

PROPOSAL RISET INFORMATIKA
Deteksi Tumor Otak pada Citra MRI
menggunakan Vision Transformer



Oleh : Aliefian Ramadhan (22081010171)

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA
2025

1 - PERUMUSAN MASALAH PENELITIAN

1.1 Latar Belakang

Tumor otak merupakan salah satu penyakit yang memiliki tingkat kematian tinggi dan memerlukan diagnosis dini yang akurat. Dalam beberapa dekade terakhir, penggunaan teknologi pengolahan citra medis berbasis Deep Learning telah menunjukkan hasil signifikan dalam mendeteksi berbagai jenis kelainan otak dari citra MRI (Magnetic Resonance Imaging). Namun, sebagian besar metode konvensional masih bergantung pada Convolutional Neural Network (CNN) yang memiliki keterbatasan dalam menangkap hubungan global antar piksel pada citra (Dosovitskiy et al., 2021).

Vision Transformer (ViT) hadir sebagai pendekatan baru yang mengadaptasi arsitektur Transformer dari ranah Natural Language Processing (NLP) ke bidang visi komputer. ViT mampu memahami konteks global dari citra melalui mekanisme self-attention, sehingga menghasilkan representasi fitur yang lebih kaya dibandingkan CNN tradisional (Khan et al., 2022).

Dalam konteks deteksi tumor otak, kemampuan ViT untuk memahami hubungan spasial secara global menjadi keunggulan utama, karena struktur otak memiliki pola kompleks yang tidak selalu dapat dipahami oleh filter konvolusional lokal. Beberapa penelitian terbaru menunjukkan bahwa ViT mampu meningkatkan akurasi deteksi tumor otak hingga 5–10% dibandingkan model CNN pada dataset seperti BraTS (Brain Tumor Segmentation) (Hatamizadeh et al., 2022).

Dengan demikian, penelitian ini bertujuan untuk menerapkan dan menganalisis performa Vision Transformer dalam mendeteksi tumor otak dari citra MRI, serta membandingkannya dengan metode berbasis CNN untuk mengetahui peningkatan akurasi dan efisiensi model.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana penerapan arsitektur Vision Transformer dalam proses deteksi tumor otak dari citra MRI?
2. Bagaimana performa Vision Transformer dibandingkan dengan metode berbasis CNN dalam hal akurasi dan efisiensi?
3. Seberapa besar pengaruh teknik augmentasi data terhadap hasil deteksi tumor otak menggunakan Vision Transformer?

1.3 Tujuan Penelitian

Penelitian ini memiliki tujuan sebagai berikut:

1. Menerapkan arsitektur Vision Transformer pada dataset MRI untuk mendeteksi keberadaan tumor otak.
2. Menganalisis dan membandingkan performa model Vision Transformer dengan CNN dalam hal akurasi dan waktu pelatihan.
3. Mengevaluasi pengaruh augmentasi data terhadap hasil deteksi tumor otak pada model Vision Transformer.

1.4 Manfaat Penelitian

Manfaat dari penelitian ini antara lain:

1. Memberikan kontribusi terhadap pengembangan metode deteksi otomatis tumor otak berbasis *Deep Learning*.
2. Menjadi referensi bagi penelitian lanjutan dalam penerapan Vision Transformer pada bidang medis, khususnya pengolahan citra MRI.
3. Menunjukkan potensi penggunaan arsitektur Transformer sebagai alternatif dari CNN untuk data medis yang kompleks.

1.5 Ruang Lingkup Penelitian

Ruang lingkup penelitian ini dibatasi agar fokus penelitian dapat tercapai secara optimal. Adapun batasan ruang lingkupnya sebagai berikut:

1. Penelitian ini hanya berfokus pada deteksi keberadaan tumor otak (tumor vs. non-tumor) menggunakan citra MRI dan tidak mencakup segmentasi area tumor secara detail.
2. Dataset yang digunakan bersumber dari dataset publik seperti Brain Tumor MRI Dataset atau BraTS, yang berisi citra MRI dengan berbagai jenis tumor otak.
3. Arsitektur model utama yang digunakan adalah Vision Transformer (ViT) dan akan dibandingkan dengan arsitektur Convolutional Neural Network (CNN) sebagai baseline.
4. Penelitian ini hanya menggunakan teknik augmentasi data dasar seperti rotasi, flipping, zoom, dan brightness adjustment untuk meningkatkan variasi data pelatihan.
5. Evaluasi model dilakukan menggunakan metrik akurasi, presisi, recall, F1-score, serta waktu pelatihan untuk menilai efisiensi dan performa sistem.
6. Penelitian ini dilakukan dalam lingkungan pemrograman berbasis Python dengan pustaka TensorFlow dan Keras, tanpa pengujian pada perangkat medis nyata.

