

Analisi di Immagini e Video (Computer Vision)

Giuseppe Manco

Outline

- Introduzione al corso
 - Docenti
 - Obiettivi
 - Programma di massima
 - Materiale didattico
 - Modalità d'esame
- Concetti introduttivi

Introduzione al corso

Admin

- Lezioni

- Giuseppe Manco

- <http://gmanco.github.com>
 - Via Bucci 8-9c.
 - giuseppe.manco@icar.cnr.it
 - Ricevimento: mercoledì 14:30-16:30
 - O su appuntamento

- Esercitazioni

- Francesco Pisani

- Via Bucci 8-9c
 - francescosergio.pisani@icar.cnr.it
 - Ricevimento: lunedì 15-17
 - o su appuntamento

Obiettivi del corso

- Acquisire e sperimentare le tecniche di base ed avanzate per l'analisi di immagini e video
 - Concetti fondamentali per l'analisi delle immagini
 - Object detection, object tracking e action detection
- Focus su deep learning e architetture neurali
 - CNN, RNN, Residual and Attention network
 - Aspetti avanzati
 - Modelli generativi e adversarial networks
- Linguaggio Python
- Pytorch/Tensorflow

Programma

- Introduzione alla computer vision.
 - Concetti fondamentali su Image processing e analysis: Image Basics, Python per Image Processing, Manipolazione di immagini.
 - Trasformazioni: Normalizzazione, filtri, Edge detection, morfologia, thresholding e segmentazione.
- Classificazione di immagini e video
 - Introduzione alla classificazione: applicazioni; approcci classici; scikit-Learn per la classificazione; limitazioni.
 - Deep Learning: Review su Neural Networks per l'analisi di immagini e video. Convolutional Neural Networks. Reti ricorrenti. R-CNN. Gestione dell'overfitting.
 - Architetture avanzate. Principali architetture di rete e loro caratteristiche: VGG, AlexNet, Inception and Residual Networks, Attention Networks. Transfer Learning.
- Concetti avanzati
 - Object Detection: Sliding windows, boundary boxes e anchors. Region Proposal Networks. Yolo e Darknet. Applicazioni.
 - Object tracking e action recognition. Optical flow; Single/multiple objects tracking; Action classification and localization.
 - Image segmentation e synthesis. UNet. Neural style transfer.
 - Modelli Generativi: Probabilistic Modeling, Autoencoders. Generative Adversarial Networks. Applicazioni: Colorization, Reconstruction, Super-Resolution, Synthesis, Text-to-image.
 - Adversarial Machine Learning. Principali attacchi e contromisure. Adversarial-free deep networks.

Materiale didattico

- Sito web del corso
 - <https://gmanco.github.io/courses/computervision/>
 - Lucidi delle lezioni
 - Notebooks e lucidi delle esercitazioni
- Libri:
 - E.R. Davies, Computer Vision: Principles, Algorithms, Applications, Learning. Fifth edition. Elsevier/Academic Press, 2018. **[Davies18]**
 - Jan Erik Solem, Programming Computer Vision with Python. O'Reilly Media, 2012. **[Solem12]**
 - Mohamed Elgendy, [Deep Learning for Vision Systems](#). Manning, 2020. **[Elg20]**
 - Rafael C. Gonzalez, Richard E. Woods, Digital image processing. 4th edition. Pearson, 2018. **[Gon18]**
 - I capitoli e le sezioni da considerare verranno indicati nei lucidi

Esame (1)

- Esperienza progettuale
 - Da concordare con il docente
 - Prevede una presentazione e una pagina di sommario con il codice sviluppato, da ospitare su github
- Orale
 - Nelle sedute d'appello
 - Si presentano i risultati del progetto

Esame (2)

- Donkey-car Challenge!

Esame (2)

- Donkey-car Challenge!
 - Using Deep Neural Network to Build a Self-Driving Car
- Obiettivi
 - Addestrare una rete neurale per la guida autonoma su un circuito
 - Livelli di difficoltà crescenti
 - Sfida a tempo
- Evento in programma a fine giugno
 - Esenta dall'orale
 - Al vaglio premi sui vincitori



Concetti introduttivi

Concetti fondamentali su Image processing e analysis

Crediti

- Alcune slides sono adattate dal corso di Analisi di Immagini di Fabrizio Angiulli

