

PERBANDINGAN KINERJA ALGORITMA TRANSFORMER DAN CONVOLUTIONAL NEURAL NETWORK DALAM DETEKSI DAN SEGMENTASI KANKER KULIT

Jennifer Averina Harjono¹⁾, Haura Putry Yasha¹⁾, Adrian Richardy Kurniawan¹⁾, Irmawati¹⁾*

¹⁾Sistem Informasi, Universitas Multimedia Nusantara, Tangerang, Banten, Indonesia

*Penulis korespondensi: irmawati@umn.ac.id

ABSTRAK

Kanker kulit merupakan salah satu jenis kanker yang paling umum di dunia, dengan jumlah kasus yang terus meningkat setiap tahunnya. Deteksi dini melalui segmentasi citra medis memainkan peran penting dalam meningkatkan akurasi diagnosis. Penelitian ini bertujuan untuk membandingkan kinerja model berbasis Transformer seperti Vision Transformer (ViT), Swin Transformer, dan Data-efficient Image Transformer (DeiT) dalam tugas deteksi, serta segmentasi kanker kulit, dengan arsitektur berbasis Convolutional Neural Network (CNN), yaitu U-Net, menggunakan dataset ISIC 2017. Evaluasi dilakukan menggunakan metrik sensitivitas, spesifisitas, AUC, akurasi, dan Dice Score. Hasil menunjukkan bahwa DeiT memperoleh nilai AUC (0.92) dan sensitivitas (0.90) tertinggi, menjadikannya model yang menjanjikan untuk skrining awal. Swin unggul dalam spesifisitas (0.97), sementara ViT menunjukkan kinerja yang seimbang. Pada tugas segmentasi, model U-Net berhasil meningkatkan Dice Score dari 0.83 menjadi 0.95. Temuan ini diharapkan dapat menjadi dasar pengembangan sistem diagnosis berbasis AI yang lebih efisien dan akurat, khususnya untuk perangkat cerdas dan platform medis otomatis.

Kata kunci: kanker kulit, deteksi, segmentasi, transformer, convolutional neural network

ABSTRACT

Skin cancer is one of the most common types of cancer worldwide, with an increasing number of cases each year. Early detection through medical image segmentation plays a crucial role in improving diagnostic accuracy. This study aims to compare the performance of Transformer-based models such as Vision Transformer (ViT), Swin Transformer, and Data-efficient Image Transformer (DeiT) in the detection, also segmenting skin cancer using Convolutional Neural Network (CNN)-based architecture, namely U-Net, with the ISIC 2017 dataset. The evaluation was conducted using metrics including sensitivity, specificity, AUC, accuracy, and Dice Score. The results show that DeiT achieved the highest AUC (0.92) and sensitivity (0.90), making it a promising model for early screening. Swin outperformed others in specificity (0.97), while ViT offered a balanced performance. In the segmentation task, the U-Net model improved the Dice Score from 0.83 to 0.95. These findings are expected to serve as a foundation for the development of more efficient and accurate AI-based diagnostic systems, particularly for smart devices and automated medical platforms.

Keywords: skin cancer, detection, segmentation, transformer, convolutional neural network

Pendahuluan

Kanker kulit merupakan salah satu jenis kanker yang paling umum di seluruh dunia, dengan jumlah kasus yang terus meningkat secara signifikan setiap tahunnya (American Cancer Society, 2025). Kemajuan dalam teknologi medis dan kecerdasan buatan (AI) telah mendorong perkembangan signifikan dalam deteksi dan pengobatannya. Jenis utama kanker kulit meliputi karsinoma sel basal (BCC), karsinoma sel skuamosa (SCC), dan melanoma, dengan melanoma sebagai jenis paling berbahaya karena tingkat penyebarannya yang tinggi. Pada tahun 2022, tercatat sekitar 331.722 kasus baru melanoma kulit secara global, menjadikannya kanker ke-17 paling umum (*Skin cancer statistics | World Cancer Research Fund*, 2025). Sekitar 90% kasus kanker kulit disebabkan oleh paparan radiasi ultraviolet (UV), baik dari matahari maupun sumber buatan, sehingga termasuk jenis kanker yang dapat dicegah (Meyer and Stockfleth, 2021). Faktor risiko lain mencakup riwayat keluarga, sistem imun yang lemah, dan paparan bahan kimia berbahaya (Wunderlich, Suppa and Gandini, 2024).

