

MRI Classification of Brain Tumors Using EfficientNetB0 Feature Extraction and Machine Learning Methods

Firza Findia Jiven ^{1*}, Rumini ^{2*}

*Informatika, Universitas Amikom Yogyakarta
firzafj@students.amikom.ac.id ¹, rumini@amikom.ac.id ²

Article Info

Article history:

Received 2025-07-21

Revised 2025-11-17

Accepted 2025-11-22

Keyword:

Brain Tumor,
Feature Extraction,
Machine Learning,
MRI Classification

ABSTRACT

Brain tumor classification using MRI images plays a crucial role in modern medical diagnostics, offering fast and accurate support for disease detection. This study proposes a classification approach that combines feature extraction using EfficientNet B0 with conventional machine learning algorithms. MRI brain images are preprocessed and resized to match EfficientNet B0 input dimensions. Feature vectors are extracted and subsequently processed using PCA for dimensionality reduction and SMOTE for class balancing. The resulting data are classified using various machine learning algorithms including Support Vector Machine, XGBoost, LightGBM, and others. Experimental results show that Support Vector Machine achieved the highest accuracy of 96%, followed by XGBoost and LightGBM at 94%. The combination of EfficientNet B0 feature extraction and lightweight classifiers proved to be effective, matching the performance of more complex deep learning models. This study does not focus on measuring computational cost directly, but rather demonstrates that combining EfficientNetB0 feature extraction with machine learning algorithms can achieve performance comparable to deep learning approaches. This highlights that lightweight models remain competitive in terms of accuracy without requiring highly complex architectures. Future work can explore this method on other medical imaging datasets and enhance model interpretability for clinical adoption.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Klasifikasi citra medis dalam deteksi tumor otak menggunakan citra MRI, mempunyai peran penting dalam dunia kedokteran modern, karena dapat mendukung proses diagnosa yang lebih cepat dan akurat serta membantu perencanaan pengobatan. Perkembangan teknologi pemrosesan citra sangat pesat, dengan penerapan teknik deep learning seperti *Convolution Neural Networks* (CNN) yang telah terbukti efektif dalam analisis citra medis[1]. Model CNN yang populer seperti ResNet50 dan VGG16 telah digunakan untuk mendeteksi berbagai jenis kelainan dalam gambar medis[2]. Meskipun akurasi tinggi, metode ini sering memerlukan komputasi yang sangat besar, hal ini bisa menjadi tantangan dalam penerapan klinis yang memerlukan efisiensi dan kecepatan[3]. Oleh karena itu, penting untuk terus mengembangkan penelitian yang bertujuan

meningkatkan efisiensi dalam klasifikasi citra medis dengan mempertahankan akurasi yang optimal[4].

Dalam studi yang dilakukan oleh sen dkk., penerapan CNN dengan arsitektur ResNet50 dan MobileNetV2 berhasil mencapai akurasi sekitar 96% pada klasifikasi tumor otak berbasis MRI. Walaupun performanya cukup baik, kedua arsitektur tersebut memiliki tingkat kompleksitas yang tinggi karena jumlah parameternya besar dan membutuhkan komputasi intensif. Di sisi lain, EfficientNetB0 mampu menghasilkan akurasi hingga 97% dengan jumlah parameter lebih kecil, menjadikannya lebih efisien namun tetap bersaing[5]. Pertimbangan ini menjadi dasar penggunaan EfficientNetB0 dalam penelitian ini.

Meskipun Penerapan deep learning dalam klasifikasi citra medis terbukti efektif, namun masih terdapat celah dalam penggunaan model *EfficientNet* untuk citra MRI tumor otak. Banyak penelitian yang hanya memanfaatkan EfficientNet

untuk ekstraksi fitur, kemudian melanjutkan proses klasifikasi dengan menggunakan deep learning seperti CNN[3]. Namun, belum banyak yang meneliti penerapan EfficientNet B0 untuk ekstraksi fitur diikuti dengan penggunaan algoritma machine learning untuk klasifikasi. Pendekatan ini dapat mengurangi kebutuhan data komputasi besar, namun masih jarang dijelajahi dalam studi mengenai citra medis. Celah ini memunculkan pertanyaan apakah algoritma machine learning dapat memberikan hasil yang setara dengan model deep learning yang lebih berat, namun lebih efisien dalam hal waktu dan sumber daya[6].

Untuk mengatasi celah, penelitian ini mengusulkan pendekatan yang memadukan penggunaan EfficientNet B0 untuk ekstraksi fitur dengan penerapan algoritma machine learning dalam klasifikasi citra MRI tumor otak. Dengan memanfaatkan metode seperti Ridge Classifier, Support Vector Machine (SVM), dan Logistic Regression, tujuan penelitian ini adalah untuk menunjukkan bahwa klasifikasi dapat dilakukan dengan akurasi tinggi tanpa memerlukan daya komputasi yang besar. Pendekatan ini bertujuan untuk menyediakan solusi yang lebih hemat sumber daya, yang sangat penting untuk aplikasi di dunia medis dalam diagnosis klinis[7]. Penggunaan algoritma machine learning setelah proses ekstraksi fitur dengan EfficientNet juga dapat menyederhanakan model dibandingkan dengan penerapan full deep learning[8]. Selain itu, sebagaimana dijelaskan oleh Ghosh dan Kole[9], algoritma machine learning seperti Logistic Regression dan Random Forest tidak hanya efisien secara komputasi, tetapi juga lebih mudah dipahami dan diinterpretasikan dibandingkan model deep learning yang kompleks. Hal ini sangat penting dalam konteks medis, karena interpretabilitas hasil dapat meningkatkan kepercayaan dokter terhadap sistem. Model yang sederhana dan ringan ini juga memungkinkan integrasi yang lebih mudah ke dalam sistem klinis yang memiliki keterbatasan perangkat keras.

Selain faktor efisiensi, interpretabilitas juga menjadi aspek penting dalam pengembangan model machine learning untuk aplikasi medis. Beberapa penelitian sebelumnya [10], [11] menekankan bahwa meskipun akurasi tinggi krusial, model dengan sifat black-box cenderung sulit diterapkan dalam praktik klinis karena tenaga medis tidak dapat menelusuri alasan di balik keputusan model. Pada kasus deteksi tumor otak, interpretabilitas diperlukan untuk menyoroti bagian citra yang paling memengaruhi prediksi, sehingga meningkatkan transparansi serta kepercayaan dokter terhadap sistem berbasis AI[12]. Oleh karena itu, penelitian ini tidak hanya menekankan efisiensi komputasi, tetapi juga membuka ruang pengembangan di masa depan melalui penerapan teknik interpretabilitas agar model lebih mudah diterima dan diandalkan di ranah medis.

Dalam penelitian ini EfficientNet B0 digunakan untuk mengekstraksi fitur dari citra MRI tumor otak. Sebelum ekstraksi fitur, citra terlebih dahulu di proses dengan teknik resize dan pre-processing yang sesuai dengan persyaratan model EfficientNet B0. Setelah fitur diekstraksi, algoritma

machine learning seperti Ridge Classifier, Support Vector Machine (SVM), dan Logistic Regression diterapkan untuk melakukan klasifikasi. Untuk mengatasi masalah ketidakseimbangan kelas dalam dataset, digunakan teknik SMOTE, serta diterapkan PCA untuk mereduksi dimensi fitur yang telah diekstraksi. Hasil dari eksperimen ini diharapkan dapat memberikan pemahaman lebih dalam tentang apakah pendekatan menggunakan algoritma machine learning dapat sebanding dengan metode deep learning[13].

