

# Klasifikasi Kanker Kulit menggunakan Metode *Convolutional Neural Network* dengan Arsitektur VGG-16

REGITA AGUSTINA, RITA MAGDALENA, NOR KUMALASARI CAECAR  
PRATIWI

Fakultas Teknik Elektro, Universitas Telkom, Indonesia  
Email: regitagustina@gmail.com

*Received* 22 November 2021 | *Revised* 12 Desember 2021 | *Accepted* 3 Januari 2022

## ABSTRAK

*Kanker kulit merupakan penyakit yang ditimbulkan oleh perubahan karakteristik sel penyusun kulit dari normal menjadi ganas, yang menyebabkan sel tersebut membelah secara tidak terkendali dan merusak DNA. Deteksi dini dan diagnosis yang akurat diperlukan untuk membantu masyarakat mengidentifikasi apakah kanker kulit atau hanya kelainan kulit biasa. Pada studi ini, dirancang sebuah sistem yang dapat mengklasifikasi kanker kulit dengan memanfaatkan citra kulit pasien yang kemudian diolah menggunakan metode Convolutional Neural Network (CNN) arsitektur VGG-16. Dataset yang digunakan berupa citra jaringan kanker sebanyak 4000 gambar. Proses diawali dengan input citra, pre-processing, pelatihan model dan pengujian sistem. Hasil terbaik diperoleh pada pengujian tanpa pre-processing CLAHE dan Gaussian filter, dengan menggunakan hyperparameter optimizer SGD, learning rate 0,001, epoch 50 dan batch size 32. Akurasi yang diperoleh sebesar 99,70%, loss 0,0055, presisi 0,9975, recall 0,9975 dan f1-score 0,9950.*

**Kata kunci:** Kanker kulit, CNN, VGG-16

## ABSTRACT

*Skin cancer is a disease caused by changes in the characteristics of skin cells from normal to malignant, which causes the cells to divide uncontrollably and damage DNA. Early detection and accurate diagnosis are necessary to help the public identify whether skin cancer or just a common skin disorder. In this study, a system was designed that can classify skin cancer by utilizing images of patients' skin which is then processed using the Convolutional Neural Network (CNN) method of VGG-16 architecture. Dataset used in the form of cancer tissue imagery as many as 4000 images. The process begins with image input, pre-processing, model training and system testing. The best results were obtained on testing without pre-processing CLAHE and Gaussian filters, using hyperparameters, SGD optimizer, learning rate 0.001, epoch 50 and batch size 32. Accuracy obtained by 99.70%, loss 0.0055, precision 0.9975, recall 0.9975 and f1-score 0.9950.*

**Keywords:** Skin cancer, CNN, VGG-16

## 1. PENDAHULUAN

Kanker kulit merupakan penyakit yang ditimbulkan oleh perubahan karakteristik sel penyusun kulit dari normal menjadi ganas, yang menyebabkan sel tersebut membelah secara tidak terkendali dan merusak DNA (**Hendaria, dkk, 2013**). Dalam tiga tipe terbanyak, kanker kulit tergolong menjadi tiga jenis yaitu melanoma maligna, karsinoma sel basal dan karsinoma sel skuamosa (**Wilvestra, dkk, 2018**). Terdapat 5,4 juta kasus baru kanker kulit di Amerika Serikat setiap tahunnya, satu dari lima orang Amerika didiagnosis mengidap penyakit kulit semasa hidup mereka. Meskipun melanoma mewakili kurang dari 5% dari semua kanker kulit di Amerika Serikat, tetapi Melanoma bertanggung jawab sekitar 75% dari semua kematian akibat kanker kulit (**Esteva, dkk, 2017**). Kanker kulit di Indonesia menempati urutan ketiga setelah kanker serviks dan kanker payudara. Beberapa faktor yang dapat menjadi pencetus timbulnya kanker kulit antara lain, faktor genetik, peningkatan radiasi sinar UV (**Silpa & V, 2013**), infeksi *human papillomavirus* dan pola hidup yang tidak sehat (**Hendaria, dkk, 2013**).

Dalam bidang medis, diagnosis penyakit kanker kulit umumnya dilakukan dengan proses biopsi. Yaitu bagian jaringan kulit diambil untuk kemudian dicek secara mendetail apakah jaringan tersebut merupakan sel kanker atau bukan. Namun, teknik pengujian ini memerlukan waktu yang cukup lama untuk seorang ahli dermatologis dan membutuhkan biaya besar (**Munthe, 2018**). Sedangkan perkiraan kelangsungan hidup kanker kulit bila terlambat terdeteksi hanya sekitar 14%, tetapi dapat meningkat lebih dari 99% jika terdeteksi pada tahap awal (**Esteva, dkk, 2017**). Deteksi dini dan diagnosis yang akurat dengan memanfaatkan citra jaringan kanker pasien yang kemudian diolah oleh sistem dengan waktu komputasi yang cepat sangat diperlukan untuk membantu masyarakat mengidentifikasi apakah kanker kulit atau hanya kelainan kulit biasa, serta membantu bidang medis dalam mengurangi risiko keterlambatan dalam penanganan kanker kulit.

