

IDENTIFIKASI PENYAKIT MATA PADA CITRA FUNDUS MENGUNAKAN MODEL HIBRIDA TRANSFER LEARNING–XGBOOST

Fajri Kurniawan Adiza¹, Yuhandri², Sumijan³

¹Informatika Dan Statistik Provinsi Sumatera Barat, Padang

^{2,3}Universitas Putra Indonesia YPTK Padang, Sumatera Barat

e-mail: ¹fajribiza@gmail.com, ²yuyu@upiypk.ac.id, ³sumijan@upiypk.ac.id

Abstract: Eye diseases constitute a major health concern that may lead to visual impairment or blindness if not detected and treated at an early stage. Manual analysis of fundus images requires considerable time, accuracy, and ophthalmological expertise; therefore, a computational method capable of supporting automated and objective classification is needed. This study aims to develop a hybrid transfer learning and XGBoost model for classifying fundus images into four ocular condition categories. The methodology involved applying Contrast Limited Adaptive Histogram Equalization (CLAHE), resizing the images to 224×224 pixels, performing retina-specific augmentation on the training data, and stratifying the dataset into training, validation, and testing sets. An ImageNet-pretrained EfficientNetV2S model was employed through transfer learning and fine-tuning to extract image features. The resulting 256-dimensional embeddings, model output probabilities, and statistical image features were subsequently combined, standardized, and classified using XGBoost optimized through five-fold cross-validation. The dataset comprised 4,217 fundus images distributed across four classes: 1,038 cataract images, 1,098 diabetic retinopathy images, 1,007 glaucoma images, and 1,074 normal images, including 2,949 training, 631 validation, and 637 testing images. Evaluation on the 637 test images demonstrated that the proposed transfer learning–XGBoost model achieved an accuracy of 92.78%, macro-precision of 0.93, macro-recall of 0.93, macro F1-score of 0.9265, and one-vs-rest ROC-AUC of 0.9871. The proposed model outperformed the standalone EfficientNetV2S model, which achieved an accuracy of 90.27%, by 2.51 percentage points, with class-specific F1-scores of 0.96 for cataract, 0.99 for diabetic retinopathy, 0.87 for glaucoma, and 0.88 for normal. These findings demonstrate that integrating transfer learning with XGBoost effectively improves eye disease classification and has the potential to support early identification using fundus images, although it is not intended to replace clinical diagnosis by ophthalmologists.

Keyword: EfficientNetV2S, eye disease classification, fundus images, transfer learning, XGBoost.

Abstrak: Penyakit mata dapat menyebabkan gangguan penglihatan hingga kebutaan apabila tidak dideteksi dan ditangani sejak dini. Analisis citra fundus secara manual memerlukan waktu, ketelitian, dan keahlian dokter, sehingga dibutuhkan metode klasifikasi otomatis yang objektif. Penelitian ini bertujuan mengembangkan model hibrida transfer learning dan XGBoost untuk mengklasifikasikan citra fundus ke dalam empat kelas kondisi mata. Metode penelitian mencakup penerapan CLAHE untuk meningkatkan kontras, resize citra menjadi 224×224 piksel, augmentasi data pelatihan, serta pembagian data secara terstratifikasi. EfficientNetV2S berbobot ImageNet digunakan sebagai pengekstrak fitur, sedangkan embedding 256 dimensi, probabilitas keluaran model, dan fitur statistik citra digabungkan, distandardisasi, kemudian diklasifikasikan menggunakan XGBoost yang dioptimasi melalui validasi silang lima fold. Dataset terdiri atas 4.217 citra fundus dalam kelas cataract, diabetic retinopathy, glaucoma, dan normal, yang dibagi menjadi 2.949 data latih, 631 data validasi, dan 637

data uji. Hasil pengujian model transfer learning–XGBoost pada 637 citra uji menghasilkan accuracy sebesar 92,78%, macro precision 0,93, macro recall 0,93, macro F1-score 0,9265, dan ROC-AUC OvR 0,9871. Kinerja tersebut melampaui model EfficientNetV2S standalone yang memperoleh accuracy 90,27% sebesar 2,51 poin persentase, dengan F1-score per kelas sebesar 0,96 untuk cataract, 0,99 untuk diabetic retinopathy, 0,87 untuk glaucoma, dan 0,88 untuk normal. Model ini berpotensi mendukung tenaga medis dalam melakukan identifikasi awal penyakit mata secara lebih cepat, efisien, dan objektif, tetapi tidak dimaksudkan untuk menggantikan diagnosis dokter.

Kata kunci: citra fundus, EfficientNetV2S, klasifikasi penyakit mata, transfer learning, XGBoost.

PENDAHULUAN

Gangguan penglihatan tetap menjadi masalah kesehatan publik karena beban penyakit meningkat seiring pertumbuhan penduduk, penuaan, diabetes, dan ketimpangan akses terhadap pelayanan mata. Foto fundus menyediakan dokumentasi noninvasif mengenai papil optik, makula, pembuluh darah, serta perubahan warna dan lesi retina, sehingga merupakan sumber data penting untuk pengembangan skrining berbantuan kecerdasan buatan. Tinjauan sistematis terhadap algoritma berbasis foto fundus menunjukkan bahwa pemanfaatan AI berkembang untuk mendukung deteksi penyakit retina, namun kinerja tetap dipengaruhi kualitas citra, variasi populasi, dan karakteristik perangkat akuisisi (Grzybowski et al., 2024). Dalam konteks retinopati diabetik, evaluasi dunia nyata juga menegaskan bahwa performa sistem otomatis perlu dipahami bersama karakteristik data dan alur klinis, bukan hanya melalui satu nilai accuracy (Tahir et al., 2025).

Katarak, retinopati diabetik, dan glaukoma memiliki karakteristik visual yang berbeda tetapi sebagian dapat menunjukkan variasi intra-kelas dan kemiripan antar-kelas. Kondisi tersebut membuat klasifikasi multikelas lebih menantang dibandingkan deteksi satu penyakit. Transfer learning menjadi pilihan yang relevan karena representasi visual yang telah dipelajari dari dataset berskala besar dapat digunakan sebagai

titik awal untuk data medis yang jumlahnya lebih terbatas. Penelitian klasifikasi penyakit mata berbasis model prelatih menunjukkan bahwa fine-tuning lapisan akhir dapat menyesuaikan fitur umum ImageNet menjadi representasi yang lebih spesifik terhadap citra fundus (Manne palli et al., 2024). EfficientNetV2S dipilih sebagai backbone karena menyediakan representasi fitur yang kuat dengan desain Fused-MBConv dan MBConv yang efisien, serta dapat diadaptasi melalui fine-tuning bertahap.

Prapemrosesan juga memengaruhi kualitas representasi. Pada penelitian ini, CLAHE digunakan untuk meningkatkan kontras lokal sebelum citra disesuaikan ke ukuran masukan model. CLAHE membatasi penguatan histogram pada area lokal sehingga perbedaan intensitas pada struktur retina dapat dibuat lebih jelas tanpa melakukan equalization global yang agresif. Pendekatan pengayaan citra fundus berbasis CLAHE telah digunakan untuk memperbaiki visibilitas detail retina (Madhavi & Kumari, 2024). Augmentasi berupa rotasi ringan, pergeseran, zoom, perubahan kecerahan, dan horizontal flip diterapkan hanya pada data pelatihan untuk menambah variasi tanpa mencemari data validasi maupun pengujian.

Selain classifier softmax pada CNN, fitur hasil pembelajaran mendalam dapat dipisahkan dari kepala klasifikasi dan digunakan sebagai masukan algoritma machine learning. XGBoost relevan untuk skenario tersebut karena dapat membangun batas keputusan nonlinier

pada representasi numerik dan menyediakan regularisasi melalui pengendalian kedalaman pohon, learning rate, subsampling, serta penalti model. Pendekatan CNN–XGBoost telah digunakan pada klasifikasi retinopati diabetik (Sushith et al., 2025), sedangkan Duan dan Tu (2025) menunjukkan bahwa deep feature dari DenseNet121 dapat digunakan oleh XGBoost untuk klasifikasi banyak kondisi retina. Laksana dan Yohannes (2026) kemudian menggabungkan fitur DenseNet121 dan EfficientNetV2S sebelum klasifikasi XGBoost, dan Ejaz et al. (2025) melaporkan manfaat concatenation beberapa representasi untuk diagnosis penyakit retina. Temuan tersebut menunjukkan potensi fusi fitur, sekaligus membuka ruang untuk mengevaluasi representasi yang lebih ringkas dengan satu backbone.

RESEARCH GAP DAN KEBARUAN PENELITIAN

Research gap penelitian ini terletak pada masih terbatasnya bukti mengenai apakah satu backbone EfficientNetV2S dapat menghasilkan representasi yang cukup kuat untuk klasifikasi empat kondisi mata ketika deep embedding tidak digunakan sendirian, tetapi diperkaya dengan probabilitas keluaran CNN dan statistik warna, kemudian dipisahkan dari classifier softmax dan diklasifikasikan menggunakan XGBoost.

Penelitian hibrida sebelumnya ada yang berfokus pada satu penyakit (Sushith et al., 2025), menggunakan backbone dan jenis citra yang berbeda (Duan & Tu, 2025), atau menggabungkan fitur dari lebih dari satu CNN (Laksana & Yohannes, 2026). Dengan demikian, pertanyaan yang belum dijawab secara langsung pada konfigurasi penelitian ini adalah apakah fusi fitur ringkas dari satu EfficientNetV2S dapat meningkatkan performa dibandingkan EfficientNetV2S standalone pada data uji yang sama, terutama pada kelas yang sulit seperti

glaukoma.

Kebaruan (novelty) penelitian ini bukan pada klaim bahwa kombinasi CNN dan XGBoost belum pernah digunakan, melainkan pada rancangan representasi 275 dimensi yang menggabungkan tiga sumber informasi: embedding 256 dimensi dari EfficientNetV2S, empat probabilitas keluaran CNN, dan 15 statistik warna RGB.

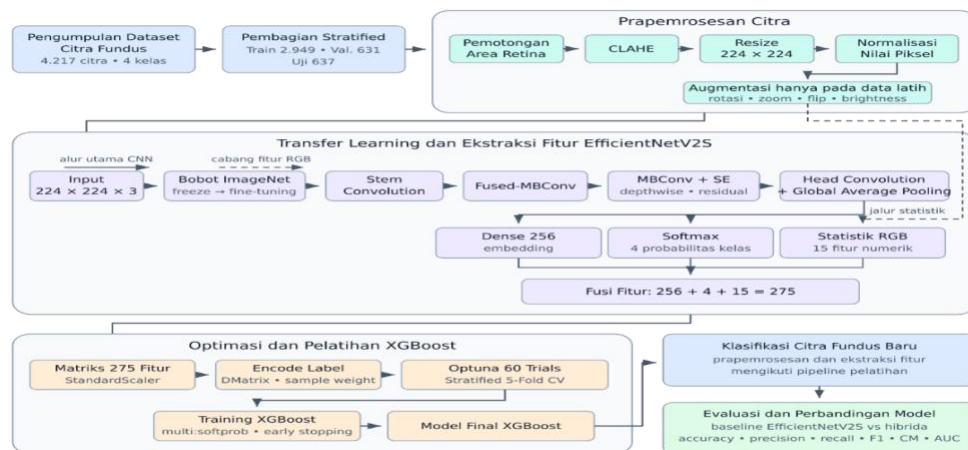
Representasi heterogen tersebut distandardisasi dan digunakan oleh XGBoost yang dioptimasi dengan Optuna serta stratified five-fold cross-validation. Kontribusi eksperimentalnya adalah perbandingan langsung terhadap EfficientNetV2S standalone dengan data holdout yang sama sehingga tambahan manfaat classifier XGBoost dapat diukur secara lebih jelas.

Berdasarkan gap tersebut, penelitian bertujuan membangun dan mengevaluasi model hibrida EfficientNetV2S–XGBoost untuk klasifikasi katarak, retinopati diabetik, glaukoma, dan normal, serta menilai peningkatan kinerja model hibrida terhadap baseline EfficientNetV2S standalone menggunakan accuracy, precision, recall, F1-score, dan ROC-AUC.

METODE

Kerangka penelitian dimulai dari pemeriksaan struktur dataset dan label, kemudian data dibagi secara terstratifikasi agar proporsi setiap kelas tetap terjaga (Ejaz et al., 2024). Tahap prapemrosesan mencakup pemotongan area retina, peningkatan kontras lokal, perubahan ukuran, normalisasi, dan augmentasi khusus data pelatihan (Härtinger & Steger, 2024).

Tahap pemodelan meliputi pelatihan serta fine-tuning EfficientNetV2S, pembentukan fitur gabungan, optimasi XGBoost, dan pengujian final menggunakan data holdout (Elkenawy et al., 2026). Berikut adalah kerangka alur penelitian.

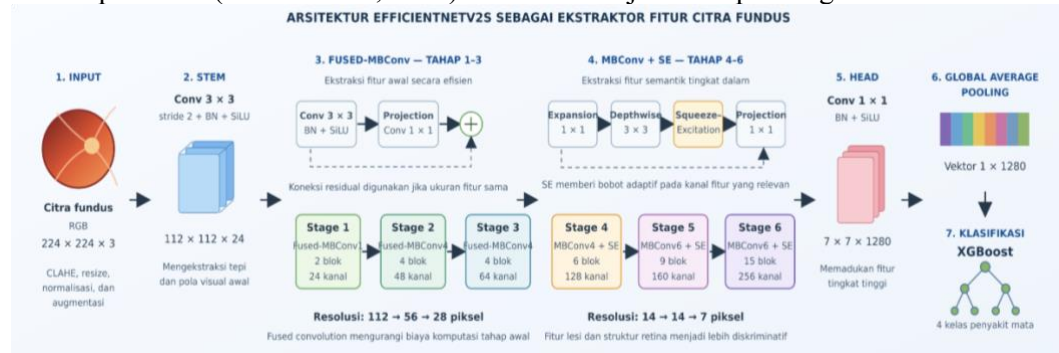


Gambar 1 Kerangka Penelitian

Arsitektur penelitian pada gambar 1 menghubungkan tahap prapemrosesan, pembelajaran fitur, dan klasifikasi dengan menggabungkan *embedding* 256 dimensi, empat probabilitas keluaran CNN, serta lima belas statistik warna RGB menjadi vektor berukuran 275 fitur yang selanjutnya menjadi masukan bagi XGBoost (Laksana & Yohannes, 2026). Selanjutnya pada blok EfficientNetV2S, citra diproses melalui *stem convolution*, beberapa *stage Fused-MBConv*, *stage MBConv* dengan *squeeze-and-excitation*, *head convolution*, dan *global average pooling*, kemudian diteruskan menuju lapisan *dense* 256 unit sebagai representasi numerik citra fundus (Alsohemi & Dardouri, 2025). Pada blok klasifikasi, fitur gabungan distandardisasi dan dikonversi ke DMATRIX, kemudian XGBoost dioptimalkan menggunakan Optuna dan *stratified five-fold cross-validation* sebelum model terbaik diuji pada data yang tidak pernah digunakan selama pelatihan (Zannah et al., 2024).

EfficientNetV2S

EfficientNetV2S adalah varian kecil dari keluarga EfficientNetV2 yang dirancang untuk memperoleh representasi visual kuat dengan proses pelatihan lebih cepat dan penggunaan parameter yang lebih efisien dibandingkan banyak CNN konvensional (Laksana & Yohannes, 2026). Dalam skema transfer learning, bobot hasil prapelatihan pada ImageNet digunakan sebagai titik awal agar pola umum seperti tepi, warna, tekstur, dan bentuk dapat ditransfer ke citra fundus, kemudian lapisan akhir disesuaikan dengan empat kelas penelitian (Mannepalli et al., 2024). EfficientNetV2S digunakan sebagai backbone berbasis transfer learning untuk mengekstraksi karakteristik penting dari citra fundus, seperti perubahan warna retina, bentuk diskus optik, pembuluh darah, dan pola lesi (Cho et al., 2024). Susunan arsitektur EfficientNetV2S yang diterapkan dalam penelitian ini ditunjukkan pada gambar 1 berikut:



Gambar 2 Arsitektur efficientnetv2s sebagai ekstraktor.

Berdasarkan Gambar 2, citra fundus berukuran $224 \times 224 \times 3$ diproses melalui lapisan stem, Fused-MBConv, dan MBConv dengan mekanisme Squeeze-and-Excitation untuk memperoleh fitur visual yang relevan (Sah et al., 2026). Bagian head menghasilkan peta fitur berukuran $7 \times 7 \times 1.280$, kemudian Global Average Pooling mengubahnya menjadi vektor fitur berdimensi 1.280 (Laksana & Yohannes, 2026). Vektor tersebut selanjutnya digunakan oleh XGBoost untuk mengklasifikasikan citra fundus ke dalam kelas katarak, retinopati diabetik, glaukoma, atau normal (Laksana & Yohannes, 2026).

EfficientNetV2S dirancang untuk meningkatkan efisiensi parameter dan mempercepat proses pelatihan melalui penggunaan blok Fused-MBConv dan MBConv (Ejaz et al., 2026). Secara matematis, proses transformasi fitur pada Fused-MBConv dapat dinyatakan melalui Persamaan 2 berikut:

$$Y = X + Wp * \phi(BN(Wf * X)) \quad (2)$$

Fused-MBConv menggabungkan proses ekspansi dan *depthwise convolution* menjadi konvolusi biasa sehingga proses pelatihan menjadi lebih cepat (Laksana & Yohannes, 2026). Blok ini terutama digunakan pada tahap awal EfficientNetV2S untuk mengekstraksi pola dasar citra fundus, seperti tekstur, tepi, dan pembuluh darah (Alsohemi & Dardouri, 2025).

Setelah melalui rangkaian blok Fused-MBConv dan MBConv, EfficientNetV2S menghasilkan *feature map* yang merepresentasikan karakteristik tingkat tinggi dari citra fundus, seperti pola pembuluh darah, tekstur retina, struktur saraf optik, dan karakteristik kelainan mata (Alsohemi & Dardouri, 2025).

Feature map tersebut kemudian diproses menggunakan *Global Average Pooling* dengan menghitung nilai rata-rata pada setiap kanal fitur, sehingga representasi citra menjadi lebih ringkas

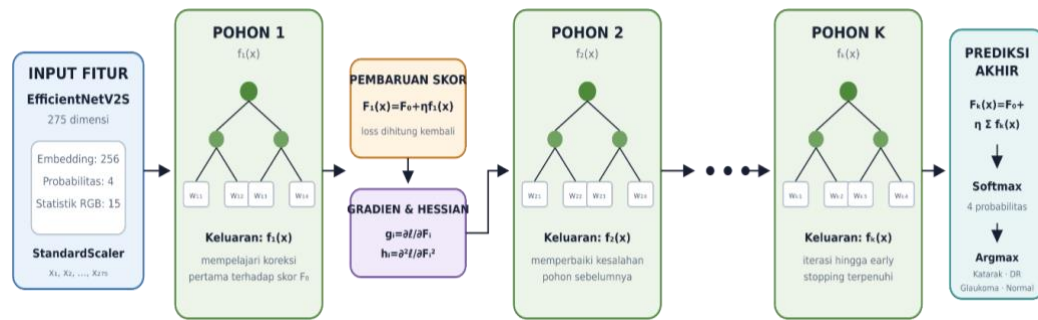
tanpa menambah parameter pelatihan (Laksana & Yohannes, 2026). Secara matematis, proses *Global Average Pooling* dinyatakan pada Persamaan 3:

$$f_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_{i,j,c} \quad (3)$$

Global Average Pooling menghitung rata-rata seluruh nilai pada setiap kanal feature map (Laksana & Yohannes, 2026). Vektor tersebut merepresentasikan karakteristik global citra fundus dan selanjutnya dapat digunakan sebagai masukan bagi. Secara arsitektural, citra berukuran $224 \times 224 \times 3$ memasuki stem convolution, dilanjutkan oleh beberapa stage Fused-MBConv pada bagian awal, stage MBConv dengan *depthwise convolution* dan *squeeze-and-excitation* pada bagian lebih dalam, kemudian head convolution serta global average pooling yang merangkum feature map menjadi vektor numerik (Laksana & Yohannes, 2026).

Extreme Gradient Boosting

Extreme Gradient Boosting merupakan algoritma ensemble berbasis decision tree yang membangun pohon secara berurutan, dengan setiap pohon baru diarahkan untuk memperbaiki kesalahan prediksi gabungan pohon-pohon sebelumnya (Qureshi, 2025). Proses pembelajarannya meminimalkan fungsi objektif yang menggabungkan loss prediksi dan penalti kompleksitas model dengan memanfaatkan informasi gradien orde pertama serta kedua, sehingga pemilihan split dapat dilakukan secara efisien sekaligus mengendalikan overfitting (Qureshi, 2025). XGBoost digunakan setelah ekstraksi fitur CNN karena mampu memodelkan hubungan nonlinier pada vektor deep feature, memberikan bobot pada sampel, dan telah diterapkan pada klasifikasi penyakit retina berbasis citra fundus (Duan & Tu, 2025). Berikut ini arsitektur proses pembelajaran pada XGBoost yang dapat dilihat pada Gambar 3:



Gambar 3 Arsitektur XGBoost.

Fungsi objektif Gambar 3 pada XGBoost digunakan untuk mengukur tingkat kesalahan prediksi model sekaligus mengontrol kompleksitas model dan mengurangi resiko overfitting melalui regularisasi (Qureshi, 2025). Berikut ini fungsi objektif yang dapat dilihat pada persamaan 4 berikut:

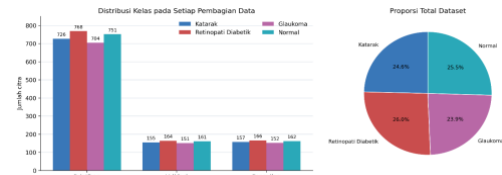
$$Obj^{(t)} = \sum_{i=1}^n l(y_i, y_i) + \sum_{k=1}^t \Omega(f_k) \quad (4)$$

Pada persamaan 4 menyatakan fungsi loss atau kesalahan yang digunakan untuk mengukur selisih nilai prediksi model dengan nilai target, menunjukkan fungsi regularisasi agar dapat mengurangi risiko overfitting (Qureshi, 2025). Regularisasi pada XGBoost digunakan untuk mengontrol kompleksitas decision tree agar model tidak terlalu menyesuaikan data training (Qureshi, 2025).

HASIL DAN PEMBAHASAN

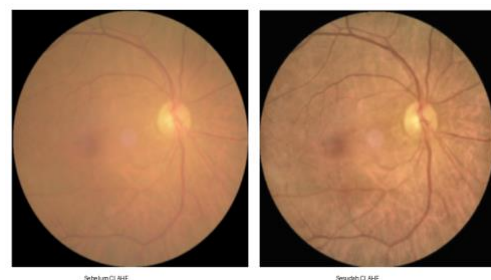
Dataset dan Prapemrosesan

Dataset berjumlah 4.217 citra fundus: 1.038 katarak, 1.098 retinopati diabetik, 1.007 glaukoma, dan 1.074 normal. Pembagian stratified menghasilkan 2.949 data latih, 631 validasi, dan 637 uji. Rasio kelas terbesar terhadap terkecil hanya sekitar 1,09, sehingga distribusi relatif seimbang dan evaluasi macro F1 tetap relevan untuk menilai performa lintas kelas.



Gambar 4 Distribusi citra pada data latih, validasi, dan uji.

Visualisasi pada Gambar 4 memperlihatkan bahwa jumlah citra pada setiap kelas memiliki proporsi yang hampir seimbang. Selanjutnya pada tahapan CLAHE meningkatkan kontras lokal tanpa mengubah struktur utama retina. Setelah pemotongan area retina dan CLAHE, citra diubah menjadi 224×224 piksel. Augmentasi hanya diterapkan pada data latih untuk menambah variasi sekaligus menjaga validasi dan data uji tetap bebas dari transformasi acak. Berikut citra fundus sebelum dan sesudah CLAHE.



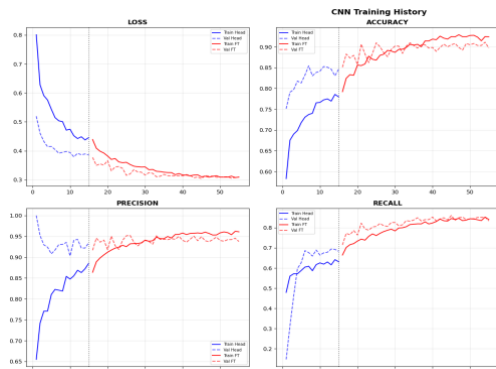
Gambar 5 Citra fundus sebelum dan sesudah CLAHE.

Gambar 5 menunjukkan bahwa penerapan CLAHE menghasilkan peningkatan kontras pada bagian-bagian penting citra fundus tanpa mengubah struktur utama objek. Pembuluh darah retina menjadi lebih jelas dan perbedaan intensitas antara area retina, saraf optik,

serta bagian yang mengalami kelainan dapat diamati dengan lebih baik

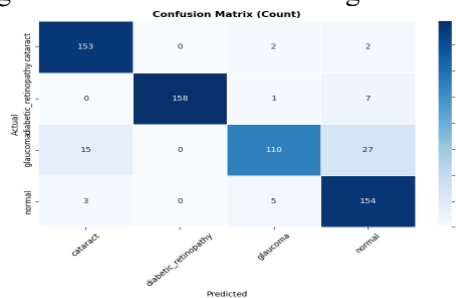
Pelatihan dan Baseline EfficientNetV2S

Pada tahap awal dengan backbone dibekukan, akurasi validasi tertinggi mencapai 85,42%. Setelah 30% lapisan akhir dibuka untuk fine-tuning dengan learning rate lebih kecil, akurasi validasi meningkat menjadi 91,13%, sedangkan akurasi pelatihan berada pada 92,47%. Selisih sekitar 1,34 poin persentase menunjukkan penyesuaian backbone meningkatkan representasi citra fundus tanpa kesenjangan train-validation yang besar.



Gambar 6 Hasil pelatihan dan validasi EfficientNetV2S.

Pada 637 citra uji, EfficientNetV2S standalone memperoleh accuracy 90,27%, macro precision 0,91, macro recall 0,90, dan macro F1-score sekitar 0,90. F1-score per kelas adalah 0,93 untuk katarak, 0,98 untuk retinopati diabetik, 0,81 untuk glaukoma, dan 0,88 untuk normal. Glaukoma menjadi kelas tersulit: confusion matrix menunjukkan recall hanya 0,72, sehingga sekitar 28% citra glaukoma belum dikenali dengan benar.

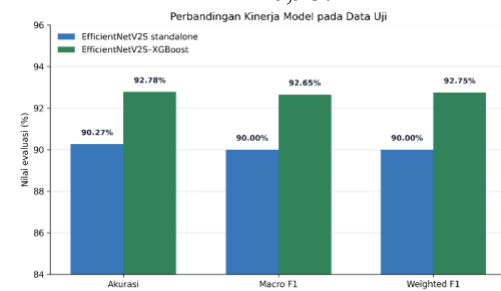


Gambar 7 Confusion matrix EfficientNetV2S standalone pada data uji.

Hasil diatas memperlihatkan dominasi nilai pada diagonal utama confusion matrix, yang menunjukkan bahwa sebagian besar citra berhasil ditempatkan pada kelas yang benar

Kinerja Model Hibrida EfficientNetV2S–XGBoost

Model hibrida menggunakan vektor 275 fitur yang terdiri atas 256 deep embedding, empat probabilitas CNN, dan 15 statistik warna RGB. Optimasi 60 percobaan Optuna dengan stratified five-fold cross-validation menghasilkan macro F1 terbaik 0,9620 pada data pengembangan. Pada data uji yang tetap ditahan, model hibrida mencapai accuracy 92,78%, macro F1-score 0,9265, weighted F1-score 0,9275, dan ROC-AUC one-versus-rest 0,9871.



Gambar 8 Perbandingan kinerja baseline dan model hibrida.

Dibandingkan baseline, accuracy meningkat 2,51 poin persentase, macro F1 sekitar 2,65 poin, dan weighted F1 sekitar 2,75 poin. Kenaikan accuracy tersebut setara dengan sekitar 16 citra tambahan yang diklasifikasikan dengan benar. Kedekatan macro dan weighted F1 pada model hibrida juga menunjukkan bahwa peningkatan tidak hanya didorong oleh satu kelas tertentu.

Pembahasan terhadap Research Gap

Temuan tersebut menjawab research gap pada pendahuluan. Satu backbone EfficientNetV2S ternyata dapat menghasilkan representasi yang cukup informatif untuk empat kondisi ketika deep embedding diperkaya probabilitas CNN dan statistik warna, kemudian dipisahkan oleh XGBoost. Berbeda dari

kerangka hibrida yang hanya berfokus pada retinopati diabetik (Sushith et al., 2025) atau menggabungkan beberapa backbone sebelum XGBoost (Laksana & Yohannes, 2026), rancangan ini menggunakan fusi heterogen yang lebih ringkas dari satu backbone.

Secara komputasional, EfficientNetV2S berfungsi membentuk representasi visual tingkat tinggi, sedangkan XGBoost mempelajari interaksi nonlinier pada fitur gabungan yang tidak harus identik dengan batas keputusan softmax. Peningkatan 2,51 poin persentase mendukung manfaat empiris kombinasi tersebut terhadap baseline pada dataset ini. Namun, hasil belum membuktikan keunggulan universal karena belum dilengkapi ablation study, uji signifikansi berpasangan, dan validasi eksternal; aspek tersebut dibahas sebagai keterbatasan dan arah penelitian berikutnya.

Evaluasi Model

Kinerja model diukur menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, *macro F1-score*, *weighted F1-score*, ROC-AUC, dan *confusion matrix* untuk memberikan penilaian yang menyeluruh terhadap kemampuan klasifikasi pada empat kelas. Evaluasi dilakukan terhadap EfficientNetV2S *standalone* sebagai model dasar dan EfficientNetV2S–XGBoost sebagai model hibrida yang diusulkan.

EfficientNetV2S *standalone* memperoleh akurasi sebesar 90,27%, *macro F1-score* sebesar 90,00%, dan *weighted F1-score* sebesar 90,00%. Setelah fitur hasil ekstraksi EfficientNetV2S diklasifikasikan menggunakan XGBoost, akurasi meningkat menjadi 92,78%, *macro F1-score* menjadi 92,65%, *weighted F1-score* menjadi 92,75%, dan ROC-AUC mencapai 98,71%. Peningkatan akurasi sebesar 2,51 poin persentase tersebut setara dengan tambahan sekitar 16 citra yang berhasil diklasifikasikan dengan benar, sedangkan selisih yang kecil antara

macro F1-score dan *weighted F1-score* menunjukkan bahwa performa model relatif konsisten pada kelas dengan jumlah sampel yang berbeda.

Berdasarkan hasil evaluasi per kelas, model EfficientNetV2S–XGBoost menghasilkan *precision* dan *recall* masing-masing sebesar 0,96 dan 0,97 pada kelas katarak, 0,99 dan 0,99 pada retinopati diabetik, 0,92 dan 0,82 pada glaukoma, serta 0,85 dan 0,93 pada kelas normal. Kelas glaukoma memiliki nilai *recall* paling rendah, yang menunjukkan bahwa sebagian citra glaukoma masih diklasifikasikan sebagai kelas lain karena kemiripan karakteristik visual, terutama dengan citra normal. Secara keseluruhan, hasil tersebut membuktikan bahwa penggunaan XGBoost mampu meningkatkan pemisahan kelas berdasarkan fitur yang diekstraksi EfficientNetV2S.

SIMPULAN

Penelitian ini berhasil menerapkan pendekatan transfer learning menggunakan EfficientNetV2S dan XGBoost untuk mengklasifikasikan citra fundus ke dalam empat kelas, yaitu katarak, retinopati diabetik, glaukoma, dan normal. Tahap prapemrosesan menggunakan CLAHE membantu meningkatkan kontras citra, sedangkan EfficientNetV2S berperan dalam mengekstraksi karakteristik visual yang selanjutnya diklasifikasikan menggunakan XGBoost. Berdasarkan hasil pengujian, model hibrida memperoleh akurasi sebesar 92,78%, sehingga menunjukkan kemampuan yang baik dalam membedakan setiap kelas penyakit mata. Pendekatan ini berpotensi digunakan sebagai sistem pendukung untuk membantu proses skrining awal penyakit mata berbasis citra fundus. Penelitian berikutnya dapat dikembangkan menggunakan dataset yang lebih besar dan beragam serta melakukan optimasi hiperparameter untuk

meningkatkan kemampuan generalisasi model.

Saran Penelitian Selanjutnya

Penelitian selanjutnya disarankan melakukan validasi eksternal pada dataset dari institusi atau kamera yang berbeda untuk menguji generalisasi; menerapkan split berbasis pasien apabila identitas pasien tersedia agar citra dari pasien yang sama tidak masuk ke set berbeda; melakukan ablation study untuk mengukur kontribusi embedding, probabilitas CNN, dan statistik RGB; serta menguji signifikansi perbedaan model menggunakan McNemar atau bootstrap confidence interval. Peningkatan sensitivitas glaukoma perlu menjadi prioritas melalui eksplorasi region-of-interest pada diskus optik, resolusi masukan yang lebih tinggi, atau pendekatan multi-scale.

DAFTAR PUSTAKA

- Alsohemi, R., & Dardouri, S. (2025). Fundus image-based eye disease detection using EfficientNetB3 architecture. *Journal of Imaging*, 11(8), 279. <https://doi.org/10.3390/jimaging11080279>
- Cho, Y.-S., Song, H.-J., Han, J.-H., & Kim, Y.-S. (2024). Attention mechanism-based glaucoma classification model using retinal fundus images. *Sensors*, 24(14), 4684. <https://doi.org/10.3390/s24144684>.
- Duan, M.-M., & Tu, X. (2025). Deep learning-based classification of multiple fundus diseases using ultra-widefield images. *Frontiers in Cell and Developmental Biology*, 13, 1630667. <https://doi.org/10.3389/fcell.2025.1630667>
- Ejaz, S., Baig, R., Ashraf, Z., Alnfai, M. M., Alnahari, M. M., & Alotaibi, R. M. (2024). A deep learning framework for the early detection of multi-retinal diseases. *PLOS ONE*, 19(7), e0307317. <https://doi.org/10.1371/journal.pone.0307317>
- Ejaz, S., Zia, H. U., Majeed, F., Shafique, U., Carvajal-Altamiranda, S., Lipari, V., & Ashraf, I. (2025). Fundus image classification using feature concatenation for early diagnosis of retinal disease. *Digital Health*, 11, 20552076251328120. <https://doi.org/10.1177/20552076251328120>.
- Grzybowski, A., Jin, K., Zhou, J., Pan, X., Wang, M., Ye, J., & Wong, T. Y. (2024). Retina fundus photograph-based artificial intelligence algorithms in medicine: A systematic review. *Ophthalmology and Therapy*, 13(8), 2125–2149. <https://doi.org/10.1007/s40123-024-00981-4>
- Härtinger, P., & Steger, C. (2024). Adaptive histogram equalization in constant time. *Journal of Real-Time Image Processing*, 21, 93. <https://doi.org/10.1007/s11554-024-01465-1>
- Khalid, M., Sajid, M. Z., Youssef, A., Khan, N. A., Hamid, M. F., & Abbas, F. (2024). CAD-EYE: An automated system for multi-eye disease classification using feature fusion with deep learning models and fluorescence imaging for enhanced interpretability. *Diagnostics*, 14(23), 2679. <https://doi.org/10.3390/diagnostics14232679>
- Laksana, J. P., & Yohannes (2026). Feature-level fusion of DenseNet121 and EfficientNetV2 with XGBoost for multi-class retinal classification. *Sinkron: Jurnal dan Penelitian Teknik Informatika*, 10(1), 439–448. <https://doi.org/10.33395/sinkron.v10i1.1567>

- Madhavi, V. V., & Kumari, P. L. S. (2024). Enrichment of retinal fundus images using EN-CLAHE and Auto-CLAHE methods. *International Journal of Intelligent Systems and Applications in Engineering*, 12(3), 1213.1221.<https://ijisae.org/index.php/IJISAE/article/view/5401>
- Mannepalli, P. K., Baghela, V. D. S., Agrawal, A., Johri, P., Dubey, S. S., & Parmar, K. (2024). Transfer learning for automated classification of eye disease in fundus images from pretrained model. *Traitement du Signal*, 41(5), 24592470.<https://doi.org/10.18280/ts.410520>
- Qureshi, I. (2025). AdaptiveSwin-CNN: Adaptive Swin-CNN framework with self-attention fusion for robust multi-class retinal disease diagnosis. *AI*, 6(2), 28.<https://doi.org/10.3390/ai6020028>
- Sushith, M., Sathiya, A., Kalaipoonguzhali, V., & Sathya, V. (2025). A hybrid deep learning framework for early detection of diabetic retinopathy using retinal fundus images. *Scientific Reports*, 15, 15166.
<https://doi.org/10.1038/s41598-025-99309-w>
- Tahir, H. N., Ullah, N., Tahir, M., Domnic, I. S., Prabhakar, R., Meerasa, S. S., Abdelneam, A. I., Tahir, S., & Ali, Y. (2025). Artificial intelligence versus manual screening for the detection of diabetic retinopathy: A comparative systematic review and meta-analysis. *Frontiers in Medicine*, 12, 1519768.
<https://doi.org/10.3389/fmed.2025.1519768>