Al telefono sembra un parente, un amico, oppure un personaggio famoso, però non è lui. L'Intelligenza artificiale riesce a riprodurne fedelmente le caratteristiche sonore. E il pericolo truffa è più che mai concreto.

Non fidatevi di quella voce

di Stefano Piazza
e Luciano Tirinnanzi

Riprodurre con incredibile fedeltà la voce umana è l'ultima frontiera delle truffe telefoniche. O meglio, una loro evoluzione. Ricordate quando l'oppositore russo Alexey Navalny truffò i servizi segreti russi imitando la voce del loro capo, e riuscì a farsi dire chi e come lo aveva avvelenato? Il caso gettò grave imbarazzo al Cremlino e più di una testa rotolò per quella umiliazione. I funzionari dell'Fsb, erede del famigerato Kgb, si giustificarono affermando che la voce era proprio come quella del loro capo. Il dissidente era dunque un buon imitatore, ma oggi il suo talento non sarebbe servito.

Già, perché appropriarsi dell'identità e del timbro vocale di qualcun altro, con le nuove funzionalità dell'intelligenza artificiale è un gioco da ragazzi. Ne sa qualcosa Massimo Moratti: l'ex presidente dell'Inter lo scorso febbraio ha versato ben 890 mila euro sul conto di una banda di truffatori che lo avevano convinto grazie a una telefonata di un falso ministro della Difesa Guido Crosetto, che chiedeva all'imprenditore quel denaro per un fantomatico riscatto finalizzato alla liberazione di inesistenti giornalisti rapiti in Medio Oriente. La Guardia di finanza ha poi recuperato il bottino, ma nella maggior parte dei casi le truffe vanno a buon fine.

Online esistono ormai numerose piattaforme a disposizione di chiunque che - grazie a banali software di voice cloning o text-to-speech avanzati - consentono di riprodurre digitalmente una voce pressoché identica a quella originale, con tanto di inflessioni, accenti e quant'altro sia sufficiente a ingannare il malcapitato. Per clonare la voce bastano brevi clip audio autentici, anche di pochi secondi: all'Ia è sufficiente un messag-

gio vocale inviato via WhatsApp o, nel caso di figure pubbliche, un intervento registrato durante un evento. Questi strumenti - disponibili sia in versione gratuita che a pagamento - sono sin troppo facili da usare e permettono di generare frasi vocali artificiali nel giro di pochi istanti e a prezzi irrisori.

Un esempio è Speechify Voice Cloning, che sfrutta sofisticati algoritmi basati sul deep learning per riprodurre audio estremamente realistici da usare per adescare le vittime designate. Alla macchina è sufficiente analizzare una registrazione di appena 30 secondi per ottenere una copia sintetica e credibile della voce, con infinite possibilità accessorie: modificarne l'intonazione, il ritmo, così come aggiungere o eliminare pause, simulando persino affanno o emozione per personalizzare ogni messaggio e indurre l'ascoltatore a credere alle parole «sintetiche». Vale ad esempio per il programma Veed.io, pensato apposta per catturare la voce: basta inserire un testo e si ottiene un *voice over* che riproduce fedelmente il timbro registrato. Così fa anche Vidnoz AI Voice Changer, che in più offre una vasta libreria con oltre 100 voci predefinite e supporta più di 140 idiomi differenti.

Evidentemente, queste soluzioni tecnologiche non sono state immaginate per truffare. Ma l'uso improprio spopola. Sono lontani i tempi degli innocenti scherzi telefonici per evitare la scuola o il lavoro, come quelli immortalati nelle pellicole di Paolo Villaggio: «Fantozzi, faccia l'accento svedese!». Oggi l'Ia può imitare alla perfezione la voce di Tom Cruise o Giorgia Meloni e tradurla in tutte le lingue del mondo, al prezzo di un abbonamento alla metro e in pochi istanti. In gergo, si chiama «voice scam».

In Italia, attualmente, le truffe del voice scam più sfruttate sono le «estorsioni emotive» che mirano specificamente ai familiari delle vittime. Il raggiro più classico per questa tipologia di reato consiste nel chiamare oppure nel mandare un vocale pre-registrato a un parente stretto, simulando un'urgenza: «Mamma mi hanno rapinato, per favore mandami subito dei soldi a questo conto poi ti spiego meglio...». In altri casi, la leva emotiva è una finta emergenza lavorativa nei confronti di dipendenti che si occupano di amministrazione: «Ciao, sono io (il capo, *ndr*), ho bisogno che tu faccia subito un bonifico a queste coordinate… è un cliente importante e non possiamo perderlo». Ancora, la richiesta di accesso a conti bancari per un falso incidente stradale: «Mi servono i codici della carta di credito per pagare il carro attrezzi, sbrigati per favore sono in autostrada...».

