



Klasifikasi Tumor Otak pada Citra MRI Menggunakan Transfer Learning EfficientNetB1 dan Visualisasi Grad-CAM

Suyanti¹, Chandy Ophelia S², Lies Aryani³, Prayitno⁴

¹⁻³ Sistem Informasi, Fakultas Ilmu Komputer, Universitas Dinamika Bangsa, Indonesia

⁴ Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dinamika Bangsa, Indonesia

Email: suyanti272@gmail.com¹, chandyophelia94@gmail.com², liesaryani6@gmail.com³, prayitno4704@gmail.com⁴

Alamat: Jl. Jenderal Sudirman, Tambak Sari, Thehok, Kota Jambi, Jambi, 36138

*Penulis Korespondensi: suyanti272@gmail.com

Abstract. *Magnetic resonance imaging (MRI) provides rich anatomical contrast for brain tumor assessment, yet routine interpretation remains time-intensive and demands high precision. This work develops a pipeline for four-class brain MRI image classification (glioma, meningioma, pituitary tumor, and no tumor) by combining automated brain-region cropping, data augmentation, and transfer learning with EfficientNetB1. Experimental results demonstrate exceptional performance, achieving an overall accuracy of 0.99 (99%) on the test set. Specifically, the model reached an F1-score of 1.00 for the no tumor class, 0.99 for pituitary, and 0.98 for both glioma and meningioma classes. Beyond reporting numerical performance, the study utilizes Grad-CAM heatmaps to verify that predictions rely on clinically plausible regions rather than spurious background cues. These results indicate that an efficiency-oriented backbone, paired with systematic preprocessing, can achieve reliable and interpretable performance for brain tumor classification tasks.*

Keywords: *brain tumor, MRI, EfficientNet, transfer learning, Grad-CAM, pipeline.*

Abstrak: *Magnetic resonance imaging (MRI) menyediakan kontras anatomi yang kaya untuk penilaian tumor otak, tetapi interpretasi rutin tetap memerlukan waktu dan ketelitian tinggi. Penelitian ini membangun pipeline klasifikasi empat kelas citra MRI otak (glioma, meningioma, pituitary, dan tidak ada tumor) dengan menggabungkan pemotongan area otak otomatis, augmentasi data, serta transfer learning menggunakan EfficientNetB1. Hasil eksperimen menunjukkan performa yang sangat kuat dengan raihan akurasi keseluruhan mencapai 0,99 (99%) pada data uji. Secara spesifik, model mencapai skor F1 sebesar 1,00 pada kelas tidak ada tumor, 0,99 pada kelas pituitary, serta 0,98 untuk kelas glioma dan meningioma. Selain laporan kuantitatif, penggunaan visualisasi Grad-CAM membuktikan bahwa prediksi model berlandaskan pada wilayah anatomi yang masuk akal secara klinis. Temuan ini menunjukkan bahwa kombinasi arsitektur yang efisien dan preprocessing yang tepat mampu menghasilkan sistem deteksi tumor otak yang akurat sekaligus dapat diinterpretasikan secara visual*

Kata kunci: tumor otak, MRI, EfficientNet, transfer learning, Grad-CAM, pipeline.

1. LATAR BELAKANG

Perkembangan terbaru dalam klasifikasi tumor sistem saraf pusat menegaskan bahwa diagnosis modern tidak lagi mengandalkan histologi semata, melainkan mengintegrasikan informasi molekuler untuk meningkatkan konsistensi klasifikasi dan relevansi klinisnya (Louis et al., 2021). Perubahan paradigma tersebut berdampak langsung pada praktik klinis karena penentuan subtype yang lebih presisi memengaruhi pilihan terapi, stratifikasi risiko, serta perencanaan tindak lanjut pasien (Horbinski et al., 2022). Dalam konteks ini, MRI menjadi modalitas utama untuk evaluasi tumor otak karena mampu menggambarkan perbedaan jaringan lunak secara detail tanpa paparan radiasi pengion, sehingga cocok untuk pemantauan berulang (Abd-Ellah et al., 2019).

