



# Capstone Project @GA

Google Landmark recognition challenge

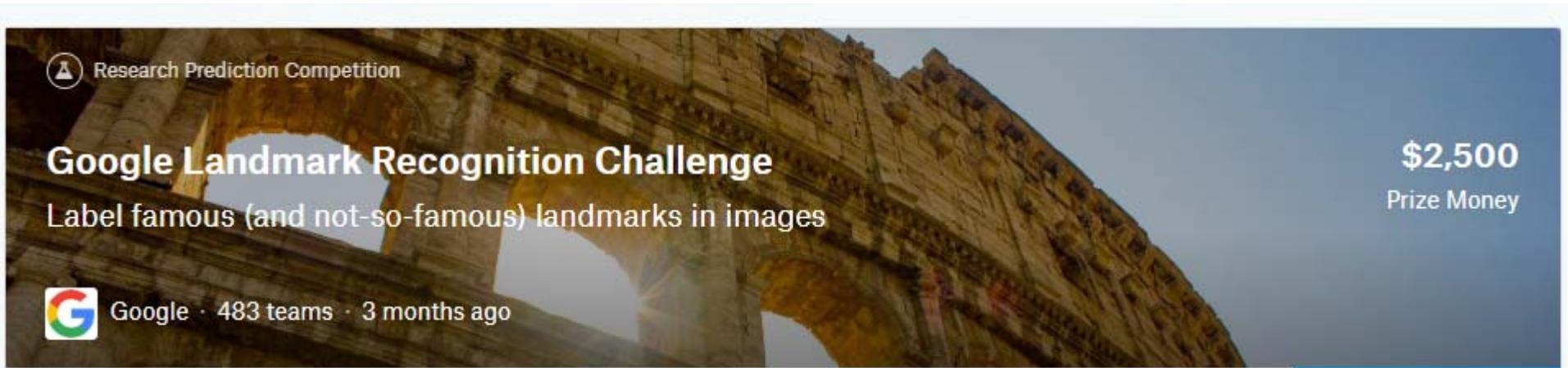
--by Wei Fu

August 2018



## Outline

- Introduction
- Exploratory data analysis
- Image classification model
- Classification result output
- Future work



A competition posted at [Kaggle.com](https://www.kaggle.com) in May 2018

**Goal:**

to predict landmark labels directly from image pixels

**Challenge:**

there are a total of 15K different famous or infamous landmarks(classes) involved and the number of training examples per class may not be very large



## Data source

Link to the datasets: <https://www.kaggle.com/google/google-landmarks-dataset>

### **Train.csv:**

- This file contains information regarding the training images to be used in the recognition task associated to the Google-Landmarks dataset. These images may also be useful to train models for the retrieval task. The first column contains the train image ID, the second column contains its URL, and the third column contains its label.

### **Test.csv:**

- This file contains information regarding the test images to be used in the recognition and retrieval tasks associated to the Google-Landmarks dataset. The first column contains the test image ID, and the second column contains its URL.



## Success Metrics

Correctly identify in which class the each test image is and the corresponding confidence level and the output format is:

[{cl1 for class1},{cl2 for class2},...,{clk for classk}], k=1,2,..10



## Project value

It will be applied to

- landmark recognition and classification
- landmark collection reorganization

It can be transferred to

- facial recognition
- Handwriting recognition and verification
- Conventional publications digitalization



## Train data

```
1 train = pd.read_csv('train.csv')
```

```
1 train.head()
```

	<b>id</b>	<b>url</b>	<b>landmark_id</b>
0	97c0a12e07ae8dd5	http://lh4.ggpht.com/-f8xYA5I4apw/RSziSQVaABI/...	6347
1	650c989dd3493748	https://lh5.googleusercontent.com/-PUUnMrX7oOyA...	12519
2	05e63ca9b2cde1f4	http://mw2.google.com/mw-panoramio/photos/medi...	264
3	08672eddcb2b7c93	http://lh3.ggpht.com/-9fgSxDYwhHA/SMvGEoltKTI/...	13287
4	fc49cb32ef7f1e89	http://lh6.ggpht.com/-UGAXxvPbr98/S-jGZbyMIPi/...	4018

```
1 train.shape
```

