

Prediksi Tumor Otak Menggunakan Metode Convolutional Neural Network (CNN) dan Algoritma Decision Tree

Azis Nurhidayat¹, Wildan Ahmad Arroseyid², Riza Samsinar³

^{1,2,3}Teknik Elektro, Fakultas Teknik, Universitas Muhammadiyah Jakarta

E-mail: ¹2021041020004@student.umj.ac.id ²20210410200011@student.umj.ac.id, ³riza.samsinar@umj.ac.id*

Abstrak – Berkembangnya teknologi memberikan banyak pengaruh signifikan di berbagai sektor, termasuk di bidang kesehatan. Salah satu permasalahan yang menjadi fokus adalah deteksi dan klasifikasi tumor otak. Klasifikasi tumor otak merupakan penelitian yang sangat penting untuk membantu memprediksi apakah seseorang terinfeksi tumor atau tidak, sehingga dapat mendukung pengambilan keputusan klinis yang lebih baik. Dalam klasifikasi citra medis, algoritma Machine Learning telah diterapkan secara luas, di mana metode seperti Decision Tree dan Convolutional Neural Network (CNN) menjadi yang paling sering digunakan untuk analisis citra medis. Penelitian ini bertujuan untuk mengevaluasi kinerja CNN dalam mendeteksi tumor otak dari citra MRI. Dataset yang digunakan terdiri dari dua kelas: "No Tumor" dan "Yes Tumor". Model CNN dirancang untuk melakukan klasifikasi secara end-to-end, sementara fitur yang diekstraksi dari CNN digunakan untuk melatih model Decision Tree sebagai pendekatan hybrid. Proses pelatihan melibatkan teknik optimasi dan evaluasi untuk memastikan hasil yang maksimal. Hasil penelitian menunjukkan bahwa model CNN sebagai model end-to-end berhasil mencapai akurasi sebesar 96% pada data validasi, sedangkan pendekatan hybrid dengan ekstraksi fitur menggunakan CNN dan klasifikasi oleh Decision Tree menghasilkan akurasi sebesar 92%. Studi ini menyoroti potensi besar CNN dalam diagnosis berbasis citra medis, khususnya untuk tumor otak, sekaligus merekomendasikan eksplorasi lebih lanjut menggunakan metode seperti augmentasi data, transfer learning, dan pengujian pada dataset yang lebih besar dan beragam.

Kata Kunci — CNN, Decision Tree, Pendeteksi Tumor Otak

1. PENDAHULUAN

Tumor otak adalah suatu penyakit yang menyerang bagian saraf pada lesi ekspansif otak dengan pembentukan sel yang abnormal pada otak dan membentuk tumor dalam lesi tengkorak kepala (intra cranial) atau di sumsum tulang belakang (medula spinalis). Berdasarkan asal usulnya, tumor otak dibagi menjadi dua jenis utama yaitu tumor primer, yang berkembang langsung dari jaringan otak, dan tumor sekunder atau metastasis otak, yang disebabkan oleh penyebaran kanker dari organ tubuh lainnya. Tumor otak dapat diklasifikasikan menjadi dua jenis utama: tumor primer dan tumor sekunder. Tumor primer adalah tumor yang berasal dari jaringan otak itu sendiri, sedangkan tumor sekunder, yang lebih sering disebut sebagai metastasis otak, merupakan tumor yang berasal dari organ tubuh lain dan menyebar ke otak.[1]

Secara umum, tumor otak dapat dibedakan menjadi beberapa jenis berdasarkan asal usul sel-selnya. **Glioma**, yang melibatkan sel glial sebagai sumber tumor, adalah salah satu jenis yang paling umum, dan dapat meliputi berbagai sub tipe seperti astrositoma, oligodendroglioma, dan glioblastoma multiforme (GBM), yang dikenal dengan tingkat keganasannya yang tinggi. Selain itu, **Meningioma**, yang berkembang dari membran pelindung otak dan sumsum tulang belakang, serta **Schwannoma**, yang berasal dari sel Schwann pada saraf perifer, juga merupakan jenis yang sering ditemukan. Setiap jenis tumor memiliki karakteristik dan tingkat keparahan yang berbeda, mulai dari tumor jinak yang tidak menyebar hingga tumor ganas yang dapat menyebabkan kerusakan otak permanen atau kematian jika tidak segera ditangani. Diagnosis tumor otak memerlukan teknologi pencitraan medis, dengan Magnetic Resonance Imaging (MRI) menjadi metode utama untuk mendeteksi dan mengkarakterisasi lesi tumor. Meskipun teknik MRI memiliki akurasi tinggi, proses analisis citra seringkali dilakukan secara manual oleh ahli radiologi, yang memakan waktu, bersifat subjektif, dan rentan terhadap kesalahan manusia. Tantangan ini semakin meningkat seiring dengan kebutuhan akan diagnosis yang lebih cepat dan akurat untuk menangani kasus tumor otak yang kompleks.[2], [3]

