MACHINE LEARNING APPROACH TO THE CLASSIFICATION AND IDENTIFICATION OF HAND GESTURE RECOGNITION USING PYTHON

Article *in* International Journal of Recent Scientific Research · December 2023

DOI: 10.24327/ijrsr.20231411.0821

Dr Mahanthesha U.
B. N. M. Institute of Technology

Available Online at http://www.recentscientific.com

CODEN: IJRSFP (USA)

International Journal of Recent Scientific Research
Vol. 14, Issue, 11, pp.4372-4377, November, 2023

ISSN: 0976-3031

International Journal of Recent Scientific Research

DOI: 10.24327/IJRSR

Research Article

MACHINE LEARNING APPROACH TO THE CLASSIFICATION AND IDENTIFICATION OF HAND GESTURE RECOGNITION USING PYTHON

[1]Dr. Anjaneya L H., [2]Dr.Mahanthesha U and [3]Banumathi K L

[1]Associate Professor, Department. of Electrical and Electronics Engineering,
Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India
[2]Associate Professor, Department of Artificial Intelligence & Machine Learning,
B N M Institute of Technology, Bangalore, Karnataka, India
[3]Assistant Professor, Department of Electronics and Communication Engineering,
Bapuji Institute of Engineering and Technology, Davanagere, Karnataka, India

DOI: http://dx.doi.org/10.24327/ijrsr.20231411.0821

ARTICLE INFO

Article History:

Received 30th September, 2023
Received in revised form 11th October, 2023
Accepted 17th November, 2023
Published online 28th November, 2023

Keywords:

Gesture recognition, OpenCV, human-computer interaction, python, machine learning

ABSTRACT

People who are deaf or dumb express themselves through gestures. Numerous applications utilizing gesture detection, computer vision, machine learning, etc. have been created. Python-based hand gesture recognition is an interesting area of computer vision that seeks to read and comprehend hand motions and movements recorded by cameras. It is utilized in a variety of fields, including robotics, virtual reality, sign language interpretation, and human-computer interface.The classification and interpretation of hand gestures is the aim of hand gesture recognition.In this paper, a straightforward method for gesture recognition is provided. Opencv and Numpy are two Python libraries used here. The method entails capturing real-time gestures and utilizing OpenCV tools to recognize them. The screen for output displays the output as a computer webcam stream.

INTRODUCTION

Python offers a potent environment for creating hand gesture detection systems thanks to its extensive collection of libraries and frameworks, which includes OpenCV.There are usually multiple steps in the procedure. Prior to segmenting or detecting the hand region, background subtraction, skin colour identification, or hand tracking algorithms may be used. Several features can be retrieved from the hand region once it has been isolated.

Next, a data set of hand gestures with labels can be used to train an algorithm that uses machine learning or a model that uses deep learning .Once trained, the model will be used to anticipate and identify hand movements instantaneously or on captured films. Gesture-based management of devices is then made possible by mapping the recognised gestures to certain commands or actions. Python's adaptability and large library of resources make it ideal for developing and implementing a variety of hand motion recognition methods. It offers a convenient setting for academics, programmers, and hobbyists to investigate and create cutting-edge applications that make use of the capabilities of computer vision and machine learning.

LITERATURE SURVEY

Machine learning can be used in a variety of ways to apply picture identification. Work has been done on video processing and depth camera sensing in addition to static picture identification. Utilizing vibrant other programming languages, a variety of procedures were built into the system to apply the procedural methods for the ultimate system's greatest efficiency.

The issue can be resolved and completely categorized into three similar approaches, namely the initial use of static picture identification methods and preprocessing techniques, the second use of deep learning models, and the third use of retired Markov Models. Colourful exploration papers used a variety of preprocessing methods and feature extraction techniques, including MATLAB as the primary programming language, Python, scale-invariant feature transform (SIFT), and Histogram of Gradients (HOG) to extract the image's feature descriptors. Ashish Sharma et al./Support Vector Machine was then used to classify these

Corresponding author:* **Dr. Mahanthesha U
Associate Professor, Department of Artificial Intelligence and Machine Learning, B N M Institute of Technology, Bengaluru.

feature descriptors into colourful rudiments [1]. Gestures have also been recognised using webcam images [2]. Xbox Kinect camera was used to recover the photos, and as part of the image pre-processing procedure, these images were converted to YCbCr format. HOG was used to extract features from the images, and SVM was used to classify the images according to their individual letters [3]. Combination of an OpenCV library-supported technique that uses videotape sequencing to obtain images and conversion into the YCbCr image format, followed by additional segmentation of the hand to produce images with less noise for improved recognition [4].

