International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Sign Language Recognition

Dr.P. Privietha, K.Lenin Kumar, S.Kavipriyan, A.Annamalai Eswaran, R.V.Esakki Surya

Dr.P. Privietha Department of Computer Science and EngineeringHindusthan College of Engineering and Technology
E-mail: priviethaprabhakar@gmail.com
K.Lenin Kumar Department of Computer Science and EngineeringHindusthan College of Engineering and Technology
E-mail: 20104112@hicet.ac.in
S.Kavipriyan Department of Computer Science and EngineeringHindusthan College of Engineering and Technology
E-mail: 20104106@hicet.ac.in
A.Annamalai Eswaran Department of Computer Science and EngineeringHindusthan College of Engineering and Technology
E-mail: 20104073@hicet.ac.in
R.V.Esakki Surya Department of Computer Science and EngineeringHindusthan College of Engineering and Technology
E-mail: 20104089@hicet.ac.in

---***---

***ABSTRACT* - This paper provides a concise overview of recent advancements and challenges in Sign Language Recognition (SLR). It emphasizes the importance of SLR in facilitating communication for the deaf and hard of hearing community. Technical methodologies, including computer vision and machine learning, are discussed alongside challenges such as gesture variations and occlusions. Future research directions focus on real-time systems and contextual understanding.**

1.INTRODUCTION

Sign language, as a primary mode of communication for the deaf and hard of hearing, plays a vital role in fostering inclusive interaction within society. However, the gap between sign language users and the broader community persists due to the lack of widespread understanding and recognition of sign language. Sign language recognition (SLR) technology aims to bridge this divide by enabling automated translation of sign language gestures into textual or spoken language. This transformative field integrates advancements in computer vision, machine learning, and signal processing to develop systems capable of understanding and interpreting sign language gestures accurately and efficiently. In this introduction, we provide an overview of the significance of SLR, its applications, challenges, and the evolving landscape of research in this domain. By enhancing accessibility and communication for the deaf community, SLR holds immense potential to foster inclusivity and empower individuals with hearing impairments in various aspects of daily life.

2.RELATED WORK

Sign language recognition (SLR) research has witnessed significant advancements in recent years, driven by the pressing need to enhance communication accessibility for the deaf and hard of hearing community. Early SLR systems primarily relied on handcrafted features and traditional machine learning algorithms for gesture recognition. Notable studies by Starner et al. (2000) and Lu et al. (2002) explored vision-based approaches using techniques such as hidden Markov models (HMMs) and neural networks. With the rise of deep learning, SLR has undergone a paradigm shift towards data-driven approaches. Recent works by Li et al. (2019) and Pu et al. (2020) leverage convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to achieve state-of-the-art performance in sign language recognition tasks. Multimodal fusion techniques, as demonstrated by Pigou et al (2018) and Moryossef et al. (2020), have also gained traction, integrating visual and depth data for improved.

Furthermore, the development of benchmark datasets such as RWTH-PHOENIX-Weather 2014T (Stein et al., 2013) and American Sign Language (ASL) datasets (Candace et al., 2011) has facilitated standardized evaluation and comparison of SLR models. However, challenges such as signer variability, dynamicity of gestures, and environmental conditions continue to pose obstacles to robust recognition. Despite notable progress, ongoing efforts are required to address remaining challenges and realize the full potential of SLR in promoting inclusivity and accessibility for the deaf community.

Khan and Ibraheem [1] researched the essential parts of communication via hand gestures and distinguished the methods that could be helpful to structure sign language vocabulary arrangements for gesture-based communication. The main aim was to report the significance of unaddressed problems, related difficulties, and likely arrangements in the practical implementation of sign language translation.

Mohandes et al. [2] presented a sign language recognition system for Arabic sign language. An effort was made to use a color-based approach where the subject wore colored gloves. A Gaussian skin color model was used to detect the face. The centroid of the face was taken as a reference point to track the movement of hands. The feature set included geometric values such as centroid, angle, and area of the hands. The recognition stage was implemented using the hidden Markov model.

