International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 03 | March - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

# Hand Gesture Recognition System Using Thermal Images: A Systematic Literature Review

Aishwarya Bhandari R
*Department of Information Science & Engineering*
*Malnad College of Engineering*
Hassan, India
aishwaryabhandari05@gmail.com

Ashika R
*Department of Information Science & Engineering*
*Malnad College of Engineering*
Hassan, India
ashikaramesh626@gmail.com

Mr. Krishna Swaroop A
Assistant Professor
*Department of Information Science & Engineering*
*Malnad College of Engineering*
Hassan, India
ksa@mcehassan.ac.in

Monika B S
*Department of Information Science & Engineering*
*Malnad College of Engineering*
Hassan, India
monikabsecomm@gmail.com

Rakshitha D
*Department of Information Science & Engineering*
*Malnad College of Engineering*
Hassan, India
rakshithadoddegowda@gmail.com

***Abstract*— Leading the way in human-computer interaction (HCI) advancements are hand gesture recognition (HGR) systems, which provide touchless, intuitive control for a wide range of applications, including healthcare, smart devices, assistive technologies, and virtual environments. In order to demonstrate the development of HGR methodologies, such as vision-based, skeleton-based, and hybrid approaches, this paper summarizes the results of recent studies. The field is dominated by vision-based systems, which use depth sensors and RGB cameras to recognize both static and dynamic gestures. Particularly for continuous gesture recognition, skeleton-based methods such as the TD-Net architecture offer computationally effective and lightweight solutions. While multimodal systems that combine RGB, thermal, and depth data improve robustness, thermal imaging expands the applicability of HGR to low-light conditions. Despite advancements, problems still exist, such as signer-independent performance, environmental adaptability, and gesture spotting in continuous streams. Directions for the future.**

***Keywords*—** *Machine learning, Deep learning, Hand Gesture, image processing, Convolutional neural networks, Thermal images.*

## I. INTRODUCTION

In many fields, hand gesture recognition has become an essential technology that connects accessibility and human-computer interaction (HCI). Significant progress has been made over time in utilizing a variety of hardware, computational methods, and sensing modalities to create reliable gesture recognition systems.

The core of hand gesture recognition is the ability to record and decipher human hand movements for use in robotics, healthcare, and other fields. Deep convolutional networks have been used in static gesture recognition research, and CNNs, TD-Nets, and YOLO-based architectures have been used in dynamic gesture tracking. The integration of depth sensors, RGB cameras, thermal imaging, and infrared imagery has guaranteed flexibility in response to various operational and environmental limitations. While some systems, like the Live Demonstration: Home Appliance Control with 3D Hand Skeletons [13], investigate 3D skeletal data for dynamic control, others, like the Robust Hand Gesture Recognition Using Deep CNN and Thermal Images [1], focus on thermal imaging for low-light conditions. Energy-efficient edge applications are the subject of papers like An Ultra-Low-Power Real-Time Gesture Recognition System [17], which demonstrate how versatile hand gesture recognition is across a range of hardware platforms.

Interactive interfaces in gaming and assistive technologies have been made possible by dynamic gesture recognition for real-time applications, as covered in papers such as Dynamic Hand Gesture Recognition [9] and An Approach for Human Machine Interaction [19]. Furthermore, systems designed for particular environments, like space exploration or underwater environments, demonstrate how versatile gesture recognition is in unusual contexts.

With projects like Hand Sign Recognition Using Infrared Imagery [3] and Real-Time Hand Gesture Recognition and Sentence Generation for Deaf Communication [16], the field also places a strong emphasis on inclusivity and improving communication for people with disabilities.

Not with standing the advancements, problems still exist in areas like real-time scalability, complex gesture detection, lighting variability, and dataset diversity. Future developments in adaptive learning algorithms, robust preprocessing, and multimodal sensing are probably going to overcome these constraints.

## II. RELATED WORKS

*A. Paper Title: Robust Hand Gestures Recognition Using a Deep CNN and Thermal Images. Year of Publication:2021*

*Description:* The limitations of RGB cameras in low light are addressed in this paper by presenting a novel hand gesture recognition system that combines deeplearning and high-resolution thermal imaging. The method uses thermal imaging to guarantee robustness in all lighting conditions. In fusion and grayscale formats, a dataset of 14,400 photos with 10 different gestures was produced. For high accuracy and quick inference times, a lightweight CNN model issued. Deployment of the system on edge devices such as the Nvidia Jetson AGX Xavier and Raspberry Pi 4 is optimized.

*Methodology:* Using the FLIR Lepton 3.5 thermal camera, the methodology starts by building a dataset that captures a variety of hand shapes and motions for resilience. Preprocessed images were used to enhance training and balance the distribution of classes. Using batch normalization, max-pooling, and ReLU activations, a CNN comprising four convolutional layers—two of which had dilation rate 2—was created to efficiently capture spatial features. The Adam optimizer was used to train the model, and MobileNetV3 architectures were used as benchmarks. TensorFlow Lite was used to convert the final models for testing on edge devices such as the Nvidia Jetson AGX Xavier and Raspberry Pi 4.

*Limitations:* One of the system's drawbacks is its dependence on a regulated dataset with consistent backgrounds, which limits its effectiveness in dynamic real-world situations. It is sensitive to changes in hand temperature brought on by outside influences and concentrates on static gestures, excluding dynamic ones. The lightweight model may have trouble scaling to larger datasets or more complex gestures, even though it is optimized for edge devices.

*Key Insights:* The study demonstrates the benefits of thermal imaging for gesture recognition, which performs exceptionally well in dim and fluctuating lighting. Thermal cameras guarantee reliable performance without requiring outside illumination. With a small 2 MB TensorFlow Lite size and 98.81% accuracy, the lightweight CNN model can be used in real-time on edge devices. Applications in human-computer interaction, smart home automation, and assistive technologies are best suited for this system. It provides a foundation for future advancements, including dynamic gesture analysis and handling complex scenarios.

*Citation:* [1] Daniel Skomedal Breland , Aveen Dayal , Ajit Jha , Phaneendra K. Yalavarthy , Senior Member, IEEE, Om Jee Pandey , Senior Member, IEEE, and Linga Reddy Cenkeramaddi , Senior Member, IEEE (2021).Robust Hand Gestures Recognition using a Deep CNN and Thermal Images. IEEE Access,DOI:[10.1109/JSEN.2021.3119977]

*B. Paper Title: Deep Learning-Based Sign Language Digits Recognition From Thermal Images With Edge Computing System, Year of publication : 2021*

*Description:* This aims to integrate a Raspberry Pi and a 32x32 pixel thermal camera into an edge computing system. Robust performance in dim or fluctuating environments is ensured by the lightweight model's 99.52% accuracy and lighting invariance. For real-time applications, the system prioritizes efficiency and portability. This invention shows how thermal imaging can be used for precise, affordable classification.

