Global Research journal of Natural Science and Technology (Grjnst)
Volume: 03 - Issue 3 (2025)
p.ISSN: 2790-7643 e.ISSN: 2790-7651
www.grjnst.com

An Intelligent Sign Language Interpretation System for Patient Communication

Syed Muhammad Hassan Zaidi

Department of AI and Mathematical Sciences, Sindh Madressatul Islam University, Karachi, Pakistan

m.hassan@smiu.du.pk

Syed Asif Ali

Department of AI and Mathematical Sciences, Sindh Madressatul Islam University, Karachi, Pakistan

aasyed@smiu.edu.pk

Imtiaz Hussain

Department of AI and Mathematical Sciences, Sindh Madressatul Islam University, Karachi, Pakistan

imtiaz@smiu.edu.pk

Basit Hasan

Department of Software Engineering, Sindh Madressatul Islam University, Karachi, Pakistan

basithq@smiu.edu.pk

Muhammad Ali

Department of AI and Mathematical Sciences, Sindh Madressatul Islam University, Karachi, Pakistan

mali@smiu.edu.pk

Corresponding author: m.hassan@smiu.du.pk **(Syed Muhammad Hassan Zaidi)**

Abstract: *Effective communication is a cornerstone of high-quality healthcare; however, individuals who are deaf or hard of hearing frequently encounter significant barriers in clinical environments due to the absence of qualified interpreters and limited proficiency in sign language among healthcare providers. To address this challenge, this study proposes a real-time sign language interpretation system that translates hand gestures into textual outputs, thereby facilitating inclusive and equitable patient–provider communication. The system combines computer vision and deep learning techniques, utilizing OpenCV for image preprocessing and MediaPipe for extracting 21-point hand landmarks. The gesture recognition model was trained using Google's Teachable Machine on a curated dataset of seven medically relevant gesture classes (e.g., Doctor, Cough, Injection), with data augmentation and normalization enhancing robustness. The final model, based on convolutional neural networks (CNNs), achieved an overall accuracy of 82%, with the Doctor gesture attaining a precision of 0.93 and Chest gesture showing 97.31% recognition accuracy in real-time tests. The modular architecture, ease of deployment, and strong performance across diverse environmental conditions make it a practical solution for medical settings. This research contributes to the advancement of assistive technologies in healthcare and underscores the role of AI-driven tools in promoting accessible, responsive, and high-quality patient care for the deaf and hard-of-hearing community.*

Key words: *Sign Language Recognition, Real-Time Gesture Translation, Assistive Communication Technology, Medical Human-Computer Interaction, Hand Landmark Detection, MediaPipe, Teachable Machine, Convolutional Neural Networks (CNN), Deaf and Hard-of-Hearing Accessibility, AI in Healthcare.*

Introduction

Sign language recognition (SLR) has emerged as a pivotal domain in the advancement of assistive technologies, aiming to mitigate communication barriers between deaf or hard-of-hearing individuals and the hearing population. Over the past decade, numerous deep learning frameworks have been proposed to enhance the accuracy, speed, and adaptability of gesture recognition systems, particularly for real-time applications.

Recent studies have demonstrated significant progress in continuous sign language recognition (CSLR). Ahn, Jang, and Chung (2023) introduced a two-pathway SlowFast network, incorporating Bi-directional Feature Fusion (BFF) and Pathway Feature Enhancement (PFE), which collectively enable simultaneous extraction of spatial and temporal features. Their model achieved state-of-the-art performance on benchmark datasets such as PHOENIX14 and CSL-Daily (Anderson et al., 2017). Similarly, Ryumin, Ivanko, and Ryumina (2023) proposed a multimodal architecture that integrates audio-visual cues captured via mobile device sensors, achieving recognition accuracies exceeding 98%. This approach underscores the potential of multimodal fusion for enhancing responsiveness in portable healthcare technologies.

To address the persistent challenge of recognizing visually similar signs (VISigns), Zuo, Wei, and Mak (2023) presented a Natural Language-Assisted SLR framework with label smoothing and inter-modality mix-up strategies. Their work improved generalization and robustness across complex datasets like MSASL and WLASL. In parallel, Hu et al. (2021) proposed SignBERT, a BERT-inspired transformer that encodes hand pose sequences as language tokens, enabling advanced transfer learning capabilities. Their subsequent work, SignBERT+ (Hu et al., 2023), integrated self-supervised pretraining informed by hand model awareness, further enhancing performance in settings with limited labelled data, a scenario frequently encountered in regional or domain-specific sign languages.

