IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12624

MOTION TRAJECTORY BASED HUMAN HAND TRACKING FOR SIGN LANGUAGE RECOGNITION

Jagdale Mrudula Dattatraya[1], Kasture Rushikesh Sunil[2], Jadhav Rushabh Pratap[3], Prof.M.M.Jadhav[4]

Bachelor Of Engineering, Electronics and Telecommunication, Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering and Technology, Baramati, India[1,2,3]

Assistant Professor, Electronics and Telecommunication ,Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering and Technology, Baramati, India[4]

Abstract*:* Sign language recognition is an important area of research with uses in a range of industries, including education, communication, and human-computer interaction. Trajectory-based human hand tracking is a popular technique for sign language recognition because it allows capturing dynamic hand movement during signing. This paper presents an overview of different trajectory-based human hand tracking methods, including template matching, feature-based tracking, and model-based tracking. This framework aims to accurately identify various one-handed dynamic isolated characters and interpret their intended meaning. Emphasis is placed on designing a system that can effectively recognize and interpret a wide variety of sign gestures, enabling effective communication and understanding between sign language users and individuals who do not know sign language.

Keywords: pre-processing, feature extraction, recognition.

I. INTRODUCTION

In the last two decades, man has made tremendous progress in the field of artificial intelligence and data processing. It has been beneficial to all areas in every aspect such as growth of company sales, easy and fast customer service and clarity of transactions. We can use this type of technology to make life easier and more enjoyable for people with disabilities like normal people. With this idea in mind, we decided to develop a system that can facilitate communication for mute or disabled people. An explanation of the problem involves the development of a computer vision system that can precisely monitor a user's hands in motion in real time and recognize the corresponding gestures in sign language. The goal is to create a system that can work in different lighting conditions and camera angles, while still being able to detect and track hands even when they overlap or occlude each other.

The proposed work aims to develop a machine learning model for predicting trajectories of moving objects. The goals of this project include obtaining or creating a dataset for the generated trajectories that can be used to develop and evaluate the model. The dataset will be pre-processed to ensure it is suitable for model building, which may include tasks such as normalization, feature selection, and data augmentation. Once the dataset is ready, the test dataset will be used to assess the model's performance once it has been trained using the train dataset.

II. LITERATURE SURVEY

This project includes an examination of relevant research papers and contributions by respective authors that are consistent with the objectives of the system. Observations and findings from these research works are thoroughly analysed in the proposed work. These documents closely match the objectives of the system and provide valuable insights into the subject. Recognition of sign language has come a long way, and one significant study in this field is the work. of N. Adaloglou et al. titled "A Comprehensive Study on Deep Learning-Based Sign Language Recognition Methods" (2022) [1] . The study evaluates computer vision-based approaches for recognising signs, employs cutting-edge deep learning algorithms, and examines a number of publicly accessible datasets. The primary aim of the study is to contribute to the identification of sign language, with a special emphasis on the mapping of unsegmented video streams to glosses.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12624

The sign language recognition system underwent an extensive investigation in 2019 by Rajesh George Rajan et al. [2]. The system has three key steps: feature extraction, pre-processing, and classification. During the pre-processing step, the system detects the hand region from images or videos, which is key to isolating the hand and eliminating unnecessary background information. In the feature extraction step, various features are extracted from the feature image or video to create a feature vector that represents the feature. These signs capture the basic characteristics of the sign, including hand shape, movement and position.

In 2016, Joyeet Singha et al. presented an approach for hand gesture recognition that used the Karhunen-Loeve Transform (KLT) specifically for two-handed character recognition [3]. Their study focused on the recognition of 23 alphabets using a two-stage recognition approach. To allow for proper segmentation, the signers wore red gloves on both hands. The segmented hand images were then used as input for subsequent feature extraction and recognition stages. In the first phase, features capturing the overall shape of the gestures were calculated. The recognition process involved training feature vectors without the need for a classifier. [3] Joyeeta Singh et al., “Hand Gesture Recognition Based on Karhunen-Loeve Transform”, 2016.

