`

**TITLE**

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING,**

**SHARDA SCHOOL OF ENGINEERING AND TECHNOLOGY, SHARDA UNIVERSITY, GREATER NOIDA**

**Real Time Sign Language Detection for Deaf and Mute Using OpenCV and CNN**

***A project submitted***

***in partial fulfillment of the requirements for the degree of Bachelor of Technology in Computer Science and Engineering***

**by**

**Kartik Tripathi (2019554678)**

**Md Zunad Alam (2019001796)**

**Shursti Sanket (2019501386)**

**Supervised by:**

**Ms. Gunjan Aggarwal**

**Assistant Professor**

May, 2023

# CERTIFICATE

This is to certify that the report entitled **“Real Time Sign Language Detection for Deaf and Mute”** Submitted by **“**Kartik Tripathi (2019554678), Md Zunad Alam (2019001796) and Shursti Sanket (2019501386)”to Sharda University, towards the fulfillment of requirements of the degree of **“Bachelor of Technology”** is record of bonafide final year Project work carried out by them in the **“**Department of Computer Science & Engineering, Sharda School of Engineering and Technology, Sharda University**”**.

The results/findings contained in this Project have not been submitted in part or full to any other University/Institute forward of any other Degree/Diploma.

**Signature of the Guide Signature of Head of Department**

**Name:** Ms. Gunjan Aggarwal **Name:** Prof. (Dr.) Nitin Rakesh

**Designation:** Assistant Professor ( **HOD of CSE)**

**Signature of External Examiner**

**Date:**

# ACKNOWLEDGEMENT

A major project is a unique chance for learning and self-development. This project would not have been possible without the great leadership and support of so many wonderful people involved in this project.

Our first and foremost gratitude goes to Dr. Nitin Rakesh, Head of CSE, who gave us this opportunity.

We would like to thank Mrs. Gunjan Aggarwal for her guidance in our project work. Despite being extraordinarily busy with academics, Mrs. Gunjan Aggarwal took time out of her busy schedule to listen, guide, and keep us on the right path. With her help, we would not have been able to accomplish what we did.

As part of the CSE department's monitoring of our progress, all facilities were arranged to make our lives easier. We would like to take this opportunity to express our gratitude for their contributions.

GitHub Project Link: https://github.com/Zunad1/FinalYearProject

Name and signature of Students:

Kartik Tripathi (2019554678)

Md Zunad Alam (2019001796)

Shursti Sanket (2019501386)

# ABSTRACT

# According to World Health Organization, over 5% of the world’s population have hearing and speaking disabilities. Sign language is the primary method of communication for persons who are deaf and mute. This may make it difficult for them to communicate on a daily basis. To overcome this difficulty, sign language may be translated into text. American Sign Language is intended to be recognized by the proposed system, which then converts it to text. The system receives as input an image of a hand spelling out the required alphabet. The Bhattacharyya Distance Metric is then used to compute the input image's histogram and compare it to the histograms of previously saved images. The proposed system processes images using OpenCV as a tool. The associated alphabet of the image whose histogram is most similar to the input image's is then checked, and the associated alphabet is printed. The system's implementation will be a modest step towards removing the social communication barrier that occurs between sign language non-users and deaf-mute people.

**CONTENTS**

[TITLE i](#_bookmark0)

[CERTIFICATE ii](#_bookmark1)

[ACKNOWLEDGEMENT iii](#_bookmark2)

[ABSTRACT iv](#_bookmark3)

[LISTOF FIGURES v](#_bookmark4)

[LISTOF TABLES vi](#_bookmark5)

[CHAPTER1: INTRODUCTION](#_bookmark6) 8

* 1. [Problem Statement](#_bookmark7) 9
  2. [Project Overview](#_bookmark8) 10
  3. [Expected Outcome](#_bookmark9) 11
  4. Possible risks 12
  5. [Hardware & Software Specifications](#_bookmark10) 13
  6. [Other Non-Functional Requirements](#_bookmark11) 15
  7. Report Outline 16

[CHAPTER2: LITERATURESURVEY](#_bookmark13) 17

* 1. [Existing Work](#_bookmark14) 17
  2. [Proposed System](#_bookmark16) 22
  3. [Feasibility Study](#_bookmark17) 23

[CHAPTER3: SYSTEM DESIGN& ANALYSIS](#_bookmark18) 25

* 1. [Project Perspective](#_bookmark19) 25
  2. [Performance Requirements](#_bookmark20) 26
  3. [Methodology](#_bookmark22) 28
  4. Testing Process 30

[CHAPTER4: RESULTS AND OUTPUTS](#_bookmark24) 34

* 1. [Implementations](#_bookmark25) 34
  2. [Proposed Model Outputs](#_bookmark25) 35

[CHAPTER5: CONCLUSION 3](#_bookmark29)7

* 1. Conclusion and Future Scope [. 3](#_bookmark30)7

[CHAPTER6: REFERENCES](#_bookmark32) 38

[ANNEXURE I](#_bookmark32) 40

[ANNEXURE II](#_bookmark32) 41

# LISTOFFIGURES

|  |  |  |
| --- | --- | --- |
| Figure No. | Table Name | Page No. |
| Fig 2.2.2 | Proposed Work | 22 |
| Fig.3.3.1 | Architecture of CNN | 26 |
| Fig.3.4.1 | Basic Classification of sign language | 28 |
| Fig.3.5.1 | Use Case Diagram | 32 |

# LISTOFTABLES

|  |  |  |
| --- | --- | --- |
| Table2.1.1 | Project Outline | 16 |
| Table2.1.1 | Comparative Analysis | 20 |
| Table 3.5.2 | Use Case Descriptions | 33 |
| Table 4.2.1 | Results | 36 |

1. **INTRODUCTION**

Sign language detection is the process of recognizing and interpreting sign language gestures in order to translate into written language. This technology is particularly useful for facilitating communication between deaf or hard of hearing individuals and those who do not know sign language. There are many different sign languages in use around the world, each with its own grammar and vocabulary. Some of the most widely used sign languages include American Sign Language (ASL), British Sign Language (BSL), and Indian Sign Language (ISL). There are several different methods for sign language detection, including using computer vision techniques to detect and recognize hand and body movements, and machine learning algorithms to interpret and translate the gestures into written or spoken language. Computer vision techniques involve analyzing video footage or images of the signer and tracking the movement of the hands and body. This can be done using various types of sensors such as cameras and depth sensors, and by applying techniques such as image processing and pattern recognition. Machine learning algorithms are trained on large datasets of sign language gestures to learn how to recognize and interpret different signs. These algorithms use various techniques such as deep learning and natural language processing to recognize and translate the signs into text or speech. Sign language detection has the potential to revolutionize the way that deaf and hard of hearing individuals communicate with the hearing world[1]. Additionally, it can be applied to increase accessibility in a range of contexts, including healthcare, education, and public transportation. However, there are still challenges in the development of this technology, such as the need for more robust and accurate detection algorithms, and the need for greater standardization of sign language vocabulary and grammar. Learning sign language is a valuable skill for anyone looking to communicate with deaf or hard of hearing individuals, and it can also be a way to enhance communication skills in general. Many resources are available to help individuals learn sign language, including books, online courses, and classes offered through community organizations and schools.

