Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 45 No. 2 (2024)

Action Recognition System for Sign Language Using TensorFlow & OpenCV

Abhinav Yadav[1], Yash Saini[2], Yatharth Goswami[3], Abhash Chauhan[4], Mr. Md. Shahid[5]
[1,2,3,4,5] *Dept. of CSE, MIET, Meerut*

Abstract- To confirm the communication between two people, both parties However, for the deaf and dumb, it is a means of communication They use poetic language to communicate with each other and with ordinary people, however Most people ignore the importance of poetic language. Not everyone knows and knows the language of poetry Communication between the hearing and the deaf is difficult. It is possible to train a people to interpret different sign language symbols into English. This facilitates communication for both deaf and non-deaf. A unified system for recognising Indian signal languages through machine learning algorithms is an approach that uses a webcam to generate an Indian Sign Language dataset, which is then used to train TensorFlow models through cross-learning to generate real-time programming languages.

Keywords – Indian Sign Language, Feature Extraction, Key point Matching, Sign/Gesture/Action Recognition.

1. Introduction

Communication requires a speaker, message, and listener. It can be divided into several categories, including written and oral, formal and informal, non-verbal, grapevine, feedback, visual, and active listening. While grapevine communication is impromptu, formal communication adheres to predetermined channels. Oral communication involves spoken words, while written communication relies on letters or emails. Non-verbal communication involves body language, and feedback happens when commenting on a product or service. Visual communication relies on visual sources and active listening involves understanding what is being conveyed for effective communication.

Technologies like data gloves, motion capture systems, or sensors are developed to make this easier. Systems for Sign Language Recognition (SLR) based on vision have also been developed. Current research has led to the creation of SLR systems using machine learning that can translate sign language with promising accuracy, but real-time applications remain a challenge. The goals moving forward are to develop a real-time SLR system using TensorFlow and train it with data from a webcam.

This adaptable system is crucial for the deaf and mute community, breaking down communication barriers with its customizability. Making interaction comprehensible for disabled participants in global meetings through sign language detection is a standout feature. The system's flexibility goes beyond this, even in education. It can be introduced to elementary school kids, instilling inclusivity early on. The potential does not stop there though. The system can be harnessed to create apps that recognize specific actions.

2. Literature Review

Sign languages, as visual languages, are used by the hearing impaired to communicate using hand, face, and body movements. They comprise over 300 worldwide variations, but their prevalence is low, limiting effective communication. Here, Sign Language Recognition (SLR) steps in, translating gestures into conventional languages like English/Hindi. SLR, an intense field of research, harnesses the power of Machine Learning to help systems learn and make decisions based on data [1][2]. To properly train and test these systems, two distinct data sets are employed. Direct measurement techniques, which use motion data gloves, sensors, or capturing systems, are examples of data acquisition methodologies. Vision-based techniques, on the other hand, concentrate mostly on the spatial and temporal extraction from RGB images. Vision-based systems in their early stages use skin

colour detection to track and identify hand representations [3]. However, advanced systems leverage face detection and subtraction, alongside background subtraction, to avoid misrecognition of non-hand body parts.

Various researchers have explored different methodologies in Hand Gesture Recognition and Sign Language to help individuals with hearing impairments. Examples include Pallavi Gurjal and Kiran Kunnur who focused on video capture and feature extraction techniques [4]. Archana S Ghotkar used Cam shift and HSV models, but faced obstacles with MATLAB compatibility [5]. P Subha Rajan and Dr G Balakrishnan proposed a method utilizing a 7-bit orientation and generation process for Indian Sign Language gestures [6]. Byung-woo Min et al. introduced a device-free method using image moments and Hidden Markov Models [7]. T. Sahana, Soumi Paul, Subhadip Basu, A. F. Mollah's proposal used multiscale density features for American Sign Language numerals [8]. Conditional Random Fields for Kinect Sensor-Based Sign Language Recognition by Seosuk-dong and Dong-ku from Chosun University, Korea [9]. Image Processing for the Recognition of Sign Language by Sawant Pramada, Deshpande Saylee, Nale Pranita, Nerkar Samiksha, Mrs.Archana, S. Vaidya from R. H. Sapat College of Engineering, Nashik [10]. These studies collectively contribute to the advancement of sign language recognition, aiming to improve communication and inclusivity.

