I Have Seen Enough: Transferring Parts Across Categories

David Novotny[1,2], Diane Larlus[2], Andrea Vedaldi[1]

[1]Visual Geometry Group, University of Oxford

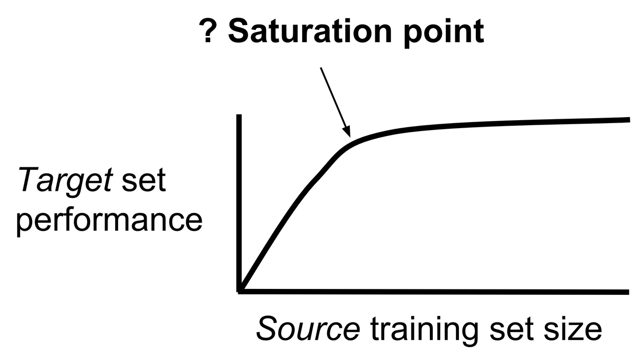
[2]Computer Vision Group, Xerox Research Centre Europe

An analysis of transferability of object parts

Main question: What is the minimal amount of supervision needed to perform transfer of object parts?

Overview

- Recent deep learning successes due to large amounts of data
- Question: **When is supervision enough?**
- An analysis of **part detector performance** on a target set w.r.t. source training set size
  - Presence of source/target sets → a domain adaptation problem



Analysis on animal parts

Parts are shareable among object categories:

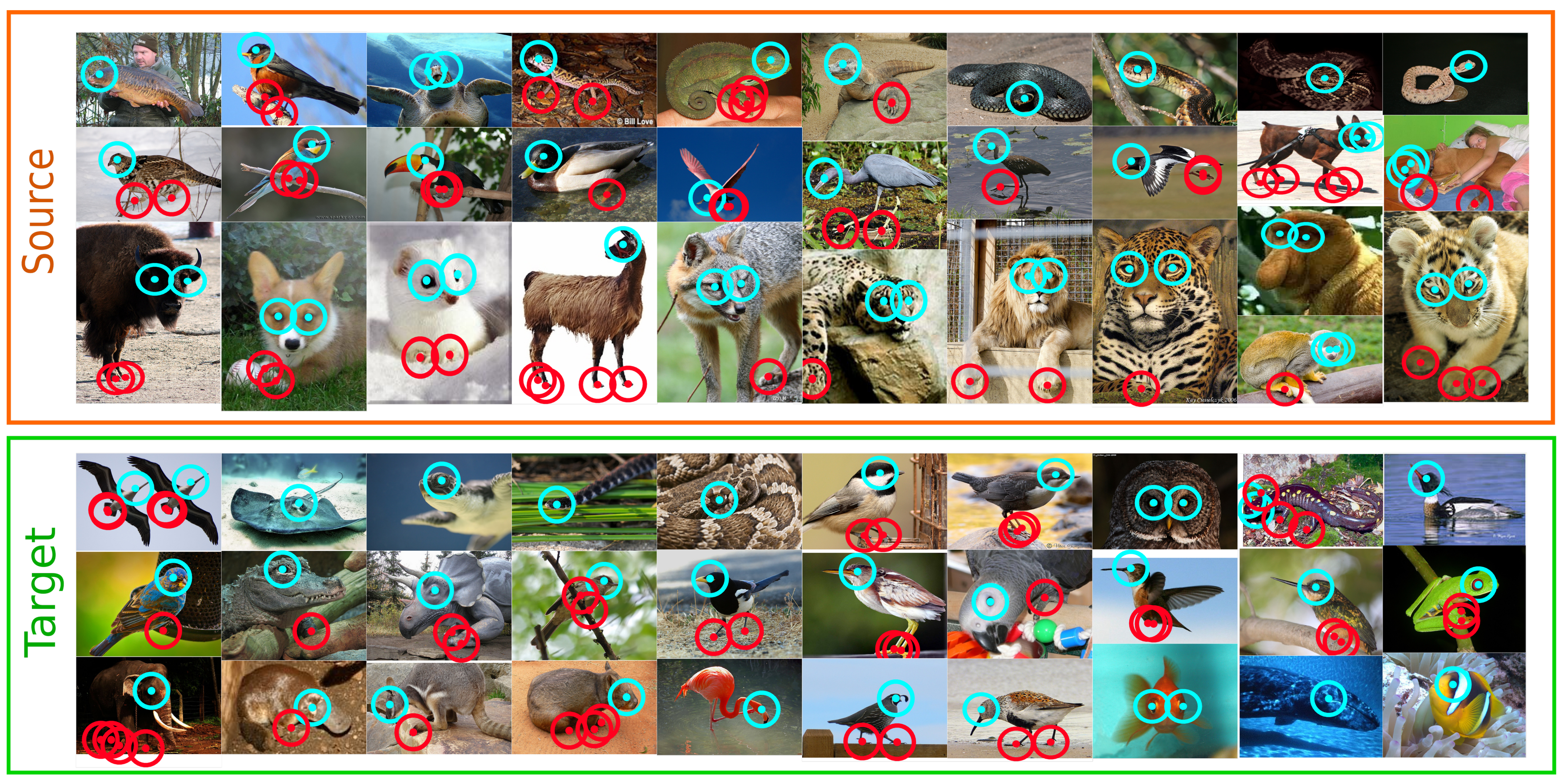
→ Train a part detector on **source categories**, test on held-out **target categories**

