
01TUJOV IMAGE PROCESSING AND COMPUTER VISION

ADVANCED FACE AND OBJECT RECOGNITION

Durata prevista dell'esercitazione: 3 ore (2 settimana)

Lo scopo della prima parte di questa esercitazione (prima settimana) è quello di testare due reti neurali, di cui la prima per il rilevamento dei volti e la seconda per quello degli oggetti, facendo ricorso al modulo DNN di OpenCV. L'esercizio 3, invece, permette di utilizzare YOLO, un sofisticato modello di rilevazione degli oggetti (seconda settimana).

WARM UP

La settimana scorsa si è visto come realizzare un sistema di riconoscimento e di tracciamento dei volti usando i classificatori basati su Haar e LBP. Queste tecniche si sono rivelate essere molto efficienti dal punto di vista computazionale, al punto da essere integrate nelle prime fotocamere compatte già nella prima metà del decennio 2000-2010.

Per effettuare tale operazione sono oggi presenti delle tecniche più accurate. Una di queste riguarda l'utilizzo di reti neurali, aventi una precisione molto elevata. Tuttavia, ciò comporta una maggiore richiesta computazionale sia per la fase di addestramento che per quella di test.

ESERCIZIO 1: RILEVAMENTO DEL VOLTO ATTRAVERSO LE RETI NEURALI

Il primo esercizio riguarda il rilevamento dei volti presenti in un'immagine tramite reti neurali. Per addestrarne una in grado di effettuare tale operazione, occorre fornire un dataset contenente un numero elevato di immagini aventi un'adeguata etichettatura dei volti presenti.

Si richiede quindi di effettuare il rilevamento dei volti delle immagini fornite nella cartella del laboratorio tra il materiale del corso e di altre a piacere. A tale scopo, viene messa a disposizione una rete neurale ResNet-10-SSD (Single-Shot Detector) pre-addestrata con il WIDER FACE dataset, composto da più di 32K immagini etichettate e raffiguranti più di 393K volti diversi.

Fare riferimento al primo link presente nella sezione "Link Utili" per scrivere il codice necessario a svolgere tale compito. Fare alcune prove con varie immagini e poi rispondere alle seguenti domande:

- Con che framework è stata addestrata la rete neurale? Fai una ricerca su internet al riguardo. Individua anche altri framework per l'addestramento di reti neurali.
- Cosa contiene il file `res10_300x300_ssd_iter_140000.caffemodel`? E il file `deploy.prototxt`?
- Tutte le rilevazioni avvengono con successo?
- Confronta le rilevazioni ottenute tramite rete neurale con quelle dei sistemi visti la scorsa settimana. Quale dà i risultati migliori?

Successivamente, prova ad estendere il codice al fine di effettuare delle predizioni sia per i frame provenienti dalla webcam del computer che per quelli di un eventuale video fornito dall'utente. A tal proposito, aggiungi due nuovi argomenti al programma per permettere all'utente di specificare il tipo di dato in ingresso. Nel materiale fornito è inoltre disponibile un video, necessario ad effettuare i dovuti test.

Suggerimento: con i video, è possibile notare come la rete neurale effettui le predizioni molto rapidamente. È quindi suggerito limitare la frequenza dei frame per secondo o attendere un input da tastiera per procedere nella prossima predizione.

ESERCIZIO 2: RILEVAMENTO DEGLI OGGETTI ATTRAVERSO LE RETI NEURALI

Le reti neurali, oltre ad essere efficaci nel rilevamento dei volti, possono essere utilizzate con successo anche nel rilevamento degli oggetti. Tale problema è più generale del primo e consiste non solo nell'individuare la presenza di un oggetto nell'immagine e nel racchiuderlo in una bounding box, ma anche nel riconoscerne la relativa tipologia. In fase di addestramento, occorre quindi fornire alla rete neurale un dataset contenente un numero elevato di immagini e un'adeguata etichettatura relativa alle classi che si vogliono individuare.

Si richiede quindi di effettuare il rilevamento degli oggetti delle immagini fornite nella cartella del laboratorio tra il materiale del corso e di altre a piacimento. A tale scopo, viene messa a disposizione una rete neurale MobileNet-SSD pre-addestrata con COCO (Common Objects in Context), un dataset composto da più di 328K immagini etichettate e raffiguranti un totale di 91 diversi oggetti. La lista completa delle classi etichettate è riportata in fig. 5, pag. 7 del file 1405.0312.pdf.

Fare riferimento al secondo link presente nella sezione "Link Utili" per scrivere il codice necessario a svolgere tale compito. Fare delle prove e poi rispondere alle seguenti domande:

- Tutte le rilevazioni avvengono con successo?
- In caso negativo, è possibile effettuare un'operazione di filtro del risultato in modo tale da rimuovere eventuali rilevazioni errate? Come?

ESERCIZIO 3: ANALISI DEL TRAFFICO ATTRAVERSO LE RETI NEURALI

Al link https://github.com/roboflow/supervision/blob/develop/examples/traffic_analysis/README.md è possibile trovare un git che esegue l'analisi del flusso di traffico utilizzando YOLOv8, un modello di rilevamento degli oggetti e di segmentazione delle immagini in tempo reale.

L'esercizio richiede di installare il necessario per far funzionare il sistema, seguendo la guida fornita nel README.md.

SULL'ADDESTRAMENTO DI RETI NEURALI...

L'addestramento di una rete neurale richiede solitamente molto tempo e molta potenza computazionale. Per questo motivo, durante il corso non verrà richiesto di effettuarne alcuno. Se tuttavia dovete trovare tale argomento interessante e vorreste apprendere come effettuare tale operazione, potreste provare a verificare eventuali tesine disponibili al riguardo o, in futuro, eventuali tesi.

LINK UTILI

- Tutorial sulla face detection tramite il modulo DNN di OpenCV:
<https://pyimagesearch.com/2018/02/26/face-detection-with-opencv-and-deep-learning/>
- Tutorial sulla object detection tramite il modulo DNN di OpenCV:
<https://pyimagesearch.com/2017/09/11/object-detection-with-deep-learning-and-opencv/>
- OpenCV, Getting Started with Videos:
https://docs.opencv.org/4.x/dd/d43/tutorial_py_video_display.html