Jurnal Kreativitas Mahasiswa Informatika **E-ISSN: 3164-1989**
Penerbit: Fakultas Ilmu Komputer Universitas Pamulang **P-ISSN: 2797-6327**
Vol. 7, No. 2, Agustus 2026 (265-275) **https://doi.org/10.32493/jatimika.v7i1**

Analisa Pengaruh Augmentasi Terhadap Akurasi CNN Dalam Klasifikasi Mata Normal dan Katarak

Adriansyah[1], Atang Susila[2]

[1,2]Teknik Informatika, Universitas Pamulang, Jl. Surya Kencana No.1 Pamulang Barat, Tangerang Selatan, Indonesia
e-mail: [1]ancaadri9@gmail.com, [2]atang.g66@gmail.com

Submitted Date: Januari 18, 2026 Reviewed Date: April 23, 2026
Revised Date: Mei 27, 2026 Accepted Date: Juli 23, 2026

Abstract

Cataracts are a leading cause of blindness worldwide. Various solutions have been developed for cataract detection through the application of Convolutional Neural Networks (CNN). However, CNN development presents challenges such as limited data. Image augmentation offers potential solution, but augmentation does not guarantee improvement. This study aims to analyze the effect of various augmentations on CNN accuracy in classifying normal eyes and cataracts. The CNN model was trained using 498 fundus images from the Grand Challenges platform. The dataset was divided into normal and cataract classes, distributed into three subsets: training, validation, and test data with a ratio of 70:15:15 for normal class and 80:10:10 for cataract class. Various augmentations included flip, rotation, zoom, translation, contrast, brightness, hue, and saturation. Evaluation metrics included accuracy, precision, recall, and F1-score. The evaluation shows that augmentations such as rotation, zoom, and brightness were able to improve the performance of the baseline model with an F1-score of 94.4% increasing to 96.15%. Conversely, hue and saturation augmentation actually decreased model performance, with the lowest F1-score reaching 50%. The analysis shows that to improve CNN model quality, proper augmentation configurations are essential. Incorrect configurations could results negative impact on model quality.

Keywords: Cataract, Convolutional Neural Network, Digital Image Processing

Abstrak

Katarak merupakan penyebab utama kebutaan di seluruh dunia. Berbagai solusi dikembangkan untuk deteksi katarak melalui penerapan *Convolutional Neural Network* (CNN). Namun, pengembangan CNN mendatangkan tantangan seperti keterbatasan data. Augmentasi citra menjadi solusi potensial untuk mengatasi masalah ini, namun augmentasi tidak selalu memberikan pengaruh positif. Penelitian ini bertujuan menganalisis pengaruh berbagai augmentasi terhadap akurasi CNN dalam klasifikasi mata normal dan katarak. Model CNN dilatih menggunakan 498 citra fundus yang tersedia di platform Grand Challenges. Dataset terbagi menjadi kelas normal dan katarak yang didistribusikan menjadi 3 *subset* yaitu data latih, validasi, dan uji dengan rasio 70:15:15 pada kelas normal dan 80:10:10 pada kelas katarak. Berbagai augmentasi yang dianalisis meliputi *flip, rotation, zoom, translation, contrast, brightness, hue*, dan *saturation*. Metrik evaluasi meliputi *accuracy, precision, recall*, dan *F1-score*. Evaluasi menunjukkan augmentasi seperti *rotation, zoom*, dan *brightness* mampu meningkatkan performa model *baseline* dengan *F1-score* sebesar 94,4% yang meningkat hingga 96,15%. Sebaliknya, augmentasi *hue* dan *saturation* justru menurunkan performa model dengan *F1-score* terendah mencapai 50%. Analisis menunjukkan bahwa untuk meningkatkan kualitas model CNN, perlu dipertimbangkan augmentasi dan parameter yang tepat. Konfigurasi yang salah justru akan mengakibatkan efek negatif yang fatal terhadap kualitas model.

Kata kunci: Citra Digital, *Convolutional Neural Network*, Katarak

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Author1, Author2, Author3

1. Pendahuluan

Penglihatan berperan krusial dalam menunjang aktivitas harian dan kualitas hidup manusia. Gangguan pada fungsi ini dapat berdampak langsung pada kemandirian, produktivitas, dan interaksi sosial. Salah satu bentuk gangguan paling serius adalah kebutaan, yang telah menjadi isu kesehatan global dengan dampak luas terhadap individu maupun masyarakat. Di antara berbagai penyebab kebutaan, katarak merupakan yang paling umum. Katarak merujuk pada kondisi ketika lensa mata mengalami kekeruhan secara bertahap sehingga menurunkan ketajaman visual dan berpotensi menyebabkan kebutaan (Lapp et al., 2023). Berdasarkan data yang diteliti pada tahun 2020 pada pengidap penyakit mata usia 50 tahun keatas, katarak menjadi penyebab utama kebutaan dan gangguan penglihatan dengan jumlah kasus mencapai 94 juta secara global (Steinmetz et al., 2021). Berdasarkan data survey *Rapid Assessment of Avoidable Blindness* (RAAB) yang dilakukan pada usia 50 tahun keatas di 15 provinsi di Indonesia, menunjukkan angka kebutaan di Indonesia yang mencapai 3%, dengan katarak sebagai penyebab utama sebesar 81% (Noviada Kwartawati et al., 2023).