* 1. **Problem Statement**

In order to facilitate communication between Deaf & Mute individuals and hearing people, a language system is formed by the structure of sign language, which is distinct from written language. In order to engage, they rely on vision-based communication.

The actions can be clearly understood by other individuals if there is a standard interface that transforms sign language to text. As a result, research has been done on a vision-based communication system that will allow Deaf & Mute persons to communicate without actually speaking the same language.

The goal is to create a human computer interface (HCI) that is simple to use and can recognize human sign language. Worldwide, there are many different sign languages, including the American Sign Language (ASL), French Sign Language, British Sign Language (BSL), Indian Sign Language, and Japanese Sign Language. Work has also been done on other sign languages.

**1.2 Project Overview**

Sign language is the most widely used mode of communication between deaf and mute person and a normal human being. It is not easy to converse with deaf and mute people so to reduce the gap a bridge was created in terms of Sign language. A widely accepted mode of communication that is easy to learn and also have some consistency everywhere. While normal being might not find a need to learn sign language which make it is extremely difficult for any deaf and mute person to understand the. So, we are preparing a model which will take in the video stream of the deaf person using the sign language gestures and converting them to the appropriate text. The model uses CNN and RNN also TensorFlow for detection. This way the normal person is also able to understand the language. This is useful as the conversation between the normal and deaf and mute person becomes easier. It is suggested to create a model with perfect precision, which would enhance the previously provided functionalities in the current systems by including new features and greater functionality under difficult circumstances. We analyzed different algorithms and prepared a comparative analysis as well.

**1.3 Expected Outcome**

I. To develop a model or system which will detect the sign language with higher accuracy.

II. Working on different module in the future.

III, Easier communication between a normal and a deaf and mute person.

1. Time saving with better functionality.
2. High precision using histogram images.
   1. **Possible risks**

* **Availability and Quality of Data**

One of the main limitations of sign language detection project is the availability and quality of data. Deep learning algorithms such as CNNs require large amounts of high-quality data to train and validate the model. If there is a limited amount of data available, or if the data is of poor quality, this can lead to unreliable results.

* **Interpretability of Results**

Another limitation of sign language detection project is the interpretability of the results. While deep learning algorithms such as CNNs can often achieve high levels of accuracy, it can be difficult to interpret how the model arrived at its conclusions.

* **Ethical and Legal Concerns**

There are also ethical and legal concerns associated with the project. For example, the use of sensitive medical data raises concerns about privacy and security, and the use of deep learning algorithms can lead to concerns about algorithmic bias and accountability.

* **Technical Challenges**

There are also technical challenges associated with the project. Developing and training deep learning algorithms can be computationally intensive, requiring significant amounts of computing power and storage. While sign language detection project has the potential to improve the accuracy of sign language. These include the availability and quality of data, the interpretability of results, ethical and legal concerns, and technical challenges. To address these limitations, it is important to carefully design the study, use high- quality data, validate and test the model regularly.

**1.5 Hardware & Software Specifications**

**Software**

1. **Python:**

The programming language Python is high-level and versatile. It is open-ended and can be used to create a wide variety of programmers because it is not focused on any specific problems. Deep indentation is a key component of Python's design philosophy, which emphasizes the importance of code readability. Due to its open-source nature, Python is available to use, fairly simple to understand, and easy to learn.

1. **Pip:**

Pip is a Python-based package management tool that is used to install and manage software packages. The Python Software Foundation recommends using pip for installing Python applications and its dependencies during deployment. To install a specific version of the package, instead of the latest one, we may specify the version directly with the pip command. It will install the package with the mentioned version.

1. **NumPy:**

NumPy is the abbreviation for numerical Python. Huge, multi-dimensional arrays and matrix are supported by the Python library NumPy, as are a wide range of high-level arithmetic operations that can be carried out on these arrays.

1. **TensorFlow:**

TensorFlow is an available as an open software library for artificial intelligence and machine learning. Neural network-based training and inference are given specific attention, despite the fact that it can be employed for many different tasks. TensorFlow helps to accelerate and simplify the creation of neural network models and machine learning algorithms.

**Hardware**

1. Processor: intel core i5 or above
2. 64-bit, quad-core, 2.5GHz minimum per core
3. Ram: 4Gb or more
4. Hard disk: 10 Gb or above
5. Operating system: windows, Mac, etc.
   1. **Other Non-Functional Requirements**

Here, we shall discuss the system's non-functional requirements.:

1. **Availability** - The software for Extraction of sign language from images can be available in all the systems.
2. **Reliability** - This software attempts to ensure appropriate content but assume no responsibility for external manipulations.
3. **Usability** - The proposed systems easy to use for user.
4. **Scalability** - The capability of proposed system to handle a growing number of deaf and mute person.

**1.6 Report Outline**

|  |  |  |  |
| --- | --- | --- | --- |
| **Project Name** | | **Start Date** | **Deadline Date** |
| Real Time Sign Language Detection for Deaf and Mute | | August, 2022 | May, 2023 |
| PROJECT SUMMARY | We are going to build an easy to use and accurate sign language detection system which helps in the bridging the gap between a normal and a deaf and mute person. The model use CNN, RNN and computer vision libraries for the detection of sign language using the addition layers of histogram and Gaussian filters. | | |
| PROJECT OBJECTIVES | * Develop a model for easy sign language detection * Higher accuracy for the system * Less budgeted system. * Not highly data and memory demanding * Reducing elapsed time compared to the existing systems. | | |
| SCOPE OF WORK | * A good model should work for most of the test cases and edge cases. * The data stored can be used for future improvement of the dataset. * A mobile app too can be created at some point for cross platform usage. | | |
| PROJECT TIMELINE | Estimated Duration: 8 months (July,2022 – March,2023)  Phase 1: Identification of Problem.  Phase 2: Detailed analysis of Feasibility  Phase 3: Work on review Paper  Phase 4: Detailed Design & Core Implementation  Phase 5: Work on Research Paper | | |
| PROJECT TEAM | Kartik Tripathi (2019554678)  Md Zunad Alam (2019001796)  Shursti Sanket (201900) | | |

Table 1.6.1

1. **LITERATURESURVEY**

**2.1 Existing Work**

**Mr. J. J et ai.,** [2] The authors discussed various hand gesture and sign language detection and recognition that have been put forth in the past by numerous researchers. For the dumb and the deaf, sign language is their sole form of communication. These persons with physical disabilities use sign language to express their emotions and ideas to others.

