

## Early Detection Of Melanoma Skin Cancer Using Gray Level Co-Occurrence Matrix And Ensemble Support Vector Machine

Mustagfirin\*<sup>1</sup>, Rony Wijanarko<sup>2</sup>, Arif Rifan Rudiyanto<sup>3</sup>, Abdullah Afnil Hisbana<sup>4</sup>, Fitrotin Na'imul Farida<sup>5</sup>

<sup>1,2,3,4,5</sup>Department of Informatics Engineering , Universitas Wahid Hasyim, Indonesia

Email: [mustagfirin@unwahas.ac.id](mailto:mustagfirin@unwahas.ac.id)

Received : Oct 26, 2025; Revised : Dec 10, 2025; Accepted : Dec 10, 2025; Published : Apr 15, 2026

### Abstract

Skin cancer is a major global health problem with incidence rates increasing every year. Melanoma, the most aggressive form of skin cancer, requires accurate early detection to reduce mortality risk. Conventional diagnostic methods such as visual examination and biopsy still face limitations in precision and consistency, highlighting the need for more objective and efficient technological approaches. This study proposes a classification method for melanoma using an ensemble of Support Vector Machine (SVM) and Random Forest (RF), supported by feature extraction through the Gray Level Co-occurrence Matrix (GLCM) and dimensionality reduction using Linear Discriminant Analysis (LDA). The research stages include image preprocessing using grayscale conversion to reduce data complexity, followed by GLCM-based texture feature extraction, and LDA transformation to enhance class separability. The classification model is developed using an ensemble voting mechanism that combines predictions from SVM and RF to produce a more stable and robust decision. Experimental results with a 60:40 train-test ratio show that the proposed method achieves an accuracy of 88.75%, outperforming each individual model tested. These findings indicate that the integration of GLCM-LDA features with the SVM-RF ensemble effectively improves melanoma detection performance. Overall, this study provides a significant contribution to the development of early detection systems in health informatics, offering potential improvements in patient safety and survival rates for individuals affected by skin cancer.

**Keywords :** *Gaussian Mixture Models, Gray Level Co-occurrence Matrix, machine learning, melanoma, skin cancer, Support Vector Machine, Random Forest.*

This work is an open access article and licensed under a Creative Commons Attribution-Non Commercial 4.0 International License



### 1. PENDAHULUAN

Kanker adalah penyakit degeneratif yang dapat terjadi pada semua usia dan seluruh bagian tubuh. Sel kanker tumbuh secara tidak terkontrol dan dapat menyebabkan komplikasi bahkan kematian[1]. Ada banyak faktor yang mempengaruhi timbulnya sel kanker, termasuk gaya hidup, radiasi, mutasi genetik, virus, dan faktor keturunan. Salah satu jenis kanker kulit adalah melanoma, yang paling ganas dan menyerang melanosit[2]. Ada beberapa ciri kanker kulit, termasuk timbulnya titik atau bintik yang semakin membesar, berubah warna, memiliki bagian tepi yang berbeda, dan muncul ke permukaan kulit dengan bentuk yang tidak beraturan[3].

Penyakit ini adalah salah satu jenis kanker yang paling sering terjadi di seluruh dunia. Di Amerika Serikat saja, pada tahun 2020 diperkirakan akan terdapat 196.060 kasus baru kanker kulit non-melanoma dan 100.350 kasus baru melanoma[1]. Di Indonesia, kanker kulit juga menjadi masalah kesehatan yang cukup serius. Menurut data dari Kementerian Kesehatan Indonesia, pada tahun 2018 terdapat 11.033 kasus kanker kulit, dengan jumlah kasus terbanyak terjadi di Provinsi Jawa Timur. Oleh karena itu, deteksi dini kanker kulit sangat penting untuk meningkatkan kesembuhan dan mengurangi risiko kematian[4].

Pentingnya deteksi dini dalam mengatasi kanker kulit telah diakui secara global. Metode deteksi dini yang paling umum dilakukan adalah dengan pemeriksaan kulit secara visual oleh dokter dan biopsi jaringan kulit atau dengan menggunakan teknik ABCD (Asymetry, Border, Color, Diameter) untuk mengetahui tanda-tanda awal kanker kulit[3]. Namun, metode ini memiliki kelemahan dalam hal keakuratan dan ketepatan waktu dalam mendeteksi kanker kulit.

Oleh karena itu, penggunaan teknologi komputer dan metode analisis citra digital telah menjadi semakin penting dalam deteksi kanker kulit. Teknologi komputer dan pengolahan citra dapat digunakan dalam deteksi kanker kulit[5]. Beberapa peneliti melakukan percobaan dengan menggunakan metode berbasis teknologi machine learning. Berdasarkan hal tersebut maka penelitian ini mengusulkan metode machine learning yaitu ensemble Support Vector Machine (SVM)[6] dan Random Forest (RF) berbasis ekstraksi fitur untuk mendeteksi kanker kulit secara akurat dan efisien. Metode ini akan memanfaatkan fitur-fitur seperti ekstraksi fitur Red Green Blue (RGB), tekstur Gray Level Co-occurrence Matrix (GLCM), dan ekstraksi ciri Linear Discriminant Analysis (LDA)[7][8][9].

Deteksi kanker kulit yang akurat dan efisien melibatkan teknologi machine learning dengan metode ensemble SVM dan RF[10]. Fitur-fitur seperti ekstraksi fitur RGB, tekstur GLCM, dan ekstraksi ciri LDA dimanfaatkan untuk mengidentifikasi pola-pola yang berkaitan dengan kanker kulit[7]. Dengan pendekatan ini, sistem dapat menghasilkan deteksi yang cepat dan akurat, serta memungkinkan penanganan dini yang lebih efektif.

Deteksi kulit memang merupakan tugas yang mudah bagi penglihatan manusia, tetapi bagi komputer terdapat beberapa faktor yang membatasi, khususnya dalam keperluan biometrik dan analisis objek. Faktor-faktor tersebut antara lain adalah pencahayaan yang tidak merata, peralatan optik, pseudo-skin background, serta fitur individual subjek seperti asal kelompok etnis yang dapat menyebabkan perbedaan warna kulit[11]. Tujuan segmentasi kulit adalah mengisolasi daerah kulit pada citra masukan yang tidak terbatas dalam lingkungan yang sulit, misalnya pada proses sejumlah besar citra web yang memiliki resolusi, kualitas, iluminasi, dan tingkat noise yang beragam[12]. Namun demikian, metode segmentasi kulit tradisional belum cukup kuat menghadapi lingkungan yang kompleks[13]. Variasi warna kulit dapat disebabkan oleh perbedaan ras, usia, kondisi iluminasi, serta objek lain yang memiliki warna mirip kulit. Hal tersebut menjadi masalah utama yang harus diselesaikan dalam segmentasi kulit[14]. Oleh karena itu, pemilihan ruang warna dapat dianggap sebagai langkah utama dalam proses segmentasi kulit.

