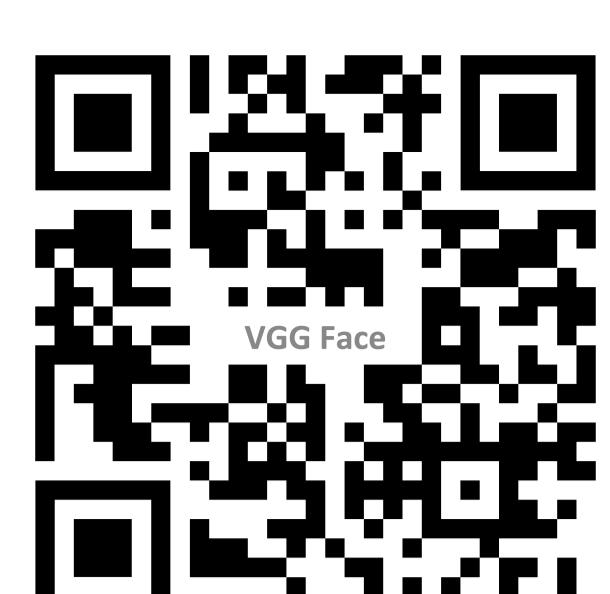




UNIVERSITY OF
OXFORD

Deep Face Recognition

Omkar M. Parkhi Andrea Vedaldi Andrew Zisserman
Visual Geometry Group, Department of Engineering Science, University of Oxford



Objective

- Build a large scale face dataset with minimal human intervention
- Train a convolutional neural network for face recognition which can compete with that of the Internet giants like Google and Facebook

Contributions

- New face dataset of 2.6 Million Faces with minimal manual labeling
- State-of-the-art results on YouTube faces in the wild dataset (EER 3%) [Better than Google's FaceNet and Facebook's DeepFace]
- Comparable to the state-of-the-art results on the Labeled faces in the wild dataset (EER 0.87%) [Comparable to Google's FaceNet and better than Facebook's DeepFace]