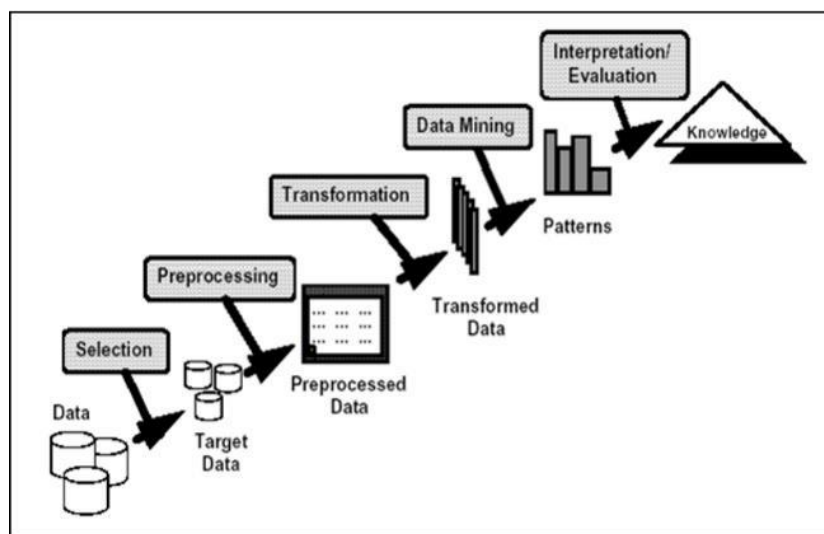
Deteksi dini sangat penting dalam membedakan kanker kulit dari kelainan kulit lainnya untuk mencegah keterlambatan penanganan (Gusti, Haerani and Syafria, 2024). Saat ini, diagnosis masih banyak mengandalkan pemeriksaan visual atau biopsi, yang membutuhkan waktu dan biaya (Serenit, Laksana and Nasution, 2024). Untuk mengatasi keterbatasan ini, teknologi mulai digunakan dalam dunia medis, terutama dalam analisis gambar kulit (Rasyid, Khotimah and Suaedah, 2025). Dalam konteks ini, *Knowledge Discovery in Databases* (KDD) berperan penting melalui teknik *data mining* untuk menemukan pola tersembunyi dalam data besar untuk meningkatkan akurasi dan efisiensi deteksi (Nofianti, Yawan and Nazar, 2023). KDD mencakup tahapan mulai dari pemilihan data hingga evaluasi hasil. Seiring kemajuan deep learning, model berbasis CNN dan transformer semakin banyak digunakan. Pada penelitian ini, digunakan algoritma U-Net serta ViT, Swin, dan DeiT untuk mendukung tugas segmentasi kanker kulit.

Penelitian ini akan membandingkan tiga model, yaitu ViT, DeiT, dan Swin yang berbasis transformer, dengan demikian dapat diketahui model mana yang dapat melakukan tugasnya dalam mendeteksi kanker kulit menggunakan dataset ISIC 2017 dan implementasi U-NET yang berbasis CNN dalam segmentasi kanker kulit. Beberapa jurnal yang akan menjadi acuan untuk dibandingkan dengan penelitian ini adalah penelitian yang dilakukan oleh (Shajimon, Ufumaka and Raza, 2023) yang mengimplementasikan ViT-B16 dengan *layer normalization* dan *focal loss*, serta mencapai akurasi 86% dan AUC 0.83 pada klasifikasi melanoma. Selain itu, penelitian oleh (Xie, 2021) mengusulkan model Swin-SimAM yang menggabungkan Swin Transformer dan modul perhatian SimAM, dengan hasil AUC sebesar 0.90. Kedua penelitian ini menjadi tolak ukur penting untuk

mengevaluasi performa model dalam penelitian ini. Sebagai pengembangan dari penelitian-penelitian sebelumnya yang mengevaluasi ViT dan Swin, penelitian ini turut menyertakan DeiT untuk memperluas analisis kinerja model Vision Transformer. Selanjutnya untuk segmentasi dengan menggunakan model U-Net yang menjadi acuan pembandingan yaitu penelitian yang dilakukan oleh (Cheng., 2024), Penelitian ini membuat pengembangan arsitektur baru bernama MSGU-Net, yang merupakan varian ringan dan efisien dari U-Net untuk tugas segmentasi citra medis, model U-Net pada penelitian ini mendapatkan *Dice Score* sebesar 82.53, yang akan dioptimalisasikan dalam penelitian ini.

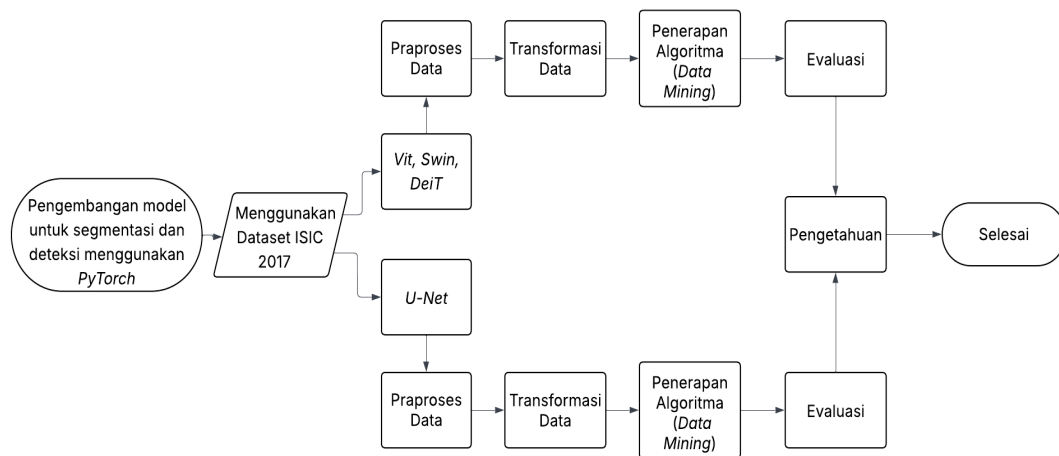
Metode

Didalam penelitian ini, peneliti menggunakan metode KDD. KDD merupakan pendekatan yang umum digunakan untuk menemukan pola tersembunyi dalam kumpulan data berukuran besar (Zai, 2022). Gambar metode KDD ditampilkan pada Gambar 1.



