

KLASIFIKASI TINGKAT KEPARAHAN DIABETIC RETINOPATHY MENGUNAKAN DENSENET201 DENGAN AUGMENTASI DATA ADAPTIF

Fadhli Rajwaa Rahmana, Jati Sasongko Wibowo

Teknik Informatika, Universitas Stikubank Semarang

Jl. Tri Lomba Juang, Mugassari, Kec. Semarang Sel., Kota Semarang, Indonesia

fadhlirajwaarahmana@gmail.com

ABSTRAK

Diabetic Retinopathy (DR) adalah komplikasi diabetes melitus yang menjadi penyebab utama kebutaan pada populasi usia produktif. Diagnosis DR saat ini dilakukan melalui pemeriksaan fundus retina manual oleh dokter spesialis mata, yang memiliki keterbatasan dalam efisiensi waktu, subjektivitas penilaian, dan tidak feasible untuk skrining massal pada populasi diabetes yang terus meningkat. Penelitian ini bertujuan membangun model klasifikasi otomatis untuk memprediksi lima tingkat keparahan DR (No_DR, Mild, Moderate, Severe, Proliferative_DR) dari citra fundus retina. Metode yang digunakan adalah Convolutional Neural Network berbasis arsitektur DenseNet201 dengan transfer learning dua tahap dan augmentasi data adaptif per kelas pada dataset APTOS 2019 (3.662 citra). Strategi augmentasi adaptif mengurangi ketidakseimbangan kelas dari rasio 9,35:1 menjadi 3,32:1. Model mencapai test accuracy 95,35% dengan F1-Score rata-rata 0,955, recall tinggi untuk kelas Severe (0,980) dan Proliferative_DR (0,928), serta peningkatan akurasi sebesar 15–27 poin persentase dibanding penelitian terdahulu. Augmentasi adaptif berhasil meningkatkan sensitivitas model terhadap kelas-kelas kritis tanpa mengorbankan performa kelas lainnya. Model ini berpotensi digunakan sebagai sistem bantu skrining DR otomatis di fasilitas kesehatan untuk deteksi dini dan prioritas pasien yang memerlukan intervensi medis segera.

Kata kunci : *Diabetic Retinopathy, DenseNet201, Transfer Learning, Augmentasi Data Adaptif, Deep Learning, Klasifikasi Citra Medis*

1. PENDAHULUAN

Diabetes Melitus merupakan penyakit kronis yang ditandai dengan tingginya kadar gula darah akibat gangguan produksi atau fungsi insulin. Salah satu komplikasi kronis yang serius dari diabetes adalah Diabetic Retinopathy (DR), yaitu kerusakan pembuluh darah retina yang dapat menyebabkan kebutaan permanen jika tidak terdeteksi sejak dini. DR merupakan penyebab utama kebutaan pada populasi usia produktif secara global, dengan risiko meningkat seiring durasi dan tingkat keparahan diabetes. Diagnosis DR saat ini dilakukan melalui pemeriksaan fundus retina oleh dokter spesialis mata untuk mengidentifikasi lesi-lesi patologis seperti mikroaneurisma, perdarahan, dan eksudat. Namun, proses diagnosis manual ini memiliki keterbatasan signifikan: memerlukan waktu lama, bersifat subjektif tergantung pengalaman dokter, rentan terhadap kelelahan visual, dan ketersediaan dokter spesialis mata yang terbatas sehingga tidak feasible untuk skrining massal pada populasi diabetes yang terus meningkat.

Teknologi Deep Learning, khususnya Convolutional Neural Network (CNN), menawarkan solusi potensial untuk mengatasi keterbatasan tersebut dengan memungkinkan deteksi dan klasifikasi DR secara otomatis dari citra fundus retina. Tantangan utama yang dihadapi dalam mengembangkan model klasifikasi DR adalah ketidakseimbangan kelas ekstrem (class imbalance) pada dataset citra fundus retina. Pada dataset APTOS 2019, kelas retina normal (No_DR) mendominasi dengan 1.805 sampel (49,3%),

sedangkan kelas dengan tingkat keparahan tinggi seperti Severe hanya memiliki 193 sampel (5,3%), menghasilkan rasio ketidakseimbangan sebesar 9,35:1. Kondisi ini menyebabkan model cenderung bias memprediksi kelas mayoritas dan mengabaikan kelas minoritas yang secara klinis paling kritis untuk dideteksi [1].

