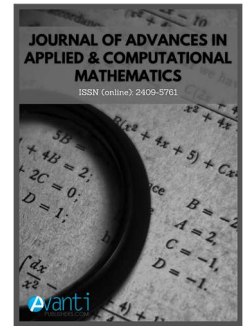




Published by Avanti Publishers
**Journal of Advances in Applied &
Computational Mathematics**

ISSN (online): 2409-5761



Recognition of Sign Language Based on Hand Gestures

Bala Murali Gunji*, Nikhil M. Bhargav, Amrita Dey, Isahak Karajagi Zeeshan Mohammed and Sachdev Sathyajith

School of Mechanical Engineering, Department of Design and Automation, Vellore Institute of Technology, Vellore-632014, Tamilnadu, India.

ARTICLE INFO

Article Type: Research Article

Keywords:

KNN

Edge detection

Neural network

Grey conversion

Sign language recognition

Timeline:

Received: June 06, 2021

Accepted: August 11, 2021

Published: November 02, 2021

Citation: Gunji BM, Bhargav NM, Dey A, Mohammed IKZ, Sathyajith S. Recognition of Sign Language Based on Hand Gestures. J Adv App Comput Math. 2021; 8: 21-32.

DOI: <https://doi.org/10.15377/2409-5761.2021.08.3>

ABSTRACT

The target of SLR or sign language recognition is to interpret the sign language into text, respectively. So the deaf and mute people can communicate with ordinary people easily. Sign language recognition has a tremendous social impact; however, it is challenging due to the significant variations and complexity in the hand actions. There are many existing methods for recognizing sign language that uses handcraft features for describing the motion of sign language and then, based on the features it makes the classification models. To approach the problem, we have discussed considering the KNN that can conveniently extract the features. The proposed model can be validated on a real data set.

*Corresponding Author
Email: bmgunji@gmail.com
Tel: +91 8093831101

1. Introduction

This paper applied K-Nearest Neighbor (KNN) algorithm for sign language recognition. The design of the algorithm is such that it works as a first-level detection and the images captured are converted into spellings using a series of steps. Since the gestures are treated as curve blocks that must be extracted with the help of a predefined character set while teaching the algorithm, KNN forms the guideline for the recognition system. KNN is used from the histogram transition table as a string of hashed transformation blocks for image pre-processing. The idea proposed in this project is to design a system for the accurate recognition of numeric signs. The signs were acquired using a regular camera. 5000 signs were created i.e., 500 per number. The desired features were extracted using hierarchical centroid and direct pixel techniques. The signs were classified using the KNN classifier and neural network techniques; 97.1% accuracy is achieved using these techniques. Ordinary people can communicate through speech and express their thoughts and ideas. However, for the community of the hearing impaired, communication is carried out by sign language. There were several methods introduced in the past for the translation of signs using the features and the gestures of the signer. Ko and Yang have successfully developed the finger mouse that allows and enables the user to specify the commands with the help of fingers [1].

A deep learning-based convolution neural network (CNN) was constructed using convolution layers followed by other layers. Instead of using complex handcrafted features, CNNs can automate the feature construction process. By using this method, 20 Italian gestures were identified with high accuracy of 91.7% when tested with experimental data [2].

Islam *et al.* identify the alphabets appropriately even for 48% percent noise in the image during processing. Convolutional Neural Network (CNN) is used to extract features to train the machine, and Multi-class Support Vector Machine (MCSVM) is used to identify the hand signs. In the training model, hand gesture images were given as input. A non-linear MVSVM is used for the classification of extracted images and regression. Kernel function was used in MVSVM for the classification of non-linear data. Finally, In the experiment, a total of 9360 images were used as a dataset, with 30 % images for training and 70% images for testing [3].

G. Anantha Rao *et al.* proposed a methodology for recognizing Indian Sign Language using CNN, used for deaf persons. For the experiment, the dataset was created using two hundred hand sign images captured at different angles. Also, these images were captured in various backgrounds, and the system architecture was built using different convolution layers, rectified linear units, stochastic pooling layers, and output layers [4]. The algorithm proposed by Rahaman *et al.* discusses Bengali sign language recognition using real-time computer vision. The system uses a cascade classifier for detecting the hands from the captured image. The hand sign from the captured image is extracted based on the color of the skin using Hue and saturation value. KNN classifier is used to compare the captured image with the trained data set of 3600 images [5].

Djamila Dahmani *et al.* [6] proposed a new sign language recognition system using hand postures. Hand orientation was taken into account in the proposed method. A KNN Classifier was used in the algorithm owing to its simplicity and high accuracy. The authors used a new segmentation approach based on skin color that can be used in backgrounds having a color similar to skin. Feature extraction was done using KNN and SVM Classification. The results showed an accuracy closed to 94%.

