



Università
di Genova

RICONOSCIMENTO DI IMMAGINI TRAMITE COMPUTER

PROF. ING. GIANNI VERNAZZA

30 MARZO 2022

INDICE

- Acquisizione ed estrazione di parametri delle immagini
- Prime elaborazioni (contorni, regioni, tessitura,...)
- Architettura delle elaborazioni
- Feature Pattern e classificazione statistiche
- Prestazioni
- Reti neurali
- Esempi di applicazioni

Acquisizione di immagini

Tipologia di immagini:

- statiche/dinamiche
- b/n
- Colori (RGB)
- 2D/3D/4D/...
- Lunghezza d'onda/frequenza (visibile, infrarosso, multifrequenza)
(Trasformata di Fourier-FFT)

Parametri immagine

Immagine (digitalizzata) matrice:

- Risoluzione spaziale (640x480; 1920x1080 pixel;...)
- Nr pixel/Voxel/...
- Risoluzione ampiezza (livelli di grigio)
- Nr immagini al secondo

Problemi preliminari:

- Calibrazione (geometrica)
- Rumore telecamera
- Risposta stabile nel tempo
- Illuminazione
- Ottica/lenti
- Etc.

Qualità immagine

- sensore
- Nr. Pixel (punti video/matrice)
- Nr. Livelli di grigio/colore
- Nr. Quadri al secondo
- Acquisizione a differenti lunghezze d'onda
- Etc.

Alla fine una 'buona' immagine richiede una 'buona' **banda**.
(La banda è come la portata/sezione di un tubo per l'acqua).

Esempi di immagini con diversi livelli di grigio



256 livelli



8 livelli
Spalla

Esempi di immagini con diversi nr. di pixel:



512x512 pixel



128x128 pixel
Dettagli contorni

Prime Elaborazioni

- Estrazione di contorni (Filtri maschere) dove c'è una differenza elevata di livelli di grigio (lunghezza, curvature,...)

Immagine in ingresso: Filtro per estrazione contorni LPF-HPF → immagine dei contorni

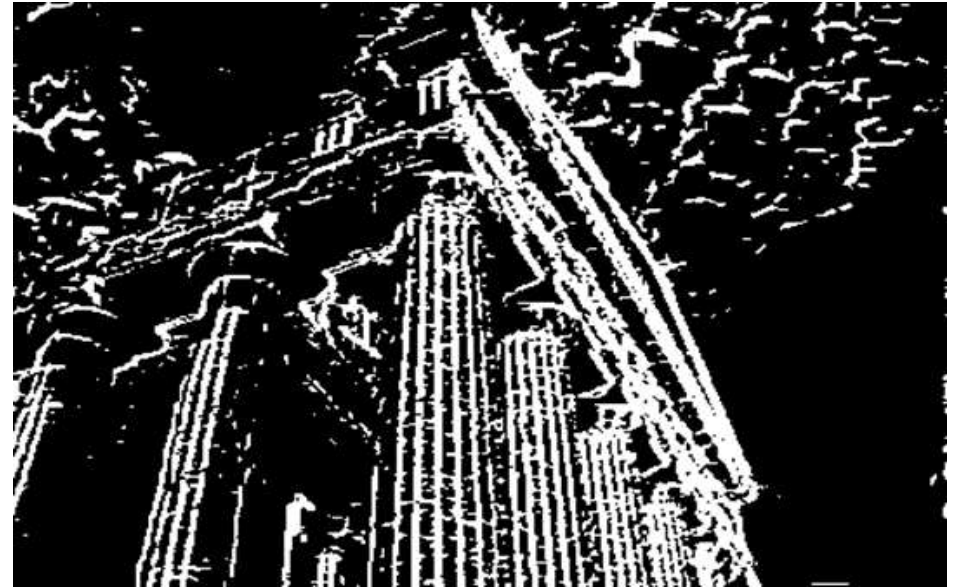
Problemi:

- Sensibilità al rumore (alcuni punti di grigio sono alterati dal rumore)
- Contorni non sottili (a volte)
- Contorni spezzettati non connessi
- Sensibilità alla soglia

Esempi di immagini filtrate



immagine del modulo



mappa dei contorni (soglia = 40)

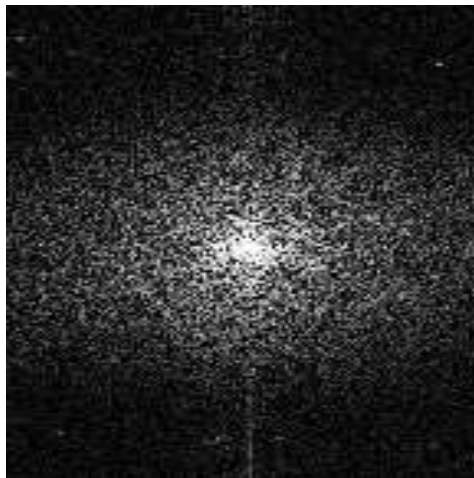
CRISPENING – HIGH-PASS



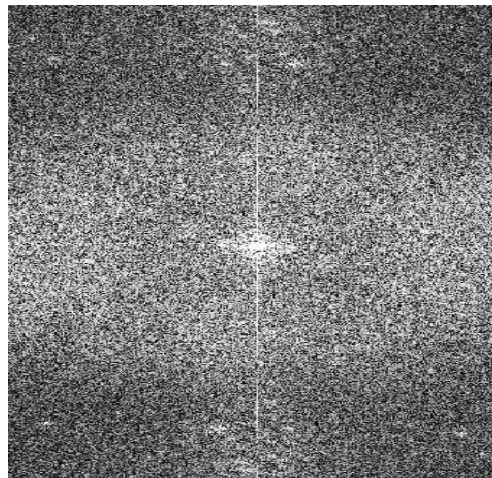
Immagine originale



Immagine filtrata



DFT immagine originale



DFT immagine filtrata

$$h = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

Maschera di convoluzione per
esaltare le alte
frequenze:
filtro HIGH-PASS

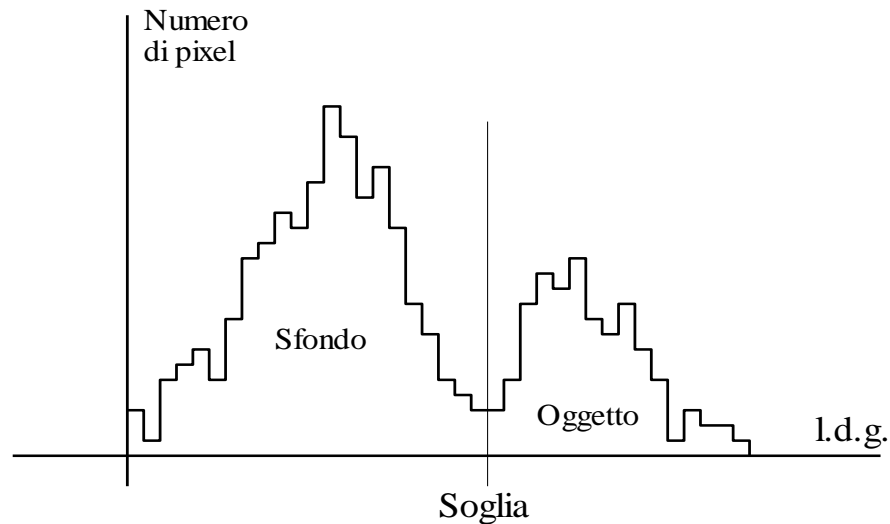
Prime elaborazioni (2)

Elaborazione di regioni (segmentazioni)

(area, perimetro, fattore di forma, baricentro, livello medio di grigio, etc.)

