A Deep Learning Approach Utilizing Convolutional Neural Networks for the Detection of Generated Human Face Images

W.G.M.G.N. Wanigasekara*, T. Jeyamugan and S. Sobana
Department of Physical Science, Faculty of Applied Science, University of Vavuniya, Sri Lanka

Abstract: The proliferation of Artificial intelligence-generated human face images has posed significant challenges to the detection and authentication of digital content, particularly within social networks. These synthetic faces, produced by advanced Artificial Intelligence techniques such as Generative Adversarial Networks (GANs), are often indistinguishable from real faces, complicating traditional detection methods. The ability to generate realistic facial images has broad implications in media, security, digital forensics, and identity verification. Traditional detection methods, which rely on manually created features and rule-based algorithms, are becoming ineffective against advanced generated images, as they struggle to detect subtle patterns that differentiate them from real ones. The urgency of developing more robust detection methods is driven by the risks posed by synthetic images, including misinformation, identity fraud, and the erosion of content integrity. In the literature, researchers used one image generator to generate the images for training and used that same generator for testing and validating the model. Most of the images were in JPG format with the same quality for all the images. This research aims to develop an advanced detection methodology, using deep learning and convolutional neural networks, to differentiate between real and generated face images that can adapt to different generators from ProGAN, StyleGAN, and Stable Diffusion, with resolutions ranging from 128x128 to 768x768 pixels, and both PNG and JPG formats. The study introduces a customized model named "CCNNgenFace," designed for detecting generated face images. The model achieved an accuracy of 89.45% on the test dataset, demonstrating its ability to generalize across different generated images and formats. This approach enhances detection accuracy, helping safeguard digital media authenticity and preventing the misuse of synthetic facial images.

Keywords:Artificial intelligence-generated, CCNNgenFace, Convolutional Neutral Network, deep learning, Generative Adversarial Networks, Synthetic faces

*Corresponding author's email:
gayeshiwanigasekara2018@gmail.com