| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704125 |

Real Time Sign Language Recognition Using Machine Learning

R.KAVISHREE, JULAPPAGARI CHIRANJEEVI REDDY, KOMMA HEEMASHWAR REDDY, PADARTHI SAITEJA

Assistant Professor, Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

Department of CSE, Muthayammal Engineering College (Autonomous), Rasipuram, Tamil Nadu, India

ABSTRACT: Sign language is a vital form of communication for the deaf and hard of hearing community. To bridge the communication gap between sign language users and non-signers, we present a novel system for real-time sign language detection using a standard web camera. This system aims to recognize sign language gestures performed in front of the camera, subsequently converting them into voice output and displaying the corresponding text on screen. The proposed system leverages computer vision techniques, including deep learning models, to capture and analyze sign language gestures. It first identifies the signer's hand and facial expressions, recognizing key markers that represent signs. The system then employs machine learning algorithms to translate these markers into sign language vocabulary. Upon successful detection and translation of the sign, the system provides simultaneous output in two ways: voice and on-screen text. The voice output enables real-time interpretation for users who may not be familiar with sign language, while the on-screen text serves as a visual reference. This dual output mechanism ensures accessibility and inclusivity for a wider audience. By integrating this system into webcams and other devices with cameras, we aim to enhance the communication capabilities of the deaf and hard of hearing community, enabling them to interact more effectively with hearing individuals. Additionally, this technology can find applications in education, healthcare, and other domains, fostering better understanding and accessibility for sign language users.

KEYWORDS: Sign language, Deaf and hard of hearing, Real-time, Sign language detection, Web camera, Computer vision

INTRODUCTION

The project presented here is a groundbreaking endeavor that addresses a significant challenge faced by the deaf and hard of hearing community – the communication gap between sign language users and those who are not proficient in sign language. In this modern age where technology plays a pivotal role in enhancing our lives, a team of innovators has developed a novel system to bridge this gap using readily available tools – a standard web camera. Sign language is a vital and beautiful form of communication, allowing the deaf and hard of hearing to express themselves, connect with others, and access information. However, it remains a barrier for those who do not understand it, limiting inclusivity and access to vital services and information. This project aims to change that by introducing a real-time sign language detection system. The core of this system lies in cutting-edge computer vision techniques, particularly deep learning models, which enable it to capture and analyze sign language gestures accurately.

When a signer interacts with the camera, the system identifies the signer's hand movements and facial expressions, recognizing key markers that represent specific signs. These markers are then translated into sign language vocabulary through the use of advanced machine learning algorithms. What makes this project truly exceptional is its dual output mechanism. Upon successful detection and translation of a sign, it provides simultaneous output in two ways: voice and onscreen text. The voice output enables real-time interpretation, making it accessible for individuals who may not be familiar with sign language. At the same time, the on-screen text serves as a visual reference, providing an additional layer of accessibility. This dual output mechanism ensures that a wider audience can engage and interact with sign language users fostering inclusivity and understanding.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704125 |

Fig 1: Sign Language Recognition with Advanced Computer

The potential impact of this project is substantial. By integrating this system into webcams and various devices with cameras, it has the power to enhance the communication capabilities of the deaf and hard of hearing community, allowing them to interact more effectively with the hearing world. Furthermore, this technology can extend its reach to multiple domains, such as education, healthcare, and many others, contributing to better understanding and accessibility for sign language users. In doing so, it represents a significant step forward in promoting inclusivity and empowering the deaf and hard of hearing community, showcasing the transformative power of technology in modern society.

LITERATURE SURVEY

Sandrine Tornay et al. [1], Sign language recognition involves modeling of multichannel information such as, hand shapes, hand movements. This requires also sufficient sign language specific data. This is a challenge as sign languages are inherently under-resourced. In the literature, it has been shown that hand shape information can be estimated by pooling resources from multiple sign languages. Such a capability does not exist yet for modeling hand movement information. In this paper, we develop a multilingual sign language approach, where hand movement modeling is also done with target sign language independent data by derivation of hand movement subunits. We validate the proposed approach through an investigation on Swiss German Sign Language, German Sign Language and Turkish Sign Language, and demonstrate that sign language recognition systems can be effectively developed by using multilingual sign language resources.

