

BAB II

TENTANG LOMBA

2.1 Deskripsi Pelaksanaan Lomba

Lomba Karya Tulis Ilmiah Nasional (LKTIN) Carbon 7.0 merupakan kompetisi ilmiah tahunan yang diselenggarakan oleh Himpunan Mahasiswa Kimia Universitas Jambi (HIMKI UNJA). Kompetisi ini menjadi wadah pengembangan potensi dan kreativitas mahasiswa di seluruh Indonesia dalam bidang sains dan teknologi.

Tahun 2025, Carbon 7.0 mengusung tema “*Shaping a Better World Through Science and Technology*”, dengan enam subtema, yaitu: Energi Terbarukan, Pendidikan, Pangan, Lingkungan, Kesehatan, dan Teknologi. Tujuan utama dari LKTIN Carbon 7.0 adalah untuk mendorong mahasiswa menghasilkan gagasan ilmiah dan inovasi yang aplikatif dalam menjawab tantangan pembangunan nasional secara berkelanjutan.



Gambar 2. 1 Logo Resmi LKTIN Carbon 7.0

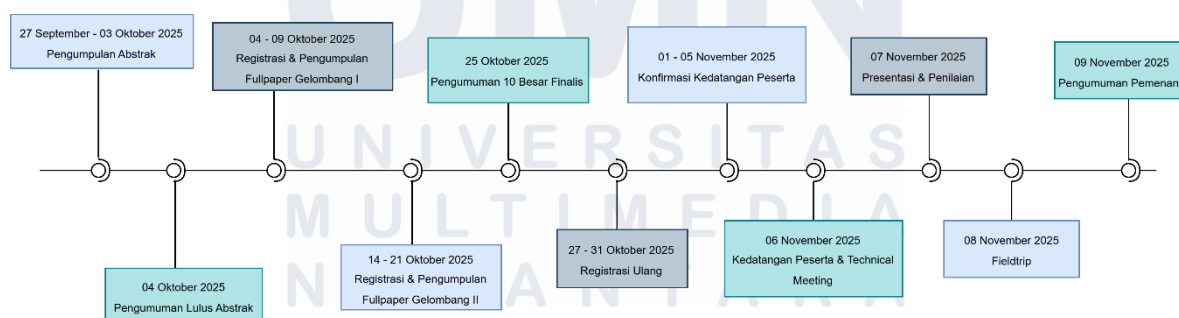
Gambar 2.1 di atas merupakan logo resmi Carbon 7.0 – *Chemistry Fair Born for Nationalism* melambangkan semangat kolaborasi, inovasi, serta rasa nasionalisme mahasiswa Indonesia dalam mengembangkan ilmu pengetahuan, khususnya di bidang kimia. Setiap elemen dalam logo menggambarkan sinergi antar mahasiswa sebagai agen perubahan, kekuatan sains sebagai landasan kemajuan bangsa, dan semangat membara untuk terus berinovasi demi negeri. Logo ini menjadi identitas visual yang mencerminkan nilai-nilai intelektual, kreatif, dan patriotik yang diusung

oleh LKTIN Carbon 7.0

Sebagai bagian dari kegiatan PRO-STEP: *Road to Champion Program*, mahasiswa mengikuti subtema Teknologi dan Kesehatan dengan karya ilmiah berjudul “Klasifikasi Tingkat Keparahan Retinopati Diabetik pada Citra Fundus Berbasis *Vision Transformer* dengan Pemrosesan Citra.” Karya ini berfokus pada penerapan teknologi, khususnya *Convolutional Neural Network* (CNN), dalam membantu deteksi dini retinopati diabetik (komplikasi dari diabetes melitus) melalui citra fundus retina pasien.

Kompetisi ini berskala nasional dan dilaksanakan secara luring di Universitas Jambi. Peserta yang berhasil masuk 10 besar finalis mendapatkan kesempatan untuk mempresentasikan karya ilmiahnya di hadapan dewan juri serta mengikuti kegiatan *field trip* dan *networking session*. Selain memperoleh pengalaman ilmiah, peserta juga berpeluang mendapatkan sertifikat nasional, uang pembinaan, dan trofi penghargaan bagi juara 1, 2, dan 3. Kompetisi ini juga memberikan manfaat besar bagi pengembangan kemampuan akademik, riset, dan komunikasi ilmiah bagi seluruh peserta [6].