Melihat tingginya prevalensi dan dampak katarak terhadap kesehatan mata, berbagai inovasi telah dikembangkan untuk mendukung proses diagnosis, salah satunya melalui pemanfaatan teknologi seperti citra digital. Sejumlah studi telah mengeksplorasi teknik pengolahan citra dan algoritma *machine learning* untuk mengenali pola visual khas pada penyakit medis. Salah satu pendekatan yang kini semakin banyak digunakan adalah *Convolutional Neural Network* (CNN), yang menawarkan solusi efektif dalam kemampuannya mengekstraksi fitur citra secara otomatis dan efisien. Penelitian oleh Naufal et al. (2021) menunjukkan bahwa CNN memiliki performa yang baik dalam klasifikasi penyakit medis seperti pada citra *X-ray* paru-paru. Dalam penelitiannya, CNN memberikan hasil terbaik dibandingkan dengan metode *machine learning* lain seperti *K-Nearest Neighbor* (KNN) dan *Support Vector Machine* (SVM), dengan akurasi tertinggi mencapai 95,91%. Hal ini menunjukkan potensi CNN dalam mengekstraksi fitur-fitur kompleks pada jenis citra medis lain seperti katarak. Sebagaimana ditunjukkan dalam penelitian oleh Marcella & Devella (2022) yang berhasil mencapai akurasi sebesar 92% pada kelas katarak menggunakan model VGG-19 dan penelitian oleh Ramadhani et al. (2023) dengan model dasar yang mencapai akurasi sebesar 92,5%. Temuan ini menunjukkan bahwa arsitektur CNN seperti VGG-19 dan model dasar sekalipun dapat mencapai performa yang mengesankan. Akan tetapi, dalam upaya meningkatkan akurasi model, sangat penting untuk memperhatikan faktor-faktor seperti teknik *pre-processing*, pemilihan *optimizer*, dan yang terpenting, kuantitas serta kualitas data latih.

Dengan mempertimbangkan pengaruh kualitas dan kuantitas data latih terhadap akurasi model, berbagai penelitian kemudian menyoroti peran augmentasi citra sebagai upaya peningkatan performa CNN. Augmentasi merupakan metode untuk menghasilkan data sintesis sehingga meningkatkan keragaman data latihan (Yang et al., 2025). Sejumlah penelitian menunjukkan bahwa penerapan augmentasi menghasilkan peningkatan performa model secara umum. Augmentasi yang diterapkan dengan cermat, seperti yang dilakukan oleh penelitian Krisnawan & Setiawan (2024) terbukti mampu meningkatkan kualitas generalisasi model secara keseluruhan. Teknik-teknik seperti *rotation, horizontal flip, dan zoom* menunjukkan dampak positif yang konsisten pada dataset kecil dengan pencapaian akurasi setinggi 97% pada data uji. Adapun Penelitian Sasongko & Amrullah (2023) yang menguji efek augmentasi pada model *transfer learning* dengan hasil peningkatan sebesar 0,03-0,1%. Kasus yang berbanding terbalik terjadi pada penelitian Al-Fahrezi (2025), augmentasi yang gagal dioptimalkan terhadap dataset justru dapat menurunkan performa model secara signifikan akibat parameter yang terlalu agresif dan durasi pelatihan yang tidak memadai. Konfigurasi yang salah ini mengakibatkan penurunan akurasi validasi pada model dari 91,13% menjadi 84,55%.

Berdasarkan hasil penelitian terdahulu, diketahui bahwa pemilihan dan penerapan teknik augmentasi yang tepat bukanlah tugas yang sederhana, melainkan memerlukan pertimbangan yang cermat terhadap strategi pelatihan yang digunakan. Penerapan yang tidak sesuai justru dapat menghasilkan dampak yang merugikan, sementara konfigurasi yang optimal terbukti mampu bermanfaat bagi peningkatan kinerja model. Penelitian ini bertujuan untuk mengkaji secara mendalam pengaruh berbagai teknik augmentasi citra terhadap performa model CNN dasar dalam klasifikasi mata normal dan katarak pada citra fundus retina. Melalui perbandingan kinerja model dasar dengan model berlapis augmentasi, dan menggunakan

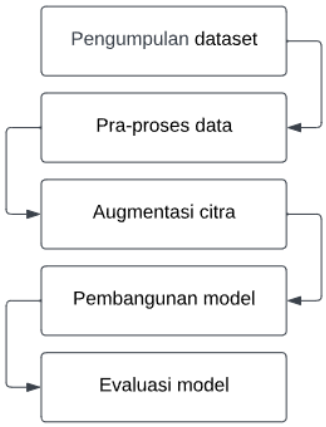
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Author1, Author2, Author3

pendekatan evaluasi perbandingan kuantitatif, diharapkan hasil penelitian ini dapat memberikan kontribusi bagi pengembangan sistem klasifikasi citra medis yang lebih akurat serta aplikatif. Manuskrip ini merupakan hasil asli penelitian yang dilakukan oleh penulis berdasarkan ide dan analisis mandiri dengan tetap mengacu pada penelitian-penelitian terdahulu yang relevan.

2. Metodologi Penelitian

a. Alur Penelitian

Dalam mendukung jalannya penelitian ini agar lebih terarah dan sistematis maka dibutuhkan suatu alur penelitian yang disusun dengan terstruktur. Secara garis besar tahapan yang dilakukan dalam penelitian ini adalah sebagai berikut:



gambar 1. Diagram alur penelitian

b. Pengumpulan Dataset

Data penelitian ini bersumber dari ODIR-5K (Ocular Disease Intelligent Recognition), yaitu dataset publik yang disediakan oleh Peking University melalui platform Grand Challenges. Dataset ini berjumlah 7.000 citra fundus retina beresolusi tinggi yang terdiri dari berbagai label penyakit mata, termasuk katarak, yang diperoleh dari pemeriksaan pasien di klinik mata dan rumah sakit. Seluruh citra dalam dataset telah dilabeli oleh tenaga medis profesional dan dapat diakses secara gratis untuk kepentingan penelitian akademik melalui proses registrasi resmi di situs penyedia.