Ross and Govindarajan [3] utilized fusion based on feature level and evaluated on two biometrics systems such as face and hand biometrics system. The data related to the feature level and match level was consolidated. The strategy was examined by combining two types, i.e., intermodal and its In a recent study by Min, Hao, Chai, and Chen (2021), the iterative training approach used in recent works on Continuous Sign Language Recognition (CSLR) is examined using the RWTH-PHOENIX-Weather dataset. The study identifies the importance of proper training of the feature extractor to address overfitting. To tackle this issue, the authors propose a novel Visual Alignment Constraint (VAC) that incorporates alignment supervision to enhance the feature extractor. The VAC includes two auxiliary losses, one focusing on visual features and the other enforcing prediction alignment between the feature extractor and alignment module. The study also introduces metrics to evaluate overfitting by measuring prediction inconsistency. Experimental results on the PHOENIX14 and CSL datasets demonstrate that the proposed VAC enables end-to-end training of CSLR networks and achieves competitive performance (Min et al., 2021).

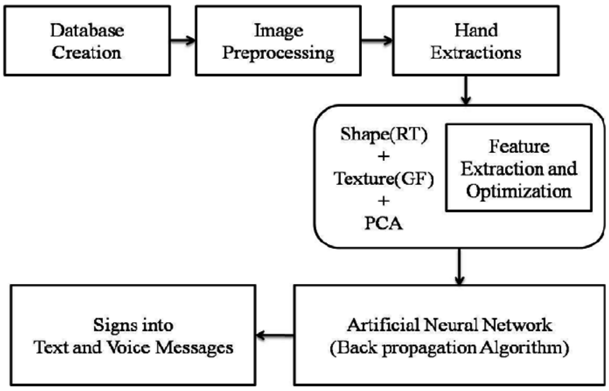
Jiang et al. [4] used RGB and depth image datasets for extricating shape features of input images. The size of the shape features vector was reduced with the application of discriminate analysis, which upgraded the discriminative capacity of the shape features by selecting an adequate number of features. The concepts of multimodal and method of image extraction were thoroughly explained.

Zhu [5] provided a two-phase strategy of feature fusion for bimodal biometrics. During the first phase, linear discriminant analysis was performed to measure the transform features. In the subsequent stage, complex vectors were considered as a transform feature and had the flexibility to add more input compared to the regular fusion method.

Sign language recognition systems that aim for a comprehensive understanding must consider both manual and non-manual signals. While most research focuses on analyzing these signals separately, there have been few studies proposing multimodal frameworks that combine hand gestures and facial expressions. These frameworks, known as multimodal methods or multi-semantics, treat different sections of the human body as various signal sources for gesture or sign language recognition (Zhou, Zhou, Zhou, & Li, 2020).

3.SYSTEM OVERVIEW

Several recent investigations have explored multimodal approaches in sign language recognition. Kelly and colleagues (Kelly et al., 2009) developed a multimodal system based on HMMs that simultaneously recognizes hand gestures and head movements. Yang and Lee (2011) proposed a multimodal framework using Hierarchical Conditional Random Fields (H-CRF) and SVMs to identify hand gestures and facial expressions, respectively. In a subsequent work, they introduced a second multimodal framework utilizing three cameras to capture different orientations (Yang & Lee, 2013). Chen, Tian, Liu, and Metaxas (2013) presented a novel multimodal framework that combines face and body gestures, employing Histogram of Oriented Gradients (HOG) features and an SVM classifier to recognize ten distinct expressions. Recent research by Huang, Zhou, Zhang, Li, and Li (2018) introduced a multimodal framework for continuous sign language recognition



The Sign Language Recognition (SLR) system comprises several interconnected components designed to accurately interpret sign language gestures and facilitate communication between signers and non-signers. The following is a high-level overview of the typical architecture of an SLR system:

Data Acquisition: The process begins with the collection of sign language data, typically in the form of videos or sequences of images capturing sign gestures. This data may be obtained from existing datasets or through custom recording setups.

