UNDERSTANDING AND PREDICTING IMAGE MEMORABILITY AT A LARGE SCALE

A. Khosla, A. S. Raju, A. Torralba and A. Oliva

Experiments by: Tyler Folkman

QUICK DEMO

- LaMem Demo

POPULARITY DATA

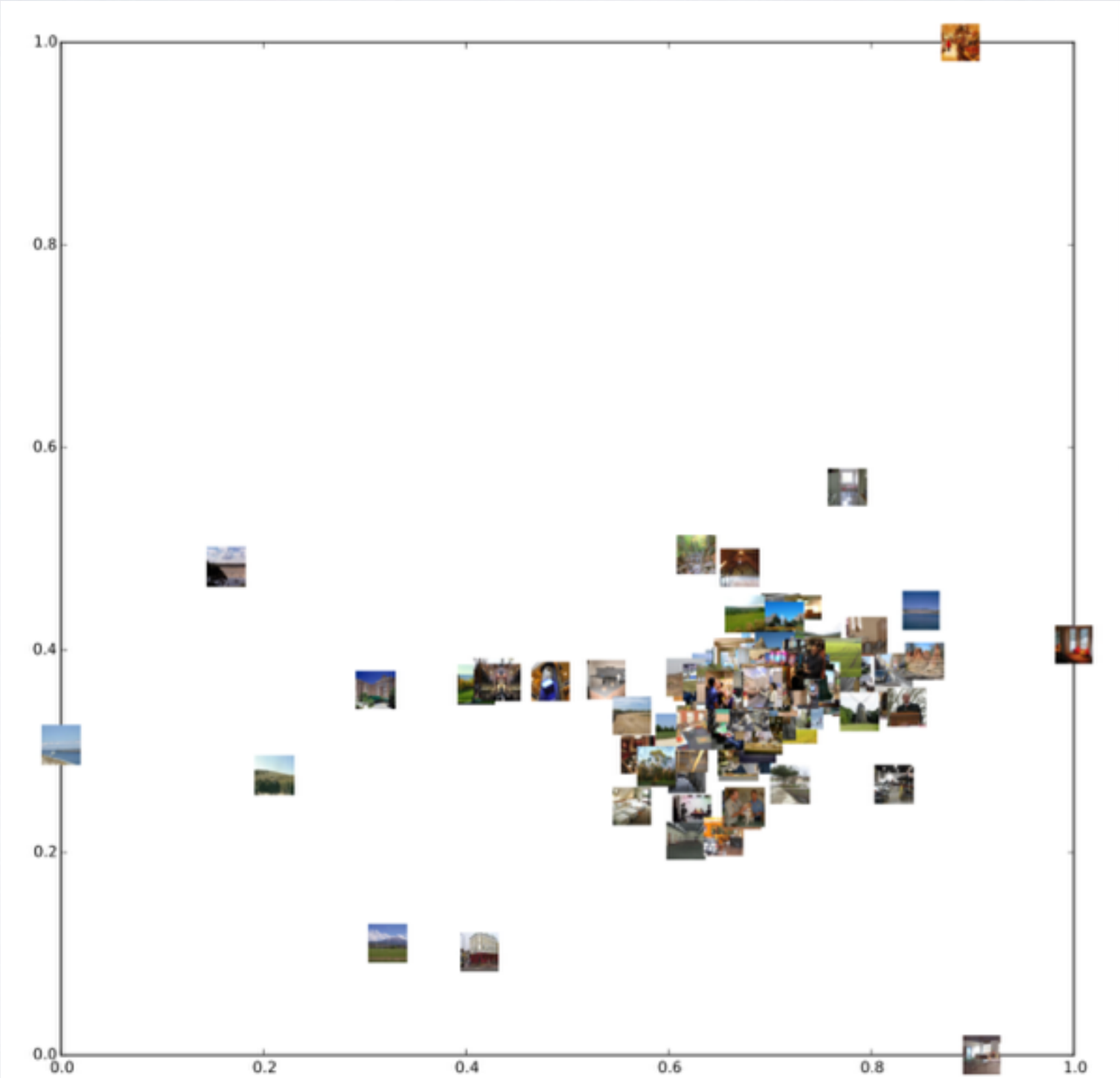
- Random sample of scene categories from SUN dataset.
- Task was to press the space bar whenever they saw an identical repeat of an image at any time in the sequence.
- Memorability score defined as percentage of correct detections.
- 2,222 target images.

