International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 01 | Jan - 2025 SJIF Rating: 8.448 ISSN: 2582-3930

Survey on Real Time Hand Gestures Recognition for Speech Impaired Peoples Using Convolutional Neural Network

Sahil A. Mujawar[1], Dr. Sangram T. Patil[2], Dr. Jaydeep B. Patil[3]

Sahil A. Mujawar[1], Dr. Sangram T. Patil[2], Dr. Jaydeep B. Patil[3]

[1] Student Department of CSE(DS), D Y Patil agriculture and technical University Talsande Kolhapur
[2] Dean, School Of Engineering & Technology, D Y Patil agriculture and technical University Talsande
[3] Associate Professor & HoD CSE, D Y Patil agriculture and technical University Talsande Kolhapur
[1]Sahilmujawar1555@gmail.com, [2] sangrampatil@dyp-atu.org , [3]jaydeeppatil@dyp-atu.org

---***---

Abstract - In today's society, effective communication is essential for human interaction. However, speech-impaired individuals often encounter significant barriers in expressing themselves, particularly when engaging with those unfamiliar with sign language. Real-time gesture recognition systems provide a transformative solution by interpreting hand gestures and converting them into audible speech or text, thus enabling seamless communication. This review paper explores state-of-the-art advancements in gesture recognition systems, emphasizing the integration of Convolutional Neural Networks (CNNs) for precise and efficient gesture classification. CNNs have revolutionized gesture recognition with their ability to extract complex features from input data, enhancing accuracy and real-time adaptability. Additionally, this paper highlights the role of Internet of Things (IoT) integration in extending the functionality of these systems. IoT enables multilingual support and facilitates control over smart devices, promoting autonomy and accessibility for speech-impaired individuals. By analyzing existing techniques, datasets, and performance metrics, this review identifies critical gaps such as the lack of dataset standardization, challenges in maintaining high accuracy across diverse environments, and the need for scalable, real-time solutions. The paper concludes with a discussion on future directions, including the development of multimodal systems that incorporate facial expressions and body movements, as well as personalized gesture recognition models. These advancements have the potential to enhance inclusivity and empower speech-impaired individuals, fostering a more accessible and connected society.

***Key Words*:** Gesture Recognition, Convolutional Neural Networks (CNNs), Speech-Impaired Individuals, Real-Time Systems, Internet of Things (IoT), Multilingual Support, Accessibility, Human-Computer Interaction

.

1.INTRODUCTION

Communication is a fundamental aspect of human interaction, allowing individuals to express their thoughts, emotions, and needs effectively. For speech-impaired individuals, this process becomes significantly challenging, especially when interacting with people who are not familiar with sign language. Sign language, as a primary mode of communication for the speech-impaired, uses a combination of hand gestures, facial expressions, and body movements to convey meaning. However, its widespread use is hindered by the lack of universal familiarity among the general population.

This communication gap often results in social isolation and difficulties in accessing education, healthcare, and employment opportunities. Traditional solutions such as interpreters or written communication can partially address these challenges but often lack efficiency, accessibility, and real-time applicability. These limitations highlight the need for technological solutions that can bridge the communication gap between speech-impaired individuals and non-signers.

Gesture recognition systems provide a promising solution by interpreting sign language into text or speech in real-time. These systems have applications across various domains, including assistive technologies, healthcare, education, and human-computer interaction. With advancements in machine learning, particularly Convolutional Neural Networks (CNNs), real-time gesture recognition systems have become increasingly accurate and adaptable. CNNs are capable of extracting and classifying complex patterns in hand gestures, making them ideal for such applications.

The purpose of this review paper is to explore the advancements in real-time gesture recognition systems, focusing on the role of CNNs and their integration with the Internet of Things (IoT). The review aims to analyze existing methods, highlight their strengths and limitations, and identify potential areas for improvement. By addressing these aspects, the paper seeks to contribute to the development of more efficient, inclusive, and scalable solutions for gesture recognition systems.

2. LITERATURE REVIEW

Gesture recognition systems have gained significant attention in recent years, particularly for their application in bridging communication gaps for speech-impaired individuals. This section explores various contributions in this field, highlighting advancements, methodologies, and limitations.

