

KLASIFIKASI RETINOPATI DIABETIK DARI CITRA FUNDUS MENGUNAKAN CNN

Oktavia Ramadhani¹, Dwi Wahyu Prabowo², Nurahman³

^{1,2,3}Program Studi Sistem Informasi, Universitas Darwan Ali, Jl. Batu Berlian No.10, Kotawaringin Timur,
Kalimantan Tengah, Indonesia, 74323

***Email korespondensi:** oktaviarmdhni1310@gmail.com

Received: 27 05 2026 – Revised: 29 07 2026 - Accepted: 04 08 2026 - Published: 05 08 2026

Abstrak. Retinopati diabetik merupakan salah satu komplikasi diabetes yang dapat menyebabkan penurunan kemampuan penglihatan hingga kebutaan apabila tidak dikenali pada tahap awal. Pemeriksaan citra fundus retina banyak digunakan untuk mengidentifikasi kondisi tersebut karena mampu menampilkan perubahan pada retina secara visual. Namun, proses skrining masih menghadapi kendala akibat tingginya jumlah pasien dibandingkan ketersediaan tenaga medis yang melakukan evaluasi citra secara manual. Penelitian ini mengusulkan model klasifikasi tingkat keparahan retinopati diabetik berbasis transfer learning dengan memanfaatkan arsitektur EfficientNet-B3. Dataset yang digunakan adalah *Indian Diabetic Retinopathy Image Dataset* (IDRiD) yang terdiri atas lima kategori tingkat keparahan penyakit. Sebelum proses pelatihan dilakukan, citra melalui tahap praproses berupa penyesuaian ukuran, peningkatan kontras menggunakan *Contrast Limited Adaptive Histogram Equalization* (CLAHE), serta augmentasi data untuk memperkaya variasi sampel pelatihan. Kinerja model dievaluasi menggunakan metrik *AUC-ROC Macro*, *accuracy*, *Cohen's Kappa*, *precision*, *recall*, dan *F1-score*. Hasil eksperimen menunjukkan bahwa model memperoleh nilai *AUC-ROC Macro* sebesar 0,8186 dengan *accuracy* sebesar 55,07% dan *Cohen's Kappa* sebesar 0,6747. Evaluasi pada masing-masing kelas menunjukkan performa yang tidak merata, terutama pada kelas dengan jumlah data yang lebih sedikit. Untuk meningkatkan pemahaman terhadap proses pengambilan keputusan model, penelitian ini menerapkan *Grad-CAM* guna memvisualisasikan area retina yang berkontribusi terhadap hasil prediksi. Secara keseluruhan, hasil penelitian menunjukkan bahwa EfficientNet-B3 memiliki kemampuan yang cukup baik dalam mengidentifikasi tingkat keparahan retinopati diabetik serta berpotensi digunakan sebagai pendukung skrining otomatis.

Kata kunci: Retinopati Diabetik, *Deep Learning*, EfficientNet-B3, *Transfer Learning*, IDRiD, *Grad-CAM*.

PENDAHULUAN

Penyakit diabetes melitus merupakan salah satu penyakit tidak menular dengan jumlah penderita yang terus meningkat di berbagai negara. Berdasarkan data dari *Global Burden of Disease Study 2021*, jumlah kasus diabetes diperkirakan akan terus bertambah hingga tahun 2050 sehingga menjadi tantangan bagi sistem kesehatan global (Ong et al., 2023). Salah satu komplikasi yang sering terjadi adalah retinopati diabetik, yaitu kerusakan pembuluh darah retina yang dapat menyebabkan gangguan penglihatan hingga kebutaan permanen apabila tidak ditangani secara tepat. Oleh karena itu, retinopati diabetik menjadi

salah satu penyebab utama kebutaan yang dapat dicegah melalui deteksi dan penanganan dini (Kropp et al., 2023).

Deteksi retinopati diabetik umumnya dilakukan menggunakan citra fundus retina yang memungkinkan identifikasi lesi seperti mikroaneurisma, eksudat, dan perdarahan retina sebagai indikator tingkat keparahan penyakit (Zhang & Juan Chen, 2024). Namun, interpretasi citra secara manual memerlukan waktu dan keahlian dokter spesialis mata sehingga menjadi tantangan ketika jumlah pasien terus meningkat (Li et al., 2024).

Perkembangan teknologi *deep learning*, khususnya *Convolutional Neural Network* (CNN), membuka peluang pengembangan sistem diagnosis otomatis berbasis citra medis. Berbagai penelitian menunjukkan bahwa CNN mampu mengklasifikasikan retinopati diabetik dengan performa yang baik, terutama ketika dikombinasikan dengan *transfer learning* untuk memanfaatkan pengetahuan dari *dataset* berskala besar (Arora et al., 2024; Baba et al., 2024; Chavan & Choubey, 2023). Salah satu arsitektur yang banyak digunakan adalah *EfficientNet*, yang menerapkan metode *compound scaling* untuk menyeimbangkan kedalaman, lebar, dan resolusi jaringan sehingga menghasilkan performa yang baik dengan efisiensi parameter yang tinggi (Tan & Le, 2020). Berdasarkan keunggulan tersebut, penelitian ini menggunakan *EfficientNet-B3* karena mampu memberikan keseimbangan antara akurasi klasifikasi dan kebutuhan komputasi.

