**Malnad College of Engineering, Hassan**



(An Autonomous Institution affiliated to VTU, Belgavi)

Mini Project Report On

**“Real time sign language to text conversion with emergency assistanc”**

*Submitted in partial fulfillment of*

*the requirements for the award of the degree of*

**Bachelor of Engineering in Computer Science and Engineering**

Submitted by

Ankitha K N 4MC22CS0

Ashwini D M 4MC22CS019 Preksha H S 4MC22CS

Vikas B 4MC22CS

Under the guidance of

**Dr. Keerthi Kumar H M**

Associate Professor

**Department of Computer Science and Engineering 2024-2025**

|  |  |  |
| --- | --- | --- |
| image001 | **Malnad College of Engineering**  **Department of Computer Science and Engineering** | C:\Users\STUDENT\AppData\Local\Microsoft\Windows\INetCache\Content.Word\image.png |

*Certificate*

Certified that the Main project phase-1 work entitled **“Real time sign language to text conversion with emergency assistance”** is a bonafide work carried out by **Ankitha K N (4MC22CS0), Ashwini D M (4MC22CS019), Preksha H S (4MC22CS), Vikas B (4MC22CS)** in partial fulfillment for the award of Bachelor of Engineering in Computer Science and Engineering of the Visvesvaraya Technological University, Belgavi during the year 2024-2025. The project report has been approved as it satisfies the academic requirements in respect of main project work prescribed for the Bachelor of Engineering Degree.

|  |  |  |
| --- | --- | --- |
| Signature of the Guide | Signature of the HOD | Signature of the Principal |
| *Dr. Keerthi Kumar H M* | *Dr. Chandrika Ramesh* | *Dr. A. J. Krishnayya* |
| Associate Professor | Professor and Head | Principal |
| Dept. of CSE, MCE | Dept. of CSE, MCE | MCE |

Examiners

Name of the Examiner Signature of the Examiner

**ABSTRACT**

Communication is a fundamental human need, and for individuals with speech and hearing impairments, sign language plays a vital role. Indian Sign Language (ISL), characterized by complex, dynamic gestures and regional variations, poses significant challenges in real-time recognition due to the lack of standardized datasets and gesture diversity. Recent advancements in deep learning and computer vision have enabled the development of automated ISL recognition systems. This survey reviews fifteen peer-reviewed research works focusing on gesture recognition using MediaPipe, BiLSTM, and CNN architectures, as well as sentence prediction through transformer-based NLP models like T5. Notably, gesture recognition models achieved accuracies above 98%, while BLEU scores for sentence formation reached 0.81. In addition, hybrid models employing YOLOv5 and CNN for emergency detection achieved mAP values up to 99.6%, with alert systems integrated via APIs like Fast2SMS for real-time SMS notifications. Despite these advancements, most existing systems specialize in only one functionality—either gesture recognition, sentence generation, or emergency alerts. This project aims to integrate all three components into a single, real-time, robust ISL-to-text conversion system with emergency assistance. The findings of this survey serve as a foundation for designing inclusive and responsive assistive communication technology.

**ACKNOWLEDGEMENTS**

Salutations to our beloved and highly esteemed institute, **Malnad College of Engineering**, for its excellent academic environment, well-qualified faculty, and laboratories equipped with modern facilities essential for technical development.

The successful completion of any task is never the result of individual effort alone. It is a product of guidance, support, and encouragement from several individuals. We take this opportunity to extend our sincere gratitude to all those who helped us complete our project and this report.

We would like to express our deepest gratitude to our mentor, **Dr. Keerthi Kumar H M**, Associate Professor, Department of **Computer Science and Engineering**, for her constant encouragement, invaluable guidance, and unwavering support throughout the project. Her insightful feedback and expert supervision played a vital role in shaping the quality and direction of our work. We consider it a privilege to have worked under her mentorship, and we have gained much more than just technical knowledge through her guidance.

We are also immensely thankful to **Dr. A. J. Krishnayya**, Principal, Malnad College of Engineering, for providing us with the opportunity and necessary facilities to carry out our project. His leadership and support created a conducive environment for academic research and innovation.