Efficiency in deployment has also been a focus in recent research. Boháček and Hrúz (2022) developed a lightweight transformer-based model utilizing normalized 2D landmarks, optimized for low-power, real-time environments such as mobile or embedded devices. Shen, Zheng, and Yang (2024) introduced StepNet, a spatial-temporal part-aware network employing both RGB and skeletal data, which demonstrated improved recognition accuracy by leveraging multimodal fusion.

Efforts to expand SLR beyond well-resourced languages like ASL are equally noteworthy. Kapitanov et al. (2023) constructed "Slovo," a crowd-sourced Russian Sign Language dataset comprising 20,000 annotated videos across 1,000 classes. Such initiatives are instrumental in supporting underrepresented sign languages and facilitating culturally inclusive machine learning applications.

A comprehensive systematic review by Wadhawan and Kumar (2021) revealed that most prior efforts concentrated on static, isolated sign recognition using camera-based approaches. Their findings emphasized the necessity for real-time, context-aware systems capable of adapting to variable environments, particularly in high-stakes domains like clinical care. Complementing this, Haldera and Tayade (2021) developed a vernacular, sensor-free recognition system based on Google's MediaPipe and support vector machines (SVM), achieving 99% accuracy in real-time translation across multiple languages, making it well-suited for deployment in healthcare settings.

Collectively, these developments indicate a growing emphasis on designing real-time, lightweight, and multimodal sign recognition systems that maintain robustness in diverse and unpredictable environments. Building upon this foundational work, the present study proposes a domain-specific SLR system tailored to the healthcare sector. Leveraging MediaPipe for landmark detection and Google's Teachable Machine for intuitive model training, this system is designed to interpret medical-related gestures in real time. The goal is to ensure efficient, scalable, and accurate patient-provider communication, thereby contributing to more inclusive and accessible healthcare delivery.

Methodology

The development of the proposed sign language interpretation system for deaf and hard-of-hearing individuals follows a structured, multi-phase pipeline designed to accurately translate hand gestures into machine-readable formats, culminating in real-time textual or spoken outputs. The methodology is informed by contemporary machine learning practices and optimized for healthcare-oriented deployment.

1. Data Acquisition and Frame Extraction

The initial phase involves the collection of video recordings demonstrating a range of medical-related sign language gestures. These videos are captured from multiple angles and under varied lighting conditions to ensure the diversity and robustness of the dataset. Subsequently, each video is decomposed into individual image frames using automated Python scripts, forming the primary corpus for feature extraction and training.

2. Data Preprocessing and Augmentation

To enhance model generalizability and prevent overfitting, the extracted frames undergo systematic preprocessing. This includes:

- **Data augmentation**: Applying geometric transformations such as rotation, horizontal flipping, scaling, and brightness adjustments to synthetically expand the dataset and simulate real-world variability.
- **Noise filtering**: Removing blurred or low-quality frames to retain only high-resolution, information-rich samples.
- **Landmark detection**: Using key-point detection algorithms—specifically via the MediaPipe framework—to extract and encode hand landmarks (x, y coordinates) that represent finger joints and palm contours. These features form the basis for classification tasks.

3. Feature Normalization and Dataset Structuring

The hand landmark data is normalized to a standard scale to ensure uniform feature distribution across samples. The wrist joint (index 0) is treated as the origin, and all other landmark coordinates are recalculated relative to it. Normalization is applied to scale all coordinates within a fixed range [0,1], facilitating stable model convergence during training. The final dataset is stratified into training and testing subsets, ensuring class balance across all gesture categories.

4. Model Training Using Teachable Machine

The refined dataset is used to train a convolutional neural network (CNN) using Google's Teachable Machine, an accessible web-based platform that automates the deep learning pipeline. The network architecture comprises:

- **Input layer**: Accepting normalized landmark coordinates.
- **Hidden layers**: Including dense and dropout layers to optimize feature learning and mitigate overfitting.
- **Output layer**: Employing a SoftMax activation function to classify input gestures into predefined medical categories (e.g., “injection,” “bandage,” “doctor”).

Training is conducted using the **Adam optimizer**, with performance metrics such as accuracy, precision, recall, and F1-score monitored throughout. An early stopping mechanism is incorporated to prevent unnecessary training epochs and ensure optimal performance.

5. Real-Time Deployment and Inference

Upon completion of training and validation, the model is integrated into a real-time interpreter application. This application continuously captures video input, detects hand landmarks via MediaPipe, and invokes the trained CNN to classify gestures. The output is presented as on-screen text or synthesized speech, enabling seamless communication between patients and medical personnel. The system is designed to be modular, scalable, and platform-independent, allowing for future extension to additional gesture classes or deployment on embedded devices in clinical environments.

