



Klasifikasi Kanker Kulit dari Citra Dermoskopi Menggunakan Fitur *Gray Level Co-occurrence Matrix* (GLCM) dengan Algoritma *Machine Learning*

Chrisman Imanuel Purba¹, Alrizal¹, dan Yoza Fendriani^{1*}

¹*Program Studi Fisika, Fakultas Sains dan Teknologi, Universitas Jambi*

Email: yozafendriani@unja.ac.id

**Corresponding Author*

Abstrak

Penelitian ini bertujuan untuk mengklasifikasikan kanker kulit berdasarkan citra dermoskopi menggunakan teknik ekstraksi fitur tekstur *Gray Level Co-occurrence Matrix* (GLCM) dengan membandingkan kinerja empat algoritma *machine learning* yakni: *Support Vector Machine* (SVM), *K-Nearest Neighbor* (K-NN), *Decision Tree*, dan *Random Forest*. Pendekatan ini dikembangkan untuk menjawab keterbatasan penelitian sebelumnya yang umumnya hanya menggunakan satu algoritma tanpa perbandingan menyeluruh. Hasil evaluasi menunjukkan bahwa *Random Forest* memberikan performa terbaik dari algoritma model lainnya dengan akurasi sebesar 92,72%, presisi 94,44%, *recall* 92,39%, dan *F1-score* 93,40%. Hal ini karena kemampuannya sebagai model *ensemble* yang menggabungkan banyak pohon keputusan melalui *voting*, sehingga efektif menangani data tidak seimbang dan tekstur citra yang kompleks. Sebaliknya, *Support Vector Machine* (SVM) menghasilkan performa terendah dengan akurasi sebesar 66,06%, presisi 84,44%, *recall* 64,40%, dan *F1-score* 73,07%, yang mengindikasikan keterbatasannya dalam mengenali pola nonlinier pada data berdimensi tinggi. Berdasarkan hasil tersebut, kombinasi GLCM dan *Random Forest* terbukti efektif dan optimal dalam klasifikasi citra medis serta berpotensi mendukung pengambilan keputusan klinis secara lebih akurat dalam deteksi dini kanker kulit.

Kata kunci: Dermoskopi, Ekstraksi Fitur, GLCM, Kanker Kulit, *Machine Learning*.

Abstract

This study aims to classify skin cancer based on dermoscopic images using texture feature extraction through the Gray Level Co-occurrence Matrix (GLCM) technique by comparing the performance of four machine learning algorithms: Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), Decision Tree, and Random Forest. This approach was developed to address the limitations of previous studies, which typically employed only a single algorithm without comprehensive comparison. The evaluation results show that Random Forest achieved the best performance, with an accuracy of 92.72%, precision of 94.44%, recall of 92.39%, and an F1-score of 93.40%. This is attributed to its ensemble nature, which combines multiple decision trees through a voting mechanism, making it effective in handling imbalanced data and complex texture patterns. Conversely, Support Vector Machine (SVM) demonstrated the lowest performance, with an accuracy of 66.06%, precision of 84.44%, recall of 64.40%, and an F1-score of 73.07%, indicating its limitations in recognizing nonlinear in high-dimensional data. Based on these results, the combination of GLCM and Random Forest has proven to be effective and optimal for medical image classification, and holds significant potential to support more accurate clinical decision-making in the early detection of skin cancer.

Keywords: *Dermoscopic, Feature Extraction, GLCM, Machine Learning, Skin Cancer.*

1. PENDAHULUAN

Seluruh bagian luar tubuh manusia dilindungi oleh sebuah sistem yang dikenal sebagai sistem integumen. Sistem integumen merupakan sistem organ yang mampu membedakan, memisahkan, dan memberikan respons terhadap rangsangan dari lingkungan sekitar [1]. Kulit sebagai komponen utama sistem integumen sekaligus organ terbesar pada tubuh manusia yang menyumbang sekitar 15% dari total berat tubuh dengan memiliki luas $\pm 1,5 \text{ m}^2$ pada orang dewasa. Kulit memiliki fungsi utama untuk mencegah kehilangan cairan berlebih dan melindungi tubuh terhadap paparan zat kimia, bakteri, serta radiasi ultraviolet (UV) dari lingkungan [2]. Saat ini, kesehatan kulit merupakan aspek penting dalam kesejahteraan manusia karena kulit sangat sensitif terhadap berbagai kondisi medis. Salah satu ancaman serius terhadap kesehatan kulit adalah kanker kulit, yang dapat berakibat fatal jika tidak segera dideteksi dan ditangani [3].

Kanker kulit salah satu kondisi ketika sel-sel kulit berubah menjadi ganas, berkembang secara tidak terkendali, merusak materi genetik (DNA), dan berpotensi menyebar ke jaringan maupun organ lain melalui proses metastasis. Berdasarkan data Globocan (*Global Cancer Observatory*) tahun 2023, terdapat sekitar 18.000 kasus kanker kulit di Indonesia dengan angka kematian mencapai 3.000 kasus. Kanker kulit tergolong sebagai salah satu jenis kanker yang cukup umum di Indonesia, menempati peringkat ketiga setelah kanker rahim dan kanker payudara, dengan prevalensi tahunan mencapai 5,9–7,8% dari seluruh kasus kanker [4].

Sejalan dengan tantangan dalam diagnosis kanker kulit, jika pasien diduga menderita penyakit ini, dokter akan melakukan biopsi dengan mengambil sampel jaringan kulit untuk diperiksa di laboratorium. Meskipun biopsi dapat memberikan informasi histopatologis, prosedur ini memiliki kekurangan, diantaranya biaya yang cukup mahal, waktu tunggu hasil hingga satu minggu, dan waktu penyembuhan luka yang lama. Selain itu, mendeteksi kanker kulit dari citra lesi melalui dermoskopi bukanlah tugas yang mudah bagi dokter kulit, karena memerlukan keahlian khusus dan ketidakcermatan dalam proses tersebut menyebabkan kesalahan diagnosis. Mengingat tantangan dalam mendiagnosis kanker kulit melalui dermoskopi cukup berat, dibutuhkan sistem deteksi dini yang lebih akurat [5].

Perkembangan teknologi *machine learning* menyediakan solusi tepat. Dalam konteks ini, *machine learning* memiliki kemampuan untuk mengungkap pola-pola tersembunyi dalam data medis dan ilmiah, yang berpotensi meningkatkan akurasi diagnosis. Penggunaan algoritma *machine learning* yang terstruktur dan sistematis diharapkan dapat meningkatkan efektivitas teknologi ini dalam dunia medis, termasuk dalam deteksi kanker kulit [6]. Pengolahan citra digital dalam klasifikasi sangat bergantung pada tekstur dan warna. Salah satu metode yang umum digunakan dalam analisis tekstur citra adalah *Gray Level Co-occurrence Matrix* (GLCM) yang berfungsi untuk menghitung frekuensi kemunculan pasangan piksel dengan nilai tertentu dan hubungan spasial spesifik [7].

