iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VIII **Month of publication:** August 2022

DOI: https://doi.org/10.22214/ijraset.2022.46470

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 10 Issue VIII August 2022- Available at www.ijraset.com

A Survey on Real Time Sign Language Detector using ML

Parag Gattani[1], Shreya Laddha[2], Sakshi Srivastava[3], Vikranth B M[4]
[1, 2, 3]*Computer Science Engineering B.M.S. College of Engineering Bengaluru, India*
[4]*Associate Professor, Dept. of CSE. B.M.S. College of Engineering Bengaluru, India*

Abstract: ***With the rapid growth in the field of object detection, developing a sign language detector as it is the main instrument of correspondence for physically challenged individuals has invoked the need for such a system. Communication via gestures is a shelter for the genuinely provoked individuals to communicate their contemplations and feelings. With the assistance of computer vision and neural organizations, we can recognize the signs and show the respective text as the output. This paper is a survey based on a broad collection of research papers on the related domain to propose an efficient and accurate model with the core functioning of a better response time and real-time detection using various machine-learning algorithms.***
Keywords: Image Processing, Machine-Learning, Neural Organization, Object Detection, Sign-Language Detector.

I. INTRODUCTION

Hand signals and gestures are used by those who are unable to speak. Ordinary people have trouble understanding their own language. As a result, a system that identifies various signs and gestures and relays information to ordinary people is required. It connects persons who are physically handicapped with others who are not.

Online communication has become the new normal standard for communication in the current circumstance, where we are all dealing with a global epidemic. It's difficult for disabled people to interact via video call or any other online method, and it's not always possible to have a third-party present to translate. The creation of a real-time sign language translator is a huge step forward in increasing communication between deaf people and the general public. The creation of a real-time sign language translator is a huge step forward in increasing communication between deaf people and the general public.

Sign language consists of motions made with the hands and other body parts, as well as facial expressions and body postures. There are a variety of sign languages, including British, Indian, and American. Users of American Sign Language (ASL) may struggle to understand British Sign Language (BSL), and vice versa. A working signing recognition system could allow the inattentive to interact with non-signing people without requiring the need of an interpreter. Our goal for this research is to create a system that can accurately classify signing.

Physically challenged people are frequently denied access to normal communication with their peers. It has been found that they find it difficult to connect with normal people with their gestures at times, as only a few of them are recognized by the majority of people. In such a community, sign language is the predominant mode of communication. It has syntax and vocabulary, just like any other language, but it communicates through visual means. The demand for a computer-based system among the disabled community is tremendous in this day and age of technology.

II. BACKGROUND

Researchers have been working on this subject for a long time, and the findings are promising. Object recognition technology is advancing at a rapid pace. The goal is to teach computers to recognize objects and create user-friendly computer interfaces (HCI). Some steps toward this goal include teaching a computer to recognize speech, facial emotions, and human gestures. Nonverbally communicated information is referred to as gestures. At any given time, a human can make an infinite number of gestures.

Computer vision researchers are particularly interested in human gestures since they are received through vision. The goal of the project is to create an HCI that can detect human motions. The conversion of these motions into machine language necessitates the use of a complicated programming procedure. For better output creation, they are focused on Image Processing and Template Matching in the project. It is understood that the previous works surveyed by researchers are primarily on image pre-processing and models centric only to few languages with no proper real-life implementation. They aren't focused on the development of new languages and with faster deployment models.

©IJRASET: All Rights are Reserved

Based on this, the current paper will address the previous researches with their inputs and outputs using various machine learning algorithms with their intended result and try to bridge the gap between running the model and deploying the same.

III. RESEARCH AIMS AND APPROACH

In the Literature survey we have gone through other similar works that are implemented in the domain of sign language recognition. Our main objective is to survey the various machine learning algorithms as well as suitable tools to implement the detection model. Main research aims were as follows:

1) The current models and algorithms available and researched by researchers.
2) Data acquisition, sign-language representation, data environment and image processing.
3) Performances of existing sign-language detection systems with its efficiency and output accuracy.

A. Search Methodology

The search criteria included in identifying papers that focused on the following keywords.

1) Sign language detection,
2) Dynamic hand gestures,
3) Real-time object detection.