Proses pengumpulan data dilakukan dengan mengunduh dataset ODIR-5K dari laman https://odir2019.grand-challenge.org/ setelah melalui tahap registrasi dan persetujuan penggunaan oleh penyedia dataset. Seluruh data yang diunduh berbentuk file citra digital berformat JPEG beserta file CSV yang memuat label diagnosis.

ID	Patient Age	Patient Sex	Left-Fundus	Right-Fundus	Left-Diagnostic Keywords	Right-Diagnostic Keywords
0	69.0	Female	0_left.jpg	0_right.jpg	cataract	normal fundus
1	57.0	Male	1_left.jpg	1_right.jpg	normal fundus	normal fundus
2	42.0	Male	2_left.jpg	2_right.jpg	laser spot, moderate non proliferative retinopathy	moderate non proliferative retinopathy
3	66.0	Male	3_left.jpg	3_right.jpg	normal fundus	branch retinal artery occlusion
4	53.0	Male	4_left.jpg	4_right.jpg	macular epiretinal membrane	mild nonproliferative retinopathy

Gambar 2. Anotasi Dataset ODIR-5K

c. Pra-proses Data

Pra-proses dilakukan untuk menyiapkan dan membersihkan data mentah agar menjadi format yang sesuai dan optimal untuk diproses oleh model *machine learning*.

1) Penyaringan data

Dataset saat ini memiliki keseluruhan data yang berjumlah 7.000 citra fundus dengan berbagai label penyakit mata. Dari keseluruhan data yang ada, dipilih data dengan label murni "*Normal*" dan "*Cataract*",

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Author1, Author2, Author3

hal ini memastikan bahwa fitur citra yang digunakan relevan dengan penelitian dan mencegah kesalahan pembelajaran fitur kelas yang tidak relevan pada model nantinya.

Berdasarkan proses seleksi ini, diperoleh 2.814 citra untuk kelas normal dan 249 citra untuk kelas katarak. Ketidakseimbangan jumlah data yang signifikan antara kedua kelas berpotensi menyebabkan model bias terhadap kelas normal yang menjadi kelas mayoritas. Maka dari itu dilakukan proses *random undersampling*, yaitu proses pemilihan data acak untuk menyeimbangkan jumlah data normal agar setara dengan jumlah data katarak yaitu 249 citra.

Selanjutnya dilakukan tahap stratifikasi, yaitu pembagian data ke dalam tiga subset yaitu data latih, validasi, dan uji. Data latih digunakan untuk melatih model, data validasi untuk mengevaluasi proses pembelajaran selama pelatihan, dan data uji untuk menilai kinerja akhir model. Pada penelitian ini, pembagian dilakukan secara terpisah pada masing-masing kelas. Rasio pembagian untuk kelas normal adalah 70:15:15 dari 262 citra, sedangkan untuk kelas katarak adalah 80:10:10 dari 262 citra. Pertimbangan rasio didasarkan pada distribusi jumlah data serta kebutuhan pembelajaran model. Mengingat dataset yang digunakan terbatas, diperlukan proporsi data latih yang lebih besar agar model dapat mempelajari pola secara optimal. Sementara itu, perbedaan rasio pada kedua kelas didasarkan pada kompleksitas karakteristik fitur serta prioritas deteksi, sehingga kelas katarak diberikan porsi data latih yang lebih besar dibandingkan kelas normal.

Tabel 1. Stratifikasi dataset

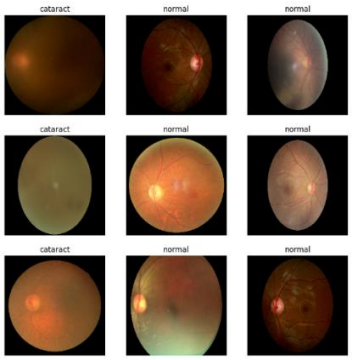
	Normal	**Katarak**
Latihan	174	199
Validasi	37	25
Tes	38	25

2) Ekstraksi data

Data yang sudah dipisah lalu diekstraksi ke berkas baru yang akan digunakan sebagai masukan untuk model. Hirarki berkas dibagi dengan berkas untuk setiap subset (latihan, validasi, tes) yang masing-masing di dalamnya terdapat dua berkas kelas (normal dan katarak).

3) Pemuatan data

Data yang sudah dipisahkan kemudian diekstraksi ke dalam berkas baru yang akan digunakan sebagai masukan bagi model. Hierarki berkas dibagi berdasarkan subset data, yaitu latihan, validasi, dan tes, di mana masing-masing subset memiliki dua kelas, yaitu normal dan katarak. Selanjutnya, seluruh citra yang telah diseleksi dimuat ke dalam program dan menjalani proses standardisasi pra-proses. Tahap pertama adalah *label encoding*, yaitu mengubah label ke dalam format biner untuk kebutuhan klasifikasi biner, dengan nilai 0 merepresentasikan kelas katarak dan nilai 1 merepresentasikan kelas normal. Representasi biner kelas ini akan ditukar pada tahap pengujian untuk keperluan evaluasi. Tahap berikutnya adalah *resizing*, yaitu mengubah ukuran seluruh citra menjadi 224×224 piksel. Dimensi ini dipilih karena merupakan standar yang umum digunakan pada arsitektur CNN modern, sehingga dapat menyeragamkan input sekaligus mengurangi beban komputasi. Setelah itu dilakukan *batching*, yaitu membagi data ke dalam batch berukuran 32 agar proses pelatihan dapat berlangsung lebih efisien. Dengan demikian, data telah siap digunakan dalam proses pelatihan model.



