iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XII **Month of publication:** December 2023

DOI: https://doi.org/10.22214/ijraset.2023.57691

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 XII Dec 2023- Available at www.ijraset.com

ASL Classification using Deep learning

Ronit Bakshi[1], Satwik Pandey[2], Tanmay Parnami[3], Utkarsh Jain[4], Lokesh Kumar Meena[5]
Dr. Akhilesh Das Gupta Institute of Professional Studies

Abstract: ***In this study, the ever-evolving landscape of American Sign Language (ASL) classification is examined through the lens of cutting-edge deep learning techniques. The main objective of this research is to improve accuracy and overcome real-time constraints. For millions of individuals who are deaf or hard of hearing, sign language serves as a vital form of communication that bridges the gap between them and the rest of the world. Deep learning, specifically the use of convolutional neural networks (CNNs) and recurrent neural networks (RNNs), has shown incredible potential in numerous computer vision and natural language processing tasks. This paper offers a comprehensive exploration of the emerging field of sign language classification through deep learning, delving into the motivations behind this research.***

Keywords: ***American Sign Language, Natural Language Processing, Real-time Object Detection, Machine Learning, Convolutional Neural Network (CNN)***

I. INTRODUCTION

Sign language is a crucial means of communication for millions of individuals worldwide who are deaf or hard of hearing. It bridges the gap between them and the rest of the world, enabling them to convey thoughts, emotions, and ideas effectively. However, the accessibility of sign language relies heavily on the proficiency of interpreters and the availability of resources. In recent years, deep learning has emerged as a transformative technology with the potential to revolutionize the way we approach sign language recognition and classification. Deep learning techniques, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have demonstrated remarkable success in various computer vision and natural language processing tasks. By harnessing the power of deep learning, we can create robust and accurate systems for recognizing and classifying sign language gestures. This paper presents a comprehensive overview of the ever-evolving field of sign language classification using deep learning. We delve into the motivations behind this research, the challenges it addresses, and the potential impact it can have on the lives of individuals who use sign language. Additionally, we will explore the key components of deep learning models tailored for sign language classification. The significance of this work lies not only in its potential to improve accessibility for the deaf and hard of hearing community but also in its broader implications for the field of computer vision and deep learning. As we unravel the complexities of sign language and create models capable of understanding and classifying it, we take a significant step toward a more inclusive and connected world.

II. LITERATURE REVIEW

[1] Rakowski et al.; critically assessed the application of contemporary neural network architectures in the domain of hand shape recognition. The investigated models demonstrated notable enhancements, exhibiting an up to 10% improvement in accuracy, precision, recall, and F1-scores for the ASL Fingerspelling classification task. Furthermore, a 4% increase in accuracy was observed for hand shape classification on the 1 million Hands dataset. Notably, the utilization of pre-trained weights significantly improved generalization to unseen subjects in the ASL Fingerspelling task, while random weight initialization resulted in overfitting. Despite a narrower performance gap in hand shape classification, the Inception-ResNet-v2 architecture emerged as the most proficient, underscoring its efficacy in advancing the field of hand shape recognition.

[2] Tolentino et al.; introduced a system translating static sign language into corresponding word equivalents, encompassing letters, numbers, and basic signs to acquaint users with sign language fundamentals. Extensive testing ensured the system's significance for non-signers, resulting in notable usability and learning impact. Collaborating with sign language experts, the study sought to surpass prior deep learning accuracy standards. The system achieved remarkable training accuracy of 99%, testing accuracy averaging 93.67% across letter, number, and word recognition. Distinguishing itself from earlier studies, it expanded gesture recognition to include numbers, enhancing Sign Language Recognition's (SLR) utility. Despite comparable accuracy, the system outperforms by recognizing larger vocabularies without external aids like gloves or hand markings, differentiating it in the landscape of existing systems.