*Methodology:* The Omron D6T camera in the thermal imaging system is linked to the Raspberry Pi through a specially designed shield. Multiple people contributed 3200 thermal images that covered 10 digit classes to the dataset. The residual learning-inspired lightweight CNN model is tailored for edge devices with 10 MB of memory and 851,978 parameters. With a batch size of eight and cross-entropy loss, the Adam optimizer was used for training. Superior accuracy and efficiency on embedded systems were confirmed through benchmarking against BiT models.

*Limitations*: When there is little temperature difference between the hand and background, low-resolution thermal images make classification difficult. The system has not yet been tested on complex backgrounds. Due to its exclusive focus on particular hand orientations and gestures, the dataset lacks diversity. It does not recognize dynamic gestures or more general sign language; it is restricted to static digit recognition. These restrictions point to the need for more advancements and practical testing.

*Key Insights:* On edge devices, the lightweight CNN outperforms heavier models like BiT while achieving high accuracy. Consistent performance in dim or fluctuating lighting is guaranteed by thermal imaging. Integration with Raspberry Pi demonstrates that real-time AI can be implemented on low-resource, portable devices. The system can be used in difficult environments because it is not dependent on visible light. Applications where RGB cameras might not work include robotics, healthcare, and assistive technology.

*Citation:* [2] Daniel S. Breland, Simen B. Skriubakken, Aveen Dayal, Ajit Jha, Phaneendra K. Yalavarthy, Senior Member, IEEE, and Linga Reddy Cenkeramaddi , Senior Member, IEE (2021). Deep Learning-Based Sign Language Digits Recognition From Thermal Images With Edge Computing System. IEEE Access, DOI: [10.1109/JSEN.2021.3061608]

*C. Paper title: Hand Sign Recognition using Infrared Imagery Provided by Leap Motion Controller and Computer Vision,Year of publication : 2021*

*Description*: The study presents a hand sign recognition system for assistive technology and HCI applications that uses infrared images from the Leap Motion Controller. To guarantee quality in a variety of settings, it preprocesses photos using noise reduction, lighting normalization, and grayscale conversion. Using depth information for 3D gesture differentiation, feature extraction locates keypoints, hand edges, and finger joints. SVMs and CNNs are examples of machine learning models that use labeled datasets to classify gestures in real time. Gestures that are recognized are mapped to particular actions and post-processed for accuracy, allowing for gesture-controlled applications.

*Methodology:* The Leap Motion Controller, which has depth-sensing capabilities, is used to first take infrared pictures of hand gestures. In addition to background subtraction, noise reduction methods like Gaussian or median filtering are used to eliminate artifacts. In addition to lowering processing requirements, grayscale conversion streamlines the photos while maintaining important details like finger joints and hand contours. By standardizing image intensity, lighting normalization guarantees consistent performance in a range of environmental conditions. In order to increase feature clarity for further analysis, enhanced edges and contours are highlighted.

*Limitations:* The system's drawbacks include difficulties with preprocessing, like managing sharp changes in lighting or extremely noisy settings that could degrade image quality. The accuracy of gesture recognition may be decreased by feature extraction's inability to handle unclear or obscured hand poses. The quality and diversity of the training dataset determine classification, which may restrict generalization to invisible gestures. For applications that need quick responses, post-processing can add latency and affect real-time performance. For users with distinct hand movements or in complex situations, mapping gestures to commands might not work as well.

*Key Insights:* The paper highlights the use of infrared imagery from the Leap Motion Controller for precise hand tracking and gesture recognition, even in low-light settings. A multistage process includes image preprocessing, feature extraction using depth data, and classification with machine learning models like SVMs and CNNs. Depth data enables 3D pose understanding, capturing intricate hand movements for improved accuracy. Supervised learning trains the system to recognize gestures in real-time, essential for dynamic applications. Real-time performance ensures responsiveness, making it suitable for gesture-controlled interfaces and assistive technologies.

*Citation:* [3] Tathagat Banerjee; K.V. Pavan Srikar; S. Ashvith Reddy; Krishna Sai Biradar; Rithika Reddy Koripally; Gummadi. Varshith. (2021). Hand Sign Recognition using Infrared Imagery Provided by Leap Motion Controller and Computer Vision. IEEE Access, DOI: [10.1109/ICIPTM52218.2021.9388334].

*D. Paper title: Smart Healthcare Hand Gesture Recognition Using CNN-Based Detector and Deep Belief Network, Year of publication : 2023*

*Description:* The paper introduces a hand gesture recognition system tailored for healthcare applications. The system integrates Convolutional Neural Networks (CNN) for gesture detection and Deep Belief Networks (DBN) for classification. The goal is to achieve high accuracy in recognizing hand gestures under complex environments using RGB data.

*Methodology:* The methodology starts with extracting video frames and enhancing image quality using gamma correction and non-local mean filtering. Hand detection is performed using a CNN with convolution, max-pooling, and fully connected layers. Spatial features are captured through Neural Gas clustering, while Locomotion Thermal Mapping highlights gesture dynamics based on hand movement speed. Fuzzy logic optimization refines the feature set, which is classified into gestures by a Deep Belief Network (DBN).

*Limitations:* The system's drawbacks include difficulties with preprocessing, like managing sharp changes in lighting or extremely noisy settings that could degrade image quality. The accuracy of gesture recognition may be decreased by feature extraction's inability to handle unclear or obscured hand poses. The quality and diversity of the training dataset determine classification, which may restrict generalization to invisible gestures. For applications that need quick responses, post-processing can add latency and affect real-time performance. For users with distinct hand movements or in complex situations, mapping gestures to commands might not work as well.

*Key Insights:* For healthcare applications, the combination of CNN and DBN offers strong gesture recognition performance. The system's ability to capture dynamic gesture details is improved by novel features like neural gas and locomotion thermal mapping. Fuzzy logic optimization raises the quality of features, which raises classification accuracy. On the Egogesture and Jester datasets, comparative analysis shows that the suggested system performs better than other cutting-edge models.

*Citation:* [4] Mohammed Alonazi; Hira Ansar; Naif Al Mudawi; Saud S. Alotaibi; Nouf Abdullah Almujally; Abdulwahab Alazeb(2023). Smart Healthcare Hand Gesture Recognition Using CNN-Based Detector and Deep Belief Network.

IEEE Access, DOI:[10.1109/ACCESS.2023.3289389]

*E. Paper title: Real-Time Virtual Mouse Using Hand Gestures for Unconventional Environments.Year of publication : 2023*

*Description:* In this paper, a touchless substitute for conventional physical mice is presented: a virtual mouse system that uses hand gestures as an input mechanism. The system uses image processing libraries like OpenCV and MediaPipe to detect and interpret hand gestures, allowing users to execute a variety of mouse actions. It is designed for unusual environments like space stations, underwater exploration, and extreme weather.

