PREMIER UNIVERSITY, CHATTOGRAM

Department of Computer Science & Engineering

Project Proposal Report

On

"Sign Language Recognition Using CNN"

SUBMITTED BY

Name: Sohela Showrin

ID: 2104010202199

Name: Rahin Toshmi Ohee

ID: 2104010202204

Name: Jahirul Islam Rakib

ID: 2104010202209

Name: Durjoy Nath

ID: 2104010202222

SUBMITTED TO

Noortaz Rezoana

Lecturer

Department of Computer Science & Engineering

Premier University, Chattogram

10 February, 2025

Table of Contents

Title Page i

Table of Contents ii

List of Figures iii

1. Introduction 1

    1.1 Purpose 1

    1.2 Motivation 2

2. Background and Present State of the System 3

3. Objectives 4

4. Literature Review 5

5. Outline Methodology 7

    5.1 Architecture 7

    5.2 Required Library 8

6. Conclusion 9

List of Figures

5.1 System Architecture Using CNN 7

1 Introduction

Communication is an important part of human life. People use language to share thoughts, feelings, and information. However, individuals with hearing or speech disabilities face difficulties in communicating with others. Sign language is a way for them to express themselves, but not everyone understands it. This creates a communication gap between people who use sign language and those who do not. Technology can help solve this problem. With the use of deep learning, computers can recognize hand gestures and translate them into letter. This project focuses on using Convolutional Neural Networks (CNN) to recognize sign language. The goal of this project is to make communication easier for people with hearing and speech disabilities.

1.1 Purpose

In the modern world, effective communication is essential for daily interactions, education, and professional environments. However, individuals with hearing and speech disabilities often face challenges in communicating with those who do not understand sign language. The main purpose of this project is to develop a system that can recognize sign language using CNN. The system will take images of hand gestures and convert them into letter. This will help bridge the communication gap between people who use sign language and those who do not. The project also aims to make sign language learning easier for beginners. CNNs are particularly well-suited for this task due to their ability to

automatically extract spatial features from visual data, such as hand gestures and facial expressions, which are integral components of sign language.

1.2 Motivation

Many people who cannot hear or speak find it hard to communicate because most people do not understand sign language. This can make them feel lonely, frustrated, and left out in schools, workplaces and daily life. They may also miss out on good opportunities because of this communication problem. That is why there is a strong need for a simple and helpful tool that can make communication easier for them.

Convolutional Neural Networks (CNNs) are very good at recognizing hand movements and facial expressions which are important in sign language. This project motivates to help people with hearing and speech disabilities by creating a system that allows them to communicate easily in schools, offices, and public places. This can remove barriers and help everyone feel included. This project will use a CNN-based model to learn from pictures of sign language, such as the American Sign Language (ASL) alphabet or other available datasets. This will help people with hearing disabilities communicate better and live more freely.

2 Background and Present State of the System

Sign language recognition is an important area of research because it helps people with hearing impairments communicate. This field has grown with the use of machine learning, especially Convolutional Neural Networks (CNNs), which have shown good results in automatically recognizing sign language. These systems can identify hand gestures, facial expressions, and body movements, turning them into text or speech. Recent progress in sign language recognition has used transfer learning and CNNs to improve accuracy and speed. Sadik *et al.* [1] developed a system for recognizing Bangla sign language using transfer learning.This method combines pre-trained models with specific data to improve recognition without needing large datasets. This is especially helpful for languages with fewer available data.

In addition, Indian Sign Language recognition has also improved with machine learning. Katoch *et al.* [2] used SURF along with SVM and CNNs to recognize ISL signs. This technique works better than older methods by combining spatial features with deep learning models to improve performance.

Barbhuiya *et al.* [3] focused on CNN-based feature extraction and classification for sign language. Their research shows how CNNs can learn important details from raw image data, making the recognition system faster and more accurate.

Rastgoo *et al.* [4] reviewed many methods for sign language recognition and suggested improvements, including combining different data sources like vision, motion capture, and sensors. Current systems still face challenges.

3 Objectives

The proposed work will be done to reach the following objectives:

- To build a CNN model that can recognize sign language accurately.
- To help people with hearing and speech difficulties communicate easily.
- To test how well the CNN model works in recognizing signs.
- To develop a simple system that can recognize sign language.
- To use a collection of hand signs to train and improve the CNN model.
- To make the system easy to use for both learners and people with disabilities.
- To reduce errors in sign recognition by improving the model's accuracy.
- To explore ways to make the system work faster and more efficiently.