Our search covered popular databases such as:

a) Science Direct,
b) IEEE Xplore Digital Library,
c) Springer Link,
d) Google Scholar.

To filter our search, we applied the following criteria:

- Publication date: Between 2010 and 2021 inclusive.
- Search domain: science, technology, or computer science.
- Publication types: journals, researches.
- Language: English.

IV. FINDINGS OF THE REVIEW

A. The Current Models And Algorithms Available And Researched By Researchers

Researchers have their own solidarity contrast with different strategies and specialists are as yet involving various techniques in fostering their own Sign Language Recognition. Every strategy too has its own restrictions contrasted with different techniques. The point of this paper is to audit the communication through signing acknowledgment approaches and observe the best technique that has been utilized by analysts. Subsequently, different analysts can get more data about the techniques utilized and could foster better communication through signing Application Systems later on.[1]

Table 1. Classification Method from Researchers

Methods used
HMM
Convolutional Neural Network
SOFM, SRN, HMM
SVM, and HMM
Kohonen Self Organizing Map
Simple Support Vector Machine
Wavelet Family Method
Accelerometer and surface electromyography
Multilayer Perceptron
Naïve Bayes Classifier

The vast majority of the articles center around three basic parts of the vision-based hand signal acknowledgment framework, in particular: information securing, information climate, and hand motion portrayal. It has likewise inspected the exhibition of the vision-based hand signal acknowledgment framework as far as acknowledgment precision. For the recognition systems, the acknowledgment exactness goes from 69% to 98%, with a normal of 88.8% among the chosen studies.[2]

A technique utilizing the HOG and SVM calculations with the Kinect programming libraries to perceive communication through signing by perceiving the hand position, hand shape, and hand activity highlights is proposed. To understand this technique, a unique 3D sign language dataset containing 72 words is gathered with Kinect, and analyzes are directed to assess the strategy. It is displayed in the exploratory outcomes that the utilization of the HOG and SVM calculations altogether builds the acknowledgment exactness of the Kinect, and is uncaring toward the foundation and different elements.[4]

Based on human pose estimation they extracted optical flow features and used a linear classifier which depicted an accuracy of 80% when assessed with Public DGS Corpus (German Sign Language). Used a Temporal Model (Uni-directional LSTM).Also a unidirectional LSTM with one layer and 64 hidden units, on the input, which is normalized for frame rate, and gives out the output to a 2-Dimensional array.[6]

An architecture using neural networks identification and tracking to translate the sign to text format. The Point of Interest (POI) and the track point introduction provided novelty and reduced storage memory requirement. [7]

Using a DataGlove, a huge thesaurus sign language evaluator is introduced with real time continuous gesture recognition of sign language. In a spurt of gesture input, the end-point detection and statistical analysis are handled in accordance with 4 frameworks: posture, position, orientation, and motion. [8]

The multi-layered random forest (MLRF) –It is a classification and regression technique that has become popular due to its efficiency and simplicity.[9]

$$t(\boldsymbol{x}) = \frac{1}{T}\sum_{i} t_i(\boldsymbol{x}),$$

Fig1 . The prediction of the whole RF t for the $\boldsymbol{x}$ is usually formed by averaging the predictions of the separate trees

$$P_{ij} = P_{ji} = \frac{|\{m|\boldsymbol{x}_i, \boldsymbol{x}_j \in L_k^m\}|}{T}.$$

Fig2. Proximity matrix P of size $N \times N$, where N is the size of training set.

A very efficient first stage of a self-governed fingerspelling recognition system using CNN from depth maps.[11]. CNN(Convolutional Neural Network) structure is used for feature extraction and classifier, and to construct the real time system, the hand locating process was applied. Skin color detection and convex hull algorithms have been used together in determining hand position. [12]

B. Data Acquisition, Sign-Language Representation, Data Environment And Image Processing

The camera is the most common equipment utilised as an input procedure in Sign Language Recognition (SLR). The method employs a skin filter to extract the skin region, after which the image is transformed into HSV colour space for each frame. MATLAB was also used to capture the image, which was then saved in a directory. The data was collected by another system using a Leap Motion Controller (LMC). [1]

