I.J. Image, Graphics and Signal Processing, **2022, 2, 61-73**
Published Online April 2022 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2022.02.06

Indian Sign Language Recognition Using 2-D Convolution Neural Network and Graphical User Interface

Shashidhar R
Department of Electronics and Communication Engineering, Sri Jayachamarajendra College of Engineering, JSS Science and Technology University, Mysuru, India
Email: shashidhar.r@sjce.ac.in

Arunakumari B. N.
Department of Computer Science and Engineering, BMS Institute of Technology and Management, Bengaluru, India
Email: arunakumaribn@bmsit.in.

A S Manjunath
Department of Computer Application, JSS Science and Technology University, Mysuru, India
Email: as.manjunath@sjce.ac.in

Roopa M
Department of Electronics and Communication , Dayananda Sagar College of Engineering, Bengaluru, India
Email: surajroopa@gmail.com

Received: 01 January 2022; Accepted: 06 March 2022; Published: 08 April 2022

Abstract: The emergence of the sign Language recollection method has a great effect on the day-to-day livings of human beings with hearing disabled individuals utilizing signs to speak with others. Much the same as verbally communicated in dialects, there is no general language as each nation has its communication in language, so every nation has its vernacular of gesture-based communication and in India, they utilize Indian Sign Language (ISL). Over the most recent couple of years, analysts have investigated the computerization of ISL. Here we developed the custom database for 26 English letters and each Letter narrates the 5 times by each person. Train the dataset using 2D CNN and create GUI for recognition. A few endeavors have been made in India and different nations. We attempt to investigate and dissect the ISL that has been made with the mechanization of communication through signing and motion acknowledgment. We attempted to investigate the difficulties that come in the ongoing sign acknowledgment framework. The testing accuracy of the proposed work is 95% and 95% for the validation accuracy.

Index Terms: Indian Sign language (ISL), Convolution Neural Network (CNN), Database Design, 2D, and Sign Language.

1. Introduction

Hand gesticulations are the greatest important form of communiqu é used by the deafness persons in their day-to-day life. Communication through signing is a characteristic linguistic which utilizes foresight motions and signal to pass on significant data among individuals. Gesture-based communication use goes back to 60 A.D, even though first recorded in the seventeenth century and it dominatingly began in western nations. Sign normally speaks to fingerspelling yet sometimes they additionally speak to phrases, thoughts. It is utilized by tough to hearing and inept to communicate people groups to express between themselves as well as other people by utilizing hand signals or developments or signs. Sign-based communication utilizes manual or nonmanual signs. It includes gestures of head, hand, and neck. There is no all-inclusive gesture-based communication to be utilized by everybody, except portrayed by areas just as nations. A portion of the significant gesture-based communications that are generally well known in various areas of the world are Indian gesticulation Language (IGL), American gesticulation Language (AGL), French Gesticulation Language (FGL), and Japanese Gesticulation Language (GSL). The gesture-based communication utilized in India is regularly known as IGL. Even though it is accepted that a similar communication through signing is utilized in Nepal, Sri Lanka, Bangladesh, and fringe areas of Pakistan, the vernaculars of IGL may change by and large with expansive lexical variety and some linguistic structure that is found in various pieces of the Indian subcontinent. There are nearly 4 million hard-of-hearing individuals and according to the All India Federation of the Deaf, there are more than 10 million people in India who have hearing problems. The Indian gesture-based communication for English letters in order is as appeared in the underneath Figure 1.

Communication through signing helps individuals who have issues identified with voice and discover it hard to discuss ordinarily with others. Thus, this has created a need to structure a framework that permits hard-of-hearing individuals to discuss effectively with ordinary individuals.

One of the primary frameworks that empowered gesture-based communication correspondence between them was made conceivable by American Telephone and Telegraph video phone additionally called picture telephone which was brought in to showcase for open in 1964 at New York World's fair. It illustrated correspondence between a deaf person at reasonable and another deaf person in another city. Sign language helps individuals who have issues identified with voice and can't discuss typically with others. Henceforth, this has created a need to plan a framework that permits a deaf person to discuss effectively with typical individuals, since most typical individuals don't know about gesture-based communication framework.

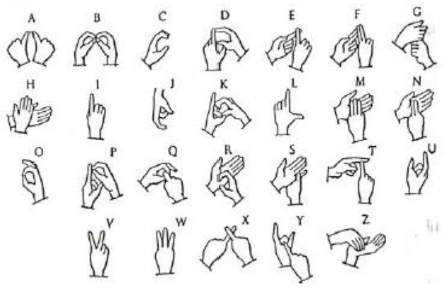


Fig. 1. Indian Sign Language symbols

The main strategies which intended to give gesticulation language recognition (GLS) were completely conceivable via information gloves next through the appearance of advanced online communication, picture handling methods, and advancements in foresight-based frameworks which received complex information science algorithms. Gesticulation language has different steps from gaining pictures to characterization till testing the results. The core objective of the proposed work is to make our database for Indian sign language (ISL) and develop a GUI-based recognition system.

The main contribution and objectives of this work are given below.

