

Segmentasi Tangan dan Wajah dengan U-Net untuk Pengenalan Isyarat SIBI (Sistem Isyarat Bahasa Indonesia) = Hand and Face Segmentation with U-Net for SIBI (Indonesian Sign System) Sign Recognition

Mahdia Aliyya Nuha Kiswanto, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920531369&lokasi=lokal>

Abstrak

Skripsi ini membahas mengenai penggunaan model segmentasi semantik UNet sebagai alternatif metode segmentasi wajah dan tangan gerakan isyarat SIBI (Sistem Isyarat Bahasa Indonesia) pada latar belakang kompleks. Penelitian dilakukan terhadap dataset gerakan isyarat SIBI milik Lab MLCV Fakultas Ilmu Komputer Universitas Indonesia. Dalam penelitian ini, dilakukan percobaan dengan tiga jenis konfigurasi UNet, yaitu UNet 4-level tanpa Batch Normalization, UNet 5-level tanpa Batch Normalization, dan UNet 4-level dengan Batch Normalization. Hasil segmentasi dari UNet konfigurasi terbaik kemudian dilakukan tahap pengenalan selanjutnya, yaitu ekstraksi fitur dengan MobileNetV2, penghapusan gerakan transisi dengan TCRF, dan gesture recognition dengan 2-layer biLSTM untuk mendapatkan hasil translasi serta evaluasi akhir. Selain itu, performa sistem dengan menggunakan metode segmentasi UNet dibandingkan dengan performa sistem dengan menggunakan metode segmentasi RetinaNet+Skin Color Segmentation. Hasil dari penelitian didapatkan bahwa konfigurasi UNet 4-level dengan Batch Normalization menghasilkan segmentasi yang sedikit lebih baik dibandingkan konfigurasi lainnya, yaitu dengan nilai IOU 0,9178% pada dataset berlatar belakang kompleks. Performa UNet terlihat baik pada saat kedua tangan berada di depan badan, dan menurun ketika tangan berada di posisi yang berdekatan dengan area kulit lainnya (lengan, leher, wajah). Didapatkan juga bahwa sistem pengenalan isyarat SIBI ke teks bahasa Indonesia dengan menggunakan metode segmentasi UNet berhasil memiliki performa yang lebih baik dibandingkan menggunakan metode segmentasi RetinaNet+Skin Color Segmentation, dengan nilai WER 2,703% dan SAcc 82,424% pada latar belakang kompleks. Didapatkan juga waktu komputasi UNet yang lebih cepat dibandingkan RetinaNet dengan waktu segmentasi 0,19643 detik per frame pada CPU NVIDIA DGX A100

This thesis discusses the use of the UNet semantic segmentation model as an alternative to hand and face segmentation methods for SIBI (Indonesian Signing System) on complex backgrounds. This research was conducted on SIBI gesture dataset by MLCV Lab (Faculty of Computer Science, Universitas Indonesia). In this study, experiments were conducted with three types of UNet configurations, namely 4-level UNet without Batch Normalization, 5-level UNet without Batch Normalization, and 4-level UNet with Batch Normalization. Segmentation results from the best UNet configuration is then carried out in the next stage of the system, namely feature extraction with MobileNetV2, epenthesis removal with TCRF, and gesture recognition with 2-layer biLSTM to obtain translation results and the final evaluations. In addition, system performance using the UNet segmentation method is compared to system performance using the RetinaNet+Skin Color Segmentation method. The results of the study showed that the 4-level UNet configuration with Batch Normalization produces slightly better segmentation than the other configurations, with an IOU of 0.9178% on a dataset with a complex background. Based on the sample results, UNet performance is good when both hands are on the front of the body, and it decreases when the hands are in close proximity to other skin areas (arms, neck, face). It was also found that the SIBI gesture recognition system to Indonesian text using the UNet segmentation method managed to have better performance than

using the RetinaNet+Skin Color Segmentation, with a WER value of 2.703% and a SAcc of 82.424% on a complex background. It was also found that UNet processing time was faster than RetinaNet with a segmentation rate of 0.19643 seconds per frame on the NVIDIA DGX A100 CPU.