SIGN LANGUAGE: A TRANSLATION USING MEDIAPIPE AND CNN

MRS.A.NANDHINI MCA, (PH.D),
ASSISTANT PROFESSOR,
DEPARTMENT OF COMPUTER APPLICATIONS,
NEHRU COLLEGE OF MANAGEMENT.

MR.AFSAL.N II MCA,
DEPARTMENT OF COMPUTER APPLICATIONS,
NEHRU COLLEGE OF MANAGEMENT.

ABSTRACT - **Conversation with a person having a hearing or speech disability has always been a challenge. Sign language has undoubtably become the major tool for the people with hearing and speech disability to communicate their thoughts and feeling to the world. It helps them to integrate themselves into the society. With the advances in machine learning, it is possible to detect sign language in real time. The project aims at building a machine learning model that will be able to classify the various hand gestures used for fingerspelling in sign language. In this user independent model, classification machine learning algorithms are trained using a set of image data and testing is done on a completely different set of data.**

INDEX : Sign Language, Mediapipe, Co-ordinates conversion using Numpy, Sign Gesture Detection using CNN Algorithm, Sign Gesture Dataset.

I.INTRODUCTION

Conversation with a person having a hearing or speech disability has always been a challenge. Sign language has undoubtably become the major tool for the people with hearing and speech disability to communicate their thoughts and feeling to the world. It helps them to integrate themselves into the society. However, sign language itself is not enough. With this boon comes a problem, the signs are often mixed up and are confusing to someone who has never learnt it or has learnt it in a different language. Throughout the long term, correspondence has played an indispensable job in return of data and sentiments in one's day to day existence. Sign language is the main medium through which deaf and mute individuals can interact with rest of the world through various hand motions. With the advances in machine learning, it is possible to detect sign language in real time. The project aims at building a machine learning model that will be able to classify the various hand gestures used for fingerspelling in sign language. In this user independent model, classification machine learning algorithms are trained using a set of image data and testing is done on a completely different set of data. In this project we are presenting Sign Language Recognition System using Sign Language. Using this system, the user will be able to use the webcam or camera to predict the sign language gestures in real-time. We are using various images of the Sign Language gestures and processing them with various Computer Vision techniques and creating a Machine Learning model using it. We have used Deep Learning model in this project, as the name suggests, Deep Learning Neural Networks, or Artificial Neural Networks, attempts to mimic the human brain through a combination of data inputs, weights, and bias. These elements work together to accurately recognize, classify, and describe objects within the data. Deep Neural Networks consist of multiple layers of interconnected nodes, each building upon the previous layer to refine and optimize the prediction or categorization. . This progression of computations through the network is called forward propagation. The input and output layers of a deep neural network are called visible layers. The input layer is where the deep learning model ingests the data for processing, and the output layer is where the final prediction or classification is made. We have utilized the OpenCv python library, Tensorflow Object Detection pipeline and transfer learning to train a deep learning model that detects sign languages in Real time.

II. LITERATURE REVIEW :

The purpose of performing this systematic literature review was to select and categorize the best available approaches to sign language detection using neural networks (NNs). Systematic literature reviews collect and analyze existing studies according to predefined evaluation criteria. Such reviews help to determine what is already known in the concerned domain of study.

Sign language is a form of communication that uses visual-manual language to convey meaning through hand gestures, facial expressions, and body movements. It is used by deaf and hard-of-hearing individuals as well as people with hearing abilities who may communicate with them. Sign languages are fully developed natural languages with their own grammar, syntax, and lexicon, distinct from spoken languages.

There are many different sign languages around the world, with some of the most widely known including American Sign Language (ASL), British Sign Language (BSL), Auslan (Australian Sign Language), and LSF

(Langue des Signes Française or French Sign Language). Each sign language has its own vocabulary and grammar, and they are not mutually intelligible.

Learning sign language can be a valuable skill for enhancing communication and inclusivity, as it enables individuals to connect with deaf or hard-of-hearing people more effectively.

MediaPipe is an open-source framework developed by Google that provides a comprehensive solution for building multimodal (e.g., vision and audio) applied machine learning (ML) pipelines. It offers a variety of pre-built components and models for tasks such as object detection, face detection, hand tracking, pose estimation, and more. MediaPipe aims to simplify the development of complex multimedia processing pipelines by providing ready-to-use building blocks that developers can integrate into their applications.

MediaPipe is widely used in a range of applications, including augmented reality (AR), virtual reality (VR), gesture recognition, sign language recognition, video analysis, and content creation. Its flexibility, performance, and ease of integration make it a popular choice for developers working on multimedia projects.

The NumPy's main object is the homogeneous multidimensional array. These arrays can be of any dimension, and all elements are of the same data type. NumPy arrays are more efficient in terms of memory usage and performance compared to Python lists.