Flow Chart

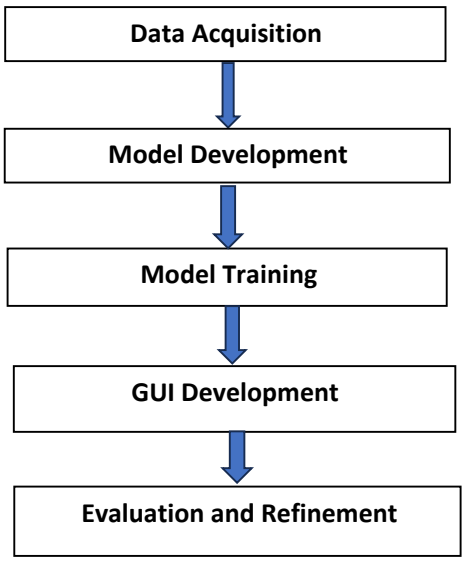

Fig. 1 Flow Chart of Sign Language Word Recognition

Data Gathering:

To take webcam photos, OpenCV and Python are used. OpenCV, created especially for real-time computer vision, offers a standard platform for computer vision applications in addition to expediting machine perception in commercial products. OpenCV includes a wide range of features, including face detection and recognition, object identification, human actions classification, tracking camera and object movements, and the extraction of 3D object models, among other tasks. It has over 2500 efficient computer vision and machine learning algorithms.

3. Methodology:

After gathering data, a labelled map is produced which identifies each object in the model by alphabet sign from A-Z. Each of these 26 labels (representing an alphabet) gets a unique id from 1 to 26. The next step is to create TF records of both training and test data by employing the generate_tfrecord tool used for training the TensorFlow object detection API. TensorFlow stores its data in a binary format (TF record) which enhances the import pipeline, thus saving time in model training. Binary files read from the disk more quickly, take up less disk space, and duplicate more quickly.

The creation, training, and application of object detection models is made easier by the open-source TensorFlow object detection API. It provides a platform with a range of pre-trained detection models available, called the TensorFlow detection model zoo. The SSD MobileNet v2 320x320 model is used in this situation; it is merged with the shared box predictor, FPN-lite feature extractor, and a focus loss technique for 320x320 images. The previous model is modified in order to train it using the new dataset. Text_format, pipeline_pb2, config_util, TensorFlow, and other libraries are imported for configuration. The main modification is that there are now 26 classes instead of the original 90, which corresponds to the amount of alphabet signs the model will be trained to identify.

4. Result Analysis

Real-time alphabet detection in Indian Sign/Action Language is possible with the created method. The TensorFlow object detection API was used in the development of the system. Transfer learning has been used to train it on the generated dataset, which comprises 650 photos overall—25 images for each letter.

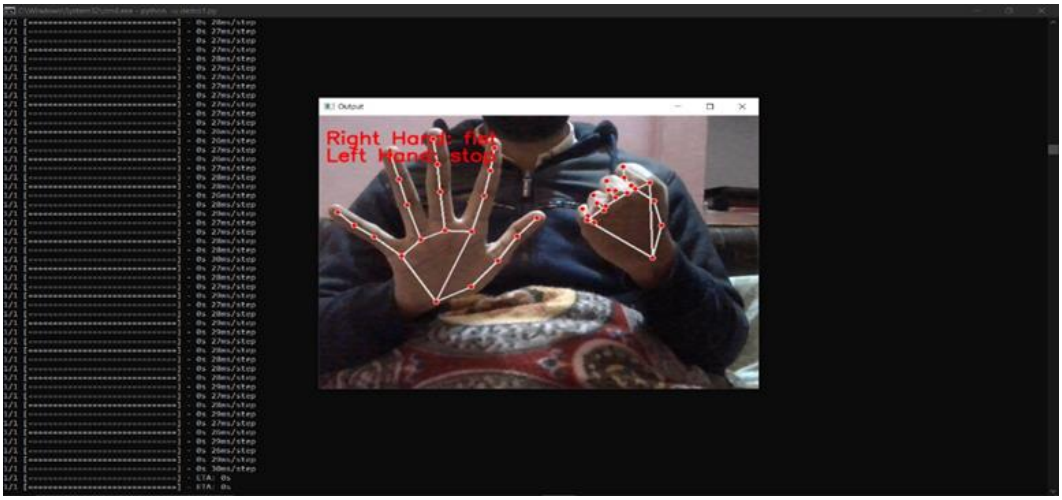

Fig.2 Right hand show fist and left-hand show stop action.

The Action Recognition System, utilizing TensorFlow and OpenCV, excels with a 95% detection rate, showcasing its prowess in real-time hand gesture recognition. The nuanced associations between gestures and actions enhance its applicability in human-computer interaction, security, and similar domains where precision matters. The system's high accuracy proves critical in practical scenarios, facilitating informed decision-making. Its success hints at versatile integration across fields, affirming its reliability as an effective solution for precise action recognition in varied contexts.

