



(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 4, April 2025

⊕ www.ijirset.com ⊠ ijirset@gmail.com 🖄 +91-9940572462 🕓 +91 63819 07438





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

DOI: 10.15680/IJIRSET.2025.1404454

Automated Sign Language Recognition with Machine Learning

Prashanthi Regonda, Eppili Jatin, Attem Varun Yadav, Javadi Adarsh Kumar

Associate Professor, Department of Computer Engineering, Sreenidhi Institute of Science and Technology, JNTUH,

Hyderabad, India

U.G. Student, Department of Computer Engineering, Sreenidhi Institute of Science and Technology, JNTUH,

Hyderabad, India

U.G. Student, Department of Computer Engineering, Sreenidhi Institute of Science and Technology, JNTUH,

Hyderabad, India

U.G. Student, Department of Computer Engineering, Sreenidhi Institute of Science and Technology, JNTUH,

Hyderabad, India

ABSTRACT: The purpose of this research is to design a system using machine learning to automate sign language gesture recognition and facilitate unproblematic interaction between the hearing/speech impaired community and society as a whole. This research works towards solving the communication disability faced by the deaf and mute community through a computer vision framework for recognizing hand gestures and mapping them to relevant text or speech. The work employs a rich dataset of static American Sign Language (ASL) hand signs and uses Convolutional Neural Networks (CNNs) for efficient feature extraction and gesture classification. The suggested model is trained and evaluated on a huge collection of labeled gesture images and exhibits high accuracy with robust generalization to novel data. One of the primary contributions of the work is the implementation of the trained model into an interactive, real-time graphical user interface developed with Python's Tkinter library. The interface captures live input from a webcam, processes each frame to identify gestures, and shows the identified character on the screen. The results of the evaluation confirm the ability of CNNs to accurately identify ASL alphabets with little delay. The proposed system presents an effective and scalable approach towards real-time sign language interpretation, supporting inclusiveness and accessibility of the disabled community. The present research also creates a scope for extension in development, i.e., inclusion of dynamic gesture detection and portability to mobile or embedded platforms to provide enhanced use in real-life situations.

KEYWORDS: Sign Language Recognition, Convolutional Neural Networks, American Sign Language, Real-Time Gesture Detection, Human-Computer Interaction

I. INTRODUCTION

Sign language is the primary means of communication for individuals with speech and hearing impairments. With limited public proficiency, they often rely on interpreters, who are not always available or affordable. An automated Sign Language Recognition (SLR) system would be able to effectively translate gestures, facilitating communication barriers[1] and improving social integration. SLR systems require huge data to translate sign languages appropriately, which consist of extremely intricate regional variation of grammar and vocabulary. It becomes difficult to implement a global model because of its complexity, and localized SLR models are more effective. SLR systems transform Human-Computer Interaction using gesture recognition, bridging sign language speakers to society. Machine learning methods for creating these systems are presented in this paper.

II. INTELLIGENT SIGN LANGUAGE RECOGNITION USING MACHINE LEARNING

An intelligent Sign Language Recognition (SLR) system identifies hand gestures, facial expressions, and posture to interpret signs. As a signer stands before a camera, the gestures of the signer are recognized by the interpreter model to





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404454|

deliver messages to non-signers. The model requires extensive pre-training over many datasets for high accuracy. Image size can differ depending on the concept. Figure 1 shows the system processing input through stages to recognize the sign. The SLR model has difficulty with the irregular hand movement, occlusion, changing signing speeds, and non-uniform lighting, which impair the accuracy of recognition.





A. Preprocessing

Preprocessing is the first step in a Sign Language Recognition (SLR) system, wherein image quality is enhanced by the elimination of blurring, noise, and lighting problems. The accuracy of recognition is increased by brightness correction and conversion to grayscale to better define edges and contours for gesture analysis. J. Singha and A. Karen Das (2013)[4] used three preprocessing techniques: skin segmentation, eigenvector extraction, and classification using the Euclidean Distance metric. Illumination[5] variation is one of the most significant issues in SLR systems that influence performance. In order to address this, R. Gross and V. Brajovic implemented an algorithm[5] that constructs a brightness field from a reference frame in order to normalize brightness under varying lighting, which improves recognition accuracy.









www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404454|

B. Segmentation

Segmentation in Sign Language Recognition (SLR) separates an image into tiny sub-regions according to shared features such as texture or color to derive meaningful information. Sub-regions are the original image and are most crucial for effective feature extraction. Hand segmentation in SLR is a significant process because hand movement supplies essential information during gesture recognition. Traditional segmentation methods are color-based, thresholding, and edge-based segmentation depending on the type of image and processing requirements. 2D grayscale and 3D RGB images are both segmentable to identify boundaries. Different segmentation techniques and their effectiveness in separating hand movement for proper recognition are discussed in this paper.