- We have developed the database for Indian Sign Language.
- We have developed an algorithm for the feature extraction of hand signals.
- We have developed an algorithm for classification.
- We have developed an algorithm for sign language recognition.
- We have to compare the proposed result with the existing result.

The rest of the manuscript is planned as follows in Section 2 explained the previous works associated with ISL, Section 3 explained in detail about proposed methodology. Section 4 explained the outcome of the suggested technique and discussion in various parameters. Section 5 is concluded our proposed work.

2. Literature Review

Communication through signing recognition can be essentially ordered into 2 strategies, for example (i) Glove Based and (ii) Vision-Based. Even though glove-based strategies have a high ground when it comes to hand division and highlight extraction since all the information is legitimately accessible from gloves and in addition glove-based strategy has a high likelihood of motion blames or flex sensor issues and so forth. Vision-based strategies are simpler to actualize as they are non-intrusive strategies, and the division and highlight extraction part should be possible utilizing picture handling with software. Hand division is a procedure to confine hands and other highlights from the remainder of the picture in vision-based frameworks. Many hand division techniques have been proposed in PC vision [1, 2]. To differentiate the edges of hands from a photograph, a Canny edge detector is used.

The canny edge detector is known for its ideal execution in recognizing edges and low error rate [3, 4]. The other technique for hand division is elliptical Fourier descriptors which are specific for separating the outline of shapes [5][6]. Feature extraction is utilized to secure highlights from the pictures caught. Background data, translation, scale, form, rotation, angle, coordinates, and movements are all included in this feature. Different techniques in feature extraction are linear discrimination analysis (LDA), edge detection, HSV (hue, saturation, value) edge detection, support vector machine, multi-dimensional Hidden Markov Model, Binary large object (BLOB) method, and Scale Invariant Feature Transform (SIFT) algorithm [7][8]. Classification is the last stage and a significant stage in motion acknowledgment.

Moreover, there are two kinds of gesture recognition approaches. A few analysts utilized the extracted features for signal acknowledgment, for example, format coordinating and some pre-owned AI classifiers, for example, HMM, Compressed Neural networks (CNN), K nearest node (KNN) [8, 9, 10]. At last, the outcomes are dissected for various signs and characters it is seen that HMM and neural system give higher acknowledgment proficiency contrasted with other techniques [7,11], such huge numbers of the works portrayed above-utilized mixture models to get higher acknowledgment rates to expand framework execution, these frameworks give dataset memory efficiencies by size decrease and utilized for movement and static signs. On account of the inclusion of programming examination and procedure the event of shortcomings and disappointments can be decreased because, in the vast majority of the huge frameworks, the equipment sub-framework will have the more propensity to get broken as opposed to the programming sub-framework because the attainability to intricacy proportion is higher in programming sub-frameworks than the equipment.

A. Using a Convolutional Neural Network with a Region of Interest

The framework was prepared to utilize 300 pictures of every Indian gesticulation language numeral caught utilizing an RGB model camera [12, 13]. The pictures are prepared on the graphic processing unit framework of the next version. The framework takes twenty-eight mins to prepare a prototype utilizing the 300 pictures of number signs. The framework was prepared to utilize a group size of 16 and starting learning rate of 0.001. The framework achieved 99.56% precision in 22 epochs [14]. The framework has tried different things with various taking in an amount varying from 0.01 and enactments have been refreshed through the preparation state. Abhishek Dudhal et al proposed a Hybrid SIFT feature extraction method to recognize the Indian sign language [19]. Sakshi et al. Proposed the Indian sign Language for publicly available datasets and got an accuracy of 92.43%[20]. Ashok et al. proposed automatic Indian sign language detection in real time atmosphere[21].

B. Sensor-Based System

This framework is a sensor-based signal acknowledgment framework that uses flex sensors for detecting hand developments [15]. To communicate with the normal person with hearing impaired persons use the ISL. Zigbee is working as transmitter and receiver and output should be displayed in the LCD and speech output come from the microcontroller [16].

In the proposed framework, bend sensors are utilized to quantify the level to which the fingers are twisted. The accelerometer inside the signal acknowledgment framework is utilized as a tilt detecting component, which consequently finds how much the finger is inclined. A tactile sensor is utilized to detect the physical association among the fingers [17]. The equivalent gestures are directed to the text to speech conversion module in the form of text. The result of text to speech amalgamation framework is heard using a speaker. The primary highlights of this framework remember its pertinence for everyday life, compactness, and minimal effort. Different Indian Sign Languages techniques approach for preprocessing techniques, feature extraction, Recognition is explained in detail and explained the various application of Indian sign language [19].

3. Methodology

Indian gesture-based communication is to presents a structure that can sense precise human gesticulation and utilize them to transfer info or for authority and manage functions. The productive path is intended to complete convolutional neural structures on the image in order to increase order productivity and for real-world application. The following are the major advancements, including the proposed framework.

Make a database for Indian sign-based communication.

- Make a model for training using CNN.
- Recognition
- Design the GUI.

