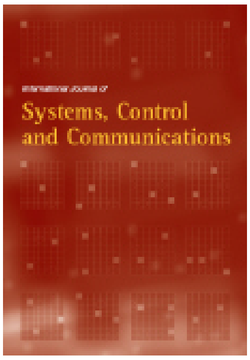


International Journal of Systems, Control and Communications

ISSN online: 1755-9359 - ISSN print: 1755-9340
https://www.inderscience.com/ijscc

Tamil sign language using relational bilevel aggregation graph convolutional network

Shashi Kumar Gowdagere Siddaramaiah, R. Vinoth

DOI: 10.1504/IJSCC.2025.10069425

Article History:

Received:	28 May 2024
Last revised:	23 August 2024
Accepted:	06 September 2024
Published online:	18 February 2025

Tamil sign language using relational bilevel aggregation graph convolutional network

Shashi Kumar Gowdagere Siddaramaiah and R. Vinoth*

Department of Electronics and Communication Engineering,
Manipal Institute of Technology,
Manipal Academy of Higher Education (MAHE),
Manipal, Karnataka-576104, India
Email: shashi.gs@manipal.edu
Email: vinoth.nair@manipal.edu
*Corresponding author

Abstract: Automatic translation of sign language to text facilitates communication between deaf or mute persons and others, including those who are not comfortable with sign language. In this manuscript, Identification of Tamil sign language utilising relational bilevel aggregation graph convolutional network (IDN-TSL-RBAGCN) is proposed. Initially, the data is collected through Tamil sign language gesture images. Artificial lizard search optimisation algorithm (ALSOA) is employed to enhance the weight parameter of relational bilevel aggregation graph convolutional network classifier (RBAGCN), which precisely classifies the Tamil sign language from the identified pattern. The proposed IDN-TSL-RBAGCN method attains 28.76%, 33.68% and 21.78% higher accuracy when compared with existing methods, like Indian sign language recognition utilising wearable sensors with multiple label categorisation (ISL-MLC), deep learning-dependent sign language recognition system for static signs (SLR-DL), and real-time vernacular sign language recognition utilising media pipe and machine learning (SLR-ML) respectively.

Keywords: artificial lizard search optimisation algorithm; Tamil sign language; adaptive-noise augmented Kalman filter; multi-objective matched synchrosqueezing chirplet transform.

Reference to this paper should be made as follows: Siddaramaiah, S.K.G. and Vinoth, R. (2025) 'Tamil sign language using relational bilevel aggregation graph convolutional network', *Int. J. Systems, Control and Communications*, Vol. 16, No. 1, pp.1–16.

Biographical notes: Shashi Kumar Gowdagere Siddaramaiah is currently working as an Assistant Professor-Selection Grade, Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education (MAHE), Manipal, India. His research interests in the area of healthcare applications, biomedical signal processing applications with a focus on emotion and mood profiling and machine learning.

R. Vinoth is an Additional Professor of Electronics and Communication Engineering Department in Manipal Institute of Technology, MAHE, Manipal, Karnataka, India. He obtained his PhD in Computational Biomechanics. He received his MTech in Biomedical Signal Processing and Instrumentation and

BE in Electronics and Communication Engineering. His research interest includes power electronics, biomechanics, medical electronics and social science.

1 Introduction

Language serves as the fundamental mode of communication for individuals. One of the initial methods by which humans communicated before formal languages were developed was sign language, which persists to this day and is particularly valued by the deaf and mute community (Sharma and Singh, 2022; Vidhyalakshmi et al., 2024; Soji and Kamalakannan, 2023). In various settings such as banks and booking counters, understanding sign language, particularly among those who are not familiar with it, poses challenges. The translation of sign language to text or voice has emerged as a solution to alleviate these difficulties. Sign language recognition (SLR) serves as an essential tool in this context, facilitating the translation of sign language gestures to either written text or spoken words. The exploration of SLR began globally about two decades ago, with significant emphasis on American as well as Japanese Sign languages (Sabharwal and Singla, 2024; Sarankumar et al., 2024). Researchers have determined on capturing, recognising, classifying sign language gestures, contributing to the field. Various sign languages, including American, Australian, Korean, and Japanese, have been extensively studied. Numerous methods and algorithms are proposed, leveraging sensor fusion signal processing, image processing, and pattern recognition modes. While applications for international sign languages, like Arabic and Chinese have been made in various fields, contributions to Tamil SLR have been significantly lacking (Jayanthi et al., 2023; Haputhanthri et al., 2023). The aim of the proposed paper is to address this gap by introducing a system designed for the consideration of Tamil sign language, fostering human-computer interaction. The proposed approach focuses on identifying a collection of 31 signs that correspond to the letters in Tamil, comprising 12 vowels, 18 consonants, and one Aayutha Ezhuthu. These signs are depicted through static or dynamic pictures captured from the palm side of the right hand. The aim is to improve the accessibility for people who experience speech or hearing impairments, providing a channel for efficient communication (Ghorai et al., 2023; Das et al., 2023).