©IJRASET: All Rights are Reserved

[3] Kulhandjian, et al.; conducted a study focused on the detection and classification of ten fundamental American Sign Language (ASL) short phrase gestures. These gestures, encompassing expressions such as "Help me", "Call 911", "Danger", "Don't touch", "Do you need help?", "Call an Ambulance", "How are you?", "Nice to meet you", "Yes", and "No" were captured utilizing an X-band Doppler radar transceiver. Employing National Instruments (NI) data acquisition (DAQ) device and LabVIEW SignalExpress software, the gestures' spectrogram images were extracted. The study utilized deep convolutional neural network (DCNN) and VGG16 algorithms implemented in Matlab for training and classifying these spectrogram images. The findings highlighted the efficacy of the DCNN algorithm in accurately classifying ASL gestures based on spectrogram data, achieving an average validation accuracy of 87.5%. Furthermore, the VGG-16 algorithm demonstrated even higher accuracy, averaging at 95% in accurately categorizing these ASL gestures.

[4] Abdulhussein et al.; introduced a novel approach to American Sign Language (ASL) gesture recognition by employing Deep Learning techniques. Employing 227 X 227 RGB images resized through Bicubic interpolation, binary ASL hand edge images are derived via edge detection. This dataset, comprising 240 images representing 24 letters across 10 individuals, undergoes training with four Convolutional Neural Networks (CNNs). Notably, the achieved accuracy stands at an impressive 99.3%, with a minimal loss error of 0.0002. The training process proves efficient, characterized by a swift elapsed time. Comparative analysis highlights superior classification results when juxtaposed with related works. Emphasizing the proficiency in recognizing static ASL letters with akin hand configurations, the study underscores the effectiveness of deep learning in achieving high accuracy in static letter gesture recognition.

[5] Sakshi Sharma and colleagues introduced a system integrating hand detection and segmentation methods with Convolutional Neural Networks (CNN) for ISL alphabet recognition. The proposed algorithm demonstrates resilience to varying light conditions and effectively addresses occlusion challenges. Leveraging the CNN obviates the need for manual feature extraction, allowing automatic feature extraction and sign gesture classification. Real-time experiments affirm the system's high accuracy in recognizing static ISL alphabets. Noteworthy is the system's limitation in handling dynamic and continuous signs, focusing solely on static ISL signs. Future endeavors aim to extend the system for continuous sentence interpretation, incorporating non-manual signs alongside manual signs.

III. PROPOSED FRAMEWORK

1) *Python:* Python is a high-level, general-purpose programming language known for its readability and simplicity. It supports multiple programming paradigms and has a vast ecosystem of libraries and frameworks, making it widely used in various fields, including web development, data science, artificial intelligence, and more.
2) *MediaPipe:* MediaPipe is an open-source framework developed by Google that provides tools and components for building applications related to media processing, particularly in the areas of computer vision and machine learning. It offers pre-built solutions for tasks like face detection, hand tracking, pose estimation, and more.
3) *Tkinter:* Tkinter is the standard GUI (Graphical User Interface) toolkit for Python. It provides a set of tools and libraries for creating desktop applications with graphical interfaces. Tkinter is based on the Tk GUI toolkit and is included with most Python installations, making it accessible for developers to create interactive and user-friendly applications.
4) *OpenCV (Open Source Computer Vision Library):* OpenCV is an open-source computer vision and machine learning software library. It provides a wide range of tools and functions for image and video analysis, including features like image processing, object detection, facial recognition, and more. OpenCV is extensively used in computer vision applications and is compatible with both C++ and Python.

IV. METHODOLOGY

The methodology for developing a Sign Language Classification System using Deep Learning involves a series of well-structured steps. Below is an outline of the key phases and activities involved

A. *Problem Definition and Understanding*

1) Define the scope of the research, including target sign languages and gestures.
2) Identify potential applications and user requirements.
3) Understand the challenges in sign language classification, such as variations in sign gestures, signer-dependent gestures, and lighting conditions.

B. Model Selection and Design

1) The linear classification using the SGD (Stochastic Gradient Descent) classifier proved to be unsuccessful.

2) The support vector machine (SVM) approach exhibited low accuracy and faced challenges.

