

BAB V PENUTUP

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa pemilihan arsitektur model dan strategi pelatihan memberikan pengaruh yang signifikan terhadap kinerja klasifikasi penyakit mata berbasis citra retina. Penelitian ini membandingkan dua arsitektur yaitu *SqueezeNet* dan *ShuffleNet*, dengan menerapkan dua strategi pelatihan berupa *fixed feature extraction* dan *partial fine-tuning*. Evaluasi dilakukan menggunakan dataset internal dan dataset eksternal untuk mengevaluasi kemampuan model dalam mempelajari pola data serta melakukan generalisasi pada distribusi data yang berbeda.

Pertama, dari aspek performa klasifikasi, *ShuffleNet* dengan pendekatan *partial fine-tuning* menunjukkan hasil terbaik dengan akurasi sebesar 0.91 pada pengujian dataset internal. Sementara itu, *SqueezeNet* memperoleh akurasi sebesar 0.84 pada skenario *fixed feature extraction* dan meningkat menjadi 0.90 pada *partial fine-tuning*. Pada pengujian dataset eksternal, seluruh model mengalami penurunan performa akibat perbedaan distribusi data, namun pendekatan *partial fine-tuning* tetap memberikan kemampuan generalisasi yang lebih baik dibandingkan *fixed feature extraction*. *SqueezeNet* dengan *partial fine-tuning* memperoleh akurasi sebesar 0.73, sedangkan *ShuffleNet* mencapai 0.69. Hasil analisis *confusion matrix* dan *classification report* menunjukkan bahwa model memiliki performa yang baik pada pengujian dataset internal, terutama pada kelas *cataract* dan *diabetic retinopathy* dengan tingkat klasifikasi yang tinggi. Namun, pada pengujian dataset eksternal, performa model mengalami penurunan, khususnya pada kelas *diabetic retinopathy* dan *normal*, yang masih sering mengalami kesalahan klasifikasi akibat kemiripan karakteristik visual antar kelas.

Kedua, dari aspek efisiensi komputasi, *SqueezeNet* memiliki keunggulan karena menggunakan jumlah parameter yang lebih kecil, yaitu 724.548 parameter dengan ukuran model 2.78 MB. Pada skenario *fixed feature extraction*, model memerlukan waktu pelatihan selama 15.68 menit, sedangkan pada *partial fine-*

tuning memerlukan waktu 17.27 menit. Sementara itu, *ShuffleNet* memiliki total parameter sebanyak 1.257.704 dengan ukuran model 4.97 MB. Waktu pelatihan *ShuffleNet* pada *fixed feature extraction* berlangsung selama 11.25 menit dan meningkat menjadi 36.05 menit pada *partial fine-tuning*. Meskipun membutuhkan sumber daya komputasi yang lebih tinggi, *ShuffleNet* mampu menghasilkan performa klasifikasi yang lebih baik. Hal ini menunjukkan adanya *trade-off* antara efisiensi komputasi dan performa klasifikasi.

Ketiga, dari aspek interpretasi visual menggunakan pendekatan *Explainable Artificial Intelligence (XAI)* melalui *Grad-CAM*, model dengan strategi *partial fine-tuning* menunjukkan kualitas fokus perhatian yang lebih baik dibandingkan *fixed feature extraction*. Berdasarkan hasil visualisasi, *ShuffleNet* menghasilkan area aktivasi yang lebih terpusat pada area retina yang relevan dan lebih konsisten dalam menyoroti bagian yang berkontribusi terhadap proses klasifikasi. Sebaliknya, *SqueezeNet* masih menunjukkan pola aktivasi yang lebih menyebar dan pada beberapa kasus meluas hingga area di luar retina. Meskipun demikian, kedua model masih mengalami kesalahan interpretasi pada citra dengan karakteristik visual yang kompleks, terutama pada kelas *diabetic retinopathy* dan *normal*.

Temuan penelitian ini menunjukkan bahwa strategi *partial fine-tuning* merupakan pendekatan pelatihan yang paling efektif dalam meningkatkan performa model dibandingkan *fixed feature extraction*. Berdasarkan seluruh aspek pengujian, *ShuffleNet* dengan strategi *partial fine-tuning* merupakan model yang paling optimal dalam penelitian ini karena mampu memberikan performa klasifikasi tertinggi, kemampuan generalisasi yang baik, serta interpretasi visual yang lebih relevan melalui pendekatan *Grad-CAM*. Meskipun *SqueezeNet* lebih unggul dari sisi efisiensi komputasi, *ShuffleNet* tetap menunjukkan performa keseluruhan yang lebih baik dalam klasifikasi penyakit mata berbasis citra retina.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa hal yang masih dapat dikembangkan dan ditingkatkan pada penelitian selanjutnya, antara lain sebagai berikut:

1. Evaluasi performa model pada lingkungan implementasi nyata, seperti aplikasi berbasis *web* atau perangkat dengan keterbatasan sumber daya, perlu dilakukan lebih lanjut. Hal ini bertujuan untuk memperoleh gambaran yang lebih akurat terkait kinerja model, khususnya dalam hal waktu respons dan stabilitas saat digunakan secara langsung oleh pengguna.
2. Penelitian ini belum mempertimbangkan aspek efisiensi komputasi secara menyeluruh dalam konteks penggunaan nyata, seperti waktu inferensi pada berbagai spesifikasi perangkat. Oleh karena itu, penelitian selanjutnya dapat menambahkan pengukuran waktu proses secara lebih detail untuk mengevaluasi kesiapan model dalam implementasi praktis.
3. Variasi karakteristik data, terutama pada dataset eksternal, masih memengaruhi performa model. Oleh karena itu, penggunaan dataset dengan distribusi yang lebih beragam serta peningkatan kualitas data dapat menjadi fokus pengembangan untuk meningkatkan kemampuan generalisasi model.
4. Analisis interpretasi model pada penelitian ini masih terbatas pada metode *Grad-CAM*. Penelitian selanjutnya dapat mempertimbangkan penggunaan metode *Explainable Artificial Intelligence (XAI)* lainnya untuk memperoleh pemahaman yang lebih komprehensif terhadap pola yang dipelajari model.