

ABSTRAK

Bahasa Isyarat Indonesia (BISINDO) merupakan instrumen komunikasi visual yang sangat krusial bagi komunitas tunarungu. Keterbatasan pemahaman masyarakat umum terhadap isyarat ini memicu kesenjangan interaksi sosial. Penelitian ini bertujuan untuk merancang dan mengimplementasikan sistem pengenalan 26 kelas alfabet BISINDO secara otomatis dan *real-time* berbasis *deep learning*. Integrasi metode yang diusulkan menggabungkan komponen *MediaPipe Hands* untuk lokalisasi objek tangan dan algoritma *Convolutional Neural Network (CNN)* sebagai pengklasifikasi. Masalah penurunan performa akibat gangguan lingkungan diselesaikan melalui pipa prapemrosesan data terpusat, di mana citra mentah diekstraksi menjadi *Region of Interest (ROI)* visualisasi rangka tangan beresolusi 224×224 piksel di atas kanvas hitam murni. Ekspansi dataset dilakukan secara offline melalui teknik *Data Augmentation* yang komprehensif, mencakup *horizontal flipping*, rotasi, *shifting*, dan *zooming* guna memfasilitasi pengenalan dua arah tangan sekaligus mereduksi risiko *Overfitting*. Pengujian arsitektur CNN model *Sequential* dengan tiga tingkat blok konvolusi menghasilkan tingkat akurasi validasi makro sebesar 73,08% dan nilai *validation loss* 0,93. Pada tahap uji *Live Inference* berbasis *webcam*, integrasi penstabil metode *Majority Voting* (10 frame, *confidence threshold* $> 0,75$) terbukti andal dalam menjaga konsistensi translasi huruf di layar. Sistem ini berhasil diwujudkan sebagai solusi penerjemah *end-to-end* yang adaptif terhadap fluktuasi spasial lingkungan nyata.

Kata kunci: BISINDO, Convolutional Neural Network, Deep Learning, *Majority Voting*, *MediaPipe Hands*.

ABSTRACT

Indonesian Sign Language (BISINDO) serves as a vital visual communication instrument for the deaf community. However, the limited understanding of the general public regarding these signs often triggers interactive barriers. This study aims to design and implement an automated, real-time recognition system for 26 BISINDO alphabet classes based on deep learning. The proposed integrative framework combines MediaPipe Hands for hand object localization and a Convolutional Neural Network (CNN) algorithm as the core classifier. Performance degradation issues caused by environmental disturbances are resolved via a centralized data preprocessing pipeline, where raw images are extracted into a hand skeleton topology Region of Interest (ROI) with a standardized resolution of 224×224 pixels overlaid on a pure black canvas. Furthermore, dataset expansion is executed offline using comprehensive Data Augmentation techniques, including horizontal flipping, rotation, shifting, and zooming to facilitate two-handed flexibility while reducing the risk of local Overfitting. Evaluation of the Sequential CNN architecture with three hierarchical convolution blocks yields a macro validation accuracy of 73.08% and a validation loss of 0.93. During the live webcam inference testing phase, the integration of a Majority Voting stabilization mechanism (10 frames, confidence threshold > 0.75) proves effective in maintaining consistent on-screen letter translation. The system successfully operates as an end-to-end translation solution that adapts to spatial fluctuations in real-world environments.

Keywords: BISINDO, Convolutional Neural Network, Deep Learning, Majority Voting, MediaPipe Hands.