ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR1337

Computer Vision Gestures Recognition System Using Centralized Cloud Server

Md Tanvir Hossain
Jain (Deemed-to-be University), Bangalore, India

Abstract:- Gesture recognition is a progressive technology that enables communication between humans and computers without any physical interaction. In recent years, gesture recognition technology has provided remarkable advancements in the technology world. A computer vision cloud-based gesture recognition system represents the potential of cloud computing to process and analyze the collected data. This paper examines the various growing aspects of gesture recognition using computer vision and cloud technology. Here we introduce a computer vision-based gesture processing unit that works on a cloud server and is used to enable real-time interaction between the physical world and the virtual server. It will help to use high-end gesture processing technology for every kind of device user. All the data will be collected using vision-based devices and monitored and processed on the centralized server. So the user doesn't have to be a highly efficient device. While using this system, the user can experience both the recent cutting-edge cloud infrastructure as well as the recent gesture recognition system. Not only humans but also object gestures can be recognized by using this proposed system. So this paper aims to represent a significant step toward a cloud-based interaction between human gestures and digital devices.

Keywords:- *Gesture Recognition, Centralized Cloud Server, Artificial Intelligence, Object Gesture, Human Gesture, Sensor Technology, Computer Vision, Deep Learning, Human-Computer Interaction.*

I. INTRODUCTION

In terms of gesture recognition techniques, there are multiple types of gesture processing techniques available. We have a myth that gesture recognition is only for human hand gesture recognition techniques. But actually, it can be used to identify everything from head nods moving head to different walking gaits. An object gesture recognition system is also available for objects. Like car movement tracking systems, automatic working machine factories, etc.

Gesture processing can be done through many sensors and cameras. In this cloud-based gesture recognition system, all the data we collect from various sources will be collected through a computer vision camera or sensor and transferred to a centralized server to process and show the output to the desired device or person.

Nowadays, almost everywhere, people are using digital devices in their day-to-day lives. From daily labor to IT engineers, all kinds of professionals use various digital devices to make their lives easy and hassle-free. However, because of the availability and high cost of the updated devices, all kinds of people are not able to use cutting-edge technology in their lives. Most of their devices are not up-to-date to perform highly efficient work. That is why most people are missing the benefits of today's technology, although they are connected to devices.

So our system will help all kinds of device users use the various gesture recognition techniques just by using their device camera or fundamental sensor. Nowadays, most smartphones have cameras and very basic, fundamental sensors to process various kinds of activity. However, the processing power of those devices needs to be higher because the chip or GPU CPU used on that device is low. So those devices can't process all that data and give an appropriate result.

By using this technique, these devices will collect the data from the physical world and send it to the cloud server. The cloud server will process all the data and provide the output to the user's device.

This cloud-integrated architecture provides a cost-effective, adaptable, and safe solution for managing and storing data. It can be useful for data storage and processing and allows users to use high-end technology without interruption.

This paper will tell us about the organized way to use cloud technology and connect with gesture recognition systems by using vision-based computer systems. It will also highlight the benefits of using a cloud-based gesture recognition system.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR1337

