

Laurea Magistrale in Informatica

Servizi cognitivi e analisi visiva



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

DIPARTIMENTO
MATEMATICA

Marco Romanelli

marco.romanelli.1@studenti.unipd.it

Corso di Intelligenza Artificiale

Università di Padova

6 luglio 2017

Servizi cognitivi

- Creare applicazioni in grado di analizzare e interpretare la realtà
- Aree principali:
 - Visione
 - Linguaggio
 - Ricerca
 - Conoscenza



Servizi cognitivi - 2

- Microsoft Corporation:
 - Microsoft Cognitive Services
- IBM:
 - Watson Services (Bluemix)
- Amazon.com, Inc:
 - Amazon Rekognition
- Google Inc.:
 - Google Cloud Machine Learning Services



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Microsoft Cognitive Services



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



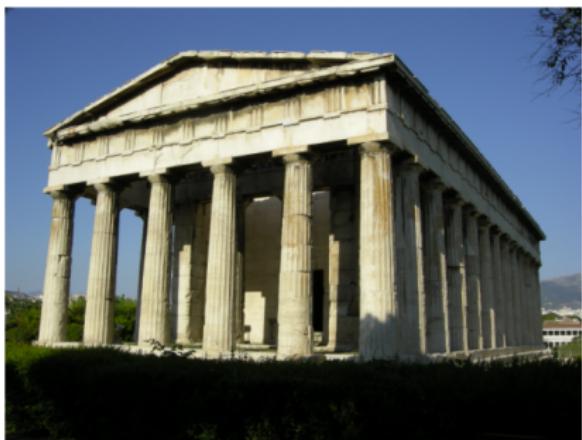
DIPARTIMENTO
MATEMATICA

Microsoft 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": 0.7426711 }, { "Name": "skateboarding", "Confidence": 0.7426709 }, { "Name": "male", "Confidence": 0.151113316 }] |

Microsoft Computer Vision API - Esempio (2)



```
"categories": [
  {
    "name": "building_pillar",
    "score": 0.98828125,
    "detail": {
      "landmarks": [
        {
          "name": "Temple of Hephaestus",
          "confidence": 0.998190939
        }
      ]
    }
]
```



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



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 Face API - Esempio



| | |
|-----------|-----------|
| GENDER | male |
| AGE | 58.6 |
| GLASSES | NoGlasses |
| SMILE | 0.052 |
| ANGER | 0.002 |
| CONTEMPT | 0 |
| DISGUST | 0 |
| FEAR | 0 |
| HAPPINESS | 0.052 |
| NEUTRAL | 0.938 |
| SADNESS | 0.008 |
| SURPRISE | 0.001 |

IBM Watson Services (Bluemix)



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

IBM Visual Recognition

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



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

IBM Visual Recognition - Esempio



| Classes | Score |
|---------------------|-------|
| artilleryman | 0.68 |
| serviceman | 0.72 |
| person | 0.72 |
| weapon | 0.57 |
| foxhole | 0.55 |
| fortification | 0.55 |
| defensive structure | 0.55 |

Type Hierarchy

/person/serviceman/artilleryman

/defensive structure/fortification/foxhole

Amazon Artificial Intelligence



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Amazon Rekognition

- Rilevamento scene e oggetti
- Moderazione immagini (contenuto esplicito)
- 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% |



Amazon Rekognition - Esempio (2)



▼ Results

Suggestive 83.5%

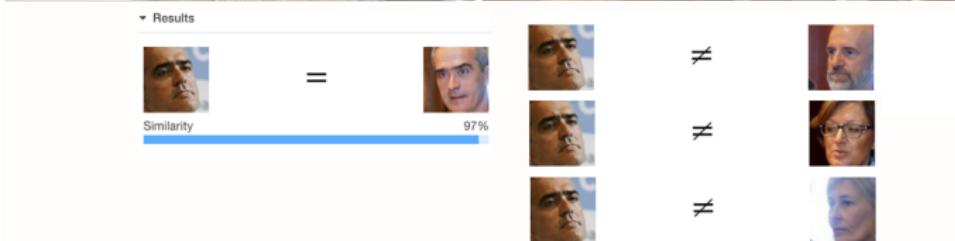
Female Swimwear Or Underwear 83.5%

► Request

► Response



Amazon Rekognition - Esempio (3)



