GLA University

SYNOPSIS

SIGN LANGUAGE RECOGNITION

Submitted By Submitted to:

Name: Khushi Gupta Faculty Name: Mr. Akash Kumar Choudhary

Id -201500342 Technical Trainer

DECLARATION

I **Khushi Gupta** of B.Tech (CSE) in 3rd Year hereby declare that the work submitted for the project of **SIGN LANGUAGE RECOGNITION** is my original work. I have not copied from any other student’s work or any other sources except where due reference or acknowledgment is made explicitly, and the work has been carried out under the guidance **of Mr. Akash Kumar Choudhary,** Department of Computer Science, GLA University. I further declare that the work on this project in this project has not been submitted and will not be submitted, either in part or in full, for any other degree in this institute or any other institute or university.

Candidate Name:

Khushi Gupta

Uni Roll No.- 201500342

Certification

Certified that this project report “**SIGN LANGUAGE RECOGNITION”** is the bonafide work of “**KHUSHI GUPTA”** who carried out the project work under **Mr. Akash Kumar Choudhary’s** supervision. Certified further that to the best of my knowledge the work reported herein does not form part of any other thesis or dissertation based on which a degree or award was conferred on an earlier occasion on this or any other candidate.

INDEX

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |  |  |  | | --- | --- | --- | | Serial No. | Name |  | | 1. | Introduction |  | | 2. | Justification |  | | 3. | Problem Statement |  | | 4. | Objective |  | | 5. | Future Scope |  | | 6. | Feasibility Study |  | | 7. | Methodology |  | | 8. | Testing |  | | 9. | Software |  | | 10. | Scope for Extension into a Major Project |  | | 11. | Conclusion |  | | |

# **INTRODUCTION: -**

The majority of persons who are deaf (hard of hearing) and people who utilise sign language to communicate inside and outside of their own community. People who are unable to talk or hear use hand gestures in this language to communicate. Sign Language Recognition (SLR) starts with the recognition of acquired hand motions and continues until text or speech is produced for comparable hand gestures. Here, there are static and dynamic hand gestures used for sign language. Although dynamic hand gesture recognition is more complex than static gesture recognition, both recognitions are crucial to the human community.

# By creating Deep Neural Network designs (Convolution Neural Network Architectures), where the model will learn to detect the hand motions images throughout an epoch, we may use Deep Learning Computer Vision to recognise the hand gestures. Once the motion has been successfully recognised by the model, a text file in English is generated and can later be converted to speech. Communication will be simpler for the deaf (hard of hearing) and disabled people thanks to this model's increased efficiency. In this work, we will examine Deep Learning's application to Sign Language Recognition.

# **Justification:-**

Even though new, easily usable technology are being developed to aid persons who have hearing impairments, more work has to be done. For instance, improvements in machine learning algorithms may make it possible for the hearing-impaired and hard-of-hearing to converse more effectively. That is what our effort attempts to do.

We set out to develop a system that could recognise hand motions used in American Sign Language (ASL). We had to develop a system that can recognise both static and dynamic hand motions because ASL uses both of them.

# **Problem Statement:-**

Many gestures are used in sign language to make it appear like movement language, which is made up of a series of hand and arms movements. There are various sign languages and hand gestures for various nations. Additionally, it should be emphasised that certain unfamiliar words can be translated by only demonstrating motions for each alphabet. Additionally, sign language has particular motions for each letter of the English alphabet and for every number from 0 to 9.

These sign languages can be divided into two categories: static gesture and dynamic gesture. The dynamic gesture is utilised for specific concepts, whereas the static gesture is used to symbolise the alphabet and numbers. Additionally, dynamic encompasses phrases, clauses, etc. The difference between the static and dynamic gestures lies in the movements of the hands, the head, or both. The three main elements of sign language—finger spelling, word-level vocabulary, and non-manual features—make it a visual language. Instead of using keywords, finger spelling is utilized to spell words letter by letter and deliver the message. Nevertheless, despite numerous research efforts over the past few decades, designing a sign interpreter is rather difficult.

Additionally, even identical signs seem very differently to different signers and from various vantage points. This research focuses on using a convolutional neural network to create a american sign language translator. We developed a lightweight network that can be utilised with low-resource embedded devices, independent programmes, and web applications.

# **Objective:-**

# The major goals of this research are to advance the field of text or speech translation and recognition of American sign language. My approach focuses on hand movements used in static sign language. This study employed convolutional neural networks (CNN) to recognise hand motions that included 26 English alphabet letters (A–Z) and 10 digits (0–9).

