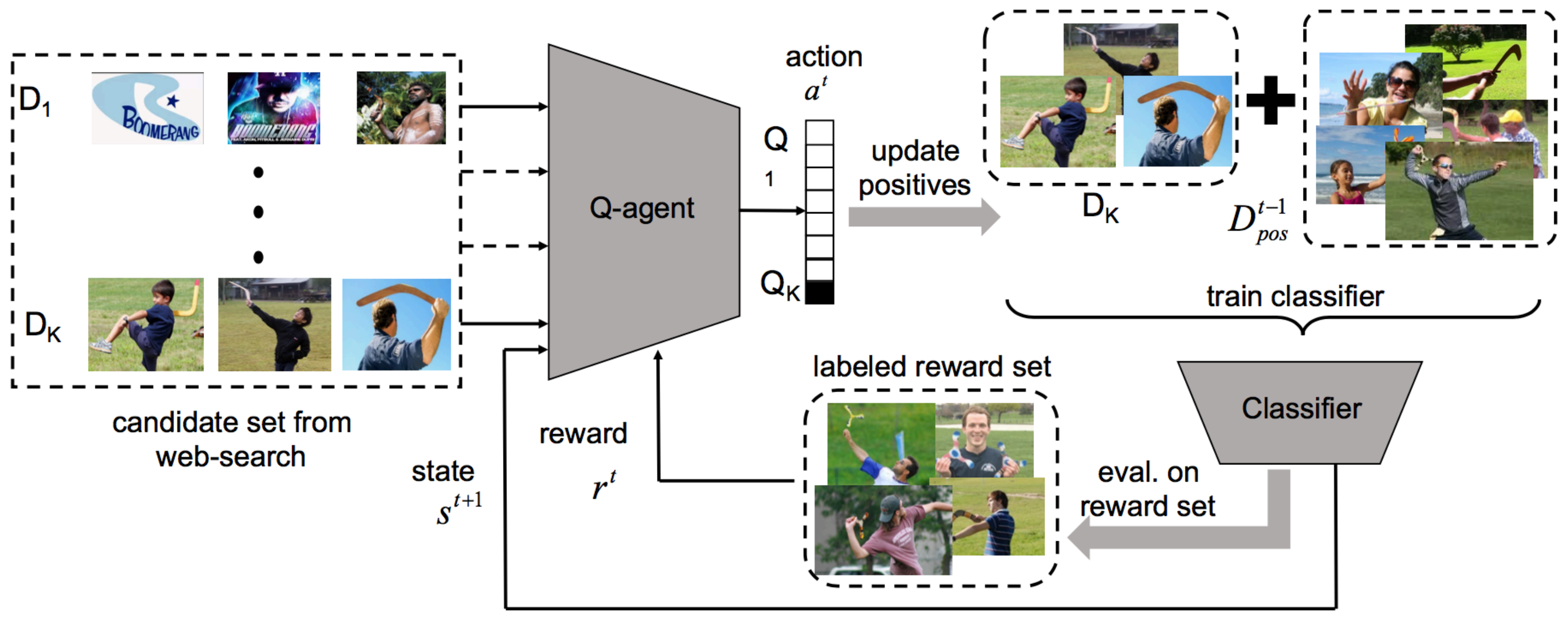
Learning to learn from noisy web videos

Serena Yeung[1], Vignesh Ramanathan[1], Olga Russakovsky[2], Liyue Shen[1] , Greg Mori[3], Li Fei-Fei[1]
[1]Stanford University [2]Carnegie Mellon University [3]Simon Fraser University

Introduction

- Manually labeling training videos for action recognition is impractical to scale to the long-tail of action categories: e.g. fine-gained, rare, or niche classes.
- We can leverage noisy data from web queries to learn new actions, using semi-supervised or "webly-supervised" approaches. However, existing methods typically do not learn and leverage domain-specific knowledge, or rely on iterative hand-tuned data labeling policies.
- Our insight is that good labeling policies can be learned from existing annotated datasets. A good policy should label noisy data such that a classifier trained on the labels would achieve high classification accuracy on the existing datasets.
- We propose a reinforcement learning-based formulation for learning data labeling policies from noisy web data. Concretely, we introduce a joint formulation of a Q-learning agent and a class recognition model. The agent selects web search examples to label as positives, which are then used to train the recognition model.

