International Journal of Applied Mathematics
Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

KARPALLAVI (HAND GESTURES) TO TEXT AND SPEECH CONVERSION: MARATHI SIGN LANGUAGE RECOGNITION

First Auther[1] Rohini Vijay Daund and Second Auther[2] Dr. Monali Y. Khachane

[1]Assistant Professor

Dept. of Computer Science,

KRT Arts, BH Commerce and AM Science (K.T.H.M) College, Nashik.

daund.rohini.32@gmail.com

[2]Assistant Professor

Dept. of Computer Science,

Dr. Annasaheb G.D. Bendale Mahila Mahavidhyalaya, Jalgaon, India

monalikhachane@gmail.com

Abstract

In present research work 'Karapallvi' (Marathi Sign Language) converted into Marathi text. Karpallavi (MSL) is a hand gestures Language used for communication with individuals and hearing Speech impairments, allowing them to express their thoughts and ideas effectively. In Nashik district each linguistic community has its self-variant of hand gesture language. Instead of its significances, 'Karapallvi'(MSL) requires rigorous research due to lacks of technological and digital support as compared to other sign languages. The 'Karapallvi' (MSL) consist 'Mulakshare'(मुळाक्षरे), which lays down the foundation of Marathi script. Effective application of Deep Learning techniques gives automatic identification of these hand gestures. The research methodology contains Dataset Preparation, Preprocessing, Feature extraction, Model train and built techniques. It is found that the present research work not only increases the accessibility but also helps to preserve 'Karapallvi' (MSL) in the era of digital disruption. The ANN model gives (approx.91%) highest accuracy as compared to KNN, CNN, SVM.

Keywords: Karpallavi, Media Pipe, ANN, TensorFlow, Marathi Sign Language.

1.Introduction:

Marathi Sign Language (MSL) is an essential medium of communication for a large section of the deaf and hard-of-hearing population in Maharashtra. For many individuals, sign language is not merely a supplementary tool but their primary and most natural form of expression. However, despite the widespread use of Marathi as a spoken language, technological solutions that support Marathi Sign Language remain scarce. Most existing sign language recognition systems focus on globally dominant languages, which creates a significant accessibility gap for users of regional sign languages. Recognizing and addressing this gap is crucial for ensuring equal access to communication, information, and

Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
opportunities. In India, a nation having linguistic diversity. So, in India primarily identifies 12 spoken languages, each language having its own distinct sign language. In Maharashtra, where Marathi is the native tongue, there are over 900,000 deaf and mute individuals, the vast majority of whom rely exclusively on Marathi Sign Language (MSL) to interact with the world. So that it emphasizes on development of real time Marathi sign language system. In this proposed work the researcher has developed a Marathi sign language hand gestures (Karpallavi) recognition which is commonly used in communication. The Proposed system will translate Marathi sign language hand gestures in to the Marathi text. This text convert into audio format. This system employed the Media pipe framework using python. The Marathi sign language recognition system will help accessibility and domestic inclusivity, improving the system's use for native Marathi speakers. In the fields of health services, education and public services, it will highly support to the digital inclusion of hearing disabled people. The effectiveness of Marathi sign language recognition will also enhance and it will support to the formation of assistive technologies in the context of Marathi Sign language recognition.

1. Objectives

1. To create the dataset for Marathi sign language.
2. To develop hand tracking module through Media Pipe.
3. To translate Marathi sign language hand gestures into Marathi text.
4. To convert Marathi text into audio.

