

## DILATION LAYER MENGGUNAKAN EFFICIENTNETV2 UNTUK MENDETEKSI PENYAKIT RETINOPATI DIABETIK

Aiman Mukhlisah<sup>\*1</sup>, Ema Utami<sup>2</sup>

<sup>1,2</sup>Universitas Amikom Yogyakarta, Yogyakarta

Email: <sup>1</sup>Aimanmukhlisah@students.amikom.ac.id, <sup>2</sup>ema.u@amikom.ac.id

<sup>\*</sup>Penulis Korespondensi

(Naskah masuk: 21 Juli 2025, diterima untuk diterbitkan: 16 Desember 2025)

### Abstrak

*Diabetic Retinopathy (DR)* merupakan salah satu komplikasi serius akibat diabetes melitus yang dapat menyebabkan kebutaan jika tidak dideteksi dan ditangani secara dini. Deteksi otomatis berbasis citra fundus retina telah menjadi fokus utama dalam penelitian *deep learning*. Namun, fitur-fitur kecil seperti *microaneurysma* dan lesi halus seringkali sulit dikenali, terutama pada citra dengan kualitas rendah. Penelitian ini mengusulkan modifikasi arsitektur EfficientNetV2 dengan menambahkan *dilation layer* guna meningkatkan kemampuan model dalam menangkap fitur kontekstual dengan cakupan spasial yang lebih luas tanpa menambah kompleksitas secara signifikan. *Dilated convolution* memungkinkan perluasan *receptive field* tanpa mengurangi resolusi spasial, sehingga lebih efektif dalam mendeteksi fitur-fitur kritis pada *DR*. Pengujian dilakukan menggunakan dataset *APROS 2019 Blindness Detection* yang menyediakan citra fundus retina beranotasi untuk lima tingkat keparahan *DR*. Model dilatih menggunakan *optimizer Adam* dengan *learning rate* sebesar 0,001 untuk mencapai konvergensi yang optimal. Evaluasi dilakukan terhadap beberapa varian EfficientNetV2, yaitu *B0*, *B1*, *B2*, *B3*, dan *S*. Hasil eksperimen menunjukkan bahwa model EfficientNetV2-B0 dan B1 memberikan akurasi terbaik, masing-masing sebesar 97,14%. Di antara keduanya, EfficientNetV2-B0 menghasilkan nilai presisi dan *f1-score* yang lebih tinggi. Temuan ini menunjukkan bahwa integrasi *dilation layer* pada arsitektur EfficientNetV2 berpotensi meningkatkan performa sistem klasifikasi *DR* otomatis secara signifikan dibandingkan model *baseline* tanpa modifikasi.

**Kata kunci:** *Diabetic Retinopathy, EfficientNetV2, dilated convolution, klasifikasi citra, CNN*

## DILATION LAYERS WITH EFFICIENTNETV2 FOR DIABETIC RETINOPATHY

### Abstract

*Diabetic Retinopathy (DR)* is a serious complication of diabetes mellitus that can lead to blindness if not detected and treated early. Automated detection based on retinal fundus images has become a primary focus in deep learning research. However, small features such as microaneurysms and subtle lesions are often difficult to identify, particularly in low-quality images. This study proposes a modification to the EfficientNetV2 architecture by incorporating dilation layers to enhance the model's ability to capture contextual features over a broader spatial range without significantly increasing complexity. Dilated convolutions allow the expansion of the receptive field without reducing spatial resolution, making them more effective in detecting critical features associated with *DR*. The evaluation was conducted using the *APROS 2019 Blindness Detection* dataset, which provides annotated retinal fundus images across five levels of *DR* severity. The model was trained using the Adam optimizer with a learning rate of 0.001 to achieve optimal convergence. The study assessed several variants of EfficientNetV2 - *B0*, *B1*, *B2*, *B3*, and *S*. Experimental results indicate that the EfficientNetV2-B0 and *B1* models achieved the highest accuracy, each reaching 97.14%. Among these, EfficientNetV2-B0 yielded higher precision and *F1-score* values. These findings suggest that integrating dilation layers into the EfficientNetV2 architecture can significantly enhance the performance of automated *DR* classification systems compared to the unmodified baseline models.

**Keywords:** *Diabetic Retinopathy, EfficientNetV2, Dilated Convolution, Image Classification, CNN*

### 1. PENDAHULUAN

*Diabetic Retinopathy (DR)* merupakan kondisi medis yang dapat merusak pembuluh darah di retina atau menyebabkan sejumlah luka pada mata yang bisa

merusak retina, sehingga dapat mengakibatkan kebutaan jika tidak ditangani dengan benar. Kondisi ini terjadi akibat tingginya kadar gula darah yang dapat memengaruhi struktur dan fungsi retina (Qummar et al., 2019). Asia menyumbang lebih dari

60% populasi diabetes global pada tahun 2021. Indonesia berada di peringkat ke-5 dengan jumlah penderita diabetes yaitu 19,5 juta jiwa (Ditjen P2P, 2024)(Yazid, 2021). Berdasarkan penelitian *The Jogjakarta Eye Diabetic Study in Community* di Yogyakarta, didapatkan kejadian penyakit DR sebesar 46,1% pada orang dewasa, 26,3%nya sudah mengancam penglihatan. Ini artinya, 1 dari 2 orang diabetes menderita retinopati diabetika.

Deteksi dini DR sangat penting untuk dilakukan karena semakin cepat penyakit ditemukan, semakin besar peluang kesembuhan. Salah satu arsitektur *deep learning* yang sering digunakan untuk pengenalan dan identifikasi gambar yakni *Convolutional Neural Network (CNN)* (Citra et al., 2023) yang diterapkan untuk mengklasifikasikan DR (Jordi de la Torre, 2020) karena kemampuannya dalam mengekstraksi fitur-fitur kompleks, khususnya pengenalan pola-pola tersembunyi (Zhang et al., 2019). Ada berbagai jenis arsitektur CNN yang dapat digunakan dalam penelitian, seperti AlexNet, ResNet, VGG, DenseNet, MobileNetV2 Efficientnet dan EfficientnetV2 (Yudistira et al., 2020).