(1225029, 3)

Note: 1.2million images are supplied as train data



## Test data

```
1 test = pd.read_csv('test.csv')
```

```
1 test.head()
```

	<b>id</b>	<b>url</b>
0	cb9998b8cdaf6235	<a href="https://lh3.googleusercontent.com/-q8B91vDIQZY...">https://lh3.googleusercontent.com/-q8B91vDIQZY...</a>
1	30728cf6e50a6bc6	<a href="https://lh3.googleusercontent.com/-91gJSKTgv5Q...">https://lh3.googleusercontent.com/-91gJSKTgv5Q...</a>
2	16afbc86b710337d	<a href="https://lh3.googleusercontent.com/-GHZdXuf2wMg...">https://lh3.googleusercontent.com/-GHZdXuf2wMg...</a>
3	d29b2166cf522450	<a href="https://lh3.googleusercontent.com/-cWDnYNQhyws...">https://lh3.googleusercontent.com/-cWDnYNQhyws...</a>
4	dd5c03b20c21cfba	<a href="https://lh3.googleusercontent.com/-PSLN6BloM-k...">https://lh3.googleusercontent.com/-PSLN6BloM-k...</a>

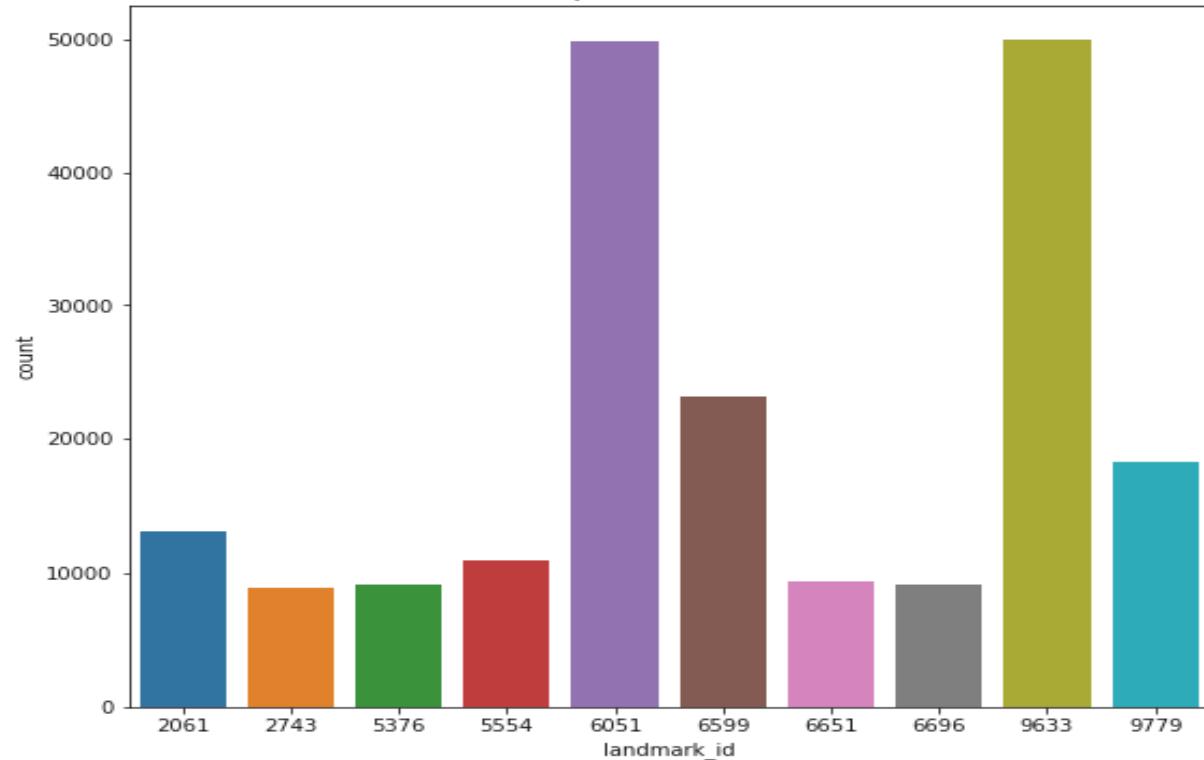
```
1 test.shape
```

(117703, 2)