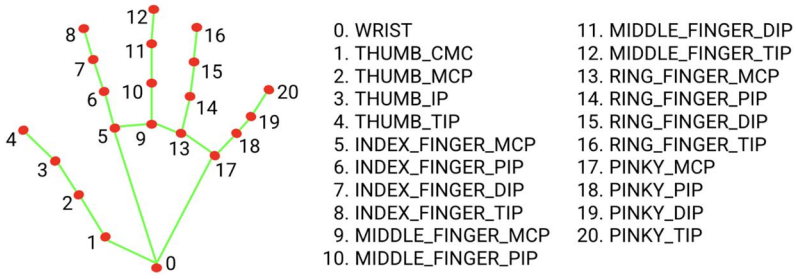
3) The Convolutional Neural Network (CNN) was deployed, utilizing the Keras module within TensorFlow. It achieved high accuracy on the dataset but encountered difficulties in real-time scenarios.

4) The MediaPipe model developed by Google was utilized for hand recognition, successfully identifying landmarks on the hand and subsequently classifying each sign based on specific alphabets in real-time.

C. Working

1) *Checking Existence of Hand Landmarks:* The specified land landmarks in the Mediapipe model serve as key reference points for accurate human pose estimation. These 20 points enable the model to analyse and understand the spatial configuration of a person's body, facilitating precise and reliable skeletal tracking in various applications.

Fig.1: Hand Landmarks



2) *Calculate Rotation Angle:* If hand landmarks exist, calculate the angle between the pinky finger and the wrist in a coordinate system with an inverted y-axis, using the formula:

$$tan\ (\theta) = \frac{y_{pinky\ MCP} - y_{wrist}}{x_{pinky\ MCP} - x_{wrist}}$$

If $\tan(\theta)$ is between -0.9 and 0.9, classify the hand as sideways; otherwise, classify it as straight.

3) *Finger Orientation:* When the hand is classified as sideways, the finger orientation typically aligns horizontally or perpendicular to the ground, extending sideways from the palm. This positioning denotes a lateral extension of the fingers, maintaining a parallel alignment with the hand's horizontal plane. If the hand is classified as sideways the finger's orientation is given as:

- Pinky Finger (PFU):

$$x_{Pinky\ Finger\ Pip} - x_{Pinky\ Finger\ Tip}$$

- Ring Finger (RFU):

$$x_{Ring\ Finger\ Pip} - x_{Ring\ Finger\ Tip}$$

- Middle Finger (MFU):

$$x_{Middle\ Finger\ Pip} - x_{Middle\ Finger\ Tip}$$

- Index Finger (IFU):

$$x_{Index\ Finger\ Pip} - x_{Index\ Finger\ Tip}$$

If the hand is classified as straight the finger's orientation is given as:

- Pinky Finger (PFU):

$$y_{Pinky\ Finger\ Pip} - y_{Pinky\ Finger\ Tip}$$

- Ring Finger (RFU):

$$y_{Ring\ Finger\ Pip} - y_{Ring\ Finger\ Tip}$$

- Middle Finger (MFU):

$$y_{Middle\ Finger\ Pip} - y_{Middle\ Finger\ Tip}$$

- Index Finger (IFU):

$$y_{Index\ Finger\ Pip} - y_{Index\ Finger\ Tip}$$

4) *Thumb Orientation:* When the hand is categorized as being in a straight position, the thumb typically aligns parallel to the other fingers, pointing upward or outward. This orientation denotes a lack of bending or curvature in the hand's overall structure, maintaining the thumb's position in line with the fingers.

- Thumb Up (TU):

$$x_{Thumb\ Tip} - x_{Thumb\ MCP}$$

The results of these checks are stored as Boolean values in the variables PFU, RFU, MFU, IFU, and TU.

5) *Classification:* In our meticulous classification of American Sign Language (ASL) symbols for each alphabet, we employed a thorough analysis that focused on discerning distinct features. The internal breakdown, as illustrated in the accompanying figure, highlights the intricate details considered during this manual categorization process. By scrutinizing and identifying key characteristics, we aimed to create a comprehensive and nuanced understanding of ASL symbols for each alphabet, ensuring accuracy and clarity in their representation. This methodical approach enhances communication and comprehension, contributing to a more effective utilization of ASL in diverse contexts.