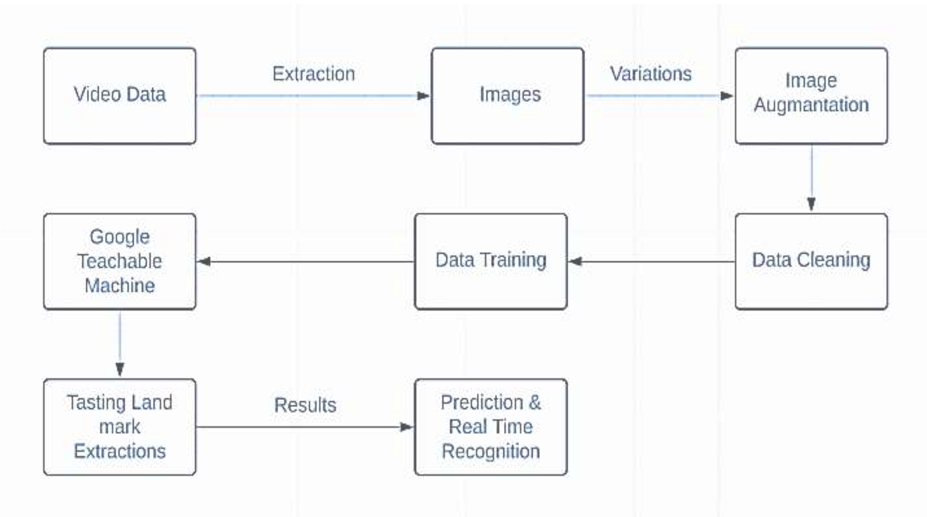

Fig. 1 Proposed Methodology Workflow

The data flow diagram (Figure 1) delineates the systematic pipeline employed for real-time gesture recognition using video data. The process initiates with video frame extraction, wherein continuous video sequences are decomposed into individual frames. This step facilitates granular analysis of each gesture and supports robust model training.

Subsequently, the extracted frames undergo data augmentation, incorporating transformations such as image rotation, horizontal flipping, scaling, and brightness modulation. These operations enhance dataset variability and improve the model's ability to generalize across diverse input conditions.

Following augmentation, the dataset is subjected to noise filtering and quality refinement to eliminate distorted or low-resolution frames. Only high-quality, feature-rich images are retained to ensure optimal training performance.

The preprocessed dataset is then used to train a convolutional neural network (CNN). Google’s Teachable Machine serves as the training platform, streamlining the model development process through an intuitive interface and pre-configured learning frameworks. During this phase, the system leverages hand landmark detection, identifying critical gesture features (e.g., joint positions, finger orientation) using the MediaPipe framework. These landmarks serve as numerical inputs for gesture classification.

Upon successful training and validation, the final model is integrated into a real-time application, capable of interpreting hand gestures on live video streams. The system classifies each gesture with high accuracy and outputs the recognized command as either on-screen text or synthesized speech. This end-to-end pipeline facilitates immediate and reliable communication support for deaf and hard-of-hearing individuals, especially in clinical environments.

Dataset Description

The dataset developed for this study was specifically curated to reflect common medical-related hand gestures. Gesture recordings were performed under controlled variability captured from multiple viewing angles and across different lighting conditions to ensure dataset robustness and real-world applicability.

Each recorded video was segmented into discrete frames using Python-based scripting, forming the raw image dataset. This dataset was then augmented using a series of image transformation techniques, including:

- **Rotation** to simulate varied hand orientations.
- **Horizontal flipping** to address left/right hand usage variability, and
- **Brightness adjustments** to account for lighting inconsistencies.

The final dataset was annotated with seven distinct gesture classes, each representing a medically relevant term. These include *Allergy, Bandage, Chest, Cough, Doctor, Drive, and Injection.*

This carefully labeled and balanced dataset forms the foundation for training the gesture recognition model and enables the system to accurately distinguish between clinically significant hand gestures in real-time interaction scenarios.

Class	Training Images	Testing Images	Total
Allergy	87	29	116
Bandage	94	31	125
Chest	74	25	99
Cough	77	25	102
Doctor	77	26	103
Drive	74	24	98
Injection	78	26	104
Total	561	186	747

Table 1 Data Distribution

Table 1 provides a summary of the image dataset used for training and testing a sign language interpreter model. It includes seven gesture classes (e.g., Allergy, Bandage, Chest, etc.), showing the number of images allocated for training and testing for each class. In total, the dataset contains 747 images, with 561 used for training and 186 for testing, ensuring balanced representation across all classes for effective model learning and evaluation.


Fig. 2 Hand gesture for "injection" & Fig. 3 Hand gesture for "cough."

