National Institute of Technology Rourkela

Progress 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 (New Seminar Room)
Date and Time	: 21 Mar 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 work is to develop efficient deep learning frameworks to overcome the above challenges in facial expression recognition (FER). In this report, we will explore both traditional and recent state-of-the-art methods and present three contributions. In the first contribution, we present a multi-path multi-scale attention network (MPMA-Net) that uses different kernel sizes in each parallel path to extract rich features at different receptive fields. As a deeper and wider network, PMMA-Net can extract sufficiently diverse attention-enhanced multi-scale features from different parallel paths, which can reduce the susceptibility of intra-class and inter-class variations due to external factors. 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. Additionally, attention networks are employed to emphasize important regions and boost the discriminative capability of the multi-scale features. 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). FCMSA-AF consists of parallel structured two-branch multi-scale attention module (MSA), feature complementing module (FCM), and attention fusion and classification module. The MSA of FCMSA-AF contains multi-scale block and lightweight attention modules in a cascaded fashion in two paths to learn diverse features. The FCM uses the correlation between the feature maps in two paths to make the multi-scale attention features complementary to each other. The experimental results verify the effectiveness of the suggested models among the state of the art methods.