

**IMPLEMENTASI ARSITEKTUR *UNET-TRANSFORMER*
DALAM SEGMENTASI SEMANTIK CITRA PEMBULUH
DARAH ARTERI DAN VENA RETINA**

DRAFT SKRIPSI

Oleh:

GIOVILLANDO

NIM 08011182126006



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SRIWIJAYA**

2025

LEMBAR PENGESAHAN

**IMPLEMENTASI KOMBINASI ARSITEKTUR UNET-TRANSFORMER
DALAM SEGMENTASI SEMANTIK PEMBULUH DARA ARTERI-VENA
PADA RETINA**

SKRIPSI

**Sebagai Salah Satu Syarat untuk Memperoleh
Gelar Sarjana Matematika**

Oleh

**GIOVILLANDO
NIM.08011182126006**

Pembimbing Kedua



**Irmeilyana, S.Si., M.Si
NIP. 197405171999032003**

**Indralaya, 19 Maret 2025
Pembimbing Utama**



**Dr. Anita Desiani, S.Si., M.Kom
NIP. 197712112003122002**

**Mengetahui,
Ketua Jurusan Matematika**



**Dr. Dian Cahyawati S. S.Si., M.Si
NIP. 197303212000122001**

PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : Giovillando

NIM : 08011182126006

Fakultas/Jurusan : Matematika dan Ilmu Pengetahuan
Alam/Matematika

Menyatakan bahwa skripsi ini adalah hasil karya ilmiah saya sendiri dan karya ilmiah ini belum pernah diajukan sebagai pemenuhan persyaratan untuk memperoleh gelar kesarjanaan strata satu (S1) dari Universitas Sriwijaya maupun perguruan tinggi lain.

Semua informasi yang dimuat di dalam skripsi ini yang berasal dari penulis lain baik yang dipublikasi atau tidak telah diberikan penghargaan dengan mengutip nama sumber penulis secara benar. Semua isi dari skripsi ini sepenuhnya menjadi tanggung jawab saya sebagai penulis.

Demikianlah surat pernyataan ini saya buat dengan sebenarnya.

Indralaya, 20 Maret 2025



Giovillando

NIM. 08011182126006

HALAMAN PERSEMBAHAN

Kupersembahkan skripsi ini untuk:

Tuhan yang Maha Esa,

Orang tuaku yang sangat kucinta,

Kakak-adikku tersayang,

Keluarga Besarku,

Semua Dosenku,

Teman baik yang Kubanggakan,

Almamaterku

Motto

“Time Is Money”

KATA PENGANTAR

Puji syukur atas kehadiran Tuhan Yang Maha Esa yang telah memberikan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi berjudul “Implementasi Arsitektur *U-Net-Transformer* dalam Segmentasi Semantik Pembuluh Darah Arteri-Vena Retina”. Skripsi ini ditulis sebagai salah satu syarat memperoleh gelar sarjana sains studi Matematika di Fakultas MIPA Universitas Sriwijaya.

Penulis menyadari bahwa skripsi ini tidak mungkin terselesaikan tanpa adanya semangat, dukungan, bantuan, bimbingan dan nasihat yang diberikan berbagai pihak selama proses penyusunan ini berlangsung. Penulis mengucapkan terima kasih yang sebesar-besarnya kepada orang tua tercinta, Ibuku **Lu Ling** dan Ayahku **Tjen Soen On** yang tidak pernah berhenti berjuang dan memberikan yang terbaik untukku sebagai putranya. Terima kasih karena tak pernah lelah mendidik, menasehati, membimbing, mendukung dan terus mendo'akan. Penulis juga ingin menyampaikan ucapan terima kasih kepada sebesar-besarnya dan penghargaan setinggi-tingginya kepada:

1. Bapak **Prof. Hermansyah, S.Si., M.Si., Ph.D** selaku Dekan Fakultas MIPA Universitas Sriwijaya. Ibu **Dr. Dian Cahyawati Sukanda, S.Si., M.Si.** selaku Ketua Jurusan Matematika dan Ibu **Des Alwine Zayanti, S.Si., M.Si.** selaku Sekretaris Jurusan Matematika yang telah membimbing dan mengarahkan dalam urusan akademik selama menempuh perkuliahan di Jurusan Matematika FMIPA Universitas Sriwijaya.
2. Ibu **Dr. Anita Desiani, S.Si., M.Kom.** selaku Dosen Pembimbing Pertama

dan Ibu **Irmeilyana, S.Si., M.Si.** selaku Dosen Pembimbing Kedua yang telah bersedia meluangkan waktu, tenaga, dan pikiran untuk memberikan bimbingan, arahan dan didikan yang berharga selama pembuatan skripsi, perlombaan dan proses perkuliahan. Bapak **Dr. Bambang Suprihatin, S.Si, M.Si.** dan Ibu **Novi Rustiana Dewi, S.Si., M.Si.** selaku dosen pembahas, telah memberikan respons, kritik, dan saran yang sangat berguna untuk perbaikan dan penyelesaian skripsi ini.

3. Bapak **Drs. Endro Setyo Cahyono, M.Si.** selaku dosen pembimbing akademik dan **Seluruh Dosen di Jurusan Matematika FMIPA** yang telah memberikan ilmu, nasihat, motivasi, serta bimbingan selama proses perkuliahan. Bapak **Irwansyah** dan Ibu **Hamidah** selaku staf administrasi di Jurusan Matematika FMIPA Universitas Sriwijaya yang telah membantu penulis selama perkuliahan.
4. Semua sahabat seperjuangan **Komputasi 2021** selama masa perkuliahan dan proses skripsi. **Kakak-kakak tingkat angkatan 2020 dan 2019** yang telah membantu dan membagikan ilmunya kepada penulis, serta **adik-adik tingkat** yang selama masa perkuliahan dan proses skripsi telah memberikan kesan yang luar biasa.
5. Kakak dan adikku tersayang, **Sintia Crist Tjen, Aneria** dan **Giovil Wijaya** yang senantiasa memberikan semangat dan doa terbaik untuk penulis.
6. Semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah memberikan bantuan dalam menyelesaikan skripsi ini hanya ucapan terima kasih yang dapat penulis berikan.

Semoga skripsi ini dapat menambah pengetahuan dan bermanfaat bagi mahasiswa/i Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya dan semua pihak yang membutuhkan.