A. Working Flow Chart:

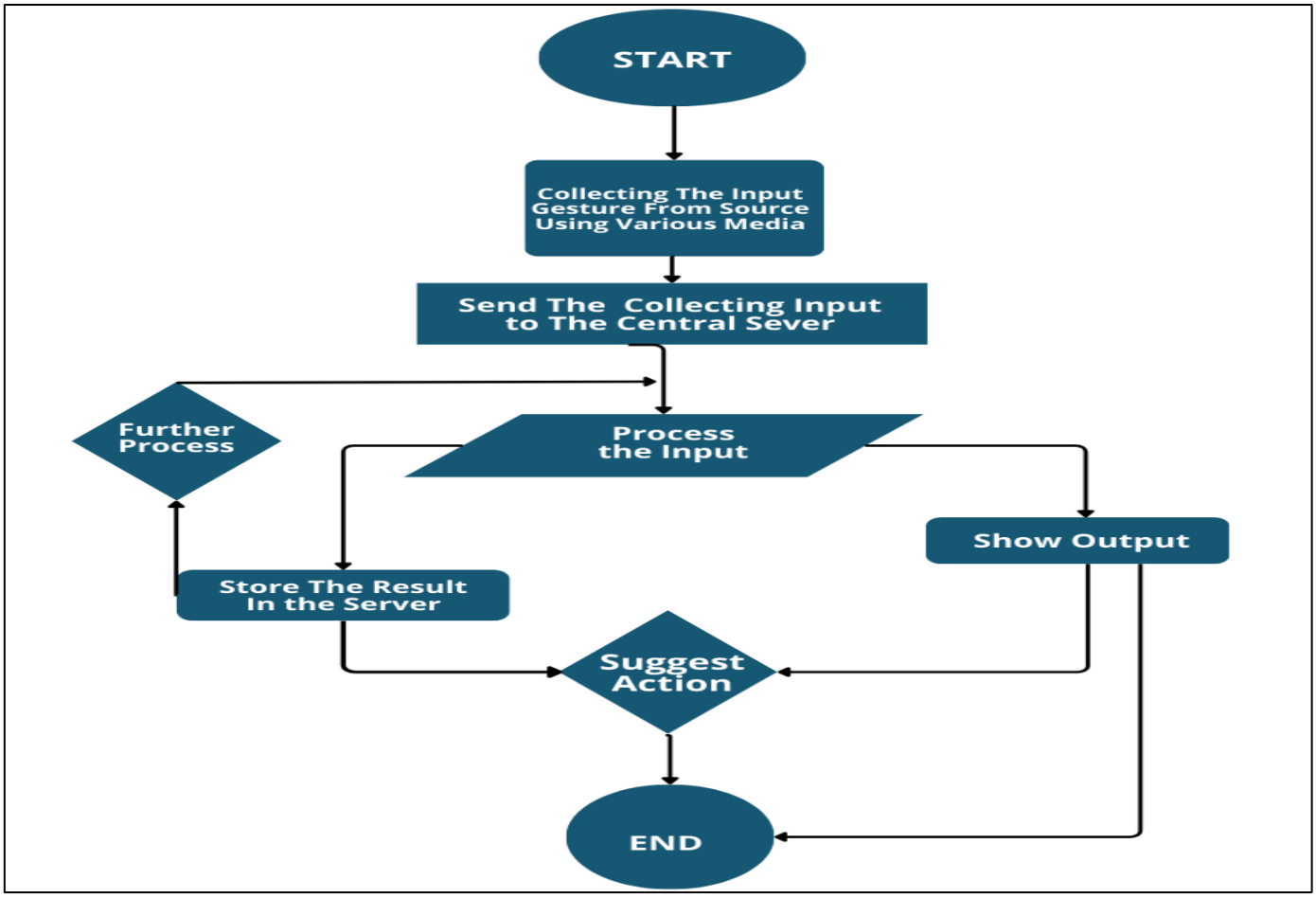

Fig 1: Working Flowchart of the System

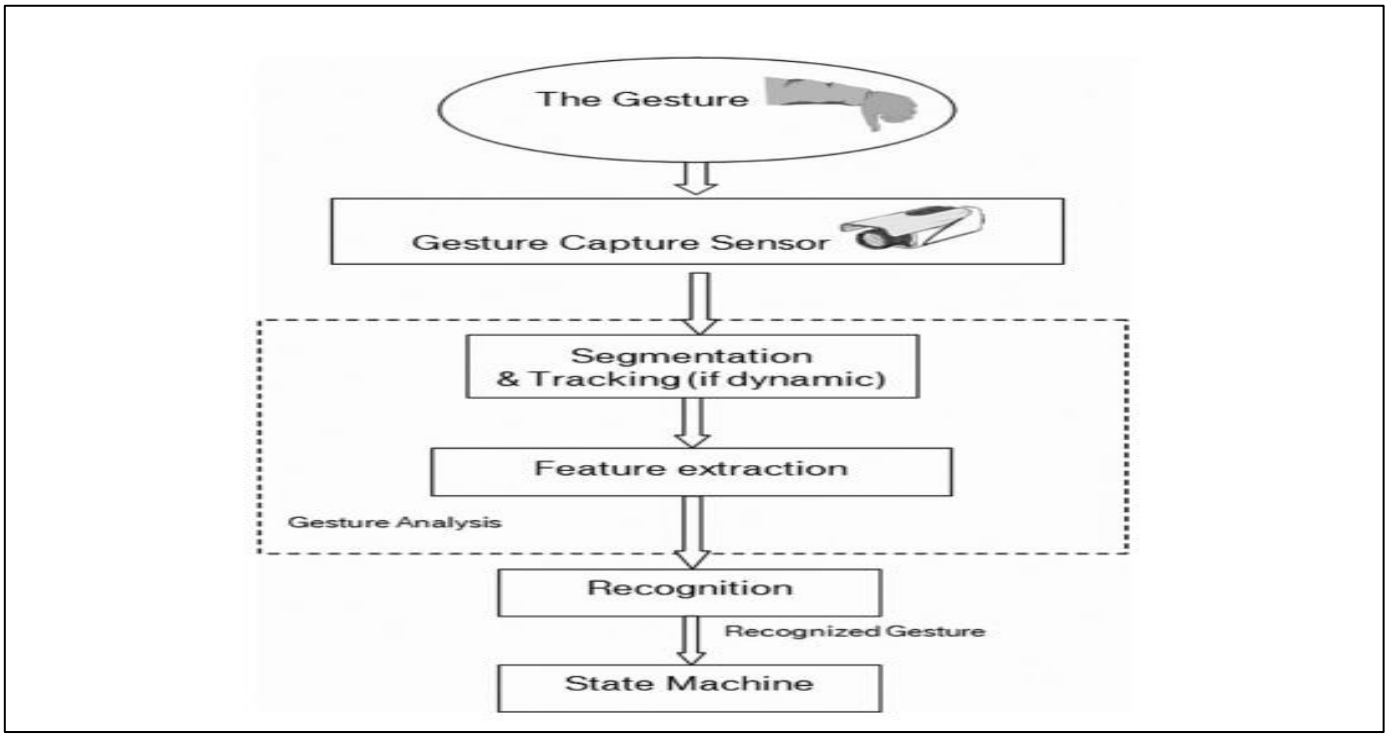

Fig 2: Gesture Recognition Process

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR1337

➢ *Let's Discuss the Above Working Flow Chart for the Proposed Method:*

Once the process begins, the sensors and camera will start collecting the data that is collected from the physical gesture. Then the collected data from the gesture shall be delivered to the main data center. In the central data center, the data will guide the process and take the necessary steps to produce a meaningful, full output. After the data is processed, the output will be shown to the user, and it will also be stored on the central server for further processing. From the admin side, it will suggest some action for the user to make it more useful.

