Complementi di biometria

University of Milan
Master's Degree in Computer Science
A.Y. 2022/2023

Stefano Vallodoro

Indice

1 Introduzione

La biometria è l'insieme di tecniche automatiche per il riconoscimento degli individui basato sulle loro caratteristiche fisiche e comportamentali.

Il riconoscimento biometrico può essere suddiviso in due categorie:

- Verifica dell'identità: Autenticazione, si conferma o si nega l'identità dichiarata dall'utente (1:1)
- Ricerca dell'identità: Identificazione (1:N)

I metodi di autenticazione si basano su due principali modalità: possesso (token based) e conoscenza. Nei metodi biometrici invece il riconoscimento avviene in base a caratteristiche fisiche e/o comportamentali dell'individuo

Vantaggi metodi biometrici

- Solo i metodi biometrici possono realizzare una identificazione/autenticazione negativa (il sistema dice che io non sono lui)
- Riducono la possibilità di reclami di ripudiazione

Svantaggi metodi biometrici

- Hanno un costo maggiore
- Alcune persone li vedono come un'invasione della propria privacy
- Rispondono con un livello di matching e non con una decisione binaria

Le 7 proprietà del tratto biometrico

1. *Universalità*, ogni persona deve possedere questo tratto o caratteristica
2. *Unicità*, due persone non devono avere lo stesso tratto uguale
3. *Permanenza*, deve essere invariante nel tempo
4. *Misurabilità*, il tratto deve poter essere esaminato quantitativamente. Devo poter estrarre dei dati quantitativi
5. *Performabilità*
6. *Accettabilità*
7. *Circonvenzione*, grado di difficoltà nell'ingannare il sistema

Enrollment e riconoscimento

- *Enrollment*: Il tratto biometrico viene per la prima volta acquisito dal sistema e registrato. Da questo posso estrarre le caratteristiche che vengono codificate in un template e salvate in un database
- *Riconoscimento* (Identificazione/Verifica): Il tratto biometrico viene nuovamente acquisito, se risulta sufficientemente aderente alle informazioni registrate nel sistema biometrico l'accesso è consentito. Vengono ripetute le azioni viste in enrollment (acquisizione, feature extraction e coding). Successivamente avverrà il matching fra il template creato e quello presente nel database

  L'utente dichiara la propria identità che è sfruttata dal DB per andare a prendere solo il template associato a quelle informazioni.

Impronta L'impronta digitale può essere acquisita tramite sensori termici, ottici, scanner tradizionali. Il riconoscimento avviene attraverso tre approcci:

1. *Correlation-based*, correlazione pixel per pixel. Trasla le immagini
2. *Ridge feature-based*, orientamento, distribuzione e posizione dei ridge
3. *Minutia-based*, dove terminano i ridge e dove ci sono biforcazioni

Il sample della impronta è un'immagine in toni di grigio. Le impronte si dividono in base ai core e delta e all'orientamento dei ridge in *Arch, Loop, Whorl*

Un'impronta può essere esaminata su 3 livelli:

- *Livello I*, pattern generale dell'impronta

  Ridge counting: Misura dei ridge che attraversano una linea immaginaria passante tra due minutiae

  Filtro di Gabor: Toglie il rumore, esalta i ridge aumentando il contrasto

  FingerCode: Lo utilizziamo quando si lavora a livello 1 e non si hanno le minutiae ma solo gli orientamenti. Genera un codice univoco per ogni impronta digitale
- *Livello II*, minutiae. Terminazioni e biforcazioni.
- *Livello III*, posizione di pori

Volto Tratto biometrico tra i meno intrusivi, usato normalmente dalle persone per riconoscersi. Il viso può essere acquisito tramite telecamere, webcam, fotocamere, smartphone, scanner 3D. Il matching avviene tramite due approcci:

- *Eigenfaces*, l'autovettore della matrice delle covarianze fra molti volti da allenamento
- *Attributi*, si misurano delle caratteristiche come la distanza fra gli occhi, la lunghezza del naso

Mano Molto ben accettato dagli utenti perchè poco invasivo. Le mani possono essere riconosciute tramite i sensori come CCD di tipo visibile o infrarosso. Gli algoritmi utilizzati rilevano i contorni e la forma o analizzano le vene della mano

Iride L'iride è considerato come il tratto biometrico più accurato, dopo il DNA, e performante anche se poco accettato perchè considerato invasivo. I sensori utilizzati possono essere CCD ad alta definizione di tipo visibile o infrarosso oppure ottiche speciali

Le feature che maggiormente sono interessanti per il sistema biometrico si vedono meglio con luce IR piuttosto che con luce visibile

