

Segmentasi Citra Dermoskopi Kanker Kulit Menggunakan Metode VGG-SegNet

DOI: <http://dx.doi.org/10.35889/progresif.v22i1.3464>

Creative Commons License 4.0 (CC BY –NC)



Muhammad Rafi' Ramadhani^{1*}, Chastine Fatichah²

Magister Teknik Informatika, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

*e-mail *Corresponding Author*: rafiramadhani1999@gmail.com

Abstract

Skin cancer, particularly melanoma, has a high mortality rate, necessitating reliable early detection based on dermoscopic images. Accurate lesion segmentation is a crucial preprocessing step prior to deep learning-based classification. However, challenges remain in preserving lesion boundary details, handling variations in lesion size, and addressing limited data availability and hyperparameter optimization. This study proposes and evaluates various configurations of a hybrid VGG–SegNet model for skin lesion segmentation using the ISIC 2017 and ISIC 2018 datasets. The methodology includes image and ground truth validation, VGG16-based normalization, input image resolution variations (128×128, 160×160, and 256×256), and training–validation data splits of 50:50 and 70:30. The model is fine-tuned using skip connections and a multi-stage training scheme with a combination of Binary Cross-Entropy, Dice, and Focal Tversky Loss. The best-performing model achieves a Dice Coefficient of 0.899 and an Intersection over Union of 0.8177 on the validation set, demonstrating precise and efficient lesion segmentation.

Keywords: Skin Cancer Segmentation; VGG-SegNet; ISIC; Fine Tuning

Abstrak

Kanker kulit, khususnya melanoma, memiliki tingkat mortalitas tinggi sehingga memerlukan deteksi dini berbasis citra dermoskopi yang andal. Segmentasi lesi yang akurat merupakan tahap penting sebelum proses klasifikasi berbasis *deep learning*. Namun, tantangan masih muncul pada detail tepi lesi, variasi ukuran, serta keterbatasan data dan optimasi *hyperparameter*. Penelitian ini mengusulkan dan mengevaluasi berbagai konfigurasi model hibrida VGG–SegNet untuk segmentasi lesi kulit menggunakan dataset ISIC 2017 dan ISIC 2018. Tahapan meliputi validasi citra dan *ground truth*, normalisasi berbasis VGG16, variasi resolusi citra (128×128, 160×160, 256×256), serta pembagian data latih dan validasi dengan rasio 50:50 dan 70:30. Model di *fine tuning* menggunakan *skip connection* dan skema pelatihan bertahap dengan kombinasi *Binary Cross-Entropy*, *Dice*, dan *Focal Tversky Loss*. Model terbaik mencapai *Dice Coefficient* 0,899 dan *IoU* 0,8177 pada data validasi, menunjukkan segmentasi yang presisi dan efisien.

Kata kunci: Segmentasi Kanker Kulit; VGG–SegNet; ISIC; Fine Tuning

1. Pendahuluan

Kanker kulit merupakan kanker yang umum ditemui diseluruh dunia. Salah satu varian dari kanker kulit yaitu melanoma, menyebabkan sebagian besar kematian terkait kanker kulit meskipun jumlah kasusnya lebih sedikit jika dibandingkan dengan kanker kulit lainnya. Kasus *melanoma* secara global terus meningkat setiap tahun, dengan perkiraan terdapat lebih dari 150.000 kasus baru dan sekitar 48.000 kematian setiap tahun secara global. Kanker kulit yang paling banyak ditemui di Indonesia adalah karsinoma sel basal (65%), kemudian karsinoma sel skuamosa (23%), *melanoma maligna* (7,9%) dan kanker kulit lainnya [1]. Kanker kulit dapat mengakibatkan kecacatan pada bagian kulit, merusak penampilan dan pada stadium lanjut dapat berakibat fatal pada penderitanya [2].

Dermoskopi sebagai teknik pencitraan non invasif telah menjadi alat penting bagi dermatolog mendeteksi dini kanker kulit yang terbukti mampu meningkatkan diagnosis dan peluang kesembuhan pasien [3]. Namun proses analisis citra dermoskopi secara manual memiliki beberapa keterbatasan yaitu proses tersebut memakan waktu, membutuhkan tingkat keahlian yang tinggi dan tidak mudah diskalakan karena jumlah kasus yang banyak [4]. Penelitian tentang kanker kulit telah menjadi perhatian pada beberapa bidang, khususnya penelitian medis dan teknologi informasi. Pada teknologi informasi, kebanyakan penelitian yang telah dilakukan adalah penelitian tentang deteksi dini kanker kulit menggunakan teknologi deep learning. Namun, pembuatan model deep learning untuk deteksi dini kanker kulit akan tergantung dengan bagaimana kualitas data yang dipakai [5]. Proses pengembangan model *deep learning* juga mendapat beberapa tantangan seperti membutuhkan *resource* komputasi yang besar, proses komputasi yang memakan waktu, proses pengembangan model yang dipengaruhi oleh kualitas data, variasi kondisi data citra (noise, pencahayaan, artefak rambut atau gelembung air dan variasi warna) serta kompleksitas batas lesi dengan kulit normal yang samar sehingga terdapat potensi *overfit* pada proses *training* datanya [6][7]. Sehingga dengan adanya tantangan dalam proses pelatihan tersebut, diusulkan salah satu metode preprocessing yang berguna untuk mengefisienkan proses pelatihan model, mengurangi potensi *overfit* dan menambah akurasi model [8]. Yaitu proses segmentasi lesi kulit untuk membantu memisahkan area lesi kulit dari latar belakang. Penelitian ini akan berfokus pada teknik segmentasi untuk mengolah dataset lesi kulit agar dapat digunakan untuk proses klasifikasi pada penelitian lebih lanjut. Proses segmentasi diperlukan karena dapat meningkatkan performa training pada model *convolution neural network* jika dilakukan menggunakan jumlah data yang sesuai [9].

