Intelligent Automation & Soft Computing
DOI: 10.32604/iasc.2023.036354
Article

Tech Science Press

Hand Gesture Recognition for Disabled People Using Bayesian Optimization with Transfer Learning

Fadwa Alrowais[1], Radwa Marzouk[2,3], Fahd N. Al-Wesabi[4,*] and Anwer Mustafa Hilal[5]

[1]Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O.Box 84428, Riyadh, 11671, Saudi Arabia
[2]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O.Box 84428, Riyadh, 11671, Saudi Arabia
[3]Department of Mathematics, Faculty of Science, Cairo University, Giza, 12613, Egypt
[4]Department of Computer Science, College of Science & Art at Mahayil, King Khalid University, Saudi Arabia
[5]Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia
*Corresponding Author: Fahd N. Al-Wesabi. Email: falwesabi@kku.edu.sa
Received: 27 September 2022; Accepted: 14 November 2022

Abstract: Sign language recognition can be treated as one of the efficient solutions for disabled people to communicate with others. It helps them to convey the required data by the use of sign language with no issues. The latest developments in computer vision and image processing techniques can be accurately utilized for the sign recognition process by disabled people. American Sign Language (ASL) detection was challenging because of the enhancing intraclass similarity and higher complexity. This article develops a new Bayesian Optimization with Deep Learning-Driven Hand Gesture Recognition Based Sign Language Communication (BODL-HGRSLC) for Disabled People. The BODL-HGRSLC technique aims to recognize the hand gestures for disabled people's communication. The presented BODL-HGRSLC technique integrates the concepts of computer vision (CV) and DL models. In the presented BODL-HGRSLC technique, a deep convolutional neural network-based residual network (ResNet) model is applied for feature extraction. Besides, the presented BODL-HGRSLC model uses Bayesian optimization for the hyperparameter tuning process. At last, a bidirectional gated recurrent unit (BiGRU) model is exploited for the HGR procedure. A wide range of experiments was conducted to demonstrate the enhanced performance of the presented BODL-HGRSLC model. The comprehensive comparison study reported the improvements of the BODL-HGRSLC model over other DL models with maximum accuracy of 99.75%.

Keywords: Deep learning; hand gesture recognition; disabled people; computer vision; bayesian optimization

1 Introduction

People with speaking and hearing disabilities utilize various communication models, and sign language is a commonly employed mode of communication. Sign language enables people to communicate with human body language; every individual word has a collection of human actions defining a specific expression. Individuals who cannot hear or speak use sign language (SL) to express their opinions and interact with ordinary people or among them [1]. SL involves the movement of various parts, such as the mouth, hand, lips, and many others denoted as gestures. Such languages comprise certain broadly accepted SL, namely Indian Sign Language (ISL), American Sign Language (ASL), Sign language in japan (JSL), and so on [2]. But the difficulty faced by these people who were deaf and hard of hearing is that ordinary persons could not understand the sign they used to communicate as they did not have proper knowledge of various signs used for communication [3].

Hand gesture recognition (HGR) is one such most advanced field where artificial intelligence (AI) and computer vision (CV) has assisted in enhancing transmission with deaf people while also supporting gesture-related signalling system [4]. Sub-fields of HGR add physical exercise monitoring, SL recognition, human action recognition, pose and posture detection, detection of special signal language utilized in sports, and controlling smart home or supported living applications with HGR [5]. For the last few years, computer scientists have utilized various computational methods and algorithms to solve issues and ease our lives [6]. The usage of hand gestures (HGs) in various software applications had more contribution to enhancing human and computer communications. The development of the gesture recognition system serves an important play in the progression of human and computer communications, and the usage of HGs from several fields was rising more frequently [7]. The application of HGs is realized in games, cognitive development assessment, virtual and augmented reality, assisted living, etc. The expansion of HGR in various sectors has gained the interest of industry and also human-robot communication in the industrial and control of autonomous cars [8]. The ultimate purpose of this practical HGR application was to recognize and classify the gestures. Hand recognition was an approach that utilized various concepts and algorithms of numerous methods, such as NNs and image processing, for understanding hand movement [9]. Generally, there were countless applications of HGR. For instance, deaf people that cannot hear it could interact with their standard SL [10].

This article develops a new Bayesian Optimization with Deep Learning-Driven Hand Gesture Recognition Based Sign Language Communication (BODL-HGRSLC) for Disabled People. The BODL-HGRSLC technique aims to recognize the HGs for disabled people's communication. The presented BODL-HGRSLC technique integrates the concepts of computer vision (CV) and DL models. In the presented BODL-HGRSLC technique, a deep convolutional neural network-based residual network (ResNet) model is applied for feature extraction. Besides, the presented BODL-HGRSLC model uses Bayesian optimization for the hyperparameter tuning process. At last, a bidirectional gated recurrent unit (BiGRU) model is exploited for the HGR procedure. A wide range of experiments was conducted to demonstrate the enhanced performance of the presented BODL-HGRSLC model.

2 Literature Review

Su et al. [11] examine a gesture detection approach with EMG signals dependent upon deep multi-parallel CNN that resolves the issue of which standard ML approaches can fall into too much helpful data in extracting features. The CNNs offer an effective manner for constraining the difficulty of FFNNs by weighted sharing and limiting local connections. The sophisticated extracting feature is that avoided, and HG is classified directly. A multi-parallel and multi-convolutional layer convolutional infrastructure were presented for classifying HG. Wang et al. [12] examine a deformable convolutional network (DCN) for optimizing the standard convolution kernels to obtain optimum performances of sEMG-based gesture

detection. The DCN initially executes a standard convolutional layer for obtaining a lower-dimensional feature map and then utilizes a deformable convolution layer to obtain a higher-dimensional feature map. In addition, the authors present and relate two novel image representation approaches dependent upon standard extracting features that allow DL structures for extracting understood correlations betwixt distinct channels in the sparse multi-channel sEMG signals.

