

BAB V

KESIMPULAN

5.1 Kesimpulan

Berdasarkan eksperimen pada empat skenario klasifikasi citra bahasa isyarat, dapat disimpulkan bahwa performa model sangat dipengaruhi oleh representasi dataset dan konfigurasi hyperparameter. Dataset A (*overlay landmark* tunggal) menunjukkan performa tinggi dengan macro F1-Score mendekati 0,98, serta mampu mengenali sebagian besar kelas minoritas. Dataset B (*landmark* pada kanvas hitam) memberikan performa sedikit lebih rendah, dengan macro F1-Score antara 0,96-0,97, karena meskipun fitur struktur tangan cukup efektif, beberapa informasi visual penting dari citra asli tidak sepenuhnya tersedia, sehingga membatasi kemampuan model dalam membedakan gestur yang mirip.

Dataset C (citra asli) memberikan performa terbaik secara keseluruhan dengan macro F1-Score dan akurasi mendekati 0,99-1,00, baik pada Validation maupun Test Set, serta mampu mengenali seluruh kelas, termasuk minoritas, dengan sangat baik. Hal ini menegaskan bahwa citra asli menyediakan representasi fitur gestur tangan yang paling lengkap. Dataset D (heterogen *overlay*) menunjukkan performa baik dengan macro F1-Score sekitar 0,97-0,98 dan menawarkan kemampuan menarik dalam mengenali variasi gestur minoritas. Dataset ini juga bermanfaat pada kondisi citra tidak ideal, misalnya ketika deteksi *landmark* gagal atau pencahayaan rendah. Secara keseluruhan, EfficientNetB0 dengan fine-tuning 40 layer terakhir dan Global Average Pooling terbukti efektif, dan meskipun *landmark* dapat membantu stabilitas model pada kondisi tertentu, citra asli tetap menjadi faktor kunci dalam mencapai akurasi dan generalisasi tertinggi.

5.2 Saran

Penelitian lanjutan disarankan untuk memperluas dan memperkaya dataset, baik dari sisi jumlah subjek, variasi latar belakang, kondisi pencahayaan, sudut pengambilan gambar, ukuran tangan, maupun keragaman gestur, guna meningkatkan ketahanan dan kemampuan generalisasi model pada kondisi dunia nyata. Pengembangan sistem pengenalan bahasa isyarat berbasis data temporal, seperti citra berurutan atau video, perlu dipertimbangkan dengan mengintegrasikan arsitektur recurrent seperti Long Short-Term Memory (LSTM)

atau *Temporal Convolutional Network (TCN)*, sehingga sistem mampu mengenali rangkaian gerakan secara dinamis dan tidak terbatas pada klasifikasi citra statis.

Selain itu, evaluasi model berbasis Dataset C dapat diterapkan pada sistem pengenalan bahasa isyarat real time dengan kamera standar, sedangkan Dataset D dengan heterogen *overlay* bermanfaat untuk kondisi citra tidak sempurna atau *landmark* yang gagal terdeteksi. Integrasi model dengan aplikasi interaktif, seperti pembelajaran bahasa isyarat atau perangkat bantu komunikasi, akan meningkatkan utilitas praktis dan memungkinkan penyesuaian terhadap variasi pengguna di lapangan