**Dynamic Hand Gesture Recognition using Neural Network**

**A Project Work Synopsis**

*Submitted in the partial fulfilment for the award of the degree of*

**BACHELOR OF ENGINEERING**

**IN**

**COMPUTER SCIENCE WITH SPECIALIZATION IN**

**ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**Submitted by:**

20BCS6741 Ansh Pandey

**Under the Supervision of:**

**Mr. Gagan Tiwari (E13616)**

**Mr. Mahadev (E13868)**



**CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,**

**PUNJAB**

**September, 2022**

# Abstract

Dynamic Hand Gesture Recognition is the most popular and widely used technology in the field of Computer Vision. It let the computer understand the hand gestures performed by the user according to the data stored in its database. In this paper, a recognition system is developed which can recognize dynamic hand gestures. It has two parts: (I) Virtual Keyboard (II) Virtual Mouse. In this recognition system, we are basically recognizing the actions performed by the hand. By using CNN and RNN for recognizing the continuous motion of hand.

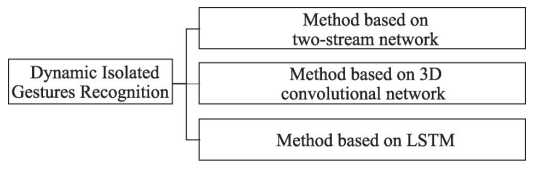
**Keywords:** hand gesture recognition; deep learning; convolutional neural networks, recurrent neural networks, computer vision, neural networks, human-computer interaction

# Table of Contents

|  |  |
| --- | --- |
| Title Page |  |
| Abstract |  |
| 1.    Introduction |  |
| 1.1  Problem Definition |  |
| 1.2 Project Overview |  |
| 1.3 Hardware Specification |  |
| 1.4 Software Specification |  |
| 2.    Literature Survey |  |
| 2.1 Existing System |  |
| 2.2 Proposed System |  |
| 2.3 Literature Review Summary |  |
| 3.    Problem Formulation |  |
| 4.    Research Objective |  |
| 5.    Methodologies |  |
| 6.    Experimental Setup |  |
| 7.    Conclusion |  |
| 8.    Tentative Chapter Plan for the proposed work |  |
| 9.    Reference |  |

# INTRODUCTION

Hand Gesture Recognition is one of the most interesting field of human-computer interface. Hand Gesture Recognition System recognizes the actions performed by the user and then send some information or perform some action as described or stated in its database. Human actions can be sensed using either camera or sensors. There can be two types of hand gestures, one is Static and second one is Dynamic. In static hand gesture, there we need to put the hand stable and its all about the shape of the hand. While in Dynamic hand gestures, there is a continuous movement of hand.



*Figure Source: https://pdf.sciencedirectassets.com/*

## 1.1 Problem Definition

Dynamic Hand Gesture Recognition using Neural Network is a system that will recognize the movements done by hand continuously and will perform the task stored in the database for that action.

This system has two parts:

* Virtual Keyboard
* Virtual Mouse

Using both these features we can access the apps, open them and perform different task as performed by the mouse and with the help of keyboard we can type virtually.

## 1.2 Problem Overview

We are developing a system that will recognize the dynamic movements of hand and will then perform some tasks like typing, opening an app, drawing, etc. It has two features Virtual Keyboard and Virtual Mouse. In Virtual Keyboard, we can type anything anywhere. Virtual Keyboard is necessary in a computer/laptop as sometimes we have faced issues related to keyboard that some keyboard buttons are missing or some buttons don’t work. In virtual Mouse, we can use our finger the same as a mouse but virtually. The camera will detect the movements of hand and detect its actions and will response accordingly. Both Virtual Keyboard and Virtual Mouse features will help in better human – computer interaction and will improve user Experience.

## 1.3 Hardware Specification

* Hardware

## 1.4 Software Specification

* Machine Learning -
* Vs Code
* Jupyter Notebook
* Data Cleaning -
* Excel
* Data Anlysis –
* Power BI

# 2. LITERATURE SURVEY

## 2.1 Existing System

The researches on gesture recognition were mainly carried in and after 1964, when a scientist named Ivan E. Sutherland uses a pen with button to draw a line in the year 1964.