1.6 Metodologi Penelitian (Umum)

Secara umum, metodologi penelitian ini terdiri atas beberapa tahapan utama sebagai berikut:

A. Pengumpulan Data

Dataset citra MRI tumor otak dikumpulkan dari sumber publik (misalnya BraTS atau Kaggle Brain MRI Dataset). Data kemudian diseleksi untuk memastikan kualitas dan konsistensi format citra.

B. Preprocessing Data

Citra MRI dilakukan pra-pemrosesan berupa resizing, normalisasi piksel, dan augmentasi data untuk memperbanyak variasi citra serta mengurangi overfitting selama pelatihan.

C. Pembangunan Model

Dua model utama dikembangkan:

Model Vision Transformer (ViT) sebagai model utama.

Model CNN konvensional (misalnya DenseNet atau ResNet) sebagai pembandingan.

Arsitektur ViT diimplementasikan sesuai konfigurasi dasar yang disesuaikan dengan ukuran citra dan jumlah kelas.

D. Pelatihan Model

Kedua model dilatih menggunakan dataset yang sama dengan pembagian data training, validation, dan testing. Parameter seperti batch size, learning rate, dan jumlah epoch disesuaikan berdasarkan hasil uji coba awal.

E. Evaluasi dan Analisis

Setelah pelatihan selesai, kedua model dievaluasi menggunakan metrik akurasi, presisi, recall, dan F1-score. Hasil evaluasi kemudian dibandingkan untuk menilai keunggulan Vision Transformer terhadap CNN.

2 - GAP RESEARCH ANALYSIS

2.1 Tinjauan Penelitian Terdahulu

Sukandar (2024) Klasifikasi Citra MRI Tumor Otak Menggunakan Metode Hibrida CNN-ViT
Penelitian ini dilakukan oleh Ivan Christopher Sukandar di UPN “Veteran” Jawa Timur.
Fokus: klasifikasi citra MRI otak untuk membedakan tumor vs non-tumor menggunakan metode hibrida CNN-ViT.
Hasil utama: akurasi terbaik diperoleh oleh model hibrida CNN-ViT ~ 93% rata-rata, sedangkan CNN saja ~ 90,80%, dan ViT saja ~ 84,80%.
Konfigurasi spesifik: pembagian data 80:10:10, optimizer Adam, learning rate 0.0001
Kelemahan yang bisa dicatat: walaupun hibrida meningkatkan akurasi, penelitian ini tampaknya hanya menguji satu dataset, dan belum secara mendalam membahas aspek efisiensi (waktu pelatihan, parameter), atau pengaruh augmentasi data.
Catatan: penelitian masih berskala tugas klasifikasi (tumor vs non-tumor) dan kombinasi CNN & ViT, tapi belum eksplorasi ViT murni atau analisis lengkap augmentasi.

Laksono, Harliana & Prabowo (2023) — Deteksi Tumor Otak Melalui Penerapan GLCM dan Naïve Bayes Classification
Penelitian oleh Puji Laksono, Harliana & Tito Prabowo di Universitas Nahdlatul Ulama Blitar.
Fokus: klasifikasi citra tumor otak menggunakan metode tekstur (GLCM: Gray Level Co-occurrence Matrix) dan klasifikasi Naïve Bayes.
Hasil: akurasi ~ 80% dengan dataset 253 citra (split training 80% / testing 20%). Presisi dan recall sekitar 85%.
Kelemahan: metode tradisional (fitur tekstur + Naïve Bayes) relatif sederhana dibandingkan deep learning, dan belum mengeksplorasi konteks spasial citra secara mendalam. Juga belum membandingkan beberapa arsitektur deep learning ataupun efektivitas augmentasi data.

Ardiansyah, Qodri, Al Banna & Al-Baihaqi (2025) — Implementasi Deteksi Tumor Otak Menggunakan YOLOv11 dan Flask
<ul style="list-style-type: none"> • Penelitian oleh AS Widagdo, K N Qodri, D Al Banna & M Z Al-Baihaqi
<ul style="list-style-type: none"> • Fokus: deteksi tumor otak dengan menggunakan algoritma deteksi objek YOLOv11 plus integrasi frontend menggunakan Flask. Teknik augmentasi data juga diterapkan (flip, rotasi 90°, noise) untuk memperluas dataset.
<ul style="list-style-type: none"> • Hasil: F1-score ~ 0,951 untuk empat kelas (0.902 Glioma, 0.989 Meningioma, 0.915 Pituitary, dan 0.997 Non tumour).
<ul style="list-style-type: none"> • Kelemahan: model deteksi objek (lokalisasi) berbeda dari tugas klasifikasi sederhana (tumor vs non-tumor) yang menjadi fokusmu; meskipun augmentasi data dibahas, namun arsitektur Transformer (ViT) tidak dibahas

