

GestureSpeak: An Accessibility-Focused Android Application for Real-Time Hand Gesture Recognition and Speech Conversion

Swarangi Vivekanand Gaikwad¹, Kalpesh Ramkrushna Patil², Sakshi Prashant Ahirrao³

Samruddhi Mahesh Beliskar⁴, Prof. C. A. Shewale⁵

Students, Department of Information Technology¹⁻⁴

Faculty, Department of Information Technology⁵

Sandip Polytechnic, Nashik, Maharashtra, India

ORCHID ID :0009-0009-0365-8434, ORCHID ID :0009-0005-9069-0715

ORCHID ID :0009-0004-5897-7786, ORCHID ID :0009-0007-8410-8228

swarangigaikwad87@gmail.com, patilkalpesh7822@gmail.com, ahirraosakshi475@gmail.com

samruddhibeliskar08@gmail.com, chetan.shewale@sandippolytechnic.org

Abstract: *Hand gesture recognition plays a vital role in assistive technologies, especially for individuals with speech or hearing impairments. GestureSpeak is an Android-based accessibility application developed using Google's MediaPipe Gesture Recognizer that enables real-time detection, classification, and interpretation of hand gestures. The application supports live camera streams, image-based gesture recognition, and video-based analysis. Recognized gestures are converted into meaningful text and synthesized speech using Android Text-to-Speech (TTS), while Speech-to-Text (STT) enables voice-controlled operation. The system integrates CameraX for efficient frame processing, GPU-accelerated inference with CPU fallback, confidence-based filtering, and intelligent sentence formation. GestureSpeak aims to bridge the communication gap by providing a hands-free, real-time gesture-to-speech solution optimized for mobile devices.*

Keywords: Gesture Recognition, MediaPipe, Accessibility, Android Application, Computer Vision, Text-to-Speech, Speech-to-Text, CameraX