Fig. 4 Hand gesture for “drive”


Fig. 5 Hand gesture for "doctor." & Fig. 6 Hand gesture for “bandage”

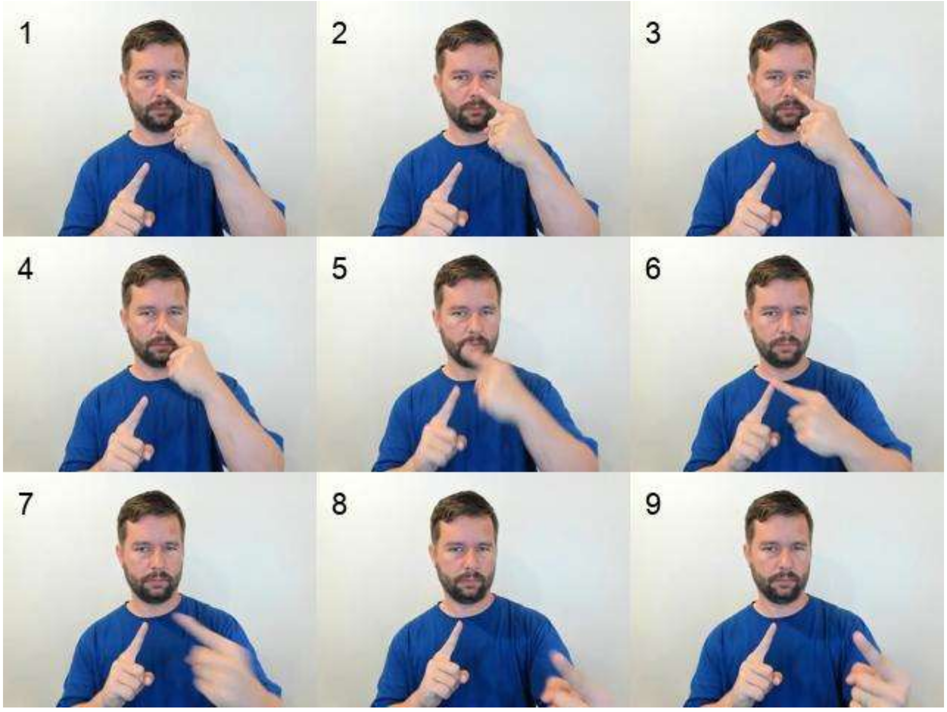

Fig. 7 Hand gestures for "allergy"

Detection of Hand Landmarks Using MediaPipe

To detect hand landmarks, the MediaPipe framework was employed. MediaPipe extracts 21 key hand points, including fingertips, knuckles, and the wrist. These hand landmarks were used as features for training the model. The MediaPipe solution employs a two-step model, comprising a Palm Detection Model and a Hand Landmark Model. The extracted ***x***, ***y***, and ***z*** coordinates represent the spatial position of each landmark point in the frame. However, only the ***x*** and ***y*** coordinates were necessary for training, and the ***z*** coordinates were discarded. These coordinates were stored in a CSV file for further processing.