Beberapa penelitian yang telah dilakukan untuk mengklasifikasi DR dan pengembangannya (El-Ateif & Idri, 2022; Phridviraj et al., 2023; Dewi et al., 2021) diantaranya penelitian yang dilakukan oleh (Qomariah et al., 2019) yang menggunakan lapisan *fully connected CNN* sebagai fitur ekstraksi dan menggunakan *Support Vektor Machine (SVM)* sebagai fitur klasifikasi. Dari hasil eksperimen, didapatkan akurasi tertinggi mencapai 95,83% dan 95,24%. Pada tahun 2023, (Tummala et al., 2023) Melakukan deteksi citra gambar untuk klasifikasi DR menggunakan model *EfficientnetV2* dengan teknik ensemble pada EfficientnetV2-S, L dan M. Uji data yang dilakukan menghasilkan nilai 75% dengan AUC sebesar 74,9% dari teknik *ensemble*. Penelitian (RIZAL et al., 2020) juga melakukan deteksi DR menggunakan metode Efficientnet dimana penelitian ini menghasilkan akurasi sebesar 79%.

Selain itu pada 2024, (Muntas et al., 2025) melakukan Klasifikasi *Leukemia Limfoblastik Akut* dengan Integrasi Model Pralatih dan *Grid Search* menggunakan efficientnetV2-S, Hasil eksperimen pelatihan menunjukkan akurasi antara 98,89% hingga 99,80%. Penelitian (Karthik et al., 2022) menggunakan model EfficientNetV2 untuk mengklasifikasikan gambar penyakit kulit ke dalam empat kategori. Model ini menunjukkan performa yang baik dengan akurasi pengujian sebesar 84,70%. Selain itu, (Ye et al., 2022) juga melakukan penelitian menggunakan arsitektur EfficientNetV2 dan mengembangkan versi baru bernama PDRNet. Hasil menunjukkan bahwa EfficientNetV2 memiliki akurasi sebesar 98,53%, sementara PDRNet mencapai 99,32%, meningkat menjadi 99,56% setelah *fine-tuning*. Namun, karena kompleksitas arsitektur dengan jumlah lapisan dan operasi konvolusi yang tinggi, PDRNet membutuhkan lebih

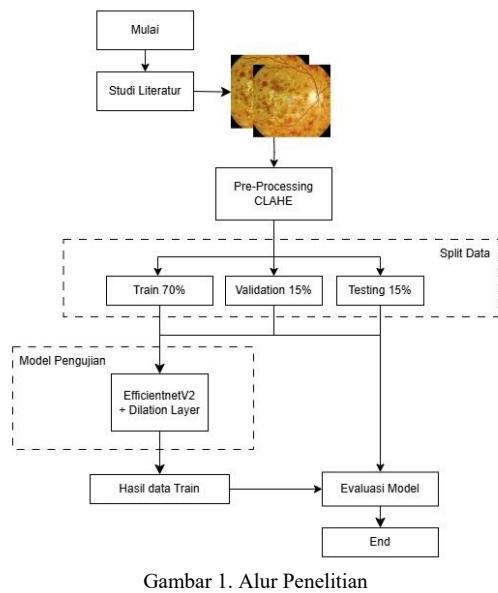
banyak waktu pelatihan, sehingga kurang efisien dibandingkan EfficientNetV2.

Meskipun teknologi *deep learning* telah membawa kemajuan signifikan dalam klasifikasi *Diabetic Retinopathy (DR)* (Fu et al., 2024), Namun demikian, tantangan yang terjadi yaitu keterbatasan dalam menangkap fitur penting yang bersifat halus dan tersebar, seperti *mikroaneurisma* dan lesi kecil, yang merupakan penanda awal DR. Fitur-fitur ini sering kali tidak terdeteksi secara optimal oleh jaringan konvolusional standar karena keterbatasan *receptive field* (Hu et al., 2024). Akibatnya, informasi kontekstual yang krusial dapat hilang. Kondisi ini diperparah oleh kualitas citra fundus retina yang sering kali terpengaruh oleh variasi pencahayaan (Roy et al., 2022). Masalah tersebut dapat memberikan pengaruh yang besar terhadap ketepatan hasil analisis.

Penelitian ini berkontribusi dengan mengintegrasikan *dilation layer* ke dalam arsitektur EfficientNetV2, yang dikenal efisien dan kuat sebagai *feature extractor*, serta menggabungkan metode CLAHE sebagai tahapan *preprocessing* untuk meningkatkan kualitas citra. Dengan menggunakan filter konvolusional ber-dilatasi, jaringan memperoleh kemampuan untuk memperluas *receptive field* secara signifikan tanpa menambah jumlah parameter atau mengurangi resolusi spasial (Liu et al., 2022). Sejumlah penelitian terdahulu telah membuktikan bahwa penggunaan *dilation layer* efektif dalam meningkatkan performa klasifikasi, sebagaimana ditunjukkan dalam studi (Shang et al., 2024) (Madarapu et al., 2024) Sebagai contoh (Das & Saha, 2022) melakukan komparasi metode *Context-Gated Dilated Convolution Network (CGDCN)* dengan ResNet50 dan akurasi tertinggi dicapai CGDCN 98,178%. Diharapkan penelitian ini membantu dokter dalam mendiagnosis penyakit DR pada penderita diabetes secara lebih akurat, serta mengurangi beban pada sistem kesehatan, sehingga memberikan dampak signifikan dalam pengembangan penelitian.

