THE MODEL INSTITUTE OF ENGINEERING AND TECHNOLOGY (Autonomous)
Kot Bhalwal, Jammu

SYNOPSIS

SIGN LANGUAGE CONVERSION

Submitted in partial fulfilment of the requirement for the award of the Degree of

Masters of Computer Applications

SESSION: 2022-2024

SUBMITTED BY

Divyansh Gupta (2022d1r015)

Aniket Sharma (2022d1r013)

M.CA - IVTH SEM

UNDER THE GUIDANCE OF

Mr. Vishal Gupta
(Head M.C.A Department)

TABLE OF CONTENT

SYNOPSIS

TITLE OF THE PROJECT:

SIGN LANGUAGE CONVERSION (SLC)

HARDWARE REQUIREMENTS

- **Processor:** AMD or Intel Core i3 or higher
- **Ram:** 4GB or higher
- **Camera or webcam:** Capable of capturing live video feed
- **Speakers or headphones:** For audio output

SOFTWARE REQUIREMENTS

- **Operating System:** Windows OS 8 or above
- **IDE:** Anaconda IDE or Visual Studio Code
- **Machine Learning Library:** OpenCV, TensorFlow, Kera's, NumPy
- **Language:** Python

INTRODUCTION

Communication is a fundamental aspect of human interaction, yet for individuals who are **deaf-mute**, traditional means of communication can present significant challenges. **Sign language** serves as a vital mode of expression for this community, but its interpretation by those unfamiliar with it remains a hurdle. In response to this need, the development of a machine learning model using Python to convert **sign language videos and photos** into **text and audio formats** offers a promising solution.

This project harnesses the power of technology to bridge the communication gap for the **deaf-mute** community. By leveraging **deep learning models**, specifically **convolutional neural networks (CNNs)**, the intricate movements and gestures of **sign language** are decoded and translated into understandable text. **Python**, a versatile and widely-used programming language, serves as the foundation for this model, providing flexibility and ease of implementation.

Key libraries such as **TensorFlow** and **Kera's** are integral to the development process, facilitating the creation and training of the neural networks. Through a combination of image processing techniques and advanced algorithms, **sign language videos and photos** are analysed, enabling the model to accurately interpret gestures and generate corresponding textual representations.

Moreover, the incorporation of **text-to-speech libraries** such as **Google Api** enhances accessibility by converting the generated text into audio output, further empowering individuals with **hearing impairments** to communicate effectively.

In summary, this project aims to revolutionize communication accessibility for the **deaf-mute** community through the innovative use of machine learning and Python programming. By transforming **sign language into text and audio**, this model strives to promote inclusivity and facilitate meaningful interactions for all individuals, regardless of their abilities or communication preferences.

ABSRACT

Sign language is one of the oldest and most natural form of language for communication, hence we have come up with a real time method using neural networks for finger spelling based Indian sign language. This project introduces a groundbreaking solution aimed at enhancing communication accessibility for individuals who are **deaf-mute** through the development of a **machine learning model** using **Python**. The model converts **sign language videos and photos** into **text and audio formats**, facilitating seamless interaction for the **deaf-mute** community.

Utilizing **deep learning models** including **convolutional neural networks (CNNs)**, intricate **sign language** gestures are accurately decoded and translated into understandable text. The project relies on key libraries such as **TensorFlow** and **Kera's** for the creation and training of neural networks. Through advanced algorithms and image processing techniques, **sign language videos and photos** are analysed to interpret gestures effectively.

Additionally, **text-to-speech libraries** such as **Google Api** are integrated to convert the generated text into audio, further enhancing accessibility for individuals with **hearing impairments**. This **machine learning**-based approach revolutionizes communication accessibility, promoting inclusivity and enabling meaningful interactions for individuals within the **deaf-mute** community.

Key Words: - Machine Learning, Image Processing, Convolutional Neural Networks model, Google Api, OpenCV, Python.

ABOUT SLC

"Sign Language Conversion" is an innovative application of Human-Computer Interaction and Machine Learning technologies. It aims to bridge the communication gap between the hearing-impaired community and those who may not understand sign language. This system is designed to recognize and interpret sign language gestures, converting them into written text. But it doesn't stop there. To make communication even more seamless, the system also includes a feature to convert this text into audio output. This allows individuals who are visually impaired or prefer auditory communication to understand the message as well.