Beberapa penelitian sebelumnya telah menggunakan klasifikasi dengan ekstraksi fitur GLCM dalam algoritma *machine learning*. Penelitian oleh Wisudawati (2021), mengklasifikasi kanker payudara pada citra mammogram dengan algoritma *Support Vector Machine* (SVM), diperoleh nilai evaluasi performa akurasi 83,59%, presisi 87,58%, dan

recall 76,84% [8]. Sementara itu, penelitian Suharyana et al. (2023), mengidentifikasi pneumonia pada dua & tiga kelas dalam algoritma *K-Nearest Neighbor*, didapatkan evaluasi pada nilai akurasi dua kelas 90% dan tiga kelas 83,3% [9]. Selanjutnya, penelitian oleh Yunianto et al. (2024), mengembangkan model *Decision Tree* pada citra CT-Scan paru-paru, diperoleh performa yang sangat baik dengan tingkat akurasi 92,5% dan *recall* 95% [10]. Terakhir, penelitian oleh Wijaya (2023), mengklasifikasikan citra monkeypox dan non-monkeypox dalam algoritma *Random Forest*, dimana nilai akurasi, presisi, *recall* dan *F1-score* berada dalam rentang 76,5 - 77% [11].

Meskipun banyak penelitian telah membuktikan kemampuan analisis tekstur menggunakan GLCM yang dikombinasikan dengan beragam algoritma *machine learning*, terdapat beberapa keterbatasan yang menjadi celah penelitian (*research gap*) dan perlu dieksplorasi lebih lanjut. Pertama, umumnya penelitian terdahulu hanya menggunakan satu jenis algoritma klasifikasi tanpa membandingkan performanya secara langsung dengan algoritma lain, sehingga belum tersedia bukti empiris yang cukup mengenai algoritma mana yang paling optimal untuk klasifikasi citra berbasis tekstur. Kedua, penelitian yang secara relevan mengkombinasikan GLCM dan pendekatan komparatif beberapa algoritma *machine learning* dalam klasifikasi citra dermoskopi kanker kulit masih sangat terbatas. Padahal, klasifikasi kanker kulit memerlukan pendekatan yang presisi dan efisien untuk mendukung diagnosis dini. Ketiga, pengaruh tahapan *pre-processing* citra terhadap performa klasifikasi juga belum banyak diteliti secara mendalam. Berdasarkan uraian di atas, penelitian ini untuk mengevaluasi dan membandingkan performa metode *Support Vector Machine*, *K-Nearest Neighbors*, *Decision Tree*, dan *Random Forest* dalam klasifikasi kanker kulit menggunakan ekstraksi fitur GLCM. Dengan pendekatan ini, diharapkan dapat diidentifikasi algoritma klasifikasi yang optimal dalam mendukung diagnosis kanker kulit dengan akurat dan efisien.

2. METODE PENELITIAN

2.1 Alat dan Bahan

Penelitian ini memerlukan bahan sebagai objek penelitian. Bahan yang diolah dalam penelitian ini adalah dataset dermoskopi. Data citra dermoskopi diperoleh dari sumber terbuka (*open public access*) pada situs Kaggle dalam tautan "www.kaggle.com" [12], dengan judul "*Skin Cancer: Malignant and Benign*" yang disusun oleh Claudio Fanconi. Dataset ini berasal dari persatuan *The International Skin Imaging Collaboration* (ISIC), lembaga yang menyediakan koleksi citra kulit dan dapat diakses oleh pengguna melalui URL berikut ini: "<https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign>". Pemilihan dataset ini didasarkan pada beberapa kriteria, antara lain kualitas citra yang tinggi dan jelas, ukuran dataset yang memadai dengan jumlah data yang seimbang antara kelas *malignant* dan *benign*, serta keberagaman jenis lesi kulit yang memberikan representasi tekstur dan pola visual yang bervariasi. Selain itu, dataset ini juga dilengkapi dengan label klasifikasi yang valid dan dokumentasi yang baik, sehingga mendukung proses pelatihan dan evaluasi model klasifikasi secara optimal. Pada penelitian ini, citra dermoskopi yang digunakan untuk klasifikasi terdiri dari 3.297 citra, terbagi menjadi dua bagian: 2.637 citra untuk data latih dan 660 citra untuk data uji.

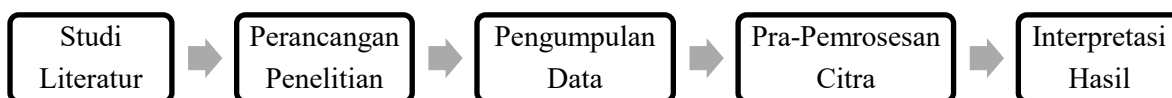
Peralatan yang digunakan dalam penelitian ini mencakup perangkat keras (*hardware*) berupa PC, dan perangkat lunak (*software*) yaitu sistem operasi *Windows*, dan *Jupyter Notebook* sebagai *editor* dalam bahasa pemrograman Python. Selain itu, penelitian ini juga menggunakan berbagai *library* utama, antara lain: *Pandas*, *Numpy*, *OpenCV2*, *Matplotlib*, dan *Scikit-Learn* untuk menunjang proses pengolahan dan analisis data.

2.2 Prosedur Penelitian

Penelitian ini dilakukan secara sistematis untuk mengevaluasi kinerja empat algoritma *machine learning* dalam mengklasifikasikan jenis kanker kulit berdasarkan citra dermoskopi, sebagaimana diadaptasi dari penelitian oleh Ananda et al. (2024) [13]. Tahap awal dimulai dengan studi literatur untuk menelusuri dan mempelajari sumber-sumber ilmiah yang relevan mengenai klasifikasi kanker kulit dengan menggunakan ekstraksi fitur *Gray Level Co-occurrence Matrix* (GLCM) dan berbagai metode seperti *Support Vector Machine* (SVM), *K-Nearest Neighbor* (K-NN), *Decision Tree*, dan *Random Forest*. Selanjutnya, dilakukan perancangan metodologi yang mencakup pemilihan algoritma, teknik evaluasi, serta penentuan sumber dan kriteria pemilihan data. Data yang digunakan merupakan citra dermoskopi yang terdiri dari dua kelas, yaitu *benign* (jinak) dan *malignant* (ganas).

Tahapan pra-pemrosesan citra dilakukan untuk memastikan kualitas data optimal sebelum proses ekstraksi fitur. Proses ini mencakup normalisasi citra guna menyamakan skala intensitas piksel, serta augmentasi data dilakukan untuk meningkatkan generalisasi model dengan cara mengenalkan variasi modifikasi citra, yakni *random crop* dan *zoom in*, yang disesuaikan hingga menghasilkan ukuran akhir 200×200 piksel dengan tetap memperhatikan detail pada area lesi kulit. Selain itu, dilakukan penanganan *data imbalance* menggunakan teknik *undersampling* mengatasi distribusi data antar kelas lebih seimbang.