A. Creation of database

Indian gesture-based communication is extraordinary and complex when contrasted with American gesture-based communication since Indian gesture-based communication incorporates the contribution of both hands however ASL is solitary. We have made a database for preparing furthermore, testing which incorporates 26 classes of 26 English letters in order. These pictures of Indian communication via gestures are taken with various foundations and from various people with fluctuating skin compositions and a model has appeared in Figure 2 below. We have a sum of 3658 pictures having a place with 26 classes.

Fig. 2 Example of database image creation

B. Convolutional Neural Network

Convolution architecture with different layers is used to generate the learning model, and the following layers are listed below:

i. 2D Convolution Layer

CNN comprises a convolution layer to distinct peaks from a data picture. It compresses the link among pixels by studying image details using tiny blocks of info. This is a scientific activity that receives two information sources viz. picture matrix and a frequency or part or filter. Contortion of a picture with several frequencies carries out actions viz., edge identification, blur, and sharpening by applying more filters honed by channels. A dot product is applied between a filter-sized patch of the input and the filter when the filter is smaller than the input data. The convolution of a filter-sized patch of the input and filter into a single value is known as a dot product. The activity is frequently referred to as the scalar item because it produces a single value. Solitary esteem is the result of duplicating the channel with the information cluster once. When the channel is applied to the information cluster multiple times, the result is a two-dimensional cluster of output esteems that indicate input sifting. As a result, the two-dimensional output of this procedure is referred to as a "feature map." The architecture of the 2-dimensional convolution neural networks is shown in Figure 3. One input layer, two convolution layers, two pooling layers, one hidden layer, and one output layer make up a 2D CNN.

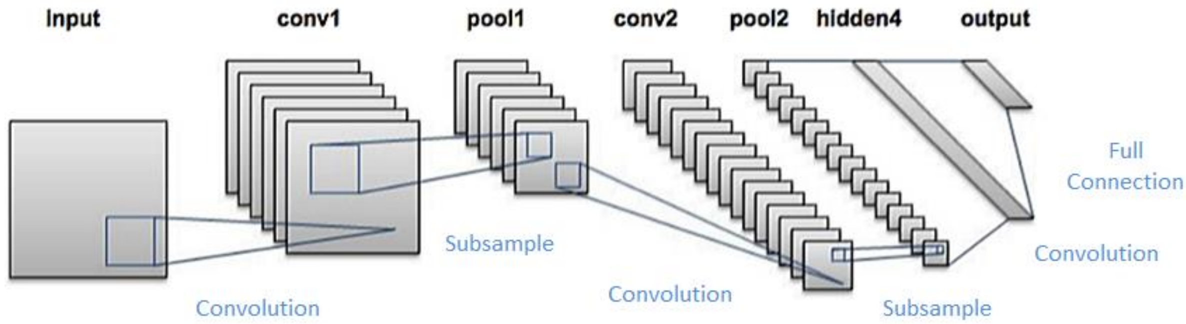


Fig. 3. Show the 2D Convolutional neural network

ii. Pooling layer

When the images are excessively large, the pooling layer minimizes the number of boundaries. Spatial pooling, also known as sub or downsampling, reduces the size of each guide while retaining critical information.

The largest element from the revised element map, as shown in Figure 4, is used in max-pooling. The average pooling could be obtained by taking the largest element. Sum pooling refers to the sum of all elements in an element map.

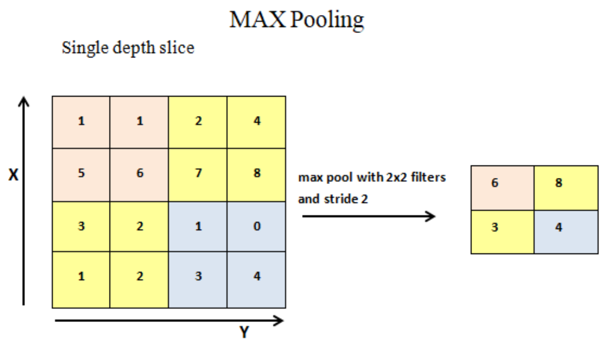


Fig. 4. Max pooling

iii. Flatten layer

As shown in Figure 5, a straightening layer transforms a two-dimensional matrix of data into a vector that can be fed into a completely associated neural system classifier.

Fig. 5. Flatten Layer operation

iv. Dense layer

A completely associated layer otherwise called the thick layer, wherein the outcomes of the convolution layers are taken care of through at least one neural layer to produce an expectation.

v. Dropout layer

Dropout refers to the act of neglecting units (such as neurons) during the preparation stage of a random arrangement of neurons. It means that these units aren't taken into account during a specific forward or backward pass. Individual hubs are dropped out of the net with probability 1-p or kept with probability p at each preparation stage, resulting in a reduced arrangement; approaching and active edges to a dropped-out hub are furthermore evacuated. Because a completely associated layer contains a considerable amount of the borders, neurons develop co-dependency during preparation, which checks the individual intensity of each neuron, resulting in over-fitting of data sets. The model kind that we will utilize is Sequential. Sequential is the least demanding approach to assemble a model in Keras. It permits you to assemble a model layer by layer. Our model comprises 4 convolution layers, 4 pooling layers, 2 thick layers, 1 flatten, and the dropout layer. The input layer is the primary layer, after which the 248 $\times$ 248 images are passed via a 2D convolution layer with a 3 x 3 x 3 kernel and 32 channels, as well as a rectified linear unit (relu) activation layer. Apply pooling of size 2 x 2 to the tangled image on the following layer, then reduce the size to 124 x 124. Following that, the size is reduced to 13 x 13 after passing through the four convolutional layers and four pooling layers. Between the convolutional and pooling layers, an actuation work called RELU is used.

