

# Klasifikasi Kanker Kulit menggunakan Custom CNN dengan SMOTE-Tomek dan Optimizer Nadam

MAYLINDA CHRISTY YOSEFINA TALAN, SYAFRIAL FACHRI PANE, RD.  
NURAINI SITI FATHONAH

Universitas Logistik dan Bisnis Internasional, Indonesia  
Email: 1214057@std.ulbi.ac.id

## ABSTRAK

Deteksi dini kanker kulit sangat penting untuk meningkatkan harapan hidup, namun diagnosis konvensional seringkali subjektif. Berbeda dari penelitian sebelumnya, penelitian ini mengusulkan sebuah konfigurasi optimal yang menggabungkan tiga komponen kunci secara simultan: arsitektur *Custom Convolutional Neural Network* (CNN) yang ringan, penanganan ketidakseimbangan data menggunakan SMOTE-Tomek, dan optimisasi pelatihan dengan *optimizer* Nadam. Pendekatan terintegrasi yang dievaluasi pada dataset HAM10000 ini terbukti mampu mencapai efisiensi komputasi dan akurasi yang tinggi. Hasil eksperimen menunjukkan model mencapai akurasi validasi hingga 95.63% dan nilai F1-score  $\geq 0.90$ , bahkan pada kelas minoritas seperti melanoma. Model ini juga berhasil diimplementasikan dalam aplikasi web dengan *confidence score* di atas 90%, membuktikan bahwa pendekatan yang diusulkan mampu memberikan solusi diagnosis yang objektif dan terukur untuk klasifikasi otomatis kanker kulit.

**Kata kunci:** Klasifikasi Kanker Kulit; Citra Dermoskopi; Custom CNN; SMOTE-Tomek; Optimizer Nadam.

## ABSTRACT

*Early detection of skin cancer is critical to improving survival rates, yet conventional diagnosis is often subjective. This study develops an objective dermoscopic image classification system using deep learning. The proposed model utilizes a lightweight Custom Convolutional Neural Network (CNN) architecture, combined with the SMOTE-Tomek method to handle data imbalance in the HAM10000 dataset. The training process was optimized using the Nadam optimizer with a 90:10 data split and 64x64 pixel image inputs. Experimental results show the model achieved a validation accuracy of up to 95.63% and an F1-score  $\geq 0.90$ , even on minority classes like melanoma. The model, successfully implemented in a web application with confidence scores above 90%, proves to be an effective solution for automatic skin cancer classification.*

**Keywords:** Skin Cancer Classification; Dermoscopic Images; Custom CNN; SMOTE-Tomek; Nadam Optimizer.

## 1. PENDAHULUAN

Kanker kulit merupakan masalah kesehatan global yang serius, dengan lebih dari 1,5 juta kasus baru pada tahun 2022, di mana melanoma menjadi jenis paling mematikan. Deteksi dini sangat krusial, karena tingkat kelangsungan hidup lima tahun dapat mencapai >99% jika terdeteksi awal, namun menurun drastis menjadi 35% jika sudah bermetastasis (**Wang et al., 2025**). Diagnosis konvensional masih bergantung pada pemeriksaan visual menggunakan aturan ABCDE, namun metode ini rentan terhadap subjektivitas dan kesalahan, terutama pada tahap awal. Oleh karena itu, dibutuhkan alat bantu diagnosis yang objektif dan efisien untuk mendukung akurasi deteksi klinis (**Mamun et al., 2025**).

Sebagai respons terhadap tantangan ini, sistem *Computer-Aided Diagnosis* (CAD) telah muncul sebagai bidang penelitian yang menjanjikan (**Kadhim et al., 2022**). Secara khusus, kemajuan dalam *Deep Learning* (DL), dan terutama *Convolutional Neural Networks* (CNNs), telah merevolusi bidang analisis citra medis (**AYENI, 2022**). Kemampuan revolusioner ini pertama kali dibuktikan oleh arsitektur seminal seperti AlexNet (**Alzubaidi et al., 2021**) dan kemudian diperdalam oleh arsitektur yang lebih kompleks seperti VGG19 (**Jaworek-Korjakowska et al., 2019**), yang menetapkan standar baru dalam klasifikasi citra skala besar. CNN adalah kelas arsitektur jaringan saraf yang terinspirasi oleh korteks visual manusia, yang mampu secara otomatis mempelajari fitur-fitur hierarkis langsung dari data citra (**Radha Jayalakshmi & Rajakumar, 2024**). Kekuatan DL dalam mempelajari fitur secara otomatis ini tidak hanya terbatas pada citra, tetapi juga telah menunjukkan efektivitas pada data sekuensial seperti teks untuk tugas analisis sentimen (**Pane & Ramdan, 2022**) dan deteksi emosi (**Pane et al., 2024**). Dengan arsitektur berlapisnya, CNN dapat mendeteksi pola-pola kompleks, mulai dari tepi dan tekstur sederhana pada lapisan awal hingga bentuk dan objek yang lebih rumit pada lapisan yang lebih dalam, menjadikannya sangat cocok untuk tugas klasifikasi citra dermoskopi (**Amartama et al., 2024**).

Salah satu sumber daya paling penting untuk melatih model DL dalam dermatologi adalah dataset HAM10000 (*Human Against Machine with 10,000 training images*), yang berisi 10.015 citra dari tujuh kategori lesi kulit (**Shetty et al., 2022**). Namun, dataset ini memiliki tantangan yang signifikan: ketidakseimbangan kelas yang parah, di mana kelas jinak seperti *Melanocytic Nevi* mendominasi lebih dari 66% data. Distribusi yang timpang ini menjadi tantangan komputasional yang krusial, karena model yang dilatih pada data ini akan cenderung bias dan gagal mengenali kelas minoritas yang langka namun kritis secara medis, seperti melanoma. Oleh karena itu, perumusan masalah dalam penelitian ini adalah untuk menjawab pertanyaan: "Bagaimana merancang sebuah model klasifikasi berbasis *Custom CNN* yang optimal, dengan mengintegrasikan teknik penyeimbangan data dan optimizer yang tepat, untuk dapat meningkatkan performa klasifikasi secara signifikan pada seluruh kelas, terutama pada kelas minoritas?"

Dalam kurun waktu lima tahun terakhir, pendekatan *Deep Learning* (DL), khususnya *Convolutional Neural Networks* (CNN), telah menjadi metode yang dominan dalam tugas deteksi kanker kulit berbasis citra dermoskopi. (**Albahar, 2019**) fokus pada pengembangan teknik regularisasi baru untuk mencegah *overfitting* pada data medis yang terbatas. (**Patil et al., 2022**) mengimplementasikan arsitektur CNN sederhana untuk klasifikasi melanoma pada dataset HAM10000 serta melakukan perbandingan terhadap beberapa algoritma optimisasi, meskipun belum mencakup eksplorasi teknik penyeimbangan data seperti SMOTE-Tomek. Sementara itu, (**Swapno et al., 2025**) mengembangkan pipeline hybrid berbasis *EfficientNetV2L* dan *LightGBM* yang dilengkapi strategi augmentasi lanjutan, dan berhasil mencapai akurasi tertinggi sebesar 99,90%. Di sisi lain, (**Wei et al., 2020**) mengembangkan