- Il sistema di acquisizione deve poter risolvere con almeno 70 pixel il raggio dell'iride
- Si usano CCD capaci di acquisire nel vicino infrarosso (NIR)
- E' necessario usare telecamere con ottiche variabili per trovare l'occhio nel volto e poi zoomare verso l'occhio
- Se manca più del 50% dell'iride occorre acquisire ancora una volta l'iride
- Di solito bastano 256 byte per rappresentare una iride, IRISCODE
- La comparazione è effettuata fra Iriscode di 256 byte attraverso il calcolo della distanza di Hamming, è uno XOR. Se prendo iridi di persone diverse e le confronto, sono talmente diverse che è come confrontare stringhe di bit casuali

L'exploit di Daugman, riconoscimento con iride da foto ad alta risoluzione nel visibile a 18 anni di distanza, mostra la possibilità del pericolo di screening di massa dagli archivi di foto (governativi, social, ...). Enorme problema di privacy nel futuro

Firma La firma è un metodo molto diffuso e semplice ma ha una bassa accuratezza. La variabilità della firma è molto alta. Il riconoscimento è basato su coordinate x, y, pressione, inclinazione della penna.

Difficilmente falsificabile, nel caso di firma online, ma elevata similitudine interclasse con le firme brevi o troppo semplici

Voce

- *Speech recognition*, riconoscimento di quello che si è detto
- *Speaker recognition*, riconoscimento della persona che sta parlando

Sistemi multimodali Più tecnologie biometriche in un sistema. Posso utilizzare tratti biometrici completamente diversi, impressioni multiple, estrarre informazioni con due tecniche diverse, utilizzare sensori diversi

Soft biometrics Alcuni tratti biometrici non posseggono le 7 caratteristiche necessarie quindi non permettono di identificare in modo univoco una persona, ma di caratterizzarla o di rendere più robusto il sistema. Ad esempio: genere, colore della pelle, colore degli occhi, peso, altezza

Variabilità intraclasse Si intende la variazione del sample o delle feature dello stesso individuo. Questa variazione può essere dovuta a rumore, variazione dello sfondo, variazioni del tratto, parziali occlusioni

Variabilità interclasse Variazione del sample o delle feature acquisiti da individui diversi

Acquisizione La cura nel processo di acquisizione influenza pesantemente l'accuratezza finale del sistema. Il sample deve essere di buona qualità per poter estrarre le caratteristiche e per salvarlo nel db, i sistemi di controllo della qualità producono un indice di qualità del sample acquisito. Se l'indice è sufficientemente alto si può procedere all'estrazione delle caratteristiche, altrimenti si richiede una nuova acquisizione.

Ad esempio per il volto si fa riferimento alle regole ICAO.

Il processo di acquisizione si suddivide in due fasi:

- *Valutazione della qualità*
- *Segmentazione*, selezionare la regione di interesse dell'immagine acquisita

Rappresentazione Visualizzazione del problema della rappresentazione in uno spazio delle feature N-dimensionale. Vogliamo che si sia *bassa variabilità intraclasse* e *alta variabilità interclasse*

Genuini e impostori

- *Genuino*, indica un individuo che accede al sistema e ha titolo per farlo
- *Impostore*, chi prova ad accedere senza averne titolo

Problema della verifica Dato in ingresso (query) un insieme di caratteristiche X_Q e la dichiarata identità I occorre determinare se (I, X_Q) appartengono a

w_1(genuino) o w_2(impostore). Tipicamente le caratteristiche X_Q vengono confrontate con quelle X_I, il template memorizzato nel sistema associato all'identità I. Si tratta di una comparazione con soglia

$$(I, X_Q) \in \begin{cases} w_1 & \text{se } S(X_Q, X_I) \geq T \\ w_2 & \text{altrimenti} \end{cases}$$

Dove S è il similarity score o match score che misura la similitudine tra X_Q e X_I e T è la soglia prefissata

Problema dell'identificazione Il sistema controlla se i tuoi dati biometrici corrispondono ad un insieme di identità registrate. Dato in ingresso (query) un insieme di caratteristiche X_Q, devo determinare l'identità I_k con k appartenente all'insieme $1, 2, 3, \ldots, M, M+1$

Dove $1, 2, 3, \ldots, M$ sono le M identità registrate nel sistema e $M+1$ rappresenta il caso si *reiezione*. Si tratta di M comparazioni con soglia

Distanza fra i template I template non sono mai uguali. Esiste sempre una distanza nello spazio delle feature che separa i template anche della stessa persona. Se si riscontrasse una distanza fra X_Q e X_I nulla (il valore di S è il massimo valore ammissibile) probabilmente saremmo di fronte ad un replay attack, ovvero una copia illecita di un template memorizzato che viene riproposto in ingresso per frodare il sistema.

