International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

IJISAE

ISSN:2147-6799 www.ijisae.org Original Research Paper

Unveiling Gesture Language: Advancements in Deep Backpropagation Neural Networks for Image-Based Sign Language Recognition

G.K. Vaidhya*[1], Dr. G. Paavai Anand[2]

Submitted: 26/01/2024 **Revised**: 04/03/2024 **Accepted**: 12/03/2024

Abstract: A defined set of languages called sign language (SL) uses both manual and visual methods to communicate information. It is frequently used to communicate with non-verbal individuals. The majority of people can only comprehend if they acquire knowledge of universal sign language. Thus, not everyone is unable to understand when we use sign language to communicate with hearing-impaired individuals. It is difficult to communicate with peoplewithout a translator. The prior approach to this problem concentrated on reading sign language. However, because it takes longer and providesincorrect precision performance, it is difficult to detect sign language in it. To solve this issue, this book offers an improved solution based on deep learning methods. For Sign Language Identification (SLI), we introduce a Deep Backpropagation Neural Network (DBNNet) method using a softmaxactivation function. First, collect the ASL dataset of SL images. Furthermore, we use the GaussianSmoothing Histogram Filter (GMHF) method to improve contrast and image quality. Additionally, based on an enhanced sign image, the Intensity Gradients Sign Edge Detection (IGSED) method locates the edges. Next, we use a Cluster-Based Watershed Segmentation (CBWS) algorithm to examine the region of interest (ROI) for a sign. Subsequently, the DBNNet method with the softmax function is suggested to efficiently identify the sign. As a result, the suggested algorithm outperformedearlier techniques in terms of sign identification accuracy, sensitivity, specificity, and F-measure performance.

Keywords: *Sign Language Identification (SLI), pre- processing, segmentation, histogram, impaired people, and deep learning.*

1. Introduction

Visible gestures and signs are used in sign language (SL) [1].Gestures, which can be made with the body, hands, arms, faces, or lips, are a type of non-verbal communication that take the place of spoken communication [2].It includes about 300 different types of SL depending on the natural language process (NLP) and culture of each country region [3]. Deaf people use sign language (SL), which is difficult for those with visual impairments to understand.

A hand sign language example is shown in Figure 1. Speaking with people is made simpler when a translator is present. The prior approach to this problem concentrated on reading sign language. However, because it takes longer and provides incorrect precision performance, it is difficult to detect sign language in it. This article employs a deep learning-based method known as Deep Backpropagation Neural Network (DBNNet) to address this proposal. The primary contribution of this paper is increasing the accuracy of SL identification through the use of segmentation and classification techniques.

[1] *SRM Institute of Science and Technology, Vadapalani, Chennai – 600026, INDIA*
ORCID ID: 0009-0007-6488-887X
[2] *SRM Institute of Science and Technology, Vadapalani, Chennai – 600026, INDIA*
ORCID ID : 0000-0002-2574-1548
** Corresponding Author Email: vaidhyashankar123@gmail.com*

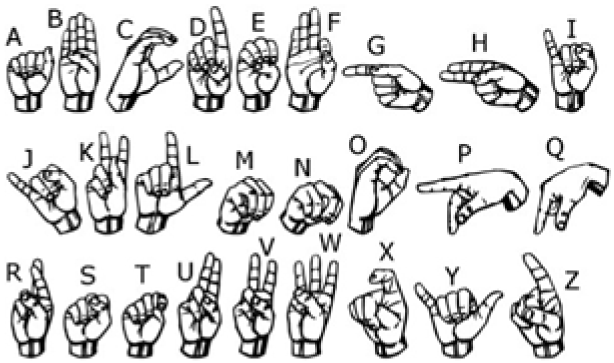


Fig. 1. Example of sign language (SL)

2. Summary of this Paper

The remainder of this article, chapter 2, is devoted to current assessments of sign detection. The suggested methodology for sign detection is examined in Chapter 3. The analysis of the experimental results is explained in Chapter 4, and Chapter 5 presents the conclusion.

3. Related works

American Sign Language is primarily some sort of default reorganization problem. Universal application is hampered by similarity and sign trajectory problems. The confused sign transaction algorithm is known as ALS. Quick support has been extended to the fast fisher vector (FFV) and bi-directional Long-Short Term Memory (Bi-LSTM) methods, which are used for quick memory processing and simple sign and signature verification recognition. The system is readily identifiable by the motion detection-based evaluation system [4].

American Sign Language (ASL) to enhance the verification of signs and signatures. To increase user security, voice-based recognition has been the foundation for sign verification. The best micro Dupler is ASL, and motion detection ought to be used to identify a safety situation correctly [5].

Motion detection and recognition's default task is dynamic gesture recognition. In order to enable quick motion detection and identification in the security system, real-time dynamic non-repetitive gestures are introduced as a solution to the gesture detection problem. The user identification systems have been improved as a result [6].

Sign Language (SL) videos are automatically generated to test the faction recognition. Based on moving and running video motion detection labels and credit of the cyber security system, the SL-based face identification system operates. Certain disadvantages of recognition accuracy [7].

Motion and facial recognition using signals and surface electromyography (sEMG) in the security system. Gestures should be used with signal and surface electromyography (sEMG) in order to achieve the highest accuracy possible for fast and dynamic user identification with security systems. [8].

Gesture Detection to Sign Language Classification is currently the most widely used machine learning technique. High-level freedom techniques are the crucial step to use there. Businesses use sensors and microcontrollers to classify data through usage. A few defaults happen while classifying. [9].

As a quick and safe way to identify signs and gesture patterns, a new dynamic hand gesture recognition algorithm (DGDL-GR) is suggested. Check patterns and user identities dynamically, then use gestures in the future to swiftly identify security systems. There, human identification is quickly made [10].

Our ideas are not empty words; sign language is an art that changes. As a crucial idea in communication, SL stands for Deaf/Mute Community. For the 90 million deaf and mute people who live in the world, sign language is vital. Translation and sign language recognition (SLR) are therefore rich fields for research. It uses one or more cameras to take a series of pictures. After the signature is completed, image processing is used to identify the signature. Instrumented gloves with sensors are used in a sensor-based technique to expose the hands [11].

Gestures have been the most basic form of communication since humans first learned to speak. Nonverbal communication is still crucial, though. These hand gestures have a variety of applications, such as games, human-computer interaction, and life development [12].

A text-to-speech mobile application for Android has been developed that translates text into speech that can be heard without the pressure sensor, which is missing from the proposed system [13]. As demonstrated by our experiments, the proposed LSTM algorithm performed exceptionally well in classification, which could benefit the ongoing research in HCI and hand gesture recognition [14]. Computer vision is a very active study area in traditional vision-based methods. Additionally, we have compiled a fairly extensive database of ten hand movement sequences from ten participants, including five static gestures and five dynamic gestures.

