International Journal on Computational Modelling Applications
Vol. 01, Iss. 01, S. No. 004, pp. 45-62, July 2024
ISSN (E): 3048-8516

Artificial Intelligence with Optimization Algorithm Based Robust Sign Gesture Recognition for Visually Impaired People

Ameer N. Onaizah

Beijing Institute Of Technology, School Of Automation, zhongguancun,Beijing,100811, China
ameer@bit.edu.cn*

How to cite this paper: Ameer N. Onaizah, "Artificial Intelligence based Automated Sign Gesture Recognition Solutions for Visually Challenged People", International Journal on Computational Modelling Applications , Vol. no. 01, Iss. No. 01, S No. 004, pp. 45–62, July 2024.

Received: 08/06/2024
Revised: 30/06/2024
Accepted: 10/07/2024
Published: 31/07/2024

Abstract

Sign language is an effective means of communication with visually impaired individuals, as it can be utilized anywhere. So, gestures play a significant part in communication between deaf and dumb people. They are a form of non-verbal data exchange that has garnered significant attention in the development of Human-Computer Interaction (HCI) models, as they permit consumers to state themselves intuitively and naturally in dissimilar contexts. Sign gesture detection is the main need of any HCI application, e.g. gaming, virtual reality, and monitoring methods. At present, Computer vision (CV) and artificial intelligence (AI) have been the efficient areas of dynamic research and growth with the advances in the assistive technology field. In this manuscript, we design and develop an Enhanced Sign Gesture Recognition Model for Disabled People Using Advanced Optimization Models (ESGRM-DPAOM). The proposed ESGRM-DPAOM system is to improve the sign gesture recognition solutions for visually impaired individuals. To accomplish that, the proposed ESGRM-DPAOM model initially applies image preprocessing using wiener filtering (WF) to eliminate the noise in input image data. For the feature extraction process, the SqueezeNet model has been employed and an optimal parameter tuning model utilizes pigeon-inspired optimization (PIO). In addition, the proposed ESGRM-DPAOM models involve the classification process with the aid of the elman neural network (ENN) model. At last, the golden jackal optimizer (GJO) algorithm adjusts the hyperparameter values of the ENN model optimally and outcomes in greater classification performance. Extensive experimentation led to authorizing the performance of the ESGRM-DPAOM approach. The simulation outcomes specified that the ESGRM-DPAOM system emphasized advancement over other existing methods.

Keywords

Artificial Intelligence; Sign Gesture Recognition; Visually Challenged People; Image Preprocessing; Golden Jackal Optimization

1. Introduction

Blind and visually impaired individuals have significant features of our community; a substantial growth in eye-related illnesses and a decrease in the human vision of visually impaired individuals became a helpful dimension for the society, medical institutions, and government agencies to oppose post-impairment challenge [1]. Vision loss undermines the capability of

blind people to accomplish multiple actions in everyday life. This growth in the blind individual population will affect the living standards and cause social differences in the future, if not considered sufficiently [2]. These concerns have driven examiners to explore novel avenues of study through multiple disciplines like assistive cognitive psychology, technologies, sensory processing, computer vision (CV), accessibility, and rehabilitation of inclusive human-computer interaction (HCI) [3]. These Aided technologies enable blind individual data admittance, stimulate safety, and assist their flexibility together with enhanced living standards, having straightforward effects on social inclusion [4].

Hand gestures play a significant role in Sign language (SL). Hand gesture recognition (HGR) is a major need for any HCI application, for example, gaming, monitoring systems, and virtual reality [5]. Indian sign language (ISL) performs a bridging between deaf and dumb people. ISL is widely different from other SLs, as it utilizes either hand. It has its phonological framework, grammar, and syntax [6]. This survey demonstrated that not much study is done on ISL identification compared with other SLs, for example, Brazilian sign language American Sign Language (ASL), and more. Each SL is generally created from non- and manual gestures [7]. The manual sign gestures contain hand postures, hand orientation, and hand movements. The non-manual sign gestures contain eye gaze, lip movements, and facial expressions. The hand gesture detection model is categorized into dual classes like vision-based and data-glove approaches [8]. Usually, hand gestures are vital and rudimentary devices in communication for sign language recognition (SLR). Absolutely, HGR is the initial stage to enhance the SLR method. Multiple investigations require acknowledging and recognizing the usage of hand gestures [9]. Artificial intelligence (AI) has the possibility to develop availability in surroundings, including the home. When it comes to those with impairments, a virtual assistant could make a massive divergence in their living standards [10]. Due to the developments in AI, people with disabilities have been admitted to a more extensive setting.

This manuscript designs and develops an Enhanced Sign Gesture Recognition Model for Disabled People Using Advanced Optimization Models (ESGRM-DPAOM). To accomplish that, the proposed ESGRM-DPAOM model initially applies image preprocessing using Wiener filtering (WF) to eliminate the noise in input image data. For the feature extraction process, the SqueezeNet has been employed and an optimal parameter tuning model utilizes pigeon-inspired optimization (PIO). In addition, the proposed ESGRM-DPAOM models involve a classification process with the aid of the Elman neural network (ENN) model. At last, the golden jackal optimizer (GJO) algorithm adjusts the hyperparameter values of the ENN model optimally and outcomes in greater classification performance. Extensive experimentation led to authorizing the performance of the ESGRM-DPAOM model.

2. Literature Review

Chang et al. [11] studied the developed system's model in dual stages. The initial stage is Region of Interest (ROI), depending on the color space segmentation model, with a preset range of colors that will extract ROI pixels from the circumstances. Alashhab et al. [12] introduce an interactive model for mobile gadgets deliberated by hand gestures that permit the consumer to regulate the gadgets and utilize multiple assistive devices by creating dynamic and static hand gestures. This method is dependent on a multi-head NN that primarily classifies and detects the gestures, and then, based on the identified gesture, accomplishes a next phase that implements the equivalent activities. This structure enhances the source needed to implement diverse tasks, it captures the benefit of data attained from a primary backbone to achieve dissimilar methods in the next phase. Lindner et al. [13] developed the design of an independent humanoid robot intended to enrich and improve customer service in shops. This control method allows independent navigation in either uncharted or known settings, with a special aim on different, cluttered, and crowded areas.

Gupta et al. [14] introduce a method for SL communication between a user and a computer in complicated settings. Still complex surroundings, the recommended method demonstrated greater precision and effectually identified motions in lower-resolution picture mode. Amangeldy et al. [15] introduced a model for automatically gathering the spatiotemporal gesture aspects by determining the coordinates of the initial field of hand and pose, normalizing them, and creating an optimum multi-layer perceptron for multi-class classification. By analyzing and extracting spatiotemporal data, the presented approach makes it probable to recognize not only static attributes but also the spatial and dynamic gesture features that resulted in to rise in the gesture recognition precision.

De Oliveira et al. [16] develop an examination of model attributes to assist design of more reasonable automated mobile home applications for visually impaired consumers. Regardless of the restrictions in the proof of concept prototype utilized in this work, consumers originate the application highlighted and promising the necessity for enhancements in the domain, emphasizing the chances to offer automatic control of appliances to improve autonomous living settings. Sruthi and Lijiya [17] introduce a dynamic ISL detection method without complex sensors or expensive gadgets to sense hand movements.

3. Proposed Methodology

In this manuscript, we design and develop an ESGRM-DPAOM algorithm. The main objective of the proposed SGRM-DPAOM technique is to improve the sign gesture recognition solutions for visually impaired individuals. To accomplish that, the proposed ESGRM-DPAOM model has distinct stages involved in image preprocessing, feature extractor, classification, and parameter optimizer approaches. Fig. 1 exemplifies the entire process of the ESGRM-DPAOM algorithm.

