International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Sign Language Prediction Using Deep Learning

K.Kavitha Sri, Mrs.Abirami

Student(IT) – Assistant Professor(IT)

Francis Xavier Engineering College, Tirunelveli, India

ABSTRACT

This project aims to bridge the communication barrier for individuals who are deaf or hard of hearing by creating a real-time sign language recognition system. The system utilizes a webcam to interpret American Sign Language (ASL), Indian Sign Language (ISL), and British Sign Language (BSL), converting hand gestures into both text and speech to aid those who are not familiar with sign language.

To achieve accurate gesture recognition, the system integrates advanced artificial intelligence techniques. It captures live video input and processes it using MediaPipe for detecting hand landmarks. A Convolutional Neural Network (CNN) model is then employed to classify the gestures. The recognized signs are displayed as text on the screen and converted into speech through a text-to-speech (TTS) engine, making communication more seamless.

For an improved user experience, the system includes an interactive dashboard that offers real-time hand-tracking visualization, prediction history, user feedback for correction, and performance analysis, including accuracy rates and response times. It also allows for a comparative evaluation of recognition accuracy across ASL, ISL, and BSL.

Designed for scalability, the system has the potential to support additional sign languages in the future. By combining computer vision, deep learning, and an intuitive user interface, this project introduces an innovative assistive technology that enhances communication for the deaf and hard-of-hearing community.

I. INTRODUCTION

Communication challenges between deaf individuals and those who do not understand sign language continue to be a major obstacle. While sign language serves as the primary means of communication for the deaf community, its complexity makes it difficult for non-signers to interpret. Automated sign language recognition systems can help bridge this gap by translating gestures into text and speech in real-time.

This project presents a deep learning-based system designed to recognize multiple sign languages, including American Sign Language (ASL), Indian Sign Language (ISL), and British Sign Language (BSL), using a webcam. The system captures live video input, detects hand landmarks, and classifies gestures using a Convolutional Neural Network (CNN) model. The identified signs are then displayed as text and converted into speech, making communication more accessible. Additionally, an interactive dashboard provides real-time analytics, allows user feedback, and tracks performance metrics.

The main contributions of this work include: A real-time sign language recognition system that supports multiple sign languages. Integration of MediaPipe for precise hand landmark detection. A CNN-based model for accurate gesture classification. A Power BI dashboard to monitor performance and enhance user interaction. A scalable framework that allows for the future inclusion of additional sign languages.

II. LITERATURE SURVEY

Wu, Y., & Ouhyoung, M. (1999). "A Survey on Vision-Based Hand Gesture Recognition." This paper provides an overview of various computer vision techniques for recognizing hand gestures. It examines different methods used for extracting features and classifying gestures.

Rautaray, S. S., & Agrawal, A. (2015). "Vision-Based Hand Gesture Recognition for Human-Computer Interaction." This study focuses on the use of vision-based gesture recognition for human-computer interaction. It outlines different machine learning approaches for gesture classification and their practical applications.

Simonyan, K., & Zisserman, A. (2014). "Very Deep Convolutional Networks for Large-Scale Image Recognition." This research introduces deep convolutional neural network (CNN) architectures that significantly improve image recognition. The insights from this study are valuable for applying CNNs to sign language recognition.

Chollet, F. (2017). "Xception: Deep Learning with Depthwise Separable Convolutions." This paper presents an advanced deep learning model that enhances the efficiency of image classification. The concepts discussed help in designing high-accuracy sign language recognition systems.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). "Deep Residual Learning for Image Recognition." This research introduces the ResNet architecture, which enhances the training of deep neural networks. The study demonstrates how residual learning techniques improve the accuracy of image classification models, making them useful for gesture recognition.

Google AI Research. (2020). "MediaPipe: A Framework for Building Perception Pipelines." This paper explores MediaPipe, a framework that enables real-time hand tracking and gesture recognition. It explains how landmark detection enhances applications that rely on gesture-based inputs.

Kingma, D. P., & Ba, J. (2014). "Adam: A Method for Stochastic Optimization." This study introduces the Adam optimization algorithm, which is widely used in training deep learning models. The research highlights how Adam improves model convergence and efficiency in training CNNs for sign language recognition.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). "Deep Learning." This book provides a comprehensive introduction to deep learning concepts, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks. These concepts are essential for developing sign language recognition systems.

IEEE Transactions on Pattern Analysis and Machine Intelligence. (2021). "Advances in Gesture Recognition." This research examines the latest advancements in gesture recognition using artificial intelligence and deep learning. It discusses various algorithms and datasets that contribute to improved accuracy in sign language recognition.

III. METHODOLOGY

III.1. Dataset Collection and Preprocessing

The initial stage in building the sign language recognition system involves collecting datasets that include images and video sequences of hand gestures from American Sign Language (ASL), Indian Sign Language (ISL), and British Sign Language (BSL). These datasets are obtained from publicly available sources and further enhanced using data augmentation techniques such as flipping, rotation, and brightness adjustments. These modifications help improve the model's ability to generalize across different conditions. Each image undergoes preprocessing steps, including grayscale conversion, normalization, and resizing, to ensure uniform input dimensions for the deep learning model.