Namun, volume data MRI yang besar dan keragaman tampilan lesi menempatkan beban kognitif yang tinggi pada pembacaan manual dan dapat memunculkan variasi antar pembaca, terutama pada kasus dengan batas lesi yang tidak tegas (Abd-Ellah et al., 2019). Pendekatan *computer-aided* diagnosis berbasis *deep learning* kemudian berkembang pesat karena *CNN* mampu mengekstraksi fitur hierarkis langsung dari citra dan secara konsisten menunjukkan kinerja kuat pada berbagai tugas analisis citra medis (Litjens et al., 2017). Pada klasifikasi tumor otak, studi-studi terkini memperlihatkan bahwa *CNN* efektif untuk membedakan pola lesi yang kompleks pada MRI, meskipun tantangan seperti kemiripan morfologi antar kelas masih sering muncul (Saeedi et al., 2023).

keterbatasan ukuran dataset berlabel pada citra medis menjadikan *transfer learning* sebagai strategi yang praktis, karena bobot pra-latih dapat mempercepat konvergensi dan meningkatkan performa ketika data pelatihan tidak sebesar dataset natural (Saeedi et al., 2023). Pemanfaatan representasi dari ImageNet, sebagai salah satu benchmark terbesar untuk pembelajaran visual, sering kali menjadi titik awal yang efektif untuk berbagai tugas turunan, termasuk pada ranah medis (Listiani et al., 2025). Atas pertimbangan efisiensi, *EfficientNet* menawarkan kerangka penskalaan yang seimbang sehingga dapat mencapai akurasi tinggi dengan biaya komputasi yang relatif rendah dibandingkan peningkatan parameter yang tidak terkontrol (Tan & Le, 2019).

Meski performa klasifikasi penting, adopsi di ranah klinis juga menuntut interpretabilitas agar pengguna dapat menilai alasan di balik keputusan model dan mendeteksi kemungkinan bias atau artefak (Tjoa & Guan, 2021). Grad-CAM merupakan pendekatan yang populer karena menghasilkan peta panas berbasis gradien yang menunjukkan lokasi kontribusi fitur terhadap prediksi kelas target pada *CNN*, sehingga lebih mudah diaudit secara visual. Penelitian ini menggunakan dataset MRI terbuka untuk mengimplementasikan klasifikasi empat kelas menggunakan *EfficientNetB1* dan mengaitkan hasilnya dengan analisis penjelasan Grad-CAM sebagai langkah menuju model yang tidak hanya akurat, tetapi juga dapat dipercaya (Selvaraju et al., 2020).

2. KAJIAN TEORITIS

Dalam tinjauan umum deep learning untuk analisis citra medis, *CNN* dianggap sebagai pendekatan dominan karena mampu menggeneralisasi pola visual melalui representasi hierarkis yang dipelajari dari data, sehingga mengurangi ketergantungan pada rekayasa fitur manual (D A et al., 2024). Konsep ini sejalan dengan fondasi teoretis deep learning yang

menjelaskan bagaimana jaringan berlapis mempelajari representasi dari fitur sederhana menuju konsep yang lebih abstrak, terutama pada data citra (Mada Sanjaya, 2023). Keberhasilan awal *CNN* pada klasifikasi berskala besar menunjukkan bahwa ketika arsitektur dan optimisasi tepat, model dapat mempelajari representasi yang cukup general untuk ditransfer ke domain lain (Krizhevsky et al., 2017).