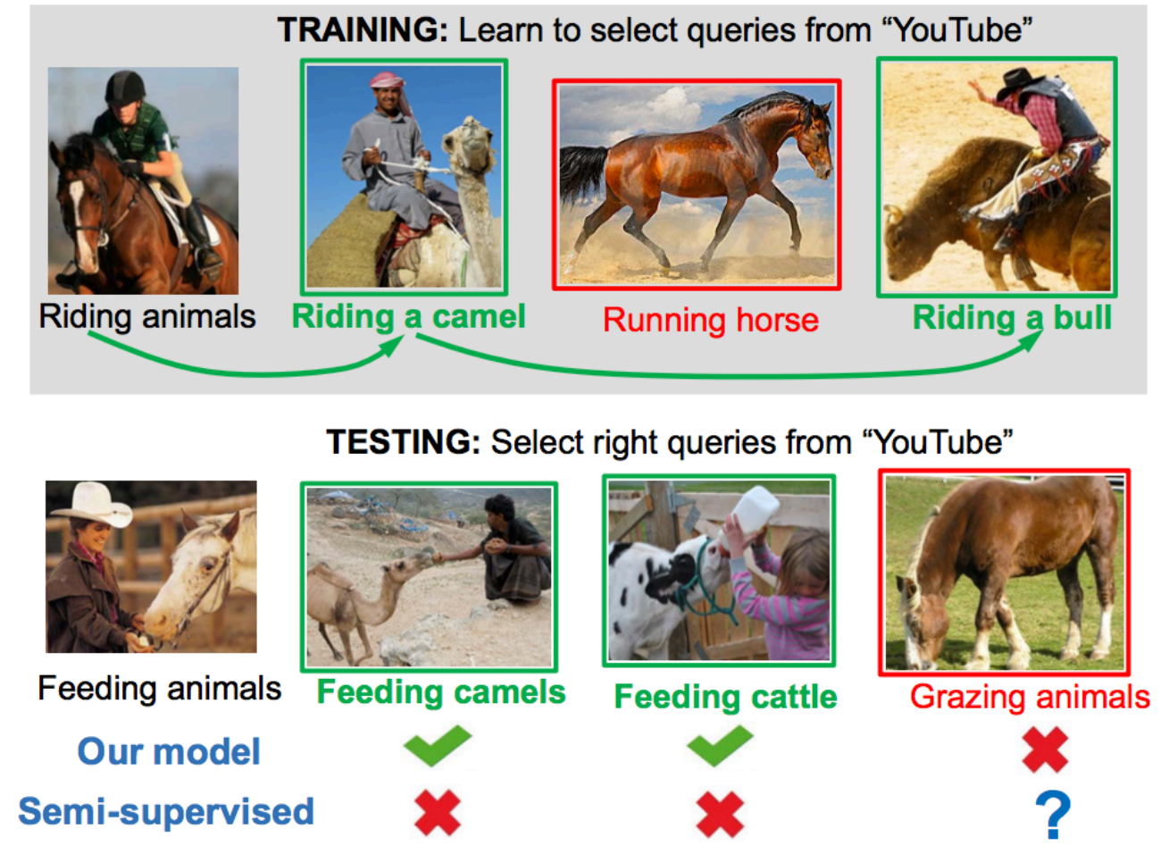
Model



Overview of model. We learn a classifier for a given visual concept using a candidate set of examples obtained from web search. At each time step t we use the Q-learning agent to select examples, e.g., D_K, to add to the existing set of positive examples D_{pos}^{t-1}.The examples are then used to train a visual classifier. The classifier both updates the agent's state s^{t+1} and provides a reward r^t . At test time the trained agent can be used to automatically select positive examples from web search results for any new visual concept.

