

In-The-Wild deepfake detection using adaptable CNN models with visual class activation mapping for improved accuracy

Saealal, Muhammad Salihin^a; Ibrahim, Mohd Zamri^b; Shapiai, Mohd Ibrahim^c; Fadilah, Norasyikin^b

^a Universiti Teknikal Malaysia Melaka, Faculty of Electric and Electronic Engineering Technology, Melaka, Malaysia

^b Universiti Malaysia Pahang, Faculty of Electrical and Electronics Engineering Technology, Pahang, Pekan, Malaysia

^c Malaysia-Japan International Institute of Technology, Utm, Centre for Artificial Intelligence and Robotics, Kuala Lumpur, Malaysia

ABSTRACT

Deepfake technology has become increasingly sophisticated in recent years, making detecting fake images and videos challenging. This paper investigates the performance of adaptable convolutional neural network (CNN) models for detecting Deepfakes. In-the-wild OpenForensics dataset was used to evaluate four different CNN models (DenseNet121, ResNet18, SqueezeNet, and VGG11) at different batch sizes and with various performance metrics. Results show that the adapted VGG11 model with a batch size of 32 achieved the highest accuracy of 94.46% in detecting Deepfakes, outperforming the other models, with DenseNet121 as the second-best performer achieving an accuracy of 93.89% with the same batch size. Grad-CAM techniques are utilized to visualize the decision-making process within the models, aiding in understanding the Deepfake classification process. These findings provide valuable insights into the performance of different deep learning models and can guide the selection of an appropriate model for a specific application.

KEYWORDS

Batch size; Convolution neural network; Deep learning; Deepfake; Grad-CAM visualization

ACKNOWLEDGEMENTS

This research is financially supported by the Fundamental Research Grant Scheme (FRGS/1/2021/ICT07/UMP/02/1) with the RDU number RDU210136 which is awarded by the Ministry of Higher Education (MOHE) via the Research and Innovation Department, Universiti Malaysia Pahang (UMP) Malaysia.