Research WORK

Ai based Information retrieval, indexing and neural search using Elastic Search Backend

Note

- This is a pure results of our ongoing research work and not a Commercial product.

OCR
Model 1


```
{
  "billboard1.jpg": {
    "text": [
      "refreshing",
      "gb",
      "refreshed",
      "bud",
      "light",
      "the",
      "budlight",
      "party",
      "media",
      "street",
      "o"
    ]
}
```

OCR
Teseract OCR

BUD
PARTY
nal

OCR
Model 2


```
{
  "text": [
    "REFRESHING",
    "ID",
    "REFRESHED",
    "BUD",
    "IGHT",
    "THE",
    "BUDLIGHI",
    "naalGs",
    "PARTY",
    "STREETe|[Dli"
  ]
}
```

Scene Reco – Model 1

{ "Environment_type": "outdoor",

"scene_recog": {
"labels": [
"gas_station",
"booth/indoor",
"motel",
"industrial_area",
"desert_road"], },

"scene_atrributes": ["natural light", "open area", "man-made", "sunny", "clouds", "far-away horizon", "natural", "dry", "no horizon"
]
}

Scene Recog - Keras – Places – 365 : hybrid


```
{
  "labels": [
    [
      "street",
      "sign"
    ],
    [
      "scoreboard"
    ],
    [
      "bakery",
      "bakeshop",
      "bakehouse"
    ],
```

Scene Recog - Keras-places 365 : base

- {
- "labels": [
- "/h/hot_spring",
- "/o/ocean",
- "/h/hospital",
- "/i/ice_shelf",
- "/i/ice_floe"
-],

Image Recog – largest Open Source Image Data set - **Tencent – Ml Images**

```
{
  "labels": [
    "street",
    "vending",
    "scoreboard",
    "bobsled",
    "trailer“
]
}
```

Keras-Models

- **Resnet-50**

```
"labels": [
    "traffic_light",
    "street_sign",
    "pole"
  ],
```

Resnet-152

```
 "labels": [
    "scoreboard",
    "pole",
    "trailer_truck"
  ],
```

Resnet-101v2

```
"labels": [
    "birdhouse",
    "water_bottle",
    "street_sign"
  ],
```

Resnet-152v2

```
 "labels": [
    "traffic_light",
    "switch",
    "vending_machine"
  ],
```

vgg16

```
"labels": [
  "scoreboard",
  "traffic_light",
  "street_sign"
],
```

xception

```
"labels": [
  "street_sign",
  "mailbox",
  "pole"
],
```

inceptionv3

```
"labels": [
  "traffic_light",
  "trailer_truck",
  "scoreboard"
],
```

nasnet_large

```
"labels": [
  "scoreboard",
  "barbershop",
  "street_sign"
],
```

mobilenet

```
"labels": [
  "trailer_truck",
  "scoreboard",
  "street_sign"
],
```

mobilenet_large

```
"labels": [
  "scoreboard",
  "street_sign",
  "mailbox"
],
```

Image Captioning – image_captioning (show and tell model)

- {
- "caption": "a sign on a pole in front of a building.",
- "prob": "8.202121480135207e-05"
- }

Dense Cap Model

"a sign is on the pole",
"a man holding a phone",
"a building with a blue and white sign",
"green and white street sign",
"tall building with many windows",
"mountains in the distance",
"a pole in the background",
"a mountain in the distance",
"the sign is blue",
"the letter is white",
"a red and white sign",
"a blue and white sign",
"the word <UNK> on the sign",
"a mountain in the distance",
"a blue and white sign",
"a blue sky with white clouds",
"white clouds in blue sky",
"the sign is red and white",
"the word <UNK> on the sign",

