Journal of Intelligent Systems and Internet of Things
Vol. 18, No. 02, PP. 304-314, 2026
DOI: https://doi.org/10.54216/JISIoT.180221

Integrating Artificial Intelligence Driven Computer Vision Framework for Enhanced Sign Language Recognition in Hearing and Speech-Impaired People

Inderjeet Kaur[1], P. Udayakumar[2], B. Arundhati[3], M. V. Rajesh[4], Naif Almakayeel[5,*], Elvir Akhmetshin[6]

[1]Department of Computer Science and Engineering, Ajay Kumar Garg Engineering College Ghaziabad, India

[2]Department of Computer Science and Engineering, Akshaya College of Engineering and Technology, Kinathukadavu, Coimbatore, India

[3]Department of EEE, Vignan's Institute of Engineering for Women, Visakhapatnam, Andhra Pradesh, India

[4]Department of Information Technology, Aditya University, Surampalem, India

[5]Department of Industrial Engineering, College of Engineering, King Khalid University, Abha, Saudi Arabia

[6]Department of Science, Innovations and Technology, Mamun University, 220900, Khiva, Uzbekistan

Emails: inderjeetk@gmail.com; udayakumarp@acetcbe.edu.in; b_arundhati@rediffmail.com; rajesh.masina@adityauniversity.in; halmakaeel@kku.edu.sa; akhmetshin@mamunedu.uz

Abstract

Sign language (SL) detection and classification for deaf persons is an essential application of machine learning (ML) and computer vision (CV) techniques. It covers emerging forms, which acquire SL implemented by entities and convert them into auditory or textual output. It is highly significant to understand that determining a correct and robust SL detection approach is a very challenging due to many tasks such as alterations in occlusions, and lighting states in hand actions and forms. Consequently, the CV and ML models is must for testing and training. A Hand gesture detection method discovers beneficial for hearing and speaking-impaired individuals by creating usage of convolutional neural network (CNN) and human-computer interface (HCI) for classifying the constant signals of SL. In this article, an Improved Fennec Fox Algorithm for Deep Learning-Based Sign Language Recognition in Hearing and Speaking Impaired People (IFFADL-SLRHSIP) technique is proposed. The presented IFFADL-SLRHSIP technique main intention is to provide effectual communication between deaf and dumb persons and normal persons utilizing CV and artificial intelligence techniques. In the IFFADL-SLRHSIP model, the enhanced SqueezeNet model is used to capture the intricate patterns and nuances of SL gestures. For detection of the SL classification process, the recurrent neural network (RNN) method is used. To optimize model performance, the improved fennec fox algorithm (IFFA) is applied for parameter tuning, enhancing the model's precision and efficiency. The experimental outputs of the IFFADL-SLRHSIP algorithm are legalized on the SL dataset. The simulation outcomes demonstrate the greater outcomes of the IFFADL-SLRHSIP approach in terms of diverse measures.

Received: April 07, 2025 Revised: June 26, 2025 Accepted: August 18, 2025

Keywords: Sign Language Recognition; Fennec Fox Algorithm; Computer Vision; Deep Learning; Hearing-Impaired Person

1. Introduction

As per the world health organization (WHO), 466 millions of people suffered from deaf in 2020. SL is a vital device for them to interconnect with one another. In recent times, analysis of deafness has accepted more difficult socio-cultural perceptions, increasing the problem of community formation, preservation, and identity language

system [1]. As a means to build individual communication, SLs do not segment the similar standards all over the world. As an alternative, cultural variance along with another factor creates a massive variance among the SL. The physical module of SLs generally contains the drive of the hand motion and forearm [2]. The creation of sentences is also several in alphabets, grammar, and vocabulary amongst several SLs. SL constantly considers the major mode of verbal communication among both deaf and dumb individuals. Communicating with such kind of people is very helpless and thus based on hand signs [3]. Visual gestures and signs are an active portion of automated SL (ASL) that deals with hearing-impaired individuals' easy and reliable communication. It contains exact code signs in which each signal brings a precise meaning relating to communication [4]. Numerous approaches are accessible to get gestural data. Controlling only the main 2 important methods are Sensor-related and Vision-related methods [5]. This method treats the image qualities such as coloring and texture portion was essential to the particular hand signal [6]. Hence, this paper aims to propose an SL recognition approach for deaf people that combines many advanced recognition algorithms and inertial sensors to complete more precise SL detection outcomes [7].

Recently, important evolution was made in research on gesture recognition or SL. Recently, typical SL detection approaches usually depend on sensor or vision technology to take related data from hand actions [8]. This data is united with both ML and deep learning (DL) models for specifying the equivalent meaning of SL. The most commonly used approach for SL detection depends on CV [9]. This technique mainly utilizes a camera as its major device to take data and then it is studied to specify the meaning of SL. It is reached over the usage of one or many cameras. Additionally conventional cameras are dedicated cameras with particular functions, which could be employed for SL detection [10].

