

**ARSITEKTUR *VOLUMETRIC U-NET* DAN *TRANSFORMER*
UNTUK SEGMENTASI TUMOR OTAK PADA CITRA
HASIL *MAGNETIC RESONANCE IMAGING***

DRAFT SKRIPSI

Oleh:

FAISHAL FITRA RAMADHAN

NIM 08011282126060



**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SRIWIJAYA
2025**

LEMBAR PENGESAHAN

**ARSITEKTUR *VOLUMETRIC U-NET* DAN *TRANSFORMER*
UNTUK SEGMENTASI TUMOR OTAK PADA CITRA
*HASIL MAGNETIC RESONANCE IMAGING***

SKRIPSI

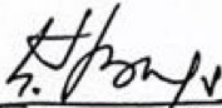
**Sebagai Salah Satu Syarat untuk Memperoleh Gelar
Sarjana Sains**

Oleh

**FAISHAL FITRA RAMADHAN
NIM.08011282126060**

Pembimbing Kedua

**Indralaya, 22 Mei 2025
Pembimbing Utama**



**Dr. Bambang Suprihatin, S.Si., M.Si.
NIP. 197101261994121001**



**Drs. Ali Amran, M.T.
NIP. 196612131994021001**

**Mengetahui,
Ketua Jurusan Matematika**



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

HALAMAN PERSEMBAHAN

Kupersembahkan skripsi ini untuk:

Yang Maha Kuasa Allah Subhanahu Wa Ta'ala,

Kedua orang tuaku tercinta,

Saudara laki-laki dan perempuanku,

Keluarga besarku,

Semua guru dan dosenku,

Sahabat-sahabatku,

Almamaterku.

Moto

“The only reason for time is so that everything doesn't happen at once.”

PERNYATAAN KEASLIAN KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : Faishal Fitra Ramadhan

NIM : 08011282126060

Fakultas/Jurusan : Matematika dan Ilmu Pengetahuan Alam/Matematika

Menyatakan bahwa skripsi ini adalah hasil karya 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 didalam skripsi ini yang berasal dari penulis lain baik yang dipublikasi atau telah diberikan penghargaan dengan mengutip nama sumber penulis yang secara benar. Semua isi dari skripsi ini sepenuhnya menjadi tanggung jawab saya sebagai penulis.

Demikianlah surat pernyataan ini saya buat dengan sebenarnya.

Indralaya, 22 Mei 2025
Penulis



Faishal Fitra Ramadhan
NIM. 08011282126060

KATA PENGANTAR

Puji syukur kehadirat Allah Subhanahu wa Ta'ala yang telah memberikan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Arsitektur *Volumetric U-Net* dan *Transformer* Untuk Segmentasi Tumor Otak Pada Citra Hasil *Magnetic Resonance Imaging*” sebagai salah satu syarat memperoleh gelar sarjana matematika di Fakultas MIPA Universitas Sriwijaya.

Penulis menyadari bahwa proses pembuatan skripsi ini merupakan proses pembelajaran yang sangat berharga serta tak lepas dari kekurangan dan keterbatasan. Dengan segala hormat dan kerendahan hati, penulis mengucapkan terima kasih dan penghargaan kepada:

1. Teristimewa kepada kedua orang tuaku tercinta, **Mohd. Yamin dan Siti Lunah**, yang dengan penuh kesabaran, kasih sayang, doa, serta dukungan yang tiada henti, selalu menjadi sumber kekuatan dalam setiap langkah perjalanan ini.
2. Ibu **Dr. Dian Cahyawati, S.Si., M.Si.**, selaku Ketua Jurusan Matematika, dan Ibu **Des Alwine Zayanti, M.Si.**, selaku Sekretaris Jurusan Matematika FMIPA Universitas Sriwijaya sekaligus Dosen Pembimbing Akademik penulis, yang telah memberikan arahan dan bimbingan kepada penulis selama masa perkuliahan dan dalam pengurusan berbagai urusan akademik.
3. Bapak **Drs. Ali Amran, M.T.**, selaku Dosen Pembimbing Utama, dan Bapak **Dr. Bambang Suprihatin, S.Si., M.Si.**, selaku Dosen Pembimbing Kedua, yang telah bersedia meluangkan waktu, tenaga, dan pikiran untuk

memberikan bimbingan, arahan, serta motivasi kepada penulis selama proses penyusunan skripsi ini.

4. Ibu **Dr. Anita Desiani, S.Si., M.Kom.**, dan Ibu **Irmeilyana, S.Si., M.Si.**, selaku Dosen Pembahas yang telah memberikan masukan, kritik, dan saran berharga, serta mendampingi penulis dengan bimbingan yang tulus selama perjalanan penyusunan skripsi ini. Terima kasih atas semua ilmu, waktu, dan perhatian yang begitu berarti.
5. **Seluruh Dosen di Jurusan Matematika FMIPA Universitas Sriwijaya** yang telah memberikan ilmu, nasihat, motivasi, serta bimbingan yang berharga kepada penulis selama masa perkuliahan.
6. Bapak **Irwansyah** selaku Admin dan Ibu **Hamidah** selaku Pegawai Tata Usaha Jurusan Matematika FMIPA Universitas Sriwijaya yang telah banyak membantu penulis dalam berbagai keperluan akademik selama masa perkuliahan.
7. Kepada **Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya beserta jajarannya**, terima kasih atas segala fasilitas, dukungan, dan kesempatan yang telah diberikan kepada penulis untuk belajar, berkembang, dan meraih berbagai pengalaman berharga selama masa perkuliahan.
8. Abang dan adikku tercinta, **Mukhammad Syamsul Bachtiar** dan **Damara Bai Haki**, yang selalu hadir di saat penulis membutuhkan bantuan. Terima kasih atas segala bentuk dukungan moril dan materiil, serta doa yang telah menjadi kekuatan bagi penulis selama menjalani perjalanan ini.