This paper proposes a glove that measures pressure at 1052 points on the human hand, based on a twig-shaped tactile sensor sheet. The impact of the suggested glove design on the variation in joint movement that takes place when the glove is worn is assessed through an experiment. As a final, preliminary conclusion, the measured data for everyday items are presented. The robot's skill and strategy will be enhanced by a glove designed for human talent evaluation [15].

This paper presents a method for translating Mexican Sign Language into Spoken Spanish. Two major challenges are developing continuous sign problem solving methods and signer-independent methods. No particular clothing, gloves, or accessories are needed in order to work. [16].

Several deep learning algorithms were employed in the novel [17] to recognize dynamic sign language. Both local and global features of the SL image were examined by the suggested method. To create an effective recognition system, however, issues with hand segmentation, local hand shape representation, global body structure modeling, and gesture flow modeling must be resolved.

The first two fundamental models are constructed with this issue in mind. One is the Conv EMG model, which has a deeply separable convolution and an initial block that is densely connected. Subsequently, the fundamental models are enhanced through the application of a multi-stream fusion technique that leverages the relationship between muscles and gestures as well as the advantages of complementary models [18].

The hand gesture recognition process for proper progress and future directives was covered in the review [19]. Additionally, the accuracy of term identification for vision-based hand gesture recognition is found in this review.

Smaller performance gains were obtained from higher-order synergies, suggesting a trade-off between accuracy and control complexity. The data presented here is used to determine how many synergies to use in control, taking into account the accuracy of the controlled robot system and the precision of the controller. The resultant system achieved high capture efficiency with minimal computational or manual effort from assist devices [20].

Every application can benefit from the lightweight and expanding personal identification and user verification features of the automatic signature verification system. This document employs date- wise user signature verification. User-provided date- wise guarantees are more secure identification advanced enhancements [21]. [22] developed Optimized Keyframe-Centered Clips (OKCC) for sign language recognition. In order to assign significance to temporal and spatial attributes extracted from skeletal data, this study also used attention-based networks.

The signature identification model framework approach uses the sigma-lognormal model; gesture is presented in the framework along with handwritten signatures. When it comes to user identification security, the sigma-lognormal model performs exceptionally well when compared to different datasets, natural language, and multiple author identifiers [23].

A major barrier to statistical analysis and pattern recognition in automated signature verification is the small number of reference signatures per user. The difficulty of normalizing scores is due to the absence of information on intra- user variability. In this paper, we propose a novel two-stage normalization process that first detects simple fakes and then handles more sophisticated fakes in the second stage. Furthermore, it talks about different approaches to normalize dynamic time warp scores [24].

Convolutional Block Attention Module (CBAM) algorithm for dynamic sign recognition was presented by the author [25]. In order to obtain feature vectors with extremely small dimensions, the model first created a pre-trained convolutional auto- encoder network. In order to accelerate the training process and improve the network's inference, it makes use of BN networks.

The article focused on the use of the boundary adaptive encoder method for Chinese SL recognition [26]. In the decoding step, a position- based windowed attention model is employed to improve the modeling performance of lengthy sign language sequences. This makes it possible to create weighting factors that are more effective.

In order to improve the speed of the authentication process, the study makes use of a multi-gesture user authentication system, a touch sensor system, a multi-gesture touching algorithm, and a well-defined user authentication system for the multi-gesture touching algorithm [27].

The transformer Encoder tool for recognizing Indian sign language (ISL) was introduced in the novel [28]. The method proposed by the author outperformed other cutting-edge convolutional structures in standard ISL recognition.

Furthermore, using their IISL2020 dataset, the paper introduced a DL-based method for ISL recognition, similar to LSTM and GRU [29]. However, the accuracy performance for sign identification is lower with this method.

The preceding issue of sign identification was the main topic of the article. Consequently, the Deep CNN algorithm for ASL identification was introduced in the novel [30]. The assigned image size improved the Deep CNN model's performance. They used data augmentation techniques to purposefully increase the size of the training data set from the available data. ASLs are difficult to distinguish due to similarities within their class.

In order to identify Arabic sign language, the study used a transfer learning algorithm [31]. In order to interpret visual data with text, a structure based on Arabic Sign Language is created to take an image dataset of visual motions.

In a similar vein, the study employed a standard ASL dataset and a Resnet-based DNN algorithm for SL recognition [32]. Several augmentation strategies are first applied to the dataset. The input image is divided into four classes using the ResNet50's stage 1 method, which is the next stage approach. An image is fed as input to the corresponding second-level model for hand gesture prediction once it has been classified into one of the sets.

Our suggested Intensity Gradients Sign Edge Detection (IGSED) technique is used in the second step to identify the sign's edges. In order to effectively identify the pixels that represent the hand SL regions using the Cluster-Based Watershed Segmentation (CBWS) approach, segmentation is a necessary second step. Following the ROI of the hand from the SL image, the wrist line is traced in order to extract the forearm portion of the SL hand, which is then used to shape the correct hand section. In order to efficiently generate the output classes of various characters, we employ DBNNet to automatically detect the invariant features from segmented depth sensing images.

SL recognition using a hierarchical sign learning algorithm created by [33]. By concatenating the primary postures of the landmarks back and forth to affect the overall video movement frames, the suggested approach preserves the spatiotemporal information of the landmarks.

Using the Pakistan SL dataset, Support Vector Machine (SVM) and the MKL algorithm were developed by [34] for SL identification. Four vision- based elements are extracted in the following steps: edge orientations, directional gradients, local binary pattern, and essential components.

Similarly, a hybrid deep neural method for Russian and

Indian sign recognition was presented in the paper [35]. Furthermore, a 3D-based DNN with convolution was used in the study to extract spatial elements from hand gestures. The vision-based recognition frameworks, however, only took into account a small number of the spatial details required for accurate sign recognition.

4. Proposed methodology for sign language identification

The suggested deep learning-based method used for Sign Language Identification (SLI) is demonstrated in this section. Three steps make up the suggested method for identifying signs: edge detection, segmentation, and classification; the first stage involves pre-processing the sign image using histogram equalization. The DL technique's structure for sign language identification (SLI) is depicted in Figure 2.The first step in SLI is pre-processing, which uses the Gaussian Smoothing Histogram Filter (GMHF) technique to reduce noise, boost brightness, and resize the image to improve image quality.