Gambar 1. Proses *Knowledge Discovery in Databases (KDD)*

Gambar 1 menampilkan proses bahwa KDD terdiri dari beberapa tahapan utama, dimulai dari pemilihan data yang relevan, dilanjutkan dengan tahap prapemrosesan untuk membersihkan dan menyiapkan data, kemudian transformasi data ke dalam format yang sesuai untuk analisis, proses *data mining* untuk menemukan pola yang signifikan, dan diakhiri dengan evaluasi terhadap hasil yang diperoleh. *Flowchart* alur penelitian dari penelitian ini ditampilkan pada Gambar 2.



Gambar 2. Alur Penelitian

Gambar 2 menggambarkan alur pengembangan model untuk segmentasi dan deteksi menggunakan *PyTorch* dengan dataset ISIC 2017. Proses dimulai dengan penggunaan dataset yang kemudian dibagi ke dalam dua bagian pengembangan model, yaitu menggunakan model ViT, Swin, DeiT di bagian atas dan U-Net di bagian bawah. Setiap di bagian menjalani tahapan praproses data, transformasi data, penerapan algoritma *data mining*, dan evaluasi. Hasil evaluasi kemudian menghasilkan pengetahuan yang menjadi keluaran akhir dari proses ini, yang kemudian menandai selesai atau selesainya pengembangan model. Proses ini menunjukkan pendekatan yang sistematis dan berulang untuk mengembangkan dan mengevaluasi model segmentasi dan deteksi. Rincian alur penelitian yang ditunjukkan pada Gambar 3 akan dijelaskan secara terperinci dalam Tabel 1.

Tabel 1. Perincian Tahapan Alur Penelitian

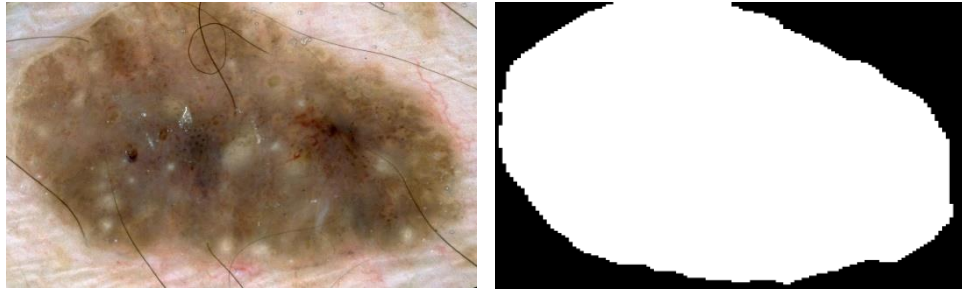
Tahapan	Aktivitas	Deskripsi Aktivitas	
		U-Net	ViT, Swin dan DeiT
Seleksi Data	Menseleksi data sesuai dengan kebutuhan penelitian.	Menggunakan Dataset ISIC 2017.	Menggunakan Dataset ISIC 2017.
Praproses Data	Praproses Data yang akan digunakan didalam penelitian.	Melakukan <i>resize</i> data menjadi 224x224 piksel.	1. Melakukan <i>resize</i> data menjadi 224x224 piksel. 2. Augmentasi data.

Transformasi Data	Melakukan Transformasi data.	1. Data mask diproses menjadi tensor biner (0 atau 1). 2. Optimasi <i>Hyperparameter</i>	1. Konversi ke Tensor 2. Normalisasi Data 3. Optimasi <i>Hyperparameter</i>
Penerapan Algoritma (<i>Data Mining</i>)	Melatih, memvalidasi dan uji coba model.	Melatih, memvalidasi dan uji coba model <i>U-Net</i> .	Melatih, memvalidasi dan uji coba model <i>ViT</i> , <i>Swin</i> dan <i>DeiT</i> .
Evaluasi	Melakukan evaluasi dari model yang telah dikembangkan.	Akurasi dan <i>Dice score</i>	1. <i>Specificity</i> , Sensitifitas, dan <i>Area Under the Curve (AUC)</i> 2. <i>Confusion Matrix</i>
Pengetahuan	Menganalisis hasil evaluasi dari model.	1. Menganalisis hasil evaluasi 2. Menyusun laporan ilmiah dan artikel ilmiah	1. Menganalisis hasil evaluasi 2. Menyusun laporan ilmiah dan artikel ilmiah

Tabel 1 menjelaskan perbedaan langkah-langkah dalam alur penelitian antara model U-Net dan model ViT, Swin, serta DeiT terlihat jelas pada beberapa tahapan. Gambaran menyeluruh mengenai pendekatan yang digunakan dalam pengembangan masing-masing model, yang sudah disesuaikan dengan karakteristik dan kebutuhan spesifik dari setiap arsitektur, mulai dari seleksi data, praproses, transformasi data, hingga evaluasi dan analisis pengetahuan adalah sebagai berikut.

A. Seleksi Data

Penelitian ini menggunakan dataset sekunder dari International Skin Imaging, kumpulan data citra kulit yang cukup terkenal di kalangan peneliti medis dan teknologi. Dataset ini berisi dua kelas utama, yaitu melanoma yang merupakan jenis kanker kulit berbahaya, dan seborrheic keratosis yaitu kondisi kulit non-kanker yang biasanya muncul seiring bertambahnya usia. Contoh gambar citra kulit yang ada di dalam dataset ISIC 2017 ditampilkan pada Gambar 3.



