

Perbandingan Kinerja ResNet50, VGG16, dan MobileNetV2 untuk Klasifikasi Tumor Otak pada Citra MRI

Performance Comparison of ResNet50, VGG16, and MobileNetV2 for Brain Tumor Classification on MRI Images

^{1,2}Muhammad Bayu Kurniawan*, ²Ema Utami

^{1,2}Program Studi Magister Teknik Informatika, Universitas Amikom Yogyakarta

^{1,2}Jl. Padjajaran, Ring Road Utara, Condongcatur, Kec. Depok, Kab. Sleman, Daerah Istimewa Yogyakarta, 55283, Indonesia

*e-mail: bayu.x@students.amikom.ac.id

(received: 1 February 2025, revised: 3 February 2025, accepted: 5 February 2025)

Abstrak

Klasifikasi tumor otak menggunakan citra MRI merupakan salah satu tantangan besar dalam diagnosis medis, yang memerlukan model dengan akurasi tinggi dan efisiensi pelatihan. Penelitian ini bertujuan untuk membandingkan kinerja tiga model Convolutional Neural Networks (CNN), yaitu ResNet50, VGG16, dan MobileNetV2, dalam klasifikasi tumor otak berbasis citra MRI. Dataset yang digunakan terdiri dari empat kategori tumor otak: glioma, meningioma, pituitary, dan notumor, dengan data yang dibagi menjadi set pelatihan, validasi, dan pengujian. Masing-masing model diuji menggunakan metrik evaluasi yang meliputi akurasi, presisi, recall, F1-score, specificity, serta waktu pelatihan untuk menilai efektivitas model dalam memprediksi tumor otak dengan akurasi tinggi dan efisiensi yang optimal. Hasil eksperimen menunjukkan bahwa VGG16 memberikan performa terbaik di antara ketiga model dengan akurasi 94.93%, presisi 94.68%, dan specificity 98.33%, serta waktu pelatihan terpendek, yaitu 47.15 menit. Sementara itu, MobileNetV2 menunjukkan kinerja yang solid dengan recall 94.08%, meskipun membutuhkan waktu pelatihan lebih lama (79.53 menit). ResNet50 mencatatkan akurasi terendah (91.67%), meskipun unggul dalam presisi (91.79%), namun memiliki kekurangan pada recall (91.25%) dan specificity (97.2%). Secara keseluruhan, penelitian ini menegaskan bahwa VGG16 adalah model yang paling efisien dan efektif untuk klasifikasi tumor otak berbasis citra MRI.

Kata kunci: resnet50, VGG16, mobilenetv2, klasifikasi tumor otak, citra MRI

Abstract

Brain tumor classification using MRI images is a significant challenge in medical diagnosis, requiring models with high accuracy and efficient training. This study aims to compare the performance of three Convolutional Neural Network (CNN) models—ResNet50, VGG16, and MobileNetV2—for brain tumor classification based on MRI images. The dataset consists of four brain tumor categories: glioma, meningioma, pituitary, and no tumor, with data split into training, validation, and testing sets. Each model was evaluated using metrics including accuracy, precision, recall, F1-score, specificity, and training time to assess their effectiveness in predicting brain tumors with optimal accuracy and efficiency. Experimental results indicate that VGG16 achieved the best overall performance, with an accuracy of 94.93%, precision of 94.68%, and specificity of 98.33%, while also having the shortest training time of 47.15 minutes. MobileNetV2 demonstrated strong performance with a recall of 94.08% but required a longer training time of 79.53 minutes. ResNet50 recorded the lowest accuracy (91.67%) despite excelling in precision (91.79%), but it underperformed in recall (91.25%) and specificity (97.2%). Overall, this study confirms that VGG16 is the most efficient and effective model for MRI-based brain tumor classification.

Keywords: resnet50, VGG16, mobilenetv2, brain tumor classification, MRI image

1 Pendahuluan

Tumor otak adalah pertumbuhan sel abnormal yang terjadi di dalam otak atau sistem saraf pusat, yang bisa bersifat jinak (non-kanker) atau ganas (kanker) [1]. Tumor otak dapat muncul di berbagai bagian otak, dan gejalanya sangat bergantung pada lokasi serta ukuran tumor. Beberapa jenis tumor otak yang paling umum adalah glioma, meningioma, dan pituitary adenoma, yang masing-masing memiliki dampak yang berbeda terhadap fungsi otak dan tubuh [2]. Gejala yang sering muncul akibat tumor otak mencakup sakit kepala, gangguan penglihatan, mual, hingga perubahan mental atau kepribadian. Meskipun tumor otak jinak seringkali tumbuh lebih lambat dan dapat diobati dengan pembedahan, tumor otak ganas cenderung lebih agresif dan dapat menyebar ke bagian lain dari otak atau tubuh [3]. Deteksi dini sangat penting untuk meningkatkan tingkat keberhasilan pengobatan dan meningkatkan peluang kesembuhan pasien. Namun, tantangan utama dalam diagnosis tumor otak adalah kesulitan dalam membedakan antara tumor dan jaringan otak yang sehat, terutama pada stadium awal.

Citra *Magnetic Resonance Imaging* (MRI) merupakan metode pencitraan yang paling sering digunakan untuk mendeteksi tumor otak karena kemampuannya dalam menghasilkan gambaran rinci tentang struktur otak tanpa melibatkan radiasi [4]. Teknologi ini memungkinkan visualisasi yang jelas dari jaringan otak, sehingga memudahkan identifikasi massa atau kelainan yang mencurigakan. Meskipun MRI sangat efektif dalam mendeteksi tumor otak, analisis citra tersebut membutuhkan keahlian tinggi dan waktu yang cukup lama. Proses ini umumnya dilakukan secara manual oleh ahli radiologi yang harus memeriksa dan menganalisis setiap citra untuk mencari tanda-tanda tumor. Proses manual ini tidak hanya memakan waktu, tetapi juga sangat bergantung pada pengalaman dan ketelitian individu, yang meningkatkan risiko kesalahan diagnosis [5]. Oleh karena itu, diperlukan sistem otomatis yang dapat mempercepat proses analisis citra, mengurangi ketergantungan pada pengalaman manusia, serta meningkatkan konsistensi dan akurasi hasil diagnosis.

