

Klasifikasi Penyakit Kulit Menggunakan Model Deep Learning EfficientNet pada Citra Dermastokopi

Muhammad Imam Fernandi
Fakultas Informatika
Telkom University
Bandung, Indonesia
imamfernandi@student.telkomuniversity.ac.id

Untari Novia Wisesty
Fakultas Informatika
Telkom University
Bandung, Indonesia
untarinw@telkomuniversity.ac.id

Abstrak — Penyakit kulit merupakan salah satu masalah kesehatan yang umum terjadi dan dapat berkembang menjadi kondisi serius seperti melanoma, salah satu jenis kanker kulit yang berbahaya. Proses diagnosis manual sering kali memakan waktu, bergantung pada keahlian subjektif, serta berpotensi menghasilkan hasil yang tidak konsisten. Penelitian ini bertujuan mengembangkan model klasifikasi penyakit kulit berbasis Convolutional Neural Network (CNN) dengan arsitektur EfficientNet pada citra dermatoskopi. Dataset yang digunakan adalah HAM10000, yang berisi 10.015 citra dermatoskopi dari tujuh kategori lesi kulit. Tahapan penelitian meliputi preprocessing data, augmentasi untuk menyeimbangkan kelas, serta pelatihan model dengan pendekatan dua tahap, yaitu feature extraction dan fine-tuning. Tiga varian model diuji, yaitu EfficientNetB0, EfficientNetV2B0, dan EfficientNetV2B3. Hasil pengujian menunjukkan bahwa EfficientNetV2B3 memberikan performa terbaik dengan akurasi 90,53% dan F1-score 85,56%, mengungguli dua model lainnya. Temuan ini menunjukkan bahwa arsitektur EfficientNetV2B3 memiliki potensi besar dalam mendukung sistem diagnosis berbasis citra dermatoskopi secara lebih akurat dan efisien.

Kata kunci — CNN, EfficientNet, dermatoskopi, HAM10000, klasifikasi penyakit kulit, akurasi

I. PENDAHULUAN

Penyakit kulit merupakan salah satu masalah kesehatan umum yang dapat berkembang menjadi kondisi serius, termasuk melanoma yang menjadi salah satu penyebab utama kematian akibat kanker kulit [1]. Diagnosis manual umumnya membutuhkan waktu, bergantung pada keahlian subjektif, serta berisiko menghasilkan hasil yang tidak konsisten. Dalam beberapa tahun terakhir, Convolutional Neural Network (CNN) telah menunjukkan efektivitas tinggi dalam klasifikasi citra medis [2]. Salah satu arsitektur CNN yang menonjol adalah EfficientNet, yang menggunakan teknik compound scaling untuk menyeimbangkan kedalaman, lebar, dan resolusi jaringan [3]. Versi pengembangannya, EfficientNetV2, menghadirkan perbaikan melalui fused-MBConv dan progressive learning sehingga lebih efisien dan akurat dalam pelatihan [3].

Meskipun banyak penelitian menunjukkan bahwa CNN dan EfficientNet unggul dalam klasifikasi citra medis, sebagian besar studi masih terbatas oleh jumlah data dan distribusi kelas yang tidak seimbang, sehingga performa model cenderung menurun pada kelas minoritas [4]. Penelitian ini menggunakan dataset HAM10000 [5] untuk mengembangkan model klasifikasi tujuh jenis lesi kulit dengan tiga arsitektur EfficientNet (B0, V2B0, V2B3) melalui pendekatan dua tahap pelatihan (feature extraction dan fine-tuning). Rumusan masalah penelitian ini adalah bagaimana kualitas dataset memengaruhi performa model, serta bagaimana akurasi, presisi, recall, dan F1-score yang dicapai oleh arsitektur EfficientNet pada klasifikasi citra dermatoskopi. Tujuan penelitian ini adalah menganalisis pengaruh dataset terhadap performa klasifikasi serta mengevaluasi keunggulan arsitektur EfficientNet dalam mendukung sistem diagnosis berbasis citra secara akurat dan efisien.