Beberapa penelitian terdahulu telah mengusulkan beberapa model segmentasi yang digunakan untuk membantu proses klasifikasi lesi kulit. Dapeng cheng et al dalam penelitiannya mengusulkan tentang metode segmentasi berbasis *U-Net* yang disebut *EA-Net*. Pada metode tersebut, *encoder* yang sama dengan model *U-Net* diberikan tambahan *dedicated convolutional blocks* untuk memberikan *pixel-level attention guidance* pada *feature map* [10]. Setelah dilakukan *training* dan evaluasi, model *EA-Net* mendapatkan hasil yang cukup baik untuk proses segmentasi pada dataset *ISIC 2017* dan *ISIC 2018*. Penelitian lain yang dilakukan oleh Ravi et al yang memperkenalkan *FDUM-Net* yaitu kombinasi dari arsitektur *feature pyramid network* (FPN) dengan *encoder DenseNet21* yang terintegrasi dengan kerangka *U-Net*. Model tersebut dievaluasi dengan delapan varian CNN yang berbeda termasuk *U-Net* dan FPN dengan *encoder MobileNet*, *InceptionV3*, *DenseNet* dll) untuk menemukan konfigurasi terbaik. Pendekatan menggunakan *Transfer Learning* dengan *Encoder* pada *ImageNet* yang terbukti efektif meningkatkan kinerja segmentasi pada dataset dermoskopi yang relatif kecil [11]. Penelitian selanjutnya adalah penelitian yang dilakukan oleh Zhihui et al yang mengembangkan model yang diberi nama *MRP-UNet* (*Multiscale Residual Pyramid U-Net*) yang menggabungkan modul fusion *multiscale input*, mekanisme *ResNet-SE Attention* dan *Pyramid Dilated Convolution* di dalam kerangka *U-Net* [12]. Pada pengujiannya, *MRP-UNet* meraih akurasi sebesar 93,51% dan *Dice Coefficient* 93,14%. Hal ini membuat model *MRP-UNet* merupakan model yang efektif dalam meningkatkan detail segmen lesi sekaligus menjaga *False Positif* agar tetap rendah. Penelitian yang dilakukan oleh Le et al juga berfokus pada pembuatan model segmentasi yang ringan dan efisien yang bernama *Mobile Anti Aliasing Attention U-Net* (MAAU). pada model, ditambahkan *layer Anti-Aliasing* pada *Encoder* untuk stabilitas translasi serta penggunaan *Attention Block* di *decoder* untuk highlighting fitur-fitur yang penting di tiap kanal [13]. Penggunaan model ini berfokus untuk optimalisasi arsitektur segmentasi apabila digunakan pada perangkat bersumber daya ringan seperti ponsel dan *Mini PC*. Yuan et al juga mengusulkan pendekatan menggunakan metode *MUCM-Net* yang mengintegrasikan *U-Net* dengan komponen MLP (*Mamba*) untuk meningkatkan *Feature Learning* [14]. *MUCM-Net* memiliki desain yang ringan dan cocok untuk implementasi secara *Mobile* namun tetap memiliki akurasi yang tinggi. penelitian menggunakan *MUCM-Net* mengedepankan pentingnya inovasi pada penyusunan arsitektur jaringan dan *Fine Tuning Parameter* untuk menyeimbangkan antara akurasi dan kompleksitas arsitektur model. Penelitian selanjutnya dilakukan oleh Sun et al. yang mengusulkan *MSCA-Net* (*Multi-Scale Contextual Attention Network*) untuk segmentasi lesi kulit dengan menggunakan kerangka *U-Net* dengan *ResNet34* sebagai *Encoder*, kemudian mereka menambahkan tiga komponen utama yaitu *Multi Scale Bridge* untuk menggantikan *Skip Connection* standar agar fitur multi skala tidak membawa *noise*, *Global-local Channel Spatial Attention Module* (GL-CSAM) untuk menangkap *Global Context* termasuk informasi *local* melalui mekanisme *Attention* serta *Scale Aware Deep*

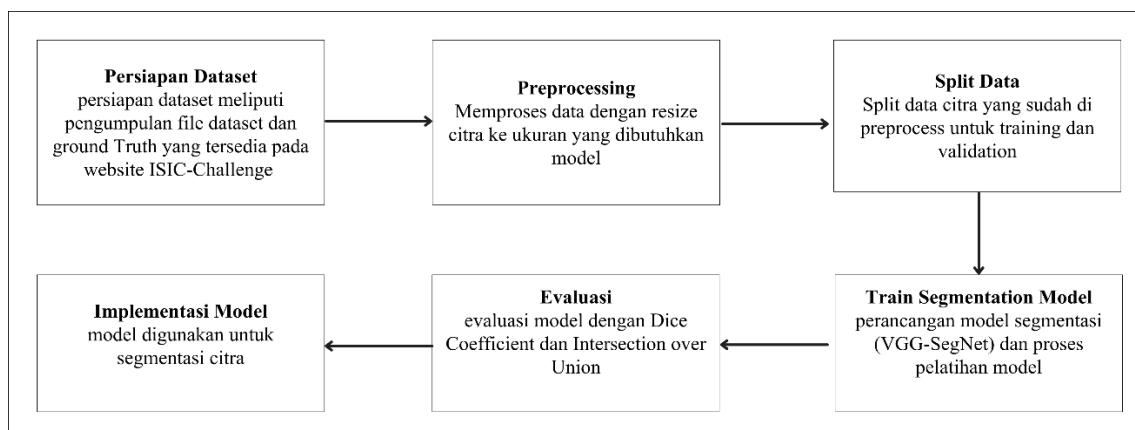
Supervision (SADS) untuk menggabungkan *Output Decoder* secara adaptif[15]. MSCA-NET menggunakan kombinasi *Binary Cross Entropy* dan *Soft Dice Loss* yang dievaluasi menggunakan ISIC 2017 memperoleh *Jaccard Index* (IoU) sebesar 79.26% dan *Dice Coefficient* sebesar 87.31%. Penelitian selanjutnya dilakukan oleh Wu et al. yang mengusulkan FAT-Net (*Feature Adaptive Transformers*) untuk segmentasi citra dermoskopi lesi kulit dengan mengganti *Encoder* tunggal menjadi *Dual Encoder* yang terdiri dari cabang CNN untuk mengekstrak fitur *local* dan cabang *Transformer* untuk menangkap *Global Context* [16]. Sehingga informasi detail dan dependensi jarak jauh dapat dipelajari secara bersamaan. Penelitian ini menambahkan *Feature Adaptation Module* pada *Skip Connection* untuk menyesuaikan distribusi fitur antara *Encoder-Decoder* dan membuat desain yang *Memory-Efficient* pada *Decoder* dengan menggabungkan fitur melalui *Element-Wise Summation*. FAT-Net menggunakan kombinasi *Binary Crossentropy* dan *Dice Loss* untuk proses *training* modelnya. Pada pengujian *dataset* ISIC 2017, FAT-Net mendapat *Dice Coefficient* sebesar 85% dan IoU sebesar 76,53%. Penelitian selanjutnya adalah penelitian yang dilakukan oleh Tong et al. yang mengembangkan model ASCU-Net (*Attention Gate, Spatial and Channel Attention U-Net*) untuk segmentasi lesi kulit pada citra dermoskopi. Inti dari pengembangan model adalah mekanisme *Triple Attention* yang menggabungkan *Attention Gate* pada *Skip Connection* untuk menekan respon yang tidak relevan dan *Spatial Attention* dan *Channel Attention* pada jalur decoding untuk meningkatkan representasi fitur dan korelasi spasial antar fitur[17]. Model ini diuji dengan menggunakan beberapa *dataset* (ISIC-2016, ISIC- 2017 dan PH2). pada ISIC-2017, ASCU-Net mendapat *Dice Coefficient* sebesar 83% dan *Jaccard Index* sebesar 74,2%.

