

Segmentasi Semantik *Optic Disk* dan *Optic Cup* pada Citra Fotografi Fundus Retina dengan Transformer

Fachry Dennis Heraldi (13520139)
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jalan Ganesha 10 Bandung
E-mail (gmail): 13520139@std.stei.itb.ac.id

Abstrak— Makalah ini menyajikan metode canggih untuk segmentasi semantik *optic disk* (OD) dan *optic cup* (OC) pada citra fotografi fundus retina, yang penting dalam deteksi dini glaukoma. Makalah ini mengintegrasikan teknik pemrosesan citra dengan fokus khusus pada penggunaan arsitektur Transformer, khususnya model SegFormer. Dataset yang digunakan adalah subset dari Dhristi-GS, yang terdiri dari citra fotografi fundus retina RGB dan citra biner segmentasi manual OD dan OC. Metodologi ini melibatkan pengumpulan data, pelabelan dataset menggunakan platform Segments.ai, rekonstruksi dataset untuk HuggingFace, prapemrosesan data, pelatihan model dengan pretrained SegFormer, dan evaluasi model. Hasil yang dicapai menunjukkan akurasi mendekati 93%, dengan *mean IoU* yang tinggi pada citra uji, menandakan efektivitas metode untuk digunakan dalam deteksi glaukoma.

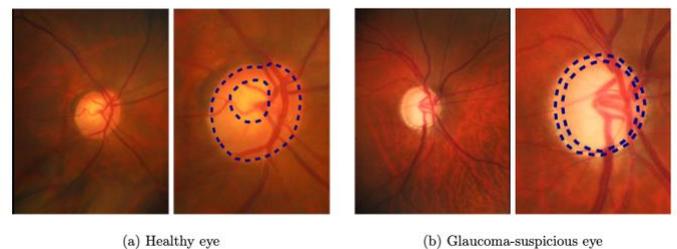
Kata kunci—segmentasi semantik; *optic disk*; *optic cup*; fotografi fundus retina; glaukoma; transformer; SegFormer

I. PENDAHULUAN

Glaukoma adalah penyakit mata kronis yang ditandai dengan peningkatan tekanan di dalam bola mata, yang menyebabkan kerusakan serius pada saraf optik [1]. Glaukoma juga dikenal sebagai penyakit neurodegeneratif mata, yang mengubah kepala saraf optik. Neuropati ini ditandai dengan penurunan progresif sensitivitas penglihatan, yang berpotensi menyebabkan kebutaan. Pada tahap awal, glaukoma tidak menimbulkan gejala atau perubahan pada lapangan visual. Namun, seiring berkembangnya penyakit, terjadi penyempitan lapangan pandang secara bertahap, dimulai dari perifer dan berlanjut ke arah tengah [2]. Jika tidak terdiagnosis, glaukoma dapat menyebabkan kerusakan permanen pada saraf optik yang mengarah pada kebutaan [3].

Deteksi dini glaukoma dapat dilakukan dengan beberapa metode, seperti Optical Coherence Tomography (OCT) dan Heidelberg Retinal Tomography (HRT), namun metode tersebut hanya tersedia di kota besar di Malaysia. Tidak hanya itu, metode tersebut juga melibatkan alat yang canggih dan memerlukan spesialis yang terlatih, serta memakan waktu dan biaya yang besar. Fotografi fundus retina adalah salah satu metode *screening* mata yang paling mungkin dilakukan dengan keterbatasan biaya [1]. Melalui fotografi fundus retina, dengan memperhatikan aspek tertentu pada anatomi mata, dapat dideteksi apakah seseorang berpotensi glaukoma atau tidak.

Pemeriksaan saraf optik termasuk tes fundus retina mata, yang memerlukan dokter untuk menemukan area *optic disk* (cakram optik, OD) dan *optic cup* (bagian tengah *optic disk*, cangkir optik, OC) serta menentukan batas-batasnya. Kehadiran glaukoma dapat diidentifikasi dengan memperhatikan cangkir saraf optik, yaitu peningkatan ukuran *optic cup*. Salah satu indikator utama penyakit ini adalah rasio *cup-to-disc* (CDR) - rasio antara tinggi cangkir dan cakram. Hal ini dianggap sebagai salah satu ciri paling representatif dari area *optic disk* dan *optic cup* untuk deteksi glaukoma, dan menurut literatur, mata dengan CDR setidaknya 0,65 biasanya dianggap mengalami glaukoma dalam praktik klinis. Gambar 1 menunjukkan contoh mata yang sehat dan mata yang dicurigai glaukoma [3]. Gambar di sisi kanan dari contoh menunjukkan area *optic disk* yang diperbesar. Batas *optic disk* ditandai dengan garis putus-putus luar, dan batas *optic cup* ditandai dengan garis putus-putus dalam. Perhatikan bahwa CDR lebih besar untuk mata yang dicurigai glaukoma.