Penelitian ini bertujuan membangun model klasifikasi tingkat keparahan Diabetic Retinopathy yang lebih akurat dengan mengatasi ketidakseimbangan kelas melalui strategi augmentasi data adaptif per kelas. Model dibangun menggunakan arsitektur DenseNet201 yang memiliki konektivitas padat untuk aliran gradien yang lebih baik, dikombinasikan dengan transfer learning dua tahap. Strategi penyelesaian masalah meliputi: (1) menggunakan augmentasi data adaptif yang memberikan intensitas transformasi berbeda untuk setiap kelas sesuai dengan jumlah sampel dan karakteristik klinisnya untuk menyeimbangkan distribusi data, (2) menerapkan transfer learning dua tahap (frozen training dan fine-tuning) untuk memaksimalkan pemanfaatan pengetahuan dari model pre-trained ImageNet, (3) merancang custom classifier head yang optimal untuk tugas klasifikasi lima kelas, dan (4) mengevaluasi model secara komprehensif menggunakan multiple metrik untuk memastikan sensitivitas tinggi pada kelas-kelas kritis (Severe dan Proliferative_DR) yang paling membutuhkan intervensi medis segera.

2. TINJAUAN PUSTAKA

Penelitian klasifikasi Diabetic Retinopathy menggunakan *deep learning* telah berkembang pesat dalam beberapa tahun terakhir, namun masih menghadapi tantangan signifikan terkait ketidakseimbangan data dan variasi visual citra fundus retina. Bagian ini membahas empat komponen landasan teori yang saling terkait: pertama, konsep medis Diabetic Retinopathy dan sistem klasifikasi lima tingkat keparahan yang menjadi target prediksi; kedua, arsitektur *Convolutional Neural Network* dengan fokus pada mekanisme *dense connectivity* pada DenseNet201 yang membedakannya dari arsitektur CNN konvensional; ketiga, teknik *transfer learning* sebagai strategi pelatihan efisien untuk dataset medis berukuran terbatas dan augmentasi data adaptif sebagai solusi ketidakseimbangan kelas; keempat, tinjauan sistematis penelitian-penelitian terdahulu yang menggunakan dataset APTOS 2019 untuk mengidentifikasi gap penelitian dan posisi kontribusi penelitian ini.

2.1. Diabetic Retinopathy

Diabetic Retinopathy adalah komplikasi mikrovaskular dari diabetes melitus yang menyebabkan kerusakan progresif pada pembuluh darah retina. Berdasarkan *International Classification of Diabetic Retinopathy* (ICDR), DR diklasifikasikan menjadi lima tingkat keparahan: No_DR (retina normal tanpa kelainan), Mild (mikroaneurisma minimal), Moderate (mikroaneurisma multipel dengan perdarahan dan eksudat), Severe (perdarahan ekstensif, *cotton wool spots*, IRMA), dan Proliferative_DR (neovaskularisasi, perdarahan vitreus). Klasifikasi yang akurat sangat penting karena menentukan urgensi tindakan medis yang diperlukan untuk mencegah kehilangan penglihatan permanen [2].

2.2. Convolutional Neural Network dan DenseNet201

Convolutional Neural Network (CNN) adalah arsitektur *deep learning* yang dirancang khusus untuk memproses data berbentuk grid seperti citra. CNN terdiri dari lapisan konvolusi yang secara otomatis mempelajari representasi fitur hierarkis dari citra input, mulai dari fitur tingkat rendah (tepi, tekstur) hingga fitur tingkat tinggi (objek kompleks), sehingga CNN mampu mendeteksi pola-pola halus yang mungkin terlewat oleh pengamat manusia. DenseNet201 merupakan varian arsitektur CNN dengan konsep konektivitas padat (*dense connectivity*) di mana setiap lapisan dalam satu blok densitas menerima input dari semua lapisan sebelumnya melalui mekanisme konkatensi. Struktur ini memiliki keunggulan dalam mengurangi masalah *vanishing gradient*, mendorong *feature reuse* sehingga lebih efisien dalam penggunaan parameter, serta memperkuat propagasi fitur dan gradien. Penelitian menunjukkan bahwa DenseNet201 mencapai akurasi

96% untuk deteksi biner DR, setara dengan arsitektur *state-of-the-art* lainnya [1].