Our suggested Intensity Gradients Sign Edge Detection (IGSED) technique is used in the second step to identify the sign's edges. In order to effectively identify the pixels that represent the hand SL regions using the Cluster-Based Watershed Segmentation (CBWS) approach, segmentation is a necessary second step. Following the ROI of the hand from the SL image, the wrist line is traced in order to extract the forearm portion of the SL hand, which is then used to shape the correct hand section. In order to efficiently generate the output classes of various characters, we employ DBNNet to automatically detect the invariant features from segmented depth sensing images.

4.1. Gaussian Smoothing Histogram Filter (GSHF)

Using the Gaussian Smoothing Histogram Filter (GSHF) technique, this stage involves pre- processing the images to remove noise, brightness, and scaling individually from the gathered SL dataset. Precise undesired background details are removed from the input SL image. Using a Gaussian smoothing filter, noise deduction entails removing noise and discontinuities. This method uses a smoothing process to get rid of the noise. The process of adjusting and normalizing the processed image's brightness and contrast is known as histogram equalization.

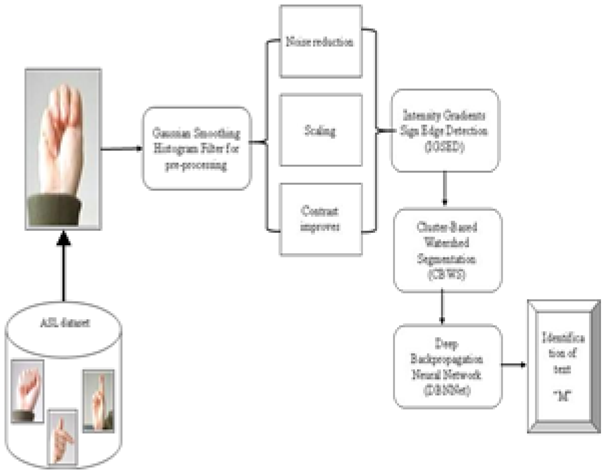


Fig. 2. SLI identification based on deep learning

$$Mean\ M^n = \sum_X T_{k,l} \tag{1}$$

Based on the total number of pixels in the input sign image for background removal, equation 1 estimates the mean in the input image T coordinates pixel (k, l).

$$N^{Reduction}(k,l) = \sum_X M^n\{T(k,l)\} \tag{2}$$

The input sign image's noise is removed by expression 2. Next, we compute the smoothing image that is determined by equation 3.

$$G_{Smooth}(k,l) = N^{Reduction}\left[Exp\left(\frac{k'^2 + A^r l'^2}{2} + \psi\right)\right] \tag{3}$$

In this case, it is assumed that the noise reduction sign image's aspect ratio and central frequency correspond to a smooth process.

$$k' = kcos\theta + lsin\theta \tag{4}$$

$$l' = -ksin\theta + lcos\theta \tag{5}$$

Assuming that θ refers to the angle of sign language, it is expressed by equation 6.

$$\theta = \frac{\pi}{d_c} * (n-1) \quad n = 1,2,3 \ldots d_c,\ d_c \in n \tag{6}$$

Where d_c refers to the amount of direction in the input SL image, the below equation estimates the aspect ratio $A^{r.}$

$$A^r = \frac{M^n(k)}{M^n(l)} \tag{7}$$

$$Standard\ \text{deviation}\sigma(\mathrm{k,l}) = \frac{1}{X}\left[\sum_{k,l=1}^{X}(T(k,l) - G_{Smooth}(k,l)^2)\right]^{1/2} \tag{8}$$

The above equation 8 finds the color variation in the sign language.

$$B_H(k,l) = \frac{1}{X}\sum_{k,l=1}^{X}\sigma(\mathrm{k,l}) \tag{9}$$

Based on the smooth Gaussian image and the total number of pixels X, Expression 9 increase the brightness of the sign image.

$$HC_{Range} = Maximum\left(B_H(k,l)\right) - minimum\left(B_H(k,l)\right) \quad (10)$$

Determine the input image's minimum and maximum brightness image and histogram color range HC_{Range}using expression 10.

$$P_{enhance} = P = \sum_{k,l}^{X} HC_{Range}(k-l)^2 \quad (11)$$

From the input image, the equation is used to produce a pre -processed sign image with improved contrast. The suggested Filters effectively remove noise and improve image contrast.

4.2. Intensity Gradients Sign Edge Detection (IGSED)

The preprocessed sign image with enhanced contrast is obtained from the input image through the use of an equation. Image contrast is enhanced and noise is effectively removed by the suggested filters.

-1	0	1
-2	0	2
-1	0	1

Fig. 3a. X – axis Kernel representation

-1	-2	-1
0	0	0
1	2	1

Fig. 3b. Y – axis Kernel representation

Figure 3 depicts the a) X-axis kernel representation and b) y-axis kernel representation in the sign image. This proposed method finds gradient intensity at pixel coordinates points k, l for a processed image P

$$\frac{\partial P}{\partial k} = P(k+1,l) - P(k-1,l) \quad (12)$$

$$\frac{\partial P}{\partial l} = P(k,l+1) - P(k,l-1) \quad (13)$$

Sign image X-axis kernel and Y-axis kernel gradient estimation can be defined by equations 14 and 15:

$$S_X = h_X * P(k,l) \quad (14)$$

$$S_Y = h_Y * P(k,l) \quad (15)$$

Let assuming that S_X denotes horizontal direction and S_Y refers to vertical direction on the enhanced image P.

$$A_G = tan^{-1}\left(\frac{S_X}{S_Y}\right) \quad (16)$$

Equation 16 is used to find the hand sign angleA_G.

$$E_{detection} = \sqrt{{S_X}^2 + {S_Y}^2} \quad (17)$$

To estimate edge detection $E_{detection}$ i in sign language, use equation 17 above. For sign edge detection, the Intensity Gradients Sign Edge Detection (IGSED) algorithm has been proposed. In order to produce clean edge detection, this method seeks to extract the image's salient edge features.