Dalam beberapa tahun terakhir, *deep learning* (DL) telah terbukti menjadi teknologi yang sangat efektif dalam menganalisis citra medis [6], [7], [8]. Model *Convolutional Neural Networks* (CNN), khususnya ResNet50, VGG16, dan MobileNetV2, telah banyak digunakan dalam berbagai aplikasi pengolahan citra medis, termasuk deteksi tumor otak. Ketiga model ini memiliki karakteristik arsitektur yang berbeda, yang dapat memengaruhi performa dalam klasifikasi tumor otak. ResNet50 [9], dengan arsitektur residual learning, terbukti efektif dalam mengatasi masalah vanishing gradient yang sering terjadi pada model dengan banyak lapisan, sehingga mampu menangani dataset berskala besar dan kompleks dengan baik. VGG16 [10], meskipun memiliki struktur yang lebih sederhana, tetap mampu melakukan ekstraksi fitur secara mendalam pada citra medis. Sementara itu, MobileNetV2 [11] menawarkan efisiensi komputasi yang lebih tinggi, menjadikannya pilihan yang lebih sesuai untuk aplikasi dengan sumber daya terbatas, seperti perangkat mobile atau sistem berbasis cloud.

Perbandingan kinerja antara model-model ini menjadi sangat penting karena pemilihan model yang tepat tidak hanya berpengaruh terhadap akurasi diagnosis, tetapi juga terhadap efisiensi sistem secara keseluruhan dalam lingkungan medis. Meskipun ketiga model ini telah digunakan dalam berbagai penelitian terkait pengolahan citra medis, perbandingan kinerjanya dalam klasifikasi tumor otak pada citra MRI masih terbatas. Oleh karena itu, penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja ResNet50, VGG16, dan MobileNetV2 dalam mendeteksi tumor otak menggunakan dataset dengan empat kategori tumor: pituitary, meningioma, glioma, dan notumor. Evaluasi dilakukan menggunakan metrik akurasi, presisi, recall, F1-score, specificity, dan waktu komputasi untuk mengidentifikasi model yang paling efektif dan efisien. Penelitian ini diharapkan dapat memberikan rekomendasi model terbaik untuk aplikasi medis dan berkontribusi dalam pengembangan sistem diagnosis otomatis yang lebih cepat, akurat, dan efisien, terutama untuk fasilitas dengan keterbatasan sumber daya komputasi.

2 Tinjauan Literatur

Berbagai penelitian sebelumnya telah menyoroti penerapan teknologi DL dalam klasifikasi tumor otak, dengan CNN sebagai model utama untuk analisis citra medis. Sebagai contoh, penelitian [12] membandingkan beberapa model CNN, termasuk ResNet50, Xception, EfficientNetB0, MobileNetV2, dan VGG16, dalam tugas klasifikasi tumor otak. Hasil eksperimen menunjukkan

<http://sistemasi.ftik.unisi.ac.id>

bahwa EfficientNetB0 memberikan performa terbaik dengan akurasi mencapai 97,61%, mengungguli model lainnya. Penelitian lain [13] mengembangkan pendekatan berbasis Deep CNN untuk identifikasi tumor otak, dengan model yang diusulkan mencapai akurasi 96%, presisi 93%, dan F1-Score 97%, yang menandakan efektivitasnya dalam mengklasifikasikan tumor otak. Studi lainnya [14] secara khusus mengevaluasi ResNet50 dalam deteksi tumor otak dengan menerapkan *global average pooling* untuk mengurangi *overfitting*, yang menghasilkan akurasi sebesar 97,48%.

Penelitian [15] mengusulkan pendekatan *transfer learning* (TL) berbasis *block-wise fine-tuning* untuk memanfaatkan model CNN yang telah dilatih sebelumnya. Studi ini menggunakan *5-fold cross-validation* dengan prapemrosesan minimal pada dataset berbobot T1 MRI, dan menemukan bahwa VGG19 mencapai akurasi 94,82%, VGG16 mencatatkan akurasi 94,65%, dan AlexNet mencatatkan 89,95%. Dalam penelitian lain [16], metode berbasis CNN diterapkan untuk mendeteksi tumor otak dari dataset yang berisi 3263 citra dengan berbagai kategori tumor. Model yang diusulkan mencapai akurasi 95,99%, meskipun masih terdapat peluang untuk peningkatan melalui optimasi hiperparameter. Penelitian oleh [17] mengusulkan penggunaan model berbasis DenseNet121, VGG16, MobileNetV2, dan Xception untuk klasifikasi tumor otak. Berdasarkan evaluasi kinerja menggunakan akurasi dan loss function, ditemukan bahwa DenseNet201 memberikan kinerja terbaik dengan akurasi pelatihan 97,48% dan akurasi validasi 96,42%.

Penelitian lain oleh [18] mengevaluasi sembilan model CNN dengan pendekatan TL, termasuk ResNet50, ResNet101, ResNet18, MobileNetV2, DenseNet201, Xception, InceptionResNetV2, InceptionV3, dan ShuffleNet. Dari perbandingan ini, InceptionResNetV2 menunjukkan performa terbaik dengan akurasi 98,91%. Penelitian lainnya [19] membandingkan lima model DL (VGG16, AlexNet, ResNet18, GoogleNet, dan ResNet50) dengan lima model machine learning (ML) tradisional (Naïve Bayes, KNN, Decision Trees, dan Linear Discriminant Analysis), menggunakan pendekatan *ensemble learning* berbasis *majority voting*. Hasil eksperimen menunjukkan bahwa pendekatan *ensemble* berbasis DL menghasilkan kinerja yang sangat baik, dengan akurasi rata-rata 98,16% dan AUC-ROC 98,02%. Dalam penelitian lainnya [20], klasifikasi keganasan tumor otak (meningioma, glioma, dan tumor hipofisis) dilakukan dengan menggunakan GoogLeNet, AlexNet, dan VGGNet, yang kemudian diikuti dengan *fine-tuning* dan *freezing layers*. Teknik data augmentasi diterapkan untuk mengurangi *overfitting*, dan hasilnya menunjukkan bahwa model VGG16 yang telah disesuaikan mencapai akurasi tertinggi sebesar 98,69%.

