

InfoTourist

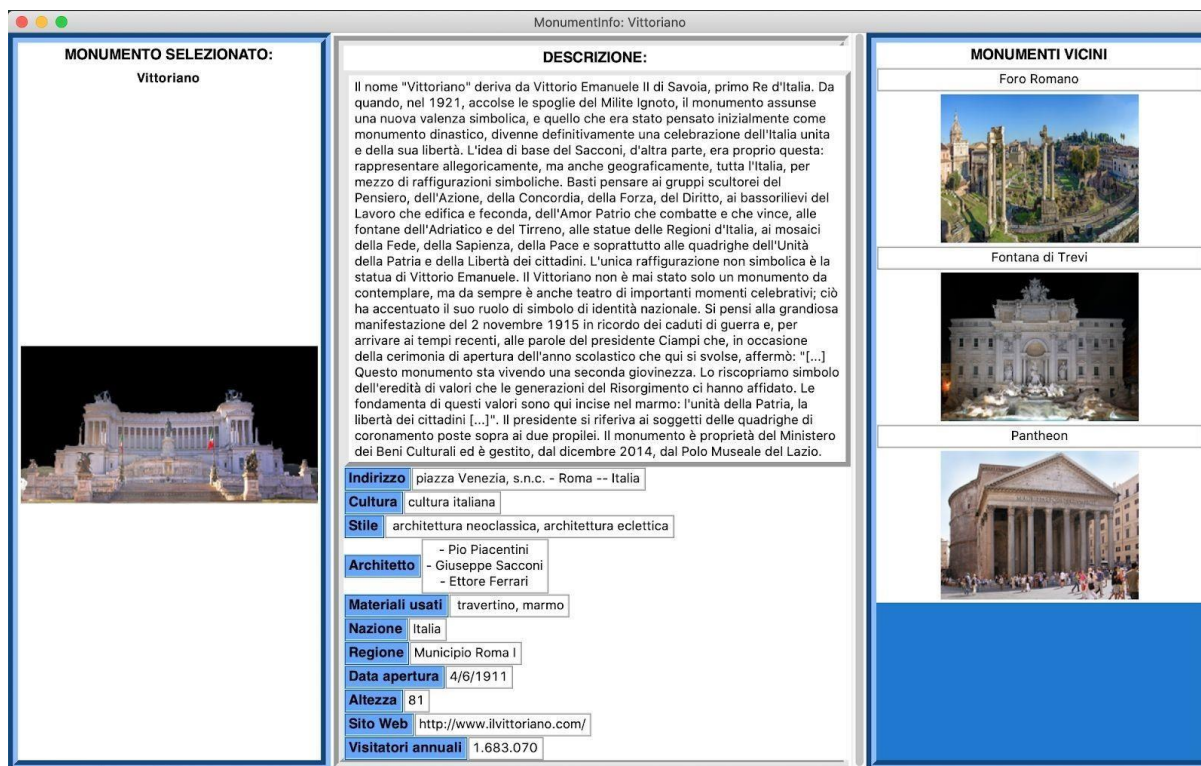
Gianluca D'Aniello - 727663

repository github: <https://github.com/gianluca2000-daniello/InfoTourist-Icon.git>

Introduzione

InfoTourist è un software che ha lo scopo di riconoscere un monumento, a partire da un'immagine. Infatti, data un'immagine fornita dall'utente, il programma restituisce una descrizione del monumento rappresentato nell'immagine e consiglia altri monumenti/luoghi e/o attrazioni vicine visitabili.

Gui All'avvio del programma è possibile selezionare un monumento che sarà oggetto del riconoscitore. Nella cartella *test* del repository sono presenti i monumenti tra i quali è possibile effettuare questa scelta.



Il programma implementa un'interfaccia grafica: sulla colonna di sinistra è rappresentata l'immagine del monumento riconosciuto, sulla colonna centrale è presente una breve descrizione di quest'ultimo e, infine, sulla colonna a destra sono riportati i tre luoghi più consigliati da visitare nella zona.

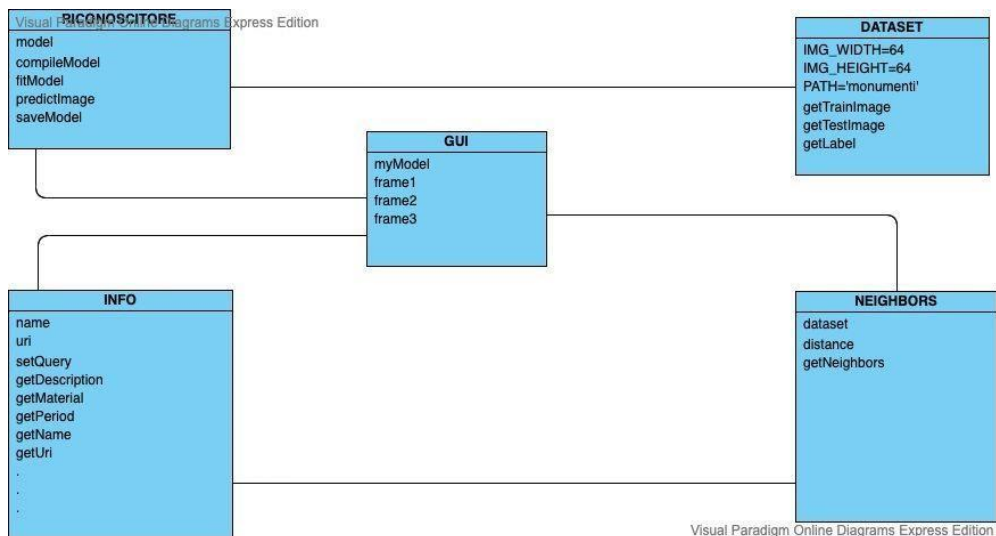
Descrizione progettuale

repository

modello	14/02/2020 11:48	Cartella di file	
monumenti	14/02/2020 11:48	Cartella di file	
test	14/02/2020 11:48	Cartella di file	
dataset.py	14/02/2020 11:48	File PY	2 KB
gui.py	14/02/2020 11:48	File PY	8 KB
infoMonumento.py	14/02/2020 11:48	File PY	11 KB
kfCV.py	14/02/2020 11:48	File PY	5 KB
myModel.py	14/02/2020 11:48	File PY	2 KB
neighbors.py	14/02/2020 11:48	File PY	3 KB
trainMyModel.py	14/02/2020 11:48	File PY	1 KB

- la cartella modello contiene file "myModel.h5" corrispondente al modello salvato del riconoscitore;
- la cartella monumenti contiene altre sottocartelle di immagini di monumenti classificati utilizzati come train e due file .csv:
 - label.csv che contiene i nomi dei monumenti, utilizzato per effettuare le query;
 - monuments.csv contiene le (LAT, LON) dei monumenti italiani con la proprietà "attrazione turistica" su WikiData. Il file è utilizzato nel k-NN;
- la cartella test contiene le immagini utilizzabili per il test programma.

Classi



- La classe **GUI (gui.py)**, che implementa l'interfaccia, richiama le altre classi a supporto dell'esecuzione del programma.

- La classe **RICONOSCITORE (myModel.py)** implementa il riconoscitore delle immagini, addestrato su un training set composto da immagini di monumenti già classificati. La classe compila e crea un modello, il quale potrà essere salvato e usato come riconoscitore.
- La classe **INFO (infoMonumento.py)** implementa una serie di query per estrarre informazioni relative al monumento riconosciuto.
- La classe **DATASET (dataset.py)**, dato un path, estrae le immagini dalle sotto cartelle per creare i subset di train e validation, utili all'addestramento del riconoscitore.
- La classe **NEIGHBORS (neighbors.py)** implementa il recommender system.

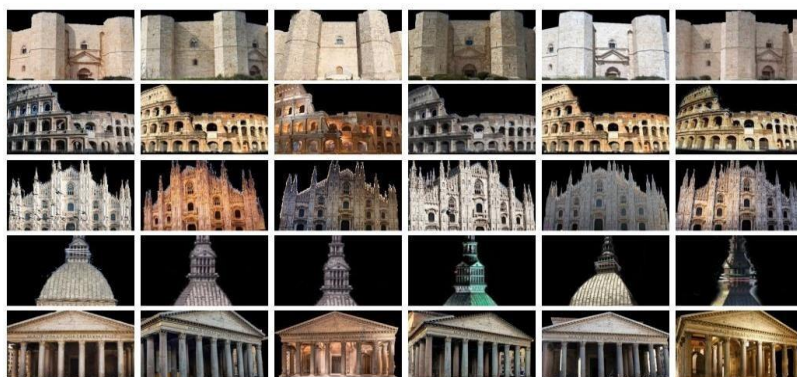
Dettagli di implementazione

Il programma è stato sviluppato in python 3.7

Esso implementa una **rete neurale convoluzionale (CNN)** per la classificazione delle immagini. Inoltre, grazie all'uso di **ontologie online** e **Linked Open Data**, è possibile ritrovare informazioni strutturate, estratte dal **Semantic Web**, per poter fornire all'utente una completa descrizione del monumento riconosciuto. Inoltre, all'utente viene fornita una serie di monumenti vicini visitabili. Utilizzando l'algoritmo di **apprendimento supervisionato case-based K-NN** ($k=3$) ritroviamo la lista dei neighbors sulla base di un nuovo dataset appositamente creato, utilizzando altre informazioni estratte da un' ontologia online (**WIKIDATA**).

Dataset

Di seguito è mostrato un esempio di alcune immagini presenti nel dataset



TRAINING SET

Il dataset è stato creato utilizzando delle immagini prese dal web, le quali sono state successivamente scontornate per ovviare al numero limitato di immagini.

Per ogni monumento ci sono circa 15 foto. Il 90% di esse è utilizzato per il training e il restante 10% per validation. Le immagini sono ridotte ad una dimensione di 64x64 e sono trasformate in array (matrice) di pixel con valori scalati tra 0 e 1.

Per fare ciò è utilizzata la libreria *preprocessing.image* di Keras (sottolibreria di tensorflow) che fornisce metodi per il preprocessing delle immagini e contenitori (**ImageDataGenerator**) che saranno istanziati con l'insieme di immagini di train e quelle di validation.

Riconoscitore

La rete neurale è stata creata utilizzando le librerie di keras.

In particolare con *keras.models* è stata importata la classe fondamentale che rappresenta la sequenza di layers del nostro modello. Il nostro modello, ossia l'insieme dei "pezzi" che andranno a comporre la nostra rete neurale artificiale, sarà composto da una sequenza lineare di layers. **Sequential** è la classe che permette di aggiungere via via layer dopo layer dall'input all'output della sequenza. Dalla libreria "*tensorflow.keras.layers*" sono stati importati i vari livelli: **Dense, Conv2D, Flatten, MaxPooling2D**

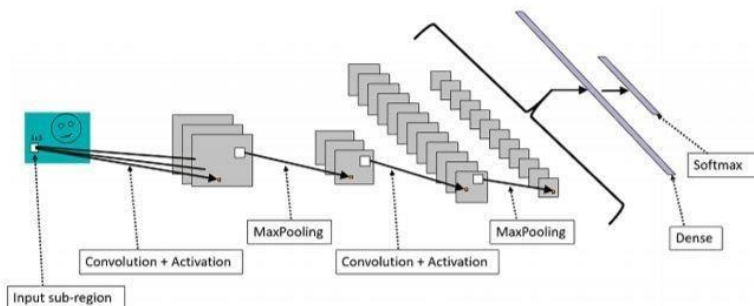
- **Conv2D:** è il livello che si occupa di individuare schemi, come ad esempio curve, angoli, circonferenze o quadrati raffigurati in un'immagine con elevata precisione. Per attuare il processo di Convoluzione si fa scorrere un certo Filtro (chiamato Kernel) sul tensore (l'immagine). Un filtro è un insieme di pesi rappresentati sotto forma di un tensore di dimensione $M \times N$ e la cui profondità è la medesima dell'immagine di input. Quando questo filtro viene fatto scorrere sull'immagine, la regione su cui si trova in quel momento è detta *Campo Recettivo*. Durante questo processo, i valori del filtro vengono moltiplicati ai pixel originali, ed infine sommati in un singolo valore scalare che rappresenta l'output di quel passo di convoluzione.

Questo processo viene ripetuto fino ad arrivare alla fine dell'immagine. Il tensore risultante è la rappresentazione di tutti gli output del processo di convoluzione e si chiama *Activation Map* (Mappa di attivazione)

- **MaxPooling2D:** dopo ogni layer convoluzionale si può applicare un'operazione chiamata Pooling che consiste nel creare una versione ridimensionata dell'output dello strato precedente in modo da ridurre le dimensioni del risultato ed ottenere una rappresentazione più compatta delle feature iniziali.

- **Flatten:** è un livello che si occupa di ridurre la dimensionalità dell'output dei vari strati convoluzionali in modo tale da ottenere in uscita un unico vettore. Questo vettore sarà l'input dell'ultima parte della rete neurale.
- **Dense:** Il vettore finale del Flatten viene usato come input per una rete neurale 'classica' in cui vengono utilizzati strati completamente interconnessi (dense layers) di neuroni che creano la rappresentazione finale dell'output.

Esempio di rappresentazione grafica CNN



MyModel

Il modello di rete neurale è il seguente:

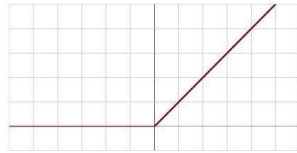
```
self.model = Sequential([
    Conv2D(16, 3, padding='same', activation='relu', input_shape=(64, 64, 3)),
    MaxPooling2D(),
    Conv2D(32, 3, padding='same', activation='relu'),
    MaxPooling2D(),
    Conv2D(64, 3, padding='same', activation='relu'),
    MaxPooling2D(),
    Flatten(),
    Dense(512, activation='relu'),
    Dense(train.num_classes, activation='softmax')
])
```

Esso è formato da 9 layers:

- Il primo layer convoluzionale prende come input un'immagine nelle dimensioni specificate (64x64x3), ha 16 neuroni di output e crea una finestra kernel 3x3 per creare la mappa di attivazione. La funzione di attivazione utilizzata è la ReLU:

- *Rectified Linear Unit (ReLU)*:

$$f(x) = \max(0, x)$$



- A seguire c'è un livello di Pooling.
- Ci sono, poi, altre 2 coppie di layers: <Conv2D e MaxPooling2D> che aumentano progressivamente le dimensioni dell'output, da 16 a 32 e da 32 a 64.
- E' presente, poi, un livello Flatten per appiattire l'output del livello precedente.
- Seguono, poi, livelli completamente connessi: col fine di connettere tutti i neuroni del livello precedente e di stabilire le varie classi identificative secondo una determinata probabilità. L'ultimo layer Dense ha come output il numero di classi in cui classificare gli esempi e la funzione di attivazione utilizzata è la *softmax* che restituisce un vettore di probabilità. Esso contiene, per ogni esempio, il valore di probabilità che l'esempio ha di appartenere ad una certa classe.

Una volta creato il modello della CNN bisogna compilarlo:

```
def compileModel(self):
    self.model.compile(optimizer='adam',
                       loss='binary_crossentropy',
                       metrics=['accuracy'])
```

Nel caso in questione, è stato usato l'ottimizzatore *ADAM*, e come metrica di errore da minimizzare la *binary_crossentropy*.

Utilizzo, poi, *l'accuratezza* per confrontare l'efficacia della rete sul validation set.

- **L'OTTIMIZZATORE** specifica l'algoritmo scelto per effettuare la **discesa di gradiente**
- La **LOSS** utilizzata è la log loss: $\log \text{ loss} = -(y \log(p) + (1-y) \log(1-p))$ dove y è la classe reale, mentre p è quella predetta dalla CNN

Per effettuare il train richiamo il metodo *fit*:

```
def fitModel(self):
    total_train = self.train.samples
    batch_train = self.train.batch_size

    total_val = self.val.samples
    batch_val = self.val.batch_size

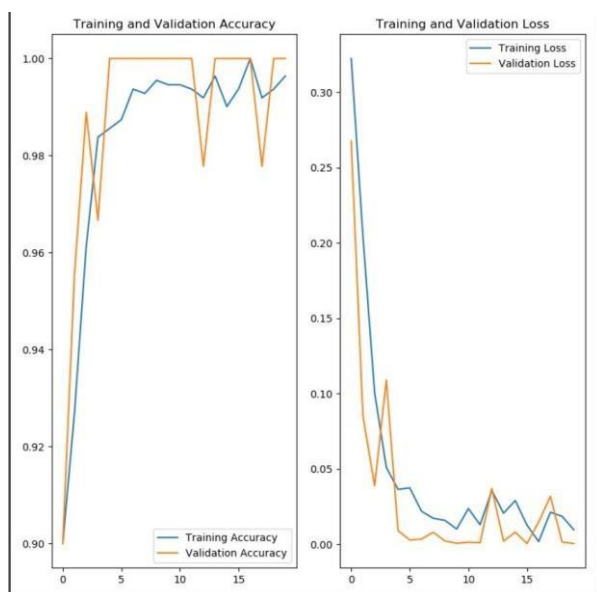
    self.model.fit_generator(
        self.train,
        steps_per_epoch=total_train // batch_train,
        epochs=self.epochs,
        validation_data=self.val,
        validation_steps=total_val // batch_val
    )
```

Qui è specificato al modello l'insieme di train, l'insieme di validazione, e il parametro *epochs* che indica il numero di 'epoche' (volte) in cui effettuare il train sui dati.

Impostazione parametri

Dopo aver effettuato una **K-fold Cross Validation**, con $k = 5$, ho scelto i parametri migliori per *batch-size* e *epochs* (rispettivamente 3 e 20), ottenendo una accuratezza media di circa 98%.

Di seguito è presente un grafico, ricavato direttamente dal programma, sull'andamento dell'errore del modello sull'insieme di validazione e su quello di train:



Query

Dato il monumento riconosciuto, per estrapolare le informazioni necessarie sono state utilizzate le basi di conoscenza di WikiData e DBpedia.

WikiData è un base di conoscenza online che raccoglie dati strutturati, il cui archivio è costituito principalmente da triple RDF soggetto-predicato-oggetto.

Per estrapolare le informazioni dal web semantico uso la libreria **SPARQLWrapper** che permette di effettuare query **SPARQL** specificando un apposito *endpoint* da cui ricavare le informazioni strutturate. Gli endpoint utilizzati sono quello di DBPedia e quello di WikiData.

```
def setQuery(self, query, wrapper):
    sparql = SPARQLWrapper(wrapper)
    sparql.setQuery(query)
    sparql.setReturnFormat(JSON)
    while True:
        i = 1
        try:
            results = sparql.query().convert()
            return results
        except urllib.error.HTTPError:
            time.sleep(i)
            i += 1
```

Il metodo **setQuery** prende in input l'endpoint che specifica la fonte da cui prelevare le informazioni e la query sotto forma di stringa. Attraverso il metodo **setQuery (query)** è possibile effettuare le query, mentre con **setReturnFormat(JSON)** si specifica in che formato verranno restituiti i risultati, per poi trattare l'output di conseguenza.

k-NN

Utilizzando l'algoritmo di apprendimento supervisionato KNN, nella fase del calcolo dei neighbors di una dato esempio, è possibile ritrovare luoghi vicini all'esempio, da poter visitare.

La similarità usata nel calcolo dei neighbors è data dalla distanza geografica del monumento riconosciuto dagli altri elementi del dataset.

Il dataset è stato creato precedentemente, sulla base di una query che restituisce tutti i luoghi/monumenti presi da WikiData con la proprietà "attrazione turistica" presenti in Italia. Il dataset ha come campi la locazione geografica del luogo (LAT e LON) e l'URI, in modo tale da poter derivare le altre informazioni con opportune query sull'URI.