National Institute of Technology Rourkela

Progress Seminar

Seminar Title	: Enhanced Dynamic Hand Gesture Recognition with 3DCNN-LSTM Fusion and Sliding Window Technique
Speaker	: Vikas Bhatt (Rollno : 521cs1002)
Supervisor	: Prof. Ratnakar Dash
Venue	: New Conference Hall, CSE Department
Date and Time	: 15 Apr 2024 (17:00 Hrs)

Abstract : Dynamic hand gesture recognition has attracted increasing research interests in computer vision, pattern recognition, and human-computer interaction. Static and dynamic hand gestures used for HCI in real-time systems are the area of vital analysis, with various feasible applications. Vision-based hand gesture recognition using a webcam has many advantages over hardware devices like data gloves. This report discusses methods based on static and dynamic gesture recognition. Some depth images and standard datasets are implemented using a convolutional neural network. Although a wide range of recognition systems based on gestures is available, they are expensive and complex. Some existing gesture recognition methods are intended for static or dynamic gestures. Detecting a gesture in real-time is a tedious task considering the time constraint in which the gesture should be correctly recognized. This research introduces an innovative method for recognising hand gestures in real-time. It combines a 3D Convolutional Neural Network (3D-CNN) with Long Short-Term Memory (LSTM) networks, emphasising single-time activations. The intended design is to tackle the difficulties given by the unknown time and geographic variability of when gestures begin and end inside continuous video feeds. The proposed model effectively extracts features and performs classification by utilising the spatial-temporal information recorded by the 3D-CNN and the sequence modelling capabilities of LSTM. Hand gesture recognition is beneficial for individuals who are deaf and mute who communicate using sign language. To address this, we have developed a Convolutional Neural Network (CNN) model employing mediapipe and a convexity method specifically for American Sign Language (ASL).