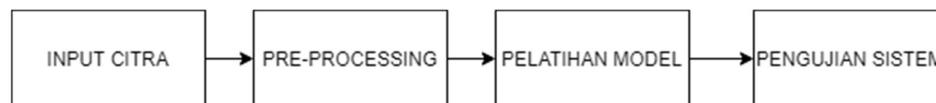
Telah terdapat beberapa penelitian terdahulu mengenai klasifikasi untuk kanker kulit menggunakan *Convolutional Neural Network* (CNN). Berdasarkan penelitian yang dilakukan oleh (**Esteva, dkk, 2017**) menguji dengan membandingkan algoritma dan dermatologis yang membedakan *keratinocyte carcinoma versus tumor seborrheic keratosis* dan *malignant melanoma versus benign nevus*. Penulis menggunakan arsitektur *Convolutional Neural Network* (CNN) GoogleNet *Inception V3* yang telah dilatih sebelumnya pada sekitar 1,28 juta gambar (1.000 kategori objek) dari Tantangan Pengenalan Visual Skala Besar ImageNet 2014. Penelitian tersebut berhasil mencapai akurasi sebesar 93,9%. Penelitian selanjutnya oleh (**Munthe, 2018**) melakukan klasifikasi citra kanker kulit berdasarkan tingkat keganasan kanker pada melanosit menggunakan *Deep Convolutional Neural Network* (DCNN). Penelitian tersebut menggunakan 1500 citra dengan tiga kategori yaitu *nevus*, *melanoma in situ* dan *malignant melanoma*. Akurasi yang didapat berdasar penelitian adalah sebesar 84%. Selanjutnya penelitian oleh (**Savera, dkk, 2020**) yakni menggunakan dua metode untuk mendeteksi dini kanker kulit, yaitu deteksi dengan klasifikasi secara regresi dan metode *Convolutional Neural Network* (CNN). Akurasi deteksi dengan ekstraksi fitur dan regresi k-NN sebesar 75% dengan *true positive rate (precision)* sebesar 85% dan *true negative rate (specificity)* sebesar 72%. Sementara, akurasi yang dihasilkan dengan menggunakan CNN adalah sebesar 76,56%. Kemudian penelitian berikutnya oleh (**Fu'adah, dkk, 2020**) mengklasifikasi kanker kulit menjadi empat kelas menggunakan metode CNN dengan 3 *hidden layer* memperoleh hasil akurasi sebesar 99% dengan *loss* sebesar 0,0346.

Merujuk pada permasalahan dan pemaparan singkat beberapa jurnal penelitian diatas, metode *Convolutional Neural Network* (CNN) terbukti memiliki hasil akurasi yang baik walau masih

terdapat *gap* dalam meningkatkan hasil akurasi. Pada penelitian ini diinginkan lebih memperoleh hasil akurasi yang lebih maksimal dengan nilai *loss* yang lebih rendah dari penelitian sebelumnya. Banyaknya gambar dengan empat kategori klasifikasi yang digunakan, maka penelitian ini akan menggunakan metode CNN dengan arsitektur VGG-16. Dimana arsitektur model VGG-16 merupakan top-5 dalam tantangan ImageNet ILSVRC tahun 2014 yang memperoleh akurasi sebesar 92,6% dengan kumpulan data lebih dari 14 juta gambar yang termasuk dalam 1000 kelas. **(Simonyan & Zisserman, 2015)**. Arsitektur VGG-16 banyak digunakan dalam teknik klasifikasi gambar *deep learning* karena berkinerja baik dan kemudahan dalam implementasinya. Penelitian akan mengklasifikasi kanker kulit menjadi empat kelas, yaitu melanoma, karsinoma sel skuamosa, dermatofibroma dan nevus pigmentosus. Diantaranya dua kelas merupakan dua jenis golongan kanker kulit terbanyak, dan dua kelas merupakan tumor jinak yang 50% berpotensi berubah menjadi melanoma maligna **(Tsaniyah, dkk, 2015)**.

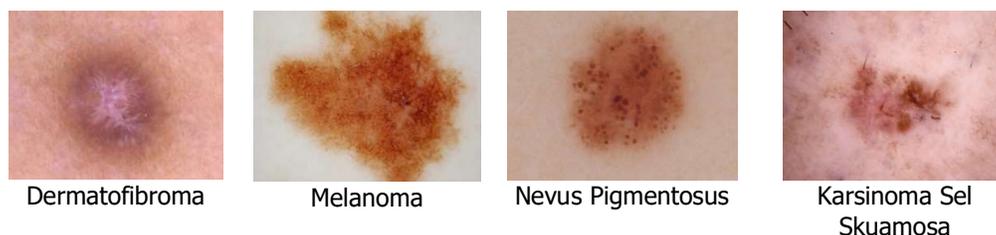
## 2. METODE PENELITIAN

Berdasarkan blok diagram pada Gambar 1 penelitian akan dibagi menjadi 4 tahap. Tahap pertama diawali dengan *input* citra kanker kulit yang kemudian masuk pada tahap *pre-processing*. Tahap *pre-processing* bertujuan untuk memperbaiki kualitas citra sehingga diharapkan menghasilkan hasil keluaran yang maksimal. Tahap ini mencakup *resizing* citra, CLAHE dan Gaussian *filter*. Setelah itu citra memasuki tahap pelatihan model menggunakan CNN dengan arsitektur VGG-16, sehingga didapatkan *output* model terbaik untuk klasifikasi kanker kulit. Lalu tahap pengujian sistem terdiri dari 4 skenario pengujian.



**Gambar 1. Blok Diagram Sistem**

*Dataset* dalam penelitian ini diperoleh dari *open-access dermatologists repository International Skin Imaging Collaboration* (ISIC) (<https://www.isic-archive.com>). Data yang digunakan sejumlah 4000 citra, yang terdiri atas 1000 citra Melanoma, 1000 citra Karsinoma Sel Skuamosa, 1000 citra Dermatofibroma, dan 1000 citra Nevus Pigmentosus dalam format \*.JPG. *Dataset* yang digunakan merupakan citra jaringan kanker. Pada proses pelatihan dan pengujian, digunakan sebanyak 75% data latih dan 25% data uji. Gambar 2 merupakan contoh data citra jaringan kanker kulit yang diperoleh dari ISIC.

