Klasifikasi Penyakit Mata melalui Citra Fundus Retina Mata dengan Arsitektur *Convolutional Neural Network (CNN)*

Muhammad Mukhlashyn Az Zahrawi
Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Indonesia
mukhlashyn.azzahrawi@gmail.com

Dr. Annisa Darmawahyuni, S.Kom., M.Kom
Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Indonesia
riset.annisadarmawahyuni@gmail.com

Abstrak*—Penelitian ini membahas pengembangan sebuah model untuk mengklasifikasikan berbagai kondisi penyakit pada mata, seperti katarak, retinopati diabetik, dan glaukoma. Dalam penelitian ini, metode Deep Learning diterapkan pada citra fundus retina. Kumpulan data yang digunakan diperoleh dari platform Kaggle, terdiri dari total 4217 citra fundus retina. Tahap pra-pemrosesan mencakup penghapusan gambar yang tidak layak, penyesuaian ukuran gambar menjadi 224x224 piksel, serta augmentasi dengan rotasi horizontal, vertikal, dan rotasi sebesar 90 derajat. Data tersebut kemudian dipecah menjadi data yang digunakan untuk pelatihan (80%), validasi (10%), dan pengujian (10%). Dalam penelitian ini, arsitektur Convolutional Neural Network (CNN) yang diterapkan adalah ResNet-50, VGG-19, DenseNet-121, dan EfficientNet-B0, dengan semua model dilatih tanpa bobot yang telah dilatih sebelumnya. Proses pelatihan model dilakukan sebanyak 50 epoch menggunakan optimizer Adam dan fungsi kerugian CrossEntropyLoss. Setiap arsitektur yang diterapkan memberikan hasil performa dan tingkat akurasi yang berbeda. Model yang dilatih dalam penelitian ini mencapai akurasi tertinggi sebesar 91,25%. Berbeda dengan sebagian besar penelitian dahulu yang mengandalkan transfer learning, pendekatan pada penelitian ini dilakukan agar dapat memahami setiap model dengan lebih objektid secara komparatif. Penelitian ini bertujuan untuk memberikan referensi mengenai arsitektur Convolutional Neural Network (CNN) yang paling sesuai dalam mengklasifikasikan berbagai penyakit mata berdasarkan citra fundus. Selain itu, hasil penelitian ini berpotensi diintegrasikan ke dalam sistem skrining medis otomatis untuk mendukung deteksi dini di daerah dengan keterbatasan tenaga dokter spesialis mata.

***Kata kunci*—Klasifikasi Gambar, Penyakit Mata, Deep Learning, Convolutional Neural Network (CNN), ResNet-50, VGG-19, DenseNet-121, EfficientNet-B0.**

I. PENDAHULUAN

Gangguan penglihatan merupakan isu kesehatan global yang signifikan, dengan lebih dari 2,2 miliar individu terpengaruh di seluruh dunia. Dari jumlah tersebut, sekitar 1 miliar kasus dapat dicegah atau dikelola melalui diagnosis dan intervensi yang tepat waktu [1]. Gangguan mata seperti katarak, glaukoma, dan retinopati diabetik adalah contoh umum yang dapat menimbulkan kebutaan apabila tidak diobati dengan segera [2].

Umumnya, dokter spesialis mata mendiagnosis penyakit mata dengan cara menganalisis citra fundus retina. Meski metode ini cukup efektif, pelaksanaannya sering kali tergantung pada keterampilan klinisi serta ketersediaan fasilitas medis, yang merupakan tantangan besar di daerah dengan sumber daya terbatas [3].

Dengan kemajuan dalam teknologi kecerdasan buatan, khususnya di bidang *Deep Learning*, otomatisasi dan efisiensi dalam proses diagnosis medis berbasis citra telah meningkat. *CNN* atau *Convolutional Neural Network* telah menunjukkan efektivitasnya dalam sejumlah penelitian untuk klasifikasi citra medis [4], [5], [6]. Namun, performa *CNN* sangat tergantung pada struktur model, metode pelatihan, dan preprocessing data. Variasi klasifikasi dapat terjadi tergantung pada pemilihan model, jenis augmentasi yang diterapkan, dan apakah bobot awal disesuaikan dengan karakteristik dataset local [3], [7], [8].

Penelitian ini bertujuan untuk mengembangkan model klasifikasi citra fundus retina guna mendeteksi tiga jenis penyakit mata yaitu katarak, glaukoma, dan retinopati diabetik. Dataset yang digunakan diperoleh dari Kaggle, yang terdiri dari 4217 gambar retina. Proses pra-pemrosesan melibatkan penghapusan gambar yang tidak valid, penyesuaian ukuran gambar menjadi 224×224 piksel, dan augmentasi citra melalui rotasi horizontal dan vertikal serta rotasi 90 derajat.

Empat arsitektur CNN yang diuji dalam penelitian ini meliputi *ResNet-50, VGG-19, DenseNet-121,* dan *EfficientNet-B0*. Semua model dilatih dengan konfigurasi yang sama agar dapat dievaluasi secara komparatif dengan lebih objektif.

