BAB V
PENUTUP

5.1 Kesimpulan

Berdasarkan hasil implementasi dan analisis pada Bab IV, penelitian ini berhasil mencapai tujuan utama, yaitu melakukan komparatif kinerja algoritma *Multi-Layer Perceptron* (MLP) dan *Convolutional Neural Network (CNN)* 1D dengan menggunakan ekstraksi fitur *landmark* MediaPipe untuk pengenalan 24 huruf alfabet statistik Sistem Isyarat Bahasa Indonesia (SIBI). Kesimpulan yang diperoleh secara umum dan khusus adalah sebagai berikut:

1. Efektivitas *Pipeline*Ekstraksi Fitur dan *Preprocessing* Data

   Proses ekstraksi *landmark* dengan MediaPipe berhasil mentransformasikan 5.267 citra mentah menjadi 1.490 sampel fitur numerik (63 fitur per sampel yang berasal dari 21 *landmark* × 3 koordinat x, y, z) yang sangat diskriminatif dan konsisten. Meskipun terdapat penurunan yang signifikan dalam jumlah sampel data hasil ekstraksi fitur, tahapan *preprocessing* yang meliputi normalisasi skala berbasis MCP jari tengah, *augmentasi* data *landmark*, penanganan ketidakseimbangan kelas melalui *class weights*, serta *One-Hot Encoding* terbukti sangat efektif. Teknik-teknik ini efektif dalam mengeliminasi variasi skala dan orientasi, meningkatkan variasi data pelatihan, serta mengurangi bias terhadap kelas minoritas kelas P dan Q dengan nilai bobot tertinggi 4.9667 dan 3.1042, sehingga kedua model mencapai kinerja tinggi meskipun distribusi sampel data setelah ekstraksi fitur *landmark* tidak merata

2. Kinerja Klasifikasi pada Data Uji

   Kedua model MLP dan CNN menunjukkan kinerja yang hampir sama dan sangat baik pada data uji (298 sampel *support*) dengan akurasi 98,32%, *macro precision* 0,9878, *macro recall* 0,9873, dan *macro F1-score* 0,9871.

Confusion matrix mengungkapkan pola kesalahan yang konsisten dan minimal dengan total sekitar 5 sampai 6 sampel yang salah klasifikasi, terkonsentrasi di kelas dengan kesamaan geometri yang tinggi seperti E-F (*recall* 0,9474), I-M-T (*recall* 0,9333), serta *false positive* pada kelas S (presisi 0,8667). *Learning Curve* stabil tanpa indikasi adanya *overfitting* maupun *underfitting*. ROC-AUC yang mendekati nilai sempurna AUC dengan nilai 1,00 pada hampir semua kelas, kecuali kelas T dengan nilai 0,98. Dengan adanya temuan ini mengkonfirmasi bahwa tahapan *preprocessing* telah menghasilkan fitur *landmark* yang cukup informatif untuk klasifikasi yang akurat.

3. Perbandingan Komparatif Kinerja MLP dan CNN
   A. Akurasi dan metrik klasifikasi kedua model menunjukkan kesetaraan pada data uji sebesar 98,32% dengan pola kesalahan yang serupa, mengindikasikan bahwa representasi *landmark* setelah *preprocessing* telah cukup *diskriminatif*, sehingga perbedaan arsitektur *fully-connected* dan *convolution* tidak berpengaruh signifikan terhadap akurasi individu.
   B. Generalisasi dan stabilitas pada 10-*fold cross validation*, CNN menunjukkan keunggulan sedikit lebih tinggi dengan rata-rata akurasi 99,75% dan simpangan baku tersebarnya data dari nilai rata-rata sebesar ± 0,0026. MLP yang mencapai akurasi 99,72% dengan simpangan baku tersebarnya data dari nilai rata-rata sebesar ± 0,0035. Standar *deviasi* yang lebih rendah pada model CNN (1D) menunjukkan stabilitas dan kemampuan generalisasi yang sedikit lebih baik, berkat kemampuan konvolusi dalam menangkap pola spasial lokal pada *sekuens landmark*.
   C. Efisiensi komputasi pada model MLP menunjukkan tingkat yang lebih tinggi dengan waktu pelatihan 42,46 detik (kurang dari satu menit). Hal ini terjadi karena kompleksitas MLP yang lebih rendah tanpa adanya lapisan operasi konvolusi yang intensif, sedangkan

CNN memerlukan waktu yang lebih lama dengan waktu pelatihan 98,36 detik (1 menit 38,36) karena adanya lapisan konvolusi dan *pooling* layer yang lebih kompleks pada model CNN dibandingkan MLP. Secara keseluruhan, CNN (1D) direkomendasikan sebagai model pilihan utama karena sedikit lebih unggul dalam stabilitas generalisasi dan ketahanan, sementara MLP tetap menjadi alternatif yang sangat kompetitif untuk aplikasi dengan keterbatasan sumber daya komputasi. Pendekatan berbasis *landmark* MediaPipe dengan *preprocessing* yang intensif telah terbukti sangat efektif untuk tugas pengenalan Bahasa Isyarat Indonesia (SIBI) statis terhadap data dengan jumlah yang terbatas, mengatasi pernyataan masalah dengan menunjukkan perbandingan kinerja dalam hal akurasi, stabilitas, efisiensi waktu, dan kompleksitas model.