Talking about the technologies in Deep Learning and Neural Network, there are many approaches defined to perform gesture recognition. For identifying extracted features, we can use Support Vector Machine, Decision Trees, K – Nearest Neighbor, etc.

Following are the few existing systems –

1. Wearable armband: It is a real – time gesture recognition system, developed using electromyography sensors and algorithms such as k – Nearest Neighbor and Decision Trees were also used.
2. Vision – based Devices: In such devices, recognition of hands and its poses and of face is done using feature based template matching techniques.

## 2.2 Proposed System

The system proposed by us will have features like Virtual Keyboard and Virtual Mouse that will help us to access the apps, searching of apps, typing, etc. on the system.

In Virtual Keyboard, we can type anything anywhere in the text area. Virtual Keyboard is necessary in a computer/laptop as sometimes we have faced issues related to keyboard that some keyboard buttons are missing or some buttons don’t work, at that moment Virtual Keyboard will help us to use the system. In virtual Mouse, we can use our finger the same as a mouse but virtually. The camera will detect the movements of hand and fingers and will detect its actions and will response accordingly.

## 2.3 Literature Review Summary (Minimum 7 articles should refer)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Year and**  **Citation** | **Article/ Author** | **Tools/ Software** | **Technique** | **Source** | **Evaluation Parameter** |
| 2011 | Mahanta C., Srinivas Yadav T. and Medhi H. | - | background subtraction method, particle filter, Radial basis function, MPEG-7 ART based shape descriptors | - | - |
| 2018 | Nuwan Munasinghe | - | Gaussian mixture based background/foreground segmentation algorithm, feedforward neural network | - | - |
| 2018 | Kenneth Lai and Svetlana N. Yanushkevich | - | CNN+LSTM network | - | - |
| 2019 | Noorkholis Luthfil Hakim, Timothy K. Shih, Sandeli Priyanwada Kasthuri Arachchi, Wisnu Aditya, Yi-Cheng Chen and  Chih-Yang Lin 3 | - | Three-dimensional Convolutional Neural Network, y Long Short-Term Memory (LSTM) model, | - | - |
| 2021 | Yuanyuan SHI, Yunan LI, Xiaolong FU, Kaibin MIAO, Qiguang MIAO | 20BN-JESTER dataset, IsoGD dataset, e Montalbano dataset | two-stream recurrent neural networks, canonical correlation analysis | - | - |
| 2022 | JiminYu, Maowei Qin & Shangbo Zhou | GeForce GTX 1080 Ti GPU, 2.40GHz 6-core CPU, Python 3.6, cuda 10.1, cuDNN 7.6, Tensorfow-GPU 2.3.0. | Fractional HS optical fow model, Key frame extraction, Data augmentation | - | - |
| 2022 | Manoj. H. M, Amrutha D Bhumika M Shivani Hosangadi, Shravya | Datasets - UCF-102 and HMDB-53 | CNN, RNN, LSTM (Long short - term memory) model | - | - |

.

# 3. PROBLEM FORMULATION

Dynamic Hand Gesture Recognition is the most popular human to human communication via continuous hand movements and widely used technology in the field of Computer Vision. In this project we are creating Hand Gesture Recognition System having functions such as Virtual Keyboard and Virtual Mouse.

The importance of hand gesture recognition system is that it improves user experience and is also helpful for people having disability such as mute people. Not only this, it also used in advancement of technology as per growing human needs. People finds it interesting and exciting to work on such systems and it more convenient for them to use these advance systems.

Need of this project of virtual keyboard and virtual mouse is that it not only helpful for impairments of people but also we sometimes face issues like the keyboard and mouse of the system are not in working conditions, keys are broken or misplaced, so these (virtual keyboard and virtual mouse) are more convenient to use and are preferred to use.