To identify gestures, an improved and optimized inheritable algorithm and adaptive filtering have been used [5]. Intricate signal processing techniques have also been applied to create a precise model that resembles a DWT and F-rate combination [6]. A novel recognition system using k-nearest neighbours (KNN) and the Bayesian approach to classify the images was demonstrated using videotape as the input[7]. Principal Component Analysis (PCA) and self-developed hand segmentation techniques were used to classify the images [8]. To interpret the change in kernel effect on the image recognition system, different types of SVM kernels and Multi-Class SVM with their comparison were used [9]. Various forms of segmentation Additionally, graph kernels have been contrasted to determine their impact on the dataset of American sign language-related images[10]. To analyze the behaviour, a similar method that uses a labelled graph kernel is also available [11]. Modern methods focused on using deep learning models to improve delicacy and speed up prosecution. To obtain information about the depth variation in the image and locate a new point for comparison, a model incorporating special tackling factors similar to depth cameras has been employed. A Convolutional Neural Network (CNN) has also been built to obtain the results [12].

By using a capsule network and adaptive pooling, a novel method of not using a pretrained model for system execution was developed [13]. Additionally, it was established that creating a deep belief network and lowering the layers of CNN-which uses a greedy strategy to accomplish this-produced better results than other methods [14]. To achieve the desired results, feature extraction using SIFT and classification using neural networks (CNN) were developed [15]. One of the models involved converting the photos to an RGB color scheme, developing the data using the motion depth channel, and afterwards creating the functioning system using 3D recurrent convolutional neural networks (3DRCNN) [16]. Another well-known method is HMM models for the recognition of sign language [7],[17],[18]. Since each of these publications uses a separate dataset with various figures and image kinds, comparing them becomes laborious. The majority of the systems were configured to use nearly noise-free images.

Research Gap

Illumination and Background Variations: It can be challenging to precisely separate the hand region due to shadows, reflections, and changes in illumination levels.

Palm Articulation and Occlusion: It might be difficult to precisely identify movements of the fingertips because fingers can occlude one another.

Limited Gesture Vocabulary: Current methods frequently concentrate on identifying a small number of predetermined hand motions. It can be challenging to expand the system's ability to recognise more motions or to dynamically learn new movements.

Sensitivity to Camera Position and Observation Angle: The size and shape of the hand may change depending on the camera angle, which may impact the precision of hand segmentation and extraction of features.

False positives must be kept to a minimum and robustness to noise must be ensured for dependable results.

PROPOSED SOLUTION

A. Block Diagram

Utilizing computer vision techniques and machine learning algorithms is necessary for hand gesture recognition. The system records and analyses hand gestures in the form of pictures or video frames, extracting pertinent aspects including hand motion, shape, and spatial relationships using PCA and ORB. In order to identify and recognise various hand movements, these attributes are then fed into a machine learning model, such as a convolutional neural network (CNN). To achieve precise recognition, the model is trained using a labelled collection of hand gesture photos. Real-time and reliable hand gesture detection is the goal of the suggested approach, which has applications in virtual reality, sign language translation, and human-computer interaction.