Aditi Kalsh *et al.* [7] use the Human-Computer Interface system with K-nearest neighbor algorithm for understanding the sign language more correctly to eliminate the use of an interpreter for deaf and hard to hear people. A camera and a computer system were used to capture and process the images. A fuzzy rule set was used for the image classification. The experimental results showed improved accuracy and a faster recognition rate compared to previous algorithms. Anup Nandy *et al.* [8] studied the Indian sign language and experimented with the K -nearest neighbor algorithm to recognize the sign language. The Videos captured for the process are divided into frames, and greyscale conversion was done. The experimental results carried with Euclidean distance and K-nearest neighbor matrices show a very high level of recognition accuracy using the image sets.

Galvão, R. K. H., & Yoneyama, T. proposed a novel signal clustering method using unsupervised training of wavelet networks. In the proposed algorithm, the weights are optimized using the wavelets basis functions. The developed algorithm is applied for the electrocardiographic signals obtained from the MIT-BIH database [9].

Parvathy, Priyanka, *et al.* proposed a vision-based hand gesture recognition system using machine learning for identifying the gestures in the image. The authors use the discrete wavelet transform and speed up robust feature extraction techniques to extract the features of the captured image. Later, SVM was used to identify the hand gestures in the image. The proposed method identifies the hand gestures in the image with an accuracy of 96.5% with a response time of 0.024 sec [10].

Tsai, T. H., Huang, C. C., & Zhang, K. L. proposed a low-cost human-computer interface system to hand gesture recognition. The proposed system mainly uses the skin and motion technique to capture the region interest from the captured image. In addition to the skin and motion technique, a convex hull algorithm was used to capture the exact area of the hand gesture by removing the hand gesture from the image. The proposed system is applied to the real-time problem to identify the accuracy of the algorithm [11].

Bhuvaneshwari, C., & Manjunathan A. proposed an embedded system with image processing for identifying the sign language using hand gesture posture by the common person. The proposed device converts the hand gesture image to the voice, which can be understood by the common person [12].

Al-Hammadi, Muneer, *et al.* proposed a deep convolutional neural network for hand gesture recognition. In this method, authors use large data sets for training the algorithm for effective identification of the hand gestures postures. The proposed algorithm identifies hand gestures images with an accuracy of 98.12% [13]. Rojasara, D., & Chitaliya, N. uses the Euclidean distance classifier to identify the various Indian sign languages for understanding the sign communication by the deaf people. The authors applied the Euclidean distance classifier on different signs and found be identifying with an accuracy of 72% [14]. Sharma, S., & Singh, S proposed a CNN (Convolution Neural Network) based deep learning algorithm to identify the hand gestures for a visual form of communication. The proposed system is trained by considering large data set of Indian sign language (ISL) and American sign language (ASL). The trained system is used for identifying hand gestures based-sign language for visual communication and the system is performing with an accuracy of 99.9%. The proposed system is compared with the existing state of art approaches [15].

Ghule, S., & Chavaan, M. developed a time-based system to identify the gestures and convert them to text and voice. The built setup is tested with the 800 samples out of which 760 samples are detected correctly with an accuracy of 95% [16]. Tasmere, D., Ahmed, B., & Das, S. R. proposed a system to analyze the identification of different hand movements as part of visual communication with deaf and dumb people using CNN. The proposed model has achieved an accuracy of 94.61% in recognizing the hand gestures symbols [17].

Vanaja, S., Preetha, R., & Sudha, S. has developed a CNN network to identify the different ISL. The developed network is trained with 3500 static images data sets with 4 layers 16 filter CNN network. Adam optimizer is used in CNN to optimize the weights while training the data sets. The developed algorithm is identifying the sign languages with an accuracy of 99.76% [18]. As the developments in recognizing the hand gesture movements are increased, the researchers are not only concentrating on recognizing the hand gesture movements with different techniques but they focused on developing on entire system with vision assistance [19, 20].

By keeping the above-discussed literature in view, in this research work, authors proposed KNN based image recognition concept to identify the hand gesture postures, and the same is applied to real-world hand gesture recognition. The proposed methodology identifies the hand gestures and converts them into sign language with an accuracy of 83%.

1.1. ISL Gestures for Numeric and Alphabets

Sign language is an essential means of communication for physically abled (especially deaf and dumb) people. Steps followed to implement sign language recognition are:

- Image collection
- Image pre-processing
- Feature extraction (using K-means clustering, visual words collection)
- Classification

Figure 1 shows different hand gestures positions for various letters and alphabets. Using this data and with the help of image processing techniques, in this research article, alphabets and numbers are identified and generate the output as sound to hear the blind persons.