4.3. Cluster-Based Watershed Segmentation (CBWS)

After edge detection, the Cluster-Based Watershed Segmentation (CBWS) algorithm is used in this section to estimate the sign language ROI. The recommended technique is dividing a digital image into regions or clusters of pixels. a proposed pixel-by-pixel CBWS method for image segmentation. After the image has been segmented, a set of regions that contain the entire image are created and numerous contours are extracted from the image. By using edge detection$E_{detection}$, this suggested technique mines the sign object boundary. Equation 18 represents image.,

$$D^E = E_{detection} - (E_{detection} \ominus S_{Element}) \quad (18)$$

Assuming that D^E refers to boundary extraction, $\ominus$symmetric difference, and $S_{Element}$ denotes sign structure element.

$$k = \frac{\sum_X ki}{D^E} \quad (19)$$

$$l = \frac{\sum_X li}{D^E} \quad (20)$$

Equations 19 and 20 estimate centroid points of every boundary imageD^E.

$$s^s = \sum_{k,l=1}^{X} (k-l)^3 D^E(k,l) \quad (21)$$

From equation is used to hand sign skewness s^S based on boundary extraction imageD^E.

$$E_{ccen}(k,l) = \left\{\left(\frac{s^s majoraxis}{2}\right)^2 - \left(\frac{s^s minoraxis}{2}\right)^2\right\}^{1/2} \quad (22)$$

The above equation can be used to find eccentricity based on the major and minor axes in the hand sign image. The ROI of sign language is effectively analyzed by this algorithm by utilizing the CBWS technique.

4.4. Deep Backpropagation Neural Network (DBNNet)

This section applies our proposed Deep Backpropagation Neural Network (DBNNet) algorithm to identify hand signs based on segmentation imagesE_{ccen}. This method accurately classifies the segmentation image for text-to-sign identification. The segment module receives the resized image as input. The output of this stage is appropriately classified as text. A thorough description of the proposed DBNNet method is given in Figure 4.

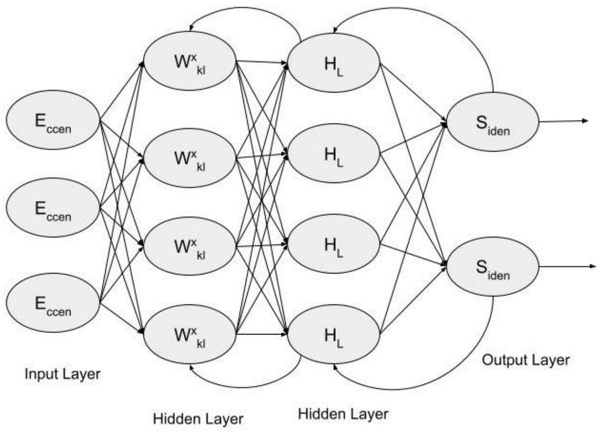


Fig.4. Process of DBNNet Approach

The obtained segmentation sign input is first taken and passed through an input layer in this classifier. The features in the first layer image are extracted by this neuron.

$$I_{Layer} = \sigma\left(\sum_{k,l=0} E_{ccen}(k,l) * K_F + bias\right) \quad (23)$$

Assuming that σ is soft-max activation function, K_F input layer kernel filter and bias. Here we calculate the σ soft-max activation function designed by equation 24.

$$Softmax\ activation\ (\sigma) = \frac{Exp^{in}}{\sum_{n}^{cl} Exp^{ot}} \quad (24)$$

Let's assume that in denotes the exponential function for input, Ot presents the exponential function for sign output, and cl is several classes. The next layer is the hidden layer, which decreases the number of elements when the sign image is enormous.

$$W_{kl}^{X} = \{1 \quad if\left(IoU > 0.5\ among\ default\ true\ box\ on\ class\ (cl)\right) \quad (25)$$

The above equation is to calculate sign weight values from the Input layer. Let us assume IoU denotes the interest of the unit.

$$A_e(\alpha,\beta) = \sum_{k,l=0}^{256} W_{kl(\alpha,\beta)}^{X} \quad (26)$$

The above equation estimates the density of the sign angles $\alpha\ and\ \beta$ for each axis.

$$H_L = \sigma \sum_{k,l=0}^{256} I_{Layer} * A_e(\alpha,\beta) + bias \quad (27)$$

The output is converted to a vector and sent to an output layer after being fed through a number of hidden layers. To create a sign-to-text format, the Softmax layer is the last layer.

$$M_{den1} = \underset{\gamma,\beta}{\arg max}\ H_L(\alpha,\beta) \quad (28)$$

$$M_{den2} = \underset{\gamma,\beta/M_{den1}}{\arg max}\ H_L(\alpha,\beta) \quad (29)$$

Equations 28 and 29 identify the maximum axis density of the sign image.

$$W_{up} = H_L \cup \left\{Exp^{-\left|\left|M_{den1(k,l)} - M_{den2(k,l)}\right|\right|}\right\} \quad (30)$$

The above equation is used to update the weight from the hidden layer's output layer.

$$\text{Error rate}(\omega) = \frac{1}{2}\sum(T_t - P_o)^2 \quad (31)$$

Expression 31 is used to find the error rate for backpropagation of the hidden layer to the hidden layer. Assuming that ω denotes the error rate, T_t refers to the target class, and P_o represents the prediction class.

$$\frac{\partial\omega}{\partial W_{up}} = \frac{\partial\omega}{\partial T_t} * \frac{\partial T_t}{\partial H_L} * \frac{\partial T_t}{\partial W_{up}} \quad (32)$$

The above equation is used to back-propagate the sign identification process.

$$Sign_{ident(k,l,\alpha,\beta)} = CR_{k,l} W_{up} \quad (33)$$

The above equation is used to identify the sign language to convert text format based on conjugate (connection)C, R pure color of the sign image and update weight R.

$$Loss = -\sum_{cc}^{cl} Sign_{ident(k,l,\alpha,\beta)} \ln(p_p, cc) \quad (34)$$

The above equation finds the sign language identification loss function based on predicted probability p_p, correct class cc. This section efficiently identifies the sign language and converts it to text format from segmented sign images.

5. Experimental analysis

In this section, we evaluate the performance of the proposed approach for sign language identification using a Deep Backpropagation Neural Network (DBNNet). First, we look at the design of the experiment, performance assessment, and results related to sign identification.

5.1. Experiment setup

This section is included in Simulation Table 1 for sign identification based on the collected dataset and a Jupyter notebook with an Anaconda tool in order to train the proposed model.

Classification performance in %				
No of sign images/approaches	Bi-LSTM	CNN	CBAM	DBNNet
1	65.25	69.45	71.68	76.24
2	73.50	76.32	79.26	81.78
3	79.69	80.48	84.62	88.14
4	84.15	86.24	89.90	94.88

Table 2. Performance classification results for the suggested and prior approaches

The proposed implementation hardware platform is Windows 10-64bit OS, 16GB Ram with processor I5.

Table 1. The SLI model's simulation parameters

Parameter	**Values**
Dataset name	ASL Alphabet dataset
Number of images	87,028
Tool	Anaconda jupyter notebook
Language	Python
Training images	87,000
Testing images	28

5.2. Dataset collection

Indian Sign Language (ISL) character images are gathered in this dataset and arranged into 29 folders that correspond to various genres. Figure 5 displays the 87,000 200x200 pixel gesture images that are contained in the training dataset folder.