Gambar 1. Contoh (a) mata yang sehat dan (b) mata yang dicurigai glaukoma

CDR adalah parameter kunci dalam menentukan glaukoma melalui segmentasi area *optic cup* dan *optic disk* dari gambar fundus. Namun, deteksi CDR menantang karena tekstur warna yang tidak jelas antara *optic cup* dan *optic disk*, serta rendahnya visibilitas batas antara keduanya dan kepadatan vaskular yang tinggi di wilayah optik. Penelitian sebelumnya telah menggunakan metode pemrosesan gambar seperti *region growing*, deteksi tepi, *thresholding*, dan *operasi morfologi*. Solusi yang lebih akurat melalui pemrosesan gambar dan teknik kecerdasan buatan termasuk Algoritma Genetik dan Jaringan Saraf Buatan (ANN) [1].

Segmentasi *optic disk* dan *optic cup* serta penentuan CDR adalah tugas yang memakan waktu dan saat ini hanya dilakukan oleh profesional [3]. Dalam makalah ini, telah dieksplorasi

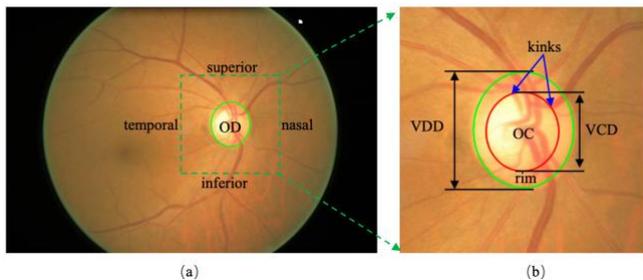
pendekatan algoritma *computer vision* untuk deteksi glaukoma berdasarkan gambar fundus mata, dengan fokus pada segmentasi *optic disk* dan *optic cup*.

Untuk meningkatkan metode ini, makalah ini mengarah pada penggunaan metode segmentasi semantik *state-of-the-art*, yaitu penggunaan model Transformer, yang dapat meningkatkan akurasi dan efisiensi dalam segmentasi *optic disk* dan *optic cup*. Metode ini diharapkan dapat mengurangi waktu yang diperlukan untuk analisis otomatis dan penilaian glaukoma, yang sangat berharga dalam situasi seperti *screening* massal dan perawatan medis di negara-negara dengan kekurangan spesialis kualifikasi. Pendekatan ini juga dapat memberikan solusi yang lebih transparan dan dapat diandalkan bagi dokter.

II. LANDASAN TEORI

A. Optic Disk (OD) dan Optic Cup (OC) pada Citra Fotografi Fundus Retina

Struktur dari *optic disk* dan *optic cup* diperlihatkan pada Gambar 2. *Optic disc* (ditandai dengan lingkaran hijau pada gambar) merupakan lingkaran berwarna kuning terang pada retina yang menunjukkan lokasi pangkal serabut saraf mata atau Optic Nerve Head (ONH). *Optic cup* (ditandai dengan lingkaran merah pada gambar) merupakan sebagian serabut syaraf mata di dalam OD yang tampak lebih terang (hampir putih) karena secara topografi lebih menonjol, akibat adanya tekanan internal bola mata (intraokular). Pada penderita penyakit glaukoma, tekanan intraokular akan meningkat sehingga OC meluas dan mengakibatkan peningkatan rasio dimensi OC terhadap OD. Peningkatan rasio ini dapat dipergunakan sebagai indikator glaukoma [4].



Gambar 2. Struktur *optic disk* dan *optic cup* pada citra fotografi fundus retina

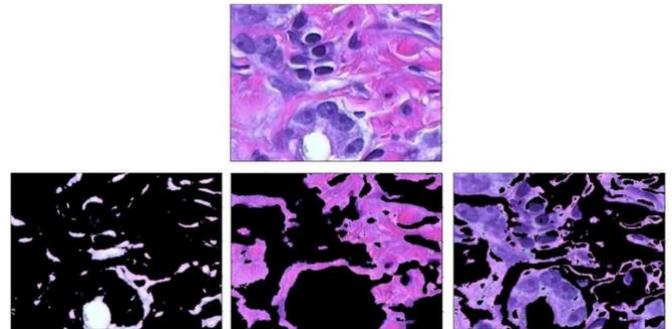
B. Segmentasi Citra

Segmentasi citra adalah proses membagi citra menjadi berbagai region atau objek, dengan setiap region terdiri dari kumpulan piksel yang terhubung. Tujuan utama dari segmentasi citra adalah untuk membedakan antara objek dan latar belakang dalam sebuah gambar. Dengan cara ini, kita dapat lebih fokus pada pengolahan segmen-segmen tertentu yang penting dalam citra, tanpa harus memproses keseluruhan gambar [5].

