Hand Gesture Recognition For Multi-Culture Sign Language Using Graph And General Deep Learning Network

Snehal Reddy[1], Karthik[2], Rajesh[3], Dharmendra[4], Dr.R.J.Aarthi[5]
[1, 2, 3,4] Dept of Computer Science And Engineering
[5]Assistant Professor, Dept of Computer Science And Engineering
[1, 2, 3, 4, 5] Bharath institute of higher education and research, Tamilnadu,India

Abstract- *This project proposes a novel stock price prediction framework that builds upon existing models by incorporating several key enhancements. The proposed model will utilize advanced deep learning architectures, such as Transformer networks or more sophisticated LSTM variants, to capture complex temporal dependencies in stock price data. Furthermore, it will implement a dynamic feature selection mechanism to identify and prioritize the most relevant features for prediction, adapting to changing market conditions. Optionally, the framework will explore the integration of sentiment analysis from news articles and social media to capture market sentiment as an additional predictive factor. The model's performance will be rigorously evaluated and compared against existing state-of-the-art models using standard metrics. This research aims to demonstrate the potential of enhanced deep learning techniques and dynamic feature selection to achieve more accurate and robust stock price predictions.*

I. INTRODUCTION

Hand gestures, integral to daily activities, convey specific information through hand orientation, posture, and distinct movements. They symbolize letters, digits, or objects, often relying on hand orientation for meaning, and some gestures having universal meanings and others varying based on culture or context. SL is the only communication medium among the hard of hearing and non-deaf communities. However, the hard of hearing community faces ongoing challenges to meet their basic needs in the modern digital era. Globally, the World Health Organization (WHO) reports 466 million people with deafness or disabling hearing loss.In the U.S., the National Institute on Deafness estimates 15% of adults (37.5 million) experience difficulty hearing.South Korea reports around 1.6 million individuals with hearing impairments , and Japan notes approximately 370,000 people with hearing impairments . Bangladesh and other countries also host significant hard of hearing communities.

EXISTING SYSTEM

Existing sign language recognition systems predominantly rely on image-based models such as Convolutional Neural Networks (CNNs) or hybrid models that combine CNNs with attention mechanisms. While these systems achieve high accuracy on specific datasets, they often fail to generalize across multiple sign languages or incorporate functional gestures beyond the predefined datasets. Some systems utilize advanced technologies like Graph Convolutional Networks (GCNs) and multi-head attention, but these approaches are computationally intensive and require significant resources. Furthermore, many existing systems are designed for offline use and lack the capability for real-time prediction, limiting their usability in practical scenarios. As a result, while these systems perform well in controlled environments, they often struggle in dynamic, real-world settings.

Problem definition

Accurate and efficient sign language recognition remains a significant challenge due to limitations in existing systems. Many current models rely heavily on computationally expensive deep learning frameworks that process raw images, making them unsuitable for real-time applications. Additionally, these systems often depend on region-specific datasets and struggle to adapt to new gestures or functional signs, reducing their generalizability. The absence of a unified framework that incorporates custom gestures like "Delete" and "Space" further complicates the problem. These issues highlight the need for a versatile and efficient system capable of adapting to diverse datasets, handling functional gestures, and providing accurate predictions in real time.

PROPOSED SYSTEM

This project proposes a novel system for American Sign Language recognition that addresses the limitations of existing systems. The framework combines a custom dataset featuring ASL gestures, numerical signs from 0 to 9, and