The primary contributions of this work are given below:

- In this manuscript, identification of Tamil sign language utilising relational Bilevel aggregation graph convolutional network (IDN-TSL-RBAGCN) is proposed.
- Here, adaptive-noise augmented Kalman filter (ANAKF) is employed for pre-processing, multi-objective matched synchrosqueezing chirplet transform (MOMSCT) is utilised for feature extraction, RBAGCN is proposed to identify Tamil sign language gestures and ALSOA is introduced for enhancing the weight parameters of RBAGCN.
- The proposed model is evaluated with existing models using Tamil sign language gesture images dataset to demonstrating significant improvements in various performance metrics, including recall, F1-score, specificity, error rate.

- The simulation results show the efficacy of the IDN-TSL-RBAGCN approach in recognising Tamil sign language gestures.

Following paper is designed as: Section 2 reveals the literature survey, Section 3 designates the proposed technique, Section 4 shows the results and Section 5 presents the conclusions.

2 Related work

Some papers related to identification of Tamil sign language using deep learning methods are given below.

Gupta and Kumar (2021) have introduced a multi-label classification approach for categorising signs based on their lexical attributes, leading to the final categorisation of each sign. The study provides the findings of its classification of 100 distinct Indian sign language signs. On the forearms of ten different signers, various surface electromyogram and inertial measurement devices were used to record the signals. The identification of the hands' static or dynamic states was made possible by the integrated processing of the signals from both hands. The model improved classification accuracy. But it attains less precision.

Wadhawan and Kumar (2020) have presented deep learning using convolutional neural networks to present robust modelling of static signs in the field of SLR. 35,000 sign images or 100 static signs were gathered for the study from a variety of users. The system efficacy was evaluated using over 50 different CNN models. It attains high accuracy but less precision.

Halder and Tayade (2021) have presented real-time vernacular SLR depending on MediaPipe and machine learning. The aim was to determine the method that simplified SLR depending on machine learning approach. The predictive model was portable and compatible with various smart device types. The technology was more convenient and comfortable since it provides real-time precise identification utilising the support vector machine without the need for wearable sensors. It has high F1-score but also has less recall.

Priya et al. (2023) have presented device designed to serve as a translation system, specifically aimed at converting sign gestures to text. To address the communication gap, have undertaken the task of enhancing communication for disabled individuals through a 'Tamil sign language translator'. This innovative device translates sign gestures into the Tamil language, providing a localised solution. The system was designed to process 31 Tamil alphabets, including 12 vowels, 18 consonants, one Aayudha Ezhuthu. It attains high precision but less recall.

Ranasinghe and Zampieri (2021) have presented offensive language in situations with limited resources by evaluating multilingual transformers' efficiency. It assesses offensive language identification in Indian languages, which include several language families, such as Dravidian (Tamil, Malayalam, and Kannada) together with Indo-Aryan (Bengali, Hindi, and Urdu). The goal was to provide these languages with useful technologies. According to the results, multilingual offensive language identification models out perform their monolingual equivalents. It has high F1-score but also has less recall.

Table 1 tabulates the related work comparison table.

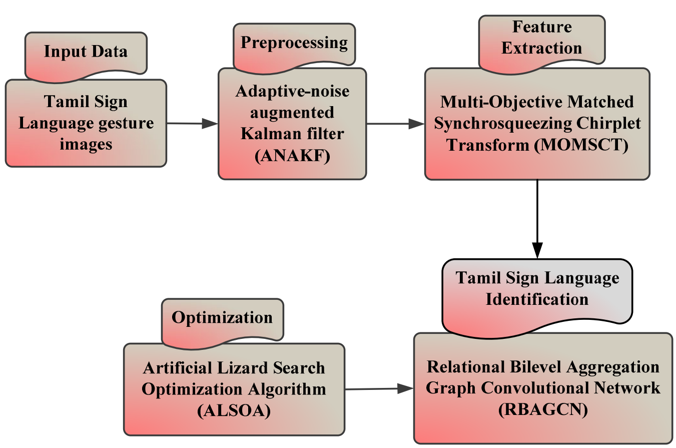
Table 1 Related work comparison table

Author	*Objectives*	*Methods*	*Advantages*	*Disadvantages*
Gupta and Kumar (2021)	To develop a system for recognising Indian sign language gestures	Classic tree-based categorisation	This model improved classification accuracy	It attains less precision
Wadhawan and Kumar (2020)	Develop a deep learning-dependent system for accurate recognition	Convolutional neural networks	It reaches higher accuracy	It attains lesser precision.
Halder and Tayade (2021)	Develop a real-time vernacular sign language identification	Support vector machine	It has high F1-score	It has less recall
Priya et al. (2023)	Develop an efficient Tamil SLR for deaf-mute individuals	Scalar interference feature transform	It attains high precision	It attains less recall
Ranasinghe and Zampieri (2021)	To assess the effectiveness of various multilingual language identification	XLM-R	It has high F1-score	It has less recall

3 Proposed methodology

This segment describes the IDN-TSL-RBAGCN method. The proposed IDN-TSL-RBAGCN is a novel approach for recognising Tamil sign language gestures. Figure 1 shows the block diagram of the proposed methodology.