Pada penelitian yang telah direview, dapat dilihat bahwa hasil terbaik segmentasi dapat diperoleh dengan adanya ketersediaan data yang memadai, pemrosesan data awal yang baik dan model dengan performa yang efisien. namun, masih terdapat beberapa keterbatasan yang menjadi *gap* untuk diteliti lebih lanjut. integrasi informasi multi skala dan detail tepi lesi masih dapat ditingkatkan. model seperti *MRP-Unet* telah mencoba modul piramida dan dilatasi, namun tidak semua model dapat menangani lesi dengan ukuran yang bervariasi atau batas yang tidak jelas. Kemudian terdapat keterbatasan jumlah data latih yang membuat model mengalami potensi *overfitting* sehingga perlu adanya peningkatan data melalui augmentasi atau teknik lain yang efektif. Sebagian penelitian telah melakukan augmentasi atau *transfer learning* namun optimasi *Hyperparameter* seperti *Learning Rate Scheduling* dan *Early Stopping* belum dibahas secara mendalam. Ketiga beberapa penelitian seperti *TransUnet* memiliki ukuran yang besar dan memperberat beban komputasi jika di implementasikan ke sumber daya yang terbatas. Di sisi lain, model dengan ukuran yang ringan akan sangat berkompromi dengan akurasi.

Untuk menjawab *gap* yang ditemukan pada berbagai penelitian diatas, penelitian ini mengusulkan pengembangan model segmentasi lesi kulit berbasis model hybrid SegNet-VGG16 yang dirancang untuk mempertahankan detail spasial. SegNet memiliki karakteristik *encoder* dan *decoder* yang efisien dengan mekanisme *pooling layer* sedangkan Blok konvolusi berbasis VGG efektif dalam mengekstrak fitur hierarkis pada citra. Selain rancangan arsitektur model, penelitian ini akan menggunakan proses fine tuning parameter pelatihan yang terstruktur (seperti *scenario* pengujian 3 fase, pengaturan *blok convolution*, konfigurasi loss yang sesuai) yang berfungsi untuk meningkatkan stabilitas pelatihan model dan menekan *overfitting* pada model. Kinerja model juga akan dievaluasi menggunakan metrik segmentasi seperti *Dice Coefficient* (DSC) dan *Intersection over Union* (IoU) sehingga perbandingan performa akan lebih objektif. Kebaruan riset serta Kontribusi utama penelitian ini Adalah pengembangan model segmentasi lesi kulit yang memperoleh hasil evaluasi yang lebih tinggi dibanding model yang sudah ada dengan penggunaan *hybrid algorithm* dari SegNet dan VGG. Model tersebut juga akan dilakukan *Fine Tuning* untuk dicari mana tuning terbaik untuk menghasilkan output model yang maksimal. Model tersebut diharapkan dapat berguna baik untuk diteliti ulang maupun diimplementasikan ke sistem medis saat ini.

2. Metodologi

Penelitian ini berfokus untuk pembuatan dan *Fine Tuning* model segmentasi citra Dermoskopi Kanker kulit menggunakan model hibrida VGG-SegNet. Proses segmentasi dilakukan secara sistematis yang dimulai dari pengumpulan data hingga evaluasi model. Tahapan utama pada proses penelitian ini adalah sebagai berikut: persiapan dataset, preprocessing data, *split* dataset menjadi *training* dan *validation* ke rasio sesuai parameter pengujian, *training segmentation model* dan evaluasi.



Gambar 1. Diagram Alur Metode Segmentasi Kanker Kulit Menggunakan Metode VGG-SegNet

Metode penelitian terdiri dari beberapa tahapan yaitu persiapan dataset dengan memilih dan mendownload dataset yang relevan dengan penelitian, Kemudian data dilakukan *Preprocessing* dengan mengubah ke size 256x256 dan dilakukan split dataset ke 70% *training data* dan 30 % *validation data* pada data *mask ground truth* dilakukan *converting ke grayscale*. Selanjutnya data akan dilatih pada model yang telah didefinisikan. Hasil dari proses pelatihan tersebut akan dievaluasi menggunakan beberapa metode seperti *Dice Coefficient* dan *Intersection over Union*.

2.1. Dataset

Penelitian ini menggunakan data yang tersedia di internet yaitu *ISIC 2017* dan *ISIC 2018*. *ISIC (International Skin Imaging Collaboration)* merupakan dataset besar untuk tugas segmentasi dan deteksi kanker kulit. Pada *isic 2018*, terdapat 7 kelas utama, yaitu *Melanoma*, *Melanocytic Nevus*, *Basal Cell Carcinoma*, *Actinic Keratosis*, *Benign Keratosis*, *Dermatofibroma* dan *Vascular Lesion*.



Gambar 2. Contoh Raw Data Lesi Kulit pada *Dataset ISIC*



Gambar 3. Contoh Dataset Ground Truth pada ISIC

Pada dataset *ISIC-2017* memiliki 3 kelas utama. Yaitu *Melanoma*, *Nevus* dan *Seborrheic Keratosis*. Pada tiap *folder ISIC 2017* dan *ISIC 2018*, terdapat juga file *Ground Truth* pada tiap file *dataset*. *Ground Truth* merupakan data referensi yang dianggap benar dan menjadi patokan evaluasi yang biasa digunakan untuk melatih, validasi dan menilai hasil model [18]. Total jumlah data yang digunakan pada *dataset ISIC 2017* adalah 2000 data citra dan 2000 citra

Ground Truth. Dan pada dataset ISIC 2018 total jumlah datanya adalah 2594 data dengan *Ground Truth* berjumlah sama.

2.2. Preprocessing

Preprocessing merupakan salah satu step awal yang penting dalam proses penelitian ini untuk memastikan data yang siap untuk di *consume* ke model segmentasi. Tahapan *preprocessing* yang dilakukan di penelitian ini adalah data citra akan di *resize* ke skala 256x256 dan dinormalisasi menggunakan *preprocess input VGG16* agar selaras dengan *statistic imageNet*.

Mask Ground Truth dibaca dalam *grayscale* yang berskala [0,1] lalu dibinarisasi pada ambang 0,5. Proses selanjutnya adalah melakukan split data untuk *train* dan *test*. Terdapat beberapa parameter yang diuji pada split data. Yaitu pembagian data *train* dan *test* dengan rasio 50:50 dan 70:30.