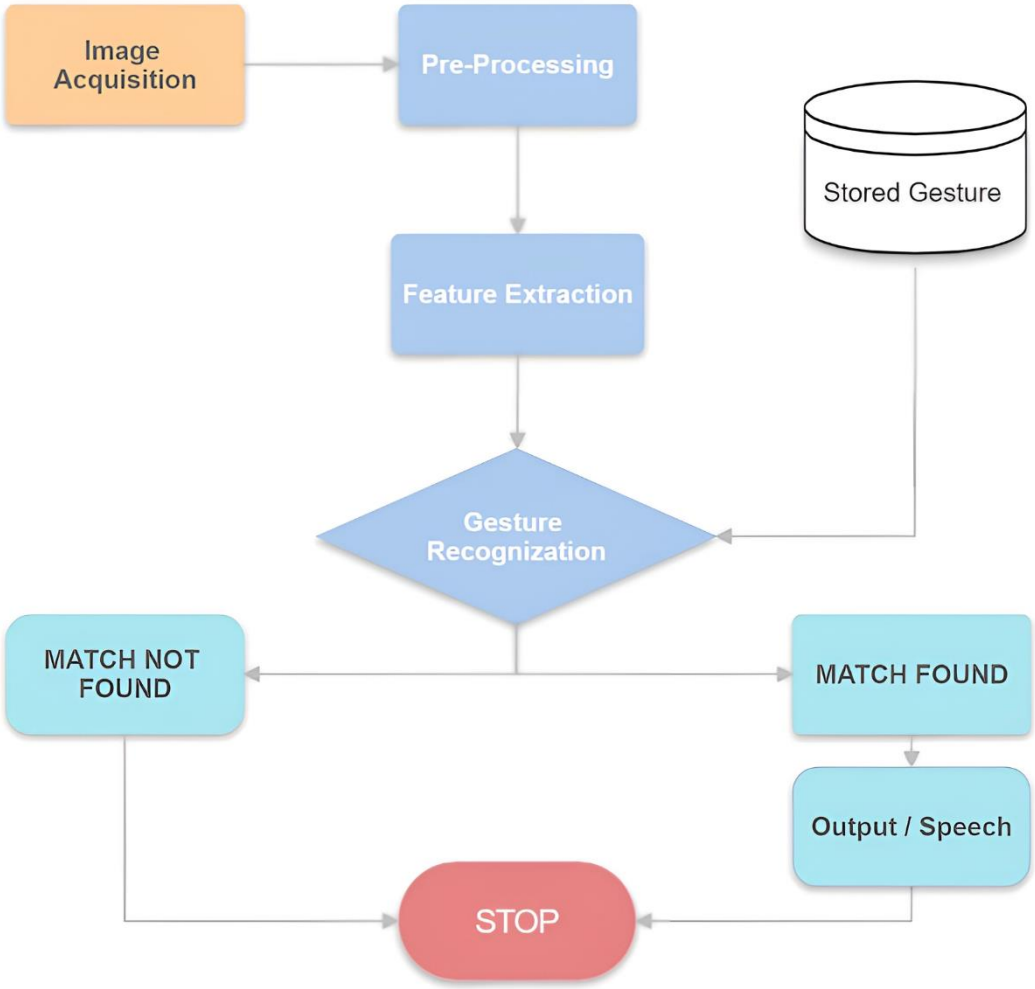
USES OF SLC

1. **Accessibility in Public Spaces:** Public places such as libraries, banks, hospitals, and government offices can integrate the system to provide accessible communication options for individuals with hearing impairments, ensuring they can access services independently.
2. **Education and Learning:** Educational institutions can utilize the system to facilitate communication between students who are deaf or hard of hearing and their peers, teachers, and support staff.
3. **Emergency Response and Public Safety:** Emergency response services can benefit from the system by enabling communication with individuals who are deaf or hard of hearing during emergencies or crisis situations, ensuring they receive timely information and assistance.
4. **Remote Communication:** The system can be utilized for remote communication, enabling individuals who are deaf or hard of hearing to participate in virtual meetings, conferences, and online interactions with real-time sign language interpretation and audio output.

WORKING OF SLC

1. **Image Acquisition:** The system captures live video feed or images containing sign language gestures using a camera or webcam.
2. **Gesture Detection:** The captured images or video frames are processed using computer vision techniques, such as object detection and segmentation. This step focuses on detecting and isolating sign language gestures from the background.
3. **Gesture Recognition:** Utilizing machine learning algorithms, particularly convolutional neural networks (CNNs) or other models trained on sign language datasets, the system recognizes and interprets the detected gestures. This involves mapping each gesture to its corresponding meaning or textual representation.
4. **Text Generation:** Once the gestures are recognized, the system converts them into written text using natural language processing (NLP) algorithms. This step involves encoding the recognized gestures into textual transcriptions, representing the spoken words associated with the sign language gestures.
5. **Audio Synthesis:** In addition to textual output, the system synthesizes audio representations of the recognized words using speech synthesis techniques. This allows for auditory communication of the interpreted sign language gestures, enhancing accessibility for individuals who are visually impaired or prefer auditory communication.
6. **Output Display:** The system may display the textual transcriptions alongside the captured video feed, providing real-time feedback of the interpreted sign language gestures. Additionally, the synthesized audio may be played back through speakers or headphones, enabling users to hear the communicated message.

STEPS INVOLVED FOR SLC SYSTEM



METHODOLOGY

OpenCV: - OpenCV is an Open-source computer vision library used as an interface between the user and a machine and it is used for image and video analysis, like facial detection, license photo editing, advanced robotic vision, and many more. The aims is to recognize Hand Gestures and interpret sign language gestures in real-time. Implement algorithms to identify and isolate hand movements and gestures from the background of captured images or video frames.

Image Processing: - Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. It is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that image. Features are extracted from pre-processed images using techniques such as edge detection and image segmentation.

