ABSTRACT

Facial expression classification is an important problem in computer vision that supports the development of adaptive Human-Computer Interaction systems. However, challenges such as variations in lighting, pose, and data imbalance still affect model performance. This study proposes a facial expression classification model based on MobileNetV3 with the integration of the Simple Attention Module (SimAM), a cosine classifier, and data augmentation techniques, namely MixUp and CutMix, on the FERPlus dataset. The proposed model is evaluated through eight experimental scenarios, ranging from the baseline model to the combination of all proposed components. The methodology includes data preprocessing, augmentation, model training using the AdamW optimizer and Cosine Annealing learning rate scheduler, and evaluation using accuracy, precision, recall, F1-score, as well as macro and weighted averages. The results show that the proposed combination improves feature representation quality and model generalization capability. The best model is achieved by combining MobileNetV3 with SimAM and applying MixUp and CutMix, resulting in a validation accuracy of 88.57% and a test accuracy of 88.11%, along with a weighted precision of 88.03%, weighted recall of 88.11%, and a weighted F1-score of 0.88. These results indicate an improvement compared to several previous studies, although performance on minority classes remains a limitation. Therefore, the proposed approach is effective in improving facial expression classification performance on the FERPlus dataset.

Keywords: Facial expression classification, MobileNetV3, SimAM, Cosine classifier, MixUp, CutMix