Mujahid et al. [13] examine a lightweight technique dependent upon YOLOv3 and DarkNet-53 are CNNs for gesture detection without another improvement of images, pre-processing, and image filtering. The presented technique obtained maximum accuracy even in difficult environments and effectively-identified gestures from lower-resolution picture mode. The presented technique has been estimated on the labelled dataset of HG in either Pascal VOC or YOLO format. Wong et al. [14] introduce an effectual, low-cost capacitive sensor device for HGR. Particularly, the authors planned a prototype of wearable capacitive sensor units for capturing the capacitance values in the electrodes located in finger phalanges. Several examines followed to provide further insights into sensing data. The authors implemented and related 2 ML approaches such as Error Correction Output Code SVMs (ECOC-SVM) and KNN classifications. In [15], an HGR model utilizing a single patchable six-axis IMU involved at the wrist using RNN was projected. The presented patchable IMU with soft procedure factors is worn from close contact with a human body, easily adjusting for skin deformation. So, signal distortion (that is, motion artefacts) created by vibration in the motion has been minimized. Besides, the patchable IMU is a wireless communication (for instance, Bluetooth) component for always sending the sensor signals to some processing devices.

Yang et al. [16] present an HGR technique dependent upon range-Doppler-angle paths and a reused LSTM (RLSTM) network utilizing a 77 GHz frequency modulated continuous wave (FMCW) multiple-input-multiple-output (MIMO) radar. To overcome the count of gesture interference, a gesture desktop was planned, and a potential HGR technique was followed to define whether it could be possible HG. Once the gesture ensues in the considered gesture desktop, the range-Doppler-angle paths were extracted by utilizing the multiple signal classification (MUSIC), discretized Fourier transform (DFT), and Kalman filtering (KF). Nasri et al. [17] examine an sEMG-controlled 3D game which leverages a DL-based structure for real-time gesture detection. The 3D game knowledge established in the case concentrated on rehabilitation movements, permitting persons with particular disabilities to utilize low-cost sEMG sensors to control the game experiences.

3 The Proposed Gesture Recognition Model

This article develops a new BODL-HGRSLC technique to recognize the HGs for disabled people's communication. The presented BODL-HGRSLC technique integrates the concepts of CV and DL models. It encompasses ResNet feature extraction, BO hyperparameter tuning, and BiGRU recognition processes. Fig. 1 represents the block diagram of the BODL-HGRSLC system.

3.1 Module I: Feature Extraction

In the presented BODL-HGRSLC technique, the ResNet model is applied for feature extraction. A research worker demonstrates a DRL architecture. The study aims to facilitate network training that is deeper when compared to the former [18]. This presented architecture helps to address the degradation problem. Instead of expecting each small number of stacked layers to fit the mapping directly, researchers permitted this layer to fit a residual mapping.

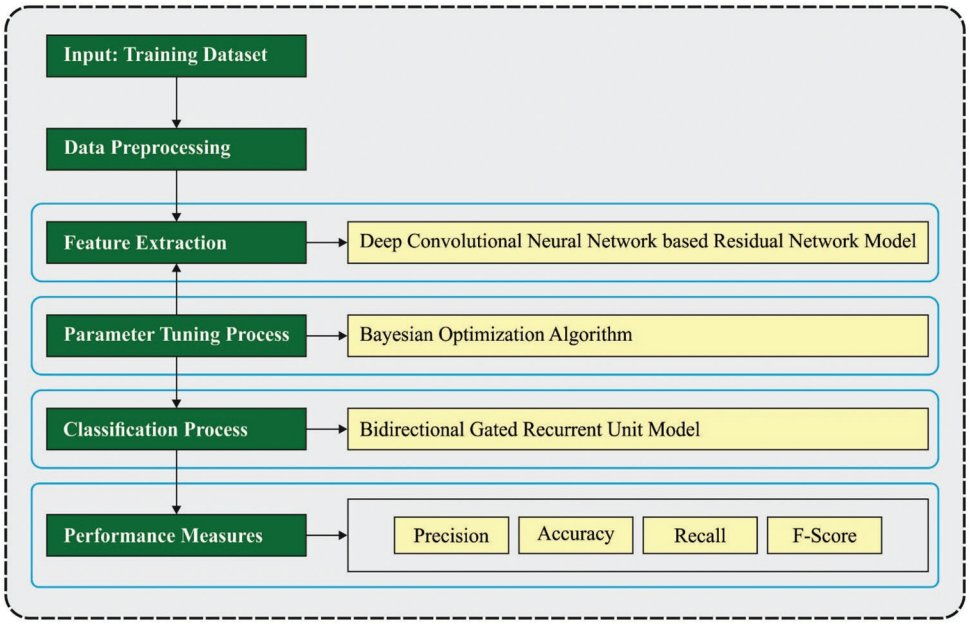

Figure 1: Block diagram of BODL-HGRSLC system

The concept of ResNet is to construct a bypass connection; it connects the deep layer by bypassing the nonlinear conversion layer. The block output is the addition of this connection output and network stacked layer output as (1).

$$H(X) = F(X) + X \tag{1}$$

Now, X represents the connection output, $F(X)$ indicates the network stacked layer output, and $H(X)$ represents the block output.

Researchers are trained in the ResNet model using ImageNet data; the test results were 3.57% errors in prediction. Each convolution filter has a stride of 2. The study used ResNet 50, where the global average pooling shapes $(1,\ 2048)$.

Next, the presented BODL-HGRSLC model uses Bayesian optimization (BO) for the hyperparameter tuning process. Generally, there exist two kinds of hyperparameter optimization techniques: automatic and manual search. Manual hyperparameter optimization is a challenging process to regenerate because it depends on various trial and error efforts [19]. Grid search isn't adaptable for high dimensions. Random search servers are a greedy method, which settles for local optimal and therefore does not obtain global optimal. Other evolutionary optimized techniques need a broad range of trained cycles and could be noisy. From the above mentioned, the BO technique could resolve this constraint by proficiently defining the global optimization of the black box function of NNs, and it can be based on the Bayes theorem. BO technique is used to resolve computationally costly function to discover the extrema.

The major components in the optimized technique are given below:

- The gaussian approach of $f(x)$.
- The bayesian updating process is used to modify the Gaussian approach at every novel assessment of $f(x)$.
- An acquisition function $an\ (x)$ depends on the Gaussian approach of f that is maximized for determining the subsequent point χ for estimation.

This model determines where the function attains the optimum value, thereby maximizing the model's accuracy and reducing loss. As already mentioned, the optimization aims at finding a minimal value of loss in the sampling point to unknown function f:

$$x_{opt} = \underset{x\epsilon D}{\arg\min}\, f(x), \tag{2}$$

whereas D represents the searching domain of χ.