9. Kepada **seluruh teman-teman Minat Komputasi angkatan 2018, 2019, 2020, 2021, 2022, dan 2023**, terima kasih atas kebersamaan, bantuan, dan ilmu yang telah kalian bagikan sepanjang perjalanan penyusunan skripsi ini. Juga kepada **seluruh Keluarga Mahasiswa Jurusan Matematika**, terima kasih telah menjadi bagian penting yang mewarnai perjalanan, perjuangan, dan kenangan indah penulis selama masa perkuliahan.
10. **BPH Himastik Laskaria dan Sinergi Cita, teman-teman organisasi seperjuangan, serta tim perlombaan**, terima kasih atas kebersamaan, semangat, serta pengalaman berharga yang telah menemani dan mewarnai perjalanan penulis selama masa perkuliahan.
11. Semua pihak yang tidak dapat penulis sebutkan satu per satu. Semoga segala kebaikan yang diberikan mendapatkan balasan terbaik dari Allah.
12. Terima kasih yang sebesar-besarnya penulis sampaikan kepada diri sendiri, yang telah berjuang dan tidak menyerah di segala keterbatasan dan rintangan. Untuk setiap doa, usaha tanpa henti, dan keyakinan yang terus dijaga hingga akhirnya dapat sampai di titik ini. Terima kasih telah percaya bahwa setiap langkah kecil membawa arti besar dalam perjalanan ini.

Semoga skripsi ini bermanfaat bagi mahasiswa/mahasiswi Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya dan seluruh pihak yang membutuhkan.

Indralaya, Juni 2025

VOLUMETRIC U-NET AND TRANSFORMER ARCHITECTURE FOR BRAIN TUMOR SEGMENTATION IN MAGNETIC RESONANCE IMAGING IMAGES

By:

FAISHAL FITRA RAMADHAN

080112821260606

ABSTRACT

Brain tumors are abnormal tissue growths within the brain that can lead to death. The components of a brain tumor can be classified into background, enhancing tumor, peritumoral edema, and non-enhancing tumor, which are identified in three-dimensional Magnetic Resonance Imaging (MRI) scans. The separation of these components can be achieved through automatic segmentation. This study proposes a combination of Vision Transformer (ViT) and Volumetric U-Net architectures for the segmentation of brain tumor components in MRI images. ViT is utilized in the encoder to capture global spatial relationships, while the decoder maintains the Volumetric U-Net structure to preserve local spatial details. The performance of the proposed architecture achieved accuracy, sensitivity, specificity, IoU, and f1-score values of 98.78%, 80.33%, 97.01%, 74.6%, and 83.3%, respectively. These results indicate a good performance in brain tumor segmentation from MRI images. At the label level, the model achieved accuracy, sensitivity, specificity, IoU, and f1-score values ranging from 78% to 98% for background, enhancing tumor, and peritumoral edema. However, the performance for the non-enhancing tumor label was relatively low, with an accuracy of 98.6%, sensitivity of 43.6%, specificity of 99.9%, IoU of 43.2%, and f1-score of 60.2%. This lower performance is attributed to the relatively small feature size and unclear boundaries of the non-enhancing tumor region. Based on these findings, future studies are encouraged to explore new approaches capable of better detecting small regions in 3D brain tumor segmentation.

Keywords: Brain tumor segmentation, MRI, Volumetric U-Net, Vision Transformer, Deep Learning

ARSITEKTUR *VOLUMETRIC U-NET* DAN *TRANSFORMER* UNTUK SEGMENTASI TUMOR OTAK PADA CITRA HASIL *MAGNETIC RESONANCE IMAGING*

Oleh:

FAISHAL FITRA RAMADHAN

08011282126060

ABSTRAK

Tumor otak merupakan pertumbuhan jaringan abnormal dalam otak yang dapat menyebabkan kematian. Bagian-bagian tumor otak dapat dipisahkan menjadi *background*, *enhancing tumor*, *peritumoral edema*, dan *non-enhancing tumor* yang terdapat pada citra *Magnetic Resonance Imaging* (MRI) dalam bentuk tiga dimensi. Pemisahan sel pada otak dapat dilakukan menggunakan segmentasi otomatis. Penelitian ini mengusulkan kombinasi arsitektur *Vision Transformer* (ViT) dan *Volumetric U-Net* untuk segmentasi sel tumor otak pada citra MRI. ViT diletakkan pada bagian *encoder* dengan tujuan menangkap hubungan spasial global, sedangkan bagian *decoder* dari penelitian ini tetap mempertahankan *Volumetric U-Net* untuk mempertahankan detail spasial lokal. Hasil kinerja dari penerapan kombinasi arsitektur menunjukkan nilai akurasi, sensitivitas, spesifisitas, IoU, dan *f1-score* berturut-turut sebesar 98.78%, 80.33%, 97.01%, 74.6%, dan 83.3%. Hasil tersebut menunjukkan performa yang baik dalam segmentasi citra MRI. Hasil kinerja untuk tiap label menghasilkan nilai akurasi, sensitivitas, spesifisitas, IoU, dan *f1-score* di rentang 78-98% pada label *background*, *enhancing tumor*, dan *peritumoral edema*. Namun, hasil kinerja pada label *non-enhancing tumor* masih relatif kecil, yaitu 98.6% pada akurasi, 43.6% pada sensitivitas, 99.9% pada spesifisitas, 43.2% pada IoU, dan 60.2% pada *f1-score*. Hal ini disebabkan karena ukuran fitur yang dimiliki label tersebut relatif kecil dan memiliki tepi yang kurang jelas. Berdasarkan hal tersebut, disarankan pendekatan baru yang bisa mendeteksi bagian kecil pada segmentasi tumor otak 3D.

Kata Kunci: Segmentasi tumor otak, MRI, *Volumetric U-Net*, *Vision Transformer*, *Deep Learning*

DAFTAR ISI

LEMBAR PENGESAHAN	ii
HALAMAN PERSEMBAHAN	iii
KATA PENGANTAR.....	v
ABSTRACT	viii
ABSTRAK	ix
DAFTAR ISI.....	x
DAFTAR GAMBAR.....	xii
DAFTAR TABEL	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Pembatasan Masalah	4
1.4 Tujuan	5
1.5 Manfaat	5
BAB 2 TINJAUAN PUSTAKA.....	6
2.1 Citra MRI Otak	6
2.2 Citra Digital 3 Dimensi	7
2.3 <i>Preprocessing</i>	7
2.4 Segmentasi Citra Semantik	8
2.5 <i>Encoder Vision Transformer (ViT)</i>	8
2.6 <i>Convolutional Neural Network (CNN)</i>	18
2.7 <i>Decoder Volumetric U-Net</i>	18
2.8 <i>Loss Function: Dice Loss</i>	22
2.9 <i>Optimization Function: Adaptive Moment Estimation (Adam)</i>	23
2.10 <i>Confusion Matrix</i>	24
BAB III METODOLOGI PENELITIAN	28
3.1 Tempat	28
3.2 Waktu	28
3.3 Alat.....	28
3.4 Metode Penelitian	28
3.4.1 Pengumpulan Data	29

3.4.2	<i>Preprocessing Data</i>	29
3.4.3	Rancangan Kombinasi Arsitektur <i>Volumetric U-Net Transformer</i>	30
3.4.4	Evaluasi Kinerja Model	34
3.4.5	Analisis dan Interpretasi Hasil	35
BAB IV HASIL DAN PEMBAHASAN		36
4.1.	Deskripsi Data	36
4.2.	<i>Preprocessing Data</i>	37
4.3.	Penerapan Kombinasi Arsitektur <i>Volumetric U-Net</i> dan <i>Transformer</i>	39
4.4.	Operasi Perhitungan Manual	39
4.4.1.	Perhitungan Manual <i>Encoder</i>	40
4.4.2.	<i>Reshaping</i>	59
4.4.3.	Perhitungan Manual <i>Decoder</i>	60
4.4.4.	<i>Loss Function: Dice Loss</i>	75
4.4.5.	<i>Optimization Function: Adaptive Moment Estimation (Adam)</i>	76
4.5.	Hasil	79
4.5.1.	<i>Training</i>	79
4.5.2.	<i>Testing</i>	81
4.5.3.	Evaluasi Kinerja Model	85
4.5.4.	Analisis dan Interpretasi Hasil	89
4.6.	Pembahasan	91
BAB V KESIMPULAN DAN SARAN		93
5.1.	Kesimpulan	93
5.2.	Saran	93
DAFTAR PUSTAKA		94

DAFTAR GAMBAR

Gambar 2.1. Hasil MRI citra otak 3D: (a) sisi samping, (b) sisi sawah, (c) sisi belakang (BraTS 2020 <i>dataset</i>).....	6
Gambar 2.2. Ilustrasi arsitektur ViT	8
Gambar 2.3. <i>Transformer encoder</i>	11
Gambar 2.4. Ilustrasi <i>Multihead Self-attention</i>	13
Gambar 2.5. Ilustrasi <i>scaled dot product attention</i>	14
Gambar 2.6. Ilustrasi struktur MLP	16
Gambar 2.7. Ilustrasi proses <i>upsampling</i>	19
Gambar 2.8. Ilustrasi proses <i>convolutional layer 3D</i>	20
Gambar 2.9. Ilustrasi operasi <i>concatenate layer</i>	22
Gambar 2.10. <i>Confusion matrix</i>	25
Gambar 3.1. Diagram alir penelitian.....	29
Gambar 3.2. Ilustrasi kombinasi arsitektur <i>Volumetric U-Net</i> dan <i>Transformer</i> ..	31
Gambar 4.1. Ilustrasi input citra 3 dimensi	37
Gambar 4.3. Ilustrasi pemotongan <i>patch</i>	40
Gambar 4.4. Ilustrasi perataan <i>patch</i>	40
Gambar 4.5. Ilustrasi <i>reshaping</i>	59
Gambar 4.6. Grafik akurasi proses <i>training</i>	79
Gambar 4.7. Grafik loss proses <i>training</i>	80
Gambar 4.8. Evaluasi kinerja model	90

