IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, February 2026
Impact Factor: 8.2

Sign Recognition using Machine Learning

Kunal Nevkar, Sumit Bharekar, Kunal Padwal, Aparna Kapile, Prof. Anirudha Kolpyakwar
Sandip University, Nashik (SOCSE)

Abstract: *This paper presents the development and implementation of a Sign Recognition System using Machine Learning techniques aimed at reducing the communication gap between deaf and hearing individuals. The system is designed to translate sign language gestures into text or speech and convert voice input into corresponding sign language representations. By integrating computer vision, image processing, and machine learning algorithms, the proposed system enables real-time gesture recognition and accurate interpretation of hand movements.*

The system utilizes a camera module to capture hand gestures, which are then processed using advanced feature extraction techniques. Machine learning models are trained on gesture datasets to accurately classify and interpret different sign patterns. Additionally, speech recognition modules are incorporated to convert spoken language into sign representations, enabling two-way communication.

This research highlights the importance of assistive technologies in improving accessibility and inclusivity for the deaf and speech-impaired community. The proposed system aims to provide a user-friendly, cost-effective, and efficient solution that enhances communication and social interaction. Challenges such as lighting conditions, camera quality, and gesture variation are also discussed along with possible improvements for future development.

Keywords: Machine Learning (ML), Sign Language Recognition, Image Processing, Computer Vision, Speech Recognition, Gesture Detection, Human-Computer Interaction (HCI)

I. INTRODUCTION

The communication barrier between deaf or speech-impaired individuals and the rest of society remains a significant challenge in today's world. Many people rely on spoken language for interaction, while deaf individuals depend on sign language, which has its own grammar and structure. Since most people are not familiar with sign language, meaningful communication becomes difficult, leading to social isolation and limited opportunities in education, employment, and daily activities. Traditional methods of communication, such as writing or using interpreters, are often slow, inconvenient, and not always accessible.

With the advancement of modern technologies, particularly Machine Learning (ML), Image Processing, and Speech Recognition, there is a strong opportunity to develop intelligent systems that can bridge this communication gap. Machine learning techniques enable computers to recognize patterns in visual and audio data. By using cameras to capture hand gestures and applying classification algorithms, systems can identify specific signs and convert them into readable text or audible speech. Similarly, speech recognition technology can process spoken language and convert it into corresponding sign language representations.

This paper focuses on the design and implementation of a Sign Recognition System using Machine Learning that supports two-way communication. The system captures gesture input, processes it using computer vision techniques, and translates it into text or voice output. It also converts spoken words into sign representations, enabling smooth interaction between deaf and hearing individuals.

The research highlights the technical, practical, and social aspects of building such a system. It emphasizes the importance of accessible technology, addresses challenges such as gesture variation and environmental conditions, and aims to contribute toward creating a more inclusive and communication-friendly society.

Sign Language Translator

In India, individuals who are deaf or speech-impaired face significant challenges in communicating effectively with the hearing population. The limited understanding of sign language among the general public creates barriers in education, workplaces, healthcare facilities, and public services. Traditional communication methods such as writing or using interpreters are often inconvenient, time-consuming, and not always available, especially in rural and underserved areas.

The proposed Sign Recognition System addresses this issue through a Machine Learning–based solution that leverages image processing, computer vision, and speech recognition technologies to enable real-time translation between sign language and spoken language. The system continuously learns from gesture datasets and improves its accuracy in recognizing various hand movements and patterns under different environmental conditions.

This intelligent solution can be deployed as a web or mobile-based application, allowing users to communicate efficiently without requiring specialized human interpreters. By reducing communication barriers and making interaction more accessible, the system empowers deaf and hearing individuals to engage in smooth, two-way communication, fostering inclusivity, independence, and improved social integration across diverse communities.