Colour Based segmentation

Segmentation of the hand is significant in hand recognition for SLR systems and can be accomplished using techniques such as color-based segmentation, thresholding, and edge detection. A glove-free method of segmentation was proposed by Paulraj et al.,[8] optimized through instructions given to the signers on what clothing to wear. Tofighi et al. suggested an adaptive histogram template with skin color and texture for effective hand extraction, though it does not perform well with complicated backgrounds. Pansare and Ingle's[11] framework includes noise elimination, morphological processing, skin detection, and grayscale thresholding for proper hand segmentation. They also created a skin segmentation model based on SVM active learning and Kalman filters for better tracking. Zieren and Kraiss[14] developed a person-independent system based on statistical models of color for skin detection, but it is troubled by cluttered backgrounds, which may be overcome by incorporating a probability threshold for different skin colors and lighting.

Edge-based segmentation

Threshold-based segmentation works well under optimal lighting but can cause size biases and edge distortions. Edgebased segmentation improves feature extraction but struggles with noise, leading to incorrect or missed edges[15]. To address this, newer SLR models use advanced edge detection. Hemayed[16] and Hassanien's Arabic Sign Language Recognition model uses the Prewitt edge detector to preserve important edge details for better gesture distinction. Similarly, Prasad et al.'s[17][18] hybrid model combines morphological subtraction and Canny edge detection to refine hand and face segmentation and improve recognition accuracy.

C. Feature Extraction

Preprocessed images have high dimensionality, and hence classification is expensive. Feature extraction decreases dimensionality without sacrificing important information and produces a concise data representation that improves classification efficiency and prediction accuracy. Feature selection chooses relevant features but might lose useful information. Feature extraction retains useful information and reduces dimensions. Features of images are termed as important, useless, or redundant[19] and highlight the most informative to enhance recognition and performance.

PCA

PCA is a prevalent unsupervised feature extraction method used extensively, ideal to handle redundant data. PCA converts high-dimensional data into lower-dimensional space retaining the significant variations and patterns. PCA makes dimensionality reduction possible by eliminating the less significant principal components, enhancing computational efficiency and classifying performance by focusing on the significant features.

Convex Hull

Convex hull is an important image processing method applied for its flexibility and strength, particularly applied in sign language recognition (SLR). It finds and links the most distant points of an object to create a convex area[20] employed for subsequent analysis. In SLR systems, convex hull is employed to track finger and palm orientations important for hand gesture recognition[21]. By defining hand features such as fingertips and edges, it assists in correct gesture recognition, with each finger processed individually to enhance model accuracy.





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025 |DOI: 10.15680/IJIRSET.2025.1404454|

D. Classification

Classification entails the examination of images and the classification of these images under predefined labels. In SLR systems, it discovers the appropriate sign from a list of gestures and anticipates new signs based on established models. Classifiers differ through learning methods—some are statistical models, others deep learning. Supervised classifiers such as SVM, Decision Trees, and ANN employ labeled data, while unsupervised classifiers such as K-Means cluster signs without labels. Deep learning algorithms like CNNs automate feature extraction and enhance accuracy by identifying intricate patterns, allowing real-time recognition. The classifier choice relies on dataset size, computational power, and gesture complexity. Good classification enhances SLR systems, facilitating bridging of communication gaps for the hearing and speech impaired.

Lazy learner classifier

Lazy learners don't build a model during training but store all data and classify new inputs by comparing them to past instances. KNN, a common lazy learner, assigns labels based on the majority of nearest neighbors. Case-Based Reasoning (CBR) reuses similar past cases to solve new ones. Though slow at prediction time, lazy learners are flexible and adapt easily to new data without retraining.



Figure 3 - Hand signs used in ASL

K- Nearest Neighbour

KNN is a general-purpose classification algorithm and can be applied to regression too. KNN classifies images based on similarity of feature vectors by means of distance metrics such as Euclidean, Cosine, Minkowski, or Chi-Square. In multimodal applications, KNN proves to be highly effective. KNN creates clusters by voting by the K-nearest neighbors by distance. Computational cost can be reduced with a subset dataset or fast-KNN. KNN is good to use in image-based applications using local features[26].