Our entire image is treated as a multidimensional array in the convolutional layer, which employs convolution activity via a convolution matrix or kernel. Along with its loads, a convolution operation expands its nearby components. Every convolution layer in our model has a filter size of 3×3.

The pooling layer reduces the size of the picture component based on the pool size; in each picture, a single pixel from the specified cover is chosen. We've used a pooling size of 2 x 2, which will cut the image's real size in half when we move on to step 2. A max-pooling layer was used to perform the pooling. In the 2 x 2 bit, this action has the highest value. We used a rectified linear unit activation function to enable the positive values (RELU). It is known as a positive capacity since it strengthens desirable traits while returning zero for negative qualities, as demonstrated in Figure 6.

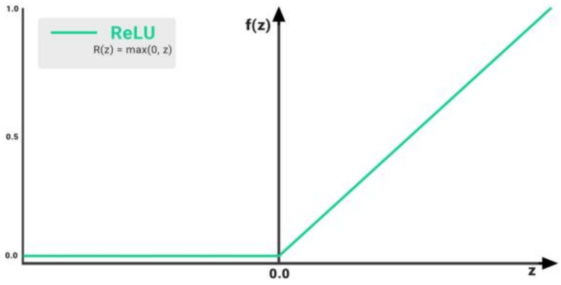


Fig. 6. Behavior of rectified linear unit function

Right now, the Rectified Linear Unit is the most often utilized initiation function in the domain. Then it's used in almost every CNN or deep learning system. Partially accurate is the Rectified Linear Unit (from the bottom). When z is less than 0, f(z) is nil, and when z is overhead or equal to zero, f(z) is equivalent to z. [0 to z] is the range. After all convolution and pooling layers, the framework is gone through the level layer. This layer converts the 2D image into a 1D array. Each layer is associated with the next layer with related loads which is known as completely associated layer or thick layer or dense layer. A dropout layer is added to avoid overfitting. The numerical model of the proposed CNN modular is as appeared in Figure 7.

$$f(z) = R(z) = \max(0, z) \tag{1}$$

$$f(z) = R(z) = \begin{matrix} 0, for\ z\ > 0 \\ z, for\ z \leq 0 \end{matrix} \tag{2}$$

This section explains the training model, testing model, accuracy curve, loss curve, number of epochs used, and GUI design discussed in detail. Finally, we compare the proposed result with the existing result.

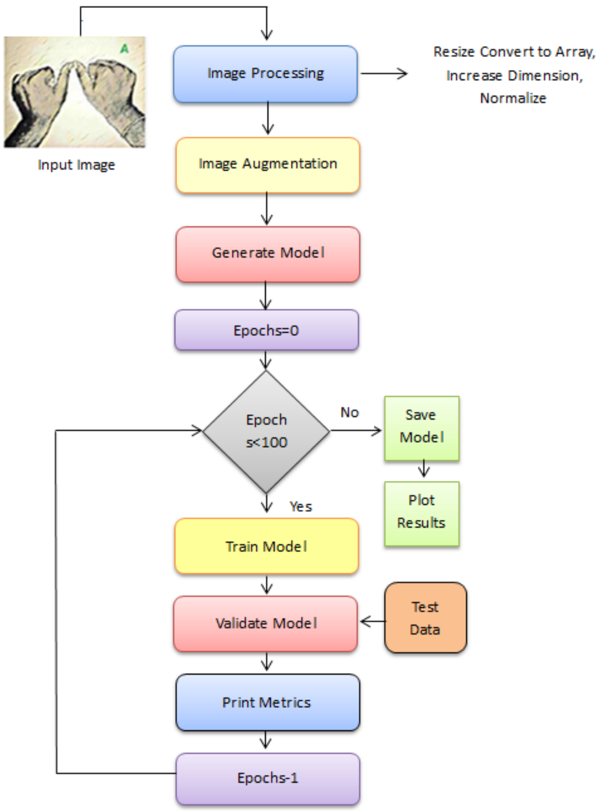


Fig. 7. Flow diagram of model training

Figure 7 shows the progression of information in preparing the CNN model and saving the model. It includes many steps such as image processing and augmentation, generating, training, validating, and saving the model.

Steps engaged with the stream chart:

1. Image Processing: Here the input pictures are resized to a specific measurement (250*250 in our case), converted to an array, measurement is expanded and pictures are standardized.
2. Augmentation: The pictures are expanded i.e, varieties of the pictures are created by inclining and flat moves. This increases the dataset.
3. Generate Model: The CNN model is produced with Epochs starting with zero.
4. Train and Validation*:* Here the model is prepared with each data and approved with test information for 100 epochs and measurements are printed each time.
5. *S*ave Model: The CNN model in the wake of being prepared is saved in the ".h5" document for use in expectation.
6. Plot Results: Here the diagram of precision and misfortune chart for preparing and approval are plotted.