Pada pengembangan arsitektur, ResNet memperkenalkan *residual connection* yang membantu pelatihan jaringan yang jauh lebih dalam dengan menjaga aliran *gradien*, sehingga menjadi salah satu *baseline* kuat pada banyak studi visi komputer (Shafiq & Gu, 2022). Di sisi lain, *VGG* menunjukkan bahwa tumpukan konvolusi kecil yang konsisten dapat menghasilkan performa kompetitif dan memberikan desain yang sederhana namun efektif sebagai pembanding (Insani & Santoso, 2024). Untuk kasus ketika segmentasi diperlukan, *U-Net* menjadi arsitektur klasik di biomedis karena memadukan informasi konteks global dan detail lokal melalui *skip connection* (Sasongko et al., 2024).

Di bidang segmentasi yang lebih modern, *nnU-Net* memperlihatkan bahwa konfigurasi yang mampu menyesuaikan diri terhadap dataset dapat mengurangi kebutuhan *tunning* manual, sehingga sering dipakai sebagai tolok ukur pada tantangan segmentasi biomedis (Isensee et al., 2021). Walaupun penelitian ini berfokus pada klasifikasi, perkembangan segmentasi relevan karena dapat menjadi tahap lanjutan untuk memperkirakan lokasi serta batas lesi, yang berpotensi meningkatkan interpretasi klinis (Isensee et al., 2021). Dari sisi efisiensi model, *EfficientNet* mengusulkan *compound scaling* yang mengatur kedalaman, lebar, dan resolusi secara terkoordinasi, sehingga menghasilkan *trade-off* akurasi-komputasi yang kuat pada berbagai ukuran model (Harahap & Husein, 2024).

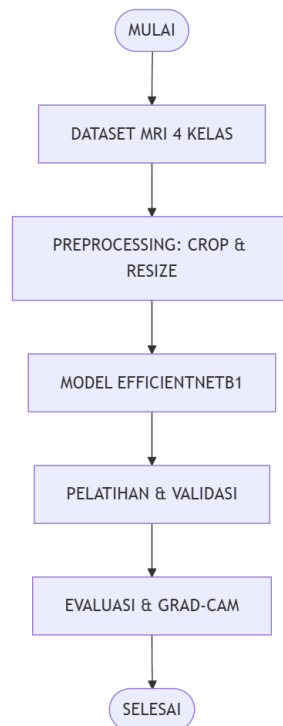
Transfer learning pada citra medis telah didokumentasikan secara luas, termasuk bagaimana karakteristik dataset dan strategi *fine-tuning* memengaruhi performa pada sistem *CAD*, terutama saat data berlabel tidak besar (Shin et al., 2016). Untuk mengurangi *overfitting*, *dropout* tetap menjadi teknik regularisasi yang umum karena mendorong jaringan membangun representasi yang lebih robust melalui pemadaman neuron secara acak selama pelatihan (Rumala, 2024). Selain itu, *Adam* sering dipilih pada pelatihan jaringan dalam karena memadukan momentum dan adaptasi laju pembelajaran per-parameter, yang secara *empiris* memudahkan optimisasi pada banyak konfigurasi (Das & Das, 2025).

Augmentasi data merupakan komponen penting ketika variasi data terbatas, karena transformasi yang wajar dapat memperkaya distribusi data dan meningkatkan generalisasi tanpa menambah biaya anotasi (Yang et al., 2022). Evaluasi model juga perlu memperhatikan ketidakseimbangan kelas, sebab dominasi kelas tertentu dapat membuat akurasi keseluruhan

terlihat tinggi meskipun performa pada kelas minoritas buruk (Haibo He & Garcia, 2009). Karena itu, metrik seperti *precision*, *recall*, dan *F1-score* disarankan untuk menilai perilaku klasifier secara lebih komprehensif pada tugas klasifikasi (Sokolova & Lapalme, 2009).