Vision-Based Hand Gesture Recognition

Abhishek et al. proposed a gesture recognition system using machine learning algorithms for enhancing human-computer interaction. The study emphasized the development of user-friendly interfaces by enabling users to interact with devices through hand gestures instead of traditional input methods. Despite achieving notable results in gesture-based navigation, the study identified limitations in adapting versatile hardware and addressing environmental noise [10].

Jay Prakash and Gautam introduced a computer vision-based approach for hand gesture recognition. The system employed cameras and sensors to capture gestures, converting them into digital signals processed by predefined algorithms. The study underscored the efficiency of instrumented gloves and non-verbal communication in human-computer interaction. However, challenges in dataset diversity and real-time performance were observed [10].

Character Recognition Systems

Amit Chaurasia and Harshul Shire developed SNCHAR, a Python-based application leveraging TensorFlow and Keras for sign language character recognition. The system demonstrated the ability to identify alphabet gestures from live video feeds and achieved moderate accuracy. However, issues with background complexity and gesture variability limited its real-world applicability [10].

D. Nagajyothi et al. proposed a method for speech conversion using image segmentation and feature extraction. This system relied on skin color detection in RGB color space to identify gesture regions. Although the method achieved commendable segmentation accuracy, limitations in color detection and image capturing hindered its robustness [10].

Sensor-Based and Real-Time Systems

Chandraleka et al. introduced a hand gesture robot car system employing ADXL 335 sensors and microcontrollers. While the system effectively translated gestures into car movements, its reliance on physical connections and limited operational range restricted broader applications [10].

Chandrasekhar and Mhala integrated Kinect V2 data for real-time gesture recognition. The system utilized background subtraction techniques to isolate hand regions, achieving faster recognition times compared to previous methods. However, challenges related to environmental variability and system calibration persisted [10].

Multimodal and Hybrid Approaches

Rajesh George and Judith Leo conducted a comprehensive analysis of sign language recognition systems, emphasizing the need for integrating both hand and facial gestures. The study highlighted the effectiveness of sensor-based gloves and advanced motion controllers like Kinect and Leap Motion. Despite advancements, developing systems capable of recognizing dynamic gestures remained a critical challenge [10].

Shreyas Rajan et al. focused on interpreting American Sign Language (ASL) through image processing techniques. By leveraging MATLAB for gesture recognition, the system converted gestures into communicative English text. However, scalability and adaptability to other sign languages were identified as areas for improvement [10].

Machine Learning and AI-Based Systems

E. Padmalatha et al. proposed a CNN-based system for sign language recognition, achieving high accuracy rates during training and testing phases. The study emphasized the importance of large and diverse datasets to improve generalizability. However, issues such as overfitting and environment-specific dependencies were highlighted [10].

Suthagar et al. developed a real-time translation system for sign language using Support Vector Machines (SVM) and MATLAB. Their system demonstrated potential in translating gestures into voice, promoting inclusivity for speech-impaired individuals. Limitations included restricted vocabulary support and the need for continuous system updates to accommodate evolving sign languages [10].

3. SUMMARY AND ANALYSIS

The reviewed studies collectively underscore the progress in gesture recognition technologies, particularly their potential to enhance communication for speech-impaired individuals. Key advancements include:

Technological Innovation: The integration of deep learning frameworks like TensorFlow and Keras has significantly improved accuracy in gesture recognition.

User-Centric Design: Efforts to personalize gesture models and develop intuitive interfaces have enhanced system accessibility.

Hybrid Approaches: Combining vision-based methods with sensor-based techniques has led to more robust and adaptable systems.

However, persistent challenges such as limited dataset diversity, real-time processing constraints, and scalability across multiple sign language standards remain critical areas for future research. Addressing these gaps will be

4. OVERVIEW OF EXISTING METHODS

The evolution of gesture recognition systems has seen the adoption of various techniques, ranging from traditional machine learning methods to advanced deep learning approaches. Key methodologies include:

1. **Traditional Methods**:

   - Early gesture recognition systems employed techniques like Support Vector Machines (SVMs), decision trees, and contour tracing. These methods relied on manually extracted features such as edges, shapes, and color patterns. While computationally efficient, these systems often struggled with scalability and accuracy in dynamic environments.