2.2 Analisis Kesenjangan (Gap Analysis)

Dari ketiga penelitian di atas, berikut kesenjangan riset yang muncul:

1. Eksplorasi arsitektur Transformer (ViT) untuk deteksi tumor otak masih terbatas

Penelitian pertama mengombinasikan CNN + ViT, tetapi ViT murni belum dieksplorasi secara maksimal. Penelitian lainnya memakai metode tekstur/klasik atau deteksi objek (YOLO) , tidak menggunakan ViT.

Kesenjangan: kurang studi yang fokus pada ViT sebagai model utama untuk klasifikasi tumor otak.

2. Perbandingan yang kurang menyeluruh antara model (akurasi dan efisiensi)

Penelitian pertama membandingkan CNN, ViT, dan hibrida, tapi hanya akurasi yang dilaporkan; detail efisiensi (waktu pelatihan, parameter, sumber daya) kurang. Penelitian YOLO fokus pada F1-score dan deteksi, bukan perbandingan klasifikasi CNN vs Transformer.

Kesenjangan: kebutuhan untuk evaluasi performa menyeluruh (akurasi + efisiensi) antara model.

3. Augmentasi data dan pengaruhnya terhadap model Transformer belum dibahas mendalam

Penelitian YOLO melakukan augmentasi data, namun tidak fokus pada ViT. Penelitian CNN-ViT tidak secara eksplisit mendalami teknik augmentasi dan pengaruhnya terhadap ViT.

Kesenjangan: perlu analisis sistematis tentang bagaimana augmentasi mempengaruhi performa ViT.

4. **Tugas klasifikasi sederhana (tumor vs non-tumor) dengan ViT belum banyak dilakukan**

Penelitian CNN-ViT melakukannya, namun hasil untuk ViT sendiri rendah (~84.80%) dan belum dioptimalkan/analisis mendalam. Penelitian YOLO lebih ke deteksi objek multis-kelas, bukan klasifikasi biner.

Kesenjangan: penelitian spesifik untuk klasifikasi biner dengan ViT masih terbatas.

5. **Interpretabilitas dan deployment kurang dibahas untuk model ViT**

YOLO+Flask membahas integrasi pengguna, namun bukan ViT. CNN-ViT tidak menyebut interpretabilitas atau bagaimana model membuat keputusan.

Kesenjangan: perlu integrasi aspek interpretabilitas (attention maps, saliency) dan potensi deploy-able model ViT.

2.3 Solusi & Posisi Penelitian (Our Solution & Positioning)

Berdasarkan gap di atas, posisi dan solusi penelitian adalah sebagai berikut:

- Posisi penelitian: Kamu akan fokus pada penerapan arsitektur ViT murni (atau varian dengan penyesuaian untuk citra medis) untuk klasifikasi tumor otak (tumor vs non-tumor) menggunakan citra MRI. Sebagai pembandingan akan digunakan model CNN konvensional sebagai baseline.
- Solusi teknis utama:
 1. Implementasi ViT yang disesuaikan (patch size, depth, head count) untuk citra MRI kecil/menengah.
 2. Pembandingan langsung antara CNN vs ViT dalam hal akurasi *dan* efisiensi (waktu pelatihan, parameter, memori).
 3. Analisis pengaruh teknik augmentasi data (rotasi, flipping, zoom, brightness) terhadap performa ViT.

4. Tambahan: interpretabilitas model ViT – visualisasi self-attention atau heatmap untuk menjelaskan prediksi.
- Mengisi gap: Dengan cara ini, penelitianmu mengisi kekosongan riset tentang ViT murni dalam klasifikasi tumor otak, membandingkan efisiensi dan akurasi, serta memberikan analisis augmentasi dan interpretabilitas.

2.4 Kontribusi Utama Penelitian

Jika dilaksanakan seperti rencana, penelitianmu akan memberi kontribusi sebagai berikut:

1. Kontribusi ilmiah: Menyediakan bukti empiris bahwa arsitektur ViT dapat digunakan efektif untuk klasifikasi tumor otak dari citra MRI, serta bagaimana performanya dibanding CNN.
2. Kontribusi metodologis: Menghadirkan pipeline eksperimen yang reproducible (preprocessing, augmentasi, model, evaluasi efisiensi) untuk domain medis.
3. Kontribusi praktis: Memberikan rekomendasi teknis kepada peneliti/praktisi medis atau engineer pengolahan citra tentang kapan dan bagaimana menggunakan ViT vs CNN untuk tugas klasifikasi tumor otak.
4. Kontribusi aplikasi: Dengan interpretabilitas model, memberikan alat yang lebih dapat dipercaya untuk potensi adopsi klinis atau implementasi sistem pendukung keputusan medis.