3. METODOLOGI PENELITIAN

Penelitian ini menggunakan dataset citra MRI otak empat kelas (glioma, meningioma, *pituitary*, dan tidak ada tumor) yang tersedia secara terbuka dan disusun dalam struktur folder per kelas. Berikut ini flowchart alur penelitian

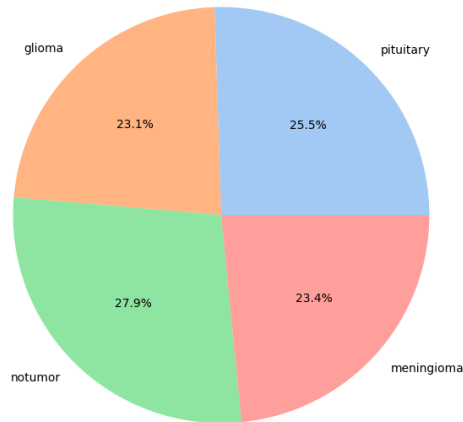


Gambar 1. Flochart Alur Penelitian

Gambar 1 mengilustrasikan alur sistematis penelitian yang dimulai dari pengolahan dataset MRI 4 kelas melalui tahapan *preprocessing* yang komprehensif. Pada tahap ini, dilakukan pemotongan ROI berbasis kontur untuk memfokuskan masukan pada area anatomi otak dan *resizing* menjadi 240x240 piksel guna menyeimbangkan detail visual dengan efisiensi komputasi. Selanjutnya, data diproses menggunakan arsitektur EfficientNetB1 melalui fase pelatihan dan validasi yang dilengkapi dengan mekanisme *callback* serta *early stopping* untuk mencegah *overfitting*. Alur penelitian ini diakhiri dengan tahap evaluasi yang tidak hanya mengandalkan metrik kuantitatif, tetapi juga menyertakan visualisasi Grad-CAM untuk memberikan interpretasi visual yang transparan terhadap hasil prediksi model. Tabel 1 konfigurasi pelatihan

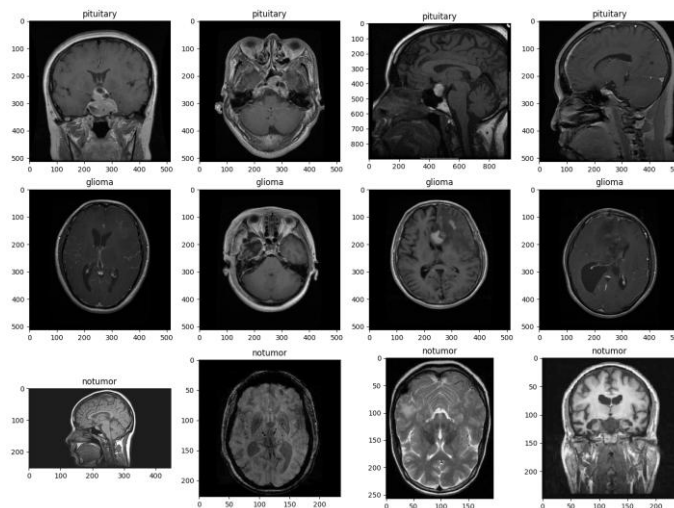
4. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil eksperimen beserta pembahasan yang menekankan hubungan antara karakteristik data, desain pipeline, perilaku pelatihan, dan pola kesalahan model.



Gambar 2. Distribusi kelas pada dataset (glioma, meningioma, notumor, pituitary).

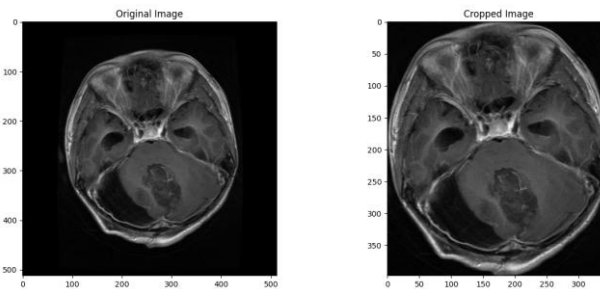
Gambar 2 memperlihatkan bahwa jumlah sampel antar kelas berada pada rentang yang relatif berdekatan, sehingga model tidak langsung terdorong untuk selalu memilih kelas tertentu demi mengoptimalkan akurasi. Meski demikian, perbedaan proporsi tetap dapat memengaruhi pembelajaran misalnya kelas dengan sampel sedikit lebih banyak berpotensi memberi sinyal gradien lebih sering selama epoch. Dalam praktiknya, kondisi ini mendorong penggunaan metrik per-kelas pada tahap evaluasi, sebab akurasi agregat dapat menyamarkan penurunan performa pada kelas yang lebih sulit atau lebih jarang.