2.3. VGG-SegNet

Penelitian ini menggunakan metode hibrida VGG-SegNet untuk model segmentasinya. *Encoder* yang digunakan menggunakan *backbone* dari VGG sedangkan *decoder* nya adalah *decoder* asli dari SegNet. Kemudian terdapat tambahan *Skip Connection* pada level menengah (blok 2, 3 dan 4). *Encoder-Decoder* bekerja dengan prinsip *Analysis-Synthesis*, yaitu *Encoder* mengekstrak *Semantic Feature* dan memperbesar *Receptive Field* sehingga jaringan memahami konteks global. *Decoder* kemudian mensintesis Kembali *Feature Map* ke resolusi asli agar dapat melakukan klasifikasi per piksel. Fungsi dari tahapan *Skip Connection* adalah untuk mempertahankan detail batas lesi yang kerap hilang akibat *Pooling*. Tiap blok *Decoder* memiliki *Layer Convolution*, *Batch Normalization* dan *ReLU* dengan *Kernel 3x3* dan *Padding* bernilai "Same" untuk menjaga optimisasi. Untuk *Output Head* nya menggunakan *layer* konvolusi 1x1 dengan *Sigmoid* untuk menghasilkan peta probabilitas biner.

2.4. Fine Tuning

Pelatihan model pada penelitian ini disusun bertahap dengan *Adam Optimizer*, *Batch Size* sebesar 32 dan total 50 *epoch*, serta *Callbacks* berupa *ModelCheckpoint* untuk menyimpan bobot terbaik, *ReduceLROnPlateau* untuk menurunkan *learning rate* saat metrik validasi stagnan, dan *EarlyStopping* untuk mencegah *overfitting*. Skema *Fine Tuning* mengikuti prinsip *Progressive Unfreezing* dalam tiga fase yaitu *Encoder Freeze* membekukan seluruh blok VGG16 sehingga hanya *Decoder* yang belajar, *Unfreeze Block5* membuka lapisan dalam *Encoder* agar *Semantic Feature* menyesuaikan domain lesi dan fase kombinasi *Unfreeze Block4+5* dan *Focal-Tversky* memperluas adaptasi ke level menengah serta mengganti *Loss Function* dari kombinasi *BCE + Dice* ke *Focal-Tversky*. Penggunaan *Focal-Tversky* dimungkinkan karena pada segmentasi lesi, bagian positif (lesi) sering kecil dan tepinya samar sehingga penggunaan *Focal-Tversky* dapat membuat model lebih fokus ke piksel yang sulit dan menekan kesalahan *false negative* agar lesi kecil tidak terlewat.

2.5. Evaluasi Model

Selama pelatihan dan validasi, performa model dipantau dengan *dice coefficient* dan *Intersection over Union*. *Dice Coefficient* mengukur tingkat kesesuaian (*Similarity*) antara dua *Region* atau *mask* melalui perbandingan luas *Overlap* relatif terhadap total luas kedua *region* tersebut. Nilai DSC berada pada rentang 0–1, di mana nilai mendekati 1 menunjukkan kesesuaian yang sangat baik. Sesuai literatur, nilai DSC antara 0.8 hingga 0.9 umumnya dianggap sebagai segmentasi yang baik. Berikut adalah rumus implementasi dari *Dice Similarity Coefficient* [19]:

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \in [0, 1] \quad (1)$$

Dimana A = *Area/RegionMask* prediksi dan B = *Area Ground Truth*, dan $|A \cap B|$ = jumlah *pixel* yang *overlap*. Sedangkan *Intersection over Union (IoU)* atau bisa disebut *Jaccard Index* digunakan untuk evaluasi seberapa baik hasil segmentasi dengan menggunakan *Ground Truth* dengan cara membagi luas irisan antara peta segmentasi hasil prediksi dan *Ground Truth* dengan luas gabungan dari keduanya [20]. Berikut adalah rumus *Intersection over Union*:

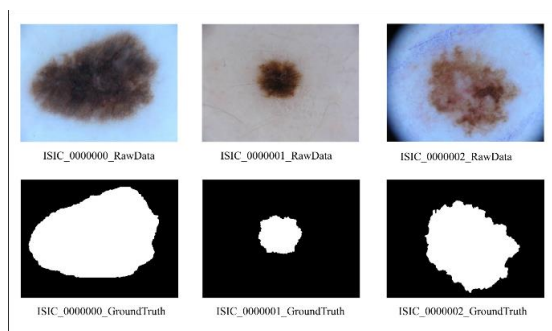
$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

Pada rumus tersebut, nilai A dan B masing-masing merepresentasikan *Ground Truth* dan peta hasil segmentasi hasil prediksi. Setelah model selesai dilatih, dilakukan *Threshold Tuning* pada set validasi dengan menjajal rentang ambang probabilitas untuk memilih nilai yang maksimal antara rata-rata *Dice/IoU*. Nilai ambang terbaik digunakan untuk menghitung metrik akhir.

3. Hasil dan Pembahasan

3.1. Persiapan Dataset

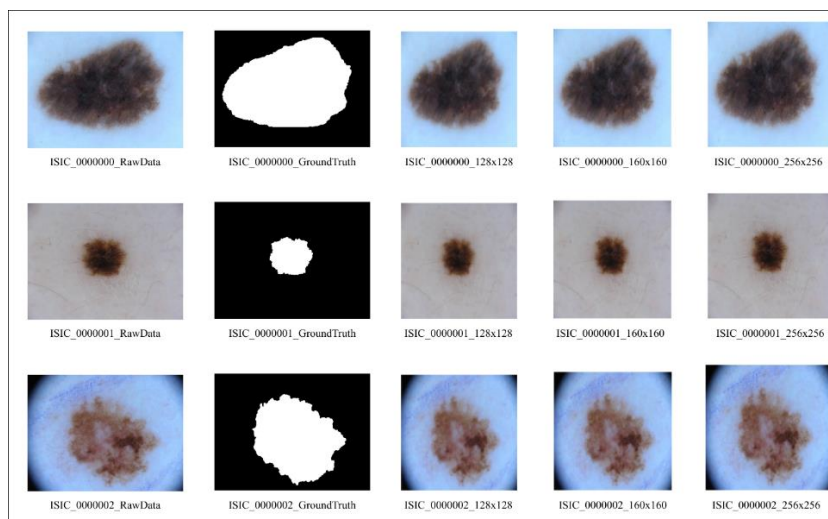
Dataset utama yang digunakan dalam pelatihan ini adalah dataset ISIC 2017. Pada penelitian ini akan menggunakan total jumlah 2000 data yang akan dibagi train dan validationnya dengan rasio 70:30. Selain itu terdapat juga file *ground truth* dari masing masing file dataset yang digunakan sebagai label referensi untuk acuan segmentasi. Berikut adalah contoh dataset yang digunakan untuk proses pelatihan model:



Gambar 4. Contoh Data Training dan Ground Truth

3.2. Preprocessing dan Split Data

Setiap data citra yang masih raw akan di *preprocessing* dengan dicari dan dipasangkan dengan *ground truth* untuk kebutuhan segmentasi dan di *resize* ukurannya menjadi 128x128, 160x160 dan 256x256 (pada model yang dibahas ini menggunakan 256x256) piksel sesuai kebutuhan *input* model-model yang di *define*.



