

Pendekatan Augmentasi Citra Fundus pada Model EfficientNet untuk Klasifikasi Tingkat Keparahan Retinopati Diabetik dengan *Dataset* Tidak Seimbang

CHALIFA CHAZAR , MUHAMMAD ARKAN ADLI, JASMAN PARDEDE, MUHAMMAD ICHWAN

Informatika, Institut Teknologi Nasional Bandung, Indonesia
Email : chalifa@itenas.ac.id

Received 8 Juli 2025 | *Revised* 17 November 2025 | *Accepted* 24 November 2025

ABSTRAK

Retinopati diabetik (RD) adalah komplikasi diabetes mellitus yang menyerang pembuluh darah retina dan berpotensi menyebabkan kebutaan jika tidak terdeteksi dini. Citra fundus retina berperan penting dalam mendeteksi serta mengklasifikasikan tingkat keparahan RD karena mampu menampilkan kelainan secara jelas. Tantangan utama dalam klasifikasi RD adalah ketidakseimbangan data antar kelas. Penelitian ini mengusulkan penggunaan EfficientNet-B0 dengan augmentasi gambar terarah pada dataset APTOS 2019. Hasil evaluasi menunjukkan peningkatan akurasi dari 73,84% menjadi 82,56% serta F1-score 0,8241. Peningkatan signifikan terlihat pada kelas minoritas, misalnya Mild dari 0,1429 menjadi 0,65 dan Severe dari 0,087 menjadi 0,4211. Temuan ini membuktikan bahwa augmentasi terarah efektif dalam mengurangi bias kelas mayoritas dan meningkatkan keandalan model.

Kata kunci: *augmentasi, EfficientNet, ketidakseimbangan kelas, retinopati diabetik*

ABSTRACT

Diabetic retinopathy (DR) is a complication of diabetes mellitus that affects the retinal blood vessels and may lead to blindness if not detected early. Fundus images play a crucial role in detecting and classifying the severity of DR as they clearly reveal pathological abnormalities. The main challenge in DR classification lies in the imbalance across severity classes. This study proposes the use of EfficientNet-B0 combined with targeted image augmentation on the APTOS 2019 dataset. The evaluation results show an improvement in accuracy from 73.84% to 82.56% and a F1-score of 0.8241. Significant gains are observed in minority classes, such as Mild (from 0.1429 to 0.65) and Severe (from 0.087 to 0.4211). These findings demonstrate that targeted augmentation is effective in reducing majority-class bias and improving model reliability.

Keywords: *class imbalance, data augmentation, diabetic retinopathy, EfficientNet*

1. PENDAHULUAN

Diabetes Mellitus (DM) merupakan salah satu masalah kesehatan global yang terus meningkat secara signifikan. Menurut **(International Diabetes Federation, 2024)**, pada tahun 2024 diperkirakan terdapat sekitar 589 juta orang dewasa berusia 20–79 tahun di seluruh dunia yang hidup dengan diabetes. Jumlah ini diproyeksikan akan meningkat menjadi 853 juta orang pada tahun 2050. Peningkatan ini turut memperbesar risiko komplikasi kronis, salah satunya adalah *diabetic retinopathy* (DR), yang dapat menyebabkan kebutaan permanen jika tidak terdeteksi dan ditangani secara dini. Dalam praktik klinis, citra fundus retina menjadi alat diagnostik utama untuk mendeteksi dan mengklasifikasikan tingkat keparahan DR. Citra ini mampu merekam kondisi retina secara detail sehingga berbagai kelainan seperti perdarahan kecil, bercak eksudat, dan perubahan vaskular dapat diidentifikasi dengan jelas.

Dalam beberapa tahun terakhir, pendekatan berbasis *deep learning* semakin banyak digunakan dalam bidang medis, termasuk untuk mendeteksi DR melalui citra fundus retina. *Convolutional Neural Network* (CNN) menjadi arsitektur yang paling umum digunakan. Berbagai studi menunjukkan potensi besar dari arsitektur seperti *EfficientNet* dan *VGG16* dalam mengklasifikasikan tingkat keparahan DR, terlebih dengan bantuan *transfer learning* dan *fine-tuning* **(Guerbai, dkk., 2024)**. Model *RA-EfficientNet* yang dikembangkan oleh **(Yi, dkk., 2021)**, berhasil mencapai akurasi hingga 0,98 untuk klasifikasi dua kelas. Pada penelitian lainnya, model *EfficientNet* dibanding *ResNet* dan *VGG* mencatat *F1-score* sebesar 0,84 yang menunjukkan performa lebih baik pada klasifikasi DR **(Chilukoti, dkk., 2022)**, berkat metode *compound scaling* yang mengoptimalkan kedalaman, lebar, dan resolusi jaringan secara bersamaan **(Tan & Le, 2019)**.

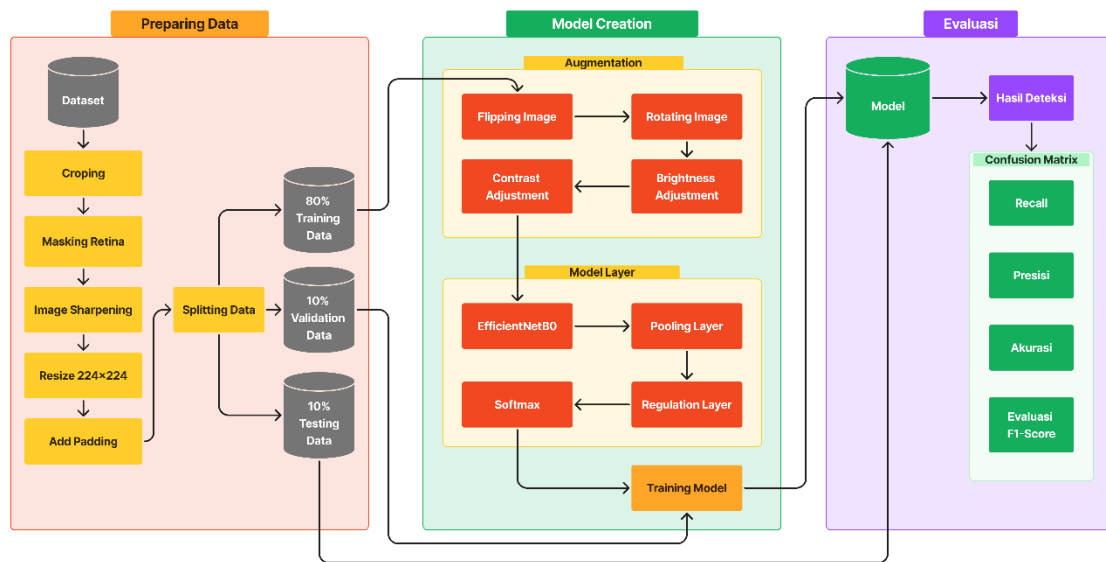
Meskipun model *EfficientNet* menunjukkan kinerja yang menjanjikan, tantangan signifikan masih muncul dari ketidakseimbangan distribusi data antar kelas. Hal ini dapat menyebabkan bias pada prediksi, terutama terhadap kelas mayoritas. Penelitian oleh **(Zahir & Adi Saputra, 2024)** menegaskan bahwa data yang tidak seimbang berpengaruh langsung terhadap penurunan akurasi dan keandalan klasifikasi. Studi lain menunjukkan bahwa model pada kelas minoritas dalam *dataset APTOS 2019*, hanya mampu mencapai *F1-score* sebesar 0,35 **(Abdullah, dkk., 2022)**, yang menandakan perlunya strategi penanganan khusus untuk mengatasi permasalahan ini.

