Journal of Intelligent Systems and Internet of Things
Vol. 14, No. 02, PP. 127-139, 2025
DOI: https://doi.org/10.54216/JISIoT.140211

Artificial Intelligence based Automated Sign Gesture Recognition Solutions for Visually Challenged People

Khalid Hamed Allehaibi[1,*]

[1]Computer Science Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Emails: kallehaibi@kau.edu.sa

Abstract

Gesture recognition is employed in human-machine communications, enhancing human life with impairments or who depend on non-verbal instructions. Hand gestures role an important role in the field of assistive technology for persons with visual impairments, whereas an optimum user communication design is of major importance. Many authors with substantial development for gesture recognition modeled several methods by using deep learning (DL) methods. This article introduces a Robust Gesture Sign Language Recognition Using Chicken Earthworm Optimization with Deep Learning (RSLR-CEWODL) approach. The projected RSLR-CEWODL algorithm majorly focuses on the recognition and classification of sign language. To accomplish this, the presented RSLR-CEWODL technique utilizes a residual network (ResNet-101) model for feature extraction. For optimal hyper parameter tuning process, the presented RSLR-CEWODL algorithm exploits the CEWO algorithm. Besides, the RSLR-CEWODL technique uses a whale optimization algorithm (WOA) with deep belief network (DBN) method for the sign language recognition method. The simulation result of the RSLR-CEWODL algorithm is tested using sign language datasets and the outcome was measured under various measures. The simulation values demonstrated the enhancements of the RSLR-CEWODL technique over other methodologies.

Received: March 17, 2024 Revised: June 28, 2024 Accepted: October 15, 2024

Keywords: Sign language recognition; Computer vision; Metaheuristics; Hyper parameter tuning; Deep belief network

1. Introduction

Gestures are a major part of our communication. They are a process of non-verbal interchange of information, which stimulated major concerns about the proposal of Human-Computer Interaction (HCI) models, while it permits users for expressing themselves naturally and intuitively in several contexts. In any situation, a single gesture could be more efficient than words (for instance, a pinch gesture creates it simpler for expressing the desired zoom level, than clarifying it with words). In many backgrounds, hand gestures role a vital play in the area of useful technologies for persons with visual impairments, but an optimal user communication design is of major importance. Sign language is a computer vision (CV)-related convoluted language that involves signs created by hand movement with facial languages [1]. And it was a natural language utilized by an individual with no or low hearing sense for interaction. Sign language is utilized for communicating sentences, and letters utilizing distinct hand signs [2]. This type of interaction makes it very easy for persons with hearing disabilities to share their views and even aids in connecting the interaction gap between persons with hearing disabilities and other individuals [3]. Humans were adapting to sign languages for communication since ancient times. Hand signs will be helpful to express any feeling or word. Thus, people across the world use hand signals continuously to convey themselves despite the formulation of writing convention [4]. The automated detection of human signs was a multidisciplinary issue that has not been resolved till now. In the past few decades, many methods were employed which include the utility of machine learning (ML) approaches for sign language detection [5].

With the arrival of deep learning (DL) approaches, there were efforts to identify human signs. Networks that depend on DL paradigms deal with learning algorithms and architectures that are biologically inspired, dissimilar to classical networks [6]. Gesture detection was further classified into two parts they are dynamic and static [7]. The pattern detection issue is related to static gesture detection, in which the feature extraction was the part of pre-processing level [8]. Feature extraction was an indispensable phase in every traditional pattern-recognizing task. The static gestures need just one image for processing inputs to classifiers, and it will take minimal computational cost. Conversely, dynamic gesture was the most difficult task in CV [9]. It necessitates a series of gestures and images identified related to features derived from the devised feature extraction technique. The hearing-impaired person had a concentration on learning hand signs for digits and alphabets for communicating with others; therefore, this work exhibits an accurate analysis among various classes for detecting the accurate hand gesture letters of ASL [10]. Twenty-four gestures of ASL MNIST data have been employed for classification; few have inter-class similarities. A deep neural network (DNN) was typically employed for ASL detection.

This article introduces a Robust Gesture Sign Language Recognition Using Chicken Earthworm Optimization with Deep Learning (RSLR-CEWODL) model. The presented RSLR-CEWODL system utilizes a residual network (ResNet-101) algorithm for feature extraction. For optimal hyper parameter tuning process, the presented RSLR-CEWODL technique exploits the CEWO algorithm. Besides, the RSLR-CEWODL technique uses a whale optimization algorithm (WOA) with deep belief network (DBN) method for the sign language recognition method. The simulation value of the RSLR-CEWODL approach can be tested utilizing a sign language database and the outcome was measured under various measures.

2. Related Works

Mannan et al. [11] utilized a deep CNN for ASL alphabet detection for overcoming the ASL detection challenge. This work proposes an ASL detection system utilizing a deep CNN. The efficiency of DCNN approach enhances with a count of provided data; for this drive, the authors executed the data augmentation system for increasing the dimension of trained data in recent data artificially. In [12], deep transfer learning was projected to detect sign series in sentences always signed in the Indian sign language (ISL) utilizing adequate labelled information of isolated signs and restricting the count of labelled sentence datasets. The presented DL approach comprises CNN, 2 BiLSTM layers, and connectionist temporal classification (CTC) for allowing end-to-end sentence detection without needing the data of sign boundaries. Kothadiya et al. [13] examine a DL-based method which identifies and detects the words in a person's gestures. DL techniques like LSTM and GRU (feedback-related learning approaches), were utilized for recognizing signs in isolated ISL video frames.