B. Proposed Method:

In a computer vision-based gesture collection system, a computer-aided vision system will collect the targeted image, video, or movement of an object to proceed with its step. The goal of this system is to capture a gesture from a human or any other object and show the output to its administrator.

Three steps are usually involved in vision-based gesture recognition: gesture representation and feature extraction, recognition, pre-processing, and gesture detection. [7]

To implement the system, we have to design three different systems that work together to make the system possible. They are:

- Computer vision system to understand the type of object
- Gesture recognition system to complete the gesture processing
- The last one is the centralized cloud server, where the data will be processed and stored.

While using the computer vision system, we have to choose the system that has already been trained to process the gesture. Using various artificial intelligence systems and with the help of machine learning algorithms and deep learning techniques, the system will figure out what the gesture is defining. As the system will be retained, it can easily predict what this gesture is trying to say. The system will use a multilayer neural network to process its data.

The gesture capture will be done through various cameras and sensors. Using a device's camera or an external camera, the image or video will be captured and sent to the server. In this part, we will also use some sensors that can be implemented into the device to capture other gestures and movements, or maybe the presence of the object and substances. Many sensors, like infrared sensors, accelerometers, gyroscopes, and other touchscreen gestures, are also used to collect data from the physical world.

Once the data is collected, we will send it to the central server, where our computer vision-pre-trained machine learning system is available to proceed further. The cloud server will be connected to the local server using communication protocols and APIs.

In the server processing part, we will use computer vision and its multilayer neural network system, CNN (convolutional neural network), to make the computational process timeless. For that, the time complexity will be less. It will use the pre-trained pattern recognition technique.

➢ *Basic Algorithm for Processing the Data*

- Import the library
- Set the parameter and define the kernel
- Load the input and plot it according to our needs
- Apply the neural network operation and plot the output

After completing the process on the server, the system will take the automatic action. Sometimes the suggestion will appear on the device screen and store the output.

While the captured gesture data is transmitted or sent from a local device to a cloud infrastructure, security should be provided by using encryption tools. The communication medium should be secured properly to prevent data loss and any changes to the original data.

C. Diagram and Picture:

Below is the Visual diagram of how The whole system works:

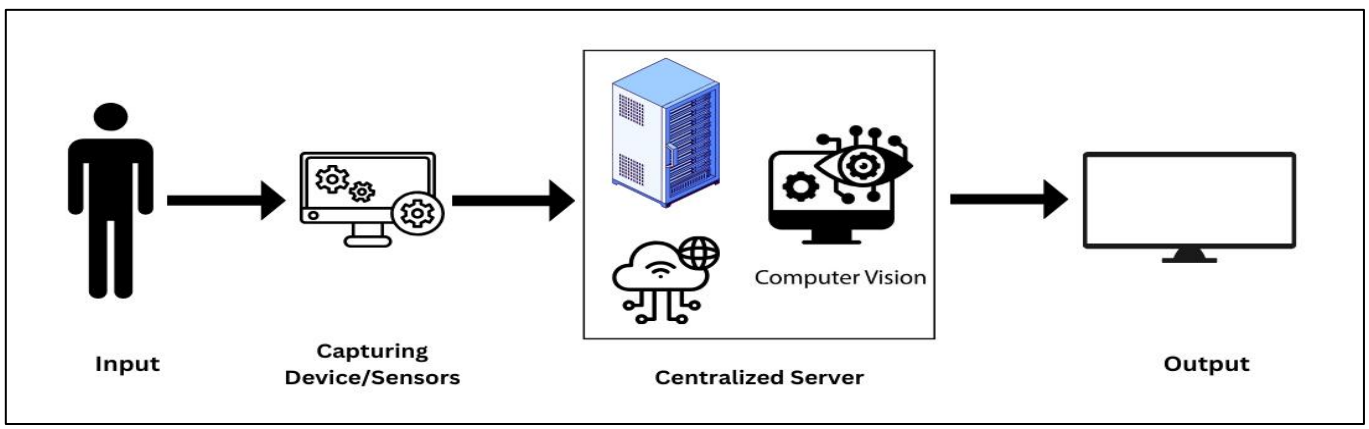

Fig 3: Visual Representation of How System Works

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR1337

➢ *How Computer Vision Works Inside the Server:*

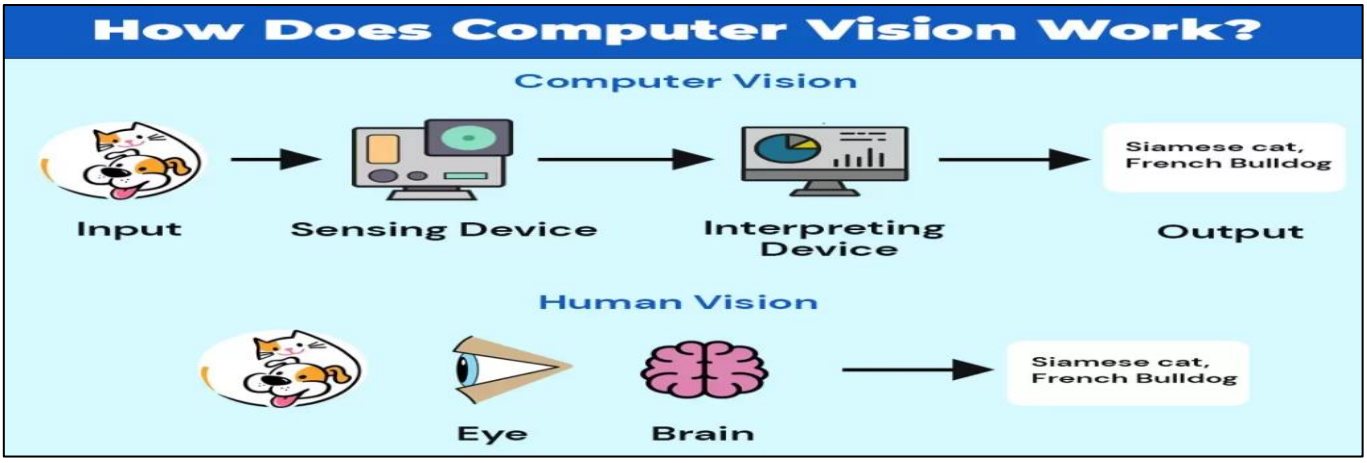

Fig 4: Computer Vision Working Process

➢ *The Inside Working Procedure of CNN (Convolutional Neural Network)*

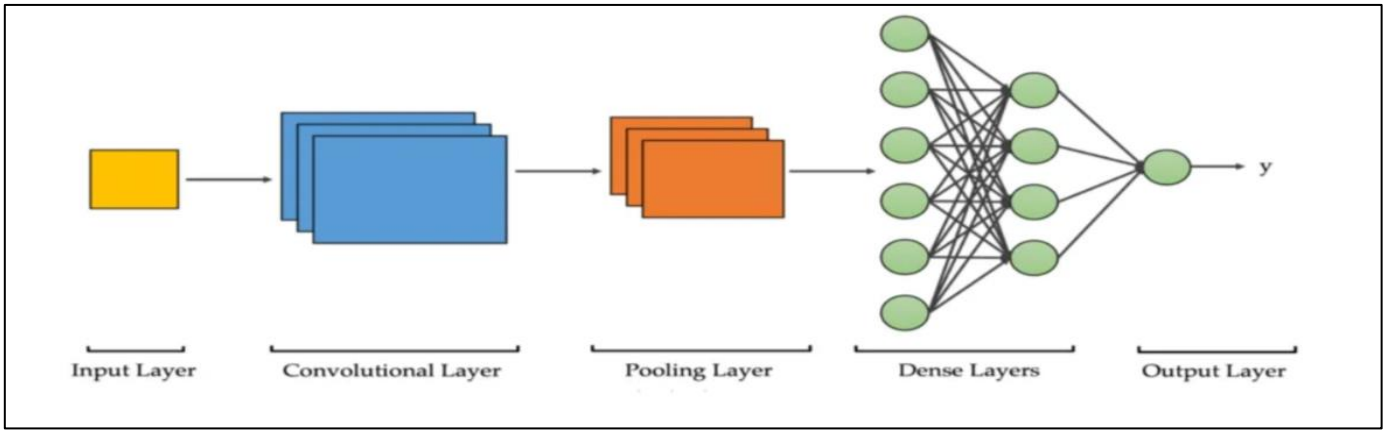

Fig 5: CNN (Convolutional Neural Network) Technology Inside Computer Vision

The method employed by the Convolutional Neural Network (CNN) operates in that it takes an input image, weighs its different features according to the unique items in the image, and then uses those weights to determine and separate the objects from each other.