DAFTAR TABEL

Tabel 2.1. Kategori nilai evaluasi kinerja model	26
Tabel 4.1 Contoh data citra dan <i>ground truth</i> MRI otak.....	36
Tabel 4.2. Hasil perhitungan rata-rata setiap baris.....	46
Tabel 4. 3. Hasil perhitungan variansi setiap baris matriks	47
Tabel 4.4. Nilai bobot <i>hidden layer</i>	57
Tabel 4. 5. Nilai bobot <i>output layer</i>	58
Tabel 4.6. Perbandingan citra asli, hasil segmentasi, dan <i>ground truth</i>	81
Tabel 4. 7. <i>Confusion matrix</i> hasil segmentasi arsitektur kombinasi <i>Volumetric U-Net</i> dan <i>Transformer</i>	82
Tabel 4.8. Perbandingan hasil evaluasi kinerja dengan penelitian lain	91

BAB I

PENDAHULUAN

1.1 Latar Belakang

Tumor otak merupakan salah satu penyakit yang memiliki tingkat kematian tinggi di dunia. Berdasarkan data dari *International Agency for Research on Cancer*, setiap tahun terdapat lebih dari 126.000 kasus tumor otak di seluruh dunia dengan angka kematian melebihi 97.000 jiwa (Taufiqurrahman *et al.*, 2023). Salah satu cara untuk mengurangi angka kematian yang tinggi akibat tumor otak adalah dengan melakukan pendeteksian dini terhadap penyakit tumor otak menggunakan *Magnetic Resonance Imaging* (MRI). MRI adalah teknologi medis yang memungkinkan visualisasi struktur internal tubuh dengan hasil berupa citra tiga dimensi (3D) yang dapat dilihat dari sisi *sagital*, *axial*, dan *coronal* (Magadza & Viriri, 2021). Bagian-bagian pada tumor otak dapat dipisahkan menjadi 4, *background* yaitu area yang tidak mengandung tumor, *enhancing tumor* yaitu perubahan sirkulasi pada tumor, *peritumoral edema* yaitu pembengkakan jaringan di sekitar tumor, dan *non-enhancing tumor* yaitu bagian yang terkena tumor (Magadza & Viriri, 2021). Bagian-bagian tumor otak pada MRI dapat dipisahkan dengan melakukan segmentasi.

Segmentasi dilakukan dengan memisahkan detail fitur penting dan tidak penting pada citra (Desiani *et al.*, 2022). Segmentasi citra MRI otak biasanya dilakukan secara manual oleh para ahli. Namun, kompleksitas struktur otak membuat segmentasi manual membutuhkan waktu yang lama dan ketelitian yang tinggi, sehingga para ahli dapat mengalami kesalahan akibat berbagai faktor seperti

gangguan lingkungan, kelelahan mata, serta tekanan psikologis (Desiani *et al.*, 2022). Metode segmentasi otomatis berbasis *deep learning* yang sering digunakan untuk mengatasi keterbatasan pada segmentasi manual adalah *Convolutional Neural Network* (Trimpl *et al.*, 2022).

CNN memiliki kemampuan untuk mengekstraksi fitur secara otomatis dan efisien sehingga sangat berguna untuk mendeteksi objek dan melakukan segmentasi (Trimpl *et al.*, 2022). Salah satu arsitektur yang menggunakan CNN dan sering digunakan untuk segmentasi citra 3D adalah *Volumetric U-Net*. Arsitektur *Volumetric U-Net* merupakan adaptasi dari *U-Net* yang menggunakan konvolusi 3D untuk meningkatkan efektivitas dalam pemrosesan data volumetrik (Radiuk, 2020). Sama seperti *U-Net*, Struktur arsitektur *Volumetric U-Net* terdiri dari dua komponen utama yaitu *encoder* dan *decoder*. *Encoder* digunakan untuk mengekstraksi fitur dari citra, sedangkan *decoder* digunakan untuk mengembalikan citra ke ukuran asli berdasarkan fitur yang telah diekstrak untuk mengeluarkan hasil akhir yang sesuai (Djohar *et al.*, 2022).