The fundamental probabilistic method for *the f* objective function is a Gaussian model prior with additional Gaussian noise in the observation that is thorough. The gaussian approach works in a way that anticipates output approximating input and therefore assumes a statistical algorithm.

$$P(M|E) \infty P(E|M)P(M). \tag{3}$$

The above equation reflects the concept of the BO technique. By observing the sampling dataset E, $P(M|E)$, the posterior probability of *the M* model is proportionate to the probability $P(E|M)$ of detecting E provided model M multiplied with (M) prior probability. It is determined that Bayes optimization enhances unknown functions by merging the previous distribution according to the Gaussian model of *the* $f(x)$ function with present sample data used to acquire the posterior function. Then, the posterior data for finding whereby the criterion values minimalize the f(x) function. This condition has characterized by *an* acquisition or utility function. Function *a* describes the next sampling point for maximizing the expected function. There is some widely employed acquisition function. In the study, we used the expected improvement (EI) function, as it assesses the expected quantity of improvement in the main function, which ignores the value that causes a rise in the objective function. *EI* evaluates the expectation of the degree of enhancement that point could attain while discovering the neighbourhood of the present optimal value. When x_{best} refers to the position of the low posterior mean and $\mu_Q(x_{best})$ indicates the lowermost value of the posterior mean, then the expected development is

$$EI(x,\ \ Q) = E_Q\big[\ \max\big(0,\ \mu_Q(x_{best}) - f(x)\big)\big] \tag{4}$$

In other words, when the development of function values is lesser than the expected values afterwards, the procedure is implemented, the present ideal value point might be the local optimum solution, and the process defines the optimal value point in another position. Searching the sampling areas include exploitation (sample from the highest value) and exploration (sample from the area of highest uncertainty) that assist in reducing the sampling count. At last, improve the performance although the function has more than one local maxima. Besides the sampling data, the BO technique depends on the previous distribution of function *f*, an essential part of the statistical inference of the posterior distribution of functions f.

The major steps in the optimization are given below:

- For existing iterations t
- Estimate $y_i = f(x)$ for determining the amount of points x_i randomly taken within the parameter bound.
- Upgrade the Gaussian f(x) approach to achieving posterior distribution over function $(f|xi,\ yi\ for\ i = 1,\ \ldots t)$.
- Define the novel point χ that maximizes acquisition function (x).

The process ends after the number of iterations limits or reaches a time. In the presented method, tuning the CNN architecture and hyperparameter can be accomplished using the BO technique.

3.2 Module II: Gesture Recognition

Finally, the BiGRU model is exploited for the HGR procedure. An RNN is effectively utilized for handling data series from distinct regions [20]. In RNN, assume that input series $x = (x_1, \ldots, x_T)$, hidden vector series $h = (h_1, \ldots, h_T)$, and resultant vector sequence $y = (y_1, \ldots, y_T)$ are derivative by the provided formulas:

$$h_\tau = \Phi(Ux_t + WWh_{\tau-1} + b) \tag{5}$$

$$y_t = Vh_t + c \tag{6}$$

Assume that Φ is the activation function, and the popular activation function was commonly an element-wise application of sigmoid functions. U represents the input-hidden weighted matrixes, W indicates the hidden-hidden weight matrixes, Eq. (6) b indicates the hidden bias vector, V stands for the hidden-output weighted matrixes, and c implies the resultant bias vector. It can be difficult to capture the long-term dependences of RNN as the gradient inclines explode or vanish. Thus, in several types of research, workers have made every effort to develop a further difficult activation function for resolving the faults. The LSTM unit was primarily presented for the sample to capture the long-term dependencies. Recently, another variation of recurrent units, the GRU, is also developed that is simpler for calculating and taking optimum generalised efficiency than LSTM units. For LSTM, it allowed the utilization of the resultant gate to control the exposure of the number of memory contents.

$$h_t = o_t tanh(c_t) \tag{7}$$

In Eq. (7), the resultant gate was demonstrated as 0_t is computed as:

$$0_t = \sigma(W_0 \cdot [h_{t-1}, x_t, c_l] + b_0) \tag{8}$$

In Eq. (8), the logistic function was represented as σ. The memory cell c_t was conserved by adding some novel memory and removing (forgetting) present memory:

$$c_t = f_t c_{t-1} + i_\tau \tilde{c} + b_c \tag{9}$$

The $\tilde{c}_t$ novel memory is provided as:

$$\tilde{c} = tanh(W_c \cdot [h_{t-1}, x_t]) \tag{10}$$

The range to remove and add memory is measured by the input gate i_t and forget gate f_t. The forget gate is computed by the following formula:

$$f_t = \sigma\big(W_f \cdot [h_{t-1}, x_t, c_{t-1}] + b_f\big) \tag{11}$$

and i_t is computed as:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t, c_{t-1}] + b_i) \tag{12}$$

In the formula, the matching bias vector was represented as b. Like the LSTM unit, GRU utilizes a gate to influence the data stream inside units, but there are no memory cells. The h_r hidden layer (HL) is a linear combination of novel HLs $\tilde{h}_t$ and earlier HL h_{t-1}:

$$h_t = (1 - z_t)h_{t-1} + z_t \tilde{h}_t \tag{13}$$

In Eq. (13), the upgrade gate z_t control that several its novel activation becomes upgrade. It can be computed as:

$$z_t = \sigma(W_z \cdot [h_{t-1},\ x_t]) \tag{14}$$

The $\tilde{h}_t$ novel activation was computed as:

$$\tilde{h} = tanh(W_h \cdot [r_t \odot h_{t-1},\ x_t]) \tag{15}$$

In Eq. (15), the forget gate r_t is similar to the upgrade unit from the LSTM:

$$r_l = \sigma(W_r \cdot [h_{t-1},\ x_t]) \tag{16}$$

But the standard RNN activities the previous data, the bi-directional RNN (BRNN) procedure data in 2 directions. The y outcome of BRNN can be obtained with measured the $\overrightarrow{h_r}$ forward hidden series and $\overleftarrow{h_t}$ backward series as follows:

$$\vec{h}_t = \Phi\left(WW_{\vec{\chi h}} x_t + W_{\vec{hh}}\vec{h}_{t-1} + b_{\vec{h}}\right) \tag{17}$$

$$\overleftarrow{h_t} = \Phi(W_{\underleftarrow{x h}} x_t xh + W_{\underleftarrow{h}\,\underleftarrow{h}}\overleftarrow{h}_{t-1} + b_{\underleftarrow{h}}) \tag{18}$$

$$y_t = W_{\vec{hy}}\vec{h}_t + W_{\underleftarrow{h y}}\overleftarrow{h}_{t-1} + b_y \tag{19}$$

Combining BRNN with GRU offers BIGRU, which is employed to access the long-term data series in 2 directions. While a fault analysis problem is commonly assumed a classifier problem, cross-entropy was modified as a loss function. The weighted cross entropy was defined as:

$$f(\theta) = -\sum_{n=1}^{N} w_n \sum_{i=1}^{M} y_i log(\hat{y}_i) \tag{20}$$

In Eq. (20), θ stands for the NN parameter, N defines the instance number, the count of faults is demonstrated as M, the true label is referred to as y_i, and the forecasted probability signifies $\hat{y}_i$. Fig. 2 demonstrates the framework of the BiGRU technique.

