

A Gesture-based Tool for Sterile Browsing of Radiology Images

INTRODUCTION

Sign language is the basic communication method for those who suffer from hearing impairment. The primary component of a sign language is hand gestures. Gesturing is an instinctive way of communication to present a specific meaning. Sign language is the visual manual modality to convey meaning which is quite similar to the hand gestures .

Language is expressed via the manual sign-stream in combination with non-manual elements. Sign languages are full-fledged natural languages with their own grammar and lexicon. In India over 21 million of people are suffering some form of disability out of that about 7.5% people are dumb and 5.8% people are both dumb and deaf . Gesture involves the hands, head or face to communicate and sign involves only the hands to communicate.

Humans express the idea or meaning by using the sign or gesture in order to convey the message to the mute people. A computer can detect and reconstruct the message addressed by the human gestures and translate approximately. Regionally different sign languages are used in different part of the world such as ASL (American Sign Language), CSL (Chinese Sign Language) and ISL (Indian Sign Language) etc, .

They are two types of sign language recognition

- i) Alternate Sign Language: Hand signs developed by speakers for a conversation in a specific context used have the standard signs and
- ii) Primary Sign Language: it is used by the group of people to communicate with each other who belong to the same family or locality.

The decomposition of the signs can be simplified into single handed or double handed signs .

Researchers have recently focused their attention on vision-based hand gesture recognition. However, due to several constraints, achieving an effective vision-driven hand gesture recognition system in real time has

remained a challenge. This paper aims to uncover the limitations faced in image acquisition through the use of cameras, image segmentation and tracking, feature extraction, and gesture classification stages of vision-driven hand gesture recognition in various camera orientations. This paper looked at research on vision-based hand gesture recognition systems from 2012 to 2022. Its goal is to find areas that are getting better and those that need more work. We used specific keywords to find 108 articles in well-known online databases. In this article, we put together a collection of the most notable research works related to gesture recognition. We suggest different categories for gesture recognition-related research with subcategories to create a valuable resource in this domain. We summarize and analyse the methodologies in tabular form. After comparing similar types of methodologies in the gesture recognition field, we have drawn conclusions based on our findings. Our research also looked at how well the vision-based system recognized hand gestures in terms of recognition accuracy. There is a wide variation in identification accuracy, from 68% to 97%, with the average being 86.6 percent. The limitations considered comprise multiple text and interpretations of gestures and complex non-rigid hand characteristics. In comparison to current research, this paper is unique in that it discusses all types of gesture recognition techniques.

Hand gesture recognition plays a significant part in delivering diverse messages using hand gestures in the digital domain. Real-time hand gesture identification is now possible because of advancements in both imaging technology and image processing algorithmic frameworks. This has enabled natural interactivity previously unattainable by the use of the two-dimensional mouse. Due to the real-time nature of gesture recognition, it should be accomplished without overburdening the computing element. Moreover, image processing plays a critical role in segmentation, feature extraction of hand gestures and ultimate recognition of the gestures. Numerous computer vision algorithmic frameworks based on image processing concepts have been developed and are being improved.

Hand motions may vary from static to dynamic, depending on their use. Hand gesture recognition technologies each have their own set of benefits and drawbacks, which are dependent on the platforms on which they are implemented. Due to numerous difficulties encountered during foreground separation from the background, there are many current obstacles to achieving realistic and effective real-time hand gesture recognition. The hand that needs to be identified is represented by the foreground. Changing picture

luminance, such as pixel colour of the hand skin and background in vision-based systems, as well as cumbersome, expensive gear in glove-enabled and depth-enabled systems, are the most common problems

1.Sahoo, J.P.; Prakash, A.J.; Pławiak, P.; Samantray, S. Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network. Sensors 2022.

This study focuses on a literature review of hand gesture strategies and discusses their pros and limits in various situations. In addition, the performance of these methods is tabulated, with an emphasis on computer vision techniques that deal with similarity and difference points; hand segmentation techniques algorithms and limitations; number and types of gestures; dataset used detection range (distance); and camera type.

2.Gadepalli, T.R.; Srivastava, G.; Liyanage, M.; Aparejo, M.; Chowdhary, C.L.; Koppa, S.; Maddi Kunta, P.K.R. Hand gesture recognition based on a Harris hawks optimized convolution neural network. Compute. Electra. Eng. 2022.