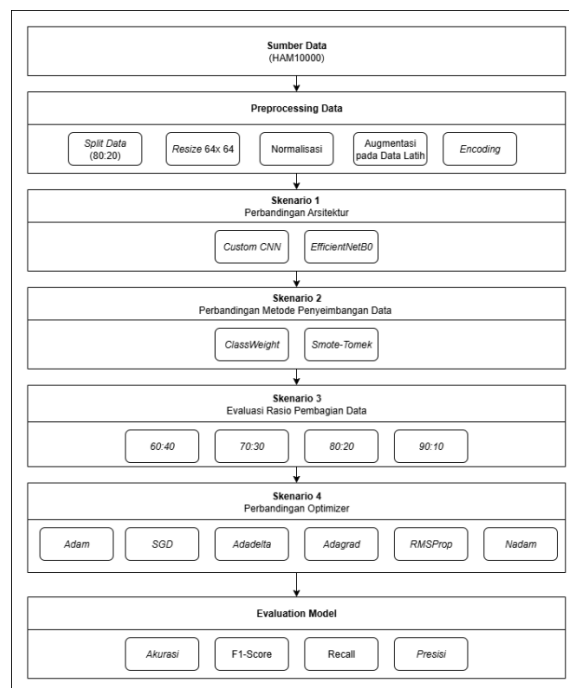
arsitektur CNN ringan berbasis ensemble yang dirancang untuk perangkat dengan keterbatasan sumber daya, selaras dengan penelitian oleh **(Putrada et al., 2023)** yang menerapkan teknik knowledge distillation dan berhasil menunjukkan bahwa efisiensi komputasi tidak harus mengorbankan akurasi, dengan capaian performa di atas 96%. Berdasarkan tinjauan tersebut, terlihat adanya *trade-off* antara arsitektur kompleks yang berat namun akurat dan arsitektur ringan yang efisien namun terkadang mengorbankan performa. Oleh karena itu, penelitian ini berfokus pada perancangan arsitektur *Custom CNN* yang bertujuan untuk menemukan keseimbangan optimal antara efisiensi komputasi (model ringan) dan performa klasifikasi yang tinggi, sebuah justifikasi yang belum dieksplorasi secara sistematis oleh penelitian-penelitian sebelumnya yang dibandingkan.

Secara spesifik, tujuan dari penelitian ini adalah: (1) mengembangkan model *Custom CNN* yang akurat, efisien, dan andal dalam mengatasi masalah ketidakseimbangan kelas; (2) menganalisis secara komparatif pengaruh dari setiap komponen untuk membuktikan bahwa kombinasi yang diusulkan merupakan konfigurasi optimal; dan (3) mengimplementasikan model terbaik ke dalam aplikasi web fungsional sebagai bukti konsep. Adapun ruang lingkup penelitian ini dibatasi pada penggunaan dataset HAM10000, dengan arsitektur *Custom CNN* dan teknik *resampling* SMOTE-Tomek, rasio data optimal (90:10), dan optimizer Nadam sebagai metode utama, yang menghasilkan klasifikasi untuk tujuh kelas lesi kulit dalam bentuk purwarupa aplikasi web.

## 2. METODE

### 2.1 Tahapan Penelitian

Metodologi penelitian ini dirancang secara sistematis untuk menemukan konfigurasi terbaik dalam model klasifikasi citra medis. 1 mengilustrasikan keseluruhan tahapan penelitian ini secara rinci.



**Gambar 1. Tahapan Penelitian yang diusulkan**

## 2.2 Sumber Data

Penelitian ini menggunakan dataset publik HAM10000 (*Human Against Machine with 10,000 training images*), yang tersedia secara terbuka melalui platform *Kaggle* (Mader, 2018). Dataset ini dikembangkan oleh tim dari Departemen Dermatologi Universitas Kedokteran Wina, Austria, dan didukung oleh praktik kanker kulit di Queensland, Australia (Tajerian et al., 2023). Dataset terdiri atas 10.015 citra dermoskopi berwarna (RGB) yang terbagi ke dalam tujuh kategori lesi kulit berpigmen, yaitu: Melanocytic Nevi (nv), Melanoma (mel), Benign keratosis-like lesions (bkl), Basal cell carcinoma (bcc), Actinic keratoses (akiec), Vascular lesions (vasc), dan Dermatofibroma (df). Seperti yang telah diuraikan pada Bab Pendahuluan, tantangan utama dari dataset ini adalah ketidakseimbangan kelas yang signifikan. Distribusi jumlah sampel untuk setiap kelas disajikan secara rinci pada Tabel 1.

**Tabel 1. Distribusi Kelas Dataset pada HAM10000**

Kelas	Singkatan	Jumlah Citra	Persentase(%)
Melanocytic Nevi	nv	6705	66.95
Melanoma	mel	1113	11.11
Benign Keratosis-like Lesions	bkl	1099	10.97
Basal Cell Carcinoma	bcc	514	5.13
Actinic Keratoses	akiec	327	3.27
Vascular Lesions	vasc	142	1.42
Dermatofibroma	df	115	1.15

## 2.3 Pra-pemrosesan Data

Seluruh data dalam penelitian ini melalui serangkaian tahap persiapan sebelum dimasukkan ke model. Untuk data citra, semua gambar (baik latih maupun validasi) yang awalnya berukuran 600x450 piksel diubah ukurannya menjadi 64x64 piksel. Selanjutnya, dilakukan normalisasi nilai intensitas piksel dengan menyekalakannya dari rentang [0, 255] menjadi [0, 1], dengan tetap mempertahankan format warna RGB. Kemudian, khusus pada data latih, diterapkan teknik augmentasi data *on-the-fly* untuk meningkatkan keragaman, yang mencakup transformasi acak seperti rotasi, pergeseran, zoom, dan flip. Selain itu untuk data label kelas yang berupa teks (misalnya 'nv', 'mel') dikonversi menjadi format numerik *one-hot vector* untuk disesuaikan dengan fungsi *loss* pada model klasifikasi multi-kelas.

## 2.4 Metode yang Diusulkan

### 2.4.1 Arsitektur Model Custom CNN

Desain arsitektur ini terinspirasi dari studi sebelumnya yang menunjukkan bahwa arsitektur Convolutional Neural Network yang sederhana tetapi mampu memberikan performa klasifikasi yang kompetitif terhadap citra lesi kulit (Patil et al., 2022). Arsitektur Custom CNN yang dirancang bersifat ringan dan disesuaikan untuk input beresolusi 64x64 piksel, dengan tujuan menguji efektivitas model ringkas dalam menghasilkan performa tinggi sekaligus efisiensi komputasi. Konfigurasi spesifik dari arsitektur ini dirancang berdasarkan prinsip-prinsip umum dalam desain CNN dan penyesuaian empiris. Peningkatan jumlah filter secara bertahap (32 → 64 → 128) pada lapisan konvolusi merupakan praktik standar untuk memungkinkan model mempelajari fitur dari yang paling sederhana (tepi, tekstur) hingga yang paling kompleks. Lapisan *Batch Normalization* ditambahkan setelah blok konvolusi untuk menstabilkan proses pelatihan dan mempercepat konvergensi. Sementara itu, lapisan *Dropout* dengan *rate* yang umum digunakan (misalnya 0.25 atau 0.5) disertakan sebagai teknik regularisasi kunci untuk mengurangi *overfitting*, yang merupakan tantangan umum pada dataset citra medis yang terbatas. Secara keseluruhan, arsitektur ini dirancang untuk mencapai keseimbangan antara kapasitas model yang cukup untuk mempelajari pola kompleks dan penggunaan regularisasi yang efektif untuk memastikan generalisasi yang baik. Struktur lengkap arsitektur disajikan pada Tabel 2.