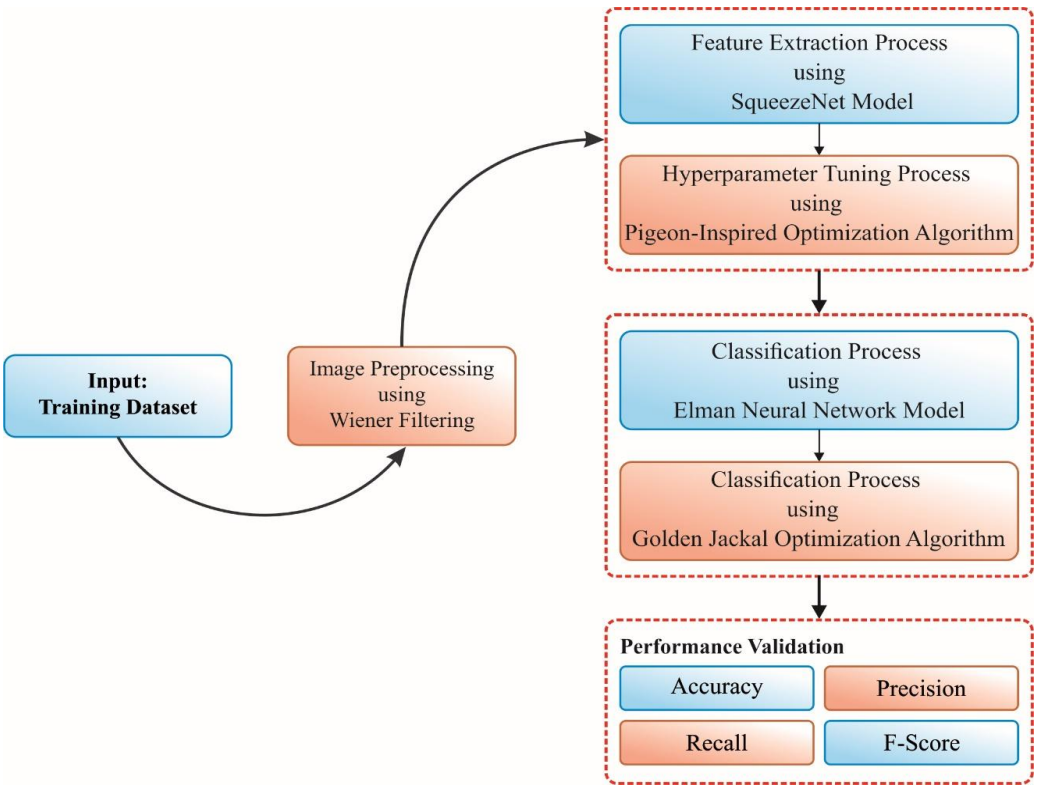

Fig. 1. Overall flow of ESGRM-DPAOM technique

3.1. Image Preprocessing

Initially, the proposed ESGRM-DPAOM model applies image preprocessing using WF to eliminate the noise in input image data. WF is a noise reduction model, employed in sign gesture recognition to enhance the simplicity of input signals by minimalizing the mean square error among the filtered and original signals [18]. It efficiently upholds gesture features while overpowering redundant noise, enhancing detection accuracy.

3.2. SqueezeNet Feature Extractor

For the feature extraction process, the SqueezeNet model has been employed. SqueezeNet was constructed according to the concept of CNN, by certain changes [19]. For SqueezeNet, the components in the pooling and convolution layers are frequently substituted with the blocks of the fire module. A fire module contains expand and squeeze layers. The activation function of $ReLU$ is additionally applied amongst these layers to improve the depth of the network. The squeeze layer comprises simply 1x1 convolutional filters, whereas the expand layer contains a connection of 1x1 and 3x3 convolution filters. SqueezeNet further stops utilizing fully connected (FC) layers. Numerous changes were generated to the original architecture of SqueezeNet, presented regarding the location of the SqueezeNet layers. The terms $fire$, $conv$and $pool$ individually characterize the fire, convolution, and pooling module layers. It is identified that the SqueezeNet was constructed utilizing 2 $conv$ and layers, 3 maximal $pool$ layers, and 8 $fire$ module layers. In general, the final layer with learnable weights is the FC layer. For SqueezeNet, the final learnable layer corresponds to the last $conv$ layer incorporated with the global average $pool$ and Softmax activation function. These configurations permit SqueezeNet to be extremely precise with smaller-size models. It creates SqueezeNet a choice in biometric systems, due to the short-term usage of time for recognition and feature extraction methods.

Next, an optimal parameter tuning model utilizes the PIO algorithm. The PIO determines a mathematical approach by mimicking the pigeon's homing behavior and is normally applied to solving optimizer problems [20]. The iron crystals on the pigeon's beak could identify the power of magnetic fields to assist pigeons in differentiating their path. Back in the olden era of Rome, people found that pigeons might recognize the path and come back to the nest and have been applied as a communication device. In coming back to its destination, the pigeon makes use of the magnetic field to sense their path in the initial phase. The PIO model suggests a map and compass operator according to the sun and geomagnetic field, presented landmark operator based on the landmark.

Compass operator

It is constructed according to the sun and the geomagnetic field. During this study, V_i^T and X_i^T characterize the velocity and location vector of the $i\,th$ pigeon, correspondingly, and their updated equation is

$$V_i^T = V_i^{T-1} \cdot e^{-RT} + rand.\left(X_{gb} - X_i^{T-1}\right) \tag{1}$$

$$X_i^T = X_i^{T-1} + V_i^T \tag{2}$$

Whereas T signifies the iteration counts of the PIO model, X_{gb} represents the location vector of the pigeons in the group demonstrating the optimum solution in that iteration, and R refers to the compass operator.

As discussed earlier, it is discovered that the velocity of consistent pigeon of preceding group and the location of the optimum result of the group in this present iteration mutually define the vector velocity of the $i\,th$ pigeon. Additionally, the location of the $i\,th$ pigeon is established by the preceding location and the present velocity. Every pigeon in the group fine-tuned their flight path toward the optimum solution based on Eq. (1) and modified their location based on Eq. (2).

Landmark operator

It is constructed according to landmarks. In the flight of the pigeon, if it is near the destination, the pigeon would establish its flight route according to the flight conditions of nearby pigeons and landmarks. N_u represents the quantity of one-half pigeons in all iterations, X_C^T refers to the central location of the group in the T^{th} generation, and fitness(x) can be described as the pigeon's mass. Assume that each pigeon flies through a straight line among their location and the destination:

$$N_u^T = \frac{N_u^{T-1}}{2} \tag{3}$$

$$X_c^T = \frac{\sum X_i^T \cdot fitness(X_i^T)}{N_u \cdot \sum f\,itness(X_i^T)} \tag{4}$$

$$X_i^T = X_i^{T-1} + rand.(X_c^T - X_c^{T-1}) \tag{5}$$

In the iteration, the destination of each pigeon is the central location in that iteration, such that half of pigeons, which is distant from the destination would establish their velocity and location vector by mentioning the pigeon's flight condition about the destination, whereas another half can achieve the destination using the highest speed.