2.3. Transfer Learning dan Augmentasi Data

Transfer learning adalah teknik di mana model yang telah dilatih pada dataset besar seperti ImageNet digunakan sebagai titik awal untuk tugas baru. Pendekatan ini sangat bermanfaat pada domain citra medis yang umumnya memiliki dataset berukuran terbatas, karena bobot *pre-trained* sudah mengandung representasi fitur visual universal yang dapat ditransfer ke tugas spesifik. Augmentasi data adalah teknik untuk memperbanyak data pelatihan secara artifisial dengan menerapkan transformasi seperti rotasi, pergeseran, *zoom*, dan perubahan kecerahan. Berbeda dengan *oversampling* statis, augmentasi menghasilkan variasi citra baru secara dinamis sehingga mengurangi risiko *overfitting*. Untuk mengatasi *class imbalance*, augmentasi adaptif memberikan intensitas transformasi yang berbeda untuk setiap kelas sesuai dengan jumlah sampel dan karakteristik datanya, yang terbukti penting untuk menyeimbangkan distribusi data dan meningkatkan kemampuan generalisasi model.

2.4. Penelitian Terkait

Beberapa penelitian telah menggunakan *deep learning* untuk klasifikasi DR pada dataset APTOS 2019. Rizal menggunakan arsitektur EfficientNet dengan resolusi 224×224 dan mencapai akurasi 79,8% untuk klasifikasi lima kelas DR [3]. Syahrul menerapkan CNN custom dengan lima blok konvolusi dan mencapai akurasi 91,10%, serta secara eksplisit menyatakan bahwa augmentasi data sangat penting untuk mencapai kinerja yang memuaskan [4]. Hendra menggunakan EfficientNetB3 dengan augmentasi standar dan mencapai validation accuracy 75,7%, serta menyoroti pentingnya augmentasi untuk mengatasi *overfitting* dan kinerja buruk pada kelas minoritas [1]. Pratiwi menggunakan EfficientNet-B0 untuk klasifikasi penyakit mata dan mencapai validation accuracy 94% dengan F1-Score untuk DR sebesar 0,98 [5].

Alifian mengembangkan model CNN dengan arsitektur VGG-16 yang dimodifikasi untuk klasifikasi DR pada dataset APTOS 2019 dan berhasil mencapai akurasi 96,45%, namun membutuhkan waktu komputasi yang lebih lama [6]. Chazar menerapkan pendekatan augmentasi citra fundus terarah pada EfficientNet-B0 untuk mengatasi ketidakseimbangan kelas dan berhasil meningkatkan akurasi dari 73,84% menjadi 82,56% dengan peningkatan signifikan pada kelas minoritas [7]. Penelitian ensemble CNN oleh Faurina menggunakan kombinasi DenseNet201, InceptionV3, dan MobileNetV2 dengan metode average voting dan mencapai akurasi 95,6% dengan F1-Score terendah 91,3% untuk klasifikasi level non-proliferasi DR [8].

Mengidentifikasi ketidakseimbangan data sebagai penyebab utama *overfitting* dan performa

yang sangat buruk pada kelas minoritas . Penelitian-penelitian ini menunjukkan bahwa arsitektur CNN modern dapat mencapai performa yang baik, namun masih terdapat ruang untuk peningkatan terutama dalam menangani ketidakseimbangan kelas dan meningkatkan sensitivitas pada kelas dengan tingkat keparahan tinggi [9] .

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan eksperimental kuantitatif berbasis *deep learning* untuk membangun model klasifikasi tingkat keparahan Diabetic Retinopathy. Metodologi dirancang secara sistematis untuk menjawab dua tantangan utama pada dataset citra medis retinopati diabetik: variasi visual yang tinggi akibat perbedaan perangkat akuisisi dan kondisi pencahayaan, serta ketidakseimbangan kelas ekstrem yang dapat menurunkan kemampuan generalisasi model secara signifikan apabila tidak ditangani dengan strategi yang tepat .

3.1. Dataset

Penelitian ini menggunakan dataset APTOS 2019 Blindness Detection yang terdiri dari 3.662 citra fundus retina berwarna (RGB) dengan resolusi bervariasi . Setiap citra telah dilabeli oleh dokter spesialis mata ke dalam lima kelas tingkat keparahan DR. Distribusi kelas pada dataset asli sangat tidak seimbang sebagaimana ditampilkan pada Tabel 1.