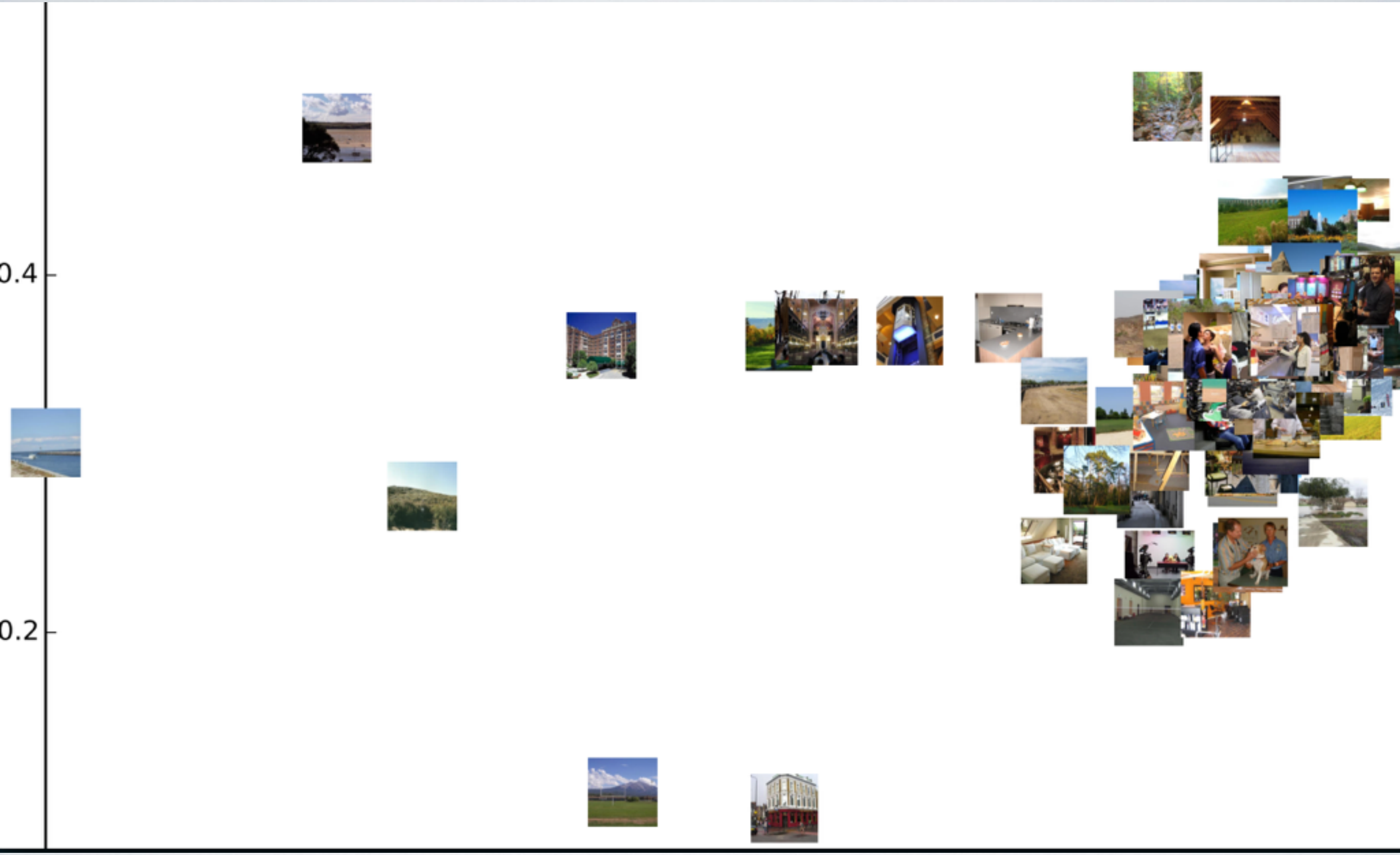
RANK CORRELATION

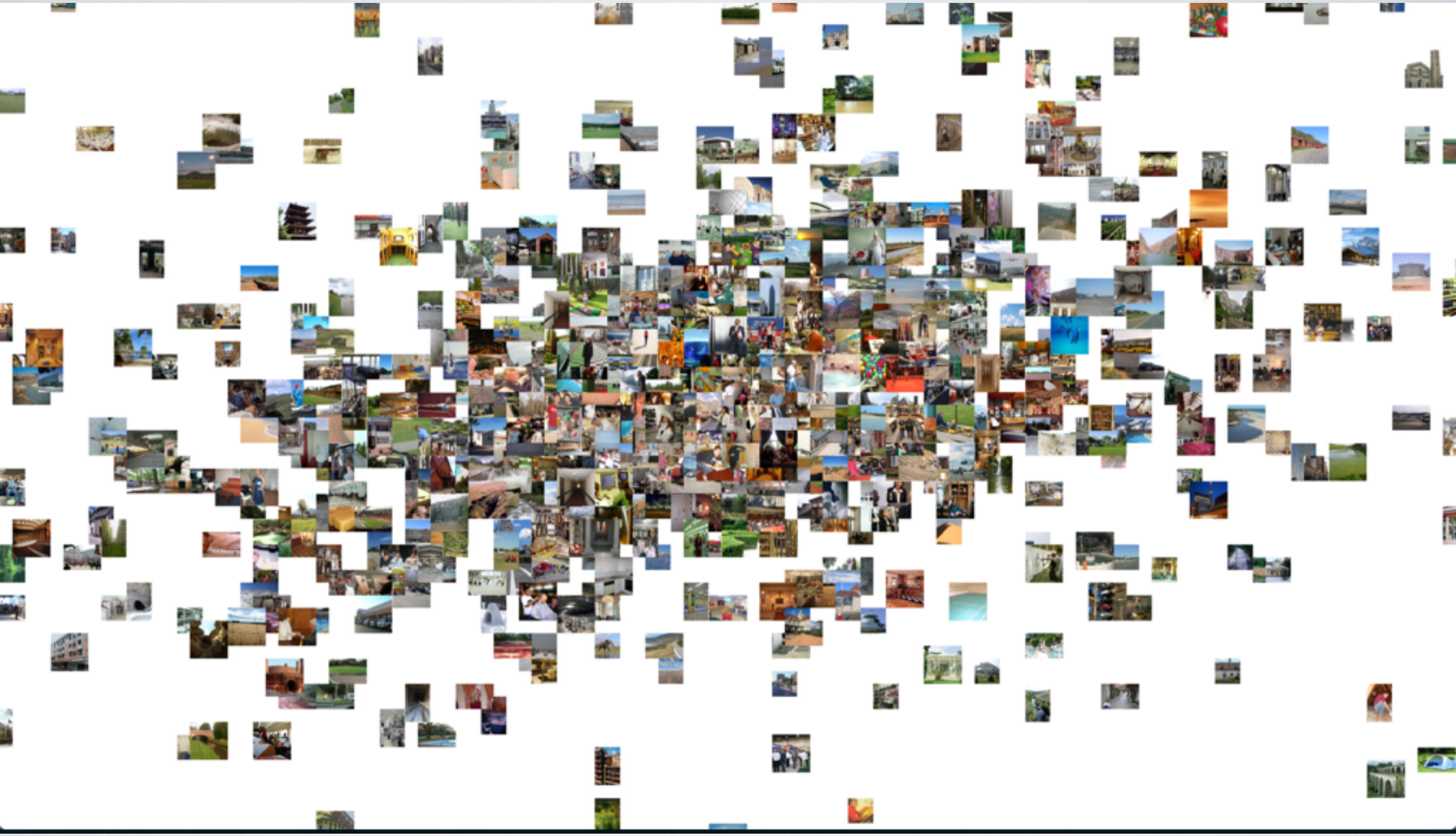
	Human Performance	State of the Art	MemNet
Popularity Data	0.75	0.54*	0.52