*Methodology:* Using hand gestures and Python, OpenCV, MediaPipe, and PyAutoGUI, the system mimics mouse actions such as drag-and-drop, clicks, and cursor movement. Using MediaPipe landmarks, color thresholding, and contour detection, video frames are processed to improve quality and identify hands. A single-shot detector model uses hand posture and movement to identify gestures. PyAutoGUI maps recognized gestures to mouse operations like drag-and-drop and left-click. Real-time performance with low latency across a range of mouse functions was verified through testing.

*Limitations:* Poor or inconsistent lighting can cause performance to suffer, particularly in underwater settings. Accurate gesture recognition may be hampered by complicated or cluttered backgrounds. A preset set of gestures are supported by the system, which restricts personalization and adaptability. For best recognition, users might need to modify their gestures to fit the needs of the system.

*Key Insights:* For special situations, the virtual mouse system provides a cost-effective and user-friendly substitute for traditional input devices. Accurate and effective gesture detection is ensured when OpenCV and MediaPipe are combined. It improves hygiene and accessibility, which makes it useful in settings with little opportunity for face-to-face interaction. The system exhibits potential for enhanced performance in difficult situations and additional gesture customization. Its potential for useful, touchless interaction solutions is demonstrated by this innovation.

*Citation:* [5] Neha Sabrin TK;Aarti Karande(2023). Real-Time Virtual Mouse using Hand Gestures for Unconventional Environment.IEEE Access,

DOI:[ 10.1109/ICCCNT56998.2023.10308331]

F. *Paper title: Low-latency hand gesture recognition with a low resolution thermal image,Year of publication : 2020*

*Description:* The study investigates the use of a low-resolution thermal camera as an affordable method for low-latency hand gesture recognition. The goal is to create an algorithm that can operate with the 32x24 pixel MLX90640 sensor to allow for gesture control in automobiles. The technique maintains high accuracy and low latency while being tailored for inexpensive hardware, like microcontrollers.

*Methodology:* Using sensors at dashboard and ceiling viewpoints, the methodology created a dataset of more than 1300 gesture videos that covered nine gestures and non-gesture actions. In order to improve low-latency performance, a novel "mixed causal" configuration was added to a 1D TCN for temporal modeling and a 2D CNN for spatial feature extraction. Two phases of training were carried out, employing CE loss and CTC loss, with substantial data augmentation to increase robustness. Classification accuracy, mean Average Precision (mAP), and latency were among the evaluation metrics, and they were compared to cutting-edge models such as LSTM and 3D CNNs. The method placed a strong emphasis on high-performance, low-cost recognition for practical uses.

*Limitations:* One of the study's drawbacks is its reliance on particular thermal sensors, which might limit its applicability to different modalities or sensor kinds. Even though the model is optimized for reasonably priced hardware, it might need to be further simplified in order to operate on very low-power microcontrollers. The system's capacity to identify more intricate gestures is constrained by the use of thermal images with low resolution (32 x 24 pixels). Sensitivity to sensor placement and viewpoint is also highlighted by performance differences between dashboard and ceiling-mounted sensors. Last but not least, the method mainly concentrates on thermal imaging, ignoring potential integration opportunities with other modalities.

*Key Insights:* The study shows that low-latency and accurate hand gesture recognition is possible with low-resolution thermal cameras, achieving 95.9% classification accuracy and 83% mAP detection with one-frame latency. Mixed causal convolutions enhance low-latency performance, making the system suitable for real-time use. While higher-quality sensors improve results, even the low-cost MLX90641 performs well, highlighting the approach's scalability and cost-effectiveness. This work emphasizes lightweight solutions for affordable real-time gesture recognition.

*Citation:* [6] Maarten Vandersteegen; Wouter Reusen; Kristof Van Beeck; Toon Goedemé. (2020). Low-latency hand gesture recognition with a low resolution thermal image. IEEE Access, DOI:[ 10.1109/CVPRW50498.2020.00057]

G. *Paper title: Hand Sign and Gesture Recognition System,Year of publication : 2023*

*Description:* In order to improve communication between people with hearing impairments and others, the paper presents a *Hand Sign and Gesture Recognition System, a nonverbal communication tool. The system uses shape-based features such as orientation, centroid, finger state, and thumb position to interpret sign language gestures in real-time. A **Convolutional Neural Network (CNN)* is used to classify gestures and extract features. Because it doesn't require external hardware, this software-focused method is affordable and widely available.

*Methodology:* Using the Haar_Cascade model, the system extracts the Region of Interest (ROI) from video frames, converts RGB ROIs to grayscale, and applies Gaussian blur to enhance them. For training and testing, OpenCV was used to create custom datasets of ASL signs that were resized to 50x50 pixels and annotated. To handle gesture classification, a CNN with dropout layers, max pooling, and ReLU activation was created. In the first step, a two-tier algorithm recognizes letters; in the second step, it clarifies unclear classifications. To attain high accuracy, the model was trained over 40 epochs.

*Limitations:* The system's use in varying lighting conditions is limited because it depends on sufficient lighting and a well-lit background for accurate recognition. It lacks dynamic movement support and is limited to static hand gestures. Furthermore, the lack of pre-existing datasets required the creation of custom datasets, which might have limited generalizability. Although it works well for recognizing letters and numbers, adding the ability to combine gestures into meaningful words is still a work in progress.

*Key Insights:* CNNs are effective at recognizing gestures, as evidenced by the model's remarkable 99.89% test accuracy. Its practical utility is increased by its hardware independence, which guarantees affordability and portability. Integration with video platforms for automated gesture-based text synthesis and use in video conferences to increase accessibility are examples of possible uses. In the future, the system hopes to expand its real-world applicability and facilitate smooth communication by combining recognized gestures into words.

*Citation:* [7] Priyal Raj; Sakshi Pandey; Yuvraj Singh; Monu Singh (2023). Hand Sign and Gesture Recognition System. IEEE Access, DOI:[ 10.1109/IC3I59117.2023.10398116]

H. *Paper Title:Development of a Real Time Vision-Based Hand Gesture Recognition System for Human-Computer Interaction,Year of Publication: 2023*

*Description:* With American Sign Language (ASL) as the dataset and YOLO-v5 as the main algorithm, the paper introduces a vision-based hand gesture recognition system for Human-Computer Interaction (HCI). To illustrate its uses, the system incorporates a bot control feature and translates sign language gestures into text. 700 pictures of ASL gestures taken in various settings make up the dataset.

*Methodology:* The first step in the Dynamic Hand Gesture Recognition system's methodology is dataset preparation, which involved taking 700 pictures of ASL gestures in various backgrounds and lighting scenarios. Three subsets of the dataset—500 for training, 100 for testing, and 100 for validation—were created. The model architecture of the system uses CSP-Darknet53 as the foundation for reliable feature extraction and YOLO-v5 for real-time object detection. The model uses Path Aggregation Network (PANet) and Spatial Pyramid Pooling (SPP) to improve localization and information flow. In order to facilitate smooth and effective control, the hardware components—an Arduino UNO, DC motors, and a motor driver—were manipulated using numerical gestures instead of ASL signs.

