

V. PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilakukan, maka dapat disimpulkan bahwa :

1. Algoritma *Convolutional Neural Network* (CNN) dan *Random Forest* berhasil diterapkan untuk klasifikasi penyakit kanker payudara berdasarkan dataset citra mammogram. Kedua algoritma dapat mengidentifikasi citra dengan baik, CNN menunjukkan hasil yang lebih unggul dalam memahami pola visual dari data citra. CNN menunjukkan keunggulan dalam memahami pola visual kompleks berdasarkan kemampuan ekstraksi fitur hierarkis melalui layer konvolusi yang dapat mendeteksi fitur tingkat rendah seperti tepi dan tekstur hingga fitur tingkat tinggi seperti bentuk massa dan kalsifikasi mikro pada mammogram, namun juga memiliki kelemahan yaitu perlu melakukan augmentasi pada dataset untuk menyeimbangkannya. Sementara itu, *Random Forest* memiliki kelebihan dalam kemampuannya mengatasi ketidakseimbangan kelas dan mengenali pola rumit pada citra tanpa memerlukan struktur berlapis seperti CNN, karena dapat menggunakan teknik *ensemble* dengan *multiple decision trees* dan *bootstrap sampling* yang secara alami memberikan bobot seimbang pada kelas minoritas serta dapat menangani *noise* dan *outlier* dengan baik, menjadikannya solusi yang lebih efisien pada dataset yang tidak seimbang.
2. Hasil perbandingan kedua algoritma *Convolutional Neural Network* (CNN) dan *Random Forest*, menunjukkan bahwa CNN lebih unggul dibandingkan *Random Forest*. Model CNN memperoleh nilai *accuracy* sebesar 97%, sedangkan *Random Forest* mencapai 95%. Pada aspek *precision (macro average)*, model CNN mencapai nilai 90% dan model *Random Forest* mencapai nilai 86%. Untuk *recall (macro average)*, model CNN berhasil mencapai nilai 94%, lebih tinggi daripada model *Random Forest* mencapai nilai 91%. Sementara pada metrik *F1-score (macro average)*, model CNN mencapai nilai 92%, sedangkan model *Random Forest* mencapai nilai 88%. Hasil ini menunjukkan bahwa model CNN lebih unggul dalam mengenali dan membedakan citra mammogram kanker payudara *benign* dan *malignant* dibandingkan model *Random Forest* pada dataset yang digunakan.

5.2 Saran

Penulis menyadari masih banyak terdapat kekurangan dan keterbatasan dalam penelitian ini. Adapun saran pada penelitian ini adalah sebagai berikut :

1. Disarankan menggunakan dataset yang seimbang, agar hasil klasifikasi lebih akurat. Ketidakseimbangan kelas dalam dataset dapat mengakibatkan model dan performa tidak optimal pada kelas minoritas.
2. Melakukan variasi lebih luas dengan menguji model pada dataset eksternal untuk memastikan generalisasi model pada data yang berbeda karakteristik.