Hira Hameed et al. [2] Sign language is a means of communication between the deaf community and normal hearing people who use hand gestures, facial expressions, and body language to communicate. It has the same level of complexity as spoken language, but it does not employ the same sentence structure as English. The motions in sign language comprise a range of distinct hand and finger articulations that are occasionally synchronized with the head, face, and body. Existing sign language recognition systems are mainly camera-based, which have fundamental limitations of poor lighting conditions, potential training challenges with longer video sequence data, and serious privacy

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704125 |

concerns. This study presents a first of its kind, contact-less and privacy-preserving British sign language (BSL) Recognition system using Radar and deep learning algorithms. Six most common emotions are considered in this proof of concept study, namely confused, depressed, happy, hate, lonely, and sad.
The collected data is represented in the form of spectrograms. Three state-of-the-art deep learning models, namely, InceptionV3, VGG19, and VGG16 models then extract spatiotemporal features from the spectrogram. Finally, BSL emotions are accurately identified by classifying the spectrograms into considered emotion signs. Comparative simulation results demonstrate that a maximum classifying accuracy of 93.33 is obtained on all classes using the VGG16 model. Aman Patel et al. [3]Sign language recognition plays an important role in real-time sign language translation, communication for deaf people, education and human-computer interaction. However, visionbased sign language recognition faces difficulties such as insufficient data, huge network models and poor timeliness. We use VTN (Video Transformer Net) to construct a lightweight sign language translation network. We construct the dataset called CSLBS (Chinese Sign Language-Bank and Station) and two-way VTN to train isolated sign language and compares it with I3D (Inflated three Dimension).

METHODS

In imaging science, image processing is processing of images using mathematical operations by using any form of signal processing for which the input is an image, a series of images, or a video, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it. Images are also processed as three-dimensional signals with the third-dimension being time or the z-axis. Image processing usually refers to digital image processing, but optical and analog image processing also are possible. This article is about general techniques that apply to all of them. The acquisition of images (producing the input image in the first place) is referred to as imaging. Closely related to image processing are computer graphics and computer vision. In computer graphics, images are manually made from physical models of objects, environments, and lighting, instead of being acquired (via imaging devices such as cameras) from natural scenes, as in most animated movies. Computer vision, on the other hand, is often considered high-level image processing out of which a machine/computer/software intends to decipher the physical contents of an image or a sequence of images (e.g., videos r 3D full-body magnetic resonance scans). [6-11] In modern sciences and technologies, images also gain much broader scopes due to the ever-growing importance of scientific visualization (of often large-scale complex scientific/experimental data). Examples include microarray data in genetic research, or real-time multi-asset portfolio trading in finance. Image analysis is the extraction of meaningful information from images; mainly from digital images by means of digital image processing techniques. Image analysis tasks can be as simple as reading bar coded tags or as sophisticated as identifying a person from their face.

Machine learning algorithms [1-5] are often categorized as supervised or unsupervised. Supervised algorithms require a data scientist or data analyst with machine learning skills to provide both input and desired output, in addition to furnishing feedback about the accuracy of predictions during algorithm training. Data scientists determine which variables, or features, the model should analyze and use to develop predictions. Once training is complete, the algorithm will apply what was learned to new data. Unsupervised algorithms do not need to be trained with desired outcome data. Instead, they use an iterative approach called deep learning to review data and arrive at conclusions. Unsupervised learning algorithms — also called neural networks are used for more complex processing tasks than supervised learning systems, including image recognition, speech-to-text and natural language generation. These neural networks work by combing through millions of examples of training data and automatically identifying often subtle correlations between many variables. Once trained, the algorithm can use its bank of associations to interpret new data. These algorithms have only become feasible in the age of big data, as they require massive amounts of training data. Machine learning algorithms are often categorized as supervised or unsupervised.

RESULT ANALYSIS

Tensorflow Object detection API is the framework for creating a deep learning network that solves object detection problems. This module consists of a variety of detection models which are pre-trained on the COCO 2017 dataset. SSD MobileNet V2 FPNLite 320x320 and SSD ResNet50 V1 FPN 640x640 are used for the detection of static and real-time signs in this research. Transfer learning is used to train the two models on the dataset generated in this research. The SSD (Single Shot Detector) architecture is a single convolution network that learns to predict bounding box locations and

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704125 |

classify these locations in one pass. The step by step procedure is divided into 5 steps which are- data acquisition, training the network, testing the network, sign recognition in real-time and results.