Note: 117 thousands images are supplied as test data

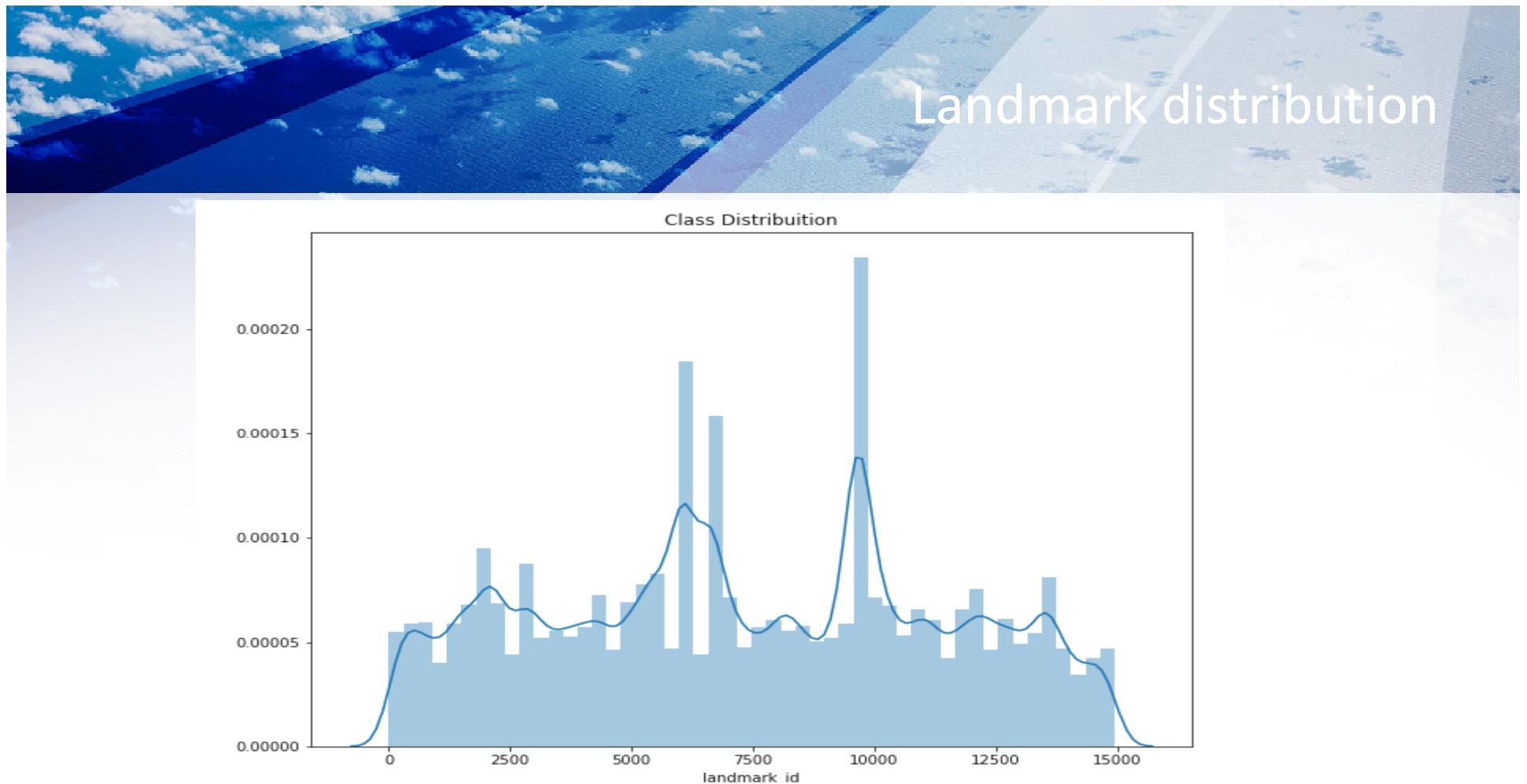
# Top 10 most popular landmarks in train data

