Public Datasets For Facial Recognition

CTI One Corporation

Date: June 9, 2020
Advisor: Dr. Harry Li, Ph.D.

Team Member:
Rachana Bumb

https://www.kaggle.com/

https://www.kaggle.com/datasets

Example 1. Search facial recognition datasets and code sample (Python notebook using data for Facial Keypoints Detection) 2019, gpu

Karan Jakhar
Data Scientist at Reinvent Systems
Chandigarh, Punjab, India

Performance on Youtube Faces DB

training.csv coordinates of facial keypoints, e.g., left eye, rigth eye etc and the image.

test.csv   image, you have to give coordinates of facial keypoints by ID lookup table csv file

IdLookupTable.csv

```
Train_Dir = '../input/training/training.csv'
Test_Dir = '../input/test/test.csv'
lookid_dir = '../input/IdLookupTable.csv'
train_data = pd.read_csv(Train_Dir)
test_data = pd.read_csv(Test_Dir)
lookid_data = pd.read_csv(lookid_dir)
os.listdir('../input')
```

Harry Li, Ph.D. 2018, 2020

https://www.kaggle.com/

https://www.kaggle.com/datasets

Example 2. Search rain forest imaging

https://www.kaggle.com/datasets?search=satellite

Dataset for CNN Training for Facial Recognition
https://www.face-rec.org/databases/

- There are almost 97 datasets on internet, freely available to public strictly for non-commercial research purposes.
- Many of these databases have specific requirements in regards to referencing.
- Certain data bases require school identity proof, some require higher authority signature from the graduate school, rest all others are free to use for research purpose with proper citations.
- The choice of an appropriate database to be used should be made based on the task given

Harry Li, Ph.D.

Image Datasets for Facial Recognition

https://www.kaggle.com/c/deepfake-detection-challenge/discussion/121594

No.	Dataset	Description	Resolution
1.	CelebA Dataset	200,000 celebrity images.	• 1024x1024 • **10,177** number of **identities**, • **5 landmark locations**, • **40 binary attributes** annotations per image.
2.	Flickr Faces	210,000 images.	• 70,000 original images, • 70,000 images cropped at 1024 x 1024 pixels, • 70,000 cropped at 128 x 128 pixels.

Harry Li, Ph.D.

3.	**Google Facial Expression Comparison**	**156,000 facial images.**	
4.	Labeled Faces in the Wild	13,000 images of over 5,700 people	
5.	Tufts Face Database	100,000 images	
6.	UMDFaces	367,000 face of over 8,200 different subjects	
7.	Yale Face Database	165 GIF images of 15 different subjects	