Figure 1 Block diagram of proposed IDN-TSL-RBAGCN model (see online version for colours)



It uses ANAKF for pre-processing and MOMSCT for feature extraction. The core classification is performed by RBAGCN, where Tamil sign gestures are identified (Moorthy et al., 2024). To enhance RBAGCN's performance, ALSOA is introduced for optimising the weight parameters. The detailed explanation is specified below.

3.1 Data acquisition

The data is collected through Tamil sign language gesture images (https://universe.roboflow.com/vijayalakshmi/tamil-sign/). The dataset used for recognising Tamil sign language gestures consists of images that depict a collection of 31 signs corresponding to the letters in Tamil, which involves 12 vowels, 18 consonants, and one Aayutha Ezhuthu. These signs are represented through both static and dynamic images captured from the palm side of the right hand. The dataset is separated as three main portions: training set, validation set, and test set. The training set has 60% data to train the model, allowing it to learn patterns and relationships. The validation set making up about 20% data. Finally, the test set has 20% data for evaluating the model's performance on unseen data.

3.2 Pre-processing by ANAKF

The captured images undergo pre-processing using the ANAKF (Vettori et al., 2023) to remove noise, ensuring cleaner input for subsequent analysis. ANAKF is a variant of the traditional Kalman filter designed to handle situations where the noise characteristics in the system may vary over time. In the context of image processing, ANAKF can be employed to improve the quality of pictures by adaption of changing noise levels.

3.2.1 Steps for pre-processing

Define a state vector representing the image, where the elements of the vector correspond to the pixel values. Then predicting the next state of the image using a dynamic model that describes the evolution of pixel values over time in the image. It is expressed in equations (1) and (2),

$$\hat{x}_{k|k-1} = F_k \hat{x}_{k-1|k-1} + B_k u_k \tag{1}$$

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k \tag{2}$$

where $\hat{x}_k$ indicates predicted image state at time k, F_k is indicating the uncertainty in the predicted image, B_k denotes input matrix at time k which represents external influences on the image, u_k represents control input at time k, P_k indicates error covariance at time k and Q_k represents noise characteristics affecting the image. Then, compare the predicted state with the actual noisy measurement (captured image). Calculate the Kalman gain, which determines how much the prediction should be corrected based on the measurement. In the ANAKF, the process as well as measurement noise covariance matrices are adapted based upon local characteristics of the image. This adaptation allows the filter to adjust to changes in noise levels. Update the state estimate utilising the Kalman gain and the difference between the predicted and the measured state. It is expressed in equations (3), (4) and (5),

$$K_k = P_{k|k-1} H_k^T \left(H_k P_{k|k-1} H_k^T + R_k \right)^{-1} \tag{3}$$

$$\hat{x}_{k|k-1} = \hat{x}_{k-1|k-1} + K_k \left(z_k - H_k \hat{x}_{k|k-1} \right) \tag{4}$$

(5)

Here, z_k denotes measurement, H_k represents measurement matrix, R_k signifies measurement noise covariance and I symbolises identity matrix. In the context of ANAKF, the key difference lies in Q_k and R_k adaptation of based on local image characteristics, allowing the filter to dynamically adjust to changing noise levels. The specific adaptation equations would depend on the chosen criteria for assessing local noise characteristics. After pre-processing the output image is supplied to the feature extraction.

3.3 Feature extraction with MOMSCT

MOMSCT (Dong et al., 2023) is employed for feature extraction. This step retrieves essential features, such as bounding box, area, centroid, perimeter, equidistance, roundness, number of edges, angles, distance.

3.3.1 Steps for feature extraction process are given below

MOMSCT enhances the time-frequency representation of data, which is crucial for capturing essential characteristics of images. MOMSCT transforms the image data into a time-frequency domain using chirplet transforms, allowing for better identification of features (Kulandaivelu et al., 2024). The synchrosqueezing technique further sharpens the representation by reassigning energy back to the original time domain, making features more distinct. The time-frequency transformation is expressed in equation (6),

$$M_{\hat{a}}(h,\eta)=\sum_{K=1}^{k}T_k(h)f^{i\partial k(h)}\sqrt{\frac{2\sigma\pi}{1-i\sigma\left[\left(\partial'''(h)-\hat{a}\right)\right]^f}-\frac{\sigma\left[\eta-\partial_k(h)\right]^2}{2\left[1-i\sigma\left[\eta-\partial_k(h)\right]\right]}} \tag{6}$$

where $M_{\hat{a}}$ indicates the matrix representing the image features, h capturing the frequency components relevant to the features, η is the covariance value, T_k is the transformation operator applied to the image features, σ is the standard deviation, $\hat{a}$ coefficients and i is the random value ∂ denotes the image features and f is the frequency component. The feature map of MOMSCT is measured using equation (7),

$$\sum_{k=1}^{k}T_k^2(h)\left(\sqrt{\pi}\partial'''(h)+\sqrt{2}(1+\left(\left(\partial'''(h)-\hat{a}\right)^2\right)\right. \tag{7}$$

where ∂''' denotes the extracted features from the pre-processed images. Finally, MOMSCT captured the important features of Tamil sign language gesture images. The extracted features are,