**Tabel 2. Arsitektur Custom CNN**

Type Layer	Bentuk Keluaran	Jumlah Parameter
conv2d (Conv2D)	(None, 64, 64, 32)	896
max_pooling2d (MaxPooling2D)	(None, 32, 32, 32)	0
conv2d_1 (Conv2D)	(None, 32, 32, 64)	18,496
max_pooling2d_1 (MaxPooling2D)	(None, 16, 16, 64)	0
conv2d_2 (Conv2D)	(None, 16, 16, 128)	73,856
max_pooling2d_2 (MaxPooling2D)	(None, 8, 8, 128)	0
batch_normalization (BatchNormalization)	(None, 8, 8, 128)	512
dropout (Dropout)	(None, 8, 8, 128)	0
flatten (Flatten)	(None, 8192)	0
dense (Dense)	(None, 128)	1,048,704
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 7)	903

Lapisan utama dari model terdiri atas beberapa blok Conv2D yang berfungsi mengekstraksi pola-pola visual penting dari citra, seperti tepi, tekstur, dan bentuk. Proses konvolusi dua dimensi pada setiap lapisan ini secara matematis dirumuskan sebagai berikut:

$$S(i, j) = (X * K)(i, j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(i+m, j+n) \cdot K(m, n) \quad (1)$$

Adapun penjelasan dari rumus (1) adalah sebagai berikut:

- $X$  : citra masukan (input image),
- $K$  : kernel atau filter,
- $S(i, j)$ : hasil konvolusi pada posisi  $(i, j)$ ,
- $M, N$ : ukuran filter (Younesi et al., 2024).

Rumus ini menunjukkan bahwa setiap posisi  $(i, j)$  pada hasil konvolusi merupakan hasil dari operasi perkalian elemen per elemen antara patch citra masukan dan kernel, lalu dijumlahkan. Setiap lapisan konvolusi dalam arsitektur diikuti oleh aktivasi *ReLU* untuk menangkap hubungan non-linear antar fitur. Arsitektur ini juga memanfaatkan beberapa lapisan fungsional yang tidak memiliki parameter untuk dilatih (parameter = 0) karena hanya melakukan operasi standar yang telah ditentukan. Lapisan-lapisan tersebut adalah *MaxPooling2D* yang berfungsi mengurangi dimensi spasial dan kompleksitas komputasi, serta Dropout yang digunakan sebagai regularisasi guna mencegah overfitting. Setelah lapisan Flatten mengubah output menjadi vektor satu dimensi, proses dilanjutkan ke lapisan Dense. Lapisan Batch Normalization juga ditambahkan untuk menstabilkan proses pelatihan. Lapisan output akhir memiliki tujuh neuron dengan aktivasi Softmax untuk mengklasifikasikan citra ke dalam tujuh kelas lesi kulit.

#### 2.4.2 Penyeimbangan Data dengan SMOTE-Tomek

Untuk mengatasi ketidakseimbangan kelas yang parah pada dataset HAM10000, penelitian ini menerapkan teknik *resampling* hibrida SMOTE-Tomek. Metode SMOTE-Tomek dipilih dibandingkan teknik *oversampling* canggih lainnya seperti ADASYN atau Borderline-SMOTE karena pendekatannya yang komprehensif. Sementara ADASYN dan Borderline-SMOTE berfokus pada sintesis data di sekitar batas keputusan atau sampel yang sulit, SMOTE-Tomek menggabungkan dua strategi: (1) *oversampling* umum melalui SMOTE untuk memperkaya representasi kelas minoritas secara keseluruhan, dan (2) *undersampling* pembersihan melalui Tomek Links untuk menghilangkan ambiguitas pada batas antar kelas. Kombinasi ini dianggap lebih robust untuk citra medis yang variasinya tinggi, karena tidak hanya menambah data minoritas tetapi juga secara aktif membersihkan *noise* dan tumpang tindih antar kelas,

sehingga menciptakan batas keputusan yang lebih jelas dan meningkatkan potensi generalisasi model.

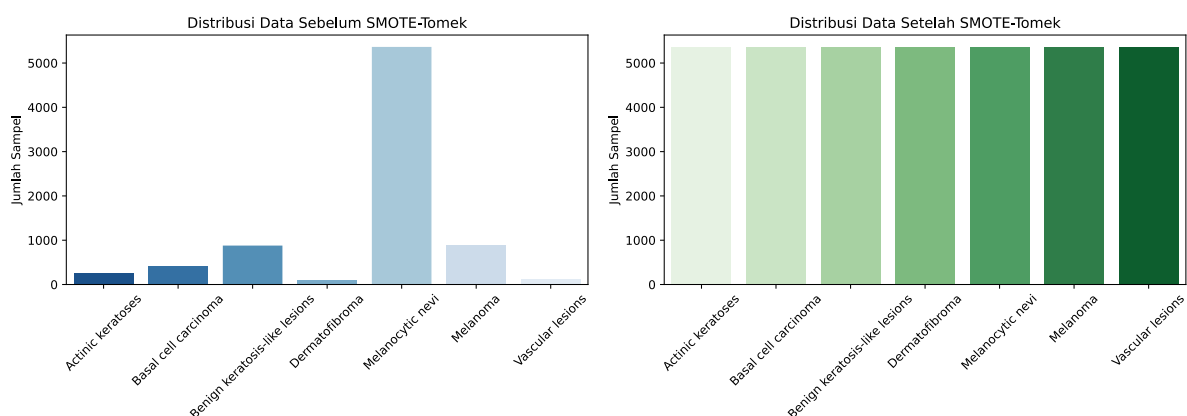
SMOTE bekerja dengan cara membuat sampel data sintetis baru untuk kelas minoritas melalui interpolasi, yang memperkaya representasi dan variasi data tanpa hanya melakukan duplikasi sederhana.

$$x_{\text{new}} = x_i + \delta \cdot (x_{\text{NN}} - x_i) \quad (2)$$

Adapun penjelasan dari rumus (2) adalah sebagai berikut:

- $x_i$  = sampel minoritas asli
- $x_{\text{NN}}$  = tetangga terdekat dari  $x_i$
- $\delta \in [0,1]$  = bilangan acak
- $x_{\text{new}}$  = sampel sintetis yang dihasilkan **(Yi et al., 2022)**.

Setelah proses sintesis oleh SMOTE, tahap berikutnya adalah penerapan Tomek Links untuk melakukan *undersampling* atau pembersihan data. Tomek Link didefinisikan sebagai pasangan sampel dari kelas yang berbeda, yang merupakan tetangga terdekat satu sama lain **(Aminullah & others, 2021)**. Kehadiran Tomek Links sering kali mengindikasikan potensi noise atau tumpang tindih pada batas antar kelas **(Nurhopipah & Magnolia, 2023)**. Oleh karena itu, dengan menghapus sampel dari kelas mayoritas dalam setiap Tomek Link yang teridentifikasi, metode ini membantu menciptakan batas keputusan yang lebih bersih dan terdefinisi dengan baik. Perlu ditekankan bahwa proses *undersampling* menggunakan Tomek Links ini bersifat adaptif; tidak ada persentase reduksi yang ditentukan sebelumnya untuk kelas mayoritas. Jumlah sampel mayoritas yang dihapus bergantung sepenuhnya pada berapa banyak pasangan Tomek Link yang teridentifikasi di batas antar kelas setelah proses SMOTE, sehingga pembersihan data lebih fokus pada area yang ambigu. Kombinasi kedua teknik ini menghasilkan distribusi data yang lebih seimbang dan memfasilitasi pelatihan model yang lebih akurat dan generalisasi yang lebih baik pada data nyata. Proses penyeimbangan data dengan SMOTE-Tomek ini diterapkan *setelah* dataset dibagi menjadi data latih dan data validasi. Teknik ini hanya diterapkan pada data latih untuk mencegah kebocoran data (*data leakage*), di mana informasi dari data sintetis dapat bocor ke dalam data validasi. Dengan demikian, data validasi tetap murni dan mencerminkan distribusi data asli, sehingga evaluasi performa model menjadi lebih objektif dan representatif terhadap data dunia nyata.



