

KLASIFIKASI KATARAK MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK DENGAN ARSITEKTUR EFFICIENTNETB5 PADA CITRA FUNDUS RETINA

Wahyu Putra Pratama^{1*}, Dwi Arman Prasetya², Andri Fauzan Adziima³

¹Sains Data, UPN Veteran Jawa Timur, Surabaya, Indonesia

²Sains Data, UPN Veteran Jawa Timur, Surabaya, Indonesia

³Sains Data, UPN Veteran Jawa Timur, Surabaya, Indonesia

*22083010092@student.upnjatim.ac.id

ABSTRAK

Katarak merupakan penyebab utama kebutaan yang masih banyak terlambat terdeteksi karena pemeriksaan manual membutuhkan keahlian dokter mata dan berpotensi menimbulkan subjektivitas. Pada banyak wilayah dengan keterbatasan tenaga medis, keterlambatan diagnosis menjadi tantangan besar yang berdampak pada meningkatnya risiko kebutaan. Untuk itu, dibutuhkan sistem deteksi otomatis berbasis kecerdasan buatan yang mampu memberikan hasil cepat dan akurat. Penelitian ini mengkaji permasalahan rendahnya akurasi dan kemampuan generalisasi model CNN konvensional dalam klasifikasi citra fundus. Sebagai solusi, penelitian ini menerapkan arsitektur *EfficientNet-B5* yang dikenal memiliki efisiensi tinggi dan kemampuan ekstraksi fitur yang kuat melalui *compound scaling*. *Dataset* citra fundus diproses melalui tahapan *preprocessing*, augmentasi, pembagian data, pelatihan model, serta evaluasi menggunakan metrik akurasi, presisi, *recall*, dan *F1-score*. Hasil pengujian menunjukkan bahwa *EfficientNet-B5* memberikan performa yang sangat baik pada data validasi, dengan akurasi terbaik mencapai 96,26%, diikuti metrik precision 0.9720, recall 0.9541, specificity 0.9714, dan F1-score 0.9630. Confusion matrix juga menunjukkan bahwa sebagian besar citra berhasil diklasifikasikan dengan benar, menandakan kemampuan model dalam melakukan generalisasi pada data baru. Temuan ini memperlihatkan bahwa *EfficientNet-B5* efektif digunakan sebagai model tunggal untuk klasifikasi katarak tanpa memerlukan metode interpretabilitas tambahan. Urgensi penelitian ini terletak pada kebutuhan sistem deteksi katarak yang ringan, cepat, dan mudah diimplementasikan di berbagai fasilitas kesehatan. Inovasi penelitian ini adalah pemanfaatan *EfficientNet-B5* sebagai model tunggal yang mampu meningkatkan stabilitas dan akurasi dibandingkan model CNN konvensional, meskipun tanpa teknik interpretabilitas tambahan. Tujuan penelitian ini adalah mengembangkan model klasifikasi katarak berbasis *EfficientNet-B5* yang akurat dan dapat digunakan sebagai alat bantu skrining awal.

Kata kunci: Klasifikasi, Katarak, *Deep Learning*, CNN, *EfficientNet*.

1 PENDAHULUAN

Katarak merupakan salah satu penyebab utama kebutaan di dunia dan dapat dicegah melalui deteksi dini serta penanganan yang tepat. *World Health Organization* melaporkan lebih dari 65 juta kasus kebutaan akibat katarak secara global (Health Organization, 2019). Di Indonesia, katarak menyumbang sekitar 80% kasus kebutaan nasional sehingga menjadi masalah kesehatan masyarakat yang sangat mendesak (PROFIL KESEHATAN INDONESIA 2022,.). Secara klinis, katarak ditandai oleh kekeruhan progresif pada lensa mata yang menghalangi cahaya masuk ke retina, dan sering berkembang akibat penuaan, diabetes, trauma, atau penggunaan kortikosteroid jangka panjang (Klein & Klein, 2018). Meskipun dapat diobati melalui pembedahan, banyak kasus tidak terdeteksi pada tahap awal karena keterbatasan akses terhadap dokter mata dan rendahnya kesadaran masyarakat untuk melakukan pemeriksaan rutin.

Proses diagnosis katarak secara konvensional masih bergantung pada pemeriksaan *slit-lamp* oleh dokter spesialis, yang sangat dipengaruhi oleh tingkat keahlian dan pengalaman klinis (Mardjuki et al., 2025). Kondisi ini menyebabkan adanya variasi *inter-observer* dan potensi keterlambatan diagnosis, terutama di daerah dengan rasio dokter mata yang masih rendah ((Health Organization, 2019)). Kondisi ini menimbulkan tantangan serius di negara berkembang seperti Indonesia, di mana rasio dokter spesialis mata hanya sekitar 1:250.000 penduduk, jauh di bawah standar WHO yang merekomendasikan 1 dokter mata per 50.000 penduduk (Salsabila et al., 2024). Seiring meningkatnya kebutuhan akan deteksi dini yang lebih cepat dan akurat, citra fundus retina mulai digunakan sebagai alternatif non-invasif untuk membantu identifikasi berbagai penyakit mata, termasuk katarak. Citra fundus memberikan representasi visual retina dan lensa mata, yang dapat menjadi sumber informasi penting dalam proses klasifikasi otomatis.

Kemajuan teknologi kecerdasan buatan, khususnya *deep learning*, telah membuka peluang untuk mengembangkan sistem deteksi katarak otomatis berbasis *Convolutional Neural Network* (CNN). *Convolutional Neural Network* (CNN) merupakan salah satu model paling representatif dalam bidang *deep learning* yang telah mencapai banyak terobosan, khususnya dalam visi komputer. CNN termasuk jenis *feedforward neural network* yang mampu mengekstraksi fitur langsung dari data melalui struktur konvolusi tanpa perlu dilakukan *feature extraction* secara manual, berbeda dengan metode tradisional (Gulshan et al., 2016). Dalam beberapa tahun terakhir, *Convolutional Neural Network* (CNN) telah menjadi pendekatan yang dominan dalam analisis citra medis, termasuk dalam klasifikasi penyakit mata. CNN terbukti mampu mengekstraksi fitur spasial dari citra secara hierarkis dan efisien, memungkinkan model mengenali pola-pola visual yang sulit terdeteksi oleh manusia (Lecun et al., 2015).

