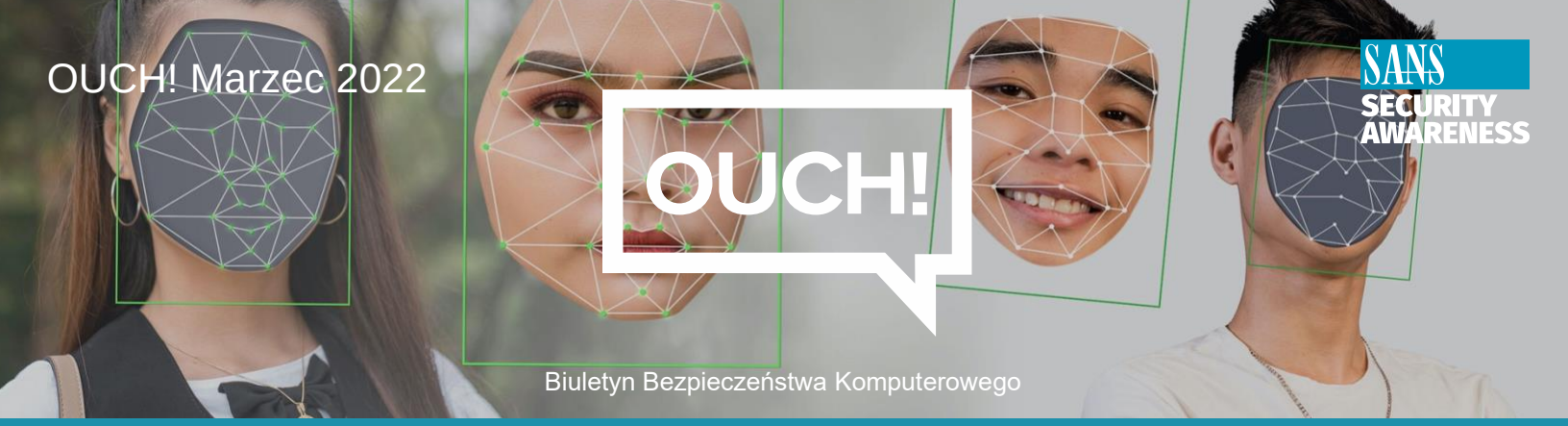


Naucz się nowej umiejętności: wykrywanie Deepfake

Czym jest deepfake?