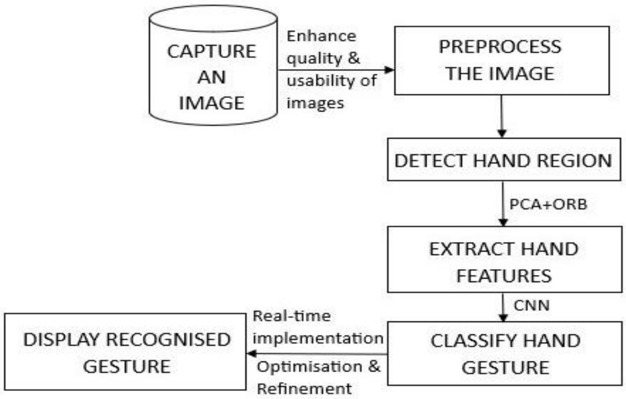

Fig. 1: Block Diagram

METHODOLOGY

B. Image Acquisition

Fig B.1.1: Trained dataset for A

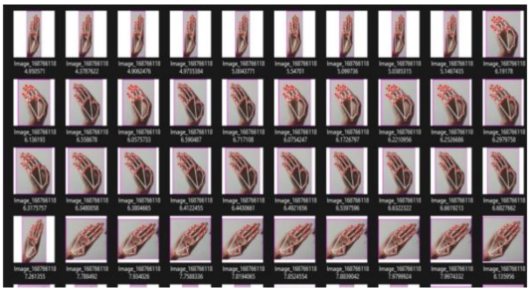

Fig B.1.2 Trained dataset for B

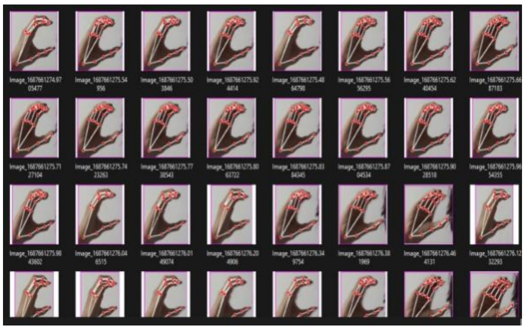

Fig B.1.3 Trained dataset for C

B.2 Image pre-Processing

Image Resizing: Resizing the input images to a standardized resolution helps maintain consistency and reduces computational complexity during processing. Open CV's resize () function can be used for resizing images. Image Normalization: Normalizing the pixel values of images ensures that the data has a consistent range, making it easier for the model to learn. Common normalization techniques include min-max scaling and mean normalization.

Background Subtraction: If the hand gesture is performed against a static or cluttered background, applying background subtraction techniques can help isolate the hand region.

Hand Segmentation: Hand segmentation techniques aim to separate the hand region from the rest of the image. Techniques such as thresholding, color-based segmentation, or skin tone detection can be used to identify and extract the hand region. Noise reduction: Image noise can negatively impact how well hand gestures are recognised.

B.1Hand Region detection

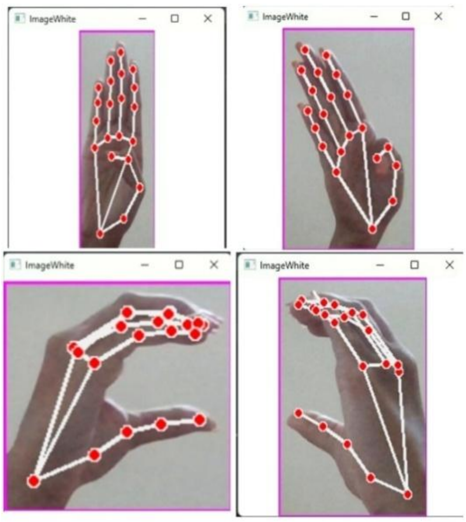

Fig B.3.1 Hand Region Detection

B.4 Feature Extraction

PCA

In image processing, PCA (principal component analysis) is used frequently. It includes preserving essential details while converting visuals into a lower-dimensional representation. The covariance matrix is computed using vectors that PCA creates by treating images as matrices. The associations between pixel values are captured by the covariance matrix. PCA determines the major components of the picture data by breaking down the matrix of covariance into eigenvectors and eigenvalues. The highest variance directions are represented by these components. PCA decreases the overall dimensionality of the picture by choosing a subset of eigenvectors with the largest eigenvalues. Feature extraction can be done using this reduction. It is possible to reassemble the reduced-dimensional representation back into an image.