**Gambar 2. Distribusi jumlah sampel pada dataset HAM10000 sebelum dan sesudah dilakukan teknik SMOTE-Tomek.**

Gambar 2 menunjukkan pada grafik sebelah kiri terlihat distribusi data sebelum diterapkannya metode SMOTE-Tomek, di mana kelas Melanocytic Nevi mendominasi secara signifikan dengan lebih dari 5.000 sampel, sementara kelas minoritas seperti Dermatofibroma dan Vascular Lesions hanya memiliki kurang dari 200 sampel. Ketimpangan ini berisiko menyebabkan bias model terhadap kelas mayoritas. Setelah dilakukan penyeimbangan dengan SMOTE-Tomek, sebagaimana ditampilkan pada grafik kanan, seluruh kelas berhasil disamakan distribusinya. Hasil akhir dari proses ini adalah distribusi data yang seimbang sempurna, di mana setiap dari tujuh kelas kini menyumbang sekitar 14.3% dari total data latih, dengan jumlah sampel masing-masing kelas ditingkatkan menjadi sekitar 5.300. Ini menunjukkan bahwa proses sintesis dan pembersihan data berjalan efektif, menciptakan distribusi kelas yang setara sehingga meningkatkan potensi generalisasi model terhadap seluruh kelas secara adil.

## 2.5 Skenario Eksperimen dan Pelatihan Model

Untuk menemukan konfigurasi model yang paling optimal, penelitian ini dirancang dengan empat skenario eksperimen utama. Pendekatan yang digunakan adalah dengan mengubah satu variabel pada setiap skenario, sementara variabel lainnya diatur ke konfigurasi *default* atau konfigurasi terbaik dari skenario sebelumnya.

Langkah-langkah eksperimen yang dilakukan adalah sebagai berikut:

1. Skenario 1: Perbandingan Arsitektur.  
Skenario ini bertujuan untuk menguji apakah arsitektur Custom CNN yang ringan mampu bersaing dengan model *pre-trained* yang kompleks. Model Custom CNN dilatih dari awal (*from scratch*). Sebagai pembandingan, arsitektur EfficientNetB0 yang telah dilatih pada dataset ImageNet digunakan, di mana proses *fine-tuning* diterapkan pada 20 lapisan terakhirnya agar dapat beradaptasi dengan tugas klasifikasi kanker kulit.
2. Skenario 2: Perbandingan Metode Penyeimbangan Data.  
Skenario ini membandingkan efektivitas metode *resampling* SMOTE-Tomek dengan metode pembobotan kelas (*class weights*) dalam menangani masalah ketidakseimbangan data.
3. Skenario 3: Evaluasi Rasio Pembagian Data.  
Empat rasio data latih dan validasi (90:10, 80:20, 70:30, dan 60:40) dievaluasi untuk menemukan alokasi data yang paling optimal. Rentang rasio pembagian data ini dipilih untuk secara sistematis menginvestigasi *trade-off* antara jumlah data untuk melatih model dan jumlah data untuk validasi. Rasio 80:20 digunakan sebagai *baseline* standar, sementara rasio 90:10 diuji untuk melihat dampak dari memaksimalkan data latih. Sebaliknya, rasio 70:30 dan 60:40 dievaluasi untuk menguji apakah alokasi data validasi yang lebih besar memberikan estimasi performa yang lebih stabil.
4. Skenario 4: Perbandingan Optimizer.  
Enam algoritma optimisasi (Nadam, Adam, SGD, Adadelta, Adagrad, dan RMSprop) diuji untuk menentukan optimizer yang paling efektif dalam mencapai konvergensi dan performa tertinggi.

Semua model dilatih menggunakan TensorFlow dan Keras selama 15 *epoch* dengan *batch size* 32 dan fungsi *loss*. Fungsi *loss* yang digunakan untuk melatih model adalah *categorical cross-entropy*, yang secara spesifik dirancang untuk masalah klasifikasi multi-kelas di mana setiap sampel hanya dimiliki oleh satu kelas (*mutually exclusive*). Fungsi ini mengukur 'jarak' antara label ground truth (kelas yang benar) dengan distribusi probabilitas prediksi yang dihasilkan oleh lapisan *Softmax* pada model. Secara matematis, rumusnya adalah sebagai berikut:

$$H(p, q) = - \sum_{i=1}^C p(x_i) \log(q(x_i)) \quad (3)$$

Adapun penjelasan dari rumus (3) adalah sebagai berikut:

- $C$  = jumlah total kelas (dalam kasus ini, 7).
- $p(x_i)$  = probabilitas sebenarnya, yang bernilai 1 untuk kelas yang benar (ground truth) dan 0 untuk kelas lainnya (sesuai format one-hot vector).
- $q(x_i)$  = probabilitas prediksi yang dihasilkan oleh model untuk kelas ke- $i$ .

Tujuan dari proses pelatihan adalah untuk meminimalkan nilai *loss* ini. Dengan meminimalkan *categorical cross-entropy* model 'dipaksa' untuk menghasilkan vektor probabilitas prediksi () yang sedekat mungkin dengan vektor *ground truth* (). Artinya, makna dari probabilitas prediksi yang dihasilkan adalah tingkat kepercayaan (*confidence*) model, di mana nilai probabilitas akan didorong mendekati 100% untuk kelas yang diprediksi benar dan mendekati 0% untuk kelas-kelas lainnya.

## 2.6 Metrik Evaluasi

Kinerja model dalam setiap skenario eksperimen dievaluasi menggunakan metrik standar dalam klasifikasi *machine learning*. Metrik-metrik ini memberikan pandangan komprehensif tentang kemampuan model, terutama dalam konteks medis di mana kesalahan klasifikasi memiliki konsekuensi yang berbeda. Berikut perbandingan rumus (4), (5), (6), dan (7) yang digunakan :

$$\text{Akurasi} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}} \quad (4)$$

$$\text{Presisi} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (5)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (6)$$

$$F1 = 2 \cdot \frac{\text{Presisi} \cdot \text{Recall}}{\text{Presisi} + \text{Recall}} \quad (7)$$

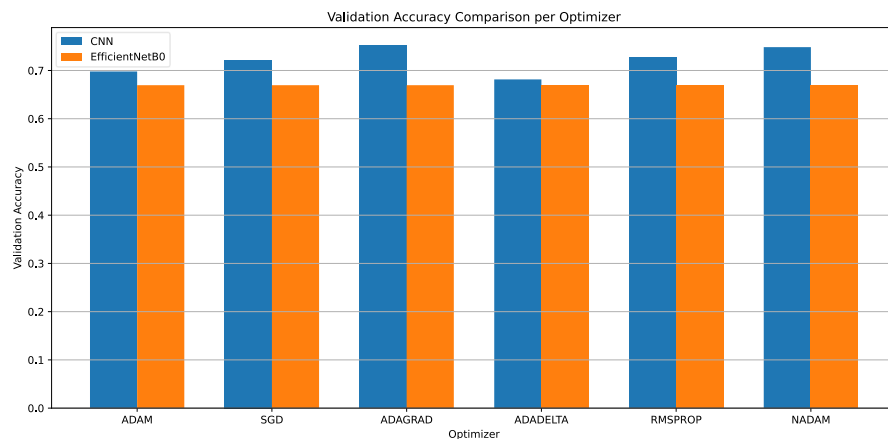
Setiap metrik evaluasi memiliki fungsi dan interpretasi yang berbeda. Akurasi mengukur kebenaran prediksi secara keseluruhan namun bisa menyesatkan pada data tidak seimbang. Presisi berfokus pada minimalisasi *false positive*—penting untuk menghindari salah diagnosis—sementara Recall (sensitivitas) krusial untuk menekan *false negative* agar tidak ada kasus aktual yang terlewat. Terakhir, F1-score menjadi penyeimbang antara presisi dan recall, yang sangat berguna untuk dataset tidak seimbang karena nilainya hanya akan tinggi jika kedua metrik tersebut juga tinggi (Al Abrori & Subhiyakto, 2025).