Together with 29 categories and 26 letters, A through Z, there are three folders with the contents space, delete, and nothing. The test dataset only contains 29 hand gesture images in order to facilitate the use of real test images.

Fig. 5. Sample images of ASL images

4.3 Evaluation metrics

Using a confusion matrix, we take into account sign identification accuracy, sensitivity, specificity, and F-measure in the evaluation models. The equation for the parameters used is defined as

$$\text{Sensitivity} = \sum \frac{T_p}{T_p + F_n} * 100 \tag{35}$$

$$\text{Specificity} = \sum \frac{T_n}{F_p + T_n} * 100 \tag{36}$$

$$\text{Accuracy} = \sum \frac{T_p + T_n}{T_p + F_n + F_p + T_n} * 100 \tag{37}$$

$$\text{F1-score} = \sum 2 * \frac{\text{Sensitivity} * \text{Specificity}}{\text{Sensitivity} + \text{Specificity}} * 100 \tag{38}$$

The confusion matrix for hand sign language identification is defined in Figure 6. It serves as an example of how well classification works.

4.4 Comparative result analysis for SLI

We compared the suggested approach with recently created methods such as CNN [30], Convolutional Block Attention Module (CBAM) [25], and Bi-directional Long-Short Term Memory (Bi-LSTM) [4].

The suggested method yielded high accuracy identification, as seen by the sign identification of accuracy result displayed in Figure 7 and Table 2. A total of 756 sign images are used for training and 244 for testing in the proposed DBNNet. The input image is trained into each of the three layers of this suggested method one at a time in order to extract crucial features. Therefore, when compared to other methods, the suggested method performs the best for sign language identification.

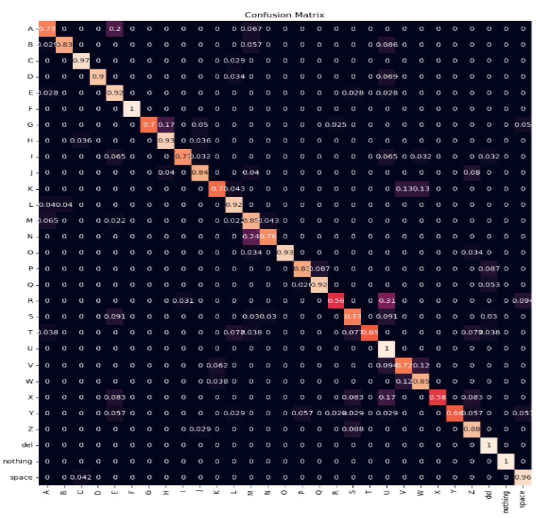


Fig. 6. Confusion Matrix

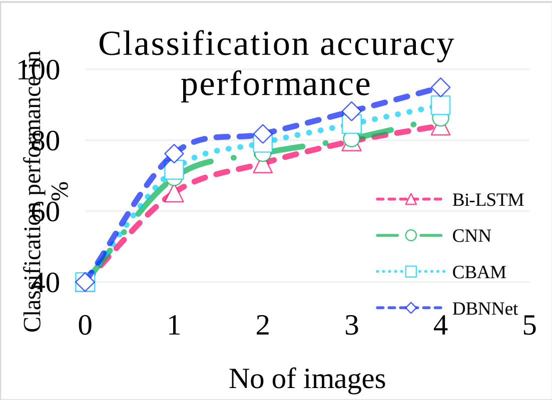


Fig. 7. Classification result for the SLI

Sensitivity performance in %				
No of sign images /Approaches	Bi-LSTM	CNN	CBAM	DBNNet
1	64.59	67.32	69.12	74.38
2	72.89	75.98	77.98	79.35
3	78.98	79.21	84.19	87.26
4	83.59	85.23	88.38	93.45

Table 3. Sensitivity performance comparison results

The robust DBNNnet algorithm SL identification process is described by the sensitivity result displayed in Figure 8 and Table 3. Comparing the suggested algorithm to other techniques like Bi-LSTM and CBAM, it achieved a sensitivity performance of 93.45%.

The suggested and current methods for specificity performance in sign language identification are shown in Figure 9 and Table 4. Comparing the proposed method to previous techniques such as CNN, CBAM, and Bi-LSTM, the specificity performance was 93.14% higher.

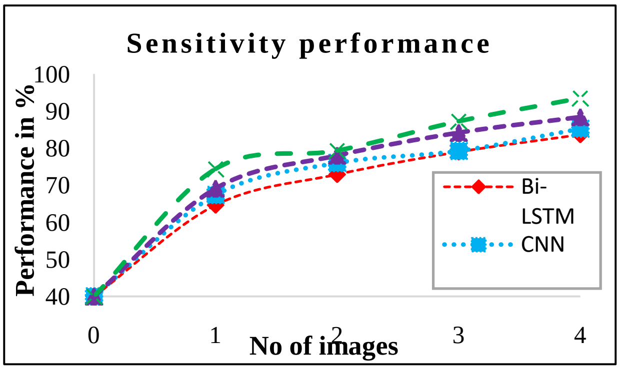


Fig. 8. Results for Sensitivity performance

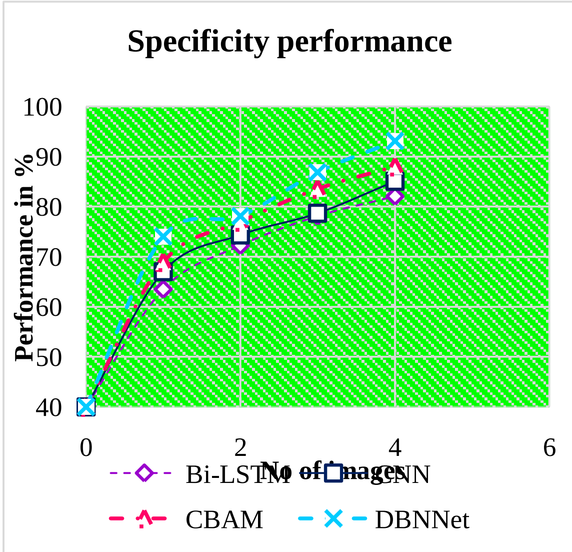


Fig. 9. Specificity performance for detection

Figure 10's sign identification of the F-measure performance provides an explanation of the results of the comparison between the proposed method and alternative methods. The suggested approach provides superior SL identification since it uses the intensity gradients sign edge detection (IGSED) method for hand sign edge detection after first improving the SL image through pre-processing.

Afterwards, divide the ROI using the segmentation technique. Lastly, the SL is found using the suggested DL-based DBNNet technique.

