

# **Explainable AI: Perbandingan Model ResNet Dalam Image Classification Pada Retinopati Diabetik**

## **I. Pendahuluan**

Retinopati diabetik (RD) adalah salah satu komplikasi *mikrovaskular* yang muncul akibat diabetes *melitus*, terutama jika kadar gula darah tidak terkontrol dalam jangka waktu lama. Komplikasi ini menyerang mata, ditandai dengan kerusakan pada pembuluh darah di retina. Kadar glukosa yang tinggi secara terus-menerus dapat menghambat suplai nutrisi ke retina, yang kemudian memicu pertumbuhan pembuluh darah baru yang abnormal. Pembuluh darah ini rentan mengalami kebocoran, sehingga mengganggu penglihatan dan berpotensi menyebabkan kebutaan permanen [1]. RD menyebabkan gangguan penglihatan dan bahkan dapat menyebabkan kebutaan jika tidak dilakukan diagnosis pada tahap awal. Dengan demikian, pemeriksaan retina tetaplah menjadi hal penting untuk menangani penyakit RD, dan pemeriksaan tersebut penting untuk mengurangi konsekuensi jangka panjang dari penyakit tersebut. Untuk mengenali dan mengobati penyakit RD memanglah sulit, namun sangatlah penting untuk menghindari resiko kebutaan permanen yang lebih serius pada individu yang memiliki resiko [11]. Penelitian sebelumnya telah menunjukkan bahwa model *deep learning* memberikan akurasi yang lebih unggul ketika digunakan untuk melakukan deteksi gangguan penglihatan [7].

Penelitian ini bertujuan untuk membandingkan berbagai model ResNet dalam mengklasifikasikan citra fundus retina dan menemukan model yang menghasilkan kinerja terbaik. Model yang diterapkan dalam penelitian ini terdiri dari ResNet-18, ResNet-34, ResNet-50, ResNet-101, dan ResNet-152. ResNet dipilih sebagai model utama dalam penelitian ini karena strukturnya yang terbukti efektif dan efisien dalam klasifikasi citra, termasuk dalam mendeteksi retinopati diabetik dari gambar fundus retina. Selain memiliki kemampuan generalisasi yang baik, ResNet juga dapat dilatih tanpa perlu perubahan arsitektur yang rumit, namun tetap memberikan performa yang setara atau bahkan lebih baik dibandingkan model *deep learning* yang lebih kompleks. Model ResNet dalam klasifikasi citra fundus retina memiliki sejumlah keuntungan, termasuk kemampuan untuk mencapai kinerja tinggi meskipun dengan struktur yang sederhana, efisiensi dalam pengolahan, serta generalisasi yang baik terhadap data baru dari berbagai asal. Di samping itu, ResNet bisa dilatih tanpa perlu melakukan perubahan pada arsitektur tertentu, sehingga menjadikannya model yang stabil dan mudah untuk direplikasi. Hal ini menjadikan ResNet sebagai pilihan yang efisien sebagai *baseline* dalam pengembangan sistem otomatis untuk mendeteksi retinopati diabetik [4]. Pendekatan *Explainable AI (XAI)* juga diterapkan dalam penelitian ini untuk meningkatkan performa klasifikasi sekaligus mendukung interpretasi terhadap prediksi yang dihasilkan oleh model.

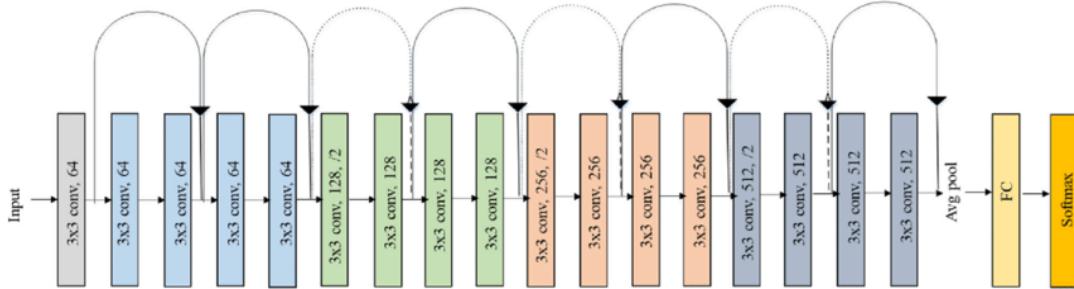
## **II. Tinjauan Pustaka**

### **2.1. Penelitian Terdahulu**

Pada penelitian [2] Menggunakan model ResNet-18 untuk mengidentifikasi retinopati diabetik yang menggunakan dataset dari Kaggle. Dari hasil penelitian tersebut mendapatkan *train accuracy* sebesar 99.91% dan *testing accuracy* sebesar 96.65%. Penelitian yang dilakukan oleh [23] mengusulkan *Disc-aware Ensemble Network* (DENet). Metode baru untuk melakukan skrining otomatis terhadap penyakit *glaukoma* dengan mendapatkan nilai AUC 0.9183 pada dataset SCES dan 0.8173 pada dataset SINDI. Penelitian yang dilakukan oleh [24] membuat suatu arsitektur deep learning bernama M-Net, yang mampu mencapai hasil segmentasi OD dan OC terbaik pada dataset ORIGA. Selain itu, metode ini juga menunjukkan kinerja yang memuaskan dalam skrining *glaukoma*, dengan perhitungan nilai CDR yang baik pada kedua dataset, yaitu ORIGA dan SCES. Dari penelitian yang dilakukan [25] mengimplementasikan arsitektur *deep learning* Inception-v3 dalam mendekripsi retinopati diabetik yang dapat dirujuk (RDR). Model tersebut menunjukkan performa yang sangat tinggi dengan sensitivitas mencapai 97.5% dan spesifisitas 93.4% pada titik operasi dengan prioritas sensitivitas tinggi pada dataset EyePACS-1. Dalam penelitian yang dilakukan [26] menggunakan *Convolutional Neural Network* (CNN) dengan 10 lapisan untuk secara otomatis dan simultan melakukan segmentasi serta klasifikasi terhadap eksudat, pendarahan, dan *mikroaneurisma*. Hasil sensitivitas sebesar 0.8758 untuk eksudat, sensitivitas 0.7158 untuk lesi gelap berdasarkan database CLEOPATRA, sensitivitas 0.257 untuk pendarahan, dan 0.4606 untuk *mikroaneurisma*.

### **2.2. ResNet-18**

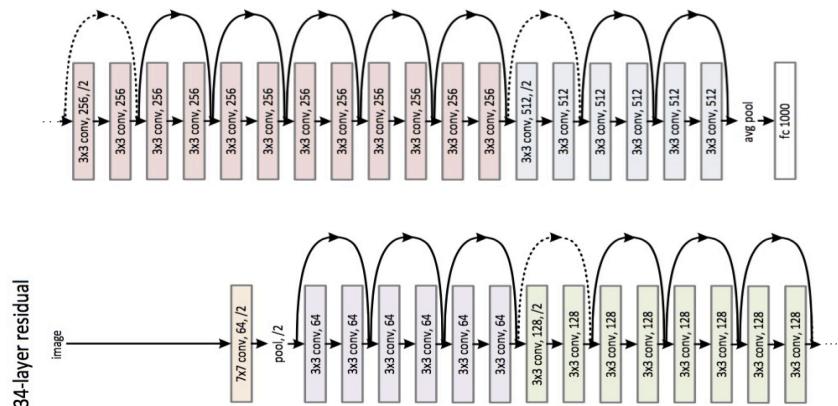
Arsitektur ResNet-18 memiliki ciri khas pada *residual block* dalam model. Pada gambar 1 inputan gambar akan melewati blok konvolusi dengan filter 64, 128, 256, 512. Dan masuk ke dalam *Average Pooling* menghitung rata-rata dari keseluruhan *pooling*, hal ini membantu untuk mencegah *overfitting*. Lapisan *Fully Connected Layer* menghubungkan *neuron* agar memberikan hasil prediksi dari model [17].



