

Pengenalan Gerakan untuk Bahasa Isyarat Argentina menggunakan CNN dan LSTM

Kartina Halawa¹, Wikky Fawwaz Al Maki²

^{1,2}Fakultas Informatika, Universitas Telkom, Bandung

artinahalawa@student.telkomuniversity.ac.id ,

wikkyfawwaz@telkomuniversity.ac.id

Abstrak

Komunikasi masih sulit bagi orang-orang yang memiliki kesulitan mendengar, terutama karena tidak ada penerjemah terlatih yang tersedia. Percakapan visual dimungkinkan dengan bahasa isyarat, tetapi teknologi yang lebih otomatis diperlukan untuk membuatnya lebih mudah digunakan oleh semua orang. Terima kasih banyak kepada jaringan saraf konvolusional (CNN) dan jaringan memori jangka pendek panjang (LSTM), dua ide baru di bidang pembelajaran mendalam, keduanya telah menunjukkan janji dalam tugas-tugas yang memerlukan pengenalan gerakan. CNN lebih baik dalam mendapatkan informasi spasial dari gambar daripada LSTM dalam menemukan hubungan temporal antara bingkai video. Dataset LSA64 memiliki 3.200 video dari 10 orang yang membuat 64 tanda yang berbeda. Studi ini menggunakan dataset ini untuk menyarankan model CNN-LSTM untuk mengenali Bahasa Isyarat Argentina. Menyelaraskan waktu dan memastikan semua bingkai sama sebagai langkah persiapan. Arsitektur hibrida mencakup CNN untuk mengekstraksi fitur spasial dan LSTM untuk memodelkan urutan temporal. Hal ini memungkinkan untuk memiliki pemahaman yang kuat tentang ruang dan waktu. Kami membandingkan kinerja model CNN dan LSTM secara terpisah dengan kinerja model kombinasi. Model CNN-LSTM hibrida menunjukkan hasil yang menjanjikan dalam pengenalan bahasa isyarat otomatis, dengan performa yang lebih baik daripada kedua model terpisah dalam eksperimen, dengan skor F1 sebesar 92,9% dan tingkat akurasi 92,9%.

Kata kunci—bahasa isyarat, deteksi gerakan, pembelajaran mendalam, CNN, LSTM