

# Analysis of Deep Learning Algorithms Using ConvNeXt and Vision Transformer for Brain Tumor Disease

Gilang Ekayanda<sup>1\*</sup>, Majid Rahardi<sup>2\*</sup>

\* Informatika, Universitas AMIKOM Yogyakarta

[gilangekayanda@students.amikom.ac.id](mailto:gilangekayanda@students.amikom.ac.id)<sup>1</sup>, [majid@amikom.ac.id](mailto:majid@amikom.ac.id)<sup>2</sup>

## Article Info

### Article history:

Received 2025-10-06

Revised 2025-10-31

Accepted 2025-11-08

### Keyword:

*Brain Tumor,  
Classification,  
ConvNeXt,  
Deep Learning,  
Vision Transformer.*

## ABSTRACT

This study aims to conduct a comparative analysis and identify the most effective deep learning architecture between ConvNeXt and Vision Transformer (ViT) for the automated classification of brain tumors from MRI imagery. Rapid and accurate brain tumor diagnosis is crucial; however, the manual interpretation of MRI scans is time-consuming and reliant on specialist expertise, creating an urgent need for reliable automation in brain tumor diagnosis. This research utilizes a dataset of 4,600 images, balanced between 2,513 'Brain Tumor' and 2,087 'Healthy' instances. A robust 5-Fold Cross-Validation methodology was employed to evaluate model performance, wherein the data was divided into five folds, each consisting of 920 images, ensuring every image served as both training and testing data. The quantitative results demonstrated high efficacy from both models, although ConvNeXt achieved a slight, consistent advantage. ConvNeXt obtained an accuracy of 99.13%, precision of 99.13%, recall of 99.13%, and an F1-Score of 99.13%. In comparison, the ViT model scored an accuracy of 98.13%, precision of 98.14%, recall of 98.13%, and an F1-Score of 98.13%. This quantitative superiority was validated through qualitative analysis using saliency maps, which confirmed that the models' computational attention was accurately focused on the anatomical locations of the actual tumor lesions.



This is an open access article under the [CC-BY-SA](#) license.

## I. PENDAHULUAN

Tumor otak merupakan salah satu masalah kesehatan global yang sangat serius, ditandai dengan pertumbuhan sel abnormal di dalam otak yang dapat berakibat fatal [1]. Di Indonesia sendiri, tingkat kasus dan angka kematian akibat tumor otak menunjukkan data yang signifikan, sehingga menegaskan urgensi penanganan yang cepat dan akurat [2]. Deteksi dini menjadi faktor krusial yang dapat menekan tingkat kematian, karena memungkinkan intervensi medis dilakukan pada tahap awal perkembangan penyakit [3]. Dalam praktik klinis, salah satu modalitas pencitraan medis utama yang digunakan untuk mendiagnosis tumor otak adalah Magnetic Resonance Imaging (MRI) karena kemampuannya menghasilkan gambar detail struktur jaringan lunak otak [4]. Meskipun demikian, interpretasi citra MRI secara manual oleh dokter spesialis masih menghadapi tantangan, seperti subjektivitas dan potensi kesalahan diagnosis akibat

kelelahan atau pengalaman yang bervariasi [5]. Oleh karena itu, pengembangan teknologi berbasis kecerdasan buatan sangat diperlukan untuk membantu proses diagnosis menjadi lebih objektif dan efisien.

Perkembangan teknologi pengolahan citra dan pembelajaran mesin telah membuka jalan bagi berbagai metode deteksi tumor otak otomatis. Pada tahap awal, penelitian berfokus pada teknik pemrosesan citra konvensional, seperti segmentasi menggunakan metode watershed dan thresholding untuk memisahkan area tumor dari jaringan sehat [6]. Selain itu, beberapa pendekatan juga memanfaatkan metode peningkatan kualitas citra seperti embossed untuk mempertegas kontur dan detail lesi pada citra MRI [7]. Seiring dengan perkembangan waktu, metode klasifikasi berbasis fitur mulai banyak digunakan. Salah satu pendekatan yang diterapkan adalah penggunaan Gray Level Co-occurrence Matrix (GLCM) untuk mengekstraksi fitur tekstur, yang kemudian diklasifikasikan menggunakan

algoritma Naïve Bayes [8]. Selain itu, metode Support Vector Machine (SVM) juga menjadi pilihan populer karena kemampuannya yang efektif dalam membangun model klasifikasi dengan akurasi tinggi untuk membedakan berbagai jenis tumor otak [9].

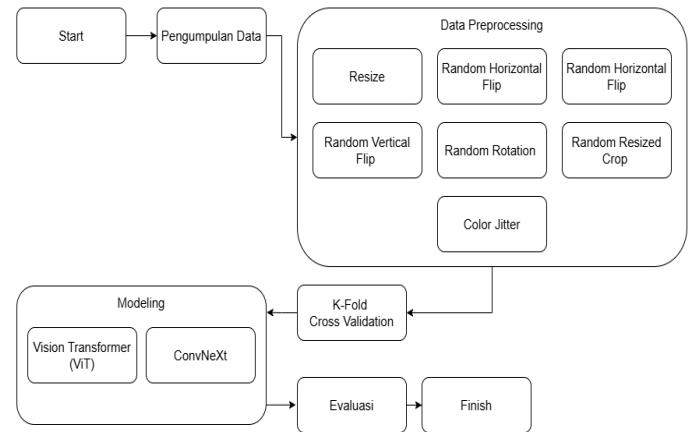
Dalam beberapa tahun terakhir, pendekatan deep learning, khususnya Convolutional Neural Network (CNN), telah terbukti memberikan kinerja yang jauh lebih baik dibandingkan metode klasik dalam analisis citra medis [10]. Beragam arsitektur CNN telah berhasil diterapkan, mulai dari VGG-16 yang dimodifikasi untuk meningkatkan ketepatan deteksi [11], hingga model yang lebih kompleks seperti ResNet-50 yang mampu mengatasi permasalahan vanishing gradient pada jaringan yang dalam [12]. Selain itu, penggunaan arsitektur modern seperti MobileNetV2 juga memperlihatkan hasil klasifikasi yang sangat potensial [13]. Di sisi lain, berkembang pula pendekatan yang tidak hanya berfokus pada klasifikasi, tetapi juga pada deteksi lokasi tumor secara akurat dengan memanfaatkan model seperti YOLOv7 [14]. Perkembangan ini menegaskan bahwa eksplorasi terhadap arsitektur deep learning yang lebih inovatif terus menjadi perhatian utama dalam upaya meningkatkan akurasi diagnosis.

Meskipun model berbasis CNN telah mendominasi penelitian dalam beberapa tahun terakhir, arsitektur baru yang disebut Vision Transformer (ViT) mulai menunjukkan kinerja yang sangat kompetitif, bahkan mampu mengungguli CNN pada beberapa tugas pengenalan gambar yang kompleks. Seiring dengan itu, evolusi arsitektur CNN juga melahirkan model-model hibrida seperti ConvNeXt, yang mengadopsi prinsip-prinsip desain dari Transformer ke dalam kerangka kerja CNN konvensional untuk mencapai performa terdepan. Namun, analisis perbandingan secara langsung antara performa arsitektur ConvNeXt dan Vision Transformer dalam konteks klasifikasi tumor otak masih sangat terbatas. Kesenjangan inilah yang menjadi motivasi utama penelitian ini untuk mengeksplorasi dan mengevaluasi kedua model canggih tersebut.

