Deep Fake: realismo virtuale o manipolazione digitale?

Agnese Rondelli
Riccardo Midali
Gkreis Savva

Repository: https://github.com/riccardomidali/deepfake
Presentazione:
https://www.canva.com/design/DAF2AYi_mxw/Zr2LEoWrNHFRFIuW-U0I9Q/edit?utm_content=DAF2AYi_mxw&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton

1. **Introduzione**

I deepfake sono delle **tecniche utilizzate per la sintesi di immagini umane basata sull'intelligenza artificiale** che utilizzano la tecnologia di apprendimento automatico chiamata **Generative Adversarial Networks** (**GAN**) per combinare e sovrapporre immagini e video esistenti con video o immagini originali. Sebbene la fabbricazione e la manipolazione di immagini e video digitali non siano una novità, i rapidi sviluppi delle reti neurali (Deep Neural Network) negli ultimi anni hanno reso il processo di creazione di immagini/audio/video falsi convincenti sempre più facili e veloci da realizzare. Il termine deepfake deriva dalla combinazione dalla tecnologia sottostante "deep learning", che è una forma di intelligenza artificiale, e dal termine “fake”, falso. Gli algoritmi di deep learning, che insegnano da soli a risolvere i problemi quando vengono forniti grandi set di dati, vengono utilizzati per scambiare volti nei contenuti video e digitali per creare supporti falsi dall'aspetto realistico. La parola "deepfake" finisce sempre più spesso per essere associata allo sconcertante problema della disinformazione online. In realtà, la tecnologia deepfake è molto di più dell’accezione negativa che gli è stata etichettata. Nei prossimi anni, sarà uno dei driver principali nella creazione ed elaborazione delle informazioni.

2. **Come nasce un deepfake?**

La storia degli deepfake ha radici principalmente nel progresso delle tecnologie di intelligenza artificiale (IA) e di deep learning.

Anni 1990-2000:

- Gli inizi della manipolazione digitale coinvolgono principalmente la grafica al computer e la post-produzione video, con l'uso di strumenti come Adobe **Photoshop** e software di editing video per modificare contenuti visivi.

Anni 2010:

- L'evoluzione delle reti neurali profonde (**deep neural networks**) e delle tecniche di **deep learning** ha portato a un significativo avanzamento nell'elaborazione dell'immagine e del riconoscimento facciale.

2015:

- Le Generative Adversarial Networks (**GAN**), un tipo di architettura di rete neurale, vengono introdotte da Ian Goodfellow e colleghi. Questa innovazione è fondamentale per la creazione di deepfake, in quanto consente la generazione di contenuti realistici attraverso la competizione tra un generatore e un discriminatore all'interno della rete.

2017-2019:

- Nei primi mesi del 2017, il termine "**deepfake**" diventa popolare quando un utente di **Reddit**, soprannominato "Deepfakes", comincia a pubblicare video di alcuni spezzoni di film dove vengono sostituiti i volti degli attori originali con i volti di altri attori, realizzando così dei video fake, abbastanza realistici, che suscitano fin da subito il divertimento e l’interesse di molti utenti. Il successo è tale che Deepfakes crea una

"subreddit", una specifica categoria all'interno di Reddit, dove i post sono organizzati in argomenti e seguiti da migliaia di utenti. Nel giro di poche settimane si passa, però, da video curiosi e divertenti, a spezzoni di video hard dove il volto dei pornoattori viene sostituito con quello di celebrità del mondo dello spettacolo, cominciando ad ingenerare perplessità e allarme. Questo ha portato a una maggiore attenzione sulla tecnologia deepfake e sulle sue possibili conseguenze negative. Negli anni successivi, la tecnologia deepfake è stata utilizzata per scopi sempre più sinistri, come la diffusione di notizie false, la manipolazione delle elezioni e la diffusione di pornografia non consensuale. Tuttavia, la tecnologia deepfake può anche essere utilizzata per scopi positivi, come la creazione di effetti speciali nei film e la creazione di contenuti educativi.

2020-2022:

- L'attenzione verso i deepfake si estende a livello globale, coinvolgendo governi, organizzazioni e istituzioni internazionali nella ricerca di soluzioni per affrontare le sfide etiche e di sicurezza associate.

2023 (Attuale):

- Gli sforzi nella creazione di deepfake avanzano, ma parallelamente vengono sviluppati e perfezionati anche **strumenti di rilevamento**. La regolamentazione e le normative sull'uso dei deepfake diventano argomenti di discussione chiave per affrontare le potenziali minacce che possono rappresentare per la società.