Gambar 3. Image and Masking

Gambar 3 merupakan contoh data yang digunakan dalam mengembangkan model *U-Net*. Data ini dipilih karena memiliki variasi gambar dan cukup representatif untuk dijadikan bahan analisis dalam pengembangan model segmentasi penyakit kulit berbasis *U-Net*. Sedangkan dalam pengembangan model ViT, Swin dan DeiT, menggunakan cara yang berbeda, yaitu dengan memisahkan data menjadi Melanoma dan Seborrheic Keratosis, gambar masking yang ada didalam tidak digunakan dalam proses pengembangan ViT, Swin dan DeiT. Dengan demikian, melalui penggunaan dataset ini, peneliti dapat melakukan diferensiasi antara dua kondisi kulit yang secara visual memiliki kemiripan tetapi memiliki implikasi klinis yang sangat berbeda. Rincian data yang telah di pisahkan secara manual ada pada Tabel 2.

Tabel 2. Pembagian *Dataset*

Jenis Data	Melanoma	Seborrheic Keratosis	Total Gambar
Train	327	242	569
Test	53	39	92
Validation	141	105	246
Total			907

Tabel 2 memberikan rincian dari data yang digunakan dalam pengembangan model ViT, Swin dan DeiT. yang meliputi jumlah gambar untuk kategori Melanoma dan Seborrheic Keratosis pada tahap pelatihan, pengujian, dan validasi, dengan total keseluruhan sebanyak 907 gambar. Dengan menggunakan data yang berbeda, diharapkan setiap model dapat dioptimalkan sesuai dengan karakteristik metode yang digunakan, sehingga model U-Net dapat lebih efektif dalam melakukan segmentasi berbasis mask, sedangkan model ViT, Swin, dan DeiT dapat lebih akurat dalam melakukan klasifikasi diferensial antara Melanoma dan Seborrheic Keratosis. Pendekatan ini memungkinkan peneliti untuk memperoleh hasil yang lebih komprehensif dan spesifik dalam mendiagnosis serta membedakan kondisi kulit yang secara visual serupa namun memiliki implikasi klinis yang berbeda.

B. Praproses Data

Pada model U-Net dilakukan *resize* data menjadi 224x224 piksel. Standarisasi ukuran input agar konsisten dan sesuai dengan arsitektur model untuk proses pelatihan yang optimal, sedangkan ViT, Swin, DeiT melakukan *resize* data menjadi 224x224 piksel serta tambahan augmentasi data untuk menstandarisasi ukuran input dan memperkaya variasi data pelatihan, sehingga model dapat belajar dari berbagai variasi, meningkatkan generalisasi, dan mengurangi risiko *overfitting*.

C. Transformasi Data

Model U-Net, ViT, Swin, dan DeiT menggunakan *hyperparameter* yang berbeda untuk mengoptimalkan pekerjaannya masing-masing, sehingga *Optimization Space* yang digunakan berbeda. Rincian dari perbedaan hyperparameter tersebut ada pada Tabel 3.

Tabel 3. Optimasi *Hyperparameter* U-Net, ViT, Swin, DeiT

<i>Hyper-parameter</i>	<i>Optimization Space</i>			
	U-Net	ViT	Swin	DeiT
Epoch	100	100	100	100
Batch Size	64	10	16	10
Learning Rate	0.0001	0.0001	0.0001	0.0001
Optimizer	Adam	Adam	Adam	AdamW
Patience	5	10	10	5
Loss Function	CrossEntropy	CrossEntropy	CrossEntropy	CrossEntropy

Tabel 3 menjelaskan optimasi hyperparameter yang digunakan pada beberapa model deep learning, yaitu U-Net, ViT, Swin Transformer, dan DeiT. Tabel 3 menampilkan ruang pencarian atau nilai hyperparameter yang diatur selama proses pelatihan untuk masing-masing model. Parameter seperti jumlah epoch yang sama-sama ditetapkan sebanyak 100 kali iterasi pada semua model, batch size yang berbeda-beda mulai dari 10 hingga 64, serta learning rate yang seragam sebesar 0,0001. Optimizer yang digunakan juga bervariasi, dimana U-Net, ViT, dan Swin memakai Adam, sedangkan DeiT menggunakan AdamW. Selain itu, nilai patience yang mengatur berapa lama proses pelatihan berhenti ketika tidak terjadi peningkatan performa diatur antara 5 sampai 10 epoch, dan fungsi loss yang digunakan pada semua model adalah CrossEntropyLoss. Tabel ini memberikan gambaran tentang konfigurasi penting dalam tuning hyperparameter untuk memperoleh performa optimal dari masing-masing arsitektur model yang diuji.