In [14], a full self-attention infrastructure to word-level SLR was presented for tackling the above problem that combines Vision Transformer as spatial encoding and an enhanced temporal Transformer. Also, the authors establish the masking future function for improving Transformer for the temporal component. Afterwards, masking future Transformer improves this series by creating the following time invisible at all the moment of frames and makes the last detection. Lu et al. [15] examine the recent DL approaches for SLR. For achieving this purpose, the CapsNet has been presented in the work that illustrates positive outcomes. The authors also present a Selective Kernel Network (SKNet) for extracting spatial data. Sign language is a basic approach of communication, the issues of SLR in digital videos from realtime is developed a novel challenge for investigations.

In [16], the CNN method of DL was employed to train the hand sign language image database. Along with the Adam optimizer system that is recognized for leveraging the adaptive learning system to figure out the rate of learning all the parameters are employed here to define the optimizing value of hyperparameters. In [17], a new system for context-aware continuous SLR utilizing a GAN structure termed SLRGAN was presented. The presented network infrastructure comprises generators, which SLR glosses with removing temporal and spatial features in video series, in addition to a discriminator which estimates the quality of generators predictive with modelling text data at gloss and sentence levels.

3. The Proposed Model

In this article, we have developed a RSLR-CEWODL technique for the recognition and identification of sign language for visually challenged persons. Initially, the presented RSLR-CEWODL technique employed the ResNet-101 model for feature extraction. For optimal hyperparameter tuning process, the presented RSLR-CEWODL technique exploits the CEWO algorithm. Besides, the RSLR-CEWODL technique used the WOA with the DBN method for the sign language recognition process. Fig. 1 demonstrates the working flow of RSLR-CEWODL method.

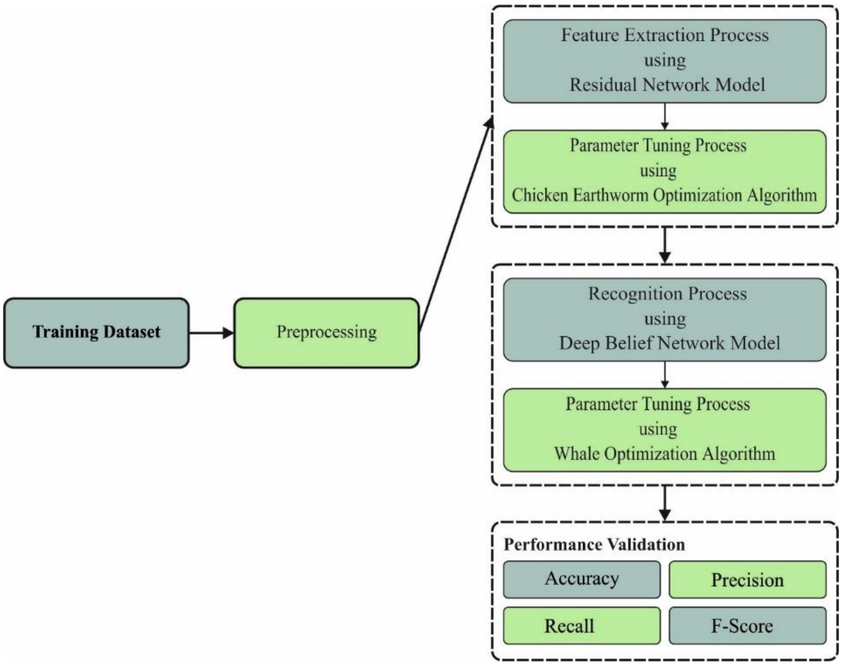

Figure 1. Working flow of RSLR-CEWODL method

A. Feature Extraction Process

In this study, the sign language images undergo a feature extraction procedure by the ResNet-101 model. The representative residual learning infrastructure ResNet101 and ResNet50 are chosen [18]. By utilizing "Sigmoid" activation, the output layer has been altered for the binary classifications. The final numerous layers of ResNet are trained on this dataset. To augment the Bayesian hyperparameter optimization technique, many trainable layers, the optimizer, and the activation are selected. We compared the impacts that either adding or not adding pre-trained weights of "ImageNet". Consider a single block of ResNet as an instance, H(x) was chosen to map to fit certain stacked layers. F(x) was a residual function.

$$F(x) = H(x) - x \tag{1}$$

Since H(x) was very tough for learning compared to F(x). It may change the original mapping to $F(x) + x$.

$$H(x) = F(x) + x \tag{2}$$

Here, the CEWO algorithm can optimally decide the ResNet-101 hyperparameters. CEWO is incorporated into CSO, and EWA [19]. EWA can be a metaheuristic approach stimulated by the reproduction behaviors of earthworms in producing offsprings. EWA was a renowned one that offers trade-offs among exploration and exploitation in meta-heuristic algorithms. EWA implements a local and global search and can be better for parallel computation and presents a balance among intensification and divergence. EWA will be considered as a potential one for parallel computation and allows an efficient balance amongst the exploitation and exploration stages. The CSO can be a biologically inspired optimization technique that stimulates the hierarchy of chicken swarming and the behaviours of finding the food, where every chicken represents a possible outcome for optimization problems. Now, the chicken swarm is classified into chicks, several hens, and a single rooster. Assume $RN, HN, CN, and\ MN$ as the count of roosters, hens, chicks, and mother hens, correspondingly. The Algorithmic steps of CEWO are discussed in the following:

- Initialization: initially, the optimization parameters involving population were initialized: $\{X_{AC}, 1 \leq A \leq H; 1 \leq C \leq S\}$, whereas, H refers to size of populations and C signifies dimensional. $X \in \{w_{gn}, w_j, w_L, b, a\}$.
- Evaluation of objective function: choice of the best chicken place was carried out because of the minimized problems. The least value of main function defines the best solution and thus, the solution with minimal values of the error is selected as the better solution.

$$MSE = \frac{1}{m}\left[\sum_{h=1}^{m} R_{target} - 0_c\right] \tag{3}$$

In Eq. (3), $X_{\tau+1}(A,C)$ denotes the novel position of the rooster $X_\tau$at $\tau+1$iteration, and $X_\tau(A,C)$ indicates the existing position of the rooster at τ iteration, $rand\ (O,\sigma^2)$ signifies the random distribution with mean O and standard deviation σ. The selection of a random rooster can be represented by r, and j_A indicates the fitness of relevant rooster X.