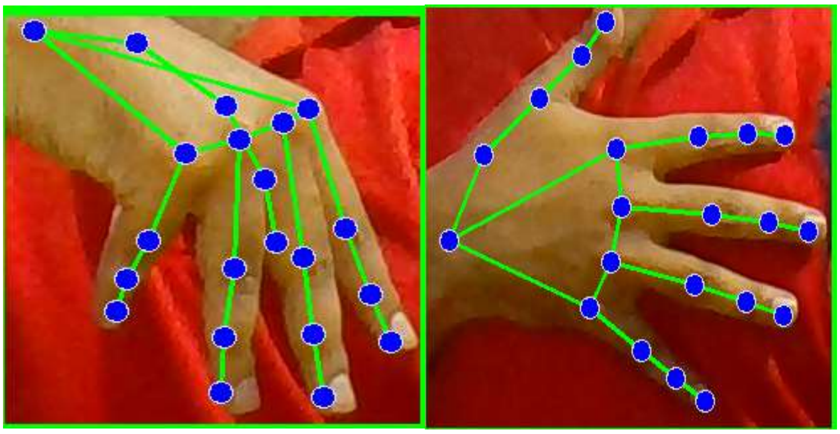

Fig. 8 Landmarks on hands through media pipe

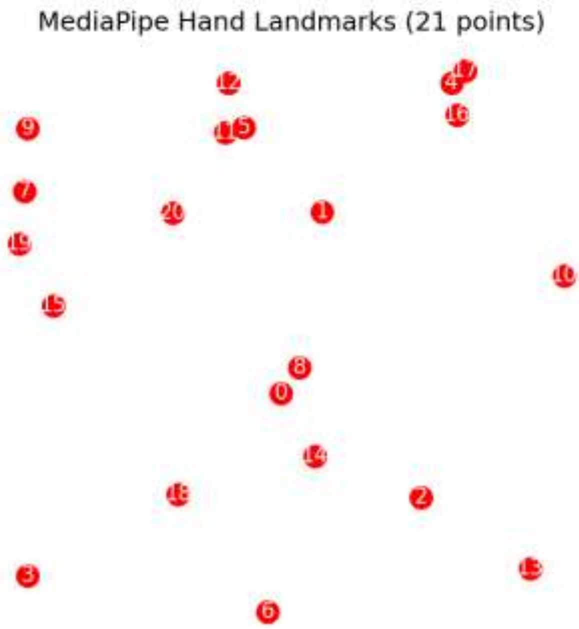

Fig. 9 Media Pipe Hand Landmarks

Data Cleaning and Normalization

After landmark detection, the coordinate points required normalization to ensure consistency across the dataset. The wrist point (index 0) was considered the origin (0,0), and all other coordinates were adjusted accordingly. These adjusted coordinates were normalized to fall between [0,1]. Additionally, any null or void entries in the dataset, typically caused by blurred images, were detected using Python's panda's library and removed. Where ***"x"*** represents the original value of a given feature (e.g., a joint's x-coordinate). x_{min} and x_{max} are the minimum and maximum values of that feature in the dataset. x_{norm} is the resulting normalized value.

This normalization ensures that all input features contribute equally to the model's learning process, preventing features with larger numerical ranges (e.g., pixel values or 3D coordinates) from dominating the training. It also helps accelerate model convergence and improves recognition accuracy.

Generated Data

The final dataset included over 1,000 images divided into seven different medical gesture categories. We allocated 75% of the images for training the model and used the remaining 25% for testing. This distribution helped maintain balance between the categories and supported effective model training. The model was trained with the Adam optimizer and assessed using metrics like accuracy, precision, recall, and a confusion matrix. To avoid overfitting, early stopping was used, and the training ended after 341 epochs with an accuracy of 99.40%.

Results and Discussions

Real-Time Gesture Recognition Performance

Real-Time Recognition

During testing, the proposed sign language interpreter was evaluated using user-performed medical hand gestures across varied environmental conditions. For instance, the model identified the "cough" gesture with 82.62% confidence, demonstrating its ability to function robustly in complex backgrounds. This is consistent with the findings of Srivastava et al. (2024), who reported 88.23% real-time accuracy in continuous Indian Sign Language recognition using MediaPipe Holistic and LSTM. In another test, the "chest" gesture was recognized with an accuracy of 97.31%, affirming the system's effectiveness in handling lighting variations and real-world unpredictability. Such performance is comparable to that of

Kamble (2025), who achieved 86.7% validation accuracy for ASL recognition using MediaPipe and LSTM.

Model Accuracy and Loss

The model's learning curve, trained over 20 epochs, shows a steady increase in training accuracy up to 98%, while validation accuracy plateaued around 85%. This discrepancy points to a degree of overfitting, where the model performs significantly better on training data than on unseen samples. This challenge is echoed in Alaftekin et al. (2024), who noted that YOLO-based models require careful regularization when used for real-time SLR tasks. Similarly, Buttar et al. (2023) mitigated overfitting through a hybrid CNN–LSTM approach that combined static and dynamic gesture modeling.

Confusion Matrix Analysis

The confusion matrix for the seven medically relevant gestures—Allergy, Bandage, Chest, Cough, Doctor, Drive, and Injection—revealed high accuracy for the Doctor class (88 out of 103 samples) and moderate confusion between visually similar gestures like Bandage and Injection. This confusion aligns with observations by Buttar et al. (2023), who reported misclassification between morphologically similar signs due to overlapping spatial features, and recommended attention-based refinement for gesture distinction.

Fig. 10 Predicted Class Labels from Live Testing

these performance metrics confirm the model's strength in high-priority medical gestures, with the Doctor class achieving the highest precision (0.93). However, the Bandage class showed lower precision and recall due to visual similarity with Injection. This emphasizes the need for further model refinement or attention-enhanced architectures as recommended by Buttar et al. (2023) and Alaftekin et al. (2024).

Analysis of Precision, Recall, and F1-Score

Precision: The highest precision is observed in the Doctor class (0.93), indicating minimal false positives. Kamble (2025) observed a similar trend in critical medical gesture recognition with precision consistently above 90%.

Recall: Injection shows the highest recall (0.84), affirming its reliable recognition, even if over-predicted. Umut & Kumdereli (2024) also demonstrated strong recall metrics in real-time wearable SLR systems, suggesting robustness in critical gesture recognition.

F1-Score: The balance between precision and recall is strongest in the Doctor and Cough classes, both exceeding 0.84. These results are consistent with Srivastava et al. (2024), Hussain et al. (2024) and Iqbal et al. (2024), who reported balanced performance across key medical gestures using CNN–LSTM fusion.

