



Università degli Studi di Verona
Dipartimento di Informatica
Corso di Laurea in Informatica

Tesi di Laurea in
Riconoscimento statistico di immagini raffiguranti bottiglie di vino

Laureanda: Alessia Bodini

Relatore: Marco Cristani

Correlatore: Pietro Lovato

Anno accademico 2018/2019
20 Novembre 2019

Scopo del progetto

Il progetto consiste nella realizzazione di un software di riconoscimento per bottiglie di vino. Si utilizzano tre dataset diversi:

- 🍷 *ground-truth* (o **gt**) dataset: immagini ufficiali dal sito di Vinitaly;
- 🍷 **raw** dataset: foto di vini recuperate tramite social;
- 🍷 **bottles** dataset: ritagli delle bottiglie dal precedente dataset.



Stato dell'arte

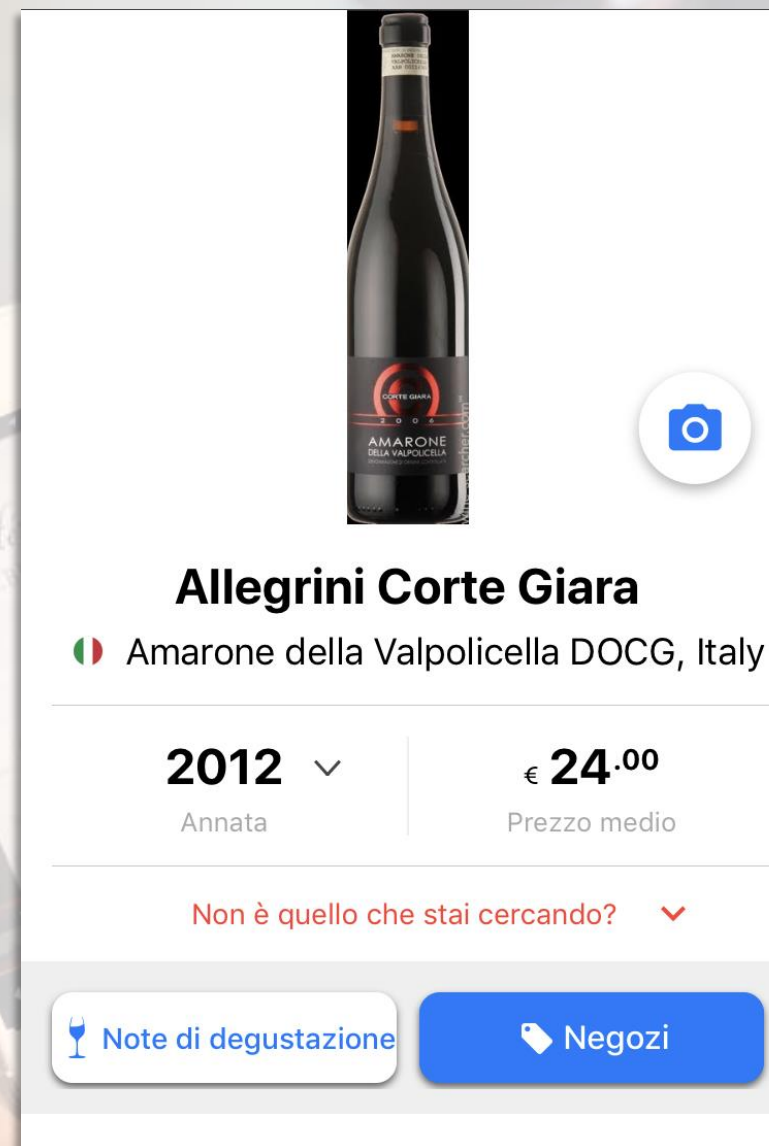
Diverse sono le applicazioni di ricerca di vini che negli ultimi anni hanno implementato un qualche tipo di riconoscimento visivo delle etichette:





Wine-Searcher

- ❖ L'etichetta deve essere centrata e messa a fuoco
- ❖ Velocità elevata ma riconoscimenti non perfetti
- ❖ Comprende riconoscimento dell'anno con relativo prezzo
- ❖ Non utilizzabile per immagini *in the wild*



The screenshot displays the Wine-Searcher app interface. At the top, there is a vertical image of a bottle of Allegrini Corte Giara Amarone della Valpolicella. To the right of the bottle is a circular camera icon. Below the bottle image, the text "Allegrini Corte Giara" is displayed in a bold, black font. Underneath this, the text "Amarone della Valpolicella DOCG, Italy" is shown with a small Italian flag icon. A horizontal line separates this from the next section, which contains two columns of information: "2012" with a dropdown arrow and the word "Annata" below it, and "€ 24.00" with the text "Prezzo medio" below it. Another horizontal line follows. Below this line is a red text prompt "Non è quello che stai cercando?" with a red dropdown arrow. At the bottom, there are two buttons: a white button with a blue wine glass icon and the text "Note di degustazione", and a blue button with a white tag icon and the text "Negozi".


Allegrini Corte Giara


🇮🇹 Amarone della Valpolicella DOCG, Italy

2012 ▼
Annata

€ **24.00**
Prezzo medio

Non è quello che stai cercando? ▼

 Note di degustazione

 Negozi



Delectable

- ❖ Riconoscimento buono anche per immagini *in the wild*
- ❖ Velocità minore rispetto al precedente

CORTE GIARA
Amarone Della Valpolicella
Corvina Blend

9,1 56 ratings **9,4** 2 pro ratings Vinous ratings

Verona in Veneto, Italy

Corvina Blend

[Check price](#)



Vivino

- ❖ Riconoscimento tramite pattern recognition e OCR
- ❖ Bassa qualità ed effetti luci/ombre possono compromettere il risultato
- ❖ Sulle bottiglie ritagliate dal dataset *raw* si ha un'accuratezza del 69.1%

Amarone della Valpolicella

CORTE GIARA
AMARONE DELLA VALPOLICELLA
DENOMINAZIONE DI ORIGINE CONTROLLATA E GARANTITA
2012
Maggini

4,0 1.371 Voti
★★★★☆

28,31 €
PREZZO MEDIO

Corte Giara [Modifica](#)

Amarone della Valpolicella 2012

 **Vino rosso da Amarone della Valpolicella, Italia**

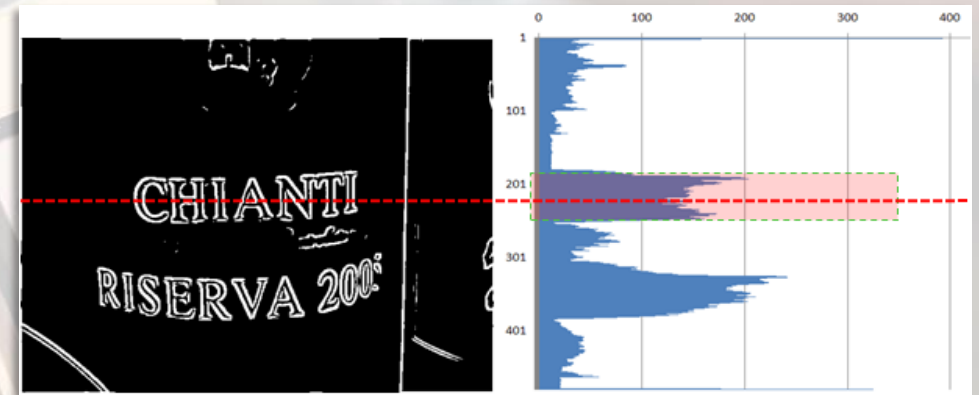
Abbinamento sbagliato? Premi [Modifica](#) per correggere