Model warna kulit dapat dianggap sebagai distribusi warna kulit dalam suatu ruang warna. Gagasan utamanya adalah bahwa warna kulit menyatu dalam ruang warna tersebut dan dapat dipisahkan oleh satu atau beberapa batas[11]. Untuk meningkatkan akurasi deteksi warna kulit, salah satu metode yang dapat digunakan adalah Gaussian Mixture Models (GMM), yaitu metode pembelajaran mesin berbasis statistik[15][16]. Pengklasifikasi GMM menghasilkan deteksi yang lebih akurat dengan menggunakan estimasi kemungkinan maksimum untuk melatih pengklasifikasi sesuai model klasifikasi[11]. Model ini terdiri atas tiga bagian utama, yaitu transduser, ekstraktor fitur, dan pengklasifikasi[17]. Selain itu, transformasi wavelet merupakan teknik transformasi 2-D yang efektif sebagai alat dalam pengenalan pola maupun aplikasi lainnya. Transformasi wavelet menyediakan sarana untuk menganalisis sinyal input dalam beberapa tingkat resolusi yang berbeda secara hierarkis, sehingga dikenal juga sebagai analisis multiresolusi[18][19]. Dalam penelitian ini, segmentasi warna kulit dilakukan dengan menggabungkan fitur warna dan tekstur kulit untuk mensegmentasi daerah kulit pada citra berwarna[12][13]. Model warna kulit diekstrak oleh pengklasifikasi GMM untuk mengestimasi probability density function (pdf). Selanjutnya, parameter mixture model diestimasi melalui pelatihan

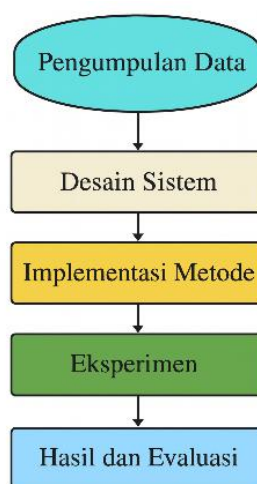
GMM pada dataset dengan menggunakan algoritma Expectation Maximization (EM). Setelah proses pelatihan selesai, model dapat digunakan pada citra baru untuk melakukan segmentasi kulit manusia.

Hasil studi komparatif menunjukkan bahwa Artificial Neural Network (ANN)[20] umumnya memberikan kinerja lebih tinggi dibandingkan Random Forest (RF)[21] pada data dengan pola non-linear dan fitur kontinu[22]. ANN mampu mencapai akurasi hingga 97.5% dengan preprocessing BFS, jauh melampaui Chi-Square ANN yang hanya memperoleh 91.4%, serta mengungguli RF yang rata-rata berada pada kisaran 79–81% dengan F1-score 74–83% pada rasio data 80:20. Keunggulan ANN[23] terletak pada kemampuannya sebagai universal function approximator yang efektif dalam menangkap hubungan kompleks pada data seperti deteksi penyakit atau prediksi energi bangunan dengan RMSE lebih rendah (4.97 vs 6.10 dibanding RF)[24]. Sebaliknya, RF lebih unggul dalam hal robustness terhadap overfitting dan bekerja baik pada data tabular, namun performanya dapat turun pada dataset tidak seimbang, terutama tanpa teknik resampling, sebagaimana terlihat dalam berbagai studi Indonesia yang hanya mencapai akurasi 77–80%. Gap penelitian muncul karena masih minimnya eksplorasi metode ensemble lanjutan (stacking, bagging, hybrid ANN-RF) pada data Indonesia yang bersifat heterogen dan kompleks seperti cuaca ekstrem[25], stunting multiclass[26], atau data kesehatan masyarakat. Model tunggal ANN/RF sering gagal menangani ketidakseimbangan kelas dengan gap akurasi hingga 13%, sementara kombinasi ensemble dengan teknik seperti SMOTE berpotensi meningkatkan generalisasi hingga 85–98% namun belum banyak diterapkan secara komprehensif.

Urgensi penelitian ini adalah pentingnya deteksi dini kanker kulit secara akurat untuk meningkatkan kesembuhan dan mengurangi risiko kematian. Hal ini sesuai dengan prioritas riset nasional (PRN) 2022-2024 khususnya dibidang Teknologi Informasi untuk kesehatan.

## 2. METODOLOGI PENELITIAN

Proses penelitian ini dimulai dengan pengumpulan data, diikuti oleh desain sistem. Setelah itu dilakukan implementasi metode, eksperimen dan evaluasi serta analisis terhadap hasil yang telah diperoleh, proses alur tahapan penelitian dapat dilihat pada Gambar 1. Model segmentasi warna kulit yang akan dikembangkan akan menggunakan fitur generasi warna kulit, fitur tekstur kulit, dan segmentasi tekstur kulit yang akan mensegmentasi daerah kulit dalam citra berwarna.



Gambar 1. Diagram Alur Tahapan Penelitian

Penelitian ini menggunakan data dari website kaggle.com. kemudian proses klasifikasi citra melanoma dimulai dengan melakukan pelatihan model, yang bertujuan agar model dapat mengenali karakteristik yang relevan dalam data pelatihan. Model dilatih untuk melakukan klasifikasi terhadap

kelas yang sesuai. Selama pelatihan, nilai kesalahan dihitung dan digunakan untuk mengoptimalkan parameter alpha dan bias dalam model. Setelah pelatihan selesai nilai tersebut disimpan dan siap diterapkan pada tahap pengujian, dimana model akan melakukan klasifikasi terhadap data baru yang belum pernah dilihat sebelumnya.

### Proses Training

Untuk mencapai klasifikasi yang akurat pada citra melanoma, model harus menjalani tahap pelatihan terlebih dahulu. Tujuan dari tahap ini adalah mendapatkan alpha dan bias yang optimal. Nilai ini akan berfungsi untuk membedakan antara kelas citra jinak dan ganas.