Tujuan utama dari segmentasi citra adalah untuk menemukan bagian-bagian citra yang koheren atau objek spesifik. Segmentasi dilakukan berdasarkan berbagai properti yang dipilih, seperti kecerahan, warna, tekstur, dan lain-lain. Proses ini membagi citra menjadi beberapa segmen yang

terhubung, di mana setiap segmen bersifat homogen berdasarkan properti yang telah ditentukan [5].

Segmentasi citra merupakan langkah penting yang mendahului proses pengenalan citra atau objek, pemahaman citra, dan sebagainya. Salah satu contoh aplikasi dari segmentasi citra adalah dalam bidang pencitraan medis, dapat dilihat pada Gambar 3. Misalnya, dalam proses diagnosis kanker, para ahli patologi biasanya menginjeksi jaringan tubuh dengan hematoxylin dan eosin (H&E) untuk membedakan jenis jaringan. Mereka kemudian menggunakan teknik segmentasi citra, seperti *clustering*, untuk mengidentifikasi jenis jaringan tersebut dalam gambar. Segmentasi memainkan peran penting dalam proses ini untuk mempermudah identifikasi dan analisis lebih lanjut [5].

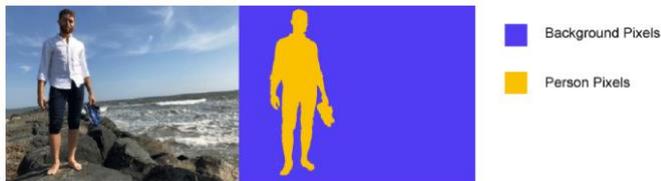


Gambar 3. Menggunakan *clustering* untuk membedakan jenis jaringan [5].

C. Segmentasi Semantik

Segmentasi semantik merupakan sebuah algoritma dari pembelajaran mendalam yang bertujuan untuk menetapkan label atau kategori pada setiap piksel dalam suatu gambar. Algoritma ini penting dalam mengidentifikasi dan mengelompokkan piksel-piksel yang termasuk dalam kategori yang berbeda-beda. Contohnya, dalam kendaraan otonom, algoritma ini diperlukan untuk mengenali berbagai elemen seperti kendaraan lain, pejalan kaki, rambu lalu lintas, trotoar, dan berbagai aspek lainnya dari jalan. Selain itu, segmentasi semantik juga memegang peran penting dalam berbagai aplikasi lain termasuk pengemudian otomatis, pencitraan medis, dan inspeksi industri. Kemampuannya untuk mengkategorikan dan membedakan komponen-komponen penting dalam gambar menjadi sangat berguna [5].

Sebuah ilustrasi dasar dari segmentasi semantik dapat dilihat dalam membagi gambar menjadi dua kategori berbeda. Sebagai contoh, dalam Gambar 4, terdapat gambar seseorang di pantai yang diiringi dengan versi lain dimana piksel-piksel dalam gambar tersebut telah dipecah menjadi dua kelas yang berbeda, yaitu orang tersebut dan latar belakangnya. Karena segmentasi semantik memberikan label pada setiap piksel dalam sebuah gambar, dianggap lebih akurat dibandingkan metode deteksi objek lainnya. Keakuratan ini menjadikan segmentasi semantik sangat bermanfaat dalam berbagai aplikasi industri yang membutuhkan pemetaan gambar yang detail dan tepat. Contoh penggunaannya termasuk di bidang pencitraan medis, teknik ini digunakan untuk menganalisis dan mendeteksi anomali seperti kanker dalam sel [5].



Gambar 4. Hasil segmentasi semantik untuk memisahkan objek orang dan latar belakang

Proses pelatihan jaringan dalam segmentasi semantik dimulai dengan menganalisis sekelompok gambar yang telah dilabeli pada tingkat piksel. Langkah pertama ini sangat penting karena memberikan data dasar yang akan digunakan oleh jaringan untuk belajar dan mengenali pola. Setelah kumpulan gambar berlabel ini siap, langkah selanjutnya adalah pembuatan jaringan segmentasi semantik itu sendiri. Jaringan ini dirancang khusus untuk dapat mengklasifikasikan setiap piksel dalam gambar ke dalam kategori yang relevan, berdasarkan label yang telah ditentukan sebelumnya [5].

Setelah jaringan segmentasi semantik terbentuk, proses pelatihan dimulai. Dalam tahap ini, jaringan diajarkan untuk mengklasifikasikan gambar ke dalam kategori piksel yang berbeda. Pelatihan ini melibatkan penyesuaian berulang-ulang berdasarkan data gambar berlabel, dengan tujuan agar jaringan dapat akurat mengidentifikasi dan mengklasifikasikan piksel dalam gambar baru yang belum pernah dilihat sebelumnya [5].