Indralaya, 06 Desember 2024

Penulis

IMPLEMENTATION OF UNET-TRANSFORMER ARCHITECTURE IN SEMANTIC SEGMENTATION OF RETINAL ARTERY-VEIN BLOOD VESSEL IMAGES

By:

GIOVILLANDO

08011182126006

ABSTRACT

Retinal blood vessels are crucial components of the eye's circulatory system, functioning to supply oxygen and nutrients while removing waste from retinal tissues. Retinal blood vessels are divided into two types: arteries and veins (A/V). Arteries and veins are often located close to each other, necessitating clear separation to assist medical professionals in identifying specific diseases and preventing errors in analyzing related conditions. The separation of arteries and veins can be performed using image segmentation-based technology. This study aims to conduct semantic segmentation of retinal blood vessels by combining the U-Net architecture, Vision Transformer (ViT), and Attention Gate. The proposed model employs ViT as an encoder to capture global spatial relationships, while U-Net acts as a decoder to restore image spatial details. An Attention Gate is integrated to filter relevant information from generated features. Segmentation performance was evaluated across five label classes: Background, Artery, Crossings, Vein, and Uncertain, using metrics including accuracy, sensitivity, specificity, F1-Score, and IoU. Evaluation results indicate that the proposed model achieved an average accuracy of 99.13%, demonstrating its ability to classify pixels with high alignment to ground truth. A sensitivity of 80.78% is classified as good, reflecting adequate balance in detecting True Positives (TP). Specificity of 91.69% indicates excellent performance in identifying Background pixels or True Negatives (TN). An F1-Score of 78.64% shows the model's reasonable balance in performance across classes. An average IoU of 77.34% suggests the model has not yet reached optimal performance in predicting overlapping areas with ground truth for artery and vein labels. This study demonstrates that the combination of U-Net, Vision Transformer, and Attention Gate effectively enhances the performance of semantic segmentation for retinal blood vessels, though improvements are still needed for specific labels.

Keyword: Semantic Segmentation, Retinal Blood Vessel, Vision Transformer, U-Net, Attention Gate

**IMPLEMENTASI ARSITEKTUR *UNET-TRANSFORMER* DALAM
SEGMENTASI SEMANTIK CITRA PEMBULUH DARAH ARTERI-
VENA RETINA**

Oleh:

GIOVILLANDO

08011182126006

ABSTRAK

Pembuluh darah retina adalah komponen penting dalam sistem peredaran darah mata yang berfungsi untuk menyediakan oksigen dan nutrisi serta mengeluarkan zat sisa dari jaringan retina. Pembuluh darah retina terbagi menjadi dua jenis yaitu arteri dan vena (A/V). Arteri dan vena sering terletak berdekatan sehingga perlu dilakukan pemisahan yang jelas untuk membantu tenaga medis dalam mengidentifikasi penyakit khusus dan mencegah kesalahan dalam menganalisis penyakit yang terkait. Pemisahan arteri dan vena dapat dilakukan dengan pendekatan teknologi berbasis segmentasi citra. Penelitian ini bertujuan untuk melakukan segmentasi semantik pembuluh darah retina menggunakan kombinasi arsitektur U-Net, Vision Transformer (ViT), dan Attention Gate. Model yang diusulkan memanfaatkan ViT sebagai encoder untuk menangkap hubungan spasial global, sedangkan U-Net berperan sebagai decoder untuk memulihkan detail spasial citra. Attention Gate ditambahkan untuk menyaring informasi relevan dari fitur yang dihasilkan. Evaluasi kinerja segmentasi dilakukan pada lima kelas label, yaitu Background, Arteri, Crossings, Vena, dan Uncertain dengan mengukur akurasi, sensitivitas, spesifisitas, F1-Score dan IoU. Hasil evaluasi menunjukkan bahwa model yang diusulkan menghasilkan rata-rata akurasi sebesar 99,13% menunjukkan bahwa model mampu mengklasifikasikan piksel dengan tingkat kesesuaian yang tinggi terhadap ground truth. Sensitivitas sebesar 80,78% dapat diklasifikasikan sebagai baik, mencerminkan keseimbangan yang memadai dalam mendeteksi True Positive (TP). Spesifisitas sebesar 91,69% yang mengindikasikan performa yang sangat baik dalam mengidentifikasi piksel yang termasuk kategori Background atau true negative (TN). F1-Score sebesar 78,64% menunjukkan model cukup baik dalam mempertahankan keseimbangan kinerja pada tiap kelasnya. IoU dengan rata-rata 77,34% menunjukkan bahwa model belum optimal dalam memprediksi area yang benar-benar tumpang tindih dengan ground truth untuk label arteri dan vena. Hasil penelitian ini menunjukkan bahwa kombinasi arsitektur U-Net, Vision Transformer, dan Attention Gate efektif dalam meningkatkan kinerja segmentasi semantik pembuluh darah retina, meskipun masih terdapat ruang untuk perbaikan pada label tertentu.

Kata Kunci: *Attention Gate*, Pembuluh Darah Retina, Segmentasi Semantik, *U-Net*, *Vision Transformer*

DAFTAR ISI

LEMBAR PENGESAHAN	ii
PERNYATAAN KEASLIAN KARYA ILMIAH.....	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR.....	v
ABSTRACT.....	viii
ABSTRAK	ix
DAFTAR ISI.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
BAB I PENDAHULUAN.....	14
1.1 Latar Belakang.....	14
1.2 Rumusan Masalah	19
1.3 Batasan Masalah	19
1.4 Tujuan.....	20
1.5 Manfaat.....	20
BAB II TINJAUAN PUSTAKA.....	Error! Bookmark not defined.
2.1 Retinal Vessel	Error! Bookmark not defined.
2.2 Citra Digital.....	Error! Bookmark not defined.
2.3 Segmentasi Citra Semantik	Error! Bookmark not defined.
2.4 <i>Preprocessing Data</i>	Error! Bookmark not defined.
2.5 Augmentasi Data	Error! Bookmark not defined.
2.6 Perbaikan Citra.....	Error! Bookmark not defined.
2.7 <i>Vision Transformer</i>	Error! Bookmark not defined.
2.8 <i>Convolutional Neural Network (CNN)</i>	Error! Bookmark not defined.
a. <i>Convolutional layer</i>	Error! Bookmark not defined.
b. <i>Batch Normalization</i>	Error! Bookmark not defined.
c. Fungsi Aktivasi ReLU.....	Error! Bookmark not defined.
d. <i>Concatenate Layer</i>	Error! Bookmark not defined.
e. <i>Upsampling Layer</i>	Error! Bookmark not defined.
2.9 <i>U-Net</i>	Error! Bookmark not defined.
2.10 <i>Attention Gate</i>	Error! Bookmark not defined.
2.11 Fungsi Aktivasi <i>Softmax</i>	Error! Bookmark not defined.

