

# Survey on Vision Based Hand Gesture Interface for Controlling Multimedia Player

Mr Rohan Agrawal<sup>1</sup>, Aditya Kuchekar<sup>2</sup>, Vaibhav Suryawanshi<sup>3</sup>,  
Vihang Lawand<sup>4</sup>, Dr. M.L. Bangare<sup>5</sup>

Students, Department of Information Technology<sup>1,2,3,4</sup>

Associate Professor, Department of Information Technology<sup>5</sup>

Smt. Kashibai Navale College Engineering, Pune, Maharashtra, India

**Abstract:** As computers become additional pervasive in society, facilitating natural human-computer interaction (HCI) can have a positive impact on their use. Hence, there has been growing interest within the development of recent approaches and technologies for bridging the human-computer barrier. the last word aim is to bring HCI to a regime wherever interactions with computers are going to be as natural as associate degree interaction between humans, and to the current finish, incorporating gestures in HCI is a crucial analysis space. Gestures have long been thought-about as associate degree interaction technique that may doubtless deliver additional natural, creative, and intuitive strategies for human activity with our computers. Hand gesture recognition is one amongll|one amongst|one in every of the systems that may notice the gesture of the hand in a period of time video. The gesture of hand is classed inside a definite space of interest. during this study, planning hand gesture recognition is one among the difficult jobs that involves 2 major issues. foremost is that the detection of the hand. Another drawback is to form an indication that's appropriate to be used one hand at a time. This project concentrates on however a system might notice, acknowledge and interpret hand gesture recognition through computer vision with the difficult factors that variability within the create, orientation, location, and scale. To perform well for developing this project, differing kinds of gestures like numbers and sign languages got to be created during this system. The image taken from the period of time video is analyzed via Haar-cascade Classifier to notice the gesture of hand before the image process is finished or in different words to notice the looks of hand in a very frame. during this project, the detection of hand are going to be done mistreatment the theories of Region of Interest (ROI) via Python programming. the reason of the results are going to be targeted on the simulation half since the distinction for the hardware implementation is that the ASCII text file to scan the period of time input video. the event of hand gesture recognition mistreatment Python, OpenCV, and YOLO V3 will be enforced by applying the theories of hand segmentation and also the hand detection system that uses the Haar-cascade classifier.

**Keywords:** Hand Gesture, OpenCv, Python, Machine Learning

## I. INTRODUCTION

With the event in Computer Vision and Human-Machine Interaction, the computer holds the foremost vital role in our everyday life. Human-Computer Interaction will give many benefits with introducing the various natural styles of device-free communication. Gesture recognition is one among the many kinds of them to move with humans Gestures square measure the natural variety of action that we regularly employed in our daily life. except for computer applications to move humans with the machine, interaction with devices like keyboard, mouse, etc. should be needed. because the varied hand gestures square measure of times employed by humans, this project aims to scale back external hardware interaction that is needed for computer application, and thus this causes the system additional reliable to be used with ease. The task of recognising hand gestures is one among the most and vital problems in computer vision. With the newest advances in info and media technology, human computer interaction (HCI) systems that involve hand process tasks like hand detection and hand gesture recognition.

The first step in any hand process system is to notice and find the hand within the period of time video from the digital camera. The detection of hand is difficult thanks to variation in create, orientation, location and scale. Image acquisition involve capturing image within the video frame by frame employing a digital camera. The captured pictures undergo the

image pre-processing method that involves color filtering, smoothing and thresholding. After with success sleuthing the gesture it's recognised and is employed to present commands to the transmission player. Through this project, we have a tendency to aim to supply a natural device-free interface for dominant a transmission player mistreatment Vision-based hand gestures recognition interface for dominant a transmission player.