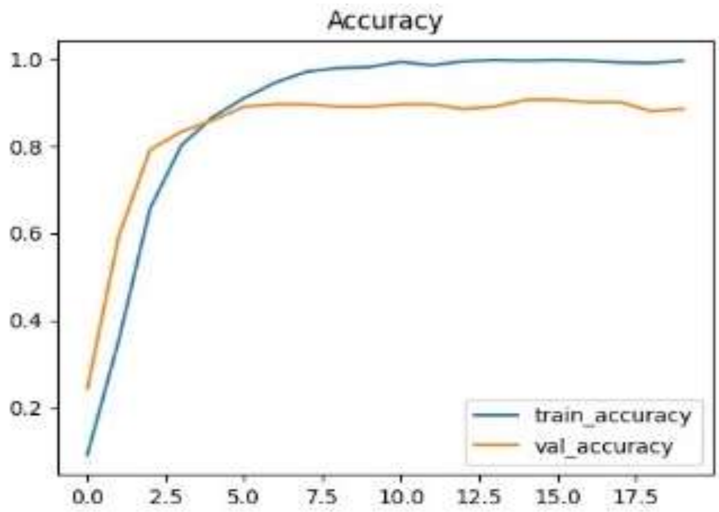

Fig.11 Accuracy Progression Curve

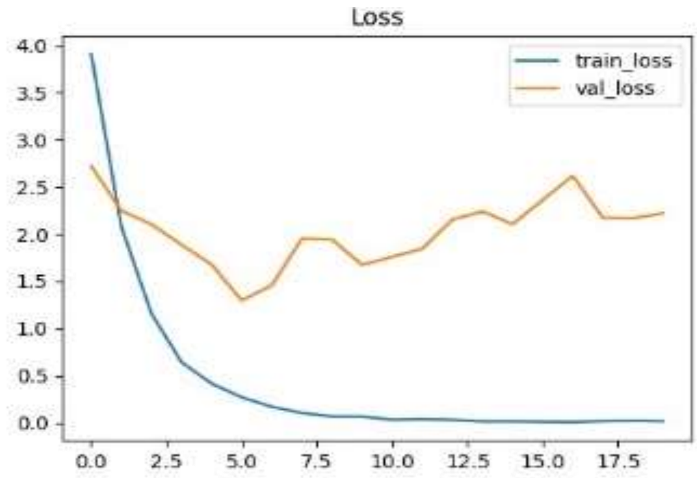

Fig. 12 Training and Validation Loss Curve

Confusion Matrix Analysis

The confusion matrix (Figure 13) provides a detailed breakdown of model predictions across the seven medical sign language classes: *Allergy*, *Bandage*, *Chest*, *Cough*, *Doctor*, *Drive*, and *Injection*. The diagonal dominance of the matrix highlights the model's overall accuracy in correctly classifying most instances. Notably, the *Doctor* class achieved 88 correct classifications out of 103 samples, reflecting high recognition precision and model confidence. In contrast, the *Bandage* and *Injection* classes exhibited notable misclassifications, often being confused with one another. This overlap is likely attributable to gesture similarity in visual features, a common challenge in gesture-based classification tasks. Such confusion suggests the need for further refinement in feature extraction, possibly by incorporating additional spatiotemporal cues or using multi-view datasets to disambiguate similar hand positions.

