iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IV **Month of publication:** April 2023

DOI: https://doi.org/10.22214/ijraset.2023.50527

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IV Apr 2023- Available at www.ijraset.com

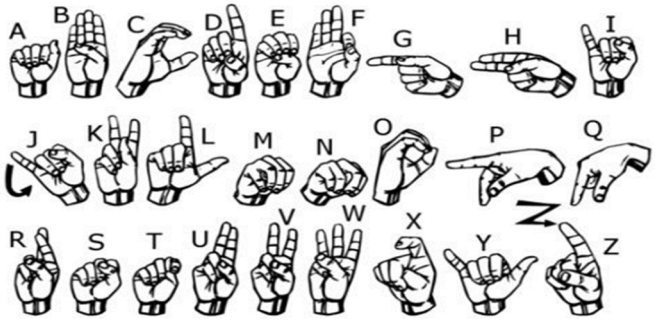
Sign Language Interpreter

Koyel Sarkar[1], Chetan Shivramwar[2], Diptanshu Mohabe[3]

Abstract: Sign language is one of the oldest and most natural form of language for communication, but since most people do not know sign language and interpreters are very difficult to come by, we have come up with a real time method using neural networks for fingerspelling based American sign language.
In our method, the hand is first passed through a filter and after the filter is applied the hand is passed through a classifier which predicts the class of the hand gestures. Our method provides 95.7 % accuracy for the 26 letters of the alphabet.

I. INTRODUCTION

In our project we primarily focus on producing a model which can recognize Fingerspelling based hand gestures in order to form a complete word by combining each gesture. The gestures we aim to train are as given in the image below.



American sign language is a predominant sign language Since the only disability Deaf and Dumb (hereby referred to as D&M) people have is communication related and since they cannot use spoken languages, the only way for them to communicate is through sign language.

Communication is the process of exchange of thoughts and messages in various ways such as speech, signals, behavior and visuals. D&M people make use of their hands to express different gestures to express their ideas with other people. Gestures are the non-verbally exchanged messages and these gestures are understood with vision. This nonverbal communication of deaf and dumb people is called sign language. A sign language is a language which uses gestures instead of sound to convey meaning combining hand-shapes, orientation and movement of the hands, arms or body, facial expressions and lip-patterns. Contrary to popular belief, sign language is not international. These vary from region to region.

A. Problem Statement

Dumb people use hand signs to communicate, hence normal people face problem in recognizing their language by signs made. Hence there is a need of the systems which recognizes the different signs and conveys the information to the normal people.

Thus, the system implemented has the following objectives :

1) *Objective 1:* Translate sign language to text/speech. The framework provides a helping-hand for speech-impaired to communicate with the rest of the world using sign language.
2) *Objective 2:* Elimination of the middle person who generally acts as a medium of translation and contain a user-friendly environment for the user by providing speech/text output for a sign gesture input.

II. SURVEY

As the project uses image processing for detecting and translation of sign language to text, it can have a wide variety of implementation. There are many more works that can be carried out as an extension of this project. This system predicts the need of the mute person but future systems may be developed that could communicate to the mute person's mobile device, allowing the system to learn the needs of the user, thereby provisioning the development of recommendatory systems as they have the relevant data related to the mute person that can easily be learned thought the neural network model.

©IJRASET: All Rights are Reserved

In vision-based methods, the computer webcam is the input device for observing the information of hands and/or fingers. The Vision Based methods require only a camera, thus realizing a natural interaction between humans and computers without the use of any extra devices, thereby reducing cost. These systems tend to complement biological vision by describing artificial vision systems that are implemented in software and/or hardware. The main challenge of vision-based hand detection ranges from coping with the large variability of the human hand's appearance due to a huge number of hand movements, to different skin-color possibilities as well as to the variations in viewpoints, scales, and speed of the camera capturing the scene.

A. *Advantages*

Sign language interpreters provide a valuable service for our community, with benefits that extend to both hearing and hearing-impaired individuals:

1) *Facilitating Meaningful Communication:* Being able to translate between spoken and sign language can connect two individuals that may not otherwise be able to effectively communicate. Because neither party has to struggle to make their needs, wants, and ideas are known, it opens the door to a more well-rounded form of communication.
2) *Enhance Accessibility:* Even though there are laws in place that are designed to support accessibility, there are still many ways in which the hearing world falls short. Hearing-impaired individuals face numerous obstacles on a daily basis, which can be understandably frustrating and even disheartening. An interpreter in virtually any setting supports the needs of a hearing-impaired individual in a way that can improve many life experiences.
3) *Support Education and Awareness:* For hearing individuals that don't have first-hand experience with a friend or family member with a hearing impairment, there can sometimes be a certain discomfort with the unknown. But with an interpreter available, many hearing people are able to realize that individuals with hearing impairments aren't really that "different" at all – they simply need some small accommodations at times.