- *Bounding box:* the smallest rectangular box that completely encloses the object in the image. It is defined by the minimum and maximum coordinates of the object.
- *Area:* the total count of pixels within the region of interest in the image. It provides information about the size of the object.
- *Centroid:* the centre point of the object, calculated as the average position of all pixels in object.
- *Perimeter*: the total boundary length of object in the image.

- *Equidistance:* the uniform distribution of points along the boundary or edges of the object.
- *Roundness:* a measure of how much the shape of the object matches an entire circle. It is frequently calculated using the area to square of the object's perimeter ratio.
- *Number of edges:* the count of distinct edges or contours in the object. Edges represent transitions between different intensity levels.
- *Angles:* the angles formed by the edges or contours of the object.
- *Distance:* the spatial separation between different parts of the object.

The extracted features are then fed to RBAGCN for categorisation.

3.4 Tamil sign gestures identification using relational bilevel aggregation graph convolutional network

The extracted features are given into RBAGCN (Yuan et al., 2023). RBAGCN is designed to recognise patterns in the extracted features and identify them corresponding to the Tamil letters. The RBAGCN uses graph convolutional networks to effectively model the relationships between different signs, enhancing its ability to classify gestures accurately. One of its key strengths is its ability to effectively handle relational information, as it operates on graph-structured data, allowing it to capture and utilise the relationships between different features extracted from sign language gestures. This relational understanding enhances the classification process by considering how features interact with one another. Additionally, RBAGCN employs hierarchical representation learning, which enables it to learn both spatial and contextual dependencies crucial for accurately distinguishing between different gestures. The bilevel aggregation mechanism further improves the model's performance by capturing both local and global context information, enhancing its ability to understand complex patterns associated with sign language.

3.4.1 Steps for Tamil sign gestures identification

RBAGCN contains three modules:

1. graph generation module (GGM)
2. similarity-based cluster building module (SCBM)
3. bilevel aggregation module (BiAM).
   - Graph generation module: in the GGM, each gesture or sign can be represented as a node in a graph. For Tamil sign gestures identification, nodes representing different gestures, which could be categorised based on various features such as hand shape, movement, and expressions. Let's denote the input graph representing the features extracted from the Tamil sign language gestures as in equation (8),

$$G = (V, E) \tag{8}$$

Here V indicates set of nodes representing features, and E denotes set of edges representing relationships between these features. Each node in the graph can be initialised with features extracted from the gesture data. The edges in the graph represent the relationships between different gestures.

- Similarity-based cluster building module: the SCBM calculates the similarity between a target gesture and its neighbouring gestures. This could involve using a similarity metric, such as cosine similarity, to quantify how alike two gestures are based on their feature representations. Gestures that are dissimilar to the target gesture, (i.e., those with low similarity scores) would be filtered out. This helps to eliminate irrelevant or distracting gestures that could confuse the recognition process. The remaining gestures, which are similar to the target gesture, would be grouped into clusters. Each cluster would represent a set of gestures that share common characteristics or convey similar meanings. Assign features to each node in the graph. If X_i represents the feature node vector i, the set of all node features can be denoted as in equation (9)

$$X = \{X_1, X_2, ..., X_N\} \tag{9}$$

Here, N denotes number of nodes. The input feature matrix is representing the extracted features from MOMSCT, and be the adjacency matrix representing relationships between these features.

- Bilevel aggregation module: the aggregated features from the lower level and combine them into a more comprehensive representation. This step focuses on understanding the broader context and relationships between different gestures. The RBAGCN performs a graph convolution operation, which is given in equation (10)

$$H - \sigma\left(\hat{D}^{-\frac{1}{2}} \hat{A} \hat{D}^{-\frac{1}{2}} XW\right. \tag{10}$$

where $\widehat{A} = A + I$ denotes adjacency matrix augmented with self-connections, $\hat{D}$ implies degree matrix of $\widehat{A}$, , W signifies weight parameters, and σ is the activation function. The output H undergoes a bilevel aggregation process, capturing both local and global context information. The classifier then utilises a soft max activation to map the aggregated features into probabilities, enabling the final classification decision based on the highest probability. The classification is expressed in equation (11),

$$Y = soft\max(H) \tag{11}$$

where Y indicates the output probability distribution value. Finally, the Tamil sign gestures are identified by RBGCNN. In general, RBAGCN does not express some adaption of optimisation strategies for evaluating optimal parameters to ensure accurate categorisation of Tamil Sign language from the identified pattern. Therefore, ALSOA is proposed to increase the weight parameter (w_i and w_j) of RBAGCN classifier, which precisely identify the Tamil Sign language from the input image.