2. Related work

Review of existing literature was conducted and summarized in the Table No 1. Dahibavkar Swaraj et. al. was developed a model for deaf people to interact with another by translating Marathi Sign Language into the alphabet. [1]. Shinde Amitkumar, Ramesh Kagalkar et. al. works on static gesture detection only for alphabets of Marathi Sign Language (MSL) was developed [2,3]. Muhammad and Vinod Kumar only Works on 10 Persian alphabet using CNN technique [4]. Mohana, R Madana focuses on sign language recognition and retrieval using machine learning and data mining with PCA and SVM [5]. M. Sharma et. al.used KNN classifier for classification [6]. Chandwani Laveen et. al. worked on static data recognition only for Alphabets, Number and words for American Sign Language (ASL) [7-8]. J. Bora, S. Dehingia, A. et. al. worked with MediaPipe and KNN classifier for Assamese Language (ASL) recognition [9]. R Kumar, A Bajpai et. al. only works on alpha - numeric data [10-13]. Moustafa, A. M. et. al. works on Arabic Sign (ArSL) Language recognition for Arabic 28 Alphabets applying Media pipe and CNN [14]. K. Gomase, A. Dhanawade, P. et. al. Works on only A-Z Alphabets with ASL [15]. There's a recognized gap in research for systems that can interpret live feed Marathi Sign Language (karpallavi) recognition via reactive computing.

Table 1. Related Work

No.	Purpose	Sign Language	Accuracy	Techniques	Limitations

Received: August 12, 2025

Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

				Used	
1	Conversion of Marathi Sign Language into the alphabet	Marathi Sign Language (MSL) Dataset size:8600	95%	Canny Edge-Based Segmentation Technique, CNN	colored cloths, and non-skin colored background are required.
2	Static Marathi language sign recognition using image processing.	Marathi Sign Language (MSL) Dataset size:1000	--	Computer vision, edge detection	Only works static data
3	System works on Marathi Sign Language recognition offline and Web camera mode.	Marathi Sign Language (MSL) Dataset size:1000	90%	Pattern matching process, based on skin filtering	Static gesture recognition for Alphabets.
4	Gesture translation from one language to another.	Persian Sign languages (ISL) Dataset size:200	98%	Background Elimination, CNN	Only work on 10 Persian alphabet
5	The current research focuses on sign language recognition and retrieval using machine learning and data mining	American Language (ASL) Dataset size:900	62.37%	PCA, SVM	background, which is typically expected to be basic, non-reflective, and starkly different from skin tone
6	This system uses a hierarchical centroid feature extractor, a direct pixel value classifier, and a neural	Regional Language (RSL) Dataset size:5000	97.10%	Neural networks and KNN classifier.	It works only for regional sign language recognition system.

Received: August 12, 2025

Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

	network				
7	A simplified methodology applying Media Pipe library and ML algorithms	American Language (ASL) Dataset size:4000	99%	Media Pipe, SVM	System works on static gesture recognition only for Alphabets, Number and some words
8	Sign Language using Media Pipe and recognition through Computer vision alphabets.	American Language (ASL) Dataset size:1800	98%	Media Pipe, KNN	Only Works on 26 alphabets.
9	Detection of Assamese Sign Language.	Assamese Language (ASL) Dataset size:2094	99%	feed-forward neural network	Limited only for Assamese Sign Language.
10	Detection of American Sign Language applying dynamic approach	American Language (ASL) Dataset size:4500	99.95%	CNN, CV Zone, Media Pipe	Only Works on American Sign Language.
11	Multiple Language Recognition system for alphabets and numbers	American Language (ASL) Dataset size:15600 Indian Language (ISL) Dataset size:4972 Italian Sign Language (Ita.SL) Dataset size:12856 Turky Sign	99%	Media Pipe Naive Bayes, SVM, KNN	Only Works on alphabets and numbers

Received: August 12, 2025

Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

		Language (TSL) Dataset size:4124			
12	Applying dynamic approach for complex hand gesture detection	American Language (ASL) Dataset size:15600 Indian Language (ISL) Dataset size:4972 Italian Language (Ita.SL) Dataset size:12856	99%	Media Pipe's	Works on Alphabets and Numbers
13	Works on numbers only	American Sign Language (ASL)	95.7%	Media Pipe, KNN	Works on Numbers and only some hand Sign
14	Proposed system works on Arabic Sign (ArSL) Language recognition for Arabic 28 Alphabets	Arabic Sign (ArSL)Language Dataset size:7057 +5600	97.1%	Media Pipe, CNN	Works on only Arabic Alphabets
15	Make a recognizer for A-Z alphabet using MediaPipe	American Sign Language (ASL)	(86-91) %	Media Pipe, KNN	Works on only A-Z Alphabets

Gaps or Limitations of Existing System:

1. Non-Skin colored cloths, Dark background, hand gloves are required
2. Static gesture recognition for Alphabets and Numbers.
3. Not working on regional languages.