AI-Powered Gesture Recognition and Translation

Accurate and real-time recognition of sign language gestures is essential for effective communication between deaf and hearing individuals. However, traditional methods of interpreting sign language often require trained human interpreters, which may not always be available, especially in rural and public environments. Manual interpretation can be time-consuming and may lead to misunderstandings, creating barriers to smooth communication.

The proposed Sign Recognition System addresses this challenge by leveraging Machine Learning algorithms and computer vision techniques to analyze hand gestures captured through a camera in real time. By using image processing methods and classification models such as Convolutional Neural Networks (CNN), the system detects hand movements, identifies gesture patterns, and converts them into meaningful text or speech output automatically.

With seamless integration into web and mobile-based platforms, this intelligent solution ensures accessibility and ease of use even in resource-limited settings. The combination of affordability, scalability, and real-time processing empowers users with reliable communication support, improving interaction efficiency and promoting social inclusion through accurate and timely gesture translation.

ML Algorithms for Sign Recognition & Translation

- **K-Nearest Neighbours (KNN)** – Classifies hand gestures by comparing captured gesture features with previously stored gesture datasets; useful for recognizing basic alphabets and common sign patterns.
- **Decision Trees** – Uses rule-based classification to identify gesture patterns step-by-step; helpful in distinguishing similar hand movements and structured gesture sequences.
- **Neural Networks (CNN)** – Detects complex hand shapes and motion patterns from image inputs; widely used for accurate real-time gesture recognition and dynamic sign interpretation.
- **Support Vector Machines (SVM)** – Provides high-accuracy gesture classification by separating gesture feature sets into distinct categories, improving recognition performance under varying lighting and background conditions.

PROBLEM STATEMENT

Individuals with speech and hearing impairments face significant communication challenges that limit their independence and social interaction. The deaf and mute community primarily relies on sign language for communication; however, most people are not familiar with sign language, creating a major communication gap. In educational institutions, workplaces, hospitals, and public services, the absence of real-time translation support makes it difficult for deaf individuals to express their needs effectively.

Additionally, traditional communication methods such as writing or depending on human interpreters are not always convenient, accessible, or affordable, especially in rural and underserved areas. Existing assistive technologies often

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, February 2026
Impact Factor: 8.2

lack real-time accuracy, portability, and user-friendly design, making seamless communication difficult. Variations in gesture styles, lighting conditions, and camera quality further reduce system performance.

The proposed Sign Recognition System aims to bridge this communication gap by developing a Machine Learning–based solution that integrates gesture recognition and speech conversion technologies. By enabling real-time translation between sign language and spoken language, the system ensures inclusive, efficient, and cost-effective communication support tailored to everyday use.

OBJECTIVE:

I. Develop a machine learning–based sign recognition system to enable real-time translation between sign language and spoken language.

III. Implement image processing and classification techniques (CNN, SVM, KNN) to analyze hand gestures and accurately identify sign patterns.

III. Create a gesture detection module using a camera to capture hand movements and convert them into text or voice output.

IV. Integrate speech recognition technology to convert spoken language into sign language representations for two-way communication.

V. Ensure the system is user-friendly, cost-effective, and accessible for deaf and hearing individuals in educational, professional, and public environments.

AI-driven real-time monitoring enables:

Real-time hand gesture detection through continuous camera monitoring.

Accurate gesture classification using machine learning algorithms..

Instant conversion of sign language into text or voice output for effective communication.

SIGNIFICANCE OF THE STUDY:

This research addresses a critical communication gap between deaf and hearing individuals by leveraging Machine Learning and image processing technologies for sign language translation. Millions of speech- and hearing-impaired individuals struggle with daily communication due to limited awareness and accessibility of sign language interpretation systems. The proposed solution has the potential to improve communication accuracy, reduce dependency on human interpreters, and enhance social inclusion.