3. Processo di creazione di un deepfake

Sono molteplici i modi in cui è possibile generare un deepfake. Il più utilizzato si basa sull'utilizzo di **reti neurali** profonde di tipo **autoencoder,** che generano una rappresentazione compressa dei dati immessi come input. Ad esempio, per sostituire il volto di un soggetto ripreso in un video con immagini del personaggio desiderato. Per ottenere risultati precisi e realistici, quando si lavora con foto o clip già esistenti, si utilizza una tecnica di sostituzione del volto. Questo consente di creare una base appositamente dedicata a tale scopo, assicurando una precisione e un realismo senza pari. L'autoencoder, più nello specifico, *studia i* video del soggetto target e lo mappa nell'ambiente, applicando le caratteristiche fisionomiche. In tal modo, si ricrea, partendo anche da poche immagini statiche, un video artificioso che riproduce il volto e la voce della persona designata. Esattamente come se stesse realmente venendo ripresa. I movimenti della testa e del viso, infatti, seguiranno quelli del soggetto originariamente ripreso, rendendo il tutto molto più realistico e credibile. In linea con quanto detto precedentemente, un altro metodo utilizzato per creare deepfake è l'utilizzo di una **rete generativa avversaria (GAN)**. L'impiego di questo sistema comporta una maggiore realismo degli effetti deepfake. Grazie a una continua ricerca di difetti nel materiale video, si rende più difficile la decodifica tramite software appositamente sviluppati.

4. Reti neurali e CNN

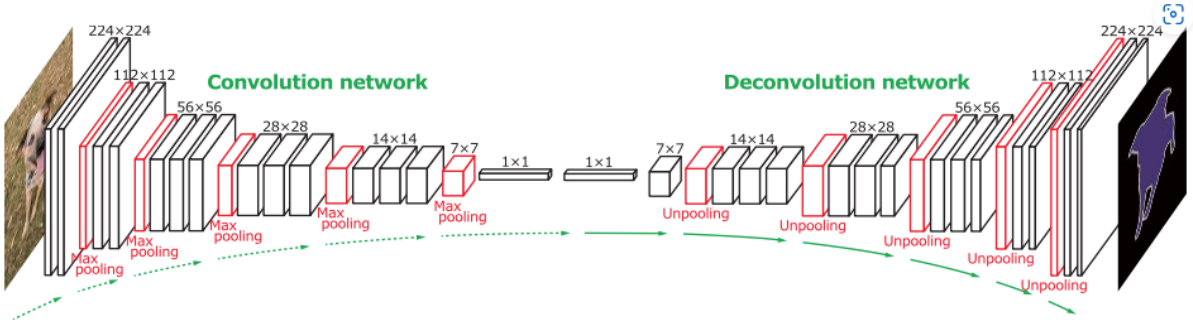
Le **reti neurali** sono modelli computazionali ispirati al funzionamento del cervello umano. Sono composte da unità chiamate neuroni, organizzate in strati (input, strati nascosti e output).Ogni connessione tra i neuroni ha un peso, e durante l'addestramento della rete, questi pesi vengono regolati per ottimizzare la risposta della rete a determinati input. Le reti neurali possono essere utilizzate per una vasta gamma di compiti, inclusi problemi di classificazione, regressione, e altre attività di apprendimento automatico.
Le **Convolutional Neural Networks** sono un tipo di rete neurali progettate per il riconoscimento delle immagini: grazie a strati che filtrano le informazioni in modo gerarchico. Nelle CNN ogni strato svolge un ruolo specifico nell'elaborazione delle informazioni visive:

- **Strato di input**: Questo strato rappresenta l'immagine di input che la rete neurale elaborerà.
- **Strati convoluzionali:** Questi strati eseguono operazioni di convoluzione applicando filtri o kernel sull'immagine. I filtri individuano caratteristiche specifiche come linee, bordi, texture, etc. Inoltre, questi strati sono responsabili dell'estrazione delle caratteristiche dell'immagine.
- **Strati di pooling**: Dopo la convoluzione, vengono solitamente inseriti strati di pooling (come il max pooling) che riducono la dimensione spaziale delle feature map, conservando le informazioni più rilevanti e riducendo la complessità computazionale.
- **Strati completamente connessi**: Questi strati ricevono le informazioni elaborate dalle fasi precedenti e le utilizzano per eseguire le decisioni finali. Questi strati sono simili a quelli di una rete neurale tradizionale e vengono spesso utilizzati per classificare l'oggetto presente nell'immagine.
- **Strato di output**: È l'ultimo strato della rete e produce l'output finale, che potrebbe essere una classificazione (ad esempio, gatto, cane, barca, ecc.) o altri tipi di output in base al compito specifico della rete.

