International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 11 | Nov - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Reverse Sign Language Recognition System Using Machine Learning

Aaditya Mahindrakar [1], Anushka Misal [2], Vaibhavi Dyavarkonda [3] ,Aishawarya Kadam [4], Prof. Smita Shendge [5].

[1]*Department of Computer Engineering,Vidya Vikas Pratishthan Institute Of Engineering And Technology,Solapur.*
[2]*Department of Computer Engineering,Vidya Vikas Pratishthan Institute Of Engineering And Technology,Solapur.*
[3]*Department of Computer Engineering,Vidya Vikas Pratishthan Institute Of Engineering And Technology,Solapur.*
[4]*Department of Computer Engineering,Vidya Vikas Pratishthan Institute Of Engineering And Technology,Solapur.*
[5]*Department of Computer Engineering,Vidya Vikas Pratishthan Institute Of Engineering And Technology,Solapur.*

---***---

Abstract - Deaf and hard-of-hearing people communicate with each other and their society by using sign language. As this is a very natural method for pupils to stay in touch with computers, many academics are working on it to make it less complex and more convenient for use. So the main objective of gesture recognition research is to make systems which can interact and communicate while understanding human gestures and use them. In simple words to communicate information. Fast and extremely accurate hand detection and realtime hand gestures.
Identification should be possible with vision-based hand gesture circumstances and interfaces. Learning and knowing sign movements and gestures is the kick start in making words and sentences for computer assisted sign language interpretation. Both dynamic and static sign actions are open and available. Both ways of gesture recognition are crucial to human culture, even if static gesture recognition is much easier than dynamic gesture recognition. When a human enters the value of alphabets and numerical value as input, the system immediately displays or outputs the appropriate recognised character shows the gesture on the monitor screen. In the following, research projects that have led to a proper system that uses convolutional neural networks to identify handwriting on the basis of the depth pictures and Hand Languages (Brain lipi) the collects.

Keywords- Convolutional neural network,Text recognition ,Convert to image,StackGAN,simultaneous Tracking text and converting to image, gesture recognition,display output as sign gesture,training machine to(A-z to 1-0).

1. INTRODUCTION

Communication is fundamental to human interaction, facilitating the exchange of thoughts, ideas, and emotions essential for daily life. For deaf and mute individuals, sign language serves as the primary means of expressing themselves within their communities, using structured gestures and body language to convey meaning. As society becomes increasingly aware of the importance of inclusive communication, the significance of sign language as a bridge between people of different abilities is widely recognized. Sign language is a structured visual language, where each gesture corresponds to a specific meaning. This framework enables deaf and mute individuals to communicate effectively, fostering connection and understanding. Recent advancements in text and image recognition have opened up possibilities for bridging communication gaps by converting text or spoken language into sign language gestures—a process known as "reverse sign language recognition." The goal of reverse sign language recognition is to identify written or spoken words

and translate them into the corresponding signs, creating a more accessible communication tool for the deaf and mute community. With ongoing progress in this field, new developments continue to refine and enhance existing systems, supporting more accurate and efficient interpretations between sign language and other forms of communication. Each advancement not only enhances accessibility but also promotes inclusivity, empowering individuals to engage fully in diverse communicative interactions.

2. BODY OF PAPER

This research utilizes StackGAN for generating realistic hand gesture images based on text descriptions of signs, specifically for the A-Z alphabets and 0-9 numbers in sign language. The methodology encompasses data collection, data preprocessing, model training, testing, and evaluation. StackGAN's generative approach allows for detailed image synthesis based on textual input, enhancing sign language accessibility through accurate visual representations of gestures.