This article proposes an Improved Fennec Fox Algorithm for Deep Learning-Based Sign Language Recognition in Hearing and Speaking Impaired People (IFFADL-SLRHSIP) technique. In the IFFADL-SLRHSIP model, the enhanced SqueezeNet method is employed to get the intricate patterns and hints of SL gestures. For detection of the SL classification process, the recurrent neural network (RNN) method is used. To optimize model performance, the improved fennec fox algorithm (IFFA) is applied for parameter tuning, enhancing the model's precision and efficiency. The experimental outcomes of the IFFADL-SLRHSIP algorithm are authorized on the dataset of SL.

2. Literature Survey

Kothadiya et al. [11] presented a progressive convolutional-based hybrid Inception structure to increase the detection of remote signs. The major assistances are to improve Inception-V4 with enhanced backpropagation (BP) over homogeneous networks. Moreover, an ensemble learning architecture with several CNNs is exploited and introduced to additionally upsurge the detection accuracy and strength of remote SL recognition systems. Buttar et al. [12] presented a DL-based approach. This work utilizes dual various DL-based techniques. A real-time American SL indicator was generated by utilizing the skeleton method, which constantly classifies constant indications in SL in the majority of cases with a DL Method. In the next DL technique, a SL indicator was generated for fixed signs by YOLO-v6. These applications are extremely useful for users of SL and beginners to practice SL in the real world. Once training either model individually for constant or static signs, a solitary approach was generated by utilizing a hybrid method. In [13], a system that identifies hand gestures and poses from the ISL in real-time. Thirty-three hand poses and several signals is recognized with features of grid-based. Smartphone cameras is applied to seize the SL and are communicated to locate unknown servers in a variety of frames. It is easy to understand and hardware peripherals practice is hence prevented. For hand tracking and detection, methods like segmentation of skin color, face recognition, and object stabilization are applied. The image grid-feature removal method was applied to the image. The k-nearest neighbours' method is utilized to categorize the hand poses. The hidden Markov Model chain is struggling through the observation and motion sequence of in-between hand poses to categorize the gestures. The authors [14] developed a hybrid technique, which contains deep TL-based CNNs with the classifier of RF for the automatic detection of Bangla SL. Moreover, a background removal method that extracts redundant features from the sign images has also been suggested. Palanivel et al. [15] proposes an AVRS-based Lip-to-Speech convertor. This device utilizes a webcam for recording the lip movements of the speaker, which was formerly converted into speech by a microcontroller of AVR. A microprocessor of AVR, a camera, and a speaker module are elements of the organization framework. The microcontroller of AVR handles real-time footage, which was taken by the camera unit. Podder et al. [16] implemented a SL detection system that may identify Arabic SL (ASL) from noted RGB videos. Additionally, operational layer-based multi-layer perceptron SelfMLP is presented to construct CNN-LSTM-SelfMLP techniques for ASL detection. MobileNet-V2 and ResNet-18 based CNN supports and 3 SelfMLPs is applied to build six dissimilar techniques for performance compared to ASL detection. Das et al. [17] designed a new method for converting UrSL with the UrSL-CNN approach, a CNN framework specially intended for this resolution. The work showed research using dual datasets comprising 1500 and 78,000 images, using a method including four components: pre-processing, data collection, prediction, and categorization.

