A Literature Survey on Hand Gesture Recognition

Krishna Swaroop A, Sudarshan Shweth Nazare, Suhas S K, Suman S Y, Uma C M
Malnad College of Engineering,Hassan

---***---

Abstract - Thermal imaging presents a promising approach for hand gesture recognition, offering enhanced robustness compared to traditional RGB camera-based systems in challenging environments. Unlike RGB camera-based solutions, thermal imaging offers robustness in diverse environments, particularly under varying lighting conditions. The study reviews advanced tools and techniques, including CNN-based feature extraction, real-time gesture mapping, and efficient model training methodologies. Key findings highlight the potential of thermal imaging in achieving precise and intuitive control of media functionalities like play, pause, and volume adjustments. The research emphasizes resource efficient designs for seamless operation on devices with constrained computational resources. This work provides a foundation for developing adaptive, reliable, and user-friendly gesture-controlled systems.

Key Words: Thermal Imaging, Gesture Recognition, Media Control, Deep Learning, Real-Time Systems, Human-Computer Interaction

1.INTRODUCTION

Hand gesture recognition is an important research domain with transformative potential for human-computer interaction. Conventional approaches have extensively utilized computer vision for applications such as human-robot collaboration, but they face challenges in dynamic and uncontrolled environments[1]. Integrating machine vision with other modalities, such as electromyographic signals, has proven effective in enhancing recognition accuracy and robustness across diverse scenarios[2]. The availability of diverse and high-quality datasets plays a crucial role in advancing gesture recognition systems, as standardized datasets are essential for training and benchmarking models[3]. In healthcare, hand gesture recognition powered by deep learning techniques like CNNs and Deep Belief Networks has facilitated efficient interaction in complex, real-world settings[4]. To address computational limitations, researchers have explored ensemble methods and machine learning algorithms, assessing their performance across different tasks[5].

Real-time gesture tracking has also been implemented for unconventional use cases, such as virtual environments, to ensure seamless user interaction[6]. Thermal imaging has emerged as a robust alternative to traditional RGB-based systems, as it is less susceptible to issues caused by lighting variations and background noise[7]. The fusion of thermal imaging with millimeter wave data has shown significant promise for achieving reliable hand gesture recognition in challenging conditions[8]. Efforts toward resource-efficient designs have led to the development of systems using low-resolution thermal cameras. These approaches leverage advanced methods, such as spiking neural networks and sparse segmentation, to reduce computational requirements without compromising accuracy[9].

Deep CNNs integrated with thermal imaging have further demonstrated the potential for robust and precise gesture recognition under varying conditions[10]. Applications of these advancements extend to controlling media players, smart healthcare, and other interactive technologies, underscoring the versatility of thermal-based gesture recognition.

This literature survey examines the progress made in hand gesture recognition systems with a focus on the use of thermal imaging. It highlights the benefits of multimodal approaches and resource-efficient models, particularly for real-time applications. By analyzing insights from recent studies, this work aims to uncover challenges and opportunities for improving the design and implementation of thermal-based hand gesture recognition systems.

2. Body of Paper

2.1 About Hand Gesture

Hand gestures are a fundamental mode of nonverbal communication, widely studied for their potential in human-computer interaction. They involve movements or static postures of the hands and fingers to convey instructions, emotions, or information. In technological applications, hand gestures are increasingly used for controlling devices, interacting with virtual and augmented reality systems, and aiding accessibility for individuals with disabilities. Gesture recognition systems rely on capturing hand movements using various sensors, such as cameras, thermal imaging devices, or electromyographic signals. These systems utilize machine learning and deep learning algorithms to process and classify gestures into meaningful actions. Despite their potential, challenges persist in achieving robust recognition under varying lighting, background conditions, and user diversity. Recent advancements, such as integrating thermal imaging and multimodal approaches, have shown promise in addressing these limitations. With applications ranging from smart Authorized licensed use limited to environments to healthcare and robotics, hand gesture recognition continues to be a dynamic and evolving field of research.