Gambar 1. Arsitektur ResNet-18

### 2.3. Resnet-34

ResNet-34 merupakan arsitektur dari model *deep learning* yaitu *Convolutional Neural Network*. Arsitektur ini mempunyai 34 lapisan neural, termasuk 32 lapisan dan 2 lapisan *fully connected*. ResNet-34 memulai arsitekturnya dengan lapisan konvolusi awal yang menggunakan 64 filter berukuran  $7 \times 7$  dan *stride* 2, yang kemudian dilanjutkan dengan serangkaian blok residual. Setiap blok ini terdiri atas dua lapisan konvolusi  $3 \times 3$  dan fungsi aktivasi ReLU, serta dilengkapi dengan *skip connection* untuk mengatasi masalah *vanishing gradient*. Untuk mencegah *overfitting*, digunakan teknik *dropout* pada lapisan *fully connected* dan dilakukan proses *tuning hyperparameter* untuk meningkatkan kinerja model. ResNet-34 dikenal sebagai salah satu model CNN yang berhasil dan sering dijadikan landasan dalam pengembangan arsitektur CNN yang lebih kompleks [18].

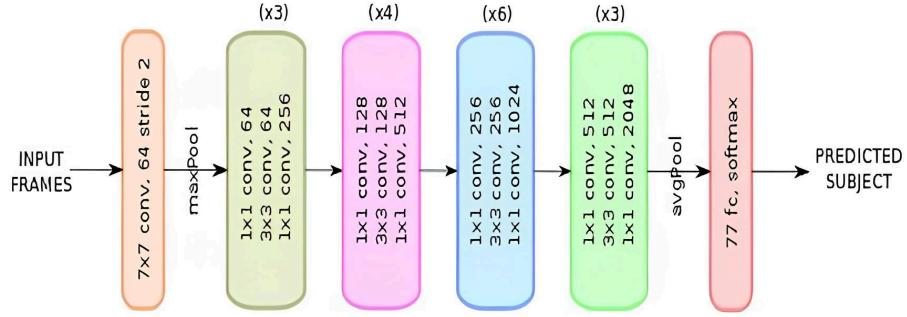


Gambar 2. Arsitektur ResNet-34

### 2.4. ResNet-50

ResNet-50 adalah salah satu arsitektur dari model CNN yang memiliki 50 layer dan pada ResNet-50 ini mempunyai *skip connection* sebanyak 3 layer dan terdapat  $1 \times 1$  convolution layer [19]. Struktur umum ResNet-50 diawali dengan

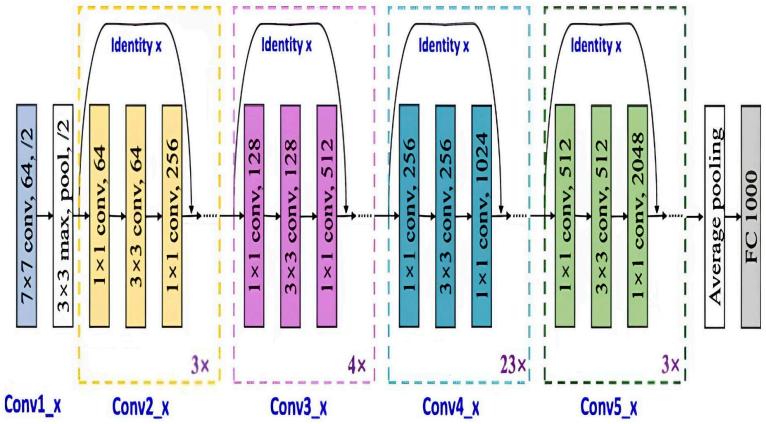
lapisan konvolusi dasar, kemudian dilanjutkan dengan sejumlah *residual block*. Setiap *residual block* terdiri dari beberapa lapisan konvolusi yang dilengkapi dengan *shortcut connection* untuk menjaga aliran informasi. Setelah seluruh blok residual dilalui, model menggunakan *global pooling layer* guna mengelokan dimensi spasial dari fitur yang telah diekstraksi. Selanjutnya, fitur tersebut diteruskan ke lapisan *fully connected* untuk menghasilkan output berupa hasil klasifikasi [20].



Gambar 3. Arsitektur ResNet-50

## 2.5. ResNet-101

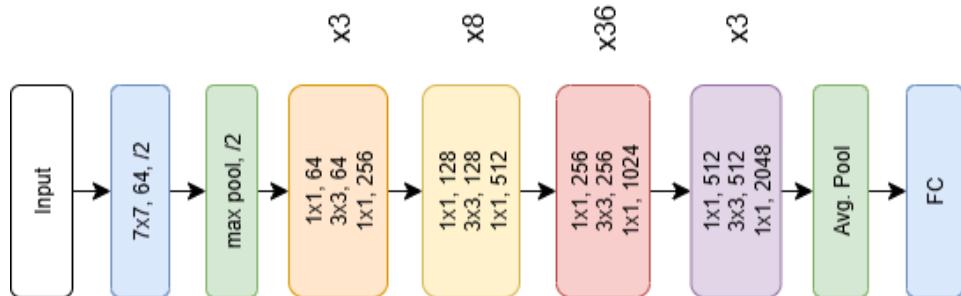
ResNet-101 merupakan salah satu arsitektur dari CNN dengan memiliki lapisan sebanyak 101 lapisan [19]. ResNet-101 memiliki struktur inti berupa *residual block* yang dilengkapi dengan *shortcut connection* sebagai komponen utamanya. Setiap blok residual terdiri atas tiga lapisan konvolusi berurutan, yaitu 1x1, 3x3, dan 1x1 yang dikenal sebagai pendekatan *Deeper Bottleneck Architecture*. Desain ini dirancang untuk meningkatkan efisiensi komputasi dengan mengurangi jumlah parameter, tanpa mengurangi kualitas dalam representasi fitur. Selain itu, ResNet-101 juga mengadopsi *identity shortcut* tanpa parameter tambahan (*parameter-free*), yang memungkinkan koneksi langsung antar layer tanpa beban komputasi ekstra, sehingga cocok digunakan dalam implementasi jaringan berskala besar [22].



Gambar 4. Arsitektur ResNet-101

## 2.6. ResNet-152

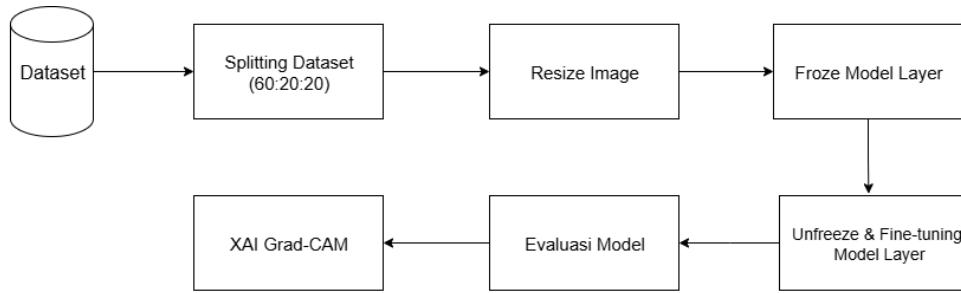
ResNet-152 merupakan salah satu varian dari arsitektur ResNet dengan kedalaman mencapai 152 lapisan dalam jaringan. Model ini dikenal karena keberhasilannya memenangkan kompetisi ILSVRC tahun 2015, dengan tingkat kesalahan top lima hanya 3.6%, yang setara dengan kemampuan manusia dalam melakukan klasifikasi gambar [9]. Struktur dari ResNet-152 terdiri atas lima tahapan utama konvolusi yang mencakup sejumlah besar blok konvolusi, serta diakhiri dengan sebuah lapisan *fully connected* untuk menghasilkan output klasifikasi.