For recording photos or video of hand motions with a video camera, there are four types of vision-based approaches:

1) A single camera, such as a video camera, digital camera, Webcam, or smartphone camera, is used at a time.
2) Active techniques, such as the Microsoft Kinect camera and the Leap Motion Controller, use light projection to find and detect hand movement.
3) Body markers such as wristbands or coloured gloves are used in invasive procedures.
4) A stereo camera captures images using multiple monocular cameras at the same time to offer depth information.[2]

To capture photos with the help of a Webcam, camera interfacing is required. The embedded camera can identify hand movements and position by capturing gestures. Capturing 30 frames per second will suffice to process photos; but, more input images will increase processing time, making the system slow and insecure. [3]

The Hand Position and Hand Action Feature Extraction Methods, as well as the Hand Shape Feature Extraction Method [4]
Orthogonal and non-orthogonal moment characteristics are used. Moments have the ability to reflect the image shape's global features. Moments based on discrete orthogonal polynomials like the Tchebichef and Krawtchouk polynomials, as well as a non-orthogonal moment called the Geometric moment, were used as features in this study. These are defined directly in image coordinate space and do not require any numerical approximation, such as continuous moments.[5]
To gauge the movement vector of every pixel on the human body as it moves from one frame to the next. Using full-body human position estimation as a proxy for such data, ignoring a set of points discovered in every video frame that denotes informative landmarks like joints and other moving parts (mouth, eyes, eyebrows, and others).[6]
A camera can be used to collect the video sequence of the signer, or the person communicating in sign language. In this project, we assume that the camera is facing the signer in order to capture the front view of the signer's hand gestures. Manually starting the acquisition is the only option. The image acquisition block of the ASLR system design is seen in Figure 1. A camera sensor is required to capture the signer's features and motions. [7]

a) Hidden Markov models for posture, orientation, and motion are used to extract the following aspects of a hand gesture.
b) For posture detection, DataGlove measures the flexion of 10 finger joints on one hand.
c) For orientation recognition, a Polhemus 3D tracker's azimuth, elevation, and palm roll are used.
d) The normalised motion trajectory is partitioned into ten linked vectors. The cosine of each pair of neighbouring vectors, the number of turning points in the trajectory, and the relative orientation between the start and end points are all factors to consider.[8]

Image with depth The hand is recognised and segmented by thresholding depth values, and a point cloud is produced for further processing. Descriptor of the ESF The ESF descriptor is made up of a series of histograms that have been concatenated together. The first histogram depicts the distances between two randomly selected points in the point cloud (function D2). The distribution of angles contained by two lines formed by randomly picking three points from a point cloud is described by the second histogram (function A3) The distribution of regions contained by three randomly sampled points is depicted in the third histogram (function D3). [9]
In ISL, an average of 1450 photos per digit were collected for the numbers 0 to 9. Approximately 300 pictures per letter in ISL were collected, except 'h', 'j', and 'v'. About 500 photos were acquired for each of the 9 gesture-related intermediate hand postures, such as Thumbs Up and Sun Up. There are a total of 24,624 photos in the dataset. The sign demonstrator wears a full-sleeved shirt in all of the photos. The majority of these photos were taken with a regular webcam, although a handful were taken with a smartphone camera. The resolutions of the photographs vary. 15 gesture films were taken for each of the 12 one-handed pre-selected gestures outlined in the study to train HMMs (After, All The Best, Apple, Good Afternoon, Good Morning, Good Night, I Am Sorry, Leader, Please Give Me Your Pen, Strike, That is Good, Towards). To make the HMMs more robust, these films feature modest variations in sequences of hand positions and hand motion. The sign demonstrator is wearing a full sleeve shirt in these films, which were taken with a smartphone camera. [10]
We divided the datasets into training and validation by volunteers since there was little to no variation across the photos for the same class of each signer. The remaining volunteer from each dataset was used to validate, and four of the five volunteers from each dataset were used to train. We decided against separating a test set because doing so would force us to remove one of the four hands from the training set, which would have a major impact on generalizability. Instead, we used the web application to evaluate the classifier by logging in and looking at the classification probabilities generated by the models. [11]
31,000 depth maps were collected using a depth sensor and a Creative Senz3D camera with a resolution of 320 240. Each of the 31 possible hand signs has 1,000 photos from five different participants in the dataset. Except for J and Z, which require temporal information for classification, 31 hand signals encompass all finger spellings of both alphabets and integers. We only have one class to represent both one letter and one number because (2/V) and (6/W) are differentiated based on context. We follow formal signals to avoid uncertainty between signers, even when some informal signs are clearer and easier to detect. The dataset is collected while subjects move their hands about on the image plane and along the z-axis to capture data from multiple views. [12]
It's the process of extracting a picture from a source, usually a hardware-based source, for image processing. In our project, WebCamera is the hardware-based source. Because no processing can be done without a picture, it is the first stage in the workflow sequence. The image that is produced has not been altered in any manner. [13]
Data collecting is an indisputable component of this study, as our outcome is greatly dependent on it. As a result, we developed our own ASL dataset with 2000 photos of 10 static alphabet signs. A, B, C, D, K, N, O, T, and Y are the ten classes of static alphabets. Two distinct signers created two different datasets.