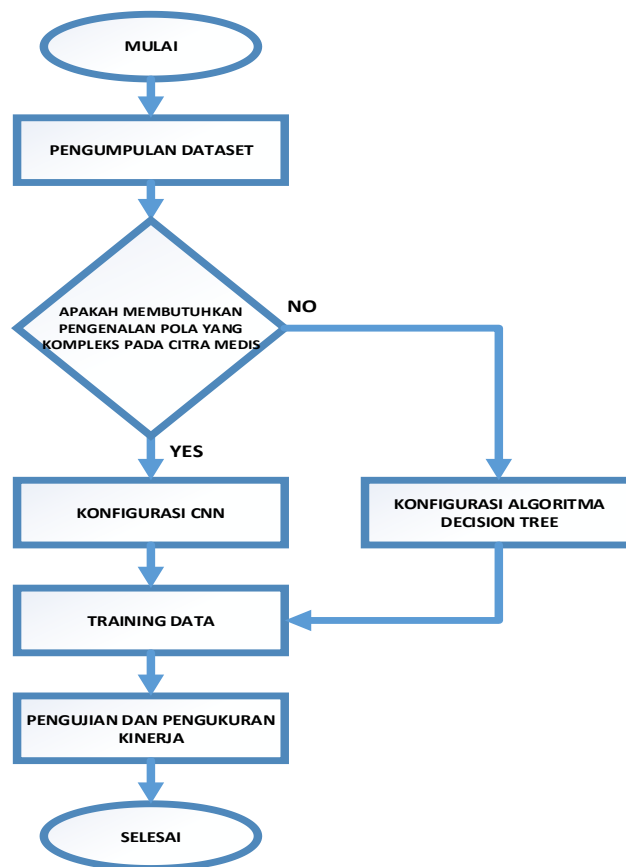
Perkembangan teknologi kecerdasan buatan (Artificial Intelligence/AI), khususnya Deep Learning, telah membuka peluang baru dalam pengolahan citra medis. Salah satu metode paling menonjol dalam Deep Learning adalah Convolutional Neural Network (CNN), yang dirancang untuk mengekstrak fitur-fitur penting dari citra secara otomatis. CNN memiliki keunggulan dalam menangkap pola-pola kompleks pada data citra, sehingga sangat efektif untuk mendeteksi dan mengklasifikasikan tumor otak berdasarkan hasil pencitraan MRI. Namun, meskipun CNN menawarkan akurasi yang tinggi, model ini memiliki keterbatasan, seperti kebutuhan akan dataset besar untuk pelatihan, kompleksitas komputasi yang tinggi, dan potensi overfitting jika tidak diimbangi dengan teknik regulasi yang tepat. Oleh karena itu, pendekatan hybrid, yang menggabungkan CNN sebagai alat ekstraksi fitur dengan algoritma klasifikasi tradisional seperti Decision Tree, muncul sebagai solusi potensial. Pendekatan ini memungkinkan pengurangan kompleksitas komputasi sambil tetap mempertahankan performa yang baik dalam

tugas klasifikasi. CNN bertugas menghasilkan representasi fitur yang kaya dari citra MRI, sedangkan Decision Tree digunakan untuk mengambil keputusan klasifikasi berdasarkan fitur-fitur tersebut. [4]

Studi sebelumnya menunjukkan bahwa penggunaan CNN pada berbagai jenis diagnosis medis telah menghasilkan akurasi yang signifikan dibandingkan dengan metode tradisional. Namun, model berbasis CNN memerlukan sumber daya komputasi tinggi dan dataset besar untuk pelatihan optimal. Di sisi lain, metode hybrid seperti kombinasi CNN dan Decision Tree menawarkan potensi pengurangan kompleksitas komputasi tanpa mengorbankan performa secara signifikan. Pendekatan ini memungkinkan penggunaan CNN sebagai alat ekstraksi fitur untuk memberikan representasi data yang lebih baik sebelum dilanjutkan dengan klasifikasi berbasis algoritma yang lebih ringan seperti Decision Tree.[2]