Tecniche di riconoscimento

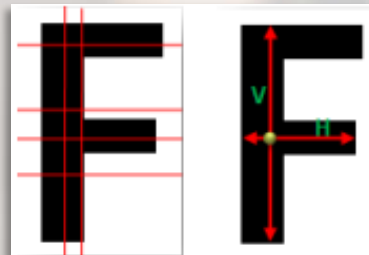
Diverse sono le tecniche di riconoscimento già sviluppate, sia nel campo delle bevande sia in quello farmaceutico. Quella ad aver riportato i migliori risultati nell'applicazione a bottiglie di vino è stata realizzata da alcuni studenti della Chonnan National University. La lettura di un'etichetta avviene in due passaggi fondamentali:

❖ **text detection**

- ❖ ricerca della regione di testo tramite *canny edge detector*
- ❖ rilevazione dei singoli caratteri
- ❖ unione dei caratteri nella sequenza corretta



- ❖ **text recognition** ricercando in ogni carattere le differenze con gli altri presenti (approccio statistico)



L'accuratezza ha raggiunto il 97.5%.

Cognitive Services di Microsoft

Microsoft mette a disposizione dell'utente una serie di **Cognitive Services** tra i quali due **API** che possono essere utilizzate in combinazione per ottenere testo scritto a mano o stampato da un'immagine:

- 🍷 Recognize Text: ritorna l'indirizzo al quale reperire i risultati;
- 🍷 Get Recognize Text Operation Results: ritorna un file .json con i risultati ottenuti dal riconoscimento.

Riconoscimento ottenuto da una foto di Amarone di Corte Giara ←

```
"recognitionResult": {  
  "lines": [  
    {  
      "boundingBox": [  
        21,  
        348,  
        114,  
        347,  
        115,  
        360,  
        21,  
        361  
      ],  
      "text": "CORTE GIARA",  
      "words": [  
        {  
          "boundingBox": [  
            22,  
            348,  
            66,  
            348,  
            66,  
            361,  
            22,  
            361  
          ],  
          "text": "CORTE"  
        },  
        ...  
      ]  
    }  
  ]  
}
```


Cognitive Services di Microsoft → Esempio di riconoscimento



Cognitive Services di Microsoft

Dall'analisi dei risultati ottenuti per l'intero dataset si possono definire alcuni punti necessari per permettere un riconoscimento soddisfacente:

- ❖ Font di semplice lettura con scrittura in stampato minuscolo o (meglio) maiuscolo
- ❖ Poche distorsioni/rotazioni
- ❖ Ambiente chiari e luminosi
- ❖ Buona qualità della foto
- ❖ Ambienti con ulteriori richiami alla marca della bottiglia
- ❖ Mancata presenza di altre bottiglie all'interno



Classificazione dei prodotti

Per la classificazione di una foto X si procede con i seguenti passaggi:

- ❖ Riconoscimento (tramite API) del testo dalla foto X e da tutte le immagini nel dataset gt
- ❖ Comparazione dei risultati ottenuti per la foto X con quelli dalle immagini in gt .
 - ❖ Per ogni immagine *ground-truth* confrontata si esaminano le distanze tra una parola riconosciuta in X e ognuna di quelle ritrovate nell'immagine gt tramite **distanza di Damerau-Levenshtein** (DL), tenendo solo il minor punteggio raggiunto per ogni parola in X . Ad esempio:

CORTE → CORTESE (2)
→ CANTINA (5)

- ❖ La somma ottenuta dall'insieme delle distanze calcolata per ogni singola parola rappresenta un valore di somiglianza con quel vino
- ❖ L'immagine in gt che ha ottenuto il punteggio minimo rappresenta il vino X nella foto indicata in partenza