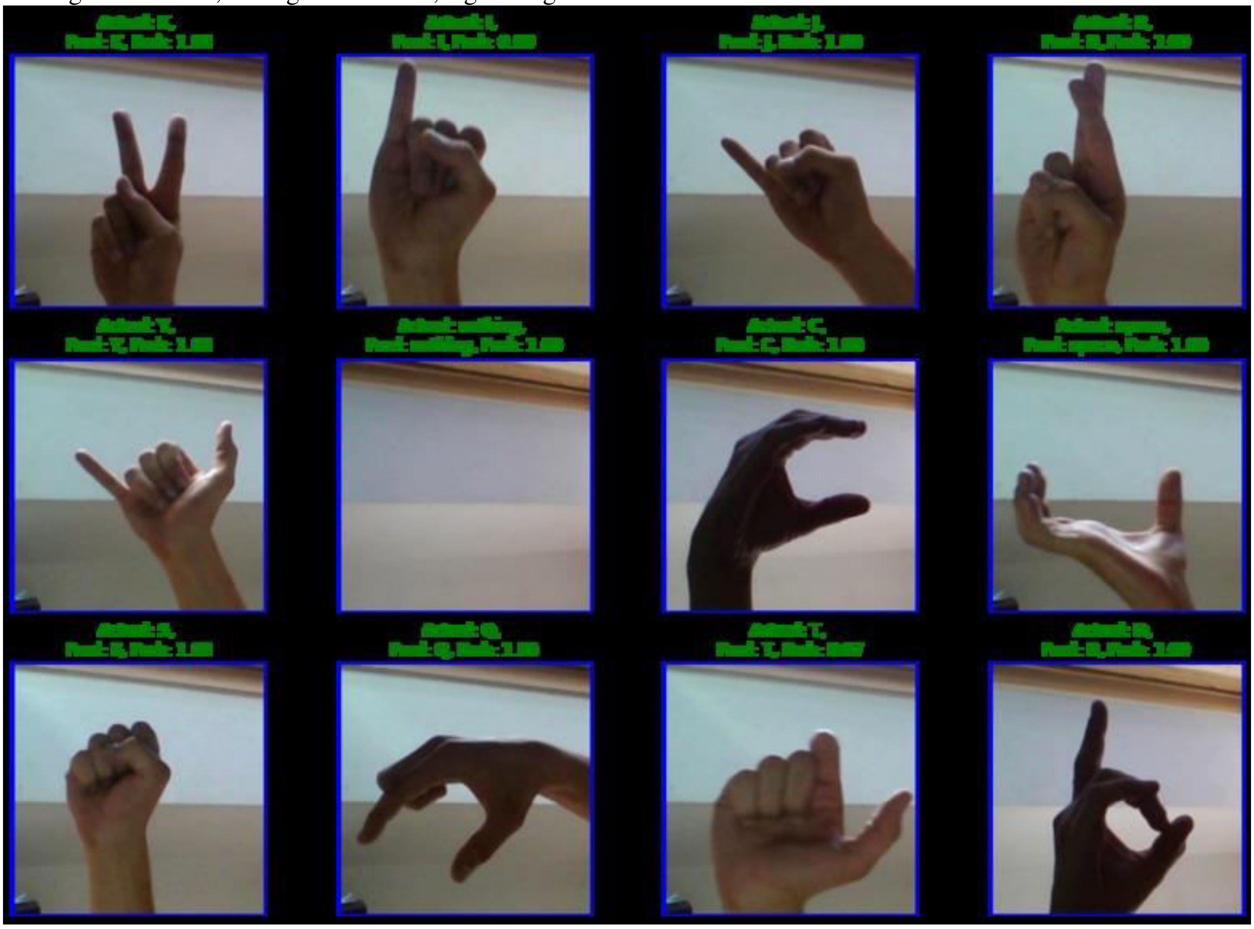

Fig 2: Result analysis Sign Language Recognition

A dataset of signs of five words ‘hello’, ‘yes’, ‘no’, ‘thank you’ ‘thumbs up’, ‘thumbs down’ and ‘one’ was created using openCV. Images were captured using web camera. 100 images of each word were captured so the dataset consisted of 100x8 = 800 images for training from a single user and testing dataset consisted of 15x8 = 120 images. Labeling of images was done using the graphical image annotation tool labelImg where only the hand part was annotated