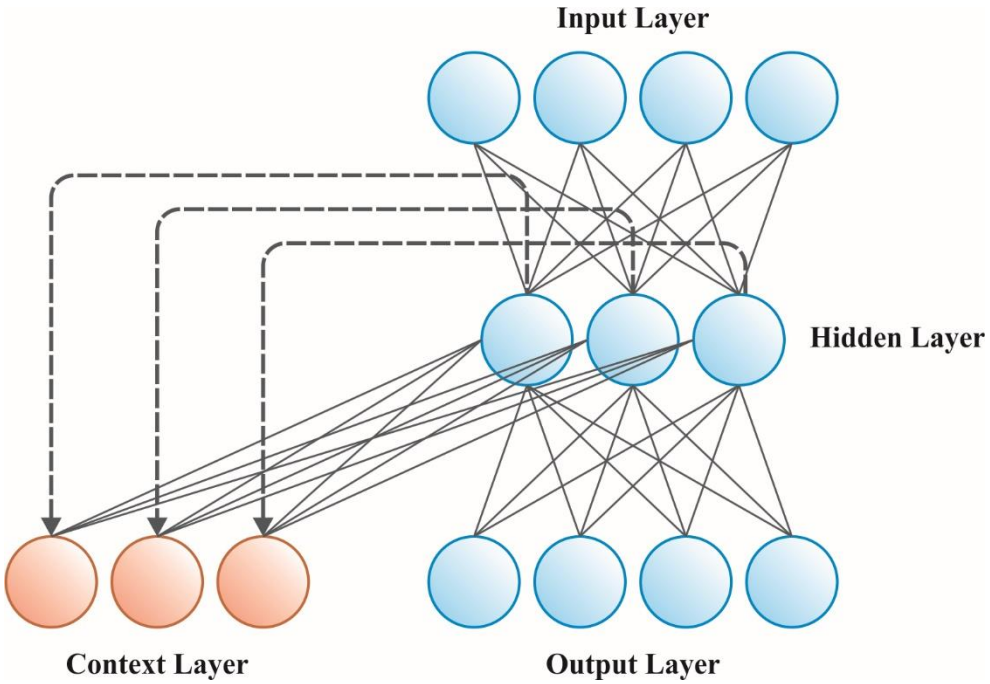

Fig. 2. Architecture of ENN

3.3. Classification using ENN Model

In addition, the proposed ESGRM-DPAOM models involve classification process with the aid of the ENN model. ENN is a recursive system highlighted by the internal self-referential layer [21]. It comprises 4 modules: context layer (CL), input layer, output layer, and hidden layer (HL). The CL is tailored for storing or memorizing the output values previously the HL. The HL's outcome is transmitted to the input layer over the CL.

The ENN's non-linear expression is as demonstrated:

$$y(k) = g\big(w^3 x(k)\big) \tag{6}$$

$$x(k) = f\big(w^1 x_c(k) + w^2 u(k-1)\big) \tag{7}$$

$$x_c(k) = x(k-1) \tag{8}$$

Whereas k denotes the ENN training number; y signifies the vector of n-dimensional output; x embodies the vector of HL neuron output; x_c signifies the vector of the feedback state; u refers to the input vector; g characterizes the output

neuron's transfer function; w^1, w^2 and w^3 denotes connection weighted matrices, and f characterizes the HL neuron's transfer function. Fig. 2 depicts the structure of ENN.

3.4. Hyperparameter Tuning Process

At last, the GJO algorithm adjusts the hyperparameter values of the ENN model optimally and outcomes in greater classification performance. The elementary GJO stimulated by the golden jackal's cooperative attacking behavior is a new optimizer model [22]. In GJO, all GL signifies the search agent or candidate solution. The cooperative attacking behavior of the golden jackal pair.

Search space expression

During GJO, initialization of the population of arbitrary prey to attain the candidate solution is distributed uniformly in the searching space. The primary solution is calculated as:

$$Y_0 = Y_{min} + rand(Y_{\max} - Y_{min}) \tag{9}$$

Whereas Y_0 states the locations of the first population of GJ, $rand$ states a uniform randomly formed vector in [0, 1], $Y_{\min}$ and $Y_{\max}$ indicate the lower and upper limitations of the solution.

The primary and next fitting is the jackal pair, the first individual matrix is calculated as:

$$Prey = \begin{bmatrix} Y_{1,1} & Y_{1,2} & L & Y_{1,d} \\ Y_{2,1} & Y_{2,2} & L & Y_{2,d} \\ M & M & M & M \\ Y_{n,1} & Y_{n,2} & L & Y_{n,d} \end{bmatrix} \tag{10}$$

Whereas $Y_{i,j}$ signifies the $j\,th$ size of $i\,th$ prey, n indicates the complete prey quantity, d demonstrates the problem variables. The matrix is calculated as:

$$F_{OA} = \begin{bmatrix} f(Y_{1,1}; Y_{1,2}; Y_{1,d}) \\ f(Y_{2,1}; Y_{2,2}; Y_{2,d}) \\ M \\ f(Y_{n,1}; Y_{n,2}; Y_{n,d}) \end{bmatrix} \tag{11}$$

Here F_{OA} states a matrix, which comprises the prey's fitness values, f displays the fitness function (FF). The primary and next fittest is taken as a male and female jackal. The pair of jackals recognizes the commensurate individual position.

Exploration stage or prey searching

The GJs might forecast and seize the prey depending on their individual assaulting qualities, however, the prey on occasion quickly escapes and evades the look of jackals. As a result, the female jackals emulate the male jackals to attain searching and waiting for other prey within the searching area. The locations are calculated as:

$$Y_1(t) = Y_M(t) - E \cdot |Y_M(t) - rl \cdot Prey(t)| \tag{12}$$

$$Y_2(t) = Y_{FM}(t) - E \cdot |Y_{FM}(t) - rl \cdot Prey(t)| \tag{13}$$

Here t states the present iteration, $Prey\,(t)$ indicates the location vector, $Y_{FM}(t)$ and $Y_M(t)$ demonstrates the present places of the female and male jackals individually. $Y_2(t)$ and $Y_1(t)$ shows the refreshed places of the female and male jackals.

The prey's evading energy E is measured as:

$$E = E_1 * E_0 \tag{14}$$

Whereas E_1 exhibits the prey's reducing energy, E_0 reveals the original condition of the energy.

$$E_0 = 2 * r - 1 \tag{15}$$

Here r indicates random values in [0,1].

$$E_1 = c_1 * \left(1 - \left(\frac{t}{T}\right)\right) \tag{16}$$

Now T demonstrates the maximal iteration, c_1 indicates a constant valued at 1.5, E_1 linearly drops from 1.5 to 0 depending on the iteration.

The rl states an arbitrary vector in accordance with the Lévy distribution that is measured as:

$$rl = 0.05 * LF(y) \tag{17}$$

The LF expresses the FF of the levy flight that can be calculated as:

$$LF(y) = 0.01 \times (\mu \times \sigma)/(|v^{(1/\beta)}|); \sigma = \left[\frac{\Gamma(1+\beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left(\frac{1+\beta}{2}\right) \times \beta \times \left(2^{\frac{\beta-1}{2}}\right)}\right]^{1/\beta} \tag{18}$$

Whereas μ and v exhibit the random values in $(0,1)$, β shows a constant value of 1.5. The refreshed location of the GJs is measured as:

$$Y(t+1) = \frac{Y_1(t) + Y_2(t)}{2} \tag{19}$$

Exploitation step or pouncing and enclosing the prey

The prey's evading energy will quickly diminish after it is infected by the jackal pair, the GJs suddenly encircle and seize the prey. The locations are calculated as:

$$Y_1(t) = Y_M(t) - E \cdot |rl \cdot Y_M(t) - Prey(t)| \tag{20}$$

$$Y_2(t) = Y_{FM}(t) - E \cdot |rl \cdot Y_{FM}(t) - Prey(t)| \tag{21}$$

Whereas t indicates the present iteration, $Prey\,(t)$ shows the location vector, $Y_{FM}(t)$ and $Y_M(t)$ demonstrate the existing places of the female and male jackals individually. $Y_2(t)$ and $Y_1(t)$ exhibits the refreshed places of the female and male jackals. Several parameters were given in the preceding sections, namely E and rl.

The fitness choice is the substantial factor, which effects the outcome of the GJO approach. The parameter range procedure contains the solution-encoded system to assess the effectiveness of the candidate results. Here, the GJO system reflects accuracy as the main measure to project the fitness function. Its mathematical formulation is given below:

$$Fitness = \max(P) \tag{22}$$

$$P = \frac{TP}{TP + FP} \tag{23}$$

While TP defines the positive value of true and FP denotes the positive value of false.

4. Performance Validation

The performance validation of ESGRM-DPAOM model is inspected on American sign language recognition dataset [23] under two aspects such as alphabets (a-z) and number (0-9). Fig. 3 represents the sample images.