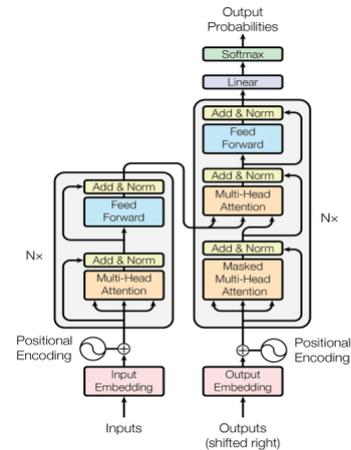
Langkah terakhir dalam proses ini adalah menilai keakuratan dari jaringan yang telah dilatih. Penilaian ini sangat penting untuk memastikan bahwa jaringan mampu melakukan segmentasi dengan tepat dan efisien. Keakuratan jaringan ini diukur dengan menguji jaringan pada set data baru yang belum pernah diolah sebelumnya. Dengan mengikuti langkah-langkah ini, jaringan segmentasi semantik dapat efektif digunakan untuk mengklasifikasikan gambar secara akurat ke dalam kategori piksel yang sesuai [5].

D. Transformer

Transformer adalah model berbasis *deep learning* yang pertama kali diperkenalkan oleh Vaswani dkk. pada tahun 2017. Awalnya, Transformer dirancang untuk tugas pemrosesan bahasa alami, tetapi sejak itu telah berhasil diterapkan dalam berbagai bidang pemrosesan data, termasuk dalam pemrosesan citra. Arsitektur Transformer berbeda dari arsitektur jaringan saraf konvensional seperti Convolutional Neural Networks (CNN) atau Recurrent Neural Networks (RNN) karena tidak melibatkan operasi konvolusi atau urutan tampilan. Arsitektur Transformer dapat dilihat pada Gambar 5. Transformer mengandalkan dua komponen utama: *self-attention* dan lapisan *feed forward*.

Self-attention memungkinkan model untuk memahami hubungan antara setiap elemen dalam urutan data masukan. Misalnya, dalam konteks pemrosesan bahasa, Transformer dapat mengidentifikasi hubungan antara kata-kata dalam sebuah kalimat. Transformer bekerja dengan menghitung bobot untuk setiap pasangan elemen dalam urutan, menunjukkan seberapa penting hubungan antara mereka. Hasilnya adalah matriks bobot yang digunakan untuk menggabungkan informasi dari semua

elemen dalam urutan. Setelah perhitungan *Self-attention*, hasilnya melewati lapisan *feed forward*, yang terdiri dari beberapa perceptron. Lapisan ini mengubah representasi data masukan menjadi representasi yang lebih tingkat dan lebih abstrak.



Gambar 5. Arsitektur Transformer [6].

Pemrosesan paralel menjadi keunggulan utama Transformer, mengatasi batasan pada RNN yang tidak dapat melakukan hal tersebut. Dengan arsitektur yang terdiri dari 6 layer pada *encoder* dan *decoder*, Transformer memanfaatkan *multi-head attention* untuk memproses berbagai aspek dari urutan input secara independen. Melalui *positional encoding* dan *embedding*, Transformer mempertahankan informasi urutan dalam teks, memungkinkan representasi yang lebih dinamis dan kontekstual.

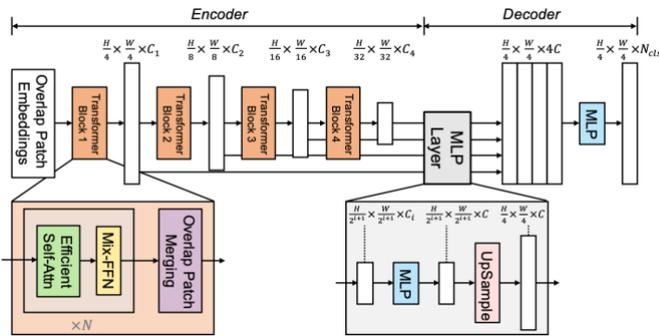
Jika pada awalnya Transformer ditujukan untuk pemrosesan Teks, terdapat varian dari Transformer agar dapat memroses citra, yaitu Vision Transformer (ViT). Vision Transformer membagi citra menjadi potongan-potongan kecil (misalnya, *patch*) dan memperlakukan setiap *patch* sebagai elemen urutan. Dengan menggunakan *self-attention*, Vision Transformer dapat memahami hubungan antara *patch* untuk menghasilkan representasi citra yang lebih baik [7].

E. Segmentasi Semantik dengan Transformer

SegFormer adalah metode segmentasi semantik canggih, diajukan oleh Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, Ping Luo pada tahun 2021. SegFormer memiliki fitur-fitur kunci yang membuatnya unggul. Diantaranya adalah penggunaan *encoder* Transformer hierarkis yang menghasilkan fitur multi-skala tanpa perlu *positional encoding*, menghindari penurunan kinerja saat resolusi citra uji berbeda dari pelatihan. Selain itu, SegFormer menggunakan *decoder* MLP yang ringan untuk mengagregasi informasi dari berbagai lapisan, yang memadukan *attention* lokal dan global untuk menghasilkan representasi yang kuat. Decoder ini dianggap sebagai kunci efisiensi segmentasi pada Transformer.

