

ABSTRACT

The research on facial expression recognition tasks has far-reaching significance for improving the understanding of human behavior. The development of facial expression recognition has been accompanied by the emergence of many technologies, especially the use of deep learning techniques. Complex features in images are given unique feature extraction capabilities and spatial hierarchy by automatic learning through multi-layer convolution operations, making it one of the most commonly used techniques for FER and other tasks. However, previous studies on deep learning models used relatively complex data sets, low generality of model architecture, no explicit comprehensive comparative study on feature extraction technology of attention mechanism, and limited selection of hyperparameter optimization. Therefore, this study provides a thorough description of facial expression recognition task and deep learning model. Different deep learning models VGG19, ResNet50 and InceptionV3 are compared and analyzed. Convolutional Block Attention Module is introduced to extract features for comparative analysis. Using Generative Adversarial Network to enhance the FER-2013 data set before model input, the imbalance of the data set is solved, and better generalization performance is provided for the training model. The experimental results show that before adding Convolutional Block Attention Module, the VGG19 model has the highest accuracy of 71.70%, while ResNet50 and Inception V3 have 71.5% and 70.40% accuracy, respectively. With the addition of Convolutional Block Attention Module, ResNet50 has the highest accuracy of 72.4%, VGG19 and Inception V3 have an accuracy of 71.9% and 71.1% respectively. Through 5×2 cross-validation and t-tests before and after hyperparameter tuning, the performance of the three models was significantly improved after the addition of Convolutional Block Attention Module. After hyperparameter tuning, the six models showed significant differences compared with that before tuning, and the performance with Convolutional Block Attention Module of ResNet50 was significantly different from that of VGG and Inception V3. In this study, it is concluded that the deep learning model can be used to obtain higher model performance. After the introduction of the attention mechanism, Convolutional Block Attention Module can make the deep learning model perform better, and the tuning of hyperparameters can significantly improve the model performance.