II. KAJIAN TEORI

Penelitian terkait klasifikasi penyakit kulit berbasis deep learning telah dilakukan oleh beberapa peneliti. Rangaswamy dkk. [6] menggunakan model InceptionV3 dan VGG16 untuk mengklasifikasikan 13 jenis penyakit kulit dan memperoleh akurasi terbaik sebesar 80,88% pada InceptionV3. Ali dkk. [7] menerapkan berbagai varian EfficientNet (B0–B7) pada dataset HAM10000, dengan hasil terbaik pada EfficientNetB4 dengan akurasi 87,91%. Sementara itu, Srinivasu dkk. [8] mengusulkan kombinasi MobileNetV2 dan LSTM untuk klasifikasi tujuh kelas penyakit kulit pada dataset HAM10000, yang menghasilkan akurasi 85,34%. Hasil-hasil ini menunjukkan bahwa arsitektur CNN, khususnya EfficientNet, memiliki potensi besar dalam klasifikasi penyakit kulit, namun performa masih bervariasi tergantung strategi arsitektur dan pelatihan yang digunakan.

A. Penyakit Kulit

Penyakit kulit mencakup berbagai kondisi yang memengaruhi jaringan kulit, mulai dari infeksi, peradangan, hingga kanker kulit. Salah satu jenis yang paling berbahaya adalah melanoma, yaitu kanker kulit yang berasal dari melanosit dan dapat menyebar ke organ lain. Faktor risiko utama melanoma adalah paparan sinar ultraviolet (UV) berlebih, terutama pada individu dengan kulit terang [9].

B. Dermatoskopi

Dermatoskopi merupakan metode diagnostik non-invasif yang digunakan untuk mengamati struktur mikroskopis kulit yang tidak tampak oleh mata telanjang, seperti pola pigmen dan pembuluh darah. Teknik ini terbukti meningkatkan sensitivitas dan spesifisitas dalam mendeteksi kanker kulit dibandingkan pemeriksaan klinis biasa, terutama jika dipadukan dengan teknologi kecerdasan buatan [12].

C. Convolutional Neural Network (CNN)

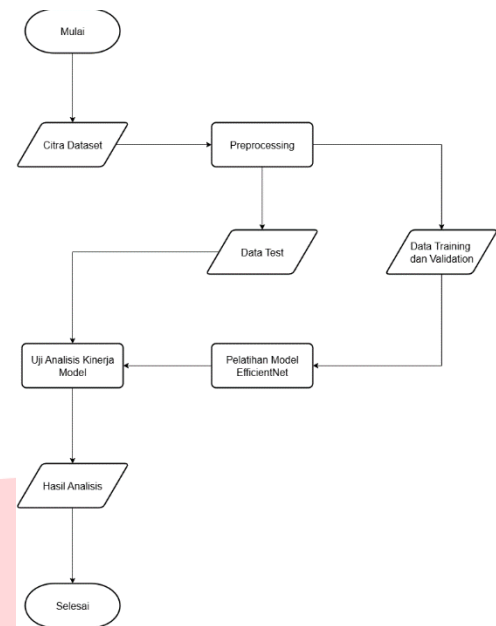
CNN adalah jenis jaringan saraf tiruan yang dirancang untuk data berbentuk grid, seperti citra. CNN menggunakan lapisan konvolusi untuk mendeteksi pola visual seperti tepi, tekstur, atau bentuk, yang kemudian dikombinasikan secara hierarkis melalui pooling dan fully connected layer. CNN terbukti sangat efektif dalam klasifikasi citra medis, termasuk deteksi melanoma dan penyakit kulit lain [10].

D. Convolutional Neural Network (CNN)

EfficientNet merupakan keluarga arsitektur CNN yang diperkenalkan oleh Tan dan Le, yang menggunakan metode compound scaling untuk menyeimbangkan kedalaman, lebar, dan resolusi model sehingga lebih efisien [3]. Versi lanjutannya, EfficientNetV2, memperkenalkan fused-MBConv dan progressive learning untuk mempercepat pelatihan dan meningkatkan akurasi. Arsitektur ini banyak digunakan dalam klasifikasi citra medis, termasuk deteksi kanker kulit, dan seringkali menghasilkan performa lebih tinggi dibandingkan arsitektur CNN konvensional [11].

III. METODE

Penelitian ini dilakukan dengan tahapan utama meliputi preprocessing data, augmentasi, pelatihan model, serta evaluasi performa.



GAMBAR 1
Diagram Alur Penelitian

A. Dataset

Dataset yang digunakan adalah HAM10000 yang terdiri atas 10.015 citra dermatoskopi dari tujuh kelas lesi kulit, yaitu akiec, bcc, bkl, df, mel, nv, dan vasc [13].