3.5 Optimisation by artificial lizard search optimisation algorithm (ALSOA)

A metaheuristic optimisation algorithm named ALSOA (Kumar et al., 2021) was inspired by lizards' methods for foraging. It aims to improve the adaptability of optimisation strategies and enhance the weight parameters of the RBAGCN for more accurate classification.

Step 1 Initialisation

Initialise a population of artificial lizards representing potential solutions. Each lizard corresponds to a set of parameters that need to be optimised. These parameters could include weights, biases, or any other relevant parameters in the RBAGCN. It is expressed in equation (12),

$$F = \begin{bmatrix} F_{1,1} & F_{1,2} & \dots & F_{1,b} \\ F_{2,1} & F_{2,2} & \dots & F_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ F_{n,1} & F_{n,2} & \dots & F_{n,b} \end{bmatrix} \tag{12}$$

where F is the population of artificial lizards, $F_{n,b}$ denotes the n^{th} dimension of i^{th} leaping lizard.

Step 2 Random generation

After initialisation process, create the weight parameter is generated randomwise with the help of ALSOA.

Step 3 Fitness function

Estimate the fitness of every lizard in the population. The fitness function is typically based on the performance of the RBAGCN model. It is labelled in equation (13),

$$Fitness\ Function = optimizing\left(w_i \text{ and } w_j\right) \tag{13}$$

Step 4 Exploration phase

This is employed to assess the weight parameter in optimisation. Implement the exploration and exploitation phases inspired by the foraging behaviour of lizards. Lizards explore their surroundings (search space) to discover new areas and exploit known areas with promising solutions, which is given in equation (14),

$$m\left(\alpha_j(h)\right) = \left[\alpha_{jd}(h), \frac{\alpha_{jd}{}^2(h) - d}{b - g}, \alpha_{jn}(h), \frac{-\alpha_{jd}{}^2(h) + u}{b - g}\right] \tag{14}$$

Here m denotes objective of image, α_j denotes value identified in the convolutional complex, $b - g$ specifies decrease time taken, $\alpha_{jn}(h)$ denotes image identification in the network. The exact value spatial modulation is calculated by equation (15),

$$e\left(\alpha_j(h)\right) = \left[0, -\frac{m}{b-g}, 0, \frac{e}{b-g}\right] \tag{15}$$

Here α_j the spatial modulation constant, e is the foraging behaviour of lizards, $\alpha_j(h)$ is the computational complexity, m is the objective value.

Step 5 Exploitation phase for optimising (w_i and w_j).

The exploitation phase of artificial lizards is given in equation (16),

$$b = \left(F_{jd}^2 F_{jn}^2 w_{jd}^2 w_{jn}^2 \left(\alpha_{jd}(h)\right) - \left(n_{ab}(h)\right)\right)^2 \tag{16}$$

Consider b as detection rate, F_{jd}^2 as error rate on the network, F_{jn}^2 as number of error rate on the network, w_{jd}^2 specifies detection of the error rate, w_{jn}^2 specifies weight computation in the network system.

Step 6 Termination

Finally, the parameter (w_i and w_j) is increased with the help of ALSOA or repeat step 3 until it reaches the halting criteria $F = F + 1$. RBAGCN is optimised with ALSOA effectively for classifying Tamil sign language from the identified pattern with better accuracy. Thus the proposed IDN-TSL-RBAGCN methods effectively enhance the weight parameters of the RBAGCN for more accurate classification.

4 Result and discussion

The proposed IDN-TSL-RBAGCN is executed in Python. The performance is examined with the mentioned metrics. The proposed IDN-TSL-RBAGCN technique is analysed with existing techniques, like CRP-TE-FSET (Gupta and Kumar, 2021), CRP-RF-FSET (Wadhawan and Kumar, 2020) and CRP-SVM-FSET (Halder and Tayade, 2021) respectively.

- *Accuracy:* the rate of appropriately calculated instances to the total instances. This is measured by equation (17),

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

Here, TP indicates true positive, TN defines true negative, FP indicates false positive, FN indicates false negative.

- *Precision:* the rate of appropriately predicted positive observations to total predicted positives. This is measured by equation (18),

$$Precision = \frac{TP}{TP + FP} \tag{18}$$

- *Specificity:* the proportion of appropriately predicted negative instances to the total actual negatives. This is computed by equation (19),

$$Specificity = \frac{TN}{TN + FP} \tag{19}$$

- *Recall:* the rate of appropriately estimated positive observations to the total true positives. This is measured by equation (20),

$$Recall = \frac{TP}{TP + FN} \tag{20}$$

- *F1 score:* this is the harmonic mean of precision as well as recall. This is measured by equation (21),

$$F1\text{-}score = 2\frac{Precision \times Recall}{Precision + Recall} \tag{21}$$

- *Error rate:* the rate of inappropriately predicted instances to the total samples. This is measured by equation (22),

$$Error\ Rate = \frac{FP + FN}{TP + TN + FP + FN} \tag{22}$$

- *Computational time:* the time consume by the method to complete the process of identification. It can be measured in seconds depending on the scale of the application.