NumPy's main object is the homogeneous multidimensional array. These arrays can be of any dimension, and all elements are of the same data type. NumPy arrays are more efficient in terms of memory usage and performance compared to Python lists. NumPy is a fundamental library in the Python scientific computing ecosystem, providing the building blocks for numerical computation and data manipulation. Its simplicity, efficiency, and extensive capabilities make it an essential tool for researchers, scientists, and engineers.

Dataset Building a sign language detection system using deep learning typically requires a dataset consisting of images or videos of sign language gestures along with corresponding labels indicating the signs being performed.

American Sign Language (ASL): This dataset consists of images of the ASL alphabet gestures. Each image represents a letter from A to Z, captured by various signers. It's commonly used for gesture recognition and finger spelling tasks.

The Convolutional Neural network (CNN) is an extension to the existing neural network[4]. Convolutional neural networks are preferred over fully connected feedforward neural layers in digital imaging applications due to the spark connectivity and weight sharing properties of image pixels. CNNs can also be tuned for different mathematical learning methods such as back propagation, learning algorithms and regularisation techniques.

CNN's hidden layer consists of convolution layers, nonlinear pooling layers and fully connected layers. CNN contains multiple convolution layers that are followed by several fully connected layers. Three important layers involved in making CNN are convolution layers, pooling layers and full-connected layers. In CNN, the convolution layer manages a set of weights which are reduced by pooling layers to give output from the convolution and reduces the input size ratio. After the convolutional layer, the output from the pooling layer is used and fed to the fully connected layer. An essential section of CNN is the Convolutional layer which consists of a variety of weights for different applications like image segmentation and multiple 2D matrices .

METHODOLOGY :

Understanding sign language is difficult task for people who never learned sign language. With the extensive use of deep learning procedures helps to understand sign gestures. Deep learning methods like Convolutional Neural Networks (CNN), has surpassed all others in hand gesture detection and classification tests. English alphabets sign representation shown in , Fig .1 below:

Fig 1. Different Sign alphabets

To identify the user hand gesture by using above mentioned methodology, we uses the following process shown in fig 2.

It is used for classifying system requirements to major transformation that will become programs in system design. This is starting point of the design phase that functionally decomposes the required specifications down to the lower level of details. It consists of a series of bubbles joined together by lines.

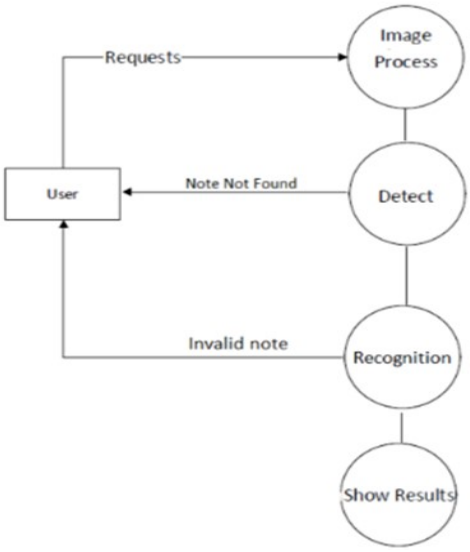

Fig 2. Flow of methodology

DATASET DESCRIPTION :

In this paper we use sign language dataset for predicting the hand gesture given by user. The American Sign Language (ASL) Alphabet dataset consists of images or videos of hand gestures representing individual letters of the ASL alphabet. Here's how you can obtain or create such a dataset for sign language translation tasks:

OpenSLR is a repository of open-source datasets for speech and language processing tasks. While it primarily focuses on speech-related datasets, you may find sign language datasets such as the ASL Alphabet dataset.

Kaggle hosts various datasets, competitions, and kernels. You can search Kaggle for ASL Alphabet datasets, and you may find publicly available datasets containing images or videos of ASL hand gestures representing each letter of the alphabet.

Create your own dataset if you cannot find a suitable dataset, you can create your own by recording videos or capturing images of ASL alphabet gestures using a camera or webcam. Ensure that you have a diverse set of hand configurations, lighting conditions, backgrounds, and signers to make the dataset more robust.

Once you have obtained or created the ASL Alphabet dataset, ensure that it is appropriately labeled with each image or video corresponding to the correct letter of the ASL alphabet. This labeled dataset can then be used to train deep learning models for sign language translation tasks, where the goal is to translate ASL gestures into text or spoken language

EXISTING AND PROPOSED METHODOLOGY

EXISTING:

Sign language is a visual language and consists of major components. There is no such existing system that implements deep learning techniques that have been proposed in this project, some attempts to use ML for sign recognition have been made but are less accurate. Developing software applications to assist sign language communication with deaf people is an important area of work. Applications of this type could be used to aid deaf people so that they can interact with computer systems using sign language. Translation systems built on automated sign language recognition (ASLR) could assist- 3 -communication between deaf people and people who may not know sign language. Additionally, ASLR systems could be used to aid in the teaching of sign languages.

PROPOSED:

images. The system uses the web camera for capturing the hand gestures of the hearing and speech impaired people. The raw images taken are given as an input to the system. The image frames are resized to maintain the equality among all the images. The proposed system includes several steps namely, Image Acquisition, Image Pre-Processing, Segmentation, Feature Extraction, Recognition and Text Output.

