



**Università degli Studi di Verona**

---

DIPARTIMENTO DI INFORMATICA  
Corso di Laurea in Ingegneria e scienze informatiche

PROGETTO DI TEORIE E TECNICHE DEL RICONOSCIMENTO

## **Food Recognition**

Candidato:

**Alessia Bodini**

Matricola VR451051

---

Anno Accademico 2019–2020

# Indice

<b>1</b>	<b>Motivazioni e fondamento logico</b>	<b>2</b>
<b>2</b>	<b>Stato dell'arte</b>	<b>2</b>
<b>3</b>	<b>Obbiettivi</b>	<b>2</b>
<b>4</b>	<b>Metodologia</b>	<b>3</b>
4.1	Ricerca del dataset . . . . .	3
4.2	Estrazione delle features . . . . .	3
4.3	Metodi di riconoscimento usati . . . . .	3

## 1 Motivazioni e fondamento logico

Il seguente progetto si pone lo scopo di identificare una serie di cibi facendo uso di modelli visti durante il corso di studio (KNN, SVM e NN). Tale tipo di riconoscimento può risultare molto utile per quanto riguarda la classificazione di piatti in tutto il mondo, ad esempio per viaggiatori o stranieri che vogliono avere maggiori informazioni sul piatto o per coloro che sono interessati ad conoscere i valori nutrizionali del cibo proposto, il tutto con una sola foto.

## 2 Stato dell'arte

L'applicazione maggiormente conosciuta per quanto riguarda il riconoscimento di cibi è al momento *Calorie Mama* [2]. Tale applicazione è disponibile per Apple e Android e permette non solo di riconoscere i cibi ma anche di mostrarne i valori nutrizionali e di far gestire all'utente le calorie assunte giornalmente e relativi programmi di fitness. La funzione di *istant food recognition* viene alimentata da *Food AI API* [4] basata sulle ultime innovazioni in campo di deep learning e in grado di riconoscere ad oggi 756 cibi diversi (gran parte cibi tipici di Singapore). Ogni piatto viene poi legato a specifici valori nutrizionali che l'utente utilizza per controllare le proprie diete direttamente dall'app.

## 3 Obbiettivi

Il mio progetto non si pone di superare i risultati già raggiunti dall'applicazione nè da *Food AI API*, ma di eseguire un'analisi sulle migliori tecniche di classificazione conosciute e decretare la più efficiente tra queste. In particolare il mio lavoro si è concentrato sull'analisi di tre principali metodi per la classificazione: KNN (*K-Nearest Neighbors*), SVM (*Support Vector Machine*) e reti neurali.

## 4 Metodologia

Il lavoro si è suddiviso nella ricerca di un dataset e relativa estrazione dei dati e delle features poi utilizzate e nell'implementazione di alcuni dei modelli di riconoscimento visti durante il corso. Si spiegano di seguito nei dettagli tali processi.

### 4.1 Ricerca del dataset

Il dataset scelto denominato *Food-101* [3] è disponibile sul sito `kaggle.com` e presenta un totale di 10100 fotografie di piatti e cibi diversi. In particolare, il dataset è suddiviso in 101 categorie di cibi, ognuno già etichettato, e presenta alcuni file HDF5 dai quali è possibile estrarre direttamente training e testing set con una risoluzione minore rispetto all'originale, così da velocizzare le operazioni. I dati utilizzati nelle vari modelli sono stati presi tutti dagli stessi due file, così da mantenere una certa coerenza con i risultati raggiunti:

- *food\_c101\_n10099\_r64x64x3.h5* per il training set, con 10099 immagini (almeno una per categoria) con risoluzione 64x64x3 (RGB, uint8);
- *food\_test\_c101\_n1000\_r64x64x3.h5* per il testing set, con 1000 immagini della stessa risoluzione indicata per il training set.

### 4.2 Estrazione delle features

Per l'estrazione delle features si è fatto uso di una rete neurale disponibili tra i modelli di Torchvision e già richiamata tramite il file *resnet.py* rilasciato per questo progetto. Tale modello è il ResNet-50, definito a partire dalla ricerca *Deep Residual Learning for Image Recognition* [1].

ResNet-50 è stato usato per i primi due metodi di riconoscimento usati (KNN e SVM), mentre alla rete neurale definita successivamente sono state date direttamente in pasto le immagini del dataset (nel formato specificato sopra).

### 4.3 Metodi di riconoscimento usati

I metodi di riconoscimento implementati sono i seguenti.

**KNN** Il metodo dei *K-Nearest Neighbors* è stato costruito utilizzando diversi tipi di metriche e un diverso numero di vicini ( $K$ ) considerati per l'attribuzione a una certa categoria. In particolare si è fatto uso delle seguenti metriche per il calcolo delle distanze tra le features:

- distanza euclidea;
- correlazione;
- distanza di Minkowski.

Per ognuna delle precedenti se ne è calcolata l'efficienza per  $K$  pari a 1, 3 e 7.

**SVM** Per l'implementazione della *Support Vector Machine* si sono presi in considerazione anche in questo caso di kernel diversi:

- polinomiale;
- RBF *Radial Basis Function*
- sigmoide.

Per tutti i casi si è testato il modello su 10 iterazioni totali.

NN

## Riferimenti bibliografici

- [1] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.
- [2] Azumio Inc. Calorie mama, 2017.
- [3] K Scott Mader. Food-101, 2018.
- [4] Prof. Steven HOI R&D team. Foodai.