Penelitian ini bertujuan untuk mengevaluasi efektivitas model berbasis CNN dalam mendeteksi tumor otak melalui dua pendekatan utama: **pendekatan end-to-end**, di mana CNN digunakan secara langsung untuk klasifikasi, dan **pendekatan hybrid**, di mana CNN digunakan sebagai ekstraktor fitur sebelum dikombinasikan dengan Decision Tree untuk klasifikasi. Kedua pendekatan ini dibandingkan berdasarkan metrik performa, seperti akurasi, presisi, recall, dan f1-score, untuk menentukan pendekatan yang paling efisien dan akurat dalam diagnosis tumor otak. Selain itu, penelitian ini juga berusaha menjawab tantangan utama dalam implementasi AI di bidang medis, seperti interpretabilitas model, kebutuhan akan data yang beranotasi, serta efisiensi komputasi. Dengan harapan, hasil penelitian ini dapat memberikan kontribusi nyata dalam meningkatkan kualitas diagnosis tumor otak, mempercepat proses pengambilan keputusan medis, serta menjadi langkah awal bagi pengembangan sistem diagnosis berbasis AI yang lebih canggih di masa depan.[5]

2. METODE PENELITIAN



Gambar 1. Tahapan Penelitian

2.1 Dataset

Penelitian ini menggunakan Dataset yang terdiri dari dua kelas, "No Tumor" dan "Yes Tumor", dengan citra yang dinormalisasi dan diubah ukurannya menjadi 128x128 piksel. Data dibagi menjadi data latih dan uji dengan rasio 80:20 yang berisi citra gambar dari MRI untuk menguji kedua metode.[4]

2.2 Implementasi

CNN dirancang dengan tiga lapisan convolusi menggunakan fungsi aktivasi ReLU dan pooling layer, diikuti oleh lapisan fully connected. Model dilatih selama 10 epoch dengan optimizer Adam dan fungsi loss sparse categorical crossentropy. Fitur dari lapisan sebelum output CNN diekstraksi dan digunakan untuk melatih Decision Tree sebagai metode klasifikasi tambahan. Model ini dievaluasi secara terpisah untuk membandingkan kinerja dengan CNN end-to-end.[6]

2.3 Pengukuran Kinerja

Kinerja dievaluasi berdasarkan akurasi, confusion matrix, classification report, serta grafik akurasi dan loss selama pelatihan. Hybrid CNN dan Kinerja Decision Tree dievaluasi menggunakan metrik yang sama dengan CNN end-to-end. Hasil dari kedua pendekatan dibandingkan untuk menilai keunggulan dan kelemahan masing-masing metode.[6]

3. HASIL DAN PEMBAHASAN

3.1 Kinerja CNN End-to-End:

Akurasi pelatihan mencapai 98%, dengan akurasi validasi sebesar 96% dan Confusion matrix menunjukkan tingkat kesalahan rendah, Model menunjukkan generalisasi yang baik pada data uji. Keberhasilan model ini didukung oleh arsitektur CNN yang dirancang dengan baik. Arsitektur terdiri dari tiga lapisan convolusi dengan fungsi aktivasi ReLU diikuti oleh lapisan pooling untuk ekstraksi fitur, dan diakhiri dengan fully connected layer untuk klasifikasi. Pendekatan ini memastikan bahwa model dapat menghasilkan representasi fitur yang kuat dan relevan untuk tugas deteksi tumor. Selain itu, penggunaan optimisasi seperti Adam optimizer membantu mempercepat konvergensi model, sehingga meningkatkan efisiensi pelatihan. Namun, model ini tetap memiliki kekurangan kecil, terutama pada kasus False Positives yang menunjukkan prediksi tumor pada citra yang sebenarnya normal. Hal ini bisa menjadi fokus pengembangan lebih lanjut untuk meningkatkan akurasi prediksi secara keseluruhan.[7]