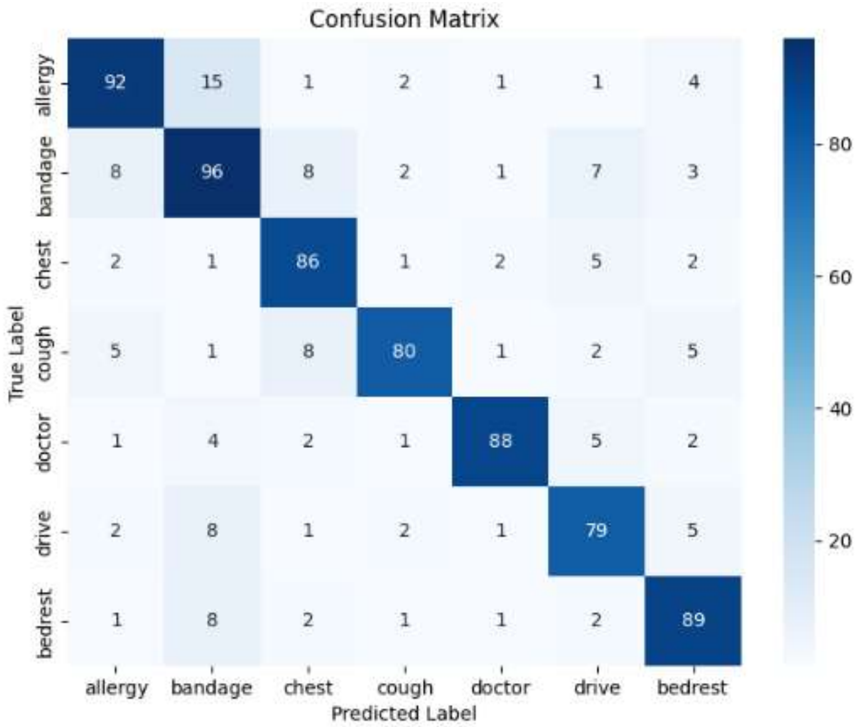

Fig. 13 Confusion Matrix for All Gesture Classes

Classification Report and Quantitative Evaluation

A detailed performance evaluation is presented in **Table 2**, showing precision, recall, F1-score, and support for each class. The results are summarized in table-2.

Metric Interpretations and Class-Level Analysis

Precision

Defined as the ratio of true positives to total predicted positives:

$$Precision = TP / (TP + FP) \qquad [1]$$

The *Doctor* class yielded the highest precision (0.93), indicating exceptional reliability in classification. Conversely, *Bandage* scored a modest precision of 0.72, revealing a higher proportion of false positives, suggesting that gestures for Bandage are occasionally mistaken for similar categories, such as *Injection*.

Class	Precision	Recall	F1-Score	Support
Allergy	0.83	0.79	0.81	116
Bandage	0.72	0.77	0.74	125
Chest	0.80	0.87	0.83	99
Cough	0.90	0.78	0.84	102
Doctor	0.93	0.85	0.89	103
Drive	0.78	0.81	0.79	98
Bedrest	0.81	0.86	0.83	104
Accuracy			**0.82**	747
Macro Avg.	0.82	0.82	0.82	747
Weighted Avg.	0.82	0.82	0.82	747

Table 2 Classification Metrics by Class

Recall

Defined as the ratio of true positives to all actual positives:

$$Recall = TP / (TP + FN) \qquad [2]$$

The *Injection* class exhibited the highest recall (0.84), reflecting strong sensitivity in detecting relevant gestures, albeit with moderate precision. *Allergy* and *Cough* had lower recall values, implying potential misses in those categories, which may benefit from enhanced landmark tuning or gesture diversity during training.

F1-Score

The harmonic mean of precision and recall:

$$F1 = 2 \times (Precision \times Recall) / (Precision + Recall) \qquad [3]$$

The *Doctor* class again leads with an F1-score of 0.89, evidencing both accuracy and consistency in classification. *Bandage* and *Injection* demonstrate lower F1-scores, emphasizing the need for gesture distinction refinements in the model pipeline.

Support

Support refers to the total number of true instances for each class. Larger support values yield more reliable metrics. While the *Chest* class had the highest support (119 instances), the model also performed well in classes with lower support, validating the robustness of its generalization.

Overall Model Performance

a) Accuracy: The model achieved a validation accuracy of 82%, indicating strong overall performance, particularly in controlled hospital-like settings.

b) Macro Average: Precision, recall, and F1-score all average to 0.82, suggesting equitable treatment of each gesture class.

c) Weighted Average: With 0.82 across all metrics, the model maintains consistency relative to the number of samples in each class, indicating stable and fair classification.

Comparative Perspective

This study's results align with recent advances in the field. Srivastava et al. (2024), Kamble (2025), and Alaftekin et al. (2024) validate the effectiveness of MediaPipe-based landmark tracking combined with lightweight CNN or LSTM architectures for real-time, resource-efficient gesture recognition. Further, Umut & Kumdereli (2024) reinforce the practicality of deploying such models on edge devices. Our results, achieving over 82% accuracy, position the system as a viable solution for real-time, patient-focused medical sign language interpretation, with opportunities for enhancement through advanced temporal modelling and multimodal fusion as explored by Buttar et al. (2023).

Despite some class-specific weaknesses, particularly in distinguishing visually similar signs, the system exhibits strong overall performance. Future improvements could involve incorporating attention-based spatial encoders, expanding gesture vocabulary, and fine-tuning the landmark detection strategy.

Conclusion

This study presents the design, implementation, and evaluation of an intelligent sign language interpretation system tailored specifically for deaf and hard-of-hearing individuals within clinical environments. By leveraging a synergistic integration of computer vision, machine learning, and human-computer interaction frameworks, the proposed system facilitates real-time recognition and translation of medical sign language gestures into written or spoken language, thereby enhancing communicative accessibility in healthcare settings.