In different lighting situations, each of them has made one alphabetical motion 200 times. The alphabetic sign motions dataset folder is further divided into two folders, one for training and the other for testing. 1600 photographs are utilised for training and the remaining images are used for testing. To ensure consistency, we took images with a webcam in the same background each time a command was issued. The resulting images are saved in the png format. It should be noted that when a png image is opened, closed, and saved again, there is no loss of quality. PNG is also capable of handling images with high contrast and detail. The images captured by the webcam will be in the RGB colour space. [14]

Extraction of hands bound convex hull points is one of the steps. The hand image is then classified using a convolutional neural network. When there are similar hand signs, a choice will be made based on the results of those phases. The image is shrunk to (28 28) pixels after hand extraction, and the colour space is changed to grey. The image is shrunk to (28 28) pixels after hand extraction, and the colour space is changed to grey. [15]

C. Performances Of Existing Sign-Language Detection Systems With Its Efficiency And Output Accuracy

From the HMM method, there are three kinds of implementations: light HMM, multi stream HMM and tied-density HMM.Light-HMM method has an accuracy of 83.6%. While Multi-Stream HMM achieved 99.4% for training dataset and 98.9% for testing dataset recognition rate of overall sub-word segment detection, 99% for training dataset and 96.5% for testing dataset recognition rate of overall sub-word detection, and 96.9% for training dataset and 86.7% for testing dataset recognition rate of overall sentence detection. TDHMM acquired an accuracy of 91.3%. SVM and HMM algorithm achieved 85.14% for demonstrating a Taiwanese sign language recognition.3D-CNN has an accuracy of 94.2%, while 3D-CNN has the accuracy of 78.8% and has an accuracy of 83% in a contest dataset. SOFM, SRN, and HMM combined method has achieved an accuracy of 91.3%. While the Kohonen SOM has 80% accuracy. 98.09% accuracy is achieved by using SimpSVM method and 97.5% accuracy is achieved by using SVM. 97% accuracy is achieved by using Eigen value weighted Euclidean distance method. Meanwhile, 95.1% accuracy is achieved by EFD and ANN and 100% accuracy is achieved by the wavelet family in simple recognition.[1]For the signer dependent, the recognition accuracy ranges from 69% to 98%, with an average of 88.8% among the selected studies. On the other hand, the signer independent's recognition accuracy reported in the selected studies ranges from 48% to 97%, with an average recognition accuracy of 78.2%. The lack in the progress of continuous gesture recognition could indicate that more work is needed towards a practical vision-based gesture recognition system.[2]Firstly capture the live stream images with the help of a webcam.Applying image preprocessing steps to remove unwanted noise and adjust the brightness of images is the next step. Then perform image analysis and apply a convexity algorithm to extract contour of hand position. Using a 2.40 GHz Intel® coreTM processor windows based Opens image processing software and IDE used to analyze 640*480 capture image size at frame rate 30 frames per second.[3]The normal recognition rate is up to 89.8%, which implies the Kinect-based recognition strategy proposed in this paper can adequately and productively perceive sign language, and it has extraordinary importance to the exploration and advancement of gesture-based communication acknowledgment innovation.[4] Krawtchouk moment is the best in terms of all the performance metrics. The recognition performances in terms of sensitivity using Krawtchouk moment, Tchebichef moment, and Geometric moment are 91.53%, 82.67%, and 76.20% respectively.[5]Based on human pose estimation they extracted optical flow features and, used a linear classifier which depicted an accuracy of 80% when assessed with Public DGS Corpus (German Sign Language).Using linear classifier with a fixed number of context frames achieves an accuracy between 79.9% to 84.3% on the test set.[6]All the letters have signs except "J" and "Z", which have gestures, which require additional motion vectors to identify them. The algorithm was able to identify all the alphabets with a 100% recognition rate. In the case of noise corruption, the algorithm resulted in an accuracy of 48%. The advantage of using neural network architecture is high processing speed.[7] The model uses Hidden Markov Model (HMM) with 51 fundamental postures, 6 orientations, and 8 motion primitives. A stretch of gestures can be recognized in real-time with an average recognition rate – 80.4%.The recognition rate achieved is 94.8%. The recognition rate for short sentences observed was 75.4%, whereas for long sentences is 84.7%.[8]Achieved an accuracy of 85% and in the h-h experiment, the achieved accuracy was 97.8%. However, the observed training time is 4000s per tree on a quad-core.[9].The postures are distinguished using the KNN algorithm. The system then achieved a precision of 99.7% for static hand poses and an accuracy of 97.23% for gesture recognition.[10] A subset of depth data has been used to train CNNs for the arrangement of 31 alphabets and numbers. American Sign Language is implemented in this system.99.99% accuracy has been achieved by the system.[11]The CNN was trained using a dataset acquired by a university and has achieved 100% test accuracy. The Real-Time System has a 98.05 percent accuracy. Only letters are employed in this study.[12]The hand gesture has been detected with the HSV color algorithm and set the background to black. The photos are divided after going through a variety of processing processes that involve various computer vision techniques. To classify and train the model CNN has been implemented.