→ Find the saturation point of the detector performance on held-out classes

→ Additional analysis: Evaluation of part transfer between categories

A novel Animal Parts dataset for studying transfer learning problems:

- Animal "**eye**" and "**foot**" keypoint annotations
- Annotated ∼15K ImageNet images of "vertebrate" animals



Related problems

Active Learning [1] Studies how many images should be annotated s.t. performance saturates ASAP

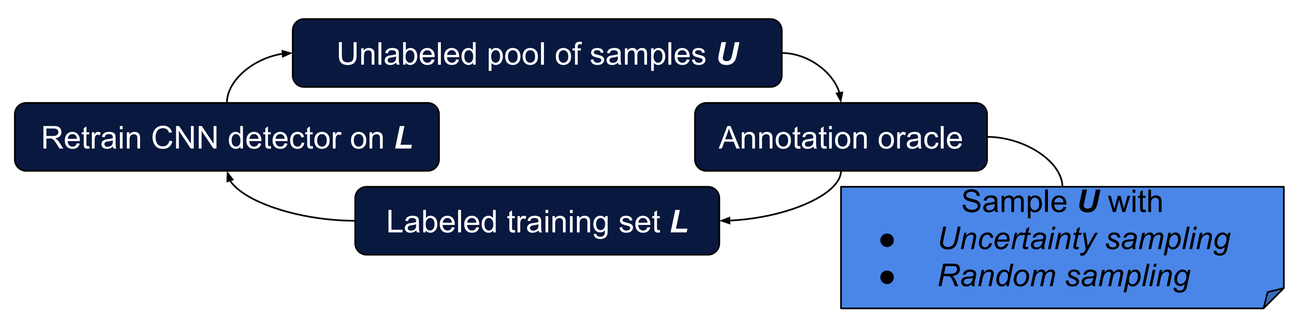
Domain adaptation [2] Animal classes ∼ domains; detecting parts ∼ the shared task

→ use the domain knowledge to improve the part detector

Proposed methods

All proposed methods utilize the CNN keypoint detector from [3]

Active learning with uncertainty/random sampling [4]