Sementara itu, penelitian [21] menggunakan metode evolusioner untuk mengidentifikasi citra MRI dalam tiga kategori tumor otak, dengan model Xception untuk ekstraksi fitur. Hasil penelitian ini menunjukkan performa yang sangat baik dengan akurasi mencapai 99,06%. Penelitian lain [22] menggunakan model TL seperti InceptionV3, ResNet101, dan DenseNet201 untuk klasifikasi tiga jenis kanker. Masing-masing model menunjukkan akurasi rata-rata sebesar 96,21%, 97%, dan 96,5%. Penelitian lain [23] mengusulkan model CNN paralel yang terdiri dari jaringan AlexNet dan VGGNet, yang memiliki struktur lapisan yang berbeda antara kedua jenis jaringan tersebut. Fitur yang diekstraksi dari kedua jaringan ini kemudian digabungkan pada satu titik dan diklasifikasikan menggunakan fungsi softmax. Model ini menunjukkan kinerja yang luar biasa, dengan akurasi sebesar 99,14% untuk klasifikasi biner dan 98,78% untuk klasifikasi multi-kelas.

Disisi lain, penelitian [24] mengadaptasi *fine-tuning* pada model yang telah dilatih sebelumnya, seperti InceptionV3 dan Xception, yang kemudian digabungkan dengan berbagai metode ML seperti Softmax, SVM, Random Forest, dan KNN, yang menghasilkan akurasi sebesar 94,34% dengan *ensemble* InceptionV3. Selain itu, penelitian [25] juga mengusulkan sebuah arsitektur CNN yang dirancang khusus untuk identifikasi tumor otak menggunakan citra MRI. Model ini menunjukkan kinerja yang lebih baik dibandingkan beberapa model lainnya, termasuk ResNet50, VGG16, dan InceptionV3, dengan mencatatkan akurasi sebesar 93,3%, recall 91,19%, AUC 98,43%, dan loss 0,25. Penelitian [26] mengkaji kinerja metode TL dengan menggunakan model VGG16, VGG19, ResNet50, Inceptionv3, DenseNet, dan MobileNet pada dataset MRI, yang menghasilkan akurasi mencapai 97%.

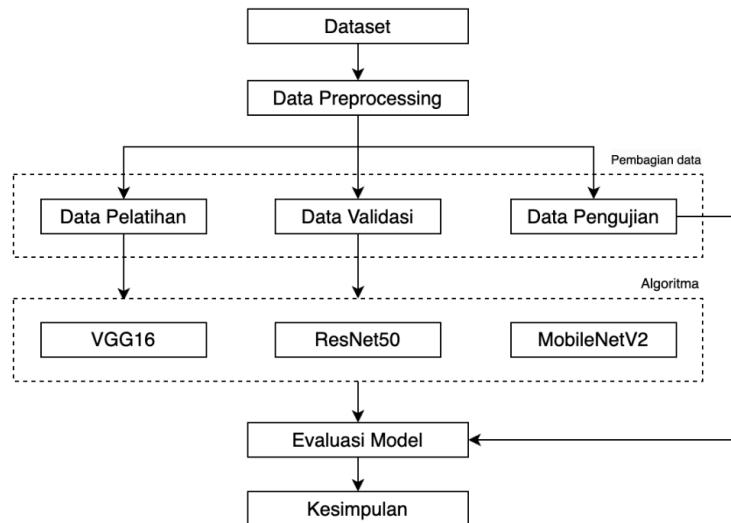
Meskipun banyak penelitian yang mengeksplorasi penggunaan CNN untuk deteksi tumor otak, perbandingan kinerja model-model seperti ResNet50, VGG16, dan MobileNetV2 dalam konteks ini masih terbatas. Selain itu, banyak studi yang lebih menekankan aspek akurasi tanpa mempertimbangkan efisiensi komputasi, yang sangat krusial untuk penerapan di lingkungan medis dengan keterbatasan sumber daya. Oleh karena itu, penelitian ini bertujuan untuk mengisi celah

<http://sistemasi.ftik.unisi.ac.id>

tersebut dengan melakukan analisis komparatif guna menentukan model yang paling optimal dalam mendeteksi tumor otak, baik dari segi akurasi maupun efisiensi.

3 Metode Penelitian

Gambar 1 menampilkan diagram alur metodologi yang diterapkan dalam penelitian ini. Tahapan utama dalam metodologi ini mencakup pengumpulan dataset, preprocessing data, pelatihan model, evaluasi model, dan analisis hasil. Penjelasan lebih lanjut mengenai setiap tahapan dijelaskan dalam bagian berikut, dimulai dari deskripsi dataset yang digunakan.



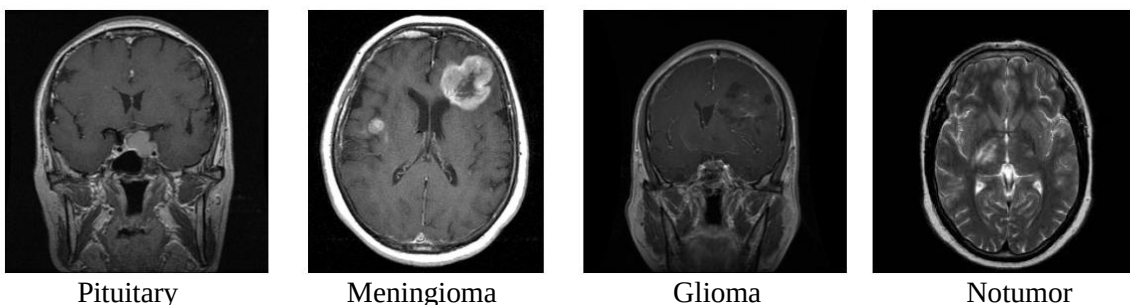
Gambar 1. Alur penelitian

3.1 Dataset

Dataset yang dipakai dalam penelitian ini bersifat publik dan diperoleh dari repositori Kaggle [27], yang berisi citra MRI untuk tugas pelatihan dan pengujian dalam klasifikasi tumor otak. Dataset ini merupakan hasil penggabungan dari tiga sumber publik, yaitu Figshare, Sartaj, dan Br35H, dan tersedia dalam format JPG dengan skala grayscale. Setiap citra dalam dataset dikategorikan ke dalam empat kelas, yaitu meningioma, glioma, pituitary, dan notumor. Dari data pelatihan, sebanyak 10% dialokasikan sebagai data validasi untuk meningkatkan kemampuan generalisasi model. Dengan demikian, penelitian ini menggunakan tiga subset data: pelatihan, validasi, dan pengujian. Tabel 1 merangkum distribusi dataset berdasarkan kelas dan subsetnya, sedangkan Gambar 2 menunjukkan contoh citra dari setiap kategori tumor dalam dataset ini.