2.2 Related Work

Hand gesture recognition systems often face significant challenges in maintaining performance under varying lighting conditions, which is a primary limitation of RGB-based imaging methods. These systems are particularly prone to errors in environments with shadows, excessive brightness, or insufficient illumination, reducing their effectiveness in practical applications[1]. As a solution, researchers have investigated thermal imaging, which offers greater reliability as it is unaffected by lighting changes. However, integrating thermal imaging into recognition systems introduces complexities in data processing and feature extraction[7]. Hybrid approaches, such as combining thermal imaging with millimeter wave sensors, have shown enhanced robustness in challenging conditions but come with increased computational demands[7]. Furthermore, machine learning models used in gesture recognition require extensive training to handle diverse lighting scenarios, which significantly increases resource requirements and development time[2]. These limitations emphasize the importance of developing innovative, efficient, and adaptive methods to improve gesture recognition systems for use in dynamic lighting environments.

A. Advancements in Hand Gesture Recognition for Human Robot Interaction

Hand gesture recognition plays a vital role in human-robot interaction by enabling intuitive and contactless communication between humans and machines. Vision-based methods are widely adopted due to their straightforward setup, relying solely on visual input to interpret gestures, making them applicable to diverse fields such as smart homes and robotic systems. Advances in deep learning have significantly enhanced the precision and speed of these systems, with some models now capable of achieving exceptional accuracy in recognizing both static and dynamic gestures. Techniques like convolutional neural networks, including MobileNetV2, are commonly employed to process hand gestures in real-time and classify them efficiently. Machine learning methods, such as Support Vector Machines and Hidden Markov Models, have further improved gesture interpretation by providing robust classification capabilities.

Personalized systems that adapt to individual users and new gestures are emerging, demonstrating the potential to enhance humanoid robot interactions by tailoring responses to specific needs. For applications requiring immediate feedback, such as robotic operations or industrial tasks, real-time recognition remains a critical component. Depth sensor cameras and 2D landmark extraction are frequently utilized to capture intricate hand features for more precise classification. By integrating artificial intelligence, computer vision, and robotics, researchers are driving innovation in human-computer interaction, unlocking new possibilities for robotic applications. Ongoing advancements in this field focus on improving the reliability and flexibility of gesture recognition systems, ensuring their effectiveness in diverse and dynamic scenarios.

B. Machine Vision and EMG-Based Hand Gesture Recognition

This study investigates the integration of electromyography (EMG) signals with machine vision to enhance the performance of hand gesture recognition systems. Such an approach is particularly valuable for applications like hand rehabilitation and prosthetics, where accurate tracking of hand and wrist movements is essential. The system captures hand gestures using a camera, while open-source libraries process data to assess hand status and wrist motion. Machine learning algorithms are applied to classify gestures based on the collected EMG signals. By combining EMG with machine vision, the system provides a reliable representation of the user's motor intentions, especially for everyday tasks. Initial testing showed strong results, with the model achieving an accuracy of 81.70% for predicting hand status and 72.05% for real-time gesture sequences.

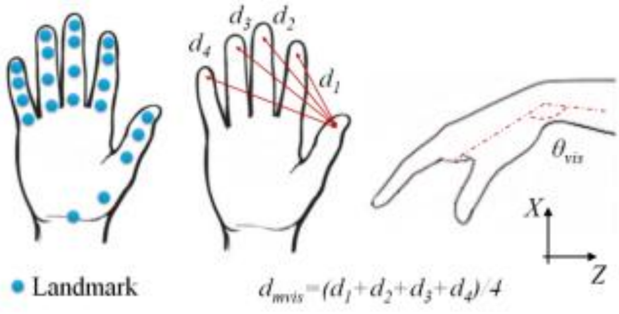

Fig 2.2.1 Parameters defined at the image processing phase

These findings demonstrate the potential of this hybrid approach to improve the precision and usability of gesture recognition systems, paving the way for advancements in assistive technologies and human-computer interaction.

C. Hand Gesture Recognition in Smart Healthcare

Gesture recognition in dynamic images is a complex challenge in computer vision, with significant applications in healthcare. In smart healthcare, these systems enhance human-computer interaction by enabling reliable gesture tracking and recognition in complex environments. Models combining CNNs and Deep Belief Networks (DBNs) have demonstrated high accuracy in recognizing gestures by processing video frames through noise reduction, feature extraction, and classification.

