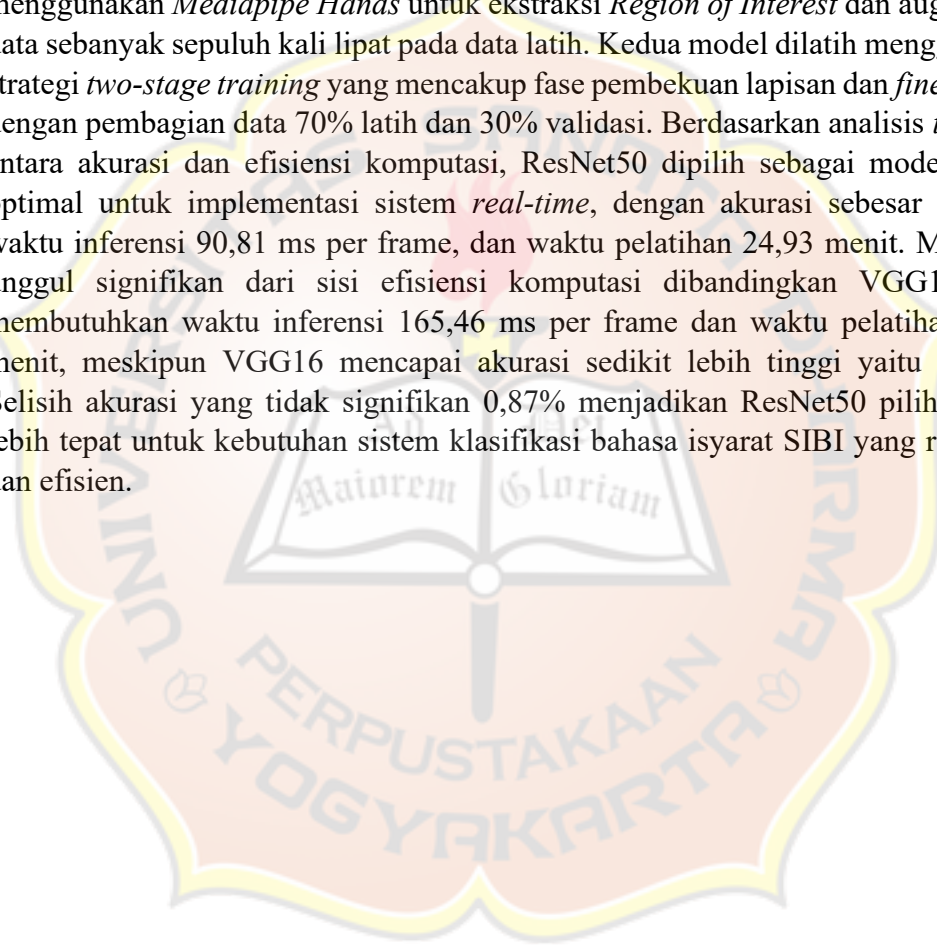


ABSTRAK

Bahasa isyarat merupakan alat komunikasi utama bagi penyandang tunarungu, namun keterbatasan pemahaman masyarakat umum terhadap Sistem Isyarat Bahasa Indonesia (SIBI) menjadi hambatan komunikasi yang signifikan. Penelitian ini bertujuan untuk membandingkan performa arsitektur VGG16 dan ResNet50 berbasis transfer learning dalam mengklasifikasikan 24 kelas alfabet SIBI, menganalisis *trade-off* antara akurasi dan efisiensi komputasi, serta mengembangkan sistem penerjemah bahasa isyarat berbasis website. Dataset terdiri dari 4.608 citra yang bersumber dari tiga dataset Kaggle, kemudian diproses menggunakan *Mediapipe Hands* untuk ekstraksi *Region of Interest* dan augmentasi data sebanyak sepuluh kali lipat pada data latih. Kedua model dilatih menggunakan strategi *two-stage training* yang mencakup fase pembekuan lapisan dan *fine-tuning*, dengan pembagian data 70% latih dan 30% validasi. Berdasarkan analisis *trade-off* antara akurasi dan efisiensi komputasi, ResNet50 dipilih sebagai model paling optimal untuk implementasi sistem *real-time*, dengan akurasi sebesar 97,32%, waktu inferensi 90,81 ms per frame, dan waktu pelatihan 24,93 menit. Model ini unggul signifikan dari sisi efisiensi komputasi dibandingkan VGG16 yang membutuhkan waktu inferensi 165,46 ms per frame dan waktu pelatihan 79,89 menit, meskipun VGG16 mencapai akurasi sedikit lebih tinggi yaitu 98,19%. Selisih akurasi yang tidak signifikan 0,87% menjadikan ResNet50 pilihan yang lebih tepat untuk kebutuhan sistem klasifikasi bahasa isyarat SIBI yang responsif dan efisien.



ABSTRACT

Sign language is the primary communication tool for individuals with hearing impairments, yet limited public understanding of the Indonesian Sign Language System (SIBI) poses a significant communication barrier. This study aims to compare the performance of VGG16 and ResNet50 architectures based on transfer learning in classifying 24 SIBI alphabet classes, analyze the trade-off between accuracy and computational efficiency, and develop a website-based sign language translation system. The dataset consists of 4,608 images sourced from three Kaggle datasets, processed using *Mediapipe* Hands for Region of Interest extraction, with tenfold data augmentation applied to the training set. Both models were trained using a two-stage training strategy comprising a layer-freezing phase and fine-tuning, with a 70:30 train-validation split. Based on the trade-off analysis between accuracy and computational efficiency, ResNet50 was selected as the most optimal model for real-time system implementation, achieving an accuracy of 97.32%, an inference time of 90.81 ms per frame, and a training time of 24.93 minutes. This model significantly outperformed VGG16 in computational efficiency, which required an inference time of 165.46 ms per frame and a training time of 79.89 minutes, despite VGG16 achieving a slightly higher accuracy of 98.19%. The negligible accuracy difference of 0.87% makes ResNet50 the more suitable choice for a responsive and efficient SIBI sign language classification system.

