Demaskowanie głębokich fałszywek

Prawdopodobnie myślisz: „Steve Buscemi zwykle nie nosi czerwieni, gdy zakłada coś bez rękawów?" Rzeczywiście, ten obraz nie jest prawdziwy. Pochodzi z wygenerowanego przez komputer filmu wideo znanego jako deepfake (głęboka fałszywka). Ze względu na wzrost mocy obliczeniowej i postępy w uczeniu maszynowym, filmy z głębokimi fałszywkami są teraz, niestety, łatwiejsze do wykonania i trudniejsze do zidentyfikowania. Jednak nie wszystko stracone. Podobnie jak komputery z pomocą ludzkiego kierownictwa, tworzą głębokie fałszywki, tak i w ten sam sposób można je wykryć. Obecne podejścia do problemu identyfikacji fałszywych filmów wykorzystują wiele technik, w tym geometrię (ruchy głowy i warg), algebrę liniową (aby wykryć rozbieżności wynikające z przekształcenia jednej twarzy na drugą) oraz rachunek prawdopodobieństwa (aby zmierzyć prawdopodobieństwo, że film nie jest prawdziwy). Jednak najważniejszą bronią w tej walce z oszustwami może być nie branie wszystkiego za  wartościowe.

Naukowcy pracują obecnie nad bardziej niezawodną, w porównaniu do istniejących rozwiązań, metodą do wykrywania głębokich fałszywek. Metoda ta wykorzystuje bity pliku wideo w celu przypisania do tego pliku matematycznie zaszyfrowanej liczby, która posłuży jako jego podpis cyfrowy. Podpis staje się częścią łańcucha bloków podobnego do tego, który jest używany w walucie cyfrowej do uwierzytelniania transakcji i wykrywania manipulacji. Wszelkie zafałszowanie w pliku wideo zmieni bity oryginalnego pliku, ale nie jego oryginalny podpis tak, że podpis nowego pliku nie będzie pokrywać się z podpisem oryginału. Dzięki tej metodzie sprawdzania poprawności, korzystanie z głębokiej fałszywki wygeneruje alert, podobny do tego, który wyskakuje, gdy próbujesz uzyskać dostęp do niezabezpieczonej witryny, więc będziesz wiedział, że to co widzisz, nie jest tym co powinieneś dostać!

Image: Screen grab from video by VillainGuy.

Tłumaczenie: Wojciech Pałubicki, Uniwersytet im. Adama Mickiewicza w Poznaniu, dzięki uprzejmości Polskiego Towarzystwa Matematycznego.

Więcej informacji:

"Protecting World Leaders Against Deep Fakes," by Agarwal, Farid, Gu, He, Nagano, and Li, 2019.

MM/146/PL

AMS American Mathematical Society

*Program **Mathematical Moments** promuje znaczenie i rozumienie roli, jaką matematyka odgrywa w nauce, przyrodzie, technice i kulturze.*

www.ams.org/mathmoments