FM e FNM

- *False match, errore di tipo I*, l'impostor score è maggiore della soglia T impostata. Il ladro entra in casa perché il sistema biometrico lo ha scambiato per voi
- *False non match, errore di tipo II*, il genuine score è minore della soglia T impostata. Voi non entrate in casa

Andando a rapportare questi valori rispettivamente con il totale dei genuini e il totale degli impostori otterremo i tassi FMR(T) e FNMR(T), che variano in funzione della soglia T scelta.

DET e ROC Per descrivere le performance del sistema è necessario disporre di un insieme di dati e curve di funzionamento

Le performance vengono espresse mostrando come variano i tassi di errore al variare di T. Posso fare un plot, dove ogni punto corrisponde al valore che ha come x il valore di FMR e come y il valore di FNMR. Con la soglia a $-\infty$ entrano anche gli impostori, con la soglia a $+\infty$ non entra nessuno. Sicuramente la curva passa da i punti $(1, 0)$ e $(0, 1)$ che rappresentano i casi estremi.

La curva DET e la curva ROC mostrano le stesse informazioni. Con un sistema ideale collassano sugli assi. Regolando la soglia T possiamo regolare il livello di sicurezza ed individuare le regioni di funzionamento.

EER L'Equel Error Rate è il tasso di errore corrispondente all'unico punto nel quale abbiamo FNMR = FMR. E' l'unico numero singolo che può riassumere il funzionamento del sistema

Scalabilità I sistemi che devono gestire una grande quantità di identità dovrebbero essere in grado di operare efficacemente quando il numero di utenti registrati nel DB aumenta. In più si richiede che il tasso di peggioramento delle prestazioni sia minore del tasso di nuovi utenti immesso

L'obiettivo di gestire efficacemente la complessità delle ricerche rispetto all'incremento del numero di template nel DB del sistema può essere raggiunto solo con un'attenta organizzazione dei DB. Un DB organizzato permette di non confrontare un template in ingresso con tutti i template ma solo con quelli contenuti in una partizione

Quando il DB viene creato, i template vengono disposti nelle partizioni (bins). Si ha *binning error* quando un individuo presenta i propri tratti biometrici al sistema e l'algoritmo di classificazione del tratto sbaglia il bin.

1.1 Reporting data

Dobbiamo fare il *report* dei dati e delle distribuzioni analizzate

1. Curva DET
2. EER
3. CMC, Cumulative Match Characteristic
4. Probabilità d'errore p
5. FTE, FTA
6. Attacchi
7. Performance results

Analizzare curva DET Esistono due diverse strategie per calcolare e analizzare DET

- *Inferenza statistica*, si inducono le caratteristiche di una popolazione dall'osservazione di una parte di essa detta campione
- *Calcolo delle probabilità*, si conoscono le curve

Supponiamo di poter variare la soglia s e di fissarla ad un valore T in mezzo fra il picco degli impostori e quello dei genuini. Un certo numero di persone appartenenti al gruppo dei genuini sono sotto la soglia T e quindi non saranno autorizzati dando luogo ad errori di False Non-Match (FNM). Al contrario una parte degli impostori hanno valori di match sopra la soglia quindi saranno autorizzati. Nel caso di indentificazione positiva i tassi vengono ripettivamente chiamati FAR, FNAR

Regola dei 3 Il tasso di errore p per il quale si ha la probabilità di zero errori in N prove è circa $p = \frac{3}{N}$, per un intervallo di confidenza del 95%. Un altro modo di leggere la cosa: se abbiamo un sistema che commette zero errori su N prove non dobbiamo pensare di avere un sistema con $p = 0$, ma con il 95% di confidenza abbiamo un sistema che ha $p = \frac{3}{N}$

Regola dei 30 Per essere sicuro con intervallo di confidenza del 90% che il tasso di errore vero sia tra il $\pm 30\%$ del tasso di errore osservato, ci devono essere almeno 30 errori

Failure To Enroll rate E' la percentuale della popolazione per i quali il sistema non è in grado di generare templates. Quanti utenti in media non riescono a completare la fase di enrollment

Failure To Acquire rate E' la percentuale di transazioni per le quali il sistema non è in grado di acquisire o individuare un'immagine di qualità sufficiente. Utenti che hanno problemi sul modulo di acquisizione dovuto magari a tratti biometrici rovinati o non validi

Attacchi Gli attacchi possono essere

- *Zero-Effort attempts*
- *Presentation Attack*, l'impostore crea dei falsi o artefatti

2 Image acquisition

I sensori si dividono principalmente in due famiglie: *CCD* e *CMOS*. Entrambi sono costituiti da una matrice di fotosensori

