

Reverse Image Search

Piyush Kansal and Vinay Krishnamurthy

Dataset and Algo

- Dataset: 400k images/tags from MIRFLICKR_[1]
- Used LDA_[2] on text and Spectral Hashing_[3] on images instead of Perceptual Hash_[4]

Text Processing

- 450k unique words/tags in lexicon

Find similar text using:

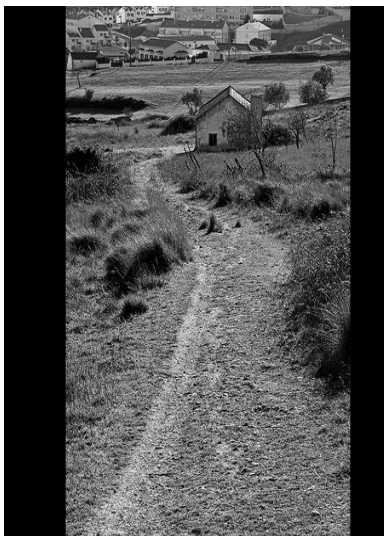
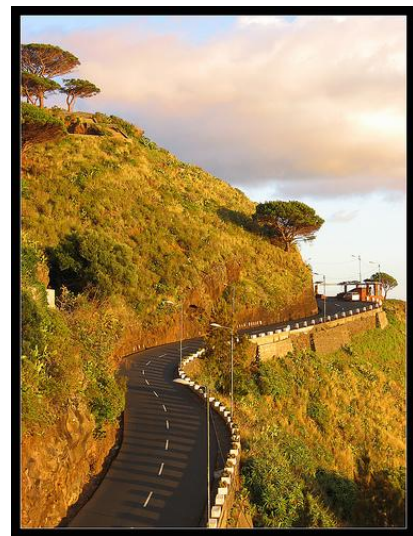
- SSD
 - output 1000 images
- LDA
 - created 10, 20, 30 and 40 topics
 - 10 words in each topic
 - selected 2 topics of highest weight
 - weighted average on topics and file count
 - output 2000 images

Image Processing

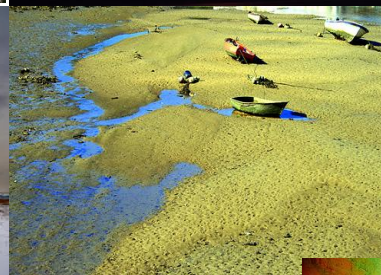
Find similar images using:

- Spectral Hashing
- 3000 images i/p to this step
- Used Hamming distance of 7 to generate final result