- Drive of hen: Hens will follow the group mate for hunting. Furthermore, it can steal their food arbitrarily, viz., established by another chicken, however, another chicken represses them. Another dominant hen might have an advantage in competing for food over a submissive one.

$$X_{\tau+1}(A,C)=X_\tau(A,C)+U_1*rand*\begin{pmatrix} X_\tau(s_l,C)-X_\tau(A,C)+ \\ U_2*rand(X_\tau(s_2,C)-X_\tau(A,C)\end{pmatrix} \quad (4)$$

Whereas, $U_1=\frac{\exp(0_A-j_{s1})}{abs(j_A)+\in}$

$$U_2=\exp\,(i_{s2}-j_A)$$

In Eq. (4), $rand$ represents the random integer within [0, 1]. s_1 represents the index of rooster, and s indicates the index of chicken that can be arbitrarily selected in the swarm. Now, $s_1\neq s_2$.

The updating formula of EWA is formulated as follows,

$$X_{\tau+1}(\mathrm{A},C)=X_\tau(A,C)+\alpha(X_{best}(A)-X_\tau(A,C))\times K \quad (5)$$

$$X_{\tau+1}(\mathrm{A},C)=X_\tau(A,C)[1-\alpha K]+\alpha\times X_{best}(A)\times K \quad (6)$$

$$X_\tau(A,C)=\frac{1}{(1-\alpha K)}[X_{\tau+1}(\mathrm{A},C)-\alpha\times X_{best}(\tau)\times K] \quad (7)$$

Now, α specifies the similarity factor to determine the distance among parent, and child earthworm, $\alpha\in[0,1]$, and K signify randomly generated integer.

$$X_{\tau+1}(A,C)=\frac{1}{(1-\alpha K)}[X_{\tau+1}(A,C)-\alpha\times X_{best}(\tau)\times K]+U_1*rand*(X_\tau(s_1,C)-X_\tau(A,C)) \quad +U_2*rand*\big(X_\tau(s_2,C)-X_\tau(A,C)\big) \quad (8)$$

Whereas, R_{target} and o_c denote the target and estimated outcome of the classifiers.

- Movement of rooster: Rooster with better fitness might search the food larger than poor fitness value. Therefore, the performance of rooster movement can be formulated by,

$$X_{\tau+1}(A,C)=X_\tau(A,C)*\big(1+rand(0,\sigma^2)\big) \quad (9)$$

$$\sigma^2=\begin{cases}1\ ;\ if j_A\leq f_r \\ \exp\,(\dfrac{j_A-j_{s1}}{(j_A)+\in})\ ;\ otherwise,\end{cases} r\in[1,M],r\neq A \quad (10)$$

$$X_{\tau+1}(A,C)\left[1-\frac{1}{1-\alpha K}\right]=U_1*rand*\big(X_\tau(s_1,C)-X_\tau(A,C)\big)$$
$$+U_2*rand*\big(X_\tau(s_2,C)-X_\tau(A,C)\big)-\alpha\times K\times X_{best}(A) \quad (11)$$

$$X_{\tau+1}(A,C)=\frac{(1-\alpha K)}{-\alpha K}\begin{Bmatrix} U_l*rand*\big(X_\tau(s_l,C)-X_\tau(A,C)\big) \\ +U_2*rand*\big(X_\tau(s_2,C)-X_\tau(A,C)\big)-\alpha\times K\times X_{best}(A)\end{Bmatrix} \quad (12)$$

- Movement of the chick: The chick moves nearby their mother for finding food, and was given as follows:

$$X_{\tau+1}(A,C)=X_\tau(A,C)+RS*\big(X_\tau(u,C)-X_\tau(A,C)\big) \quad (13)$$

In Eq. (13), $X_\tau(u,C)$ indicates the position of C^{th} chicks' mother. RS denotes the velocity of the chick that obeys mother and was selected at random within [0, 2].

- Verifying the feasibility of the solution: This can be calculated according to the main function. When the recently generated solution was superior to the prior one, then it can be changed by a novel solution.
- Termination: The process will terminate, and the optimum solution was retained.

B. Sign Language Recognition Process

To accurately recognize sign language, the DBN model is used. Finally, the WOA is applied to tune the DBN parameters. is a heuristic approach created to mimic the predation behaviors of humpback whales [20]. The process simulates spiral bubble-net hunting and foraging models of humpback whales: random, encircling hunting, and bubble hunting. Specifically, the WOA involves the abovementioned three searching processes. In this work, every decision parameter represents the coordinate of every humpback whale, and the individual probability selecting of 3 behaviors is the same. The presented model uses the optimum outcome in the present group to be the principal whale place and the remaining individual uses the topmost whale location as the prey location. They surround the target by approaching the prey's location and relocating their location:

$$\vec{X}(t+1) = -\vec{D}\cdot\vec{A} + \vec{X}(t) \tag{14}$$

$$\vec{D} = \left|\vec{C}\cdot\vec{X}*(t) - \vec{X}(t)\right| \tag{15}$$

Whereas, t indicates the amount of iterations; $\vec{X}(t+1)$ denotes the individual afterwards $t+1$-generation updating location; $\vec{X}^*(t)$ refers to the leader whale of t-generation location; $\vec{A}$ shows the convergence aspect; $\vec{X}(t)$ denotes the location of the individual afterwards t-generation was upgraded; $\vec{C}$ represent swing factor and it could be formulated as:

$$\vec{C} = 2\vec{r} \tag{16}$$

$$\vec{A} = -\vec{a} + 2\vec{r}.\vec{a} \tag{17}$$

In Eq. (17), $\vec{r}$ symbolizes the randomly generated integer ranges from 0 to 1, $\vec{a}$ slowly rises in the iteration method and linearly reduces from 2 to 0. The spiral updating equation realizes the bubble predation of whales. The location updating equation can be given in the following:

$$\vec{X}(t+1) = \vec{X}*(t) + \cos(2\pi l)\cdot e^{b1}\cdot\vec{D}' \tag{18}$$

$$\vec{D}' = \left|\vec{X}*(t) - \vec{X}(t)\right| \tag{19}$$

The distance among a few whales is resolved in the iteration method and an optimum individual; is a spiral constant. During hunting, whales randomly change their position based on the location of other whales, thus extending the search range to find the best target.

If $|A| > 1$, the individual whale extends the searching range for arbitrary predation and implements global search:

$$\vec{X}(t+1) = -\vec{D}\cdot\vec{A} + \vec{X}_{rand}(t) \tag{20}$$

$$\vec{D} = \left|\vec{C}\cdot\vec{X}_{rand}(t) - \vec{X}(t)\right| \tag{21}$$

Whereas, $\vec{D}$ indicates the distance amongst a specific whale resolved in the iteration and arbitrarily chosen individual whale; $\vec{X}$ represent the random individual whale in the existing position of the population. Fig. 2 illustrates the structure of DBN.

$$\vec{X}(t+1) = \begin{cases} -\vec{D}\cdot\vec{A} + \vec{X}*(t) & ,p > 0.5 \\ -\vec{D}\cdot\vec{A} + \vec{X}_{rand}(t) & ,p < 0.5 \cup A < 1 \\ \vec{X}*(t) + \cos(2\pi l)\cdot e^{b1}\cdot\vec{D} & ,p < 0.5 \cup A \geq 1 \end{cases} \tag{22}$$

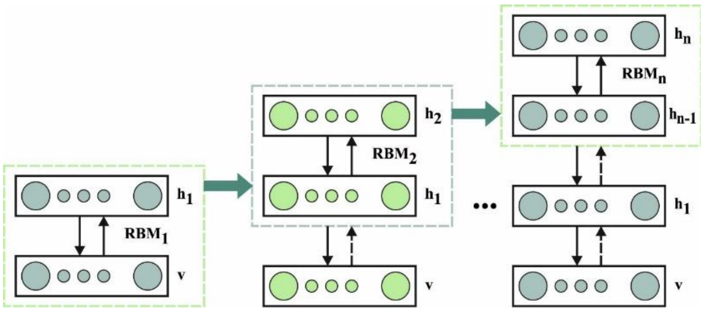

Figure 2. Architecture of DBN

The WOA method establish a FF to have superior classifier performance. It sets a positive value for suggesting higher solutions for the candidate's performance. The decrease of the classifier rate of error was selected as the FF in this case, as expressed in Eq. (23).

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{No.of\ misclassified\ instances}{Total\ No.of\ instances} * 100 \tag{23}$$

4. Results and Discussion

The experimental outcome of RSLR-CEWODL method on the sign language recognition method is tested using two aspects: alphabets (a-z) and numerals (0-9). Fig. 3 represents some sample images.

Figure 3. Sample images

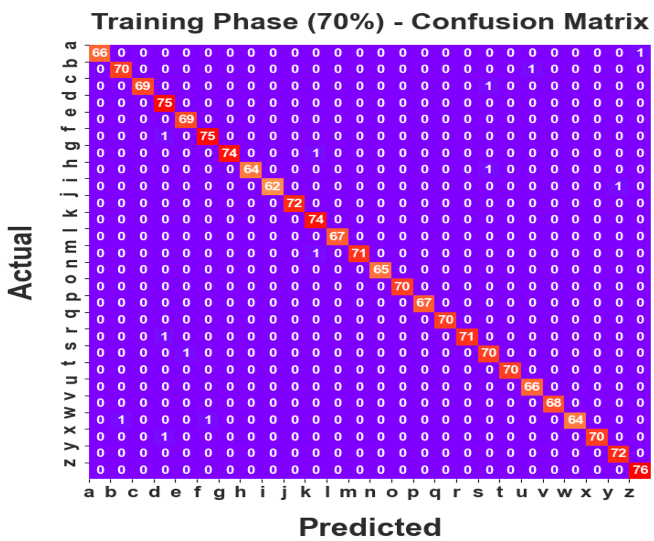

Figure 4. Confusion matrix of RSLR-CEWODL algorithm with 70% of TR database under ASR

The confusion matrix of the RSLR-CEWODL algorithm with 70% of the TR database on alphabet sign recognition (ASR) process is revealed in Fig. 4. The outcome implied the RSLR-CEWODL methodology has capably categorized all 26 characters.

Table 1 reveals the overall classifier outcomes of the RSLR-CEWODL algorithm with 70% of the TR database on ASR. The outcomes designated the RSLR-CEWODL method have obtained effectual recognition performance in all class labels.