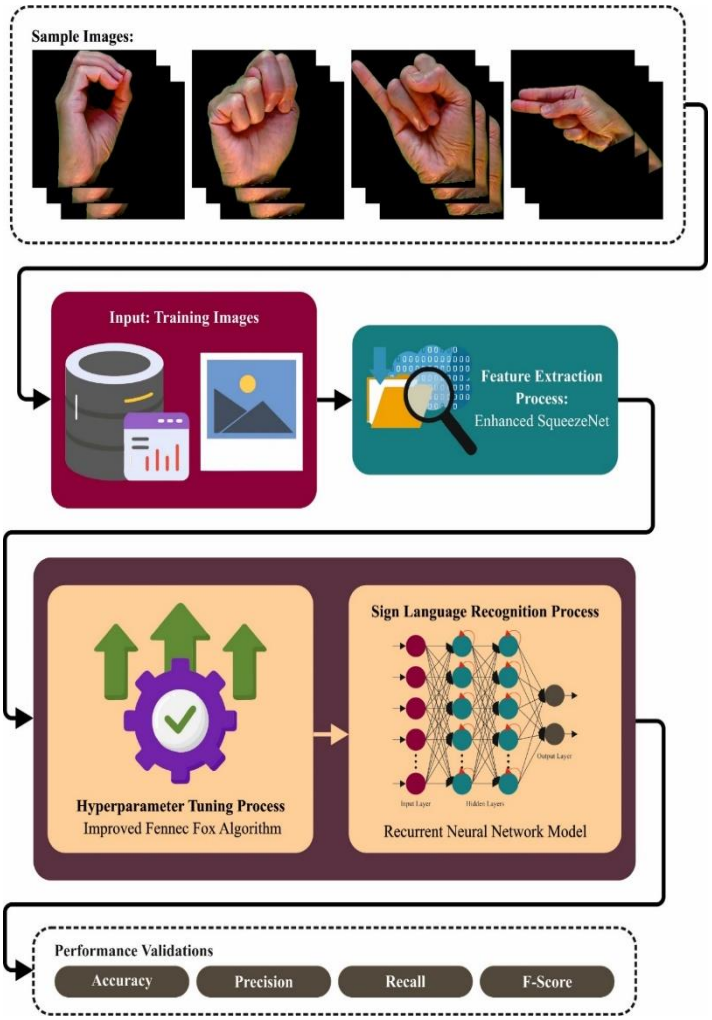

Figure 1. Working flow process of IFFADL-SLRHSIP technique

3. Materials and Methods

In this article, the IFFADL-SLRHSIP technique is proposed. The presented IFFADL-SLRHSIP model main goal is to provide effectual communication among deaf and dumb persons and normal persons utilizing CV and AI techniques. To accomplish that, the IFFADL-SLRHSIP model contains three various stages namely feature extraction, classification, and parameter optimizer. Fig. 1 portrays the overall workflow of IFFADL-SLRHSIP model.

A. Stage I: Enhanced SqueezeNet Feature Extraction

Primarily, the IFFADL-SLRHSIP technique takes place in the enhanced SqueezeNet model and is used to get the intricate patterns and hints of SL gestures [18]. SqueezeNet is a lightweight network architecture, which attains a recognition accuracy in comparison with that of AlexNet in the dataset of ImageNet. Simultaneously, the model parameter counts is decreased. The SqueezeNetvl.l needs 2.4 x less calculation and a somewhat smaller amount of parameters than the unique SqueezeNet without losing precision. This model contains dual convolutional layers, 8 Fire modules, 3 maximal softmax, pooling, and an average global pooling layer. Between others, the Fire module is the basic SqueezeNet that utilizes a larger amount of 1x1 convolution kernels and a combination of 1x1 and 3x3 convolutional kernels rather than 3x3 convolution kernels. This substitution significantly decreases the parameter counts and model's calculation but improving the model depth and the detection accuracy has no effect. Fig. 2 represents the infrastructure of SqueezeNet.

SqueezeNet enables utilization on right-hand platforms with source and calculating power restrictions, SqueezeNet is chosen as the foundation of detection system. Additionally, the enhanced SqueezeNet approach is utilized by improving the fire module. Furthermore, the mixture of 3x1 and 1x3 convolution is utilized to get a similar sensory domain as that of the 3x3 convolutional that results in a 33% decline in the amount of parameters. Meanwhile, the efficient channel attention (ECA) method is presented.

Especially the intention of the entire unit is defined. The Inception model's series a mixture of 1xN and Nx1 convolution kernels is employed rather than NxN convolutional kernels that not only attain the similar sensation region but also reduce the similar parameter counts. Hence, the 3x3 convolutional in the unique fire module into similar 3x1 and 1x3 complications is divided. These decompositions not only decrease the parameters of the model but also increase the network's depth, hence enhancing the feature extraction proficiency. Initially, the input features are channel compacted by 1x1 convolutional in the squeeze phase. Next, in the expand stage, the channel counts is expanded by 1x1, 3x1, and 1x3 convolutions, correspondingly; lastly, the outcomes of various sizes of the convolutional kernel are output to the following layer.

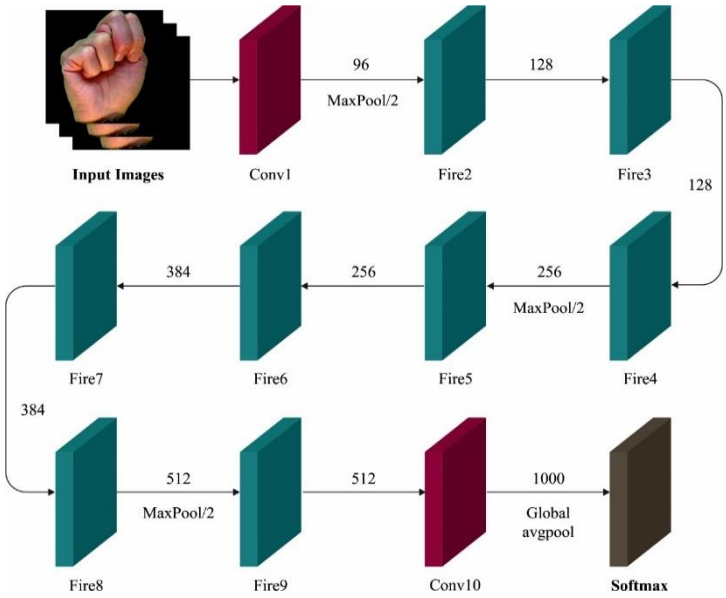

Figure 2. Structure of SqueezeNet model

Then, the ECA module is presented. The mechanism of ECA improves the feature representation of CNNs by taking inter-channel dependency, employing 1D convolutions rather than fully connected (FC) layers, thus considerably decreasing the parameters of the network.