Gambar 5. Hasil Resize Citra

Pada gambar diatas adalah hasil pemasangan raw data dengan *ground truth* dan contoh hasil resize citra dari ukuran 128x128, 160x160 dan 256x256. Tahapan selanjutnya adalah membagi data latih dan data validasi dengan rasio perbandingan 70:30 dan 50:50 sesuai parameter pengujian model. Pada model yang dijelaskan ini menggunakan rasio *split* data 70:30.

3.3. Pelatihan Model

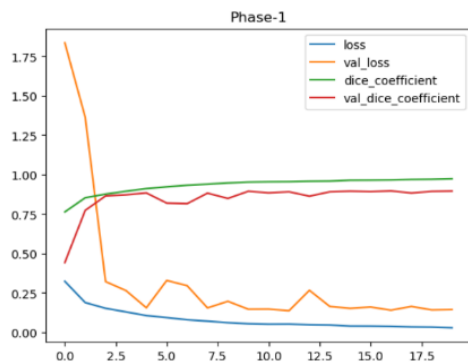
Pada tahap ini akan dilakukan *train* pada model VGG-SegNet pada input citra dermoskopi kanker kulit yang sudah di *preprocessing* dan di *split*. Pada penelitian ini akan menggunakan variasi *hybrid VGG-SegNet* dengan rincian model sebagai berikut:

Tabel 1. Rincian Model VGG-SegNet

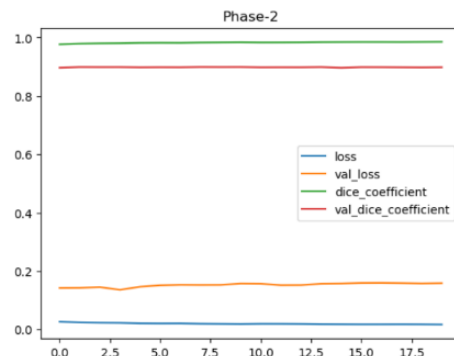
No	Nama Block	Deskripsi
1	<i>Input</i>	Citra masukan berukuran 128x128x3, 160x160x3 atau 256x256x3 (sesuai model).
2	<i>Encoder – VGG16 Pretrained (Conv 3x3 + ReLU)</i>	<i>Encoder</i> memakai backbone VGG16 <i>pretrained</i> dengan rangkaian <i>Conv2D 3x3</i> beraktivasi ReLU pada tiap blok.
3	<i>block_pool2</i>	Blok <i>encoder</i> dengan <i>MaxPooling</i> ; jumlah filter 128 (<i>output</i> fitur b2).
4	<i>block_pool3</i>	Blok <i>encoder</i> dengan <i>MaxPooling</i> ; jumlah filter 256 (<i>output</i> fitur b3).
5	<i>block_pool4</i>	Blok <i>encoder</i> dengan <i>MaxPooling</i> ; jumlah filter 512 (<i>output</i> fitur b4).
6	<i>block_pool5 (b5)</i>	Blok <i>encoder</i> terdalam dengan <i>MaxPooling</i> ; jumlah filter 512, menghasilkan fitur b5 beresolusi 8x8x512 (fitur paling abstrak/global).
7	<i>Decoder Stage 1 (Upsample + Skip b4)</i>	b5 (8x8x512) di- <i>upsample bilinear</i> 16x16, lalu <i>concat</i> dengan b4; diproses blok 2x(<i>Conv2D 3x3 + BatchNorm</i>) dengan 512 filter.
8	<i>Decoder Stage 2 (Upsample + Skip b3)</i>	Hasil <i>stage 1</i> di- <i>upsample</i> 32x32, <i>concat</i> dengan b3, diproses blok 2x(<i>Conv2D 3x3 + BatchNorm</i>) dengan 256 filter.
9	<i>Decoder Stage 3 (Upsample + Skip b2)</i>	Hasil <i>stage 2</i> di- <i>upsample</i> 64x64, <i>concat</i> dengan b2, diproses blok 2x(<i>Conv2D 3x3 + BatchNorm</i>) dengan 128 filter.
10	<i>Decoder Stage 4 (Upsample no-skip)</i>	<i>Upsample</i> 64 menjadi 128 tanpa <i>skip connection</i> , diproses blok 2x(<i>Conv2D 3x3 + BatchNorm</i>) dengan 64 filter.
11	<i>Decoder Stage 5 (Upsample no-skip)</i>	<i>Upsample</i> 128 menjadi 256 tanpa <i>skip connection</i> , diproses blok 2x(<i>Conv2D 3x3 + BatchNorm</i>) dengan 64 filter.
12	<i>Output Layer (Sigmoid)</i>	<i>Conv2D</i> (1, 1x1, <i>sigmoid</i>) memproyeksikan fitur 256x256x64 menjadi peta probabilitas <i>mask</i> 256x256x1 (segmentasi <i>biner</i>).

Setelah model dibangun, selanjutnya adalah melakukan *train* model. Model dilatih dengan *fine tuning* yang dilakukan secara bertahap (3 fase). Hal ini dimaksudkan agar model dapat beradaptasi dengan citra dermoskopi tanpa merusak model VGG16 yang sudah dilatih sebelumnya di *ImageNet*. Fase pertama pelatihan model adalah *encoder VGG16* dibebukan (*freeze layer*). Jadi yang dilatih adalah *decoder* dan *layer output* agar model cepat belajar pola dasar pada dataset lesi kulit. Fase ini menggunakan *optimizer Adam* dengan *learning rate* 1e-3 dan *Loss BCE+Dice* agar model lebih stabil menghadapi piksel yang tidak seimbang (area terdeteksi lesi jauh lebih sedikit dibanding *background*). Kemudian pada fase kedua, hanya *block 5* (blok terdalam) yang di *unfreeze* dan *training* dilanjutkan dengan *learning rate* yang lebih kecil (1e-4) yang membuat model dapat mulai menyesuaikan fitur tingkat tinggi (*global feature*) untuk