Sejumlah penelitian menunjukkan bahwa CNN mampu mempelajari pola visual secara efektif dari citra medis dan memberikan hasil klasifikasi yang kompetitif (Junayed et al., 2021). Di antara berbagai arsitektur CNN, *EfficientNet* merupakan salah satu model modern yang menawarkan keseimbangan optimal antara akurasi dan efisiensi komputasi melalui *compound scaling* (Tan & Le, 2019). Namun, penelitian yang secara khusus mengevaluasi performa *EfficientNetB5* sebagai model tunggal untuk klasifikasi katarak pada citra fundus masih terbatas, sehingga diperlukan studi yang lebih fokus untuk menguji potensi arsitektur tersebut.

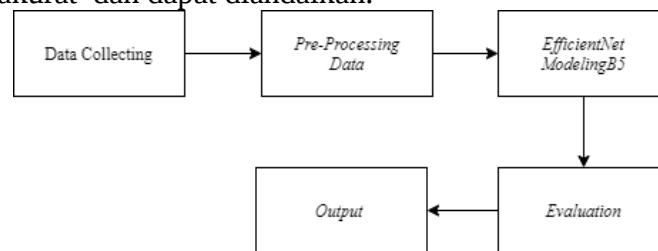
Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengembangkan dan mengevaluasi model klasifikasi katarak menggunakan arsitektur *EfficientNet-B5* pada citra fundus retina. Penelitian ini diharapkan mampu memberikan kontribusi dalam meningkatkan akurasi klasifikasi katarak dan memperluas pemanfaatan model CNN yang efisien pada bidang oftalmologi. Dengan demikian, penelitian ini tidak hanya berupaya menjawab keterbatasan

metode diagnosis manual, tetapi juga memberikan dasar ilmiah untuk pengembangan sistem pendukung keputusan medis berbasis kecerdasan buatan di masa depan.

2 METODE

2.1 Jenis Penelitian

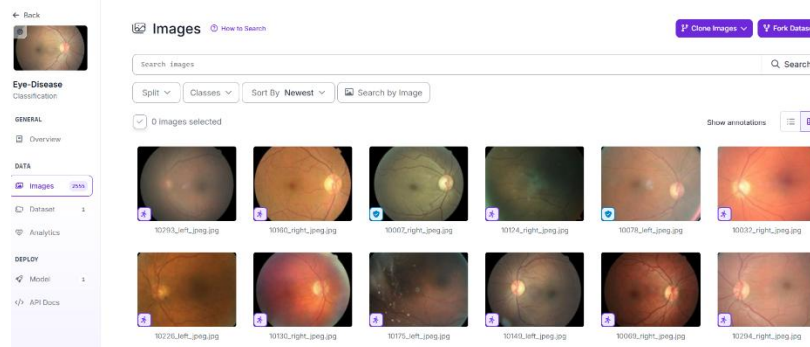
Pada penelitian ini digunakan pendekatan kuantitatif eksperimental untuk menguji performa model algoritma EfficientNetB5 dalam mengklasifikasikan fundus citra mata normal dan mata katarak. Penelitian ini dilakukan melalui beberapa tahapan utama sebagaimana ditunjukkan pada Gambar 2.1. Tahapan-tahapan tersebut dirancang untuk memastikan hasil klasifikasi yang akurat dan dapat diandalkan.



Gambar 2.1 Alur Pengerjaan

2.2 Data Collecting

Tahapan pengumpulan data dalam penelitian ini diawali dari pasien pengidap penyakit mata yang menjadi sumber utama pengambilan data citra medis. Citra fundus retina dari pasien-pasien tersebut kemudian dikumpulkan dan dihimpun oleh kontributor melalui proses kurasi data medis, lalu dipublikasikan dalam bentuk dataset terbuka di platform Roboflow dengan judul Eye Disease Classification 2024.



Gambar 2.2. Data Citra fundus Mata

Gambar 2.2. menunjukkan tampilan publikasi *dataset* pada platform Roboflow. *Dataset* ini berisi citra fundus mata yang telah diklasifikasikan berdasarkan kondisi kesehatan mata, seperti katarak, glaukoma, retinopati diabetik, dan kondisi normal, yang selanjutnya digunakan sebagai data utama dalam penelitian ini.

2.3 Pre-Processing Data

Tahap berikutnya adalah data preprocessing, Tahap pra-pemrosesan data merupakan langkah krusial dalam penelitian ini karena berfungsi sebagai fondasi untuk memastikan kualitas data sebelum dimasukkan ke dalam proses pelatihan model. Tujuan dari tahap ini adalah untuk mengolah data mentah agar menjadi lebih terstruktur, bersih, dan siap digunakan oleh model pembelajaran mesin. Data yang tidak diolah dengan benar dapat menyebabkan bias, kesalahan klasifikasi, atau performa model yang rendah. Oleh karena itu, tahapan pra-

pemrosesan dirancang untuk meningkatkan konsistensi, kelengkapan, dan distribusi data secara optimal.

Langkah pertama adalah data cleaning, yaitu menghapus citra yang buram, terlalu gelap, atau rusak. Hal ini dilakukan agar hanya data dengan kualitas baik yang digunakan, sehingga model dapat belajar dengan lebih efektif. Selanjutnya *Class balancing* adalah proses untuk menyamakan jumlah data pada setiap kelas dalam *dataset* agar model tidak bias terhadap kelas yang dominan. Dalam penelitian ini, *dataset* terdiri dari dua kelas yaitu normal dan katarak.

