

ABSTRAKSI

Mahesa Tirta Panjalu, 51420345

PEMBACAAN GERAK BIBIR MENGGUNAKAN CNN, BI-LSTM DAN CTC LOSS FUNCTION PADA DATASET BAHASA INGGRIS

Skripsi, Jurusan Informatika, Fakultas Teknologi Industri, Universitas Gunadarma, 2024

Kata Kunci: Membaca Gerak Bibir, LipNet, CNN, LSTM, CTC, WER, CER
(xiv + 77 Halaman + Lampiran)

Tunarungu adalah orang yang mengalami gangguan pendengaran. Dampak utama dari kondisi ini adalah hambatan dalam komunikasi verbal atau lisan, sehingga menyulitkan komunikasi dengan orang yang mendengar. Bagian bibir adalah bagian yang biasa digunakan untuk berbicara atau berkomunikasi. Gerakan bibir saat berkomunikasi akan menghasilkan gerakan yang berbeda-beda setiap kata atau huruf yang diucapkan. Bibir dapat digunakan untuk memprediksi kata dari gerak bibir yang akan terdeteksi saat berbicara. Teknologi yang semakin berkembang dapat membantu permasalahan tersebut dalam membaca gerak bibir. *Convolutional Neural Network* atau CNN telah berkembang pesat dan menjadi salah satu metode yang paling populer dalam bidang pengenalan citra dan pemrosesan video karena kemampuannya untuk secara otomatis mempelajari fitur dari data masukan. Penelitian ini bertujuan melakukan pembacaan gerak bibir menggunakan metode CNN, *Long Short-Term Memory* (LSTM) dan *Connectionist Temporal Classification* (CTC) dalam bahasa Inggris. Penelitian ini menggunakan dataset dari The Grid audiovisual sentence corpus sebanyak 1000 video dan 1000 teks. Pada tahapan *preprocessing* terdiri dari dua bagian yaitu *preprocessing* video dan *preprocessing* teks. Tahapan *preprocessing* video meliputi konversi *grayscale*, *cropping frame*, augmentasi dan normalisasi. Tahapan *preprocessing* teks dilakukan proses *encoding* pada dataset *alignments*. Tahapan klasifikasi menggunakan metode *Convolutional Neural Networks*, *Long Short-Term Memory* dan *Connectionist Temporal Classification Loss Function*. Hasil evaluasi mendapatkan nilai akurasi terbaik sebesar 96,9%, *Word Error Rate* (WER) sebesar 0,66%, dan *Character Error Rate* (CER) sebesar 0,16% dengan menggunakan model yang dengan skenario data 80:20 dan *batch size* 2.

Daftar Pustaka (2016-2024)