*Limitations:* The system has some drawbacks, such as a lower detection accuracy of 88.5% for hazy images, which compromises its dependability in less-than-ideal visual circumstances. Additionally, it has trouble identifying gestures in extremely different lighting or backgrounds, though the sophisticated features of YOLO-v5 help to some extent. Additionally, the system's capacity to adapt to more complicated or varied sign languages may be limited by the 700-image dataset, which could affect its robustness and generalizability in larger applications.

*Key Insights:* The system outperformed other models like Zhang et al. (90%) and Saxena et al. (89.6%) with an astounding overall accuracy of 93%. The YOLO-v5 model showed remarkable efficiency, achieving high performance in real-time gesture recognition with fewer computational resources. The system's versatility was further demonstrated by real-world uses, such as improving communication for the hard of hearing and facilitating bot control. This demonstrates its potential for future incorporation into domains like home automation and robotics, thereby increasing its impact and usability.

*Citation:*[8] Arijit Das;Kaulik Maitra;Shayan Roy;Biswarup Ganguly;Meghna Sengupta;Shreya Biswas(2023). Development of a Real Time Vision-Based Hand Gesture Recognition System for Human-Computer Interaction.IEEE Access, DOI:[10.1109/ASPCON59071.2023.10396583]

I. *Paper Title: Dynamic Hand Gesture Recognition, Year of Publication:2022*

*Description:* This paper focuses on a Dynamic Hand Gesture Recognition (HGR) System that assigns tasks based on real-time gesture interpretation from a webcam using a CNN. The system is intended to improve human-computer interaction, especially for touchless interfaces, gaming, and accessibility applications. With a 99.84% accuracy rate, it is a reliable option for dynamic gesture recognition.

*Methodology:* Three modules make up the methodology's structure. In the first, a CNN model is constructed using Keras and trained on a collection of gesture image data. The second module employs preprocessing methods such as background subtraction, Gaussian blur, and grayscale conversion to predict gestures in real-time from live video feeds using OpenCV. In order to control devices such as keyboards and mice, the last module uses the PyAutoGUI library to assign tasks to recognized gestures.

*Limitations:* The computational demands of real-time processing and difficulties with background segmentation are among the system's limitations. Even though the system manages dynamic gestures well, there is still room for improvement in terms of integrating more gestures and making it more resilient in a variety of settings.

*Key Insights:* Important findings demonstrate how the system can do away with hardware dependencies and provide an affordable substitute for gadgets like Kinect. It is versatile for a range of applications, such as robotics, surveillance, and accessibility solutions, thanks to its high accuracy and modular design. Expanding the gesture library, improving resilience, and integrating with IoT devices for wider applications are possible future developments.

*Citation:* [9] Subash Chandra Bose Jaganathan; Kesavan R; Thevaprakash P; Krishna Basak; Shinjan Verma; Anisha Mital (2022). Dynamic Hand Gesture Recognition. IEEE Access, DOI:[ 10.1109/iSSSC56467.2022.10051259]

J. *Paper Title: Hand Gesture Recognition in Low-Intensity Environment Using Depth Images, Year of Publication:2017*

*Description:* The suggested vision-based system makes use of depth images captured by devices such as the Microsoft Kinect. It can identify gestures with varying numbers of raised fingers, ranging from 0 to 5. Appliances and other devices can be controlled with gestures, which work flawlessly in a variety of lighting situations. The processing has a recognition accuracy of more than 90% and is tuned for real-time performance.

*Methodology:* Using thresholding and depth images, the hand gesture recognition system separates the hand. OpenCV is used to detect edges, and the hand is identified as the largest contour. Finger positions are determined by analyzing the convex hull and convexity defects. These characteristics are combined with the aspect ratio of the hand to recognize and display gestures that represent the numbers 0–5.

*Limitations:* The system can only identify gestures that match fingers 0–5. The availability of depth cameras like Kinect is crucial. For robustness, more testing with various hand shapes and in various environments might be necessary.

*Key Insights:* Depth-based segmentation enables the system to function well in both dark and well-lit conditions. A simple and hands-free way to interact with computers and appliances is through gesture recognition. The methodology is useful for real-world applications because it guarantees high accuracy and real-time responsiveness.

Citation: [10] Dinesh Kumar Vishwakarma; Varun Grover(2017). Hand Gesture Recognition in Low-Intensity Environment Using Depth Images. IEEE Access, DOI:[ 10.1109/ISS1.2017.8389446]

*K. Paper Title: AHD: Thermal-Image Based Adaptive Hand Detection for Enhanced Tracking System, Year of Publication:2018*

*Description:* This work introduces a new technique for tracking and detecting hands using thermal images, which detect heat emitted by objects such as human hands and are not impacted by lighting. The Guidance Framework for Tracking by Detection (GFTD) is a tracking framework that incorporates the suggested Adaptive Hand Detection (AHD) technique. The tracking accuracy is significantly improved by this method.

*Methodology:* For precise shape and location detection, the system models hand temperature using Adaptive Hand Detection (AHD) and produces a binary image. For reliable tracking, this directs the Guidance Framework for Tracking by Detection (GFTD), which combines AHD with trackers such as KCF. It demonstrated high accuracy using IOU and OTE metrics when tested in a variety of lighting conditions.

*Limitations:* Consistent detection is difficult because the hand's heat can fluctuate. Although useful for speed, thermal images' basic features might not make the most of more sophisticated trackers. The quality of the initial temperature models determines the detection and tracking accuracy.

*Key Insights:* In hand detection, the AHD algorithm had an 86.8% success rate. When combined with the GFTD framework, tracking accuracy increased by 15% in IOU and 16.3% in OTE, with KCF showing the best results. Compared to image fusion techniques, thermal imaging was less computationally demanding and demonstrated resilience to changes in illumination. By decreasing errors and permitting automatic updates while tracking, GFTD further improved tracker performance.

*Citations:* [11] Eungyeol Song; Hyeongmin Lee; Jaesung Choi; Sangyoun Lee(2018). AHD: Thermal-Image Based Adaptive Hand Detection for Enhanced Tracking System.IEEE Access, DOI:[ 10.1109/ACCESS.2018.2810951]

*L. Paper Title: A Review of the Hand Gesture Recognition System: Current Progress and Future Directions, Year of Publications:2021*

*Description:* The study examines developments in hand gesture recognition technology, with a focus on vision-based techniques. It investigates camera-based gesture recognition methods and contrasts them with device-based methods. Vision-based systems are appropriate for applications in communication, education, and rehabilitation because they provide a more natural interaction without the need for wearable technology.

*Methodology:* The review analyzes data acquisition, gesture representation, and recognition accuracy by combining results from 98 studies carried out between 2014 and 2020. Along with feature extraction techniques like CNNs and HOG, it highlights a variety of data capture methods, such as single cameras and active techniques like Kinect. Signer-dependent and independent systems were examined in experimental evaluations.