**Pramada, Sawant** **et al.,** **[3]** The barrier that is put up between people who are disabled or handicapped and the average person is one of our society's biggest flaws. The only method we can disseminate ideas or a statement is through interaction, yet those who have disabilities (such those who are deaf or foolish) find it challenging to interact with others in a typical way. Many people who are deaf or dumb rely on sign language as their primary form of communication. To make it easier for the deaf to interact with the hearing community, sign language recognition (SLR) tries to automatically translate sign languages. With the aid of the SURF method and image processing, this method recognizes and extracts hand 23 motion features. MATLAB software is used for all of this work. This method makes it simple to instruct a deaf and mute individual.

**Sruthi, R Rao et al., [4]** With the help of this study's work, average people can use a sample hand gesture detection system to better connect with special persons. The aforementioned research work is concerned with the problem of real-time gesture identification in the sign language used by the deaf community. The problem is based on the processing of images and utilizes Color Classification, Skin Detection, Segmentation Methods, Picture Filtering, and Feature Matching techniques. This system recognizes gestures from American Sign Language (ASL), including the letters and some of its sentences.

**Arora et al.,**[5] A technique for image processing (the identification of skin and marker pixels) has been utilized to recognize English alphabetic sign language without needing the hand to be exactly aligned to the camera. This system uses a simple and highly accurate coordinate computation, color calibration, and pattern matching algorithm for segmentation, the threshold used to extract features, and recognition. However, users must wear a particular color band on their fingers.

**Rugia S.T. Kamal et al.,**[6] The overall goal of the project is to make communication simpler between regular people and persons who are deaf or dumb by accurately translating sign language to voice or text. The mute and the deaf communicate through sign language, but it can be confusing to anyone who are not familiar with it. Therefore, a gadget that can convert gestures into speech and text must be created. This will be a big step in enabling deaf and mute persons to interact with the general public.

**Tahir Khan** **et al.,**[7] The application of contemporary methods considerably broadens the potential applications of conventional microscopic techniques in the forensic area, enabling the collection of essential quantitative data in forensic studying the pedological phases and identifying the mineral phases, or the option of organic phase analysis directly in the SEM chamber.

**Cao dong et al.,**[8] In this study, the researcher recognized American sign language using Microsoft Kinect. To detect the ASL alphabet, a depth camera is used as a Kinect sensor. The feature extraction process used a distance adaptive scheme. For classification purposes, support vector machines and RF classifier algorithms are used. An ANN network was used to train the data. 90% of the time, the system was accurate.

**R Rumana et al.,**[9] In the paper, it was discussed how body language and manual gestures have been used to develop sign language recently. Pre-processing, feature extraction, and classification are the three steps that most sign language recognition systems include. Scale-invariant feature transforms (SIFT), Hidden Markov models (HMM), neural networks (NN), support vector machines (SVM), and others are classification techniques used for recognition.

**K.Dixit et al.,** [10] The approach for automatic sign recognition utilizing shape-based characteristics is described in this study. The hand region is separated from the images using Otsu's thresholding method, which selects an optimal threshold to reduce the variance of threshold black and white pixels within a class. Hu's invariant moments are used to split the hand region, and its features are computed. These features are then fed into an artificial neural network for categorization. The system's performance is evaluated using Accuracy, Sensitivity, and Specificity.

**Y. Dawod et al.,** [11] Using Indian sign language, the author of this essay suggested a system that would convert hand gestures into the appropriate text messages to enable communication between hearing people and deaf and dumb people (ISL). After testing is finished, the system will be put into use on the Android operating system and made accessible as an application for mobile devices like smartphones and tablets. The main objective is to develop an algorithm that can instantly translate dynamic gesture to text.

**P. R. Kumar et al.,** [12] The approach suggested in includes separating the hand motions from the original color visuals. Using the Chan-Vese (CV) active contour model, the segmented hand positions' shape was manipulated, yielding a 92.1% identification rate.

**P. V. V. Kishore et al.,** [13] A feature matrix with elliptical Fourier descriptors was constructed from the segmented hand motions and extracted forms, and an artificial neural network trained on the back propagation approach was used to classify the features. The suggested 4 camera model's average recognition rate for sign language recognition is around 92.23%.

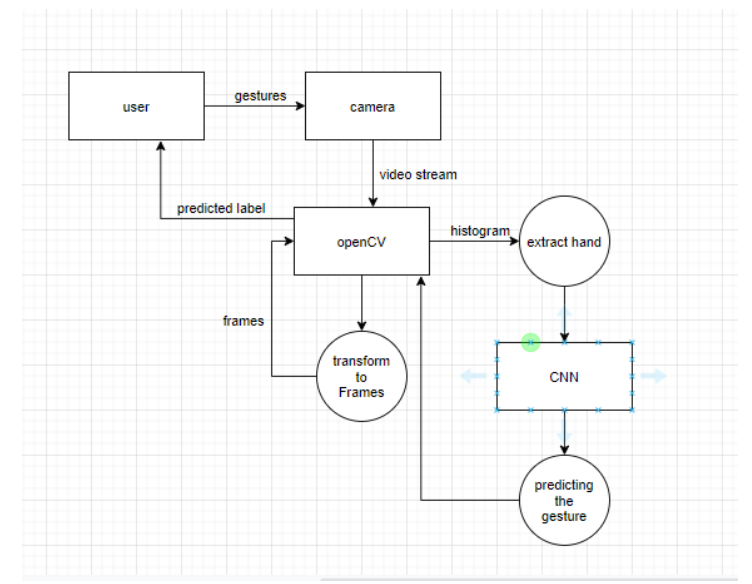
**M. Kumar et al.,** [14] The most effective approach for transliterating 24 static ASL alphabet movements are demonstrated in this research study (“Letter J and Z are not included since they need hand movement”). The translation of 10 fixed ASL number motions into English text and handling of video pixels are therefore required. The main goal is to make a humanoid or machine-readable English document out of 24 fixed ASL alphabets and numerals.