## II. MOTIVATION

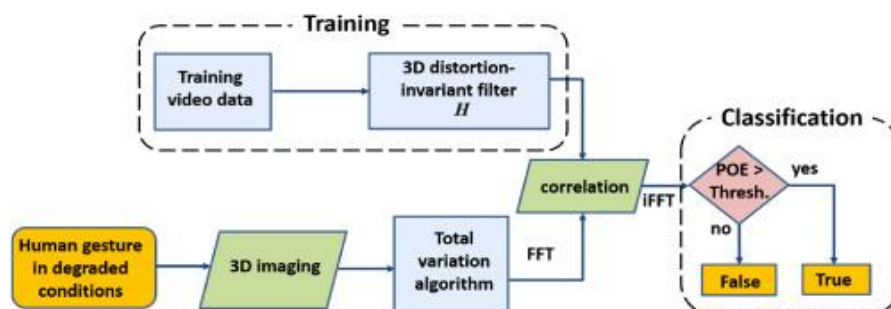
Nowadays, Human-Computer Interaction is increasing. to manage the transmission player while not employing a mouse or keyboard, we are able to use hand gestures to present commands to the computer, which can command the media player per instructions: given by North American country mistreatment hand gestures.

Many times, once our elders watch or hear music/video, they're unable to move with the media player as they're not accustomed to the interface during this state of affairs, A Vision-based Hand Gesture Interface for dominant transmission Player will facilitate them to move with media play mistreatment hand gestures.

Many times, once we square measure at a celebration or off from a mouse or keyboard whereas paying attention to music or observation videos, it becomes troublesome to move with the media player to alter, stop, or pause music/video. In these case, usage of hand gestures will build our life straightforward by dominant media player through hand gestures.

## III. LITERATURE SURVEY

Hand gesture recognition has become a crucial analysis space. Text Mining is that the discovery by computer of recent, antecedently unknown info, by mechanically extracting info from totally different written resources. during this paper, a Survey of Hand gesture recognition mistreatment OpenCV and applications are given. a number of the technologies that are developed and may be employed in the Hand gesture recognition method square measure making ready the binary mask, Computing the contour and it's bell-shaped hull, and sleuthing the fingertips as shown below in fig(1)



**Figure 1: Gestures recognition using OpenCV**

### 3.1 Recent Studies

Several papers and comes have targeted the difficulty of hand gesture recognition. Francis et al[1] compared strategies for gesture recognition in cars, evaluating accelerometers-based, glove-based, and Kinect-based approaches. Mitra et al[2] analyzed additional computationally serious strategies mistreatment hidden Andrei Markov models and finite state machines. Ghotkar et al[3] given a completely unique approach handy segmentation and gesture recognition mistreatment totally different color areas. The strategies planned by Francis et al needed further hardware, whereas those planned by Mitra et al were computationally serious, requiring classification and interval.

Our goal was thence to follow the instance of Ghotkar et al, and explore the additional basic strategies of hand segmentation and gesture recognition accessible, applying them to execute easy controls on a media player.

## IV. EXISTING METHODOLOGY

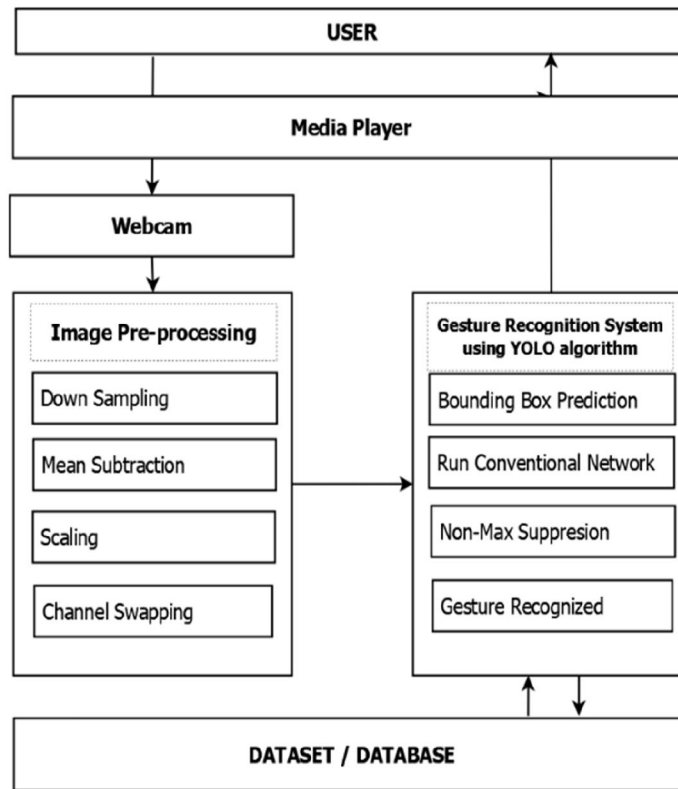
The exact meaning of object detection, tracking and identification and point out the general problems regarding the object detecting, tracking and identification. To track multiple objects, different methods are in use such methods have shown very good performance, considering more frames before making association decisions should generally help better overcome ambiguities caused by longer-term occlusions and false or missed detections.

