International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 5 May 2023, pp: 942-946 www.ijaem.net ISSN: 2395-5252

Gesture Recognition through Machine Learning

Yukta Bisht ,Vineet Kumar, Dr. Atul Kumar, Dr. Vibha Srivastava
Department of Information technology,
Shri Ramswaroop Memorial College of Engineering and Management,Lucknow, Uttar Pradesh

Date of Submission: 15-05-2023 Date of Acceptance: 30-05-2023

ABSTRACT— Recognizing hand gestures is an important way for people to interact with computers, especially in areas like gaming, virtual reality, and robotics. This article provides a thorough review of various methods for identifying hand gestures, including traditional machine learning and newer deep learning techniques. It also discusses the challenges in this field and suggests directions for future research. Gesture is a crucial part of how people communicate, and it can express our intentions. For example, when someone makes a movement or action without really meaning it, we call that an "empty gesture." The researchers behind this project aim to create an affordable device that uses computer vision and gesture recognition to allow people to interact with virtual objects using hand gestures.

I. INTRODUCTION

Hand gesture recognition has become a vital research topic in recent years due to its significant potential applications in HCI. Hand gestures are a natural way for humans to communicate, and recognizing them accurately can improve the user experience in various applications. Moreover, it can enable more intuitive and efficient control of devices.

➢ Traditional Approaches:

The traditional approaches for hand gesture recognition involve extracting features from the input image and using machine learning techniques to classify them. The most common feature extraction methods include histogram of oriented gradients (HOG), local binary patterns (LBP), and scale- invariant feature transform (SIFT). The classifiers used with these features include support vector machines (SVM), k-nearest neighbours (k-NN), and decision trees.

➢ Deep Learning Approaches:

In recent years, deep learning approaches have shown remarkable success in various computer vision tasks, including hand gesture recognition. Convolutional neural networks (CNN) are the most commonly used deep learning architecture for this task. These networks can automatically extract features from raw input images and classify them. Moreover, recurrent neural networks (RNN) can handle temporal information and have been used for recognizing dynamic hand gestures.

➢ Challenges and Future Directions:

Despite the success of hand gesture recognition techniques, several challenges still need to be addressed. These include recognizing gestures in low light conditions, occlusions, and with varying hand shapes and sizes. Additionally, recognizing dynamic gestures accurately remains a challenge, and real-time performance is critical in many applications. Future research directions include developing techniques that can handle these challenges, incorporating multimodal input for more robust recognition, and improving the interpretability of deep learning models.

• STATE-OF-THE-FIELD

The Sign Language Recognition Prototype is a vision-based system designed to recognize the American Sign Language alphabet shown in Figure 1 in real-time. The primary goal of the prototype is to validate the effectiveness of a vision-based system for recognizing sign language and to identify hand features that can be used with machine learning algorithms in real-time sign language recognition systems. The solution uses only one camera and operates under certain assumptions, including that the user is within a defined distance and area in front of the camera, that the user's hand is not obstructed by other objects, and that the system is used indoors due to cameralimitations in bright sunlight.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 5 May 2023, pp: 942-946 www.ijaem.net ISSN: 2395-5252

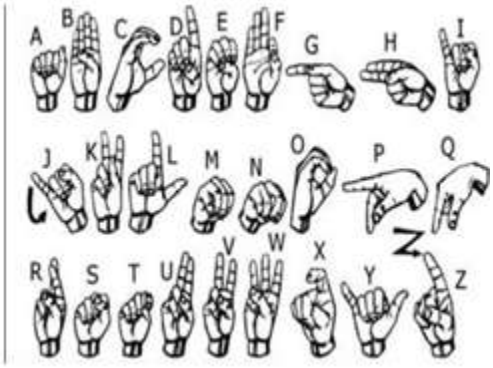

Fig. 1. American Sign Language

The various stages of gesture recognition systemare shown in Fig.2: -

Fig.2 Phases of Gesture Recognition

II. RELATED WORK

Most research on hand gesture recognition has relied on glove-based systems, which attach sensors like potentiometers and accelerometers to each finger. These sensors detect movements and identify the corresponding alphabet. Christopher Lee and Yangsheng Xu developed a system that could recognize 14 hand alphabet letters, learn new gestures, and update each gesture model in real-time. Advanced glove devices have since been designed, such as the Sayre Glove, Dexterous Hand Master, and Power Glove. However, a significant drawback of glove-based systems is the need for recalibration whenever a new user is introduced, as the image processing unit must identify the user's fingertips. Our project, by contrast, uses image processing and has the advantage of being versatile in terms of background color and not requiring color bands. Therefore, security systems could use it to identify individuals based on "who they are" instead of "what they have" or "what they remember". Generally, biometrics can be classified into two main categories.