Gambar 5. Arsitektur ResNet-152

## III. Metode Penelitian

Dalam penelitian ini melakukan perbandingan model untuk melakukan klasifikasi mata dengan menggunakan dataset fundus retina. Arsitektur yang diusulkan mencakup berbagai tahapan, seperti akuisisi data, pemartisian data, *preprocessing* data, *fine-tuning*, *testing*, dan evaluasi. Performa model dinilai dalam matrik akurasi dan dibandingkan dengan model lainnya, dalam melakukan performa klasifikasi pada kelas label. Alur kerja dalam penelitian ini digambarkan pada Gambar 6 berikut.



Gambar 6. Metodologi penelitian

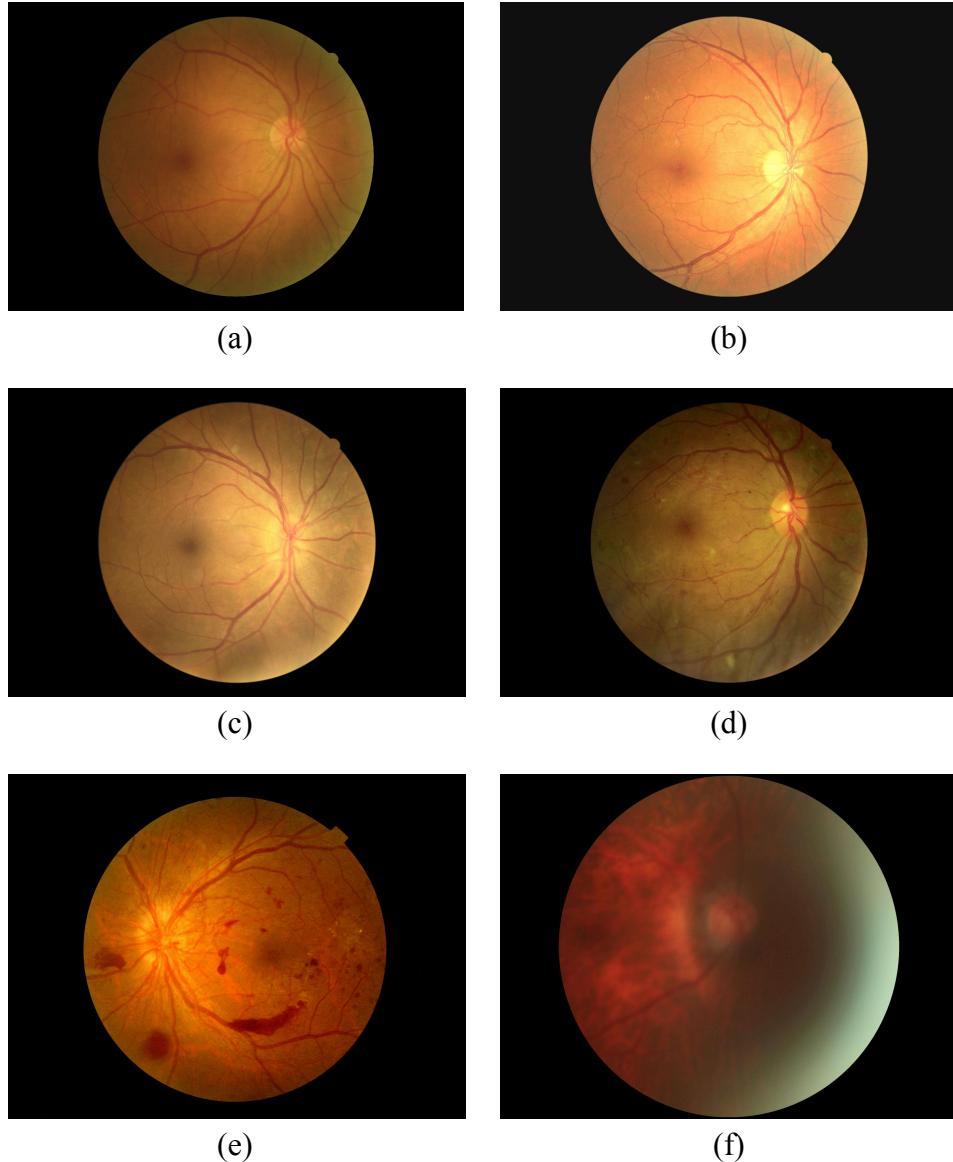
### 3.1 Dataset

Kumpulan data yang digunakan dalam penelitian ini dikumpulkan dari berbagai sumber yang memiliki reputasi baik, yaitu kumpulan citra retinopati diabetik (OIA-DDR) dari 147 rumah sakit, yang mencakup 23 provinsi di Tiongkok. Kumpulan data tersebut mencakup enam kelas berbeda yaitu tidak ada, RD ringan, sedang, berat, RD proliferatif, dan pada kategori ke enam adalah kumpulan data dengan gambar-gambar kualitas buruk yang menyebabkan gambar tidak dapat diklasifikasikan. Dataset berisi dengan total gambar sebanyak 13.673 gambar fundus retina dari 9.598 pasien [27]. Pada penelitian ini dimanfaatkan hanya 1.632 gambar fundus retina, dengan kelas label tidak ada penyakit RD sebanyak 439, ringan sebanyak 156, sedang sebanyak 387, berat sebanyak 30, proliferatif RD sebanyak 274, dan terakhir gambar dengan kualitas buruk sebanyak 346. Distribusi kelas label pada gambar fundus retina dapat diamati pada Gambar 2, sementara pada Tabel 1 memberikan rincian komprehensif tentang alokasi kumpulan dataset. Gambar 7 (a) menampilkan fundus retina normal, Gambar 7 (b) menunjukkan tanda-tanda penderita RD ringan, Gambar 7 (c) menunjukkan tanda-tanda penderita RD sedang, Gambar 2 (d) menunjukkan tanda-tanda penderita dengan RD berat, Gambar 7 (e) menunjukkan tanda-tanda penderita dengan RD proliferatif , dan terakhir pada Gambar 7 (f) menunjukkan gambar dengan kualitas buruk.

Nomor Kelas	Kelas	Jumlah Gambar
0	Retina normal	439
1	RD ringan	156
2	RD sedang	387
3	RD berat	30
4	RD proliferatif	274
5	Gambar kualitas buruk	346

<b>Total</b>	1.632
--------------	-------

Tabel 1. Distribusi kelas pada gambar fundus retina



Gambar 7. Retina fundus (a) normal, (b) RD ringan, (c) RD sedang, (d) RD berat, (e) RD proliferasif, dan (f) Gambar dengan kualitas buruk

Dataset kemudian dibagi menjadi tiga subset untuk menjawab dari tujuan penelitian. Dataset tersebut mengalami *splitting data* / pemisahan, dimana dialokasikan 60% data untuk set *training* / pelatihan, 20% untuk set *validation* / validasi, dan sisanya 20% untuk set *testing* / pengujian. Set pelatihan / *training* digunakan untuk tahapan *fine-tuning* pada model, memungkinkan melakukan konfigurasi dalam berbagai model untuk mengoptimalkan kinerja klasifikasi. Set validasi / *validation* digunakan untuk

menilai dan melakukan validasi model yang sudah dilatih, sementara pada set pengujian / *testing* memberikan evaluasi pada model secara independen untuk mengukur kinerja model akhir [7].