Classificazione dei prodotti

Il risultato ottenuto per la foto di una bottiglia di Falesia, ad esempio, è il seguente:

```
Enter the name of an image in 'images_winebottles\test': Falesia.jpg
```

```
Sorting bottles for similarities with the chosen one:
```

```
['Falesia' 'Primitivo' 'LaSala' 'Jorche' 'Valiano' 'TelosWhite' 'Rinaldi'  
'Pinkfluid' 'Colvendra' 'CorteGiaraRipasso' 'CorteGiaraAmarone'  
'CorteMajoli' 'CorteMajoliAmarone' 'TelosRed' 'Antinori' 'Caparzo'  
'PassoDelleMule' 'Fongaro' 'Ferrari' 'Lvnae' 'PraSoave' 'EtnaRosso'  
'LeFontiASanGiorgio' 'Lungarotti' 'MarcheseAntinori' 'Damusa' 'Revi'  
'Gulfi' 'CadelBosco']
```

```
The name of the bottle is: Falesia
```



Validazione del sistema

Per studiare l'accuratezza raggiunta dal sistema descritto si è calcolato il numero di classificazioni corrette avvenute all'interno dei dataset *raw* e *bottles*. Il processo completo consiste in:

- ❧ Importazione dei nomi delle immagini da analizzare dai dataset *raw* e *gt*
- ❧ Riconoscimento (tramite API) delle immagini dei dataset indicati
- ❧ Per ogni foto in *raw*:
 - ❧ Dal file .json vengono estratte le parole riconosciute e inserite in un array, poi passato alla funzione *getScores* che si occupa di calcolare i punteggi (come già visto)
 - ❧ A partire dagli score ottenuti si delinea un nuovo array *scores* contenente i nomi delle bottiglie (gli stessi delle immagini in *gt*) in ordine di somiglianza
 - ❧ Se il nome della corretta bottiglia di riferimento si trova in posizione *x* dell'array *scores* si registra un aumento del numero di immagini riconosciute in posizione *x* (in un nuovo array *ranks*)

Lo stesso procedimento può essere utilizzato per l'analisi sul dataset bottles.