5. General Adversarial Network

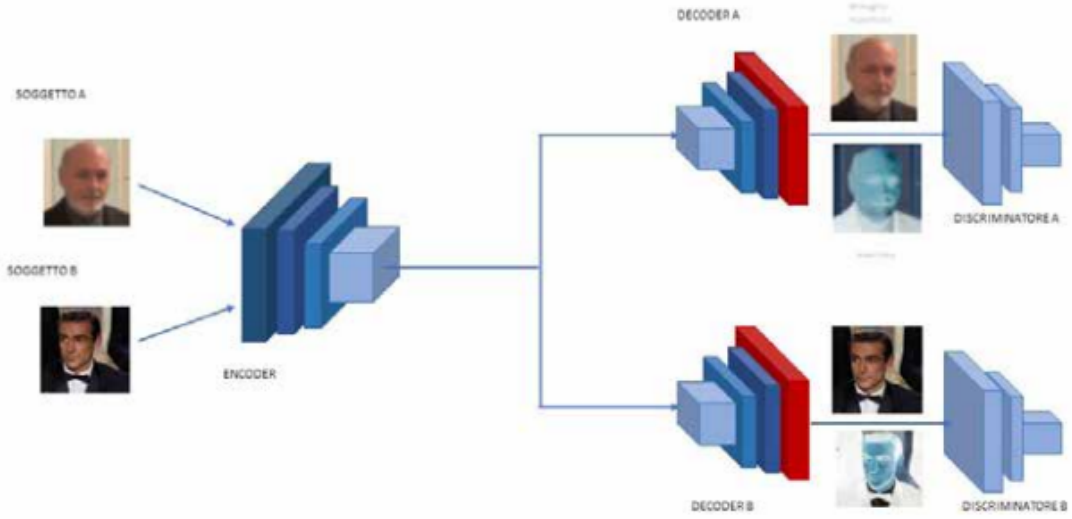
La tecnologia deepfake, utilizza algoritmi di Intelligenza Artificiale che richiedono grandi capacità computazionali e di risorse, in generale la tecnica più utilizzata è chiamata **Generative Adversarial Network** (GAN).
La tecnica del GAN fa parte di un ramo di apprendimento automatico basato sulle reti neurali. Queste reti sono progettate per emulare i processi neuronali del cervello umano e possono essere addestrate a riconoscere o manipolare specifiche attività e informazioni.
La caratteristica di un modello GAN risiede nella rivalità tra due o più reti neurali, nel caso del modello GAN per la creazione del deepfake le reti si chiamano generatore e discriminatore.Il modello di GAN utilizzato per la generazione dei deepfake, sfrutta due reti neurali messe l'una contro l'altra con l'obiettivo di generare un output realistico. Lo scopo è garantire che i deepfake creati siano il più reali possibile.
Il **generatore** ha come compito compito di generare un'immagine contraffatta utilizzando vettori di rumore - un elenco di numeri casuali - che sembrino il più realistici possibile. Il risultato è una rappresentazione dimensionale, di qualità inferiore, di quella stessa faccia che, a volte, viene definita vettore di base o faccia latente.
La seconda rete, il **discriminatore**, determina la veridicità delle immagini generate. Confronta l'immagine contraffatta, generata dal generatore, con le immagini autentiche nel set di dati per determinare quali immagini sono reali e quali false. Sulla base di questi risultati, il generatore, varia il parametro per la generazione delle immagini. Questo ciclo

continua fino a quando il discriminatore non riesce ad accertare che un'immagine generata è falsa e viene quindi utilizzata nell'output finale.
Per poter generare un deepface e fare un face-swap dal soggetto A a soggetto B - è necessario addestrare una rete di codificatori basata su reti neurali convoluzionali, anche dette CNN (Convolutional Neural Networks).



Una CNN utilizza molti strati convoluzionali, ovvero strati dove viene impiegata un'operazione matematica chiamata convoluzione che filtra gli input per trovare le informazioni più utili.