Train-validation-test split merupakan salah satu tahapan penting dalam *pipeline* pembelajaran mesin yang bertujuan untuk membagi *dataset* ke dalam tiga *subset* berbeda, yaitu data latih, data validasi, dan data uji (Oymak et al., 2022). Data latih digunakan sebagai bahan utama untuk membentuk dan menyesuaikan bobot model selama proses pelatihan. Sementara itu, data validasi berperan dalam mengevaluasi kinerja model secara berkala pada data yang belum pernah dilihat selama pelatihan, sehingga dapat digunakan untuk mendeteksi *overfitting* atau *underfitting*. Terakhir, data uji digunakan sebagai acuan akhir untuk menilai seberapa baik model dapat menggeneralisasi performanya terhadap data baru dan tidak dikenal sebelumnya.

Augmentasi citra merupakan teknik yang digunakan dalam pembelajaran mesin untuk memperluas dan memperkaya variasi data pelatihan dengan cara menerapkan berbagai transformasi terhadap citra yang sudah ada tanpa mengubah labelnya. Metode ini berfungsi sebagai bentuk regularisasi implisit yang membantu mencegah *overfitting*, yaitu kondisi ketika model terlalu menyesuaikan diri dengan data latih dan gagal menggeneralisasi pada data baru (Alimisis et al., 2025). Dengan menciptakan variasi citra dari data yang terbatas, augmentasi memungkinkan model untuk belajar dari representasi yang lebih luas tanpa memerlukan pengumpulan data tambahan.

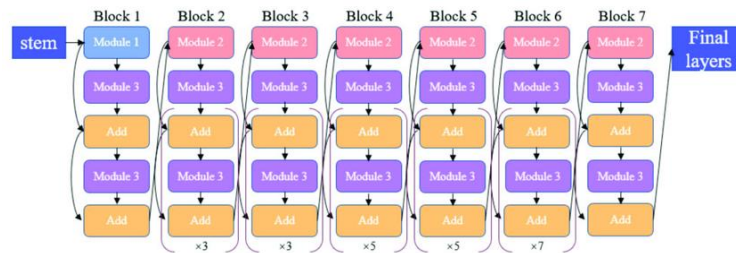
Dalam penelitian ini, augmentasi citra diterapkan secara khusus hanya pada data latih, dengan menggunakan transformasi seperti rotasi, pembalikan horizontal atau vertikal, serta penyesuaian tingkat kecerahan. Teknik augmentasi ini dirancang untuk meningkatkan keragaman pola visual yang dilihat oleh model selama proses pelatihan, sehingga model menjadi lebih tangguh terhadap variasi nyata di dunia nyata. Dengan demikian, augmentasi berperan penting dalam mengurangi kemungkinan *overfitting* serta meningkatkan kemampuan generalisasi model dalam mengklasifikasikan citra fundus retina yang belum pernah dilihat sebelumnya.

2.4 EfficientNetB5 Modelling

EfficientNetB3 merupakan salah satu varian dari keluarga arsitektur *EfficientNet* yang dikembangkan untuk mengoptimalkan performa model dengan efisiensi parameter yang tinggi. Arsitektur ini termasuk dalam kategori *Convolutional Neural Network* (CNN) dan dirancang untuk menangani tugas klasifikasi citra dengan kompleksitas tinggi, termasuk dalam bidang medis seperti analisis citra fundus retina (Dudeja et al., 2023). Keunggulan utama *EfficientNetB5* terletak pada metode penskalaan terkoordinasi yang disebut *compound scaling*, yang secara seimbang meningkatkan kedalaman jaringan (*depth*), lebar jaringan (*width*), dan resolusi *input* gambar. Pendekatan ini menghasilkan keseimbangan antara akurasi yang tinggi dan efisiensi komputasi, menjadikan *EfficientNetB5* pilihan ideal untuk aplikasi klasifikasi yang membutuhkan performa tinggi tanpa membebani sumber daya secara berlebihan.

Pendekatan ini terbukti lebih optimal dibandingkan metode konvensional yang hanya meningkatkan salah satu dimensi saja (Tan & Le, 2019.). Metode ini berbeda dengan metode konvensional yang umumnya hanya melakukan peningkatan pada satu aspek arsitektur, sehingga sering kali berujung pada *overfitting* atau peningkatan beban komputasi yang tidak proporsional. Salah satu keunggulan *EfficientNet* adalah penggunaan Mobile Inverted Bottleneck Convolution (MBConv) sebagai blok utama. MBConv merupakan pengembangan dari *depthwise separable convolution* yang sebelumnya digunakan pada MobileNetV2. Blok MBConv terdiri dari tiga tahap utama, yaitu *expansion layer*, *depthwise convolution*, serta

projection layer. Selain itu, MBConv dilengkapi dengan *squeeze-and-excitation (SE) block* yang berfungsi memberikan bobot perhatian pada saluran (*channel attention*), sehingga jaringan dapat lebih fokus pada fitur penting dari citra masukan (Sandler et al., 2018).



Gambar 2.3 Struktur Arsitektur *EfficientNetB5*

Struktur *EfficientNet-B5* diawali dengan *stem layer* yang melakukan ekstraksi fitur dasar melalui konvolusi awal, kemudian dilanjutkan ke tujuh blok utama yang tersusun dari blok-blok MBConv (*Mobile Bottleneck Convolution*), yang merupakan varian *inverted residual block* dengan *depthwise separable convolution*, *squeeze-and-excitation*, dan *skip connection* guna meningkatkan representasi fitur sekaligus mempertahankan efisiensi gradien. Setiap blok mengandung pengulangan modul yang meningkat sesuai dengan kebutuhan kompleksitas fitur, sehingga blok-blok berikutnya lebih dalam dan lebih lebar dibanding blok sebelumnya. Setelah seluruh blok dilewati, representasi fitur dirangkum melalui *global average pooling* dan diakhiri oleh lapisan klasifikasi yang berupa *fully connected layer*, dengan kemungkinan penerapan *dropout* untuk mitigasi *overfitting*. Struktur ini memungkinkan model menangkap pola visual pada berbagai level abstraksi sambil menjaga penggunaan sumber daya tetap efisien. Proses fine-tuning dilakukan dengan membekukan sebagian lapisan awal untuk mempertahankan fitur low-level, kemudian membuka lapisan lebih dalam untuk melatih fitur high-level yang lebih relevan terhadap karakteristik citra fundus retina.