We also express our heartfelt appreciation to all the faculty members and staff of the **Department of Computer Science and Engineering** for their constant help, motivation, and encouragement throughout this journey.

Ankitha K N

Ashwini D M

Preksha H S

Vikas B

**TABLE OF CONTENTS**

1. [Introduction](#_bookmark0) 1
   1. [Introduction to Leukemia Detection 1](#_TOC_250002)
   2. [About Project](#_bookmark1) 2
      1. [Problem Statement 2](#_TOC_250001)
      2. [Objective 2](#_TOC_250000)

|  |  |  |
| --- | --- | --- |
| **2** [**Literature Survey**](#_bookmark2) | **4** |  |
| 2.1 Review of Relevant Literature. . . . . . . . . . . . . . . . . . |  | 4 |
| 2.2 Current Models, Methods, and Methodologies . . . . . . . . . | 4 |  |

2.3 Dataset Description . . . . . . . . . . . . . . . . . . . . . 4

**3 Conclusion 10**

**References 11**

**Chapter 1 Introduction**

# Introduction to Sign Language Recognition and Emergency Assistance

Sign language serves as a vital mode of communication for individuals with speech and hearing impairments, enabling them to express thoughts and interact with others when verbal communication is not possible. Among the various systems, Indian Sign Language (ISL) is widely used in India and is characterized by complex, dynamic hand gestures, often involving both hands. However, recognizing ISL poses significant challenges due to regional variations, gesture ambiguity, and the lack of standardized datasets.

Recent advancements in computer vision and deep learning have introduced new possibilities for automating gesture recognition, sentence generation, and emergency response. Tools such as MediaPipe, BiLSTM, and transformer-based Natural Language Processing (NLP) models have shown high accuracy in recognizing gestures and forming meaningful sentences. Furthermore, integrating real-time emergency alert systems using models like YOLOv5 and APIs such as Fast2SMS enables immediate assistance during critical situations. These technologies collectively pave the way for a unified assistive solution that addresses both communication and safety concerns.

# About Project

## Problem Statement

Individuals with speech and hearing impairments face significant barriers in communication and are especially vulnerable during emergencies. Existing sign language systems often lack real-time capabilities, contextual sentence generation, or integrated emergency assistance. This limitation creates an urgent need for a robust and inclusive system that can recognize dynamic ISL gestures in real time, generate coherent text output, and respond promptly to emergency gestures with appropriate alerts.

## Objective

 To develop an automated system for **real-time recognition of Indian Sign Language (ISL)** gestures using deep learning techniques.

 To implement **MediaPipe and BiLSTM** for accurate detection and classification of dynamic hand gestures.

 To integrate a **transformer-based NLP module (T5)** for generating **contextual and grammatically correct sentences** from recognized sign sequences.

 To incorporate an **emergency gesture detection mechanism** using **CNN and YOLOv5,** enabling prompt identification of critical situations.

 To establish a **real-time SMS alert system** through APIs like **Fast2SMS** for emergency notifications with minimal delay.

 To design a **user-friendly, interactive interface** that ensures smooth communication and system usability **on consumer-grade hardware**.

 To improve accessibility, inclusivity, and safety for individuals with hearing and speech impairments through a **unified assistive communication platform**.

Chapter 2 **Literature Survey**

* 1. **Review of Relevant Literature**

Real-time sign language recognition has emerged as a crucial application in assistive technology, particularly for bridging communication gaps faced by individuals with speech and hearing impairments. The field has significantly evolved over the past decade, with advancements in deep learning, natural language processing (NLP), and computer vision. Numerous studies have contributed to the development of gesture recognition systems capable of converting Indian Sign Language (ISL) into meaningful text or speech.

Traditional machine learning approaches such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF) were initially used for classifying static gestures by extracting handcrafted features like contour shapes, orientation histograms, and geometric descriptors. However, these methods were limited in their ability to process continuous or dynamic gestures and often struggled with background noise and signer variation.