For input feature $X_{Input} \in \mathbb{R}^{C\times H\times W}$, the average feedback of every channel was gained afterward average pooling globally to achieve related information globally to get the features of the $(C, 1,1)$ size. Formerly, the feature learning is carried out with 1D convolution by an adaptable convolutional kernel to take the local dependency between channels. According to the activation function of Sigmoid, the convolution layer output is non-linearly converted such that each parameters within the range of (0-1), and the attention weight vector is gained. The vector of attention weight is calculated based on the succeeding Eq. (1):

$$X_W = \sigma\left(Conv_{lDk}\left(GAP\left(X_{Input}\right)\right)\right), \quad (1)$$

whereas GAP signifies average global pooling and $Conv_{lDk}$ means 1D convolutional with convolutional kernel k. In this work, k is fixed to be 3x3. Furthermore, σ denotes $Sigmoid$. Now, the output feature X_{Output} was formulated below:

$$X_{Output} = X_W \otimes X_{Input}, \quad (2)$$

Here $\otimes$ signifies the multiplication of inter-element; $X_W \in \mathbb{R}^{C\times 1\times 1}$ represents the weighted feature mapping.

B. Stage II: RNN-based SL Recognition

For detection of SL classification process, the RNN approach is used. RNNs exceed the conventional feed-forward neural network, enhanced by the combination of recurrent edges that summarize adjacent temporal stages [19]. This expansion informs temporal dimensions of the structural design, presenting a subtle theory of time inside the network architecture. However, RNNs are not obligatory to show clear temporal stages inside standard limits, the combination of recurrent edges can actively provide growth to different stages, including auto connections in the process. After getting a sequence of vectors (x1,x2,..,xn) as input, they give a next series (h1,h2,..,hn) with data on

the received order at every phase. RNNs handle variable-length classifications by effective as a recurrent hidden layer (HL), whereas every starting is subjective by the activation from the preceding iteration.

$$h_t^R = v(S.xt + J.h_{t-1}^R + b) \quad (3)$$

$$K_t^R = v(S^0 \cdot h_t^R + b^0) \quad (4)$$

Whereas, v means smooth, bounded function, such as a hyperbolic tangent or logistic function and K_t^R refers to the outcome of a RNN, or its predicted value. The recurrent concealed state h_t^R at every time step, t is defined by the input vector x_t, its previous HL h_{t-1}^R, and the bias as b inside the NN framework. The weight matrices signified as S and J, perform as discriminating filters, determining the magnitude of importance recognized to current input and the previous HL. This iterative feedback loop directs the change of their parameters, iteratively enhancing the weights till an additional decline of the error becomes unachievable.

C. Stage III: IFFA-based Parameter Optimizer

To optimize model performance, the IFFA is applied for parameter tuning, enhancing the model's accuracy and efficiency. FFA is a metaheuristic technique that is dependent upon the capability of Fennec foxes to escape and tunnel from harsh hunters [20]. Generally, the Fennec fox's powerful dig skill and escape behaviour are the main motivation and resource of their projected FFO. They typically live in desert and semi-desert regions and choose constant humps, which are very simple to tunnel. It typically hunts at night, consumes extensively, survives in clusters, and has an energetic character.

1. Initialize

In the initialize stage, the Fennec foxes were positioned at random in the searching space by using the Eq. (5):

$$Y_i : y_{i,j} = lb_j + r \cdot \left(upb_j - lowb_j\right), i = 1,2,\cdots,N, j = 1,2,\cdots,m \quad (5)$$

Here, Y_i means the ith Fennec fox, N represents the overall Fennec foxe counts, m means the amount of decision variables, r is the produced value at random among [0,1], $lowb$ and upb represents the lower and upper limits, correspondingly.

In Eq. (6), Y refers to the matrix of the population composed of every Fennec fox:

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_i \\ \vdots \\ Y_N \end{bmatrix}_{N\times m} \quad N \times \begin{bmatrix} y_{1,1} & \cdots & y_{1j} & \cdots & y_{1,m} \\ \vdots & \ddots & \vdots & \iddots & \vdots \\ y_{i,1} & \cdots & y_{ij} & \cdots & y_{i,m} \\ \vdots & \iddots & \vdots & \ddots & \vdots \\ y_{N,1} & \cdots & y_{Nj} & \cdots & y_{N,m} \end{bmatrix}_{N\times m} \quad (6)$$

While $Y_i = (y_{i,1}, y_{i,2}, \ldots, y_{i,m})$ denotes the i^{th} Fennec fox; the vector of column signifies the candidate value. To resolve the values of an objective function, the vector model set in Eq. (7) is employed for exhibiting:

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix}_{N\times 1} = \begin{bmatrix} F(Y_1) \\ \vdots \\ F(Y_i) \\ \vdots \\ F(Y_N) \end{bmatrix}_{N\times 1} \quad (7)$$

While F signifies the vector comprising the values of an objective function, F_i indicates the value of main function of Fennec fox i.

2. Location update

The location renewal phase of Fennec foxes is mostly implemented as per the Fennec foxes tunneling victim and escaping hunters.