Fig. 4 states the confusion matrix generated by the ESGRM-DPAOM method below 70%TRPH on the alphabet recognition (ASR) process. The performances indicate that the ESGRM-DPAOM algorithm has effectual detection and identification of all the class labels specifically.

Fig. 3. Sample Images

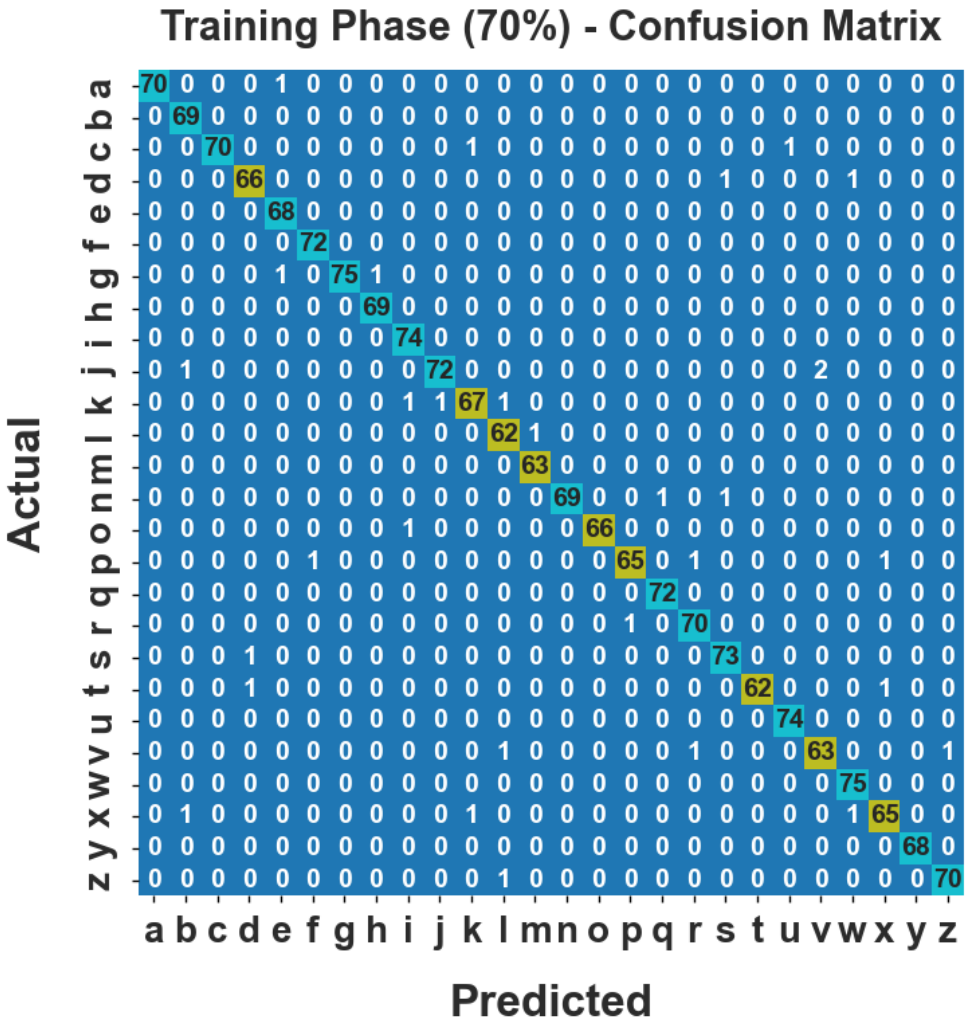

Fig. 4. Confusion matrix of ESGRM-DPAOM under 70% TRPH on ASR

Table 1 illustrates the classifier result of the ESGRM-DPAOM system using 70%TRPH on ASR. The performances imply that the proposed ESGRM-DPAOM method has attained efficacious performance in every class.

Table 1 Classifier outcome of ESGRM-DPAOM model with 70%TRPH on ASR

TRPH (70%)									
Class	**Acc.**	**PPV**	**DR**	**F-Score**	**Class**	**Acc.**	**PPV**	**DR**	**F-Score**
a	99.95	100.00	98.59	99.29	n	99.89	100.00	97.18	98.57
b	99.89	97.18	100.00	98.57	o	99.95	100.00	98.51	99.25
c	99.89	100.00	97.22	98.59	p	99.78	98.48	95.59	97.01
d	99.78	97.06	97.06	97.06	q	99.95	98.63	100.00	99.31
e	99.89	97.14	100.00	98.55	r	99.84	97.22	98.59	97.90
f	99.95	98.63	100.00	99.31	s	99.84	97.33	98.65	97.99
g	99.89	100.00	97.40	98.68	t	99.89	100.00	96.88	98.41
h	99.95	98.57	100.00	99.28	u	99.95	98.67	100.00	99.33
i	99.89	97.37	100.00	98.67	v	99.73	96.92	95.45	96.18
j	99.78	98.63	96.00	97.30	w	99.89	97.40	100.00	98.68

k	99.73	97.10	95.71	96.40	x	99.73	97.01	95.59	96.30
l	99.78	95.38	98.41	96.88	y	100.00	100.00	100.00	100.00
m	99.95	98.44	100.00	99.21	z	99.89	98.59	98.59	98.59

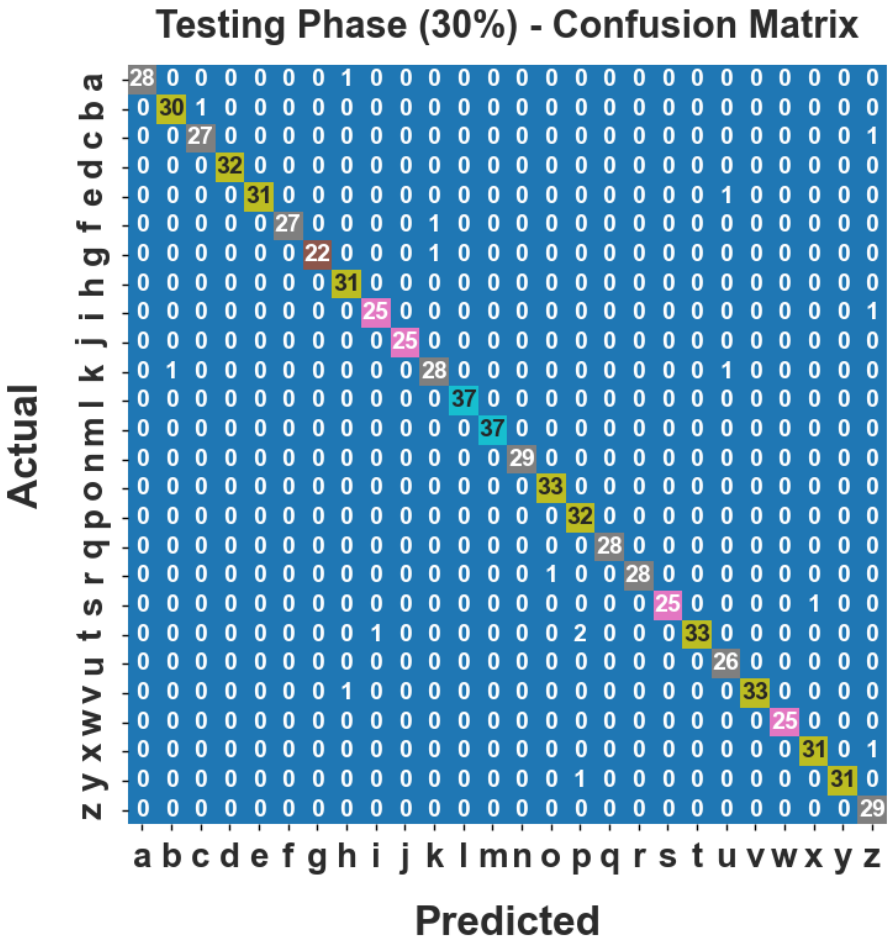

Fig. 5. Confusion matrix of ESGRM-DPAOM under 30% TSPH

Fig. 5 determines the confusion matrix produced by the ESGRM-DPAOM method below 30%TSPH on ASR. The performances suggest that the ESGRM-DPAOM approach has accurate detection and identification of all the class labels accurately.

