

CHAPTER 1 INTRODUCTION

1.1 Research Background

People with disabilities can be defined as individuals with physical, intellectual, mental, and/or sensory limitations for an extended period, interacting with the environment (Ashar et al., 2019). Sensory disability is the disruption of one of the functions of the five senses, such as visual, or people with deaf or dumb disabilities. People with disabilities often communicate through speaking, but to learn to write, they use Braille, a system of letters that combines dots. Deaf and dumb disabilities refer to limitations in hearing and speaking experienced by individuals (Allo and Alsih, 2022).

Based on data from the World Health Organization (WHO), in March 2020, it was reported that around 5% of the world's population experienced deafness, comprising approximately 432 million adults and 34 million children (Mua'wwanah et al., 2015). According to data from the World Health Organization (WHO), in 2021, around 1.5 billion people worldwide were affected by deafness, with the highest prevalence found among adolescents. This condition is strongly associated with lifestyle habits, particularly the tendency of adolescents to listen to music at excessively high volumes. The Ministry of Social Affairs noted that 7.03% of the 30.38 million people with disabilities are classified as deaf (Yustianingsih et al., 2024). People with sensory disabilities need a place to communicate with interlocutors. There are limitations to date, causing communication to be hindered.

Verbal communication is the primary medium for most people in their daily interactions. For people with deaf and dumb disabilities, reliance on alternative communication methods such as sign language, lip reading, or hearing aids is often the solution. Unfortunately, not everyone in the social environment understands sign language and can adapt to their communication needs. This need becomes a barrier and creates a gap in social interaction.

In the literature, several studies have been conducted to develop speech-to-text-based communication systems for deaf individuals. One of the studies by

Kurniawan (2017) developed a speech-to-text communication system that uses Google Cloud speech-to-text software. This system can be used to convert voice into text in real-time. This research utilizes the Speech Recognition Application on Google Cloud Speech, examining three recognition rate factors: tone, pronunciation, and speaking speed. In future development, a more attractive and neat application user interface will be created, and iOS should be used.

Dhake et al. (2020) have conducted research on sign language communication systems with deaf and dumb people. This system provides two-way communication with deaf and dumb people by using a Raspberry Pi to process hand gesture images and also generate text words and speech. The research built an easy method of communication with deaf and mute people by utilizing image processing algorithms. The language used is Indian sign language, which contains letters in the application of the system model. The constraints of using the language and realizing the data in one sentence do not exist.

Alsulaiman et al. (2023) develop and facilitate communication with deaf people. Building a large Saudi Sign Language dataset. The research was conducted to introduce a Convolutional Neural Network (CNN) for sign language recognition and implement the proposed architecture. The research was used to facilitate the integration of people with disabilities into their society. Future research is expected to create opportunities for people with disabilities using two-way communication.

Another research project conducted by Arrozaq (2024) developed a computer vision implementation to translate the SIBI sign language alphabet on an Android application. This system can be used to detect motion in room conditions with varying light using computer vision. Further research is expected to improve accuracy in dim places in the variation data.

The above studies have shortcomings that become obstacles to development. This obstacle is the motivation for the research to be carried out. This research is used to solve the constraints of communication and learning sign language. The obstacle is when dumb people who communicate with normal people through writing take a long time to write and understand. There is also little speaking space for dumb people due to discrimination from the interlocutor, which

causes a big obstacle. These obstacles will be solved with the advancement of digital technology. Digital technology is used as a means of communication that helps people with disabilities communicate. One of the communication technologies that can be used is the speech-to-text (STT) based communication room system. STT itself is the process of processing a sound signal, extracting features from it, which are then compared with the extraction results from other sound signals to recognize similarities (Dinata et al, 2017). STT is a technology that can convert voice into text.

This technology is used to translate the voice of a person with deaf and dumb disabilities into text that a person can read without a dumb, and vice versa. This text is shared with others through the medium of text on a personal computer (PC). This STT will be made into a system that will be operated using Computer Vision. Computer Vision is a field of science with the potential to have a profound impact on human life. It does not merely replicate human vision but provides the ability to digitally capture, process, and interpret visual information for specific applications (Putra, 2010).

In this research, Computer Vision plays a crucial role in bridging the communication gap by recognizing, processing, and translating hand gestures into digital information that can be combined with speech-to-text results. Unlike STT, which focuses on converting audio signals into text. Computer Vision ensures that visual gestures, particularly sign language from the deaf and dumb individuals, are accurately captured and interpreted.

Based on these issues, a Computer Vision system is used to process data combined with speech-to-text (STT), enabling two-way communication. This system allows messages from people without disabilities to be translated from speech to text, while messages from deaf and mute individuals can be translated from sign language to written text or speech. This tool is designed to serve as a medium that enhances the confidence of deaf and mute individuals to communicate effectively in social environments without discrimination.

1.2 Research Purpose

The objectives of this research are as follows:

1. To develop two-way communication for people with deaf and dumb disabilities with people who do not have sensory disabilities.
2. To develop a speech-to-text-based communication system equipped with Computer Vision (CV) technology.
3. To improve speech recognition accuracy by combining STT and hand gesture image analysis.

1.3 Research Benefits

The benefits obtained are as follows:

1. To provide insight and learning opportunities for people who do not understand sign language.
2. To serve as a communication medium that facilitates smooth interaction between people with deaf and dumb disabilities and people without sensory disabilities in everyday life.
3. To increase the social inclusion of people with deaf and dumb disabilities in daily activities.

1.4 Research Scope and Limitations

The scope of this research focuses on developing an STT-based communication system equipped with Computer Vision to analyze hand gestures. This study uses Indonesian Sign System (SIBI) and Indonesian Sign Language (BISINDO), as the target sign language and Indonesian speech recognition.

The limitations of this study are as follows:

1. The image data used is 350 Indonesian Sign Language System (SIBI) or Indonesian Sign Language (BISINDO), sign language images with model classes, namely 26 letters (A-Z) and 9 numbers (1-9).
2. The CNN algorithm is used to train the image model to be detected.
3. The camera used is a webcam to capture images.
4. Tests were conducted on a small group of people with speech impairment