2.2 Alur Pendaftaran Lomba



Gambar 2. 2 Alur Pendaftaran dan Pelaksanaan LKTIN Carbon 7.0

Sumber: Himpunan Mahasiswa Kimia Universitas Jambi. (2025). Guidebook Lomba Karya Tulis Ilmiah Nasional Carbon 7.0. Universitas Jambi

Berdasarkan Gambar 2.2, alur pelaksanaan Lomba Karya Tulis Ilmiah Nasional (LKTIN) Carbon 7.0 dimulai dari tahap pengumpulan abstrak yang dilaksanakan pada 27 September–03 Oktober 2025. Selanjutnya, peserta yang dinyatakan lolos abstrak pada 04 Oktober 2025 melanjutkan ke tahap registrasi dan pengumpulan fullpaper

Gelombang I pada 04–09 Oktober 2025, serta Gelombang II pada 14–21 Oktober 2025. Setelah seluruh fullpaper terkumpul, panitia melakukan proses seleksi hingga diumumkannya 10 besar finalis pada 25 Oktober 2025, yang kemudian diwajibkan melakukan registrasi ulang pada 27–31 Oktober 2025 serta konfirmasi kedatangan peserta pada 01–05 November 2025. Tahap final dilaksanakan secara luring yang diawali dengan kedatangan peserta dan technical meeting pada 06 November 2025, dilanjutkan dengan sesi presentasi dan penilaian karya ilmiah pada 07 November 2025. Rangkaian kegiatan final juga mencakup field trip pada 08 November 2025, dan ditutup dengan pengumuman pemenang pada 09 November 2025.

2.3 Portfolio Hasil Karya Lomba

Lomba Karya Tulis Ilmiah Nasional (LKTIN) Carbon merupakan kegiatan ilmiah tahunan yang diselenggarakan oleh Himpunan Mahasiswa Kimia Universitas Jambi (HIMKI UNJA). Sejak awal penyelenggaraannya, kompetisi ini telah menjadi wadah bagi mahasiswa di seluruh Indonesia untuk mengembangkan inovasi berbasis sains dan teknologi yang berorientasi pada keberlanjutan dan kesejahteraan masyarakat. Pada periode Carbon 5.0 (2023), beberapa universitas yang berhasil meraih juara antara lain [6]:

- Juara 1: Universitas Udayana
- Juara 2: Universitas Andalas
- Juara 3: Universitas Jambi
- *Best Poster*: Institut Teknologi Calvin
- *Best Presentation*: Universitas Udayana

Karya-karya yang dihasilkan menunjukkan keberagaman ide dan kemampuan riset mahasiswa Indonesia dalam mengintegrasikan ilmu pengetahuan dengan inovasi teknologi. Kompetisi ini juga mendorong peserta untuk mengembangkan kemampuan berpikir kritis, analitis, serta berkomunikasi secara ilmiah di tingkat nasional.

Melalui konsistensi penyelenggaraan dari tahun ke tahun, LKTIN Carbon terus berperan sebagai ajang prestisius bagi mahasiswa Indonesia dalam menghasilkan gagasan dan inovasi ilmiah yang relevan dengan perkembangan teknologi dan kebutuhan masyarakat.

2.4 Output Lomba yang Akan Dihasilkan

Kegiatan PRO-STEP: Road to Champion Program melalui Lomba Karya Tulis Ilmiah Nasional (LKTIN) Carbon 7.0 menghasilkan beberapa output utama dan pendukung yang menjadi luaran resmi dari keikutsertaan mahasiswa dalam kompetisi akademik tingkat nasional.

1. Karya Tulis Ilmiah

Output utama dari kegiatan lomba ini adalah karya tulis ilmiah (full paper) yang berjudul “Klasifikasi Tingkat Keparahan Retinopati Diabetik pada Citra Fundus Berbasis CNN dengan Pemrosesan Citra”. Karya ilmiah ini disusun berdasarkan tahapan penelitian yang sistematis, mulai dari studi literatur, pengolahan dataset citra fundus, pembangunan dan pelatihan model *deep learning*, hingga analisis hasil menggunakan visualisasi Grad-CAM. Full paper ini menjadi luaran akademik utama yang dinilai oleh dewan juri dan dipresentasikan pada tahap final lomba. Dataset yang digunakan dalam penelitian ini bersumber dari Kaggle *Diabetic Retinopathy Detection* Dataset, yang berisi ribuan gambar citra fundus retina dengan berbagai tingkat keparahan retinopati diabetik. Setiap citra telah diberi label oleh tenaga medis profesional berdasarkan tingkat kerusakan retina.

