Distributed Deep Search Engine

Based on Image representations & Text

Bharathi, Laura & Pratheeeksha
New York University
Introduction

- Image search engine supporting multimodal queries
- Dataset: 50k images MVSO and Flickr image database
- Deep learning and Information retrieval based search
System Design

MVSO/Flickr Source

Images → Feature extractor → KD-tree generator → Image indexers

Text → Doc shard creator → Doc shards → Main backend service

Text index generator → Text indexers

Web backend → Web frontend
System Architecture - Feature Generation

Image source: AlexNet Architecture, Leonardo Araujo dos Santos, Artificial Intelligence
https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/image_folder_7/AlexNet_0.jpg
System Architecture - Image Indexers

- Image feature vectors stored in KD trees
- \( K = 4096 \)
- Min leaf size 20
- Trees sharded by doc id

Image source: Stanford, CS 368, Geometric Algorithms, Search Structures notes
http://graphics.stanford.edu/courses/cs368-00-spring/TA/manuals/CGAL/ref-manual2/SearchStructures/kdtree.gif
Demo
Results

- What types of similarity can be represented?
- What happens with different image sizes?
- How fast is our Search Engine?
- Where does it fail?
Results - What types of similarity can be represented?
Results - What happens with different image sizes?

- Assumptions about Image sizes
- Alexnet takes (227,227)
- Smaller images upscaled
- Larger images downscaled
- PIL Lanczos filter
- Lossless scaling
### Results - How fast is our Search Engine?

<table>
<thead>
<tr>
<th>Number of Index servers</th>
<th>Avg. querying speed (in seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.1796</td>
</tr>
<tr>
<td>15</td>
<td>1.621</td>
</tr>
<tr>
<td>10</td>
<td>1.9182</td>
</tr>
<tr>
<td>15</td>
<td>2.4872</td>
</tr>
<tr>
<td>25</td>
<td>2.8902</td>
</tr>
</tbody>
</table>
Issues we encountered

- Query slow down as number of trees increases
- Tornado timeout bug
- Images with four channels
- Black images - moons
We filtered out mostly black images
Discussion on Failure Cases - Case of Evil Dog

(a) Query image
(b) 1st result
(b) 2nd result
(b) 3rd result
(c) 4th result
Discussion on Failure Cases - Case of Wreck boat

(a) Query image

(b) 1st result

(b) 2nd result

(b) 3rd result

(c) 4th result
Discussion on Failure Cases—Curious case of missing pizza

(a) Query image

(b) 1st result

(b) 2nd result

(b) 3rd result

(c) 4th result
Demo
Further Work

Improving Scalability

- Increasing number of Image indices
- Reduce dimensionality of the feature vectors
- Distributed KD-Trees.

Improving Quality

- Use more expressive neural networks
- Try different types of spatial tree
- Redesign combined image and text search
- Reduce title boost and/or effect of normalization for document length
Further Work - Additional features

- Image Segmentation
- Extract image description from image, query with this in absence of text
- Querying with an image and style
- Multiple Image Types

Our ambition is to scale this to 5m images
References

3. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou,
Questions?
Appendix
System Architecture

- **Feature Generation**
  - Low dimensional Image representation
  - Pretrained Alexnet

- **Image Indexers**
  - KD-Trees shared by doc id
  - Map-reduce framework

- **Text Indexers**
  - Inverted index
  - Sharded by doc id

- **Doc Shards**
  - Containing image metadata
  - Sharded by doc id

- **Image Store**
  - A folder on disk

- **Retrieval**
  - Features extracted from query image
  - Query both sets of indices
  - Combine results and display