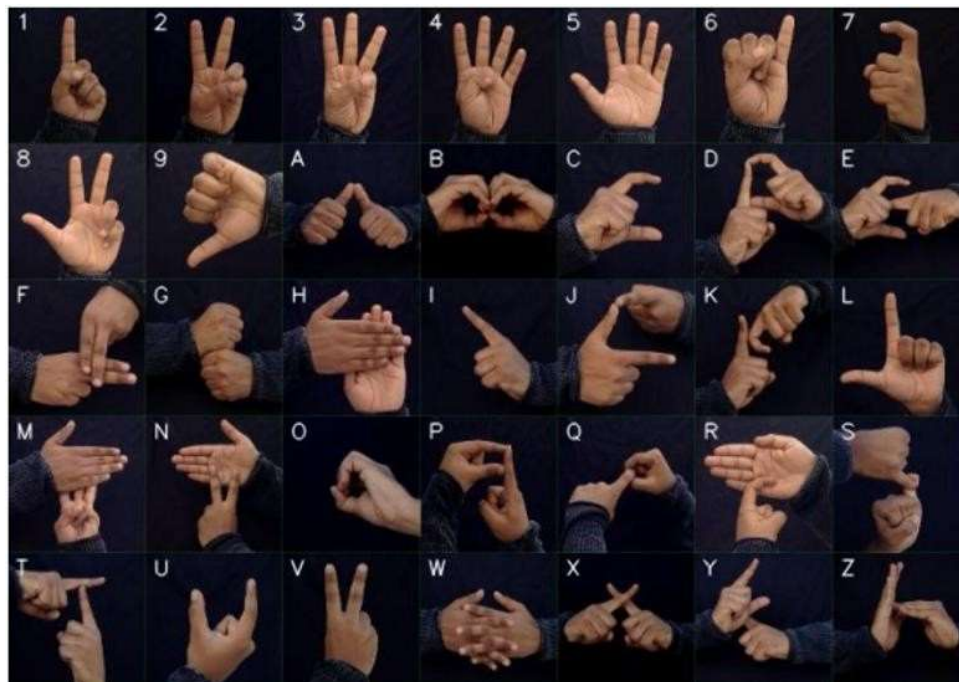


Figure 1: Indian sign language for numbers and alphabets.

1.2. Real-Time Recognition of Gestures

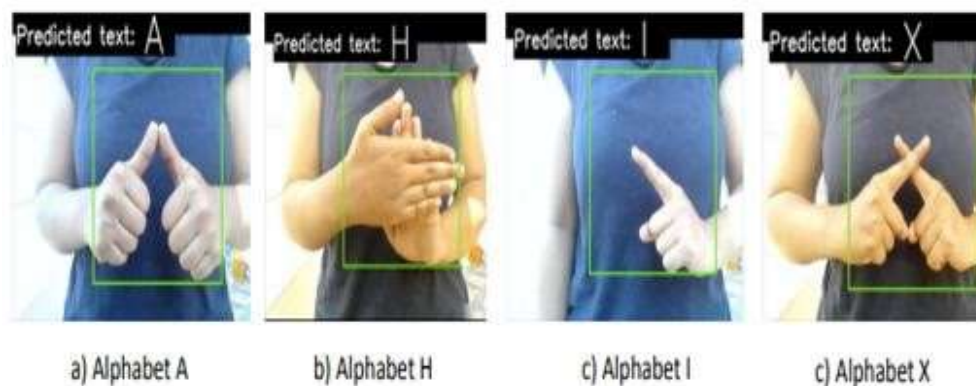


Figure 2: Symbols for various alphabets for real-time recognition.

Figure 2 shows the real-time hand gesture symbols for identifying the alphabets and numbers. These are processed using image processing techniques to identify the symbols and generate the output as audio signals for blind persons.

1.3. Overview of the System



Figure 3: Steps involved in sign language recognition.

Step-1: In this step, the image is captured through the webcam and stored in the database for further process.

Step-2: This step involves processing. Pre-processing of images is done so that they are ready for feature extraction. The following are the operations that are performed to extract the information from the image.

- Image segmentation (skin masking)
- Skin detection
- Edge detection

Image segmentation: The image is converted into HSV color space to obtain a masked skin image. Skin pixels range from (0,40,30) to (43,255,254). Skin segmentation can be done using a skin mask.

Edge detection: The edges in the images are found by the Canny Edge technique. This is done by detecting sharp discontinuities in the obtained image.

2. Methodology Implementation

2.1. Image Pre-Processing

The pre-processing of the video data content is very important for satisfying the environmental scene and taking into consideration the memory requirement [11]. Factors like background, viewpoint, illumination and the camera location are crucial to address the complexity of the scene, which affect the images of the object. The first step involved in the pre-processing block is the filtering of the image shown in Figure 4. A median filter or a moving average filter can be effectively used to remove the noise of the image. The next step involves the background subtraction of the image, which can be done by running a Gaussian average filter [12].

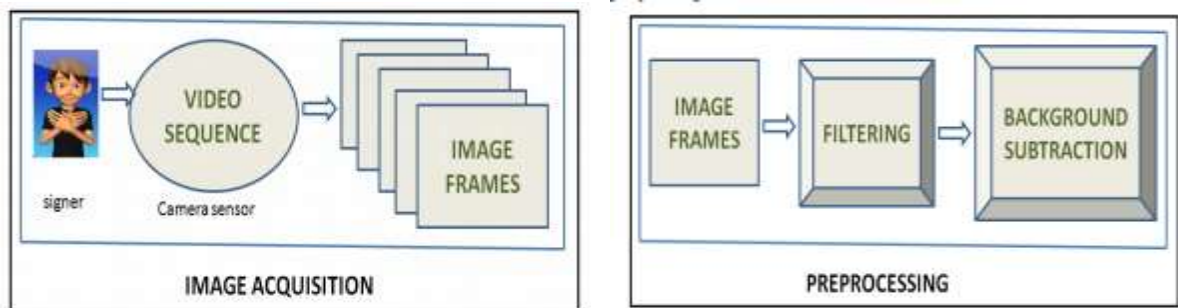


Figure 4: Image acquisition and processing.