Vijay *et al* (2023) melakukan segmentasi citra tumor otak menggunakan *Volumetric U-Net* dengan tambahan *residual Spatial Pyramid Pooling* dan menghasilkan nilai *f1-score* 87%. Rehman *et al* (2020) menerapkan arsitektur *Volumetric U-Net* untuk segmentasi sel tumor otak dan menghasilkan nilai *f1-score* sebesar 79%. Shaukat *et al* (2022) menerapkan arsitektur *Volumetric U-Net* untuk segmentasi citra tumor otak dan menghasilkan nilai *f1-score* sebesar 95%. Penelitian tersebut tidak mengukur evaluasi kinerja yang lain seperti akurasi, sensitivitas, dan ukuran evaluasi lainnya. Bagian *encoder* pada arsitektur

Volumetric U-Net melibatkan operasi konvolusi menggunakan kernel yang digunakan untuk menangkap informasi pada citra. Namun, kernel pada *Volumetric U-Net* kesulitan menangkap informasi spasial secara global sehingga karakteristik citra seperti bentuk, tekstur, atau warna yang serupa dapat dikenali sebagai bagian yang sama meskipun berada di posisi yang berbeda (Vasiliuk *et al.*, 2023). Arsitektur yang dapat menangkap informasi spasial secara global pada bagian yang kompleks adalah *Vision Transformer* (ViT).

ViT merupakan arsitektur berbasis *transformer* yang bekerja dengan cara membagi citra menjadi *patch-patch* yang diproses secara paralel menggunakan mekanisme *self-attention* (Li *et al.*, 2023). Mekanisme *self-attention* pada ViT mampu menangkap hubungan dari tiap bagian citra dengan menghitung pengaruh setiap *patch* terhadap *patch* lainnya dalam satu citra, sehingga lebih efektif dalam menangani informasi spasial global (Xie *et al.*, 2021). ViT awalnya digunakan untuk tugas klasifikasi citra, namun penerapannya sebagai *encoder* pada tugas segmentasi telah dilakukan di beberapa penelitian. Jia & Shu (2021) menerapkan kombinasi antara ViT dengan CNN untuk melakukan segmentasi tumor otak dan menghasilkan *f1-score* sebesar 92%. Hatamizadeh *et al* (2022) menggabungkan arsitektur ViT dengan *decoder U-Net* 2D dan menghasilkan *f1-score* sebesar 71%. Penelitian-penelitian tersebut tidak melibatkan evaluasi kinerja model lain seperti akurasi, sensitivitas, spesifisitas dan *Intersection over Union* (IoU).

Penelitian yang diusulkan membahas tentang pengembangan arsitektur dengan menggabungkan arsitektur *Volumetric U-Net* dengan ViT untuk segmentasi tumor otak pada citra hasil MRI otak. ViT digunakan sebagai *encoder* untuk

menangkap hubungan spasial global dalam citra agar model mampu memahami konteks keseluruhan dari struktur citra otak 3D, sedangkan bagian *decoder* tetap menggunakan *Volumetric U-Net* untuk mempertahankan detail kecil dan kompleks. Penelitian yang diusulkan bertujuan untuk mengekstraksi fitur segmentasi pada empat label, yaitu *background*, *peritumoral edema*, *non-enhancing tumor*, dan *enhancing tumor*. Kinerja model yang diusulkan dievaluasi menggunakan metrik akurasi, sensitivitas, spesifisitas, IoU, dan *f1-score*.

1.2 Rumusan Masalah

Rumusan masalah dari penelitian ini yaitu bagaimana menerapkan kombinasi arsitektur ViT pada *encoder* dan *Volumetric U-Net* pada *decoder* untuk mendapatkan hasil yang optimal dalam memisahkan bagian *background*, *enhancing tumor*, *peritumoral edema*, dan *non-enhancing tumor* dengan mengukur kinerja evaluasi berupa akurasi, sensitivitas, spesifisitas, IoU, dan *f1-score*.

1.3 Pembatasan Masalah

Pembatasan masalah dalam penelitian yang diusulkan antara lain yaitu sebagai berikut:

1. Penelitian yang diusulkan membahas segmentasi citra MRI tumor otak menggunakan empat label, yaitu *background*, *peritumoral edema*, *non-enhancing tumor*, dan *enhancing tumor*.
2. Kinerja evaluasi yang digunakan untuk mengukur hasil segmentasi yaitu akurasi, sensitivitas, spesifisitas, IoU, dan *f1-score*.

1.4 Tujuan

Tujuan dari penelitian yang diusulkan yaitu untuk mendapatkan hasil yang optimal dari penerapan kombinasi arsitektur ViT sebagai *encoder* dan *Volumetric U-Net* sebagai *decoder* untuk melakukan segmentasi tumor otak dengan mengukur akurasi, sensitivitas, spesifisitas, IoU, dan *f1-score*.

1.5 Manfaat

Manfaat dari penelitian yang diusulkan adalah:

1. Memberikan alternatif arsitektur yang dapat digunakan untuk melakukan segmentasi citra medis 3D.
2. Hasil yang diperoleh dapat digunakan dalam segmentasi otomatis untuk mendukung diagnosis di bidang kedokteran.