Dilanjutkan dengan pembuatan desain sistem untuk mendukung proses klasifikasi kanker kulit melanoma. Desain sistem tersebut memberikan gambaran menyeluruh mengenai sistem yang akan dikembangkan. Penerapan algoritma Gray Level Co-occurrence Matrix (GLCM) dan Ensemble Support Vector Machine (SVM).

Algoritma dalam penelitian ini akan menggunakan ensemble SVM dengan bantuan Gray Level Co-occurrence Matrix untuk melakukan ekstraksi fitur. Sebelum melakukan ekstraksi fitur dilakukan preprocessing untuk terlebih dahulu. Penelitian ini dilakukan prosedur latih dan prosedur uji. Dimana prosedur latih digunakan untuk memahami pola informasi yang dimasukkan sebagai input. Sedangkan prosedur uji digunakan untuk melakukan evaluasi peroforma SVM yang telah dilatih sebelumnya.

### Input Image

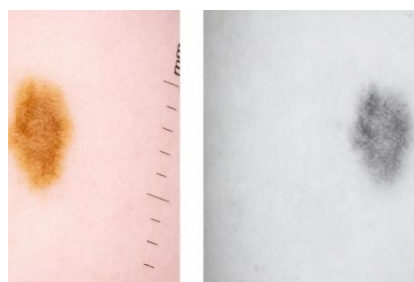
Data citra kanker kulit melanoma tersebut berformat .JPG dengan resolusi 300x300 piksel. Total citra yang digunakan terbagi dalam dua kelas, yaitu melanoma jinak dan ganas, dengan masing-masing kelas terdiri dari 3000 citra.

### Preprocessing Citra

Preprocessing citra adalah tahap yang dilakukan untuk mempersiapkan citra sebelum memasuki proses pelatihan.

### Grayscale

Citra grayscale adalah gambar yang terdiri atas variasi tingkat keabuan mulai dari hitam (0) hingga putih (255). Dengan berbagai tingkat abu-abu di antaranya. Proses grayscaling dilakukan untuk mengurangi dimensi dan kompleksitas gambar berwarna dengan mempertahankan informasi intensitas cahaya dari setiap piksel tanpa memperhitungkan warna, proses tersebut dapat dilihat pada Gambar 2. Proses grayscaling sangat penting dikarenakan dapat mempermudah perhitungan tekstur dan mengurangi kompleksitas yang muncul jika gambar RGB diproses langsung. Setiap piksel pada citra grayscale memiliki nilai intensitas 8-bit. Proses ini diperlukan untuk memastikan bahwa hanya intensitas yang dianalisis, sehingga informasi warna yang tidak diperlukan dapat dihilangkan, membuat perhitungan lebih efisien. Sehingga proses ini merupakan salah satu syarat sebelum melakukan ekstraksi fitur gray level co-occurrence matrix[27].



Gambar 2. Citra hasil *grayscale*

$$grayscale = 0.2989R + 0.578G + 0.114B$$

Keterangan:

$R$ ,  $G$ , dan  $B$  = intensitas warna merah, hijau, dan biru di koordinat piksel  $(x, y)$  pada citra warna. Koefisien 0.2989, 0.578, dan 0.114 = kontribusi relatif dari masing-masing saluran warna terhadap kecerahan piksel dalam citra grayscale.

**Gray Level Co-occurrence Matrix**

Ekstraksi fitur merupakan teknik untuk mendapatkan karakteristik yang mewakili sifat utama dengan memisahkannya dari fitur yang tidak diperlukan. Gray Level Co-occurrence Matrix (GLCM) merupakan teknik ekstraksi fitur yang cukup umum digunakan untuk melakukan analisis tekstur. GLCM bekerja dengan mempresentasikan hubungan antara dua piksel yang bertetangga[28].

$$glcmNorm = \frac{glcmValue}{\sum_i^N glcmValue}$$

Keterangan:

$glcmNorm$  = Nilai normalisasi dari elemen GLCM terkait.

$glcmValue$  = Nilai elemen GLCM pada koordinat  $(x, y)$ .

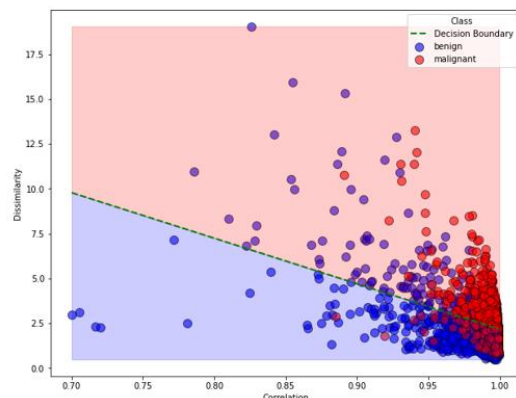
$original(x, y)$  = Nilai piksel asli pada koordinat  $(x, y)$  dalam citra yang digunakan untuk menghasilkan GLCM.

**Support Vector Machine**

Support Vector Machine (SVM) merupakan salah satu metode pembelajaran mesin dengan kategori terpadu (supervised). SVM bekerja dengan mencari hyperplane terbaik yang memaksimalkan margin (jarak) antara masing- masing kelas. Hyperplane merupakan batas pemisah antara dua kelas data berbeda[29].

SVM terbagi menjadi dua tipe, yakni SVM Linier dan Non-Linier. SVM Linier menangani data linier. Data terpisahkan secara linier oleh garis hyperplane melalui soft margin. Sedangkan, SVM Non-Linier menangani data tidak linier dengan kata lain data jauh lebih kompleks. Sehingga dalam mengatasi data kompleks dapat menggunakan “kernel trick”. Kernel trick merupakan metode mengubah data pada dimensi yang lebih tinggi. Sehingga dapat menghasilkan hyperplane yang lebih optimal.

Kernel Trick menggunakan fungsi matematis yang dapat mengubah data dari dimensi tertentu ke dimensi yang lebih tinggi. Pada penelitian ini, agar mengetahui kernel mana yang cocok untuk digunakan pada setiap fitur data adalah dengan cara membuat sebuah grafik perbandingan dari dataset yang digunakan. Akan tetapi penelitian ini menggunakan lebih dari 2 fitur, sehingga akan ditampilkan grafik setiap dua fitur yang digunakan. Adapun tujuan dari melakukan visualisasi grafik agar dapat mempermudah klasifikasi pada SVM.