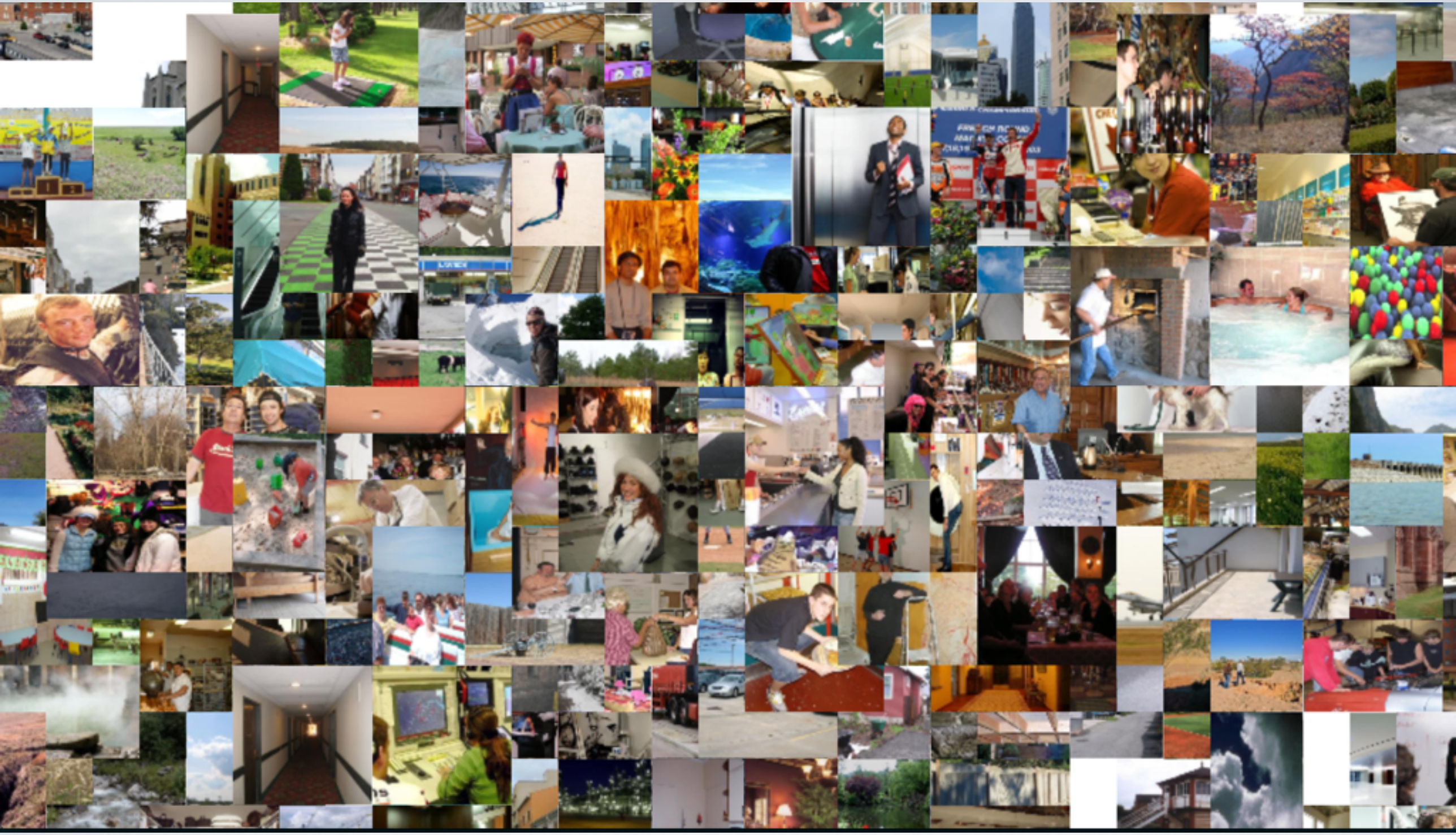
* Isola, P., Xiao, J., Torralba, A., Oliva, A. What makes an image memorable? IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011. Pages 145-152.

T-SNE EMBEDDINGS




0.4
0.2





DO PEOPLE MAKE IMAGES MEMORABLE?

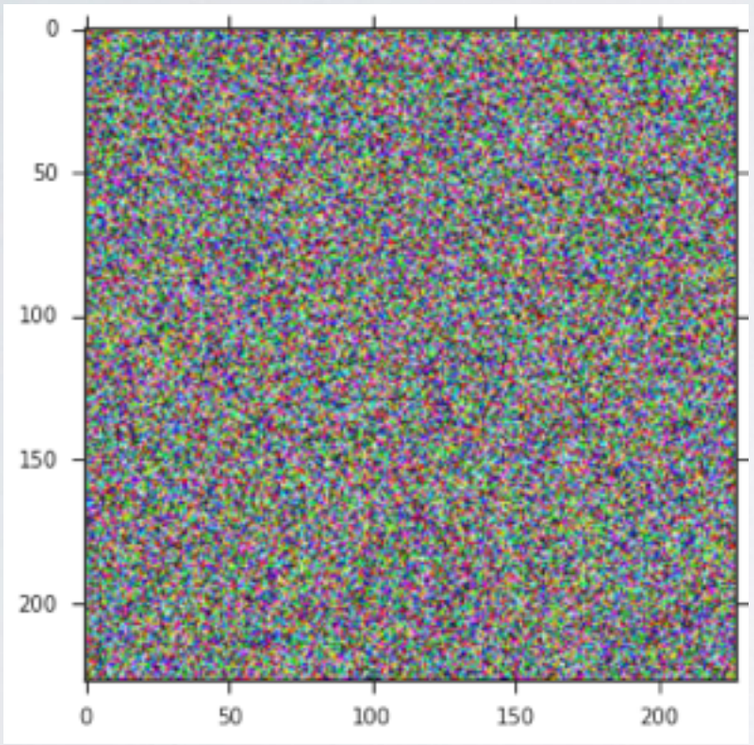
- 17.8% of the data have pedestrians detected in them.

- 2.4% have faces detected.

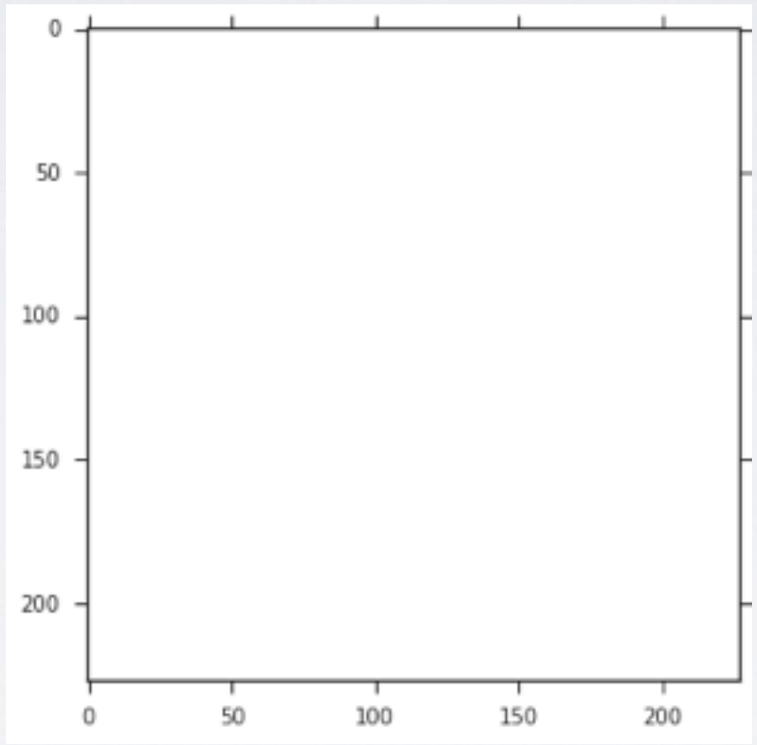
- Pedestrians detected using HOG features and faces using Haar feature-based cascade classifiers.

- What if these people were blurred?