Figure 8 shows the progression of anticipating the output letter. It comprises of basically 3 steps:

Image processing*:* Here the picture is enlarged, resized to 250 x 250, and changed over to array, at that point the measurement is extended, and toward the end it is normalized.

Load model: The saved model as ".h5" in the preparation is loaded to be used for acknowledgment.

Predict output: The output is anticipated dependent on the model produced previously, furthermore, the output class name is printed i.e., the letter corresponding to the class.

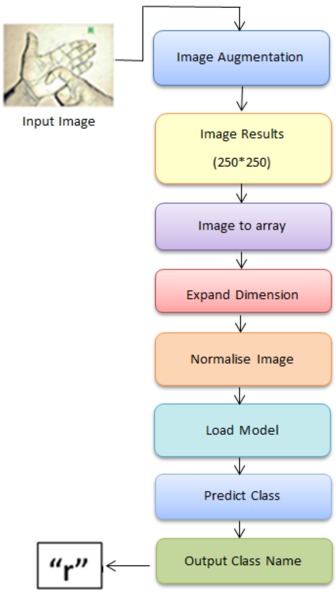
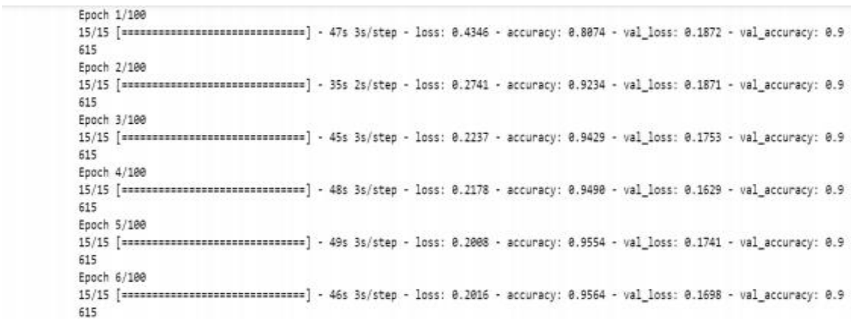


Fig. 8. Flow diagram of predicting the output

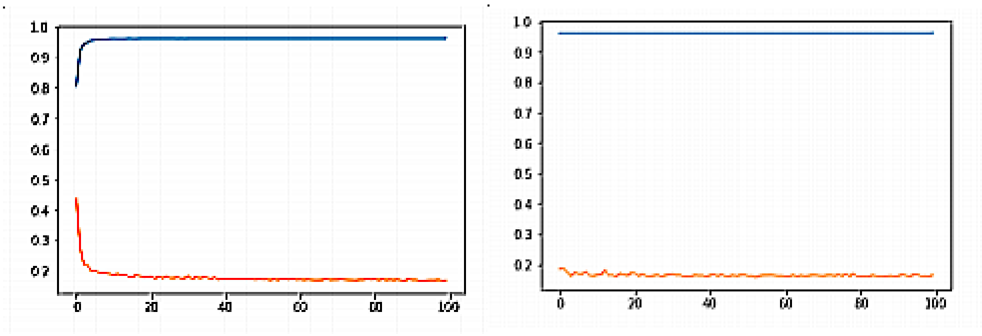
4. Result and Discussion

To prepare the model, we characterized 100 epochs. A complete presentation of the data set to be taught to a learning system is called an epoch. The number of epochs must be defined before a model can be built.



```
Epoch 1/100
15/15 [==============================] - 47s 3s/step - loss: 0.4346 - accuracy: 0.8074 - val_loss: 0.1872 - val_accuracy: 0.9
615
Epoch 2/100
15/15 [==============================] - 35s 2s/step - loss: 0.2741 - accuracy: 0.9234 - val_loss: 0.1871 - val_accuracy: 0.9
615
Epoch 3/100
15/15 [==============================] - 45s 3s/step - loss: 0.2237 - accuracy: 0.9429 - val_loss: 0.1753 - val_accuracy: 0.9
615
Epoch 4/100
15/15 [==============================] - 48s 3s/step - loss: 0.2178 - accuracy: 0.9490 - val_loss: 0.1629 - val_accuracy: 0.9
615
Epoch 5/100
15/15 [==============================] - 49s 3s/step - loss: 0.2008 - accuracy: 0.9554 - val_loss: 0.1741 - val_accuracy: 0.9
615
Epoch 6/100
15/15 [==============================] - 46s 3s/step - loss: 0.2016 - accuracy: 0.9564 - val_loss: 0.1698 - val_accuracy: 0.9
615
```

Fig. 9. Training and validation



(a) (b)

Fig.10. (a) Training Result, (b)Validation Result

In Figure 8, we can see the precision we got for the primary epoch is 80% and for the second epoch, it is 92%. At that point for the next epoch the exactness will continue expanding, furthermore, gets consistent at some level. If we attempt to train the model with more epochs, then it will result in overfitting. The accuracy and misfortune result chart for preparing and approval information is given below. Figure 9 displays the correctness curve of the proposed architecture and Figure 10 displays the validation arch of the architecture. The exactness for training and testing information is 96% as plotted in the above-mentioned charts.