Gambar 3. Contoh citra MRI untuk masing-masing kelas pada dataset.

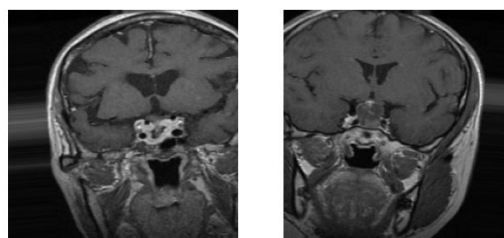
Contoh citra pada Gambar 3 menegaskan bahwa perbedaan antar kelas tidak hanya muncul sebagai “ada/tidaknya” massa, melainkan juga perbedaan lokasi, bentuk, intensitas, serta hubungan lesi dengan struktur sekitar. Sebagai ilustrasi, lesi dapat tampak sebagai area hiperintens atau hipointens tergantung sekuens dan kondisi, dan pada beberapa kasus batasnya

tidak jelas sehingga terlihat ‘membraur’ dengan jaringan sekitar. Variasi semacam ini mengapa model berbasis CNN dibutuhkan: jaringan harus menggabungkan petunjuk tekstur lokal dan konteks global secara simultan. Pada sisi lain, keragaman ini juga menjelaskan mengapa kelas tertentu lebih mudah tertukar khususnya ketika pola intensitas atau lokasi lesi memiliki kemiripan yang kuat.



Gambar 4. Contoh citra sebelum pemotongan ROI (sebelum *cropping*).

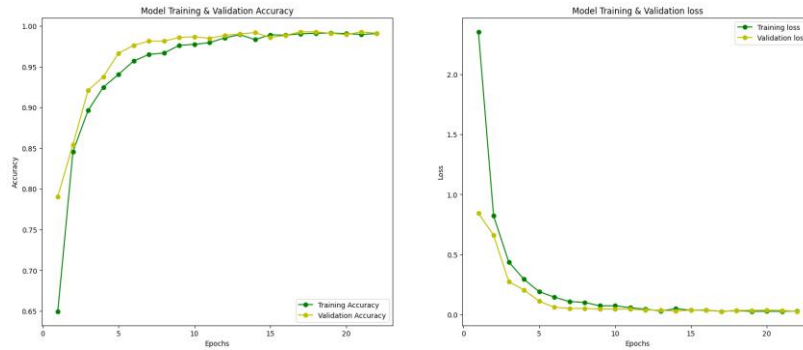
Perbandingan sebelum dan sesudah *cropping* pada Gambar 4 menunjukkan perubahan yang substantif pada komposisi masukan: area otak menjadi lebih dominan, sementara ruang kosong di luar kepala yang sebelumnya memenuhi sebagian besar piksel berkurang drastis. Secara pembelajaran, hal ini penting karena CNN belajar dari statistik piksel; jika background terlalu dominan, model berisiko mempelajari ‘*shortcut*’ yang tidak relevan (misalnya kontras tepi atau pola noise tertentu pada background) alih-alih struktur anatomi. *Cropping* juga membuat skala relatif struktur otak lebih konsisten, sehingga jaringan tidak perlu menghabiskan kapasitas untuk beradaptasi dengan variasi framing yang sebenarnya tidak terkait label. Namun, perlu dicatat bahwa *cropping* berbasis kontur dapat gagal pada citra dengan artefak kuat atau kontras buruk; karena itu, pemeriksaan kualitas hasil *cropping* tetap disarankan sebelum pelatihan skala besar.