2.2. Feature Extraction

In this section, from the captured image through the camera source, features (symbols) should be identified. The image is read as a matrix, and the computation involves time and memory. There are two subcategories of the features hand movement and hand shape. The point of interest of hands is the attributes of the hand gestures. In Figure 5, one can see there are two points of interest to represent the direction of movement and shape. A, B, C, D, E is the fingertips, and the hand is considered the track point denoted by TP. The Track point

indicates the motion direction. To extract the features in the image, discrete wavelet transforms and Gray Level Co-occurrence Matrix are used.

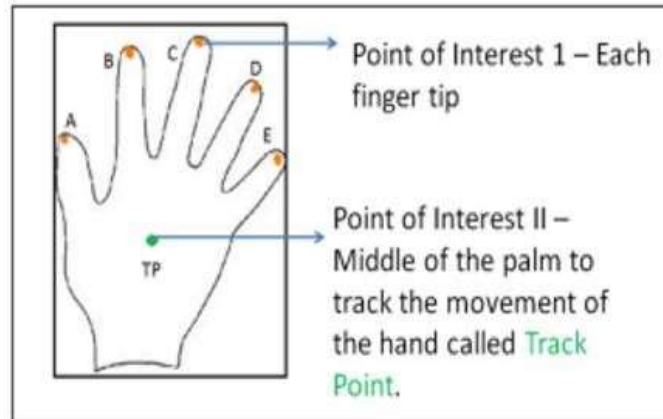


Figure 5: Point of interest and midpoint of palm.

2.3. Discrete Wavelet Transforms (DWT)

DWT is a computerized technique for computing the wavelet transform of a signal at a quick pace. It is the best solution for computational time overload and can also be used for image processing. They are used to identify and filter out all white Gaussian noise. Wavelet transformation is done on 2-D images. Individual signals are obtained by applying DWT for 10 levels. By evaluating parameters like variance, entropy, energy, standard deviation, etc., the one-dimensional feature can be extracted.

2.4. GLCM and Textures

GLCM is Gray Level Co-occurrence Matrix calculated by the occurrence of combinations of gray levels in an image, and texture feature measures image texture using contents of GLCM at the pixel of interest. GLCM is also offered by Echoy view, which generates a virtual variable that represents, on a single beam echogram, specified texture calculation.

2.5. K-Nearest Neighbor (KNN) Algorithm

KNN is one of the supervised learning algorithms used to find similar things in a particular scene. This algorithm is mainly used in pattern recognition of the images to find a similar pattern in the images. The pseudo-code of the KNN algorithm is shown below, which is implemented for the symbol identification in the captured image.

%%-----pseudo code of KNN algorithm-----%%

Step-1: Load the extracted data for training.

Step-2: Calculate the distance between the data using the Euclidean distance formula.

$$\text{Euclidean distance } n = \sqrt{\sum_{i=1}^k (x_i - y_i)^2}$$

Step-3: Arrange the n value on the non-decreasing side

Step-4: Find the optimal value for K

Step-5: Form the sorted n values select the top K rows

Step-6: Find the most frequent class from these chosen 'K' rows. This will be the predicted class.

A complete step-by-step process of the proposed methodology, starting from the image acquisition to extracting features and identifying the symbols using KNN, is shown in Figure 6.

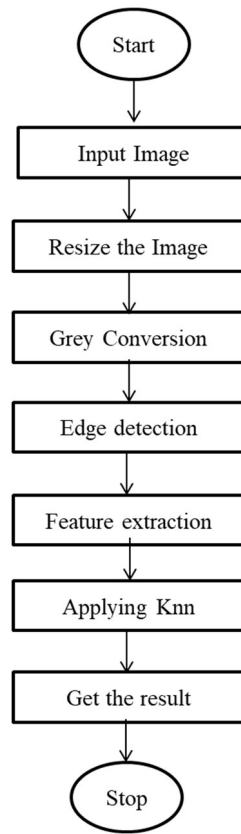


Figure 6: Flow chart of the proposed methodology.

3. Results and Discussions

In this section, the identification of symbols using the proposed methodology is explained. For this, the letters 'C', letter 'J' and letter 'Z' and number '7' are identified with the proposed methodology.

Alphabet C

To identify the symbols in the image, the captured image undergoes Resizing. After resizing, the grey conversion of the image is obtained. The Grayscale images are sufficient for several tasks, and also, there is no need to use a more complicated process like color images.

The edge detection permits the user to observe several features by the change in gray level at a significant amount. This helps in reducing the data of the image and also helps in retaining the structural properties of the image, respectively.