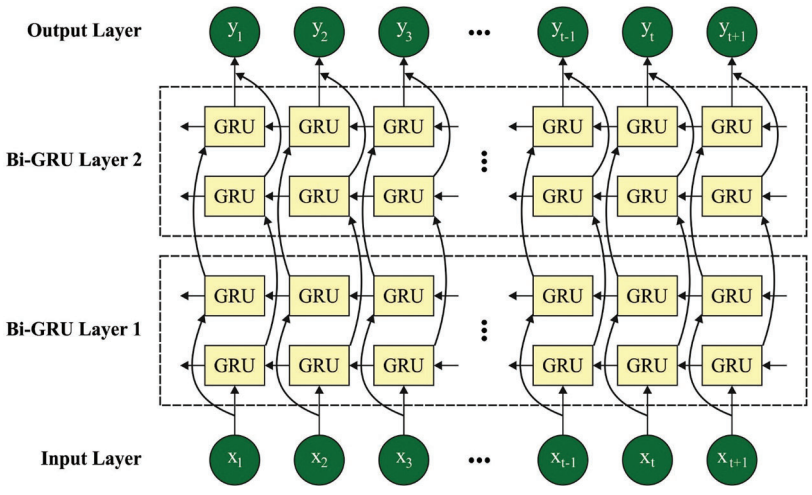

Figure 2: Architecture of BiGRU

4 Results Analysis

The experimental validation of the BODL-HGRSLC model is tested using a dataset comprising different hand gesture images. The parameter settings are learning rate: 0.01, dropout: 0.5, batch size: 5, epoch count: 50, and activation: ReLU. Fig. 3 demonstrates the sample images for alphabets A-Z.

Figure 3: Sample images for alphabets A-Z

The confusion matrix derived by the BODL-HGRSLC model on the entire dataset is depicted in Fig. 4. The figure indicates that the BODL-HGRSLC model has proficiently categorized 26 class labels.

Table 1 and Fig. 5 demonstrate the overall gesture recognition outcomes of the BODL-HGRSLC model. The experimental results demonstrated that the BODL-HGRSLC model had effectively categorized all the hand gestures. For instance, in class 'a', the BODL-HGRSLC model has obtained $accu_y$ of 99.70%, $prec_n$ of 92.92%, $reca_l$ of 99.80%, and F_{score} pf 96.24%. Meanwhile, in class 'b', the BODL-HGRSLC model has obtained $accu_y$ of 99.73%, $prec_n$ of 96.59%, $reca_l$ of 96.40%, and F_{score} pf 96.50%. Eventually, in class 'c', the BODL-HGRSLC model obtained $accu_y$ of 99.75%, $prec_n$ of 97.56%, $reca_l$ of 95.80%, and F_{score} pf 96.67%. Simultaneously, on class 'd', the BODL-HGRSLC model has obtained $accu_y$ of 99.75%, $prec_n$ of 96.80%, $reca_l$ of 96.80%, and F_{score} pf 96.80%. Finally, in class 'e', the BODL-HGRSLC model has obtained $accu_y$ of 99.77%, $prec_n$ of 98.55%, $reca_l$ of 95.40%, and F_{score} pf 96.95%.

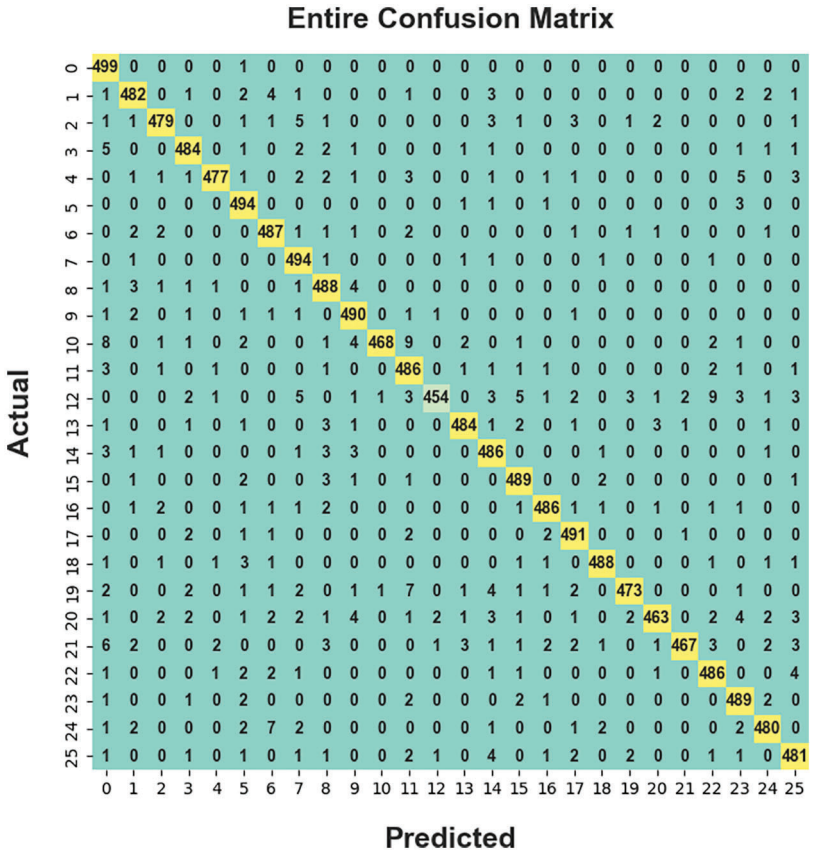

Figure 4: Confusion matrices of the BODL-HGRSLC approach under the Entire dataset

Table 1: Result analysis of BODL-HGRSLC algorithm with distinct class labels under the entire dataset