D. Evaluasi

Model U-Net mengevaluasi menggunakan metrik akurasi dan *Dice score*, untuk menilai kualitas segmentasi dengan mengukur kesesuaian prediksi terhadap data *ground truth* secara spasial, sedangkan ViT, Swin, DeiT mengevaluasi menggunakan metrik spesitifitas, sensitivitas, *Area Under the Curve* (AUC), dan *confusion matrix* untuk memberikan gambaran lebih lengkap mengenai performa model dalam hal deteksi dan klasifikasi, termasuk keseimbangan antara true positive dan true negative serta kemampuan membedakan kelas dengan lebih baik. Dengan perhitungan seperti:

$$Akurasi = (TP + TN)/(TP + FP + FN + TN) \quad (1)$$

$$Sensitifitas = TP/(TP + FN) \quad (2)$$

$$Spesifisitas = TN/(TN + FP) \quad (3)$$

$$AUC = Sensitifitas + Spesifisitas/2 \quad (4)$$

Rumus tersebut digunakan untuk menilai kinerja model. Akurasi mengukur prediksi benar secara keseluruhan, sensitivitas fokus pada deteksi kasus positif, spesifisitas pada kasus negatif, dan AUC mencerminkan kemampuan model membedakan dua kelas secara umum (Istifarinta, Prajitno and Soejoko, 2022).

E. Pengetahuan

Hasil evaluasi dari model U-Net digunakan untuk memahami kelebihan dan kekurangan model dalam membedakan jenis penyakit kulit, seperti melanoma dan seborrheic keratosis. Sementara itu, evaluasi terhadap model ViT, Swin, dan DeiT memberikan pemahaman mengenai keunggulan arsitektur transformer dalam melakukan segmentasi citra medis secara lebih adaptif. Kedua pengetahuan yang diperoleh menjadi dasar dalam penyusunan laporan dan artikel ilmiah.

Hasil dan Pembahasan

Penelitian ini bertujuan untuk mengkaji dan membandingkan kinerja tiga arsitektur model yaitu ViT, Swin, dan DeiT, dalam mendeteksi kanker kulit serta mengevaluasi kinerja U-Net CNN dalam segmentasi kanker kulit menggunakan dataset ISIC 2017. Evaluasi dilakukan dengan mengacu pada empat metrik utama: AUC, Spesitifitas, Sensitifitas, dan Akurasi. Hasil pengujian model pada data validasi dan data uji ditampilkan pada Tabel 4.

Tabel 4. Hasil *Testing* Model Transformer Dalam Segmentasi Kanker Kulit

Model	AUC	Spesitifitas	Sensitifitas	Akurasi
ViT	0.91	0.91	0.79	0.86
Swin	0.91	0.97	0.80	0.90
DeiT	0.92	0.79	0.90	0.84

Tabel 4 menunjukkan bahwa setiap model memiliki keunggulan masing-masing. ViT mencatat AUC sebesar 0.91 dengan specificity 0.91 dan sensitivity 0.79, menunjukkan performa yang seimbang. Swin memiliki akurasi tertinggi (0.8984) dan specificity paling tinggi (0.97), sehingga lebih unggul dalam mengurangi false positive. Sementara itu, DeiT menonjol dalam sensitivity (0.90) dan AUC tertinggi (0.92), yang membuatnya lebih tepat digunakan untuk skrining awal karena mampu mendeteksi lebih banyak kasus positif. Pemilihan model perlu mempertimbangkan trade-off antara sensitivity dan specificity yang ditawarkan masing-masing model. Selanjutnya perbandingan hasil pengujian model ViT dan Swin dalam penelitian ini dengan model sejenis dari studi terdahulu yang dianggap sebagai *state-of-the-art* pada tugas deteksi kanker kulit yang ditampilkan pada Tabel 5.

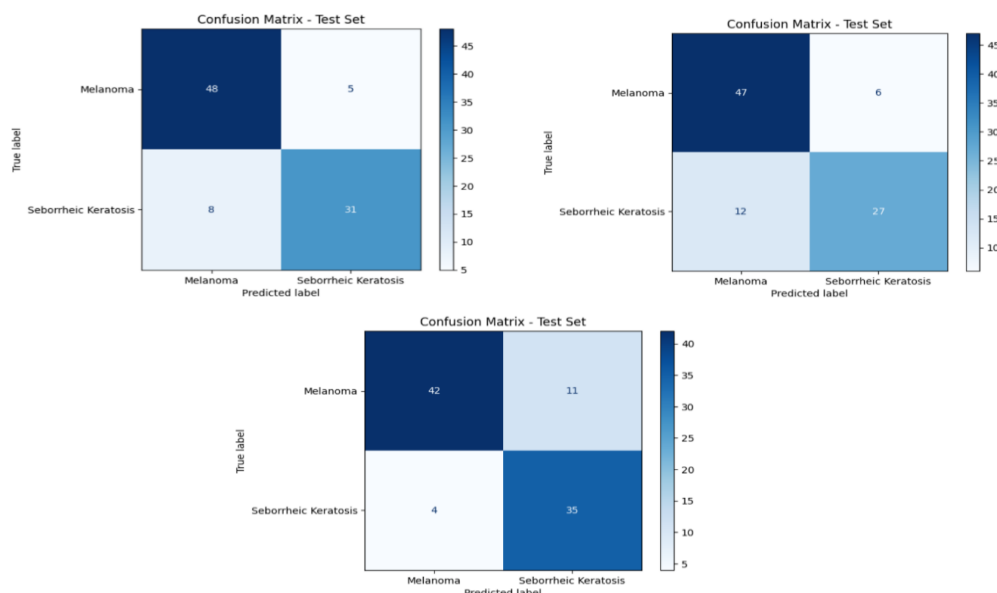
Tabel 5. Hasil testing yang dibandingkan berdasarkan *state-of-the-art*