Table 4. Specificity performance for SLI

Specificity performance in %				
No of Sign images/Approaches	Bi-LSTM	CNN	CBAM	DBNNet
1	63.47	67.03	68.93	74.01
2	72.25	74.35	76.98	78.12
3	78.12	78.68	83.56	86.89
4	82.1	85.09	88.01	93.14

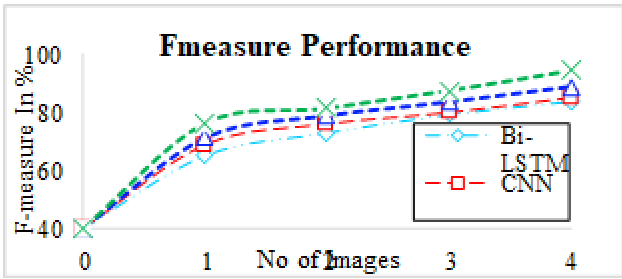


Fig. 10. Comparison of F-measure performance

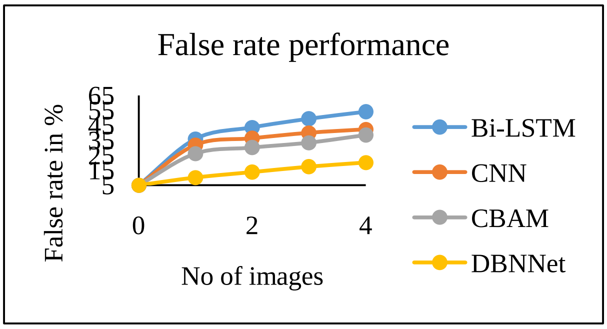


Fig. 11. Comparison of False rate performance

The false performance rating using the suggested and earlier methods is contrasted in Figure 11. The suggested method's efficiency is shown in the figure to have a lower false rate due to its three layers, which allow the input image to be trained into each layer one at a time in order to extract crucial features.

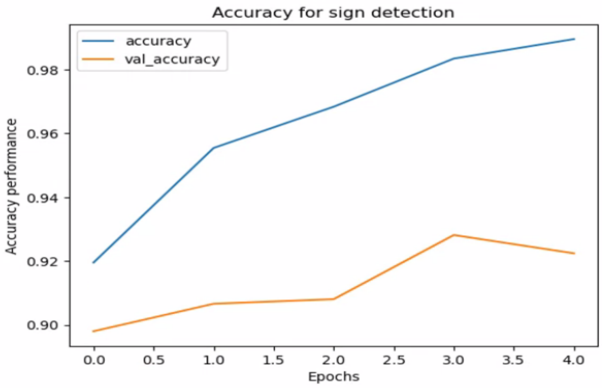


Fig.12. Result for accuracy of sign identification

Figure 12's result demonstrates how sign detection accuracy reaches a high level of performance accuracy. Since extending the epochs enhances the performance in terms of accuracy for sign identification.

The loss for hand sign identification with validation loss is described in Figure 13. The hand sign identification loss graph analysis shows the validation performance in blue and the loss performance in orange. The validation loss performance loss is 0.24, while the suggested loss performance result is 0.065.

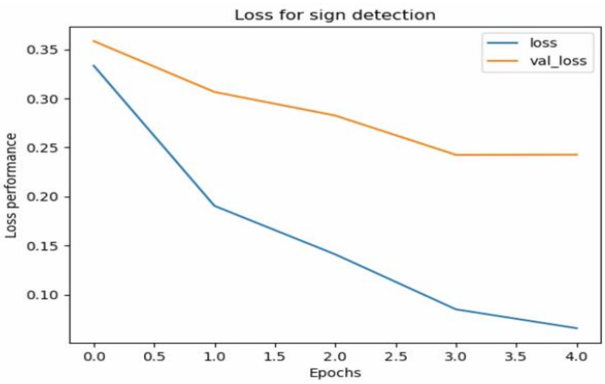
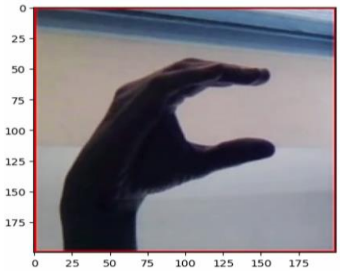
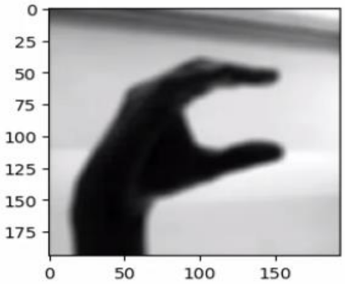
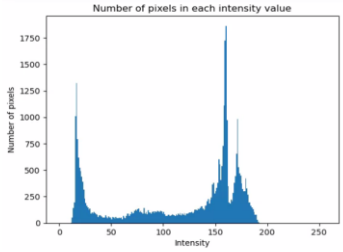
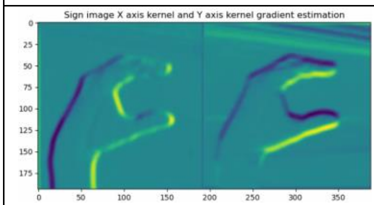


Fig. 13. Result for loss of sign identification

4.5 Discussion

This paper employs the ASL dataset to identify hand gesture signs using the deep learning-based DBNNet algorithm. Table 5's simulation results showed that the suggested algorithm was capable of accurately and proficiently identifying the sign. As of right now, the accuracy of CBAM sign classification is 89.90%, with sensitivity at 88.38%, specificity at 88.01%, F-measure at 88.59%, and false classification at 38.47%.

Table 5. Simulation screenshots for hand sign identification

Figure	Description
 **Fig. 14.** Input image	The hand gesture image from the gathered dataset is defined in Figure 14.
 **Fig. 15.** Pre-processed image	The pre-processed image is examined in Figure 15 using the Gaussian Smoothing Histogram Filter (GMHF) method. The brightness and quality of the sign image are improved by this technique.
 **Fig. 16.** Intensity values in the image	The intensity values in the hand sign from the processed image are defined in Figure 16.
 **Fig. 17.** X and Y axis kernelrepresentation	The X and Y axis kernel representation for determining the horizontal and vertical direction of signs is defined in Figure 17.

