iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60828

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 12 Issue IV Apr 2024- Available at www.ijraset.com

Development of Dynamic Image Recognition System for Hand Sign Language into Audio and Visual Output using Artificial Neural Networks

Prof. M. C. Chavan[1], Prof. D.N. Chaudhari[2], Pratham J[3], Vinay M[4], Lalit C[5], Ashish T[6]
Department of Robotics and Automation Robotics and Automation Engineering, Zeal College of Engineering and Research

Abstract: ***This system provides a simple solution to a complex problem faced by deaf and dumb community. This paper visualizes the development web application of a real-time image recognition system whose main aim is to convert or translate the hand signs used in sign language in textual or visual and audio form output by using artificial neural networks. This system used computer vision to see signs as an input for processing and makes accurate predictions by using deep learning methods in real-time. By using artificial neural networks (ANN), the system achieves better performance in predicting or recognizing hand signs. The system makes communication between people using sign language and people who don't know sign language effective by providing output as an audio and texted / visual, which even helps people in both ways. The project aims only to provide the society a solution by helping people who use sign language to express themselves to make it more effective and give them an ability to speak by using the technologies.***

Keywords: Hand sign language, Dynamic image recognition, Artificial neural networks, Audio output, Visual output, Communication.

I. INTRODUCTION

For deaf or signers, hand sign language plays an important role by becoming a bridge of communication for them, it also provides them a way to express themselves by using hand signs. However, for non-signers or the people who don't understand sign language it is very difficult for them to understand. To provide solution to this problem, this study introduces with and web application that provides functionality of dynamic sign recognition system which is able to convert or translate the input hand signs of hand sign language into audio and visual outputs.

By specially using artificial neural networks for the system over convolutional neural networks, recurrent neural networks, for more accurate hand signs prediction.

The main aim of the system is to provide (Specially for deaf / signers) an accessibility of this web application to every individual which will enhance the communication and will give them an ability to express themselves more effectively and they can stand with everyone in the surrounding as an equal. There are about 14% of the worlds population who can speak or have problems while speaking and 14% of the worlds population is a large number so, we have provided a concept that can become a helping hand to the hand sign users by converting their hand signs in text and speech, which will make them feel like they are speaking or simply it gives them an ability to speak and express themselves more effectively.

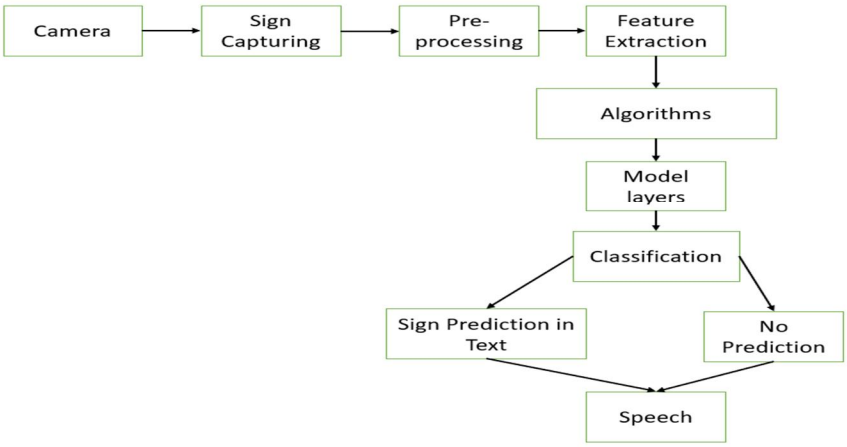
II. METHODOLOGY

The system uses various technologies that helped in building a efficient and effective model for sign language recognition. Following figure presents or explains the flowchart of the system working and how it predicts hand signs and also below there is the explanation of all modules and libraries that are used for building the model.

It explains how image is captured using feature extraction by media pipe library of python programming then it is compared to already trained datasets which helps the models predict the hand signs which has more similarities and less errors and the printed label or output is converted in audio output.