Applications of Proposed System:

1. Effective communication leads to the better Educational and Social development.
2. Real time communication for hard of hearing impairment community.

Received: August 12, 2025

International Journal of Applied Mathematics
Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3. Methodology:

Following flowgraphs gives outline for both training and testing Marathi sign language recognition system with MediaPipe, ML and DL techniques.

A: Model Training flowgraph:

1.Data Collection and Preparation:

a. Data collection:

The researcher has collected the hand gesture(Karpallavi) image data of Marathi alphabet(मराठी मुळाक्षरे) i.e. अ, आ, इ, ई, ए, ऐ, ओ, औ, क, ख, ग, घ, च, छ, ज, झ, ट, ठ, ड, ढ, ण, त, थ, द, ध, न, प, फ, ब, भ, म, य, र, ल, व, श, स, ह, ळ, क्ष, & ज्ञ (मराठी मुळाक्षरे) . Images are collected from the students and teachers of पडसाद कर्णबधीर विद्यालय, सिडको, नाशिक and श्रीमती माई लेले श्रवण व विकास विद्यालय, गंगापूर रोड, नाशिक, and also collected the data from individuals. In Marathi hand gestures (Karpallavi) 41 alphabet, each hand gestures having 50 images respectively. In this way the researcher has collected the 2050 (41*50 = 2050) images for preparing dataset.

The researcher cleaned and segregated the raw data (Hand gestures images) into clear and precise data for high accuracy. The researcher has collected the 2050 hand gesture images of 41 Marathi sign hand gesture alphabets into the raw format, which is needed to clean and segregate the data as per the order of the Marathi alphabets. This cleaned and segregated dataset of 2050 hand gesture images was used for feature extraction and landmark detection using Media Pipe and TensorFlow with Machine learning and Deep Learning algorithm.

International Journal of Applied Mathematics
Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
The researcher had referred a book i.e. परिसर अभ्यास, (भाग – १), for collection of authentic sign language. The signs of 41 Marathi alphabets referred by महाराष्ट्र राज्य पाठ्यपुस्तक निर्मिती व अभ्यासक्रम संशोधन मंडळ, पुणे. This book comprises the signs of Marathi language mostly used by deaf and dumb community.

b. Data Cleaning and Preprocessing:

Data Cleaning and Segregation of raw data (Hand gestures images) into clear and precise data for high accuracy. The collected the 2050 hand gesture images of 41 Marathi sign hand gesture alphabets into the raw format, which is needed to clean and segregate the data as per the order of the Marathi alphabets. This cleaned and segregated dataset of 2050 hand gesture images was used for feature extraction and landmark detection using Media Pipe with ML and DL technique. Cleaned images are flipping for removing mirror view of image. Then images are converted from BGR to RGB for compatibility with MediaPipe

The proposed framework is shown in Fig. 1.

