

Deteksi Tingkat Keparahan Retinopati Diabetik Menggunakan EfficientNet V2-S

Muhammad Iqmal Basori¹, I Gede Susrama Mas Diyasa^{2*}, Hendra Maulana³

¹ Informatika, Universitas Pembangunan Nasional “Veteran” Jawa Timur

² Magister Teknologi Informasi, Universitas Pembangunan Nasional “Veteran” Jawa Timur

³ Bisnis Digital, Universitas Pembangunan Nasional “Veteran” Jawa Timur

121081010297@upnjatim.ac.id

3hendra.maulana.if@upnjatim.ac.id

*Corresponding author email: igsusrama.if@upnjatim.ac.id

Abstrak—Retinopati diabetik merupakan komplikasi mikrovaskular diabetes melitus yang menjadi penyebab utama kebutaan pada populasi usia produktif, sehingga deteksi dini melalui screening rutin sangat penting untuk mencegah kehilangan penglihatan permanen. Penelitian ini mengembangkan sistem deteksi otomatis retinopati diabetik menggunakan arsitektur *deep learning* EfficientNetV2-S dengan pendekatan klasifikasi biner untuk membedakan citra retina normal dan yang terindikasi retinopati diabetik. Dataset APTOS 2019 yang terdiri dari 3662 citra fundus retina dipraprocessing dan dibagi menjadi *training set* sebanyak 2929 citra dan *validation set* sebanyak 733 citra dengan *stratified splitting ratio* 80:20. Model diimplementasikan menggunakan framework PyTorch dengan strategi *transfer learning* dari bobot *pre-trained* ImageNet. Untuk memvalidasi efektivitas arsitektur, dilakukan evaluasi komparatif terhadap EfficientNet-B0 dan EfficientNet-B3 dengan konfigurasi training identik. Training selama 20 epoch dengan batch size 16 menghasilkan performa optimal pada epoch ke-8 dengan validation accuracy 99.32%, mengungguli EfficientNet-B0 (99.05%) dan EfficientNet-B3 (98.91%). Keunggulan signifikan EfficientNetV2-S terletak pada minimalisasi false positive dengan hanya 2 kasus (0.55%), berkurang 50% dibanding EfficientNet-B0 dan 60% dibanding EfficientNet-B3. Model memiliki sensitivitas tinggi dengan hanya 3 false negative dari 372 kasus positif. Hasil penelitian mendemonstrasikan bahwa EfficientNetV2-S memberikan hubungan antara akurasi dan efisiensi komputasi untuk implementasi *screening automated* retinopati diabetik pada *telemedicine platforms* di fasilitas kesehatan dengan keterbatasan akses spesialis oftalmologi.

Kata Kunci—Diabetic Retinopathy, EfficientNetV2-S, Deep Learning, Transfer Learning, Computer Visual

I. PENDAHULUAN

Retinopati Diabetik (DR) adalah komplikasi yang disebabkan oleh penyakit Diabetes Mellitus, yang menjadi fokus utama dalam penelitian ini. DR merupakan salah satu komplikasi serius dari diabetes yang dapat berujung pada peningkatan risiko penyakit lain seperti gangguan ginjal, penyakit jantung, hingga kematian. Diperkirakan jumlah penderita DR akan terus meningkat dengan signifikan di seluruh dunia dalam beberapa tahun mendatang. Proyeksi jumlah pasien DR *global* pada tahun 2040 diperkirakan mencapai 642 juta orang [1]. Di

Indonesia sendiri, prevalensi DR cukup tinggi, yakni sekitar 43,1% dari populasi penderita diabetes. Namun, meskipun angka penderita DR tergolong besar, tingkat pengobatan untuk kondisi ini masih sangat rendah, hanya sekitar 26,1%. Mayoritas dari penderita DR berasal dari kelompok masyarakat dengan keterbatasan finansial, yang tentunya mempengaruhi akses mereka terhadap pengobatan dan perawatan yang memadai. Penelitian ini mengungkapkan bahwa sebagian besar responden yang menderita DR adalah perempuan, yang mencakup 77,5% dari total sampel. Rata-rata usia responden adalah 55,2 tahun, dengan durasi penyakit diabetes yang sudah berjalan rata-rata 4,4 tahun. Selain itu, sebagian besar pasien atau sekitar 59,6% tidak memiliki riwayat keluarga dengan diabetes [2].

Pemanfaatan kecerdasan buatan (AI) dalam deteksi tingkat keparahan retinopati diabetik di Indonesia semakin berkembang dan menjadi solusi mengingat masih terbatasnya tenaga spesialis dan sarana pemeriksaan retinal manual. AI sangat dibutuhkan untuk meningkatkan cakupan dan efisiensi skrining serta menurunkan beban kebutaan akibat deteksi yang terlambat. Di Indonesia, keterbatasan dokter spesialis mata, kendala geografis, serta biaya pemeriksaan membuat cakupan deteksi dini RD masih rendah [3].

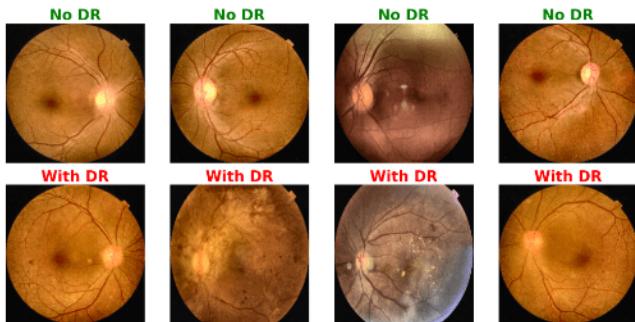
Menurut Dr. dr. Gitalisa Andayani, Sp.M(K), yang menjabat sebagai Ketua Tim Pengabdian Masyarakat FKUI sekaligus Dosen di Departemen Ilmu Kesehatan Mata RSCM, memerlukan pengembangan program skrining *Diabetic Retinopathy* yang lebih cepat dan akurat. Karena keberhasilan program skrining ini juga sangat bergantung pada ketersediaan peralatan dan fasilitas yang memadai, yang dapat mendukung proses diagnosis secara optimal. Maka dari itu, penting sekali dalam pengembangan *model* algoritma untuk program skrining agar deteksi dini untuk pasien bisa lebih cepat dan tepat dalam penanganan risiko retinopati diabetik [4].

II. METODOLOGI PENELITIAN

A. Landasan teori

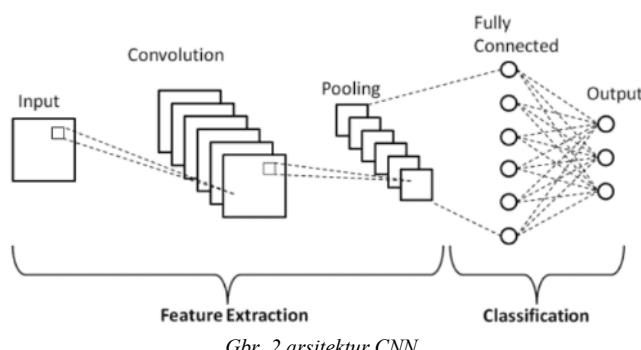
1) *Diabetik retinopati*: merupakan salah satu komplikasi mikrovaskular yang muncul pada pasien diabetes mellitus, yang

mempengaruhi pembuluh darah pada retina mata dan dapat menyebabkan masalah penglihatan hingga kebutaan [5]. Penyakit ini berkembang secara bertahap seiring berjalanannya waktu, dengan tingkat keparahannya bisa dilihat pada Gbr. 1, yang dimana tahapan tersebut meliputi No DR (tanpa retinopati diabetik), dan With DR (dengan penyakit retinopati diabetik) [6].