5.2 Saran

Berdasarkan hasil penelitian yang sudah dilakukan, terdapat beberapa rekomendasi yang dapat dipertimbangkan untuk penelitian mendatang serta untuk penerapan praktis.

1. Optimalisasi Tingkat Keberhasilan Ekstraksi *Landmark*

   Pengurangan data dari 5.267 citra menjadi 1.490 sampel data menunjukkan kurangnya kemampuan deteksi MediaPipe untuk variasi citra tertentu. Akibat timbulnya keterbatasan proses ekstraksi citra ini tambahkan *preprocessing* citra lanjutan seperti segmentasi tangan (menggunakan model seperti U^2-Net atau SAM) atau penghapusan latar belakang agar meningkatkan akurasi deteksi dan mengurangi kehilangan data.

2. Perluasan ke Gestur Dinamis dan Pengenalan Kalimat

   Mengingat keterbatasan penelitian ini yang fokus pada statistik gestur, pengembangan selanjutnya dapat diarahkan pada pemrosesan gestur dinamis. Pemanfaatan arsitektur *deep learning* temporal seperti LSTM, GRU, atau *Transformer* sangat disarankan untuk memproses data video.

Integrasi ini diharapkan dapat mengakomodasi alfabet dan urutan kata yang dinamis di SIBI, sehingga meningkatkan fungsionalitas sistem dalam konteks interaksi dunia nyata.

3. Implementasi *Real-time* dan Pengujian Lapangan

   Hasil evaluasi penelitian saat ini masih dilakukan secara mandiri, pengembangan aplikasi *mobile real-time* menggunakan *TensorFlow Lite* merupakan langkah krusial yang dapat dicoba dalam penelitian selanjutnya. Pengujian langsung dengan komunitas tunarungu di SLB sangat dianjurkan untuk mengamati model akurasi dalam kondisi lingkungan yang bervariasi seperti pencahayaan dan latar belakang, serta untuk memastikan bahwa latensi dan kenyamanan penggunaan memenuhi kebutuhan komunikasi penyandang tunarungu dalam aktivitas sehari-hari.

4. Pembangunan Dataset Lebih Besar dan Representatif

   Dataset yang digunakan saat ini relatif kecil dan seragam, disarankan untuk penelitian selanjutnya mengumpulkan dataset baru dengan skala yang lebih besar (melebihi 10.000 sampel). Kumpulan data harus mencakup variabilitas yang lebih luas, termasuk keragaman etnis, rentang usia, warna kulit, dan gaya gerak tubuh yang berbeda. Selain itu, akuisisi data dalam kondisi dan latar belakang pencahayaan yang ekstrem sangat penting untuk meningkatkan ketahanan model terhadap keragaman pengguna di Indonesia dalam skenario dunia nyata.

5. Eksplorasi Arsitektur *Hybrid* dan Model Mutakhir

   Untuk eksplorasi lebih lanjut, disarankan untuk melakukan studi komparatif dengan arsitektur terkini seperti Vision *Transformer* (ViT), model *hybrid* MLP-CNN, serta pendekatan *end-to-end* pada citra mentah. Hal ini bertujuan untuk memvalidasi efektivitas ekstraksi fitur otomatis dan

mengidentifikasi peluang untuk meningkatkan akurasi dan efisiensi model dalam sistem pengenalan isyarat SIBI.

6. Penelitian Dampak Sosial dan *Inklusivitas*

Penelitian lebih lanjut diperlukan mengenai dampak penerapan teknologi ini terhadap peningkatan komunikasi dengan sistem yang terbuka dan aksesibilitas pendidikan bagi individu dengan tunarungu dan tunawicara. Selain itu, perlu dilakukan studi komparatif dengan sistem BISINDO (Bahasa Isyarat Indonesia) untuk menghindari bias budaya dan memastikan bahwa teknologi yang dikembangkan dapat mengisi celah keragaman bahasa isyarat yang digunakan oleh komunitas tunarungu di Indonesia secara lebih luas.