La rete CNN deve essere addestrata prima di poter generare deepfake e per questo utilizza centinaia di immagini, set dati, del soggetto A e del soggetto B.



L'encoder impara a codificare le caratteristiche come forma, colore, espressioni di entrambi le immagini tra cui vogliamo i face swap. Dopo l'uscita dell'encoder, le immagini vengono presentate a due decodificatori sulla rete CNN. I decodificatori non solo imparano come ricostruire le due immagini ma anche come creare delle maschere che aiutano a produrre delle immagini più realistiche dopo lo scambio dei volti.

La combinazione encoder-decoder è comunemente chiamata autoencoder e costituisce la rete di generazione del GAN.

Pertanto, si hanno due GAN: GAN A (costituito da encoder e decoder A) e GAN B (costituito dall'encoder e decoder B). Due discriminatori separati per A e B, imparano a distinguere meglio tra immagini reali e false. Quando inseriamo le immagini generate nel rispettivo discriminatore, la rete GAN spinge il generatore a realizzare immagini più realistiche, un ciclo continuo che termina quando le immagini generate non sono distinguibili da quelle reali.

Concluso l'addestramento, per generare il deepfake inviamo l'immagine del soggetto A all'encoder e al decoder B per ricostruire l'immagine. Siccome il decodificatore B ha imparato a generare il volto del soggetto B, genererà il volto del soggetto B con le caratteristiche del volto del soggetto A.

Immagine generata dal GAN:

Oltre alle immagini i modelli GAN sono in grado di elaborare anche segnali audio. La differenza nell'analisi da parte dei GAN tra audio e immagini è molto simile: le immagini possono essere viste come matrici, mentre i segnali audio possono essere considerati come semplici vettori.

6. Deepfake Training Process

La creazione di un deepfake coinvolge diverse fasi complesse, che sfruttano algoritmi di intelligenza artificiale, in particolare le Generative Adversarial Networks (GAN). Ecco una panoramica del processo di creazione di un deepfake:

1. **Raccolta dei Dati:**
    - Raccogliere un ampio set di dati (dataset) contenente immagini o video della persona di destinazione, il cosiddetto "attore". Questo set di dati è fondamentale per insegnare al modello a imitare tratti facciali, espressioni e movimenti della persona.
2. **Pre-elaborazione dei Dati:**
    - Pre-elaborare i dati per garantire che siano uniformi e di alta qualità. Ciò può includere il ritaglio del viso, l'allineamento delle immagini e la normalizzazione dei colori. La qualità dei dati di addestramento influisce direttamente sulla capacità del modello di generare deepfake credibili.
3. **Creazione del Modello GAN:**
    - Creare un modello di GAN, composto da due parti principali:
        - Generatore (Generator): Crea nuove immagini o video sintetici.

        - Discriminatore (Discriminator): Valuta se un'immagine è reale o sintetica.
4. **Addestramento del Modello:**
    - Addestrare il modello GAN utilizzando il set di dati raccolto. Durante questo processo, il generatore cerca di produrre immagini sintetiche sempre più realistiche, mentre il discriminatore cerca di migliorare la sua capacità di distinguere tra immagini reali e sintetiche.Uso di set di dati di addestramento e di validazione per monitorare le prestazioni del modello. Utilizzo di algoritmi di ottimizzazione per regolare i pesi del modello. Iterazioni continue di addestramento e validazione fino a raggiungere risultati soddisfacenti.

