Fake Image Detection Using Deep Learning

Raidah S. Khudeyer, Noor M. Al-Moosawi
College of Computer Science & Information Technology, University of Basrah, Iraq
E-mail: raidah.khudayer@uobasrah.edu.iq, almoosawinoor2@gmail.com

Keywords: fake detection, CNN, efficientNetB0, synthetic media, deep learning

Received: March 13, 2023

With the emergence of numerous electronic communication programs and image processing programs, as well as an increase in the number of people who use them with a zeal for publishing everything related to their lives and their special pictures and their fear of those who might use these pictures for malicious or humorous purposes, it has become necessary to have specialized and precise systems to determine whether a picture is real or fake. Our work aims to detect real and fake faces by using and modifying one of the most efficient CNN architectures, EfficientNetB0, after improving the architecture with additional fully connected layers and efficiently training the model by using the Adam optimizer and a scheduler learning rate technique. Our findings on the well-known 140k-real-and-fake-faces Kaggle dataset showed state-of-the-art accuracy with the lowest error rate. We achieved 99.06% accuracy, and 0.0569 error rate respectively.

Povzetek: Predstavljena je metoda globokih mrež, ki z uporabo EfficientNetB0 in optimizatorja Adam na Kagglovih 140.000 obrrazih ločuje prave in lažne obraze z 99,06% točnostjo.

1 Introduction

A fast-expanding field, synthetic media is media created by technology. Because of this, artificial media may also be referred to as "AI-generated media". Some examples of synthetic media include music composed by AI, text generation, imagery and video, voice synthesis, and fake images.

In recent years more research has excelled and increased attention on CNN in several areas, such as the classification of images, object detection, facial recognition, fingerprint analysis, computer-aided diagnosis, and facial expressions [1], [2].

The appearance of and the rapid advancement in deep learning makes the detection of authentic and manipulated facial images and video clips more difficult, which is called Deep Fake. Subsequently, the need for techniques that can discover the integrity of digital visual content as images, videos, text so on, is important.

Machine-based algorithms are crucial in detecting fake images, which can take many different forms. Modeling these methods as binary classification issues. It acquires hand-crafted features, investigates hidden information, and separates fake images from editing procedures like an improvement (histogram equalization, color alteration, etc.), geometry modifications (rotation, cropping, shearing), and content changes (copy-move, cut-paste, etc.) [3].

End-to-end learning solutions based on Deep Learning (DL), mainly using Deep Convolution Neural Networks (DCNNs), are created to benefit from automated learning and feature extracting [4-8].

The related research in fake detection started using deep learning and the CNN model since 2018, as summarized in Table 1, in [9] the researchers mention methods that are used to fake detection. A standard CNN consists of multiple components such as convolutional layers, pooling layers, and fully-connected layers. It is created to automatically pick up on spatial feature hierarchies by learning them using a backpropagation algorithm. CNN architecture receives an input image after going through several building blocks and can tell the difference between real and fake faces. The hyperparameters that need to be optimized during CNN training include learning rate, batch size, activation layer, regularization method, and optimizer. in 2019 proposed an efficient network called Efficient Net [10], although the rapid development of convolutional neural network gradually its deficiencies was reason to replace it with pertained models such as Resent, mobile Net, ..., which are required to get more accuracy accordingly the network depth, network width, and input image resolution that need to be manually adjusted [11].

Training a large deep-learning model is a difficult optimization task due to many difficulties including overfitting, underfitting, finishing derivatives, and choosing suitable hyperparameters [12]. The learning rate is one of the most difficult hyperparameters to set and it has a big impact on how well models perform. The traditional neural network training technique can improve performance and speed up training on some issues by utilizing a learning rate that modifies throughout training [13].

In 2018 [3], the researchers used GANs to create fake faces with multiple resolutions and sizes, then used different DCNN models to detect fake images. They apply a deep-face recognition system to transfer weight, and the network is fine-tuned suitable by using real or fake images in the AI Challenge.