Analisi di immagini e video

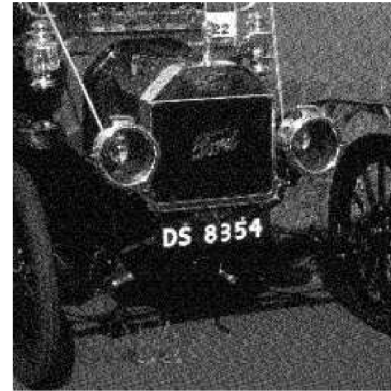
- Vari concetti
 - Digital image processing
 - Image Processing
 - Image Analysis
 - Computer Vision
- Discipline le cui definizioni e confini non sono univocamente riconosciuti
- La distinzione tra tipo di output prodotto è esplicativa, ma in pratica le discipline si sovrappongono ampiamente, “sconfinando” l'una nell'altra

Image Processing

- L'input e l'output sono immagini, di solito della stessa dimensione
- Miglioramento della qualità ed interpretabilità dell'immagine (image enhancement)
 - Aumento del contrasto
 - Eliminazione del rumore

Image processing

Blur (offuscamento, nebulosità) removal



Noise (random errors) removal

Image Analysis

- L'input è un'immagine; l'output è un insieme di features (caratteristiche)
 - la posizione di un difetto, la posizione e il numero di cellule presenti, i bordi degli oggetti presenti, ...
- Forti connessioni con Pattern Recognition e Machine Learning
- Identificazione e riconoscimento di forme caratteristiche nell'immagine, ad es.
 - Spigoli, linee, bordi
 - Optical Character Recognition (OCR)
 - Riconoscimento di oggetti e scene

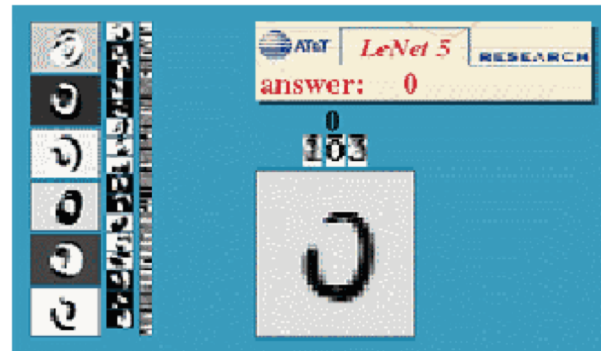
Computer Vision

- L'input è un'immagine; l'output è una descrizione o annotazione semantica del suo contenuto
- Accoppiamento tra image analysis e tecniche di AI
- Obiettivo finale: abilitare i computer ad emulare la visione umana (**percepire**), inclusa la capacità di **apprendere**, di eseguire **inferenze** e di prendere **azioni** sulla base dell'input visuale (Intelligenza Artificiale)