## 2. METODE PENELITIAN

Penelitian ini merancang *system* untuk menklasifikasikan *Diabetic Retinopathy* berdasarkan citra retina menggunakan EfficientnetV2. Metode yang digunakan pada penelitian ini mengusulkan metode EfficientnetV2 sebagai *feature extraction* yang dibekukan bobotnya, kemudian ditambahkan beberapa lapisan untuk mendukung proses klasifikasi. penerapan lapisan konvolusional dengan teknik *dilation*, yang menggunakan filter berukuran  $3 \times 3$  dan *dilation rate* sebesar 2 untuk memperluas area tangkap (*receptive field*) secara efektif tanpa menambah jumlah parameter atau menurunkan resolusi spasial. Alur pada penelitian ini dijelaskan melalui bagan alur penelitian pada gambar 1.



Gambar 1. Alur Penelitian

## 2.1. Studi Literatur

Dasar informasi penelitian ini diambil dari jurnal-jurnal yang menelaah penyakit diabetes. Banyak peneliti memakai data terkait *diabetic retinopathy* untuk mengkaji persoalan, memprediksi kondisi pasien, dan mengevaluasi performa algoritma guna menemukan metode terbaik.

## 2.2. Dataset



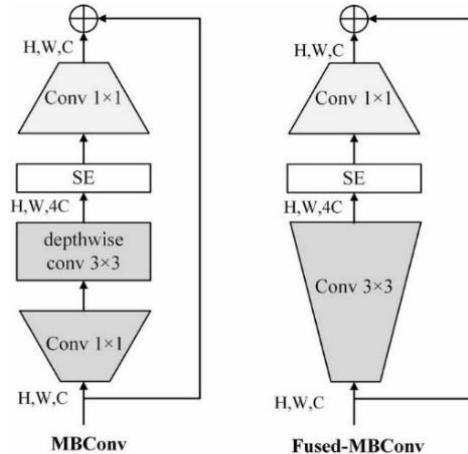
Gambar 2. Dataset DR

Penggunaan data penelitian ini menggunakan data public didapatkan dari situs web Kaggle.com yang memungkinkan akses ke kumpulan data dalam berbagai ukuran. Citra digital retina dikontribusikan oleh *APTOs 2019 Blindness Detection*. Dataset gambar retina *diabetic retinopathy* adalah gambar retina dari penderita diabetes. Citra retina pada penelitian ini telah digunakan pada penelitian sebelumnya (Raiaan et al., 2023), dan terbagi dalam 5 kelas yaitu *Mild*, *Moderate*, *No\_DR*, *Proliferate\_DR*, dan *Severe*.

## 2.3. Preprocessing

Pada tahap ini proses pemodelan dilakukan untuk membangun model klasifikasi menggunakan EfficientNetV2. EfficientNetV2 merupakan generasi

terbaru dari keluarga EfficientNet dirilis pada tahun 2021 dan dirancang untuk pengolahan gambar (Tan & Le, 2019) yang memiliki keunggulan signifikan dalam klasifikasi gambar. Dengan efisiensi pelatihan mencapai 11 kali lebih cepat serta ukuran model yang 6,8 kali lebih kecil dibandingkan pendekatan sebelumnya. EfficientNetV2 menggunakan blok *Mobile Inverted Bottleneck Convolution* (MBConv) dengan rasio ekspansi lebih kecil serta *Fused-MBConv* yang ditambahkan di lapisan awal. Selain itu, digunakan kernel berukuran lebih kecil, yaitu  $3 \times 3$ , di beberapa lapisan. Gambar 3 menampilkan struktur MBConv dan *Fused-MBConv*, yang memiliki perbedaan pada lapisan yang digunakan. MBConv menggunakan *depthwise conv*  $3 \times 3$  dan *Conv*  $1 \times 1$ , sedangkan *Fused-MBConv* hanya memakai *Conv*  $3 \times 3$ . Rasio ekspansi yang lebih kecil memberikan manfaat berupa pengurangan overhead pada akses memori. (Tan & Le, 2021)



Gambar 3. Struktur MBConv dan Fused-MBConv (Mei et al., 2024)

Penerapan *dilation layer* yang merupakan salah satu strategi efektif untuk memperluas *receptive field* tanpa meningkatkan jumlah parameter atau kehilangan resolusi spasial. Pada penelitian ini, *dilation layer* diintegrasikan setelah *backbone* EfficientNetV2 yang dibekukan, dengan menggunakan lapisan konvolusi 2D berukuran kernel  $3 \times 3$  dan *dilation rate* sebesar 2. Penggunaan teknik ini memungkinkan model untuk menangkap informasi kontekstual yang lebih luas dari fitur spasial gambar retina tanpa memperbesar ukuran filter atau kedalaman jaringan. Integrasi *dilated convolution* pada tahap awal setelah ekstraksi fitur dasar dari EfficientNetV2 meningkatkan sensitivitas model terhadap pola yang lebih kompleks dan beresolusi rendah, tanpa mengorbankan efisiensi komputasi.

## 2.4. Evaluasi

Evaluasi performa model dilakukan dengan memanfaatkan metrik *accuracy*, *precision*, *recall*, dan *F1-score* sebagaimana dirumuskan dalam Persamaan (1) digunakan untuk mengkalkulasi

*accuracy* sebagai rasio antara jumlah prediksi yang tepat terhadap total data, persamaan (2) digunakan mendefinisikan *precision* sebagai proporsi prediksi positif yang benar dari seluruh prediksi positif, Persamaan (3) yaitu *recall* digunakan untuk mengukur sejauh mana model berhasil mengidentifikasi seluruh data positif. Adapun Persamaan (4) memberikan nilai *F1-score*, yang merupakan harmonisasi antara *precision* dan *recall*, dan sangat berguna dalam konteks data yang tidak seimbang. Analisis lebih lanjut terhadap kesalahan prediksi dilakukan melalui *confusion matrix*, sedangkan kemampuan model dalam membedakan antar kelas dievaluasi menggunakan kurva ROC dan nilai AUC. Nilai *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN) dihitung guna memperoleh pemahaman yang lebih komprehensif terhadap performa model (Majumdar et al., 2023).