Table 2 demonstrates the classifier result of the ESGRM-DPAOM technique using 30%TSPH on ASR. The performances signify that the proposed ESGRM-DPAOM algorithm has reached efficient performance in every class.

Table 2 Classifier outcome of ESGRM-DPAOM model with 30%TSPH on ASR

TSPH (30%)									
Class	**Acc.**	**PPV**	**DR**	**F-Score**	**Class**	**Acc.**	**PPV**	**DR**	**F-Score**
a	99.87	100.00	96.55	98.25	n	100.00	100.00	100.00	100.00
b	99.74	96.77	96.77	96.77	o	99.87	97.06	100.00	98.51
c	99.74	96.43	96.43	96.43	p	99.62	91.43	100.00	95.52
d	100.00	100.00	100.00	100.00	q	100.00	100.00	100.00	100.00
e	99.87	100.00	96.88	98.41	r	99.87	100.00	96.55	98.25

f	99.87	100.00	96.43	98.18	s	99.87	100.00	96.15	98.04
g	99.87	100.00	95.65	97.78	t	99.62	100.00	91.67	95.65
h	99.74	93.94	100.00	96.88	u	99.74	92.86	100.00	96.30
i	99.74	96.15	96.15	96.15	v	99.87	100.00	97.06	98.51
j	100.00	100.00	100.00	100.00	w	100.00	100.00	100.00	100.00
k	99.49	93.33	93.33	93.33	x	99.74	96.88	96.88	96.88
l	100.00	100.00	100.00	100.00	y	99.87	100.00	96.88	98.41
m	100.00	100.00	100.00	100.00	z	99.62	90.62	100.00	95.08

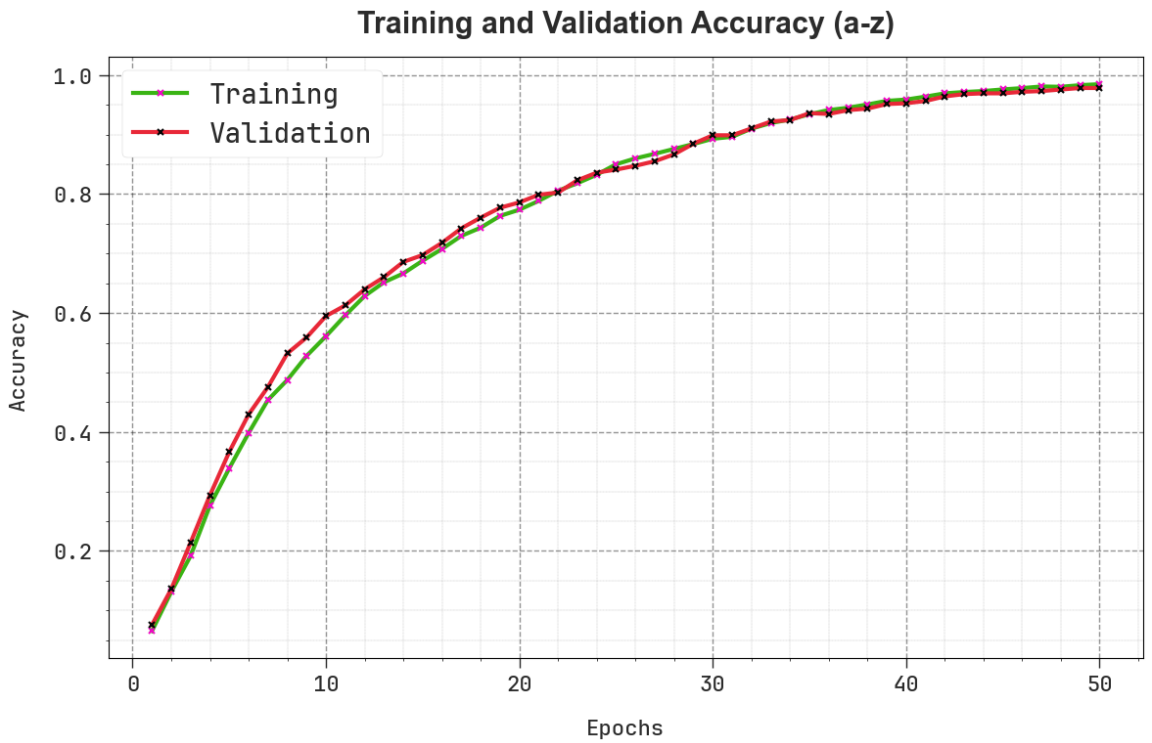

Fig. 6. $Accu_y$ the curve of ESGRM-DPAOM technique under ASR

In Fig. 6, the training accuracy (TRAC) and validation accuracy (VDAC) performances of the ESGRM-DPAOM technique below ASR are settled. The performance underscored that the values of TRAC and VDAC display an increasing trend indicating the proficiency of the ESGRM-DPAOM model with modified performance through multiple repetitions.

In Fig. 7, the training loss (TRLS) and validation loss (VDLS) graph of the ESGRM-DPAOM approach below ASR is depicted. It is signified that the values of TRLS and VDLS reveal a diminishing tendency, notifying the capacity of the ESGRM-DPAOM method in equalizing a tradeoff among data fitting and generalization.