ADVANTAGES OF PROPOSED SYSTEM

- In our project we basically focus on producing a model which can recognize
- In finger spelling based hand gestures in order to form a complete word by combining each gesture.
- The gestures we aim to train are as given in the output fast
- User can easy to use

CONCLUSION:

In conclusion, sign language detection systems hold tremendous potential for transforming communication and fostering inclusivity for individuals who use sign language. These systems have the ability to bridge the gap between sign language and spoken or written language, enabling seamless interaction with non-signers in various settings.By continually improving sign language detection systems, we can empower individuals who use sign language in education, employment, healthcare, and social interactions. Breaking down communication barriers and promoting inclusivity will create a more equitable society where everyone can communicate effectively and participate fully in various aspects of life. As technology continues to evolve, it is crucial to prioritize research, development, and collaboration to realize the full potential of sign language detection systems. Through ongoing innovation and efforts, we can contribute to a more inclusive and accessible future for individuals who use sign language.

REFERENCES

[1] T. Starner, J. Weaver, A. Pentland. Real-Time American Sign Language Recognition from Video Using Hidden Markov Models, M.I.T. Media Laboratory Perceptual Computing Section Technical Report No. 375, 1996.

[2] H. Hienz, K. Grobel, Automatic Estimation of Body Regions from Video Images, in I. Wachsmuth and M. Fröhlich (Eds.) Gesture and Sign Language in Human Computer Interaction (Berlin: Springer-Verlag,1998), pp. 135-145.

[3] S. He. (2019). Research of a Sign Language Translation System Based on Deep Learning, pp. 392-396. 10.1109/AIAM48774.2019.00083.

[4] Herath, H.C.M. & W.A.L.V. Kumari, &Senevirathne, W.A.P.B & Dissanayake, Maheshi. (2013). IMAGE BASED SIGN LANGUAGE RECOGNITION SYSTEM FOR SINHALA SIGN LANGUAGE.

[5] M. Geetha and U. C. Manjusha, "A Vision Based Recognition of Indian Sign Language Alphabets and Numerals Using B-Spline Approximation", International Journal on Computer Science and Engineering (IJCSE), vol. 4, no. 3, pp. 406-415. 2012.

[6] Pigou L., Dieleman S., Kindermans PJ., Schrauwen B. (2015) Sign Language Recognition Using Convolutional Neural Networks. In: Agapito L., Bronstein M., Rother C. (eds) Computer Vision - ECCV 2014 Workshops. ECCV 2014. Lecture Notes in Computer Science, vol. 8925. Springer, Cham.

[7] Escalera S., Baró X., Gonzàlez J., Bautista M.A., Madadi, M., Reyes M., Ponce-López V, Escalante H.J., Shotton J., Guyon, I. (2014). ChaLearn Looking at People Challenge 2014: Dataset and Results. Workshop at the European Conference on Computer Vision (pp. 459-473). Springer, Cham.

[8] Huang J., Zhou W., & Li H. (2015). Sign Language Recognition using 3D convolutional neural networks. IEEE International Conference on Multimedia and Expo (ICME) (pp. 1-6). Turin: IEEE.

[9] Edon Mustafa. Sign Language Interpretation using Kinect, MSc in Software Engineering and Telecommunications, The University of Sheffield. 2014.

[10] https://www.kaggle.com/ - Sign Language MNIST Dataset.(Accessed on 25th July, 2022)

[11] C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of ELECTRICAL ENGINEERING, Vol.63 (6), pp.365-372, Dec.2012.

[12] C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011.

[13] C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor & Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011.

[14] C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.

[15] Nagarajan C., Neelakrishnan G., Akila P., Fathima U., Sneha S. "Performance Analysis and Implementation of 89C51 Controller Based Solar Tracking System with Boost Converter" Journal of VLSI Design Tools & Technology. 2022; 12(2): 34–41p.

[16] C. Nagarajan, G.Neelakrishnan, R. Janani, S.Maithili, G. Ramya "Investigation on Fault Analysis for Power Transformers Using Adaptive Differential Relay" Asian Journal of Electrical Science, Vol.11 No.1, pp: 1-8, 2022.

[17] G.Neelakrishnan, K.Anandhakumar, A.Prathap, S.Prakash "Performance Estimation of cascaded h-bridge MLI for HEV using SVPWM" Suraj Punj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:750-756

[18] G.Neelakrishnan, S.N.Pruthika, P.T.Shalini, S.Soniya, "Perfromance Investigation of T-Source Inverter fed with Solar Cell" Suraj Punj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:744-749

[19] C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007

[20] M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022

[21] M Suganthi, N Ramesh, CT Sivakumar, K Vidhya, "Physiochemical Analysis of Ground Water used for Domestic needs in the Area of Perundurai in Erode District", International Research Journal of Multidisciplinary Technovation, pp: 630-635, 2019