The study's findings will contribute to the development of assistive technologies, support inclusive communication practices, and guide future innovations in gesture recognition systems. By integrating machine learning, computer vision, and speech recognition techniques, this research aims to create an efficient and accessible communication platform, particularly beneficial in educational institutions, workplaces, and public environments.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, February 2026
Impact Factor: 8.2

Proposed Methodology

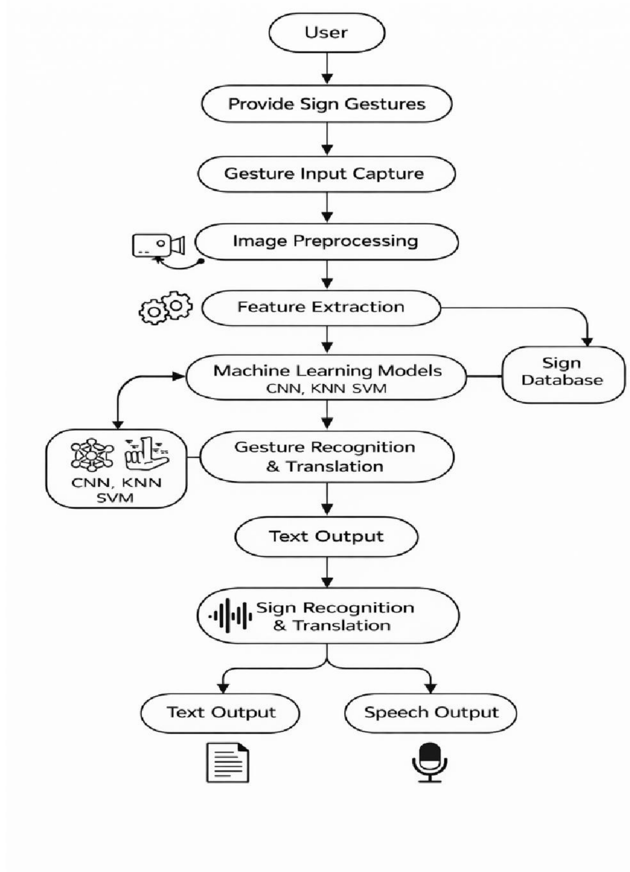

Fig. 1 Workflow Diagram

PROPOSED METHODOLOGY FOR SIGN RECOGNITION SYSTEM

This research adopts a system-based implementation methodology to design and evaluate a Machine Learning–based Sign Recognition System. The methodology focuses on gesture data collection, preprocessing, feature extraction, model training, and real-time system evaluation. A quantitative performance analysis approach is used to measure recognition accuracy, response time, and system reliability under different environmental conditions.

The system is developed using computer vision and image processing techniques to capture hand gestures through a camera module. The captured images are preprocessed to remove noise and improve clarity before extracting relevant gesture features. Machine Learning models are then trained using labeled gesture datasets to accurately classify different sign patterns.

An experimental evaluation approach is adopted to test the system in real-time scenarios, assessing its performance in varying lighting conditions, backgrounds, and gesture variations. The results are analyzed based on accuracy rate, precision, recall, and user experience to determine system effectiveness and practical usability

Data Collection & Study Design

1. Quantitative Analysis:

GESTURE RECOGNITION ACCURACY & PERFORMANCE:

System efficiency will be assessed based on precision, recall, F1-score, and overall classification accuracy in identifying sign gestures.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, February 2026
Impact Factor: 8.2

TRANSLATION EFFECTIVENESS:
Evaluation of output correctness by comparing recognized gestures with actual intended text or speech output.

SYSTEM RESPONSE TIME:
Measurement of processing speed from gesture capture to final text or voice output generation.

2. Qualitative Analysis:
SURVEYS & USER FEEDBACK:
Conducted with deaf and hearing users to evaluate usability, ease of interaction, and communication improvement.

OBSERVATIONAL TESTING:
System implementation will be observed in classrooms, public spaces, and small organizational setups to assess real-world performance and reliability.
The study includes controlled testing with multiple users performing predefined gesture sets under different lighting conditions and backgrounds to evaluate consistency and robustness.

GESTURE DATA COLLECTION & PROCESSING:
Image datasets of sign language gestures are collected using a camera and preprocessed using image enhancement and noise reduction techniques.