Model	AUC	Specificity	Sensitivity	Akurasi
ViT	0.83	0.37	0.77	0.86
ViT	0.91	0.91	0.79	0.86
Swin	0.90	-	-	-
Swin	0.91	0.97	0.80	0.90

Tabel 5 menunjukkan bahwa model ViT dan Swin dalam penelitian ini memiliki performa yang lebih baik dibandingkan model sejenis dari studi terdahulu dimana teks yang ditebalkan merupakan penelitian saya sendiri. ViT versi penelitian ini mencapai AUC 0.91 dengan spesifisitas 0.91 dan sensitivitas 0.79, lebih unggul dari ViT milik Shajimon et al. (2023) yang hanya memiliki AUC 0.83 dan spesifisitas 0.37. Sementara itu, Swin juga menunjukkan peningkatan, dengan AUC 0.91 dan spesifisitas tinggi 0.97, melampaui model Swin milik Khan dan Khan (2023) yang hanya melaporkan AUC 0.90 tanpa metrik lain. Hal ini menegaskan bahwa model dalam studi ini lebih andal dan seimbang dalam klasifikasi kanker kulit. Sebagai pengembangan dari penelitian-penelitian sebelumnya yang hanya mengevaluasi ViT dan Swin, penelitian ini turut menyertakan DeiT untuk memperluas analisis kinerja model Vision Transformer.

Hasilnya menunjukkan bahwa DeiT memiliki nilai AUC dan sensitivity tertinggi, sehingga potensial digunakan pada tahap skrining awal untuk mendeteksi kanker kulit secara lebih akurat. Hal ini menunjukkan bahwa DeiT dapat menjadi alternatif model yang kompetitif dan relevan untuk diaplikasikan dalam sistem deteksi berbasis deep learning.

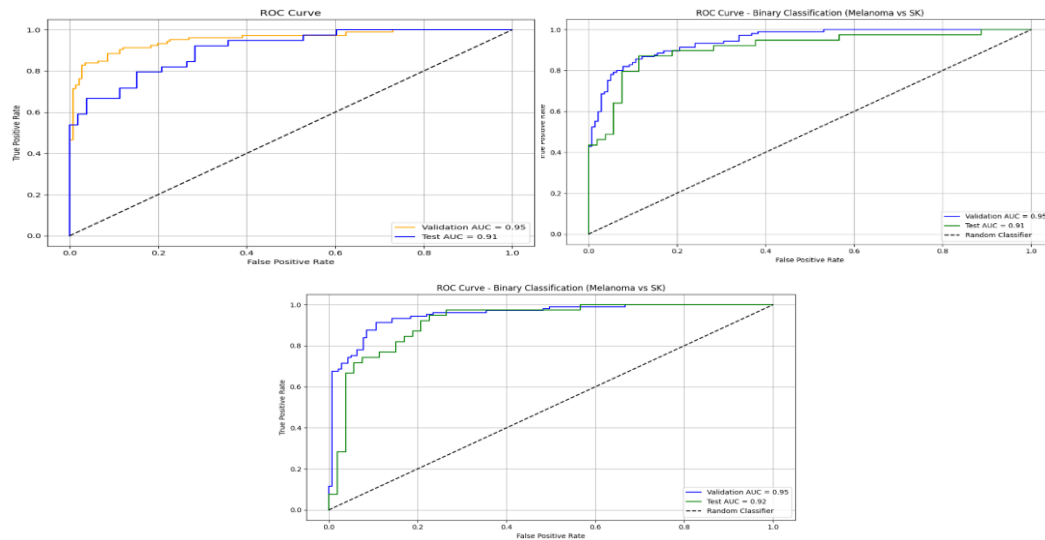
Evaluasi klasifikasi dilakukan menggunakan confusion matrix untuk mengetahui tingkat akurasi, sensitivitas, dan spesifisitas dari model ViT, Swin, dan DeiT sebagaimana ditunjukkan pada Gambar 4.



Gambar 4. Hasil Confusion Matrix dari ViT, Swin dan DeiT

Gambar 4 menunjukkan hasil confusion matrix dari tiga model deteksi lesi kulit, yaitu ViT, Swin, dan DeiT. Model ViT memiliki nilai true negative yang cukup baik ($TN = 31$), menunjukkan kemampuannya dalam mengenali Seborrheic Keratosis dengan benar serta false negative yang rendah ($FN = 5$), menandakan sensitivitas tinggi terhadap Melanoma. Sementara itu, model Swin menunjukkan peningkatan false positive ($FP = 12$), yang berarti lebih banyak kasus non-Melanoma diklasifikasikan salah sebagai Melanoma. Adapun model DeiT memiliki true negative tertinggi ($TN = 35$), menunjukkan spesifisitas yang kuat, namun jumlah false negative-nya meningkat ($FN = 11$), sehingga beberapa kasus Melanoma tidak terdeteksi.

