iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** I **Month of publication:** January 2023

DOI: https://doi.org/10.22214/ijraset.2023.48557

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 I Jan 2023- Available at www.ijraset.com

Hand Gesture Recognition

Swetha Kotavenuka[1], Harshitha Kodakandla[2], Nimmakayala Sai Krishna[3], Dr. S P V Subba Rao[4]
[1, 2, 3]*JNTUH Affiliated, Department of Electronics and Communication Engineering, Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India*
[4]*Project Guide, Professor and Head of Department, Department of Electronics and Communication Engineering, Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India*

Abstract: ***This work presents a computer-vision-based application for recognizing hand gestures. A live video feed is captured by a camera, and a still image is extracted from that feed with the aid of an interface. At least once per count hand gesture (one, two, three, four, and five), the system is trained. After that, the system is given a test gesture to see if it can identify it. Several algorithms that are capable of distinguishing a hand gesture were studied. It was determined that the highest rate of accuracy was achieved by using the computational neural network known as the Alexnet algorithm. Traditionally, systems have used data gloves or markers as a means of input. We are free to use the system however we like. In this way, the user can make natural hand gestures in front of the camera. The system implemented serves as an extendable basis for future work toward a fully robust hand gesture recognition system, which is still the subject of intensive research and development.***
Keywords: HAND GESTURE, MATLAB, Alexnet.

I. INTRODUCTION

Users have benefited from recent advancements in computer software technology as well as related hardware technology, which provides a value-added service. In the course of normal life, gestures of the body serve as an effective mode of communication. They are able to convey a wealth of information and emotions in a concise manner. Take, for instance: It is possible to communicate anything from "happy goodbye" to "caution" by swaying one's hand from side to side. Another aspect that is lacking in the majority of human-computer dialogues is the utilization of the full potential of physical gesture. Hand gesture recognition is one of the fundamental difficulties in computer vision. Modern advances in computing and communication have made it possible to build fully automated systems for human interaction. Computers with this capability can detect hands, identify them, and follow their movements.

Natural Image contains many technical and digital specifics that can be used in various areas of computer vision. Visual text detection and recognition have recently become patentable due to their widespread use in areas such as content-based image searching, automatic number plate recognition, data extraction from documents like passports, business cards, and bank statements, and the translation of handwriting into real-time control of computers. However, text must be reliably detected from any image of a natural scene, despite the presence of an unpredictable background, variations in font style, size, colour, and orientation, and geometric and photometric distortion. It is the first step of any hand processing system to detect and localise a hand in an image. However, hand detection proved challenging due to the wide range of possible hand poses, orientations, locations, and sizes. The varying degrees of illumination also contribute to the richness of the scene. Using filtering, the system will initially determine the colour of the subject's skin. In order to provide an accurate count of fingers, the image must first be put through a series of image pre-processing steps. The system locates the point that is closest to the contour point.

The image is degraded by the system in accordance with the centroid point. On the final image, we applied additional image pre-processing steps in order to ensure that the fingers appear correctly. The number of fingers is finally detected by the system, and the user is shown the count.

II. LITERATURE REVIEW MATERIALS AND METHODOLOGY

A paper by Chung et.al (2019) use of a machine learning method for gesture recognition was proposed. To solve this problem, the authors of this research built a deep convolutional neural network. When compared to traditional approaches, this one makes better use of context and geographic focus. Hand gestures include many procedures, including monitoring, tracking, and detection. There are two main types of deep CNN architecture, and they are Alex net and VGG net. The major motivation for this technique is the photographs of the hand that remain after tracking and recognition have been completed. The motion of the hand goes through several stages. Phase separation is the first step in reducing noise in a picture, followed by interference suppression, and finally background image reduction. Home use and human-computer communication both benefit from this app's features.