Berdasarkan latar belakang, tujuan utama dari penelitian ini adalah untuk membangun, menganalisis, dan membandingkan secara komprehensif kinerja model deep learning ConvNeXt dan Vision Transformer (ViT) dalam tugas klasifikasi tumor otak dari citra Magnetic Resonance Imaging (MRI). Manfaat yang diharapkan dari penelitian ini adalah memberikan bukti empiris mengenai keunggulan dan kelemahan masing-masing arsitektur pada kasus medis spesifik ini. Hasil penelitian diharapkan dapat menjadi referensi penting bagi para peneliti dan praktisi dalam memilih arsitektur yang paling efektif dan efisien untuk pengembangan sistem diagnosis berbantuan komputer di masa depan, sehingga dapat meningkatkan akurasi dan kecepatan diagnosis tumor otak secara signifikan. Selain itu, penelitian ini juga berpotensi memberikan kontribusi terhadap pengembangan teknologi analisis citra medis berbasis kecerdasan buatan yang lebih adaptif dan andal.

## II. METODE

Proses penelitian dirancang sesuai dengan Gambar 1, yang menampilkan tahapan penelitian. Langkah-langkah yang ditempuh untuk mencapai tujuan penelitian dijelaskan sebagai berikut.



Gambar 1. Diagram Alur Penelitian

### A. Pengumpulan Data

Pengumpulan dan persiapan dataset yang digunakan dalam penelitian ini berasal dari dataset publik berjudul "Brain MRI Images for Brain Tumor Detection" yang diperoleh dari platform Kaggle. Dataset ini terdiri dari 4.600 citra MRI otak yang telah dikategorikan ke dalam dua kelas, yaitu "Brain Tumor" sebanyak 2.513 citra dan "Healthy" sebanyak 2.087 citra, seperti ditunjukkan pada Tabel 1.

TABEL 1  
DETAIL KELAS DATASET

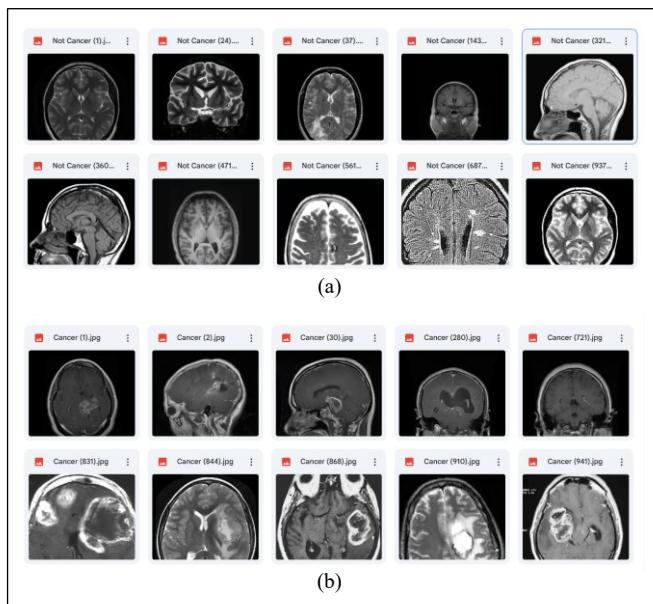
Kelas	Jumlah Citra
Brain Tumor	2.513
Healthy	2.087
Total	4.600

Setiap citra dalam dataset merepresentasikan kondisi otak manusia yang diperoleh melalui hasil pemindaian Magnetic Resonance Imaging (MRI), yaitu teknik pencitraan medis yang mampu menampilkan struktur internal otak secara detail tanpa prosedur invasif. Citra-citra tersebut terdiri atas dua kelas utama yang merepresentasikan kondisi berbeda pada struktur otak. Citra-citra terdiri dari kelas Healthy, yang menunjukkan struktur otak dalam kondisi normal tanpa adanya kelainan jaringan, dan Brain Tumor, yang menampilkan adanya indikasi pertumbuhan jaringan abnormal atau massa tumor di area tertentu pada otak. Setiap kelas memiliki karakteristik visual yang berbeda, baik dari segi tekstur, intensitas piksel, maupun bentuk morfologisnya. Perbedaan tersebut menjadi dasar penting bagi model deep learning untuk mempelajari fitur-fitur diskriminatif yang dapat digunakan dalam proses identifikasi dan klasifikasi.

Dataset ini memiliki peranan yang sangat krusial sebagai fondasi utama dalam tahap pelatihan (training) dan pengujian (testing) model yang dikembangkan. Melalui pemanfaatan

dataset ini, diharapkan model mampu mengenali pola dan ciri khas visual dari masing-masing kelas secara akurat serta menunjukkan kemampuan generalisasi yang baik terhadap data baru. Tampilan visual pada Gambar 2 memperlihatkan perbedaan karakteristik antara otak normal dan otak yang mengalami tumor, sekaligus menggambarkan kompleksitas pola yang harus dikenali oleh model dalam proses klasifikasi. Link dataset:

[https://www.kaggle.com/datasets/preetviradiya/brian-tumor-dataset?select=metadata\\_rgb\\_only.csv](https://www.kaggle.com/datasets/preetviradiya/brian-tumor-dataset?select=metadata_rgb_only.csv)



Gambar 2. (a) Gambar dari kelas Healthy, (b) Gambar dari kelas Brain Tumor

### B. Data Preprocessing

Tahap pra-pemrosesan bertujuan membersihkan, menstandarisasi, dan memperkaya variasi citra MRI agar metode klasifikasi dapat bekerja lebih efektif. Proses ini juga meningkatkan kualitas data sehingga model dapat belajar lebih optimal dari pola visual yang relevan. Tahapan yang diterapkan dijabarkan sebagai berikut:

- 1) *Preprocessing* dilakukan untuk memastikan semua citra berada pada format dan rentang nilai yang konsisten sebelum dimasukkan ke model. Langkah-langkahnya:
  - a) **Resize (Standarisasi ukuran)**  
Semua citra diubah ukurannya menjadi  $224 \times 224$  piksel agar dimensi input seragam untuk model [15] [16].
  - b) **Konversi ke format tensor**  
Citra (PIL / numpy) dikonversi menjadi tensor PyTorch melalui `ToTensor()` sehingga dapat diproses oleh framework deep learning.
  - c) **Normalisasi (Standarisasi nilai piksel)**  
Tensor citra dinormalisasi menggunakan  $\text{mean} = [0.485, 0.456, 0.406]$  dan  $\text{std} = [0.229, 0.224, 0.225]$  (nilai standar ImageNet). Normalisasi menstabilkan distribusi intensitas piksel sehingga training lebih konsisten.

2) *Augmentasi* dilakukan untuk menambah variasi pada data latih guna mengurangi overfitting dan meningkatkan kemampuan generalisasi model. Dalam penelitian ini, proses augmentasi data diterapkan melalui berbagai transformasi citra yang bertujuan memperkaya variasi visual pada dataset tanpa menambah jumlah data secara eksplisit. Dengan demikian, model diharapkan menjadi lebih robust terhadap perbedaan posisi, orientasi, dan pencahayaan citra MRI. Teknik-teknik augmentasi yang digunakan dijelaskan sebagai berikut:

#### a) Random Horizontal Flip ( $p = 0.5$ )

Transformasi ini membalik citra secara horizontal (kiri-kanan) pada 50% sampel secara acak. Tujuannya untuk menambah variasi orientasi lateral pada citra MRI tanpa mengubah makna medis, sehingga model dapat mengenali pola spasial secara lebih umum.