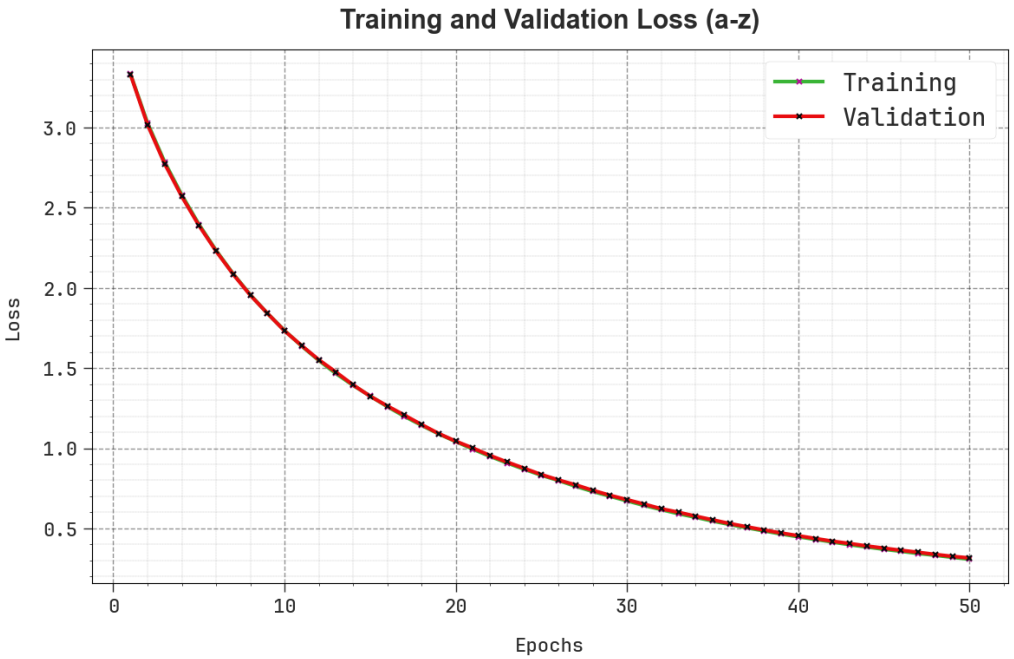

Fig. 7. Loss curve of ESGRM-DPAOM technique under ASR

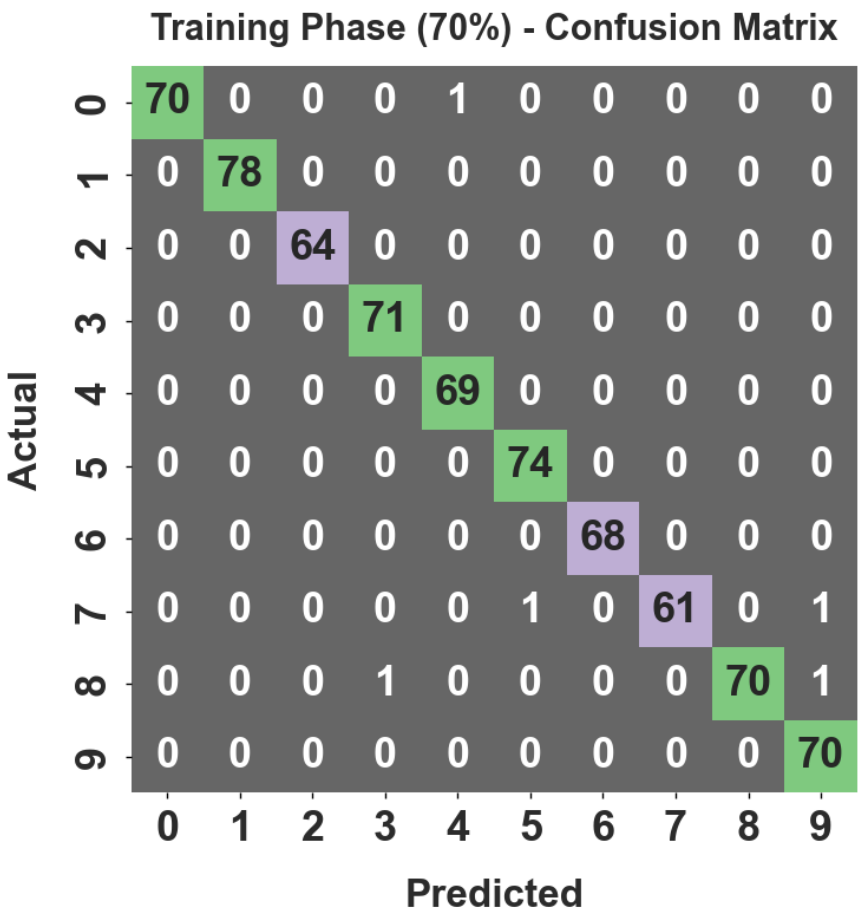

Fig. 8. Confusion matrix of ESGRM-DPAOM under 70% TRPH on NSR

Fig. 8 verifies the confusion matrix generated by the ESGRM-DPAOM algorithm below 70%TRPH on the numerical recognition (NSR) process. The performances indicate that the ESGRM-DPAOM approach has effectual detection and identification of all the class labels specifically.

Table 3 exemplifies the classifier result of the ESGRM-DPAOM system using 70%TRPH on NSR. The performances denote that the proposed ESGRM-DPAOM model has accomplished efficient performance in each class.

Table 3 Classifier outcome of ESGRM-DPAOM model with 70%TRPH on NSR

Training Phase (70%)				
Class	**Acc.**	**PPV**	**DR**	**F-Score**
0	99.86	100.00	98.59	99.29
1	100.00	100.00	100.00	100.00
2	100.00	100.00	100.00	100.00
3	99.86	98.61	100.00	99.30
4	99.86	98.57	100.00	99.28
5	99.86	98.67	100.00	99.33
6	100.00	100.00	100.00	100.00
7	99.71	100.00	96.83	98.39
8	99.71	100.00	97.22	98.59
9	99.71	97.22	100.00	98.59

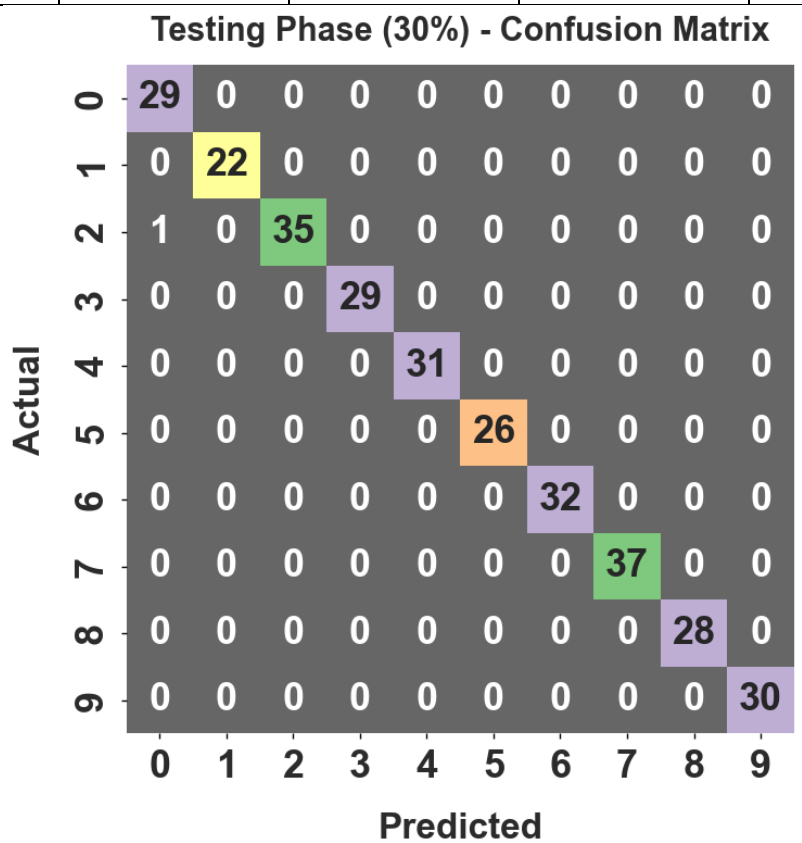

Fig. 9. Confusion matrix of ESGRM-DPAOM under 30% TSPH on NSR

Fig. 9 demonstrates the confusion matrix produced by the ESGRM-DPAOM approach below 30%TSPH on NSR. The performances suggest that the ESGRM-DPAOM algorithm has effectual detection and identification of all the class labels accurately.

Table 4 signifies the classifier result of the ESGRM-DPAOM method using 30%TSPH on NSR. The performances imply that the proposed ESGRM-DPAOM technique has gained efficient performance in every class.

Table 4 Classifier outcome of ESGRM-DPAOM model with 30%TSPH on NSR

TSPH (30%)				
Class	**Acc.**	**PPV**	**DR**	**F-Score**
0	99.67	96.67	100.00	98.31
1	100.00	100.00	100.00	100.00
2	99.67	100.00	97.22	98.59
3	100.00	100.00	100.00	100.00
4	100.00	100.00	100.00	100.00
5	100.00	100.00	100.00	100.00
6	100.00	100.00	100.00	100.00
7	100.00	100.00	100.00	100.00
8	100.00	100.00	100.00	100.00
9	100.00	100.00	100.00	100.00