Figure 2 Accuracy analysis (see online version for colours)

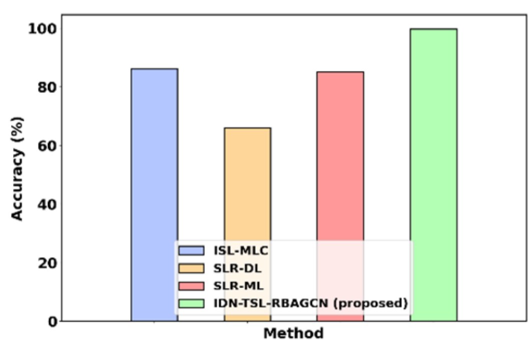


Figure 2 represents accuracy analysis. The IDN-TSL-RBAGCN reaches 50.56%, 20.76%, 35.97%, 20.67% better when evaluated to the existing ISL-MLC, SLR-DL and SLR-ML models.

Figure 3 displays precision analysis. The proposed IDN-TSL-RBAGCN attains 30.56%, 24.76%, 20.67% greater when evaluated to the existing ISL-MLC, SLR-DL and SLR-ML models.

Figure 4 shows recall analysis. The proposed IDN-TSL-RBAGCN achieves 30.66%, 23.56% and 20.47% higher when evaluated to the existing ISL-MLC, SLR-DL and SLR-ML models.

Figure 5 displays F1-score analysis. The IDN-TSL-RBAGCN reaches 20.26%, 24.56% and 30.47% higher when evaluated to the existing ISL-MLC, SLR-DL and SLR-ML models.

Figure 3 Precision estimation (see online version for colours)

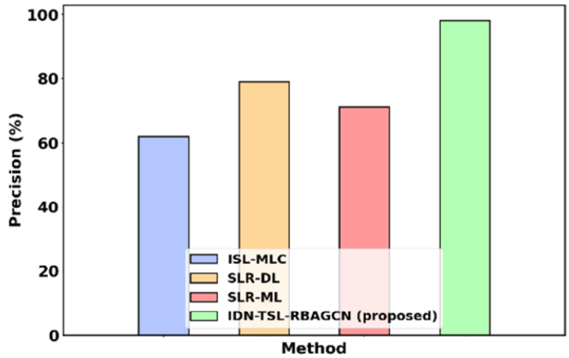


Figure 4 Recall analysis (see online version for colours)

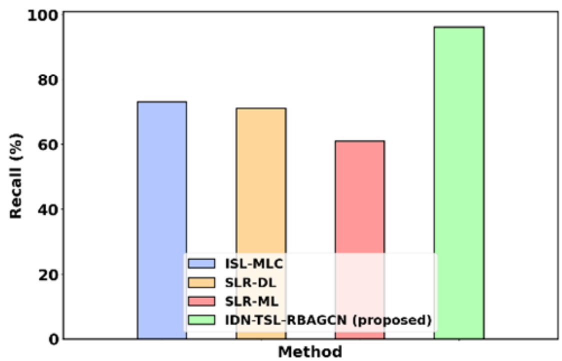


Figure 5 F1-score analysis (see online version for colours)

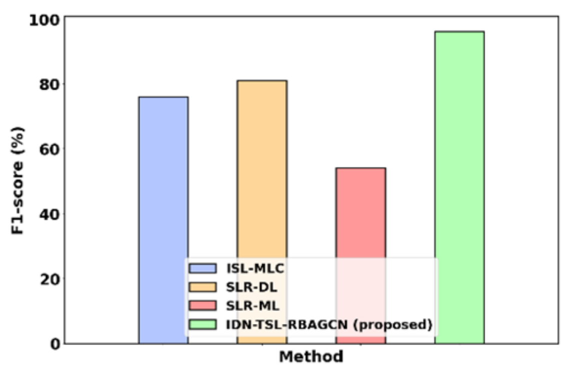


Figure 6 shows specificity analysis. The proposed IDN-TSL-RBAGCN technique results in Phi coefficient that are 20.26%, 24.56% and 30.47% higher when evaluated to the existing ISL-MLC, SLR-DL and SLR-ML models.