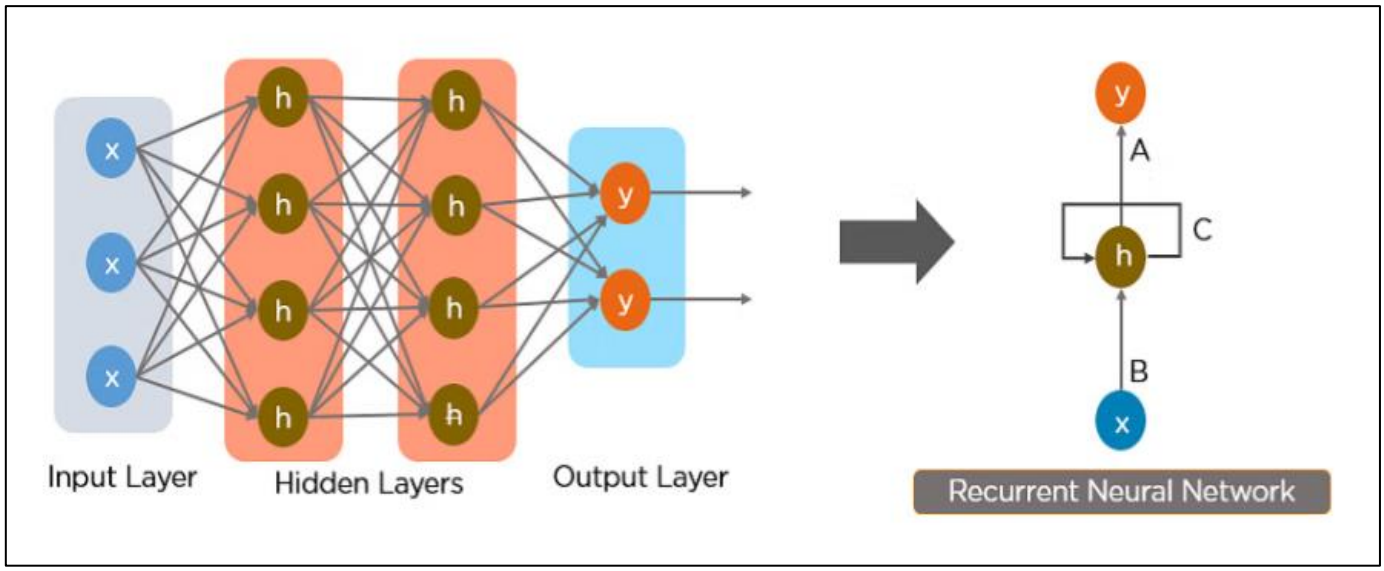

Fig 6: RNN (Recurrent Neural Networks) Technology Inside Computer Vision

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR1337

Also Inside the computer vision, the Recurrent Neural Networks (RNNs) operate by taking a particular layer's output and applying it as an input cyclically to predict outputs that come after. In the recurrent neural network, the nodes from various neural network layers are combined to form a single layer.

II. TECHNOLOGY USED

While developing we can use many recent cutting-edge technologies that work together to deploy this system successfully. From capturing gestures from the physical world to showing output to the user there are many sub-technology involved that work at the same time to provide the right output. Let's see some of the potential topics that are involved in this system.

While starting to capture the gesture we use sensors and a camera to capture the gesture. After that, the communication media takes place to send the collected data to the server. There the machine learning algorithm uses an image processing technique to complete the gesture recognition and recognize the gesture to show output. Here we use cloud computing technology to store data and also process them in the centralized server. Here is also the real-time data processing technology used.

At last, the system will suggest some action to take for the user. There we use the feedback system using artificial intelligence and a pre pre-trained model. So overall we can say we use the below technology to make this system successful:

- Sensor technology
- Pattern matching algorithms
- Image processing
- Gesture recognition technique
- Cloud server
- Real-time data processing

III. LITERATURE REVIEW

Table 1: Discussion

Title	Authors	Discussion
Gesture recognition using recurrent neural networks	K Murakami, H Taguchi	This paper discusses the system of how Japanese sign language is being used in daily communication and from there how we can develop a gesture recognition system. It also tells about the different kinds of posture recognition systems that are less complicated to proceed in the gesture recognition system. Here they are using RNN to develop their model. The author showed the connection between sign language and gesture recognition systems using recent technology.
Review of constraints on vision-based gesture recognition for human-computer interaction	Biplab Ketan Chakraborty, Debajit Sarma, M.K. Bhuyan, Karl F MacDorman	The authors are deploying the process of how gesture recognition is helping human-computer interaction nowadays. It is a long and complex process to develop the system where the interaction takes place using gestures only. Mainly it focuses on the hand gesture process and how it works in computer vision technology. It also discusses how vision vision-guided robot system works for the HCL process. But it works under a very limited amount of boundaries.
Space-time event clouds for gesture recognition: From RGB cameras to event cameras	Q Wang, Y Zhang, J Yuan, Y Lu	The space-time cloud for gesture recognition is the process of capturing images through the light that comes from the images and it will detect the color light and object shapes. The following paper discusses the process of how the point net system is used to capture the motion of an object whose gesture is being captured. It also tells how the fastest-moving object gestures are being captured and processed using the frame rate.
A Survey on Intelligent Gesture Recognition Techniques.	JJ Ojeda-Castelo, MLM Capobianco-Uriarte, JA Piedra-Fernandez, R Ayala	Without having physical interaction with the system the gesture recognition system can allow users to use the system. The authors of this paper show how artificial intelligence is helping the gesture recognition system in the process of completing the task and how the system is helping in the pandemic time, where touching things can be a major issue. The authors have developed their recognition model that helps to interact between gesture recognition and the AI system. They also showed the growing aspect of research and work that is going on with this topic based on the data that are available till 2000-2020.
Pantomime: Mid-air gesture recognition with sparse millimeter-wave radar point clouds	S Palipana, D Salami, LA Leiva, S Sigg	Motion gesture recognition using millimeter radio waves and high-frequency sensing is the main point of this paper. The authors developed a system called “Pantomine” which has an accuracy level of (95-97)% in terms of moveable gesture recognition. As a result, it can detect not only human gestures but also it can recognize the objects and elements that are continuously moving from one palace to another.
Gelicit: a cloud platform for distributed gesture	N Magrofuoco, J Vanderdonckt	This paper mainly focuses on the area of the advantages and disadvantages of the gesture recognition system. It also focuses mainly on three parts. The

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR1337

elicitation studies		model-based and the cloud platform can be used in the system. They introduce a cloud platform called "Gelicit" claiming that it is the first ever cloud platform for the gesture identification system.
Hardware/Software Codesign for Hand gesture recognition using a Convolutional Neural Network.	S Said, L Kalms, D Göhringer, MAA El Ghany	The following paper discusses the focusing area of what needs to be considered when it comes to detecting the gesture of an object. The factors that can lead to a complex situation for the system. They also focusing the time complexity of the system which will help to improve the consistency. They claimed that their accuracy level is 94% in terms of overcoming the challenges.
Survey on Emotional Body Gesture Recognition	Fatemeh Noroozi, Ciprian Adrian Corneanu, Dorota Kaminska, Tomasz Sapinski, Sergio Escalera, Gholamreza Anbarjafari,	Emotion body gesture detection is one the important topics nowadays for the researcher. This paper tells us the pipeline of how an emotional body gesture recognition determines the situation of a person in real-time. This research shows how the body gesture defines one current situation of his mind. In criminology and psychology, this technique can help to detect the situation of the victim. This research mainly determines that body language can be a source of anyone's emotional outcome based on the gender and culture that the person belongs to.