Ricerca di contorni che circoscrivano (aree omogenee) ovvero ricerchino i pixel in gruppi omogenei (o quasi) (es. per tono di grigio, tessitura,)

Esempio figure:



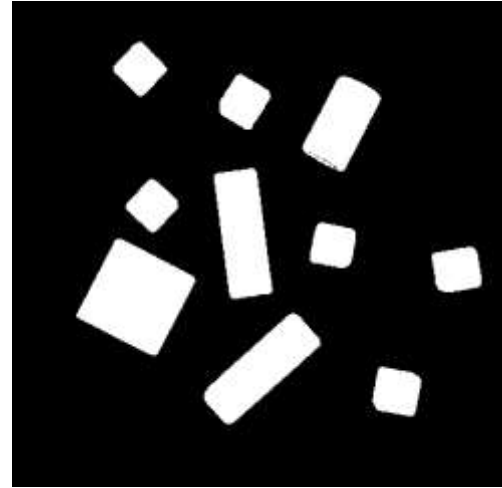
$$I_{out}(m, n) = \begin{cases} 1 & \text{if } I_{in}(m, n) \geq \text{Soglia} \\ 0 & \text{if } \text{Input}(m, n) < \text{Soglia} \end{cases}$$

Problemi quale il livello di 'omogeneità'.

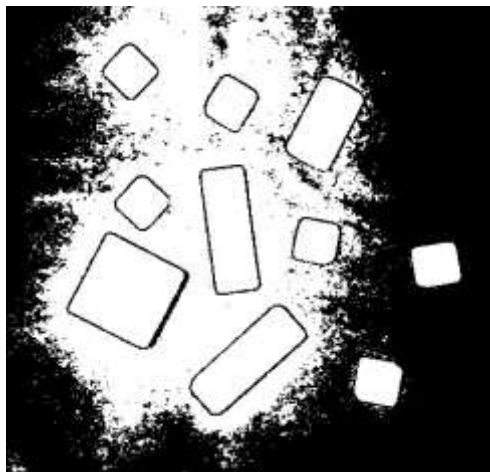
Esempi di thresholding (sogliatura):



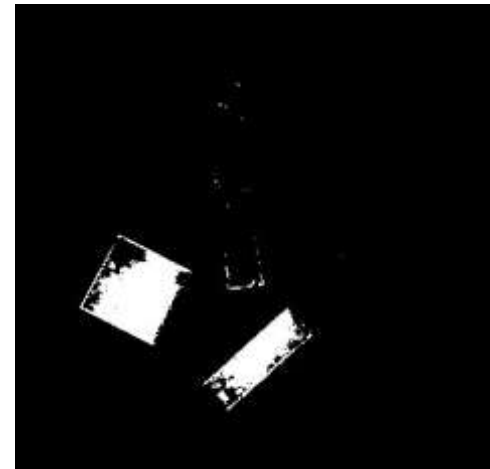
Immagine originale



Soglia pari a 95



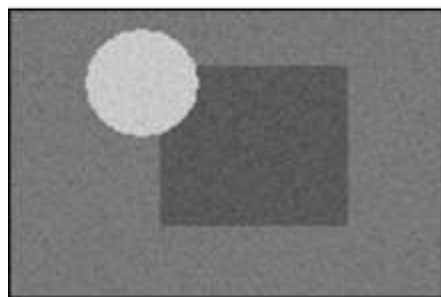
Soglia pari a 25



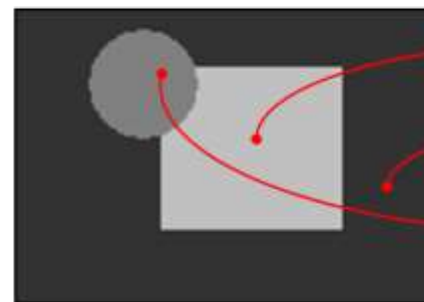
Soglia pari a 225

Segmentazione: esempio

Immagini originali



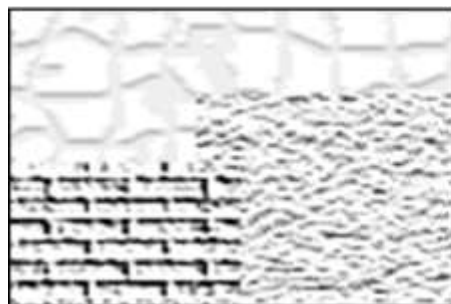
Immagini segmentate



REGIONE 2

REGIONE 1

REGIONE 3



REGIONE 1

REGIONE 2

REGIONE 3

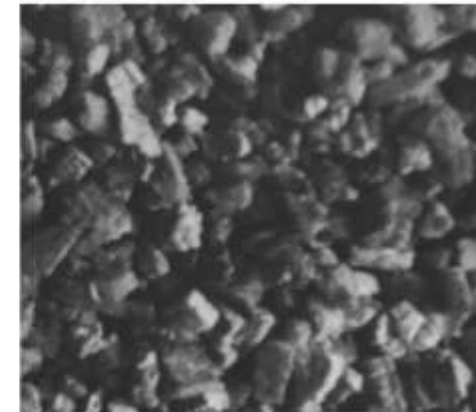
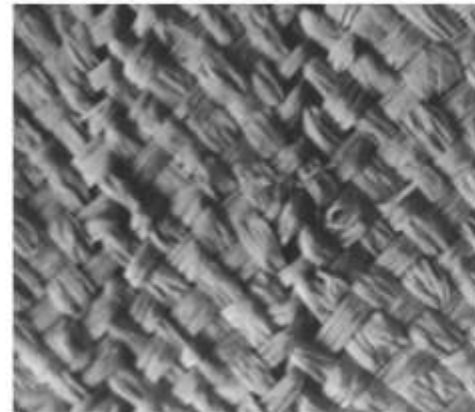
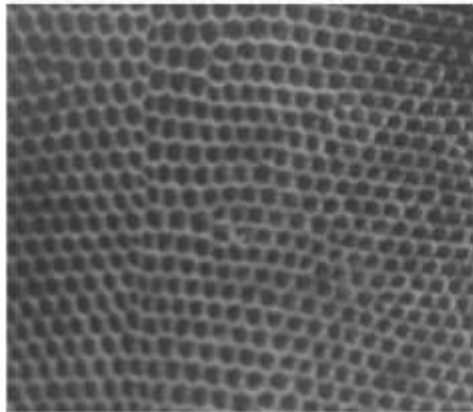
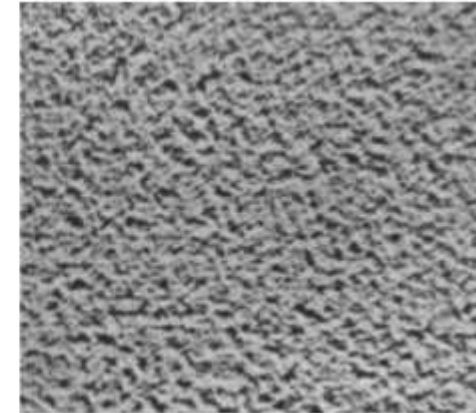
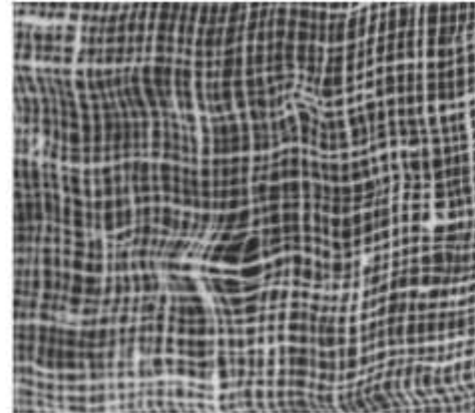
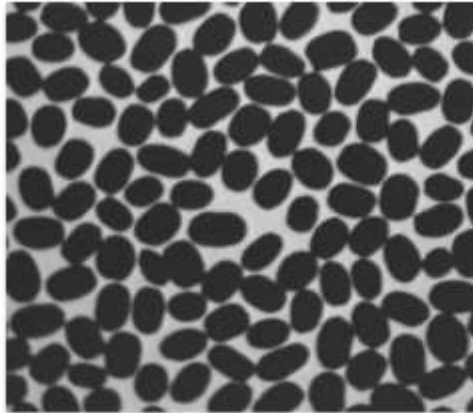
Analisi di tessitura