Entire dataset				
Class labels	Accuracy	Precision	Recall	F-Score
a	99.70	92.92	99.80	96.24
b	99.73	96.59	96.40	96.50
c	99.75	97.56	95.80	96.67
d	99.75	96.80	96.80	96.80
e	99.77	98.55	95.40	96.95
f	99.75	95.00	98.80	96.86
g	99.74	95.87	97.40	96.63
h	99.74	94.64	98.80	96.67
i	99.72	95.13	97.60	96.35
j	99.75	95.70	98.00	96.84
k	99.74	99.57	93.60	96.49
l	99.63	93.46	97.20	95.29

(Continued)

Table 1 (continued)

Entire dataset				
Class labels	Accuracy	Precision	Recall	F-Score
m	99.61	98.91	90.80	94.68
n	99.79	97.78	96.80	97.29
o	99.67	94.37	97.20	95.76
p	99.78	96.45	97.80	97.12
q	99.80	97.59	97.20	97.39
r	99.79	96.46	98.20	97.32
s	99.85	98.39	97.60	97.99
t	99.72	98.13	94.60	96.33
u	99.64	97.89	92.60	95.17
v	99.72	99.15	93.40	96.19
w	99.72	95.67	97.20	96.43
x	99.72	95.14	97.80	96.45
y	99.74	97.17	96.00	96.58
z	99.68	95.63	96.20	95.91
Average	**99.73**	**96.56**	**96.50**	**96.50**

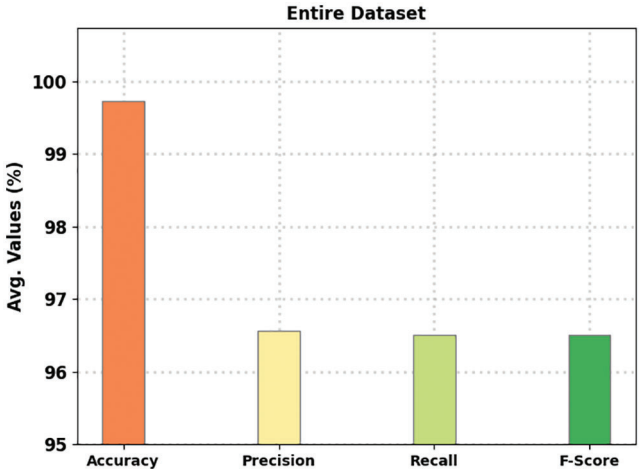

Figure 5: Average analysis of the BODL-HGRSLC algorithm under the entire dataset

Table 2 and Fig. 7 illustrate the complete gesture recognition outcomes of the BODL-HGRSLC method under 70% of TR dataset. The experimental outcomes manifested by the BODL-HGRSLC approach have effectively categorized all the hand gestures. For example, in class 'a', the BODL-HGRSLC technique has acquired $accu_y$ of 99.69%, $prec_n$ of 92.78%, $reca_l$ of 99.71%, and F_{score} pf 96.12%. In the meantime, in class 'b', the BODL-HGRSLC approach has achieved $accu_y$ of 99.75%, $prec_n$ of 96.37%, $reca_l$ of 96.67%, and F_{score} pf 96.52%. Parellerly, on class 'c', the BODL-HGRSLC method has attained $accu_y$ of

99.71%, $prec_n$ 96.70%, $reca_l$ of 95.55%, and F_{score} pf 96.12%. At the same time, on class 'd', the BODL-HGRSLC algorithm has gained $accu_y$ of 99.75%, $prec_n$ of 96.31%, $reca_l$ of 97.13%, and F_{score} pf 96.72%. At last, in-class 'e', the BODL-HGRSLC technique has reached $accu_y$ of 99.74%, $prec_n$ of 98.626%, $reca_l$ of 94.94%, and F_{score} pf 96.94%.

Table 2: Result analysis of BODL-HGRSLC algorithm with distinct class labels under 70% of TR dataset

Training phase (70%)				
Class labels	Accuracy	Precision	Recall	F-Score
a	99.69	92.78	99.71	96.12
b	99.75	96.37	96.67	96.52
c	99.71	96.70	95.55	96.12
d	99.75	96.31	97.13	96.72
e	99.74	98.26	94.94	96.57
f	99.73	94.72	98.63	96.64
g	99.75	96.03	97.41	96.72
h	99.68	93.32	98.87	96.01
i	99.75	96.08	97.44	96.76
j	99.76	95.28	98.18	96.71
k	99.69	99.41	92.84	96.01
l	99.64	93.81	97.59	95.66
m	99.54	98.49	89.84	93.97
n	99.77	98.08	96.23	97.14
o	99.69	94.64	96.95	95.78
p	99.81	97.09	97.95	97.52
q	99.80	97.44	97.44	97.44
r	99.76	96.33	97.43	96.88
s	99.85	98.57	97.46	98.01
t	99.71	98.18	94.19	96.14
u	99.69	97.80	93.67	95.69
v	99.74	98.77	94.13	96.40
w	99.70	95.44	97.27	96.35
x	99.71	95.51	97.14	96.32
y	99.71	97.35	95.11	96.22
z	99.69	95.32	96.92	96.11
Average	**99.72**	**96.46**	**96.41**	**96.40**

The confusion matrix derived by the BODL-HGRSLC method on 70% of TR dataset is shown in Fig. 6. The figure represents the BODL-HGRSLC algorithm has effectively classified 26 class labels.