CCD: Charge-Coupled Device E' un sensore che ha un funzionamento simile ad un *fotodiodo*. Un fotodiodo è un dispositivo semiconduttore che converte la luce in corrente elettrica. Acquisisce luce elettricamente, immagazzina la carica in ogni pixel e poi legge la carica dei pixel per creare l'immagine. È abbastanza costoso ma con qualità altissima

EMCCD: Electron-Multiplying CCD Variazione del CCD in cui la carica accumulata in ogni pixel viene amplificata più volte attraverso un moltiplicatore di elettroni, prima di essere letta. Ciò aumenta la sensibilità del sensore e riduce il rumore del segnale

CMOS: Complementary Metal-Oxide Semiconductor Ogni singolo fotodiodo è accoppiato ad un convertitore, riduttore di rumore, e circuiti di digitalizzazione. Permette di gestire meglio il singolo pixel

Non ci sono 3 sensori uno per canale (RGB) ma un solo sensore e ogni singolo pixel può vedere o blu o verde o rosso

Shutter Componente tipico di una fotocamera, sia meccanica che elettronica. Il modo in cui il sensore della fotocamera legge il segnale di determinati pixel può essere diverso

Rolling shutter Il rolling shutter acquisisce l'immagine scansionando il sensore una linea alla volta, dal lato sinistro a quello destro. Ciò significa che l'immagine non viene acquisita istantaneamente, ma viene registrata in modo sequenziale.

Gli effetti negativi del rolling shutter sono:

- *Wobble-jello effect*, effetto gelatina
- *Skew*, l'immagine si piega diagonalmente in una direzione o un'altra
- *Spatial Aliasing*, pixel verticali adiacenti si spostano
- *Temporal Aliasing*, quando un'immagine che contiene un movimento viene catturata con una frequenza di campionamento troppo bassa

Global shutter Tutte le righe di pixel dell'immagine sono acquisite nello stesso istante, senza differenze di tempo tra di esse.

Le fotocamere CCD utilizzano tipicamente shutter globali. Nonostante questa modalità di shutter non abbia differenze temporali sull'immagine, la lettura è tipicamente lenta a causa dell'avere solo un ADC.

Filtri ottici Strumento che trasmette selettivamente la luce con particolari proprietà come: una o più lunghezza d'onda (colore); polarizzazione, selezionando solo le onde luminose che vibrano in una determinata direzione, bloccando quelle che vibrano in altre direzioni; attenua l'intensità.

Ad esempio nel day/night surveillance hanno un filtro che permette di tagliare l'infrarosso in modo tale da migliorare la visualizzazione notturna, di giorno viene inserito per poter filtrare la luce IR

Multispectral imaging è la scansione della superficie e della sub-superficie con diverse lunghezze d'onda e colori da diverse angolazioni fino ad una profondità di 4 mm

Focalizzazione Cosa significa a fuoco per un sensore digitale? Che i raggi di 1 punto arrivino almeno nello stesso pixel

Profondità di campo Range delle distanze dell'oggetto su cui l'immagine è sufficientemente ben focalizzata. Una maggiore profondità di campo implica che una vasta gamma di distanze nell'immagine appaia nitida, mentre una profondità di campo più ridotta comporta che solo un'area limitata dell'immagine sarà a fuoco.

Diaframma Man mano che il diaframma si chiude, passeranno sempre meno raggi all'interno della camera

Tempo di esposizione Il tempo di esposizione si riferisce alla durata in cui il sensore dell'immagine viene esposto alla luce durante lo scatto di una foto. Un tempo di esposizione più lungo permette di catturare più luce e quindi immagini più luminose, mentre un tempo di esposizione più breve può catturare meno luce, ma può consentire di congelare i movimenti o di ottenere un'immagine più nitida

WDR *Wide Dynamic Range* è un algoritmo che combina più esposizioni di un'immagine a diverse lunghezze di esposizione, segue poi la fusione dell'immagine che mostra dettagli sia nelle aree più luminose che in quelle più scure. In questo modo, il WDR consente di ottenere immagini più bilanciate e dettagliate anche in situazioni di contrasto estremo

HDR *High Dynamic Range* è una tecnologia utilizzata per gestire un'ampia gamma di luminosità. L'HDR combina più immagini scattate con diverse esposizioni, permettendo di avere un'immagine finale con maggiori dettagli sia nelle zone di ombra che nelle zone di luce. In pratica, l'HDR crea un'immagine finale

che ha una gamma dinamica più ampia rispetto a quella che l'occhio umano può percepire

Macro In fotografia sono degli obiettivi che ci permettono di avvicinarci al soggetto a tal punto che ci sia un rapporto 1:1 tra il soggetto fotografato e la dimensione dello stesso proiettata sul sensore

Frame rate Il framerate è la frequenza di cattura o riproduzione dei fotogrammi che compongono un filmato

2.1 Lenti

Le lenti possono essere utilizzate in diversi sistemi di visione per migliorare la qualità dell'immagine, regolare la quantità di luce che entra nel sistema e controllare la profondità di campo. Senza lenti, la luce entrerebbe in modo non controllato, producendo immagini poco chiare e fuori fuoco.