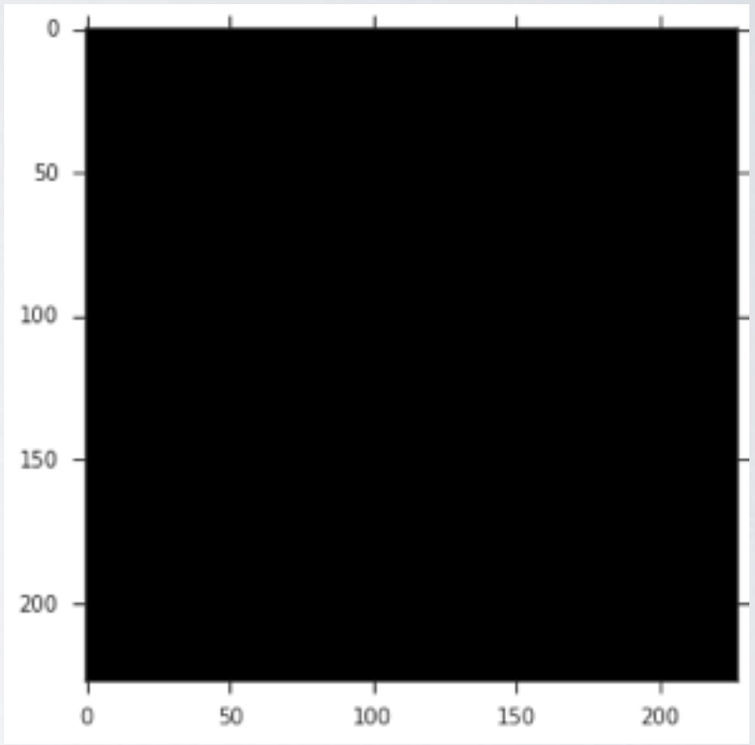
BASELINES



0.73

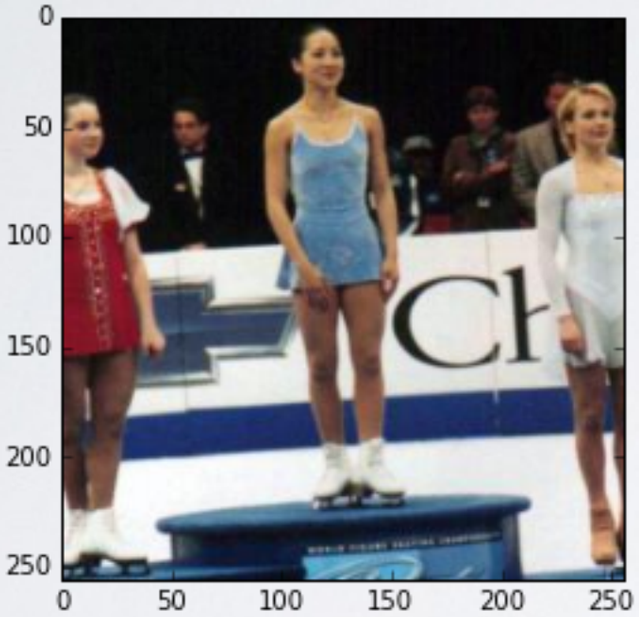


0.73

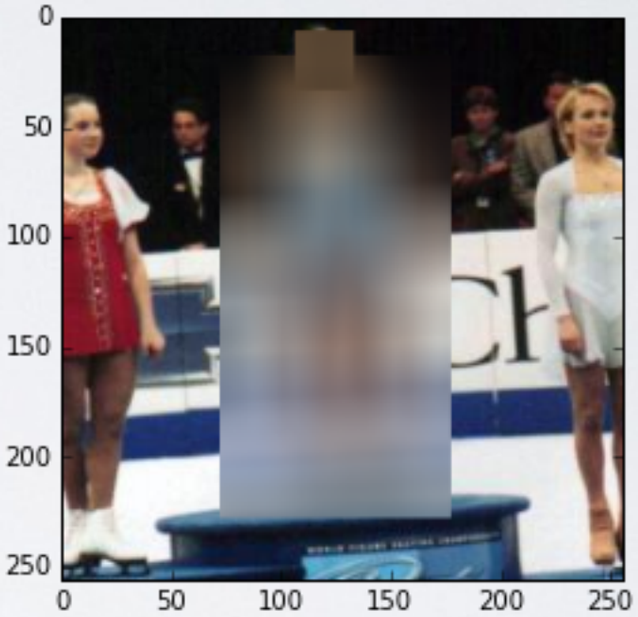


0.72

BLURRING PEOPLE



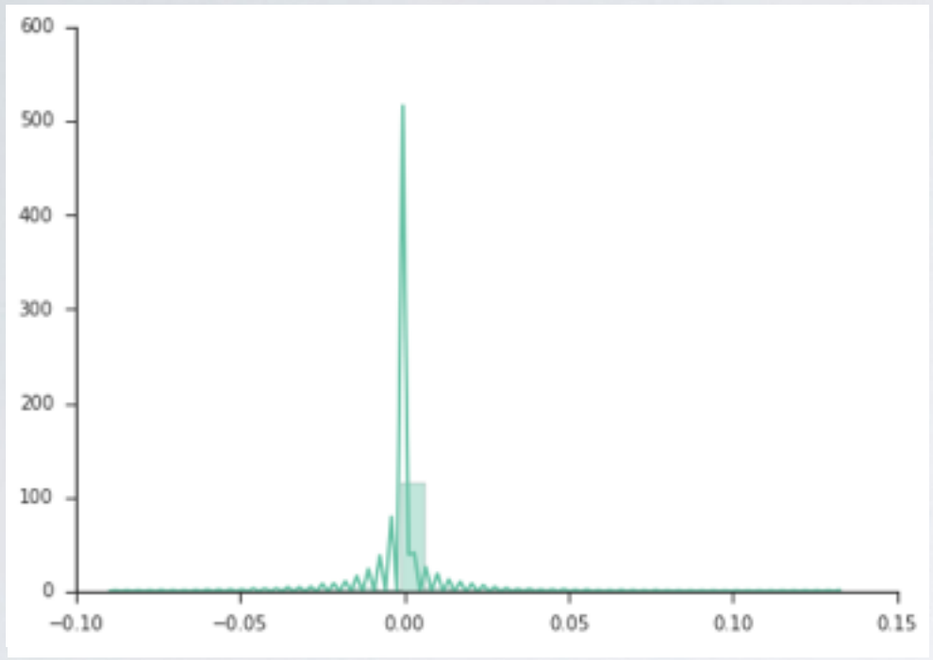
0.80



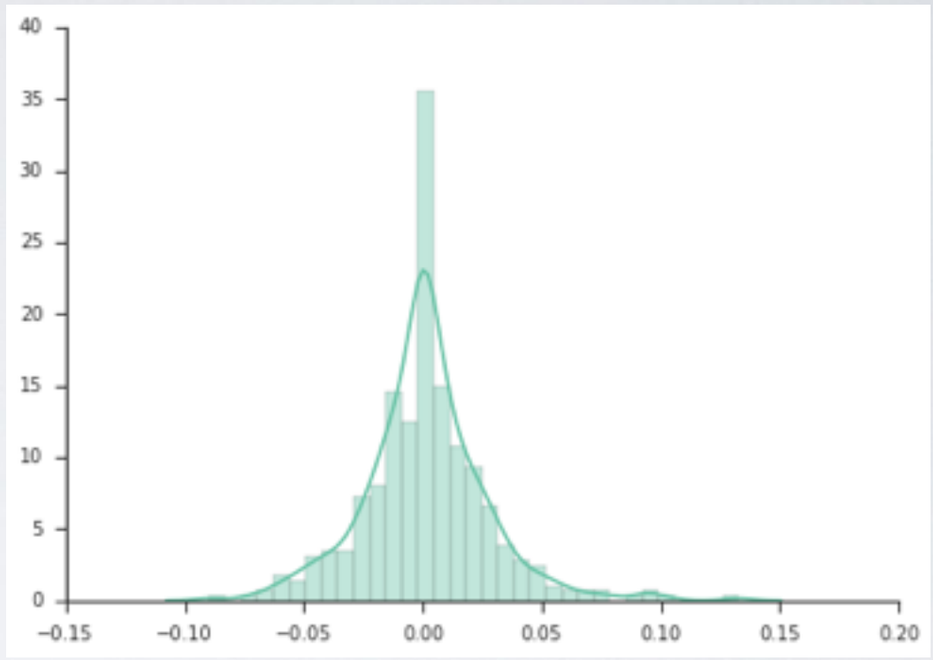
0.70

Actual Memorability: 0.90

