National Institute of Technology Rourkela

Synopsis Seminar

Seminar Title	: Development of Multi-scale Deep Learning Models for Facial Expression Recognition
Speaker	: Chakrapani Ghadai (Rollno : 520ee7014)
Supervisor	: Dipti Patra
Venue	: EE-401, Seminar Hall
Date and Time	: 28 Jul 2025 (5.15 PM)
Abstract	: Deep learning based methods have been widely used in facial expression recognition (FER) due to their significantly higher performance than traditional methods. FER is challenging due to three main challenges associated with it. Firstly, large intra-class variations and subtle inter-class visual differences which can be largely varied by very small variations in pose, occultation or illumination, etc. Secondly, it is complex to identify features at a granular level to represent minute discriminating dynamic facial region parts (e.g., nose, eyes, mouth, lips, etc.). Finally, the FER datasets are small, given the complexity involved in the problem. Conventional CNN with increased depth has better generalization capabilities, but it can lead to overfitting when used in such problems. The objective of this thesis is to develop efficient deep learning frameworks to overcome the above challenges in facial expression recognition (FER). In this thesis, we explore both traditional and recent state-of-the-art methods and present four contributions. The thesis&rsquos contribution begins with a multi-branch multi-scale attention network (MBMSA-Net) that uses different kernel sizes in each parallel path to extract rich features at different receptive fields. The second contribution presents a parallel structured multi-scale attention network (MSA-Net). Each parallel branch in MSA-Net utilizes channel complementary multi-scale blocks to broaden the effective receptive field and capture features having diversity. The third contribution presents a novel method for facial expression recognition using the proposed feature complementation and multi-scale attention model with attention fusion (FCMSA-AF). The fourth contribution introduces a parallel-structured Multi-Path Multi-Scale Attention Network with Element Recalibration (MPMSA-ER) for facial expression recognition.