Tabel 1. Distribusi Dataset APTOS 2019 Sebelum Augmentasi

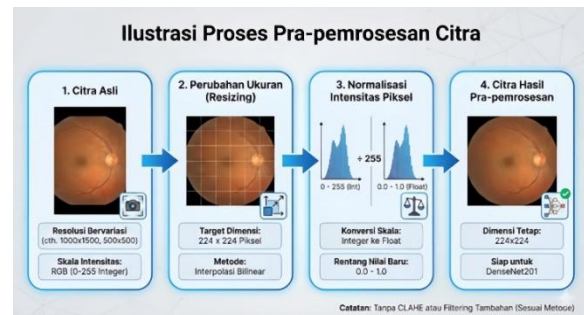
Kelas	Label	Jumlah Sampel	Persentase
No_DR	0	1.805	49,3%
Moderate	2	999	27,3%
Mild	1	370	10,1%
Proliferative_DR	4	295	8,1%
Severe	3	193	5,3%
TOTAL		3.662	100%

Dari Tabel 1 dapat dilihat bahwa kelas No_DR mendominasi dengan 1.805 sampel (49,3%), sedangkan kelas Severe hanya memiliki 193 sampel (5,3%), menghasilkan rasio ketidakseimbangan sebesar 9,35:1. Kondisi ini berpotensi menyebabkan model cenderung bias memprediksi kelas mayoritas dan mengabaikan kelas dengan tingkat keparahan tinggi yang secara klinis justru paling kritis untuk dideteksi.

3.2. Pra-pemrosesan Citra

Tahapan pra-pemrosesan citra yang dilakukan secara berurutan dilustrasikan pada Gambar 1. Proses ini meliputi dua langkah utama: pertama, seluruh citra input dengan berbagai resolusi diubah ukurannya (*resizing*) menjadi dimensi standar 224×224 piksel menggunakan metode interpolasi bilinear agar sesuai dengan spesifikasi input arsitektur *pre-trained* ImageNet . Kedua, dilakukan normalisasi intensitas piksel dengan mengonversi nilai dari

rentang 0–255 menjadi rentang 0–1. Langkah normalisasi ini krusial untuk mempercepat konvergensi algoritma optimasi dan menjaga stabilitas gradien selama proses pelatihan .



Gambar 1. Ilustrasi Proses Pra-pemrosesan Citra

3.3. Augmentasi Data Adaptif

Untuk mengatasi ketidakseimbangan kelas, diterapkan strategi augmentasi data adaptif yang memberikan intensitas transformasi berbeda untuk setiap kelas. Kelas No_DR tidak diberikan augmentasi tambahan untuk mempertahankan karakteristik aslinya, sedangkan kelas-kelas lain diaugmentasi dengan target sekitar 5.000 sampel per kelas. Parameter transformasi dirancang sesuai karakteristik klinis lesi yang menjadi ciri khas kelas tersebut, sebagaimana ditampilkan pada Tabel 2.

Tabel 2. Parameter Augmentasi Adaptif per Kelas

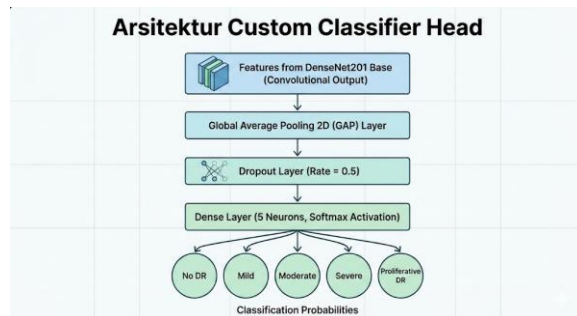
Kelas	Rotasi	Shift (W/H)	Shea/Zoom	Flip	Brightness
No_DR	—	—	—	—	—
Mild	10°	0,1	0,1/0,15	Ya	[0,9–1,1]
Moderate	15°	0,1	0,1/0,10	Ya	[0,9–1,1]
Severe	20°	0,1	0,1/0,10	Ya	[0,8–1,2]
Proliferative	10°	0,1	0,1/0,15	Ya	[0,9–1,1]

Dari Tabel 2 dapat dilihat bahwa kelas Severe diberikan transformasi paling agresif dengan rotasi hingga 20° dan rentang kecerahan yang lebih lebar untuk menghasilkan lebih banyak variasi. Hasil augmentasi memperluas dataset menjadi 23.012 citra dengan distribusi: Moderate (5.994), Mild (5.180), Severe (5.018), Proliferative_DR (5.015), dan No_DR (1.805). Rasio ketidakseimbangan berkurang menjadi sekitar 3,32:1.