## 3. HASIL DAN PEMBAHASAN

### 3.1 Perbandingan Arsitektur Model : Custom CNN vs EfficientNet80

Eksperimen pertama dilakukan untuk membandingkan performa model Custom CNN yang dibangun dan dilatih dari awal dengan arsitektur *pre-trained EfficientNet80* yang telah melalui proses *fine-tuning* pada 20 lapisan terakhir. Tujuan dari eksperimen ini adalah untuk menguji apakah arsitektur ringan dan sederhana yang dilatih dari awal dapat memberikan hasil yang kompetitif dibandingkan model kompleks yang telah melalui *pre-training* pada dataset skala besar seperti *ImageNet*.



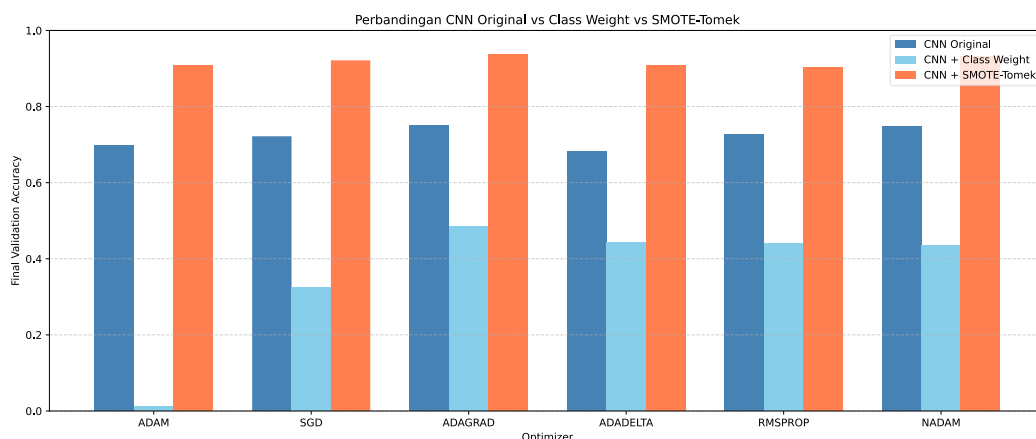


**Gambar 3. Perbandingan Custom CNN & EfficientNetB0**

Gambar 3 menunjukkan bahwa model Custom CNN secara konsisten menghasilkan akurasi validasi lebih tinggi dibandingkan *EfficientNetB0* di semua konfigurasi *optimizer* (Adam, SGD, Adagrad, Adadelata, RMSprop, dan Nadam), dengan performa terbaik dicapai saat menggunakan Nadam ( $\sim 0.76$ ). Sebaliknya, *EfficientNetB0* cenderung stagnan di kisaran  $\sim 0.67$ – $0.68$ , meskipun telah dilakukan *fine-tuning* pada 20 layer terakhir. Hasil ini mengindikasikan bahwa arsitektur ringan seperti Custom CNN, yang fleksibel dan mudah dilatih dari awal, mampu menangkap pola relevan dalam citra dermoskopi secara lebih efektif. Dengan kompleksitas rendah dan kinerja yang kompetitif, Custom CNN menjadi solusi yang lebih praktis untuk sistem deteksi kanker kulit berbasis citra, terutama di lingkungan dengan sumber daya terbatas.

### 3.2 Perbandingan Metode Penyeimbangan Data: SMOTE-Tomek vs *Class Weights*

Eksperimen kedua bertujuan untuk mengevaluasi pengaruh teknik penyeimbangan data terhadap performa klasifikasi model CNN, dengan membandingkan metode SMOTE-Tomek dan *Class Weights*, serta hasil *baseline* dari model CNN tanpa penyeimbangan (original). Penyeimbangan data menjadi aspek krusial dalam penelitian ini mengingat ketimpangan kelas yang signifikan pada dataset HAM10000, di mana kelas minoritas seperti "Melanoma" dan "Dermatofibroma" jauh lebih sedikit dibanding kelas mayoritas seperti "Melanocytic Nevi".



**Gambar 4. Perbandingan CNN Original, CNN+Class Weight, dan CNN+SMOTE-Tomek**

Berdasarkan Gambar 4, metode SMOTE-Tomek secara konsisten memberikan akurasi validasi tertinggi ( $\sim 0.89$ – $0.93$ ) di seluruh jenis *optimizer*, melampaui model original ( $\sim 0.68$ – $0.75$ ) dan

*Class Weights* ( $\sim 0.30\text{--}0.45$ ). Selain itu, dilakukan analisis juga terhadap F1-Score, Presisi, dan Recall berdasarkan *macro average* dari tiap model dengan hasil sebagai berikut:

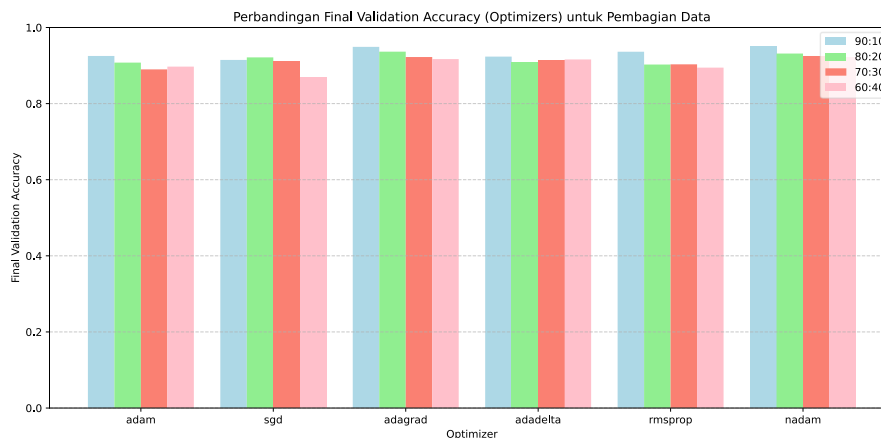
**Tabel 3. Classification Report Metode Penyeimbangan Data**

Model	Presisi	Recall	F1-Score
CNN Original	0.53	0.38	0.40
CNN+Class Weight	0.17	0.20	0.16
CNN+SMOTE-Tomek	0.44	0.56	0.44

Berdasarkan Tabel 3, perbandingan tiga varian model CNN menunjukkan bahwa penerapan metode penyeimbangan data SMOTE-Tomek memberikan hasil paling optimal. Model CNN+SMOTE-Tomek berhasil mencapai F1-Score tertinggi sebesar 0.44, dengan peningkatan signifikan pada nilai Recall menjadi 0.56. Kinerja ini melampaui model CNN Original yang memiliki F1-Score 0.40 dan jauh lebih unggul dibandingkan model CNN+Class Weight yang justru mengalami penurunan performa drastis dengan F1-Score hanya 0.16. Dengan demikian, dapat disimpulkan bahwa SMOTE-Tomek adalah metode yang paling efektif untuk meningkatkan keseimbangan performa model, terutama dalam kemampuannya mengidentifikasi kelas positif. Perbandingan ini dilakukan dalam kondisi eksperimen yang terkontrol dan adil, di mana arsitektur model, optimizer, jumlah epoch, dan rasio pembagian data dijaga tetap sama untuk ketiga skenario pengujian. Hal ini memastikan bahwa perbedaan performa yang diamati murni disebabkan oleh pengaruh dari metode penyeimbangan data yang digunakan.

### 3.3 Perbandingan Rasio Data

Eksperimen ketiga bertujuan mengevaluasi pengaruh variasi rasio pembagian data latih dan validasi terhadap performa model CNN. Empat konfigurasi yang diuji adalah 90:10, 80:20, 70:30, dan 60:40, untuk menentukan rasio paling optimal dalam menghasilkan akurasi validasi tertinggi sekaligus menjaga kemampuan generalisasi model terhadap data baru.