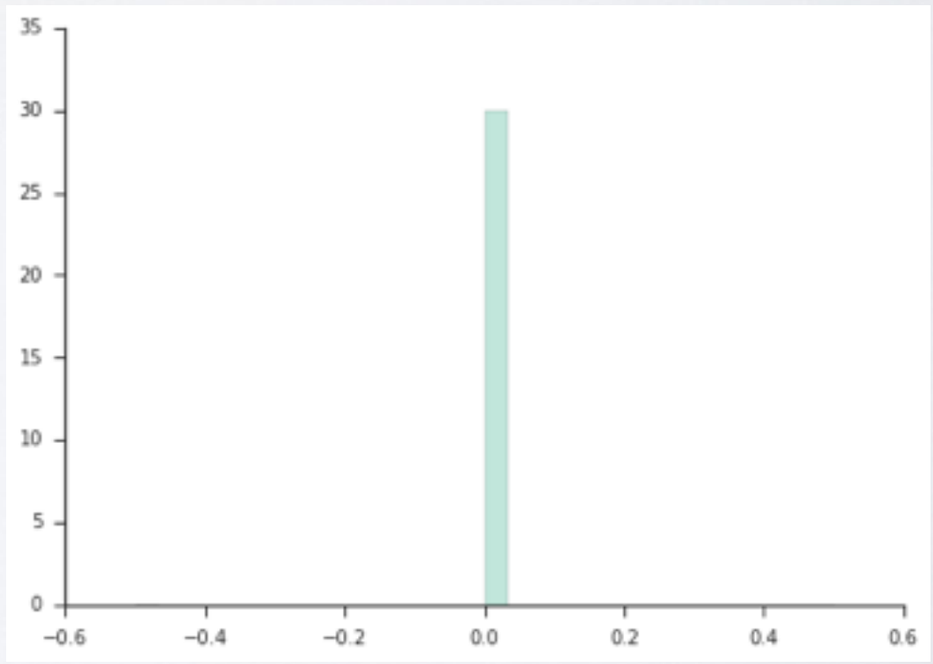
All Images

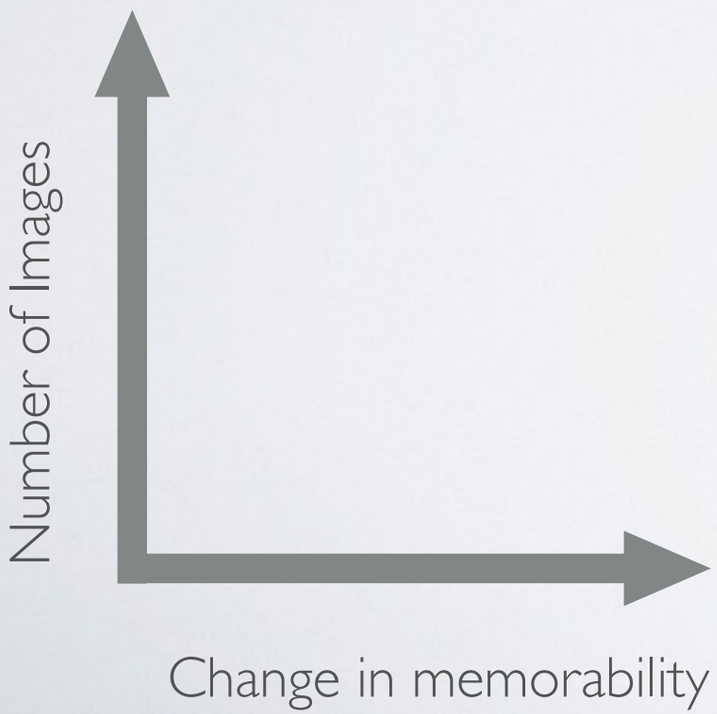


Images with People



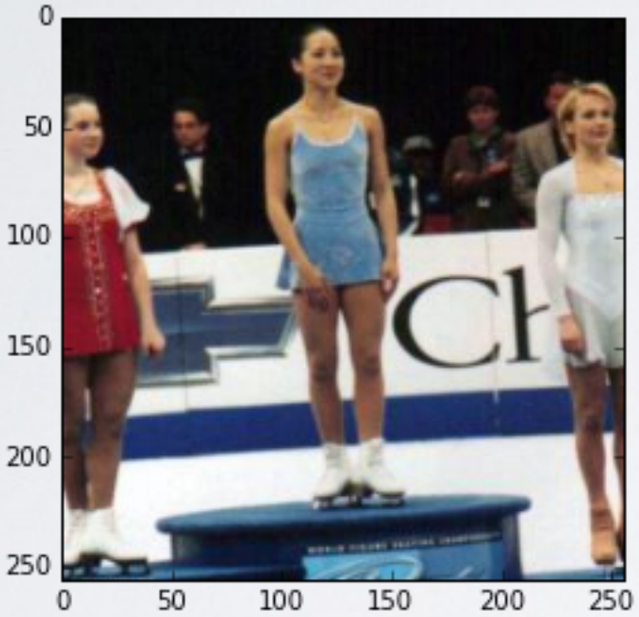
Images without People



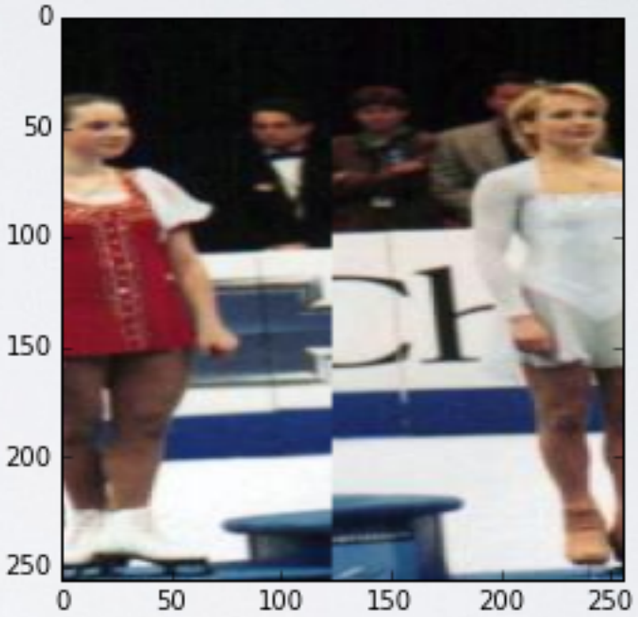


Blurring detected pedestrians or faces doesn't seem to consistently decrease memorability

