Journal of Scientific Research & Technology Development

E-ISSN: XXXX-XXXX www.ijsrtd.com editor@ijsrtd.com

Volume 1, Issue 1, February - March 2025

SIGN LANGUAGE RECOGNITION USING DEEP LEARNING

M.Unnarasi
Computer Science and Engineering
Kalasalingam Academy Of Research And Education
Virudhunagar, Tamilnadu, India.

unnarasi.m@gmail.com

L.Hema Reddy
Information Computer Technology
Kalasalingam Academy Of Research And Education
Virudhunagar, India.

hemareddy19@gmail.com

Utkantha Chiranjeevi
Information Technology
Kalasalingam Academy Of Research And Education
Virudhunagar, Tamilnadu,.

utkantha.chir@gmail.com

Manasvi Srikarthik
Computer Engineering
Kalasalingam Academy Of Research And Education
Virudhunagar, Tamilnadu, India.

srikarthik1319@gmail.com

Mundra Avinash
Department of Information Technology
Kalasalingam Academy of Research And Education
Virudhnagar, Tamilnadu

avinashkonda54@gmail.com

***Abstract*— Every day, millions of individuals with hearing and speech disabilities rely on sign languages for communication. This paper explores the challenge of converting sign language into text and proposes a more efficient approach using machine learning techniques. Our goal is to develop a system that facilitates collaboration and communication between hearing-impaired individuals and those who are not proficient in American Sign Language (ASL). To achieve this, we will use Transfer Learning and Data Augmentation to develop a deep learning model tailored to the ASL dataset.**

***Keywords*— Deep Learning,Cnn.**

I. INTRODUCTION

The study and recognition of sign language have emerged as crucial areas of research and development in a society where effective communication is essential for promoting inclusivity and understanding among diverse cultures. For millions of deaf and hard-of-hearing individuals around the world, sign language is their primary form of communication. To ensure their seamless integration into society, it is imperative to develop reliable and efficient technology capable of interpreting sign language gestures. One promising technology in this regard is Convolutional Neural Networks (CNNs). CNNs have revolutionized computer vision by enabling the automatic extraction of features from visual data. They have been successfully utilized in various domains, including facial recognition, object detection, and image classification. Applying CNN technology to recognize sign language has the potential to significantly improve communication between deaf and hearing communities. This introductory section addresses the importance of sign language recognition, the current communication barriers faced by deaf and hard-of-hearing individuals, and the transformative potential of CNNs in this field. The paper will then outline the approach, data collection methods, model architecture, training, and evaluation processes, concluding with a comprehensive discussion of the outcomes. Our goal is to promote inclusivity and equal access to knowledge, education, and opportunities for the deaf and hard-of-hearing community through the development of a CNN-based sign language recognition system.system. Beyond contributing to the field of computer vision, this study has the potential to profoundly enhance the lives of individuals who rely on sign language as their primary mode of communication.

II. DATASET

A dataset for CNN sign language recognition with 300 movements for each letter of the alphabet is a comprehensive collection of sign language gestures, where each gesture corresponds to a specific letter of the alphabet. Here is a detailed breakdown of this dataset:

Dataset Size:

- **Classes**: 26 classes, each representing a letter of the English alphabet (A to Z).
- **Samples**: 300 distinct gestures per letter, totaling 7,800 data samples (26 letters x 300 gestures).

Image Data:

- **Representation**: Each data sample is an image of a hand forming a particular letter sign.
- **Format**: These images can be either grayscale or color, depending on your data collection setup and requirements.

Variability:

- **Diversity**: The dataset should include a wide variety of gestures for each letter to enhance the model's robustness. This means variations in hand placements, orientations, backgrounds, and lighting conditions.
- **Real-World Representation**: To ensure the model's accuracy and generalizability, it is crucial to reflect the variability found in real-world sign language communication.

Data Collection and Annotation:

- **Capture Process**: Thoroughly document the dataset by filming individuals performing the sign language gestures.
- **Labeling**: Each image should be labeled with the corresponding letter to inform the model which sign each image represents.