3.2 Kinerja Hybrid CNN + Decision Tree

Metode hybrid yang menggabungkan CNN sebagai ekstraktor fitur dan Decision Tree sebagai algoritma klasifikasi menunjukkan performa yang kompetitif. Akurasi model ini mencapai 92%, sedikit lebih rendah dibandingkan pendekatan CNN end-to-end, namun dengan keunggulan dalam efisiensi komputasi. Fitur yang diekstraksi memberikan representasi yang baik untuk klasifikasi, namun hasilnya bergantung pada distribusi data. [8]

3.3 Perbandingan Model

Berdasarkan hasil evaluasi, terdapat perbedaan signifikan antara dua pendekatan yang diuji, baik dalam hal akurasi maupun efisiensi. Berikut adalah perbandingan rinci kedua metode:

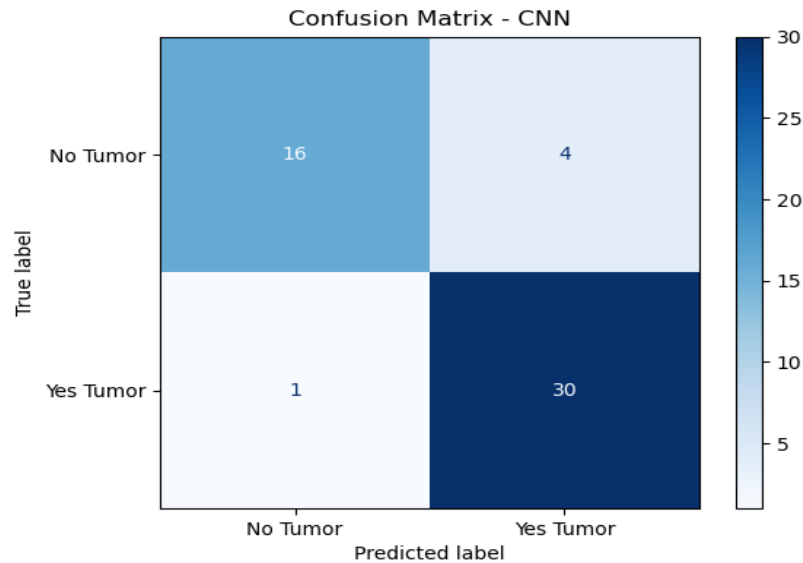
Keunggulan CNN End-to-End:

- Memiliki akurasi yang lebih tinggi (96%) dengan kesalahan prediksi yang lebih rendah.
- Mampu menangkap pola-pola non-linear yang kompleks dari citra MRI.
- Cocok untuk aplikasi dengan dataset besar dan kebutuhan akurasi tinggi.
- Arsitektur yang end-to-end memungkinkan integrasi langsung dengan sistem prediksi tanpa perlu langkah tambahan untuk ekstraksi fitur.[6]

Keunggulan Hybrid CNN + Decision Tree:

- Lebih hemat sumber daya komputasi karena proses klasifikasi dilakukan oleh Decision Tree yang relatif sederhana.
- Menyediakan interpretabilitas hasil yang lebih baik, memungkinkan analisis lebih mendalam terhadap keputusan klasifikasi.
- Cocok untuk aplikasi dengan keterbatasan infrastruktur atau dataset yang lebih kecil.[6]

Secara keseluruhan, pemilihan metode tergantung pada kebutuhan spesifik aplikasi. Pendekatan CNN end-to-end lebih disarankan untuk skenario yang membutuhkan akurasi tinggi dan kemampuan menangkap pola-pola kompleks. Sebaliknya, pendekatan hybrid lebih cocok untuk kondisi yang mengutamakan efisiensi dan interpretabilitas. Studi lanjutan dapat difokuskan pada pengembangan metode hybrid yang lebih canggih, misalnya dengan mengganti Decision Tree dengan algoritma klasifikasi yang lebih kuat seperti Random Forest atau Gradient Boosting untuk meningkatkan performa tanpa mengorbankan efisiensi.