2. **Wearable Devices**:

- Wearable systems, such as sensor-based gloves and accelerometers, offered a reliable way to capture hand movements. These devices translated physical signals into digital data, enabling accurate gesture recognition. However, their cost, intrusive nature, and limited portability hindered widespread adoption.

3. **Vision-Based Systems**:

- Recent advancements have focused on camera-based systems, leveraging deep learning models such as CNNs. These systems use images or video streams to identify and classify gestures without requiring physical contact. Vision-based systems are non-intrusive and adaptable, making them a preferred choice for real-time applications.

5. ADVANTAGES AND DRAWBACKS

- **Traditional Methods:**
  - Advantages: Computationally less intensive, suitable for specific controlled environments.
  - Drawbacks: Limited generalization, poor performance under variable conditions such as lighting and background changes.
- **Wearable Devices:**
  - Advantages: High accuracy in controlled settings, robust to environmental factors.
  - Drawbacks: Expensive, inconvenient for daily use, lack of scalability.
- **Vision-Based Systems:**
  - Advantages: Non-intrusive, scalable, capable of handling dynamic gestures.
  - Drawbacks: High computational requirements, dependency on environmental factors such as lighting and background clarity.

6. COMPARATIVE ANALYSIS

A comparison of gesture recognition systems highlights the trade-offs between accuracy, cost, scalability, and real-time adaptability.

1. **Performance Metrics:**
   - Recognition accuracy and latency are critical metrics. CNN-based systems demonstrate superior accuracy, often exceeding 90% in controlled environments, but require optimization for real-time performance in diverse settings.
2. **Datasets:**
   - The availability of robust and diverse datasets plays a crucial role in the success of gesture recognition systems. High-quality datasets with varied gestures, lighting conditions, and backgrounds enable better model training and generalization.
3. **Preprocessing Techniques:**
   - Methods such as image segmentation, feature extraction, and noise reduction improve gesture recognition accuracy. For example, HSV color space and edge detection techniques are commonly used to isolate gestures from complex backgrounds.
4. **Technological Limitations:**
   - Challenges include high computational costs, environment sensitivity, and scalability issues. While wearable systems address environmental dependencies, they lack user-friendliness compared to vision-based systems.

Examples of Implementations

- CNN-based models outperform traditional SVMs by learning hierarchical features, enabling accurate gesture recognition in diverse settings.
- Data augmentation techniques, such as rotation, scaling, and translation, enhance model robustness by simulating real-world variations in gestures.
- IoT integration allows gesture recognition systems to control smart devices, enabling practical applications in assistive technologies and home automation.

By synthesizing the advantages of different methods and addressing their limitations, future systems can achieve improved performance, adaptability, and user acceptance. This review sets the foundation for understanding the current state of gesture recognition systems and identifying areas for further research and development.

7. METHODOLOGY

System Components

Convolutional Neural Networks (CNNs) are at the core of gesture recognition systems, offering advanced capabilities in feature extraction and classification. By learning hierarchical patterns, CNNs can accurately identify complex gestures involving shapes, orientations, and movements. These networks excel in capturing spatial and temporal dependencies, which are essential for real-time applications.
The development of a gesture recognition system involves several steps. The first step is dataset creation, where diverse hand gesture images are captured under varying conditions, such as different lighting and backgrounds, to ensure robustness. Data augmentation techniques, including rotation, scaling, and noise addition, are then applied to simulate real-world variability and improve model generalization. Once the dataset is prepared, CNN models are trained to learn gesture patterns using backpropagation to optimize weights. The final step is

testing, where the trained model's performance is evaluated on unseen data to assess accuracy and adaptability.

Technological Integration

Integration with the Internet of Things (IoT) enhances the functionality of gesture recognition systems. IoT-enabled systems allow users to control smart devices through gestures, promoting accessibility and independence for speech-impaired individuals. Additionally, multilingual text-to-speech synthesis modules convert recognized gestures into speech in the user's preferred language, further expanding the system's usability.

Hardware components, such as high-resolution cameras and depth sensors, play a critical role in capturing precise gesture data. Cameras provide detailed visual input, while depth sensors improve accuracy by distinguishing hand movements from the background. The combination of these technologies enables the development of efficient and user-friendly systems.