Gambar 3. Visualisasi citra masukan

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Author1, Author2, Author3

E-ISSN: 3164-1989
P-ISSN: 2797-6327
https://doi.org/10.32493/jatimika.v7i1

d. Augmentasi Citra

Teknik augmentasi citra diterapkan untuk meningkatkan keragaman data latih dan mencegah overfitting. Setiap teknik augmentasi yang dipilih memiliki tujuan spesifik dalam meningkatkan kemampuan generalisasi model. Berikut beberapa augmentasi yang diterapkan pada penelitian ini:

Tabel 2. Daftar teknik augmentasi

Teknik Augmentasi	Deskripsi	Parameter
Random Flip	Membalik citra secara horizontal, vertikal, atau keduanya.	Horizontal, Vertikal, Horizontal dan Vertikal
Random Rotation	Memutar citra dalam rentang sudut tertentu.	10-30%
Random Zoom	*Zoom in/out* acak pada citra	10-30%
Random Translation	Menggeser citra di sumbu x/y.	10-30%
Random Contrast	Mengubah kontras citra.	10-30%
Random Brightness	Mengubah tingkat kecerahan.	10-30%
Random Hue	Menggeser hue warna.	10-30%
Random Saturation	Mengubah saturasi warna.	10-30%

Teknik augmentasi diintegrasikan langsung sebagai lapisan pada model, sehingga transformasi citra dilakukan secara dinamis selama proses pelatihan agar menghasilkan variasi data yang sangat beragam tanpa memerlukan tambahan ruang penyimpanan. Nilai parameter pada setiap teknik augmentasi ditentukan berdasarkan rentang parameter yang tersedia. Khusus untuk teknik augmentasi yang menggunakan skala 0–100%, nilai parameter dibatasi pada kisaran 10–30% guna memastikan augmentasi yang tidak berlebihan. Dengan menerapkan teknik augmentasi, diharapkan model dapat memiliki performa yang lebih baik pada data yang belum pernah dilihat sebelumnya.

e. Pembangunan Model

Model dikembangkan menggunakan framework TensorFlow dan Keras dengan bahasa pemrograman Python. Arsitektur dan *hyperparameter* pada model ini ditentukan berdasarkan hasil pengujian paling optimal yang dilakukan selama penulisan penelitian ini.

1) Arsitektur model

Arsitektur model CNN penelitian ini dirancang secara kustom dengan konfigurasi yang dipertimbangkan agar model mencapai hasil yang optimal.

Tabel 3. Arsitektur model

Lapisan	Output
Input	(224, 224, 3)
Augmentasi	(224, 224, 3)
Rescaling	(224, 224, 3)
Conv2D	(224, 224, 32)
Conv2D	(224, 224, 32)
MaxPooling2D	(112, 112, 32)
Conv2D	(112, 112, 64)
Conv2D	(112, 112, 64)
MaxPooling2D	(56, 56, 64)
Conv2D	(56, 56, 128)
Conv2D	(56, 56, 128)
MaxPooling2D	(28, 28, 128)
Conv2D	(28, 28, 256)
GlobalAveragePooling2D	(256)
Dense	(256)
Dense	(1)

Keterangan:

- Lapisan *Input:* Menerima masukkan gambar berukuran 224x224 dengan 3 channel warna (RGB).

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Author1, Author2, Author3

- Lapisan Augmentasi: Memodifikasi gambar secara acak bedasarkan jenis augmentasi yang dimasukkan.
- Lapisan *Rescaling:* Melakukan normalisasi pada citra asli yang memiliki rentang nilai 0-255 menjadi 0-1.
- Lapisan *Conv2D*: Melakukan ekstraksi fitur dari citra melalui operasi konvolusi.
- Lapisan *MaxPooling2D*: Mengurangi dimensi spasial pada fitur citra dengan mengambil nilai maksimum pada setiap wilayah tertentu untuk menurunkan kompleksitas komputasi dan mempertahankan fitur dominan pada citra.
- Lapisan *GlobalAveragePooling2D:* Mengumpulkan setiap map fitur pada citra menjadi satu nilai rata-rata.
- Lapisan *Dense*: Menghasilkan keputusan klasifikasi akhir.

Secara keseluruhan, arsitektur ini dirancang untuk mencapai keseimbangan antara efisiensi komputasi dan kemampuan generalisasi, sehingga dapat memberikan performa optimal pada dataset citra fundus retina yang terbatas.

2) *Hyperparameter* model

Hyperparameter merupakan parameter yang ditentukan sebelum proses pelatihan dan tidak diperbarui secara otomatis oleh algoritma. Penentuan nilai *hyperparameter* dilakukan melalui percobaan bertahap untuk memperoleh konfigurasi yang memberikan akurasi optimal.