©IJRASET: All Rights are Reserved

Pre-processing the picture captured by the camera is crucial for ensuring safety and minimising the amount of computing work required for processing. There are also a plethora of other aspects that have a major bearing on the outcome, such as the lighting, the setting, the backdrop of the photograph, the signer's hand and body position and orientation, the camera's settings, and the depth of field.

Predicting a gesture from a series of video frames, such as a wave, sign language movements, or a clap, is the essence of vision-based human gesture recognition. Using gestures instead of a mouse or a remote control to interact with computers and other devices is one of the many advantages of this technology. In addition to being useful in the contexts already mentioned (consumer electronics and mechanical systems control, robot learning, and video games), gesture recognition from films offers a wide range of potential uses. Learning robots can benefit, for instance, from the online prediction of numerous actions for incoming movies from multiple cameras. Human gesture recognition in videos is difficult to model because of several factors, including inaccurate ground truth data for video data sets, a large number of possible gestures performed by actors in a video, heavy class imbalance in training data, and a large amount of data needed to train a robust classifier from scratch. Methods of deep learning, include Slow Fast two-path convolutional networks. network pre-training based on huge datasets of video activity recognition been demonstrated to boost performance on smaller datasets through transfer learning.

A. Taking Input

First, we extract the hand component responsible for the gesture from the input video by cropping the picture to a frame with predetermined bounds. To emphasise the gesture's lone arm, we employ linear image filtering techniques such as smothered sharpening and edgr enhancement to further improve the cropped image.

B. Image Refining

To aid in skin detection, we use a grey algorithm to correct for lighting. The three primary colors—red, green, and blue—represent the light source, or the brightness levels of the image as measured by an electronic camera (TRed Green and Blue filters). The primary goal of colour segmentation is to isolate certain features, such as lines, curves, and so on.

C. Detecting Skin Pixels

Several methods exist for transforming colours into skin-recognizable patterns. There are several colour spaces that might be useful for the skin detection procedure:

1) CIEXYZ
2) YCPCr
3) YIQ
4) YUV

Scatter matrices have been employed by various colour spaces as a performance metric for calculating skin and non-skin classes. Another issue is that following colour space transformation, comparing pixels that are skin and those that aren't skin has a different histogram. Three out of the four performance measures tested show that the YCBCr colour space works admirably. Therefore, the YCBCr colour space was used for the skin detection method. The Y component of YCBCr represents luminance data, whereas the Cb and CR components hold information about two colour differences: the difference between blue and a reference value and the difference between red and a reference value, respectively.

D. Cross-Correlation Technique

When comparing pixel intensities, cross correlation is employed. An improved picture from the input video is cross-correlated with all other images in our b=database. Ten photos with the highest corr2()values are saved. The image at the top of the file with the largest corr2() value indicates the greatest degree of resemblance to the input image. The picture identification process is complete, and the appropriate characters are being employed at last.

III. MODELING AND ANALYSIS

This system design demonstrates that edge detection is an integral part of the phase necessary to determine the hand's ultimate form. The second stage, which is the bulk of this study, is devoted to the feature extraction problem, and it makes use of two feature extraction techniques, namely hand contour and complex moments. In this research, we employed two distinct feature extraction methods: a boundary-based method for hand contour extraction and a region-based method for complicated moment extraction.

In order to solve issues like scaling, translation, and rotation that arise in the context of hand gesture detection, feature extraction techniques are implemented. We examine several issues with recognition and convergence of the Alexnet method used in the classification phase, where neural networks are employed to detect the gesture image based on its extracted feature. As can be seen in Fig. 1, the hand gesture recognition system may be broken down into the following steps. The initial phase is capturing photographs of hand gestures using a digital camera in a variety of settings (including variations in magnification, translation, and rotation). In the second phase, a pre-processor performs preliminary filtering and enhancements such edge identification and smoothing. Next, either hand contour or complex moment are used to extract characteristics from hand gesture photos.