Dataset

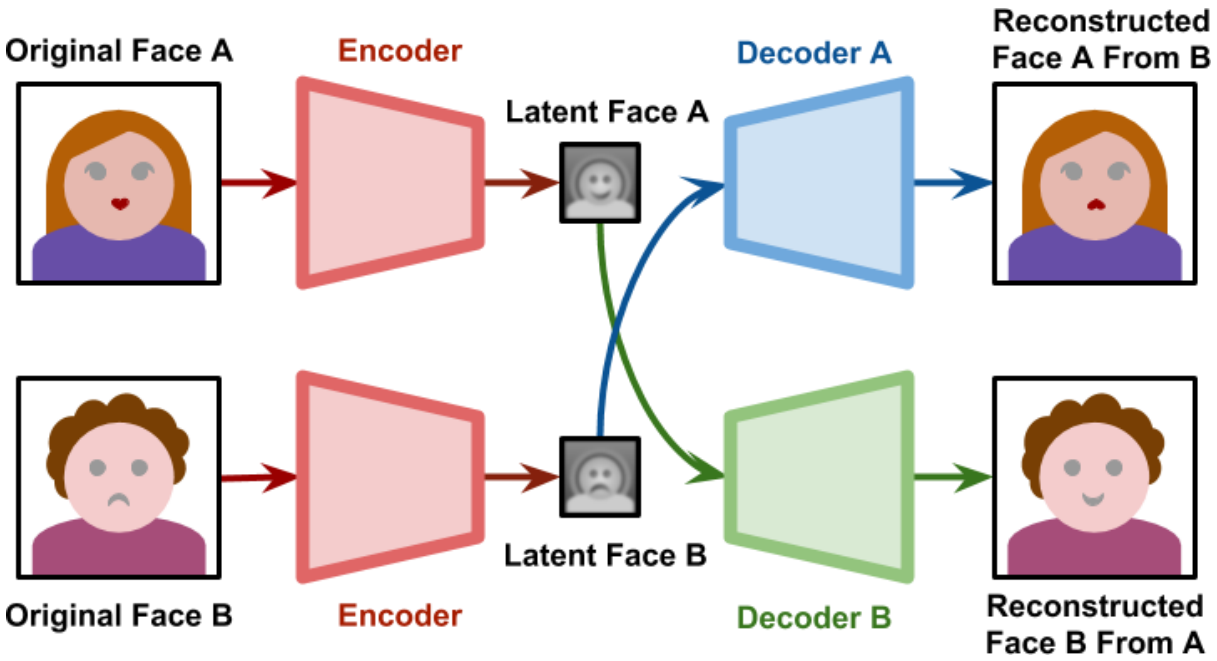
Creare deepfake richiede l'uso di dataset ampi e diversificati per addestrare reti neurali in grado di generare contenuti realistici. Tuttavia, è importante sottolineare che l'uso responsabile di tali dataset è fondamentale, e molte piattaforme online hanno politiche che vietano la creazione e la diffusione di contenuti dannosi o manipolati senza consenso.

Ecco alcuni dataset comunemente utilizzati per addestrare modelli di deepfake:

- **CelebA-HQ:** CelebA-HQ è un dataset di volti ad alta qualità che contiene immagini di celebrità. Questo dataset è spesso utilizzato per addestrare modelli di deepfake che coinvolgono il volto umano.
- **FFHQ (Flickr-Faces-HQ):** Questo dataset contiene immagini di volti raccolte da Flickr con una risoluzione elevata. È stato creato per la generazione di immagini realistiche di volti umani e può essere utilizzato per addestrare modelli di deepfake.
- **DeepFakeTIMIT**: Basato sul dataset TIMIT, che è un dataset di riconoscimento vocale, DeepFakeTIMIT è stato esteso per includere video di espressioni facciali. È stato utilizzato per addestrare modelli di deepfake che coinvolgono sia il volto che la voce.
- **VoxCeleb**: Questo dataset contiene registrazioni audio e video di discorsi provenienti da varie fonti, inclusi discorsi di celebrità. Può essere utilizzato per addestrare modelli di deepfake che coinvolgono la sincronizzazione labiale e la manipolazione della voce.
- **DeepFake Challenge Dataset (DFDC):** Creato nell'ambito di una competizione organizzata da Facebook, il DFDC contiene un ampio numero di video deepfake insieme ai corrispondenti video originali. È stato utilizzato per lo sviluppo di tecniche di rilevamento dei deepfake.
- **UADFV (Utrecht Audio-Visual Deepfake Dataset):** Questo dataset include video deepfake con corrispondenti video reali. È stato creato per la ricerca sulla manipolazione di contenuti audio e visivi.

7. **I tools per la creazione**

DeepFaceLab (https://github.com/iperov/DeepFaceLab)
Si tratta di uno dei software più utilizzati per la creazione di contenuti raffiguranti sostituzioni di volti. Utilizza una combinazione di GAN e algoritmi di deep learning per eseguire le attività.
Processo di funzionamento:



In questo schema viene illustrato come vengono utilizzate le reti neurali combinate per effettuare operazioni di codifica e decodifica sulle immagini facciali, il processo di cui si occupa questo tool.

Encoding delle facce: si utilizza una rete neurale per codificare le caratteristiche chiave di un volto in un vettore latente. Questo rappresenta l'essenza dell'aspetto di quella faccia in forma numerica.