©IJRASET: All Rights are Reserved

A. Flow Chart:



Firstly the camera, OpenCV library in python is used due to which we were able to integrate with web cam which allows to system to capture performed hand signs of sign language. After capturing the image there are particular co-ordinated occurred on hand by using Media Pipe which helps system to predict hand sign by comparing the image points with already stored data set. There are about 6 layers of neural networks in the system which helps system predict the output more accurately. Once prediction is made the dedicated label is shown and the texted label is converted in the speech.

B. Modules Used for Implementation:

1) OpenCV: For Camera or web cam integration from which image is captured in the model.
2) Media Pipe: for feature extractions and for identifying hand co-ordinates to system for sign prediction.
3) TensorFlow: For creating model and its neural networks, in which there are nodes which acts as or can be said as mimic as a human brain in which there are neuron which transfer the data from one node to another it is also said as forward and backward propagation for the error solving and get the data with less error.
4) Different algorithms are used for the system to work efficiently and the algorithm is chosen on the bases of requirement.
5) web-speech API: This converts the predicted output in speech which enhances the communication for the people.

III. IMPLEMENTATION

A. Web Application Interface

The Web application provides the interactive user interface to its users for performing hand signs. The web application is made by keeping an aim in mind that it should be accessible to every individual who need our system. It has a try now button, after clicking it renders to new window and opens the camera on which users can perform the hand signs and the output can be taken as per the hand signs performed. The web application users interface can seen in fig. 1 and fig. 2.