Most frequent landmarks shown

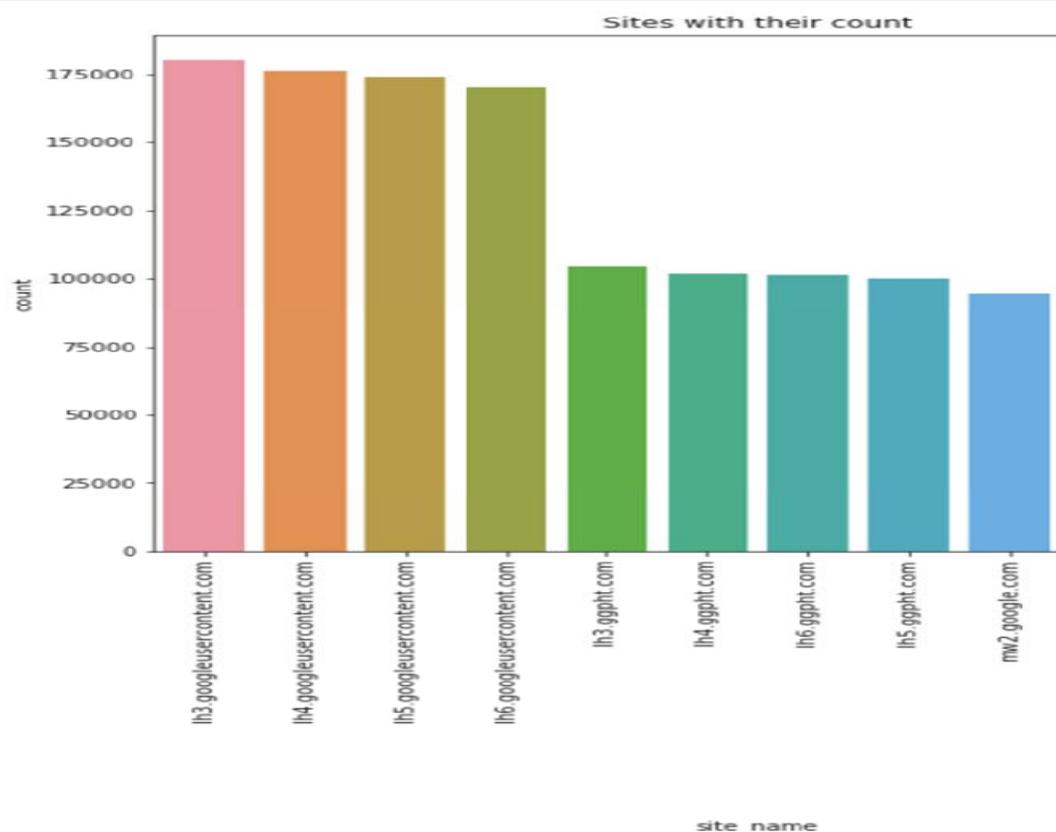


- 1- 9633
- 2- 6051
- 3- 6599
- 4- 9779
- 5- 2061
- 6- 5554
- 7- 6651
- 8- 5376
- 9- 6696
- 10- 2743

# Landmark distribution



# Image storage websites



lh3.googleusercontent.com  
lh4.googleusercontent.com  
lh5.googleusercontent.com  
lh6.googleusercontent.com  
lh3.ggpht.com  
lh4.ggpht.com  
lh6.ggpht.com  
lh5.ggpht.com  
mw2.google.com  
Static.panoramio.com  
Commondatastorage.googleapis.com  
0-focus-  
opensocial.googleusercontent.com



## The most popular landmark: 9633



St. Peter's Basilica (Latin: Basilica Sancti Petri), is an Italian Renaissance church in Vatican City, the papal enclave within the city of Rome.