2.12	<i>Loss Function</i>	Error! Bookmark not defined.
2.13	<i>Optimization Function: Adaptive Moment Estimation</i>	Error! Bookmark not defined.
2.14	<i>Confusion Matrix</i>	Error! Bookmark not defined.
BAB III METODOLOGI PENELITIAN		Error! Bookmark not defined.
3.1	Tempat	Error! Bookmark not defined.
3.2	Waktu	Error! Bookmark not defined.
3.3	Alat	Error! Bookmark not defined.
3.4	Tahapan Penelitian	Error! Bookmark not defined.
3.4.1	Pengumpulan Data	Error! Bookmark not defined.
3.4.2	<i>Preprocessing data</i>	Error! Bookmark not defined.
3.4.3	Tahap <i>Training</i>	Error! Bookmark not defined.
3.4.4	Tahap <i>Testing</i>	Error! Bookmark not defined.
3.4.5	Evaluasi Kinerja Model	Error! Bookmark not defined.
3.4.6	Analisis dan Interpretasi Model	Error! Bookmark not defined.
BAB IV METODOLOGI PENELITIAN		Error! Bookmark not defined.
4.1	Deskripsi Dataset	Error! Bookmark not defined.
4.2	<i>Preprocessing Data</i>	Error! Bookmark not defined.
4.3	Arsitektur Unet-Transformer Attention	Error! Bookmark not defined.
4.3.1	<i>Encoder Vision Transformer</i>	Error! Bookmark not defined.
4.3.2	<i>Decoder dan Attention Gate</i>	Error! Bookmark not defined.
4.3.3	<i>Upsampling Layer</i>	Error! Bookmark not defined.
4.3.4	Perhitungan Manual Attention Gate	Error! Bookmark not defined.
4.3.5	<i>Concatenate Layer</i>	Error! Bookmark not defined.
4.3.6	Fungsi Aktivasi <i>Softmax</i>	Error! Bookmark not defined.
4.3.7	<i>Categorical Cross Entropy Loss Function</i>	Error! Bookmark not defined.
4.3.8	<i>Optimization Function: Adaptive Moment Estimation (Adam)</i>	Error! Bookmark not defined.
4.4	Implementasi Arsitektur Unet-Transformer Attention dalam Penelitian Error! Bookmark not defined.	
4.4.1	Hasil.....	Error! Bookmark not defined.
4.4.2	Pembahasan.....	Error! Bookmark not defined.
BAB V KESIMPULAN DAN SARAN		Error! Bookmark not defined.
5.1	Kesimpulan	Error! Bookmark not defined.

5.2 SaranError! Bookmark not defined.
 DAFTAR PUSTAKA.....Error! Bookmark not defined.

DAFTAR TABEL

Tabel 2.1. *Confusion Matrix***Error! Bookmark not defined.**
 Tabel 2.2. Kategori Kinerja (*Nath et al.* 2014).**Error! Bookmark not defined.**
 Tabel 4.1. Sampel Data Citra pada Dataset DRIVE dan LES-AV **Error! Bookmark not defined.**
 Tabel 4.2. Perhitungan rata-rata (μ_j) untuk setiap *mini batch***Error! Bookmark not defined.**
 Tabel 4.3. Perhitungan varians σ_j^2 setiap *mini batch* **Error! Bookmark not defined.**
 Tabel 4.4. Nilai Bobot *Hidden Layer* dan *Output*..**Error! Bookmark not defined.**
 Tabel 4.5. Perbandingan Citra Asli, *Ground truth*, dan Hasil Segmentasi **Error! Bookmark not defined.**
 Tabel 4.6. *Confusion Matrix* dari Proses *Testing*...**Error! Bookmark not defined.**
 Tabel 4.7. Evaluasi kinerja pada setiap label**Error! Bookmark not defined.**
 Tabel 4.8. Perbandingan hasil rata-rata evaluasi kinerja dengan penelitian lain
**Error! Bookmark not defined.**

DAFTAR GAMBAR

Gambar 2.1. Bagian Arteri dan Vena pada Citra Retina..... **Error! Bookmark not defined.**

Gambar 2.2. Teknik augmentasi data *flipping* dan *rotate* **Error! Bookmark not defined.**

Gambar 2.3. Ilustrasi Arsitektur *Vision Transformer* **Error! Bookmark not defined.**

Gambar 2.4. Struktur *Transformer Encoder***Error! Bookmark not defined.**

Gambar 2.5. Ilustrasi *Multi-Head Self-Attention* ...**Error! Bookmark not defined.**

Gambar 2.6. Ilustrasi *MultiLayer Perceptron***Error! Bookmark not defined.**

Gambar 2.7. Contoh Proses *Convolution Layer*.....**Error! Bookmark not defined.**

Gambar 2.8. Ilustrasi *Upsampling***Error! Bookmark not defined.**

Gambar 2.9. Ilustrasi Arsitektur *U-Net***Error! Bookmark not defined.**

Gambar 4.1. Segmentasi Semantik Pembuluh darah Arteri dan Vena Retina**Error! Bookmark not defined.**

Gambar 4.2. Ilustrasi Arsitektur *Unet-Transformer Attention*....**Error! Bookmark not defined.**

Gambar 4.3. Pemotongan Citra Input menjadi *Patch* **Error! Bookmark not defined.**

Gambar 4. 4. Ilustrasi *Flat Patches*.....**Error! Bookmark not defined.**

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pembuluh darah retina adalah komponen penting dalam sistem peredaran darah mata yang berfungsi untuk menyediakan oksigen dan nutrisi serta mengeluarkan zat sisa dari jaringan retina (O'Leary and Campbell, 2023). Penelitian pembuluh darah retina banyak dilakukan karena dapat diamati tanpa perlu prosedur pembedahan. Pembuluh darah retina dapat memberikan gambaran tentang kondisi kesehatan pembuluh darah (Hu *et al.*, 2021). Pembuluh darah retina terbagi menjadi dua jenis yaitu arteri dan vena (A/V). Letak Arteri dan vena sering terletak berdekatan sehingga perlu dilakukan pemisahan yang jelas untuk membantu tenaga medis dalam mengidentifikasi penyakit khusus dan mencegah kesalahan dalam menganalisis penyakit yang terkait (Remeseiro *et al.*, 2021). Pemisahan arteri dan vena dapat dilakukan dengan pendekatan teknologi berbasis segmentasi citra (Morano *et al.*, 2021).

Segmentasi citra adalah teknik pemrosesan yang digunakan untuk membagi suatu citra menjadi bagian-bagian atau wilayah tertentu, dengan memisahkan pola penting dari pola yang tidak penting seperti *Background* (Ghosh *et al.*, 2019). Namun, segmentasi pembuluh darah retina sebagian besar masih bergantung pada segmentasi manual (C. Chen *et al.*, 2021). Tingkat kerumitan dan variasi pola pembuluh darah retina yang kompleks, mengakibatkan segmentasi manual memiliki kelemahan. Kelemahan tersebut berupa memerlukan penggunaan tenaga ahli medis dalam hal memilah area arteri dengan vena, konsumsi waktu yang cukup lama dan penggunaan tenaga yang tinggi (Kumar *et al.* 2023; Soomro *et al.* 2019). Kelemahan segmentasi manual dapat diatasi dengan menggunakan segmentasi otomatis. Segmentasi otomatis dari citra retina diperlukan untuk mengenali perubahan struktural seperti penyempitan, dan pelebaran

pembuluh darah secara efisien (Girard *et al.* 2019). Salah satu metode segmentasi otomatis adalah *Convolutional Neural Network* (CNN) (X. Han *et al.* 2022).

