

TESIS

**PENGARUH PROSES AUGMENTASI TERHADAP MODEL TRANSFER
LEARNING RESNET50 PADA KLASIFIKASI KANKER MELANOMA**



Disusun oleh:

Nama : Ahmad Febri Diansyah
NIM : 22.51.1188
Konsentrasi : Business Intelligence

PROGRAM STUDI S2 INFORMATIKA
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2024

TESIS

**PENGARUH PROSES AUGMENTASI TERHADAP MODEL TRANSFER
LEARNING RESNET50 PADA KLASIFIKASI KANKER MELANOMA**

**EFFECT OF AUGMENTATION PROCESS ON THE RESNET50
TRANSFER LEARNING MODEL IN MELANOMA CANCER
CLASSIFICATION**

Diajukan untuk memenuhi salah satu syarat memperoleh derajat Magister



Disusun oleh:

Nama : Ahmad Febri Dlansyah
NIM : 22.51.1188
Konsentrasi : Business Intelligence

PROGRAM STUDI S2 INFORMATIKA
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2024

HALAMAN PENGESAHAN

**PENGARUH PROSES AUGMENTASI TERHADAP MODEL TRANSFER
LEARNING RESNET50 PADA KLASIFIKASI KANKER MELANOMA**

**EFFECT OF AUGMENTATION PROCESS ON THE RESNET50 TRANSFER
LEARNING MODEL IN MELANOMA CANCER CLASSIFICATION**

Dipersiapkan dan Disusun oleh

Ahmad Febri Diansyah

22.51.1188

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tesis
Program Studi S2 Informatika
Program Pascasarjana Universitas AMIKOM Yogyakarta
pada hari Selasa, 9 Juli 2024

Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer

Yogyakarta, 9 Juli 2024
Rektor

Prof. Dr. M. Suyanto, M.M.
NIK. 190302001

HALAMAN PERSETUJUAN

**PENGARUH PROSES AUGMENTASI TERHADAP MODEL TRANSFER
LEARNING RESNET50 PADA KLASIFIKASI KANKER MELANOMA**

**EFFECT OF AUGMENTATION PROCESS ON THE RESNET50 TRANSFER
LEARNING MODEL IN MELANOMA CANCER CLASSIFICATION**

Dipersiapkan dan Disusun oleh

Ahmad Febri Diansyah

22.51.1188

Telah Diujikan dan Dipertahankan dalam Sidang Ujian Tesis
Program Studi S2 Informatika
Program Pascasarjana Universitas AMIKOM Yogyakarta
pada hari Selasa, 9 Juli 2024

Pembimbing Utama

Anggota Tim Penguji

Prof. Dr. Ema Utami, S.Si., M.Kom
NIK. 190302037

M. Hanafi, S.Kom., M.Eng., Ph.D.
NIK. 190302024

Pembimbing Pendamping

Dr. Kumara Ari Ynana, S.T., M.T.
NIK. 190302575

Dhani Ariatmanto, M.Kom., Ph.D
NIK. 190302197

Prof. Dr. Ema Utami, S.Si., M.Kom.
NIK. 190302037

Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer

Yogyakarta, 9 Juli 2024

Direktur Program Pascasarjana

Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : Ahmad Febri Diansyah
NIM : 22.51.1188
Konsentrasi : Business Intelligence

Menyatakan bahwa Tesis dengan judul berikut:
Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 pada Klasifikasi Kanker Melanoma

Dosen Pembimbing Utama : Prof. Dr. Ema Utami, S.Si., M.Kom
Dosen Pembimbing Pendamping : Dhani Ariatnanto, M.Kom., Ph.D

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian **SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta
5. Pernyataan ini **SAYA** buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi

Yogyakarta, 9 Juli 2024
Yang Menyatakan,



Ahmad Febri Diansyah

Ahmad Febri Diansyah

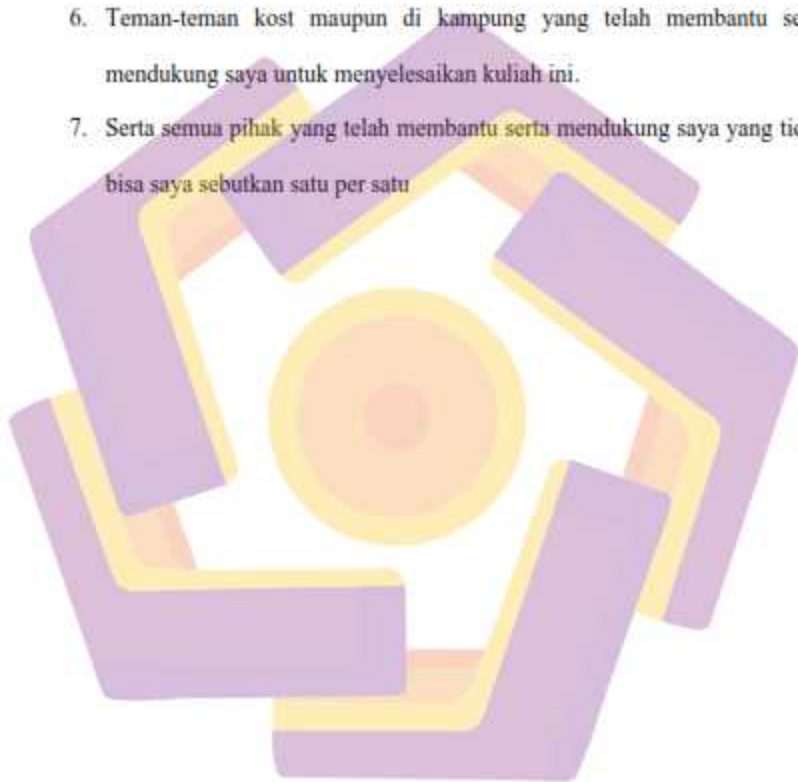
HALAMAN PERSEMBAHAN

Puji Syukur saya panjatkan kehadirat Allah SWT yang telah memberikan nikmat dan berkat yang luar biasa kepada saya, sehingga saya bisa menyelesaikan tesis ini dengan baik. Saya juga sangat berterima kasih kepada orang-orang yang secara langsung maupun tidak langsung telah membantu saya dalam menyelesaikan tesis ini. Tesis ini saya persembahkan kepada:

1. Ayah dan Ibu saya yang selalu mendokakan saya, selalu mendukung baik finansial maupun dukungan lainnya. Terima kasih sudah mengorbankan banyak hal untuk keberhasilan ini.
2. Ibu Prof. Dr. Ema Utami, S.Si., M.Kom dan Bapak Dhani Ariatmanto, M.Kom, Ph.D selaku dosen pembimbing yang selalu memberikan masukan serta bimbingan yang positif dalam menyelesaikan tesis ini. Terima kasih juga atas ilmu yang telah diberikan.
3. Bapak dan Ibu Dosen yang selalu memberikan ilmu yang bermanfaat selama saya kuliah.
4. Teman-teman Angkatan 28 untuk memori indah yang pernah kita rajut bersama selama perkuliahan. Terima kasih atas bantuan dan ilmu yang pernah kalian bagi.
5. Kepada seseorang yang namanya tak lagi bisa saya sebutkan. Meski kini kita berada di jalan yang berbeda, saya tak akan pernah melupakan dukungan dan kasih sayang yang pernah Anda berikan. Setiap bantuan, semangat, dan kehadiran Anda menjadi fondasi penting dalam perjalanan

akademis saya. Berkat kebaikan hati Anda, saya berhasil melewati banyak tantangan dan mencapai titik ini. Meskipun waktu telah memisahkan kita, kenangan akan momen-momen indah yang pernah kita bagi tetap terukir dalam ingatan, dan untuk itu, saya selamanya bersyukur.

6. Teman-teman kost maupun di kampung yang telah membantu serta mendukung saya untuk menyelesaikan kuliah ini.
7. Serta semua pihak yang telah membantu serta mendukung saya yang tidak bisa saya sebutkan satu per satu



HALAMAN MOTTO

"The first cause of Absurd conclusions I ascribe to the want of Method"

Penyebab pertama dari kesimpulan absurd, adalah kurangnya metode

Thomas Hobbes

"God does not play dice with the universe"

Tuhan tidak bermain dadu dengan alam semesta.

Albert Einstein

"A computer would deserve to be called intelligent if it could deceive a human into believing that it was human"

Sebuah komputer layak disebut cerdas jika ia dapat menipu manusia untuk mempercayai bahwa ia adalah manusia

Alan Turing

KATA PENGANTAR

Puji Syukur kehadiran Allah SWT yang telah melimpahkan rahmat dan hidayah sehingga penulis dapat menyelesaikan tugas akhir dengan judul " Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 pada Klasifikasi Kanker Melanoma" dengan baik dan sesuai waktu yang diharapkan.

Tesis ini untuk memenuhi salah satu syarat untuk memperoleh gelar Magister Informatika di Universitas Amikom Yogyakarta. Penulis menyadari bahwa penulisan ini tidak akan terwujud tanpa adanya dukungan, bantuan dari berbagai pihak. Untuk itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada yang terhormat:

1. Kedua Orang tua saya, Ayah Sukoco dan Ibu Nanik Lestari yang telah memberikan saya semangat untuk segera menyelesaikan kuliah S2 saya ini.
2. Adik saya, Masryagita Putri Utari yang telah memberikan saya dukungan untuk menyelesaikan kuliah S2 ini.
3. Ibu Prof. Dr. Kusriani, M.Kom selaku Direktur Program Pascasarjana.
4. Ibu Prof. Dr. Ema Utami, S.Si., M.Kom selaku Pembimbing Utama saya yang telah membimbing saya sehingga terselesaikannya naskah tesis ini.
5. Bapak Dhani Ariatmanto, M.Kom, Ph.D selaku Pembimbing Pendamping saya yang telah membimbing saya sehingga terselesaikannya naskah tesis ini.
6. Bapak Dr. Ferry Wahyu Wibowo, S.Si., M.Cs. dan M. Rudyanto Arief, M.T. selaku Penguji Seminar Proposal Tesis.

7. Bapak Dr. Kumara Ari Yuana, S.T., M.T. dan Emha Taufiq Luthfi, S.T., M.Kom. selaku Penguji Seminar Hasil Proposal Tesis dan Ujian Tesis.
8. Bapak M. Hanafi, S.Kom., M.Eng., Ph.D. dan Dr. Kumara Ari Yuana, S.T., M.T.. selaku Penguji Ujian Tesis.
9. Bapak dan Ibu Dosen Universitas Amikom Yogyakarta yang telah memberikan ilmu, pengetahuan, motivasi, dan pengalaman setiap mengajar selama penulis menempuh kuliah.
10. Kepada seseorang yang tak lagi bisa saya sebutkan. Berkatmu, aku bisa mencapai titik ini dalam perjalanan akademisku. Kenangan indah bersamamu akan selalu ada dalam ingatan, dan aku tetap bersyukur untuk itu.
11. Teman-teman Angkatan 28 yang telah menemani penulis selama masa perkuliahan dan berbagi canda tawa bersama.

Penulis menyadari bahwa dalam pembuatan tesis ini masih banyak kekurangan dan kelemahannya. Oleh karena itu penulis berharap kepada semua pihak agar dapat menyampaikan kritik dan saran yang membangun untuk menambah kesempurnaan tesis ini. Namun, penulis tetap berharap tesis ini akan bermanfaat bagi semua pihak yang membacanya.

Yogyakarta, 9 Juli 2024

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PENGESAHAN.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PERNYATAAN KEASLIAN TESIS.....	v
HALAMAN PERSEMBAHAN.....	vi
HALAMAN MOTTO.....	viii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xv
INTISARI.....	xvi
<i>ABSTRACT</i>	xvii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah.....	5
1.3. Batasan Masalah.....	5
1.4. Tujuan Penelitian.....	6
1.5. Manfaat Peneliti.....	7
BAB II TINJAUAN PUSTAKA.....	8
2.1. Tinjauan Pustaka.....	8
2.2. Keaslian Penelitian.....	11

2.3. Landasan Teori.....	22
2.3.1. Kanker Melanoa	22
2.3.2. Klasifikasi.....	25
2.3.3. Augmentasi.....	26
2.3.4. Oversampling	27
2.3.5. Gaussian Blur.....	28
2.3.6. Resnet-50.....	29
2.3.7. Confussion Matrix.....	32
BAB III METODE PENELITIAN.....	35
3.1. Jenis, Sifat, dan Pendekatan Penelitian.....	35
3.2. Metode Pengumpulan Data.....	35
3.3. Metode Analisis Data.....	35
3.4. Alur Penelitian	36
BAB IV HASIL PENELITIAN DAN PEMBAHASAN.....	40
4.1. Membangun Dataset	40
4.4.1. Dataset.....	40
4.4.2. Persebaran Data	41
4.2. Analisis Data.....	42
4.3. Preprocessing.....	46
4.3.1. Pengelompokan Data.....	46
4.3.2. Preprocessing.....	46

4.3.3. Augmentasi.....	47
4.4. Membangun Model dan Implementasi	52
4.5. Hasil Analisis dan Pembahasan	53
4.5.1. Hasil Pelatihan Dari Model	53
4.5.2. Hasil Perbandingan Pengujian.....	62
4.5.3. Faktor yang Mempengaruhi Hasil.....	65
4.5.4. Perbandingan dengan Penelitian Sebelumnya.....	69
4.5.5. Upaya Perbaikan Kinerja Algoritma.....	72
BAB V PENUTUP.....	75
5.1. Kesimpulan	75
5.2. Saran	76
DAFTAR PUSTAKA	79

DAFTAR TABEL

Tabel 1.1. Penelitian Terdahulu yang Menggunakan Proses Augmentasi.....	3
Tabel 2.1. Matriks literature review Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 Pada Klasifikasi Kanker Melanoma	11
Tabel 4.1. Detail Dataset Kanker Melanoma	41
Tabel 4.2. Pembagian Data Augmentasi Normal.....	48
Tabel 4.3. Pembagian Data Augmentasi + Oversampling.....	49
Tabel 4.4. Pembagian Data Augmentasi + Gaussian Blur.....	51
Tabel 4.5. Hasil Laporan Skenario Pertama.....	54
Tabel 4.6. Hasil Laporan Skenario Kedua.....	57
Tabel 4.7. Hasil Laporan Skenario Ketiga.....	59
Tabel 4.8. Hasil Akurasi Tiap Skenario.....	62
Tabel 4.9. Perbandingan Penelitian Sebelumnya.....	70

DAFTAR GAMBAR

Gambar 2.1 Melanoma Benign	23
Gambar 2.2 Melanoma Malignant	23
Gambar 2.3 Arsitektur Resnet50.....	32
Gambar 2.4 Confusion Matrix	33
Gambar 3.1 Alur Penelitian.....	37
Gambar 3.2 Alur Penelitian dengan Oversampling	38
Gambar 3.3 Alur Penelitian dengan Gaussian Blur	39
Gambar 4.1 Persebaran Jumlah Data Kanker Melanoma	42
Gambar 4.2 Hasil Evaluasi Matrix Tanpa Augmentasi	54
Gambar 4.3 Hasil Akurasi Tiap Skenario	64
Gambar 4.4 Perbandingan Train Loss dan Validation Loss pada skenario 1 Augmentasi Flip Horizontal	83
Gambar 4.5 Perbandingan Train Loss dan Validation Loss pada skenario 3 Augmentasi Flip Horizontal	83
Gambar 4.6 Confusion Matix Skenario 1 Augmentasi Flip Horizontal	85
Gambar 4.7 Confusion Matix Skenario 3 Augmentasi Flip Horizontal	85

INTISARI

Kanker melanoma merupakan salah satu jenis kanker kulit yang paling mematikan dan memerlukan diagnosis dini untuk meningkatkan peluang kesembuhan. Dalam konteks ini, penelitian ini bertujuan untuk mengevaluasi dampak proses augmentasi data terhadap kinerja model ResNet50 dalam mengklasifikasikan kanker melanoma. Penelitian ini membatasi variabelnya pada penggunaan dataset Melanoma Detection dari Kaggle yang terdiri dari gambar benign dan malignant. Metode analisis yang digunakan adalah analisis performansi dengan metrik akurasi, presisi, recall, dan F1-score, serta analisis visualisasi menggunakan confusion matrix.

Hasil penelitian menunjukkan bahwa kemampuan model ResNet50 untuk mengklasifikasikan kanker melanoma lebih baik dengan menambahkan data. Analisis visualisasi menunjukkan bahwa kombinasi augmentasi dengan rotasi, zoom, dan flip menghasilkan hasil yang lebih akurat dibandingkan dengan tanpa augmentasi. Selain itu, analisis skenario augmentasi yang diuji menunjukkan perbedaan yang signifikan dalam performa model setelah dilakukan augmentasi, dengan heatmap dan confusion matrix menunjukkan distribusi prediksi yang lebih baik.

Penelitian ini menemukan bahwa peningkatan data dapat meningkatkan kinerja model ResNet50 dalam tugas klasifikasi kanker melanoma. Peningkatan data membantu model mengenali pola-pola pada gambar melanoma, meningkatkan akurasi, presisi, recall, dan skor F1-model. Penelitian ini membuat kemajuan signifikan dalam penggunaan teknik augmentasi untuk meningkatkan kinerja model deep learning dalam tugas klasifikasi gambar.

Kata kunci: Augmentasi Data, ResNet50, Kanker Melanoma, Klasifikasi Gambar, Deep Learning.

ABSTRACT

Melanoma cancer is one of the deadliest types of skin cancer and requires early diagnosis to increase the chances of cure. In this context, this study aims to evaluate the impact of the data augmentation process on the performance of the ResNet50 model in classifying melanoma cancer. This study limits its variables to the use of Kaggle's Melanoma Detection dataset consisting of benign and malignant images. The analysis methods used are performance analysis with accuracy, precision, recall, and F1-score metrics, and visualization analysis using confusion matrix.

The results show that the ability of the ResNet50 model to classify melanoma cancer is better with augmented data. Visualization analysis showed that the combination of augmentation with rotation, zoom, and flip produced more accurate results compared to no augmentation. In addition, analysis of the augmentation scenarios tested showed significant differences in model performance after augmentation, with the heatmap and confusion matrix showing better prediction distribution.

This study found that data enhancement can improve the performance of the ResNet50 model in melanoma cancer classification tasks. Data augmentation helps the model recognize patterns in melanoma images, improving the accuracy, precision, recall, and F1-model score. This research makes significant progress in the use of augmentation techniques to improve the performance of deep learning models in image classification tasks.

Keywords: Data Augmentation, ResNet50, Melanoma Cancer, Image Classification, Deep Learning.

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Kanker adalah salah satu penyakit paling serius yang dapat menyebabkan kematian di seluruh dunia. Menurut Organisasi Kesehatan Dunia, hingga 160.000 orang di seluruh dunia terkena kanker kulit setiap tahun (Milton, 2019). Australia memiliki salah satu tingkat kanker kulit tertinggi, empat kali lipat dari Amerika Serikat. Berdasarkan data dari Biro Statistik Australia, 32,6% dari semua orang Australia yang menderita kanker menderita kanker kulit. 971.279 kasus kanker kulit didiagnosis ditahun 2012, 2.162 di antaranya berakibat fatal. Sedangkan menurut Centers for Disease Control and Prevention, pada tahun 2015 terdapat 80.442 kasus kanker kulit dan 8.885 kematian di Amerika Serikat. Relatif sedikit penderita kanker kulit di Indonesia dibandingkan ketiga negara tersebut. Namun, perhatian harus diberikan pada kanker kulit ini sedini mungkin, karena tidak hanya dapat menyebabkan kelainan bentuk yang dapat mempengaruhi penampilan, tetapi juga dapat menyebabkannya. kematian stadium lanjut (Bisla et al., 2019). Menurut kutipan dari Pusat Penelitian Sumber Daya Kesehatan dan Pelayanan Kesehatan Indonesia, terdapat sekitar 14 juta kasus kanker baru dan 8,2 juta kematian akibat kanker di seluruh dunia pada tahun 2012.