Gambar 3. Grafik fitur *Dissimilarity* dan *Correlation*

Dapat dilihat perbandingan antara kelas jinak dan ganas pada fitur Dissimilarity dan Correlation. Grafik tersebut menampilkan Correlation pada sumbu x dan Dissimilarity pada sumbu y. Dari grafik tersebut dapat diamati bahwa penyebaran data dari kedua kelas membentuk kelas dan dapat terpisah secara garis lurus. Sehingga kernel linier mungkin merupakan pilihan yang tepat untuk melakukan klasifikasi.

**Ensemble SVM**

Dalam penelitian yang dilakukan oleh melakukan klasifikasi terhadap kanker kulit melanoma menggunakan Support Vector Machine dengan tiga kernel, yakni Linier, Polynomial dan RBF. Ketiga model kernel tersebut dapat memberikan akurasi yang memuaskan. Dengan demikian, pendekatan ensemble menggunakan majority vote diterapkan untuk meningkatkan akurasi secara lebih optimal[30].

$$f_{ensemble}(x) = mode(f_{linier}(x), f_{poly}(x), f_{rbf}(x))$$

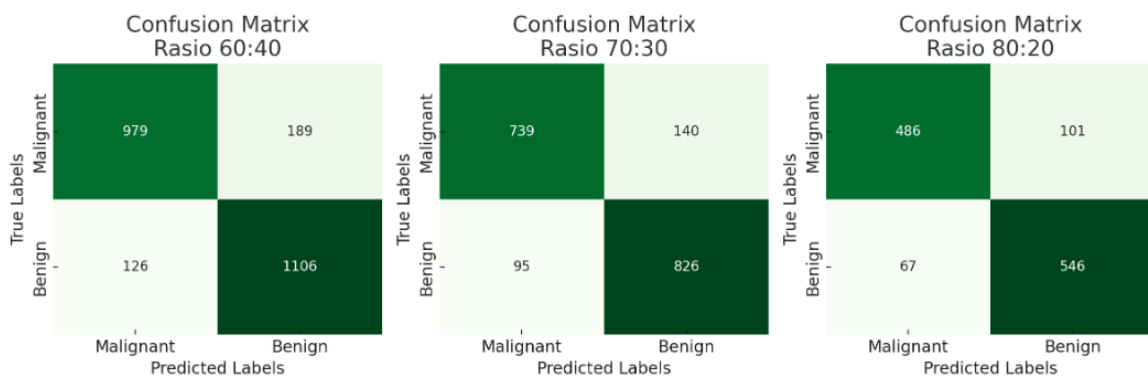
Keterangan:

mode = fungsi yang mengambil nilai mayoritas dari tiga prediksi.

**3. HASIL DAN PEMBAHASAN**

**Hasil**

Pengujian pertama dilakukan pada kernel individual, yakni kernel linier, rbf, dan polynomial yang terlihat pada Gambar 4.



Gambar 4. Confusion matrix kernel linier

Maka dapat diperoleh hasil perhitungan seperti ditunjukkan pada Tabel 1 confusion matrix untuk precision, recall, dan f1-score.

Tabel 1. Nilai precision, recall, dan f1-score pada kernel linier

Rasio	Kelas	Precision (%)	Recall (%)	F1-Score (%)
60:40	Ganas	88,6	83,82	86,14
	Jinak	85,41	89,77	87,53
70:30	Ganas	88,61	84,07	86,28
	Jinak	85,51	89,69	87,55
80:20	Ganas	87,88	82,79	85,26
	Jinak	84,39	89,07	86,67

Dari hasil uji coba terhadap setiap kernel individual 3 rasio data. Dapat diketahui bahwa kernel polynomial dengan degree 3 menghasilkan akurasi yang paling tinggi pada setiap rasio data seperti ditunjukkan pada Tabel 2.

Tabel 2. Hasil *accuracy*, *precision*, *recall* dan *f1-score* kernel individual

Rasio	Kernel	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
60:40	Linier	86,88	88,6	83,82	86,14
	RBF	86,75	85,41	89,77	87,53
	Polynomial $d = 3$	88,75	89,59	86,99	88,27
	Polynomial $d = 5$	85,29	83,82	86,47	85,12
	Polynomial $d = 7$	76,42	72,63	82,71	77,34
70:30	Linier	86,94	88,61	84,07	86,28
	RBF	87	87,81	85,21	86,49
	Polynomial $d = 3$	88,28	89,39	86,23	87,78
	Polynomial $d = 5$	85,39	84,94	85,32	85,08
	Polynomial $d = 7$	81,89	79,76	84,3	81,97
80:20	Linier	86	87,88	82,79	85,26
	RBF	86,75	87,28	85,35	86,3
	Polynomial $d = 3$	88,42	89,16	86,88	88,01
	Polynomial $d = 5$	84,92	85,24	83,65	84,44
	Polynomial $d = 7$	79	75,26	85,01	79,84

Evaluasi performa sistem klasifikasi sangat penting untuk menilai efektivitas model dalam memproses data dan menghasilkan prediksi yang akurat. Pada penelitian ini, pengujian dilakukan menggunakan beberapa kernel Support Vector Machine (SVM), yakni linier, Radial Basis Function (RBF), serta polynomial dengan berbagai derajat (3, 5, dan 7). Selain itu, dilakukan pula pendekatan ensemble yang menggabungkan kernel linier, RBF, dan polynomial derajat 3. Pengujian dilakukan pada berbagai skenario rasio pembagian data, yaitu 60:40, 70:30, dan 80:20.

Hasil pengujian menunjukkan bahwa kernel linier mencapai akurasi tertinggi pada rasio 70:30, yakni sebesar 86,94%, dengan nilai presisi 88,61%, recall 84,07%, dan f1-score 86,28%. Hal ini mengindikasikan bahwa kernel linier memiliki kemampuan generalisasi yang baik pada dataset dengan pembagian data tersebut. Kernel RBF menghasilkan akurasi tertinggi sebesar 86,75% pada dua rasio, yaitu 60:40 dan 80:20, meskipun nilai presisi, recall, dan f1-score bervariasi. Presisi tertinggi terdapat pada rasio 80:20 dengan nilai 87,28%, sementara recall dan f1-score tertinggi berada pada rasio 60:40, masing-masing 83,82% dan 86,14%. Hasil ini konsisten dengan literatur yang menyatakan bahwa kernel RBF efektif untuk menangani data non-linear, meski performanya dipengaruhi oleh parameter dan pembagian data[10].

