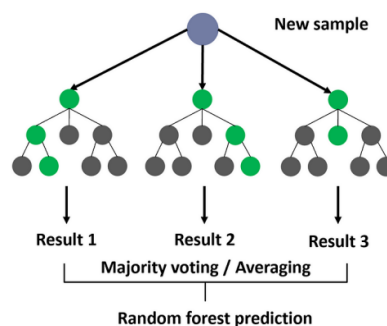
diambil adalah nilai statistik dari setiap kanal warna HSV, meliputi: *mean* (rata-rata), standar deviasi, *skewness* (kemiringan distribusi), *kurtosis* (keruncingan distribusi). Ekstraksi ini bertujuan untuk mendeteksi perubahan warna yang dapat menjadi indikator adanya alergi pada gusi. Proses ekstraksi dilakukan menggunakan pustaka *NumPy* dan *OpenCV*.

2.4 Normalisasi Fitur

Normalisasi Data adalah proses dilakukan transformasi sebuah atribut numerik ke bentuk yang lebih sederhana berskala 0 sampai 1 [12]. Untuk memastikan bahwa seluruh fitur memiliki skala yang sebanding dan tidak mendominasi satu sama lain, dilakukan normalisasi fitur menggunakan dua metode yaitu *min-max scaling* untuk meratakan nilai ke rentang 0–1 dan *Z-Score normalization* untuk menjadikan distribusi data memiliki rata-rata 0 dan standar deviasi 1. Proses ini bertujuan untuk meningkatkan stabilitas dan akurasi model saat proses pelatihan.

2.5 Implementasi *Random Forest*

Tahap utama dalam penelitian ini adalah mengimplementasikan *Random Forest* untuk membuat model klasifikasi. Pada tahap ini, dataset dibagi menjadi dua bagian yaitu, data pelatihan (*training set*) dan data pengujian (*test set*) dengan persentase 80% dari data untuk latihan dan 20% untuk pengujian.



Gambar 2. *Random Forest*

Pada Gambar 2 menggambarkan algoritma *Random Forest* yang merupakan beberapa pohon keputusan dihasilkan dari sampel set pelatihan, menjadikan metode *Random Forest* sebagai pendekatan berbasis pohon keputusan [13]. *Random Forest* membangun pohon keputusan *ensemble* menggunakan teknik *bagging*. Penelitian Breiman menyoroti sejumlah manfaat *Random Forest*, seperti kemampuannya untuk menghasilkan hasil klasifikasi akhir berkualitas tinggi dengan kesalahan yang berkurang, kemampuannya untuk menangani sejumlah besar data secara lebih efektif, dan efektivitasnya sebagai salah satu algoritma untuk masalah data yang hilang [14].

Keuntungan dari menggunakan *Random Forest* adalah model yang dihasilkan dapat memiliki tingkat akurasi yang tinggi dan dapat mengatasi *overfitting*. Selain itu, *Random Forest* juga dapat digunakan untuk mengevaluasi pentingnya fitur (*feature importance*) pada dataset [15]. Hal ini dapat membantu dalam pemilihan fitur yang relevan dan meningkatkan kinerja model. Namun, kelemahan dari *Random Forest* adalah kompleksitas yang tinggi dan waktu pelatihan yang relatif lambat pada dataset yang sangat besar [15]. Setelah *preprocessing*, data pelatihan akan dimasukkan ke dalam algoritma *Random Forest* untuk proses pelatihan. Pada tahap ini:

- Model akan mempelajari pola dalam data dengan menggunakan fitur-fitur warna yang diekstraksi, dan memetakan pola ini untuk membedakan kondisi gusi yang sehat dan yang menunjukkan gejala alergi.
- Model juga akan melakukan penyesuaian pada parameter seperti jumlah pohon (*trees*) dan kedalaman (*depth*) untuk mengoptimalkan kinerjanya.
- Pemrosesan ini bertujuan untuk membuat model mengenali pola perubahan warna gusi yang dapat mengindikasikan adanya gejala alergi.

Klasifikasi dilakukan menggunakan algoritma *Random Forest*, yaitu metode *ensemble* berbasis *decision tree* yang bekerja dengan teknik *bagging*. Proses pelatihan melibatkan:

- Pemilihan parameter jumlah pohon (*n_estimators*) dan kedalaman pohon (*max_depth*).
- Pembentukan beberapa pohon keputusan secara acak.
- Voting mayoritas dari hasil masing-masing pohon.

Implementasi sistem dilakukan dengan menggunakan bahasa pemrograman *python* dan pustaka seperti *OpenCV* untuk pemrosesan citra dan *Scikit-learn* untuk penerapan algoritma *Random Forest*. Tahap ini meliputi penulisan kode untuk setiap langkah dalam alur kerja sistem.