Proses augmentasi sering digunakan para peneliti terdahulu untuk memberikan variasi dataset yang mereka punya. Augmentasi gambar adalah proses yang digunakan untuk meningkatkan kualitas atau kuantitas dari data gambar yang

digunakan dalam pembelajaran mesin (Ghosh et al., 2020). Dalam proses augmentasi, gambar asli diubah dengan berbagai transformasi seperti *rotasi*, *flip*, *zoom*, *crop*, translasi, perubahan *brightness* dan *contrast*, perubahan warna, penambahan noise dll. (Bisla et al., 2019). Seperti yang dilakukan oleh (Milton, 2019), dia menggunakan banyak sekali proses augmentasi seperti *rotation*, *flip*, *random crop*, *adjust brightness*, *adjust contrast*, *pixel jitter*, *Aspect Ratio*, *random shear*, *zoom*, dan *vertical* dan *horizontal shift* dengan tingkat akurasi yang didapat sebesar 76%. Ada juga penelitian yang dilakukan oleh (Teresia R. Savera, Winsya H. Suryawan, 2020) menggunakan proses augmentasi *vertical flipping* dan *random cropping* dengan tingkat akurasi sebesar 82%. Namun dari kedua penelitian tersebut tidak menjelaskan apakah dengan proses augmentasi bisa mempengaruhi tingkat akurasi. Pada tabel 1 merupakan rangkuman beberapa penelitian kanker melanoma yang menggunakan proses augmentasi dan bisa menjadi acuan augmentasi apa saja yang bisa digunakan dalam penelitian kali ini.

Tabel 1.1. Penelitian Terdahulu yang Menggunakan Proses Augmentasi

Penelitian	Rotation	Crop	Shear	Noise	Zoom	Shift	Flip	Saturation	Contrast	Brightness	Color	Sprinkled Cutout	Pixel Jitter	Aspect Ratio	Microscopic
(Adepu et al., 2023)	x		x		x	x	x	x	x	x		x			x
(López de la Rosa et al., 2022)	x				x	x	x								
(Bansal et al., 2022)	x			x	x		x		x	x					
(Nancy Jane et al., 2023)	x			x			x								
(Bisla et al., 2019)	x	x			x	x	x		x	x	x				
(Milton, 2019)	x	x	x		x	x	x		x	x			x	x	

Dari Tabel 1.1 bisa dilihat proses augmentasi yang sering digunakan adalah Rotation, Shift, Flip, Contrast dan Brightness.

Klasifikasi citra merupakan metode yang dapat mengklasifikasikan citra menurut kategori tertentu(Zhang et al., 2020). Convolutional neural network (CNN) merupakan algoritma yang sering dipakai akhir-akhir ini untuk proses klasifikasi gambar(Kutlu et al., 2020). Melihat hal tersebut tentunya banyak para peneliti berlomba-lomba untuk menciptakan CNN lain yang bisa disebut sebagai Transfer Learning dalam upaya mencari akurasi terbaik dengan komputasi yang lebih cepat(Soegeng et al., 2021). Algoritma yang berhasil diciptakan para peneliti antara lain XceptionNet, VGG16, Resnet50 dan masih banyak lagi. Dengan adanya ini tentunya menguntungkan banyak peneliti selanjutnya untuk mencoba berbagai macam uji coba metode(Wolf et al., 2019).

Pada penelitian kali ini akan menggunakan salah satu model transfer learning yaitu Resnet50. ResNet50 adalah deep convolutional neural network yang ditraining pada kumpulan data ImageNet dan banyak digunakan untuk klasifikasi gambar(Pangestu & Bunyamin, 2018). Keuntungan menggunakan model ini adalah akurasinya yang tinggi dan menggunakan koneksi residu yang memungkinkan untuk mempelajari fungsi yang lebih kompleks(He et al., 2016). Tidak ada model yang sempurna oleh sebab itulah Resnet50 mempunyai kelemahan yaitu ukuran yang besar, fleksibilitas yang terbatas dan kemungkinan mengandung bias pada dataset yang dilatih(He et al., 2016). Kelebihan dan kelemahan tersebut membuat model transfer learning ResNet50 dari segi kelemahan model ini bisa digunakan. Melihat penelitian sebelumnya (Saputro et al., 2022; Teresia R. Savera, Winsya H.

Suryawan, 2020), mereka mengusulkan untuk menggunakan dataset yang tidak banyak variasi dan menggunakan model CNN.

Dari penelitian ini akan diketahui apakah dengan menggunakan proses augmentasi bisa mempengaruhi tingkat akurasi dan apakah algoritma ResNet50 dapat digunakan untuk mengklasifikasikan kanker melanoma atau tidak. Penggunaan proses augmentasi dan algoritma Resnet50 akan memberikan akurasi dalam klasifikasi kanker melanoma berdasarkan data yang telah tersedia dari dataset yang akan digunakan dalam pemrosesan dan pemodelan.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka perlu rumusan masalah dalam penelitian ini yaitu :

- a. Bagaimana pengaruh augmentasi data terhadap akurasi, presisi, recall, dan f1-score model ResNet50 dalam tugas klasifikasi gambar?
- b. Manakah skenario augmentasi yang menghasilkan performa terbaik (akurasi, presisi, recall, dan f1-score) untuk model ResNet50 dalam tugas klasifikasi gambar?

1.3. Batasan Masalah

Bagian Batasan masalah dalam Penelitian ini adalah:

- a. Dataset yang digunakan berasal dari Kaggle dengan nama Melanoma Detection Dataset dari Pablo Lopez Santori (ISIC 2017).
- b. Dataset yang diperoleh berupa gambar dengan format JPG.

- c. Tipe kanker kulit tentunya ada berbagai macam seperti basal cell carcinoma, squamous cell carcinoma, melanoma, merkel cell carcinoma dan kaposi sarcoma, pada penelitian ini hanya melanoma (malignant dan benign) yang dipilih.
- d. Preprocessing data menggunakan proses augmentasi yaitu Rotation, Zoom dan Flip.
- e. Parameter augmentasi yang digunakan antara lain:
 - Rotation: 45
 - Zoom: 50
 - Flip: Horizontal dan Vertical
- f. Platform yang digunakan adalah Google Colaboratory dan Jupyter Lab.
- g. Framework yang digunakan adalah Tensorflow.
- h. Tool/Library yang digunakan adalah Numpy, Pandas, Matplotlib, Seaborn, Sklearn dan Keras.
- i. Algoritma yang digunakan adalah ResNet50.
- j. Pengujian model menggunakan Confussion Matrix

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah :

- a. Menguji apakah dengan menggunakan proses augmentasi akan mempengaruhi tingkat akurasi, presisi, recall dan f1-score.
- b. Mencari skenario augmentasi yang mempunyai performa terbaik (akurasi, presisi, recall, dan f1-score).

1.5. Manfaat Penelitian

a. Manfaat Teoritis :

1. Mampu mengembangkan metode Resenet50 untuk membuat klasifikasi kanker melanoma.
2. Mengetahui pengaruh proses augmentasi dalam model Resnet50.
3. Mampu untuk menganalisis dan membahas bagaimana kerja metode Resnet50 yang diusulkan.

b. Manfaat Praktis :

1. Memudahkan seseorang untuk mengidentifikasi penyakit kanker melanoma.
2. Penanganan kanker melanoma akan menjadi lebih cepat.
3. Membantu dalam mengembangkan terapi baru yang lebih efektif untuk menangani kanker melanoma.
4. Meningkatkan pemahaman tentang faktor-faktor yang mempengaruhi perkembangan dan penyebaran kanker tersebut.

BAB II

TINJAUAN PUSTAKA

2.1. Tinjauan Pustaka

Pengembangan convolutional neural network (CNN), secara luas dianggap sebagai titik balik dalam penelitian deep learning (Bisla et al., 2019). Pendekatan berbasis CNN kemudian mengungguli manusia dalam beberapa tugas klasifikasi gambar. Akibatnya, teknik analisis otomatis lesi kulit baru-baru ini didominasi oleh algoritma deep learning. Namun diagnosis komputasi kanker kulit sulit karena kelangkaan relatif data lesi yang berlabel dan kualitas data pelatihan yang buruk. Masalah seperti ini juga ditemui pada penelitian sebelumnya seperti dataset yang terlalu sedikit (Naeem et al., 2020; Saputro et al., 2022), terlalu bervariasi (Teresia R. Savera, Winsya H. Suryawan, 2020) dan bagaimana perbedaan jumlah dataset dapat mempengaruhi hasil dari model yang dibuat (Adegun & Viriri, 2020). Walaupun akurasi yang didapat sudah lumayan bagus namun masalah ini masih menjadi tantangan dalam mendeteksi kanker kulit.

Beberapa peneliti tersebut juga memaparkan beberapa masukan terhadap penelitian selanjutnya untuk mengatasi masalah yang mereka temukan. Naeem et al (Naeem et al., 2020) menyarankan menggunakan *fine-tuning* ke *hyper-parameter* yang dapat mengurangi kemungkinan hasil yang *overfitting* sedangkan Khasanah et al (Khasanah et al., 2021a) menyarankan menggunakan arsitektur CNN yang lain, mereka menyimpulkan kalau bisa saja yang bermasalah bukan jumlah datanya melainkan arsitektur CNN yang belum tepat. Adapun saran yang diutarakan oleh

(Perez & Wang, 2017) yaitu menggunakan proses augmentasi, yang mana dengan menggunakan proses augmentasi dataset yang ada bisa bertambah jumlahnya bergantung pada jumlah augmentasi yang digunakan. Tentu saja ini didukung oleh penelitian yang dilakukan oleh (Shorten & Khoshgoftaar, 2019) Penelitian ini meninjau dan menganalisis secara menyeluruh berbagai pendekatan untuk meningkatkan data untuk klasifikasi gambar. Mereka juga mengevaluasi berbagai pendekatan untuk meningkatkan data dan menganalisis hubungannya dengan tingkat akurasi klasifikasi.

Pada penelitian yang dilakukan oleh (Teresia R. Savera, Winsya H. Suryawan, 2020) dimana mereka menggabungkan proses segmentasi dan augmentasi. Dengan menggunakan model CNN hasil dari penelitian tersebut tidak begitu bagus dengan akurasi sebesar 76%. Pada penelitian (Khasanah et al., 2021a) lebih menekankan pada preprocessing menggunakan ekstraksi fitur dan algoritma yang digunakan adalah *Random Forest*. Ekstraksi fitur yang digunakan adalah *histogram, haralick, hue moment*. Dari ketiga ekstraksi fitur tersebut hue moment menghasilkan akurasi yang lebih baik yaitu 0,842, sedangkan histogram mendapat akurasi sebesar 0,834 dan haralick sebesar 0,619.

Adapun yang mengusulkan menggabungkan proses ekstraksi fitur dan augmentasi seperti yang dilakukan oleh (Bansal et al., 2022). Mereka menggunakan tahap preprosing hair removal dengan 3 cara yaitu Black-hat based method (BHM), LoG and Top-hat based method (LoG_THM) dan Black-hat and top-hat based method (BH_THM). Selanjutnya mereka menggunakan proses augmentasi Horizontal Flip, Vertical Flip, Rotation, Zoom, Random Noise, Increase Brightness,

Increase Contrast untuk memperkaya dataset. Algoritma yang digunakan adalah ResNet dan EfficientNet dengan dataset yang mereka gunakan adalah HAM 10000. Hasil yang didapat pada algoritma ResNet Log_THM mendapat akurasi terbaik yaitu 0,927 dan pada algoritma EfficientNet LoG_THM juga mendapat akurasi terbaik yaitu 0.940.



2.2. Keaslian Penelitian

Tabel 2.1. Matriks *literature review* Pengaruh Proses Augmentasi Terhadap Model *Transfer Learning* Resnet50 Pada Klasifikasi Kanker Melanoma

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
1	Deteksi Dini Kanker Kulit Menggunakan K-NN Dan Convolutional Neural Network	Teresia R. Savera, Winsya H. Suryawan, Agung W. Setiawan, Jurnal Teknologi Informasi dan Ilmu Komputer (JTHK), (2020)	Sebuah program perangkat lunak yang memungkinkan diagnosis dini kanker kulit diperlukan karena merupakan jenis kanker yang berpotensi fatal. Deteksi dini kanker	Kedua pendekatan menghasilkan hasil yang hampir identik. Tingkat klasifikasi k-NN adalah 75%, sedangkan tingkat klasifikasi CNN adalah 76,56. Dataset yang digunakan berdampak pada akurasi yang ditemukan. Dataset yang digunakan terlalu bervariasi dan	Kumpulan data yang digunakan berdampak pada akurasi yang ditemukan, namun karena kumpulan data terlalu beragam dan volumenya terlalu kecil, sulit untuk membuat generalisasi.	Menggunakan dataset yang lebih baik dengan memfokuskan penelitian hanya pada jenis kanker melanoma benign dan malignant, ditambah data yang didapat akan diperbanyak.

Tabel 2.1. Matriks literature review Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 Pada Klasifikasi Kanker Melanoma (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			<p>kulit diyakini akan lebih cepat ditemukan.</p>	<p>jarang untuk digeneralisasikan. Agar teknik k-NN klasifikasi kanker kulit memperhitungkan empat parameter, kemajuan studi selanjutnya memerlukan penggunaan lebih banyak data dan kumpulan data standar untuk klasifikasi. Metode yang ditemukan juga dapat digunakan sebagai alat deteksi</p>		

Tabel 2.1. Matriks literature review Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 Pada Klasifikasi Kanker Melanoma (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			kulit diyakini akan lebih cepat ditemukan.	dini kanker kulit karena komputasinya relatif ringan dan dapat dimasukkan ke dalam sistem deteksi berbasis smartphone.		
2	Deep Learning-Based System for Automatic Melanoma Detection	Adegun, Adekanmi A. Viriri, Serestina, IEEEAcces, (2020)	Pendekatan ini berupaya menyederhanakan arsitektur pembelajaran mendalam untuk	Untuk deteksi dan segmentasi lesi melanoma yang andal, arsitektur jaringan konvolusional yang dalam telah dikembangkan dalam penelitian ini. Berdasarkan	Model klasifikasi langsung bekerja lebih baik dengan kumpulan data yang lebih kecil.	Menggunakan model yang berbeda dengan yang digunakan dan akan dibandingkan lagi dengan akurasi yang sudah didapat

Tabel 2.1. Matriks literature review Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 Pada Klasifikasi Kanker Melanoma (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			deteksi melanoma. Ini juga berupaya menciptakan sistem produktif yang dapat melakukan tugas diagnosis medis real-time dari diagnosis kanker melanoma dari diagnosis kanker melanoma.	hasil pengklasifikasian softmax, sebuah teknik baru dikembangkan untuk mengkategorikan lesi melanoma dan non-melanoma. Pendekatan ini berupaya menyederhanakan arsitektur pembelajaran mendalam untuk deteksi melanoma. Ini juga bermaksud untuk membuat sistem produktif yang dapat		pada penelitian tersebut.

Tabel 2.1. Matriks literature review Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 Pada Klasifikasi Kanker Melanoma (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
				melakukan tugas diagnosis medis real- time dari diagnosis kanker melanoma.		
3	Klasifikasi Penyakit Kanker Kulit Menggunakan Metode Convolutional Neural Network	Saputro, Reynaldi Rio Junaidi, Apri Saputra, Wahyu Andi, Journal of Dinda Kelompok Keahlian Rekayasa Data Institut	Penelitian ini bertujuan untuk membantu masyarakat umum dalam membedakan antara tanda lahir yang khas dan keganasan melanoma.	Akurasi tes ini terbilang baik, menurut hasil model CNN dari studi yang dilakukan. Epoch berdampak pada keluaran model. Hasil ditingkatkan dengan menggunakan lebih banyak zaman.	Penelitian lebih lanjut menggunakan arsitektur CNN yang berbeda masih dapat menghasilkan hasil akurasi yang lebih baik.	menggunakan arsitektur transfer learning yaitu ResNet50 akan dibandingkan dengan dataset yang sama dan dibandingkan lagi

Tabel 2.1. Matriks literature review Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 Pada Klasifikasi Kanker Melanoma (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	(Studi Kasus: Melanoma)	Teknologi Telkom Purwokerto, (2022)				dengan dataset yang telah di augmentasi.
4	Skin Cancer Classification Using Random Forest Algorithm	Khasanah, Nurul Komartudin, Rachman Afni, Nurul Maulana, Yana Iqbal Salim, Agus, JURNAL SISFOTENIKA, (2021)	Agar dapat merawat pasien dengan baik, temuan penelitian ini diharapkan dapat menunjukkan nilai akurasi yang lebih tinggi untuk identifikasi kelainan kulit pada manusia.	Ekstraksi fitur momen rona yang memiliki nilai akurasi sebesar 0,842 menghasilkan hasil akurasi terbaik dari ketiga ekstraksi fitur yang diuji. Studi ini inovatif karena menggunakan metode ekstraksi fitur yang lebih beragam dan mencapai	Arsitektur jaringan saraf convolutional (CNN) dan algoritma pembelajaran mendalam kemungkinan akan digunakan dalam penelitian di masa depan untuk mencapai akurasi lebih tinggi.	menggunakan arsitektur transfer learning yaitu ResNet50 akan dibandingkan dengan dataset yang sama dan dibandingkan lagi dengan dataset yang telah di augmentasi.

Tabel 2.1. Matriks literature review Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 Pada Klasifikasi Kanker Melanoma (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
				<p>tingkat akurasi yang lebih tinggi daripada penelitian sebelumnya. Keterbatasan penelitian ini adalah algoritma tersebut tidak diimplementasikan dalam sebuah aplikasi, sehingga temuannya tidak dapat langsung diterapkan pada bidang kedokteran.</p>		

Tabel 2.1. Matriks literature review Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 Pada Klasifikasi Kanker Melanoma (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
5	Malignant Melanoma Classification Using Deep Learning: Datasets, Performance Measurements, Challenges and Opportunities	Naeem, Ahmad Farooq, Muhammad Shoaib Khelifi, Adel Abid, Adnan (2020), IEEE Access	Menyatukan penelitian mutakhir yang mengidentifikasi tren penelitian terbaru, masalah, dan prospek diagnosis melanoma dan mencari solusi yang sudah ada untuk diagnosis deteksi melanoma menggunakan	Telah dicatat bahwa saat menggunakan algoritme pembelajaran mendalam, metode persiapan komposit dan canggih seperti perubahan ukuran gambar, pemangkasan, dan normalisasi nilai piksel tidak terlalu diperlukan. Dengan meninjau makalah terkait, taksonomi yang diusulkan dan model yang diusulkan telah	Untuk menghindari overfitting, peneliti harus menggunakan kumpulan data yang lebih besar dan menyesuaikan dengan hyper-parameter.	Menggunakan dataset yang lebih terfokus pada kanker melanoma benign dan malignant ditambah dengan meningkatkan dataset dengan teknik augmentasi.