Computer vision

- Optical Character recognition

Mano libera stampatello: ~98%



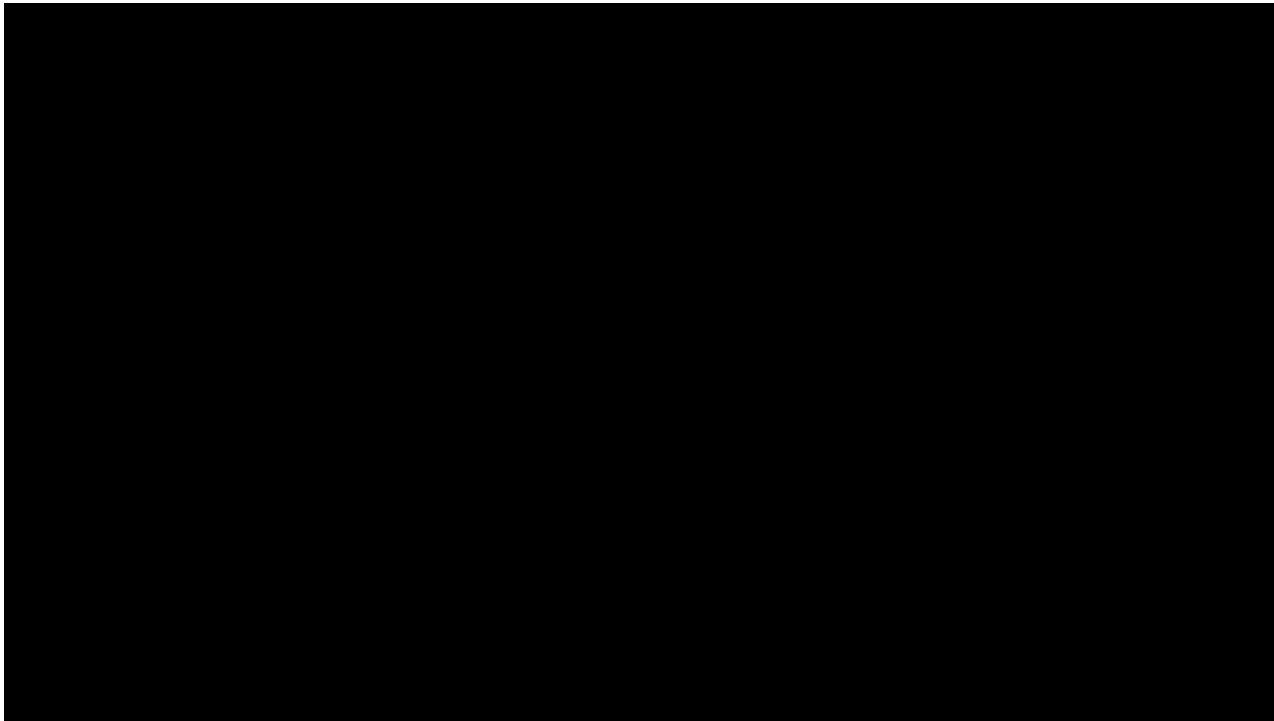
Stampati: ~100%



Mano libera corsivo: difficile

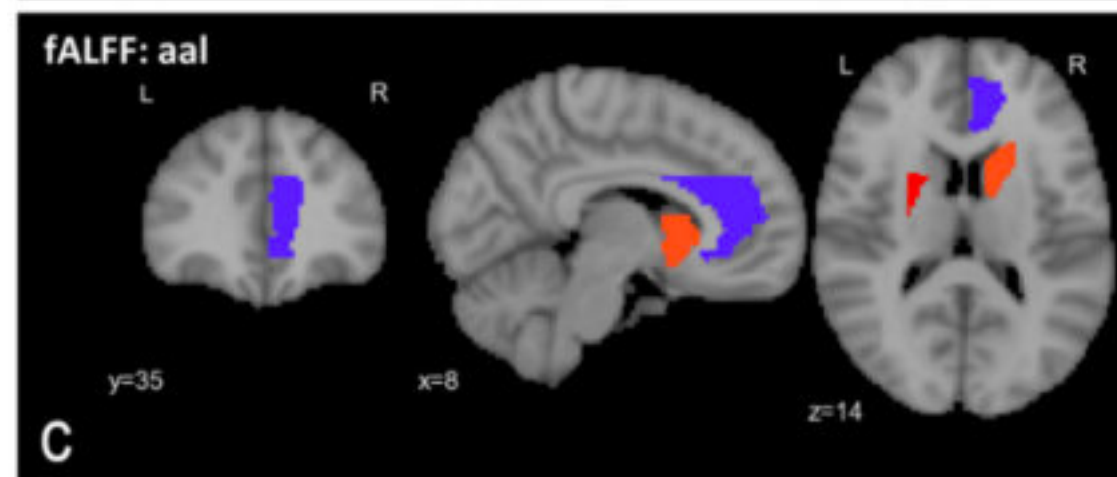
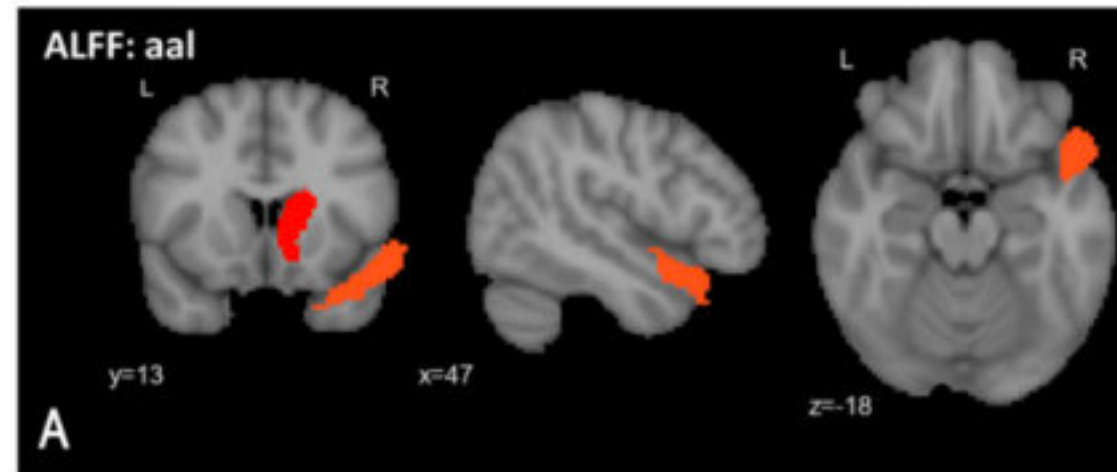
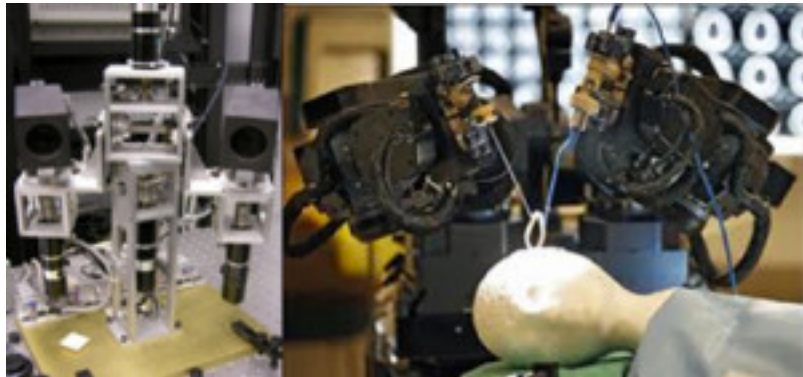
Computer vision