**COMPARATIVE ANALYSIS**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Authors** | **Year** | **Algorithm/ Methodology** | **Accuracy** | **Observation** |
| Paulo Trigueiros  Luís Paulo Reis [15] | 2017 | Convolutional Neural Networks (CNN), Faster R-CNN, modified VGGNet models. | 93.3% | A dataset comprising twenty-six letter signs is retrieved from various video footages. Each letter sign has 1000 images on average with total 28000 images. |
| Meita Chandra Ariesta, Ricky Anderson [16] | 2017 | DTW/ISODATA algorithms, Welch algorithm, Viterbi algorithm | 78% | Sign acquisition, vision-based systems, and sign recognition methods all benefit from artificial neuron networks |
| Ashok K Sahoo, Gouri Sankar Mishra [17] | 2012 | Examination is done of the data collection, pre-processing, transformation processes, feature extraction, categorization, and outcomes. | 88.6% | The model performs well in at the ChaLearn 2014 Observing at Persons movement detection contest, earning a joint Jaccard Index of 0.789. |
| Dipalee Golekar,  Sidheshwar Katare[18] | 2022 | Skin-Colour Segmentation, Edge Segmentation, L-BFGS | 98.2% | A real-time sign language gesture identification system was constructed using computer vision and M.L techniques. A sophisticated segmentation method was used for extracting finest skin marks |
| Mahesh Kumar N B[19] | 2020 | Segmentation, dilation. erosion, feature extraction | 95% | This SLR system recognises a group of generated signals and converts them into text or voice that includes the relevant factors. |
| Sander Dieleman,  Pieter-Jan Kindermans[20] | 2015 | We make use of the ChaLearn Looking at People data set  issue from 2014 (CLAP14) in this work | 78% | The model performs well in at the ChaLearn 2014 Observing at Persons movement detection contest, earning a joint Jaccard Index of 0.789. |
| Zakariah, Mohammed  Alotaibi, Yousef Ajmi  Koundal[21] | 2022 | Otsu's Algorithm-Based Indian recognition of sign languages System for Deaf Persons. | 88.6% | The majority of these advancements are predicated on the concept of interaction fragments, which stand for more compact portions of an encompassing interaction. |
| Shwetha S kulloli , R H Goudar [22] | 2018 | K-Nearest Neighbour Analysis (KNN), Probabilistic Neural Networks **(**PNN) | 87% | Recognition of Words, Sentences using Vision Based Approach |
| Amrita Thakur, Pujan Budhathoi [23] | 2020 | K-Nearest Neighbour Analysis (K-NNA) and Principal Component Analysis Algorithm (PCA) | 66.6% | The most effective way for translating 24 static ASL alphabet movements is demonstrated in this study. The handling of video pixels and the translation of 10 predetermined ASL numerical movements into English text are both necessary. |

Table 2.1.1

**2.2 Proposed System**



**Fig 2.2.2 Proposed Work**

* 1. **Feasibility Study**

1. **Technical Feasibility:**

Technical feasibility examines and assesses available hardware, software, and technological resources for the project's development. This technical feasibility analysis reveals whether the necessary tools and technologies are available for use in project development.

1. **Operational Feasibility:**

Operating potential examines the simplicity of operating and maintaining the product after deployment, as well as the degree of service delivery in accordance with the criteria. Examine whether consumers feel at ease using new software. Check to see if the solution the software development team has offered is acceptable.

1. **Economic Feasibility:**

A profitability learning analyses costs and advantages. This indicates that a thorough review of the development project's costs will be done as part of this feasibility study. This sums up all expenses required for the final development, such as those for design and development, running expenses, required hardware and software resources, etc. The project's potential financial value to the organization is then examined.

1. **Legal Feasibility:**

A legal feasibility assessment examines the proposal from the standpoint of legality. Analysis of potential legal barriers to the project's implementation, including, but not limited to, data protection or social media regulations, project certificates, licenses, and copyrights. A legal feasibility study often looks into the legal and moral compliance of a prospective business.

1. **Schedule Feasibility:**

A schedule feasibility study primarily examines the suggested project deadlines, including the amount of time the team will need to finish the final product. This has a big effect on the company since if the project isn't finished on time, its goal might not be achieved.

1. **SYSTEM DESIGN & ANALYSIS**

**3.1 Project Perspective**

1. **Software Interface:**

* Our model will work on python language with the requirement such as TensorFlow, OpenCV, NumPy and Histogram.
* Minimum requirement of RAM is 4 GB
* It captures the images and convert them into grayscale images.
* The feature extraction is done and all noise is removed

1. **Dataset Creation**

* The interface stores the data in the dataset folder created by the model.
* Minimum 2gb storage should be available for image storage.

1. **Communication interface**

* After the storage in data set model is set to extract the features from the live video feed and give the output.
  1. **Performance Requirement**

1. **Response Time:**

* Measure how quickly the model responds when the user gives the live video feed.

1. **Workload:**

* Extracting only the useful part of the images and edges are more visible using the histogram and Gaussian filter.

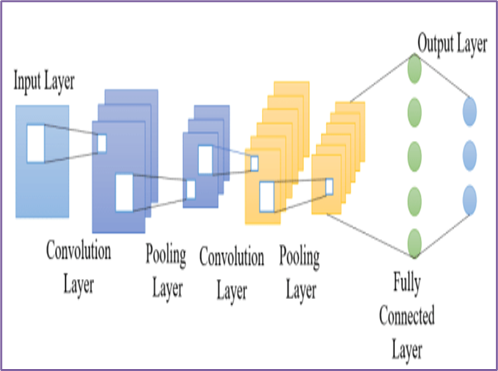
1. **Scalability:**

* Scalability is the increase in the system’s workload that the system should be able to process.

1. **Platform:**

* Creating a GUI for easy access of the program.
  1. **Methods for Sign Language:**

**Convolutional Neural Networks (CNN):** Convolutional neural networks, sometimes known as deep neural networks artificial neural networks, largely utilized for categorization of photos, grouping them in accordance with their resemblance and is capable of recognizing objects. These Algorithms can recognize people, street signs, objects, cancers, and other features of visual information. Using convolutional networks to capture text, use the optical character recognition (OCR). It enables analogue and digital devices to process natural language papers that have been handwritten and contain symbols for pictures be written down. The introduction of LeNet architecture will be executed using Convolutional architecture [23]. CNN represents significant developments in computer vision (CV), and applications for robots, drones, security, and self-driving automobiles medical evaluations and therapies for those who are blind.