"the word <UNK> on the sign",
"a yellow line of trees",
"a blue and white sign",
"mountains in the distance",
"the letter is white",
"the word <UNK> on the bus",
"mountains in the distance",
"wing of the plane",
"blue and white hat on the head",
"the letter is white",
"mountains in the distance",
"a mountain in the distance",
"the word <UNK> on the sign",
"wing of a plane",
"a pole with a sign",
"white clouds in blue sky",
"white clouds in blue sky",
"yellow flowers in the grass",
"a metal railing",
"mountains in the distance",
"white clouds in blue sky",
"mountains in the distance",
"the letter is white",

"the letter is white",
"a small hill in the distance",
"a red and white sign",
"the shirt is blue",
"a hill in the distance",
"a small green bush",
"a small tree in the background"
"white clouds in blue sky",
"a small white sign",
"a sign on a pole",
"white paper on the wall",
"a small patch of green grass",
"a small tree in the background"
"white letters on a sign",
"the word <UNK> on the sign",
"trees in the distance",
"a small green bush",
"a large white building",
"white clouds in blue sky",
"a sign on a metal pole",
"white clouds in blue sky",
"a metal railing",
"the wing of a plane",

"a pole in the background",
"a mountain in the distance",
"a tall building with a red roof",
"a blue and white sign",
"a clear blue sky",
"a black and white sign",
"the sign is blue",
"the sign is black",
"a pole in the background",
"the sign is blue",
"white clouds in blue sky",
"the wing of a plane",
"a black and white plane",
"a patch of grass",
"white clouds in blue sky",
"wing of a plane",
"white clouds in blue sky",
"a metal pole",
"white clouds in blue sky",
"white clouds in blue sky",
"a tall building",
"white clouds in blue sky",
"white clouds in blue sky",

"white clouds in blue sky",
"a window on a building",
"white clouds in blue sky",
"the wing of a plane",
"black and white sign on building",
"white sign with black letters",
"the wing of the plane",
"a sign on the pole",
"a white sign on a building",
"white clouds in blue sky",
"wing of a plane",
"white wall in the background",
"wing of the plane",
"white paper on wall",
"a metal pole",
"white clouds in blue sky",
"white clouds in blue sky"

Image Caption Pytorch

[
"a sign that is on a metal pole",
"a sign that is on the side of a street",
"a sign that is on top of a pole",
"a sign that is on top of a metal pole",
"a sign that is on the side of a pole"
]

Named Entity Recognition

- The following output is by taking above image captioning input.
- {'PERSON': ['Charles Angles'], 'NORP': ['American'], 'ORG': ['Udacity', 'Recode'], 'DATE': ['8229', 'earlier this week', 'less than a decade later'], 'CARDINAL': ['dozens', '9966']}

- text =

When Charles Angles started working on self-
driving cars at Google in 2007, few people outside of the company took him seriously. “I can t
ell you very senior CEOs of major American car companies would shake my hand and turn aw
ay because I wasn’t worth talking to,” said Thrun, now the co-
founder and CEO of online higher education startup Udacity, in an interview with Recode earli
er this week. He has worked in dozens of projects like this in less than a decade and have 9966
customers from them.

Face Recognition

Single Image Trained

Trained on

Recognized Successfully

Celebrity Recognition

Single Image Trained

Trained on

Recognized Successfully

Audio Finger Printing =>(shazam search)

Source Audio

Audio Result of search

{
'total_time': 0.2674314975738525 4,
'fingerprint_time': 0.1720659732818603 5',
'query_time': 0.0852406024932861 3,
'align_time': 0.0084052085876464 84,
'results': [
{
'song_id': 3,
'song_name': b'Aladin-New-World-JFla',
'input_total_hashes': 5966,
'fingerprinted_hashes_in_db': 73345,
'hashes_matched_in_input': 1857,
'input_confidence': 0.31,
'fingerprinted_confidence': 0.03,
'offset': 1571,
'offset_seconds': 72.9571,
'file_sha1':
b'2873A6B127854CE747746958B054BD5ACAA
B94DA'
},
]
}