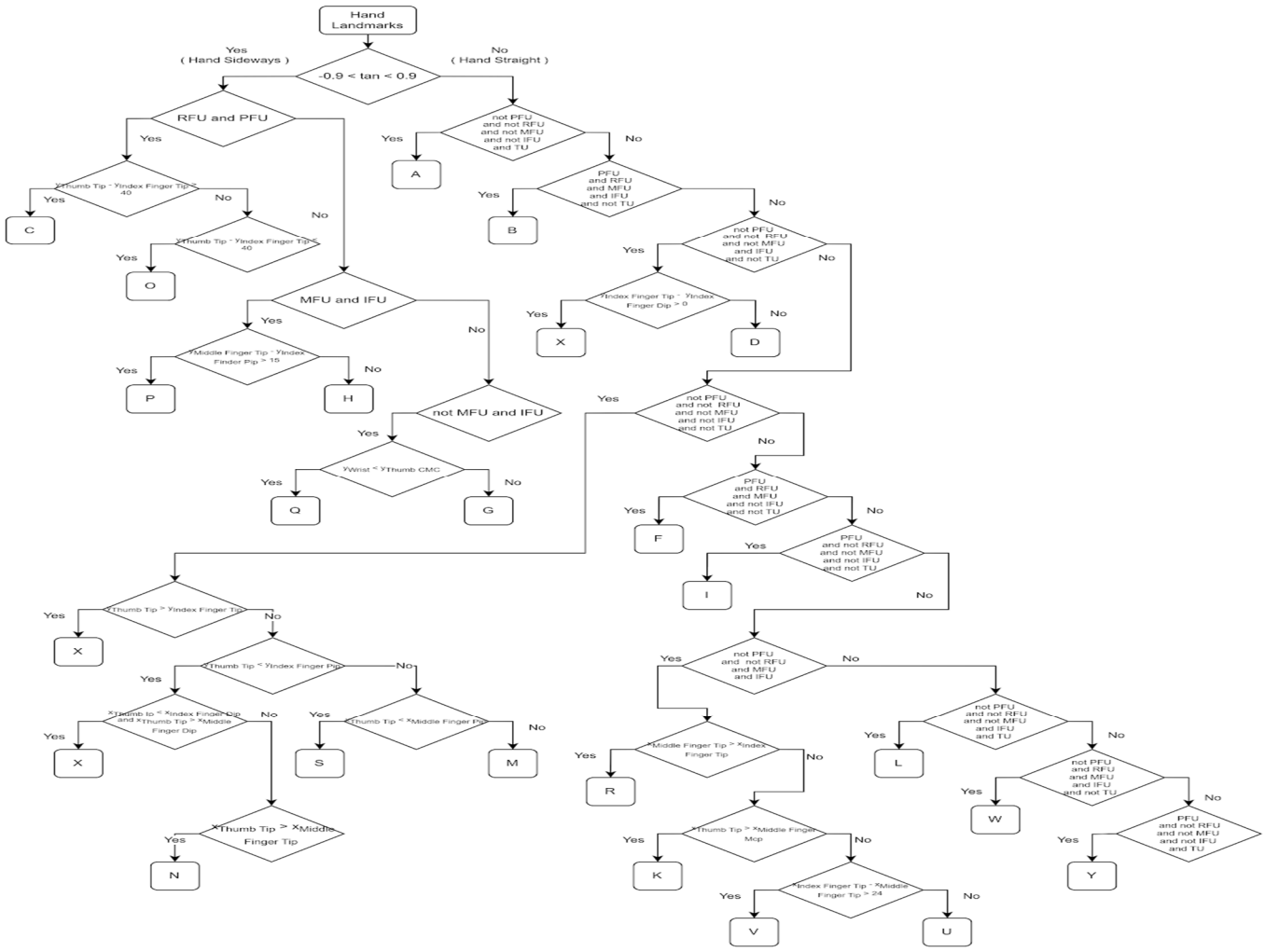


Fig.2: Classification Implementation

V. RESULT ANALYSIS

As we delved into the MNIST ASL dataset, we encountered limitations due to its low resolution of 28 x 28 pixels. It became apparent that accurately capturing the intricacies of American Sign Language gestures would be challenging with such constrained image quality. To achieve more nuanced and reliable results, we made the decision to supplement our dataset with real-time data. With our own dataset at hand, we could overcome the drawbacks of the MNIST ASL dataset and gather higher-quality images that more accurately portray the complexity of hand gestures. This use of real-time data greatly enhances the robustness and adaptability of our model, equipping it to handle a wider range of hand poses and variations in lighting conditions. Ultimately, the incorporation of real-time data heightens the effectiveness of our project.