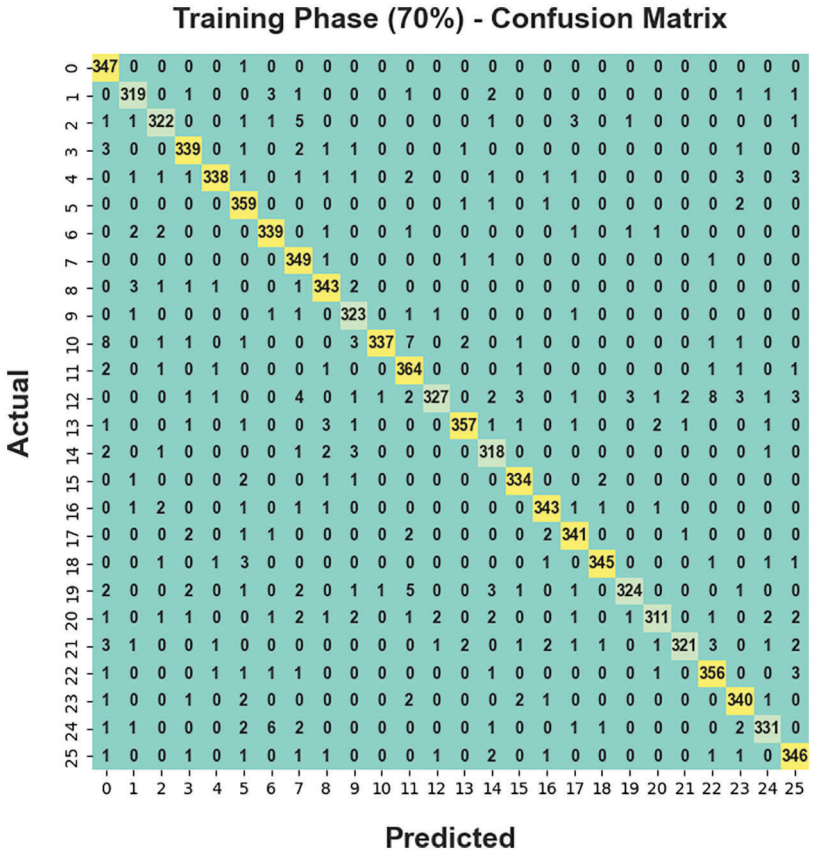

Figure 6: Confusion matrices of BODL-HGRSLC approach under 70% of TR dataset

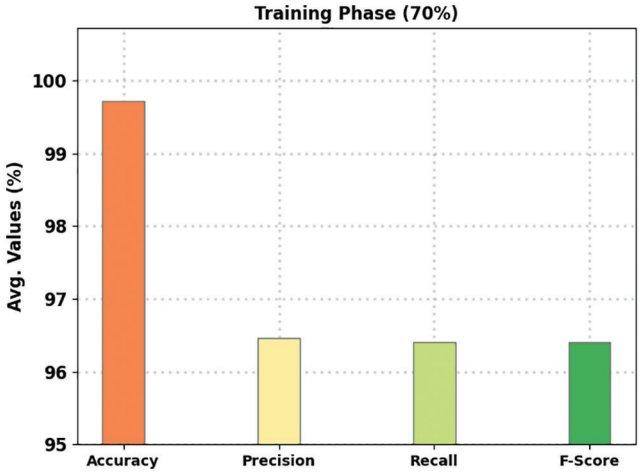

Figure 7: Average analysis of the BODL-HGRSLC algorithm under 70% of TR dataset

The confusion matrix derived by the BODL-HGRSLC approach on 30% of TS dataset is displayed in Fig. 8. The figure denotes the BODL-HGRSLC methodology has proficiently categorized 26 class labels.

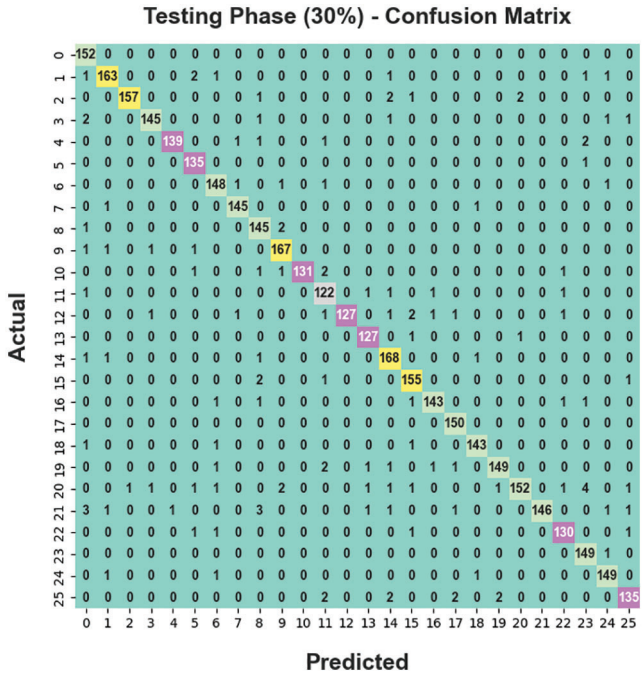

Figure 8: Confusion matrices of BODL-HGRSLC approach under 30% of TS dataset

Table 3 and Fig. 9 portray the overall gesture recognition outcomes of the BODL-HGRSLC model under 30% of TS dataset. The experimental outcomes of the BODL-HGRSLC method have effectively categorized all the hand gestures. For example, in class 'a', the BODL-HGRSLC approach has reached $accu_y$ of 99.72%, $prec_n$ of 93.25%, $reca_l$ of 100%, and F_{score} pf 96.51%. In the meantime, in class 'b', the BODL-HGRSLC technique has reached $accu_y$ of 99.75%, $prec_n$ of 96.37%, $reca_l$ of 96.67%, and F_{score} pf 96.52%. In Parallel, on class 'c', the BODL-HGRSLC method has attained $accu_y$ of 99.71%, $prec_n$ 96.70%, $reca_l$ of 95.55%, and F_{score} pf 96.12%. Concurrently, on class 'd', the BODL-HGRSLC approach has attained $accu_y$ of 99.75%, $prec_n$ of 96.31%, $reca_l$ of 97.13%, and F_{score} pf 96.72%. At last, in class 'e', the BODL-HGRSLC algorithm has acquired $accu_y$ of 99.74%, $prec_n$ of 98.626%, $reca_l$ of 94.94%, and F_{score} pf 96.57%.

Table 3: Result analysis of BODL-HGRSLC algorithm with distinct class labels under 30% of TS dataset

Testing phase (30%)				
Class labels	Accuracy	Precision	Recall	F-Score
a	99.72	93.25	100.00	96.51
b	99.69	97.02	95.88	96.45
c	99.82	99.37	96.32	97.82
d	99.77	97.97	96.03	96.99
e	99.85	99.29	96.53	97.89
f	99.82	95.74	99.26	97.47
g	99.72	95.48	97.37	96.42
h	99.87	97.97	98.64	98.31

(Continued)

Table 3 (continued)