Gbr. 1 Gambar Fundus No DR dan With DR

2) *Convolutional Network*: adalah jenis *deep learning* yang dirancang khusus untuk pengolahan data berbentuk *grid*, seperti citra. Terinspirasi oleh struktur visual mamalia, CNN mengikuti prinsip yang ditemukan oleh Hubel dan Wiesel, yang menunjukkan bahwa neuron dalam otak merespons stimulus visual dalam area terbatas (*receptive field*). CNN terdiri dari beberapa lapisan utama yang bekerja secara hierarkis untuk memproses informasi secara efisien [7].



Gbr. 2 arsitektur CNN

Bisa dilihat pada Gbr. 2, lapisan pertama, yaitu *Convolutional Layer*, merupakan komponen utama yang melakukan operasi konvolusi, di mana hubungan antara *pixel* dipelajari melalui operasi matematis antara citra dan kernel atau *filter* berukuran kecil (misalnya 3×3 atau 5×5) yang berisi nilai acak. Selanjutnya, *Pooling Layer* berfungsi untuk mengurangi dimensi *feature map* sambil tetap mempertahankan informasi penting dari citra. *Fully Connected Layer* mengubah *feature map* yang telah diproses menjadi vektor yang siap untuk klasifikasi akhir, dan akhirnya, *Output Layer* menghasilkan prediksi untuk setiap kelas. Seiring dengan proses berjalan, kompleksitas CNN meningkat secara bertahap, di mana lapisan-lapisan awal berfokus pada fitur sederhana seperti

warna dan tepi, sedangkan lapisan-lapisan selanjutnya mulai mengenali elemen atau bentuk objek yang lebih besar, hingga pada akhirnya objek yang dimaksud dapat teridentifikasi dengan akurat [8].

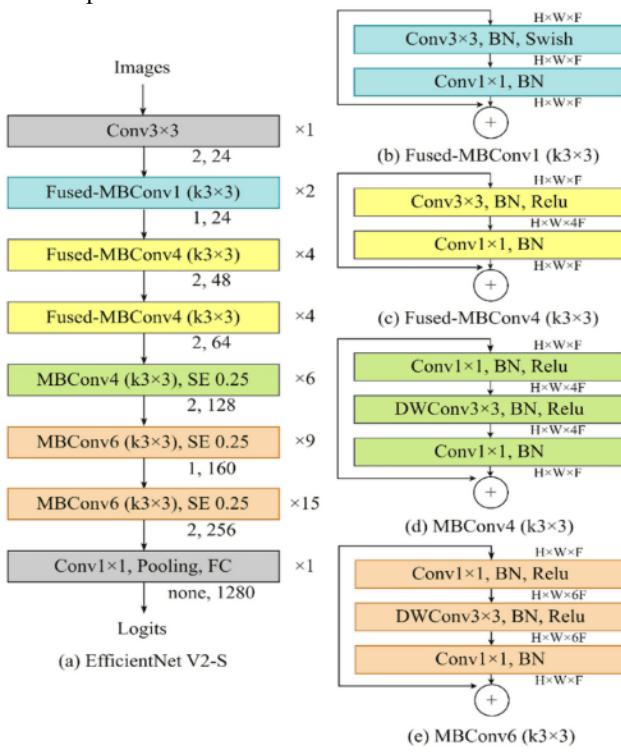
B. Studi Pustaka

Perkembangan penelitian Arsitektur CNN dan *Transfer Learning* di Indonesia telah mengembangkan berbagai arsitektur CNN untuk deteksi dan klasifikasi RD dengan hasil yang semakin meningkat. Model yang umum digunakan antara lain *ResNet*, *VGG*, *DenseNet*, dan *EfficientNet*. Penelitian terbaru menunjukkan bahwa *ResNet-152* mencapai akurasi tertinggi 99,41% pada dataset Kaggle untuk diagnosa DR dari fundus image menggunakan *improved activation function* dalam CNN models [9]. Pada model *RSG-Net* yang menggabungkan *multiple architectures*, *DenseNet201* mencapai akurasi tertinggi 82,7% pada *test set* untuk *binary classification*. Pada *2-stage binary classification* khusus, model mencapai akurasi 99,37%, *sensitivity* 100%, *specificity* 98,62%, dan *F1-score* 99,42% [10]. Studi perbandingan menunjukkan *VGG19* mencapai akurasi 80,40% untuk *binary classification* DR (2 classes: dengan DR dan tanpa DR). Dalam keluarga arsitektur EfficientNet, terdapat beberapa varian yang telah digunakan untuk deteksi penyakit retina. EfficientNet-B0 sebagai model baseline terkecil dalam family EfficientNet dengan 4 juta parameter telah menunjukkan efisiensi komputasi yang baik untuk aplikasi *mobile health* [11]. Sementara itu, EfficientNet-B3 dengan 10,7 juta parameter menawarkan *trade-off* yang seimbang antara kompleksitas model dan performa klasifikasi [12]. Dengan finding bahwa performa model terkait langsung dengan jumlah *convolutional* dan *pooling layers*. *EfficientNetB4* dengan enhanced cross-entropy loss mencapai akurasi 85,44%, *sensitivity* 97,36%, *specificity* 70,61%, *precision* 90,17%, dan *F1-score* 93,62% pada *binary classification* DR. *EfficientNetB5* dengan E-CE loss mencapai akurasi 84,88%, *sensitivity* 98,09%, *specificity* 65,49%, *precision* 88,72%, dan *F1-score* 93,17% [13]. Dalam studi lain penggunaan *ensemble* menggabungkan *pretrained EfficientNet Series* (B0-B7), *ResNet*, dan *DenseNet* dari berbagai sumber (*ImageNet*, *RadImageNet*) untuk meningkatkan *robustness* dan akurasi. Hasilnya, meningkat secara signifikan performa deteksi penyakit mata, mencapai akurasi 96-98% dengan sensitivitas dan spesifisitas tinggi [14]. Selain algoritma yang sudah disebutkan, banyak penelitian menggunakan *model pretrained* lain seperti *InceptionV3*, *MobileNetV2*, *Xception*, *NASNet*, dan *DenseNet*. Penggunaan *transfer learning* dari model *pretrained* tersebut telah membantu mengurangi waktu pelatihan, meningkatkan akurasi deteksi, dan mempercepat *deployment model* dengan *dataset* terbatas.

C. Pre-trained Model EfficientNetV2-S

EfficientNetV2-S yang merupakan hasil evolusi dari arsitektur *EfficientNet*, terdiri dari serangkaian blok konvolusi yang dirancang secara optimal menggunakan teknik *Neural*

Architecture Search. Arsitektur ini menggabungkan dua jenis *building blocks* utama yaitu *MBConv* (*Mobile Inverted Bottleneck Convolution*) yang telah terbukti efektif pada *EfficientNet* generasi pertama, serta *Fused-MBConv blocks* yang diperkenalkan pada *EfficientNetV2* untuk meningkatkan kecepatan *training* dengan mengurangi *overhead* dari operasi *depthwise convolution* pada layer-layer awal. Setiap *MBConv block* menggunakan *expansion layer* yang memperluas jumlah *channel*, diikuti oleh *depthwise convolution* untuk ekstraksi fitur spasial yang efisien, dan *projection layer* yang mengurangi dimensi channel kembali, dengan *skip connection* yang menghubungkan *input* dan *output block* untuk memfasilitasi *gradient flow* yang lebih baik selama *backpropagation* [15]. Untuk gambaran arsitektur *model* ini bisa dilihat pada Gbr. 3.