1. Data Collection:-

Dataset Acquisition: A comprehensive dataset is collected that includes sign language gestures for all A-Z alphabets and 0-9 numbers. The dataset should contain labelled images of each gesture from multiple angles and variations. Textual Descriptions: Each gesture is paired with detailed text descriptions to act as inputs for the StackGAN model. Descriptions include key details, such as finger positioning, hand orientation, and other distinguishing characteristics to enhance the realism of generated images. Data Augmentation: To increase dataset diversity, various data augmentation techniques are applied, such as rotation, scaling, flipping, and contrast adjustments. This helps StackGAN generate accurate images under diverse realworld conditions.

2. Data Preprocessing:-

Image Processing: Images are preprocessed by resizing to a consistent resolution (e.g., 64x64 or 128x128 pixels) for StackGAN input. Noise reduction and normalization techniques are applied to ensure high-quality data.

Text Vectorization: Text descriptions are preprocessed using tokenization and vectorization. Techniques like Word2Vec or GloVe embeddings are employed to convert text descriptions into vector formats compatible with StackGAN.

Feature Extraction: In cases where additional features may aid the model, advanced feature extraction methods, such as HOG or SIFT, may be applied to assist in accurate gesture differentiation during image generation.

3. StackGAN Model Architecture:-

StackGAN is selected for its two-stage architecture, allowing progressive generation of high-resolution images from textual descriptions.

The model architecture comprises two main stages: Stage I Generator: This stage generates a low-resolution (64x64) image based on the text description. This initial image captures basic shape and outline details, focusing on the overall structure of the hand gesture.

Stage II Generator: The low-resolution image is refined into a high-resolution (128x128 or higher) image, adding fine details like finger joints, contours, and shading. Stage II also incorporates text embedding to ensure high fidelity between the image and the input description. The dual-stage process of StackGAN enables it to generate images with both accurate structure and intricate details, crucial for distinguishing different sign language gestures.

4. Training and Optimization:-

Training Procedure: The model is trained on paired textimage datasets, where each image corresponds to a specific alphabet or number gesture. Both stages of the StackGAN are trained iteratively with the text descriptionsto learn an accurate mapping from text to image.

Loss Functions: A combination of adversarial loss and text-image matching loss is used to optimize the model. Adversarial loss encourages realistic image generation, while the matching loss ensures generated images are consistent with the textual description.

Optimization Techniques: Adam optimizer is employed to update model parameters, using a tuned learning rate to balance convergence speed and model accuracy.

Techniques like batch normalization and dropout are also applied to avoid overfitting and improve model generalization.

5. Testing and Evaluation:-

Validation: A portion of the dataset is set aside for validation to assess the model's performance on unseen data. K-fold cross-validation is applied to further evaluate model robustness.

Testing on New Descriptions: The model is tested with novel text descriptions for each gesture to evaluate its ability to generalize to new descriptions while maintaining accurate gesture generation.

Performance Metrics: Metrics such as Inception Score (IS) and Fréchet Inception Distance (FID) are used to evaluate image quality and similarity to real images. Additionally, human evaluation is conducted to assess the subjective quality of generated images, especially in terms of accurate gesture representation.

Accuracy in Recognition:Generated images are assessed for accuracy in depicting specific A-Z alphabet and 0-9 number gestures. Precision, recall, and F1-score are calculated to gauge the success rate in producing recognizable sign language gestures.

6. System Implementation and Real-World Testing:-

Real-Time Testing: The model is deployed for real-time testing, where users input text descriptions to generate corresponding sign language gesture images. Testing focuses on response time, accuracy, and user experience.

User Interface (UI) Design: A user-friendly interface is designed to allow users to input text for the desired alphabet or number and receive the generated image as output. The UI provides real-time feedback and supports smooth interaction.

Integration with Assistive Devices: The system can be integrated with assistive devices or mobile applications to facilitate access for the deaf and mute community, providing visual support for learning and communication.

A) Text Embedding

Taking input from user

B) Stage 1 – Low Resolution

C) The initial stage generates a basic, lowresolution image based on the text input, capturing primary shapes, colors, and layout. This rough sketch provides a foundation for adding finer details in the high-resolution stage that follows.