### 3.2 Preprocessing

Tahap *preprocessing* melibatkan transformasi data sebelum dapat digunakan untuk melakukan pelatihan model, dengan tujuan meningkatkan dataset dalam memperkuat kinerja model pada klasifikasi gambar. Dalam kumpulan data ini, yang terdiri dari total 1.632 gambar di enam kelas, beberapa langkah diambil untuk memastikan keseragaman dan kompatibilitas pada dataset. Gambar diubah menjadi dimensi 224x224 *pixel* untuk memastikan ukuran yang konsisten pada seluruh dataset. Selain itu, normalisasi ke format *ImageNet* dilakukan menggunakan nilai rata-rata [0.485, 0.456, 0.406] dan standar deviasi [0.299, 0.224, 0.225]. Perlu untuk dicatat bahwa tidak ada segmentasi yang dilakukan pada kumpulan data tersebut [7].

### 3.3 Matrik Evaluasi Model

Digunakan beberapa matrik evaluasi dalam penelitian ini, diantaranya seperti *accuracy*, *precision*, *recall*, *F1-Score*, *loss*. Semua matrik evaluasi memiliki fungsinya tersendiri seperti *accuracy* digunakan untuk mengukur kinerja model dalam melakukan klasifikasi data dengan benar, *precision* digunakan untuk menilai akurasi hasil positif yang dapat diprediksi oleh model, *recall* digunakan untuk menilai kelengkapan model dalam menemukan semua data positif, *F1-score* digunakan untuk memberikan keseimbangan antara *precision* dan *recall*, dan terakhir *loss* digunakan untuk memberi tahu seberapa jauh prediksi model dari label yang benar [9]. Rumus dari matrik evaluasi yang dipakai dalam penelitian ini, dapat dilihat pada Rumus 1, Rumus 2, Rumus 3, Rumus 4, dan Rumus 5.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Rumus 1. *Accuracy*

$$Precision = \frac{TP}{TP+FP}$$

Rumus 2. *Precision*

$$Recall = \frac{TP}{TP+FN}$$

Rumus 3. *Recall*

$$F1Score = 2 * \frac{Precision*Recall}{Precision+Recall}$$

Rumus 4. *F1-Score*

Persamaan dalam rumus-rumus 1 sampai 4 memiliki komponen-komponen sebagai berikut :

- *TP (True Positive)* adalah jumlah contoh positif yang diklasifikasikan dengan benar.
- *TN (True Negative)* adalah jumlah contoh negatif yang diklasifikasikan dengan benar.

- *FP (False Positive)* adalah jumlah contoh negatif yang salah diklasifikasikan sebagai positif.
- *FN (False Negative)* adalah jumlah contoh positif yang salah diklasifikasikan sebagai negatif.

$$loss(crossentropy) = \frac{-1}{N} \sum_1^N y_i * \log(\hat{y}_i) + (1 - y_i) * \log(1 - \hat{y}_i)$$

Rumus 5. Loss

Dalam Rumus 5 memiliki komponen sebagai berikut :

- $N$  adalah jumlah total data (sampel).
- $y_i$  adalah label asli dari data ke-i (0 atau 1).
- $\hat{y}_i$  adalah probabilitas prediksi oleh model bahwa data ke-i termasuk kelas 1.
- $\log$  adalah logaritma natural (basis e).

### 3.4 XAI Grad-CAM

Grad-CAM, atau *Gradient-weighted Class Activation Mapping* adalah teknik yang digunakan dalam *deep learning* untuk visualisasi wilayah penting dari gambar masukan yang berkontribusi pada prediksi kelas tertentu di CNN. Selama forward pass, dihitung gradien skor kelas yang diprediksi sehubungan dengan peta fitur yang menunjukkan sensitivitasnya. Gradien ini kemudian dibobot berdasarkan *global average pooling* (GAP) gradien dalam setiap peta fiturnya, yang menghasilkan jumlah bobot. Jumlah bobot yang dihasilkan dilewatkan melalui aktivitas ReLU untuk membuat peta aktivasi yang menyoroti wilayah berpengaruh dari gambar masukan untuk kelas yang diprediksi. Grad-CAM memberikan interpretabilitas yang berharga dengan menawarkan representasi visual dari area penting untuk melakukan proses pengambilan keputusan dengan model CNN, dalam membantu memahami dan melakukan validasi perilaku jaringan. Rumus Grad-CAM dapat dinyatakan sebagai Rumus 6 berikut [6].

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial Y_c}{\partial A_{ij}^k}$$

Rumus 6. Grad-CAM

Persamaan dalam Rumus 6 memiliki komponen-komponen sebagai berikut :

- $\alpha_k^c$  representasi dari pentingnya peta fitur k untuk kelas c.
- $Z$  adalah faktor normalisasi, seringkali merupakan elemen dalam peta fitur.
- $Y_c$  adalah skor kelas sebelum dilakukan aktivasi *softmax*.

- $A_{ij}^k$  merepresentasikan aktivitas peta fitur ke-k pada lokasi spasial  $(i,j)$ .

## IV. Hasil Pembahasan

### 4.1 Froze Model Layer

*Froze* model layer, digunakan dalam penelitian untuk memanfaatkan fitur-fitur umum yang telah dipelajari dari dataset besar seperti *ImageNet*, *edge detection*, deteksi pola, dan deteksi tekstur. Dengan membekukan layer awal atau seluruh bagian *feature extractor* pada model, proses pelatihan dapat menjadi lebih efisien dikarenakan hanya bagian akhir model (*classifier*) yang akan dilatih ulang pada tahap *unfreeze*. Pendekatan ini juga membantu untuk menghindari *overfitting*, terutama ketika dataset target berukuran besar agar dapat menghemat sumber daya komputasi serta waktu pelatihan / *training*. Model yang dipakai dalam penelitian ini adalah versi dari berbagai ResNet dimana diantaranya adalah ResNet-18, ResNet-34, ResNet-50, ResNet-101, dan ResNet-152. Pada Tabel 2 memberikan rincian dari proses *freeze* model layer pada penelitian ini, dapat dilihat bahwa hasil dari kelima model masih belum optimal ini dikarenakan tahap *freeze* model layer masih merupakan tahapan pertama dalam pelatihan / *training* dengan dataset. Total *epoch* yang digunakan dalam melakukan pelatihan / *training* dataset adalah 20, namun dalam penelitian ini digunakan fungsi *early stopping* dimana pelatihan dataset akan dihentikan jika nilai validasi akurasi tidak mengalami kenaikan pada 7 kali tahapan *epoch* kedepan.