Table 1: Overall classification of RSLR-CEWODL system with 70% of TR database under ASR

Training Phase (70%)									
Class	Acc.	PPV	DR	F-Score	Class	Acc.	PPV	DR	F-Score
a	99.95	100.00	98.51	99.25	n	100.00	100.00	100.00	100.00
b	99.89	98.59	98.59	98.59	o	100.00	100.00	100.00	100.00
c	99.95	100.00	98.57	99.28	p	100.00	100.00	100.00	100.00
d	99.84	96.15	100.00	98.04	q	100.00	100.00	100.00	100.00
e	99.95	98.57	100.00	99.28	r	99.95	100.00	98.61	99.30
f	99.89	98.68	98.68	98.68	s	99.84	97.22	98.59	97.90
g	99.95	100.00	98.67	99.33	t	100.00	100.00	100.00	100.00
h	99.95	100.00	98.46	99.22	u	99.95	98.51	100.00	99.25
i	99.95	100.00	98.41	99.20	v	100.00	100.00	100.00	100.00
j	100.00	100.00	100.00	100.00	w	99.89	100.00	96.97	98.46
k	99.89	97.37	100.00	98.67	x	99.95	100.00	98.59	99.29
l	100.00	100.00	100.00	100.00	y	99.95	98.63	100.00	99.31
m	99.95	100.00	98.61	99.30	z	99.95	98.70	100.00	99.35

The confusion matrix of the RSLR-CEWODL methodology with 30% of the TS database on the ASR process is shown in Fig. 5. The outcome demonstrated the RSLR-CEWODL algorithm has proficiently categorized all 26 characters.

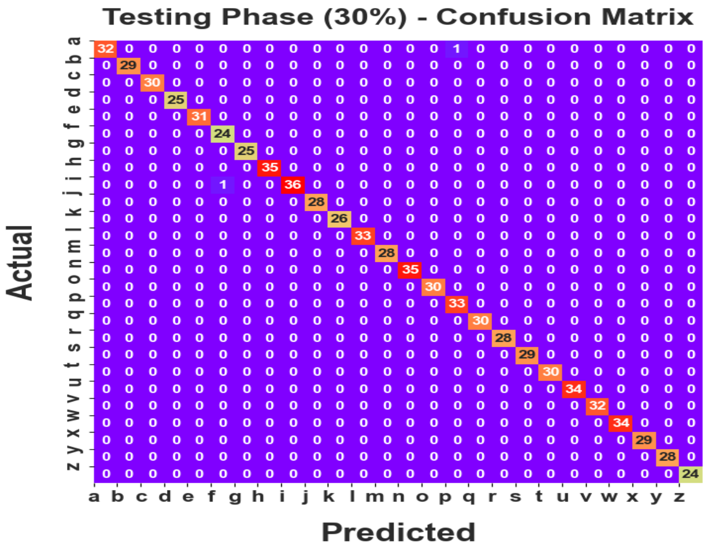

Figure 5. Confusion matrix of RSLR-CEWODL algorithm with 30% of TS database under ASR

Table 2 demonstrated the classification outcome of the RSLR-CEWODL algorithm with 30% of the TS database on ASR. The outcome displayed the RSLR-CEWODL approach has gained effectual recognition performance in all class labels.

Table 2: Overall classification of RSLR-CEWODL system with 30% of TS database under ASR

Testing Phase (30%)									
Class	Acc.	PPV	DR	F-Score	Class	Acc.	PPV	DR	F-Score
a	99.87	100.00	96.97	98.46	n	100.00	100.00	100.00	100.00
b	100.00	100.00	100.00	100.00	o	100.00	100.00	100.00	100.00
c	100.00	100.00	100.00	100.00	p	99.87	97.06	100.00	98.51

d	100.00	100.00	100.00	100.00	q	100.00	100.00	100.00	100.00
e	100.00	100.00	100.00	100.00	r	100.00	100.00	100.00	100.00
f	99.87	96.00	100.00	97.96	s	100.00	100.00	100.00	100.00
g	100.00	100.00	100.00	100.00	t	100.00	100.00	100.00	100.00
h	100.00	100.00	100.00	100.00	u	100.00	100.00	100.00	100.00
i	99.87	100.00	97.30	98.63	v	100.00	100.00	100.00	100.00
j	100.00	100.00	100.00	100.00	w	100.00	100.00	100.00	100.00
k	100.00	100.00	100.00	100.00	x	100.00	100.00	100.00	100.00
l	100.00	100.00	100.00	100.00	y	100.00	100.00	100.00	100.00
m	100.00	100.00	100.00	100.00	z	100.00	100.00	100.00	100.00

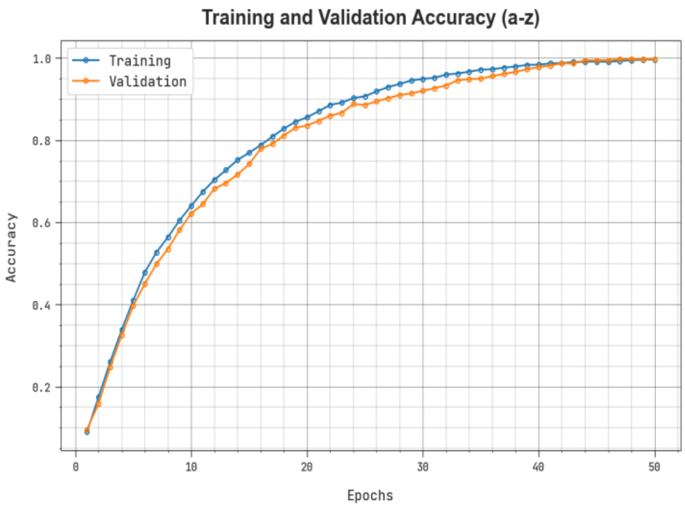

Figure 6. TRAC and VDAC curve of RSLR-CEWODL approach under ASR

The TRAC and VDAC of the RSLR-CEWODL methodology on ASR solution in Fig. 6. The outcome displays that the RSLR-CEWODL technique has enhanced the outcome with maximal values of TRAC and VDAC. It could be obvious that the RSLR-CEWODL method has maximal TRAC results.