#### b) Random Vertical Flip ( $p = 0.5$ )

Transformasi ini membalik citra secara vertikal (atas-bawah) pada sebagian sampel secara acak. Penerapannya membantu model beradaptasi terhadap perubahan orientasi citra dan meningkatkan kemampuan generalisasi terhadap posisi fitur yang berbeda.

#### c) Random Rotation ( $\pm 30^\circ$ )

Citra dirotasi secara acak dalam rentang  $\pm 30$  derajat untuk mensimulasikan variasi sudut pemindaian MRI. Teknik ini melatih model agar tetap mampu mengenali pola tumor meskipun orientasi gambar sedikit berubah [17].

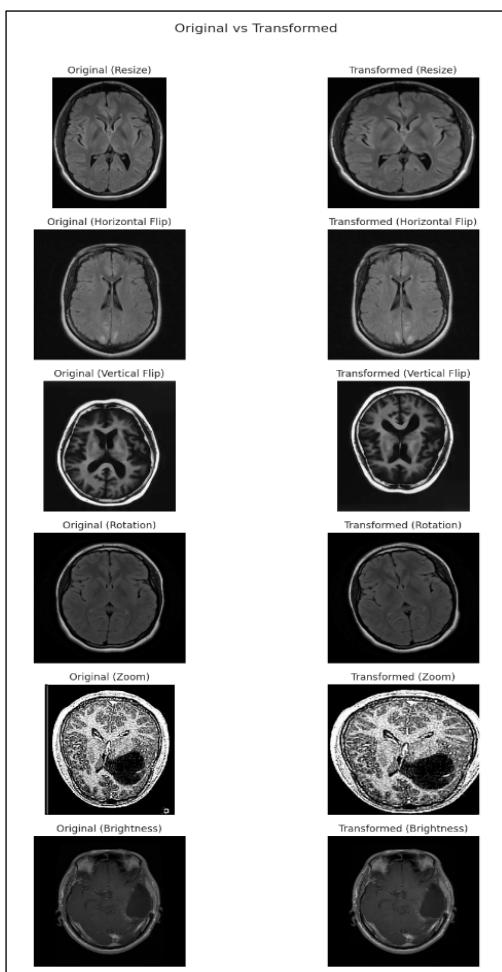
#### d) Random Resized Crop (skala $0.8\text{--}1.0$ $224 \times 224$ )

Transformasi ini memotong bagian acak dari citra dengan skala 80–100% kemudian mengubah ukurannya menjadi  $224 \times 224$  piksel. Proses ini secara adaptif menciptakan efek zoom acak yang membantu model fokus pada area berbeda dari citra otak.

#### e) Color Jitter (brightness = 0.3)

Teknik ini mengubah tingkat kecerahan citra secara acak hingga  $\pm 30\%$  untuk mensimulasikan perbedaan pencahayaan yang mungkin terjadi akibat variasi perangkat pemindaian maupun kondisi lingkungan saat pengambilan gambar. Pendekatan ini membantu meningkatkan kemampuan model dalam mengenali objek dengan tingkat pencahayaan yang tidak seragam.

Proses preprocessing dan augmentasi juga memiliki peran penting dalam memperkuat ketahanan model terhadap variasi visual yang muncul. Dengan melakukan perubahan terkontrol pada citra, seperti variasi intensitas kecerahan, rotasi, atau pergeseran kecil, model dapat belajar untuk mengenali pola yang sama meskipun tampil dalam kondisi berbeda. Pendekatan ini membantu mengoptimalkan proses pembelajaran model agar lebih adaptif dan tidak mudah bias terhadap kondisi pencahayaan tertentu. Contoh hasil penerapan augmentasi dapat dilihat pada Gambar 3 di bawah ini.



Gambar 3. Contoh tahap augmentasi

#### C. K-Fold Cross Validation

K-Fold Cross Validation merupakan metode statistikal yang digunakan untuk mengevaluasi performa model yang dikembangkan dengan cara membagi dataset menjadi beberapa subset atau lipatan (fold) berukuran sama. Teknik ini memberikan estimasi performa model yang lebih stabil dan akurat dibandingkan dengan pembagian data secara tunggal menjadi training dan testing saja [18] [19]. sehingga hasil evaluasi menjadi lebih representatif terhadap model.

Pada penelitian ini diterapkan metode 5-Fold Cross Validation, di mana total 4.600 data dibagi secara seimbang menjadi 5 subset (fold), masing-masing berjumlah 920 data [20]. Proses pelatihan dan pengujian dilakukan sebanyak 5 iterasi. Pada setiap iterasi, satu fold berfungsi sebagai data uji (testing data), sementara empat fold lainnya digunakan sebagai data latih (training data). Dengan cara ini, setiap data akan bergantian menjadi data uji satu kali dan data latih empat kali, sebagaimana ditunjukkan pada Tabel 2.

TABEL 2  
TABEL SKEMA PEMBAGIAN DATA

Fold ke	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
1	Test	Train	Train	Train	Train
2	Train	Test	Train	Train	Train
3	Train	Train	Test	Train	Train
4	Train	Train	Train	Test	Train
5	Train	Train	Train	Train	Test

Pada setiap iterasi, model dilatih menggunakan 3.680 data latih dan diuji menggunakan 920 data uji. Setelah seluruh iterasi selesai, hasil evaluasi performa dari kelima iterasi akan dirata-ratakan untuk memperoleh nilai presisi, recall, F1-score, dan akurasi secara keseluruhan. Proses ini bertujuan untuk menghasilkan estimasi performa model yang lebih general dan mengurangi risiko overfitting. Pada tahap ini, dataset yang telah diproses dibagi menjadi dua subset untuk keperluan pelatihan dan pengujian.

#### D. Modeling

Proses ini merupakan tahapan inti dalam pengembangan model. Proses ini merupakan tahapan inti dalam pengembangan model, di mana seluruh komponen pembelajaran mesin diimplementasikan untuk menghasilkan sistem klasifikasi yang optimal. Pada tahap ini, model dirancang agar mampu mengenali pola dan karakteristik visual dari citra MRI otak secara akurat, sekaligus mampu melakukan generalisasi terhadap data yang belum pernah dilihat sebelumnya.

Dua arsitektur utama yang digunakan dalam penelitian ini adalah ConvNeXt dan Vision Transformer (ViT). Kedua model ini dipilih karena mewakili dua pendekatan modern dalam pengolahan citra digital terkini, ConvNeXt dengan desain konvolusional yang telah dioptimalkan untuk kinerja tinggi, sedangkan ViT mengandalkan mekanisme self-attention yang mampu memahami hubungan spasial antarpatch secara lebih global. Kombinasi keduanya diharapkan dapat memberikan wawasan komparatif yang mendalam terhadap efektivitas arsitektur konvolusional dan transformer dalam klasifikasi citra medis.

Kedua model dilatih menggunakan dataset citra otak yang telah melalui tahap pre-processing untuk menjamin kualitas dan keseragaman data. Pelatihan dilakukan dengan memanfaatkan model ConvNeXt dan ViT yang dilatih pada dataset, kemudian disesuaikan dengan kebutuhan penelitian melalui pengaturan sejumlah hyperparameter utama seperti batch size, learning rate, num epochs, optimizer, weight decay, criterion/loss, image size, drop rate, drop path rate, num workers, scheduler, seed, dan num classes, sebagaimana ditunjukkan pada Tabel 3.

