深度偽造技術

文／林一平 講座教授

Nikola Tesla。

Nikola Tesla（1856 ～ 1943）在 1890 年代預言「21 世紀時，機器人將取代古代文明中奴隸勞動所佔據的位置。」

這項預言在當今的人工智慧（AI）技術的發展下似乎正在實現。早期的 AI 技術已經能夠大致準確地分辨狗和貓的圖片，隨著生成式人工智慧（generative AI）的突破性發展，它逐漸深入我們的生活並引領著社會變革。當 AI 技術進入深度偽造（DeepFake）的層次，將會加速我們進入元宇宙世界，實現 Tesla 的預言。

深度偽造是一種透過電腦生成的影片技術，用於創造看似真實的虛假影像。它使用 AI 技術將一個人的臉替換為另一個人的臉，同時匹配微小的面部表情，從說話到皺眉都能保持一致。

這項技術利用深度學習算法和大量訓練數據生成逼真的影片，使觀眾難以區分真實和偽造的影像。製作一個臉部交換的視頻通常需要以下步驟：首先，使用編碼器處理 2 個人數百萬張的照片。編碼器是一個 AI 系統，用於尋找並學習 2 個臉部之間的相似之處，並將這些相似之處簡化為共同的特徵，並壓縮圖像。

然後，使用一個名為解碼器的第二個 AI 系統，從壓縮的照片中恢復出臉部。你訓練一個解碼器來恢復第一個人的臉部，另一個解碼器來恢復第二個人的臉部，因為這兩張臉是不同的。當需要進行臉部交換時，只需將編碼的照片輸入「另一個」解碼器。

例如，將某甲的臉部壓縮圖像輸入已經訓練過某乙的解碼器。然後，解碼器使用某甲的表情和面部定位來重建某乙的臉部。為了製作逼真的影片，這個步驟必須處理每一幀畫面。

現今，訓練某甲與某乙臉部的模型，以及在影片中合併臉部的過程，幾乎可以即時完成。

早期有名的深度偽造例子包括 2 個假影片：美國前總統 Barack Obama 稱呼川普（Donald Trump）為「完全蠢貨」和 Mark Zuckerberg 吹噓對數十億人的被盜數據擁有絕對控制。我們在烏克蘭的戰爭中也見證假影片的應用，以及使用知名人物臉孔的成人內容。

然而，深度偽造技術也可能被用於音頻和圖像，大部分國家禁止未經同意且具有邪惡目的的深度偽造使用。

不過，除了潛在危險性，深度偽造技術在一些有趣且輕鬆的應用中也顯示出潛力。例如，將深度偽造應用於教育領域，可以使課堂更有趣。想像一下，在英文課堂上，可以邀請虛構的名人來講解課程，例如劉德華。雖然使用真實人物的深度偽造可能會被視為非法，但是使用不存在的人物則可以避免法律問題。企業也開發並銷售深度偽造服務，以實現自動化新聞播報，甚至減少演員的參與，節省成本。例如，TikTok 上就有一個深度偽造的阿諾史瓦辛格（Arnold Schwarzenegger），使用俄語講話，省去了他學習俄文的功夫。

深度偽造技術的應用範圍廣泛且多樣，但我們必須謹慎使用，以避免濫用和潛在的負面影響。只有在合法、道德且有創意的方式下，才能充分發揮深度偽造技術的潛力。

林一平

國立陽明交通大學資工系終身講座教授暨華邦電子講座

現為國立陽明交通大學資工系終身講座 教授暨華邦電子講座，曾任科技部次長，為 ACM Fellow、IEEE Fellow、AAAS Fellow 及 IET Fellow。研究興趣為物聯網、行動計算及系統模擬，發展出一套物聯網系統 IoTtalk，廣泛應用於智慧農業、智慧教育、智慧校園等領域 / 場域。興趣多元，喜好藝術、繪畫、寫作，遨遊於科技與人文間自得其樂，著有 < 閃文集 >、< 大橋驟雨 >。

Deepfake Technology

Nikola Tesla (1856–1943) prophesied in the 1890s, "In the twenty-first century, robots will take the place that slave labor occupied in ancient civilization."