2.6 Evaluasi Model

Setelah implementasi, sistem akan diuji dengan data uji untuk mengukur performa model. Pengujian ini menggunakan matrik evaluasi seperti akurasi, *presisi*, *recall*, dan *F1-score* guna mengetahui seberapa baik metode *Random Forest* dalam mendeteksi gejala alergi berdasarkan warna gusi. Ini adalah langkah penting untuk memastikan performa model dalam mendeteksi gejala alergi dengan akurat.



		Actual Values	
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	TP (True Positive)	FP (False Positive) <i>Type I Error</i>
	0 (Negative)	FN (False Negative) <i>Type II Error</i>	TN (True Negative)

Gambar 3. Confusion Matrix

Pada Gambar 3 menunjukkan *confusion matrix* yang merupakan metode untuk melihat rincian klasifikasi, serta analisis fitur penting yang mengidentifikasi elemen-elemen kunci dalam prediksi [7]. Tabel yang digunakan untuk menilai kinerja model klasifikasi *machine learning* disebut *matrix confusion*. Matriks ini menyajikan ringkasan hasil prediksi model pada sebuah dataset, memungkinkan kita mengevaluasi seberapa akurat atau salah model dalam mengklasifikasikan kumpulan data [7]. Akurasi Klasifikasi merupakan kinerja total dari suatu teknik klasifikasi ditunjukkan oleh akurasi klasifikasi [16].

Ukuran dalam mengevaluasi kinerja model berdasarkan *confusion matrix* ada berbagai macam, diantaranya adalah *accuracy*, *precision*, dan *F1 score*. *Accuracy*, *precision*, dan *F1 score* dapat diperoleh sebagai persamaan (1),(2),(3) dan (4) berikut [16]:

- a. Akurasi (*Accuracy*) : Merupakan proporsi dari total data yang diprediksi dengan benar oleh model, baik positif maupun negatif.

$$\frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

- b. Presisi (*Precision*) : Merupakan proporsi dari data positif yang diprediksi benar oleh model dari total data yang diprediksi positif.

$$\frac{TP}{TP+FP} \quad (2)$$

- c. *Recall* : ini mengukur seberapa banyak dari semua kasus positif yang diidentifikasi oleh model. Perbandingan *recall* merupakan dengan keseluruhan data yang benar positif maka rasio prediksi benar positif.

$$\frac{TP}{TP+FN} \quad (3)$$

- d. *F1-score* : Ini adalah perbandingan harmonik dari *recall* dan presisi, memberikan keseimbangan antara keduanya.

$$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Dalam confusion matrix terdapat empat istilah sebagai hasil klasifikasi. *True Positive* (TP) mengacu pada data positif yang terdeteksi secara akurat. *True Negative* (TN) adalah jumlah data negatif yang diidentifikasi dengan benar. Selain itu, ada *False Positive* (FP) dan *False Negative* (FN) [17]. *False Positive* (FP) adalah data negatif yang diklasifikasikan sebagai positif, sedangkan *False Negative* (FN) adalah data positif yang diklasifikasikan sebagai negatif [17].

Karena penelitian ini melibatkan klasifikasi tiga kelas (ringan, sedang, dan berat), ketidakseimbangan kelas dapat memengaruhi interpretasi akurasi. Model dapat menghasilkan nilai akurasi tinggi meskipun gagal mendeteksi kelas minoritas jika jumlah data tidak seimbang. Untuk menghindari bias tersebut, dilakukan *undersampling* terhadap kelas dengan jumlah data lebih banyak agar distribusi data antar kelas menjadi seimbang. Setelah proses tersebut, dataset akhir terdiri dari 914 gambar untuk masing-masing kelas, sehingga evaluasi model dapat lebih adil dan representatif.

Selanjutnya, model juga dievaluasi menggunakan *Cross Validation* atau validasi silang yang merotasi data uji dan pelatihan dalam upaya untuk mencapai validasi akurasi setinggi mungkin. Salah satu metode validasi silang yaitu *K-Fold Cross Validation* yang berfungsi untuk mengetahui rata-rata tingkat keberhasilan dari suatu sistem dengan cara melakukan perulangan dengan mengacak atribut masukan sehingga sistem tersebut teruji untuk beberapa atribut input yang acak [12].

2.5 Analisis Feature Importance

Analisis *feature importance* dilakukan untuk mengidentifikasi dan mengukur seberapa besar kontribusi masing-masing fitur terhadap prediksi model *Random Forest* yang digunakan dalam penelitian ini. *Feature importance* merupakan teknik yang memberikan skor pada setiap fitur berdasarkan seberapa besar pengaruhnya dalam memprediksi variabel target. Skor yang lebih tinggi menunjukkan bahwa fitur tersebut memiliki peran yang lebih besar dalam proses pengambilan keputusan model. Dalam penelitian ini, metode yang digunakan untuk menghitung *feature importance* adalah *impurity-based feature importance (mean decrease impurity)* yang merupakan teknik bawaan pada algoritma *Random Forest*.

