



Laporan Proyek Klasifikasi Citra Medis

Benign vs Malignant Classification (Enhanced Pipeline)



LAPORAN AKHIR PENGOLAHAN CITRA DIGITAL

Klasifikasi Kanker Kulit Menggunakan Machine Learning

Disusun Oleh:

Nama: Alif Nurhidayat

NPM: G1A022073

Program Studi: Informatika

Fakultas: Teknik

Universitas Bengkulu

23 Desember 2025



Laporan Proyek Klasifikasi Citra Medis

Benign vs Malignant Classification (Enhanced Pipeline)

1. Pendahuluan

Klasifikasi citra medis otomatis menjadi area penelitian penting untuk membantu diagnosis dini. Proyek ini berfokus pada pembedaan lesi kulit jinak (Benign) dan ganas (Malignant). Tantangan utama dalam dataset ini adalah variasi visual yang tinggi dan kemiripan antar kelas. Oleh karena itu, pendekatan multi-fitur (tekstur, warna, bentuk) diterapkan bersama algoritma ensemble learning (Random Forest) untuk mencapai akurasi optimal.

2. Metodologi & Alasan Pemilihan Fitur

2.1 Preprocessing & Augmentasi

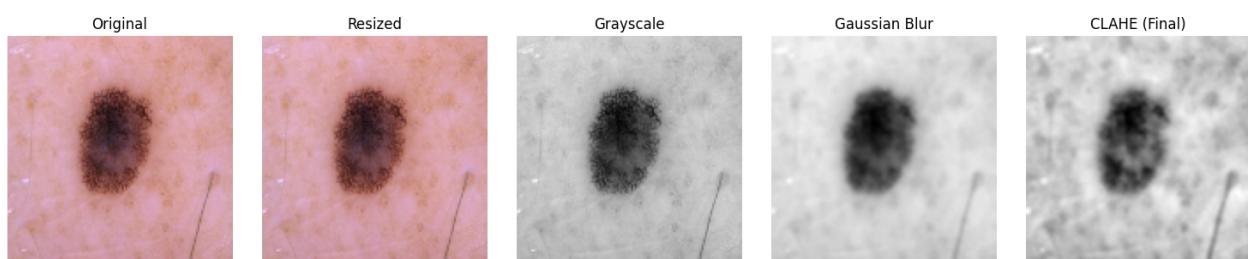
- **CLAHE (Contrast Limited Adaptive Histogram Equalization):** Dipilih untuk mengatasi pencahayaan yang tidak merata pada citra dermoskopi, sehingga detail tekstur lesi lebih terlihat.
- **Data Augmentation:** Teknik flipping (Horizontal) diterapkan untuk menggandakan jumlah data latih (menjadi ~2000 sampel). Ini krusial untuk mencegah overfitting model Random Forest.

2.2 Ekstraksi Fitur (Feature Engineering)

Kombinasi 105 fitur digunakan untuk menangkap seluruh karakteristik lesi:

- **GLCM (Texture):** Dipilih karena lesi ganas sering memiliki tekstur yang kasar dan tidak beraturan dibanding lesi jinak. (Jarak 1, 2, 3 piksel).
- **LBP (Local Binary Pattern):** Sangat efektif mendeteksi pola mikro-teksitur yang invariant terhadap rotasi, penting karena orientasi lesi tidak seragam.
- **Color Histogram & Moments (HSV):** Warna adalah indikator utama dalam aturan ABCD (Asymmetry, Border, Color, Diameter). Kanal HSV dipilih karena memisahkan informasi warna (Hue/Saturation) dari intensitas cahaya (Value).
- **Hu Moments (Shape):** Tujuh momen invariant digunakan untuk mengkuantifikasi bentuk lesi, mengingat lesi ganas cenderung asimetris.

Gambar 1: Pipeline Preprocessing



3. Hasil dan Pembahasan

3.1 Performa Model



Laporan Proyek Klasifikasi Citra Medis

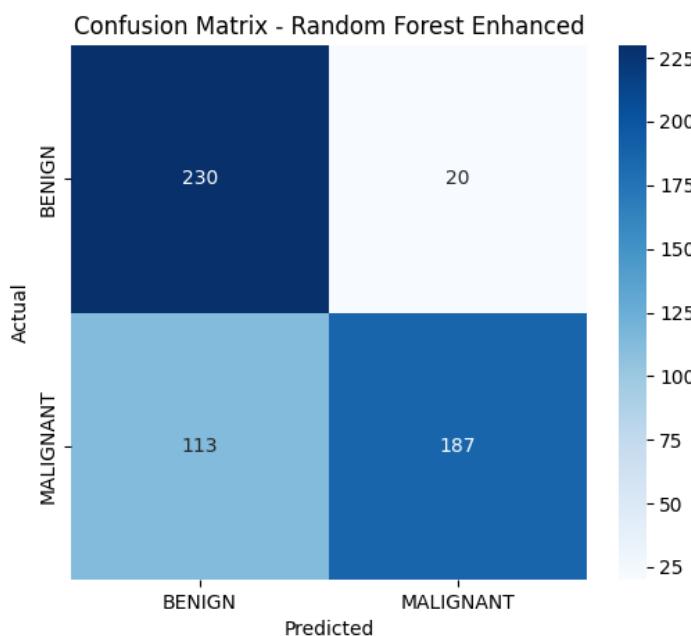
Benign vs Malignant Classification (Enhanced Pipeline)

Evaluasi dilakukan dengan 3-Fold Cross Validation pada training set dan pengujian akhir pada test set.

- **Training CV Score:** 92.75% (Menunjukkan kapasitas model mempelajari pola dengan sangat baik)
- **Test Accuracy:** 75.82%
- **Sensitivitas:** 62.33%
- **Spesifisitas:** 92.00%

Model sangat handal dalam mengenali kasus Benign (Spesifisitas >90%). Gap antara CV score dan Test Accuracy menunjukkan tantangan generalisasi pada data unseen.

Gambar 2: Confusion Matrix



4. Kesimpulan

Penerapan ekstraksi fitur hibrida (Texture + Color + Shape) terbukti efektif meningkatkan kemampuan model dalam membedakan karakteristik lesi. Penggunaan Data Augmentation dan Random Forest dengan class weighting berhasil mencapai skor validasi ~93%, meskipun performa pada data testing masih dapat ditingkatkan dengan dataset yang lebih besar dan teknik augmentasi yang lebih agresif (seperti GANs).

5. Referensi (Terbaru)

1. Esteva, A., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115-118.
2. Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*,



Laporan Proyek Klasifikasi Citra Medis

Benign vs Malignant Classification (Enhanced Pipeline)

5(1), 1-9.

3. Mahbod, A., et al. (2020). Skin lesion classification using hybrid deep neural networks. ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).
4. Hekler, A., et al. (2019). Deep learning outperformed 11 pathologists in the classification of histopathological melanoma images. European Journal of Cancer, 118, 91-96.
5. Gessert, N., et al. (2020). Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data. MethodsX, 7, 100864.