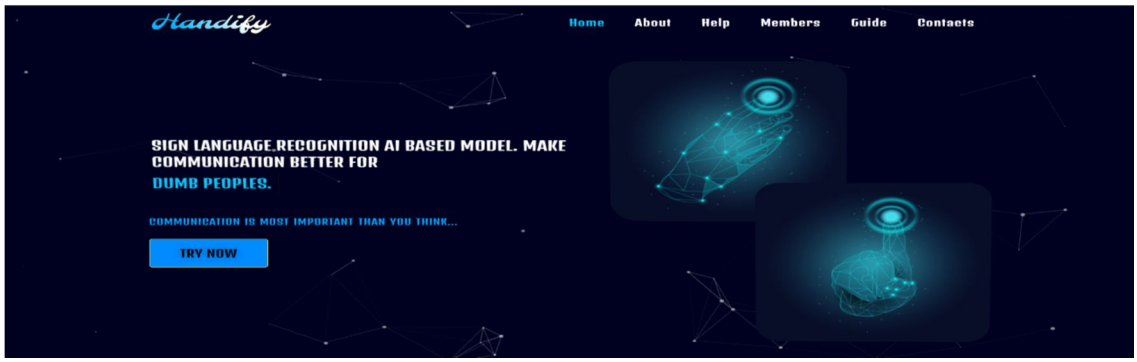


Fig. 1 Web Application – Home Page

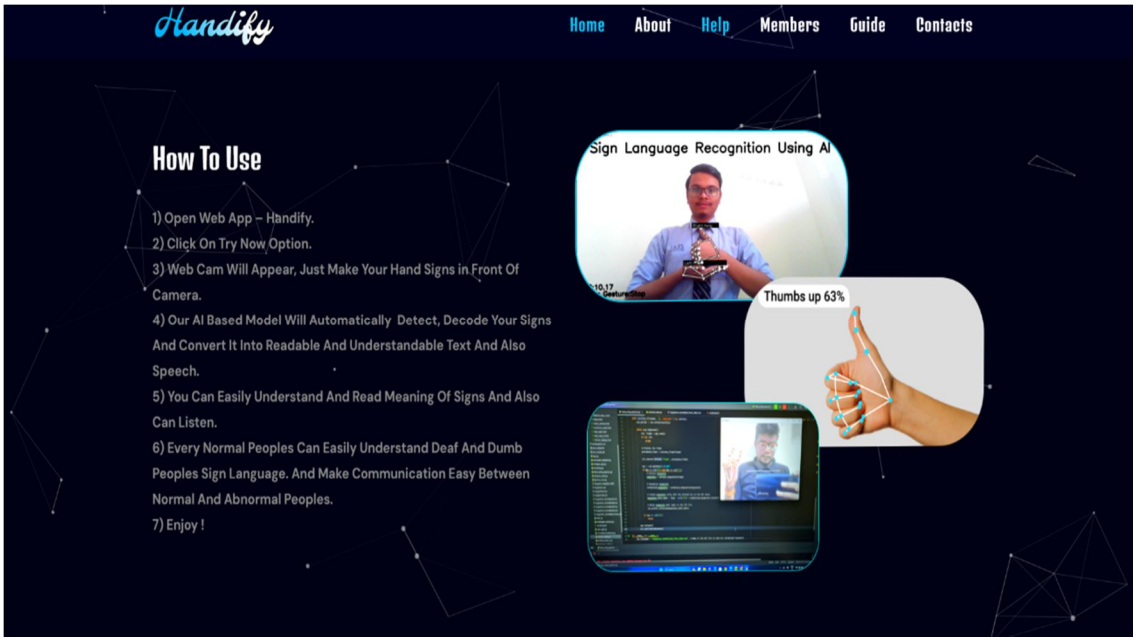


Fig. 2 Web Application – Help Page

B. Performance Result / Analysis

Following table represents the value of the parameters on which the system's performance is determined, the parameters are explained below.

Table.1 System Performance Result

	Precision	Recall	F1-score	Support
0	0.99	0.99	0.99	197
1	0.99	0.94	0.96	228
2	0.88	0.74	0.80	342
3	0.98	1.00	0.99	252
4	0093	1.00	0.96	394
5	0.62	0.96	0.75	241
6	0.93	1.00	0.96	230
7	0.98	0.97	0.98	263
8	1.00	0.84	0.91	210
9	0.92	0.95	0.94	259
10	1.00	1.00	1.00	314
11	0.99	1.00	1.00	754
12	0.98	0.72	0.83	218
13	1.00	0.89	0.92	254
14	0.99	0.59	0.74	405
15	0.99	0.80	0.89	270
16	1.00	1.00	1.00	616
17	1.00	0.98	0.99	205
18	0.82	1.00	0.90	227
19	0.61	0.95	0.75	307
Accuracy			0.92	6186
Macro accuracy	0.93	0.91	0.91	6186
Weighted accuracy	0.94	0.92	0.92	6186

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

Here are some explanation of the parameters:

1) accuracy: In our hand sign recognition system the models provides an accuracy of about 92% approximately in predicting any hand sign. It shows that almost 8 out of 10 are predicted almost correctly.

2) Precision: Precision measures the proportion of true positive predictions among all positive predictions. It helps evaluate the system's ability to avoid false positives. Precision is simple the study of measure of the number of true predictions in all of the positive predictions. Our system has precision value calculated which is between macro average 93% and weighted average 94%.

3) Recall (Sensitivity): Sensitivity is termed as Recall or the rate of getting true positive, it can be said as it is the mease of how much positive output are being identified. It is important in assessing the ability of the system to capture and predict all relevant hand signs. Our system has the recall of macro average is 91% and weighted average is 92% in the system for predicting hand signs.

4) F1 Score: It is simply said as it is the harmonic mean of the recall and precision. It also provides the balance measures in the false positive and false negatives. It is mainly useful when there is an imbalance in the number of negative instances and positive instances. Our system provides the F1 Score of macro average is 91% and weighted average is 92% which can be referred in the table 1.

5) Confusion matrix: In our system the fig.2 Confusion Matrix provides the diagrammatic presentation of the systems predictions. It helped in visualizing the performance of the model in predicting hand signs by the number of correct and incorrect predictions. Fig.2 confusion matrix can be seen for more understanding.