Tabel 1 Distribusi Data

Kelas	Data Pelatihan	Data Validasi	Data Pengujian
pituitary	1311	146	300
meningioma	1205	134	306
glioma	1189	132	300
notumor	1435	160	405
Total	5140	572	1311



Gambar 2. Contoh citra dalam dataset

3.2 Prapemrosesan Data

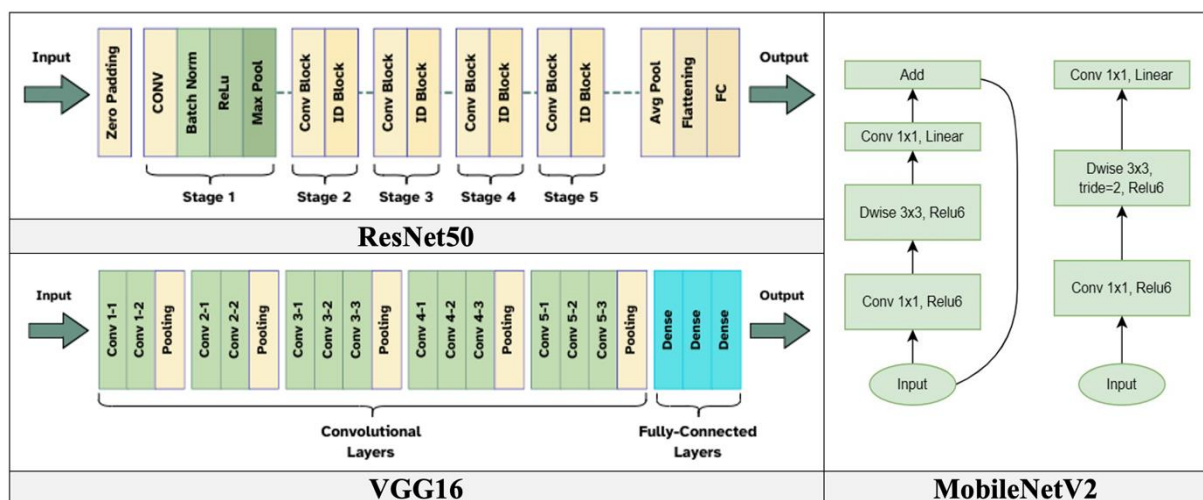
Sebelum digunakan dalam proses pelatihan, dataset MRI mengalami tahap *preprocessing* untuk memastikan kualitas dan konsistensi data. Seluruh citra diubah ukurannya menjadi 224×224 piksel agar sesuai dengan arsitektur model yang digunakan. Selanjutnya, peningkatan kontras diterapkan menggunakan *Contrast Limited Adaptive Histogram Equalization* (CLAHE) dengan ukuran grid 8×8 dan nilai batas klip 3.0 guna memperjelas struktur tumor tanpa menyebabkan distorsi berlebihan. Terakhir, dilakukan normalisasi dengan mengonversi nilai piksel ke dalam rentang 0-1 untuk meningkatkan stabilitas model selama proses pelatihan.

3.3 Arsitektur Model

Penelitian ini menggunakan tiga arsitektur CNN, yaitu ResNet50, VGG16, dan MobileNetV2, yang telah terbukti efektif dalam tugas klasifikasi citra. Model-model ini diterapkan dengan pendekatan *transfer learning*, di mana bobot awal diperoleh dari ImageNet, kemudian dilakukan penyesuaian pada *fully connected layer* untuk klasifikasi tumor otak. ResNet50 memiliki 50 lapisan dengan konsep *residual learning*, yang memungkinkan informasi melewati beberapa lapisan melalui *skip connections*, sehingga mengatasi masalah *vanishing gradient*. Model ini diawali dengan *convolutional layer* (7×7 , stride 2) dan *max-pooling layer* (3×3 , stride 2), diikuti oleh empat blok residual utama yang terdiri dari kombinasi *bottleneck layers* (1×1 , 3×3 , dan 1×1 convolutions).

VGG16 adalah model dengan 16 lapisan, yang menggunakan konvolusi berukuran kecil (3×3) dalam setiap lapisan untuk ekstraksi fitur bertahap. Arsitektur ini terdiri dari lima blok konvolusi, masing-masing berisi dua hingga tiga *convolutional layers*, diikuti oleh *max-pooling layer* (2×2 , stride 2). Setelah ekstraksi fitur, model memiliki dua *fully connected layers* dengan 4096 neuron sebelum memasuki lapisan *output*. MobileNetV2 dirancang sebagai model efisien dan ringan, menggunakan teknik *depthwise separable convolutions* untuk mengurangi jumlah parameter tanpa mengorbankan akurasi. Model ini terdiri dari *convolutional layer* awal (3×3 , stride 2) dan *batch normalization*, diikuti oleh 17 *inverted residual blocks* yang mengoptimalkan jalur informasi melalui *shortcut connections*.

Dalam penelitian ini, ketiga model dimodifikasi dengan menghapus *fully connected layer* asli dan menambahkan *Flatten layer* untuk mengubah keluaran fitur menjadi vektor satu dimensi sebelum masuk ke *fully connected layers*. Selanjutnya, ditambahkan *fully connected layer* dengan 512 neuron (ReLU) dan lapisan dropout untuk mengurangi *overfitting*. Lapisan *output* menggunakan softmax dengan 4 neuron, yang masing-masing mewakili kategori glioma, meningioma, pituitary, dan notumor. Gambar 3 menampilkan arsitektur model yang dipakai dalam penelitian ini.



Gambar 3. Arsitektur model

3.4 Parameter Pelatihan

Pada tahap pelatihan, pemilihan hyperparameter memainkan peran penting dalam menentukan performa akhir model. Daftar Hyperparameter yang diterapkan dalam penelitian ini dapat dilihat pada Tabel 2.