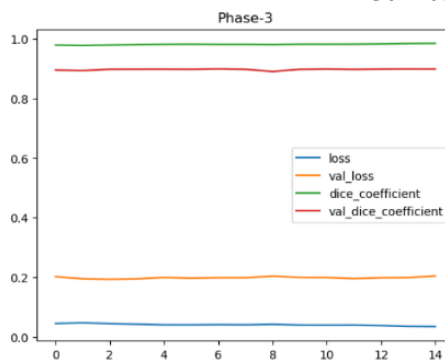
mencegah *overfitting* dan agar lebih cocok dengan karakteristik citra lesi kulit. Fase ketiga *encoder* dibuka lebih luas, yaitu block4 dan block5 dilatih bersamaan dengan *learning rate* yang lebih kecil ($5e-5$) untuk fitur menengah hingga tinggi yang berkaitan dengan batas dan tekstur lesi. Pada fase ketiga *loss* diganti dengan *BCE+Focal Tversky* agar model lebih peka terhadap batas lesi yang samar. Selama proses *training*, model menyimpan bobot terbaik berdasarkan *Validation Dice Coefficient*, pada prosesnya model juga menurunkan *learning rate* secara otomatis jika performa stagnan dengan fungsi *ReduceLROnPlateau* dan menghentikan training jika angka *Validation Dice Coefficient* tidak membaik (*EarlyStopping*) agar menghindari terjadinya *overfit*. History proses training model yang berisikan pemantauan *loss* dan *dice coefficient* tiap *epoch* disimpan dalam grafik sebagai berikut:



Gambar 5. Train Model Fase 1.



Gambar 6. Train Model Fase 2.



Gambar 7. Train Model Fase 3.

Pada fase *training* tahap pertama dan *Epoch* awal, nilai *validation loss* masih berada pada angka yang tinggi hingga mencapai diatas 1.75. hingga akhirnya turun pada 0.30 dan fluktuatif pada *Epoch* 3 hingga 8 namun berangsur angsur normal hingga berada pada angka 0.15. tahap kedua dan ketiga training model. *Dice Coefficient* dan *Validation Dice Coefficient* terlihat stabil pada angka 0.97 dan 0.89 dan hanya menurun sedikit pada fase 3. Namun karena adanya fungsi save hasil evaluasi yang terbaik berdasarkan *Dice Coefficient*, maka model yang disave bukan dari akhir tahap tiga namun berada pada tahap pertengahan di fase kedua.

3.4. Evaluasi

Hasil model yang memiliki *Validation Dice Coefficient* sebesar 0.89899 dan *Validation Intersection over Union* sebesar 0.8177 menunjukkan arsitektur tersebut sudah cocok terhadap data latih dengan lebih baik dan tidak mengalami *overfitting*. hal ini selaras dengan tujuan penelitian yaitu membuat model segmentasi yang dapat bekerja dengan akurat dibuktikan dengan hasil *Dice Coefficient* yang tinggi. Selain itu untuk menguji efektifitas penggunaan *fine tuning*, pada penelitian ini kami juga melatih model-model lain dengan setingan *fine tuning* yang berbeda beda.

Tabel 2. Hasil Pengujian model Segmentasi

No	Model	Fine Tuning	Dataset	Val_loss	Val_DSC	Val_IoU	Rank
1	VGG-SegNet	50 Epoch + input 256x256	ISIC-2017 (split 70-30)	0.1524	0.89899	0.8177	#1
		Skip Connection + BCE					
		Focal-Tversky					
		3 Phase Training					
2	VGG-SegNet	Add B2 Block	ISIC-2017 (split 70-30)	0.2074	0.89807	0.8164	#2
		50 Epoch + input 256x256					
		Skip Connection + BCE					
		Focal-Tversky					
3	VGG-SegNet	3 Phase Training	ISIC-2017 (split 70-30)	0.1761	0.89696	0.8144	#3
		100 Epoch + input 256x256					
		Skip Connection + BCE					
		100 Epoch + input 256x256					
4	VGG-SegNet	Skip Connection + BCE	ISIC-2017 (split 70-30)	0.2134	0.89642	0.8137	#4
		Focal-Tversky					
		3 Phase Training					
		50 Epoch + input 256x256					
5	VGG-SegNet	Skip Connection + BCE	ISIC-2017 (split 50-50)	0.1985	0.8954	0.8116	#5
		Focal-Tversky					
		3 Phase Training					
		Add B2 Block					
6	VGG-SegNet	50 Epoch + input 256x256	ISIC-2018 (split 70-30)	0.2572	0.8833	0.7926	
		Skip Connection + BCE					
		Focal-Tversky					
		100 Epoch + input 160x160					
7	VGG-SegNet	100 Epoch + input 160x160	ISIC-2018 (split 50-50)	4.8769	0.8416	0.7313	
8	SegNet	50 Epoch + input 128x128	ISIC-2017	0.3509	0.78	0.6479	
9	VGG-SegNet	100 Epoch + input 256x256	ISIC-2018 (split 50-50)	3.0668	0.8053	0.6787	
10	VGG-SegNet	100 Epoch + input 160x160	ISIC-2018 (split 70-30)	0.3434	0.8559	0.7493	

Pada tabel perbandingan diatas Perbedaan performa dari sisi *Validation Dice Coefficient* pada model dengan Rank 1 dan Rank 2 relatif sangat kecil yaitu 0.00092 yang mengindikasikan kontribusi *block2* hanya menambah sedikit peningkatan pada model namun tetap memiliki kontribusi untuk menaikkan performa model. Namun pada nilai *validation loss* cukup tinggi yaitu 0.055. semakin kecil nilai *validation loss* maka semakin bagus karena *error ratenya* makin kecil. Pada rank 5 terbawah terlihat model model yang tidak menggunakan *fine tuning* seperti *Skip Connection + BCE*, *Focal-Tversky* dan *3 Phase Training* memiliki hasil *Validation Dice Coefficient* yang lebih rendah.

Pada penelitian ini juga membandingkan model yang telah ada sebelumnya. Model yang dibuat perbandingan terbatas pada model segmentasi citra lesi kulit yang dilatih menggunakan

dataset yang sama dengan model terbaik pada penelitian ini. Berikut adalah tabel perbandingannya.