III.2. Hand Landmark Detection and Feature Extraction

To effectively capture and recognize hand gestures, the system employs MediaPipe, a machine learning framework designed for real-time hand tracking. MediaPipe identifies key hand landmarks, such as finger positions and palm orientation, to extract relevant features for classification. This approach reduces computational load by focusing on essential hand movements rather than processing the entire image, thereby improving recognition accuracy.

III.3. Model Training and Classification

Once hand landmarks are extracted, they are fed into a Convolutional Neural Network (CNN), a deep learning model optimized for image classification. The CNN is trained on the preprocessed dataset to recognize unique patterns in different sign language gestures. Techniques such as the Adam optimizer and cross-entropy loss function are used to enhance model efficiency and accuracy. After training, the model is capable of making real-time predictions, categorizing gestures into their respective sign language classes.

III.4. Real-Time Prediction and Interactive Dashboard

After the system recognizes a sign, it is converted into both text and speech using a Text-to-Speech (TTS) engine, making communication more accessible. To provide real-time insights, a Power BI dashboard is integrated, displaying system performance metrics such as recognition accuracy, response time, and prediction history. The dashboard also allows users to compare accuracy levels across ASL, ISL, and BSL, track trends, and offer feedback for further model improvements. This interactive design enhances user experience while ensuring the system remains scalable for future enhancements.

IV. EXISTING SYSTEM

IV.1. Conventional Methods of Communication for the Deaf

Deaf and hard-of-hearing individuals primarily communicate with the hearing community through human interpreters, text-based messaging, and pre-recorded sign language videos. While these approaches improve accessibility to some extent, they have notable limitations. Human interpreters are not always available, text messaging lacks the expressiveness of sign language, and pre-recorded videos do not allow for real-time conversations, making spontaneous communication difficult.

IV.2. Challenges with Existing Gesture Recognition Systems

Some sign language recognition technologies rely on glove-based sensors that track hand movements. However, these devices are often costly, uncomfortable to wear for extended periods, and require frequent calibration. Other systems utilize computer vision techniques, but they struggle with environmental factors such as lighting conditions, variations in hand positioning, and background clutter. As a result, recognition accuracy can be inconsistent. Additionally, many of these models are designed for only one sign language, such as ASL, limiting their applicability on a global scale.

IV.3. Absence of Real-Time and Interactive Capabilities

Many current sign language recognition systems lack real-time processing, making them unsuitable for natural conversations. Additionally, they do not incorporate interactive features such as user feedback mechanisms, prediction history tracking, or performance monitoring dashboards. The absence of these features reduces their effectiveness and usability. This underscores the necessity for an AI-powered system that can recognize multiple sign languages in real time while offering an interactive and user-friendly experience.

V. The Proposed Sign Language Recognition System

The proposed system addresses these challenges by developing an AI-powered, real-time sign language recognition solution that supports American Sign Language (ASL), Indian Sign Language (ISL), and British Sign Language (BSL). Using advanced deep learning and computer vision techniques, the system accurately identifies and classifies hand gestures, converting them into both text and speech. To enhance the user experience, a Power BI dashboard provides real-time performance tracking and recognition history visualization.

V.1. Recognition of Multiple Sign Languages

Unlike traditional systems that focus on a single sign language, this solution is designed to support ASL, ISL, and BSL, ensuring broader accessibility. A Convolutional Neural Network (CNN) is trained on a diverse dataset that includes gestures from these languages, ensuring high recognition accuracy. The system is also designed to be scalable, allowing additional sign languages to be incorporated in the future.

V.2. Real-Time Gesture Detection and Classification

The system leverages MediaPipe, an advanced computer vision framework, to track and analyze hand landmarks in real time. These extracted features are processed by a CNN model that accurately classifies gestures. Once classified, the recognized sign is displayed as text on the screen and converted into speech using a Text-to-Speech (TTS) engine, enabling fluid communication between deaf individuals and non-signers.