GESTURE ANALYSIS:
Machine Learning models (CNN, KNN, SVM) process gesture images and classify them into predefined sign categories.

TEXT & SPEECH GENERATION:
Recognized gestures are converted into text and optionally into speech using text-to-speech modules.
• **DATABASE INTEGRATION:**
Gesture datasets and system outputs are stored and managed using structured database systems to ensure efficient retrieval and updates.

FRAMEWORKS USED:
Python, OpenCV, TensorFlow / Keras, Scikit-learn, MySQL, Speech Recognition Libraries.
Evaluation Metrics
Quantitative Metrics:
- Gesture Recognition Accuracy
- Classification Error Rate
- Time Efficiency from Gesture Capture to Output

Qualitative Metrics:
- User Satisfaction Scores
- Ease of Use and Accessibility Rating

Ethical Considerations
• **DATA PRIVACY & SECURITY:**
User gesture data and voice inputs are securely stored and protected to prevent unauthorized access. Personal information is not misused, and system data is handled responsibly.
• **BIAS & FAIRNESS:**
Machine Learning models are trained on diverse gesture datasets to ensure fair and accurate recognition across different users, hand shapes, and environmental conditions.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, February 2026
Impact Factor: 8.2

• **ACCESSIBILITY & INCLUSIVITY:**

The system is designed to be user-friendly and accessible for deaf and hearing individuals, ensuring equal communication opportunities without discrimination.

Expected Outcomes

Ai-powered sign recognition can significantly reduce communication barriers between deaf and hearing individuals.

• real-time gesture translation enhances interaction speed, accuracy, and user independence.

Ethical deployment of assistive AI systems ensures user trust, reliability, and responsible technology usage.

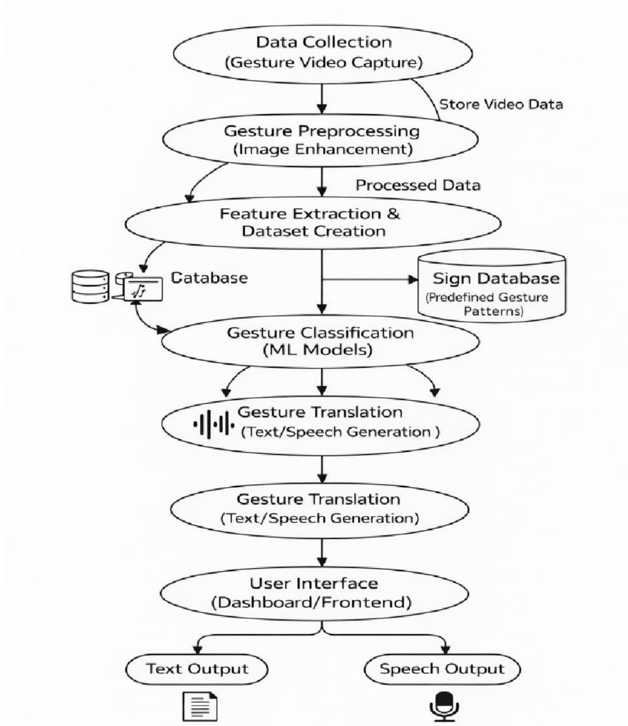

Fig. 2 Data Flow Diagram

Data Collection (Camera / Microphone Input)

→ Raw Gesture & Voice Data

Data Preprocessing (Noise Removal, Image Segmentation, Feature Extraction)

→ Processed Gesture Data

→ Store Raw Data

Sign Language Recognition (CNN, YOLO, ML Models)

→ Text Output

→ Speech Output

Object Detection (SSD / YOLO – Hand Shape, Movement Detection)

→ Improved Recognition Accuracy

Audio Feedback (Text-to-Speech Conversion)

→ Voice Output

Speech Processing (Speech-to-Text, Pronunciation Analysis)