To overcome these limitations, researchers shifted toward deep learning models—especially Convolutional Neural Networks (CNNs)—which can automatically extract spatial features from hand images and videos. CNNs have been successfully used for detecting both single-hand and double-hand gestures with improved accuracy and robustness. More recent models like BiLSTM (Bidirectional Long Short-Term Memory) have shown great promise in processing sequential gesture data, capturing both past and future temporal dependencies, which is vital for recognizing dynamic ISL sequences.

Moreover, transformer-based NLP models such as T5 have been adopted to enhance the contextual understanding of identified gestures and generate coherent, grammatically correct sentences. Rashmi and Lalita achieved BLEU scores above 0.8 using such architectures, indicating effective sentence generation from sign inputs.

In the domain of emergency gesture detection, object detection models like YOLOv5 combined with CNN classifiers have shown high performance, achieving mean average precision (mAP) scores exceeding 99%. These systems have been integrated with SMS APIs such as Fast2SMS to send real-time alerts, proving effective in critical scenarios like fire, theft, or accidents.

Despite these advancements, most existing systems focus only on isolated components—either gesture recognition or emergency detection. Few provide an integrated, end-to-end pipeline capable of real-time ISL-to-text conversion along with emergency response functionality. Additionally, challenges such as signer variability, limited datasets, inconsistent lighting conditions, and lack of generalized models still hinder real-world deployment.

In summary, the reviewed literature strongly supports the use of deep learning and NLP models in ISL recognition, with BiLSTM, CNN, and T5 transformers forming the foundation of high-accuracy systems. However, the integration of emergency assistance and context-aware sentence prediction in a unified, real-time solution remains an open research opportunity—one which this project aims to address.

* 1. **Current Models, Methods, and Methodology**

Various machine learning (ML) and deep learning (DL) techniques have been developed for sign language recognition, particularly for Indian Sign Language (ISL) translation systems. Traditional ML models such as Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Random Forest (RF) rely on manually extracted features such as hand contour shapes, orientation histograms, or edge-based descriptors. While these methods are interpretable and computationally efficient, they often lack the robustness to handle continuous or complex gesture sequences, especially under diverse lighting and background conditions.

Recent progress in deep learning has led to the adoption of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) for automatic feature extraction and sequential learning from raw video frames. CNN-based architectures are widely used for spatial pattern recognition in static gestures, while RNN variants like Bidirectional Long Short-Term Memory (BiLSTM) models are employed for modeling temporal dependencies in dynamic gestures. This dual approach enhances recognition accuracy and adaptability across different users and hand motions.

For sentence-level translation, transformer-based Natural Language Processing (NLP) models such as T5 have been fine-tuned to generate grammatically correct sentences from a sequence of recognized gestures. This integration adds context-awareness and linguistic fluency to the system, a significant improvement over word-by-word output.

In addition to gesture translation, the proposed system incorporates real-time emergency detection using a hybrid model combining CNN with YOLOv5. YOLOv5 offers rapid and accurate detection of predefined emergency gestures, and APIs such as Fast2SMS are utilized to deliver immediate SMS alerts to designated contacts in critical situations.

Preprocessing techniques—such as background removal, hand landmark detection using MediaPipe, and normalization of hand joint coordinates—play a crucial role in improving system performance and model stability. These methods reduce noise and increase sensitivity to subtle gesture variations, enabling more reliable recognition in real-time scenarios.

Data augmentation techniques such as mirroring, rotation, and lighting variation are applied to artificially expand training datasets and improve model generalization. However, challenges such as signer variability, inconsistent gesture speeds, limited labeled datasets, and generalization across different physical environments remain active areas for further research.

In summary, the combination of CNNs, BiLSTM networks, and NLP transformers forms the backbone of modern sign language systems. The proposed integration of emergency detection and real-time SMS alerting introduces a novel, unified pipeline aimed at enhancing accessibility and safety for users with speech and hearing impairments.