Model	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Epoch
ResNet-18	1.1247	60.53%	1.0326	61.47%	9/20
ResNet-34	0.9536	67.28%	0.9187	65.14%	14/20
ResNet-50	0.8708	71.37%	0.8096	68.20%	18/20
ResNet-101	0.8678	70.14%	0.9103	66.36%	20/20
ResNet-152	0.8359	71.68%	0.8488	67.89%	19/20

Tabel 2. Hasil *freeze* model layer

### 4.2 Unfreeze & Fine-tuning Model Layer

Setelah melakukan tahapan *freeze* pada model layer, proses *unfreeze* dilakukan untuk memungkinkan pelatihan kembali / proses *fine-tuning* pada keseluruhan layer dari yang sebelumnya dilakukan *freeze layer* digunakan semua layer dalam model. Tujuannya untuk model tidak hanya mengandalkan fitur umum dari *pre-trained* tetapi juga dapat menyesuaikan bobotnya secara lebih spesifik terhadap karakteristik dari dataset yang

diberi. Dengan membuka layer-layer kembali, model dapat belajar dari fitur yang lebih relevan dan meningkatkan hasil akurasi pada tugas klasifikasi terutama jika dataset yang diolah cukup besar. Proses *unfreeze* dilakukan secara bertahap untuk menjaga stabilitas dan menghindari *overfitting* dalam model. Tabel 3, menggambarkan hasil *fine-tuning* dengan melakukan *unfreeze* model layer secara terperinci. Dimana tetap dilakukan perbandingan model, dengan cara melakukan *fine-tuning* pada model ResNet-18, ResNet-34, ResNet-50, ResNet-101, dan ResNet-152. Tahapan *epoch* ditentukan sama seperti proses *freeze* model layer sebelumnya yaitu 20, dan digunakan fungsi *early stopping* dimana saat proses pelatihan / *training* dataset tidak mengalami kenaikan pada validasi akurasi dalam 7 kali tahapan *epoch* setelah akurasi tertinggi maka tahapan pelatihan dataset dihentikan.

<b>Model</b>	<b>Train Loss</b>	<b>Train Accuracy</b>	<b>Validation Loss</b>	<b>Validation Accuracy</b>	<b>Epoch</b>
ResNet-18	0.2953	91.82%	0.5530	78.59%	20/20
ResNet-34	0.2422	92.74%	0.5388	78.29%	6/20
ResNet-50	0.2947	90.39%	0.5809	79.20%	19/20
ResNet-101	0.2283	92.02%	0.5876	76.45%	15/20
ResNet-152	0.1922	93.97%	0.6044	77.06%	10/20

Tabel 3. Hasil *unfreeze* model layer

### 4.3 Evaluasi Model

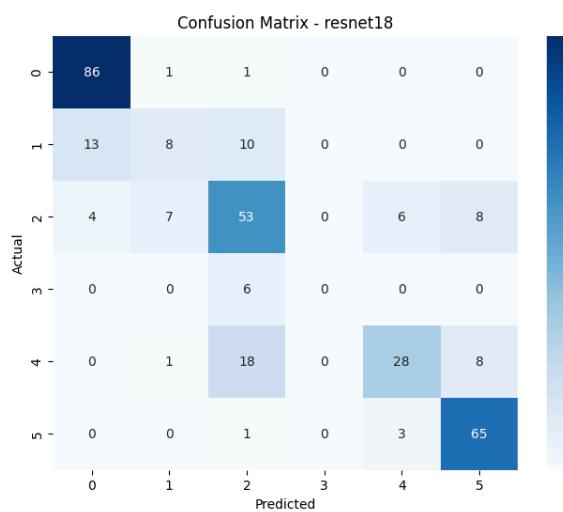
Tahap evaluasi menggunakan data *testing* dari *splitting data* yang sama di awal tahap penelitian. Dimana penilaian nya menggunakan *classification report*, dalam kelima kelas label yang ada yaitu 0 sebagai gambar dengan fundus mata normal, 1 sebagai tanda-tanda RD ringan, 2 sebagai tanda-tanda RD sedang, 3 sebagai tanda-tanda RD berat, 4 sebagai tanda-tanda RD proliferatif, dan terakhir 5 sebagai gambar dengan kualitas yang buruk. Pada *classification report* didapat 5 hasil berdasarkan 5 model yang digunakan. Berikut tabel dari kelima model :

<b>Model</b>	<b>Label</b>	<b>precision</b>	<b>recall</b>	<b>f1-score</b>	<b>support</b>
RestNet-18	0	0.8350	0.9773	0.9005	88
	1	0.4706	0.2581	0.3333	31
	2	0.5955	0.6795	0.6347	78

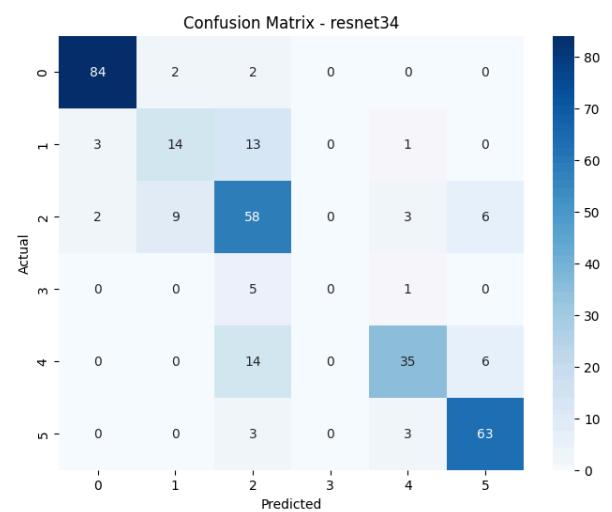
	3	0.0000	0.0000	0.0000	6
	4	0.7568	0.5091	0.6087	55
	5	0.8025	0.9420	0.8667	69
	<i>accuracy</i>			0.7339	327
	<i>macro avg</i>		0.5767	0.5610	0.5573
	<i>weighted avg</i>		0.7080	0.7339	0.7106
<hr/>					
RestNet-32	0	0.9438	0.9545	0.9492	88
	1	0.5600	0.4516	0.5000	31
	2	0.6105	0.7436	0.6705	78
	3	0.0000	0.0000	0.0000	6
	4	0.8140	0.6364	0.7143	55
	5	0.8400	0.9130	0.8750	69
	<i>accuracy</i>			0.7768	327
	<i>macro avg</i>		0.6281	0.6165	0.6182
	<i>weighted avg</i>		0.7669	0.7768	0.7675
<hr/>					
RestNet-50	0	0.8485	0.9545	0.8984	88
	1	0.3636	0.2581	0.3019	31
	2	0.5765	0.6282	0.6012	78
	3	0.0000	0.0000	0.0000	6
	4	0.6667	0.5455	0.6000	55
	5	0.8553	0.9420	0.8966	69
	<i>accuracy</i>			0.7217	327
	<i>macro avg</i>		0.5518	0.5547	0.5497
	<i>weighted avg</i>		0.6929	0.7217	0.7039
<hr/>					
RestNet-101	0	0.8557	0.9432	0.8973	88

	1	0.4000	0.2581	0.3137	31
	2	0.5595	0.6026	0.5802	78
	3	0.0000	0.0000	0.0000	6
	4	0.6667	0.6182	0.6415	55
	5	0.8400	0.9130	0.8750	69
	<i>accuracy</i>			0.7187	327
	<i>macro avg</i>	0.5536	0.5558	0.5513	327
	<i>weighted avg</i>	0.6910	0.7187	0.7022	327
<hr/>					
ResNet-152	0	0.8660	0.9545	0.9081	88
	1	0.5714	0.5161	0.5424	31
	2	0.5978	0.7051	0.6471	78
	3	0.0000	0.0000	0.0000	6
	4	0.7838	0.5273	0.6304	55
	5	0.8493	0.8986	0.8732	69
	<i>accuracy</i>			0.7523	327
	<i>macro avg</i>	0.6114	0.6003	0.6002	327
	<i>weighted avg</i>	0.7409	0.7523	0.7404	327