Tabel 2. 1 Kategori Tingkat Keparahan Retinopati Diabetik

Kelas	Kategori	Deskripsi
0	No_DR	Tidak terdapat tanda-tanda retinopati diabetik pada retina
1	Mild	Terlihat mikroaneurisma awal, namun belum ada perdarahan atau eksudat yang signifikan
2	Moderate	Terdapat perdarahan dan eksudat keras, namun belum terjadi neovaskularisasi
3	Severe	Perdarahan dan kerusakan pembuluh darah menyebar luas di seluruh retina

4	Prolifera _{te} _DR	Terjadi pertumbuhan pembuluh darah baru (neovaskularisasi) dan risiko tinggi kebutaan permanen
---	-----------------------------	--

Berdasarkan Tabel 2.1, dataset yang digunakan dalam penelitian ini terdiri dari lima kelas tingkat keparahan retinopati diabetik, yaitu No_DR, Mild, Moderate, Severe, dan Prolifera_{te}_DR. Setiap kelas mempresentasikan kondisi klinis retina yang berbeda, mulai dari retina normal tanpa tanda kerusakan hingga kondisi prolifera_{tif}e yang memiliki risiko tinggi menyebabkan kebutaan. Pembagian kelas ini menjadi acuan utama dalam proses pelabelan data dan pelatihan model *deep learning*, sehingga model dapat mempelajari perbedaan karakteristik visual pada setiap tingkat keparahan retinopati diabetik secara lebih terstruktur dan sistematis.

Pemilihan metode *Convolutional Neural Network* (CNN) dalam penelitian ini didasarkan pada hasil kajian literatur yang relevan. Salah satu penelitian acuan berjudul “Evaluasi Performasi Ruang Warna pada Klasifikasi *Diabetic Retinopathy* Menggunakan *Convolution Neural Network*” menunjukkan bahwa CNN memiliki kemampuan yang baik dalam mengekstraksi fitur citra fundus retina dan memberikan performa klasifikasi yang optimal. Hasil kajian tersebut menjadi dasar dan justifikasi penggunaan CNN pada penelitian ini, serta memperkuat relevansi metode yang digunakan dengan permasalahan klasifikasi retinopati diabetik.

2. Poster

Selain karya tulis ilmiah, output lain yang dihasilkan adalah poster ilmiah dalam bentuk standing banner.



Gambar 2. 3 Poster Lomba

Gambar 2.3 menunjukkan poster lomba yang berfungsi sebagai media visual untuk merangkum isi penelitian secara ringkas dan informatif, meliputi latar belakang masalah, metode yang digunakan, arsitektur model, hasil eksperimen, serta kesimpulan penelitian. Poster dipresentasikan dan dipamerkan pada tahap final sebagai bagian dari proses penilaian lomba.

2.4.1 Convolutional Neural Network (CNN) dan Transformer

Model pertama yang digunakan yaitu kombinasi antara *Convolutional Neural Network* (CNN) dan *Transformer*. Model CNN + *Transformer* digunakan karena kedua arsitektur tersebut punya kelebihan yang saling melengkapi ketika menganalisis citra fundus. CNN bagus dalam mengenali fitur-fitur kecil seperti mikroaneurisma, perdarahan, atau titik eksudat yang biasanya muncul di retina penderita retinopati diabetik. Namun, CNN kurang kuat dalam melihat hubungan antararea retina yang jaraknya jauh. Di sisi lain, Transformer punya kemampuan memahami pola global melalui mekanisme *self-attention*. Mekanisme ini membantu model “melihat” keseluruhan citra secara lebih menyeluruh, sehingga model bisa mendeteksi pola kerusakan yang lebih besar, misalnya penyebaran perdarahan luas atau

neovaskularisasi yang biasanya ada pada DR tingkat berat.

Menurut penelitian [7], gabungan CNN dan Transformer terbukti memberikan performa yang lebih baik dibandingkan hanya menggunakan CNN saja. CNN akan fokus mengambil detail penting, sementara Transformer menghubungkan informasi tersebut menjadi gambaran utuh sehingga model lebih akurat dalam membedakan tingkat keparahan DR. Penelitian tersebut juga menunjukkan bahwa arsitektur hybrid ini lebih stabil terhadap perbedaan kualitas citra fundus, variasi pasien, dan kondisi pencahayaan. Karena alasan tersebut, penggunaan CNN + *Transformer* dianggap cocok untuk klasifikasi retinopati diabetik yang tingkat keparahannya sering terlihat mirip, tetapi sebenarnya berbeda secara klinis. Pendekatan ini membantu sistem bekerja lebih teliti dan lebih konsisten dalam proses deteksi.

Model CNN + *Transformer* pada penelitian ini dirancang untuk menggabungkan kemampuan CNN dalam mengekstraksi fitur spasial citra dengan kemampuan *transformer* dalam memahami hubungan global antar fitur tersebut. Proses pelatihan dilakukan selama 40 epoch dengan *learning rate* awal sebesar 1×10^{-4} . Hasil evaluasi performa model CNN + *Transformer* ditunjukkan pada Tabel 2.2.