Salah satu pendekatan yang banyak digunakan untuk mengatasi ketidakseimbangan data adalah teknik augmentasi gambar. Teknik ini memungkinkan penciptaan variasi data baru dari citra yang sudah ada melalui transformasi seperti rotasi, *cropping*, pencahayaan, dan lainnya. Penelitian **(Sanjaya & Ayub, 2020)** menunjukkan bahwa augmentasi dengan metode *random crop*, *rotate*, dan *mixup* mampu meningkatkan akurasi model *ResNet-152* dari 76% menjadi 82,67% dalam klasifikasi mobil. **(Santoso, 2024)** juga berhasil menerapkan augmentasi seperti rotasi, translasi, *zoom*, dan *blur* untuk meningkatkan performa deteksi aksara Kawi pada *dataset* kecil, yang berkontribusi pada pencegahan *overfitting* dan peningkatan generalisasi model. Hal ini diperkuat oleh analisis komparatif dari **(Pak, dkk., 2020)**, yang secara spesifik mendemonstrasikan dampak prapemrosesan dan augmentasi, dalam studi mereka, akurasi model *EfficientNet* melonjak dari 65,5% pada data asli menjadi 79% setelah penerapan prapemrosesan menggunakan *CLAHE* dan augmentasi data.

Penelitian ini bertujuan untuk mengembangkan model klasifikasi DR dengan mengintegrasikan arsitektur *EfficientNet* dan teknik augmentasi gambar. Diharapkan pendekatan ini dapat meningkatkan kemampuan klasifikasi, terutama pada kelas minoritas, serta memberikan kontribusi dalam pengembangan sistem deteksi dini yang lebih akurat dan andal dalam dunia medis.

2. METODOLOGI PENELITIAN

Bab ini menjelaskan metodologi yang diterapkan dalam penelitian, dimulai dari tahapan pengumpulan dan pra-pemrosesan data citra fundus retina, pembentukan model klasifikasi menggunakan arsitektur *EfficientNet-B0* dengan implementasi augmentasi data, hingga proses validasi dan evaluasi kinerja model akhir menggunakan berbagai metrik akurasi.



Gambar 1. Blok Diagram Sistem

Gambar 1. menunjukkan alur penelitian yang terdiri dari tiga tahapan utama, yaitu persiapan data, pembuatan model, dan evaluasi. Tahap persiapan data bertujuan untuk menghasilkan data citra retina yang siap digunakan dalam pelatihan model melalui proses prapemrosesan dan pembagian data. Tahap pembuatan model berfokus pada pengembangan dan pelatihan model klasifikasi menggunakan arsitektur *EfficientNetB0* untuk mengenali tingkat keparahan retinopati diabetik. Sementara itu, tahap evaluasi dilakukan untuk menilai kinerja model yang telah dilatih menggunakan metrik seperti *recall*, presisi, akurasi, dan *F1-score*.

2.1. Pengumpulan Data

Penelitian ini menggunakan *dataset APTOS 2019* yang diunduh dari situs Kaggle yang disediakan oleh Rumah Sakit Mata Aravind di India (Karthik, dkk, 2019). *Dataset* ini terdiri dari 3662 gambar beresolusi tinggi yang telah didiagnosis oleh dokter yang terlatih, dan diklasifikasikan ke dalam lima tingkat keparahan DR, dari tingkat nol hingga empat (Karthik, dkk, 2019).



Gambar 2. Contoh Data dari *Dataset APTOS 2019* (a) No DR, (b) Mild, (c) Moderate, (d) Severe, dan (e) Proliferative DR. (Karthik dkk., 2019)

Gambar 2. menunjukkan contoh citra fundus retina yang digunakan sebagai *dataset* penelitian, yang terdiri dari lima tingkat keparahan retinopati diabetik, yaitu *No DR*, *Mild*, *Moderate*, *Severe*, dan *Proliferative DR*. Setiap tingkat menggambarkan perbedaan kondisi retina berdasarkan tingkat kerusakan akibat penyakit.

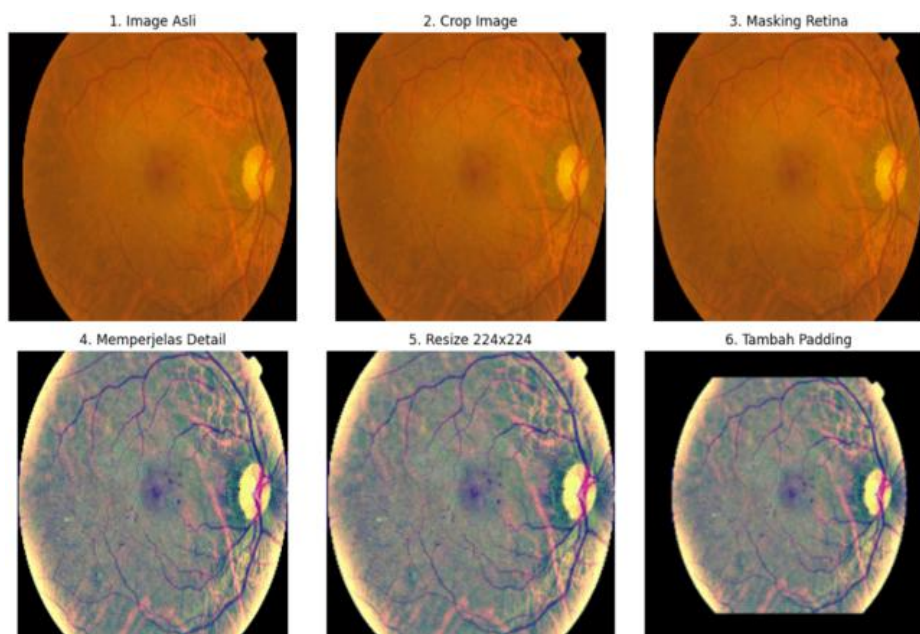
Tabel 1. Distribusi *Dataset*

Kelas	Jumlah Gambar	Persentase (%)
<i>No_DR</i>	1805	49,3
<i>Mild</i>	370	10,1
<i>Moderate</i>	999	27,3
<i>Severe</i>	193	5,3
<i>Proliferate_DR</i>	295	8,1

Tabel 1. menunjukkan distribusi *dataset* APTOS 2019 yang terdiri dari 3662 gambar fundus retina dengan lima tingkat keparahan DR. Kelas *No_DR* (normal) mendominasi *dataset* dengan 1805 gambar (49,3% dari total), diikuti oleh kelas *Moderate* dengan 999 gambar (27,3%). Kelas minoritas meliputi *Mild* dengan 370 gambar (10,1%), *Proliferate_DR* dengan 295 gambar (8,1%), dan *Severe* dengan jumlah terkecil yaitu 193 gambar (5,3%). Ketidakeimbangan data yang signifikan terlihat jelas dengan rasio antara kelas mayoritas (*No_DR*) dan kelas minoritas (*Severe*) mencapai 9 banding 1. Ketidakeimbangan ini dapat memengaruhi performa model dalam mengenali kelas minoritas karena model cenderung lebih mudah mempelajari pola dari kelas mayoritas.