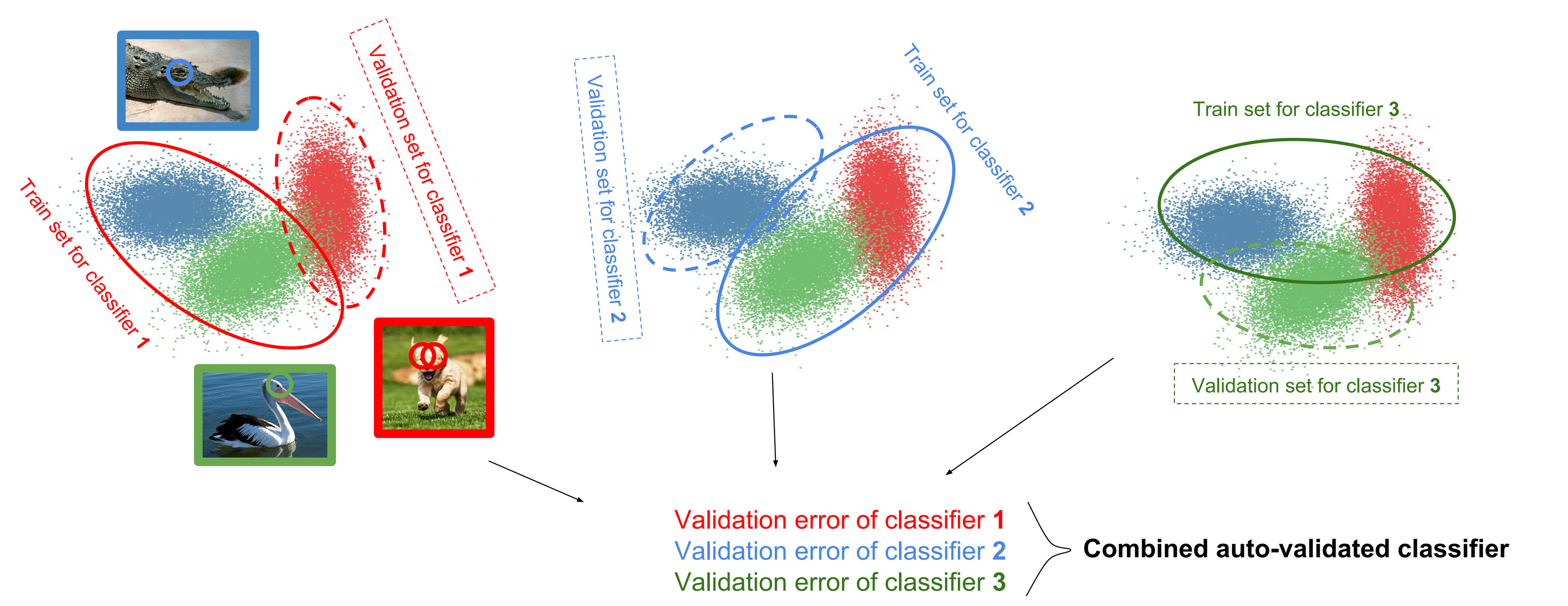


- Originally designed for simple classification problems
- Our problem is in fact an instance of Multiple Instance Active Learning
- We adapt uncertainty sampling for training CNNs in our scenario:
  - The image with the most uncertain pixel-level prediction is selected for annotation
- The first use of an actively trained CNN for image data

Active-transfer learning by auto-validation

During SGD iterations of CNN training perform simultaneously:

Train domain specific classifiers C & Online validation error estimation on C's complementary domains



- The combined classifier obtained with [5]
- The ensemble classifiers tend to disagree

→ **active sampling by picking samples on which the ensemble disagrees the most** (aka query-by-committee - QBC)