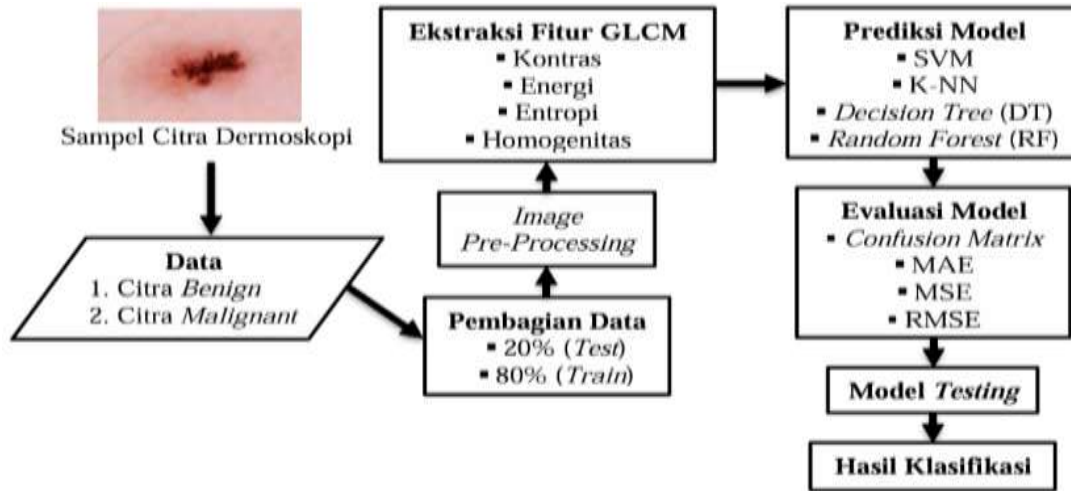
Terakhir, interpretasi hasil klasifikasi dilakukan untuk mengidentifikasi algoritma terbaik dan faktor-faktor utama yang memengaruhi klasifikasi kanker kulit dalam citra dermoskopi, sehingga dapat memberikan diagnosis dini yang akurat bagi pasien.



Gambar 1. Diagram Alir Prosedur Penelitian

2.3 Tahap Pengolahan dan Teknik Analisis Data

Penelitian dimulai pada tahap pengolahan data berdasarkan pembagian data ke dalam set pelatihan dan set pengujian. Penelitian sebelumnya oleh Fata & Donny (2024) [14] menunjukkan bahwa pembagian data dengan proporsi 80% untuk pelatihan dan 20% untuk pengujian memberikan hasil yang optimal. Peneliti membagi dataset menjadi dua bagian, yaitu data latih sebanyak 80% dan data uji sebanyak 20%. Tujuan dari pembagian ini adalah untuk menilai seberapa baik model yang telah dilatih dapat bekerja pada data baru yang belum pernah dilihat sebelumnya. Dengan cara ini, model dapat dievaluasi kemampuannya dalam mengenali pola dan menerapkannya pada situasi yang berbeda.



Gambar 2. Diagram Alir Tahap Pengolahan dan Teknik Analisis Data

Penelitian ini menggunakan teknik ekstraksi fitur *Gray Level Co-occurrence Matrix* (GLCM) untuk menganalisis tekstur citra berdasarkan distribusi intensitas piksel pada jarak (d) dan sudut (θ) tertentu. GLCM dipilih karena mampu merepresentasikan karakteristik tekstur secara statistik dalam membedakan lesi kulit *benign* dan *malignant*. Dibandingkan metode modern seperti *Convolutional Neural Networks* (CNN), GLCM lebih ringan secara komputasi, tidak memerlukan data latih dalam jumlah besar, serta hasilnya lebih mudah diinterpretasikan. Sejalan dengan hal ini, GLCM tetap efektif dan efisien untuk penelitian dengan keterbatasan data dan sumber daya, serta memberikan hasil yang kompetitif [15]. Beberapa fitur GLCM yang digunakan dalam penelitian ini meliputi kontras, energi, entropi, dan homogenitas dengan masing-masing dirumuskan dalam persamaan sebagai berikut [16]:

$$\text{Kontras} = \sum_{i=1}^N \sum_{j=1}^N (i - j)^2 p(i, j) \quad (1)$$

$$\text{Energi} = \sum_{i=1}^N \sum_{j=1}^N p(i, j)^2 \quad (2)$$

$$\text{Entropi} = - \sum_{i=1}^N \sum_{j=1}^N p(i, j) \log_2 (p(i, j)) \quad (3)$$

$$\text{Homogenitas} = \sum_{i=1}^N \sum_{j=1}^N \frac{p(i, j)}{1 + |i - j|} \quad (4)$$

Keterangan

i dan j : indeks baris dan kolom pada matriks GLCM

N : jumlah tingkat keabuan (*gray levels*) dalam citra

$p(i, j)$: probabilitas kemunculan bersama pasangan piksel dengan intensitas i dan j

Setelah dataset siap, model dilatih menggunakan empat algoritma klasik, yaitu *Support Vector Machine* (SVM), *K-Nearest Neighbor* (K-NN), *Decision Tree*, dan *Random Forest*. Meskipun saat ini telah berkembang berbagai metode klasifikasi terbaru, seperti *deep learning* dan *advanced ensemble*, pemilihan keempat algoritma ini didasarkan pada efektivitasnya dalam studi klasifikasi berbasis fitur statistik serta efisiensinya dalam konteks

dataset dengan ukuran dan kompleksitas sedang. Pada algoritma SVM, model dilatih untuk membagi kelas dengan menggunakan model atau pola menggunakan garis batas keputusan optimal sehingga menghasilkan jarak terbaik antara kedua kelas. Selanjutnya, algoritma K-NN mengelompokkan data baru berdasarkan mayoritas kelas dari 'K' tetangga terdekat dalam ruang fitur. Algoritma *Decision Tree* membangun model klasifikasi dengan membuat struktur pohon keputusan yang memilih atribut terbaik untuk memisahkan data pada setiap simpul. Terakhir, algoritma *Random Forest* membangun beberapa pohon keputusan secara acak dan menggabungkan prediksi dari setiap pohon melalui *voting* atau *averaging* untuk meningkatkan nilai akurasi model dan mengurangi risiko *overfitting* [17].

2.4 Evaluasi Model

Kinerja model algoritma *machine learning* dievaluasi lebih lanjut menggunakan metrik *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE), dan *Root Mean Squared Error* (RMSE). Ketiga metrik ini membantu menilai sejauh mana model mampu memprediksi nilai sebenarnya dengan mengukur rata-rata kesalahan prediksi yang dihasilkan. Nilai MAE, MSE, dan RMSE yang semakin mendekati 0 mengindikasikan performa model semakin baik dalam memprediksi data.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (5)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

dimana n menyatakan jumlah data, y_i sebagai nilai aktual data ke-I, and $\hat{y}_i$ menyatakan nilai prediksi dari model untuk data ke-i.

Untuk tugas klasifikasi, akurasi digunakan sebagai metrik evaluasi. Akurasi dihitung sebagai rasio antara jumlah prediksi yang benar dibagi dengan jumlah total data. Rumus untuk akurasi adalah:

$$Akurasi = \frac{\text{Jumlah prediksi benar}}{\text{Jumlah total data}} \times 100\% \quad (8)$$

Tabel 1. Tingkat Akurasi [18]

Kategori	Nilai Akurasi
Sangat Bagus	90% - 100 %
Bagus	80% - 90%
Cukup	70% - 80%
Buruk	60% - 70%
Sangat Buruk	≤60%

Confusion Matrix salah satu teknik evaluasi kinerja model yang menyajikan visualisasi hasil prediksi dibandingkan dengan kelas sebenarnya. Teknik ini menjelaskan seberapa baik model melakukan prediksi dengan membandingkan hasil prediksi dengan kelas asli, yang mencakup informasi aktual dari data tersebut. Kinerja *Confusion Matrix* diukur

menggunakan metrik *True Positive* (TP), *False Negative* (FN), *False Positive* (FP), dan *True Negative* (TN) [19], sebagaimana ditunjukkan pada Tabel 2.