2.2. *Preprocessing Data*

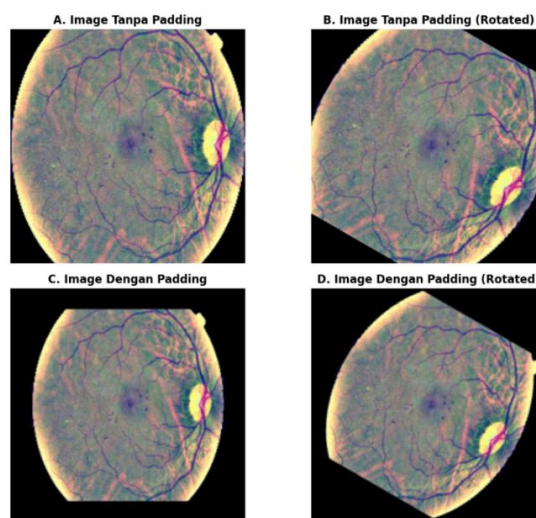
Sebelum data citra retina dapat digunakan untuk melatih model deteksi retinopati diabetik, diperlukan tahapan prapemrosesan untuk mengoptimalkan kualitas dan konsistensi data. Proses ini bertujuan meningkatkan kemampuan model dalam mengidentifikasi fitur-fitur penting pada retina dengan menghilangkan *noise* dan memastikan kesesuaian data input.



Gambar 3. Tahapan Prapemrosesan *Dataset* Citra Fundus

Gambar 3. menunjukkan tahap pertama yakni *cropping* untuk menghilangkan area hitam di sekitar retina yang tidak mengandung informasi diagnostik relevan. Area ini dapat mengganggu proses pembelajaran model sehingga perlu dieliminasi. Selanjutnya, dilakukan pembuatan masker retina berdasarkan analisis tingkat kecerahan untuk memisahkan region retina dari *background*. Setelah masker terbentuk, *background* non-retina diganti dengan warna hitam untuk memfokuskan analisis pada *region of interest*.

Tahap selanjutnya adalah memperjelas gambar retina agar lebih mudah dianalisis. Dalam proses ini, bagian gambar yang terlalu gelap dibuat lebih terang, sedangkan bagian yang terlalu terang dibuat lebih redup sehingga seluruh gambar memiliki tingkat pencahayaan yang seimbang. Proses ini juga mempertajam detail-detail penting seperti pembuluh darah dan bintik-bintik abnormal yang menjadi tanda penyakit. Setelahnya, citra dilakukan perubahan ukuran menjadi dimensi standar 224×224 piksel untuk memastikan konsistensi ukuran input yang diperlukan dalam arsitektur *EfficientNet-B0*.



Gambar 4. Perbandingan Rotasi Gambar dengan *Padding* dan Tidak

Tahap akhir yang krusial adalah penambahan *padding* beberapa piksel berwarna latar belakang dengan memberikan perhatian khusus pada aspek *positioning* retina. *Padding* ini berfungsi memastikan retina berada pada posisi tepat di tengah citra untuk memberikan ruang aman. Gambar 4. menunjukkan posisi retina yang terpusat ini bertujuan mencegah terjadinya *cropping* yang tidak diinginkan ketika dilakukan augmentasi rotasi selama proses *training*. Tanpa *padding* yang memadai, rotasi citra dapat menyebabkan bagian penting retina terpotong atau memunculkan artefak di tepian citra yang mengganggu proses pembelajaran model. Dengan memastikan retina berada di pusat koordinat rotasi, seluruh struktur anatomis retina tetap terjaga dalam *boundary* citra meskipun mengalami transformasi rotasi dengan berbagai sudut.

Serangkaian prapemrosesan ini menghasilkan *dataset* yang homogen dengan kualitas citra yang optimal dan *robustness* terhadap augmentasi geometris, sehingga menghasilkan *dataset* yang lebih berkualitas untuk proses pelatihan model selanjutnya.

2.3. *Splitting Data*

Dalam penelitian ini, *dataset APTOS 2019* dibagi menjadi tiga bagian dengan rasio 80:10:10, masing-masing untuk tahap *training*, *validation*, dan *testing*. Proses pembagian menerapkan teknik *stratified sampling* guna menjaga keseimbangan distribusi kelas DR di setiap bagian data.

2.4 Augmentasi Data

Ketidakseimbangan distribusi kelas pada *dataset APTOS 2019* menjadi hambatan utama dalam klasifikasi retinopati diabetik, di mana model pada kelas minoritas hanya mencapai *F1-score* sebesar 0,35 (**Abdullah, dkk, 2022**). Ketidakseimbangan *dataset* juga dapat menurunkan akurasi klasifikasi yang terbukti pada penelitian (**Zahir & Adi Saputra, 2024**). Untuk mengatasi hal tersebut, penelitian ini menerapkan augmentasi data selektif dengan melakukan *oversampling* pada kelas minoritas hingga jumlahnya seimbang dengan kelas mayoritas. Tujuannya adalah memastikan proporsi yang setara untuk setiap tingkat keparahan retinopati diabetik dalam pelatihan model *EfficientNet*.

Pemilihan teknik augmentasi didasarkan pada karakteristik citra fundus retina dan bukti empiris sebelumnya, augmentasi yang tepat terbukti dapat meningkatkan akurasi dan mencegah *overfitting* (**Sanjaya & Ayub, 2020; Santoso, 2024**). Untuk meningkatkan keberagaman data tanpa mengubah informasi fitur, digunakan kombinasi penyesuaian intensitas kecerahan, kontras, *horizontal flip*, dan *rotation*. Penyesuaian ini membantu model beradaptasi terhadap variasi pencahayaan, kualitas gambar, dan orientasi pengambilan gambar.