Penelitian ini bertujuan mengembangkan model klasifikasi tingkat keparahan retinopati diabetik menggunakan *EfficientNet-B3* berbasis *transfer learning* pada *dataset* IDRiD. Untuk meningkatkan kualitas citra digunakan metode *Contrast Limited Adaptive Histogram Equalization* (CLAHE), sedangkan permasalahan *class imbalance* ditangani menggunakan *WeightedRandomSampler* dan *class weight*. Selain itu, metode *Grad-CAM* diterapkan untuk meningkatkan interpretabilitas model. Kinerja model dievaluasi menggunakan metrik *AUC-ROC*, *accuracy*, *Cohen's Kappa*, *precision*, *recall*, dan *F1-score*.

MASALAH

Meskipun berbagai penelitian menunjukkan bahwa *deep learning* mampu memberikan performa yang baik dalam klasifikasi retinopati diabetik, masih terdapat beberapa kendala yang memengaruhi kinerja model pada data medis. Salah satu permasalahan utama adalah ketidakseimbangan distribusi data (*class imbalance*), di mana jumlah data pada kelas tertentu jauh lebih banyak dibandingkan kelas lainnya. Kondisi ini

dapat menyebabkan model lebih mudah mempelajari pola dari kelas mayoritas sehingga performa klasifikasi pada kelas minoritas menjadi menurun (Ahmed, 2025; Pamungkas et al., 2025).

Selain distribusi data, pemilihan arsitektur model juga berpengaruh terhadap hasil klasifikasi. Berbagai penelitian telah menerapkan kombinasi CNN dan *transfer learning* untuk mendeteksi retinopati diabetik dari citra fundus retina (Chavan & Choubey, 2023; Vij & Arora, 2023). Namun, sebagian penelitian masih berfokus pada peningkatan akurasi tanpa mempertimbangkan kemampuan model dalam mengklasifikasikan seluruh tingkat keparahan penyakit secara seimbang. Di sisi lain, efisiensi komputasi juga perlu diperhatikan agar model dapat diterapkan secara lebih efektif. Salah satu arsitektur yang memiliki potensi memenuhi kebutuhan tersebut adalah *EfficientNet-B3* (Arora et al., 2024).

Permasalahan lain adalah rendahnya interpretabilitas model *deep learning* yang umumnya bekerja sebagai *black box*. Dalam bidang kesehatan, kemampuan menjelaskan alasan di balik suatu prediksi sangat penting karena dapat memengaruhi pengambilan keputusan klinis. Oleh karena itu, diperlukan metode yang mampu menunjukkan bagian citra yang berkontribusi terhadap hasil klasifikasi. Salah satu pendekatan yang banyak digunakan adalah *Explainable Artificial Intelligence* (XAI) melalui *Gradient-weighted Class Activation Mapping* (*Grad-CAM*) (Ikram & Imran, 2025; Muntaha et al., 2024).

Berdasarkan telaah literatur, penelitian yang mengombinasikan *EfficientNet-B3* berbasis *transfer learning*, penanganan *class imbalance* menggunakan *WeightedRandomSampler* dan *class weight*, serta visualisasi *Grad-CAM* pada klasifikasi tingkat keparahan retinopati diabetik menggunakan *dataset* IDRiD masih relatif terbatas. Oleh karena itu, penelitian ini bertujuan mengembangkan model klasifikasi retinopati diabetik yang dievaluasi menggunakan *AUC-ROC*, *Cohen's Kappa*, *precision*, *recall*, dan *F1-score*, serta didukung visualisasi *Grad-CAM* untuk meningkatkan interpretabilitas hasil prediksi.

METODE PELAKSANAAN

Penelitian ini menggunakan pendekatan *deep learning* untuk melakukan klasifikasi tingkat keparahan retinopati diabetik berdasarkan citra fundus retina. Tahapan penelitian meliputi pengumpulan dataset, *preprocessing* citra, pembagian dataset, pelatihan model menggunakan *EfficientNet-B3* berbasis *transfer learning*, evaluasi performa model, serta interpretasi hasil prediksi menggunakan *Grad-CAM*.

A. Dataset Penelitian

Dataset yang digunakan dalam penelitian ini adalah Indian Diabetic Retinopathy Image Dataset (IDRiD) yang berisi citra fundus retina dengan label tingkat keparahan retinopati diabetik (Porwal et al., 2018). Dataset ini tersedia secara publik melalui platform *Grand Challenge* dan dapat diakses pada laman <https://idrid.grand-challenge.org>. Klasifikasi dilakukan ke dalam lima kategori, yaitu Grade 0 (*No Diabetic Retinopathy*), Grade 1 (*Mild*), Grade 2 (*Moderate*), Grade 3 (*Severe*), dan Grade 4 (*Proliferative Diabetic Retinopathy*). Dataset kemudian dibagi menggunakan metode *stratified split* menjadi data pelatihan sebesar 70%, data validasi sebesar 15%, dan data pengujian sebesar 15% untuk menjaga proporsi distribusi kelas pada setiap subset data.