Stage 1: Exploitation

Throughout the victim-hunting phase, the Fennec fox travels an area with R radius. This property allows the method to tactic the global optimum outcome more thoroughly. During the growth phase, the computation method equivalent to the location of Fennec fox upgrade was given below:

$$y_{i,j}^1 = y_{ij} + (2 \cdot r - 1) \cdot R_{ij}, \quad (8)$$

$$R_{i,j} = \alpha \cdot \left(1 - \frac{t}{T}\right) \cdot y_{ij}, \quad (9)$$

$$Y_i = \begin{cases} Y_i^{P1}, & F_i^{P1} < F_i \\ Y_i, & else \end{cases} \quad (10)$$

While Y_i^{P1} denotes the novel location of the ith Fennec fox, F_i^{P1} refers to the corresponding value of an objective function, t represents the present amount of iterations, T means the maximum iteration count, α refers to a fixed constant, which is set as 0.2.

Stage 2: Discover

In this phase, the Fennec foxes' excellent capability to escape permits the method to evade being stuck in local goals. In this stage, the equivalent mathematical method for upgrading the location of Fennec foxes' is given below:

$$Y_i^{rand} : y_i^{rand} = y_{kj}, k \in \{1,2,\cdots,N\}, i = 1,2,\cdots,N \quad (11)$$

$$y_{ij}^{p2} = \begin{cases} y_{ij} + r \cdot \left(\gamma_{ij}^{rand} - I \cdot y_{ij}\right), & F_i^{rand} < F_i \\ y_{ij} + r \cdot (Jij - y_{ij}^{rand}), & else \end{cases} \quad (12)$$

$$Y_i = \begin{cases} Y_i^{P2}, & F_i^{P2} < F_i \\ Y_i, & else \end{cases} \quad (13)$$

Here, Y_i^{rand} denotes that Fennec fox has run away, F_i^{rand} and F_i^{P2} represents a value of the objective function, Y_i^{P2} refers to the upgraded location of Fennec Fox; I specifies a produced value at random within the interval of [1, 2]. If the method is completely set, in stages 1 and 2, the algorithm finishes an iteration.

The FFA provides major benefits in resolving optimizer issues; it will face assured tasks, the probability of becoming trapped in local goals, and restricted performance in definite states. To overwhelm these restrictions and additionally improve the model a multi-strategy improved FFA named IFFA is developed depending upon the novel FFA. In IFFA, many strategies are united to tackle these challenges. At first, the sin chaotic mapping tactic was combined into the initialization stage in order to improve the uniform distribution of early population. This aids in enhancing exploration abilities and evades convergence of premature.

3. Sin chaotic mapping strategy

The strategy is detected to own an advanced level of chaotic behaviour when equated to the Logistic chaotic map method. In the initialization phase, inserting a sin chaotic map creates the distribution of the population more even. The sin chaotic mapping mathematical formulation is mentioned below:

$$\begin{cases} y_{n+1} = \sin\dfrac{2}{y_n}, & n = 0,1,\cdots,N \\ -1 < y_n < 1, & y_n \neq 0 \end{cases} \quad (14)$$

The mapping shows a uniform distribution. Therefore, by using this, the initialization of the FFA technique can outcome in a further uniformly distributed population of Fennec fox. So, it improves the performance and mains to enhance the speed of convergence.

The IFFA model raises a fitness function (FF) to achieve improved classifier performance. It adjusts a positive value to suggest the improved efficiency of candidate outcomes. In this study, the reduction of the classifier ratio of error is dignified as FF.

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{no.of\ misclassified\ instances}{Total\ no.of\ instances} \times 100 \quad (15)$$

4. Experimental Result and Analysis

In this section, the validation of the IFFADL-SLRHSIP approach is performed under the American SL dataset [21].

Fig. 3 signify the SL detection of IFFADL-SLRHSIP method with 70%TRPH and 30%TSPH. The performances presented that the IFFADL-SLRHSIP technique reached effectual detection of all samples. On 70%TRPH, the IFFADL-SLRHSIP attains average $accu_y$ of 99.52%, $prec_n$ of 99.40%, $reca_l$ of 99.47%, and F_{score} of 99.50%. Also, on 30%TSPH, the IFFADL-SLRHSIP technique attains average $accu_y$ of 98.60%, $prec_n$ of 98.59%, $reca_l$ of 98.55%, and F_{score} of 98.51%.

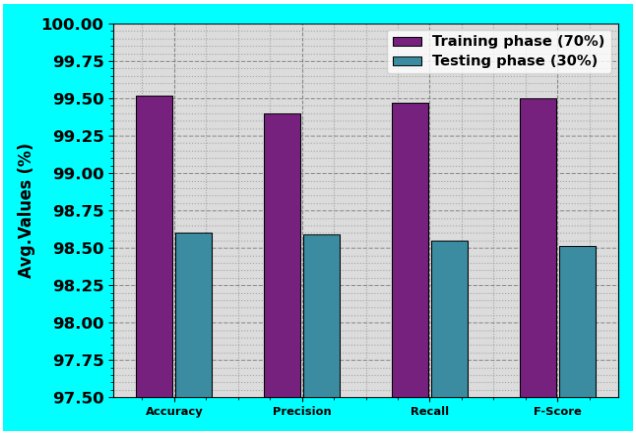

Figure 3. Average of IFFADL-SLRHSIP method under 70% TRPH and 30% TSPH

In Fig. 4, the training accu_y (TRAAY) and validation accu_y (VLAAY) performances of the IFFADL-SLRHSIP model over 0-25 epochs are represented. The output highlighted that the TRAAY and VLAAY outcomes display an increasing trend that reported the competence of the IFFADL-SLRHSIP model with maximum performance beyond several repetitions. Moreover, the TRAAY and VLAAY remain firm above the epochs that indicate least overfitting and reveal the greater performance of the IFFADL-SLRHSIP technique.