Tabel 2.1. Matriks literature review Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 Pada Klasifikasi Kanker Melanoma (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			Pembelajaran mendalam.	diberikan. Analisis ini juga mengklasifikasikan kelemahan utama dalam pendekatan saat ini dan menyoroti area yang masih membutuhkan pengembangan. Hasil dari pendekatan buatan mengungguli hasil dari teknik pembelajaran mendalam tradisional.		

Tabel 2.1. Matriks literature review Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 Pada Klasifikasi Kanker Melanoma (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
6	Melanoma Detection Using Convolutional Neural Network	Zhang, Runyuan (2021)	Penulis melakukan studi menyeluruh tentang melanoma dan menemukan bahwa jaringan saraf convolutional dengan kedalaman, cakupan, dan resolusi yang lebih besar bekerja lebih baik.	Secara sistematis meneliti sejarah dan situasi deteksi melanoma saat ini. Berdasarkan temuan ini, penulis melihat B6 berkapasitas Efisien untuk menangkap detail yang lebih halus dan gambar dermoskopi klinis yang lebih kompleks dari lesi kulit.	Itu dapat mencrapkan jaringan yang diusulkan ke jenis kanker kulit yang lebih luas dengan mengeksplorasi lebih jauh hubungan unik antara kanker kulit dan melanoma.	pada penelitian ini akan berfokus pada kanker kulit jenis melanoma benign dan malignant ditambah dengan memperbesar jumlah dataset dengan proses augmentasi.

Tabel 2.1. Matriks literature review Pengaruh Proses Augmentasi Terhadap Model Transfer Learning Resnet50 Pada Klasifikasi Kanker Melanoma (Lanjutan)

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
				<p>Hasil eksperimen menunjukkan bahwa jaringan yang diusulkan cenderung berfokus pada wilayah yang lebih relevan dengan detail melanoma berkat jaringan resolusi yang lebih dalam, lebih luas, dan lebih tinggi. Oleh karena itu, saya mencapai akurasi klasifikasi yang lebih baik daripada metode populer sebelumnya.</p>		

2.3. Landasan Teori

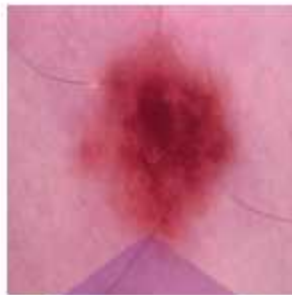
2.3.1. Kanker Melanoa

Melanoma adalah jenis kanker yang muncul dari sel-sel pigmen di kulit. Ini adalah salah satu jenis kanker kulit yang paling ganas dan dapat menyebar ke organ lain dalam tubuh jika tidak diobati dengan cepat. Melanoma biasanya muncul sebagai benjolan hitam atau coklat pada kulit, tetapi juga dapat muncul sebagai benjolan berwarna pink, ungu, atau putih.

Faktor risiko utama untuk melanoma termasuk riwayat kerusakan DNA kulit dari paparan sinar matahari atau radiasi UV dari sumber lain, genetik, dan riwayat keluarga kanker kulit. Deteksi dini dan pemeriksaan kulit reguler oleh dokter dapat membantu dalam mendeteksi dan mengobati melanoma sebelum itu menyebar. Ada beberapa jenis melanoma yang dikenal seperti melanoma klasik, nodular, lentigo maligna, *acral lentiginous* dan *amelanotic melanoma*. Semua jenis ini memerlukan perhatian medis segera dan beberapa dapat menyebar dengan cepat jika tidak diobati tepat waktu. Ada beberapa jenis melanoma yang dikenal :

a. *Melanoma Benign* (Jinak)

Pertumbuhan sel-sel melanosit yang tidak invasif dan cenderung tidak berbahaya dikenal sebagai melanoma benign, juga dikenal sebagai *nevus melanocytic*. Lesi melanositik benigna biasanya berwarna seragam dan tidak berubah secara signifikan dari waktu ke waktu. Mereka biasanya berbentuk bulat atau oval dengan tepi yang teratur, simetris, dan ukuran yang relatif kecil. *Melanoma benign* tidak berisiko metastasis atau menyebar ke bagian tubuh lainnya.



Gambar 2.1. *Melanoma Benign*

b. *Melanoma Malignant (Ganas)*

Melanoma malignant adalah jenis melanoma yang invasif yang dapat metastasis atau menyebar ke jaringan dan organ tubuh lainnya. Biasanya tumbuh lebih cepat dan dapat menyebar ke lapisan kulit yang lebih dalam dan bahkan ke sistem peredaran darah atau limfatik. Bentuk yang tidak teratur, warna yang tidak merata, tepi yang tidak teratur, dan ukuran yang lebih besar adalah ciri khas penyakit melanoma malignant. Perubahan pada ukuran, bentuk, warna, atau simetri lesi melanoma dapat menjadi tanda bahwa Anda harus segera mendapatkan bantuan medis.



Gambar 2.2. *Melanoma Malignant*

Penyebab utama melanoma adalah kerusakan DNA pada sel-sel kulit yang disebabkan oleh paparan sinar ultraviolet (UV) dari matahari atau sumber lain, seperti lampu tanning. Risiko melanoma juga meningkat dengan banyaknya tanda-tanda kerusakan kulit, seperti bekas luka atau keriput, serta faktor genetik dan riwayat keluarga yang positif untuk kanker kulit. Untuk mengobati melanoma, dokter akan mengevaluasi stadium kanker dan menentukan strategi pengobatan yang sesuai. Beberapa opsi pengobatan yang mungkin digunakan meliputi:

- a. Operasi: Melanoma dikeluarkan dari kulit dengan operasi. Jika melanoma terdeteksi pada tahap awal, operasi mungkin cukup untuk menyelamatkan pasien.
- b. Terapi sistemik: Terapi sistemik, seperti kemoterapi atau terapi imunologi, digunakan jika melanoma telah menyebar ke organ lain dalam tubuh.
- c. Radioterapi: ini digunakan ketika operasi atau terapi sistemik tidak mungkin atau tidak efektif.
- d. Terapi Targeted: terapi seperti inhibitor kinase menargetkan protein yang spesifik dalam sel kanker untuk menghentikan perkembangannya

Kombinasi dari pengobatan di atas mungkin digunakan bersama-sama untuk meningkatkan efektivitas pengobatan. Kemoterapi, radioterapi, terapi sistemik dan terapi targeted juga digunakan dalam kondisi yang memerlukan pengobatan bertahap atau paliatif. Prognosis tergantung pada tahap kanker saat diagnosa dan respons pasien terhadap terapi yang diterima.

2.3.2. Klasifikasi

Secara Klasifikasi adalah proses mengelompokkan data ke dalam kategori-kategori atau kelas-kelas yang telah ditentukan sebelumnya(Kusumawardani & Karningsih, 2021). Dalam pembelajaran mesin, klasifikasi sering digunakan untuk membuat model yang dapat mengklasifikasikan data baru berdasarkan data yang telah diterima sebelumnya(Erdiansyah et al., 2022).

Ada beberapa jenis klasifikasi yang dikenal, di antaranya :

- a. Klasifikasi biner: hanya ada 2 kelas, contoh : spam atau bukan spam, tumor jinak atau ganas.
- b. Klasifikasi multi-kelas: lebih dari 2 kelas, contoh : memprediksi jenis bunga dari suatu foto.
- c. Klasifikasi multi-label: sebuah data dapat di label dengan lebih dari satu kelas, contoh : tag pada foto instagram.

Klasifikasi digunakan dalam berbagai bidang, seperti pengenalan wajah, analisis teks, pemrosesan gambar, dan analisis data(Yunial, 2020). Dalam pembelajaran mesin, klasifikasi dapat dilakukan dengan menggunakan algoritma seperti naive bayes, decision tree, random forest, atau support vector machine (SVM)(Nurfaizah et al., 2019). Pemilihan algoritma tergantung pada karakteristik dari data yang akan diklasifikasikan.

2.3.3. Augmentasi

Augmentasi gambar adalah proses yang digunakan untuk meningkatkan kualitas atau kuantitas dari data gambar yang digunakan dalam pembelajaran mesin (Walawalkar et al., 2020). Proses ini digunakan untuk membuat model yang lebih robust dan mampu mengatasi masalah *overfitting* (Wu et al., 2020).

Dalam proses augmentasi, gambar asli diubah dengan berbagai transformasi seperti rotasi, flip, zoom, crop, translasi, perubahan *brightness* dan *contrast*, perubahan warna, dan penambahan *noise* (Rana et al., 2022). Transformasi ini digunakan untuk memperluas variasi dari gambar yang digunakan dalam training model. Hal ini membuat model lebih baik dalam menangani data yang tidak terlihat sebelumnya saat training dan meningkatkan performa model saat digunakan dalam kondisi sebenarnya (Ghosh et al., 2020).

Contohnya, jika kita memiliki data gambar yang hanya mencakup objek dalam sudut pandang yang sama, model yang dibangun dari data tersebut mungkin kurang baik dalam mengenali objek dari sudut pandang yang berbeda (Bird et al., 2022). Dengan menggunakan augmentasi gambar, kita dapat memperluas variasi dari sudut pandang gambar sehingga model dapat belajar untuk mengenali objek dari berbagai sudut pandang (Bisla et al., 2019).

Selain itu, dengan menambahkan *noise*, mengubah *brightness* dan *contrast*, atau mengubah warna dari gambar, kita dapat membuat model lebih robust terhadap perbedaan kondisi pencahayaan atau warna yang mungkin terjadi pada data asli (Rana et al., 2022). Hal ini sangat penting dalam aplikasi praktis seperti pengenalan objek dalam lingkungan nyata yang kompleks (Mandal et al., 2020).

Dalam ilustrasi pengolahan gambar, augmentasi digunakan untuk menambah jumlah data yang digunakan dalam training model dan membuat model yang lebih robust, yang membuat model mampu mengatasi overfitting dan memperkuat performa dalam kondisi sebenarnya (Walawalkar et al., 2020).

2.3.4. Oversampling

Teknik oversampling digunakan dalam pembelajaran mesin dan analisis data untuk menangani masalah ketidakseimbangan kelas. Ketika jumlah contoh dalam satu kelas jauh lebih besar daripada contoh dalam kelas lainnya, itu disebut ketidakseimbangan kelas. Hal ini dapat menyebabkan model yang dilatih menjadi bias terhadap kelas yang lebih dominan. Teknik oversampling meningkatkan jumlah contoh yang diambil dari kelas minoritas untuk membuat distribusi kelas lebih seimbang.

Synthetic Minority Over-sampling Technique (SMOTE) adalah teknik oversampling yang populer yang memanfaatkan interpolasi untuk menghasilkan contoh sintesis baru dari kelas minoritas. Dua contoh acak dari kelas minoritas dipilih, dan titik baru dibuat di antara kedua contoh tersebut. Dengan demikian, SMOTE tidak hanya menambah jumlah data tetapi juga memperkenalkan variasi baru ke dalam data yang dapat membantu model belajar pola yang lebih baik.

Secara matematis, rumus sederhana dapat digunakan untuk menjelaskan SMOTE. Misalkan dua contoh dari kelas minoritas adalah x_i dan x_j , dan λ adalah angka acak antara 0 dan 1. Kemudian, contoh sintesis x_{new} dapat dihitung sebagai:

$$x_{new} = x_i + \lambda \cdot (x_j - x_i) \dots\dots\dots(2.1)$$

Proses ini diulang hingga jumlah contoh dalam kelas minoritas mencapai jumlah yang diinginkan. Dengan demikian, SMOTE membantu mengurangi ketidakseimbangan kelas dan memungkinkan model pembelajaran mesin untuk mengenali pola yang lebih baik dari kelas minoritas.

Oversampling acak, yang digunakan untuk menyalin contoh-contoh dari kelompok minoritas secara acak hingga distribusi kelas menjadi seimbang, adalah salah satu metode oversampling lainnya selain SMOTE. Meskipun teknik ini sederhana, ada kelemahannya, yaitu kemungkinan overfitting karena model sering menghafal contoh yang diulang. Oleh karena itu, metode seperti SMOTE yang menghasilkan contoh sintesis digunakan lebih sering untuk mengatasi ketidakseimbangan kelas.

2.3.5. Gaussian Blur

Dalam pemrosesan citra, Gaussian Blur adalah teknik penyaringan yang digunakan untuk menghaluskan dan mengurangi detail atau suara dalam gambar (Yusuf et al., 2024). Metode ini berasal dari fungsi Gaussian, yang menghasilkan efek kabur yang halus dan alami. Fungsi Gaussian adalah fungsi distribusi probabilitas dalam bentuk lonceng, dan ketika diterapkan pada gambar, nilai piksel berubah sesuai dengan kontribusi piksel tetangga dengan bobot yang ditentukan oleh fungsi Gaussian (Grasela Tatuin et al., 2023).

Kernel Gaussian diterapkan ke setiap piksel gambar dalam proses Blur Gaussian. Kernel ini adalah matriks berukuran ganjil, seperti 3×3 , 5×5 , atau 7×7 , yang elemennya dihitung dengan menggunakan fungsi Gaussian. Dalam gambar

baru, setiap piksel berasal dari rata-rata tertimbang piksel-piksel tetangganya dalam gambar awal, dengan kernel Gaussian memberikan bobotnya (Ekoputra, 2019). Secara matematis, proses ini dapat dinyatakan dengan rumus:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \dots\dots\dots(2.2)$$

Dimana $G(x, y)$ adalah nilai fungsi Gaussian di titik (x, y) , dan σ adalah standar deviasi yang menentukan seberapa lebar distribusi Gaussian.

Untuk berbagai aplikasi, seperti deteksi tepi dan segmentasi gambar sebelum pemrosesan, Gaussian Blur sangat bermanfaat. Ini mengurangi noise dengan sangat baik tanpa kehilangan banyak detail penting dalam gambar. Dalam fungsi Gaussian, standar deviasi σ mengatur jumlah blur yang diterapkan. Nilai σ yang lebih tinggi menghasilkan blur yang lebih halus dan tersebar, sementara nilai σ yang lebih rendah mempertahankan lebih banyak detail asli. Selain itu, karena fungsi Gaussian isotropik, blur yang dihasilkan seragam ke segala arah.

2.3.6. Resnet-50

ResNet50 adalah arsitektur jaringan saraf yang dikembangkan oleh Microsoft Research untuk memecahkan masalah pengenalan objek dalam kumpulan data ImageNet (Kusumawardani & Karningsih, 2021). Arsitektur ini dikenal karena kemampuannya menangani masalah pengenalan objek dengan baik, bahkan pada gambar berkualitas rendah atau berisik (Pangestu & Bunyamin, 2018).

ResNet50 memiliki 50 lapisan jaringan saraf termasuk lapisan konvolusi, lapisan normalisasi batch, dan lapisan aktivasi ReLU(Kusumawardani & Karningsih, 2021). Lapisan-lapisan ini dapat dipentaskan menggunakan teknik hop join yang memungkinkan informasi yang dibuang oleh proses konvolusi untuk dikirim kembali ke lapisan yang lebih tinggi(Pangestu & Bunyamin, 2018). Ini memungkinkan ResNet50 menangani masalah gradien bocor yang sering terjadi di jaringan saraf yang terlalu dalam.

ResNet50 juga memiliki transfer learning yang baik, sehingga dapat digunakan untuk memecahkan masalah pengenalan objek di kumpulan data lain hanya dengan menyetel beberapa lapisan atasnya(Chu et al., 2020). Arsitektur ini telah terbukti memberikan performa yang sangat baik dalam berbagai kompetisi pengenalan objek, seperti ImageNet *Large Scale Image Recognition Challenge* (ILSVRC) dan Microsoft Common Object in CONtext (MS COCO)(He et al., 2016).

ResNet50 bekerja dengan memproses input gambar melalui serangkaian lapisan jaringan saraf termasuk lapisan konvolusi, lapisan normalisasi batch, dan lapisan aktivasi ReLU. Masing-masing lapisan ini bertanggung jawab untuk mengekstrak fitur dari citra yang berguna untuk pengenalan objek. Berikut lapisan yang digunakan dalam Resnet50 :

a. Lapisan Konvolusi

Lapisan ini bertugas untuk mengekstrak fitur-fitur spasial dari gambar dengan cara menggunakan filter yang disebut kernel.

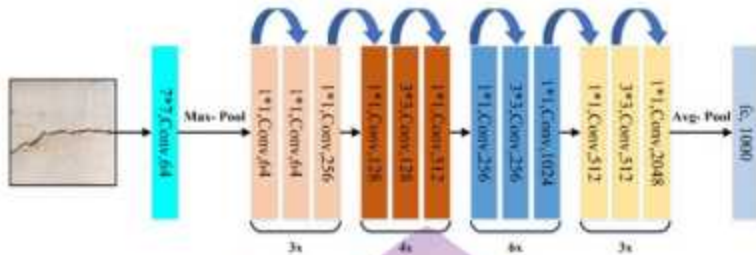
b. Lapisan *Batch Normalization*

Bertanggung jawab untuk menormalkan data yang dihasilkan oleh lapisan konvolusi agar lebih mudah diproses oleh lapisan berikutnya. Lapisan ini juga dapat mempercepat konvergensi jaringan saraf dan mengurangi redundansi.

c. Lapisan ReLU

Bertanggung jawab untuk mengganti semua nilai negatif pada data dengan 0, sehingga hanya nilai positif yang diproses oleh lapisan berikutnya. Lapisan ini berguna untuk mengurangi waktu komputasi dan mempercepat konvergensi jaringan saraf.

Jaringan saraf residual (ResNet), yang memenangkan Tantangan Pengenalan Visual Skala Besar ImageNet, diusulkan oleh (He et al., 2016). (ILSVRC 2015). ResNet memperkenalkan koneksi sisa antar lapisan, mengurangi kerugian, mempertahankan perolehan pengetahuan, dan meningkatkan kinerja selama fase pelatihan. Ketika sebuah layer memiliki residual link, outputnya adalah konvolusi dari input dan inputnya. Gambar 2.3 menampilkan diagram blok arsitektur model ResNet.



Gambar 2.3. Arsitektur Resnet50

Arsitektur ResNet-50 pada gambar 2.3 dimulai dengan mengambil input gambar dan mengoperasikannya melalui lapisan konvolusi berukuran 7×7 dengan 64 filter, diikuti oleh lapisan Max-Pooling untuk mengurangi dimensi. Kemudian, data melewati beberapa tahap blok residu. Tahap pertama terdiri dari 3 blok dengan konvolusi 1×1 dan 3×3 . Tahap kedua memiliki 4 blok dengan pola serupa, tetapi dengan filter yang lebih banyak (128 filter). Tahap ketiga memiliki 6 blok, meningkatkan jumlah filter menjadi 256, dan tahap keempat memiliki 3 blok dengan 512 filter. Setiap blok residu mengandung shortcut connections yang menghubungkan input langsung ke output untuk menjaga aliran informasi dan mengatasi masalah vanishing gradient. Setelah semua tahap blok residu, data diproses melalui lapisan Average Pooling dan lapisan fully connected yang menghasilkan output akhir berukuran 1000, siap untuk klasifikasi.