Figure 6 Specificity analysis (see online version for colours)

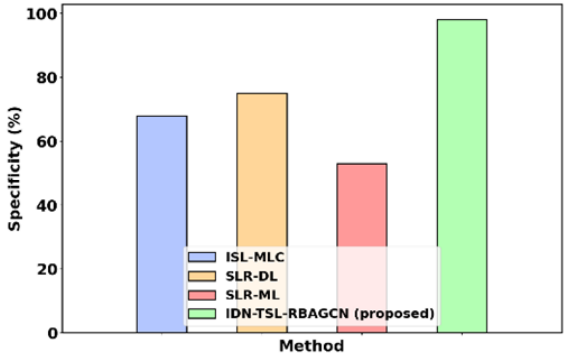


Figure 7 Error rate analysis (see online version for colours)

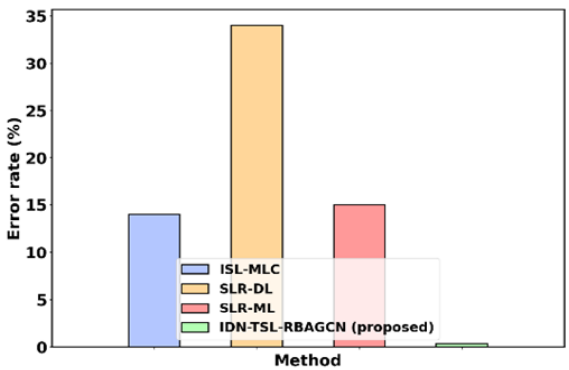


Figure 7 shows error rate analysis. The IDN-TSL-RBAGCN technique results in error rate 10.26%, 14.56% and 10.47% lower when evaluated to the existing ISL-MLC, SLR-DL and SLR-ML models.

Figure 8 Computational time analysis (see online version for colours)

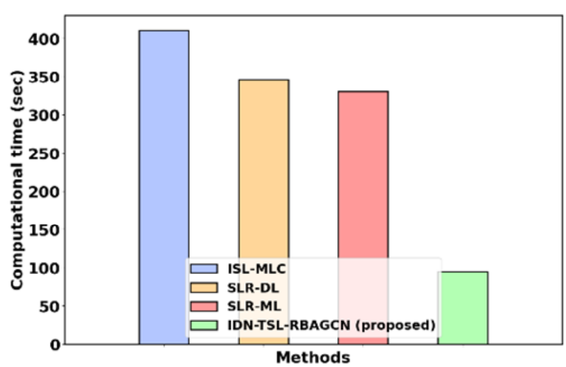


Figure 8 shows computational time analysis. The proposed IDN-TSL-RBAGCN technique results in computational time that are 15.26%, 24.56% and 30.47% higher for the classification of Tamil sign language identification when evaluated to the existing ISL-MLC, SLR-DL and SLR-ML models.

4.1 Ablation study

Table 2 tabulates ablation experiments. Baseline model does not use any pre-processing techniques and it attains low accuracy and low sensitivity. The noise reduction only configuration includes only the noise reduction pre-processing method. It provides better results than baseline model. From the table it clearly shows the proposed method with all processing steps provides high accuracy and sensitivity.

Table 2 Ablation experiments

Configuration	*Noise reduction*	*Feature extraction*	*Optimisation*	*Accuracy (%)*	*Sensitivity (%)*
Baseline model	No	No	No	92.67	90.97
Noise reduction only	Yes	No	No	93.18	92.52
Feature extraction only	No	Yes	No	96.14	94.96
Optimisation only	No	No	Yes	94.39	94.71
Noise reduction + Feature extraction	Yes	Yes	No	97.65	95.86
Proposed method with all processing steps	Yes	Yes	Yes	99.48	98.92

4.2 Discussion

The proposed IDN-TSL-RBAGCN technique is a promising approach to bridge communication barriers for the deaf or mute community. The use of RBAGCN for classification, coupled with optimisation through ALSOA, shows significant potential for accurate recognition of Tamil sign language gestures. However, there remains room for discussion regarding scalability, real-time implementation feasibility, and the extension of this methodology to other sign languages. Additionally, future research could explore enhancing the strength of the model against variations in lighting conditions, hand orientation, and gestures' speed. Furthermore, integrating user feedback mechanisms to continuously refine and improve the system's performance would be a valuable avenue for exploration. Overall, this work lays a foundation for advancing assistive technologies for the deaf or mute community and opens doors for further innovation in SLR systems.

5 Conclusions

In conclusion, the IDN-TSL-RBAGCN offers a promising approach to bridge communication gaps between deaf or mute individuals and others. By utilising a combination of pre-processing techniques, feature extraction algorithms, and a specialised classification network, the system demonstrates efficient recognition of TSL gestures. Moreover, the integration of the ALSOA enhances the classification precision further. Through comprehensive evaluation against existing methods, the IDN-TSL-RBAGCN method exemplifies its potential as an important tool for facilitating inclusive communication and interaction within the deaf community and beyond.

References

Das, S., Imtiaz, M.S., Neom, N.H., Siddique, N. and Wang, H. (2023) 'A hybrid approach for Bangla sign language recognition using deep transfer learning model with random forest classifier', *Expert Systems with Applications*, Vol. 213, p.118914.

