PuntoSicuro

ARTICOLO DI PUNTOSICURO

Anno 26 - numero 5590 di Venerdì 29 marzo 2024

Un interessante documento sui metodi individuazione di immagini Deepfake

Il fenomeno della contraffazione delle immagini fotografiche o video sta letteralmente esplodendo, con conseguenze potenzialmente assai gravi sia sulle persone coinvolte, sia sui destinatari dei messaggi.

Il General accounting Office americano ha pubblicato un prezioso documento che dà indicazioni su come sia possibile rilevare indizi di immagini contraffatte. La stessa intelligenza artificiale che viene utilizzata per realizzare video, audio ed immagini contraffatte può essere utilizzata per mettere in evidenza delle anomalie, che possono mettere sul chi vive chi questi file osserva.

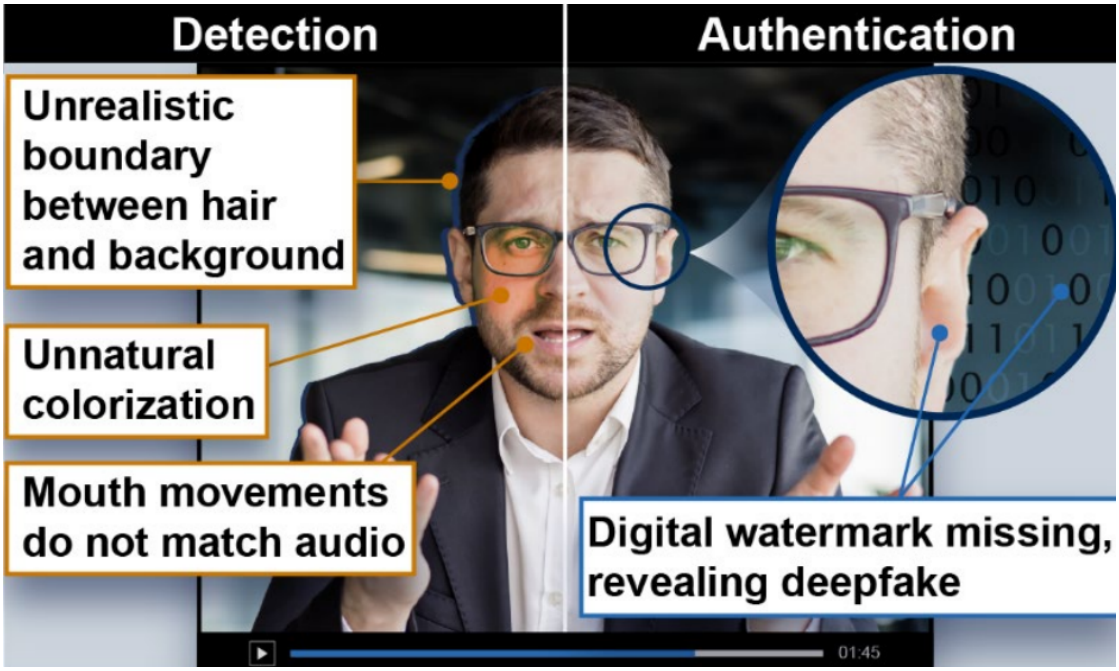
Esistono vari metodi, che possono esaminare:

- inconsistenze facciali e vocali,
- segni evidenti di generazione di file Deepfake,
- anomalie sui colori.

A queste tecnologie di individuazione di file contraffatti si dovrebbero abbinare anche le tecnologie, che garantiscono le genuinità dell'immagine. Queste tecnologie di autentica sono articolate a vari livelli, come ad esempio:

- l'applicazione di contrassegni digitali, che permettono di mettere in evidenza successive alterazioni dell'immagine, inizialmente contraffatta,
- l'utilizzo dei metadati, che descrivono le caratteristiche del file e che possono essere inseriti in modo crittografico, a prova di sostituzione. La mancanza od alterazione di questi contrassegni indica che il file è stato alterato,
- La distribuzione delle immagini mediante sistemi blockchain, che crea una rete relativamente sicura e che rende assai difficile l'alterazione dell'immagine, in quanto è possibile effettuare sempre un confronto fra un'immagine e la successiva e rilevare alterazioni.