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Google Cloud Machine Learning Services



Google Cloud Platform



Google Cloud Vision API

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

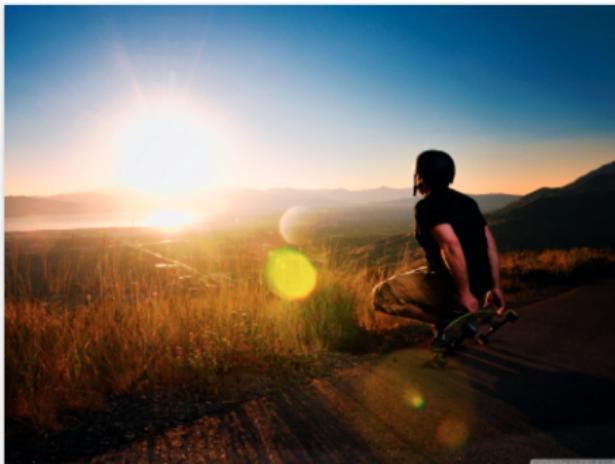


UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

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% |

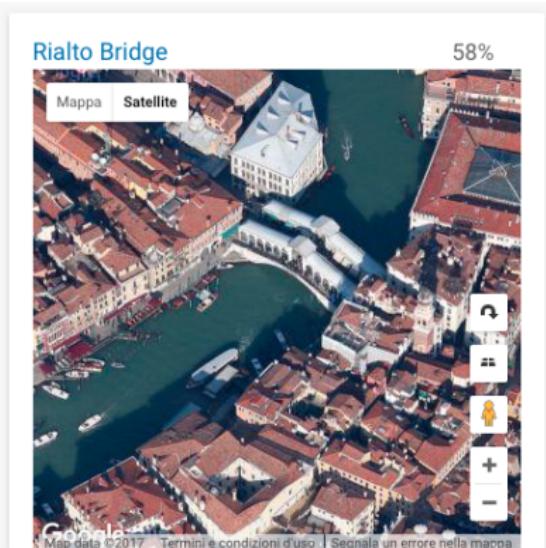
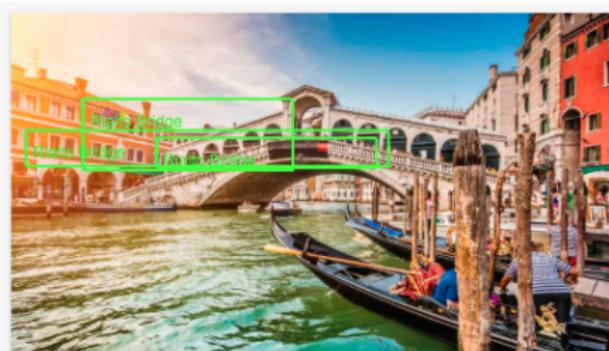


UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Google Cloud Vision API - Esempio (2)



Google Cloud Vision API - Esempio (3)

Web Entities

| | |
|--------------------------|---------|
| Red Bull Rampage | 1.48892 |
| Red Bull | 0.78122 |
| Mountain bike | 0.73475 |
| Freeride | 0.66016 |
| Bicycle | 0.65046 |
| Downhill mountain biking | 0.43498 |
| Cycling | 0.34226 |
| Photograph | 0.32909 |
| Extreme sport | 0.31287 |
| Image | 0.30529 |
| BMX | 0.30446 |
| Red Bull Rampage | 0.27605 |
| Mountain biking | 0.26627 |



stock-photo-142869191-1-1500x1000.jpg

| | |
|-----------------------|-----|
| Sky | 94% |
| Soil | 92% |
| Mountainous Landforms | 89% |
| Rock | 83% |
| Extreme Sport | 83% |
| Atmosphere Of Earth | 79% |
| Mountain | 77% |
| Terrain | 77% |
| Cycling | 75% |



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.