Tabel 2. *Confusion Matrix* [19]

<i>Confusion Matrix</i>		Prediksi	
		<i>Benign</i>	<i>Malignant</i>
Kelas Sebenarnya	<i>Benign</i>	TP	FP
	<i>Malignant</i>	FN	TN

True Positive (TP) = jumlah kasus *benign* yang benar diklasifikasikan sebagai *benign*

True Negative (TN) = jumlah kasus *malignant* yang benar diklasifikasikan sebagai *malignant*

False Positive (FP) = jumlah kasus *benign* yang salah diklasifikasikan sebagai *malignant*

False Negative (FN) = jumlah kasus *malignant* yang salah diklasifikasikan sebagai *benign*

Confusion Matrix berdasarkan evaluasi kinerja model terdiri dari metrik utama, yaitu *presisi*, *recall*, dan *F1-score* yang masing-masingnya dirumuskan sebagai berikut:

$$\text{Presisi} = \frac{TP}{TP + FP} \quad (9)$$

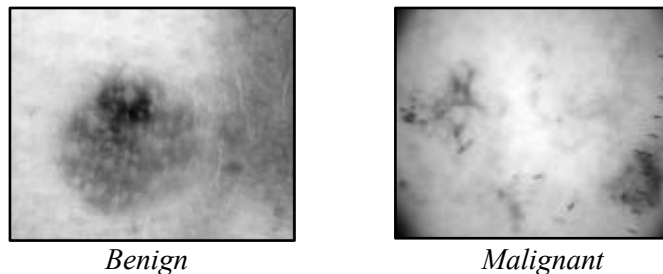
$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

$$\text{F1-score} = \frac{2 (\text{Presisi} \times \text{Recall})}{(\text{Presisi} + \text{Recall})} \quad (11)$$

3 HASIL DAN PEMBAHASAN

3.1 Uji Visualisasi Citra

Visualisasi merupakan tahap fundamental dalam analisis citra yang bertujuan untuk mengeksplorasi distribusi dan karakteristik data visual. Melalui tahapan ini, jenis kanker kulit berdasarkan tingkat keganasannya dapat dikenali lebih jelas, termasuk tekstur dan ciri pendukung lainnya. Pada tahap ini, masing-masing satu gambar dari kategori *benign* dan *malignant* diambil dari folder data pelatihan dan pengujian. Gambar yang awalnya berwarna (RGB) diubah menjadi citra abu-abu (*grayscale*). Konversi ini dilakukan untuk menghilangkan elemen warna yang tidak diperlukan dan menonjolkan tekstur serta kontras guna memudahkan pengamatan pola khas antara jinak dan ganas.

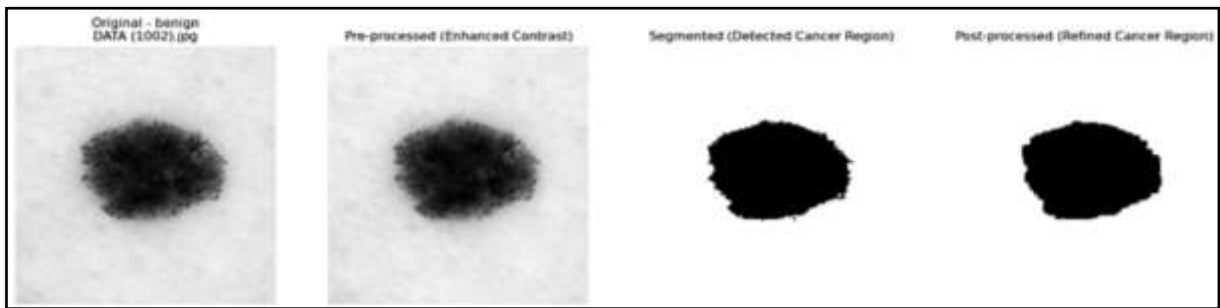


Gambar 3. Contoh Citra Dermoskopi pada Kanker Kulit [12]

3.2 Image Pre-Processing

Image pre-processing merupakan serangkaian teknik awal dalam pemrosesan citra digital yang berfungsi untuk meningkatkan kualitas citra serta menyiapkannya agar lebih optimal untuk dianalisis oleh sistem atau algoritma. *Pre-processing* citra dilakukan sebagai langkah awal dalam alur klasifikasi untuk menyesuaikan format dan struktur data citra sesuai kebutuhan model. Tahapan ini mencakup konversi citra ke format *grayscale* guna menghilangkan informasi warna yang tidak diperlukan, sehingga fokus analisis tertuju pada tekstur kulit. Citra kemudian diubah ukurannya menjadi 200x200 piksel untuk menjaga keseragaman dimensi dalam dataset. Intensitas piksel juga dinormalisasi ke dalam rentang 0 hingga 1 agar menghindari masalah perbedaan skala nilai piksel pada citra.

Langkah selanjutnya adalah segmentasi citra, yang bertujuan memisahkan area lesi kulit dari latar belakang berdasarkan perbedaan intensitas. Setelah objek diisolasi, diterapkan proses *post-processing* untuk mengurangi *noise* dan menyempurnakan hasil segmentasi.



Gambar 4. Hasil Serangkaian Proses *Image Pre-Processing* pada Citra Dermoskopi [12]

3.3 Ekstraksi Fitur *Gray Level Co-occurrence Matrix* (GLCM)

Ekstraksi fitur dilakukan sebagai tahap pemrosesan data dalam mengubah data citra mentah menjadi bentuk angka yang lebih bermakna. Pada penelitian ini, metode *Gray-Level Co-Occurrence Matrix* (GLCM) diterapkan untuk menganalisis pola dan struktur tekstur citra dengan mengevaluasi hubungan statistik antar pasangan piksel pada tingkat keabuan tertentu. GLCM menghitung frekuensi pasangan piksel yang muncul pada empat arah (θ) berbeda (0° , 45° , 90° , dan 135°) dengan jarak (d) satu piksel [20]. Hasil keseluruhan ekstraksi fitur menghasilkan 2.637 data latih dan 660 data uji dalam setiap sampel direpresentasikan oleh 16 fitur tekstur GLCM, dengan menghasilkan total 52.752 fitur untuk proses klasifikasi.

Tabel 3. Rata-rata Keseluruhan Dataset dalam Fitur GLCM

No	Fitur GLCM	Arah Orientasi Sudut				Rata-rata
		0°	45°	90°	135°	
1	Kontras	0.0108	0.0157	0.0089	0.0156	0.0127
2	Energi	0.8256	0.8223	0.8265	0.8223	0.8242
3	Entropi	0.7540	0.7821	0.7430	0.7820	0.7653
4	Homogenitas	0.9946	0.9922	0.9956	0.9922	0.9936

Berdasarkan data yang ditampilkan dalam Tabel 3, menunjukkan dari setiap nilai rata-rata fitur GLCM yang diekstraksi untuk keseluruhan dataset. Nilai tersebut menggambarkan tekstur citra dalam dataset. Fitur kontras, yang mengukur variasi intensitas antara piksel yang berdekatan, memiliki rata-rata 0,0127, menunjukkan adanya perbedaan tingkat kecerahan yang mencerminkan keberagaman pola tekstur. Fitur energi, yang merepresentasikan tingkat keseragaman atau kekompakan tekstur, mencatat nilai rata-rata 0,8242, menandakan bahwa sebagian besar citra memiliki pola yang relatif seragam. Sementara itu, entropi, yang mengukur tingkat keragaman atau kompleksitas tekstur, memiliki rata-rata 0,7653, mengindikasikan adanya variasi ketidakteraturan dalam pola visual lesi. Adapun homogenitas, yang menilai kedekatan nilai piksel yang serupa dalam citra, mencatat nilai rata-rata tertinggi yaitu 0,9936, mengindikasikan bahwa sebagian besar tekstur pada citra bersifat halus dan teratur [21].