DAFTAR PUSTAKA

- Aytekin, A. (2021). Comparative analysis of the normalization techniques in the context of MCDM problems. *Decision Making: Applications in Management and Engineering*, 4(2), 1–25. <https://doi.org/10.31181/dmame210402001a>
- Bai, Y. (2022). ReLU-function and derived function review. *SHS Web of Conferences*, 144, 1–5. <https://doi.org/10.1051/shsconf/202214402006>
- Bashmal, L., Bazi, Y., & Al Rahhal, M. (2021). Deep Vision Transformers for remote sensing scene classification. *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2815–2818. <https://doi.org/10.1109/IGARSS47720.2021.9553684>
- Bhandari, A., Koppen, J., & Agzarian, M. (2020). Convolutional Neural Networks for brain tumour segmentation. *Insights into Imaging*, 11(1), 77. <https://doi.org/10.1186/s13244-020-00869-4>
- Bracci, S., & de Beeck, H. P. O. (2023). Understanding human object vision: a picture is worth a thousand representations. *Annual Review of Psychology*, 74, 113–135. <https://doi.org/10.1146/annurev-psych-032720-041031>
- Chen, D., O’Bray, L., & Borgwardt, K. M. (2022). Structure-aware Transformer for graph representation learning. *International Conference on Machine Learning*, 3469–3489. <https://proceedings.mlr.press/v162/chen22r.html>
- Cingolani, S., Gallo, M., & Tanaka, K. (2021). Normalized solutions for fractional nonlinear scalar field equations via lagrangian formulation. *Nonlinearity*, 34, 4017–4056. <https://doi.org/10.1088/1361-6544/ac0166>
- Cortiñas-Lorenzo, B., & Pérez-González, F. (2020). Adam and the ants: on the influence of the optimization algorithm on the detectability of DNN watermarks. *Entropy*, 22(12), 1379. <https://doi.org/10.3390/e22121379>
- Daimary, D., Bora, M. B., Amitab, K., & Kandar, D. (2020). Brain tumor segmentation from MRI images using hybrid Convolutional Neural Networks. *Procedia Computer Science*, 167, 2419–2428. <https://doi.org/10.1016/j.procs.2020.03.295>
- Davis, J., & Frank, L. (2022). Revisiting batch normalization. *European Conference on Computer Vision*, 212–228. https://doi.org/10.1007/978-3-031-19803-8_13
- 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. *2022 International Conference on Informatics, Multimedia, Cyber*

- and Information System (ICIMCIS), 423–428.
<https://doi.org/10.1109/ICIMCIS56303.2022.10017519>
- Desiani, A., Suprihatin, B., Yahdin, S., Putri, A. I., & Husein, F. R. (2021). Bi-path architecture of CNN segmentation and classification method for cervical cancer disorders based on pap-smear images. *IAENG International Journal of Computer Science*, 48(3).
https://www.iaeng.org/IJCS/issues_v48/issue_3/IJCS_48_3_37.pdf
- Djohar, M. A., Desiani, A., Amran, A., Yahdin, S., Putri, D. L. D., Zayanti, D. A., & Dewi, N. R. (2022). Liver Segmentation Using Convolutional Neural Network Method with U-Net Architecture. *Journal Of Informatics And Telecommunication Engineering*, 6(1), 221–234.
- Fang, Y., Huang, H., Yang, W., Xu, X., Jiang, W., & Lai, X. (2022). Nonlocal convolutional block attention module VNet for gliomas automatic segmentation. *International Journal of Imaging Systems and Technology*, 32(2), 528–543. <https://doi.org/10.1002/ima.22639>
- Fei, Z., Wu, Z., Xiao, Y., Ma, J., & He, W. (2020). A new short-arc fitting method with high precision using adam optimization algorithm. *Optik*, 212, 1–12.
<https://doi.org/10.1016/j.ijleo.2020.164788>
- Foroozesh, J., Khosravani, A., Mohsenzadeh, A., & Mesbahi, A. H. (2013). Application of Artificial Intelligence (AI) modeling in kinetics of methane hydrate growth. *American Journal of Analytical Chemistry*, 4(11), 616–622.
<https://doi.org/10.4236/ajac.2013.411073>
- Ghafari, S., Tarnik, M. G., & Yazdi, H. S. (2021). Robustness of Convolutional Neural Network models in hyperspectral noisy datasets with loss functions. *Computers & Electrical Engineering*, 90, 107009.
<https://doi.org/10.1016/j.compeleceng.2021.107009>
- Hafiz, A. M., & Bhat, G. M. (2020). A survey on instance segmentation: state of the art. *International Journal of Multimedia Information Retrieval*, 9, 171–189. <https://doi.org/10.1007/s13735-020-00195-x>
- Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y., Yang, Z., Zhang, Y., & Tao, D. (2022). A survey on Vision Transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1), 87–110. <https://doi.org/10.1109/TPAMI.2022.3152247>
- Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., Roth, H. R., & Xu, D. (2022). UNETR: Transformers for 3D medical image segmentation. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 574–584.
<https://doi.org/10.48550/arXiv.2103.10504>

