International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Survey on Hand Gesture Recognition Using Deep Learning

Anusha S[1], Basavaraj[2], Divya Sundar S[3], C R Sudhanva[4], G N Samrudda[5]

[1]*Assistant Professor, Department of Computer Science and Engineering, MVJ College of Engineering, Bangalore, Karnataka, India.*
[2,3,4]*Undergraduate Scholar, Department of Computer Science and Engineering, MVJ College of Engineering, Bangalore, Karnataka, India.*
[1]*anusha@mvjce.edu.in*, [2]*salibasava84310@gmail.com*, [3]*mailtosundar01@gmail.com*, [4]*crs.holla@gmail.com*, [5]*samruddagn@gmail.com*.

A B S T R A C T

Nowadays technology is developing at a great pace where humans are getting closer and closer to interacting with computers every day. Human Computer Interaction will reach a stage where the users will have to instruct the system in a more dynamic way than the current mouse and keyboard methods. Computer programs like Gesture Recognition can greatly support such kind of dynamic control with the help of Deep Learning and various recent trending technologies.

Keywords: Hand Gesture recognition, Hand Gesture Using Deep Learning, Image Identification, Image Filtering using CNN.

1. Introduction

Gesture Recognition has been a growing field of research widely for more than two decades in human computer interaction. Recent developments in computer software and related hardware technology have provided a value-added service to the users. In everyday life, Physical gestures are a powerful means of communication. They can economically be able to convey a rich set of facts and feelings through just simple actions. The main goal of Human Computer Interaction is to improve the user experience and the communication between the Human and the Computer. There have been many great strides that has brought users much closer to interacting with the system, in those gesture recognition is about to one of them. For example, showing one's thumbs up to another person can represent anything from a "All the best" to "Good luck". Being able to utilize the full potential of physical gesture is also something that most human computer dialogues lack.

The implementation of hand gesture recognition is one of the most significant and fundamental problems in computer vision. With recent advances in information technology and media, automated human interaction systems are being created that include hand processing tasks such as hand detection, hand recognition, and hand movement tracking. Gesture recognition technology is a vision-based system that uses a camera and a motion sensor to track the user's movements and recognize movements in real time.

The applications of gesture recognition are vast and span across various industries. In gaming, it enables users to control virtual characters or environments using hand gestures, creating immersive and interactive experiences. In healthcare, it can be used for rehabilitation exercises, remote patient monitoring, and surgical procedures. In the automotive industry, it can enhance driver safety by enabling drivers to control infotainment systems or settings with hand gestures, reducing distractions while driving. Newer and more advanced camera systems and programs also allow or have the ability to track depth data, which can greatly help improve gesture tracking. Using real-time image processing, users can instantly interact with the application program and achieve the desired results. As technology continues to advance, gesture recognition has the potential to transform how humans interact with machines, making interactions more intuitive and seamless. With ongoing developments in sensor technology, machine learning, and applications, gesture recognition is poised to play a significant role in shaping the future of human-computer interaction.

2. Literature Review

2.1 Hand Gesture Recognition Using Deep Learning

The paper "Hand Gesture Recognition Using Deep Learning" presents an overview of the use of deep learning techniques for hand gesture recognition. The authors highlight the increasing demand for natural and intuitive human-computer interaction, and the potential of hand gesture recognition for various applications, such as gaming, virtual reality, and smart homes. The paper discusses different deep learning approaches, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), used for hand gesture recognition.

The authors review various datasets and pre-processing techniques used in hand gesture recognition research, and discuss the challenges associated with hand gesture recognition, such as variations in hand shapes, lighting conditions, and occlusions. The paper presents a comprehensive review of recent studies and experimental results in the field, including accuracy, speed, and robustness of different deep learning models. Finally, the authors discuss potential future research directions and applications of hand gesture recognition using deep learning, highlighting the importance of continued advancements in sensor technology, data collection, and algorithm development for improving the accuracy and practicality of hand gesture recognition systems.