Dalam proses pelatihan, setiap model dievaluasi menggunakan skema 5-Fold Cross Validation sebagaimana dijelaskan pada bagian sebelumnya. Pendekatan ini digunakan agar performa model dapat diukur secara lebih objektif dan tidak bergantung pada satu pembagian data tertentu. Setiap fold memberikan gambaran performa model pada subset data yang berbeda, sehingga hasil akhirnya

mencerminkan kemampuan model dalam melakukan generalisasi terhadap seluruh distribusi data. Dengan demikian, nilai akurasi, presisi, recall, dan F1-score yang diperoleh merupakan rata-rata dari kelima proses pelatihan dan pengujian tersebut.

TABEL 3  
DETAIL HYPERPARAMETER PADA MODEL

Hyperparameter	ConvNeXt	Vision Transformer (ViT)
batch size	32	32
learning rate	1.00E-04	1.00E-04
num epochs	10	10
optimizer	AdamW	AdamW
weight decay	0.01	0.01
criterion / loss	CrossEntropyLoss	CrossEntropyLoss
img size	224	224
drop rate	0.1	0.1
drop path rate	0.1	0.1
num workers	2	2
scheduler	ReduceLROnPlateau	ReduceLROnPlateau
seed	42	42
num classes	2	2

Selanjutnya, hasil evaluasi dari setiap fold dikompilasi dan dirata-ratakan untuk memperoleh performa keseluruhan model yang paling representatif dan konsisten. Proses ini tidak hanya memberikan ukuran performa yang lebih stabil, tetapi juga membantu mengidentifikasi variasi hasil antar fold yang dapat menunjukkan seberapa sensitif model terhadap variasi data yang ada di seluruh dataset.

Setiap hyperparameter yang digunakan pada model ConvNeXt dan Vision Transformer (ViT) memiliki fungsi dan peran tersendiri dalam proses pelatihan model. Penjelasan mengenai masing-masing hyperparameter tersebut disajikan pada Tabel 4.

TABEL 4  
PENJELASAN MENGENAI MASING-MASING HYPERPARAMETER

Komponen dan Parameter	Penjelasan
Batch size	Merupakan jumlah citra yang diproses secara bersamaan dalam satu iterasi pelatihan
Learning rate (lr)	Menentukan besarnya langkah pembaruan bobot pada setiap iterasi, sehingga memengaruhi kecepatan konvergensi model
Num epochs	Menunjukkan banyaknya siklus pelatihan penuh terhadap seluruh dataset
Optimizer	Digunakan untuk memperbarui bobot model berdasarkan hasil perhitungan gradien
Weight decay	Berfungsi sebagai parameter regularisasi untuk mengurangi risiko overfitting dengan menekan nilai bobot yang terlalu besar
Criterion / loss	Fungsi kerugian yang digunakan untuk menghitung perbedaan antara nilai prediksi dan label sebenarnya
Img size	Menunjukkan ukuran resolusi citra masukan yang digunakan dalam proses pelatihan
Drop rate	Merupakan tingkat probabilitas penghapusan neuron secara acak selama pelatihan guna mencegah overfitting

Drop path rate	Mengatur probabilitas penghapusan jalur koneksi antar blok jaringan, yang membantu meningkatkan kemampuan generalisasi model
Num workers	Menentukan jumlah proses paralel yang digunakan dalam pemutaman data untuk mempercepat proses pelatihan
Scheduler	Digunakan untuk mengatur penyesuaian learning rate secara dinamis
Seed	Digunakan untuk memastikan proses pelatihan bersifat deterministik dan dapat direproduksi
Num classes	Menunjukkan jumlah kelas keluaran yang diprediksi oleh model

Proses training model dilakukan pada platform Google Colaboratory dengan dukungan GPU NVIDIA Tesla T4 untuk mempercepat proses komputasi dan efisiensi waktu pelatihan. Selama tahap pelatihan, kinerja model pada data validasi dipantau secara berkala guna mendeteksi kemungkinan overfitting serta melakukan penyesuaian dan optimasi hyperparameter agar diperoleh performa terbaik. Selain itu, model checkpoint disimpan secara berkala untuk memastikan bahwa hasil pelatihan dengan performa optimal tidak hilang akibat gangguan proses. Hasil akhir dari proses training berupa file model berformat .pth, yang berisi bobot terbaik dari masing-masing arsitektur, yaitu ConvNeXt dan Vision Transformer (ViT), yang kemudian digunakan untuk tahap evaluasi dan pengujian guna menilai kemampuan generalisasi model terhadap data baru secara menyeluruh serta memastikan kestabilan performa pada berbagai kondisi dan variasi dataset, termasuk pengujian terhadap data yang belum pernah dilihat sebelumnya.

#### E. Evaluasi

Setelah proses pelatihan selesai, model yang telah dikembangkan kemudian dievaluasi secara menyeluruh untuk menilai sejauh mana kemampuannya dalam melakukan generalisasi terhadap data. Proses evaluasi ini bertujuan untuk mengukur performa model secara objektif melalui berbagai metrik yang relevan. Hasil evaluasi disajikan dalam bentuk tabel dan grafik agar lebih mudah dianalisis dan dibandingkan. Metrik yang digunakan meliputi Precision, Recall, F1-Score, dan Accuracy, yang masing-masing akan dijelaskan beserta rumus serta interpretasinya pada bagian berikut.

1) *Precision*: merupakan metrik yang digunakan untuk mengukur tingkat ketepatan model dalam membuat prediksi positif. Metrik ini menunjukkan seberapa banyak dari seluruh prediksi positif yang dihasilkan oleh model benar-benar termasuk ke dalam kelas positif sebenarnya. Nilai precision yang tinggi menunjukkan bahwa model memiliki tingkat keakuratan tinggi dalam mengidentifikasi data positif, sehingga hasil prediksinya dapat lebih dipercaya, terutama dalam kasus di mana kesalahan prediksi positif memiliki konsekuensi besar. Rumusnya:

$$\text{Precision} = \frac{TP}{(TP + FP)} \quad (1)$$

Dimana:

- (TP) True Positive = Objek yang ada dan terdeteksi dengan benar  
 (FP) False Positive = Objek yang diprediksi positif oleh model tetapi sebenarnya negatif

2) *Recall*: merupakan metrik yang menilai sejauh mana model mampu mendeteksi seluruh data positif yang benar dalam suatu dataset. Dengan kata lain, recall berfokus pada kemampuan model untuk menemukan semua kasus positif tanpa melewatkannya satupun. Nilai recall yang tinggi menunjukkan bahwa model dapat mengenali sebagian besar data positif yang ada, meskipun terkadang dapat meningkatkan risiko munculnya prediksi positif palsu. Metrik ini sangat penting dalam konteks di mana kehilangan satu kasus positif dapat berdampak besar, seperti dalam deteksi penyakit atau sistem keamanan. Rumusnya:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

Dimana:

- (TP) True Positive = Objek yang ada dan terdeteksi dengan benar  
 (FN) False Negative = Objek yang ada namun tidak terdeteksi oleh model

3) *F1-Score*: Merupakan metrik yang digunakan untuk menyeimbangkan Precision dan Recall, dengan menggabungkan keduanya menjadi satu nilai yang merepresentasikan kinerja model secara keseluruhan. F1-Score dihitung menggunakan rata-rata harmonis dari dua metrik tersebut, sehingga memberikan penilaian yang lebih adil terhadap model yang memiliki ketidakseimbangan antara kemampuan mendeteksi data positif (recall) dan ketepatan memprediksi data positif (precision). Metrik ini berguna ketika data tidak seimbang antara kelas positif dan negatif, karena mampu memberikan gambaran lebih akurat terhadap performa model dalam mengatasi perbedaan distribusi kelas. Rumusnya:

$$F1 = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} = \frac{2 \cdot TP}{(2 \cdot TP + FP + FN)} \quad (3)$$

Dimana:

- (TP) True Positive = Objek yang ada dan terdeteksi dengan benar

- (FP) False Positive = Objek yang diprediksi positif oleh model tetapi sebenarnya negatif  
 (FN) False Negative = Objek yang ada namun tidak terdeteksi oleh model

4) *Accuracy*: Merupakan metrik yang digunakan untuk mengukur proporsi prediksi yang benar dari keseluruhan data, baik untuk kelas positif maupun negatif. Metrik ini menghitung seberapa banyak prediksi model yang sesuai dengan label sebenarnya dibandingkan dengan seluruh data. Accuracy sering digunakan karena memberikan gambaran umum mengenai performa model, terutama ketika distribusi data antara kelas positif dan negatif seimbang. Namun, pada kasus di mana data tidak seimbang, metrik ini dapat menyesatkan karena model mungkin tampak memiliki performa tinggi padahal hanya menguasai kelas mayoritas. Rumusnya:

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (4)$$

Dimana:

- (TP) True Positive = Objek yang ada dan terdeteksi dengan benar  
 (TN) True Negative = Objek yang tidak ada dan diprediksi tidak ada oleh model  
 (FP) False Positive = Objek yang diprediksi positif oleh model tetapi sebenarnya negatif  
 (FN) False Negative = Objek yang ada namun tidak terdeteksi oleh model

Proses evaluasi model dilakukan menggunakan metode K-Fold Cross Validation. Hasil pengujian, yang dirangkum pada Tabel 5, secara konsisten menunjukkan bahwa arsitektur ConvNeXt mengungguli Vision Transformer (ViT) pada keempat metrik evaluasi utama. Secara rinci, ConvNeXt mencatatkan rata-rata precision 99,14%, recall 99,13%, F1-score 99,13%, dan akurasi 99,13%. Sebagai perbandingan, model ViT memperoleh rata-rata precision 98,14%, recall 98,13%, F1-score 98,13%, dan akurasi 98,13%. Keunggulan ini menegaskan bahwa ConvNeXt memiliki kemampuan generalisasi yang lebih superior pada dataset dan konfigurasi pelatihan yang digunakan.

TABEL 5  
TABEL EVALUASI

K-Fold	ConvNeXt				Vision Transformer (ViT)			
	Precision	Recall	F1	Accuracy	Precision	Recall	F1	Accuracy
1	0.986957	0.986957	0.986957	0.986957	0.986095	0.985870	0.985887	0.985870
2	0.994618	0.994565	0.994562	0.994565	0.969628	0.969565	0.969578	0.969565
3	0.990222	0.990217	0.990218	0.990217	0.989135	0.989130	0.989128	0.989130
4	0.997836	0.997826	0.997826	0.997826	0.974273	0.973913	0.973927	0.973913
5	0.987314	0.986957	0.986967	0.986957	0.988044	0.988043	0.988042	0.988043
Mean	<b>0.991389</b>	<b>0.991304</b>	<b>0.991306</b>	<b>0.991304</b>	<b>0.981435</b>	<b>0.981304</b>	<b>0.981312</b>	<b>0.981304</b>
Std	<b>0.004733</b>	<b>0.0048</b>	<b>0.004797</b>	<b>0.0048</b>	<b>0.00888</b>	<b>0.008943</b>	<b>0.008938</b>	<b>0.008943</b>

Lebih jauh dari sekadar nilai rata-rata, analisis stabilitas antar-fold juga menyoroti keunggulan ConvNeXt. Model ini menunjukkan variasi (standar deviasi) yang sangat kecil dan konsisten, yang tercatat pada tabel sebesar 0,004733 untuk precision, 0,0048 untuk recall, 0,004797 untuk F1-score, dan 0,0048 untuk akurasi. Sebaliknya, ViT menampilkan variabilitas yang jauh lebih tinggi pada setiap metrik: 0,00888 (precision), 0,008943 (recall), 0,008938 (F1-score), dan 0,008943 (akurasi). Nilai standar deviasi ViT ini secara konsisten hampir dua kali lipat (sekitar 1,86 kali) lebih tinggi daripada ConvNeXt. Instabilitas ViT ini terlihat jelas jika meninjau performa per-fold: akurasi ViT mengalami penurunan signifikan pada fold 2 (96,96%) dan fold 4 (97,39%). Sementara itu, performa terendah ConvNeXt tidak pernah jatuh di bawah 98,69% (pada fold 1 dan 5). Standar deviasi yang rendah dan performa yang stabil di setiap fold ini mengindikasikan bahwa ConvNeXt tidak hanya lebih akurat, tetapi juga lebih dapat diandalkan (reliable) dan tidak sensitif terhadap cara pembagian data latih/uji.

Selisih rata-rata akurasi antara ConvNeXt (0,991304) dan ViT (0,981304) adalah tepat 1,00 poin persentase. Meskipun margin ini mungkin tampak kecil secara numerik, dalam implementasi praktis pada aplikasi kritis, perbedaannya menjadi sangat signifikan. Perbedaan ini berimplikasi langsung pada pengurangan jumlah kesalahan klasifikasi. Sebagai ilustrasi pada skala yang lebih besar, untuk setiap 10.000 sampel data, model ViT diprediksi akan menghasilkan rata-rata 187 kesalahan (tingkat kesalahan 1,87%), sedangkan model ConvNeXt hanya akan menghasilkan 87 kesalahan (tingkat kesalahan 0,87%). Ini berarti, dengan memilih ConvNeXt, potensi misklasifikasi dapat dikurangi lebih dari 53% dibandingkan dengan ViT. Dalam konteks diagnosis medis berbantuan komputer, pengurangan puluhan atau ratusan kesalahan diagnosis memiliki dampak langsung pada peningkatan akurasi diagnostik dan hasil perawatan pasien. Oleh karena itu, pemilihan arsitektur yang menawarkan mean performance lebih tinggi dan lower variance dalam hal ini ConvNeXt merupakan keputusan yang lebih aman secara teknis untuk menjamin keandalan operasional.

Terdapat beberapa hipotesis yang dapat menjelaskan mengapa ViT menunjukkan performa yang sedikit lebih rendah dan lebih variatif dalam penelitian ini. Salah satu faktor utamanya adalah kebutuhan ViT akan data latih dalam jumlah yang jauh lebih besar untuk membangun representasi fitur yang kuat, mengingat arsitektur ini tidak memiliki inductive bias konvolusional bawaan seperti ConvNeXt. Selain itu, model Transformer umumnya sangat sensitif terhadap pengaturan hyperparameter dan strategi augmentasi data. Ada kemungkinan bahwa inductive bias pada ConvNeXt memang lebih cocok untuk karakteristik dataset ini, sehingga menghasilkan generalisasi yang lebih baik.

