

ABSTRACT

Inadequate feature extraction remains a major challenge in facial expression classification using traditional methods. This study aims to improve the accuracy of facial expression recognition by applying a deep learning-based approach that specifically focuses on key facial regions. The proposed method combines a Siamese Neural Network (SNN) to measure similarity between expressions and a multi-level feature extraction (HFT-CNN) for detailed feature extraction from three main facial areas: the entire face, the eyes and eyebrows, and the mouth. These three channels are integrated to form a richer and more informative feature representation. The implementation results show that the HFT-CNN architecture achieves an accuracy of up to 99%, while the SNN Triplet model records an accuracy of 91%. However, the prediction performance of both models remains inconsistent across different testing conditions, indicating limitations in their generalization ability. Moreover, the data collection and preprocessing stages also significantly influence model performance, requiring manual selection to ensure the quality and relevance of the training data.

Keywords: *Convolutional neural networks, facial expressions, machine learning, Siamese networks.*