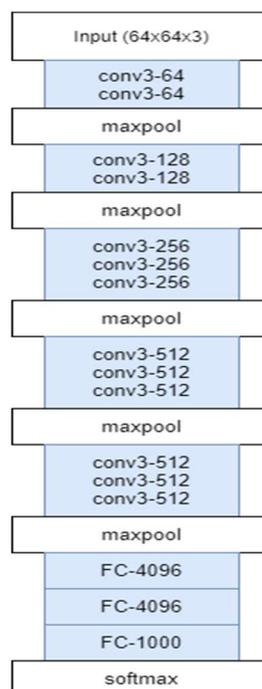


**Gambar 2. Dataset Citra Jaringan Kanker Kulit**

Tahap *pre-processing* dilakukan sebagai salah satu skenario untuk mengetahui performa terbaik. Tahap *pre-processing* mencakup beberapa proses yaitu *resizing* citra, CLAHE dan Gaussian *filter*. Pada tahap *resizing* citra terjadi proses perubahan ukuran citra digital dari ukuran yang beragam menjadi ukuran 64x64 piksel. Pada *pre-processing* CLAHE, terjadi peningkatan nilai kontras dengan memberikan nilai *clip limit* (nilai batas maksimal tinggi suatu

histogram) sehingga kontras dapat meningkat namun tidak berlebihan walaupun bekerja pada area lokal (**Prabowo & Premunendar, 2019**). Sedangkan Gaussian *filter* merupakan *filter* penghalusan linier yang memilih bobot sesuai dengan bentuk fungsi Gaussian, baik dalam domain spasial atau dalam domain frekuensi (**Wang, dkk, 2014**).

Pelatihan model yang digunakan pada penelitian ini adalah CNN dengan model arsitektur VGG-16. CNN merupakan pengembangan dari *Multilayer Perceptron* (MLP) yang dirancang untuk mengolah suatu data dua dimensi. CNN termasuk kedalam tipe *Deep Neural Network*, karena sistem memiliki kedalaman jaringan yang tinggi dan banyak digunakan pada data gambar (**Fu'adah, dkk, 2020**). Metode CNN sendiri secara umum terbagi menjadi *feature extraction layer* dan *classification layer*. Pada *feature extraction layer*, *input* citra masuk pada *convolutional layer*, *layer* aktivasi ReLU dan *pooling layer*. Selanjutnya, pada *classification layer* meliputi *fully-connected layer* dan lapisan aktivasi *softmax*. Gambar 3 menunjukkan arsitektur pelatihan model VGG-16.



**Gambar 3. Pelatihan Model VGG-16**

Pada model VGG-16, tahap pelatihan model terdiri 13 *convolutional layer*, 3 *fully-connected layer*, 5 *maxpooling layer* beserta *activation layer* ReLU dan *softmax*. Input citra berukuran 64x64 piksel dengan 3 *channel* RGB. Model ini melalui 5 kali proses konvolusi dengan ukuran filter 3x3 dan setiap prosesnya memakai jumlah filter yang berbeda, yaitu 64 *filter*, 128 *filter*, 256 *filter* dan 512 *filter*. Untuk *maxpooling* menggunakan ukuran filter 3x3 dengan jumlah stride 2. Pada *fully-connected layer*, 2 *layer* terdiri dari 4096 unit dan 1000 unit (**Simonyan & Zisserman, 2015**).

*Confusion matrix* adalah suatu metode yang digunakan untuk mengukur performansi metode klasifikasi, dimana berisi informasi yang membandingkan hasil klasifikasi oleh sistem dengan yang seharusnya (**Prasetyo, 2012**). Terdapat empat istilah representasi hasil klasifikasi saat pengukuran performansi menggunakan *confusion matrix*. Diantaranya yaitu Total *True Positive* (TTP), Total *True Negative* (TTN), Total *False Positive* (TFP) dan Total *False Negative* (TFN). Dimana berdasarkan dari nilai TTP, TTN, TFP dan TFN tersebut dapat diperoleh nilai akurasi, presisi, *recall* dan *f1-score* (**Pratiwi, dkk, 2021**).

Akurasi mempresentasikan seberapa akurat sistem dapat mengklasifikasi data secara benar. Nilai akurasi adalah perbandingan antara data yang terklasifikasi benar dengan jumlah keseluruhan data. Nilai akurasi dapat diperoleh dengan Persamaan (1) berikut.

$$Akurasi = \frac{TTP_{all} + TTN_{all}}{Jumlah\ total\ pengujian} \quad (1)$$

Presisi adalah perbandingan nilai dari jumlah data kategori terklasifikasi benar dengan total keseluruhan data kategori terklasifikasi benar. Untuk mencari nilai presisi dapat digunakan Persamaan (2) berikut.

$$Presisi = \frac{TTP_{all}}{TTP_{all} + TFP_i} \quad (2)$$

*Recall* dilakukan untuk mengetahui perbandingan jumlah data kategori terklasifikasi benar oleh sistem dengan jumlah data kategori terklasifikasi benar dan salah. Persamaan (3) berikut dapat digunakan untuk memperoleh nilai *recall*.

$$Recall = \frac{TTP_{all}}{TTP_{all} + TFN_i} \quad (3)$$

*F1-Score* merupakan penggabungan dari presisi dan *recall*. Nilai *f1-score* dapat diperoleh dengan Persamaan (4) berikut.

$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (4)$$

Dimana  $TTP_{all}$  berarti total data yang menunjukkan kondisi saat data positif dan terprediksi benar sebagai positif.  $TTN_{all}$  menunjukkan kondisi saat data negatif dan terprediksi benar sebagai negatif.  $TFP_i$  menunjukkan kondisi saat data negatif tetapi terprediksi salah sebagai positif.  $TFN_i$  menunjukkan kondisi saat data positif tetapi salah terprediksi sebagai negatif.

