Mateusz Nowak
Wydział Zarządzania

Deepfake: zabawa czy narzędzie do manipulacji?

Deepfake to technologia pozwalająca na tworzenie sztucznie wygenerowanych materiałów dźwiękowych, jak i wideo, w których występujący mogą przeobrazić się w zupełnie inną osobę. Technologia ta może być wykorzystywana w wielu celach, w tym w rozrywce i edukacji, ale także w celach manipulacji. Są dwie metody tworzenia materiałów deepfake. Pierwsza wykorzystuje dwa algorytmy: enkoder i dekoder. Ten pierwszy odpowiada za wyszukanie wspólnych cech dwóch obrazów, które mają zostać połączone w jeden, by potem dekoder mógł je przenieść na nowy, fałszywy obraz. Drugim sposobem są sieci Generative Adversarial Network (GAN), w której dwa algorytmy współpracują ze sobą. Pierwszy z nich – generator, korzystając z ogólnych informacji tworzy obrazy (np. człowieka) określając kolejne jego cechy wyróżniające (sylwetka, twarz, oczy). Drugi algorytm to dyskryminator, który ocenia czy dostarczone mu przez generator obrazy są prawdziwe lub nie.

Deepfakey szybko stają się popularne jako narzędzie twórcze dla artystów, którzy chcą tworzyć niezwykłe i niewiarygodne treści. Popularne serwisy społecznościowe, takie jak Tiktok, czy Youtube, są pełne filmów wytworzonych przy użyciu technologii deepfake. Najczęściej używany jest w celach humorystycznych, gdzie autor filmu przeobraża się w sławną osobę i wypowiada zdania lub wykonuje czynności, których ta sławna osoba nigdy by nie wypowiedziała.

Zyskuje na tym również branża filmowa, która w ten sposób może wygenerować twarz zmarłego lub starego już aktora w zupełnie nowym filmie. Zabieg ten użyto na przykład w filmie „Rouge One: A Star Wars Story", w którym mogliśmy podziwiać cyfrową wersję Petera Cushinga w roli wielkiego moffa, którego zanimowali pracownicy z Industrial Light & Magic. Największą zaletą tego zabiegu jest to, że do stworzenia w miarę przyzwoitego efektu potrzeba kilkuset zdjęć, co w przypadku celebrytów nie jest zbyt dużym problemem. Wystarczy wykorzystać fotografie dostępne w sieci, aby podmienić twarze aktorów w nagraniu.

Niestety mamy do czynienia również z negatywnym wykorzystywaniem deepfake. Idealnym tego przykładem jest polityka, gdzie fałszywe informacje tego typu mogą powodować rażące konsekwencje. Wyobraźmy sobie deepfake z prezesem Narodowego Banku Polskiego, który po raz kolejny podniósłby stopy procentowe. Wywołałoby to kolosalne skutki na nasz kraj i gospodarkę. Dobrze przygotowany film z wykorzystaniem tej technologii ułatwiłby też manipulowanie opinią społeczną. Postępujący rozwój technologii sprawi zaś, że pojawi się problem z rozróżnieniem materiałów prawdziwych od fałszywych. Stąd już krok do spadku zaufania społecznego i chaosu informacyjnego.

Filmy deepfake są również szeroko rozpowszechnione w branży pornograficznej. Zmiany najczęściej polegają na zastąpieniu twarzy oryginalnej aktorki/aktora na twarze celebrytów, oczywiście bez ich zgody.

Zagrożenie tym oprogramowaniem nie dotyczy tylko elity społeczeństwa. Dotyka ono również nas. Deepfake otwiera wielkie pole do szantażów i wymuszeń. Coraz łatwiej będzie można spreparować film na którym dokonujemy przestępstwa lub znajdujemy się w niedwuznacznej sytuacji. Wzrośnie przez to nieufność społeczna i wiara w przekazywanie

informacji. Jak bowiem można zaufać czemukolwiek, gdy na przykład film, będący materiałem w sprawie przed sądem może być fałszywy?

Zalew tak groźnych treści przyczynił się do zainwestowania przez duże korporacje m.in. „Microsoft" czy „Amazon" w oprogramowania, które w łatwy sposób będzie wykrywało zmanipulowane materiały. Jednak w erze gdzie każdy może stworzyć deepfake, walka z nim jest bardzo trudna. W Polsce nie pojawiły się jeszcze żadne regulacje prawne, które opisywałyby jakie konsekwencje grożą za wykorzystanie oprogramowania w zły sposób.

Narzędzie deepfake jest coraz bardziej popularniejsze i trudniejsze do odróżnienia od prawdy. Technologia ta stale się rozwija, naśladując przy tym coraz lepiej ludzkie gesty, mimikę twarzy, czy głos. Przeglądając treści w internecie nie powinniśmy wierzyć we wszystko co jest tam zamieszczane. Kluczem jest filtrowanie informacji i porównywanie ich w wielu źródłach. W taki sposób możemy w maksymalny sposób zabezpieczyć się przed fałszywymi informacjami stworzonymi poprzez deepfake.