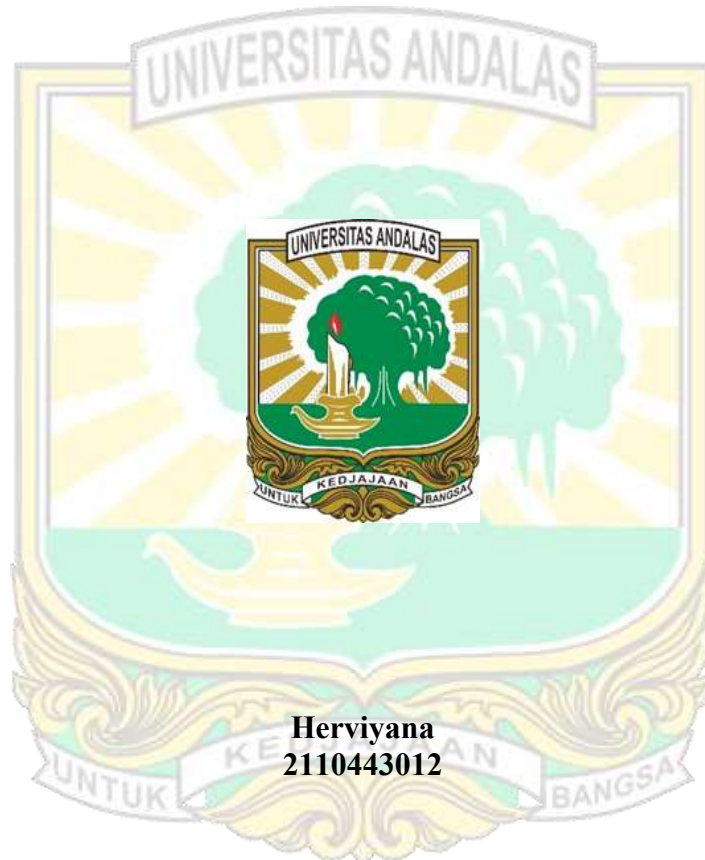


**DEVELOPMENT OF A SPEECH-TO-TEXT COMMUNICATION
SYSTEM FOR THE DEAF AND DUMB USING COMPUTER VISION**

UNDERGRADUATE THESIS

**A paper is one of the requirements
To obtain a Bachelor of Science Degree
From the Universitas Andalas**



**Herviyana
2110443012**

**DEPARTMENT OF PHYSICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCE
UNIVERSITAS ANDALAS
PADANG**

2025

PENGEMBANGAN SISTEM KOMUNIKASI *SPEECH-TO-TEXT* UNTUK ORANG TUNARUNGU DAN TUNAWICARA MENGGUNAKAN TEKNOLOGI *COMPUTER VISION*

ABSTRAK

Komunikasi verbal masih menjadi tantangan utama bagi penyandang disabilitas rungu wicara, sehingga dibutuhkan sistem yang menjembatani interaksi dua arah secara efektif. Penelitian ini mengembangkan sistem komunikasi yang mengintegrasikan *Computer Vision* (CV) untuk klasifikasi bahasa isyarat dan *speech-to-text* (STT) untuk pengenalan ucapan, sehingga komunikasi dapat tercapai secara lebih efektif antara penyandang disabilitas dan masyarakat umum. Sistem ini bekerja dengan memanfaatkan *Computer Vision*, kamera sebagai alat sensor untuk mendeteksi gestur tangan, kemudian diproses menggunakan MediaPipe untuk ekstraksi *keypoints* dan diklasifikasikan dengan model *Convolutional Neural Network* (CNN). Model latih dari 71 kelas gambar bahasa isyarat, terdiri dari 24 kelas gambar SIBI dan 47 kelas gambar BISINDO, dengan jumlah total gambar data mentah sebanyak 2.160 citra sebelum *augmentasi* dan 6.390 citra setelah *augmentasi*. Hasil pengujian menunjukkan bahwa CNN mampu mengenali gestur dengan akurasi tinggi, yakni 95,91% untuk SIBI dan 92,64% untuk BISINDO. Sistem dilengkapi *speech-to-text* agar dapat mengubah ucapan menjadi teks. Integrasi antara CV dan STT ini, akan ditampilkan melalui website yang responsif baik pada PC atau *Smartphone*. Sementara itu, hasil deteksi gestur akan ditampilkan secara *real-time* melalui layar LCD 16x2 yang terhubung ke NodeMCU ESP8266 dengan menggunakan protokol MQTT. Penelitian ini menunjukkan bahwa integrasi CV dan STT mampu memberikan solusi dalam mendukung komunikasi verbal bagi penyandang disabilitas rungu wicara. Implementasi ke depan diarahkan pada pengembangan perangkat portable berbasis IoT yang lebih efisien.

Kata Kunci: Bahasa Isyarat, Computer Vision, MediaPipe, Model CNN, Speech-to-Text,

DEVELOPMENT OF A SPEECH-TO-TEXT COMMUNICATION SYSTEM FOR THE DEAF AND DUMB USING COMPUTER VISION

ABSTRACT

Verbal communication remains a significant challenge for people with hearing impairments, so a system is needed to effectively bridge two-way interaction. This research develops a communication system that integrates Computer Vision (CV) for sign language classification and speech-to-text (STT) for speech recognition, enabling more effective communication between people with disabilities and the general public. The system operates by utilizing Computer Vision, with a camera as a sensor to detect hand gestures, which are then processed using MediaPipe for keypoint extraction and classified using a Convolutional Neural Network (CNN) model. The model was trained using 71 classes of sign language images, consisting of 24 courses of SIBI images and 47 classes of BISINDO images, with a total of 2,160 raw images before augmentation and 6,390 images after augmentation. Test results show that the CNN can recognize gestures with high accuracy, namely 95.91% for SIBI and 92.64% for BISINDO. The system is equipped with speech-to-text (STT) to convert speech into text. The integration between Computer Vision (CV) and STT will be displayed through a responsive website on both PCs and smartphones. Meanwhile, gesture detection results will be displayed in real-time on a 16x2 LCD screen connected to a NodeMCU ESP8266 using the MQTT protocol. This study shows that the integration of CV and STT can provide solutions to support verbal communication for people with speech and hearing disabilities. Future implementations will focus on developing more efficient IoT-based portable devices.

Keywords: Sign Language, Computer Vision, MediaPipe, CNN Model, speech-to-text