# 4. OBJECTIVES

Objective of this project, The Dynamic hand gesture recognition is to make an application that controls some specific functionalities of computer using hand gesture via integrated webcam.

Before the rapid development of deep learning, research on dynamic gesture recognition mainly relied on manual extraction of features and then building sequence models for recognition but the accuracy of these models was low and inefficient, so as the demand increases for accuracy, the model is now mostly depend on deep learning. As dynamic gesture data is generally presented in video form, it is difficult for 2D convolutional neural networks (CNNs) to extract sufficient feature information, which has prompted researchers to explore new directions. With Du Tran et al.8 proposing the C3D network model, which solved the problem of retaining both temporal and spatial features, the dynamic gesture recognition problem was widely solved by deep learning methods. the C3D model was also the earliest 3D CNN model.

Most current algorithms on dynamic gesture recognition using 2D CNN serialize the video datasets as a chart or a single image, which loses the information on the variation of key spatio-temporal features. Our proposed algorithm intuitively extracts the temporal and spatial information in the video datasets and fuses the two, making full use of the key features in the video datasets. To summarize, the main contributions of this paper are:

• An improved the HS model is proposed with the fractional order method, in which fractional-order and deep learning are creatively combined;

• An improved clustering algorithm is proposed based on a tradition model, which can effectively extract the keyframes of complex actions;

• A strategy for network input is proposed to use the original keyframe mosaic image and the optical fow keyframe mosaic image instead of the video data, which effectively reduces the size of the data set and the difficulty of training

In our application we are covering three functions using hand gesture. These functions are Virtual keyboard and AI Virtual mouse system.

Virtual Keyboard: It is an in-air tapping keyboard for AR/VR HMDs that supports ten figure typing experience. In designing virtual keyboard, we utilized the conventional US keyboard as our default layout. It allows users to freely change the size of virtual objects that we can apply or designed our keyboard to be mobile. The physical keyboard consists of 42 cubes, each with embedded collision interface defining the action upon fingertips collision with the keys. Hand models are driven with the coordinates of the fingertips, palm centre and finger joints, which is read from the hand tracking sensor.

Virtual mouse: The main objective of the proposed AI virtual mouse system is to develop an alternative to the regular and traditional mouse system to perform and control the mouse functions, and this can be achieved with the help of a web camera that captures the hand gestures and hand tip and then processes these frames to perform the particular mouse function such as left click, right click, and scrolling function

# 5. METHODOLOGY

This section is concerned with the recognition of dynamic hand gestures. Since the movement of the hand conveys important information about the message users try to communicate, in addition to the spatial information, temporal features which represent the motion of the hand are necessary for recognizing dynamic hand gestures.

Pre-processing

The main aim of pre-processing is to eliminate noise present in the captured images. Noise creates problems in hand segmentation and tracking phase, and hence reduces the performance of the overall system. Mostly salt and pepper noise and additive Gaussian noise are present in the images. These types of noises can be removed by using median filter and mean filter respective.

Segmentation

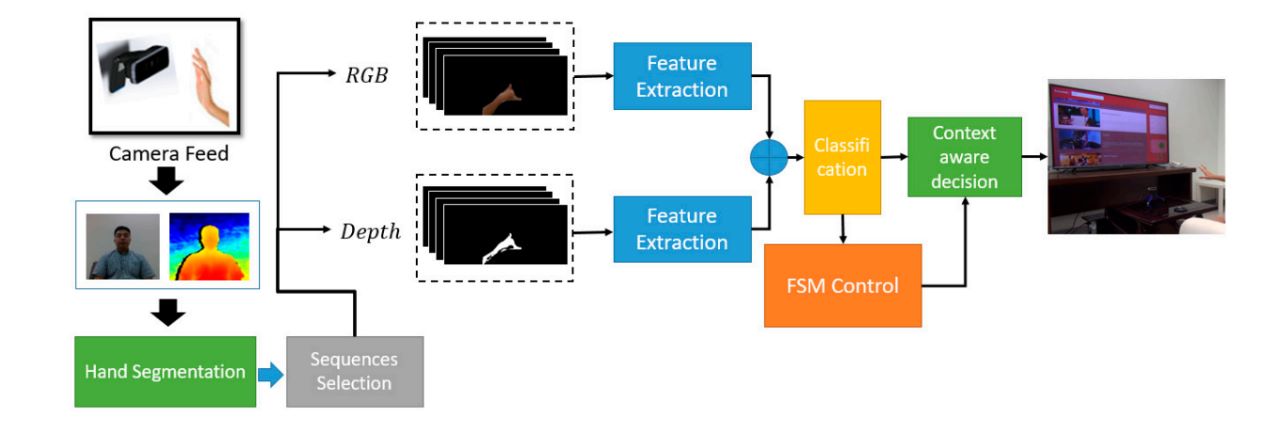
In order to realize high-precision hand gesture recognition, it is necessary that the hand region is extracted correctly from the background. The main goal of this step is to segment the moving hand from the back ground. Hand region extraction method is implemented by using background subtraction algorithm. One major advantage of this method is that it extracts the gesture region without future frame images.