*Limitations:* Unsustainable environments, like changing lighting and complex backgrounds, present difficulties for vision-based hand gesture recognition. Furthermore, it is still very difficult to discern between continuous gestures and the transitions between them. The majority of current research concentrates on signer-dependent systems, which restricts its applicability to larger populations.

*Key Insights:* Signer-dependent systems have an average recognition accuracy of 88.8%, whereas signer-independent systems have an average of 78.2%. While vision-based systems are excellent at natural interaction, they need more reliable solutions for continuous and dynamic gestures. Larger multilingual datasets, sophisticated 3D recognition, and real-world application systems are the main focus of future research.

*Citations:* [12] Noraini Mohamed; Mumtaz Begum Mustafa; Nazean Jomhari (2021). *A Review of the Hand Gesture Recognition System: Current Progress and Future Directions.* IEEE Access, *DOI:[10.1109/ACCESS.2021.3129650].*

*M. Paper Title: Live Demonstration: Home Appliance Control System with Dynamic Hand Gesture Recognition base on 3D Hand Skeletons,Year of Publications:2022*

*Description:* The system employs a lightweight CNN with DetNet to detect hands and create 3D hand skeletons, processing them with a dynamic gesture recognition model. Using cost-effective RGB cameras, it achieves 99.4% accuracy on the test dataset. Implemented on an Nvidia Jetson AGX Xavier, it demonstrates appliance control like fans and lights.

*Methodology:* T3D hand skeletons created from RGB images are processed by the system. It avoids the complexity of RNNs or LSTMs by using a one-dimensional CNN model to analyze temporal features of skeleton movements. There are a total of 1215 gestures in the dataset, which consists of 200 clips per movement for six different gesture types. At a processing speed of 15 frames per second (fps), the recognition model generates control signals for appliances.

*Limitations:* The system's reliance on a preset set of gestures restricts the range of possible interactions. Because it uses RGB cameras, its performance may change in difficult lighting or background situations. Even with optimization, the hardware specifications still rely on specialized equipment such as the Nvidia Jetson AGX Xavier.

*Key Insights:* The study shows that inexpensive gesture recognition systems utilizing RGB cameras and thin neural networks are feasible. It is a promising option for non-contact human-computer interaction because of its high accuracy and real-time performance, particularly in smart home applications.

*Citations:* [13] Tsung-Han Tsai; Yi-Jhen Luo; Wei-Chung Wan(2022). Live Demonstration: Home Appliance Control System with Dynamic Hand Gesture Recognition base on 3D Hand Skeletons. IEEE Access, DOI:[ 10.1109/AICAS54282.2022.9870006]

*N. Paper Title: Hand Gesture Recognition System as an Alternative Interface for Remote Controlled Home Appliances, Year of Publications:2018*

*Description:* An alternate interface for managing household appliances is presented in the paper: a hand gesture recognition system. Static hand gestures are detected by an Android application with a camera, processed, and compared to a gesture database. Recognized gestures provide an easy-to-use and affordable remote-control solution for controlling lights, fans, and other appliances through infrared signals. *Methodology:* Using the built-in camera on the Android device, the system records hand gestures, processes them with OpenCV, and compares them to preset gesture templates stored in a database. Corresponding infrared commands are sent to the appliances upon recognition. During testing, environmental factors like lighting and infrared range were assessed along with recognition accuracy and response times.

*Limitations:* Lighting conditions have an impact on the system's performance; accurate detection requires at least 120 lux. For accurate recognition, the hand must contrast with the background. Usability in some setups is limited by the infrared communication's requirement for a direct line of sight and its six-meter maximum range. *Key Insights:* The Samsung Galaxy S4 displayed a 100% recognition success rate, demonstrating the system's high accuracy. Efficiency was demonstrated by response times of less than three seconds. Enhancing dynamic gesture support, increasing lighting adaptability, and extending gesture functionality for more comprehensive appliance control are some of the recommendations.

*Citations:* [14] Marvin S. Verdadero; Celeste O. Martinez-Ojeda; Jennifer C. Dela Cruz(2018). Hand Gesture Recognition System as an Alternative Interface for Remote Controlled Home Appliances. IEEE Access, DOI:[10.1109/HNICEM.2018.8666291]

*O. Paper Title: A Continuous Real-time Hand Gesture Recognition Method based on Skeleton, Year of Publications:2022*

*Description:* The study suggests a continuous, real-time hand gesture recognition system that uses skeleton data to improve computational efficiency and accuracy. It detects gestures and divides them into pointing and non-pointing categories using the TD-Net architecture. The system is made to overcome the difficulties in identifying and identifying gestures in continuous streams by processing skeleton sequences that are taken from RGB images.

*Methodology:* Three TD-Net models are used for gesture detection and classification, and Mediapipe is used to extract hand skeletons from RGB images. TD-Net processes skeleton sequences to detect gestures in real-time, differentiating between non-pointing, pointing, and no gesture actions. The IPN dataset was used for evaluations, which showed that the system could identify 13 different gestures with low latency and high efficiency.

*Limitations:* The technique has trouble telling what is happening and what isn't, especially when hands are moving outside of the camera frame. Short-duration or perpendicular gestures to the screen result in a decrease in recognition accuracy. Furthermore, although skeleton-based data is resistant to changes in lighting and background, it lacks the comprehensive context that RGB-based techniques offer, which restricts the ability to recognize subtle gestures.

*Key Insights:* The system outperforms RGB-based techniques in terms of speed, achieving 40.1% Levenshtein accuracy for continuous gesture recognition with an inference time of 0.1ms. Skeletal data is used to guarantee robustness against background clutter and lighting. End-to-end models may be used in future developments to increase the accuracy of gesture spotting and recognition.

*Citations:* [15] Tien-Thanh Nguyen; Nam-Cuong Nguyen; Duy-Khanh Ngo; Viet-Lam Phan; Minh-Hung Pham; Duc-An Nguyen(2022). A Continuous Real-time Hand Gesture Recognition Method based on Skeleton. IEEE Access,DOI:[10.1109/ICCAIS56082.2022.9990122]

*P. Paper Title: Real-Time Hand Gesture Recognition and Sentence generation for Deaf and Non-Verbal Communication, Year of Publications:2024*

*Description:* In order to help deaf and non-verbal people communicate, the paper presents a system that can recognize hand gestures in real time and generate sentences. The system recognizes gestures and transforms them into grammatically correct sentences using deep learning techniques such as CNNs and RNNs. It encourages inclusivity for non-verbal communities by supporting a broad variety of gestures and facilitating efficient communication through generated text or speech.

*Methodology:* The system uses a CNN that has been trained on a large dataset of hand gestures to classify them in real time. In order to produce coherent and contextually relevant sentences, recognized gestures are subsequently processed using natural language processing techniques, such as sequence-to-sequence models. Metrics like accuracy, fluency, and usability are used to assess the system, showing how it can improve communication situations in the real world.