Tabel 2. Pengaturan hyperparameter

Hyperparameter	Nilai
Learning Rate	0.001
Batch Size	32
Optimizer	Adam
Epochs	25
Dropout Rate	0.5
Fully Connected Neurons	512

3.5 Evaluasi Model

Evaluasi model dilakukan untuk mengukur performa ResNet50, VGG16, dan MobileNetV2 dalam klasifikasi tumor otak. Pengujian dilakukan menggunakan data pengujian dengan enam metrik evaluasi: akurasi, presisi, recall, F1-score, dan specificity. Semua metrik dihitung berdasarkan Confusion Matrix, yang terdiri dari True Positive (TP), False Positive (FP), True Negative (TN), dan False Negative (FN). Persamaan (1) hingga (6) menunjukkan rumus yang digunakan untuk menghitung setiap metrik berdasarkan *confusion matrix*.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Presisi = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1 - Score = 2 \times \frac{(Recall \times Precision)}{(Recall + Precision)} \quad (4)$$

$$Specificity = \frac{TN}{TN+FP} \quad (5)$$

4 Hasil dan Pembahasan

Bagian ini menjelaskan hasil eksperimen klasifikasi tumor otak menggunakan tiga arsitektur CNN. Dataset yang digunakan telah dibagi menjadi set pelatihan, validasi, dan pengujian untuk memastikan evaluasi model dilakukan secara terpisah. Proses pelatihan dilakukan selama 25 epoch dengan *learning rate* sebesar 0.001 dan menggunakan optimizer Adam. Tabel 3 menampilkan *confusion matrix* yang dihasilkan dari eksperimen menggunakan tiga model CNN yang diuji.

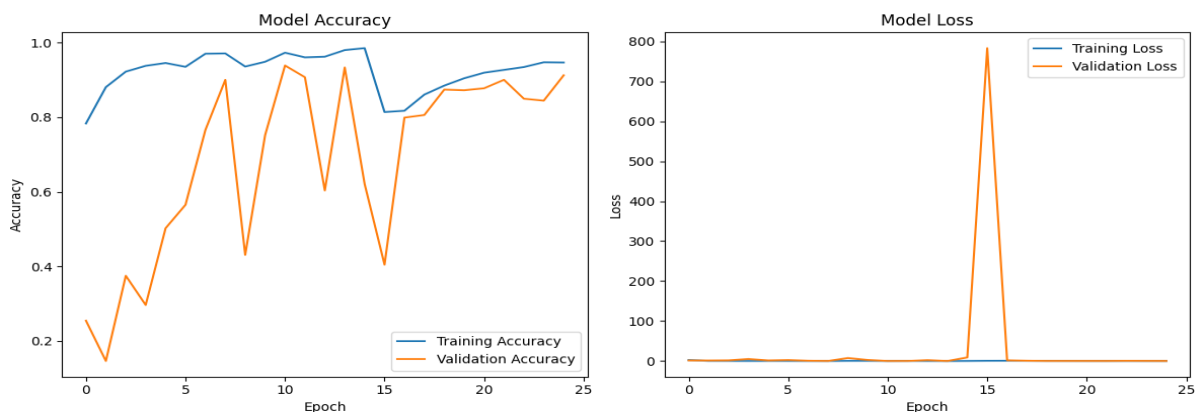
Tabel 3. Confusion matrix model CNN

Model	Kelas	TP	FN	FP	TN
ResNet50	pituitary	284	16	41	980
	meningioma	237	79	18	987
	glioma	401	4	36	880
	notumor	289	11	15	1006
VGG16	pituitary	283	17	32	989
	meningioma	277	39	27	978
	glioma	405	0	5	911
	notumor	289	11	3	1018
MobileNetV2	pituitary	271	29	18	1003
	meningioma	278	38	34	971
	glioma	405	0	10	906
	notumor	294	6	11	1010

Model ResNet50 dilatih selama 25 epoch dengan batch size 32. Setelah pelatihan selesai, dilakukan evaluasi terhadap performa model dalam mengklasifikasikan tumor otak menggunakan *confusion matrix*. Tabel 4 merangkum hasil evaluasi model ini, sedangkan Gambar 4 memperlihatkan perubahan akurasi dan loss selama pelatihan.

Tabel 4. Evaluasi performa resnet50

Model	Kelas	Presisi	Recall	F1-Score	Specificity
ResNet50	pituitary	95,07	96,33	95,7	95,98
	meningioma	92,94	75	83,01	98,21
	glioma	87,38	94,67	90,88	96,07
	notumor	91,76	99,01	95,25	98,53



Gambar 4. Grafik akurasi dan loss model resnet50

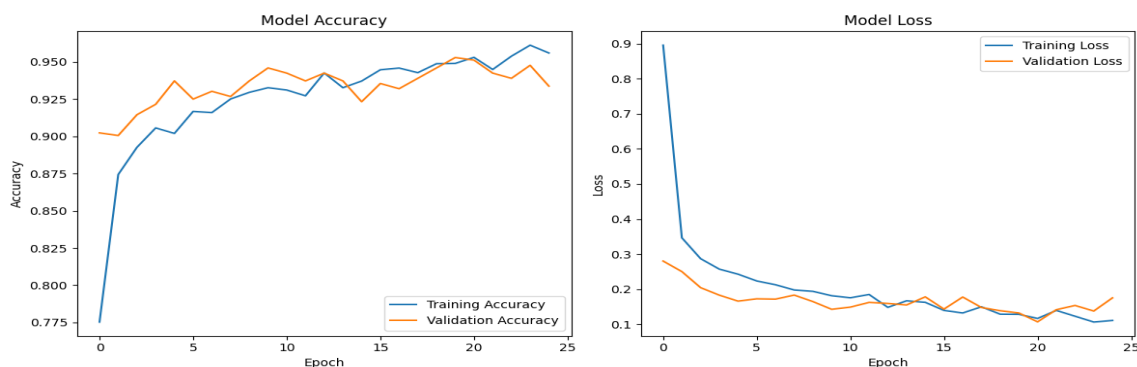
Berdasarkan Tabel 4, ResNet50 memiliki performa yang cukup baik dalam mendeteksi kelas notumor dan pituitary, dengan recall masing-masing sebesar 99,01% dan 96,33%. Namun, model masih kesulitan mengklasifikasikan meningioma dengan recall terendah (75%), yang menunjukkan bahwa banyak sampel dari kelas ini salah diklasifikasikan sebagai kelas lain. Meskipun glioma memiliki recall tinggi (94,67%), presisinya lebih rendah (87,38%), yang mengindikasikan model cenderung memberikan prediksi positif yang lebih luas tetapi kurang akurat dalam beberapa kasus.