Gambar 5. Contoh hasil augmentasi citra selama pelatihan.

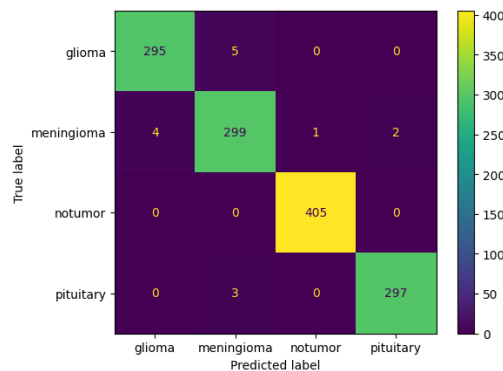
Gambar 5 menampilkan contoh transformasi yang diterapkan selama pelatihan. Secara konseptual, augmentasi membantu model ‘melihat’ variasi yang mungkin terjadi pada data nyata, seperti pergeseran posisi kepala, perbedaan orientasi, dan variasi framing. Transformasi yang digunakan sengaja dibuat ringan agar label klinis tidak berubah. misalnya rotasi kecil masih merepresentasikan lesi yang sama, hanya dengan orientasi berbeda. Dampak praktisnya

adalah model menjadi lebih robust, karena ia tidak mengasosiasikan label hanya pada posisi absolut lesi dalam citra, tetapi juga belajar pola yang lebih stabil seperti tekstur dan bentuk relatif. Pada tahap ini, keseimbangan penting: augmentasi yang terlalu agresif dapat merusak struktur halus yang sebenarnya krusial untuk membedakan kelas.



Gambar 6. Kurva akurasi dan loss pelatihan serta validasi selama *epoch*.

Kurva akurasi dan loss pada Gambar 6 memperlihatkan konvergensi yang stabil: *loss* menurun tajam pada epoch awal, lalu melandai ketika model mulai mencapai representasi yang memadai untuk membedakan kelas. Jarak antara kurva *training* dan *validation* relatif kecil pada fase akhir, yang mengindikasikan bahwa regularisasi dan strategi penghentian pelatihan bekerja dengan baik dalam mencegah *overfitting*. Pola ini juga menunjukkan bahwa model tidak sekadar menghafal contoh pelatihan; ia mampu mempertahankan performa ketika dihadapkan pada data validasi yang tidak dilihat selama update bobot. Jika jarak *training-validation* membesar pada fase akhir, itu biasanya tanda bahwa model mulai menghafal *noise* atau artefak; pada eksperimen ini, fenomena tersebut tampak terkendali.



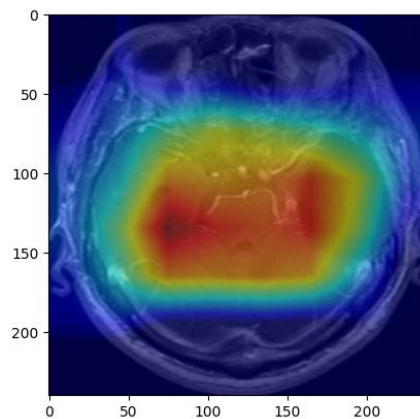
Gambar 7. Confusion Matrix

Confusion matrix pada Gambar 7 menunjukkan mayoritas prediksi berada pada diagonal utama, sehingga menegaskan performa model yang sangat baik pada data uji; kesalahan yang muncul relatif kecil dan terutama berupa pertukaran prediksi antara glioma dan meningioma, yang mengindikasikan adanya kemiripan karakteristik visual pada kedua kelas tersebut sehingga pemisahan menjadi lebih menantang bagi model

Tabel 1. menyajikan ringkasan metrik n pada evaluasi data uji.