Edge detection is followed by image segmentation. Here DWT feature extraction and FCM segmentation are used to extract the information from the image. The segmentation helps partition the image into various subgroups called image objects that can reduce the complexity of the image, and thus, the analysis of the image becomes easier. The network is finally trained properly in KNN and can identify the sign language correctly with a good accuracy percentage. A confusion matrix is also calculated to know the performance of the algorithm. There is no confusion between selections, as witnessed. Figure 7 represents the conversion of the color image to a grey image with a reduced scale. Figure 8 represents the image enhancement, data segmentation, and feature extraction. Figure 9 represents the confusion matrix along with the output result of finding the alphabet letter 'C'.



Figure 7: Conversion of input image to grey scale.



Figure 8: Image Enhancement.

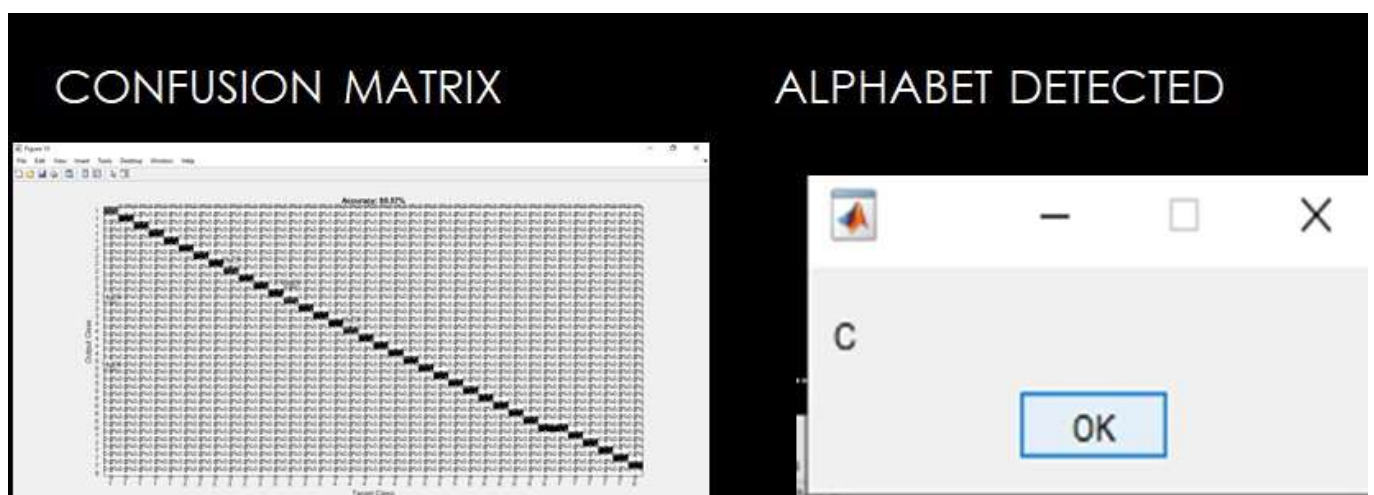


Figure 9: Confusion matrix.