## The second most popular landmark: 6051



The Colosseum or Coliseum also known as the Flavian Amphitheatre is an oval amphitheatre in the centre of the city of Rome, Italy

# The most popular landmark: 9633



000a3c7c32743ef  
1



000a6ac40648d52  
f



000b3b51ffa7273  
0



000dcb10ff3d387  
2



000ece7f564b977  
c



000edceb2acb41  
4b



000eeab943d869e  
3



00a1a8b9419c90d  
0



00a8c23bca2f9ca  
7



00a8e9e9ff219ef6  
b2



00a23d4037841b  
8



00a31eb80e44804  
4



00a688963a80c01  
0



00a62303087c74f  
9



00aa9e19c038eb6  
e



00ab208dbf885ee  
7



00af00d7437e7a6  
e



00af5e78e04ad6d  
6



00afe2ec722b101  
8



00b27bca9188819  
9



00b60d682fae6db  
c



00b97cc85a02e17  
0



00ba6eba7cbe04  
dd



00ba83ec15ea930  
7



00be1477d88c975  
4



00bf0b212f64921  
3



00c0da3f139b74c  
3



00c2e653cb11d84  
f



00c3efa79f5b6c5  
a



00c9e255a1ca57a  
9



00c9ead0f8b87ad  
7



00c10c9e99f4907  
2

# The second most popular landmark: 6051



00bffd02d9d8c  
3



00c3be48fb01fef4  
8



00c7a8ecf770bd3  
7



00c9d49e0edc223  
7



00c15cda207db0  
62



00c72ce00b43ed9  
b



00c8455f992f03b  
8



00c93821ad25ff9  
3



00d25dbb7fa5829  
8



00d3981bcbf2a26  
e



00d5127e145f7d5  
b



00d255998afded0  
7



00da79295043bf0  
e



00dceaa8c2b2e7c  
e



00dd4a737d6327  
41



00ddd1597fdbdc  
32



00e1f16ce16a20d  
1



00e4eff937d2bf50  
9



00e5ea8317ecfe4  
b4



00e6d86d8c1754  
3



00e57cebb8f1cb0  
c



00e0633aef3c43b  
c



00e6675b4452827  
c



00e9873d589a50f  
5



00ea52cc4922c99  
c



00eb7ebe9eee609  
8



00eb6215d54926e  
6



00ebd65f28fad3c  
e



00ee97c570b47f1  
9



00ef301cc32a282  
8

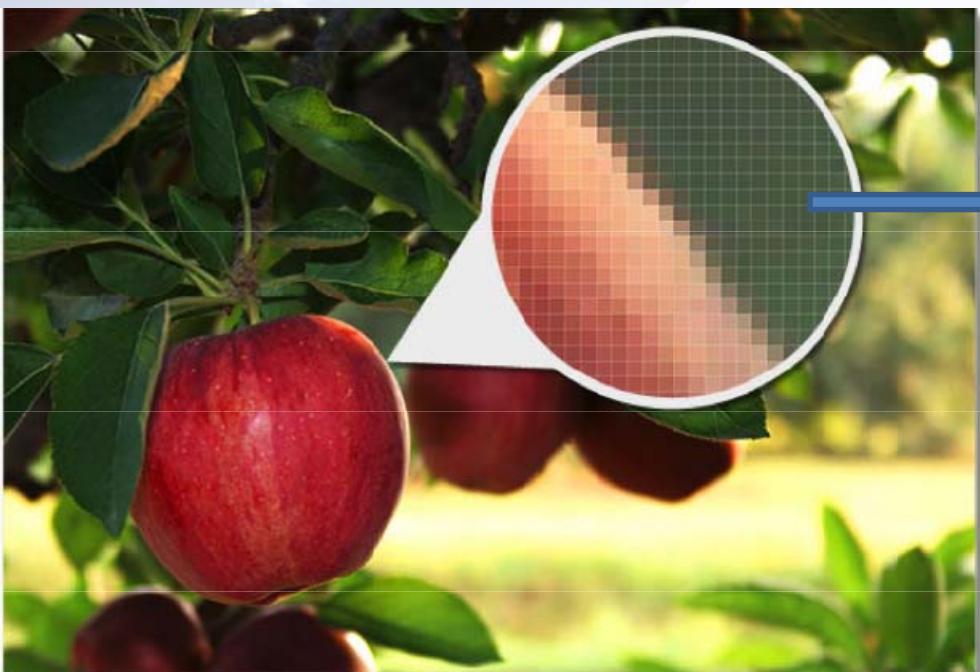


00f5b22726e8b83  
5



00f6e31580a180b  
7

# What is a image?



B	0.95	0.05	0.98	0.98
G	0.95	0.05	0.98	0.99
R	0.95	0.05	0.98	0.97
	0.60	0.40	0.37	0.99
	0.94	0.06	0.99	1.00
	0.82	0.18	0.97	0.99
	0.99	0.01	0.99	0.02
	0.92	0.08	1.00	0.06
	0.76	0.24	0.98	0.00
	0.94	0.06	0.02	0.07
	0.88	0.12	0.06	0.11
	0.62	0.38	0.00	0.07
	1.00	0.00	0.07	0.11
	0.97	0.03	0.11	0.90
	0.78	0.22	0.90	