3 - Metodologi Penelitian

3.1 Jenis Penelitian

Penelitian ini merupakan penelitian eksperimen kuantitatif di bidang Computer Vision yang berfokus pada pengembangan dan pengujian model Deep Learning berbasis Vision Transformer (ViT) untuk mendeteksi tumor otak dari citra MRI. Pendekatan ini digunakan untuk mengevaluasi performa ViT dalam membedakan citra MRI dengan tumor dan tanpa tumor secara akurat

3.2 Tahapan Penelitian

Secara umum, metode penelitian ini terdiri atas beberapa tahapan utama sebagai berikut:

1. Pengumpulan Data (Dataset Acquisition)

Dataset yang digunakan adalah Brain MRI Dataset yang berisi citra MRI otak dengan dua kelas: tumor dan non-tumor. Dataset ini diperoleh dari sumber publik Kaggle.

Sumber Dataset :

<https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>

Setiap citra akan berformat .jpg atau .png dengan ukuran yang bervariasi. Dataset dibagi menjadi dua bagian:

- Data latih (train set): 80% dari total data
- Data uji (test set): 20% dari total data

2. Preprocessing Data

Tahap ini bertujuan menyiapkan citra agar sesuai dengan kebutuhan model Vision Transformer. Langkah-langkah yang dilakukan meliputi:

- Resize citra menjadi ukuran standar (misalnya 224×224 piksel).
- Normalisasi piksel ke rentang [0, 1].
- Augmentasi data seperti rotasi, flipping, zoom, dan shifting untuk memperkaya variasi citra serta mencegah overfitting.
- Encoding label menjadi format kategorikal (0 untuk non-tumor, 1 untuk tumor).

3. Perancangan Model (Model Design)

Model yang digunakan dalam penelitian ini adalah Vision Transformer (ViT) dengan arsitektur dasar yang diadaptasi dari model pra-terlatih (pre-trained model) seperti ViT-B/16. Langkah perancangan:

- Menggunakan transfer learning dengan pre-trained weights dari ImageNet.

- Mengganti classification head dengan lapisan Dense berjumlah dua neuron dan fungsi aktivasi softmax.
- Menggunakan optimizer Adam dan loss function categorical cross-entropy.

4. Pelatihan Model (Training Phase)

Model dilatih menggunakan train set dengan beberapa parameter:

- Epoch: 30
- Batch size: 32
- Learning rate: 0.0001
- Early stopping berdasarkan validation accuracy

Proses pelatihan dilakukan menggunakan framework TensorFlow atau PyTorch dengan GPU acceleration agar mempercepat komputasi.

5. Pengujian dan Evaluasi Model (Evaluation)

Setelah pelatihan, model diuji menggunakan test set untuk mengukur performa. Metrik evaluasi yang digunakan antara lain:

- Accuracy
- Precision
- Recall
- F1-Score
- Confusion Matrix

Hasil evaluasi kemudian dibandingkan dengan model lain dari penelitian terdahulu (misalnya CNN-ViT hybrid, YOLOv11, dan GLCM-Naïve Bayes) untuk menilai keunggulan Vision Transformer.

6. Analisis Hasil

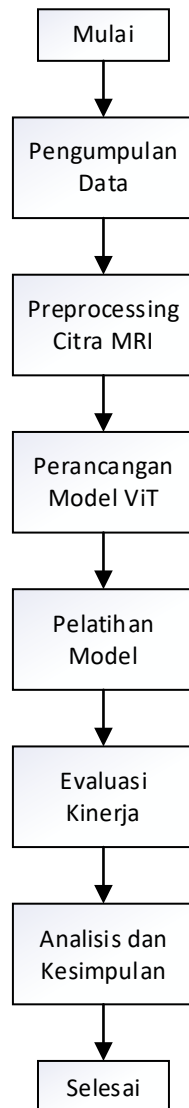
Tahap ini mencakup interpretasi hasil pengujian model serta pembahasan kelebihan dan keterbatasan metode Vision Transformer dalam mendeteksi tumor otak.

7. Kesimpulan dan Saran

Tahap akhir berisi penarikan kesimpulan terhadap hasil eksperimen serta saran untuk pengembangan lebih lanjut, seperti peningkatan dataset, fine-tuning model, atau integrasi ke aplikasi medis berbasis web.

3.3 Alur Penelitian

Secara keseluruhan, alur penelitian dapat digambarkan sebagai berikut:



4 - Mindmap Penelitian