2.3.7. Confusion Matrix

Evaluation matrix atau confusion matrix adalah sebuah tabel yang digunakan untuk mengevaluasi performa dari suatu model klasifikasi (Algayres et al., 2020).

Tabel tersebut menampilkan hasil dari perbandingan antara data yang dihasilkan oleh model dengan data yang sebenarnya (ground truth)(Kim & Cho, 2022). Umumnya, matriks tersebut digunakan untuk mengukur performa dari model dengan metrik seperti akurasi, presisi, dan recall(Khoiruzzaman et al., 2021).

Setiap baris dari matriks tersebut mewakili hasil yang dihasilkan oleh model, sedangkan setiap kolom mewakili hasil yang sebenarnya (ground truth)(Kim & Cho, 2022). Di pojok kiri atas matriks, kita akan menemukan jumlah dari data yang terprediksi sebagai kelas positif dan sebenarnya merupakan kelas positif. Di pojok kanan atas matriks, kita akan menemukan jumlah dari data yang terprediksi sebagai kelas positif tapi sebenarnya merupakan kelas negatif. Demikian juga pada pojok kiri bawah dan kanan bawah. Bisa dilihat pada gambar 2.4.



		Actual Values	
		Positive [1]	Negative [0]
Predicted Values	Positive [1]	TP	FP
	Negative [0]	FN	TN

Gambar 2.4. Confusion Matrix

Dalam praktiknya menggunakan evaluation matrix untuk menentukan performa dari model klasifikasi yang digunakan, dan juga untuk mengevaluasi apakah suatu model dapat digunakan untuk aplikasi yang ditentukan(Ransi et al., 2020). Untuk menghitung akurasi, presisi, recall dan f1-score bisa dilihat pada persamaan 2.3 – 2.6.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \dots\dots\dots(2.3)$$

$$\text{Precision} = \frac{TP}{FP+TP} \dots\dots\dots(2.4)$$

$$\text{Recall} = \frac{TP}{TP+FN} \dots\dots\dots(2.5)$$

$$F1 = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \dots\dots\dots(2.6)$$

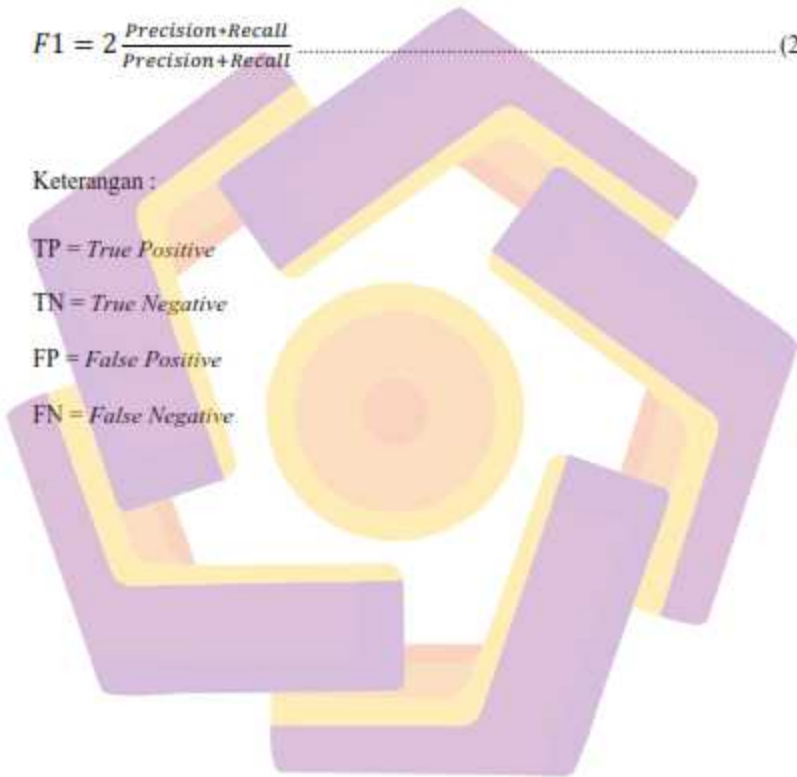
Keterangan :

TP = *True Positive*

TN = *True Negative*

FP = *False Positive*

FN = *False Negative*



BAB III

METODE PENELITIAN

3.1. Jenis, Sifat, dan Pendekatan Penelitian

Penelitian eksperimental yang menggunakan pendekatan kuantitatif digunakan untuk menguji pengaruh proses augmentasi terhadap model Resnet50 dengan objek kanker melanoma. Hasilnya kemudian dipresentasikan dalam bentuk penelitian deskriptif yang mendetail untuk memberikan gambaran komprehensif tentang pengaruh proses augmentasi.

3.2. Metode Pengumpulan Data

Dataset yang digunakan berasal dari Kaggle dengan judul Melanoma Detection Dataset dari Pablo Lopez Santori (ISIC 2017) , pada setiap jenis kanker melanoma mempunyai jumlah data gambar yang hampir sama banyak yaitu sekitar 5500 gambar benign dan 5105 gambar malignant.

3.3. Metode Analisis Data

Analisis data dalam penelitian ini dapat dilakukan dengan beberapa metode seperti Analisis performansi untuk mengevaluasi seberapa baik model transfer learning ResNet50 yang digunakan dalam klasifikasi kanker melanoma. Metrik yang digunakan untuk mengevaluasi performansi dapat berupa akurasi, presisi, recall, F1-score. Dan Analisis visualisasi untuk mengevaluasi perbedaan

performansi model transfer learning ResNet50 setelah dilakukan proses augmentasi. Analisis visualisasi ini dapat berupa confusion matrix.

3.4. Alur Penelitian

Pada penelitian ini terdiri dari 4 tahap, yaitu:

1. Data Collection

Data diambil dari kaggle dengan nama Melanoma Detection Dataset dari Pablo Lopez Santori (ISIC 2017).

2. Data Cleaning

Data yang diperoleh akan ditinjau ulang dan yang akan digunakan hanya gambar kanker kulit melanoma.

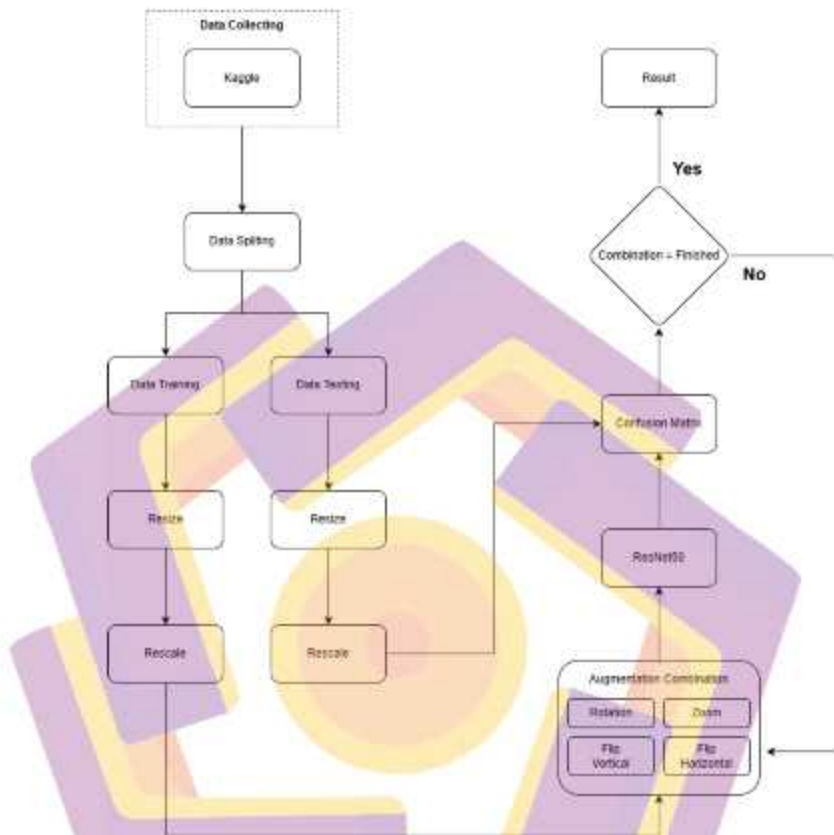
3. Preprocessing

Data dibagi menjadi 2 bagian yaitu training dan testing selanjutnya data training akan masuk ke dalam tahap preprocessing menggunakan augmentasi dengan kombinasi Rotation, Zoom, Flip.

4. Model Training dan Testing

Data yang telah di preprocessing selanjutnya akan dilakukan proses model training menggunakan algoritma ResNet50. Setelah itu model training akan dilakukan uji coba dan akan menghasilkan nilai akurasi, presisi recall dan F1 Score.

Setelah semua kombinasi augmentasi habis akan direkap dan dilaporkan dalam bentuk tabel dan bisa dilakukan penarikan kesimpulan penelitian. Semua alur penelitian bisa dilihat pada Gambar 3.1.

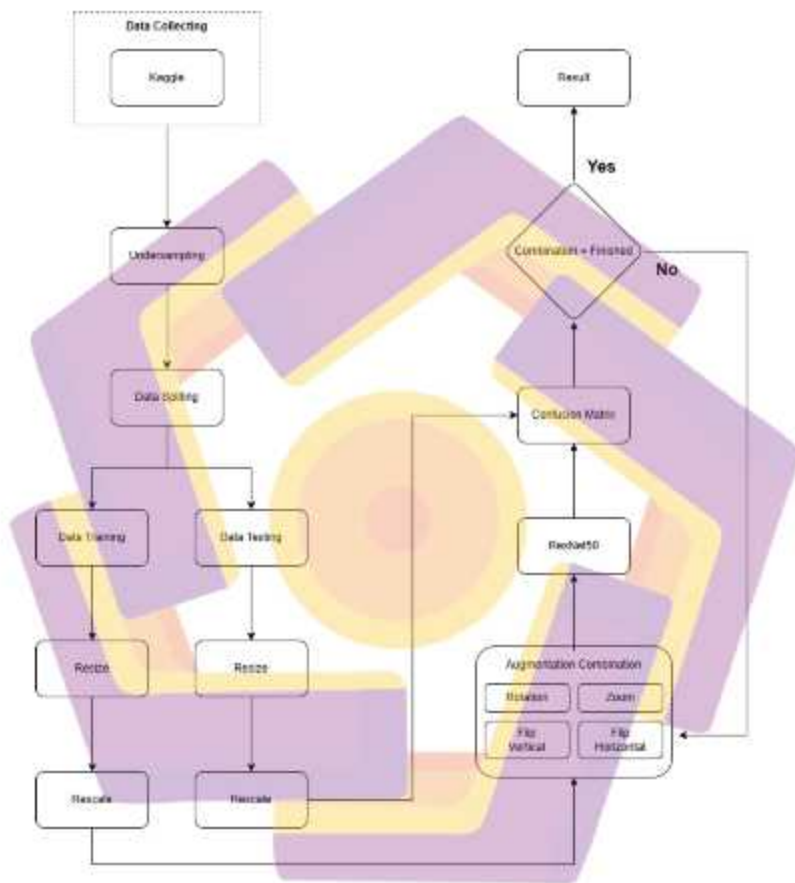


Gambar 3.1. Alur Penelitian

5. Skenario Lain

Jika hasil akurasi tidak lebih baik dari penelitian sebelumnya, ada beberapa langkah yang dapat diambil untuk meningkatkan performa model. Pertama, penambahan proses oversampling yang dilakukan sebelum proses pembagian data (Data Splitting) untuk menangani ketidakseimbangan kelas. Langkah ini bertujuan untuk memastikan bahwa setiap kelas memiliki representasi yang lebih seimbang

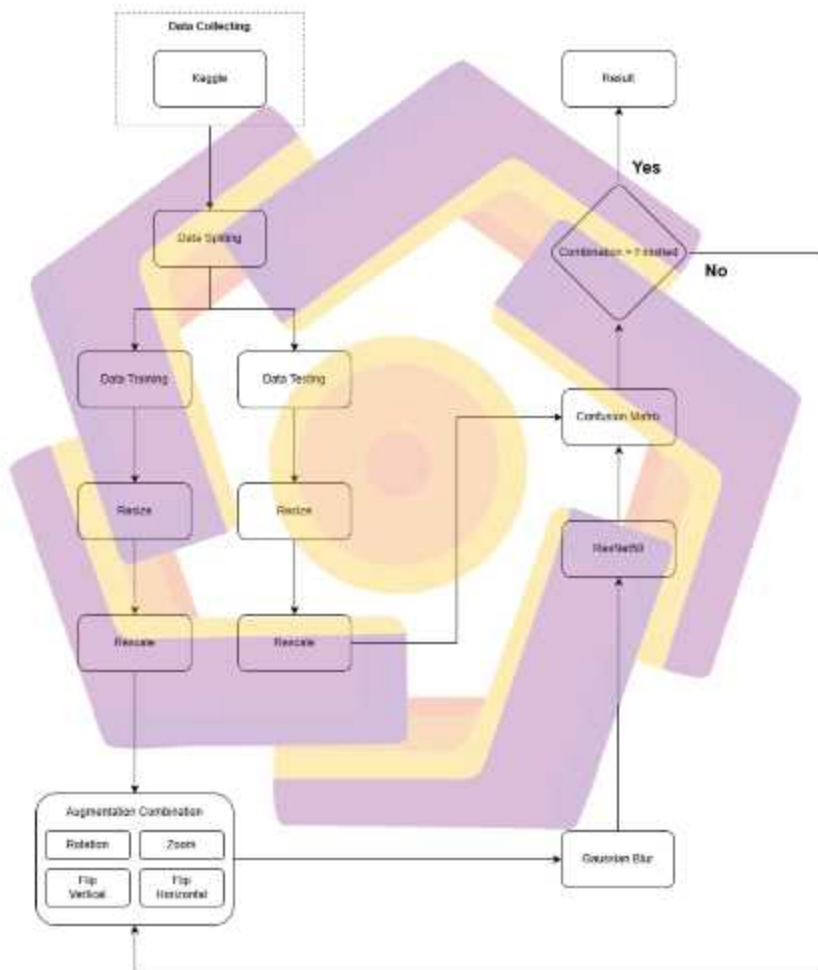
dalam dataset, sehingga model dapat belajar dengan lebih baik. Alur lengkap dari langkah ini dapat dilihat pada gambar 3.2.



Gambar 3.2. Alur Penelitian dengan Oversampling

Kedua, penambahan proses Gaussian Blur tanpa menggunakan oversampling. Proses ini diterapkan setelah augmentasi data dan bertujuan untuk membuat gambar dalam data training menjadi blur. Penambahan Gaussian Blur

diharapkan dapat membantu model menjadi lebih tangguh terhadap variasi dan noise dalam gambar, sehingga meningkatkan kemampuan generalisasi model. Alur dari proses ini bisa dilihat pada gambar 3.3.



Gambar 3.3. Alur Penelitian dengan Gaussian Blur

BAB IV



HASIL PENELITIAN DAN PEMBAHASAN

4.1. Membangun Dataset

4.4.1. Dataset

Dataset Kaggle yang terdiri dari dua kelas, Benigna dan Malignant, terdiri dari 10.605 gambar, dengan 5.105 gambar dari kelas Malignant dan 5.500 gambar dari kelas Benigna. Data ini digunakan dalam penelitian tentang kanker kulit. Dataset [Hasnainjaved/melanoma-skin-cancer-dataset-of-10000-images](https://www.kaggle.com/datasets/hasnainjaved/melanoma-skin-cancer-dataset-of-10000-images) dapat diakses melalui tautan berikut: <https://www.kaggle.com/datasets/hasnainjaved/melanoma-skin-cancer-dataset-of-10000-images>. Penelitian ini memanfaatkan keragaman gambar untuk analisis kanker kulit dan dapat memberikan wawasan yang berharga tentang deteksi dini kanker kulit serta pemahaman lebih lanjut tentang jenis kanker kulit yang berbeda. Karena jumlah data yang besar, diharapkan penelitian ini dapat meningkatkan akurasi dan keandalan hasil analisis kanker kulit. Platform penyedia dataset Kaggle menawarkan sumber daya penting untuk penelitian medis dan pengembangan model deteksi kanker kulit yang lebih baik. Untuk detail datasetnya bisa dilihat pada tabel 4.1.

Tabel Tabel 4.1. Detail Dataset Kanker Melanoma

Jenis	Contoh citra Kanker Melanoma	Jumlah
Benign		5.500
Malignant		5.105

4.4.2. Persebaran Data

Penelitian ini menggunakan data yang tidak seimbang dalam persebaran, dengan sekitar 500 gambar yang berbeda, seperti yang ditunjukkan dalam Gambar 4.1 yang menunjukkan persebaran data kanker melanoma. Grafik ini sangat penting untuk menganalisis dan memahami hasil penelitian.



Gambar 4.1. Persebaran Jumlah Data Kanker Melanoma

4.2. Analisis Data

Sebelum Dalam rangka menganalisis proses implementasi, perlu disusun skenario untuk memperoleh perbandingan hasil evaluasi model. Skenario ini akan dibagi menjadi tiga tahap utama, masing-masing terdiri dari 15 skenario. Pada tahap pertama, kita akan menjalankan model dengan augmentasi tanpa melakukan oversampling data. Pada tahap kedua, augmentasi akan dilakukan bersamaan dengan proses oversampling, dan pada tahap ketiga akan ditambahkan fitur Gaussian Blur tanpa oversampling. Pemilihan oversampling dan gaussian blur tidak dipilih secara acak melainkan dipilih berdasarkan beberapa pendalaman penelitian terdahulu. Beberapa penelitian telah mengamati manfaat dari oversampling dalam meningkatkan akurasi model, dengan (Mohammed et al., 2020) menekankan keunggulannya dibanding undersampling, sementara (Tanti, 2023) menyoroti efektivitas Random

Oversampling, terutama dalam meningkatkan recall. (Diantika et al., 2024) juga menyarankan penggunaan Random Oversampling untuk meningkatkan kinerja model. Demikian pula, dalam konteks penggunaan Gaussian Blur, penelitian oleh (AGUSTINA et al., 2022; Ekoputra, 2019) menunjukkan bahwa teknik ini membantu memperhalus citra, meningkatkan akurasi tepi citra, serta memperbaiki kualitas citra secara keseluruhan. Penting untuk memahami bahwa evaluasi model tidak hanya berfokus pada nilai akurasi, tetapi juga melibatkan aspek presisi, recall, dan f1-score. Untuk melihat hasil secara holistik, kita akan menggunakan metode seperti confusion matrix dan classification report.