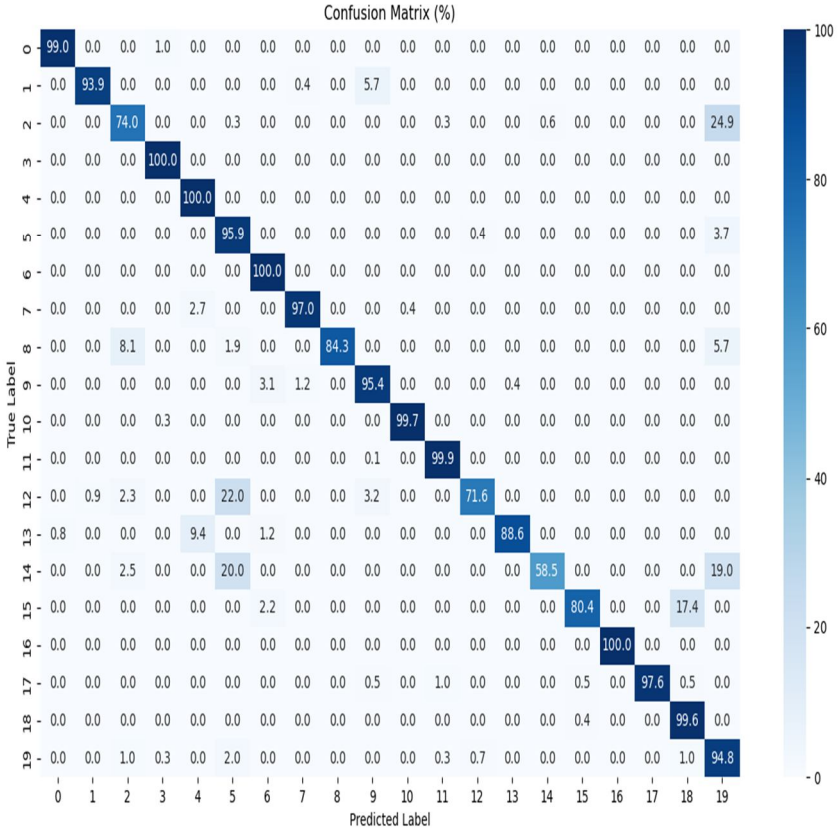


Fig.2 Confusion Matrix

C. Labels and Sentences

Following are the labels or can be said as hand signs that currently the system can predict accurately and there about approximately 20 to 24, sentences that can be created using hand signs that can be seen in the table 2 and table 3.

Table. 2 Sign Labels

1. Hii
2. How are you
3. Name
4. Need
5. Water
6. Food
7. Help
8. To go
9. I am
10. Hospital
11. What are
12. You
13. Doing
14. My home
15. Medicine
16. Emergency
17. Have
18. Sorry
19. Please
20. Bathroom
21. Fine
22. Contact
23. My father
24. My mother

Above Table.2 contains the labels or words on which whole model is trained, these are the hand signs which are trained to the system.

Sentences that could be formed using Table. 2 labels

1.	Hii
2.	How are you
3.	I am “name”
4.	I want to go hospital
5.	I need water
6.	I need food
7.	I want help
8.	Please help
9.	I want to go home
10.	What are you doing
11.	I need medicine
12.	I have emergency
13.	Sorry “name”
14.	I am fine
15.	I want to go to bathroom
16.	Please contact home
17.	Please contact my father
18.	Please contact my mother

Table. 3 Sentences made from hand sign labels

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

Above Table. 3 contains the example of few sentences that can be created using the labels of signs in Table. 2 which are given as an example to make sentences using the above sentences and labels.

D. Validation Standards

Validating a hand sign recognition system using AI involves assessing its performance and ensuring its accuracy, robustness, and reliability.

Here are some standard validation points:

1) Range: Hand sign recognition system hand a range for performing hand signs which lies between (15 cm to 1 m). Within this the system can accurately analyze and predict performed desired hand sign.

2) Number of hands: Hand sign recognition system can accurately recognize hand signs performed by one or two hands. Maximum **2** hand can be predicted accurately by the system.

3) Light area: Person performing hand signs should be performing the hand signs in the lighted area which will help the system to track hand signs more accurately. Dark areas should be avoided while performing hand signs.

4) Number of hand signs: Currently we have trained the model about 25 hand signs which can be utilized for making various sentences.

5) Web application: Hand sign recognition system using AI is deployed in a web application which is accessible to the user through internet.

By following these validation standards, you can ensure the reliability and effectiveness of an AI-based hand sign recognition system.

IV. RESULTS

In real time images are captures from the camera by the system thwn the captured images are given to algorithm or model for prediction or processing. The media pipe library of the python is used for accurate fearure extraction on hand co-ordinates. The captured co-ordinates are then compared with the already trained co-ordinates dataset for every label or word. Then the most compaired or similar co-ordinated label value will be predicted. The predicted sign will be visualized with the text and the text will be converted in to the speech.