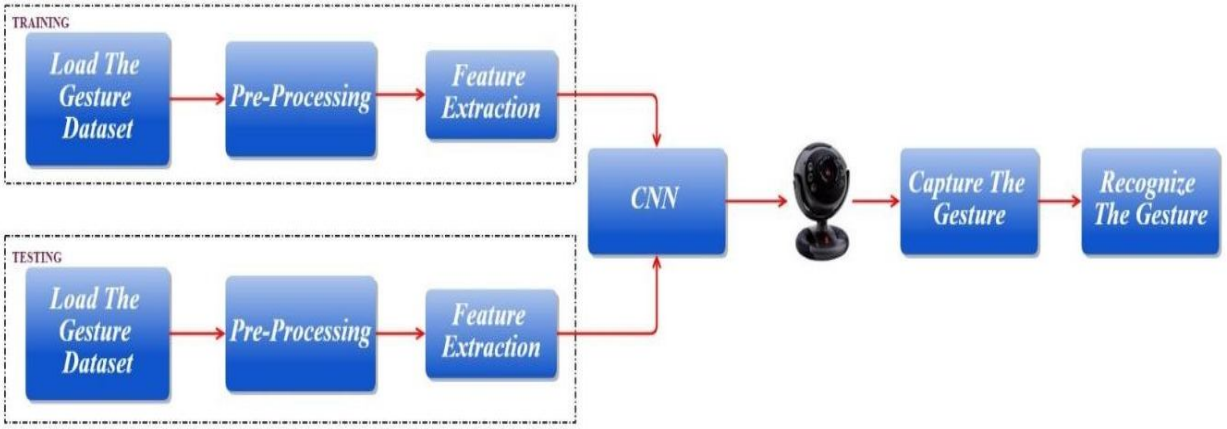


Figure 1: Architecture of hand gesture recognition

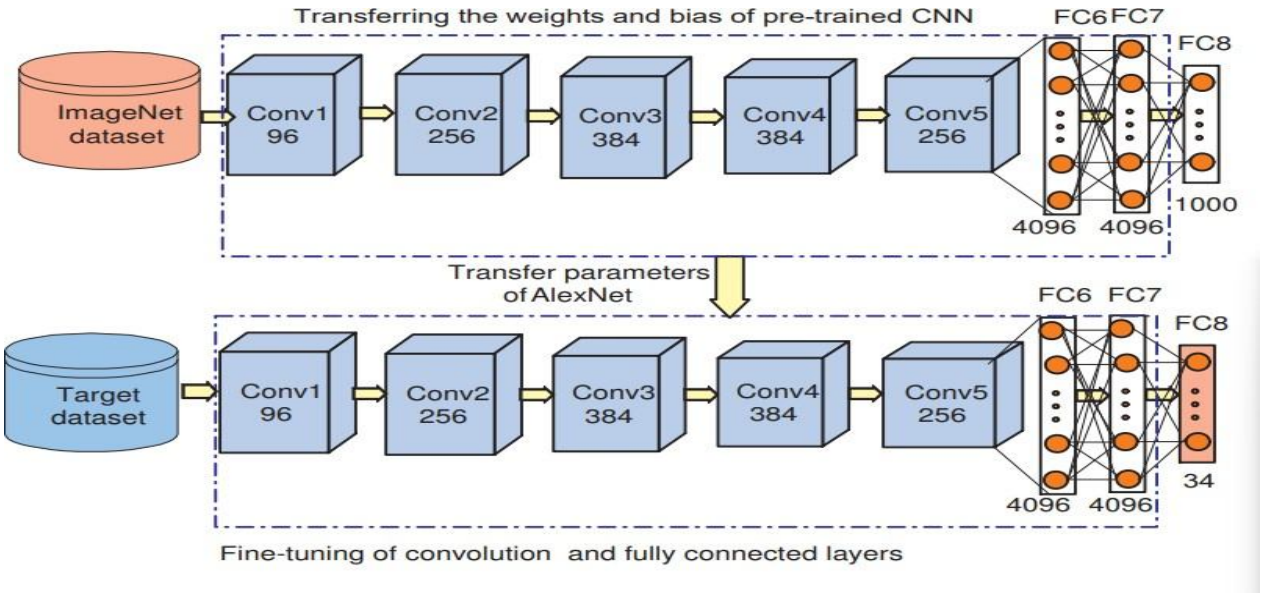


Figure 2: Fine-tuning process using pre-trained AlexNet on target hand gesturedataset.