Gbr. 3 Arsitektur Umum EfficientNetV2-S [16]

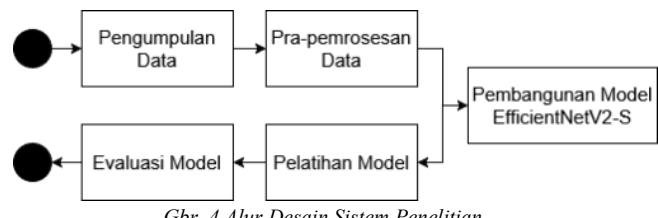
Penelitian terbaru telah membuktikan bahwa arsitektur *EfficientNetV2-S* menunjukkan performa unggul pada berbagai masalah klasifikasi citra medis, termasuk deteksi penyakit mata seperti katarak. Dalam studi deteksi katarak, *EfficientNetV2-S* mampu mencapai akurasi pengujian hingga 92,75% dengan *recall* dan *precision* serupa, membuktikan kemampuan model mengenali ciri unik katarak secara akurat dari citra fundus mata [17]. Penelitian lain pada klasifikasi penyakit malaria berbasis citra darah juga mendapatkan *EfficientNetV2-S* bisa mencapai akurasi validasi hingga 96%, semakin menegaskan keunggulan dan fleksibilitas model untuk transfer learning pada ragam permasalahan citra medis [18]. Keunggulan *EfficientNetV2-S* terletak pada arsitektur CNN yang telah dioptimasi sehingga menghasilkan model yang ringkas namun memiliki kapasitas

representasi fitur visual yang sangat kuat. Model ini menggunakan teknik *depthwise separable convolutions* dan *compound scaling* yang memaksimalkan efisiensi dan akurasi, memungkinkan pelatihan lebih cepat dengan jumlah parameter lebih sedikit dibandingkan arsitektur seperti *ResNet* atau *VGG*, namun tetap mempertahankan performa klasifikasi yang tinggi [12].

Untuk mengukur arsitektur *EfficientNetV2-S* dalam deteksi retinopati diabetik, penelitian ini juga mengimplementasikan dua model baseline pembanding dari keluarga *EfficientNet*, yaitu *EfficientNet-B0* dan *EfficientNet-B3*. *EfficientNet-B0* dipilih sebagai baseline ringan dengan 4 juta parameter, sedangkan *EfficientNet-B3* dengan 10,7 juta parameter. Ketiga model dilatih dengan konfigurasi hyperparameter yang sama, meliputi learning rate 0.0001, optimizer AdamW dengan weight decay 0.01, batch size 16, dan durasi training 20 epoch dengan augmentasi data yang sama juga. Evaluasi dilakukan pada validation set yang sama menggunakan *metrik accuracy*, *precision*, *recall*, *F1-score*, *confusion matrix*, serta *inference time* untuk menilai efisiensi komputasi.

III. DESAIN DAN IMPLEMENTASI SISTEM

Penelitian ini menggunakan pendekatan kuantitatif terapan dengan fokus pada penerapan model *deep learning* menggunakan algoritma *EfficientNetV2-S* untuk mendeteksi diabetik retinopati. Jenis penelitian ini merupakan eksperimen yang bertujuan untuk menerapkan solusi berbasis teknologi dalam bidang medis, khususnya dalam mendeteksi penyakit DR melalui dataset citra retina dan penggunaan arsitektur algoritma *pre-trained* yang tersedia. Metrik alur desain dan implementasi sistem dalam penelitian ini melibatkan beberapa tahapan utama, yang dapat dilihat pada Gbr. 4, termasuk pengumpulan data, pra-pemrosesan data, pembangunan model *EfficientNetV2-S*, pelatihan model, dan evaluasi model.



Gbr. 4 Alur Desain Sistem Penelitian

A. Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari dataset APTOS 2019 *Blindness Detection* yang disediakan oleh *Asia Pacific Tele-Ophthalmology Society* (APOTOS) dan dapat diakses melalui platform Kaggle. Dataset ini mencakup 3662 gambar fundus retina yang dikategorikan dalam lima label kelas yang telah ditentukan, yaitu: Retina Normal (1805 gambar), Ringan (370 gambar), Sedang (999 gambar), Parah (193 gambar), dan Proliferatif (295 gambar). Dataset ini

disimpan dalam format .PNG [19]. Persebaran data berdasarkan kategori keparahan dapat dilihat pada Tabel I berikut.

Tabel I persebaran data

Kelas	Klasifikasi	Jumlah Sampel
0	Tidak ada Diabetic Retinopathy	1805
1	Non-ploriferatif Diabetic Retinopathy Ringan	370
2	Non-ploriferatif Diabetic Retinopathy Sedang	999
3	Non-ploriferatif Diabetic Retinopathy Parah	193
4	Ploriferatif Diabetic Retinopathy	295

Berdasarkan pendekatan klinis dari ARIA System (*Automated Retinal Image Analysis*), klasifikasi tingkat keparahan retinopati diabetik dapat dikelompokkan secara *binary* menjadi "*non-referable*" dan "*referable*". "*Non-referable*" mencakup kategori No DR (tanpa kelainan) dan *Mild DR* (hanya mikroaneurisma), sedangkan "*referable*" meliputi *grade Moderate DR* (lebih dari mikroaneurisma namun belum kategori berat), *Severe DR* (≥ 20 perdarahan *intraretina*, *venous beading*, *IRMA*), dan *Proliferative DR* (terdapat neovaskularisasi atau perdarahan vitreous/preretinal). Dengan demikian, seluruh grade 1-4 (*Mild, Moderate, Severe NPDR*, dan *PDR*) secara klinis masuk dalam kelompok "*referable*" yang memerlukan rujukan ke spesialis mata [20]. Maka dari itu, peneliti akan menggabungkan seluruh sampel diabetik retinopati dari grade 1-4 seperti pada Tabel II untuk dijadikan sampel klasifikasi *binary* untuk kebutuhan penelitian ini.

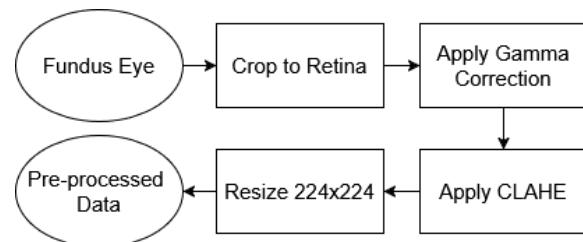
Tabel II persebaran data setelah dikelompokkan menjadi data binary

Kelas	Klasifikasi	Jumlah Sampel
0	No DR	1805
1	With DR	1857

B. Pra-pemrosesan Data

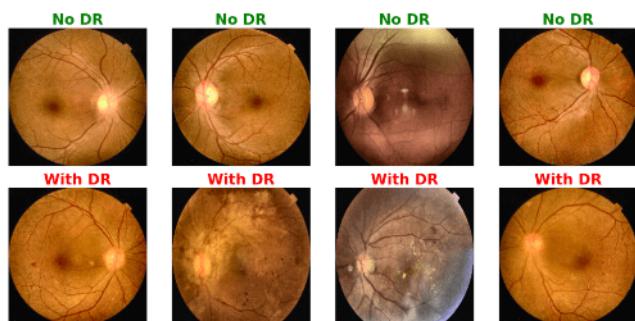
Pra-pemrosesan data menggunakan empat teknik (bisa dilihat pada Gbr. 5. *Crop to retina* untuk fokus pada area retina, *Gamma correction* ($\gamma=1.2$) untuk meningkatkan kontras, *CLAHE (Contrast Limited Adaptive Histogram Equalization)* untuk normalisasi cahaya, dan *Resize* ke 224x224 piksel. Dataset dibagi dengan *stratified split* 80-20 menjadi *training set*

(2,929 gambar) dan *validation set* (733 gambar). Semua pra-pemrosesan data menggunakan pustaka *OpenCV* dan *NumPy* untuk kecepatan waktu dan efisiensi komputasi selama training [20].