- Veicoli autonomi



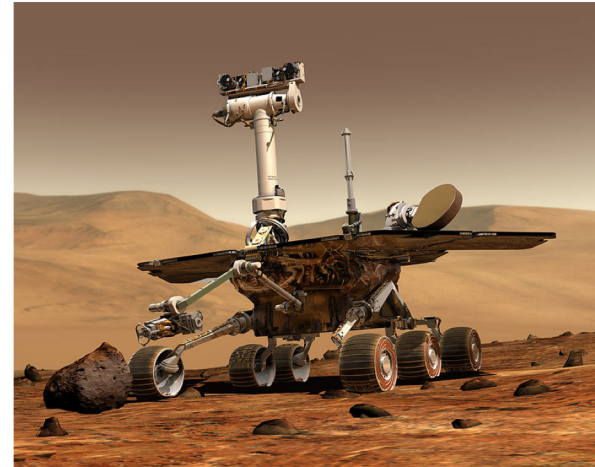
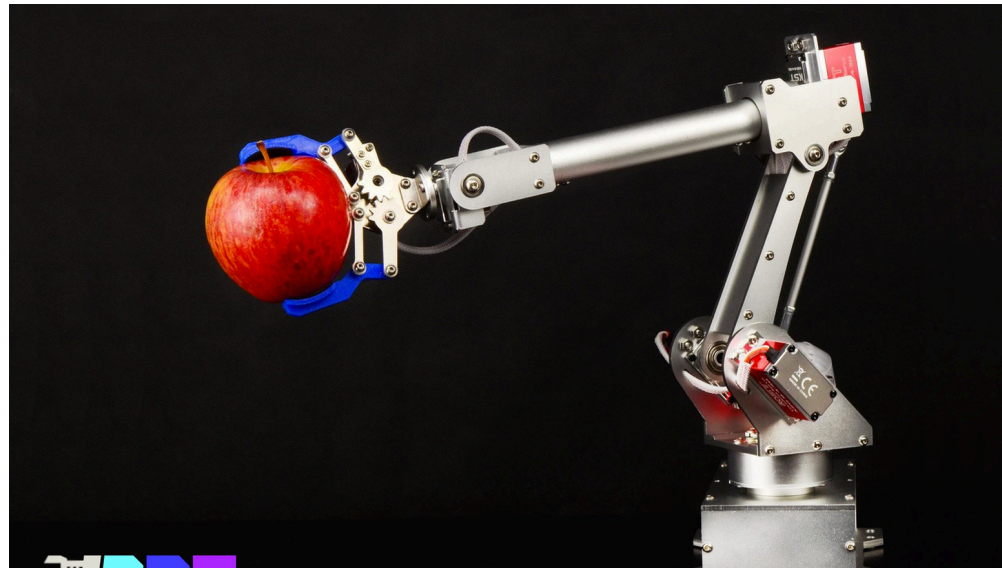
Computer vision

