

Tugas Mata Kuliah

Rekayasa Fitur dan Pengenalan Pola (RFPP)

Klasifikasi Melanoma dengan Pendekatan *Hand-crafted Features*
dan Pemelajaran Mesin

Khalilullah Al Faath

NIU: 566643

Magister Ilmu Komputer



Universitas Gadjah Mada
Semester Gasal 2025/2026

Daftar Isi

1 Pendahuluan	1
2 Metode	2
2.1 Akuisisi	3
2.2 <i>Pre-processing</i>	4
2.3 Segmentasi	6
2.3.1 <i>Thresholding</i> dengan Otsu	6
2.3.2 Cek Intensitas Pusat dan Inversi Masker	7
2.3.3 Memilih Objek Terbesar di Pusat	7
2.3.4 Simpan Mask Akhir	7
2.4 Ekstraksi Fitur	7
2.5 Seleksi Fitur	9
2.5.1 Principal Component Analysis (PCA)	9
2.5.2 Chi-Square	10
2.6 Klasifikasi	10
2.6.1 Support Vector Machine (SVM)	11
2.6.2 Random Forest	11
2.6.3 Evaluasi Kinerja Klasifikasi	12
3 Hasil dan Pembahasan	12
3.1 Dataset yang digunakan	12
3.2 Lingkungan pengujian	13
3.3 Hasil	13
3.3.1 Statistik Dataset	13
3.3.2 Hasil Ekstraksi Fitur	13
3.3.3 Seleksi Fitur dan Hasil Klasifikasi	15
3.3.4 Hasil Eksperimen Lanjutan	18
3.4 Diskusi	21
3.4.1 Pengaruh Seleksi Fitur	21
3.4.2 Performa Model dan Perbandingan SVM vs Random Forest . . .	21
3.4.3 Dampak Ketidakseimbangan Kelas	21
3.4.4 Keterbatasan dan Peluang Pengembangan	22
4 Kesimpulan	22

Daftar Gambar

1	Karakteristik Melanoma	1
2	Diagram alir metode yang dilakukan	3
3	Contoh citra yang ada pada Dataset ISIC 2016 Task 3.	3
4	Contoh citra <i>benign</i> dan <i>malignant</i>	4
5	Diagram alir pra-pemrosesan.	5
6	(1) Konversi ke Grayscale; (2) Menambahkan <i>salt and pepper noise</i> ; (3) Menerapkan <i>Iterative median filtering</i>	6
7	Diagram alir segmentasi.	6
8	Diagram alir Otsu.	7
9	Contoh hasil segmentasi citra dermoskopik: (a) berhasil, (b) gagal.	7
10	Pembagian Ekstraksi fitur citra dermoskopik berdasarkan mask yang di- gunakan	8
11	Distribusi tiap kelas	13
12	Matriks korelasi tiap kolom	14
13	Boxplot beberapa fitur penting	14
14	Distribusi tiap fitur berdasarkan kelas	15
15	Hasil <i>confusion matrix</i> kedua seleksi fitur.	18

Daftar Tabel

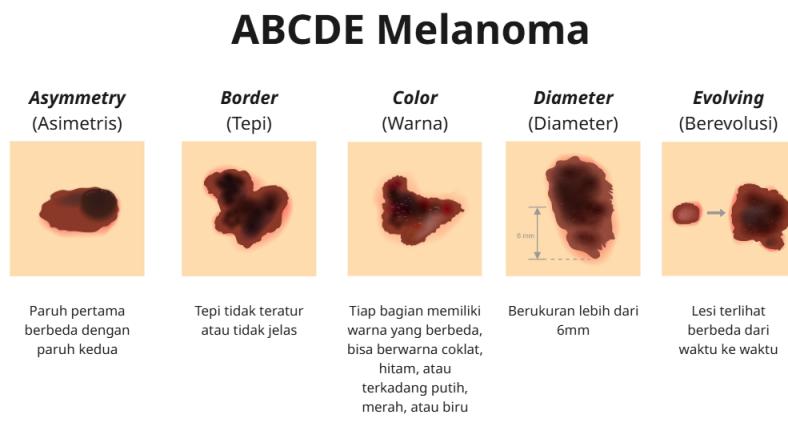
1	Ringkasan fitur ekstraksi dari dataset ISIC 2016.	12
2	Perbandingan Metrik Evaluasi Chi-square vs PCA pada Kelas Malignant	17
3	Perbandingan hasil seluruh skenario eksperimen yang dilakukan.	20

1 Pendahuluan

Melanoma maligna merupakan bentuk neoplasma kulit yang paling mematikan yang merupakan hasil transformasi melanosit, yaitu sel yang berasal dari *neural crest*. Karena asal-usul embriologisnya, melanoma tidak hanya terbatas pada kulit, tetapi juga dapat berkembang di lokasi lain tempat sel *neural crest* bermigrasi, seperti traktus gastrointestinal, otak, dan mata (melanoma okular). Secara patofisiologis, perkembangan melanoma sangat dipengaruhi oleh faktor lingkungan (terutama paparan radiasi UV), genetik (seperti mutasi pada gen BRAF, CDKN2A, CDK4), dan imunologis [1].

Urgensi pengembangan sistem deteksi dan klasifikasi melanoma didasari oleh tren peningkatan insiden dan risiko fatalitas yang yang tinggi. Berdasarkan data *National Cancer Institute (NCI) Surveillance, Epidemiology, and End Results (SEER)*, (1) melanoma kini merupakan kanker ganas yang paling umum yang terjadi di pria dan wanita, (2) pada tahun 2023, diperkirakan terdapat 97.000 kasus baru di Amerika Serikat dengan estimasi 7.990 kematian [2], (3) Tingkat kelangsungan hidup relatif 5 tahun (5-year relative survival rate) sangat bergantung pada stadium saat diagnosis: pasien dengan melanoma stadium 0 (in-situ) memiliki tingkat kelangsungan hidup mencapai 97% hingga 100%. Sebaliknya, angka ini menurun drastis menjadi sekitar 30% hingga 52% bagi pasien yang terdiagnosis pada stadium IV (metastatik) [1], (4) di Indonesia, pada tahun 2020, angka kasus kanker kulit mencapai 18.000 kasus dengan angka kematian sekitar 3.000 [3].

Deteksi dini melanoma dapat dilakukan dengan menggunakan metode jembatan keledai "ABCDE", yaitu "A" untuk *Asymmetry* atau asimetris, "B" untuk *Border* atau tepi, "C" untuk *color* atau warna, "D" untuk *Diameter*, dan "E" untuk *Evolving* atau berevolusi [1]. Gambar 1 menggambarkan karakteristik khusus dari melanoma.



Gambar 1: Karakteristik Melanoma

Pengembangan sistem *Computer-Aided Diagnosis (CAD)* untuk deteksi melanoma telah menjadi fokus penelitian intensif dalam beberapa dekade terakhir. Secara garis besar, penelitian-penelitian sebelumnya dapat dikelompokkan menjadi dua pendekatan

utama: metode berbasis ekstraksi fitur manual atau (*hand-crafted features*) dan metode berbasis pemelajaran dalam atau *Deep Learning*.