Many global approaches that use more information have been explored to overcome errors of detections. Most proposed hand gesture systems can be divided into two categories of computer vision techniques. First, a simple approach is to use image processing techniques via OpenCV library and possibly other tools to provide interaction in real time, which considers time consumption because of real-time processing.

## V. PROPOSED METHODOLOGY

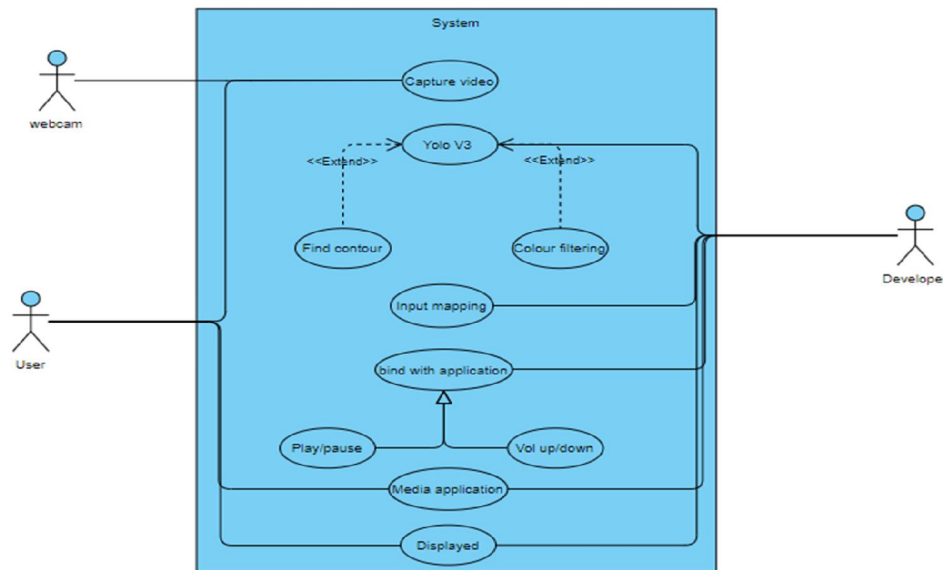
In our proposed system, we are taking into consideration the need for improvement of hand gesture recognition accuracy as well as keeping the processing time as low as possible while implementing Machine Learning and Deep learning algorithms.

