RALF: A Reinforced Active Learning Formulation for Object Class Recognition

Sandra Ebert, Mario Fritz, and Bernt Schiele

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:
  \[ \text{Gra}(z_i) = \sum_j W_{ij} \]
- Sum over edges \( W \) normalized by the number of all edges \( P \) per node:
  \[ \text{Gra}(z_i) = \frac{\sum_j W_{ij}}{\sum_j P_{ij}} \]
- Down weighting of the neighboring edges after selection to avoid oversampling of the same regions:
  \[ \text{Gra}(z_j) = \text{Gra}(z_j) - \beta \text{Gra}(z_i) P_{ij} \]

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

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 table serves as a knowledge base and is updated after each iteration
- Reward \( r \) based on entropy minimization
- Parameter learning rate \( \lambda \) and discount factor \( \gamma \) 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 \( U \in \{ \text{Ent, Mar} \} \) with Entropy [1,2,5] and Margin [4],
- Exploration \( D \in \{ \text{Nod, Ker, Gra} \} \) with Kernel farthest first [1], Node potential [2], and our novel Graph Density
- Comparison of several mixtures of criteria and different trade-offs \( \beta \in [0, 1] \)

Results with RALF

- ETH-80
- C-PASCAL
- CALTECH 101

Conclusion:

- Our new exploration criteria Graph 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

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

References:

[1] Y. Baram et al., JMLR, 2004

QR-Code and references: http://www2.mpi-inf.mpg.de/content/ralf