A. GUI Design

A wonderful graphical user interface (GUI) is intended to make our venture user. This GUI is structured with assistance from PyQt5 python restricting library. In this GUI, the user ought to enter the way of the picture that he needs to check, click on approve. The outcome will be shown on GUI after approval. Figure 11 shows the front page of the Indian sign language recognition Graphical User Interface.

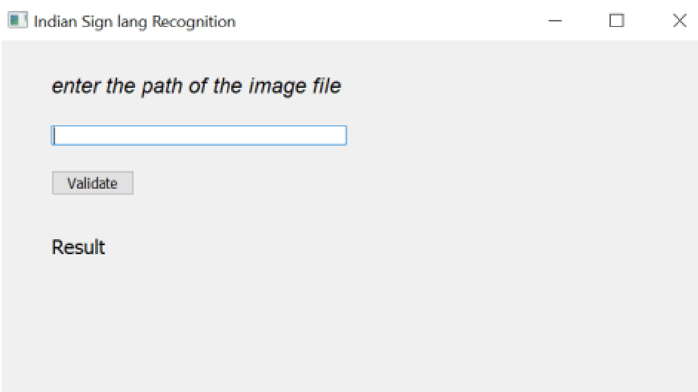


Fig. 11. Front page of GUI

Three of the main classes used to design the GUI:

1. Q Label
2. Q Line Edit
3. Q Push Button

Q Label: An object of Q Label can be utilized to show non-editable content or pictures. Utilizing the Q Label object, we showed some valuable writings.

Q Line Edit: An object of Q Line Edit gives a crate wherein one line of text can be entered. We utilized the object of Q Line Edit to give a line alert box to the user to enter the way of the picture he needs to approve.

Q Push Button: An object of Q Push Button presents a catch which when clicked can be modified to summon a specific capacity. We have characterized a catch called validate. After the user enters the way of the picture and taps on validation the outcome will be shown.

B. Recognition

An excellent graphical user interface (GUI) is intended to acknowledge Indian gesture-based communication intelligence. This GUI is structured with assistance from the PyQt5 python restricting library. In this GUI the user ought to enter the way of the picture that he needs to check and snap-on validate. The outcome will be shown on GUI after validation. At first, an object of the primary UI called the main window is characterized. When we are characterized with the object of the primary UI, we can resize the vibe of the main window and place traits inside. Three of the main classes used to design the GUI viz. Q Label, Q Lin edit, and Q Push Button

Q Label: An object of Q Label can be utilized to show non-editable content or pictures. Utilizing the Q Label object, we showed some valuable writings.

Q Line Edit: An object of Q Line Edit gives a crate wherein one line of text can be entered. We utilized the object of Q Line Edit to give a line alter box to the user to enter the way of the picture he needs to approve.

Q Push Button: An object of Q Push Button presents a button which when clicked can be modified to conjure a specific capacity. We have characterized a button called validate. After users enter the way of the picture and snaps on validating the outcome is as shown in Figure 12.

Figure 13 is the GUI intended for Indian gesture-based communication acknowledgment. The user ought to enter the way of the picture document which he needs to approve and click on to validate. The result will be shown on a similar GUI. As shown in Figure 16, Indian gesture-based communication letter set A is given as the information, and the result is shown as A. Similarly, the input image (Figure 14 and Figure 15) and validated results are demonstrated (Figure 16 and Figure 17).

Fig. 12. Image for Input

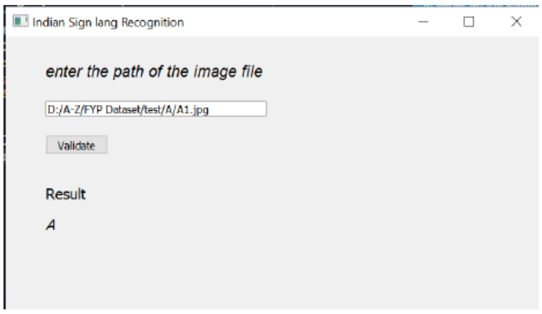


Fig. 13. Validated output

Fig. 14. Image for Input

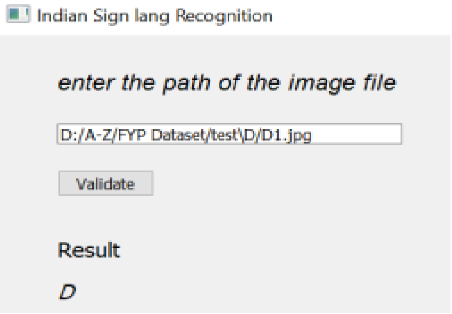


Fig. 15. Validated result

Fig. 16. Image for Input

Fig. 17. Validated result

Table 1 Comparison with au-courant methods using various parameters. The applications of the Indian sign language can be utilized in fingerspelled communication, smart gadgets for help with getting text/discourse information from signs and this technique gives a comprehension of communication through signing to normal people.