Tabel 4. *Hyperparameter* model

Hyperparameter	Nilai	Keterangan
Optimizer	Adam	Optimizer efisien dengan kecepatan konvergensi yang baik.
Learning Rate	0,001	Nilai awal yang umum digunakan dan stabil untuk pelatihan.
Loss Function	Binary Crossentropy	Fungsi yang sesuai untuk masalah klasifikasi biner antara dua kelas.
Batch Size	32	Ukuran yang seimbang antara kecepatan pelatihan dan stabilitas gradien.
Epoch	100	Nilai dipilih berdasarkan rata-rata performa model dalam mencapai epoch terbaiknya.
Metrics	Binary Accuracy, Precision, Recall, F1-score	Memberikan evaluasi kinerja model yang komprehensif dari berbagai perspektif.
Callbacks	EarlyCallback	Menghentikan pelatihan ketika performa model tidak lagi meningkat secara signifikan agar performa lebih ringan dan model berakhir dengan epoch terbaik.

3. Hasil dan Diskusi

a. Prosedur Evaluasi

Hasil pelatihan seluruh model dievaluasi menggunakan metrik yang telah ditentukan untuk membandingkan performanya serta menganalisis pengaruh penerapan berbagai teknik augmentasi pada model CNN. Proses evaluasi diawali dengan pengujian model *baseline* tanpa augmentasi sebagai acuan pembanding. Selama proses pelatihan, dilakukan pemantauan dan perbandingan nilai akurasi pada data latih dan data validasi. Perbandingan ini bertujuan untuk menilai kemampuan model dalam mempelajari pola dari data latih serta mengevaluasi kemampuan generalisasi model terhadap data baru yaitu validasi.

Setelah akurasi pada data validasi dinilai memadai, model kemudian diuji menggunakan data tes. Pada proses pengujian, kelas katarak diperlakukan sebagai kelas positif, sedangkan kelas normal sebagai kelas negatif. Oleh karena itu, evaluasi difokuskan pada kemampuan model dalam mendeteksi kasus katarak secara tepat dan konsisten. Keempat keluaran dari pengujian, yaitu *true positive* (TP), *false positive* (FP), *true negative* (TN), dan *false negative* (FN), disajikan dalam bentuk *confusion matrix* untuk menghitung performa model menggunakan metrik umum evaluasi yaitu akurasi, presisi, *recall*, dan *F1-score*. Rumus

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Author1, Author2, Author3

perhitungan metrik yang digunakan dalam penelitian ini mengacu pada Rashed & Popescu (2023) yang ditunjukkan sebagai berikut:

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

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

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

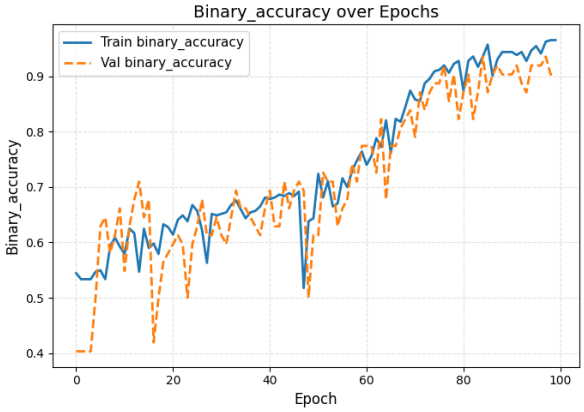
$$F1\ Score = \frac{(2TP)}{(2TP + FN + FP)}$$

Keterangan:
TP = True Positive
FP = False Positive
TN = True Negative
FN = False Negative

Setelah pengujian pada model *baseline* dilakukan, tahap selanjutnya adalah menguji model dengan berbagai model augmentasi menggunakan *confusion matrix* dan metrik evaluasi yang sama. Arsitektur dan parameter pada model augmentasi dirancang identik dengan model *baseline*, dengan perbedaan pada penambahan lapisan augmentasi sebelum lapisan konvolusi yang hanya aktif selama proses pelatihan. Hasil pengujian dari seluruh model kemudian dibandingkan untuk menganalisis pengaruh masing-masing teknik augmentasi terhadap kinerja model. Dalam penelitian ini, metrik *F1-score* dijadikan acuan utama dalam membandingkan performa model, karena metrik ini mampu merepresentasikan keseimbangan klasifikasi antara kedua kelas.

b. Evaluasi Pelatihan Model Dasar

Hasil pelatihan model *baseline* divisualisasikan melalui grafik akurasi pelatihan dan validasi, sebagaimana ditunjukkan pada Gambar 4.



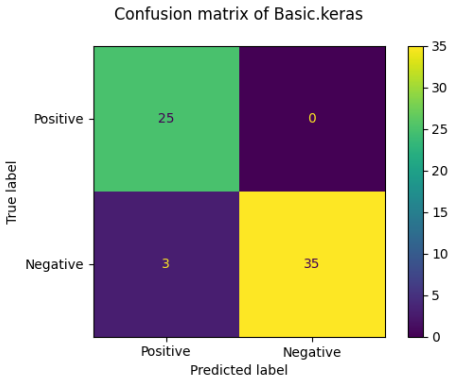
Gambar 4. Visualisasi akurasi model dasar selama pelatihan

Berdasarkan diagram pelatihan model dasar pada Gambar 4, terlihat bahwa model CNN tanpa augmentasi mampu mencapai tingkat akurasi yang tinggi, dengan nilai akurasi pelatihan dan validasi pada *epoch* akhir berada di atas 90%. Meskipun pada tahap awal pelatihan terlihat fluktuasi akurasi yang cukup signifikan, khususnya pada data validasi, tren keseluruhan menunjukkan peningkatan akurasi yang