B. *Disadvantage*

The biggest disadvantage is speed. Interpreted code runs slower than compiled code. This is because the interpreter has to analyse and convert each line of source code (or bytecode) into machine code before it can be executed.
Disadvantages of signs language in non-verbal communication: Vague and imprecise: Non-verbal communication 1. is quite vague and imprecise. Since in this communication, there is no use of words or language which expresses clear meaning to the receiver. No dictionary can accurately classify them.

III. METHODOLOGY AND EXPERIMENTATION

A. *Objectives*

The main objective is to translate sign language to text/speech. The framework provides a helping-hand for speech-impaired to communicate with the rest of the world using sign language.
This leads to the elimination of the middle person who generally acts as a medium of translation.
This would contain a user-friendly environment for the user by providing speech/text output for a sign gesture input.

B. *Description & Design*

American sign language is a predominant sign language Since the only disability D&M people have is communication related and they cannot use spoken languages hence the only way for them to communicate is through sign language.
Communication is the process of exchange of thoughts and messages in various ways such as speech, signals, behavior and visuals.Deaf and Mute(Dumb)(D&M) people make use of their hands to express different gestures to express their ideas with other people.
Gestures are the nonverbally exchanged messages and these gestures are understood with vision. This nonverbal communication of deaf and dumb people is called sign language.
In this project we basically focus on producing a model which can recognize Fingerspelling based hand gestures in order to form a complete word by combining each gesture.
Steps of building this project:
The first Step of building this project was of creating the folders for storing the training and testing data. As, in this project we have built our own dataset.
The second step, after the folder creation is of creating the training and testing dataset.

©IJRASET: All Rights are Reserved

We captured each frame shown by the webcam of our machine .
In each frame We defined a region of interest (ROI) which is denoted by a blue bounded square .
After capturing the image from the ROI, we applied gaussian blur filter to the image(to mute noises) which helps for extracting various features of the image.
After the creation of the training and testing data. The third step is of creating a model for training. Here, We have used Convolutional Neural Network(CNN) for building this model(CNN is a deep learning neural network sketched for processing structured arrays of data.)
The final step after the model has been trained is of creating a GUI that will be used to convert Signs into text and form sentence, which would be helpful for communicating with D&M people.

C. *Features*

1) Capture face images via webcam or external USB camera.
2) Hand gestures on an image must be detected.
3) The hand gestures must be detected in bounding boxes.
4) System should have a graphic card to process the images and videos on real time.
5) Reliable
6) Ease of use
7) Secure

IV. APPLICATIONS

A. *Effective Communication*

In accordance with the Americans with Disabilities Act (ADA), the District of Columbia must provide sign language interpreters upon request to individuals with disabilities when necessary to ensure effective communication. The DC government must provide sign language interpreters at its expense.

B. *Forms of Communication*

Sign language interpreters use their hands, fingers, and facial expression to translate spoken English into American Sign Language (ASL) and other signed languages. Interpreters may also serve clients who use transliterated Signed English (use of ASL signs, structured in English word order), Tactile interpreting (for individuals who are Deaf-Blind), Oral method (use of silent lip movements to repeat the spoken word so that the client can lip read), and Cued Speech modes of communication.

C. *Availability of Interpreter Services*

DC Government agencies can advise individuals of their willingness to provide interpreter services by stating that “Auxiliary aids and services will be provided upon request. Please contact [provide name and contact information (e-mail address is preferred)] at least 3 days in advance” and displaying an interpreting symbol on publications, notices, and flyers.

