

BAB V

PENUTUP

5.1 Kesimpulan

Berdasarkan hasil skenario pengujian terhadap model VGG-16 dan VGG-16 yang dikombinasikan dengan *Extreme Learning Machine* (ELM), yang diuji berdasarkan variasi ukuran input citra 150×150, 200×200, 224×224, 256×256, dan 300×300 piksel serta metode *balancing* data dengan *Random Over Sampling* (ROS) dan tanpa *balancing* data maka diperoleh beberapa kesimpulan sebagai berikut:

1. Pengaruh Ukuran Input Citra terhadap Performa Model.
Ukuran input citra berpengaruh terhadap performansi model. Pada pengujian secara umum, ukuran 200×200 memberikan hasil terbaik pada seluruh metrik evaluasi, baik untuk model VGG-16 maupun VGG-16-ELM. Misalnya, pada model VGG-16 tanpa *balancing* data, ukuran 200x200 menghasilkan akurasi 96.59% dan *F1-Score* 0.9769, kemudian dengan *balancing* data akurasi diperoleh sebesar 93.70% dan *recall* 99.53% lebih tinggi dibanding ukuran 150x150, 224x224, 256x256, dan 300x300. Hal serupa juga terlihat pada model VGG-16-ELM.
2. Pengaruh *Balancing* Data terhadap Performa Model.
Penerapan metode *balancing* data *Random Over Sampling* (ROS) memiliki dampak yang bervariasi tergantung pada model yang digunakan. Pada model VGG-16, penggunaan ROS justru menyebabkan penurunan akurasi dan *F1-Score* pada seluruh ukuran input. Sebaliknya, pada model VGG-16-ELM, penggunaan ROS dapat meningkatkan performa akurasi pada seluruh ukuran input.
3. Perbandingan Performa antara Model VGG-16 dan VGG-16-ELM.
Model VGG-16 secara umum menunjukkan performa yang lebih tinggi dibanding VGG-16-ELM pada kondisi tanpa *balancing* data. Namun, model VGG-16-ELM menunjukkan peningkatan performa yang lebih baik setelah diterapkan ROS, terutama dalam hal akurasi dan *F1-Score*.
4. Model dan Performa Terbaik
Berdasarkan seluruh pengujian, performa terbaik diperoleh pada model

VGG-16 dengan input 200×200 tanpa *balancing* data, yang menghasilkan akurasi sebesar 96.59% dan *F1-Score* sebesar 0.9769. Pada penggunaan *balancing* data performa terbaik diperoleh pada model VGG-16-ELM dengan input 200×200 , yang menghasilkan akurasi sebesar 95.91% dan *F1-Score* sebesar 0.9723.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, berikut beberapa saran yang dapat dipertimbangkan untuk pengembangan lebih lanjut:

1. Penggunaan Model Arsitektur yang Beragam.
Penelitian selanjutnya disarankan untuk mengeksplorasi arsitektur CNN lainnya seperti ResNet, DenseNet, atau EfficientNet guna mengetahui perbandingan performansi dalam klasifikasi pneumonia, serta mengidentifikasi model dengan efisiensi dan akurasi terbaik.
2. Penerapan Teknik *Balancing* Data Lain.
Selain metode *Random Over Sampling* (ROS), dapat digunakan teknik *balancing* data lainnya untuk melihat pengaruhnya terhadap performa model, terutama dalam menghadapi ketidakseimbangan data yang lebih kompleks.
3. Perluas Kategori Klasifikasi.
Penelitian ini hanya fokus pada dua kelas (NORMAL dan PNEUMONIA). Penelitian lanjutan dapat mempertimbangkan klasifikasi *multi-kelas*, misalnya membedakan jenis pneumonia berdasarkan penyebabnya (bakteri, virus, atau jamur), untuk memperluas kontribusi dalam sistem diagnosis medis otomatis.
4. Integrasi ke Sistem Klinis atau Aplikasi Nyata.
Model yang telah dikembangkan dapat diintegrasikan ke dalam sistem pendukung keputusan medis atau aplikasi berbasis web/mobile untuk membantu tenaga medis dalam diagnosis awal berbasis citra X-ray dada.
5. Pengujian dengan Dataset Berbeda dan Lebih Besar.

Untuk memastikan generalisasi model, sebaiknya dilakukan pengujian menggunakan dataset lain yang lebih besar atau berasal dari berbagai rumah sakit dengan kondisi pemindaian yang beragam.

Halaman ini sengaja dikosongkan