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Author1, Author2, Author3

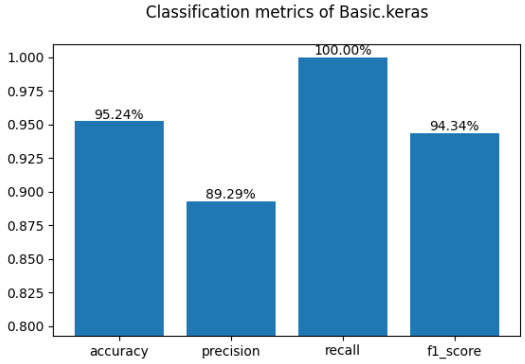
konsisten seiring bertambahnya epoch. Pada fase akhir pelatihan, kurva akurasi pelatihan dan validasi menunjukkan kecenderungan yang relatif sejalan tanpa perbedaan yang mencolok. Hal ini mengindikasikan bahwa model dasar telah mempelajari pola utama pada data latih dan melakukan generalisasi yang cukup baik pada data validasi. Dengan demikian, pada skenario tanpa augmentasi ini tidak ditemukan indikasi *underfitting* maupun *overfitting* yang signifikan, sehingga model dasar dapat dijadikan *baseline* yang layak untuk dibandingkan dengan skenario pelatihan menggunakan augmentasi citra.

Untuk mengevaluasi performa klasifikasi model secara lebih rinci, digunakan *confusion matrix* sebagaimana ditunjukkan pada Gambar 5 serta evaluasi nilai metrik yang ditampilkan pada Gambar 6.



Gambar 5. *Confusion matrix* model dasar

Confusion matrix pada Gambar 5 menunjukkan performa model CNN tanpa augmentasi berhasil mengklasifikasikan 25 citra katarak dan 35 citra normal secara akurat. Namun demikian, masih ditemukan 3 kasus *false positive*, yang mengindikasikan bahwa model masih berpotensi kecil untuk salah mengklasifikasikan kelas normal sebagai katarak.



Gambar 6. Evaluasi model dasar

Pada diagram metrik evaluasi pada Gambar 6, terlihat bahwa model CNN tanpa augmentasi secara umum mampu memberikan prediksi yang baik dari total 63 data tes yang diuji dengan nilai akurasi sebesar 95,24%, dan presisi sebesar 89,29% yang mengindikasikan masih adanya kekeliruan dalam mengklasifikasikan mata normal sebagai katarak, sebagaimana juga ditunjukkan oleh *confusion matrix*. Sementara itu, model menunjukkan performa yang sangat baik dalam mendeteksi mata katarak, ditunjukkan oleh nilai *recall* sebesar 100% yang menandakan tidak terdapat kasus mata katarak yang salah diklasifikasikan. Nilai *F1-score* sebesar 94,34% menunjukkan keseimbangan yang relatif baik dalam generalisasi kedua kelas. Secara keseluruhan, model dasar memiliki kinerja klasifikasi yang baik dan seimbang, namun masih memiliki ruang untuk peningkatan.

c. Evaluasi Perbandingan Model Augmentasi

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Author1, Author2, Author3

Setelah dilakukan evaluasi kinerja model *baseline* serta. Dilakukan analisis komparatif model *baseline* terhadap model yang menerapkan setiap augmentasi secara tunggal. Analisis ini dilakukan dengan membandingkan nilai akurasi, presisi, *recall*, dan *F1-score* yang disajikan dalam bentuk tabel ablasi, dengan model *baseline* sebagai acuan. Pendekatan ini memungkinkan pengaruh setiap modifikasi atau teknik augmentasi terhadap kinerja model dianalisis secara terstruktur. Metrik yang digunakan sebagai acuan utama dalam komparasi adalah *F1-score*, karena metrik ini tidak hanya mencerminkan tingkat ketepatan model, tetapi juga menunjukkan keseimbangan kinerja dalam mengklasifikasikan kedua kelas.

Tabel 5. Komparasi evaluasi model augmentasi

Model	Parameter terbaik	Akurasi (%)	Presisi (%)	Recall (%)	F1-score (%)	Pengaruh (F1-score) (%)
Base	-	95,24	89,29	100	94,34	0
Flip	*Horizontal & Vertical*	95,24	89,29	100	94,34	0
Rotation	20%	96,83	92,59	100	96,15	+ 1,81
Zoom	20%	96,83	92,59	100	96,15	+ 1,81
Translation	10%	96,83	96	96	96	+ 1,66
Contrast	20%	95,24	92,31	96	94,12	- 0,22
Brightness	30%	96,83	92,59	100	96,15	+ 1,81
Hue	10%	85,71	80,77	84	82,35	- 11,99
Saturation	10%	68,25	66,67	40	50	- 44,34

Hasil evaluasi menunjukkan performa yang beragam. Beberapa augmentasi dapat meningkatkan generalisasi model walaupun relatif tidak terlalu besar secara persen karena performa pada model dasar yang sudah sangat baik. Hasil terbaik ditunjukkan oleh augmentasi *rotation, zoom,* dan *brightness* yang mampu meningkatkan nilai *F1-score* sebesar 1,81%, dari 94,34% menjadi 96,15%. Hasil positif lainnya juga diperlihatkan oleh augmentasi *translation* dimana augmentasi ini menghasilkan peningkatan *F1-score* sebesar 1,66% dengan nilai 96% yang tidak jauh berbeda dengan hasil terbaik lainnya. Keempat augmentasi positif ini menunjukkan kebutuhan variasi yang realistis pada kondisi pengambilan citra fundus, seperti perbedaan orientasi kamera, jarak fokus, pencahayaan, dan posisi mata pasien.