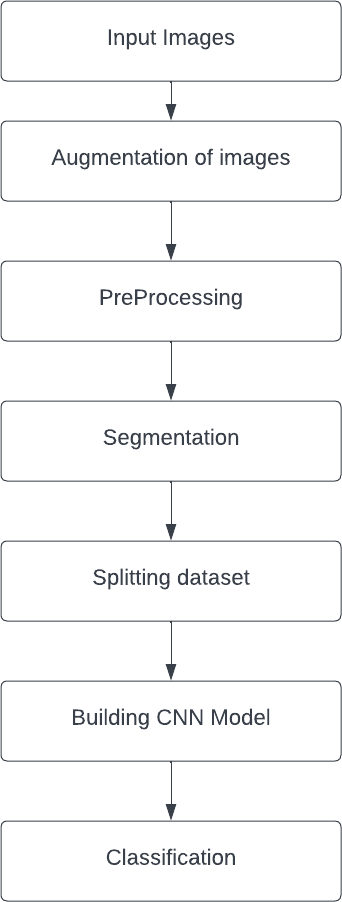


**Fig.3.3.1. Architecture of CNN**

**K-Nearest Neighbor (KNN**): The K-nearest neighbors (KNN) algorithms are supervised machine learning algorithms that can handle relatively difficult classification tasks while being extremely simple to develop. While attempting to categorize a new information instance, all data is used for training. It is a non-parametric technique used for classification and regression since it makes no assumptions about the underlying data [24]. The stages for implementing the KNN algorithm in Python are as follows: processing the input, calculating the distance, finding the k closest points, classifying the data, and evaluating correctness. Sample vectors are categorized using the MATLAB tool "knn-classify". Unsupervised classifier KNN is used.

**Probabilistic Neural Networks (PNN)**: An effective multi-class classifier uses probabilistic neural networks. A multi-layered feedforward network with four layers—input layer, pattern layer, summation layer, and output layer—that organizes the operations is used to implement the statistical algorithm known as kernel discriminant analysis, or PNN. Principally utilized in pattern recognition and classification. On the basis of a feed-forward neural network [24]. Network creation is done using the MATLAB command "newpnn". Predictor variables are present for each neuron in the input layer. N-1 neurons are employed for every N category. By taking the median out of the equation and dividing the result by the midspread range, the range of the values is standardized. The hidden layer then receives input neurons. For each case in the training dataset, there is one neuron in the pattern layer, which has Gaussian functions created with the provided set of information points as its centers. The sigma values are used to calculate the RBF kernel, which is then applied to the test case's Euclidean distance from the center of the neuron. The summation layer does exactly what its name implies: it adds the outcomes from the two layers for each class.

**3.4 METHODOLOGY**



**Fig.3.4.1 Basic Classification**

### Input Images:

This is the first step of our proposed model in which we give input video stream.

### Augmentation of images:

The quantity of pictures is less so in order to increase. The quantity of images in dataset, we need to perform the augmentation process. It is one of the most useful techniques which helps to increase the size of data without need of new images.

### Preprocessing:

In order to enhance the quality of the hand gesture image, pre-processing is a crucial and initial step. Impulsive noise reduction and image scaling are crucial pre-processing procedures. In the initial stage, we crop the images and convert the hand image into its equivalent gray-scale image. This raises the accuracy rate of classification and the diagnosis.

### Segmentation:

Image segmentation is a method for dividing an image into various components. The main goal of this division is to keep the quality of the images while making it simple to examine and interpret them. In order to analyses the size, volume, position, texture, and shape of the extracted image, segmentation methods are useful since they have the ability to discover or identify the anomalous component from the image. The simplest technique for segmenting images is thresholding. The process of turning a greyscale image into a binary image, where the two levels are allocated to pixels with values above or below the set threshold value, is a non-linear operation.

**Classification:**

Classification is one of the best ways to identify images, including any kind of medical imaging. Every algorithm for categorizing things is predicated on the notion that an image has one or more features, each of which is associated with a specific class. The constructed CNN model is used to categorize the words and display the appropriate results.

**3.5 Testing Process**

**Data Collection**: Collecting a diverse dataset of sign language videos is crucial to training and testing the accuracy of the detection system. The dataset should include different signers, genders, ethnicities, lighting conditions, camera angles, and signing styles.

**Annotation**: Annotation involves adding metadata to the dataset, such as the gloss or meaning of each sign, the signer's identity, and the timestamp. This process is essential to train the machine learning algorithms that detect the signs.

**Pre-processing**: The videos may need to be pre-processed to standardize the lighting, cropping, and frame rate. This step aims to remove noise and irrelevant information from the video and improve the quality of the input.

**Training**: The pre-processed data is used to train a machine learning model, which can learn patterns and relationships between the input (video frames) and output (sign language signs).

**Validation**: A validation set is used to test the accuracy and reliability of the trained model. This step involves testing the model on a dataset that the model has not seen before and measuring its performance using metrics such as precision, recall, and F1-score.

**Testing**: The final step involves testing the model on a real-time sign language input stream and evaluating its performance in real-world scenarios.

**Iteration:** Based on the results of the testing phase, the model can be retrained and improved, and the testing process repeated until the desired level of accuracy is achieved.

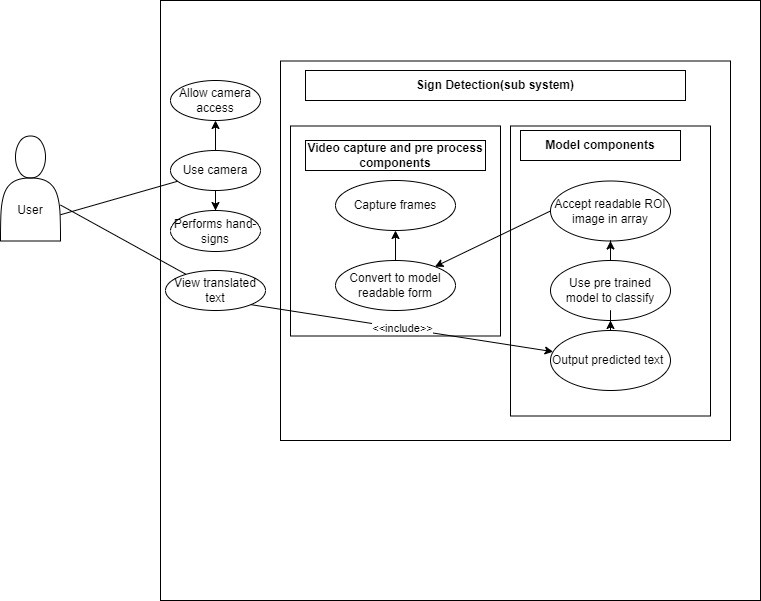
* + 1. **Use case diagram**

A use case diagram is a visual depiction of a system's functionality that demonstrates the connections between actors and use cases. A use case diagram may be used to show the numerous actors who engage with the system and the various use cases that the system must serve in the context of sign language recognition.

The purpose of a use case diagram for sign language recognition is to deliver a clear, concise depiction of the system's functionality that can be easily understood by stakeholders. Having a common grasp of the system's needs and the ability to collaborate efficiently to produce a high-quality output may be ensured as a result.

In a use case diagram for sign language recognition, the actors might include the user, who interacts with the system through a user interface, and the sign language interpreter, who provides feedback to the user. The use cases might include tasks such as capturing video of sign language gestures, analyzing the video to recognize the signs, and providing feedback to the user based on the recognized signs.