B. Preprocessing Citra

Pada tahap praproses, seluruh citra fundus retina diubah ukurannya menjadi 300×300 piksel agar sesuai dengan masukan model EfficientNet-B3. Selanjutnya diterapkan metode CLAHE pada kanal hijau (*green channel*) untuk meningkatkan kontras citra dan memperjelas karakteristik pembuluh darah retina. Setelah itu dilakukan normalisasi nilai piksel dan augmentasi data menggunakan pustaka *Albumentations* untuk meningkatkan variasi data pelatihan serta mengurangi risiko *overfitting*.

C. Penanganan Ketidakseimbangan Data

Distribusi data pada setiap tingkat keparahan retinopati diabetik tidak seimbang, sehingga diperlukan strategi untuk mengurangi bias model terhadap kelas mayoritas. Pada penelitian ini digunakan dua pendekatan, yaitu *WeightedRandomSampler* dan *class weight*. *WeightedRandomSampler* digunakan untuk meningkatkan peluang pengambilan sampel dari kelas minoritas selama proses pelatihan, sedangkan *class weight* diterapkan pada fungsi kehilangan (*loss function*) sehingga kesalahan prediksi pada kelas minoritas memperoleh bobot yang lebih besar dibandingkan kelas mayoritas.

Metode CLAHE bekerja dengan mendistribusikan kembali intensitas piksel berdasarkan histogram lokal sehingga kontras citra dapat ditingkatkan. Probabilitas kemunculan setiap tingkat intensitas dapat dihitung menggunakan Persamaan (1).

$$P(i) = \frac{h(i)}{N} \quad (1)$$

Pada Persamaan (1), $p(i)$ menunjukkan probabilitas kemunculan intensitas ke- i pada citra. Nilai n_i menyatakan jumlah piksel pada intensitas ke- i , sedangkan N merupakan jumlah seluruh piksel dalam citra.

D. Arsitektur Model

Pembuatan model klasifikasi dikembangkan berdasarkan arsitektur EfficientNet-B3 yang sebelumnya dilatih pada dataset ImageNet (Tan & Le, 2020). Dalam pembuatan model, arsitektur EfficientNet-B0 dan EfficientNet-B3 dijadikan kandidat model yang akan dibandingkan. Konsep transfer learning diterapkan sebagai usaha untuk memanfaatkan fitur visual yang sudah terlatih pada dataset beskala besar tersebut. Layer klasifikasi di bagian akhir model dimodifikasi untuk mengeluarkan output sejumlah lima sesuai dengan banyak kelas tingkat keparahan retinopati diabetik.

Tabel 1. Arsitektur Model EfficientNet-B3

Tahapan	Komponen	Keterangan
<i>Preprocessing</i>	Citra Fundus Retina	Ukuran $300 \times 300 \times 3$ piksel
	<i>CLAHE, Normalisasi, Augmentasi</i>	Meningkatkan kualitas dan variasi data
<i>Feature Extraction</i>	<i>EfficientNet-B3 (Pretrained ImageNet)</i>	Ekstraksi fitur otomatis dari citra retina
<i>Pooling</i>	<i>Global Average Pooling</i>	Mengurangi dimensi fitur
<i>Regularization</i>	<i>Dropout (0,3)</i>	Mengurangi risiko <i>overfitting</i>
<i>Classification Layer</i>	<i>Fully Connected Layer</i>	Menghasilkan 5 neuron keluaran
<i>Output</i>	<i>Softmax</i>	Probabilitas Grade 0–4

Arsitektur model yang digunakan disajikan pada Tabel 1. Citra fundus retina hasil praproses digunakan sebagai masukan ke *EfficientNet-B3* yang telah dilatih pada *ImageNet*. Fitur yang diekstraksi kemudian diproses melalui *Global Average Pooling* dan lapisan klasifikasi dengan fungsi aktivasi *Softmax* untuk menghasilkan probabilitas lima tingkat keparahan retinopati diabetik sebagaimana dirumuskan pada Persamaan (2).

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (2)$$

Pada Persamaan (2), P_i merupakan probabilitas kelas ke- i yang dihasilkan model. Nilai z_i adalah keluaran neuron ke- i sebelum proses aktivasi, sedangkan K menyatakan jumlah kelas yang digunakan dalam proses klasifikasi.

E. Proses Pelatihan

Pelatihan model dilakukan melalui penggunaan *optimizer AdamW* dengan nilai *learning rate* 0,0001. Batas maksimum epoch yang digunakan adalah sebanyak 50 *epoch* dengan *batch size* 8. Proses pembelajaran dirancang optimal dengan *learning rate scheduler* menggunakan *CosineAnnealingLR*. Selain itu, konsep *early stopping* diterapkan dengan batas 10 *epoch* untuk mencegah terjadinya *overfitting* ketika performa validasi tidak lagi mengalami peningkatan.