Dato che la voce è davvero identica a quella reale, la credibilità e il successo della truffa dipendono principalmente dall'abilità nell'uso dei termini opportuni, dalla rapidità del messaggio e dall'«effetto sorpresa»: sfruttando la pressione del tempo e la suggestione di un familiare che riceve una cattiva notizia e si vuole rendere utile, è facile colpire nel segno. Vale soprattutto per le persone più vulnerabili e meno esperte di tecnologia, come gli anziani. Nel 2023, non a caso, gli ultra 65enni sono stati vittime predilette di truffe finanziarie per un totale stimato in milioni euro. La cifra esatta è 559,4 milioni di euro ed è stata calcolata dalla Fabi, il sindacato dei bancari. Secondo un loro studio, le truffe online hanno fruttato ai criminali digitali ben 114 milioni nel 2022, lievitati a 181 nel 2024 (+58 per cento). Una tendenza purtroppo destinato a proseguire, date le sconvolgenti potenzialità dell'Ia.

Come proteggersi dunque? Anzitutto, riconoscere i campanelli d'allarme: richieste urgenti di denaro, massima segretezza (cioè non informare altri parenti) e richieste di pagamento insolite (bonifico, buoni regalo o criptovalute) sono quasi sempre un segno di frode. Utile è anche riagganciare e richiamare il numero del parente, in caso di sospetto (un numero sconosciuto è spesso sinonimo di truffa). Infine, è utile usare una parola in codice prestabilita per verificare la fondatezza della chiamata e del suo contenuto. ■

L'esperta: «Tuteliamo la voce umana»

Elisa Garfagna, voice designer, doppiatrice e podcaster ne sa qualcosa di voce. E forte della sua esperienza racconta l'allarme.

La voce generata da la è un pericolo per il cittadino comune?
Sì, purtroppo è un pericolo tangibile. La voce è un dato sensibile, ognuno di noi ha una voce unica, che può essere utilizzata, per esempio, come dato biometrico per le banche e consentire a potenziali truffatori di accedere ai nostri risparmi. Se la nostra voce può essere campionata, significa che tutti possiamo dire tutto, anche il falso: con la nostra voce potrebbero farci ammettere un crimine, un omicidio, magari anche un semplice tradimento, che però non è mai avvenuto.
Quanto è avanti l'Ia nell'imitare la voce delle persone?
Le voci campionate sono fedeli, ormai, al 99 per cento rispetto all'originale. E fa passi da gigante. È una tecnologia spaventosamente potente che nelle mani sbagliate può generare disastri.
Le è capitato di sentire di truffe o di usi sbagliati della voce, come imitazioni di quella dei doppiatori?
Sì, mi è capitato di parlare con un collega a cui hanno campionato la voce, senza che lui ne fosse a conoscenza, e che una casa di produzione abbia utilizzato la sua voce, o meglio quella artificiale, per il voice over di uno spot. Per quanto mi riguarda, ho sentito, inoltre, la mia voce campionata, per scherzo, ed è stato traumatico per l'estrema vicinanza all'originale.
Cosa pensa che dovrebbero fare i doppiatori e le persone che lavorano nel tuo settore per proteggersi dalle voci create dall'intelligenza artificiale e dalle truffe che possono incorrere?
Innanzitutto, trasmettere all'esterno il valore e la bellezza del doppiaggio come forma d'arte. Noi italiani siamo i numeri uno in questo settore! Poi servirebbe che, attraverso leggi ad hoc che tutelano i dati biometrici individuali, venisse inserita anche la voce umana. Infine, richiedere, come già sta accadendo con le immagini, che obbligatoriamente venga applicato un'etichetta che segnali che le voci presenti in un progetto audiovisivo sono generate da Intelligenza artificiale.■

A sinistra, Hong Kong, dove spesso finiscono i soldi delle truffe vocali. Qui, Marco Tronchetti Provera, ingannato con altri dalla falsa voce del ministro della Difesa Guido Crosetto (sopra).

GETTY IMAGES (3), ISTOCK