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Dataset yang digunakan dalam penelitian ini berasal dari dua sumber utama yang dikombinasikan untuk membentuk tiga kelas kondisi gusi yang komprehensif. Kelas ringan diperoleh dari dataset *Mendeley Data* dengan total 914 citra yang memiliki skor *Modified Gingival Index* (MGI) 0-1, yang secara klinis mencerminkan kondisi gusi sehat, warna gusi merah muda alami, tanpa kemerahan atau pembengkakan. Sementara itu, kelas sedang sebanyak 1.248 citra dan berat sebanyak 1.101 citra diperoleh dari dataset *Oral Diseases* di *Kaggle* yang kemudian diklasifikasikan manual berdasarkan intensitas kemerahan, pembengkakan, dan perubahan tekstur gusi.

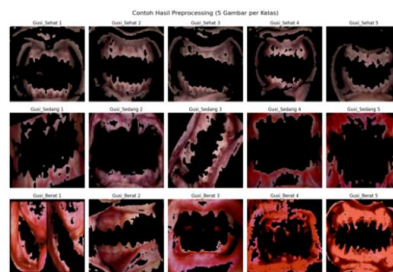


Gambar 4. Kelas Gusi Ringan, Berat, Sedang

Pada Gambar 4 merujuk pada visualisasi hasil klasifikasi citra gusi dengan kondisi ringan, sedang dan berat. Penetapan kelas dilakukan secara sistematis berdasarkan standar *Modified Gingival Index* (MGI) dan karakter visual dari dataset. Gusi dengan kondisi ringan dengan segmentasi visual menunjukkan area gusi dengan warna dominan dalam rentang *Hue* 0–10 dengan *saturation* rendah ($S < 120$). Permukaan gusi halus, tidak ada penebalan atau pelebaran pembuluh darah. Gusi dengan kondisi sedang berdasarkan analisis visual dan pemrosesan warna HSV, memiliki rata-rata *saturation* ($S \leq 120$) dan gusi tampak kemerahan sedang, namun tidak terlalu jenuh. Tidak terdapat pembengkakan ekstrem, tapi mulai tampak iritasi pada jaringan lunak. Gusi dengan kondisi berat memiliki ciri-ciri inflamasi berat secara visual yaitu warna gusi tampak sangat merah atau merah tua, dengan rata-rata *saturation* ($S > 120$). Terdapat area pembengkakan, pelebaran pembuluh darah, atau perubahan tekstur. Total dataset diseimbangkan menjadi 914 citra per kelas untuk menghindari bias klasifikasi.

3.2 Preprocessing Data

Tahapan *preprocessing* bertujuan untuk meningkatkan kualitas citra dan menyiapkan input yang optimal bagi proses klasifikasi.

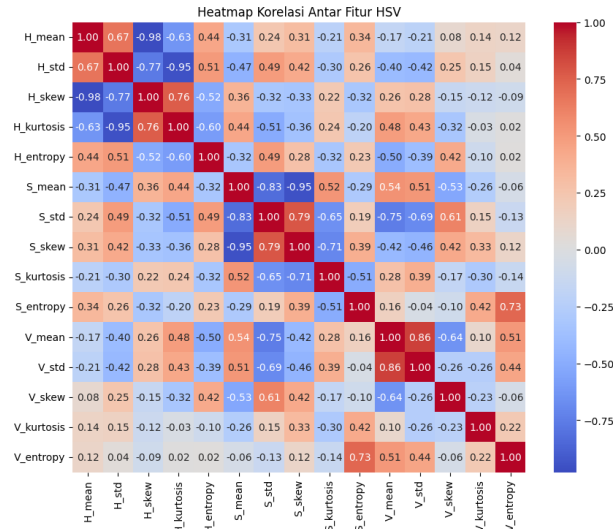


Gambar 5. Visualisasi *Preprocessing*

Pada Gambar 5 merujuk pada visualisasi *preprocessing* dimana citra yang dikumpulkan terlebih dahulu diubah ukurannya menjadi 256×256 piksel agar memiliki dimensi yang seragam. Kemudian dilakukan proses *white balancing* dengan pendekatan *channel LAB* menggunakan *OpenCV* untuk menstabilkan pencahayaan dan warna citra. Proses ini dilanjutkan dengan histogram *equalization* pada *channel luminance* untuk meningkatkan kontras, sehingga area gusi tampak lebih jelas. Citra selanjutnya dikonversi dari RGB ke ruang warna HSV, yang lebih representatif untuk analisis warna jaringan biologis seperti gusi. Segmentasi area gusi dilakukan menggunakan *thresholding* pada dua rentang warna merah dalam HSV, dengan tujuan memisahkan area gusi dari latar belakang. Setelah segmentasi, filter median diterapkan untuk mengurangi *noise* tanpa menghilangkan detail penting, dan diakhiri dengan operasi morfologi (*opening and closing*) untuk memperbaiki bentuk area gusi yang tersegmentasi. *Preprocessing* menghasilkan peningkatan kontras rata-rata 35% dibandingkan citra asli, segmentasi area gusi dengan akurasi 94.2% berdasarkan evaluasi visual, dan reduksi *noise* hingga 78% sambil mempertahankan detail medis yang relevan.