Gbr. 5 Alur pra-pemrosesan Data

Dengan hasil gambar fundus dengan Pra-pemrosesan datanya bisa dilihat pada Gbr. 6 berikut. Setelah melalui tahapan pra-pemrosesan data, bisa dilihat pada gambar memperjelas mikroneusma dan pembuluh darah retina yang menjadi indikator utama dalam mendeteksi retinopati diabetik. Dengan fokus yang lebih tajam pada area retina, detail mikrostruktur yang sebelumnya mungkin tersembunyi kini lebih terlihat, meningkatkan kemampuan *model* untuk mengenali tanda-tanda awal penyakit [20].



Gbr. 6 hasil pra-pemrosesan data fundus

C. Pembangunan Model

Penelitian ini mengimplementasikan arsitektur *deep learning EfficientNetV2-S* menggunakan *framework PyTorch* sebagai fondasi utama dalam membangun model deteksi retinopati diabetik. *PyTorch* dipilih karena menyediakan pengembangan arsitektur *neural network* dengan akselerasi GPU melalui *CUDA*. Model dibangun dengan memanfaatkan teknik *transfer learning*, di mana bobot *pre-trained* dari *dataset ImageNet* yang terdiri dari lebih dari 14 juta gambar dengan 1000 kelas digunakan sebagai inisialisasi awal pada *backbone EfficientNetV2-S*. Implementasi *backbone EfficientNetV2-S* diambil dari modul *torchvision.models* yang menyediakan arsitektur *pre-trained state-of-the-art* untuk berbagai tugas *computer vision*.

Input layer menerima citra dengan dimensi 224x224 piksel dalam format RGB yang terdiri dari 3 channel warna, yang diproses melalui serangkaian transformasi menggunakan library *torchvision.transforms*. Pada tahap *preprocessing*, citra retina yang berasal dari berbagai resolusi dan kualitas terlebih dahulu di-*resize* ke ukuran standar 224x224 piksel menggunakan fungsi *Resize* dengan metode interpolasi bilinear untuk mempertahankan kualitas visual. Proses normalisasi kemudian diterapkan menggunakan nilai *mean* [0.485, 0.456, 0.406] dan *standard deviation* [0.229, 0.224, 0.225] dari dataset *ImageNet*.

Pada tahap modifikasi arsitektur untuk mengadaptasi model *pre-trained* ke tugas klasifikasi biner yang spesifik, lapisan *classifier original* dari *EfficientNetV2-S* yang awalnya dirancang untuk 1000 kelas *ImageNet* diganti dengan arsitektur *classifier* baru menggunakan modul *torch.nn.Sequential*. Arsitektur *classifier* baru ini terdiri dari *Dropout layer* dengan probabilitas 0.3 yang secara *random* menonaktifkan 30 persen *neuron* selama *training* untuk mencegah *overfitting* dengan memaksa model belajar representasi fitur yang lebih *robust* dan tidak terlalu bergantung pada *neuron-neuron* spesifik. Setelah *Dropout*, sebuah *fully connected layer* yang diimplementasikan menggunakan *torch.nn.Linear* memetakan *feature vector* dengan dimensi 1280 yang merupakan *output* dari *global average pooling backbone* *EfficientNetV2-S* ke dalam 2 neuron *output* yang merepresentasikan *logic* untuk dua kelas klasifikasi, yaitu “*No DR*” untuk citra retina normal dan “*With DR*” untuk citra retina yang terindikasi retinopati diabetik. Total *parameter* yang dapat dilatih dalam model ini mencapai 20.180.050 *parameter* yang mencakup seluruh bobot dan bias pada setiap layer konvolusi, *batch normalization*, dan *fully connected layer*, yang seluruhnya dilakukan selama proses *training*.

D. Pelatihan Model

Model dilatih menggunakan *optimizer AdamW* yang diimplementasikan melalui *torch.optim.AdamW*, sebuah varian dari *Adam optimizer* yang menambahkan *weight decay* secara terpisah dari *gradient update* untuk memberikan regularisasi yang lebih efektif. *Learning rate* awal ditetapkan sebesar 0.0001 yang merupakan nilai konservatif namun aman untuk *fine-tuning* model *pre-trained*, dengan *weight decay* 0.01 yang berperan sebagai teknik regularisasi L2 untuk mencegah bobot model menjadi terlalu besar dan menyebabkan *overfitting*, sehingga membantu model mencapai konvergensi optimal sambil mempertahankan kemampuan generalisasi pada data yang belum pernah dilihat sebelumnya. Fungsi *loss* yang digunakan adalah *Cross-Entropy Loss* yang diimplementasikan melalui *torch.nn.CrossEntropyLoss*, yang merupakan kombinasi dari *Softmax activation* dan *Negative Log-Likelihood Loss*.

Untuk mengoptimalkan proses pembelajaran dan meningkatkan konvergensi model, strategi *learning rate scheduling* diterapkan menggunakan *Cosine Annealing* dengan parameter *T_max* sebesar 20 *epoch* yang sesuai dengan total durasi *training*. *Scheduler* ini bekerja dengan secara bertahap menurunkan *learning rate* dari nilai maksimal di awal *training* mengikuti fungsi *cosine* yang *smooth*, hingga mencapai nilai mendekati nol di akhir *training*, kemudian melakukan *fine-tuning* dengan langkah kecil menjelang akhir *training* untuk *konvergen* ke minimum lokal atau *global* yang lebih baik. Model dilatih selama 20 *epoch* dengan *batch size* 16 yang dipilih, dimana setiap *epoch* model akan melihat keseluruhan dataset training sebanyak satu kali dengan update *parameter* dilakukan setiap *mini-batch*.

Untuk meningkatkan *robustness* dan kemampuan generalisasi *model*, augmentasi data diterapkan secara ekstensif pada *training set* menggunakan berbagai transformasi yang diimplementasikan melalui *torchvision.transforms.Compose* yang menggabungkan *multiple transformations* dalam sebuah *pipeline*. Augmentasi yang diterapkan meliputi *random horizontal flip* dengan *probabilitas* 0.5 yang mensimulasikan variasi orientasi kamera saat pengambilan gambar retina, *random vertical flip* dengan *probabilitas* 0.5 yang menambah variasi rotasi data, *random rotation* hingga 20 derajat dalam rentang -20 hingga +20 derajat untuk mensimulasikan perbedaan sudut pengambilan gambar fundus *camera*, serta *color jitter* yang secara *random* mengubah *brightness* dan *contrast* dalam rentang ± 0.2 atau 20 persen dari nilai original untuk mensimulasikan variasi kondisi pencahayaan dan kualitas citra yang berbeda-beda dari berbagai perangkat *medical imaging*. Seluruh teknik augmentasi ini diterapkan secara *random* pada setiap *epoch* dengan probabilitas yang telah ditentukan, sehingga *model* akan melihat variasi yang berbeda dari gambar yang sama pada setiap *epoch*, yang bertujuan meningkatkan generalisasi *model* dan mencegah *overfitting* dari data training yang terbatas. Selama proses *training*, *library PIL (Python Imaging Library)* digunakan untuk *loading* dan manipulasi dasar citra.