Berdasarkan data Tabel 3 tersebut, terlihat bahwa fitur homogenitas memiliki nilai rata-rata tertinggi sebesar 0,9936, diikuti oleh energi sebesar 0,8242, entropi sebesar 0,7653, dan kontras sebesar 0,0127. Nilai-nilai tersebut menunjukkan bahwa homogenitas merupakan fitur yang paling dominan dalam karakteristik tekstur citra dermoskopi secara keseluruhan. Hal ini mengindikasikan bahwa mayoritas lesi kulit dalam dataset cenderung memiliki pola tekstur yang seragam, yang merupakan ciri umum dari lesi jinak (*benign*). Sebaliknya, variasi nilai entropi dan kontras yang lebih tinggi biasanya ditemukan pada lesi ganas (*malignant*), karena struktur permukaannya yang lebih kompleks dan tidak teratur.

3.4 Modelling

Tahap pemodelan ini disusun untuk menjawab pertanyaan penelitian yang tercantum pada bagian rumusan masalah dan tujuan penelitian, serta bertujuan memperoleh solusi terhadap permasalahan melalui implementasi model klasifikasi yang telah dirancang. Setiap model dilatih menggunakan data latih untuk melakukan prediksi, kemudian dievaluasi kinerjanya. Evaluasi dilakukan untuk menilai kemampuan model dalam mengklasifikasikan pola data dan menemukan aspek yang perlu diperbaiki atau optimasi lebih lanjut.

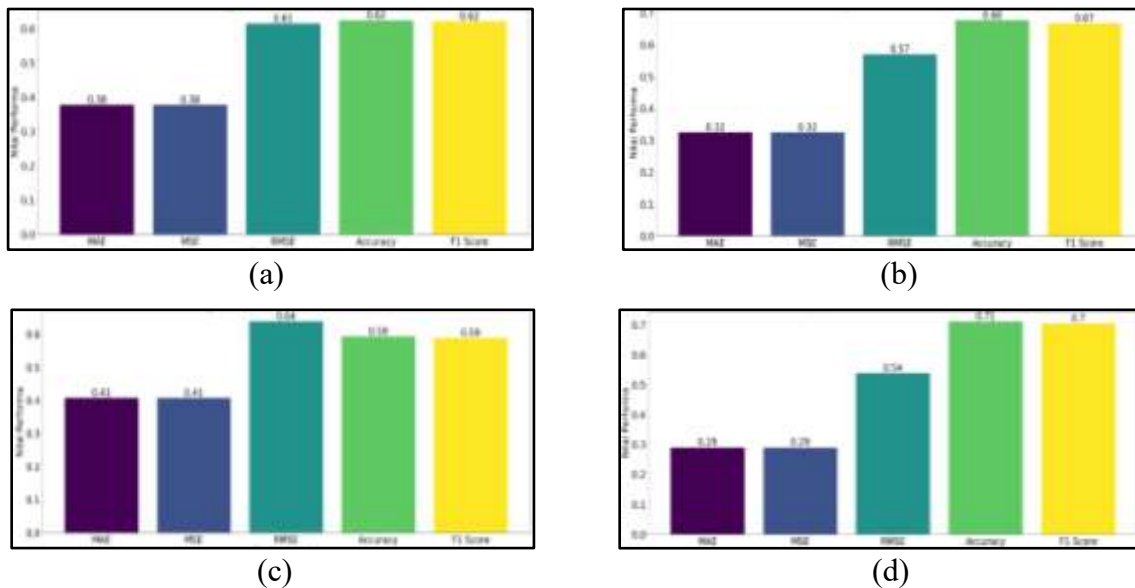
Tabel 4. Hasil Aktual dan Prediksi Empat Algoritma pada Proses Pelatihan

No.	Hasil Aktual	Hasil Prediksi Algoritma <i>Machine Learning</i>			
		SVM	K-NN	DT	RF
0	<i>Malignant</i>	<i>Benign</i>	<i>Malignant</i>	<i>Benign</i>	<i>Benign</i>
1	<i>Malignant</i>	<i>Benign</i>	<i>Malignant</i>	<i>Malignant</i>	<i>Malignant</i>
2	<i>Benign</i>	<i>Malignant</i>	<i>Benign</i>	<i>Benign</i>	<i>Benign</i>
...	...	...	...	...	...
525	<i>Benign</i>	<i>Malignant</i>	<i>Malignant</i>	<i>Benign</i>	<i>Malignant</i>
526	<i>Malignant</i>	<i>Malignant</i>	<i>Benign</i>	<i>Benign</i>	<i>Malignant</i>
527	<i>Malignant</i>	<i>Benign</i>	<i>Malignant</i>	<i>Malignant</i>	<i>Malignant</i>

Hasil prediksi menunjukkan bahwa SVM mengklasifikasikan 275 kasus sebagai *benign* dan 253 sebagai *malignant*, namun dibandingkan data aktual (294 *benign*, 234 *malignant*). Hal ini terjadi karena SVM bekerja optimal pada data dengan batas kelas yang

kelas. Pada dataset ini, terdapat tumpang tindih tekstur antar kelas yang membuat *margin* optimal sulit ditentukan, dimana SVM sensitif terhadap distribusi dan *outlier*. Model K-NN cenderung *over-predict benign* (321 *benign*, 207 *malignant*). Hal ini disebabkan oleh sifat K-NN yang sangat bergantung pada distribusi data sekitar. Karena kelas *benign* lebih dominan, model cenderung bias ke kelas tersebut, hal ini K-NN sensitif terhadap ketidakseimbangan kelas lokal. Sementara itu, *Decision Tree* (287 *benign*, 241 *malignant*) dapat dijelaskan oleh kemampuannya dalam melakukan pemisahan data berdasarkan atribut yang paling informatif. Terakhir, *Random Forest* (313 *benign*, 215 *malignant*) menunjukkan kemampuannya mengurangi *overfitting* dengan mengombinasikan hasil dari banyak pohon secara acak dan menjaga kestabilan prediksi meskipun terdapat fitur bervariasi dalam data.

Penelitian ini juga menerapkan berbagai metrik evaluasi untuk menilai performa model, khususnya dalam membandingkan hasil prediksi dengan nilai aktual. Metrik yang digunakan meliputi MAE, MSE, dan RMSE dengan mengukur seberapa dekat prediksi model pada nilai sebenarnya berdasarkan tingkat kesalahan setiap model algoritma.