Selama proses pelatihan, fungsi kerugian yang digunakan adalah *Cross Entropy Loss* untuk mengukur perbedaan antara hasil prediksi model dan label sebenarnya. Persamaan *Cross Entropy Loss* ditunjukkan pada Persamaan (3).

$$L = - \sum_{i=1}^K w_i y_i \log \log (p_i) \quad (3)$$

Pada Persamaan (3), L menyatakan nilai loss yang digunakan untuk mengukur kesalahan prediksi model. Nilai y_i merupakan label sebenarnya, p_i adalah probabilitas hasil prediksi model, dan K menunjukkan jumlah kelas yang diklasifikasikan.

F. Evaluasi Model

Kinerja model dievaluasi menggunakan beberapa metrik, yaitu AUC-ROC (*Area Under the Receiver Operating Characteristic Curve*), *accuracy*, *Cohen's Kappa*, *precision*, *recall*, *F1-score*, *confusion matrix*, dan kurva ROC. Penggunaan beberapa metrik evaluasi bertujuan untuk memberikan gambaran yang lebih komprehensif mengenai kemampuan model dalam mengklasifikasikan seluruh tingkat keparahan retinopati diabetik, terutama pada kondisi distribusi data yang tidak seimbang.

Salah satu metrik yang digunakan untuk mengevaluasi kinerja model adalah *accuracy*. Nilai *accuracy* dihitung menggunakan Persamaan (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

Pada Persamaan (4), nilai *accuracy* dihitung berdasarkan perbandingan antara jumlah prediksi yang benar dengan jumlah seluruh data. Nilai TP menunjukkan True Positive,

TN menunjukkan True Negative, FP menunjukkan False Positive, dan FN menunjukkan False Negative.

Selain *accuracy*, penelitian ini juga menggunakan *Cohen's Kappa* untuk mengukur tingkat kesesuaian antara hasil prediksi model dan label sebenarnya dengan mempertimbangkan kemungkinan kesesuaian yang terjadi secara acak. Nilai *Cohen's Kappa* dihitung menggunakan Persamaan (5).

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (5)$$

Pada Persamaan (5), κ menunjukkan nilai Cohen's Kappa yang digunakan untuk mengukur tingkat kesesuaian antara hasil prediksi dan label sebenarnya. Nilai P_o merupakan tingkat kesesuaian yang diamati, sedangkan P_e menunjukkan tingkat kesesuaian yang dapat terjadi secara acak.

G. Interpretasi Model Menggunakan Grad-CAM

Untuk meningkatkan interpretabilitas model, penelitian ini menerapkan metode *Gradient-weighted Class Activation Mapping (Grad-CAM)* (Selvaraju et al., 2017). Metode ini menghasilkan peta aktivasi yang menunjukkan area citra retina yang paling berpengaruh terhadap keputusan klasifikasi model. Dengan demikian, hasil prediksi yang dihasilkan tidak hanya dapat dinilai dari sisi performa, tetapi juga dapat dianalisis berdasarkan area retina yang menjadi fokus perhatian model.

HASIL DAN PEMBAHASAN

A. Hasil

1. Perbandingan Arsitektur Model

Tabel 2. Perbandingan Performa *EfficientNet-B0* dan *EfficientNet-B3*

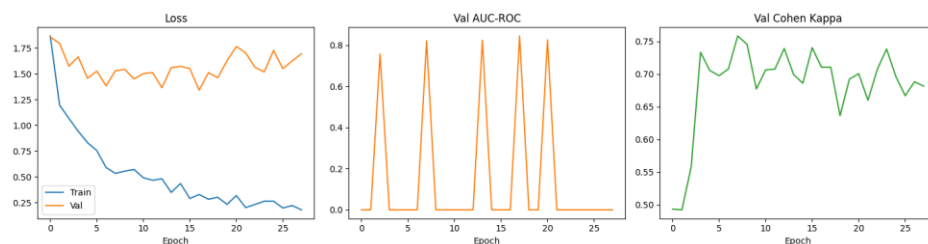
Model	AUC-ROC Macro	<i>Accuracy</i>	<i>Cohen's Kappa</i>
<i>EfficientNet-B0</i>	0,7332	39,13%	0,5037
<i>EfficientNet-B3</i>	0,8186	55,07%	0,6747

Pada penentuan model yang akan digunakan pada penelitian, dilakukan dengan melakukan uji coba dengan menggunakan dua jenis arsitektur *EfficientNet*, di antaranya adalah *EfficientNet-B0* dan *EfficientNet-B3*. Berdasarkan hasil evaluasi, *EfficientNet-B3* memiliki kinerja yang lebih baik dari pada *EfficientNet-B0* pada

semua metrik. Perbedaan kinerja dapat dilihat dari nilai *Accuracy*, *AUC-ROC Macro*, serta *Cohen's Kappa*.