Hasil eksperimen menunjukkan bahwa penggunaan EfficientNet B0 untuk ekstraksi fitur diikuti oleh algoritma *Support Vector Machine* (SVM), XGBoost, dan LightGBM menghasilkan akurasi yang sangat baik, setara dengan metode *deep learning* lainnya. Algoritma Support Vector Machine (SVM) juga menunjukkan kinerja yang kompetitif meskipun dengan akurasi sedikit lebih rendah. Penggunaan SMOTE terbukti efektif dalam meningkatkan kinerja model pada dataset yang tidak seimbang, sementara PCA mempercepat proses pelatihan tanpa mengorbankan akurasi. Eksperimen ini membuktikan bahwa pendekatan yang menggabungkan algoritma machine learning dengan ekstraksi fitur dari EfficientNet B0 dapat memberikan efisiensi lebih besar dalam hal waktu komputasi dan penggunaan sumber daya. Secara keseluruhan, temuan ini mendukung bahwa model berbasis machine learning dapat bersaing dengan metode deep learning dalam klasifikasi citra medis[14].

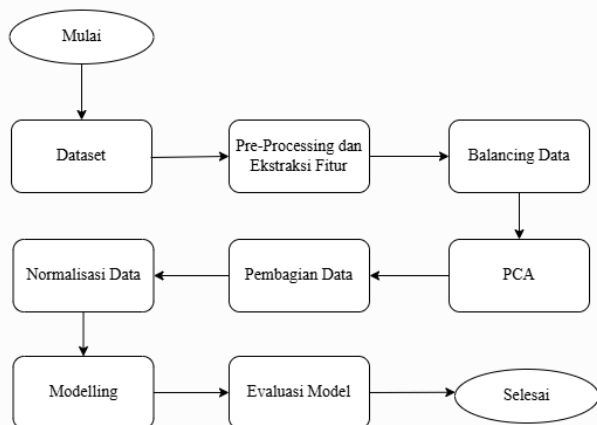
Hasil eksperimen berhasil mengisi celah yang ada dengan menunjukkan bahwa kombinasi *EfficientNet B0* untuk ekstraksi fitur dan algoritma machine learning untuk klasifikasi dapat menghasilkan kinerja yang setara dengan deep learning. Pendekatan ini tidak hanya mencapai akurasi yang tinggi, tetapi juga mengurangi beban komputasi, yang sangat penting untuk penerapan di dunia nyata. Dengan menggunakan *Support Vector Machine* (SVM), XGBoost, dan LightGBM, terbukti bahwa model yang lebih ringan dapat beroperasi secara efektif dengan sumber daya yang lebih terbatas tanpa mengurangi kualitas prediksi. Selain itu, penggunaan SMOTE dan PCA meningkatkan kinerja model pada dataset yang tidak seimbang, mengatasi masalah umum yang ditemui dalam citra medis. Temuan ini memberikan wawasan baru bahwa untuk beberapa aplikasi klinis, machine learning setelah ekstraksi fitur bisa menjadi solusi yang lebih efisien [15].

Kesimpulan dari penelitian ini membuka kemungkinan untuk pendekatan yang lebih efisien dalam klasifikasi citra medis, khususnya pada kasus MRI tumor otak. Pendekatan yang menggabungkan EfficientNet B0 dengan algoritma machine learning ini dapat diterapkan di lingkungan medis yang memerlukan waktu respons cepat dan keterbatasan sumber daya komputasi. Penelitian ini juga menunjukkan bahwa model deep learning yang kompleks tidak selalu diperlukan untuk mencapai hasil yang optimal. Ekstraksi fitur menggunakan model ringan diikuti dengan algoritma machine learning bisa menjadi alternatif yang lebih praktis. Di masa depan, penelitian ini dapat diperluas dengan menguji metode serupa pada jenis citra medis lain serta

memperkenalkan teknik interpretabilitas untuk memahami keputusan model dengan lebih baik. Hal ini dapat meningkatkan adopsi machine learning dalam aplikasi medis.

II. METODE

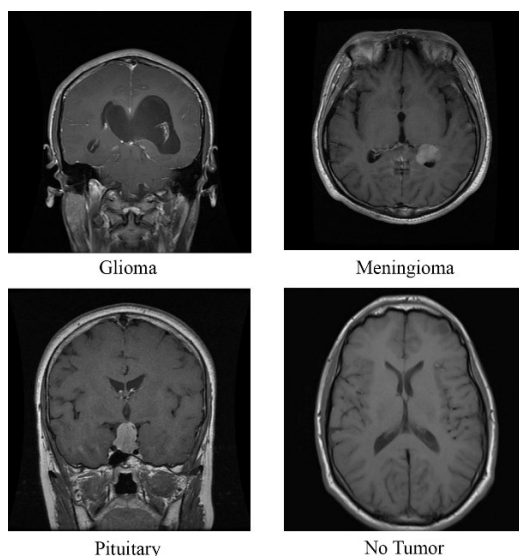
Penelitian ini melibatkan beberapa tahapan, seperti yang ada di gambar 1. Penelitian ini terdiri dari beberapa tahapan utama yaitu pengambilan dataset tumor otak, pre-processing, balancing data, reduksi dimensi, pembagian data, normalisasi data, pelatihan model, serta evaluasi performa model.



Gambar 1 Tahapan Penelitian

A. Dataset

Penelitian ini menggunakan dataset citra MRI otak yang mencakup beberapa jenis tumor, yaitu glioma, meningioma, pituitary, serta kategori tanpa tumor (no tumor). Dataset tersebut diunduh dari platform Kaggle melalui link: <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset/data>.



Gambar 2 Sample Dataset

File disediakan dalam format .zip dan berisi gambar dengan resolusi tinggi yang telah dikelompokkan ke dalam direktori sesuai dengan masing-masing label kelas. Seluruh gambar kemudian diproses dengan menyesuaikan ukurannya menjadi 224×224 piksel agar sesuai dengan format input dari arsitektur EfficientNet B0 yang digunakan dalam tahap ekstraksi fitur. Sample dari dataset dapat dilihat pada Gambar 2.

Citra pada bagian kiri atas menampilkan jenis glioma, yang tampak sebagai massa tidak beraturan pada area parenkim otak dengan batas tepi yang kurang tegas. Pada sisi kanan atas terlihat meningioma, yang umumnya muncul di dekat permukaan otak dan memiliki bentuk lebih teratur dengan kontras jelas terhadap jaringan di sekitarnya. Gambar 2 kiri bawah memperlihatkan pituitary tumor yang berlokasi di kelenjar hipofisis di dasar otak dan menunjukkan peningkatan intensitas pada area tersebut. Adapun citra kanan bawah menggambarkan kondisi otak normal atau no tumor, tanpa indikasi adanya kelainan struktural.

Jumlah data pada setiap kategori kelas secara rinci disajikan pada Tabel 1.

TABEL I
JUMLAH CITRA DALAM DATASET CITRA MRI TUMOR OTAK

Kelas Tumor	Jumlah Citra
Glioma	1321
Meningioma	1339
No Tumor	1595
Pituitary	1457

B. Preprocessing dan Ekstraksi Fitur

Setiap gambar diproses terlebih dahulu menggunakan fungsi `preprocess_input` dari pustaka Keras, kemudian dikonversi ke dalam bentuk array numerik. Proses ekstraksi fitur dilakukan dengan memanfaatkan arsitektur EfficientNet B0 yang dikonfigurasi tanpa lapisan fully connected (dengan pengaturan `include_top=False` dan `pooling='avg'`). Pengaturan ini memungkinkan model menghasilkan representasi vektor fitur berdimensi 1280 untuk setiap citra. Seluruh vektor hasil ekstraksi, beserta label kelasnya, disimpan dalam file berformat .csv untuk keperluan tahap klasifikasi berikutnya.