Output: Hii / Hello



Output: call

V. FUTURE SCOPE

The Sign Language Recognition System using AI has laid a foundation for future advancements and applications. As technology continues to evolve, there are several exciting avenues for exploration and improvement in the field of sign language recognition.

1) Multi-Modal Integration: Future systems can explore the integration of multiple modalities, combining visual information with additional sensory inputs like facial expressions and contextual cues. This holistic approach can enhance the richness and accuracy of sign language interpretation.

2) Enhanced Gesture Vocabulary: Expanding the gesture vocabulary covered by the system can lead to more comprehensive communication. Research and development efforts can focus on capturing a broader range of signs, including regional variations and specialized signs in specific domains

3) Edge Computing and IoT Integration: Optimizing the system for edge computing and integration with Internet of Things (IoT) devices can enhance accessibility in various environments. Real-time, on-device processing can reduce latency and improve the system's applicability in different scenarios.

4) Cross-Cultural Adaptability: Adapting the system to recognize and accommodate different sign language variations worldwide enhances its cross-cultural applicability. Collaborative efforts can involve the creation of extensive datasets representing diverse signing styles.

5) Continuous Learning and Adaptation: Implementing mechanisms for continuous learning and adaptation will ensure the system stays relevant over time. This involves the ability to update the model with new signs, expressions, and linguistic nuances.

VI. CONCLUSIONS

We conclude, that the development of the dynamic sign recognition system for hand signs language which represents such development as an advancement in technology. By Utilizing ANN (Artificial Neural Networks), the system’s performance is enhanced for real-time sign recognition and prediction of the hand sign. The audio and visual output enhances the communication for every individual with problem, by this it provides a real-time interactive environment to the needy users. In future more works or research could be done on this, new signs could be added, user interface could be enhanced. For making people express themselves more effectively and communication would be enriched.