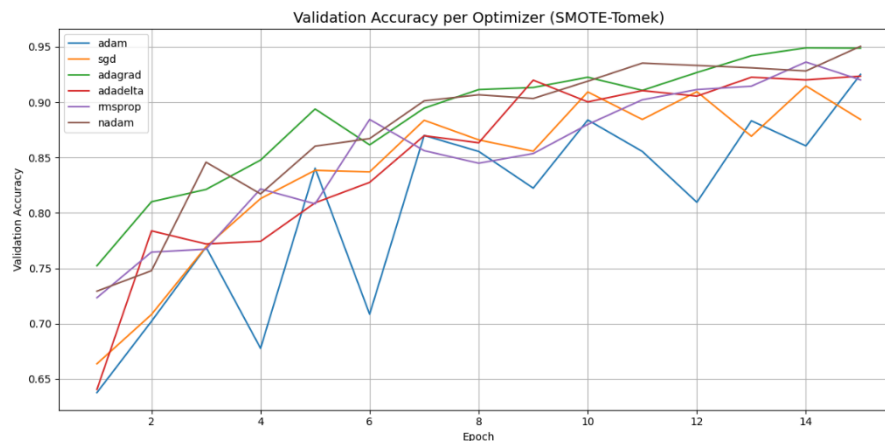
**Gambar 5. Perbandingan Rasio Data**

Gambar 5 menunjukkan bahwa rasio 90:10 konsisten menghasilkan akurasi validasi tertinggi ( $\sim 0.93\text{--}0.96$ ) di hampir semua *optimizer*, mengungguli rasio lainnya. Meskipun proporsi data validasi lebih kecil, jumlah data latih yang lebih besar mampu memperkaya proses pembelajaran dan meningkatkan generalisasi tanpa *overfitting*. Sebaliknya, rasio 60:40 menurunkan akurasi akibat keterbatasan data latih. Temuan ini menegaskan pentingnya pemilihan rasio data yang optimal, khususnya dalam klasifikasi citra medis, dan mendukung

penggunaan rasio 90:10 sebagai konfigurasi akhir terbaik untuk model Custom CNN pada dataset HAM10000.

### 3.4 Perbandingan Optimizer

Eksperimen keempat bertujuan menganalisis pengaruh berbagai algoritma optimizer terhadap performa akhir model Custom CNN yang telah diseimbangkan dengan SMOTE-Tomek. Analisis dilakukan berdasarkan hasil validasi dari eksperimen ketiga dengan rasio data 90:10, tanpa pelatihan ulang. Fokus utamanya adalah mengevaluasi kontribusi masing-masing *optimizer* dalam mempercepat dan mengarahkan konvergensi model menuju performa optimal.

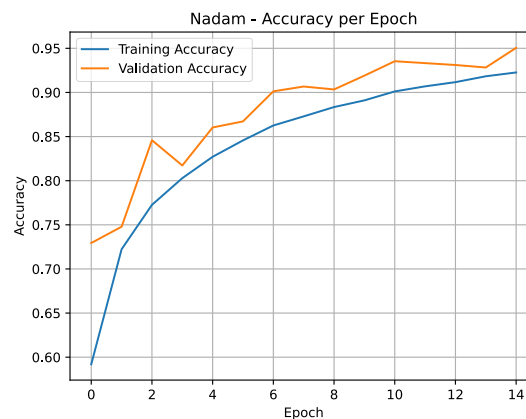


**Gambar 6. Perbandingan tiap *optimizer* dari epoch 1 sampai epoch 15**

Berdasarkan analisis pada Gambar 6, terlihat karakteristik yang berbeda dari setiap optimizer. Dari segi kecepatan konvergensi, Nadam dan Adagrad terbukti paling cepat, karena keduanya mampu mencapai akurasi validasi di atas 0.85 dalam kurang dari 6 *epoch*. Dalam hal stabilitas, Nadam juga menunjukkan kurva yang paling konsisten menaik, menjadikannya yang paling stabil. Sebaliknya, Adam dan SGD bersifat lebih fluktuatif, dengan perubahan akurasi yang drastis antar *epoch*, terutama pada Adam yang mengalami penurunan tajam setelah mencapai puncaknya. Performa yang menurun di *epoch-epoch* akhir pada optimizer seperti Adam dan SGD juga dapat mengindikasikan adanya potensi *overfitting*, di mana model mulai kehilangan kemampuan generalisasinya setelah mencapai performa puncak. Secara keseluruhan, analisis ini menunjukkan bahwa Nadam adalah *optimizer* yang paling unggul karena kombinasi kecepatan konvergensi, stabilitas, dan performa puncak yang tinggi.

### 3.5 Kinerja Model Optimal

Setelah melalui serangkaian eksperimen komparatif, konfigurasi model terbaik yang diperoleh dalam penelitian ini adalah arsitektur CNN kustom yang dilatih menggunakan optimizer Nadam, rasio pembagian data latih dan validasi sebesar 90:10, serta strategi penyeimbangan data SMOTE-Tomek. Kombinasi ini terbukti menghasilkan performa yang paling unggul di antara konfigurasi lainnya. Untuk memvalidasi performa model secara menyeluruh, dilakukan analisis mendalam terhadap metrik pelatihan dan pengujian, termasuk grafik akurasi selama epoch pelatihan serta classification report yang memuat nilai presisi, recall, dan F1-score untuk masing-masing kelas. Evaluasi ini memberikan gambaran yang komprehensif terhadap kemampuan generalisasi model dalam mengklasifikasikan berbagai jenis lesi kulit secara akurat.



**Gambar 7. Grafik akurasi per epoch oleh model terbaik**

Gambar 7 menunjukkan tren peningkatan akurasi selama proses pelatihan model. *Training accuracy* meningkat secara stabil dari ~0.58 pada epoch pertama hingga >0.92 pada epoch ke-15. Sementara itu, *validation accuracy* juga mengalami lonjakan signifikan sejak awal pelatihan, dari ~0.73 pada epoch 1 hingga menyentuh nilai maksimum ~0.95 di akhir epoch. Meskipun demikian, disadari bahwa metrik akurasi saja dapat memberikan gambaran yang kurang lengkap, terutama pada dataset yang tidak seimbang di mana performa pada kelas minoritas sangat krusial. Oleh karena itu, untuk evaluasi performa yang lebih robust dan detail per-kelas, analisis merujuk pada metrik Presisi, Recall, dan F1-Score yang disajikan secara komprehensif pada Tabel 4 (Classification Report). Tabel tersebut menyajikan bukti kuantitatif kemampuan model dalam mengenali setiap kelas, termasuk kelas-kelas kritis seperti Melanoma, secara adil dan seimbang.

**Tabel 4. Classification Report oleh model terbaik**

Nama Kelas	Presisi	Recall	F1-Score
Melanocytic Nevi	0.94	0.85	0.90
Melanoma	0.89	0.92	0.90
Benign Keratosis-like Lesions	0.91	0.95	0.93
Basal Cell Carcinoma	0.98	0.99	0.98
Actinic Keratoses	0.98	0.99	0.98
Vascular Lesions	1.00	1.00	1.00
Dermatofibroma	0.99	1.00	1.00
(macro average)	0.956	0.957	0.953
(weighted average)	0.956	0.957	0.953
(micro average)	0.956	0.956	0.956

Terakhir, Tabel 4 menyajikan evaluasi performa model klasifikasi kanker kulit berdasarkan presisi, recall, dan F1-score untuk tiap kelas. Secara keseluruhan, model menunjukkan performa sangat baik dengan F1-score  $\geq 0.90$  di hampir semua kelas. Kelas Melanocytic Nevi mencatat F1-score 0.90 dengan presisi 0.94 dan recall 0.85, menunjukkan adanya beberapa kasus yang belum terdeteksi secara optimal. Untuk kelas Melanoma, model mencatat presisi 0.89 dan recall 0.92, menandakan kemampuan tinggi dalam mengenali kasus kanker kulit paling mematikan tersebut. Kinerja terbaik ditemukan pada kelas minoritas seperti Vascular Lesions dan Dermatofibroma, dengan nilai presisi, recall, dan F1-score mendekati atau sama dengan 1.00. Kelas lain seperti Benign Keratosis-like Lesions, Basal Cell Carcinoma, dan Actinic

Keratosis juga menunjukkan *F1-score* tinggi (0.93–0.98), menegaskan keandalan model dalam membedakan jenis lesi kulit yang beragam.