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

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

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

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

### 3. HASIL

Eksperimen ini dilakukan dengan menerapkan beberapa varian arsitektur EfficientNetV2, yakni B0, B1, B2, B3, dan S masing masing menggunakan *dilation layer*, dengan tujuan untuk mengevaluasi performa masing-masing model dalam mendeteksi *diabetic retinopathy*. Melalui Penerapan model dilakukan perbandingan terhadap metrik evaluasi seperti *akurasi*, *presisi*, *recall*, dan *F1-score* guna mengidentifikasi model mana yang paling optimal dalam mengklasifikasikan *diabetic retinopathy* saat dikombinasikan dengan teknik *dilated convolution*.

#### 3.1. Skenario Pengujian

Nomor dan judul tabel dituliskan di tengah kolom (*center alignment*). Implementasi penelitian ini dilakukan menggunakan menggunakan 25 *epoch* dengan ukuran *batch size* 32 dengan Nilai *learning rate* sebesar 1e-4 (0.0001) yang dikonfigurasikan menggunakan *Adam optimizer* untuk memastikan proses pembaruan bobot berlangsung secara stabil selama pelatihan. Proses pelatihan dilakukan dengan menggunakan *dilation layer* dan *dilation rate* sebesar 2 pada efficientnetv2 dan hasil evaluasinya ditampilkan pada Tabel 1.

Berdasarkan hasil evaluasi pada Tabel 1 menunjukkan bahwa penerapan *dilation layer* 2x2 pada arsitektur EfficientNetV2 memberikan performa yang kompetitif dalam melakukan klasifikasi DR. Model EfficientNetV2-B0 dan B1 menunjukkan

performa tertinggi, dengan akurasi sebesar 97,14%, nilai presisi masing-masing 97,26% dan 97,18%, serta nilai *recall* dan *F1-Score* sebesar 97,14%. hasil ini berindikasi bahwa penggunaan *dilation convolution* berkontribusi secara signifikan dalam memperluas *receptive field* tanpa meningkatkan jumlah parameter secara drastis, sehingga model mampu menangkap pola lokal dan global pada citra retina dengan lebih baik.

Tabel 1. Evaluasi EfficientnetV2

Efficientnetv2	Akurasi	Presisi	Recall	Skor F1
<b>B0</b>	97,14%	97,26%	97,14%	97,14%
<b>B1</b>	97,14%	97,18%	97,14%	97,13%
<b>B2</b>	95,83%	96,00%	95,83%	95,83%
<b>B3</b>	94,66%	94,77%	94,66%	94,62%
<b>S</b>	94,92%	95,40%	94,92%	94,91%

Model EfficientNetV2-B2, menghasilkan sedikit penurunan pada akurasi 95,83% dengan presisi tinggi mencapai 96,00%, model ini cenderung menghasilkan prediksi yang sangat akurat meskipun terjadi sedikit penurunan sensitivitas. Performa model B3 dan S dengan akurasi masing-masing 94,66% dan 94,92%, berada pada kategori yang masih cukup baik. Integrasi *dilation layer* ke dalam *backbone* EfficientNetV2 berpotensi meningkatkan kemampuan model dalam mengenali fitur penting pada gambar fundus retina, yang sangat krusial dalam mendeteksi kondisi DR secara dini dan akurat.

#### 3.2. Visualisasi Hasil

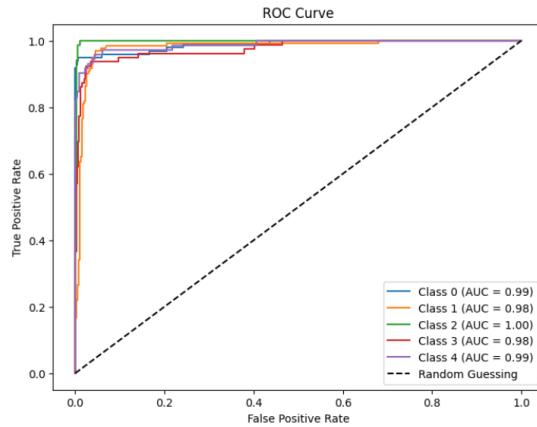
Berdasarkan hasil klasifikasi, model EfficientNetV2-B0 menunjukkan performa terbaik. Hal ini ditunjukkan pada Tabel 2 yang memuat rincian akurasi klasifikasi untuk masing-masing kelas.

	precision	recall	f1-score	support
Mild	1.00	0.89	0.94	98
Moderate	0.84	0.98	0.90	132
No_DR	0.96	1.00	0.98	151
Proliferate_DR	0.93	0.82	0.87	79
Severe	0.97	0.86	0.91	72

Gambar 4. Performa Model EfficientnetV2-B0 Berdasarkan Presisi, Recall, dan Skor F1

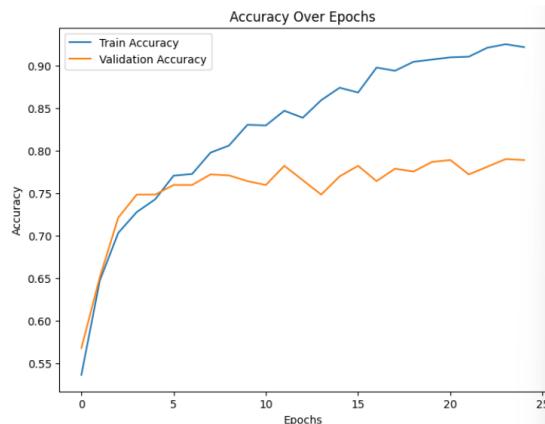
Model EfficientNetV2-B0 berhasil menunjukkan performa yang kuat dalam mengklasifikasikan berbagai tingkat keparahan *diabetic retinopathy*, seperti yang tergambar dalam metrik evaluasi pada gambar 4. Untuk kelas No DR, model mencatat *recall* paling tinggi sebesar 100% dan *precision* sebesar 96%, menghasilkan *F1-score* tertinggi yakni 98%, yang mengindikasikan kemampuan deteksi yang sangat andal pada kondisi retina normal. Kelas Mild dengan *precision* mencapai 100%, meskipun *recall*-nya sedikit menurun di angka