Untuk kernel polynomial, derajat 3 menghasilkan akurasi terbaik, yakni 88,75% pada rasio 60:40, dengan presisi 89,59%, recall 86,99%, dan f1-score 88,27%. Hal ini menunjukkan bahwa polynomial  $d=3$  mampu menangkap pola data lebih kompleks dibanding kernel linier maupun RBF. Kernel polynomial derajat 5 mencapai akurasi tertinggi sebesar 85,39% pada rasio 70:30, sementara polynomial derajat 7 hanya mampu mencapai 81,89% pada rasio 70:30. Semakin tinggi derajat polynomial, kompleksitas model meningkat sehingga berpotensi menimbulkan overfitting, yang berdampak pada penurunan akurasi[31].

Selanjutnya, dilakukan pengujian dengan metode ensemble yang menggabungkan kernel linier, RBF, dan polynomial derajat 3. Hasil ensemble ditampilkan dalam Tabel 4.12 dan 4.13, dengan performa terbaik diperoleh pada rasio 60:40, yaitu akurasi 88,17%, presisi 89,25%, recall 86,04%, dan

f1-score 87,62%. Pada rasio 70:30, akurasi yang dicapai hampir setara, yakni 88,11%, sementara pada rasio 80:20 diperoleh akurasi 87,42%. Nilai ini menunjukkan bahwa ensemble mampu meningkatkan stabilitas performa dengan mengurangi kelemahan masing-masing kernel. Hal ini sejalan dengan penelitian[13] yang menegaskan bahwa metode ensemble dapat mengurangi variansi dan meningkatkan akurasi dibanding model tunggal.

Secara keseluruhan, hasil penelitian ini memperlihatkan bahwa kernel polynomial derajat 3 memberikan akurasi tertinggi pada model tunggal, sementara metode ensemble menghasilkan performa yang lebih konsisten pada berbagai rasio data. Dengan demikian, dapat disimpulkan bahwa pendekatan ensemble merupakan strategi yang lebih handal dalam penerapan klasifikasi berbasis SVM, khususnya pada data yang memiliki kompleksitas tinggi.

### Pembahasan

Berdasarkan hasil pengujian yang telah dilakukan terhadap keempat skenario penelitian dengan pembagian data menjadi rasio 60:40, 70:30, dan 80:20, dapat diketahui bahwa setiap skenario menunjukkan performa yang bervariasi sesuai dengan karakteristik kernel yang digunakan sesuai pada Gambar 5. Pada skenario pertama, pengujian difokuskan pada performa kernel individual sebelum dilakukan ensemble. Hasil uji coba menunjukkan bahwa kernel polynomial degree 3 menghasilkan akurasi tertinggi, yakni 88,75% pada rasio 60:40. Kinerja yang tinggi ini diduga karena kernel polynomial degree 3 memiliki kemampuan dalam memetakan pola non-linear secara lebih baik, sehingga mampu menangkap kompleksitas data melanoma yang sulit dijangkau oleh kernel linier atau RBF sederhana.

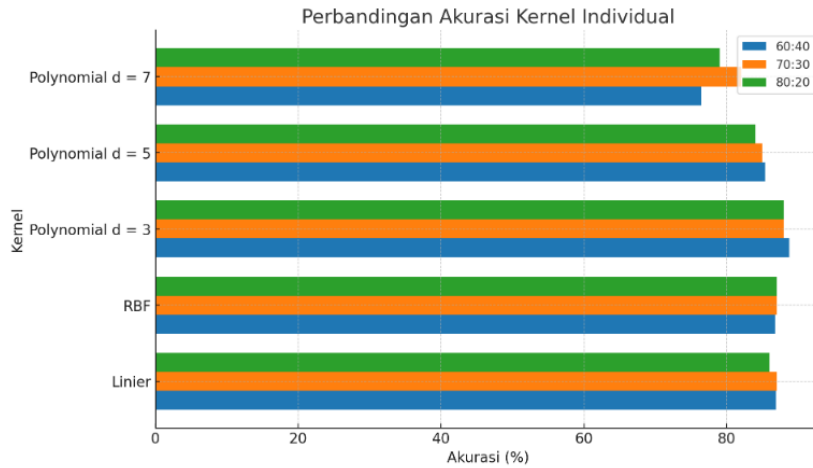
Kernel RBF juga menunjukkan performa yang cukup baik dengan akurasi tertinggi sebesar 87% pada rasio 70:30. Hal ini sejalan dengan keunggulan kernel RBF yang mampu bekerja secara efektif dalam data berdimensi tinggi dan kompleks[32]. Akan tetapi, kernel polynomial dengan degree lebih tinggi, yakni degree 5 dan degree 7, justru mengalami penurunan performa signifikan. Kernel polynomial degree 5 hanya menghasilkan akurasi tertinggi sebesar 85,39%, sedangkan degree 7 bahkan turun hingga 81,89%. Penurunan akurasi ini kemungkinan besar disebabkan oleh meningkatnya kompleksitas model yang mengarah pada masalah overfitting, di mana model terlalu menyesuaikan diri dengan data latih sehingga mengurangi kemampuan generalisasi terhadap data uji[33][34].

Selain itu, penggunaan nilai parameter regularisasi  $C = 1$  turut memengaruhi hasil. Nilai ini menjaga keseimbangan antara margin yang besar dan jumlah kesalahan klasifikasi. Pada kernel polynomial degree 3, pengaturan tersebut terbukti optimal, sedangkan pada polynomial degree lebih tinggi, nilai  $C$  yang sama kurang efektif sehingga berdampak pada performa yang menurun. Dengan demikian, dapat disimpulkan bahwa pemilihan kernel serta parameter yang tepat sangat krusial dalam meningkatkan akurasi klasifikasi melanoma.

Berdasarkan hasil pengujian pada berbagai skenario ensemble, dapat dilihat bahwa pemilihan kernel dan parameter sangat memengaruhi performa klasifikasi melanoma. Pada skenario A, yang menggunakan kernel linier, RBF, dan polynomial degree 3, diperoleh akurasi tertinggi sebesar 88,17% pada rasio data 60:40, meskipun sedikit menurun menjadi 87,42% pada rasio 80:20. Dibandingkan kernel individual, hasil ini lebih baik dibanding kernel linier (86,88%) dan RBF (86,75%), tetapi masih lebih rendah dibanding polynomial degree 3 (88,75%). Temuan ini menunjukkan bahwa ensemble memberikan stabilitas pada precision, recall, dan F1-score, meskipun performa polynomial degree 3 secara individual tetap lebih unggul.