Tahap pelatihan model *EfficientNetB5* menjadi langkah kunci dalam membangun sistem klasifikasi citra fundus retina yang andal. Pada tahapan ini, model dilatih menggunakan data latih yang telah dibersihkan, diseimbangkan, dan diperluas melalui proses augmentasi guna meningkatkan keberagaman pola visual. Pelatihan dilakukan selama 25 *epoch*, yang berarti model menjalani 25 siklus penuh dalam memproses seluruh data latih. Selama proses ini, bobot-bobot pada jaringan saraf diperbarui secara bertahap agar model dapat secara akurat membedakan karakteristik dari masing-masing kelas, yaitu katarak, glaukoma, retinopati diabetik, dan mata normal. Seluruh proses pelatihan dieksekusi menggunakan GPU untuk mempercepat komputasi dan meningkatkan efisiensi waktu. Visualisasi hasil pelatihan model sangat penting untuk memahami proses belajar model selama *training*.

Dalam penelitian ini, grafik *loss* dan akurasi per *epoch* digunakan untuk memantau performa model terhadap data latih dan validasi selama 25 *epoch*. Grafik *loss* menunjukkan seberapa baik model mengurangi kesalahan prediksi, sementara grafik akurasi menampilkan kemampuan model dalam mengklasifikasi citra secara benar. Penurunan *loss* dan peningkatan akurasi yang konsisten menandakan proses pelatihan berjalan efektif. Sebaliknya, jika *loss* validasi meningkat sementara akurasi tidak membaik, hal ini dapat menunjukkan *overfitting*. Dengan mengamati kedua grafik ini, peneliti dapat mengevaluasi stabilitas pelatihan serta menentukan langkah optimalisasi lebih lanjut.

2.5 Evaluation

Tahap evaluasi dan prediksi model menjadi langkah penutup dalam mengukur kinerja model *EfficientNetB5* yang telah selesai dilatih. Evaluasi dilakukan dengan menggunakan data uji, yaitu data yang tidak digunakan selama pelatihan maupun validasi, guna mengukur kemampuan model dalam mengenali pola baru secara umum. Evaluasi ini bertujuan untuk memastikan bahwa model tidak hanya mampu mencapai akurasi tinggi pada data pelatihan,

tetapi juga memiliki kemampuan generalisasi yang baik terhadap data baru. Proses evaluasi dilakukan menggunakan dataset pengujian (*testing set*) yang tidak pernah digunakan selama proses pelatihan, sehingga dapat memberikan gambaran objektif mengenai performa akhir model.

Metrik utama yang digunakan dalam penelitian ini meliputi akurasi, precision, recall, dan F1-score, yang umum digunakan pada permasalahan klasifikasi medis karena mampu memberikan penilaian yang lebih komprehensif terhadap performa model. Akurasi dihitung sebagai proporsi prediksi benar terhadap keseluruhan sampel, sedangkan precision dan recall digunakan untuk menilai bagaimana model membedakan kelas positif dan negatif. Precision mengukur sejauh mana prediksi positif benar-benar relevan, sementara recall mengukur sejauh mana model mampu menemukan seluruh instance dari kelas positif.

Selain metrik yang telah disebutkan penelitian ini menggunakan evaluasi *confusion matrix*, yang menyajikan jumlah prediksi yang tepat dan keliru untuk setiap kelas. Melalui *confusion matrix*, tidak hanya terlihat tingkat akurasi keseluruhan model, tetapi juga efektivitasnya dalam mengklasifikasikan penyakit mata seperti mata normal dan katarak. *Confusion matrix* adalah sebuah tabel evaluasi yang digunakan untuk menilai performa model klasifikasi, khususnya dalam pendekatan *supervised learning*. Tabel ini menunjukkan kemampuan model dalam membedakan berbagai kelas, dengan cara mencatat jumlah prediksi yang benar dan salah pada masing-masing kategori. Untuk klasifikasi biner, *confusion matrix* memiliki format 2x2, sementara pada klasifikasi multi-kelas, tabel ini berbentuk $n \times n$ sesuai jumlah kelas yang dianalisis (Vanacore et al., 2024).

Tabel 2.1 *Confusion Matrix*

<i>Actual Class</i>	<i>Predicted Class</i>		Total
	<i>Yes</i>	<i>No</i>	
<i>Yes</i>	TP	FN	P
<i>No</i>	FP	TN	N
Total	P'	N'	P + N

Dalam evaluasi model klasifikasi, terdapat empat komponen utama yang digunakan untuk mengukur performa model, yaitu *True Positive* (TP), *False Positive* (FP), *False Negative* (FN), dan *True Negative* (TN). *True Positive* (TP) merujuk pada jumlah data yang diprediksi sebagai positif oleh model dan memang benar secara aktual. Sebaliknya, *False Positive* (FP) menggambarkan jumlah data yang sebenarnya negatif, namun salah diklasifikasikan sebagai positif. *False Negative* (FN) menunjukkan kasus di mana data yang seharusnya diklasifikasikan sebagai positif justru diprediksi sebagai negatif oleh model. Sementara itu, *True Negative* (TN) adalah jumlah data yang diprediksi sebagai negatif dan memang benar bernilai negatif secara aktual (Piegorsch, 2020).