As thematically depicted in a flowchart, a well-executed segmentation step results in a flawless features extraction step, both of which are crucial to accurate recognition. The process of extracting the features vector from a segmented image varies from one use case to the next. The characteristics have been represented in a number of different ways, and extraction methods have been similarly varied. The fingertips' location, the palm's centre, etc. were used in several techniques that relied on hand form, together with hand contour and silhouette. developed a feature vector consisting of 5 parameters; the first parameter represents the ratio aspect of the hand's bounding box, while the other parameters are the mean values of brightness pixels. The gesture is identified using a classification strategy once the input hand picture has been modelled and analysed. The accuracy of the recognition process can be improved by using a well-tuned Alexnet classification algorithm with well chosen feature parameters.

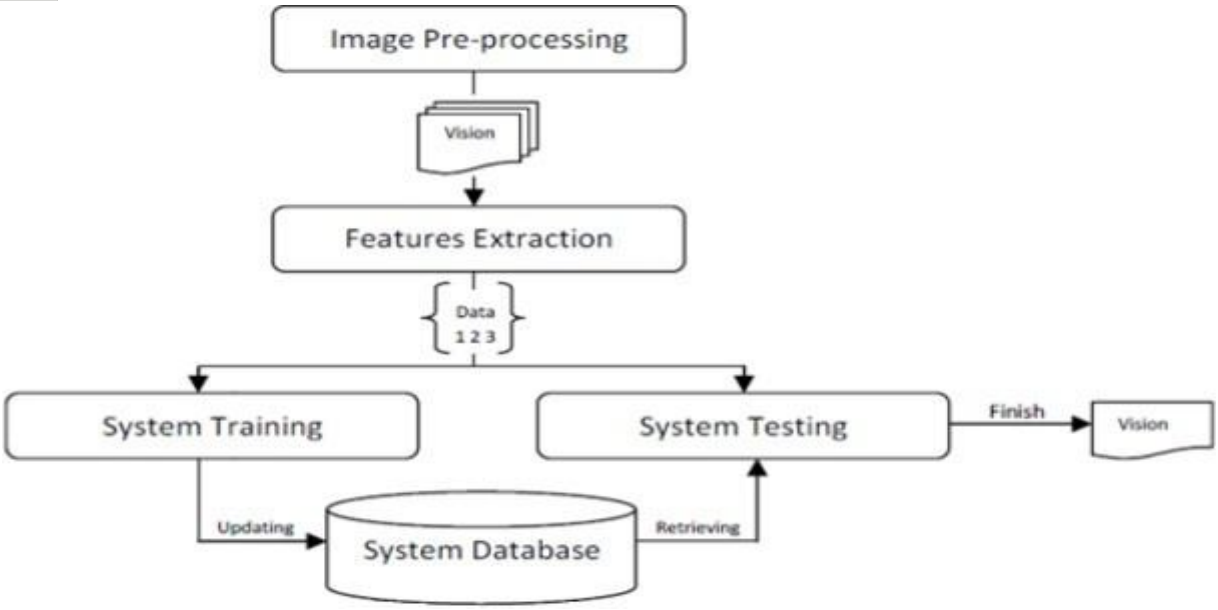


Figure 3: FLOWCHART OF THE PROJECT

IV. RESULTS AND DISCUSSION

In order to interpret hand gestures, computer vision-based systems first record the hand's motion on camera. The computer vision method of identifying a hand movement and then understanding what it was used to accomplish involves three primary procedures: detection, tracking, and recognition. The first step in the process is for the system to filter out any background and identify the skin tone. For an exact count of fingers, the image must first go through a series of image pre-processing steps. Closest point to the contour point is determined by the system.