Manipolazione dei vettori latenti: Una volta ottenuti i vettori latenti delle due facce, puoi manipolarli in diverse modalità. Ad esempio, puoi combinare o modificare questi vettori per ottenere risultati diversi, come la creazione di un ibrido tra due volti o la sostituzione di un volto con un altro.

Decoding delle facce: Dopo aver manipolato i vettori latenti, DeepFaceLab utilizza il decodificatore della rete neurale per convertire nuovamente questi vettori in immagini facciali. Questo processo di decodifica produce le immagini modificate o create, basate sulle modifiche apportate ai vettori latenti.

8. Algoritmi e librerie

DeepFace
DeepFace è un efficiente framework per il riconoscimento facciale e l'analisi degli attributi progettato per Python. Questo strumento sfrutta modelli all'avanguardia come VGG-Face, Google FaceNet, OpenFace, Facebook DeepFace, DeepID, ArcFace, Dlib e SFace per riconoscere i volti e analizzare attributi come età, genere, emozioni e razza. Gli esperimenti dimostrano che mentre gli esseri umani raggiungono una precisione media del 97,53% nei compiti di riconoscimento facciale, i modelli racchiusi all'interno di DeepFace superano e superano questo livello di precisione.

Sebbene la libreria DeepFace non crei direttamente deepfake, le sue analisi costituiscono una parte essenziale del processo di verifica per identificare eventuali manipolazioni nei contenuti. I rilevatori dei tratti facciali possono essere utilizzati come indicatori di potenziali manipolazioni. Inoltre, combinando tali analisi con altri strumenti, potrebbero costituire i primi passi verso la creazione di deepfake.

Il nostro progetto utilizza la libreria DeepFace per eseguire analisi e riconoscimento facciale su immagini. È composto da 3 codici:

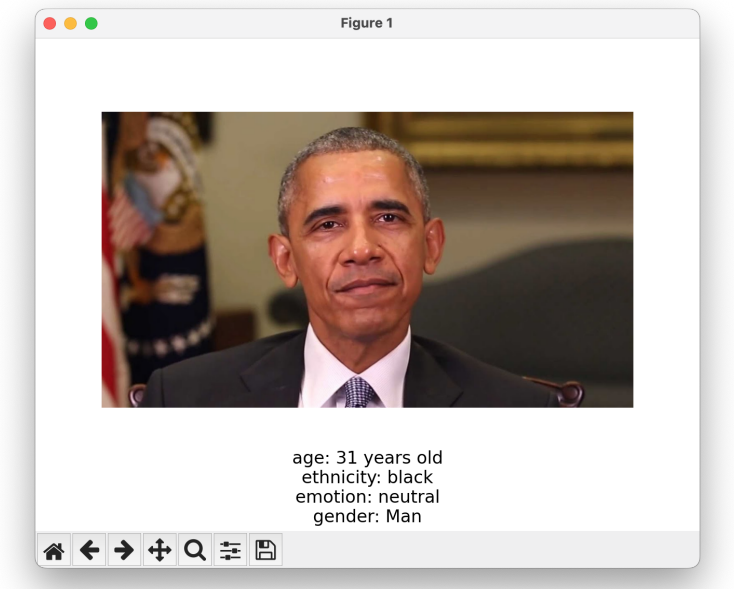
- **Face Recognition**: permette di valutare se due immagini in input raffigurano lo stesso volto. Utilizzando la libreria DeepFace, esegue un confronto tra le immagini per determinare la corrispondenza facciale.
- **Face Attribute**: accetta in input un'immagine contenente un volto e restituisce informazioni riguardanti età, genere, etnia e stato emotivo. Sfruttando le funzionalità di DeepFace, analizza i tratti facciali per fornire una descrizione.
- **Real Time Recognition**: offre le stesse funzionalità di Face Attribute, ma funziona in tempo reale utilizzando la webcam. Riconosce e analizza i volti inquadrati dalla telecamera per fornire informazioni.