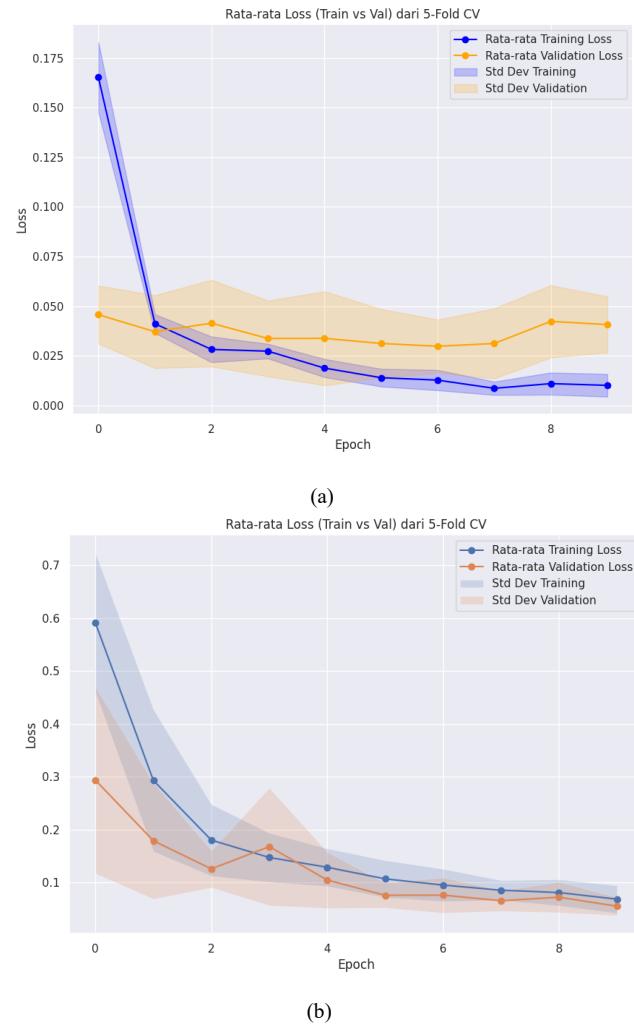
Sebagai kesimpulan, berdasarkan konfigurasi eksperimen yang dilaporkan, arsitektur ConvNeXt terbukti memberikan hasil yang lebih superior dan stabil. Hal ini terlihat dari nilai rata-rata metrik yang lebih tinggi (sekitar 99,13%-99,14%) serta variasi antar-fold yang lebih rendah (sekitar 0,47%-

0,48%). Meskipun demikian, Vision Transformer tetap merupakan arsitektur yang menjanjikan. Potensinya dapat dieksplorasi lebih lanjut melalui tuning hyperparameter yang lebih intensif, penggunaan teknik augmentasi data yang lebih agresif, atau penambahan jumlah data latih untuk meningkatkan stabilitas dan performanya.

### III. HASIL DAN PEMBAHASAN

#### A. Analisis Gap Generalisasi pada Model

Hasil pelatihan memperlihatkan pola konvergensi yang konsisten pada kedua arsitektur.



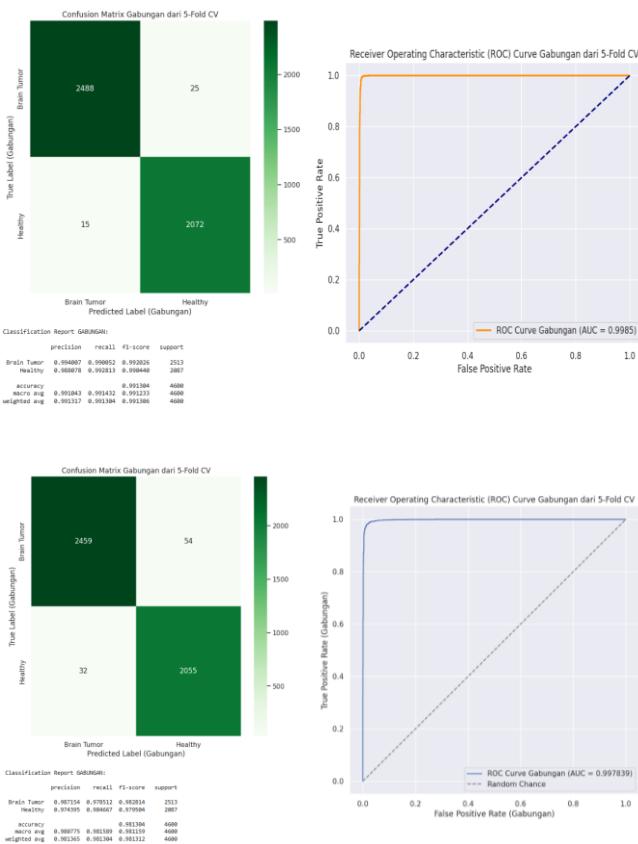
Gambar 4. (a)Gambar gabungan epoch convnext, (b) Gambar gabungan epoch Vision Transformer (ViT)

Pada fase inisialisasi pelatihan (epoch awal) terjadi penurunan nilai loss yang tajam pada kedua set (pelatihan dan validasi), yang mengindikasikan adaptasi bobot yang cepat terhadap korelasi dasar dalam dataset. Setelah penurunan awal tersebut, loss pada data pelatihan terus menurun hingga mencapai nilai yang sangat rendah, sedangkan loss validasi menurun lebih lambat, mengalami fluktuasi ringan pada

beberapa epoch awal, lalu memasuki fase stabilisasi. Pola ini menandakan bahwa kedua model berhasil mempelajari representasi fitur yang relevan sehingga penurunan loss tidak lagi bersifat eksponensial pada epoch-epoch akhir. Visualisasi dapat dilihat pada Gambar 4.

Perbedaan antara kurva training dan validation loss menunjukkan bahwa kedua arsitektur mampu menangkap pola penting dalam data pelatihan. Nilai training loss yang rendah menandakan kemampuan model mengekstraksi fitur pembeda antara tumor dan jaringan sehat, sedangkan selisih kecil dengan validation loss menunjukkan pembelajaran yang intens dan representasi yang kuat. Stabilisasi validation loss pada nilai rendah memperlihatkan performa yang robust dan konsisten, menegaskan bahwa kedua arsitektur memiliki proses pelatihan yang konvergen, stabil, dan impresif.

### B. Performa Model Akhir



Gambar 5. (a) Hasil akhir dari Model ConvNeXt, (b) Hasil akhir dari Model Vision Transformer (ViT)

Berdasarkan Gambar 5, hasil evaluasi 5-Fold Cross-Validation menunjukkan bahwa kedua model, ConvNeXt dan Vision Transformer (ViT), memiliki performa klasifikasi yang sangat tinggi. Analisis confusion matrix memperlihatkan bahwa ConvNeXt memprediksi 2488 kasus "Brain Tumor" dengan benar (True Positive), hanya 25 kasus

yang terlewat (False Negative), serta 2072 kasus "Healthy" (True Negative) dengan 15 kesalahan prediksi (False Positive). Secara keseluruhan, model ConvNeXt hanya menghasilkan 40 kesalahan klasifikasi.

Sebagai perbandingan, model ViT mencatat 2.459 prediksi "Brain Tumor" yang benar (True Positive) dan 54 kasus yang terlewat (False Negative). Untuk kelas "Healthy", ViT menghasilkan 2.055 prediksi benar (True Negative) dan 32 prediksi keliru (False Positive). Total kesalahan klasifikasi model ViT mencapai 86 kasus jumlah yang lebih dari dua kali lipat dibandingkan kesalahan yang dibuat oleh ConvNeXt.

Keunggulan performa ConvNeXt juga terkonfirmasi melalui analisis kurva Receiver Operating Characteristic (ROC). Visualisasi menunjukkan kedua model memiliki kemampuan diskriminasi yang sangat baik, dengan kurva yang mendekati sudut kiri-atas; namun ConvNeXt sedikit lebih unggul menurut nilai Area Under the Curve (AUC), yaitu 99,85% dibandingkan ViT yang memperoleh 99,78%. Nilai AUC yang nyaris sempurna pada kedua model ini menegaskan kapabilitas keduanya dalam membedakan kelas di berbagai tingkat threshold. Untuk rincian metrik performa akhir, lihat Tabel 6 di bawah ini.