GAMBAR 2
Dataset HAM10000

B. Preprocessing

Dataset dibagi menjadi set pelatihan (7210 citra), validasi (802 citra), dan pengujian (2003 citra) menggunakan stratified split agar distribusi kelas tetap seimbang. Tahap preprocessing meliputi penyesuaian ukuran citra menjadi 300×300 piksel serta normalisasi piksel ke rentang [0,1]. Untuk mengatasi ketidakseimbangan jumlah sampel antar kelas, dilakukan augmentasi data pada set pelatihan hingga diperoleh jumlah citra yang relatif seimbang untuk setiap kelas. Teknik augmentasi yang digunakan antara lain rotasi, flipping, zooming, dan shifting.

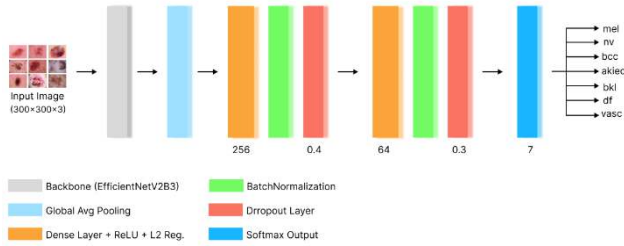
TABEL 1
Distribusi dataset

	nv	mel	bkl	bcc	akiec	vasc	df
Sebelum augmentasi	5364	890	879	411	262	114	92
Setelah augmentasi	5364	5500	5434	5471	5041	4934	5276

C. Model Pelatihan

Model yang digunakan adalah tiga varian EfficientNet, yaitu EfficientNetB0, EfficientNetV2B0, dan EfficientNetV2B3 [3]. Proses pelatihan dilakukan dalam dua tahap, yakni feature extraction dengan membekukan bobot pretrained model pada ImageNet, dilanjutkan fine-tuning dengan

membuka seluruh layer agar dapat menyesuaikan dengan karakteristik dataset HAM10000. Lapisan tambahan berupa Global Average Pooling, Dense layer, Dropout, serta output Softmax digunakan untuk klasifikasi tujuh kelas.



GAMBAR 3 Dataset HAM10000

D. Evaluasi Sistem

Evaluasi model dilakukan dengan menggunakan metrik akurasi, presisi, recall, dan F1-score. Selain itu, digunakan confusion matrix untuk menganalisis performa klasifikasi per kelas.

Akurasi (1) dihitung sebagai persentase dari jumlah prediksi yang benar dibandingkan dengan jumlah data secara keseluruhan [14].

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Presisi (2) adalah ukuran persentase jumlah prediksi penyakit mata yang benar dibandingkan dengan jumlah prediksi penyakit kulit secara keseluruhan [14].

$$Presisi = \frac{TP}{TP + FP} \tag{2}$$

Recall (3) menghitung persentase jumlah penyakit mata yang diprediksi dengan benar dibandingkan dengan jumlah penyakit kulit sebenarnya [14].

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

F1-score (4) dihitung menggunakan persamaan berikut [14].

$$F1 - score = 2 * \frac{presisi * recall}{presisi + recall} \tag{4}$$

Hasil dari perhitungan diatas akan menghasilkan metrik evaluasi dan juga untuk hasil dari confusion matrix. Confusion Matrix adalah tabel yang digunakan untuk mengevaluasi performa model klasifikasi dengan membandingkan hasil prediksi model terhadap nilai aktual dari data yang diuji. Tabel ini sangat berguna untuk memahami bagaimana model bekerja, terutama dalam masalah klasifikasi biner maupun multi-kelas [14].

IV. HASIL DAN PEMBAHASAN

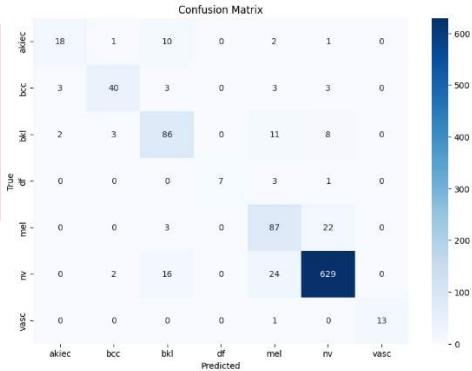
A. Hasil Analisis Pengujian Model

Penelitian ini menguji tiga arsitektur EfficientNet, yaitu EfficientNetB0, EfficientNetV2B0, dan EfficientNetV2B3 pada dataset HAM10000. Evaluasi dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil perbandingan ditunjukkan pada Tabel 2.