Gambar 5. Metrik Evaluasi Pelatihan (a) SVM (b) K-NN (c) *Decision Tree* (d) *Random Forest*

Evaluasi performa menunjukkan bahwa algoritma *Random Forest* menunjukkan performa paling unggul dengan nilai MAE dan MSE sebesar 0,289772 serta RMSE 0,538305, menandakan tingkat kesalahan prediksi yang paling rendah di antara seluruh model. K-NN menyusul dengan MAE dan MSE sebesar 0,323863 dan RMSE 0,569090 yang mencerminkan kestabilan model meskipun masih terdapat bias terhadap kelas mayoritas. Sementara itu, SVM mencatat MAE dan MSE sebesar 0,376893 serta RMSE 0,613916, menunjukkan performa sedang dengan kesalahan yang cukup tinggi. *Decision Tree* memiliki performa paling rendah dengan MAE dan MSE sebesar 0,407196 serta RMSE 0,638119, mengindikasikan sensitivitas tinggi terhadap struktur dan distribusi data.

3.5 Hasil Klasifikasi Performa pada Algoritma *Machine Learning*

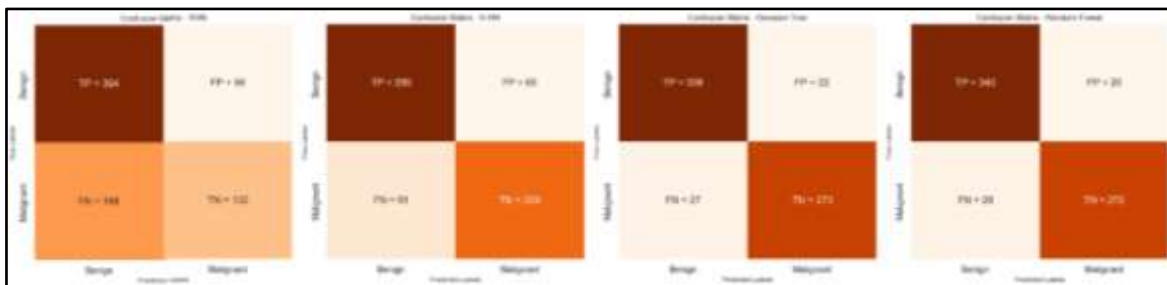
Hasil aktual mengacu pada kategori asli yang terdapat dalam dataset, yakni jenis kanker kulit yang sebenarnya berdasarkan diagnosis medis atau pemeriksaan sebelumnya. Hasil prediksi model algoritma *machine learning* adalah *output* yang dihasilkan oleh model klasifikasi. Hal ini perlu dilakukan untuk membandingkan hasil prediksi dengan hasil aktual guna menilai kinerja model. Sebuah model dianggap efektif dan dapat diandalkan jika hasil prediksinya sangat sesuai dengan hasil aktual dengan tingkat akurasi yang tinggi.

Model *Support Vector Machine* (SVM) memprediksi 472 kasus sebagai *benign* dan 188 kasus sebagai *malignant*, sementara *K-Nearest Neighbor* (K-NN) memprediksi 386 kasus sebagai *benign* dan 274 kasus sebagai *malignant*. Model *Decision Tree* (DT) memprediksi 365 kasus sebagai *benign* dan 295 kasus sebagai *malignant*, sedangkan *Random Forest* (RF) memprediksi 368 kasus sebagai *benign* dan 292 kasus sebagai *malignant*. Meskipun hasilnya bervariasi ditinjau secara keseluruhan, maka model-model ini berhasil membedakan lesi *benign* dan *malignant* dengan tingkat akurasi yang berbeda-beda.

Tabel 5. Hasil Aktual dan Prediksi Empat Algoritma pada Proses Pengujian

No.	Hasil Aktual	Hasil Prediksi Algoritma <i>Machine Learning</i>			
		SVM	K-NN	DT	RF
0	<i>Benign</i>	<i>Benign</i>	<i>Benign</i>	<i>Benign</i>	<i>Benign</i>
1	<i>Benign</i>	<i>Benign</i>	<i>Benign</i>	<i>Benign</i>	<i>Benign</i>
2	<i>Benign</i>	<i>Benign</i>	<i>Benign</i>	<i>Benign</i>	<i>Benign</i>
...					
658	<i>Malignant</i>	<i>Malignant</i>	<i>Malignant</i>	<i>Malignant</i>	<i>Malignant</i>
659	<i>Malignant</i>	<i>Malignant</i>	<i>Malignant</i>	<i>Malignant</i>	<i>Malignant</i>

a) *Confusion Matrix*



Gambar 6. Evaluasi *Confusion Matrix* untuk Setiap Algoritma *Machine Learning*

Berdasarkan evaluasi *Confusion Matrix* sebagaimana ditunjukkan pada Gambar 6, algoritma SVM memperoleh performa klasifikasi yang cukup. Tercatat sebanyak 304 *True Positive* (TP), yang berarti model berhasil mengklasifikasikan citra kulit *benign* secara tepat. Kemudian, terdapat 132 *True Negative* (TN), menunjukkan bahwa model mampu mengenali citra kanker kulit *malignant* dengan benar. Namun, terdapat 168 *False Negative* (FN), yaitu citra kanker kulit yang sebenarnya *malignant* tetapi salah diklasifikasikan sebagai *benign*, yang dapat berisiko menyebabkan keterlambatan diagnosis. Sementara itu, 56 *False Positive*

(FP) menunjukkan bahwa model salah mengklasifikasikan citra *benign* sebagai *malignant*, yang berpotensi menimbulkan kekhawatiran atau pengobatan yang tidak diperlukan. Model K-NN memiliki performa yang berbeda dengan 295 TP, 209 TN, 91 FN, dan 65 FP, menunjukkan sensitivitas yang lebih tinggi terhadap *malignant* namun tetap mengalami kesulitan dalam mengenali kanker kulit ganas. *Decision Tree* berhasil mengklasifikasikan 338 TP dan 273 TN, dengan hanya 27 FN dan 22 FP, tetap efektif meskipun terdapat beberapa kesalahan klasifikasi. *Random Forest* menunjukkan hasil terbaik, di mana 340 TP, 272 TN, 28 FN, dan hanya 20 FP, dengan kesalahan yang lebih rendah sehingga mampu menghasilkan akurasi tertinggi dan klasifikasi yang lebih seimbang.

