

Laurea Magistrale in Informatica

Servizi cognitivi e analisi visiva



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



Marco Romanelli

marco.romanelli.1@studenti.unipd.it

Corso di Intelligenza Artificiale

Università di Padova

22 giugno 2017

Servizi cognitivi

- Creare applicazioni in grado di analizzare e interpretare la realtà
- A livello tecnico:
 - chiamate REST (endpoint http)
 - a cui passo dati (es. immagini) e parametri
- Aree principali:
 - Visione
 - Linguaggio
 - Ricerca
 - conoscenza



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Servizi cognitivi - 2

- Microsoft Cognitive Services (Microsoft Corporation)
 - Computer Vision API
 - Content Moderator API
 - Face (Emotion) API



Servizi cognitivi - 2

- Microsoft Cognitive Services (Microsoft Corporation)
 - Computer Vision API
 - Content Moderator API
 - Face (Emotion) API
- Watson Services (Bluemix) (IBM)
 - Visual Recognition



Servizi cognitivi - 2

- Microsoft Cognitive Services (Microsoft Corporation)
 - Computer Vision API
 - Content Moderator API
 - Face (Emotion) API
- Watson Services (Bluemix) (IBM)
 - Visual Recognition
- Amazon Artificial Intelligence (Amazon.com, Inc)
 - Amazon Rekognition



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Servizi cognitivi - 2

- Microsoft Cognitive Services (Microsoft Corporation)
 - Computer Vision API
 - Content Moderator API
 - Face (Emotion) API
- Watson Services (Bluemix) (IBM)
 - Visual Recognition
- Amazon Artificial Intelligence (Amazon.com, Inc.)
 - Amazon Rekognition
- Google Cloud Machine Learning Services (Google Inc.)
 - Cloud Vision API



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Microsoft Computer Vision API

- Riconoscimento elementi dell'immagine
- Classificazione
- Riconoscimento volti
- Riconoscimento del testo
- Generazione di descrizioni
- Riconoscimento contenuti non adatti ai minori
- Altro:
 - Creazione anteprime
 - Identificazione tipo, colori e qualità immagine
 - Estensione classificatore



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Microsoft Content Moderator API

- Tre tipi di operazioni:
 - Rilevare la presenza di pornografici e osé in generale
 - Rilevare la presenza di volti di persone
 - Contenuti personalizzati
 - Confronto fra immagini
 - Identificazione contenuto: alcool, nudità, armi, violenza, volgarità, eccetera



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Microsoft Face API e Emotion API

- Rilevamento volti
- Confronto fra due volti
- Confronto/ricerca in un insieme di volti
- Emozioni: rabbia, paura, felicità, espressione neutra, tristezza, sorpresa, disprezzo e disgusto



Microsoft Computer Vision API - Esempio



NOME DELLA FUNZIONALITÀ:	VALORE
Descrizione	{ "Tags": ["skating", "person", "man", "outdoor", "riding", "sport", "skateboard", "young", "board", "air", "shirt", "black", "boy", "jumping", "side", "park", "trick", "ramp", "doing", "hill"], "Captions": [{ "Text": "a young man riding a skateboard", "Confidence": 0.8099058 }] }
Tag	[{ "Name": "skating", "Confidence": 0.9999515 }, { "Name": "person", "Confidence": 0.9605775 }, { "Name": "roller skating", "Confidence": 0.945730746 }, { "Name": "man", "Confidence": 0.918820739 }, { "Name": "outdoor", "Confidence": 0.9107821 }, { "Name": "riding", "Confidence": 0.9000071 }, { "Name": "sport", "Confidence": 0.8697403 }, { "Name": "skateboard", "Confidence": 0.742671 }, { "Name": "skateboarding", "Confidence": 0.7426709 }, { "Name": "male", "Confidence": 0.151113316 }] }







IBM Visual Recognition

- Riconoscimento elementi (classificazione in categorie)
- Riconoscimento volti
- Gestione classificatore:
 - Creazione
 - Aggiornamento
 - Ricerca



IBM Visual Recognition - Esempio



Food	Score
chocolate chip pancakes	1.00 
sweet pancakes	1.00 
sweet cooked dough	1.00 
sweet treat	1.00 
baked Alaska	0.50 
dessert	0.50 

Amazon Rekognition

- Rilevamento scene e oggetti
- Operazioni su volti di persone:
 - analisi
 - confronto
 - ricerca (collezione personalizzata)



Amazon Rekognition - Esempio



▼ Results

Skateboard	99.2%
Sport	99.2%
People	99.2%
Person	99.2%
Human	99.2%
Parking	97.4%

Google Cloud Vision API

- Rilevamento:
 - oggetti e scene
 - volti
 - testo
 - luoghi d'interesse
 - loghi
 - contenuti non adatti ai minori
- Entità Web



Google Cloud Vision API - Esempio



People__Entertainment_and_recreation__037655_.jpg

Sunset	83%
Sunrise	81%
Light	80%
Morning	78%
Sunlight	73%
Evening	70%
Sun	68%
Dawn	61%

Servizi a confronto - Riconoscimento oggetti



Servizi a confronto - Riconoscimento oggetti

	Microsoft C.S	IBM W.S.	Amazon A. I.	Google C.M.L.S.
Etichette	plane (97,41%) indoor (96,20%) floor (96,10%) airplane (91,19%) airport (91,11%) aircraft (72,66%) transport (67,56%)	hangar (97,9%) blue color (85,9%) steel blue color (75,5%)	Hangar (95,74%) Aircraft (89,96%) Airplane (89,96%) Warplane (66,30%) Jet (57,09%) Landing (52,80%)	Airliner (96%) Airline (95%) Airplane (95%) Vehicle (91%) Air travel (90%) Aircraft (87%) Aviation (85%)
Descrizione	a large airplane at an airport	-	-	-

Tabella: Tabella riassuntiva per il riconoscimento oggetti e ambientazione.