- Diagnostica medica, chirurgia robotica



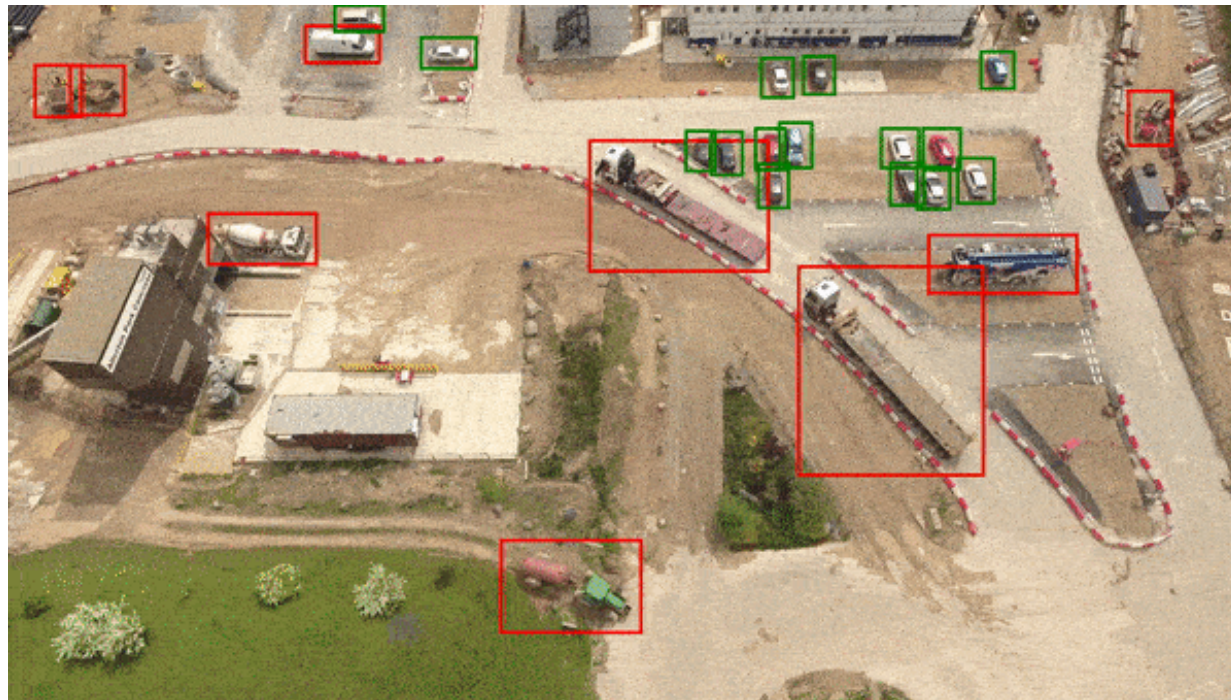
Computer vision

- Robotics



Computer vision

- Droni, sorveglianza



Computer vision

Mobile Visual Search



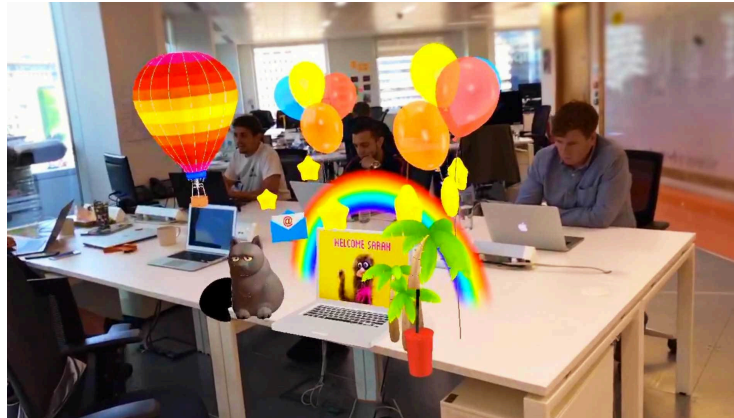
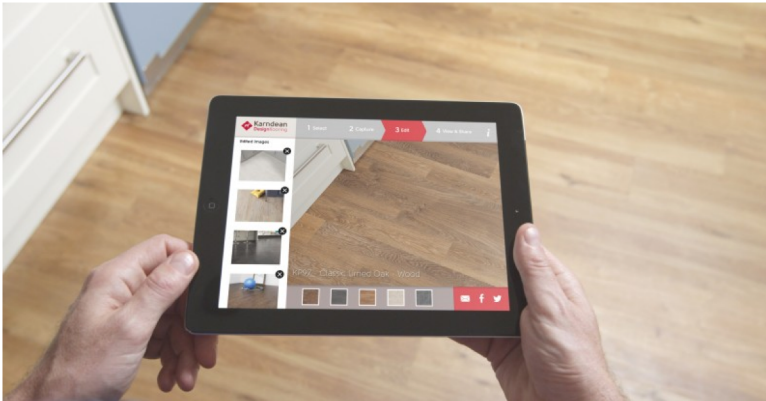
Image retrieval