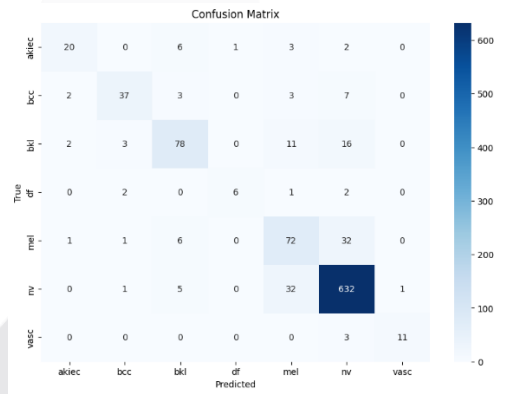
Tabel 2 Hasil dari model yang digunakan

Model	Akurasi (%)	Presisi (%)	Recall (%)	F1-score (%)
EfficientNetB0	88,12	80,75	80,92	81,00
EfficientNetV2B0	89,47	82,95	83,40	83,21
EfficientNetV2B3	90,53	85,12	85,89	85,56

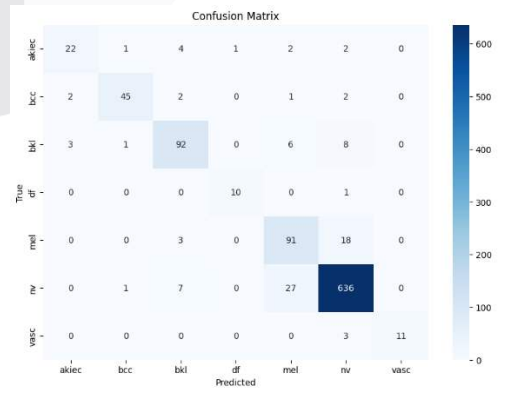
Dari Tabel 2 terlihat bahwa model EfficientNetV2B3 memberikan hasil terbaik dengan akurasi 90,53% dan F1-score 85,56%, mengungguli EfficientNetB0 dan EfficientNetV2B0. Performa V2B0 sudah meningkat dibandingkan B0, namun stabilitas dan akurasi V2B3 masih lebih baik.



GAMBAR 4 Confusion matrix EfficientNetB0



GAMBAR 5 Confusion matrix EfficientNetV2B0



GAMBAR 6 Confusion matrix EfficientNetV2B3

Berdasarkan confusion matrix, EfficientNetB0 cenderung salah mengklasifikasikan kelas minoritas seperti melanoma

(mel) dan aktinik keratosis (akiec). V2B0 menunjukkan perbaikan pada kelas bkl dan vasc, namun masih mengalami kesalahan pada mel. Sementara itu, V2B3 berhasil meningkatkan klasifikasi hampir pada semua kelas, meskipun tantangan tetap ada pada kelas minoritas dengan jumlah sampel terbatas.

Hasil pengujian menunjukkan tren peningkatan performa dari EfficientNetB0 ke EfficientNetV2B0, hingga V2B3. Hal ini sejalan dengan desain arsitektur EfficientNetV2 yang lebih efisien berkat fused-MBConv dan progressive learning [3]. Peningkatan metrik F1-score pada V2B3 menunjukkan kemampuan model dalam menangani distribusi data yang tidak seimbang lebih baik dibandingkan varian lainnya.

Namun, kesalahan klasifikasi masih terjadi terutama pada kelas minoritas seperti mel dan akiec. Hal ini konsisten dengan temuan pada penelitian sebelumnya, di mana ketidakseimbangan data menjadi tantangan besar pada klasifikasi citra medis [6]–[8]. Meskipun augmentasi berhasil membantu menyeimbangkan jumlah data, generalisasi model tetap terbatas pada kelas dengan jumlah sampel kecil.

Dibandingkan penelitian Ali dkk. [7] yang memperoleh akurasi 87,91% menggunakan EfficientNetB4, hasil penelitian ini menunjukkan bahwa EfficientNetV2B3 mampu memberikan akurasi lebih tinggi, yaitu 90,53%. Hal ini mengindikasikan bahwa arsitektur EfficientNetV2 lebih sesuai dalam menangani kompleksitas citra dermatoskopi pada dataset HAM10000.