Presenza di una 'trama ricorrente' (v. figure)

Analisi con metodi:

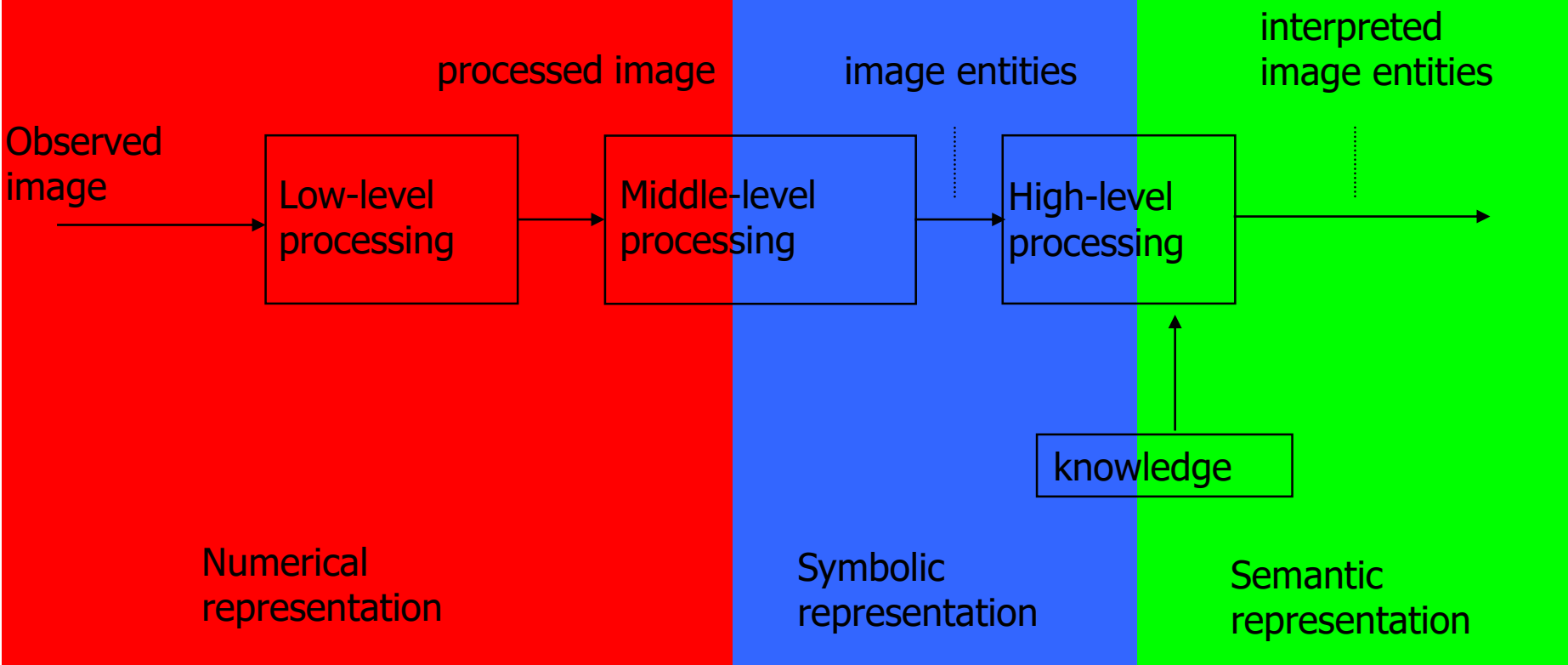
- Strutturali (elemento di base texel/grammatiche)
- Statistici (ricorrenze statistiche spaziali)
- Modelli (markoviani)
- frattali