The TRLS and VDLS of the RSLR-CEWODL algorithm under ASR solution in Fig. 7. The outcome illustrates the RSLR-CEWODL method has showed higher results with lesser values of TRLS and VDLS. Note that the RSLR-CEWODL system has resulted in lesser VDLS performances.

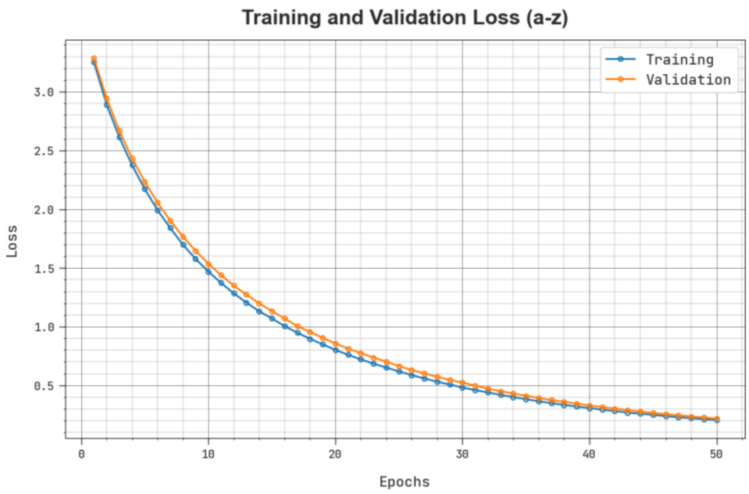

Figure 7. TRLS and VDLS outcome of RSLR-CEWODL method under ASR

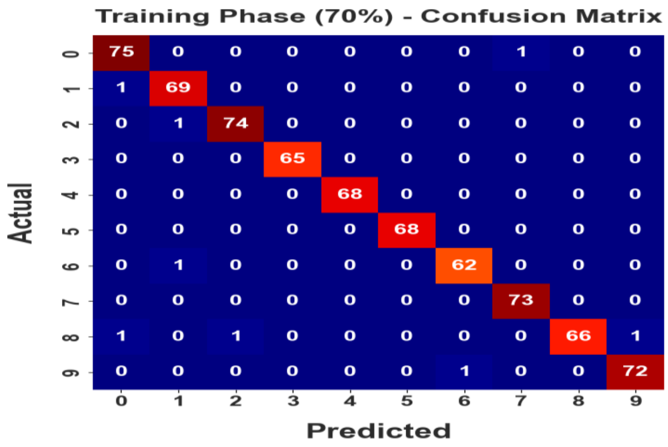

Figure 8. Confusion matrix of RSLR-CEWODL method with 70% of TR database under NSR

The confusion matrix of the RSLR-CEWODL algorithm with 70% of the TR database on the number sign recognition (NSR) process is represented in Fig. 8. The outcome specified the RSLR-CEWODL system has proficiently categorized all 10 numbers.

Table 3 establishes the overall classifier outcome of the RSLR-CEWODL algorithm with 70% of the TR database on NSR. The outcome designated the RSLR-CEWODL method has gained effectual recognition performance under all class labels.

Table 3: Overall classification of RSLR-CEWODL system with 70% of TR database under NSR

Training Phase (70%)				
Class	Acc.	PPV	DR	F-Score
0	99.57	97.40	98.68	98.04
1	99.57	97.18	98.57	97.87
2	99.71	98.67	98.67	98.67
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	99.71	98.41	98.41	98.41
7	99.86	98.65	100.00	99.32
8	99.57	100.00	95.65	97.78
9	99.71	98.63	98.63	98.63

The confusion matrix of the RSLR-CEWODL methodology with 30% of the TS database on the NSR process is given in Fig. 9. The outcome outperformed the RSLR-CEWODL method and has proficiently categorized all 10 classes.

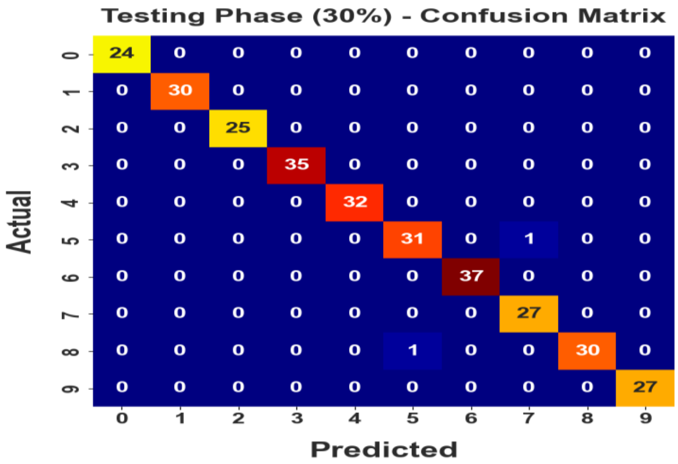

Figure 9. Confusion matrix of RSLR-CEWODL algorithm with 70% of TS database under NSR

Table 4 exhibits the overall classifier outcome of the RSLR-CEWODL method with 30% of the TR database on NSR. The outcomes designated the RSLR-CEWODL approach have attained effectual recognition performance under all classes.

Table 4: Overall classification of RSLR-CEWODL system with 30% of TS database under NSR

Testing Phase (30%)				
Class	Acc.	PPV	DR	F-Score
0	100.00	100.00	100.00	100.00
1	100.00	100.00	100.00	100.00
2	100.00	100.00	100.00	100.00
3	100.00	100.00	100.00	100.00
4	100.00	100.00	100.00	100.00
5	99.33	96.88	96.88	96.88
6	100.00	100.00	100.00	100.00
7	99.67	96.43	100.00	98.18
8	99.67	100.00	96.77	98.36
9	100.00	100.00	100.00	100.00

The TRAC and VDAC of the RSLR-CEWODL algorithm on NSR solution in Fig. 10. The outcome implied that the RSLR-CEWODL methodology has depicted a greater solution with greater values of TRAC and VDAC. It could be obvious that the RSLR-CEWODL methodology has accomplished maximal TRAC performances.