Developers may better grasp the demands of the system and build a solution that satisfies everyone's needs by producing a use case diagram for sign language recognition. It also helps to ensure that the system is intuitive and easy to use, improving the overall user experience.



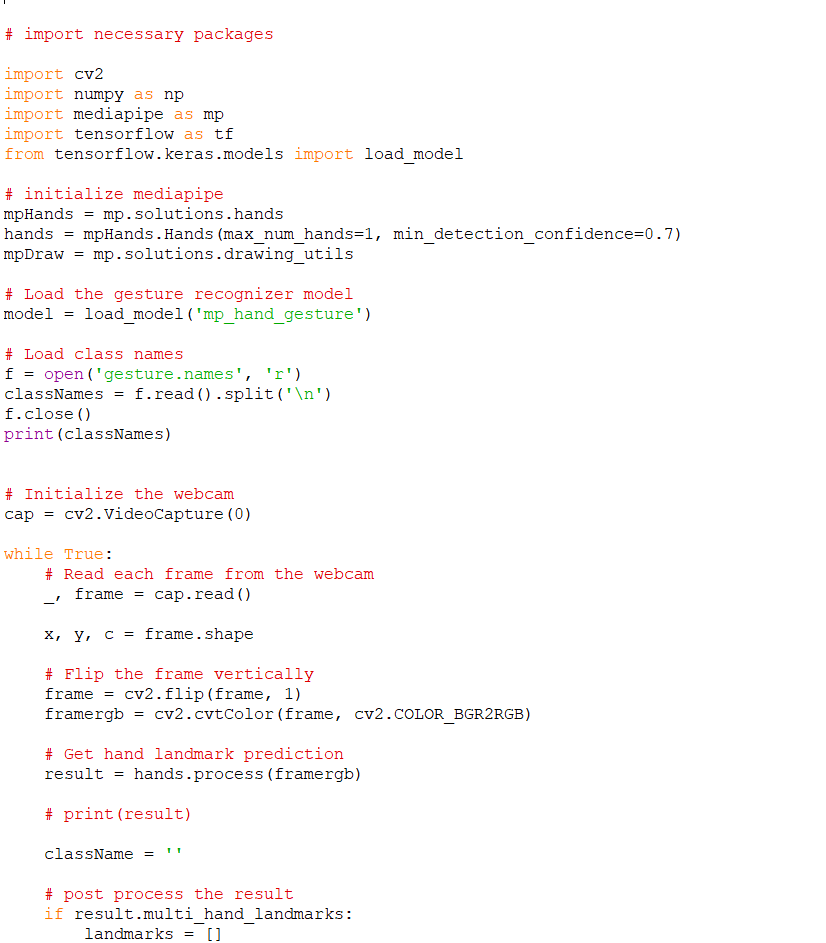
**Fig.3.5.1 Use Case Diagram**

**3.5.2 Use case description**

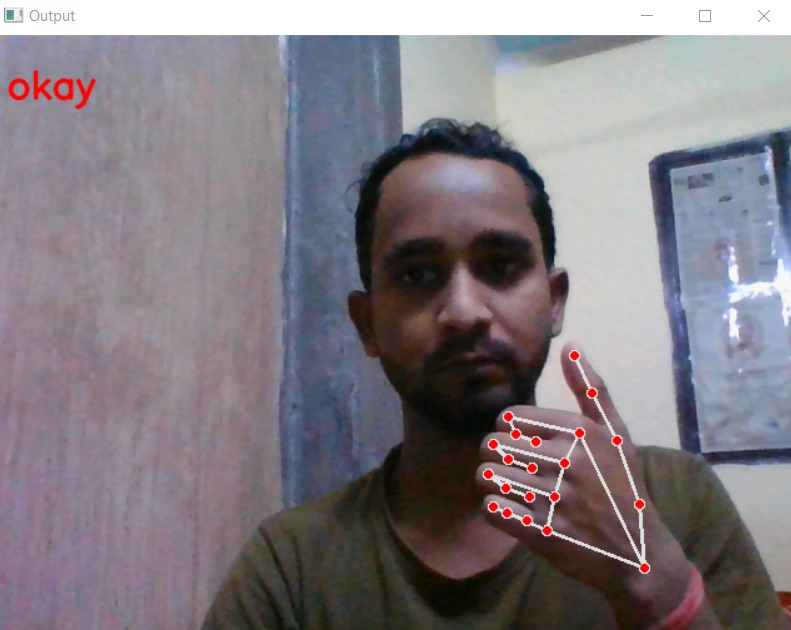
|  |  |
| --- | --- |
| **Use Case ID:** | S2T |
| **Use Case Name:** | Sign 2 Text |
| **Users:** | Normal User, Deaf User, Mute user |
| **Description:** | A user opens the application and navigate to camera window to capture the video stream of signs of deaf people and application will translate these signs to text. |
| **Trigger:** | A deaf/mute user will act as an input for machine learning model and he provides signs and gestures to model for training and evaluation. |
| **Preconditions:** | PRE-1. User should provide camera access.  PRE-2. Application must have internet access to work correctly.  PRE-3. Signs should appear in the certain position in camera window for better results and accuracy. |
| **Post conditions:** | POST-1. All the generated results should be stored in database so that it can be viewed. |
| **Normal Flow:** | User will have to open camera in application and capture the hand signs and gestures. Then these signs will give to the model to predict the result against signs and result will be shown. Then user save the output results to database so that it can be viewed later. |
| **Alternative Flows:** | E1. If user has not given camera access to application  1. It will display error message.  E2. If application has no internet connection.  1. It will generate predicted text but accuracy would be low. |

**Table 3.5.2 Use Case Descriptions**

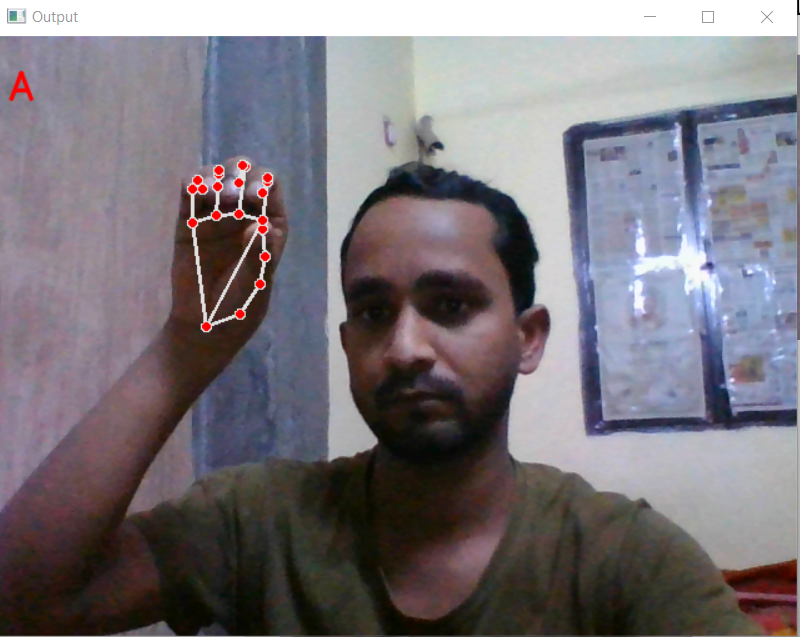
1. **RESULTS AND OUTPUTS**
   1. **Implementations:**