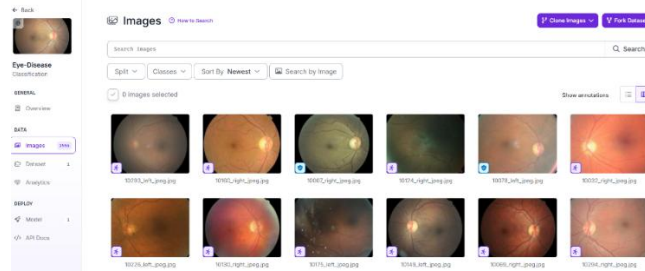
3 HASIL DAN PEMBAHASAN

Bagian hasil dan pembahasan merupakan komponen krusial dalam penelitian ini, yang berfungsi untuk menyampaikan serta mengevaluasi kinerja model klasifikasi EfficientNetB3 dalam mengidentifikasi penyakit mata melalui citra fundus retina. Pada bagian ini, disajikan berbagai temuan dari proses pelatihan dan pengujian model, termasuk akurasi, nilai loss, dan metrik evaluasi lainnya. Hasil tersebut kemudian dianalisis untuk menilai sejauh mana model dapat mengenali karakteristik visual dari masing-masing kategori penyakit. Selain itu, pembahasan difokuskan pada efektivitas modifikasi yang diterapkan terhadap arsitektur

EfficientNetB3 serta analisis terhadap keunggulan dan keterbatasan model, khususnya dalam konteks penerapan pada data medis.

3.1 Data Collecting

Berdasarkan hasil tahap pengumpulan data, diketahui bahwa *dataset* yang dimanfaatkan dalam penelitian ini berasal dari platform Roboflow dengan nama *Eye Diseases Classification* yang dipublikasikan pada tahun 2024. *Dataset* ini mencakup ribuan gambar fundus mata yang telah diklasifikasikan ke dalam empat kelas, yaitu mata normal, katarak, glaukoma, dan retinopati diabetik. Setiap gambar disimpan dalam folder sesuai dengan kategorinya dan mencerminkan kondisi mata kanan maupun kiri dari masing-masing pasien. Ilustrasi salah satu contoh citra fundus yang digunakan dapat dilihat pada Gambar 3.1

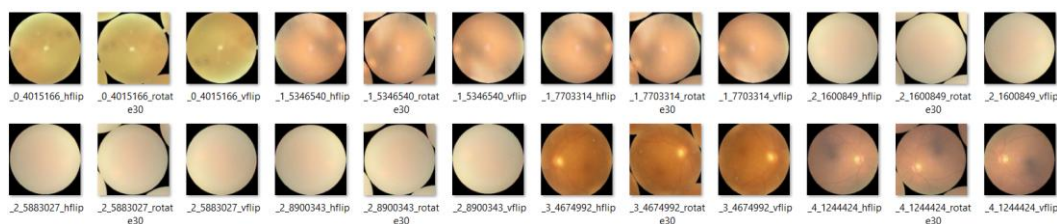


Gambar 3.1 Dataset Citra Fundus Mata

Dataset ini bersifat *open-source* dan tersedia untuk umum, sehingga dapat dimanfaatkan oleh peneliti dalam pengembangan model klasifikasi penyakit mata. Untuk mempermudah dan mempercepat proses pengambilan serta pengelolaan data, penelitian ini memanfaatkan pustaka *kagglehub* yang memungkinkan koneksi langsung antara Kaggle dan lingkungan pemrograman. Setelah proses unduh selesai, seluruh citra disimpan secara lokal dan siap digunakan pada tahapan berikutnya, termasuk pra-pemrosesan, pelatihan, hingga evaluasi model.

3.2 Pre-Processing Data

Tahapan pra-pemrosesan data memainkan peran krusial dalam memastikan bahwa data yang digunakan dalam pelatihan model memiliki kualitas yang optimal, karena kualitas *input* sangat berpengaruh terhadap performa sistem klasifikasi secara keseluruhan. Setelah data dinyatakan bersih dan seimbang, tahap selanjutnya adalah melakukan pembagian *dataset* ke dalam tiga bagian, yakni 70% untuk pelatihan, 20% untuk validasi, dan 10% untuk pengujian. Tujuan dari pembagian ini adalah agar model dapat memperoleh pola dari sebagian besar data, menguji performa awalnya pada data yang belum pernah dilihat saat pelatihan, dan mengevaluasi kemampuannya secara objektif pada data yang benar-benar baru. Proses ini menjadi langkah penting dalam menilai sejauh mana model mampu melakukan generalisasi. Tahapan terakhir dalam pra-pemrosesan adalah augmentasi pada data latih, yang divisualisasikan pada Gambar 3.2



Gambar 3.2 Gambar Dataset Yang Sudah Diaugmentasi

Teknik augmentasi yang digunakan mencakup transformasi seperti rotasi citra, pembalikan secara horizontal dan vertikal, serta penyesuaian tingkat kecerahan. Tujuannya adalah untuk menciptakan keragaman data secara buatan tanpa menambahkan gambar baru, sehingga model memiliki kemampuan yang lebih baik dalam mengenali berbagai variasi visual pada citra fundus retina. Pendekatan ini juga efektif dalam meminimalkan kemungkinan *overfitting*. Luaran dari seluruh proses pra-pemrosesan ini adalah *dataset* yang telah siap digunakan dalam pelatihan model klasifikasi berbasis arsitektur *EfficientNetB5*.