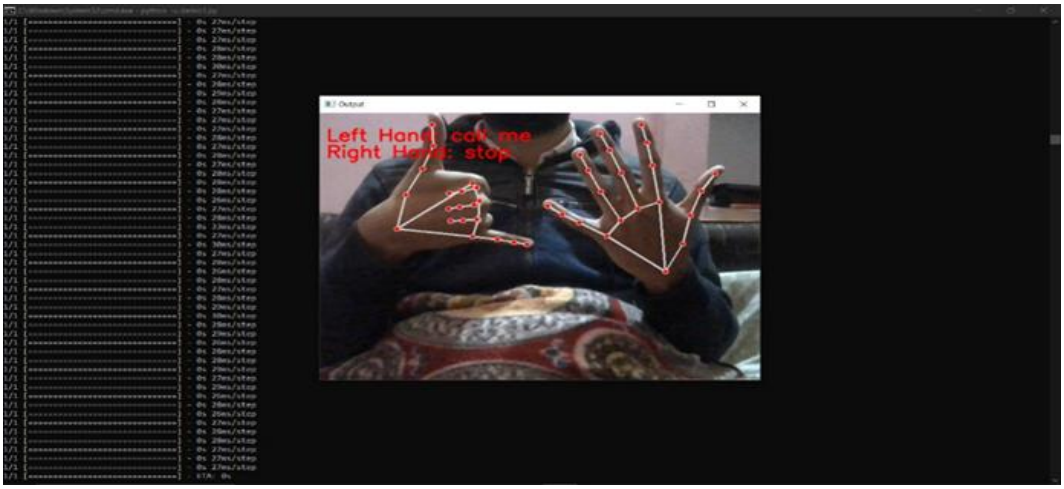

Fig.3 Right hand show stop and left-hand show call me action.

5. Conclusion And Future Scope

Sign languages use hand, body, and facial movements for communication, crucial for specially-abled people. However, not everyone knows sign language, which hinders communication. To solve this, an automated Sign Language Recognition system was created using TensorFlow object detection API. The Indian Sign Language alphabet dataset was used to train this system, which translates sign language into spoken words. The model uses a webcam and Python & OpenCV to capture images for data, making it cost-effective. The system boasts an impressive average confidence rate of around 85.45%, despite being trained on a small dataset. Looking ahead, enlarging the dataset will enable recognition of more gestures. It is flexible too - other models can replace the current TensorFlow one, and changing the dataset can facilitate recognition of different sign languages.

Reference

[1] "Sign Language Recognition System using TensorFlow Object Detection API" https://arxiv.org/ftp/arxiv/papers/2201/2201.01486.pdf
[2] **"**Sign Language Recognition: State of The Art"
[3] https://www.researchgate.net/publication/262187093_Sign_language_recognition_State_of_the_art
[4] "Machine Learning Methods for Sign Language Recognition: A Critical Review and Analysis" https://www.sciencedirect.com/science/article/pii/S2667305321000454#abs0002
[5] Pallavi Gurjal, Kiran Kunnur, "Real Time Hand Gesture Recognition using SIFT", International Journal for Electronics and Engineering, 2012
[6] Ghotkar, Archana S., "Hand Gesture Recognition for Indian Sign Language", *International Conference on Computer Communication and Informatics (ICCCI),* 2012. https://www.semanticscholar.org/paper/Hand-gesture-recognition-for-Indian-Sign-Language-Ghotkar-Khatal/995b10bcea63962c0a758c58ed85134bccad4fa1
[7] Dr. S. Nandagopal S. Priyanga journal on "Sign Language Detection using Python and OpenCV" https://ijetms.in/Vol-7-issue-2/Vol-7-Issue-2-55.html
[8] Byung-woo min, Ho-sub yoon, Jung soh, Takeshi ohashi and Toshiaki jima," Visual Recognition of Static/Dynamic Gesture: Gesture-Driven Editing System", *Journal of Visual Languages & Computing Volume10, Issue3,* June 1999
[9] "Hand sign recognition from depth images with multi-scale density features for deaf mute persons"
[10] Hee-Deok Yang "Sign Language Recognition with the kinetic sensor based on conditional random field" https://www.mdpi.com/1424-8220/15/1/135
[11] "Intelligent Sign Language Recognition System Using Image Processing"
[12] https://www.iosrjen.org/Papers/vol3_issue2%20(part-2)/H03224551.pdf