V. RESULTS AND DISCUSSION

We convert our input images (RGB) into grayscale and apply gaussian blur to remove unnecessary noise. We apply adaptive threshold to extract our hand from the background and resize our images to 128 x 128.
The prediction layer estimates how likely the image will fall under one of the classes. So, the output is normalized between 0 and 1 and such that the sum of each value in each class sums to 1. We have achieved this using SoftMax function.
At first the output of the prediction layer will be somewhat far from the actual value. To make it better we have trained the networks using labelled data.
The cross-entropy is a performance measurement used in the classification. It is a continuous function which is positive at values which is not same as labelled value and is zero exactly when it is equal to the labelled value. Therefore, we optimized the cross-entropy by minimizing it as close to zero. To do this in our network layer we adjust the weights of our neural networks. TensorFlow has an inbuilt function to calculate the cross entropy.
As we have found out the cross-entropy function, we have optimized it using Gradient Descent in fact with the best gradient descent optimizer is called Adam Optimizer.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

VI. CONCLUSION

In this report, a functional real time vision based American Sign Language recognition for D&M people have been developed for asl alphabets.

We achieved final accuracy of 95.0% on our data set. We have improved our prediction after implementing two layers of algorithms wherein we have verified and predicted symbols which are more similar to each other.

This gives us the ability to detect almost all the symbols provided that they are shown properly, there is no noise in the background and lighting is adequate

We achieved final accuracy of 90.0% on our data set. We have improved our prediction after implementing two layers of algorithms wherein we have verified and predicted symbols which are more similar to each other.

This gives us the ability to detect almost all the symbols provided that they are shown properly, there is no noise in the background and lighting is adequate.

VII. ACKNOWLEDGEMENT

We thank the almighty Lord for giving me the strength and courage to sail out through the tough and reach on shore safely.

There are number of people without whom this projects work would not have been feasible. Their high academic standards and personal integrity provided me with continuous guidance and support.

We owe a debt of sincere gratitude, deep sense of reverence and respect to our guide and mentor Ms. Ambrish Srivastav, Professor, AITR, Indore for his motivation, sagacious guidance, constant encouragement, vigilant supervision and valuable critical appreciation throughout this project work, which helped us to successfully complete the project on time.

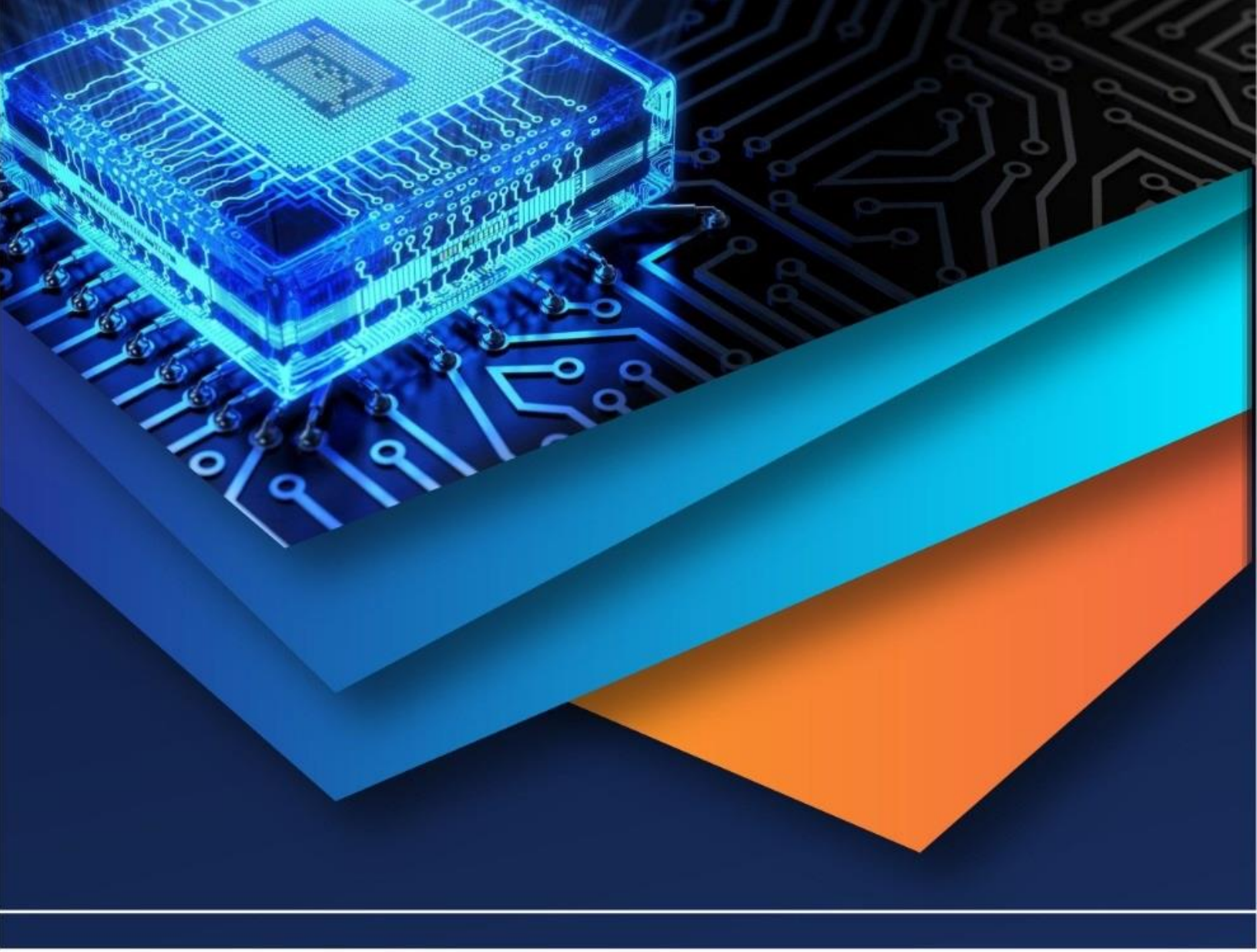
We express profound gratitude and heartfelt thanks to Dr Kamal Kumar Sethi, HOD CSE, AITR Indore for his support, suggestion and inspiration for carrying out this project. I am very much thankful to other faculty and staff members of CSE Dept, AITR Indore for providing me all support, help and advice during the project. We would be failing in our duty if do not acknowledge the support and guidance received from Dr S C Sharma, Director, AITR, Indore whenever needed. We take opportunity to convey my regards to the management of Acropolis Institute, Indore for extending academic and administrative support and providing me all necessary facilities for project to achieve our objectives.