"A real-time system for recognising isolated Swedish Sign Language characters utilising binary coefficients and a rigorous recognition criterion is presented in "A Real-Time Gesture Recognition System for Isolated Swedish Sign Language Characters"[4] (2016) by Kalin Stefano vet Alan approach based on Hidden Markov Models (HMM) and Principal Component Analysis (PCA) functions is proposed by Gerald Krell et al. in "Gesture recognition for alphabets from hand trajectories using hidden Markov models"[5] (2007) for recognising alphabets in sign language. Iker Vazquez Lopez et al.'s "Work on Two-Handed Signs" (2020) [6] expands on earlier work and incorporates a number of components for ambidextrous character identification, including hand shape detection, texture analysis, finger feature extraction, and hand region colour segmentation. (2016) [7] used "A Real-time Gesture Recognition System for Isolated Swedish Sign Language Characters" ,Hidden Markov models (HMMs) and principal component analysis (PCA) techniques were used in "Gesture recognition for alphabets from hand motion trajectories using hidden Markov models" (2007) to achieve high recognition rates for alphabets in signed language.

III. METHODOLOGY

A vision forms the system's foundation. The use of any artificial gadgets for interaction is not a concern because all signs are expressed by naked hands.

Figure illustrates the architectural design of our system, presenting a comprehensive overview of the process flow. It provides a clear understanding of how the system will progress, leading to the generation of the desired end-result. The diagram encapsulates the entire process, outlining the various stages and interactions involved. By studying this architecture, stakeholders gain valuable insights into the system’s functionality, enabling them to visualize the systematic execution of tasks that culminate in the final outcome.

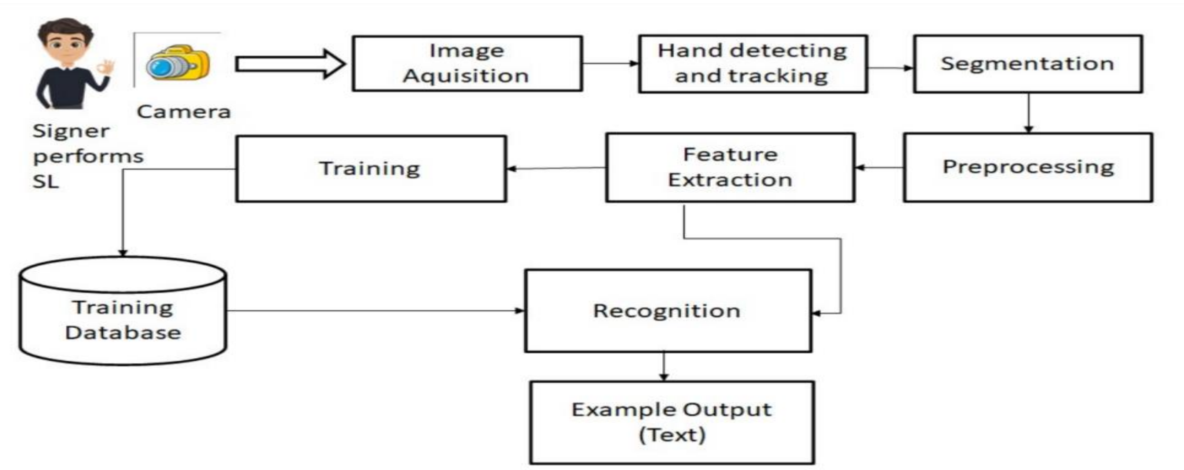

Fig. 1 Proposed System

DOI: 10.17148/IJARCCE.2023.12624

Convolutional Neural Networks (CNNs) use their capacity to extract pertinent elements from visual data to recognise sign language very effectively. There are several important layers and processes in the CNN architecture for understanding sign language. A broad dataset of sign language photos or videos is gathered at the outset of the process, spanning a wide range of movements and variations. The pre-processing of the acquired data includes downsizing the photos, making them grayscale, and normalising the pixel values. Training and testing versions of the dataset are separated.