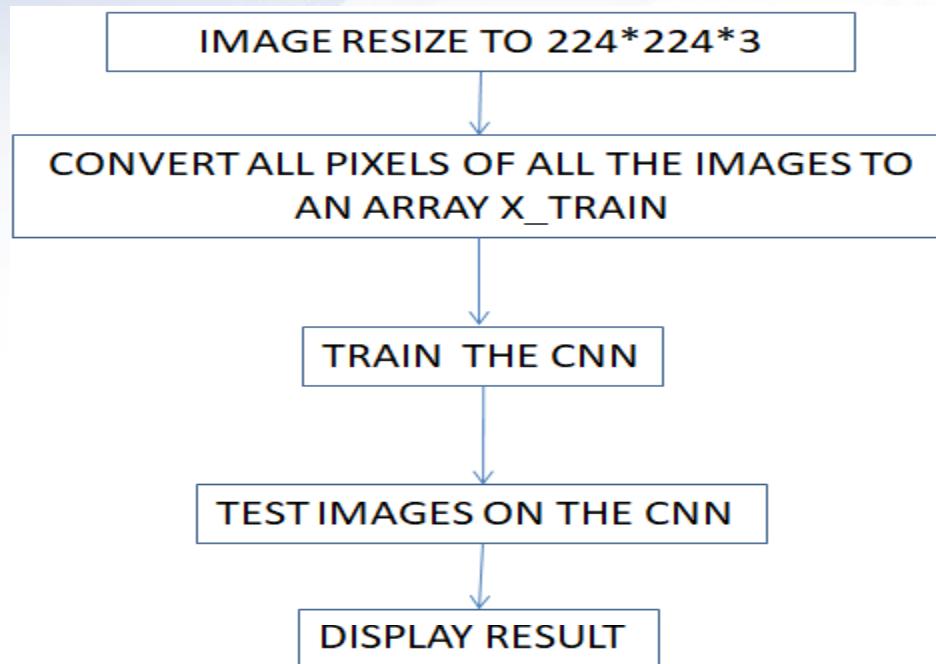


## Plans

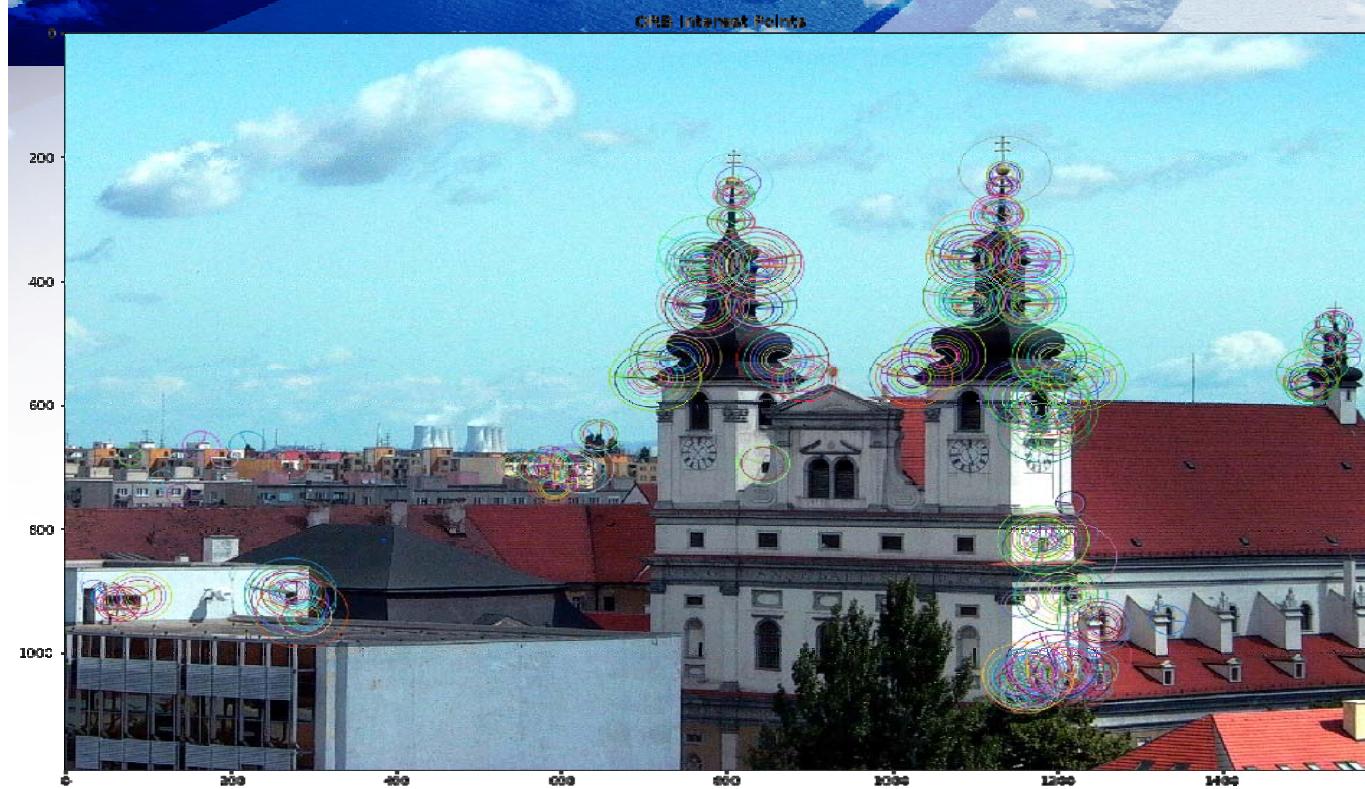
- To build or retrain a convolutional neural network (CNN) model
- To feed the model with the train data
- To predict the landmark and evaluate the outcome
- To optimize the model based on the outcome evaluations