E. Evaluasi Model

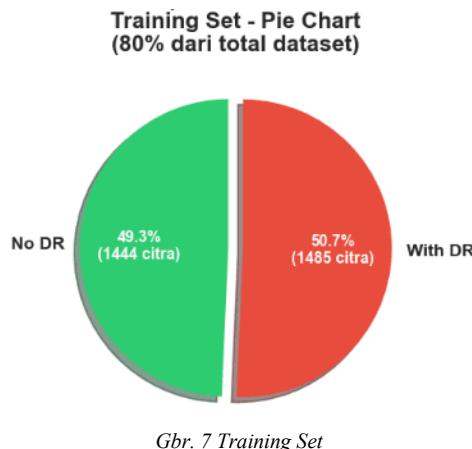
Evaluasi performa *model* diimplementasikan melalui *library scikit-learn* dalam mendeteksi retinopati diabetik pada data validasi yang terdiri dari 733 citra retina. Metrik *accuracy_score* digunakan untuk menghitung persentase prediksi benar dari keseluruhan prediksi yang dibuat oleh model, memberikan gambaran umum performa klasifikasi secara keseluruhan. *Confusion matrix* yang divisualisasikan menggunakan *heatmap* dari *library seaborn* menampilkan matriks 2x2 yang menganalisis empat komponen penting yaitu *True Positive* untuk sampel dengan retinopati diabetik yang berhasil dideteksi, *True Negative* untuk retina normal yang diidentifikasi dengan tepat, *False Positive* yang merepresentasikan alarm palsu di mana retina normal salah

diklasifikasikan sebagai retinopati diabetik, dan *False Negative* yang merupakan kesalahan paling berbahaya karena *model* gagal mendeteksi retinopati diabetik yang sebenarnya ada sehingga pasien berisiko tidak mendapatkan penanganan yang diperlukan. *Classification_report* dari *scikit-learn* menghasilkan laporan evaluasi mencakup tiga metrik untuk setiap kelas yaitu *precision* yang mengukur proporsi prediksi positif yang benar untuk mengetahui tingkat kepercayaan model, *recall* yang mengukur kemampuan model menangkap kasus retinopati diabetik yang ada untuk meminimalkan *missed diagnosis*, dan *F1-score* sebagai *harmonic mean* dari *precision* dan *recall*. Visualisasi *learning curve* juga dibuat menggunakan *matplotlib* untuk memplot *training accuracy* dan *validation accuracy* sepanjang 20 *epoch*, yang berguna untuk mendeteksi fenomena seperti *overfitting*.

IV. HASIL DAN PEMBAHASAN

A. Hasil Pra-pemrosesan Dataset

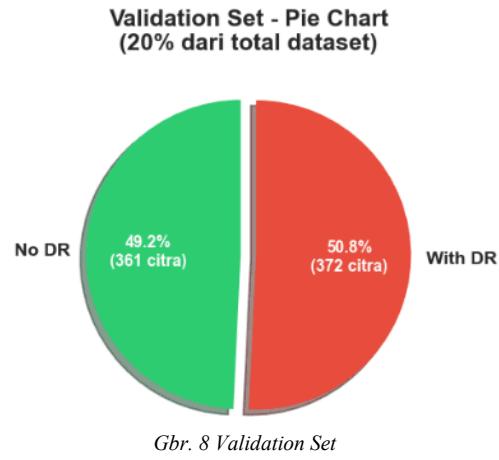
Proses *preprocessing* citra dilakukan menggunakan *library OpenCV*. Setiap citra fundus retina yang awalnya memiliki dimensi dan kualitas yang berbeda akibat perbedaan perangkat dan kondisi pengambilan gambar, dengan langkah pertama adalah *resizing* citra ke dimensi 224x224 piksel menggunakan metode interpolasi. Normalisasi intensitas piksel dengan rentang nilai dari 0-255 menjadi 0-1 agar sesuai dengan arsitektur model *deep learning*. Serta beberapa jenis *preprocessing* seperti *crop to retina*, *apply gamma*, *apply CLAHE* bertujuan untuk penghapusan noise dan *artifacts* yang mungkin ada pada citra *original*, serta penyesuaian kontras untuk meningkatkan visibilitas fitur-fitur penting seperti pembuluh darah, mikroaneurisma, dan eksudat dari penyakit retinopati diabetik.



Gbr. 7 Training Set

Setelah *preprocessing* data selesai, *dataset* kemudian dibagi menjadi *training set* dan *validation set* menggunakan *stratified splitting* dengan rasio 80:20. Hasil pembagian dataset menunjukkan bahwa *training set* terdiri dari 2929 citra dengan komposisi 1444 citra untuk kelas "No DR" dan 1485 citra untuk

kelas "With DR", *Validation set* yang digunakan untuk evaluasi performa *model* selama dan setelah *training* terdiri dari 733 citra dengan 361 citra kelas "No DR" dan 372 citra kelas "With DR", untuk selengkapnya bisa dilihat hasil visual dari Gbr. 7 dan Gbr. 8.



Gbr. 8 Validation Set

B. Hasil Training Model EfficientNetV2-S

Proses training *model EfficientNetV2-S* dilakukan selama 20 *epoch* dengan menggunakan konfigurasi *hyperparameter* yang telah ditentukan. Training dilakukan pada *hardware* dengan akselerasi *GPU CUDA* dari *NVIDIA RTX 3050* untuk operasi konvolusi. Didapatkan untuk setiap *epoch* diselesaikan dalam waktu rata-rata sekitar 50-60 detik untuk *training set* dan 4-5 detik untuk *validation set*. Seluruh proses *training* yang mencakup 20 *epoch* diselesaikan dalam waktu total 19.06 menit. Pada setiap *epoch*, *training set* dengan 2929 citra diproses dalam 184 mini-batch, sedangkan *validation set* dengan 733 citra diproses dalam 46 mini-batch berdasarkan batch size 16 yang telah ditetapkan.



Gbr. 9 Proses awal model dilatih

Dinamika pembelajaran model dapat diamati melalui perubahan *loss* dan *accuracy* sepanjang *epoch training*, yang menunjukkan pola konvergensi yang sangat baik tanpa indikasi *overfitting* atau *underfitting*. Pada Gbr. 9 dan Gbr. 10, progress bar menunjukkan angka 184/184 untuk training dan 46/46 untuk validation, yang merupakan jumlah mini-batch yang diproses pada setiap epoch berdasarkan pembagian dataset dengan batch size 16. Lalu pada *epoch pertama*, model yang dimulai dengan bobot *pre-trained* dari *ImageNet* langsung menunjukkan performa yang cukup baik dengan *training accuracy* mencapai 92.11 persen dan *validation accuracy* 96.18. *Training loss* pada *epoch pertama* tercatat sebesar 0.2069 yang kemudian menurun secara konsisten pada *epoch-epoch* berikutnya, menunjukkan bahwa model terus mempelajari pola yang lebih kompleks dari data training dengan efektif. *Validation loss* juga menunjukkan penurunan yang stabil dari 0.1047 pada *epoch pertama* menjadi 0.0503 pada *epoch kedua*. Peningkatan performa terlihat pada *epoch-epoch* awal *training*, di mana *validation accuracy* melonjak dari 96.18 persen pada *epoch pertama* menjadi 98.36 persen pada *epoch kedua*, kemudian terus meningkat mencapai 98.91 persen pada *epoch ketiga*. Puncak performa model tercapai pada *epoch ke-8*, di mana model mencatat performa terbaik dengan *training accuracy* 99.32 persen dan *validation accuracy* yang juga mencapai 99.32 persen dengan *validation loss* hanya 0.0325 pada Gbr. 11. Model pada *epoch* ini kemudian disimpan sebagai *best model* yang akan digunakan untuk evaluasi *final* dan *deployment*.

```
Epoch 5/20
-----
Training: 100%|██████████| 184/184 [00:50<00:00, 3.64it/s]
Validation: 100%|██████████| 46/46 [00:04<00:00, 10.39it/s]
Train - Loss: 0.0538 | Acc: 98.33%
Val - Loss: 0.0502 | Acc: 98.77%

Epoch 6/20
-----
Training: 100%|██████████| 184/184 [00:50<00:00, 3.67it/s]
Validation: 100%|██████████| 46/46 [00:04<00:00, 10.46it/s]
Train - Loss: 0.0556 | Acc: 98.40%
Val - Loss: 0.0400 | Acc: 98.50%

Epoch 7/20
-----
Training: 100%|██████████| 184/184 [00:51<00:00, 3.60it/s]
Validation: 100%|██████████| 46/46 [00:04<00:00, 9.61it/s]
Train - Loss: 0.0382 | Acc: 98.87%
Val - Loss: 0.0536 | Acc: 98.64%

Epoch 8/20
-----
Training: 100%|██████████| 184/184 [00:51<00:00, 3.58it/s]
Validation: 100%|██████████| 46/46 [00:04<00:00, 9.84it/s]
Train - Loss: 0.0262 | Acc: 99.32%
Val - Loss: 0.0325 | Acc: 99.32%
✓ Best model saved! (Acc: 99.32%)
```