* 1. **Dataset Description**

The performance and generalization capability of sign language recognition systems largely depend on the quality, size, and diversity of datasets used for training and evaluation. For Indian Sign Language (ISL), dataset availability is limited, and most existing datasets suffer from small sample sizes, lack of signer diversity, and inconsistent gesture representation. However, several datasets have been used or created in past research efforts, each contributing unique features to the development of gesture recognition and emergency alert systems.

Below are the most commonly referenced datasets in ISL research:

**1. Adithya & Rajesh Emergency ISL Dataset**

* Description: A curated video dataset capturing ISL emergency gestures such as "help", "thief", "pain", and "fire".
* Usage: Specifically developed for training emergency detection modules in real-world scenarios.
* Highlights: Enables the development of responsive systems for critical conditions.
* Limitation: Limited signer diversity and background variation.

**2. DeepSign Dataset (Kothadiya et al.)**

* Description: Includes video sequences for 11 ISL signs with frame-wise annotation.
* Model Tested: GRU + LSTM hybrid architecture.
* Accuracy Achieved: ~97% on single signer data.
* Limitation: Does not include emergency gestures or sentence-level labeling.

**3. MediaPipe Hand Landmark Dataset (Generated In-House)**

* Description: Hand landmark points (21 keypoints per hand) extracted using MediaPipe Holistic framework.
* Usage: Serves as input for BiLSTM-based gesture classification.
* Advantage: Enables frame-wise modeling without relying on full image classification.
* Flexibility: Can be generated in real-time, allowing for continuous data expansion.

**4. Custom ISL Dataset (Self-Curated)**

* Description: A tailored dataset recorded by the project team to train the BiLSTM model. Includes multiple gesture classes across different users and lighting conditions.
* Scope: Covers both commonly used ISL gestures and pre-defined emergency actions.
* Format: Sequences of MediaPipe landmarks with associated gesture labels.
* Limitation: Limited number of participants; ongoing expansion is planned.

**5. Kaggle Gesture Datasets (Non-ISL)**

* Description: Publicly available hand gesture datasets on Kaggle used for initial experimentation.
* Limitation: Do not follow ISL standards or include Indian-specific gestures.

In conclusion, dataset limitations remain a major bottleneck for ISL recognition research. This project uses a hybrid approach, combining real-time MediaPipe landmark extraction with self-curated gesture samples to overcome data scarcity and enable real-time adaptability. For emergency detection, the use of Adithya & Rajesh’s dataset provides a foundation for building robust, responsive alert systems.

**Chapter 6 Conclusion**

This survey highlights the rapid progress and effectiveness of deep learning and natural language processing techniques in real-time sign language recognition and emergency assistance systems. Traditional machine learning models such as SVM and KNN have shown moderate performance when applied to isolated gestures using handcrafted features. However, the shift toward deep learning architectures—particularly CNNs for spatial feature extraction and BiLSTM for temporal gesture analysis—has significantly improved recognition accuracy and system reliability.

The integration of transformer-based NLP models like T5 has enabled contextual sentence generation, bridging the gap between gesture recognition and natural language communication. Additionally, the use of object detection algorithms such as YOLOv5, in combination with real-time messaging APIs like Fast2SMS, has made emergency gesture detection and response both practical and timely.

Preprocessing steps like hand landmark extraction via MediaPipe and structured gesture data collection have proven essential for building robust models. Despite challenges such as dataset limitations, signer variability, and environment-dependent performance, the findings from this survey provide a strong foundation for implementing an inclusive, real-time ISL recognition system.

While the current phase is limited to a literature survey and methodology design, the insights gathered point toward the feasibility of developing a unified platform that supports accurate gesture translation, sentence generation, and emergency alerting. Future work will focus on dataset expansion, multi-signer support, and deployment on mobile platforms to enhance accessibility for users in real-world scenarios.