Problemi delle lenti I problemi che possiamo avere utilizzando le lenti sono:

- *Vignettatura*, si notano i lati della foto scuri
- *Compound thick lens*, quando una lente è troppo spessa e quindi la sua curvatura non è sufficiente per correggere completamente l'aberrazione sferica e l'aberrazione cromatica.
- *Aberrazione sferica*, quando le onde di luce provenienti dal centro della lente sono focalizzate in modo diverso rispetto alle onde provenienti dai bordi della lente, causando un effetto di sfocatura dell'immagine
- *Aberrazione cromatica*, quando le lenti non sono in grado di focalizzare tutti i colori dello spettro luminoso nello stesso punto. I colori arrivano su pixel diversi. La soluzione è quella di aggiungere un'altra lente con le cavità opposte in modo tale che l'effetto venga annullato con un'altra aberrazione cromatica

Lente asferica Una lente che ha una forma non sferica ma è progettata per correggere l'aberrazione sferica

Lenti liquide Composta da due *fluidi isodensi* (uno più denso e uno meno denso) contenuti al suo interno ne modificano la curvatura a seconda della tensione elettrica che li attraversa.

Sono composte da due superfici in vetro o plastica trasparente e un fluido trasparente e immiscibile, come l'olio. Quando il fluido viene spostato all'interno della lente, la superficie della lente si curva, producendo un cambiamento nel

potere diottrico della lente. Consentono di regolare rapidamente la messa a fuoco senza dover spostare fisicamente l'obiettivo della fotocamera

L'uso di lenti liquide permette oltre alla messa a fuoco, anche magnificazione ottica in uno spazio molto compatto

Magnificazione La magnificazione si riferisce all'aumento dell'immagine di un oggetto ed è definita come il rapporto tra la dimensione dell'immagine e la dimensione dell'oggetto.

Con una *singola lente*, la magnificazione può essere calcolata in base alla distanza focale della lente e alla distanza dell'oggetto dalla lente stessa. Con *due lenti*, la magnificazione può essere aumentata utilizzando un'oculare. In questo caso, la prima lente crea un'immagine dell'oggetto, che viene poi ingrandita ulteriormente dalla seconda lente, l'oculare.

Magnificazione variabile In ottica un obiettivo zoom è un obiettivo complesso la cui lunghezza focale può variare. Lo zoom può essere ottico oppure digitale

- *Zoom ottico*, una lente (lente di fuoco) che avrà l'obiettivo di mettere a fuoco e un insieme di lenti chiamate sistema di lenti afocali che saranno utilizzate per effettuare la magnificazione variabile
- *Zoom digitale*, solamente lo zoom ottico ingrandisce veramente, quello digitale scala l'immagine senza aggiungere informazioni allo scatto

3 Dataset partitioning

Non bisogna usare tutti i dati per l'allenamento, devono essere correttamente usati per evitare bias nel risultato. Ci sono due principali problemi che possono sorgere utilizzando gli stessi dati per addestrare e valutare il modello

- *Underfitting*, il modello è poco adeguato ai dati
- *Overfitting*, il modello offre alte performance con i dati noti ma si comporta male con i dati mai visti

k-Fold Cross Validation E' una specifica tecnica di cross-validation in cui i dati di training vengono suddivisi in k parti uguali, il modello viene addestrato su $k-1$ parti e testato sulla parte rimanente. Questo processo viene ripetuto k volte, in modo che ogni parte sia usata come set di test una volta. Alla fine, si calcolano le medie delle prestazioni del modello sui k set di test per ottenere una stima della performance siu dati non visti

Stratified k-FCV Inserisce un numero uguale di campioni di ogni classe su ogni partizione in modo da mantenere la distribuzione delle classi uguali in tutte le partizioni, in modo da evitare che un particolare fold contenga solo esempi di una sola classe.

5x2 Cross Validation L'intero set di dati è diviso in modo random in due subset A e B. Il modello viene prima costruito utilizzando A e validato utilizzando il subset B. Successivamente, il processo viene invertivo

Leave one out Prevede di utilizzare ogni singola osservazione come set di test, e il resto del dataset come set di training. Per ogni si rimuove una sola osservazione dal dataset e si addestra il modello sulle rimanenti.

4 Apprendimento induttivo

Machine Learning in biometria Il sistema biometrico è un *classificatore*

Nel modo tradizionale abbiamo una serie di if-then per andare al risultato. In ML invece abbiamo un modello generato da una prima fase di learning contenente dei dati di allenamento e dei risultati che ci si aspetta.

