POLITECNICO MILANO 1863

COMUNICATO STAMPA

—

Il Politecnico di Milano contro i deepfake: due progetti per scovare video e audio falsi

Con FF4ALL i ricercatori del Politecnico analizzano come vengono generati immagini e video sintetici, con FUN-Media studiano invece i deepfake vocali, minaccia emergente e ancora poco conosciuta

Milano, 28 aprile 2026 – Si sono conclusi due progetti di ricerca europei per il rilevamento dei **deepfake**, contenuti digitali falsi, e il contrasto alla loro diffusione: **FF4ALL** e **FUN-Media**. L'**Image and Sound Processing Lab** (ISPL) del Politecnico di Milano, finanziato da fondi PNRR, per FF4ALL ha analizzato fenomeni emergenti legati alla **generazione di immagini e video sintetici**, mentre per FUN-Media si è concentrato sulla rilevazione di **deepfake vocali**, una delle minacce emergenti più rilevanti nel panorama della sicurezza digitale. I risultati dei progetti fanno un passo importante verso lo sviluppo di tecnologie affidabili per la tutela dell'informazione digitale, il contrasto alla disinformazione e la protezione degli utenti in un ecosistema mediatico sempre più complesso e dinamico.

L'Image and Sound Processing Lab (ISPL) del **Dipartimento di Elettronica, Informazione e Bioingegneria** del Politecnico di Milano da anni è impegnato nello sviluppo di tecniche avanzate per l'analisi forense multimediale. Le attività in questo ambito sono coordinate dai professori Stefano Tubaro e Paolo Bestagini, con il contributo dei ricercatori Sara Mandelli e Luca Comanducci.

FF4ALL

I ricercatori dell'ISPL hanno studiato i modi in cui vengono ingegnerizzati e diffusi immagini e video falsi. Sono state studiate per esempio tecniche che consentono di trasformare immagini reali in versioni sintetiche estremamente realistiche, rendendo più complessa la verifica della loro autenticità e mascherando tracce fondamentali per l'analisi forense. Parallelamente, il laboratorio ha sviluppato **nuovi strumenti per il rilevamento di volti sintetici**, combinando informazioni geometriche tridimensionali e caratteristiche strutturali del volto. Queste soluzioni migliorano la capacità di generalizzare i modelli che individuano i falsi, e mantengono buone prestazioni anche in presenza di operazioni di post-processing, come compressione o editing.

«Un ulteriore contributo riguarda lo studio dei sistemi usati per rilevare i falsi» spiega il professor **Stefano Tubaro**. «Comprendere su quali elementi basino le proprie decisioni è infatti cruciale per aumentarne l'affidabilità». I modelli basati su intelligenza artificiale, addestrati su grandi quantità di dati, risultano spesso difficili da interpretare e non è sempre chiaro quali caratteristiche vengano utilizzate per classificare un contenuto. In questa direzione, sono state sviluppate tecniche per identificare le **regioni del volto più rilevanti** per la classificazione, rendendo più trasparenti i processi decisionali dei detector.

Le attività del progetto hanno incluso l'analisi dell'impatto di nuove **tecnologie di compressione** basate su intelligenza artificiale, che possono generare artefatti manipolati difficili da distinguere dai contenuti autentici e manipolati.

In collaborazione con le università partner del progetto è stato infine realizzato il **dataset WILD, che raccoglie immagini false** generate da venti modelli di ultima generazione: è una risorsa importante per identificare la tecnologia generativa utilizzata per sintetizzare un'immagine.

FUN-MEDIA

Con FUN-Media, il focus è stata la rilevazione di **deepfake vocali**. Per affrontare questa sfida, i ricercatori dell'ISPL hanno sviluppato nuove architetture basate sui cosiddetti modelli *Mixture of Experts* per la rilevazione dei falsi, in grado di combinare più sistemi specializzati per migliorare le prestazioni anche in presenza di tecniche generative mai osservate durante l'addestramento. Questi approcci offrono maggiore flessibilità e adattabilità rispetto ai detector tradizionali, risultando particolarmente efficaci in scenari complessi e in continua evoluzione.

Un'ulteriore linea di ricerca ha esplorato l'utilizzo di detector forensi basati sul **rilevamento di anomalie**. In questo caso, i modelli vengono addestrati esclusivamente su segnali vocali autentici, apprendendone le caratteristiche distintive, e sono quindi in grado di identificare i contenuti sintetici come deviazioni rispetto al comportamento atteso.

«Accanto alla rilevazione, il progetto ha affrontato anche il problema dell'attribuzione, ovvero l'identificazione della tecnologia generativa responsabile della creazione di un contenuto audio» dice il docente **Paolo Bestagini**. Sono stati sviluppati detector in grado di stabilire se due tracce vocali siano state prodotte dallo stesso modello generativo. Ulteriori contributi riguardano lo sviluppo di tecniche per l'analisi dettagliata del **segnale vocale**, per esempio a livello di fonemi, e l'impiego di modelli capaci di evidenziare le caratteristiche acustiche più rilevanti.

IL CONTESTO

Negli ultimi anni, i rapidi progressi nei modelli di **intelligenza artificiale generativa** hanno profondamente trasformato il modo in cui contenuti digitali come immagini, video e audio vengono prodotti, condivisi e fruiti. Se da un lato queste tecnologie aprono nuove opportunità creative e applicative, dall'altro introducono rischi rilevanti in termini di sicurezza, disinformazione e manipolazione dei contenuti. In particolare, esse consentono di impersonare individui con crescente efficacia e di generare contenuti estremamente realistici, spesso **privi di tracce evidenti di alterazione**. Questo scenario amplifica il rischio di attacchi di social engineering e la diffusione di fake news su larga scala, rendendo i contenuti manipolati sempre più credibili e difficili da rilevare.

LINK AI VIDEO E ALLE IMMAGINI

PER INFORMAZIONI:

Martina Pagani, +39 345 116 6210, relazionimedia@polimi.it