3.4. Arsitektur Model

Model dibangun menggunakan arsitektur DenseNet201 *pre-trained* ImageNet sebagai *backbone* (ekstraktor fitur), dikombinasikan dengan *custom classifier head* yang dirancang khusus untuk tugas klasifikasi lima kelas DR. Struktur *custom head* terdiri dari lapisan-lapisan berikut secara berurutan: Global Average Pooling 2D untuk mereduksi dimensi *feature map*, Conv2D dengan 64 filter dan aktivasi ReLU yang diregularisasi dengan L2 (1e-4), MaxPooling2D untuk *downsampling*, Flatten untuk konversi ke vektor 1D, Dense layer dengan 128

neuron dan regularisasi L2, Batch Normalization untuk stabilisasi distribusi aktivasi, Dropout dengan *rate* 0,5 untuk mencegah *overfitting*, Dense layer kedua dengan 64 neuron, Batch Normalization, dan akhirnya output Dense layer dengan 5 neuron dan aktivasi Softmax untuk menghasilkan probabilitas prediksi setiap kelas. Total parameter model sekitar 20 juta dengan bobot pretrained pada backbone DenseNet201.



Gambar 2. Arsitektur Model DenseNet201 dengan Custom Classifier Head

3.5. Strategi Pelatihan

Pelatihan model dilakukan dalam dua tahap menggunakan strategi transfer learning untuk memaksimalkan pemanfaatan pengetahuan dari ImageNet. Tahap 1 – Frozen Training (15 epoch): Seluruh lapisan DenseNet201 dibekukan dengan mengatur *trainable=False*, sehingga hanya custom classifier head yang dilatih menggunakan optimizer Adam dengan learning rate 1×10^{-4} . Tujuan tahap ini adalah menstabilkan bobot head yang diinisialisasi secara acak sebelum melakukan fine-tuning pada backbone. Tahap 2 – Fine-Tuning (25 epoch): Setelah head mencapai stabilitas, sebanyak 30 lapisan terakhir DenseNet201 dibuka untuk ikut dilatih dengan learning rate yang jauh lebih kecil (1×10^{-5}). Lapisan-lapisan awal yang menangkap fitur universal tetap dibekukan untuk mempertahankan pengetahuan yang sudah optimal.

Fungsi loss yang digunakan adalah Categorical Crossentropy dengan metrik evaluasi accuracy. Tiga callback diterapkan: ModelCheckpoint untuk menyimpan model dengan validation accuracy terbaik, EarlyStopping dengan *patience=5* untuk menghentikan pelatihan jika validation loss tidak membaik selama 5 epoch berturut-turut, dan ReduceLROnPlateau dengan *factor=0,5* dan *patience=2* untuk menurunkan learning rate secara adaptif jika validation loss stagnan.

3.6. Pembagian Data dan Evaluasi

Dataset hasil augmentasi (23.012 citra) dibagi menjadi tiga subset menggunakan *stratified split*: data latih (80% = 18.409 citra), data validasi (10% = 2.301 citra), dan data uji (10% = 2.302 citra). Data uji dipisahkan sejak awal dan tidak pernah digunakan selama pelatihan untuk memastikan evaluasi yang objektif. Model dievaluasi menggunakan beberapa

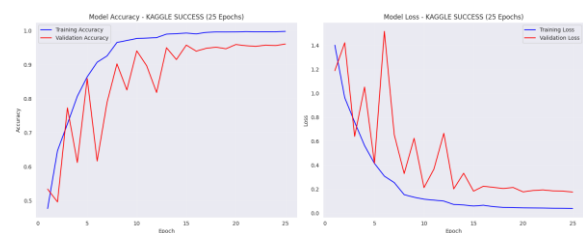
metrik: *accuracy* untuk mengukur proporsi prediksi benar, *precision* untuk mengukur proporsi prediksi positif yang benar, *recall* untuk mengukur proporsi sampel aktual yang terdeteksi, F1-Score sebagai rata-rata harmonik *precision* dan *recall*, serta *confusion matrix* untuk analisis detail pola kesalahan klasifikasi antar kelas.

4. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil eksperimen yang telah dilakukan, meliputi hasil pelatihan model selama 40 epoch, evaluasi performa pada data uji yang independen, analisis metrik per kelas, analisis pola kesalahan melalui *confusion matrix*, perbandingan dengan penelitian terdahulu, serta pembahasan mendalam mengenai efektivitas pendekatan yang diusulkan dan keterbatasan model.