89%, sehingga *F1-score* tercatat sebesar 94%. Pada kelas *Moderate* menunjukkan kemampuan deteksi yang baik dengan *recall* 98% dan *precision* 84%, menghasilkan *F1-score* sebesar 90%. Untuk kelas *Severe* dan *PDR*, *precision* yang tinggi masing-masing sebesar 97% dan 93% menunjukkan ketepatan model dalam mengidentifikasi kedua kelas ini, meskipun *recall* berada di angka 86% yang menandakan masih ada beberapa kasus yang tidak terdeteksi. Nilai *F1-score* untuk kedua kelas tersebut tetap berada pada tingkat yang baik, yaitu 91% untuk *Severe* dan 87% untuk *PDR*. Hasil ini mencerminkan bahwa model memiliki ketajaman yang tinggi dalam mengenali ciri khas tiap kategori retinopati diabetik dan berpotensi sebagai alat pendukung diagnosis yang efektif.

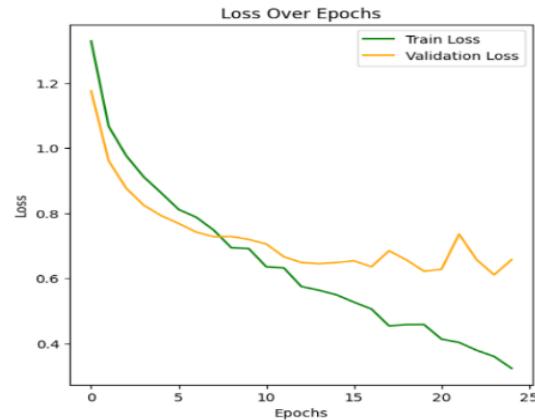


Gambar 5. Grafik Kurva ROC

Gambar 5 menunjukkan kurva ROC (*Receiver Operating Characteristic*) untuk masing-masing kelas target. Dari grafik tersebut terlihat bahwa seluruh kelas memiliki nilai *Area Under Curve* (AUC) yang sangat tinggi, yakni antara 0,98 hingga 1,00. Hal ini menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat baik, dengan tingkat deteksi positif yang tinggi dan tingkat kesalahan klasifikasi yang sangat rendah untuk semua kelas. Posisi kurva yang dekat dengan sudut kiri atas grafik menunjukkan performa klasifikasi yang sangat optimal, jauh dari kurva acak (*random guessing*) yang ditunjukkan oleh garis putus-putus.



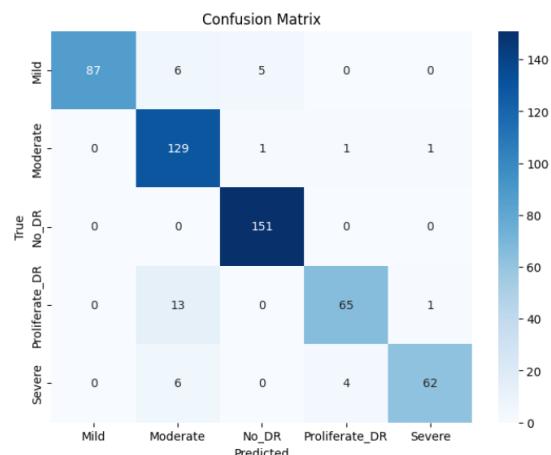
Gambar 6. Grafik Akurasi Model



Gambar 7. Grafik Kerugian Model

Performa model klasifikasi dalam penelitian ini ditunjukkan pada gambar 6 dimana terdapat perubahan pada data pelatihan dan validasi selama proses pelatihan. akurasi data pelatihan menunjukkan peningkatan yang konsisten dan signifikan, mencapai nilai mendekati 97% pada akhir pelatihan. Sementara itu, akurasi validasi mengalami peningkatan pada *epoch* awal namun cenderung stagnan dan sedikit fluktuatif pada kisaran 77% hingga 80% setelah epoch ke-10. Pola ini mengindikasikan bahwa model belajar cukup baik terhadap data pelatihan, meskipun terdapat *overfitting* ringan karena performa validasi tidak mengikuti peningkatan performa pelatihan secara proporsional.

Gambar 7 menggambarkan nilai *loss* pada data pelatihan dan validasi sepanjang proses training. *Loss* pada data pelatihan terus menurun secara bertahap, yang menunjukkan bahwa model semakin mampu meminimalkan kesalahan dalam memprediksi data pelatihan. Sebaliknya, *loss* pada data validasi mengalami penurunan hanya di awal *epoch*, namun setelahnya cenderung tidak stabil dan mengalami fluktuasi. Hal ini kembali menguatkan indikasi bahwa model mengalami *overfitting* ringan, di mana model terlalu menyesuaikan diri dengan data pelatihan sehingga performanya terhadap data baru (validasi) tidak mengalami peningkatan yang sepadan.



Gambar 8. Grafik Confusion Matrix

Berdasarkan confusion matrix yang ditampilkan pada gambar 8 distribusi prediksi model terhadap lima kelas menunjukkan dominasi prediksi yang benar disepanjang diagonal utama, yang merepresentasikan kecocokan antara label asli dan hasil prediksi. Pada kelas *Mild*, terdapat 87 gambar yang berhasil diprediksi dengan benar, meskipun masih ada kesalahan klasifikasi ke kelas *Moderate* sebanyak 6 gambar dan ke *No\_DR* sebanyak 5 gambar. Pada kelas *Moderate*, model mencatat 129 prediksi yang sesuai, sementara masing-masing satu gambar salah diarahkan ke *No\_DR*, *Proliferate\_DR*, dan *Severe*. Kelas *No\_DR* menunjukkan performa paling akurat, dengan seluruh 151 gambar berhasil diklasifikasikan tanpa kesalahan. Sebaliknya, kelas *Proliferate\_DR* menunjukkan sedikit kelemahan, dengan 65 prediksi yang benar namun terdapat 13 gambar yang salah diklasifikasikan sebagai *Moderate* dan satu sebagai *Severe*. Untuk kelas *Severe*, sebanyak 62 gambar diklasifikasikan dengan tepat, namun masih terdapat beberapa kesalahan prediksi ke kelas *Moderate* (6 gambar), *No\_DR* (4 gambar), dan *Proliferate\_DR* (4 gambar). Secara umum, kesalahan klasifikasi paling banyak terjadi antara kelas-kelas dengan tingkat keparahan yang berdekatan, yang kemungkinan disebabkan oleh kemiripan karakteristik *visual* pada citra fundus mata di masing-masing kelas.