Tabel 2. 2 Hasil Evaluasi Model CNN + Transformer

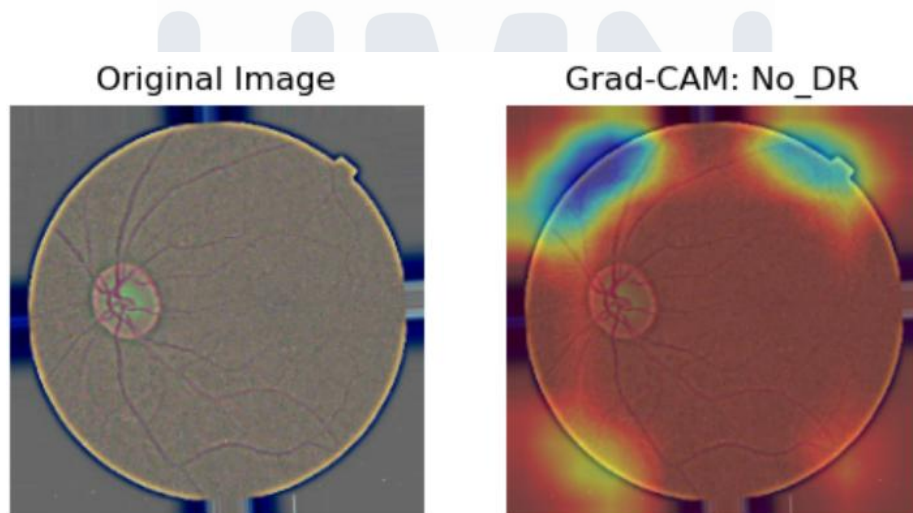
Parameter	Nilai
Epoch terbaik	39
Akurasi pelatihan	86,72%
<i>Loss</i> pelatihan	0,3880
Akurasi validasi	63,21%
<i>Loss</i> validasi	1,1834
<i>Learning rate</i> akhir	1×10^{-4}

Tabel 2.2 menunjukkan bahwa model mencapai epoch terbaik pada epoch ke-39 dengan akurasi pelatihan sebesar 86,72% dan *loss* pelatihan sebesar 0.3880. Sementara itu, akurasi validasi yang diperoleh sebesar 63,21% dengan *loss* validasi sebesar 1,1834. Perbedaan antara akurasi pelatihan dan validasi menunjukkan bahwa model telah mampu mempelajari pola pada data latih, namun masih terdapat tantangan dalam melakukan generalisasi pada data validasi.

Secara umum, hasil model CNN + *Transformer* menunjukkan bahwa integrasi

attention mechanism mampu memperbaiki kemampuan model dalam mengenali hubungan spasial antar fitur. Untuk memahami bagaimana model *Convolutional Neural Network* (CNN) dan *Transformer* melakukan proses klasifikasi pada citra fundus, digunakan teknik visualisasi Grad-CAM (*Gradient-weighted Class Activation Mapping*). Teknik ini membantu menunjukkan area-area pada retina yang menjadi perhatian utama model ketika menentukan label prediksi, seperti No_DR, *Mild*, *Moderate*, *Severe*, atau *Proliferate_DR*.

Pada dasarnya, Grad-CAM menghitung gradien dari kelas yang ingin diprediksi terhadap feature map pada lapisan terakhir CNN atau pada bagian encoder visual pada *Transformer*. Gradien tersebut kemudian dipadukan untuk menghasilkan sebuah *heatmap* yang menyoroti bagian-bagian citra yang paling berkontribusi terhadap keputusan model. Dengan adanya visualisasi ini, mahasiswa dapat memahami bagaimana model membaca pola kerusakan retina dan memastikan bahwa prediksi yang dihasilkan sesuai dengan karakteristik klinis pada tiap tingkat keparahan retinopati diabetik. Gambar berikut memperlihatkan contoh hasil visualisasi Grad-CAM untuk beberapa tingkat keparahan *Diabetic Retinopathy* (DR):



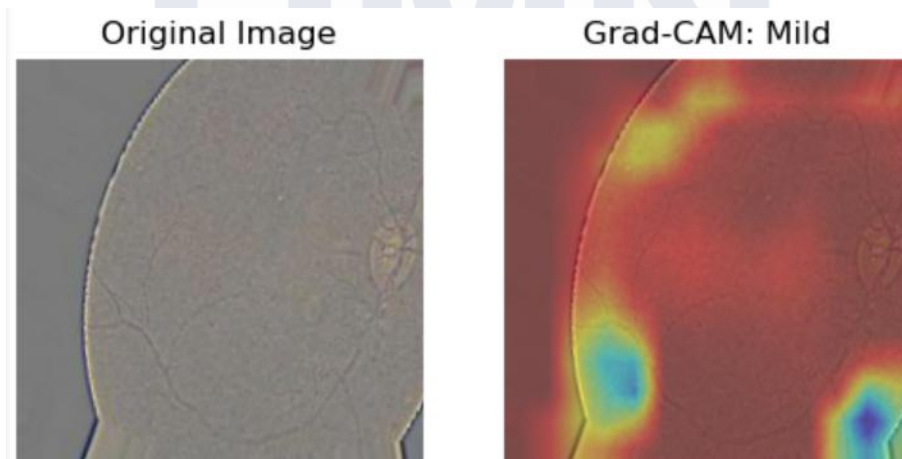
Gambar 2. 4 Grad-CAM No_DR

Gambar 2.4 ditujukan hasil visualisasi Grad-CAM untuk kelas No_DR. Hasil visualisasi Grad-CAM menunjukkan bahwa model CNN + *Transformer* memusatkan perhatian pada area saraf optik (disk optik) serta pola pembuluh darah utama yang masih terlihat normal dan teratur. Fokus model muncul dalam bentuk warna merah hingga kuning tipis di area tertentu, terutama pada bagian atas retina. Warna ini

