VareseNews

Generatori di volti umani con AI: tra creatività e il rischio deepfake

Pubblicato: Mercoledì 26 Giugno 2024

Negli ultimi anni l'**intelligenza artificiale generativa**, o **generative AI**, ha compiuto progressi straordinari, fino ad arrivare alla possibilità di **generare volti umani** perfettamente realistici. Questi strumenti possono offrire grandi opportunità creative, specialmente negli ambiti del **marketing** e della **comunicazione**, ma presentano anche significativi rischi in termini di **sicurezza** e **privacy**.

Secondo la ricerca di ExpressVPN la possibilità che ormai ognuno di noi abbia un **gemello virtuale online**, non è più roba da film di fantascienza. Considerando che i risultati sono sempre più simili alla realtà, imparare a difendersi dai pericoli insiti in questa tecnologia non è più qualcosa di rimandabile.

Come fa un'intelligenza artificiale a generare volti umani?

Ma **come si genera volto umano iperrealistico**? Le intelligenze artificiali che creano immagini, tra cui proprio i volti umani, si basano su **reti neurali generative avversarie**, anche dette **GANs** o **Generative Adversarial Networks**.

Le GANs consistono in sostanza in due reti neurali che lavorano in tandem svolgendo compiti opposti. Il **generatore** ha la funzione creativa e dà vita alle immagini artificiali; queste vengono successivamente analizzate da un **discriminatore** che valuta quanto siano realistiche, confrontandole con dati reali. Le due reti lavorano in competizione, così da migliorare continuamente le **capacità di creazione del**

generatore, che finirà per migliorare in maniera costante la verosimiglianza delle immagini.

Il processo inizia con l'**addestramento dell'intelligenza artificiale** su un vasto database di immagini e video di volti umani. Durante l'addestramento, il generatore impara a produrre immagini che ingannano il discriminatore fino a diventare talmente abile che i volti creati risultano indistinguibili da quelli reali.

Questo sistema, grazie a continui **feedback** e miglioramenti, raggiunge un livello di perfezione tale da creare **immagini fotorealistiche di persone** che non esistono, o repliche perfettamente credibili di foto di persone reali. Alcune agenzie di comunicazione stanno ad esempio creando delle **virtual influencer** per pubblicizzare i loro prodotti, con risultati fino a qualche anno fa nemmeno lontanamente immaginabili.

I rischi connessi all'uso dell'AI: i casi di deepfake

I deepfake sono una delle applicazioni più controverse dell'**AI generativa**. Essi utilizzano GANs per creare video e immagini falsi in cui le persone sembrano dire o fare cose che in realtà non hanno mai fatto. I rischi legati alla proliferazione in rete dei **deepfake** sono enormi, almeno tanto quanto le opportunità creative.

Uno dei principali pericoli è la **disinformazione**. I deepfake possono essere utilizzati per creare **video falsi di politici o personaggi pubblici**, diffondendo informazioni errate e manipolando l'opinione pubblica attraverso fake news.

Un altro rischio significativo è legato alla **violazione della privacy e della sicurezza personale**. I deepfake possono essere impiegati per creare **contenuti pornografici non consensuali**, mettendo a rischio la reputazione e la vita delle persone coinvolte. Possono inoltre essere utilizzati per **estorsioni** e altre **attività criminali**, al punto che è stato stimato che proprio i deepfake saranno una delle cyberminacce più impattanti del 2024.

Come proteggersi dai deepfake

La **consapevolezza pubblica** e **l'educazione digitale** sono sicuramente tra le armi di difesa più potenti contro i deepfake. Bisogna innanzitutto educare le persone ai rischi e alle caratteristiche delle AI generative, favorendo la divulgazione e lo sviluppo di un senso critico nei confronti dei contenuti digitali. Le **campagne di sensibilizzazione** possono ad esempio insegnare agli utenti come verificare l'autenticità delle informazioni e come segnalare contenuti sospetti.

Sul fronte legale, molti paesi stanno adottando normative specifiche per **contrastare i deepfake**, alla luce anche di un pericoloso aumento nella frequenza di diffusione di contenuti falsi. **Leggi che puniscono la creazione e la distribuzione di deepfake** non consensuali possono sicuramente dissuadere i malintenzionati, anche se l'applicazione delle stesse norme rimane un ambito controverso e non sempre facile.

Anche le **piattaforme online** possono dare il loro contributo nel rilevare, segnalare e rimuovere contenuti deepfake per proteggere gli utenti. Sono inoltre in fase di sviluppo diversi strumenti utili per **riconoscere immagini e video falsi**. Questi strumenti utilizzano **tecniche di analisi avanzate**, in grado di individuare anomalie impercettibili a occhio nudo, tra cui incongruenze nella frequenza dei fotogrammi o nella resa delle ombre.

Va però considerato che la rapida evoluzione delle **tecnologie di AI generativa** richiede un'altrettanta costanza nello sviluppo di contromisure valide, dal momento che difese anti-deepfake potrebbero diventare obsolete nel giro di pochi mesi.

Redazione VareseNews
redazione@varesenews.it