ORB

A popular approach for detection of features in image processing is ORB (Oriented FAST and Rotated BRIEF). For purposes including object recognition, picture registration, and image stitching, it recognises different key points in images. Key point detection algorithm, FAST and feature description algorithm, BRIEF are combined in ORB. Pixel intensity comparisons are used to locate corners and edges during key point detection. In order achieve rotational invariance, ORB assigns rotations to key points. This process creates binary attribute descriptors by analysing sets of pixels in each point's immediate vicinity. With the help of these descriptors, tasks like picture registration and object recognition are made possible by the ability to match up key points in various images. Typical matching methods like nearest neighbour search and RANSAC are used. Combination of PCA and ORBFor feature extraction in image processing, Oriented FAST and Rotated BRIEF (ORB) and Principal Component Analysis (PCA) can be used together. While PCA helps lessen the dimensionality of these attributes, ORB can be used to identify and define different key points in images. The most important features can be kept while decreasing the feature space by using PCA for the descriptors produced by ORB. This combined approach of ORB and PCA enables efficient feature extraction, which can be beneficial for tasks such as object recognition, image classification, and image retrieval.

B.5Classification

Convolution Neural Network (CNN)- For classifying hand gestures, a pre-trained CNN model is utilized, and the model weights are kept in a file called "Model/keras_model.h5" that is commonly related to the Keras deep learning library. CNN-like neural network models are frequently built and trained using Keras. The class named Classifier from CVZ one. Classification Module, The pre-trained CNN model which generates forecasts based on input photos of hands. The hand motion in the img White image is classified using the CNN model by the Classifier class's get Prediction function.

$$Y(i,j) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(i-m, j-n).W + b \text{ ------1}$$

Equation B.5.1 Convolution Equation

The convolution operation applies a filter (kernel) to the input image to extract features. The equation for the convolution operation can be represented as follows:

Here,
Y(i, j) is the output feature map at position (i, j)
X(i-m, j-n) is the input value at position (i-m, j-n),
W(m, n) is the filter weight at position (m, n),
b is the bias term.
k = imgSize / w
hCal = math.ceil(k * h)

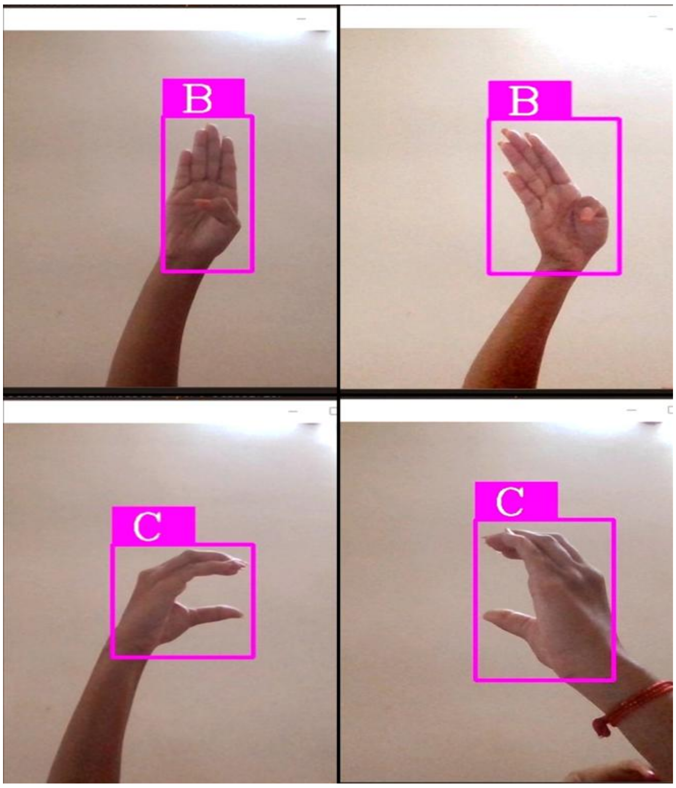

Fig B.5.1: Classification and identification

RESULTS AND ANALYSIS

The hand gesture recognition systemdemonstrates the system's overall competence in correctly extracting features using the PCA and ORB combination and categorising hand motions using CNN.

Gesture-wise Performance: For each distinct hand gesture, the system's performance was examined.

Real-time Performance: The system's effectiveness at recognising hand gestures in real-time was assessed, taking into account elements like processing time per frame and the capacity to handle a range of hand positions and movements.

Limitations & Future Improvements: Any restrictions or flaws in the system were noted, including its sensitivity to lighting conditions and the requirement for a controlled environment. Possible avenues for development included adding more sophisticated algorithms, investigating multi-modal input sources, and broadening the dataset diversity.