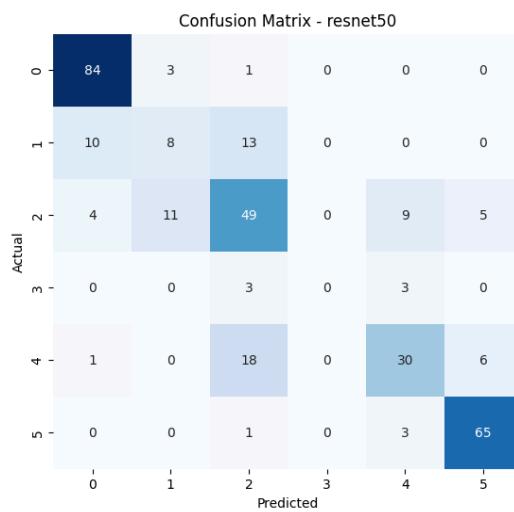
Tabel 4. Evaluasi Model ResNet-18, ResNet-34, ResNet-50, ResNet-101, ResNet-152



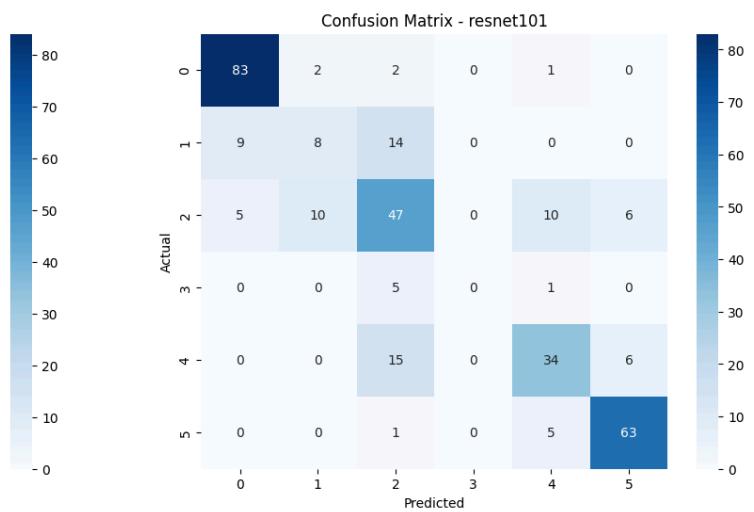
Gambar 8. *Confusion matrix ResNet-18*



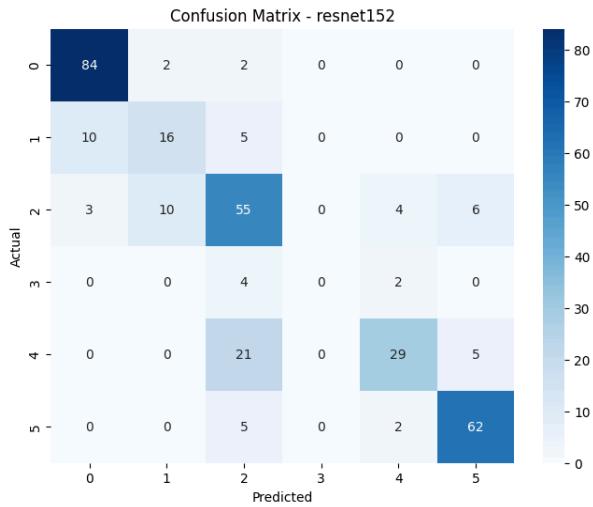
Gambar 9. *Confusion matrix ResNet-34*



Gambar 10. *Confusion matrix ResNet-50*



Gambar 11. *Confusion matrix ResNet-101*

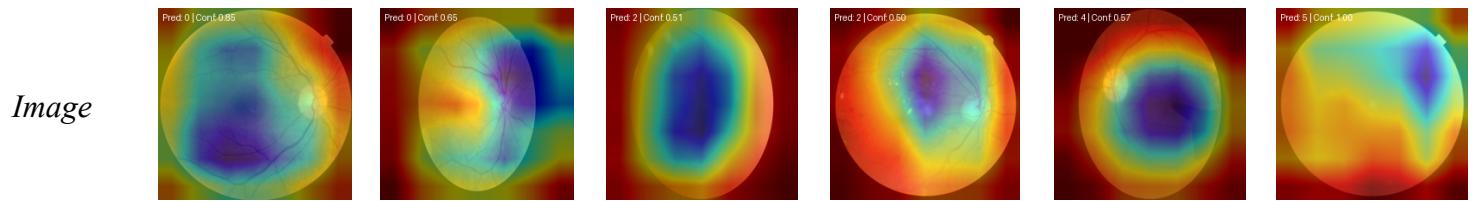


Gambar 12. *Confusion matrix ResNet-152*

#### 4.4 Explainable AI: Grad-CAM

*Explainable AI / XAI* pada model ResNet digunakan dalam penelitian ini untuk memberikan transparansi dan kejelasan dari hasil prediksi model *Fine-tuned* ResNet, terutama saat digunakan dalam bidang kritis seperti kesehatan. Ini juga dapat digunakan dalam melakukan analisis kesalahan dan pengambilan keputusan yang lebih akurat dari model ResNet. Penelitian ini memanfaatkan XAI untuk menjelaskan fokus pada bagian gambar fundus retina dari masing-masing model dengan hasil *fine-tuned* saat prediksi. Model yang digunakan dalam tahap tambahan XAI, yaitu model dari *frozen* dan *unfrozen* model layer. Diantaranya adalah model dari ResNet-18, ResNet-34, ResNet-50, ResNet-101, dan terakhir ResNet-152. Tabel 5, dijelaskan hasil dari XAI dari kelima model *fine-tuned* ResNet secara rinci pada tahapan awal *freeze model layer*. Lalu pada Tabel 6, dijelaskan hasil *unfreeze model layer* dimana tahapan ini melakukan *fine-tuned* secara keseluruhan pada model layer dengan cara melakukan *unfreeze* pada tahapan sebelumnya.

Model		ResNet-18 ( <i>Freeze model layer</i> )					
Label	0	1	2	3	4	5	
Prediksi	0	0	2	2	4	5	
Confidence	0.85	0.65	0.51	0.50	0.57	1.00	



**Model**

**ResNet-34 (*Freeze model layer*)**

<b>Label</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Prediksi	0	1	5	2	2	5
<i>Confidence</i>	0.99	0.39	0.98	0.67	0.48	1.00
<i>Image</i>						

**Model**

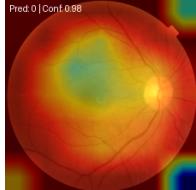
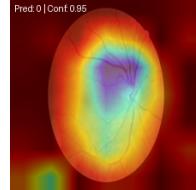
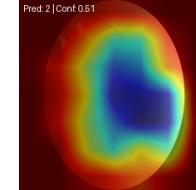
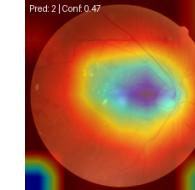
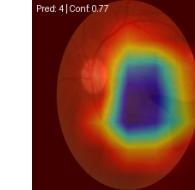
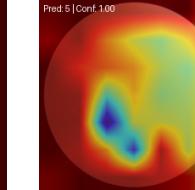
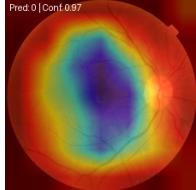
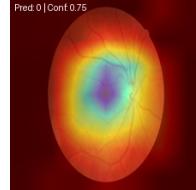
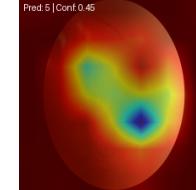
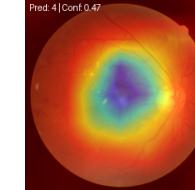
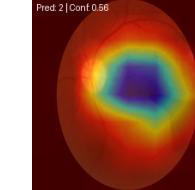
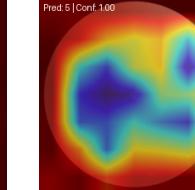
**ResNet-50 (*Freeze model layer*)**