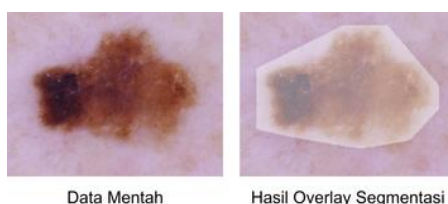
Tabel 3. Perbandingan Evaluasi Model Terbaik dengan Model Penelitian Lain

Nama Model	<i>Dataset</i>	<i>DSC</i>	<i>IoU (Jaccard)</i>
<i>MSCA-Net</i> [15]	ISIC 2017	0.8731	0.7926
<i>FAT-Net</i> [16]	ISIC 2017	0.85	0.7653
<i>ASCU-Net</i> [17]	ISIC 2017	0.83	0.742
Penelitian ini	ISIC 2017	0.8989	0.8177

Berdasarkan tabel perbandingan diatas, model yang dilatih pada penelitian ini menunjukkan performa tertinggi pada dataset ISIC 2017 untuk kedua metrik utama segmentasi, yaitu DSC dan IoU, model tersebut memiliki selisih 0,0258 DSC dan 0,0251 IoU dibanding *MSCA-Net*; selisih 0,0489 DSC dan 0,0524 IoU dibanding *FAT-Net*; serta 0,0689 DSC dan 0,0757 IoU dibanding *ASCU-Net*. Hal ini mengindikasikan model mampu menghasilkan segmentasi yang lebih akurat.

3.5. Implementasi Model

Model dengan hasil evaluasi terbaik kemudian di *deploy* untuk mensegmentasi citra dermoskopi kanker kulit. Dalam tahapan *deployment*nya ada beberapa citra dermoskopi yang diambil dari dataset lain diluar dataset yang digunakan untuk pengujian yaitu ISIC 2019. Berikut adalah contoh *input* data mentah yang berhasil di segmentasi menggunakan model dengan *Rank* pertama:



Gambar 8. Data yang Berhasil disegmentasi

Segmentasi dilakukan dengan melakukan *mapping* citra lesi kulit dengan mendeteksi bentuk global dengan lesi dengan baik. Namun masih terdapat *over segmentation* di tepi luar lesi kulit dengan model tetap mendeteksi Sebagian kulit yang sehat di sekitar lesi. Namun secara *overall* model dapat bekerja dengan mensegmentasi lesi kulit secara keseluruhan. Nantinya model segmentasi lesi kulit ini akan di implementasikan untuk memisahkan daerah yang dideteksi kanker kulit dan kulit yang sehat dan di *resize* untuk kebutuhan input citra klasifikasi.

3.6. Pembahasan

Hasil penelitian yang dilakukan menunjukkan bahwa model *VGG-SegNet* dengan tambahan konfigurasi fine tuning seperti skip connection, pelatihan 3 fase bertahap, kombinasi *loss BCE* dan *Focal-Tversky* dan penambahan *block b2* pada modelnya mampu menghasilkan kinerja yang baik pada dataset *ISIC 2017*. Model terbaik memperoleh *Validation Dice Coefficient* sebesar 0.89899 dan *Validation IoU* sebesar 0.8177. Jika dilihat dari skema pengujian yang dijelaskan pada bab Hasil Pengujian, peningkatan performa segmentasi tidak hanya ditentukan oleh arsitektur, namun strategi *fine tuning* yang sesuai juga dapat meningkatkan performa arsitektur. Temuan ini menunjukkan bahwa representasi fitur dan penjagaan terhadap informasi spasial pada proses decoding berperan dalam meningkatkan performa segmentasi area terdeteksi lesi kulit.

Secara konsep, hasil penelitian ini sejalan dengan kajian dari Mirikharaji et al yang mengungkapkan bahwa tantangan utama segmentasi lesi kulit terletak pada batas lesi yang kurang jelas, variasi bentuk dan ukuran lesi, dan keberadaan artefak citra, sehingga model membutuhkan kemampuan *feature extraction* yang kuat namun tetap mempertahankan detail spasial [9]. Dengan demikian, penelitian ini menguatkan teori terhadap penelitian terdahulu

bahwa kompleksitas arsitektur bukan hanya menjadi satu-satunya faktor keberhasilan hasil segmentasi, namun terdapat faktor lain seperti efektivitas skema pelatihan dan *fine tuning*.

Jika dibandingkan dengan beberapa penelitian sejenis dengan dataset yang sama (ISIC 2017), model pada penelitian ini menunjukkan kinerja yang lebih baik dibandingkan model lain seperti *MSCA-Net* [15], *FAT-Net* [16], dan *ASCU-Net* [17] yang masing-masing memperoleh *Dice Coefficient* dan *IoU* sebesar 87.31% dan 79,26%, 85% dan 76,53%, 83% dan 74,20%. Oleh karena itu, temuan ini berimplikasi bahwa optimalisasi model berbasis *VGG-SegNet* masih relevan untuk dikembangkan sebagai pendekatan segmentasi kulit yang lebih akurat. Kontribusi ilmiah pada penelitian ini terletak pada penguatan gagasan bahwa integrasi *backbone pretrained*, *skip connection*, pelatihan bertahap dan kombinasi fungsi *loss* dapat menjadi alternatif yang efektif untuk peningkatan akurasi segmentasi serta membuka peluang pengembangan model yang lebih adaptif. Namun, hasil implementasi di luar data citra latih masih menunjukkan adanya *over segmentation* di area tepi lesi yg terdapat pada beberapa output data citra. Hal ini menunjukkan bahwa perlu Adanya kajian terhadap teknik-teknik lain seperti augmentasi data atau post processing pada penelitian selanjutnya yang bertujuan untuk meningkatkan robustness model.

5. Simpulan

Penelitian ini berhasil mencapai tujuan dalam mengidentifikasi konfigurasi model *VGG-SegNet* terbaik untuk segmentasi citra dermoskopi kanker kulit pada dataset *ISIC 2017* dan *ISIC 2018*. Lima model teratas menunjukkan kinerja yang cukup baik dengan nilai *Validasi Dice Coefficient* diatas 0,895. Dengan konfigurasi dan strategi pelatihan yang telah dijabarkan pada bab sebelumnya, model mampu mendeteksi area lesi dengan presisi yang sesuai target penelitian.

Meskipun demikian terdapat beberapa keterbatasan dalam penelitian ini. Pelatihan model dilakukan dengan sumber daya komputasi dan waktu yang terbatas sehingga jumlah *Epoch* dan cakupan eksplorasi *hyperparameter* belum optimal. Kemudian augmentasi data seperti *flip*, *rotate*, *crop* dan lain-lain belum diterapkan pada penelitian ini. Ketiadaan augmentasi dapat membuat model kurang *robust* terhadap variasi posisi atau orientasi lesi baru. Kemudian penelitian yang telah dilakukan belum melakukan *Post Processing* pada output segmentasi. Tanpa Langkah *Post Processing*, hasil segmentasi memiliki potensi masih mengandung kesalahan kecil seperti prediksi area lesi kulit yang terputus atau *False Positive* diluar lesi utama.