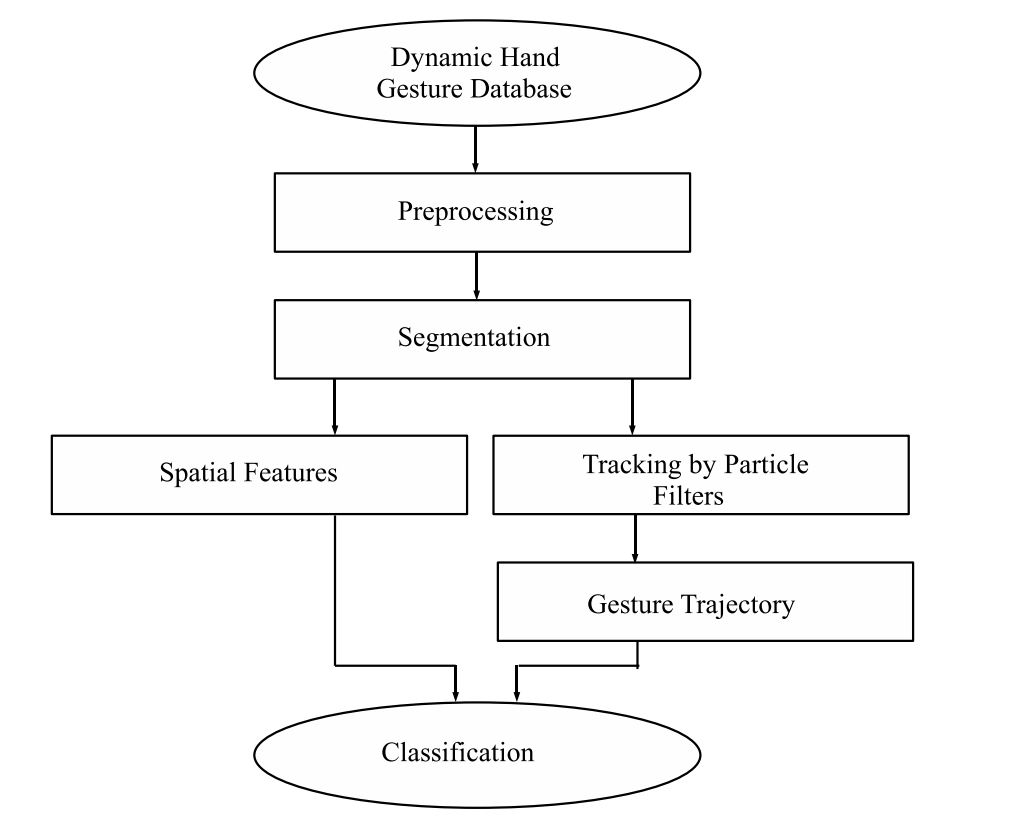
Feature Extraction

Feature extraction is a crucial module in any Computer vision (CV) system. The implementation of this module has a considerable effect on the robustness of the system. For recognizing global hand motions, in addition to the spatial information of the hand, temporal features are also necessary.