- To report the final results



# How do we classify images?



# What features can be extracted?

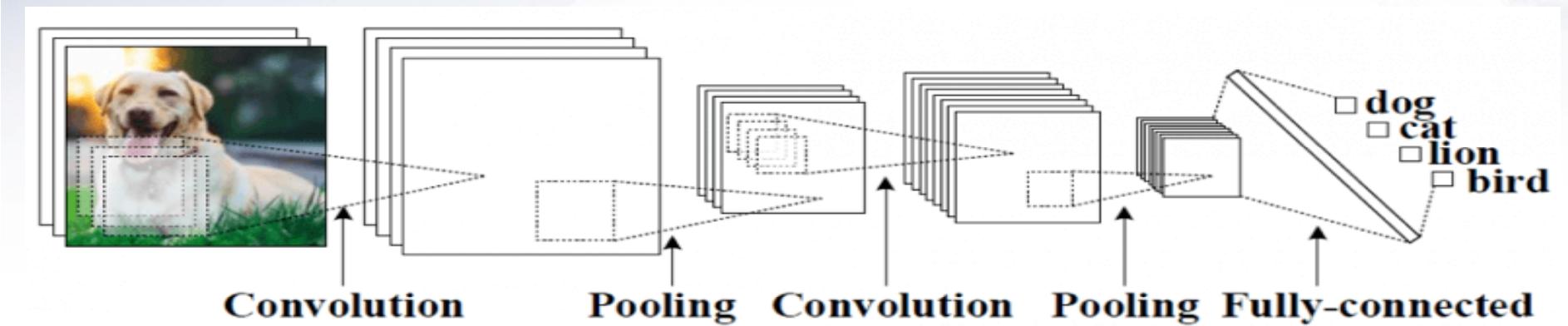


- To extract
- Corners
  - Edges
  - Tips
  - Holes

Picture source: <https://www.kaggle.com/wesamelshamy/image-feature-extraction-and-matching-for-newbies>