- **Vital**– The concept is related to the physical characteristics of a person, which encompasses attributes like bodyshape, facial features, palm print, iris, fingerprints, and other physical traits
- **Performance**– Signatures have long been a popular method of identifying individuals based on their behavioral traits, but newer techniques like keystroke dynamics and voice analysis are also gaining traction.

1. Interpreter for communication with deaf andmute individuals- A Review

This article will examine various methods for facilitating communication between deaf-mute individuals and non-deaf-mute individuals. There are two main categories of approaches: wearable communication devices and online learning systems. Wearable communication devices include glove-based systems, keypad methods, and Handicom touch screens. These devices use a variety of sensors, accelerometers, microcontrollers, text-to-speech conversion modules, keypads, and touch screens to interpret and convey messages. Online learning systems, on the other hand, eliminate the need for external communication devices. There are five subcategories of online learning systems: the SLIM module, TESSA, Wi-See technology, SWI_PELE system, and web sign technology. These systems use various methods to interpret sign language and convey messages between deaf-mute andnon-deaf-mute individuals.

2. Hand Gesture Recognition Using PCA in :

The paper presents a gesture recognition system that employs a skin color model approach and thresholding method, along with an efficient template matching algorithm. The system is intended for use in various applications, such as human- robotics and gaming. At first, the system segments the hand area using a skin color model in the YCbCr color space. Then, it applies thresholding to differentiate the foreground and background. Lastly, a template-based matching technique is developed using Principal Component Analysis (PCA) for identification. This system has the potential to accuratelyrecognize hand gestures, which can be usedin diverse applications, including human- robot interaction and sign language recognition.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 5 May 2023, pp: 942-946 www.ijaem.net ISSN: 2395-5252

3. Gesture Recognition System for People with Speech Impairment

We have introduced a system for identifying static hand gestures based on digital image processing. To extract the necessary features from the hand gestures, we applied the Scale-Invariant FeatureTransform (SIFT) algorithm. SIFT is known for its capability to identify and describe image features that are unaffected by changes in scale, rotation, and noise addition. Our system computes SIFT features at the edges of the hand gestures to create a feature vector. This feature vector is then utilized to train a machine learning model that can accurately recognize specific hand gestures.

4. Hand Gesture Recognition for Sign LanguageRecognition: A Review in –

The paper discusses different approaches torecognizing hand gestures and sign language that have been proposed by researchers in the past. For people who are deaf or mute, sign language is the primary means of communication. It enables them to convey their thoughts and emotions through hand movements, facial expressions, and body language. Various techniques, including computer vision, machine learning, and neural networks, have been explored to interpret and analyze sign language and hand gestures. The development of sign language recognition technology has the potential to improve the communication abilities of people with disabilities and facilitate their interactionwith the world around them.

5. The paper describes a system that enables real-time detection and recognition of hand gestures used in ISL and ASL, utilizing the Scale Invariant Feature Transform (SIFT) technique.

The authors proposed an innovative real-time vision-based system for recognizing hand gestures, which can be applied to a variety of human-computer interaction scenarios. The system is capable of detecting 35 distinct hand gestures used in Indian and American Sign Language (ISL and ASL) with high accuracy and at a faster rate. To minimize the risk of false detections, the system employed an RGB- to-GRAY segmentation technique. Furthermore, the authors developed a novelmethod of Scale Invariant Feature Transform (SIFT) that was utilized to extract relevant features. The system was modeled using MATLAB, and to enhance its usability and effectiveness, a graphical user interface (GUI) model was implemented.

III. WORK DONE-

The two figures here show the different hand gestures used.

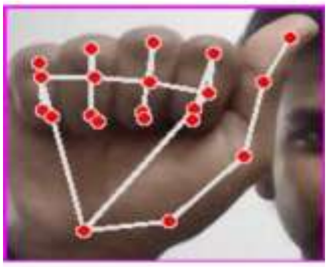

Fig. 3 Hand Gesture for letter “A”

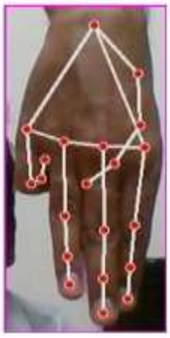

Fig. 4 Hand Gesture for letter “M”

IV. FRAMEWORK

The approach utilized in this system is based on vision. The hand signs are captured directly, without the need for any artificial devices, thereby eliminating thelimitations and complexities that arise fromsuch devices. The system relies on the detection and interpretation of natural hand movements and gestures, which are thentranslated into a form of interaction that facilitates communication between the userand the device.