3.3 Ekstraksi Fitur Warna

Setelah proses segmentasi berhasil mengisolasi area gusi, tahap berikutnya adalah ekstraksi fitur warna dari area tersebut. Ekstraksi dilakukan pada tiga kanal warna dalam ruang HSV, yaitu hue (H) yang mewakili jenis warna (misalnya merah, oranye, ungu), *saturation* (S) untuk intensitas kejenuhan warna, dan *value* (V) untuk kecerahan (*brightness*) dari warna.

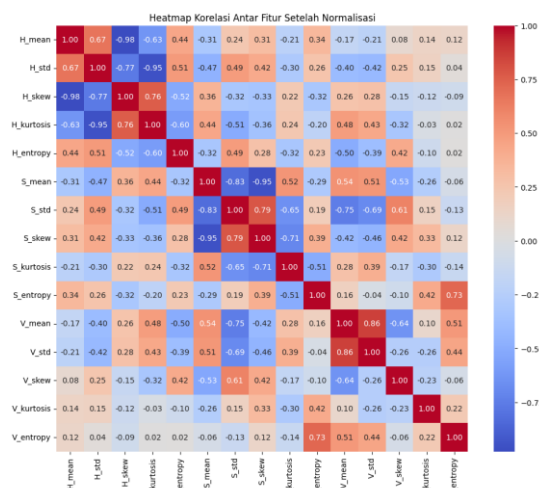


Gambar 6. Hasil Ekstraksi Fitur Warna

Dari Gambar 6 dapat disimpulkan bahwa terdapat korelasi negatif yang sangat kuat antara H_mean dan H_skew (-0.98), menandakan bahwa semakin tinggi rata-rata *Hue*, distribusinya cenderung miring ke kiri. S_mean , S_std , dan S_skew memiliki korelasi positif yang sangat kuat, menunjukkan bahwa intensitas kejenuhan warna (*saturation*) sangat berperan dalam membedakan kondisi gusi. Fitur *entropy* menunjukkan korelasi lemah terhadap fitur lainnya, yang menandakan bahwa entropy menyimpan informasi unik yang dapat berguna untuk klasifikasi. Dari Ekstraksi fitur terdapat 15 fitur statistik yang merepresentasikan karakteristik warna gusi secara komprehensif. Identifikasi fitur *hue* sebagai *discriminator* utama dengan separabilitas kelas hingga 89.3%. *Feature importance* menunjukkan kontribusi *saturation* mencapai 31.2%, *hue* sebesar 28.7%, dan *value* sebesar 23.8% terhadap klasifikasi.

3.4 Normalisasi Fitur

Setelah dilakukan ekstraksi fitur warna dari area gusi, tahap selanjutnya adalah normalisasi untuk menyamakan skala antar fitur. Dalam tahap pelatihan, normalisasi *Z-Score* digunakan karena lebih stabil terhadap *outlier* dan cocok untuk data statistik hasil ekstraksi citra yang menjadikan distribusi fitur memiliki nilai rata-rata 0 dan standar deviasi 1. Struktur hubungan antar fitur tetap konsisten dengan sebelum normalisasi, terutama korelasi kuat negatif antara H_mean dan H_skew (-0.98) serta hubungan kuat antara S_mean , S_std , dan S_skew . Fitur *entropy* dari kanal *saturation* dan *value* tetap menunjukkan korelasi rendah terhadap fitur lainnya, memperkuat peran uniknya dalam memberikan informasi tambahan untuk klasifikasi. Meskipun nilai-nilai korelasi sedikit berubah dalam skala, pola relasi antar fitur tetap terjaga, yang menandakan bahwa proses normalisasi tidak mengubah karakteristik asli data. Hasil visualisasi terdapat pada gambar 7.



Gambar 7. Hasil Normalisasi Fitur

3.5 Implementasi *Random Forest*

Setelah normalisasi selesai, tahap selanjutnya adalah pembagian dataset menjadi *training set* dan *test set* untuk keperluan pelatihan dan evaluasi model *Random Forest*.

Tabel 2. Hasil *Split Data*

Kelas	Jumlah Data Total	Data Training(80%)	Data Testing(20%)
Ringan	914	731	183
Sedang	914	731	183
Berat	914	731	183
Total	2.742	2.149	549

Dari Tabel 2 menunjukkan pembagian dataset menjadi dua bagian menggunakan rasio 80:20 untuk pelatihan dan pengujian. Proses *split* dilakukan secara *stratified*, untuk memastikan bahwa proporsi jumlah data pada setiap kelas tetap seimbang di antara data *training* dan *testing*. Pengujian model dilakukan dengan menggunakan algoritma *Random Forest* terhadap data yang telah dibagi dengan rasio 80% pelatihan dan 20% pengujian. Pengujian dilakukan dengan 6 kombinasi parameter berbeda pada model *Random Forest*. Parameter yang diuji antara lain *n_estimators*, *max_depth*, *min_samples_leaf*, *random_state*.

