Allarme deepfake, così l’intelligenza artificiale ci aiuterà (forse) a combattere i falsi creati dall’intelligenza artificiale

L’ultimo esempio è quello più innocuo, ma che però ha creato più scalpore: un finto Tom Cruise, praticamente identico all’originale, che scherza e ride in una serie di **video molto condivisi su TikTok**. Talmente condivisi da suscitare un ampio dibattito online e da spingere il loro creatore a rimuoverli temporaneamente dal social network, dimostrazione pratica del livello raggiunto dai cosiddetti deepfake, quei falsi (immagini, video, audio) **realizzati grazie all’utilizzo di algoritmi basati sull’intelligenza artificiale**, che partendo da un volto è in grado di simularne un altro, anche ricreando la mimica facciale e le espressioni.

Quelle clip erano (sono) **fatte per divertire** e infatti su

TikTok sono tornate e sono rimaste, ma il timore di molti è che siano un antipasto di quello che ci attende in futuro, quando i deepfake saranno utilizzati per imitare in maniera incredibile (anzi, molto credibile) **un esponente politico, un personaggio pubblico, un vicino di casa**, la maestra di nostra figlia. E farle dire qualsiasi cosa. Come faranno le persone a capire che cosa è vero e che cosa no? Come faranno i giornalisti? Facendosi aiutare dalla tecnologia, ovviamente.

Le IA usate contro le IA

Alcuni ricercatori dell'Università di **Buffalo, negli Stati Uniti**, hanno trovato un modo per distinguere i volti umani da quelli generati da un computer, analizzando il riflesso negli occhi. Nel **documento stilato dagli scienziati (che è questo, in pdf)** viene ricordato che la cornea funziona un po' come uno specchio e riflette la luce che si trova di fronte: nel caso degli esseri umani, **quello che si vede riflesso nei due occhi** è pressoché uguale, perché hanno davanti gli stessi oggetti e le stesse fonti luminose; nel caso dei deepfake questo non succede, o comunque non succede quasi mai, perché (semplificando) **i volti artificiali (come questo)** vengono creati da database di facce che vengono combinate insieme per ottenere il risultato desiderato e gli occhi possono anche arrivare da due visi diversi.

Per trovare i falsi, i ricercatori hanno utilizzato **un software che ha imparato a riconoscere gli umani** dopo avere studiato decine di migliaia di occhi e i loro riflessi, cioè un'intelligenza artificiale per contrastare un'altra intelligenza artificiale. E i risultati sono piuttosto soddisfacenti, visto che sfiorano il 95% di affidabilità.

Uno dei video del falso Tom Cruise su TikTok

> *@deeptomcruise*
> *My □ impression…*

♬ origineel geluid – Tom

Una battaglia che è appena iniziata

Con qualche controindicazione, evidenziata dagli stessi ricercatori: il sistema funziona (molto) bene se davanti al viso c'è una fonte di luce abbastanza chiara ed evidente da generare un riflesso sulle cornee e soprattutto **se entrambi gli occhi sono visibili**, così che l'IA possa metterli a confronto; inoltre, un successivo lavoro di post-produzione sul "falso" potrebbe intervenire anche a livello di questi dettagli, così da armonizzare fra loro i riflessi su occhio destro e occhio sinistro.

Col tempo, comunque, è probabile che queste contromisure diventino **ancora più efficaci (nell'aiutarci)**, cosa che però faranno **anche i deepfake (nell'ingannarci)**. Insomma, è solo l'inizio dell'ennesima battaglia fra buoni e cattivi… solo che questa volta riguarda le macchine.