MAGIC HANDS

(HAND RECOGNITION DETECTION)

# PROJECT SYNOPSIS

OF MINOR PROJECT

# BACHELOR OF TECHNOLOGY

COMPUTER SCIENCE ENGINEERING

Submitted By:

Submitted To:

Prof. Jaswant Singh

Nitin Goyal(1905026)

Harmehak Singh(1905103)

Arshdeep Singh(2004692)



**GURU NANAK DEV ENGINEERING COLLEGE,**

**LUDHIANA**

## Table of Contents:

|  |  |
| --- | --- |
| **Content** | **Page no.** |
| 1. **Introduction**    1. **Type of Hand Gestures**   **1.2Programming Language used**  **1.3Python as scripting language** | **3**  **4**  **4**  **4** |
|
| **2.Rationale**  **3.Objective**  **4.Literature review**  **5.Feasibility study** | **4**  **5**  **5**  **6** |
| **6.Methodology**   1. **Facilities required for proposed work** | **6**  **8** |
| **8.Expected outcome**  **9.References** | **8**  **8** |

1. **INTRODUCTION**:

The minor project “Magic hands” is based upon the idea that users can use simple gestures to simply guess the emotions via hands and communicate with each other using hand gestures or sign language. Humans are able to recognize body and sign language easily. Gesture recognition is a topic in computer science and language technology with the goal of interpreting human [gestures](https://en.wikipedia.org/wiki/Gesture) via mathematical algorithms. It is a sub-discipline of [computer vision](https://en.wikipedia.org/wiki/Computer_vision).

Gestures can originate from any bodily motion or state but commonly originate from the face or hand. Gestures are nothing but the nonverbal means of communication which includes bodily actions to convey some messages, these messages can be used as commands to the system. Hand gestures provide a separate complementary modality to speech for expressing ones ideas. Information associated with hand gestures in a conversation is degree, discourse structure, spatial and temporal structure.

The automated recognition of hand gestures may have applications in various area such as in design, robotics, virtual reality and most importantly in sign language. The key problem is how to make a computer able to understand the hand gestures. Machine learning is used to solve various real-time problems. It is commonly used for classification, detection, recognition, and predictions problems.

The Vision-Based gesture recognition system uses a camera for capturing the gesture in the image or video format. The vision-based approach considers properties such as texture and color and the limitations of this approach are lightning, location, image noise.

The key problem in gesture interaction is how to make hand gestures understood by computers. The approaches present can be mainly divided into “Data-Glove based” and “Vision Based” approaches.

The Vision Based methods require only a camera, thus realizing a natural interaction between humans and computers without the use of any extra devices.

We choose this type of work for doing better with development cycle, development period, deep learning and neural networks, computer vision, etc. Our end product is a software that will use the camera either of our laptops to identify our hand gestures.

## Type of Hand Gestures

The gaming industry is divided into two parts: a) Static Gestures b) Dynamic Gestures

* + 1. Static Gestures: Static gestures are postures that do not consider movements into account, e.g. thumbs up. Static gestures are those that only require the processing of a single image at the input of the classifier, the advantage of this approach is the lower computational cost.
    2. Dynamic Gestures: Dynamic gestures are postures that take into consideration the angles between fingers during certain start and end times, e.g. drawing letters in the air. A dynamic hand gesture is a spatial-temporal pattern and has four basic features: velocity, shape, location, and orientation. The motion of the hand can be described as a temporal sequence of points with respect to the hand centroid of the person performing the gesture.

**1.2Technology Used:**

* Python
* Mediapipe Holistic
* Neural Networks
* Keras
* Long-Short Term Memory (LSTM)

## 1.3Python as scripting languages

* Ease to use
* Runs faster and complete task efficiently.
* Easy to learn
* Suitable for automation tasks

## 2.Rationale:

It becomes very difficult for the people to interact with each other who can’t either speak or hear. oreover, it becomes difficult for them to interact with people in the normal world. The purpose of Sign Language Recognition (SLR) systems is to provide an efficient and accurate way to recognise sign language has aids for the hearing impaired for example, or enabling very young children to interact with computers (recognizing sign language), among others.

## 3.OBJECTIVES

* 1. To extract Holistic Keypoints from hands.
  2. To train an LSTM model and maximum number of gestures.
  3. To change computer response corresponding to different gestures.

## Literature Review:

[1]Hand Gesture Detection and Recognition System by Muhammad Inayat Ullah Khan suggested Hand gesture recognition research is classified in three categories. First “Glove based Analysis” attaching sensor with gloves mechanical or optical to transduces flexion of fingers into electrical signals for hand posture determination and additional sensor for position of the hand. This sensor is usually an acoustic or a magnetic that attached to the glove. Look-up table software toolkit provided for some applications to recognize hand posture.

The second approach is “Vision based Analysis” that human beings get information from their surroundings, and this is probably most difficult approach to employ in satisfactory way. Many different implementations have been tested so far. One is to deploy 3-D model for the human hand. Several cameras attached to this model to determine parameters corresponding for matching images of the hand, palm orientation and joint angles to perform hand gesture classification. Lee and Kunii developed a hand gesture analysis system based on a three-dimensional hand skeleton model with 27 degrees of freedom. They incorporated five major constraints based on the human hand kinematics to reduce the model parameter space search. To simplify the model matching, specially marked gloves were used.

The Third implementation is “Analysis of drawing gesture” use stylus as an input device. This drawing analysis led to recognition of written text. Mechanical sensing work has used for hand gesture recognition at vast level for direct and virtual environment manipulation. Mechanically sensing hand posture has many problems like electromagnetic noise, reliability and accuracy. By visual sensing gesture interaction can be made potentially practical but it is most difficult problem for machines.