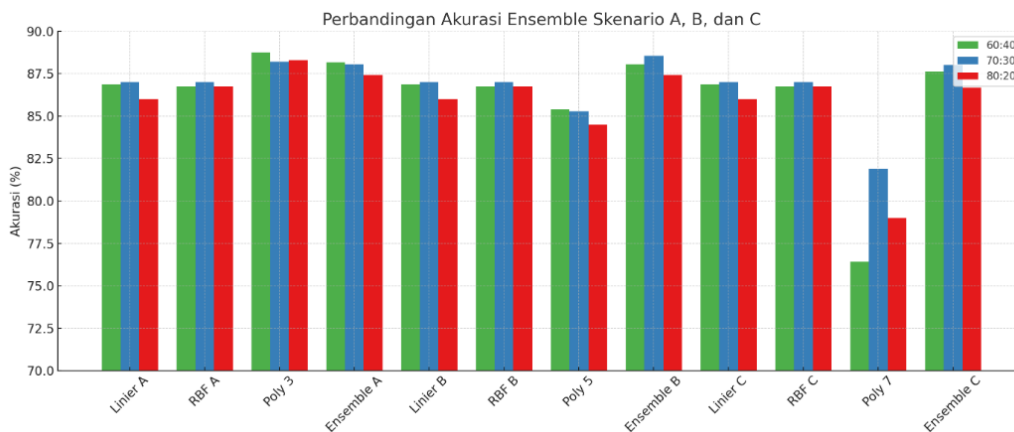
Skenario B dapat dilihat pada Gambar 6. yang menggabungkan kernel linier, RBF, dan polynomial degree 5 menghasilkan akurasi sebesar 88,04%, lebih tinggi dibanding semua kernel individual, termasuk polynomial degree 5 yang hanya mencapai 85,29%. Pada rasio 70:30, skenario ini bahkan mencatatkan akurasi tertinggi, yakni 88,56%, menegaskan bahwa kombinasi kernel mampu

mengoptimalkan performa meskipun salah satu kernel lemah. Namun, pada rasio 80:20, akurasi menurun menjadi 87,42%, meski tetap melampaui kernel individual.



Gambar 5. Perbandingan Kernel

Skenario C dengan kernel linier, RBF, dan polynomial degree 7 menunjukkan akurasi 87,62% pada rasio 60:40, lebih tinggi dari kernel individual. Namun, pada rasio 80:20 akurasinya turun menjadi 86,67%, lebih rendah dibanding kernel RBF (86,75%). Penurunan ini kemungkinan besar dipengaruhi buruknya performa polynomial degree 7.



Gambar 6. Perbandingan Akurasi Esemble Skenario A, B dan C

Secara keseluruhan, hasil ini menegaskan pentingnya pemilihan kernel dan parameter yang tepat. Kernel polynomial degree 3 memberikan kontribusi terbesar karena kemampuannya memetakan data non-linear secara optimal ke ruang berdimensi tinggi[32]. Dengan demikian, ensemble SVM terbukti efektif dalam meningkatkan kestabilan klasifikasi melanoma dibandingkan penggunaan kernel tunggal.

Metode Artificial Neural Network (ANN) mampu menghasilkan akurasi yang cukup tinggi yaitu 97,5%. Akan tetapi, metode ini memiliki kompleksitas komputasi yang cukup tinggi sehingga kurang sesuai jika dibandingkan dengan Support Vector Machine (SVM) yang lebih sederhana dalam implementasi. Berbeda dengan metode k-Nearest Neighbor (KNN), hasil penelitian menunjukkan akurasi sebesar 85,22%, lebih rendah dibandingkan ensemble SVM yang mampu mencapai akurasi rata-rata di atas 87%. Temuan ini juga melampaui penelitian yang dilakukan oleh [35] menggunakan KNN,

di mana hasil akurasi lebih rendah dibandingkan dengan SVM. Hal ini membuktikan bahwa penggunaan kernel polynomial pada SVM lebih efektif dalam menangani pola data non-linear pada citra melanoma, sebagaimana juga dijelaskan oleh [10] bahwa SVM memiliki keunggulan dalam mempelajari batasan non-linear secara optimal.

Penelitian sebelumnya oleh [36] yang menggunakan metode Random Forest dengan rasio data 70:30 menghasilkan akurasi sebesar 81%. Random Forest memang cukup kuat dalam menangani data berdimensi tinggi, namun hasil penelitian ini menunjukkan bahwa ensemble SVM dengan kombinasi kernel linier, RBF, dan polynomial degree 3 mampu menghasilkan akurasi sebesar 88,17% pada rasio data 60:40, yang secara signifikan lebih tinggi dibandingkan Random Forest.

Sementara itu, penelitian yang dilakukan oleh [30] menggunakan dataset melanoma dengan dua kelas, yaitu jinak dan ganas, serta menerapkan metode SVM. Hasil akurasi yang diperoleh adalah 80% untuk kernel linier, 67% untuk kernel RBF, dan 90% untuk kernel polynomial. Jika dibandingkan, kernel polynomial degree 3 pada penelitian ini menghasilkan akurasi 88,75% pada rasio data 60:40, hanya sedikit lebih rendah dari penelitian Wahaningar. Perbedaan ini diduga dipengaruhi oleh variasi preprocessing maupun fitur yang digunakan. Dengan penggunaan ensemble, kelemahan dari penelitian terdahulu yang hanya mengandalkan satu kernel dapat diatasi, sehingga menghasilkan performa klasifikasi yang lebih stabil dan konsisten.