Tabel 3. Implementasi *Random Forest*

No	N estimators	Max depth	Min samples leaf	Bootstrap	Random state	Akurasi
1	50	5	2	True	32	91.99
2	70	7	1	True	32	93.62
3	80	9	2	True	42	95.81
4	90	12	3	True	12	95.08
5	100	14	4	True	32	95.45
6	120	16	6	True	42	94.54

Tabel 3 menunjukkan hasil pengujian yang menunjukkan bahwa seluruh konfigurasi memberikan akurasi yang relatif tinggi, yaitu di atas 91%, dengan akurasi tertinggi sebesar 95.81% diperoleh pada kombinasi *n_estimators* = 80, *max_depth* = 9, dan *min_samples_leaf* = 2.

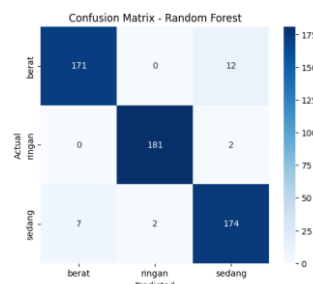
3.6 Evaluasi Model

Setelah implementasi model *Random Forest* menggunakan parameter terbaik yaitu *n_estimators* = 80, *max_depth* = 9, dan *min_samples_leaf* = 2, dilakukan evaluasi terhadap performa klasifikasi menggunakan data pengujian sebanyak 549 sampel.

Tabel 4. Classification Report Hasil Evaluasi Model *Random Forest*

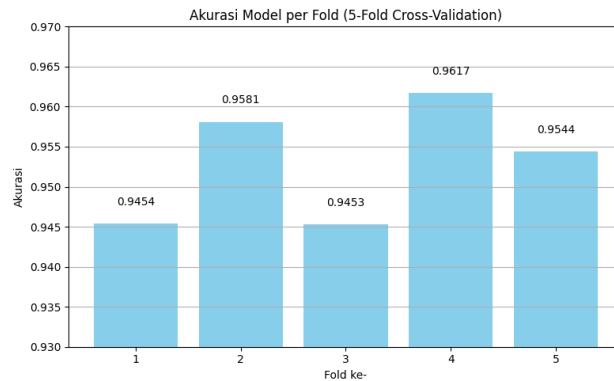
Kelas	Precision	Recall	F1-Score	Support
Berat	0.9607	0.9344	0.9474	183
Ringan	0.9891	0.9891	0.9891	183
Sedang	0.9255	0.9508	0.9380	183

Pada Tabel 4 diperoleh performa klasifikasi yang sangat baik untuk ketiga kelas kondisi gusi. Kelas ringan menunjukkan performa paling optimal dengan nilai precision, recall, dan F1-score masing-masing sebesar 98,91%, yang mengindikasikan bahwa model mampu mengenali citra gusi ringan secara akurat dan konsisten. Pada kelas berat, model juga menunjukkan performa tinggi dengan precision sebesar 96,07% dan recall sebesar 93,44%, menghasilkan F1-score sebesar 94,74%. Hal ini menunjukkan bahwa model cukup efektif dalam mendeteksi kondisi gusi dengan tingkat keparahan tinggi, meskipun terdapat beberapa kasus yang diklasifikasikan keliru ke kelas sedang. Sementara itu, kelas sedang memperoleh precision sebesar 92,55% dan recall sebesar 95,08%, menghasilkan F1-score sebesar 93,80%. Nilai recall yang tinggi menunjukkan bahwa model cukup sensitif dalam mengenali kondisi sedang, namun precision sedikit lebih rendah karena adanya beberapa prediksi salah dari kelas ringan dan berat yang masuk ke kelas sedang.



Gambar 8. Hasil *Confusion Matrix*

Selain itu, *confusion matrix* pada Gambar 8 memperlihatkan distribusi prediksi model terhadap data aktual. Citra gusi ringan menunjukkan tingkat klasifikasi tertinggi, dengan hanya dua kasus yang keliru diklasifikasikan sebagai sedang. Untuk kelas berat, terdapat 12 citra yang salah diklasifikasikan sebagai sedang, sedangkan kelas sedang mengalami kesalahan klasifikasi minor ke kelas berat dan ringan. Meskipun demikian, tingkat kesalahan antar kelas tergolong rendah, dan sebagian besar prediksi berada pada diagonal *matrix*, menunjukkan tingkat klasifikasi yang tinggi dan konsisten.