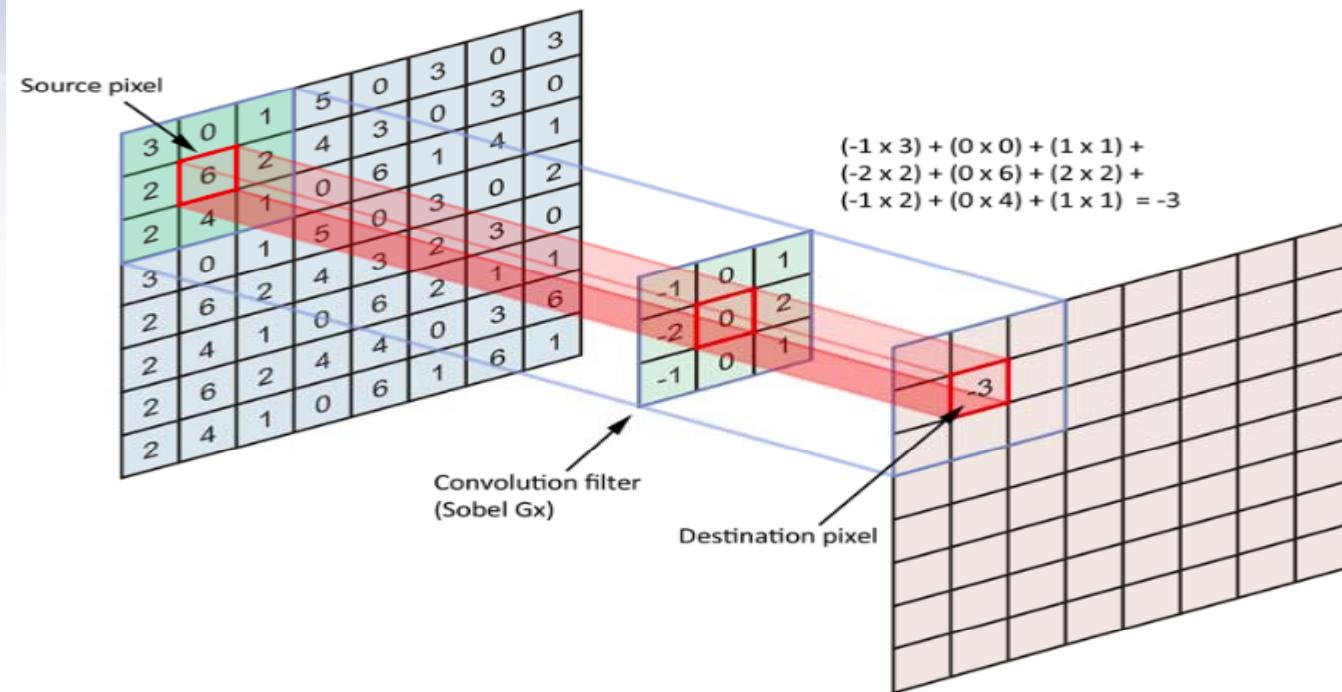


What does CNN\* look like?



\* CNN: convolutional neural network

# What does convolution do?



# Classification results on two classes



[0.96, 0.04]

(9633, 6051)



[0.81, 0.19]

(9633, 6051)

# Classifying five classes

-landmark id: 2743, 5376, 5554, 6051, 9633



2743

↑  
Pantheon  
(Rome, Italy)



Petronas Towers  
(Kuala Lumpur, Malaysia)



5554





## Classifying five classes

landmark id: 2743, 5376, 5554, 6051, 9633



[0.87, 0.01, 0.02, 0.05, 0.05]  
(2743, 5376, 5554, 6051, 9633)



[0.62, 0.03, 0.11, 0.08, 0.16]  
(2743, 5376, 5554, 6051, 9633)



## Accuracy summary

Accuracy	Training on 500 images per class	Training on 1000 images per class
2 classes	Up to 85%	Up to 96%
5 classes	Up to 65%	Up to 77%
10 classes	10%	N/A

CNN Model performance



## Future work

To improve the model's performance:

- Training the model on more images
- Training the model on more classes
- To optimize image preparation



*Thank you*

*Questions?*