The model has achieved 90% accuracy.[13]A method for using deep convolutional networks in American Sign Language has been implemented to label images of alphabets and digits.It was observed that 82.5% accuracy was on the alphabet gestures and 97% validation set accuracy on digits.[14]CNN is used to train the model and identify the pictures. The model has achieved accuracy about 95%.[15]

V. PROPOSED METHODOLOGY

A. System Architecture

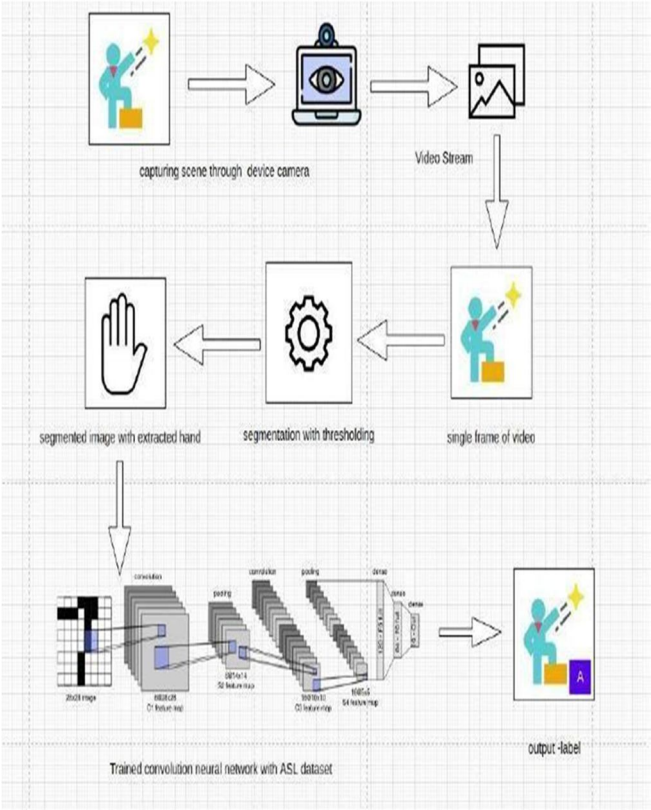


Fig 3: Architecture of Sign Language recognition System

B. Abstract Specification Of Subsystem

Table 4.1: Verification of test case

		Input description	Expected output	Test status
1	Loading model	Initializing trained model and load it	Loaded model without errors	pass
2	Converting video to frames	Capturing video and converting it into frames	Image frames of captured video stream	pass
3	Recognize hand gesture	Image frame that contains hand object	label	Pass