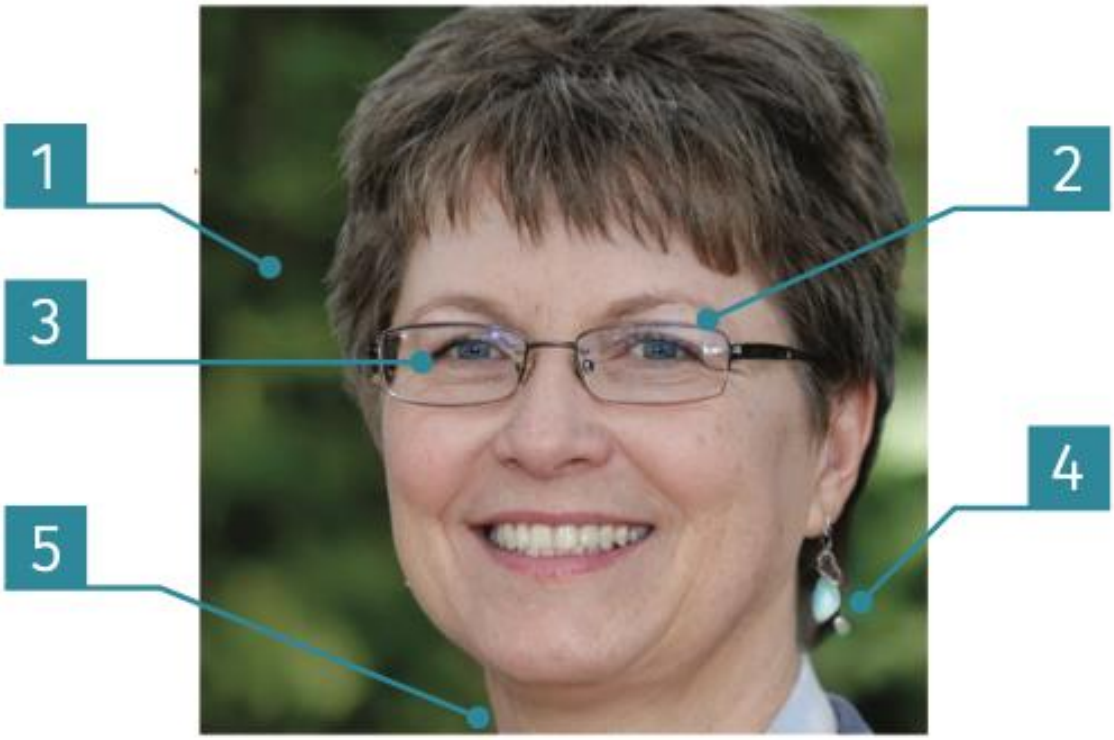
Słowo „deepfake” to połączenie „deep learning” (głębokie uczenie) i „fake”(fałszywy). Deepfake to sfałszowane zdjęcia, filmy lub nagrania dźwiękowe. Czasami postaci w nich są generowane komputerowo, a wyglądają i brzmią tak, jakby mogli być prawdziwymi ludźmi. Zdarza się, że ludzie są prawdziwi, ale ich twarze i głosy są manipulowane, aby robić i mówić rzeczy, których nie zrobili lub nie powiedzieli. Na przykład sfałszowany film wideo może zostać wykorzystany do odtworzenia wizerunku znanej osoby lub polityka mówiącego coś, czego nie powiedział. Korzystając z tych bardzo realistycznych podróbek, napastnicy mogą tworzyć alternatywną rzeczywistość, w której nie zawsze możesz zaufać swoim oczom i uszom.

Niektóre deepfake mają uzasadnione cele, takie jak filmy przywracające życie zmarłym aktorom, aby odtworzyć sławną postać. Jednak cyberprzestępcy zaczynają wykorzystywać potencjał deepfake'ów. Wykorzystują je, aby oszukiwać, dzięki czemu mogą kraść pieniądze, nękać ludzi lub tworzyć fałszywe wiadomości. W niektórych przypadkach tworzone są nawet fałszywe firmy złożone z pracowników, którzy nie istnieją. W świetle tych ataków musisz jeszcze bardziej uważać na to, w co wierzysz, czytając wiadomości lub media społecznościowe.

FBI ostrzega, że w przyszłości deepfake będą miały „bardziej dotkliwy i powszechny wpływ ze względu na stopień zaawansowania użytych mediów”. Naucz się rozpoznawać oznaki deepfake, aby uchronić się przed tymi wiarygodnymi symulacjami. Każda forma deepfake — obraz, wideo i dźwięk — ma swój własny zestaw wad, które mogą go zdradzić.

Nieruchome obrazy

Deepfake, który możesz zobaczyć najczęściej, to fałszywe zdjęcie profilowe w mediach społecznościowych. Poniższy obrazek jest przykładem fałszywego zdjęcia ze strony taosobanieistnieje.com. Poniżej ilustracji znajduje się pięć różnych wskazówek sugerujących, że może to być deepfake. Wskazówki nie są łatwe do zauważenia i mogą być trudne do zidentyfikowania:



1. Tło: tło jest często rozmyte lub krzywe i może mieć niespójne oświetlenie, takie jak wyraźne cienie skierowane w różne strony.
2. Okulary: Przyjrzyj się dokładnie połączeniu między oprawkami a zausznikami w pobliżu skroni. Deepfake często mają niedopasowane połączenia o nieco innych rozmiarach lub kształtach.
3. Oczy: zdjęcia używane obecnie do fałszywych zdjęć profilowych wydają się mieć oczy w tym samym miejscu w kadrze, co powoduje to, co niektórzy nazywają „głębokim spojrzeniem".
4. Biżuteria: Kolczyki mogą być bezkształtne lub dziwnie przymocowane. Naszyjniki mogą być osadzone w skórze.
5. Kołnierze i ramiona: Ramiona mogą być zniekształcone lub niedopasowane. Kołnierzyki koszuli mogą być różne z każdej strony.