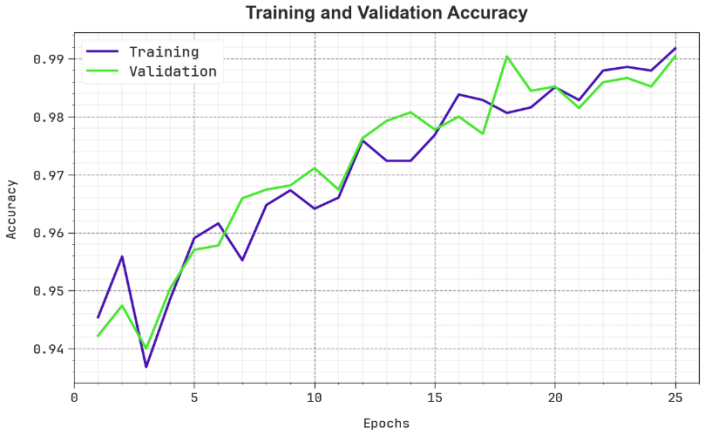

Figure 4. $Accu_y$ Curve of IFFADL-SLRHSIP method

Figure 5. Loss graph of IFFADL-SLRHSIP method

In Fig. 5, the TRA loss (TRALOS) and VLA loss (VLALOS) outcome of the IFFADL-SLRHSIP over 0-25 epochs is presented. It is represented by the TRALOS and VLALOS values represent a trend to reduce, reporting the proficiency of the IFFADL-SLRHSIP technique in weighing an exchange among data fitting and generalization. The frequent reduce in loss values furthermore guarantees the superior solution of the IFFADL-SLRHSIP model and fine-tune the predictive results after a while.

In Fig. 6, the precision-recall (PR) analysis of the IFFADL-SLRHSIP demonstrates its efficiency by computing PR values across all classes. The outcome shows that the IFFADL-SLRHSIP consistently achieves robust PR results through all three classes, maintaining high precision and recall. This steady improvement highlights its efficiency in accurate classification.

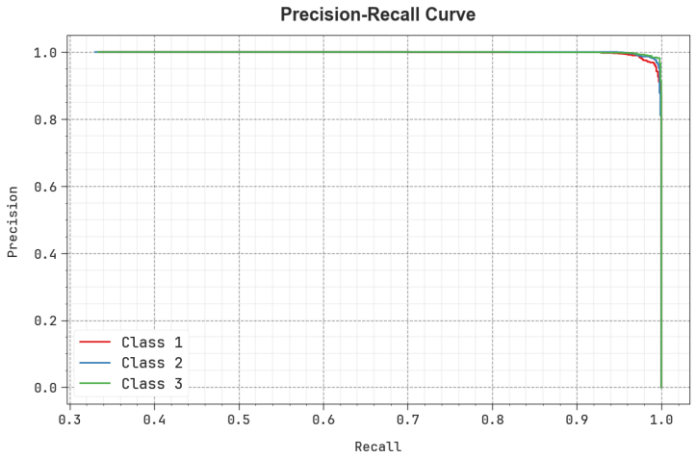

Figure 6. PR curve of IFFADL-SLRHSIP method

In Fig. 7, the ROC outcome of IFFADL-SLRHSIP model is deliberate. The performances indicate that the IFFADL-SLRHSIP technique ranges maximum ROC performances above each class, representing substantial competency of selective classes. This dependable tendency of higher ROC across numerous classes implies the skilled presentation of the IFFADL-SLRHSIP technique for forecasting classes, featuring the robust behaviour of classification method.

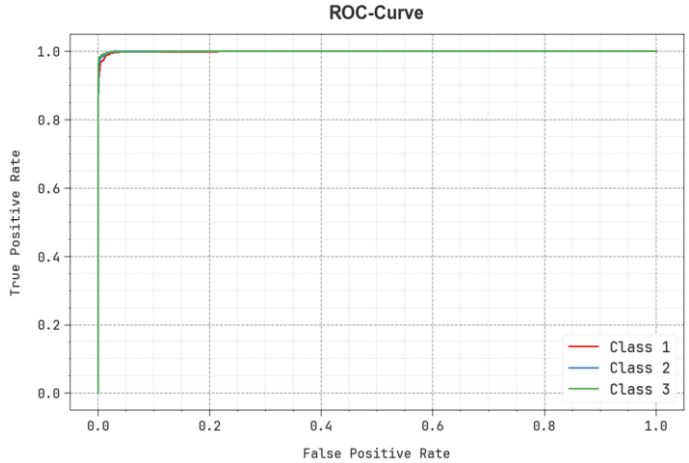

Figure 7. ROC curve of IFFADL-SLRHSIP method

Table 1 and Fig. 8 observe the comparison results of the IFFADL-SLRHSIP model with the existing systems [22]. The performances emphasized that the KNN, SVM, ANN, CNN, and ODTL-SLRC methods reported poorer performance. In the meantime, the EBESO-TLSLR technique has obtained somewhat earlier results. Followed by, the IFFADL-SLRHSIP approach reported enhanced performance with higher $prec_n$, $reca_l$, $accu_y$, and F_{score} of 99.40%, 99.47%, 99.52%, and 99.50%, correspondingly.