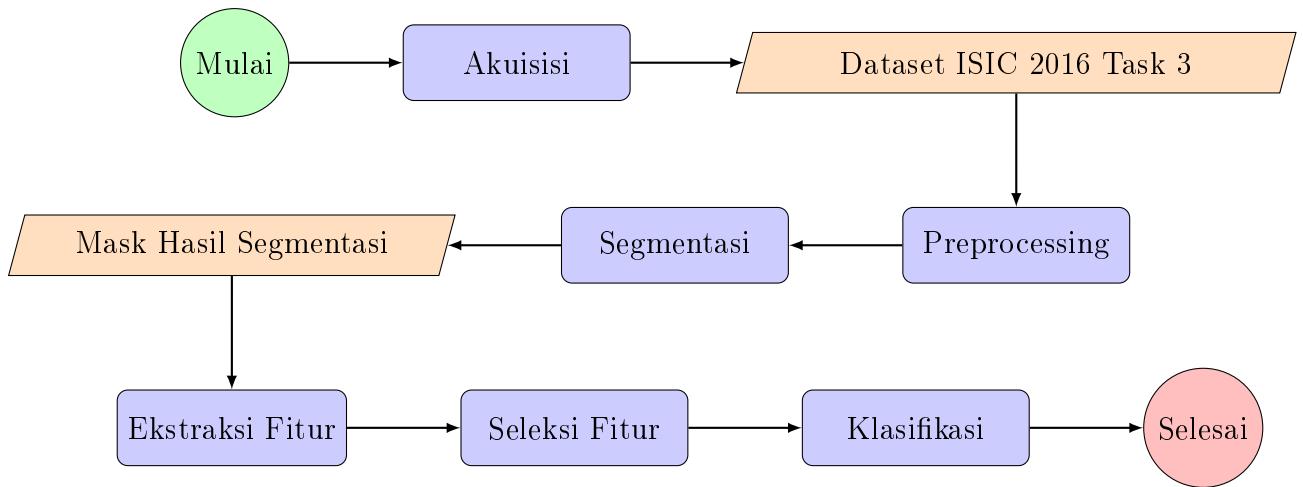
Pendekatan ekstraksi manual Pendekatan pertama, yaitu metode ekstraksi fitur manual, berfokus pada pemodelan karakteristik visual melanoma berdasarkan pengetahuan klinis dan aturan morfologi. Pada pendekatan ini, proses analisis citra dilakukan melalui tahapan pra-pemrosesan, segmentasi lesi, dan ekstraksi fitur berbasis ciri geometris, tekstur, atau warna. Contohnya termasuk penggunaan fitur statistik orde pertama, *Histogram of Oriented Gradients (HOG)*, *Local Binary Patterns (LBP)*, momen bentuk, hingga parameter geometrik seperti asimetri, *compactness*, dan *eccentricity*. Pendekatan ini memiliki keunggulan berupa interpretabilitas yang tinggi karena parameter yang dianalisis selaras dengan kriteria klinis seperti aturan ABCDE. Namun, performanya sangat bergantung pada kualitas segmentasi dan sensitivitas model terhadap variasi perangkat fotografi, pencahayaan, serta keberadaan artefak seperti rambut, bayangan, dan variasi kulit [4], [5], [6].

Pendekatan pemelajaran dalam Sebaliknya, pendekatan berbasis pemelajaran dalam (*Deep Learning*) memanfaatkan jaringan saraf tiruan berlapis banyak, khususnya *Convolutional Neural Network (CNN)*, untuk mengekstraksi representasi fitur secara otomatis tanpa memerlukan desain fitur manual oleh pakar. Model seperti ResNet, EfficientNet, dan Vision Transformer (ViT) telah menunjukkan performa unggul dalam kompetisi internasional seperti International Skin Imaging Collaboration (ISIC). Keunggulan utama pendekatan ini adalah kemampuannya untuk belajar dari data berskala besar, mengurangi ketergantungan pada pra-pemrosesan, serta meningkatkan robustnes terhadap variasi pola, warna, dan tekstur lesi kulit. Namun, tantangannya mencakup kebutuhan dataset besar yang teranotasi dengan baik, risiko bias algoritmik akibat ketidakseimbangan distribusi data (misalnya warna kulit pasien), serta keterbatasan interpretabilitas (*black-box problem*) [7], [8], [9].

Klasifikasi melanoma yang dilakukan pada tugas ini dilakukan dengan menggunakan ekstraksi fitur manual dengan model pemelajaran mesin *Support Vector Machine (SVM)*.

2 Metode

Bagian ini menjelaskan alur langkah yang dilakukan pada tugas ini. Gambar 2 menunjukkan alur langkah metode yang diusulkan.



Gambar 2: Diagram alir metode yang dilakukan

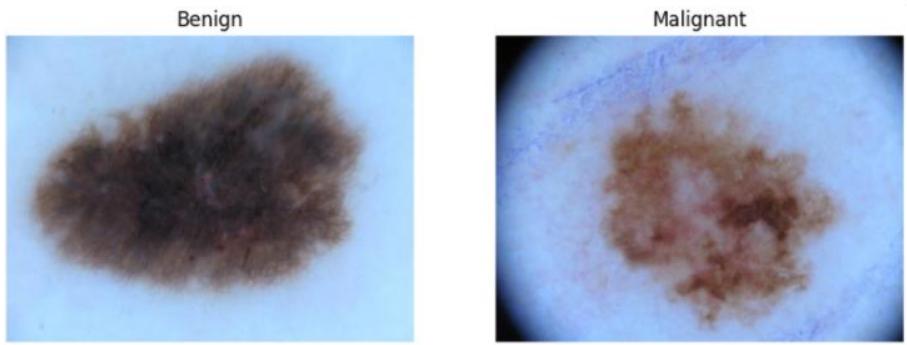
2.1 Akuisisi

Dataset yang digunakan pada tugas ini adalah Dataset ISIC 2016 Task 3.¹ Dataset ini merupakan himpunan citra dermoskopik yang digunakan untuk menyelesaikan permasalahan klasifikasi lesi kulit menjadi dua kategori, yaitu *benign* (non-malignant) dan *malignant*. Gambar 3 menunjukkan contoh data yang ada pada Dataset, sementara Gambar 4 menunjukkan contoh citra berlabel *benign* dan *malignant*.



Gambar 3: Contoh citra yang ada pada Dataset ISIC 2016 Task 3.

¹Dataset ini dapat diunduh di pranala berikut <https://challenge.isic-archive.com/landing/2016/39/>



Gambar 4: Contoh citra *benign* dan *malignant*.

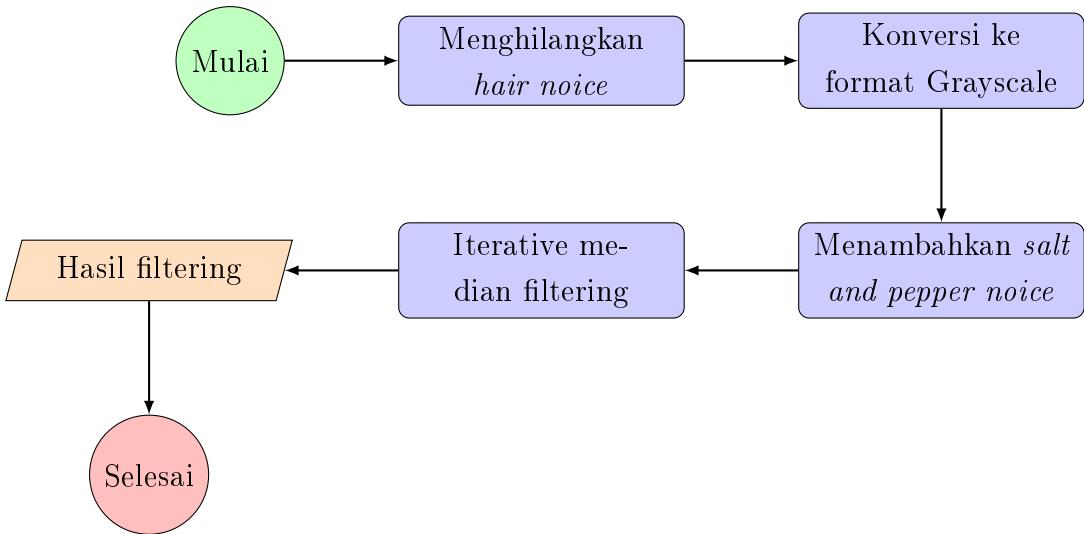
Dataset pelatihan terdiri atas 900 citra lesi kulit dalam format JPEG dengan resolusi bervariasi. Setiap citra dinamai menggunakan pola `ISIC_<image_id>.jpg` dengan tujuh digit pengenal unik. Berkas pelatihan dilengkapi dengan *Training Ground Truth* berupa sebuah berkas CSV yang memuat dua kolom: (1) `image_id` yang sesuai dengan nama berkas citra, dan (2) label diagnosis definitif (*benign* atau *malignant*) yang ditentukan berdasarkan konsensus pakar serta laporan patologi.

Selain itu, disediakan pula seperangkat *Test Data* yang terdiri dari 379 citra dengan format yang sama. Berbeda dengan data pelatihan, data uji tidak disertai label kebenaran untuk menjaga objektivitas evaluasi. Pada proses kompetisi aslinya, peserta diminta mengunggah hasil prediksi dalam format CSV dua kolom berisi `image_id` dan skor probabilistik dalam rentang [0.0, 1.0], dengan nilai di atas 0.5 mengindikasikan prediksi *malignant* [10].

Dapat dilihat pada Gambar 3 terdapat beberapa jenis *noise* yang dapat diidentifikasi, yaitu *border noise*, *hair noise*, dan *object noise*.