Cara kerja dari SegFormer terdiri atas dua modul utama, yakni *encoder* Transformer hierarkis yang menghasilkan fitur kasar resolusi tinggi dan fitur halus resolusi rendah, serta *decoder* MLP yang ringan untuk menggabungkan fitur multi-

level ini dan menghasilkan *mask* segmentasi semantik akhir. Kerangka kerja dari SegFormer dapat dilihat pada Gambar 6. Dengan pendekatan desain yang sederhana namun efektif, SegFormer berhasil menghindari desain kompleks yang seringkali ada pada metode sebelumnya, menghasilkan efisiensi serta kinerja yang tinggi. Selain mencapai hasil terbaru yang terbaik pada dataset umum, SegFormer juga menunjukkan kinerja yang baik dengan pendekatan *zero-shot*, memperlihatkan SegFormer sebagai terobosan penting dalam dunia segmentasi semantik.

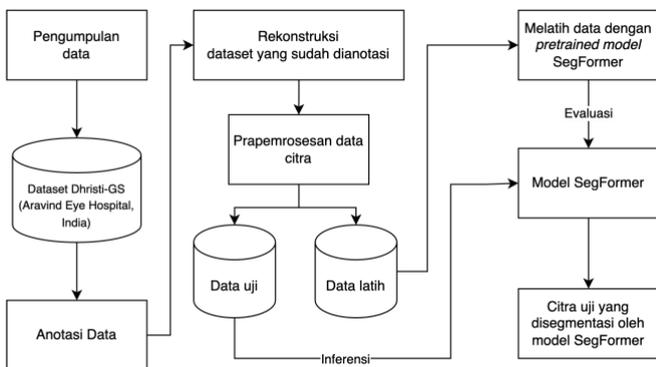


Gambar 6. Kerangka kerja SegFormer yang terdiri atas modul *encoder* dan *decoder*

SegFormer telah berhasil menetapkan standar baru dalam hal kinerja dan efisiensi. Misalnya, model SegFormer-B4 mencapai 50,3% mIoU di ADE20K dengan 64 juta parameter, yang 5 kali lebih kecil dan 2,2% lebih baik daripada metode terbaik sebelumnya. Model terbaik, SegFormer-B5, mencapai 84,0% mIoU pada set validasi Cityscapes dan menunjukkan kinerja pendekatan *zero-shot* yang luar biasa pada Cityscapes-C.

III. METODE PENYELESAIAN MASALAH

Untuk melakukan segmentasi semantik *optic disk* dan *optic cup* pada citra fotografi fundus retina, dirancang langkah penyelesaian yang diilustrasikan oleh skema pada Gambar 7.



Gambar 7. Skema rancangan penyelesaian masalah segmentasi semantik

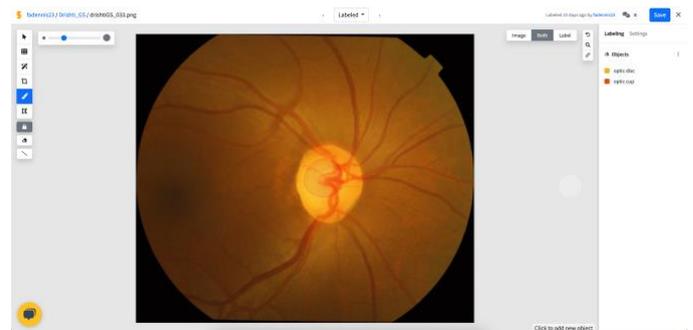
Proses pada setiap tahapan penyelesaian masalah akan dirincikan lebih lanjut pada tiap subbab pada bab ini.

A. Pengumpulan Data

Dataset yang digunakan pada makalah ini adalah subset data Dhristi-GS dari Aravind Eye Hospital, India. Terdapat 30 pasang citra fotografi fundus retina RGB dan citra biner segmentasi manual OD dan OC oleh dokter. Tiga puluh pasang citra dibagi menjadi 20 pasang data latih dan 10 pasang data uji. Data latih digunakan untuk pengembangan model, sedangkan data uji digunakan untuk menguji kinerja model [4].

B. Pelabelan Dataset

Untuk mendapatkan label segmentasi, perlu dilakukan penandaan kelas dari region yang ada pada citra. Proses pelabelan data dilakukan menggunakan alat dari platform Segments.ai. Platform ini dapat memudahkan proses segmentasi citra dan dapat diekstrak hasilnya menjadi HuggingFace Datasets. Gambar 8 menunjukkan halaman platform Segments.ai untuk melakukan pelabelan data.



Gambar 8. Penggunaan platform Segments.ai dalam pelabelan data

Proses pelabelan diawali dengan penentuan kelas dari region yang akan dilabeli. Dua kelas telah ditetapkan: *optical disc* dan *optical cup*. Pelabelan ini dilaksanakan secara manual, dengan mengacu pada label yang sudah diberikan oleh dokter sebagai referensi. Untuk melakukan pelabelan, region pada citra ditandai sesuai dengan kelasnya. Hal ini dilakukan menggunakan alat 'brush' pada Segments.ai, yang memungkinkan pelabelan yang tepat dan akurat pada setiap region yang diinginkan.