4.1. Hasil dan Pelatihan Model

Proses pelatihan model berlangsung selama 40 epoch total (15 epoch *frozen training* + 25 epoch *fine-tuning*) dengan total waktu pelatihan sekitar 6 jam pada GPU Tesla T4. Kurva pelatihan menunjukkan pola konvergensi yang stabil dan konsisten tanpa fluktuasi besar. Akurasi *training* meningkat secara bertahap dan mencapai 99,78% pada epoch terakhir, sementara akurasi validasi mencapai 96,05%. *Validation loss* stabil di sekitar 0,1766. Gap antara akurasi *training* dan validasi sebesar 3,73% mengindikasikan adanya *overfitting* moderat, namun masih dalam batas wajar untuk model *deep learning* yang kompleks dengan 20 juta parameter. Mekanisme regularisasi yang diterapkan (Dropout 0,5, L2 regularization, Batch Normalization) berhasil mencegah *overfitting* yang parah.



Gambar 3. Kurva Training dan Validation Accuracy per Epoch

4.2. Hasil Evaluasi pada Data Uji

Evaluasi pada 2.302 citra data uji yang independen menghasilkan performa yang sangat baik sebagaimana ditampilkan pada Tabel 3.

Tabel 3. Hasil Evaluasi Model pada Data Uji

Metrik	Nilai
Test Accuracy	95,35%
Test Loss	0,1850
Macro F1-Score	0,955
Weighted F1-Score	0,953
Prediksi Benar	2.195 dari 2.302
Prediksi Salah	107 dari 2.302

Dari Tabel 3 dapat dilihat bahwa model mencapai *test accuracy* 95,35%, yang berarti dari 2.302 citra pada data uji, model berhasil mengklasifikasikan 2.195 citra dengan benar dan hanya salah pada 107 citra (4,65%). Konsistensi antara akurasi validasi (96,05%) dan akurasi uji (95,35%) dengan gap hanya 0,70% menunjukkan bahwa model memiliki kemampuan generalisasi yang baik dan tidak hanya menghafal pola pada data latih.

4.3. Analisis Performa Per Kelas

Metrik evaluasi per kelas ditampilkan pada Tabel 4, yang menunjukkan performa yang seimbang di seluruh kelas dengan F1-Score berkisar antara 0,942 hingga 0,965.

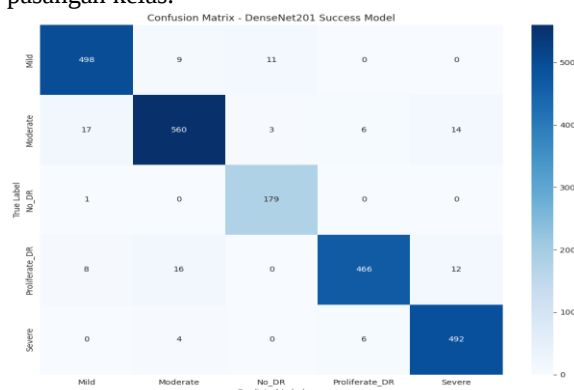
Tabel 4. Metrik Evaluasi Per Kelas (Precision, Recall, F1-Score)

Kelas	Precision	Recall	F1-Score	Jumlah Sampel
Mild	0,950	0,961	0,956	518
Moderate	0,951	0,933	0,942	600
No_DR	0,927	0,994	0,960	180
Proliferative_DR	0,975	0,928	0,951	502
Severe	0,950	0,980	0,965	502
Macro Average	0,951	0,959	0,955	2.302

Dari Tabel 4 dapat dilihat bahwa kelas No_DR menunjukkan performa terbaik dengan *recall* 0,994, yang berarti 99,4% retina normal berhasil terdeteksi dengan benar. Kelas Severe memiliki F1-Score tertinggi (0,965) dengan *recall* 0,980, menunjukkan bahwa model sangat sensitif terhadap kasus-kasus parah yang memerlukan intervensi segera. Kelas Proliferative_DR memiliki *precision* tertinggi (0,975), yang berarti 97,5% prediksi Proliferative_DR adalah benar. Hasil ini menunjukkan bahwa strategi augmentasi adaptif berhasil meningkatkan sensitivitas model terhadap kelas-kelas minoritas yang secara klinis paling kritis, tanpa mengorbankan performa pada kelas lainnya.

4.4. Analisis Confusion Matrix

Analisis *confusion matrix* mengungkapkan pola kesalahan klasifikasi model secara detail untuk setiap pasangan kelas.