Gambar 4 menunjukkan bahwa akurasi pelatihan dan validasi mengalami peningkatan yang cukup stabil, tetapi pada epoch ke-15 terjadi lonjakan tajam pada loss validasi. Hal ini menunjukkan adanya outlier atau kesalahan adaptasi model pada subset data tertentu. Setelah lonjakan tersebut, loss validasi stabil kembali, menunjukkan model mengatasi gangguan dan melanjutkan pelatihan dengan baik.

Model VGG16 dilatih selama 25 epoch dengan batch size 32. Setelah pelatihan selesai, confusion matrix digunakan untuk mengevaluasi kinerja model dalam mengklasifikasikan setiap kelas tumor. Hasil evaluasi disajikan dalam Tabel 5, sementara Gambar 5 menampilkan grafik akurasi dan loss selama proses pelatihan.

Tabel 5. Evaluasi performa vgg16

Model	Kelas	Presisi	Recall	F1-Score	Specificity
VGG16	pituitary	98,97	96,33	97,64	99,70
	meningioma	91,12	87,66	89,35	97,31
	glioma	89,84	94,33	92,03	96,87
	notumor	98,78	100	99,39	99,45



Gambar 5. Grafik akurasi dan loss model vgg16

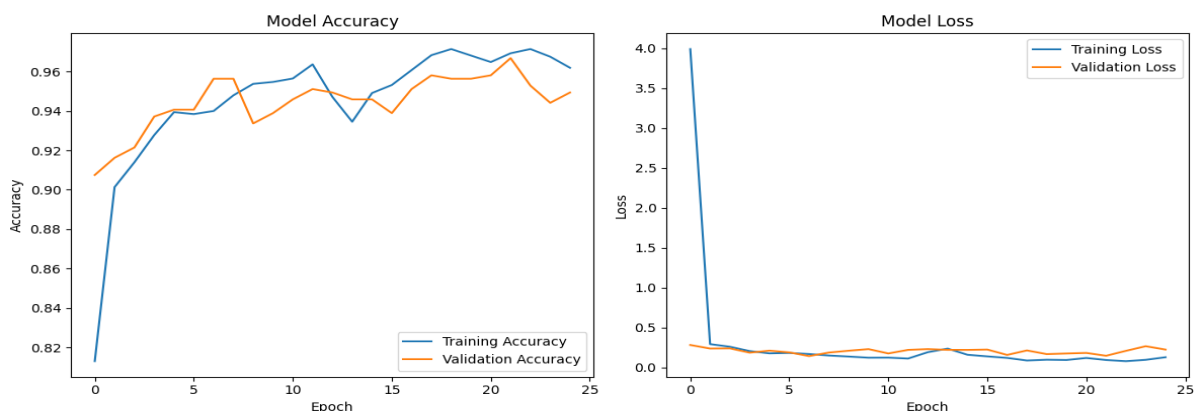
Berdasarkan Tabel 5, VGG16 menunjukkan kinerja sangat baik dalam mendeteksi kelas notumor dengan recall 100% dan presisi 98,78%, menunjukkan bahwa model sangat akurat dalam mengenali sampel tanpa tumor. Pituitary juga memiliki performa sangat baik, dengan presisi 98,97%. Namun, meningioma menunjukkan tantangan, dengan recall yang lebih rendah (87,66%) dan presisi 91,12%, mengindikasikan adanya kesulitan model dalam membedakan tumor jenis ini dengan kelas lainnya. Meskipun demikian, specificity untuk semua kelas tetap tinggi, mencapai 99,70% pada pituitary, yang menunjukkan kemampuan model dalam membedakan kelas negatif dengan sangat baik.

Gambar 5 menunjukkan bahwa akurasi pelatihan meningkat stabil, namun akurasi validasi sedikit berfluktuasi. Loss pelatihan dan validasi menunjukkan penurunan yang tajam, dengan stabilitas yang tercapai pada akhir pelatihan, mengindikasikan konvergensi yang baik pada model.

Pelatihan MobileNetV2 dilakukan selama 25 epoch dengan batch size 32. Setelah pelatihan selesai, model diuji menggunakan *confusion matrix* untuk menilai keakuratan klasifikasi pada setiap kategori tumor. Tabel 6 menampilkan hasil evaluasi model, dan Gambar 6 merepresentasikan kurva akurasi serta loss selama proses pelatihan.

Tabel 6. Evaluasi performa mobilenetv2

Model	Kelas	Presisi	Recall	F1-Score	Specificity
MobileNetV2	pituitary	96,39	98	97,19	98,24
	meningioma	89,1	87,97	88,54	96,62
	glioma	93,77	90,33	92,02	98,91
	notumor	97,59	100	98,78	98,92