Setelah proses pelabelan selesai, terbentuk koleksi data yang berikutnya dapat dikonstruksi menjadi dataset yang lebih sesuai untuk digunakan dalam proses pembangunan model.

C. Rekonstruksi Dataset

Data yang telah dilabeli menggunakan platform Segments.ai kemudian diolah dan direkonstruksi untuk disesuaikan dengan platform HuggingFace. Setelah proses rekonstruksi, dataset ini akan mencakup representasi citra yang diubah menjadi matriks piksel. Dalam matriks ini, setiap piksel pada citra diberi label sesuai dengan kategorinya: `0` mewakili *background*, `1` untuk *optical disk*, dan `2` untuk *optical cup*. Pendekatan ini memungkinkan identifikasi yang lebih mudah dan presisi dalam proses segmentasi semantik, dengan memberikan informasi label pada setiap elemen dalam citra tersebut.

D. Prapemrosesan Data

Model SegFormer mengharapkan input memiliki bentuk tertentu sehingga data latih perlu ditransformasi menggunakan

SegformerImageProcessor tagar sesuai dengan bentuk yang diharapkan. Proses transformasi dilakukan secara *on-the-fly* di dalam proses pelatihan dengan memanfaatkan fungsi ColorJitter dari torchvision untuk secara acak mengubah kecerahan, kontras, saturasi, dan *hue* dari citra.

E. Proses Pelatihan Data dengan Pretrained Model SegFormer

Proses pelatihan data dilakukan menggunakan model yang telah dilatih sebelumnya (*pretrained model*). Model spesifik yang digunakan dikembangkan oleh Xu Sun [8], dan telah menjalani proses *fine-tuning* untuk segmentasi semantik dengan menggunakan dataset dari REFUGE Challenge. Dataset ini khusus berisi gambar fotografi fundus retina, yang menjadi salah satu domain dari makalah ini. Penggunaan model yang telah terlatih dengan dataset khusus ini diharapkan dapat meningkatkan efisiensi dan akurasi dalam pensegmentasian *optic disk* dan *optic cup*.

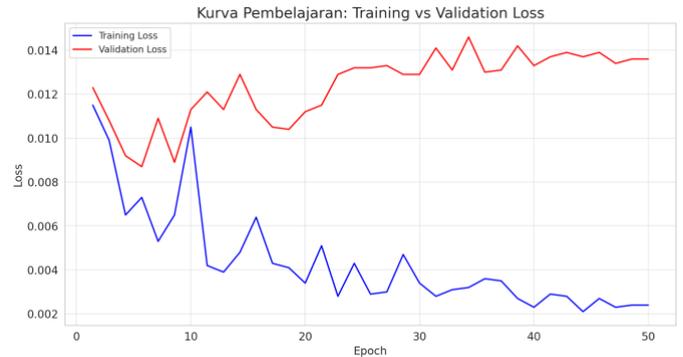
Selain itu, pelatihan model dilakukan dengan memanfaatkan pustaka Transformer yang disediakan oleh HuggingFace. Dalam proses pelatihan model, berbagai *hyperparameter* ditetapkan untuk memastikan efisiensi dan efektivitas pelatihan. *Learning rate* ditetapkan sebesar 0.00006, memungkinkan model untuk belajar secara bertahap tanpa melewatkan fitur penting dalam data. Ukuran batch untuk pelatihan dan evaluasi keduanya ditetapkan dengan nilai 2, memastikan data diproses dalam kelompok kecil, yang memungkinkan penggunaan memori yang efisien dan stabilitas dalam proses pelatihan. *Seed* ditetapkan pada nilai 42 untuk memastikan reproduktivitas eksperimen.

Optimizer yang dipilih adalah Adam, dengan efisiensi yang diketahui dalam pembaruan bobot jaringan. Parameter *betas* untuk optimizer ini ditetapkan pada (0.9, 0.999), dan *epsilon* sebesar 0.00000001, memberikan keseimbangan yang baik antara momentum dan stabilisasi dalam proses optimisasi. Jenis penjadwalan *learning rate* yang dipilih adalah linear, yang membantu dalam penyesuaian nilai learning rate secara bertahap sepanjang proses pelatihan. Proses pelatihan dilakukan selama 50 *epoch*, memberikan waktu yang cukup bagi model untuk belajar secara menyeluruh dari data tanpa *overfitting*.

