Chapter 10
Demystifying Deepfakes: Understanding Its Technological Landscape

Pooja Dehankar
Ajeenkya D.Y. Patil School of Engineering, India

Kashvi Chaturvedi
https://orcid.org/0009-0005-3192-3400
Ajeenkya D.Y. Patil University, India

Susanta Das
https://orcid.org/0000-0002-9314-3988
Ajeenkya D.Y. Patil University, India

ABSTRACT

Deepfakes are emerging topic at the nexus of multimedia & AI that have experienced a significant surge in recent years. AI is experiencing a technical upsurge with the creation of hyper-realistic video, audio, & images produced by algorithms known as deepfakes. The world was enthralled by the introduction of deepfake, a new method of media manipulation. Deep learning algorithms enable the manipulation & creation of digital information that is incredibly realistic & difficult to distinguish from actual content, leading to the creation of fake media. Although deepfakes can be useful for study, education, & entertainment, they also present a number of serious issues in a number of different fields, including fraud, political manipulation, disinformation, & reputational harm. This chapter examines how deepfakes are created & detected. It investigates ways to counteract deepfake's adversarial purposes. It provide readers a better grasp of the trends & breakthroughs that are currently occurring in this field, & the areas that still need to be researched/addressed.

DOI: 10.4018/979-8-3373-7011-8.ch010

1. INTRODUCTION

Deep learning-based Artificial Intelligence (AI) technologies have advanced significantly over the past few years, improving both the form and quality of content production as well as automatic content generating. With deepfake technology, images, sounds, and movies can be automatically generated, manipulated, and forged to create incredibly realistic and challenging-to-distinguish effects. The public news, democracy, and human society are now seriously threatened by fake information (Borges et al., 2019).

Among the DL-powered apps that have been making an appearance lately is Deepfake. Therefore, the main way that deepfake systems can produce fake visuals is by substituting scenes or videos, images and sounds which are indistinguishable from genuine ones by people. Our fingertips now have the ability to manipulate artificial speech, images, and videos thanks to a variety of technologies. Today, video and image forgeries can be indistinguishable from genuine work with the naked eye. This could result in several problems, from portraying fake evidence in court to influencing public perception. Given these considerations, we need to have some tools in order to determine reality (Heidari et al., 2024). With the evolving technology, it is possible to create forgeries or altered images, which are nearly identical to real images, but undetectable to naked eye. The generative adversarial network (GAN) is most popular and trusted artificial intelligence models for creating fake images. It modifies and reorders pixels in unique ways, making images that have never been created before. Initially intended to produce fake images, these models are becoming more and more popular. Two neural networks, discriminator network and generator network, support GAN approaches; the discriminator is the one that tells the difference between the generated images (from generator) & original images (used for training the discriminators). The generator's objective is to minimize the chance of being wrong, while the discriminator's is to maximize the chance of being right, in a minimax game. Resulting fake images have propagated misinformation online and violated our privacy. Majority of these images are posted on social media in order to help their false validity and create misinformation. This illegal and immoral method can also be used to carry out a lot of frauds (Arora & Soni, 2021).

Deep learning algorithms are causing a rapid revolution in the production of audiovisual material. The term “deepfakes” refers to the term used to refer to artificial audiovisual media produced by deep learning. These media are characterized by their remarkable qualities, including being incredibly easy to create and having the ability to blend in with actual sounds and images captured by sensors. This technological progress has brought up ethical issues that have received a lot of attention (Millière, 2022).