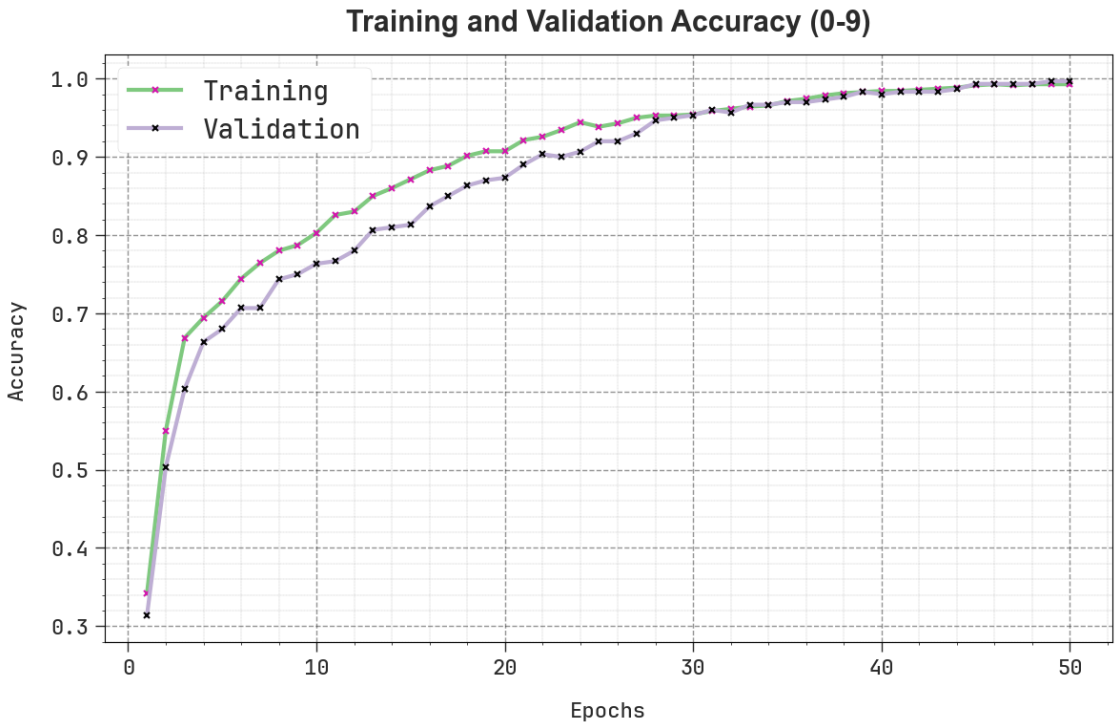

Fig. 10. $Accu_y$ curve of ESGRM-DPAOM technique under NSR

In Fig. 10, the TRAC and VDAC performances of the ESGRM-DPAOM technique below NSR are constituted. The performance underscored that the values of TRAC and VDAC represent an increasing trend notifying the competency of the ESGRM-DPAOM model with modified performance across various repetitions.

In Fig. 11, the TRLS and VDLS graph of the ESGRM-DPAOM algorithm below NSR is demonstrated. It is implied that the values of TRLS and VDLS show a diminishing trend, notifying the ability of the ESGRM-DPAOM system in corresponding a tradeoff among data fitting and generalization.

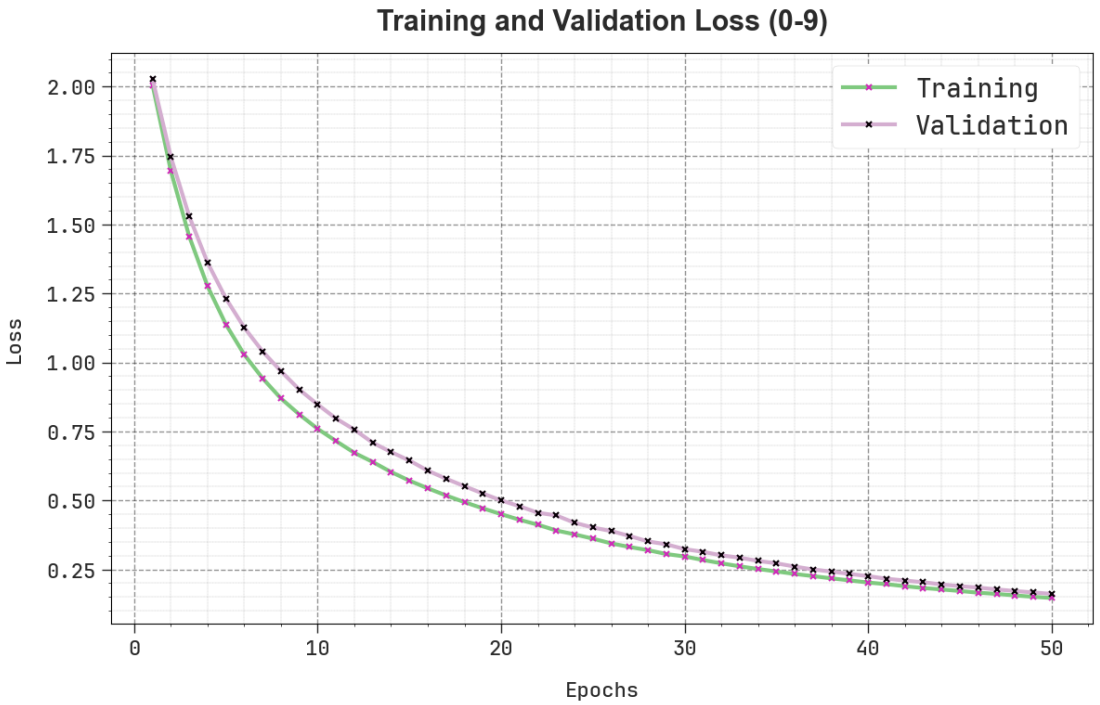

Fig. 11. Loss curve of ESGRM-DPAOM technique under NSR

Table 5 depicts the sign language recognition performance of the ESGRM-DPAOM approach on ASR (a-z) and NSR (0-9). The performances denote that the ESGRM-DPAOM method has reached efficient performance. According to ASR, the proposed ESGRM-DPAOM system has attained a typical *accuy* of 99.85%, PPV of 98.10%, DR of 98.06%, and *F-score* of 98.05%. Moreover, according to NSR, the ESGRM-DPAOM technique has gained an average *accuy* of 99.90%, PPV of 99.49%, DR of 99.49%, and *F-score* of 99.49%.

Table 5 Overall average sign language recognition outcome of the ESGRM-DPAOM model

	Acc.	**PPV**	**DR**	**F-Score**
(a-z)				
Training Phase	99.87	98.30	98.29	98.28
Testing Phase	99.83	97.90	97.82	97.82
Average	**99.85**	**98.10**	**98.06**	**98.05**
(0-9)				
Training Phase	99.86	99.31	99.26	99.28
Testing Phase	99.93	99.67	99.72	99.69
Average	**99.90**	**99.49**	**99.49**	**99.49**

Table 6 and Fig. 12 provide a comparative analysis of the ESGRM-DPAOM model with existing models [24, 25]. The values in the table indicate that the proposed ESGRM-DPAOM system has got higher accuracy of 99.90%, whereas the existing models Hypertuned DCNN, CNN, ANN, KNN, American SLR-LMS, Arabic HGR-LMC, SLSHG-LMC-RNN, and SVM techniques have obtained lesser accuracy of 99.75%, 99.68%, 98.05%, 96.02%, 79.88%, 91.36%, 96.46%, and 97.96%, respectively.

Table 6 Accuracy outcome of the ESGRM-DPAOM model with existing methodologies

Methodology	Accuracy (%)
ESGRM-DPAOM	99.90
Hypertuned DCNN	99.75
Convolutional Neural Network	99.68
Artificial Neural Networks	98.05
KNN Algorithm	96.02
American SLR-LMS	79.88
Arabic HGR-LMC	91.36
SLSHG-LMC-RNN	96.46
Support Vector Machine	97.96