Audio Classification

Source Audio

Result of Classification


```
{
  "status": "ok",
  "predictions": [
    {
      "label_id": "/m/04rlf",
      "label": "Music",
      "probability": 0.7622102499008179
    },
    {
      "label_id": "/m/05r5c",
      "label": "Piano",
      "probability": 0.10926598310470581
    },
    {
      "label_id": "/t/dd00034",
      "label": "Tender music",
      "probability": 0.10484001040458679
    },
    {
      "label_id": "/m/05148p4",
      "label": "Keyboard (musical)",
      "probability": 0.10307112336158752
    },
    {
      "label_id": "/m/01s0ps",
      "label": "Electric piano",
      "probability": 0.08495655655860901
    }
  ]
}
```

Object Detection (yolo v4)

cell phone: 38%

Object Detection (Retina Net)

Traffic light : 37.69156038761139 [713, 16, 908, 442]

---

bottle : 41.481840610504 15 : [713, 105, 905, 437]

---

person : 78.07514071464539 : [703, 100, 927, 450]

Our State of the art model
model1- Open images trained on 1 million images


```
{
  "objects": [
    "Billboard",
    "Drink",
    "Person",
    "Tin can"
  ],
```

Our model trained on COCO dataset


```
{
  "objects": [
    "person",
    "cell phone"
  ],
  "score": [
    0.9828922210693,
    0.2182917590966
  ]
}
```

Our model trained on VOC dataset


```
{
  "objects": [
    "bottle"
  ],
  "score": [
    0.921953558921814
  ]
}
```

Our YOLO algorithm model


```
{
  "objects": [
    "person",
    "cell phone"
  ],
  "score": [
    0.1007056354866,
    0.3237611420593
  ]
}
```

State of the art fastest object detection model 2 millisecond recognition time


```
{
  "objects": [
    "bottle",
    "person"
  ],
```

Image 2 image search

Browse... No file selected.

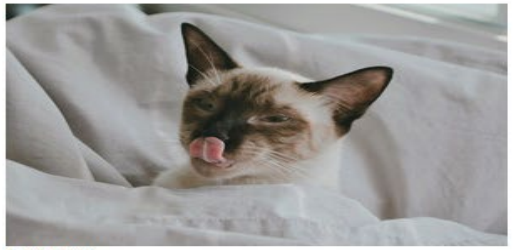
Submit Query

Query:

Results:


0.868823


0.965242


1.08808


1.11124


1.15912


1.16995


1.18847

1.19261


1.19476

Sentence Embedding and Cluster Search

- **DataSet or demo Corpus**
- corpus = ['A man is eating food.',
- 'A man is eating a piece of bread.',
- 'The girl is carrying a baby.',
- 'A man is riding a horse.',
- 'A woman is playing violin.',
- 'Two men pushed carts through the woods.',
- 'A man is riding a white horse on an enclosed ground.',
- 'A monkey is playing drums.',
- 'A cheetah is running behind its prey.'
-]

Sentence Embedding and Cluster Search

- **Query: Someone in a gorilla costume is playing a set of drums.**

  Result of search:

- A monkey is playing drums. (Score: 0.7985)
- A cheetah is running behind its prey. (Score: 0.2860)
- The girl is carrying a baby. (Score: 0.2351)
- A man is riding a horse. (Score: 0.2023)
- A man is riding a white horse on an enclosed ground. (Score: 0.1963)

Speech to text

Youtube video for speech to text - https://youtu.be/v9arM_agKFA

WORD Embedding

- **pd.DataFrame(vectors.k_neighbors('california'))**
- **pd.DataFrame(vectors.k_neighbors('conference'))**

cosim	word			cosim	word
0 1.000	california			1.0000	conference
0.5060	university			0.4320	international
0.4239	berkeley			0.4063	secretariat
0.4103	barbara			0.3857	jcdl
0.3941	santa			0.3798	annual
0.3899	southern			0.3708	conferences
0.3673	uc			0.3705	forum
0.3542	johns			0.3629	presentations
0.3396	indiana			0.3601	workshop
0.3388	melvy			0.3580	...