*Limitations:* Due to the intricacies of hand movement and fluctuating lighting, the system has trouble accurately detecting gestures. The range of gestures that can be recognized is limited due to its limited dataset. Furthermore, accessibility and adaptability for a variety of users may be hampered by the hardware specifications and lack of user customization options.

*Key Insights:* The reliability of the suggested system is demonstrated by its high performance, which includes an accuracy of 99.1% and an F1-score of 98.56%. It improves autonomy for non-verbal people and fills communication gaps by combining gesture recognition with sentence generation. Expanding the gesture dataset, enhancing adaptability to various environments, and adding more dynamic customization features are possible future improvements.

*Citations:* [16] K Lavanya Manikanta; N Shyam; Saraswathi S(2024). Real-Time Hand Gesture Recognition and Sentence generation for Deaf and Non-Verbal Communication. IEEE Access, DOI:[10.1109/ADICS58448.2024.10533502]

*Q. Paper Title: An Ultra-Low-power Real-Time Hand-Gesture Recognition System for Edge Applications, Year of Publications:2021*

*Description:* With a sequence analyzer for dynamic gestures and a shallow decision tree and Edge-CNN for static gestures, this paper presents an ultra-low-power real-time hand gesture recognition system for edge applications. It lowers memory and computation requirements by concentrating on hand edge data and low-resolution binary images. The system uses only 184 µW of power and achieves 92.6% accuracy for 24 dynamic gestures.

*Methodology:* To produce binary images, the system preprocesses input images using morphological operations, filtering, segmentation, and subsampling. A sequence analyzer with majority voting detects dynamic gestures, and a decision tree and Edge-CNN are used to extract features. A 65nm TSMC chip is used to implement the design, and an FPGA setup is used to test its performance in real time.

*Limitations:* Flexibility for wider use cases is limited by the system's recognition of gestures within particular hand movement speeds. Its dependence on motion detection makes it difficult to use for subtle or stationary gestures. The binary preprocessing method's FPGA-based configuration limits its platform versatility, and it may not work well in complex backgrounds or with changing lighting.

*Key Insights*: The system's 184 µW power consumption makes it perfect for wearable and energy-efficient edge devices. Without sacrificing accuracy, the hybrid Edge-CNN and decision tree design lowers memory and processing demands. Its modular design improves stability and noise resistance while demonstrating the possibility of touchless control and real-time human-robot interaction.

*Citations:* [17] Yuncheng Lu; Zehao Li; Tony Tae-Hyoung Kim(2021). An Ultra-Low-power Real-Time Hand-Gesture Recognition System for Edge Applications. IEEE Access, DOI:[10.1109/AICAS51828.2021.9458436]

*R. Paper Title: Research on the Hand Gesture Recognition Based on Deep Learning, Year of Publications:2018*

*Description:* The deep learning-based hand gesture recognition system that can track, segment, and identify hand gestures in real time is the main topic of this paper. It separates the hand from complicated backgrounds by combining the AdaBoost classifier with Haar features with a Gaussian mixture model for skin color segmentation. The LeNet-5 convolutional neural network is used to recognize 10 digits (1–10), and the CamShift algorithm is used for gesture tracking. The system's accuracy is an impressive 98.3%.

*Methodology:* Tracking, recognition, and segmentation are the three main processes in the system. AdaBoost classifiers are used for accurate localization in segmentation, while Gaussian mixture models are used for initial hand detection. The CamShift algorithm uses motion and distortion to track movements in real time. Lastly, a dataset of 20,000 photos is used for training and testing the LeNet-5 CNN for classification, ensuring reliable performance in a range of interior situations.

*Limitations:* Because the system relies on skin color segmentation, it may produce false positives because it is sensitive to objects with similar hues, such faces or arms. Because the technique has 3D modeling capabilities, its accuracy declines for hand rotations or gestures with complicated shapes. Furthermore, the method's scalability for larger applications or dynamic motions is limited because it can only recognize ten predetermined gestures.

*Key Insights:* This work shows how to recognize hand gestures in a practical and effective way by combining deep learning and conventional methods to achieve high accuracy. LeNet-5 guarantees accurate classification in challenging situations, while CamShift improves real-time tracking. The system presents a viable option for robotics and human-computer interaction applications, despite its limitations in handling 3D motions and scalability.

*Citations:* [18] Jing-Hao Sun; Ting-Ting Ji; Shu-Bin Zhang; Jia-Kui Yang; Guang-Rong Ji(2018). Research on the Hand Gesture Recognition Based on Deep Learning. IEEE Access, DOI:[10.1109/ISAPE.2018.8634348]

*S. Paper Title: An Approach for Human Machine Interaction using Dynamic Hand Gesture Recognition, Year of Publications:2019*

*Description:* Using cost-effective webcams for real-time applications, this paper presents a dynamic hand gesture recognition system for human-machine interaction (HMI). It uses region of interest (ROI) and contour analysis to identify gestures and combines RGB and HSV color spaces for skin color segmentation. Slide navigation and other user-friendly HMI applications are made possible by the system's 95% accuracy in recognizing eight predefined dynamic gestures.

*Methodology:* Video frames are preprocessed by the system using morphological operations to improve binary images, median filtering for noise reduction, and skin color segmentation. ROI extraction and contour analysis are used to detect hands, and then convex hull defects are used to count fingers. In order to classify eight dynamic gestures, gesture recognition entails tracking the centroid of the hand and examining motion trajectories.

*Limitations:* The accuracy of segmentation is decreased by the system's difficulties with complex backgrounds, but it performs well in controlled lighting and with a range of skin tones. Because of its lack of rotational invariance, it can't recognize gestures with different orientations. Furthermore, the framework's use in more extensive interaction scenarios is limited because it can only recognize predefined gestures.

*Key Insights:* The framework offers a reliable, affordable, and highly accurate solution for dynamic hand gesture recognition with low hardware requirements. Its lightweight design guarantees effective performance through the use of contour analysis and median filtering. It has potential for HMI applications, despite being restricted to predefined gestures and basic environments; future developments will concentrate on complex backgrounds and rotational invariance.

*Citations*: [19] Mohd Aquib Ansari; Dushyant Kumar Singh(2019). An Approach for Human Machine Interaction using Dynamic Hand Gesture Recognition. IEEE Access, DOI:[ 10.1109/CICT48419.2019.9066173]

## III. METHODOLOGY

Data collection is the first of several crucial steps in the methodology for hand gesture recognition (HGR) systems. The majority of input methods are vision-based, using infrared cameras, RGB cameras, and depth sensors to record both static and dynamic gestures. The cost-effectiveness of single cameras makes them popular, but active devices like the Leap Motion Controller and depth sensors like Microsoft Kinect improve spatial resolution. In low-light conditions, thermal imaging is used to record the heat signatures of hands, and skeleton-based data extraction—which is frequently made possible by programs like Mediapipe—is used to separate joint coordinates in order to effectively depict hand motion. Techniques for preprocessing are essential for improving the obtained data. While segmentation techniques isolate the Region of Interest (ROI) using methods like skin color segmentation in HSV color space, noise reduction techniques like Gaussian blur are used to reduce artifacts. In order to accommodate differences in hand size, position, and orientation, normalization further standardizes skeleton data.