# **Future Scope:-**

Future scope includes:

* This established model can be applied to other sign languages, such Indian Sign Language, at this time it is confined to American Sign Language.
* The model could be further developed with a dataset to automatically separate the gesture from the collected frame by automatically removing the background.
* Fine-tuning and enhancement of the model to recognise frequent words and idioms.
* Additionally, it takes two arms to train the neural network model to recognise symbols in a well-organized manner.
* Additionally, it takes two arms to train the neural network model to recognise symbols in a well-organized manner. Add dynamic hand movements as an enhancement to the current static finger spelling.
* Adding the improved model to already-existing Artificial intelligence systems such Amazon Alexa to improve visual identification.

# **Feasibility Study:-**

An extensive evaluation of the operational (need), financial, and technological components of a proposal is typically part of a feasibility analysis. The purpose of a feasibility study is to evaluate a system proposal to see if it can be built within the limits of the proposed budget while still meeting the needs of the users. It also determines whether the system will be commercially viable. The earliest available time should be given for a moderately inexpensive feasibility study. The choice is taken whether to move forward with a more in-depth analysis depending on the study. A new project's feasibility is typically evaluated when it is suggested.

A feasibility study is conducted to ascertain whether the proposed system can be developed with the resources at hand and what should be the cost factor. Factors taken into account in the feasibility study were

1. Technical Feasibility
2. Economic Feasibility
3. Behavioural Feasibility

For this project, we require at least approx. 1-2 months for it to work. For developing or publishing we require at least economical feasibility at least of 5-6 lakhs.

During testing, developers find differences between the system and its

models by executing the system with sample input data sets. During unit testing,

developers compare the object design model with each object and subsystem.

During integration testing, combinations of subsystems are integrated and

Compared with the system design model. During system testing, typical and

exception cases are run through the system and compared with the requirements

model. The goal of testing is to discover as many faults as possible such that they

can be repaired before the delivery of the system

During testing, developers find differences between the system and its

models by executing the system with sample input data sets. During unit testing,

developers compare the object design model with each object and subsystem.

During integration testing, combinations of subsystems are integrated and

Compared with the system design model. During system testing, typical and

exception cases are run through the system and compared with the requirements

model. The goal of testing is to discover as many faults as possible such that they

can be repaired before the delivery of the system

# **Methodology:-**

CNN is concentrated on extracting temporal features. In this article, we will draw a connection between the two disciplines and use the CNN technique to identify specific gestures in images,as I implemented multiple classification algorithms to predict whether the person is using sign language or not as a mode of communication.

Two networks, each with three layers of convolution and three layers of max-pooling, were utilized by Lionel Pigou and colleagues. Both of the CNNs were trained to recognise features on the upper body and the hand.

# **Testing:-**

During testing, developers find differences between the system and its

models by executing the system with sample input data sets. During unit testing,

developers compare the object design model with each object and subsystem.

During integration testing, combinations of subsystems are integrated and

Compared with the system design model. During system testing, typical and

exception cases are run through the system and compared with the requirements

model. The goal of testing is to discover as many faults as possible such that they

can be repaired before the delivery of the system.

models by executing the system with sample input data sets. During unit testing,

developers compare the object design model with each object and subsystem.

During integration testing, combinations of subsystems are integrated and

Compared with the system design model. During system testing, typical and

exception cases are run through the system and compared with the requirements

model. The goal of testing is to discover as many faults as possible such that they

can be repaired before the delivery of the system

# We input the retrained model to the trained CNN to predict labels for certain gestures. Below is a description of the procedure. If we break down a gesture into frames once more, each frame has a corresponding prediction and label connected to it. The second method of solving the issue was capturing the CNN's pool layer output instead of relying on the predicted label it produced. Both systems used the same preprocessing, however the output differed and was stored separately.

# **Software:-**

* Python IDE
* Google Collab

# **Scope for Extension into a Major Project:-**

# The model will work effectively even when taught on a small dataset if it is pre-trained on a larger dataset, such as the roughly 14,000-class ILSRVC, and fine-tuned with the ISL dataset. In the case of models that are user-dependent, the user will give the model a series of images for development so that it could get to understand the user. The model will effectively serve a particular user in this way.

# **Conclusion:-**

In this report, we presented a machine learning method for recognising American Sign Language. It is a method for addressing the issues that those with speech and hearing impairments confront. Its two main parts are classifying images and understanding the movements in photos. A larger dataset might yield better results because the dataset we are working with is smaller.