4 Literature Review

Our project aims to create a sign language prediction system using Convolutional Neural Networks (CNNs) with Python and TensorFlow/Keras. The goal is to recognize American Sign Language (ASL) signs from images or videos. The model will focus on real-time prediction and be optimized for mobile devices. We will use data preprocessing methods like resizing, normalization, and data augmentation (flipping, rotation, and contrast adjustments). To make it work on mobile devices, we will use TensorFlow Lite and OpenCV for video detection. Kasapbaşı et al. (2022) [5] trained a CNN model on an ASL dataset with different lighting and distances, reaching 99.38helps improve accuracy. We will also use data augmentation but focus on real-time processing with OpenCV. Murali et al. (2022) [6] built a real-time ASL system using HSV color segmentation and OpenCV for feature extraction. Their method was fast and used less computing power. We will use a similar real-time approach but improve accuracy with a custom CNN designed for mobile use. Garcia & Viesca (2020) [7] used transfer learning with GoogLeNet for better accuracy. However, transfer learning needs high computing power, which is not ideal for mobile devices. We will develop a lightweight CNN that works well in real-time on mobile phones. Das et al. (2020) [8] focused on static ASL signs and got 94.34validation accuracy. Their model worked well but could not recognize moving signs. Our model will work with both static and moving signs for better real-time use. Huang et al. (2014) [9] used 3D CNNs with Kinect depth data for better ASL accuracy. Their method needed special hardware, but we will use 2D CNNs with normal cameras to make our system

more accessible. Rathia et al. (2019) [10] used ResNet50 for ASL recognition and got 99.03will create a lightweight CNN for real-time mobile applications. Misra et al. (2011) [11] used Histogram of Oriented Gradients (HOG) with CNNs to improve recognition. HOG features are useful, but we will focus on deep learning-based feature extraction for better accuracy. Dong et al. (2015) [12] used a Kinect-based ASL system with depth and color data. It performed well but required extra hardware. Our system will work with regular cameras to make it easier to use and cheaper. Kala et al. (2014) [13] used Sparse Autoencoders for learning ASL features. While autoencoders are useful, CNNs are better for large ASL datasets. We will use CNNs for handling complex ASL signs. Bheda & Radpour (2017) [14] showed that deep CNNs work better than older machine learning methods for ASL recognition. We will follow their work but make sure our model is fast and optimized for mobile devices. Kulkarni & Lokhande (2010) [15] worked on gesture segmentation to improve ASL recognition. We will also use gesture segmentation but combine it with CNNs for better accuracy in recognizing different signs. Research shows that CNNs are very good for sign language recognition. Kasapbaşı et al. (2022) and Murali et al. (2022) showed that techniques like flipping, rotation, and contrast improvements help accuracy. We will use these methods to make our model more flexible. Many models, like Das et al. (2020), focus only on static signs. Our model will also recognize moving signs for better real-time use. Unlike hardware-based models like Huang et al. (2014), our model will work with normal cameras, making it easier to use. By combining the best ideas from these studies, we will create a simple, fast, and accurate sign language recognition system for real-world use.

5 Outline Methodology

5.1 Architecture

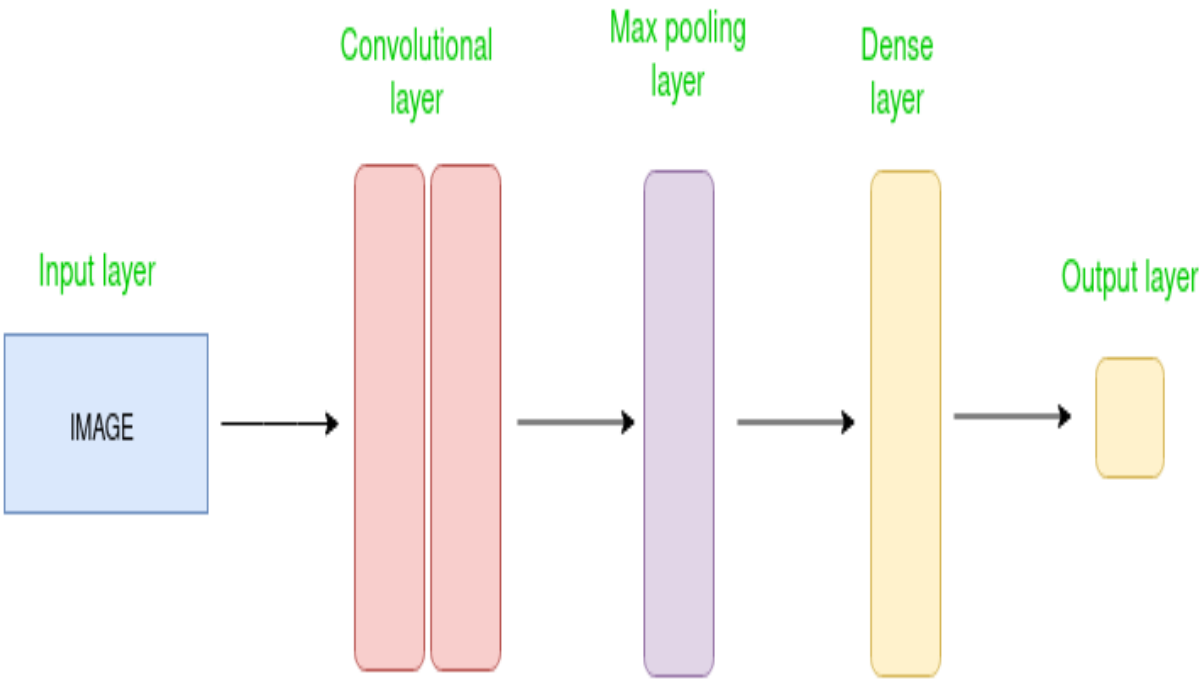

Figure 5.1: System Architecture Using CNN

5.2 Required Library

The following libraries are required to develop and implement the sign language recognition system:

1. **TensorFlow/Keras** - For building and training the model.
2. **OpenCV** - For image preprocessing.
3. **NumPy** - For numerical operations and handling arrays.
4. **Matplotlib** - For visualizing data and model performance.
5. **Scikit-learn** - For dataset preprocessing and evaluation metrics.
6. **Pandas** - For handling and analyzing structured data.

6 Conclusion

To conclude, it can be said that the sign language recognition project convert sign language gestures into letter using Convolutional Neural Networks (CNNs). By using the American Sign Language (ASL) dataset, the system will help people who don't know sign language communicate with those who use it. The system will carefully analyze ASL gestures and give accurate translations, making it easier for everyone to understand each other. This project will be especially useful for people with hearing impairments, allowing them to communicate better in schools, workplaces, and daily life. It also shows how deep learning can be used to solve real-world problems and make technology more helpful. The main goal is to create a tool that removes communication barriers and makes society more inclusive for everyone.

Bibliography

[1] M. R. Sadik, R. I. Sony, N. N. I. Prova, *et al.*, "Computer vision based bangla sign language recognition using transfer learning," in *2024 Second International Conference on Data Science and Information System (ICDSIS)*, IEEE, 2024, pp. 1–7.

[2] S. Katoch, V. Singh, and U. S. Tiwary, "Indian sign language recognition system using surf with svm and cnn," *Array*, vol. 14, p. 100 141, 2022.

[3] A. A. Barbhuiya, R. K. Karsh, and R. Jain, "Cnn based feature extraction and classification for sign language," *Multimedia Tools and Applications*, vol. 80, no. 2, pp. 3051–3069, 2021.

[4] R. Rastgoo, K. Kiani, and S. Escalera, "Sign language recognition: A deep survey," *Expert Systems with Applications*, vol. 164, p. 113 794, 2021.

[5] A. Kasapbaşi, A. E. A. Elbushra, A.-H. Omar, and A. Yilmaz, "Deepaslr: A cnn based human computer interface for american sign language recognition for hearing-impaired individuals," *Computer methods and programs in biomedicine update*, vol. 2, p. 100 048, 2022.

[6] R. S. L. Murali, L. Ramayya, and V. A. Santosh, "Sign language recognition system using convolutional neural network and computer vision," *Int. J. Eng. Innov. Adv. Technol*, vol. 4, pp. 138–141, 2022.

[7] B. Garcia, S. A. Viesca, *et al.*, "Real-time american sign language recognition with convolutional neural networks," *Convolutional Neural Networks for Visual Recognition*, vol. 2, no. 225-232, p. 8, 2016.

[8] S. Das, S. K. Biswas, and B. Purkayastha, "Occlusion robust sign language recognition system for indian sign language using cnn and pose features," *Multimedia Tools and Applications*, pp. 1–20, 2024.

[9] J. Huang, W. Zhou, H. Li, and W. Li, "Sign language recognition using 3d convolutional neural networks," in *2015 IEEE international conference on multimedia and expo (ICME)*, IEEE, 2015, pp. 1–6.

[10] P. Rathi, R. Kuwar Gupta, S. Agarwal, and A. Shukla, "Sign language recognition using resnet50 deep neural network architecture," in *5th International Conference on Next Generation Computing Technologies (NGCT-2019)*, 2020.

[11] R. K. Singh, A. K. Mishra, and R. Mishra, "Hand gesture identification for improving accuracy using convolutional neural network (cnn)," *The Scientific Temper*, vol. 13, no. 02, pp. 327–335, 2022.

[12] H.-D. Yang, "Sign language recognition with the kinect sensor based on conditional random fields," *Sensors*, vol. 15, no. 1, pp. 135–147, 2014.

[13] V. Kumar, G. C. Nandi, and R. Kala, "Static hand gesture recognition using stacked denoising sparse autoencoders," in *2014 seventh international conference on contemporary computing (IC3)*, IEEE, 2014, pp. 99–104.

[14] M. A. Ridwan, H. Mubarok, *et al.*, "The the recognition of american sign language using cnn with hand keypoint," *International Journal on Information and Communication Technology (IJoICT)*, vol. 9, no. 2, pp. 86–95, 2023.

[15] V. S. Kulkarni and S. Lokhande, "Appearance based recognition of american sign language using gesture segmentation," *International Journal on Computer Science and Engineering*, vol. 2, no. 03, pp. 560–565, 2010.