CONCLUSION

The goal of our survey paper has been to shed lighton the unmet technological needs of the Deaf and Hard of Hearing (DHH) community, with a particular focuson Indian Sign Language recognition. In this pursuit, our study proposes the creation of a Real-Time Indian Sign Language Recognition system using a hybrid CNN approach. Although this system is currently in the conceptual phase, our comprehensive analysis of existing methodologies and their limitations provides in- valuable insights for its future implementation. By advocating for technology that aligns with the unique requirements of the DHH community, our survey paper underscores the significance of socially responsible technological advancements. By demonstrating the importance of tailored solutions catering to the diverse needs of individuals, especially marginalized communities like the DHH, our research paves the way for a more inclusive technological landscape. This approach emphasizes the necessity of considering social responsibility in technology design, encouraging future researchers and developers to create solutions that prioritize accessibility and inclusivity. Collaborative efforts between researchers, technologists, and the DHH community will be crucial in bringing this vision to fruition.

REFERENCES

[1] Goyal, Kaushal. ”Indian Sign Language Recognition Using Mediapipe Holistic.” arXiv preprint arXiv:2304.10256 (2023).

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704125 |

[2] Mohammedali, Atyaf Hekmat, Hawraa H. Abbas, and Haider Ismael Shahadi. ”Real-time sign language recognition system.” International Journal of Health Sciences 6 (2022): 10384-10407.
[3] Shenoy, Kartik, et al. ”Real-time Indian sign language (ISL) recognition.” 2018 9th international conference on computing, communication and networking technologies (ICCCNT). IEEE, 2018.
[4] Murali, Romala Sri Lakshmi, L. D. Ramayya, and V. Anil Santosh. ”Sign language recognition system using convolutional neural network and computer vision.” (2020).
[5] Puranik, V., Gawande, V., Gujarathi, J., Patani, A., Rane, T.(2022, August). Video-based Sign Language Recognition using Recurrent Neural Networks. In 2022 2nd Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-6). IEEE.
[6] Katoch, Shagun, Varsha Singh, and Uma Shanker Tiwary. ”Indian Sign Language recognition system using SURF with SVM and CNN.” Array 14 (2022): 100141.
[7] Kasapbas¸i, Ahmed, Ahmed Eltayeb AHMED Elbushra, AL- HARDANEE Omar, and Arif Yilmaz. ”DeepASLR: A CNN based human computer interface for American Sign Language recognition for hearingimpaired individuals.” Computer Methods and Programs in Biomedicine Update 2 (2022): 100048.
[8] Sundar, B., and T. Bagyammal. ”American Sign Language Recognition for Alphabets Using MediaPipe and LSTM.” Procedia Computer Science 215 (2022): 642-651.
[9] Kothadiya, Deep, Chintan Bhatt, Krenil Sapariya, Kevin Patel, AnaBele´n Gil-Gonza´lez, and Juan M. Corchado. ”Deepsign: Sign language detection and recognition using deep learning.” Electronics 11, no. 11 (2022): 1780.
[10] Piyush Kapoor, Hema N:Sign Language and Common Gesture Using CNN. Electronics (2021):2278-3091 [11] Yu Liu, Parma Nand, Md Akbar Hossain, Minh Nguyen and Wei Qi Yan. Sign language recognition from digital videos using feature pyramid network with detection transforme. Springer(2022):21673–21685
[12] Niloy Sikder, Abu Shamim Mohammad Arif, Md. Sanaullah Chowdhury and Abdullah-Al Nahid. Human Activity Recognition Using Multichannel Convolutional Neural Network. (2019) IEEE
[13] R.S. Sabeenian, S. Sai Bharathwaj and M. Mohamed Aadhil. Sign Language Recognition Using Deep Learning and Computer Vision(2020) ISSN: 1943-023X
[14] Sharvani Srivastava, Amisha Gangwar, Richa Mishra and Sudhakar Singh. Sign Language Recognition System using TensorFlow Object Detection API. Springer 2021.