CONCLUSION

Successful recognition of hand motions using computer vision techniques and machine learning algorithms is demonstrated by the Python-based hand gesture identification system. The system's ability to extract informative information and employ machine learning models allowed for robust and real-time recognition. Due to the attained recognition accuracy and processing efficiency, the system is perfect for applications in virtual reality, sign language translation, and human-computer interaction. The employment of more advanced algorithms, research into multi-modal input sources, and dataset extension for

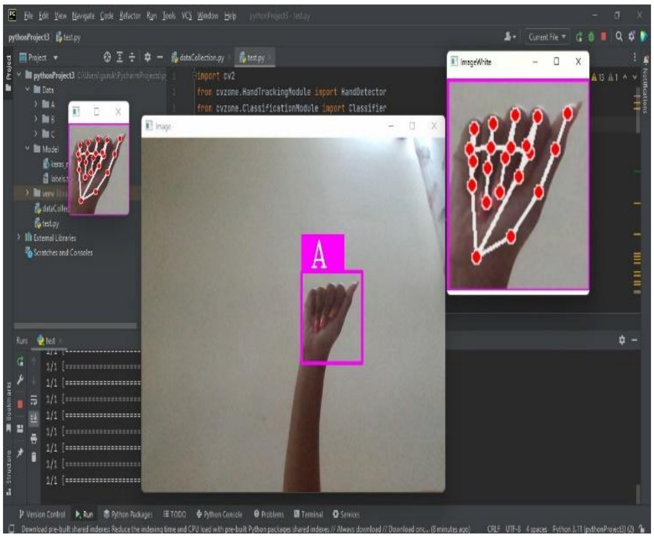

Fig IV.1 Project Demonstration for A

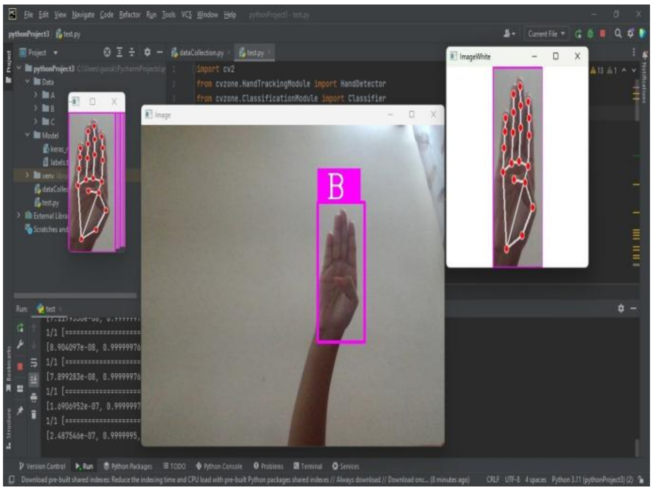

Fig IV.2 Project Demonstration for B

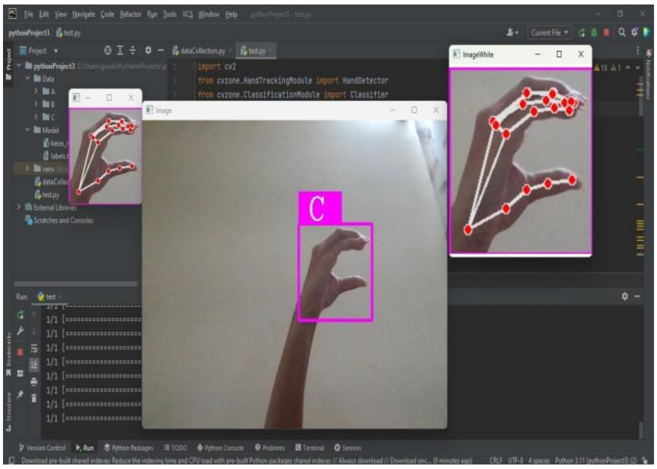

Fig IV.3 Project Demonstration for C

ACKNOWLEDGMENT

We are grateful to the friends, well wishers of Department of Artificial Intelligence and Machine Learning, BNMIT, BIET

Davangere for their insightful feedback and constant encouragement. Without them and their collective contributions, this research project would not have been possible.