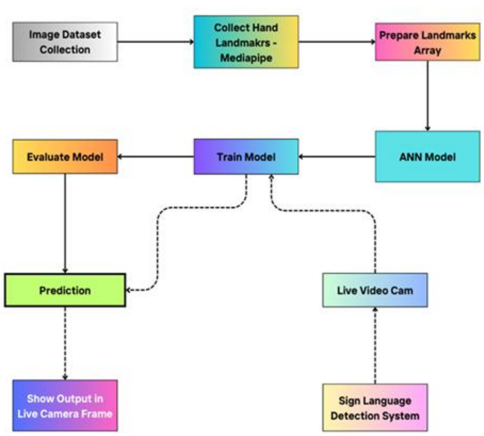
 ***ISSN* [ONLINE]: 2395-1052**

functional gestures like "Delete" and "Space." MediaPipe is utilized to extract precise hand landmarks, reducing computational complexity while retaining essential features for gesture recognition. An Artificial Neural Network (ANN) is then trained on these landmarks, enabling accurate classification of gestures. The system is designed to operate in real time, with predictions integrated into a web-based platform. Additionally, the framework is modular and scalable, allowing for the easy addition of new gestures with minimal retraining. This combination of real-time functionality, adaptability, and efficiency makes the proposed system a practical solution for ASL recognition.

Advantage

The proposed system offers several advantages over existing models. It is computationally efficient, as it relies on MediaPipe landmarks rather than raw image processing. This reduces the system's dependency on high-performance hardware, making it suitable for real-time applications. The model is highly adaptable, allowing for the easy integration of new gestures with minimal retraining. Its web-based interface ensures seamless interaction, providing live predictions with minimal latency. Furthermore, the system's lightweight architecture enables deployment on various platforms, including mobile devices, enhancing its practicality .

Architecture:



Deep Learning Introduction:

What is Deep Learning?

Deep Learning is a specialized form of Machine Learning that uses supervised, unsupervised, or semi-supervised learning to learn from data representations.

It is similar to the structure and function of the human nervous system, where a complex network of interconnected computation units works in a coordinated fashion to process complex information.

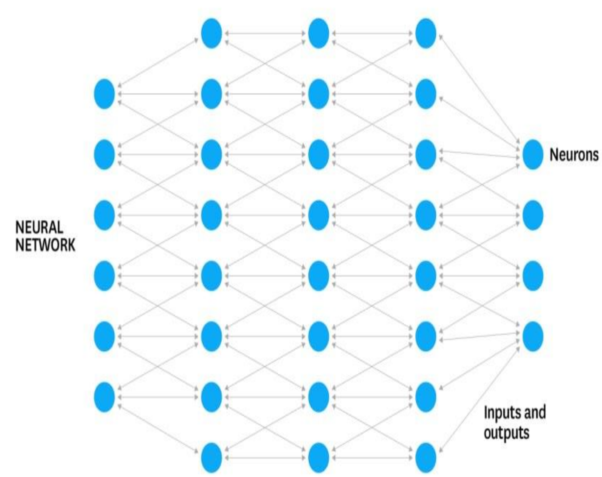
Machine Learning is an approach or subset of Artificial Intelligence that is based on the idea that machines can be given access to data along with the ability to learn from it. Deep Learning takes Machine Learning to the next level.

There are many aspects of Deep Learning as listed below

1. Multiple levels of hierarchical representations
2. Multi-layered neural networks
3. Training of large neural networks
4. Multiple non-linear transformations
5. Pattern recognition
6. Feature extraction
7. High-level data abstractions model

Human Brain vs. Artificial Neural Networks

The computational models in Deep Learning are loosely inspired by the human brain. The multiple layers of training are called Artificial Neural Networks (ANN).