Berbeda dengan sebagian besar penelitian terdahulu yang mengandalkan *transfer learning* dari *dataset* umum seperti *ImageNet*, penelitian ini secara khusus melakukan pelatihan model dari nol (*training from scratch*). Pendekatan ini diambil supaya dapat memberikan evaluasi dan analisis yang lebih objektif terhadap kemampuan asli dari setiap arsitektur *CNN* dalam memahami dataset citra fundus retina mata.

Diharapkan penelitian ini dapat memberikan rekomendasi mengenai arsitektur *CNN* yang paling efektif untuk klasifikasi citra fundus retina yang pada akhirnya mendukung pengembangan sistem diagnosis dini penyakit mata yang lebih tepat dan efisien, khususnya di daerah yang mengalami kesulitan dengan tenaga medis.

II. KAJIAN PUSTAKA

Dalam beberapa tahun terakhir, berbagai studi telah menghasilkan metode untuk mengklasifikasikan penyakit mata yang berorientasi pada citra fundus retina dengan memanfaatkan teknik *Deep Learning*, khususnya *Convolutional Neural Network (CNN).*

Metode ini telah terbukti efektif dalam mengidentifikasi penyakit seperti katarak, retinopati diabetik, dan glaukoma, yang dapat menyebabkan kebutaan apabila tidak diobati sejak awal.

Model arsitektur *VGG-19* yang tidak menggunakan bobot yang telah dilatih dalam pelatihannya, telah diaplikasikan untuk klasifikasi citra fundus, dan penelitian menunjukkan bahwa model *CNN* yang dikembangkan dari nol mampu mencapai tingkat akurasi yang bersaing, terutama ketika dilakukan preprocessing secara optimal [4]. Selain itu, sistem klasifikasi sederhana berbasis *CNN* juga berhasil mengidentifikasi retinopati diabetik serta glaukoma, terutama dalam situasi terbatasnya sumber daya medis [3].

Sebuah versi modifikasi dari *ResNet-50* telah dirancang untuk klasifikasi retinopati diabetik melalui penyesuaian parameter dan struktur jaringan dengan karakteristik citra fundus, yang menghasilkan performa lebih baik dibandingkan pendekatan tradisional [2].

Berbagai penelitian juga menyoroti signifikansi teknik augmentasi data dan preprocessing dalam meningkatkan performa model. Augmentasi yang berbasis rotasi dilaporkan dapat menaikkan akurasi model hingga 10 persen [7], [9]. Di samping itu, augmentasi seperti rotasi *horizontal,* rotasi *vertikal*, dan rotasi 90 derajat terbukti meningkatkan kemampuan generalisasi model *CNN* terhadap variasi citra retina [8], [10].

Pemilihan arsitektur juga memegang peranan penting dalam proses klasifikasi. *Residual connection* yang terdapat pada *ResNet-50* efektif dalam mengatasi penurunan akurasi pada jaringan yang sangat dalam [11], sedangkan *DenseNet-121* memungkinkan untuk mengulangi fitur dari *layer* sebelumnya di seluruh jaringan [12]. *VGG19* memiliki desain yang sederhana namun dalam, hal ini menjadi salah satu keunggulannya [13], sementara *EfficientNet-B0* dirancang dengan pendekatan *compound scaling* yang mengedepankan efisiensi parameter [14].

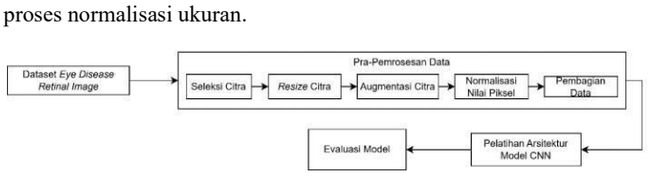
Citra fundus sebagai input untuk klasifikasi penyakit mata telah terbukti penting secara klinis karena memungkinkan visualisasi kondisi retina secara non-invasif. Citra-citra ini banyak digunakan dalam skrining terhadap penyakit kronis seperti glaukoma dan retinopati diabetic [15].

Dengan kata lain, keberhasilan klasifikasi penyakit mata berbasis *CNN* sangat dipengaruhi oleh kombinasi antara pilihan arsitektur, strategi pelatihan, serta teknik augmentasi dan preprocessing yang sesuai dengan karakteristik dataset di lapangan [10], [16].

III. METODOLOGI PENELITIAN

A. DATASET

Data yang dipakai dalam studi ini diambil dari platform Kaggle, dengan total 4. 217 gambar fundus retina yang dikategorikan ke dalam tiga jenis penyakit mata yakni katarak, glaukoma, dan retinopati diabetik, serta satu kategori untuk mata yang sehat. Karena setiap gambar memiliki resolusi yang berbeda-beda, maka dilakukan proses normalisasi ukuran.



Gambar 1. Diagram Skema Metodologi Penelitian

B. PRA-PEMROSESAN DATA

Tahapan pra-pemrosesan bertujuan untuk menjamin kualitas citra sebelum memasukkannya ke dalam model CNN. Proses awal mencakup seleksi gambar, yang melibatkan penghapusan citra yang tidak relevan, rusak, atau tidak jelas. Penelitian yang dilakukan oleh Abramoff dan rekan-rekannya pada tahun 2010 [15]) menunjukkan bahwa kualitas gambar berpengaruh besar terhadap akurasi sistem deteksi yang berbasis citra medis.