Le tre componenti del Machine Learning

- *Rappresentazione*, spazio delle ipotesi
- *Valutazione*, come decidiamo se il modello sta lavorando bene
- *Ottimizzazione*, modificare i parametri in modo da minimizzare l'errore

Deductive reasoning *Top-down*, parto dalla teoria

Inductive reasoning *Bottom-up*, parto dalle osservazioni. Principalmente le reti neurali e il machine learning si basano sull'indittuvità imparando dall'esperienza che matura nel tempo

Tipologie di apprendimento

- *Supervised*, dati di allenamento includono gli output desiderati
- *Unsupervised*, dati di allenamento non includono gli output desiderati
- *Semi-supervised*, dati di allenamento includono alcuni output desiderati
- *Reinforcement learning*, decisioni prese in funzione della reward

Transfer learning Il modello viene utilizzato e allenato anche per altri task, modificando nel caso i pesi necessari per farlo funzionare anche in altri domini applicativi

Data augmentation Applico trasformazioni come rotazioni, traslazioni e scalature, per far imparare alla rete neurale a individuare oggetti in differenti posizioni, rotazioni e scale.

Rasoio di Occam Principio secondo il quale, in caso di diverse spiegazioni possibili per un fenomeno, la spiegazione più semplice e con meno assunzioni è quella da preferire

4.1 Metodi induttivi

4.1.1 Artificial Neural Networks

Le reti feed-forward hanno una struttura a strati che consiste in un numero omogeneo (ma non lineare) di elementi di elaborazione. Il segnale parte da sinistra e si propaga a destra. Il neurone elabora come un punto di vista, attraverso i suoi pesi guarda cosa hanno fatto i neuroni precedenti. Se cambiano di pesi, lo stesso neurone vede le informazioni sotto un altro punto di vista. L'uscita sarà una decisione

Funzione di attivazione E' utilizzata per determinare l'output di un neurone in base al suo input. Esistono diverse funzioni di attivazione, ma tutte accettano un input numerico e generano un output in base a una regola specifica

Back-propagation Consiste nell'aggiornare i pesi della rete attraverso il calcolo dell'errore tra l'output prodotto dalla rete e il valore desiderato. Si parte dall'output della rete e si calcola l'errore rispetto al valore desiderato, che viene poi propagato all'indietro nella rete per calcolare gli aggiornamenti da applicare ai pesi. L'obiettivo è minimizzare l'errore di predizione della rete.

4.2 Metodi tradizionali

K-Nearest Neighbors

1. Calcolo la distanza da gli altri record di training, distanza euclidea
2. Identifico i vicini k
3. Uso le etichette delle classi dei vicini per determinare l'etichetta della classe del record. Prendo la classe che ha il numero maggiore di etichette fra gli elementi k

Decision tree Non prendo decisioni in parallelo

4.3 Deep Learning

L'approccio funziona come una rete neurale classica. Costituita da una fase di training e una fase di deploy. Si aggiungono strati per diminuire il numero di nodi totali necessari

Shallow vs Deep La differenza tra reti shallow e deep riguarda la profondità della rete, ovvero il numero di layer di cui essa è composta

4.3.1 Convolutional Neural Network

E' una rete neurale con alcuni strati convoluzionali. Fa quello che fa la fully connected network (e probabilmente ne ha una al suo interno). L'informazione viaggia dagli ingressi alle uscite in modo diretto. L'idea è quella di usare dei *kernel* di convoluzione.

Convoluzione Spostiamo di pixel in pixel il nostro kernel. Se trovo la forma nell'immagine che è simile a quella nel kernel, nell'uscita della convoluzione, *feature map*, troveremo un valore alto. Avrò una feature map per ogni filtro

Rectified Linear Unit Il rettificatore è un modulo che dato un segnale, elimina la parte negativa. Voglio una risposta solo se hai trovato un'immagine simile a quella del kernel, se il risultato della convoluzione è negativo o basso non mi interessa.

Modulo di pooling Semplificare il problema e guardarlo in uno spazio più piccolo. Se in una certa zona ho trovato un massimo, per me quella zona corrisponde al valore di massimo. Ottengo una nuova immagine ma più piccola

Flattening Ho più feature map, creo il vettore unico, mando nella rete fully

Esempio end-to-end Parto da un'immagine, faccio la convoluzione, applico max pooling, poi di nuovo convoluzione, lo faccio tante volte. La matrice di feature che ottengo, verrà rianalizzata da un altro kernel, per creare un'interpretazione che sale di livello. Alla fine creo un unico vettore chiamato *flat* e dovrò classificarlo, per farlo userò una rete tradizionale. In base al tipo di vettore verrà associato un oggetto

Autoencoders Sono una classe di algoritmi di apprendimento che utilizzano una rete neurale per codificare e decodificare dati. L'obiettivo è quello di apprendere una rappresentazione compatta di un insieme di dati, in modo da poterli ricostruire con una bassa perdita di informazione.