Kelas	Precision	Recall	F1-score	Support
Glioma	0.99	0.98	0.98	300
Meningioma	0.97	0.98	0.98	306
No tumor	1.00	1.00	1.00	405
Pituitary	0.99	0.99	0.99	300
Akurasi (overall)	-	-	0.99	1311
Macro avg	0.99	0.99	0.99	1311
Weighted avg	0.99	0.99	0.99	1311

Berdasarkan Tabel 1, model klasifikasi menunjukkan performa yang sangat impresif dengan akurasi keseluruhan mencapai 0,99. Kelas "No tumor" berhasil diidentifikasi dengan sempurna (*F1-score 1,00*), yang mengindikasikan bahwa sistem sangat andal dalam membedakan kondisi normal tanpa risiko *false alarm*. Nilai *precision* dan *recall* yang seimbang dan tinggi di seluruh kelas, termasuk skor 0,99 untuk Pituitary serta 0,98 untuk Glioma dan Meningioma, membuktikan kemampuan model dalam meminimalkan kesalahan diagnosis. Hal ini sangat krusial dalam konteks klinis, di mana kesalahan kecil pada Glioma dan Meningioma yang hanya berupa "salah tukar" antar jenis tumor jauh lebih aman dibandingkan kesalahan fatal yang mengategorikan tumor sebagai kondisi normal.



Gambar 9. Visualisasi Grad-CAM pada citra uji.

Visualisasi Grad-CAM pada Gambar 9 memperdalam pemahaman tentang ‘alasan’ model. Idealnya, peta panas menyoroti area lesi atau struktur sekitar yang relevan dengan diagnosis, bukan wilayah acak atau pinggiran citra. Pada contoh yang baik, *heatmap* terkonsentrasi pada area yang secara visual tampak abnormal, mengindikasikan bahwa model benar-benar memanfaatkan sinyal yang bermakna. jika *heatmap* konsisten menyorot tepi kepala, teks *overlay*, atau artefak tertentu, maka perlu dicurigai bahwa model belajar *shortcut*. Dalam penelitian ini, peta yang dihasilkan secara umum menunjukkan fokus yang lebih masuk akal, sehingga mendukung interpretasi bahwa performa tinggi tidak semata-mata berasal dari kebetulan pada background.

5. KESIMPULAN DAN SARAN

Implementasi *EfficientNetB1* dengan optimasi ROI dan augmentasi data berhasil menghasilkan model klasifikasi tumor otak yang sangat akurat dengan akurasi 0,99. Model ini menunjukkan performa luar biasa dengan F1-score sempurna 1,00 pada kelas "No tumor", serta skor 0,99 untuk Pituitary dan 0,98 untuk Glioma serta Meningioma. Keunggulan ini diperkuat oleh visualisasi Grad-CAM yang secara akurat menyoroti area lesi secara klinis, memberikan transparansi visual yang memastikan bahwa setiap prediksi didasarkan pada fitur anatomi yang relevan dan dapat diandalkan bagi diagnosis medis. Penelitian lanjutan disarankan menggunakan dataset multi-sumber (berbeda rumah sakit/sekuens MRI) agar model lebih *robust* dan generalisasi meningkat.

6. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Dinamika Bangsa serta pihak-pihak yang telah memberikan dukungan selama proses penyusunan penelitian ini.