Gambar 4. Confusion Matrix pada Data Uji (Heatmap 5x5)

Berdasarkan analisis *confusion matrix*, sebagian besar kesalahan klasifikasi terjadi pada kelas-kelas yang berdekatan secara klinis. Kesalahan terbesar terjadi pada pasangan Moderate ↔ Severe (14+9=23 kesalahan) dan Proliferative_DR ↔ Severe (16+16=32 kesalahan). Pola kesalahan ini dapat dipahami karena perbedaan visual pada kasus *borderline* antara tingkat keparahan yang berdekatan memang sangat halus, bahkan untuk dokter mata berpengalaman sekalipun. Yang penting dicatat adalah bahwa model sangat jarang melakukan kesalahan ekstrem seperti mengklasifikasikan Proliferative_DR sebagai No_DR atau sebaliknya, menunjukkan bahwa model telah mempelajari representasi fitur yang bermakna secara klinis.

4.5. Perbandingan dengan Penelitian Terdahulu

Tabel 5 menampilkan perbandingan performa model yang dikembangkan dengan penelitian-penelitian terdahulu yang menggunakan dataset APTOS 2019 yang sama.

Tabel 5. Perbandingan Hasil dengan Penelitian Terdahulu pada Dataset APTOS 2019

Penelitian	Arsitektur	Resolusi	Akurasi	F1-Score
Rizal dkk. (2020)	EfficientNet	224×224	79,8%	—
Syahrul & Sasongko (2022)	CNN Custom	—	91,10%	—
Kusuma & Poetro (2025)	EfficientNetB3	—	75,7%	—
Penelitian Ini	DenseNet201	224×224	95,35%	0,955

Dari Tabel 5 dapat dilihat bahwa hasil penelitian ini menunjukkan peningkatan signifikan dibanding penelitian terdahulu: peningkatan sebesar +15,55 poin persentase dibanding Rizal dkk., +4,25 poin persentase dibanding Syahrul & Sasongko, dan +19,65 poin persentase dibanding Kusuma & Poetro. Peningkatan ini dapat dikaitkan dengan penggunaan DenseNet201 yang memiliki *dense connectivity* untuk *feature reuse* dan aliran gradien yang lebih baik, strategi augmentasi adaptif per kelas yang menangani ketidakseimbangan secara lebih efektif, serta *transfer learning* dua tahap yang sistematis.

4.6. Pembahasan

Hasil penelitian mendemonstrasikan bahwa kombinasi arsitektur DenseNet201, augmentasi data adaptif, dan transfer learning dua tahap merupakan pendekatan yang sangat efektif untuk klasifikasi tingkat keparahan Diabetic Retinopathy. Strategi augmentasi yang tidak mengaugmentasi kelas No_DR ternyata tidak menurunkan performa pada kelas tersebut, justru mencapai recall 0,994. Di sisi lain, augmentasi intensif pada kelas minoritas berhasil meningkatkan sensitivitas model terhadap kasus-kasus

kritis dengan recall tinggi untuk Severe (0,980) dan Proliferative_DR (0,928).

Gap minimal antara validation accuracy (96,05%) dan test accuracy (95,35%) sebesar 0,70% mengindikasikan bahwa model memiliki kemampuan generalisasi yang baik. Dalam konteks skrining medis, sensitivitas tinggi pada kasus kritis lebih diutamakan dibanding precision, karena satu kasus parah yang terlewat dapat berakibat pada kebutaan permanen, sedangkan false positive masih dapat dikoreksi melalui pemeriksaan lanjutan oleh dokter.

Model ini berpotensi besar untuk diintegrasikan sebagai sistem bantu skrining (computer-aided diagnosis) di fasilitas kesehatan, terutama di klinik primer yang tidak memiliki dokter spesialis mata. Sistem dapat membantu mengidentifikasi pasien dengan DR tingkat parah untuk dirujuk segera, serta memprioritaskan pemeriksaan lanjutan berdasarkan tingkat keparahan yang diprediksi. Namun, model masih memiliki keterbatasan seperti indikasi overfitting moderat (gap 3,73%), kompleksitas komputasi yang memerlukan GPU, dan belum melalui validasi klinis prospektif.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil membangun model klasifikasi tingkat keparahan Diabetic Retinopathy menggunakan arsitektur DenseNet201 dengan augmentasi data adaptif dan transfer learning dua tahap. Model mencapai test accuracy sebesar 95,35% dengan F1-Score rata-rata 0,955, menunjukkan performa yang sangat baik dan seimbang pada semua kelas. Strategi augmentasi adaptif efektif mengurangi ketidakseimbangan kelas dari rasio 9,35:1 menjadi 3,32:1 dan meningkatkan sensitivitas model terhadap kelas kritis dengan recall 0,980 untuk Severe dan 0,928 untuk Proliferative_DR. Perbandingan dengan penelitian terdahulu menunjukkan peningkatan akurasi sebesar 4–20 poin persentase pada dataset APTOS 2019 yang sama. Model ini berpotensi digunakan sebagai komponen sistem bantu skrining DR otomatis di fasilitas kesehatan untuk membantu deteksi dini dan prioritas pasien yang memerlukan pemeriksaan lanjutan atau tindakan medis segera.