Experiments

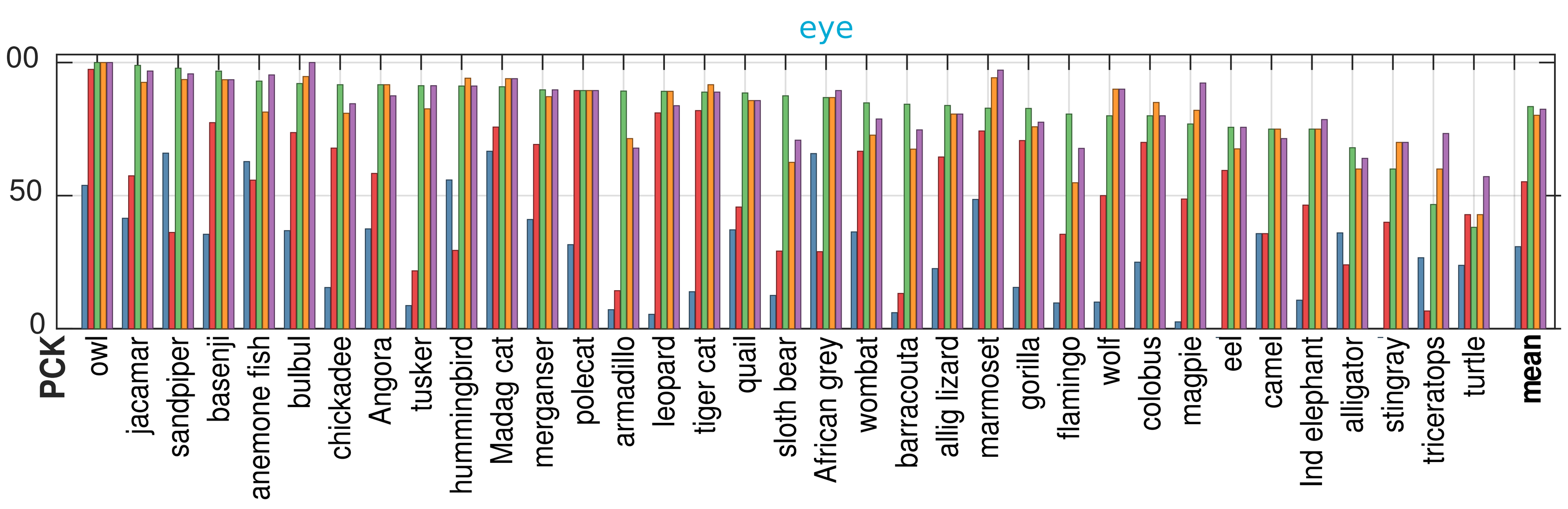
Part transferability

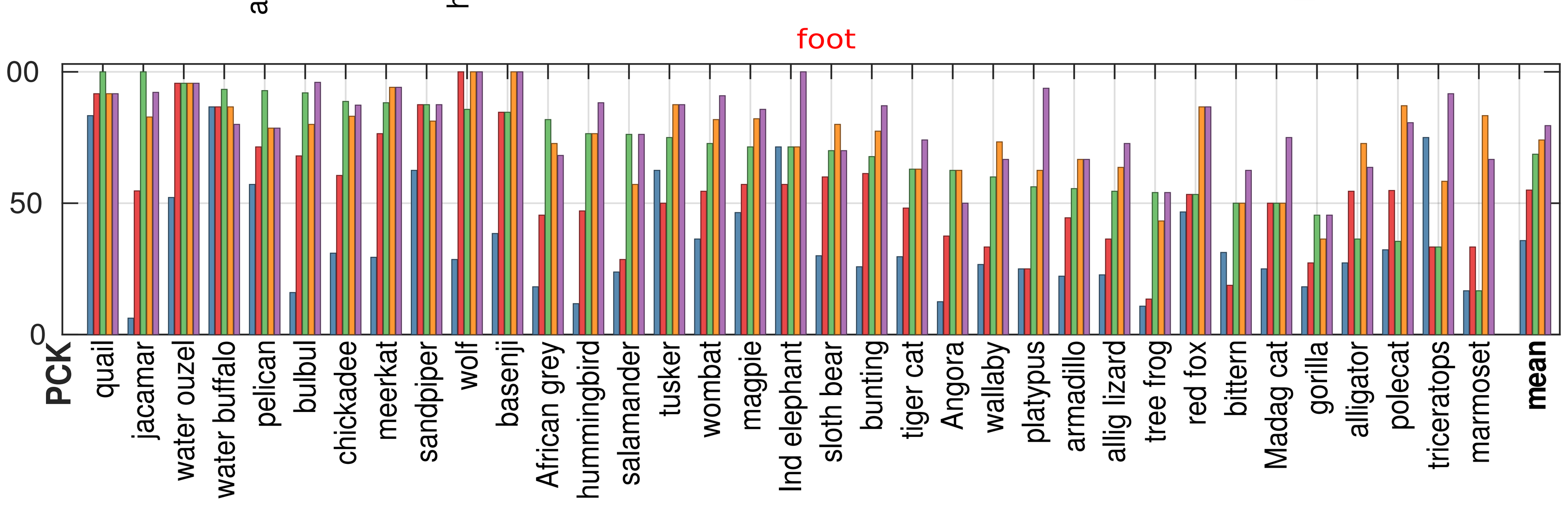
Relative difficulty of part detection and part transfer

1. **Given a target class** T, train a CNN part detector on:
   - T - the same class
   - the T's nearest class (in semantic distance)
   - the T's farthest class
   - the source classes
   - all source and target classes