Esempi di tessiture

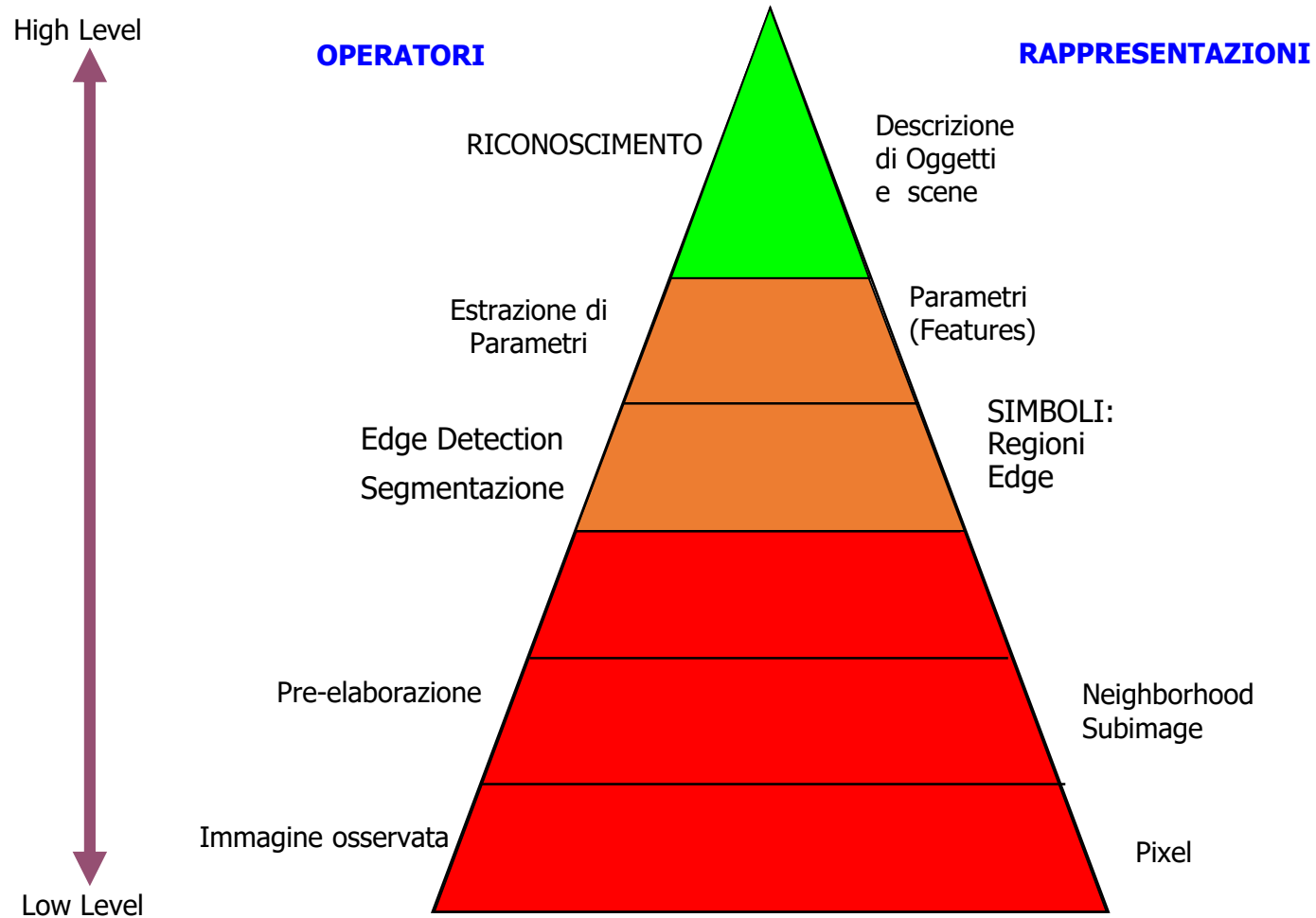


Esempi di tessiture naturali (immagini di test raccolte da Brodatz)

Livelli di Elaborazione

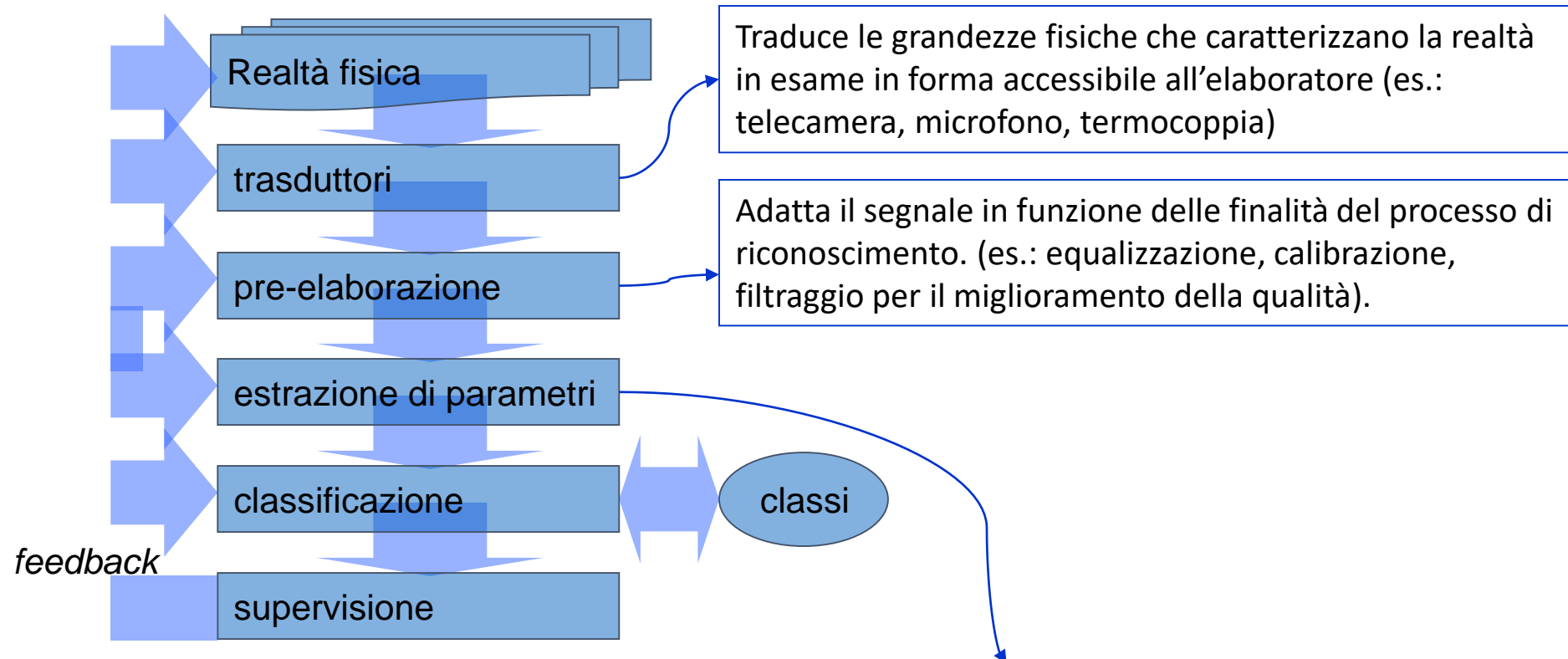


Livelli di Elaborazione



Estrazione di parametri (Feature Pattern)

Rivediamo la sequenza del processo di riconoscimento



È uno dei problemi principali, e consiste nell'estrarre l'informazione contenuta nel segnale e renderla utilizzabile per la successiva analisi: dipende pertanto dalle caratteristiche del processo di classificazione. Un parametro estratto da un segnale deve essere significativo, ossia avere un alto potere discriminante, ma al tempo stesso deve poter essere ottenuto con un'adeguata semplicità computazionale.

Classificatori supervisionati

- Esistono due tipologie principali di classificatore:
 - **supervisionato** (*supervised*);
 - **non supervisionato** (*unsupervised*).
- Le fasi di un processo di **classificazione supervisionata**:
 - addestramento (*training*): il sistema opera su un insieme di dati di addestramento (*training set*), cioè di campioni preclassificati.
 - verifica (*test*): tale fase è svolta operando su un *test set* composto da dati anch'essi preclassificati, ma non utilizzati in fase di *training*.
 - La scelta dei dati di *training* e di *test* è molto importante: idealmente essi dovrebbero rappresentare tutte le varie casistiche, ossia dovrebbero essere “**statisticamente completi**”.

Classificatori non supervisionati

- Nei classificatori **non supervisionati**:
 - **non viene usato alcun *training set***, tipicamente perchè non sono note nè facilmente identificabili le classi fisiche dell'applicazione in esame (**classi naturali**).
 - la **complessità** è sovente maggiore che nei classificatori supervisionati ed il risultato necessita di una **validazione** (*cluster analysis*).

