Learning Hierarchical Visual Representations in Deep Neural Networks Using Hierarchical Linguistic Labels

Joshua Peterson, Paul Soulos, Aida Nematzadeh, & Tom Griffiths, University of California, Berkeley

**CATEGORY LEARNING APPROACHES**

Image classification learning models in computer vision use single labels for each object class.

- Pug
- Dalmatian
- Zebra
- Tow Truck

Children learn multiple labels for a concept:

- Animal
- Level
- Superordinate
- Basic
- Subordinate
- Pug
- Dalmatian
- Dog
- Poodle

Which level of abstraction results in representations that better capture human generalization behavior?

**EXPLORING REPRESENTATIONS**

Which labels result in representations that match people?

We evaluated the correlation between the neural representations and human similarity judgements.

Training with basic labels only performs the same as previous work using only subordinate labels (Peterson, et al., 2016).

**GENERALIZATION**

Training - Conditions

- 1 subordinate
- 3 subordinate
- 3 basic

Test - Pick everything that is a “dax”

Generalization Model

\[ g(q_i, c) = e^{-d(q_i, c)} \]

<table>
<thead>
<tr>
<th>Model</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>basic training</td>
<td>57</td>
</tr>
<tr>
<td>sub training</td>
<td>0</td>
</tr>
<tr>
<td>basic training</td>
<td>38</td>
</tr>
<tr>
<td>sub training</td>
<td>57</td>
</tr>
<tr>
<td>Peterson et al., 2016</td>
<td>58</td>
</tr>
</tbody>
</table>

**TRAINING**

We train a CNN to predict classes at multiple levels of a hierarchy, each weighted using the hyperparameter \( \alpha \).

Subordinate- and basic-level labels were taken from ILSVRC12 and Wang and Cottrell (2015) respectively.

\[ \alpha L_{basic} + (1 - \alpha) L_{subordinate} \]

**CONCLUSION**

Subordinates are already perceptually similar, so basic level supervision appears to be sufficient to learn the taxonomy. Future work should explore how few-shot learning strategies might benefit these biases.