Tabel 2. Ciri Khusus Tingkat Keparahan Retinopati Diabetik (Eimoufidi & Ammoun, 2021)

Tahapan DR	Rekaman Oftalmoskop (Perekaman dari Alat Oftalmoskop)
Normal	Tanpa adanya kelainan apa pun
NPDR Ringan	Hanya terdapat mikroaneurisma
NPDR Sedang	Mikroaneurisma ada tetapi dalam jumlah yang lebih kecil daripada NPDR berat
NPDR Berat	Penebalan vena (Venous beading) di dua wilayah atau lebih. Kelainan mikrovaskuler intraretinal menonjol (Prominent intraretinal microvascular abnormality) di satu wilayah atau lebih
PDR	Pendarahan vitreous/pra-retina (Vitreous/pre-retinal hemorrhage). Neovaskularisasi (pembentukan pembuluh darah baru)

Untuk memastikan integritas citra data latih hasil augmentasi, dilakukan evaluasi manual terhadap ciri khusus retinopati diabetik. Evaluasi mengacu pada Tabel 2. klasifikasi klinis: Normal, NPDR (*Mild, Moderate, Severe*), dan PDR (**Eimoufidi & Ammoun, 2021**). Verifikasi ini memastikan ciri khas setiap tingkat keparahan tetap terlihat. Misalnya, pada NPDR ringan dan sedang, *mikroaneurisma* harus tetap terlihat. Untuk NPDR berat, *venous beading* dan *abnormalitas mikrovaskuler intraretinal (IRMA)* harus jelas. Sedangkan pada PDR, *neovaskularisasi* dan pendarahan *vitreoretinal/preretinal* harus tetap dikenali. Hasil augmentasi menunjukkan citra tetap mempertahankan ciri-ciri khas sesuai standar medis, sehingga layak untuk pelatihan model.

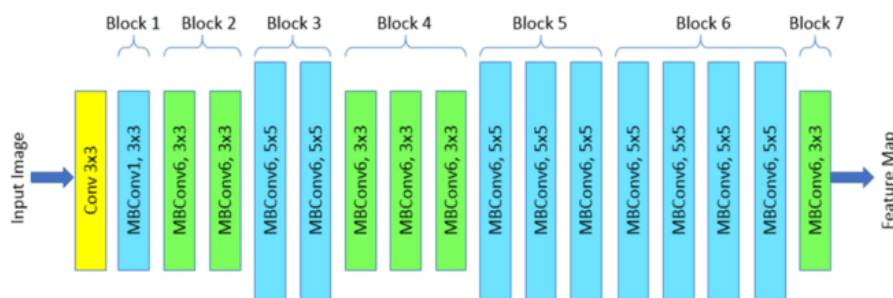
2.5 Pengembangan Model

Penelitian ini menggunakan arsitektur model *EfficientNet* berdasarkan keunggulan *compound scaling method* yang mengoptimalkan kedalaman, lebar, dan resolusi jaringan secara simultan (Tan & Le, 2019). Pendekatan ini memungkinkan pencapaian performa tinggi dengan efisiensi komputasi yang optimal, sebagaimana terbukti dalam penelitian (**Chilukoti, dkk, 2022**) yang mencatat *F1-score* 0,84 untuk klasifikasi retinopati diabetik, mengungguli arsitektur *ResNet* dan *VGG*.

Pemilihan varian *EfficientNet-B0* didasarkan pada hasil evaluasi komparasi yang dilakukan (Yi, dkk, 2021) dimana *EfficientNet-B7* mencapai akurasi tertinggi 98,36% namun memiliki ukuran *input* yang besar dan jumlah parameter tertinggi. Sebaliknya, *EfficientNet-B0* mampu mencapai akurasi 97,95% dengan ukuran *input* yang lebih kecil (224×224) dan parameter yang lebih rendah. Mengingat keterbatasan sumber daya *hardware*, *EfficientNet-B0* menjadi pilihan optimal yang menyeimbangkan performa dan efisiensi komputasi.

2.5.1 Arsitektur EfficientNet

Arsitektur *backbone EfficientNet-B0* yang digunakan untuk mengekstraksi fitur dari citra fundus, seperti pada Gambar 5. Model ini dimulai dengan lapisan konvolusi 3×3 yang berfungsi untuk menangkap fitur dasar seperti tepi dan tekstur dari gambar. Selanjutnya, fitur tersebut diproses melalui tujuh blok utama yang tersusun dari *MBConv* (*Mobile Inverted Bottleneck Convolution*), yaitu jenis konvolusi yang dirancang agar efisien dan ringan secara komputasi.



Gambar 5. Arsitektur EfficientNet-B0 (Naeeni Davarani, dkk, 2024)

Setiap blok *MBConv* memiliki jumlah lapisan dan ukuran *kernel* yang berbeda, seperti 3×3 atau 5×5 , yang membantu model memahami informasi pada berbagai skala. Blok-blok awal fokus pada fitur sederhana, sedangkan blok-blok selanjutnya menangkap pola yang lebih kompleks. Dengan struktur ini, *EfficientNet-B0* dapat menghasilkan *feature map* akhir yang kaya informasi, namun tetap efisien untuk dilatih dan digunakan.

2.5.2 Arsitektur Model

Untuk mengadaptasi EfficientNet pada tugas klasifikasi DR, dikembangkan arsitektur model gabungan yang mengintegrasikan *feature extractor EfficientNet-B0* dengan *custom classification head*. Arsitektur ini dirancang untuk mengoptimalkan pembelajaran fitur dari citra fundus retina sambil mencegah *overfitting* pada *dataset* yang telah diperbanyak variasinya menggunakan teknik augmentasi.

Tabel 3. Arsitektur Model Klasifikasi Berbasis EfficientNet-B0

Layer (type)	Output Shape	Param #
efficientnetb0 (Functional)	(None, 7, 7, 1280)	4,049,571
global_average_pooling2d (GlobalAveragePooling2D)	(None, 1280)	0
dropout (Dropout)	(None, 1280)	0
dense (Dense)	(None, 128)	163,968
batch_normalization (BatchNormalization)	(None, 128)	512
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 5)	645

Pada Tabel 3. menunjukkan arsitektur model dalam penelitian ini yang menggunakan *EfficientNet-B0* sebagai *backbone* untuk mengekstraksi fitur dari citra fundus, menghasilkan representasi berukuran (*None*, 7, 7, 1280). Selanjutnya, fitur ini diproses melalui *Global Average Pooling 2D* untuk mereduksi dimensi menjadi vektor 1280 dimensi, disertai dengan *Dropout* sebagai langkah awal regularisasi. Bagian *classification head* terdiri dari *Dense layer* dengan 128 neuron untuk transformasi fitur, dilengkapi *Batch Normalization* guna menstabilkan proses pelatihan, serta *Dropout* tambahan untuk mengurangi risiko *overfitting* (Duque-Arribas, dkk, 2025). Lapisan *output* menggunakan *Dense layer* dengan 5 neuron dan aktivasi softmax untuk klasifikasi multikelas DR (Normal hingga PDR). Arsitektur *EfficientNet-B0* memiliki 4.049.571 parameter, dan hanya 165.125 parameter ditambahkan melalui *classification head*. Desain ini menggabungkan kekuatan ekstraksi fitur *EfficientNet* dengan regulasi.