2.2 *Pre-processing*

Setelah mendapatkan Dataset, dilakukan *pre-processing* atau pra-pemrosesan untuk menghilangkan sebagian *noise*. Gambar 5 menunjukkan langkah-langkah pra-pemrosesan yang dilakukan.

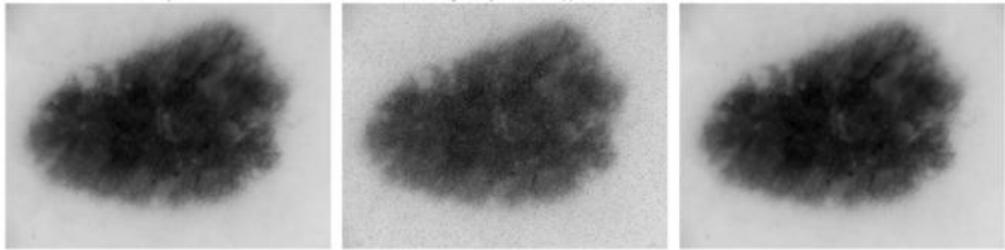


Gambar 5: Diagram alir pra-pemrosesan.

Setiap tahap pra-pemrosesan memiliki tujuan spesifik untuk mempersiapkan citra sebelum dilakukan segmentasi dengan mengikuti alur dari penelitian Jaseema Yasmin dan Mohamed Sadiq [11].

1. **Menghilangkan *hair noise*:** Pada citra dermoskopik, rambut yang menutupi lesi dapat mengganggu analisis dan segmentasi. Oleh karena itu, rambut dihilangkan menggunakan operasi morfologi *black-hat* yang diikuti dengan *inpainting* untuk menutupi area rambut dengan piksel sekitarnya, sehingga citra menjadi lebih bersih.
2. **Konversi ke format Grayscale:** Citra RGB dikonversi menjadi citra intensitas (grayscale) untuk menyederhanakan data dan mengurangi kompleksitas komputasi. Dengan ini, analisis selanjutnya hanya fokus pada intensitas piksel, sehingga pengaruh variasi warna yang tidak relevan dapat diminimalkan.
3. **Menambahkan *salt and pepper noise*:** Tahap ini dilakukan untuk menyiapkan citra dalam kondisi yang memungkinkan pengujian ketahanan metode filter. Noise buatan ditambahkan pada citra grayscale agar metode *iterative median filtering* dapat diuji efektivitasnya dalam menghilangkan noise dan mempertahankan tepi lesi.
4. **Iterative Median Filtering:** Citra yang telah diberi noise kemudian difilter menggunakan metode median secara iteratif. Filter median efektif untuk menghilangkan noise impulsif (*salt and pepper*) tanpa merusak kontur objek penting. Proses iteratif memastikan noise tersisa dapat dihilangkan secara bertahap, sambil mempertahankan detail lesi.
5. **Hasil Filtering:** Tahap ini menghasilkan citra yang bersih dari noise, dengan kontur lesi tetap terjaga. Citra hasil ini siap digunakan untuk tahap segmentasi dan ekstraksi fitur selanjutnya.

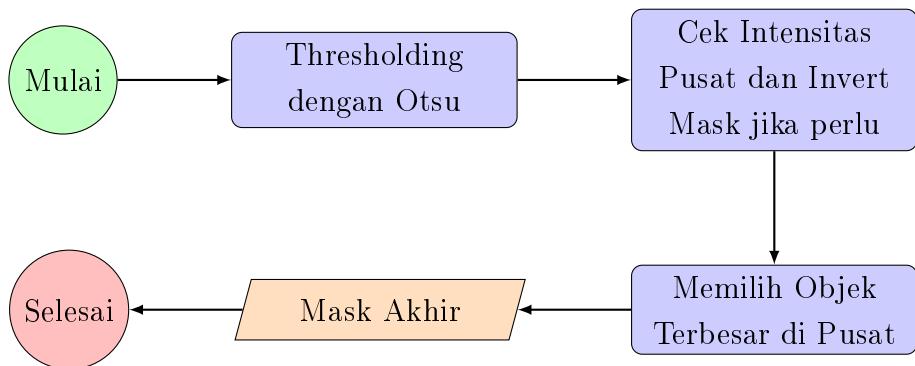
Gambar 6 memperlihatkan contoh hasil tiap tahap filtering pada citra dermoskopik.



Gambar 6: (1) Konversi ke Grayscale; (2) Menambahkan *salt and pepper noise*; (3) Menerapkan *Iterative median filtering*

2.3 Segmentasi

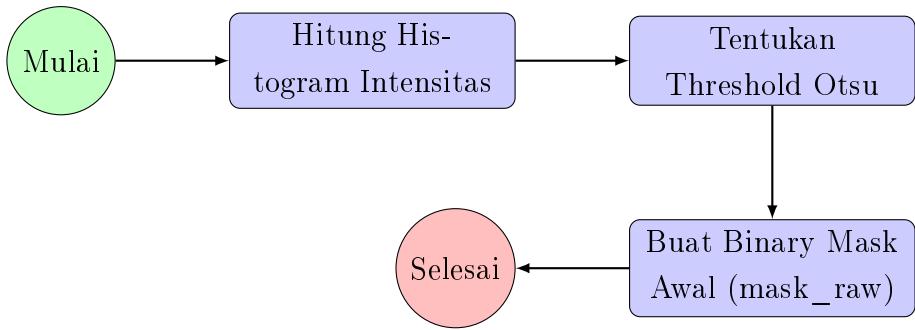
Setelah pra-pemrosesan, didapatkan citra hasil filtering yang siap untuk disegmentasi. Gambar 7 menunjukkan tahapan segmentasi yang dilakukan.



Gambar 7: Diagram alir segmentasi.

2.3.1 *Thresholding* dengan Otsu

Citra hasil pra-pemrosesan difilter menggunakan metode threshold otomatis Otsu (*Otsu's method*) untuk menghasilkan *binary mask* awal (*mask_raw*). *Threshold* Otsu secara adaptif menentukan nilai ambang yang memisahkan piksel gelap (lesi) dan terang (kulit) berdasarkan histogram intensitas, sehingga tidak perlu menentukan *threshold* secara manual.



Gambar 8: Diagram alir Otsu.

2.3.2 Cek Intensitas Pusat dan Inversi Masker

Untuk memastikan lesi berada di area foreground (nilai putih pada mask), intensitas piksel di pusat citra dicek. Jika rata-rata intensitas pusat kurang dari 127, maka *mask* dibalik (`cv2.bitwise_not`) sehingga objek lesi menjadi foreground. Langkah ini mencegah kesalahan orientasi mask yang bisa memengaruhi tahap berikutnya.

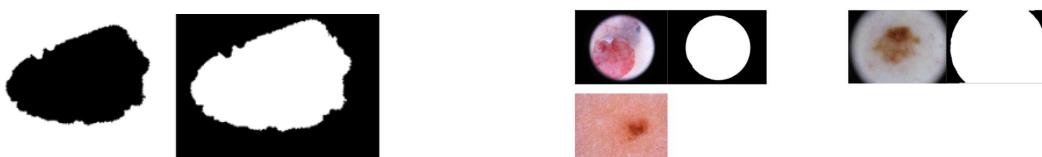
2.3.3 Memilih Objek Terbesar di Pusat

Setelah inversi, semua objek pada mask diperiksa. Fungsi `keep_center_object_only` digunakan untuk memilih objek terbesar yang berada di sekitar pusat citra, sementara objek kecil atau noise di pinggiran diabaikan. Hal ini memastikan hanya lesi utama yang akan dianalisis.

2.3.4 Simpan Mask Akhir

Mask akhir (`final_mask`) yang sudah dibersihkan dan hanya menyisakan lesi utama kemudian disimpan sebagai file gambar. Selain itu, luas area objek utama dicatat (`contour_area`) sebagai informasi tambahan untuk analisis selanjutnya.

Gambar 9 menunjukkan hasil dari tahap segmentasi yang dilakukan.



(a) Segmentasi berhasil.