Setelahnya, setiap citra diubah ukurannya menjadi 224×224 piksel agar kompatibel dengan input layer dari arsitektur CNN yang digunakan. Penelitian oleh Marcella & Devella pada tahun 2022 [4] menunjukkan bahwa mengubah ukuran citra menjadi 224×224 efektif dalam mencocokkan input dengan arsitektur CNN tanpa mengakibatkan distorsi signifikan pada pola visual yang penting seperti makula atau diskus optikus.

Selanjutnya, dilakukan augmentasi data untuk meningkatkan variasi citra serta menurunkan risiko overfitting. Beberapa teknik augmentasi yang diterapkan mencakup rotasi horizontal, rotasi vertikal, dan rotasi sebesar 90 derajat, yang memungkinkan model belajar dari beragam sudut pandang. Penggunaan teknik rotasi ini dapat meningkatkan performa klasifikasi citra medis hingga 10 persen, sementara tetap mempertahankan konsistensi label [7], [9].

Setelah proses augmentasi dataset, nilai piksel *RGB* dinormalisasi berdasarkan rata-rata dan deviasi standar. Meskipun penelitian ini tidak memanfaatkan pretrained weights, metode ini tetap efektif untuk menstabilkan nilai input, mempercepat pelatihan, dan menjaga agar rentang data tetap konsisten[16], [17]. Normalisasi sangat penting untuk mencegah masalah exploding gradient dan mempercepat proses konvergensi jaringan saraf [17].

Akhirnya, dataset dibagi menjadi tiga bagian: 80 persen untuk data pelatihan, 10 persen untuk data validasi, dan 10 persen untuk data uji. Proses pemisahan ini dilakukan dengan tetap mempertahankan proporsi label agar distribusi antar kelas tetap seimbang. Metode ini sejalan dengan praktik umum di dalam penelitian klasifikasi citra medis [8].

C. ARSITEKTUR *CNN*

Dalam studi ini, dipilih empat arsitektur CNN untuk mengklasifikasikan citra fundus retina, yaitu *ResNet-50*, *VGG-19*, *DenseNet-121*, dan *EfficientNet-B0*. Setiap model dilatih dari dasar tanpa memanfaatkan berat yang telah dilatih sebelumnya agar sesuai dengan karakteristik dataset lokal.

- *ResNet-50*

ResNet atau *Residual Network*, diperkenalkan oleh He dan rekan-rekannya pada tahun 2016 [11]. Konsep pokok dari *ResNet* adalah residual learning, di mana masing-masing blok tidak hanya mempelajari fitur baru, tetapi juga residu dari fitur-fitur yang ada sebelumnya. Untuk mengatasi masalah vanishing gradient dan melatih jaringan yang sangat dalam dengan efektif, *ResNet-50* menggunakan pendekatan ini. Selain itu, model ini telah direvisi dan dioptimalkan untuk tujuan deteksi retinopati diabetik oleh Lin dan Wu pada tahun 2023 [2].

- *VGG-19*

VGGNet, yang diperkenalkan oleh Simonyan & Zisserman pada tahun 2014 [13], mengadopsi pendekatan yang sederhana tetapi efektif. Dengan menyusun konvolusi kecil dengan ukuran 3×3 secara berurutan, *VGGNet* mampu meningkatkan kedalaman dari jaringan. *VGG-19* memiliki struktur yang tidak rumit tetapi cukup dalam berkat penggunaan lapisan konvolusi yang berulang. Model ini telah diterapkan oleh Marcella & Devella (2022) [4] untuk klasifikasi penyakit pada mata yang berbasis pada citra fundus.

• *DenseNet-121*

DenseNet, atau *Densely Connected Convolutional Networks*, diperkenalkan oleh Huang dan rekan-rekannya pada tahun 2017 [12]. Konsep dasar dari *DenseNet* adalah hubungan antara setiap lapisan. Setiap lapisan dalam *DenseNet* mendapatkan input dari semua lapisan yang ada sebelumnya. *DenseNet-121* memanfaatkan koneksi yang rapat antar lapisan, yang membantu dalam penggunaan fitur dengan lebih baik serta mengurangi jumlah parameter yang tidak perlu, sehingga menjadikan kinerjanya sangat baik untuk mengklasifikasikan gambar medis, menurut Huang dan timnya pada tahun 2017 [12].

• *EfficientNet-B0*

EfficientNet, yang diciptakan oleh Tan dan Le pada tahun 2019 [14], memperkenalkan metode pengukuran yang komprehensif, yaitu sebuah sistem penskalaan jaringan yang seimbang dalam hal kedalaman, lebar, dan resolusi input. EfficientNet-B0 menggunakan metode penskalaan komprehensif ini untuk memaksimalkan ukuran jaringan, mengoptimalkan kedalaman, lebar, dan resolusi input, yang menghasilkan kinerja yang bagus dengan efisiensi komputasi yang lebih tinggi.