REMOVING PEOPLE



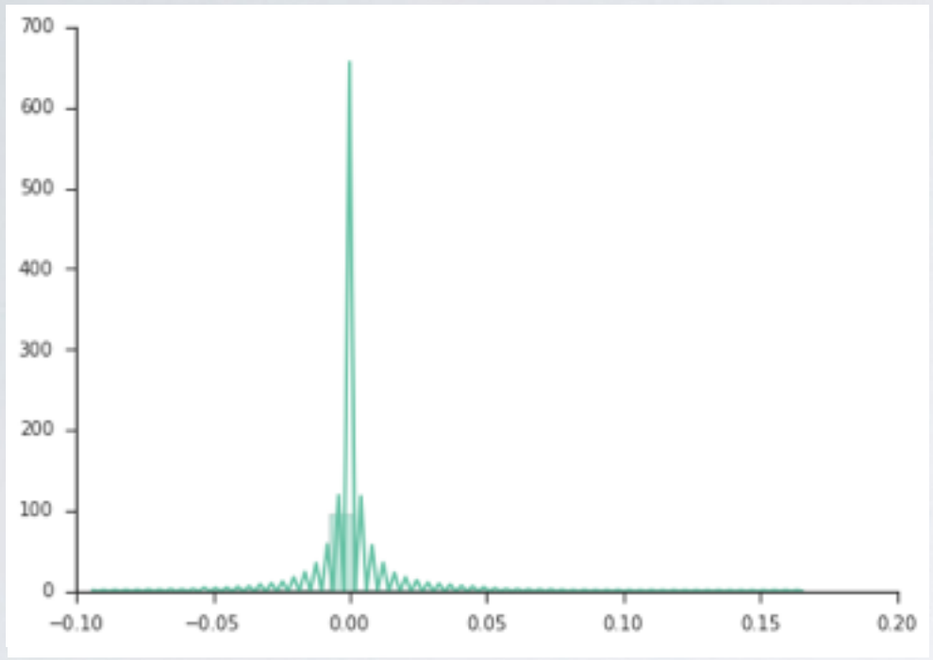
0.80



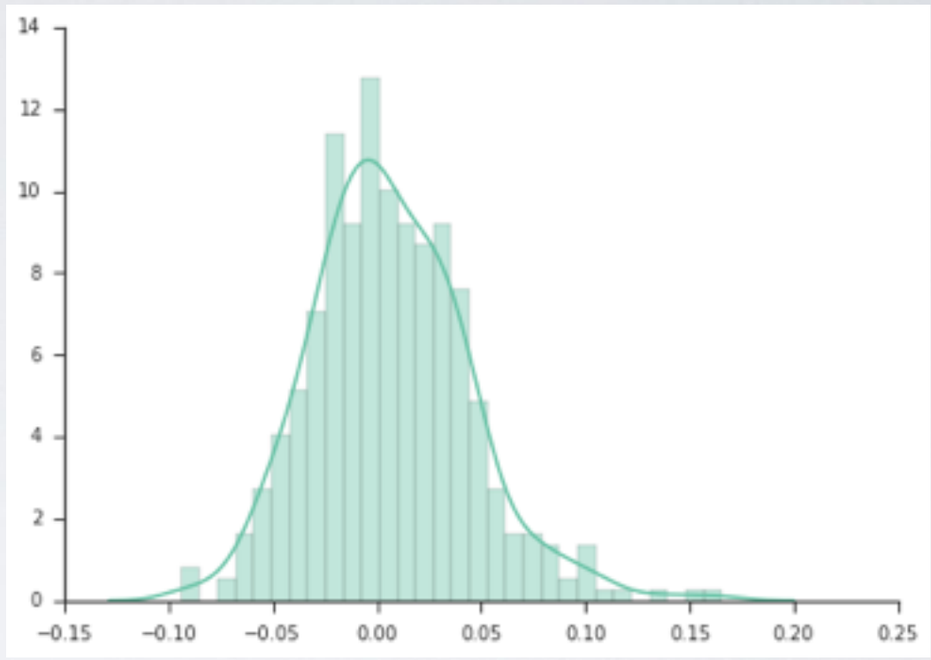
0.75

Actual Memorability: 0.90

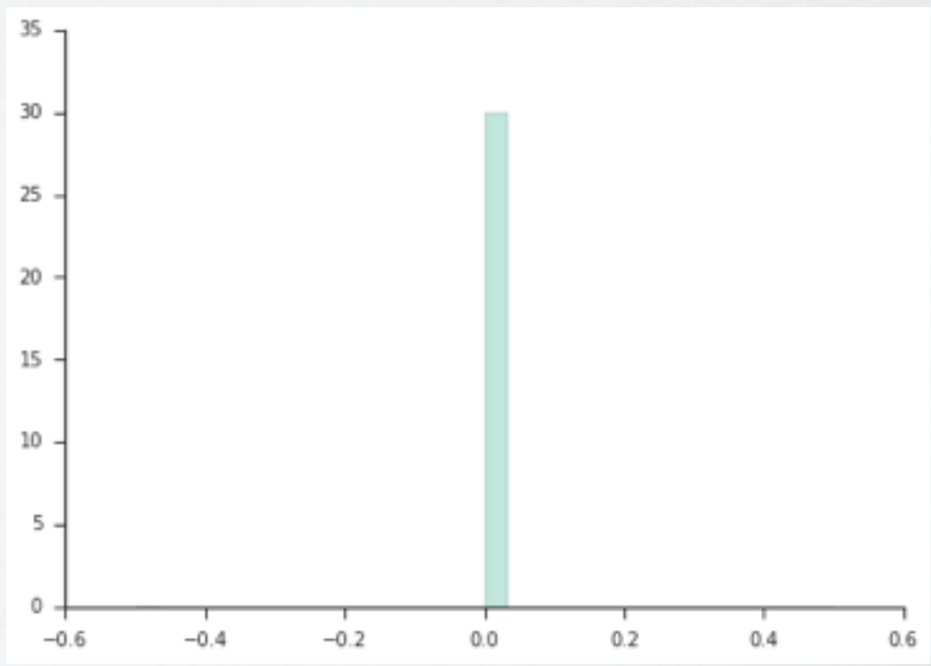
All Images

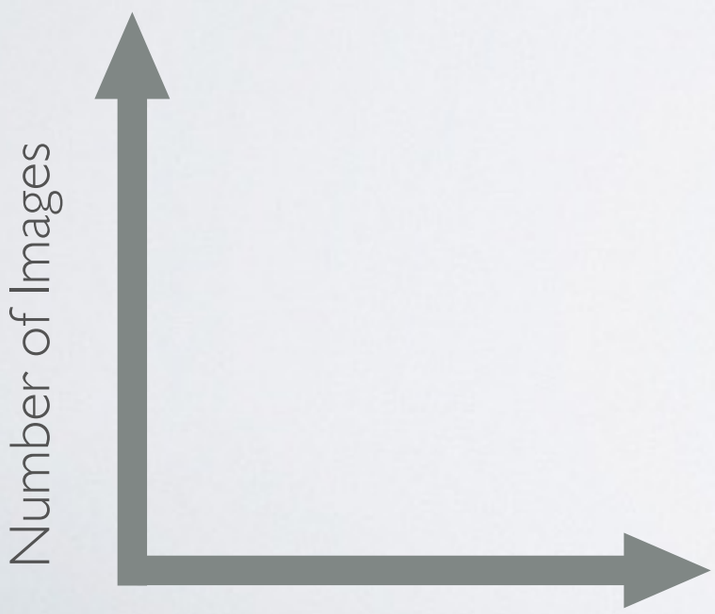


Images with People