Seluruh skenario yang akan digunakan dalam penelitian ini dirinci dengan cermat, memastikan bahwa setiap langkah dapat memberikan wawasan yang komprehensif terkait kinerja model. Dengan demikian, tahap analisis sebelum implementasi dapat menjadi landasan yang kuat untuk memahami potensi keberhasilan dan perbaikan yang mungkin diperlukan selama tahap implementasi yang sesungguhnya. Berikut skenario yang digunakan pada penelitian ini :

- a. Augmentasi tanpa oversampling
 1. Resnet50 dengan augmentasi (Zoom)
 2. Resnet50 dengan augmentasi (Rotation)
 3. Resnet50 dengan augmentasi (Flip Horizontal)
 4. Resnet50 dengan augmentasi (Flip Vertical)
 5. Resnet50 dengan augmentasi (zoom + Rotation)

6. Resnet50 dengan augmentasi (zoom + Flip Horizontal)
 7. Resnet50 dengan augmentasi (zoom + Flip Vertical)
 8. Resnet50 dengan augmentasi (Rotation + Flip Horizontal)
 9. Resnet50 dengan augmentasi (Rotation + Flip Vertical)
 10. Resnet50 dengan augmentasi (Flip Horizontal + Flip Vertical)
 11. Resnet50 dengan augmentasi (Zoom + Rotation + Flip Horizontal)
 12. Resnet50 dengan augmentasi (Zoom + Rotation + Flip Vertical)
 13. Resnet50 dengan augmentasi (Zoom + Flip Horizontal + Flip Vertical)
 14. Resnet50 dengan augmentasi (Rotation + Flip Horizontal + Flip Vertical)
 15. Resnet50 dengan augmentasi (Zoom + Rotation + Flip Horizontal + Flip Vertical)
- b. Augmentasi dengan oversampling
1. Resnet50 dengan augmentasi (Zoom)
 2. Resnet50 dengan augmentasi (Rotation)
 3. Resnet50 dengan augmentasi (Flip Horizontal)
 4. Resnet50 dengan augmentasi (Flip Vertical)
 5. Resnet50 dengan augmentasi (zoom + Rotation)
 6. Resnet50 dengan augmentasi (zoom + Flip Horizontal)
 7. Resnet50 dengan augmentasi (zoom + Flip Vertical)
 8. Resnet50 dengan augmentasi (Rotation + Flip Horizontal)
 9. Resnet50 dengan augmentasi (Rotation + Flip Vertical)

10. Resnet50 dengan augmentasi (Flip Horizontal + Flip Vertical)
11. Resnet50 dengan augmentasi (Zoom + Rotation + Flip Horizontal)
12. Resnet50 dengan augmentasi (Zoom + Rotation + Flip Vertical)
13. Resnet50 dengan augmentasi (Zoom + Flip Horizontal + Flip Vertical)
14. Resnet50 dengan augmentasi (Rotation + Flip Horizontal + Flip Vertical)
15. Resnet50 dengan augmentasi (Zoom + Rotation + Flip Horizontal + Flip Vertical)

c. Augmentasi ditambah Gaussian Blur

1. Resnet50 dengan augmentasi (Zoom)
2. Resnet50 dengan augmentasi (Rotation)
3. Resnet50 dengan augmentasi (Flip Horizontal)
4. Resnet50 dengan augmentasi (Flip Vertical)
5. Resnet50 dengan augmentasi (zoom + Rotation)
6. Resnet50 dengan augmentasi (zoom + Flip Horizontal)
7. Resnet50 dengan augmentasi (zoom + Flip Vertical)
8. Resnet50 dengan augmentasi (Rotation + Flip Horizontal)
9. Resnet50 dengan augmentasi (Rotation + Flip Vertical)
10. Resnet50 dengan augmentasi (Flip Horizontal + Flip Vertical)
11. Resnet50 dengan augmentasi (Zoom + Rotation + Flip Horizontal)
12. Resnet50 dengan augmentasi (Zoom + Rotation + Flip Vertical)
13. Resnet50 dengan augmentasi (Zoom + Flip Horizontal + Flip Vertical)

14. Resnet50 dengan augmentasi (Rotation + Flip Horizontal + Flip Vertical)
15. Resnet50 dengan augmentasi (Zoom + Rotation + Flip Horizontal + Flip Vertical)

4.3. Preprocessing

4.3.1. Pengelompokan Data

Pembuat dataset telah membagi data, seperti yang ditunjukkan dalam Gambar 6 Kelas malignant memiliki 4605 gambar data training dan 500 gambar data testing, sementara kelas benign memiliki 5000 gambar data training dan 500 gambar data testing. Beragaman dan representativitas dalam pelatihan dan pengujian model klasifikasi dijamin oleh distribusi data ini. Diharapkan bahwa model dapat belajar dengan lebih baik dari berbagai kasus dan menghasilkan hasil yang lebih akurat saat diuji pada data yang belum pernah dilihat sebelumnya dengan pembagian yang proporsional. Karena rasio data yang seimbang antara kelas malignant dan benign, evaluasi kinerja model dalam mengenali kedua kelas ini sangat penting.

4.3.2. Preprocessing

Sebelum dataset dapat digunakan dalam model klasifikasi, proses preprocessing data sangat penting. Pada awal proses, langkah pengukuran dilakukan untuk standarisasi ukuran file menjadi 224 x 224 piksel. Ini sangat penting saat menggunakan arsitektur Resnet50 yang mengharapkan input dengan

dimensi ini. Hal lain yang perlu diperhatikan adalah bahwa data harus memiliki tiga channel warna (RGB) untuk memenuhi format yang diterima oleh model. Untuk memastikan efisiensi pengolahan data saat melatih model, pengaturan ukuran batch sebesar 32 juga diterapkan. Semua langkah ini digabungkan membentuk tahap awal yang penting untuk membuat lingkungan yang ideal untuk belajar dan menghasilkan hasil yang akurat bagi model klasifikasi.

4.3.3. Augmentasi

Preprocessing Strategi penting dalam penelitian kanker melanoma adalah augmentasi data, yang berfokus pada peningkatan jumlah gambar data dengan mengubah gambar yang sudah ada. Penelitian ini menjamin keseimbangan jumlah data melanoma untuk setiap jenis kanker yang diteliti. Diharapkan variasi data pelatihan yang lebih kaya akan menjadi tujuan utama dari peningkatan ini. Rescale, rotasi, zoom, dan flip vertikal dan horizontal adalah beberapa metode augmentasi yang digunakan, dengan fill mode menggunakan metode terdekat. Diharapkan dengan menerapkan peningkatan ini, penelitian dapat lebih baik memahami dan mengenali pola gambar melanoma secara lebih komprehensif.

Untuk augmentasi gambar, ImageDataGenerator digunakan. Ini adalah teknik yang secara acak mengubah gambar aslinya untuk meningkatkan variasi data pelatihan. Dengan menggunakan perubahan seperti rotasi, zoom, pergeseran horizontal dan vertikal, serta pemutaran gambar, kita dapat mencegah overfitting dan meningkatkan generalisasi model. Memproses data latih, validasi, dan uji dilakukan setelah menentukan augmentasi. Objek generator gambar `train_gen`,

`valid_gen`, dan `test_gen` bertanggung jawab untuk memuat dan memproses gambar dari masing-masing dataset.

`ImageDataGenerator` akan menetapkan konfigurasi untuk mengolah data. Ini termasuk menggunakan `rescaling` untuk mengubah rentang nilai piksel dan menerapkan augmentasi lainnya. Fungsi `flow_from_dataframe` sangat penting untuk melaksanakan generator, terutama ketika data dimuat dari data frame seperti pandas `DataFrame`. Fungsi ini mempermudah proses pengolahan data untuk pelatihan, validasi, dan pengujian model dengan menyesuaikan kolom gambar, kolom label, dan elemen lainnya. Oleh karena itu, penggunaan `ImageDataGenerator` dan generator gambar sangat penting dalam mempersiapkan dataset untuk pelatihan model dengan variasi terbaik.

Jumlah data training yang akan diperoleh dalam tahap augmentasi ini dapat dilihat pada tabel sebagai berikut :

a. Augmentasi Normal

Tabel 4.2. Pembagian Data Augmentasi Normal

Augmentasi	Persebaran Data	
	Malignant	Benign
Original	4605	5000
Zoom	9210	10000
Rotation	9210	10000
Flip Horizontal	9210	10000
Flip Vertical	9210	10000
Zoom + Rotation	13815	15000
Zoom + Flip Horizontal	13815	15000
Zoom + Flip Vertical	13815	15000
Rotation + Flip Horizontal	13815	15000
Rotation + Flip Vertical	13815	15000

Tabel 4.2. Pembagian Data Augmentasi Normal (Lanjutan)

Augmentasi	Persebaran Data	
	Malignant	Benign
Flip Horizontal + Flip Vertical	13815	15000
Zoom + Rotation + Flip Horizontal	18420	20000
Zoom + Rotation + Flip Vertical	18420	20000
Zoom + Flip Horizontal + Flip Vertical	18420	20000
Rotation + Flip Horizontal + Flip Vertical	18420	20000
Zoom + Rotation + Flip Horizontal + Flip Vertical	23025	25000

Tabel 4.2 menunjukkan bahwa peningkatan jumlah data pelatihan dipengaruhi oleh jumlah proses Augmentasi yang diterapkan. Dengan satu proses Augmentasi, setiap kelas data akan dilipatgandakan menjadi dua. Misalnya, jumlah data kelas malignant menjadi 9210 dan kelas benign menjadi 10000. Pada tambahan dua proses Augmentasi berikutnya, jumlah data akan menjadi tiga kali lipat dari jumlah aslinya, dan seterusnya. Sebagai contoh, dengan tiga proses Augmentasi, jumlah data akan menjadi empat kali lipat, dan dengan empat proses Augmentasi, jumlah data akan menjadi lima kali lipat.

b. Augmentasi + Oversampling

Tabel 4.3. Pembagian Data Augmentasi + Oversampling

Augmentasi	Persebaran Data	
	Malignant	Benign
Original	4605	5000
Oversampling	5000	5000
Oversampling (Zoom)	10000	10000
Oversampling (Rotation)	10000	10000

Tabel 4.3. Pembagian Data Augmentasi + Oversampling (Lanjutan)

Augmentasi	Persebaran Data	
	Malignant	Benign
Oversampling (Flip Horizontal)	10000	10000
Oversampling (Flip Vertical)	10000	10000
Oversampling (Zoom + Rotation)	15000	15000
Oversampling (Zoom + Flip Horizontal)	15000	15000
Oversampling (Zoom + Flip Vertical)	15000	15000
Oversampling (Rotation + Flip Horizontal)	15000	15000
Oversampling (Rotation + Flip Vertical)	15000	15000
Oversampling (Flip Horizontal + Flip Vertical)	15000	15000
Oversampling (Zoom + Rotation + Flip Horizontal)	20000	20000
Oversampling (Zoom + Rotation + Flip Vertical)	20000	20000
Oversampling (Zoom + Flip Horizontal + Flip Vertical)	20000	20000
Oversampling (Rotation + Flip Horizontal + Flip Vertical)	20000	20000
Oversampling (Zoom + Rotation + Flip Horizontal + Flip Vertical)	25000	25000

Pada tabel 4.3, dapat dilihat bahwa sebelum penambahan data training dilakukan, data diseimbangkan terlebih dahulu karena terdapat ketidakseimbangan antara kelas malignant dan benign. Keseimbangan ini dicapai dengan menggunakan teknik random oversampling, di mana data dari kelas yang kurang (malignant) diperbanyak secara acak hingga jumlahnya setara dengan kelas yang lebih banyak (benign), sehingga masing-masing kelas memiliki 5000 gambar. Setelah data disetarakan, dilakukan proses augmentasi. Dalam proses ini, penambahan data dilakukan dengan metode augmentasi, di mana 1 augmentasi menggandakan data menjadi 2 kali lipat, 2 augmentasi menjadi 3 kali lipat, dan seterusnya hingga 4 augmentasi yang menggandakan

data menjadi 5 kali lipat. Teknik random oversampling membantu mengatasi ketidakseimbangan data dengan memperbanyak sampel dari kelas minoritas, sehingga model dapat belajar dengan lebih baik dari data yang seimbang.

c. Augmentasi + Gaussian Blur

Tabel 4.4. Pembagian Data Augmentasi + Gaussian Blur

Augmentasi	Persebaran Data	
	Malignant	Benign
Original	4605	5000
Zoom	9210	10000
Rotation	9210	10000
Flip Horizontal	9210	10000
Flip Vertical	9210	10000
Zoom + Rotation	13815	15000
Zoom + Flip Horizontal	13815	15000
Zoom + Flip Vertical	13815	15000
Rotation + Flip Horizontal	13815	15000
Rotation + Flip Vertical	13815	15000
Flip Horizontal + Flip Vertical	13815	15000
Zoom + Rotation + Flip Horizontal	18420	20000
Zoom + Rotation + Flip Vertical	18420	20000
Zoom + Flip Horizontal + Flip Vertical	18420	20000
Rotation + Flip Horizontal + Flip Vertical	18420	20000
Zoom + Rotation + Flip Horizontal + Flip Vertical	23025	25000

Pada tabel 4.4, penambahan data mengikuti pola yang sama seperti penjelasan sebelumnya, yaitu dengan setiap penambahan 1 proses augmentasi, jumlah data training dikali 2, penambahan 2 augmentasi datanya dikali 3, dan

seterusnya. Yang membedakan dari skenario sebelumnya adalah penambahan proses Gaussian Blur, yang membuat seluruh gambar data training menjadi blur, baik gambar asli maupun hasil augmentasi. Gaussian Blur bekerja dengan menghaluskan gambar menggunakan filter Gaussian, yang mengurangi detail dan noise dengan cara menghitung rata-rata bobot piksel di sekitar setiap piksel berdasarkan distribusi Gaussian.

4.4. Membangun Model dan Implementasi

Peneliti memilih model transfer learning Resnet50 untuk digunakan selama tahap implementasi pembangunan model CNN. Untuk karakteristik data input, lapisan input menggunakan dimensi 224 x 224 piksel. Lapisan output terdiri dari lapisan Fully Connected, yang berfungsi sebagai lapisan klasifikasi. Peneliti dapat mengubah model klasifikasi mereka sesuai kebutuhan, dan untuk penelitian ini memilih fungsi softmax CNN sebagai metode klasifikasi. Konfigurasi ini memungkinkan model untuk menggunakan informasi yang diperoleh dari Resnet50 untuk menghasilkan representasi fitur yang kompleks dari data input. Implementasi ini memberikan kerangka kerja yang kuat untuk penelitian, yang memungkinkan modifikasi model klasifikasi untuk memenuhi tujuan penelitian dan memenuhi kebutuhan analisis data tertentu.

Model yang dibuat merupakan model neural network dengan framework keras dan struktur sequential. Model ini terdiri dari lapisan yang saling terhubung. Pertama, digunakan ResNet50 yang telah dilatih sebelumnya dengan data imagenet tanpa penyesuaian khusus. Lapisan GlobalAveragePooling2D() kemudian

mengambil hasil ResNet50 untuk menghasilkan output satu dimensi yang direduksi. Selanjutnya, empat lapisan Dense ditambahkan, masing-masing berisi 256, 128, 64, dan 32 neuron dengan menggunakan fungsi aktivasi ReLU. Terakhir, menggunakan fungsi aktivasi sigmoid, ditambahkan lapisan Dense dengan jumlah neuron yang sesuai dengan jumlah kelas yang ada dalam dataset. Untuk model ini, optimizer Adam menggunakan rate learning $1e-5$. Lost function adalah `categorical_crossentropy` dengan metrik evaluasi akurasi.

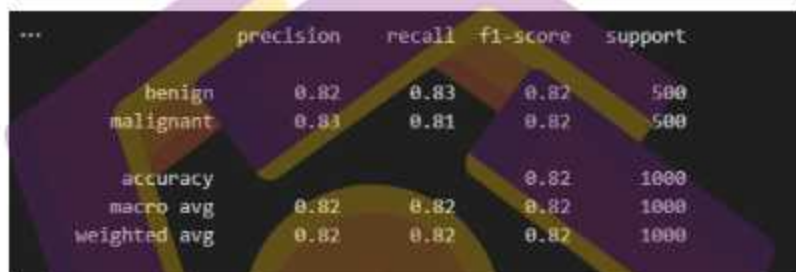
4.5. Hasil Analisis dan Pembahasan

Pada Pada tahap ini, skenario yang telah ditentukan pada penjelasan sebelumnya digunakan untuk mengklasifikasikan kanker melanoma. Hasil dari skenario ini akan dijelaskan secara rinci dan dianalisis. Metode pengujian Confusion Matrix dan Classification Report digunakan. Metode ini tidak hanya menunjukkan akurasi model klasifikasi tetapi juga memberikan nilai precision, recall, dan f1-score yang dapat digunakan untuk memeriksa performa skenario di atas.

4.5.1. Hasil Pelatihan Dari Model

Sebagai langkah awal dalam penelitian ini, dilakukan pelatihan data training tanpa menerapkan proses augmentasi, oversampling, atau gaussian blur. Dataset yang digunakan terdiri dari 5000 gambar benign dan 4605 gambar malignant. Hasil dari pelatihan ini dapat dilihat pada gambar 4.2. Berdasarkan hasil tersebut, model menunjukkan kinerja yang cukup baik dengan akurasi keseluruhan sebesar 82%.

Untuk klasifikasi gambar benign, model mencapai presisi sebesar 82%, recall sebesar 83%, dan f1-score sebesar 82%. Sementara itu, untuk klasifikasi gambar malignant, model memperoleh presisi sebesar 83%, recall sebesar 81%, dan f1-score sebesar 82%. Hasil ini menjadi landasan penting dalam penelitian ini untuk mengevaluasi apakah penerapan teknik augmentasi dapat meningkatkan performa model dalam hal akurasi, presisi, recall, dan f1-score.



```

...
      precision    recall  f1-score   support

 benign         0.82      0.83      0.82     500
 malignant      0.83      0.81      0.82     500

 accuracy              0.82     1000
 macro avg           0.82      0.82      0.82     1000
 weighted avg        0.82      0.82      0.82     1000
  
```

Gambar 4.2. Hasil Evaluasi Matrix Tanpa Augmentasi

4.5.1.1. Skenario Pertama

Tabel 4.5. Hasil Laporan Skenario Pertama

No	Augmentasi	Benign			Malignant			Accuracy
		Recall	Precision	F1-Score	Recall	Precision	F1-Score	
1	Zoom	85	91	88	91	84	87	88
2	Rotation	80	95	87	94	77	84	86
3	Flip Horizontal	73	98	84	97	65	77	81
4	Flip Vertical	85	91	88	90	84	87	87
5	Zoom + Rotation	76	97	85	95	70	80	83
6	Zoom + Flip Horizontal	74	98	84	97	65	78	81
7	Zoom + Flip Vertical	83	93	88	92	81	86	87

Tabel 4.5. Hasil Laporan Skenario Pertama (Lanjutan)

No	Augmentasi	Benign			Malignant			Accuracy
		Recall	Precision	F1-Score	Recall	Precision	F1-Score	
8	Rotation + Flip Horizontal	73	98	84	98	64	77	81
9	Rotation + Flip Vertical	75	97	85	96	68	80	83
10	Flip Horizontal + Flip Vertical	89	82	85	83	90	87	86
11	Zoom + Rotation + Flip Horizontal	84	91	88	91	83	87	87
12	Zoom + Rotation + Flip Vertical	86	89	87	88	85	87	87
13	Zoom + Flip Horizontal + Flip Vertical	85	90	87	90	84	87	87
14	Rotation + Flip Horizontal + Flip Vertical	84	91	87	90	83	86	87
15	Zoom + Rotation + Flip Horizontal + Flip Vertical	89	79	84	81	90	85	85

Tabel 4.5 merupakan hasil augmentasi pada klasifikasi kanker melanoma untuk skenario augmentasi dengan data normal dan menunjukkan bahwa berbagai teknik augmentasi data berbeda dalam memprediksi kelas benign (jinak) dan malignant (ganas), yang diukur dengan metrik akurasi, presisi, recall dan f1-score. Untuk Teknik augmentasi tunggal Zoom memiliki recall tinggi untuk kelas benign dan akurasi tinggi untuk kelas benign, masing-masing dengan akurasi keseluruhan 88% dan rata-rata skor F1 yang kuat yaitu 87,5%, menunjukkan bahwa zoom membantu model membedakan fitur antara kelas benign dan malignant. Meskipun demikian, Flip Horizontal memiliki skor yang rendah untuk varian benign dengan nilai akurasi 81%, recall 73% dan F1-score 84%, yang menunjukkan bahwa teknik ini mungkin kurang efektif untuk varian benign.