C. Pemuatan Fitur dan Balancing Data

Setelah fitur dan label berhasil disimpan, data tersebut dimuat kembali untuk menjalani tahap pemrosesan lanjutan. Label kelas diubah menjadi format numerik dengan memanfaatkan `LabelEncoder`, agar kompatibel dengan model klasifikasi yang memerlukan input dalam bentuk angka. Salah satu permasalahan umum dalam pengolahan data medis adalah ketimpangan distribusi antar kelas, di mana sebagian kelas memiliki jumlah sampel yang jauh lebih sedikit dibandingkan kelas lainnya. Untuk mengatasi isu ini, digunakan pendekatan Synthetic Minority Oversampling Technique (SMOTE), yang menghasilkan data sintetik dari kelas minoritas dengan cara melakukan interpolasi antar data

yang berdekatan. Tujuan dari penerapan SMOTE adalah untuk mengurangi bias model terhadap kelas mayoritas serta memastikan model mampu mengenali karakteristik kelas minoritas secara seimbang.

D. Reduksi Dimensi

Untuk menekan kompleksitas komputasi sekaligus meminimalkan risiko overfitting, dilakukan langkah reduksi dimensi menggunakan teknik Principal Component Analysis (PCA). Metode ini berfungsi untuk memetakan data fitur berdimensi tinggi ke dalam ruang berdimensi lebih rendah, dengan tetap mempertahankan sebanyak mungkin variansi dari data asli. Pada penelitian ini, jumlah komponen utama dipilih sedemikian rupa sehingga mampu mencakup lebih dari 95% total variansi. Proses reduksi ini tidak hanya mempercepat waktu pelatihan model, tetapi juga membantu meningkatkan kemampuan generalisasi dengan cara mengurangi noise yang terdapat pada fitur.

E. Pembagian dan Normalisasi Data

Data yang telah melalui proses SMOTE dan PCA selanjutnya dibagi menjadi data pelatihan dan data pengujian menggunakan metode stratified train-test split dengan perbandingan 70:30. Teknik stratifikasi ini diterapkan agar proporsi distribusi label tetap seimbang pada kedua subset data. Setelah proses pembagian selesai, dilakukan normalisasi fitur menggunakan StandardScaler dari pustaka Scikit-learn. Proses ini mentransformasikan setiap fitur agar memiliki rata-rata nol dan standar deviasi satu, sehingga perbedaan skala antar fitur tidak memengaruhi performa model klasifikasi dan proses pelatihan dapat berjalan secara optimal.

F. Klasifikasi dengan Algoritma Machine Learning

Tahap klasifikasi dalam penelitian ini dilakukan dengan menerapkan beragam algoritma machine learning, masing-masing dengan keunikan dan tingkat kompleksitas tersendiri. Beberapa model yang digunakan meliputi *Logistic Regression*, *Ridge Classifier*, *Support Vector Machine* (SVM), *Random Forest*, *Decision Tree*, *Gradient Boosting*, *AdaBoost*, *Naive Bayes*, serta *K-Nearest Neighbors* (KNN). Selain itu, dua algoritma ensemble modern, yaitu XGBoost dan LightGBM, juga diikutsertakan dalam evaluasi. Seluruh model dilatih menggunakan data pelatihan yang telah diproses, kemudian diuji pada data pengujian untuk menilai kemampuannya dalam mengklasifikasikan empat kategori citra MRI otak: glioma, meningioma, pituitary, dan no tumor.

Pemilihan algoritma dalam penelitian ini didasarkan pada hasil penelitian terdahulu dan pertimbangan teknis yang relevan. Logistic Regression dan Ridge Classifier dipilih karena mampu menangani data berdimensi tinggi sekaligus memberikan interpretasi hasil yang jelas [9]. SVM digunakan karena terbukti memiliki performa stabil pada data non-linear, seperti dilaporkan oleh Sen dkk. [5]. Algoritma ensemble modern, yaitu XGBoost dan LightGBM, dipertimbangkan karena menunjukkan akurasi dan kemampuan generalisasi yang tinggi pada data citra medis kompleks [16]. Sebagai pembanding, model sederhana seperti Decision Tree, Naive

Bayes, dan KNN turut dilibatkan sebagai baseline sesuai pendekatan pada penelitian sebelumnya [17].

Penggunaan berbagai algoritma ini dimaksudkan untuk melakukan perbandingan performa antar model secara sistematis, guna menemukan pendekatan paling efektif dalam klasifikasi citra medis yang berbasis pada fitur hasil ekstraksi deep learning. Dengan mengevaluasi beragam metode, diharapkan dapat diidentifikasi model yang tidak hanya unggul dari segi akurasi, tetapi juga mampu memberikan generalisasi yang baik pada data baru.

G. Evaluasi Model

Kinerja model klasifikasi dievaluasi menggunakan berbagai metrik yang lazim digunakan dalam konteks klasifikasi multikelas. Metrik-metrik ini dirancang untuk mengukur seberapa efektif model dalam mengidentifikasi dan membedakan setiap kelas secara adil, khususnya pada dataset medis yang sering menghadirkan tantangan berupa ketidakseimbangan jumlah data antar kelas. Beberapa indikator utama yang digunakan dalam evaluasi mencakup accuracy, precision, recall, F1-score, dan ROC-AUC. Kelima metrik ini secara kolektif memberikan pemahaman menyeluruh tentang kemampuan model dalam mengurangi kesalahan klasifikasi serta menjaga keseimbangan antara prediksi positif dan negatif yang akurat.

Akurasi digunakan untuk mengukur seberapa besar proporsi prediksi yang tepat dibandingkan dengan keseluruhan jumlah prediksi yang dilakukan. Rumus akurasi dinyatakan sebagai:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN}$$

di mana TP (*true positive*) adalah jumlah prediksi positif yang benar, TN (*true negative*) adalah jumlah prediksi negatif yang benar, FP (*false positive*) merupakan prediksi positif yang salah, dan FN (*false negative*) adalah prediksi negatif yang salah. Dalam konteks klasifikasi multikelas, akurasi dihitung dengan membagi total jumlah prediksi yang benar dengan keseluruhan jumlah data sampel.

Precision merupakan indikator yang menunjukkan seberapa akurat model dalam mengidentifikasi sampel sebagai kelas positif. Metrik ini sangat penting dalam situasi di mana kesalahan false positive harus ditekan seminimal mungkin, seperti pada konteks diagnosis medis. Precision dihitung menggunakan rumus berikut:

$$\text{Precision} = \frac{TP}{TP+FP}$$

Recall, atau yang juga dikenal sebagai sensitivitas, mengukur kemampuan model dalam menemukan semua kasus yang benar-benar termasuk dalam kelas positif. Metrik ini sangat krusial dalam bidang medis karena berkaitan langsung dengan potensi model untuk tidak mengabaikan atau melewatkan diagnosis terhadap kondisi penyakit yang

seharusnya terdeteksi. Nilai recall dihitung dengan rumus sebagai berikut:

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1-score merupakan rata-rata harmonik dari precision dan recall, yang digunakan sebagai ukuran kinerja model ketika diperlukan keseimbangan antara kedua metrik tersebut. F1-score sangat bermanfaat dalam situasi di mana baik ketepatan prediksi maupun kemampuan mendeteksi kasus positif memiliki peran yang sama penting. Rumus perhitungannya dituliskan sebagai berikut:

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Nilai F1-score tinggi menunjukkan bahwa model mampu memberikan hasil prediksi yang baik dengan meminimalkan kesalahan baik dari sisi false positive maupun false negative.