IV. RESULTS AND ANALYSIS

We evaluate the performance of this system by using the accuracy level and the time it takes to reach its final output that will reflect the user. In the diagram section, it shows how the system will work and the final output.

The accuracy level is dependent on the system training. How many data sets are trained to complete the training process. On the other hand, the cloud server system is also essential to completing the task. The best result is when the server works seamlessly and provides faster data processing. The storage system plays a vital role in the output result.

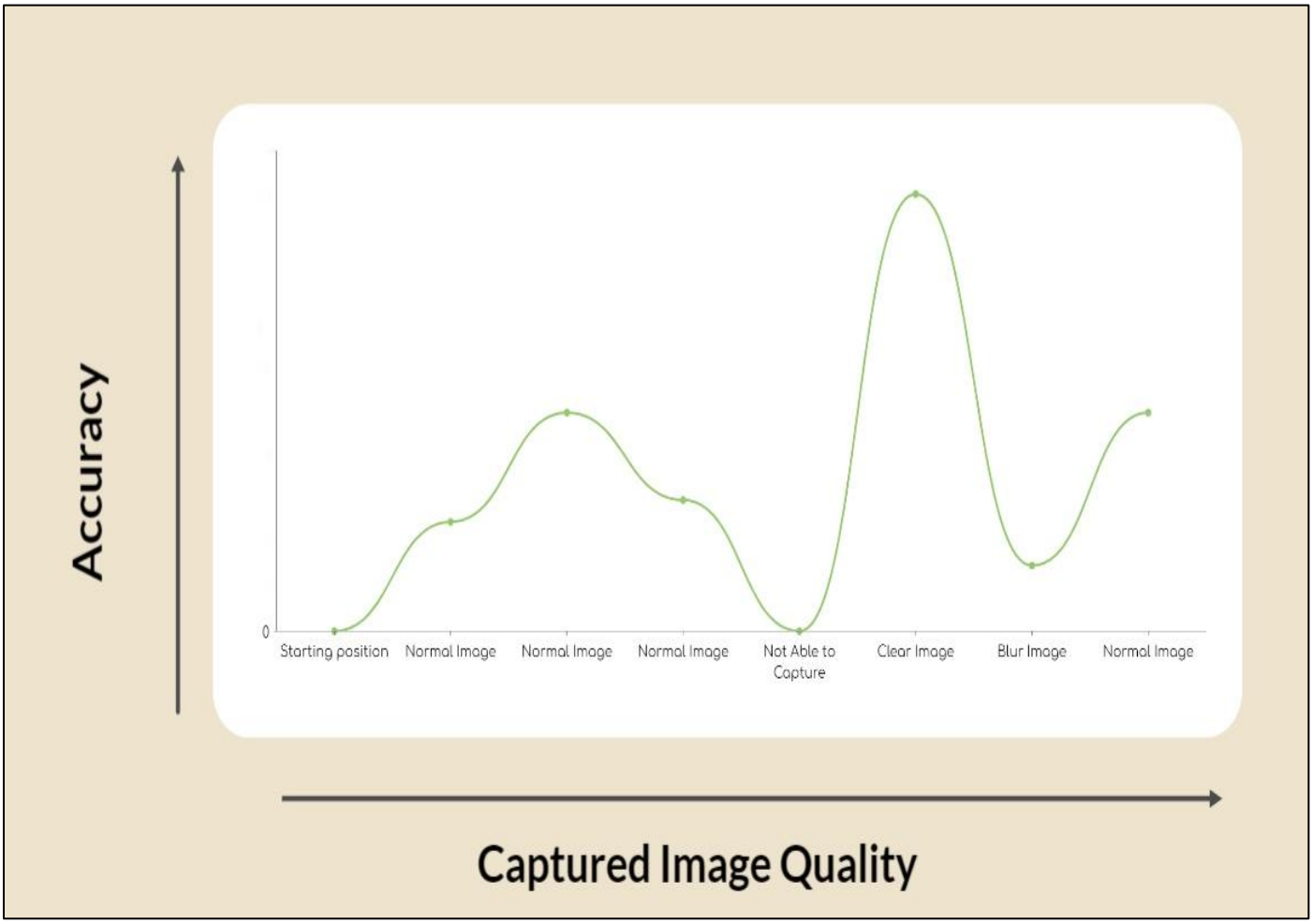

Fig 7: Accuracy Graph based on Captured Image

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR1337

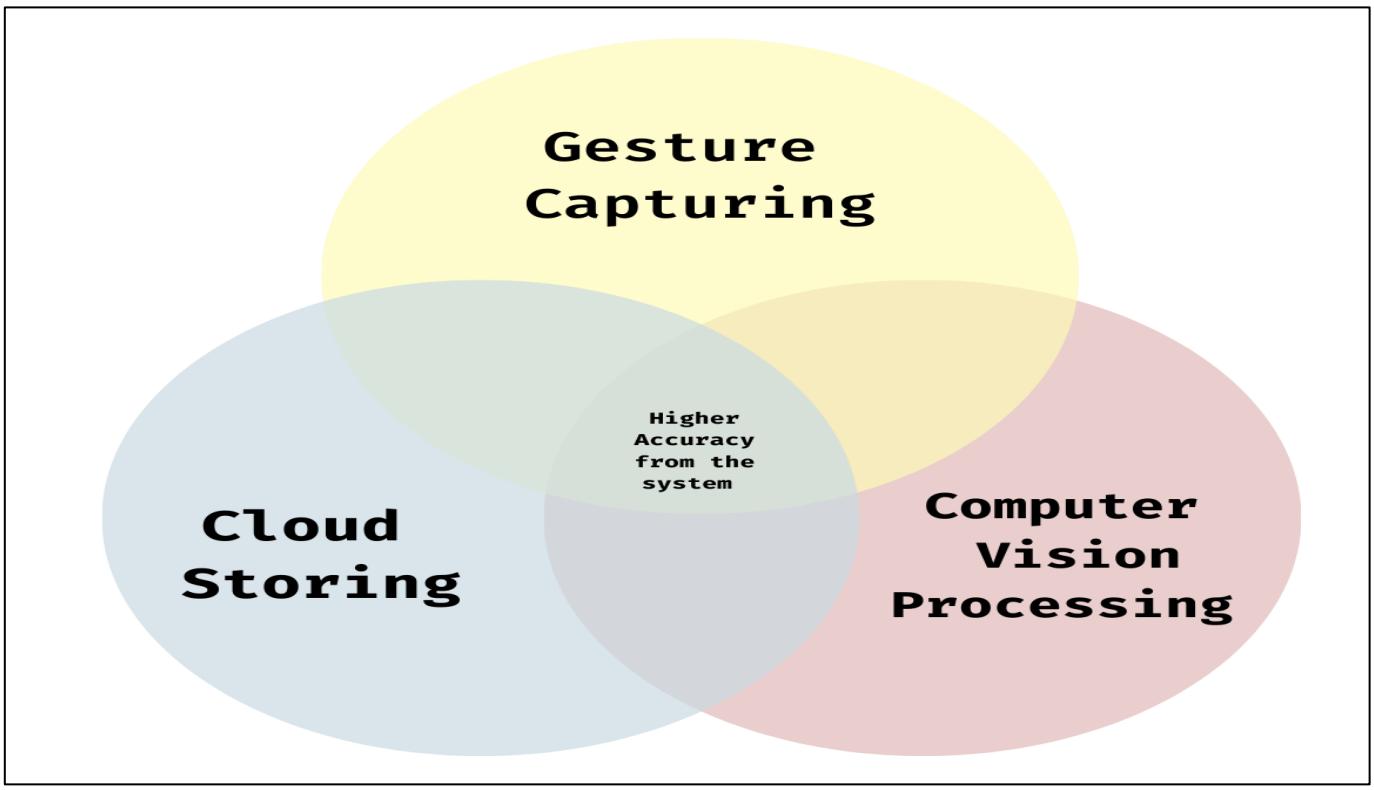

Fig 8: High Accuracy Factor Representation using Ven Diagram

Finally, we can show in this paper that using a computer vision technology that includes CNN & RNN to process the collected image in a recognition system it's possible to get an output from this system with an accuracy level of almost (95%-97%).

V. CHALLENGES / LIMITATION

One of the most challenging study areas in the field of computer vision is human motion gesture detection. Applications for it can be found in many different domains, including virtual reality, intelligent monitoring, human behavior analysis, and human-computer interaction. The complex nature of understanding human gestures adds another level of difficulty, therefore developments in this field are essential to unlocking creative solutions and improving human-machine interaction. [25]

While developing a cloud-based computer vision system for recognition, a lot of limitations and challenges will take place. This is a rapidly growing field in computer technology. However, developing and using this properly is a tough task. As it is a high cost and professionals are needed to manage and maintain the system.

While proceeding with the images occlusion, background clutter, variations in illumination, busy backdrops, light sensitivity, form, speed, size, and self-occlusion. can happen in computer vision which is used to detect the gesture. [24] Using CNN, we obtain an accuracy of up to 97.06%. [26]