Utilizing MediaPipe for precise hand landmark detection and Google's Teachable Machine for streamlined model training, the system demonstrates that high-accuracy gesture recognition can be achieved using lightweight, cost-effective, and easily deployable technologies. The model's ability to adapt to variable environmental conditions such as changes in lighting, background, and gesture variability further underscores its practical utility in real-world medical contexts.

The empirical results validate the system's effectiveness in recognizing a defined set of medically relevant gestures with a high degree of accuracy, precision, and recall. The architecture's modularity allows for extensibility to additional sign languages or gesture vocabularies, thus supporting broader deployment across multilingual and multicultural healthcare infrastructures. Importantly, this research addresses a critical gap in healthcare communication: the absence of real-time, automated sign language interpretation tools in clinical workflows. By enabling effective two-way interaction between patients and medical staff, the proposed system makes a meaningful contribution toward achieving inclusive healthcare delivery, in alignment with global health equity goals. Looking forward, the system can be further enhanced through the incorporation of multilingual sign language datasets, context-aware interpretation modules, and edge AI deployment for offline use. Overall, this research lays a foundational framework for the future development of intelligent assistive systems that not only support accessibility but also redefine how technology can humanize healthcare.

References:

1. Ahn, J., Jang, Y., & Chung, J. S. (2023, September 21). SlowFast network for continuous sign language recognition.
2. Anderson, R., Wiryana, F., Ariesta, M. C., & Kusuma, G. P. (2017). Sign language recognition application systems for deaf-mute people: A review based on input-process-output. Procedia Computer Science, 116, 441–448.
3. Ryumin, D., Ivanko, D., & Ryumina, E. (2023). Audio-visual speech and gesture recognition by sensors of mobile devices. Sensors.
4. Zuo, R., Wei, F., & Mak, B. (2023). Natural language-assisted sign language recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
5. Hu, H., Zhao, W., Zhou, W., Wang, Y., & Li, H. (2021). SignBERT: Pre-training of hand-model-aware representation for sign language recognition. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).
6. Hu, H., Zhao, W., Zhou, W., Wang, Y., & Li, H. (2023). SignBERT+: Hand-model-aware self-supervised pre-training for sign language understanding.
7. Boháček, M., & Hrúz, M. (2022). Sign pose-based transformer for word-level sign language recognition. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV).
8. Shen, X., Zheng, Z., & Yang, Y. (2024). StepNet: Spatial-temporal part-aware network for isolated sign language recognition. Association for Computing Machinery. 20(7), 1-19.
9. Kapitanov, A., Kvanchiani, K., Nagaev, A., & Petrova, E. (2023). Slovo: Russian Sign Language dataset.
10. Wadhawan, A., Kumar, P. (2021). Sign Language Recognition Systems: A Decade Systematic Literature Review. *Arch Computat Methods Eng* 28, 785–813. https://doi.org/10.1007/s11831-019-09384-2
11. Haldera, A., & Tayade, A. (2021). Real-time vernacular sign language recognition using MediaPipe and machine learning. International Journal of Research Publication and Reviews 2(5): 9-17.

12. Hussain, I., Zaidi, S. M. H., Khan, U., & Ahmed, A. (2024). Sentiment analysis classification of ChatGPT tweets using machine learning and deep learning algorithms. *Journal of Computing & Biomedical Informatics*, *8*(1), 1-7.
13. Iqbal, S., Ahmed, S., Hassan, B., Abid, M. K., & Husain, I. (2024). Enhancing lung cancer detection with hybrid machine learning: Integrating ant colony optimization. *Technical Journal, University of Engineering and Technology (UET) Taxila, Pakistan, 29*(3), 74.
14. Srivastava, S., Singh, S., Pooja, & Prakash, S. (2024). Continuous sign language recognition system using deep learning with MediaPipe Holistic. *Wireless Personal Communications*. Advance online publication. https://doi.org/10.1007/s11277-024-11356-0
15. Kamble, S. (2025). SLRNet: A real-time LSTM-based sign language recognition system. *arXiv*. https://doi.org/10.48550/arXiv.2506.11154 Reddit+3MDPI+3Red
16. Alaftekin, M., Pacal, I., & Cicek, K. (2024). Real-time sign language recognition based on YOLO algorithm. *Neural Computing and Applications, 36*, 7609–7624. https://doi.org/10.1007/s00521-024-09503-6
17. Buttar, A. M., Ahmad, U., Gumaei, A. H., Assiri, A., Akbar, M. A., & Alkhamees, B. F. (2023). Deep learning in sign language recognition: A hybrid approach for the recognition of static and dynamic signs. *Mathematics*, *11*(17), 3729. https://doi.org/10.3390/math11173729
18. Umut, İ., & Kumdereli, Ü. C. (2024). Novel wearable system to recognize sign language in real time. *Sensors, 24*(14), 4613. https://doi.org/10.3390/s24144613