ROC-AUC adalah metrik yang digunakan untuk menilai sejauh mana sebuah model mampu membedakan kelas positif dan negatif pada berbagai nilai threshold. Semakin mendekati angka 1, semakin baik kemampuan model dalam melakukan pemisahan antar kelas. Untuk kasus klasifikasi multikelas, perhitungan ROC-AUC dilakukan dengan metode one-vs-rest sehingga performa masing-masing kelas dapat dianalisis secara individual. Metrik ini memiliki peran penting dalam aplikasi medis karena dapat membantu mengidentifikasi potensi kesalahan prediksi, terutama false negative, yang berisiko tinggi dalam proses diagnosis.

Dalam konteks klasifikasi multikelas, perhitungan metrik evaluasi dapat dilakukan menggunakan pendekatan macro average atau weighted average, tergantung pada tujuan analisis. Pendekatan macro average menghitung nilai metrik secara individual untuk setiap kelas, kemudian mengambil rata-rata dari semua kelas tanpa mempertimbangkan jumlah sampel. Sebaliknya, weighted average menghitung rata-rata metrik dengan mempertimbangkan proporsi sampel pada masing-masing kelas, sehingga memberikan bobot lebih besar pada kelas dengan jumlah data yang lebih banyak.

Selain menggunakan metrik numerik, evaluasi model juga dilakukan secara visual melalui confusion matrix. Matriks dua dimensi ini menampilkan jumlah prediksi yang tepat maupun yang keliru untuk setiap kombinasi antara kelas sebenarnya dan kelas hasil prediksi. Confusion matrix sangat membantu dalam mengidentifikasi pola kesalahan klasifikasi, misalnya apakah model cenderung salah mengklasifikasikan kelas "glioma" sebagai "meningioma" atau sebaliknya. Dalam penelitian ini, *confusion matrix* divisualisasikan dalam bentuk heatmap menggunakan pustaka Seaborn, sehingga hubungan antar kelas dapat diamati dengan lebih intuitif melalui representasi visual.

Gabungan antara evaluasi berbasis angka dan visualisasi memberikan wawasan yang lebih komprehensif mengenai kinerja dan karakteristik model. Pendekatan ini juga menjadi landasan dalam menentukan algoritma paling efektif yang layak diterapkan dalam sistem klasifikasi citra medis berbasis

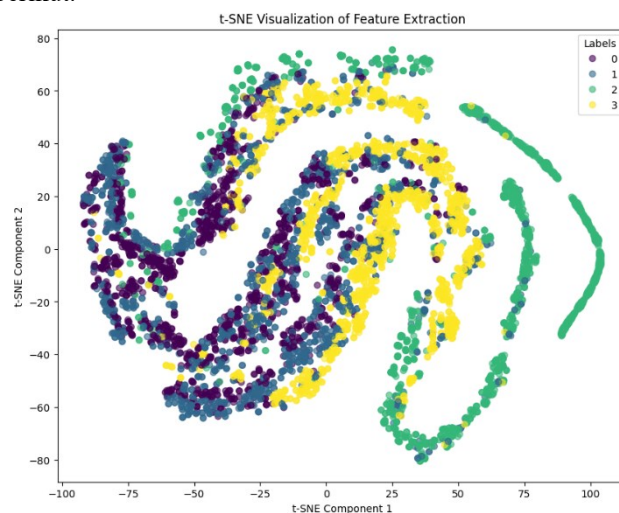
MRI di dunia nyata. Evaluasi secara menyeluruh sangatlah krusial, terutama dalam konteks klinis, karena kesalahan dalam prediksi berpotensi menimbulkan konsekuensi serius terhadap proses diagnosis maupun pengambilan keputusan terkait penanganan pasien.

III. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil yang diperoleh dari tiap tahapan yang telah diuraikan dalam metode penelitian sebelumnya. Penjelasan berikut menggambarkan secara sistematis temuan yang dihasilkan dari keseluruhan proses yang telah dilaksanakan.

A. Ekstraksi Fitur

Penelitian dimulai dengan tahap awal pemrosesan dataset citra MRI otak yang terbagi dalam empat kategori yaitu glioma, meningioma, pituitary, serta no tumor. Masing-masing gambar disesuaikan ukurannya menjadi 224×224 piksel guna memenuhi spesifikasi input dari arsitektur EfficientNet B0. Ekstraksi fitur dilakukan dengan memanfaatkan model EfficientNet B0 yang telah dilatih sebelumnya pada dataset ImageNet, menggunakan konfigurasi tanpa lapisan akhir (*include_top=False*) dan strategi pooling rata-rata global, sehingga dihasilkan vektor fitur sepanjang 1280 elemen untuk tiap gambar. Sebagai pelengkap informasi visual, karakteristik tekstural citra juga diekstraksi melalui pendekatan Gray-Level Co-occurrence Matrix (GLCM). Hasil dari kedua metode ekstraksi ini kemudian dikombinasikan dan disimpan dalam format .csv untuk digunakan dalam tahap klasifikasi berikutnya. Visualisasi dari hasil ekstraksi fitur dapat dilihat pada gambar berikut.



Gambar 3 Visualisasi Ekstraksi Fitur

B. Pemuatan Fitur

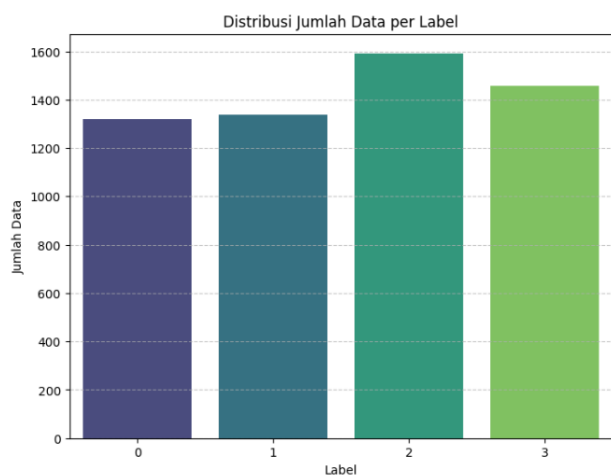
Setelah proses ekstraksi fitur dari citra MRI selesai dan hasilnya disimpan dalam format .csv, data tersebut dimuat kembali untuk digunakan dalam tahap pelatihan model. Label

kelas yang awalnya berbentuk teks seperti glioma, meningioma, pituitary, dan no tumor diubah menjadi format numerik menggunakan fungsi *LabelEncoder* dari library Scikitlearn. Konversi ini bertujuan untuk memastikan data dapat diterima oleh algoritma klasifikasi yang membutuhkan input dalam bentuk angka, sekaligus mempermudah dalam perhitungan metrik evaluasi. Selain itu, transformasi label ke angka juga mempermudah proses stratifikasi saat membagi dataset menjadi data pelatihan dan pengujian, sehingga distribusi kelas tetap seimbang di kedua subset data.

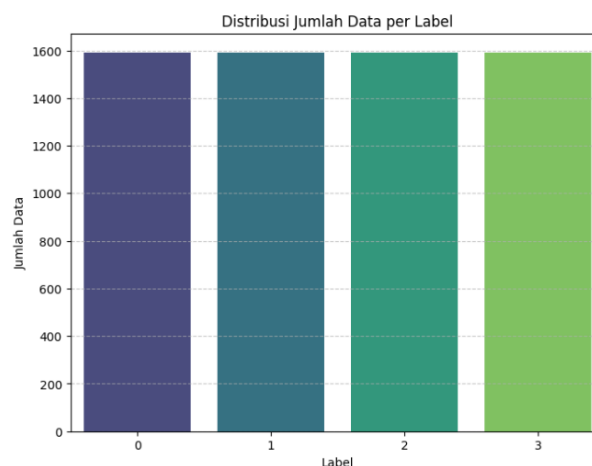
C. Balancing Data (SMOTE)

Sebelum memulai proses pelatihan model, dilakukan evaluasi awal terhadap sebaran jumlah data pada masing-masing kelas dalam dataset. Hasil analisis awal memperlihatkan adanya ketidakseimbangan, di mana kelas 2 memiliki jumlah sampel lebih dominan dibandingkan kelas lainnya seperti pada Gambar 4. Ketimpangan ini berisiko menyebabkan model cenderung bias terhadap kelas mayoritas. Untuk mengurangi permasalahan tersebut, diterapkan metode *Synthetic Minority Oversampling Technique* (SMOTE), yang bekerja dengan membuat data sintetis melalui interpolasi antar titik dari kelas minoritas yang berdekatan pada ruang fitur. Dalam penelitian ini, SMOTE dijalankan dengan parameter default ($k_neighbors=5$, $random_state=42$), sehingga setiap sampel minoritas memiliki tetangga terdekat yang digunakan sebagai acuan interpolasi.