References

1. Anita Jadhav, Rohit Asnani, RolanCrasto, OmprasadNi-lande&Anamol Ponkshe. (2015). Gesture Recognition using Support Vector Machine. International Journal of Electrical, Electronics and Data Communication, 36-41.
2. RuslanKurdyumov, Phillip Ho & Justin Ng. (2011). Sign Language Classification Using Webcam Images.
3. Lahoti, S. K. (2018). Android based American Sign Language Recognition System with Skin Segmentation and SVM. International Conference on Computing, Communication and Networking Technologies (ICCCNT) (p. 6). IEEE.
4. Rajeshree S. Rokade & Dharmpal D. Doye. (2015). Spelled sign word recognition using key frame. IET Image Processing, 381-388.
5. Rajesh Kaluri& Ch. Pradeep Reddy. (2016). A framework for sign gesture recognition using improved genetic algorithm and adaptive filter. Cogent Engineering.
6. Jaya Prakash Sahoo , Samit Ari & Dipak Kumar Ghosh. (2018). Hand gesture recognition using DWT and F-ratio based feature descriptor. IET Image Processing, 1780-1787.
7. Tamer Shanableh, KhaledAssaleh& M. Al-Rousan. (2007). Spatio-Temporal Feature-Extraction Techniques for Isolated Gesture Recognition in Arabic Sign Language. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 641-650.K. Elissa, “Title of paper if known,” unpublished.
8. Marwa Elpeltagy, MoatazAbdelwahab, Mohamed E. Hussein; Amin Shoukry; AsmaaShoala & MoustafaGalal. (2018). Multi-modalitybased Arabic sign language recognition. IET Computer Vision, 1031-1039. 190 Ashish Sharma et al. / Procedia Computer Science 173 (2020) 181–190 10 Author name / Procedia Computer Science 00 (2019) 000–000
9. Islam, M. R. (2018). Hand Gesture Feature Extraction Using Deep Convolutional Network for Recognizing American Sign Language. International Conference on Frontiers of Signal Processing (ICFSP) (p. 5). IEEE.
10. Harchaoui, Z., & Bach, F. (2007). Image Classification with Segmentation Graph Kernels. 2007 IEEE Conference on Computer Vision and Pattern Recognition.
11. Ruiqi Zhao & Aleix M. Martinez. (2015). Labeled Graph Kernel for Behavior Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1640-1650.
12. Bin Xie, Xiaoyu He & Yi Li. (2018). RGB-D static gesture recognition based on convolutional neural network. The Journal of Engineering, 1515-1520.
13. Jalal, M. A. (2018). American Sign Posture Understanding with Deep Neural Networks. International Conference on Information Fusion (FUSION) (p. 7). IEEE.
14. Ma, M. X. (2018). Design and Analyze the Structure based on Deep Belief Network for Gesture Recognition. International Conference on Advanced Computational Intelligence (ICACI) (p. 5). IEEE. [15] Shanta, S. S. (2018). Bangla Sign Language Detection Using SIFT and CNN. International Conference on Computing, Communication and Networking Technologies (ICCCNT) (p. 6). IEEE.
15. Ye, Y., Tian, Y., Huenerfauth, M., & Liu, J. (2018). Recognizing American Sign Language Gestures from Within Continuous Videos. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2145-214509, IEEE.
16. Daniel Kelly, John Mc Donald & Charles Markham. (2011). Weakly Supervised Training of a Sign Language Recognition System Using Multiple Instance Learning Density Matrices. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 526-541.
17. Jeroen F. Lichtenauer, Emile A. Hendriks & Marcel J.T. Reinders. (2008). Sign Language Recognition by Combining Statistical DTW and Independent Classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2040-2046.
18. Dr. H S Mohana and Mr. Mahanthesha U Published a Technical Book “ROBOTICS MADE SIMPLE-Beginners Guide to Building Robots” by Mantech Publications Pvt. Ltd. Ghaziabad, Dehli/NCR, 2021, ISBN: 978-81-948050-8-3. The book is available in Amazon link: https://www.amazon.in/Robotics-Simple-Beginners-Building-Robots/dp/8194805082
19. Mr. Mahanthesha U and Dr. H S Mohana, Published a research article "Covid-19 Social Distance Identification and Suspicious Activity Detection in Local and Remote Environments" International Journal Of Advance Research And Innovative Ideas In Education (IJARIIE), Volume 7 Issue 6 2021 Page 1003-1011.
20. Mr. Mahanthesha U, Anusha AU, Anusha C, Bhumika S and Harshini S, “Development of Real Time Facial Emotional Video Analysis with Speech Data Fusion using Random Forest Algorithm”, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Published by,www.ijert.org NCCDS-2021 Conference Proceedings, Volume 9, Issue 12.
21. Dr. H S Mohana and Mr. Mahanthesha U Published Research Article titled “Human Action Recognition Using Stip Evaluation Techniques” in Progress in Advanced Computing and Intelligent Engineering. Advances in Intelligent Systems and Computing, vol 1198. Springer, Singapore. https://doi.org/10.1007/978-981-15-6584-7_38, Print ISBN: 978-981-15-6583-0, Online ISBN: 978-981-15-6584-7, 30 October 2020.
22. Dr. H S Mohana and Mr. Mahanthesha U, published a research paper titled “Human Action Recognition Using Stip Evaluation Techniques” in Progress in Advanced Computing and Intelligent Engineering. Advances in Intelligent Systems and Computing, vol 1198. Springer, Singapore. https://doi.org/10.1007/978-981-15-6584-7_38, Print ISBN: 978-981-15-6583-0, Online ISBN: 978-981-15-6584-7, 30 October 2020.