2. Hasil Pelatihan Model

Model *EfficientNet-B3* dilatih menggunakan dataset IDRiD dengan jumlah maksimum 50 epoch. Proses pelatihan dihentikan pada *epoch* ke-28 menggunakan mekanisme *early stopping* karena tidak terjadi peningkatan nilai AUC validasi selama 10 *epoch* berturut-turut. Model terbaik diperoleh pada epoch ke-18 dengan nilai AUC validasi sebesar 0,8444. Berdasarkan grafik pelatihan pada Gambar 1, nilai *training loss* mengalami penurunan yang konsisten dari 1,8651 menjadi sekitar 0,18. Hal tersebut menunjukkan bahwa model mampu mempelajari pola dan karakteristik citra fundus retina dengan baik. Sementara itu, *validation loss* masih mengalami fluktuasi selama proses pelatihan. Kondisi ini menunjukkan adanya perbedaan karakteristik antara data pelatihan dan data validasi yang umumnya ditemukan pada dataset medis dengan distribusi kelas yang tidak seimbang. Selain itu, nilai *Cohen's Kappa* validasi berada pada rentang 0,63–0,76 selama proses pelatihan. Hasil tersebut menunjukkan bahwa model memiliki tingkat konsistensi yang cukup baik dalam melakukan klasifikasi tingkat keparahan retinopati diabetik.



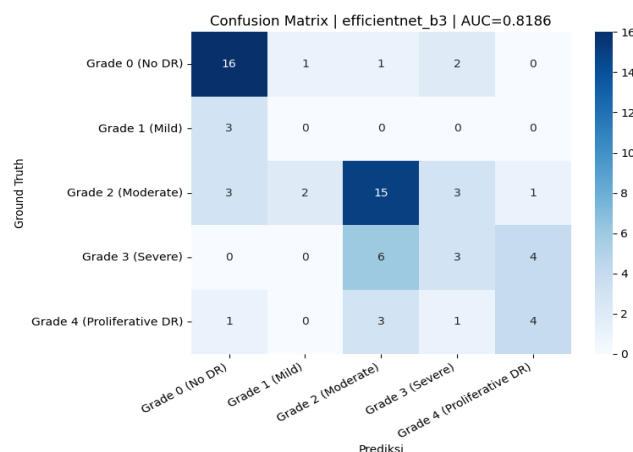
Gambar 1. Grafik *Training Loss*, *Validation Loss*, *AUC-ROC*, dan *Cohen's Kappa*

3. Evaluasi Performa Model

Tabel 3. Hasil Evaluasi Model *EfficientNet-B3*

Kelas	Precision	Recall	F1-Score
Grade 0 (No DR)	0,6957	0,8000	0,7442
Grade 1 (Mild)	0,0000	0,0000	0,0000
Grade 2 (Moderate)	0,6000	0,6250	0,6122
Grade 3 (Severe)	0,3333	0,2308	0,2727
Grade 4 (Proliferative DR)	0,4444	0,4444	0,4444

Berdasarkan Tabel 3, model menunjukkan performa terbaik pada *Grade 0* dengan nilai *recall* sebesar 0,80. Sebaliknya, performa terendah ditemukan pada *Grade 1* karena jumlah data yang relatif sedikit. Hal ini menunjukkan bahwa ketidakseimbangan kelas masih memengaruhi kemampuan model dalam mengenali seluruh kategori secara merata.



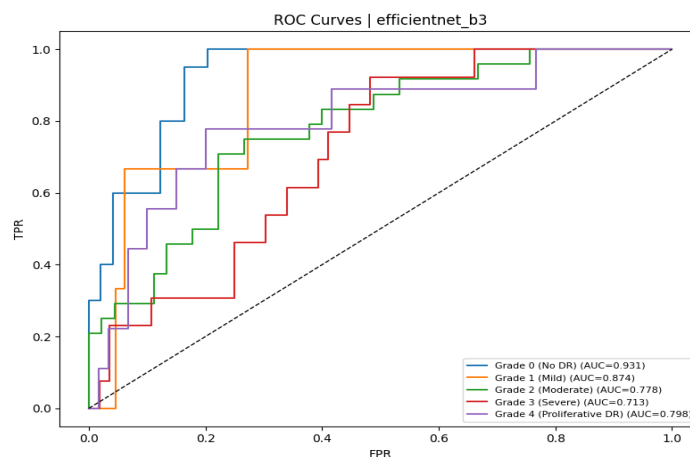
Gambar 2. *Confusion Matrix* Hasil Klasifikasi

Berdasarkan *confusion matrix* yang disajikan pada Gambar 2, model berhasil menunjukkan performansi klasifikasi yang cukup bagus untuk kelas *Grade 0* (No DR) dan *Grade 2* (Moderate). Model mampu mengklasifikasikan sebanyak 16 dari 20 gambar dalam kelas *Grade 0* dan 15 dari 24 gambar dalam kelas *Grade 2*. Ini mengindikasikan bahwa model sudah mampu mengidentifikasi sifat visual dalam dua kelas ini. Namun demikian, semua gambar dalam kelas *Grade 1* (Mild) gagal dikelompokkan secara benar dan cenderung dievaluasi sebagai *Grade 0*. Karena itu, akurasi *precision*, *recall*, dan *F1-score* dalam kelas ini mendekati angka 0 (≈ 0). Selain itu, kesalahan dalam proses klasifikasi masih terjadi pada kelas