IV. PELATIHAN MODEL DAN EVALUASI

A. PELATIHAN MODEL

Keempat arsitektur *CNN* dilatih selama 50 epoch dengan menggunakan *optimizer Adam* dan fungsi *loss CrossEntropyLoss*. Dataset dibagi menjadi data latih, data uji, dan data validasi dengan perbandingan data latih sebanyak 80%, data uji sebanyak 10%, dan data validasi sebanyak 10%. Proses pelatihan dilakukan dari awal tanpa menggunakan bobot yang telah dilatih sebelumnya. Untuk mengevaluasi performa, akurasi klasifikasi dari setiap model diukur. Penelitian ini bertujuan untuk menemukan arsitektur CNN terbaik untuk mengklasifikasikan penyakit mata berdasarkan citra fundus, mengikuti penelitian sebelumnya seperti Tashkandi (2025) dan Babaqi serta rekan-rekan (2023) [5], [7], [9].

B. EVALUASI MODEL

Evaluasi model klasifikasi dilakukan untuk menilai kinerja dan ketepatan prediksi dari struktur CNN yang diaplikasikan. Penilaian ini memiliki peranan penting dalam klasifikasi gambar medis, di mana kesalahan dalam klasifikasi dapat mempengaruhi keputusan diagnosis. Oleh karena itu, studi ini menerapkan berbagai metrik penilaian yang sering digunakan dalam klasifikasi gambar medis yang berfokus pada deep learning.

1) CONFUSION MATRIX

Confusion matrix merupakan sebuah tabel yang berfungsi untuk menggambarkan kinerja algoritma pengklasifikasian. Tabel ini menunjukkan jumlah prediksi yang tepat dan tidak tepat berdasarkan label yang sebenarnya serta label yang diprediksi oleh model. *Confusion matrix* terdiri dari empat komponen utama, yaitu *True Positive (TP), True Negative (TN), False Positive (FP),* dan *False Negative (FN)* [17].

Pada situasi multi-kelas seperti klasifikasi untuk empat jenis penyakit mata (normal, katarak, glaukoma, dan retinopati diabetik), *confusion matrix* berbentuk tabel N×N, dengan N mewakili jumlah kelas yang ada. *Confusion matrix* juga berfungsi sebagai dasar untuk menghitung metrik evaluasi lainnya, seperti akurasi, presisi, dan sensitivitas.

2) METRIK EVALUASI

Berdasarkan data pada *confusion matrix*, metrik evaluasi berikut digunakan untuk menilai performa model:

• Akurasi

Akurasi berfungsi untuk mengukur proporsi prediksi yang benar terhadap keseluruhan jumlah prediksi. Akurasi adalah metrik yang paling umum digunakan untuk menilai kinerja model klasifikasi secara keseluruhan [17] .

$$Aku = \frac{T+T}{T+T+T+T} \quad (1)$$

dimana; TP adalah jumlah data positif yang diklasifikasikan sebagai benar, TN adalah jumlah data negatif yang diklasifikasikan sebagai salah, FP adalah jumlah data negatif yang diklasifikasikan sebagai positif, dan FN adalah jumlah data positif yang diklasifikasikan sebagai negatif.

• Presisi

Presisi mengukur seberapa baik model dapat memprediksi kelas positif. Dari semua data yang dianggap positif, seberapa banyak yang benar-benar positif. Presisi sangat penting terutama dalam bidang data medis, untuk mengurangi kesalahan positif palsu. [7].

$$uu = \frac{T}{T+T} \quad (2)$$

• Sensitivitas

Sensitivitas mengukur seberapa baik model dalam mendeteksi seluruh data yang seharusnya diklasifikasikan sebagai positif. Sensitivitas penting untuk memastikan bahwa kasus penyakit tidak terlewat [14] .

$$unsnu = \frac{T}{T+T} \quad (3)$$

• Spesifisitas

Spesifisitas mengukur kemampuan model dalam mengidentifikasi data negatif secara benar. Dalam klasifikasi medis, spesifisitas penting untuk menghindari kesalahan diagnosis pada individu sehat [15] .

$$= \frac{T}{T+T} \quad (4)$$

• *F1-Score*

F1-score adalah rata-rata antara presisi dan sensitivitas, yang memberikan keseimbangan antara kedua metrik tersebut. *F1-score* berguna ketika distribusi kelas tidak seimbang atau saat kesalahan klasifikasi memiliki risiko tinggi [17] .

$$1 - Sk = 2 \times \frac{uu \times unsnu}{uu + unsnu} \quad (5)$$

Pada klasifikasi multi-kelas, evaluasi dilakukan menggunakan *macro average*, yaitu menghitung metrik untuk setiap kelas secara independen, lalu merata-ratakannya. Pendekatan ini adil untuk dataset yang seimbang dan memberikan penilaian yang menyeluruh terhadap performa model di seluruh kelas [12], [18], [19] .