mengindikasikan bahwa model sedang memvalidasi struktur retina normal sebagai dasar untuk memastikan bahwa citra tidak mengandung tanda-tanda retinopati diabetik.

Pada citra retina sehat, tidak terdapat mikroaneurisma, perdarahan, eksudat keras maupun eksudat lunak. Karena itu, sebagian besar area visualisasi Grad-CAM didominasi warna biru dan hijau, yang menunjukkan bahwa area tersebut tidak dianggap relevan dalam proses klasifikasi [8]. Model hanya memberikan perhatian ringan pada bagian saraf optik dan pembuluh darah utama untuk mengevaluasi pola aliran darah yang wajar dan tidak mengalami penyempitan maupun pelebaran abnormal.

Aktivasi yang rendah ini sejalan dengan karakteristik klinis retina normal, di mana struktur retina tampak bersih, dengan permukaan dan tekstur yang seragam [7]. Model CNN menangkap detail spasial kecil seperti ketebalan pembuluh darah, sedangkan *Transformer* memperhatikan hubungan global antar-area retina untuk memastikan tidak ada pola kerusakan yang tersembunyi. Hasil visualisasi ini memperkuat bahwa model tidak hanya menebak secara acak, tetapi benar-benar mengevaluasi aspek-aspek penting dari retina sehat sebelum memutuskan klasifikasi No_DR.



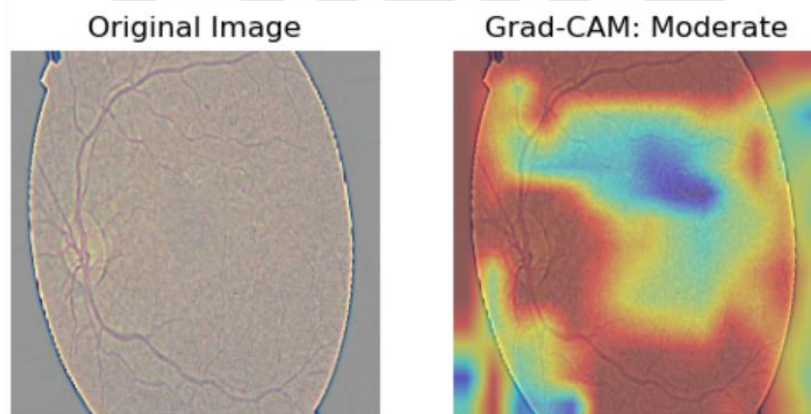
Gambar 2. 5 Grad-CAM Mild

Gambar 2.5 ditujukan hasil visualisasi Grad-CAM untuk kelas *Mild*. Hasil visualisasi Grad-CAM menunjukkan area berwarna merah dan kuning di sekitar pembuluh darah kecil pada retina. Warna merah menandakan tingkat aktivasi paling tinggi, artinya bagian tersebut dianggap sangat penting oleh model dalam menentukan

bahwa gambar termasuk kategori retinopati diabetik ringan. Aktivasi biasanya muncul pada area yang memiliki microaneurysm, yaitu titik-titik kecil yang menjadi tanda paling awal DR. Microaneurysm merupakan pelebaran kecil pada pembuluh darah kapiler yang sering tidak terlihat secara jelas oleh mata manusia, tetapi dapat ditangkap oleh model *deep learning* melalui pola tekstur dan perubahan warna pada citra fundus.

Area berwarna biru dan hijau menunjukkan bagian citra yang tidak terlalu berpengaruh terhadap prediksi model. Distribusi aktivasi ini sesuai dengan ciri klinis DR tahap mild, di mana perubahan retina masih minimal dan hanya terjadi di beberapa titik tertentu. Melalui visualisasi Grad-CAM, dapat terlihat bahwa model mampu mengenali fitur-fitur halus ini dan memusatkan perhatian pada area dengan kemungkinan lesi awal.

