COSTA: Co-occurrence statistics for zero-shot classification

Goals and Novelties

Goal
COSTA: a zero-shot classifier using a linear combination of existing classifiers, weighted by co-occurrences statistics

Existing classifiers
Co-occurrences
Zero-shot prediction

Novelties
1. Allows for multi-label zero-shot prediction
2. For certain concepts, COSTA outperforms supervised SVMs
3. Co-occurrences can be obtained easily from the web
4. COSTA is a natural prior in few-shot classification

Hypothesis: concept-to-concept inter-dependencies reveal a significant part of the latent image semantics

Advantages over Attributes:
1. Many visual concepts can be described as an open set of concept-to-concept relations
2. Mapping between known and unknown labels only require co-occurrence statistics
3. All concepts are treated equally, no distinction between classes and attributes

Conclusion: COSTA leverages, by design, the bias of natural co-occurrences of visual concepts.

COSTA model

Zero-shot classifier
• Set of k known labels and trained classifiers
• Co-occurrences between known labels and unknown labels
• Approximate new classifier \( \hat{w} \) using of the known classifiers:

\[
\hat{w} = \sum_{k} a_k w_k s_{ik}
\]

with: co-occurrence weight \( a_k \) and regression weight \( a_i \)

Define co-occurrence weights
1. Normalized co-occurrences \( \hat{s}_{ij} = \frac{c_{ij}}{c_i} + c_j \)
2. Dice’s coefficient \( \hat{s}_{ij} = \frac{c_{ij}}{c_i + c_j} \)

Define an object by what it is not

\[
\hat{w} = \sum_{k} w_k s_{ik}^{++} - w_k s_{ik}^{+-} - w_k s_{ik}^{-+} + w_k s_{ik}^{--}
\]

Classifier regression
• Leave one out optimization over the known classifiers

\[
L_{reg} = \sum_{i} \left\| \hat{w}_i - \sum_{k} a_k w_k s_{ik} \right\|_2^2
\]

Co-occurrences

Hypothesis: concept-to-concept inter-dependencies reveal a significant part of the latent image semantics

Advantages over Attributes:
1. Many visual concepts can be described as an open set of concept-to-concept relations
2. Mapping between known and unknown labels only require co-occurrence statistics
3. All concepts are treated equally, no distinction between classes and attributes

Conclusion: COSTA leverages, by design, the bias of natural co-occurrences of visual concepts.

Example

Objects co-exist with other objects in a scene.

Sink is usually in the same visual space as a cupboard, a stove, and a dishwasher.

This research is supported by the STW STORY project and the Dutch national program COMMIT.

Experiments

Datasets:
- H-SUN (poster), iCLEF10, CUB Attributes (paper)
- Evaluation: Mean Average Precision over 25% of the labels
- Features: Fisher Vectors, SIFT, PCA 96, K=16, SSR, L2 normalization
- Settings: Supervised SVM (SUB), Leave-one-out (L1O), Zero-shot with 25% unknown labels (ZS75) and 50% unknown labels (ZS50)

Zero-shot classification

<table>
<thead>
<tr>
<th>Setting</th>
<th>SUB</th>
<th>L1O</th>
<th>ZS75</th>
<th>ZS50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>13.5</td>
<td>13.5</td>
<td>12.3</td>
<td>12.3</td>
</tr>
<tr>
<td>COSTA</td>
<td>12.9</td>
<td>13.9</td>
<td>12.9</td>
<td>12.9</td>
</tr>
<tr>
<td>Cover Dice</td>
<td>13.5</td>
<td>13.5</td>
<td>12.9</td>
<td>12.9</td>
</tr>
<tr>
<td>P/N Dice</td>
<td>13.5</td>
<td>13.5</td>
<td>12.9</td>
<td>12.9</td>
</tr>
</tbody>
</table>

COSTA: Jumpsstarts learning with just a few examples

Comparison to SVM

For showcase, rug, vase, oven, armchair, poster, COSTA outperforms SVMs by over 5%