C. Use Case Diagram

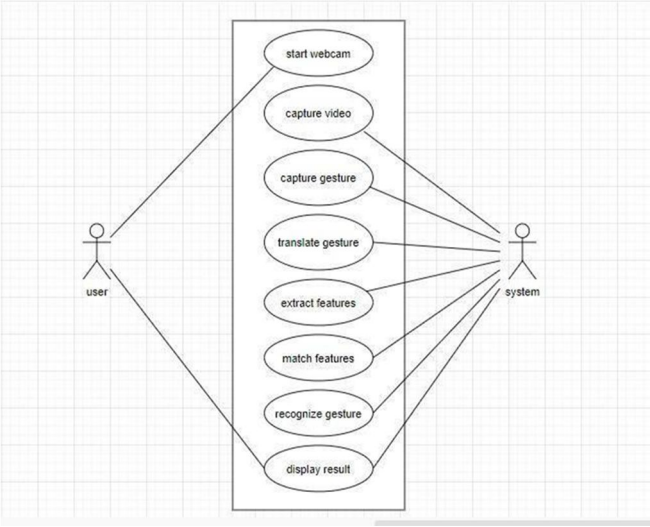


Fig 4: Use case diagram of sign language recognition System

Table 2: Use Case Scenario for sign language recognition system

Use case name	Sign language recognition
Participating actors	User, System
Flow of events	Start the system(u) Capturing video(s) Capture gesture(s) Translate gesture(s) Extract features(s) Match features(s) Recognizing gesture(s) Display Result
Entry condition	Run the code
Exit condition	Displaying the label
Quality requirements	Cam pixels clarity, good light condition

D. Methodology

Supervised machine learning: It is one of the ways of machine learning where the model is trained by input data and expected output data. To create such model, it is necessary to go through the following phases:

1) *Model construction:* It depends on machine learning algorithms. In our project, we will be using Faster R- Convolutional Neural Networks.
2) *Model Training:* The model is trained using training data, where the dataset will be divided into 2 sets: The Training dataset and the Testing dataset respectively.
3) *Model Testing:* The second set of data is loaded. This data set gives true accuracy and will be verified.

E. Preprocessing

1) *Uniform Aspect Ratio:* Understanding aspect ratios: Aspect ratio refers to the proportional connection between the width and height of a picture. It essentially describes the shape of a picture. Aspect ratios are expressed as a width-to-height formula, such as this: A square image, for example, has a 1:1 aspect ratio since the height and breadth are equal. The image may be 500px x 500px or 1500px x 1500px and still have a 1:1 aspect ratio. A portrait-style image, for example, might have a 2:3 ratio. The height is 1.5 times longer than the width in this aspect ratio. As a result, the image could be 500px by 750px, 1500px by 2250px, and so on.
2) *Cropping To An Aspect Ratio:* Aside from using built-in site style options, you may want to manually crop an image to a certain aspect ratio. For example, if you use product images that have the same aspect ratio, they'll all crop the same way on your site.

Image scaling:

a) Image scaling is the process of resizing a digital image in computer graphics and digital imaging. Upscaling or resolution improvement is a term used in video technology to describe the amplification of digital content.
b) When scaling a vector graphic image, geometric transformations can be used to scale the graphic primitives that make up the image without sacrificing image quality. A new image with a larger or lower number of pixels must be created when scaling a raster graphics image. When the pixel number is reduced (scaling down), the quality of the image is frequently compromised. Scaling raster graphics is a two-dimensional example of sample-rate conversion, which is the conversion of a discrete signal from one sampling rate (in this case, the local sampling rate) to another in the context of digital signal processing.