Temuan ini juga sejalan dengan penelitian sebelumnya yang menggunakan model CNN dan *Transformer* untuk mendeteksi retinopati diabetik. Studi oleh [7] menunjukkan bahwa microaneurysm merupakan salah satu fitur utama yang di-highlight oleh Grad-CAM pada DR tahap ringan. Selain itu, penelitian [8] menemukan bahwa model berbasis *Vision Transformer* cenderung memusatkan aktivasi pada area pembuluh darah kecil dan titik lesi mikro saat melakukan klasifikasi DR tingkat mild. Hal ini memperkuat bahwa aktivasi yang muncul pada visualisasi milik model sudah sesuai dengan karakteristik medis dari kondisi tersebut.



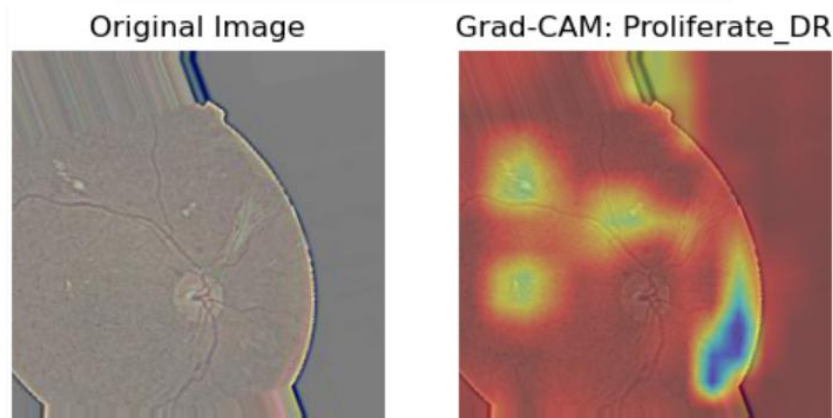
Gambar 2. 6 Grad-CAM Moderate

Gambar 2.6 ditujukan hasil visualisasi Grad-CAM untuk kelas *Moderate*. Visualisasi Grad-CAM memperlihatkan area panas (merah-oranye-kuning) yang terpusat pada bagian tengah retina. Warna-warna tersebut menandakan bahwa model memberikan perhatian tinggi pada area yang mengandung eksudat keras (*hard*

exudates), perdarahan kecil (dot/blot hemorrhages), dan perubahan mikrostruktural retina [9]. Pola ini merupakan ciri khas retinopati diabetik tingkat sedang, di mana kerusakan retina mulai menyebar tetapi belum mencapai tahap proliferaatif.

Aktivasi merah pada heatmap menunjukkan wilayah yang paling menentukan keputusan model. Biasanya area tersebut memiliki kumpulan lesi yang bentuk dan distribusinya lebih kompleks dibandingkan tahap *Mild*. Sementara itu, area biru dan hijau menunjukkan bagian yang tidak signifikan atau tidak digunakan model saat melakukan prediksi.

Distribusi Grad-CAM yang muncul pada stage *Moderate* ini sesuai dengan temuan penelitian lain yang menyebutkan bahwa CNN sangat efektif mengidentifikasi fitur lokal seperti perdarahan dan eksudat kecil. Selain itu, mekanisme attention pada arsitektur modern seperti *Vision Transformer* dapat membantu model melihat pola penyebaran lesi secara lebih menyeluruh [10]. Oleh karena itu, aktivasi heatmap yang muncul menggambarkan pola klinis DR *Moderate* yang juga ditemukan dalam banyak studi terkait citra fundus.



Gambar 2. 7 Grad-CAM Prolifera_DR

Gambar 2.7 ditujukan hasil visualisasi Grad-CAM untuk kelas Prolifera_DR. Hasil visualisasi Grad-CAM menunjukkan aktivasi berwarna merah, oranye, dan kuning yang menyebar hampir ke seluruh bagian retina. Penyebaran aktivasi ini menandakan bahwa model mendeteksi banyak area yang dianggap penting untuk proses klasifikasi, terutama area yang mengindikasikan neovaskularisasi, yaitu pertumbuhan pembuluh darah baru yang abnormal pada retina [11]. Neovaskularisasi merupakan ciri utama dari tahap proliferaatif, serta menjadi penanda bahwa kerusakan

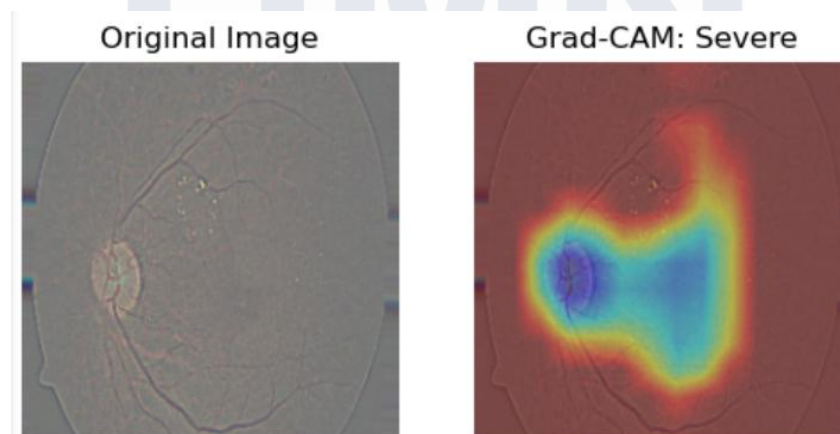
retina telah mencapai tingkat paling parah.