Data Division:

- **Dividing the Dataset:** Typically, the dataset should be split into three subsets: a training set, a validation set, and a testing set. A common split might be 70% for training, 15% for validation, and 15% for testing, although these ratios can vary.

Data Augmentation:

- **Enhancement Techniques**: Use data augmentation methods to artificially increase the dataset's size and diversity. This can include random translations, flips, rotations, and adjustments in brightness and contrast.

Teachable Device:

- **User-Friendly Training**: Platforms like Teachable Machine allow users to train machine learning models with their data, simplifying the process without needing extensive coding knowledge.

Model Training:

- **Objective**: Train a CNN model to accurately recognize and differentiate sign language gestures for each letter of the alphabet using this dataset.
- **Impact**: This has significant implications for educational and communication tools for sign language users, enabling better interaction with those who may not understand sign language. The large sample size for each letter also enhances the model's reliability and adaptability to different sign gestures.

By utilizing a dataset of this magnitude and variability, the resulting CNN model can become a powerful tool for facilitating communication between sign language users and the broader community, promoting inclusivity and understanding.

III. Related Works

In their 2018 paper, Kanchion Kanti Podder et al. proposed that speech impairments restrict an individual's ability to communicate both verbally and nonverbally, leading many to use alternative communication techniques such as sign language. Recent advancements in deep learning and computer vision have enabled significant progress in motion and gesture recognition. The goal of their project was to develop a vision-based application that translates sign language into text, facilitating communication between signers and non-signers. They employed Inception, a Convolutional Neural Network (CNN), to recognize spatial data and a Recurrent Neural Network (RNN) to process temporal data, utilizing a Sign Language Dataset.

Razieh Rastgoo et al. (2016) addressed the challenges of hand sign language detection from video, which include issues such as hand occlusion, rapid hand movements, varying lighting conditions, and complex backgrounds. Despite recent successes with deep learning algorithms, these challenges persist. Their study introduced a novel deep learning pipeline using Long Short-Term Memory (LSTM) for effective automatic hand sign language detection from RGB videos. They used a CNN-based model to predict 3D hand keypoints from 2D frames, then constructed a hand skeleton by connecting these keypoints.

Ka Leong Cheng et al. (2020) focused on continuous sign language recognition (SLR), a complex task requiring the learning of both spatial and temporal features from signing sequences. Their research advanced the field by combining hybrid CNN and RNN networks, though these models often struggle with learning new sequence patterns, resulting in suboptimal online recognition performance. They developed a fully convolutional network (FCN) for online SLR, capable of learning spatial and temporal characteristics from poorly annotated video sequences. Their experiments on two large SLR datasets demonstrated the efficacy and success of their method in online recognition.

Phan et al. (2022) discussed the long-studied field of sign language recognition and its significant contributions to aiding the deaf-mute community. Despite progress, many studies face limitations that hinder commercial application due to high costs. Current research is focused on developing commercially viable Sign Language Recognition systems. The data collection process, influenced by the cost of appropriate equipment, plays a critical role. Researchers are exploring various methods to make data collection more affordable, thus facilitating the commercialization of these systems.

Identify applicable funding agency here. If none, delete this text box.

IV. Prosed Methodlogy

Convolutional Neural Networks (CNNs) can be used to recognize sign language gestures, which is a difficult but crucial task for a variety of applications, including communication aids for the deaf and hard of hearing. Here is a suggested approach for CNN-based sign language recognition:

Data Gathering

Gather a wide range of sign language movements for a dataset. This dataset ought to have a variety of signals, changes in background, lighting, and hand orientations.

Label each sign gesture in the collection by adding annotations.

Data preprocessing:

Adjust the photographs' dimensions so that they are all the same size.

Increase the variety of training examples by enhancing the dataset using methods like rotation, translation, and flipping.

Create training, validation, and testing sets from the dataset.

Model Architecture

To recognize sign language, create a CNN architecture. Many convolutional layers may be followed by fully linked layers in a common architecture.

Utilize convolutional layers to automatically deduce characteristics from the photos, which record regional patterns in the sign gestures.