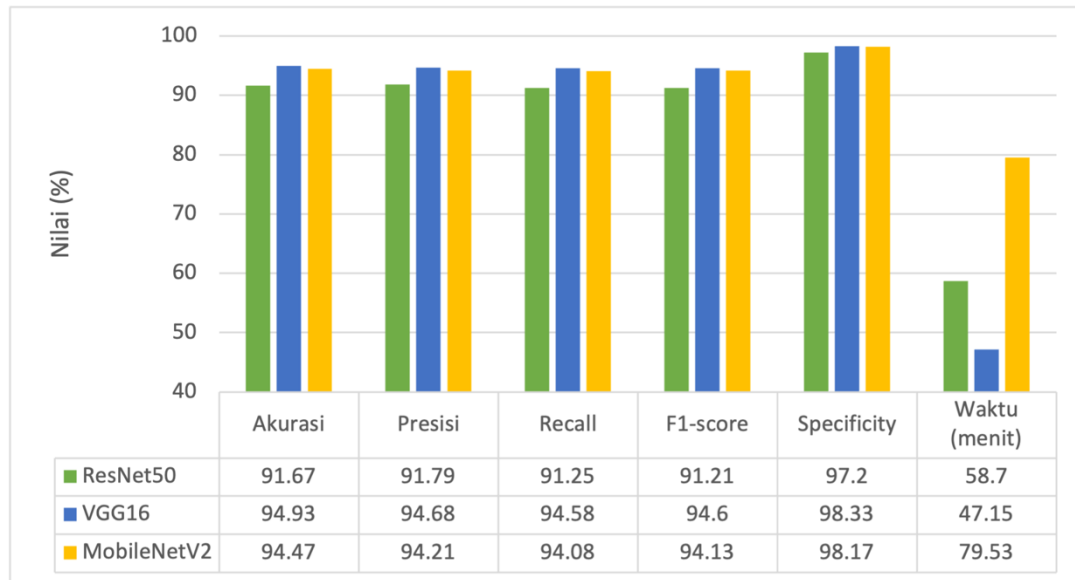


Gambar 6. Grafik akurasi dan loss model mobilenetv2

Berdasarkan Tabel 6, MobileNetV2 menunjukkan hasil yang sangat baik pada kelas notumor dengan recall 100% dan presisi 97,59%, yang menunjukkan kemampuan model dalam mengenali sampel tanpa tumor secara akurat. Kelas pituitary juga memiliki presisi tinggi (96,39%) dan recall (98%), yang menandakan model sangat baik dalam mengenali tumor jenis ini. Meningioma menunjukkan performa sedikit lebih rendah dengan recall 87,97% dan presisi 89,1%, yang menunjukkan bahwa model masih menghadapi kesulitan dalam mengidentifikasi meningioma dengan sangat akurat. Specificity untuk setiap kelas cukup tinggi, dengan nilai tertinggi pada kelas notumor (98,92%) dan glioma (98,91%), menunjukkan kemampuan model dalam membedakan sampel negatif dengan sangat baik.

Gambar 6 menunjukkan bahwa akurasi pelatihan meningkat stabil selama proses pelatihan, dengan sedikit fluktuasi pada akurasi validasi yang menunjukkan kestabilan model. Loss model juga mengalami penurunan tajam pada awal pelatihan dan tetap stabil pada nilai yang rendah di epoch berikutnya, yang mengindikasikan konvergensi yang baik dan rendahnya risiko overfitting.

Penelitian ini mengevaluasi performa tiga model CNN, yaitu ResNet50, MobileNetV2, dan VGG16, dalam tugas klasifikasi tumor otak. Setiap model dilatih menggunakan dataset yang terdiri dari citra MRI otak dan diuji untuk memprediksi empat kategori tumor: glioma, meningioma, pituitary, dan notumor. Gambar 7 menunjukkan perbandingan kinerja dari ketiga model, mengukur metrik evaluasi utama seperti akurasi, presisi, recall, F1-score, specificity, serta waktu pelatihan yang diperlukan oleh masing-masing model.



Gambar 7. Perbandingan kinerja model

5 Kesimpulan

Berdasarkan hasil evaluasi yang diperoleh, VGG16 menunjukkan kinerja terbaik di antara ketiga model yang diuji dalam klasifikasi tumor otak berbasis citra MRI. Model ini mencatatkan akurasi tertinggi (94,93%), diikuti oleh presisi (94,68%), recall (94,58%), dan specificity (98,33%) yang juga menunjukkan hasil yang unggul. VGG16 juga memiliki waktu pelatihan yang lebih efisien dibandingkan dengan model lainnya, yaitu 47.15 menit, menjadikannya pilihan optimal dalam hal efisiensi dan efektivitas. Meskipun MobileNetV2 memiliki akurasi sedikit lebih rendah (94,47%) dan membutuhkan waktu pelatihan yang paling lama (79.53 menit), model ini tetap menunjukkan kinerja yang solid, dengan recall yang baik (94,08%) dan specificity (98,17%), menunjukkan efisiensi dalam mengidentifikasi tumor otak secara keseluruhan. ResNet50, meskipun unggul dalam presisi (91.79%), mencatatkan akurasi terendah di antara ketiga model (91,67%) dan menunjukkan kelemahan dalam hal recall (91,25%) dan specificity (97,2%). Waktu pelatihan ResNet50 adalah 58.7 menit, lebih lama dibandingkan dengan VGG16, namun lebih cepat dibandingkan dengan MobileNetV2. Secara keseluruhan, VGG16 terbukti sebagai model yang lebih seimbang dan efisien dalam klasifikasi tumor otak. Penelitian lebih lanjut dapat mempertimbangkan penerapan teknik tambahan seperti data augmentation, ensemble learning, atau penggunaan Vision Transformer untuk meningkatkan kinerja klasifikasi, khususnya pada kategori tumor yang lebih sulit.

Referensi

- [1] H. Mehnatkesh, S. M. J. Jalali, A. Khosravi, and S. Nahavandi, "An Intelligent Driven Deep Residual Learning Framework for Brain Tumor Classification using MRI Images," *Expert Syst Appl*, vol. 213, 2023, doi: 10.1016/j.eswa.2022.119087.
- [2] S. Lapointe, A. Perry, and N. A. Butowski, "Primary Brain Tumours in Adults," *The Lancet*, vol. 392, no. 10145, pp. 432–446, 2018, doi: 10.1016/S0140-6736(18)30990-5.
- [3] D. V. Gore and V. Deshpande, "Comparative Study of Various Techniques using Deep Learning for Brain Tumor Detection," *2020 International Conference for Emerging Technology, INCET 2020*, no. June 2020, 2020, doi: 10.1109/INCET49848.2020.9154030.
- [4] A. Rovira and C. Auger, "Beyond McDonald: Updated Perspectives on MRI Diagnosis of Multiple Sclerosis," *Expert Rev Neurother*, vol. 21, no. 8, pp. 895–911, 2021, doi: 10.1080/14737175.2021.1957832.
- [5] M. Havaei et al., "Brain Tumor Segmentation with Deep Neural Networks," *Med Image Anal*, vol. 35, pp. 18–31, 2017, doi: 10.1016/j.media.2016.05.004.
- [6] M. Bintah, T. Noor, N. Z. Zenia, M. S. Kaiser, S. Al Mamun, and M. Mahmud, "Application of Deep Learning in Detecting Neurological Disorders from Magnetic Resonance Images : A