An essential step in ensuring precise classification is feature extraction. While image-based approaches use descriptors like Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), and Convolutional Neural Networks (CNNs) for spatial feature extraction, skeleton-based approaches rely on features like Joint Collection Distance (JCD) and Normalized Coordinates of Joints (NCJ). Using sophisticated algorithms designed to meet particular requirements, classification and recognition are accomplished. While temporal models like TD-Net use spatio-temporal features to handle dynamic gestures in continuous streams, CNNs are frequently used for image-based recognition because of their powerful spatial representation capabilities. For smaller datasets in controlled settings, Support Vector Machines (SVMs) are preferred, and hybrid models combine features from images and skeletons to increase accuracy.

System performance must be evaluated and benchmarked using datasets such as the IPN for continuous gesture recognition. To guarantee suitability for real-time applications, standard metrics like accuracy, F1-score, and Levenshtein distance are used in conjunction with measurements of inference speed. For example, TD-Net is very effective for real-world deployment, achieving an impressive 0.1 ms inference time. Adaptability to a variety of situations is ensured by robustness testing against changing lighting, background complexity, and occlusion.

Lastly, deployment entails modifying lightweight models to formats like TensorFlow Lite and optimizing them for particular hardware platforms, such as edge devices like the Raspberry Pi and Nvidia Jetson. For extremely efficient HGR solutions in energy-constrained settings, some systems also use custom hardware. This thorough approach demonstrates how cutting-edge sensing technologies, reliable feature extraction, and effective algorithms work together to meet the demands and difficulties of contemporary HGR systems.

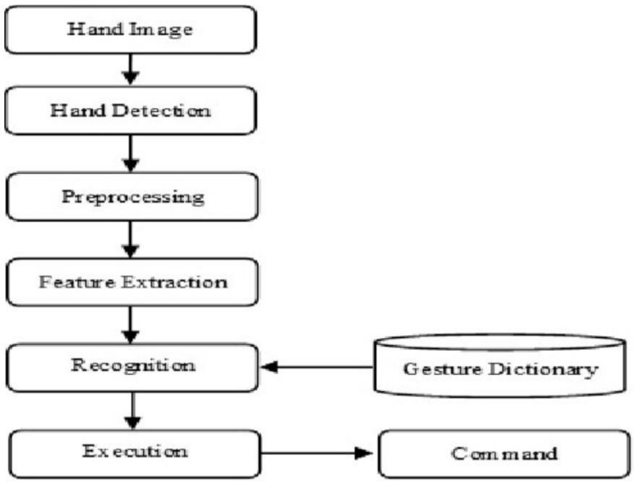


Fig. 1. Methodology

## IV. ALGORITHMS USED

### A. *Support Vector Machine (SVM)*

Classification tasks, including binary and multi-class problems, are the main application for SVM, a reliable supervised learning algorithm. It works by determining which hyperplane best divides data points into discrete groups. SVM works best in datasets with a distinct margin of separation and in high-dimensional spaces. SVM can also handle non-linear relationships in the data by using kernel functions like the radial basis function (RBF), linear, and polynomial. It is a popular option for applications such as text classification, bioinformatics, and image recognition because of its exceptional generalization ability.

1. Normalize the data set
2. For Each *C*, $\gamma$:
   - 2.1. Cross Validate using leave-one-out.
     - 2.1.1. Train and test the SVM.
     - 2.1.2. Store the success rate.
   - 2.2. Compute the average success rate.
   - 2.3. Update the best *C* and $\gamma$ if needed.
   - 2.4. Return to 2.1 with next *C*, $\gamma$.
3. Choose *C*, $\gamma$ with best average success rate, and perform step (2) using fine scale around the selected parameters.

Fig. 2. SVM Algorithm

### B. Convolutional Neural Networks (CNN)

With a primary focus on image and video processing, Convolutional Neural Networks (CNNs) are sophisticated deep learning models designed for structured data analysis. By using fully connected layers for decision-making, pooling layers to lower computational complexity, and convolutional layers to extract features, they mimic the human visual system by spotting hierarchical patterns. CNNs excel in a variety of applications, including object detection, medical diagnostics, autonomous vehicles, and facial recognition, because of their ability to detect edges, textures, shapes, and objects. Their capacity to automatically extract features from unprocessed data, eliminating the need for feature engineering, has transformed computer vision and is propelling advancements in artificial intelligence.

```
for iteration = 1 to Maximum iteration do
    for i = 1 to N do
        while f(k) ⇐ f(s) do
            for all k ε - neighbours(s) do
                Generates an k ⇐ ε - neighbour(s)
                if fitness(k) > fitness(s) then
                    Replace k with s;
                end if
            end for
        end while
    end for
end for
```

Fig. 3. CNN Algorithm

### C. LSTM, or long short-term memory

The temporal component is essential when LSTM networks are used to recognize dynamic hand gestures. In order to recognize gestures that span multiple frames in a video, LSTMs—a type of Recurrent Neural Network (RNN)—are made to handle long-term dependencies in sequence data. In contrast to conventional RNNs, LSTMs are very good at extracting patterns from sequences of different lengths because they control the information flow across time steps using memory cells and gates (input, forget, and output gates). In order to improve performance in continuous gesture recognition tasks, LSTMs are frequently used in hand gesture recognition to process time-series data (such as hand movement sequences) and comprehend the relationships between successive gestures.

### D. Triple Feature Double Motion, or TD-Net

For continuous gestures in particular, the TD-Net architecture was created for skeleton-based gesture recognition. Joint Collection Distances (JCD), Normalized Coordinates of Joints (NCJ), and motion features are the three primary feature types used by the network. While NCJ concentrates on the relative positions of joints, especially the thumb, to represent hand shapes, JCD computes the Euclidean distances between joints to capture spatial relationships. Using both slow and fast motion metrics to distinguish between different gesture speeds, the motion features capture the dynamics of hand gestures. With these features, TD-Net's light weight architecture enables accurate and efficient real-time gesture recognition. Continuous gesture recognition, where identifying the beginning and ending of gestures, has demonstrated the effectiveness of the architecture.

### E. PCA, or principal component analysis

High-dimensional data can be made simpler while maintaining as much variance as possible by using PCA, a dimensionality reduction technique. By reducing the number of features extracted from skeleton data or images, PCA improves the efficiency of the classification process in hand gesture recognition. PCA preserves the most significant features (those with the highest variance) and eliminates less pertinent information by converting the original data into a set of orthogonal (uncorrelated) principal components. This lessens the possibility of overfitting and facilitates data processing. To increase gesture classification accuracy while preserving computational efficiency, PCA is sometimes used in conjunction with other feature extraction techniques, such as Histogram of Oriented Gradients (HOG), in hand gesture recognition systems.