Table 1: Comparative study of IFFADL-SLRHSIP approach with existing models

Technique	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}
KNN Algorithm	96.25	96.34	98.97	95.14
SVM Classifier	98.10	96.72	97.55	96.79
ANN Method	98.11	98.15	98.40	98.00
CNN Technique	99.09	96.89	96.04	95.24
ODTL-SLRC	99.03	97.63	95.33	97.95
EBESO-TLSLR	99.42	97.26	95.32	98.02
IFFADL-SLRHSIP	99.52	99.40	99.47	99.50

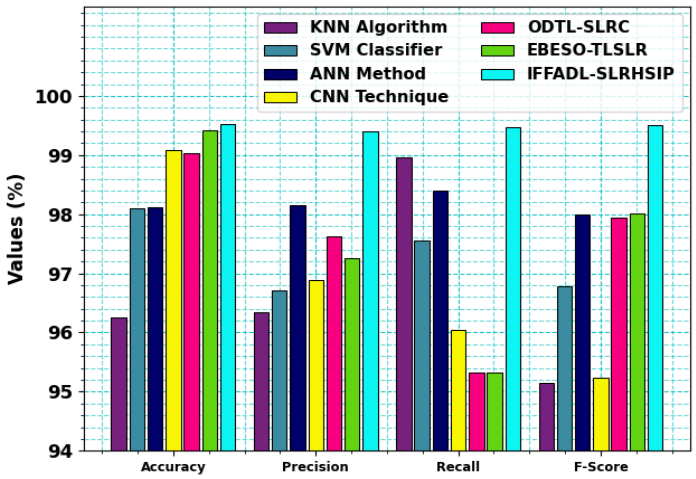

Figure 8. Comparative analysis of IFFADL-SLRHSIP with existing models

In Table 2 and Fig. 9, the comparative solutions of the IFFADL-SLRHSIP approach are illustrated in computational time (CT). The results mention such the IFFADL-SLRHSIP model gets improved performance. According to CT, the IFFADL-SLRHSIP model provides a lower CT of 1.11min whereas the KNN, SVM, ANN, CNN, ODTL-SLRC, and EBESO-TLSLR methods accomplish better CT values of 16.60min, 14.40min, 15.47min, 11.27min, 6.51min, and 3.08min, respectively.

Table 2: CT result of IFFADL-SLRHSIP technique with existing approaches

Technique	CT (min)
KNN Algorithm	16.60
SVM Classifier	14.40
ANN Method	15.47
CNN Technique	11.27
ODTL-SLRC	6.51
EBESO-TLSLR	3.08
IFFADL-SLRHSIP	1.11

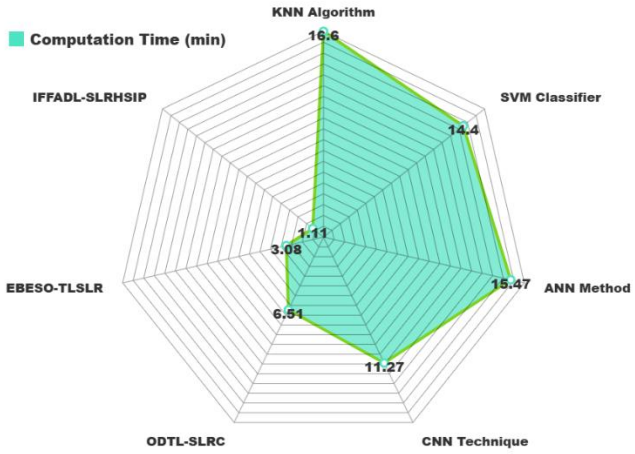

Figure 9. CT outcome of IFFADL-SLRHSIP with existing approaches

5. Conclusion

In this study, the IFFADL-SLRHSIP technique is proposed. The presented IFFADL-SLRHSIP technique main goal is to provide real communication among dumb and deaf persons and normal peoples utilizing CV and AI techniques. To accomplish that, the IFFADL-SLRHSIP technique contains three various stages namely feature extraction, classification, and parameter optimizer. Initially, the IFFADL-SLRHSIP technique takes place in the enhanced SqueezeNet model and is used to collect the intricate forms and shades of SL gestures. For detection of the SL classification process, the RNN model is used. To optimize model performance, the IFFA is applied for parameter tuning, enhancing the model accuracy and efficiency. The experimental outputs of the IFFADL-SLRHSIP approach are authenticated on SL dataset. The results exemplify the greater outcomes of the IFFADL-SLRHSIP approach across various measures.

Data Availability Statement: The data that support the findings of this study are openly available in Kaggle repository at https://www.kaggle.com/datasets/ayuraj/asl-dataset, reference number [21].