2.5.3 Hyperparameter

Selama proses pelatihan model, *Categorical Crossentropy* dipilih sebagai fungsi *loss* untuk mengukur perbedaan antara prediksi model dan label sebenarnya, yang sangat sesuai untuk tugas klasifikasi multikelas. Untuk mengoptimalkan bobot model, *Adam optimizer* digunakan dengan *learning rate* sebesar 0.0001. Setiap iterasi pelatihan memproses *batch size* sebanyak 32 citra. Model dilatih selama 50 *epoch*, namun untuk mencegah *overfitting* dan menghemat sumber daya komputasi, *Early Stopping* diimplementasikan dengan *patience* 15, yang berarti pelatihan akan berhenti jika validasi *loss* tidak menunjukkan penurunan selama 15 *epoch* berturut-turut.

2.6 Evaluasi Model

Evaluasi model dilakukan menggunakan data *testing* yang telah dipisahkan sebesar 10% dari total *dataset*. Proses evaluasi menggunakan *confusion matrix* sebagai alat utama untuk menilai kinerja model klasifikasi yang telah dibangun.

		Actual Values	
		1 (Positive)	0 (Negative)
Predicted Values	1 (Positive)	TP (True Positive)	FP (False Positive) <i>Type I Error</i>
	0 (Negative)	FN (False Negative) <i>Type II Error</i>	TN (True Negative)

Gambar 6. Ilustrasi *Confusion Matrix*

Pada Gambar 6. menunjukkan *Confusion matrix* atau matriks kesalahan merupakan tabel yang menyajikan perbandingan antara hasil prediksi model dengan label sebenarnya dari data uji (Syurfi, 2021). Matriks ini terdiri dari empat komponen utama yaitu *True Positive* (TP) dimana model memprediksi dengan benar kelas positif sebagai positif, *True Negative* (TN) dimana model memprediksi dengan benar kelas negatif sebagai negatif, *False Positive* (FP) dimana model salah memprediksi kelas negatif sebagai positif, dan *False Negative* (FN) dimana model salah memprediksi kelas positif sebagai negatif. Berdasarkan empat komponen yang didapat, dilakukan perhitungan metrik evaluasi untuk mengukur kinerja model secara komprehensif pada Persamaan (1) - (4).

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Akurasi mengukur proporsi prediksi yang benar dari keseluruhan prediksi yang dibuat oleh model. Metrik ini memberikan gambaran umum tentang kemampuan model dalam mengklasifikasikan seluruh data dengan tepat **(Istiqomah, 2022)**

$$Presisi = \frac{TP}{TP + FP} \quad (2)$$

Presisi mengevaluasi ketepatan model dalam menghasilkan prediksi positif yang benar. Presisi tinggi menunjukkan bahwa model menghasilkan sedikit *False Positive*, sehingga mayoritas prediksi positif adalah benar **(Istiqomah, 2022)**

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Recall menilai kemampuan model dalam mendeteksi semua instance positif yang sebenarnya ada dalam data. Metrik ini penting untuk meminimalkan *False Negative*, terutama dalam kasus di mana kehilangan deteksi positif memiliki konsekuensi yang signifikan **(Istiqomah, 2022)**

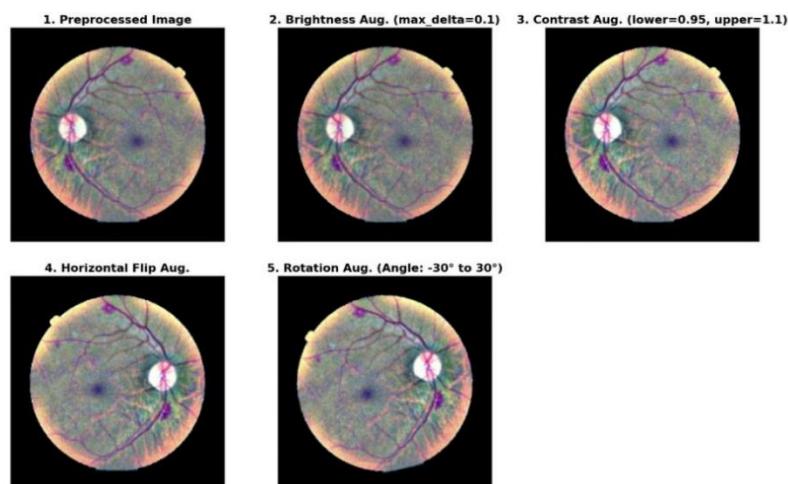
$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (4)$$

F1-score merupakan rata-rata harmonis dari presisi dan *recall* yang memberikan keseimbangan antara kedua metrik tersebut. Metrik ini sangat berguna untuk mengevaluasi model pada *dataset* yang tidak seimbang, di mana akurasi saja tidak cukup representatif untuk menggambarkan kinerja model secara keseluruhan **(Istiqomah, 2022)**

3. HASIL DAN PEMBAHASAN

3.1 Hasil Augmentasi Data

Visualisasi tahapan implementasi augmentasi data pada citra retina, dapat dilihat pada Gambar 7. Proses dimulai dengan citra yang telah melalui tahap prapemrosesan. Berbagai transformasi, seperti penyesuaian kecerahan dan kontras, *horizontal flip*, serta rotasi diterapkan secara sekuensial. Setiap langkah augmentasi menghasilkan citra baru yang bervariasi, namun tetap mempertahankan fitur penting DR.



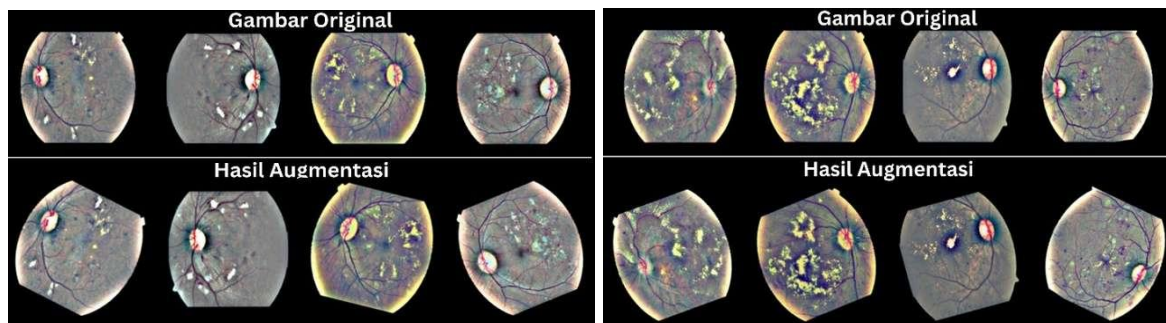
Gambar 7. Visualisasi Augmentasi Citra Retina Fundus

