INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
SHORT ABSTRACT OF THESIS

Name of the Student	:	H Pallab Jyoti Dutta
Roll Number	:	186102008
Programme of Study	:	Ph.D.

Thesis Title: Hand Gesture Detection and Recognition for Gesture-based Patient Rehabilitation and Assistance Systems.

Name of Thesis Supervisor(s)	:	Prof. M. K. Bhuyan
Thesis Submitted to the Department/ Center	:	EEE
Date of completion of Thesis Viva-Voce Exam	:	18/12/2024
Key words for description of Thesis Work	:	Hand Gesture, Hand Segmentation, Hand Gesture Recognition, Hand Gesture Detection, Hand Keypoint Detection, Hand Gesture-based Interfaces.

SHORT ABSTRACT

Hand gestures serve as a natural and widely used means of human interaction, playing a crucial role in establishing seamless human-computer interactions. However, to facilitate effortless interactions, precise decoding of the hand gestures is essential, which is hindered by background clutter, variations in illumination, the presence of skin areas, such as hands or faces in the vicinity, occlusion, and variable hand shapes and sizes. Much work has been done in the literature to address these issues and recognize the gestures, but the generalization is yet to be achieved. This dissertation aims to address these concerns by developing a robust method for hand gesture recognition and applying it to create interfaces that enable human-computer interaction tailored to specific human needs. A method is proposed to segment the hand region in an image that removes the irrelevant information from the background. For this, two segmentation models were proposed; one model utilizes spatial and channel attention and the other benefits from combining a convolution neural network and a linearized transformer unit. Moreover, a novel loss function optimizes the models to resolve class imbalance, ensure boundary smoothness, and retain the hand's shape. These segmented results were further utilized to obtain hand gesture recognition results in a two-stage arrangement. A novel adaptive kernel channel attention layer assists the recognition network in achieving accurate results. The recognition accuracy for two benchmark datasets was 93.8% and 98.0%, which highlights the preciseness of the proposed approach. This two-stage approach is not very suitable for online applications. Therefore, three hand detection methods that localized the hand region and gesture class con currently were proposed. The first method is an anchor boxes-based RetinaNet CBAM hand gesture detection model. The second and third methods are anchor less and detect hand gestures using a detection transformer and CenterNet-based model, respectively. The best-performing model, i.e., the second method, achieved a recognition rate of 89.6% and 100% for two benchmark datasets. Once a robust detection model is available, it can be used to model gesture-operated interfaces for specific tasks. Hand keypoint detection also plays an important role in these interfaces, and hence, a robust keypoint detection model with a multiscale attention block is proposed. In this work, two interfaces were designed that cater to patients undergoing hand rehabilitation and patients in hospitals communicating with medical staff. The proposed methods underwent thorough qualitative and quantitative analysis, revealing state-of-the-art performance even under challenging conditions. The seam less integration of the hand detection algorithm into the interfaces was also success fully accomplished. Patients using the rehabilitation interface reported noticeable improvements in hand functioning, while those utilizing the communication inter face experienced smooth and efficient communication with medical staff. These outcomes underscore the effectiveness of the proposed methods, demonstrating their practical applicability in real-world scenarios.