*Loss function* merupakan bagian penting ketika melakukan pelatihan model. Fungsi ini mengukur seberapa baik model melakukan tugas dalam menghitung *loss* dari *output* model dan target yang diinginkan. Jika prediksi model salah, maka nilai *loss* akan tinggi. Jika prediksi model cukup bagus, maka nilai *loss* akan mendekati nilai nol.

*Categorical crossentropy* merupakan salah satu dari *loss function* yang digunakan untuk mengklasifikasi data *multi-class*. Singkatnya, terdapat data satu dari banyak kategori yang mungkin, namun model harus memutuskan satu. Secara formal, ini dirancang untuk mengukur perbedaan diantara dua probabilitas (**Bruballa, 2018**).

Model CNN dilatih mendalam dengan memperbaharui semua lapisan dalam jaringan secara iteratif dan *optimizer* sangat memegang peranan penting. Algoritma *gradient descent* merupakan pilihan umum untuk melakukan optimasi pada *neural network* (**Bera & Shrivastava, 2020**). Untuk meminimalkan fungsi tujuan, parameter diperbarui dalam arah kebalikan dari gradien fungsi tujuan. Adapun *optimizer* yang digunakan pada penelitian adalah SGD, RMSprop, Adam dan Nadam.

### 3. HASIL DAN PEMBAHASAN

Pengujian sistem dilakukan untuk mengetahui hasil performansi terbaik serta menganalisis parameter apa saja yang mempengaruhi kinerja sistem. Pada tahap ini pengujian sistem dibagi menjadi 4, yaitu pengujian *resizing* citra, pengujian tanpa *pre-processing* CLAHE dan Gaussian *filter* dan pengujian menggunakan *pre-processing* CLAHE dan Gaussian *filter*. Dimana dalam setiap pengujian dilakukan skenario pengaruh *hyperparameter* yang mencakup skenario *optimizer*, *learning rate*, *epoch* dan *batch size*.

Pengujian *resizing* Citra dilakukan dengan melakukan perubahan ukuran citra, bertujuan menemukan ukuran citra yang sesuai untuk menghasilkan hasil akurasi yang maksimal. Adapun ukuran citra yang akan diuji adalah 64×64, 128×128 dan 224×224 piksel, dengan beberapa parameter lain yang sudah ditentukan yaitu *optimizer* SGD, nilai *learning rate* 0,001, *epoch* 50 dan *batch size* 32. Tabel 1 menunjukkan hasil akurasi pengujian *resizing* citra.

**Tabel 1. Pengujian *Resizing* Citra**

Ukuran Citra	Akurasi
32×32	99,80%
64×64	100%
128×128	98,59%

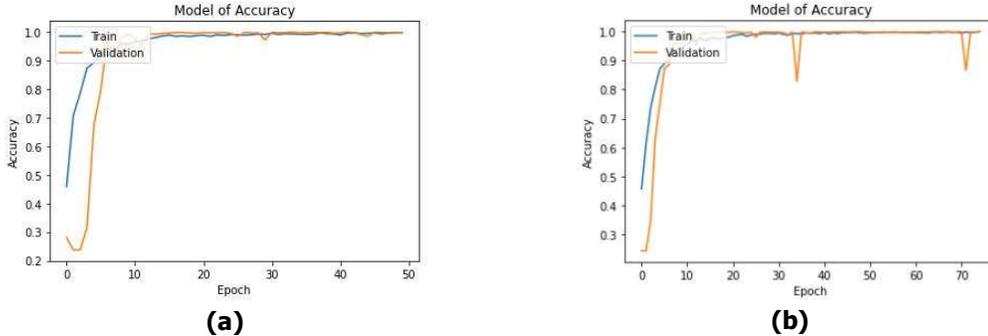
Skenario selanjutnya pengujian tanpa menggunakan *pre-processing* CLAHE dan Gaussian *filtering*. Pada pengujian ini terdapat 4 skenario pengujian *hyperparameter*, yaitu melihat pengaruh penggunaan *optimizer*, *learning rate* (LR), *epoch* dan *batch size* (BC). Hasil skenario pengujian ditunjukkan pada Tabel 2 berikut.

**Tabel 2. Hasil pengujian tanpa *pre-processing* CLAHE dan Gaussian *filter***

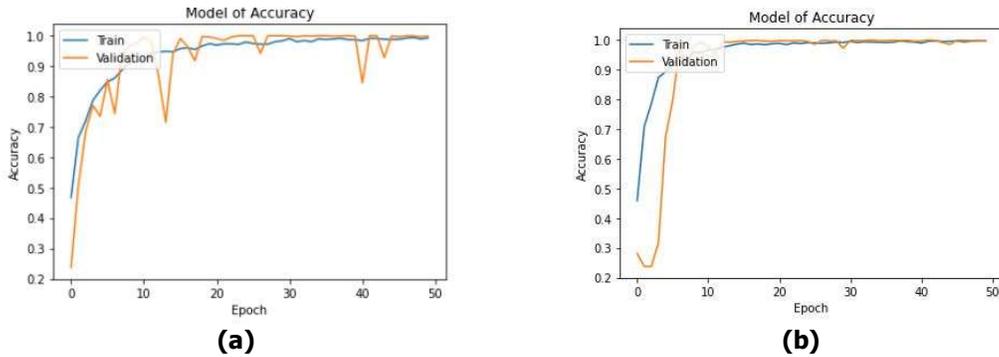
Skenario 1: <i>Optimizer</i> (LR 0,001, <i>epoch</i> 15, BC 32)			Skenario 2: <i>Learning rate</i> ( <i>Optimizer</i> SGD, <i>epoch</i> 15, BC 32)		
<i>Optimizer</i>	Akurasi	<i>Loss</i>	<i>Learning Rate</i>	Akurasi	<i>Loss</i>
Adam	92,40%	0,3434	0,1	87,70%	0,4599
Nadam	92,00%	0,1966	0,01	94,90%	0,1920
RMSprop	79,20%	0,9475	0,001	95,20%	0,1294
SGD	95,20%	0,1294	0,0001	84,70%	0,3609
Skenario 3: <i>Epoch</i> ( <i>Optimizer</i> SGD, LR 0,001, BC 32)			Skenario 4: <i>Batch Size</i> ( <i>Optimizer</i> SGD, LR 0,001, <i>epoch</i> 50)		
<i>Epoch</i>	Akurasi	<i>Loss</i>	<i>Batch size</i>	Akurasi	<i>Loss</i>
15	95,20%	0,1294	16	99,80%	0,0120
25	97,10%	0,0601	32	99,70%	0,0055
50	99,70%	0,0055	64	99,40%	0,0168
75	100%	9,2684e-04	128	99,70%	0,0098

