

IMPLEMENTASI COMPUTER VISION DALAM KLASIFIKASI KALIMAT BISINDO BERBASIS WEBCAM DENGAN MENGGUNAKAN MODEL YOLOV8 DAN MEDIAPIPE

Nama : Suhardianto Rimanda
NIM : 6103230046
Dosen Pembimbing : Afis Julianto, S.Kom., M.Kom

ABSTRAK

Keterbatasan pemahaman masyarakat terhadap Bahasa Isyarat Indonesia (BISINDO) menjadi hambatan komunikasi dengan penyandang tunarungu. Penelitian ini membangun sistem penerjemah BISINDO berbasis *Computer Vision* yang mengenali kata isyarat secara *real-time* melalui *webcam* dan merangkainya menjadi kalimat. Sistem disusun sebagai rangkaian berjenjang, yaitu *YOLOv8* sebagai pendeteksi *Region of Interest* (ROI) tangan, *MediaPipe Holistic* sebagai pengekstraksi *landmark*, LSTM dan BiLSTM sebagai pengklasifikasi kata, serta modul *Natural Language Processing* (NLP) berbasis aturan sebagai penyusun kalimat, kemudian diimplementasikan sebagai *prototype* berbasis *website* untuk komunikasi satu arah dengan keluaran teks dan suara melalui gTTS. *Dataset* terdiri atas 5.000 video pada 40 kelas kata, yaitu 2.000 video publik *Kaggle* dan 3.000 video rekaman mandiri, dengan rasio 70:15:15. Pengujian dilakukan melalui empat skema yang dibedakan menurut komposisi penutur dan arsitektur model. Pada skema satu sumber penutur, LSTM dan BiLSTM sama-sama memperoleh *accuracy* 99,64% dengan *macro F1-score* 0,996, sedangkan pada skema dengan komposisi penutur lebih beragam diperoleh 85,92% (LSTM) dan 88,03% (BiLSTM). Model *YOLOv8n* memperoleh mAP@50 sebesar 0,9947 dan mAP@50–95 sebesar 0,8593. Pengujian penyusunan kalimat memperoleh tingkat keberhasilan 40%, dengan kecepatan pemrosesan sistem di 15 FPS. Hasil ini menunjukkan bahwa kinerja pada skema satu sumber penutur belum mencerminkan kemampuan generalisasi model, sehingga *accuracy* 88,03% dijadikan acuan kinerja model pada penelitian ini.

Kata Kunci: BISINDO, *YOLOv8*, LSTM, NLP, *Computer Vision*.

IMPLEMENTATION OF COMPUTER VISION FOR WEBCAM-BASED BISINDO SENTENCE CLASSIFICATION USING YOLOV8 AND MEDIAPIPE

Name : Suhardianto Rimanda
NIM : 6103230046
Supervisor : Afis Julianto, S.Kom., M.Kom

ABSTRACT

The limited public understanding of Indonesian Sign Language (BISINDO) remains a communication barrier for deaf individuals. This research develops a *Computer Vision*-based BISINDO translation system that recognises sign words in real time through a webcam and arranges them into sentences. The system is built as a staged pipeline comprising YOLOv8 as a hand Region of Interest (ROI) detector, MediaPipe Holistic as a landmark extractor, LSTM and BiLSTM as word classifiers, and a rule-based Natural Language Processing (NLP) module as a sentence composer, implemented as a website-based prototype for one-way communication with text and speech output through gTTS. The dataset comprises 5,000 videos across 40 word classes, consisting of 2,000 public Kaggle videos and 3,000 self-recorded videos, divided at a ratio of 70:15:15. Testing was conducted through four schemes distinguished by signer composition and model architecture. In the single-signer scheme, both LSTM and BiLSTM achieved an accuracy of 99.64% with a macro F1-score of 0.996, while the scheme with a more diverse signer composition yielded 85.92% (LSTM) and 88.03% (BiLSTM). The YOLOv8n model achieved an mAP@50 of 0.9947 and an mAP@50–95 of 0.8593. Sentence composition testing achieved a success rate of 40%, with a system processing speed of 14 FPS. These results indicate that performance under the single-signer scheme does not yet reflect the model's generalisation capability; therefore, the accuracy of 88.03% is taken as the reference performance of the model in this research.

Keywords: BISINDO, YOLOv8, LSTM, NLP, *Computer Vision*.