- Survey on the Detection of Alzheimer's Disease , Parkinson's Disease and Schizophrenia," Brain Inform*, 2020, doi: 10.1186/s40708-020-00112-2.
- [7] T. D. Nguyen, D. T. Le, J. Bum, S. Kim, S. J. Song, and H. Choo, "Retinal Disease Diagnosis using Deep Learning on Ultra-Wide-Field Fundus Images," *Diagnostics*, vol. 14, no. 1, pp. 1–19, 2024, doi: 10.3390/diagnostics14010105.
- [8] S. Natha, U. Laila, I. A. Gashim, K. Mahboob, M. N. Saeed, and K. M. Noaman, "Automated Brain Tumor Identification in Biomedical Radiology Images: A Multi-Model Ensemble Deep Learning Approach," *Applied Sciences*, vol. 14, no. 5, p. 2210, 2024, doi: 10.3390/app14052210.
- [9] M. B. Kurniawan and E. Utami, "Comparative Analysis of Contrast Enhancement Methods for Classification of Pekalongan Batik Motifs using Convolutional Neural Network," *Jurnal Teknik Informatika (JUTIF)*, vol. 5, no. 6, pp. 1779–1787, 2024, doi: 10.52436/1.jutif.2024.5.6.2621.
- [10] A. Younis, L. Qiang, C. O. Nyatega, M. J. Adamu, and H. B. Kawuwa, "Brain Tumor Analysis using Deep Learning and VGG-16 Ensembling Learning Approaches," *Applied Sciences (Switzerland)*, vol. 12, no. 14, 2022, doi: 10.3390/app12147282.
- [11] L. Xu and M. Mohammadi, "Brain Tumor Diagnosis from MRI based on Mobilenetv2 Optimized by Contracted fox Optimization Algorithm," *Heliyon*, vol. 10, no. 1, 2024, doi: 10.1016/j.heliyon.2023.e23866.
- [12] P. S. L. Veeranki, G. L. Banavath, and P. R. Devi, "Detection and Classification of Brain Tumors using Convolutional Neural Network," *7th International Conference on Trends in Electronics and Informatics, ICOEI 2023 - Proceedings*, pp. 780–786, 2023, doi: 10.1109/ICOEI56765.2023.10125652.
- [13] M. A. B. Siddiaue, S. Sakib, M. M. R. Khan, A. K. Tanzeem, M. Chowdhury, and N. Yasmin, "Deep Convolutional Neural Networks Model-based Brain Tumor Detection in Brain MRI Images," *Proceedings of the 4th International Conference on IoT in Social, Mobile, Analytics and Cloud, ISMAC 2020*, pp. 909–914, 2020, doi: 10.1109/I-SMAC49090.2020.9243461.
- [14] R. L. Kumar, J. Kakarla, B. V. Isunuri, and M. Singh, "Multi-Class Brain Tumor Classification using Residual Network and Global Average Pooling," *Multimed Tools Appl*, vol. 80, no. 9, pp. 13429–13438, 2021, doi: 10.1007/s11042-020-10335-4.
- [15] Z. N. K. Swati et al., "Brain Tumor Classification for Mr Images using Transfer Learning and Fine-Tuning," *Computerized Medical Imaging and Graphics*, vol. 75, pp. 34–46, 2019, doi: 10.1016/j.compmedimag.2019.05.001.
- [16] A. E. Minarno, M. H. C. Mandiri, Y. Munarko, and H. Hariyady, "Convolutional Neural Network with Hyperparameter Tuning for Brain Tumor Classification," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, vol. 4, 2021, doi: 10.22219/kinetik.v6i2.1219.
- [17] P. Rajak, A. S. Jangde, and G. P. Gupta, "Towards Design of Brain Tumor Detection Framework using Deep Transfer Learning Techniques," *IGI Global*, pp. 90–103, 2023, doi: 10.4018/978-1-6684-5264-6.ch004.
- [18] N. Ullah et al., "An Effective Approach to Detect and Identify Brain Tumors using Transfer Learning," *Applied Sciences (Switzerland)*, vol. 12, no. 11, 2022, doi: 10.3390/app12115645.
- [19] G. S. Tandel, A. Tiwari, and O. G. Kakde, "Performance Optimisation of Deep Learning Models using Majority Voting Algorithm for Brain Tumour Classification," *Comput Biol Med*, vol. 135, no. June, p. 104564, 2021, doi: 10.1016/j.compbimed.2021.104564.
- [20] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, "A Deep Learning-based Framework for Automatic Brain Tumors Classification using Transfer Learning," *Circuits Syst Signal Process*, vol. 39, no. 2, pp. 757–775, 2020, doi: 10.1007/s00034-019-01246-3.
- [21] M. A. Hamza et al., "Optimal and Efficient Deep Learning Model for Brain Tumor Magnetic Resonance Imaging Classification and Analysis," *Applied Sciences (Switzerland)*, vol. 12, no. 15, 2022, doi: 10.3390/app12157953.
- [22] M. Wu, Q. Liu, C. Yan, and G. Sen, "Multi-Classification of Brain Tumors on Magnetic Resonance Images using an Ensemble of Pre-Trained Convolutional Neural Networks," *Curr Med Imaging*, vol. 19, pp. 65–76, 2023, doi: 10.2174/1573405618666220415122843.

- [23] A. Kazemi, M. E. Shiri, A. Sheikahmadi, and M. Khodamoradi, “Classifying Tumor Brain Images using Parallel Deep Learning Algorithms,” *Comput Biol Med*, vol. 148, p. 105775, 2022, doi: 10.1016/j.compbiomed.2022.105775.
- [24] N. Noreen, S. Palaniappan, A. Qayyum, I. Ahmad, and M. O. Alassafi, “Brain Tumor Classification based on Fine-Tuned Models and the Ensemble Method,” *Computers, Materials and Continua*, vol. 67, no. 3, pp. 3967–3982, 2021, doi: 10.32604/cmc.2021.014158.
- [25] M. I. Mahmud, M. Mamun, and A. Abdelgawad, “A Deep Analysis of Brain Tumor Detection from MR Images using Deep Learning Networks,” *Algorithms*, vol. 16, no. 4, pp. 1–19, 2023, doi: 10.3390/a16040176.
- [26] N. Shamshad et al., “Enhancing Brain Tumor Classification by a Comprehensive Study on Transfer Learning Techniques and Model Efficiency using MRI Datasets,” *IEEE Access*, vol. 12, no. May, pp. 100407–100418, 2024, doi: 10.1109/ACCESS.2024.3430109.
- [27] N. Masoud, “Brain Tumor MRI Dataset,” Kaggle. Accessed: Sep. 16, 2024. [Online]. Available: <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>