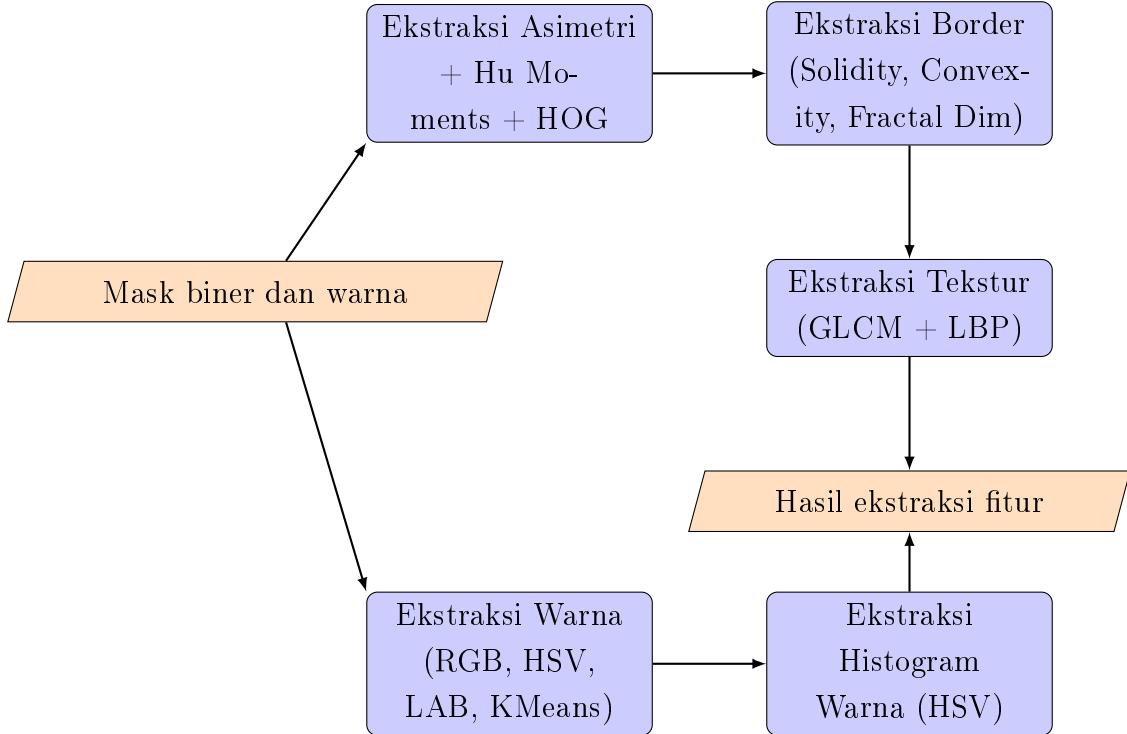
(b) Segmentasi gagal.

Gambar 9: Contoh hasil segmentasi citra dermoskopik: (a) berhasil, (b) gagal.

2.4 Ekstraksi Fitur

Setelah citra telah melalui tahap segmentasi, langkah selanjutnya adalah mengekstraksi fitur yang relevan untuk klasifikasi melanoma. Fitur yang diambil mencakup beberapa

kategori: Asimetri, Border, Warna, Diameter, Tekstur, dan Histogram Warna. Setiap tahap dilakukan menggunakan mask dari hasil segmentasi untuk memastikan fitur hanya diambil dari area lesi. Gambar 10 menunjukkan tahapan ekstraksi fitur apa saja yang dilakukan.



Gambar 10: Pembagian Ekstraksi fitur citra dermoskopik berdasarkan mask yang digunakan

1. **Asimetri (Asymmetry)** Citra lesi dipotong berdasarkan bounding box mask dan diubah ukurannya menjadi 128×128 piksel. Kemudian dilakukan perbandingan dengan versi citra yang dibalik horizontal dan vertikal, lalu dihitung selisih absolut rata-rata untuk menilai tingkat asimetri. Selain itu, *Hu Moments* dan *HOG (Histogram of Oriented Gradients)* juga dihitung untuk merepresentasikan bentuk dan tekstur lokal lesi.
2. **Border (Batas Lesi)** Menggunakan kontur dari mask, dihitung *convex hull*, luas dan keliling asli serta keliling hull. Dari sini dihitung fitur *solidity* dan *convexity*. Fractal dimension juga dihitung dari kontur tepi untuk menangkap kompleksitas bentuk lesi.
3. **Warna (Color)** Fitur warna diekstraksi dari piksel di dalam mask dalam ruang warna RGB, HSV, dan LAB. Statistik dasar seperti standar deviasi tiap channel dihitung. Selain itu, K-Means clustering dengan 3 cluster digunakan untuk mengukur sebaran warna (color spread) sebagai representasi heterogenitas warna lesi.
4. **Diameter dan Area** Luas, keliling, dan diameter maksimum lesi dihitung dari kontur untuk mendukung karakteristik ukuran dan bentuk.

5. Tekstur (GLCM + LBP) Citra dikonversi ke grayscale.

- **GLCM (Gray Level Co-occurrence Matrix)**: Fitur kontras, dissimilarity, homogeneity, energy, korelasi, dan ASM dihitung dari area mask untuk menangkap pola tekstur lokal.
- **LBP (Local Binary Pattern)**: Histogram LBP dihitung dengan metode “uniform” untuk invarian terhadap rotasi, ditambah fitur statistik seperti energi dan entropi dari distribusi LBP.

6. Histogram Warna

Histogram warna HSV dihitung dengan 8 bins per channel untuk menangkap distribusi warna pada lesi. Nilai histogram dinormalisasi agar independen terhadap ukuran lesi.

Implementasi dan Penyimpanan Semua fitur digabungkan dalam satu *DataFrame* dan disimpan ke file CSV (`features_ABCD_final.csv`) untuk tahap selanjutnya. Fitur-fitur ini siap digunakan untuk klasifikasi melanoma, baik menggunakan model *machine learning* tradisional maupun metode *deep learning* berbasis tabular input.

2.5 Seleksi Fitur

Seleksi fitur bertujuan untuk memilih subset fitur yang paling relevan dengan target klasifikasi, sekaligus mengurangi dimensi dataset agar model lebih efisien dan mengurangi risiko *overfitting*. Pada penelitian ini, digunakan dua pendekatan seleksi fitur: **Principal Component Analysis (PCA)** dan **Chi-Square**.

2.5.1 Principal Component Analysis (PCA)

PCA adalah metode reduksi dimensi yang mentransformasikan fitur asli ke dalam himpunan baru *principal components* yang saling ortogonal. Komponen-komponen ini disusun berdasarkan varian data, sehingga komponen pertama menyimpan varian terbesar, komponen kedua menyimpan varian terbesar kedua, dan seterusnya.

Langkah-langkah PCA secara umum adalah:

1. Standarisasi fitur agar memiliki rata-rata nol dan variansi satu.
2. Hitung matriks kovarians dari fitur standar.
3. Lakukan dekomposisi nilai eigen (*eigen decomposition*) untuk mendapatkan nilai eigen dan vektor eigen.
4. Pilih sejumlah komponen utama yang menangkap persentase varian tertentu (misal 95%) sebagai fitur akhir.

Keuntungan PCA:

- Mengurangi redundansi antar fitur yang saling berkorelasi.
- Mempercepat proses pelatihan model karena jumlah fitur berkurang.
- Mempertahankan informasi penting meskipun dimensi dikurangi.

2.5.2 Chi-Square

Uji chi-square digunakan untuk mengukur hubungan antara setiap fitur dengan kelas target. Konsep dasarnya adalah membandingkan frekuensi pengamatan dengan frekuensi yang diharapkan jika tidak ada hubungan antara fitur dan kelas.

Langkah-langkah seleksi fitur dengan chi-square:

1. Untuk setiap fitur, buat tabel kontingensi yang menunjukkan distribusi nilai fitur terhadap kelas target.
2. Hitung nilai chi-square untuk masing-masing fitur menggunakan rumus:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

di mana O_i adalah frekuensi yang diamati dan E_i adalah frekuensi yang diharapkan.

3. Fitur dengan nilai chi-square tinggi dianggap memiliki hubungan signifikan dengan kelas target dan dipilih untuk model.

Keuntungan chi-square:

- Memungkinkan pemilihan fitur yang paling berpengaruh terhadap klasifikasi.
- Sederhana dan cepat untuk dataset dengan banyak fitur.
- Efektif untuk fitur diskrit atau numerik yang telah dikategorikan.

Dengan kombinasi ini, diharapkan fitur akhir yang digunakan memiliki informasi maksimal, bebas dari redundansi, dan siap digunakan untuk proses klasifikasi dengan efisiensi tinggi.

2.6 Klasifikasi

Setelah fitur diekstraksi dan dilakukan seleksi fitur, tahap selanjutnya adalah klasifikasi untuk membedakan citra lesi menjadi kelas *benign* dan *malignant*. Pada penelitian ini, digunakan dua metode klasifikasi: **Support Vector Machine (SVM)** dan **Random Forest**.