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] R. Rastgoo, K. Kiani, and S. Escalera, "Hand sign language recognition using multi-view hand skeleton," *Expert Syst. Appl.*, vol. 150, p. 113336, 2020.

[2] R. Selvanambi, M. Karuppiah, and M. Islabudeen, "Mobile application-based sign language detector for deaf people," in *Designing and Developing Innovative Mobile Applications*, IGI Global, 2023, pp. 329–350.

[3] M. M. Alnfiai, "Deep learning-based sign language recognition for hearing and speaking impaired people," *Intell. Autom. Soft Comput.*, vol. 36, no. 2, pp. 1653–1669, 2023.

[4] Jiang, Y. Wu, and A. Demosthenous, "Hand Gesture Recognition Using Three-Dimensional Electrical Impedance Tomography," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 67, no. 5, pp. 1554–1558, 2020.

[5] Y. Li *et al.*, "Finger gesture recognition using a smartwatch with integrated motion sensors," *Web Intell.*, vol. 16, no. 2, pp. 123–129, 2018.

[6] P. Kumar, R. Saini, P. P. Roy, and D. P. Dogra, “A position and rotation invariant framework for sign language recognition (SLR) using Kinect,” *Multimedia Tools Appl.*, vol. 77, no. 7, pp. 8823–8846, 2018.

[7] Alrowais *et al.*, "Sign language recognition and classification model to enhance quality of disabled people," *Comput. Mater. Continua*, vol. 73, no. 2, pp. 3419–3432, 2022.

[8] H. M. Alshahrani, A. A. Alharthi, M. S. Alzahrani, and R. A. Alhassan, "An Enhanced Framework for Real-Time Sign Language Recognition Using Deep Learning Techniques," *Comput., Mater. & Continua*, vol. 69, no. 1, pp. 123–140, 2023.

[9] S. Hossain *et al.*, "Bengali hand sign gestures recognition using convolutional neural network," in *Proc. 2020 Second Int. Conf. Inventive Res. Comput. Appl. (ICIRCA)*, Coimbatore, India, Jul. 15-17, 2020, pp. 636–641.

[10] J. Shi and Z. Dai, "Research on Gesture Recognition Method Based on EMG Signal and Design of Rehabilitation Training System," in *Proc. IEEE 3rd Adv. Inf. Technol., Electron. Autom. Control Conf.*, Chongqing, China, Oct. 12–14, 2018, pp. 835–838.

[11] D. R. Kothadiya, C. M. Bhatt, H. Kharwa, and F. Albu, "Hybrid InceptionNet based Enhanced Architecture for Isolated Sign Language Recognition," *IEEE Access*, 2024.

[12] M. Buttar *et al.*, "Deep learning in sign language recognition: a hybrid approach for the recognition of static and dynamic signs," *Mathematics*, vol. 11, no. 17, p. 3729, 2023.

[13] Padmaja, D. R. Kumari, M. Anitha, and S. Kasanagottu, "Indian sign language recognition in real time for deaf and dumb," in *AIP Conf. Proc.*, vol. 2971, no. 1, Jun. 2024.

[14] S. Das *et al.*, "A hybrid approach for Bangla sign language recognition using deep transfer learning model with random forest classifier," *Expert Syst. Appl.*, vol. 213, p. 118914, 2023.

[15] N. Palanivel *et al.*, "Visual recognition system for hearing impairment using internet of things," in *AIP Conf. Proc.*, vol. 3075, no. 1, Jul. 2024.

[16] K. K. Podder *et al.*, "Signer-independent Arabic Sign Language recognition system using deep learning model," *Sensors*, vol. 23, no. 16, p. 7156, 2023.

[17] K. Das *et al.*, "Enhancing Communication Accessibility: UrSL-CNN Approach to Urdu Sign Language Translation for Hearing-Impaired Individuals," *CMES-Computer Modeling Eng. Sci.*, vol. 141, no. 1, 2024.

[18] Y. Zhang, J. Wang, J. Chen, D. Shi, and X. Chen, "A Space Non-Cooperative Target Recognition Method for Multi-Satellite Cooperative Observation Systems," *Remote Sens.*, vol. 16, no. 18, p. 3368, 2024.

[19] Basit, J. Mohamad Zain, A. K. Jumaat, N. I. Hamdan, and H. Z. Mojahid, "Predicting COVID-19 trends: a deep dive into time-dependent SIRSD with deep learning technique," *Malaysian J. Comput. (MJoC)*, vol. 9, no. 2, pp. 1955–1978, 2024.

[20] Hu, K. Song, X. Li, and Y. Wang, "DEMFFA: a multi-strategy modified Fennec Fox algorithm with mixed improved differential evolutionary variation strategies," *J. Big Data*, vol. 11, no. 1, p. 69, 2024.

[21] "ASL Dataset". [Online]. Available: https://www.kaggle.com/datasets/ayuraj/asl-dataset.

[22] M. M. Asiri, A. Motwakel, and S. Drar, "Enhanced Bald Eagle Search Optimizer with Transfer Learning-based Sign Language Recognition for Hearing-impaired Persons," *J. Disabil. Res.*, vol. 2, no. 3, pp. 86–93, 2023.