For number 7

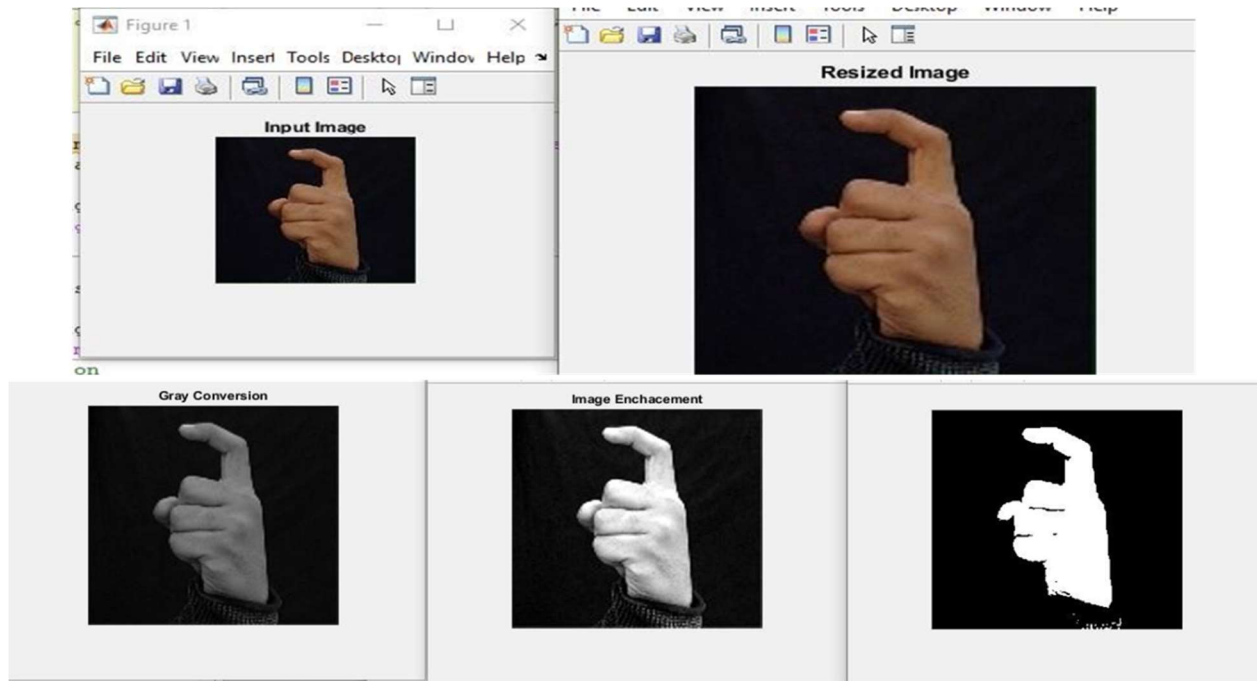


Figure 10: Image conversion and enhancement.



Figure 11: Image output.

Similar to finding the letter 'C' using the proposed methodology, number '7' is also identified, as shown in Figures 9 & 10.

For Alphabet J

Similar to finding the letter 'C' using the proposed methodology, the letter 'J' is also identified, as shown in Figure 11.

For Alphabet Z

Similar to finding the letter 'C' using the proposed methodology, the letter 'Z' is also identified, as shown in Figure 12.

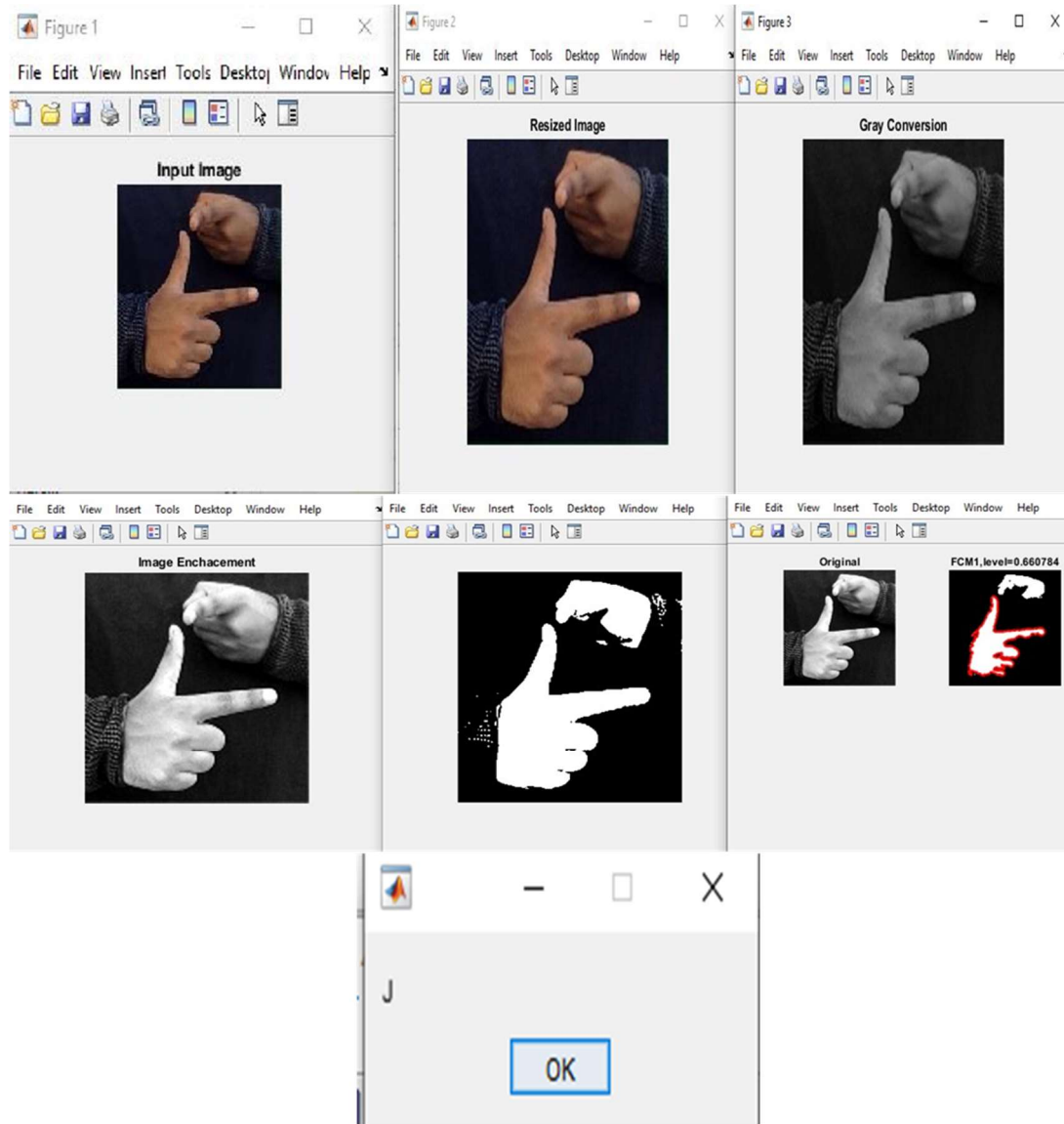


Figure 12: Identification of Alphabet J.