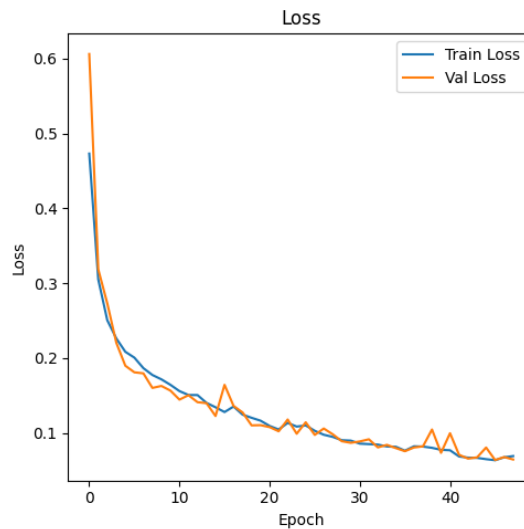
Evaluasi performa model dengan kurva Receiver Operating Characteristic (ROC) digunakan untuk mengukur kemampuan model dalam membedakan dua kelas secara menyeluruh untuk memperlihatkan performa tiga model ViT, Swin, dan DeiT dalam mengklasifikasikan melanoma dan seborrheic keratosis yang ditampilkan pada Gambar 5.



Gambar 5. ROC Curve (Melanoma vs Seborrheic Keratosis) dengan model ViT, Swin, DeiT

Gambar 5 menampilkan hasil ROC Curve dari ketiga model transformer ViT, Swin, dan DeiT menunjukkan performa klasifikasi yang sangat baik dalam membedakan antara melanoma dan seborrheic keratosis pada dataset ISIC 2017. Ketiga model menghasilkan nilai AUC validasi yang sama yaitu 0,95, namun terdapat perbedaan kecil pada hasil AUC pengujian, di mana ViT dan Swin mencatat nilai 0,91, sedangkan DeiT sedikit lebih unggul dengan AUC sebesar 0,92. Hal ini mengindikasikan bahwa DeiT memiliki kemampuan generalisasi yang lebih baik terhadap data baru dibanding dua model lainnya.

Selain pendekatan berbasis Vision Transformer, penelitian ini juga menguji performa model U-Net sebagai baseline dalam tugas segmentasi kanker kulit. Evaluasi dilakukan berdasarkan dua metrik utama, yaitu akurasi dan Dice Score. Selama proses pelatihan hingga tahap pengujian, model menunjukkan nilai *train loss* dan *validation loss* yang stabil dan cukup baik. Visualisasi kurva *train loss* dan *validation loss* ditampilkan pada Gambar 6.



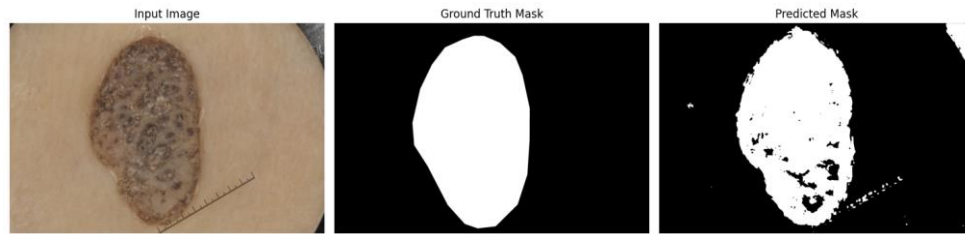
Gambar 6. *Train dan Test Loss*

Gambar 6 terlihat bahwa nilai *train loss* dan *validation loss* menurun secara konsisten seiring bertambahnya jumlah epoch. Penurunan ini menunjukkan bahwa model berhasil belajar dari data pelatihan dan mampu melakukan generalisasi dengan baik terhadap data validasi. Selain itu, tidak terlihat adanya indikasi *overfitting*, karena kedua kurva loss menunjukkan pola yang serupa dan tetap stabil hingga akhir pelatihan. Hal ini mengindikasikan bahwa model memiliki performa pelatihan yang baik dan layak untuk dilanjutkan ke tahap pengujian. Hasil akurasi dan *Dice Score* ditampilkan pada Tabel 6.

Tabel 6. Perbandingan Hasil Testing Penelitian Sebelumnya Dengan Model Usulan

Model	Akurasi	<i>Dice-Score</i>
U-Net	-	0.83
U-Net	0.90	0.95

Tabel 6 menunjukkan performa model U-Net dalam segmentasi kanker kulit, tulisan yang ditebalkan merupakan hasil dari model U-Net di penelitian ini. Pada data *testing*, U-Net mencatat akurasi sebesar 89.61 dan *Dice Score* sebesar 0.95. sedangkan penelitian yang menjadi acuan dari model U-Net yaitu adalah penelitian dari (Cheng H., Zhang, Y. and Xu., 2024), yang menghasilkan *Dice Score* sebesar 0.83. Hasil ini mengindikasikan bahwa model U-Net dalam penelitian ini dapat mengoptimalkan *Dice Score* yang lebih baik daripada penelitian terdahulunya. Nilai ini tergolong tinggi dan menunjukkan bahwa U-Net mampu melakukan generalisasi dengan cukup baik pada data yang belum pernah dilihat sebelumnya. Hasil dari prediksi model ditampilkan dalam Gambar 7.



Gambar 7. Hasil Prediksi Model

Gambar 7 dapat diamati bahwa model telah berhasil mengidentifikasi area utama dari objek yang menjadi fokus segmentasi. Hasil prediksi menunjukkan bahwa model mampu mengenali bentuk global dari lesi dengan cukup baik, yang menandakan bahwa model telah mempelajari representasi fitur secara efektif. Meskipun masih terdapat beberapa prediksi yang tidak sempurna, seperti kemunculan noise di area sekitar objek dan ketidakteraturan pada bagian dalam, hal ini merupakan bagian dari proses iteratif dalam pengembangan model segmentasi. Secara keseluruhan, hasil ini menunjukkan potensi yang menjanjikan dari model dalam melakukan segmentasi objek secara otomatis dan dapat ditingkatkan lebih lanjut melalui optimalisasi parameter, atau peningkatan kualitas data pelatihan.