Preprocessing: Raw sign language data undergoes preprocessing to enhance its quality and suitability for subsequent analysis. This may involve tasks such as noise reduction, background subtraction, image enhancement, and normalization.

Feature Extraction: Extracting discriminative features from preprocessed data is crucial for effective gesture recognition. Various techniques, including handcrafted features like Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), or deep learning-based feature extraction using Convolutional Neural Networks (CNNs), are employed to capture relevant information from sign language images or video

Gesture Recognition: The extracted features are then inputted into a gesture recognition model to classify sign gestures into their corresponding meanings or words. This stage often involves machine learning or deep learning algorithms trained on labeled sign language datasets. Common models include Support Vector Machines (SVMs), Hidden Markov Models (HMMs), Recurrent Neural Networks (RNNs), or Transformer-based architectures.

Postprocessing: After classification, postprocessing techniques may be applied to refine the recognition results. This can include temporal smoothing to improve gesture coherence over time, error correction mechanisms, or context-aware processing to consider linguistic and grammatical rules.

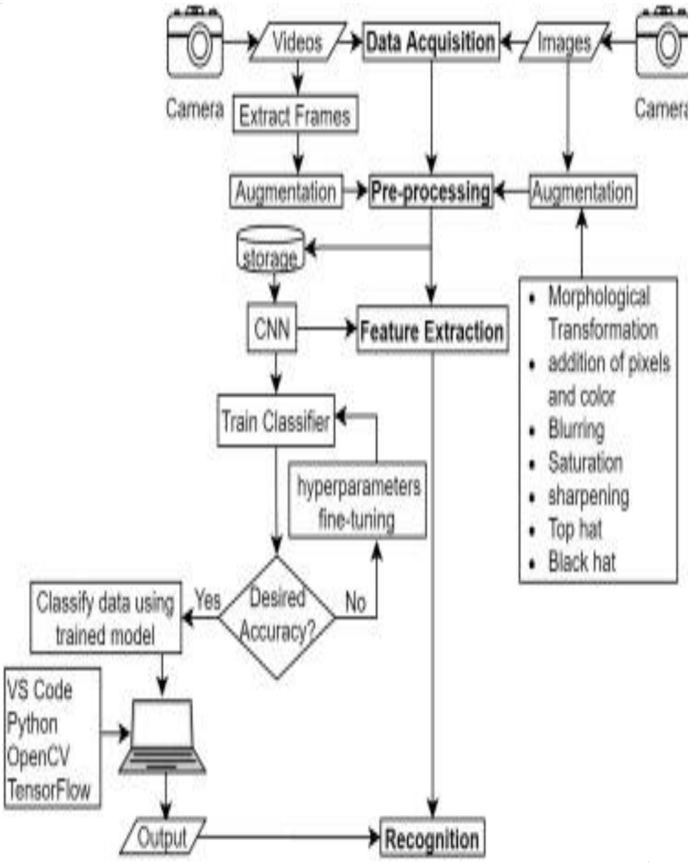
Integration and Output: The recognized sign language gestures are integrated into the desired output format, which may involve text-based transcription, speech synthesis, or visual rendering of sign language animations. This output is then presented to the intended recipient, such as a non-signer user or a device interface.

Feedback and Adaptation: SLR systems often incorporate feedback mechanisms to improve their performance over time. User feedback, corrections, and system logs may be used to update the recognition models through techniques like online learning or model retraining.