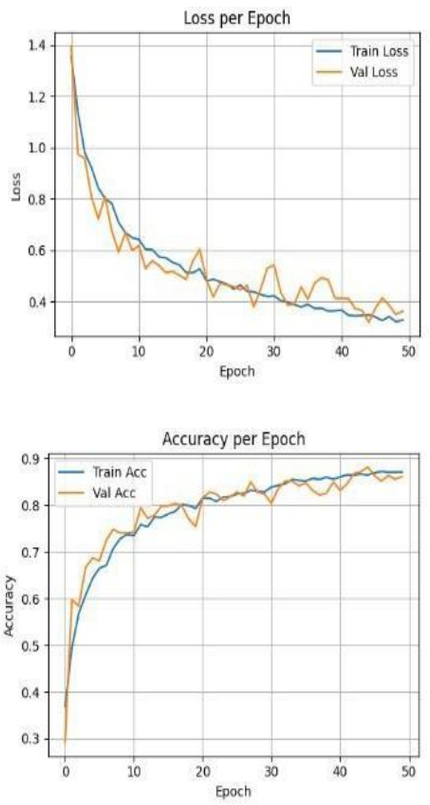
V. HASIL PENELITIAN

A. HASIL PELATIHAN MODEL

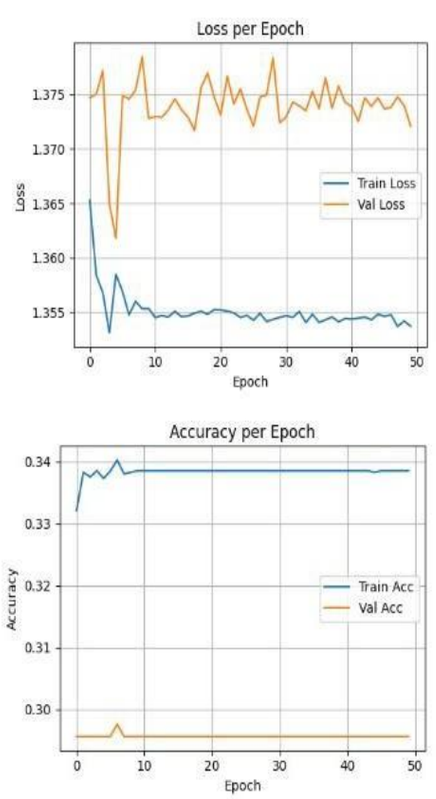
Model dilatih selama 50 *epoch* dengan dataset yang telah melalui proses pra-pemrosesan dan augmentasi. Arsitektur yang diuji pada penelitian ini meliputi *ResNet-50, VGG-19, DenseNet-121,* dan

EfficientNet-B0. Model dilatih dari awal tanpa menggunakan bobot yang sudah dilatih.

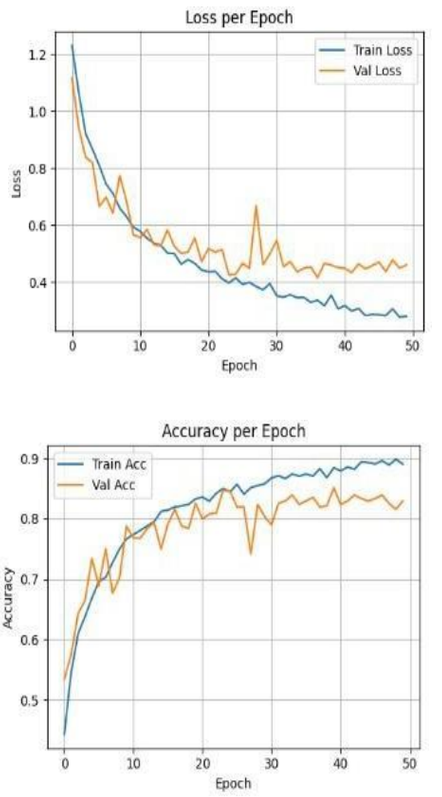
Berikut ini adalah plot hasil pelatihan yang menampilkan tren *loss* dan akurasi validasi selama proses *training*:



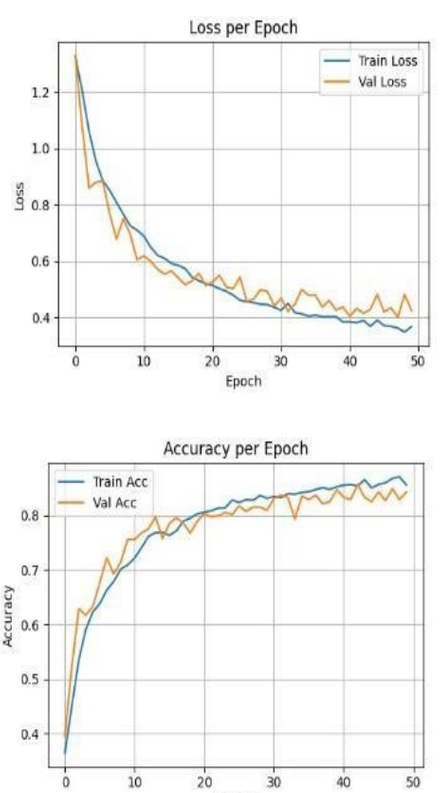
Gambar 2. *Training Plot ResNet-50*



Gambar 3. *Training Plot VGG-19*



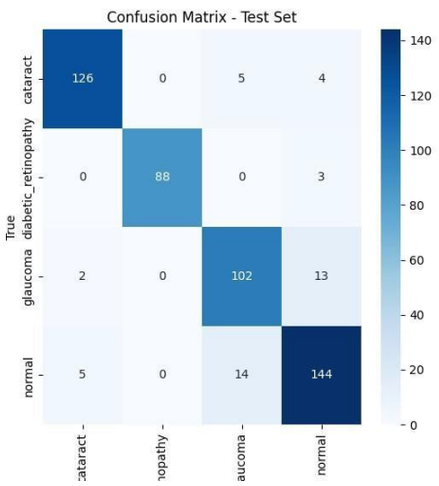
Gambar 4. *Training Plot DenseNet-121*



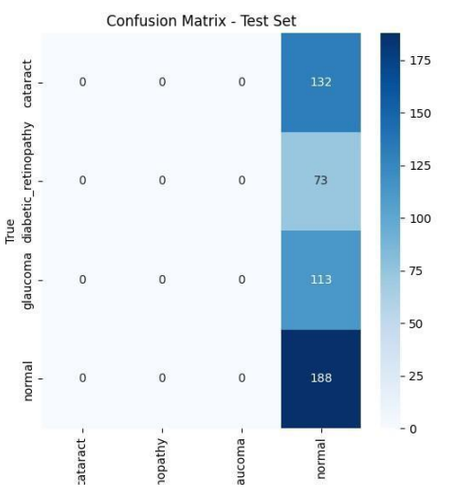
Gambar 5. *Training Plot EfficientNet-B0*

B. EVALUASI HASIL KLASIFIKASI

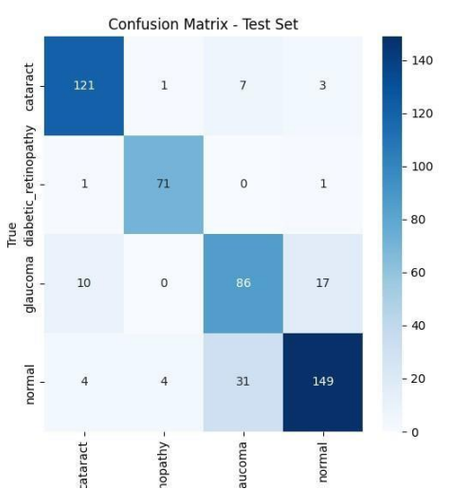
Berikut adalah *Confusion Matrix* yang didapat setelah melewati tahap pelatihan masing-masing arsitektur model:



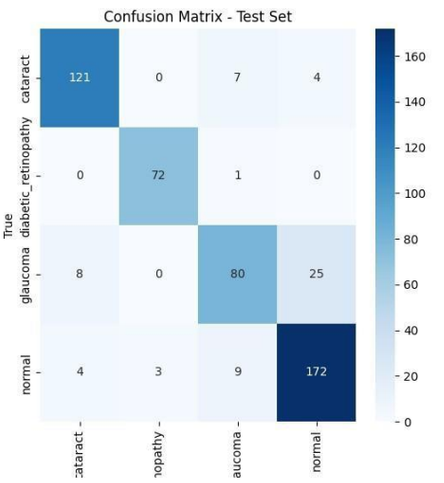
Gambar 6. *Confusion Matrix ResNet-50*



Gambar 7. *Confusion Matrix VGG-19*



Gambar 8. *Confusion Matrix DenseNet-121*



Gambar 9. *Confusion Matrix EfficientNet-B0*

Dari Gambar 6, 7, 8, dan 9, didapatkan hasil yang merepresentasikan nilai akurasi, presisi, sensitivitas, spesifisitas, dan *F1-Score* pada Tabel 1.

Model	Kelas	Akurasi	Presisi	Sensitivitas	Spesifisitas	*F1-Score*
ResNet-50	1	91,25	92,00	93,33	98,37	92,66
	2		100,00	96,70	99,36	98,32
	3		83,61	86,44	95,02	85,00
	4		85,21	87,80	91,62	86,49
VGG-19	1	37,15	0	0	73,94	0
	2		0	0	85,25	0
	3		0	0	83,18	0
	4		27,87	100	0	43,59
DenseNet-121	1	84,75	87,05	91,73	97,60	89,34
	2		93,42	97,26	99,38	95,00
	3		68,25	74,14	92,91	71,00
	4		86,13	78,84	97,99	82,30
EfficientNet-B0	1	89,90	89,63	92,37	97,99	90,98
	2		95,89	98,63	99,79	96,28
	3		82,47	72,07	95,19	76,96
	4		83,89	91,49	88,97	87,53

Tabel 1. Hasil Evaluasi Setiap Model

Dari tabel tersebut terlihat bahwa *ResNet-50* menghasilkan performa terbaik di seluruh metrik evaluasi, diikuti oleh *EfficientNet-B0* dan *DenseNet-121*. Sementara itu, *VGG-19* menunjukkan performa yang jauh lebih rendah. Perbedaan performa ini menunjukkan betapa pentingnya struktur arsitektur dalam mendukung kemampuan model mengenali pola visual kompleks pada citra retina.