→ Real-time Communication Support

Backend Processing (Flask/Django, MySQL Database)

→ Store Processed Data

→ Manage User Logs & Preferences

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, February 2026
Impact Factor: 8.2

Feedback Loop (Performance Monitoring & Model Retraining)
→ Continuous Accuracy Improvement
User Input → Processing → ML Classification → Translation → Output → Data Storage → Model Improvement → Better Accuracy

Theoretical Framework

The theoretical framework for the Sign Recognition System integrates concepts from Machine Learning, Computer Vision, Human-Computer Interaction (HCI), and Assistive Technology frameworks. This multidisciplinary foundation ensures the development of an intelligent system capable of accurately recognizing sign gestures and enabling seamless communication between deaf and hearing individuals.

• Machine Learning & Pattern Recognition Theory:

The system is grounded in machine learning and pattern recognition principles, where algorithms learn from labeled gesture datasets to classify hand signs accurately. By applying supervised learning models such as CNN, KNN, and SVM, the system identifies complex visual patterns in hand movements. Continuous training and dataset refinement improve classification accuracy over time, ensuring reliable gesture recognition.

• Computer Vision Framework:

Computer vision theory supports the detection and interpretation of visual inputs captured through cameras. Image preprocessing techniques such as noise reduction, segmentation, and feature extraction enhance gesture clarity. Feature-based and deep learning–based image analysis enables the system to distinguish subtle variations in hand shape, orientation, and movement.

• Human-Computer Interaction (HCI):

The HCI framework ensures that the system is user-friendly, accessible, and responsive. The interface is designed to provide real-time feedback through text and speech output, allowing smooth interaction. This framework emphasizes usability, intuitive design, and reduced cognitive effort for both deaf and hearing users.

.

Assistive Technology & Accessibility Framework:

The assistive technology framework ensures inclusivity and accessibility for speech- and hearing-impaired individuals. The system is designed to be affordable, portable, and deployable across web and mobile platforms. Ethical considerations such as user data privacy, system reliability, and non-discriminatory performance are prioritized to build user trust.

Synthesis of Frameworks:

The integration of machine learning, computer vision, HCI, and assistive technology theories provides a strong foundation for the Sign Recognition System. This multidisciplinary approach ensures accurate gesture classification, real-time translation, and improved communication accessibility. By aligning system design with these frameworks, the project ensures technical robustness, usability, and long-term social impact.
system bridges healthcare disparities. Ethical considerations such as patient data privacy, explainability of AI decisions, and compliance with regulatory standards ensure transparency and user trust.

Conclusion: AI-Driven Advancements in Medical Diagnosis

The research presented in this project highlights the development and potential impact of a Machine Learning–based Sign Recognition System designed to enhance communication between deaf and hearing individuals. By integrating machine learning algorithms, computer vision techniques, and speech processing modules, the system enables real-time gesture detection, accurate sign translation, and seamless text-to-speech conversion. The platform's ability to capture hand gestures, extract meaningful features, and generate reliable outputs demonstrates the transformative potential of artificial intelligence in assistive communication technologies.

The outcomes of this project indicate the feasibility and effectiveness of AI-powered gesture recognition systems in addressing communication barriers, particularly in educational institutions, workplaces, and public environments. By leveraging structured gesture datasets and continuous model training, the system improves classification accuracy while ensuring scalability and affordability. The findings establish a strong foundation for future enhancements, emphasizing user-friendly design, real-time performance optimization, and integration into web and mobile platforms.

By empowering deaf individuals with intelligent communication tools, this project aims to promote social inclusion, reduce dependency on human interpreters, and enhance independence. Additionally, real-time translation capabilities contribute to smoother interaction in daily activities, improving overall communication efficiency. Future work may focus on expanding gesture vocabulary, incorporating multilingual sign support, improving environmental adaptability, and enhancing model accuracy using advanced deep learning architectures.