Testing phase (30%)				
Class labels	Accuracy	Precision	Recall	F-Score
i	99.64	92.95	97.97	95.39
j	99.74	96.53	97.66	97.09
k	99.85	100.00	95.62	97.76
l	99.62	92.42	96.06	94.21
m	99.77	100.00	93.38	96.58
n	99.85	96.95	98.45	97.69
o	99.62	93.85	97.67	95.73
p	99.69	95.09	97.48	96.27
q	99.79	97.95	96.62	97.28
r	99.87	96.77	100.00	98.36
s	99.85	97.95	97.95	97.95
t	99.74	98.03	95.51	96.75
u	99.51	98.06	90.48	94.12
v	99.67	100.00	91.82	95.74
w	99.77	96.30	97.01	96.65
x	99.74	94.30	99.33	96.75
y	99.79	96.75	98.03	97.39
z	99.67	96.43	94.41	95.41
Average	**99.75**	**96.79**	**96.75**	**96.73**

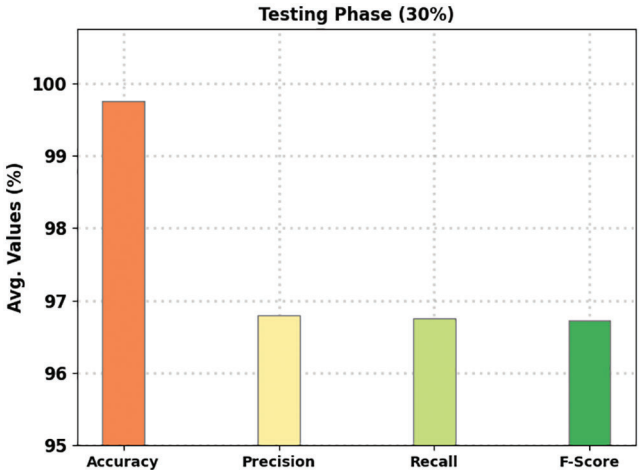

Figure 9: Average analysis of BODL-HGRSLC algorithm under 30% of TS dataset

The training accuracy (TRA) and validation accuracy (VLA) acquired by the BODL-HGRSLC approach on the test dataset is demonstrated in Fig. 10. The experimental result exhibited the BODL-HGRSLC method has attained maximal values of TRA and VLA. Seemingly the VLA is greater than TRA.

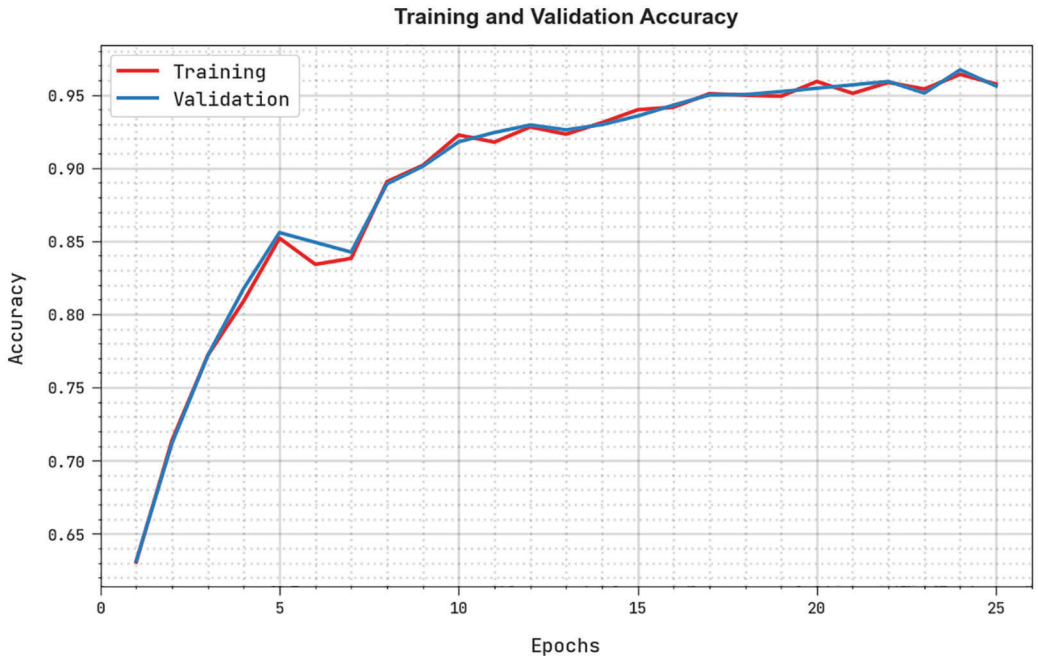

Figure 10: TRA and VLA analysis of the BODL-HGRSLC algorithm

The training loss (TRL) and validation loss (VLL) obtained by the BODL-HGRSLC algorithm on the test dataset are displayed in Fig. 11. The experimental outcome denoted the BODL-HGRSLC algorithm has outperformed minimal values of TRL and VLL. Particularly, the VLL is lesser than TRL.

Figure 11: TRL and VLL analysis of the BODL-HGRSLC algorithm

A comprehensive $accu_y$ examination of the BODL-HGRSLC with recent models is given in Table 4 and Fig. 12 [21,22]. The results indicated that the k-NN and AlexNet models have failed to show effectual results. Following, the NB and CNN models have attained slightly enhanced performance. In line with this, the ANN and LSTM models have reached reasonable outcomes.

Table 4: Comparative analysis of BODL-HGRSLC approach with recent algorithms

Methods	Accuracy
BODL-HGRSLC	99.75
Hypertuned DCNN	99.62
ANN model	93.08
k-NN model	82.31
NB model	90.86
CNN model	91.17
AlexNet model	84.92
LSTM model	95.48

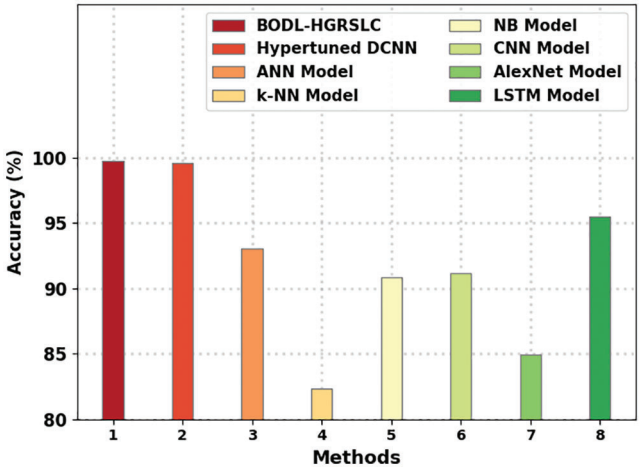

Figure 12: Comparative analysis of BODL-HGRSLC approach with recent algorithms

Though the hyperparameter-tuned DCNN model has gained near optimal $accu_y$ of 99.62%, the presented BODL-HGRSLC model has reached a higher $accu_y$ of 99.75%. Therefore, the experimental results confirmed the enhanced outcomes of the BODL-HGRSLC with a higher $accu_y$ of 99.75%. Thus, the BODL-HGRSLC can be employed as a productive approach for gesture recognition.