Perhatian khusus diberikan pada kelas minoritas seperti *Vascular Lesions* dan *Dermatofibroma* yang mencapai F1-Score sempurna (1.00), sebuah hasil yang memerlukan interpretasi hati-hati. Setelah memastikan bahwa performa ini diuji pada data uji independen yang tidak terlibat dalam proses SMOTE (sehingga menepis kemungkinan *data leakage*), penjelasan yang paling masuk akal adalah kombinasi dari jumlah sampel uji yang sangat terbatas untuk kelas-kelas ini dan kemungkinan adanya fitur visual yang sangat khas yang memudahkan model untuk mempelajarinya. Oleh karena itu, hasil ini diinterpretasikan sebagai keberhasilan model pada dataset spesifik ini, dan validasi lebih lanjut pada dataset eksternal disarankan sebagai arah penelitian di masa depan untuk mengonfirmasi temuan ini.

Secara keseluruhan, kombinasi arsitektur CNN yang fleksibel, strategi penyeimbangan data SMOTE-Tomek, serta optimasi menggunakan algoritma Nadam telah menghasilkan model klasifikasi kanker kulit berbasis citra dermoskopi yang sangat akurat dan andal. Model ini terbukti mampu mengenali pola visual kompleks dengan tingkat kesalahan minimal, menjadikannya kandidat yang kuat untuk implementasi dalam sistem pendukung keputusan klinis di bidang dermatologi.

**Tabel 5. Perbandingan Penelitian Terdahulu dengan Metode yang Diusulkan**

Author	Arsitektur Utama	Model Ringan	Penanganan Imbalance	Optimasi	Evaluasi Lengkap	Implementasi Aplikasi
(Patil et al., 2022)	CNN Sederhana	✓	✗	Adam	✗	✓
(Swapno et al., 2025)	EfficientNetV2L + LightGBM	✗	Augmentasi Data	Adam	✓	✗
(Wei et al., 2020)	Ensemble CNN Ringan	✓	Augmentasi Data	Adam	✗	✗
Metode yang diusulkan	Custom CNN	✓	SMOTE-Tomek	Nadam	✓	✓

Berdasarkan perbandingan pada Tabel 5, terlihat bahwa metode yang diusulkan dalam penelitian ini merupakan pendekatan yang paling komprehensif dibandingkan dengan tiga penelitian terdahulu. Penelitian oleh (Patil et al., 2022) dan (Wei et al., 2020) sama-sama menggunakan arsitektur model yang ringan, namun tidak secara eksplisit menangani ketidakseimbangan data, hanya menggunakan augmentasi tanpa strategi resampling, serta tidak menyertakan evaluasi metrik yang lengkap maupun bukti implementasi dalam aplikasi nyata. Sementara itu, (Swapno et al., 2025) memang menyajikan evaluasi metrik yang lengkap dan menangani ketidakseimbangan data dengan augmentasi, tetapi menggunakan arsitektur kompleks (*EfficientNetV2L + LightGBM*) yang tidak ringan dan tidak dilengkapi dengan implementasi aplikasi.

Metode yang diusulkan dalam penelitian ini menjadi satu-satunya yang memenuhi semua aspek penting secara simultan: menggunakan arsitektur Custom CNN yang ringan, menangani ketidakseimbangan data secara efektif dengan SMOTE-Tomek, melakukan optimasi dengan algoritma Nadam yang terbukti paling stabil dan akurat, menyajikan evaluasi metrik yang lengkap (akurasi, presisi, *recall*, *F1-score*), serta mengimplementasikan model ke dalam aplikasi web fungsional. Kombinasi seluruh pendekatan ini tidak hanya menghasilkan akurasi validasi yang tinggi (95.63%), tetapi juga membuktikan bahwa model yang dikembangkan siap digunakan dalam skenario nyata sebagai alat bantu klasifikasi otomatis kanker kulit berbasis citra dermoskopi.

### 3.6 Implementasi Aplikasi

Untuk mendemonstrasikan fungsionalitas dan kelayakan implementasi dari model CNN terbaik, sebuah purwarupa aplikasi klasifikasi citra berbasis web telah dikembangkan. Aplikasi ini mampu menerima input gambar dermoskopi dari pengguna dan mengembalikan hasil klasifikasi berupa nama kelas beserta nilai *confidence score*. Sebagai studi kasus fungsional, dilakukan uji coba menggunakan beberapa gambar sampel yang representatif dari setiap kelas lesi kulit. Penting untuk ditekankan bahwa tujuan dari pengujian ini adalah untuk menunjukkan kelayakan teknis aplikasi, bukan sebagai validasi performa statistik yang komprehensif.

**Tabel 6. Gambar input yang digunakan dalam pengujian**

Gambar	Nama File Gambar	Label Asli	Label Lengkap
	ISIC_0024693	nv	Melanocytic Nevi
	ISIC_0028215	mel	Melanoma
	ISIC_0027815	bkl	Benign Keratosis-like Lesions
	ISIC_0028155	bcc	Basal Cell Carcinoma
	ISIC_0029417	akiec	Actinic Keratoses
	ISIC_0031197	vasc	Vascular Lesions
	ISIC_0027008	df	Dermatofibroma

Sebagai bagian dari pengujian implementatif, dilakukan uji coba terhadap sejumlah gambar representatif dari masing-masing kelas. Penting untuk ditekankan bahwa semua gambar uji ini diambil dari *set data validasi* yang telah dipisahkan di awal dan tidak pernah digunakan dalam proses pelatihan, augmentasi, maupun SMOTE, sehingga merupakan data yang benar-benar baru (*unseen data*) bagi model. Hasil prediksi yang dihasilkan oleh aplikasi mencakup

label yang ditetapkan oleh model serta *confidence score* dalam bentuk persentase, dan seluruhnya disajikan secara rinci dalam Tabel 7.

**Tabel 7. Hasil Prediksi dan *Confidence Score***

<b>Nama File Gambar</b>	<b>Label Prediksi</b>	<b><i>Confidence Score</i></b>
ISIC_0024693	Melanocytic Nevi	98.96%
ISIC_0028215	Melanoma	93.70%
ISIC_0027815	Benign Keratosis-like Lesions	99.99%
ISIC_0028155	Basal Cell Carcinoma	97.18%
ISIC_0029417	Actinic Keratoses	99.97%
ISIC_0031197	Vascular Lesions	100.00%
ISIC_0027008	Dermatofibroma	99.88%

Hasil implementasi pada Tabel 7 menunjukkan bahwa model berhasil mengklasifikasikan berbagai jenis lesi kulit dengan akurasi tinggi, ditunjukkan oleh *confidence score* di atas 90% pada semua gambar uji. Tidak terdapat kesalahan klasifikasi, termasuk pada kelas minoritas seperti Dermatofibroma dan Vascular Lesions. Keberhasilan ini membuktikan bahwa model CNN yang dikembangkan tidak hanya efektif secara eksperimen, tetapi juga layak diterapkan dalam sistem pendukung keputusan medis berbasis aplikasi.