2. A Survey Paper on Hand Gesture Recognition

Gesture recognition is a technique used to identify human gestures and use them to transfer information. It consists of body and hand gestures, divided into two categories: static and dynamic. To recognize static gestures, a classifier or template matches is used, while a dynamic movement is used to determine the meaning of hand gestures. Dynamic gesture recognition is the recognition of hand movements by four features: velocity, shape, location and orientation. It has many uses in medical applications, entertainment applications, automated systems, and better life for the disabled people.

Guillaume Plouffe and ana-maria cretu proposed a framework for hand gesture recognition, and a DTW algorithm is used for identifying gesture identification. Rafiqul zaman khan, Rios-Sorial, Ram Rajesh J, Hamaid A. Jalab, Siddharth S. Rataray, Etsuko Ueda, and Chan Wah Ng proposed hand gesture recognition systems, using skin colour filtering, edge detection and convex hull methods to detect hand gestures. Siddharth's system uses C++ and openCV, while Chan Wah Ng's system uses image furrier information as the primary characteristic and is categorized with the support of RBF network. This paper discusses the importance of human computer interaction in hand gesture recognition, the past of hand gesture recognition and technical problems, the use of vision based, glove-based and depth based techniques, and the launch of kinect data for finger identification and hand gesture identification. It also reviews 26 openly available hand gesture databases and provides links for downloading them.

2.3 A Review of the Hand Gesture Recognition System: Current Progress and Future Directions

Developing hand gesture recognition systems such as sign language applications is extremely important to overcome the communication barrier with people who are unfamiliar with sign language. For hand gesture recognition systems, input that may not be considered during the development stage includes environmental noise, signers` variability, language variability, and so on. Due to this complexity, it appears that many researchers have placed less emphasis on continuous sign language for vision-based hand gesture recognition, which has limited practicality for real-world applications. The other challenge in this field is the development of robust signer-independent hand gesture recognition systems.

Since the vision-based hand gesture recognition system is practical for real-life applications, it must apply to any user in any environment. However, there had been no review that examined the extent of research made towards the development of the vision-based hand gesture recognition system and the possible future directions. As such, the current paper addresses that gap by reviewing current and past literature to examine the progress made thus far by vision-based hand gesture recognition systems.

This ease of use of the vision-based system was challenged by how it handled datasets made up of dynamic hand gestures in sign language, such as isolated and continuous signs. The authors reviewed the state-of-the-art technique used in recent hand gesture and sign language recognition research in areas such as data acquisition, pre-processing, segmentation, feature extraction, and classification. It appears that the past works reviewed by researchers left a gap; they had not examined the challenges and future direction of the vision-based hand gesture recognition system. Based on this, the current paper will address this gap by reviewing existing literature to identify the progress of research in vision-based hand gesture recognition systems for the present and for future directions.

2.4 Home Outlet and LED Array Lamp Controlled by a Smartphone with a Hand Gesture Recognition

This design is composed of two parts: a smartphone's application and a wireless remote-control unit. The application reads the data from the built-in accelerometer and gyroscope and provides the required hand gesture recognition. The remote control unit is composed of a NFC coil, a microcontroller unit (MCU) and a Bluetooth transmission module. The application will detect a user's hand gesture such as a rotating or a flipping movement that will dim the LED array lamp or switch the power on/off. There are two types of dimming modes: one is the white light LED dimmer; the other one is the colourful LED light adjustment mode.

The smartphone will start the extraction of the rotational angle information, by using the gyroscope to catch the rotation angle to decide whether to turn off or to dim the LED light. If the rotation angle is equal or less than 0 degrees, the power will turn on. If the angle is large or equal to 180 degrees, the power automatically turns off. Otherwise, the LED will adjust the brightness to adjust the brightness. This design includes the control mechanism for a power line outlet, which can only control whether the power line outlet is on or off. The system provides a convenient and intuitive way for users to control home appliances using hand gestures and a smartphone, offering potential applications in smart homes and Internet of Things (IoT) environments. The paper also highlights future research directions and improvements for the proposed system.