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Servizi a confronto - Riconoscimento volti

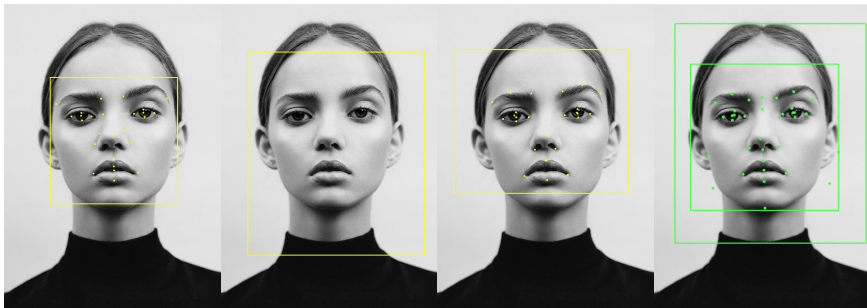


Figura: Da sinistra: Microsoft, IBM, Amazon, Google.

Tariffe - Microsoft Computer Vision API

- Piano gratuito [chiamate all'API mensili]:
 - 5000
 - max 20 chiamate al minuto
- Piani a pagamento:
 - prezzo in base al metodo utilizzato e alla quantità
 - Es. Tagging: 1\$ ogni 1000 chiamate/mese (fino a 1M)



Tariffe - IBM Visual Recognition

- Piano gratuito:
 - 250 immagini al giorno
 - un classificatore personalizzato con 5000 immagini
 - della durata di 30 giorni.
- Piano standard:
 - 25000 chiamate al giorno
 - classificazione: [singola chiamata all'API]
 - standard: 0,002\$
 - con classificatore personalizzato: 0,004\$
 - volti: 0,004\$ a chiamata
 - addestramento classificatore: 0,10\$ a immagine



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Tariffe - Amazon Rekognition

- Piano gratuito:
 - 5000 chiamate (al mese)
 - 1000 metadati facciali (al mese)
 - durata di 12 mesi
- Piano a pagamento:
 - prezzo in base al numero di chiamate mensili
 - fino a 1M: 1\$ (ogni 1000 immagini)
 - oltre i 100M: 0,40\$ (ogni 1000 immagini)



Tariffe - Google Cloud Vision API

- Piano gratuito:
 - 1000 chiamate al mese
- Piano a pagamento:
 - varia in base al metodo utilizzato e al numero di chiamate
 - Es. Label Detection: 1,50\$ ogni 1000 chiamate fino a 1M
 - Es. Web Detection: 3,50\$ ogni 1000 chiamate fino a 1M



Tariffe - Confronti

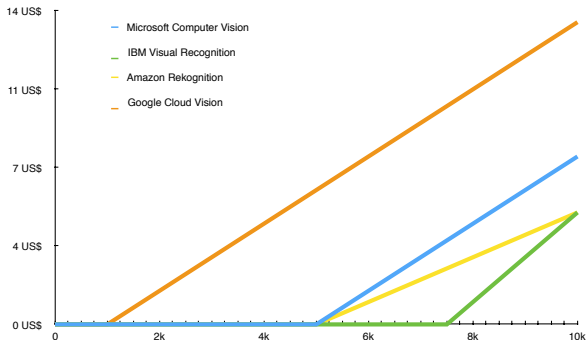


Figura: Riconoscimento oggetti con piano gratuito (da 0 a 10K immagini).

Tariffe - Confronti (2)

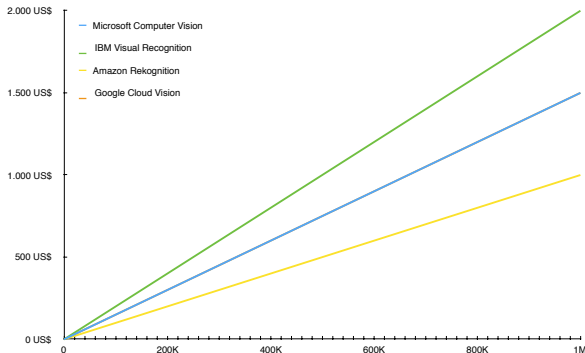


Figura: Riconoscimento oggetti senza piano gratuito (da 0 a 1M immagini).



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Tariffe - Confronti (3)

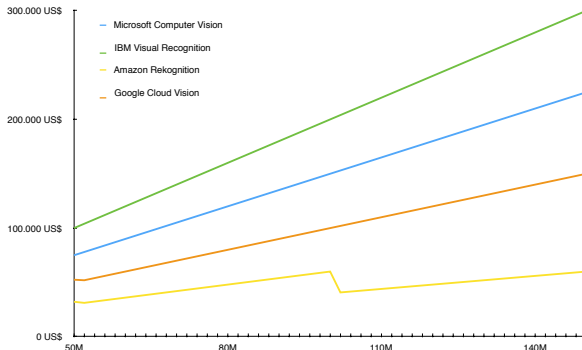


Figura: Riconoscimento oggetti (da 50M a 150M di immagini).

Conclusioni

- Le funzioni principali sono offerte da tutte le piattaforme
- Solo alcuni permettono la gestione di un classificatore
- Spesso le funzioni di ricerca/confronto sono limitate ai volti
- I risultati in risposta sono molto buoni
- Presenza di piani gratuiti



Grazie per l'attenzione



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