2.6.1 Support Vector Machine (SVM)

SVM adalah algoritma pembelajaran *supervised* yang bertujuan menemukan *hyperplane* optimal yang memisahkan kelas-kelas pada ruang fitur. Hyperplane ini memaksimalkan margin antara kelas, sehingga model lebih tahan terhadap kesalahan klasifikasi pada data baru.

Beberapa detail implementasi SVM pada penelitian ini:

- **Kernel:** RBF (*Radial Basis Function*) digunakan untuk menangani data yang tidak linier.
- **Parameter C:** Mengatur toleransi kesalahan pada data pelatihan.
- **Parameter Gamma:** Mengatur pengaruh satu data terhadap bentuk hyperplane.

Keunggulan SVM:

- Efektif untuk dataset dengan dimensi tinggi.
- Memiliki mekanisme margin yang meminimalkan risiko *overfitting*.
- Cocok untuk dataset medis yang biasanya memiliki jumlah sampel terbatas.

2.6.2 Random Forest

Random Forest adalah metode *ensemble learning* berbasis pohon keputusan, yang membangun banyak pohon secara acak dan mengambil keputusan mayoritas (*majority voting*) untuk klasifikasi.

Detail implementasi Random Forest pada penelitian ini:

- **Jumlah pohon:** Misal 100 pohon untuk mencapai kestabilan prediksi.
- **Fitur per split:** Setiap pohon hanya menggunakan subset fitur acak untuk memecah node.
- **Mekanisme voting:** Kelas akhir ditentukan berdasarkan mayoritas hasil dari semua pohon.

Keunggulan Random Forest:

- Tahan terhadap *overfitting* karena *bagging* dan penggunaan banyak pohon.
- Mampu menangani fitur numerik dan kategorikal.
- Memberikan informasi penting tentang kontribusi masing-masing fitur (*feature importance*).

2.6.3 Evaluasi Kinerja Klasifikasi

Kedua model dievaluasi menggunakan metrik:

- **Akurasi:** Persentase prediksi benar dari total prediksi.
- **Precision, Recall, F1-Score:** Mengukur performa model terhadap kelas *malignant* dan *benign*.

3 Hasil dan Pembahasan

Pada tahap ini dijelaskan hasil dari eksperimen yang dilakukan.

3.1 Dataset yang digunakan

Sebagaimana yang disebutkan pada Bagian 2.1, Dataset yang digunakan adalah Dataset ISIC 2016 Task 3 yang memang ditujukan untuk klasifikasi melanoma.

Setelah dilakukan ekstraksi fitur, didapatkan dataset tabular dalam format file CSV (`features_ABCD_final.csv`) yang terdiri dari 900 baris dan 67 kolom sebagaimana yang ditunjukkan pada Tabel 1.

Kategori Fitur	Nama Kolom
Identitas	image_id, label
Asimetri	asym_total, hu_0, hu_1, hu_2, hu_3, hu_4, hu_5, hu_6, hog_mean, hog_var
Batas / Border	solidity, convexity, fractal_dim
Warna	std_R, std_G, std_B, std_H, std_S, std_V, color_spread
Ukuran / Diameter	area, perimeter, max_diameter
Tekstur (GLCM)	glcm_contrast, glcm_dissimilarity, glcm_homogeneity, glcm_energy, glcm_correlation, glcm_ASM
Tekstur (LBP)	lbp_bin_0, lbp_bin_1, lbp_bin_2, lbp_bin_3, lbp_bin_4, lbp_bin_5, lbp_bin_6, lbp_bin_7, lbp_bin_8, lbp_bin_9, lbp_energy, lbp_entropy
Histogram Warna (HSV)	hist_H_bin_0, hist_H_bin_1, hist_H_bin_2, hist_H_bin_3, hist_H_bin_4, hist_H_bin_5, hist_H_bin_6, hist_H_bin_7, hist_S_bin_0, hist_S_bin_1, ..., hist_V_bin_7

Tabel 1: Ringkasan fitur ekstraksi dari dataset ISIC 2016.

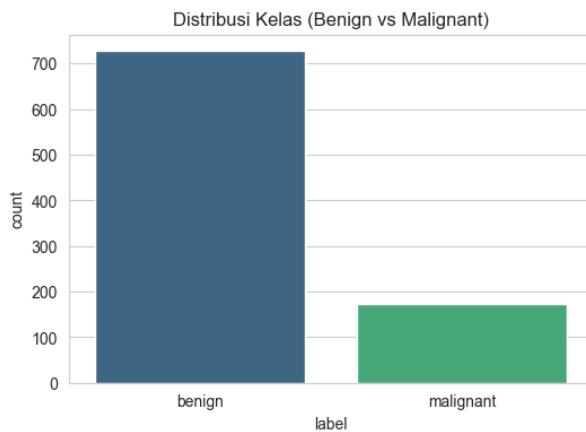
3.2 Lingkungan pengujian

Pengujian manual dilakukan menggunakan Jupyter Notebook dengan bahasa pemrograman Python 3.12.2 dengan spesifikasi *device* laptop HP Pavilion Gaming 5 adalah CPU AMD Ryzen 5 5600H, GPU RTX 3050 dan RAM 16 GB yang dijalankan di Windows 11 64 bit.

3.3 Hasil

3.3.1 Statistik Dataset

Setelah mengimpor data csv, dilakukan beberapa eksplorasi visual statistik dataset.

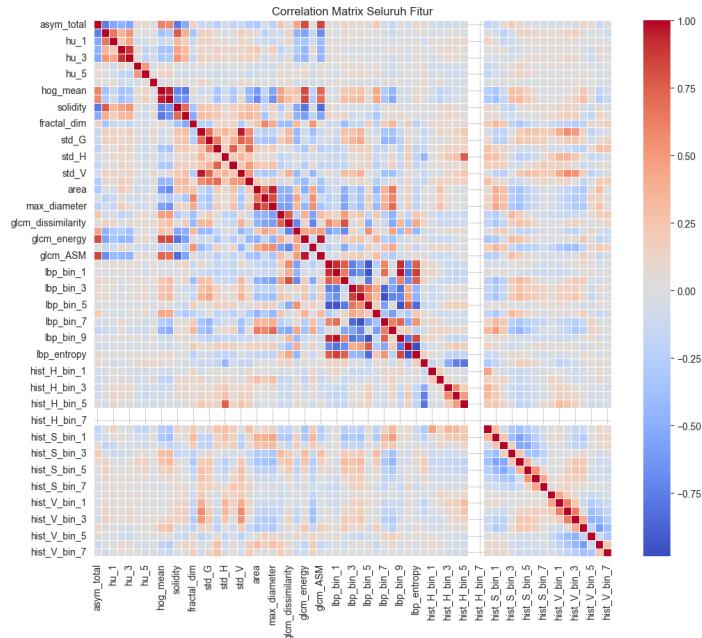


Gambar 11: Distribusi tiap kelas

Gambar 11 menampilkan distribusi tiap kelas dengan kelas *benign* memiliki 727 instans, sementara kelas *malignant* memiliki 173 instans. Kondisi ini merupakan kondisi *data imbalance* yang perlu ditangani lebih lanjut.

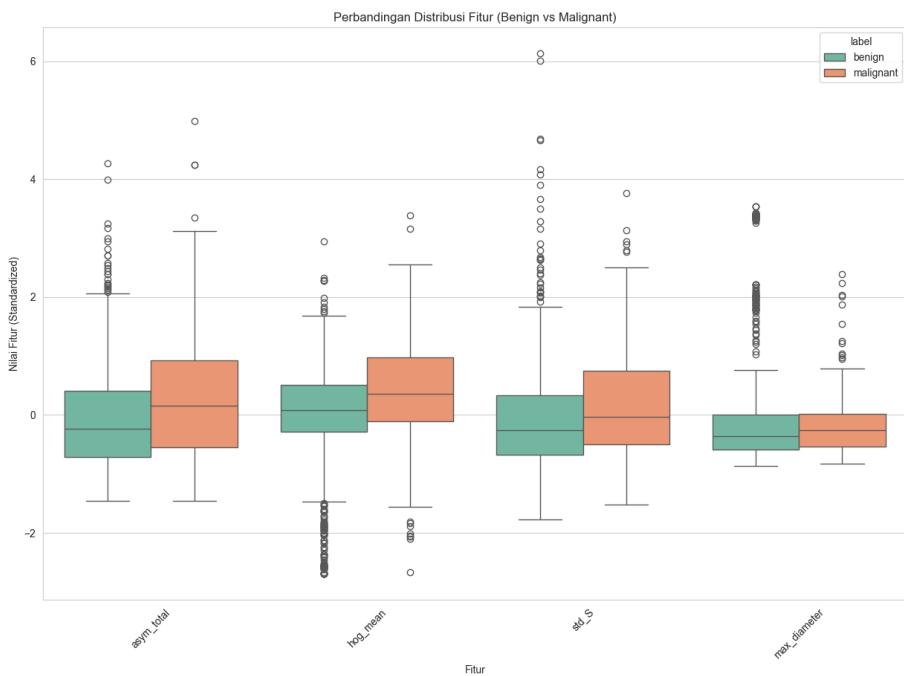
3.3.2 Hasil Ekstraksi Fitur

Setelah itu, dilakukan beberapa visualisasi untuk melihat hasil dari ekstraksi fitur.