VI. FUTURE WORK

The suggested sign language recognition system can be further expanded to recognize gestures and facial expressions in addition to sign language letters. Instead of letter labels, sentences as a more accurate translation of language will be displayed. This improves the readability of the document as well. It is possible to broaden the scope of several sign languages. To improve letter detection accuracy, more training data can be added. This concept could be expanded to include the conversion of signs into speech.

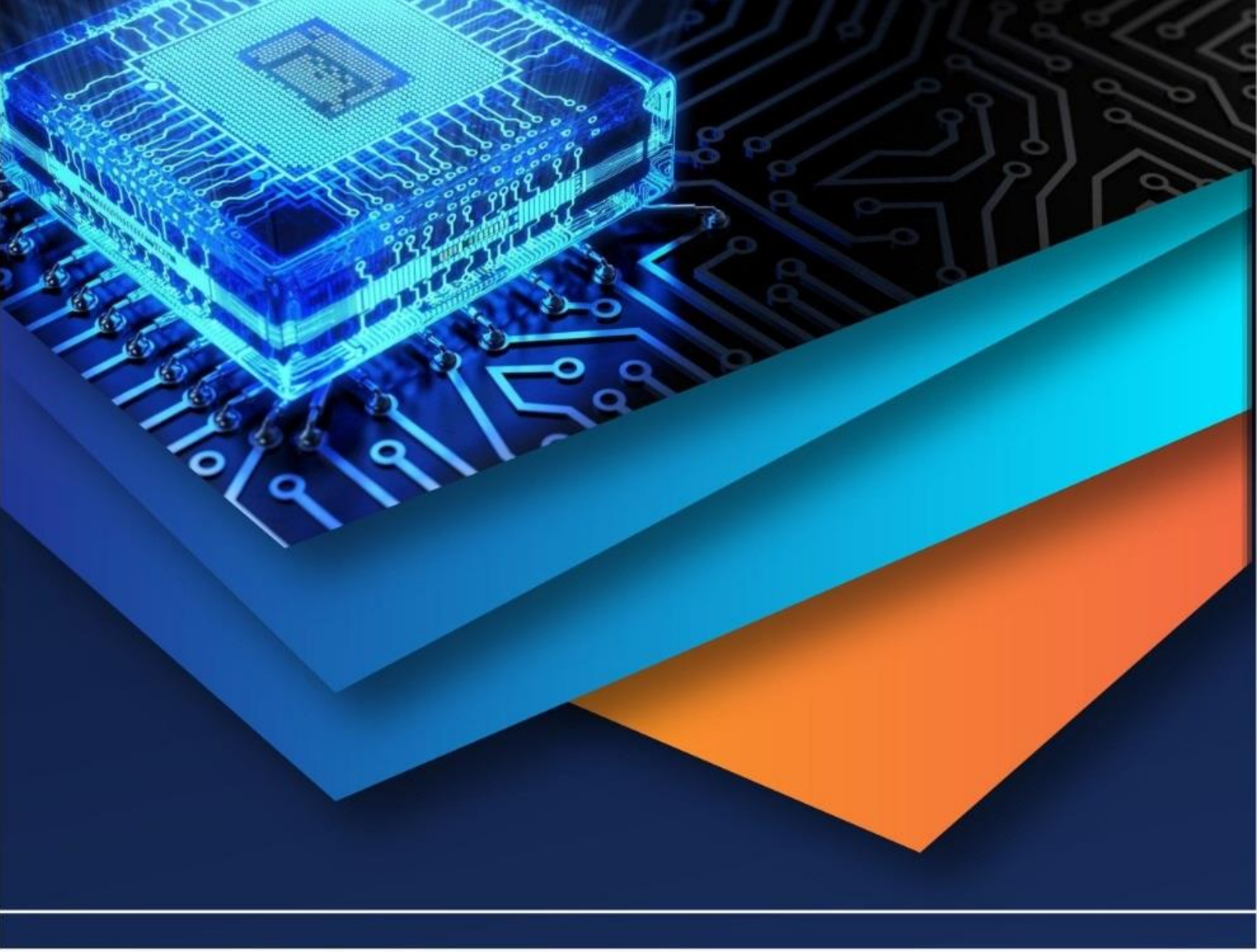
VII. CONCLUSION

Today's applications require a variety of image types as sources of data for explanation and analysis. Several characteristics must be retrieved in order to conduct various tasks. Degradation occurs when a picture is converted from one form to another, such as when digitizing, scanning, sharing, storing, and so on. As a result, the resulting image must go through an image enhancement process, which consists of a collection of approaches aimed at improving an image's visual presence. Image enhancement improves the interpretability or awareness of information in images for human listeners while also giving superior input for other automatic image processing systems. The image is then subjected to feature extraction using a variety of approaches in order to make it more computer-readable. A sign language recognition system is a useful tool for preparing an expert's knowledge, detecting edges, and combining erroneous data from several sources.