Convolutional Neural Network (CNN) merupakan algoritma *deep learning* yang banyak digunakan dalam memisahkan fitur-fitur penting dalam citra dan efektif menangkap informasi spasial dengan otomatis (Taye *et al.*, 2023). Salah satu arsitektur CNN yang populer dalam melakukan segmentasi citra adalah *U-Net* (Yin *et al.*, 2022). *U-Net* adalah arsitektur jaringan saraf yang dirancang untuk segmentasi citra (Punn & Agarwal, 2022). Arsitektur *U-Net* mampu menangkap fitur detail dari Citra input sekaligus mempertahankan informasi spasial yang penting (Williams *et al.* 2023). *U-Net* dapat melakukan segmentasi pada tingkat piksel, sehingga sangat akurat dalam mendeteksi batas-batas objek yang kecil dan kompleks, seperti pembuluh darah dan organ dalam Citra medis (Xiuqin *et al.* 2019)

Struktur dasar arsitektur *U-Net* terdiri dari 3 jalur yaitu *encoder*, *bridge* dan *decoder*. Jalur pertama adalah *encoder*, yang berperan dalam menangkap representasi fitur dari citra dengan menerapkan serangkaian operasi konvolusi dan *down-sampling* (Yin *et al.* 2022). Pada *encoder* terdapat operasi konvolusi yang digunakan untuk menangkap fitur-fitur dari input citra. Operasi konvolusi pada *encoder* diikuti oleh operasi *downsampling* yang biasanya dilakukan dengan *pooling*. Operasi *downsampling* bertujuan untuk mengecilkan resolusi Citra mempertahankan informasi penting (Huang *et al.* 2020; Nirthika *et al.* 2022). Jalur kedua adalah jalur *bridge*, yang berperan sebagai penghubung antara jalur *encoder* dengan jalur *decoder* (Zhao *et al.* 2023). Pada *bridge* digunakan operasi *concatenate* yang bertugas untuk menggabungkan fitur dari *encoder* dan *decoder* (J. Li *et al.* 2022). Jalur ketiga adalah *decoder*, digunakan untuk memulihkan resolusi asli Citra dengan menerapkan *up-convolutions* atau *upsampling* (Siddique *et al.* 2021). Proses *upsampling* bertujuan untuk meningkatkan kembali resolusi citra sehingga hasil segmentasi dapat

sesuai dengan dimensi asli Citra. Di dalam jalur *decoder*, terdapat *skip connections*, yang berfungsi untuk menggabungkan fitur dari *encoder* yang memiliki resolusi yang sama, menjaga agar informasi spasial yang penting tetap dipertahankan selama proses decoding (Xu *et al.* 2020).

Sathananthavathi & Indumathi (2021) melakukan penelitian segmentasi pembuluh darah retina menggunakan modifikasi arsitektur *U-Net* pada bagian *encoder* dengan meningkatkan proses *concatenate* menggunakan *addition layer* yaitu *Encoder Enhanced Atrous U-Net* (EEA-*UNet*). EEA-*UNet* memperoleh nilai sensitivitas yang dihasilkan masih berada di bawah 79,18%. Yu *et al.* (2021) juga mengusulkan modifikasi *U-Net* dengan *Multi-Scale Iterative Aggregation* (MIA-*Unet*). Namun, pada dataset STARE, DRIVE dan CHASE_DB1 nilai *Sensitivitas* yang diperoleh dibawah 85%. *U-Net* menunjukkan kinerja yang baik dalam segmentasi citra medis, terutama pembuluh darah retina. Meskipun efektif dalam meningkatkan fokus arsitektur dalam melakukan segmentasi, *U-Net* memiliki beberapa kelemahan terutama pada bagian encoder. *U-Net* menggunakan lapisan konvolusi untuk mengekstrak fitur, yang lebih terfokus pada informasi lokal dalam gambar mengakibatkan ketidakmampuan dalam memahami konteks yang lebih besar, seperti posisi arteri dan vena yang berdekatan namun terpisah secara visual (Ouyang *et al.*, 2023; Siddique *et al.*, 2021). Salah satu arsitektur yang dapat mengatasi keterbatasan tersebut adalah *Vision Transformer* (Dosovitskiy *et al.* 2020).

Vision Transformer (ViT) adalah arsitektur *deep learning* yang diperkenalkan untuk tugas-tugas pengolahan citra seperti segmentasi citra (H. Wu *et al.* 2021). ViT berbeda dengan CNN yang menggunakan operasi konvolusi untuk mengekstraksi fitur dari Citra, ViT memiliki fitur *Image Patch Embedding* yang membagi Citra menjadi *patch* kecil yang kemudian diubah menjadi token dan diproses oleh *transformer encoder* (Touvron *et al.* 2021). Pada *Transformer Encoder* token-token yang diterima diproses menggunakan mekanisme *Self-Attention*. *Self-Attention*

berbeda dengan konvolusi yang beroperasi secara lokal, *Self-Attention* menangkap hubungan global antar elemen input *Self-Attention* memungkinkan model untuk fokus pada bagian-bagian yang penting dari citra secara global (Carion *et al.* 2020). Chen *et al.* (2022) menerapkan arsitektur *Transformer* yang dimodifikasi dengan *U-Net* yaitu *PCAT-Unet* dalam Segmentasi Pembuluh darah retina. Penelitian ini menghasilkan sensitivitas sebesar 79%. Ruiping *et al.* (2024) mengusulkan *Vision Transformer* yang dimodifikasi dengan jaringan piramida dan piramida pooling (PPM) yaitu *ViT-UperNet*. Penelitian ini melakukan segmentasi untuk Citra pembuluh darah retina dan memperoleh akurasi adalah 93,85% dan Dice sebesar 92,61%. Namun, pada penelitian ini hanya menggunakan pengukuran Akurasi dan *Dice*.

Modifikasi ViT dengan U-Net dapat menjalankan tugas segmentasi. Hal ini dikarenakan *ViT* merupakan arsitektur yang dirancang untuk tugas klasifikasi citra, sehingga tidak dapat menjalankan tugas segmentasi (Thisanke *et al.*, 2023). *ViT* tidak memiliki komponen *decoder* yang diperlukan untuk menghasilkan prediksi *piksel* yang detail dan akurat (Xia *et al.*, 2024). *Decoder* adalah komponen yang berfungsi untuk memproses fitur yang diekstraksi oleh *encoder* yang berisi prediksi untuk setiap piksel pada citra. *Decoder* pada U-Net bekerja dengan cara menggabungkan informasi fitur dari *encoder* dengan informasi spasial dari resolusi tinggi melalui *skip connections* (Khanh *et al.*, 2020).

U-Net memiliki keterbatasan dalam mekanisme *skip connections* yang digunakan (Desiani, Erwin, *et al.* 2022). *Skip connections* merupakan lapisan yang menghubungkan fitur hasil rekonstruksi dari *encoder* ke *decoder* U-Net *Skip connections* menggabungkan semua informasi dari lapisan *encoder* ke *decoder* tanpa mempertimbangkan relevansi atau kepentingan fitur yang disalurkan (Azad *et al.*, 2024). Kelemahan ini dapat menyebabkan terjadinya penyatuan informasi yang tidak relevan, sehingga mengakibatkan terjadinya *overlapping* (Wang *et al.*, 2024).

Mekanisme tambahan di bagian *decoder* yang dapat menyaring dan memilih fitur yang relevan, agar model dapat lebih fokus pada fitur penting adalah *Attention Gate*.