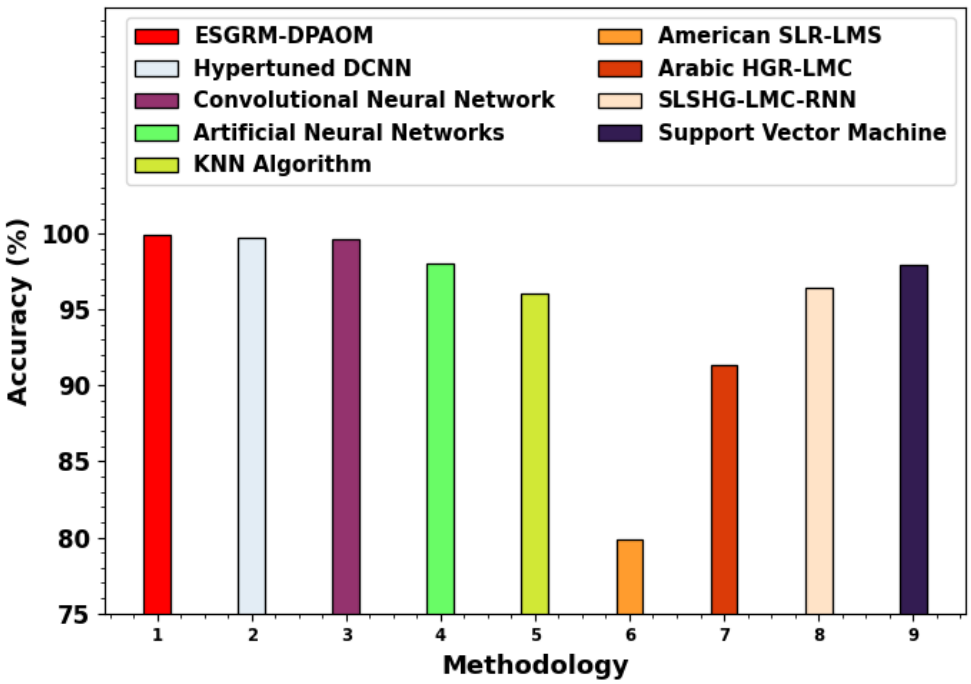

Fig. 12. Accuracy outcome of ESGRM-DPAOM model with existing methodologies

5. Conclusion

In this manuscript, we design and develop an ESGRM-DPAOM algorithm. The proposed ESGRM-DPAOM system is to improve the sign gesture recognition solutions for visually impaired individuals. To accomplish that, the proposed ESGRM-DPAOM model initially applies image preprocessing using WF to eliminate the noise in input image data. For the feature extraction process, the SqueezeNet has been employed and an optimal parameter tuning model utilizes the PIO algorithm. In addition, the proposed ESGRM-DPAOM models involve a classification process with the aid of the ENN model. At last, the GJO algorithm adjusts the hyperparameter values of the ENN model optimally and outcomes in greater classification performance. Extensive experimentation led to authorizing the performance of the ESGRM-DPAOM algorithm. The simulation outcomes specified that the ESGRM-DPAOM technique emphasized advancement over other existing methods.

References

[1] Sharma, S. and Singh, S., 2021. Vision-based hand gesture recognition using deep learning for the interpretation of sign language. *Expert Systems with Applications*, *182*, p.115657.

[2] Mujahid, A., Awan, M.J., Yasin, A., Mohammed, M.A., Damaševičius, R., Maskeliūnas, R. and Abdulkareem, K.H., 2021. Real-time hand gesture recognition based on deep learning YOLOv3 model. *Applied Sciences*, *11*(9), p.4164.

[3] Zheng, Z., Wang, Q., Yang, D., Wang, Q., Huang, W. and Xu, Y., 2022. L-sign: Large-vocabulary sign gestures recognition system. *IEEE Transactions on Human-Machine Systems*, *52*(2), pp.290-301.

[4] Juneja, S., Juneja, A., Dhiman, G., Jain, S., Dhankhar, A. and Kautish, S., 2021. Computer vision-enabled character recognition of hand gestures for patients with hearing and speaking disability. *Mobile Information Systems*, *2021*(1), p.4912486.

[5] Mohamed, N., Mustafa, M.B. and Jomhari, N., 2021. A review of the hand gesture recognition system: Current progress and future directions. *IEEE access*, *9*, pp.157422-157436.

[6] Gangrade, J. and Bharti, J., 2023. Vision-based hand gesture recognition for Indian sign language using convolution neural network. *IETE Journal of Research*, *69*(2), pp.723-732.

[7] Sadeddine, K., Chelali, F.Z., Djeradi, R., Djeradi, A. and Benabderrahmane, S., 2021. Recognition of user-dependent and independent static hand gestures: Application to sign language. *Journal of Visual Communication and Image Representation*, *79*, p.103193.

[8] Tasmere, D., Ahmed, B. and Das, S.R., 2021. Real time hand gesture recognition in depth image using cnn. *International Journal of Computer Applications*, *174*(16), pp.28-32.'

[9] Padmanandam, K., Rajesh, M.V., Upadhyaya, A.N., Chandrashekar, B. and Sah, S., 2022. Artificial intelligence biosensing system on hand gesture recognition for the hearing impaired. *International Journal of Operations Research and Information Systems (IJORIS)*, *13*(2), pp.1-13.

[10] Allehaibi, K.H., 2025. Artificial Intelligence based Automated Sign Gesture Recognition Solutions for Visually Challenged People. *Journal of Intelligent Systems and Internet of Things*, (2), pp.127-27.

[11] Chang, V., Eniola, R.O., Golightly, L. and Xu, Q.A., 2023. An Exploration into Human–Computer Interaction: Hand Gesture Recognition Management in a Challenging Environment. *SN Computer Science*, *4*(5), p.441.

[12] Alashhab, S., Gallego, A.J. and Lozano, M.Á., 2022. Efficient gesture recognition for the assistance of visually impaired people using multi-head neural networks. *Engineering Applications of Artificial Intelligence*, *114*, p.105188.

[13] Lindner, T., Wyrwał, D. and Milecki, A., 2023. An autonomous humanoid robot designed to assist a human with a gesture recognition system. *Electronics*, *12*(12), p.2652.

[14] Gupta, R., Oza, D. and Chaudhari, S., 2022. Real-Time Hand Tracking and Gesture Recognizing Communication System for Physically Disabled People. In *Inventive Communication and Computational Technologies: Proceedings of ICICCT 2021* (pp. 731-746). Springer Singapore.

[15] Amangeldy, N., Milosz, M., Kudubayeva, S., Kassymova, A., Kalakova, G. and Zhetkenbay, L., 2023. A Real-Time Dynamic Gesture Variability Recognition Method Based on Convolutional Neural Networks. *Applied Sciences*, *13*(19), p.10799.

[16] de Oliveira, G.A., Oliveira, O.D.F., de Abreu, S., de Bettio, R.W. and Freire, A.P., 2022. Opportunities and accessibility challenges for open-source general-purpose home automation mobile applications for visually disabled users. *Multimedia Tools and Applications*, *81*(8), pp.10695-10722.

[17] Sruthi, C.J. and Lijiya, A., 2023. Double-handed dynamic gesture recognition using contour-based hand tracking and maximum mean probability ensembling (MMPE) for Indian Sign language. *The Visual Computer*, *39*(12), pp.6183-6203.

[18] Göreke, V., 2023. A novel method based on Wiener filter for denoising Poisson noise from medical X-Ray images. *Biomedical Signal Processing and Control*, *79*, p.104031.

[19] Safie, S.I. and Ramli, R., 2023. Footprint biometric authentication using SqueezeNet. *Indonesian Journal of Electrical Engineering and Computer Science*, *31*(2), pp.893-901

[20] Yu, Y., Wang, Y., Xu, D., Dou, Z. and Yang, M., 2023. Research on charging and discharging strategy of electric vehicles in park micro-grid based on pigeon-inspired optimization algorithm. *International Journal of Innovative Computing, Information and Control*, *19*(3), pp.721-735.

[21] Yang, M. and Liu, Y., 2023. Research on the potential for China to achieve carbon neutrality: A hybrid prediction model integrated with elman neural network and sparrow search algorithm. *Journal of Environmental Management*, *329*, p.117081.

[22] Zhang, J., Zhang, G., Kong, M. and Zhang, T., 2023. Adaptive infinite impulse response system identification using an enhanced golden jackal optimization. *The Journal of Supercomputing*, *79*(10), pp.10823-10848.

[23] https://www.kaggle.com/datasets/ayuraj/asl-dataset

[24] Wadhawan, A. and Kumar, P., 2020. Deep learning-based sign language recognition system for static signs. *Neural computing and applications*, *32*(12), pp.7957-7968.

[25] Mannan, A., Abbasi, A., Javed, A.R., Ahsan, A., Gadekallu, T.R. and Xin, Q., 2022. Hypertuned deep convolutional neural network for sign language recognition. *Computational Intelligence and Neuroscience*, *2022*.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."