Setelah dilakukan oversampling dengan SMOTE, setiap kelas memiliki jumlah data yang sama, yaitu 1.592 gambar seperti pada Gambar 5. Keseimbangan ini membuat proses pelatihan berlangsung lebih adil karena tidak ada kelas yang mendominasi. Walaupun secara teori metode SMOTE dapat menghasilkan titik sintetis yang tidak sepenuhnya representatif dan berisiko mengurangi kemampuan generalisasi, hasil visualisasi distribusi data tidak menunjukkan adanya outlier maupun pola anomali. Hal ini mengindikasikan bahwa distribusi tetap merata sehingga dampak negatif terhadap generalisasi dapat diabaikan.



Gambar 4 Kelas Sebelum dilakukan SMOTE



Gambar 5 Kelas Sesudah dilakukan SMOTE

D. Reduksi Dimensi (PCA)

Karena penggabungan fitur dari CNN dan GLCM menghasilkan data dengan dimensi yang sangat tinggi, diterapkan teknik *Principal Component Analysis* (PCA) untuk mereduksi jumlah fitur. Tujuan utama dari penggunaan PCA adalah menjaga lebih dari 95% total variansi dari data awal agar informasi penting tetap terwakili, sekaligus mengurangi kemungkinan terjadinya overfitting dan mempercepat waktu pelatihan model. Dalam penelitian ini, PCA diterapkan dengan parameter $n_components=100$, sehingga dimensi awal yang sangat tinggi berhasil direduksi menjadi 100 komponen utama.

Visualisasi menggunakan t-SNE setelah penerapan PCA pada Gambar 3 menunjukkan bahwa data masih terklasifikasi secara alami sesuai label tanpa adanya indikasi pola anomali. Kondisi ini mengisyaratkan bahwa PCA tidak menyebabkan hilangnya informasi penting ataupun menghasilkan distribusi yang menyimpang. Dengan demikian, meskipun terdapat kemungkinan pengurangan detail, penggunaan PCA dalam penelitian ini dinilai aman karena mampu memperbaiki efisiensi komputasi tanpa menurunkan akurasi model.

E. Pembagian Data

Setelah tahap penyeimbangan data diselesaikan, dataset kemudian dipisahkan menjadi dua bagian, yakni data pelatihan dan data pengujian, dengan perbandingan 70% untuk pelatihan dan 30% untuk pengujian. Proses pembagian ini dilakukan menggunakan metode *Stratified Train-Test Split* guna memastikan bahwa proporsi jumlah sampel dari setiap kelas tetap konsisten di kedua subset. Pendekatan stratifikasi ini sangat krusial untuk mencegah terjadinya ketimpangan distribusi label yang dapat memengaruhi hasil pelatihan maupun evaluasi. Dengan sebaran data yang seimbang, model memperoleh kesempatan belajar yang adil dari seluruh kelas, serta dapat diuji secara akurat dalam mengklasifikasikan masing-masing kategori.

F. Normalisasi Data

Setelah data dibagi menjadi data latih dan data uji, dilakukan proses normalisasi terhadap fitur menggunakan fungsi *StandardScaler* dari library *Scikit-learn*. Proses ini mengubah skala setiap fitur sehingga memiliki nilai rata-rata nol dan deviasi standar satu. Normalisasi ini sangat penting, terutama bagi algoritma seperti *Support Vector Machine* (SVM) dan *K-Nearest Neighbors* (KNN), yang sangat dipengaruhi oleh skala data input. Tanpa penyesuaian skala, fitur dengan nilai numerik yang lebih besar dapat mendominasi proses pelatihan dan memengaruhi hasil klasifikasi. Oleh karena itu, normalisasi menjadi langkah krusial dalam menjaga kestabilan, keadilan, dan akurasi model selama proses pembelajaran.

G. Evaluasi Model

Penelitian ini mengevaluasi performa sebelas algoritma machine learning dalam tugas klasifikasi citra MRI otak, berdasarkan fitur yang diperoleh dari ekstraksi menggunakan model *EfficientNet B0* yang dikombinasikan dengan GLCM. Algoritma yang diuji meliputi *Logistic Regression*, *Ridge Classifier*, *Support Vector Machine* (SVM), *Decision Tree*, *Random Forest*, *Gradient Boosting*, *AdaBoost*, *Naive Bayes*, *K-Nearest Neighbors* (KNN), serta dua metode *ensemble* terkini yaitu *LightGBM* dan *XGBoost*.

Penilaian kinerja dilakukan dengan memanfaatkan metrik seperti akurasi, F1-score dalam versi macro dan weighted, serta visualisasi melalui confusion matrix. Seluruh dataset telah melalui tahap balancing menggunakan SMOTE, reduksi dimensi dengan PCA, serta normalisasi menggunakan *StandardScaler*, guna menjamin proses pelatihan berjalan secara efisien dan optimal.

TABEL 1
HASIL PERBANDINGAN MODEL

Model	Result				
	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	0.90	0.90	0.90	0.90	0.98
Decision Tree	0.81	0.81	0.81	0.81	0.87
Random Forest	0.93	0.93	0.93	0.93	0.99
Gradient Boosting	0.90	0.90	0.90	0.90	0.98
AdaBoost	0.75	0.77	0.75	0.76	0.91
Naive Bayes	0.82	0.82	0.82	0.82	0.95
K-Nearest Neighbors	0.92	0.92	0.92	0.92	0.99
Ridge Classifier	0.89	0.89	0.89	0.89	0.97
Support Vector Machine	0.96	0.96	0.96	0.96	0.99
XGBoost	0.94	0.94	0.94	0.94	0.99
LightGBM	0.94	0.94	0.94	0.94	0.99

Berdasarkan hasil evaluasi dari Tabel 1, algoritma *Support Vector Machine* (SVM) tampil paling unggul dengan capaian akurasi 96%, precision, recall, dan F1-score sebesar 0,96, serta skor ROC-AUC mencapai 0,99. Hal ini menunjukkan

bahwa model tidak hanya konsisten dalam mengenali tiap kelas, tetapi juga memiliki kemampuan diskriminasi yang sangat baik. Nilai recall yang tinggi (0,96) menjadi faktor krusial dalam konteks medis, karena mampu memastikan sebagian besar kasus tumor dapat terdeteksi dengan baik sekaligus meminimalkan risiko false negative yang berbahaya pada praktik klinis.

Di bawahnya, algoritma *XGBoost* dan *LightGBM* sama-sama menunjukkan kinerja kuat dengan akurasi 94% dan ROC-AUC 0,99, sehingga dapat dianggap sebagai opsi kompetitif untuk mendukung deteksi medis berbasis citra MRI. Sementara itu, model seperti *Random Forest* dan *K-Nearest Neighbors* (KNN) juga memperlihatkan performa stabil dengan akurasi di atas 90% dan ROC-AUC mendekati 0,99, meskipun masih sedikit di bawah tiga model terbaik.