4. Conclusion

The proposed methodology uses the KNN algorithm for sign language recognition based on classification algorithms. Initially, in the proposed methodology, hand gestures are captured with the camera and send to the image processing unit, where image processing operations are carried out, such as feature extraction and edge detection to extract the information from the captured image. To identify the sign language from the captured hand gesture positions, sign languages were categorized into five classes, each containing five images. Total 175 images for the alphabets (A-Z) and numeric (1-9), which were categorized into a total of 35 classes. To increase accuracy, pre-training was performed with a more extensive dataset. Recognition of moving signs such as letters "J" and "Z" was possible and done with the help of a motion vector. The system is trained with four subjects, and 83% accuracy was achieved for the Indian sign language dataset with the help of a depth image dataset.

As a future scope, the proposed methodology will be trained with more data sets. Even the hand gesture shown by the partially hand disabled can be recognized as sign language, and the also sees that the accuracy of identifying the hand gestures for sign language increases above 90 %.

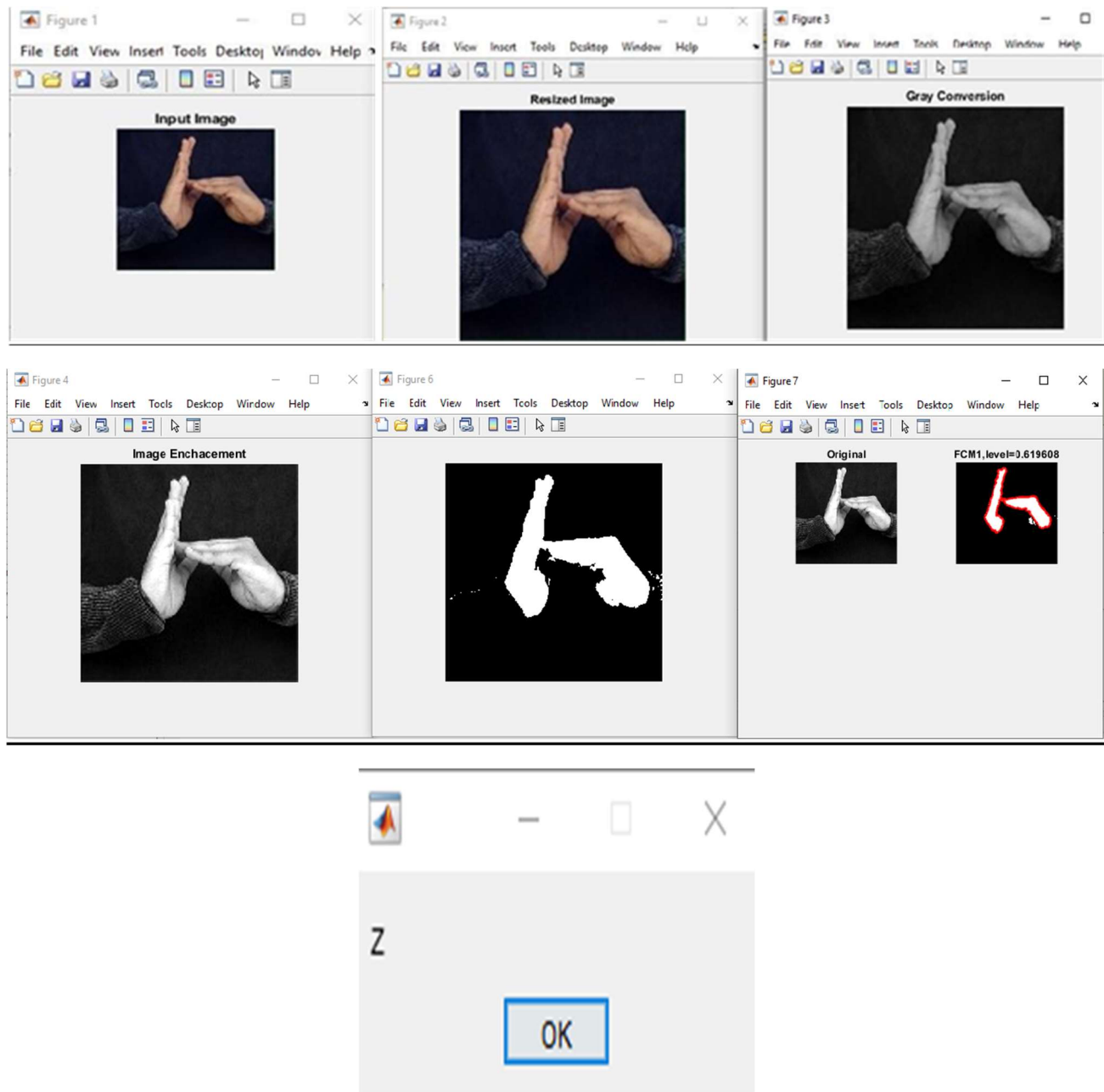


Figure 13: Identification of Alphabet Z.