23. Dr. H S Mohana and Mr. Mahanthesha U, published a research paper titled “Human Action Recognition Using Stip Techniques” in International Journal of Innovative Technology and Exploring Engineering (IJITEE), Scopus Indexed Journal ISSN: 2278-3075, Volume-9 Issue-7, May 2020.
24. Dr. H S Mohana and Mr. Mahanthesha U, published research paper titled “Smart Digital Monitoring For Attendance System” In Ieee Xplore, Doi: 10.1109/Icrieece44171.2018.9009166, 27 February 2020.
25. Dr. H S Mohana and Mr. Mahanthesha U, published research paper titled “Smart Digital Monitoring For Attendance System” in the proceedings of International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering - (ICRIEECE), School of Electrical Engineering, Kalinga Institute of Industrial Technology (KIIT), Bhubaneswar, India. (Deemed to be University U/S 3 of the UGC Act, 1956), July 27th & 28th 2018.
26. Mr. Mahanthesha U, Pooja, Rakshitha R, Sherley Vincent, Smitha S Potdar, published a technical paper “Depth Based Tracking Of Moving Objects Using Kinect” in International Journal of Engineering Research and Technology (IJERT), Volume 6, Issue 13, ISSN No: 2278-0181, 17th May 2018.
27. Mr. Mahanthesha U, Supriya K S, Supriya Pal, Sushma Bhagavathi B S, Vidya P G, published a technical paper “A THERAPY SYSTEM FOR ARTICULATION DISORDER CORRECTION” in the Journal of Engineering Research and Technology (IJERT Volume 6, Issue 13, ISSN No: 2278-0181, 17th May 2018.
28. Mr. Mahanthesha U and Dr. H S Mohana, published a technical paper “Identification Of Human Facial Expression Signal Classification Using Spatial Temporal Algorithm” in International Journal of Engineering Research in Electrical and Electronic Engineering (IJEREEE), Vol 2, Issue 5, ISSN (Online) 2395-2717,May 2016.
29. Mr. Mahanthesha U, Bindu K, Kavyashree A L, Lakshmi Kiran V, Mallika, published a technical paper “Wireless Data Transmission Through Light” in the International Research Journal of Engineering and Technology (IRJET), Volume: 03 Issue: 06, e-ISSN: 2395 -0056, p-ISSN: 2395-0072, June-2016.
30. Smt. Geetha Jayaram Bhat and Mr. Mahanthesha U, published a technical paper titled “Application Dependent Geo-Routing Protocol On Multi-Radio Networks” in “International Journal on Recent and Innovation Trends in Computing and Communication” (IJRITCC), Volume 2, Issue 7, ISSN: 2321-8169, July-2014.

How to cite this article:

Anjaneya L H., Mahanthesha U and Banumathi K L, 2023. Machine Learning Approach to the Classification and Identification of Hand Gesture Recognition Using Python. *Int J Recent Sci Res.* 14(11), pp.4372-4377.