Validazione del sistema



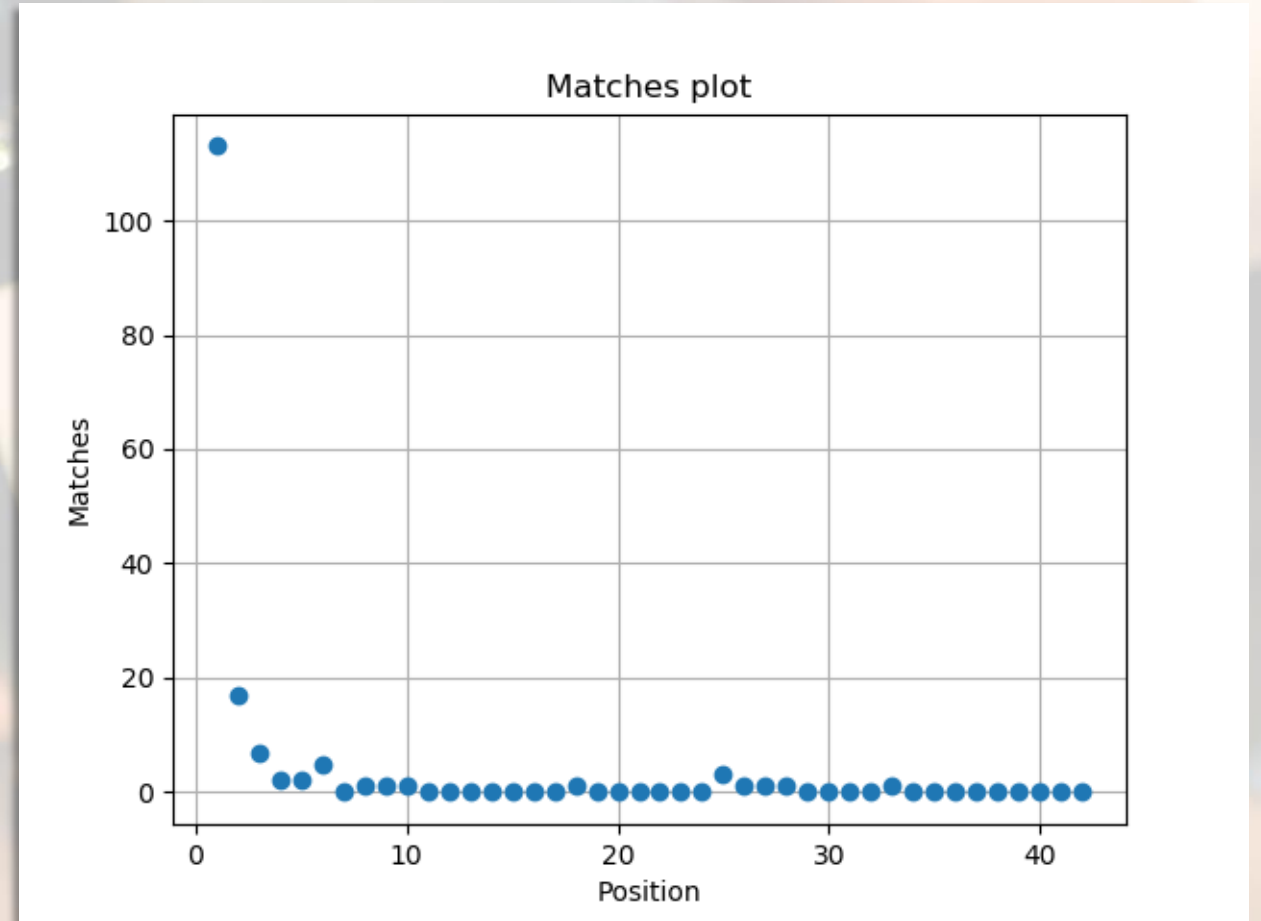
riconosciuto in posizione 3

Jorche	LaSala	Caparzo	CorteGiaraAmarone	...
1	2	3	4	5

Validazione del sistema

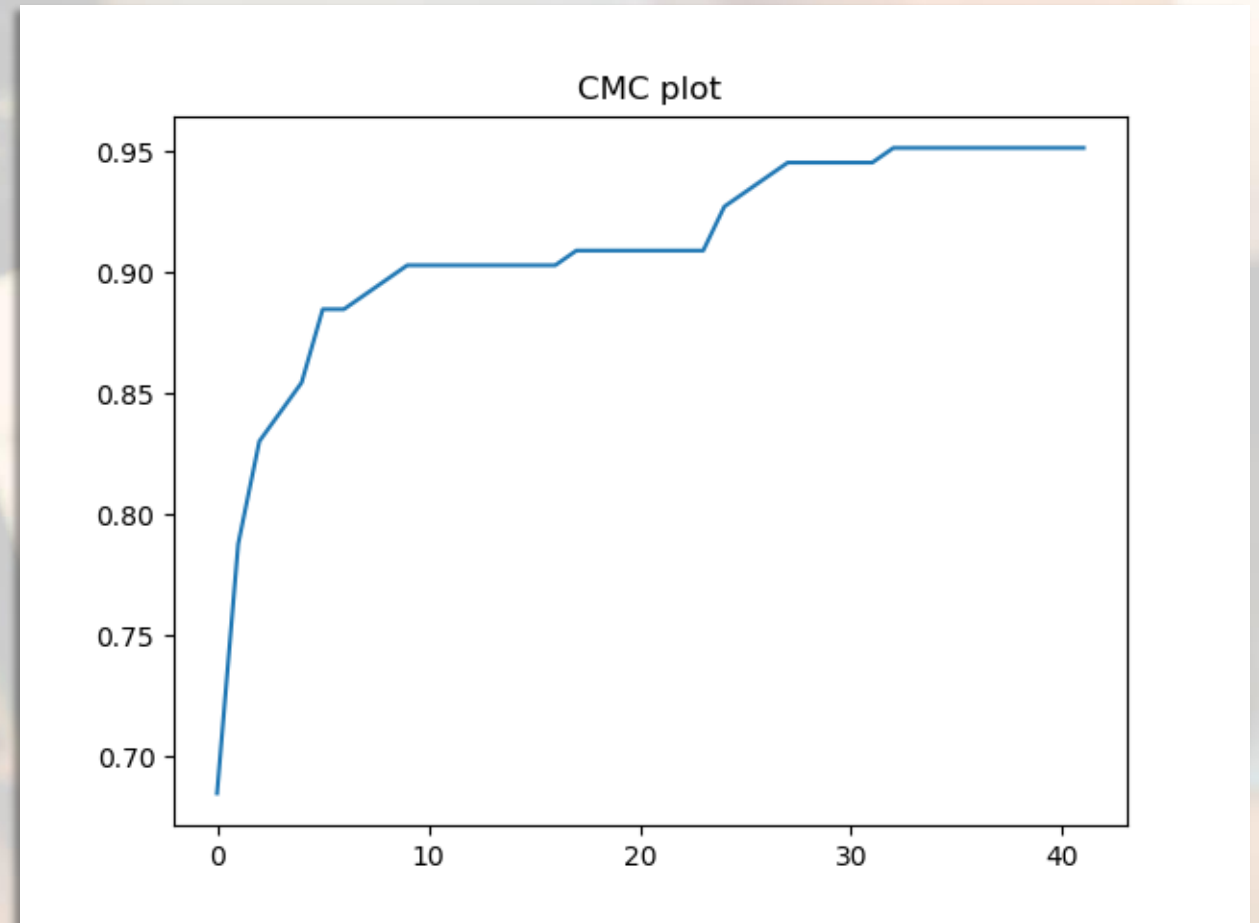
Si ottiene così il **grafico dei match**, dove:

- ☞ x = posizione in cui viene trovato il match corretto
- ☞ y = match avvenuti in quella posizione



Validazione del sistema

Dai match ottenuti si ricava il grafico relativo alla **CMC** (*Cumulative Match Curve*), ottenuto dalla somma cumulativa dei valori nell'array *ranks*, precedentemente divisi per il numero di bottiglie in *gt*.



Validazione del sistema

Ricavando l'area presente sotto le rispettive CMC si ottengono i seguenti valori di accuratezza:


	<i>Bottles</i>	<i>Raw</i>
Campioni	90	165
AUC	0.9604	0.9204

Sviluppi futuri

- ❖ Creazione di documenti contenenti informazioni di base della bottiglia trovata da stampare al momento della sua identificazione
- ❖ Aumento del dataset *gt* in modo da poter classificare più tipi di vino possibili
- ❖ Passaggio a un software di riconoscimento che non comporti l'acquisto di licenze mantenendo un simile livello di accuratezza
- ❖ Ampliamento del programma che comprenda un'interfaccia utente adeguata

Bibliografia

- 🍷 R. Hemming, «Best wine label scanning apps 2016», *JancisRobinson.com*, 04 2016.
- 🍷 J. Lim, S. Kim, J. Park, G. Lee, H. Yang, and C. Lee, «Recognition of text in wine label images», pp. 1-5, Nov 2009.
- 🍷 A. M. Arboleda and C. Arce-Lopera, «Quantitative analysis of product categorization in soft drinks using bottle silhouettes», *Food Quality and Preference*, vol. 45, pp. 1-10, 2015.
- 🍷 A. C. Tao Wang, David J. Wu and A. Y. Ng., «End-to-end text recognition with convolutional neural networks», 2012.
- 🍷 G. Fairchild, «pyxdameraulevenshtein».

A person is holding a white smartphone in their right hand. The screen of the phone displays a digital representation of a wine label. The label features the text 'Cahors' in a large, serif font, with '2008' below it. Above 'Cahors', there is a smaller, cursive-style text that appears to be 'Cahors Heritage'. The background of the image is a blurred scene of a wine tasting event, with several hands holding wine glasses filled with red wine. A bottle of wine is also visible in the background.

Grazie per l'attenzione!