Kombinasi augmentasi yang dilakukan menunjukkan bahwa penggunaan Zoom bersama dengan Rotation mengalami penurunan efektivitas, dimana akurasi turun menjadi 83%. Namun, ketika Zoom dikombinasikan dengan Flip Horizontal dan Flip Vertical, terjadi peningkatan dengan akurasi yang lebih seimbang antar kelas, mencapai 87%. Hal ini menunjukkan efektivitas dalam menyeimbangkan pengenalan fitur antar kelas. Sementara itu, kombinasi Rotation dengan Flip Horizontal dan Flip Vertical juga memberikan keseimbangan yang baik dengan akurasi yang serupa, 87%, meskipun ini sedikit cenderung sama bahkan lebih rendah dibandingkan penggunaan teknik flip vertical atau zoom secara tunggal.

Dari semua teknik yang diuji, penggunaan Zoom saja menunjukkan kinerja terbaik dengan akurasi tertinggi 88%, sementara kombinasi multi-augmentasi yang mencakup Zoom, Rotation, Flip Horizontal, dan Flip Vertical menunjukkan kinerja yang sedikit lebih rendah dengan akurasi 85% dibandingkan dengan penggunaan dua atau tiga teknik augmentasi, yang mungkin menunjukkan bahwa terlalu banyak jenis augmentasi dapat mengurangi kemampuan model untuk menggeneralisasi dengan baik.

Meskipun beberapa metode augmentasi menunjukkan hasil yang menjanjikan, mereka masih belum mencapai target akurasi 90%. Metode augmentasi yang lebih sederhana seperti Zoom dan Flip cenderung memberikan hasil yang lebih konsisten dan efektif baik secara individual maupun dalam kombinasi.

4.5.1.2. Skenario Kedua

Tabel 4.6. Hasil Laporan Skenario Kedua

No	Augmentasi	Benign			Malignant			Accuracy
		Recall	Precision	F1-Score	Recall	Precision	F1-Score	
1	Zoom	88	90	89	90	88	89	89
2	Rotation	86	93	90	93	85	89	89
3	Flip Horizontal	83	97	90	98	81	88	89
4	Flip Vertical	88	90	89	90	88	89	89
5	Zoom + Rotation	84	91	88	91	83	87	87
6	Zoom + Flip Horizontal	89	82	86	84	90	87	86
7	Zoom + Flip Vertical	50	100	67	0	0	0	50
8	Rotation + Flip Horizontal	50	100	67	0	0	0	50
9	Rotation + Flip Vertical	74	98	84	97	65	78	81
10	Flip Horizontal + Flip Horizontal	89	82	86	84	90	87	86
11	Zoom + Rotation + Flip Horizontal	90	81	85	82	91	87	86
12	Zoom + Rotation + Flip Vertical	85	95	89	94	83	88	89
13	Zoom + Flip Horizontal + Flip Vertical	50	100	67	0	0	0	50
14	Rotation + Flip Horizontal + Flip Vertical	90	81	85	82	91	86	86
15	Zoom + Rotation + Flip Horizontal + Flip Vertical	79	99	88	98	73	84	86

Tabel 4.6 merupakan hasil augmentasi pada skenario kedua klasifikasi kanker melanoma, di mana proses oversampling ditambahkan dengan harapan mampu mencapai target akurasi 90%, menunjukkan variasi kinerja pada berbagai metode augmentasi. Semua teknik augmentasi data, yaitu Rotation, Flip Horizontal, Flip Vertical, dan Zoom, menunjukkan performa yang solid dalam mendeteksi ciri-

ciri kedua kelas kanker kulit. Rotation menonjol dengan akurasi 89%, recall 93% untuk malignant, dan F1-score 90% untuk benign, menunjukkan efektivitasnya dalam meningkatkan deteksi ciri-ciri kedua kelas. Flip Horizontal juga menunjukkan hasil yang baik dengan akurasi dan F1-score 89%, terutama pada kelas malignant dengan recall 98%. Flip Vertical dan Zoom memiliki akurasi dan F1-scores yang solid di 89%, dengan keseimbangan yang baik antara recall dan precision untuk kedua kelas.

Kombinasi teknik augmentasi seperti Zoom + Rotation + Flip Vertical telah terbukti efektif, dengan akurasi dan skor F1 yang tinggi sebesar 89% dan recall dan precision yang tinggi untuk kelas malignant. Kombinasi Zoom + Rotation + Flip Horizontal juga telah terbukti efektif, dengan akurasi sebesar 86% dan skor F1 dan recall yang tinggi untuk kelas malignant. Namun, kombinasi augmentasi lain menunjukkan performa yang jauh lebih baik.

Teknik yang menggunakan rotasi, baik secara individual maupun dalam kombinasi dengan metode lain, cenderung mencapai skor yang lebih tinggi. Sebaliknya, beberapa percobaan menunjukkan bahwa kombinasi teknik yang melibatkan rotasi horizontal dan/atau vertical cenderung menghasilkan performa yang lebih rendah, terutama dalam identifikasi penyakit malignant. Dalam proyek pengenalan pola gambar medis, konsistensi ini penting untuk dipertimbangkan saat mengembangkan dan menerapkan strategi augmentasi gambar.

Hasil penelitian menyoroti bahwa, meskipun oversampling telah digunakan untuk meningkatkan kinerja deteksi, beberapa metode peningkatan, khususnya yang melibatkan rotasi dan flip horizontal, terlihat lebih efektif daripada pendekatan yang

memanfaatkan zoom dan kombinasi augmentasi yang lebih kompleks. Performa yang rendah dari beberapa kombinasi diperkuat oleh temuan (Andrianto Iskandar & Salam, 2024) yang menunjukkan bahwa hasil eksperimen mereka hanya sedikit berbeda, menunjukkan pengaruh yang tidak terlalu signifikan. Meskipun beberapa penelitian seperti yang disebutkan oleh (Mohammed et al., 2020) menyebutkan keunggulan oversampling, (Tanti, 2023) menunjukkan peningkatan nilai recall, dan (Diantika et al., 2024) menyatakan peningkatan kinerja model, namun tidak sesuai dengan harapan. Hal ini menegaskan bahwa tidak semua jenis augmentasi sesuai untuk meningkatkan akurasi model klasifikasi melanoma ditambah dengan oversampling.

4.5.1.3. Skenario Ketiga

Tabel 4.7. Hasil Laporan Skenario Ketiga

No	Augmentasi	Benign			Malignant			Accuracy
		Recall	Precision	F1-Score	Recall	Precision	F1-Score	
1	Zoom	84	99	91	99	81	89	90
2	Rotation	85	98	91	97	83	90	90
3	Flip Horizontal	86	96	91	96	84	90	90
4	Flip Vertical	87	97	92	96	86	91	91
5	Zoom + Rotation	78	100	88	100	72	83	86
6	Zoom + Flip Horizontal	85	98	91	98	83	90	91
7	Zoom + Flip Vertical	81	98	89	98	77	86	87
8	Rotation + Flip Horizontal	87	97	92	97	86	91	91
9	Rotation + Flip Vertical	82	99	90	99	78	87	89
10	Flip Horizontal + Flip Vertical	84	99	91	98	81	89	90
11	Zoom + Rotation + Flip Horizontal	83	99	90	99	79	88	89

Tabel 4.7. Hasil Laporan Skenario Ketiga (Lanjutan)

No	Augmentasi	Benign			Malignant			Accuracy
		Recall	Precision	F1-Score	Recall	Precision	F1-Score	
12	Zoom + Rotation + Flip Vertical	82	92	89	98	78	87	88
13	Zoom + Flip Horizontal + Flip Vertical	85	100	92	100	83	90	91
14	Rotation + Flip Horizontal + Flip Vertical	83	99	90	99	79	88	89
15	Zoom + Rotation + Flip Horizontal + Flip Vertical	83	99	90	99	80	88	89

Dalam skenario ketiga untuk klasifikasi kanker melanoma, Tabel 4.7 mengungkapkan peningkatan hasil yang signifikan dengan penambahan Gaussian blur, mencapai target akurasi yang diharapkan sebesar 90%. Teknik-teknik seperti flip horizontal, rotasi, dan zoom menunjukkan performa yang kuat, semuanya berhasil mencapai skor akurasi dan F1-score sebesar 90%. Lebih jauh lagi, teknik-teknik ini menyajikan recall dan precision yang seimbang untuk kedua kelas, yang berkontribusi pada F1-score yang tinggi.

Selain itu, skenario augmentasi yang menggunakan teknik Flip Vertical dan gabungan dari Rotation + Flip Horizontal dengan Gaussian Blur menunjukkan hasil yang paling mengesankan dengan akurasi model sebesar 91%, untuk presisi mendapat nilai sebesar 97%, recall sebesar 87% dan f1-score sebesar 92%. Sedangkan untuk hasil klasifikasi kelas malignant mendapat nilai presisi sebesar 86% dan f1-score sebesar 91%, untuk recall kedua metode ini mendapatkan hasil yang berbeda. Untuk recall kelas malignant pada augmentasi Flip Vertical mendapat skor sebesar 96% sedangkan pada augmentasi gabungan Rotation + Flip

Horizontal mendapat skor sedikit lebih tinggi yaitu 97% artinya gabungan Rotation + Flip Horizontal sedikit lebih baik.. Hasil ini menunjukkan bahwa metode gabungan ini sangat efektif.

Meskipun berbagai kombinasi teknik augmentasi telah dieksplorasi, tidak semua berhasil memberikan hasil yang optimal. Khusus untuk kelas malignant, kombinasi teknik zoom dengan rotasi hanya mencapai akurasi 86% dan FI-score yang relatif rendah, yang menunjukkan bahwa penambahan Gaussian blur tidak secara signifikan meningkatkan performa. Hal ini mengindikasikan bahwa tidak semua kombinasi teknik augmentasi memberikan dampak positif terhadap hasil klasifikasi.

Dari hasil penelitian ini, kita dapat menyimpulkan bahwa teknik augmentasi yang melibatkan flip vertical, baik secara mandiri maupun dalam kombinasi dengan metode lain seperti rotasi, flip horizontal, dan zoom, cenderung menghasilkan performa terbaik dalam mencapai target akurasi 90%. Sementara metode seperti rotasi, flip horizontal, dan zoom memang dapat meningkatkan kinerja model, penambahan Gaussian blur tidak selalu menghasilkan peningkatan yang signifikan. Temuan dari (Abayomi-Alli et al., 2021) menunjukkan bahwa pengujian akurasi antara gambar kualitas rendah dan tinggi dengan Gaussian blur memberikan hasil menarik. Meskipun biasanya gambar berkualitas tinggi memberikan hasil yang lebih baik dalam klasifikasi, penelitian ini menyoroti bahwa dengan menggunakan augmentasi data yang tepat, model dapat memberikan klasifikasi yang baik bahkan dengan gambar berkualitas rendah. Oleh karena itu, pemilihan dan kombinasi teknik

augmentasi yang tepat sangat penting untuk meningkatkan kinerja model klasifikasi melanoma.

4.5.2. Hasil Perbandingan Pengujian

Tabel 4.8 menunjukkan perbandingan akurasi dari berbagai teknik augmentasi dalam klasifikasi kanker melanoma dengan menggunakan metode normal, oversampling, dan Gaussian Blur. Tujuan analisis ini adalah untuk mengevaluasi seberapa efektif masing-masing teknik dalam mencapai target akurasi 90%.

Tabel 4.8. Hasil Akurasi Tiap Skenario

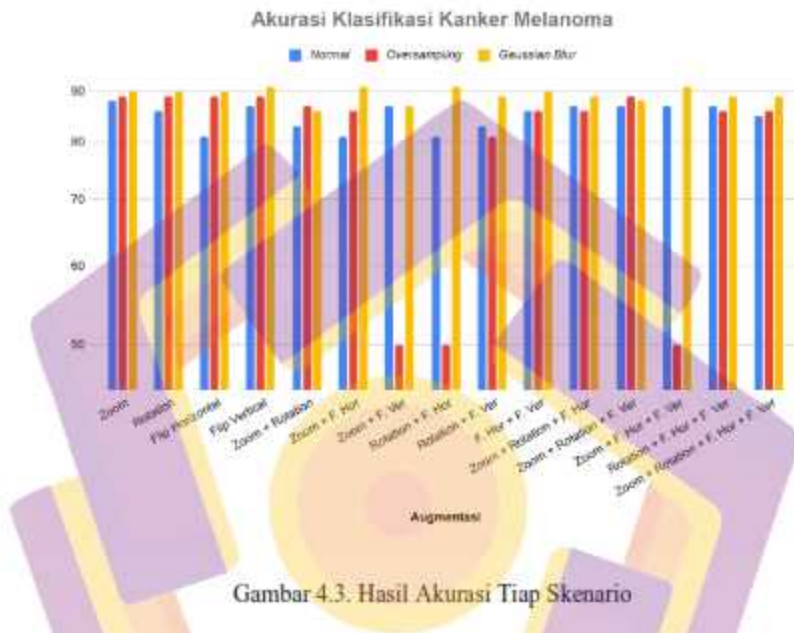
Augmentasi	Normal	Oversampling	Gaussian Blur
Zoom	88	89	90
Rotation	86	89	90
Flip Horizontal	81	89	90
Flip Vertical	87	89	91
Zoom + Rotation	83	87	86
Zoom + F. Hor	81	86	91
Zoom + F. Ver	87	50	87
Rotation + F. Hor	81	50	91
Rotation + F. Ver	83	81	89
F. Hor + F. Ver	86	86	90
Zoom + Rotation + F. Hor	87	86	89
Zoom + Rotation + F. Ver	87	89	88
Zoom + F. Hor + F. Ver	87	50	91
Rotation + F. Hor + F. Ver	87	86	89
Zoom + Rotation + F. Hor + F. Ver	85	86	89

Dari hasil perbandingan tiga skenario augmentasi, terungkap bahwa setiap metode augmentasi memiliki keunggulan tersendiri tergantung pada kombinasi augmentasi yang digunakan. Dalam skenario tanpa augmentasi, teknik Zoom, Rotation, dan Flip Horizontal menghasilkan akurasi yang mendekati atau melebihi 85%, dengan kombinasi Flip Horizontal dan Flip Vertical mencapai akurasi tertinggi sebesar 89%. Hal ini menunjukkan bahwa bahkan tanpa augmentasi tambahan, pemilihan teknik yang tepat dapat menghasilkan performa yang cukup baik.

Namun, augmentasi dengan oversampling menunjukkan variasi performa yang signifikan. Teknik Zoom, Rotation, Flip Horizontal dan Flip Vertical menunjukkan hasil yang stabil dengan akurasi 89%, tetapi beberapa kombinasi, seperti Zoom + Flip Horizontal + Flip Vertical, hanya mencapai akurasi 50%. Ini menunjukkan bahwa oversampling tidak selalu efektif untuk semua jenis peningkatan dalam konteks data ini, terutama ketika peningkatan mengubah orientasi gambar secara signifikan, yang dapat menyebabkan representasi data yang tidak akurat atau bahkan menurunkan performa model.

Sebaliknya, augmentasi dengan Gaussian Blur memberikan hasil yang paling konsisten dan tinggi di antara semua skenario. Teknik Zoom, Flip Horizontal, dan Flip Vertical yang digunakan dengan Gaussian Blur menunjukkan peningkatan akurasi yang konsisten hingga 90% atau lebih. Contoh pada kombinasi Flip Vertical dengan Gaussian Blur mencapai akurasi tertinggi sebesar 91%, menunjukkan bahwa metode ini sangat efektif dalam menangani variasi dalam data gambar. Gaussian Blur membantu model berkonsentrasi pada fitur penting dengan

menghaluskan gambar, sehingga meningkatkan kemampuan prediktifnya dalam kasus yang kompleks..



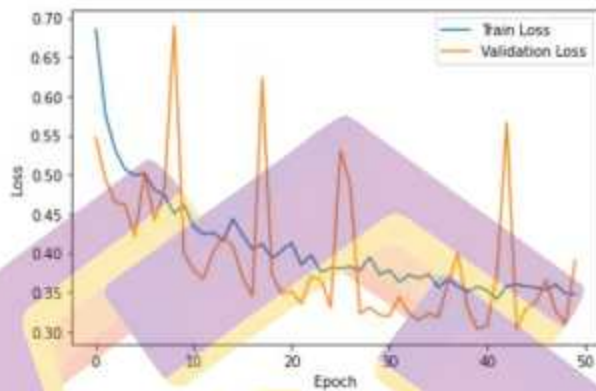
Gambar 4.3 juga menunjukkan keunggulan Gaussian Blur dibandingkan skenario lainnya, meskipun ada beberapa kombinasi augmentasi dengan oversampling yang mencapai hasil yang sebanding. Misalnya, kombinasi Zoom, Rotation, dan Flip Horizontal menunjukkan nilai akurasi yang sama-sama tinggi antara oversampling dan Gaussian Blur, yaitu 89%. Namun, hasil keseluruhan menunjukkan bahwa Gaussian Blur lebih konsisten dalam mencapai atau mendekati target akurasi 90%, sedangkan oversampling menunjukkan variasi yang lebih besar dalam performa.

4.5.3. Faktor yang Mempengaruhi Hasil

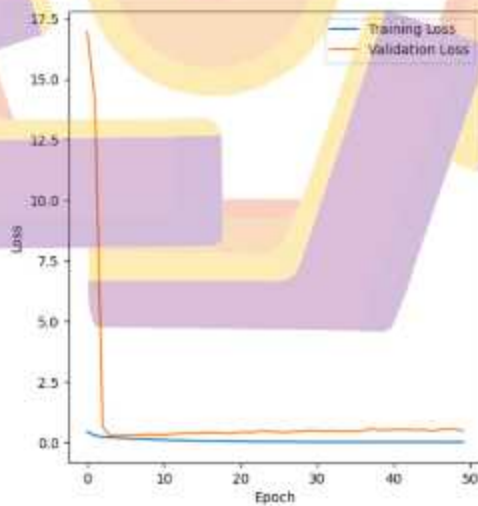
Secara keseluruhan, tampaknya menambahkan Gaussian Blur meningkatkan akurasi model klasifikasi kanker melanoma daripada hanya menggunakan oversampling atau tanpa metode tambahan. Hasil yang tepat dari berbagai kursus pelatihan dipengaruhi oleh banyak hal. Penambahan teknik Gaussian Blur adalah salah satu komponen yang paling menonjol. Teknik ini meningkatkan akurasi, dengan skor tertinggi sebesar 91% pada salah satu proses augmentasi, Rotation + Flip Horizontal. Nilai akurasi ini berkorelasi dengan skor $f1$ setiap kelas, yang dipengaruhi oleh keseimbangan antara nilai presisi dan recall masing-masing kelas. Sebagai contoh, pada skenario pertama dengan proses augmentasi Flip Horizontal, terjadi ketidakseimbangan yang jelas. Hal ini menyebabkan hasil akurasi yang didapat pada proses augmentasi Flip Horizontal hanya mencapai 81%, jauh dari target peneliti sebesar 90%.