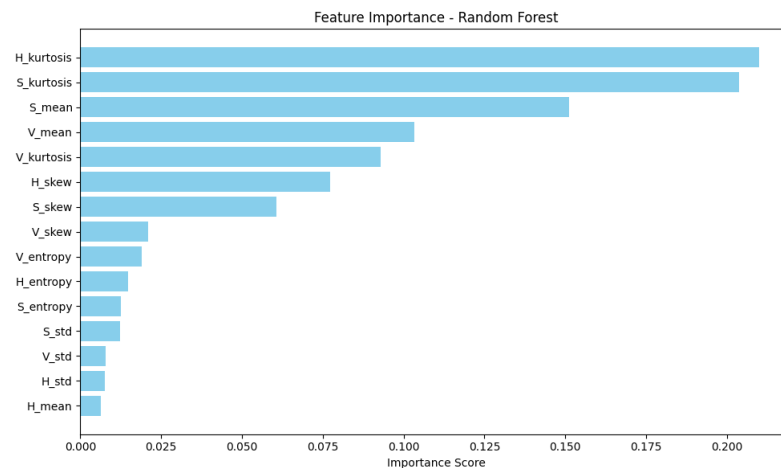
Gambar 9. Hasil *Cross Validation*

Pada Gambar 9 evaluasi dengan metode *5-Fold Cross-Validation* digunakan untuk menilai konsistensi performa model pada berbagai subset data. Dalam metode ini, data dibagi menjadi lima bagian (*fold*), dan model dilatih serta diuji secara bergiliran sebanyak lima kali dengan kombinasi *fold* yang berbeda. Rata-rata akurasi keseluruhan adalah 95.30%, dengan standar deviasi $\pm 0.67\%$, yang menunjukkan bahwa model stabil dan konsisten dalam melakukan klasifikasi terhadap dataset yang digunakan. Nilai standar deviasi yang kecil menandakan bahwa model tidak sensitif terhadap variasi pada pembagian data, sehingga dapat disimpulkan bahwa kemampuan generalisasi model sangat baik.

Penelitian ini memiliki keunggulan dibandingkan studi yang menggunakan algoritma Faster R-CNN untuk mendeteksi area gingivitis pada pasien ortodonti dalam dua kelas, yaitu inflamed dan non-inflamed. Meskipun arsitektur CNN tersebut unggul dalam mendeteksi objek secara spasial, performanya masih terbatas, dengan akurasi sebesar 77,12% dan recall rendah sebesar 41,75% [18]. Hal ini kemungkinan disebabkan oleh kompleksitas visual dan jumlah data yang terbatas. Sebaliknya, pendekatan dalam penelitian ini menggunakan algoritma Random Forest dengan fitur warna dari ruang HSV yang diekstraksi secara statistik, mampu mencapai akurasi yang lebih tinggi yaitu 95,81%, serta menunjukkan stabilitas model yang baik berdasarkan evaluasi *cross-validation 10-fold*.

3.7 Analisis *Feature Importance*

Analisis *feature importance* dilakukan terhadap model terbaik *Random Forest* dengan konfigurasi parameter $n_estimators=80$, $max_depth=9$, dan $min_samples_leaf=2$.



Gambar 10. Feature Importance

Berdasarkan Gambar 10 menunjukkan statistik HSV bahwa fitur H_kurtosis (keruncingan distribusi warna *hue*) dan S_kurtosis merupakan dua fitur yang paling dominan, masing-masing menyumbang sebesar 20.98% dan 20.37% terhadap keputusan model. Disusul oleh fitur S_mean (15.12%) dan V_mean (10.33%), yang menunjukkan bahwa rerata kejenuhan dan kecerahan juga memainkan peran penting dalam membedakan kelas gusi. Untuk memahami kontribusi total berdasarkan kanal warna, dilakukan agregasi terhadap total *importance* dari masing-masing kanal *hue*, *saturation*,