Tabel 4. Distribusi *Dataset* Setelah Augmentasi

Kelas	Train (Sebelum Augmentasi)	Train (Setelah Augmentasi)	Validation	Test
<i>No_DR</i>	1444	1444	180	181
<i>Mild</i>	296	1444	37	37
<i>Moderate</i>	799	1444	100	100
<i>Severe</i>	154	1444	20	19
<i>Prolifera_{te}_DR</i>	236	1444	29	30

Terlihat jelas dari Tabel 4. bahwa augmentasi data, khususnya melalui teknik augmentasi, berhasil menyeimbangkan jumlah sampel untuk setiap kelas dalam set data pelatihan menjadi 1444. Ini secara efektif mengatasi masalah *imbalance* data yang sering terjadi pada *dataset* medis, di mana kelas minoritas (seperti *Severe* dan *Prolifera_{te}_DR*) memiliki jumlah sampel yang jauh lebih sedikit. Dengan pemerataan ini, model pembelajaran mesin diharapkan dapat belajar dari representasi yang lebih seimbang dari setiap kondisi DR, memungkinkan identifikasi yang lebih akurat dari spektrum variasi yang lebih luas, dan mengurangi bias terhadap kelas mayoritas.

Dengan dataset yang seimbang, model diharapkan tidak hanya meningkatkan akurasi keseluruhan, tetapi yang lebih penting, menunjukkan peningkatan signifikan pada metrik *Recall* dan *F1-Score* untuk kelas minoritas seperti *Severe* dan *Prolifera_{te}_DR*. Peningkatan ini memastikan bahwa model memiliki sensitivitas tinggi dalam mendeteksi kasus penyakit yang parah, menjadikan sistem klasifikasi ini lebih andal dan relevan dalam aplikasi klinis.



Gambar 8. Validasi Visual Hasil Augmentasi

Pada Gambar 8. menunjukkan kelayakan data hasil augmentasi, dilakukan validasi oleh seorang ahli. Proses ini melibatkan *Dr. dr. Riani Erna, Sp.M(K)*, seorang dokter spesialis mata (*ophthalmologist*) konsultan dari Rumah Sakit Dr. Moh. Hoesin. Dalam validasi ini, sampel citra yang telah melalui proses augmentasi untuk setiap kelas keparahan (*Mild*, *Moderate*, *Severe*, dan *Prolifera_{te}_DR*) disajikan kepada ahli. Tujuannya adalah untuk mengonfirmasi bahwa ciri-ciri klinis utama seperti *microaneurysm* pada kelas *Mild*, *venous beading* atau *IRMA* (abnormalitas *mikrovaskuler intraretinal*) pada kelas *Severe*, serta *neovascularization* pada kelas *Prolifera_{te}_DR* tetap terlihat jelas, utuh, dan tidak terdistorsi. Hasil validasi dari ahli menunjukkan bahwa proses augmentasi yang diterapkan berhasil menjaga integritas citra medis, sehingga data yang digunakan untuk melatih model tetap memiliki kualitas yang valid secara klinis.

3.2 Eksperimen Skenario Augmentasi dan Pelatihan Model

Penelitian ini menguji beberapa skenario augmentasi data untuk mengevaluasi pengaruhnya terhadap performa model klasifikasi DR. Setiap skenario dirancang untuk mengetahui kombinasi teknik augmentasi mana yang paling efektif dalam membantu model mengenali tingkat keparahan DR. Evaluasi dilakukan menggunakan metrik *F1-score* karena metrik ini mempertimbangkan keseimbangan antara *precision* dan *recall*, yang penting terutama pada kasus dengan distribusi kelas yang tidak seimbang seperti pada *dataset APTOS 2019*.

Tabel 5. Hasil Pengujian Berdasarkan Skenario pada Data Uji

No	Metode Augmentasi	<i>No_DR</i> F1	<i>Mild</i> F1	<i>Moderate</i> F1	<i>Severe</i> F1	<i>Proliferate_DR</i> F1	F1 Score	Akurasi
1	-	0,8916	0,1429	0,7009	0,087	0,449	0,6863	0,7384
2	Rotasi, <i>Flip</i> Horizontal	0,9415	0,4932	0,67	0,4314	0,3333	0,7462	0,7466
3	Rotasi, <i>Flip</i> Horizontal, Kontras	0,9524	0,5312	0,7345	0,5263	0,5306	0,794	0,7984
4	Rotasi, <i>Flip</i> Horizontal, Kecerahan	0,9465	0,5116	0,6839	0,36	0,52	0,7659	0,7575
5	Rotasi, <i>Flip</i> Horizontal, Kecerahan, Kontras	0,9781	0,65	0,7411	0,4211	0,6415	0,8241	0,8256

Pada Tabel 5. skenario 5 menggunakan kombinasi augmentasi paling lengkap yakni rotasi, *flip horizontal*, kontras, dan kecerahan. Skenario ini menghasilkan akurasi tertinggi sebesar 82,56% dan *F1-score* sebesar 0,8241. Performa model dalam mendeteksi kelas *No_DR*, *Mild*, *Moderate*, dan *Proliferate_DR* juga meningkat secara konsisten, menjadikan skenario ini sebagai yang paling seimbang untuk klasifikasi menyeluruh.

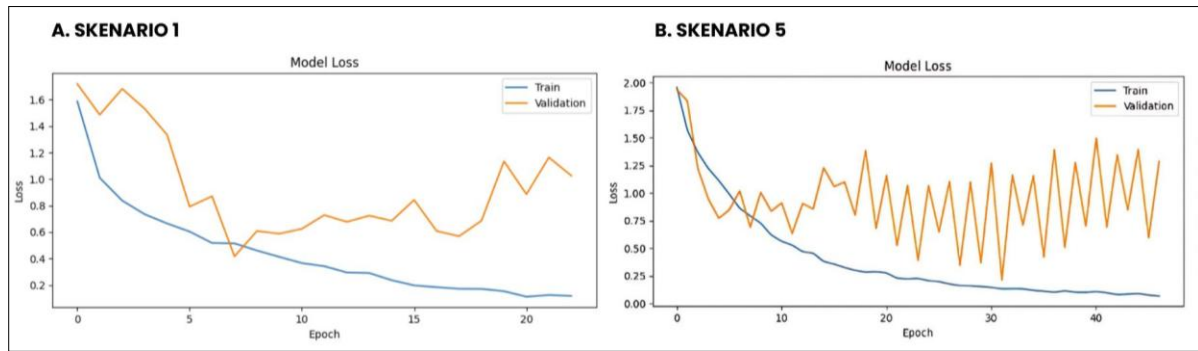
Sementara itu, skenario 3 menggunakan augmentasi tanpa kecerahan. Hasilnya menunjukkan peningkatan signifikan pada deteksi kelas *Severe*, dengan nilai *F1-score* tertinggi sebesar 0,5263. Hal ini menunjukkan bahwa adanya augmentasi kontras membantu model mengenali karakteristik khas dari DR tingkat lanjut yang cenderung memiliki kontras tinggi secara visual.