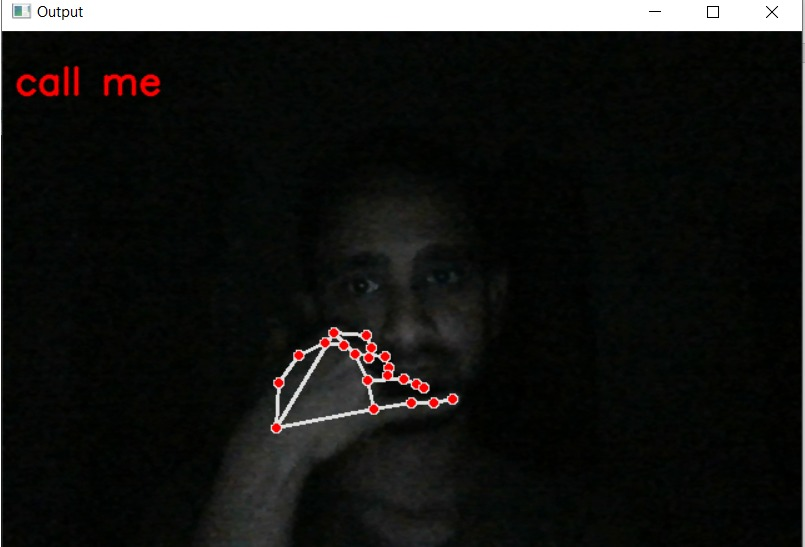
****

* 1. **Proposed Model Outputs**

****

****

****

****

**The proposed system's accuracy**

|  |  |
| --- | --- |
| Gesture Name | Accuracy (%) |
| OKAY | 87.7 |
| STOP | 84.6 |
| THANK YOU | 89.3 |
| CALL ME | 81.4 |

**Table 4.2.1 Results**

1. **CONCLUSION**

The fundamental goal of a sign language detecting system is to offer a workable method of hand gesture interaction between a normal and dumb person. By employing a webcam or other built-in camera that recognizes signs and analyses them for detection, the suggested system may be accessible. We may infer from the model's output that, given conditions of regulated light and intensity, the suggested system can produce reliable results. Also, adding new motions is simple, and the model will be more accurate if there are more photographs captured at various angles and frames. As a result, expanding the dataset makes it simple to scale up the model. The algorithm comes with different drawbacks, such as environmental conditions that reduce the accuracy of the detection, such as low light intensity and unmanaged backdrop. As a result, we will now seek to fix these issues and expand the dataset to get more exact results.

**6. REFERENCES**

[1] A. Thakur, P. Budhathoki, S. Upreti, S. Shrestha, and S. Shakya, “Real Time Sign Language Recognition and Speech Generation,” *J. Innov. Image Process.*, vol. 2, no. 2, pp. 65–76, Jun. 2020, doi: 10.36548/jiip.2020.2.001.

[2] M. J. J. Godwin, P. S, N. S, and S. S. R, “Support Vector Machine based Image Classification for Deaf and Mute People,” *Int. J. Adv. Eng. Res. Sci.*, vol. 5, no. 4, pp. 46–52, 2018, doi: 10.22161/ijaers.5.4.7.

[3] S. Pramada, “Intelligent Sign Language Recognition Using Image Processing,” *IOSR J. Eng.*, vol. 03, no. 02, pp. 45–51, 2013, doi: 10.9790/3021-03224551.

[4] R. Sruthi, B. V. Rao, P. Nagapravallika, G. Harikrishna, and K. N. Babu, “Vision Based Sign Language By Using Matlab,” pp. 3–6, 2018.

[5] S. Arora and A. Roy, “Recognition of sign language using image processing,” in *International Journal of Business Intelligence and Data Mining*, 2018, vol. 13, no. 1–3, pp. 163–176. doi: 10.1504/IJBIDM.2018.088428.

[6] “View of A Hand Gesture Recognition System for Deaf-Mute Individuals”.

[7] B. Tahir Khan under supervision of Amir Hassan Pathan, “Hand Gesture Recognition based on Digital Image Processing using MATLAB,” *Int. J. Sci. Eng. Res.*, vol. 6, no. 9, p. 338, 2015, [Online]. Available: http://www.ijser.org

[8] C. Dong, M. C. Leu, and Z. Yin, “American Sign Language alphabet recognition using Microsoft Kinect,” *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2015-Octob, no. April 2020, pp. 44–52, 2015, doi: 10.1109/CVPRW.2015.7301347.

[9] Manisha U. Kakde, Mahender G. Nakrani, and Amit M. Rawate, “A Review Paper on Sign Language Recognition System For Deaf And Dumb People using Image Processing,” *Int. J. Eng. Res.*, vol. V5, no. 03, pp. 590–592, 2016, doi: 10.17577/ijertv5is031036.

[10] K. Dixit and A. S. Jalal, “Automatic Indian Sign Language recognition system,” *Proc. 2013 3rd IEEE Int. Adv. Comput. Conf. IACC 2013*, no. February 2013, pp. 883–887, 2013, doi: 10.1109/IAdCC.2013.6514343.

[11] A. Y. Dawod, “Hand Gesture Recognition based on Isolated and Continuous Sign Language,” no. September, pp. 1–11, 2018.

[12] P. V.V and P. R. Kumar, “Segment, Track, Extract, Recognize and Convert Sign Language Videos to Voice/Text,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 3, no. 6, 2012, doi: 10.14569/ijacsa.2012.030608.

[13] P. V. V. Kishore, M. V.D. Prasad, C. R. Prasad, and R. Rahul, “4-Camera model for sign language recognition using elliptical fourier descriptors and ANN,” *Int. Conf. Signal Process. Commun. Eng. Syst. - Proc. SPACES 2015, Assoc. with IEEE*, no. April, pp. 34–38, 2015, doi: 10.1109/SPACES.2015.7058288.

[14] M. Kumar, “Conversion of Sign Language into Text,” *Int. J. Appl. Eng. Res.*, vol. 13, no. 9, pp. 7154–7161, 2018, [Online]. Available: http://www.ripublication.com

[15] P. Trigueiros, F. Ribeiro, and L. P. Reis, “Vision-based Portuguese Sign Language Recognition System.”