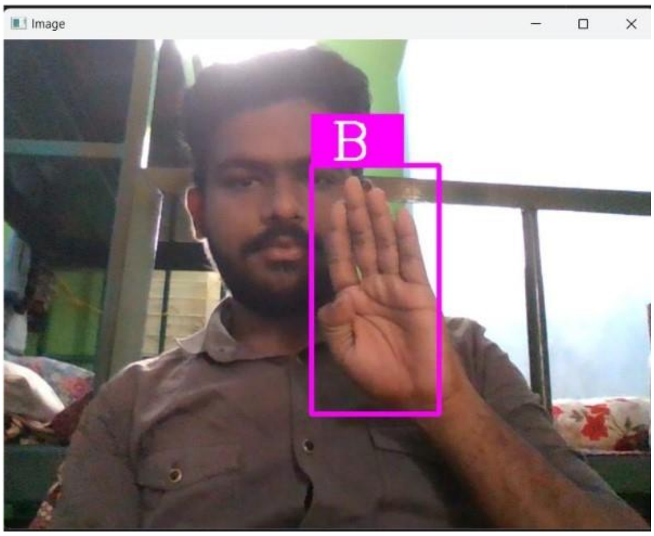
Deployment and Accessibility: Finally, the SLR system is deployed in real-world settings to facilitate communication accessibility for sign language users. This may involve integration into assistive technologies, educational platforms, or communication devices to empower individuals with hearing impairments in various contexts

4.GESTURE RECOGNITION

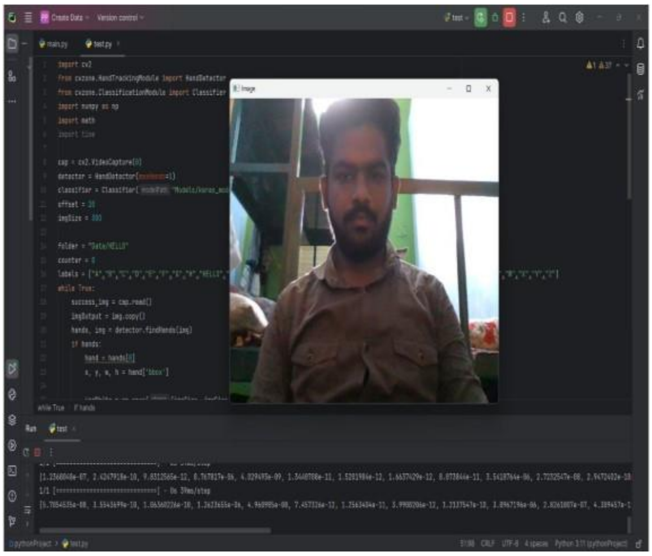
The system design for our American Sign Language (ASL) recognition project is a critical phase that lays the foundation for the entire system's architecture and functionality. This phase encompasses a comprehensive plan for the system's structure, components, and their interactions. The design process is driven by the project's objectives and requirements, focusing on creating a seamless interface between ASL users and the technology that interprets their signs. It encompasses the selection of appropriate technologies, algorithms, and hardware components, ensuring that the ASL recognition system performs efficiently and accurately. The system design will address various key aspects, including the architectural framework, data flow, user interface, and the underlying neural network models that power sign language recognition. Additionally, it will detail the system's ability to adapt to diverse sign language gestures, environmental conditions, and users' preferences. The user experience and ease of interaction are central to the design, aiming to make ASL communication more accessible and inclusive. In this section, we will delve into the detailed system design, including the architecture, data flow, user interface, and the core components that enable realtime ASL recognition. This design phase is pivotal in translating the project's objectives into a functional and user friendly ASL recognition system.



5.SYSTEM INTERFACE



The GUI for your ASL recognition prediction page should primarily focus on displaying the recognized ASL signs and words. Include a large, easily readable text area where the predicted signs or words are shown. Consider adding navigation buttons that allow users to switch between the prediction page and other sections of the application. Make sure the user interface remains visually appealing and intuitive to enhance the user experience while interacting with the ASL



The GUI for your ASL recognition system should be user-friendly and concise. It should include a clear title and logo at the top, along with a prominent video feed displaying the real-time camera input for sign language recognition. An about section can provide information about your project, and an exit button should allow users to close the application when finished. It's crucial to focus on accessibility, error handling, and responsive design to cater to a diverse user base effectively. Regular user testing and feedback are essential to refine the GUI for optimum usability which is easier for the user to go through and can able to access the image in a easier manner which is used in a effective way to know the impaired persons how to use it in an efficient way .