Il documento offre, ad esempio, l'immagine che segue, nella quale gli specialisti hanno messo in evidenza alcune anomalie.



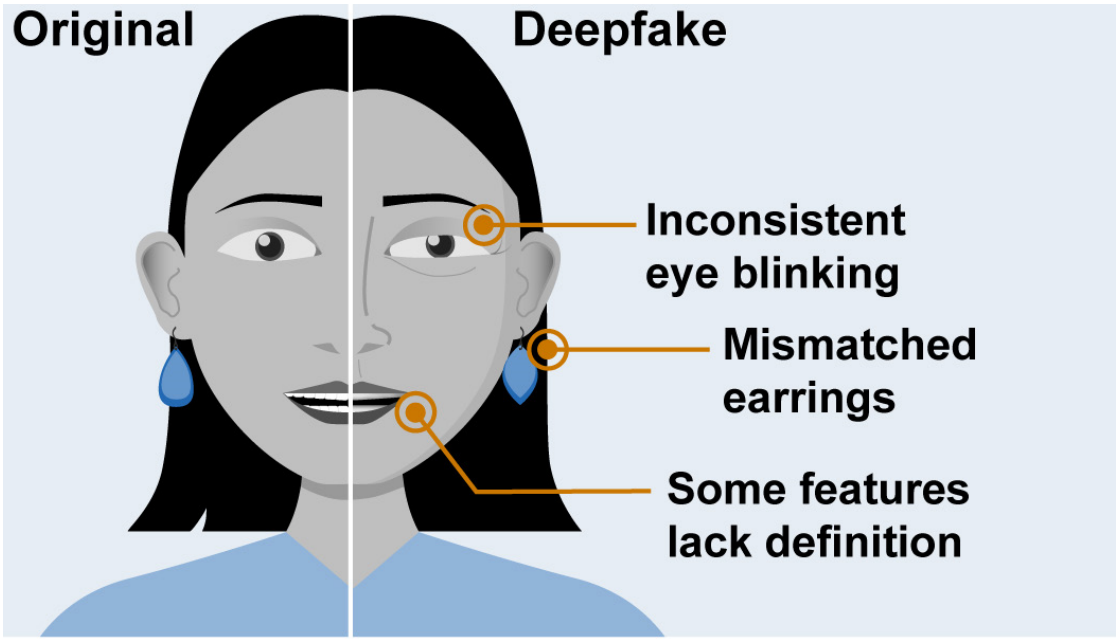
Ad esempio, la linea di demarcazione fra la capigliatura e lo sfondo può presentare caratteristiche anomale, come pure il colore della pelle. Un elemento particolarmente significativo è legato al fatto che i movimenti delle labbra possono non corrispondere al messaggio parlato, che viene pronunciato.

Per quanto riguarda invece i contrassegni di autenticità, è possibile rilevare l'assenza di un contrassegno digitale, posto sul lobo dell'orecchio, che dimostra come l'immagine sia stata alterata.

Il documento illustra anche altre tecniche assai significative, come ad esempio il fatto che il battito delle ciglia della persona che viene ripresa potrebbe avere una frequenza anomala, non coerente con una situazione di normalità.

D'altro canto, abbiamo tutti appreso, dai mezzi di comunicazione di massa, come una foto della famiglia reale inglese sia stata immediatamente percepita come alterata, quando osservata da esperti del settore.

Il documento riporta anche un'altra immagine, che può aiutare a mettere in evidenza situazioni anomale.



Source: GAO; conceived from DARPA image at https://www.darpa.mil/news-events/2019-09-03a. | GAO-20-379SP

In questo caso si può rilevare, come accennato in precedenza:

- un battito delle ciglia non normale,
- il fatto che un orecchino sia diverso dall'altro,
- il fatto che alcune caratteristiche facciali non abbiano lo stesso livello di definizione della porzione di volto originale.

I security manager dovranno cominciare a diventare, se non proprio esperti del settore, almeno un poco più a conoscenza di queste tecnologie, che potranno aiutare a separare il vero dal falso contraffatto, che potrebbe arrecare all'azienda danni non trascurabili.

COMBATING DEEPFAKES (pdf)

Adalberto Biasiotti

www.puntosicuro.it