Disamping menentukan *hyperparameter*, didefinisikan pula metrik evaluasi yang digunakan untuk menilai kinerja model. Metrik utama yang difokuskan adalah akurasi dan mean Intersection over Union (mIoU). Kedua metrik ini adalah standar dalam mengevaluasi model segmentasi. Metrik akurasi akan mengukur seberapa tepat model memprediksi kelas pada tiap piksel citra yang telah tersegmentasi, sementara itu metrik mIoU mengukur tingkat tumpang tindih antara area yang diprediksi oleh model dan area sebenarnya (*ground truth*). Untuk setiap prediksi, IoU dihitung dengan membagi area irisan (*intersection*) antara prediksi dan *ground truth* dengan area gabungan (*union*) dari keduanya. mIoU adalah rata-rata dari nilai-nilai IoU ini untuk semua contoh dalam dataset. Nilai yang lebih tinggi menunjukkan performa yang lebih baik, dengan nilai sempurna 1 menandakan prediksi yang sepenuhnya akurat dan cocok dengan *ground truth*. Metrik mIoU sangat penting dalam segmentasi semantik karena menyediakan ukuran kuantitatif yang objektif tentang seberapa baik model mengenali dan meng-*outline* area spesifik dalam gambar.

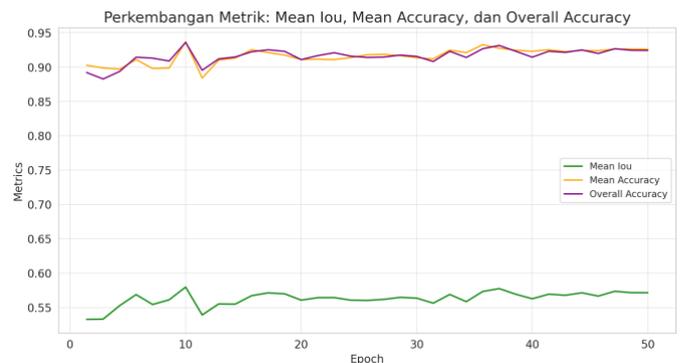
F. Evaluasi Model SegFormer

Saat melakukan proses pelatihan, model akan dievaluasi kinerjanya setiap menyelesaikan satu *epoch* secara iteratif hingga 50 *epoch*.



Gambar 9. Kurva *training loss* dan *validation loss* terhadap *epoch*

Gambar 9 menunjukkan kurva perbandingan antara *training loss* dan *validation loss* sepanjang berjalannya *epoch*. *Loss* adalah metrik yang mengukur seberapa baik model memprediksi label yang benar. *Training loss* menurun seiring dengan meningkatnya *epoch*, menunjukkan bahwa model semakin baik dalam mempelajari data latih. Sebaliknya, *validation loss* memberikan indikasi tentang bagaimana model berperforma pada data yang tidak dilihat selama pelatihan. Stabilitas atau penurunan *validation loss* seiring waktu menandakan model tidak hanya mempelajari data latih dengan baik tetapi juga mampu menggeneralisasi pembelajarannya ke data baru, suatu aspek penting dalam menghindari *overfitting*.

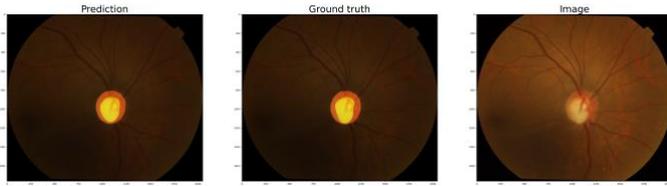


Gambar 10. Perkembangan metrik *mean IoU*, *mean accuracy*, dan *overall accuracy*

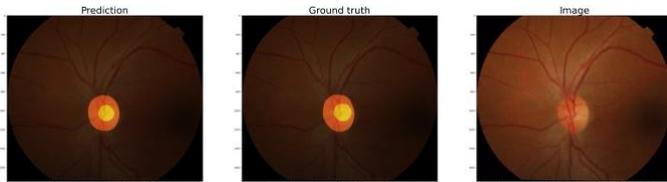
Sementara itu, Gambar 10 menampilkan bagaimana *mean IoU*, *mean accuracy*, dan *overall accuracy* berkembang sepanjang pelatihan. Peningkatan *mean IoU* menunjukkan model semakin akurat dalam memetakan batas-batas *optic disk* dan *optical cup*. *Mean accuracy* dan *overall accuracy* mengukur keseluruhan keakuratan prediksi model, dengan *mean accuracy* memberikan rata-rata akurasi untuk setiap kelas dan *overall accuracy* untuk keseluruhan gambar. Kenaikan dalam metrik-metrik ini menunjukkan peningkatan kinerja model dalam mengklasifikasikan setiap piksel secara benar. Disimpulkan pada akhir pelatihan, model memberikan akurasi mendekati 93%.

G. Inferensi Model dan Hasil Pengujian Citra

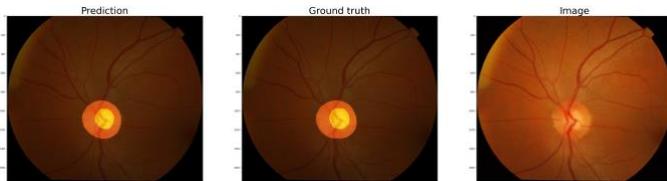
Setelah model berhasil dibangun, model siap digunakan untuk inferensi untuk segmentasi semantik. Berikut adalah tiga contoh citra uji beserta *mean IoU* dan *mean accuracy*-nya.