3.3 EfficientNetB5 Modelling

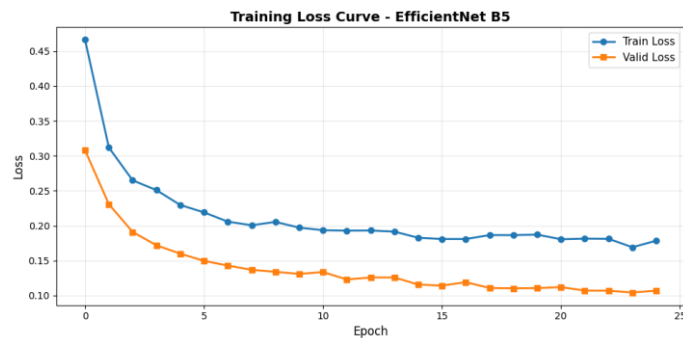
Selama tahap pelatihan, model menggunakan data latih yang telah mengalami proses augmentasi dan secara rutin dievaluasi terhadap data validasi pada setiap *epoch*. Pelatihan dilakukan selama 25 *epoch*, yang berarti seluruh kumpulan data latih dilalui sebanyak 25 kali untuk menyesuaikan bobot model menggunakan algoritma *backpropagation*. Dalam setiap siklus pelatihan, dua indikator utama yang diamati adalah nilai *loss* dan akurasi, baik pada data latih maupun validasi. Nilai *loss* merepresentasikan deviasi prediksi model dari label yang sebenarnya, sedangkan akurasi mencerminkan seberapa banyak prediksi yang tepat. Perbandingan grafik antara kedua metrik ini pada data latih dan validasi sangat penting untuk mengidentifikasi adanya gejala *overfitting* atau *underfitting*.

Tabel 3.1 Sejarah Pelatihan Model

<i>Epoch</i>	<i>Valid Loss</i>	<i>Valid Accuracy</i>
1	0.3080	0.9486
2	0.2308	0.9579
3	0.1910	0.9626
4	0.1719	0.9579
5	0.1598	0.9579
...	...	...
22	0.1070	0.9579
23	0.1068	0.9533
24	0.1042	0.9579
25	0.1068	0.9626

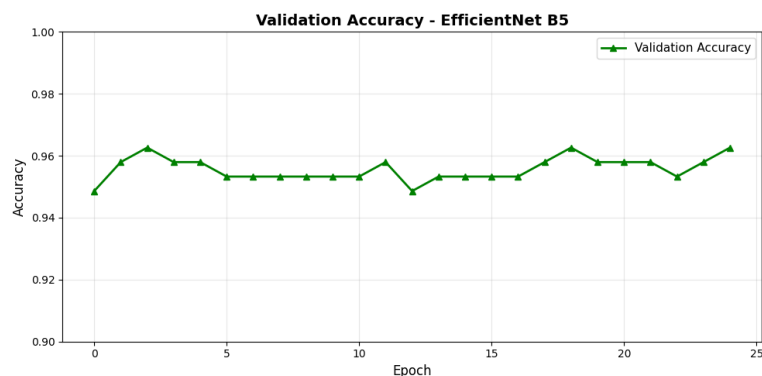
Berdasarkan Tabel 3.1, hasil pelatihan model EfficientNet-B5 menunjukkan bahwa nilai *validation loss* mengalami penurunan yang signifikan seiring bertambahnya epoch, dimulai dari 0,3080 pada epoch pertama dan mencapai nilai terendah sebesar 0,1042 pada epoch ke-24. Penurunan ini mengindikasikan bahwa model semakin mampu meminimalkan kesalahan prediksi pada data validasi. Selain itu, nilai *validation accuracy* berada pada kisaran yang tinggi dan relatif stabil, yaitu di atas 94%, sepanjang proses pelatihan. Stabilitas akurasi validasi tersebut menunjukkan bahwa model memiliki kemampuan generalisasi yang baik dalam membedakan kelas katarak dan normal, serta tidak mengalami degradasi performa yang signifikan selama pelatihan.

Pola numerik yang ditampilkan pada Tabel 3.1 selanjutnya divisualisasikan pada grafik kurva training loss dan validation loss. Grafik tersebut menunjukkan penurunan tajam pada kedua nilai loss di epoch-epoch awal, yang mencerminkan proses pembelajaran model yang berlangsung efektif. Pada epoch selanjutnya, kurva loss cenderung melandai dan stabil, menandakan bahwa model telah mencapai kondisi konvergen. Perbedaan yang relatif kecil antara training loss dan validation loss juga mengindikasikan bahwa model tidak mengalami overfitting, sehingga performa yang diperoleh pada data pelatihan dapat digeneralisasikan dengan baik pada data yang belum pernah dilihat sebelumnya. Visualisasi tren *loss* dan akurasi dari proses pelatihan ini dapat dilihat pada Gambar 3.3



Gambar 3.3 Grafik Loss Model

Selanjutnya, Gambar 3.3 menampilkan grafik *validation accuracy* selama proses pelatihan model EfficientNet-B5. Grafik tersebut menunjukkan bahwa nilai akurasi validasi berada pada tingkat yang tinggi dan relatif stabil, yaitu berkisar antara 94,8% hingga 96,3% sepanjang 25 epoch pelatihan. Meskipun terdapat fluktuasi kecil pada beberapa epoch, perubahan tersebut tidak bersifat signifikan dan tidak menunjukkan tren penurunan performa. Stabilitas *validation accuracy* ini memperkuat hasil yang ditunjukkan oleh grafik loss sebelumnya, di mana penurunan dan konvergensi nilai *validation loss* diikuti oleh kemampuan model dalam mempertahankan tingkat akurasi yang konsisten. Dengan demikian, model dapat dikatakan telah mencapai performa yang optimal dan memiliki kemampuan generalisasi yang baik terhadap data validasi.