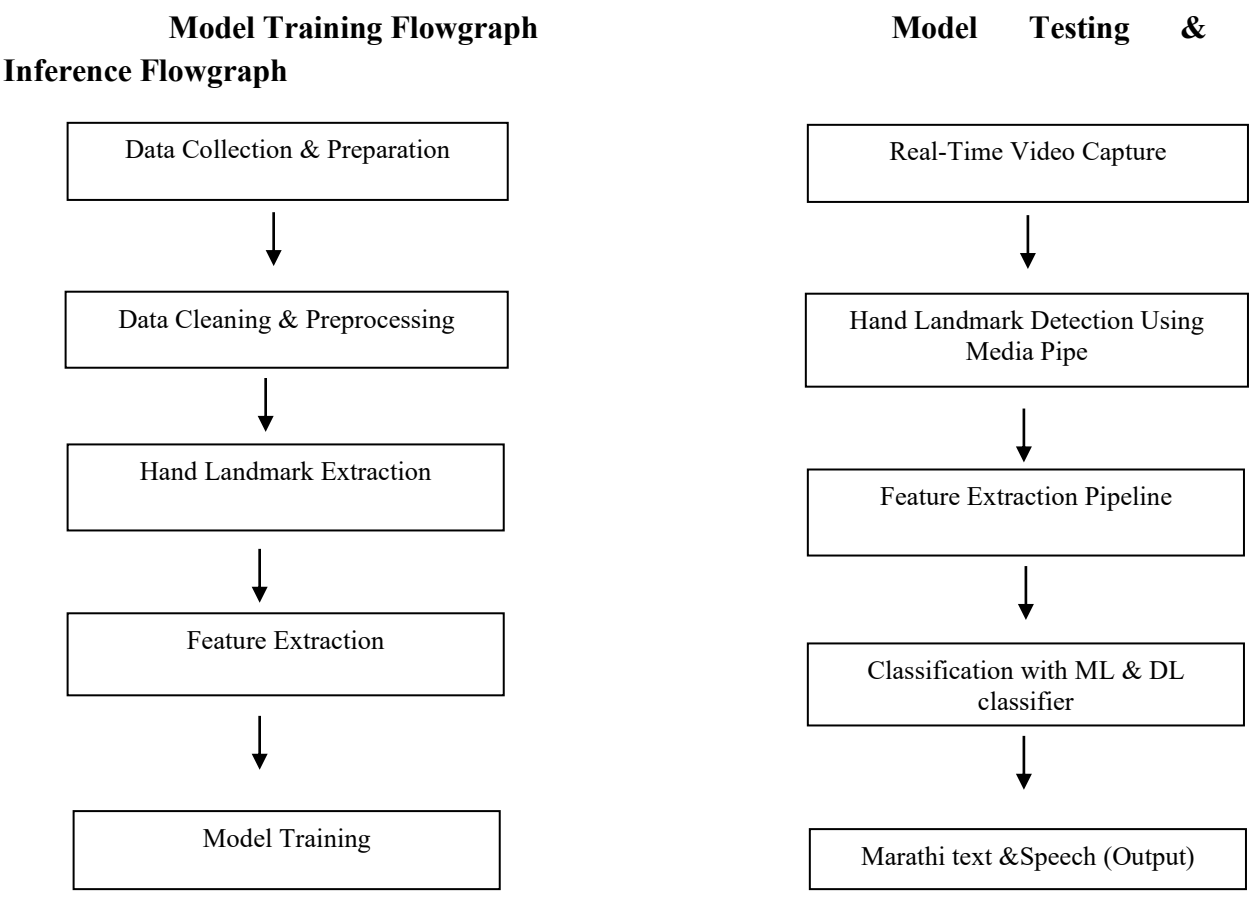


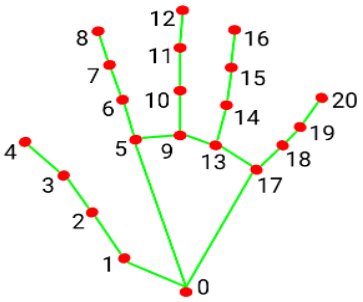
Fig.1: Research method for ISL into Marathi text and speech

2.Hand- Landmark Extraction:

The collected images or videos are processed to detect 21 key land mark for each hand. Instead of processing every pixel, the media pipelines only extract the numerical coordinates of 21 hand landmark key points per hand capture precise finger and palm

Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
positions.21 3-D landmarks include wrist is origin point (Landmark 0). Thumb having 4 points (CMC, MCP, IP, and Tip). 4 points for each of the four remaining fingers (MCP, PIP, DIP, and Tip) respectively. The 3D (x, y, z) coordinates for all detected landmarks are concatenated into a single, flattened 1D array.