Daftar Referensi

- [1] S. L. Menaldi, "Ilmu Penyakit Kulit dan Kelamin. Edisi ketujuh," 7th ed., Jakarta: Badan Penerbit FKUI, 2016.
- [2] N. Wedayani, N. A. Putri R, and D. Hidajat, "Edukasi Tentang Pengenalan Tanda Gejala, Pencegahan dan Penanganan Kanker Kulit Sebagai Dampak Paparan Sinar Matahari dan Penggunaan Kosmetik Berbahan Kimia Berbahaya di Poli Kulit Rumah Sakit Akademik Universitas Mataram," *Jurnal Pengabdian Magister Pendidikan IPA*, vol. 5, no. 3, pp. 223–226, Sep. 2022, doi: 10.29303/jpmipi.v5i3.2133.
- [3] A. Ahmed, G. Sun, A. Bilal, Y. Li, and S. A. Ebad, "Precision and efficiency in skin cancer segmentation through a dual encoder deep learning model," *Sci. Rep.*, vol. 15, no. 1, pp. 1–16, Feb. 2025, doi: 10.1038/s41598-025-88753-3.
- [4] P. Bansal, R. Garg, and P. Soni, "Detection of melanoma in dermoscopic images by integrating features extracted using handcrafted and deep learning models," *Comput. Ind. Eng.*, vol. 168, no. 3, pp. 1–15, Jun. 2022, doi: 10.1016/j.cie.2022.108060.
- [5] O. Akinrinade and C. Du, "Skin cancer detection using deep machine learning techniques," *Intell. Based. Med.*, vol. 11, no. 21, pp. 1–16, Jan. 2025, doi: 10.1016/j.ibmed.2024.100191.
- [6] R. Agrawal, N. Gupta, and A. S. Jalal, "CACBL-Net: a lightweight skin cancer detection system for portable diagnostic devices using deep learning based channel attention and adaptive class balanced focal loss function," *Multimed. Tools Appl.*, vol. 84, no. 14, pp. 12877–12900, Jun. 2024, doi: 10.1007/s11042-024-19485-1.
- [7] A. R. Priambodo and C. Fatichah, "Leveraging Convolutional Block Attention Module (Cbam) For Enhanced Performance In Mobilenetv3-Based Skin Cancer Classification," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 3, pp. 1389–1404, Jun. 2025, doi: 10.52436/1.jutif.2025.6.3.4546.

- [8] A. Tawakuli, B. Havers, V. Gulisano, D. Kaiser, and T. Engel, "Survey:Time-series data preprocessing: A survey and an empirical analysis," *Journal of Engineering Research (Kuwait)*, vol. 13, no. 2, pp. 674–711, Jun. 2025, doi: 10.1016/j.jer.2024.02.018.
- [9] Z. Mirikharaji *et al.*, "A survey on deep learning for skin lesion segmentation," *Med. Image Anal.*, vol. 88, no. 102863, pp. 1–40, Aug. 2023, doi: 10.1016/j.media.2023.102863.
- [10] D. Cheng *et al.*, "EA-Net: Research on skin lesion segmentation method based on U-Net," *Heliyon*, vol. 9, no. 12, pp. 1–12, Dec. 2023, doi: 10.1016/j.heliyon.2023.e22663.
- [11] H. Sharen, M. Jawahar, L. Jani Anbarasi, V. Ravi, N. Saleh Alghamdi, and W. Suliman, "FDUM-Net: An enhanced FPN and U-Net architecture for skin lesion segmentation," *Biomed. Signal Process. Control*, vol. 91, no. 4, pp. 1–15, May 2024, doi: 10.1016/j.bspc.2024.106037.
- [12] Z. Liu, J. Hu, X. Gong, and F. Li, "Skin lesion segmentation with a multiscale input fusion U-Net incorporating Res2-SE and pyramid dilated convolution," *Sci. Rep.*, vol. 15, no. 1, pp. 1–19, Dec. 2025, doi: 10.1038/s41598-025-92447-1.
- [13] P. T. Le *et al.*, "Anti-Aliasing Attention U-net Model for Skin Lesion Segmentation," *Diagnostics*, vol. 13, no. 8, pp. 1–15, Apr. 2023, doi: 10.3390/diagnostics13081460.
- [14] C. Yuan, D. Zhao, and S. S. Agaian, "MUCM-Net: a Mamba powered UCM-Net for skin lesion segmentation," *Explor. Med.*, vol. 5, pp. 694–708, 2024, doi: 10.37349/emed.2024.00250.
- [15] Y. Sun, D. Dai, Q. Zhang, Y. Wang, S. Xu, and C. Lian, "MSCA-Net: Multi-scale contextual attention network for skin lesion segmentation," *Pattern Recognit.*, vol. 139, no. 6, pp. 1–14, Jul. 2023, doi: 10.1016/j.patcog.2023.109524.
- [16] H. Wu, S. Chen, G. Chen, W. Wang, B. Lei, and Z. Wen, "FAT-Net: Feature adaptive transformers for automated skin lesion segmentation," *Med. Image Anal.*, vol. 76, no. 9, pp. 1–14, Feb. 2022, doi: 10.1016/j.media.2021.102327.
- [17] X. Tong, J. Wei, B. Sun, S. Su, Z. Zuo, and P. Wu, "Ascu-net: Attention gate, spatial and channel attention u-net for skin lesion segmentation," *Diagnostics*, vol. 11, no. 3, pp. 1–18, Mar. 2021, doi: 10.3390/diagnostics11030501.
- [18] E. R. Jefferson and E. Trucco, "The challenges of assembling, maintaining and making available large data sets of clinical data for research," in *Computational Retinal Image Analysis*, 1st ed., Trucco Emanuele, MacGillivray Tom, and Xu Yanwu, Eds., Oxford: Academic Press (Elsevier), 2019, ch. 20, pp. 429–444. doi: 10.1016/B978-0-08-102816-2.00021-6.
- [19] O. Rainio and R. Klén, "Modified Dice Coefficients for Evaluation of Tumor Segmentation from PET Images: A Proof-of-Concept Study," *Journal of Imaging Informatics in Medicine*, vol. 39, no. 2, pp. 785–793, May 2025, doi: 10.1007/s10278-025-01535-1.
- [20] M. E. Rayed, S. M. S. Islam, S. I. Niha, J. R. Jim, M. M. Kabir, and M. F. Mridha, "Deep learning for medical image segmentation: State-of-the-art advancements and challenges," *Inform. Med. Unlocked*, vol. 47, no. 7, pp. 1–25, Jan. 2024, doi: 10.1016/j.imu.2024.101504.