Classificazione automatica dei Landmarks



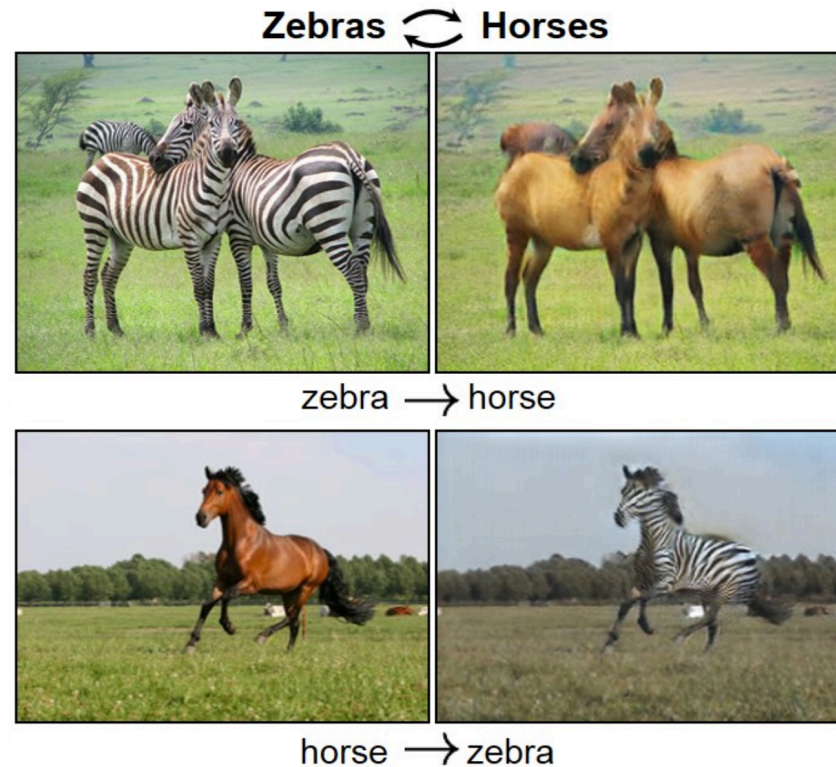
Computer vision

- Realtà Aumentata



Computer vision

- Image/video manipulation



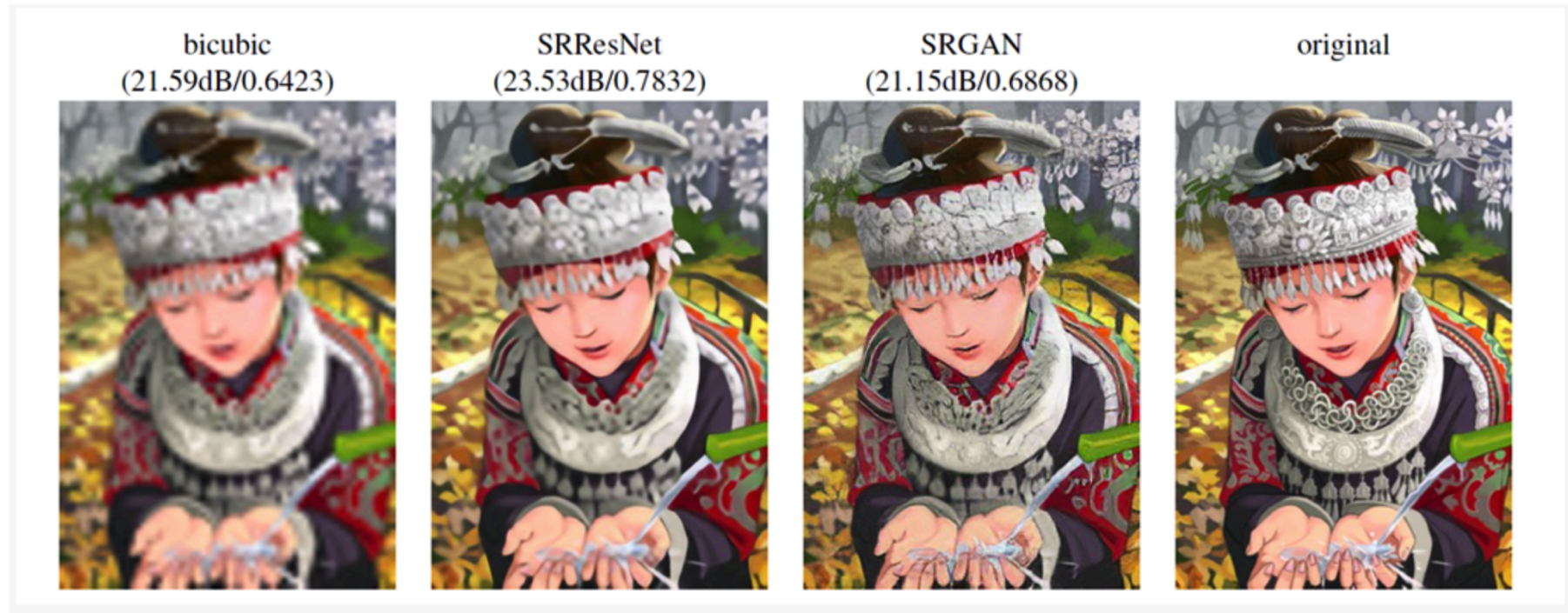
Computer vision

- Image/video manipulation



Computer vision

- Image/video manipulation



Computer vision

- Image/video manipulation



Computer Vision

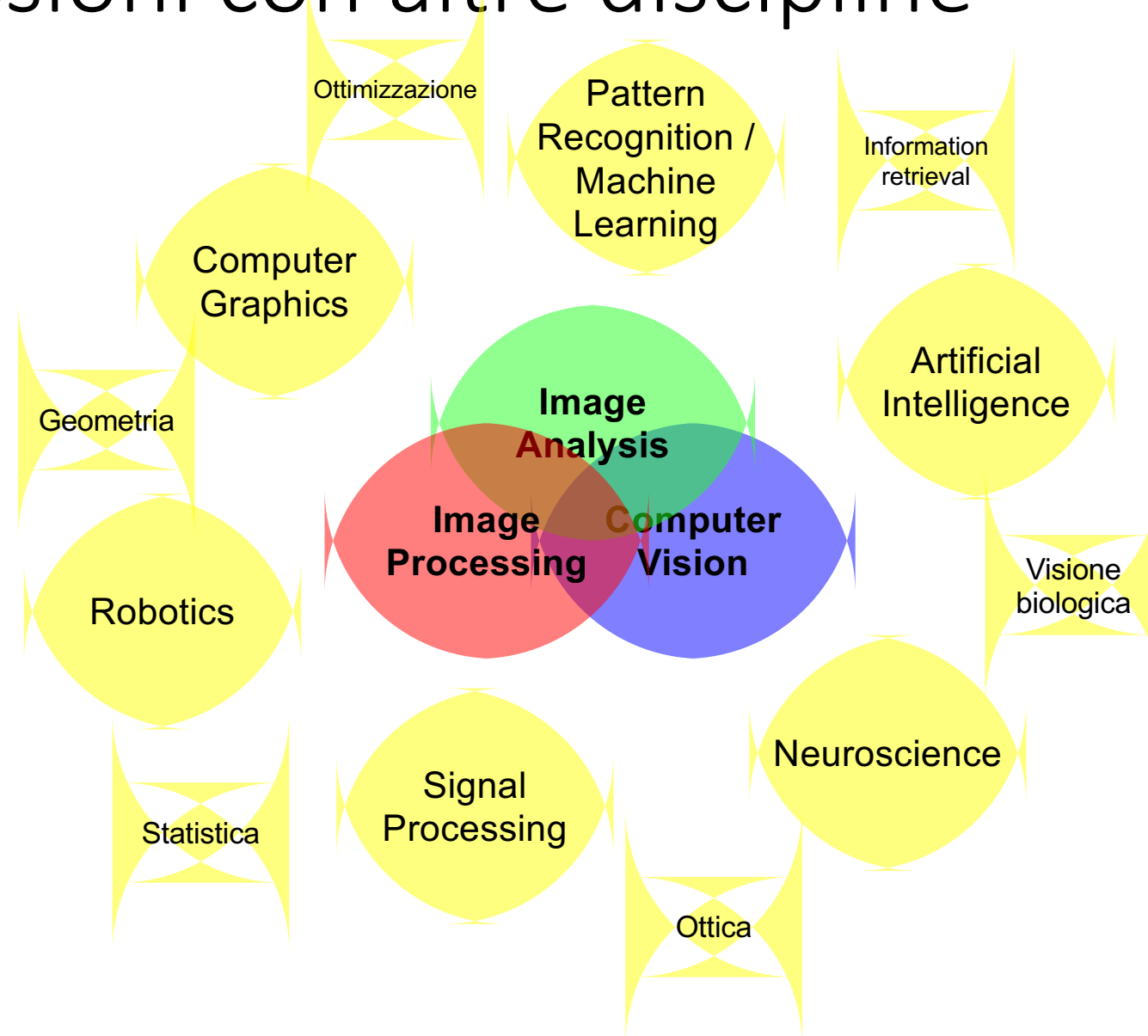
- Arte



Tecniche di elaborazione di immagini

- **Low-level vision (Image Processing)**
 - Manipolazione dell'immagine
- **Mid-level vision (Image Analysis)**
 - Individiazione di bordi, contorni, oggetti di vario tipo, classificazione di scene
- **High-level vision (Computer Vision)**
 - Percepire il contenuto ed eseguire funzioni cognitive
- Funzioni strettamente collegate tra di loro

Connessioni con altre discipline



Difficoltà

- Settore in cui la ricerca è molto attiva
 - visione artificiale non realizzata
- La visione è “naturale” per l'essere umano:
 - Capacità di comprendere “istantaneamente” scene complesse
- Algoritmo su immagine 1000 x 1000:
 - 1 punto = $256^3 = 16.777.216$ combinazioni
 - Possibili immagini = $1.000.000^{16.777.216} > 10^{7.000.000}$
 - La maggior parte saranno non interessanti, ma spazio ad altissima dimensionalità!
- Diverse tecniche sono ormai tecnologia



230	229	232	234	235	232	148
237	236	236	234	233	234	152
255	255	255	251	230	236	161
99	90	67	37	94	247	130
222	152	255	129	129	246	132
154	199	255	150	189	241	147
216	132	162	163	170	239	122

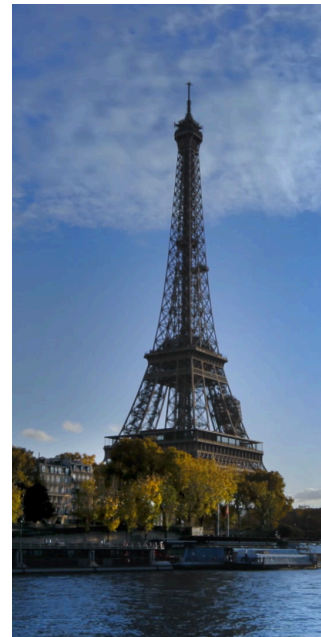
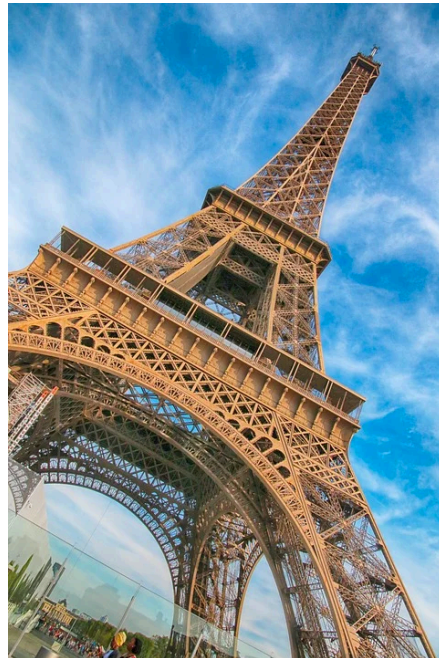
Difficoltà

- Limitazioni sui sensori



Difficoltà

- Variazioni di prospettiva



Difficoltà

- Luminosità variabile



Difficoltà

- Deformazioni



Difficoltà

- Ostacoli



Difficoltà

- Variazioni



Esercizio

- Fare una ricerca sulle applicazioni della Computer Vision
 - Documentare le 5 applicazioni che ritenete più sorprendenti e d'impatto
 - Pagine web, descrizioni, papers...
 - Inviare il materiale al docente
- L'elenco fornito verrà pubblicato sul sito web del corso