Pada skenario *epoch*, hasil pengujian menunjukkan akurasi *epoch* tertinggi berada pada *epoch* 75 dengan nilai akurasi sebesar 100% dengan *loss* sebesar 9,2684e-04. Namun, hasil terbaik yang dipilih untuk skenario *epoch* adalah 50, karena kerja model yang tidak terlalu berat tetapi sudah menunjukkan peningkatan akurasi disetiap iterasi (*epoch*) dan perbedaan akurasi antara data *training* dan validasi tidak jauh berbeda. Sedangkan untuk *epoch* 75 masih menunjukkan munculnya banyak *spikes* atau lonjakan seperti pada Gambar 4. Pada skenario *batch size*, hasil akurasi menunjukkan *batch size* tertinggi berada pada nilai 16 dengan nilai akurasi sebesar 99,80% dan *loss* 0,0120. Namun, grafik akurasi *batch size* 16 pada Gambar 5 menunjukkan

munculnya banyak *spikes*/lonjakan, sehingga *batch size* terbaik berada pada *batch size* 32. Jadi, pada tahap pengujian tanpa *pre-processing* CLAHE dan tanpa Gaussian *filter* dengan menggunakan *hyperparameter* terbaik yaitu *optimizer*SGD, *learning rate* 0,001, *epoch* 50 dan *batch size* 32 didapatkan hasil akhir nilai akurasi sebesar 99,70% dan *loss* sebesar 0,0055.



**Gambar 4. Grafik Akurasi Pengujian Epoch (a) 50 dan (b) 75**



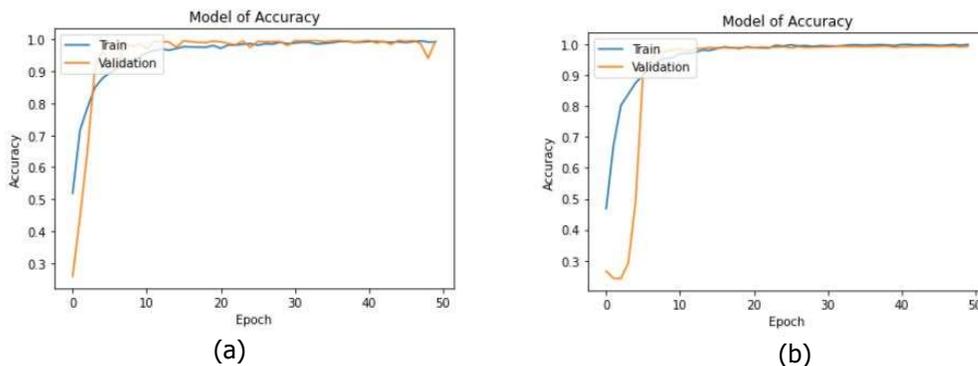
**Gambar 5. Grafik Akurasi Pengujian Batch Size (a) 16 dan (b) 32**

Pengujian selanjutnya dilakukan untuk melihat pengaruh *pre-processing* CLAHE. Dalam pengujian dengan *pre-processing* CLAHE dilakukan 4 skenario pengujian terhadap *hyperparameter*, yaitu *optimizer*, *learning rate* (LR), *epoch* dan *batch size* (BC). Hasil skenario pengujian ditunjukkan pada Tabel 3 berikut.

**Tabel 3. Hasil pengujian dengan *pre-processing* CLAHE**

Skenario 1: <i>Optimizer</i> (LR 0,001, <i>epoch</i> 15, BC 32)			Skenario 2: <i>Learning rate</i> (Optimizer SGD, <i>epoch</i> 15, BC 32)		
<i>Optimizer</i>	Akurasi	<i>Loss</i>	<i>Learning Rate</i>	Akurasi	<i>Loss</i>
Adam	90,10%	0,4043	0,1	96,20%	0,1167
Nadam	96,60%	0,0915	0,01	98,80%	0,0442
RMSprop	75,70%	1,8139	0,001	99,10%	0,0465
SGD	99,10%	0,0465	0,0001	90,10%	0,2426
Skenario 3: <i>Epoch</i> (Optimizer SGD, LR 0,001, BC 32)			Skenario 4: <i>Batch Size</i> (Optimizer SGD, LR 0,001, <i>epoch</i> 50)		
<i>Epoch</i>	Akurasi	<i>Loss</i>	<i>Batch size</i>	Akurasi	<i>Loss</i>
15	99,10%	0,0465	16	99,50%	0,0170
25	98,60%	0,0705	32	99,30%	0,0242
50	99,30%	0,0242	64	99,20%	0,0181
75	99,10%	0,0341	128	98,70%	0,0423

Pada hasil pengujian skenario 4, *batch size* tertinggi berada pada nilai 16 dengan akurasi sebesar 99,50% dan *loss* 0,0170. Namun, seperti pada Gambar 6 grafik akurasi *batch size* 16 masih menunjukkan munculnya banyak *spikes*/lonjakan sehingga *batch size* terbaik berada pada nilai 32. Jadi, pada tahap pengujian dengan *pre-processing* CLAHE menggunakan *hyperparameter* terbaik yaitu *optimizer* SGD, *learning rate* 0,001, *epoch* 50 dan *batch size* 32 didapatkan nilai akurasi sebesar 99,30% dan nilai *loss* 0,0242.