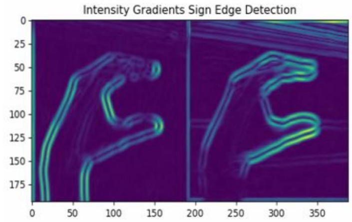
<table>
<tr><td>

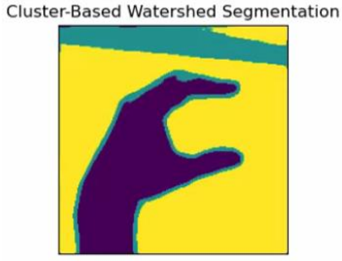
Fig. 18. Intensity Gradients Sign Edge Detection</td><td>The sign is executed based on edge detection in Figure 18. To achieve flawless edge detection, this technique extracts the essential edge features from the image.</td></tr>
<tr><td>

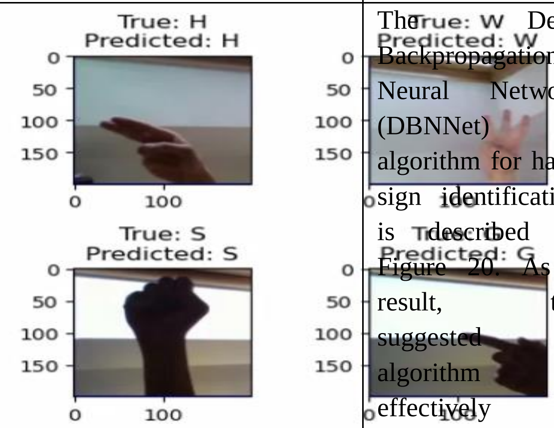
Fig. 19. Cluster-based watershed segmentation</td><td>To estimate the sign language ROI, Figure 19 investigates the use of cluster segmentation. The sign object boundary is mined using this suggested method from edge detection images.</td></tr>
<tr><td>

Fig. 20. Sign Identification</td><td>The Deep Backpropagation Neural Network (DBNNet) algorithm for hand sign identification is described in Figure 20. As a result, the suggested algorithm effectively recognizes the hand sign and formats it as text.</td></tr>
</table>

5. Conclusion

The book presents a deep learning-based approach for categorizing and dividing the dataset of sign language images. Pre-processing, the first method that has been suggested, has been carried out to improve the sign input image. Next, we examined the sign edges of the processed image using the Intensity Gradients Sign Edge Detection (IGSED) method. The sign ROI was then estimated using Cluster-Based Watershed Segmentation (CBWS). Lastly, the suggested DBNNet method for locating the sign and converting it to text. As a result, the simulation classification method identified SL with 94.88% accuracy. Performance for specificity was 93.14%, sensitivity was 93.45%, the F-measure rate was 94.21%, and the false rate was 20.14%. As a result, the suggested approach performs better than earlier ones.

Conflict of Interest: The authors declare that they have no conflict of interest.

Funding: Not applicable

Accessibility of Any Data supporting the Manuscript is provided in the Link below,

https://www.kaggle.com/datasets/grassknoted/isl-alphabet.

Acknowledgements

I thank my colleagues from SRM Institute of Science and Technology, Vadapalani, Chennai, who provided insight and expertise that greatly assisted this part of research, although they may not agree with all of the interpretations/conclusions of this paper. I thank Dr.Paavai Anand G., for comments that greatly improved the manuscript.

Author contributions

G.K.Vaidhya: Conceptualization, Methodology, Implementation, Writing-Original draft preparation, Validation, Field study. **Dr.Paavai Anand G.:** Investigation, Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Saiful, Md&Isam, Abdulla & Moon, Hamim&Tammana, Rifa& Das, Mitul &Alam, Md& Rahman, Ashifur. (2022). Real-Time Sign Language Detection Using CNN. 10.1109/ICDABI56818.2022.10041711.

[2] Dabre, Kanchan; Dholay, Surekha (2014). The machine learning model for sign language interpretation using webcam images. *[IEEE 2014 International Conference on Circuits, Systems, Communication and Information Technology Applications (CSCITA) - Mumbai, Maharashtra, India (2014.4.4-2014.4.5)]* 317–321.

[3] Long Huynh;Viet Ngo; (2020). Recognize Vietnamese Sign Language Using Deep Neural Networks. 2020 7th NAFOSTED Conference on Information and Computer Science (NICS), doi:10.1109/nics51282.2020.9335904

[4] S. B. Abdullahi and K. Chamnongthai, "American Sign Language Words Recognition Using Spatio-Temporal Prosodic and Angle Features: A Sequential Learning Approach," *in IEEE Access, vol. 10, pp. 15911-15923, 2022, doi:*

10.1109/ACCESS.2022.3148132.

[5] S. Z. Gurbuz et al., "American Sign Language Recognition Using RF Sensing," in *IEEE Sensors Journal, vol. 21, no. 3, pp. 3763-3775, 1 Feb.1, 2021, doi: 10.1109/JSEN.2020.3022376.*

[6] M. Lee and J. Bae, "Real-Time Gesture Recognition in the View of Repeating Characteristics of Sign Languages," in *IEEE Transactions on Industrial Informatics*, vol. 18, no. 12, pp. 8818-8828, Dec. 2022, doi: 10.1109/TII.2022.3152214.

[7] B. Natarajan et al., "Development of an End-to-End Deep Learning Framework for Sign Language Recognition, Translation, and Video Generation," in *IEEE Access*, vol. 10, pp. 104358-104374, 2022, doi: 10.1109/ACCESS.2022.3210543.

[8] J. Li, J. Meng, H. Gong, and Z. Fan, "Research on Continuous Dynamic Gesture, Recognition of Chinese Sign Language, Based on Multi-Model Fusion," *in IEEE Access*, vol. 10, pp. 106946-106957, 2022, doi: 10.1109/ACCESS.2022.3212064.

[9] J. DelPreto, J. Hughes, M. D'Aria, M. de Fazio and D. Rus, "A Wearable Smart Glove and Its Application of Pose and Gesture Detection to Sign Language Classification," *in IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 10589-10596, Oct. 2022, doi: 10.1109/LRA.2022.3191232.

[10] Y. Dong, J. Liu and W. Yan, "Dynamic Hand Gesture Recognition Based on Signals From Specialized Data Glove and Deep Learning Algorithms," *in IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-14, 2021, Art no. 2509014, doi: 10.1109/TIM.2021.3077967.

[11] K. Nimisha and A. Jacob, "A Brief Review of the Recent Trends in Sign Language Recognition," *2020 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, 2020, pp. 186-190, doi: 10.1109/ICCSP48568.2020.9182351.

[12] S. S. Kakkoth and S. Gharge, "Survey on Real Time Hand Gesture Recognition," 2017 *International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)*, Mysore, India, 2017, pp. 948-954, doi: 10.1109/CTCEEC.2017.8455041.

