

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian, pengujian dan pembahasan yang telah dilakukan mengenai penerapan Grad-CAM pada model *deep learning* untuk interpretasi hasil deteksi pneumonia pada citra X-Ray, maka dapat ditarik kesimpulan sebagai berikut:

1. Analisis terhadap hasil visualisasi Grad-CAM membuktikan bahwa keputusan model didasarkan pada fitur klinis yang valid, yaitu area opasitas atau kekeruhan pada paru-paru, bukan pada *noise* atau gangguan latar belakang. Hal ini menunjukkan bahwa penerapan Grad-CAM efektif dalam mengurangi sifat *black-box* pada algoritma *Deep Learning*, karena memberikan transparansi bahwa model telah mempelajari fitur patologis yang relevan secara medis (biomedis) dalam mendeteksi pneumonia.
2. Penerapan metode Grad-CAM berhasil memvisualisasikan wilayah fokus (*heatmap*) yang menjadi dasar pengambilan keputusan klasifikasi oleh model. Pada citra yang terprediksi Pneumonia, visualisasi secara konsisten menunjukkan area atensi tinggi (warna merah pekat) yang terpusat pada wilayah paru-paru. Sebaliknya, pada citra Normal, pola atensi terlihat menyebar ke area luar paru-paru atau latar belakang, yang menandakan tidak ditemukannya indikasi visual kelainan pada organ tersebut.
3. Model *deep learning* yang dikembangkan menunjukkan kinerja yang kuat dalam mengklasifikasikan citra X-Ray. Berdasarkan metrik evaluasi pada data uji, model mencapai akurasi sebesar 90,38%. Secara khusus, model memiliki sensitivitas yang sangat baik untuk kelas Pneumonia dengan nilai *Recall* sebesar 0,93 (dari skala 1,00). Nilai *recall* yang tinggi ini menunjukkan efektivitas model dalam meminimalkan kesalahan negatif palsu (*False Negative*), sehingga model dinilai aman dan potensial untuk digunakan sebagai alat bantu skrining medis awal.

5.2 Saran

Penelitian ini masih memiliki beberapa keterbatasan, yang dapat diperbaiki dengan penelitian lebih lanjut. Keterbatasan utama adalah keanekaragaman dan jumlah data citra sinar-X dada yang digunakan. Agar model dapat generalisasi yang lebih baik terhadap berbagai kondisi pasien dan perbedaan perangkat pencitraan medis, penelitian selanjutnya disarankan untuk memanfaatkan dataset yang lebih besar dan berasal dari berbagai sumber. Selain itu, penelitian ini hanya menggunakan satu arsitektur *deep learning*, DenseNet121, dan satu metode interpretabilitas, Grad-CAM. Penelitian lanjutan juga dapat membandingkan arsitektur lain, seperti EfficientNet atau ResNet, dan menggabungkan Grad-CAM dengan metode AI yang dapat dijelaskan lainnya untuk mendapatkan analisis interpretabilitas yang lebih mendalam. Terakhir, evaluasi interpretasi model penelitian ini belum melibatkan validasi dari tenaga medis. Oleh karena itu, disarankan agar penelitian selanjutnya melibatkan radiolog atau tenaga medis untuk menilai apakah hasil visualisasi Grad-CAM selaras dengan temuan klinis. Ini akan memastikan bahwa sistem yang dikembangkan memiliki tingkat kepercayaan yang lebih tinggi dan manfaat klinis yang lebih besar.