Interactive Image Labelling

- In cases as stock photography indexing, automatic image labelling is not accurate enough, and manual labelling too costly.
- An interactive approach is less costly, and much more accurate.
- Structured models encode dependencies among image labels explicitly, and are more expressive than independent classifiers.
- In a fully automatic setting (without user input), our models improve modestly over independent models.
- In an interactive setting, where user response is taken into account, the performance gain is significant.

Summary

Trees of groups of labels

- Group labels into single variable to allow more dependencies between labels
- Every state modeled explicitly, a node has $k^d$ states, $k$ offers a trade off between: computational tractability, expressiveness of the tree, and overfitting on train set.
- To determine tree: (1) Agglomerative label clustering, (2) Chow-Liu algorithm on clusters.
- Mixture of Trees: over models with different node size $k$, trees are learned independently and we use the average of the predictions of the individual trees.

Result on Image Labelling

- Three datasets:
  1. ImageCLEF'10: 6400 train / 1600 test images, 93 labels
  2. SUN'09 4300 train / 4500 test images, 107 labels
  3. Animals w Attributes 24000 train / 6000 test images, 85 labels
- Performance evaluated using
  - MAP retrieval performance of ranking images per label.
  - iMAP annotation performance of ranking labels per image.

Performance in MAP and iMAP

Comparison between fully automated setting (top), and interactive setting (bottom). Performance of independent (yellow), tree (light-red), and mixture of trees (dark-red) models.

Performance versus number of Questions

Percentage of images with at least $N$ labels for which the top $N$ predicted labels are all correct. Comparison between state-of-the-art [1] [2], which uses object bounding-boxes during train time, and our independent (yellow) and mixture of trees (red) models, which uses only image labels. Automatic prediction (dark), and interactive setting with 10 labels (light).