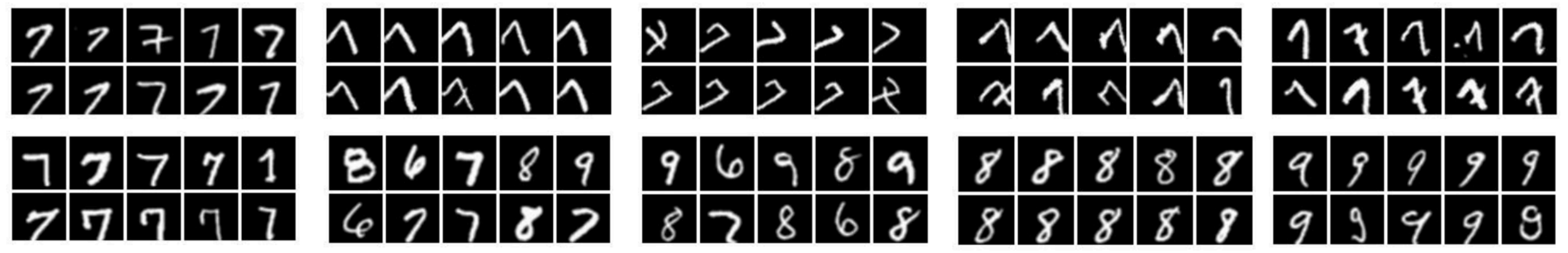
State representation: based on the distribution of classifier scores output by the recognition model. $s = \{H_{pos}, H_{pos}, \{H_{D_1}, \dots, H_{D_K}\}, \mathrm{P}\}$, where $\{H_{pos}, H_{pos}, \{H_{D_1}, \dots, H_{D_K}\}\}$ are the histograms of classifier scores for the positive set, the negative set, and each candidate subset, respectively. P is the proportion of the desired number of positives already obtained.

Reward: the change in the classifier's accuracy at time t after updating its positive set with the newly chosen examples D_{a_t}. Accuracy is computed on the held-out annotated data D_{reward}.



An annotated dataset is used to learn a policy for how to label data for new, unseen classes. This allows learning domain-specific knowledge and how to select diverse exemplars while avoiding semantic drift.

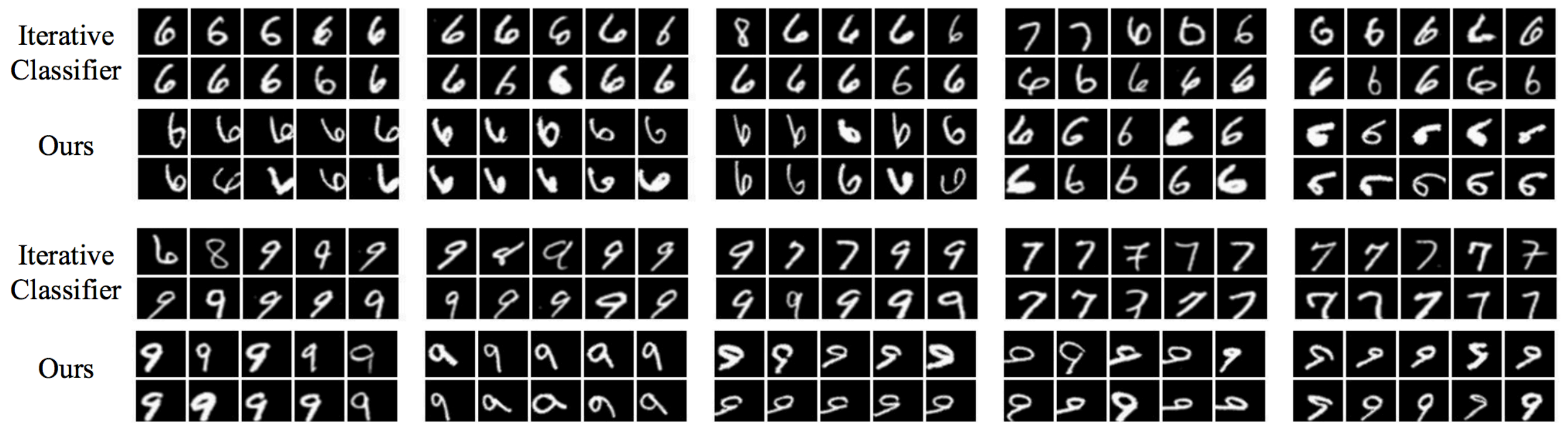
Noisy MNIST digit classification



Ten sample query subsets in Noisy MNIST for the digit 7. *Top row.* Different translation and rotation transformations. *Bottom row.* The two leftmost queries have different amounts of noise, the center one is a mixture bucket, and the rightmost two are different digits.