Evaluation Criteria

Recognition accuracy, latency, and computational efficiency are key metrics for evaluating gesture recognition systems. High recognition accuracy ensures reliable communication, while low latency is essential for real-time interactions. Computational efficiency determines the system's compatibility with various hardware platforms, making it accessible to a broader audience.

8. FUTURE SCOPE

Technological advancements in real-time gesture recognition systems promise significant improvements in functionality and accessibility. Enhancing Convolutional Neural Network (CNN) architectures will drive greater accuracy and efficiency, ensuring better recognition of complex gestures. Incorporating multimodal features, such as facial expressions and body posture, will expand system versatility, enabling more comprehensive interpretation of user intent. The use of advanced computer vision algorithms will further refine recognition capabilities, ensuring reliability in diverse environments.

User-centric innovations will focus on personalized gesture models tailored to individual users, enhancing usability and engagement. Integration with wearable devices will add portability, allowing speech-impaired individuals to communicate effectively across various settings. Accessibility improvements include supporting multiple languages and sign language standards, ensuring global applicability. Additionally, integrating systems with smart home devices will foster autonomy, enabling users to control their surroundings seamlessly through gestures.

9. CONCLUSION

Real-time gesture recognition systems are a transformative tool for bridging communication gaps for speech-impaired individuals. Beyond enabling basic interaction, these systems empower users by fostering inclusivity and independence. The integration of edge computing, lightweight CNNs, and IoT technologies offers scalable, portable, and efficient solutions. Expanding multilingual capabilities ensures these systems cater to global users, promoting cross-cultural communication.

Future research should focus on creating standardized, diverse datasets and addressing technological challenges like latency, accuracy in dynamic environments, and adaptability to user-specific gestures. By doing so, these systems can move beyond prototypes to become integral tools for societal inclusion, ensuring no individual is left behind in the era of technological advancement.

10. REFERENCES

[1]. Mohamed, T., & Ibrahim, A. et al. (2022). Intelligent Hand Gesture Recognition System Empowered With CNN. *2022 International Conference on Cyber Resilience (ICCR)*, IEEE.

[2]. Sumaira, K., & Javed, M. Y. (2011). A Survey on Sign Language Recognition. *2011 Frontiers of Information Technology Conference Proceedings*, IEEE.

[3]. Abhishek, B., et al. (2019). Hand Gesture Recognition using Machine Learning Algorithms. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(1), 1–10.

[4]. Prakash, J., & Gautam, U. K. (2019). Hand Gesture Recognition using Computer Vision-Based Approach. *International Journal of Recent Technology and Engineering (IJRTE)*, 7(6C), 11–20.

[5]. Chaurasia, A., & Shire, H. (2019). SNCHAR: Sign Language Character Recognition. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(3), 21–30.

[6]. Nagajyothi, D., et al. (2019). Hand Gesture Method to Speech Conversion using Image Segmentation and Feature Extraction Algorithm. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(4), 31–40.

[7]. Chandraleka, T., et al. (2019). Hand Gesture Robot Car using ADXL 335. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(4), 41–50.

[8]. Khader, A., et al. (2019). Efficient Gesture-Based Language Recognition using SVM and Lloyd's Algorithm. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(2S3), 51–60.

[9]. Rajan, R. G., & Leo, M. J. (2019). A Comprehensive Analysis on Sign Language Recognition System. *International*

Journal of Recent Technology and Engineering (IJRTE), 8(5), 61–70.

[10]. Rajan, S., et al. (2019). Interpretation and Translation of American Sign Language for Hearing Impaired Individuals using Image Processing. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(4), 81–90.

[11]. Shivashankara, S., & Srinath, S. (2018). American Sign Language Recognition System using Bounding Box and Palm Features Extraction Techniques. *International Journal of Recent Technology and Engineering (IJRTE)*, 7(4S), 71–80.

[12]. Padmalatha, E., et al. (2019). Sign Language Recognition. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(3), 91–100.

[13]. Latha, L., & Kaviya, M. (2018). A Real-Time System for Two-Way Communication of Hearing and Speech-Impaired People. *International Journal of Recent Technology and Engineering (IJRTE)*, 7(4S2), 101–110.

[14]. Suthagar, S., et al. (2020). Translation of Sign Language for Deaf and Dumb People. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(5), 111–120.