Penelitian selanjutnya disarankan untuk mengeksplorasi penyeimbangan semua kelas termasuk No_DR agar tidak terjadi ketidakseimbangan baru, pengurangan overfitting melalui teknik regularisasi tambahan seperti label smoothing atau early stopping yang lebih agresif, pengujian model pada citra dengan kualitas yang sangat rendah atau dengan artefak ekstrem untuk menilai robustness, implementasi teknik explainability seperti Grad-CAM untuk menunjukkan area retina yang menjadi dasar prediksi model sehingga meningkatkan kepercayaan dokter terhadap sistem, serta pelaksanaan validasi klinis prospektif pada lingkungan pelayanan kesehatan nyata

dengan melibatkan dokter spesialis mata sebagai pengguna akhir sebelum deployment secara luas.

DAFTAR PUSTAKA

- [1] Hendra Irawan Wijaya Kusuma and Bagus Satrio Waluyo Poetro, "Deteksi Diabetic Retinopathy Menggunakan Convolutional Neural Network Berbasis Efficientnet Dan Grad-Cam Pada Citra Fundus Retina," *J. Rekayasa Sist. Inf. dan Teknol.*, vol. 3, no. 1, pp. 126–138, 2025, doi: 10.70248/jrsit.v3i1.2896.
- [2] M. Zahir and R. Adi Saputra, "Deteksi Penyakit Retinopati Diabetes Menggunakan Citra Mata Dengan Implementasi Deep Learning CNN," *J. Teknoinfo*, vol. 18, no. 1, pp. 121–132, 2024, [Online]. Available: <https://www.kaggle.com/datasets/gunavenkatdodd/eye-diseases-classification>
- [3] S. RIZAL, N. IBRAHIM, N. K. C. PRATIWI, S. SAIDAH, and R. Y. N. FU'ADAH, "Deep Learning untuk Klasifikasi Diabetic Retinopathy menggunakan Model EfficientNet," *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, vol. 8, no. 3, p. 693, 2020, doi: 10.26760/elkomika.v8i3.693.
- [4] F. H. Syahrul and P. S. Sasongko, "Penerapan Convolutional Neural Network Untuk Klasifikasi Tingkat Keparahan Retinopati Diabetik Pada Penderita Diabetes Melitus," *J. Masy. Inform.*, vol. 13, no. 1, pp. 1–14, 2022, doi: 10.14710/jmasif.13.1.42354.
- [5] I. D. A. P. Pratiwi Tentrirajaya and I. G. N. L. Wijayakusuma, "Eye Disease Classification Using EfficientNet-B0 Based on Transfer Learning," *J. Appl. Informatics Comput.*, vol. 9, no. 4, pp. 1415–1422, 2025, doi: 10.30871/jaic.v9i4.9743.
- [6] K. Alifian and H. Wibowo, "Indonesian Journal of Mathematics and Natural Sciences," vol. 46, no. 1, pp. 45–52, 2023.
- [7] C. Chazar, M. A. Adli, and J. Pardede, "Pendekatan Augmentasi Citra Fundus pada Model EfficientNet untuk Klasifikasi Tingkat Keparahan Retinopati Diabetik dengan Dataset Tidak Seimbang," vol. 10, no. 2, pp. 180–194, 2025.
- [8] R. Faurina, E. P. Purwandari, M. T. Pratama, and I. Agustian, "KLASIFIKASI LEVEL NON-PROLIFERATIF RETINOPATI DIABETIK DENGAN ENSEMBLE," vol. VIII, pp. 1–10, 2021.
- [9] S. H. Abdullah, R. Magdalena, and R. Y. N. Fu'adah, "Diabetic Retinopathy Classification Based on Fundus Image Processing and Deep Learning," *J. Electr. Syst. Control Eng.*, vol. 5, no. 2, pp. 84–90, 2022.