Gambar 3.4 Grafik Akurasi Validasi

3.4 Evaluation

Hasil dari tahap evaluasi dan prediksi model terhadap data uji menunjukkan bahwa arsitektur *EfficientNetB5* yang telah dilatih mampu mengklasifikasikan citra fundus retina secara efektif ke dalam dua kategori utama: normal dan katarak. Penilaian kinerja model

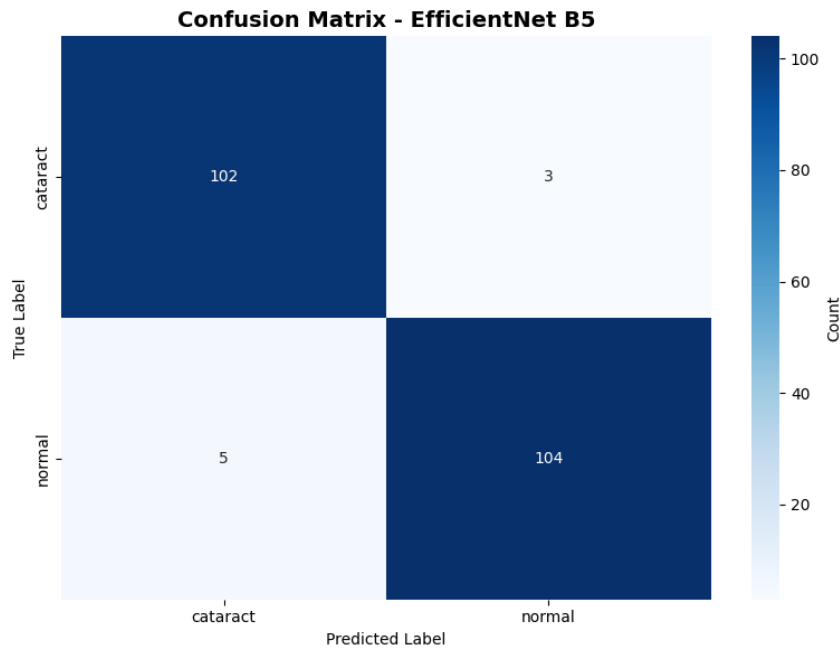
dilakukan menggunakan sejumlah metrik, yakni akurasi, presisi, *recall*, dan *f1-score*, yang dihitung berdasarkan hasil prediksi pada data uji. Evaluasi ini memberikan gambaran menyeluruh tentang seberapa baik model mengenali pola-pola visual dari citra yang belum pernah digunakan selama pelatihan. Detail metrik evaluasi dari masing-masing kelas ditampilkan pada Tabel X

Tabel 3.2 Matrix Evaluasi Model EfficientNetB5

Matrix Evaluasi			
Precision	Recall	F1-Score	Specificity
97%	95%	96%	97%

Berdasarkan Tabel 3.2, hasil evaluasi model EfficientNet-B5 menunjukkan performa yang sangat baik pada seluruh metrik pengukuran. Nilai *precision* sebesar 97% mengindikasikan bahwa sebagian besar prediksi positif yang dihasilkan model merupakan prediksi yang benar, sehingga tingkat kesalahan *false positive* relatif rendah. Selanjutnya, nilai *recall* sebesar 95% menunjukkan bahwa model mampu mendeteksi sebagian besar kasus katarak dengan baik, meskipun masih terdapat sebagian kecil kasus yang tidak teridentifikasi. Nilai *F1-score* sebesar 96% menegaskan adanya keseimbangan yang baik antara *precision* dan *recall*, yang mencerminkan konsistensi performa model dalam klasifikasi. Selain itu, nilai *specificity* sebesar 97% menunjukkan kemampuan model yang sangat baik dalam mengenali citra mata normal secara benar, sehingga meminimalkan kesalahan dalam mengklasifikasikan data normal sebagai katarak. Secara keseluruhan, hasil ini menunjukkan bahwa model EfficientNet-B5 memiliki tingkat akurasi dan kemampuan generalisasi yang tinggi dalam membedakan kondisi mata katarak dan normal.

Meskipun nilai metrik evaluasi pada Tabel 3.2 menunjukkan performa model EfficientNet-B5 yang tinggi pada aspek *precision*, *recall*, *F1-score*, dan *specificity*, metrik-metrik tersebut belum sepenuhnya menggambarkan distribusi kesalahan klasifikasi yang terjadi pada setiap kelas. Oleh karena itu, diperlukan analisis lebih lanjut menggunakan *confusion matrix* untuk melihat secara rinci jumlah prediksi yang benar dan keliru pada masing-masing kelas, baik pada citra mata katarak maupun normal. Visualisasi *confusion matrix* ini memungkinkan evaluasi yang lebih mendalam terhadap pola kesalahan model, khususnya dalam mengidentifikasi kasus *false positive* dan *false negative*, sehingga memberikan gambaran yang lebih komprehensif mengenai kinerja model dalam konteks klasifikasi medis.



Gambar 3.5 Confusion Matrix

Berdasarkan *confusion matrix* hasil pengujian model EfficientNetB5, dapat diketahui bahwa model menunjukkan kinerja klasifikasi yang sangat baik dalam membedakan citra fundus retina katarak dan normal. Dari total data uji, sebanyak 102 citra katarak berhasil diklasifikasikan dengan benar sebagai katarak (*true positive*), sementara hanya 3 citra katarak yang keliru diklasifikasikan sebagai normal (*false negative*). Pada kelas normal, model mampu mengklasifikasikan 104 citra secara tepat sebagai normal (*true negative*), dengan hanya 5 citra normal yang salah diklasifikasikan sebagai katarak (*false positive*). Distribusi kesalahan yang relatif kecil pada kedua kelas menunjukkan bahwa model memiliki kemampuan generalisasi yang baik dan tidak bias terhadap salah satu kelas tertentu. Secara keseluruhan, hasil *confusion matrix* ini mengindikasikan bahwa arsitektur EfficientNetB5 efektif dalam menangkap karakteristik visual yang membedakan citra katarak dan normal, sehingga berpotensi untuk diterapkan sebagai sistem pendukung diagnosis katarak berbasis citra fundus retina.