#### 4. PEMBAHASAN

Pada penelitian ini, penambahan *dilation layer* dengan *dilation rate* 2 pada arsitektur *EfficientNetV2* bertujuan untuk memperluas *receptive field* tanpa meningkatkan jumlah parameter secara signifikan. Mekanisme ini diterapkan setelah fitur diekstraksi oleh *backbone* *EfficientNetV2* yang dibekukan, kemudian dilanjutkan dengan normalisasi, *pooling global*, dan *lapisan fully connected* sebelum prediksi akhir dilakukan melalui fungsi aktivasi *softmax*. Dengan memperluas cakupan area yang diproses tanpa mengorbankan resolusi spasial, *dilation layer* meningkatkan sensitivitas model terhadap fitur halus seperti *mikroaneurisma* dan bercak *hemoragik* kecil yang menjadi indikator penting pada citra fundus retina. Sehingga model dapat mengenali tekstur, pola *vaskular*, dan perubahan halus pada retina yang terkait dengan tingkat keparahan *diabetic retinopathy* secara lebih akurat.

Hasil evaluasi menunjukkan bahwa penerapan *dilation layer* pada *EfficientNetV2-B0* dan *B1* memberikan performa terbaik dengan akurasi mencapai 97,14% dan metrik evaluasi lain yang sangat kompetitif. Ini mengindikasikan bahwa *dilation convolution* mampu meningkatkan efektivitas ekstraksi fitur dari citra retina sebesar 9,84% dari model *efficientnetV2*, sehingga model lebih sensitif dalam membedakan kelas-kelas yang berbeda pada DR.

Selain itu, dibandingkan dengan varian model tanpa *dilation* pada studi sebelumnya, (Xu et al.,

2023) melakukan klasifikasi diagnosis otomatis untuk klasifikasi tingkat keparahan retinopati diabetik menggunakan *EfficientNetV2* yang dikombinasikan dengan mekanisme *atensi squeeze-and-excite* (SE) dengan 2 lapisan *pooling* menghasilkan akurasi mencapai 84,2%, dan penelitian (Choirunisa et al., 2021) yang menggunakan *EfficientNet-B0* tanpa proses tambahan tambahan dapat mencapai akurasi 95%. Begitu juga penelitian yang menghasilkan akurasi sebesar 87,3%. Dengan menggunakan model *EfficientNetV2*.

Penelitian ini sejalan dengan temuan (Che et al., 2023) yang menggunakan *dilation layer* dengan berbasis FPN dengan Modul Residual Dilatasi untuk mengatasi masalah segmentasi multiskala pada citra retina tikus. Temuan ini membangun jaringan ekstraksi fitur *backbone* melalui kombinasi antara *dilated convolution* dan *residual block*, guna meningkatkan jangkauan *receptive field* sehingga dapat memperoleh informasi konteks yang lebih luas. Hasil penelitian ini mempengaruhi efektifitas model jaringan didukung dengan akurasi terbaik dengan nilai mPA sebesar 0,9732 dan mIoU sebesar 0,9213.

Hasil dari penelitian ini memiliki implikasi signifikan baik dalam ranah keilmuan maupun industri. Secara keilmuan, integrasi *dilated convolution* ke dalam arsitektur *EfficientNetV2* menunjukkan potensi besar dalam meningkatkan akurasi dan ketahanan model terhadap kompleksitas visual citra medis, khususnya fundus retina, serta membuka peluang untuk pengembangan arsitektur *lightweight* yang efektif dalam klasifikasi multikelas tanpa peningkatan parameter secara signifikan. Selain menambah literatur mengenai penggunaan *dilation layer* untuk memperluas *receptive field* tanpa mengorbankan resolusi spasial yang sebelumnya lebih banyak digunakan untuk segmentasi penelitian ini juga menunjukkan relevansi tinggi bagi industri, khususnya di bidang teknologi kesehatan (*healthtech*) dan *oftalmologi digital*.

Model yang dikembangkan memiliki potensi untuk diimplementasikan sebagai bagian dari sistem diagnosis berbantuan komputer (*computer-aided diagnosis*) dalam deteksi dini *diabetic retinopathy*, terutama di daerah dengan keterbatasan tenaga medis spesialis. Dengan performa yang kompetitif, model ini dapat diterapkan pada skala klinis, sistem berbasis *cloud*, atau perangkat mobile untuk skrining massal, sehingga meningkatkan efisiensi diagnosis, mempercepat pengambilan keputusan medis, dan memperluas jangkauan layanan kesehatan mata. Pendekatan berbasis *dilated EfficientNetV2* ini tidak hanya menyempurnakan metodologi klasifikasi citra medis, tetapi juga menghadirkan solusi praktis yang adaptif terhadap kebutuhan industri kesehatan masa kini dan masa depan.

Hasil penelitian ini memberikan kontribusi dalam bidang pengolahan citra berbasis *deep learning*. Meskipun menunjukkan peningkatan akurasi yang signifikan, penelitian ini masih memiliki

beberapa keterbatasan yaitu terdapat indikasi *overfitting* yang ditunjukkan pada Gambar 5 dan 6. Hal ini disebabkan oleh kompleksitas model yang tinggi namun diuji pada dataset yang tidak seimbang, sehingga mengurangi kemampuan generalisasi model. Selain itu, ukuran sampel yang digunakan dalam penelitian ini juga masih terbatas.