The CNN model is trained using the training set, while After convolution and pooling, the feature maps are merged into a one-dimensional vector representation. Fully connected layers learn to combine features and map them to specific sign language classes. The output layer generates predictions representing the recognized gestures or text. The CNN model goes through training using the training set, weight adjustment using backpropagation and optimization algorithms. The training process continues for several epochs until satisfactory performance is achieved. Once training is complete, the model is evaluated on a test set to assess its generalization and performance on unseen data. Metrics such as accuracy are calculated to measure effectiveness.

A CNN implementation involves specific steps: convolution perations, ReLU layer, pooling (max pooling), flattening, full join (dense layer), optimizer selection and compilation. Using CNN in sign language recognition, the model can automatically learn and extract distinguishing features. This technology can facilitate communication and bridge the gap between deaf and hearing communities, benefiting individuals with hearing impairments.

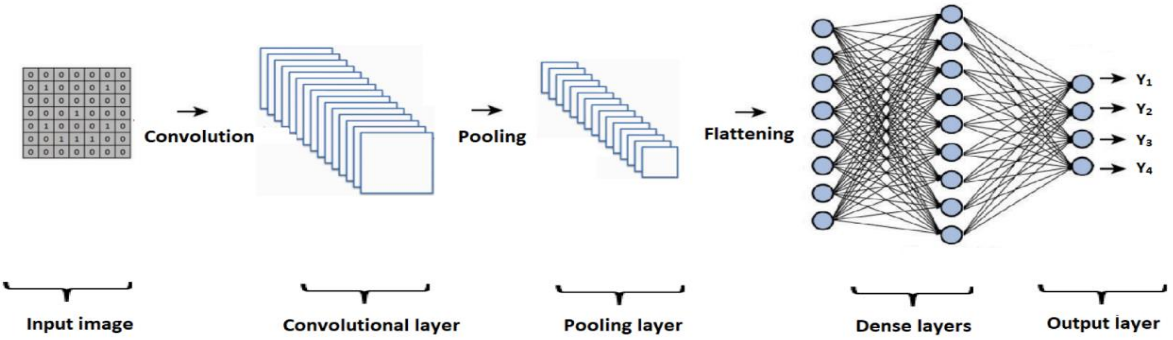

Fig. 2 CNN Architecture

Let's break down each step in more detail:
Step 1: Convolution operation (filter image):
Use filters to extract local patterns and features from an image.
Step 1(b): ReLU layer:
Introduce non-linearity using the ReLU activation function to capture complex relationships.
Step 2: Pooling (maximum pooling):
Resampling feature maps by selecting maximum values in sliding windows.
Step 3: Flattening:
Convert the feature maps to a one-dimensional vector for further processing.
Step 4: Complete Connection:
Connect all neurons to analyze and combine features for predictions.
Step 4(b): Dense ():
Fully connected layer with adjustable number of neurons.
Step 4(c): Optimizer ():
Algorithm for adjusting network weights during training (e.g. Adam, RMSprop, SGD).
Step 4(d): compile():
Prepare the model for training by setting the learning rate, metrics, and optimization settings.

IV. RESULTS

In the project, we have trained a model to achieve impressive results, as depicted by the accuracy graph. Throughout the training process, the model consistently exhibited excellent performance, with a remarkably low loss of 0.0289. Moreover, the accuracy achieved by the model on the training data stands at an impressive 0.9918, indicating its ability to correctly classify a vast majority of instances. The results obtained from this evaluation are equally noteworthy, with

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁑ Impact Factor 8.102 ⁑ Peer-reviewed / Refereed journal ⁑ Vol. 12, Issue 6, June 2023
DOI: 10.17148/IJARCCE.2023.12624

a validation loss of 0.1278. Although slightly higher than the training loss. The validation accuracy of 0.9731 further reinforces the capability to perform well on instances outside of the training set.