Sebaliknya, algoritma seperti *Decision Tree*, *AdaBoost*, dan *Naive Bayes* menunjukkan performa relatif lebih rendah, dengan akurasi berkisar 75–82% serta nilai ROC-AUC yang lebih kecil dibandingkan model unggulan. Konsistensi nilai precision, recall, dan F1-score pada model terbaik memperkuat bukti bahwa kinerja yang dicapai tidak hanya tinggi pada satu metrik, melainkan seimbang dalam mendeteksi dan mengklasifikasikan seluruh kelas tumor.

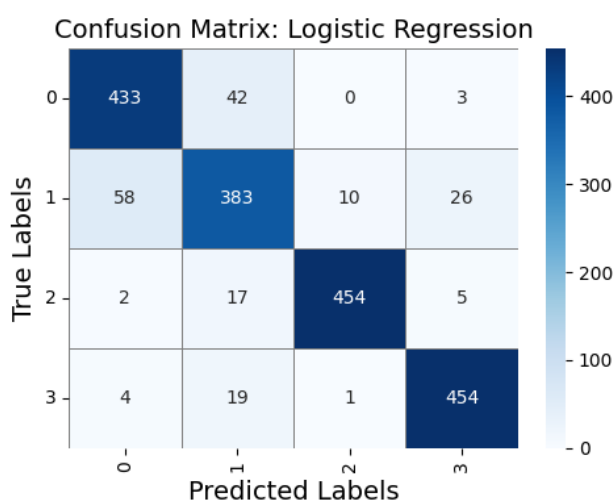
Berdasarkan Tabel 2, penelitian-penelitian terdahulu umumnya memanfaatkan arsitektur CNN end-to-end dan menghasilkan akurasi yang sangat tinggi. Sen dkk.[5] melaporkan bahwa penggunaan *EfficientNetB0* mampu mencapai akurasi 97,61% untuk klasifikasi tumor otak multi-kelas. Di sisi lain, Wibowo dkk.[18] menggunakan *EfficientNetB7* yang dioptimalkan dengan hyperparameter tuning, dan mencatatkan akurasi tertinggi yakni 98%. Sedangkan Candra dkk.[19], dengan menerapkan transfer learning berbasis VGG-16 pada dataset yang lebih kecil, berhasil memperoleh akurasi sebesar 97%.

TABEL 2
PERBANDINGAN DENGAN PENELITIAN SEBELUMNYA

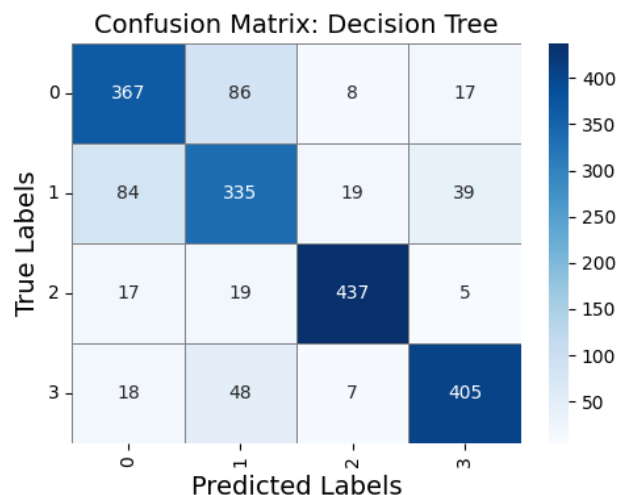
Penelitian	Metode	Akurasi	Keterangan
G. Balaji, R. Sen, dan H. Kirty[5]	CNN dengan Transfer Learning (<i>EfficientNet B0</i> , <i>ResNet50</i> , <i>Xception</i> , <i>MobileNetV2</i> , <i>VGG16</i>).	Efficient NetB0: 97.61% (tertinggi), <i>VGG16</i> dan model lain lebih rendah	Klasifikasi multi-kelas tumor otak. Preprocessing: denoising, skull stripping, augmentation.
M. A. Purnama Wibowo, M. B. Al Fayyadl, Y. Azhar, dan Z. Sari [18]	<i>EfficientNetB0</i> & <i>EfficientNetB7</i> dengan Hyperparameter Tuning, Data Augmentation	91–98% (skenario terbaik mencapai 98%)	Dataset 3264 MRI (4 kelas). Optimasi hyperparameter meningkatkan performa.
D. Candra, G. Wibisono, M. Ayu, and M. Afrad [19]	Transfer Learning menggunakan VGG-16 (16 convolutional	97%	Dataset kecil (253 MRI, 155 tumor & 98 non-tumor). Fokus

	layers, fine tuning, dropout, fully connected)		pada binary classification (tumor vs non-tumor). VGG-16 terbukti efektif meski dengan dataset terbatas.
Penelitian ini	EfficientNetB0 + Support Vector Machine	96% (F1-score 0,96)	Hybrid: ekstraksi fitur CNN + machine learning. Lebih ringan komputasi, tetap kompetitif.

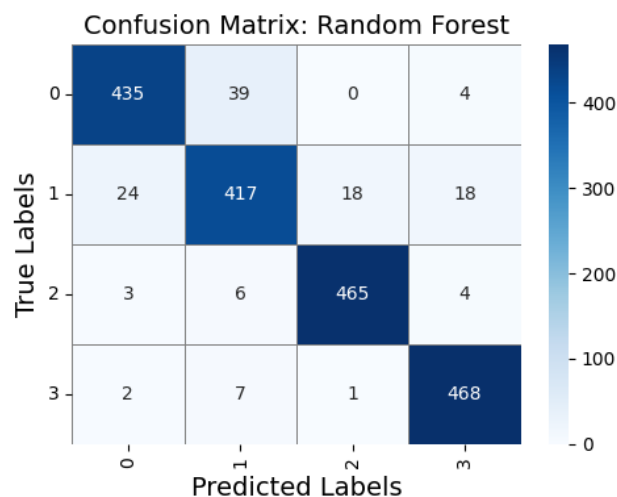
Dengan menggabungkan EfficientNetB0 untuk ekstraksi fitur dan Support Vector Machine sebagai metode klasifikasi, penelitian ini berhasil mencapai akurasi 96% dengan F1-score 0,96. Walaupun performanya sedikit di bawah beberapa studi terdahulu, pendekatan hybrid ini tetap menunjukkan daya saing serta efisiensi komputasi yang lebih baik. Meskipun belum dilakukan pengujian signifikansi statistik antar model, hasil perbandingan ini menegaskan bahwa penggunaan model ringan tetap memiliki relevansi, khususnya pada lingkungan dengan keterbatasan sumber daya.