## 5. KESIMPULAN

Berdasarkan implementasi arsitektur EfficientNetV2 yang telah dimodifikasi dengan *dilation layer* serta hasil evaluasi yang dilakukan, penelitian ini membuktikan bahwa penambahan *dilated convolution* secara signifikan mampu meningkatkan kinerja model dalam mengklasifikasikan citra *diabetic retinopathy*. Penggunaan *dilation rate* sebesar 2 pada arsitektur ini, khususnya pada varian B0 dan B1, menghasilkan tingkat akurasi tertinggi sebesar 97,14%, disertai dengan nilai *precision*, *recall*, dan *F1-score* yang sangat baik.

Penerapan metode ini terbukti dapat memperluas *receptive field* tanpa meningkatkan jumlah parameter secara mencolok, sehingga model menjadi lebih efektif dalam menangkap pola lokal maupun global pada citra retina. Visualisasi hasil, seperti kurva ROC dan *confusion matrix*, turut mendukung temuan ini dengan menunjukkan tingkat akurasi tinggi dan konsistensi prediksi di setiap kelas.

Meski demikian, penelitian ini masih menghadapi beberapa kendala, seperti indikasi *overfitting* ringan yang terlihat dari perbedaan performa antara data pelatihan dan data validasi, serta keterbatasan dalam hal variasi dan keseimbangan dataset yang digunakan. Faktor-faktor ini berpotensi membatasi kemampuan model dalam melakukan generalisasi terhadap data baru di luar sampel pelatihan.

Untuk pengembangan ke depan, disarankan agar dilakukan penerapan teknik regularisasi, pemanfaatan dataset yang lebih besar dan seimbang, serta eksplorasi terhadap berbagai teknik *preprocessing* lain, seperti CLAHE dan augmentasi berbasis adaptif. Penelitian ini memberikan sumbangsih yang berarti dalam bidang pemrosesan citra medis dengan *deep learning*, serta membuka jalan menuju pengembangan sistem diagnosis *diabetic retinopathy* yang lebih presisi dan siap diaplikasikan dalam lingkungan klinis.

## DAFTAR PUSTAKA

- CHE, Z., BI, F., SUN, Y., XING, W., HUANG, H., dan ZHANG, X. 2023. Dilated residual FPN-based segmentation for mouse retinal images. *Heliyon*, 9(8), e18605. <https://doi.org/10.1016/j.heliyon.2023.e18605>
- CHOIRUNISA, N. A., KARLITA, T., & ASMARA, R. 2021. Deteksi Ras Kucing Menggunakan Compound Model Scaling Convolutional Neural Network. *Technomedia Journal*, 6(2), 236–251. <https://doi.org/10.33050/tmj.v6i2.1704>
- CITRA, E. E., FUDHOLI, D. H., dan DEWA, C. K. 2023. Implementasi Arsitektur EfficientNetV2 Untuk Klasifikasi Gambar Makanan Tradisional Indonesia. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 7(2), 766. <https://doi.org/10.30865/mib.v7i2.5881>
- DAS, S., & SAHA, S. K. 2022. Diabetic retinopathy detection and classification using CNN tuned by genetic algorithm. *Multimedia Tools and Applications*, 81(6), 8007–8020. <https://doi.org/10.1007/s11042-021-11824-w>
- DEWI, C., SANTOSO, A., INDRIATI, I., DEWI, N. A., & ARBAWA, Y. K. 2021. Evaluasi Performasi Ruang Warna pada Klasifikasi Diabetic Retinopathy Menggunakan Convolution Neural Network. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 8(3), 619–624. <https://doi.org/10.25126/jtiik.2021834459>
- Ditjen P2P. 2024, January 11. *Saatnya Mengatur Si Manis*. Kemenkes Ditjen P2P. <https://p2p.kemkes.go.id/saatnya-mengatur-si-manis/#:~:text=Menurut%20IDF%2C%20Indonesia%20menduduki%20peringkat,merupakan%20ibu%20dari%20segala%20penyakit>
- EL-ATEIF, S., & IDRI, A. 2022. Single-modality and joint fusion deep learning for diabetic retinopathy diagnosis. *Scientific African*, 17. <https://doi.org/10.1016/j.sciaf.2022.e01280>
- FU, Y., JU, Y., & ZHANG, D. 2024. MSEF-Net: A multi-scale EfficientNet Fusion for Diabetic Retinopathy grading. *Biomedical Signal Processing and Control*, 98, 106714. <https://doi.org/10.1016/j.bspc.2024.106714>
- HU, H., YU, C., ZHOU, Q., GUAN, Q., & FENG, H. 2024. HDConv: Heterogeneous kernel-based dilated convolutions. *Neural Networks*, 179, 106568. <https://doi.org/10.1016/j.neunet.2024.106568>
- JORDI DE LA TORRE, A. V. D. P. 2020. A deep learning interpretable classifier for diabetic retinopathy disease grading. *Neurocomputing*, 396, 365–476.
- KARTHIK, R., VAICHOLE, T. S., KULKARNI, S. K., YADAV, O., & KHAN, F. 2022. Eff2Net: An efficient channel attention-based convolutional neural network for skin disease classification. *Biomedical Signal Processing and Control*, 73, 103406. <https://doi.org/10.1016/j.bspc.2021.103406>
- LIU, R., CAI, W., LI, G., NING, X., & JIANG, Y. 2022. Hybrid Dilated Convolution Guided Feature Filtering and Enhancement Strategy for Hyperspectral Image Classification. *IEEE Geoscience and Remote Sensing Letters*, 19, 1–5. <https://doi.org/10.1109/LGRS.2021.3100407>