ResNet-50 unggul karena koneksi residualnya memungkinkan propagasi gradien tetap stabil, sehingga pembelajaran tetap efektif pada jaringan yang dalam. Model ini mampu mempertahankan informasi spasial di setiap lapisan dan belajar fitur kompleks seperti pola pembuluh darah dengan baik.

EfficientNet-B0 menempati posisi kedua dengan akurasi tinggi dan efisiensi komputasi yang baik. Pendekatan *compound scaling* pada model ini membantu menyeimbangkan kedalaman dan lebar jaringan sehingga tetap efisien meskipun dijalankan pada perangkat keras terbatas.

DenseNet-121 menunjukkan performa solid dengan mekanisme *dense connectivity* yang memungkinkan pemanfaatan ulang fitur antar lapisan. Hal ini membuat model mampu mengekstraksi fitur secara efektif tanpa kehilangan fitur penting di lapisan dalam.

Sebaliknya, *VGG-19* mengalami penurunan performa yang signifikan serta indikasi *overfitting* ketika dilatih dari awal, menunjukkan bahwa model tidak tergenaralisasi dengan baik. Tanpa koneksi *residual* atau *dense*, *model* ini kesulitan mempertahankan fiturselama proses pembelajaran dan cenderung *overfitting* pada data latih.

Meskipun keempat model *CNN* dilatih dengan metode pelatihan yang sama. Performa masing-masing arsitektur menunjukkan perbedaan signifikan. Perbedaan ini terutama disebabkan oleh karakteristik dari masing-masing arsitektur, seperti dijelaskan berikut ini:

• Struktur Arsitektur Model

ResNet-50 mengimplementasikan koneksi residual supaya masalah *vanishing gradient* pada jaringan dalam sangat dapat dihindari. Walaupun desain ini memiliki kekuatan, melatih dari tahap awal tanpa bobot yang sudah ada membuat proses konvergensi berlangsung lebih lambat, dan kinerjanya sangat bergantung pada banyaknya data serta jumlah *epoch* yang cukup [11].

VGG-19 merupakan sebuah model yang memiliki struktur sekuensial yang mudah, dengan menggunakan lapisan konvolusi 3x3 secara berurutan. Model ini tidak memiliki *shortcut* atau koneksi padat, yang membuatnya kurang efisien dibandingkan dengan arsitektur terbaru dalam menjaga informasi spasial. Oleh karena itu, model ini dapat mengalami masalah *overfitting* atau *underfitting* ketika dilatih dari awal [11].

EfficientNet-B0 menerapkan metode pengukuran gabungan, yang secara terencana mengatur kedalaman, lebar, dan resolusi input. Metode ini telah terbukti memberikan kinerja yang unggul dengan parameter dan efisiensi komputasi yang lebih baik. Kinerja yang efisien ini membuat *EfficientNet* tetap berfungsi optimal meskipun dilatih tanpa bobot yang sudah dilatih sebelumnya [14]].

DenseNet-121 menerapkan hubungan erat antara lapisan, yang dikenal sebagai koneksi padat. Dalam model ini, setiap lapisan memperoleh input dari semua lapisan yang ada sebelumnya. Desain ini mendukung penggunaan ulang fitur dan mengatasi masalah penurunan akurasi yang terjadi pada jaringan yang lebih dalam, sehingga sangat cocok untuk mendeteksi pola visual yang rumit, seperti pada pembuluh darah di retina [12].

• Adaptabilitas *Dataset*

Citra fundus retina memiliki pola rumit yang halus dan tidak selalu terlihat jelas, contohnya variasi dalam ukuran pembuluh darah dan perubahan warna di makula. Model-model seperti *ResNet, EfficientNet,* dan *DenseNet* bisa menangani hal ini dengan menggabungkan fitur dari berbagai tingkat dan pemrosesan ruang yang lebih efektif [2], [9], [12], [20].

Di sisi lain, model yang lebih sederhana seperti *VGG-19* biasanya hanya dapat menangkap fitur yang paling jelas dan tidak berhasil mengenali pola kontekstual tanpa bantuan bobot yang telah dilatih sebelumnya sehingga arsitektur sederhana seperti *VGG-19* membutuhkan dataset dengan ukuran yang sangat besar untuk melakukan generalisasi model dengan baik, seperti yang dijelaskan oleh Marcella & Devella (2022) [4] dalam klasifikasi citra fundus.

• Pengaruh *Pretrained Weights*

Bobot yang sudah dilatih sebelumnya biasanya dipakai untuk memberikan bobot awal yang berasal dari dataset besar seperti *ImageNet*. Di dalam penelitian ini, setiap model dilatih dari nol. Situasi ini meningkatkan pentingnya arsitektur itu sendiri sebagai faktor utama dalam menentukan kesuksesan pelatihan awal [17], [21].

Model-model terbaru seperti *ResNet, DenseNet,* dan *EfficientNet* dapat belajar dengan efektif dari parameter yang tidak teratur karena desain mereka memang fokus pada efisiensi dalam belajar. Sebaliknya, arsitektur seperti *VGG* sangat memerlukan *transfer learning* agar bisa mendapatkan kinerja terbaik [13].

