

UNIVERSITÀ DEGLI STUDI DI CATANIA

D​IPARTIMENTO​ ​DI​ M​ATEMATICA​ ​E​ I​NFORMATICA

C​ORSO​ ​DI​ L​AUREA​ M​AGISTRALE​ ​IN​ I​NFORMATICA

Interior Room Classifier

Immagine che contiene disegnando

Descrizione generata automaticamente

Progetto Machine Learning

Studenti:

Pietro Vassallo 1000008684

Gioele Cageggi 1000008660 Anno Accademico 2019/2020

Indice

[Introduzione 3](#_Toc41383949)

[Dataset 4](#_Toc41383950)

# Introduzione

La classificazione automatica della scena (talvolta denominata "riconoscimento della scena", o analisi della scena) è un problema di ricerca ricorrente nella Computer Vision, che consiste nell’assegnare un’etichetta come «spiaggia», «camera da letto» o semplicemente «interno» o «esterno» a un’immagine presentata come input, basata sul contenuto complessivo dell’immagine.

Consideriamo il seguente problema affrontato da noi umani: il riconoscimento e la classificazione di un ambiente reale che contiene superfici e oggetti multipli, organizzati in modo significativo avviene in un decimo di secondo o meno, grazie alla nostra capacità di catturare il senso della scena.

Quanto dovrebbe imparare un sistema artificiale prima di raggiungere le capacità di riconoscimento della scena di un essere umano?

Tralasciando la densità delle possibili scene che potrebbero essere classificate il problema viene affrontato con le Convolutional Neural Network. Queste architetture ci permettono di identificare oggetti o scene utilizzando dataset di grandi dimensioni, ma nella seguente relazione ci occuperemo di «allenare» le CNN per adattarle al nostro problema: il riconoscimento di scene interne o dette anche indoor scenes. Pertanto, l’obbiettivo è quello di identificare se la scena in input è una camera da letto, una cucina, un bagno o un salone.

È importante questa tipologia di task perché sono sempre di più gli scenari in cui è necessario per un dispositivo riconoscere l’ambiente interno ad una casa in cui si trova. Consideriamo il caso di un robot che si occupa di pulire le zone di casa che gli vengono assegnate: dovrà necessariamente costruire una mappa della casa per poter capire come muoversi ed identificare le aree di interesse.

Altri esempi potrebbero le applicazioni che si occupano di aggiungere oggetti in realtà virtuale, in modo tale da consigliare in base alla scena inquadrata i corrispondenti elementi da inserire.