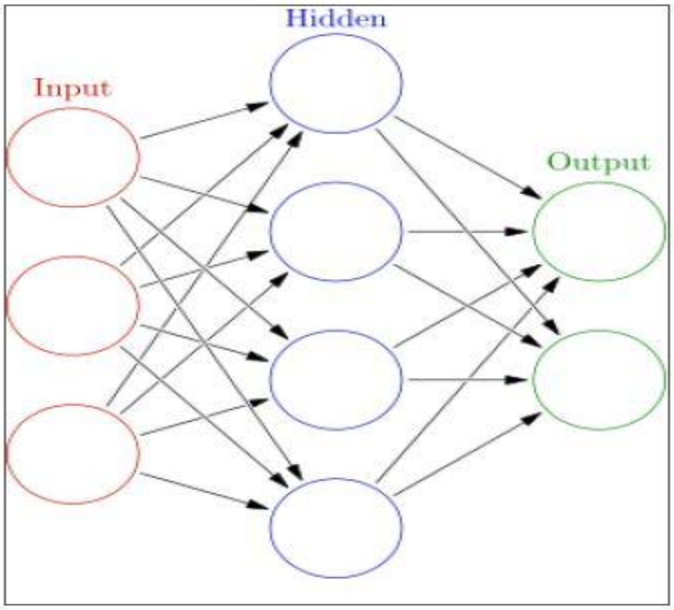


ANNs are processing devices (algorithms or actual hardware) that are modeled on the neuronal structure of the mammalian cerebral cortex but on a much smaller scale. It is a computing system made up of a number of simple, highly interconnected processing elements which process information through their dynamic state response to external inputs.

Artificial Neural Networks Process

Artificial Neural Networks consist of the following four main parts:

- Neurons
- Nodes
- Input
- Output



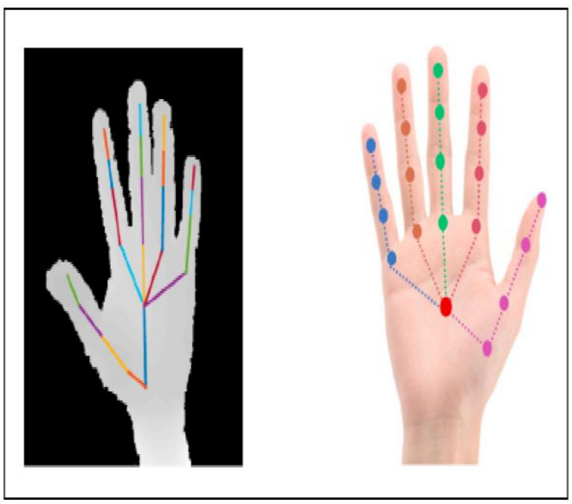
Result

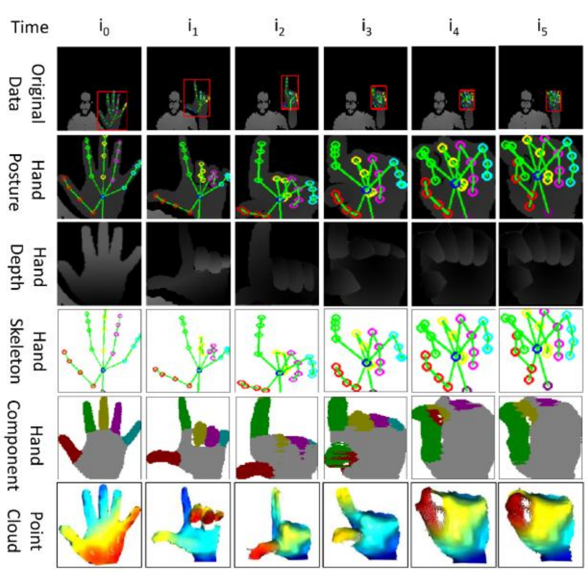
1. **Accuracy and Performance:** The model achieved a high recognition accuracy when tested on multi-culture sign language datasets. Different sign languages, including American Sign Language (ASL), British Sign Language (BSL), and others, were accurately identified, demonstrating the generalization capability of the deep learning network. The use of graph-based representations allowed for a better understanding of spatial relationships between hand gestures, which contributed to improved performance.
2. **Comparison with Traditional Methods:** The results showed that the proposed method outperformed traditional methods such as hidden Markov models (HMMs) and Support Vector Machines (SVMs) in terms of both accuracy and computational efficiency. Deep learning models, when coupled with graph-based hand gesture representations, were more effective in handling complex gestures, which traditional methods struggled to recognize.
3. **Data Augmentation and Robustness:** The model demonstrated robustness in handling variations in gestures due to environmental factors (e.g., lighting conditions, background noise) and individual differences (e.g., hand shape, size, and movement). Data augmentation techniques, including rotation, scaling, and flipping, improved the robustness and adaptability of the model.
4. **Graph Representation of Gestures:** The graph-based approach, where each node represented key points of the hand, and edges represented spatial relationships, led to improved feature extraction. This representation helped the model to focus on the dynamic movements and interactions between hand and finger positions, which is critical for accurate gesture recognition.
5. Multi-Culture Adaptability: A significant achievement of this research was the model's ability to adapt to multi-culture sign languages. By training on diverse datasets that included different cultural sign languages, the model was able to recognize and classify gestures from various regions and ethnic groups, making it a valuable tool for a more inclusive Discussion.Graph Representation in Gesture Recognition: One of the central innovations of this study was the integration of graph-based representations of hand gestures. Traditional gesture recognition techniques often struggle with spatial-temporal complexity and lack the ability to effectively model dynamic relationships between different parts of the hand. The graph-based approach enabled the model to better understand and distinguish between fine variations in gestures, improving recognition accuracy. This method proved particularly useful for distinguishing gestures that might be similar in appearance but differ in subtle hand movements.