Eager Learning Classifiers

Eager learners construct a generic model from training data prior to exposure to test data, enabling quick predictions but using more resources and time at training. They are different from lazy learners, who can have difficulties with handling new or changing data. Popular eager classifiers are Decision Trees, SVMs, ANNs[27], and Naïve Bayes. They are suitable for real-time applications such as medical diagnosis and fraud detection but are very dependent on pre-trained data, which can hamper adaptability.





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404454|

Support Vector Machines

Support Vector Machines (SVMs) find extensive applications in object and image recognition, particularly in cluttered environments[28]. SVMs operate on the principle of identifying best hyperplanes[29][30] that differentiate data classes. Training with a balanced dataset of positive and negative samples increases precision. SVMs are robust against noise and function well in tasks such as hand and face detection, even in diverse lighting and backgrounds. They also categorize color images in formats such as RGB and CMYK, where color histograms[32] assist in extracting patterns for precise recognition.

Decision Tree

A decision tree classifier is a tree model that divides data by a divide-and-conquer strategy, creating branches until a stopping criterion—such as max depth or no improvement in performance—is reached. It begins with a root node selected by information gain and proceeds to split according to feature values until leaf nodes with final choices are reached. While useful, decision trees have the tendency to overfit training data. This can be solved by employing pre-pruning (depth limiting) or post-pruning (removing redundant branches) to enhance generalization and performance in the real world.

Naïve Bayes

Naïve Bayes classifier is a probabilistic model based on Bayes' theorem under feature independence assumption for simplification. It is therefore efficient and scalable and suited for use in applications such as sentiment analysis, text classification, and spam filtering. However, it is plagued by the zero probability problem where attributes with zero probability cannot be predicted. This can be solved using Laplace smoothing, which ensures probabilities are never zero.

ANN

Artificial Neural Networks (ANN) are employed for pattern recognition, such as sign language recognition (SLR). They have an input layer, hidden layer(s), and output layer. The hidden layer acquires features by changing the weights and biases while training. When complexity grows, more hidden layers are included, resulting in deep learning models. This enhances accuracy but also rises in computational cost, training time, and complexity.

Experimental setup

The SLR model was tested with four candidates performing hand gestures under controlled conditions with stable background and lighting. Each candidate performed gestures for numbers 1 to 5 five times for consistent data. The preprocessing phase involved background removal and contour-based segmentation, followed by feature extraction using the convex hull method. Gesture classification was done with K-Nearest Neighbors (KNN) and Euclidean Distance. The model's performance was evaluated based on correct gesture identification, showing higher accuracy for numbers 1 and 2. The overall accuracy was 65%, indicating potential for improvement in feature extraction and classification methods for better performance with other gestures.





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025



|DOI: 10.15680/IJIRSET.2025.1404454|

Figure 5 - High level DFD for the system

Candidate	Numbers					
	0	1	2	3	4	5
А	3	4	4	3	3	2
В	4	4	4	2	2	3
С	3	5	3	4	3	3
D	2	4	4	4	3	2

TABLE 1: SUCCESSFUL ATTEMPTS OF EACH CANDIDATE

III. CONCLUSION AND FUTURE WORK

For the computer SLR system to be effective in usability, it must have the capability to detect gestures in real-world contexts. This paper provides a summary of a machine learning-based SLR model, where various algorithms are employed at each stage to extract maximum information with minimal computational expense. The study applied the convex hull method as the feature extraction approach, resulting in an accuracy rate of 75%. Enhancements in performance could be achieved by leveraging a larger dataset and trying out different classifiers. Although the model





www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 4, April 2025

|DOI: 10.15680/IJIRSET.2025.1404454|

currently focuses on identifying and recognizing individual signs, future developments could allow for the recognition of continuous sign language, thus increasing its practicality for real-world applications.