As it is a cloud-based vision program system that will process its program in the cloud only, the communication will happen wirelessly only. So the network issue will be the first and major issue that will arrive. Also, the security chain for the communication media will be a major issue as the data will transfer wireless only. Maintaining the cloud server and the data center will be tough as lots of data will be processed at a time. While collecting and detecting the image-based gesture the image quality and noise will be an issue. Sometimes because of the lighting background, a type of image will be an issue for the sensor to collect and process in the system. These are the limitations and challenges that will be faced based on this system while deploying in real-time in any project.

VI. FUTURE SCOPE

Computer vision data or image processing is an emerging technology that is using the gesture recognition system nowadays. By using this feature in the upcoming days gesture recognition will be possible in any sort of device. The centralized cloud server also plays a vital role where the computer vision system will be held and monitored.

In future the gesture recognition can take place without having a physical server as well as any physical interaction of people. So it's time to properly focus and cope with this technology so that we can benefit from this shortly.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR1337

Future research could concentrate on optimizing algorithms for increased accuracy, real-time processing capabilities, and seamless integration with edge computing. The future of gesture recognition systems holds innovation that will have a big impact on human-computer interaction in a variety of industries as long as technology keeps improving.

VII. CONCLUSION

The creation and deployment of a Computer Vision Gestures Recognition System that makes use of a Centralized Cloud Server is a significant advancement in improving user experience and human-computer interaction. Having a centralized cloud server boosts the system's accessibility and scalability, facilitating broad adoption and smooth interaction with other applications. To extract the actual output in the gesture recognition system we are using computer vision technology to accurately predict the data in this paper. Using the following method and technology, we can successfully implement a system that can predict the output and show some results. In this paper, we properly discuss how a recurrent neural network system is being used in computer vision that will help to recognize the gesture and store the output in a cloud server for further use. The work completely is in the research phase and further development and implementation will take place. This research presents exciting implications for the future of gesture recognition systems and their integration into other domains, given the close relationship between computer vision and cloud computing in the gesture recognition field.