We are grateful to our parent and family members who have always loved and supported us unconditionally. To all of them, we want to say “Thank you”, for being the best family that one could ever have and without whom none of this would have been possible.

REFERENCES

[1] T. Yang, Y. Xu, and “A., Hidden Markov Model for Gesture Recognition”, CMU-RI-TR-94 10, Robotics Institute, Carnegie Mellon Univ., Pittsburgh, PA, May 1994.
[2] Pujan Ziaie, Thomas M uller, Mary Ellen Foster, and Alois Knoll “A Na ̈ıve Bayes Munich, Dept. of Informatics VI, Robotics and Embedded Systems, Boltzmannstr. 3, DE-85748 Garching, Germany.
[3] https://docs.opencv.org/2.4/doc/tutorials/imgproc/gausian_median_blur_bilateral_filter/gausian_median_blur_bilateral_filter.html
[4] Mohammed Waleed Kalous, Machine recognition of Auslan signs using PowerGloves: Towards large-lexicon recognition of sign language.
[5] aeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural Networks-Part-2/
[6] http://www-i6.informatik.rwth-aachen.de/~dreuw/database.php
[7] Pigou L., Dieleman S., Kindermans PJ., Schrauwen B. (2015) Sign Language Recognition Using Convolutional Neural Networks. In: Agapito L., Bronstein M., Rother C. (eds) Computer Vision - ECCV 2014 Workshops. ECCV 2014. Lecture Notes in Computer Science, vol 8925. Springer, Cham
[8] Zaki, M.M., Shaheen, S.I.: Sign language recognition using a combination of new vision-based features. Pattern Recognition Letters 32(4), 572–577 (2011).
[9] N. Mukai, N. Harada and Y. Chang, "Japanese Fingerspelling Recognition Based on Classification Tree and Machine Learning," 2017 Nicograph International (NicoInt), Kyoto, Japan, 2017, pp. 19-24. doi:10.1109/NICOInt.2017.9
[10] Byeongkeun Kang, Subarna Tripathi, Truong Q. Nguyen” Real-time sign language fingerspelling recognition using convolutional neural networks from depth map” 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)
[11] Number System Recognition (https://github.com/chasinginfinity/number-sign-recognition)
[12] https://opencv.org/
[13] https://en.wikipedia.org/wiki/TensorFlow
[14] https://en.wikipedia.org/wiki/Convolutional_neural_nework
[15] http://hunspell.github.io/

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)