Convolutional neural networks (CNN) are used to categorize images of hand gestures. A newly developed metaheuristic technique, the Harris hawk's optimization (HHO) algorithm, is utilized to optimize the CNN's hyperparameters. Their extensive comparison research shows that the proposed HHO-CNN hybrid model outperforms current models by achieving 100 percent accuracy.

3.Amin, M.S.; Rizvi, S.T.H. Sign Gesture Classification and Recognition Using Machine Learning. Cybernet. Syst. 2022.

This article examines the flex, accelerometer, and gyroscope-based smart prototype developed to recognize sign language motions. These sensors are

put on a glove in order to record and assemble alphabetic (i.e., 0–10, A–Z) and numeric (i.e., 0–10, and A–Z datasets). The primary purpose of the proposed model is to categorize sign gestures produced by deaf–mute people and identify the true meaning of movements performed.

4.Kong, F.; Deng, J.; Fan, Z. Gesture recognition system based on ultrasonic FMCW and Comvest model. Measurement 2022.

Based on the ultrasonic frequency modulated continuous wave (FMCW) and the Conv LSTM model, a system for gesture identification was suggested in this work. It uses a hardware configuration consisting of one transmitter and three spatially separated receivers.

5.Sabo, S.; Singha, J.; Laskar, R.H. Dynamic hand gesture recognition using combination of two-level tracker and trajectory-guided features. Multimer. Syst. 2022.

In this research, the authors proposed a hand gesture detection system for a dataset of lowercase numbers and alphabets. The suggested method recognizes the hand using information on skin colour and mobility. Hand tracking is performed using a two-level tracking system and a modified Kanade–Lucas–Tomasi (KLT) tracking algorithm.

6.Oudah, M.; Al-Naji, A.; Chahl, J. Computer Vision for Elderly Care Based on Hand Gestures. Computers 2021.

Deaf–mute elderly folk use five distinct hand signals to seek a particular item, such as drink, food, toilet, assistance, and medication. Since older individuals cannot do anything independently, their requests were delivered to their smartphone.

Challenges

Microsoft Kinect v2 sensor's capability to extract hand movements in real time keeps this study in a restricted area.

Gesture-based technology may assist the handicapped, as well as the general public, to maintain their safety and requirements. Due to the significant changeability of the properties of each motion with regard to various persons, gesture detection from video streams is a complicated matter.

7.Mujahid, A.; Awan, M.J.; Yasin, A.; Mohammed, M.A.; Damaševičius, R.; Maskeliunas, R.; Abdulkareem, K.H. Real-Time Hand - Gesture Recognition Based on Deep Learning YOLOv3 Model. Appl. Sci. 2021.

This article offers a lightweight model based on the YOLO (You Look Only Once) v3 and the DarkNet-53 neural networks for gesture detection without further preprocessing, filtration of pictures and image improvement. Even in a complicated context the suggested model was very accurate, and even in low resolution image mode motions were effectively identified. Rate of high frame.

Challenges

The primary challenge of this application for identification of gestures in real time is the classification and recognition of gestures. Hand recognition is a method used by several algorithms and ideas of diverse approaches for understanding the movement of a hand, such as picture and neural networks.

8.Bhuiyan, M.R.; Abdullah, D.; Hashim, D.; Farid, F.; Uddin, D.; Abdullah, N.; Samsudin, D. Crowd density estimation using deep learning for Hajj pilgrimage video analytics. F1000Research 2021.

The purpose of this study is to offer a method for Hajj applications that is based on a convolutional neural network model. They also created a technique for counting and then assessing crowd density. The model employs an

architecture that recognizes each individual in the crowd, marks their head position with a bounding box, and counts them in their own unique dataset (HAJJ-Crowd).

Challenges

There has been a growth in interest in the improvement of video analytics and visual monitoring to better the safety and security of pilgrims while in Makkah. It is mostly due to the fact that Hajj is a one-of-a-kind event with hundreds of thousands of people crowded into a small area.