Tabel 3. Perbandingan ANN dan Ensemble SVM

Aspek yang Dibandingkan	Penelitian Anda (GLCM + Ensemble SVM)	Penelitian Sebelumnya
Tujuan	Klasifikasi melanoma berbasis tekstur GLCM dan ensemble SVM	Segmentasi + klasifikasi melanoma dan lesi jinak
Metode Ekstraksi Fitur	GLCM (dissimilarity, correlation, homogeneity, contrast, entropy, energy) Ensemble SVM (linier, RBF, polynomial)	Fitur warna, tekstur, dan geometri setelah segmentasi
Metode Klasifikasi	Polynomial degree 3 terbaik; degree terlalu tinggi menurunkan akurasi	SVM dan ANN
Pengaruh Parameter Kebutuhan Komputasi	Rendah–menengah	Parameter segmentasi & ANN sangat berpengaruh
Implikasi Klinis	Cocok untuk sistem ringan/deteksi dini	Tinggi Cocok untuk sistem klinis presisi tinggi

Berdasarkan Tabel 3. perbandingan penelitian [37] dengan hasil penelitian ini adalah metode, fitur, kinerja, dan kebutuhan komputasi. Penelitian sekarang lebih sederhana, cepat, dan menggunakan GLCM serta Ensemble SVM dengan akurasi 89,45%. Sementara itu, Dalila et al. [37] memakai segmentasi dan fitur lebih kompleks sehingga akurasi lebih tinggi, sekitar 95%. Namun, metode mereka membutuhkan komputasi lebih besar dan proses lebih rumit dibanding pendekatan sekarang.

#### 4. KESIMPULAN

Penelitian ini menunjukkan bahwa kombinasi metode Gray Level Co-Occurrence Matrix (GLCM) dan Ensemble Support Vector Machine (SVM) mampu meningkatkan performa klasifikasi citra kanker kulit melanoma. Penggunaan ensemble dengan majority vote dari kernel linier, RBF, dan polynomial efektif mengatasi keterbatasan masing-masing kernel individu dalam menghadapi pola melanoma yang kompleks. Hasil pengujian membuktikan bahwa ensemble menghasilkan akurasi lebih tinggi, yaitu 89,45%, dibandingkan kernel tunggal sebesar 88,61%. Kernel polynomial degree 3 menjadi

konfigurasi terbaik, sementara degree lebih tinggi justru menurunkan akurasi. Fitur tekstur GLCM seperti dissimilarity, correlation, homogeneity, contrast, entropy, dan energy berkontribusi signifikan dalam merepresentasikan karakteristik melanoma. Penelitian ini berhasil mengembangkan ensemble SVM-GLCM dengan akurasi 89,45%, mengatasi keterbatasan metode tunggal [24], dan berkontribusi pada deteksi dini melanoma di informatika kesehatan.

Kontribusi ini mendukung prioritas riset nasional PRN 2022–2024 di bidang TI kesehatan melalui peningkatan akurasi deteksi dini melanoma berbasis citra. Penelitian lanjutan dapat mengintegrasikan deep learning pada dataset lebih besar untuk meningkatkan generalisasi dan kinerja sistem.

## UCAPAN TERIMAKASIH

Kami mengucapkan terima kasih banyak kepada LP2M Universitas Wahid Hasyim yang telah mendukung kami baik secara finansial sehingga penelitian ini dapat terlaksana dengan baik. Kami juga mengucapkan terima kasih kepada semua pihak yang telah membantu sehingga penelitian ini dapat terwujud.

## REFERENSI

- [1] T. Tarver, “Cancer Facts & Figures 2012. American Cancer Society (ACS): Atlanta, GA: American Cancer Society, 2012. 66 p., pdf. *J Consum Health Internet*, vol. 16, no. 3, pp. 366–367, Jul. 2012, doi: 10.1080/15398285.2012.701177.
- [2] V. Merabishvili, “Review Cancer incidence in Five Continents Vol. XI / Ed. by F. Bray, M. Colombet, L. Mery, M. Piñeros, A. Znaor, R. Zanetti and J. Ferlay. IARC Sci. publ. № 166, Lyon, France, 1545 p.,” *Problems in oncology*, vol. 67, no. 6, p. 855, Dec. 2021, doi: 10.37469/0507-3758-2021-67-6-855.
- [3] N. T. & A. J. Barbara Fallon, Jennifer Ma, Kate Allan, Melanie Pillhofer, “Opportunities for prevention and intervention with young children: lessons from the Canadian incidence study of reported child abuse and neglect,” *Child Adolesc Psychiatry Ment Health*, vol. 7, no. 4, 2013.
- [4] R. L. Siegel, K. D. Miller, N. S. Wagle, and A. Jemal, “Cancer statistics, 2023,” *CA Cancer J Clin*, vol. 73, no. 1, pp. 17–48, Jan. 2023, doi: 10.3322/caac.21763.
- [5] X. Liu, S. Ouellette, M. Jamgochian, Y. Liu, and B. Rao, “One-class machine learning classification of skin tissue based on manually scanned optical coherence tomography imaging,” *Sci Rep*, vol. 13, no. 1, Jan. 2023, doi: 10.1038/s41598-023-28155-5.
- [6] K. Tank, “A Comparative Study on Mushroom Classification using Supervised Machine Learning Algorithms,” *International Journal of Trend in Scientific Research and Development (IJTSRD)*, vol. 5, no. 5, pp. 716–723, 2021.
- [7] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural Features for Image Classification,” *IEEE Trans Syst Man Cybern*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973, doi: 10.1109/tsmc.1973.4309314.
- [8] A. M. Martinez and A. C. Kak, “PCA versus LDA,” *IEEE Trans Pattern Anal Mach Intell*, vol. 23, no. 2, pp. 228–233, 2001, doi: 10.1109/34.908974.
- [9] L. Breiman, “Random Forests,” *Mach Learn*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/a:1010933404324.
- [10] C. Cortes and V. Vapnik, “Support-Vector Networks,” *Mach Learn*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1023/a:1022627411411.
- [11] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, “A survey of skin-color modeling and detection methods,” *Pattern Recognit*, vol. 40, no. 3, pp. 1106–1122, Mar. 2007, doi: 10.1016/j.patcog.2006.06.010.
- [12] S. L. Phung, A. Bouzerdoum, and D. Chai, “Skin segmentation using color pixel classification: analysis and comparison,” *IEEE Trans Pattern Anal Mach Intell*, vol. 27, no. 1, pp. 148–154, Jan. 2005, doi: 10.1109/tpami.2005.17.
- [13] R. Khan, A. Hanbury, J. Stöttinger, and A. Bais, “Color based skin classification,” *Pattern Recognit Lett*, vol. 33, no. 2, pp. 157–163, Jan. 2012, doi: 10.1016/j.patrec.2011.09.032.

