# Chapter 3
# Combating Deepfake-Generated Photos and Videos Using Generative Adversarial Network

**B. Aarthi**
https://orcid.org/0000-0001-6485-8848
*SRM Institute of Science and Technology, Ramapuram, India*

**A. Smruthi**
*SRM Institute of Science and Technology, Ramapuram, India*

**Pamireddy Thanishka**
*SRM Institute of Science and Technology, Ramapuram, India*

**G. Sakthi Prasanna**
*SRM Institute of Science and Technology, Ramapuram, India*

**P. Mahendran**
*Dhaanish Ahmed College of Engineering, India*

## ABSTRACT

*Rapid advances in artificial intelligence and machine learning have resulted in the creation of Deep Fakes, which are manipulated films, audio, and images capable of disseminating false information, fake news, and altering sensitive records. The prevalence of deepfake technology has raised significant concerns regarding the veracity of digital content, underscoring the critical need for reliable deepfake facial recognition algorithms. This study embarks on developing an advanced*

DOI: 10.4018/979-8-3373-4672-4.ch003

*deepfake detection system leveraging Generative Adversarial Networks (GANs) within a programming environment. The central focus is to create a neural network that can effectively differentiate between authentic and artificially generated media content. To accomplish this, the system undergoes extensive training using diverse datasets, enabling it to recognize subtle nuances and specific artifacts associated with GAN-generated content.*

## INTRODUCTION

Neural network algorithms are used to modify media content, which is referred to as "deepfake." The term "deepfakes" refers to the creation of synthetic and fake content that is produced by superimposing a person's voice and face on another using a variety of machine learning (ML) and deep learning (DL) techniques to make the fake content sound and look real. This fuels the propagation of false information and hatred in social media networks, distorts the public's perception, and can be used for more nefarious purposes, including identity theft, spoofing, extortion, and character assassination. Recent social media platforms have seen several regrettable instances of deepfakes, raising questions about possible public exploitation of these platforms. To create modified facial photographs that mimic identical motions and movements, deepfake approaches use sophisticated deep learning models such as autoencoders and generative adversarial networks (GANs) to evaluate an individual's facial features and actions. The ability to modify films and photographs has become quite common thanks to technological advancements. If this trend continues, evidence containing photos and videos must be reviewed before being presented in court. Primarily because of advancements in technology, particularly in machine and deep learning, the creation and modification of photos and videos have become mainstream. Fake information is simple, but fighting deepfakes and correcting records are more challenging. Deepfake can only be effectively countered by fully comprehending its underlying principles and technology (Rajest et al., 2024).

**What Are Deepfakes:** An artificial intelligence technique called "deepfake AI" produces realistic-looking photo, audio, and video hoaxes. The word, which combines fake and deep learning, refers to both the technique and the phony information that results from it. Deepfakes typically substitute one person for another in pre-existing source material. In addition, they create entirely original videos featuring actual individuals saying or acting in ways they have never done. Deepfake face manipulations are usually categorized into four primary categories.

- Face generation, which creates completely new facial images