Attention gate merupakan mekanisme dalam jaringan saraf yang dirancang untuk mengarahkan model *attention* ke area yang lebih relevan dalam citra (Shen *et al.* 2019). Mekanisme *Attention Gate* berfungsi untuk menyaring informasi yang dikirim melalui *skip connections*, sehingga hanya fitur-fitur yang relevan dan penting diteruskan ke bagian *decoder* (Zhang *et al.* 2020). Seleksi fitur pada mekanisme *Attention Gate* dilakukan dengan menghitung koefisien *Attention* berupa bobot nilai pada fitur-fitur yang relevan untuk proses segmentasi (Bhatkalkar *et al.* 2020). Koefisien ini dihitung menggunakan fungsi aktivasi *sigmoid* dan memiliki rentang nilai antara nol hingga satu. Semakin tinggi bobot nilai pada suatu fitur, maka semakin besar tingkat kepentingannya dalam proses segmentasi (Khanh *et al.* 2020). Keuntungan dari *Attention Gate* yang diusulkan adalah dapat meningkatkan sensitivitas model dan akurasi prediksi label yang padat dengan menekan aktivasi fitur di wilayah yang tidak berkorelasi (Rahman *et al.*, 2024).

Wu *et al.* (2019) menerapkan *Attention Gate* pada bagian *decoder U-Net* dan *Denseblock* dalam model *U-Net*. Pada penelitian tersebut data yang digunakan hanya dataset DRIVE sehingga pengujian yang dilakukan terbatas dengan hasil pengukuran sensitivitasnya masih rendah diantara 76%. K. Li *et al.* (2020) juga menerapkan *Attention Gate* di dalam model *U-NET* dengan nama *AGNET* pada segmentasi pembuluh darah retina dengan nilai sensitivitas *AGNET* hanya 83%. Selain itu, pada kedua penelitian tersebut segmentasi yang dilakukan hanya menggunakan 2 label saja, sehingga penggunaan mekanisme dari *attention gate* kurang maksimal. Pembuluh darah retina terdiri dari 2 macam pembuluh darah retina yaitu arteri dan vena yang mana pada tiap pembuluh darah tersebut juga dapat terpengaruh secara berbeda oleh variasi jenis penyakit

(Krestanova *et al.* 2020). *Arsitektur U-Net* yang dikombinasikan dengan *Attention gate* mampu meningkatkan fokus pada fitur-fitur penting dalam segmentasi (Khanh *et al.* 2020).

Penelitian ini mengusulkan kombinasi arsitektur *U-Net* dan *Vision Transformer* (ViT) untuk segmentasi citra dengan lima kelas yaitu *Background*, *arteries*, *veins*, *Uncertain*, dan *crossing*. Pada penelitian ini, model ViT bertindak sebagai *encoder* untuk menangkap hubungan global dalam citra. ViT akan membagi citra menjadi *patch* dan menambahkan *positional embedding* untuk mempertahankan informasi spasial. Hasil keluaran dari ViT kemudian dihubungkan dengan *decoder U-Net*, yang melakukan *upsampling* pada fitur global tersebut. Di *decoder*, terdapat *Attention Gate* yang berfungsi untuk memprioritaskan informasi penting dari fitur yang dihasilkan. Dengan memadukan ViT yang memahami konteks global dan *U-Net* yang melakukan rekonstruksi detail spasial, model ini diharapkan dapat meningkatkan akurasi dan kualitas segmentasi citra secara keseluruhan.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah bagaimana menimplementasi *Unet-Transformer* dalam melakukan segmentasi semantik pembuluh darah arteri dan vena citra retina?

1.3 Batasan Masalah

Batasan masalah yang diterapkan pada penelitian ini adalah Evaluasi kinerja pada penelitian ini menggunakan metode evaluasi kinerja *Confusion Matrix* yaitu akurasi, sensitivitas, spesifisitas, *F1-score* dan IoU.

1.4 Tujuan

Tujuan penelitian ini adalah

- a. Penelitian ini bertujuan mengevaluasi efektivitas model *Unet-Transformer* dalam segmentasi semantik pembuluh darah retina dengan fokus pada lima label yaitu *Background*, *Arteri*, *Vena*, *Crossings*, dan *Uncertain*.
- b. Penelitian ini bertujuan untuk mengukur kinerja model dalam meningkatkan akurasi, sensitivitas, spesifisitas, F1-score, dan IoU melalui pendekatan segmentasi otomatis yang mampu membedakan antara kategori tersebut secara efektif.

1.5 Manfaat

Manfaat dari hasil penelitian ini adalah memberikan bantuan kepada para ahli medis dalam melakukan segmentasi otomatis pembuluh darah arteri dan vena retina dengan hasil yang konsisten sehingga dapat melakukan analisis pada citra retina.

DAFTAR PUSTAKA

- Ali, A. M., Benjdira, B., Koubaa, A., El-Shafai, W., Khan, Z., & Boulila, W. (2023). Vision Transformers in Image Restoration: A Survey. *Journal of Sensors*, 23(5). <https://doi.org/10.3390/s23052385>
- Alomar, K., Aysel, H. I., & Cai, X. (2023). Data Augmentation in Classification and Segmentation: A Survey and New Strategies. *Journal of Imaging*, 9(2). <https://doi.org/10.3390/jimaging9020046>
- Anaya-Isaza, A., Mera-Jiménez, L., & Zequera-Díaz, M. (2021). Overview of deep learning in medical imaging. *Journal of Informatics in Medicine Unlocked*, 26(1). <https://doi.org/10.1016/j.imu.2021.100723>
- Araslanov, N., & Roth, S. (2020). Single-stage semantic segmentation from image labels. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, i, 4252–4261. <https://doi.org/10.1109/CVPR42600.2020.00431>
- Azad, R., Aghdam, E. K., Rauland, A., Jia, Y., Avval, A. H., Bozorgpour, A., Karimijafarbigloo, S., Cohen, J. P., Adeli, E., & Merhof, D. (2024). Medical Image Segmentation Review: The Success of U-Net. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12), 10076–10095. <https://doi.org/10.1109/TPAMI.2024.3435571>
- Bagchi, S., & Bathula, D. R. (2022). EEG-ConvTransformer for single-trial EEG-based visual stimulus classification. *Pattern Recognition*, 129, 108757. <https://doi.org/https://doi.org/10.1016/j.patcog.2022.108757>
- Banerjee, C., Mukherjee, T., & Pasiliao, E. (2020). Feature representations using the reflected rectified linear unit (RReLU) activation. *Big Data Mining and Analytics*, 3(2), 102–120. <https://doi.org/10.26599/BDMA.2019.9020024>
- Bashmal, L., Bazi, Y., & Al Rahhal, M. (2021). Deep Vision Transformers for Remote Sensing Scene Classification. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2815–2818. <https://doi.org/10.1109/IGARSS47720.2021.9553684>
- Bhatkalkar, B. J., Reddy, D. R., Prabhu, S., & Bhandary, S. V. (2020). Improving the Performance of Convolutional Neural Network for the Segmentation of Optic Disc in Fundus Images Using Attention Gates and Conditional Random Fields. *IEEE Access*, 8, 29299–29310. <https://doi.org/10.1109/ACCESS.2020.2972318>
- Biswas, R., Vasan, A., & Roy, S. S. (2020). Dilated Deep Neural Network for Segmentation of Retinal Blood Vessels in Fundus Images. *Iranian Journal of Science and Technology - Transactions of Electrical Engineering*, 44(1), 505–518. <https://doi.org/10.1007/s40998-019-00213-7>
- Cao, J., Li, Y., Sun, M., Chen, Y., Lischinski, D., Cohen-Or, D., Chen, B., & Tu, C. (2022). DO-Conv: Depthwise Over-Parameterized Convolutional Layer. *IEEE Transactions on Image Processing*, 31, 3726–3736. <https://doi.org/10.1109/TIP.2022.3175432>

- Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-End Object Detection with Transformers. *European Conference on Computer Vision*, 213–229. https://doi.org/10.1007/978-3-030-58452-8_13
- Chen, C., Chuah, J. H., Ali, R., & Wang, Y. (2021). Retinal vessel segmentation using deep learning: A review. *IEEE Access*, 9, 111985–112004. <https://doi.org/10.1109/ACCESS.2021.3102176>
- Chen, D., Yang, W., Wang, L., Tan, S., Lin, J., & Bu, W. (2022). PCAT-UNet: UNet-like network fused convolution and transformer for retinal vessel segmentation. *PLoS ONE*, 17(1), 1–22. <https://doi.org/10.1371/journal.pone.0262689>
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen’s Kappa and Brier Score in Binary Classification Assessment. *IEEE Access*, 9, 78368–78381. <https://doi.org/10.1109/ACCESS.2021.3084050>
- Deng, W., Shi, Q., Member, S., & Li, J. (2021). Attention-Gate-Based Encoder – Decoder Network for Automatic Building Extraction. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 2611–2620. <https://doi.org/10.1109/JSTARS.2021.3058097>
- Desiani, A., Adrezo, M., Marselina, N. C., Arhami, M., Salsabila, A., & Al-Filambany, M. G. (2022). A Combination of Image Enhancement and U-Net Architecture for Segmentation in Identifying Brain Tumors on CT-SCAN Images. *International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, 423–428. <https://doi.org/10.1109/ICIMCIS56303.2022.10017519>
- Desiani, A., Erwin, Suprihatin, B., Efriliyanti, F., Arhami, M., & Setyaningsih, E. (2022). VG-DropDNet a Robust Architecture for Blood Vessels Segmentation on Retinal Image. *IEEE Access*, 10(June), 92067–92083. <https://doi.org/10.1109/ACCESS.2022.3202890>
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *Computer Vision and Pattern Recognition*. <http://arxiv.org/abs/2010.11929>
- Duong, H. T., & Nguyen-Thi, T. A. (2021). A review: preprocessing techniques and data augmentation for sentiment analysis. *Computational Social Networks*, 8(1), 1–16. <https://doi.org/10.1186/s40649-020-00080-x>
- Fan, H., Xiong, B., Mangalam, K., Li, Y., Yan, Z., Malik, J., & Feichtenhofer, C. (2021). Multiscale Vision Transformers. *Proceedings of the IEEE International Conference on Computer Vision*, 6804–6815. <https://doi.org/10.1109/ICCV48922.2021.00675>
- Frańk, W., Wojtasińska, A., Lisińska, W., Młynarska, E., Franczyk, B., & Rysz, J. (2022). Pathophysiology of Cardiovascular Diseases: New Insights into Molecular Mechanisms of Atherosclerosis, Arterial Hypertension, and Coronary Artery Disease. *Biomedicines*, 10(8). <https://doi.org/10.3390/biomedicines10081938>

- Ghosh, S., Das, N., Das, I., & Maulik, U. (2019). Understanding deep learning techniques for image segmentation. *ACM Computing Surveys*, 52(4), 1–58. <https://doi.org/10.1145/3329784>
- Girard, F., Kavalec, C., & Cheriet, F. (2019). Joint segmentation and classification of retinal arteries/veins from fundus images. *Artificial Intelligence in Medicine*, 94, 96–109. <https://doi.org/10.1016/j.artmed.2019.02.004>
- Grimstad, B., & Andersson, H. (2019). ReLU networks as surrogate models in mixed-integer linear programs. *Computers and Chemical Engineering*, 131, 106580. <https://doi.org/10.1016/j.compchemeng.2019.106580>
- Guo, C., Szemenyei, M., Pei, Y., Yi, Y., & Zhou, W. (2019). SD-Unet: A Structured Dropout U-Net for Retinal Vessel Segmentation. *Proceedings - 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering, BIBE 2019, August 2020*, 439–444. <https://doi.org/10.1109/BIBE.2019.00085>
- Han, K., Xiao, A., Wu, E., Guo, J., Xu, C., & Wang, Y. (2021). Transformer in Transformer. *Advances in Neural Information Processing Systems*, 34, 15908–15919. https://proceedings.neurips.cc/paper_files/paper/2021/file/854d9fca60b4bd07f9bb215d59ef5561-Paper.pdf
- Han, X., Wu, X., Wang, S., Xu, L., Xu, H., Zheng, D., Yu, N., Hong, Y., Yu, Z., Yang, D., & Yang, Z. (2022). Automated segmentation of liver segment on portal venous phase MR images using a 3D convolutional neural network. *Insights into Imaging*, 13(1), 1–10. <https://doi.org/10.1186/s13244-022-01163-1>
- Hu, J., Wang, H., Cao, Z., Wu, G., Jonas, J. B., Wang, Y. X., & Zhang, J. (2021). Automatic Artery/Vein Classification Using a Vessel-Constraint Network for Multicenter Fundus Images. *Frontiers in Cell and Developmental Biology*, 9(June), 1–15. <https://doi.org/10.3389/fcell.2021.659941>
- Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y. W., & Wu, J. (2020). UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2020-May(ii)*, 1055–1059. <https://doi.org/10.1109/ICASSP40776.2020.9053405>
- Ilesanmi, A. E., Ilesanmi, T., & Gbotoso, G. A. (2023). A systematic review of retinal fundus image segmentation and classification methods using convolutional neural networks. *Healthcare Analytics*, 4(May), 100261. <https://doi.org/10.1016/j.health.2023.100261>
- Jadon, S. (2020). A survey of loss functions for semantic segmentation. *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2020*. <https://doi.org/10.1109/CIBCB48159.2020.9277638>
- Jiang, D., Qu, H., Zhao, J., Zhao, J., & Hsieh, M. Y. (2021). Aggregating multi-scale contextual features from multiple stages for semantic image segmentation. *Connection Science*, 33(3), 605–622. <https://doi.org/10.1080/09540091.2020.1862059>