TABEL 6  
TABEL PERFORMA MODEL AKHIR

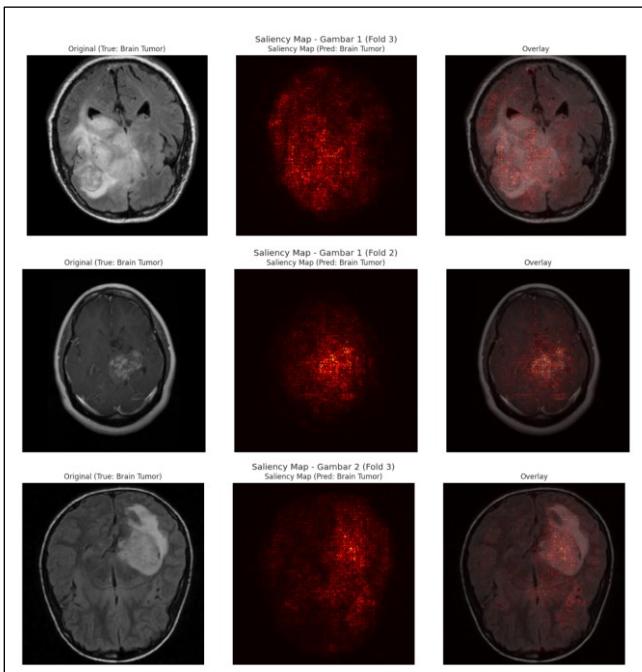
Metrik	ConvNeXt	Vision Transformer (ViT)
Precision	0.991317	0.981365
Recall	0.991304	0.981304
F1-Score	0.991306	0.981312
Accuracy	0.991304	0.981304

Berdasarkan data Tabel 6, model ConvNeXt menunjukkan keunggulan performa yang konsisten dibandingkan Vision Transformer (ViT) di semua metrik utama. ConvNeXt mencapai Akurasi sebesar 99,13%, melampaui ViT yang mencatatkan 98,13%. Keunggulan ini juga terlihat jelas pada metrik evaluasi krusial lainnya, termasuk Precision (99,13% vs 98,14%), Recall (99,13% vs 98,13%), dan F1-Score (99,13% vs 98,13%).

Laporan klasifikasi sebelumnya juga memperkuat temuan ini, mengonfirmasi bahwa keunggulan ConvNeXt tidak hanya pada akurasi keseluruhan. Model ConvNeXt juga unggul pada metrik precision, recall, dan F1-score untuk kedua kelas, menunjukkan kemampuan lebih baik dalam membedakan kondisi tersebut. Meskipun kedua arsitektur menunjukkan kinerja sangat baik, hasil validasi silang menempatkan ConvNeXt sebagai model yang lebih tangguh. Hal ini dibuktikan dengan jumlah kesalahan klasifikasi jauh lebih rendah: ConvNeXt hanya menghasilkan 25 kasus False Negative (dibandingkan 54 pada ViT) dan 15 kasus False Positive (dibandingkan 32 pada ViT). Angka kesalahan lebih minim ini menunjukkan bahwa ConvNeXt memiliki keseimbangan dan ketangguhan superior dalam mengklasifikasikan dataset yang digunakan.

### C. Feature Localization

Dalam menilai performa model deep learning pada citra medis, tidak cukup hanya melihat hasil akurasinya, tetapi juga diperlukan pemahaman mendalam mengenai interpretabilitas model. Penelitian ini menerapkan teknik lokalisasi fitur (feature localization) menggunakan Saliency Map untuk mengidentifikasi region-region pada citra Magnetic Resonance Imaging (MRI) otak yang paling berkontribusi terhadap prediksi model. Metode ini bekerja dengan menghitung gradien dari skor kelas output (dalam hal ini "Brain Tumor") terhadap piksel-piksel citra input. Nilai perubahan yang tinggi pada setiap piksel menunjukkan bahwa piksel tersebut memiliki pengaruh besar terhadap hasil prediksi model dan secara visual divisualisasikan melalui area terang pada heatmap yang menunjukkan tingkat kontribusi tertinggi terhadap keputusan model. Hasil visualisasi dari proses Saliency Map dapat dilihat pada Gambar 6.



Gambar 6. Hasil Saliency Map

Analisis visual dari hasil yang disajikan menunjukkan korelasi spasial yang kuat antara area saliensi tinggi dan lokasi anatomi lesi tumor. Panel Overlay mengilustrasikan bahwa fokus atensi model, yang ditandai oleh superimposisi heatmap (titik merah), akurat terpusat di batas massa patologis. Observasi ini konsisten di berbagai fold validasi, menunjukkan model tidak berfokus pada artefak atau jaringan otak sehat yang tidak relevan, melainkan secara efektif mengidentifikasi area patologis yang benar-benar signifikan dan bermakna dalam proses analisis.

Hasil ini memberikan validasi kualitatif yang penting, menegaskan bahwa model yang dikembangkan tidak hanya mampu menghasilkan klasifikasi yang akurat, tetapi juga mendasarkan keputusannya pada fitur-fitur yang memiliki

relevansi klinis. Kemampuan model dalam mengidentifikasi lokasi patologi secara tepat menjadi langkah krusial dalam membangun kepercayaan serta mendorong potensi penerapan sistem diagnostik berbasis AI di lingkungan medis.

### IV. KESIMPULAN

Berdasarkan data perbandingan yang disajikan, kedua arsitektur modern ConvNeXt dan Vision Transformer menunjukkan performa klasifikasi yang sangat tinggi dan mengesankan. Meskipun demikian, terdapat perbedaan kecil namun konsisten, di mana ConvNeXt menunjukkan keunggulan tipis dibandingkan Vision Transformer pada semua metrik yang dievaluasi. Keunggulan utama dan paling signifikan terlihat pada metrik akurasi, di mana ConvNeXt berhasil mencapai 99,13%, sementara Vision Transformer mencatatkan akurasi sedikit di bawahnya, yaitu 98,13%. Perbedaan ini, walaupun tampak kecil, dapat menjadi krusial dalam penerapan di lingkungan praktis.

Keunggulan ConvNeXt juga berlanjut pada metrik-metrik lain yang lebih spesifik dan mendalam. Pada metrik Precision, ConvNeXt memperoleh skor 99,13%, sedangkan Vision Transformer mencapai 98,14%, menunjukkan kemampuan sedikit lebih baik dalam meminimalkan terjadinya kesalahan positif. Demikian pula pada metrik Recall, ConvNeXt dengan nilai 99,13% mengungguli Vision Transformer yang memperoleh 98,13%, mengindikasikan kemampuan yang lebih unggul dalam mengidentifikasi seluruh kasus positif yang relevan serta meminimalkan kesalahan negatif.

Sebagai metrik yang dirancang untuk menyeimbangkan Precision dan Recall, nilai F1-Score juga menegaskan superioritas ConvNeXt dengan skor 99,13%, dibandingkan Vision Transformer yang mencatatkan 98,13%. Secara ringkas, meskipun performa kedua model ini sangat berdekatan dan keduanya berada pada tingkat yang sangat tinggi, hasil yang ditampilkan secara konsisten menunjukkan bahwa arsitektur ConvNeXt memberikan hasil yang sedikit lebih unggul di seluruh aspek evaluasi dibandingkan Vision Transformer.

Melengkapi keunggulan kuantitatif tersebut, analisis visual melalui saliency map memberikan validasi kualitatif yang bermakna. Observasi heatmap menunjukkan bahwa perhatian komputasional model terfokus pada lokasi anatomi lesi tumor yang sebenarnya, bukan pada artefak visual atau jaringan sehat di sekitarnya. Konfirmasi ini membuktikan bahwa performa klasifikasi tinggi dicapai dengan mendasarkan keputusan pada fitur relevan secara klinis, yang menjadi fondasi penting dalam membangun kepercayaan serta mendorong adopsi sistem diagnostik berbantuan kecerdasan buatan di lingkungan klinis.