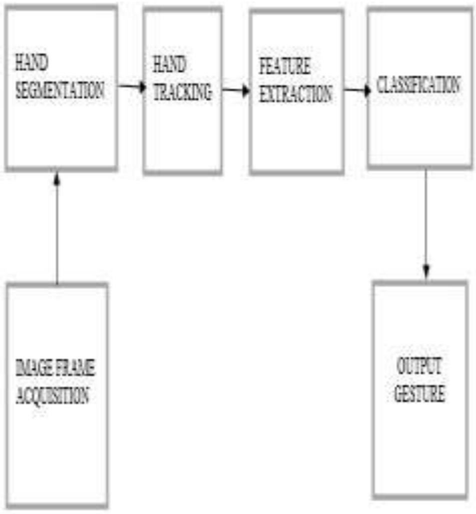

Fig. 5 Flowchart of Hand Gesture

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 5 May 2023, pp: 942-946 www.ijaem.net ISSN: 2395-5252

V. DATA COLLECTION

Creating a comprehensive database of sign language gestures is crucial for the accurate recognition and comparison of captured images during communication. To create our dataset, we followed a step-by-step process using the OpenCV library. Initially, we captured approximately 800 images of each symbol in ASL for training purposes and around 200 images per symbol for testing. To capture each frame, we utilized the webcam of our machine and defined a region of interest (ROI) denoted by a blue square as shown in the image. We then extracted the ROI from the entire image and converted it to a grayscale format. Finally, we applied a Gaussian blur filter to the image to enhance the extraction of various features.

GESTURE CLASSIFICATION

Incorporating data augmentation techniques to increase the size and diversity of the training data for the CNN model. This can improve the robustness of the model and its ability to generalize to new data. Using a more advanced feature extraction technique such as deep learning-based feature extraction or handcrafted feature extraction based on domain knowledge. Incorporating a language model to improve the prediction of the final word by taking into account the context and the probability of different word combinations. Using a more sophisticated approach to detect spaces between words, such as using a sequence labeling model or an attention-based model. Considering the use of other modalities such as audio or motion data to improve the prediction accuracy. Adding a user feedback loop to improve the system's ability to adapt to individual user behavior and preferences. Evaluating the system's performance on a larger and more diverse dataset to assess its real-world applicability and to identify potential limitations and areas forimprovement.

TRAINING AND TESTING

The first step in image pre-processing is to convert the RGB color images into grayscale. This is done to reduce thecomputational complexity and to eliminate the color information that may not be relevant for the task at hand. After converting the images to grayscale, we apply a Gaussian blur filter to smooth out any noise or irregularities in the image. This helps in improving the accuracy of the subsequent image processing operations. Next, we apply an adaptive thresholding algorithm to extract the hand region from the background. The adaptive thresholding algorithm takes into account the local variations in intensity across the image and adjusts the threshold accordingly. This helps in accurately segmenting the hand region from the background, even in cases where the lighting conditions or the background texture may vary. Once the hand region is extracted, we resize the image to a standard size of 128 x 128 pixels. This helps in reducing the variability in image size and aspect ratio, which can affect the performance of the subsequent machine learning models. Finally, we feed the pre-processed images as input to our machine learning model for training and testing.

VI. CONCLUSION

This report describes the development of a real-time vision-based American Sign Language (ASL) recognition system for alphabets, with the aim of assisting Deaf and Dumb individuals in communication through sign language. This allowed us to detect almost all the symbols provided they were shown properly, with no background noise and adequate lighting. The system was developed using various computer vision techniques, including image pre-processing, feature extraction, and classification. To pre-process the images, we converted RGB images to grayscale and applied a Gaussian blur filter to remove noise.The pre- processed images were fed into a CNN model for training and testing. The system demonstrated the potential of computer vision and machine learning in developing assistive technologies for individuals with disabilities. Further improvements could be made by incorporating more complex models, expanding the dataset to include more variations in hand gestures.

REFERENCES: -

[1]. Ren and Wu (2018) conducted a comprehensive evaluation of hand gesture recognition techniques
[2]. Murthy, Kandaswamy, and Bhowmick (2019) also reviewed various techniques, challenges, and applications related to hand gesturerecognition
[3]. Fathi and Mori (2008) proposed a technique for recognizing complex events using joint low-level features.
[4]. Cao and Kit (2017) developed atechnique for hand gesture recognition using depth-based hand skeletal points and discussed its applications.
[5]. Cai, Zhang, Liu, and Wang (2018) surveyed deep learning-based hand gesture recognition techniques.
[6]. P. Garg, N. Aggarwal and S. Sofat. (2009), "Vision Based Hand Gesture

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 5 May 2023, pp: 942-946 www.ijaem.net ISSN: 2395-5252

Recognition," World Academy of Science, Engineering andTechnology, Vol. 49
[7]. Mokhar M. Hasan, Pramod K. Mishra, (2012) "Features Pitting using Multivariate Gaussian Distribution for Hand Gesture Recognition", International Journal of Computer Science & Emerging Technologies JCSET.
[8]. W. T. Freeman and Michal R. (1995) "Orientation Histograms for Hand Gesture Recognition" IEEE International Workshop onAutomatic Face and Gesture Recognition.