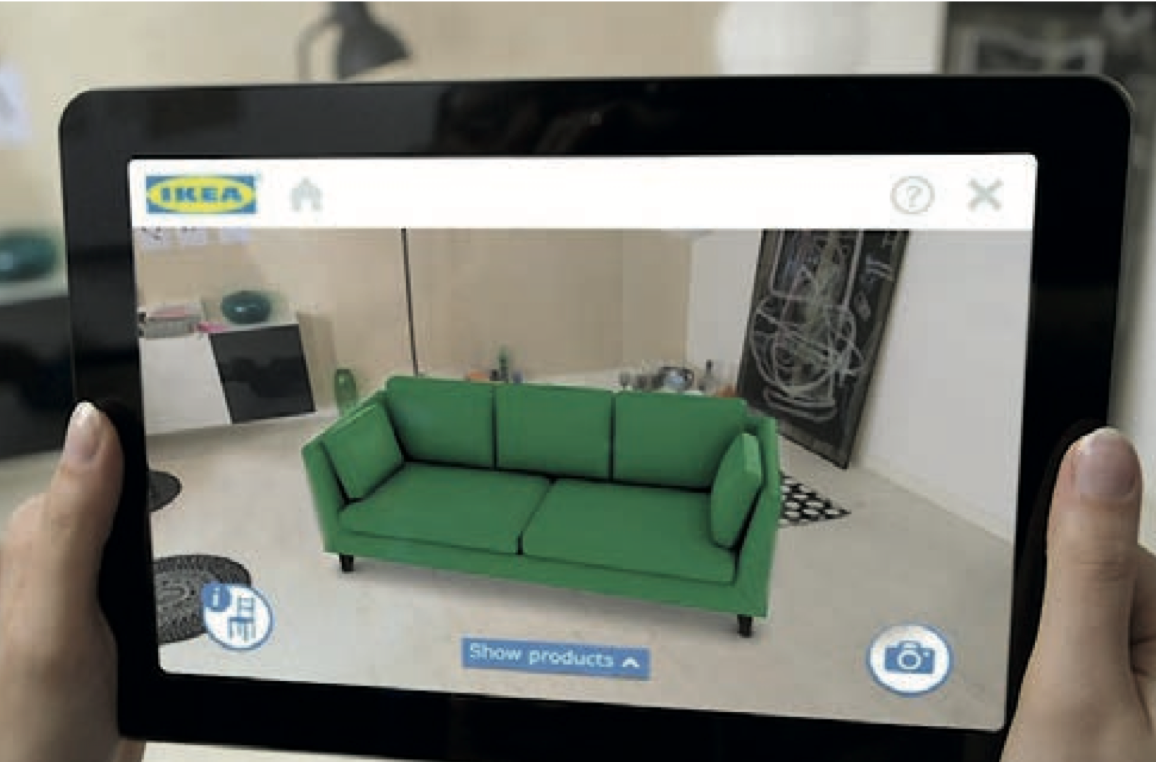


Figure 1- IKEA AR App; Source: IKEA

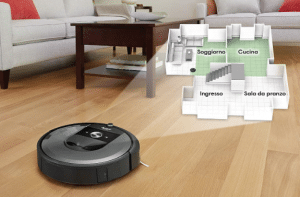


Figure 2-prova

# Acquisizione dati e creazione dataset

Per la realizzazione del dataset sono state acquisite 5821 immagini da due appartamenti.

La cattura è avvenuta mediante due smartphones: iPhone 11 pro e iPhone XR.