REFERENCES

- 1. B. L. Loeding, S. Sarkar, A. Parashar, and A. I. Karshmer, "Progress in automated computer recognition of sign language," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 3118, 2004, doi: 10.1007/978-3-540-27817-7_159.
- 2. J. Singha and A. Karen Das, "Recognition of Indian Sign Language in Live Video," 2013, arXiv preprint arXiv:1306.1301.
- R. Gross and V. Brajovic, "An Image Preprocessing Algorithm for Illumination Invariant Face Recognition, 3. presented at the International Conference on Audio-and Video-Based Biometric Person Authentication (pp. 10-18). Springer, Berlin, Heidelberg..
- M. P. Paulraj, S. Yaacob, H. Desa, C. R. Hema, and W. M.R. Wan Ab Majid, "Extraction of head and hand 4. gesture features for recognition of sign language," 2008 Int. Conf. Electron. Des. ICED 2008, no. January, 2008, doi: 10.1109/ICED.2008.4786633.
- 5. J. Ayshree and R. Pansare, "Vision-Based Approach for American Sign Language Recognition Using Edge Orientation Histogram."2016 International Conference on Image, Vision and Computing (ICIVC) (pp. 86-90). IEEE.
- J. Zieren and K.-F. Kraiss, "Robust Person-Independent Visual Sign Language Recognition." presented at Iberian 6. Conference on Pattern Recognition and Image Analysis (pp. 520-528). Springer, Berlin, Heidelberg, 2005.
- W.-X. Kang, Q.-O. Yang, and R.-P. Liang, "The Comparative Research on Image Segmentation Algorithms," 7. presented at First International Workshop on Education Technology and Computer Science (Vol. 2, pp. 703-707). IEEE. 2009, doi: 10.1109/ETCS.2009.417.
- E. E. Hemayed and A. S. Hassanien, "Edge-based recognizer for Arabic sign language alphabet (ArS2V- Arabic 8. sign to voice)," in ICENCO'2010 - 2010 International Computer Engineering Conference: Expanding Information Society Frontiers, 2010, doi: 10.1109/ICENCO.2010.5720438.
- M. V.D. Prasad, P. V. V. Kishore, E. Kiran Kumar, and D. Anil Kumar, "Indian sign language recognition system 9. using new fusion based edge operator," J. Theor. Appl. Inf. Technol., vol. 88, no. 3, 2016.
- 10. J. Serra, "Image analysis and mathematical morphology," by Academic Press, London, 1982, Cytometry, vol. 4, no. 2, 1983, doi: 10.1002/cyto.990040213.
- 11. S. S. Rautaray and A. Agrawal, "Vision based hand gesture recognition for human computer interaction : a survey," Artificial intelligence review pp. 1-54, 2015, doi: 10.1007/s10462-012-9356-9.
- 12. M. A. Jayaram and H. Fleyeh, "Convex Hulls in Image Processing: A Scoping Review," American Journal of Intelligent Systems, vol. 6, no. 2, pp. 48–58, 2016, doi: 10.5923/j.ajis.20160602.03.
- 13. Y. Wen, C. Hu, G. Yu, and C. Wang, "A robust method of detecting hand gestures using depth sensors," in Proceedings - 2012 IEEE Symposium on Haptic Audio- Visual Environments and Games, HAVE 2012, 2012, pp. 72-77, doi: 10.1109/HAVE.2012.6374441.
- 14. L. Y. Hu, M. W. Huang, S. W. Ke, and C. F. Tsai, "The distance function effect on k-nearest neighbor classification for medical datasets," Springerplus, vol. 5, no. 1, Dec. 2016,
- 15. doi: 10.1186/s40064-016-2941-7.
- 16. S.Rafatirad and M Heidari, "An exhaustive analysis of lazy vs. Eager learning methods for real-estate property investment."2018.
- 17. Praveen Kumar Maroju, Venu Madhav Aragani (2025). Predictive Analytics in Education: Early Intervention and Proactive Support With Gen AI Cloud. Igi Global Scientific Publishing 1 (1):317-332.
- 18. S. Bernhard, "SVMs-a practical consequence of learning theory.", Corpus ID: 2526394,1998.
- 19. O. Chapelle, P. Haffner, and V. N. Vapnik, "Support Vector Machines for Histogram-Based Image Classification," IEEE transactions on Neural Networks, 10(5), pp.1055-1064. 1999.
- V. Vapnik, "The Nature of Statistical Learning Theory.", Springer science & business media,2013
 S. Agrawal, N. K. Verma, P. Tamrakar, and P. Sircar, "Content Based Color Image Classification using SVM," presented at Eighth International Conference on Information Technology: New Generations (pp. 1090-1094). IEEE. 2011, doi: 10.1109/ITNG.2011.202.









INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY





www.ijirset.com