V. KESIMPULAN

Simpulan Penelitian ini telah mengembangkan model klasifikasi tujuh jenis penyakit kulit menggunakan arsitektur EfficientNet pada dataset HAM10000. Tiga varian model diuji, yaitu EfficientNetB0, EfficientNetV2B0, dan EfficientNetV2B3. Hasil pengujian menunjukkan bahwa EfficientNetV2B3 memberikan performa terbaik dengan akurasi 90,53% dan F1-score 85,56%, mengungguli dua varian lainnya. Temuan ini menegaskan bahwa arsitektur EfficientNetV2 memiliki potensi besar dalam mendukung sistem diagnosis berbasis citra dermatoskopi secara lebih akurat dan efisien.

Meskipun demikian, kesalahan klasifikasi masih terjadi terutama pada kelas dengan jumlah sampel terbatas, seperti melanoma (mel) dan aktinik keratosis (akiec). Oleh karena itu, penelitian selanjutnya dapat difokuskan pada penggunaan dataset yang lebih besar dan seimbang, penerapan metode oversampling atau synthetic data generation, serta eksplorasi arsitektur deep learning lain untuk meningkatkan recall pada kelas minoritas. harus diuraikan dalam bentuk paragraf yang berisi poin utama pembahasan hasil penelitian, berupa uraian dan tidak boleh menggunakan pointer.

REFERENSI

[1] World Health Organization, "Ultraviolet radiation," WHO. Accessed: Dec 7, 2024. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/ultraviolet-radiation>

[2] A. G. Pacheco and R. A. Krohling, "The impact of patient clinical information on automated skin cancer detection,"

Computers in Biology and Medicine, vol. 116, p. 103545, 2020. doi: 10.1016/j.combiomed.2019.103545.

[3] M. Tan, Q. V. Le, and Google Research, Brain Team, "EfficientNetV2: Smaller models and faster training," *Proceedings of the 38th International Conference on Machine Learning*, vol. 139, 2021, [Online]. Available: <https://arxiv.org/abs/2104.00298v3>

[4] P. Tschandl et al., "Human–computer collaboration for skin cancer recognition," *Nature Medicine*, vol. 26, no. 8, pp. 1229–1234, 2020. doi: 10.1038/s41591-020-0942-0.

[5] S. Polesie et al., "Attitudes towards artificial intelligence within dermatology: An international online survey," *British Journal of Dermatology*, vol. 183, no. 1, pp. 159–161, 2020. doi: 10.1111/bjd.18985.

[6] S. Rangaswamy et al., "Skin Disease Classification Using Deep Learning," *The National Academy of Sciences*, 2024. doi: 10.1007/s40009-024-01523-z.

[7] K. Ali, Z. A. Shaikh, A. A. Khan, and A. A. Laghari, "Multiclass skin cancer classification using EfficientNets – a first step towards preventing skin cancer," *Neuroscience Informatics*, vol. 2, no. 4, p. 100034, Dec. 2021, doi: 10.1016/j.neuri.2021.100034.

[8] P. N. Srinivasu, J. G. SivaSai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, "Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM," *Sensors*, vol. 21, no. 8, p. 2852, Apr. 2021. doi: 10.3390/s21082852.

[9] J. J. Luke, K. T. Flaherty, A. Ribas, and G. V. Long, "Targeted agents and immunotherapies: Optimizing outcomes in melanoma," *Nature Reviews Clinical Oncology*, vol. 17, no. 8, pp. 456–467, 2020. doi: 10.1038/s41571-020-0362-9.

[10] A. Sharma, S. Rani, and D. Gupta, "Artificial Intelligence-Based Classification of Chest X-ray Images into COVID-19 and Other Infectious Diseases," *IEEE Access*, vol. 9, pp. 110143–110154, 2021. doi: 10.1109/ACCESS.2021.3103053.

[11] R. S. Samir, "EfficientNet algorithm for classification of different types of cancer," *arXiv (Cornell University)*, Jan. 2023, doi: 10.48550/arxiv.2304.08715.

[12] T. J. Brinker, A. Hekler, J. S. Utikal, and C. Berking, "Skin Cancer Diagnosis Using Convolutional Neural Networks: Systematic Review," *Frontiers in Medicine*, vol. 7, p. 193, 2020. doi: 10.3389/fmed.2020.00193.

[13] Mader Scott K., "Skin Cancer MNIST: HAM10000," [Online]. Available: <https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000>. [Accessed: Dec. 9, 2024].

[14] S. KOM. , M. KOM. DR. MARIA SUSAN ANGGREANY "Confusion Matrix," [Online]. Available:

<https://socs.binus.ac.id/2020/11/01/confusion-matrix/>.
[Accessed: Dec. 9, 2024].