V.3. Performance Monitoring Through Power BI Dashboard

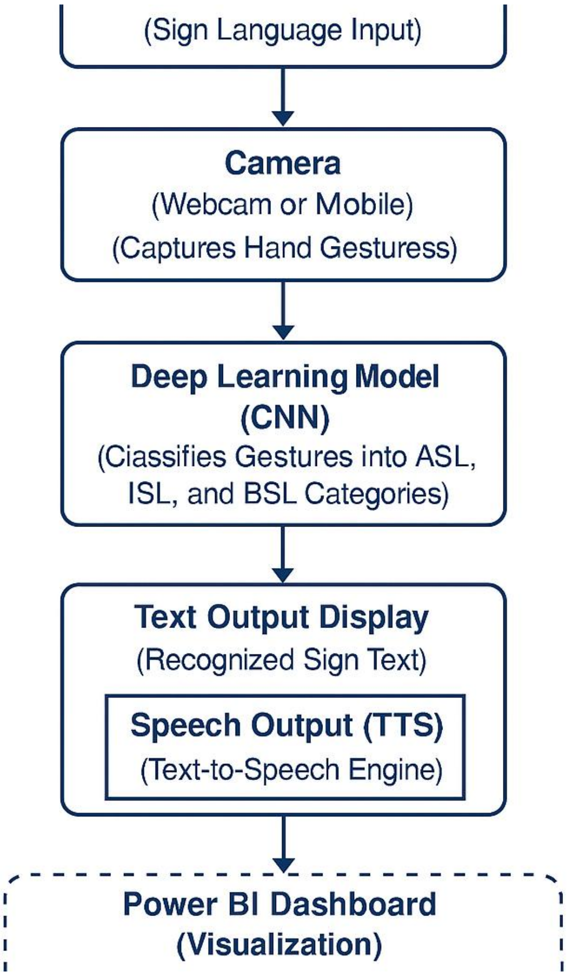
To provide users with valuable insights into system performance, an interactive Power BI dashboard is integrated. This dashboard includes: Tracking recognition accuracy and response time, Visualization of prediction history, User feedback collection for continuous model improvement, Comparison of recognition accuracy across ASL, ISL, and BSL

This ensures transparency and allows the system to be refined based on real-world user data.

V.4. Scalability and Future Developments

To improve accessibility and efficiency, the system is designed with future advancements in mind, including: Expanding support for additional sign languages, Enhancing model accuracy through continuous training with updated datasets, Optimizing deep learning algorithms for faster real-time processing, Integrating the system with mobile applications for wider accessibility, By implementing these improvements, the system aims to provide an inclusive and effective communication tool for the deaf and hard-of-hearing community.

VI. ARCHITECTURE DIAGRAM



EXPLANATION

User Input – The system captures hand gestures through a camera, which can be a webcam or a mobile device.

Hand Landmark Detection – Using the MediaPipe framework, the system identifies key hand landmarks and extracts important features necessary for gesture recognition.

Gesture Classification – A deep learning model based on Convolutional Neural Networks (CNN) processes the detected hand movements and classifies them into American Sign Language (ASL), Indian Sign Language (ISL), or British Sign Language (BSL).

Output Display – The recognized sign is displayed as text on-screen, and if required, it is also converted into speech using a Text-to-Speech (TTS) engine to facilitate communication.

Performance Monitoring – The system records predictions, accuracy levels, and user feedback, which are displayed on a Power BI dashboard. This helps in refining the model and improving future performance.

VII. FUTURE SCOPE

The sign language recognition system has significant potential for expansion and improvement. As technology advances, several future developments can enhance its functionality:

Support for Additional Sign Languages – The system can be trained to recognize more sign languages, making it inclusive for a wider range of users.

Enhanced Accuracy and Efficiency – Further advancements in deep learning models could improve recognition accuracy and reduce processing time.

Mobile and Smart Device Integration – Making the system accessible through mobile applications and smart devices would provide users with greater flexibility and convenience.

Advanced Data Analytics – The Power BI dashboard could be enhanced with more detailed analytics, helping developers understand user patterns and improve system performance.

Customization and Interactive Feedback – Allowing users to personalize their experience and provide real-time feedback could make the system more user-friendly and adaptable to individual needs.

VIII. REFERENCE

Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. This paper details the architecture of deep convolutional networks, providing a foundation for understanding image-based recognition tasks.

Chollet, F. (2017). Xception: Deep Learning with Depthwise Separable Convolutions. This work presents an efficient deep learning architecture that has influenced modern CNN design, relevant for gesture recognition models.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. Introducing the ResNet model, this research addresses challenges in training deep networks, which is applicable in improving sign language prediction accuracy.

Wu, Y., & Ouhyoung, M. (1999). A Survey on Vision-Based Hand Gesture Recognition. This survey examines various approaches to hand gesture recognition using computer vision, laying the groundwork for later deep learning methods.

Rautaray, S. S., & Agrawal, A. (2015). Vision-Based Hand Gesture Recognition for Human-Computer Interaction. This study reviews vision-based methods for gesture recognition, focusing on applications in interactive systems.

Google AI Research. (2020). MediaPipe: A Framework for Building Perception Pipelines. MediaPipe is crucial for real-time hand tracking, providing the landmark detection that supports gesture-based systems.

Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. The Adam optimizer is widely used in training deep learning models and is instrumental in improving convergence in CNNs used for sign language recognition.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. This comprehensive resource covers essential deep learning concepts and techniques that are central to developing AI-driven recognition systems.

IEEE Transactions on Pattern Analysis and Machine Intelligence. (2021). Advances in Gesture Recognition. This publication highlights recent progress in gesture recognition technologies and methodologies, offering insights into state-of-the-art approaches.

Journal of Artificial Intelligence Research. (2022). Enhancing Assistive Technology with Deep Learning. This article explores the role of deep learning in improving assistive communication technologies, underscoring its impact on accessibility solutions for the deaf community.