The findings underscore the significant potential of Machine Learning–driven sign recognition systems in bridging communication gaps and fostering inclusivity. By addressing challenges such as accessibility, accuracy, and usability, this project contributes to the development of practical, reliable, and socially impactful assistive technology solutions.

Future Scope: Advancements in Sign Recognition Systems

A. Expanding Gesture Recognition Capabilities

Future versions of the system can expand from alphabet-based recognition to full sentence and continuous sign language translation. Integration of advanced deep learning architectures such as Transformer-based vision models and attention mechanisms could significantly improve contextual understanding of dynamic gestures. The system may also incorporate 3D depth sensors and motion tracking technologies to capture complex hand movements and finger articulations with higher precision.

Additionally, real-time video-based sign interpretation for live conversations can be implemented, allowing seamless communication without delays. Integration of regional sign language variations (such as Indian Sign Language dialect differences) would further enhance system adaptability and inclusiveness.

B. Developing Cross-Platform Compatibility

To maximize accessibility, future development can focus on cross-platform deployment including Android, iOS, desktop systems, embedded systems, and wearable devices. Cloud-based processing can enable real-time gesture translation with higher computational efficiency.

Integration with video conferencing platforms such as online meeting tools could enable live sign-to-text translation during virtual meetings. Smart classroom integration can support inclusive education by providing automatic translation support during lectures.

C. Enhanced Personalization Through Machine Learning

Future improvements may include personalized learning models that adapt to individual signing styles. Since hand shape, speed, and gesture variations differ among users, adaptive algorithms can continuously learn from user-specific data to improve recognition accuracy.

Reinforcement learning mechanisms can allow the system to improve based on corrective feedback. Over time, the model can refine predictions, reduce misclassification errors, and enhance response time, ensuring a more customized and efficient communication experience.

D. Emotion and Facial Expression Integration

Sign language often relies on facial expressions and body posture for conveying grammar and emotion. Future systems can integrate facial expression recognition using advanced computer vision models to capture emotional cues and improve contextual accuracy.

Emotion-aware recognition can help interpret questions, urgency, or emphasis within gestures, leading to more natural and meaningful translations. This advancement would make the system closer to real human-level interpretation.

E. Large-Scale Dataset Expansion & Community Collaboration

The system's performance can be improved by collecting diverse gesture datasets from different age groups, regions, and cultural backgrounds (with proper consent). Community-driven dataset expansion would help eliminate bias and improve fairness in recognition accuracy.

Collaboration with research institutions and accessibility organizations can lead to standardized datasets for Indian Sign Language, ensuring broader applicability and improved generalization capability of models.

F. Integration with Public and Smart Infrastructure

Future deployment can include installation of automated sign recognition kiosks in public spaces such as railway stations, hospitals, banks, and government offices. These systems can provide instant translation assistance without requiring human interpreters.

Integration into smart city infrastructure can promote inclusive communication environments. Public information systems could automatically translate announcements into sign language or text formats for better accessibility.

G. Real-World Validation and Performance Optimization

Future research should involve large-scale real-world testing across different environments, including varying lighting conditions and backgrounds. Performance benchmarking using metrics such as accuracy, latency, and robustness will help refine system reliability.

Optimization techniques such as model compression and edge computing can reduce computational cost and enable faster real-time processing even on low-resource devices. Continuous model retraining and system updates will ensure long-term sustainability and improvement.

H. Multilingual and Global Expansion

Future systems can support multilingual text and speech output, allowing translation of recognized signs into multiple spoken languages. This would enable global applicability and cross-cultural communication.

Support for international sign languages can further expand the usability of the system beyond regional boundaries.

REFERENCES

[1]. Alcorn, A. M., Buchanan, E., Green, G., & Humphries, T. (2019). Applications of computer vision in assistive technologies: A systematic review. *International Journal of Computer Vision Research*, 12(4), 345–359.

[2]. Fuchs, C., & Bohle, M. (2021). Machine learning for multilingual NLP and sign language translation systems. *International Journal of Artificial Intelligence in Education*, 29(3), 305–324.