Generative Adversarial Network Le GAN, sono un tipo di reti neurali artificiali costituite da due parti principali: un *generatore* e un *discriminatore*, che competono tra di loro in modo da generare dati sintetici difficili da distinguere da quelli reali. Il generatore cerca di produrre immagini realistiche a partire da un input, mentre il discriminatore cerca di distinguere le immagini prodotte dal generatore da quelle reali.

5 Sistemi multimodali

I sistemi biometrici multimodali sono quelli che utilizzano o sono capaci di utilizzare più di uno tratto biometrico fisiologico o comportamentale. I sistemi multimodali sono più accurati dei sistemi che li compongono

Vantaggi

- Aumenta la copertura della popolazione riducendo l'FTE
- Sono un efficace metodo antispoofing

Svantaggi

- Più costosi
- Più lenti

Livello di fusione Per confrontare fra loro i valori di diversi matcher è necessario eseguire prima un'operazione di normalizzazione

1. *Livello di feature*, prima dei matcher e dei moduli di decisione.
2. *Livello di matchscore*, tecnica più diffusa. Possiamo fare fusione pre matching e after matching.
   - *classificatore*, è un modulo che avendo un ingresso $s_1, s_2, \ldots, s_n$, produce direttamente l'uscita impostore/genuino
   - *combinatore*, è un modulo che combina in modo lineare, non lineare, logico/combinatorio, i valori $s_1, s_2, \ldots, s_n$ e passa un unico valore S al decisore (che può essere di nuovo un classificatore o una soglia)
3. *Livello del modulo di decisione*

Sistemi multimodali gerarchici Acquisizioni biometriche in cascata a seconda del risultato dell'identificazione precedente

Fusione con qualità del tratto Uno score basso in ingresso può dipendere dalla qualità del tratto in input, non per forza un caso di impostore

6 Flusso ottico e impieghi in biometria

Tecnica che consente di analizzare il movimento di oggetti all'interno di una scena attraverso l'elaborazione di immagini. Questa tecnica viene utilizzata in diverse applicazioni, tra cui la guida autonoma dei droni e dei robot, dove è necessario rilevare gli oggetti. Viene impiegato anche in ambito di biometria, per l'anti-spoofing e per la biometria comportamentale.

Flusso ottico E' una misura della direzione e della velocità del movimento dell'oggetto nella scena ripresa dall'immagine

Motion field Proiezione nell'immagine di vettori di movimento tridimensionali

Applicazione biometriche del flusso ottico

- Videosorveglianza
- Face tracking, posso analizzare micro e macro-espressioni
- Pattern comportamentale dalla camminata

Histogram of Oriented Optical Flow Tecnica di elaborazione dell'immagine utilizzata per analizzare il flusso ottico. Divide l'immagine in regioni e calcola il flusso ottico per ciascuna regione, quindi calcola l'orientazione di ogni vettore di flusso ottico e costruisce un istogramma delle orientazioni. Questo istogramma può quindi essere utilizzato come rappresentazione dell'informazione del flusso ottico

7 Privacy

Continuous authentication Autenticazione biometrica non in modo istantaneo ma in modo continuo

Silent authentication L'utente non deve neanche accorgersi che sta facendo l'autenticazione. Esempio è l'utilizzo di smartwatch con il controllo dei movimenti

Classical password protection L'approccio tradizionale con password non è applicabile in quanto il dato biometrico cambia continuamente, diversamente è un replay attack. L'hash sarà sempre diverso

7.1 Biometria cancellabile

E' un modo per incorporare la protezione e le funzionalità sostitutive in biometria per creare un sistema più sicuro. Distorsione intenzionale e sistematicamente ripetibile delle caratteristiche biometriche

Feature transformation

- *Bio-hashing*, anche detto *salting*, la funzione di trasformazione è invertibile. Utilizzato per aggiungere un valore casuale al template biometrico prima di applicare la trasformazione con la chiave K. In questo modo, anche se un attaccante dovesse compromettere il sistema e ottenere il template trasformato $F(T, K)$, non sarebbe in grado di risalire al template originale senza conoscere il valore del salting
- *Non–invertible transformation*, una funzione che rende difficile dal punto di vista computazionare invertire un template trasformato in un template originale

Helper data Template crittato memorizzato

Key binding Associare una chiave crittografica a una identità

8 Videosorveglianza

Non ho enrollment ma guardo le persone e in caso di movimenti/azioni strane vado a vedere se sono in watchlist