Digit	Budget	Seed	Label propagation	Label spreading	TSVM	Iterative classifier	Ours
6	60	42.6	37.9	41.1	39.5	43.1	**60.9**
	80	42.6	40.8	45.6	44.4	43.2	**61.3**
	100	42.6	42.2	46.7	46.2	42.4	**71.4**
7	60	48.4	51.1	48.6	46.1	49.7	**55.1**
	80	48.4	48.8	48.5	42.6	48.5	**57.6**
	100	48.4	48.1	46.6	39.7	47.4	**55.7**
8	60	39.1	35.0	35.2	41.2	38.3	**56.2**
	80	39.1	40.0	34.1	39.6	39.6	**55.6**
	100	39.1	42.0	30.2	40.8	38.0	**55.5**
9	60	37.9	37.5	36.5	41.4	52.4	**52.4**
	80	37.9	37.9	37.4	38.9	53.5	**53.5**
	100	37.9	38.0	37.6	39.5	55.7	**55.7**
All digits	60	42.0	40.4	40.3	42.1	43.4	**56.1**
	80	42.0	41.9	41.4	41.4	43.4	**57.0**
	100	42.0	42.6	40.3	41.5	42.3	**59.5**

AP on Noisy MNIST, with budgets of 60, 80 and 100 for the numbers of positive examples selected from D_{cand}.



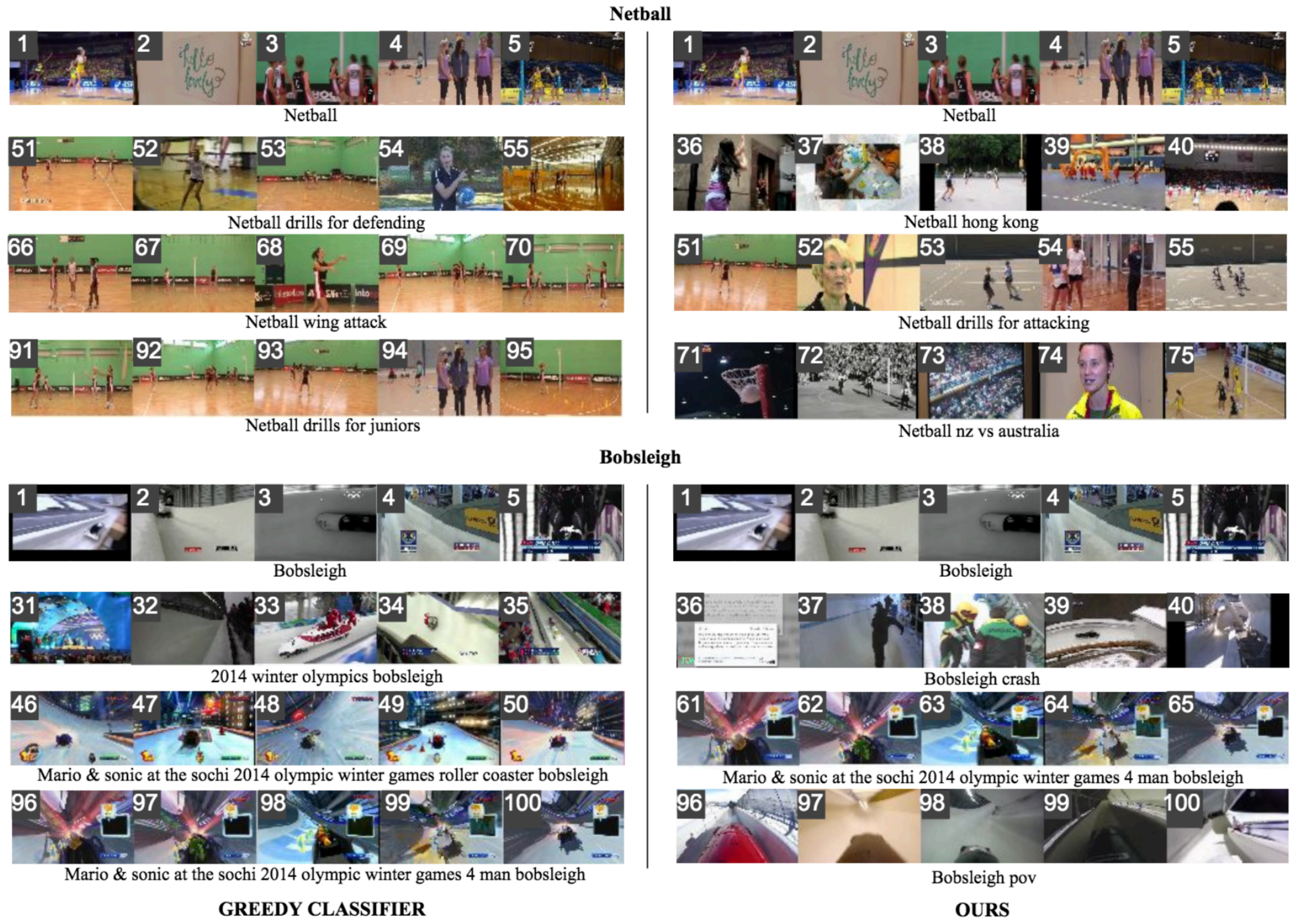
Comparison of positive query subsets selected using our method versus the greedy classifier baseline. Subsets chosen by each method are shown from left to right, for the digit 6 (top example) and the digit 9 (bottom example). Our model is better able to select useful positives with visual diversity, while avoiding semantic drift.