Codice Face attribute:

```
#import delle librerie
from deepface import DeepFace
import cv2
import matplotlib.pyplot as plt

#inserire qui il percorso dell'immagine da analizzare
img_path = 'data/img5.jpg'

#apre l'immagine
img = cv2.imread(img_path)

#la converte in spazio colore RGB per una corretta visualizzazione
img_rgb = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)

#analizza le caratteristiche utilizzando "Deepface.analyze"
obj = DeepFace.analyze(img_path, actions = ['age', 'gender', 'race', 'emotion'])
```

```
output = "age: "+str(obj[0]["age"])+" years old\nethnicity: "+str(obj[0]["dominant_race"])+"\nemotion:
"+str(obj[0]["dominant_emotion"])+"\ngender: "+str(obj[0]["dominant_gender"])

#mostra a video l'immagine e le caratteristiche
plt.imshow(img_rgb)
plt.figtext(0.5, 0.01, output, ha='center', va='bottom', fontsize=12)
plt.subplots_adjust(left=0.1, right=0.9, top=0.9, bottom=0.2)
plt.axis('off')
plt.show()
```

Esempio di utilizzo:



9. Aspetto morale e critica

Aspetti positivi dell'uso dei deepfake:
- **Intrattenimento**
I deepfake hanno trovato il loro posto anche nelle due industrie più grandi dell'intrattenimento: i videogiochi e i film.
Nel industria dei videogiochi i deepfake vocali sono stati quel metodo che ha aiutato di piu per combattere la discriminazione in base al genere e orientamento sessuale. Grazie ai skin vocali creati dai deepfake i giocatori (specialmente quelli che fanno parte di una minoranza o del genere femminile) hanno più possibilità per un gameplay piacevole rispetto ad alcuni anni fa.
Per quanto riguarda i film i deepfake vocali ma anche visivi hanno aiutato a portare di nuovo sul grande schermo personaggi importanti in saghe famose, i cui attori erano già defunti da diversi anni.

Nel 2018 il film Rogue One e nel 2019 The Rise of Skywalker, entrambi appartenenti alla saga di Star Wars, hanno usato i deepfake nelle loro tecniche di CGI e VFX per portare sullo schermo il personaggio della Principessa Leia con gli stessi aspetti che l'attrice Carrie Fisher aveva nel 1977 (periodo dell'uscita del primo film di Star Wars) e nel 2016 (l`anno della sua morte).


The rise of Skywalker (2019)


New Hope(77) vs Rouge One(18)

-**Educazione**
Un altro aspetto positivo dei deepfake è il loro uso sulle figure storiche e contemporanee importanti, rendendo le aule più interattive e coinvolgenti.
Una figura già usata per questo scopo, grazie all'AI, è stata quella dell'ex presidente americano John Kennedy e i suoi discorsi mai pronunciati per porre fine alla guerra fredda, che hanno aiutato gli studenti della storia a riconoscere il problema in un modo molto particolare.

-Assistenza sanitaria
Con gli ultimi cambi delle regole sulla privacy dentro L'Unione Europea il lavoro dei ricercatori, specialmente di quelli che fanno ricerca medica è diventato più difficile considerando che adesso per l'uso dei dati, anche a scopo di sola ricerca gli utenti/pazienti devono dare il loro consenso ogni volta, anche nei casi dei dati anonimizzati.
Qui i deepfake sono stati usati principalmente per aiutare le soluzioni di machine learning in medicina, tramite la creazione dei "pazienti artificiali". Per questi pazienti viene creato un dataset sintetico partendo dai dati reali sui quali i ricercatori hanno già accesso. In questo modo è possibile alleviare il problema della privacy e allo stesso tempo poter condividere, ai fini di ricerca dati artificiali con altri ricercatori.

Aspetti negativi dell'uso dei deepfake:

-**Pornografia**

Il primissimo caso d'uso di natura dannosa dei deepfake è stato visto nella pornografia, infliggendo violenza emotiva, reputazionale e, in alcuni casi, nei confronti dell'individuo, principalmente donne.

Secondo DeepTrace (società di sicurezza informatica) il 96% dei deepfake sono di contenuto pornografico con oltre 135 mila visualizzazioni su siti pornografici.

I video con più successo sono quelli di celebrità e quelli di revenge porn (video porno per l'umiliazione della vittima).

Più in generale, sebbene i deepfake pornografici siano invadenti per la privacy, un grande numero di comunità online sono piuttosto indifferenti. Questo può essere spiegato da diverse intuizioni comportamentali.

Primo, fintanto che un individuo ottiene un certo piacere e non vi è alcuna minaccia per i propri diritti personali, proprietà, o reputazione, l'individuo non è contro ciò di cui sta usufruendo. In secondo luogo, nel caso dei deepfake concernenti celebrità, le persone vedono ciò che vogliono sia vero, cioè la celebrità piuttosto che la persona sulla quale il volto è sovrapposto (bias di desiderabilità)

-**Influenza politica e disinformazione**

I deepfake hanno aiutato tanto nell'abbassamento della fiducia del pubblico per quanto riguarda i fatti e le notizie.