The image is distorted by the system around the image's centroid. Further image pre-processing steps were applied to the final image to ensure that fingers showed up correctly. Finally, the system is able to count the user's fingers and show them on the screen.

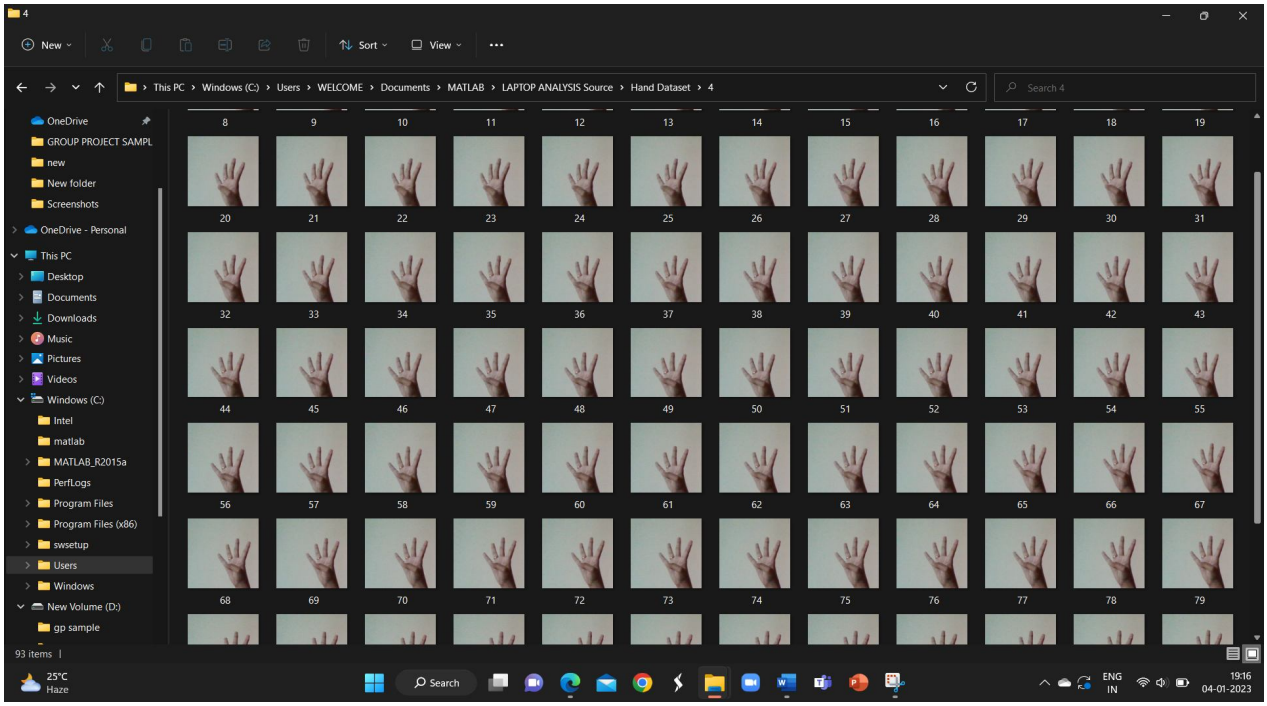


Figure1: Input image

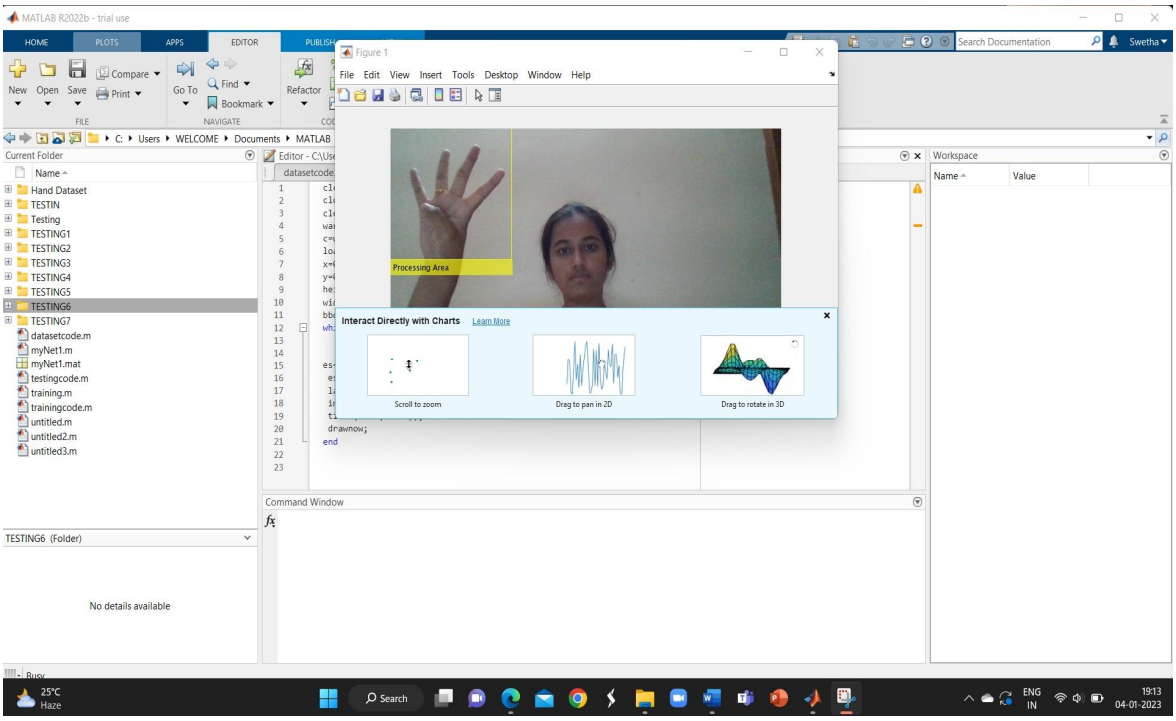


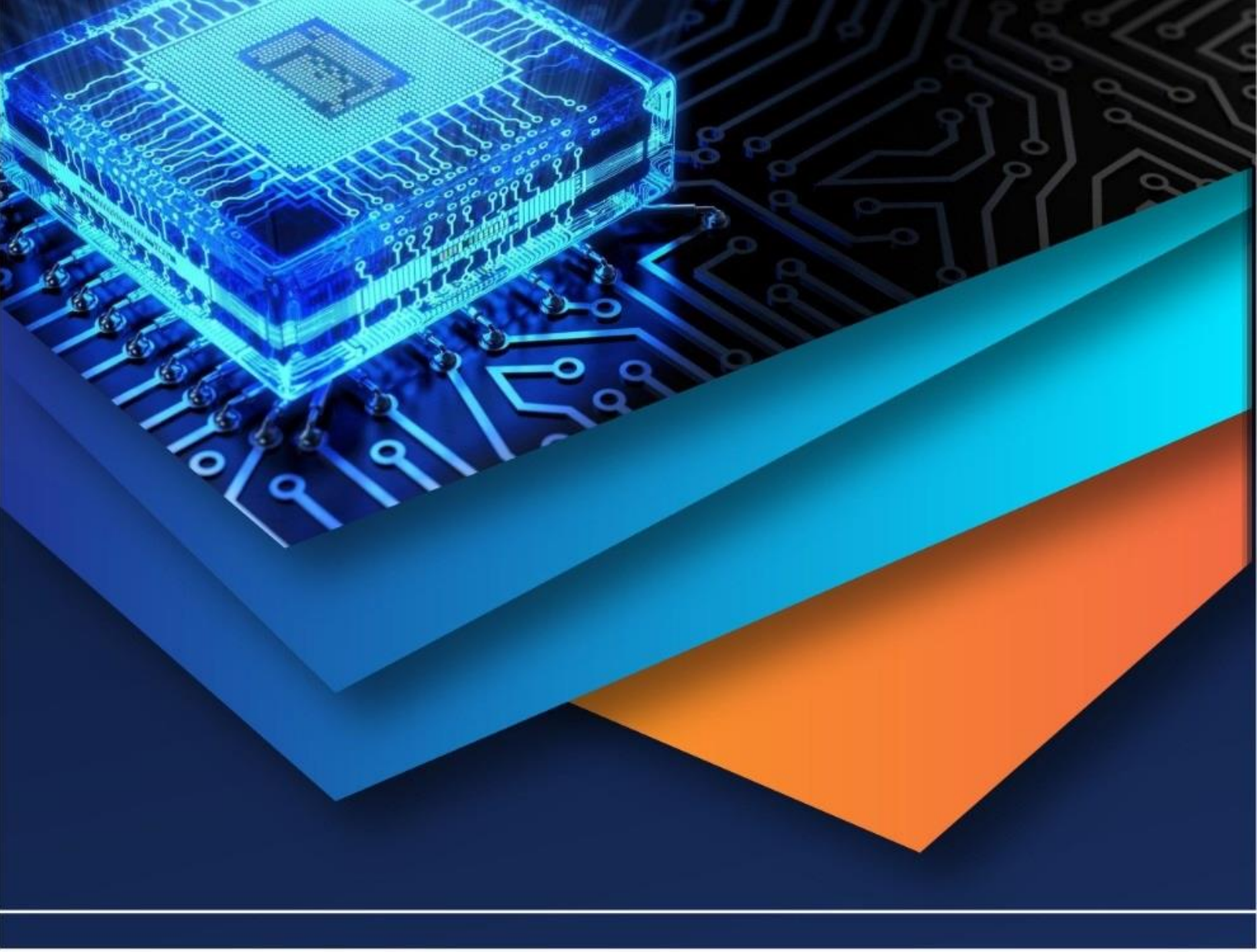
Figure2: Output image

V. CONCLUSION

When it comes to providing a natural HCI capability, hand gesture detection is crucial. It is already common knowledge that detection, segmentation, and tracking are the three most crucial elements in gesture recognition. With the help of the features extraction and classification capabilities of the