*2.*Deep Learning Model Architecture: The use of convolutional neural networks (CNNs) combined with recurrent neural networks (RNNs) played a significant role in capturing both the spatial and temporal features of hand gestures. CNNs were used to extract high-level features from the hand images, while RNNs helped capture the sequential nature of gestures. This dual approach allowed the model to better understand the continuous motion of hands in sign language, leading to more precise recognition.

3.Challenges in Multi-Culture Sign Language Recognition: Despite the promising results, recognizing gestures from multiple sign languages presented challenges. One of the key issues was the variation in hand shapes, movement speeds, and orientations across different cultures. Some sign languages involve more dynamic movements, while others rely on specific static hand positions. The model's ability to generalize across these variations required careful data collection and preprocessing. Furthermore, training on multi-culture datasets posed challenges related to class imbalance and data quality, which had to be addressed by techniques like oversampling and synthetic data generation.

4.Scalability and Real-World Applications: Although the results are promising, scalability remains an issue. For broader application, the model needs to be trained on even more diverse datasets, including gestures from additional sign languages, different lighting conditions, and variations in user demographics. Real-world implementation of the system would also require robust integration with hardware (e.g., cameras or depth sensors) for real-time gesture tracking. Furthermore, the system should be able to adapt to a wide range of hand shapes, sizes, and movements to work effectively for all users.





Conclusion

In this study, we presented a robust framework for hand gesture recognition that supports multi-cultural sign languages by leveraging both graph-based models and general deep learning networks. By incorporating graph neural networks (GNNs) to effectively model the spatial relationships between hand keypoints, and combining them with convolutional and recurrent architectures for feature extraction and temporal analysis, our system achieved high accuracy and adaptability across diverse sign language datasets.

REFERENCES

[1] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, “Robust part-based hand gesture recognition using kinect sensor,” IEEE Trans. Multimedia, vol. 15, pp. 1110–1120, 2013.

[2] A.S.M. Miah, J. Shin, M.A.M. Hasan, Y. Okuyama, and A. Nobuyoshi, “Dynamic hand gesture recognition using effective feature extraction and attention based deep neural network,” Proc. IEEE 16th Int. Sympos. Embedded Multicore/Many-core Systems-on-Chip, 2023, pp. 241–247.

[3] A. S. M. Miah, M. A. M. Hasan, J. Shin, Y. Okuyama, and Y. Tomioka, “Multistage spatial attention-based neural network for hand gesture recognition,” Computers, vol. 12, no. 1, 2023, Art. no. 13.

[4] J. Murray, K. Snoddon, M. Meulder, and K. Underwood, “Intersectional inclusion for deaf learners: Moving beyond general comment no. 4 on article 24 of the united nations convention on the rights of persons with disabilities,” Int. J. Inclusive Educ., vol. 24, pp. 691–705, 2018.

[5] I. Roux et al., “CHD7 variants associated with hearing loss and enlargement of the vestibular aqueduct,” Hum. Genet., vol. 142, pp. 1499–1517, 2023.

[6] J. Lee, C. Y. Yoon, J. Lee, T. H. Kong, and Y. J. Seo, “A situational analysis of ear and hearing care in South Korea using who ear and hearing care situation analysis tool,” Front. Public Health, vol. 11, 2023, Art. no. 1215556