

Ingegneria del Software 2022-2023

Annotatore d'Immagini

Applicazioni che incapsulano algoritmi basati sul machine-learning supervisionato, ossia allenato tramite istanze già "classificate", necessitano, al fine di risultare performanti, di dataset etichettati in maniera molto accurata.

È quindi auspicabile che il dataset di allenamento presenti entrambe le seguenti caratteristiche:

- una varietà di "samples", o campioni, sufficientemente estesa;
- la qualità delle "etichette" associate ai campioni di allenamento deve essere quanto più simile al risultato prodotto da un esperto del dominio.

È facile osservare come tali requisiti siano spesso in conflitto tra loro visto che spesso si utilizzano dataset ad alta numerosità con istanze etichettate in maniera approssimata oppure dataset costituiti di poche istanze etichettate a regola d'arte.

Nel contesto delle immagini, un task importante consiste nell'individuare aree di interesse in singole immagini. Tali aree di interesse dipendono strettamente dal contesto applicativo, per esempio, nel campo dell'automotive, potremmo essere più interessati a rilevare i bounding-box di pedoni rispetto a quello di altri veicoli, e comunque, in ogni caso, le due tipologie di bounding box vanno separate semanticamente.

Il presente progetto consente a un'esperto di dominio di etichettare un'immagine con le relative regioni di interesse al fine di creare dataset ad alta numerosità etichettati con bounding box precisi sia dal punto di vista della posizione spaziale all'interno dell'immagine sia dal punto di vista della loro classificazione semantica.

Descrizione Breve

Si realizzi un annotatore di immagini, ossia un tool per annotare manualmente immagini provenienti da diverse fonti, ad esempio: immagini mediche; foto di documenti testuali; foto di persone/oggetti.

In particolare, l'applicazione deve consentire:

- l'annotazione manuale di immagine attraverso bounding box a cui deve essere associata un'etichetta (TAG) obbligatoria e un valore di testo (NOTE) opzionale;
- il salvataggio delle annotazioni. Le annotazioni, così come le immagini sono persistenti, i.e., devono essere salvate su disco tramite la struttura a cartelle/file del sistema operativo o tramite una base di dati.

Per esempio, ogni immagine ha associato un file con le sue annotazioni in formato **.csv** con campi **tag**, **note** , **x**, **y**, **larghezza**, **lunghezza**. Un'immagine può ricevere più di un'annotazione;

- la possibilità di ingrandire/ridurre l'immagine e, inoltre, di navigare all'interno della stessa per migliorare la precisione dei bounding boxes;
- la possibilità di aggiungere/modificare/eliminare tags;

- la possibilità di caricare eliminare immagini dalla libreria dell'annotatore. A tal fine si consiglia di copiare le immagini all'interno libreria dell'annotatore e di identificarle, all'interno di tale libreria, tramite un nome univoco.

Descrizione delle Componenti

Il software si compone principalmente di due finestre. La prima, quella dove l'utente svolge la maggior parte del lavoro di annotazione, la chiameremo Working Image View (WIV).

La WIV consente di annotare la singola immagine. La seconda finestra, detta Image Manager (IM), è dedicata alla gestione della collezione di immagini annotate e del set di tags disponibili nella WIV.

Di seguito si dettagliano le componenti principali sia per WIV che per IM.

Working Image View

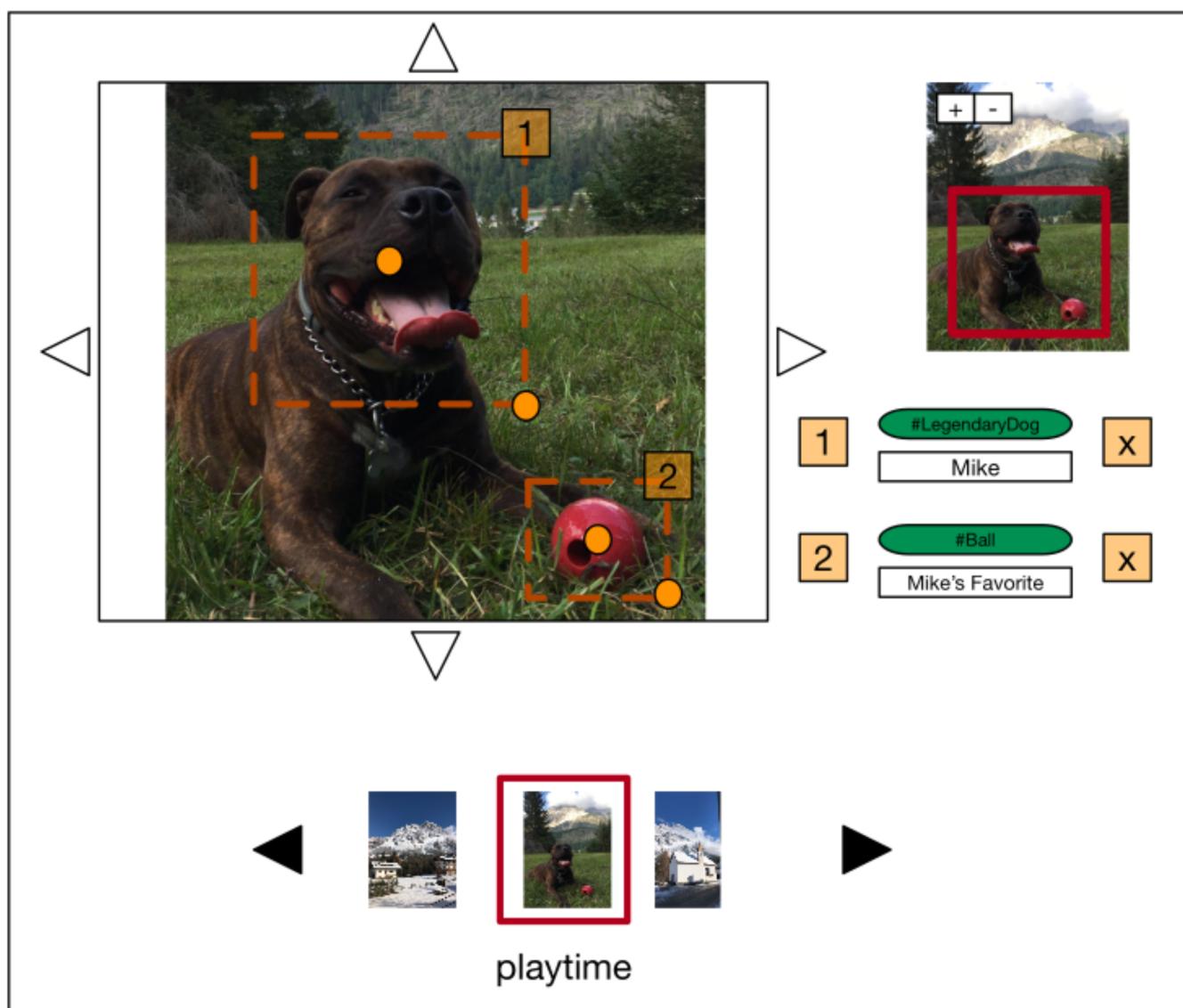


Figura 1 – un esempio di Working Image View.

Le componenti della WIV vanno organizzate all'interno di una singola finestra. Il layout di tale finestra è a discrezionalità del designer, in Figura 1 ne viene comunque proposto un esempio, che si può riprodurre o da cui si può prendere spunto. Le funzionalità della WIV sono descritte di seguito facendo riferimento all'esempio di Figura 1:

main image window

Nella regione in alto a destra di Figura 1 abbiamo la rappresentazione della porzione d'immagine in cui l'utente può introdurre nuove annotazioni. Quando l'utente clicca su una parte dell'immagine viene introdotto un nuovo rettangolo mobile e ridimensionabile a cui viene associato un numero progressivo, e la corrispondente entry nella tag list (si veda sotto-sezione corrispondente). Le quattro frecce bianche ai lati della regione servono per navigare nell'immagine muovendosi lungo le quattro direzioni cardinali senza variare il livello di zoom corrente;

image navigation map

Nella parte in alto a destra di Figura 1 abbiamo l'image navigation map che rappresenta l'immagine nella sua totalità. Trascinando il rettangolo rosso all'interno di questa componente, è possibile spostare la porzione visualizzata all'interno della main image window. Sono, infine, presenti due pulsanti +/- per decidere il livello di zoom nella main image window;

tag list

Questa componente rappresenta una lista di triple **numero-rettangolo, tag, valore**. Il valore numero-rettangolo è progressivo e viene inserito com nuovo elemento automaticamente dopo l'inserimento di un nuovo rettangolo nella main image window. In seguito sarà possibile selezionare il relativo tag tra la lista di tag disponibili (si veda tag disponibili nella descrizione della finestra Image Manager) tramite un menu drop-down. Inoltre, il valore associato al tag, se presente, può essere digitato tramite text-field. Infine, vi è un bottone **x** associato a ogni elemento della lista. Tale pulsante consente di eliminare l'elemento della lista corrispondente e il relativo rettangolo associato nella main image window;

slideshow

Questa componente consiste una lista navigabile, tramite le due frecce poste ai lati, e rappresenta l'insieme delle immagini caricate nell'Image Manager. L'immagine correntemente selezionata corrisponde all'immagine nella main image window e nelle altre componenti descritte sopra. Quando lo slideshow viene fatto scorrere a destra/sinistra l'immagine corrispondente viene caricata nelle componenti descritte sopra sostituendo quella attuale.

Image Manager

Come per la Working Image View le componenti dell'Image Manager vanno organizzate all'interno di una singola finestra. Il layout di tale finestra è a discrezione del designer, in Figura 1 ne viene comunque proposto un esempio, che si può riprodurre o da cui si può prendere spunto. Le funzionalità dell'Image Manager sono descritte di seguito facendo riferimento all'esempio di Figura 2:

image set

Nella regione in alto a sinistra di Figura 2 viene rappresentato il set di immagini caricate dall'utente. Queste corrispondono alle immagini navigabili tramite lo slideshow della Working Image View.

insert image form

Questa componente (rappresentata nella parte in alto a destra di Figura 2) consente all'utente di caricare una nuova immagine. Di conseguenza l'image set e, conseguentemente, lo slideshow della Working Image View, vengono aggiornati. Per inserire una nuova immagine l'utente trascina l'immagine nel rettangolo tratteggiato inserisce nel textfield sottostante il nome che vuole associare all'immagine e preme il tasto + a fianco di tale textfield.

available tags

Questa componente (rappresentata nella parte bassa di Figura 2) contiene l'insieme di tags consentiti nella tag list della Working Image View (i.e., costituiscono i possibili valori del drop-down menu citato nella descrizione della tag list). Per inserire un nuovo tag l'utente scrive il suo nome nel textfield contenuto nella componente (qui rappresentato sul fondo di Figura 2) e preme il tasto +. Inoltre, i tag si possono anche eliminare cliccando sul pulsante x posto a fianco di ogni tag nell'insieme.

The screenshot shows a user interface for managing images. On the left, there is a grid of six images with their corresponding tags below them:

- A snowy mountain scene labeled "sanvito"
- A brown dog labeled "gnam"
- A snowy mountain scene labeled "antelao"
- A dog playing with a ball labeled "playtime"
- A small white church in a snowy landscape labeled "villanova"
- A large mountain range labeled "pelmo"

To the right of the grid is a dashed rectangular area containing the text "+ Drag a new Image Here". Below this area is a text input field labeled "Image Name" and a small "+" button.

Below the main grid is a section titled "Available Tags" with a list of tags:

- #LegendaryDog
- #Ball
- #Mountain
- #Deer
- #Cat
- #Food
- #Cake

Each tag is enclosed in a green rounded rectangle with an orange "X" button to its left. At the bottom of this section is another text input field labeled "Tag Name" and a small "+" button.

Figura 2 - un esempio di Image Manager.

Domande Frequenti

Ecco alcune domande (con relative risposte) che sono pervenute nella precedente iterazione dell'esercizio. Se avete ulteriori domande sulla consegna contattate il docente responsabile (Pietro Sala) via mail all'indirizzo pietro.sala@univr.it per prendere un appuntamento.

- Cosa succede se un rettangolo di annotazione sta parzialmente nella working image view?

Questo può accadere, ad esempio, se il livello di zoom è troppo alto e/o la parte attualmente visualizzata non contiene tutto il rettangolo di annotazione in quanto spostata rispetto a questo. Soluzione: si visualizzino solo i rettangoli di annotazione che sono *totalmente* contenuti all'interno della working image view corrente, l'utente potrà sempre visualizzare gli altri rettangoli spostando l'immagine o variando il livello di zoom;

- Cosa succede se cancello un tag nella lista available tags di IM e vi sono annotazioni associate nella libreria dell'applicazione?

Soluzione: si eliminano tutte le annotazioni associate al tag eliminato.