Dataset construction

1. Candidate list generation: 5000 identities

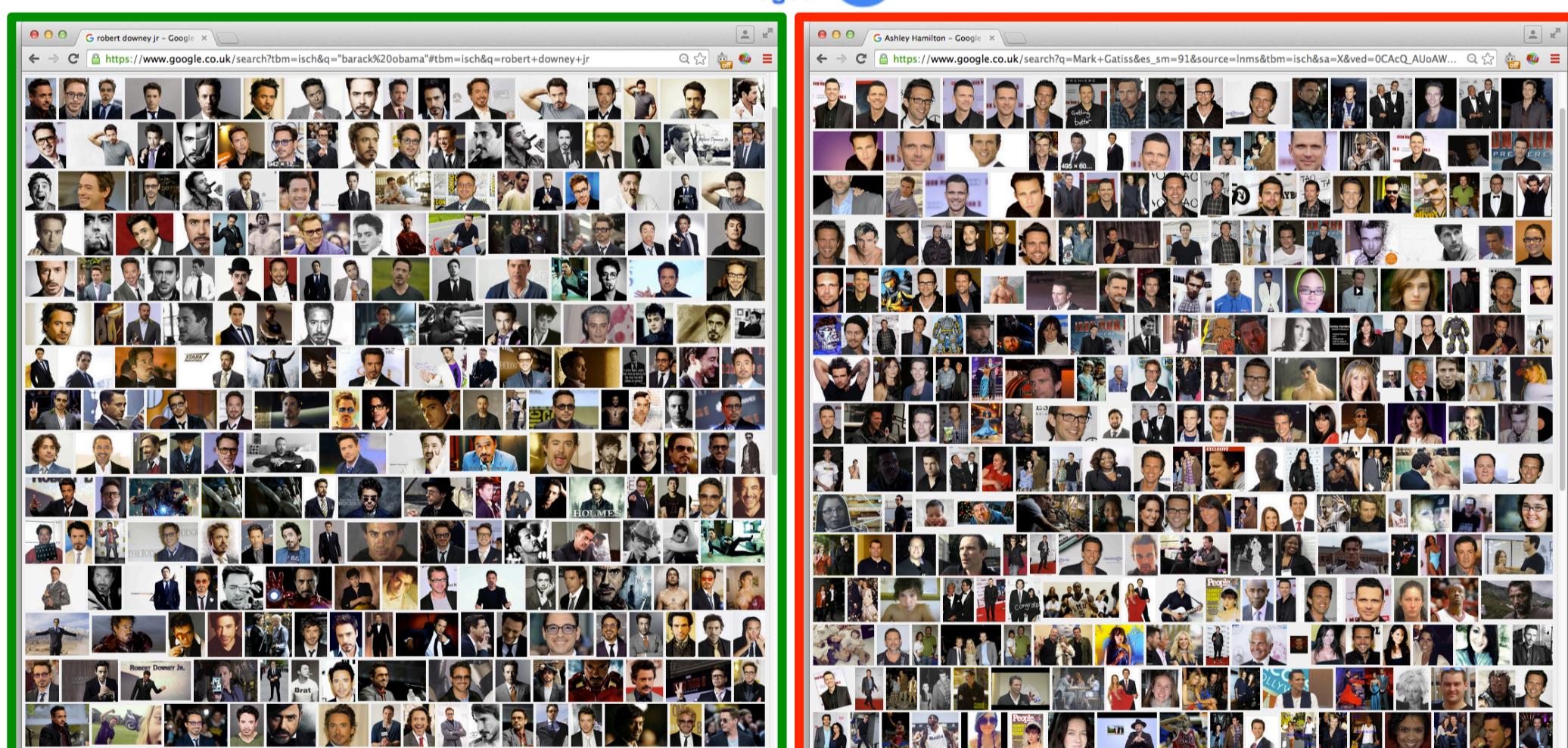
- Tap the knowledge on the web



- Collect representative images for each celebrity
 - 200 images per identity

2. Candidate list filtering

- Remove people with low representation on the Google image search.
- Remove overlap with public benchmarks
- 2622 celebrities in the final dataset



3. Rank image sets

- Download 2000 images per identity



- Learn classifier using data obtained in step 1
- Rank 2000 images and select the top 1000

4. Near duplicate removal

- VLAD descriptor based near duplicate removal

5. Manual filtering

- Curating the dataset further using manual checks

Dataset Statistics

No.	Aim	Mode	# Persons	# images / person	Total # images	Anno. effort
1	Candidate list generation	Auto	5000	200	1,000,000	-
2	Candidate list filtering	Manual	2622	-	-	4 days
3	Rank image sets	Auto	2622	1000	2,622,000	-
4	Near duplicate removal	Auto	2622	623	1,635,159	-
5	Manual filtering	Manual	2622	375	982,803	10 days

Representation Learning

- **"Very Deep" Architecture**
 - Different from previous architectures proposed for face recognition:
 - locally connected layers (Facebook)
 - inception (Google)
- **Learning a multi-way classifier**
 - Soft Max Objective
 - 2622 way classification
 - 4096d descriptor
- **Learning Task Specific Embedding**
 - The embedding is learnt by minimizing the triplet loss

$$\sum_{(a,p,n) \in T} \max\{0, \alpha - \|x_a - x_n\|_2^2 + \|x_a - x_p\|_2^2\}$$

- Learning a projection layer from 4096 to 1024 dimensions
- Online triplet formation at the beginning of each iteration
- Fine tuned on target datasets

Experiments

Labeled Faces in the Wild – Unrestricted Protocol

Effect of dataset curation

Dataset	#images	100%-EER
Curated	982,803	92.83
Full	2,622,000	95.80

Effect of embedding learning

Objective	100%-EER
Softmax classification	97.30
Embedding Learning	99.10

More experiments in the paper

Dataset

Example images from our dataset



Dataset Comparison

No.	Aim	# Persons	Total # images
1	Labeled Faces in the Wild	5,749	13,233
2	WDRef	2995	99,773
3	Celeb Faces	10177	202,599
4	Ours	2622	2.6M
5	Facebook	4030	4.4M
6	Google	8M	200M