CNN (Convolutional Neural Network): - It is most preferably used algorithm for operations upon the images. In simple word what CNN does is, it extracts the feature of image and convert it into lower dimension without losing its characteristics. So that working upon it would not take more computational power.

Classification is performed in two steps: training and testing.

- Training step involves calculating the parameters of a probability distribution using the training samples matrix.
- Testing step determines the posterior probability of untested samples belonging to each class and classifies them accordingly.

Text-to-Speech Conversion: - Google API Conversion of Text to Speech is used to convert textual descriptions generated based on class labels into synthesized speech. The API generates equivalent grammatical text descriptions to ensure natural-sounding speech output. Synthesized speech allows for auditory communication of the interpreted sign language gestures, enhancing accessibility for users.

FLOWCHART OF CNN

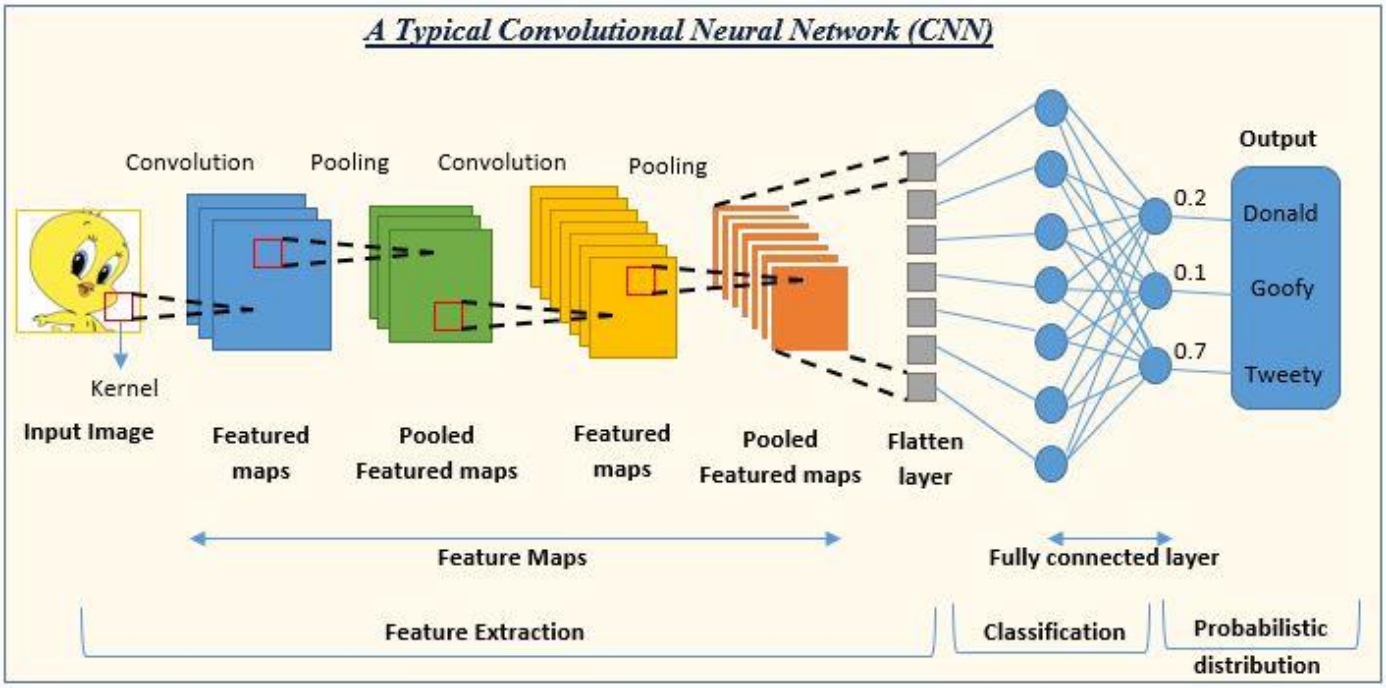
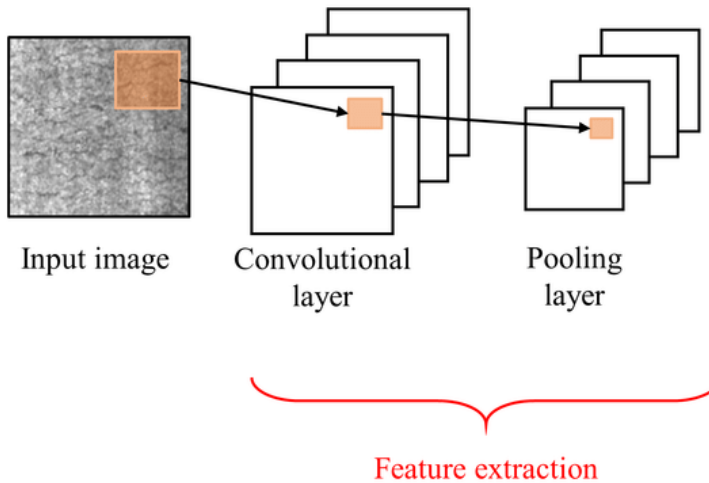


Figure 1

Feature Extraction (Figure 2)



Classification (Figure 3)