This prophecy is being realized through the continuous advancements in artificial intelligence (AI) technology. Initial AI technologies have successfully achieved a reasonably accurate distinction between images of dogs and cats. The breakthroughs and developments in generative artificial intelligence (generative AI) are increasingly infiltrating our daily lives, steering societal transformations. As AI technology delves deeper into deepfakes, it is poised to accelerate our immersion into the metaverse, ultimately bringing Tesla's prophecy to fruition.

Deepfake involves the computer-generated manipulation of videos to create deceptively realistic images. Utilizing AI technology, it seamlessly substitutes one individual's face with another's while maintaining consistency in subtle facial expressions, spanning from talking to frowning.

This technology employs deep learning algorithms and extensive training data to produce realistic videos, challenging viewers to distinguish between authentic and manipulated images. The process of crafting a face-swapping video includes the following stages: Initially, millions of photos featuring two individuals are analyzed using an encoder. This encoder, functioning as an AI system, is specifically designed to recognize and comprehend similarities between the two faces, streamlining these resemblances into shared features, and compressing the images.

Afterward, a second AI system, called a decoder, is employed to restore faces from the compressed images. You train one decoder to restore the face of the first person and another decoder to retrieve the face of the second person since these two faces differ. When face swapping is required, simply input the encoded images into the 'other' decoder.

To illustrate, compressed facial images of person A are fed into a decoder previously trained on person B. Subsequently, the decoder utilizes the facial expressions and features of person A to reconstruct the face of person B. To achieve realistic videos, this process needs to be applied to every frame in the footage.

Currently, training models for the facial features of Person A and Person B, as well as the process of blending faces in videos, can be accomplished almost instantly.

Early well-known examples of deepfakes include two fake videos: one depicting former U.S. President Barack Obama calling Donald Trump 'a total and complete dipshit,' and another featuring Mark Zuckerberg bragging about having complete control over stolen data from billions of individuals. Furthermore, deepfake videos have been observed in the Ukraine conflict, and there has also been the creation of adult content incorporating the faces of well-known celebrities.

Furthermore, deepfake technology can also be applied to manipulate both audio and images. The majority of countries prohibit its unauthorized and malicious use without consent.

Nevertheless, deepfake technology exhibits promise beyond potential risks in various intriguing and playful applications. For instance, integrating deepfakes into the education field could enhance classroom engagement. Envision a scenario in an English class where a fictitious celebrity, such as 'Andy Lau,' is invited to lecture. While the use of deepfakes featuring real individuals may be deemed unlawful, employing non-existent personas can prevent legal complications. Moreover, businesses are involved in developing and selling deepfake services not only for automated news reporting but also to reduce costs by minimizing the involvement of actors. A case in point is a TikTok video featuring a deepfake of Arnold Schwarzenegger speaking in Russian, eliminating the need for him to try learning the Russian language.

While deepfake technology has a broad and varied range of applications, its usage should be approached cautiously to prevent misuse and potential adverse impacts. Employing it in legal, ethical, and creative manners is essential to unlock the complete potential of deepfake technology.

Dr. Jason Yi-Bing Lin

Lifetime Chair Professor of the Department of Computer Science at National Yang Ming Chiao Tung University and Winbond Chair Professor

Dr. Lin is currently a lifetime chair professor of the Department of Computer Science at National Yang Ming Chiao Tung University and Winbond chair professor. He is an ACM Fellow, IEEE Fellow, AAAS Fellow and IET Fellow. His research interests include Internet of Things, mobile computing, and system simulation. He has developed an Internet of Things system called IoTtalk, which is widely used in smart agriculture, smart education, smart campus, and other fields. He has a variety of interests, such as art, painting, and writing, as well as voyaging through science, technology, and humanities.