

# FACE RECOGNITION

## Percezione Umana vs Computer

Gruppo di lavoro:

Edoardo Crosera – 986874

Alessandro Di Matteo – 01708A

Alessio Colombi – 986643

### Sommario

<b><i>Percezione umana nel Face Recognition</i></b> .....	2
<b>Introduzione</b> .....	2
<b>Il sistema visivo umano impiega risorse specializzate per i volti</b> .....	2
<b>Regioni di attivazione del cervello</b> .....	2
<b>Esperimento di Zhen, Fang, Liu (2013)</b> .....	3
<b>Rete di elaborazione dei volti</b> .....	4
<b>Ruolo della risoluzione</b> .....	5
<b>Processing olistico del sistema visivo</b> .....	6
<b><i>Face Recognition basato su Computer</i></b> .....	7
<b>Estrazione delle feature</b> .....	7
Tecnica delle autofacce.....	8
<b>Reti neurali e Face recognition</b> .....	9
<b>Utilizzo delle Reti Neurali Convoluzionali (CNN) per l'Identificazione e la Segmentazione delle Immagini</b> .....	10
Identificazione degli Oggetti .....	10
Segmentazione Semantica .....	10
Segmentazione delle Istanze.....	10
<b>Altre Tecniche di Intelligenza Artificiale</b> .....	11
<b>Utilizzo delle Reti Neurali Ricorrenti (RNN) nell'Identificazione Facciale</b> .....	11
Tracciamento del Movimento del Viso.....	11
Analisi del Contesto Sequenziale .....	11
Integrazione con Altre Fasi del Processo.....	12

# Percezione umana nel Face Recognition

## Introduzione

Il volto è sempre stato uno dei tratti biometrici con i segni più distintivi per il riconoscimento dell'identità umana, usato sia dagli umani, sia dalle macchine. Questo grazie al fatto che ogni umano ha un volto con caratteristiche uniche e distintive come la disposizione degli occhi, del naso, della bocca e delle altre caratteristiche facciali.

Queste tecnologie che ci permettono di identificare le persone tramite il volto sono state possibili soprattutto grazie alle varie scoperte fatte nell'ambito umano del riconoscimento facciale, per continuare nella ricerca di quel livello di precisione che solo l'umano molte volte riesce a raggiungere.

## Il sistema visivo umano impiega risorse specializzate per i volti

Nel corso degli anni la teoria che il nostro sistema visivo impiegasse delle risorse neurali specializzate per la percezione dei volti non era mai stata dimostrata, fino a quando si sono iniziati a fare degli esperimenti sul cosiddetto "effetto inversione", dove si iniziò a capire che esistono meccanismi cognitivi e neurali unici per l'elaborazione dei volti.

L'effetto inversione è stato definito grazie a un esperimento su bambini e adulti in cui si cercava di capire che strategie applicassero per il riconoscimento del volto, con un volto invertito di 180 gradi. I risultati ci dicono che nel corso di diversi anni di vita si verifica un cambiamento di strategia. Infatti, neonati e bambini in età prescolare adottano una strategia in gran parte pezzettuale e basata sulle caratteristiche per riconoscere i volti. Gradualmente, la persona attua una strategia olistica, ovvero che considera la totalità del volto piuttosto che i singoli particolari, che è una strategia più sofisticata che coinvolge informazioni configurazionali, che ci conferma la loro importanza per la robusta performance nel riconoscimento dei volti che gli adulti mostrano.

Questo potrebbe confermare al fatto che il nostro cervello utilizzi un modulo specializzato proprio per il riconoscimento di volti.

## Regioni di attivazione del cervello

Numerosi studi di risonanza magnetica funzionale (fMRI) hanno identificato molteplici regioni corticali che mostrano una maggiore risposta ai volti rispetto agli oggetti.

Le regioni più frequentemente individuate sono nella corteccia occipito-temporale; principalmente le zone interessate sono:

- Giro fusiforme (FG), o area facciale fusiforme, (FFA)
- Giro occipitale inferiore (IOG), o area occipitale facciale, (OFA)
- Solco temporale superiore posteriore (pSTS)

Queste tre regioni si pensa che costituiscano il sistema principale per il riconoscimento dei volti: FG e IOG analizzano gli aspetti invarianti dei volti che sottendono al riconoscimento degli individui, mentre il pSTS processa gli aspetti mutevoli dei volti come la direzione dello sguardo, l'espressione facciale e i movimenti delle labbra per facilitare le comunicazioni sociali.

Inoltre, sono state individuate altre regioni come:

- amigdala e l'insula sono sintonizzate sugli aspetti emotivi dell'espressione facciale
- una regione nel solco intraparietale (IPS) è attivata quando la direzione dello sguardo sposta l'attenzione spaziale
- le regioni situate nel polo temporale (TP) e nel giro temporale medio anteriore (aMTG) sono sensibili alla familiarità dei volti
- una regione nel giro frontale inferiore (IFG) è coinvolta nell'elaborazione degli aspetti semantici dei volti
- una regione nel cortex orbitofrontale (OFC) è coinvolta nell'estrazione delle informazioni sulla bellezza facciale

Tuttavia, il riconoscimento tipico dei volti dipende non solo dalla funzionalità intatta delle singole regioni, ma anche dall'interazione dinamica tra di loro.

### Esperimento di Zhen, Fang, Liu (2013)

Lo studio ha esaminato come il cervello elabora e riconosce i volti umani utilizzando la risonanza magnetica funzionale (fMRI) su un gruppo di 42 studenti universitari, di cui 18 donne. Durante le scansioni fMRI, ai partecipanti sono stati mostrati volti umani, oggetti familiari, scene e oggetti disordinati. I ricercatori hanno utilizzato un nuovo approccio chiamato "approccio soggetto-specifico vincolato al gruppo (GSS)" per identificare automaticamente le regioni del cervello coinvolte nel riconoscimento dei volti. Sono state individuate 25 regioni del cervello coinvolte nel riconoscimento dei volti umani, distribuite in modo bilaterale e simmetrico. Queste regioni hanno dimostrato una forte affidabilità e selettività nel riconoscimento dei volti.

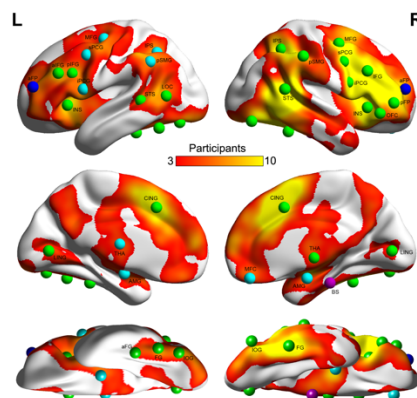


Fig.1 La figura mostra una mappa probabilistica che indica il numero di partecipanti che hanno mostrato un'attivazione selettiva per i volti, insieme alle posizioni delle regioni d'interesse a livello di gruppo.

Una volta identificate queste regioni, gli studiosi hanno esaminato come esse fossero connesse tra loro attraverso la connettività funzionale. È emerso che queste regioni formano una rete di elaborazione dei volti, caratterizzata da una struttura gerarchica e da dinamiche di rete che variano in base al compito svolto dai partecipanti, che fosse il riconoscimento dei volti o degli oggetti.

## Rete di elaborazione dei volti

La rete di elaborazione dei volti è stata analizzata in base alla connettività funzionale tra le regioni di interesse (ROI) durante un compito di riconoscimento dei volti. Le ROI sono state raggruppate in tre sub-reti relativamente indipendenti, corrispondenti a:

- riconoscimento dell'identità individuale
- recupero della conoscenza personale
- analisi dell'espressione facciale.

L'analisi di clustering ha mostrato che ci sono maggiori differenze tra la parte del cervello che aiuta a riconoscere le persone rispetto alle altre due parti coinvolte nel ricordare informazioni sulla persona e nel capire le espressioni facciali. Questo suggerisce che la parte del cervello dedicata al riconoscimento individuale sia più separata e distinta rispetto alle altre due. Questo può essere spiegato anche dalla loro distanza anatomica, ma tuttavia rimangono delle zone connesse funzionalmente che sono distanti anatomicamente, che quindi in parte scarta questa ipotesi.

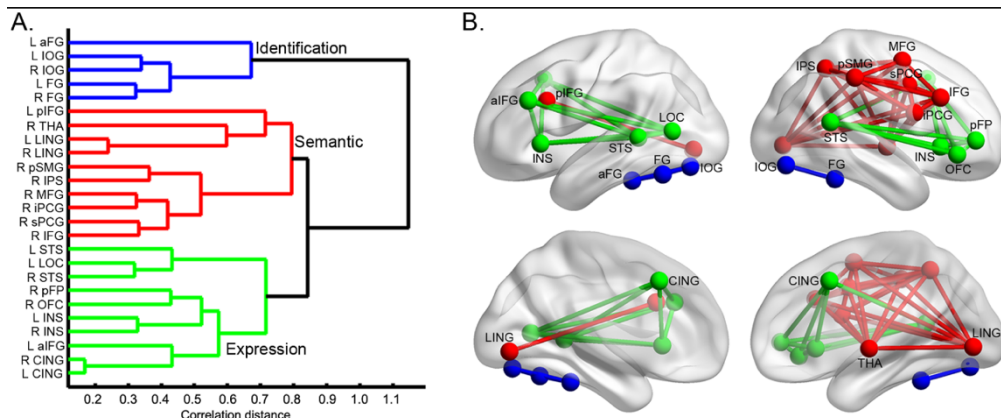


Fig.2 L'immagine mostra il dendrogramma dell'analisi di clustering gerarchico con le sub-reti colorate in base alla funzione principale. Le connessioni tra le sub-reti non sono visualizzate per chiarezza.

Il cervello quando passiamo dal riconoscere volti al riconoscere oggetti, ci sono cambiamenti significativi nella comunicazione tra le diverse parti del cervello coinvolte. Questi cambiamenti sono particolarmente evidenti nella parte del cervello che identifica i volti. In sostanza, quando cambiamo il tipo di compito che facciamo, il nostro cervello si adatta modificando la sua comunicazione interna, soprattutto nella parte che si occupa del riconoscimento dei volti.

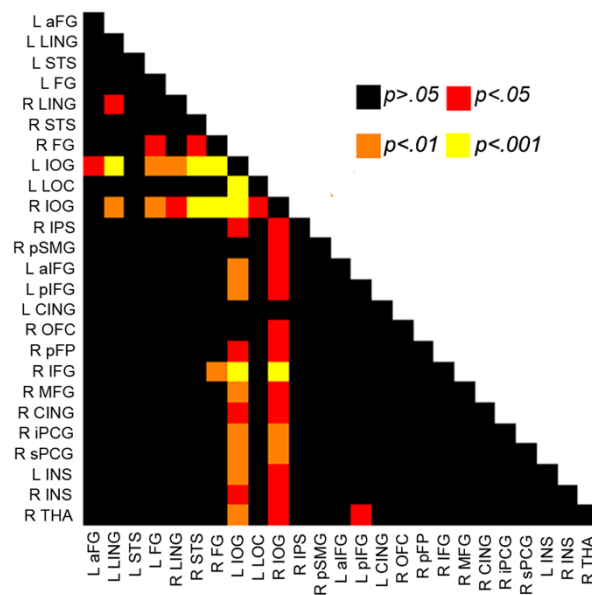


Fig.3 L'immagine illustra le modifiche nella connettività funzionale tra le regioni selettive per i volti quando i partecipanti hanno cambiato dal compito di riconoscimento dei volti a quello di riconoscimento degli oggetti. Le celle nere indicano che la variazione nella connettività funzionale non è significativa, mentre le celle colorate mostrano una significativa riduzione della connettività funzionale tra le regioni selettive per i volti.

## Ruolo della risoluzione

Al contrario delle macchine, il nostro sistema visivo riesce a riconoscere volti familiari anche con immagini a bassa risoluzione. Questo fa capire la complessità di esso, in quanto le macchine, in questo momento, non sono ancora in grado di lavorare con immagini a bassa risoluzione, in quanto lavorano sulle sottili differenze delle caratteristiche facciali.

Lavori come quello di Harmon e Julesz, e Yip e Sinha ci dimostrano che le persone riescono a riconoscere volti familiari anche con risoluzioni bassissime (16x16, 7x10, e raggiunge il picco di prestazioni a 19x27). Questo risultato dimostra almeno che dettagli caratteristici fini non sono necessari per ottenere una buona performance di riconoscimento dei volti. Tuttavia, non sappiamo ancora con certezza quali siano gli aspetti esatti per questo riconoscimento, e soprattutto come tradurli in codici per essere utilizzati dalle macchine.



## Processing olistico del sistema visivo

Spesso i volti possono essere identificati da pochissime informazioni. Sadr et al. e altri hanno dimostrato che una sola caratteristica (come gli occhi o, notevolmente, le sopracciglia) può essere sufficiente per il riconoscimento di molti volti famosi.

Tuttavia, si è notato che se si vanno a combinare la metà superiore di un volto di un soggetto 1, con la metà inferiore di un volto di un soggetto 2, risulta difficile andare ad estrapolare le due diverse identità utilizzate per creare il nuovo volto.

Questo ci fa capire che il nostro sistema visivo usa una strategia olistica, che quindi va a vedere la totalità dell'immagine, andando quindi ad influenzare l'analisi delle caratteristiche singole. Se invece si andasse a disallineare le due metà, il riconoscimento delle due identità utilizzate diventa molto più semplice.

Quindi sebbene il processo delle caratteristiche sia importante per il riconoscimento facciale, il modello ci suggerisce che il riconoscimento facciale dipende fortemente da processi olistici.



Per quanto riguarda le caratteristiche facciali, non tutte sono uguali per il riconoscimento dell'identità di un volto. Infatti, diversi esperimenti hanno indicato come caratteristiche più importanti occhi, naso, bocca, ma in modo ancora più sorprendente le sopracciglia, che secondo Sadr et al. potrebbero essere anche più importanti rispetto agli occhi.

Un loro esperimento, infatti, consisteva nel cancellare digitalmente le sopracciglia da un insieme di 50 immagini di volti di celebrità, in seguito ai soggetti sono state mostrate queste immagini singolarmente e sono stati chiesti di nominarle. Le prestazioni con le immagini prive di sopracciglia sono risultate significativamente peggiori rispetto a quelle con le immagini originali, e anche rispetto a quelle prive di occhi.

Il perché dell'importanza delle sopracciglia in questo ambito potrebbe essere ricondotto in primo luogo, al fatto che sono molto importanti per comunicare emozioni e altri segnali non verbali. Poiché il sistema visivo potrebbe già essere incline a prestare attenzione alle sopracciglia per rilevare e interpretare tali segnali, potrebbe essere che questo pregiudizio si estenda anche al compito di identificazione facciale. In secondo luogo, per una serie di motivi, le sopracciglia possono servire come una caratteristica facciale molto stabile. Poiché tendono ad essere caratteristiche facciali relativamente ad alto contrasto e di dimensioni importanti, le sopracciglia possono sopravvivere a sostanziali degradazioni dell'immagine.



## Face Recognition basato su Computer

### Estrazione delle feature

L'estrazione delle feature locali è un approccio flessibile e potente per estrarre informazioni rilevanti direttamente dai dati, senza dipendere da un modello predefinito. È particolarmente utile in situazioni in cui le caratteristiche rilevanti possono variare notevolmente da un'osservazione all'altra o quando non è disponibile un modello pre-addestrato adatto al problema specifico.

Nel riconoscimento facciale, per esempio, si hanno diversi step, ognuno dei quali ugualmente importante.

- Face detection: essenziale per l'individuazione dei volti all'interno di una scena, anche in condizioni di luce variabili o con volti di diversi colori, posizioni, scale o espressioni.
- Face segmentation: suddivide le immagini in regioni in base ai principali attributi della faccia come occhi, bocca e contorno del viso, vengono utilizzate diverse tecniche di image processing e analisi delle immagini. Ad esempio, l'analisi delle texture e dei colori può essere impiegata per identificare regioni che corrispondono a specifiche caratteristiche facciali, mentre algoritmi di rilevamento dei contorni possono essere utilizzati per delineare il contorno del viso.
- Face tracking: sfrutta le informazioni fornite dai frame successivi per prevedere lo spostamento o le deformazioni del volto nel tempo. Spesso coinvolgono l'uso di algoritmi di machine learning e di visione artificiale, come i filtri di Kalman o i modelli di apprendimento profondo (deep learning). Questi algoritmi sono in grado di analizzare i pattern di movimento del volto nel tempo e di adattarsi alle variazioni nelle espressioni facciali e nelle condizioni ambientali.
- Feature extraction: vengono identificate ed estratte le caratteristiche che contraddistinguono il soggetto utilizzando approcci come la Principal Component Analysis (PCA), la Independent Component Analysis (ICA) o altre tecniche di machine learning.
- Matching: Si possono utilizzare diverse metriche, come le distanze euclidee o algoritmi più complessi come le reti neurali, viene considerata la similitudine interclasse per il volto, che valuta quanto due volti siano simili non solo all'interno della stessa classe (ad esempio, gemelli o sosia), ma anche tra classi diverse (ad esempio, fratelli o genitori).



L'estrazione delle feature basate sul modello coinvolge l'utilizzo di un modello predefinito per estrarre le caratteristiche rilevanti dai dati. Questo approccio sfrutta le conoscenze apprese dai modelli pre-addestrati su set di dati più ampi e può essere particolarmente utile quando si dispone di un set di dati limitato o quando si desidera utilizzare conoscenze preesistenti per affrontare un nuovo problema.

## Tecnica delle autofacce

Nel campo informatico si è sempre cercato di capire come far riconoscere alle macchine che oggetto è rappresentato in una foto, soprattutto per applicazioni come il face recognition, dove dobbiamo avere anche una certa qualità e accuratezza nel riconoscimento.

Una prima possibilità consisteva nel considerare la matrice come un dettaglio ad alta dimensionalità e inferire da essa un vettore di informazioni a dimensioni inferiori, per poi cercare di riconoscere la persona in dimensioni inferiori. Ma questo fu usato principalmente quando i computer non erano ancora in grado di elaborare molti dati come ora.

Nel 1987, un articolo di Sirovich e Kirby ha provato a ipotizzare che le immagini dei volti umani siano una somma ponderata di alcune “immagini chiave”, chiamate autofacce, chiamate così in quanto sono gli autovettori della matrice di covarianza delle immagini di volti umani sottratte alla media.

Nell'algoritmo illustrato da Sirovich e Kirby, poi sviluppato ulteriormente da Turk e Pentland, vediamo coinvolto anche l'analisi per componenti principali (PCA), ovvero una tecnica statistica che ci permette di diminuire la dimensionalità dei dati per mantenere le caratteristiche più significative, ovvero le feature più distintive per il face recognition.

Nella pratica questa analisi implica un calcolo della matrice di covarianza, in quanto i dati sono sottoforma di matrice, misurando così le relazioni tra i diversi dati. In seguito, si applica una decomposizione spettrale dove si trovano:

- gli autovettori, i quali sono le direzioni nello spazio dei dati dove abbiamo le variazioni massime
- gli autovalori, i quali rappresentano l'importanza relativa di ciascuna di quelle direzioni

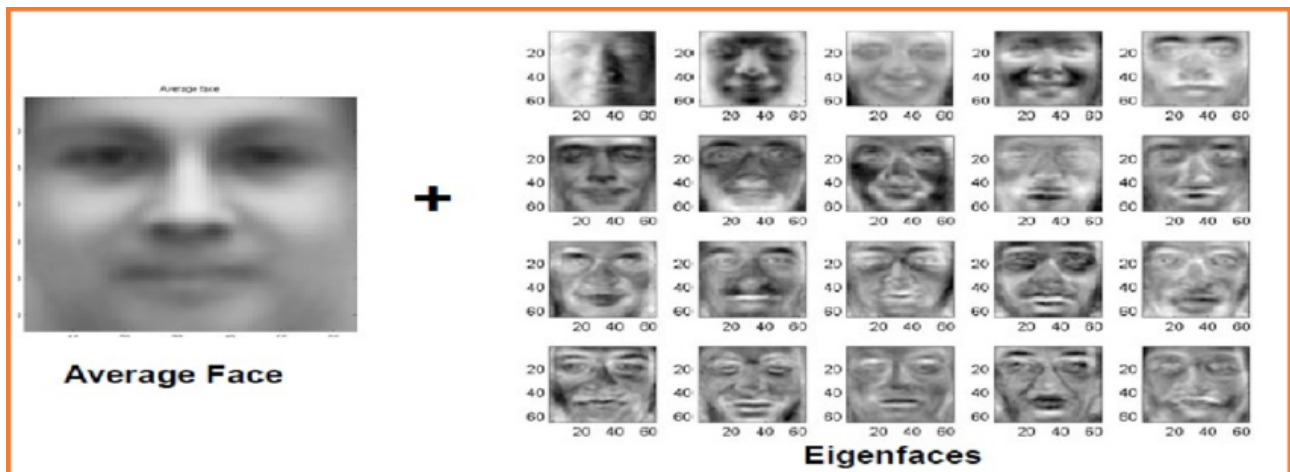
Una volta trovati autovettori e autovalori, andiamo a prendere le prime  $k$  componenti principali, in modo da poter diminuire la dimensionalità nello spazio dei dati.

Ora, che abbiamo capito il funzionamento del PCA, possiamo elencare i passaggi per il funzionamento delle autofacce:

1. Raccolta immagini, dove raccolgo il mio dataset, dove le immagini sono trasformate in matrici di pixel, le quali devono essere in scala di grigi e devono avere le stesse dimensioni.
2. Calcolo della media, calcolo la media di tutte le immagini così da avere il “volto medio” del dataset.
3. Sottrazione della media, la quale è un passaggio fondamentale, in quanto andiamo a normalizzare le immagini eliminando le variazioni dovute a luce, angolazioni e altre differenze non essenziali all'interno delle immagini.

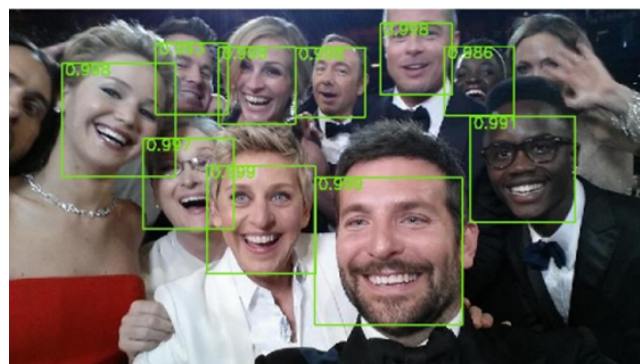


4. Calcolo delle autofacce, dove andiamo ad utilizzare l'algoritmo PCA illustrato prima, dove in questo caso gli autovettori sono chiamati autofacce.
5. Rappresentazione delle immagini, dove ogni immagine all'interno del dataset può essere rappresentata da una combinazione lineare delle autofacce, espressa come un vettore di pesi.
6. Riconoscimento facciale, dove per riconoscere un volto, andiamo a calcolare il vettore dei pesi per il volto da riconoscere, successivamente si va a fare un confronto tra il vettore calcolato e i vettori del dataset e si va a classificare in base alla somiglianza di questi due vettori.



## Reti neurali e Face recognition

L'identificazione facciale rappresenta una delle più rivoluzionarie applicazioni dell'intelligenza artificiale nell'era moderna. Grazie all'avanzamento delle reti neurali, in particolare delle reti neurali convoluzionali (CNN) e delle reti neurali ricorrenti (RNN), i sistemi di riconoscimento facciale hanno raggiunto nuovi livelli di precisione e affidabilità. Questi sistemi sono in grado di estrarre, elaborare e confrontare con estrema precisione le caratteristiche facciali, aprendo così una vasta gamma di possibilità in settori come la sicurezza, la sorveglianza, l'accesso agli edifici e la personalizzazione dell'esperienza utente.



Durante la fase di feature extraction, le CNN analizzano l'immagine facciale in input attraverso una serie di filtri convoluzionali, ciascuno progettato per rilevare specifiche caratteristiche visive. Ad esempio, uno strato convoluzionale potrebbe individuare i contorni del viso, mentre uno strato successivo potrebbe concentrarsi sui dettagli degli occhi o della bocca. Questo processo di estrazione delle caratteristiche è fondamentale per ridurre la complessità dell'immagine e catturare le informazioni più rilevanti per l'identificazione o la verifica.

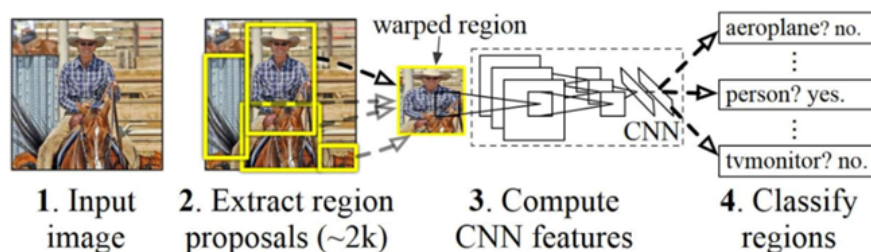
Le CNN sono particolarmente adatte per questa fase poiché mantengono la località delle informazioni all'interno dell'immagine, consentendo loro di identificare pattern spaziali complessi come bordi, texture e forme. Inoltre, il processo di apprendimento automatico delle CNN consente loro di adattarsi dinamicamente a una vasta gamma di condizioni di illuminazione, posizioni del viso e variazioni nell'aspetto individuale. Questa capacità di adattamento e generalizzazione rende le CNN strumenti potenti per l'identificazione facciale in una varietà di contesti, dall'accesso sicuro ai dispositivi alla sorveglianza pubblica.

## Utilizzo delle Reti Neurali Convoluzionali (CNN) per l'Identificazione e la Segmentazione delle Immagini

Nel campo dell'elaborazione delle immagini, le reti neurali convoluzionali (CNN) hanno rivoluzionato l'approccio all'identificazione degli oggetti e alla segmentazione delle immagini. Le CNN, ispirate al funzionamento del cervello umano, sono progettate per catturare e apprendere pattern spaziali complessi, rendendole particolarmente adatte per affrontare sfide come l'identificazione degli oggetti, la segmentazione semantica e l'individuazione delle istanze all'interno di un'immagine.

### Identificazione degli Oggetti

Nel contesto dell'identificazione degli oggetti, le CNN giocano un ruolo fondamentale nel rilevare e classificare gli oggetti presenti in un'immagine. Con architetture come YOLO (You Only Look Once), le CNN possono analizzare un'intera immagine in una sola volta, generando predizioni dettagliate su posizione, classe e confidenza degli oggetti individuati. Questo approccio consente una rapida e precisa identificazione degli oggetti, rendendolo ideale per applicazioni di sorveglianza, monitoraggio e analisi delle immagini.



### Segmentazione Semantica

La segmentazione semantica è un'altra area in cui le CNN si distinguono. Utilizzando tecniche come le reti deconvoluzionali, le CNN sono in grado di analizzare e suddividere un'immagine in regioni distinte in base al loro significato semantico. Questo processo consente di distinguere tra diverse classi di oggetti all'interno dell'immagine, consentendo una comprensione più approfondita del suo contenuto. Ad esempio, in un'immagine di una strada, la segmentazione semantica potrebbe distinguere tra strada, marciapiede, auto e pedoni, consentendo un'analisi più dettagliata dell'ambiente circostante.

### Segmentazione delle Istanze

Infine, la segmentazione delle istanze estende ulteriormente la capacità delle CNN distinguendo non solo le classi di oggetti, ma anche le singole istanze di tali classi. Questo significa che le CNN possono non solo rilevare la presenza di persone in un'immagine, ma anche distinguere tra diverse persone presenti e tracciare le loro posizioni individuali. Questa capacità è cruciale in

applicazioni come il conteggio delle persone, il tracciamento del movimento e la sorveglianza di luoghi affollati.

In sintesi, l'utilizzo delle reti neurali convoluzionali (CNN) ha rivoluzionato l'identificazione e la segmentazione delle immagini, consentendo una vasta gamma di applicazioni in settori come la sicurezza, l'automazione industriale, la medicina e molto altro ancora. La loro capacità di apprendere pattern complessi e di adattarsi a una vasta gamma di dati visivi le rende uno strumento potente per esplorare e comprendere il mondo visivo che ci circonda.

## Altre Tecniche di Intelligenza Artificiale

Oltre alle CNN, esistono altri approcci di intelligenza artificiale utilizzati nell'identificazione facciale:

- Support Vector Machines (SVM): Modello di apprendimento supervisionato noto per la sua efficacia nel separare e classificare dati complessi, adatto per situazioni in cui la separazione tra classi è non lineare.
- Reti Neurali Generative (GAN): Utilizzate per generare dati sintetici, le GAN possono essere impiegate per migliorare la robustezza e l'affidabilità dei sistemi di riconoscimento facciale generando volti sintetici da utilizzare come dati di addestramento.
- Reti Neurali Ricorrenti (RNN): Progettate per lavorare con dati sequenziali, le RNN possono essere integrate in sistemi di identificazione facciale per gestire dati sequenziali correlati, ad esempio per il tracciamento del movimento del viso nel tempo.

L'integrazione di diverse tecniche di intelligenza artificiale consente di sviluppare sistemi di identificazione facciale più robusti e affidabili, in grado di gestire una vasta gamma di scenari e condizioni operative.

## Utilizzo delle Reti Neurali Ricorrenti (RNN) nell'Identificazione Facciale

Le reti neurali ricorrenti (RNN) sono un tipo di rete neurale progettata per lavorare con dati sequenziali, come sequenze temporali o dati testuali. Sebbene siano meno comunemente utilizzate nelle fasi primarie di identificazione facciale rispetto alle CNN, le RNN possono giocare un ruolo significativo in determinati aspetti del processo.

### Tracciamento del Movimento del Viso

Una delle aree in cui le RNN possono essere integrate è nel tracciamento del movimento del viso nel tempo. Poiché le immagini di volti possono variare notevolmente a causa di cambiamenti nell'illuminazione, nell'angolazione e nelle espressioni facciali, è utile avere sistemi in grado di comprendere e seguire tali variazioni nel tempo. Le RNN possono essere addestrate su sequenze di immagini facciali per prevedere e tracciare il movimento del viso nel tempo, consentendo una migliore comprensione della dinamica facciale.

### Analisi del Contesto Sequenziale

In contesti in cui è importante considerare il contesto sequenziale delle immagini, ad esempio in sistemi di sorveglianza o monitoraggio in tempo reale, le RNN possono essere utilizzate per analizzare e interpretare le sequenze di immagini facciali. Questo può includere il riconoscimento di azioni o comportamenti specifici basati su sequenze di espressioni facciali o movimenti del viso nel tempo. Le RNN sono in grado di catturare le dipendenze temporali nei dati, consentendo

una migliore comprensione del contesto e una risposta più intelligente ai cambiamenti nell'ambiente.

### Integrazione con Altre Fasi del Processo

Le RNN possono essere integrate in sistemi più complessi di identificazione facciale insieme alle CNN e ad altre tecniche. Ad esempio, possono essere utilizzate per analizzare sequenze di feature estratte dalle CNN per individuare pattern temporali significativi nel movimento del viso o nelle espressioni facciali. Questa integrazione consente una visione più completa e dinamica delle informazioni facciali, migliorando la capacità di identificare e seguire i volti in vari contesti.

In sintesi, mentre le CNN dominano nelle fasi di estrazione delle caratteristiche e di identificazione degli oggetti, le RNN sono preziose per analizzare il contesto temporale e sequenziale delle immagini facciali, contribuendo a una comprensione più approfondita e dinamica del viso umano. La loro capacità di lavorare con dati sequenziali le rende un complemento utile alle CNN nell'identificazione facciale.