

Riconoscimento e tracciamento di elementi su video ad alta risoluzione

Studente: Davide Liu

Relatore: Prof. Lamberto Ballan

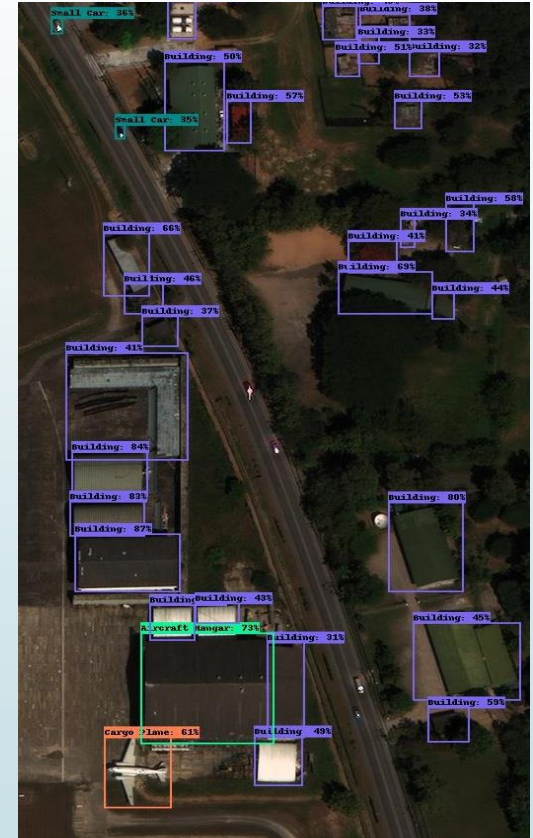
Tutor aziendale: Leonardo Dal Zovo

Università degli studi di Padova
Corso di Laurea Triennale in informatica

Esame di laurea

Azienda - Studiomapp

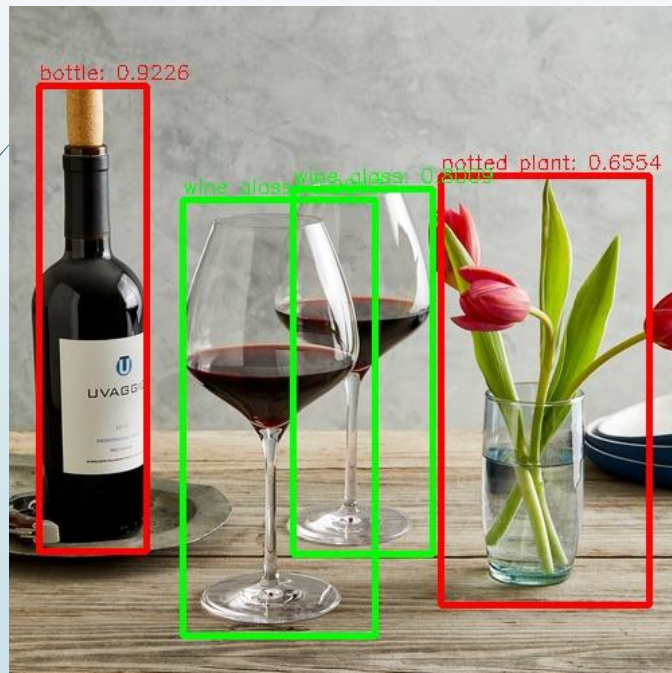
- Startup innovativa fondata a fine 2015 con sede a Ravenna
- Sviluppa algoritmi di intelligenza artificiale applicati alla computer vision
- Studio del territorio tramite analisi di immagini satellitari
- Classificata al quarto posto nella sfida di object detection xView Challenge organizzata dal Pentagono