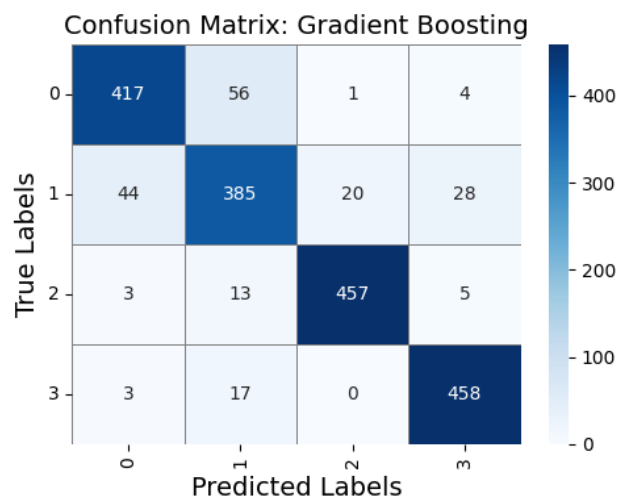
Gambar 6 Model Logistik Regression



Gambar 7 Model Decision Tree

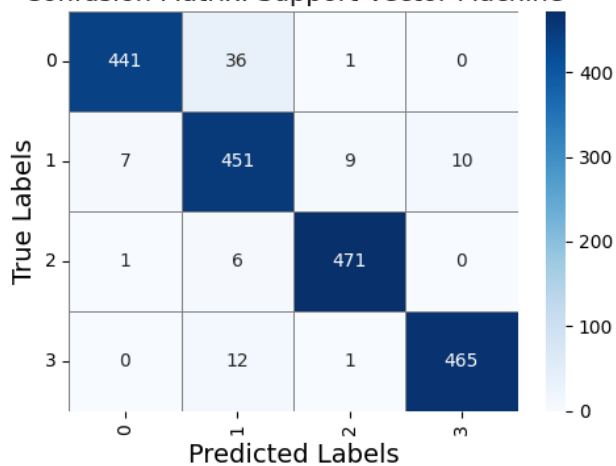


Gambar 8 Model Random Forest



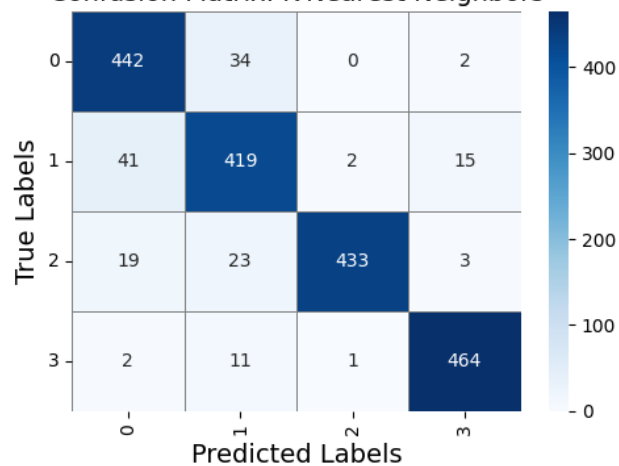
Gambar 9 Model Gradient Boosting

Confusion Matrix: Support Vector Machine



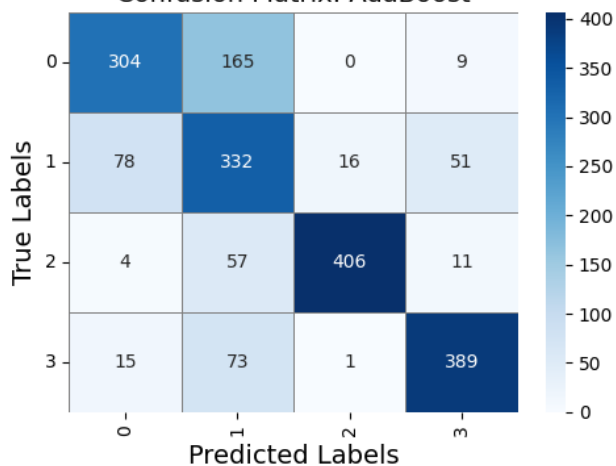
Gambar 10 Model Support Vector Machine

Confusion Matrix: K-Nearest Neighbors



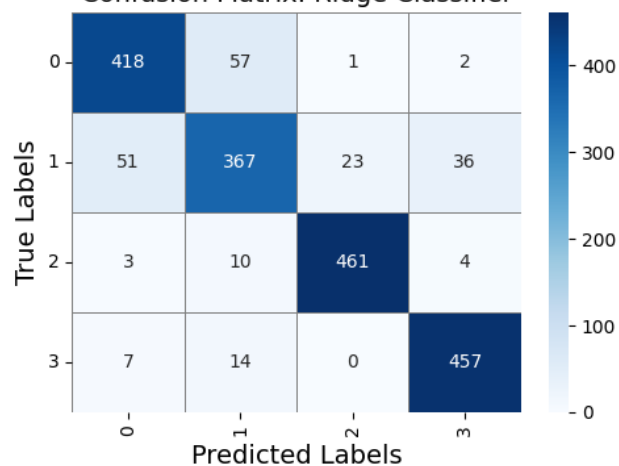
Gambar 13 Model K-Nearest Neighbors

Confusion Matrix: AdaBoost



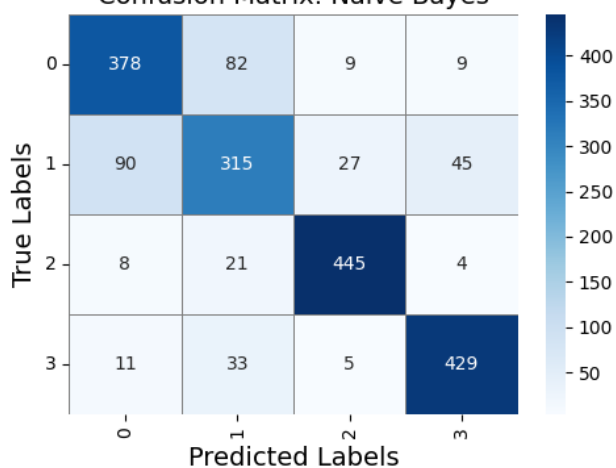
Gambar 11 Model AdaBoost

Confusion Matrix: Ridge Classifier



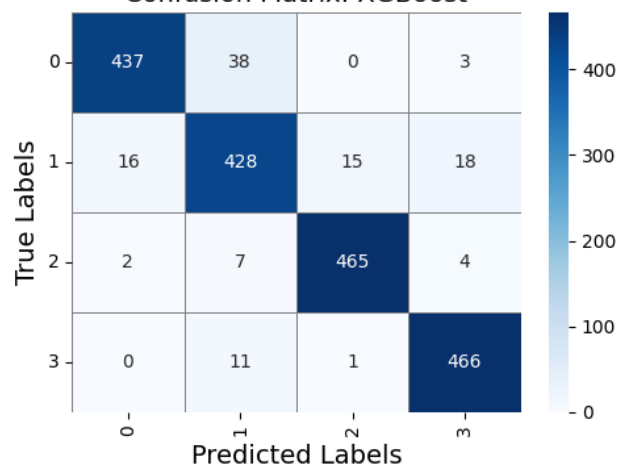
Gambar 14 Model Ridge Classifier

Confusion Matrix: Naive Bayes

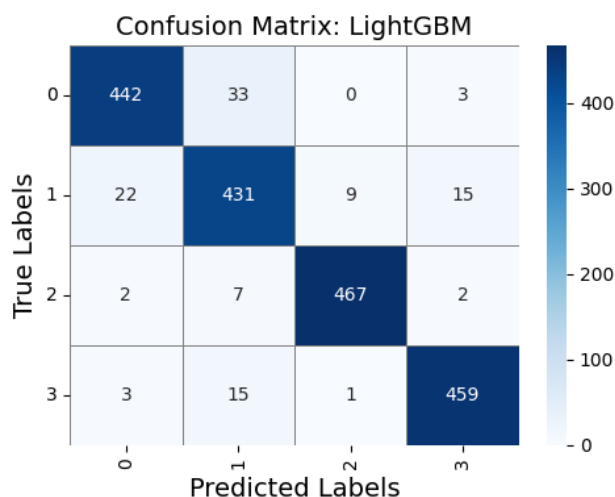


Gambar 12 Model Naive Bayes

Confusion Matrix: XGBoost



Gambar 15 Model XGBoost



Gambar 16 Model LightGBM

Hasil analisis terhadap seluruh confusion matrix memperlihatkan bahwa algoritma seperti Support Vector Machine (SVM), LightGBM, XGBoost, dan Ridge Classifier mampu mengenali setiap kelas dengan akurasi sangat tinggi serta tingkat kesalahan prediksi yang rendah. SVM, misalnya, menunjukkan distribusi hasil klasifikasi yang hampir sempurna dengan kesalahan antarkelas yang minim, khususnya pada kelas 2 dan 3. Hal serupa terlihat pada LightGBM dan XGBoost yang konsisten menghasilkan prediksi benar dominan di diagonal utama, menandakan efektivitasnya dalam membedakan jenis tumor otak yang memiliki kemiripan visual, seperti glioma dan meningioma. Ridge Classifier juga menghasilkan capaian serupa dengan akurasi tinggi di setiap kelas. Temuan ini menegaskan bahwa model-model tersebut tidak hanya unggul dalam mengenali kelas mayoritas, tetapi juga andal dalam mendeteksi kelas minoritas, yang menunjukkan keberhasilan strategi prapemrosesan seperti penggunaan SMOTE dan PCA.