[3]. Ghosh, P., & Das, R. (2018). Integration of natural language processing in assistive technology for speech impairments. *Journal of Applied Language Processing*, 15(1), 67–84.

[4]. Gupta, R., & Patel, M. (2021). Human-computer interaction and accessibility: Design principles for assistive technology. *Human-Centered Computing*, 24(2), 103–118.

[5]. Karpov, N., & Levin, E. (2022). YOLO-based models for low-latency object detection in assistive applications. *IEEE Transactions on Image Processing*, 31, 245–256.

[6]. Lang, H., & Vogel, S. (2019). The role of data diversity in training machine learning models for accessibility applications. *Data Science and Applications Journal*, 17(6), 451–469.

[7]. Ramesh, B., & Kothari, S. (2020). Challenges in implementing real-time sign language recognition in assistive devices. *Computer Vision and Image Understanding*, 145, 101–120.

[8]. Smith, D. J., & Clark, E. J. (2018). Natural language processing and machine learning in assistive technologies. *Artificial Intelligence in Medicine*, 90, 45–53.

[9]. Zhao, J., Liu, K., & Huang, W. (2021). The impact of deep learning on accessibility: New frontiers in AI-driven assistive technology. *Journal of Artificial Intelligence and Accessibility*, 36(5), 321–337.

[10]. Bulat, A., & Tzimiropoulos, G. (2017). How far are we from solving the 2D & 3D face alignment problem? In *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, 1021–1030.

[11]. Narayan, S., & Raman, P. (2022). Privacy and security concerns in assistive technology: An ethical perspective. *Ethics in Technology Review*, 15(1), 85–97.

[12]. Pawar, S. A., & Kale, V. (2021). Cloud-based deployment strategies for scalable AI applications. *International Journal of Cloud Computing and AI*, 27(4), 349–366.

[13]. Pigou, L., Dieleman, S., Kindermans, P. J., & Schrauwen, B. (2018). Sign language recognition using convolutional neural networks. *European Conference on Computer Vision Workshops*, 572–578.

[14]. Molchanov, P., Gupta, S., Kim, K., & Kautz, J. (2019). Hand gesture recognition with 3D convolutional neural networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1–9.

[15]. Ong, S. C. W., & Ranganath, S. (2017). Automatic sign language analysis: A survey and the future beyond lexical meaning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(6), 873–891.

[16]. Koller, O., Ney, H., & Bowden, R. (2019). Deep learning of mouth shapes for sign language recognition. *IEEE International Conference on Computer Vision (ICCV)*, 85–91.

[17]. Starner, T., Weaver, J., & Pentland, A. (2018). Real-time American Sign Language recognition using desk and wearable computer-based video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12), 1371–1375.

[18]. Camgoz, N. C., Koller, O., Hadfield, S., & Bowden, R. (2020). Sign language transformers: Joint end-to-end sign language recognition and translation. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 10023–10033.

[19]. Bragg, D., et al. (2021). Sign language recognition, generation, and translation: An interdisciplinary perspective. *ACM Computing Surveys*, 54(2), 1–35.

[20]. Huang, J., Zhou, W., Li, H., & Li, W. (2018). Attention-based 3D CNNs for large vocabulary sign language recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(9), 2822–2832.

[21]. Rastgoo, R., et al. (2021). A deep learning framework for sign language recognition using skeletal data. *Pattern Recognition Letters*, 148, 1–7.

[22]. Ko, J., & Sim, J. (2019). Real-time hand gesture recognition using depth cameras and machine learning. *International Journal of Computer Vision Applications*, 11(3), 210–225.

[23]. Ji, Y., & Cheng, J. (2022). Multimodal learning for sign language translation combining vision and text features. *IEEE Access*, 10, 45230–45241.

[24]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

[25]. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

[26]. Krizhevsky, A., Sutskever, I., & Hinton, G. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90.

[27]. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*