Gambar 11. Hasil prediksi segmentasi semantik pada citra uji pertama dengan perolehan *mean IoU* sebesar 90.32% dan *mean accuracy* sebesar 99.83%



Gambar 12. Hasil prediksi segmentasi semantik pada citra uji kedua dengan perolehan *mean IoU* sebesar 86.55% dan *mean accuracy* sebesar 99.77%



Gambar 13. Hasil prediksi segmentasi semantik pada citra uji ketiga dengan perolehan *mean IoU* sebesar 92.35% dan *mean accuracy* sebesar 99.77%

IV. KESIMPULAN

Telah berhasil dikembangkan metode segmentasi semantik yang efektif untuk identifikasi *optic disk* dan *optic cup* pada citra fotografi fundus retina. Penggunaan arsitektur Transformer, khususnya model SegFormer, menghasilkan peningkatan signifikan dalam akurasi dan efisiensi segmentasi. Proses ini meliputi langkah-langkah dari pengumpulan dan pelabelan data, rekonstruksi dataset, prapemrosesan, hingga pelatihan dan evaluasi model. Model yang dihasilkan mencapai akurasi hingga 93% dengan *mean IoU* di atas 86% pada citra uji, menunjukkan potensi besar dalam aplikasi klinis untuk deteksi glaukoma. Hasil dari segmentasi diharapkan dapat membantu dokter dalam memperkirakan potensi glaukoma dengan membandingkan rasio *cup-to-disc* (CDR).

VIDEO LINK YOUTUBE DAN NOTEBOOK

Untuk memberikan gambaran tentang persoalan yang diangkat dan solusinya, penjelasan dalam bentuk video dapat dilihat dengan mengakses link: <https://youtu.be/13oCx4Co9p4>. Sementara itu, untuk melihat pengembangan kode dan algoritma, dapat mengakses Google Colab Notebook pada link: <https://colab.research.google.com/drive/1tXVKxkh9VhCsfOjdPvBkJr5sN7766pW-?usp=sharing>

UCAPAN TERIMA KASIH

Penulis mengungkapkan rasa syukur kepada Tuhan Yang Maha Esa yang mengizinkan penulis menyelesaikan makalah tentang segmentasi semantik *optic disk* dan *optic cup* dalam citra fotografi fundus retina. Penghargaan juga penulis tujukan kepada orang tua penulis atas dukungan tak terhingga mereka, dan kepada Bapak Dr. Ir. Rinaldi Munir, M.T., yang telah memberikan bimbingan dan ilmu yang berharga dalam bidang interpretasi dan pengolahan citra. Tak lupa juga penulis mengucapkan terima kasih kepada teman-teman penulis yang tidak dapat penulis sebutkan satu persatu yang telah mendukung dan membantu penulis dalam mengembangkan ide untuk makalah ini. Penulis berharap makalah ini dapat memberikan kontribusi yang bermanfaat dalam penelitian medis dalam pendeteksian dini glaukoma.

REFERENSI

- [1] N. E. A. Khalid, N. M. Noor, dan N. M. Ariff, "Fuzzy c-Means (FCM) for optic cup and disc segmentation with morphological operation," dalam *Procedia Computer Science*, Elsevier B.V., 2014, hlm. 255–262. doi: 10.1016/j.procs.2014.11.060.
- [2] A. Mvoulana, R. Kachouri, dan M. Akil, "Fully automated method for glaucoma screening using robust optic nerve head detection and unsupervised segmentation based cup-to-disc ratio computation in retinal fundus images," *Computerized Medical Imaging and Graphics*, vol. 77, 2019, doi: 10.1016/j.compmedimag.2019.101643i.
- [3] A. Sevastopolsky, "Optic Disc and Cup Segmentation Methods for Glaucoma Detection with Modification of U-Net Convolutional Neural Network," Apr 2017, doi: 10.1134/S1054661817030269.
- [4] "Segmentasi Optic Disc & Optic Cup dari Citra Fotografi Fundus, Tugas Besar EB3206 Pengolahan Citra Biomedika Semester II/2022-2023," 2023.
- [5] R. Munir, "Materi Segmentasi Citra Bagian 2 IF4073 Interpretasi dan Pengolahan Citra," 2023.
- [6] A. Vaswani *dkk.*, "Attention Is All You Need," Jun 2017.
- [7] A. Dosovitskiy *dkk.*, "AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE." [Daring]. Tersedia pada: <https://github.com/>
- [8] X. Sun, Y. Xu, W. Zhao, T. You, dan J. Liu, "Optic Disc Segmentation from Retinal Fundus Images via Deep Object Detection Networks."

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 16 Desember 2023

Fachry Dennis Herald
13520139