9.Bhuiyan, M.R.; Abdullah, J.; Hashim, N.; Al Farid, F.; Samsudin, M.A.; Abdullah, N.; Uddin, J. Hajj pilgrimage video analytics using CNN. Bull. Electr. Eng. Inform. 2021.

This study presents crowd density analysis using machine learning. The primary goal of this model is to find the best machine learning method for crowd density categorization with the greatest performance

Challenges

Crowd control is essential for ensuring crowd safety. Crowd monitoring is an efficient method of observing, controlling, and comprehending crowd behaviour

10.Zamri, M.N.H.B.; Abdullah, J.; Bhuiyan, R.; Hashim, N.; Farid, F.A.; Uddin, J.; Husen, M.N.; Abdullah, N. A Comparison of ML and DL Approaches for Crowd Analysis on the Hajj Pilgrimage. In Proceedings of the International Visual Informatics Conference; Springer: Berlin/Heidelberg, Germany, 2021.

Data augmentation is a technique of expanding the data set by producing various picture shapes to increase model performance

It also helps to mitigate the over-fitting issue in the model during the training stage. The overcast issue arises when random noise or mistakes occur instead of when the underlying connection is there. Using an increase in data, additional images were produced for the model from each picture because some irrelevant patterns may occur throughout the model training process. Several methods were employed for data augmentation operations: rotational changes, vertical and horizontal rotations, and intensity disorder, including light disturbances.

A classical ANN involves a local minimal issue, which typically ends with a local optimization process rather than a globally optimal state. More overfitting issues often complicates general machine learning models. Intensive network structure optimization may address the issues of the local minima and overriding by DNNs [80,81]. Deep learning is a machine learning-based approach that educates computers to accomplish tasks similar to those performed by humans.

11.Bari, B.S.; Islam, M.N.; Rashid, M.; Hasan, M.J.; Razman, M.A.M.; Musa, R.M.; Ab Nasir, A.F.; Majeed, A.P.A. A real-time approach of diagnosing rice leaf disease using deep learning-based faster R-CNN framework. Peerj Comput. Sci. 2021.

1A classical ANN involves a local minimal issue, which typically ends with a local optimization process rather than a globally optimal state. More overfitting issues often complicates general machine learning models. Intensive network structure optimization may address the issues of the local minima and overriding by DNNs .

Deep learning is a machine learning-based approach that educates computers to accomplish tasks similar to those performed by humans. For example, deep learning is the underlying technology that enables driverless automobiles to detect traffic lights and people. It is also the underlying principle of audio and speech recognition in a variety of devices, such as mobile phones and tablets.

Deep learning is gaining popularity because it is capable of performing tasks that were previously impossible.

12.Alonso, D.G.; Teyseyre, A.; Soria, A.; Berdun, L. Hand gesture recognition in real world scenarios using approximate string matching. Multimer. Tools Appl. 2020.

investigated dynamic hand movements using spatial–temporal algorithm convolutional networks. Three types of graph edges associated with the activity of hand joints were proposed in a skeleton-based model. A deep neural network was utilized to pick semantic characteristics in order to provide an accurate output.

13.Zhang, T.; Lin, H.; Ju, Z.; Yang, C. Hand Gesture recognition in complex background based on convolutional pose machine and fuzzy Gaussian mixture models. Int. J. Fuzzy Syst. 2020

created a string-matching technique for understanding hand motions in real-time situations. The k-means technique was used to create an approximation string matching to capture the features of hand joints. It was done to enhance the precision of various motions by specifying the number of clusters.

14.Köpüklü, O.; Gunduz, A.; Kose, N.; Rigoll, G. Online dynamic hand gesture recognition including efficiency analysis. IEEE Trans. Biom. Behav. Identity Sci. 2020

The gesture matching was performed by analysing the fused gesture dataset, where the gesture frames were categorized. The SFM was then utilized to

accelerate the analysis processing. To improve the efficiency of gesture analysis

15.Alnaim, N. Hand Gesture Recognition Using Deep Learning Neural Networks. Ph.D. Thesis, Brunel University, London, UK, 2020.

In this study, image processing techniques such as wavelets and empirical mode decomposition were suggested to extract picture functionalities in order to identify 2D or 3D manual motions. Classification of artificial neural networks (ANN), which was utilized for the training and classification of data in addition to the CNN (CNN).