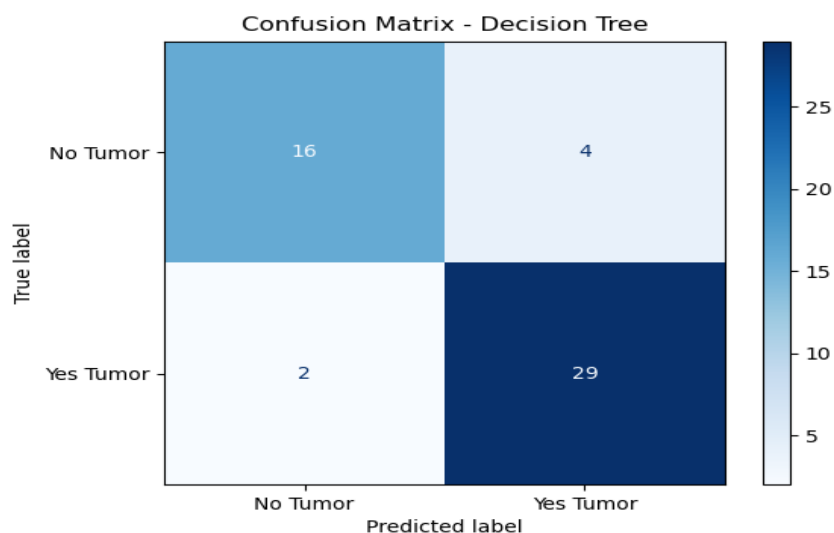
Gambar 2 Confusion Matrix Metode CNN

Hasil:

- True Positives (Yes Tumor Predicted as Yes Tumor): 30
- True Negatives (No Tumor Predicted as No Tumor): 16
- False Positives (No Tumor Predicted as Yes Tumor): 4
- False Negatives (Yes Tumor Predicted as No Tumor): 1

Kesimpulan:

Model CNN memiliki performa yang baik, dengan mayoritas prediksi benar dan sedikit kesalahan dengan True Negative 4 kali mendeteksi tumor pada gambar yang sebenarnya tidak memiliki tumor. Hasil ini menunjukkan bahwa CNN lebih sensitif terhadap mendeteksi tumor dibandingkan mengidentifikasi gambar tanpa tumor.



Gambar 3 Confusion Matrix Metode Decision Tree

Hasil:

- True Positives (Yes Tumor Predicted as Yes Tumor): 29
- True Negatives (No Tumor Predicted as No Tumor): 16
- False Positives (No Tumor Predicted as Yes Tumor): 4
- False Negatives (Yes Tumor Predicted as No Tumor): 2

Kesimpulan:

Model Decision Tree memiliki performa sedikit di bawah CNN yaitu True Negative: 2 kali gagal mendeteksi tumor (false negatives), yang dapat menjadi kritis dalam konteks medis. Namun, hasil keseluruhan cukup baik dengan distribusi prediksi yang hampir sama dengan CNN.

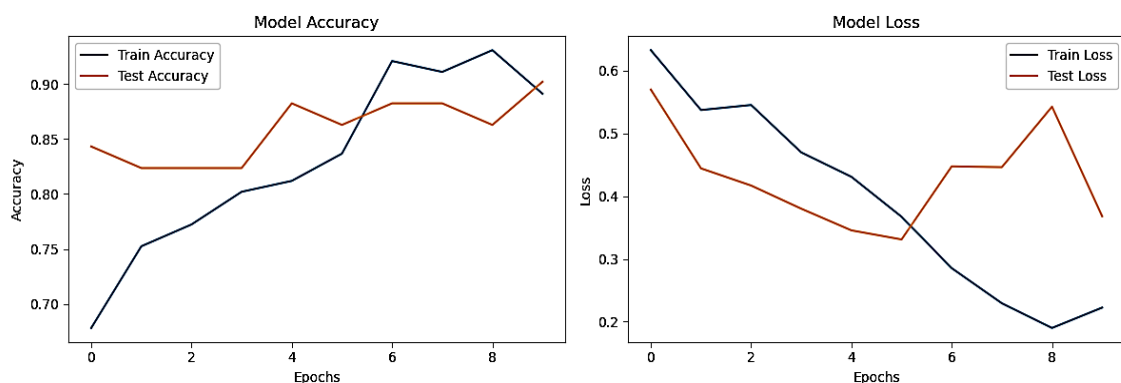
Tabel 1. Perbandingan Hasil Deteksi Menggunakan Metode CNN dan Decision Tree
Classification Report of CNN and Decision Tree:

	Precision		Recall		F1-Score		Support	
	CNN	Decision Tree	CNN	Decision Tree	CNN	Decision Tree	CNN	Decision Tree
no	0.94	0.89	0.80	0.80	0.86	0.84	20	20
yes	0.88	0.88	0.97	0.94	0.92	0.91	31	31
Accuracy	-	-	-	-	0.90	0.88	51	51
Macro Avg	0.91	0.88	0.88	0.87	0.89	0.87	51	51
Weighted Avg	0.91	0.88	0.90	0.88	0.90	0.88	51	51

Pada Tabel Perbandingan ini kita dapat melihat banyak aspek yang memiliki perbedaan pada saat melakukan prediksi kanker otak dengan menggunakan metode CNN dan Decision Tree. Berikut ini Definisi dan Komponen Utama:

- 1. Precision (Presisi)**
Presisi adalah rasio prediksi positif yang benar terhadap total prediksi positif. Ini menunjukkan seberapa banyak dari kasus yang diprediksi positif yang sebenarnya positif. Presisi yang tinggi menunjukkan bahwa model jarang membuat kesalahan dalam memprediksi kelas positif.
- 2. Recall (Pengingat)**
Recall adalah rasio prediksi positif yang benar terhadap total aktual positif. Ini mengukur kemampuan model untuk mengidentifikasi semua instance yang relevan. Recall yang tinggi menunjukkan bahwa model berhasil menangkap sebagian besar kasus positif yang sebenarnya.
- 3. F1-Score**
F1-score adalah rata-rata harmonis dari presisi dan recall. Ini memberikan keseimbangan antara kedua metrik, terutama ketika dataset tidak seimbang. F1-score yang tinggi menunjukkan bahwa model memiliki keseimbangan yang baik antara presisi dan recall.
- 4. Support**
Support adalah jumlah kejadian aktual dari setiap kelas dalam dataset. Ini memberikan konteks untuk metrik presisi dan recall. Interpretasinya Mengetahui jumlah instance untuk setiap kelas membantu dalam memahami seberapa baik model bekerja pada kelas yang berbeda.
- 5. Akurasi**
Akurasi adalah rasio dari prediksi yang benar terhadap total instance. Ini memberikan gambaran umum tentang seberapa baik model bekerja. Akurasi yang tinggi menunjukkan bahwa sebagian besar prediksi model adalah benar.
- 6. Rata-Rata Makro**
Rata-rata makro menghitung metrik secara independen untuk setiap kelas dan kemudian mengambil rata-rata, memperlakukan semua kelas sama. Rata-rata makro yang tinggi menunjukkan bahwa model bekerja dengan baik di seluruh kelas.
- 7. Rata-Rata Tertimbang**
Rata-rata tertimbang (*Weighted Average*) mempertimbangkan dukungan dari setiap kelas, memberikan pandangan yang lebih seimbang tentang kinerja model di seluruh kelas yang tidak seimbang. Rata-rata tertimbang yang tinggi menunjukkan bahwa model bekerja dengan baik secara keseluruhan, mempertimbangkan jumlah *instance* di setiap kelas.

3.4 Grafik Model Akurasi dan Model Loss



Gambar 4. Grafik Model Akurasi dan Model Loss

Keterangan Grafik:

- Sumbu X: Jumlah epoch (iterasi pelatihan).
- Sumbu Y: Untuk akurasi, nilai akurasi dari 0 hingga 1. Sedangkan untuk loss, nilai loss/error.

Penjelasan:

- Train Accuracy yang ditampilkan dengan Grafik biru menunjukkan peningkatan akurasi model pada data latih. Model berhasil belajar lebih baik seiring bertambahnya epoch.
- Test Accuracy yang ditampilkan dengan Grafik oranye menunjukkan performa model pada data uji. Stabilitas akurasi ini menunjukkan bahwa model memiliki kemampuan generalisasi yang baik terhadap data uji.
- Train Loss yang ditampilkan dengan Grafik biru menunjukkan penurunan nilai loss pada data latih. Ini menunjukkan bahwa model mempelajari pola dalam data latih.
- Test Loss yang ditampilkan dengan Grafik oranye menunjukkan loss pada data uji. Penurunan test loss ini menandakan model tidak mengalami overfitting yang signifikan.

Kesimpulannya Model CNN ini menunjukkan hasil pelatihan yang stabil, dengan akurasi di atas 90% pada data latih dan data uji, serta loss yang konsisten menurun.