4 KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa penerapan arsitektur Convolutional Neural Network (CNN) EfficientNet-B5 mampu memberikan performa yang sangat baik dalam mengklasifikasikan citra fundus retina ke dalam kategori mata normal dan katarak. Melalui tahapan pra-pemrosesan yang meliputi data cleaning, class balancing, pembagian dataset, serta augmentasi citra, model dapat dilatih secara optimal sehingga mampu mempelajari karakteristik visual yang relevan dari citra fundus retina.

Hasil evaluasi menunjukkan bahwa model EfficientNet-B5 mencapai akurasi sebesar 96,26%, dengan nilai precision 97%, recall 95%, F1-score 96%, dan specificity 97%. Nilai-nilai tersebut mengindikasikan bahwa model tidak hanya akurat dalam mendeteksi kasus katarak, tetapi juga memiliki kemampuan yang sangat baik dalam mengidentifikasi citra mata normal, sehingga risiko kesalahan klasifikasi dapat diminimalkan. Analisis confusion matrix juga memperlihatkan jumlah kesalahan prediksi yang relatif kecil pada kedua kelas, yang menandakan kemampuan generalisasi model yang baik terhadap data uji.

Selain itu, stabilitas nilai loss dan akurasi selama proses pelatihan dan validasi menunjukkan bahwa model tidak mengalami overfitting maupun underfitting secara signifikan. Hal ini menegaskan bahwa strategi pelatihan dan fine-tuning yang diterapkan pada EfficientNet-B5 sudah tepat untuk permasalahan klasifikasi citra medis, khususnya citra fundus retina. Keunggulan EfficientNet-B5 dalam memanfaatkan metode compound scaling terbukti mampu menghasilkan keseimbangan antara akurasi tinggi dan efisiensi komputasi.

Secara keseluruhan, penelitian ini membuktikan bahwa EfficientNet-B5 efektif digunakan sebagai model tunggal untuk klasifikasi katarak berbasis citra fundus retina tanpa memerlukan metode interpretabilitas tambahan. Temuan ini memiliki potensi besar untuk dikembangkan lebih lanjut sebagai sistem pendukung keputusan medis dalam skrining awal katarak, terutama di wilayah dengan keterbatasan tenaga medis spesialis. Untuk penelitian selanjutnya, disarankan agar dilakukan pengujian pada dataset yang lebih besar dan beragam, serta pengembangan sistem berbasis aplikasi agar model dapat diimplementasikan secara langsung dalam lingkungan klinis.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh pihak yang telah berkontribusi dalam pelaksanaan penelitian ini, baik dalam bentuk dukungan teknis, pendampingan, maupun peran sebagai konsultan selama proses penelitian. Apresiasi juga disampaikan kepada rekan-rekan yang telah membantu dalam pengumpulan dan analisis data serta memberikan masukan yang konstruktif, sehingga penelitian ini dapat diselesaikan dengan baik dan sesuai dengan tujuan yang diharapkan.

DAFTAR PUSTAKA

- Alimisis, P., Mademlis, I., Radoglou-Grammatikis, P., Sarigiannidis, P., & Papadopoulos, G. Th. (2025). Advances in diffusion models for image data augmentation: a review of methods, models, evaluation metrics and future research directions. *Artificial Intelligence Review*, 58(4), 112. <https://doi.org/10.1007/s10462-025-11116-x>
- Dudeja, T., Dubey, S. K., & Bhatt, A. K. (2023). Ensembled EfficientNetB3 architecture for multi-class classification of tumours in MRI images. *Intelligent Decision Technologies*, 17(2), 395–414. <https://doi.org/10.3233/IDT-220150>
- Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA - Journal of the American Medical Association*, 316(22), 2402–2410. <https://doi.org/10.1001/jama.2016.17216>
- Health Organization, W. (2019). *World report on vision*.
- Junayed, M. S., Islam, M. B., Sadeghzadeh, A., & Rahman, S. (2021). CataractNet: An automated cataract detection system using deep learning for fundus images. *IEEE Access*, 9, 128799–128808. <https://doi.org/10.1109/ACCESS.2021.3112938>
- Klein, R., & Klein, B. E. K. (2018). *Epidemiology of Ocular Functions and Diseases in Persons With Diabetes* (3rd ed.). National Institute of Diabetes and Digestive and Kidney Diseases (US), Bethesda (MD). <http://europepmc.org/books/NBK567971>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. In *Nature* (Vol. 521, Number 7553, pp. 436–444). Nature Publishing Group. <https://doi.org/10.1038/nature14539>
- Mardjuki, A., Larasandi, A., Silaen, D. H., & Ramadani, R. A. (2025). Improving Knowledge and Behavior through Cataract Education in the Pre-Prosperous Community of Tangerang. *Jurnal Abmas Negeri (JAGRI)*, 6(1), 101–109. <https://doi.org/10.36590/jagri.v6i1.1525>
- Oymak, S., Li, M., & Soltanolkotabi, M. (2022). *Generalization Guarantees for Neural Architecture Search with Train-Validation Split*.
- Piegorsch, W. W. (2020). Confusion Matrix. In *Wiley StatsRef: Statistics Reference Online* (pp. 1–4). Wiley. <https://doi.org/10.1002/9781118445112.stat08244>
- PROFIL KESEHATAN INDONESIA 2022. (2022).
- Salsabila, K. D., Fakhriinnisa, T. A., & Unari, U. (2024). Geographic distribution of ophthalmologist in East Java 2023. *Surabaya Medical Journal*, 15–24. <https://doi.org/10.59747/smjidisurabaya.v2i1.57>
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). *MobileNetV2: Inverted Residuals and Linear Bottlenecks*.

- Tan, M., & Le, Q. V. (2019). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*.
- Vanacore, A., Pellegrino, M. S., & Ciardiello, A. (2024). Fair evaluation of classifier predictive performance based on binary confusion matrix. *Computational Statistics*, 39(1), 363–383. <https://doi.org/10.1007/s00180-022-01301-9>