Mediapipe Hand Tracking:



0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP
17. PINKY_MCP
18. PINKY_PIP
19. PINKY_DIP
20. PINKY_TIP

Fig.2: Media Pipe Hand land Mark

3.Features extraction:

After extracting 21 landmarks each having x, y, z coordinates. In process of feature extraction use the mathematical combination of 21 hand key points with calculating the distance between each two points.so there are 210 combinations forms for 21 landmarks. This distance calculates by Euclidean distance method. The Euclidean distance represents the straight-line distance between two points in three-dimensional space and is widely used to measure spatial similarity. This calculation measures how far two landmarks are from each other by considering differences along all three spatial axes.

By applying this computation across multiple landmark pairs, the system captures important characteristics such as finger spacing, joint articulation, and overall hand shape, which are essential for distinguishing between different sign language gestures.

Formula for Mathematical Combination:

$$nCr = \frac{n!}{r!(n-r)!} \quad \text{.....} \quad (1)$$

n: The total number of items in the set.

r: The number of items to be chosen from the set.

Formula for Euclidean distance method:

$$WRIST\_THUMB\_CMC = \sqrt{((X_{(2)} - X_{(1)})^2 + (Y_{(2)} - Y_{(1)})^2 + (Z_{(2)} - Z_{(1)})^2)} \quad \text{.....} \quad (2)$$

4.Model training:

Received: August 12, 2025

Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
The classification of karapallavi (MSL gestures) model was built with different ML (Decision Tree, SGD, KNN, SVM) and DL (CNN, ANN) architecture in TensorFlow. It takes the input as extracted feature from landmarks. All this structured data is grouped into sequences representing full hand gestures. To train the model divide the dataset, commonly using 80% for training and 20% for testing to prevent overfitting. Split training dataset is used to teach the model the patterns in the data and testing dataset is used to evaluate how well the trained model recon evaluate how well the trained model generalizes to unseen data. The model is compiled using an optimizer and a loss function.

B. Model Testing & Inference Flowgraph:

This phase focuses on evaluating the model's accuracy and running it in real-time.

1. **Real-Time Video Capture:**

Real-time video capture is a critical component of the proposed sign language recognition system, as it serves as the primary input source for gesture analysis. Live video data is captured using a standard webcam, enabling continuous frame acquisition for real-time gesture recognition. The video capture functionality is implemented using OpenCV.

2. **Hand Landmark Detection using MediaPipe:**

Hand landmark detection is a crucial stage in the proposed sign language recognition system. This model operates in two stages: first, a palm detection model identifies the presence and approximate location of a hand in the frame; second, a landmark model precisely estimates 21 three-dimensional hand landmarks. These landmarks correspond to key points on the hand, including finger joints, fingertips, and the wrist. Each landmark is represented by normalized (x, y, z) coordinates, capturing both spatial position and relative depth information.

3. **Feature Extraction Pipeline:**

Feature extraction transforms raw 21 hand landmark data into a structured and meaningful representation suitable for machine learning. Instead of directly using absolute landmark coordinates, the system focuses on capturing geometric relationships between different parts of the hand. Extracting meaningful features from these landmarks, the system computes distances between pairs of landmark points using the Euclidean distance method with all possible combination.

4. **Classification with ML and DL classifier:**

The pre-train model takes extracted feature vector as input with Decision Tree, SGD, KNN, SVM, CNN and ANN classifier, which has been recognize dynamic Marathi Sign Language gestures. All models predict the symbol into one of the 41 classes using 210 features extracted from the live feed. Comparing all this ANN achieves an accuracy of approximately 91%. Based on the learned patterns, the ANN predicts the corresponding gesture for each processed frame, enabling accurate and efficient real-time karpallavi (MSL gesture) recognition.