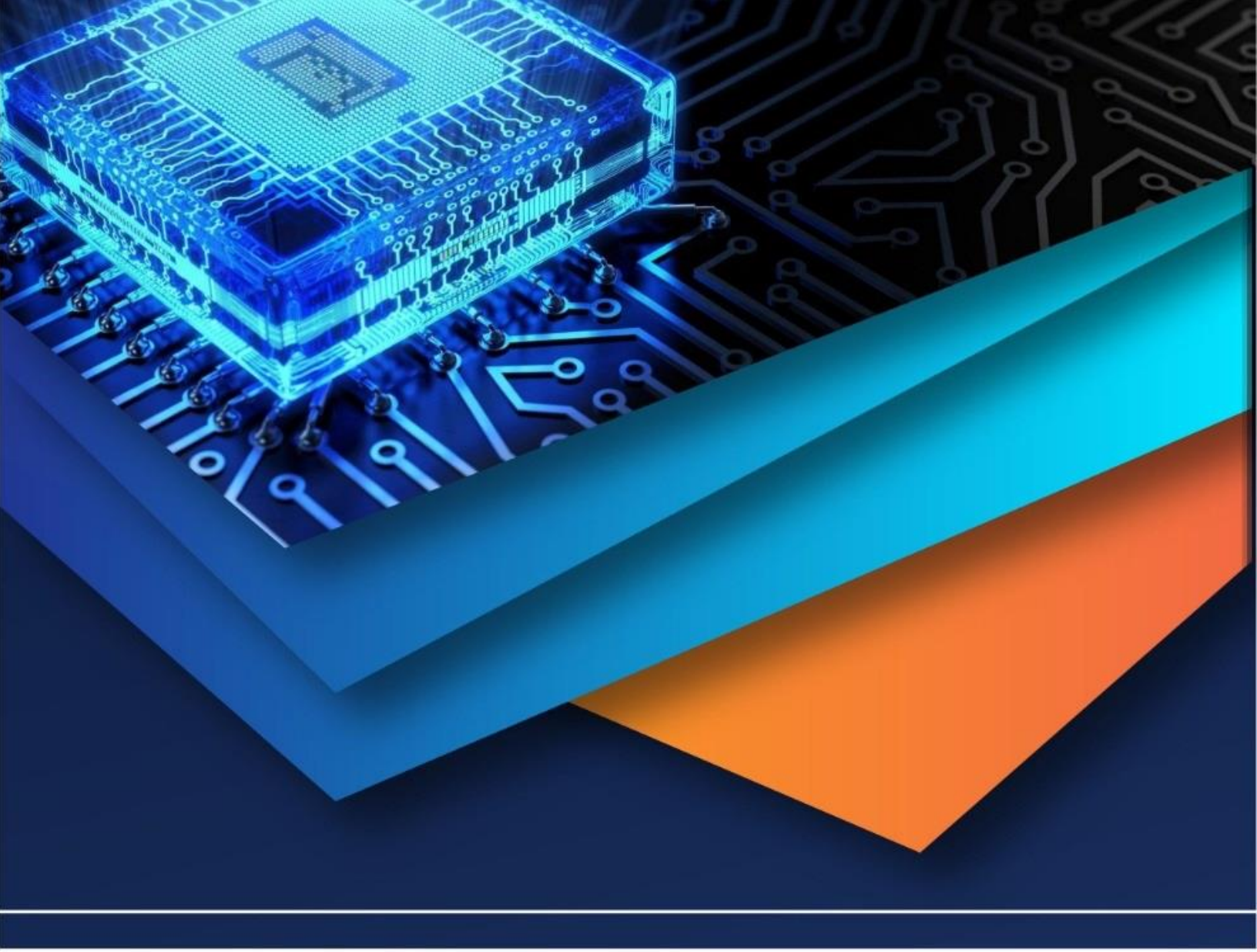
REFERENCES

[1] Ahmed B, Abdu H. G, Adel A, Muhammed A. (2023) has conducted an experiment on “Deep Learning in Sign Language Recognition: A Hybrid Approach for Recognition of Static and Dynamic Signs”, (Volume 11), (issue 17), Department of computer science.URL: https://doi.org/10.3390/math11173729

[2] Medhini P, Prasad H, Sai D, Shivam T. (2022) has conducted an experiment on “Sign Language Conversion to Text to Speech” (Volume 9), (Issue 7), Department of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bangalore, 560056. URL: https://www.jetir.org>papers

[3] Vilma V, E J Honesty P, G Maniknand, Ms. Hemalata S. (2024) has conducted an experiment on "Sign Language Detection and Recognition Using Media Pipe and Deep learning Algorithms" (Volume 11), (Issue 2), International Journal of Scientific Research in Science and Technology. URL: https://ijsrts.com/index.php/home/article/view/IJSRST5241123
[4] Jean M, Trong N, Huu H. (2013) has conducted an experiment on "Static Hand Gesture Recognition Using Artificial Neural Networks" (Volume 1), (Issue 1), Journal of Image and Graphics. URL: https://joig.1.1.34-38
[5] Pigou L, Dieleman S, Kindermans J, Schrauwen B. (2015) "Sign Language Recognition Using Convolutional Neural Networks". (Volume 8925) (issue 40) Lecture Notes in Computer Science,. Springer, Cham. URL: https://link.springer.com/chapter/10.007/978-3-319-16178-5_40
[6] Zaki M.M., Shaheen S. (2011) "Sign language recognition using a combination of new vision-based features" (Volume 32) (Issue 4).
Links:
[7] https://docs.opencv.org/2.4/doc/tutorials/imgproc/gausian_median_blur_bilateral_filter/gausian_median_blur_bilateral_filter.html
[8] http://www-i6.informatik.rwth-aachen.de/~dreuw/database.php
Git-hub Repository:
[9] Daeshpande3.github.io "A Beginner Guide to Understanding Convolutional Neural Networks"
Books:
[10] "Machine Learning for Sign Language Recognition" by A. Smith (2021): Explores machine learning techniques specifically applied to sign language recognition, covering both traditional methods and recent advancements.
[11] "Learn American Sign Language" by J. Guido (2015): Has in depth knowledge of the signs performed by signers for day-to-day communication, which helped us to study all the hand signs used were taught to the model.

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



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)