KESIMPULAN

Penelitian ini membuktikan bahwa arsitektur *CNN* modern seperti *ResNet-50, DenseNet-121,* dan *EfficientNet-B0* mampu melakukan klasifikasi citra fundus retina secara efektif meskipun dilatih dari awal tanpa *pretrained weights*. *ResNet-50* menunjukkan performa tertinggi dengan akurasi 91,25% dan keseimbangan metrik evaluasi terbaik, berkat mekanisme *residual learning* yang menjaga stabilitas gradien. *EfficientNet-B0* unggul dalam efisiensi komputasi melalui *compound scaling*, sedangkan *DenseNet-121* mempertahankan konsistensi pembelajaran melalui *dense connectivity*. Sebaliknya, *VGG-19* memiliki performa terendah akibat keterbatasan arsitektur konvensional tanpa koneksi antar-lapisan dan dataset yang kurang mumpuni untuk arsitektur sederhana seperti *VGG-19*.

Analisis *confusion matrix* menunjukkan bahwa *ResNet-50* dan *EfficientNet-B0* memiliki distribusi prediksi paling seimbang antar kelas, sementara *DenseNet-121* cenderung lebih sensitif terhadap kelas Retinopati Diabetik. Hasil ini menegaskan bahwa pemilihan arsitektur berperan krusial dalam klasifikasi multi-kelas citra medis.

Secara keseluruhan, penelitian ini memberikan kontribusi empiris terhadap pengembangan sistem skrining penyakit mata otomatis berbasis *Deep Learning*, yang berpotensi diimplementasikan di daerah dengan keterbatasan tenaga medis untuk mendukung diagnosis dini secara cepat dan akurat.

REFERENSI

[1] W. Health Organization, “World report on vision.”

[2] C. L. Lin dan K. C. Wu, “Development of revised ResNet-50 for diabetic retinopathy detection,” *BMC Bioinformatics*, vol. 24, no. 1, Des 2023, doi: 10.1186/s12859-023-05293-1.

[3] M. Nur *dkk.*, “KLASIFIKASI PENYAKIT MATA BERDASARKAN CITRA FUNDUS MENGGUNAKAN YOLO V8,” 2023.

[4] D. Marcella dan S. Devella, “Klasifikasi Penyakit Mata Menggunakan Convolutional Neural Network Dengan Arsitektur VGG-19,” vol. 3, no. 1, hlm. 60–70, 2022.

[5] A. Tashkandi, “Eye Care: Predicting Eye Diseases Using Deep Learning Based on Retinal Images,” *Computation*, vol. 13, no. 4, Apr 2025, doi: 10.3390/computation13040091.

[6] S. R. Alam dan others, “A comprehensive survey on deep learning in medical image analysis,” *Comput Biol Med*, vol. 143, 2022.

[7] T. Babaqi, M. Jaradat, A. E. Yildirim, S. H. Al-Nimer, dan D. Won, “Eye Disease Classification Using Deep Learning Techniques,” Jul 2023, [Daring]. Tersedia pada: http://arxiv.org/abs/2307.10501

[8] M. S. Qulub dan S. Agustin, “INDENTIFIKASI PENYAKIT MATA DENGAN KLASIFIKASI CITRA FOTO FUNDUS MENGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN),” 2024.

[9] T. Babaqi, S. Al-Nimer, M. Jaradat, A. E. Yildirim, S. H. AlNimer, dan D. Won, “Eye Disease Classification Using Deep Learning Techniques,” 2023, doi: 10.48550/arXiv.2307.10501.

[10] F. J. W. Ten Kate dan others, “Data augmentation for deep learning in histopathology: A review,” *Front Med (Lausanne)*, vol. 8, 2021.

[11] K. He, X. Zhang, S. Ren, dan J. Sun, “Deep Residual Learning for Image Recognition,” dalam *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun 2016, hlm. 770–778. doi: 10.1109/CVPR.2016.90.

[12] G. Huang, Z. Liu, L. Van Der Maaten, dan K. Q. Weinberger, “Densely Connected Convolutional Networks,” dalam *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul 2017, hlm. 2261–2269. doi: 10.1109/CVPR.2017.243.

[13] K. Simonyan dan A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” Sep 2014.

[14] M. Tan dan Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” Mei 2019.

[15] M. D. Abramoff, M. K. Garvin, dan M. Sonka, “Retinal imaging and image analysis,” 2010. doi: 10.1109/RBME.2010.2084567.

[16] A. K. Das dan others, “Deep learning architecture selection for automated DR grading,” *Artif Intell Med*, vol. 107, 2020.

[17] Y. LeCun, Y. Bengio, dan G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, hlm. 436–444, Mei 2015, doi: 10.1038/nature14539.

[18] R. R. Raghunath dan others, “Multi-class classification of retinal fundus images using CNN,” dalam *Procedia Computer Science*, 2020.

[19] Z. Wang dan others, “Ensembling convolutional neural networks for grading diabetic retinopathy,” *IEEE Access*, vol. 8, 2020.

[20] P. Rajalakshmi dan others, “Validation of smartphone based retinal photography for diabetic retinopathy screening,” *PLoS One*, vol. 10, no. 9, 2015.

[21] B. E. Bejnordi dan others, “Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in breast cancer,” *JAMA*, vol. 318, no. 22, 2017.