REFERENCES

[1] Anderson, Ricky, et al. "Sign language recognition application systems for deaf-mute people: a review based on input-process-output." Procedia computer science 116 (2017): 441-448.
[2] N. Mohamed, M. B. Mustafa and N. Jomhari, "A Review of the Hand Gesture Recognition System: Current Progress and Future Directions," in IEEE Access, vol. 9, pp. 157422-157436, 2021, doi: 10.1109/ACCESS.2021.3129650.
[3] A. S. Nikam and A. G. Ambekar, "Sign language recognition using image-based hand gesture recognition techniques," 2016 Online International Conference on Green Engineering and Technologies (IC-GET), 2016, pp. 1-5, doi: 10.1109/GET.2016.7916786.
[4] A. S. Nikam and A. G. Ambekar, "Sign language recognition using image-based hand gesture recognition techniques," 2016 Online International Conference on Green Engineering and Technologies (IC-GET), 2016, pp. 1-5, doi: 10.1109/GET.2016.7916786.
[5] Chatterjee, Subhamoy, Dipak Kumar Ghosh, and Samit Ari. "Static hand gesture recognition based on fusion of moments." Intelligent Computing, Communication and Devices. Springer, New Delhi, 2015. 429-434.
[6] Moryossef, Amit, et al. "Real-time sign language detection using human pose estimation." European Conference on Computer Vision. Springer, Cham, 2020.
[7] P. Mekala, Y. Gao, J. Fan and A. Davari, "Real-time sign language recognition based on neural network architecture," 2011 IEEE 43rd Southeastern Symposium on System Theory, 2011, pp. 195-199, doi: 10.1109/SSST.2011.5753805.
[8] Rung-Huei Liang and Ming Ouhyoung, "A real-time continuous gesture recognition system for sign language," Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, 1998, pp. 558-567, doi: 10.1109/AFGR.1998.671007.
[9] A. Kuznetsova, L. Leal-Taixé and B. Rosenhahn, "Real-Time Sign Language Recognition Using a Consumer Depth Camera," 2013 IEEE International Conference on Computer Vision Workshops, 2013, pp. 83-90, doi: 10.1109/ICCVW.2013.18.

[10] K. Shenoy, T. Dastane, V. Rao and D. Vyavaharkar, "Real-time Indian Sign Language (ISL) Recognition," 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2018, pp. 1-9, doi: 10.1109/ICCCNT.2018.8493808.
[11] M. Taskiran, M. Killioglu and N. Kahraman, "A Real-Time System for Recognition of American Sign Language by using Deep Learning," 2018 41st International Conference on Telecommunications and Signal Processing (TSP), 2018, pp. 1-5, doi: 10.1109/TSP.2018.8441304.
[12] B. Kang, S. Tripathi and T. Q. Nguyen, "Real-time sign language fingerspelling recognition using convolutional neural networks from depth map," 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), 2015, pp. 136-140, doi: 10.1109/ACPR.2015.7486481.
[13] Indian Sign Language Recognition using Convolutional Neural Network Rachana Patil, Vivek Patil, Abhishek Bahuguna, Gaurav Datkhile ITM Web Conf. 40 03004 (2021)DOI: 10.1051/itmconf/20214003004
[14] Pigou, Lionel, et al. "Sign language recognition using convolutional neural networks." European Conference on Computer Vision. Springer, Cham, 2014.
[15] Garcia, Brandon, and Sigberto Alarcon Viesca. "Real-time American sign language recognition with convolutional neural networks." Convolutional Neural Networks for Visual Recognition 2 (2016): 225-232.



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)