Tracking is used to obtain these motion features of a dynamic hand gesture video. It is the process of estimating the trajectory of hand, as the hand moves in the area of interest. The tracking algorithm has to estimate the state of the system at any time instant, given a set of observations.

Spatial feature vectors and features obtained from gesture trajectory are given as input to the neural network classifier to classify different dynamic gestures. Radial basis function (RBFs) neural network is used for classification.





# 6. EXPERIMENTAL SETUP

1. Setup the Vs code installed in computer.

Visual Studio Code is a lightweight but powerful source code editor which runs on your desktop and is available for Windows, macOS and Linux. It comes with built-in support for JavaScript, TypeScript and Node.js and has a rich ecosystem of extensions for other languages and runtimes (such as C++, C#, Java, Python, PHP, Go, .NET).

1. Check whether the Python is installed in Vs code.

The Python extension will automatically install the Pylance, Jupyter and isort extensions to give you the best experience when working with Python files and Jupyter notebooks.

1. Setup the Open CV library in VS code.

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today’s systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

1. Tensorflow

TensorFlow is an open-source software library. TensorFlow was originally developed by researchers and engineers working on the Google Brain Team within Google’s Machine Intelligence research organization for the purposes of conducting machine learning and deep neural networks research, but the system is general enough to be applicable in a wide variety of other domains as well

# 7. CONCLUSION

The Dynamic Hand Gesture Keyboard and Mouse using neural network is a promising technology that has the potential to revolutionize the way we interact with computers. The system uses a neural network to recognize hand gestures and translate them into computer commands, allowing users to control their computers in a more intuitive and natural way.

The use of machine learning algorithms and neural networks in this technology allows for a high level of accuracy in recognizing hand gestures, making it a reliable and efficient system. The dynamic nature of the system also makes it adaptable to different users with varying hand sizes and shapes.

While the technology is still in its early stages of development, it has the potential to greatly improve accessibility for individuals with physical disabilities and enhance the overall user experience for all computer users.

Overall, the Dynamic Hand Gesture Keyboard and Mouse using neural network is a promising technology that has the potential to transform the way we interact with computers, and further research and development in this area could lead to even more advanced and innovative applications.

## 8. TENTATIVE CHAPTER PLAN FOR THE PROPOSED WORK

**CHAPTER 1: INTRODUCTION**

This section contains introduction about Hand Gesture Recognition and its types that is Dynamic Hand Gesture Recognition and Static Hand Gesture Recognition.

**CHAPTER 2: LITERATURE REVIEW**

This section contains details about the Research Paper that were taken as references.

**CHAPTER 3: OBJECTIVE**

This section contains objective of the project.

**CHAPTER 4: METHODOLOGIES**

This section contains methodology that is used in the project.

**CHAPTER 5: EXPERIMENTAL SETUP**

This section contains information about the modules required in the setup of the environment as per project need.

**CHAPTER 6: CONCLUSION AND FUTURE SCOPE**

This section contains the conclusion of the project.

## REFERENCES

1. <https://www.nature.com/articles/s41598-022-08133-z>
2. <https://www.semanticscholar.org/paper/Human-computer-interface-using-hand-gesture-based-Jalab-Omer/1c7e28fc4855f1948026354f12d5623918d37dda>
3. <https://www.youtube.com/watch?v=NJmk1DUyyB8>
4. <https://www.youtube.com/watch?v=L6Dbtqi7SQM>
5. <https://www.youtube.com/watch?v=9zqgvsfbz8M>
6. <https://becominghuman.ai/deep-learning-hand-gesture-recognition-b265f4e6cf02>
7. <https://becominghuman.ai/deep-learning-hand-gesture-recognition-b265f4e6cf02>
8. <https://dl.acm.org/doi/10.1145/3293353.3293421>
9. <https://www.nature.com/articles/s41598-022-08133-z>
10. <https://www.mdpi.com/1424-8220/19/24/5429>