Il concetto di classificazione

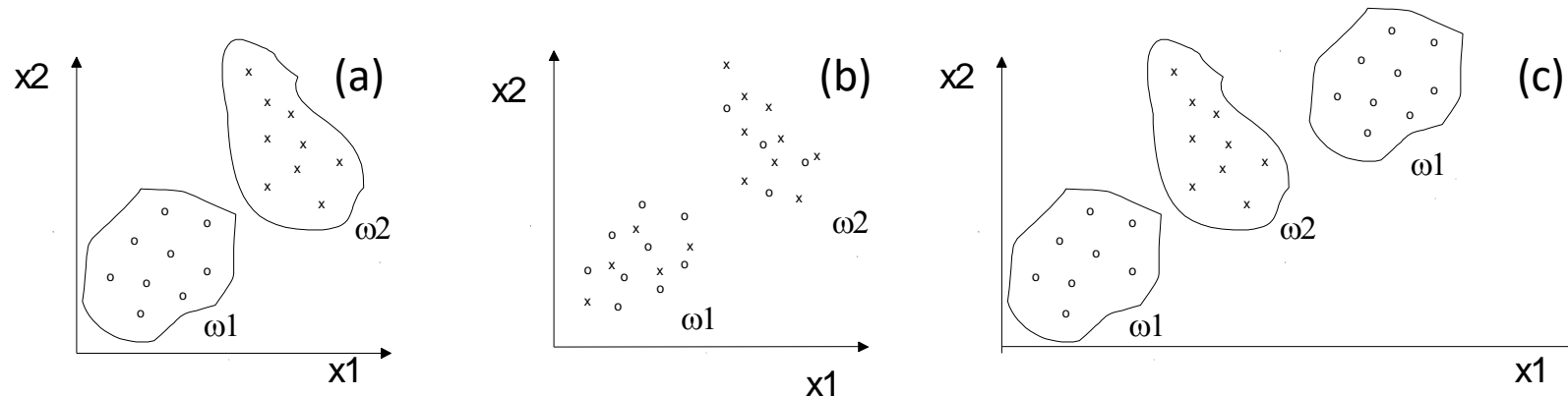
- Un **metodo di classificazione** ha il compito di decidere a quale classe di appartenenza debba essere assegnato ciascun **campione** (*pattern*), basandosi sul vettore delle misure (*feature*) fornito dai sensori.
- La **classificazione** si basa sulle seguenti ipotesi:
 - per ciascun campione è disponibile un insieme di **misure** rappresentabili mediante un vettore di caratteristiche (*feature*).
 - le **classi** di appartenenza sono note *a priori* ed in numero finito.
 - si ha una sufficiente conoscenza *a priori* sulle classi oppure si ha a disposizione un insieme di **campioni** o *pattern* preclassificati.

Misure di distanza

- Svitati classificatori, fra i quali alcuni che utilizzano esplicitamente funzioni discriminanti, sono basati sul criterio della minima distanza fra il campione incognito e le classi. È quindi importante fornire una definizione di distanza nello spazio delle *feature*.
- Una **distanza** o **metrica** nello spazio delle *feature* è una funzione $d(\cdot, \cdot)$ definita dalle seguenti proprietà:
 - $d(\mathbf{x}, \mathbf{x}) = 0$;
 - $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x}) > 0 \quad \forall \mathbf{x} \neq \mathbf{y}$;
 - $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ (disuguaglianza triangolare)

Classi multimodali

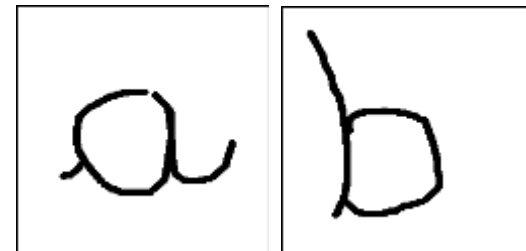
- Una classe si dice **multimodale** quando può essere pensata come unione di più sottoclassi separabili linearmente.
 - Esempio



- (a) presenta due classi separabili linearmente, (b) e (c) due classi non separabili linearmente. La classe ω_1 in (c) è bimodale.
- Nei casi (a) e (c) le tecniche statistiche funzionano bene, mentre il caso (b) presenta maggiori complessità.

Sistema di Pattern recognition classico

- Raccolta dati
- Scelta delle feature
- Scelta del modello
- Addestramento del modello
- Valutazione



Esempio guida: sistema che distingue tra le lettere scritte a mano “a” e “b”.

Raccolta dati

- Collezione di un insieme “sufficiente” e “rappresentativo” di esempi dal problema in esame.
- “sufficiente”?
- “rappresentativo”?
- Problemi di sensoristica (risoluzione, banda, ...)

Scelta delle features

- Non si possono utilizzare i dati così come sono (immagine 256x256 sono 65536 pixels)
- **Feature**: caratteristiche misurabili del fenomeno in esame (**pattern** = vettore di features):
 - semplici da calcolare;
 - invarianti a trasformazioni irrilevanti;
 - affidabili;
 - indipendenti;
 - discriminanti;
 - poche (problema della curse of dimensionality);
- In questa fase è molto utile l'utilizzo della conoscenza a priori sul problema

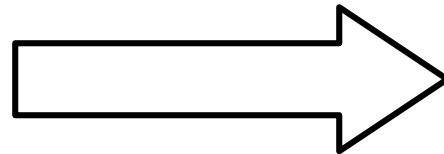
Scelta delle features: estrazione e selezione

- Il numero di feature deve essere piccolo per limitare il costo della misura e non influire sull'accuratezza del classificatore
- Estrazione di feature:
 - misura sui dati
 - creazione di nuove feature da combinazioni di feature misurate
- Selezione di feature: migliore sottoinsieme delle feature estratte
- Uso di una funzione criterio per la riduzione: tipicamente l'errore di classificazione di un sottoinsieme di feature.
- Inoltre, è importante determinare la dimensione dello spazio ridotto.

Addestramento del modello

- Sinonimi:
 - **training** del classificatore
 - **learning** del classificatore
- Processo con il quale si utilizzano i dati a disposizione (training set) per la costruzione del modello
- Esempi tratti dal problema

(Training Set)
Conoscenza a
priori



Regole che
governano il
fenomeno

Addestramento supervisionato

- Esempio:
- il training set è costituito da un insieme di immagini “a” e “b”.
- di ogni immagine conosciamo l’esatta classificazione (cioè se è “a” oppure “b”)
- queste informazioni sono utilizzate per determinare la soglia del classificatore.

Apprendimento supervisionato: esempio

- Vogliamo insegnare ad un computer a distinguere tra auto e biciclette
- Come dataset utilizziamo un insieme di 10 immagini diverse di biciclette e 10 di automobili.



Apprendimento supervisionato: esempio

- Mostriamo al computer le immagini con le biciclette e gli diciamo che quelle sono biciclette, lasciando che lui capisca da solo quali sono le caratteristiche delle biciclette.