#### 4. KESIMPULAN

Penelitian ini mengembangkan model klasifikasi kanker kulit berbasis citra dermoskopi menggunakan arsitektur Custom CNN yang ringan namun efektif. Optimalisasi dilakukan melalui teknik SMOTE-Tomek, rasio data latih-validasi 90:10, dan *optimizer* Nadam. Dataset HAM10000 yang digunakan mencakup 10.015 citra dari tujuh kelas lesi kulit, termasuk kelas minoritas seperti Melanoma dan Dermatofibroma yang menantang untuk diklasifikasi secara adil. Model yang diusulkan mencapai akurasi validasi 95,63% dengan presisi, *recall*, dan *F1-score*  $\geq 0,90$  pada hampir seluruh kelas, melampaui model pre-trained seperti *EfficientNetB0*. Selain itu, model berhasil diimplementasikan dalam aplikasi web yang mampu melakukan klasifikasi citra secara cepat dan akurat dengan *confidence score* di atas 90%, membuktikan kelayakannya untuk digunakan dalam skenario nyata. Meskipun demikian, penelitian ini memiliki beberapa keterbatasan, yaitu: (1) validasi yang terbatas pada satu dataset tanpa *k-fold cross-validation*; (2) tidak adanya uji signifikansi statistik pada perbandingan *optimizer*; (3) analisis proses pelatihan yang hanya berbasis akurasi; serta (4) kurangnya validasi klinis dari ahli medis terhadap hasil prediksi model. Oleh karena itu, penelitian selanjutnya sangat disarankan untuk fokus pada validasi eksternal, penerapan *k-fold cross-validation* untuk evaluasi yang lebih robust, dan kolaborasi klinis dengan ahli medis guna memastikan generalisasi dan kelayakan model di dunia nyata.

#### DAFTAR RUJUKAN

- Al Abrori, Z. Z. H., & Subhiyakto, E. R. (2025). Analisis Komparatif Akurasi Prediksi Kanker Payudara Menggunakan Algoritma Random Forest dan Logistic Regression. *Jurnal Algoritma*, 22(1), 300–311.
- Albahar, M. A. (2019). Skin Lesion Classification Using Convolutional Neural Network with Novel Regularizer. *IEEE Access*, 7(March 2019), 38306–38313. <https://doi.org/10.1109/ACCESS.2019.2906241>

- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. In *Journal of Big Data* (Vol. 8, Issue 1). Springer International Publishing. <https://doi.org/10.1186/s40537-021-00444-8>
- Amartama, S. N., Hidayah, A. N., Sari, P. K., & Ramadhani, R. A. (2024). Implementasi Convolutional Neural Network (CNN) dalam Pengenalan Pola Penulisan Tangan. *Seminar Nasional Teknologi & Sains*, 3(1), 133–138. <https://doi.org/10.29407/stains.v3i1.4155>
- Aminullah, M., & others. (2021). *Perbandingan performa klasifikasi machine learning dengan teknik resampling pada dataset tidak seimbang*. Fakultas Sains dan Teknologi UIN Syarif Hidayatullah Jakarta.
- AYENI, J. A. (2022). Convolutional Neural Network (CNN): The architecture and applications. *Applied Journal of Physical Science*, 4(4), 42–50. <https://doi.org/10.31248/ajps2022.085>
- Jaworek-Korjakowska, J., Kleczek, P., & Gorgon, M. (2019). Melanoma thickness prediction based on convolutional neural network with VGG-19 model transfer learning. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2019-June*(June), 2748–2756. <https://doi.org/10.1109/CVPRW.2019.00333>
- Kadhim, Y. A., Khan, M. U., & Mishra, A. (2022). *Deep Learning-Based Computer-Aided Diagnosis (CAD): Applications for Medical Image Datasets*.
- Mader, K. S. (2018). *Skin Cancer MNIST: HAM10000*. <https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000>
- Nurhopipah, A., & Magnolia, C. (2023). Perbandingan metode Resampling pada Imbalanced Dataset untuk klasifikasi komentar program MBKM. *Jurnal Publikasi Ilmu Komputer Dan Multimedia*, 2(1), 9–22.
- Pane, S. F., Abdullah, F., & Habibi, R. (2024). Deteksi Emosi Pada Teks Berbahasa Indonesia Menggunakan Pendekatan Ensemble. *JTT (Jurnal Teknologi Terapan)*, 10(2), 80. <https://doi.org/10.31884/jtt.v10i2.551>
- Pane, S. F., & Ramdan, J. (2022). Pemodelan Machine Learning : Analisis Sentimen Masyarakat Terhadap Kebijakan PPKM Menggunakan Data Twitter. *Jurnal Sistem Cerdas*, 5(1), 12–20. <https://doi.org/10.37396/jsc.v5i1.191>
- Patil, S. M., Rajguru, B. S., Mahadik, R. S., & Pawar, O. P. (2022). Melanoma Skin Cancer Disease Detection Using Convolutional Neural Network. *2022 3rd International Conference for Emerging Technology, INCET 2022, July 2022*. <https://doi.org/10.1109/INCET54531.2022.9825381>



- Putrada, A. G., Alamsyah, N., Pane, S. F., Fauzan, M. N., & Perdana, D. (2023). Knowledge Distillation for a Lightweight Deep Learning-Based Indoor Positioning System on Edge Environments. *2023 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 370–375. <https://doi.org/10.1109/ISITIA59021.2023.10220996>
- Radha Jayalakshmi, P., & Rajakumar, P. S. (2024). A Survey: Recent Advances and Clinical Applications of Deep Learning in Medical Image Analysis. *Deep Learning in Medical Image Analysis: Recent Advances and Future Trends*, 59–78. <https://doi.org/10.1201/9781003343172-4>
- Shetty, B., Fernandes, R., Rodrigues, A. P., Chengoden, R., Bhattacharya, S., & Lakshmana, K. (2022). Skin lesion classification of dermoscopic images using machine learning and convolutional neural network. *Scientific Reports*, 12(1), 1–11. <https://doi.org/10.1038/s41598-022-22644-9>
- Swapno, S. M. R., Nobel, S. M. N., Meena, P. K., Meena, V. P., Bahadur, J., & Appaji, A. (2025). Accelerated and precise skin cancer detection through an enhanced machine learning pipeline for improved diagnostic accuracy. *Results in Engineering*, 25(January), 104168. <https://doi.org/10.1016/j.rineng.2025.104168>
- Tajerian, A., Kazemian, M., Tajerian, M., & Malayeri, A. A. (2023). Design and validation of a new machine-learning-based diagnostic tool for the differentiation of dermatoscopic skin cancer images. *PLoS ONE*, 18(4 April), 1–17. <https://doi.org/10.1371/journal.pone.0284437>
- Wang, M., Gao, X., & Zhang, L. (2025). Recent global patterns in skin cancer incidence, mortality, and prevalence. *Chinese Medical Journal*, 138(2), 185–192.
- Wei, L., Ding, K., & Hu, H. (2020). Automatic Skin Cancer Detection in Dermoscopy Images Based on Ensemble Lightweight Deep Learning Network. *IEEE Access*, 8, 99633–99647. <https://doi.org/10.1109/ACCESS.2020.2997710>
- Yi, X., Xu, Y., Hu, Q., Krishnamoorthy, S., Li, W., & Tang, Z. (2022). ASN-SMOTE: a synthetic minority oversampling method with adaptive qualified synthesizer selection. *Complex and Intelligent Systems*, 8(3), 2247–2272. <https://doi.org/10.1007/s40747-021-00638-w>
- Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M., & Henkel, J. (2024). A Comprehensive Survey of Convolutions in Deep Learning: Applications, Challenges, and Future Trends. *IEEE Access*, 12, 41180–41218. <https://doi.org/10.1109/ACCESS.2024.3376441>