Challenges

Three-dimensional gesture disparities were measured utilizing the left and right 3D gesture videos

16.Min, Y.; Zhang, Y.; Chai, X.; Chen, X. An efficient pointlet for point clouds-based gesture recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020

This work formulates the recognition of gestures as an irregular issue of sequence identification and aims to capture long-run spatial correlations in points of the cloud. In order to spread information from past to future while maintaining its spatial structure, a new and effective Pointlet is suggested.

Challenges

The underlying geometric structure and distance information for the object surfaces are accurately described in dot clouds as compared with RGB data, which offer additional indicators of gesture identification

17.Al-Hammadi, M.; Muhammad, G.; Abdul, W.; Alsulaiman, M.; Bencherif, M.A.; Alrayes, T.S.; Mathkour, H.; Mekhtiche, M.A. Deep learning-based approach for sign language gesture recognition with efficient hand gesture representation. IEEE Access 2020

A new system is presented for a dynamic recognition of hand gestures utilizing various architectures to learn how to partition hands, local and global features and globalization and recognition features of the sequence

Challenges

To create an efficient system for recognition, hand segmentation, local representation of hand forms, global corporate configuration, and gesture sequence modelling need to be addressed.

18.Neethu, P.; Suguna, R.; Sathish, D. An efficient method for human hand gesture detection and recognition using deep learning convolutional neural networks. Soft Comput. 2020

This article detects and recognizes the gestures of the human hand using the method to classification for neural networks (CNN). This process flow includes hand area segmentation using mask image, finger segmentation, segmented finger image normalization and CNN classification finger identification

Challenges

SVM and the naive Bayes classification were used to recognize the conventional gesture technique and needed a large amount of data for the identification of gesture patterns.

19.Zoph, B.; Cubuk, E.D.; Ghiasi, G.; Lin, T.Y.; Shlens, J.; Le, Q.V. Learning data augmentation strategies for object detection. In Proceedings of the European Conference on Computer Vision, Glasgow, UK, 23–28 August 2020

It also helps to mitigate the over-fitting issue in the model during the training stage. The overcast issue arises when random noise or mistakes occur instead of when the underlying connection is there. Using an increase in data, additional images were produced for the model from each picture because some irrelevant patterns may occur throughout the model training process. Several methods were employed for data augmentation operations: rotational changes, vertical and horizontal rotations, and intensity disorder, including light disturbances

20.Mungra, D.; Agrawal, A.; Sharma, P.; Tanwar, S.; Obaidat, M.S. PRATIT: A CNN-based emotion recognition system using histogram equalization and data augmentation. Multimed. Tools Appl. 2020

Convolutional neural networks (CNN) are used to categorize images of hand gestures. A newly developed metaheuristic technique, the Harris hawk's optimization (HHO) algorithm, is utilized to optimize the CNN's hyperparameters. Their extensive comparison research shows that the proposed HHO-CNN hybrid model outperforms current models by achieving 100 percent accuracy

21.Rashid, M.; Sulaiman, N.; PP Abdul Majeed, A.; Musa, R.M.; Bari, B.S.; Khatun, S. Current status, challenges, and possible solutions of

EEG-based brain-computer interface: A comprehensive review.
Front. Neurorobotics 2020

A classical ANN involves a local minimal issue, which typically ends with a local optimization process rather than a globally optimal state. More overfitting issues often complicates general machine learning models. Intensive network structure optimization may address the issues of the local minima and overriding by DNNs

22.Mathew, A.; Amudha, P.; Sivakumari, S. Deep Learning Techniques: An Overview. In Proceedings of the International Conference on Advanced Machine Learning Technologies and Applications, Manipal, India, 13–15 February 2020;

Deep learning is a machine learning-based approach that educates computers to accomplish tasks similar to those performed by humans. For example, deep learning is the underlying technology that enables driverless automobiles to detect traffic lights and people. It is also the underlying principle of audio and speech recognition in a variety of devices, such as mobile phones and tablets. Deep learning is gaining popularity because it is capable of performing tasks that were previously impossible. A deep learning model is constructed by layering data, which may be images, text, or audio, into distinct and discrete categorization layers.