### 5.1 Proposed System Architecture

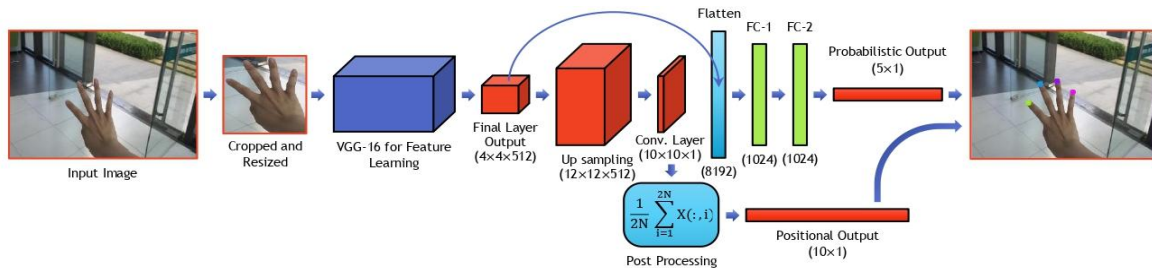


- **Webcam:** When the media player opens, the webcam automatically starts to function and captures live video which then is read by OpenCV frame by frame and fed into the YOLO neural network for object detection.
- **Image preprocessing:** The proposed architecture begins with the user performing some hand gestures facing the camera. The camera captures live video which is then divided into multiple image frames. **Gesture Recognition System using yolo V3 algorithm:** After image pre-processing, the output is fed into a gesture recognizing system that uses the YOLO algorithm. The YOLO algorithm recognizes the gesture in the following phases:
- **Bounding Box Prediction:** In this Each grid cell predicts a bounding box involving the x, y coordinate and the width and height and the confidence. **Non-Max Suppression:** In this it select one entity (e.g., bounding boxes) out of many overlapping entities. It is used to choose specific selection criteria to arrive at desired result **Gesture Recognition:** Using above method the final hand region and gesture is recognised.

## 5.2 YOLO V3



A dynamic hand gesture recognition algorithm based on You Only Look Once: Version 3 (YOLOv3) is proposed for continuous dynamic hand gesture recognition. The YOLOv3 object detection algorithm is used to train and identify the grayscale image which include the information of continuous dynamic hand gestures.



The effectiveness of the proposed method is verified by the recognition confusion matrix. And the proposed method has an High recognition accuracy for four custom dynamic hand gestures. In YOLOv3 single neural network is applied to the full image. This neural network divides the image into regions. The divide regions are then used to predict bounding boxes according to their probabilities for each region. The weight of the bounding boxes are given by the probabilities. In mAP measured at 0.5 IOU YOLOv3 is on par with Focal Loss but about 4x faster. Moreover, you can easily tradeoff between speed and accuracy simply by changing the size of the model, no retraining required. YOLOv3 is extremely fast and accurate. This algorithm will be used for gesture recognition and for training the model by bringing a balance between speed and accuracy.

## VI. TECHNOLOGY PLATFORM

- PC
- CPU - AMD® A10-7850k Radeon R7
- GPU - Nvidia GeForce GT 1030-4GB
- SSD/HDD - 256 GB
- OS - Windows 10
- Dependencies
- Python 3.7
- Camera

- 720p resolution camera
- Network configuration and training
- Assign 93 filters to all convolutional layers before every YOLO layer
- Batch size of 64 with 16 subdivisions.

## **VII. CONCLUSION**

At last, we conclude that A vision-based Hand body language gesture recognition using YOLO V3 generally refers to the process of recognising hand gestures using various processes. Hand gesture recognition using YOLO V3 is a young interdisciplinary field that draws using yolov3 algorithm, in this the gesture is recognised in three phases i.e. using yolov3 algorithm, in this the gesture is recognised in three phases i.e bounding box prediction, running conventional network, non maximum suppression. There are different kinds of techniques used for hand gesture recognition.

The effectiveness of the proposed method is verified by the recognition confusion matrix. And the proposed method has an High recognition accuracy for four custom dynamic hand gestures. This study has found that generally hand gesture recognition is made easy using YOLO V3. This work would be implemented to detect hand gestures and then further recognise and implement to control multimedia player.

## **REFERENCES**

- [1]. J. Francis and A. B K, "Significance of Hand Gesture Recognition Systems in Vehicular Automation-A Survey", International Journal of Computer Applications, vol. 99, no. 7, pp. 50-55, 2014.
- [2]. A. S.Ghotkar and G. K. Kharate, " Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction", International Journal of Human Computer Interaction, vol. 3, no. 1, pp. 15-25, 2018.
- [3]. Hand Recognition and Gesture Control Using a Laptop Web-camera Zi Xian Justin Yeo Exchange Student with Stanford University 450 Serra Mall, Stanford, CA 94305
- [4]. Hand gesture recognition on python and opencv To cite this article: Ahmad Puad Ismail et al 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1045 012043
- [5]. Image Processing Algorithm for Appearance-Based Gesture Recognition (IEEE 2019)
- [6]. Hand Gesture Modelings and Recognition using Geometric Features: Canadian Journal on Image Processing and Computer Vision, vol. 3, pp. 12-26, March 2012
- [7]. Gesture Controlled Media Player using TinyYoloV3( May 2020)