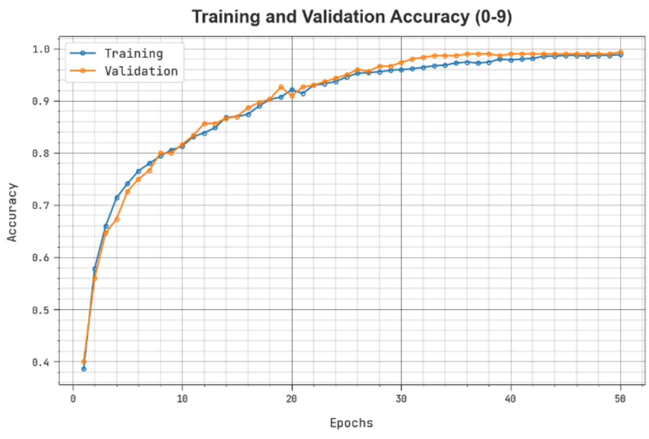

Figure 10. TRAC and VDAC outcome of RSLR-CEWODL algorithm under NSR

The TRLS and VDLS of the RSLR-CEWODL algorithm under NSR solution in Fig. 11. The outcome outperformed that the RSLR-CEWODL methodology has exposed optimum solution with reduced values of TRLS and VDLS. It could be noted that the RSLR-CEWODL method has resulted in lesser VDLS performances.

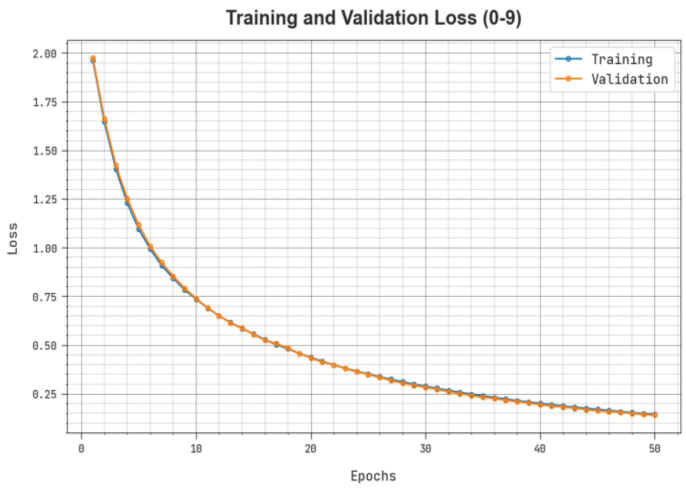

Figure 11. TRLS and VDLS outcome of RSLR-CEWODL algorithm under NSR

Table 5 reports an overall sign language recognition performance of the RSLR-CEWODL model on characters (a-z) and numerals (0-9). The outcome stated that the RSLR-CEWODL method accurately recognized all the classes. For example, on alphabet sign recognition, the RSLR-CEWODL approach has obtained an average $accu_y$ of 99.97%, PPV of 99.53%, DR of 99.53%, and F_{score} of 99.53%. Meanwhile, on number sign recognition, the RSLR-CEWODL method has gained an average $accu_y$ of 99.82%, PPV of 99.11%, DR of 99.11%, and F_{score} of 99.11%.

Table 5: Overall sign language recognition outcome of the RSLR-CEWODL system

	Acc.	PPV	DR	F-Score
(a-z)				
Training Phase	99.95	99.32	99.28	99.30
Testing Phase	99.98	99.73	99.78	99.75
Average	99.97	99.53	99.53	99.53
(0-9)				
Training Phase	99.77	98.89	98.86	98.87
Testing Phase	99.87	99.33	99.36	99.34
Average	99.82	99.11	99.11	99.11

Eventually, a comparative study of the RSLR-CEWODL algorithm with other systems in Table 6 and Fig. 12 [11, 21]. The simulation values implied that the RSLR-CEWODL system has offered an improved $accu_y$ value of 99.82%. Contrastingly, the Hypertuned DCNN, CNN, ANN, KNN, American SLR-LMS, Arabic HGR-LMC, SLSHG-LMC-RNN, and SVM models have reached reduced $accu_y$ of 99.67%, 99.62%, 98.00%, 95.95%, 79.83%, 91.30%, 96.41%, and 97.90% respectively.

Table 6: Accuracy outcome of RSLR-CEWODL algorithm with other methodologies

Technique used	Accuracy (%)
RSLR-CEWODL	99.82
Hypertuned DCNN	99.67
CNN Model	99.62
ANN Model	98.00
K-nearest neighbours	95.95
American SLR-LMS	79.83
Arabic HGR-LMC	91.30
SLSHG-LMC-RNN	96.41
SVM Model	97.90

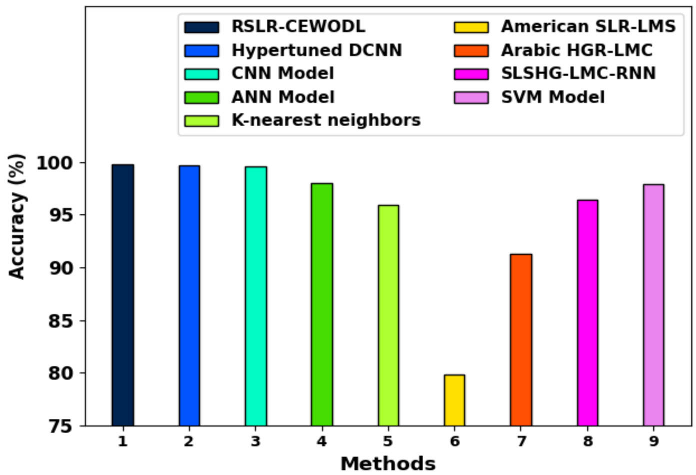

Figure 12. Accuracy analysis of RSLR-CEWODL algorithm with other methodologies

These outcomes show that the RSLR-CEWODL algorithm has exhibited maximal performance over other approaches.