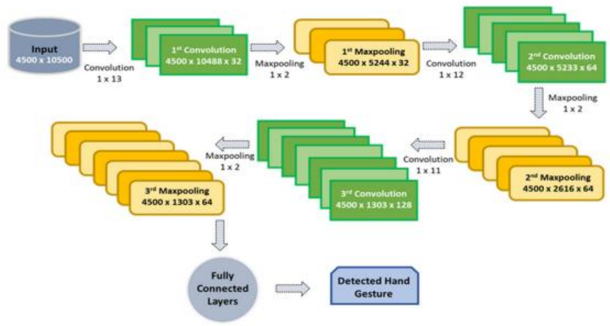

Fig 2.2.2 The architecture of proposed CNN model for hand detection

For instance, a proposed model achieved 90.73% accuracy on the Egogesture dataset, showcasing its effectiveness in comparison to other approaches. These systems are especially beneficial for paralyzed patients, allowing them to communicate their needs through gestures interpreted and displayed by computer-aided systems. Multimodal sensing technologies, such as surface electromyography (sEMG) and pressure-based force myography (pFMG), further improve recognition accuracy. By integrating multiple data inputs, such systems achieve up to 100% accuracy for specific gestures. This technology reduces the burden on healthcare providers by enabling more intuitive and efficient patient care. The use of CNNs and deep learning continues to advance gesture recognition, making these systems increasingly robust. Such advancements hold great potential for improving healthcare delivery and enhancing the quality of life for patients.

D. Real-Time Virtual Mouse for Hand Gesture-Based Computer Control

The development of a real-time virtual mouse based on hand gestures focuses on creating a software-driven interface that allows users to control a computer cursor without using a physical mouse. By utilizing sensors and machine learning algorithms, the system accurately recognizes hand gestures and translates them into cursor movements on the screen. This technology proves especially useful in environments where traditional input devices, such as underwater or space exploration, may not be practical. The virtual mouse presents an affordable, user-friendly alternative for human-computer interaction, offering an intuitive control method.

Accuracy in recognizing hand gestures, makes it a suitable for various applications. Additionally, by reducing physical contact with devices, it can help limit the spread of viruses. Incorporating technologies like OpenCV and MediaPipe boosts the system's responsiveness and precision, making it accessible to a wider range of users. This innovation opens up new possibilities for creating more natural and intuitive digital interfaces, advancing the field of human-computer interaction.

E. Gesture Recognition Using Millimeter Wave Radar and Thermal Imaging for Contactless Control

The gesture recognition system combining millimeter wave radar (mmWave) and thermal imaging was developed to address the growing demand for contactless control, especially during the COVID-19 pandemic. Unlike traditional machine vision-based systems, which can struggle in low-light conditions or complex environments, this system enhances recognition accuracy by using thermal imaging for better performance in challenging settings. The system employs a multi-color conversion algorithm to improve palm recognition on thermal images, supported by deep learning techniques to increase accuracy. As users perform gestures, mmWave radar captures point cloud data, which is then processed through a neural network model to track and monitor hand movements with over 80% accuracy. This combination of radar and thermal imaging overcomes the limitations of RGB cameras, which are less effective in low-light conditions. The result is a robust, reliable gesture recognition system suitable for diverse applications where traditional methods may fail.

This innovative approach provides a more effective solution for contactless control, improving performance and usability. By integrating mmWave radar and thermal imaging, the system opens up new possibilities for human-computer interaction in various fields. Its ability to work effectively in various environments makes it a valuable tool for modern applications. This advancement represents a significant leap in the field of gesture recognition, offering a superior alternative to conventional methods.

2.3. APPLICATION PROSPECT

⁎ *Advancements in Gesture Recognition Technology:*

Gesture recognition systems use RGB cameras to capture frames, which are processed with convolutional neural networks (CNNs) to interpret specific gestures. These systems are designed for real-time operation, allowing rapid and accurate identification of predefined hand movements. Techniques like data normalization and augmentation enhance model performance, ensuring reliability under different lighting conditions and for various users. Efficient feature extraction is achieved through pre-trained models such as ResNet and MobileNet, which also enable lightweight implementation for resource-constrained setups.