4 SIMPULAN

- Kinerja Model CNN dan Decision Tree dapat memprediksi tumor otak dengan akurasi tinggi, yaitu sebesar 90% dan 88% secara berturut-turut. Hasil ini menunjukkan bahwa model tersebut efektif dalam mendeteksi tumor otak dengan data uji sebanyak 253.
- Perbandingan Model CNN menunjukkan kinerja yang lebih baik dibandingkan dengan model Decision Tree dalam memprediksi tumor otak. Hal ini disebabkan oleh kemampuan CNN dalam mengenali pola kompleks pada citra medis.
- Faktor Penentu dan Fitur-fitur yang diekstrak dari citra MRI seperti tekstur, bentuk, dan ukuran tumor berpengaruh signifikan terhadap kinerja model. Pemilihan fitur yang tepat dapat meningkatkan akurasi prediksi.
- Keterbatasan dataset dan variabilitas citra medis dapat mempengaruhi kinerja model. Oleh karena itu, perlu dilakukan pengumpulan data yang lebih besar dan beragam untuk meningkatkan akurasi prediksi.
- Implikasi dari Penelitian ini memiliki potensi untuk membantu dokter dalam mendeteksi dan mendiagnosis tumor otak secara lebih akurat dan cepat. Integrasi model CNN dan Decision Tree dapat menjadi alat bantu yang efektif dalam pengobatan tumor otak.

5 SARAN

- Pengumpulan Dataset yang Lebih Besar dan Beragam
Pengumpulan dataset yang lebih besar dan beragam dapat meningkatkan akurasi prediksi model. Hal ini dapat dilakukan dengan mengumpulkan data dari berbagai rumah sakit dan pusat kesehatan. Dan juga peningkatan dataset melalui augmentasi data untuk meningkatkan generalisasi model.
- Penggunaan Teknik Transfer Learning
Penggunaan teknik transfer learning dapat meningkatkan kinerja model dengan memanfaatkan pengetahuan yang sudah ada dari model lain. Contohnya menggunakan model pre-trained seperti VGG16 atau ResNet50.
- Pengembangan Sistem Pendukung Keputusan
Pengembangan sistem pendukung keputusan dapat membantu dokter dalam membuat keputusan yang lebih akurat dan cepat. Hal ini dapat dilakukan dengan mengintegrasikan model prediksi dengan sistem informasi kesehatan
- Eksplorasi metode klasifikasi lain seperti Random Forest atau Support Vector Machine.

DAFTAR PUSTAKA

- [1] S. D. Kamil, D. Widiyanto, and N. Chamidah, Perbandingan Metode Decision Tree Dengan Naïve Bayes Dalam Klasifikasi Tumor Otak Citra *MRI*. 2020.
- [2] M. N. Winnarto, M. Mailasari, and A. Purnamawati, Ekstraksi Fitur Dan Implementasi Machine Learning Untuk Klasifikasi Jenis Tumor Otak. 6(1), 65-70, 2024.
- [3] V. Essianda, A. D. Indrasari, P. Widyastuti, T. Syahla, and R. Rohadi, "Brain Tumor : Molecular Biology, Pathophysiology, and Clinical Symptoms," *Jurnal Biologi Tropis*, vol. 23, no. 4, pp. 260–269, Sep. 2023.
- [4] M. N. M. Hakim, A. B. Nugroho, and A. E. Minarno, "Prediksi Tumor Otak Menggunakan Metode Convolutional Neural Network," *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, vol. 17, no. 1, p. 48, Jul. 2023.
- [5] A. Chaudhuri, "Scopus Indexed Journal A Novel CNN Model for Early Prediction of Brain Cancer," 2024.

- [6] Ananda, J. S. Analisis Klasifikasi Penyakit Tumor Otak Pada Citra Radiografi Menggunakan Metode Support Vector Machines (Svm) Dengan Python (Doctoral dissertation, Universitas Jambi), 2024.
- [7] Febrianti, A. S., Sardjono, T. A., & Babgei, A. F. Klasifikasi Tumor Otak pada Citra Magnetic Resonance Image dengan Menggunakan Metode Support Vector Machine. Jurnal Teknik ITS, 9(1), A118-A123, 2020.
- [8] Gupta, R. K., Bharti, S., Kunhare, N., Sahu, Y., & Pathik, N. Brain tumor detection and classification using cycle generative adversarial networks. Interdisciplinary Sciences: Computational Life Sciences, 14(2), 485-502, 2022.