Dari sisi augmentasi, dapat dilihat pada skenario satu, dua, dan tiga bahwa hasil yang didapat bervariasi. Uniknya, pada proses augmentasi tunggal, hasil yang didapat justru lebih baik ketimbang gabungan dari beberapa augmentasi. Sebagai contoh, pada skenario kedua, semua yang menggunakan proses augmentasi tunggal seperti Zoom, Rotation, Flip Horizontal, dan Vertical mendapatkan nilai akurasi yang sama rata yaitu 89%, sementara kombinasi dari keempat augmentasi tersebut hanya mendapatkan 86%. Meskipun belum mencapai target 90%, ini menunjukkan bahwa pemilihan kombinasi augmentasi yang tepat sangat penting dan berpengaruh signifikan terhadap tingkat akurasi. Penelitian lebih lanjut diperlukan untuk

menemukan kombinasi augmentasi yang optimal guna mencapai hasil yang diharapkan.



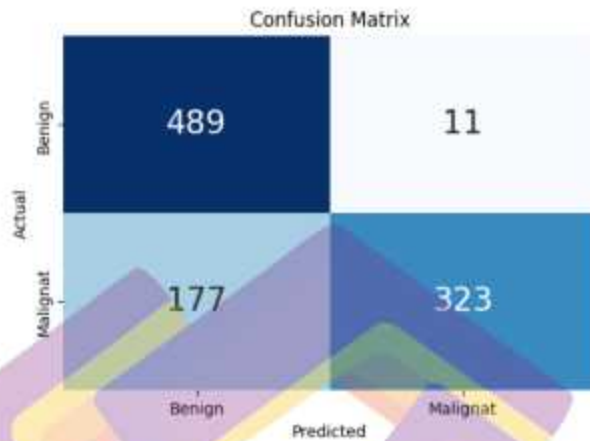
Gambar 4.4. Perbandingan Train Loss dan Validation Loss pada skenario 1 Augmentasi Flip Horizontal



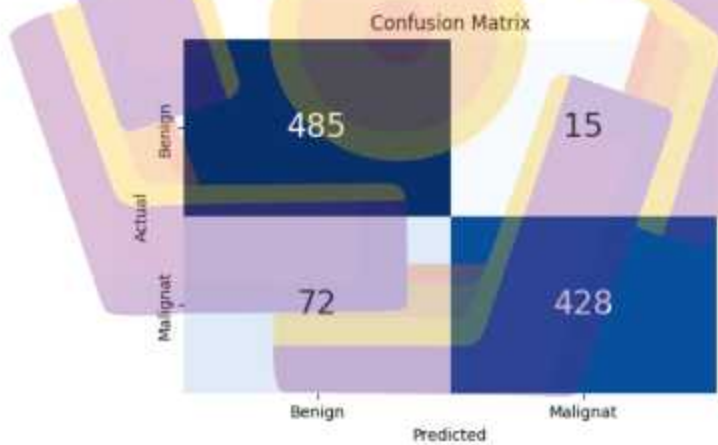
Gambar 4.5. Perbandingan Train Loss dan Validation Loss pada skenario 3 Augmentasi Flip Horizontal

Gambar 4.4 dan Gambar 4.5 menunjukkan dua grafik yang membandingkan antara "Training Loss" dan "Validation Loss" untuk dua skenario yang berbeda dalam klasifikasi kanker melanoma. Skenario 1 (Flip Horizontal) menunjukkan bahwa "Validation Loss" cenderung lebih tinggi dan fluktuatif dibandingkan dengan "Training Loss", menandakan adanya overfitting. Overfitting terjadi ketika model belajar terlalu spesifik pada data latihan sehingga kurang mampu menggeneralisasi pada data validasi. Grafik ini memperlihatkan bahwa meskipun "Training Loss" menurun, "Validation Loss" tetap tidak stabil dan lebih tinggi, menunjukkan model tidak berfungsi baik pada data baru.

Sebaliknya, Skenario 3 (Flip Horizontal) menunjukkan performa yang lebih baik dengan "Training Loss" dan "Validation Loss" yang stabil dan rendah setelah menurun cepat pada awalnya. Kedua metrik ini hampir sama, menunjukkan bahwa model mampu belajar dari data latihan tanpa kehilangan kemampuan generalisasi pada data validasi. Tidak ada tanda-tanda overfitting atau underfitting dalam skenario ini, yang berarti model memiliki performa yang baik pada data latihan dan validasi. Dalam konteks klasifikasi kanker melanoma, keseimbangan antara "Training Loss" dan "Validation Loss" sangat penting untuk memastikan model dapat bekerja dengan baik pada data baru yang belum pernah dilihat sebelumnya.



Gambar 4.6. Confusion Matix Skenario 1 Augmentasi Flip Horizontal



Gambar 4.7. Confusion Matix Skenario 3 Augmentasi Flip Horizontal

Gambar 4.6 dan Gambar 4.7 menunjukkan dua matriks kebingungan yang menggambarkan hasil klasifikasi kanker melanoma dalam dua skenario berbeda,

yaitu "Flip Vertical" dan "Flip Horizontal". Dalam Skenario 3 (Flip Horizontal), model menunjukkan kinerja yang cukup baik dengan 485 prediksi benar untuk kasus benign dan 428 prediksi benar untuk kasus malignant. Namun, ada juga beberapa kesalahan, yaitu 15 kasus benign yang diprediksi sebagai malignant dan 72 kasus malignant yang diprediksi sebagai benign. Confusion Matrix ini menggambarkan kemampuan model dalam membedakan antara kasus benign dan malignant dengan tingkat kesalahan yang relatif rendah.

Pada Skenario 1 (Flip Horizontal), hasil klasifikasi menunjukkan bahwa model melakukan 489 prediksi benar untuk kasus benign dan 323 prediksi benar untuk kasus malignant. Namun, tingkat kesalahan lebih tinggi dibandingkan dengan Skenario 3, dengan 11 kasus benign yang salah diklasifikasikan sebagai malignant dan 177 kasus malignant yang salah diklasifikasikan sebagai benign. Perbandingan antara kedua skenario ini mengindikasikan bahwa faktor-faktor seperti teknik augmentasi atau pemrosesan data (seperti flipping) dapat mempengaruhi performa model dalam mengklasifikasikan kanker melanoma, dengan Skenario 3 tampaknya memberikan hasil yang lebih akurat dibandingkan Skenario 1.

4.5.4. Perbandingan dengan Penelitian Sebelumnya

Tabel 4.9 menyajikan perbandingan mendalam antara hasil penelitian ini dan penelitian-penelitian terdahulu yang menggunakan dataset yang sama dan dianggap sebagai yang terbaik dalam klasifikasi kanker melanoma. Analisis ini sangat penting untuk memahami kemajuan yang telah dicapai dalam mengembangkan metode diagnostik yang lebih efektif dan akurat, serta untuk

mengidentifikasi kendala yang masih ada. Perbandingan ini membantu mengevaluasi seberapa jauh teknik yang dikembangkan dalam penelitian ini meningkatkan akurasi dan efektivitas diagnostik dibandingkan dengan pendekatan sebelumnya.

Tabel 4.9. Perbandingan Penelitian Sebelumnya

Peneliti	Algoritma	Accuracy
(Khasanah et al., 2021)	Random Forest	84.2%
(Savera et al., 2020)	K-NN dan CNN	76%
(Milton, 2019)	Deep Neural Network	73%
(Indraswari et al., 2021)	MobileNetV2	85%
(Jojoa Acosta et al., 2021)	ResNet152	90.4%
(Yusuf et al., 2024)	CNN	85%
(Ashari et al., 2024)	DenseNet121	90.63%
Proposed Model	ResNet-50	91%

Analisis dalam penelitian klasifikasi kanker melanoma menunjukkan bahwa model yang menggunakan ResNet-50 mencapai hasil yang sangat mengesankan dengan akurasi sebesar 91%. Ini merupakan peningkatan signifikan dibandingkan dengan metode yang digunakan sebelumnya. ResNet-50, yang merupakan arsitektur jaringan neural konvolusional yang cukup mendalam, menggunakan koneksitas residual untuk mengatasi masalah menghilangnya gradien yang sering terjadi di jaringan dengan banyak lapisan.

Ketika dibandingkan dengan algoritma lain yang digunakan dalam studi sebelumnya, keunggulan ResNet-50 menjadi semakin jelas. Algoritma Random Forest yang digunakan oleh Khasanah et al. mencapai akurasi 84.2%, sementara metode K-NN dan CNN oleh Savera et al. serta Deep Neural Network oleh Milton

masing-masing hanya mencapai akurasi 76% dan 73%. Milton menyoroti tantangan dalam generalisasi fitur visual dari lesi akibat ketidakseimbangan data dan pentingnya variasi fitur yang lebih baik. Hasil penelitian ini menunjukkan bahwa teknik oversampling yang tepat dapat mengatasi ketidakseimbangan dataset dan meningkatkan kinerja klasifikasi secara signifikan.

Algoritma MobileNetV2 oleh Indraswari et al. dan CNN oleh Yusuf et al. masing-masing mencapai akurasi 85%, menunjukkan bahwa arsitektur CNN yang lebih kompleks dapat memberikan hasil yang lebih baik dibandingkan metode yang lebih sederhana. Indraswari et al. menekankan pentingnya pemilihan arsitektur dan teknik augmentasi yang tepat untuk meningkatkan akurasi model.

Jojoa Acosta et al. menggunakan ResNet152 dan mencapai akurasi 90,4%. Ini menunjukkan bahwa arsitektur yang lebih dalam dan kompleks dapat mendekati performa ResNet-50 dengan kompleksitas yang lebih tinggi. Penelitian oleh Ashari et al. menggunakan DenseNet121 dan mencapai akurasi 90,63%, mendekati performa ResNet-50. Ini menegaskan bahwa arsitektur jaringan yang dalam dapat meningkatkan kinerja klasifikasi. Namun, dengan strategi augmentasi yang tepat, ResNet-50 mampu melampaui hasil penelitian sebelumnya, menunjukkan potensinya dalam meningkatkan akurasi klasifikasi pada dataset yang beragam dan tidak seimbang.

Kesimpulan dari analisis ini menyoroti bahwa penggunaan arsitektur CNN yang lebih canggih dan mendalam, seperti ResNet-50, dapat memberikan hasil yang lebih unggul dalam klasifikasi kanker melanoma. Performa tinggi dari model ini

membuka potensi besar untuk aplikasinya dalam pengembangan alat diagnostik yang lebih akurat dan dapat diandalkan di bidang medis, khususnya untuk deteksi dan klasifikasi kanker melanoma. Keberhasilan ini juga menegaskan pentingnya terus mendorong inovasi dan penelitian dalam algoritma pembelajaran mendalam agar bisa terus meningkatkan kualitas diagnostik dalam kedokteran.

4.5.5. Upaya Perbaiki Kinerja Algoritma

Dalam penelitian ini, akurasi sebesar 91% sudah mendekati target, tetapi tingkat salah deteksi sebesar 10% masih tidak dapat diterima dalam bidang medis. Untuk memperbaiki kinerja algoritma, beberapa upaya berikut dapat dipertimbangkan: Pertama, teknik augmentasi data yang lebih beragam dapat membantu model menjadi lebih tahan terhadap variasi data. Teknik seperti Color Jitter, Random Erasing, dan Elastic Transformations dapat meningkatkan latihan data dengan berbagai variasi, memungkinkan model untuk belajar lebih baik dalam berbagai kondisi. Selain itu, kinerja model dapat secara signifikan ditingkatkan dengan menerapkan transfer learning dengan menggunakan model pra-terlatih yang sudah dioptimalkan pada dataset besar seperti ImageNet dan kemudian menerapkan pelatihan lanjutan pada dataset khusus yang berkaitan dengan kanker melanoma. Transfer learning memungkinkan model memanfaatkan fitur-fitur yang sudah dipelajari dari dataset besar, sehingga mempercepat proses pelatihan dan meningkatkan akurasi.

Kedua, peningkatan kualitas data adalah komponen penting dalam peningkatan kinerja algoritma. Model dapat meningkatkan kemampuan deteksi dengan mendapatkan data dengan resolusi lebih tinggi. Selain itu, mendapatkan data

anotasi yang lebih akurat dapat membantu model belajar dengan lebih baik, mengurangi kesalahan klasifikasi. Untuk meningkatkan kemampuan generalisasi model, penambahan data latih dengan mengumpulkan lebih banyak data dari berbagai sumber juga penting. Dengan variasi data yang lebih besar, model lebih baik dalam menangani berbagai kondisi yang mungkin muncul dalam data baru.

Ketiga, dapat membantu mengurangi overfitting dengan menggunakan metode regularisasi seperti dropout atau L2 regularization. Overfitting terjadi ketika model terlalu berfokus pada data latih sehingga kurang mampu menggeneralisasi pada data validasi atau data baru. Teknik regularisasi membantu model untuk tidak terlalu bergantung pada fitur tertentu dari data latih, sehingga lebih mudah untuk generalisasi. Untuk menemukan konfigurasi model yang ideal, penyesuaian hyperparameter yang lebih mendalam diperlukan. Setiap model memiliki berbagai hyperparameter yang perlu disesuaikan agar model dapat bekerja dengan baik, dan menemukan hyperparameter yang paling ideal dapat secara signifikan meningkatkan performa model.

Keempat, untuk mengatasi overfitting secara lebih efektif, beberapa pendekatan tambahan dapat diterapkan. Salah satunya adalah penggunaan Cross-Validation, yang dapat membantu memastikan bahwa model tidak hanya cocok pada satu subset data, tetapi bekerja baik pada berbagai subset data yang berbeda. Menggunakan lebih banyak data latih juga bisa membantu mengurangi overfitting, karena model akan memiliki lebih banyak variasi data untuk belajar. Early Stopping adalah teknik lain yang dapat digunakan, di mana pelatihan dihentikan ketika kinerja model pada data validasi mulai memburuk, sehingga mencegah model untuk

terlalu menyesuaikan dengan data latih. Dengan kombinasi berbagai teknik ini, diharapkan model dapat mengurangi overfitting dan meningkatkan kinerjanya secara keseluruhan.

Terakhir, akurasi dan stabilitas prediksi dapat ditingkatkan melalui pengajaran kelompok dengan menggabungkan prediksi dari berbagai model. Memanfaatkan metode ensemble seperti bagging dan boosting, model dapat mengompensasi kelemahan satu sama lain, meningkatkan performa secara keseluruhan. Teknik Explainable AI (XAI) seperti Grad-CAM juga dapat membantu memahami bagian mana dari gambar yang dilihat oleh model. Ini tidak hanya membantu meningkatkan dan memperbaiki model yang didasarkan pada interpretasi ini, tetapi juga memberikan kepercayaan lebih kepada pengguna akhir, seperti dokter, bahwa model membuat keputusan berdasarkan fitur yang relevan. Metode-metode ini diharapkan dapat meningkatkan kinerja algoritma untuk mendeteksi kanker melanoma dengan akurasi yang lebih tinggi dan kepercayaan yang lebih besar.

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan Sebagai hasil dari analisis dan pembahasan yang dilakukan dalam penelitian ini, dapat disimpulkan bahwa:

- a. Dari data dan analisis yang disajikan, terlihat bahwa proses augmentasi berdampak besar pada akurasi, presisi, recall, dan skor f1 dalam klasifikasi kanker melanoma. Penggunaan metode augmentasi yang berbeda, seperti Zoom, Flip, dan Rotation, baik secara individual maupun bersamaan, telah menunjukkan perubahan signifikan dalam hasil klasifikasi. Secara khusus, semua metrik evaluasi menunjukkan peningkatan yang signifikan ketika Gaussian Blur digunakan sebagai metode augmentasi tambahan; ini menunjukkan bahwa penggunaan augmentasi yang tepat dapat meningkatkan kemampuan model untuk mengidentifikasi pola dengan lebih efektif dan mengurangi overfitting.
- b. Berdasarkan analisis skenario yang diberikan, skenario augmentasi yang menggunakan teknik Flip Vertical dan gabungan dari Rotation + Flip Horizontal dengan Gaussian Blur menunjukkan hasil yang paling mengesankan dengan akurasi model sebesar (91%), untuk presisi mendapat nilai sebesar (97%), recall sebesar (87%) dan f1-score sebesar (92%). Sedangkan untuk hasil klasifikasi kelas malignant mendapat nilai presisi sebesar (86%) dan f1-score sebesar (91%), untuk recall kedua

metode ini mendapatkan hasil yang berbeda. Untuk recall kelas malignant pada augmentasi Flip Vertical mendapat skor sebesar (96%) sedangkan pada augmentasi gabungan Rotation + Flip Horizontal mendapat skor sedikit lebih tinggi yaitu (97%) artinya gabungan Rotation + Flip Horizontal sedikit lebih baik. Hasil ini menunjukkan bahwa kombinasi teknik augmentasi tertentu, tergantung pada karakteristik data dan model yang digunakan, mungkin lebih efektif. Situasi ini berhasil mencapai tingkat akurasi di atas 90 persen, yang sesuai dengan ambang penelitian.

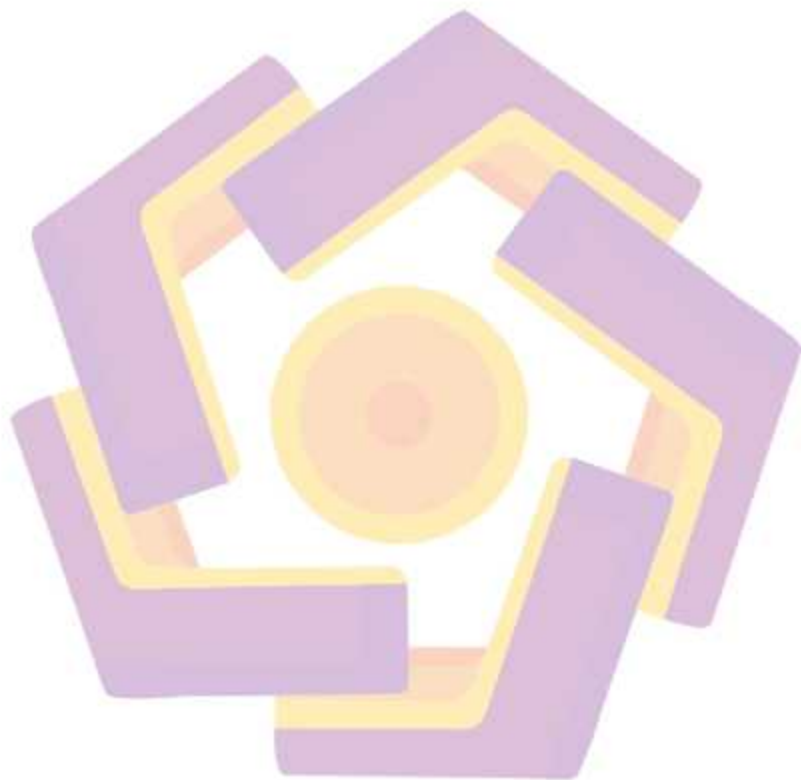
- c. Penelitian membandingkan berbagai algoritma klasifikasi untuk gambar kanker melanoma, dan menemukan bahwa model dengan arsitektur CNN Resnet-50 menghasilkan akurasi tertinggi (91%). Hal ini menunjukkan keefektifan arsitektur CNN yang mendalam dalam mengatasi permasalahan klasifikasi gambar melanoma, di mana akurasi sangatlah krusial untuk diagnosis yang tepat. Oleh karena itu, investasi dalam teknologi canggih seperti Resnet-50 berpotensi meningkatkan keandalan dan presisi sistem diagnostik medis secara signifikan.