5.Marathi text and speech (Output):

Received: August 12, 2025

International Journal of Applied Mathematics
Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
The output of the classification model is recognized hand gesture and printed on the screen with hand landmarks using OpenCV. Then, Marathi text is displayed on the screen. recognized karpallavi (MSL gesture) converted into the audio format by using the gtts library using python.

C. Results and Discussion:

The performance evaluation of all models measure on the basis of matrices like confusion matrix. Many machine learning and Deep learning models are used for hand gesture detection. The models are evaluates based-on Test accuracy, F1 Score, precision and recall. In result table ANN shows 91.10% accuracy, 91.69% for F1 Score, 91.11% for precision and 91.03% for recall. The ANN Model shows the higher accuracy in every parameter than Decision tree model, KNN model ,CNN Model ,ANN Model and SGD Model.

Predicted Output:

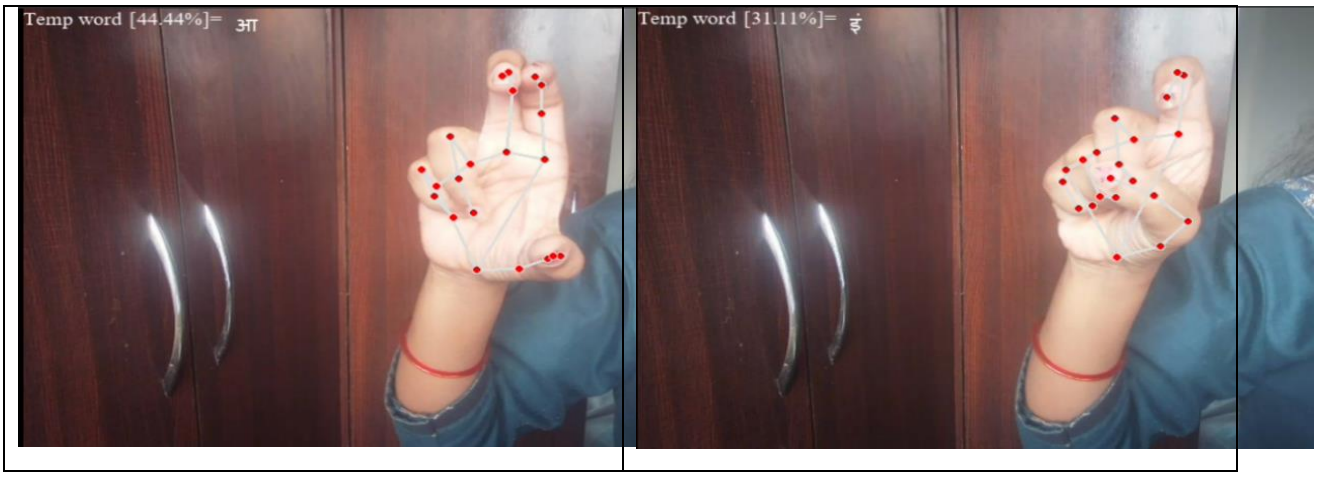


Figure 3: Images of Marathi text recognition Hand gestures

Table No 2 Result and Accuracy

Sr. No	Model Name	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	Decision Tree	60.00	62.71	60.00	60.15
2	SGD	79.72	82.14	79.72	77.71
3	KNN	80.33	81.83	80.33	89.57
4	CNN	86.89	88.71	86.89	86.74
5	ANN	91.10	91.69	91.11	91.03
6	SVM	85.95	87.16	85.95	86.05

The confusion matrix by using ANN algorithm for MSL gestures.

