Hand gesture recognition technique steps vary from simple to complex applications. Generally, the steps are usually divided as the following:

1. **Hand gesture frame acquisition**
2. **Hand tracking**
3. **Feature extraction**
4. **Classification**

Modules

Data Acquisition Module -. In this phase the image is captured using dataset. Images are saved in a folder and divided the images into 2 sections like testing and training. In our project data set we build a new dataset of 600 images of 80% training section and 20% testing section.

Data Preprocessing Module - The step of data preprocessing plays a very important role in contributing to the accuracy of any training model. In our project, we mainly perform,

- Segmentation – it is the process of separating the data into distinct groups. Image segmentation involves converting an image into a collection of regions of pixels that are represented by a labeled image. By dividing an image into segments, we can process only the important segments of the image instead of processing the entire image.

- Annotation - Data annotation is the process of labeling the data available in various formats like text, video or images. For supervised machine learning labeled data sets are required, so that machine can easily and clearly understand the input patterns. Data annotation encompasses the text, images and videos to annotate or label the content of object of interest in the images while ensuring the accuracy to make sure it can be recognized by the machine.

Classification Module - This module identifies and classifies the data given. The main aim of the proposed work is to enrich the true meaning of gesture by Faster R-CNN & Resnet50. Resnet 50 is stands like back bone of this network. Here the captured image can determine the true meaning of that particular image and display it. Also, it can classify the different type of same label images using Faster R-CNN method.

In the first stage, the images are undergone through preprocessing stage. Our application takes video inputs of the dance form and gives out the corresponding meaning of the hand mudra. CNN processes the input images and classify into different categories. The second step is feeding the preprocessed images to the CNN. The CNN model will learn from the features that have been extracted from the input images. Out of total 600 images,80% taken as training section and 20% taken as testing section. By using fast-RCNN testing section is faster than other methods. And accuracy is greater than others. The comparison among many classification methods shows that Faster R-CNN shows best performance.