Apprendimento supervisionato: esempio

- Facciamo lo stesso con le automobili

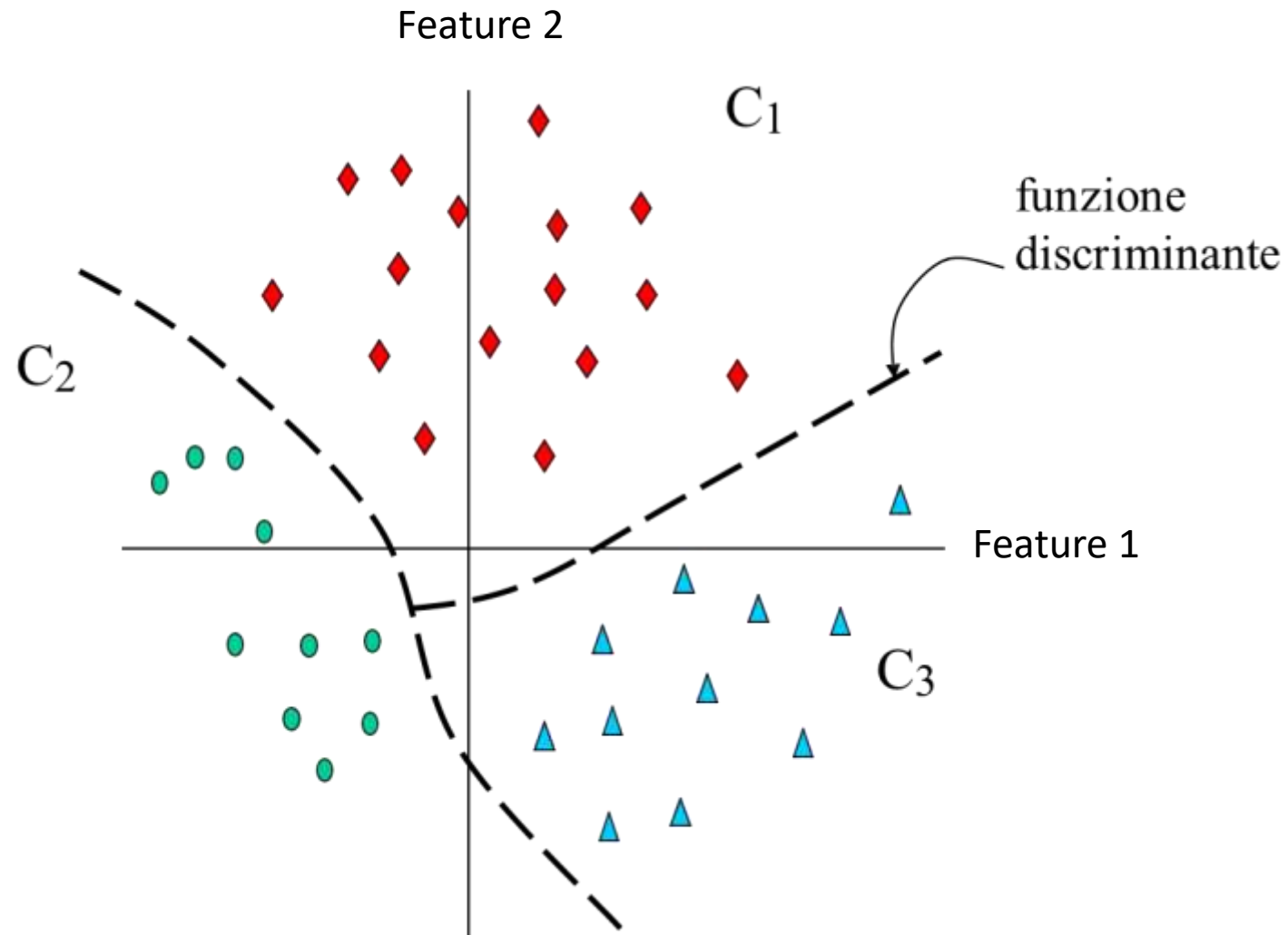


Apprendimento supervisionato: esempio PRESTAZIONI

- Infine, usiamo una nuova immagine e lasciamo il computer decidere se si tratta di un'automobile o di una bicicletta.



Classificazione statistica



Riconoscimento di numeri scritti a mano

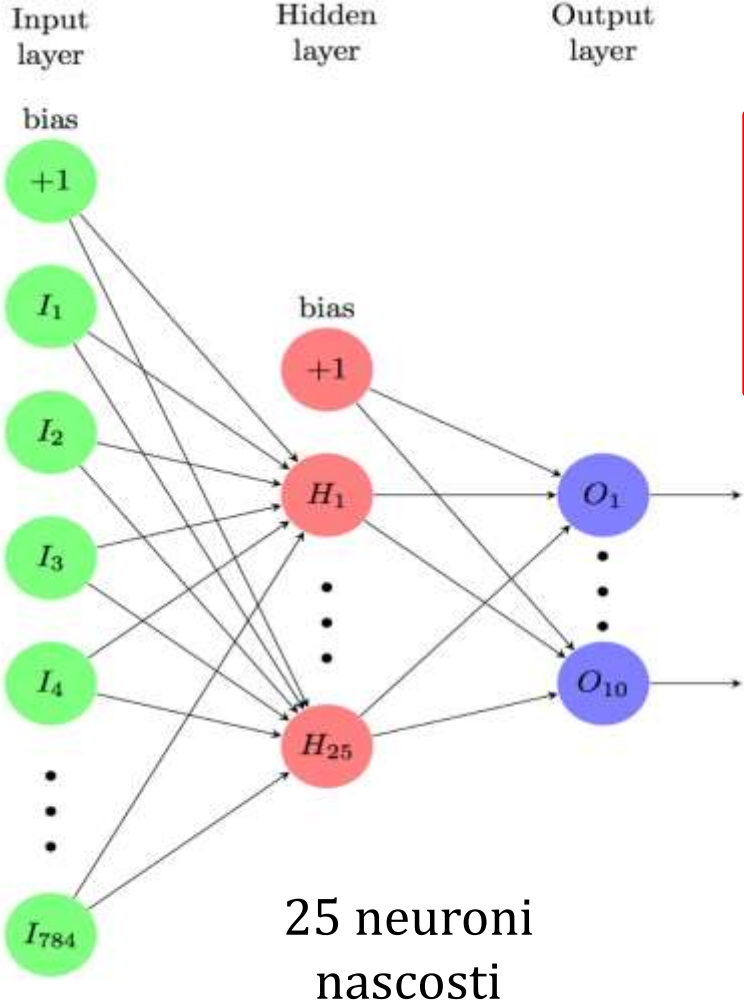
MNIST database:

3000 immagini
28x28 di numeri
scritti a mano
28



Immagine di esempio

Ogni input
corrisponde ad
1 pixel

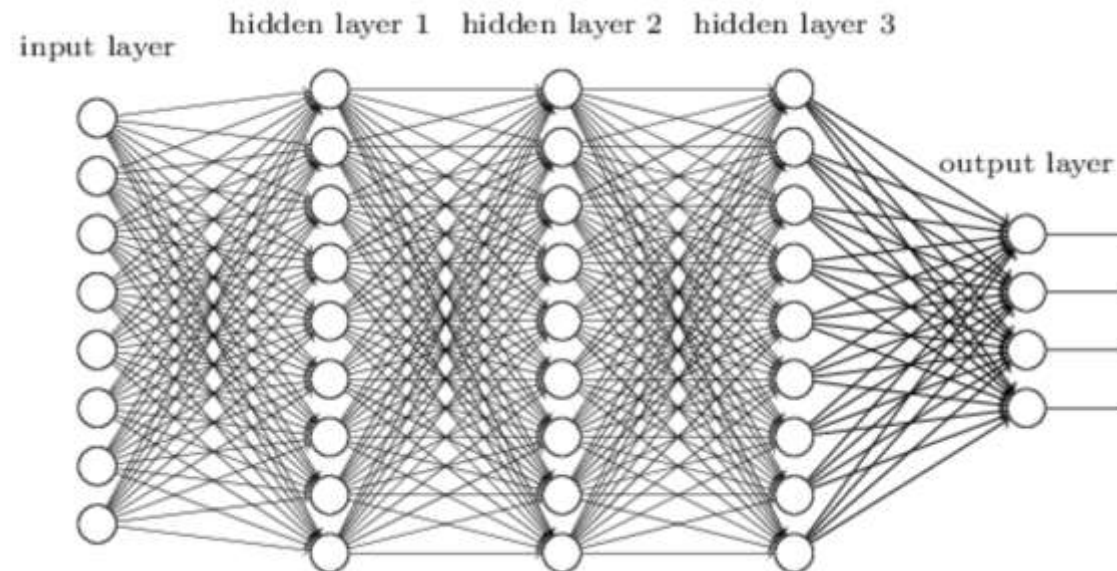


Si inizia con pesi casuali e si usa una tecnica (detta back propagation) per imparare i pesi corretti