Gambar 12: Matriks korelasi tiap kolom

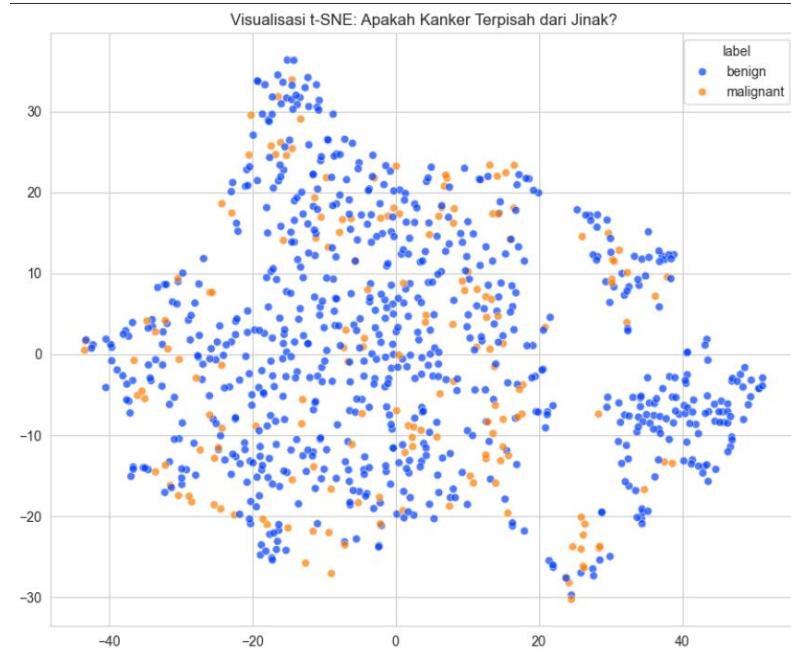
Gambar 12 menampilkan matriks korelasi tiap data numerik. Dapat dilihat bahwa kolom-kolom dengan kategori yang sama, sebagaimana pada Tabel 1, memiliki korelasi yang cukup tinggi yang ditunjukkan oleh terbentuknya kotak persegi di sepanjang diagonal.



Gambar 13: Boxplot beberapa fitur penting

Gambar 13 menampilkan boxplot beberapa fitur penting, seperti *asym_otal*, *hog_mean*, *std_S*, dan *max_diameter*. Dapat dilihat bahwa keempatnya merupakan fitur-fitur yang

lemah karena distribusinya mirip setelah distandarisasi.



Gambar 14: Distribusi tiap fitur berdasarkan kelas

Gambar 14 menampilkan proyeksi data berdimensi tinggi ke dalam 2 dimensi untuk melihat keterpisahan data antara kelas kanker (*malignant-oranye*) dan jinak (*benign-biru*). Dari hasil visualisasi dapat dilihat bahwa terjadi tumpang tindih yang signifikan (*significant overlap*) yang mana titik oranye tidak berkumpul dalam satu klaster yang terpisah. Titik-titik ini tersebar dan tercampur baur dengan titik biru. Hal ini juga mengisyaratkan bahwa tidak ada garis pemisah yang jelas (linear separability) yang memisahkan kedua kelas ini.

Selain itu, secara visual, titik biru jauh lebih banyak daripada titik oranye. Hal ini mengonfirmasi adanya *data imbalance*, sebagaimana yang telah disebutkan di Bagian 3.3.1.

Berdasarkan hal-hal di atas, sangat penting untuk melakukan seleksi fitur.

3.3.3 Seleksi Fitur dan Hasil Klasifikasi

Pada penelitian ini dilakukan dua pendekatan seleksi fitur, yaitu metode berbasis statistik (Chi-square) dan reduksi dimensi linier (Principal Component Analysis/PCA). Kedua hasil seleksi fitur tersebut kemudian diuji menggunakan model klasifikasi Support Vector Machine (SVM). Bagian ini menyajikan fitur-fitur terpilih, jumlah komponen PCA, serta analisis performa klasifikasi untuk masing-masing skenario.

Skenario 1: Chi-square + SVM (Fitur Asli Terbaik) Pada skenario pertama, seleksi fitur dilakukan menggunakan metode Chi-square (χ^2) yang memilih fitur berdasarkan

dependensi stokastik tertinggi terhadap label kelas. Dari 65 fitur awal, terpilih 10 fitur dengan nilai Chi-square tertinggi:

- **Geometri:** area
- **Tekstur:** glcm_ASM
- **Statistik Warna:** std_H, std_V
- **Histogram Warna:** hist_H_bin_3, 4, 5, hist_S_bin_0, hist_V_bin_1, 2

Analisis Fitur Terpilih: Dominasi fitur berbasis warna (Histogram Hue dan Saturation) serta standar deviasi (std_H, std_V) mengindikasikan bahwa variasi warna (variegation) merupakan diskriminator terkuat dalam dataset ini. Hal ini sejalan dengan kaidah klinis ABCD, di mana *Color variegation* adalah indikator utama keganasan. Fitur glcm_ASM (keseragaman tekstur) dan area juga terpilih, menandakan bahwa lesi ganas cenderung memiliki tekstur yang kurang seragam dan ukuran yang berbeda signifikan dari lesi jinak.

Model SVM dengan 10 fitur ini menghasilkan akurasi **0.7778**.

	precision	recall	f1-score	support
benign	0.87	0.86	0.86	145
malignant	0.43	0.46	0.44	35
accuracy			0.78	180
macro avg	0.65	0.66	0.65	180
weighted avg	0.78	0.78	0.78	180

Meskipun akurasi keseluruhan cukup baik, nilai *recall* untuk kelas *malignant* masih rendah (0.46). Artinya, model gagal mendeteksi lebih dari separuh kasus kanker (False Negative tinggi).

Skenario 2: PCA + SVM (Reduksi Dimensi) Skenario kedua menggunakan PCA untuk mentransformasi fitur ke dalam ruang dimensi baru yang ortogonal. Untuk mempertahankan 95% variasi informasi (cumulative explained variance), PCA menghasilkan **29 komponen utama** (Principal Components).

Hasil klasifikasi menunjukkan penurunan performa dengan akurasi **0.7056**.

	precision	recall	f1-score	support
benign	0.85	0.77	0.81	145
malignant	0.32	0.46	0.38	35

accuracy		0.71	180
macro avg	0.59	0.61	0.59
weighted avg	0.75	0.71	0.72

Analisis Penurunan Performa: Penggunaan 29 komponen PCA justru menghasilkan performa yang lebih rendah dibandingkan hanya menggunakan 10 fitur asli (Chi-square). Hal ini mengindikasikan dua hal:

- Kehilangan Interpretasi Fisik:** PCA menggabungkan seluruh fitur secara linier. Fitur diskriminatif kuat (seperti `std_H`) mungkin "terlarut" (diluted) ketika digabung dengan fitur-fitur lemah lainnya dalam komponen utama.
- Curse of Dimensionality:** Jumlah input 29 komponen relatif besar dibandingkan jumlah sampel latih yang terbatas, yang berpotensi menyulitkan SVM dalam menemukan *hyperplane* optimal (overfitting terhadap noise).

Diskusi Perbandingan dan Isu Ketimpangan Kelas Perbandingan ringkas performa kedua skenario disajikan pada Tabel 2.