b) Evaluasi Performa Algoritma *Machine Learning*

Tabel 6. Hasil Evaluasi Performa Empat Model Algoritma pada Proses Pengujian

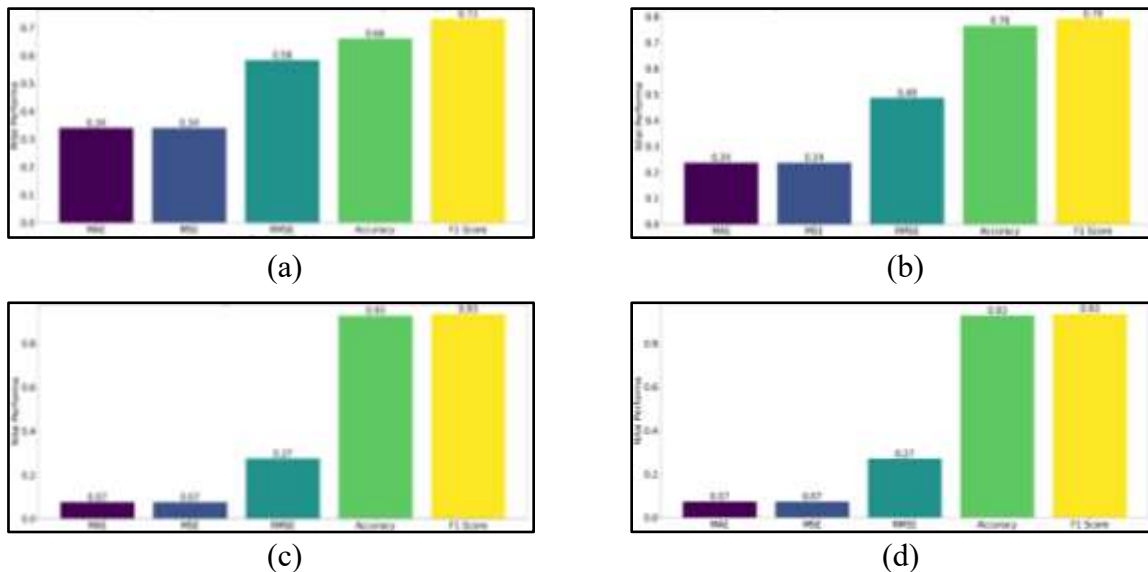
Performa	SVM	K-NN	<i>Decision Tree</i>	<i>Random Forest</i>
Akurasi	66,06%	76,36%	92,57%	92,72%
Presisi	84,44%	81,94%	93,88%	94,44%
<i>Recall</i>	64,40%	76,42%	92,60%	92,39%
<i>F1-Score</i>	73,07%	79,08%	93,24%	93,40%

Berdasarkan hasil evaluasi performa empat model algoritma dalam Tabel 6, *Random Forest* menunjukkan kinerja terbaik di antara model yang diuji. Keunggulan ini berasal dari kemampuannya sebagai model *ensemble* yang menggabungkan banyak pohon keputusan, sehingga mampu menangani variabilitas data dan kompleksitas tekstur citra secara lebih stabil dan akurat. Hal ini tercermin dari capaian akurasi tertinggi sebesar 92,72% yang menunjukkan tingkat prediksi benar paling tinggi dibandingkan total prediksi. Model ini juga unggul dalam metrik presisi, dengan nilai 94,44%, menunjukkan tingkat ketepatan tinggi dalam mengklasifikasikan kasus positif. Diikuti oleh *Decision Tree* dengan presisi 93,88%, SVM sebesar 84,44%, dan K-NN sebesar 81,94%. Dalam metrik *recall*, yang menunjukkan seberapa baik model mengenali semua kasus positif, *Decision Tree* mencatatkan nilai tertinggi sebesar 92,60%, diikuti oleh *Random Forest* (92,39%), K-NN (76,42%), dan SVM (64,40%), yang menandakan bahwa SVM memiliki kesulitan dalam mendeteksi seluruh kasus positif, karena sensitif terhadap distribusi data dan kurang efektif dalam mengenali pola kompleks dari fitur tekstur GLCM, terutama tanpa penyesuaian kernel atau parameter yang optimal. Dari sisi *F1-Score*, harmonisasi antara presisi dan *recall*, *Random Forest* kembali menjadi model terbaik dengan nilai 93,40%, disusul oleh *Decision Tree* (93,24%), K-NN (79,08%), dan SVM dengan nilai *F1-Score* terendah sebesar 73,07%.

Hasil penelitian ini sejalan dengan temuan Rahagiyanto et al. (2025) [22], yang menunjukkan bahwa algoritma *Random Forest* memiliki kinerja klasifikasi yang lebih unggul dibandingkan *Support Vector Machine* (SVM), *K-Nearest Neighbor* (K-NN), dan *Decision Tree*. Performa superior *Random Forest* ini didasarkan pada kemampuannya dalam mengelola data yang kompleks dan mengurangi *overfitting* melalui penggabungan beberapa pohon keputusan untuk menghasilkan prediksi akhir yang lebih akurat. Di sisi lain, algoritma

SVM kurang optimal ketika menghadapi data yang tidak terpisah secara linear tanpa pemilihan kernel yang tepat. Dengan demikian, *Random Forest* dinilai mampu dalam menghasilkan performa klasifikasi yang konsisten dan akurat.

c) Evaluasi Metrik MAE, MSE, dan RMSE



Gambar 7. Metrik Evaluasi Pengujian (a) SVM (b) K-NN (c) *Decision Tree* (d) *Random Forest*

Berdasarkan hasil metrik evaluasi sebagaimana ditunjukkan pada Gambar 7, menjelaskan bahwa *Random Forest* menunjukkan performa terbaik di antara seluruh algoritma dengan nilai MAE dan MSE sebesar 0,072720 serta RMSE 0,269679, yang mengindikasikan prediksi yang sangat akurat dan tingkat kesalahan yang paling rendah. Di sisi lain, SVM mencatat performa terburuk dengan MAE dan MSE sebesar 0,339393 serta RMSE 0,582575, menunjukkan bahwa model ini menghasilkan kesalahan prediksi tertinggi dikarenakan kurang menyesuaikan dengan kompleksitas dan distribusi data tekstur citra.

Dalam penelitian ini, beberapa keterbatasan yang perlu dicermati untuk menjaga objektivitas dan kredibilitas kajian. Pertama, dataset yang digunakan berasal dari satu sumber terbuka, sehingga berpotensi mengandung bias representasi terhadap populasi nyata, baik secara demografis maupun variasi jenis lesi. Kedua, metode ekstraksi fitur GLCM memiliki keterbatasan dalam menangkap pola visual kompleks seperti bentuk tepi tidak teratur, warna campuran, atau gradasi halus yang umum dijumpai pada citra medis. GLCM hanya mengandalkan hubungan antar piksel berdasarkan arah dan jarak tertentu, sehingga mungkin kurang optimal dalam menggambarkan fitur-fitur abstrak. Terakhir, jumlah data pengujian terbatas juga menjadi hambatan dalam mengukur generalisasi model secara menyeluruh.

4 SIMPULAN

Penelitian ini berhasil mengembangkan model klasifikasi kanker kulit pada citra dermoskopi menggunakan ekstraksi fitur *Gray Level Co-occurrence Matrix* (GLCM) dan empat algoritma *machine learning*. Hasil temuan utama menunjukkan bahwa *Random Forest*

memberikan performa terbaik dengan akurasi 92,72%, presisi 94,44%, *recall* 92,39%, dan F1-Score 93,40%, karena kemampuannya sebagai model *ensemble* dalam menangani variasi data secara stabil dan efektif mengurangi risiko *overfitting*. Sebaliknya, SVM menunjukkan performa terendah dengan akurasi 66,06%, presisi 84,44%, *recall* 64,40%, dan F1-Score 73,07%, akibat kurang optimal dalam menangani data dengan distribusi kompleks dan kelas yang saling tumpang tindih. Secara keseluruhan, metode GLCM terbukti efektif dalam merepresentasikan tekstur citra untuk mendeteksi lesi kulit jinak dan ganas. Secara praktis, penelitian ini berkontribusi dalam mendukung deteksi dini kanker kulit secara otomatis, sementara secara ilmiah memberikan bukti bahwa metode statistik klasik yakni GLCM tetap relevan ketika dikombinasikan dengan algoritma klasifikasi yang tepat. Meskipun demikian, keterbatasan ini terletak pada ukuran dataset dan fitur GLCM yang belum mampu menangkap pola visual kompleks secara menyeluruh.