- Kar, M. K., Nath, M. K., & Neog, D. R. (2021). A Review on Progress in Semantic Image Segmentation and Its Application to Medical Images. *SN Computer Science*, 2(5), 397. <https://doi.org/10.1007/s42979-021-00784-5>
- Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 173, 24–49. <https://doi.org/10.1016/j.isprsjprs.2020.12.010>
- Khan, A., Sohail, A., Zahoor, U., & Qureshi, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, 53(8), 5455–5516. <https://doi.org/10.1007/s10462-020-09825-6>
- Khanh, T. L. B., Dao, D. P., Ho, N. H., Yang, H. J., Baek, E. T., Lee, G., Kim, S. H., & Yoo, S. B. (2020). Enhancing U-net with spatial-channel attention gate for abnormal tissue segmentation in medical imaging. *Applied Sciences (Switzerland)*, 10(17), 1–19. <https://doi.org/10.3390/AP10175729>
- Krestanova, A., Kubicek, J., & Penhaker, M. (2020). Recent techniques and trends for retinal blood vessel extraction and tortuosity evaluation: A comprehensive review. *IEEE Access*, 8, 197787–197816. <https://doi.org/10.1109/ACCESS.2020.3033027>
- Kumar, A., Agrawal, R. K., & Joseph, L. (2023). IterMiUnet: A lightweight architecture for automatic blood vessel segmentation. *Multimedia Tools and Applications*, 82(28), 43207–43231. <https://doi.org/10.1007/s11042-023-15433-7>
- Lakshmipriya, B., Pottakkat, B., & Ramkumar, G. (2023). Deep learning techniques in liver tumour diagnosis using CT and MR imaging - A systematic review. *Artificial Intelligence in Medicine*, 141, 102557. <https://doi.org/10.1016/j.artmed.2023.102557>
- Li, J., Gao, G., Yang, L., Liu, Y., & Yu, H. (2022). DEF-Net: A Dual-Encoder Fusion Network for Fundus Retinal Vessel Segmentation. *Electronics (Switzerland)*, 11(22), 1–12. <https://doi.org/10.3390/electronics11223810>
- Li, J., Sun, H., & Li, J. (2023). Beyond confusion matrix: learning from multiple annotators with awareness of instance features. *Machine Learning*, 112(3), 1053–1075. <https://doi.org/10.1007/s10994-022-06211-x>
- Li, K., Yao, Z., Luo, Y., Qi, X., Liu, P., & Wang, Z. (2020). Retinal Blood Vessel Segmentation via Attention Gate Network. *ACM International Conference Proceeding Series*, 247–251. <https://doi.org/10.1145/3429889.3429936>
- Liu, W., Zhu, F., Ma, S., & Liu, C.-L. (2024). MSPE: Multi-Scale Patch Embedding Prompts Vision Transformers to Any Resolution. *Computer Vision and Pattern Recognition*, 1–16. <http://arxiv.org/abs/2405.18240>
- Lv, Y., Ma, H., Li, J., & Liu, S. (2020). Attention Guided U-Net with Atrous Convolution for Accurate Retinal Vessels Segmentation. *IEEE Access*, 8, 32826–32839. <https://doi.org/10.1109/ACCESS.2020.2974027>

- Maharana, K., Mondal, S., & Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 3(1), 91–99.
- Martinez-Perez, M. E., Parker, K. H., Witt, N., Hughes, A. D., & Thom, S. A. M. (2020). Automatic artery/vein classification in colour retinal images. *Twelfth International Conference on Machine Vision, 11433*, 346–353.
- Morano, J., Hervella, Á. S., Novo, J., & Rouco, J. (2021). Simultaneous segmentation and classification of the retinal arteries and veins from color fundus images. *Artificial Intelligence in Medicine*, 118. <https://doi.org/10.1016/j.artmed.2021.102116>
- Mostafa, S., & Wu, F.-X. (2021). Diagnosis of autism spectrum disorder with convolutional autoencoder and structural MRI images. In *Neural engineering techniques for autism spectrum disorder* (pp. 23–38). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-12-822822-7.00003-X>
- Nath, S. S., Mishra, G., Kar, J., Chakraborty, S., & Dey, N. (2014). A survey of image classification methods and techniques. *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies, ICCICCT 2014*, 554–557. <https://doi.org/10.1109/ICCICCT.2014.6993023>
- Nirthika, R., Manivannan, S., Ramanan, A., & Wang, R. (2022). Pooling in convolutional neural networks for medical image analysis: a survey and an empirical study. *Neural Computing and Applications*, 34(7), 5321–5347. <https://doi.org/10.1007/s00521-022-06953-8>
- O’Leary, F., & Campbell, M. (2023). The blood–retina barrier in health and disease. *FEBS Journal*, 290(4), 878–891. <https://doi.org/10.1111/febs.16330>
- Oğuz, A., & Ertuğrul, Ö. F. (2023). Introduction to deep learning and diagnosis in medicine. In *Diagnostic Biomedical Signal and Image Processing Applications with Deep Learning Methods* (pp. 1–40). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-323-96129-5.00003-2>
- Oltu, B., Akşahin, M. F., & Kibaroglu, S. (2021). A novel electroencephalography based approach for Alzheimer’s disease and mild cognitive impairment detection. *Biomedical Signal Processing and Control*, 63, 102223. <https://doi.org/https://doi.org/10.1016/j.bspc.2020.102223>
- Ouyang, J., Liu, S., Peng, H., Garg, H., & Thanh, D. N. H. (2023). LEA U-Net: a U-Net-based deep learning framework with local feature enhancement and attention for retinal vessel segmentation. *Complex and Intelligent Systems*, 9(6), 6753–6766. <https://doi.org/10.1007/s40747-023-01095-3>
- Punn, N. S., & Agarwal, S. (2022). Modality specific U-Net variants for biomedical image segmentation: a survey. *Artificial Intelligence Review*, 55(7), 5845–5889. <https://doi.org/10.1007/s10462-022-10152-1>
- Rahman, C. M. A., Bhuiyan, R. K., Shyam, S. P., Subnom, R., & Rashid, A. Bin. (2024). Attention

- Enabled MultiResUNet for Bio-Medical Image Segmentation. *International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT)*, 622–627. <https://doi.org/https://doi.org/10.1109/ICEEICT62016.2024.10534532>
- Remeseiro, B., Mendonça, A. M., & Campilho, A. (2021). Automatic classification of retinal blood vessels based on multilevel thresholding and graph propagation. *Visual Computer*, 37(6), 1247–1261. <https://doi.org/10.1007/s00371-020-01863-z>
- Reyad, M., Sarhan, A. M., & Arafa, M. (2023). A modified Adam algorithm for deep neural network optimization. *Neural Computing and Applications*, 35(23), 17095–17112. <https://doi.org/10.1007/s00521-023-08568-z>
- Ruiping, Y., Kun, L., Shaohua, X., Jian, Y., & Zhen, Z. (2024). ViT-UperNet: a hybrid vision transformer with unified-perceptual-parsing network for medical image segmentation. *Complex and Intelligent Systems*, 10(3), 3819–3831. <https://doi.org/10.1007/s40747-024-01359-6>
- Salvi, M., Acharya, U. R., Molinari, F., & Meiburger, K. M. (2021). The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. *Computers in Biology and Medicine*, 128, 104129. <https://doi.org/10.1016/j.compbimed.2020.104129>
- Sarki, R., Ahmed, K., Wang, H., Zhang, Y., Ma, J., & Wang, K. (2021). Image Preprocessing in Classification and Identification of Diabetic Eye Diseases. *Data Science and Engineering*, 6(4), 455–471. <https://doi.org/10.1007/s41019-021-00167-z>
- Sathananthavathi, V., & Indumathi, G. (2021). Encoder Enhanced Atrous (EEA) Unet architecture for Retinal Blood vessel segmentation. *Cognitive Systems Research*, 67, 84–95. <https://doi.org/10.1016/j.cogsys.2021.01.003>
- Shajun Nisha, S., & Nagoor Meeral, M. (2021). Applications of deep learning in biomedical engineering. In *Handbook of Deep Learning in Biomedical Engineering* (pp. 245–270). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-12-823014-5.00008-9>
- Shen, Y., Fang, Z., Gao, Y., Xiong, N., Zhong, C., & Tang, X. (2019). Coronary Arteries Segmentation Based on 3D FCN with Attention Gate and Level Set Function. *IEEE Access*, 7, 42826–42835. <https://doi.org/10.1109/ACCESS.2019.2908039>
- Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. *IEEE Access*, 9, 82031–82057. <https://doi.org/10.1109/ACCESS.2021.3086020>
- Singh, V., Asari, V. K., & Rajasekaran, R. (2022). A Deep Neural Network for Early Detection and Prediction of Chronic Kidney Disease. *Diagnostics*, 12(1), 1–22. <https://doi.org/10.3390/diagnostics12010116>
- Soomro, T. A., Afifi, A. J., Zheng, L., Soomro, S., Gao, J., Hellwich, O., & Paul, M. (2019). Deep Learning Models for Retinal Blood Vessels Segmentation: A Review. *IEEE Access*, 7,

71696–71717. <https://doi.org/10.1109/ACCESS.2019.2920616>

- Strudel, R., Garcia, R., Laptev, I., & Schmid, C. (2021). Segmenter: Transformer for Semantic Segmentation. *Proceedings of the IEEE International Conference on Computer Vision*, 7242–7252. <https://doi.org/10.1109/ICCV48922.2021.00717>
- Taye, M. M. (2023). Theoretical Understanding of Convolutional Neural Network : *Mdpi Ag*, 11(3), 52.
- Thisanke, H., Deshan, C., Chamith, K., Seneviratne, S., Vidanaarachchi, R., & Herath, D. (2023). Semantic Segmentation using Vision Transformers: A survey. *Engineering Applications of Artificial Intelligence*, abs/2305.0. <https://doi.org/https://doi.org/10.48550/arXiv.2305.03273>
- Tian, F., Gao, Y., Fang, Z., & Gu, J. (2021). Automatic coronary artery segmentation algorithm based on deep learning and digital image processing. *Applied Intelligence*, 51(12), 8881–8895.
- Tiwari, A. (2022). Supervised learning: From theory to applications. In: Artificial Intelligence and Machine Learning for EDGE Computing. In *Artificial Intelligence and Machine Learning for EDGE Computing* (pp. 23–32). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-12-824054-0.00026-5>
- Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., & Jégou, H. (2021). Training data-efficient image transformers & distillation through attention. *Proceedings of Machine Learning Research*, 139, 10347–10357.
- Wang, H., Cao, P., Yang, J., & Zaiane, O. (2024). Narrowing the semantic gaps in U-Net with learnable skip connections: The case of medical image segmentation. *Neural Networks*, 178. <https://doi.org/10.1016/j.neunet.2024.106546>
- Williams, C., Falck, F., Deligiannidis, G., Holmes, C., Doucet, A., & Syed, S. (2023). A Unified Framework for U-Net Design and Analysis. *Advances in Neural Information Processing Systems*, 36(NeurIPS), 1–38.
- Wu, C., Zou, Y., & Zhan, J. (2019). DA-U-Net: Densely Connected Convolutional Networks and Decoder with Attention Gate for Retinal Vessel Segmentation. *IOP Conference Series*, 533(1). <https://doi.org/10.1088/1757-899X/533/1/012053>
- Wu, H., Xiao, B., Codella, N., Liu, M., Dai, X., Yuan, L., & Zhang, L. (2021). CvT: Introducing Convolutions to Vision Transformers. *Proceedings of the IEEE International Conference on Computer Vision*, 22–31. <https://doi.org/10.1109/ICCV48922.2021.00009>
- Xia, C., Wang, X., Lv, F., Hao, X., & Shi, Y. (2024). ViT-CoMer: Vision Transformer with Convolutional Multi-scale Feature Interaction for Dense Predictions. *CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 5493–5502. <https://doi.org/https://doi.org/10.1109/CVPR52733.2024.00525>
- Xiuqin, P., Zhang, Q., Zhang, H., & Li, S. (2019). A fundus retinal vessels segmentation scheme

- based on the improved deep learning u-net model. *IEEE Access*, 7, 122634–122643. <https://doi.org/10.1109/ACCESS.2019.2935138>
- Xu, W., Deng, X., Guo, S., Chen, J., Sun, L., Zheng, X., Xiong, Y., Shen, Y., & Wang, X. (2020). High-resolution u-net: Preserving image details for cultivated land extraction. *Sensors (Switzerland)*, 20(15), 1–23. <https://doi.org/10.3390/s20154064>
- Yin, X. X., Sun, L., Fu, Y., Lu, R., & Zhang, Y. (2022). U-Net-Based Medical Image Segmentation. *Journal of Healthcare Engineering*. <https://doi.org/10.1155/2022/4189781>
- Yu, L., Qin, Z., Ding, Y., & Qin, Z. (2021). MIA-UNet: Multi-scale iterative aggregation u-network for retinal vessel segmentation. *CMES - Computer Modeling in Engineering and Sciences*, 129(2), 805–827. <https://doi.org/10.32604/cmes.2021.017332>
- Zafar, A., Aamir, M., Nawari, N. M., Arshad, A., Riaz, S., Alruban, A., Dutta, A. K., & Almotairi, S. (2022). A Comparison of Pooling Methods for Convolutional Neural Networks. *Applied Intelligence*, 12(17), 8643.
- Zhang, J., Jiang, Z., & Dong, J. (2020). Attention Gate ResU-Net for Automatic MRI Brain Tumor Segmentation. *IEEE Access*, 8, 58533–58545. <https://doi.org/10.1109/ACCESS.2020.2983075>
- Zhao, L., Ye, L., Zhang, M., Jiang, H., Yang, Z., & Yang, M. (2023). DPSDA-Net: Dual-Path Convolutional Neural Network with Strip Dilated Attention Module for Road Extraction from High-Resolution Remote Sensing Images. *Remote Sensing*, 15(15). <https://doi.org/10.3390/rs15153741>
- Zhou, X., Liu, H., Shi, C., & Liu, J. (2022). The basics of deep learning. In *Deep Learning on Edge Computing Devices* (pp. 19–36). Elsevier. <https://doi.org/https://doi.org/10.1016/B978-0-32-385783-3.00009-0>