Table 1. The proposed method is compared to the existing method.

Comparison Parameters	2019[18]	Proposed method
Recognition	Includes 50 Gestures	Includes 26 English alphabets
Number of images in the dataset	5000 images, 100 images each of 50 gestures	3658 images
Pre-processing	Uses SIFT feature extraction technique to feed CNN	Resize, convert to the array, normalization
Number of convolution layers used	10- 2D Convolution layers used	4-2D convolutional layers
Number of epochs defined	25 epochs trained	100 epochs trained
Accuracy	92.78%	96.15%

5. Conclusion

The CNN profound learning framework has been designed for letter sets from A to Z as a foundation piece of Indian gesticulation language recognition. The framework has been set up to deal with the 4000 static RGB photos captured with a standard camera. For testing, the framework used 30 images for each image. The model was created by putting together a deep learning framework that used a region-based convolutional neural network. An accuracy of 96 percent has been achieved using the framework. Incorporate additional visuals from letters in place of static images of Indian gesture-based communication for words in the future. We can further build up this strategy to catch profundity as a component for acknowledgment in complex situations.

The dataset in our undertaking is accomplished for 26 English letters in order. The limitation of the proposed work is low-resolution images are not recognized and Real-time datasets are not recognized in the proposed system. In the future, it tends to be expanded to incorporate numerals and motions. We are doing acknowledgment of static images. In the future, it tends to be executed for moving motions which show the absolute most ordinarily utilized words in our day-by-day life. In the future, we can likewise expand the number of pictures per class (or letters in order) to build the precision of the framework.

References

[1] Mike Wu and Ravin Balakrishnan. 2003. multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In Proceedings of the 16th annual ACM symposium on User interface software and technology (UIST '03). Association for Computing Machinery, New York, NY, USA, 193–202. DOI: https://doi.org/10.1145/964696.964718.

[2] Didier Casalta, Yves Guiard, and Michel Beaudouin-Lafon. 1999. Evaluating two-handed input techniques: rectangle editing and navigation. In CHI '99 Extended Abstracts on Human Factors in Computing Systems (CHI EA '99). Association for Computing Machinery, New York, NY, USA, 236–237. DOI: https://doi.org/10.1145/632716.632862.

[3] Syed Mohammad Abid Hasan, Kwanghee Ko, Depth edge detection by image-based smoothing and morphological operations, Journal of Computational Design and Engineering, Volume 3, Issue 3, July 2016, Pages 191–197, https://doi.org/10.1016/j.jcde.2016.02.002

[4] Tongping Shen, Fangliang Huang, and Li Jin. 2019. An Improved Edge Detection Algorithm for Noisy Images. In Proceedings of the 2019 International Conference on Artificial Intelligence and Computer Science (AICS 2019). Association for Computing Machinery, New York NY, USA, 84–88. DOI: https://doi.org/10.1145/3349341.3349373.

[5] Nanda Maulina Firdaus and Erdefi Rakun. 2019. Recognizing fingerspelling in SIBI (sistem isyarat bahasa Indonesia) using OpenPose and elliptical fourier descriptor. In Proceedings of the International Conference on Advanced Information Science and System (AISS '19). Association for Computing Machinery, New York, NY, USA. Article 14, 1–6. DOI: https://doi.org/10.1145/3373477.3373491

[6] Sener, S., Unel, M. Affine invariant fitting of algebraic curves using Fourier descriptors. Pattern Anal Applic 8, 72–83 (2005). https://doi.org/10.1007/s10044-005-0245-6

[7] Cheok, M.J., Omar, Z. & Jaward, M.H. A review of hand gesture and sign language recognition techniques. Int. J. Mach. Learn. & Cyber. 10, 131–153 (2019). https://doi.org/10.1007/s13042-017-0705-5.

[8] Ashutosh Samantaray, Sanjaya Kumar Nayak, Ashis Kumar Mishra"Hand Gesture Recognition using Computer Vision" International Journal of Scientific & Engineering Research, Volume 4, Issue 6, June 2013 ISSN 2229-5518

[9] Jung In Koh, Josh Cherian, Paul Taele, and Tracy Hammond. 2019. Developing a Hand Gesture Recognition System for Mapping Symbolic Hand Gestures to Analogous Emojis in Computer-Mediated Communication. ACM Trans. Interact. Intell. Syst. 9, 1, Article 6 (February 2019), 35 pages. DOI: https://doi.org/10.1145/3297277

[10] Yale Song, David Demirdjian, and Randall Davis. 2012. Continuous body and hand gesture recognition for natural human-computer interaction. ACM Trans. Interact. Intell. Syst. 2, 1, Article 5 (March 2012), 28 pages. DOI: https://doi.org/10.1145/2133366.2133371

[11] Wei Huang and Yuting Zhang. 2020. Application of Hidden Markov Chain and Artificial Neural Networks in Music Recognition and Classification. In Proceedings of 2020 the 6th International Conference on Computing and Data Engineering (ICCDE 2020). Association for Computing Machinery, New York, NY, USA, 49–53. DOI: https://doi.org/10.1145/3379247.3379276

[12] Bindu Verma and Ayesha Choudhary. 2018. Dynamic Hand Gesture Recognition using Convolutional Neural Network with RGB-D Fusion. In Proceedings of the 11th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP 2018). Association for Computing Machinery, New York, NY, USA, Article 67, 1–8. DOI: https://doi.org/10.1145/3293353.3293421

[13] Shanta, Shirin Sultana et al. “Bangla Sign Language Detection Using SIFT and CNN.” 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (2018): 1-6.