Received: August 12, 2025

International Journal of Applied Mathematics
Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

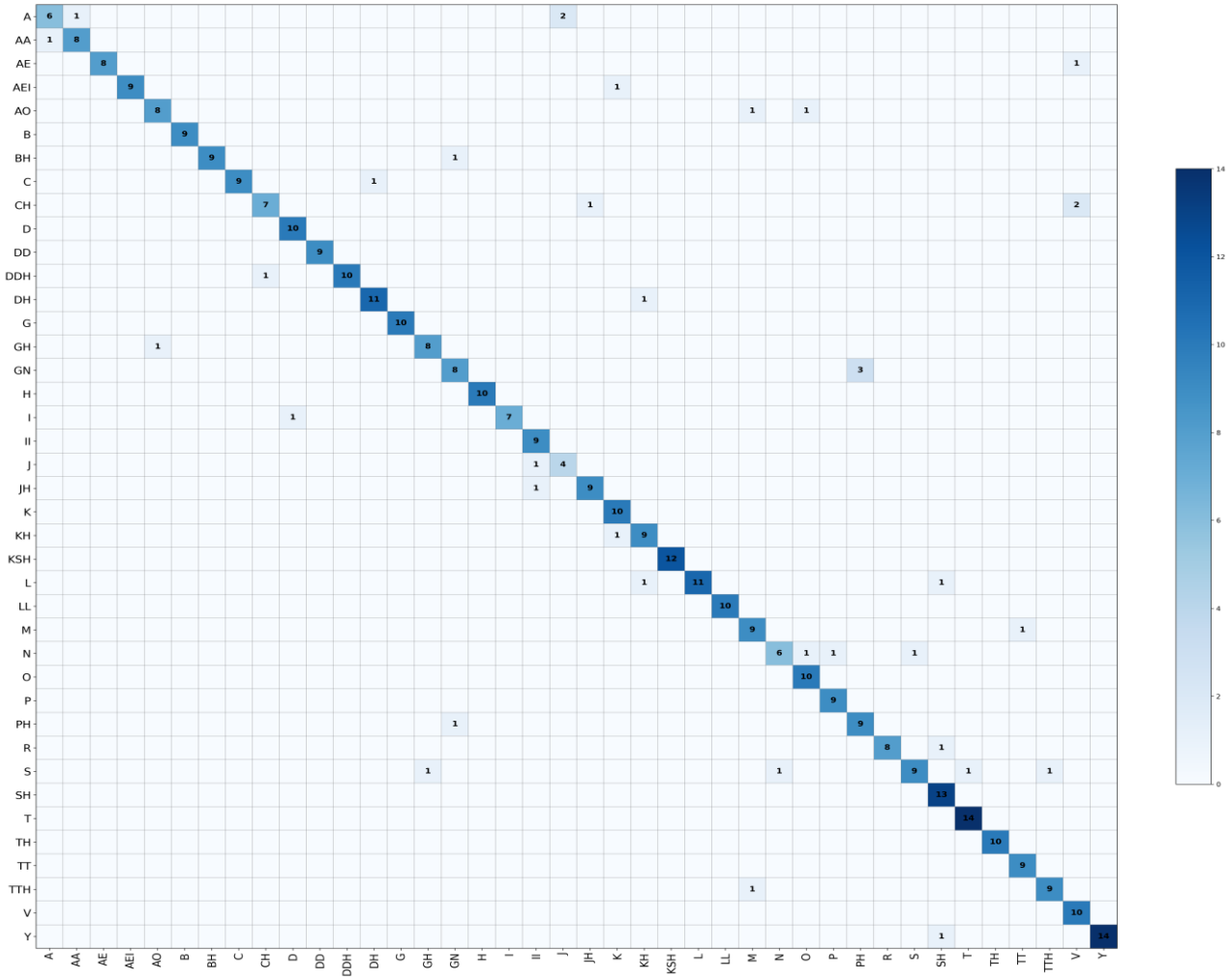


Fig.4: Confusion Matrix by using ANN Algorithm

D. Conclusion:

An automated MSL recognition system can act as a bridge between hearing-impaired users and the general population, enabling smoother interactions in everyday situations such as classrooms, offices, government services, and public spaces. This directly contributes to reducing social isolation and promoting independence. From an educational perspective, Marathi Sign Language recognition holds significant importance. The live videos are captured into the system using the OpenCV library, which extracts video frames into separate images. The researcher has successfully achieved its objectives by extracting key features from the hand gesture data of Marathi sign language for 41 alphabets (मुळाक्षरे). Classifying and recognizing the gestures using machine learning techniques. In proposed system it can be concluded that by using media Pipe and TensorFlow framework, live Marathi sign gestures are detected. ANN model shows the higher accuracy (approx. 91%) as compare to another model.