Video

Naukowcy z Massachusetts Institute of Technology opracowali listę pytań, która pomoże ustalić, czy film jest prawdziwy, biorąc pod uwagę, że deepfake często nie jest w stanie „w pełni odzwierciedlić naturalnej fizyki" sceny lub oświetlenia.

1. Policzki i czoło: Czy skóra wydaje się zbyt gładka lub zbyt pomarszczona? Czy cechy wskazujące na wiek oczu, włosów i skóry są zgodne między sobą?
2. Oczy i brwi: Czy cienie pojawiają się w miejscach, w których się ich spodziewasz?
3. Okulary: Czy są jakieś odblaski? Czy jest za dużo blasku? Czy kąt padania światła zmienia się, gdy osoba się porusza?
4. Zarost: Czy zarost wygląda na prawdziwy? Deepfakes może dodawać lub usuwać wąsy lub brodę.
5. Pieprzyki na twarzy: Czy pieprzyk wygląda naturalnie?
6. Mruganie: Czy osoba nie mruga za często?
7. Rozmiar i kolor ust: Czy rozmiar i kolor ust pasują do twarzy?

Dźwięk/głos

Naukowcy twierdzą, że technologie takie jak spektrogramy mogą pokazać, kiedy nagrania głosowe są fałszywe. Natomiast większość z nas nie ma możliwości analizowania głosu, gdy dzwoni napastnik. Zwróć uwagę na przekaz, dziwny tonu głosu i wyrażane emocje oraz braku hałasu w tle. Podrabiane głosy mogą być trudne do wykrycia. Jeśli otrzymasz dziwne połączenie z prawdziwej organizacji, możesz sprawdzić, czy połączenie jest prawdziwe- rozłączając się, a następnie oddzwaniając do tej organizacji. Pamiętaj, aby użyć zaufanego numeru telefonu, na przykład numeru telefonu z rachunku lub wyciągu z organizacji lub numer telefonu z oficjalnej witryny internetowej organizacji.

Wnioski

Należy pamiętać, że osoby atakujące aktywnie wykorzystują deepfake. Mogą tworzyć fałszywe konta w mediach społecznościowych, aby łączyć się lub tworzyć fałszywe filmy, aby wpłynąć na opinię publiczną. Niektórzy nawet sprzedają swoje usługi, aby inni napastnicy mogli zrobić to samo. Nie oczekujemy, że zostaniesz ekspertem od deepfake, ale jeśli poznasz podstawy ich identyfikacji, znacznie lepiej będziesz się bronić. Jeśli podejrzewasz, że wykryłeś deepfake, zgłoś to na stronie lub w źródle, które udostępnia treści.

Redaktor gościnny

Kerry Tomlinson (@KerryTNews) jest reporterką wiadomości w Ampere News i certyfikowaną specjalistką SANS Security Awareness Professional. Jej misją jest tłumaczenie tego, co dzieje się w cyfrowym świecie dla ludzi na wszystkich poziomach wiedzy, za pomocą wnikliwych wiadomości i przekonujących prezentacji.

Źródła

Ataki socjotechniczne: https://www.sans.org/newsletters/ouch/social-engineering-attacks/
Czy potrafisz dostrzec oszustwo? (Ampere News): https://www.amperesec.com/news/can-you-spot-the-fake
MIT's deepfake detection test (MIT): https://detectfakes.media.mit.edu/
Znajdź deepfake: https://www.spotdeepfakes.org/en-US

Polski przekład CERT Polska: Bartłomiej Wnuk, Aleksandra Węgrzynowicz