NEIL: Extracting Visual Knowledge from Web Data

Xinlei Chen, Abhinav Shrivastava, Abhinav Gupta
ICCV 2013 (cited 32 on Google scholar)

Presented by Lee YoonSeok  Nov. 04, 2014
Research Goal

- Build **world’s largest visual knowledge base** using a computer program that runs 24 hours per day and 7 days per week to automatically extract visual knowledge from Internet data

- NEIL (Never Ending Image Learner)
NEIL’s Knowledge Base

- 3 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 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

● Relationships
  ● Object-Object
    ● Partonomy, Taxonomy/Similarity
  ● Object-Scene
  ● Object-Attribute
  ● Scene-Attribute
1. No Bounding-boxes
2. Noise
3. Multiple Meanings (Polysemy)
(0) Seed Images

(1) Subcategory Discovery

Desktop Computer

Monitor

Keyboard

Television
EXEMPLARY DETECTORS

Car
POLYSEMY

Falcon
(0) Seed Images

(1) Subcategory Discovery

(2) Train Models

Desktop Computer (1)
Desktop Computer (2)
Desktop Computer (3)
...
Monitor (1)
...
Keyboard

---

- Desktop Computer
- Monitor
- Keyboard
- Television
(0) Seed Images

(1) Subcategory Discovery

(2) Train Models
   - Desktop Computer (1)
   - Desktop Computer (2)
   - Desktop Computer (3)
   - Monitor (1)
   - ...

(3) Relationship Discovery

Learned relationships:
- Keyboard is a part of Desktop Computer
- Monitor is a part of Desktop Computer
- Television looks similar to Monitor
Never Ending

(0) Seed Images
(1) Subcategory Discovery
(2) Train 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
Takes quite long!

16 iterations for 2.5 months, with 200 core
How to use relationships

- Scoring function for category $i$, given image $x$

\[
\phi_i(x) + \sum_{i,j \in \mathcal{R}_O \cup \mathcal{R}_A} \phi_j(x_l) \psi_{i,j}(x, x_l) + \sum_{i,k \in \mathcal{R}_S} \omega_k(x)
\]

- $\phi_i(x)$: Score from object detector $i$ for image $x$

- $\psi_{i,j}(x, x_l)$: Compatibility function between object category $i,j$

- $\omega_k(x)$: Score from scene classifier $k$ for image $x$
How to use relationships

- Scoring function for category $i$, given image $x$

\[ \phi_i(x) + \sum_{i,j \in \mathcal{R}_O \cup \mathcal{R}_A} \phi_j(x_l) \psi_{i,j}(x, x_l) + \sum_{i,k \in \mathcal{R}_S} \omega_k(x) \]

- Original score
- Score of object, attribute relationship
- Score of scene relationship

$\phi_i(x)$: Score from object detector $i$ for image $x$

$\psi_{i,j}(x, x_l)$: Compatibility function between object category $i,j$

$\omega_k(x)$: Score from scene classifier $k$ for image $x$
## Quantitative Result

<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><strong>0.62</strong></td>
</tr>
</tbody>
</table>
**Quantitative Result**

Table 2. mAP performance for object detection on 15 categories.

<table>
<thead>
<tr>
<th>Method</th>
<th>mAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latent SVM (50 Google Images)</td>
<td>0.34</td>
</tr>
<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><strong>0.51</strong></td>
</tr>
</tbody>
</table>
Contributions

- Never ending learning algorithm from the Internet via macro-vision
- Automatically build a large visual structured knowledge base including relationships
- Relationships can be constraints for improving semi-supervised learning
Limitations

- Rely on human (or other project) to learn a new concept
  - Published: >2500 Concepts (2.5 month)
  - Now: 2702 Concepts (1 year?)

- Are relationships meaningful?
  - Published: >1700 relationships
  - Now: 4695 relationships
Apple can be a kind of / look similar to Orange.

Banana can be a kind of / look similar to Orange.

Basketball can be a kind of / look similar to Orange.

Candle can be a kind of / look similar to Orange.

Gourd can be a kind of / look similar to Orange.

Lantern can be a kind of / look similar to Orange.

Lemon can be a kind of / look similar to Orange.

Frisbee can be a kind of / look similar to Orange.

Melon can be a kind of / look similar to Orange.

Peach can be a kind of / look similar to Orange.

Platter can have a part Orange.
Q&A

http://www.neil-kb.com/