<b>Label</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Prediksi	0	0	2	4	2	5
<i>Confidence</i>	0.90	0.57	0.65	0.69	0.38	1.00
<i>Image</i>						

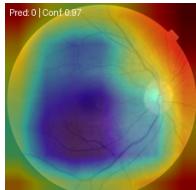
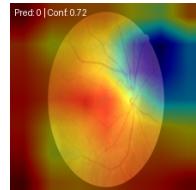
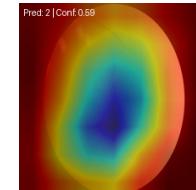
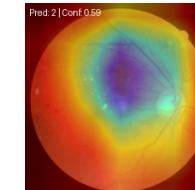
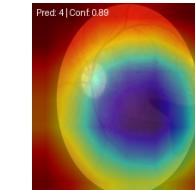
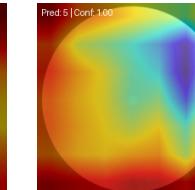
**Model**

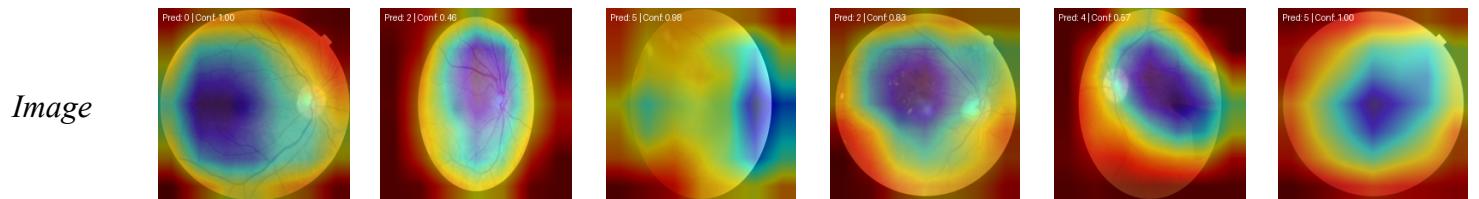
**ResNet-101 (*Freeze model layer*)**

<b>Label</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Prediksi	0	0	2	2	4	5
<i>Confidence</i>	0.98	0.95	0.51	0.47	0.77	1.00

<i>Image</i>						
<b>Model</b>	<b>ResNet-152 (<i>Freeze model layer</i>)</b>					
<b>Label</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Prediksi	0	0	5	4	2	5
<i>Confidence</i>	0.97	0.76	0.45	0.47	0.56	1.00
<i>Image</i>						

Tabel 5. Hasil XAI pada tahap *freeze model layer*

<i>Image</i>						
<b>Model</b>	<b>ResNet-18 (<i>Unfreeze model layer</i>)</b>					
<b>Label</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Prediksi	0	0	2	2	4	5
<i>Confidence</i>	0.97	0.72	0.59	0.59	0.89	1.00
<b>Model</b>	<b>ResNet-34 (<i>Unfreeze model layer</i>)</b>					
<b>Label</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Prediksi	0	2	5	2	4	5
<i>Confidence</i>	1.00	0.46	0.98	0.83	0.57	1.00

**Model****ResNet-50 (*Unfreeze model layer*)**

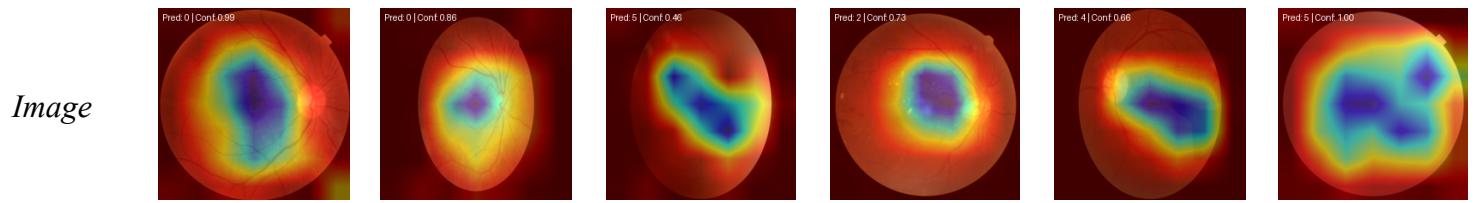
<b>Label</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Prediksi	0	0	2	4	4	5
<i>Confidence</i>	0.99	0.63	0.77	0.82	0.72	1.00

**Model****ResNet-101 (*Unfreeze model layer*)**

<b>Label</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Prediksi	0	0	2	4	4	5
<i>Confidence</i>	0.99	0.98	0.53	0.63	0.99	1.00

**Model****ResNet-152 (*Unfreeze model layer*)**

<b>Label</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>
Prediksi	0	0	5	2	4	5
<i>Confidence</i>	0.99	0.86	0.46	0.73	0.66	1.00



Tabel 6. Hasil XAI pada tahap *unfreeze model layer*

## V. Kesimpulan

Berdasarkan perbandingan model ResNet (ResNet-18, ResNet-34, ResNet-50, ResNet-101, dan ResNet-152) untuk klasifikasi citra fundus retina pada Retinopati Diabetik (RD), tahap *fine-tuning* secara signifikan meningkatkan kinerja model dibandingkan tahap *freeze*. Meskipun semua model menunjukkan peningkatan akurasi pelatihan dan validasi setelah *fine-tuning*, model ResNet-34 mencapai validasi akurasi tertinggi sebesar 77.68%, menunjukkan keseimbangan yang baik antara kinerja dan kompleksitas model untuk tugas klasifikasi ini. Namun perlu dicatat bahwa klasifikasi untuk kasus RD berat (Label 3) masih menjadi tantangan di semua model dengan *precision*, *recall*, dan *f1-score* sebesar 0.0000 yang mengindikasikan model tidak berhasil mengidentifikasi label ini; dilatarbelakangi oleh keterbatasan dataset. Pada Visualisasi Grad-CAM, seperti yang terlihat pada model *fine-tuned* ResNet-18 (*Confidence*: Label 3: 0.59), ResNet-34 (*Confidence*: Label 3: 0.83), ResNet-50 (*Confidence*: Label 3: 0.82), ResNet-101 (*Confidence*: Label 3: 0.63), dan ResNet-152 (*Confidence*: Label 3: 0.73), model dapat memberikan transparansi dan kejelasan dari hasil prediksi. Grad-CAM seharusnya memanfaatkan model dari ResNet untuk menyoroti wilayah fokus yang seharusnya, yaitu pada sekitar serabut saraf optik di retina atau dengan sebutan *optic disc*. Namun demikian, fokus *heatmap* justru menyebar ke sekitar *background* di semua label tanpa terkecuali; Dimana hal ini tidak selaras dengan nilai *classification report*; Berujung pada mayoritas prediksi label gambar yang salah.