Grade 3 (Severe) dan *Grade 4 (Proliferative Diabetic Retinopathy)*, dimana hampir seluruh gambar kelas ini dikelompokkan sebagai *Grade 2 (Moderate)*. Hal ini menunjukkan bahwa model masih kesulitan dalam melakukan identifikasi pada kelas-kelas dengan sifat visual yang mirip. Dalam hal ini, model kesulitan melakukan identifikasi terhadap tingkatan retinopati diabetik yang lebih parah. Secara umum, *confusion matrix* menunjukkan adanya pengaruh *class imbalance* pada beberapa kelas.

4. Analisis Kurva ROC

Kurva ROC digunakan untuk mengevaluasi kemampuan model dalam membedakan masing-masing kelas retinopati diabetik. Berdasarkan Gambar 3, seluruh kurva berada di atas garis diagonal yang menunjukkan bahwa model memiliki kemampuan klasifikasi yang lebih baik dibandingkan tebakan acak. Nilai AUC tertinggi diperoleh pada *Grade 0 (No DR)* sebesar 0,931, sedangkan nilai AUC pada kelas lainnya berkisar antara 0,713 hingga 0,874. Secara keseluruhan, nilai *AUC-ROC Macro* sebesar 0,8186 menunjukkan bahwa model memiliki kemampuan yang baik dalam membedakan lima tingkat keparahan retinopati diabetik.



Gambar 3. Kurva ROC Multi-Class

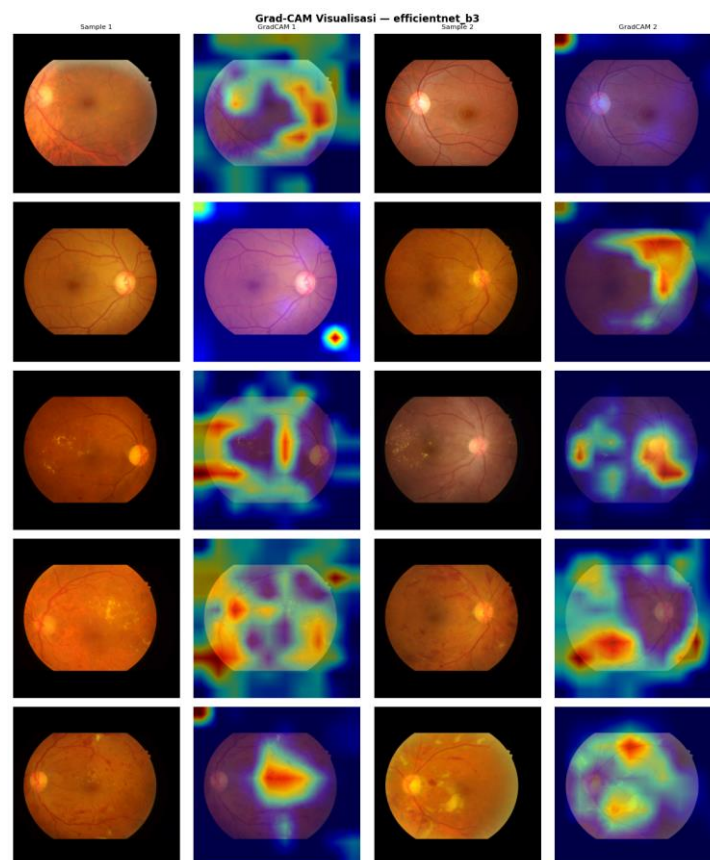
Tabel 3 menunjukkan bahwa kelas *Grade 0* memiliki nilai AUC tertinggi sebesar 0,931, sedangkan kelas *Grade 3* memiliki nilai AUC terendah sebesar 0,713. Meskipun demikian, seluruh kelas memperoleh nilai AUC di atas 0,70 yang menunjukkan kemampuan diskriminatif model yang cukup baik.

Tabel 4. Nilai AUC Tiap Kelas

Kelas	AUC
<i>Grade 0 (No DR)</i>	0,931
<i>Grade 1 (Mild)</i>	0,874
<i>Grade 2 (Moderate)</i>	0,778
<i>Grade 3 (Severe)</i>	0,713
<i>Grade 4 (Proliferative DR)</i>	0,798

5. *Grad-CAM*

Untuk meningkatkan interpretabilitas model, dilakukan visualisasi menggunakan metode *Grad-CAM*. Visualisasi ini menunjukkan area retina yang menjadi fokus model saat melakukan klasifikasi tingkat keparahan retinopati diabetik. Berdasarkan Gambar 4, model memberikan perhatian yang lebih besar pada area yang mengandung lesi seperti mikroaneurisma, perdarahan, dan eksudat, sehingga mendukung proses pengambilan keputusan yang lebih transparan.