Sementara itu, augmentasi *flip* dan *contrast* memberikan kontribusi yang tidak terlalu signifikan dengan nilai *flip* yang identik dengan *baseline* sementara itu nilai model *contrast* mengalami penurunan kecil sebesar 0,22% dengan nilai 94,12%. Penurunan kecil pada *contrast* kemungkinan disebabkan oleh perubahan kontras yang mengganggu ketajaman batas antara area lensa keruh dan jaringan sekitarnya, sehingga model sedikit kesulitan dalam membedakan kedua kelas.

Hasil negatif yang cukup signifikan ditunjukkan oleh augmentasi *hue* dan *saturation* dengan penurunan sebesar 11,9% dengan nilai 82,35% pada augmentasi *hue* dan penurunan sebesar 44,34% pada augmentasi *saturation* dengan total nilai *F1-score* terkecil sebesar 50%. Penurunan yang drastis ini dimungkinkan karena augmentasi *hue* dan *saturation* yang secara langsung mendistorsi fitur pada citra sehingga model mempelajari representasi warna yang tidak relevan atau tidak sesuai dengan kondisi nyata.

d. Interpretasi Hasil Analisis

Model CNN dasar tanpa augmentasi menunjukkan kinerja klasifikasi yang kuat. Dari 63 data tes yang diuji, model dasar mendapat akurasi 95,24%, presisi 89,29%, *recall* 100%, dan *F1-score* 94,34%. Nilai *recall* sebesar 100% menunjukkan bahwa model dapat mendeteksi fitur katarak dengan baik, sehingga tidak terjadi *false negative*. Di sisi lain, nilai presisi sebesar 89,29% mengindikasikan masih adanya tiga kasus *false positive*, yaitu citra mata normal yang salah diklasifikasikan sebagai katarak. Performa ini secara umum sebanding dengan penelitian klasifikasi katarak serupa menggunakan CNN oleh Ramadhani et al., (2023) yang mencapai akurasi sebesar 92,5%. Model dasar menunjukkan performa tinggi yang cukup untuk dijadikan acuan perbandingan yang baik.

Melalui analisis perbandingan model dengan augmentasi, teknik *rotation*, *zoom*, dan *brightness* masing-masing berhasil meningkatkan *F1-score* sebesar 1,81% menjadi 96,15%, sementara augmentasi *translation* memberikan peningkatan sebesar 1,66% menjadi 96%. Augmentasi geometris (*rotation, zoom,*

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Author1, Author2, Author3

translation) mensimulasikan variasi yang realistis pada citra, seperti perbedaan sudut kamera, jarak fokus, dan posisi mata pasien, sehingga model belajar mengekstraksi fitur katarak yang lebih invarian terhadap transformasi spasial. Temuan ini konsisten dengan hasil penelitian Krisnawan & Setiawan (2024) yang menunjukkan bahwa teknik augmentasi geometris seperti *rotation* dan *zoom* terbukti meningkatkan kemampuan generalisasi model secara konsisten khususnya pada dataset berukuran kecil.

Augmentasi *flip* menghasilkan nilai yang identik dengan model *baseline*, sementara augmentasi *contrast* mengakibatkan penurunan kecil sebesar 0,22%. Katarak pada citra fundus retina dapat diidentifikasi melalui perubahan pada lensa kristalin yang ditandai dengan pergeseran suhu warna citra retina menjadi lebih hangat serta penurunan kontras akibat hamburan cahaya intraokular (Gonzalez-Amador et al., 2022). Modifikasi kontras secara acak berpotensi mengurangi ketajaman suatu citra, sehingga akan menyulitkan model dalam membedakan kedua kelas secara konsisten. Dampak negatif paling signifikan ditunjukkan oleh augmentasi *hue* dan *saturation*, dengan penurunan F1-score masing-masing sebesar 11,99% dan 44,34%. Hal ini disebabkan oleh warna yang bukan merupakan fitur yang relevan baik pada mata katarak maupun normal sehingga adanya modifikasi warna menyebabkan model untuk mempelajari fitur yang tidak relevan pada citra asli, atau bahkan mendistorsi warna asli yang terdapat pada citra fundus.

Secara keseluruhan, analisis ini menunjukkan bahwa dalam mengatasi masalah keterbatasan data, augmentasi dapat meningkatkan kualitas model melalui terbentuknya sampel data yang lebih beragam. Namun, efektivitas augmentasi pada CNN sangat ditentukan oleh teknik, konfigurasi parameter, dan juga jenis data yang digunakan.

4. Kesimpulan

Penelitian ini menganalisis pengaruh berbagai teknik augmentasi citra terhadap performa model CNN dalam klasifikasi citra fundus mata normal dan katarak untuk mengatasi keterbatasan data. Model *baseline* tanpa augmentasi telah menunjukkan kinerja yang baik dengan akurasi 95,24% dan *F1-score* 94,34%, yang menjadikannya acuan perbandingan yang valid.

Penerapan augmentasi tertentu terbukti mampu meningkatkan kemampuan generalisasi model. Augmentasi *rotation*, *zoom*, dan *brightness* memberikan hasil terbaik dengan peningkatan *F1-score* hingga 96,15%, sedangkan *translation* juga memberikan peningkatan positif dengan *F1-score* sebesar 96%. Hasil tersebut menunjukkan bahwa augmentasi geometris dan pencahayaan mampu membantu model mengenali variasi kondisi citra fundus yang realistis.