Berdasarkan hasil evaluasi, skenario 3 menunjukkan keunggulan dalam mendeteksi kasus *Severe* DR, namun skenario 5 memberikan performa terbaik secara keseluruhan dengan akurasi dan *F1-score* tertinggi pada sebagian besar kelas. Oleh karena itu, dalam penelitian ini, skenario 5 dipilih sebagai pendekatan utama karena mampu memberikan hasil yang paling seimbang dan optimal untuk klasifikasi tingkat keparahan DR.

3.3 Analisis Pelatihan Arsitektur Model

Pelatihan model dilakukan menggunakan GPU *NVIDIA Tesla P100* yang disediakan oleh *platform Kaggle*. Total waktu yang dibutuhkan untuk pelatihan model skenario 5 adalah 1222 detik atau sekitar 20 menit. Waktu ini tergolong efisien mengingat ukuran *dataset* yang telah diperbanyak variasinya menggunakan teknik augmentasi.

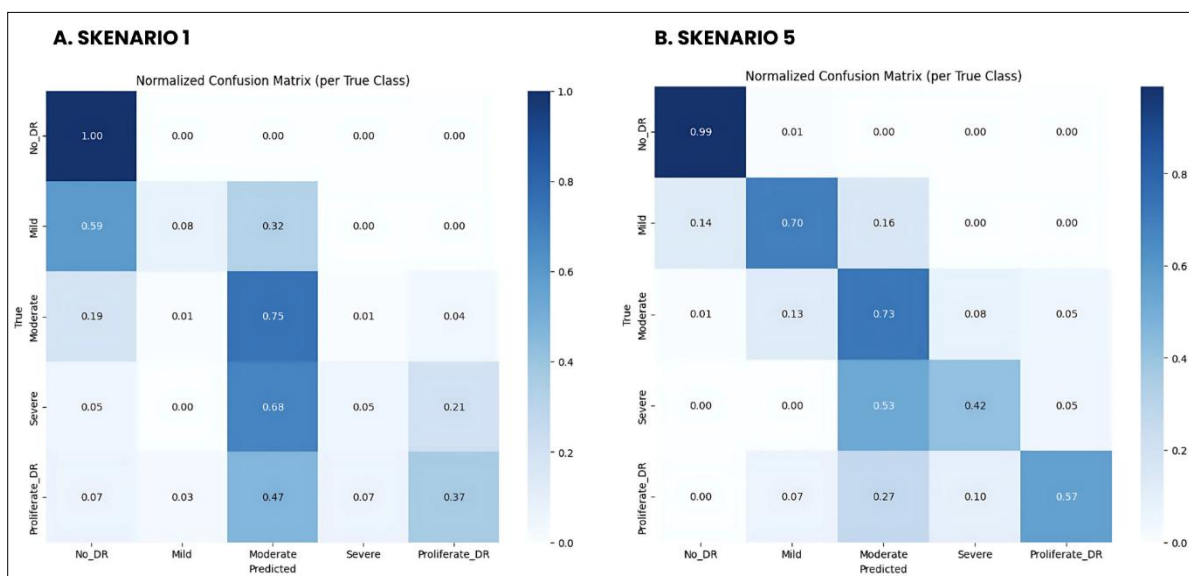
Pendekatan Augmentasi Citra Fundus pada Model EfficientNet untuk Klasifikasi Tingkat Keparahan Retinopati Diabetik dengan *Dataset* Tidak Seimbang



Gambar 9. Evaluasi *Loss* Model pada Skenario 1 dan 5

Gambar 9. menunjukkan perbandingan kurva *loss* antara Skenario 1 (tanpa augmentasi) yang berhenti pada *epoch* ke-8 dan Skenario 5 yang berhenti pada *epoch* ke-32 dengan mekanisme *early stopping* (*patience* 15). Skenario 1 menunjukkan *training loss* yang menurun konsisten dari 1,6 ke 0,4 mengindikasikan pembelajaran terbatas pada data original. Sebaliknya, Skenario 5 menunjukkan *training loss* yang turun lebih dalam dari 2,0 ke mendekati 0,2, namun *validation loss* mengalami fluktuasi signifikan setelah *epoch* ke-10 akibat strategi regularisasi dan kompleksitas data hasil augmentasi. Pola fluktuasi *validation loss* pada Skenario 5 mencerminkan proses adaptasi model terhadap variasi data yang lebih kaya, yang meskipun menyebabkan variabilitas antar *epoch*, pada akhirnya berkontribusi pada peningkatan kemampuan generalisasi dan performa klasifikasi yang lebih seimbang.

Model menggunakan *EfficientNet-B0* sebagai *feature extractor* yang dikombinasikan dengan *custom classification head*. *EfficientNet-B0* dipilih karena efisiensi *compound scaling*, penggunaan *Mobile Inverted Bottleneck Convolution* (MBConv) yang memisahkan operasi spasial dan *channel-wise* untuk efisiensi komputasi, dan *Squeeze-and-Excitation* (SE) *blocks* yang memberikan bobot adaptif pada channel berdasarkan kepentingan fitur. Dengan jumlah 4.049.571 parameter yang termasuk kecil, arsitektur ini sesuai untuk pelatihan dengan sumber daya komputasi terbatas. Arsitektur ini menghasilkan *feature map* berukuran $7 \times 7 \times 1280$ yang kemudian diproses melalui *Global Average Pooling*, regulasi, dan *dense layers* untuk klasifikasi akhir. Kombinasi arsitektur ini memungkinkan ekstraksi fitur ciri seperti *mikroaneurisma*, *pendarahan*, dan *neovaskularisasi* secara efektif.



Gambar 10. *Confusion Matrix* Data Uji Skenario 1 dan 5

Gambar 10. menyajikan dua *confusion matrix* yang telah dinormalisasi pada model skenario 1 dan 5 menggunakan data pengujian.. Model skenario 1 menunjukkan akurasi sangat tinggi pada kelas *No_DR* dan *Moderate*, menandakan dominasi prediksi pada kelas-kelas ini, namun performanya terbatas pada kelas-kelas lain seperti *Mild*, *Severe*, dan *Proliferate_DR* yang merupakan kelas minoritas. Sebaliknya, Model Skenario 5, yang menggunakan teknik augmentasi data, secara umum menunjukkan peningkatan yang signifikan dalam memprediksi setiap kelas, khususnya pada kelas *Mild* dan *Severe* yang sebelumnya sulit diidentifikasi.