D) Stage 2 – High Resolution

approach ensures that each sign image is refined and representative of real-life gestures, making it ideal for use in educational tools, assistive technology, and real-time recognition systems.

E) Module 4 – Graphic User Interface (GUI).

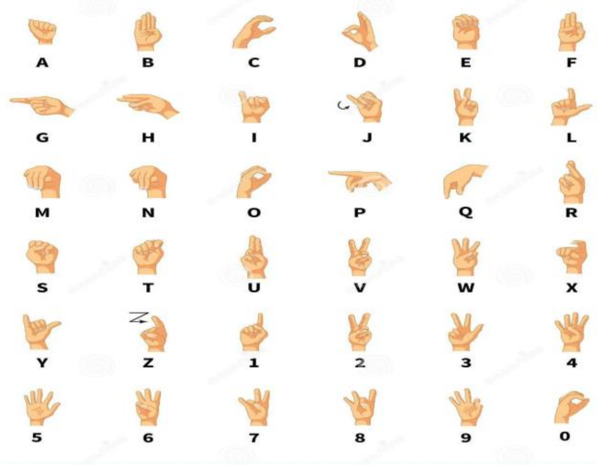

Fig1. sign languages used in the algorithm

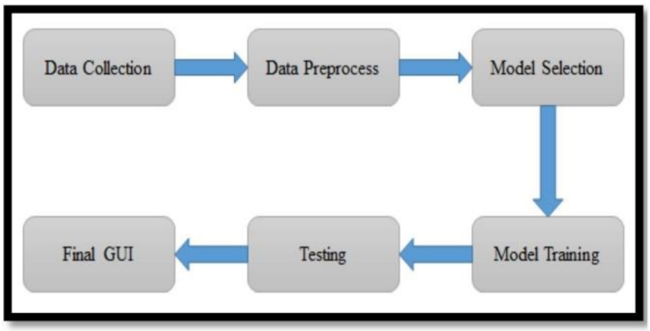

Fig2. Process Flow Diagram

3. CONCLUSIONS

The research presented underscores significant advancements and persistent challenges in the realm of sign language learning and recognition, highlighting the pivotal role of deep learning models, particularly Convolutional Neural Networks (CNNs). The studies encompass various languages, including Arabic, Pakistani, Chinese, Hindi, and Indian Sign Language, emphasizing the essential need for efficient communication methods for the deaf and mute community. CNNs have shown exceptional ability in accurately identifying hand gestures, surpassing traditional methods and delivering robust, efficient performance suited even for mobile platforms. Researchers have successfully addressed the complexities inherent in gesture recognition, including diverse hand shapes, orientations, and facial expressions, by meticulously constructing datasets and refining network layers throughout the deep learning process. These innovations represent a promising step forward in developing accessible, user-friendly solutions that enhance communication capabilities and inclusivity for the hearingimpaired community.

4. ACKNOWLEDGEMENT

I would like to express my deepest gratitude to all those who have supported and contributed to the completion of this project on Sign Language Recognition using Machine Learning and Deep Learning. This project would not have been possible without the guidance, resources, and encouragement I received.
First and foremost, I would like to thank my project supervisor [Supervisor's Name] for their invaluable guidance, insightful suggestions, and continuous support throughout the project. Their expertise in machine learning and dedication to accessibility served as a vital source of motivation.
I would also like to extend my appreciation to [Institution/Organization Name] for providing the necessary facilities, resources, and data, which were essential in building and training the model. Additionally, I am grateful to the members of [Lab or Research Group Name] for their support and feedback.
A special thank you goes to my family and friends, whose patience and encouragement have been a source of strength during this challenging yet rewarding journey.
Finally, I am incredibly grateful to the Deaf and Hard of Hearing communities for inspiring this project. Their resilience and need for inclusivity have been a driving force behind this work, and I hope this project can contribute, even in a small way, to breaking down communication barriers.

Thank you all for your support and encouragement.