Gambar 6. Grafik Akurasi Pengujian *Batch Size* (a) 16 dan (b) 32

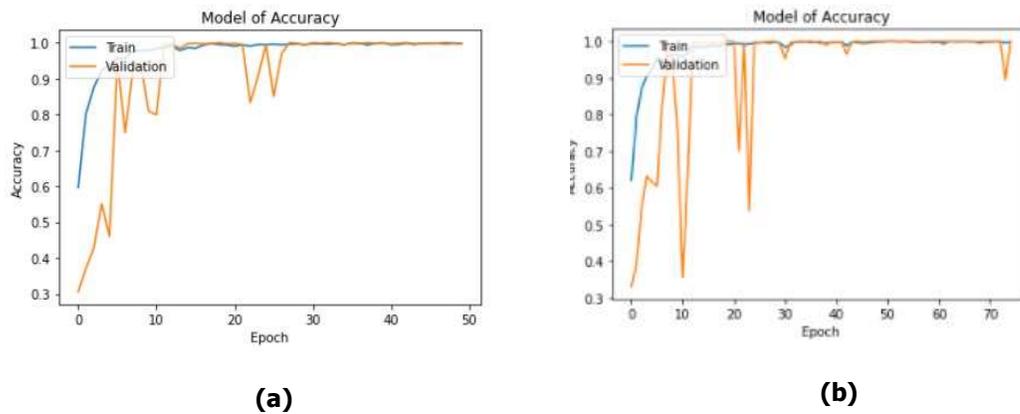
Pengujian berikutnya ialah pengaruh *pre-processing* Gaussian filtering. Dalam pengujian dengan *pre-processing* Gaussian filter dilakukan 4 skenario pengujian terhadap *hyperparameter*, yaitu *optimizer*, *learning rate* (LR), *epoch* dan *batch size* (BC). Hasil skenario pengujian ditunjukkan pada Tabel 4 berikut.

Tabel 4. Hasil pengujian dengan *pre-processing* Gaussian filter

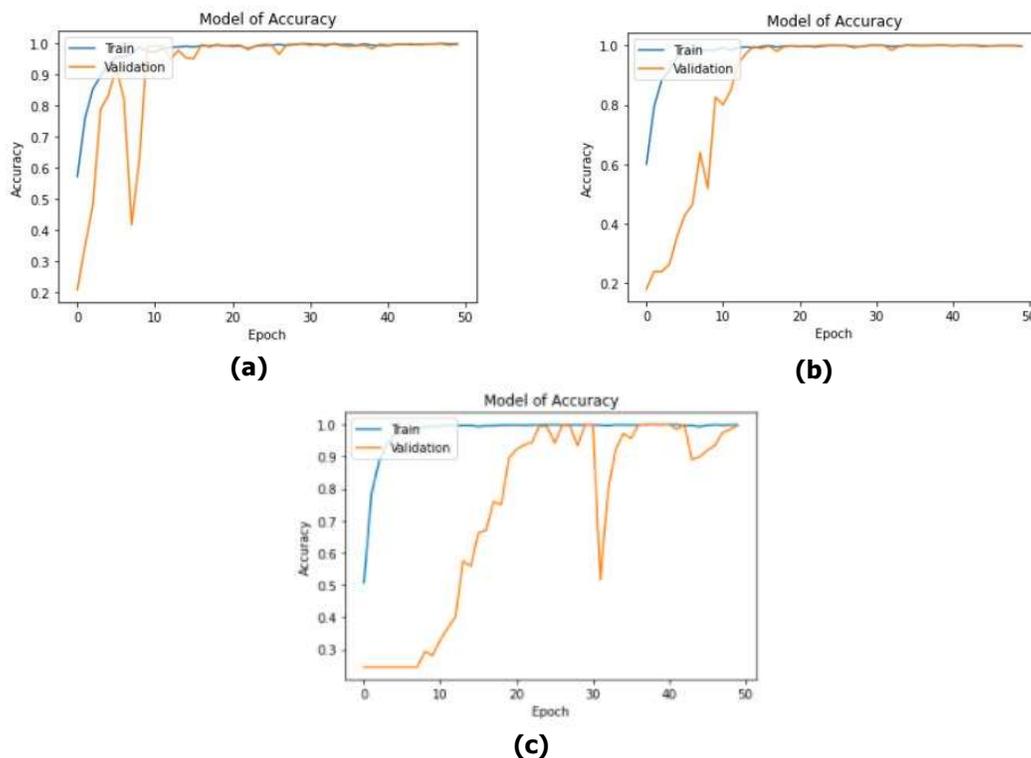
Skenario 1: <i>Optimizer</i> (LR 0,001, <i>epoch</i> 15, BC 32)			Skenario 2: <i>Learning rate</i> ( <i>Optimizer</i> SGD, <i>epoch</i> 15, BC 32)		
<i>Optimizer</i>	Akurasi	<i>Loss</i>	<i>Learning Rate</i>	Akurasi	<i>Loss</i>
Adam	86.50%	0.9560	0,1	46,20%	1,5528
Nadam	81.80%	0.4920	0,01	98,90%	0,0342
RMSprop	88.50%	0.2015	0,001	98,00%	0,0600
SGD	98.00%	0.0600	0,0001	94,50%	0,1498
Skenario 3: <i>Epoch</i> ( <i>Optimizer</i> SGD, LR 0,01, BC 32)			Skenario 4: <i>Batch Size</i> ( <i>Optimizer</i> SGD, LR 0,01, <i>epoch</i> 50)		
<i>Epoch</i>	Akurasi	<i>Loss</i>	<i>Batch size</i>	Akurasi	<i>Loss</i>
15	98,90%	0,0342	16	99,40%	0,0157
25	94,90%	0,1149	32	99,70%	0,0119
50	99,60%	0,0125	64	99,60%	0,0129
75	100%	8,2328e-04	128	99,70%	0,0195