DAFTAR PUSTAKA

- Abd-Ellah, M. K., Awad, A. I., Khalaf, A. A. M., & Hamed, H. F. A. (2019). A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned. *Magnetic Resonance Imaging*, *61*, 300–318. <https://doi.org/10.1016/j.mri.2019.05.028>
- D A, A., Shekar, M. S., Bharadwaj, A., Vineeth, N., & Neelima, M. L. (2024). Deep Learning in Medical Image Analysis: A Survey. *2024 International Conference on Innovation and Novelty in Engineering and Technology (INNOVA)*, 1–5. <https://doi.org/10.1109/INNOVA63080.2024.10847040>
- Das, S., & Das, S. (2025). *A Comparative Analysis of Optimization Methods for Classification on Various Datasets*.
- Haibo He, & Garcia, E. A. (2009). Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263–1284. <https://doi.org/10.1109/TKDE.2008.239>
- Harahap, M., & Husein, A. M. (2024). Penerapan Efficient-Net Dalam Mengklasifikasi Kanker Kulit. *PUBLIS PENERBIT UNPRI PRESS*, *1*(1). <https://jurnal.unprimdn.ac.id/index.php/isbn/article/view/5405>
- Horbinski, C., Berger, T., Packer, R. J., & Wen, P. Y. (2022). Clinical implications of the 2021 edition of the WHO classification of central nervous system tumours. *Nature Reviews Neurology*, *18*(9), 515–529. <https://doi.org/10.1038/s41582-022-00679-w>
- Insani, M. K., & Santoso, D. B. (2024). Perbandingan Kinerja Model Pre-Trained CNN (VGG16, RESNET, dan INCEPTIONV3) untuk Aplikasi Pengenalan Wajah pada Sistem Absensi Karyawan. *Jurnal Indonesia : Manajemen Informatika Dan Komunikasi*, *5*(3), 2612–2622. <https://doi.org/10.35870/jimik.v5i3.925>

- Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2), 203–211. <https://doi.org/10.1038/s41592-020-01008-z>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>
- Listiani, H., Asia, S. N., Sepriano, S., & Judijanto, L. (2025). *Deep Learning: Konsep, Arsitektur, dan Implementasi*. PT. Sonpedia Publishing Indonesia.
- Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). *A Survey on Deep Learning in Medical Image Analysis*. <https://doi.org/10.48550/ARXIV.1702.05747>
- Louis, D. N., Perry, A., Wesseling, P., Brat, D. J., Cree, I. A., Figarella-Branger, D., Hawkins, C., Ng, H. K., Pfister, S. M., Reifenberger, G., Soffietti, R., Von Deimling, A., & Ellison, D. W. (2021). The 2021 WHO Classification of Tumors of the Central Nervous System: A summary. *Neuro-Oncology*, 23(8), 1231–1251. <https://doi.org/10.1093/neuonc/noab106>
- Mada Sanjaya, W. S. (2023). *Deep Learning Citra Medis Berbasis Pemrograman Python-Penerbit Bolabot*. Bolabot.
- Rumala, D. J. (2024). *Brain Disease Classification in MRI Images Using Deep Learning Based on Deep-Stacked Models and Enhanced Feature Representation*. Institut Teknologi Sepuluh Nopember.
- Saeedi, S., Rezayi, S., Keshavarz, H., & R. Niakan Kalhori, S. (2023). MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. *BMC Medical Informatics and Decision Making*, 23(1), 16. <https://doi.org/10.1186/s12911-023-02114-6>
- Sasongko, A., Maulana, M. S., Mustopa, A., & Nugraha, W. (2024). Automatic Wound Image Segmentation with U-Net Model for Smartphone Application. *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, 10(2), 267. <https://doi.org/10.26418/jp.v10i2.78548>
- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2020). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *International Journal of Computer Vision*, 128(2), 336–359. <https://doi.org/10.1007/s11263-019-01228-7>
- Shafiq, M., & Gu, Z. (2022). Deep Residual Learning for Image Recognition: A Survey. *Applied Sciences*, 12(18), 8972. <https://doi.org/10.3390/app12188972>
- Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. M. (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298. <https://doi.org/10.1109/TMI.2016.2528162>
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>
- Tan, M., & Le, Q. V. (2019). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. <https://doi.org/10.48550/ARXIV.1905.11946>

- Tjoa, E., & Guan, C. (2021). A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI. *IEEE Transactions on Neural Networks and Learning Systems*, 32(11), 4793–4813. <https://doi.org/10.1109/TNNLS.2020.3027314>
- Yang, S., Xiao, W., Zhang, M., Guo, S., Zhao, J., & Shen, F. (2022). *Image Data Augmentation for Deep Learning: A Survey* (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2204.08610>