[2] Gesture Recognition using CNN and RNN by Rajalakshmi J, Kumar P suggested that there are various machine learning algorithms that are used for the recognition of hand gestures and few of them are discussed as follows: -

a) A real-time hand gesture recognition system with four modules such as Data acquisition, Pre-processing, Feature extraction, Gesture recognition, and the Real-time hand detection is done using Histogram of Oriented Gradients (HOG) feature in MATLAB and the k-Nearest Neighbor (KNN) algorithm is used to classify the input images.

b) A technique which commands computer using six static and eight dynamic hand gestures is proposed with three main steps hand shape recognition, tracing of detected hand (if dynamic), and converting the data into the required command, the system uses VGG16 a CNN architecture is used as the pretrained model to recognize the gestures.

[3] Hazim, Nawaf & Al-Dabbagh, Sinan Sameer Mahmood & Esam Matti, Wael. (2016) suggested face recognition has gained a great deal of popularity because of the wide range of applications such as in entertainment, smart cards, information security, law enforcement, and surveillance. It is a relevant subject in pattern recognition, computer vision, and image processing.

Face recognition is a challenge, given the certain variability in information because of random variation across different people, including systematic variations from various factors such as lightning conditions and pose. Computational methods of face recognition need to address numerous challenges. These types of difficulties appear because faces need to be represented in such a way that best utilizes the available face information to define a specific face from all the other faces in the database. Face pose is a specifically difficult problem in this aspect simply because all faces seem similar; specifically, all faces consist of two eyes, mouth, nose, and other features that are in the same location.

Regardless of the method used, the most important concern in face recognition is dimensionality. Suitable methods are needed to reduce the dimension of the studied space. Working on higher dimension cases overfitting, where the system starts to memorize. Computational complexity is also an important problem when working on large databases.

## FEASIBILITY STUDY:

In this project, we will look into all the possibilities and analyze whether it is profitable to work on it or not: -

1. **Economical:** The cost of hardware and software are normal, and it can be performed on any operating system.
2. **Efficient:** If we will deploy the software on playstore, it is going to provide the desired results.
3. **Operationally & Accessibility:** The software is easy to understand and operate and can be understood and used by people of any age groups to detect hand gestures. Anyone can easily download and access the software and can also provide review and feedback.

## METHODOLOGY

**1. Idea:** This project is based on the Hand Gesture Detection System which detects the Sign Language using techniques like Deep Learning, Image Processing, Machine Learning, Artificial Intelligence, etc.This project uses hand gesture recognition which recognises our hand gestures and displays the corresponding output on the screen.

**2. Set Camera:** The camera is the only hardware component required to capture live video feed of class.

**3.Detect Face,Hand and Pose Landmarks:** We have detected the face, hand and Pose Landmarks using OpenCv and Mediapipe Holistic.

**4. Extract Keypoints:** We extracted the different face, hand and pose landmarks according to their different x,y,z co-ordinates and then stored them in a numpy array.

**5.Setup Folders For Data Collection:** We created folders so that the numpy arrays on our device can be stored in different folders corresponding to the different training sample of our dataset.

**6.Collect keypoint sequences:** Different key point sequences were collected and were stored in the folders made in the previous step.

**7.Preprocess Data and Create Labels:** Data pre-processing states the manual or automated sorting of the data to remove any irrelevant images and to have all the images with same dimensions .Thus it becomes important if the image contains other irrelevant information then maybe the system detects something else from that image and it will decrease the overall accuracy of the system.

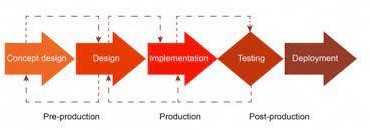
**8. Build and Train a LSTM Deep Learning Model**: Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning (DL). LSTM has feedback connections. It can process not only single data points (such as images), but also entire sequences of data (such as speech or video). LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series. We trained our dataset using this model in order to test and make sign language predictions.

**9. Make Sign Language Predictions:** After training and testing of the LSTM Model different signs of the hands are recognised by the computers. We have used 3 signs “Thank You”, “Love You”, “Hello”. Data Collection Data Preprocessing Code Generation Technical Requirement Analysis Programming and Development Testing and Deployment

**10. Save Model Weights:** We assigned weights of our input against every neuron in LSTM. Each component has weights associated with all of its input from the previous layer. Weights associated with hidden state are called kernel weights and weights associated with input are called recurrent kernel weights.

**11. Evaluation Using a Confusion Matrix:** We have used the confusion matrix to summarize the performance of our classification model. The number of correct and incorrect predictions are summarized with count values.

**12. Test in Real Time:** It is really important to test the model to check for the accuracy of the model. Therefore it is important to have test dataset just dedicated to test the accuracy. The testing and training datasets are made using some functions made in python only dedicated for the task. The main ratio followed for training and testing is generally 10% testing dataset and 90% training dataset and it can vary according to size of the dataset. There are number of researches made just for the efficient testing of the systems and different mathematical formulas used to test accuracy of the systems



## Facilities Required For Proposed Work:

* + - Pycharm
    - GoogleColab
    - WebCam

## EXPECTED OUTCOME

* It will be able to detect hand gestures
* It will also be able to detect sign language

## 9.REFERENCES:

* 1. <https://www.ijrte.org/wp-content/uploads/papers/v9i2/B3417079220.pdf>
  2. <https://www.diva-portal.org/smash/get/diva2:519237/FULLTEXT01.pdf>
  3. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6164357/