Servizi a confronto - Riconoscimento volti

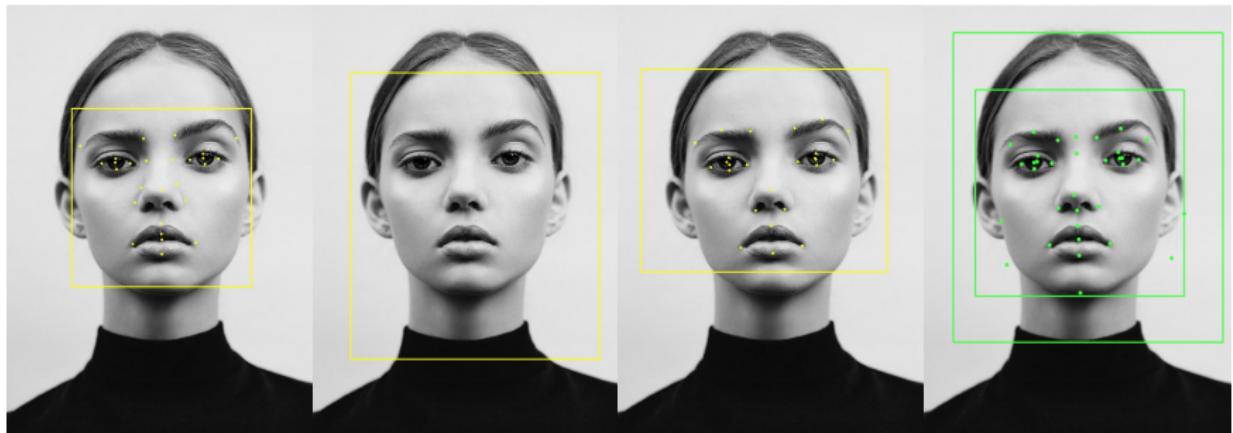


Figura: Da sinistra: Microsoft, IBM, Amazon, Google.



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Utilizzo API

- Requisiti: essere registrati all'API (chiave).
- Chiamata (REST) a un endpoint http.
- Risultati in JSON.

```
curl -v -X POST "https://westus.api.cognitive.microsoft.com/vision/v1.0/
    analyze?visualFeatures=Tags&details={Celebrities}&language=en"
-H "Content-Type: application/json"
-H "Ocp-Apim-Subscription-Key: {subscription key}"
--data-ascii "{body}"
```

Figura: Vision API con Curl.



Utilizzo API (2)

```
import http.client, urllib.request, urllib.parse, urllib.error, base64

headers = {
    # Request headers
    'Content-Type': 'application/json',
    'Ocp-Apim-Subscription-Key': '{subscription key}',
}

params = urllib.parse.urlencode({
    # Request parameters
    'visualFeatures': 'Tag',
    'details': 'Celebrities',
    'language': 'en',
})

try:
    conn = http.client.HTTPSConnection('westus.api.cognitive.microsoft.com')
    conn.request("POST", "/vision/v1.0/analyze?%s" % params, "{body}", headers)
    response = conn.getresponse()
    data = response.read()
    print(data)
    conn.close()
except Exception as e:
    print("[Errno {0}] {1}".format(e errno, e strerror))
```



Tariffe

- Unità di misura: 1000 chiamate all'API al mese.
- Costo in base:
 - alla quantità di chiamate mensili,
 - alla funzionalità richiesta.
- Tutte le piattaforme offrono piani gratuiti.



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Tariffe - Piani gratuiti

- Quantità di chiamate all'API:
 - 5000 al mese (Microsoft, Amazon)
 - 1000 al mese (Google)
 - 250 al giorno (IBM)
- Altri limiti:
 - 20 al minuto (Microsoft)
 - Durata massima di un mese (IBM)
 - Durata massima di 12 mesi (Amazon)



Tariffe - Piani a pagamento

- Variano in base alla funzionalità richiesta.
- Esempi [costo ogni 1000 chiamate all'API]:
 - Vision API - riconoscimento oggetti: 1\$
 - Vision API - descrizione: 1,5\$
 - Visual Recognition: 2\$
 - Reckognition: 1\$
 - Cloud Vision - riconoscimento oggetti: 1,5\$
 - Cloud Vision - riconoscimento Web: 3,5\$