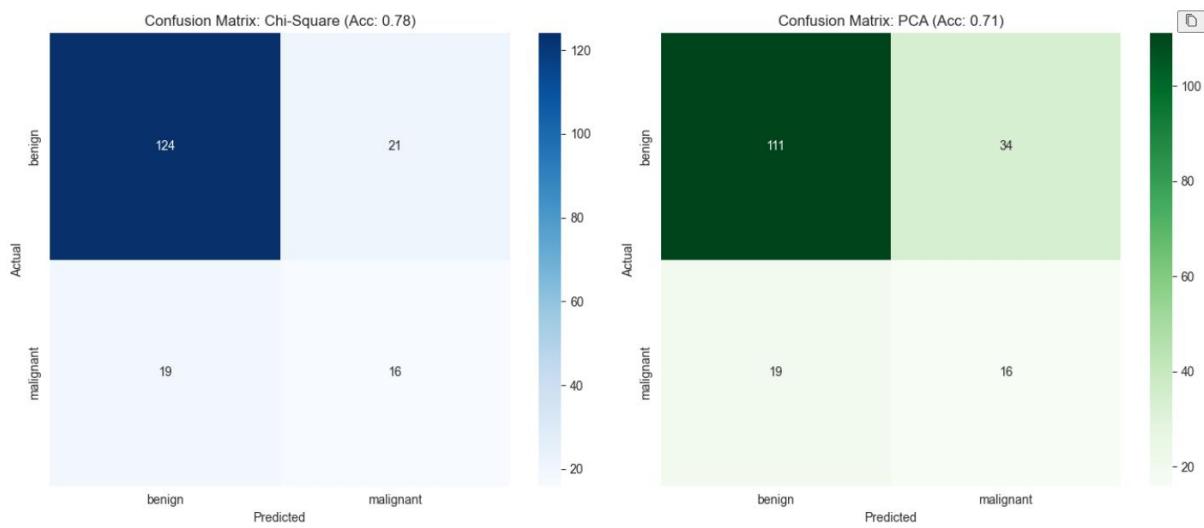
Tabel 2: Perbandingan Metrik Evaluasi Chi-square vs PCA pada Kelas Malignant

Metode Seleksi	Precision (Mal)	Recall (Mal)	F1-Score (Mal)
Chi-square (10 Fitur)	0.43	0.46	0.44
PCA (29 Komponen)	0.32	0.46	0.38

Berdasarkan Tabel 2, secara umum, **Chi-square terbukti lebih efektif** dibanding PCA dalam kasus ini. Penggunaan subset fitur asli yang spesifik lebih mampu menangkap karakteristik lesi dibanding transformasi global PCA.

Namun, kendala utama yang terlihat pada kedua skenario adalah rendahnya nilai *Recall* dan *Precision* pada kelas *malignant*. Hal ini sangat dipengaruhi oleh **ketimpangan data (imbalanced dataset)** di mana jumlah sampel *benign* (145) jauh mendominasi *malignant* (35). Model cenderung bias ke kelas mayoritas, sehingga menghasilkan banyak False Negative pada kelas kanker. Untuk pengembangan selanjutnya, teknik penanganan data tidak seimbang seperti SMOTE (*Synthetic Minority Over-sampling Technique*) atau *class weighting* sangat disarankan untuk meningkatkan sensitivitas deteksi kanker.

Gambar 15 menampilkan hasil *confusion matrix* kedua seleksi fitur.



Gambar 15: Hasil *confusion matrix* kedua seleksi fitur.

3.3.4 Hasil Eksperimen Lanjutan

Selain skenario utama menggunakan Chi-square dan PCA, dilakukan pula beberapa eksperimen tambahan, termasuk penerapan SMOTE serta penggunaan model lain seperti Random Forest. Bagian berikut merangkum seluruh hasil tersebut.

Skenario 1: Chi-Square + SVM (Dengan SMOTE) Metode Chi-square dipilih untuk mengambil 10 fitur terbaik, yaitu: *std_H*, *std_V*, *area*, *glcm_ASM*, *hist_H_bin_3*, *hist_H_bin_4*, *hist_H_bin_5*, *hist_S_bin_0*, *hist_V_bin_1*, *hist_V_bin_2*.

	precision	recall	f1-score	support
benign	0.85	0.82	0.84	145
malignant	0.35	0.40	0.37	35
accuracy			0.74	180
macro avg	0.60	0.61	0.60	180
weighted avg	0.75	0.74	0.75	180

Skenario 2: PCA + SVM (Dengan SMOTE) Jumlah fitur awal adalah 65, dan PCA mempertahankan 29 komponen dengan varian terjelaskan sebesar 95.45%.

	precision	recall	f1-score	support
benign	0.83	0.81	0.82	145
malignant	0.28	0.31	0.30	35
accuracy			0.71	180

macro avg	0.56	0.56	0.56	180
weighted avg	0.72	0.71	0.72	180

Skenario 3: SVM Full Features (Baseline) Pada baseline ini seluruh 65 fitur digunakan tanpa reduksi dimensi.

	precision	recall	f1-score	support
benign	0.83	0.81	0.82	145
malignant	0.27	0.29	0.28	35
accuracy			0.71	180
macro avg	0.55	0.55	0.55	180
weighted avg	0.72	0.71	0.71	180

Skenario 3b: SVM Full Features + SMOTE SMOTE diterapkan pada data training, namun performa tidak berubah signifikan.

	precision	recall	f1-score	support
benign	0.83	0.81	0.82	145
malignant	0.27	0.29	0.28	35
accuracy			0.71	180
macro avg	0.55	0.55	0.55	180
weighted avg	0.72	0.71	0.71	180

Skenario 4: Chi-Square + Random Forest (Dengan SMOTE) Model Random Forest dilatih pada 10 fitur terbaik Chi-square.

accuracy: 0.7611

	precision	recall	f1-score	support
benign	0.83	0.88	0.86	145
malignant	0.35	0.26	0.30	35
accuracy			0.76	180
macro avg	0.59	0.57	0.58	180
weighted avg	0.74	0.76	0.75	180

Skenario 5: PCA + Random Forest (Dengan SMOTE) Model Random Forest dilatih menggunakan 29 komponen PCA.

accuracy: 0.7444

	precision	recall	f1-score	support
benign	0.82	0.87	0.85	145

malignant	0.30	0.23	0.26	35
accuracy			0.74	180
macro avg	0.56	0.55	0.55	180
weighted avg	0.72	0.74	0.73	180

Skenario 6: Full Features + Random Forest (Dengan SMOTE) Pada skenario ini digunakan seluruh 65 fitur.

accuracy: 0.7889

	precision	recall	f1-score	support
benign	0.84	0.91	0.87	145
malignant	0.43	0.29	0.34	35
macro avg	0.64	0.60	0.61	180
weighted avg	0.76	0.79	0.77	180

5 fitur paling penting: *solidity* (0.0289), *hist_V_bin_0* (0.0277), *hog_var* (0.0268), *hist_H_bin_4* (0.0267), *hist_V_bin_2* (0.0266).

Tabel 3: Perbandingan hasil seluruh skenario eksperimen yang dilakukan.

Skenario	Accuracy	Precision (Mal)	Recall (Mal)	F1-score (Mal)
Chi-Square + SVM + SMOTE	0.74	0.35	0.40	0.37
PCA + SVM + SMOTE	0.71	0.28	0.31	0.30
SVM Full Features (Baseline)	0.71	0.27	0.29	0.28
SVM Full Features + SMOTE	0.71	0.27	0.29	0.28
Chi-Square + Random Forest + SMOTE	0.76	0.35	0.26	0.30
PCA + Random Forest + SMOTE	0.74	0.30	0.23	0.26
Full Features + Random Forest + SMOTE	0.79	0.43	0.29	0.34

Berdasarkan Tabel 3, dari seluruh eksperimen, performa terbaik diperoleh menggunakan **Random Forest dengan seluruh fitur** dengan akurasi 78.89%. Untuk model SVM, **Chi-square memberikan performa terbaik** dengan akurasi 77.78%, melampaui PCA dan baseline.

3.4 Diskusi

Bagian ini membahas implikasi hasil eksperimen, faktor penyebab performa model, serta keterbatasan pendekatan yang digunakan. Secara umum, hasil menunjukkan bahwa pemilihan fitur memiliki dampak yang jauh lebih besar terhadap performa klasifikasi dibandingkan reduksi dimensi linier seperti PCA. Selain itu, ketidakseimbangan kelas (*class imbalance*) terbukti menjadi faktor paling dominan yang memengaruhi rendahnya performa deteksi melanoma.