Sports-1M action recognition

- **Training classes (used to learn policy):** 300 Sports-1M classes
- **Test classes:** 105 Sports-1M classes
- Policy labels noisy YouTube data, using videos returned by the YouTube query suggestion feature for 30 different query expansions per class.
- At training time of learning the policy, rewards are based on classification accuracy, where classifiers are trained on the policy-labeled noisy data and evaluated on the annotated reward dataset (Sports-1M test videos for the 300 classes).
- To evaluate the learned policy, classifiers are trained on policy-labeled noisy data for the 105 previously unseen Sports-1M test classes, and evaluated on annotated Sports-1M test videos for these classes.

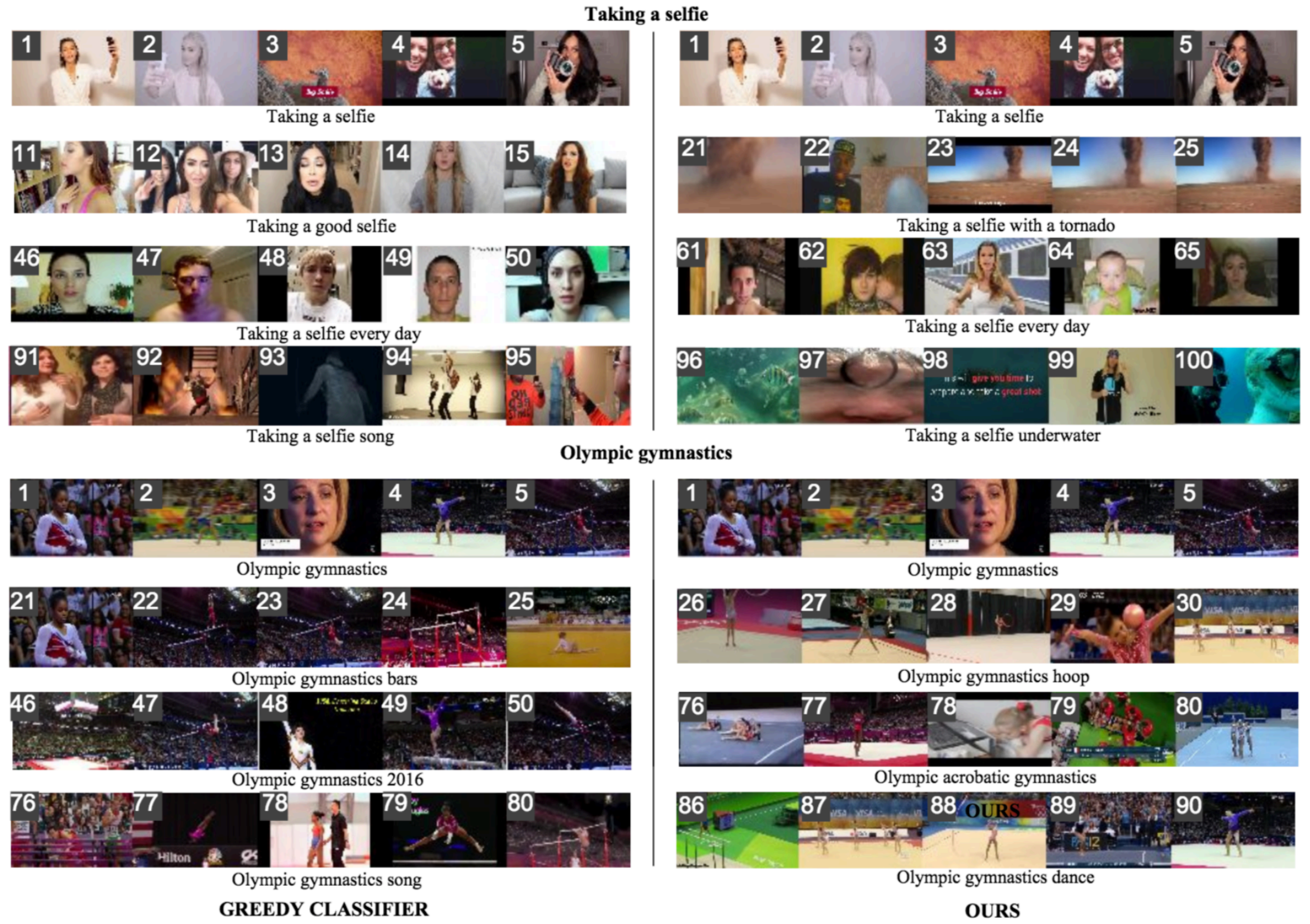
Method	Budget-60	Budget-80	Budget-100
Seed	64.3	64.3	64.3
Label propagation	65.4	65.4	67.2
Label spreading	65.4	66.6	67.3
TSVM	70.7	71.7	72.5
Greedy	71.7	73.8	74.8
Greedy-clustering	72.3	73.2	74.3
Greedy-KL	74.1	74.7	74.7
Ours	**75.4**	**76.2**	**77.0**