Images without People

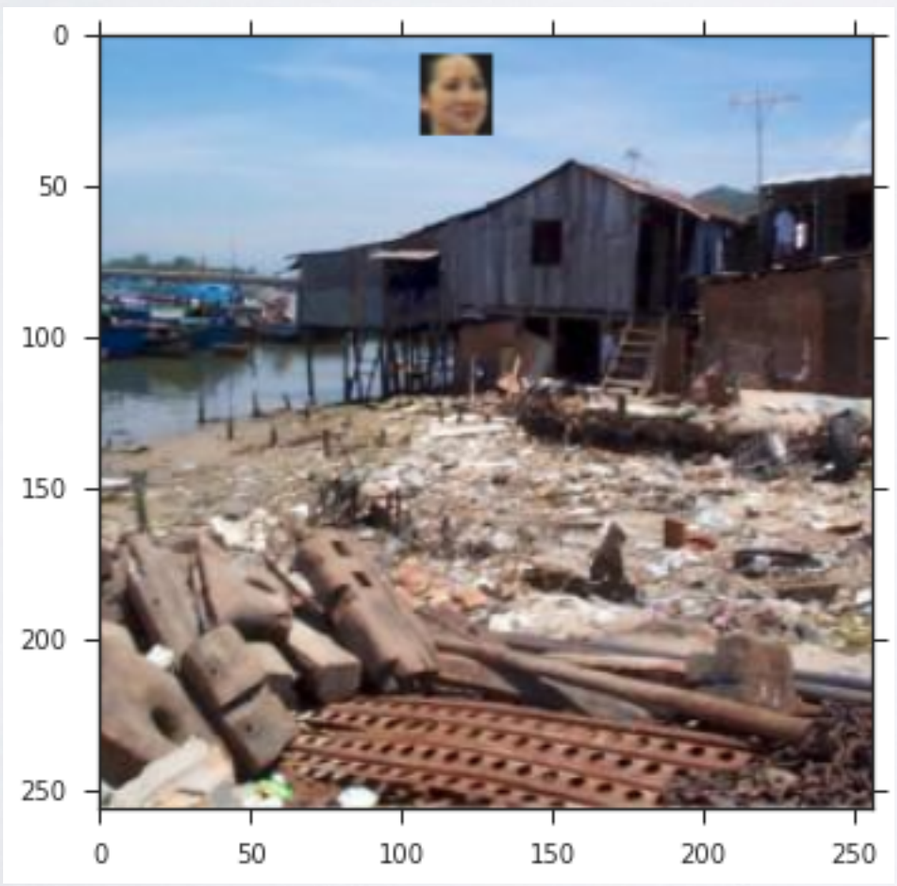




Removing people or faces from images shows stronger signs of decreasing memorability, but still not very conclusive.

ADDING FACE TO IMAGES

- What happens if we paste a face into all of our images?