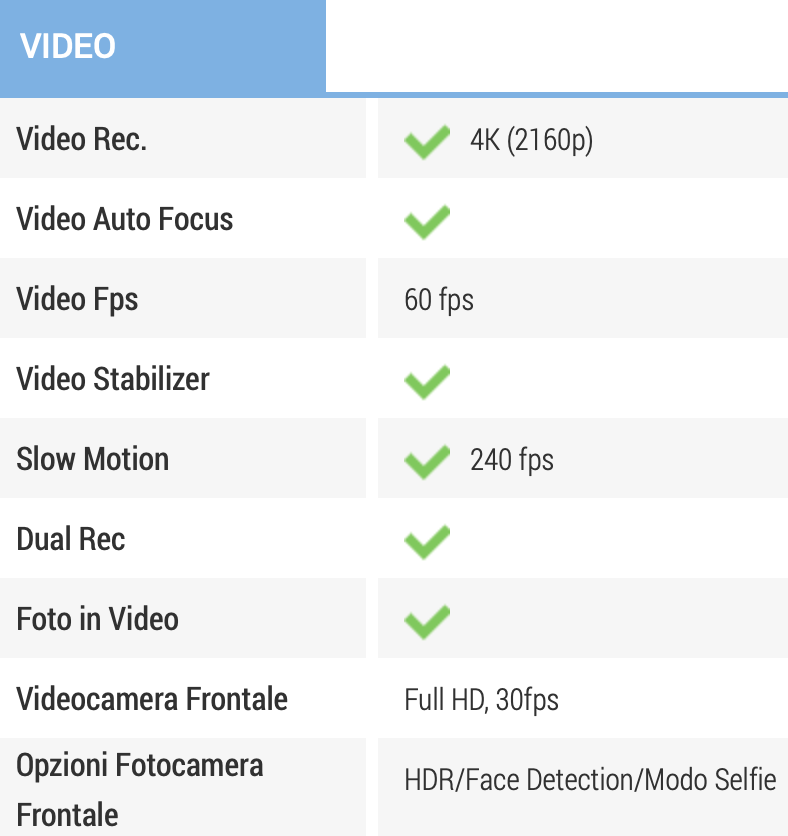


Figure 3 - Scheda Tecnica video Iphone XR

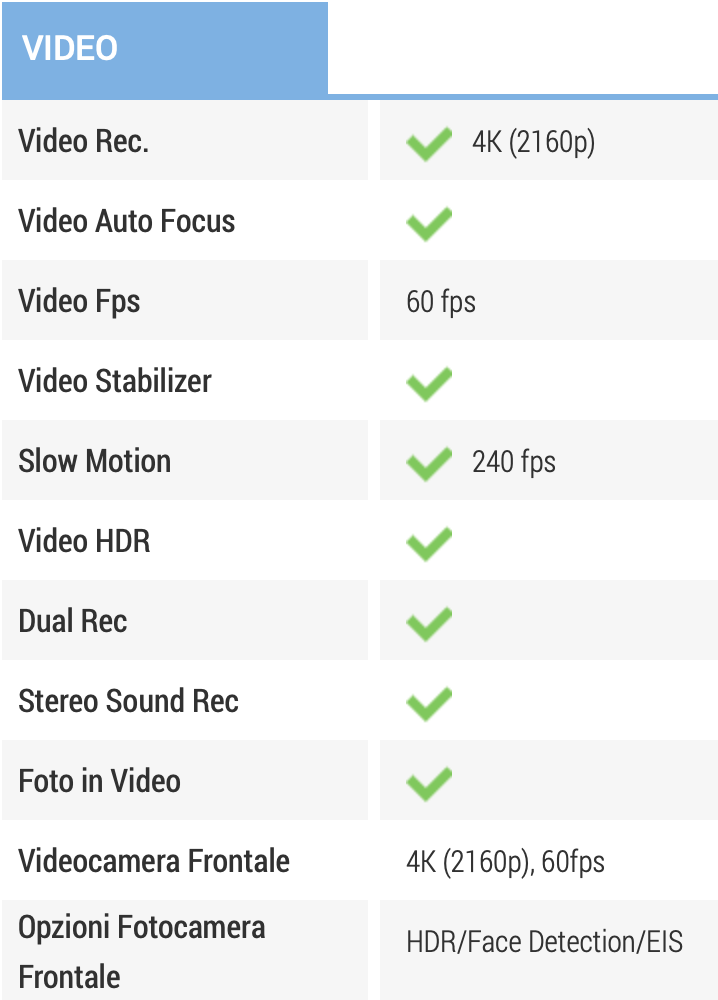


Figure 4 - Scheda Tecnica video Iphone 11 Pro

Le acquisizioni sono state effettuate attraverso la registrazione di video a 360 gradi in punti specifici delle stanze. Tali punti sono stati ottenuti mappando l’ambiente con punti di cattura a distanza di un metro circa tra loro:



Figure 5 - Esempio di mappatura delle stanze per l'acquisizione dei video

Successivamente sono stati estratti dai video i frame (10 per secondo) che sono stati utilizzati per la popolazione del dataset.

Per una prima etichettatura sono state inserite le immagini estratte all’interno di cartelle corrispondenti alla stanza di appartenenza.

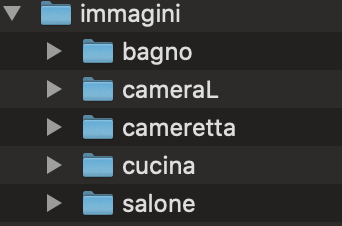


Figure 6 - Struttura classi

Le 5821 immagini contengono rispettivamente:

* 902 frame per il bagno
* 288 frame per la camera letto matrimoniale (cameraL)
* 1213 frame per cameretta
* 864 frame per la cucina
* 2553 frame per il salone

N.B.

La mole di immagini per stanza è proporzionale alla grandezza delle stanza, pertanto le stanze più grandi hanno più frame. Per la camera da letto matrimoniale è stata fatta l’acquisizione solo da un appartamento, questo giustifica il ridotto numero di frame.

Obbiettivo

Dataset

La relazione dovrà essere consegnata in formato doc e pdf e dovrà contenere le seguenti sezioni:

Problema: introduce il problema affrontato in maniera formale dandone le motivazioni e contesti applicativi. Contiene esempi visivi che descrivono il problema.

Dataset: descrive il Dataset utilizzato. In particolare deve contenere dettagli sulla fase di acquisizione (come è stato ottenuto), di etichettatura dei dati (es. tool e modalità utilizzati), informazioni generali sul numero di immagini e le statistiche sulle classi, esempi visuali delle classi, informazioni su training e test set, organizzazione della cartella del dataset (es. nomenclatura file).

Metodo: illustra i metodi proposti e mostra i dettagli tecnici delle soluzioni, motivandole le scelte. Inserire figure che mostrano l'idea generale (in maniera schematica) dei metodi presentati; E' consigliato suddividere la sezione in sottosezioni e inserire figure per agevolare la spiegazione;

Valutazione: descrive quali sono le misure di valutazione utilizzate nella successiva sezione sperimentale.

Esperimenti: presenta gli esperimenti, incluse le prove effettuate per arrivare al risultato proposto, mostra e discute i risultati. E' necessario riportare i risultati in tabella ed è possibile aiutarsi con dei grafici;

Demo: descrive la demo realizzata, come usarla e le sue funzionalità. La demo può essere accompagnata da un video dimostrativo da allegare alla relazione che ne chiarisce l'utilizzo.

Codice: descrive come è organizzato il codice consegnato.

Conclusione: riassume brevemente il lavoro fatto. Quali "lezioni" sono state imparate? Cosa si potrebbe fare per migliorare il metodo proposto?

Servizi di supporto per esperimenti: https://colab.research.google.com/ AWS