

Deep Hierarchical Attention Fusion Network for Robust Facial Emotion Recognition in Unconstrained Environments

Shruthi S

Research Scholar, Department of CSE, Sri Siddhartha Institute of Technology, SSAHE, Tumakuru

Assistant Professor, Department of AIML, Nagarjuna College of Engineering and Technology, Bengaluru, Karnataka

Manjula Y

Research Supervisor, Sri Siddhartha Institute of Technology, SSAHE, Tumakuru

Abstract:

Facial emotion recognition remains challenging due to variations in illumination, pose, occlusion, and individual expression intensity. This paper introduces a novel Deep Hierarchical Attention Fusion Network (DHAF-Net) that integrates spatial attention mechanisms with temporal consistency modeling for enhanced emotion classification. Our approach employs a multi-scale feature extraction pipeline combined with adaptive attention weighting to capture both subtle micro-expressions and prominent emotional cues. The proposed architecture incorporates a novel Cross-Modal Attention Fusion (CMAF) module that synthesizes geometric facial landmarks with deep convolutional features. Experimental evaluation on FER2013, CK+, AffectNet, and RAF-DB datasets demonstrates that DHAF-Net achieves superior performance with accuracy improvements of 4.2% to 7.8% over state-of-the-art methods. Specifically, our model attains 94.6% accuracy on CK+, 89.3% on RAF-DB, and 85.7% on AffectNet, while maintaining computational efficiency suitable for real-time applications. The results indicate that hierarchical attention mechanisms significantly enhance the model's ability to disambiguate complex emotional states in challenging real-world conditions.

Keywords:

Emotion recognition, facial expression analysis, attention mechanisms, deep learning, convolutional neural networks, affective computing.