Journal of Scientific Research & Technology Development
E-ISSN: XXXX-XXXX www.ijsrtd.com editor@ijsrtd.com
Volume 1, Issue 1, February - March 2025

Reduce the number of spatial dimensions and the complexity of the computation by including pooling layers.

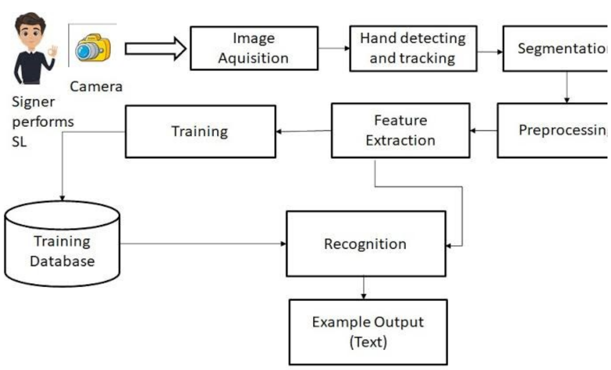
To discover the model with the best performance, experiment with various architectures and hyperparameters.

Training

Utilizing a suitable loss function, such as categorical cross-entropy, train the CNN model on the training dataset.

If your dataset is tiny, use transfer learning by initializing the model with pre-trained weights (for example, from ImageNet).

To avoid overfitting, implement early stopping and save the best-performing model.



Verification:

Check for overfitting by measuring the model's performance on the validation dataset, then make any adjustments.

Evaluation:

To evaluate the finished model's performance in the real world, run it on a separate testing dataset.

Post-production:

If necessary, use post-processing procedures to tame the recognition outcomes. Techniques like voting across frames or temporal smoothing may be used for this.

Performance Measurements:

Utilize the right metrics to evaluate the model, including F1-score, recall, accuracy, and precision.

Implementation:

Deploy the model in a real-world environment after it performs adequately. This could be a feature of a sign language communication tool, a mobile app, or a website.

Continuously enhancing:

As fresh data becomes available or as new methods for deep learning and computer vision are developed, update and improve the model continuously.