5 Conclusion

This article develops a new BODL-HGRSLC technique to recognize the HGs for disabled people's communication. The presented BODL-HGRSLC technique integrates the concepts of CV and DL models. In the presented BODL-HGRSLC technique, the ResNet model is applied for feature extraction. Besides, the presented BODL-HGRSLC model uses Bayesian optimization for the hyperparameter tuning process. At last, the BiGRU model is exploited for the HGR procedure. A wide range of experiments was

conducted to demonstrate the enhanced performance of the presented BODL-HGRSLC model. The comprehensive comparison study reported the improvements of the BODL-HGRSLC model over other DL models with maximum accuracy of 99.75%. In the future, an ensemble fusion of DL models can be derived to enhance the recognition outcomes.

Funding Statement: The authors extend their appreciation to the King Salman centre for Disability Research for funding this work through Research Group no KSRG-2022-017.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] L. Chen, J. Fu, Y. Wu, H. Li and B. Zheng, "Hand gesture recognition using compact CNN via surface electromyography signals," *Sensors*, vol. 20, no. 3, pp. 672, 2020.

[2] A. Moin, A. Zhou, A. Rahimi, A. Menon, S. Benatti *et al.,* "A wearable biosensing system with in-sensor adaptive machine learning for hand gesture recognition," *Nature Electronics*, vol. 4, no. 1, pp. 54–63, 2021.

[3] S. Z. Gurbuz and M. G. Amin, "Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring," *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 16–28, 2019.

[4] P. Wang, H. Liu, L. Wang and R. X. Gao, "Deep learning-based human motion recognition for predictive context-aware human-robot collaboration," *CIRP Annals*, vol. 67, no. 1, pp. 17–20, 2018.

[5] F. Wen, Z. Sun, T. He, Q. Shi, M. Zhu *et al.,* "Machine learning glove using self-powered conductive superhydrophobic triboelectric textile for gesture recognition in VR/AR applications," *Advanced Science*, vol. 7, no. 14, pp. 2000261, 2020.

[6] U. Côté-Allard, C. L. Fall, A. Drouin, A. C. Lecours, C. Gosselin *et al.,* "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 4, pp. 760–771, 2019.

[7] D. S. Breland, A. Dayal, A. Jha, P. K. Yalavarthy, O. J. Pandey *et al.,* "Robust hand gestures recognition using a deep CNN and thermal images," *IEEE Sensors Journal*, vol. 21, no. 23, pp. 26602–26614, 2021.

[8] D. Tasmere, B. Ahmed and S. R. Das, "Real time hand gesture recognition in depth image using CNN," *International Journal of Computer Applications*, vol. 174, no. 16, pp. 28–32, 2021.

[9] C. S. Liang and H. Li-Wu, "Using deep learning technology to realize the automatic control program of robot arm based on hand gesture recognition," *International Journal of Engineering and Technology Innovation*, vol. 11, no. 4, pp. 241, 2021.

[10] A. Dayal, N. Paluru, L. R. Cenkeramaddi and P. K. Yalavarthy, "Design and implementation of deep learning based contactless authentication system using hand gestures," *Electronics*, vol. 10, no. 2, pp. 182, 2021.

[11] Z. Su, H. Liu, J. Qian, Z. Zhang and L. Zhang, "Hand gesture recognition based on semg signal and convolutional neural network," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 35, no. 11, pp. 2151012, 2021.

[12] H. Wang, Y. Zhang, C. Liu and H. Liu, "sEMG based hand gesture recognition with deformable convolutional network," *International Journal of Machine Learning and Cybernetics*, vol. 13, no. 6, pp. 1729–1738, 2022.

[13] A. Mujahid, M. J. Awan, A. Yasin, M. A. Mohammed, R. Damaševičius *et al.,* "Real-time hand gesture recognition based on deep learning YOLOv3 model," *Applied Sciences*, vol. 11, no. 9, pp. 4164, 2021.

[14] W. K. Wong, F. H. Juwono and B. T. T. Khoo, "Multi-features capacitive hand gesture recognition sensor: A machine learning approach," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 8441–8450, 2021.

[15] E. V. Añazco, S. J. Han, K. Kim, P. R. Lopez, T. S. Kim *et al.,* "Hand gesture recognition using single patchable six-axis inertial measurement unit via recurrent neural networks," *Sensors*, vol. 21, no. 4, pp. 1404, 2021.

[16] Z. Yang and X. Zheng, "Hand gesture recognition based on trajectories features and computation-efficient reused LSTM network," *IEEE Sensors Journal*, vol. 21, no. 15, pp. 16945–16960, 2021.

[17] N. Nasri, S. Orts-Escolano and M. Cazorla, "An semg-controlled 3D game for rehabilitation therapies: Real-time time hand gesture recognition using deep learning techniques," *Sensors*, vol. 20, no. 22, pp. 6451, 2020.

[18] M. Hammad, P. Pławiak, K. Wang and U. R. Acharya, "ResNet-attention model for human authentication using ECG signals," *Expert Systems*, vol. 38, no. 6, pp. e12547, 2021.

[19] X. B. Jin, W. Z. Zheng, J. L. Kong, X. Y. Wang, Y. T. Bai *et al.,* "Deep-learning forecasting method for electric power load via attention-based encoder-decoder with Bayesian optimization," *Energies*, vol. 14, no. 6, pp. 1596, 2021.

[20] Q. Zhou, C. Zhou and X. Wang, "Stock prediction based on bidirectional gated recurrent unit with convolutional neural network and feature selection," *Plos One*, vol. 17, no. 2, pp. e0262501, 2022.

[21] A. Mannan, A. Abbasi, A. Javed, A. Ahsan, T. Gadekallu *et al.,* "Hypertuned deep convolutional neural network for sign language recognition," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–10, 2022.

[22] A. Rahagiyanto, A. Basuki, R. Sigit, A. Anwar and M. Zikky, "Hand gesture classification for sign language using artificial neural network," in *2017 21st Int. Computer Science and Engineering Conf. (ICSEC)*, Bangkok, Thailand, pp. 1–5, 2017.