6.CONCLUSION

The system is novel approach to ease the difficulty in communicating with those having speech and vocal disabilities. Since it follows an image-based approach it can be launched as an application in any minimal system and hence has near zero-cost. There are further areas of improvement such as increasing the system performance under robust and unfavorable environment (lot of clutter, poor lighting). We also need to expand the current feature set to be able recognize more gesture (like those involving two hands or facial cues). We also need to deal with co-articulation.

REFERENCES

1. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998. doi: 10.1109/5.726791.
2. T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression Database," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 12, pp. 1615-1618, Dec. 2003. doi: 10.1109/TPAMI.2003.1251154.
3. C. Szegedy et al., "Going Deeper with Convolutions," arXiv:1409.4842, 2014.
4. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for ImageRecognition," arXiv:1512.03385, 2015.
5. A. M. Martinez, "The AR Face Database," CVC Technical Report #24, 1998.
6. OpenCV Library. (2021). OpenCV Library. [Online]. Available: https://opencv.org/.
7. TensorFlow. (2021). TensorFlow. [Online]. Available: https://www.tensorflow.org/.
8. Keras Documentation. (2021). Keras. [Online]. Available: https://keras.io/.
9. Dlib C++ Library. (2021). Dlib C++ Library. [Online]. Available: http://dlib.net/.
10. Python Software Foundation. (2021). Python. [Online]. Available: https://www.python.org/.
11. Pu, Junfu, Wengang Zhou, and Houqiang Li. (2019) “Iterative alignment network for continuous sign language recognition”, Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.
12. R.H. Abiyev, M. Arslan, J.B. Idoko, (2020) “Sign language translation using deep convolutional neural networks”, KSII Transactions on Internet and Information Systems, Vol. 14 (2), pp. 631-653. 48
13. P. Anguraj and T. Krishnan, “Design and implementation of modified BCD digit multiplier for digit-by-digit decimal multiplier,” Analog Integr. Circuits Signal Process., pp. 1–12, 2021.
14. T. Krishnan, S. Saravanan, A. S. Pillai, and P. Anguraj, “Design of highspeed RCA based 2-D bypassing multiplier for fir filter,” Mater. Today Proc., Jul. 2020, doi: 10.1016/j.matpr.2020.05.803.
15. T. Krishnan, S. Saravanan, P. Anguraj, and A. S. Pillai, “Design and implementation of area efficient EAIC modulo adder,” Mater. Today Proc., vol. 33, pp. 3751–3756, 2020.
16. S.C. Agrawal, A.S. Jalal, R.K. Tripathi, (2016) “A survey on manual and non-manual sign language recognition for isolated and continuous sign” International Journal of Applied Pattern Recognition, Vol. 3 (2), pp. 99.
17. Sincan, OzgeMercanoglu, et al. (2021) “Chalearn LAP large scale signer independent isolated sign language recognition challenge: Design, results and future research”, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition
18.Thad Eugene Starner. Visual recognition of american sign language using hidden markov models. Master's thesis, Massachusetts Institute of Technology, Cambridge MA, 1995.
19. P.K. Bora M.K. Bhuyan and D. Ghosh. Trajectory guided recognition of hand gestures having only global motions. International Science Index, 2008.
20. Emil M. Petriu Qing Chen, Nicolas D. Georganas. Feature extraction from 2d gesture trajectory in dynamic hand gesture recognition, 2006.
21. Manisha M. Ingle Neelam K. Gilorkar. A review on feature extraction for indian and american sign language, 2014.
22. Chih-Chung Chang and Chih-Jen Lin. A review on feature extraction for indian and american sign language, 2013.
23. ATHIRA P K ALEENA K RAJ and DEEPA I K. Sign language conversion software for people with hearing and vocal disabilities,2013.
24. Mokhtar M. Hasan & Pramod K. Misra, HSV Brightness Factor Matching for Gesture Recognition System. International Journal of Image Processing (IJIP), Volume (4): Issue (5)
25. Xingyan. Li. "Vision Based Gesture Recognition System with High Accuracy". Department ofComputer Science, The University of Tennessee, Knoxville, TN 37996-3450, 2005