Gbr. 10 Proses awal model dilatih

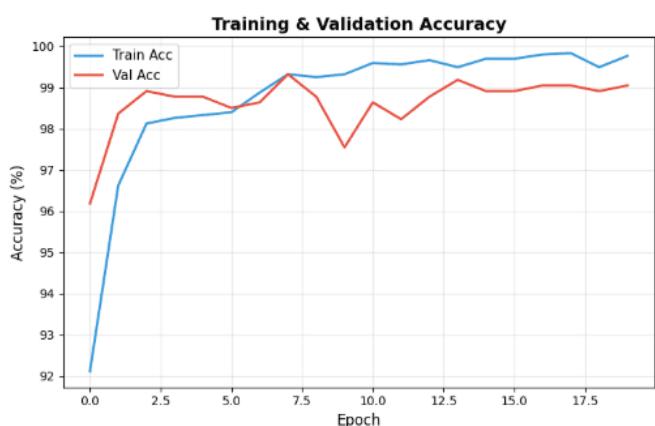
```
Epoch 18/20
-----
Training: 100%|██████████| 184/184 [00:51<00:00, 3.54it/s]
Validation: 100%|██████████| 46/46 [00:04<00:00, 10.46it/s]
Train - Loss: 0.0079 | Acc: 99.83%
Val - Loss: 0.0312 | Acc: 99.05%

Epoch 19/20
-----
Training: 100%|██████████| 184/184 [00:51<00:00, 3.56it/s]
Validation: 100%|██████████| 46/46 [00:04<00:00, 9.98it/s]
Train - Loss: 0.0157 | Acc: 99.49%
Val - Loss: 0.0335 | Acc: 98.91%

Epoch 20/20
-----
Training: 100%|██████████| 184/184 [00:55<00:00, 3.31it/s]
Validation: 100%|██████████| 46/46 [00:05<00:00, 8.79it/s]
Train - Loss: 0.0098 | Acc: 99.76%
Val - Loss: 0.0311 | Acc: 99.05%
=====
Training complete! Time: 19.06 minutes
Best Validation Accuracy: 99.32%
=====
```

Gbr. 11 Proses akhir model dilatih

Pada Gbr. 12, model tetap menunjukkan stabilitas yang sangat baik dengan *validation accuracy* yang berfluktuasi dalam rentang 97.54 persen hingga 99.18 persen, *Training accuracy* terus meningkat mencapai nilai tertinggi 99.83 persen pada *epoch ke-18*, *analisis loss curves* di mana *training loss* menurun secara monoton dari 0.2069 pada *epoch pertama* hingga mencapai nilai serendah 0.0079 pada *epoch ke-18*, *validation loss* yang dimulai dari 0.1047 pada *epoch pertama* juga menunjukkan tren penurunan yang konsisten hingga mencapai nilai terendah 0.0311 pada *epoch ke-20*.



Gbr. 12 Analisis kurva training dan loss

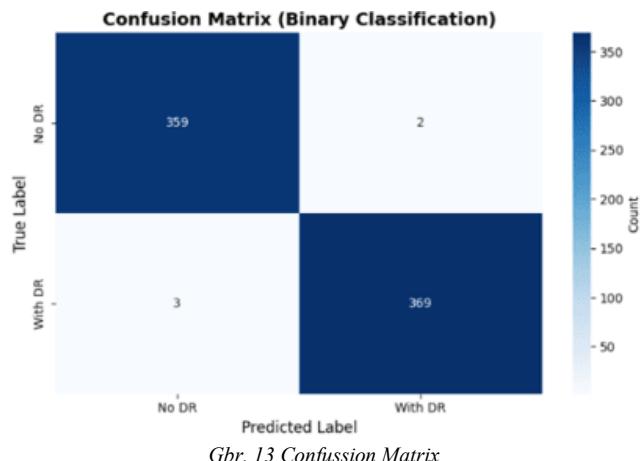
Pada Tabel III evaluasi performa *model EfficientNetV2-S* dilakukan pada *validation set* yang terdiri dari 733 citra retina menggunakan *best model* yang tersimpan dari *epoch* ke-8 untuk mengukur kemampuan *model* dalam metrik evaluasi utama yang digunakan adalah *accuracy* yang mengukur proporsi

prediksi benar dari keseluruhan prediksi, di mana model berhasil mencapai *validation accuracy* final sebesar 99.32 persen, menunjukkan bahwa dari 733 citra pada validation set, model mampu mengklasifikasikan sekitar 728 citra dengan benar dan hanya melakukan kesalahan pada sekitar 5 citra saja.

Tabel III Tabel laporan klasifikasi

	precision	recall	F1-score	support
No DR	99.17%	99.45%	99.31%	361
With DR	99.46%	99.19%	99.33%	372
macro avg	99.32%	99.32%	99.32%	733
weighted avg	99.32%	99.32%	99.32%	733

Pada Gbr. 13 dan Tabel IV analisis *False Positive* menunjukkan bahwa terdapat hanya 2 kasus dari 361 citra retina normal yang salah diklasifikasikan retinopati diabetik, menghasilkan *False Positive Rate* yang sangat rendah hanya sekitar 0.55 persen. Untuk analisis *False Negative*, berdasarkan *confusion matrix*, model menghasilkan 3 False Negative dari 372 kasus positif aktual, yang berarti ada 3 pasien yang sebenarnya memiliki retinopati diabetik namun tidak terdeteksi oleh *model* dan akan mendapatkan hasil *screening negatif*. *False Negative Rate* yang hanya 0.81 persen dalam penelitian ini menunjukkan bahwa model memiliki sensitivitas yang sangat tinggi dalam mendeteksi keberadaan retinopati diabetik.



Gbr. 13 Confusion Matrix

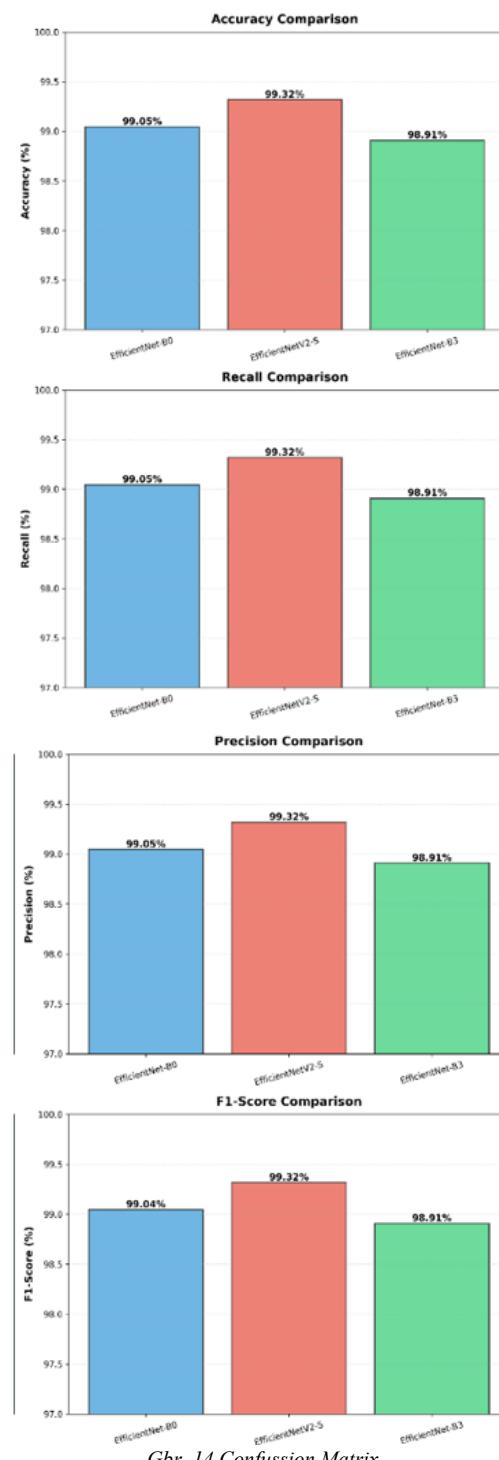
Tabel IV Breakdown Confusion Matrix

Kelas Sebenarnya	No DR	With DR	Total	Accuracy
No DR	359 (TN)	2 (FP)	361	99.45%
With DR	3 (FN)	369 (TP)	372	99.19%
Total	362	371	733	99.32%

C. Perbandingan Performa dengan Model Baseline

Untuk memvalidasi keunggulan arsitektur EfficientNetV2-S, dilakukan evaluasi terhadap dua model *baseline* dari keluarga

EfficientNet, yaitu EfficientNet-B0 dan EfficientNet-B3. Ketiga model dilatih dengan konfigurasi hyperparameter yang sama menggunakan dataset APTOS 2019, kemudian dievaluasi pada validation set yang terdiri dari 733 citra fundus retina. Perbandingan dilakukan berdasarkan empat aspek utama yaitu performa klasifikasi, kompleksitas model, efisiensi inferensi, dan pola kesalahan prediksi.



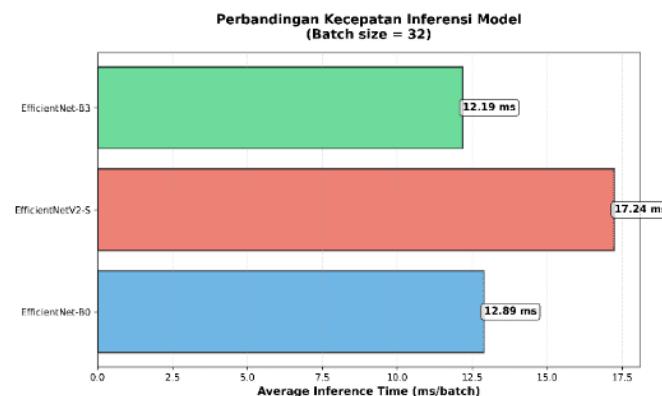
Gbr. 14 Confusion Matrix

1) *Perbandingan Performa dengan Model Baseline*: Hasil evaluasi metrik klasifikasi pada Tabel V dan visualisasi pada Gbr. 14 menunjukkan bahwa EfficientNetV2-S mencapai performa terbaik di semua metrik evaluasi. Model ini mencatat accuracy sebesar 99.32%, precision 99.32%, recall 99.32%, dan F1-score 99.32%, mengungguli EfficientNet-B0 yang mencapai accuracy 99.05% dan EfficientNet-B3 dengan accuracy 98.91%. Perbedaan akurasi terlihat 0.27% lebih tinggi dari B0 dan 0.41% dari B3.

Tabel V Perbandingan Metrik Klasifikasi

Model	Params	Acc	Precision	Recall	F1-Score
B0	4.0M	99.05%	99.05%	99.05%	99.04%
V2-S	20.2M	99.32%	99.32%	99.32%	99.32%
B3	10.7M	98.91%	98.91%	98.91%	98.91%

2) *Analisis waktu inferensi*: Analisis waktu inferensi pada Gbr. 15 menggunakan batch size 16 menunjukkan hasil yang menarik. EfficientNet-B3 memiliki waktu inferensi tercepat dengan 12.19 ms per batch, diikuti oleh EfficientNet-B0 dengan 12.89 ms per batch, sedangkan EfficientNetV2-S memerlukan 17.24 ms per batch. Meskipun EfficientNetV2-S memiliki waktu inferensi yang lebih lambat sekitar 34% dibandingkan B3, perbedaan hanya sekitar 5 milidetik per batch (32 gambar), yang berarti sekitar 0.16 ms per gambar.



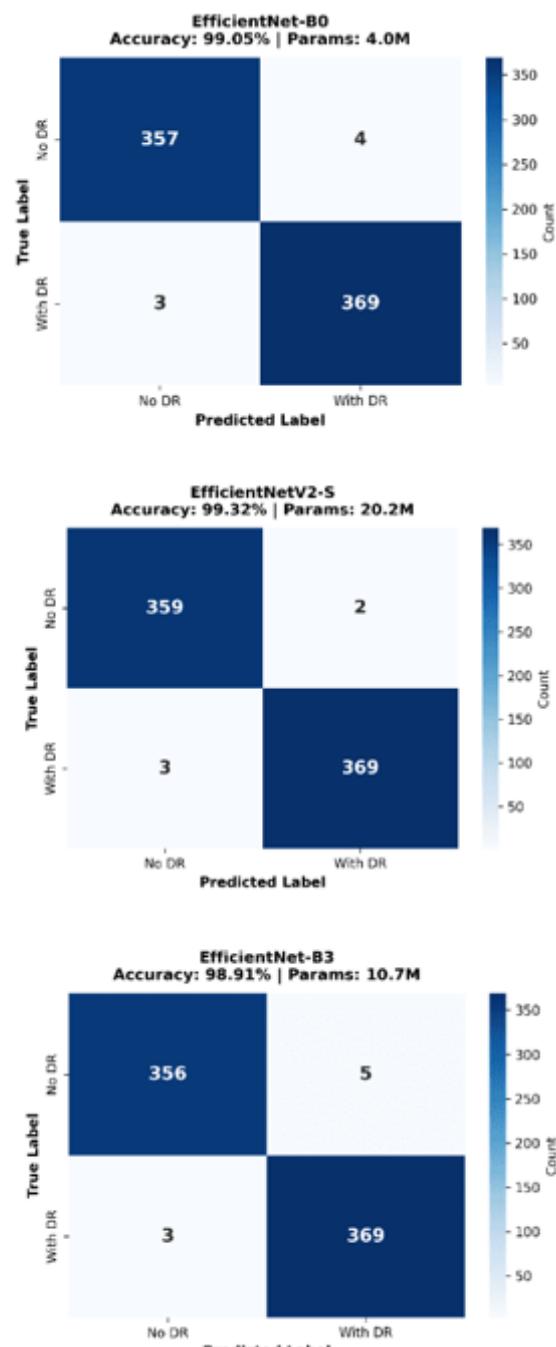
Gbr. 15 Perbandingan Kecepatan Inferensi Model

3) *Confusion Matrix*: Pada Gbr. 16 menampilkan perbandingan confusion matrix dari ketiga model yang mengungkap pola kesalahan prediksi. Ketiga model menunjukkan performa yang sangat baik dengan kesalahan prediksi yang minimal, namun distribusi error yang berbeda mengungkap karakteristik masing-masing arsitektur.