Uno dei casi più famosi di disinformazioni con gravi conseguenze politiche avvenne nel 2018, quando la popolazione di Gabon nutriva dubbi sulla salute del loro presidente, Ali Bongo, e iniziavano anche a circolare voci sulla sua presunta morte. in risposta il governo rilascia un video dove si vede Bongo che legge il discorso per Capodanno alla popolazione. Dopo una settimana, credendo che il video fosse un deepfake, i militari tentano un colpo di stato senza successo.

Non fu mai chiarito se il video fosse stato un deepfake o meno, però solo l'idea fu abbastanza per buttare un intero stato in caos per alcuni giorni.

-**Frodi e truffe**

I deepfake possono essere utilizzati per impersonare le identità di leader aziendali e dirigenti e facilitare le frodi o per la manipolazione dei mercati.

Un esempio può essere quello avvenuto nel marzo 2019, quando l'amministratore delegato di un'azienda energetica con sede nel Regno Unito, ha ricevuto una telefonata dal suo capo tedesco, che gli ordinava di trasferire 220.000 euro a un fornitore in Ungheria. L'amministratore avrebbe poi dettagliato che ha riconosciuto un "lieve accento tedesco" della voce del suo capo e ha seguito l'ordine di trasferire il denaro entro un'ora. I 220.000 euro sono stati trasferiti in Messico e incanalati su altri conti, e la società energetica, che non è stata identificata, ha denunciato l'incidente alla sua compagnia di assicurazioni. Un funzionario di quest'ultima ha affermato che i ladri hanno utilizzato l'intelligenza artificiale per creare un deepfake della voce del dirigente tedesco.

Dal punto di vista normativa in questo caso i ladri hanno anche sfruttato una "zona grigia" perché in tanti stati le leggi sulla protezione dell'identità proteggono di solito nome e immagine ma pochi stati hanno passato leggi anche sulla protezione della voce e anche in quei pochi casi le leggi garantiscono protezione solo per i vivi.

In casi come questi le situazioni non previste dalla legge permettono che i deepfake non vengano puniti e nel caso dell'azienda nel 2019 i ladri sono stati accusati solo per il furto e non per aver impersonato altre persone.

10. *Conclusioni*

I deepfake rappresentano una tecnologia straordinaria che suscita sensazioni contrastanti: sicuramente interesse ma d’altra parte preoccupazione. I contenuti manipolati che si possono creare sono estremamente veritieri e convincenti che di conseguenza hanno generato una serie di implicazioni etiche, sociali e legali. Al contempo rappresentano un grande passo avanti offrendo opportunità significative nel campo della ricerca. Il contrasto tra gli aspetti positivi e negativi dei deepfake mette in luce la necessità di affrontare queste sfide in modo **responsabile** e **ponderato**. L'implementazione di regolamentazioni e politiche volte a mitigare l'abuso dei deepfake, unitamente allo sviluppo di strumenti di rilevamento e verifica, è fondamentale per proteggere la società da potenziali danni.

11. Bibliografia

Sono stati utilizzati i seguenti siti web per ottenere informazioni:

- **Deepfakes: Trick or Treat**: https://irep.ntu.ac.uk/id/eprint/38737/1/1247050_Lee.pdf
- **Learning Self-Consistency for Deepfake Detection:** https://arxiv.org/pdf/2012.09311.pdf
- **An Introduction On Deepfake:** https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/cyber-foundry/lcf-articles/LCFArticle-Josh-Deepfakes_WEB.pdf
- **Deepfake & Cyber Intelligence:** https://www.ictsecuritymagazine.com/pubblicazioni/deepfake-cyber-intelligence/
- **Esplorazione del fenomeno dei deepfake generati dall'IA: analisi tecnica ed etica:** https://1drv.ms/b/c/d3c33ec244081988/EWXV2TEY4QVJnCQY51LstjIBI3QHgbQEyCtBhDzFe4BOEA

Per i tools e la libreria sono stati utilizzate le seguenti repository:

- **Deep Face Recognition in Python:** https://medium.com/nerd-for-tech/deep-face-recognition-in-python-41522fb47028
- **Libreria DeepFace**: https://github.com/serengil/deepface
- **DeepFaceLab:** https://github.com/iperov/DeepFaceLab