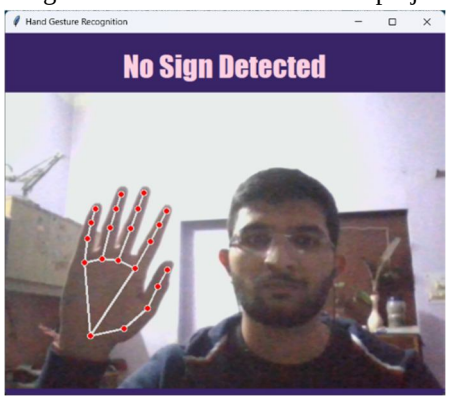


Fig.3: Model Output (No Symbol Detected)

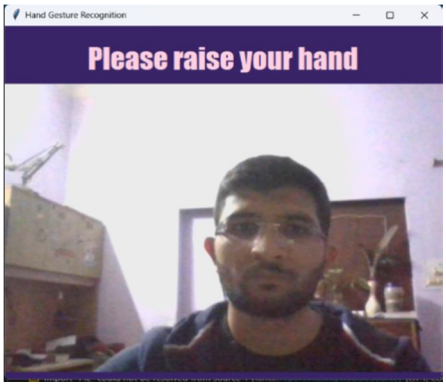


Fig.4: Model Output (No Hand Detected)

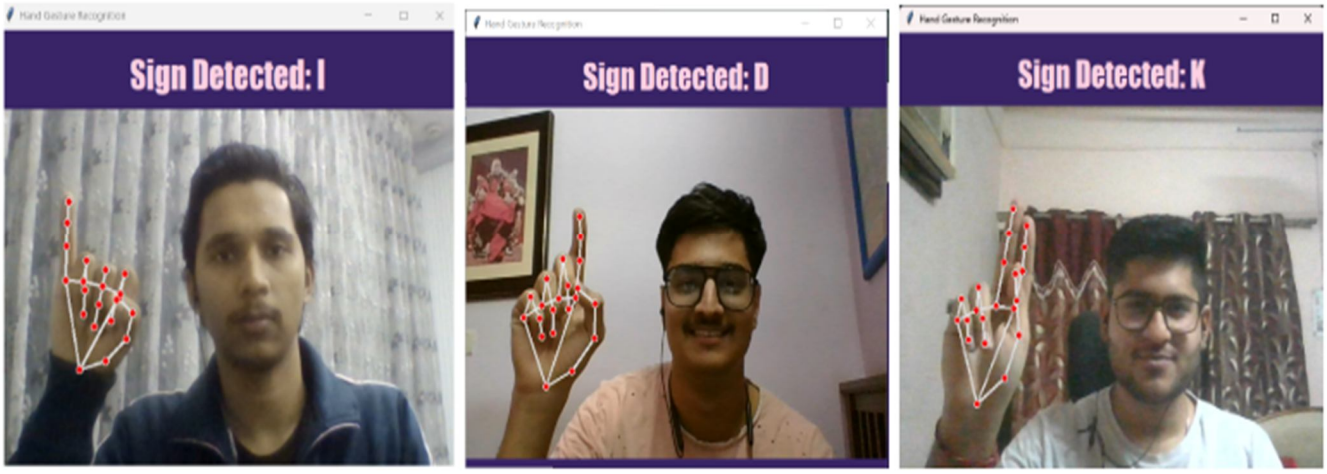


Fig.5: Model Output (Symbol Detected)

VI. APPLICATION

Sign language recognition applications have far-reaching implications across diverse fields, utilizing technology to enrich communication and accessibility for individuals who rely on sign language. In the realm of assistive technology, it fosters inclusivity through communication devices that facilitate interaction with both signers and non-signers, as well as wearable devices offering real-time translation. In education, interactive learning tools and online courses leverage sign language recognition to support deaf students in classroom participation and knowledge acquisition. The integration of sign language recognition in smartphones, tablets, virtual assistants, and healthcare settings enhances accessibility, enabling effective communication in various contexts. It plays a pivotal role in emergency services, workplace communication, and entertainment, offering captioning and subtitles in media content and enhancing interactive experiences in virtual and augmented reality applications. Social media platforms benefit from sign language recognition features, fostering inclusivity in content creation and interaction. Moreover, its incorporation into public services and government agencies ensures accessibility and effective communication with citizens using sign language. Sign language recognition technology also fuels research and development, driving innovation in assistive technologies and contributing to a more inclusive and accessible world for the deaf community across a myriad of domains.

