We are going to use **2D CNN** and **LSTM** (long short term memory) for our BE project.

system will take video as an input. Then it will generate the optical flow image of every frame.

The optical flow image and its corresponding RGB image will be given to CNN as input for extraction of feature vector of that image.

Then we will be passing these features vector through LSTM.

LSTM will then classify the gesture from given feature vectors.

After classification we can assign that gesture to particular operation.