Kesimpulan

Penelitian ini berhasil membandingkan tiga model yang berbasis transformer, yaitu ViT, Swin, dan DeiT, dalam tugas deteksi kanker kulit. Hasilnya menunjukkan bahwa masing-masing model memiliki keunggulan tersendiri, di mana DeiT memperoleh nilai AUC dan sensitivitas tertinggi, menjadikannya paling potensial untuk digunakan dalam tahap skrining awal. Swin unggul dalam spesifisitas, yang menunjukkan kemampuannya dalam meminimalkan prediksi positif palsu, sementara ViT menunjukkan performa yang seimbang antara sensitivitas dan spesifisitas. Pendekatan berbasis Transformer mampu memberikan hasil deteksi yang kompetitif. Pada segmentasi kanker kulit, model U-Net yang di optimasi mendapatkan nilai *Dice Score* sebesar 82.53 dan meningkat secara signifikan menjadi 94.51. Untuk penelitian selanjutnya, akan dilakukan penggabungan antara arsitektur CNN dan Transformer dalam pendekatan hybrid untuk memperoleh hasil yang lebih optimal.

Ucapan Terima Kasih

Para penulis mengucapkan syukur kepada Tuhan Yang Maha Esa atas berkat dan rahmat-Nya sehingga penelitian ini dapat diselesaikan. Ucapan terima kasih juga disampaikan kepada Universitas Multimedia Nusantara dan Ibu Dr. Irmawati, S.Kom., M.M.S.I. atas bimbingan dan dukungannya selama proses penelitian.

Kontribusi Penulis

Jennifer Averina Harjono dan Haura Putry Yasha berkontribusi dalam pengumpulan serta analisis data. Adrian Richardy Kurniawan pembuatan dan evaluasi model. Bimbingan dalam proses penelitian dan penulisan artikel diberikan oleh dosen pendamping.

Daftar Pustaka

- American Cancer Society (2025) ‘Cancer Facts and Figures’. American Cancer Society.
- Cheng, H., Zhang, Y. and Xu, H. (2024) ‘MSGU-Net: a lightweight multi-scale ghost U-Net for image segmentation’, *Frontiers in Neurorobotics*, 18. Cheng, H., Zhang, Y. and Xu, H.
- Gusti, G.P.H.P., Haerani, E. and Syafria, F. (2024) ‘Implementasi Algoritma Convolutional Neural Network (Resnet-50) untuk Klasifikasi Kanker Kulit Benign dan Malignant’, *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(3), pp. 984–992. Gusti, G.P.H.P., Haerani, E. and Syafria, F.
- Istifarinta, O.L., Prajitno, P. and Soejoko, D.S. (2022) ‘Computer-Aided Detection (CAD) Deteksi Nodul Paru-Paru dari Computed Tomography (CT)’, *INTERVENTIONAL Journal of Medical Physics and Biophysics*, 9(1), pp. 1–8. Istifarinta, O.L., Prajitno, P. and Soejoko, D.S.
- Meyer, T. and Stockfleth, E. (2021) ‘Light and Skin’, *Current Problems in Dermatology (Switzerland)*, 55, pp. 53–61. Meyer, T. and Stockfleth, E.
- Nofianti, A., Yawan, M.Y. and Nazar, M.A. (2023) ‘Implementasi Data Mining dalam Pengolahan Data Transaksi Toko Sembako Menggunakan Algoritma Apriori (Studi Kasus : Toko Devan Mart)’, *G-Tech: Jurnal Teknologi Terapan*, 7(1), pp. 165–173. Nofianti, A., Yawan, M.Y. and Nazar, M.A.
- Rasyid, M.R.A.A.R., Khotimah, K. and Suaedah, S. (2025) ‘Sistem Pakar Pra-Diagnosa Penyakit Kulit dengan Metode Forward Chaining dan CNN dan Sains Data’. Rasyid, M.R.A.A.R., Khotimah, K. and Suaedah, S.
- Serenit, C., Laksana, F.D. and Nasution, R.S. (2024) ‘Deteksi Kanker Kulit Menggunakan Algoritma Convolutional Neural Network’. Serenit, C., Laksana, F.D. and Nasution, R.S.
- Shajimon, G.M., Ufumaka, I. and Raza, H. (2023) ‘An Improved Vision-Transformer Network for Skin Cancer Classification’, *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 2213–2216. Shajimon, G.M., Ufumaka, I. and Raza, H.
- Skin cancer statistics | World Cancer Research Fund* (2025).
- Wunderlich, K., Suppa, M. and Gandini, S. (2024) ‘Risk Factors and Innovations in Risk Assessment for Melanoma, Basal Cell Carcinoma, and Squamous Cell Carcinoma’, *Cancers 2024, Vol. 16, Page 1016*, 16(5), p. 1016. Wunderlich, K., Suppa, M. and Gandini, S.

Xie, J. (2021) ‘Melanoma Detection based on Swin Transformer and SimAM’, *IEEE Information Technology, Networking, Electronic and Automation Control Conference, ITNEC 2021*, pp. 1517–1521. Xie, J.