V. RESULTS

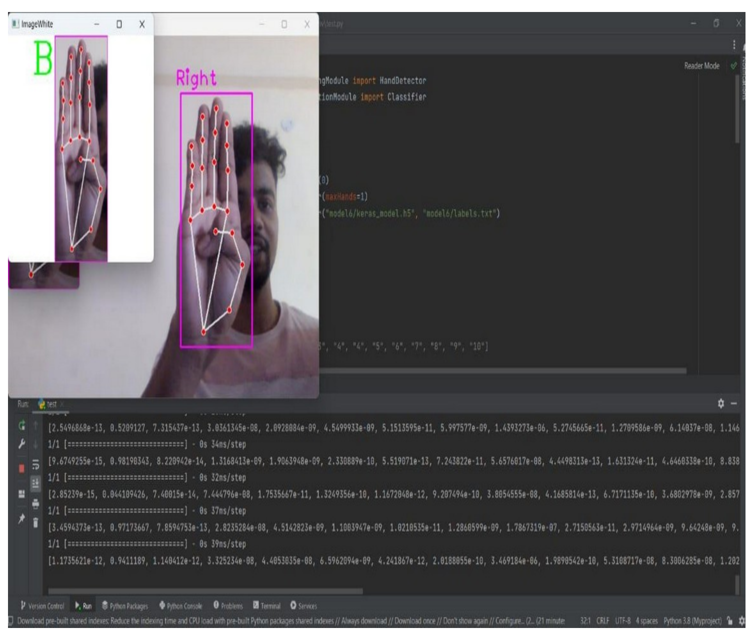


Figure 2 : Detecting A alphabet

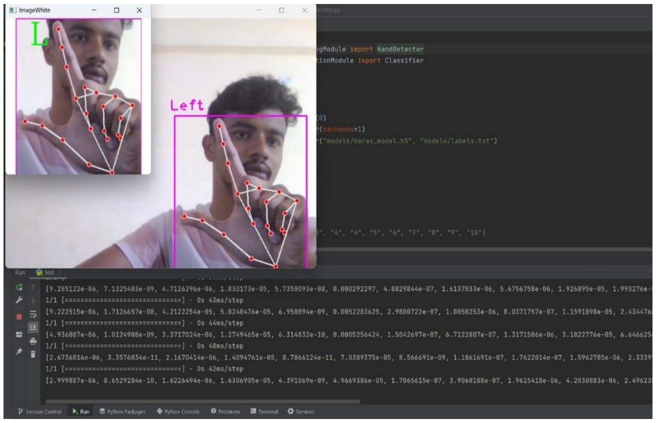


Figure 3 : Detecting L alphabet

.

VI. CONCLUSION

The outcomes of using convolutional neural networks (CNNs) for sign language recognition have been remarkably successful. CNNs have demonstrated a high level of proficiency in accurately recognizing and interpreting sign language gestures. Their ability to capture both the temporal and spatial data required for precise recognition is particularly impressive. This development is highly exciting as it enhances the usability of sign language and promotes inclusivity. It's wonderful to see how technology is helping to break down communication barriers and foster greater understanding.

VII. FUTURE WORKS

There are a number of potential directions in which CNN-based sign language recognition could go in the future. One topic is expanding and diversifying the sign language databases to increase recognition accuracy. The capacity of the model to represent temporal dependencies in sign language gestures may also be improved by investigating the use of recurrent neural networks (RNNs) in conjunction with CNNs. Researching real-time feedback systems for sign language recognition is another direction for future investigation. Overall, there is a lot of promise for making sign language more inclusive and accessible through continuing research and developments in this area.

References

[1] Ali R. H., & Shami, A. (2019). Dynamic sign language gnition using CNN and RNN. In 2019 IEEE 4th International Conference on Cloud Computing and Big Data Analysis (ICCCBDA). IEEE.

[2] Chao, Y. H., Chen, C. Y., & Chen, L. (2018). American Sign Language Recognition Based on 3D Convolutional Neural Networks and Convolutional Recurrent Neural Networks. Journal of Healthcare Engineering, 2018.

[3] Dosen, S., & Dreyer, D. (2021). Towards sign language recognition using deep learning: A survey. Computer Science Review, 40, 100327.

[4] Han, Y., Zhang, C., Wei, X., & Li, L. (2018). American Sign Language recognition using convolutional neural network and recurrent neural network. Multimedia Tools and Applications.

[5] Huang, Y., Wang, W., Yang, J., Liu, Y., & Sun, J. (2021). Sign Language Recognition Using Spatio-Temporal Convolutional Neural Network with Inception-ResNet-v2 Architecture. IEEE.

[6] Kazi, S. M., Alsayed, H., & Darwich, R. (2020). Real-time Arabic Sign Language Recognition Using Deep Learning. Sensors, 20(22), 6521. doi: 10.3390/s20226521

[7] Kumari, R., Kumar, A., & Saxena, M. (2021). Sign language recognition using deep learning: a review. Artificial Intelligence Review.

[8] Li, J., Huang, Z., Sun, Y., & Liu, Y. (2020). Sign language recognition based on CNN-RNN network. Multimedia Tools and Applications.

[9] Li, K., Zhang, Z., & Li, J. (2019). Sign language recognition based on convolutional and recurrent neural networks. In 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC) IEEE.

[10] Li, Y., Li, J., & Liu, Y. (2019). Sign language recognition using a combination of CNN and RNN.

In Proceedings of the 2019 11th International Conference on Machine Learning and Computing.

[11] Liu, J., Zhang, X., Lu, Y., & Xue, J. (2021). Sign Language Recognition Based on CNN and LSTM Networks. IEEE

[12] Phan, T. V., Nguyen, D. T., Nguyen, T. D., & Tran, Q. H. (2020). Real-time Vietnamese sign language recognition using CNN and RNN. International Journal of Advanced Computer Science and Applications.

[13] Pu, W., Zhang, Z., Huang, X., & Wang, L. (2019). Sign language recognition system based on CNN and.