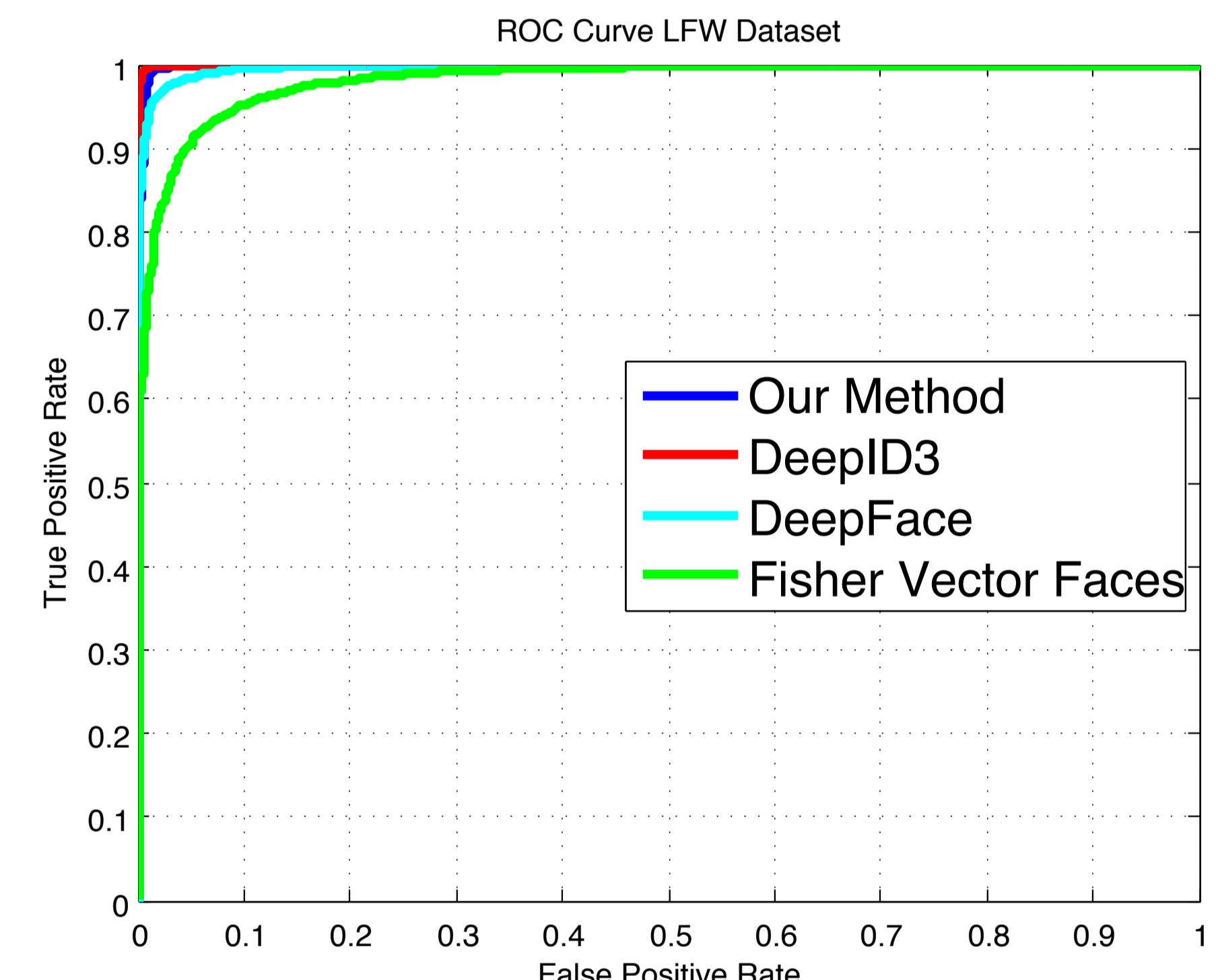
Comparison with the state of the art

YouTube Faces In the Wild: Unrestricted Protocol

No.	Method	# Training Images	# Networks	100%-EER	Accuracy
1	Video Fisher Vector Faces	-	-	87.7	83.8
2	DeepFace (Facebook)	4 M	1	91.4	91.4
4	DeepID-2,2+,3		200	-	93.2
5	FaceNet (Google)	200 M	1	-	95.1
7	Ours (VGG Face)	2.6 M	1	97.60	97.4