Funzioni avanzate per la sorveglianza

- *High Dynamic Range*
- *Gamma dinamica estrema*
- *Intelligent Auto Mode*, controllo riflessioni
- *Color Night Vision*, vedere colori anche a bassa luminosità ambientale
- *Group of Picture controll*, estrapolare informazioni video, come quelle che denotano differenze di traffico, predicendo il futuro. Se invece non ci sono persone, il frame si abbassa anche se mi sembra di vedere un video continuo
- *Compressione Auto-VIQS*, riesce a comprendere quale area dell'immagine cambia più spesso rispetto a quella che non lo fa. Si può risparmiare banda modificando solamente quella in movimento.
- *Riduzione del rumore*, applico un filtro per avere un'immagine migliore
- *Intelligent Face Compression*, la camera riesce a determinare automaticamente la porzione dell'immagine che è importante ai fini dell'identificazione
- *Rilevamento dei volti*
- *RainWash/ ClearSight coating*, rivestimento

Encryption Crittazione dei dati a riposo e durante la trasmissione. Il template biometrico non può essere convertito in un'immagine in caso di data breach

Data purging Elimina i dati non più utili presenti in watchlist

9 Vulnerabilità dei sistemi biometrici

Possiamo distinguere principalmente due tipi di attacchi

- *Attacchi indiretti*, dall'interno del sistema. Ad esempio manipolando le reference biometriche all'interno del database
- *Attacchi diretti*, attraverso il sensore

Presentation Attack Detection Determinazione automatica del presentation attack. I metodi di rilevamento della liveness sono definiti come subset dei metodi di PAD

Presentation Attack Instrument Oggetto usato nel presentation attack. Anche noto come *artefact*

Adversarial attack Mira a ingannare il matcher o l'algoritmo introducendo delle perturbazioni impercettibili all'occhio umano nell'input.

Esistono all'interno di questa categoria diverse tipologie di attacchi

- *Attacchi whitebox*, l'attaccante ha accesso ai parametri del modello
- *Attacchi blackbox*, l'attaccante non ha accesso ai parametri del modello
- *Non-targeted adversarial attack*, l'obiettivo è solo quello di generare input manipolati che siano classificati erroneamente dal modello, senza preoccuparsi di quale classe venga assegnata
- *Targeted adversarial attack*, generare input manipolati che siano classificati in una classe specifica

9.1 Variazioni delle immagini

1. *Controllo in IR*
   Le immagini stampate o le immagini mostrate tramite schermi non possono essere usati per attaccare le camere IR. Una stampa o l'utilizzo di uno schermo non vengono visualizzati da un sistema infrarosso
2. *Approcci compositi*,
   Usando diversi sistemi potremmo acquisire informazioni differenti sui soggetti diminuendo così la capacità di poter truffare un sistema di questo genere. Solitamente si utilizzano telecamere del visibile, NIR e 3D per realizzare un'immagine tridimensionale
3. *Ricerca dei Moirè*
   Con effetto moiré si indica una figura di interferenza, punteggiature dello schermo. Visti i pattern periodici nello spettro, potrebbe essere un fake

4. *Tecniche Challenge–Response*
   Il sistema sfida l'utente con alcune istruzioni casuali, la risposta viene controllata per verificare se l'utente ha seguito le istruzioni. Ad esempio il sistema chiede all'utente di dire una determinata frase davanti al sensore

9.2 Possibili attacchi su sistema voice

1. *Impersonation*, fingendosi un'altra persona
2. *Replay attack*, registrando la voce della vittima
3. *Voice conversion*
4. *Sintetizzazione del discorso*, text to speech
5. *Artificial, non speech like tones*, l'input non è una voce vera ma un segnale

LipPass Sfrutta i componenti sullo smartphone per poter descrivere il movimento della bocca utilizzando i segnali acustici che rimbalzano sul viso.

10 Ambient intelligence

Ambient Un ambiente digitale che si integra nella quotidianità, integra proattivamente oggetti e ambiente assistendo l'utente nella vita quotidiana. Le caratteristiche del sistema sono: dinamico, embedded, intelligente, adattabile, interconnesso

Intelligence Il sistema è sensibile al contesto, adattivo e impara dalle preferenze dell'utente.

L'ambiente AML è basato su miniaturizzazione e hardware low-cost. In che modo la biometria può essere utile?

- *Person identification*, per assegnare conoscenze e preferenze
- *Person classification*, per affinare servizi caratteristici
- *Person action understanding*, applicare servizi e operazioni alle azioni umane

Iride non vincolato Scanner di iridi che brandeggia (fa tilt per spostarsi) per aiutare l'acquisizione e usa un semaforo luminoso per aiutare l'utente. Minimizza FTE e FTA

Le 4 P *Performance* in continuo aumento, *Pervasività* della biometria, *Possibilità* applicative, *Privacy* a rischio