[13] X. Zhang, Z. Yang, T. Chen, D. Chen and M. -C. Huang, "Cooperative Sensing and Wearable Computing for Sequential Hand Gesture Recognition," in *IEEE Sensors Journal*, vol. 19, no. 14, pp. 5775-5783, 15 July15, 2019, doi: 10.1109/JSEN.2019.2904595.

[14] B. G. Lee and S. M. Lee, "Smart Wearable Hand Device for Sign Language Interpretation System With Sensors Fusion," in *IEEE Sensors Journal*, vol. 18, no. 3, pp. 1224-1232, 1 Feb.1, 2018, doi: 10.1109/JSEN.2017.2779466.

[15] T. Sagisaka, Y. Ohmura, Y. Kuniyoshi, A. Nagakubo and K. Ozaki, "High-density conformable tactile sensing glove," 2011 11th *IEEE-RAS International Conference on Humanoid Robots*, Bled, Slovenia, 2011, pp. 537-542, doi: 10.1109/Humanoids.2011.6100898.

[16] R. Galicia, O. Carranza, E. D. Jiménez and G. E. Rivera, "Mexican sign language recognition using movement sensor," 2015 *IEEE 24th International Symposium on Industrial Electronics (ISIE)*, Buzios, Brazil, 2015, pp. 573-578, doi: 10.1109/ISIE.2015.7281531.

[17] M. Al-Hammadi et al., "Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation," in *IEEE Access, vol. 8, pp. 192527-192542, 2020, doi: 10.1109/ACCESS.2020.3032140.*

[18] Z. Chen, J. Yang and H. Xie, "Surface-Electromyography-Based Gesture Recognition Using a Multi-stream Fusion Strategy," in *IEEE Access*, vol. 9, pp. 50583-50592, 2021, doi: 10.1109/ACCESS.2021.3059499.

[19] N. Mohamed, M. B. Mustafa and N. Jomhari, "A Review of the Hand Gesture Recognition System: Current Progress and Future Directions," in *IEEE Access*, vol. 9, pp. 157422-157436, 2021, doi: 10.1109/ACCESS.2021.3129650.

[20] M. K. Burns, D. Pei and R. Vinjamuri, "Dynamic Control of Virtual Hand Grasp Using Spatiotemporal Synergies," in *IEEE Access*, vol. 7, pp. 112327-112338, 2019, doi: 10.1109/ACCESS.2019.2932956.

[21] M. Diaz, M. A. Ferrer and J. J. Quintana, "Anthropomorphic Features for On-Line Signatures," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 12, pp. 2807-2819, 1 Dec. 2019, doi: 10.1109/TPAMI.2018.2869163.

[22] W. Pan, X. Zhang and Z. Ye, "Attention-Based Sign Language Recognition Network Utilizing Keyframe Sampling and Skeletal Features," in *IEEE Access*, vol. 8, pp. 215592-215602, 2020, doi: 10.1109/ACCESS.2020.3041115.

[23] M. A. Ferrer, M. Diaz, C. Carmona-Duarte and A. Morales, "A Behavioral Handwriting Model for Static and Dynamic Signature Synthesis," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1041-1053, 1 June 2017, doi: 10.1109/TPAMI.2016.2582167.

[24] Fischer, R. Plamondon, C. O'Reilly and Y. Savaria, "Neuromuscular Representation and Synthetic Generation of Handwritten Whiteboard Notes," 2014 *14th International Conference on Frontiers in Handwriting Recognition*, Hersonissos, Greece, 2014, pp. 222-227, doi: 10.1109/ICFHR.2014.45.

[25] YanglaiHuang , Jing Huang, "Dynamic Sign Language Recognition Based on CBAM with Autoencoder Time Series Neural Network", *Hindawi Mobile Information Systems,* Volume 2022, 10 pages.

[26] S. Huang and Z. Ye, "Boundary-Adaptive Encoder With Attention Method for Chinese Sign Language Recognition," in *IEEE Access*, vol. 9, pp. 70948-70960, 2021, doi: 10.1109/ACCESS.2021.3078638.

[27] N. Sae-Bae, N. Memon, K. Isbister and K. Ahmed, "Multitouch Gesture-Based Authentication," in *IEEE Transactions on Information Forensics and Security,* vol. 9, no. 4, pp. 568-582, April 2014, doi: 10.1109/TIFS.2014.2302582

[28] D. R. Kothadiya, C. M. Bhatt, T. Saba, A. Rehman and S. A. Bahaj, "SIGNFORMER: DeepVision Transformer for Sign Language Recognition," in *IEEE Access, vol. 11, pp. 4730-4739, 2023, doi: 10.1109/ACCESS.2022.3231130.*

[29] D. Kothadiya, C. Bhatt, K. Sapariya, K. Patel, A.-B. Gil-González and J. M. Corchado, "Deepsign: Sign language detection and recognition using deep learning", *Electronics, vol. 11, no. 11, pp. 1780, Jun. 2022.*

[30] A. Mannan, A. Abbasi, A. R. Javed, A. Ahsan, T. R. Gadekallu and Q. Xin, "Hypertuned deep convolutional neural network for sign language recognition", *Comput. Intell. Neurosci.*, vol. 2022, pp. 1-10, Apr. 2022.

[31] M. Zakariah, Y. A. Alotaibi, D. Koundal, Y. Guo and M. M. Elahi, "Sign language recognition for Arabic alphabets using transfer learning technique", *Comput. Intell. Neurosci*., vol. 2022, pp. 1-15, Apr. 2022.

[32] P. Rathi, R. K. Gupta, S. Agarwal and A. Shukla, "Sign language recognition using ResNet50 deep neural network architecture", *Proc. 5th Int. Conf. Next Gener. Comput. Technol. (NGCT)*, 2020.

[33] H. Luqman, "An Efficient Two-Stream Network for Isolated Sign Language Recognition Using Accumulative Video Motion," in *IEEE Access*, vol. 10, pp. 93785-93798, 2022, doi: 10.1109/ACCESS.2022.3204110.

[34] F. Shah, M. S. Shah, W. Akram, A. Manzoor, R. O. Mahmoud and D. S. Abdelminaam, "Sign Language Recognition Using Multiple Kernel Learning: A Case Study of Pakistan Sign Language," in *IEEE Access*, vol. 9, pp. 67548-67558, 2021, doi: 10.1109/ACCESS.2021.3077386.

[35] E. Rajalakshmi et al., "Multi-Semantic Discriminative Feature Learning for Sign Gesture Recognition Using Hybrid Deep Neural Architecture," in *IEEE Access*, vol. 11, pp. 2226-2238, 2023, doi: 10.1109/ACCESS.2022.3233671.