[14] Lamptey K. Odartey, Yonfeng Huang, Effah E. Asantewaa, and Promise R. Agbedanu. 2019. Ghanaian Sign Language Recognition Using Deep Learning. In Proceedings of the 2019 the International Conference on Pattern Recognition and Artificial Intelligence (PRAI '19). Association for Computing Machinery, New York, NY, USA, 81–86. DOI: https://doi.org/10.1145/3357777.3357784

[15] N. Krishnaraj, M. G.Kavitha, T. Jayasankar, K. Vinoth Kumar "A Glove Based Approach to Recognize Indian Sign Languages" International Journal of Recent Technology and Engineering (IJRTE), Volume-7, Issue-6, March 2019ISSN: 2277-3878

[16] Pengfei Lu, Matt Huenerfauth, Collecting a Motion-Capture Corpus of American Sign Language for Data-Driven Generation Research, Proceedings of the NAACL HLT 2010 Workshop on Speech and Language Processing for Assistive Technologies, pages 89–97, Los Angeles, California, June 2010.
[17] Markus Fritzsche, Norbert Elkmann, and Erik Schulenburg. 2011. Tactile sensing: a key technology for safe physical human robot interaction. In Proceedings of the 6th international conference on Human-robot interaction (HRI '11). Association for Computing Machinery, New York, NY, USA, 139–140.
[18] Dudhal A., Mathkar H., Jain A., Kadam O., Shirole M. (2019) Hybrid SIFT Feature Extraction Approach for Indian Sign Language Recognition System Based on CNN. In: Pandian D., Fernando X., Baig Z., Shi F. (eds) Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB). ISMAC 2018. Lecture Notes in Computational Vision and Biomechanics, vol 30. Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-00665-5_72.
[19] Mohammed. Safeel, Tejas. Sukumar, Shashank. K. S, Arman. M. D, Shashidhar. R and Puneeth. S. B, "Sign Language Recognition Techniques- A Review," 2020 IEEE International Conference for Innovation in Technology (INOCON), Bangluru, India, 2020, pp. 1-9, doi: 10.1109/INOCON50539.2020.9298376.
[20] John J., Sherif B.V. (2022) Hand Landmark-Based Sign Language Recognition Using Deep Learning. In: Chen J.IZ., Wang H., Du KL., Suma V. (eds) Machine Learning and Autonomous Systems. Smart Innovation, Systems and Technologies, vol 269. Springer, Singapore. https://doi.org/10.1007/978-981-16-7996-4_11.
[21] Ashok kumar Sahoo "Indian Sign Language Recognition Using Machine Learning Techniques" Macromolecular Symposia, 2021. https://doi.org/10.1002/masy.202000241

Authors' Profiles

Prof. Shashidhar R obtained his Master of Technology in VLSI Design and Embedded System from Visvesvaraya Technological University, Belagavi in the year 2014. He is currently pursuing his Ph.D. in JSS S&TU, Mysuru. He is currently working as an Assistant Professor in the Department of Electronics and Communication Engineering, JSS Science & Technology University, Sri Jayachamarajendra College of Engineering, Mysuru. Before this, he also worked at Dayananda Sagar College of Engineering, Bangalore. He teaches Both UG and PG students. His area of interest is Advanced Embedded systems, Signal Processing, Machine Learning, Artificial Intelligence and he published several reputed Journals and attended several conferences.

Dr. Arunakumari B N Completed Ph.D. from Visveswaraya Technological University, Belagavi, Currently Working as Assistant Professor in the Department of Computer science and Engineering at BMSIT, Bengaluru. She published four journals and attended seven International conferences. Her research interest is discrete mathematical Structure, Unified Modeling Language, and Artificial Intelligence.

Prof. A S Manjunath obtained his Master of Computer application from Mysuru University, Mysuru. He is currently pursuing his Ph.D. in V.T.U, Belagavi. He is currently working as an Assistant Professor in the Department of Computer applications, JSS Science & Technology University, Sri Jayachamarajendra College of Engineering, Mysuru. He published Scopus indexed Journals and Conferences. His area of interest is Image Processing, Pattern Recognition, Machine learning.

Dr. Roopa M currently working as Assistant Professor, Department of Electronics and Communication, Dayananda Sagar College of Engineering, Bengaluru, she published several articles in reputed journals and also conferences. Her area of interest is Embedded systems, Digital System Design, and Artificial Intelligence.

How to cite this paper: Shashidhar R, Arunakumari B. N., A S Manjunath, Roopa M, " Indian Sign Language Recognition Using 2-D Convolution Neural Network and Graphical User Interface", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.14, No.2, pp. 61-73, 2022.DOI: 10.5815/ijigsp.2022.02.06