Secara keseluruhan, arsitektur yang diusulkan menunjukkan performa yang baik dalam hal akurasi dan efisiensi komputasi. Meskipun *validation loss* tidak sepenuhnya stabil, model mampu menangkap fitur penting dari citra fundus. Hal ini membuktikan bahwa dengan augmentasi data yang tepat, strategi regularisasi dasar, dan pemanfaatan *EfficientNet-B0*, model dapat menghasilkan klasifikasi yang akurat sekaligus efisien untuk diterapkan di lingkungan terbatas sumber daya.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan model klasifikasi DR berbasis *EfficientNet-B0* yang dikombinasikan dengan strategi augmentasi data terarah untuk mengatasi ketidakseimbangan kelas pada *dataset* APTOS 2019. Evaluasi menunjukkan perbedaan performa yang signifikan: model tanpa augmentasi (skenario 1) hanya mencapai akurasi 73,84% dan *F1-score* 0,6863, sedangkan model dengan augmentasi lengkap (skenario 5) mampu meningkatkan akurasi hingga 82,56% dengan *F1-score* 0,8241. Performa pada kelas minoritas juga mengalami perbaikan, ditunjukkan dengan peningkatan *F1-score* kelas tingkat *Mild* dari 0,1429 menjadi 0,65, tingkat *Severe* dari 0,087 menjadi 0,4211, dan tingkat *Proliferative DR* dari 0,449 menjadi 0,6415. Temuan ini menegaskan bahwa ketika menghadapi keterbatasan pada jumlah dan keragaman data, penerapan augmentasi yang tepat dapat menjadi strategi yang cukup efektif. Dengan menyamakan distribusi data dan mempertahankan fitur ciri penting, model dapat dilatih secara adil dan menghasilkan klasifikasi yang lebih optimal bagi seluruh kelas. Selain itu, penggunaan teknik *padding* dalam proses augmentasi juga terbukti penting untuk menjaga agar informasi dari data asli tetap dipertahankan saat proses augmentasi, sehingga hasil augmentasi tidak menghilangkan detail penting pada citra fundus. Untuk penelitian selanjutnya, disarankan penelitian lanjutan yang dapat difokuskan secara khusus untuk mengeksplorasi pengaruh teknik augmentasi tertentu, seperti variasi penyesuaian kontras, terhadap hasil klasifikasi model. Hal ini penting untuk memahami sejauh mana manipulasi visual tersebut dapat meningkatkan atau justru mengganggu kemampuan model dalam mengenali ciri khas dari masing-masing kelas penyakit.

UCAPAN TERIMA KASIH

Ucapan terima kasih juga ditujukan kepada *Dr. dr. Riani Erna, Sp.M(K)* atas kesediaannya meluangkan waktu dan keahliannya untuk melakukan validasi terhadap citra fundus hasil augmentasi yang digunakan dalam penelitian ini. Penulis juga berterima kasih kepada keluarga dan teman-teman atas dukungan dan semangat yang diberikan sepanjang penyusunan penelitian ini.

DAFTAR RUJUKAN

- Abdullah, S. H., Magdalena, R., & Fu'adah, R. Y. N. (2022). Klasifikasi Diabetic Retinopathy Berbasis Pengolahan Citra Fundus Dan Deep Learling. *Journal Of Electrical And System Control Engineering*, 5(2), 84–90. <https://doi.org/10.31289/jesce.v5i2.5659>
- Chilukoti, S. venkatesh, Maida, A. S., & Hei, X. (2022). Diabetic Retinopathy Detection using Transfer Learning from Pre-trained Convolutional Neural Network Models. Dalam *IEEE Journal Of Biomedical And Health Informatics: Vol. XX* (hlm. 1). <https://doi.org/10.36227/techrxiv.18515357.v1>
- Duque-Arribas, C., Tabernero, H. M., Montes, D., Caballero, J. A., & Galceran, E. (2025). A neural network approach to determining photometric metallicities of M-type dwarf stars. *Astronomy & Astrophysics*, 698, L12. <https://doi.org/10.1051/0004-6361/202554722>
- Eimoufidi, A., & Ammoun, H. (2021). *EfficientNetB3 Architecture for Diabetic Retinopathy Assessment using Fundus Images*.
- Guerbai, Y., Bouaraki, W., Saibi, A., Bida, M., & Halimouche, R. (2024). Deep learning techniques for diabetic retinopathy classification: a focus on VGG16 and EfficientNetB0. *South Florida Journal of Development*, 5(10), e4517. <https://doi.org/10.46932/sfjdv5n10-027>
- International Diabetes Federation. (2024). *IDF Diabetes Atlas 11th Edition*. <https://idf.org/about-diabetes/diabetes-facts-figures/>
- Istiqomah, A. (2022). *Perbandingan Kinerja Deep Learning Alexnet Dan Efficientnet Dengan Augmentasi Data Untuk Identifikasi Kupu-Kupu*. <http://digilib.unila.ac.id/66216/3/FILE%20SKRIPSI%20FULL%20TANPA%20BAB%20PEMBAHASAN.pdf>
- Karthik, Maggie, & Sohier, D. (2019). *APTOS 2019 Blindness Detection*. Kaggle. <https://kaggle.com/competitions/aptos2019-blindness-detection>
- Naeeni Davarani, M., Arian Darestani, A., Guillen Cañas, V., Azimi, H., Havadaragh, S. H., Hashemi, H., & Harirchian, M. H. (2024). (Archi image) Efficient segmentation of active and inactive plaques in FLAIR-images using DeepLabV3Plus SE with efficientnetb0 backbone in multiple sclerosis. *Scientific Reports*, 14(1), 16304. <https://doi.org/10.1038/s41598-024-67130-6>
- Pak, A., Ziyaden, A., Tukeshev, K., Jaxylykova, A., & Abdullina, D. (2020). Comparative analysis of deep learning methods of detection of diabetic retinopathy. *Cogent Engineering*, 7(1). <https://doi.org/10.1080/23311916.2020.1805144>

- Sanjaya, J., & Ayub, M. (2020). Augmentasi Data Pengenalan Citra Mobil Menggunakan Pendekatan Random Crop, Rotate, dan Mixup. *Jurnal Teknik Informatika dan Sistem Informasi*, 6(2). <https://doi.org/10.28932/jutisi.v6i2.2688>
- Santoso, R. (2024). *Augmentasi Data pada Prasasti Logam untuk Deteksi Aksara Kawi*.
- Syurfi, I. (2021). *Penerapan Deep Learning Dengan Convolutional Neural Network Untuk Klasifikasi Citra Diabetic Dengan Arsitektur Efficientnet-B7*. <https://repository.uin-suska.ac.id/57860/1/BAB%20I%2C%20II%2C%20III%2C%20dan%20VI.pdf>
- Tan, M., & Le, Q. V. (2019). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. <http://arxiv.org/abs/1905.11946>
- Yi, S. L., Yang, X. L., Wang, T. W., She, F. R., Xiong, X., & He, J. F. (2021). Diabetic retinopathy diagnosis based on RA-efficientnet. *Applied Sciences (Switzerland)*, 11(22). <https://doi.org/10.3390/app112211035>
- Zahir, M., & Adi Saputra, R. (2024). *Deteksi Penyakit Retinopati Diabetes Menggunakan Citra Mata Dengan Implementasi Deep Learning CNN* (Vol. 18, Nomor 1).