Sebaliknya, tidak semua teknik augmentasi memberikan dampak positif. Augmentasi *hue* dan *saturation* justru menyebabkan penurunan performa yang cukup signifikan, dengan F1-score terendah mencapai 50%. Hal ini menunjukkan bahwa perubahan warna yang berlebihan dapat mendistorsi karakteristik asli citra fundus sehingga model mempelajari fitur yang tidak relevan. Augmentasi *contrast* juga menunjukkan sedikit penurunan performa sebesar 0,22%, sementara *flip* tidak memberikan perubahan yang berarti dibandingkan model *baseline*.

Penelitian ini menunjukkan bahwa dengan konfigurasi yang tepat, augmentasi dapat menjadi solusi alternatif yang patut dipertimbangkan untuk mengatasi keterbatasan data guna meningkatkan kualitas model CNN dalam klasifikasi mata normal dan katarak. Namun sebaliknya, augmentasi dengan konfigurasi yang tidak tepat justru bisa berakibat fatal terhadap kualitas dataset dan model secara keseluruhan. Diharapkan hasil penelitian ini dapat menjadi referensi yang relevan bagi pengembangan sistem klasifikasi citra medis yang lebih akurat serta aplikatif kedepannya.

Referensi

Al-Fahrezi, M. A. (2025). Pengaruh Augmentasi Data Terhadap Akurasi Pelatihan Model CNN untuk Klasifikasi Jenis Ikan. *JITSI : Jurnal Ilmiah Teknologi Sistem Informasi*, *6*(2), 177–185. https://doi.org/10.62527/jitsi.6.2.471

Gonzalez-Amador, E., Arines, J., Charlón, P., Garcia-Porta, N., Abraldes, M. J., & Acosta, E. (2022). Improvement of Retinal Images Affected by Cataracts. *Photonics*, *9*(4). https://doi.org/10.3390/photonics9040251

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Author1, Author2, Author3

Krisnawan, Y. J., & Setiawan, W. (2024). Implementasi Convolutional Neural Network (CNN) untuk Deteksi Katarak pada Citra Mata. *Jurnal Ekonomi Manajemen Sistem Informasi*, *6*(1), 443–450. https://doi.org/10.38035/jemsi.v6i1.3066

Lapp, T., Wacker, K., Heinz, C., Maier, P., Eberwein, P., & Reinhard, T. (2023). Cataract Surgery-Indications, Techniques, and Intraocular Lens Selection. *Deutsches Arzteblatt International*, *120*(21–22), 377–386. https://doi.org/10.3238/arztebl.m2023.0028

Marcella, D., & Devella, S. (2022). *Klasifikasi Penyakit Mata Menggunakan Convolutional Neural Network Dengan Arsitektur VGG-19*. *3*(1), 60–70. https://doi.org/10.35957/algoritme.v3i1.3331

Naufal, M. F., Kusuma, S. F., Tanus, K. C., Sukiwun, R. V., Kristiano, J., Lieyanto, J. O., & R., D. C. (2021). Analisis Perbandingan Algoritma Klasifikasi Citra Chest X-ray Untuk Deteksi Covid-19. *Teknika*, *10*(2), 96–103. https://doi.org/10.34148/teknika.v10i2.331

Noviada Kwartawati, N., Isti Harjanti, A., & Rahayu, S. (2023). Operasi Katarak di RS Telogorejo. *Jurnal Pengabdian Dan Pemberdayaan Masyarakat*, *3*(1), 91–98. https://doi.org/10.31331/manggali.v3i1.2433

Ramadhani, F., Satria, A., & Salamah, S. (2023). Implementasi Algoritma Convolutional Neural Network dalam Mengidentifikasi Dini Penyakit pada Mata Katarak. *Sudo Jurnal Teknik Informatika*, *2*(4), 167–175. https://doi.org/10.56211/sudo.v2i4.408

Rashed, B. M., & Popescu, N. (2023). Performance Investigation for Medical Image Evaluation and Diagnosis Using Machine-Learning and Deep-Learning Techniques. *Computation*, *11*(3). https://doi.org/10.3390/computation11030063

Sasongko, T. B., & Amrullah, A. (2023). *Analisis Efek Augmentasi Dataset Dan Fine Tune Pada Algoritma Pre-Trained Convolutional Neural Network (Cnn) Effect Analysis of Dataset Augmentation and Fine Tune on Pre- Trained Convolutional Neural Network (Cnn) Algorithm*. *10*(4), 763–768. https://doi.org/10.25126/jtiik.2023106583

Steinmetz, J. D., Bourne, R. R. A., Briant, P. S., Flaxman, S., Taylor, H. R., Jonas, J. B., Abdoli, A., Abrha, W. A., Abualhasan, A., Abu-Gharbieh, E., Adal, T. G., Afshin, A., Ahmadieh, H., Alemayehu, W., Alemzadeh, S. A., Alfaar, A. S., Alipour, V., Androudi, S., Arabloo, J., … Vos, T. (2021). Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: The Right to Sight: An analysis for the Global Burden of Disease Study. *The Lancet Global Health*, *9*(2), e144–e160. https://doi.org/10.1016/S2214-109X(20)30489-7

Yang, S. R., Yang, H. C., Shen, F. R., & Zhao, J. (2025). Image Data Augmentation for Deep Learning: A Survey. *Ruan Jian Xue Bao/Journal of Software*, *36*(3), 1390–1412. https://doi.org/10.13328/j.cnki.jos.007263

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License
Copyright © 2026 Author1, Author2, Author3