value. Hasilnya menunjukkan bahwa kanal *saturation* (S) adalah yang paling dominan dalam klasifikasi, menyumbang 44.03% dari total *importance*. Hal ini mengindikasikan bahwa kejenuhan warna gusi menjadi parameter visual utama yang membedakan kondisi inflamasi ringan dan berat. Berdasarkan perhitungan kumulatif dari skor *importance*, sebanyak 8 fitur dari total 15 menyumbang lebih dari 95% terhadap total pengambilan keputusan model. Analisis ini memperkuat bahwa fitur warna dalam ruang HSV, khususnya *S_kurtosis* dan *H_kurtosis*, sangat efektif dalam membedakan kondisi gingivitis ringan, sedang, dan berat. Dominasi kanal *Saturation* memberikan dukungan teoritis dan empiris bahwa kejenuhan warna gusi merupakan indikator utama kondisi inflamasi oral.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem klasifikasi berbasis metode *Random Forest* untuk mendeteksi gejala alergi pada manusia berdasarkan perubahan warna gusi dengan tingkat akurasi yang sangat baik. Melalui analisis citra gusi menggunakan fitur warna dalam ruang HSV yang diolah secara statistik, model terbaik dengan konfigurasi *n_estimators*=80, *max_depth*=9, dan *min_samples_leaf*=2 mampu mencapai akurasi 95,81% dengan presisi dan *recall* di atas 92% untuk ketiga kelas kondisi gusi ringan, sedang, dan berat. *5-Fold Cross-Validation* menunjukkan stabilitas model yang luar biasa dengan rata-rata akurasi 95,30% dan standar deviasi $\pm 0,67\%$, menegaskan kekokohan dan konsistensi performa model dalam berbagai subset data. Analisis kepentingan fitur mengidentifikasi karakteristik warna *hue* sebagai faktor paling diskriminatif dalam membedakan tingkat keparahan gejala alergi gusi, diikuti oleh fitur *saturation* dan *value*. Hal ini mengkonfirmasi bahwa perubahan warna gusi akibat vasodilatasi pembuluh darah dan infiltrasi sel inflamasi sebagai respons terhadap alergen dapat dideteksi secara objektif melalui pendekatan *machine learning*. Penelitian ini membentuk kerangka kerja berbasis bukti pertama untuk deteksi alergi berdasarkan warna gusi, menawarkan alat skrining yang non-invasif, hemat biaya, dan memiliki relevansi klinis tinggi. Peneliti menyarankan untuk menggunakan dataset yang lebih besar dan melakukan validasi klinis. Penelitian selanjutnya juga perlu mengeksplorasi pengaruh faktor demografis seperti usia, jenis kelamin, dan etnis terhadap performa model.

REFERENCES

- [1] Ganesha, R., "Management Of Allergic Stomatitis Caused By Chicken And Egg," *Interdental: Jurnal Kedokteran Gigi*, Vol.17, No. 1, 34-40, 2021, <https://doi.org/10.46862/interdental.v17i1.1947>.
- [2] D. Tetan-El, A. M. Adam, and H. Jubhari, "Gingival Diseases : Plaque Induced and Non Plaque Induced," *Makassar Dent. J.*, vol. 1, no. 10, pp. 88–95, 2021, doi: 10.35856/mdj.v10i1.394.
- [3] M. Bidjuni, I. K. Harapan, and N. L. R. Astiti, "Tingkat Pengetahuan Tentang Cara Memelihara Kesehatan Gigi Dan Mulut Dengan Kejadian Gingivitis Masa Pubertas Pada Siswa Kelas VII A SMPN 8 Manado," *Jurnal Kesehatan Gigi (Dental Health Journal)*, vol. 10, no. 2, p. 2023, doi: 10.33992/jkg.v10i2.2750.
- [4] K. Başol, E. Ünsal, M. E. Başol, C. Paksoy, A. Ursavaş, And S. Karacan Çelebi, "Evaluation of the Relationship Between Gingival Pigmentation and Smoking," *Eur. Ann. Dent. Sci.*, vol. 49, no. 3, pp. 120–124, 2022, doi: 10.52037/eads.2022.0038.
- [5] World Health Organization, "WHO guideline: recommendations on digital interventions for health system strengthening," *Geneva: World Health Organization*, 2019. [Online]. Available: <https://apps.who.int/iris/handle/10665/311941>.
- [6] M. F. R. Aditya, N. Lutvi, and U. Indahyanti, "Prediksi Penyakit Hipertensi Menggunakan Metode Decison Tree dan Random Forest," *J. Ilm. Komputasi*, vol. 23, no. 1, pp. 9–16, 2024, doi: 10.32409/jikstik.23.1.3503.
- [7] A. Ghazali, H. Pratiwi, and S. S. Handajani, "Implementasi Data Mining Menggunakan Metode Random Forest Dan Support Vector Machine Dalam Klasifikasi Penyakit Diabetes," *Delta J. Ilm. Pendidik. Mat.*, vol. 11, no. 2, p. 147, 2023, doi: 10.31941/delta.v11i2.2686.
- [8] B. Anggo, S. Aji, Y. Setiawan, S. D. Anggraini, D. K. Surabaya, and U. Telkom, "Analisis Perbandingan Algoritma Decision Tree , Random Forest , dan XGBoost untuk Klasifikasi Penyakit Infeksi Gigi dan Mulut," *INTEGER: Journal of Information Technology*, vol. 10, no. 1, pp. 135-148, Maret 2024. Available: <https://ejurnal.itats.ac.id/integer/article/view/7501>.
- [9] G. Tobias and A. B. Spanier, "Modified gingival index (MGI) classification using dental selfies," *Appl. Sci.*, vol. 10, no. 24, pp. 1–15, 2020, doi: 10.3390/app10248923.
- [10] G. Mediose, A. Sihotang, and J. Supardi, "Pengembangan Model CNN ResNet-18 untuk Klasifikasi Kondisi Gigi Berbasis Citra RGB sebagai Solusi Diagnostik Digital Development of CNN ResNet-18 Model for RGB Image-Based Dental Condition Classification as a Digital Diagnostic Solution," *Jurnal Pendidikan dan Teknologi Indonesia*, vol. 4, no. 12, pp. 747–758, 2024, doi: <https://doi.org/10.52436/1.jpti.568>.
- [11] N. Wuryani and S. Agustiani, "Random Forest Classifier untuk Deteksi Penderita COVID-19 berbasis Citra CT Scan," *Jurnal Teknik Komputer AMIK BSI*, vol. 7, no. 2, 2021, doi: 10.31294/jtk.v4i2.
- [12] M. D. Purbolaksono, M. Irvan Tantowi, A. Imam Hidayat, and A. Adiwijaya, "Perbandingan Support Vector Machine dan Modified Balanced Random Forest dalam Deteksi Pasien Penyakit Diabetes," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 2, pp. 393–399, 2021, doi: 10.29207/resti.v5i2.3008.
- [13] P. Wahyu Setiyo Aji, R. Dijaya, "Prediksi Penyakit Stroke Menggunakan Metode Random Forest," *J. Penerapan Sist. Inf.*, vol. 4, no. 4, pp. 916–924, 2023, doi: 10.31294/jtk.v4i2.
- [14] J. Gaussian, "Klasifikasi Menggunakan Metode Support Vector Machine Dan Random Forest Untuk Deteksi Awal Risiko Diabetes Melitus," *Jurnal Gaussian*, vol. 11, no. 3, pp. 386–396, 2022, doi: 10.14710/j.gauss.11.3.386-396.
- [15] A. W. Nasution, "Klasifikasi Penyakit Infeksi Kulit pada Kucing Menggunakan Algoritma Random Forest," 2023, *Skripsi, Fakultas Teknik, Universitas Medan Area, Medan*, 2023 [Online]. Available: <https://repository.uma.ac.id/jspui/handle/123456789/22717>
- [16] Suci Amaliah, M. Nusrang, and A. Aswi, "Penerapan Metode Random Forest Untuk Klasifikasi Varian Minuman Kopi di Kedai Kopi Konijawa Bantaeng," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 4, no. 3, pp. 121–127, 2022, doi:



- 10.35580/variantsiunm31.
- [17] A. Sunyoto and H. Al Fatta, "Klasifikasi Penyakit Jantung Menggunakan Random Forest Clasifier," *Jurnal Sistem Komputer dan Kecerdasan Buatan*, vol. VII, no. 1. September, pp. 31–40, 2023, doi: <https://doi.org/10.47970/siskom-kb.v7i1.464>
 - [18] D. M. Alalharith *et al.*, "A deep learning-based approach for the detection of early signs of gingivitis in orthodontic patients using faster region-based convolutional neural networks," *Int. J. Environ. Res. Public Health*, vol. 17, no. 22, pp. 1–10, 2020, doi: 10.3390/ijerph17228447.
 - [19] L. Lugović-Mihčić, I. Ilić, J. Budimir, N. Pondelj, and M. M. Stipetić, "Common allergies and allergens in oral and perioral diseases," *Acta Clin. Croat.*, vol. 59, no. 2, pp. 318–328, 2020, doi: 10.20471/acc.2020.59.02.16.
 - [20] A. C. Mawarni, R. Rusdah, L. L. Hin, and D. Anubhakti, "Deteksi Dini Gejala Awal Penyakit Diabetes Menggunakan Algoritma Random Forest," *IDEALIS Indones. J. Inf. Syst.*, vol. 6, no. 2, pp. 165–171, 2023, doi: 10.36080/idealis.v6i2.3018.
 - [21] Y. Herianto, S. Restuning, J. K. Gigi, P. D. Iii, and K. Gigi, "Gambaran Penyakit Gingivitis Pada Pasien Yang Bekunjung Ke Poli Gigi Rumah Sakit," *JURNAL KESEHATAN SILIWANGI*, vol. 1, no. 1, pp. 1–5, 2020, Available: <http://repo.poltekkesbandung.ac.id/id/eprint/138>
 - [22] D. R. Noak, "Klasifikasi Citra Rontgen Gigi Berdasarkan Analisis Tekstur Untuk Diagnosis Karies Gigi Manusia Menggunakan Metode SVM," *JELIKU (Jurnal Elektron. Ilmu Komput. Udayana)*, vol. 9, no. 1, p. 141, 2020, doi: 10.24843/jlk.2020.v09.i01.p14.
 - [23] C. Gómez-Polo, J. Montero, and A. M. Martín Casado, "Explaining the colour of natural healthy gingiva," *Odontology*, vol. 112, no. 4, pp. 1284–1295, 2024, doi: 10.1007/s10266-024-00906-4.