Warna merah pada Grad-CAM menggambarkan bagian retina yang menjadi fokus utama model, biasanya terletak pada:

- pembuluh darah baru yang rapuh dan mudah pecah,
- jaringan fibrosa,
- perdarahan luas,
- serta area abnormal pada diskus optikus.

Karena pembuluh darah baru ini tumbuh secara tidak teratur, model akan menganggap hampir seluruh bagian retina relevan untuk menentukan prediksi, sehingga heatmap tampak lebih luas dibandingkan kategori lain (*Normal*, *Mild*, *Moderate*, atau *Severe*).

Penelitian terdahulu juga menunjukkan bahwa pada retinopati diabetik proliferaatif, pola lesi cenderung tersebar luas, sehingga aktivasi Grad-CAM akan tampak lebih dominan pada seluruh citra retina [12]. Selain itu, model *deep learning* seperti CNN maupun *Vision Transformer* cenderung memberikan respons kuat terhadap pola-pola patologis yang kompleks dan meluas pada DR tingkat lanjut. Oleh karena itu, hasil visualisasi Grad-CAM yang muncul pada kategori Prolifera DR ini sesuai dengan karakter klinis penyakit serta temuan penelitian terkait citra fundus.



Gambar 2. 8 Grad-CAM Severe

Gambar 2.8 ditujukan hasil visualisasi Grad-CAM untuk kelas *Severe*. Visualisasi Grad-CAM menunjukkan aktivasi merah-oranye yang sangat intens pada wilayah retina yang mengalami kerusakan berat. Aktivasi ini muncul terutama pada area yang memiliki perdarahan luas (*blot hemorrhages*), *venous beading*, serta *intraretinal microvascular abnormalities* (IRMA), yang secara klinis merupakan ciri

tahap *Severe* [13].

Temuan visualisasi ini sejalan dengan penelitian [14] dalam yang menjelaskan bahwa model *deep learning* cenderung memberikan perhatian tinggi pada pola kerusakan vaskular yang tidak normal, terutama pada pembuluh darah yang menebal, pecah, atau mengalami penyempitan. Fitur-fitur tersebut merupakan indikator yang sangat kuat dalam menentukan tingkat keparahan retinopati diabetik. Aktivasi berwarna merah menunjukkan wilayah yang paling berpengaruh terhadap prediksi model. Sedangkan area biru atau hijau merupakan bagian retina yang tidak dijadikan acuan signifikan dalam proses klasifikasi. Distribusi pola Grad-CAM seperti ini menunjukkan bahwa model sudah belajar mendeteksi karakteristik klinis DR Severe secara tepat.

2.4.2 Vision Transformer

Vision Transformer (ViT) merupakan pengembangan dari model *transformer* yang mampu mempelajari hubungan jarak jauh pada citra dan terbukti memberikan kinerja yang lebih baik dibandingkan CNN. Namun demikian, ViT memerlukan dataset berukuran besar untuk dapat dilatih secara efektif, sehingga hal ini menjadi kendala dalam penerapannya pada domain DR [15]. Deteksi dini dan pencegahan DR sebagai komplikasi serius dari diabetes yang dapat menyebabkan kebutaan memerlukan metode yang otomatis, akurat, dan mampu menyesuaikan kebutuhan pasien. Berbagai teknik *deep learning* telah diterapkan, dan model *transformer* terbukti efektif dalam pemrosesan bahasa alami serta mampu memahami hubungan jarak jauh pada gambar sehingga dapat menghasilkan kinerja yang lebih baik dibandingkan model CNN [15].

Arsitektur *Vision Transformer* (ViT) merupakan pengembangan dari arsitektur *Transformer* yang saat ini menjadi standar tertinggi (*state-of-the-art*) dalam bidang pemrosesan bahasa alami [16]. *Transformer* pada awalnya digunakan dalam pemrosesan bahasa alami (NLP) dengan memanfaatkan mekanisme *self-attention*. Karena kemampuan representasi datanya yang lebih unggul dibandingkan algoritma *deep learning* lain, termasuk CNN, para peneliti mulai mengadaptasikan *Transformer* ke ranah *computer vision*. Salah satu hasil dari pengembangan tersebut adalah munculnya *Vision Transformer* (ViT) [17]. Pelatihan dilakukan selama 10 epoch menggunakan *learning rate* awal sebesar 1×10^{-4} . Berdasarkan hasil akhir pelatihan,

diperoleh performa sebagai berikut:

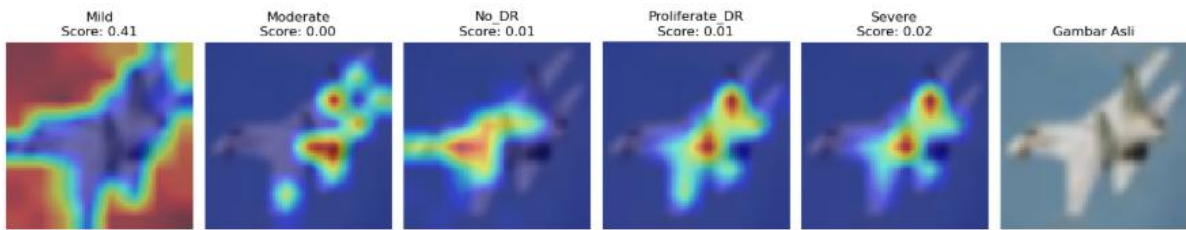
Tabel 2. 3 Hasil Evaluasi Model Vision Transformer

Parameter	Nilai
Epoch terbaik	10
Akurasi pelatihan	94,40%
Loss pelatihan	0.2073
Akurasi validasi	53,00%
Loss validasi	1,9727
Learning rate akhir	1×10^{-4}

Berdasarkan hasil evaluasi yang ditunjukkan pada Tabel 2.3, model *Vision Transformer* mampu mencapai akurasi pelatihan sebesar 94,40%. Nilai ini menunjukkan bahwa model dapat mempelajari pola pada data *training* dengan baik. Namun demikian, akurasi validasi yang hanya mencapai 53,00% mengindikasikan adanya kesenjangan kinerja (gap) antara data pelatihan dan data validasi. Kondisi ini biasanya menandakan *overfitting*, yaitu ketika model terlalu menyesuaikan diri terhadap data pelatihan sehingga kurang mampu melakukan generalisasi pada data baru.

Selain itu, nilai *loss* validasi sebesar 1,9727 yang jauh lebih tinggi dibandingkan *loss* pelatihan memperkuat indikasi bahwa model masih belum optimal dalam mempelajari variasi fitur pada citra fundus, terutama mengingat ukuran dataset yang digunakan relatif kecil. Seperti yang telah dijelaskan sebelumnya, *Vision Transformer* pada dasarnya dirancang untuk bekerja lebih efektif pada dataset berukuran besar. Oleh karena itu, performa yang kurang stabil pada data validasi merupakan konsekuensi umum dari keterbatasan jumlah data.

Meskipun demikian, penggunaan *learning rate* 1×10^{-4} terbukti mampu menjaga proses pelatihan tetap stabil dan mencegah model mengalami penurunan performa secara drastis. Namun, untuk meningkatkan hasil validasi, diperlukan strategi tambahan seperti augmentasi data yang lebih agresif, regularisasi, atau fine-tuning parameter model. Pendekatan-pendekatan tersebut berpotensi meningkatkan kemampuan generalisasi ViT dalam mendeteksi tingkat keparahan retinopati diabetik. Berikut hasil dari visualisasi *Vision Transformer*.



Gambar 2. 9 Grad-CAM Vision Transformer

Gambar 2.9 menampilkan hasil visualisasi Grad-CAM pada model *Vision Transformer* (ViT) untuk setiap kelas retinopati diabetik, yaitu Mild, Moderate, No_DR, Proliferative DR, dan Severe, disertai dengan nilai skor prediksi pada masing-masing kelas. Visualisasi ini menunjukkan area-area pada citra fundus retina yang menjadi fokus perhatian model dalam proses pengambilan keputusan klasifikasi. Berdasarkan Gambar 2.8, terlihat bahwa model memberikan perhatian pada area yang berbeda untuk setiap tingkat keparahan retinopati diabetik. Pada kelas No_DR, aktivasi cenderung rendah dan tersebar, yang mengindikasikan tidak adanya area patologis yang signifikan. Sementara itu, pada kelas Mild dan Moderate, perhatian model mulai terfokus pada area tertentu yang berpotensi mengandung indikasi awal kelainan retinan. Pada kelas Severe dan Proliferate DR, area dengan aktivasi tinggi (ditunjukkan oleh warna merah dan kuning) tampak lebih terpusat, mencerminkan adanya kerusakan retina yang lebih kompleks dan menyeluruh.

Hasil visualisasi Grad-CAM ini menunjukkan bahwa model *Vision Transformer* tidak hanya menghasilkan prediksi kelas, tetapi juga mampu mengidentifikasi pola visual yang relevan sesuai dengan karakteristik masing-masing tingkat keparahan retinopati diabetik. Dengan demikian, visualisasi ini mendukung interpretabilitas model dan memperkuat kepercayaan terhadap hasil klasifikasi yang dihasilkan.