## V. RESULTS

A number of studies offer creative approaches to gesture identification and detection. The TD-Net architecture, which is based on a skeleton-based methodology and performs exceptionally well in both isolated and continuous hand gesture recognition, is introduced in the first paper. To increase the precision and effectiveness of hand gesture recognition, it makes use of Joint Collection Distances (JCD), Normalized Coordinates of Joints (NCJ), and motion analysis (both slow and fast motion). With a Levenshtein accuracy of 40.10% and an inference time of 0.1 ms for continuous gestures, the TD-Net performs better than other models in terms of both computational speed and recognition accuracy.

When compared to models such as ResNet and ResNext, the continuous hand gesture recognition techniques based on skeleton data perform more accurately than those utilizing RGB and optical flow images. The suggested technique outperforms other models such as ST-GCN and 2s-AGCN for isolated hand gesture recognition, achieving an accuracy of 84.98% and an F1-score of 79.01%. Gesture spotting, where the model must determine the start and finish points of a gesture—a task more difficult than identifying isolated gestures—is one of the main obstacles in continuous gesture recognition. A number of approaches offer solutions, including the Levenshtein distance metric to enhance performance and Media pipe for hand pose estimation.

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 01_palm | 1.00 | 1.00 | 1.00 | 177 |
| 02_l | 1.00 | 1.00 | 1.00 | 216 |
| 03_fist | 1.00 | 1.00 | 1.00 | 187 |
| 04_fist_moved | 1.00 | 1.00 | 1.00 | 215 |
| 05_thumb | 1.00 | 1.00 | 1.00 | 231 |
| 06_index | 1.00 | 1.00 | 1.00 | 189 |
| 07_ok | 1.00 | 1.00 | 1.00 | 192 |
| 08_palm_moved | 1.00 | 1.00 | 1.00 | 213 |
| 09_c | 1.00 | 1.00 | 1.00 | 177 |
| 10_down | 1.00 | 1.00 | 1.00 | 203 |
| accuracy | | | 1.00 | 2000 |
| macro avg | 1.00 | 1.00 | 1.00 | 2000 |
| weighted avg | 1.00 | 1.00 | 1.00 | 2000 |

Fig. 4. Classification Report

Despite their effectiveness, current techniques still have drawbacks, particularly when it comes to managing continuous, dynamic gestures in uncontrolled environments. The papers recommend better handling of gesture spotting and coarticulation effects for sign language recognition, as well as enhancements to feature extraction methods (e.g., PCA, HOG, and CNNs). In the future, a more complete approach to dynamic sign language recognition in practical settings might be offered by combining deep learning models with real-time gesture spotting.

## References


[1] Robust Hand Gestures Recognition using a Deep CNN and Thermal Images(2021). Authors:Daniel Skomedal Breland,

[5] Real-Time Virtual Mouse using Hand Gestures for Unconventional Environment(2023).Authors:Neha Sabrin TK;Aarti Karande.

[6] Low-latency hand gesture recognition with a low resolution thermal image(2020).Authors:Maarten Vandersteegen; Wouter Reusen; Kristof Van Beeck; Toon Goedemé.

[7] Hand Sign and Gesture Recognition System(2023).Authors:Priyal Raj; Sakshi Pandey; Yuvraj Singh; Monu Singh.

[8] Real Time Vision-Based Hand Gesture Recognition System for Human-Computer Interaction(2023). Authors: Arijit Das;Kaulik Maitra;Shayan Roy;Biswarup Ganguly;Meghna Sengupta;Shreya Biswas.

[9] Dynamic Hand Gesture Recognition(2022). Authors:Subash Chandra Bose Jaganathan; Kesavan R; Thevaprakash P; Krishna Basak; Shinjan Verma; Anisha Mital.

[10] Hand Gesture Recognition in Low-Intensity Environment Using Depth Images(2017).Authors:Dinesh Kumar Vishwakarma; Varun Grover

[11] Thermal-Image Based Adaptive Hand Detection for Enhanced Tracking System(2018). Authors:Eungyeol Song; Hyeongmin Lee; Jaesung Choi; Sangyoun Lee.

[12] A Review of the Hand Gesture Recognition System(2021). Authors:Noraini Mohamed; Mumtaz Begum Mustafa; Nazean Jomhari.

[13] Live Demonstration: Home Appliance Control System with Dynamic Hand Gesture Recognition base on 3D Hand Skeletons(2022).Authors:Tsung-Han Tsai; Yi-Jhen Luo; Wei-Chung Wan.

[14] Hand Gesture Recognition System as an Alternative Interface for Remote Controlled Home Appliances(2018).Authors:Marvin S. Verdadero; Celeste O. Martinez-Ojeda; Jennifer C. Dela Cruz.

[15] A Continuous Real-time Hand Gesture Recognition Method based on Skeleton(2022). Authors:Tien-Thanh Nguyen; Nam-Cuong Nguyen; Duy-Khanh Ngo; Viet-Lam Phan; Minh-Hung Pham; Duc-An Nguyen.

[16] Real-Time Hand Gesture Recognition and Sentence generation for Deaf and Non-Verbal Communication(2024). Authors:K Lavanya Manikanta; N Shyam; Saraswathi S.

[17] An Ultra-Low-power Real-Time Hand-Gesture Recognition System for Edge Applications(2021) Authors:Yuncheng Lu; Zehao Li; Tony Tae-Hyoung Kim.

[18] Research on the Hand Gesture Recognition Based on Deep Learning (2018).Authors:Jing-Hao Sun; Ting-Ting Ji; Shu-Bin Zhang; Jia-Kui Yang; Guang-Rong Ji.

[19] An Approach for Human Machine Interaction using Dynamic Hand Gesture Recognition(2019).Authors:Mohd Aquib Ansari; Dushyant Kumar Singh.

Aveen Dayal , Ajit Jha , Phaneendra K. Yalavarthy , Om Jee Pandey and Linga Reddy Cenkeramaddi.

[2] Deep Learning-Based Sign Language Digits Recognition From Thermal Images With Edge Computing System(2021). Authors:Daniel S. Breland, Simen B. Skriubakken, Aveen Dayal, Ajit Jha, Phaneendra K. Yalavarthy and Linga Reddy Cenkeramaddi.

[3] Hand Sign Recognition using Infrared Imagery Provided by Leap Motion Controller and Computer Vision (2021). Authors:Tathagat Banerjee; K.V. Pavan Srikar; S. Ashvith Reddy; Krishna Sai Biradar; Rithika Reddy Koripally; Gummadi. Varshith.

[4] Smart Healthcare Hand Gesture Recognition Using CNN-Based Detector and Deep Belief Network(2023).Authors:Mohammed Alonazi; Hira Ansar; Naif Al Mudawi; Saud S. Alotaibi; Nouf Abdullah Almujally; Abdulwahab Alazeb.