5.2. Saran

Berdasarkan Berdasarkan kesimpulan yang telah dirumuskan, berikut adalah empat saran yang dapat dipertimbangkan untuk meningkatkan pengembangan dan efektivitas sistem klasifikasi kanker melanoma:

- a. Karena metode augmentasi seperti Gaussian Blur meningkatkan akurasi, disarankan untuk mengeksplorasi dan menerapkan berbagai metode augmentasi. Eksperimen yang menggabungkan berbagai metode augmentasi dapat membantu dalam menentukan set augmentasi yang paling efektif untuk berbagai jenis data klinis.
- b. Dalam penelitian ini, Resnet-50 menunjukkan hasil yang sangat baik. Namun, untuk terus meningkatkan akurasi dan efisiensi, arsitektur CNN yang mendalam harus disesuaikan dan dituning. Penelitian lebih lanjut juga bisa mencakup penggunaan arsitektur baru, yang mungkin lebih efisien atau lebih sesuai dengan kekhasan data melanoma.
- c. Untuk memastikan bahwa teknologi ini memenuhi kebutuhan klinis dan dapat digunakan secara luas dalam praktik medis, pengembang AI, peneliti medis, dan praktisi kesehatan harus bekerja sama untuk mengintegrasikan sistem klasifikasi yang telah dikembangkan ke dalam alat diagnostik medis.
- d. Perbaikan kualitas gambar dalam dataset juga harus dipertimbangkan, selain metode augmentasi dan tuning arsitektur. Metode yang digunakan untuk meningkatkan kualitas gambar termasuk perbaikan kontras menggunakan Histogram Equalization atau Adaptive Histogram Equalization, penyempurnaan keseimbangan putih dengan Algoritme Grey World atau White Patch, pengurangan suara dengan Median Filtering atau Gaussian Smoothing, peningkatan ketajaman gambar menggunakan Unsharp Masking atau High Pass Filtering, normalisasi nilai piksel dengan Z-Score Normalization atau Min-Max Scaling, dan koreksi warna dengan Color

Balance. Diharapkan bahwa dengan menggunakan metode ini, model akan memiliki performa dan akurasi yang lebih baik. Namun, penelitian lebih lanjut diperlukan untuk menemukan kombinasi metode perbaikan terbaik.



DAFTAR PUSTAKA

- Abayomi-Alli, O. O., Damaševičius, R., Misra, S., & Maskeliūnas, R. (2021). Cassava disease recognition from low-quality images using enhanced data augmentation model and deep learning. *Expert Systems*, 38(7). <https://doi.org/10.1111/exsy.12746>
- Adegun, A. A., & Viriri, S. (2020). Deep learning-based system for automatic melanoma detection. *IEEE Access*, 8, 7160–7172. <https://doi.org/10.1109/ACCESS.2019.2962812>
- Adepu, A. K., Sahayam, S., Jayaraman, U., & Arramraju, R. (2023). Melanoma classification from dermatoscopy images using knowledge distillation for highly imbalanced data. *Computers in Biology and Medicine*, 154(January), 106571. <https://doi.org/10.1016/j.compbiomed.2023.106571>
- AGUSTINA, R., MAGDALENA, R., & PRATIWI, N. K. C. (2022). Klasifikasi Kanker Kulit menggunakan Metode Convolutional Neural Network dengan Arsitektur VGG-16. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, 10(2), 446. <https://doi.org/10.26760/elkomika.v10i2.446>
- Algayres, R., Zaiem, M. S., Sagot, B., & Dupoux, E. (2020). Evaluating the reliability of acoustic speech embeddings. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2020-October, 4621–4625. <https://doi.org/10.21437/Interspeech.2020-2362>

- Andrianto Iskandar, D., & Salam, A. (2024). *Evaluasi Performa Oversampling dan Augmentasi pada Klasifikasi Penyakit Kulit Menerapkan Convolutional Neural Network*. <https://doi.org/10.30865/mib.v8i1.7119>
- Ashari, M. R., Sari, Z., & Chandranegara, D. R. (2024). Klasifikasi Kanker Kulit Menggunakan Metode Deep Learning. *REPOSITOR*, 6(1), 11–16. <https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign>
- Azis, H., Purnawansyah, P., Fattah, F., & Putri, I. P. (2020). Performa Klasifikasi K-NN dan Cross Validation Pada Data Pasien Pengidap Penyakit Jantung. *ILKOM Jurnal Ilmiah*, 12(2), 81–86. <https://doi.org/10.33096/ilkom.v12i2.507.81-86>
- Bansal, P., Garg, R., & Soni, P. (2022). Detection of melanoma in dermoscopic images by integrating features extracted using handcrafted and deep learning models. *Computers and Industrial Engineering*, 168(March), 108060. <https://doi.org/10.1016/j.cie.2022.108060>
- Bird, J. J., Barnes, C. M., Manso, L. J., Ekárt, A., & Faria, D. R. (2022). Fruit quality and defect image classification with conditional GAN data augmentation. *Scientia Horticulturae*, 293. <https://doi.org/10.1016/j.scienta.2021.110684>
- Bisla, D., Choromanska, A., Berman, R. S., Stein, J. A., & Polsky, D. (2019). Towards automated melanoma detection with deep learning: Data purification and augmentation. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2019-June, 2720–2728. <https://doi.org/10.1109/CVPRW.2019.00330>

- Chu, Y., Yue, X., Yu, L., Sergei, M., & Wang, Z. (2020). Automatic Image Captioning Based on ResNet50 and LSTM with Soft Attention. *Wireless Communications and Mobile Computing*, 2020. <https://doi.org/10.1155/2020/8909458>
- Cossio, M. (2023). *Augmenting Medical Imaging: A Comprehensive Catalogue of 65 Techniques for Enhanced Data Analysis*. <http://arxiv.org/abs/2303.01178>
- Diantika, S., Nalatissifa, H., Maulidah, N., Supriyadi, R., & Fauzi, A. (2024). Penerapan Teknik Random Oversampling Untuk Memprediksi Ketepatan Waktu Lulus Menggunakan Algoritma Random Forest. In *Computer Science (CO-SCIENCE)* (Vol. 4, Issue 1). <https://www.kaggle.com/>
- Ekoputra, G. (2019). Penghapusan Noise Pada Citra Kamera CCTV Menggunakan Metode Gaussian Filter. In *Prosiding Annual Research Seminar* (Vol. 5, Issue 1).
- Erdiansyah, U., Irmansyah Lubis, A., & Erwansyah, K. (2022). Komparasi Metode K-Nearest Neighbor dan Random Forest Dalam Prediksi Akurasi Klasifikasi Pengobatan Penyakit Kulit. *Jurnal Media Informatika Budidarma*, 6(1), 208. <https://doi.org/10.30865/mib.v6i1.3373>
- Firmansyah, W. A., Hayati, U., & Wijaya, Y. A. (2023). ANALISA TERJADINYA OVERFITTING DAN UNDERFITTING PADA ALGORITMA NAIVE BAYES DAN DECISION TREE DENGAN TEKNIK CROSS VALIDATION. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 7, Issue 1).

- Ghosh, P., Mustafi, S., & Mandal, S. N. (2020). Image-Based Goat Breed Identification and Localization Using Deep Learning. *International Journal of Computer Vision and Image Processing*, 10(4), 74–96. <https://doi.org/10.4018/ijcvip.2020100105>
- Grasela Tatuin, M., Kelen, Y. P. K., Manek, S. S., Informatika, T., Sains, P., Kesehatan, D., & Timor, U. (2023). *Pengaruh Ukuran Jendela Ketetangaan (Window) Terhadap Hasil Reduksi Noise pada Metode Median Filter dan Gaussian Filter* (Vol. 3). <https://ejournal.sidyanusa.org/index.php/jkdn>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- Indraswari, R., Rokhana, R., & Herulambang, W. (2021). Melanoma image classification based on MobileNetV2 network. *Procedia Computer Science*, 197, 198–207. <https://doi.org/10.1016/j.procs.2021.12.132>
- Isa, I. G. T., & Junedi, B. (2022). Hyperparameter Tuning Epoch dalam Meningkatkan Akurasi Data Latih dan Data Validasi pada Citra Pengendara. *Prosiding Sains Nasional Dan Teknologi*, 12(1), 231. <https://doi.org/10.36499/psnst.v12i1.6697>
- Jojoa Acosta, M. F., Caballero Tovar, L. Y., Garcia-Zapirain, M. B., & Percybrooks, W. S. (2021). Melanoma diagnosis using deep learning techniques on dermatoscopic images. *BMC Medical Imaging*, 21(1). <https://doi.org/10.1186/s12880-020-00534-8>

- Khasanah, N., Komarudin, R., Afni, N., Maulana, Y. I., & Salim, A. (2021a). Skin Cancer Classification Using Random Forest Algorithm. *Sisfotenika*, 11(2), 137. <https://doi.org/10.30700/jst.v11i2.1122>
- Khasanah, N., Komarudin, R., Afni, N., Maulana, Y. I., & Salim, A. (2021b). Skin Cancer Classification Using Random Forest Algorithm. *SISFOTENIKA*, 11(2), 137. <https://doi.org/10.30700/jst.v11i2.1122>
- Khoiruzzaman, N., Ramadhani, R. D., & Junaidi, A. (2021). Hasil Klasifikasi Algoritma Backpropagation dan K-Nearest Neighbor pada Cardiovascular Disease. *Journal of Dinda : Data Science, Information Technology, and Data Analytics*, 1(1), 17–27. <https://doi.org/10.20895/dinda.v1i1.141>
- Kim, S. C., & Cho, Y. S. (2022). Predictive System Implementation to Improve the Accuracy of Urine Self-Diagnosis with Smartphones: Application of a Confusion Matrix-Based Learning Model through RGB Semiquantitative Analysis. *Sensors*, 22(14). <https://doi.org/10.3390/s22145445>
- Kusumawardani, R., & Karningsih, P. D. (2021). Detection and Classification of Canned Packaging Defects Using Convolutional Neural Network. *PROZIMA (Productivity, Optimization and Manufacturing System Engineering)*, 4(1), 1–11. <https://doi.org/10.21070/prozima.v4i1.1280>
- Kutlu, H., Avcı, E., & Özyurt, F. (2020). White blood cells detection and classification based on regional convolutional neural networks. *Medical Hypotheses*, 135, 109472. <https://doi.org/10.1016/j.mehy.2019.109472>
- López de la Rosa, F., Gómez-Sirvent, J. L., Sánchez-Reolid, R., Morales, R., & Fernández-Caballero, A. (2022). Geometric transformation-based data

- augmentation on defect classification of segmented images of semiconductor materials using a ResNet50 convolutional neural network. *Expert Systems with Applications*, 206(January), 117731. <https://doi.org/10.1016/j.eswa.2022.117731>
- Mandal, S. N., Ghosh, P., Mukherjee, K., Dan, S., Mustafi, S., Roy, K., Hajra, D. K., & Banik, S. (2020). InceptGI: a ConvNet-Based Classification Model for Identifying Goat Breeds in India. *Journal of The Institution of Engineers (India): Series - B*, 101(5), 573–584. <https://doi.org/10.1007/s40031-020-00471-8>
- Milton, M. A. A. (2019). *Automated Skin Lesion Classification Using Ensemble of Deep Neural Networks in ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection Challenge*.
- Mohammed, R., Rawashdeh, J., & Abdullah, M. (2020). Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results. *2020 11th International Conference on Information and Communication Systems, ICICS 2020*, 243–248. <https://doi.org/10.1109/ICICS49469.2020.239556>
- Naeem, A., Farooq, M. S., Khelifi, A., & Abid, A. (2020). Malignant Melanoma Classification Using Deep Learning: Datasets, Performance Measurements, Challenges and Opportunities. *IEEE Access*, 8, 110575–110597. <https://doi.org/10.1109/ACCESS.2020.3001507>
- Nancy Jane, Y., Charanya, S. K., Amsaprabhaa, M., Jayashanker, P., & Nehemiah H., K. (2023). 2-HDCNN: A two-tier hybrid dual convolution neural network

- feature fusion approach for diagnosing malignant melanoma. *Computers in Biology and Medicine*, 152(March 2022), 106333. <https://doi.org/10.1016/j.combiomed.2022.106333>
- Nurfaizah, Hariguna, T., & Romadon, Y. I. (2019). The accuracy comparison of vector support machine and decision tree methods in sentiment analysis. *Journal of Physics: Conference Series*, 1367(1). <https://doi.org/10.1088/1742-6596/1367/1/012025>
- Pandey, R., Purohit, H., Castillo, C., & Shalin, V. L. (2022). Modeling and mitigating human annotation errors to design efficient stream processing systems with human-in-the-loop machine learning. *International Journal of Human Computer Studies*, 160. <https://doi.org/10.1016/j.ijhcs.2022.102772>
- Pangestu, M. A., & Bunyamin, H. (2018). Analisis Performa dan Pengembangan Sistem Deteksi Ras Anjing pada Gambar dengan Menggunakan Pre-Trained CNN Model. *Jurnal Teknik Informatika Dan Sistem Informasi*, 4, 337–344.
- Perez, L., & Wang, J. (2017). *The Effectiveness of Data Augmentation in Image Classification using Deep Learning*.
- Rana, P., Sowmya, A., Meijering, E., & Song, Y. (2022). Data augmentation with improved regularisation and sampling for imbalanced blood cell image classification. *Scientific Reports*, 12(1), 1–13. <https://doi.org/10.1038/s41598-022-22882-x>
- Ransi, N., Surimi, L., Tenriawaru, A., & Saidi, L. O. (2020). *Analisis Sentimen Masyarakat Terhadap Toko Online Aplikasi*. 1–8.

- Saputro, R. R., Junaidi, A., & Saputra, W. A. (2022). Klasifikasi Penyakit Kanker Kulit Menggunakan Metode Convolutional Neural Network (Studi Kasus: Melanoma). *Journal of Dinda : Data Science, Information Technology, and Data Analytics*, 2(1), 52–57. <https://doi.org/10.20895/dinda.v2i1.349>
- Savera, T. R., Suryawan, W. H., & Setiawan, A. W. (2020). DETEKSI DINI KANKER KULIT MENGGUNAKAN K-NN DAN CONVOLUTIONAL NEURAL NETWORK. 7(2), 373–378. <https://doi.org/10.25126/jtiik.202072602>
- Setia Budi, E., Nofriyaldi Chan, A., Priscillia Alda, P., & Arif Fauzi Idris, M. (2024). RESOLUSI: Rekayasa Teknik Informatika dan Informasi Optimasi Model Machine Learning untuk Klasifikasi dan Prediksi Citra Menggunakan Algoritma Convolutional Neural Network. *Media Online*, 4(5), 509. <https://djournals.com/resolusi>
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1). <https://doi.org/10.1186/s40537-019-0197-0>
- Soegeng, M. K., Liliana, L., & Noertjahyana, A. (2021). Penerapan Convolutional Neural Network untuk Klasifikasi Kanker Kulit Melanoma pada Dataset Gambar Kulit. *Jurnal Infra*, 9(1), 47–51.
- Taesiri, M. R., Nguyen, G., Habchi, S., Bezemer, C.-P., & Nguyen, A. (2023). *ImageNet-Hard: The Hardest Images Remaining from a Study of the Power of Zoom and Spatial Biases in Image Classification*. <http://arxiv.org/abs/2304.05538>

- Tanti. (2023). *Random Oversampling, Chi-Square, dan AdaBoost dalam Penanganan Ketidakeimbangan Kelas pada Klasifikasi C5.0*.
<https://doi.org/10.30865/mib.v7i2.5862>
- Teresia R. Savera, Winsya H. Suryawan, A. W. S. (2020). Deteksi Dini Kanker Kulit Menggunakan K-NN dan Convolutional Neural Network. In *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)* (Vol. 7, Issue 2). WORLD SCIENTIFIC, <https://doi.org/10.25126/jtiik.2020702602>
- Ulfah Nur Oktaviana, Ricky Hendrawan, Alfian Dwi Khoirul Annas, & Galih Wasis Wicaksono. (2021). Klasifikasi Penyakit Padi berdasarkan Citra Daun Menggunakan Model Terlatih Resnet101. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(6), 1216–1222.
<https://doi.org/10.29207/resti.v5i6.3607>
- van Zyl, C., Ye, X., & Naidoo, R. (2024). Harnessing eXplainable artificial intelligence for feature selection in time series energy forecasting: A comparative analysis of Grad-CAM and SHAP. *Applied Energy*, 353.
<https://doi.org/10.1016/j.apenergy.2023.122079>
- Walawalkar, D., Shen, Z., Liu, Z., & Savvides, M. (2020). Attentive Cutmix: An Enhanced Data Augmentation Approach for Deep Learning Based Image Classification. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2020-May*, 3642–3646.
<https://doi.org/10.1109/ICASSP40776.2020.9053994>

- Wolf, T., Sanh, V., Chaumond, J., & Delangue, C. (2019). *TransferTransfo: A Transfer Learning Approach for Neural Network Based Conversational Agents. ii.*
- Wu, M., Yue, H., Wang, J., Huang, Y., Liu, M., Jiang, Y., Ke, C., & Zeng, C. (2020). Object detection based on RGC mask R-CNN. *IET Image Processing*, 14(8), 1502–1508. <https://doi.org/10.1049/iet-ipr.2019.0057>
- Yani, V. L., Aradea, A., & Mubarak, H. (2022). Optimasi Prakiraan Cuaca Menggunakan Metode Ensemble pada Naïve Bayes dan C4.5. *Jurnal Teknik Informatika Dan Sistem Informasi*, 8(3). <https://doi.org/10.28932/jutisi.v8i3.5455>
- Yunial, A. H. (2020). Analisa Perbandingan Algoritma Klasifikasi Support Vector Machine, Deccesion Tree Dan Naive Bayes. *Prosiding Seminar Nasional Informatika Dan Sistem Informasi*, 5(2).
- Yusuf, R., Rahman, A., Sthevanie, F., & Kosala, G. (2024). Deteksi Kanker Kulit Melanoma Menggunakan Derivative of Gausstan dan Convolutional Neural Network. *Jurnal Penelitian Informatika*, 2, 1–5. <https://doi.org/10.25124/logic.v2i1.7529>
- Zhang, Y., Gao, J., & Zhou, H. (2020). Breeds Classification with Deep Convolutional Neural Network. *ACM International Conference Proceeding Series*, 24, 145–151. <https://doi.org/10.1145/3383972.3383975>