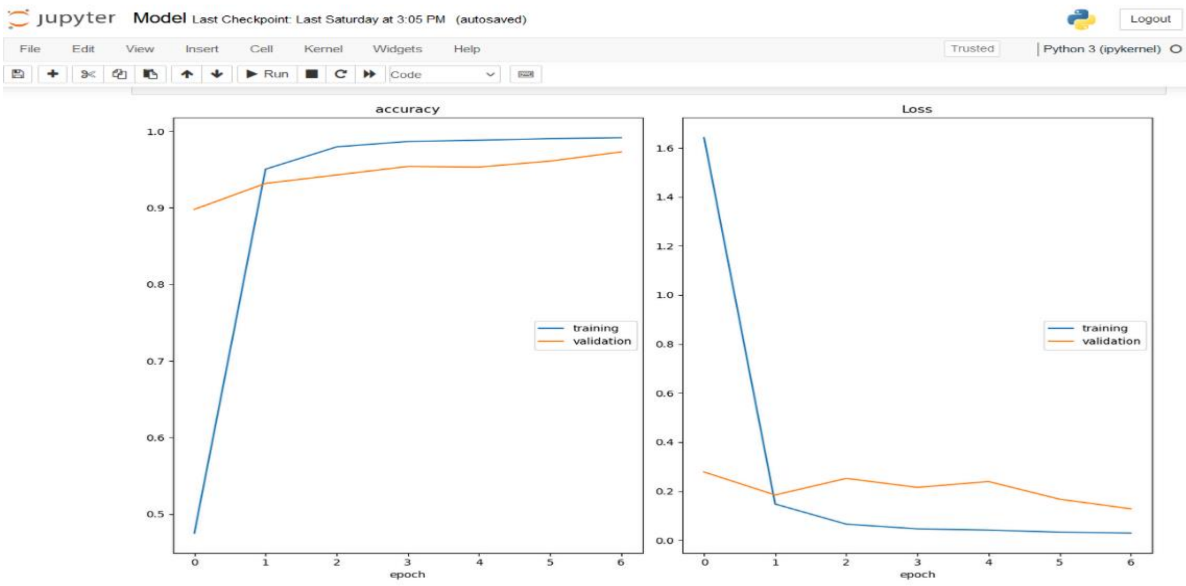

Fig .3 Accuracy and loss

Throughout the training process, a learning rate of 0.0010 was employed. This value, chosen to optimize the model's convergence and performance, contributed to the model's ability to learn and adapt to the training data effectively.

```
accuracy
        training              (min:    0.475, max:    0.992, cur:    0.992)
        validation            (min:    0.898, max:    0.973, cur:    0.973)
Loss
        training              (min:    0.029, max:    1.642, cur:    0.029)
        validation            (min:    0.128, max:    0.278, cur:    0.128)
857/857 [==============================] - 58s 67ms/step - loss: 0.0289 - accuracy: 0.9918 - val_loss: 0.1278 - val_accuracy:
0.9731 - lr: 0.0010
Wall time: 6min 48s

In [13]: score = model.evaluate(X_test, y_test, verbose = 0)
         print('Test loss:', score[0])
         print('Test accuracy:', score[1])

Test loss: 0.12777647376060486
Test accuracy: 0.9730898141860962
```

Fig. 4 Code of Accuracy and loss

Overall, the accuracy graph and the associated metrics highlight the remarkable performance of our trained model, showcasing its ability to accurately classify instances and generalize well to unseen data. These results provide confidence in the model's effectiveness and reliability for the given project.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁑ Impact Factor 8.102 ⁑ Peer-reviewed / Refereed journal ⁑ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12624

Result Outputs

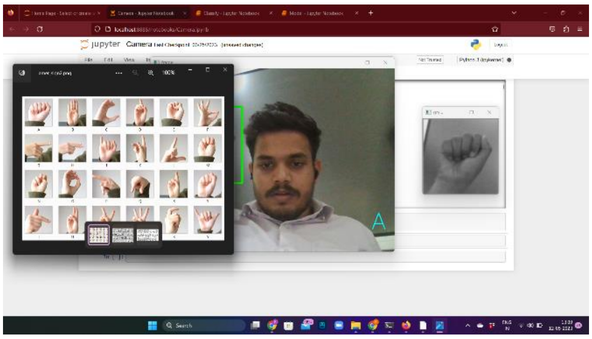

Fig.5 Result of A gesture

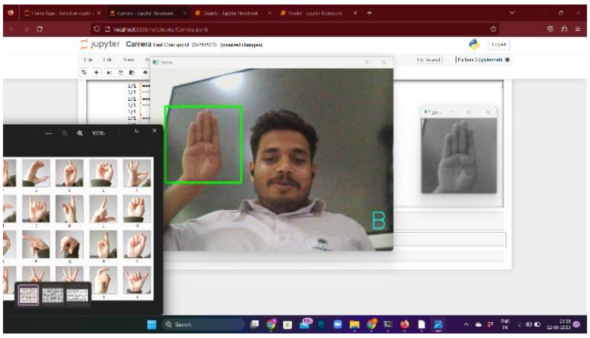

Fig.6 Result of B gesture

V. CONCLUSION

In summary, human hand tracking for motion trajectory-based sign language recognition is an innovative technology with huge potential. This approach, which focuses on capturing and analyzing the complex movement trajectories of the human hand, offers a reliable and robust solution for real-time sign language gesture interpretation.

One of this technology's primary benefits is its capacity to get beyond frequent issues with sign language recognition, like occlusions and variations in illumination. The system can capture the expressive nature and nuanced meanings sent through particular hand gestures, postures, and trajectories by precisely tracking hand movement. Accurate sign language interpretation and efficient communication depend on this level of precision.

However, there are still areas that require further research and development. Improving algorithms used in hand tracking, expanding gesture databases to include a wide variety of sign languages.

In conclusion, human hand tracking for sign language recognition based on motion trajectory represents a significant advance in assistive technologies. It offers a promising solution for improving accessibility and inclusiveness for individuals with hearing impairments. Thanks to continuous progress and research, this technology has the potential to revolutionize communication, bridge the gap between communities and contribute to a more inclusive and fair society for all.

ACKNOWLEDGMENT

We extend our sincere gratitude to our guide, **Mr. M. M. JADHAV**, for his invaluable contribution to the project. We are also grateful to **Dr. B. H. PATIL**, the Head of the Department, for his continuous support and guidance. We appreciate the support and facilities provided by our principal, **Dr. R. S. BICHKAR**. Additionally, we acknowledge the contributions of the department's faculty and support staff. Their assistance has been instrumental in the project's successful conclusion.

REFERENCES

[1]. N. Adaloglou et al., "A Comprehensive Study on Deep Learning-Based Methods for Sign Language Recognition,", IEEE Transactions on Multimedia, Vol. 24, pp. 1750-1762, 2022, doi: 10.1109/TMM.2021.3070438 \
[2]. Rajesh George Rajan 2.M Judith Leo. "A Comprehensive Analysis on Sign Language Recognition System " In IEEE Conference, 2019
[3]. Joyeeta Singha 2.Karen Das "Hand Gesture Recognition Based on KarhunenLoeve Transform"
[4]. Kalin Stefanov 2.Jonas Beskow "A Real-time Gesture Recognition System for Isolated Swedish Sign Language Signs " In (2016)
[5]. Gerald Krell 2.Bernd Michaelis "Gesture Recognition for Alphabets from Hand Motion Trajectory Using Hidden Markov Models "
[6]. Iker Vazquez Lopez "] Hand gesture recognition or sign language transcription."
[7]. Paulo Trigueiros 2. Fernando Ribeiro "Vision-based Portuguese Sign Language Recognition System"
[8]. Bhumika Nandwana 2. Satyanarayan Tazi 3. Santosh kumar "A Survey Paper on Hand Gesture Recognition "
[9]. Dinh-Son Tran, 2. Hyung-Jeong Yang "Real-Time Hand Gesture Spotting and Recognition Using RGB-D Camera and 3D Convolutional Neural Network (2020)"