2.5 Automated Hand Gesture Recognition using a Deep Convolutional Neural Network model.

Gesture recognition is a computer science technology that helps users interact with digital devices using simple and natural body gestures. Hand gesture recognition is a part of gesture recognition that is based on recognizing the movements of hands meant to be delivered. Deep learning is a type of artificial neural network that can deduce or distinguish patterns and understand them. CNN is a technique that can be applied to images to make them interpretable by machines. This paper aims to train an algorithm to classify images of various hand gestures and signs using a convolutional neural network with Keras and TensorFlow. It uses data captured using the OpenCV library to improve the accuracy score of existing hand gesture recognition techniques. A real-time anti-encroaching hand gesture recognition and hand tracking mechanism has been proposed to improve human-computer interactions and bring ease to those who rely on gestures for their day-to-day communication.

CNN is a type of neural network that is empowered with certain layers, such as an input layer, a hidden layer, a convolutional layer, a max pooling layer, a fully connected layer, and an output layer. A convolutional layer is the first hidden layer and is used to detect and extract features from an image. Edge image is useful for the initial steps and layers of CNN, as it is the first set of primitive feature sets for working in a hierarchy of all the features. Convert an image into a matrix of binary integer values, such as 0s and 1s. The proposed system achieves high accuracy in hand gesture recognition, with results surpassing previous approaches. The paper also presents an evaluation of the system's performance, including metrics such as accuracy, precision, recall, and F1-score. Additionally, the paper discusses the potential applications of automated hand gesture recognition, such as human-computer interaction, virtual reality, robotics, and healthcare.

The system's ability to recognize hand gestures automatically without the need for markers or sensors makes it versatile and applicable in various scenarios. The paper also discusses the challenges and limitations of the proposed system, such as lighting conditions, hand pose variations, and dataset size. The authors propose future research directions, including exploring different CNN architectures, improving real-time performance, and extending the system to recognize dynamic gestures. Overall, the paper presents a novel approach for automated hand gesture recognition using a deep CNN model, with potential implications for various fields and opportunities for further advancements in the field.

2.6 Home Appliance Control by a Hand Gesture Recognition Belt in LED Array Lamp Case.

This paper describes a wearable device composed of a three-axis accelerometer, a gyroscope, a micro-control unit (MCU) and Bluetooth communication modules. The accelerometer and gyroscope combination are on a module and continually send out data. The MCU obtains both the data and the human hand shaking acceleration signal by using a Kalman filter. The control command, by means of Bluetooth communication modules transmitted, controls the home appliance on/off function. The WAGC is composed of two parts: the hand gesture recognition belt (HGRB) and the receiver unit.

The HGRB is a master device which can recognize the user's hand gestures and decode them, while the receiver unit needs an additional MCU, Bluetooth module and some solid-state relays (SSR). The HGRB is tied to the hands and when the system is reading the accelerometer and gyroscope value, any hand movement becomes a very obvious noise. To improve the recognition rate, the system adds a Kalman filter and an iteration algorithm to calculate and suppress the white noise. When the user wears the HGRB and is ready to detect hand gestures, the sensors will read out a rotation signal and the system wakes up HGRB by clicking the surface of the device. The number of click times decides the detection actions. Procedure A detects the rotation angle and the direction of motion, while Procedure B only detects the palm rotation angle. The system will determine the angle of palm rotation and transform those actions to fine tune the brightness instruction.

The paper presents the design, implementation, and evaluation of the system, including the accuracy and response time of the hand gesture recognition belt. The system offers a convenient and intuitive way for users to control home appliances without the need for physical switches or buttons. The paper also discusses potential applications of the system in smart homes and IoT environments, as well as limitations and future research directions. Overall, the paper presents a unique approach for home appliance control using a hand gesture recognition belt, offering potential benefits in terms of convenience, accessibility, and user experience.

2.7 A Sliding Window Approach to Natural Hand Gesture Recognition using a Custom Data Glove

This paper presents a gesture recognition system built for recognizing 31 natural and interaction-oriented hand gestures. It achieves an accuracy of 98.5% with linear discriminant analysis for dimensionality reduction and logistic regression for classification, and accuracy does not significantly suffer when the computationally expectable FFT features are removed. The system involves capturing hand motion data from the custom data glove, which is equipped with sensors to detect hand gestures in real-time. The captured data is then processed using a sliding window approach to segment hand gestures into smaller windows for analysis. Machine learning algorithms are applied to these segmented windows to recognize hand gestures. The paper presents the design and implementation of the system, including the data glove design, sensor configuration, and feature extraction techniques. The effectiveness of the proposed approach is evaluated through experiments and performance metrics such as accuracy, precision, recall, and F1-score.

This result is relevant for gesture-based interfaces, as it means that continuous gesture detection based on continuous sensing is accurate enough and can be implemented in a computationally efficient manner. Future work on computing efficiency will include an exploration of which sensors are irrelevant and could be completely removed without degrading the recognition accuracy. It will also be interesting to design interactions with computer systems using such gestures that can be recognized well. Overall, the paper presents a novel approach to natural hand gesture recognition using a custom data glove and a sliding window approach, with potential implications in various domains and opportunities for further advancements.

2.8 3-D Hand Motion Tracking and Gesture Recognition Using a Data Glove.

This paper presents a data glove that is composed of three tri-axis accelerometers, one controller, and one Bluetooth, and implements a 3-D digital hand model for hand motion and gesture recognition using motion sensors such as accelerometers, gyroscopes, bend sensors, force sensors, and so on. It is more suitable than video and motion sensor-based recognition in the complex environment, as it is attached to a user, provides more coverage, and can be acquired wirelessly. The KHU-I data glove is composed of three tri-axis accelerometer (Free Scale Inc.) as the sensor of hand motion, an ATmegaI28 (AVR) as the controller, and a FB755 Bluetooth as the wireless communication device. The output of the accelerometer is an analog signal that has a range of 3.3V. The signal is transmitted to a computer through Bluetooth wireless communication, which supports the 9600 bit per second baud rate and is possible for 1:8 multiple communication.

The sensor and controller use rated voltage sources of 5V and 3.3V, respectively. Data acquisition and display are according to the following procedures: i) acquisition of data of hand motion via accelerometers, ii) transmission of data to a PC, iii) evaluation of data received to determine whether an error occurred or not through the received data length, iv) data acquisition and display under two conditions: (Condition I) acquisition in the case of successful transmission, data display; (Condition II) acquisition failure, data display. This paper presents a 3-D digital hand model consisting of 15 ellipsoids that represent the volume of each knuckle, 15 joints, and skeletons. The hand model is implemented by connecting each ellipsoid to a joint, which has a limited angle and is OOF as a hand. The KHU-l data glove is capable of capturing hand motion via tri-axis accelerometer sensors and wireless communication between the data glove and a PC. Future work requires faster computation times to reduce the time delay of the system and advanced recognition methods to recognize more complex gestures.

2.9 A Method for Hand Gesture Recognition .

Gestures are used for verbal or non-verbal communication and can be made with the movements of fingers, hands, arms, heads, faces, or bodies. Hand gestures are the most expressive and used more frequently, and the use of hand gestures as a natural interface serves as a motivating factor for research in hand gesture recognition. Microsoft launched a 3D depth-sensing camera known as Kinect in 2010, which can be used to capture RGB images and depth images. In 2011, Meenakshi et al. first captured the RGB images and converted them into YCbCr.

In 2008, Fong et al. proposed a method for analyzing features extracted from an image: the location of fingers. This work used Microsoft Kinect to capture images for different hand gestures using one, two, three, four, and five fingers. This paper takes hand images with Microsoft Kinect and extracts the image features on the basis of contour area and convexity defects. Kinect is a motion-sensing input device with an IR projector and two cameras, a multi-array microphone, and a motorized pivot. It produces 640 x 480 RGB and depth images at 30 frames per second. Preprocessing is needed to prepare the image for input into our method. We use the nearest neighbor interpolation algorithm to interpolate pixels and get a depth array that has meaningful values in all the pixels. We then use a median filter with 55 windows on the depth array to make the data smooth. Finally, we calculate the convex hull of a set of points, which we use to find contours.

2.10 Hand Gesture Recognition for Human Computer Interaction.

The paper "Hand Gesture Recognition for Human Computer Interaction" discusses the significance of hand gesture recognition as a means of interaction between humans and computers. They highlight the importance of natural and intuitive interfaces for human-computer interaction, and how hand gestures can provide a seamless and convenient way for users to interact with computers. They have created a gesture recognition system that does not utilize markers, making it more user-friendly and low-cost. Various machine learning algorithms used for hand gesture recognition are also discussed, including deep learning models such as convolutional neural networks (CNNs).

The paper presents experimental results to evaluate the performance of the hand gesture recognition system. The advantages of hand gesture recognition, such as its potential for accessibility and inclusivity for users with disabilities, are highlighted. They also discuss the limitations and future research directions in hand gesture recognition, including improving accuracy, robustness to variations in hand gestures, and real-time implementation. The authors compare the performance of different machine learning algorithms and discuss the impact of different factors on the accuracy of the system. The paper concludes by summarizing the importance of hand gesture recognition in human-computer interaction and its potential for shaping the future of interactive technologies.

3. Conclusion

Gesture recognition has emerged as an important and growing field of research in human-computer interaction, and recent developments in computer software and hardware technology have made it possible to provide value-added services to users. Physical gestures are a powerful means of communication in everyday life, and the goal of human-computer interaction is to improve the user experience and communication between humans and computers. Gesture recognition technology has the potential to unlock the full potential of physical gestures in human-computer dialogues, and it is one of the most significant and fundamental problems in computer vision. With recent advances in information technology and media, automated human interaction systems that include hand processing tasks such as hand detection, recognition, and movement tracking are being created. Real-time image processing using camera and motion sensor technology enables users to interact instantly with the application program and achieve the desired results. As technology continues to advance, it is expected that gesture recognition will become an even more integral part of human-computer interaction, making communication between humans and computers more seamless and intuitive.

References

[1] S. Hussain, R. Saxena, X. Han, J. A. Khan and H. Shin, "Hand gesture recognition using deep learning," 2017 International SoC Design Conference (ISOCC), Seoul, Korea (South), 2017, pp. 48-49, doi: 10.1109/ISOCC.2017.8368821.

[2] Nandwana, Bhumika & Tazi, Satyanarayan & Trivedi, Sheifalee & Kumar, Dinesh & Vipparthi, Santosh. (2017). A survey paper on hand gesture recognition. 147-152. 10.1109/CSNT.2017.8418527.

[3] N. Mohamed, M. B. Mustafa and N. Jomhari, "A Review of the Hand Gesture Recognition System: Current Progress and Future Directions," in IEEE Access, vol. 9, pp. 157422-157436, 2021, doi: 10.1109/ACCESS.2021.3129650.

[4] Hung, Chi-Huang & Bai, Ying-Wen & Wu, Hsu-Yao. (2016). Home outlet and LED array lamp controlled by a smartphone with a hand gesture recognition. 5-6. 10.1109/ICCE.2016.7430502.

[5] Dhall, Ishika & Vashisth, Shubham & Aggarwal, Garima. (2020). Automated Hand Gesture Recognition using a Deep Convolutional Neural Network model. 811-816. 10.1109/Confluence47617.2020.9057853.

[6] Hung, Chi-Huang & Bai, Ying-Wen & Wu, Hsu-Yao. (2015). Home appliance control by a hand gesture recognition belt in LED array lamp case. 599-600. 10.1109/GCCE.2015.7398611.

[7] Luzhnica, Granit & Simon, Joerg & Lex, Elisabeth & Pammer-Schindler, Viktoria. (2016). A Sliding Window Approach to Natural Hand Gesture Recognition using a Custom Data Glove. 10.1109/3DUI.2016.7460035.

[8] Kim, Ji-Hwan & Thang, Nguyen & Kim, Tae-Seong. (2009). 3-D hand motion tracking and gesture recognition using a data glove. 1013 - 1018. 10.1109/ISIE.2009.5221998.

[9] J. Shukla and A. Dwivedi, "A Method for Hand Gesture Recognition," 2014 Fourth International Conference on Communication Systems and Network Technologies, Bhopal, India, 2014, pp. 919-923, doi: 10.1109/CSNT.2014.189.

[10] Haria, Aashni & Subramanian, Archanasri & Asokkumar, Nivedhitha & Poddar, Shristi & Nayak, Jyothi. (2017). Hand Gesture Recognition for Human Computer Interaction. Procedia Computer Science. 115. 367-374. 10.1016/j.procs.2017.09.092.