[16] Suharjito, R. Anderson, F. Wiryana, M. C. Ariesta, and G. P. Kusuma, “Sign Language Recognition Application Systems for Deaf-Mute People: A Review Based on Input-Process-Output,” in *Procedia Computer Science*, 2017, vol. 116, pp. 441–448. doi: 10.1016/j.procs.2017.10.028.

[17] A. Kumar Sahoo, G. Sankar Mishra, K. Kumar Ravulakollu, and A. K. Sahoo, “Sign language recognition: State of the art Ambient Intelligence View project Latent Fingerprint segmentation and detection View project SIGN LANGUAGE RECOGNITION: STATE OF THE ART,” vol. 9, no. 2, 2014, [Online]. Available: www.arpnjournals.com

[18] S. Sharma, V. P. Saxena, and K. Satish, “Comparative analysis on sign language recognition system,” *Int. J. Sci. Technol. Res.*, vol. 8, no. 8, pp. 981–990, 2019.

[19] Nithyakalyani K, S. Ramkumar, and K. Manikandan, “Design And Implementation Of Sign Language Translator Using Microtouch Sensor,” *Int. J. Sci. Technol. Res.*, vol. 9, p. 1, 2020, [Online]. Available: www.ijstr.org

[20] L. Pigou, S. Dieleman, P.-J. Kindermans, and B. Schrauwen, “Sign Language Recognition Using Convolutional Neural Networks BT - Computer Vision - ECCV 2014 Workshops,” pp. 572–578, 2015, [Online]. Available: https://core.ac.uk/download/pdf/55693048.pdf

[21] M. Zakariah, Y. A. Alotaibi, D. Koundal, Y. Guo, and M. Mamun Elahi, “Sign Language Recognition for Arabic Alphabets Using Transfer Learning Technique,” *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/4567989.

[22] R. H. Goudar and S. S. Kulloli, “A effective communication solution for the hearing impaired persons: A novel approach using gesture and sentence formation,” *Proc. 2017 Int. Conf. Smart Technol. Smart Nation, SmartTechCon 2017*, pp. 168–172, 2018, doi: 10.1109/SmartTechCon.2017.8358363.

**ANNEXURE 1**

The presentation for the research paper entitled “Sign Language Detection and Recognition A Review of State-of-the-Art Techniques and Application” has been **accepted** in the International Conference on Demystifying Emerging Trends in Green Technology (ICDETGT-2023).

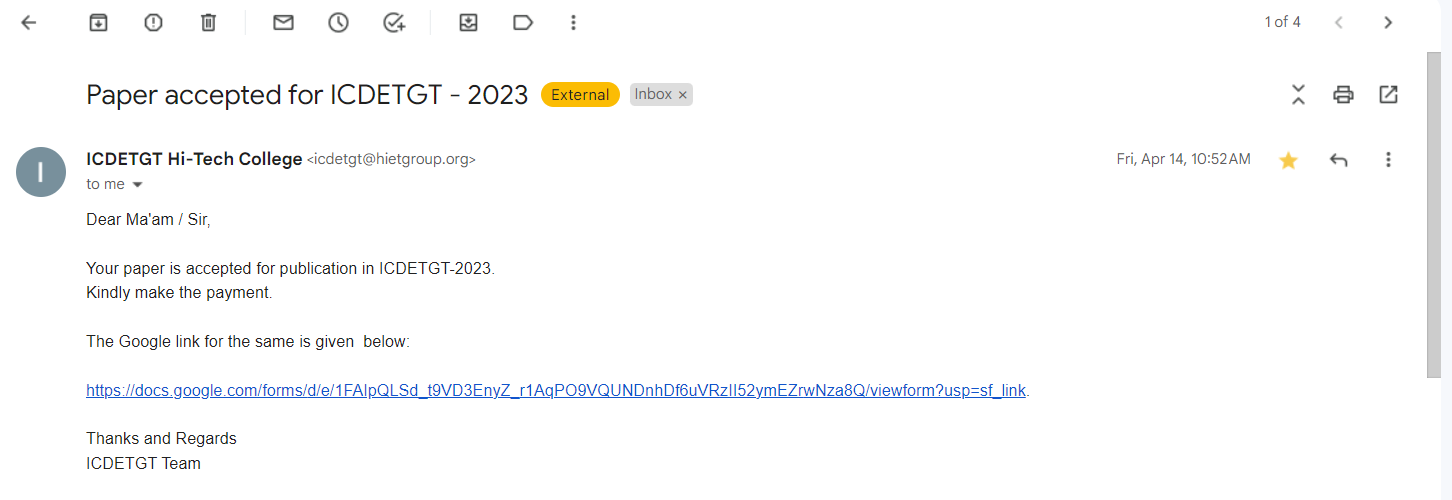
**Paper Title:**

Sign Language Detection and Recognition A Review of State-of-the-Art Techniques and Application.

**Abstract:**

Sharing or exchanging ideas, information, or emotions is referred to as communication. In order to establish communication between the parties, a common language must be understood by both. However, deaf and mute persons use distinct communication techniques. Being mute implies you cannot speak, just as being deaf means, you cannot hear. They use sign language to communicate with each other and with everyday individuals, but everyday individuals do not value the value of sign language. It can be challenging for a normal people to interaction with a deaf or mute people as not everyone is educated in or comprehends sign language. Machine learning can be used to build a model to get around this problem.

**Authors:** Kartik Tripathi, Md Zunad Alam, Shrusti Sanket, Ms. Gunjan Aggarwal

****

**ANNEXURE 2**

The presentation for the research paper entitled “Real Time Sign Language Detection for Deaf and Mute using OpenCV and CNN” has been **accepted** in the International Conference on Contemporary Engineering and Technology (ICCET 2023).

**Paper Title:**

Real Time Sign Language Detection for Deaf and Mute using OpenCV and CNN.

**Abstract:**

Around 5% of the world's population has hearing and speech impairments, according to the World Health Organization. The main form of communication for deaf and mute people is gesture language. This may make it difficult for them to communicate on a daily basis. To overcome this difficulty, The proposed system is designed to recognize American Sign Language and translate it into text. A picture of a hand writing the necessary alphabet is fed into the system as input. The histogram of the input photo is then calculated using the Bhattacharyya Distance Metric, and it is compared to the statistical features of previously saved images. The proposed system processes images using OpenCV as a tool. After checking the related letter of the images whose histogram resembles the input image's the most, the associated letter is printed. The system's adoption will be a minor step towards eradicating the social communication gap that exists between sign language non-users and deaf-mute users.

**Authors:** Kartik Tripathi, Md Zunad Alam, Shrusti Sanket, Ms. Gunjan Aggarwal