Hasil pengujian skenario ketiga menunjukkan nilai akurasi *epoch* tertinggi berada pada nilai 75 dengan akurasi sebesar 100% dengan *loss* 8,2328e-04. Namun, grafik akurasi *epoch* 75 menunjukkan munculnya banyak *spikes*/lonjakan dan *loss* meningkat pada *epoch* tertentu (Gambar 7) dengan *cost* yang tinggi. Sehingga *epoch* yang dipilih dengan performa terbaik dan *cost* yang rendah adalah *epoch* 50. Sedangkan untuk skenario 4, hasil pengujian tertinggi berada pada nilai 32 dan 128, namun pada grafiknya (Gambar 8) menunjukkan penurunan data validasi secara mendadak di beberapa *epoch* tertentu yang mengakibatkan munculnya banyak *spikes*/lonjakan maka *batch size* yang dipilih adalah 64. Sehingga pada pengujian dengan *pre-processing* Gaussian filter menggunakan *hyperparameter* terbaik, yaitu *optimizer*

SGD, *learning rate* 0,01, *epoch* 50 dan *batch size* 64 didapatkan hasil akhir nilai akurasi sebesar 99,60% dan *loss* 0,0129.



**Gambar 7. Grafik Pengujian *Epoch* (a) 50 dan (b) 75**



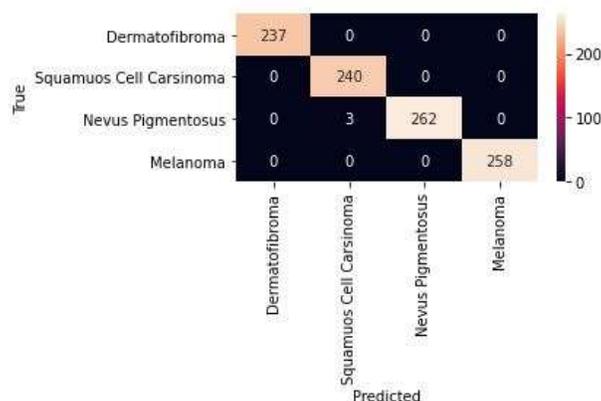
**Gambar 8. Grafik Akurasi Pengujian *Batch Size* (a) 32, (b) 64 dan (c) 128**

Berdasarkan hasil akhir skenario pengujian yang ditunjukkan pada Tabel 2, 3 dan 4 hasil optimal berada pada pengujian tanpa *pre-processing* CLAHE dan Gaussian *filter* dibandingkan pengujian lainnya. Dari hasil skenario pengujian *hyperparameter optimizer*, menunjukkan bahwa optimizer SGD dianggap paling baik dalam mengklasifikasi kanker kulit menggunakan metode CNN arsitektur VGG-16. Pada *optimizer* Adam, Nadam dan RMSprop, *optimizer* tersebut melakukan penyimpanan rata-rata gradien dari proses sebelumnya secara eksponensial, yang menyebabkan keterlambatan proses kinerja dibanding SGD. Sedangkan SGD tidak melakukan pengulangan sehingga proses kinerja lebih cepat terutama untuk jumlah data yang besar.

Pada hasil skenario *hyperparameter learning rate*, penggunaan *learning rate* yang terlalu kecil belum tentu membuat proses ketelitian jaringan menjadi bertambah, tetapi menyebabkan proses *training* yang terlalu lama. Sedangkan nilai *learning rate* yang terlalu besar pun menyebabkan pembelajaran kurang optimal karena waktu yang terlalu cepat sehingga proses *training* tidak stabil. Sehingga pada kasus penelitian ini nilai 0,001 menunjukkan hasil paling optimal.

Dari hasil pengujian skenario *hyperparameter epoch* menunjukkan bahwa semakin besar jumlah *epoch* belum tentu menghasilkan performansi yang optimal. Pada penelitian ini, *epoch* yang besar cenderung menunjukkan penurunan secara mendadak disetiap *epoch* tertentu yang mengakibatkan munculnya *spikes*/lonjakan pada grafik akurasi. Maka pada penelitian ini *epoch* terbaik adalah 50, karena *cost* yang rendah dan kerja model yang tidak terlalu berat tetapi sudah menunjukkan hasil yang optimal.

Untuk hasil pengujian skenario *hyperparameter batch size*, penggunaan *batch size* yang lebih kecil memiliki kecepatan konvergensi dalam algoritma yang lebih cepat tetapi menghasilkan *noise* dalam perhitungan yang lebih besar, sedangkan penggunaan *batch size* yang lebih besar dapat membantu mengurangi *noise* tetapi tidak menjamin peningkatan pada nilai akurasi. Sehingga pada penelitian ini, *batch size* yang menghasilkan nilai optimal adalah *batch size* 32.