Reference

- [1] Ko BK, Yang HS. Finger mouse and gesture recognition system as a new human computer interface. *Computers & Graphics*, 1997; 21(5): 555-561. [https://doi.org/10.1016/S0097-8493\(97\)00034-4](https://doi.org/10.1016/S0097-8493(97)00034-4)
- [2] Pigou L. *et al.* Sign language recognition using convolutional neural networks. *European Conference on Computer Vision*. Springer, Cham, 2014.
- [3] Islam MR, Mitu UK, Bhuiyan RA, Shin J. Hand gesture feature extraction using deep convolutional neural network for recognizing American sign language. In 2018 4th International Conference on Frontiers of Signal Processing (ICFSP) 2018; pp. 115-119. IEEE. <https://doi.org/10.1109/ICFSP.2018.8552044>
- [4] Rao GA, Syamala K, Kishore PVV, Sastry ASCS. Deep convolutional neural networks for sign language recognition. In 2018 Conference on Signal Processing And Communication Engineering Systems (SPACES) 2018; pp. 194-197. IEEE. <https://doi.org/10.1109/SPACES.2018.8316344>

- [5] Rahaman MA, Jasim M, Ali MH, Hasanuzzaman M. Real-time computer vision-based Bengali sign language recognition. In 2014 17th international conference on computer and information technology (ICCIT) 2014; pp. 192-197. IEEE. <https://doi.org/10.1109/ICCITechn.2014.7073150>
- [6] Dahmani D, Larabi S. User-independent system for sign language finger spelling recognition. Journal of Visual Communication and Image Representation, 2014; 25(5): 1240-1250. <https://doi.org/10.1016/j.jvcir.2013.12.019>
- [7] Kalsh EA, Garewal NS. Sign language recognition system. International journal of computational engineering research, 2013; 3(6): 15-21.
- [8] Nandy A, Prasad JS, Mondal S, Chakraborty P, Nandi GC. Recognition of isolated indian sign language gesture in real time. In International Conference on Business Administration and Information Processing 2010; pp. 102-107. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-12214-9_18
- [9] Galvão RKH, Yoneyama T. A competitive wavelet network for signal clustering. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2004; 34(2): 1282-1288. <https://doi.org/10.1109/TSMCB.2003.817104>
- [10] Parvathy P, Subramaniam K, Venkatesan GP, Karthikaikumar P, Varghese J, Jayasankar T. Development of hand gesture recognition system using machine learning. Journal of Ambient Intelligence and Humanized Computing, 2021; 12(6): 6793-6800. <https://doi.org/10.1007/s12652-020-02314-2>
- [11] Tsai TH, Huang CC, Zhang KL. Design of hand gesture recognition system for human-computer interaction. Multimedia Tools and Applications, 2020; 79(9): 5989-6007. <https://doi.org/10.1007/s11042-019-08274-w>
- [12] Bhuvaneshwari C, Manjunathan A. Advanced gesture recognition system using long-term recurrent convolution network. Materials Today: Proceedings, 2020; 21: 731-733. <https://doi.org/10.1016/j.matpr.2019.06.748>
- [13] Al-Hammadi M, Muhammad G, Abdul W, Alsulaiman M, Bencherif MA, Mekhtiche MA. Hand gesture recognition for sign language using 3DCNN. IEEE Access, 2020; 8: 79491-79509. <https://doi.org/10.1109/ACCESS.2020.2990434>
- [14] Rojasara D, Chitaliya N. Real time visual recognition of Indian sign language using wavelet transform and principle component analysis. International Journal of Soft Computing and Engineering (IJSCE), 2014; 4(3): 17-20.
- [15] Sharma S, Singh S. Vision-based hand gesture recognition using deep learning for the interpretation of sign language. Expert Systems with Applications, 2021; 182: 115657. <https://doi.org/10.1016/j.eswa.2021.115657>
- [16] Ghule S, Chavaan M. Implementation of Hand Gesture Recognition System to Aid Deaf-Dumb People. In Advances in Signal and Data Processing 2021; pp. 183-194. Springer, Singapore. https://doi.org/10.1007/978-981-15-8391-9_14
- [17] Tasmere D, Ahmed B, Das SR. Real time hand gesture recognition in depth image using cnn. International Journal of Computer Applications, 975: 8887.
- [18] Vanaja S, Preetha R, Sudha S. Hand Gesture Recognition for Deaf and Dumb Using CNN Technique. In 2021 6th International Conference on Communication and Electronics Systems (ICCES) 2021; pp. 1-4 IEEE. <https://doi.org/10.1109/ICCES51350.2021.9489209>
- [19] Jain R, Jain M, Jain R, Madan S. Human Computer Interaction-Hand Gesture Recognition. Advanced Journal of Graduate Research, 2022; 11(1): 1-9. <https://doi.org/10.21467/ajgr.11.1.1-9>
- [20] Gourab JH, Raxit S, Hasan A. A Robotic Hand: Controlled With Vision Based Hand Gesture Recognition System. In 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI) 2021; pp. 1-4. IEEE.