Received: August 12, 2025

Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Future Scope:

1. In future achieving higher accuracy.
2. Development of Mobile application.

References:

[1] S. Dahibavkar, J. Dhopte, M. Patole, and S. Madachane, "Marathi sign language recognition," 2020.

[2]A. Shinde, "Sign language to text and vice versa recognition using computer vision in Marathi," International Journal of Computer Applications, vol. 118, no. 13, pp. 1–7, 2014.

[3]A. Shinde and R. Kagalkar, "Advanced Marathi sign language recognition using computer vision," International Journal of Computer Applications, vol. 118, no. 13, pp. 1–7, 2015.

[4]Jafari, Muhammad Reza. n.d. "Persian Sign Gesture Translation to English Spoken Language on Smartphone." International Journal of Engineering Research 9 (06).

[5]Mohana, R Madana "Machine Learning and Data Mining Techniques for Sign Language Recognition and Retrieval System." International Journal of Engineering Computational Research and Technology 2019.

[6]M. Sharma, R. Pal, and A. K. Sahoo, "Indian sign language recognition using neural networks and KNN classifiers," International Journal of Computer Applications, vol. 9, no. 8, 2014.

[7]L. Chandwani, J. Khilari, K. Gurjar, P. Maragale, A. Sonare, S. Kakade, A. Bhatt, and R. Kulkarni, "Gesture based sign language recognition system using Mediapipe," Preprint, 2023, doi: 10.21203/rs.3.rs-3106646/v1.

[8] K. Goyal and V. G. Velmathi, "Indian sign language recognition using Mediapipe Holistic," International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 2020.

[9] J. Bora, S. Dehingia, A. Boruah, A. A. Chetia, and D. Gogoi, "Real-time Assamese sign language recognition using MediaPipe and deep learning," Procedia Computer Science, vol. 218, pp. 1384–1393, 2023.

[10] R. Kumar and A. Bajpai, "Mediapipe and CNNs for real-time ASL gesture recognition," in Proc. Int. Conf. Smart City and Emerging Technology, 2021, doi: 10.13140/RG.2.2.32364.03203.

[11] A. Haldera and A. Tayade, "Real-time vernacular sign language recognition using MediaPipe and machine learning," International Journal of Research Publication and Reviews, 2021.

[12] K. M. Kavana, "Recognization of hand gestures using Mediapipe Hands," International Research Journal of Modernization in Engineering Technology and Science, vol. 4, no. 5, 2022.

Received: August 12, 2025

Volume 38 No. 12s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
[13] R. S. Rohit, R. Kandoi, and S. Kumar, "Hand gesture recognition for human computer interaction through KNN algorithm and Mediapipe," International Journal of Computer Sciences and Engineering, vol. 11, pp. 14–18, 2023.

[14] A. M. J. Al Moustafa, M. S. Mohd Rahim, B. Bouallegue, M. M. Khattab, A. M. Soliman, G. Tharwat, and A. M. Ahmed, "Integrated Mediapipe with a CNN model for Arabic sign language recognition," Journal of Electrical and Computer Engineering, vol. 2023, pp. 1–15, 2023, doi: 10.1155/2023/8870750.

[15] K. Gomase, A. Dhanawade, P. Gurav, S. Lokare, and J. Dange, "Hand gesture identification using Mediapipe," International Research Journal of Engineering and Technology, 2022.

[16] Google, "MediaPipe," 2023. https:// github.com /google/mediapipe.

[17] TensorFlow, "TensorFlow," 2023. https:// www. Tensorflow.org.

Received: August 12, 2025