### DAFTAR PUSTAKA

- [1] F. A. A. Harahap, A. N. Nafisa, E. N. D. B. Purba, and N. A. Putri, "Implementasi Algoritma Convolutional Neural Network Arsitektur Model Mobilenetv2 Dalam Klasifikasi Penyakit Tumor Otak Glioma, Pituitary Dan Meningioma," *J. Teknol. Informasi*,

- [Komputer, dan Apl. (JTIKA)], vol. 5, no. 1, pp. 53–61, 2023, doi: 10.29303/jtika.v5i1.234.
- [2] B. S. E. Dwi and D. R. I. M. Setiadi, “Deteksi Tumor Otak Dengan Metode Convolutional Neural Network,” *J. Eksplora Inform.*, vol. 13, no. 2, pp. 188–197, 2024, doi: 10.30864/eksplora.v13i2.971.
- [3] M. Kristian, S. Andryana, and A. Gunaryati, “Diagnosa Penyakit Tumor Otak Menggunakan Metode Waterfall Dan Algoritma Depth First Search,” *JIPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.)*, vol. 6, no. 1, pp. 11–24, 2021, doi: 10.29100/jipi.v6i1.1840.
- [4] T. A. (Tika) Mutiara and Q. N. (Qudsiah) Azizah, “Klasifikasi Tumor Otak Menggunakan Ekstraksi Fitur HOG dan Support Vector Machine,” *J. Khatulistiwa Inform.*, vol. 4, no. 1, pp. 45–50, 2022, [Online]. Available: <https://www.neliti.com/publications/491268/>
- [5] R. Andre, B. Wahyu, and R. Purbaningtyas, “Klasifikasi Tumor Otak Menggunakan Convolutional Neural Network Dengan Arsitektur Efficientnet-B3,” *J. IT*, vol. 11, no. 3, pp. 55–59, 2021, [Online]. Available: <https://jurnal.umj.ac.id/index.php/just-it/index>
- [6] A. Prayogi, A. C. Siregar, and R. W. S. Insani, “Deteksi Tumor Otak Menggunakan Metode Watershed dan Thresholding Pada Citra MRI,” *Jutisi J. Ilm. Tek. Inform. dan Sist. Inf.*, vol. 12, no. 3, p. 1761, 2023, doi: 10.35889/jutisi.v12i3.1688.
- [7] T. Hidayat, D. M. Dama, and K. S. D. Irmanti, “Analisis Komparatif Metode Peningkatan Kualitas Citra Digital untuk Deteksi Area Tubercoloma pada Citra MRI,” *J-Innovation*, vol. 13, no. 2, pp. 72–77, 2024, doi: 10.55600/jipa.v13i2.289.
- [8] P. Laksono, H. Harliana, and T. Prabowo, “Deteksi Tumor Otak Melalui Penerapan GLCM dan Naïve Bayes Classification,” *J. Ilm. Intech Inf. Technol. J. UMUS*, vol. 5, no. 1, pp. 41–48, 2023, doi: 10.46772/intech.v5i1.1286.
- [9] A. J. Sinulingga, D. R. Manalu, and S. Manurung, “Klasifikasi Jenis Tumor Otak Berdasarkan Citra Glioma Menggunakan Metode Support Vector Machine,” *Method. J. Tek. Inform. dan Sist. Inf.*, vol. 9, no. 2, pp. 23–25, 2023, doi: 10.46880/mtk.v9i2.1887.
- [10] A. Azhar, B. Siswoyo, D. Pratama, K. Anam, and H. Susana, “Penerapan Algoritma Convolutional Neural Network (Cnn) Untuk Diagnosa Tumor Otak,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 1797–1801, 2024, doi: 10.36040/jati.v8i2.8242.
- [11] A. B. A. Aulia and Alamsyah, “Peningkatan Hiperparameter Framework Deep Learning VGG-16 untuk Pendekstian Tumor Otak pada Teknologi MRI,” *Indones. J. Math. Nat. Sci.*, vol. 47, no. 2, pp. 99–107, 2024, [Online]. Available: <https://journal.unnes.ac.id/journals/JM/index>
- [12] M. L. Septipalan, M. S. Hibrizi, N. Latifah, R. Lina, and F. Bimantoro, “Klasifikasi Tumor Otak Menggunakan CNN Dengan Arsitektur Resnet50,” *Semin. Nas. Teknol. Sains*, vol. 3, no. 1, pp. 103–108, 2024, doi: 10.29407/stains.v3i1.4357.
- [13] M. N. M. Hakim, A. B. Nugroho, and A. E. Minarno, “Prediksi Tumor Otak Menggunakan Metode Convolutional Neural Network,” *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 17, no. 1, p. 48, 2023, doi: 10.30872/jim.v17i1.5246.
- [14] R. S. Passa, S. Nurmaini, and D. P. Rini, “Deteksi Tumor Otak Pada Magnetic Resonance Imaging Menggunakan Yolov7,” *J. Ilm. Matrik*, vol. 25, no. 2, pp. 116–121, 2023, doi: 10.33557/jurnalmatrik.v25i2.2404.
- [15] O. A. Supriadi, E. Utami, and D. Ariatmanto, “Deteksi Tumor Otak Melalui Gambar MRI Berdasarkan Vision Transformers dengan Tensorflow dan Keras,” *J. Inform. Univ. Pamulang*, vol. 8, no. 3, pp. 385–392, 2023, doi: 10.32493/informatika.v8i3.32707.
- [16] Y. I. Salsabila, H. Mustofa, and M. A. Ulimuha, “Klasifikasi Citra Tumor Otak Menggunakan Teknik Transfer Learning Pada Arsitektur Resnet-50,” *J. Inf. Syst. Informatics Comput. Issue Period*, vol. 9, no. 1, pp. 128–139, 2025, doi: 10.5236/jisicom.v9i1.1925.
- [17] D. Husen, “Klasifikasi Citra MRI Tumor Otak Menggunakan Metode Convolutional Neural Network,” *bit-Tech*, vol. 7, no. 1, pp. 143–152, 2024, doi: 10.32877/bt.v7i1.1576.
- [18] W. Wijiyanto, A. I. Pradana, S. Sopangi, and V. Atina, “Teknik K-Fold Cross Validation untuk Mengevaluasi Kinerja Mahasiswa,” *J. Algoritm.*, vol. 21, no. 1, pp. 239–248, 2024, doi: 10.33364/algoritma.v21.i1.1618.
- [19] Y. N. Fuadah, I. D. Ubaidullah, N. Ibrahim, F. F. Taliningsing, N. K. SY, and M. A. Pramuditho, “Optimasi Convolutional Neural Network dan K-Fold Cross Validation pada Sistem Klasifikasi Glaukoma,” *ELKOMIKA J. Tek. Energi Elektr. Tek. Telekomun. Tek. Elektron.*, vol. 10, no. 3, p. 728, 2022, doi: 10.26760/elkomika.v10i3.728.
- [20] H. Hafid, “Penerapan K-Fold Cross Validation untuk Menganalisis Kinerja Algoritma K-Nearest Neighbor pada Data Kasus Covid-19 di Indonesia,” *J. Math.*, vol. 6, no. 2, pp. 161–168, 2023, [Online]. Available: <http://www.ojs.unm.ac.id/jmathcos>