**References**

1. A. I. Singh, B. Mathai, S. Silas, and J. B. Princess, “Real-Time Sign Language Translator for Deaf and Mute,” in 2023 IEEE International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), December 2023.
2. R. Gaikwad and L. Admuthe, “Sign Language Recognition of Words and Sentence Prediction Using LSTM and NLP,” Journal of Theoretical and Applied Information Technology, vol. 102, no. 4, February 2024.
3. S. Daga, A. Dusane, and D. Bobby, “With You – Indian Sign Language Detection and Alert System,” in 2024 IEEE Conference on Smart Computing and Informatics (ESCI), Pune, India.
4. Q. M. Areeb, M. Nadeem, R. Alroobaea, and F. Anwer, “Deep Learning Based Hand Gesture Recognition for Emergency Situation: A Study on Indian Sign Language,” in 2021 International Conference on Data Analytics for Business and Industry (ICDABI), pp. 245–250, 2021.
5. Q. M. Areeb and M. Nadeem, “Helping Hearing-Impaired in Emergency Situations: A Deep Learning-Based Approach,” International Journal of Scientific Research in Computer Science, vol. 12, no. 1, February 2024.
6. T. Kemkar, V. Rai, and B. Verma, “Sign Language to Text Conversion using Hand Gesture Recognition,” in Proceedings of the 8th International Conference on Communication and Electronics Systems (ICCES), pp. 1580–1587, 2023. DOI: 10.1109/ICCES57224.2023.10192820.
7. J. C. Kavitha and D. Nagarajan, “Real Time Automated Sign Language Recognition and Transcription with Audio Feedback,” in Proceedings of the IEEE International Carnahan Conference on Security Technology (ICCST), pp. 1–6, 2023. DOI: 10.1109/ICCST59048.2023.10474276.
8. U. Farooq, M. S. M. Rahim, N. Sabir, A. Hussain, and A. Abid, “Advances in Machine Translation for Sign Language: Approaches, Limitations, and Challenges,” Neural Computing and Applications, vol. 33, pp. 14357–14399, 2021. DOI: 10.1007/s00521-021-06079-3.
9. N. Buckley, L. Sherrett, and E. L. Secco, “A CNN Sign Language Recognition System with Single and Double-Handed Gestures,” in Proceedings of the 45th Annual IEEE Computer Software and Applications Conference (COMPSAC), pp. 1–8, 2021. DOI: 10.1109/COMPSAC51774.2021.0017.
10. D. V. Padmaja, V. Meghana, T. Satvika, P. N. Rohitha, and Ch. V. D. Durga, “Indian Sign Language Recognition Using OpenCV,” International Journal of Engineering Science and Advanced Technology (IJESAT), vol. 24, no. 5, pp. 362–364, May 2024.
11. M. K. N. B. Mahesh Kumar, “Conversion of Sign Language into Text,” International Journal of Applied Engineering Research, vol. 13, no. 9, pp. 7154–7161, 2018.
12. J. Rekbai, J. Bhattacharya, and S. Majumder, “Shape, Texture and Local Movement Hand Gesture Features for Indian Sign Language Recognition,” in Proceedings of the IEEE International Conference on Emerging Trends in Robotics and Communication Technologies (INTERACT), CMERI–CSIR, Durgapur, India, pp. 30–35, 2011. DOI: 10.1109/INTERACT.2011.5975432.
13. V. Adithya and R. Rajesh, “Hand Gestures for Emergency Situations: A Video Dataset Based on Words from Indian Sign Language,” Data in Brief, vol. 31, p. 106016, July 2020. DOI: 10.1016/j.dib.2020.106016.
14. D. Kothadiya, C. Bhatt, K. Sapariya, K. Patel, A.-B. Gil-González, and J. M. Corchado, “Deepsign: Sign Language Detection and Recognition Using Deep Learning,” Electronics, vol. 11, no. 11, p. 1780, June 2022. DOI: 10.3390/electronics11111780.
15. A. O. Hashi, S. Z. M. Hashim, and A. B. Asamah, “A Systematic Review of Hand Gesture Recognition: An Update from 2018 to 2024,” IEEE Access, vol. 12, pp. 143599–143618, July 2024. DOI: 10.1109/ACCESS.2024.3421992.