Gestures are mapped to specific commands, such as controlling media playback or adjusting volume, enabling intuitive device interactions. Training is conducted using frameworks like TensorFlow or PyTorch, with optimizers like Adam and RMSprop ensuring efficient model training. Real-time inference utilizes tools such as TensorFlow Lite or PyTorch Mobile for streamlined and responsive performance. Integration with external devices is achieved using automation tools and application programming interfaces (APIs), ensuring seamless functionality. Extensive testing across diverse scenarios ensures these systems are robust, scalable, and user friendly. Gesture recognition technology is a cornerstone of hands-free control, with applications spanning entertainment, healthcare, and accessibility solutions.

3. CONCLUSIONS

Gesture recognition systems that utilize thermal imaging and advanced neural network techniques provide a reliable and efficient solution for enabling contactless control in various applications. By employing low-resolution thermal sensors and Spiking Neural Networks (SNNs), these systems effectively address issues such as lighting inconsistencies and background interference, ensuring reliable performance in diverse conditions.

Techniques like Sparse Segmentation and Robust Principal Component Analysis (R-PCA) enhance the system’s accuracy while minimizing computational requirements. These developments make the technology well-suited for applications with limited resources, including automotive systems, healthcare devices, and smart home technologies. The combination of affordability, energy efficiency, and versatile functionality highlights the potential of thermal imaging-based gesture recognition as a transformative innovation in human computer interaction.

ACKNOWLEDGEMENT

We would like to express our gratitude to Malnad College of Engineering, Hassan, for providing the necessary resources and support for this research.

We also extend our appreciation to the authors and researchers whose work has significantly contributed to this survey.

Additionally, we acknowledge the valuable discussions and insights from our colleagues, which helped refine our study.

REFERENCES

1. Computer vision-based hand gesture recognition for human-robot interaction: a review Jing Qi · Li Ma · Zhenchao Cui · Yushu Yu (2023)
2. A machine vision and electromyographic-based approach for hand gesture recognition by Emilia Currò,Lorenzo Bombaci,Antonino Quattrocchi,Cristiano De Marchis,Dario Milone,Giovanni Gugliandolo and Nicola Donato(2024)
3. Diverse hand gesture recognition dataset by Zahra Mohammadi, Alireza Akhavanpour, Razieh Rastgoo ,Mohammad Sabokrou(2022)
4. Smart Healthcare Hand Gesture Recognition Using CNN-Based Detector and Deep Belief Network by Mohammed Alonazi , Hira Anser , Naif Al Mudawi,Saud S. Alotaibi , Nouf Abdullah

Almujally , Abdulwahab Alazeb, Ahmad Jalal , Jaekwang Kim And Moohong Min(2023)
5. Performance Assessment of Machine Learning Algorithms and Ensemble Techniques for Hand Gesture Recognition using Electromyographic Signals by Sushama Dhumal and Prashant Sharma (2023)
6. Real-Time virtual mouse using Hand Gestures for Unconventitonal Environment By Neha Sabrin TK and Prof. Aarti Karande (2023)
7. Gesture Recognition System based on Millimeter Wave and Thermal Imager By Wen-Hsiang Yeh, Yi-Lin Cheng and Yu-Ping Liao (2023)
8. Resource-Efficient Gesture Recognition using Low-Resolution Thermal Camera via Spiking Neural Networks and Sparse Segmentation By Ali Safa, Wout Mommen, Piet Wambacq and Lars Keuninckx (2024)
9. Robust Hand Gestures Recognition Using a Deep CNN and Thermal Images By Daniel Skomedal Breland , Aveen Dayal , Ajit Jha , Phaneendra K. Yalavarthy, Om Jee Pandey and Linga Reddy Cenkeramaddi (2021)
10. Machine Learning-Based Approach for Hand Gesture Recognition By M. Deepika, Shilpa Choudhary, Sandeep Kumar and Kodi Srinivas (2023)