Gambar 4. *Grad-CAM*

Hasil visualisasi menunjukkan bahwa model tidak hanya memanfaatkan informasi latar belakang citra, tetapi juga berfokus pada area retina yang relevan secara klinis. Hal ini mengindikasikan bahwa model memiliki kemampuan yang baik dalam mengenali karakteristik retinopati diabetik dan meningkatkan kepercayaan terhadap hasil prediksi yang dihasilkan.

B. Pembahasan

Hasil evaluasi menunjukkan bahwa *EfficientNet-B3* mampu mengklasifikasikan tingkat keparahan retinopati diabetik pada dataset IDRiD, yang ditunjukkan oleh nilai AUC_ROC Macro sebesar 0,8186. Selain itu, nilai *accuracy* sebesar 55,07% dan *Cohen's Kappa* sebesar 0,6747 menunjukkan tingkat kesesuaian yang cukup tinggi antara hasil prediksi model dan label sebenarnya. Temuan ini mendukung penelitian sebelumnya yang menyatakan bahwa keluarga *EfficientNet* efektif dalam melakukan ekstraksi fitur pada citra fundus retina karena mampu mempertahankan keseimbangan antara performa dan efisiensi komputasi.

Berdasarkan hasil evaluasi per kelas, performa terbaik diperoleh pada *Grade 0 (No Diabetic Retinopathy)* dan *Grade 2 (Moderate)*. Sebaliknya, kelas *Grade 1 (Mild)* masih menunjukkan hasil yang sangat rendah pada metrik precision, recall, maupun *F1-score*. Hasil confusion matrix memperlihatkan bahwa sebagian besar sampel *Grade 1* diprediksi sebagai *Grade 0*, sedangkan beberapa sampel *Grade 3* dan *Grade 4* masih diklasifikasikan sebagai *Grade 2*. Kondisi tersebut menunjukkan bahwa model mengalami kesulitan dalam membedakan kelas-kelas yang memiliki karakteristik visual yang serupa. Selain itu, jumlah data yang tidak seimbang antar kelas juga menjadi faktor yang memengaruhi kemampuan model dalam mengenali kategori minoritas.

Penerapan CLAHE membantu meningkatkan kualitas visual citra sehingga karakteristik retina dapat terlihat lebih jelas selama proses ekstraksi fitur. Di sisi lain, penggunaan *WeightedRandomSampler* dan *class weight* berkontribusi dalam mengurangi dominasi kelas mayoritas selama proses pelatihan. Visualisasi menggunakan Grad-CAM menunjukkan bahwa model memberikan perhatian pada area retina yang mengandung lesi seperti mikroaneurisma, eksudat, dan perdarahan, yang merupakan indikator penting retinopati diabetik. Temuan tersebut menunjukkan bahwa model tidak hanya menghasilkan prediksi yang cukup baik, tetapi juga memiliki tingkat interpretabilitas yang lebih tinggi. Meskipun demikian, peningkatan performa pada

kelas minoritas masih diperlukan agar kemampuan klasifikasi dapat menjadi lebih merata pada seluruh tingkat keparahan penyakit.

C. Kesimpulan

Penelitian ini berhasil mengembangkan model klasifikasi tingkat keparahan retinopati diabetik menggunakan arsitektur *EfficientNet-B3* berbasis transfer learning pada dataset IDRiD. Tahap praproses citra dengan CLAHE serta penerapan *WeightedRandomSampler* dan *class weight* terbukti membantu mengatasi permasalahan kualitas citra dan ketidakseimbangan distribusi data selama proses pelatihan. Berdasarkan hasil pengujian, model menghasilkan nilai *AUC-ROC Macro* sebesar 0,8186, *accuracy* sebesar 55,07%, dan *Cohen's Kappa* sebesar 0,6747. Evaluasi pada setiap kelas menunjukkan bahwa performa terbaik diperoleh pada *Grade 0*, sedangkan *Grade 1* masih menjadi tantangan karena keterbatasan jumlah sampel dan kemiripan karakteristik visual dengan kelas lainnya. Selain menghasilkan performa klasifikasi yang cukup baik, penggunaan *Grad-CAM* juga mampu memberikan gambaran mengenai area retina yang menjadi dasar pengambilan keputusan model. Secara keseluruhan, kombinasi *EfficientNet-B3*, CLAHE, *WeightedRandomSampler*, *class weight*, dan *Grad-CAM* menunjukkan potensi yang menjanjikan untuk mendukung proses skrining retinopati diabetik secara otomatis. Penelitian selanjutnya dapat difokuskan pada penerapan teknik penyeimbangan data yang lebih efektif atau penggunaan arsitektur yang lebih kompleks guna meningkatkan performa pada kelas minoritas.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga penelitian ini dapat diselesaikan dengan baik. Penulis juga menyampaikan terima kasih kepada dosen pembimbing yang telah memberikan arahan, masukan, dan bimbingan selama proses penelitian. Selain itu, penulis mengucapkan terima kasih kepada Program Studi Sistem Informasi Universitas Darwan Ali Sampit atas dukungan yang diberikan selama pelaksanaan penelitian.