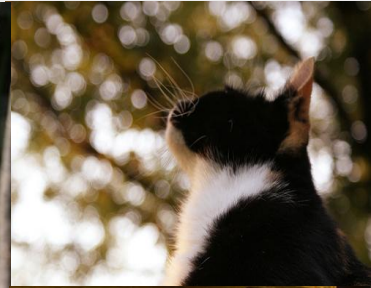
Test 1 - Scenery



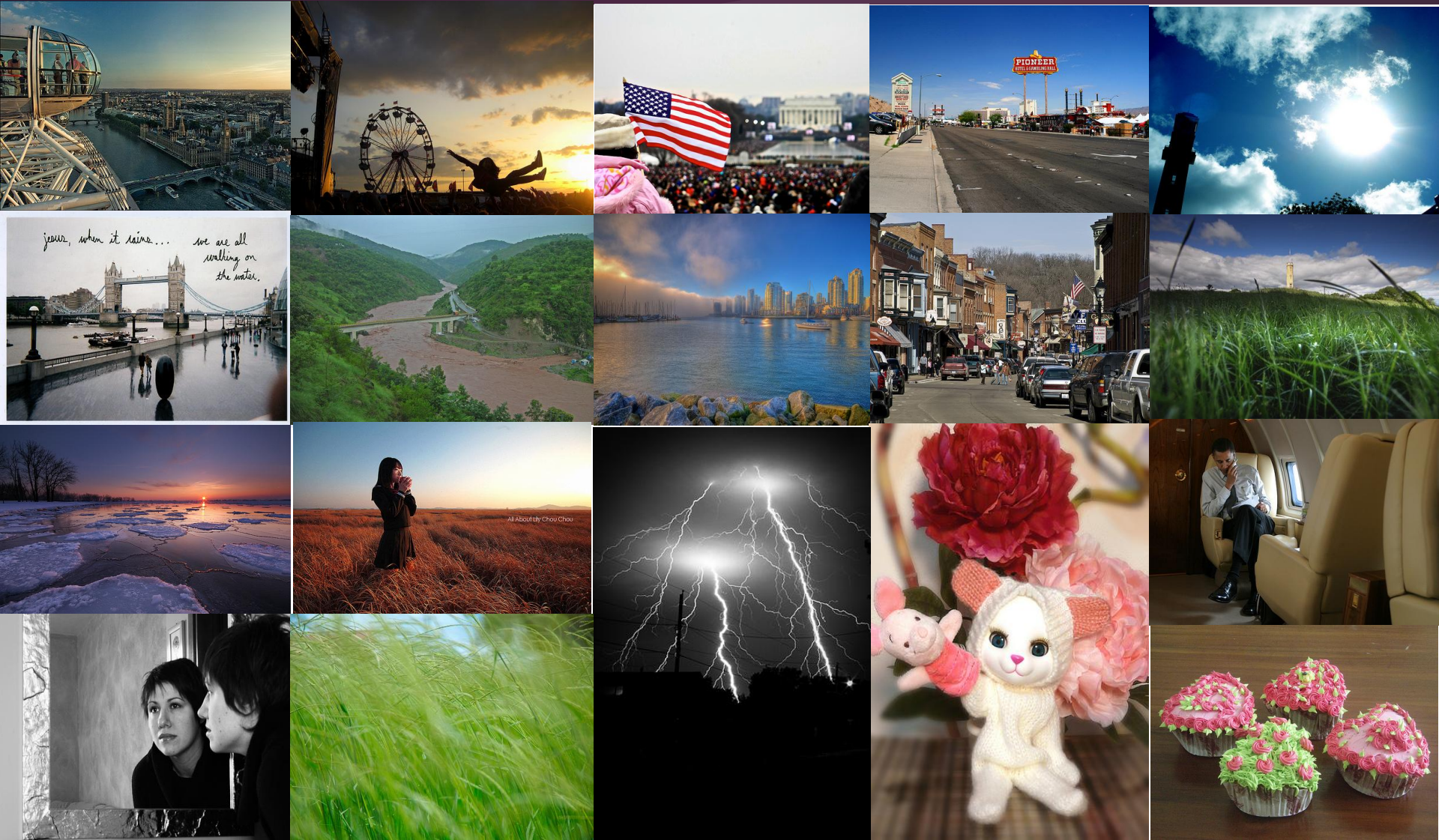
Test 2 - Water bodies



Test 3 - Animals



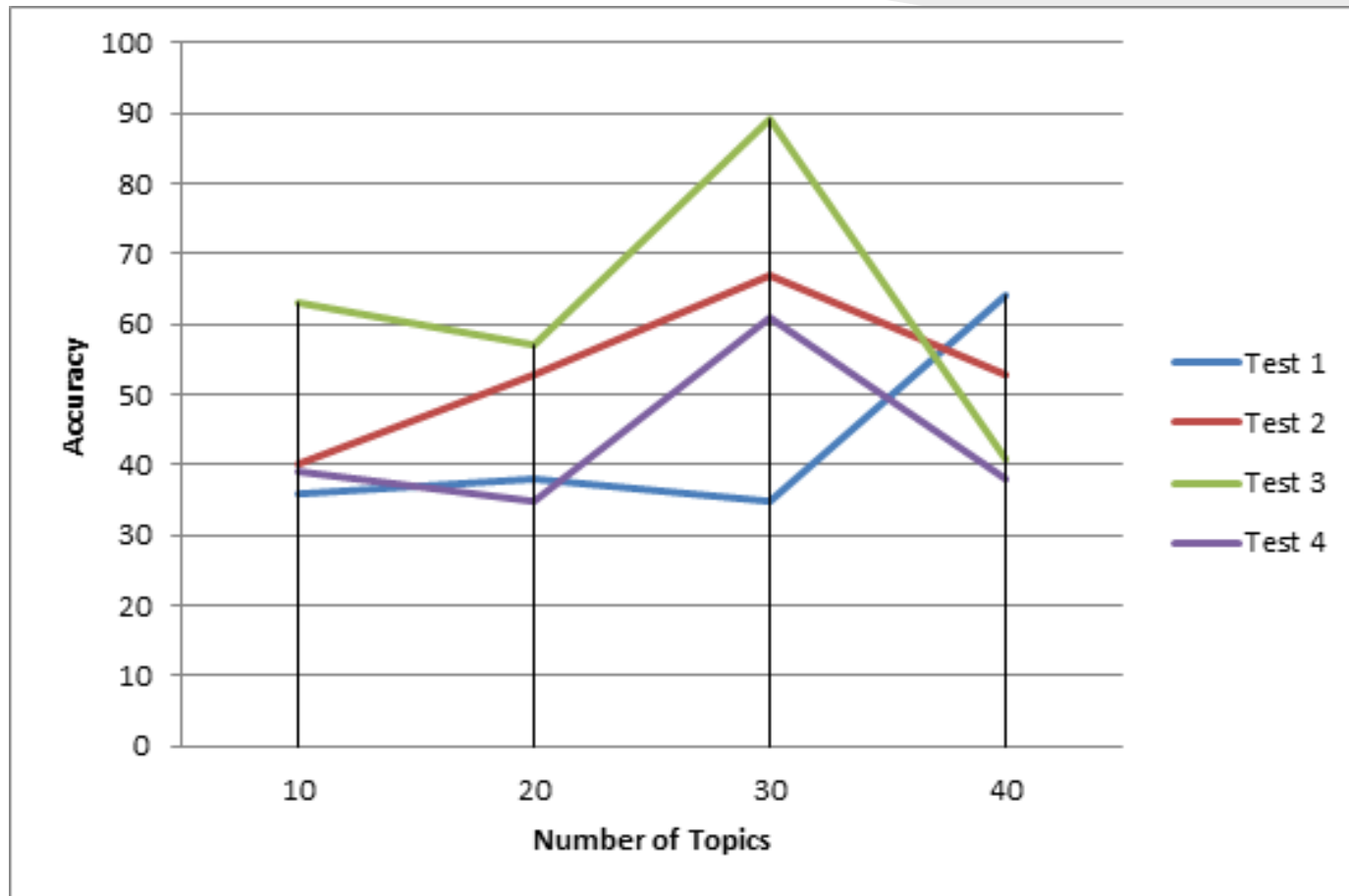
Test 4 - Cities



Test Results - Analysis

- Test 1
 - Good results. 'ilustrarportugal', 'sérieouro', 'http://www.flickr.com/photos/tags/**abigfave**/'
- Test 2
 - Much better results. Empty tag files, not so relevant words in topics: '365days', 'bw', '35mm'
- Test 3
 - Good results: text analysis - Animals, image analysis - People
 - Empty tag files
- Test 4
 - Satisfactory results: mostly from image analysis
 - Mostly empty tag files
- Test results quite dependent on how closely tags match and accuracy of tags placed on corresponding images
- Good improvement for larger number of topics, mirflickr has 10 image categories and 30 tag categories

Test Results - Analysis (cont.)



Future Work

- Multidimensional Spectral Hashing (MDSH) [5] instead of Spectral Hashing, takes care of both closer and larger neighbourhoods
- Do not consider empty tag files, will improve accuracy

References

- [1] <http://press.liacs.nl/mirflickr/>
- [2] D. Blei, A. Ng, and M. Jordan. Latent Dirichlet allocation. In: ANIPS. (2002)
- [3] Weiss, Y., Torralba, A., Fergus, R.: Spectral hashing. In: NIPS. (2008)
- [4] <http://phash.org/>
- [5] Weiss, Y., Torralba, A., Fergus, R.: Multidimensional Spectral hashing. In: ECCV. (2012)



Thank You