Una uscita per ogni numero possibile
Si seleziona l'uscita con peso maggiore
e.g. 9

Reti neurali "deep"

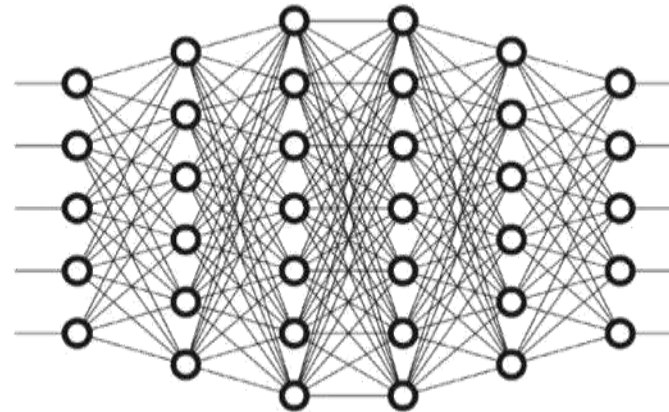
- i primi lavori estendevano semplici reti neurali per avere più livelli nascosti altamente connessi
- se tali reti potessero essere addestrate, sarebbero molto più potenti delle reti neurali "superficiali"
- ma le reti multistrato generiche sono estremamente difficili da addestrare!!



“Come si addestra una rete neurale artificiale?”

In sintonia con quello che avviene nel **cervello umano** durante le fasi dell'**apprendimento**, una rete neurale artificiale viene **addestrata mediante l'uso di esempi noti**.

Un esempio è una coppia (Immagine, Etichetta). L'etichetta identifica il contenuto dell'immagine: Cane, Gatto, Uccello, ...



Etichetta

Analisi di Immagini Telerilevate Satellitari

Un altro campo in cui le reti neurali trovano molto spazio è l'elaborazione delle immagini satellitari telerilevate su larga scala.

Nell'esempio di sinistra sono state usate per individuare le aree occupate dagli edifici (in fucsia); nell'esempio di destra sono state usate per individuare la rete stradale.



Altri esempi

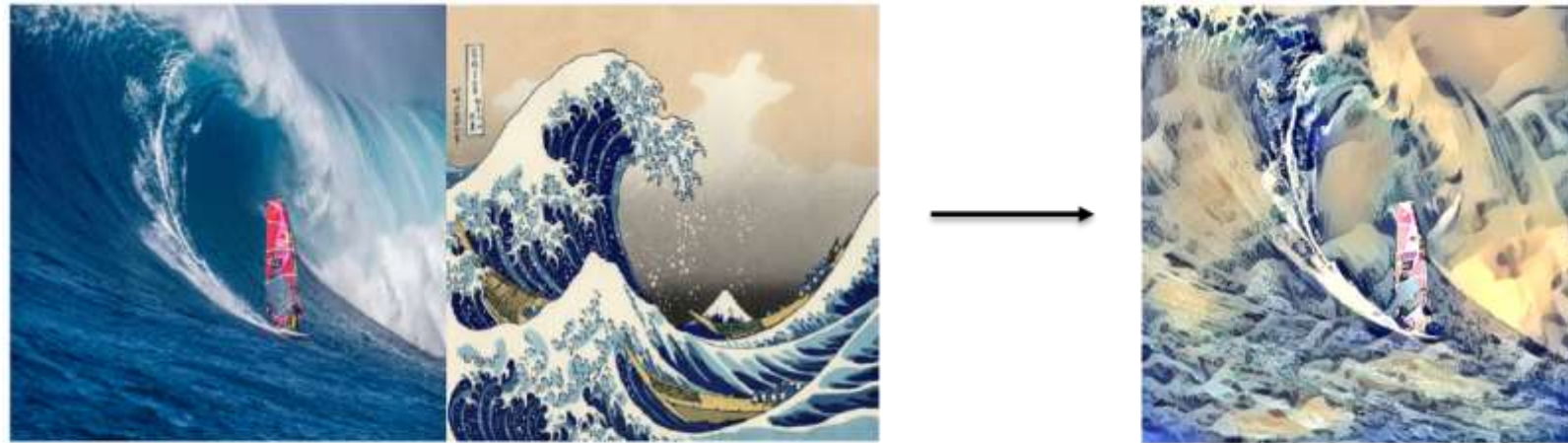
Le reti neurali sono utilizzate per la predizione dell'andamento dei mercati, per la traduzione simultanea, per il riconoscimento vocale, etc.



Sintesi di Immagini Artistiche e Style Transfer

Esempi di applicazione in ambito artistico sono la sintesi di texture e la sintesi di immagini artistiche attraverso “style transfer”.

Quest’ultima è una tecnica di *computer vision* che permette di creare, a partire da due immagini, un’immagine sintetica caratterizzata dal contenuto iconografico di un’immagine e dallo stile dell’altra.