Dong, F., Yang, J. and Cai, Y. (2023) 'Multi-objective matched synchrosqueezing chirplet transform for fault feature extraction from marine turbochargers', *IEEE Access*, Vol. 11, pp.80702–80715.

Ghorai, A., Nandi, U., Changdar, C., Si, T., Singh, M.M. and Mondal, J.K. (2023) 'Indian sign language recognition system using network deconvolution and spatial transformer network', *Neural Computing and Applications*, Vol. 35, No. 28, pp.20889–907.

Gupta, R. and Kumar, A. (2021) 'Indian sign language recognition using wearable sensors and multi-label classification', *Computers & Electrical Engineering*, Vol. 90, p.106898.

Halder, A. and Tayade, A. (2021) *Real-time Vernacular Sign Language Recognition using Mediapipe and Machine Learning*, ISSN: 2582:7421, [online] https://www.ijrpr.com.

Haputhanthri, H.H., Tennakoon, H.M., Wijesekara, M.A., Pushpananda, B.H. and Thilini, H.N. (2023) 'Multi-modal deep learning approach to improve sentence level Sinhala sign language recognition', *The International Journal on Advances in ICT for Emerging Regions*, Vol. 16, No. 2.

Jayanthi, P., Bhama, P.R. and Madhubalasri, B. (2023) 'Sign language recognition using deep CNN with normalised keyframe extraction and prediction using LSTM', *Journal of Scientific and Industrial Research*, Vol. 82, No. 7, pp.745–755.

Kulandaivelu, G., Taluja, A., Gawas, M. and Nath, R.K. (2024) 'Automated breast cancer diagnosis optimized with higher-order attribute-enhancing heterogeneous graph neural networks using mammogram images', *Biomedical Signal Processing and Control*, Vol. 97, p.106659.

Kumar, N., Singh, N. and Vidyarthi, D.P. (2021) 'Artificial lizard search optimization (ALSO): a novel nature-inspired meta-heuristic algorithm', *Soft Computing*, Vol. 25, No. 8, pp.6179–6201.

Moorthy, M.C., Sekhar, J.C., Inayatullah, S. and Agrawal, G. (2024) 'Optimized brain tumor identification via graph sample and aggregate-attention network with artificial lizard search algorithm', *Knowledge-Based Systems*, Vol. 302, p.112362.

Priya, C.B., Ibrahim, S.S., Thangam, D.Y., Francis, X. and Jency, P. (2023) 'Tamil sign language translation and recognition system for deaf-mute people using image processing techniques', in *Proceedings of the 2023 International Conference on Software Engineering and Machine Learning*.

Ranasinghe, T. and Zampieri, M. (2021) 'An evaluation of multilingual offensive language identification methods for the languages of India', *Information*, Vol. 12, No. 8, p.306.

Sabharwal, S. and Singla, P. (2024) 'Translation of Indian sign language to text – a comprehensive review', *International Journal of Intelligent Systems and Applications in Engineering*, Vol. 12, No. 14, pp.309–319.

Sarankumar, R., Ramkumar, M., Vijaipriya, K. and Velselvi, R. (2024) 'Bidirectional gated recurrent unit with auto encoders for detecting arrhythmia using ECG data', *Knowledge-Based Systems*, Vol. 294, p.111696.

Sharma, S. and Singh, S. (2022) 'Recognition of Indian sign language (ISL) using deep learning model', *Wireless Personal Communications*, Vol. 123, No. 1, pp.671–692.

Soji, E.S. and Kamalakannan, T. (2023) 'Indian sign language recognition using surf feature extraction and MDAE for patient disability discussion', in *Computational Intelligence for Clinical Diagnosis*, pp.445–459, Springer International Publishing, Cham.

Vettori, S., Di Lorenzo, E., Peeters, B., Luczak, M.M. and Chatzi, E. (2023) 'An adaptive-noise augmented Kalman filter approach for input-state estimation in structural dynamics', *Mechanical Systems and Signal Processing*, Vol. 184, p.109654.

Vidhyalakshmi, M.K., Jagadeesh, K., Dharmaraj, B.S., Sankar, N.P. and Kumar, R.G. (2024) 'Indian sign language recognition using transfer learning with efficient Net', *Smart Science*, Vol. 12, No. 2, pp.269–280.

Wadhawan, A. and Kumar, P. (2020) 'Deep learning-based sign language recognition system for static signs', *Neural Computing and Applications*, Vol. 32, No. 12, pp.7957–7968.

Yuan, L., Huang, G., Li, F., Yuan, X., Pun, C.M. and Zhong, G. (2023) 'RBA-GCN: relational bilevel aggregation graph convolutional network for emotion recognition', *IEEE/ACM Transactions on Audio, Speech and Language Processing*, Vol. 31, pp.2325–2337.

Websites

https://universe.roboflow.com/vijayalakshmi/tamil-sign/.