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

VII. CONCLUSION

In concluding our sign language recognition research, our focus on creating an understanding and interpreting system for sign language gestures has achieved significant milestones, particularly in real-time recognition capabilities with low error rates. Throughout this endeavour, we successfully navigated and overcame various challenges, showcasing adaptability and resilience. The explored technologies and methodologies extend far beyond this research's confines, finding applications in assistive devices, education tools, workplace communication, and healthcare interactions. Rooted in a user-centric design, our recognition system incorporates feedback, aligning with the needs of the deaf community.

Beyond breaking communication barriers, our paper contributes ethically by emphasizing privacy, consent, and cultural sensitivity. As we conclude, our commitment to inclusivity remains steadfast, anticipating the continued evolution of sign language recognition technologies and their positive impact on the lives of those reliant on sign language for communication.

VIII. FUTURE SCOPE

The future of sign language recognition holds exciting possibilities as technology advances. Deep learning techniques like CNNs and RNNs are set to enhance accuracy, while multi-model recognition combining visual information with data from wearables or depth sensors promises a more comprehensive understanding of gestures. Improving real-time processing capabilities is crucial for applications like video conferencing, and research into transfer learning and domain adaptation techniques will create adaptable models for diverse signing styles. Integrating facial expression recognition with hand gesture recognition can enhance accuracy, and standardized benchmarks are essential for fair assessments. Creating extensive datasets covering various sign languages and contextual factors is vital. Utilizing edge computing for on-device processing enhances privacy, particularly for wearables and mobile applications. Lastly, integrating sign language recognition into robotic systems offers exciting possibilities for seamless human-robot interaction.

IX. ACKNOWLEDGEMENT

We thank Prof. Lokesh Kumar Meena* our project's mentor, Dr. Akhilesh Das Gupta Institute of Professional Studies. Whose leadership and support have served as the compass guiding us through the challenging terrain of this research. His valuable feedback and contribution remarkably enhanced our manuscript.

REFRENCES

[1] Rakowski, Alexander & Wandzik, Lukasz. (2018). Hand Shape Recognition Using Very Deep Convolutional Neural Networks. ICCCV '18: Proceedings of the 2018 International Conference on Control and Computer Vision. 8-12. 10.1145/3232651.3232657.

[2] Tolentino, Lean Karlo & Serfa Juan, Ronnie & Thio-ac, August & Pamahoy, Maria & Forteza, Joni & Garcia, Xavier. (2019). Static Sign Language Recognition Using Deep Learning. International Journal of Machine Learning and Computing. 9. 821-827. 10.18178/ijmlc.2019.9.6.879.