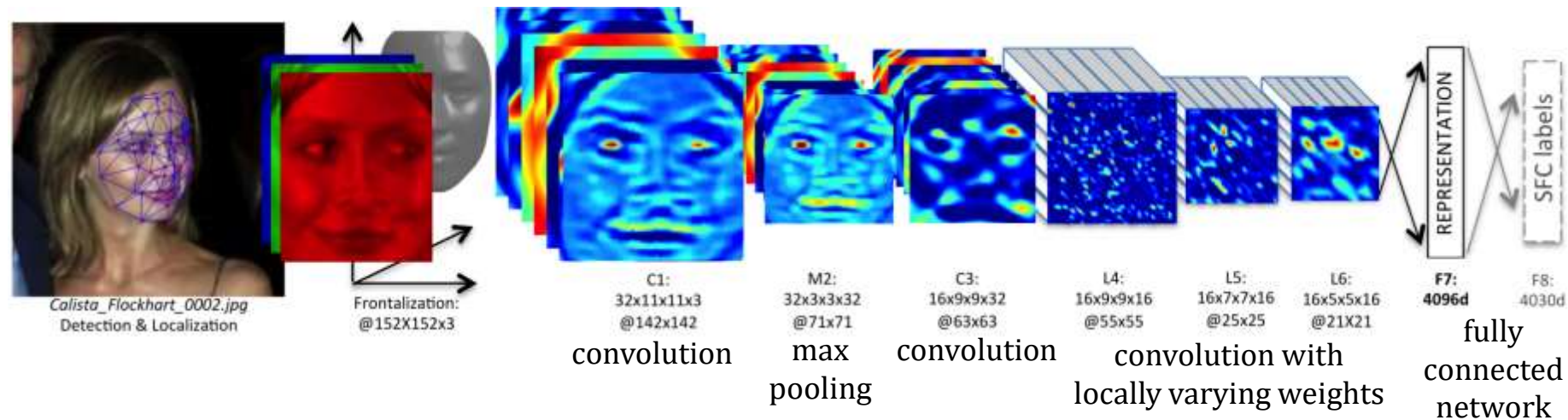
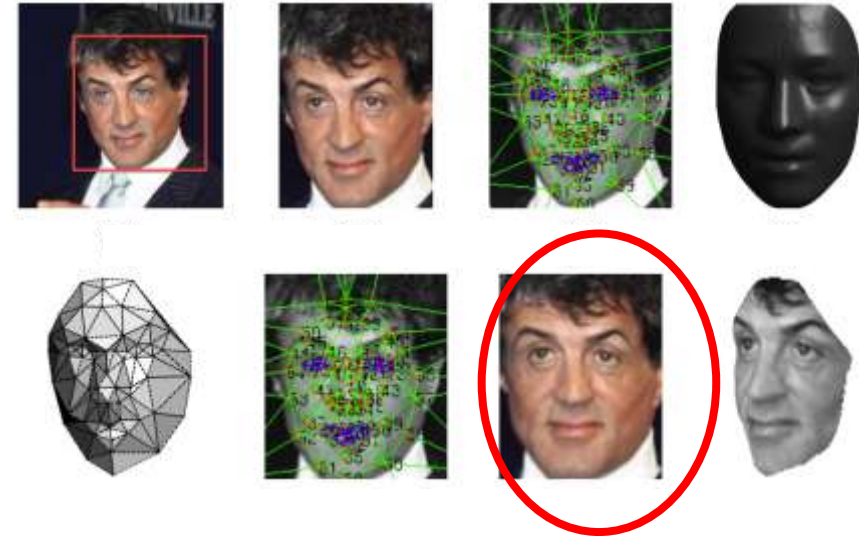


Risultato della sintesi utilizzando lo stile de “La grande onda di Kanagawa” di Hokusai su una foto.

Facebook's DeepFace system

Taigman et al., 2014

- Rilevazione del volto
- Allineamento 2D del volto rilevato
- Allineamento ad un modello 3D usando 67 punti noti
- Uso del modello 3D e dell'immagine per generare la vista frontale



Casi d'uso: attività fraudolente

- Identificazione transazioni legittime o fraudolente

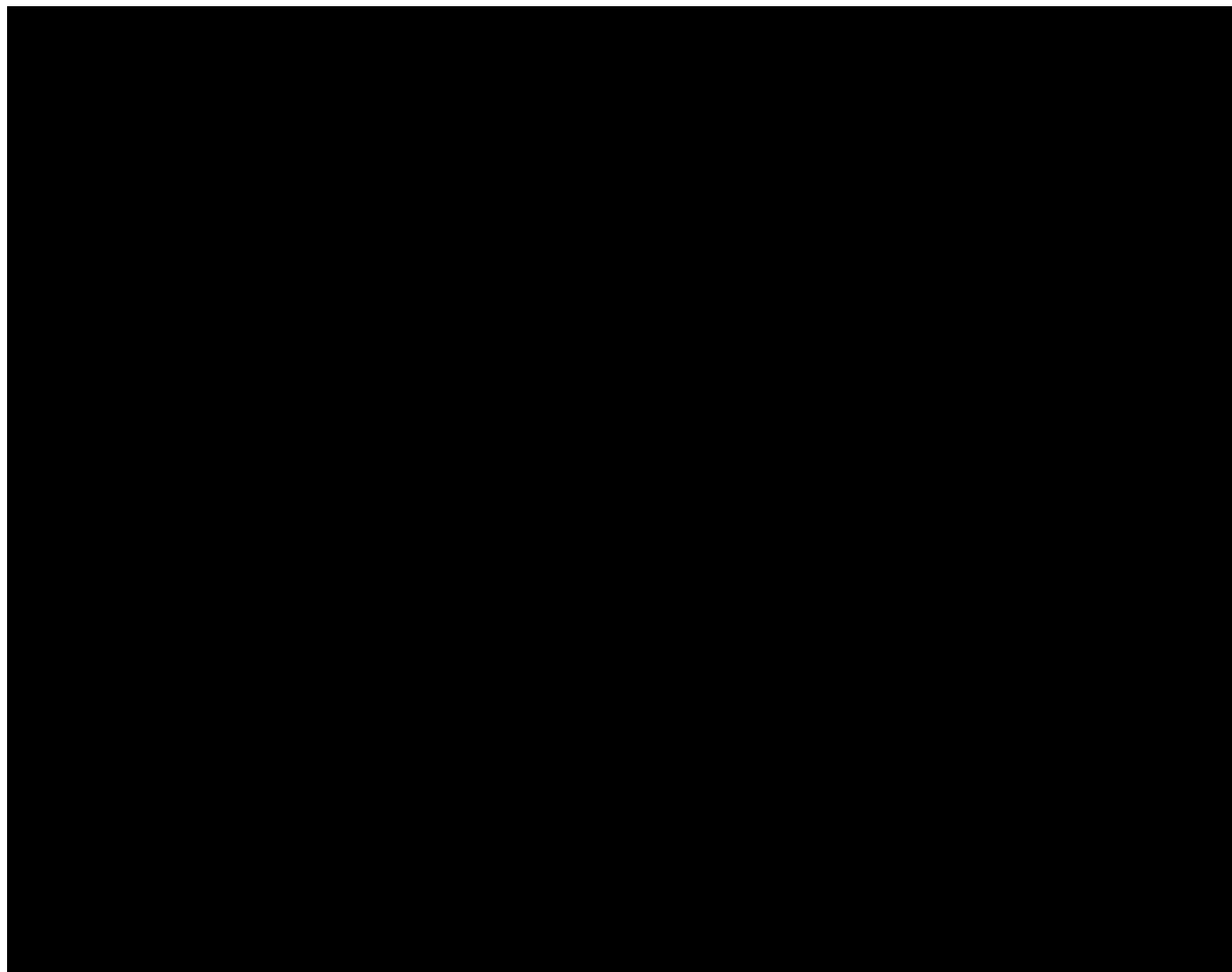


Casi d'uso: identificazione dello spam

- Sistema di classificazione per individuare le email di spam



Casi d'uso: guida autonoma



Casi d'uso: eventi anomali



Casi d'uso: tracking di persone



Casi d'uso: sorveglianza del traffico



Conclusioni

- **Machine Learning e Pattern Recognition** sono alla base di molti sistemi di elaborazione dell'informazione
- Grazie all'evoluzione degli algoritmi di pattern recognition, sistemi artificiali intelligenti vengono ormai utilizzati in un'enorme quantità di **campi applicativi**
- Le **reti neurali** sono strumenti potenti che sono quasi sempre alla base dei moderni sistemi di pattern recognition
- Il **miglioramento delle prestazioni** dei sistemi embedded rendono oggi possibile l'uso delle tecniche di pattern recognition anche in mobilità (smartphone)

**VI RINGRAZIO PER
LA CORTESE
ATTENZIONE**