- MADARAPU, S., ARI, S., & MAHAPATRA, K. 2024. A multi-resolution convolutional attention network for efficient diabetic retinopathy classification. *Computers and Electrical Engineering*, 117, 109243. <https://doi.org/10.1016/j.compeleceng.2024.109243>
- MAJUMDAR, S., PRAMANIK, P., & SARKAR, R. 2023. Gamma function based ensemble of CNN models for breast cancer detection in histopathology images. *Expert Systems with Applications*, 213. <https://doi.org/10.1016/j.eswa.2022.119022>
- MEI, H., WANG, Q., YU, L., & ZENG, Q. 2024. A deep learning-based algorithm for intelligent prediction of adverse geologic bodies in tunnels. *Measurement Science and Technology*, 35(9), 096119. <https://doi.org/10.1088/1361-6501/ad4c8a>
- MUNTASA, A., WAHYUNINGRUM, R. T., HUSNI, SUGIARTI, A., YUSUF, M., MOTWAKEL, A., DEWI, D. A., ASMARA, Y. P., TUZZAHRA, Z., & MAHMUDI, W. F. 2025. EfficientNetV2-S Enhancement to Classify Acute Lymphoblastic Leukemia: Integrating Pre-Trained Models and Grid Search for Optimal Performance. *International Journal of Intelligent Engineering and Systems*, 18(1), 409–421. <https://doi.org/10.22266/ijies2025.0229.30>
- PHRIDVIRAJ, M. S. B., BHUKYA, R., MADUGULA, S., MANJULA, A., VODITHALA, S., & WASEEM, M. S. 2023. A bi-directional Long Short-Term Memory-based Diabetic Retinopathy detection model using retinal fundus images. *Healthcare Analytics*, 3. <https://doi.org/10.1016/j.health.2023.100174>
- QOMARIAH, D. U. N., TJANDRASA, H., & FATICHAH, C. 2019. Classification of Diabetic Retinopathy and Normal Retinal Images using CNN and SVM. *2019 12th International Conference on Information & Communication Technology and System (ICTS)*, 152–157. <https://doi.org/10.1109/ICTS.2019.8850940>
- QUMMAR, S., KHAN, F. G., SHAH, S., KHAN, A., SHAMSHIRBAND, S., REHMAN, Z. U., KHAN, I. A., & JADOON, W. 2019. A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection. *IEEE Access*, 7, 150530–150539. <https://doi.org/10.1109/ACCESS.2019.2947484>
- RAIAAN, M. A. K., FATEMA, K., KHAN, I. U., AZAM, S., RASHID, M. R. U., MUKTA, M. S. H., JONKMAN, M., & DE BOER, F. 2023. A Lightweight Robust Deep Learning Model Gained High Accuracy in Classifying a Wide Range of Diabetic Retinopathy Images. *IEEE Access*, 11, 42361–42388. <https://doi.org/10.1109/ACCESS.2023.3272228>
- RIZAL, S., IBRAHIM, N., PRATIWI, N. K. C., SAIDAH, S., & FU'ADAH, R. Y. N. 2020. Deep Learning untuk Klasifikasi Diabetic Retinopathy menggunakan Model EfficientNet. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, 8(3), 693. <https://doi.org/10.26760/elkomika.v8i3.693>
- ROY, S., TYAGI, M., BANSAL, V., & JAIN, V. 2022. SVD-CLAHE boosting and balanced loss function for Covid-19 detection from an imbalanced Chest X-Ray dataset. *Computers in Biology and Medicine*, 150, 106092. <https://doi.org/10.1016/j.combiomed.2022.106092>
- SHANG, Z., YU, C., HUANG, H., & LI, R. 2024. DCNet: A lightweight retinal vessel segmentation network. *Digital Signal Processing*, 153, 104651. <https://doi.org/10.1016/j.dsp.2024.104651>
- SILVA-TINOCO, R., CABRERA-GERARDO, D., CASTILLO-MARTÍNEZ, L., & CUATECONTZI-XOCHITIOTZI, T. 2024. Diabetic retinopathy prevalence in Mexico: Results from a primary public health access initiative for screening in patients with type 2 diabetes. *Diabetes Epidemiology and Management*, 13. <https://doi.org/10.1016/j.deman.2023.100182>
- TAN, M., & LE, Q. V. 2019. *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. <http://arxiv.org/abs/1905.11946>
- TAN, M., & LE, Q. V. 2021. *EfficientNetV2: Smaller Models and Faster Training*. <http://arxiv.org/abs/2104.00298>
- TUMMALA, S., THADIKE MALLA, V. S. G., KADRY, S., SHARAF, M., & RAUF, H. T. 2023. EfficientNetV2 Based Ensemble Model for Quality Estimation of Diabetic Retinopathy Images from DeepDRiD. *Diagnostics*, 13(4). <https://doi.org/10.3390/diagnostics13040622>
- XU, X., LIU, D., WANG, M., & LEI, M. 2023. *Application of End-To-End EfficientNetV2 in Diabetic Retinopathy Grading*. 153, 182–190. [https://doi.org/10.1007/978-3-031-20738-9\\_22](https://doi.org/10.1007/978-3-031-20738-9_22)
- YAZID, R. K., 2021. *Pendeteksian Diabetik Retinopatik Menggunakan Convolutional Neural Network (CNN)*. Universitas Sriwijaya.
- YE, Y., ZHOU, H., YU, H., HU, H., ZHANG, G., HU, J., & HE, T. 2022. An Improved EfficientNetV2 Model Based on Visual Attention Mechanism: Application to Identification of Cassava Disease. *Computational Intelligence and Neuroscience*, 2022. <https://doi.org/10.1155/2022/1569911>
- YUDISTIRA, N., WIDODO, A. W., & RAHAYUDI, B. 2020. Deteksi Covid-19 pada Citra Sinar-X

- Dada Menggunakan Deep Learning yang Efisien. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 7(6), 1289. <https://doi.org/10.25126/jtiik.2020763651>
- ZHANG, Y.-D., DONG, Z., CHEN, X., JIA, W., DU, S., MUHAMMAD, K., & WANG, S.-H. 2019. Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation. *Multimedia Tools and Applications*, 78(3), 3613–3632. <https://doi.org/10.1007/s11042-017-5243-3>

*Halaman ini sengaja dikosongkan*