5. Conclusion

In this article, we have developed the presentation of RSLR-CEWODL algorithm for the detection and classification of sign language. Initially, the presented RSLR-CEWODL technique employed the ResNet-101 model for feature extraction. For optimal hyperparameter tuning process, the presented RSLR-CEWODL technique exploits the CEWO algorithm. Besides, the RSLR-CEWODL technique used the WOA with the DBN method for the sign language recognition model. The simulation outcome of the RSLR-CEWODL methodology was tested by exploiting the sign language dataset and the outcome are evaluated in varying aspects. The simulation outcome showed the developments of the RSLR-CEWODL approach over other techniques. In future, hybrid DL algorithms can increase the efficiency of the RSLR-CEWODL methodology.

Funding: "The author gratefully acknowledges technical support provided by Faculty of Computing and Information Technology at King Abdulaziz University, Jeddah, Saudi Arabia"

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

References

[1] 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.

[2] Ryumin, D., Ivanko, D. and Ryumina, E., 2023. Audio-Visual Speech and Gesture Recognition by Sensors of Mobile Devices. Sensors, 23(4), p.2284.

[3] Alrowais, F., Marzouk, R., Al-Wesabi, F.N. and Hilal, A.M., 2023. Hand Gesture Recognition for Disabled People Using Bayesian Optimization with Transfer Learning. Intelligent Automation & Soft Computing, 36(3).

[4] Lee, J.W. and Yu, K.H., 2023. Wearable Drone Controller: Machine Learning-Based Hand Gesture Recognition and Vibrotactile Feedback. Sensors, 23(5), p.2666.

[5] Lu, W.X., Fang, P., Zhu, M.L., Zhu, Y.R., Fan, X.J., Zhu, T.C., Zhou, X., Wang, F.X., Chen, T. and Sun, L.N., 2023. Artificial Intelligence–Enabled Gesture-Language-Recognition Feedback System Using Strain-Sensor-Arrays-Based Smart Glove. Advanced Intelligent Systems, p.2200453.

[6] Adaloglou, N., Chatzis, T., Papastratis, I., Stergioulas, A., Papadopoulos, G.T., Zacharopoulou, V., Xydopoulos, G.J., Atzakas, K., Papazachariou, D. and Daras, P., 2021. A comprehensive study on deep learning-based methods for sign language recognition. IEEE Transactions on Multimedia, 24, pp.1750-1762.

[7] John, J. and Sherif, B.V., 2022. Hand Landmark-Based Sign Language Recognition Using Deep Learning. In Machine Learning and Autonomous Systems (pp. 147-157). Springer, Singapore.

[8] Rastgoo, R., Kiani, K. and Escalera, S., 2021. Sign language recognition: A deep survey. Expert Systems with Applications, 164, p.113794.

[9] Gupta, R., Golaya, S. and Srinivasan, R., 2022, March. Transfer-Learning Based User-Personalization of Indian Sign Language Recognition System. In 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS) (Vol. 1, pp. 615-620). IEEE.

[10] Gao, Y., Jia, C., Chen, H. and Jiang, X., 2020. Chinese fingerspelling sign language recognition using a nine-layer convolutional neural network. EAI Endorsed Transactions on e-Learning, 7(20).

[11] 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.

[12] Sharma, S., Gupta, R. and Kumar, A., 2021. Continuous sign language recognition using isolated signs data and deep transfer learning. Journal of Ambient Intelligence and Humanized Computing, pp.1-12.

[13] Kothadiya, D., Bhatt, C., Sapariya, K., Patel, K., Gil-González, A.B. and Corchado, J.M., 2022. Deepsign: Sign Language Detection and Recognition Using Deep Learning. Electronics, 11(11), p.1780.

[14] Du, Y., Xie, P., Wang, M., Hu, X., Zhao, Z. and Liu, J., 2022. Full Transformer Network with Masking Future for Word-Level Sign Language Recognition. Neurocomputing.

[15] Lu, J., Nguyen, M. and Yan, W.Q., 2021, January. Sign language recognition from digital videos using deep learning methods. In International Symposium on Geometry and Vision (pp. 108-118). Springer, Cham.

[16] Arora, S., Gupta, A., Jain, R. and Nayyar, A., 2021. Optimization of the CNN model for hand sign language recognition using Adam optimization technique. In Micro-electronics and telecommunication engineering (pp. 89-104). Springer, Singapore.

[17] Papastratis, I., Dimitropoulos, K. and Daras, P., 2021. Continuous sign language recognition through a context-aware generative adversarial network. Sensors, 21(7), p.2437.

[18] Ji, L., Mao, R., Wu, J., Ge, C., Xiao, F., Xu, X., Xie, L. and Gu, X., 2022. Deep Convolutional Neural Network for Nasopharyngeal Carcinoma Discrimination on MRI by Comparison of Hierarchical and Simple Layered Convolutional Neural Networks. Diagnostics, 12(10), p.2478.

[19] Vijay, M. and Indumathi, G., 2021. Deep belief network-based hybrid model for multimodal biometric system for futuristic security applications. Journal of Information Security and Applications, 58, p.102707.

[20] Cheng, J., Xu, J., Chen, W. and Song, B., 2022. Locating and sizing method of electric vehicle charging station based on Improved Whale Optimization Algorithm. Energy Reports, 8, pp.4386-4400.

[21] 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.