5 DAFTAR PUSTAKA

- [1] Y. Ananda, G. Edo, dan S. Adelia. (2024). *Buku Ajar Sistem Integumen*. Purbalingga: Eureka Media Aksara.
- [2] Enggar, Taqwin, Muliani, dan F. L. Abd. (2024) *Anatomi Fisiologi Untuk Mahasiswa Kesehatan*. Purbalingga: Eureka Media Aksara.
- [3] N. Martin dan U. Daniel. (2024). Klasifikasi Kanker Kulit Pada Citra Dermatoskopi Menggunakan CNN. *Jurnal Algoritme*, Vol. 5 (1), ISSN: 2775-8796. Doi: 10.35957/algoritme.v5i1.9034.
- [4] A. O. Ardiansyah. (2024). *Kanker Kulit dan Sarkoma Jaringan Lunak - Dari Teori Preklinik Hingga Aplikasi Klinik*. Surabaya: Airlangga University Press.
- [5] A. Ameri. (2020). A Deep Learning Approach to Skin Cancer Detection In Dermoscopy Images. *J Biomed Phys Eng*, Vol. 10 (6). Doi: 10.31661/Jbpe.V0i0.2004-1107.
- [6] R. Suresh. (2024). Revolutionizing Physics: A Comprehensive Survey of Machine Learning Applications. *Frontiers in Physics*. Doi: 10.3389/Fphy.2024.1322162.
- [7] Neneng, S. P. Ajeng, dan A. A. Ahmad Ari. (2021). Perbandingan Hasil Klasifikasi Jenis Daging Menggunakan Ekstraksi Ciri Tekstur Gray Level Co-Occurrence Matrices (GLCM) dan Local Binary Pattern (LBP). *Smatika Jurnal*, Vol. 11 (1), ISSN: 2087-0256. Doi: 10.32664/Smatika.V11i01.572.
- [8] L. M. Wisudawati. (2021). Klasifikasi Tumor Jinak Dan Tumor Ganas Pada Citra Mammogram Menggunakan Gray Level Co-Occurrence Matrix (GLCM) Dan Support Vector Machine (SVM). *Jurnal Ilmiah Informatika Komputer*, Vol. 26 (2), Doi: 10.35760/Ik.2021.V26i2.4897.
- [9] S. Suharyana, A. Fuad, D. Armylia Chandra, Y. Mohtar, S. Umi, dan C. Rifai. (2023). Pneumonia Classification Based On GLCM Features Extraction Using K-Nearest Neighbor. *Indonesian Journal of Applied Physics (IJAP)*, Vol. 13 (2), ISSN: 2089 – 0133. Doi: 10.13057/ijap.v13i2.77120.
- [10] M. Yudianto, D. M. Rizka, dan S. Esti. (2024). Using *Decision Tree* With First And Second-Order Statistical Feature Extraction For Classification Of Lung Cancer. *Indonesian Journal of Applied Physics (IJAP)*. Vol. 14 (2), ISSN: 2089 – 0133. Doi: 10.13057/Ijap.V14i2.87676.

- [11] W. Wijaya, R. P. Muhammad, dan P. W. Eka. (2023). Klasifikasi Monkeypox Menggunakan Ekstraksi Fitur GLCM dan Algoritma *Random Forest*. *2nd Mdp Student Conference (MSC)*, Vol 2 (1), E-ISSN: 2985-7406. Doi: 10.35957/Mdp-Sc.V2i1.4435.
- [12] Kaggle, "Skin Cancer: Malignant vs. Benign," *Kaggle Datasets*, [Online]. Available: <https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign>.
- [13] J. S. Ananda, F. Yoza, dan P. Jesi. (2024). Classification Analysis Of Brain Tumor Disease In Radiographic Images Using Support Vector Machines (SVM) With Python. *Journal Online Physics*, Vol. 9 (3), ISSN: 2502-2016. Doi: 10.22437/Jop.V9i3.3627.
- [14] M. I. I. Fata dan A. Donny. (2024). Penerapan Metode Naive Bayes Pada Sistem Klasifikasi Kualitas Biji Kopi Robusta. *Jurnal Indonesia : Manajemen Informatika dan Komunikasi*, Vol. 5 (1), ISSN: 2723-7079. Doi: 10.35870/Jimik.V5i1.515.
- [15] H. Sumarti, S. Qolby, T. Devi, S. Fahira, dan P. D. R. Tara. Identification of Covid-19 Based on Features Texture Histogram and Gray Level Co-Occurrence Matrix (GLCM) Using K-Means Clustering Methods in Chest X-Ray Digital Images. (2023). *Jurnal Penelitian Fisika dan Aplikasinya (JPFA)*, Vol. 13 (1), ISSN: 2087-9946. Doi: 10.26740/Jpfa.V13n1.P51-66.
- [16] F. S. Lesiangi, A. Y. Mauko, dan B. S. Djahi. (2021). Feature Extraction Hue, Saturation, Value (HSV) and Gray Level Cooccurrence Matrix (GLCM) for Identification of Woven Fabric Motifs in South Central Timor Regency. *Journal of Physics: Conference Series*. Doi: 10.1088/1742-6596/2017/1/012010.
- [17] Y. Amelia. (2023). Perbandingan Metode Machine Learning untuk Mendeteksi Penyakit Jantung. *Idealis: Indonesia Journal Information System*, Vol. 6 (2), ISSN: 2684-7280. Doi: 10.36080/Idealis.V6i2.3043.
- [18] A. I. Sang, S. Edi, dan D. Irfan. (2021). Analisis Data Mining untuk Klasifikasi Data Kualitas Udara DKI Jakarta Menggunakan Decision Tree dan Support Vector Machine. *E-Proceeding Of Engineering Telkom University*, Vol. 8 (5), ISSN : 2355-9365.
- [19] D. Normawati dan A. P. Surya. (2021). Implementasi Naïve Bayes Classifier dan *Confusion Matrix* pada Analisis Sentimen Berbasis Teks pada Twitter. *Sains Komputer & Informa (J-SAKTI)*, Vol. 5 (2), ISSN: 2548-9771. Doi: 10.30645/j-sakti.v5i2.369.
- [20] D. M. Rani, D. Frastica, dan F. Yoza. (2024). Classification of Lung Disease on X-Ray Images Based on Gray Level Co-Occurrence Matrix (GLCM) Feature Extraction and Backpropagation Neural Network Using Python Gui. *Journal Online Physics*, Vol. 9 (2), ISSN: 2502-2016. Doi: 10.22437/Jop.V9i2.32806.
- [21] A. Maulida, Nurhidayah, F. Yoza, dan Haryono. (2022). Segmentasi Citra Mammogram Untuk Deteksi Dini Kanker Payudara Dengan Menggunakan Metode Otsu Thresholding. *Jurnal Fisika Unand*, Vol. 11 (2), ISSN: 2302-8491. Doi: 10.25077/Jfu.11.2.180-186.2022.
- [22] A. Rahagiyanto, H. P. Bakhtiyar, Y. Muhammad, V. Veronika, E. J. S. Gandu, dan D. Atma. (2025). Perbandingan Kinerja Algoritma K-NN-DT-RF-SVM untuk Deteksi Dini Risiko Kematian Ibu. *J-REMI : Jurnal Rekam Medik dan Informasi Kesehatan*, Vol. 6 (2), ISSN: 2721-866X. Doi: 10.25047/J-Remi.V6i2.5658.