Convolutional Neural Network (CNN) approach, a system has been developed for recognizing hand motions in this study. Seven short-distance 2-D and 3-D motions are captured using a variety of mobile cameras, backdrops, lighting, hand position, and hand form. A number of experiments were carried out to evaluate the efficacy of the CNN approach in both training and testing scenarios. Training was found to result in higher accuracy than testing did.

REFERENCES

[1] Chung, H. Y., Chung, Y. L., & Tsai, W. F. (2019, February). An efficient hand gesture recognition system based on deep CNN. In 2019 IEEE International Conference on Industrial Technology (ICIT) (pp. 853- 858). IEEE.
[2] Bao, P., Maqueda, A. I., del-Blanco, C. R., & García, N. (2017). Tiny hand gesture recognition without localization via a deep convolutional network. IEEE Transactions on Consumer Electronics, 63(3), 251- 257.
[3] Chung, Y. L., Chung, H. Y., & Tsai, W. F. (2020). Hand gesture recognition via image processing techniques and deep CNN. Journal of Intelligent & Fuzzy Systems, (Preprint), 1-14.
[4] Jayanthi, P., & Bhama, P. R. S. (2018, December). Gesture Recognition based on Deep Convolutional Neural Network. In 2018 Tenth International Conference on Advanced Computing (ICoAC) (pp. 367-372). IEEE.
[5] Zhan, F. (2019, July). Hand gesture recognition with convolution neural networks. In 2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI) (pp. 295-298). IEEE.



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)