IV. KESIMPULAN

Kesimpulan dari penelitian ini adalah keberhasilan pengembangan sistem deteksi otomatis retinopati diabetik menggunakan arsitektur EfficientNetV2-S dengan pendekatan klasifikasi biner pada dataset APTOS 2019 (3.662 citra fundus retina). Model yang diimplementasikan menggunakan framework PyTorch dengan *transfer learning* dari ImageNet

mencapai performa optimal pada epoch ke-8, dengan *validation accuracy* 99,32%, *precision* 99,32%, *recall* 99,32%, dan *F1-score* 99,32%. Evaluasi terhadap model *baseline* membuktikan superioritas EfficientNetV2-S, yang mengungguli EfficientNet-B0 (*accuracy* 99,05%) dan EfficientNet-B3 (*accuracy* 98,91%). Keunggulan paling signifikan terletak pada minimalisasi *false positive* dengan hanya 2 kasus (0,55%), berkurang 50% dibanding EfficientNet-B0 dan 60% dibanding EfficientNet-B3.



Gbr. 16 Perbandingan Confusion Matriks ketiga model

Ketiga model menunjukkan sensitivitas tinggi yang konsisten dengan 3 *false negative* dari 372 kasus positif (FN rate 0,81%). Meskipun EfficientNetV2-S memerlukan *inference time* sedikit lebih lama (17,24 ms per *batch* vs 12,19 ms pada B3), perbedaan 5 milidetik ini memberikan peningkatan akurasi yang penting. Hasil penelitian mendemonstrasikan bahwa EfficientNetV2-S menawarkan hubungan optimal antara akurasi, efisiensi komputasi, dan minimalisasi kesalahan diagnosis untuk implementasi skrining otomatis retinopati diabetik pada *telemedicine platforms* di fasilitas kesehatan dengan keterbatasan akses spesialis oftalmologi. Penelitian lanjutan dapat dieksplorasi untuk validasi pada *dataset eksternal*, implementasi klasifikasi *multi-class*, optimisasi untuk *edge devices*, dan evaluasi dalam *setting klinis real-world*.

REFERENSI

- [1] X. Wang, C. Qiu, X. Ren, Z. Xiong, V. C. M. Leung, and D. Niyato, "Research Challenges and Future Directions," *Wirel. Networks (United Kingdom)*, pp. 105–108, 2022, doi: 10.1007/978-3-031-10186-1_6.
- [2] R. Amelia, M. D. Sari, V. Virgayanti, H. Wijaya, R. M. Situmorang, and S. Nadyah, "Clinical assessment of retinopathy in type 2 diabetes: a cross-sectional study," *Fam. Med. Prim. Care Rev.*, vol. 26, no. 4, pp. 433–437, 2024, doi: 10.5114/fmcr.2024.144910.
- [3] dr. S. Evani, "Deteksi Retinopati Diabetik dengan Artificial Intelligence," <https://www.alomedika.com>. Accessed: Nov. 02, 2025. [Online]. Available: <https://www.alomedika.com/deteksi-retinopati-diabetik-dengan-ai>
- [4] Maudisha, "Preventing Diabetic Retinopathy in Diabetic Patients, FKUI Educated Health Workers and Communities on Kepulauan Seribu," www.ui.ac.id. [Online]. Available: <https://www.ui.ac.id/en/preventing-diabetic-retinopathy-in-diabetic-patients/>
- [5] É. B. Rangel, C. O. Rodrigues, and J. R. De Sá, "Micro- and Macrovascular Complications in Diabetes Mellitus: Preclinical and Clinical Studies," *J. Diabetes Res.*, vol. 2019, no. 1, pp. 96–123, 2019, doi: 10.1155/2019/2161085.
- [6] C. Suedumrong, S. Phongmoo, T. Akarajaka, and K. Leksakul, "Diabetic Retinopathy Detection Using Convolutional Neural Networks with Background Removal, and Data Augmentation," *Appl. Sci.*, vol. 14, no. 19, 2024, doi: 10.3390/app14198823.
- [7] K. F. Manampiring and U. M. Nusantara, "Gambar 2.1. Arsitektur CNN Sumber: [10]," pp. 7–24.
- [8] D. S. Candra, "Implementasi Convolutional Neural Network (CNN) Untuk Klasifikasi Citra Bunga," vol. 16, no. 1, pp. 2580–2582, 2020.
- [9] U. Bhimavarapu and G. Battineni, "Deep Learning for the Detection and Classification of Diabetic Retinopathy with an Improved Activation Function," *Healthc.*, vol. 11, no. 1, 2023, doi: 10.3390/healthcare11010097.
- [10] S. Akhtar *et al.*, "A deep learning based model for diabetic retinopathy grading," *Sci. Rep.*, vol. 15, no. 1, pp. 1–20, 2025, doi: 10.1038/s41598-025-87171-9.
- [11] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," *36th Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 10691–10700, 2019.
- [12] M. Tan and Q. V. Le, "EfficientNetV2: Smaller Models and Faster Training," *Proc. Mach. Learn. Res.*, vol. 139, pp. 10096–10106, 2021.
- [13] H. Liu, K. Yue, S. Cheng, C. Pan, J. Sun, and W. Li, "Hybrid model structure for diabetic retinopathy classification," *J. Healthc. Eng.*, vol. 2020, 2020, doi: 10.1155/2020/8840174.
- [14] M. Azmoeleh-Kalati, H. Shabani, M. S. Maghreh, Z. Barzegar, and R. Lashgari, "Leveraging an ensemble of EfficientNetV1 and EfficientNetV2 models for classification and interpretation of breast cancer histopathology images," *Sci. Rep.*, vol. 15, no. 1, pp. 1–25, 2025, doi: 10.1038/s41598-025-06853-6.
- [15] S. Tummala, V. S. G. Thadikemalla, S. Kadry, M. Sharaf, and H. T. Rauf, "EfficientNetV2 Based Ensemble Model for Quality Estimation of Diabetic Retinopathy Images from DeepDRID," *Diagnostics*, vol. 13, no. 4, 2023, doi: 10.3390/diagnostics13040622.
- [16] Y. NA, "Efficientnet v2-s Architecture," <https://www.researchgate.net>. Accessed: Nov. 02, 2025. [Online]. Available: https://www.researchgate.net/figure/Efficientnet-v2-s-Architecture-a-represents-the-schematic-of-the-model-from-image-input_fig4_373782480
- [17] S. Kevin, S. K. Singh, M. P. Kadam, and M. S. Patil, "RESEARCH IN ENGINEERING MANAGEMENT CATARACT EYE DETECTION USING DEEP LEARNING MODELS," pp. 1092–1101, 2025.
- [18] J. G. Purnomo, S. Birowo, and M. A. Maulana, "Identifikasi Malaria Pada Citra Darah Dengan Convolutional Neural Network," *bit-Tech*, vol. 7, no. 2, pp. 406–412, 2024, doi: 10.32877/bt.v7i2.1828.
- [19] S. Karthik, M., & Dane, "APOTOS 2019 Blindness Detection," <https://www.kaggle.com>. Accessed: Feb. 07, 2025. [Online]. Available: <https://kaggle.com/competitions/aptos2019-blindness-detection>
- [20] A. Noriega *et al.*, "Screening diabetic retinopathy using an automated retinal image analysis system in independent and assistive use cases in mexico: Randomized controlled trial," *JMIR Form. Res.*, vol. 5, no. 8, pp. 1–12, 2021, doi: 10.2196/25290.