- 
- [14] D. Chai and K. N. Ngan, "Face segmentation using skin-color map in videophone applications," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 9, no. 4, pp. 551–564, Jun. 1999, doi: 10.1109/76.767122.
- [15] R. Alief and E. Nurmiati, "Penerapan Kecerdasan Buatan Dan Teknologi Informasi Pada Efisiensi Manajemen Pengetahuan," *Jurnal Masyarakat Informatika*, vol. 13, no. 1, pp. 59–69, 2022, doi: 10.14710/jmasif.13.1.43760.
- [16] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data Via the EM Algorithm," *J R Stat Soc Series B Stat Methodol*, vol. 39, no. 1, pp. 1–22, Sep. 1977, doi: 10.1111/j.2517-6161.1977.tb01600.x.
- [17] C. M. Bishop, "Pattern Recognition and Machine Learning," in *Pattern Recognition and Machine Learning*, Springer New York, 2006. doi: 10.1007/978-0-387-45528-0\_1.
- [18] I. Daubechies, *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992. doi: 10.1137/1.9781611970104.
- [19] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans Pattern Anal Mach Intell*, vol. 11, no. 7, pp. 674–693, Jul. 1989, doi: 10.1109/34.192463.
- [20] B. Mutale, N. C. Withanage, P. K. Mishra, J. Shen, K. Abdelrahman, and M. S. Fnais, "A performance evaluation of random forest, artificial neural network, and support vector machine learning algorithms to predict spatio-temporal land use-land cover dynamics: a case from lusaka and colombo," *Front Environ Sci*, vol. 12, 2024, doi: 10.3389/fenvs.2024.1431645.
- [21] S. P. Nainggolan and A. Sinaga, "COMPARATIVE ANALYSIS OF ACCURACY OF RANDOM FOREST AND GRADIENT BOOSTING CLASSIFIER ALGORITHM FOR DIABETES CLASSIFICATION," *Sebatik*, vol. 27, no. 1, pp. 97–102, Jun. 2023, doi: 10.46984/sebatik.v27i1.2157.
- [22] N. A. Nugroho and A. W. Wijayanto, "Perbandingan Algoritma Machine Learning dalam Pengklasifikasian Tingkat Kemiskinan di Indonesia Tahun 2021 1".
- [23] Z. Darojah, R. Susetyoko, N. Ramadijanti, P. Elektronika Negeri Surabaya, and P. Korespondensi, "STRATEGI PENANGANAN IMBALANCE CLASS PADA MODEL KLASIFIKASI PENERIMA KARTU INDONESIA PINTAR KULIAH BERBASIS NEURAL NETWORK MENGGUNAKAN KOMBINASI SMOTE DAN ENN," vol. 10, no. 2, pp. 457–466, 2023, doi: 10.25126/jtiik.2023106480.
- [24] E. S. Rusdi *et al.*, "PERBANDINGAN ANN, RANDOM FOREST, DAN XGBOOST DALAM KLASIFIKASI ANTIBIOTIK DENGAN PENERAPAN METODE SAMPLING COMPARISON OF ANN, RANDOM FOREST, AND XGBOOST IN ANTIBIOTIC CLASSIFICATION WITH SAMPLING METHODS," vol. 12, no. 4, pp. 2355–7699, 2025.
- [25] C. Nur Azzahra *et al.*, "Sistemasi: Jurnal Sistem Informasi Klasifikasi Cuaca Jawa Barat menggunakan Ensemble Learning pada Data Meteorologi Weather Classification in West Java using Ensemble Learning on Meteorological Data." [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [26] E. Prasetyo and K. Nugroho, "Optimasi Klasifikasi Data Stunting Melalui Ensemble Learning pada Label Multiclass dengan Imbalance Data Optimizing Stunting Data Classification Through Ensemble Learning on Multiclass Labels with Imbalance Data."
- [27] A. Nazir, A. Akhyar, Y. Yusra, and E. Budianita, "Toddler Nutritional Status Classification Using C4.5 and Particle Swarm Optimization," *Scientific Journal of Informatics*, vol. 9, no. 1, pp. 32–41, 2022, doi: 10.15294/sji.v9i1.33158.
- [28] A. Ahdika, M. D. Kartikasari, S. K. Dini, and I. Ramadhani, "Diversification of agricultural areas in Indonesia using dynamic copula modeling and K-Means clustering," *Sains Malays*, vol. 50, no. 9, pp. 2791–2817, 2021, doi: 10.17576/jsm-2021-5009-24.
- [29] R. Gholami and N. Fakhari, "Support Vector Machine: Principles, Parameters, and Applications," in *Handbook of Neural Computation*, Elsevier, 2017, pp. 515–535. doi: 10.1016/b978-0-12-811318-9.00027-2.
- [30] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *Jurnal Edik Informatika*, vol. 2, no. 2, pp. 213–219, 2016.
- [31] J. Chen *et al.*, "Effects of the surface albedo on short-wave infrared detection of atmospheric CO<sub>2</sub>," *Wuli Xuebao/Acta Physica Sinica*, vol. 64, no. 23, pp. 290–301, 2015, doi: 10.7498/aps.64.239201.
-

- 
- [32] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. The MIT Press, 2001. doi: 10.7551/mitpress/4175.001.0001.
- [33] Y. Zhang, H. Du, T. Piao, H. Shi, and S. B. Tsai, “How Manufacturing Companies Can Improve Their Competitiveness: Research on Service Transformation and Product Innovation Based on Computer Vision,” *Journal of Global Information Management*, vol. 32, no. 1, pp. 1–26, 2024, doi: 10.4018/JGIM.336485.
- [34] H. Hu, Z. Zhang, Z. Xie, and S. Lin, “Local relation networks for image recognition,” *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2019-Octob, pp. 3463–3472, 2019, doi: 10.1109/ICCV.2019.00356.
- [35] F. Dalila, A. Zohra, K. Reda, and C. Hocine, “Segmentation and classification of melanoma and benign skin lesions,” *Optik (Stuttg)*, vol. 140, pp. 749–761, Jul. 2017, doi: 10.1016/j.ijleo.2017.04.084.
- [36] N. A. Winanti, D. P. Martiyaningsih, C. A. A. Soemedhy, and U. Athiyah, “Analisis Klasifikasi Citra Kanker Kulit dengan Random Forest,” *remik*, vol. 7, no. 1, pp. 506–515, Jan. 2023, doi: 10.33395/remik.v7i1.12102.
- [37] F. Dalila, A. Zohra, K. Reda, and C. Hocine, “Segmentation and classification of melanoma and benign skin lesions,” *Optik (Stuttg)*, vol. 140, pp. 749–761, Jul. 2017, doi: 10.1016/j.ijleo.2017.04.084.