mAP on Sports-1M with different budgets for the number of selected positive examples.


Netball
Netball
Netball drills for defending
Netball wing attack
Netball drills for juniors
Netball
Netball hong kong
Netball drills for attacking
Netball nz vs australia
Bobsleigh
Bobsleigh
2014 winter olympics bobsleigh
Mario & sonic at the sochi 2014 olympic winter games roller coaster bobsleigh
Mario & sonic at the sochi 2014 olympic winter games 4 man bobsleigh
Bobsleigh
Bobsleigh crash
Mario & sonic at the sochi 2014 olympic winter games 4 man bobsleigh
Bobsleigh pov
GREEDY CLASSIFIER
OURS

Comparison of positive query subsets selected using our method versus the greedy classifier baseline, for two classes. Rather than show all 100 selected videos here, we highlight interesting differences. Each row shows a selected query subset (query phrase and corresponding 5 videos), with the numerical position of the selection out of 100. The first row shows seed videos. *Top example.* The greedy classifier chooses many similar-looking examples, while our method learns that examples of the action in different environments are useful positives. *Bottom example.* The greedy classifier drifts from bobsleigh to video games, while our method is robust to semantic drift and selects useful subcategories of bobsleigh videos such as crashes and pov.

Long-tail action labeling



Comparison of positive query subsets selected using our method versus the greedy classifier baseline, for two long-tail classes. See Sports-1M figure for explanation of figure structure. *Top example.* The greedy classifier selects many similar-looking examples of taking a selfie, while our method learns domain-specific knowledge that positives in different environments are more useful, e.g. with a tornado or underwater. *Bottom example.* The greedy classifier selects similar examples of gymnastics, whereas our method selects visually distinct subcategories.

References

[1] X. Zhu and Z. Ghahramani. Learning from labeled and unlabeled data with label propagation. In International Conference on Machine Learning, 2002.
[2] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Scholkopf. Learning with local and global consistency. In Neural Information Processing Systems, 2004.
[3] T. Joachims. Transductive inference for text classification using support vector machines. In International Conference on Machine Learning, 1999.
[4] X. Chen and A. Gupta. Webly supervised learning of convolutional networks. In International Conference on Computer Vision, 2015
[5] L.-J. Li and L. Fei-Fei. OPTIMOL: automatic online picture collection via incremental model learning. International Journal of Computer Vision, 88(2):147–168, 2010.