Labeled Faces In the Wild: Unrestricted Protocol

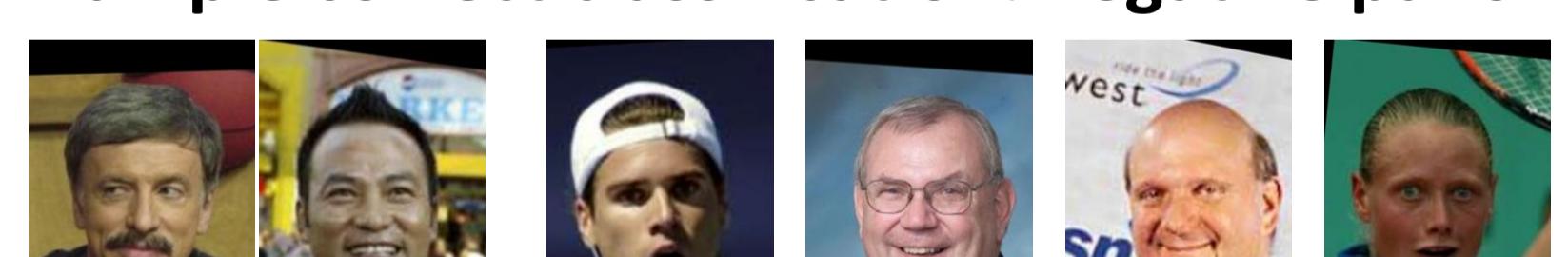
No.	Method	# Training Images	# Networks	Accuracy
1	Fisher Vector Faces	-	-	93.10
2	DeepFace (Facebook)	4 M	3	97.35
3	DeepFace Fusion (Facebook)	500 M	5	98.37
4	DeepID-2,3	Full	200	99.47
5	FaceNet (Google)	200 M	1	98.87
6	FaceNet+ Alignment (Google)	200 M	1	99.63
7	Ours (VGG Face)	2.6 M	1	98.78



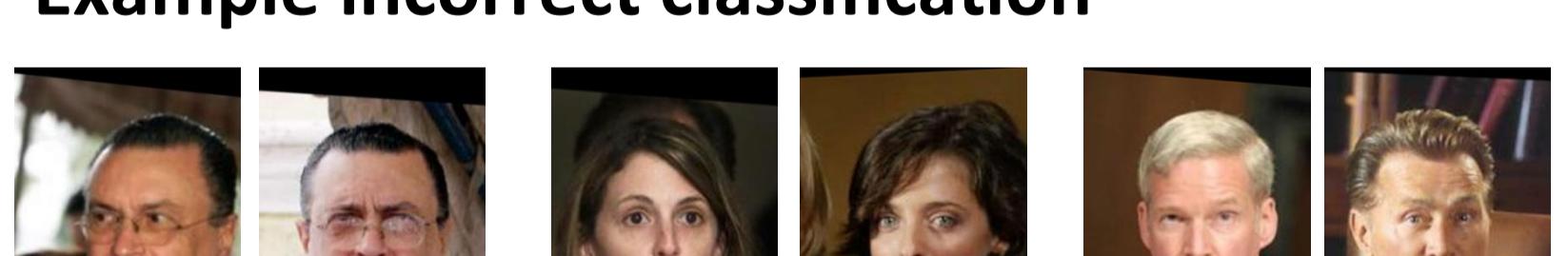
Example correct classification: positive pairs



Example correct classification: negative pairs



Example incorrect classification



False negative

False positive

Download



CNN model and face detector available online!!

Visit:
http://www.robots.ox.ac.uk/~vgg/software/vgg_face/

Acknowledgements

This research is based upon work supported by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA)