RALF: A **R**einforced **A**ctive **L**earning **F**ormulation for Object Class Recognition

max planck institut informatik

Sandra Ebert, Mario Fritz, and Bernt Schiele

UNIVERSITÄT DES SAARLANDES

Motivation

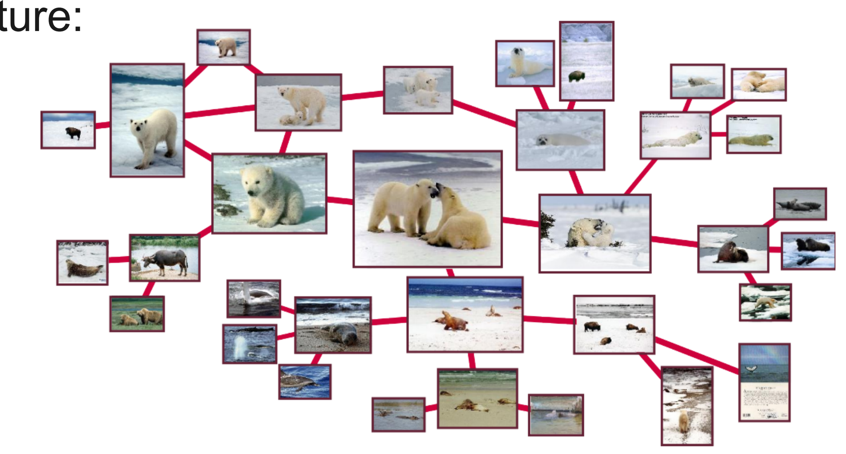
- Active learning to reduce the amount of labels
- Find representative labels for semi-supervised learning

Open questions:

- Which active learning criteria should be used?
- What is a good trade-off between exploration and exploitation?
- How can we find the right strategy for a dataset without any prior knowledge?

New exploration criteria

- Finds dense regions in a k-nearest neighbor graph structure:



- Sum over edges (W) normalized by the number of all edges (P) per node:

$$Gra(x_i) = \frac{\sum_i W_{ij}}{\sum_i P_{ij}}$$

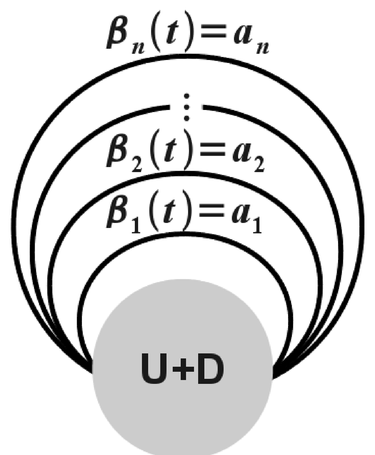
- Down weighting of the neighboring edges after selection to avoid oversampling of the same regions:

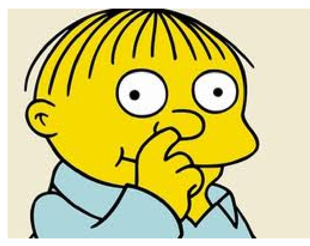
$$Gra(x_j) = Gra(x_j) - Gra(x_i)P_{ij}$$

RALF: Reinforced Active Learning Formulation

Contributions:

1) Consider active learning as a Markov decision process (MDP)
2) Any number of criteria and trade-offs possible
3) Adapts during the learning process to each specific dataset without any prior knowledge

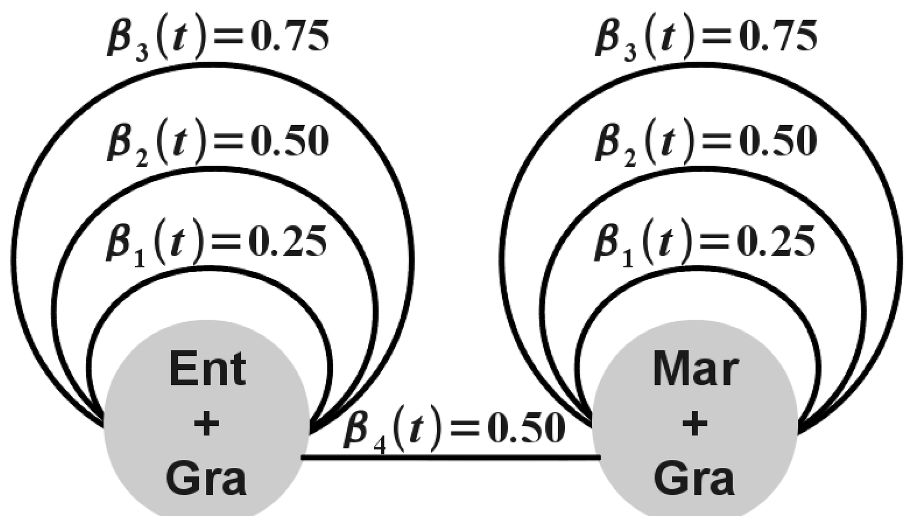




[6]

1) **Markov decision process** (MDP) to learn the best strategy for each dataset

- States: mixtures of criteria
- Actions: trade-offs or switches among states
- Any number of states and actions possible, e.g., 3 criteria and 3 different trade-offs:



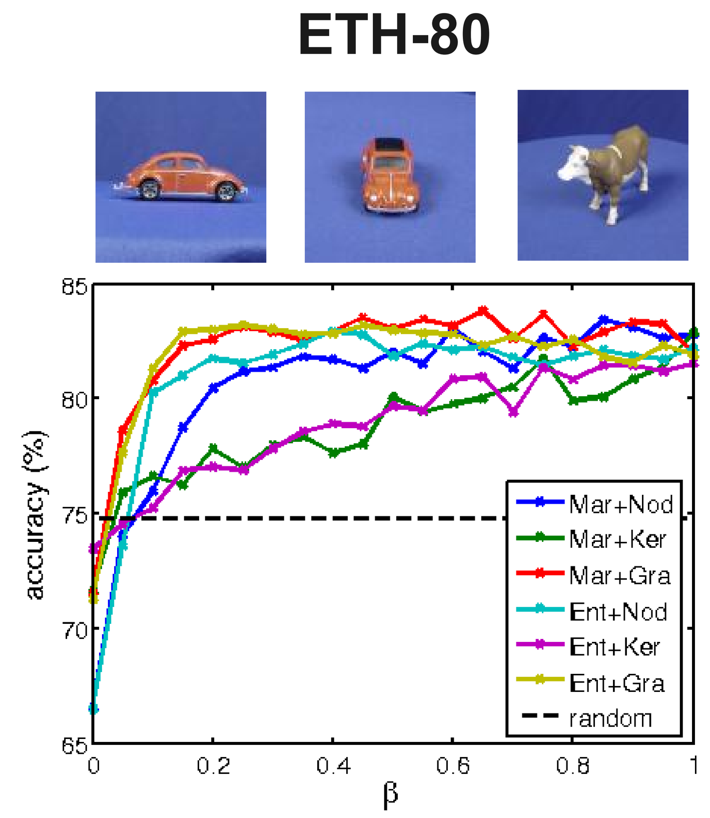
2) **Q-Learning** - a fast feedback-driven reinforcement learning algorithm to learn this MDP:

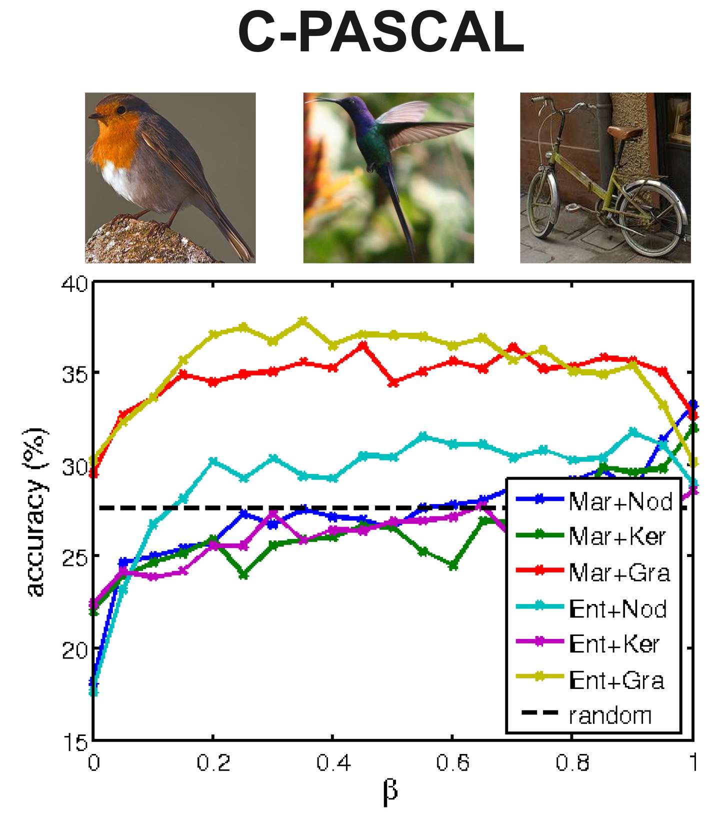
$$Q(s^{(t-1)},a) \leftarrow Q(s^{(t-1)},a) + \lambda\left(r^{(t)} + \gamma \max_{a_i} Q(s^{(t)},a_i) - Q(s^{(t-1)},a)\right)$$

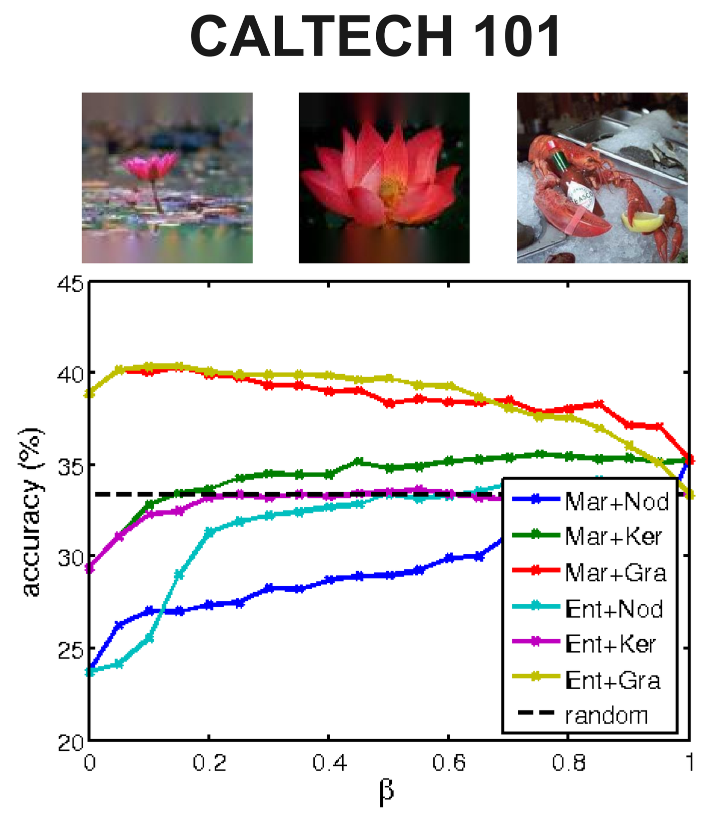
- Q table serves as a knowledge base and is updated after each iteration
- Reward r based on entropy minimization
- Parameter learning rate λ and discount factor γ are the same across all datasets