Tariffe - Confronti

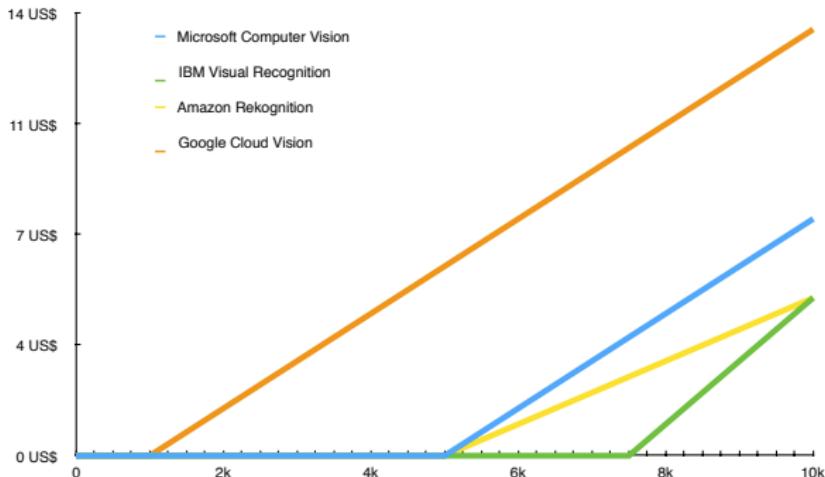


Figura: Riconoscimento oggetti con piano gratuito (da 0 a 10K immagini).



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Tariffe - Confronti (2)

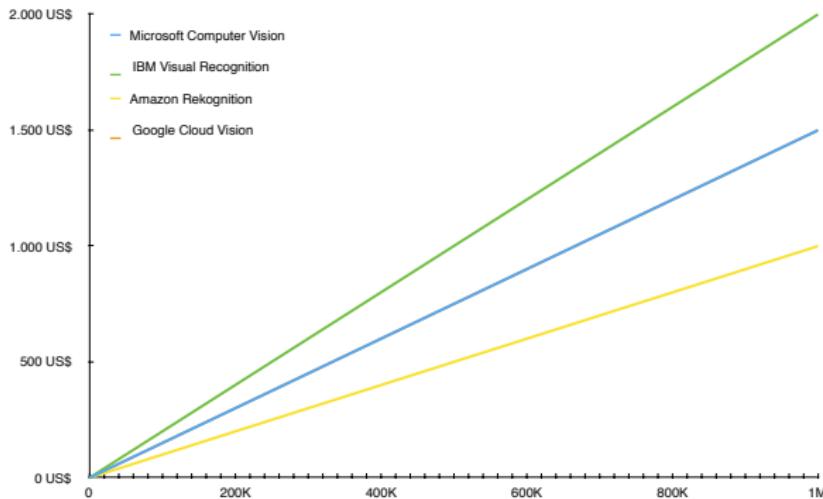


Figura: Riconoscimento oggetti senza piano gratuito (da 0 a 1M immagini).



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Tariffe - Confronti (3)

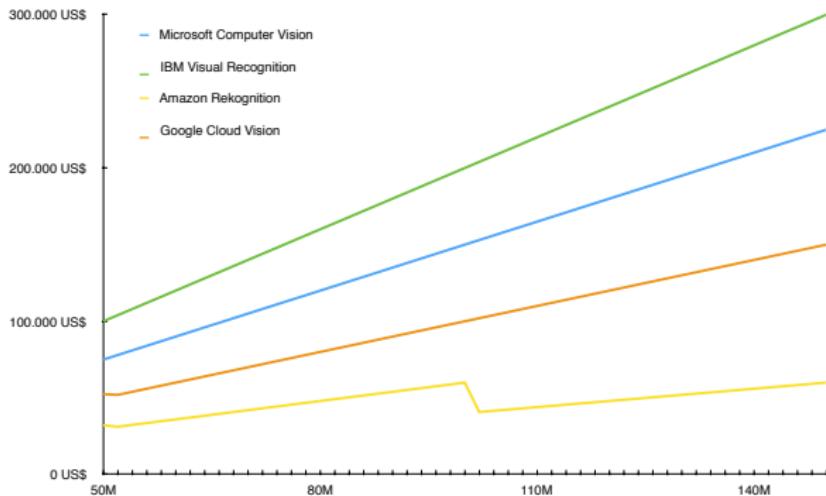


Figura: Riconoscimento oggetti (da 50M a 150M di immagini).



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Conclusioni

- PRO:
 - molte funzionalità
 - semplicità di utilizzo
 - presenza di piani gratuiti
- CONTRO:
 - funzioni di ricerca spesso limitate ai volti
 - impossibile fornire un feedback al risultato



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



DIPARTIMENTO
MATEMATICA

Grazie per l'attenzione



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



IBM Bluemix



Google Cloud Platform