Project Ideas

Fine-grained Categorization of Wild Animals ‘in the Wild’

- Build a web demo for fine-grained categorization of animals and/or their activities. You can consider a subset of images and classes of the Snapshot Serengeti data (contact the instructors to get access to the data).

Top classification results and associated confidence:

<table>
<thead>
<tr>
<th>Species</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zebra</td>
<td>0.98</td>
</tr>
<tr>
<td>Striped Hyena</td>
<td>0.01</td>
</tr>
<tr>
<td>Leopard</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standing</td>
<td>0.95</td>
</tr>
<tr>
<td>Interacting</td>
<td>0.03</td>
</tr>
<tr>
<td>Eating</td>
<td>0.02</td>
</tr>
</tbody>
</table>

- As part of your set of features, you may consider using the “Overfeat” source code (check the resources page) for extraction of features based on deep convolutional neural networks.
Face Mask Detection

- Build a system for detection of face masks or helmets. Traditional face detectors fail to detect faces under severe occlusion. The system could be applied for access control in ATM machines.

- You may also address the problem of improving face detection in challenging scenarios (e.g., a person wearing a hat + sunglasses).
Project Ideas

Cross-Domain Learning: From ImageNet to Serengeti

- Build a transfer learning/cross domain learning paradigm to leverage both clean large scale image data from ImageNet (http://www.image-net.org/) and the annotation information from the Snapshot Serengeti project for an improved recognition rate.

- In particular, the cross domain learning method should be able to match the ImageNet hierarchy in a flexible way to ensure maximum usage of the clean data.

ImageNet data

Serengeti images
Project Ideas

Fast Foreground Matching with Learned Binary Codes

- Motivated from the practical need in the Snapshot Serengeti project, the key objective is to find or match foreground animals from different scenes and cameras with the assumption that the cameras are generally static. The matching and search results can be used to help the researchers to capture and understand the dynamics and distributions of different species in the Serengeti National Park, Tanzania.

- Note that the location information of each camera and the time stamp of each shot will be useful for extracting the background and perform location related fast search.
Generalize Deep Learning to Small Datasets

A great advantage of deep learning models lies in their generalizability. Although deep learning may easily overfit on small datasets, we may employ the filters and networks trained from large datasets (e.g., ImageNet), and adapt the model for image recognition on small datasets. Because such a network benefits from prior knowledge, we can often find its superior performance over the state-of-the-art on these small datasets. Here is a good example of generalizing a deep model to dog-vs-cat recognition.

How to get existing deep learning models?
• You may borrow models from Caffe or OverFeat, both can run with/without GPU

How to choose another small dataset?
• You may consider Kaggle Dog vs. Cat, Caltech 101, Caltech 256, or a subset of Serengeti animal datasets.
Existing public deep learning demos work well in recognizing general objects but not faces. Below is a failed recognition example using deep convolutional nets. The reason for the failure is because most public deep neural network methods are trained from ImageNet LSVRC, which does not contain a “human face” category.

You are encouraged to build a demo which is more friendly with face images.

- You may borrow the deep learning module from Caffe or OverFeat
- You may use OpenCV for the face detection module.
- By combining the above two modules, you can be creative when realizing the user uploads a face image. Some funny examples are like “this is a face of famous computer vision researcher”, or “you look like a movie star”.

![Image of a failed recognition example using deep convolutional nets.]

Predicted objects:
1. Bow Tie, Bow-Tie, Bowtie (0.23)
2. Suit, Suit Of Clothes (0.08)
3. Bonnet, Poke Bonnet (0.08)
4. Hair Slide (0.05)
5. Band Aid (0.04)

Other objects:
Project Ideas

Deep Learning For Object Detection

The recent progress in object detection suggests another revolutionary change since the introduction of the prevailing deformable part-based models five years ago. In recent studies, the deep learning based approach has shown great potential in the ImageNet detection challenge. However, there are still limited studies about using deep learning for other datasets such as PASCAL or Caltech Pedestrian detection.

You are encouraged to modify the deep learning models for object detection in images or videos.
- You may refer to the OverFeat paper and part of their source code to start.
- You may choose one or more categories from PASCAL for your detection tasks.
- You may try ImageNet fine-grained object recognition images (with bounding boxes) for your detection tasks.