NEIL: Extracting Visual Knowledge from Web Data

Xinlei Chen, Abhinav Shrivastava, Abhinav Gupta
Carnegie Mellon University
Lots of labeled data & common-sense relationships have helped improve performance!

But how do we label data and collect common sense at a large scale?
Labeled Data:

Amazon

Mechanical Turk beta

LabelMe

Common Sense Relationships:

Cycorp

Home of smarter solutions
Labeled Data:

- Amazon Mechanical Turk
- LabelMe
- Instagram
- Facebook

1M Boxes
5 years

400M images
daily!

Common Sense Relationships:

- 2M Rules
- 30 years
  and still continuing…
NEIL
NEVER ENDING IMAGE LEARNER

Running
24 hours a day,
7 days a week

Trying to understand images on the web and build world’s largest Visual Knowledge Base automatically…
NEIL’S KNOWLEDGE BASE

Concepts

Relationships
OBJECTS
Camry
Scenes

Parking Lot

Raceway
ATTRIBUTES

Round Shape

Crowded
RELATIONSHIPS

Object - Object

Partonomy

Wheel is a part of Car

Taxonomy or Similarity

Corolla is a kind of/looks similar to of Car
RELATIONSHIPS

Object - Scene

Car is found in Raceway
RELATIONSHIPS

Object - Attribute

Wheel is/has Round shape
RELATIONSHIPS

Scene – Attribute

Bamboo forest is/has Vertical lines
NEIL'S KNOWLEDGE BASE

Concepts

- Objects
- Scenes
- Attributes

(Provided by Text Analysis)

Relationships

- Object-Object
  Partonomy, Taxonomy/Similarity
- Object-Scene
- Object-Attribute
- Scene-Attribute
NEIL
NEVER ENDING IMAGE LEARNER

• 4 months on 200 cores

• >2500 Concepts
  ▪ >1500 Objects
  ▪ >1034 Scenes
  ▪ >87 Attributes

• >5 million images analyzed

• Labeled
  ▪ 600K images
  ▪ 3000 relationships

www.neil-kb.com

Train Your Own Concept!
NEIL: Never Ending Image Learner
I Crawl, I See, I Learn.

WHAT HAS NEIL LEARNED TO SEE?
Here are random examples:

OBJECTS

SCENES

ATTRIBUTE

TRAIN A
CONCEPT

ABOUT

WHAT COMMON SENSE FACTS HAVE NEIL LEARNED?
Here are a few examples:
NEIL
NEVER ENDING IMAGE LEARNER

• 4 months on 200 cores  www.neil-kb.com

• >2500 Concepts
  ▪ >1500 Objects
  ▪ >1034 Scenes
  ▪ >87 Attributes

• >5 million images analyzed

• Labeled
  ▪ 600K images
  ▪ 3000 relationships
WHY DOES NEIL WORK?
MICRO-VISION
MACRO-VISION
**Structure in the Visual World**

**Car** is found on **Street**

**Sheep** are **White**

**Constrained**

**Semi-supervised Learning**
SEMANTICALLY-DRIVEN ACQUISITION

[Fergus et al. ECCV’04], [Berg et al. CVPR’06], [Snavely et al. SIGGRAPH’06], [Schroff et al. ICCV’07], [Simon et al. ICCV’07], [Hays et al., CVPR’08], [Li et al. ECCV’08], [Shrivastava et al. ToG’11], [Rubenstein et al. CVPR’13] ...
HOW DOES NEIL WORK?
1. No Bounding-boxes
2. Noise
3. Multiple Meanings (Polysemy)
(0) Seed Images

Desktop Computer

(1) Subcategory Discovery

Monitor

(1) Subcategory Discovery

Keyboard

(1) Subcategory Discovery

Television
EXEMPLAR DETECTORS

Car
POLYSEMY

Falcon
(0) Seed Images

(1) Subcategory Discovery

(2) Train Models
- Desktop Computer (1)
- Desktop Computer (2)
- Desktop Computer (3)
- Monitor (1)
- Monitor (3)
- Keyboard (1)
- Keyboard (2)
- Keyboard (3)
- Television (1)
- Television (2)
- Television (3)
Train Models

• Latent SVM
  ▪ Objects, Attributes
  ▪ CHOG

• Linear SVM
  ▪ Scenes, Attributes
  ▪ Color, Texton, HOG, SIFT, GIST

• … Your model?
(0) Seed Images

(1) Subcategory Discovery

(2) Train Models

(3) Relationship Discovery
Learned relationships:

- **Keyboard** is a part of **Desktop Computer**
- **Monitor** is a part of **Desktop Computer**
- **Television** looks similar to **Monitor**
Learned relationships:

- Keyboard is a part of Desktop Computer
- Monitor is a part of Desktop Computer
- Television looks similar to Monitor

(0) Seed Images

(1) Subcategory Discovery

(2) Train Models

(3) Relationship Discovery

Learned relationships:
(0) Seed Images

(1) Subcategory Discovery

(2) Retrain Models

(3) Relationship Discovery

(4) Add New Instances

Learned relationships:
- Keyboard is a part of Desktop Computer
- Monitor is a part of Desktop Computer
- Television looks similar to Monitor
**COMMON SENSE RELATIONSHIPS**

**Object - Object**

- **Sparrow** is a kind of/looks similar to **bird**
- **Eye** is a part of **Baby**

**Object - Scene**

- **Helicopter** is found in **Airfield**
- **Ferris wheel** is found in **Amusement park**
COMMON SENSE RELATIONSHIPS

Figure 5. Qualitative Examples of Scene-Object (rows 1-2) and Object-Object (rows 3-4) Relationships Extracted by NEIL
# Can NEIL Help Vision Tasks?

## Scene Classification

<table>
<thead>
<tr>
<th>Method</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seed Classifier (15 Google Images)</td>
<td>0.52</td>
</tr>
<tr>
<td>Bootstrapping (without relationships)</td>
<td>0.54</td>
</tr>
<tr>
<td>NEIL Scene Classifiers</td>
<td>0.57</td>
</tr>
<tr>
<td>NEIL (Classifiers + Relationships)</td>
<td>0.62</td>
</tr>
</tbody>
</table>


Can NEIL help vision tasks?

### Object Detection

<table>
<thead>
<tr>
<th>Method</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latent SVM (450 Google Images)</td>
<td>0.28</td>
</tr>
<tr>
<td>Latent SVM (450, Aspect Ratio Clustering)</td>
<td>0.30</td>
</tr>
<tr>
<td>Latent SVM (450, HOG-based Clustering)</td>
<td>0.33</td>
</tr>
<tr>
<td>Seed Detector (NEIL Clustering)</td>
<td>0.44</td>
</tr>
<tr>
<td>Bootstrapping (without relationships)</td>
<td>0.45</td>
</tr>
<tr>
<td>NEIL Detector</td>
<td>0.49</td>
</tr>
<tr>
<td>NEIL Detector + Relationships</td>
<td>0.51</td>
</tr>
</tbody>
</table>
NEIL

NEVER ENDING IMAGE LEARNER

Running
24 hours a day,
7 days a week

Forever to
Label data,
Learn relationships
THANK YOU!

All Models and Relationships will be found* on www.neil-kb.com

Opera house is found in Sydney

*20 Dec, 2013