Sebaliknya, kinerja algoritma seperti Naive Bayes dan AdaBoost relatif kurang optimal, terlihat dari banyaknya kesalahan klasifikasi antar kelas. Naive Bayes mengalami kesulitan khususnya pada kelas 0 dan 1, dengan tingkat salah prediksi yang cukup tinggi, kemungkinan besar disebabkan oleh asumsi distribusi yang terlalu sederhana untuk mengakomodasi kompleksitas fitur gabungan dari CNN dan GLCM. Pada sisi lain, AdaBoost juga menunjukkan tingkat kesalahan yang signifikan, terutama pada kelas 0 dan 1. Decision Tree menempati posisi menengah, namun masih menghasilkan mis-klasifikasi yang cukup jelas, terutama pada dua kelas yang sama.

Adapun *K-Nearest Neighbors* (KNN) cenderung sulit membedakan kelas 1 dan 2, meskipun performanya sedikit lebih baik dibandingkan Naive Bayes dan AdaBoost. Sementara itu, Random Forest dan Gradient Boosting tampil lebih stabil, dengan akurasi tinggi serta distribusi kesalahan yang lebih terkontrol. Random Forest menunjukkan konsistensi melalui dominasi prediksi diagonal, sedangkan Gradient Boosting mampu menjaga keseimbangan dalam

klasifikasi multi-kelas. Secara umum, hasil confusion matrix menegaskan bahwa pendekatan berbasis regularisasi (SVM, Ridge Classifier) maupun *ensemble boosting* (XGBoost, LightGBM, Gradient Boosting, Random Forest) lebih unggul dalam menangani klasifikasi multi-kelas berbasis fitur EfficientNetB0, dibandingkan model sederhana seperti Naive Bayes, Decision Tree, dan AdaBoost.

IV. KESIMPULAN

Penelitian ini membuktikan bahwa metode ekstraksi fitur dengan memanfaatkan *EfficientNet B0*, ketika dikombinasikan dengan algoritma machine learning seperti *Support Vector Machine*, XGBoost, dan LightGBM, mampu menghasilkan klasifikasi citra MRI tumor otak yang akurat dengan efisiensi komputasi yang lebih optimal dibandingkan dengan pendekatan deep learning secara menyeluruh.

Teknik SMOTE efektif dalam menyeimbangkan distribusi kelas, sementara penerapan PCA membantu mereduksi dimensi data sehingga proses pelatihan tidak mengorbankan kualitas prediksi model tanpa menggunakan dimensi yang besar. Ridge Classifier muncul sebagai model dengan performa tertinggi, mencatatkan akurasi hingga 95%, disusul oleh Logistic Regression dan LightGBM yang juga menunjukkan hasil yang kompeten.

Temuan ini membuka peluang besar bagi penerapan sistem deteksi berbasis machine learning di lingkungan medis yang seringkali memiliki keterbatasan dalam hal daya komputasi. Di samping itu, pendekatan ini menjadi fondasi yang solid untuk riset lanjutan, seperti pengujian pada berbagai dataset citra medis lainnya, peningkatan aspek interpretabilitas model, serta pengembangan sistem yang terintegrasi dengan platform web atau perangkat mobile guna mendukung proses diagnosis klinis secara lebih praktis.

DAFTAR PUSTAKA

- [1] M. Arabahmadi, R. Farahbakhsh, and J. Rezazadeh, "Deep Learning for Smart Healthcare—A Survey on Brain Tumor Detection from Medical Imaging," Mar. 01, 2022, *MDPI*. doi: 10.3390/s22051960.
- [2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition." [Online]. Available: <http://image-net.org/challenges/LSVRC/2015/>
- [3] M. Tan and Q. V Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks."
- [4] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks." [Online]. Available: <http://image-net.org/challenges/LSVRC/2017/results>
- [5] G. Balaji, R. Sen, and H. Kirty, "Detection and Classification of Brain tumors Using Deep Convolutional Neural Networks," Aug. 2022, [Online]. Available: <http://arxiv.org/abs/2208.13264>
- [6] C. Szegedy *et al.*, "Going Deeper with Convolutions."
- [7] I. Soesanti, H. Avizenna, and I. Ardiyanto, "Classification of Brain Tumor MRI Image using Random Forest Algorithm and Multilayers Perceptron," 2020.
- [8] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Sep. 2014, [Online]. Available: <http://arxiv.org/abs/1409.1556>

- [9] A. GHOSH and A. KOLE, "A Comparative Study of Enhanced Machine Learning Algorithms for Brain Tumor Detection and Classification," Oct. 26, 2021. doi: 10.36227/techrxiv.16863136.v1.
- [10] A. Holzinger, G. Lings, H. Denk, K. Zatloukal, and H. Müller, "Causability and explainability of artificial intelligence in medicine," Jul. 01, 2019, *Wiley-Blackwell*. doi: 10.1002/widm.1312.
- [11] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization." [Online]. Available: <http://gradcam.cloudev.org>
- [12] S. M. Lundberg, P. G. Allen, and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions." [Online]. Available: <https://github.com/slundberg/shap>
- [13] J. Seetha and S. S. Raja, "Brain tumor classification using Convolutional Neural Networks," *Biomedical and Pharmacology Journal*, vol. 11, no. 3, pp. 1457–1461, Sep. 2018, doi: 10.13005/bpj/1511.
- [14] M. Li, Y. Jiang, Y. Zhang, and H. Zhu, "Medical image analysis using deep learning algorithms," *Front Public Health*, vol. 11, 2023, doi: 10.3389/fpubh.2023.1273253.
- [15] P. K. Mall *et al.*, "A comprehensive review of deep neural networks for medical image processing: Recent developments and future opportunities," *Healthcare Analytics*, vol. 4, Dec. 2023, doi: 10.1016/j.health.2023.100216.
- [16] A. R. W. Sait and R. Nagaraj, "An Enhanced LightGBM-Based Breast Cancer Detection Technique Using Mammography Images," *Diagnostics*, vol. 14, no. 2, Jan. 2024, doi: 10.3390/diagnostics14020227.
- [17] I. Zine-dine, J. Riffi, K. El Fazazi, M. A. Mahraz, and H. Tairi, "Brain Tumor Classification using Machine and Transfer Learning," *Scitepress*, May 2022, pp. 566–571. doi: 10.5220/0010762800003101.
- [18] M. A. Purnama Wibowo, M. B. Al Fayyadl, Y. Azhar, and Z. Sari, "Classification of Brain Tumors on MRI Images Using Convolutional Neural Network Model EfficientNet," *Jurnal RESTI*, vol. 6, no. 4, pp. 538–547, Aug. 2022, doi: 10.29207/resti.v6i4.4119.
- [19] D. Candra, G. Wibisono, M. Ayu, and M. Afrad, "LEDGER: Journal Informatic and Information Technology Transfer Learning model Convolutional Neural Network menggunakan VGG-16 untuk Klasifikasi Tumor Otak pada Citra Hasil MRI," 2024.