



DATA-DRIVEN VISUAL SIMILARITY FOR CROSS-DOMAIN IMAGE MATCHING

Tomasz Malisiewicz

(Massachusetts Institute of Technology)

Abhinav Shrivastava, Abhinav Gupta and Alexei A. Efros

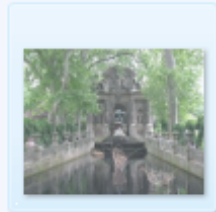
(Carnegie Mellon University)



Medici Fountain, Paris



Search by image



[→ Move](#)

Drop image here

[Watch a short video](#) to learn more.

Search

About 2 results (0.29 seconds)

Everything

Images

Maps

Videos

News

Shopping

More



Image size:
1024 × 829

No other sizes of this image found.

Visually similar





Medici Fountain, Paris (winter)



Search

About 2 results (0.29 seconds)

Everything

Images

Maps

Videos

News

Shopping

More



Image size:
713 × 600

No other sizes of this image found.

Visually similar







Search

About 2 results (0.29 seconds)

Everything

Images

Maps

Videos

News

Shopping

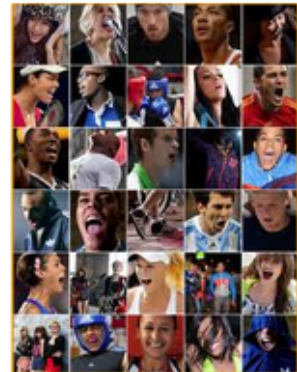
More



Image size:
319 x 482

No other sizes of this image found.

Visually similar







Search

About 2 results (0.29 seconds)

Everything

Images

Maps

Videos

News

Shopping

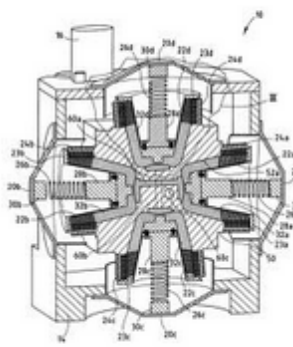
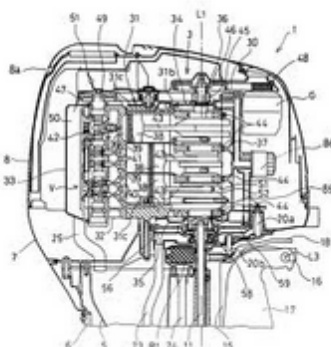
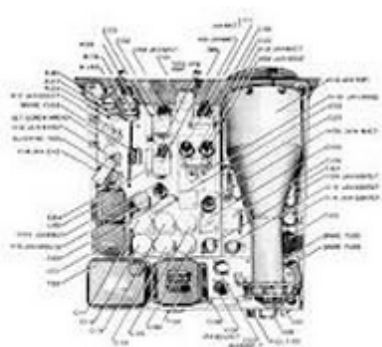
More



Image size:
443 × 482

No other sizes of this image found.

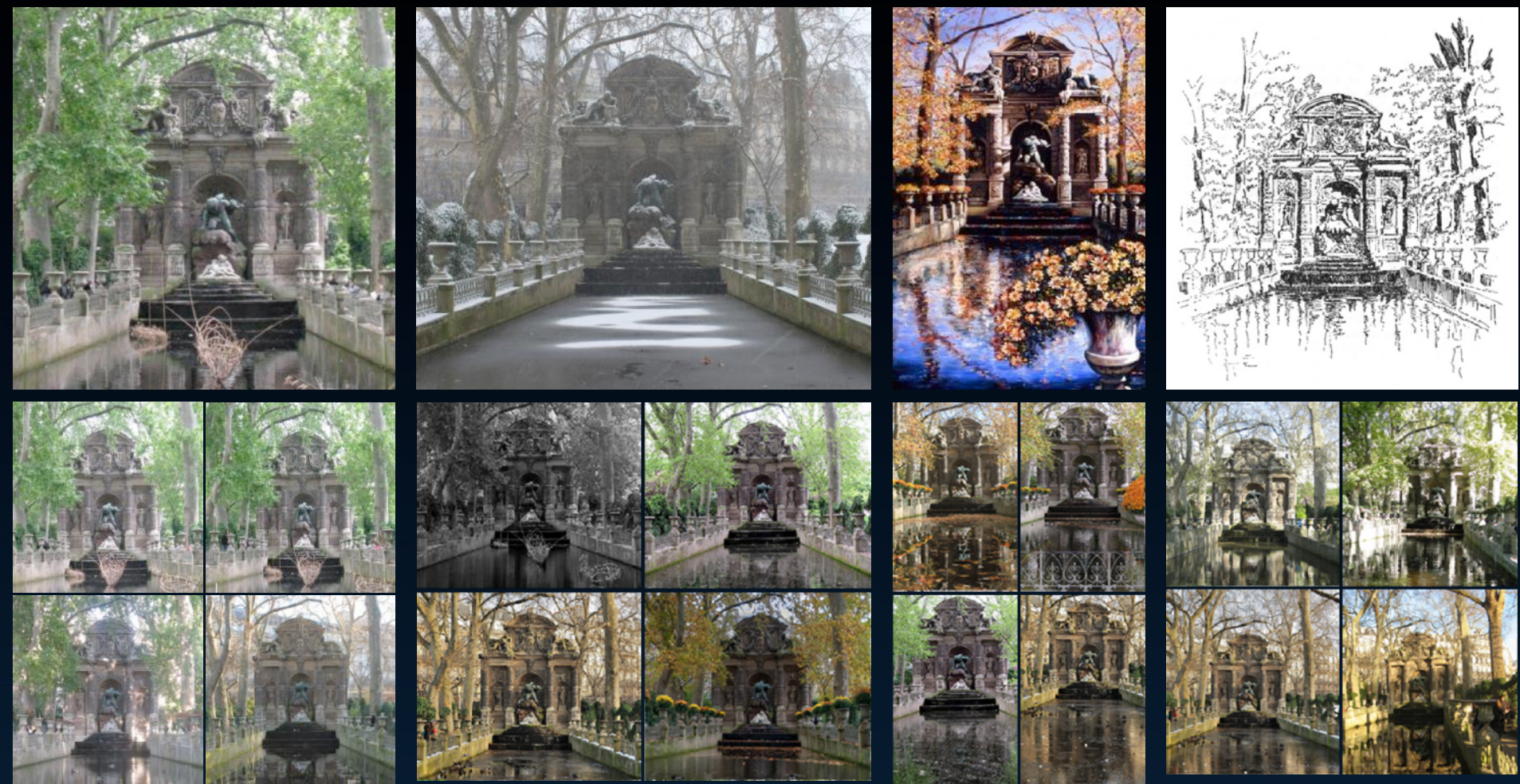
Visually similar



OUR GOAL



OUR GOAL



WHY IS THIS SO HARD?

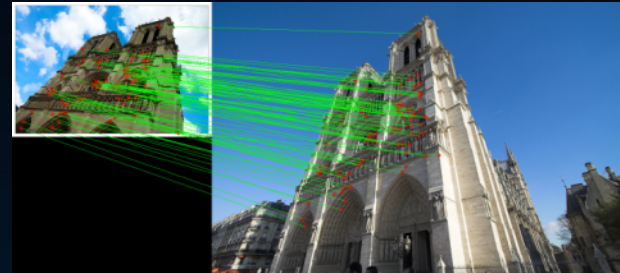


IMAGE RETRIEVAL

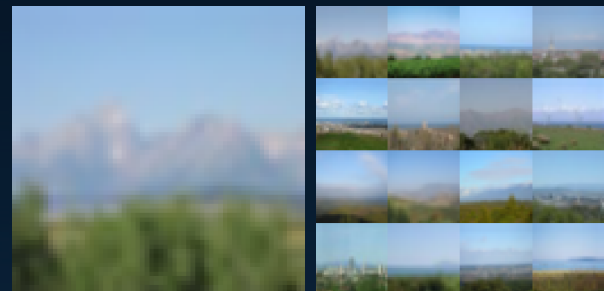
- Color-histograms
 - QBIC [Flickner et al., 1995]
 - Pentland et al., 1996
 - ...



- SIFT-based approaches
 - Lowe, 1999, 2004
 - Sivic and Zisserman, 2003
 - Chum et al., 2007-10
 - Jegou et al., 2008-10
 - Lazebnik et al., 2009
 - ...



- Gist-based approaches
 - Oliva and Torralba, 2006
 - Hays and Efros, 2007
 - Weiss et al., 2007
 - Torralba et al., 2008
 - ...

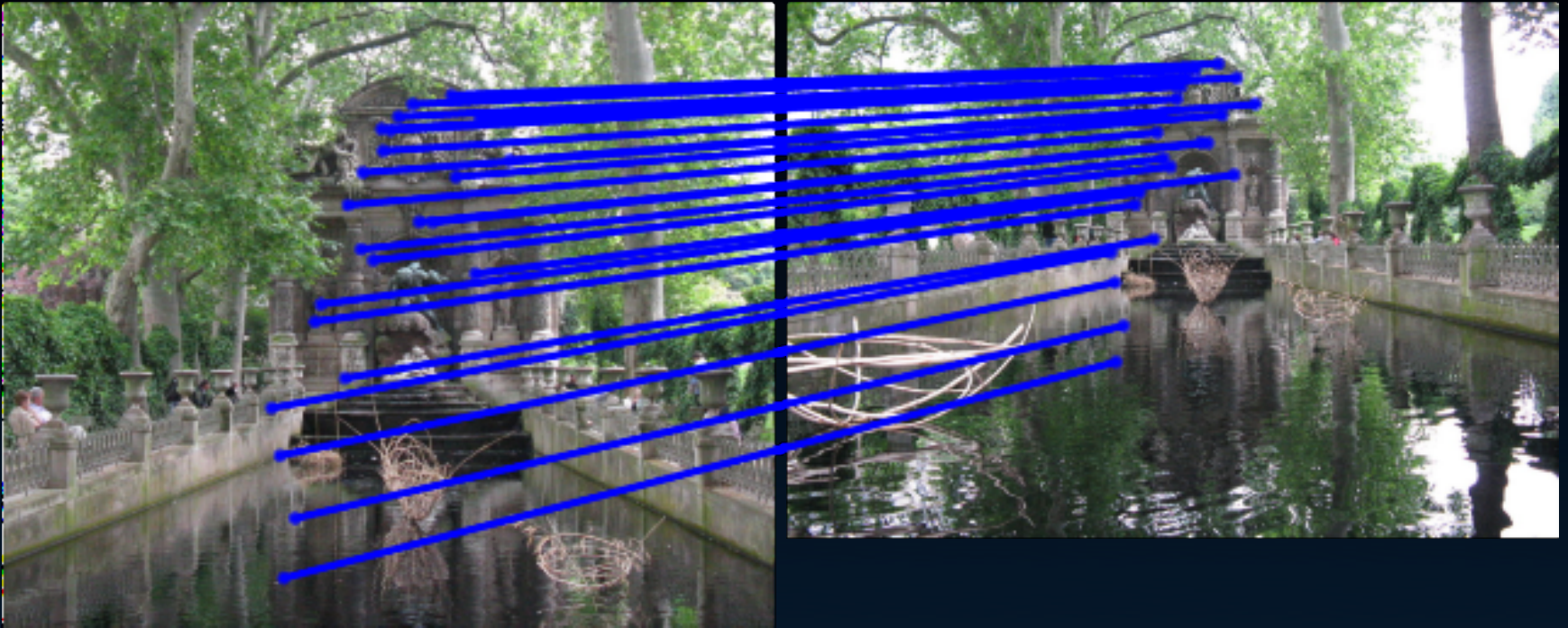


EXAMPLE: SIFT MATCHING

EXAMPLE: SIFT MATCHING



EXAMPLE: SIFT MATCHING



EXAMPLE: SIFT MATCHING

EXAMPLE: SIFT MATCHING



EXAMPLE: SIFT MATCHING

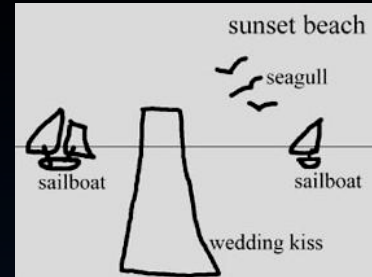


CROSS-DOMAIN MATCHING

CROSS-DOMAIN MATCHING

- Sketch to Photo

- Kato et al., 1992
- Liang et al., 2004
- Chen et al., 2009
- Eitz et al., 2010
- ...



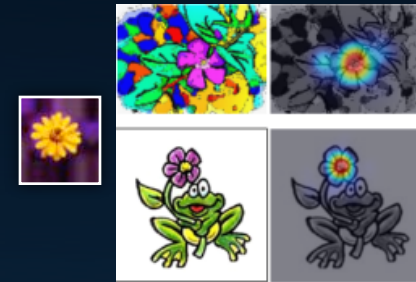
- Painting to Photo

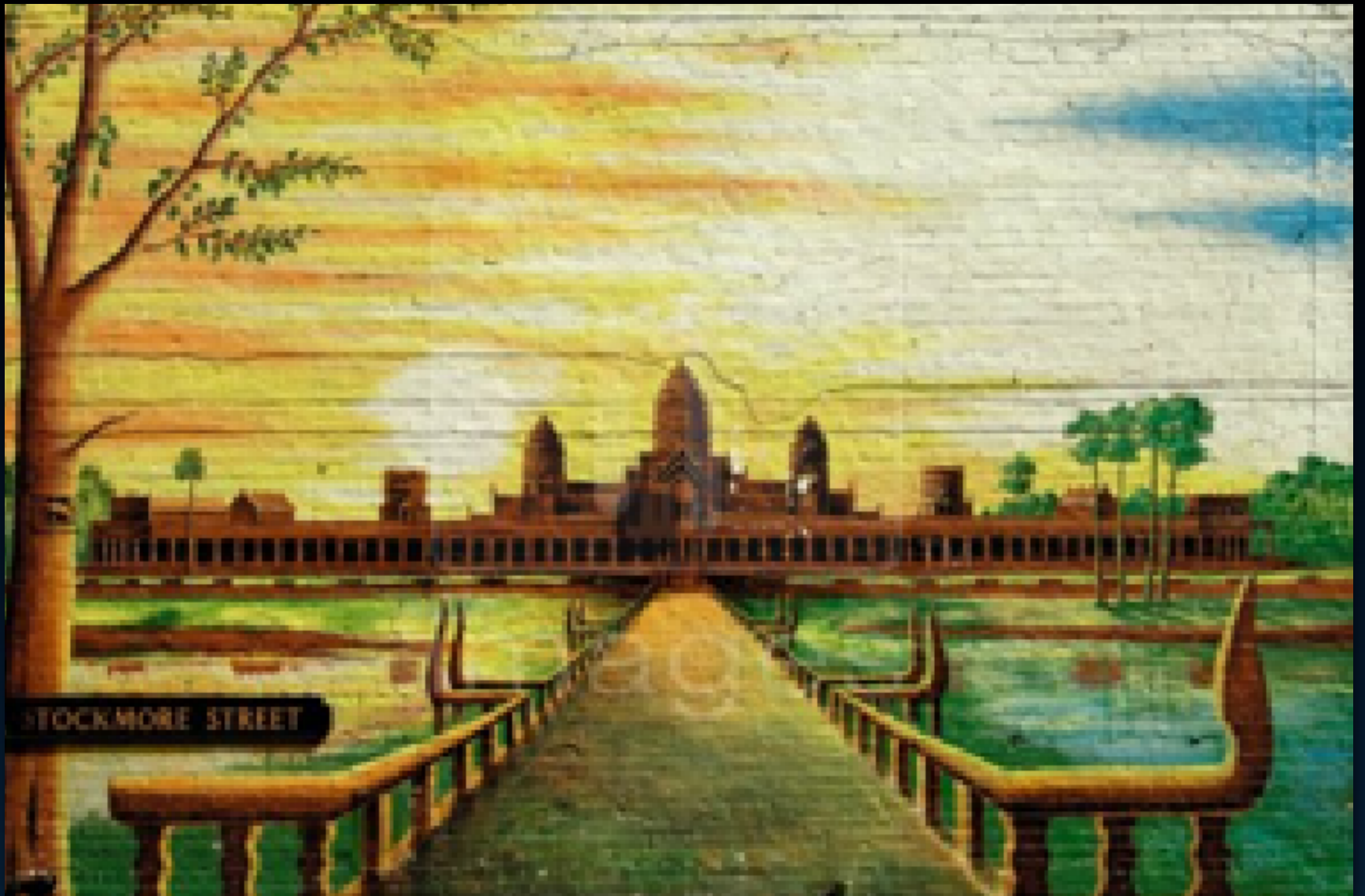
- Hirata et al., 1992
- Russell et al., 2011
- ...



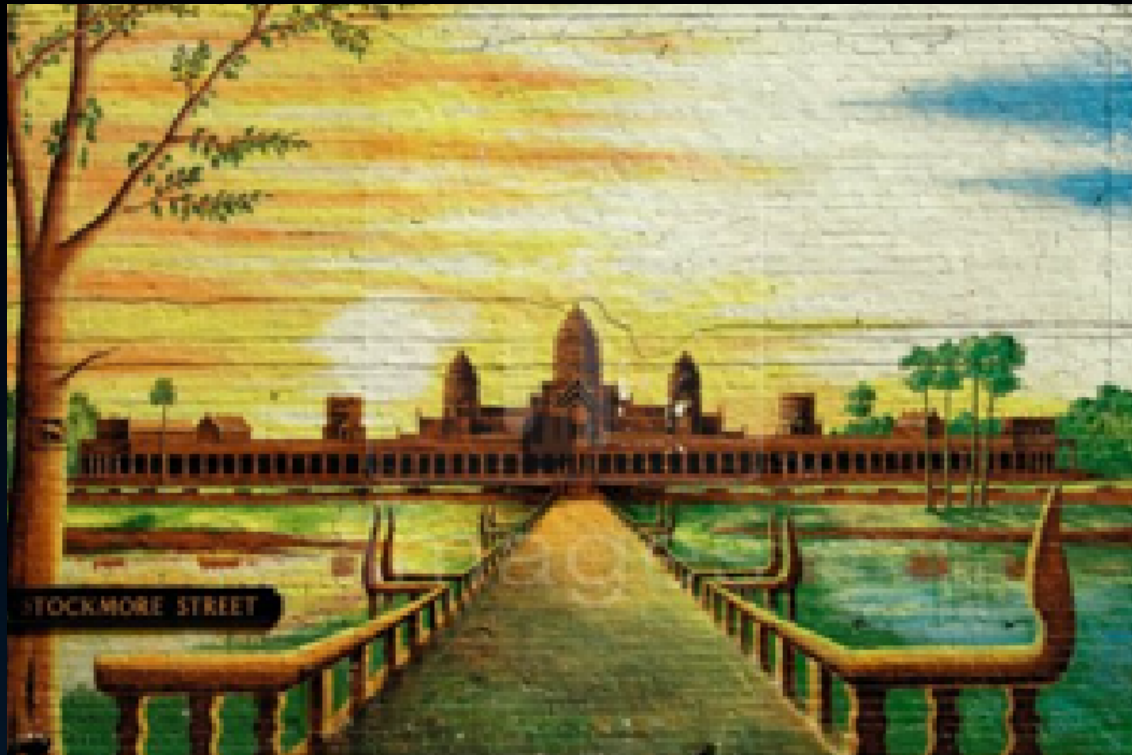
- Domain-invariant

- Shechtman and Irani, 2007
- Boiman and Irani, 2006
- ...



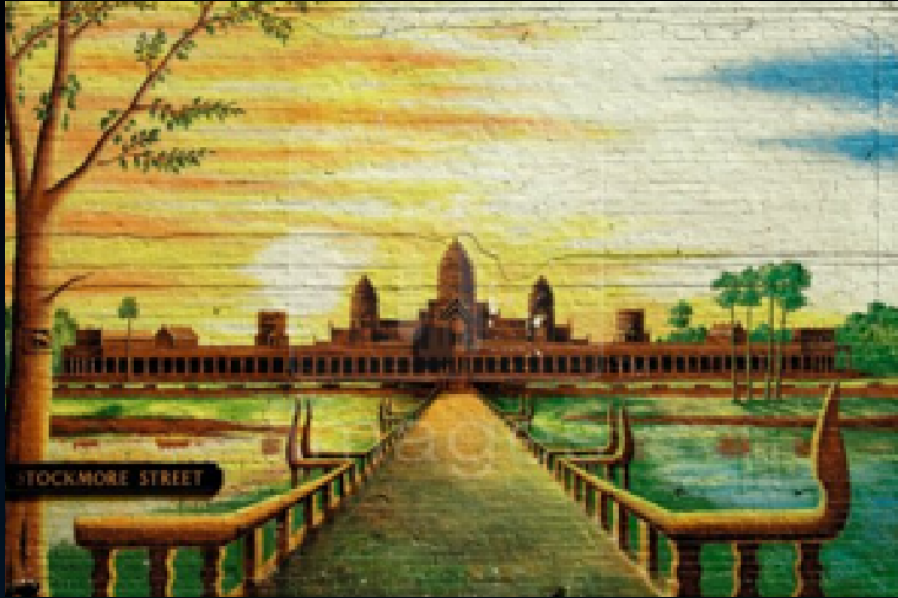


Input Query



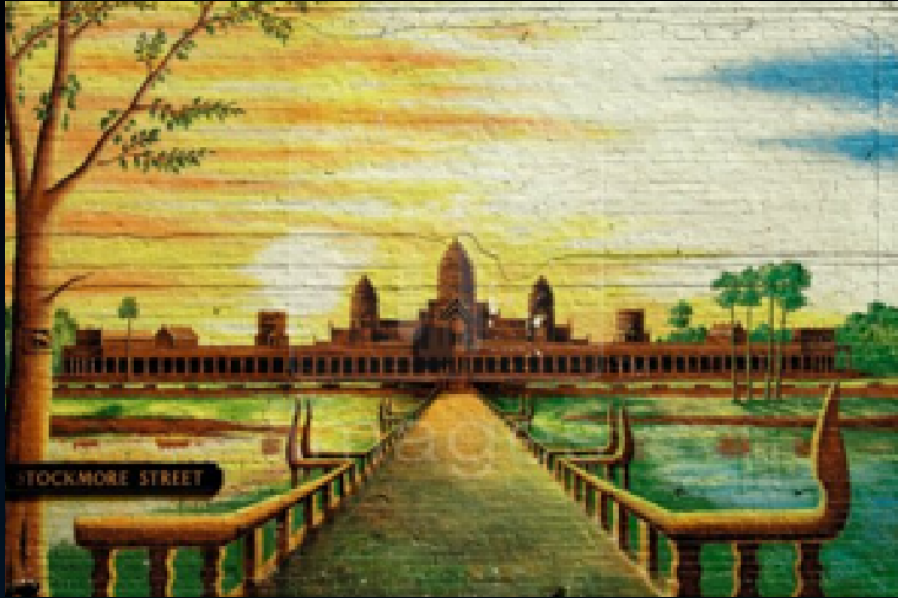
Top Matches

Input Query



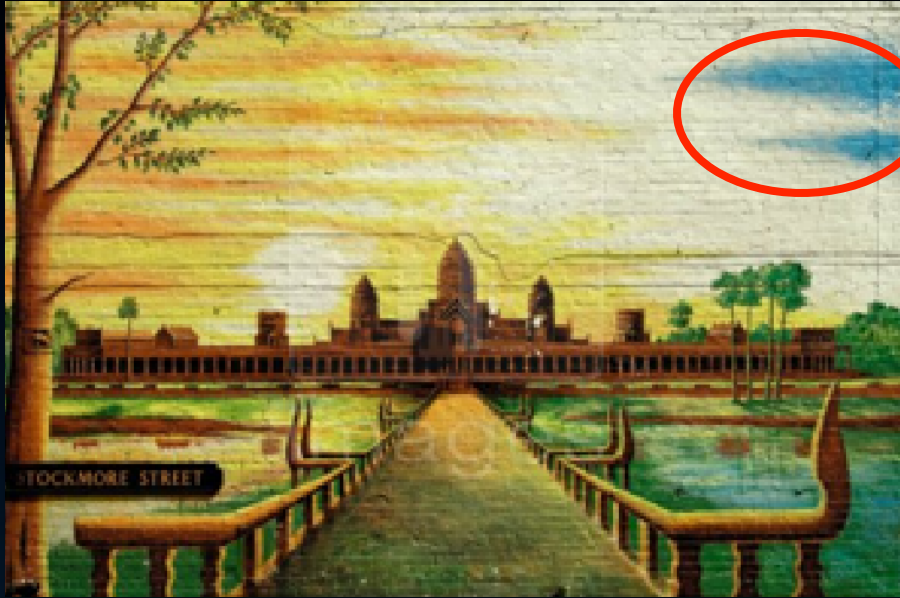
Top Matches

Input Query



Top Matches

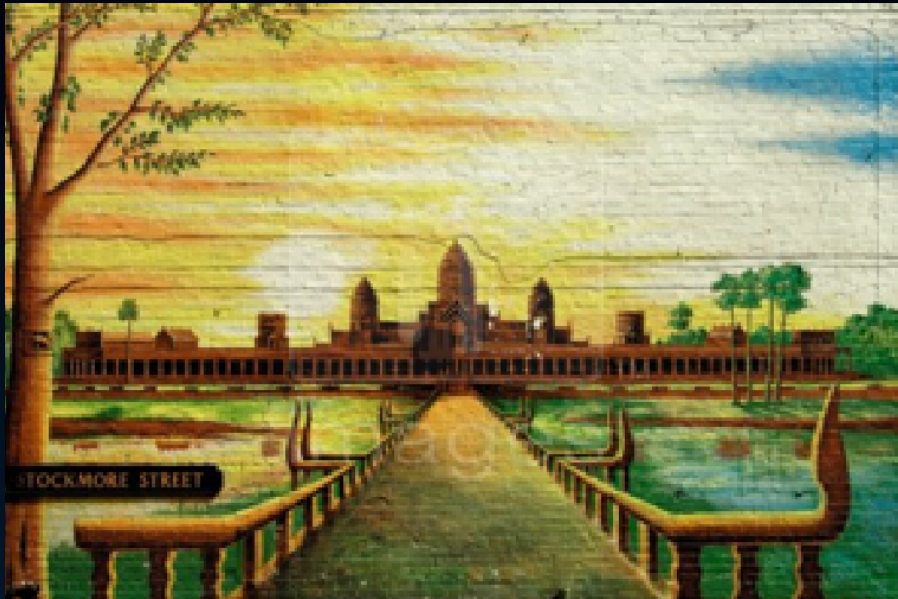
Input Query



Top Matches

IMPORTANT PARTS?

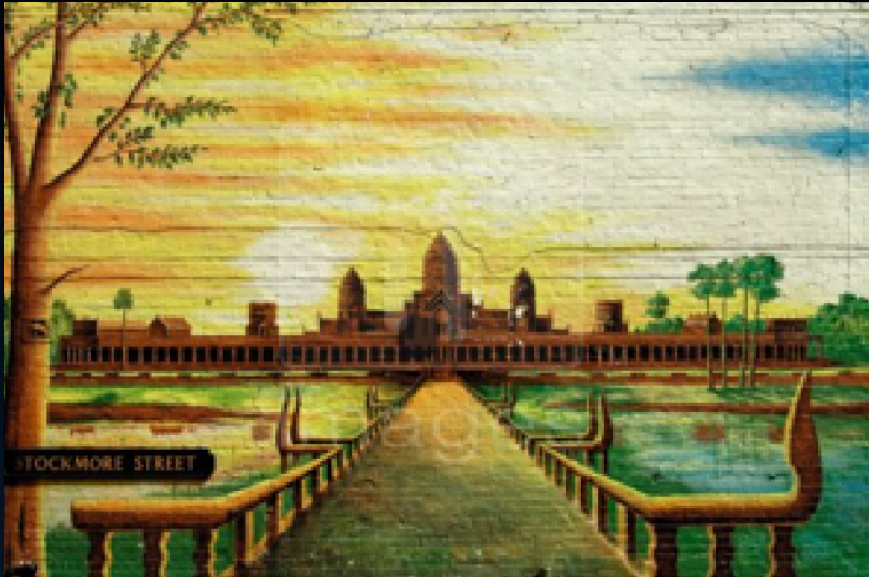
Input Query



Important Parts

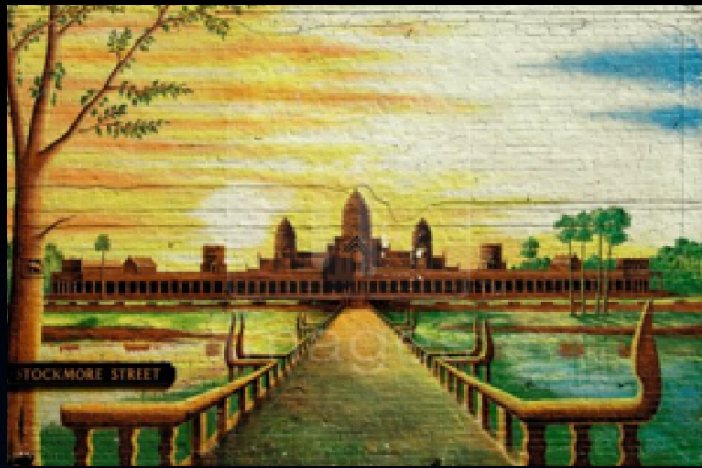


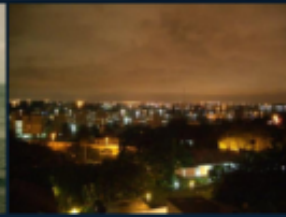
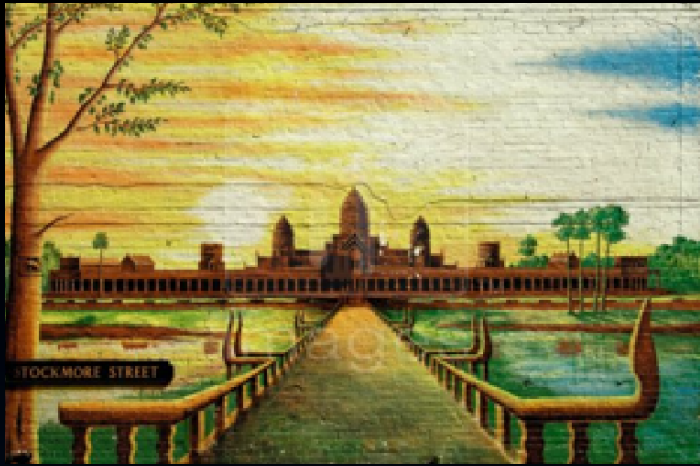
Input Query

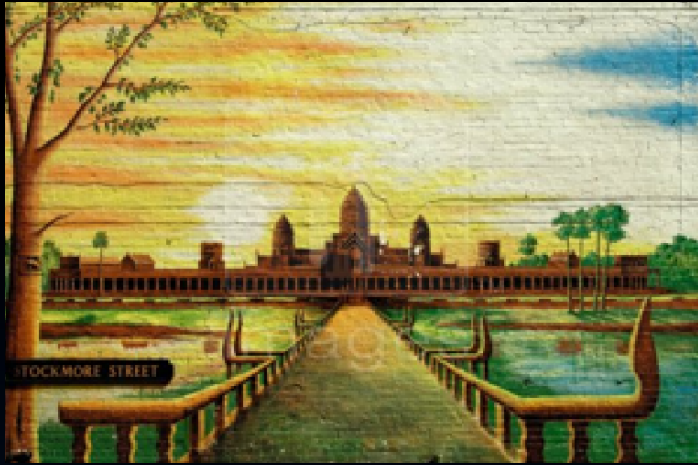


Top Matches

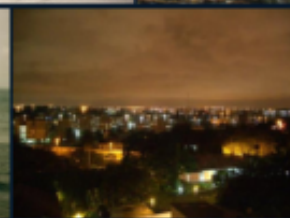


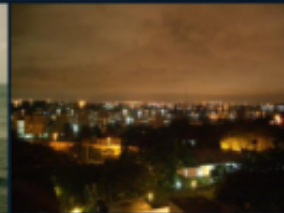
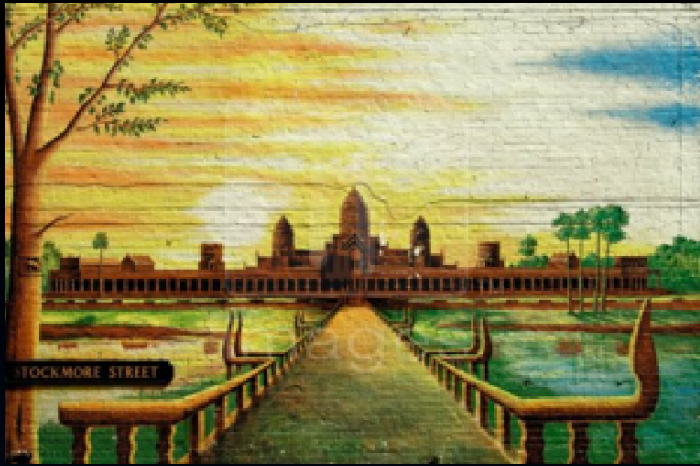


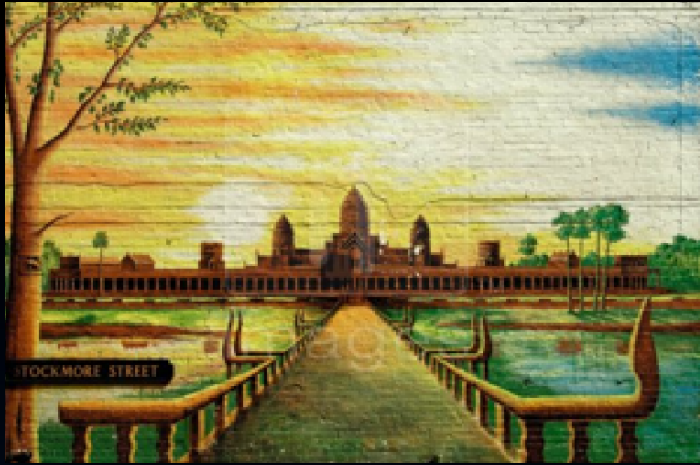


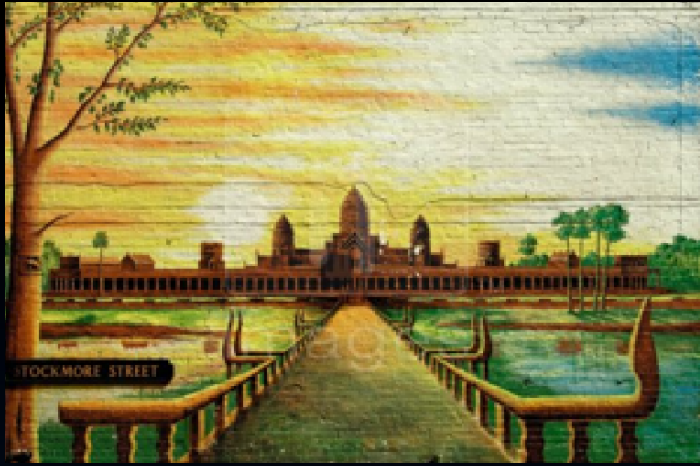


“Data-driven Uniqueness”



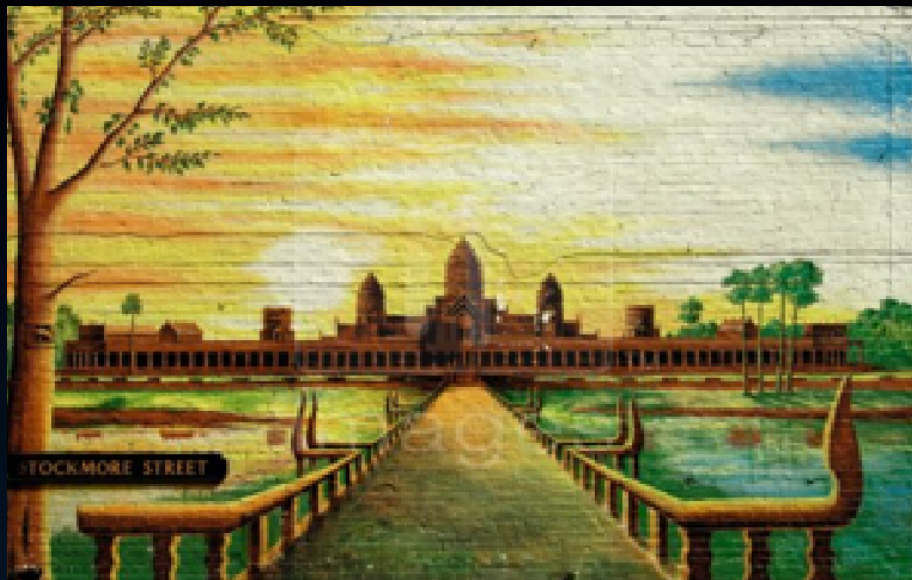






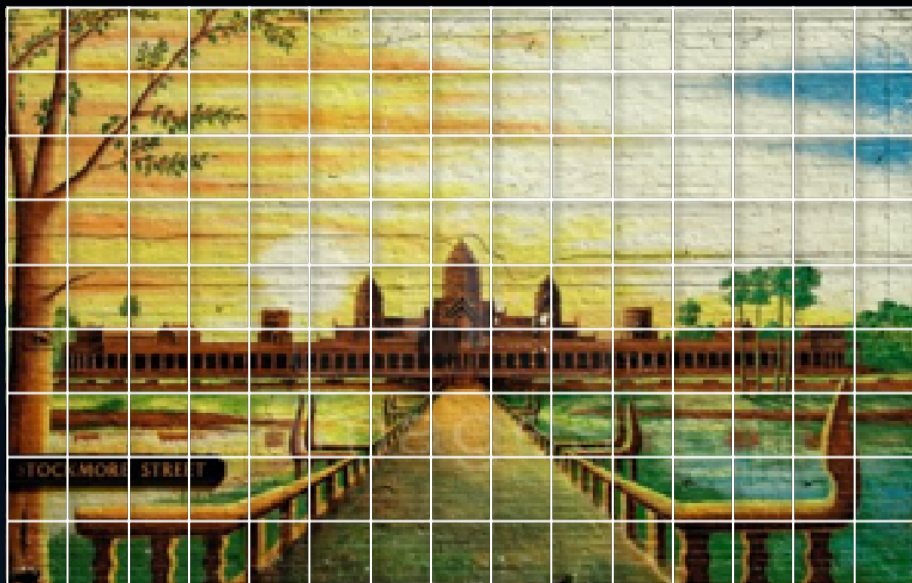
FEATURE REPRESENTATION

HISTOGRAM OF ORIENTED GRADIENTS (HOG)



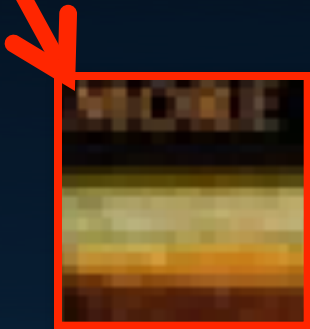
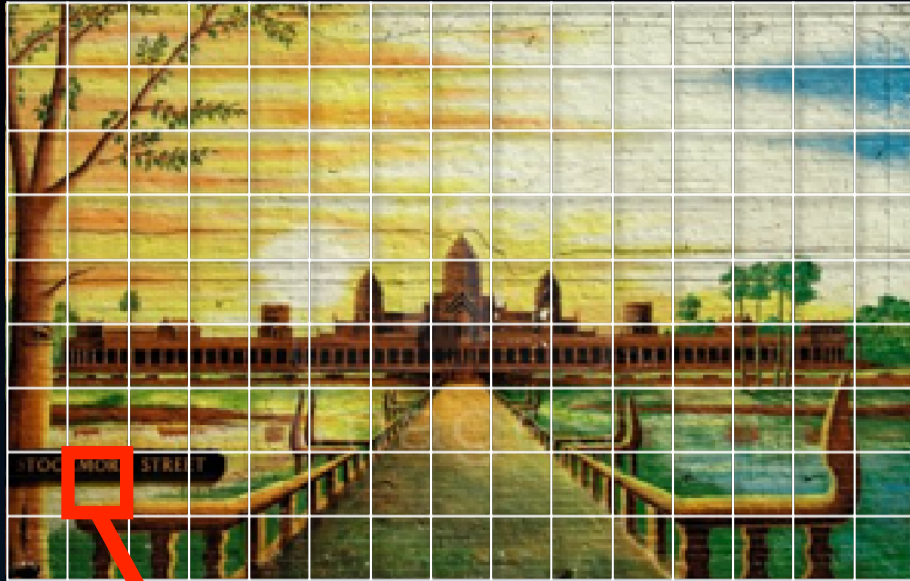
FEATURE REPRESENTATION

HISTOGRAM OF ORIENTED GRADIENTS (HOG)



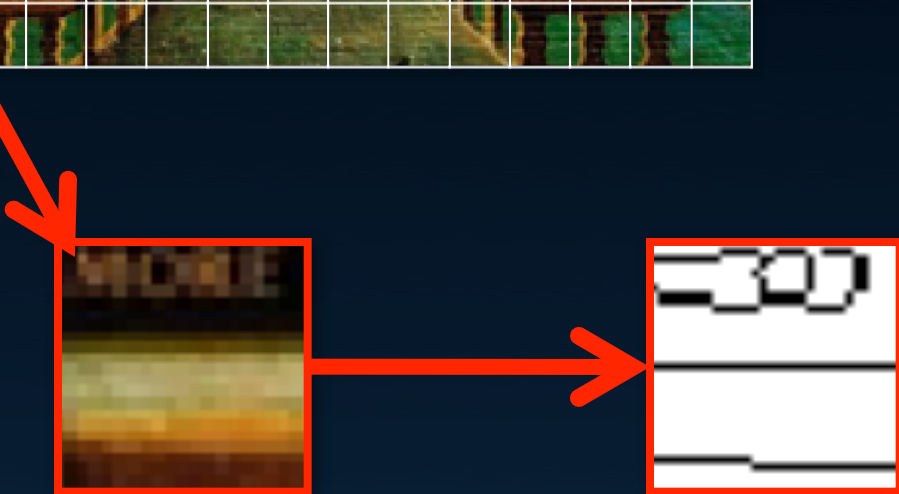
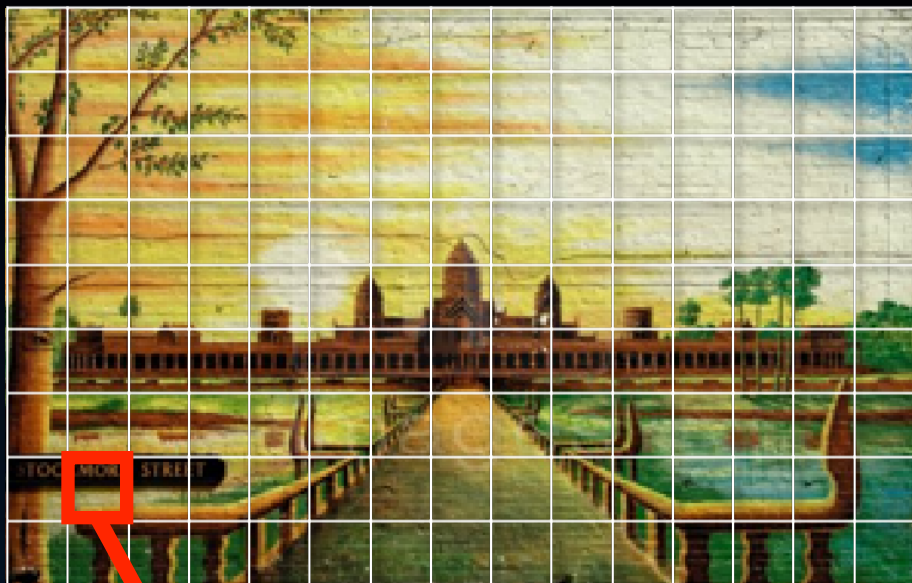
FEATURE REPRESENTATION

HISTOGRAM OF ORIENTED GRADIENTS (HOG)



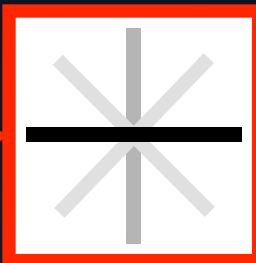
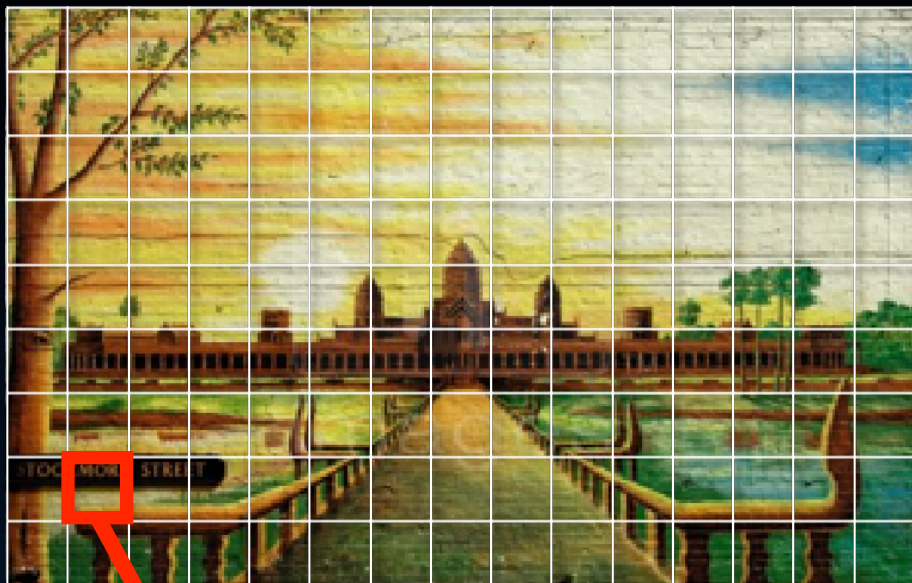
FEATURE REPRESENTATION

HISTOGRAM OF ORIENTED GRADIENTS (HOG)



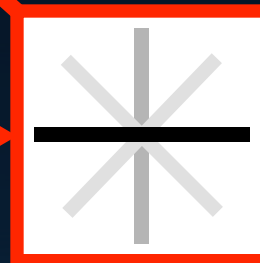
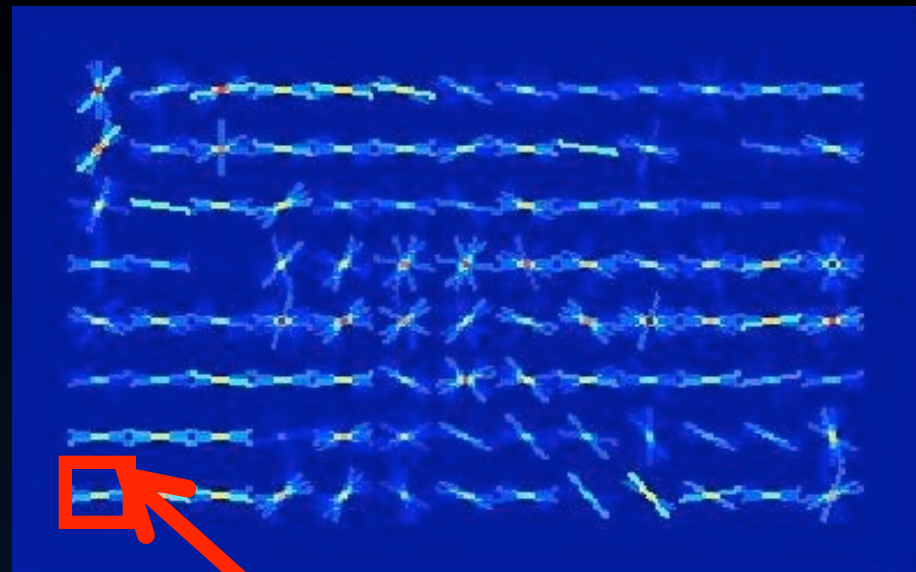
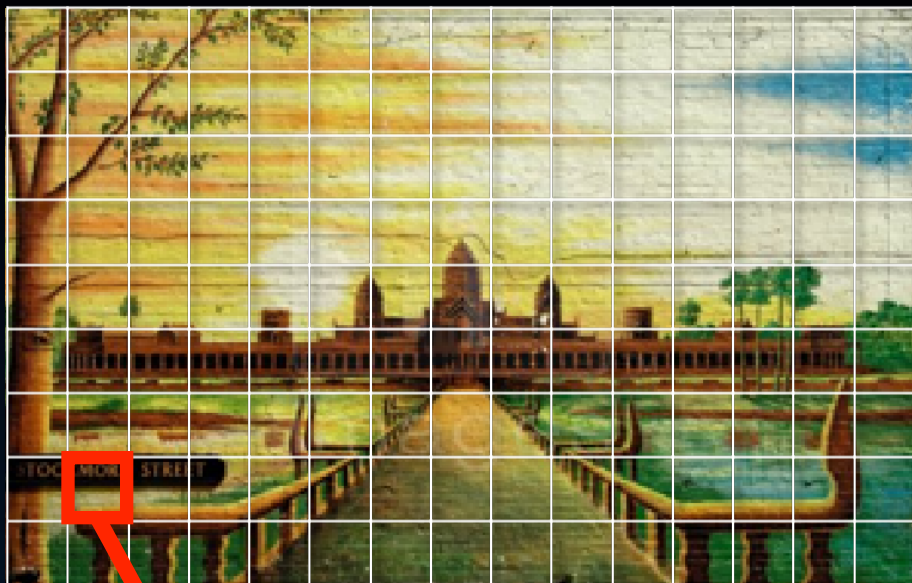
FEATURE REPRESENTATION

HISTOGRAM OF ORIENTED GRADIENTS (HOG)

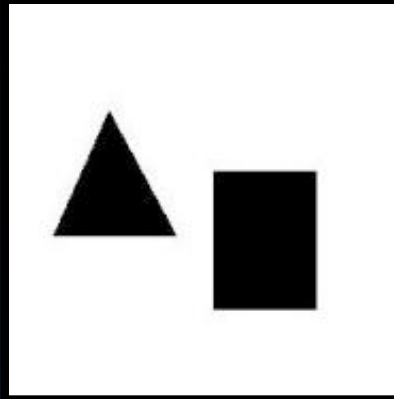


FEATURE REPRESENTATION

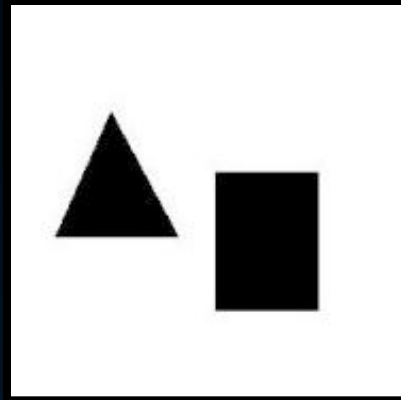
HISTOGRAM OF ORIENTED GRADIENTS (HOG)



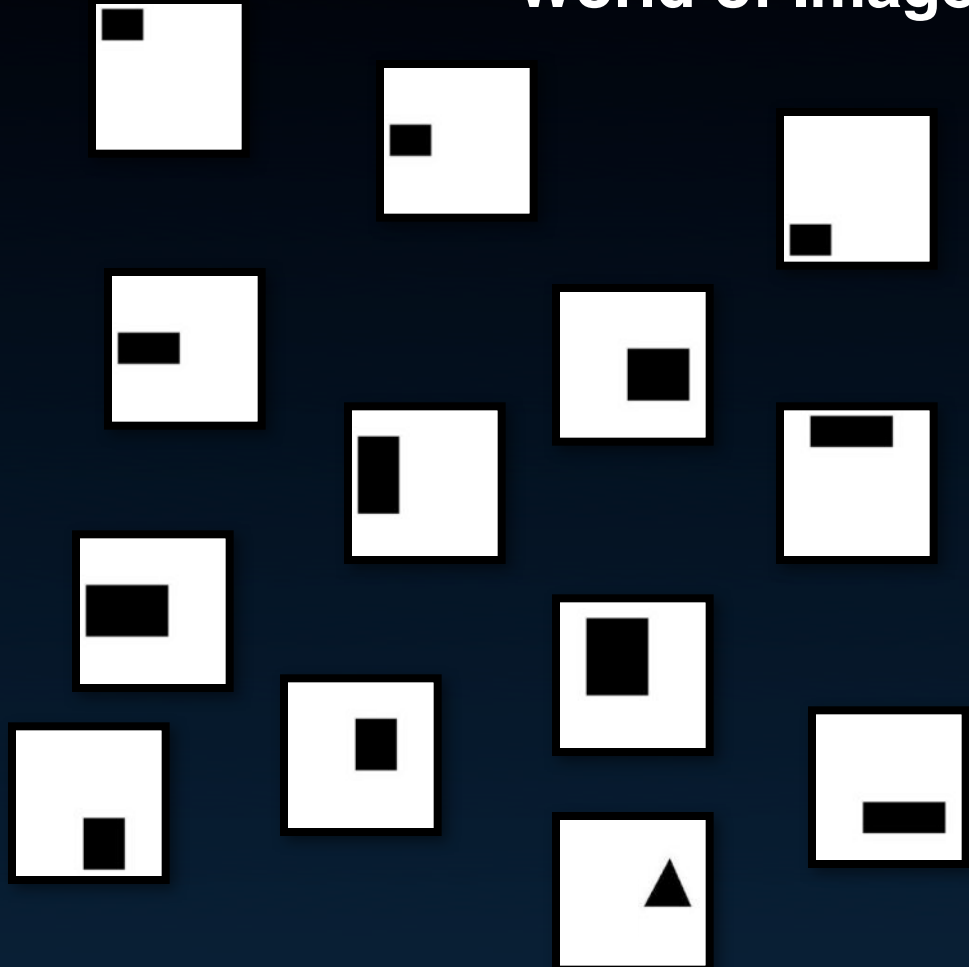
WHAT IS UNIQUE?



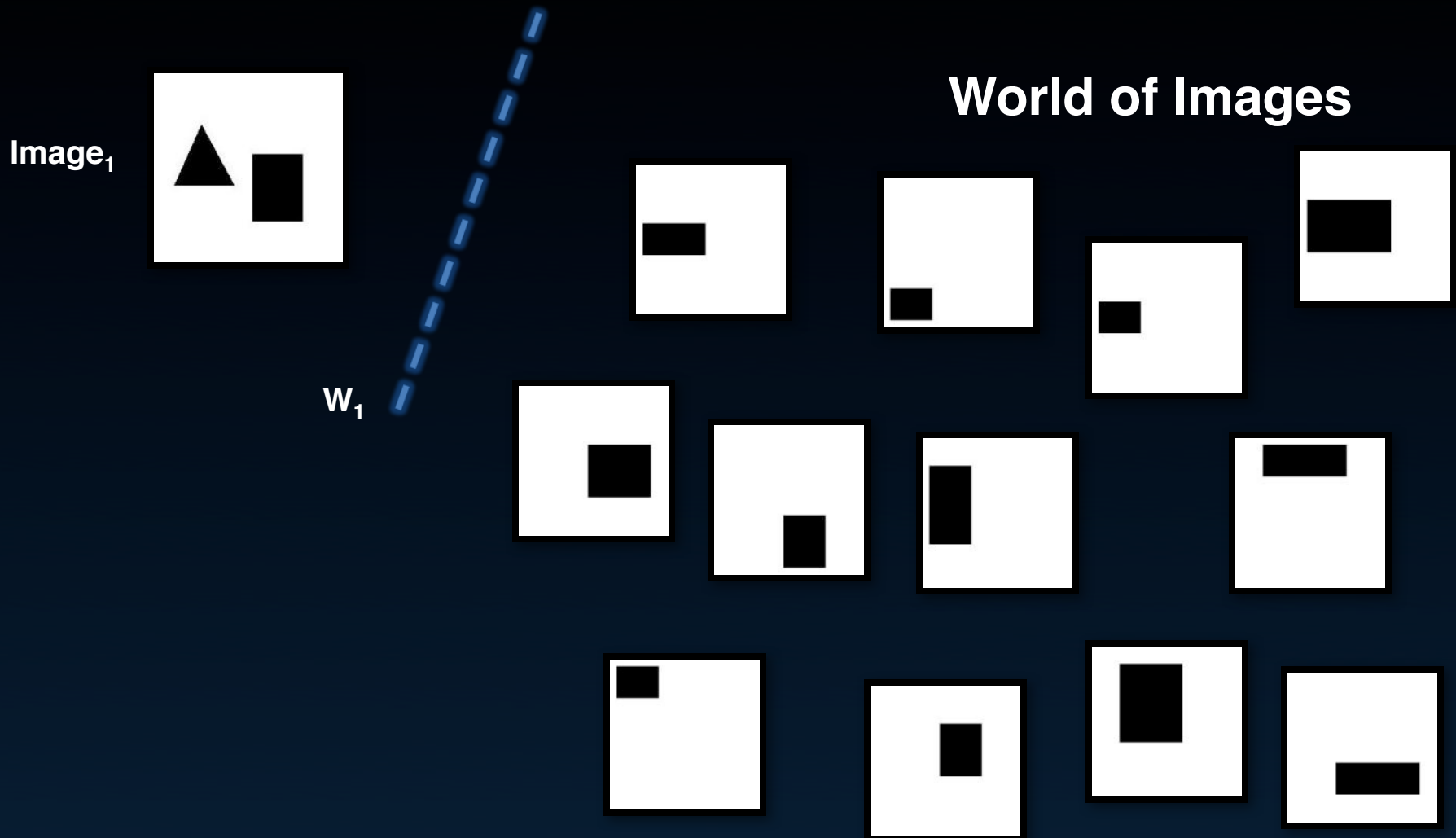
WHAT IS UNIQUE GIVEN THIS WORLD?



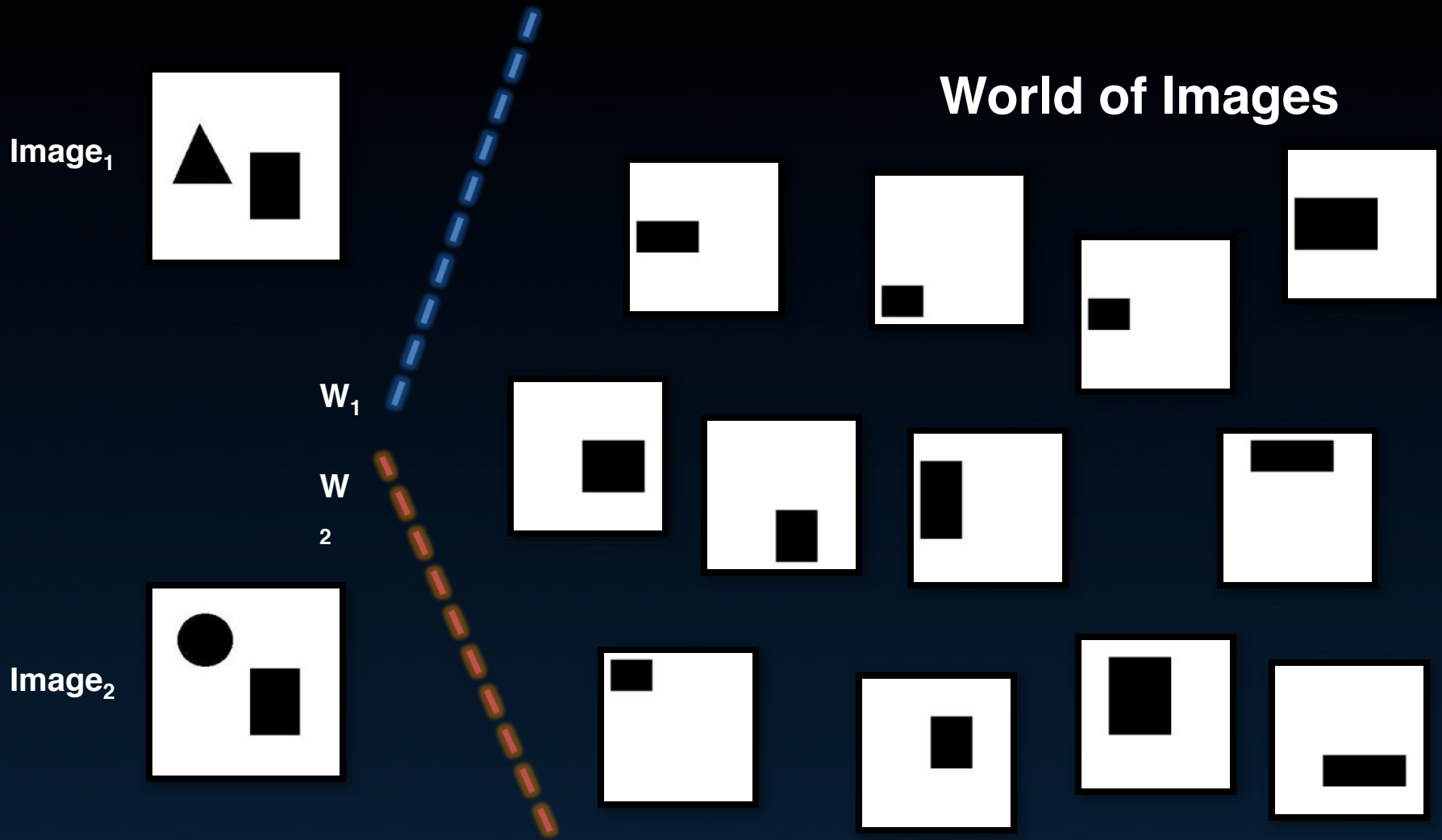
World of Images



PER-EXEMPLAR SVM



PER-EXEMPLAR SVM

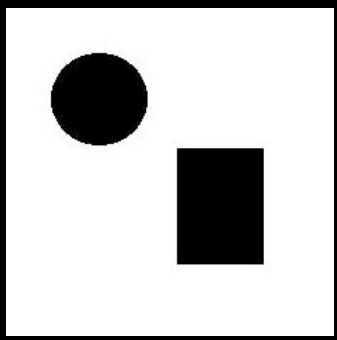
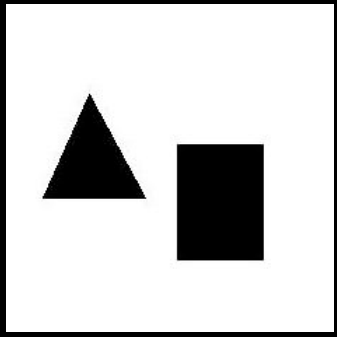


Tomasz Malisiewicz, Abhinav Gupta, Alexei A. Efros.

Ensemble of Exemplar-SVMs for Object Detection and Beyond. In ICCV, 2011.

VISUALIZING UNIQUENESS

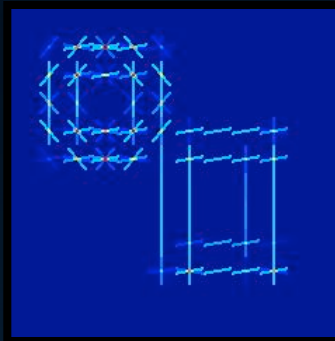
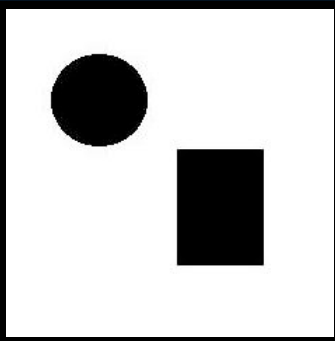
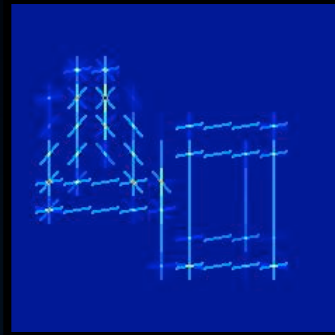
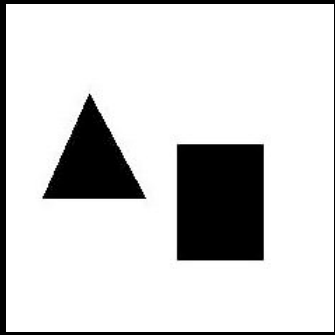
Query



VISUALIZING UNIQUENESS

Query

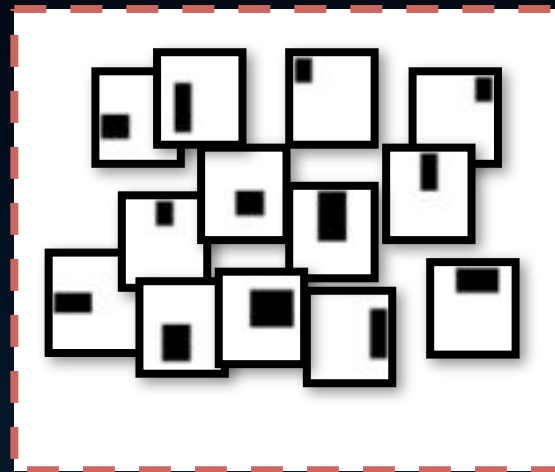
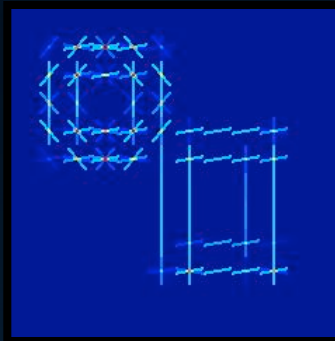
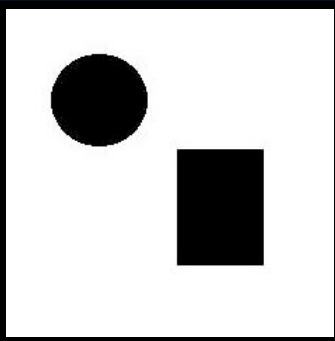
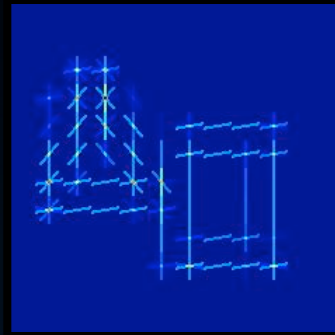
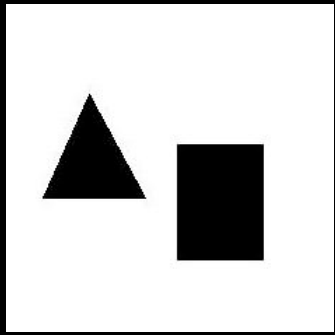
Before



VISUALIZING UNIQUENESS

Query

Before



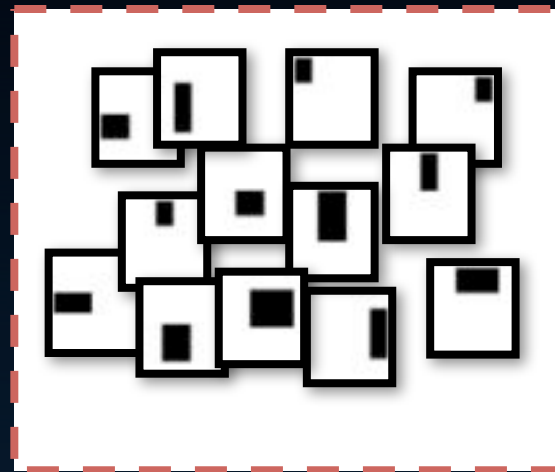
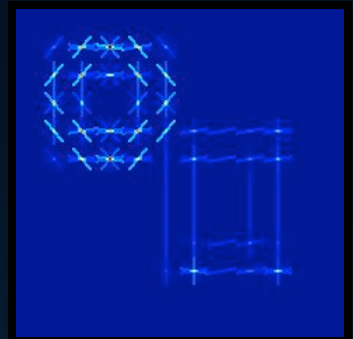
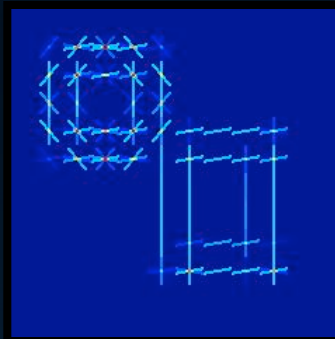
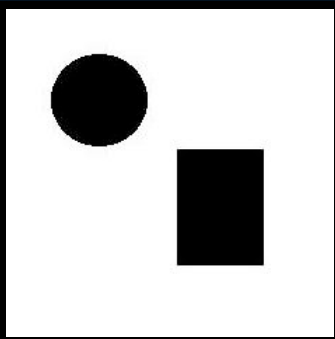
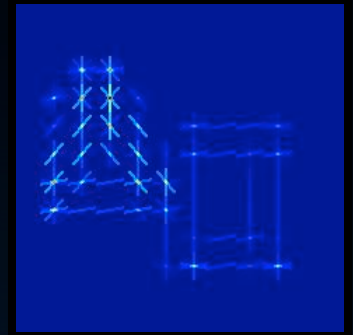
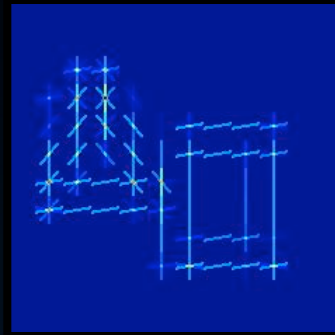
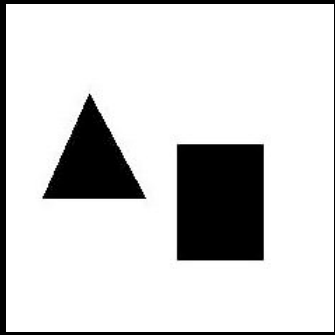
World of Images

VISUALIZING UNIQUENESS

Query

Before

After



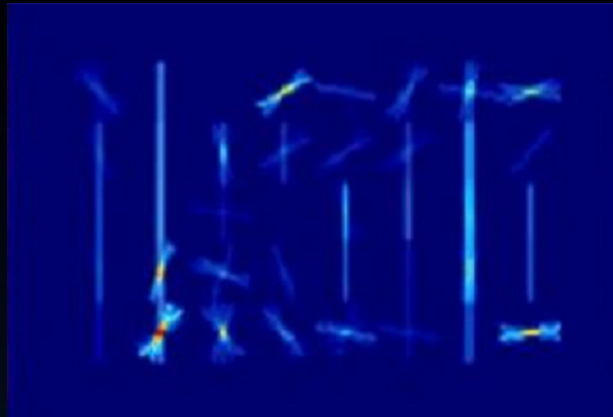
World of Images



Input Query



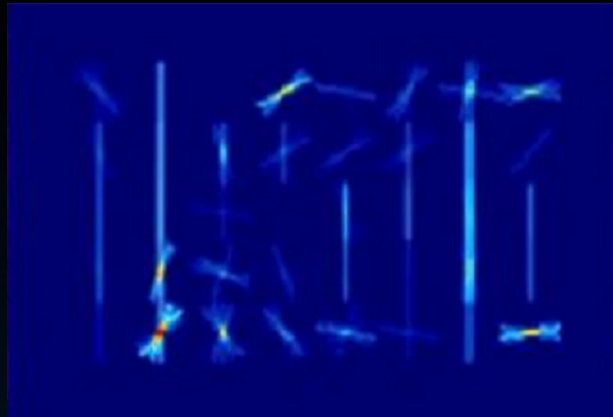
Input Query



HOG



Input Query



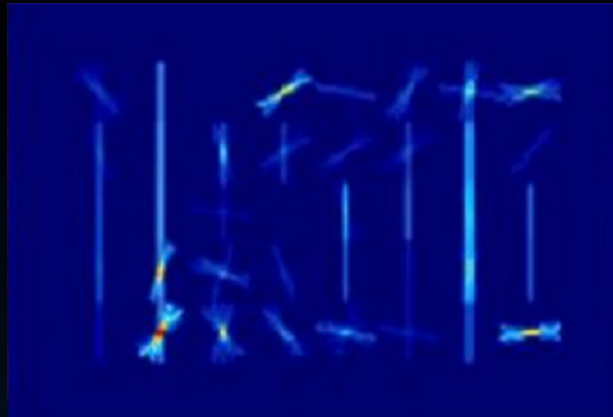
HOG



Top Match



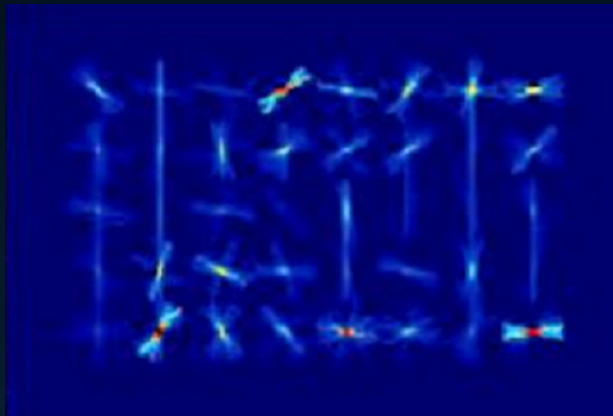
Input Query



HOG



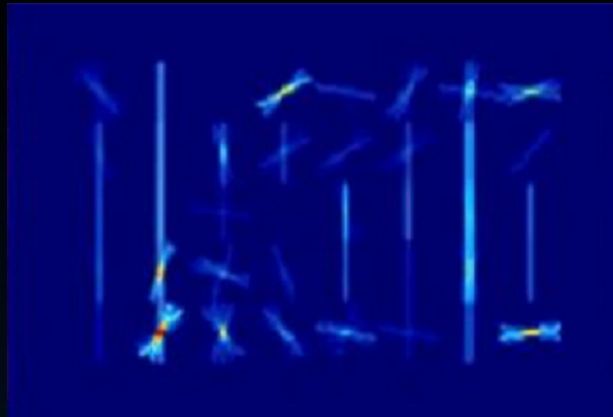
Top Match



Learnt Weights



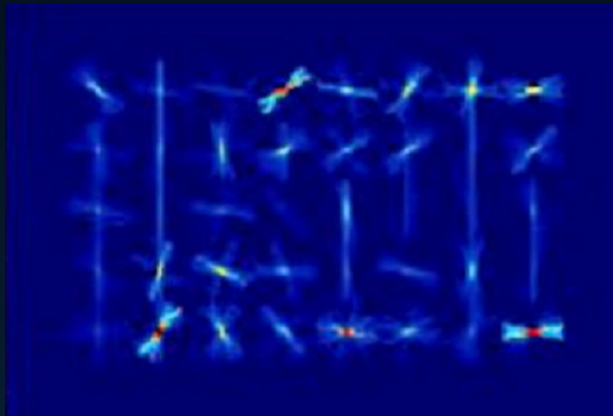
Input Query



HOG



Top Match



Learnt Weights



Top Match

SEARCH USING IMAGES

Input Query



SEARCH USING IMAGES

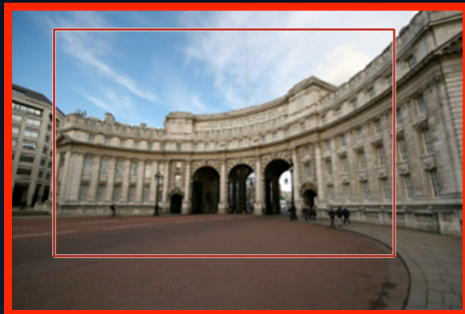
Input Query



Top Matches

SEARCH USING IMAGES

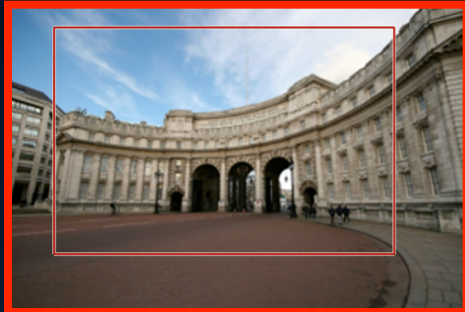
Input Query



Top Matches

SEARCH USING IMAGES

Input Query



Top Matches

SEARCH USING PAINTINGS



Input Painting

SEARCH USING PAINTINGS



Input Painting



GIST

SEARCH USING PAINTINGS



Input Painting



GIST



Bag-of-Words

SEARCH USING PAINTINGS



Input Painting



GIST



Bag-of-Words



Tiny Images

SEARCH USING PAINTINGS



Input Painting



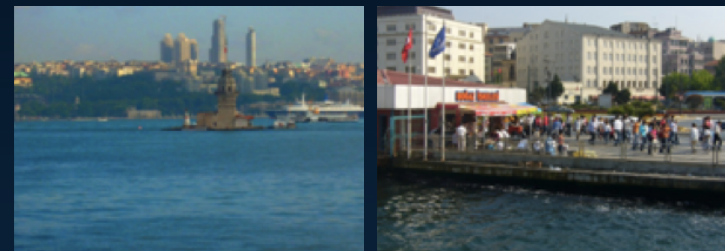
GIST



Bag-of-Words



Tiny Images



HOG

SEARCH USING PAINTINGS



Input Painting



GIST



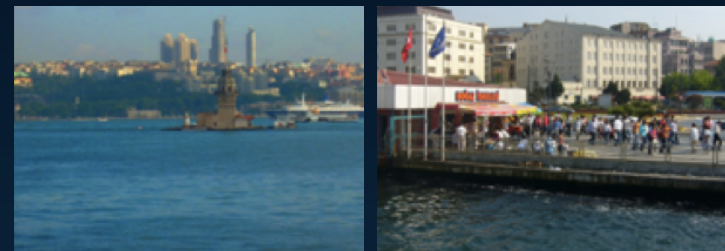
Bag-of-Words



Our Approach



Tiny Images



HOG

SEARCH USING PAINTINGS



Input Painting



Top Matches

SEARCH USING PAINTINGS

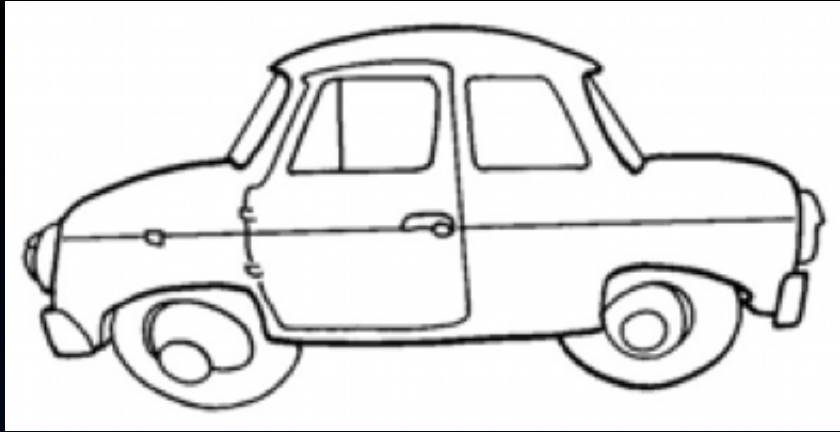


Input Painting



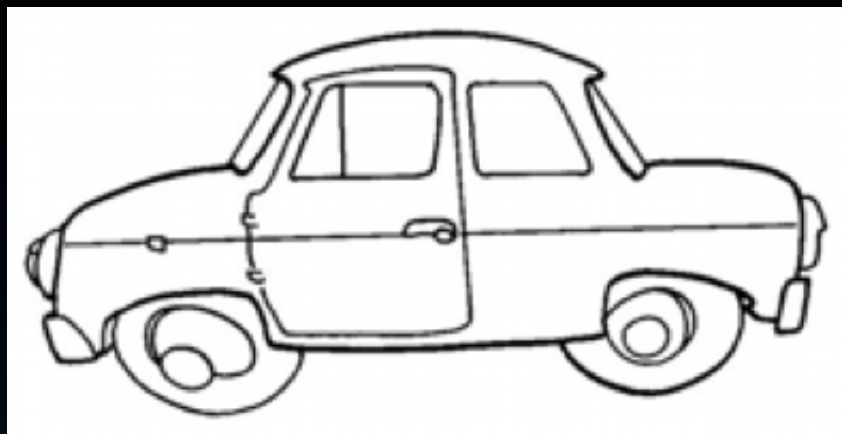
Top Matches

SEARCH USING SKETCHES



Input Sketch

SEARCH USING SKETCHES



Input Sketch



Tiny Images



GIST

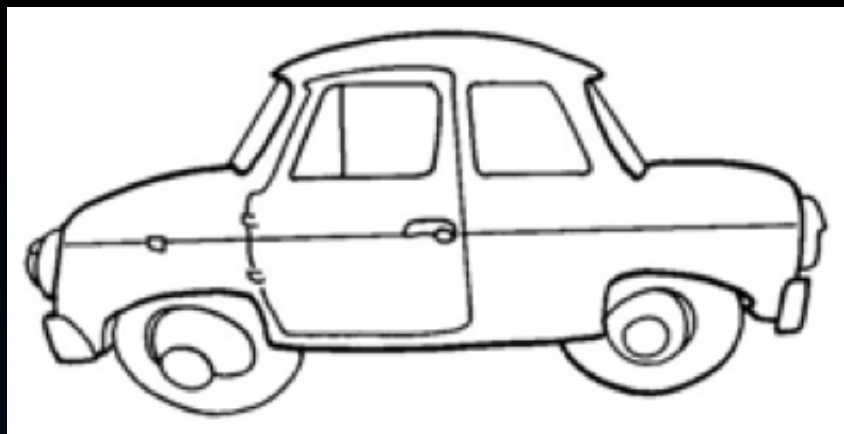


Bag-of-Words



HOG

SEARCH USING SKETCHES



Input Sketch



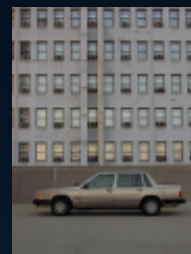
Tiny Images



GIST



Our Approach

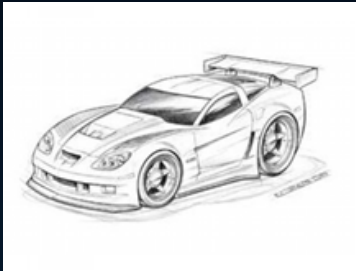
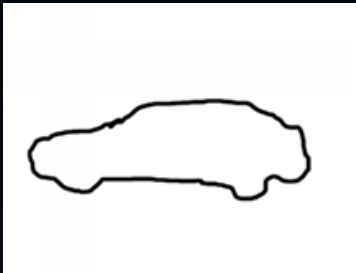


Bag-of-Words



HOG

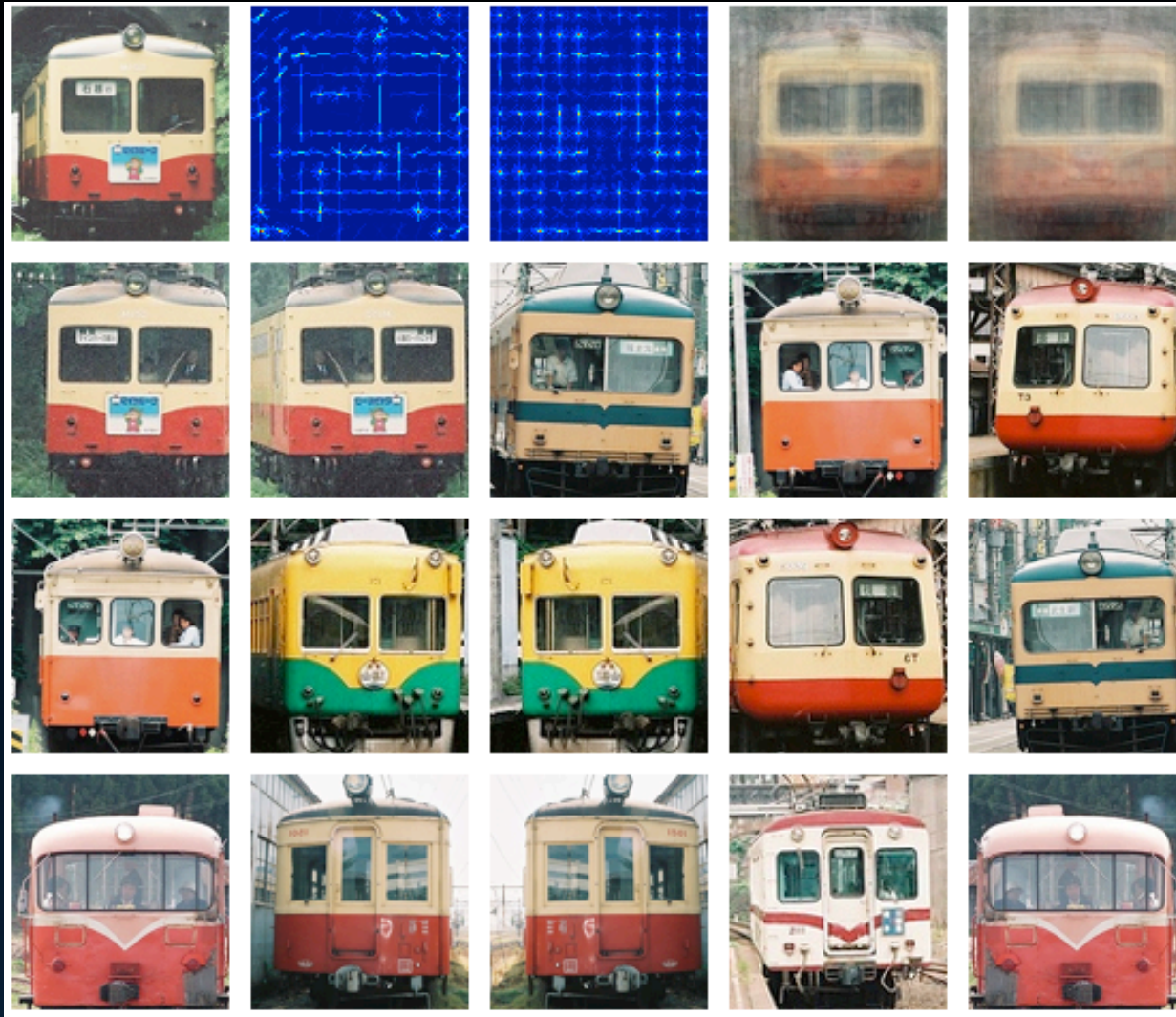
SEARCH USING SKETCHES



SEARCH USING SKETCHES



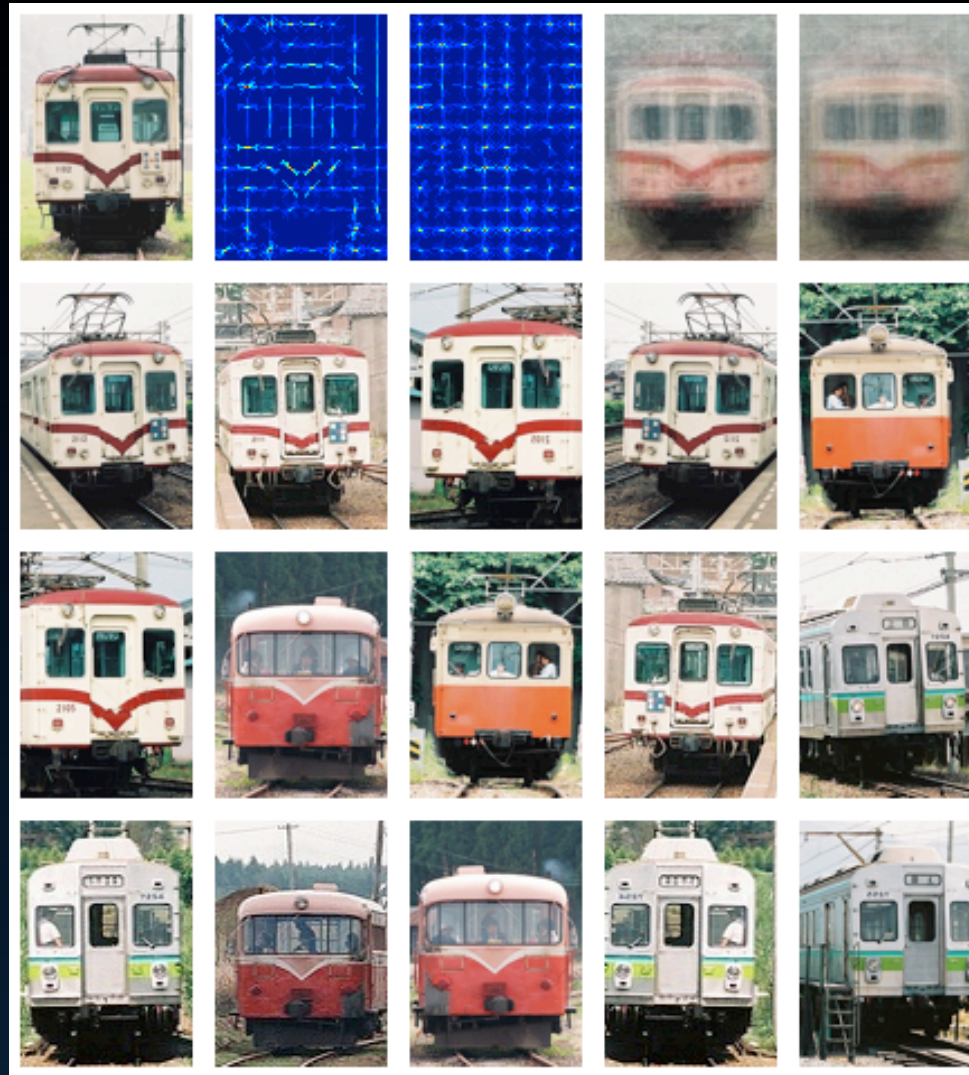
SEARCH USING OBJECTS



Tomasz Malisiewicz, Abhinav Gupta, Alexei A. Efros.

Ensemble of Exemplar-SVMs for Object Detection and Beyond. In ICCV, 2011.

SEARCH USING OBJECTS



Tomasz Malisiewicz, Abhinav Gupta, Alexei A. Efros.

Ensemble of Exemplar-SVMs for Object Detection and Beyond. In ICCV, 2011.

QUANTITATIVE EVALUATIONS

IMAGE INSTANCE RETRIEVAL

HOLIDAYS DATASET [Jegou et al., 2008]

#Query Images: 500 Photographs

#Retrieval Set Images ~ 1M Flickr Images

IMAGE (INSTANCE) MATCHING ON HOLIDAYS DATASET



IMAGE INSTANCE RETRIEVAL

HOLIDAYS DATASET

	Top-5			
Dataset Size	1,490	11,490	101,490	1,001,490
GIST	0.0106	0.0106	0.0106	0.0106
Tiny Images	0.0106	0.0106	0.0106	0.0106
Spatial Pyramid	0.3417	0.3063	0.2471	0.1967
Our Approach	0.6588	0.6393	0.5890	0.5836

Mean true positive rate of the top-5 image matches as a function of dataset size

SKETCH-BASED IMAGE RETRIEVAL

SKETCH-BASED IMAGE RETRIEVAL

Query Sketches:

25 Car & 25 Bicycle Sketches

SKETCH-BASED IMAGE RETRIEVAL

Query Sketches:

25 Car & 25 Bicycle Sketches

Retrieval Set:

SKETCH-BASED IMAGE RETRIEVAL

Query Sketches:

25 Car & 25 Bicycle Sketches

Retrieval Set:

10,000 Annotated Images

SKETCH-BASED IMAGE RETRIEVAL

Query Sketches:

25 Car & 25 Bicycle Sketches

Retrieval Set:

10,000 Annotated Images

Pascal VOC 2007 Dataset

SKETCH-BASED IMAGE RETRIEVAL

Query Sketches:

25 Car & 25 Bicycle Sketches

Retrieval Set:

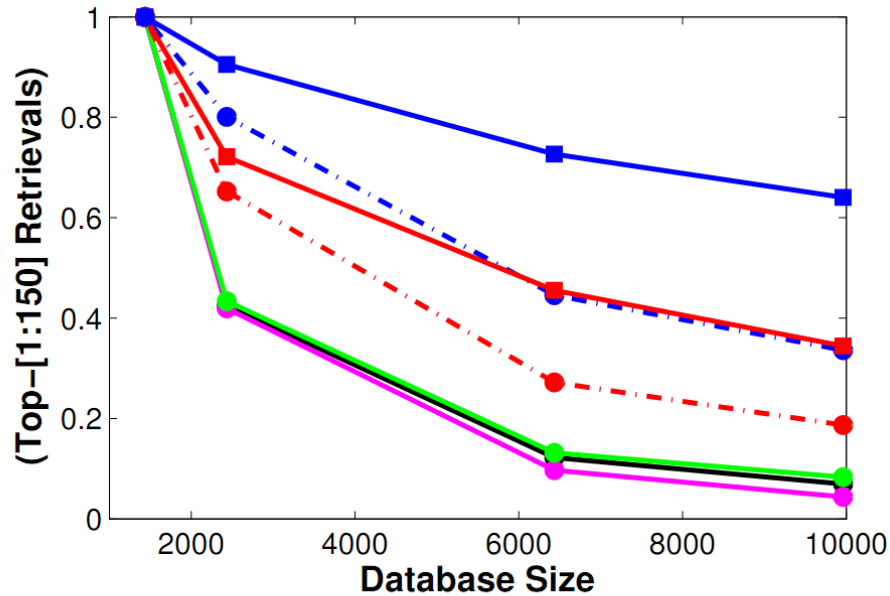
10,000 Annotated Images

Pascal VOC 2007 Dataset

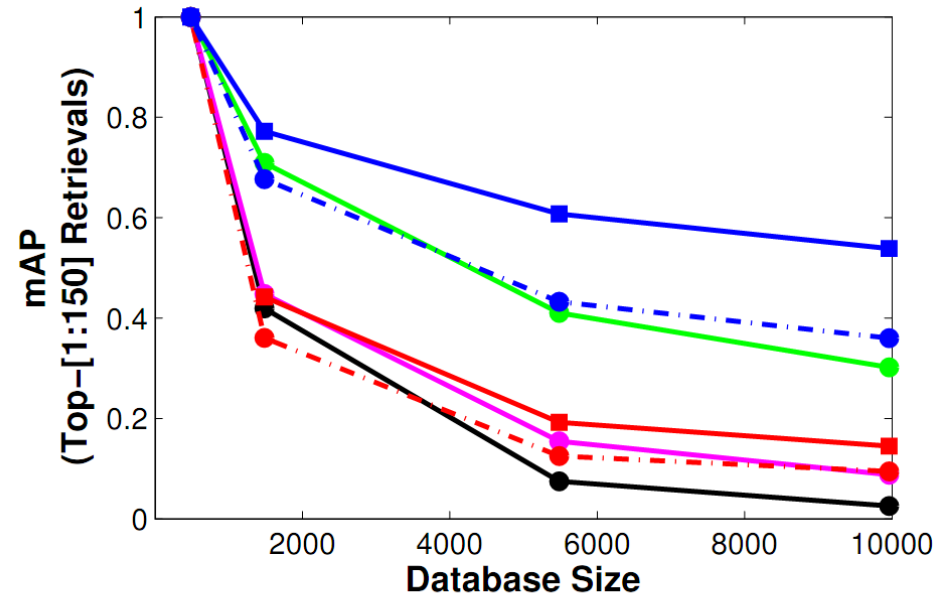
[Everingham et al., 2008]

SKETCH-BASED IMAGE RETRIEVAL

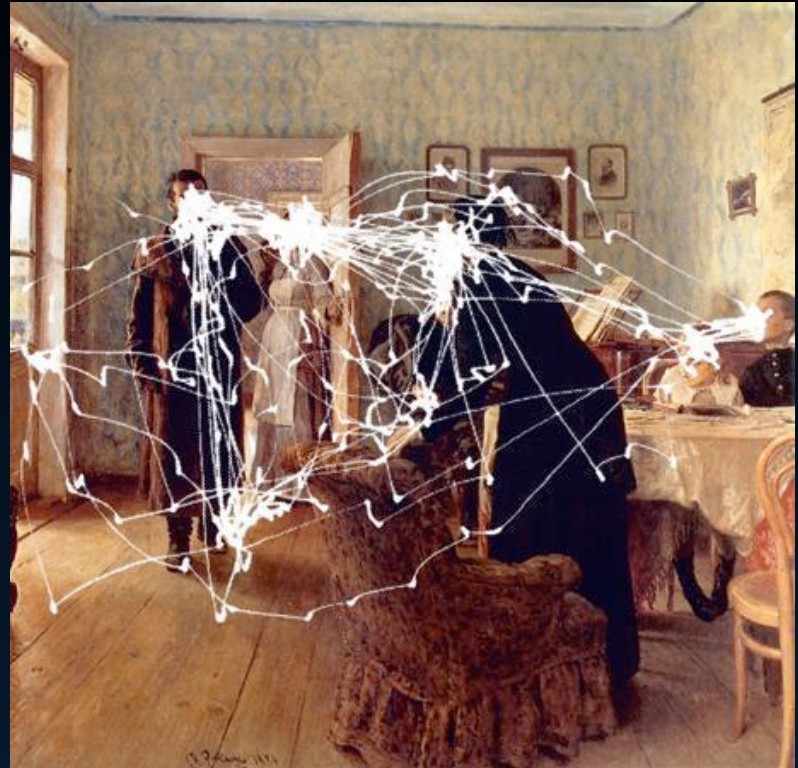
mAP for Car Sketches



mAP for Bicycle Sketches



SALIENCY



PROXY FOR SALIENCY

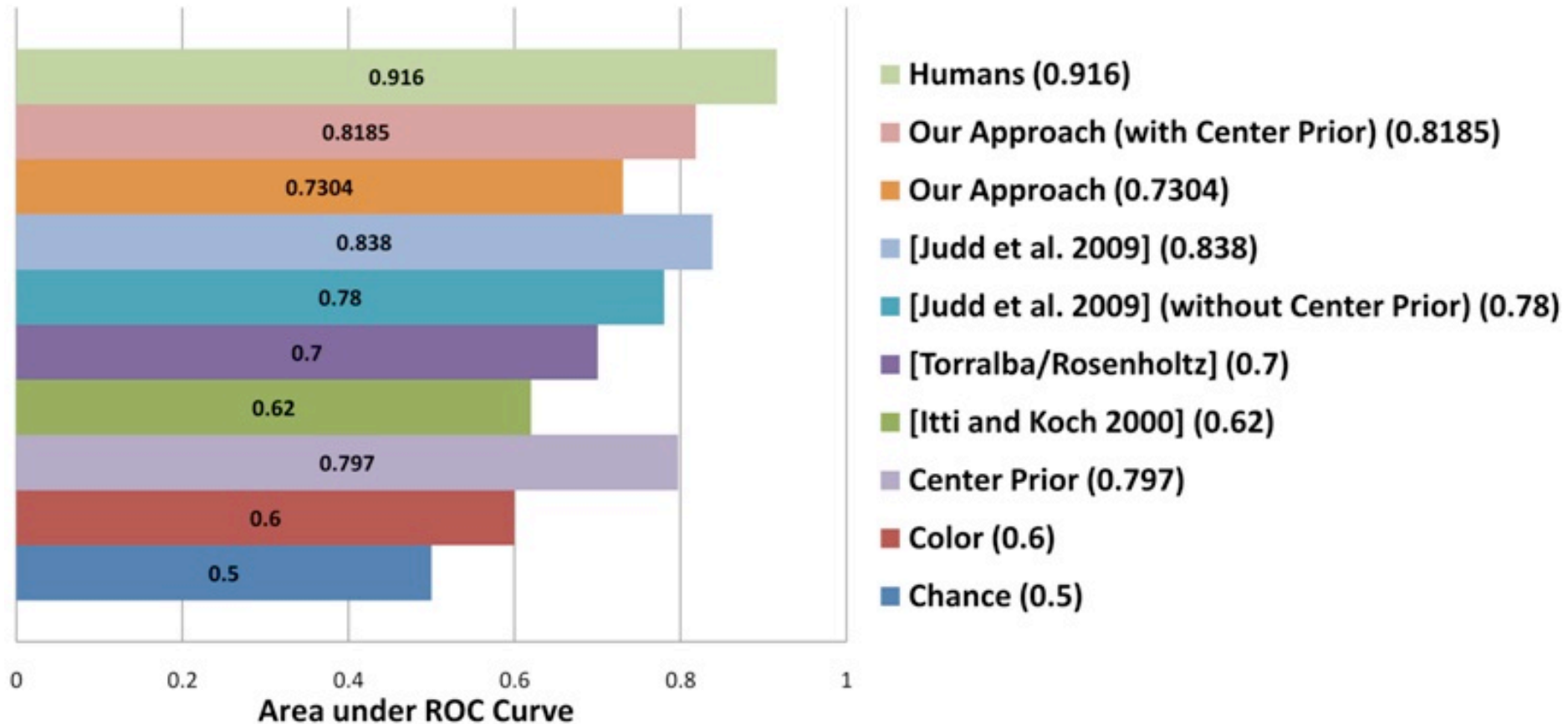


PROXY FOR SALIENCY



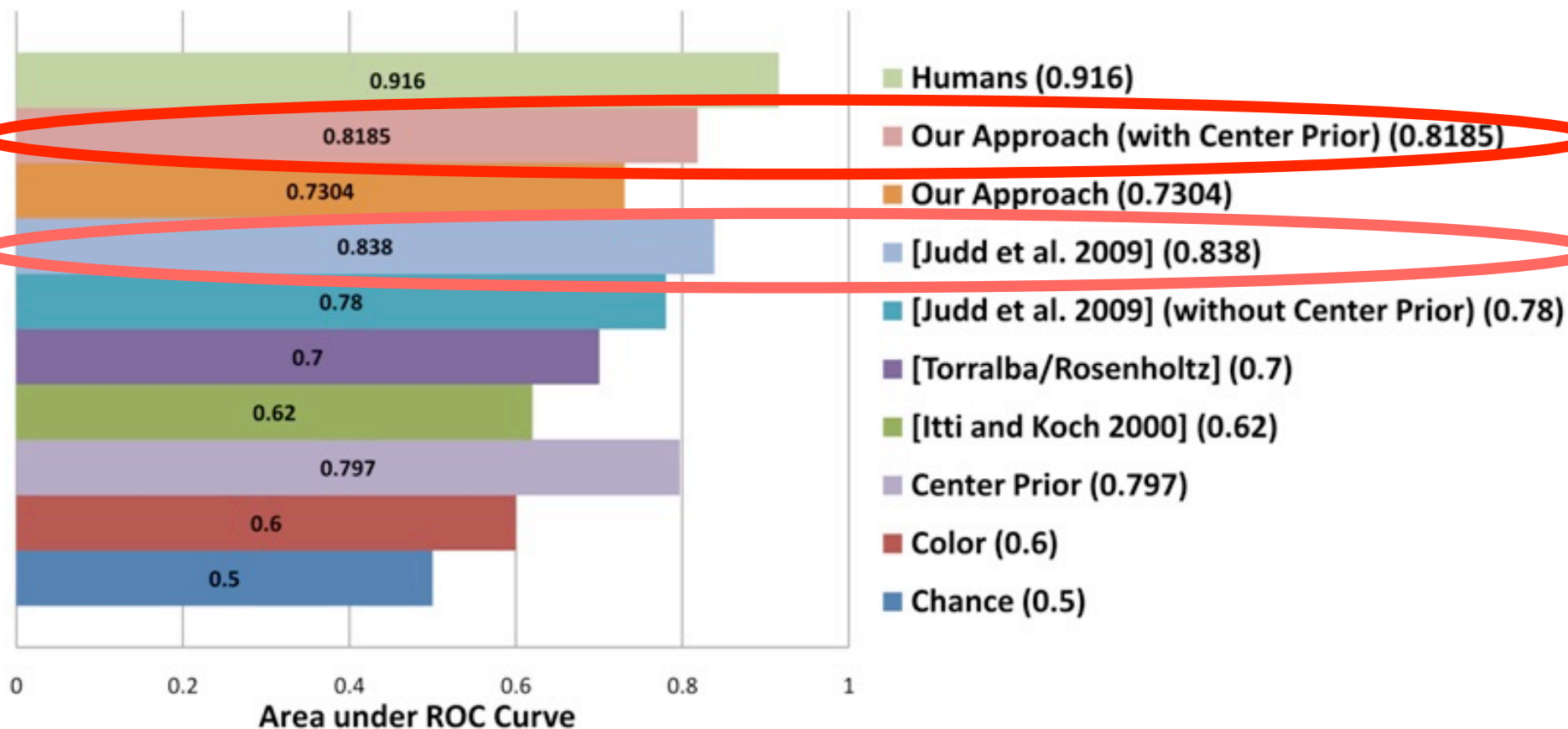
PREDICTING SALIENCY

SALIENCY DATASET [Judd et al., 2009]



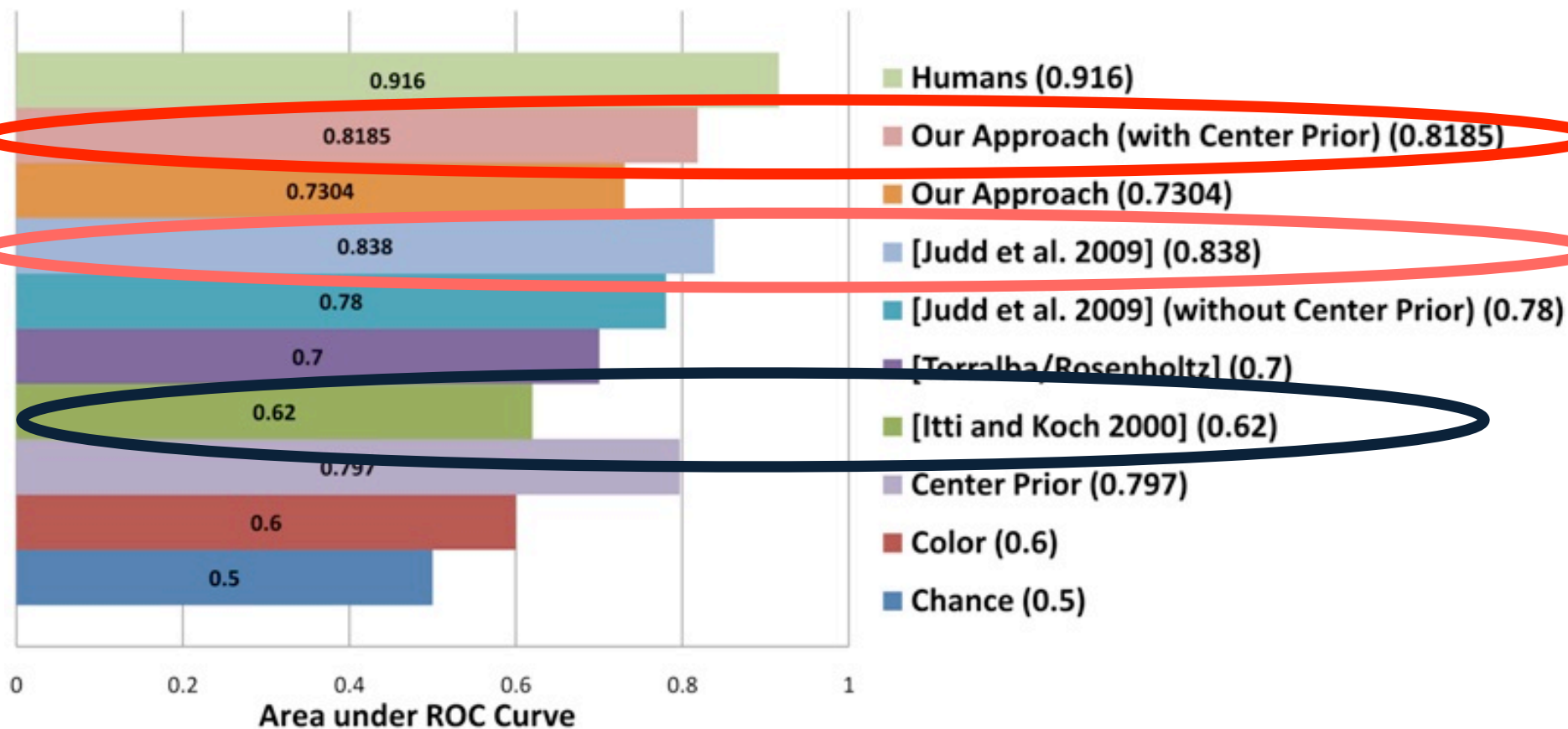
PREDICTING SALIENCY

SALIENCY DATASET [Judd et al., 2009]



PREDICTING SALIENCY

SALIENCY DATASET [Judd et al., 2009]



WHERE DOES IT FAIL?



Top Matches

APPLICATIONS

RE-PHOTOGRAPHY



**Historical Image of
Boston Station**

RE-PHOTOGRAPHY



**Historical Image of
Boston Station**



Re-photographed Image

RE-PHOTOGRAPHY

Computational Re-photography
(Bae et al., 2010)



Historical Image of
Boston Station



Re-photographed Image

RE-PHOTOGRAPHY

Computational Re-photography (Bae et al., 2010)



Historical Image of
Boston Station



Re-photographed Image



Then & Now View

INTERNET RE-PHOTOGRAPHY

Computational Re-photography (Bae et al., 2010)



Historical Image of
Boston Station



Re-photographed Image



Then & Now View



Historical Image of
Boston Station

INTERNET RE-PHOTOGRAPHY

Computational Re-photography (Bae et al., 2010)



Historical Image of
Boston Station



Re-photographed Image



Then & Now View



Historical Image of
Boston Station

Search
10,000 Flickr Images
of Boston

INTERNET RE-PHOTOGRAPHY

Computational Re-photography (Bae et al., 2010)



Historical Image of Boston Station



Re-photographed Image



Then & Now View

Our Approach



Historical Image of Boston Station

Search
10,000 Flickr Images
of Boston



Top Match

INTERNET RE-PHOTOGRAPHY

Computational Re-photography (Bae et al., 2010)



Historical Image of Boston Station



Re-photographed Image



Then & Now View

Our Approach



Historical Image of Boston Station



Top Match
From 10,000 Flickr Images



Then & Now View

INTERNET RE-PHOTOGRAPHY



Paris (1940)

INTERNET RE-PHOTOGRAPHY



Paris (1940)



Top Matches

INTERNET RE-PHOTOGRAPHY

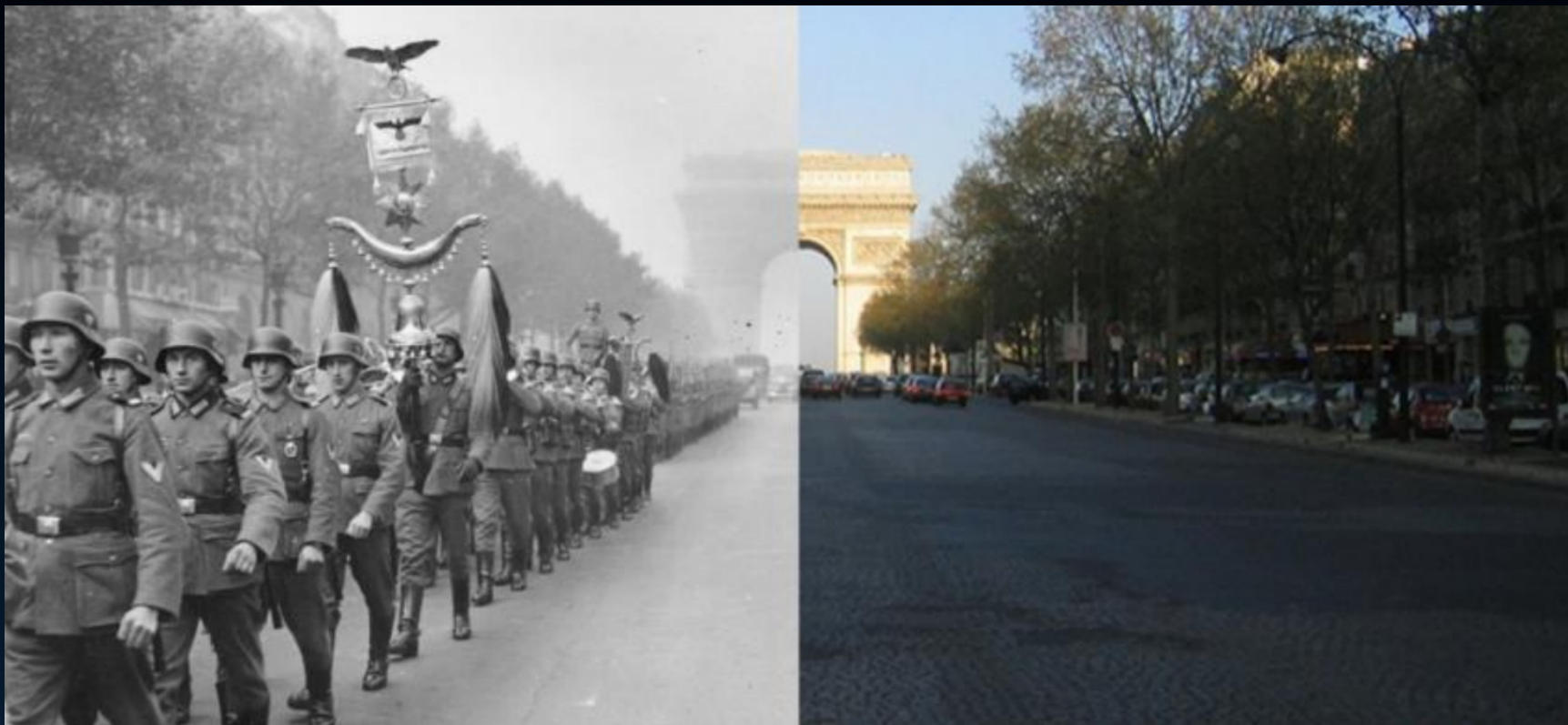


Paris (1940)



Top Matches

INTERNET RE-PHOTOGRAPHY



WHERE WAS THE PAINTER STANDING?

Input Painting



PAINTING2GPS

Input Painting



Retrieval set

10,000 Geo-tagged Flickr Images

100 top matches used to estimation

PAINTING2GPS

Input Painting



Estimated Geo-location



Estimated using 100 top matches

VISUAL SCENE EXPLORATION

VISUAL SCENE EXPLORATION

flickr with YouTube

Sign In or Register Show More Help Sign Out

Home You Organize & Create Contacts Groups Explore Upload

Search

Search

Photos Groups People

Everyone's Uploads

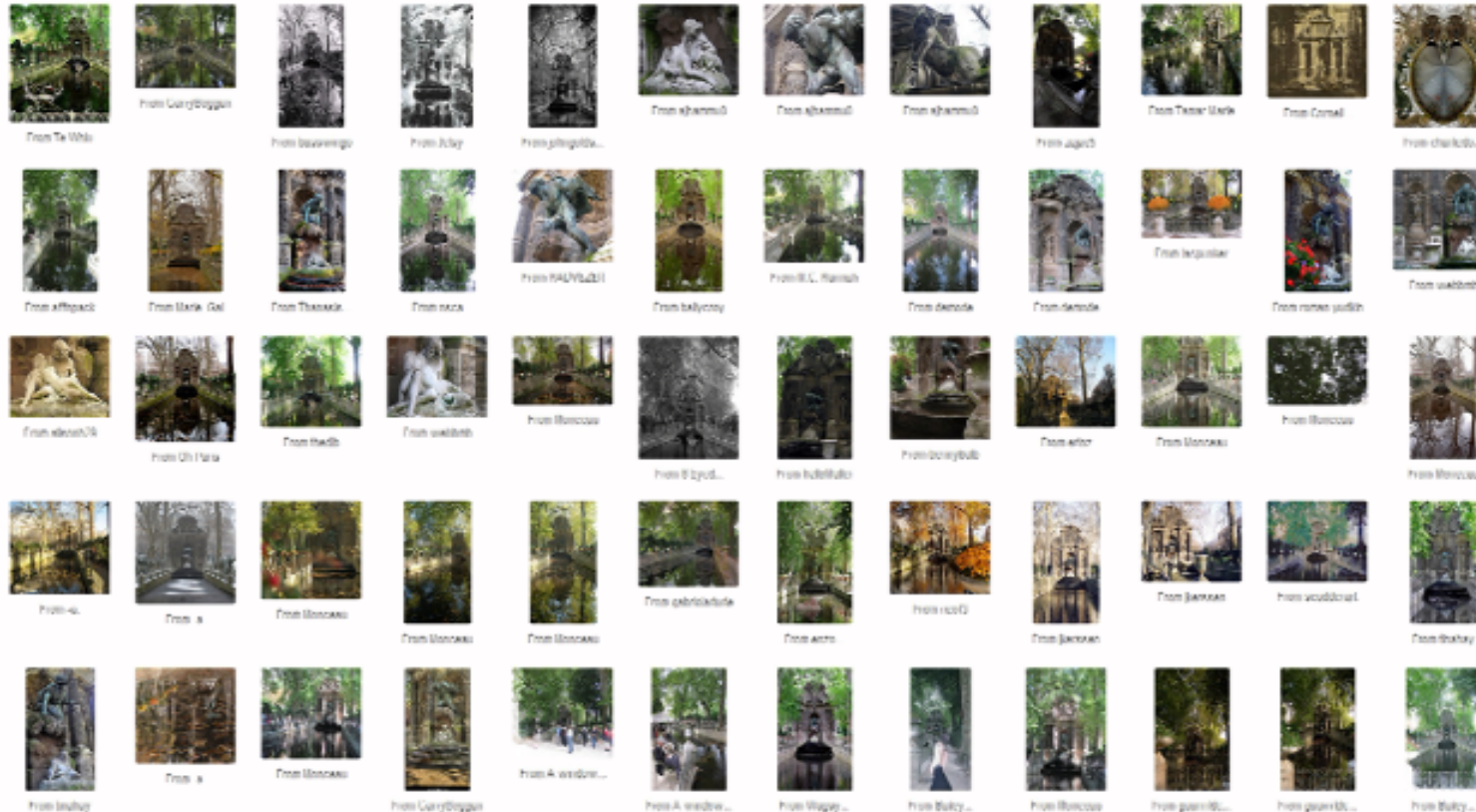
medici fountain

SEARCH

Full Text | Top Only
Advanced Search

Sort: Relevant Recent Interesting

View: Small Medium Detail Slideshow



prev 1 2 3 4 5 6 7 next

(1,773 results)



Want to go up-front? Get That

Groups

more...

- Florence - Florence
3,373 members | 765,770 photos
- Paris
26,024 members | 244,242 photos

Photographers

more...

- Emilio Geertz
Cupertino, United States
See related photos
- wallyg
jersey city, nj, us
See related photos

Tag Clusters

more...

- Photos with tags: florence, italy and florence
- Photos with tags like park, fountain and fountain

Looking to license a photo?

Looking to license an image? You can search through the licensed image of Flickr members enrolled in the Getty Images Program.

Photos

more...

PHOTOSYNTH

[Snavely et al., 2006]

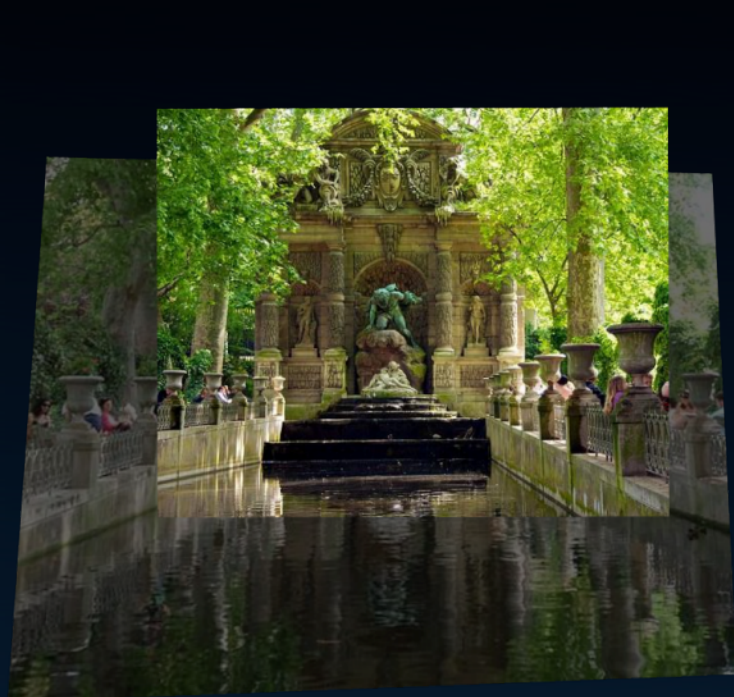
Dataset size: 136 photos (from flickr)

of discovered synths: 14

82 photos not part of any synth

PHOTOSYNTH

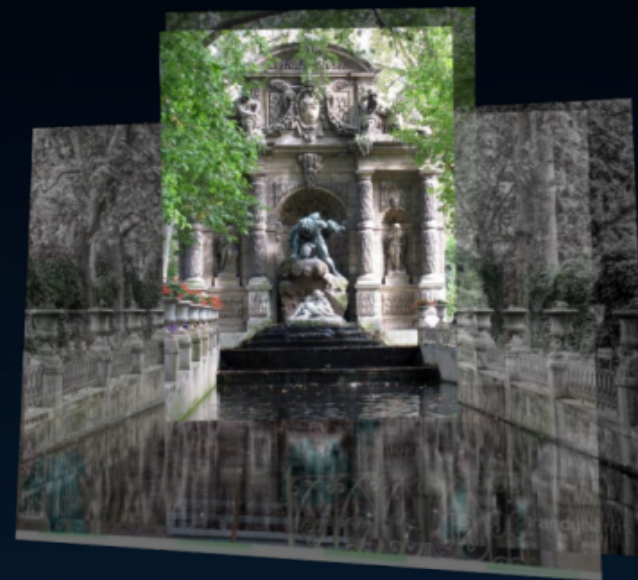
[Snavely et al., 2006]



Synth 1



Synth 2



Synth 3

Dataset size: 136 photos (from flickr)

of discovered synths: 14

82 photos not part of any synth

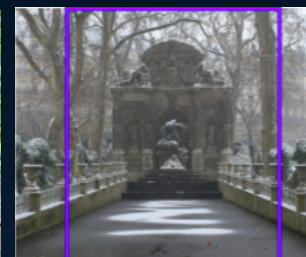
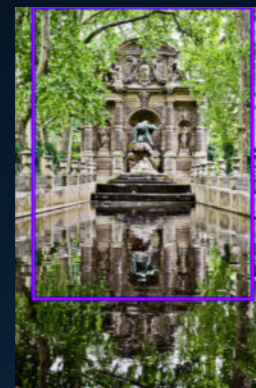
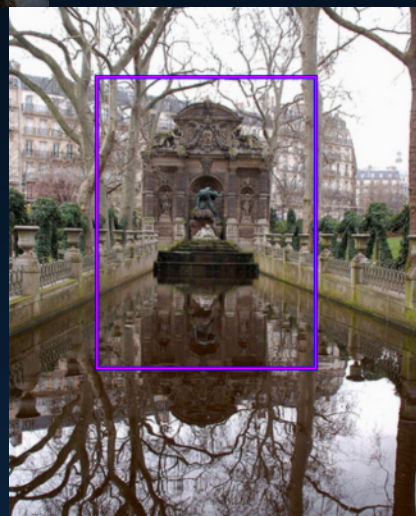
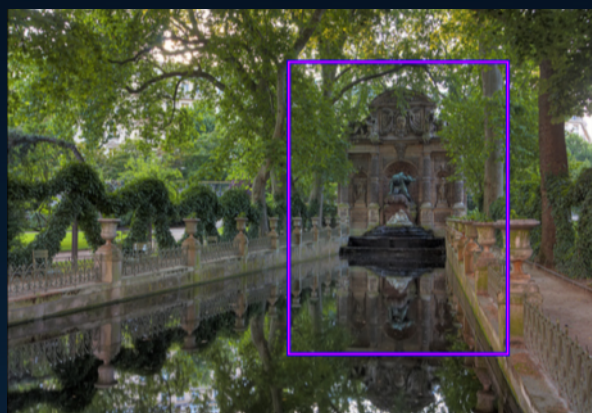
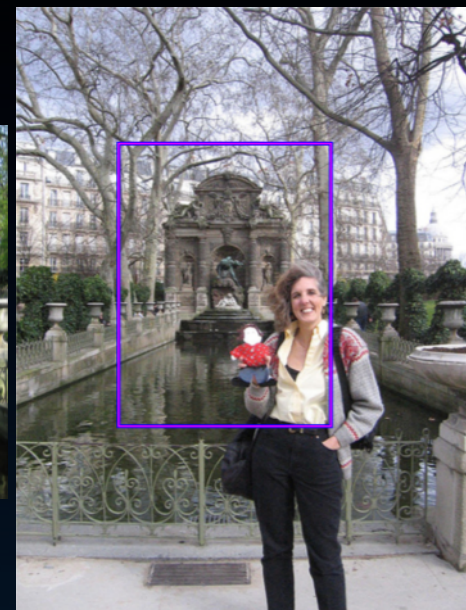
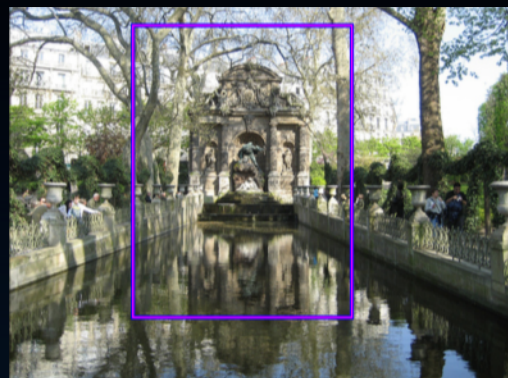
FINDING SIMILAR IMAGES

Query image



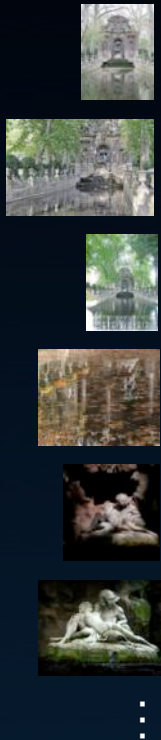
FINDING SIMILAR IMAGES

Query image

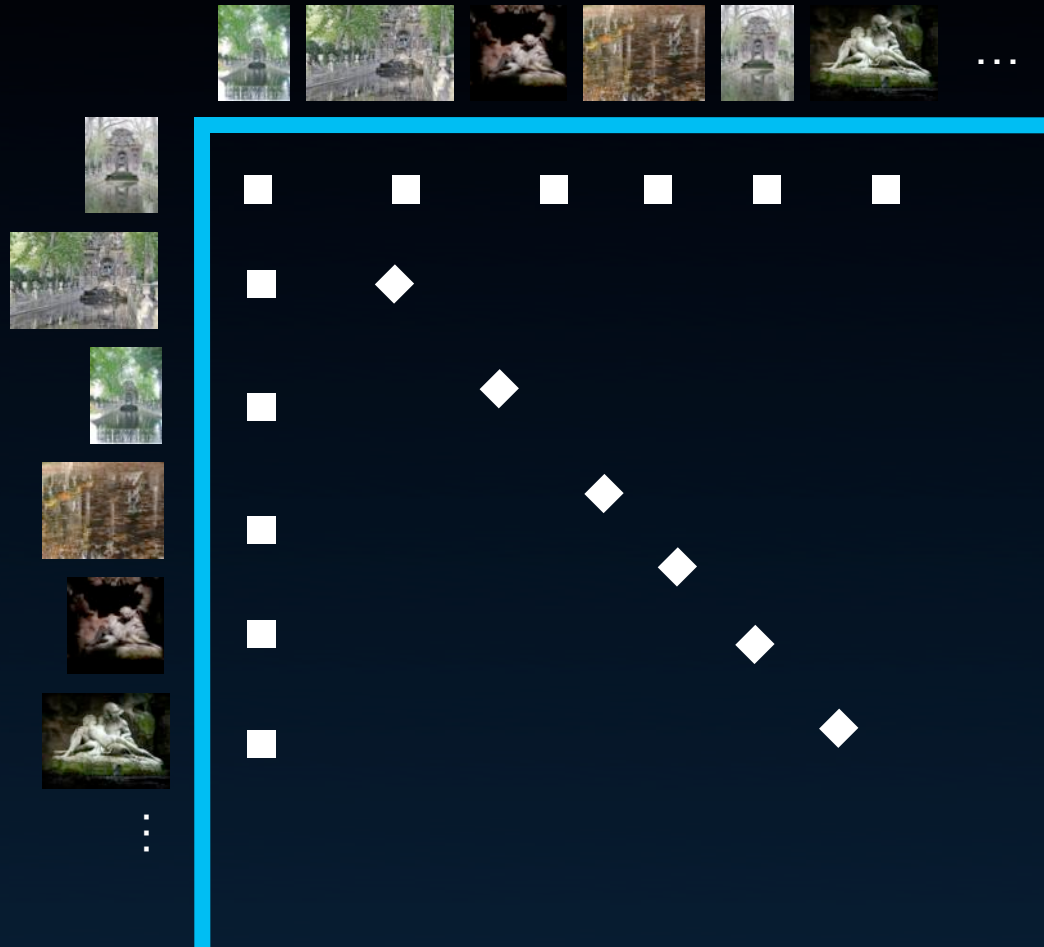


PAIRWISE SIMILARITY MATRIX

PAIRWISE SIMILARITY MATRIX



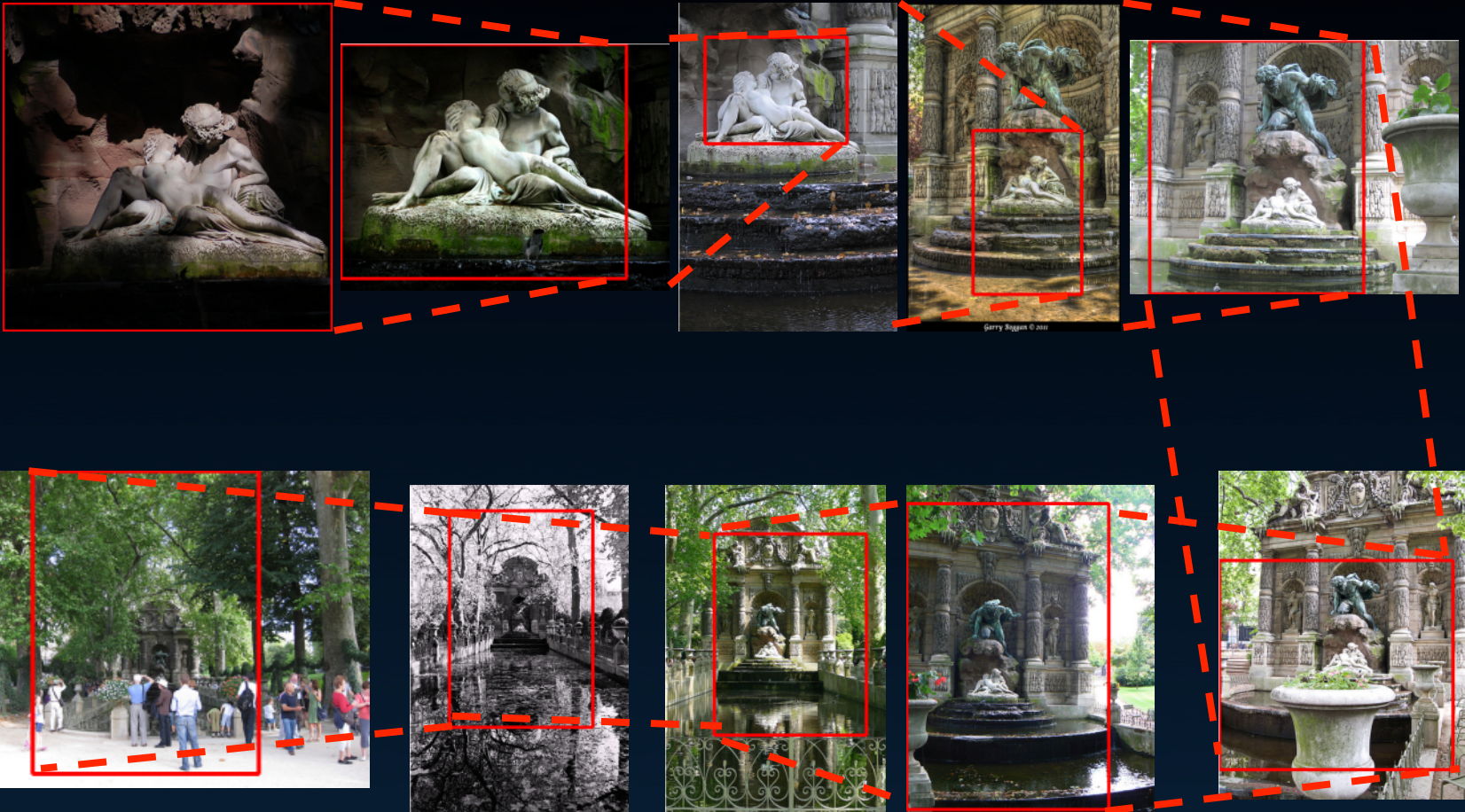
PAIRWISE SIMILARITY MATRIX



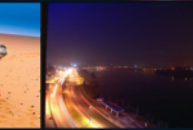
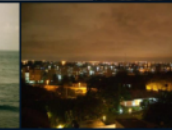
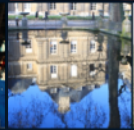
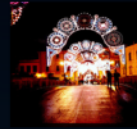
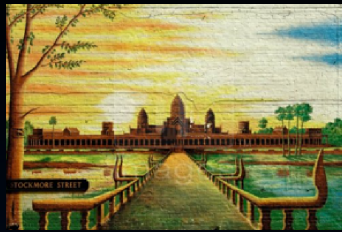
TRAVERSING THE GRAPH



TRAVERSING THE GRAPH

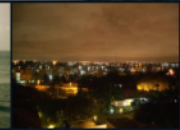
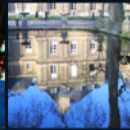
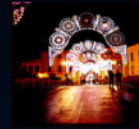
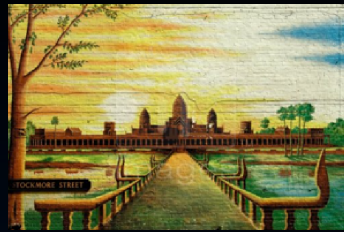


CONCLUSION



- Good News:
 - Results surprisingly nice
- Bad News:
 - Computationally expensive

CONCLUSION



Website:

<http://graphics.cs.cmu.edu/projects/crossDomainMatching/>

Code: <https://github.com/quantombone/exemplarsvm>

THANK YOU!

Abhinav Shrivastava, Tomasz Malisiewicz, Abhinav Gupta, Alexei A. Efros. **Data-driven Visual Similarity for Cross-domain Image Matching.** In SIGGRAPH ASIA, 2011.

Tomasz Malisiewicz, Abhinav Gupta, Alexei A. Efros. **Ensemble of Exemplar-SVMs for Object Detection and Beyond.** In ICCV, 2011.

THANK YOU!

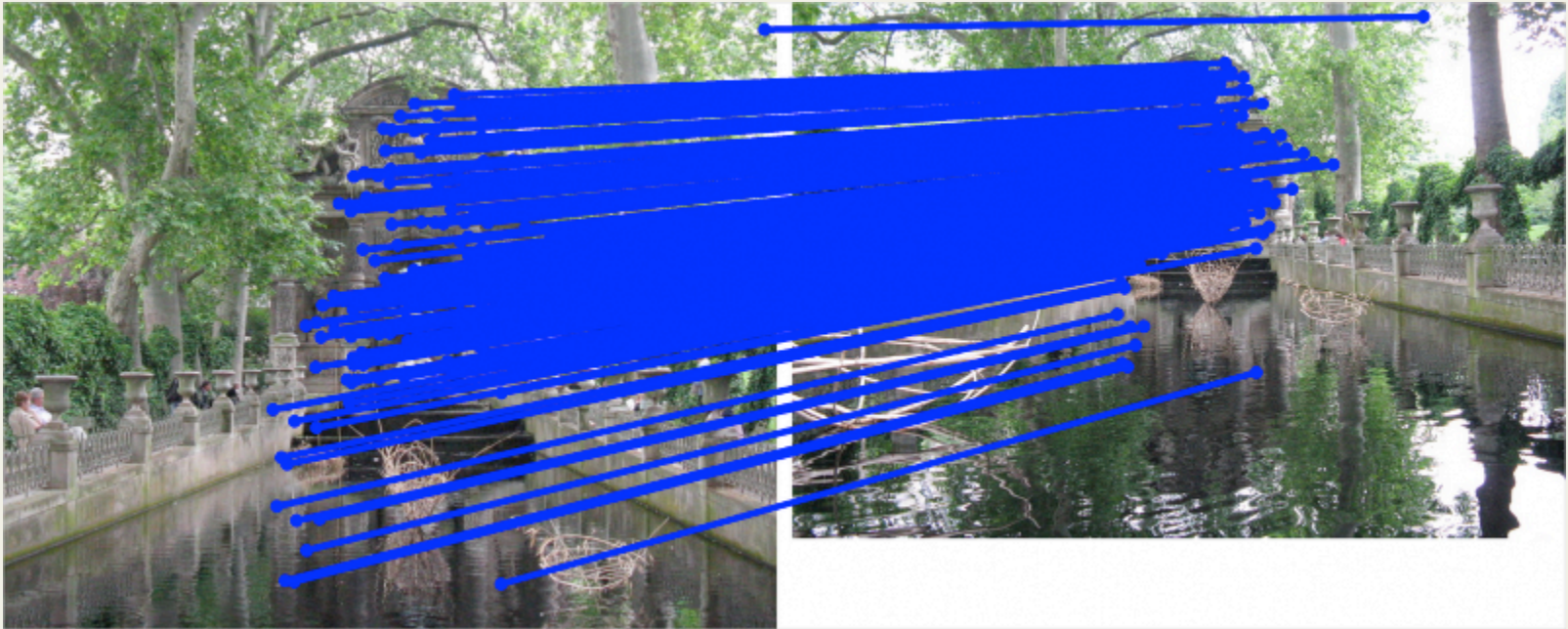


Abhinav Shrivastava, Tomasz Malisiewicz, Abhinav Gupta, Alexei A. Efros. **Data-driven Visual Similarity for Cross-domain Image Matching.** In SIGGRAPH ASIA, 2011.

Tomasz Malisiewicz, Abhinav Gupta, Alexei A. Efros. **Ensemble of Exemplar-SVMs for Object Detection and Beyond.** In ICCV, 2011.

BACKUP SLIDES...

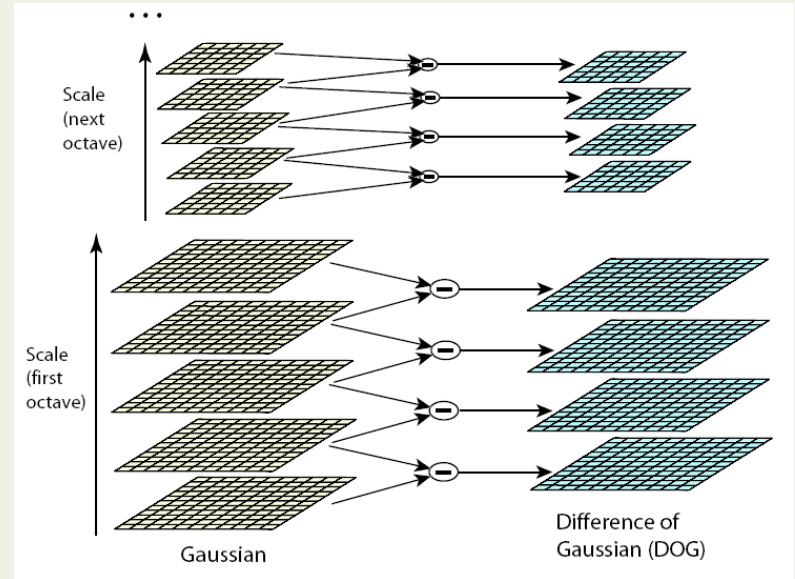
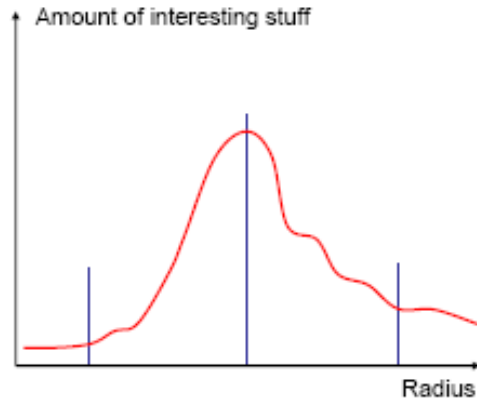
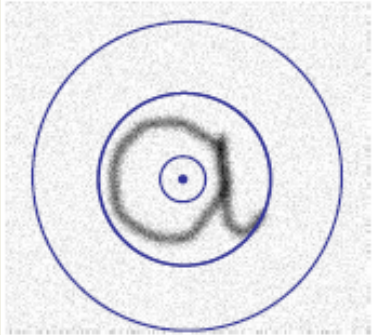
ALL SIFT MATCHES



ALL SIFT MATCHES



SIFT



Find extrema
in 3D DoG space

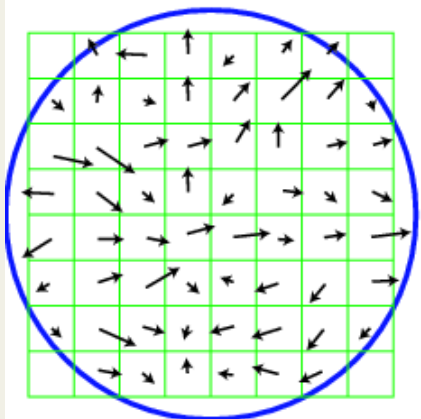
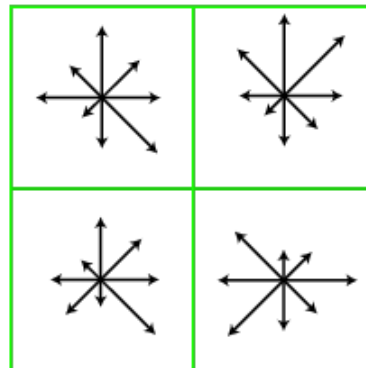
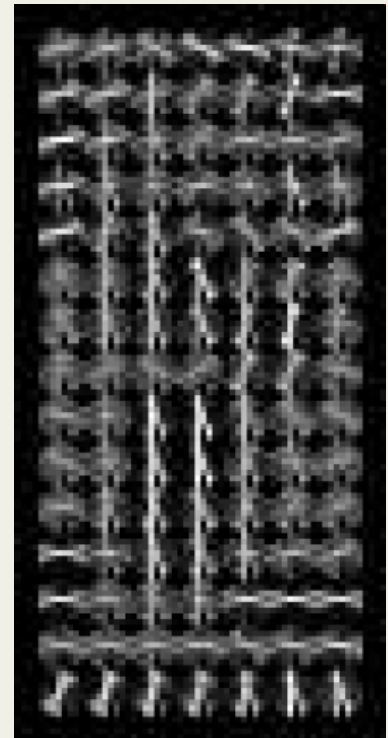
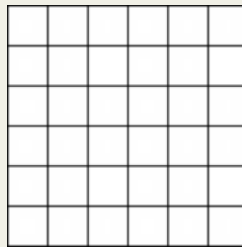
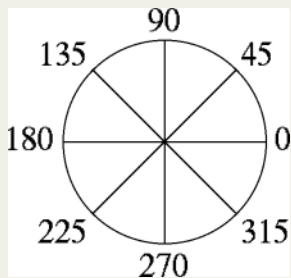
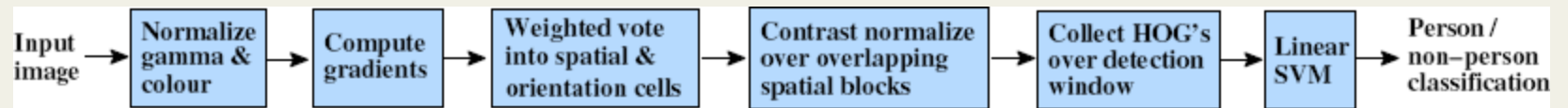


Image gradients

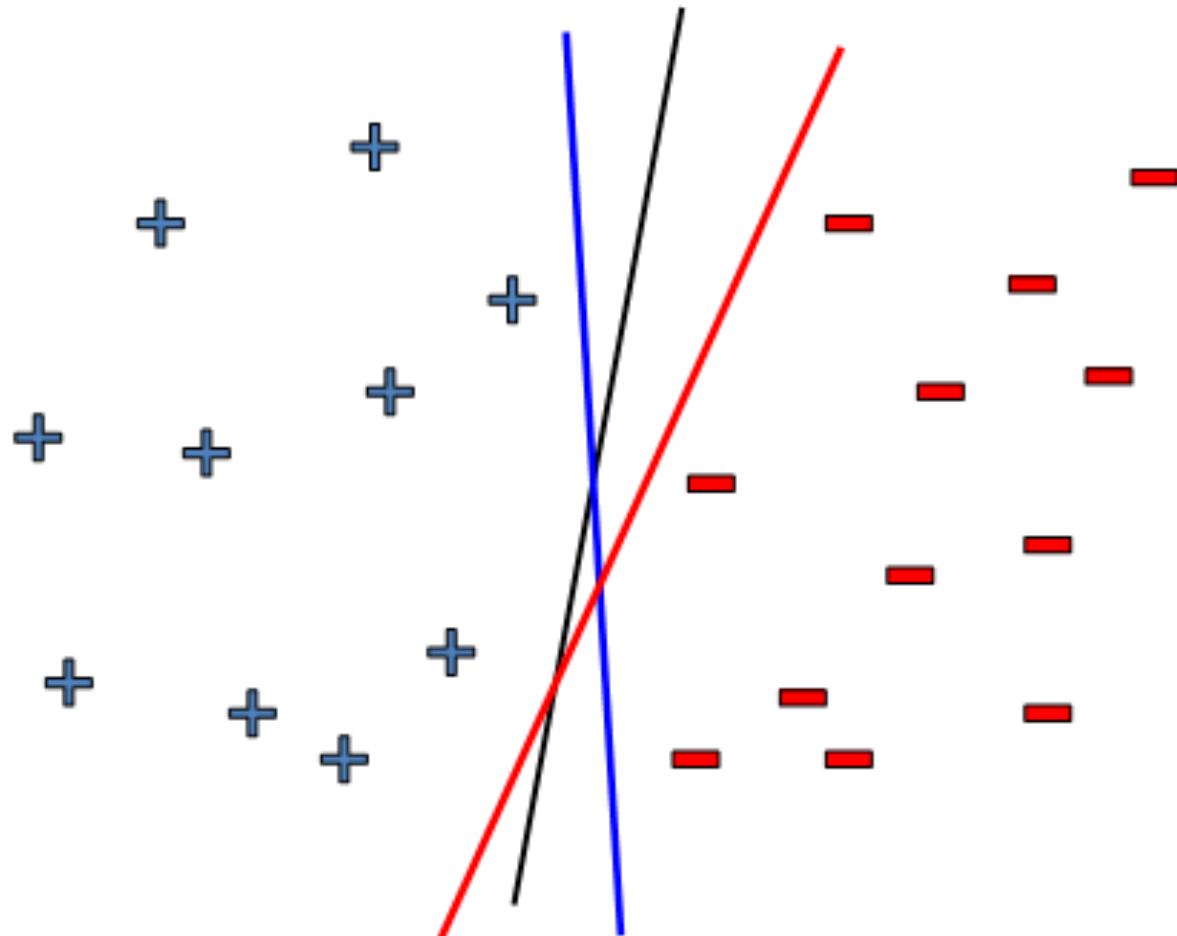


Keypoint descriptor

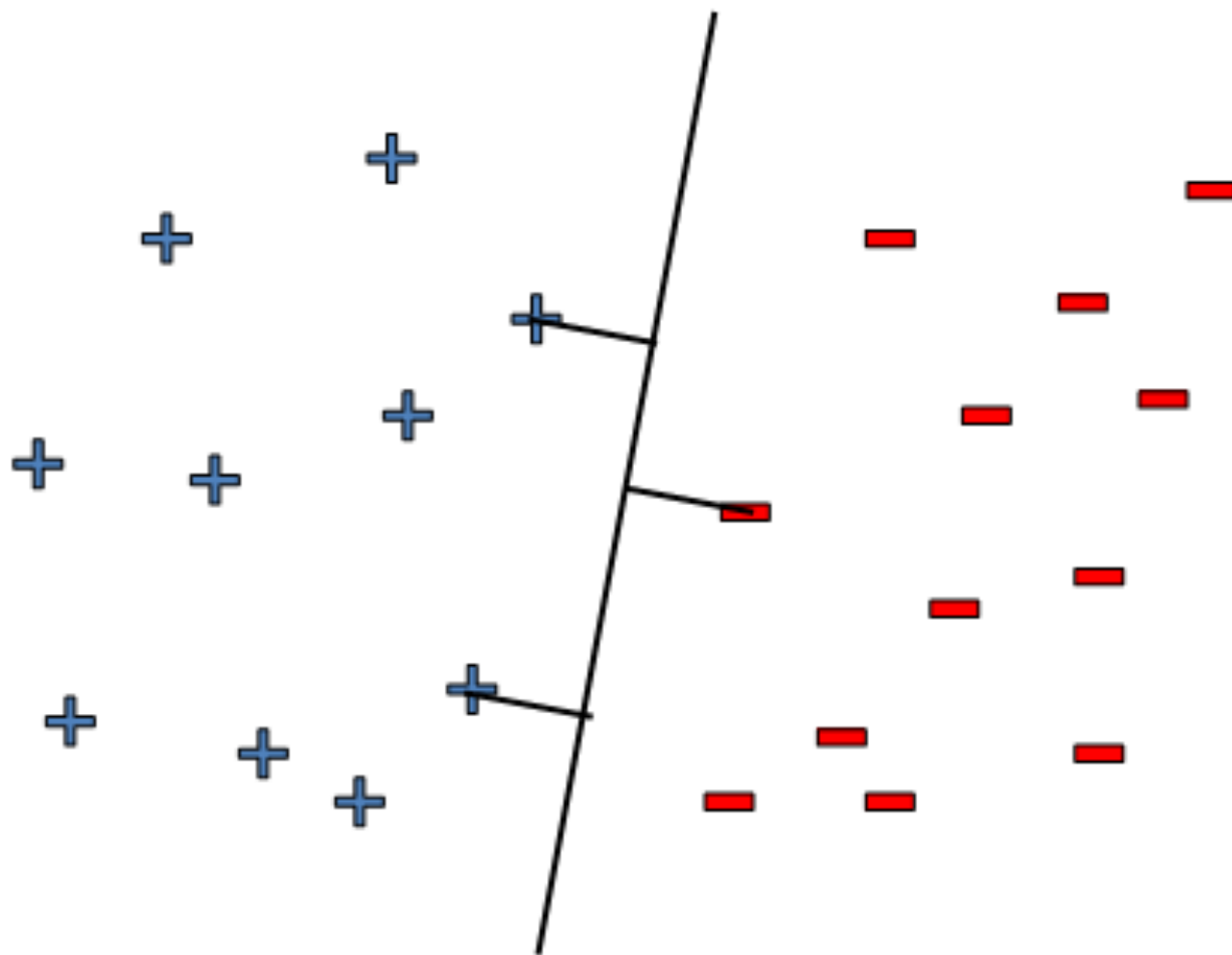
SIFT



Linear classifiers – which line is better?



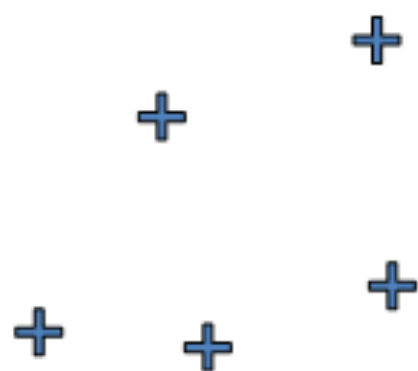
Pick the one with the largest margin!



Parameterizing the decision boundary

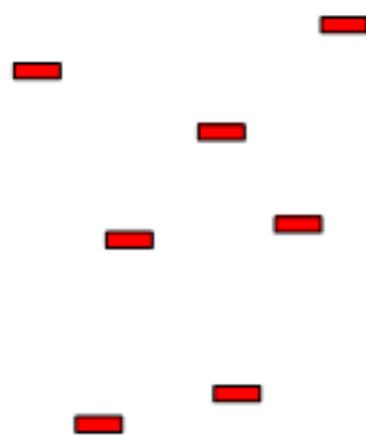
$$\mathbf{w} \cdot \mathbf{x} = \sum_j w^{(j)} x^{(j)}$$

$$\mathbf{w} \cdot \mathbf{x} + b > 0$$



$$\mathbf{w} \cdot \mathbf{x} + b = 0$$

$$\mathbf{w} \cdot \mathbf{x} + b < 0$$



Example i ($= 1, 2, \dots, n$):

$\langle x_i^{(1)}, \dots, x_i^{(m)} \rangle$ — m features

$y_i \in \{-1, +1\}$ — class

Data:

$\langle x_1^{(1)}, \dots, x_1^{(m)}, y_1 \rangle$

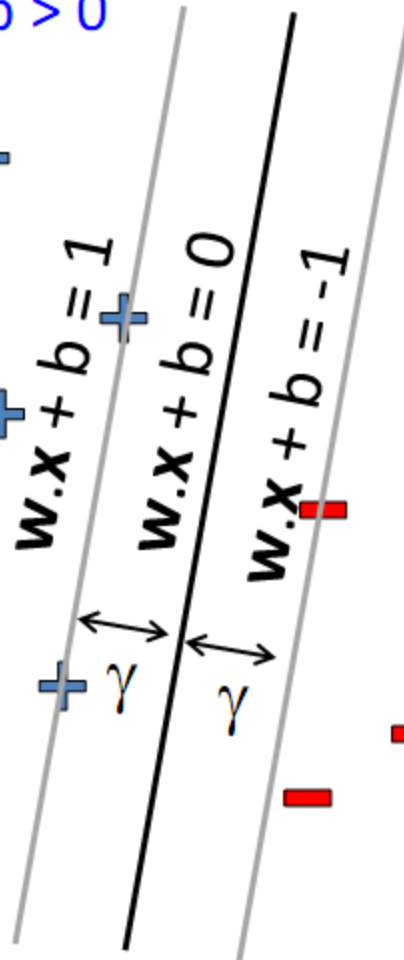
\vdots

$\langle x_n^{(1)}, \dots, x_n^{(m)}, y_n \rangle$

Support Vector Machines

$$w \cdot x + b > 0$$

$$w \cdot x + b < 0$$



$$\min_{w,b} w \cdot w$$

$$\text{s.t. } (w \cdot x_j + b) y_j \geq 1 \quad \forall j$$

Solve efficiently by quadratic programming (QP)

- Well-studied solution algorithms

Linear hyperplane defined by “support vectors”