- Huang, L., Zhu, E., Chen, L., Wang, Z., Chai, S., & Zhang, B. (2022a). A transformer-based generative adversarial network for brain tumor segmentation. *Frontiers in Neuroscience*, 16, 1054948.
- Huang, L., Zhu, E., Chen, L., Wang, Z., Chai, S., & Zhang, B. (2022b). A transformer-based generative adversarial network for brain tumor segmentation. *Frontiers in Neuroscience*, 16. <https://doi.org/10.3389/fnins.2022.1054948>
- Javidi, B., Hua, H., Stern, A., Martinez, M., Matobe, O., & Wetzstein, G. (2022). Focus issue introduction: 3D image acquisition and display: technology, perception and applications. *Optics Express*, 30(3), 4655–4658. <https://doi.org/10.1364/OE.539459>
- Jia, Q., & Shu, H. (2021). BiTr-Unet: a CNN-Transformer combined network for MRI brain tumor segmentation. *International MICCAI Brainlesion Workshop*, 3–14. https://doi.org/10.1007/978-3-031-09002-8_1
- Kalayeh, M. M., & Shah, M. (2020). Training faster by separating modes of variation in batch-normalized models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(6), 1483–1500. <https://doi.org/10.1109/TPAMI.2019.2895781>
- Kim, J.-C., & Chung, K. (2020). Hybrid multi-modal deep learning using collaborative concat layer in health bigdata. *IEEE Access*, 8, 192469–192480. <https://doi.org/10.1109/ACCESS.2020.3031762>
- Krstinić, D., Serić, L., & Slapnicar, I. (2023). Comments on “MLCM: multi-label confusion matrix.” *IEEE Access*, 11, 40692–40697. <https://doi.org/10.1109/ACCESS.2023.3267672>
- Kumar, S., Kumar, A., & Lee, D. G. (2022). Semantic segmentation of UAV images based on Transformer Framework with context information. *Mathematics*, 10(24), 4735. <https://doi.org/10.3390/math10244735>
- Latif, U., Shahid, A. R., Raza, B., Ziauddin, S., & Khan, M. A. (2021). An end-to-end brain tumor segmentation system using multi-inception-UNET. *International Journal of Imaging Systems and Technology*, 31(4), 1803–1816. <https://doi.org/10.1002/ima.22585>
- Lee, M. (2023). Mathematical analysis and performance evaluation of the GELU activation function in deep learning. *Journal of Mathematics*, 2023(1), 4229924. <https://doi.org/10.1155/2023/4229924>
- Li, Y., Wang, J., Dai, X., Wang, L., Yeh, C.-C. M., Zheng, Y., Zhang, W., & Ma, K.-L. (2023). How does attention work in Vision Transformers? A visual analytics attempt. *IEEE Transactions on Visualization and Computer*

- Graphics*, 29(6), 2888–2900. <https://doi.org/10.1109/TVCG.2023.3261935>
- Liu, Y.-C., Shahid, M., Sarapugdi, W., Lin, Y.-X., Chen, J.-C., & Hua, K.-L. (2021). Cascaded atrous dual Attention U-Net for tumor segmentation. *Multimedia Tools and Applications*, 80, 30007–30031. <https://doi.org/10.1007/s11042-020-10078-2>
- Magadza, T., & Viriri, S. (2021). Deep learning for brain tumor segmentation: a survey of state of the art. *Journal of Imaging*, 7(2), 19. <https://doi.org/10.3390/jimaging7020019>
- Mishra, S., Vanli, O. A., Huffer, F. W., & Jung, S. (2016). Regularized discriminant analysis for multi-sensor decision fusion and damage detection with lamb waves. *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2016*, 9803, 728–741. <https://doi.org/10.1117/12.2217959>
- Mzoughi, H., Njeh, I., Wali, A., Slima, M. Ben, BenHamida, A., Mhiri, C., & Mahfoudhe, K. Ben. (2020). Deep multi-scale 3D Convolutional Neural Network (CNN) for MRI gliomas brain tumor classification. *Journal of Digital Imaging*, 33, 903–915. <https://doi.org/10.1007/s10278-020-00347-9>
- Nguyen, X.-B., Lee, G., Kim, S., & Yang, H.-J. (2020). Self-supervised learning based on spatial awareness for medical image analysis. *IEEE Access*, 8, 162973–162981. <https://doi.org/10.1109/ACCESS.2020.3021469>
- 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/10.1016/j.bspc.2020.102223>
- Radiuk, P. (2020). Applying 3D U-Net architecture to the task of multi-organ segmentation in computed tomography. *Applied Computer Systems*, 25(1), 43–50. <https://doi.org/10.2478/acss-2020-0005>
- Ramzan, F., Khan, M. U. G., Iqbal, S., Saba, T., & Rehman, A. (2020). Volumetric segmentation of brain regions from MRI scans using 3D Convolutional Neural Networks. *IEEE Access*, 8, 103697–103709. <https://doi.org/10.1109/ACCESS.2020.2998901>
- Rehman, M. U., Cho, S., Kim, J. H., & Chong, K. T. (2020). Bu-net: Brain tumor segmentation using modified U-Net architecture. *Electronics*, 9(12), 2203. <https://doi.org/10.3390/electronics9122203>
- Saeed, M. U., Ali, G., Bin, W., Almotiri, S. H., AlGhamdi, M. A., Nagra, A. A., Masood, K., & Amin, R. ul. (2021). RMU-Net: A novel residual mobile U-Net model for brain tumor segmentation from MR Images. *Electronics*, 10(16), 1962. <https://doi.org/10.3390/electronics10161962>

- Saito, S., & Ueda, J. (2024). Preclinical magnetic resonance imaging and spectroscopy in the fields of radiological technology, medical physics, and radiology. *Radiological Physics and Technology*, 17, 47–59. <https://doi.org/10.1007/s12194-024-00785-y>
- Sanjaya, Y. C., Gunawan, A. A. S., & Irwansyah, E. (2020). Semantic segmentation for aerial images: a literature review. *Engineering, Mathematics and Computer Science Journal (EMACS)*, 2(3), 133–139. <https://doi.org/10.21512/emacsjournal.v2i3.6737>
- 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>
- Shaukat, Z., Farooq, Q., Tu, S., Xiao, C., & Ali, S. (2022). A state-of-the-art technique to perform cloud-based semantic segmentation using deep learning 3D U-Net architecture. *BMC Bioinformatics*, 23(1), 251. <https://doi.org/10.1186/s12859-022-04794-9>
- Shruti, P., & Rekha, R. (2023). A review of Convolutional Neural Networks, its variants and applications. *2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS)*, 31–36. <https://doi.org/10.1109/ICISCoIS56541.2023.10100412>
- Talluri, K. K., Fiedler, M.-A., & Al-Hamadi, A. (2022). Deep 3D Convolutional Neural Network for facial Micro-expression analysis from video images. *Applied Sciences*, 12(21), 11078. <https://doi.org/10.3390/app122111078>
- Taufiqurrahman, T., Rianto, S., & Setiawan, A. N. (2023). Comparative analysis of intensity modulation radiation therapy and volumetric modulation arc therapy in brain tumor cases to minimize radiation dose to organs at risk. *International Journal of Social Health*, 2(10), 718–724. <https://doi.org/10.58860/ijsh.v2i10.111>
- Tomasila, G., & Emanuel, A. W. R. (2020). MRI image processing method on brain tumors: a review. *AIP Conference Proceedings (Vol. 2296, No. 1)*. <https://doi.org/10.1063/5.0030978>
- Trimpl, M. J., Primakov, S., Lambin, P., Stride, E. P. J., Vallis, K. A., & Gooding, M. J. (2022). Beyond automatic medical image segmentation—the spectrum between fully manual and fully automatic delineation. *Physics in Medicine and Biology*, 67(12). <https://doi.org/10.1088/1361-6560/ac6d9c>
- Vasiliuk, A., Frolova, D., Belyaev, M., & Shirokikh, B. (2023). Limitations of out-of-distribution detection in 3D medical image segmentation. *Journal of Imaging*, 9(9), 191. <https://doi.org/10.3390/jimaging9090191>

- Vijay, S., Guhan, T., Srinivasan, K., Vincent, P. M. D. R., & Chang, C. Y. (2023). MRI brain tumor segmentation using residual Spatial Pyramid Pooling-powered 3D U-Net. *Frontiers in Public Health*, 11(1). <https://doi.org/10.3389/fpubh.2023.1091850>
- Vujović, Ž. (2021). Classification model evaluation metrics. *International Journal of Advanced Computer Science and Applications*, 12(6), 599–606. <https://doi.org/10.14569/IJACSA.2021.0120670>
- Xie, Y., Zhang, J., Shen, C., & Xia, Y. (2021). CoTr: efficiently bridging CNN and Transformer for 3D medical image segmentation. *Medical Image Computing and Computer Assisted Intervention 2021: 24th International Conference*, 171–180. https://doi.org/10.1007/978-3-030-87199-4_16
- Yang, X., He, S., Wu, J., Yang, Y., Hou, Z., & Ma, S. (2023). Exploring spatial-based position encoding for image captioning. *Mathematics*, 11(21), 4550. <https://doi.org/10.3390/math11214550>
- Ye, J. C., & Ye, J. C. (2022). Artificial Neural Networks and Backpropagation. *Geometry of Deep Learning: A Signal Processing Perspective*, 91–112. https://doi.org/10.1007/978-981-16-6046-7_6
- Zhang, L., Zhang, J., Shen, P., Zhu, G., Li, P., Lu, X., Zhang, H., Shah, S. A., & Bennamoun, M. (2020). Block level skip connections across cascaded V-Net for multi-organ segmentation. *IEEE Transactions on Medical Imaging*, 39(9), 2782–2793. <https://doi.org/10.1109/TMI.2020.2975347>
- Zhang, Y., Han, Y., & Zhang, J. (2023). MAU-Net: mixed attention U-Net for MRI brain tumor segmentation. *Mathematical Bioscience and Engineering*, 20, 20510–20527. <https://doi.org/10.3934/mbe.2023907>
- Zhang, Y., Li, Z., Nan, N., & Wang, X. (2023). TranSegNet: hybrid CNN-Vision Transformers encoder for retina segmentation of optical coherence tomography. *Life*, 13(4), 976. <https://doi.org/10.3390/life13040976>
- Ziaee, A., & CCano, E. (2022). Batch layer normalization a new normalization layer for CNNs and RNNs. *Proceedings of the 6th International Conference on Advances in Artificial Intelligence*, 40–49. <https://doi.org/10.1145/3571560.3571566>