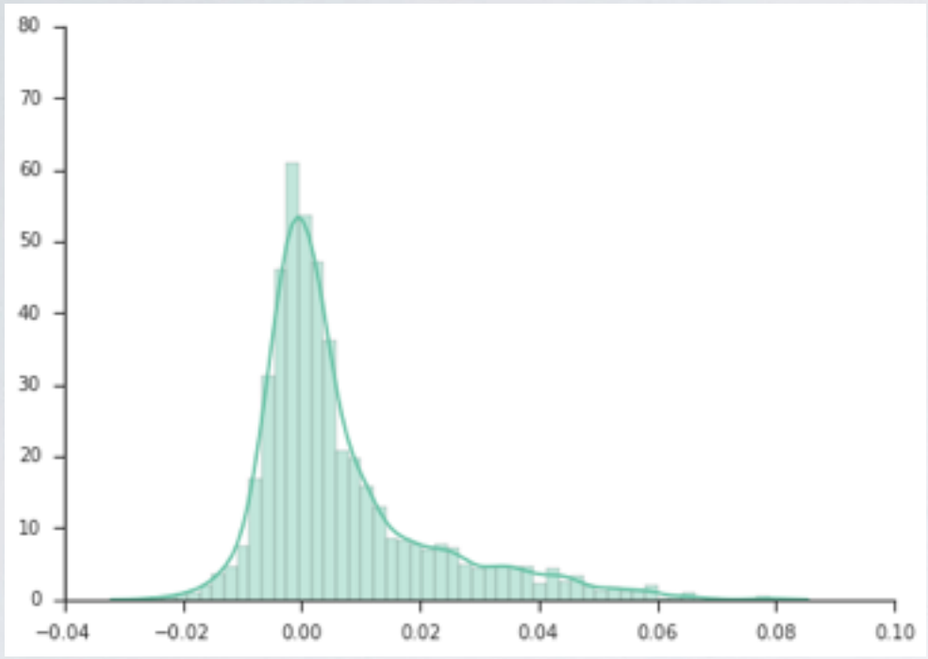


Actual: 0.61

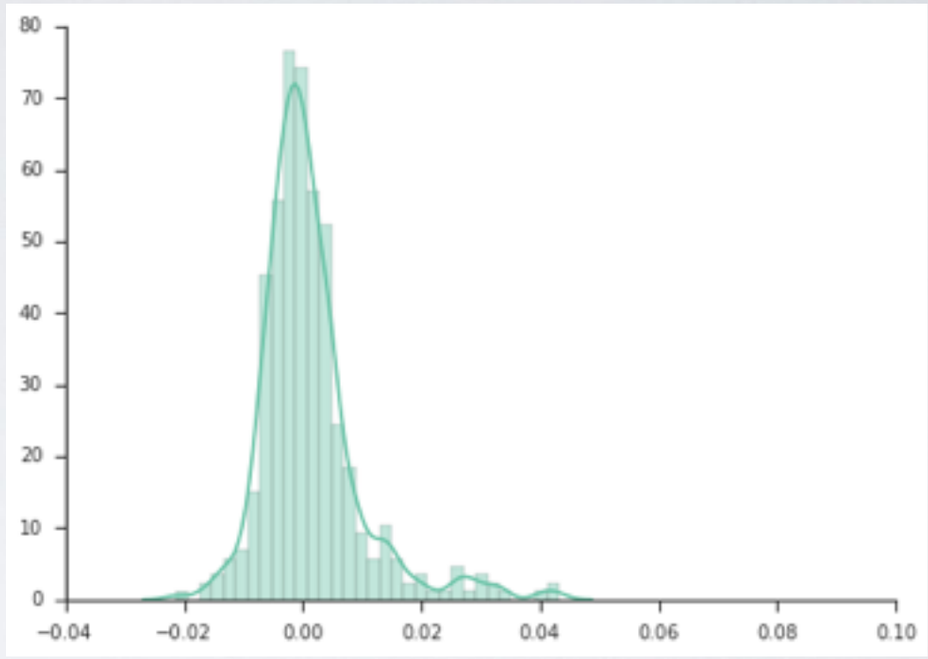
Predicted: 0.60

Predicted with Face: 0.62

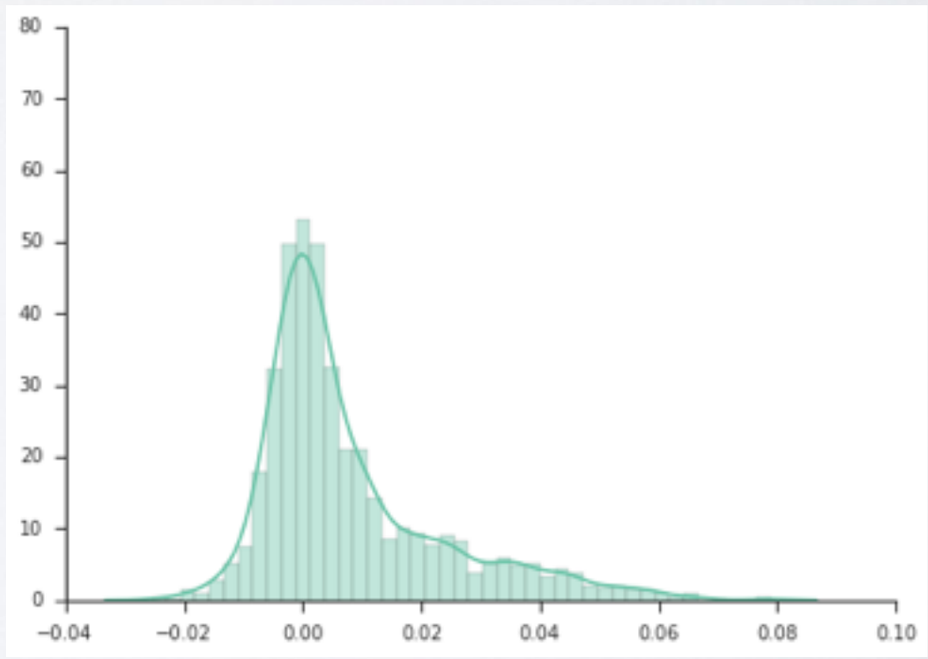
All Images

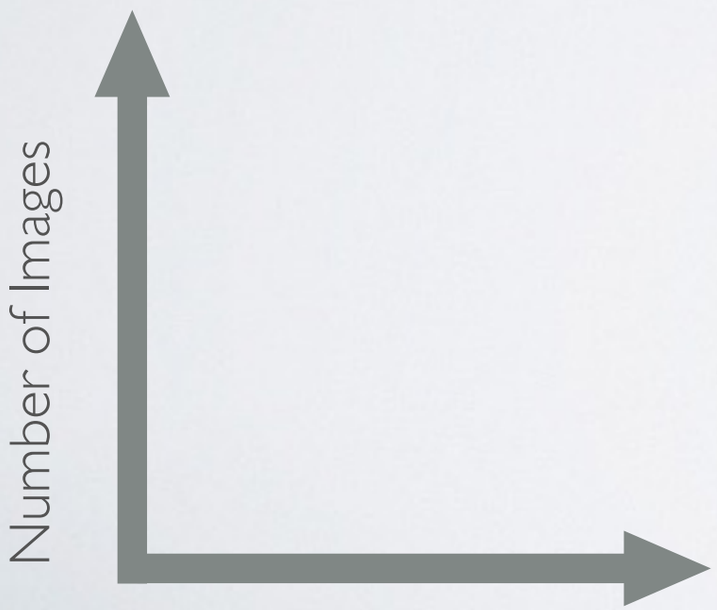


Images with People



Images without People





Change in memorability (added faces - normal)

Adding a face to images seems to increase memorability

SUMMARY

- MemNet generalizes to the popularity dataset - approaching state-of-the-art results (without fine-tuning).
- t-SNE embeddings suggest people might improve memorability while landscapes and structures are not very memorable.
- Inconclusive results when blurring/removing people in images and its effects on MemNet. Perhaps stronger results if hand blur all people.
- Adding a single 27x27 face to images looks to boost predicted memorability especially for images with no people.
- Adding or removing people from images may be changing predicted memorability for other reasons.

REFERENCES

- http://web.mit.edu/phillipi/Public/WhatMakesAnImageMemorable/

- http://memorability.csail.mit.edu/