REFERENCES

[1]. Wang, Qinyi, et al. "Space-time event clouds for gesture recognition: From RGB cameras to event cameras." 2019 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2019.
[2]. Murakami, Kouichi, and Hitomi Taguchi. "Gesture recognition using recurrent neural networks." Proceedings of the SIGCHI conference on Human factors in computing systems. 1991.
[3]. Chong, Yanwen, Jianfeng Huang, and Shaoming Pan. "Hand Gesture recognition using appearance features based on 3D point cloud." Journal of Software Engineering and Applications 9.4 (2016): 103-111.
[4]. Miranda, Leandro, et al. "Real-time gesture recognition from depth data through key poses learning and decision forests." 2012 25th SIBGRAPI Conference on Graphics, Patterns and Images. IEEE, 2012.
[5]. Berman, Sigal, and Helman Stern. "Sensors for gesture recognition systems." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 42.3 (2011): 277-290.
[6]. Bailador, Gonzalo, et al. "Real time gesture recognition using continuous time recurrent neural networks." 2nd International ICST Conference on Body Area Networks. 2007.
[7]. Chakraborty, Biplab Ketan, et al. "Review of constraints on vision-based gesture recognition for human–computer interaction." IET Computer Vision 12.1 (2018): 3-15.
[8]. Jiang, Du, et al. "Gesture recognition based on binocular vision." Cluster Computing 22 (2019): 13261-13271.
[9]. Ng, Chan Wah, and Surendra Ranganath. "Real-time gesture recognition system and application." Image and Vision computing 20.13-14 (2002): 993-1007.
[10]. Cheok, Ming Jin, Zaid Omar, and Mohamed Hisham Jaward. "A review of hand gesture and sign language recognition techniques." International Journal of Machine Learning and Cybernetics 10 (2019): 131-153.
[11]. Ojeda-Castelo, Juan Jesus, et al. "A Survey on Intelligent Gesture Recognition Techniques." IEEE Access 10 (2022): 87135-87156.
[12]. Ong, Sylvie CW, and Surendra Ranganath. "Automatic sign language analysis: A survey and the future beyond lexical meaning." IEEE Transactions on Pattern Analysis & Machine Intelligence 27.06 (2005): 873-891.
[13]. Magrofuoco, Nathan, and Jean Vanderdonckt. "Gelicit: a cloud platform for distributed gesture elicitation studies." Proceedings of the ACM on Human-Computer Interaction 3.EICS (2019): 1-41.
[14]. Suarez, Jesus, and Robin R. Murphy. "Hand gesture recognition with depth images: A review." 2012 IEEE RO-MAN: the 21st IEEE international symposium on robot and human interactive communication. IEEE, 2012.
[15]. Min, Yuecong, et al. "An efficient pointlstm for point clouds based gesture recognition." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.
[16]. Palipana, Sameera, et al. "Pantomime: Mid-air gesture recognition with sparse millimeter-wave radar point clouds." Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies 5.1 (2021): 1-27.
[17]. Nock, Matthew K., and Ronald C. Kessler. "Prevalence of and risk factors for suicide attempts versus suicide gestures: analysis of the National Comorbidity Survey." Journal of abnormal psychology 115.3 (2006): 616.
[18]. Burgess, Robin, et al. "Selection of large-scale 3d point cloud data using gesture recognition." Technological Innovation for Cloud-Based Engineering Systems: 6th IFIP WG 5.5/SOCOLNET Doctoral Conference on Computing, Electrical and Industrial Systems, DoCEIS 2015, Costa de Caparica, Portugal, April 13-15, 2015, Proceedings 6. Springer International Publishing, 2015.
[19]. Noroozi, Fatemeh, et al. "Survey on emotional body gesture recognition." IEEE transactions on affective computing 12.2 (2018): 505-523.
[20]. Ibraheem, Noor Adnan, and RafiqulZaman Khan. "Survey on various gesture recognition technologies and techniques." International journal of computer applications 50.7 (2012): 38-44.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR1337

[21]. Wu, Ying, and Thomas S. Huang. "Vision-based gesture recognition: A review." International gesture workshop. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999.
[22]. Davis, James, and Mubarak Shah. "Visual gesture recognition." IEE Proceedings-Vision, Image and Signal Processing 141.2 (1994): 101-106.
[23]. Pavlovic, Vladimir I., Rajeev Sharma, and Thomas S. Huang. "Visual interpretation of hand gestures for human-computer interaction: A review." IEEE Transactions on pattern analysis and machine intelligence 19.7 (1997): 677-695.
[24]. Said, Sarah, et al. "Hardware/Software-Codesign for Hand Gestures Recognition using a Convolutional Neural Network." Proceedings of the INTelligent Embedded Systems Architectures and Applications Workshop 2019. 2019.
[25]. Ma, Rui, Zhendong Zhang, and Enqing Chen. "Human motion gesture recognition based on computer vision." Complexity 2021 (2021): 1-11.
[26]. Martija, Mygel Andrei M., Jakov Ivan S. Dumbrique, and Prospero C. Naval Jr. "Underwater gesture recognition using classical computer vision and deep learning techniques." (2020).