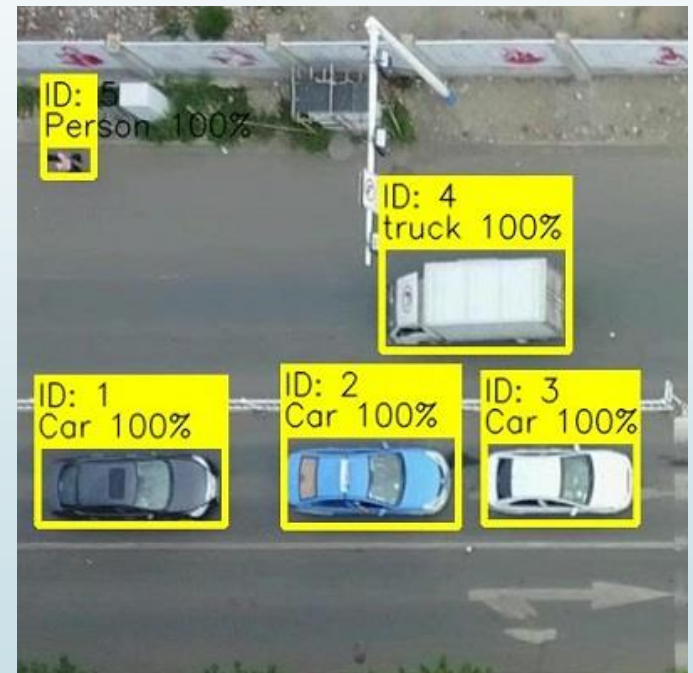
Computer vision

"If we want machines to think, we need to teach them to see" Li Fei Fei

➤ Object detection

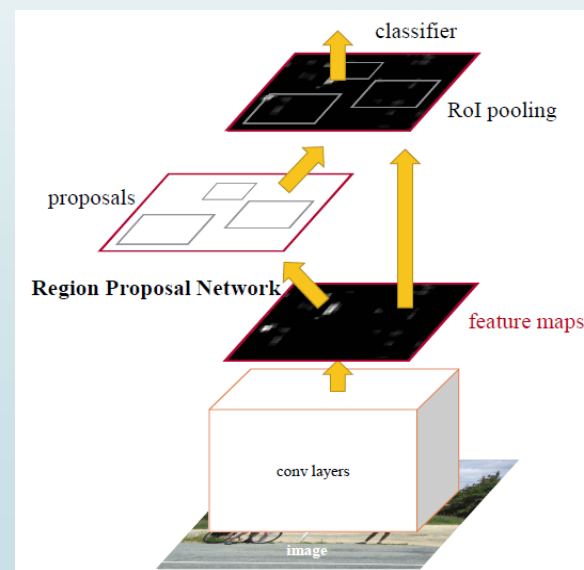
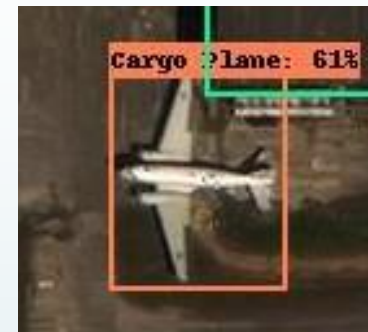


➤ Object tracking



Object detection

- Interpretare il contenuto di un'immagine
- Riconoscere gli elementi presenti in essa:
 - Bounding box
 - Categoria
 - Probabilità
- Stato dell'arte:
 - Convolutional neural network
 - Faster R-CNN



Difficoltà nell'elaborare immagini in alta risoluzione

- I modelli esistenti ridimensionano automaticamente le immagini troppo grandi per adattarle al loro input:
 - Perdita di risoluzione
- Processare immagini a bassa risoluzione è più efficiente:
 - Meno pixels da elaborare
- I datasets utilizzati per allenare i modelli sono composti da milioni di immagini in bassa risoluzione (ImageNet):
 - Immagini in alta definizione più difficili da reperire

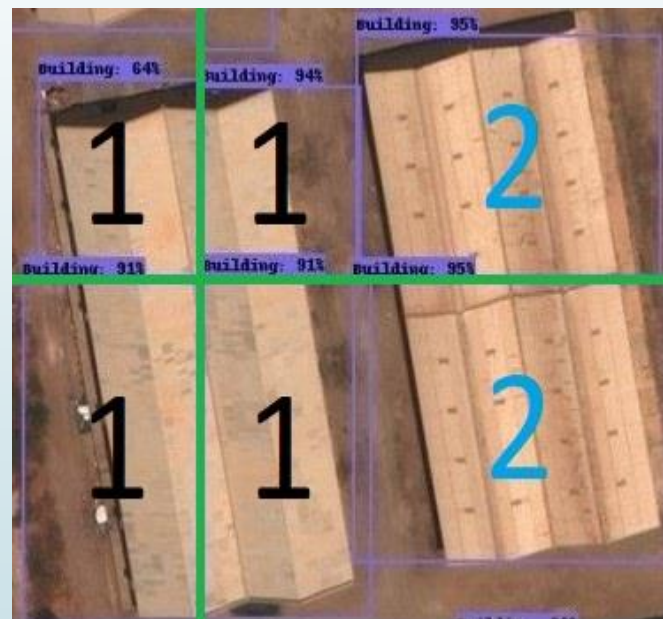
Object detection con frammentazione

- Suddividere un' immagine in regioni
- Effettuare la detection su una regione alla volta
- Eliminare le boxes molto sovrapposte
 - Non-Max Suppression
- Problema: uno stesso elemento può venire individuato più volte
- Algoritmo per ricomporre le detections frammentate



Algoritmo per ricomporre le detections frammentate

- Si raggruppano tutte le boxes che potrebbero appartenere allo stesso elemento:
 - Stessa categoria
 - Vicine tra loro o intersecate
 - Individuate su regioni diverse
 - Dimensioni compatibili
- Queste boxes vengono sostituite con una nuova box che le racchiude tutte



Risultati detection

- Dataset xView composto da immagini satellitari (60 categorie diverse)
- Molti oggetti risultano molto piccoli e sfocati
- Metriche mAP e F1 circa al 25%
 - mAP migliora fino al 1.7%
 - F1 migliora fino al 16.9%
- I miglioramenti sono dovuti ad un incremento della precision e della recall



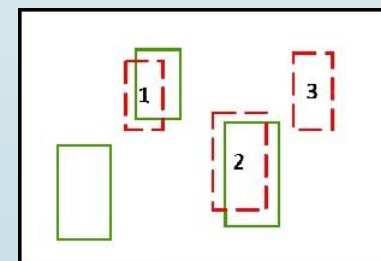
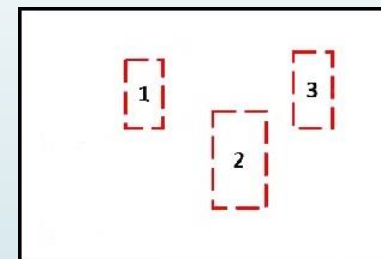
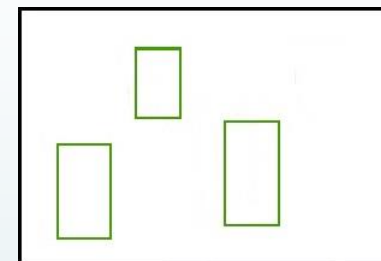
Object tracking

- ▶ Tracciare specifici elementi attraverso una sequenza di frames
- ▶ Assegnazione di un ID univoco ad ogni elemento tracciato
- ▶ Tracciando un oggetto in movimento possono sorgere alcuni dei seguenti problemi:
 - Occlusione
 - Sfocatura
 - Luminosità
 - Etc...



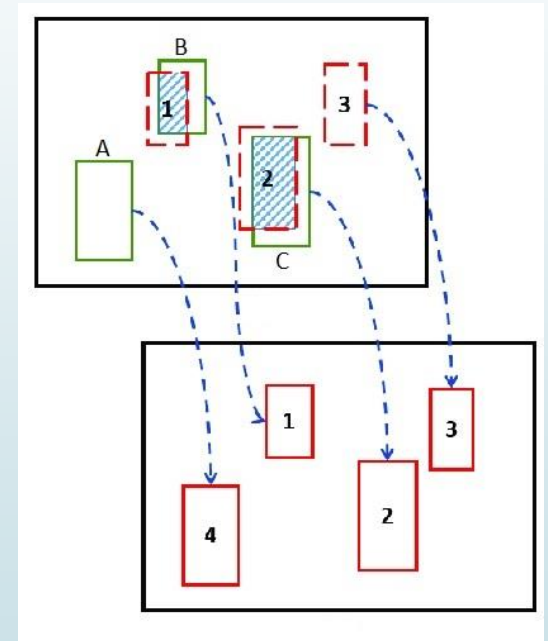
Algoritmo di tracking

- Effettuare una detection per ogni frame
- Sfruttare dei trackers per tracciare gli elementi presenti nel video
- Ogni tracker usa un filtro per prevedere la prossima locazione dell'oggetto tracciato a partire dalle locazioni precedenti
 - Filtro di Kalman
- Assegnare ad ogni tracker il rispettivo oggetto tracciato individuato tramite la nuova detection



Assegnazione tracker-detection

- Calcolare un valore di similarità per ogni coppia tracker-detection:
 - Intersection over Union (IoU)
 - Dimensione delle aree
 - Rapporto tra i lati
- Effettuare gli assegnamenti massimizzando il valore di similarità
- Creare nuovi tracker per tracciare eventuali nuovi oggetti
- Cancellare i trackers che da troppo frames non sono più stati assegnati



Risultati tracking

- Dataset VisDrone2019 formato da video ripresi da un drone
- Solo 5 categorie di elementi:
 - Detections molto più semplici rispetto al precedente dataset
- Circa il 70% degli oggetti vengono tracciati per almeno l'80% della loro durata
- Metriche IDF1 e MOTA circa al 50%
- Aumentando la qualità della detection aumenta anche la qualità del tracking



Grazie per l'attenzione!