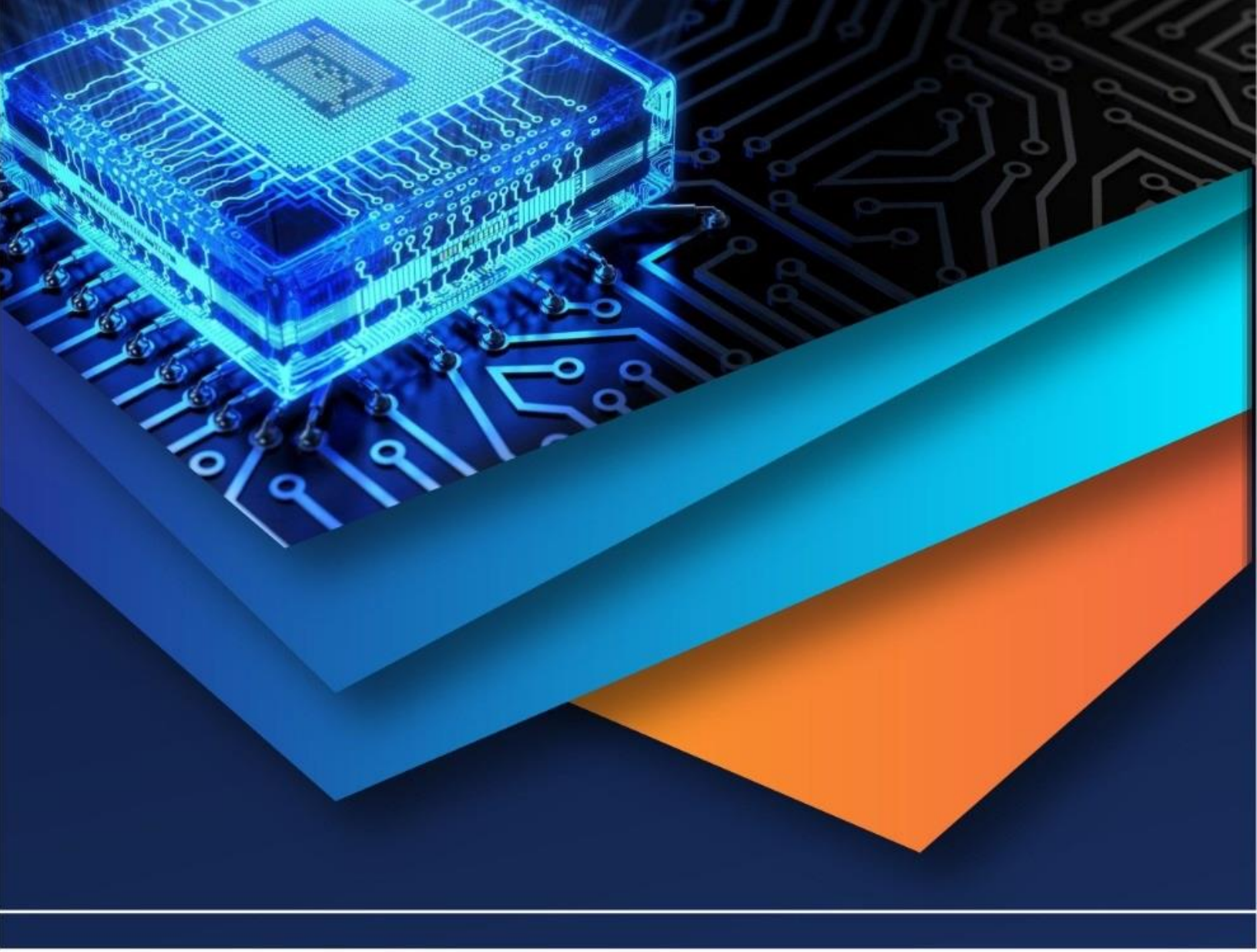
[3] Kulhandjian, Hovannes & Sharma, Prakshi & Kulhandjian, Michel & D'Amours, Claude. (2019). Sign Language Gesture Recognition Using Doppler Radar and Deep Learning. 1-6. 10.1109/GCWkshps45667.2019.9024607.

[4] A.Al, Abdulwahhab & Raheem, Firas. (2020). Hand Gesture Recognition of Static Letters American Sign Language (ASL) Using Deep Learning. 38, Part A. 926-937.

[5] Gangrade, Jayesh & Bharti, Jyoti. (2020). Vision-based Hand Gesture Recognition for Indian Sign Language Using Convolution Neural Network. IETE Journal of Research. 69. 1-10. 10.1080/03772063.2020.1838342.

[6] Kothadiya, Deep & Bhatt, Chintan & Sapariya, Krenil & Patel, Kevin & Gil, Ana & Corchado Rodríguez, Juan. (2022). Deepsign: Sign Language Detection and Recognition Using Deep Learning. Electronics. 11. 1780. 10.3390/electronics11111780.

[7] Bantupalli, Kshitij & Xie, Ying. (2018). American Sign Language Recognition using Deep Learning and Computer Vision. 4896-4899. 10.1109/BigData.2018.8622141.



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)