2. **The part transfer performance is evaluated by reporting PCK on** T**:**

(PCK - percentage of correctly localized parts)





Training on source classes ( ) comparable to all source and target classes ( )

→ *Training on a diverse set of classes brings satisfactory transfer performance*

Performance degrades significantly when training on the farthest class ( )

→ *Demonstrates the relevance of the semantic distance for cross-domain transfer*

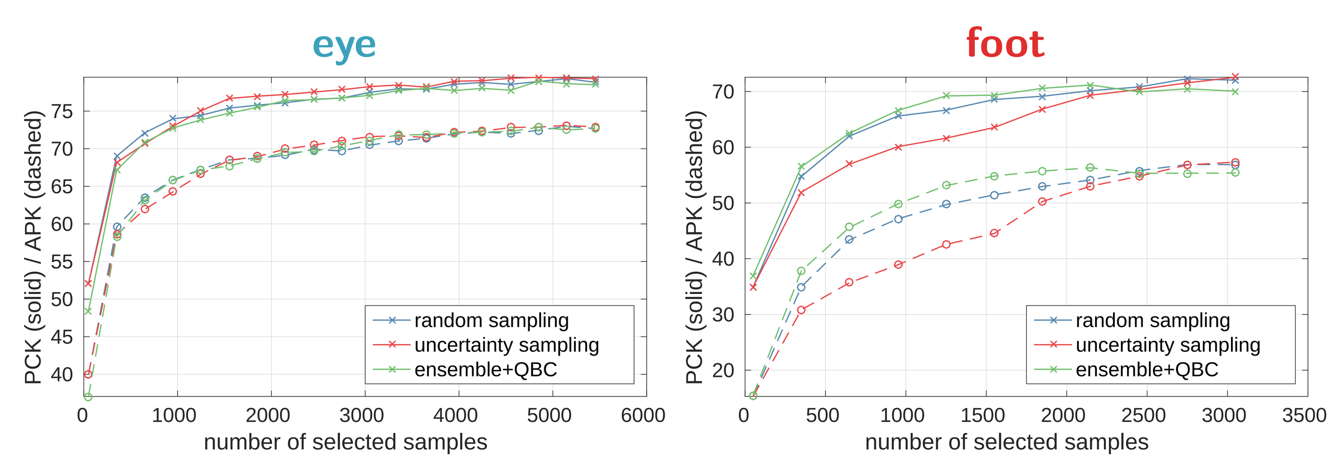
Active-transfer learning

Part detection performance as the number of annotations increases

1. Split the set of "vertebrate" classes to 50 source classes and 50 target classes
2. Gradually add source class part annotations and report detection performance on the target classes

Comparison of 3 active-transfer learning approaches: (the text in brackets denotes the label in the figure)

- Random sampling ("random sampling")
- Active learning by auto-validation ("ensemble+QBC")
- Uncertainty sampling ("uncertainty sampling")



Performance reaches 98% of the accuracy of the fully annotated scenario by providing only a few thousand examples.

→ *Excellent performance achieved with a handful of samples.*

Ensemble+QBC outperforms other methods on the "foot" keypoints, all methods on par for the "eye" keypoints

Conclusions

- Extensive testing of part transferability
  - Evaluation of part detectors in presence of a domain shift with bounded number of annotations
- Introduced a novel Animal Parts dataset
- Excellent performance achieved only from a limited number of training samples
- The proposed Ensemble+QBC active-transfer learning method outperforms other competitors on the "foot" detection task

References

[1] D. Cohn, L. Atlas, and R. Ladner, "Improving generalization with active learning," *Machine learning*, 1994.

[2] K. Saenko, B. Kulis, M. Fritz, and T. Darrell, "Adapting visual category models to new domains," in *Proc. ECCV*, 2010.

[3] S. Tulsiani and J. Malik, "Viewpoints and keypoints," in *Proc. CVPR*, IEEE, 2015.

[4] S. Tong and D. Koller, "Support vector machine active learning with applications to text classification," *JMLR*, 2002.

[5] A. Krogh and J. Vedelsby, "Neural network ensembles, cross validation, and active learning," in *Proc. NIPS*, 1995.