**Gambar 9. Confusion Matrix**

Gambar 9 menunjukkan *confusion matrix* dari hasil akhir pengujian performa terbaik. Kesalahan terjadi pada tiga gambar kelas karsinoma sel skuamosa yang terdeteksi sebagai nevus pigmentosus. Parameter lain seperti presisi, *recall*, dan *f1-score* juga digunakan untuk mengevaluasi kinerja sistem. Hasil dari nilai parameter performansi sistem tersebut mendekati satu, yang berarti sistem menunjukkan performansi terbaik dengan minimum kesalahan.

#### 4. KESIMPULAN

Berdasarkan hasil pengujian, penelitian ini mampu merancang sistem yang dapat mengklasifikasi kanker kulit menggunakan metode CNN dengan arsitektur VGG-16. Sistem dapat mengklasifikasi kanker kulit menjadi 4 kelas yaitu, melanoma, karsinoma sel skuamosa, dermatofibroma dan nevus pigmentosus. Didalam pengujian sistem dilakukan skenario pengujian dengan dan tanpa *pre-processing* dan skenario terhadap 4 *hyperparameter* yaitu *optimizer*, *learning rate*, *epoch* dan *batch size*. Kemudian hasil akhir dengan performa dan akurasi terbaik diperoleh dari hasil pengujian tanpa *pre-processing* CLAHE dan Gaussian *filter*.

Citra di-*resize* berukuran 64×64 piksel dan menggunakan *hyperparameter optimizer* SGD, *learning rate* 0,001, *epoch* 50 dan *batch size* 32. Nilai akurasi diperoleh sebesar 99,70%, *loss* 0,0055, *presisi* 0,9975, *recall* 0,9975 dan *f1-score* 0,9950. Peningkatan nilai akurasi sistem dibandingkan dengan penelitian sebelumnya ini menunjukkan bahwa arsitektur VGG-16 memiliki kemampuan yang tepat dan detail dalam mengenali dan mengklasifikasikan citra. Sistem mampu digunakan sebagai alat bantu tenaga medis untuk mengurangi risiko keterlambatan dalam penanganan kanker kulit.

## DAFTAR RUJUKAN

- Bera, S., & Shrivastava, V. K. (2020). Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification. *International Journal of Remote Sensing*, 41(7), 2664–2683. <https://doi.org/10.1080/01431161.2019.1694725>
- Bruballa, R. G. (2018). *Understanding Categorical Cross-Entropy Loss, Binary Cross-Entropy Loss, Softmax Loss, Logistic Loss, Focal Loss and all those confusing names*. Retrieved from [https://gombu.github.io/2018/05/23/cross\\_entropy\\_loss/](https://gombu.github.io/2018/05/23/cross_entropy_loss/)
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118. <https://doi.org/10.1038/nature21056>
- Fu'adah, Y. N., Pratiwi, N. C., Pramudito, M. A., & Ibrahim, N. (2020). Convolutional neural network (CNN) for automatic skin cancer classification system. *IOP Conference Series: Materials Science and Engineering*, 982(1), 0–10. <https://doi.org/10.1088/1757-899X/982/1/012005>
- Hendaria, M. P., Asmarajaya, A., & Maliawan, S. (2013). *Kanker kulit*. Universitas Udayana.
- Munthe, T. L. D. (2018). *Klasifikasi citra kanker kulit berdasarkan tingkat keganasan kanker pada melanosit menggunakan deep convolutional neural network (DCNN)*. Repositori Institusi Universitas Sumatera Utara (RI-USU), 44–48.
- Prabowo, D., & Pramunendar, R. (2019). Implementasi metode clahe menggunakan parameter distribusi untuk meningkatkan kualitas citra objek bawah air. *Prosiding SNST Ke-10*, (pp. 264–269).
- Prasetyo, E. (2012). *Data Mining: Konsep dan aplikasi menggunakan MATLAB* (1st ed.). Yogyakarta Andi Offset.
- Pratiwi, N. K. C., Fu'adah, Y. N., & Edwar. (2021). Early detection of deforestation through satellite land geospatial images based on CNN architecture. *Jurnal Infotel*, 13(2), 1–8.
- Savera, T. R., Suryawan, W. H., & Setiawan, A. W. (2020). *Deteksi dini kanker kulit menggunakan K-Nn dan early detection of skin cancer using K-Nn and convolutional*. 7

- (2), 373–378. <https://doi.org/10.25126/jtiik.202072602>
- Silpa, S. R., & V, C. (2013). A review on skin cancer. *International Research Journal of Pharmacy*, 4(8), 83–88. <https://doi.org/10.7897/2230-8407.04814>
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, (pp. 1–14).
- Tsaniyah, R. A. D., Aspitriani, & Fatmawati. (2015). Prevalensi dan Gambaran Histopatologi Nevus Pigmentosus di Bagian Patologi Anatomi Rumah Sakit Dr . Mohammad Hoesin Palembang Periode 1 Januari 2009-31 Desember 2013. *Mks*, 2, 110–114. <https://media.neliti.com/media/publications/181832-ID-prevalensi-dan-gambaran-histopatologi-ne.pdf>
- Wang, M., Zheng, S., Li, X., & Qin, X. (2014). A new image denoising method based on the bilateral filter. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, (pp. 163–167). <https://doi.org/10.1109/ICASSP.2008.4517763>
- Wilvestra, S., Lestari, S., & Asri, E. (2018). Studi retrospektif kanker kulit di poliklinik ilmu kesehatan kulit dan kelamin RS Dr. M. Djamil Padang periode tahun 2015-2017. *Jurnal Kesehatan Andalas*, 7, 47. <https://doi.org/10.25077/jka.v7i0.873>