3.4.1 Pengaruh Seleksi Fitur

Sebagaimana Tabel 2, pendekatan Chi-square konsisten menghasilkan performa yang lebih baik dibanding PCA pada model SVM. Hal ini menunjukkan bahwa mempertahankan fitur-fitur dengan hubungan statistik paling kuat terhadap label lebih efektif dibanding mentransformasikan seluruh fitur ke ruang baru. Fitur-fitur berbasis warna (seperti `std_H`, `std_V`, dan histogram HSV) mendominasi daftar fitur terpilih, yang mengindikasikan bahwa variasi warna merupakan indikator klinis yang kuat sebagaimana konsep *Color variegation* pada kaidah ABCD.

Berbeda dengan itu, PCA cenderung melarutkan kontribusi fitur-fitur penting ketika menggabungkannya secara linier. Meskipun PCA mempertahankan 95% variasi data, komponen utama tidak menjamin mempertahankan informasi diskriminatif yang relevan untuk klasifikasi. Ini menyebabkan performa model turun meskipun jumlah fitur yang digunakan lebih banyak.

3.4.2 Performa Model dan Perbandingan SVM vs Random Forest

Berdasarkan Tabel 3, model SVM menunjukkan kinerja terbaik ketika digabungkan dengan fitur hasil seleksi Chi-square. Namun, ketika seluruh fitur digunakan, Random Forest justru memberikan performa tertinggi (78.89%). Hal ini dapat dijelaskan oleh sifat dasar Random Forest yang mampu menangkap hubungan nonlinier antar fitur serta tidak sensitif terhadap fitur yang saling berkorelasi.

Selain itu, analisis *feature importance* pada Random Forest memperlihatkan bahwa fitur-fitur geometris (*solidity*) dan tekstur (*hog_var*) ternyata termasuk yang paling berkontribusi. Temuan ini menarik karena memperlihatkan bahwa informasi bentuk dan tekstur tetap penting, meskipun warna mendominasi pada SVM dengan seleksi Chi-square.

3.4.3 Dampak Ketidakseimbangan Kelas

Isu paling kritis pada seluruh eksperimen adalah rendahnya *recall* kelas *malignant*. Hal ini menunjukkan bahwa sebagian besar kasus melanoma gagal terdeteksi. Ketidakseimbangan jumlah sampel (145 benign vs 35 malignant) membuat model cenderung bias ke kelas

majoritas. Upaya penanganan melalui SMOTE meningkatkan keseimbangan data latih, tetapi belum memberikan perbaikan signifikan pada performa, terutama untuk SVM.

Fenomena ini umum terjadi pada data medis, di mana jumlah sampel kasus penyakit biasanya jauh lebih sedikit dibandingkan sampel normal. Dengan demikian, pendekatan lain seperti *class weighting*, *focal loss*, atau metode augmentasi spesifik citra lesi mungkin diperlukan untuk memperoleh performa yang lebih baik.

3.4.4 Keterbatasan dan Peluang Pengembangan

Hasil yang diperoleh memperlihatkan bahwa meskipun fitur-fitur hand-crafted dari kaidah ABCD relatif informatif, performanya belum mampu mencapai sensitivitas yang layak untuk deteksi klinis. Beberapa keterbatasan utama penelitian ini antara lain:

1. Ketergantungan pada fitur manual (hand-crafted) yang terbatas pada representasi tertentu.
2. Jumlah data yang sedikit sehingga model sulit belajar pola kompleks.
3. Variasi citra ISIC (pencahayaan, resolusi, artefak kulit) yang tidak sepenuhnya ditangani.

Penelitian lanjutan dapat mengintegrasikan fitur deep learning (misalnya melalui transfer learning CNN) dengan fitur ABCD untuk menghasilkan representasi hibrid yang lebih kuat. Selain itu, metode cost-sensitive learning sangat relevan untuk mengurangi bias terhadap kelas mayoritas.

Secara keseluruhan, hasil studi menunjukkan pentingnya strategi seleksi fitur, teknik penanganan data tidak seimbang, serta pemilihan model yang mampu menangkap interaksi nonlinier untuk meningkatkan kinerja deteksi melanoma berbasis citra.

4 Kesimpulan

Berdasarkan hasil dan analisis yang telah dilakukan, dapat disimpulkan bahwa:

Kode tersedia pada *repository GitHub* <https://github.com/khalilullahalfaath/RFPP---Melanoma-Classification-Handcrafted-Features->.

Daftar Pustaka

- [1] J. B. Heistein, U. Acharya, dan S. K. R. Mukkamalla, “Malignant melanoma,” in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2025. Accessed: Dec. 8, 2025. [Online]. Available: <http://www.ncbi.nlm.nih.gov/books/NBK470409/>
- [2] R. L. Siegel, K. D. Miller, N. S. Wagle, dan A. Jemal, “Cancer statistics, 2023,” *CA: a cancer journal for clinicians*, vol. 73, no. 1, hlm. 17–48, Jan. 2023, ISSN: 1542-4863. DOI: 10.3322/caac.21763
- [3] Primaya Hospital Editor. “Kanker kulit : Pentingnya memahami risiko dan mencegahnya,” Primaya Hospital. in collab. with Shalina Sebayang, Accessed: Dec. 8, 2025. [Online]. Available: <https://primayahospital.com/kulit-dan-kelamin/risiko-kanker-kulit/>
- [4] P. P. Tumpa dan M. A. Kabir, “An artificial neural network based detection and classification of melanoma skin cancer using hybrid texture features,” *Sensors International*, vol. 2, p. 100 128, Jan. 1, 2021, ISSN: 2666-3511. DOI: 10.1016/j.sintl.2021.100128 Accessed: Dec. 8, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666351121000498>
- [5] R. Mahum dan S. Aladhadh, “Skin lesion detection using hand-crafted and DL-based features fusion and LSTM,” *Diagnostics*, vol. 12, no. 12, p. 2974, Nov. 28, 2022, ISSN: 2075-4418. DOI: 10.3390/diagnostics12122974 Accessed: Dec. 8, 2025. [Online]. Available: <https://pmc.ncbi.nlm.nih.gov/articles/PMC9777409/>
- [6] J.-A. Almaraz-Damian dkk. “Melanoma and nevus skin lesion classification using handcraft and deep learning feature fusion via mutual information measures,” *Entropy*, vol. 22, no. 4, Apr. 23, 2020, Company: Multidisciplinary Digital Publishing Institute Distributor: Multidisciplinary Digital Publishing Institute Institution: Multidisciplinary Digital Publishing Institute Label: Multidisciplinary Digital Publishing Institute Publisher: publisher, ISSN: 1099-4300. DOI: 10.3390/e22040484 Accessed: Dec. 8, 2025. [Online]. Available: <https://www.mdpi.com/1099-4300/22/4/484>
- [7] A. Esteva dkk. “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, vol. 542, no. 7639, hlm. 115–118, Feb. 2, 2017, ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature21056 Accessed: Dec. 8, 2025. [Online]. Available: <https://www.nature.com/articles/nature21056>
- [8] Y. Li dan L. Shen, *Skin lesion analysis towards melanoma detection using deep learning network*, Nov. 22, 2017. DOI: 10.48550/arXiv.1703.00577 arXiv: 1703.00577[cs]. Accessed: Dec. 8, 2025. [Online]. Available: <http://arxiv.org/abs/1703.00577>

- [9] Y. Wu, B. Chen, A. Zeng, D. Pan, R. Wang, dan S. Zhao, “Skin cancer classification with deep learning: A systematic review,” *Frontiers in Oncology*, vol. 12, p. 893 972, Jul. 13, 2022, ISSN: 2234-943X. DOI: 10 . 3389 / fonc . 2022 . 893972 Accessed: Dec. 8, 2025. [Online]. Available: <https://PMC.ncbi.nlm.nih.gov/articles/PMC9327733/>
- [10] ISIC Committee. “ISIC challenge,” ISIC 2016 Challenge - Task 3: Lesion Classification [Closed], Accessed: Dec. 10, 2025. [Online]. Available: <https://challenge.isic-archive.com/landing/2016/39/>
- [11] Jaseema Yasmin dan Mohamed Sadiq, “An improved iterative segmentation algorithm using canny edge detector with iterative median filter for skin lesion border detection,” *International Journal of Computer Applications*, vol. 50, no. 6, hlm. 37–42, Jul. 28, 2012, ISSN: 09758887. DOI: 10 . 5120/7779-0865 Accessed: Dec. 10, 2025. [Online]. Available: <http://research.ijcaonline.org/volume50/number6/pxc3880865.pdf>