Results for different sampling criteria and trade-offs

- Active learning framework: $H(x_i) = \beta U(x_i) + (1-\beta)D(x_i)$
  - Exploitation $U \in \{Ent, Mar\}$ with **Ent**ropy [1,2,5] and **Mar**gin [4],
  - Exploration $D \in \{Nod, Ker, Gra\}$ with **Ker**nel farthest first [1], **Nod**e potential [2], and our novel **Gra**ph Density
- Comparison of several mixtures of criteria and different trade-offs $\beta \in [0,1]$



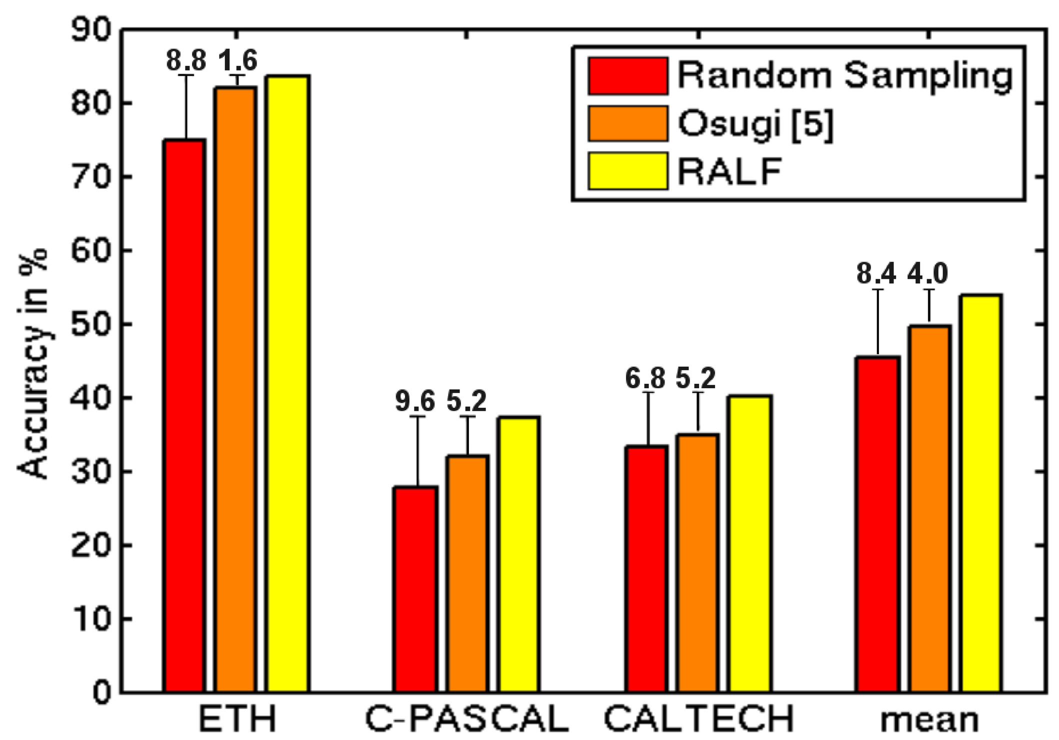




Conclusion:

- Our new exploration criteria **Gra**ph density works always best in combination with an exploitation criteria
- Single criteria < fixed trade-off < time-varying trade-off (see paper) < adaptive trade-off (see RALF)
- Each dataset need a different trade-off and different mixture of criteria

Results with RALF



- up to 9.6% improvement to random sampling
- up to 5.2% to previous work [5]

Conclusion

- New exploration criteria **graph density** that performs best among previous exploration criteria
- Best strategy is dataset dependent and time-varying
- Novel active learning formulation **RALF** that adapts the sampling strategy during the learning to each specific dataset without any prior knowledge

(QR-)Code and references:

[1] Y. Baram et.al. JMLR, 2004
[2] N. Cebron, M.R. Berthold, DMKD, 2009
[3] S. Ebert et.al., ECCV 2010
[4] A. Joshi et.al., CVPR, 2009
[5] T. Osugi, S. Scott, ICDM, 2005
[6] R. Wiggum. The Simpsons, 1989

http://www.d2.mpi-inf.mpg.de/content/ralf