5. REFERENCES

[1] Mourad Bahani, Aziza El Ouaazizi, Khalil Maalmi "Ara-BERT and DF-GAN fusion for Arabic text-to-image generation", 2022.
[2] M. Alsulaiman, M. Faisal, M. Mekhtiche, M. Bencherif, T. Alrayes, G. Muhammad, H. Mathkour, W. Abdul, Y. Alohali, M. Alqahtani, H. Al-Habib, H. Alhalafi, M. Algabri, M. Al-hammadi, H. Altaheri, and T. Alfakih, "Facilitating the communication with deaf people: Building a largest Saudi sign language dataset," 2023.
[3] K. Kudrinko, E. Flavin, X. Zhu, and Q. Li, "Wearable Sensor-Based Sign Language Recognition: A Comprehensive Review," 2020.
[4] F. Shah, M. S. Shah, W. Akram, A. Manzoor, R. O. Mahmoud, and D. S. Abdelminaam, "Sign Language Recognition Using Multiple Kernel Learning: A Case Study Of Pakistan Sign Language," 2021.
[5] S. M. Kamal, Y. Chen, S. Li, X. Shi, and J. Zheng, "Technical Approaches To Chinese Sign Language Processing," 2020.
[6] H. D. Yang, S. Sclaroff, and S. W. Lee, "Sign Language Spotting With A Threshold Model Based On Conditional," 2009.
[7] C. M. Sharma, K. Tomar, R. K. Mishra, and V. M. Chariar, "Indian Sign Language Recognition Using Fine-Tuned Deep Transfer Learning Mode," 2021.
[8] S. C. J and L. A, "Signet: A Deep Learning Based Indian Sign Language Recognition System," 2019.
[9] B. M. V and S. T. D, "Indian Sign Language Numeral Recognition Using Region Of Interest Convolutional Neural Network," 2018.
[10] G. Rao, K. Syamala, P. V. V. Kishore, and A. S. C. S. Sastry, "Deep Convolutional Neural Networks For Sign Language Recognition," 2020.
[11] N.Pugeault & R.Bowden "Spelling It Out: Real–Time ASL Fingerspelling Recognition" 2019
[12] Xu T, Zhang P, Huang Q, Zhang H, Gan Z, Huang X, et al. "Attngan: Fine- grained text to image generation with attentional generative adversarial networks",2018
[13] Yong Xuan Tan , Chin Poo Lee , Mai Neo , Kian Ming Lim , Jit Yan Lim,"Text-to-image synthesis with

self-supervised bi-stage generative adversarial network" 2023
[14] B. Natarajan , E. Rajalakshmi , R. Elakkiya , Ketan Kotecha , Ajith Abraham , (Senior Member, IEEE), Lubna Abdelkareim Gabralla and V. Subramaniyaswamy, "Development of an End-to- End Deep Learning Framework for Sign Language Recognition, Translation, and Video Generation" 2022
[15] Hyunhee Lee , Gyeongmin Kim , Yuna Hur , And Heuiseok Lim, "Visual Thinking of Neural Networks: Interactive Text to Image Synthesis" 2021.
[16] Uche Osahor , (Student Member, Ieee), And Nasser M. Nasrabadi, "Text-Guided Sketch-to- Photo Image Synthesis" 2023.
[17] Nikhil Kasukurthi ,Brij Rokad, Shiv Bidani, Aju Dennisan, "American Sign Language Alphabet Recognition using Deep Learning" 2019.
[18] Shagun Katoch, Varsha Singh, Uma Shankar Tiwary," Indian Sign Language recognition system using SURF with SVM and CNN" 2022.
[19] Satwik Ram Kodandaram, N. Pavan Kumar, Sunil G ," Sign Language Recognition using Deep learning." ,2021
[20] Yuxuan Liu, Xijun Jiang, Xingge Yu, Huaidong Ye, Chao Ma, Wanyi Wang, Youfan Hu, "A wearable system for sign language recognition enabled by a convolutional neural network", 2023