## Daftar pustaka

- [1] I. M. J. R. L. M. M. Z. M. Y. P. Herdiana Nurul Utami, “Lombok Medical Journal,” *Crohn Dis. Pathophysiol. Diagnosis Manag.*, vol. 2, no. 2, pp. 1–4, 2023.
- [2] Ahmed, F., Fatima, A., Mamoon, M., & Khan, S. (2024). Identification of the Diabetic Retinopathy Using ResNet-18. 2nd International Conference on Cyber Resilience, ICCR 2024. <https://doi.org/10.1109/ICCR61006.2024.10532925>
- [3] Asia, A. O., Zhu, C. Z., Althubiti, S. A., Al-Alimi, D., Xiao, Y. L., Ouyang, P. B., & Al-Qaness, M. A. A. (2022). Detection of Diabetic Retinopathy in Retinal Fundus Images

Using CNN Classification Models. *Electronics* (Switzerland), 11(17).  
<https://doi.org/10.3390/electronics11172740>

- [4] Castilla, T., Martínez, M. S., Leguía, M., Larrabide, I., & Orlando, J. I. (2022). A ResNet is All You Need? Modeling A Strong Baseline for Detecting Referable Diabetic Retinopathy in Fundus Images. <http://arxiv.org/abs/2210.03180>
- [5] Doshi, D., Shenoy, A., Sidhpura, D., & Gharpure, P. (2017). Diabetic retinopathy detection using deep convolutional neural networks. International Conference on Computing, Analytics and Security Trends, CAST 2016, 261–266. <https://doi.org/10.1109/CAST.2016.7914977>
- [6] Faria, F. T. J., Moin, M. Bin, Debnath, P., Fahim, A. I., & Shah, F. M. (2024). Explainable Convolutional Neural Networks for Retinal Fundus Classification and Cutting-Edge Segmentation Models for Retinal Blood Vessels from Fundus Images. <http://arxiv.org/abs/2405.07338>
- [7] Imaduddin, H., Utomo, I. C., & Anggoro, D. A. (2024). Fine-tuning ResNet-50 for the classification of visual impairments from retinal fundus images. *International Journal of Electrical and Computer Engineering*, 14(4), 4175–4182. <https://doi.org/10.11591/ijece.v14i4.pp4175-4182>
- [8] Ishihara, K., & Matsumoto, K. (2023). Comparing the Robustness of ResNet, Swin-Transformer, and MLP-Mixer under Unique Distribution Shifts in Fundus Images. *Bioengineering*, 10(12). <https://doi.org/10.3390/bioengineering10121383>
- [9] Karthikayan, P. N., Varshan, Y. S., Kattamuri, H. G., & Jayaraman, U. (2024). Explainable AI: Comparative Analysis of Normal and Dilated ResNet Models for Fundus Disease Classification. <http://arxiv.org/abs/2407.05440>
- [10] Khojasteh, P., Passos Júnior, L. A., Carvalho, T., Rezende, E., Aliahmad, B., Papa, J. P., & Kumar, D. K. (2019). Exudate detection in fundus images using deeply-learnable features. *Computers in Biology and Medicine*, 104, 62–69. <https://doi.org/10.1016/j.combiomed.2018.10.031>
- [11] Lin, C. L., & Wu, K. C. (2023). Development of revised ResNet-50 for diabetic retinopathy detection. *BMC Bioinformatics*, 24(1). <https://doi.org/10.1186/s12859-023-05293-1>
- [12] Pan, Y., Liu, J., Cai, Y., Yang, X., Zhang, Z., Long, H., Zhao, K., Yu, X., Zeng, C., Duan, J., Xiao, P., Li, J., Cai, F., Yang, X., & Tan, Z. (2023). Fundus image classification using Inception V3 and ResNet-50 for the early diagnostics of fundus diseases. *Frontiers in Physiology*, 14. <https://doi.org/10.3389/fphys.2023.1126780>

- [13] Rajeshwari, A., & Sathishkumar, M. (2025). Diabetic Retinopathy Detection Using Deep Convolution Neural Network (pp. 544–558). [https://doi.org/10.2991/978-94-6463-754-0\\_48](https://doi.org/10.2991/978-94-6463-754-0_48)
- [14] Sharma, N., & Lalwani, P. (2025). A multi model deep net with an explainable AI based framework for diabetic retinopathy segmentation and classification. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-93376-9>
- [15] Thanki, R. (2023). A deep neural network and machine learning approach for retinal fundus image classification. *Healthcare Analytics*, 3. <https://doi.org/10.1016/j.health.2023.100140>
- [16] Xiao, Z., Zhang, X., Geng, L., Zhang, F., Wu, J., & Liu, Y. (2019). Research on the method of color fundus image optic cup segmentation based on deep learning. *Symmetry*, 11(7). <https://doi.org/10.3390/sym11070933>
- [17] G. Mediose, A. Sihotang, and J. Supardi, “Pengembangan Model CNN ResNet-18 untuk Klasifikasi Kondisi Gigi Berbasis Citra RGB sebagai Solusi Diagnostik Digital Development of CNN ResNet-18 Model for RGB Image-Based Dental Condition Classification as a Digital Diagnostic Solution,” vol. 4, no. 12, pp. 747–758, 2024.
- [18] I. Arsitektur, A. Dan, and R. Pada, “Implementasi arsitektur alexnet dan resnet 34 pada klasifikasi citra penyakit daun kentang menggunakan transfer learning,” vol. 7, no. 5, pp. 3293–3301, 2023.
- [19] A. Vierisyah, R. M. Fajri, P. Studi, S. Komputer, U. Indo, and G. Mandiri, “KLASIFIKASI KANKER PARU PARU MENGGUNAKAN CNN DENGAN 5 ARSITEKTUR,” no. 629, pp. 84–91.
- [20] J. Sistem and S. Komputer, “PENERAPAN CONVOLUTIONAL NEURAL NETWORK PADA KLASIFIKASI TANAMAN MENGGUNAKAN RESNET50,” vol. 8, no. 2, pp. 221–228, 2023.
- [21] A. Performa and R.-D. A. N. Alexnet, “Analisis performa resnet-152 dan alexnet dalam klasifikasi jenis kanker kulit,” vol. 8, no. 1, 2023.
- [22] A. Fathir, R. Januar, J. Indra, D. S. Kusumaningrum, and S. Faisal, “Application of Convolutional Neural Network ( CNN ) Algorithm with ResNet-101 Architecture for Monkey Pox Detection in Human,” vol. 9, no. 3, pp. 1006–1012, 2025.
- [23] H. Fu *et al.*, “Disc-aware Ensemble Network for Glaucoma Screening from Fundus Image,” vol. 0062, no. c, 2018, doi: 10.1109/TMI.2018.2837012.

- [24] H. Fu *et al.*, “Joint Optic Disc and Cup Segmentation Based on Multi-label Deep Network and Polar Transformation,” pp. 1–9.
- [25] V. Gulshan *et al.*, “Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs,” vol. 94043, pp. 1–9, 2016, doi: 10.1001/jama.2016.17216.
- [26] J. H. Tan, S. Sivaprasad, and S. V Bhandary, “Automated Segmentation of Exudates , Haemorrhages , Microaneurysms using Single Convolutional Neural Network,” no. August, 2017, doi: 10.1016/j.ins.2017.08.050.
- [27] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, and H. Kang, “Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening,” *Inf. Sci. (Ny).*, vol. 501, pp. 511–522, 2019, doi: 10.1016/j.ins.2019.06.011.