DAFTAR PUSTAKA

- Ahmed, F. (2025). *Addressing High Class Imbalance in Multi-Class Diabetic Retinopathy Severity Grading with Augmentation and Transfer Learning*. <http://arxiv.org/abs/2507.17121>
- Arora, L., Singh, S. K., Kumar, S., Gupta, H., Alhalabi, W., Arya, V., Bansal, S., Chui, K. T., & Gupta, B. B. (2024). Ensemble *deep learning* and EfficientNet for accurate diagnosis of diabetic retinopathy. *Scientific Reports*, 14(1), 30554. <https://doi.org/10.1038/s41598-024-81132-4>
- Baba, S. M., Bala, I., Dhiman, G., Sharma, A., & Viriyasitavat, W. (2024). Automated diabetic retinopathy severity grading using novel DR-ResNet + *deep learning* model. *Multimedia Tools and Applications*, 83(28), 71789–71831. <https://doi.org/10.1007/s11042-024-18434-2>
- Chavan, S., & Choubey, N. (2023). An automated diabetic retinopathy of severity grade classification using transfer learning and fine-tuning for fundus images. *Multimedia Tools and Applications*, 82(24), 36859–36884. <https://doi.org/10.1007/s11042-023-15135-0>
- Ikram, A., & Imran, A. (2025). ResViT FusionNet Model: An explainable AI-driven approach for automated grading of diabetic retinopathy in retinal images. *Computers in Biology and Medicine*, 186, 109656. <https://doi.org/10.1016/j.compbimed.2025.109656>
- Kropp, M., Golubnitschaja, O., Mazurakova, A., Koklesova, L., Sargheini, N., Vo, T.-T. K. S., de Clerck, E., Polivka, J., Potuznik, P., Polivka, J., Stetkarova, I., Kubatka, P., & Thumann, G. (2023). Diabetic retinopathy as the leading cause of blindness and early predictor of cascading complications—risks and mitigation. *EPMA Journal*, 14(1), 21–42. <https://doi.org/10.1007/s13167-023-00314-8>
- Li, H., Jia, W., Vujosevic, S., Sabanayagam, C., Grauslund, J., Sivaprasad, S., & Wong, T. Y. (2024). Current research and future strategies for the management of vision-threatening diabetic retinopathy. *Asia-Pacific Journal of Ophthalmology*, 13(5), 100109. <https://doi.org/10.1016/j.apjo.2024.100109>
- Muntaha, S., Salam, S. S., & Mustakim, N. (2024). An Explainable AI-based *Deep Learning* Model for Classification of Diabetic Retinopathy Stages Using Retinal Fundus Images. *2024 27th International Conference on Computer and Information Technology (ICCIT)*, 3010–3015. <https://doi.org/10.1109/ICCIT64611.2024.11022403>
- Ong, K. L., Stafford, L. K., McLaughlin, S. A., Boyko, E. J., Vollset, S. E., Smith, A. E., Dalton, B. E., Duprey, J., Cruz, J. A., Hagins, H., Lindstedt, P. A., Aali, A., Abate, Y. H., Abate, M. D., Abbasian, M., Abbasi-Kangevari, Z., Abbasi-Kangevari, M., Abd ElHafeez, S., Abd-Rabu, R., ... Vos, T. (2023). Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a

- systematic analysis for the Global Burden of Disease Study 2021. *The Lancet*, 402(10397), 203–234. [https://doi.org/10.1016/S0140-6736\(23\)01301-6](https://doi.org/10.1016/S0140-6736(23)01301-6)
- Pamungkas, Y., Triandini, E., Yunanto, W., & Thwe, Y. (2025). Enhancing Diabetic Retinopathy Classification in Fundus Images using CNN Architectures and Oversampling Technique. *Journal of Robotics and Control (JRC)*, 6(1), 413–425. <https://doi.org/10.18196/jrc.v6i1.25331>
- Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., & Meriaudeau, F. (2018). Indian Diabetic Retinopathy Image Dataset (IDRiD): A Database for Diabetic Retinopathy Screening Research. *Data*, 3(3), 25. <https://doi.org/10.3390/data3030025>
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *2017 IEEE International Conference on Computer Vision (ICCV)*, 618–626. <https://doi.org/10.1109/ICCV.2017.74>
- Tan, M., & Le, Q. V. (2020). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. <http://arxiv.org/abs/1905.11946>
- Vij, R., & Arora, S. (2023). A novel deep transfer learning based computerized diagnostic Systems for Multi-class imbalanced diabetic retinopathy severity classification. *Multimedia Tools and Applications*, 82(22), 34847–34884. <https://doi.org/10.1007/s11042-023-14963-4>
- Zhang, J., & Juan Chen. (2024). Research on grading detection methods for diabetic retinopathy based on deep learning. *Pakistan Journal of Medical Sciences*, 41(1), 225–229. <https://doi.org/10.12669/pjms.41.1.9171>



© 2026 by authors. Content on this article is licensed under a Creative Commons Attribution 4.0 International license. (<http://creativecommons.org/licenses/by/4.0/>).