IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, VOL. 5, NO. 3, JUNE 2011

A Survey of Active Learning Algorithms for Supervised Remote Sensing Image Classification

Devis Tuia, *Member, IEEE*, Michele Volpi, *Student Member, IEEE*, Loris Copa, Mikhail Kanevski, and Jordi Muñoz-Marí

Josu Maiora

Overview

- Defining an efficient **training set** →Fundamental phase for classification
- Active learning aims at building efficient training sets by **iteratively improving** the model performance through sampling.
- A user-defined heuristic ranks the unlabeled pixels according to a function of the **uncertainty**
- This paper reviews and tests the main families of active learning algorithms:
  1. committee,
  2. large margin,
  3. posterior probability-based

1. CONCEPTS AND DEFINITIONS

Algorithm 1: General active learning algorithm

Inputs

— Initial training set $X^{\epsilon} = \{\mathbf{x}_i, y_i\}_{i=1}^{l}$ ($X \in \mathcal{X}$, $\epsilon = 1$).
— Pool of candidates $U^{\epsilon} = \{\mathbf{x}_i\}_{i=l+1}^{l+u}$ ($U \in \mathcal{X}$, $\epsilon = 1$).
— Number of pixels q to add at each iteration (defining the batch of selected pixels S).

repeat
  Train a model with current training set X^{ϵ}.
  for each candidate in U^{ϵ} **do**
    Evaluate a user-defined *heuristic*
  end for
  Rank the candidates in U^{ϵ} according to the score of the heuristic
  Select the q most interesting pixels. $S^{\epsilon} = \{\mathbf{x}_k\}_{k=1}^{q}$
  The user assigns a label to the selected pixels. $S^{\epsilon} = \{\mathbf{x}_k, y_k\}_{k=1}^{q}$
  Add the batch to the training set $X^{\epsilon+1} = X^{\epsilon} \cup S^{\epsilon}$.
  Remove the batch from the pool of candidates $U^{\epsilon+1} = U^{\epsilon} \setminus S^{\epsilon}$
  $\epsilon = \epsilon + 1$
until a stopping criterion is met.

2. COMMITTEE-BASED ACTIVE LEARNING

- The first family of active learning methods quantifies the uncertainty of a pixel by considering a committee of learners

3. LARGE-MARGIN-BASED ACTIVE LEARNING

- The second family of methods is specific to margin-based classifiers (SVM)
  - the points more likely to become support vectors are the ones lying within the margin of the current model

3. POSTERIOR PROBABILITY BASED ACTIVE LEARNING

- The third class of methods uses the estimation of posterior probabilities of class membership (i.e.,$p(y \mathbf{x})$) to rank the candidates.

4. DATASETS

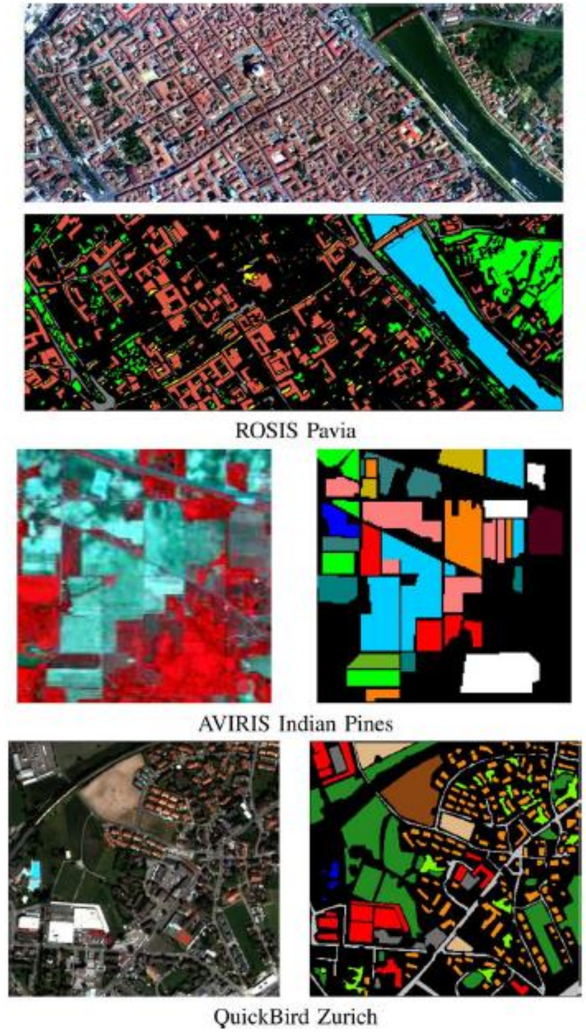
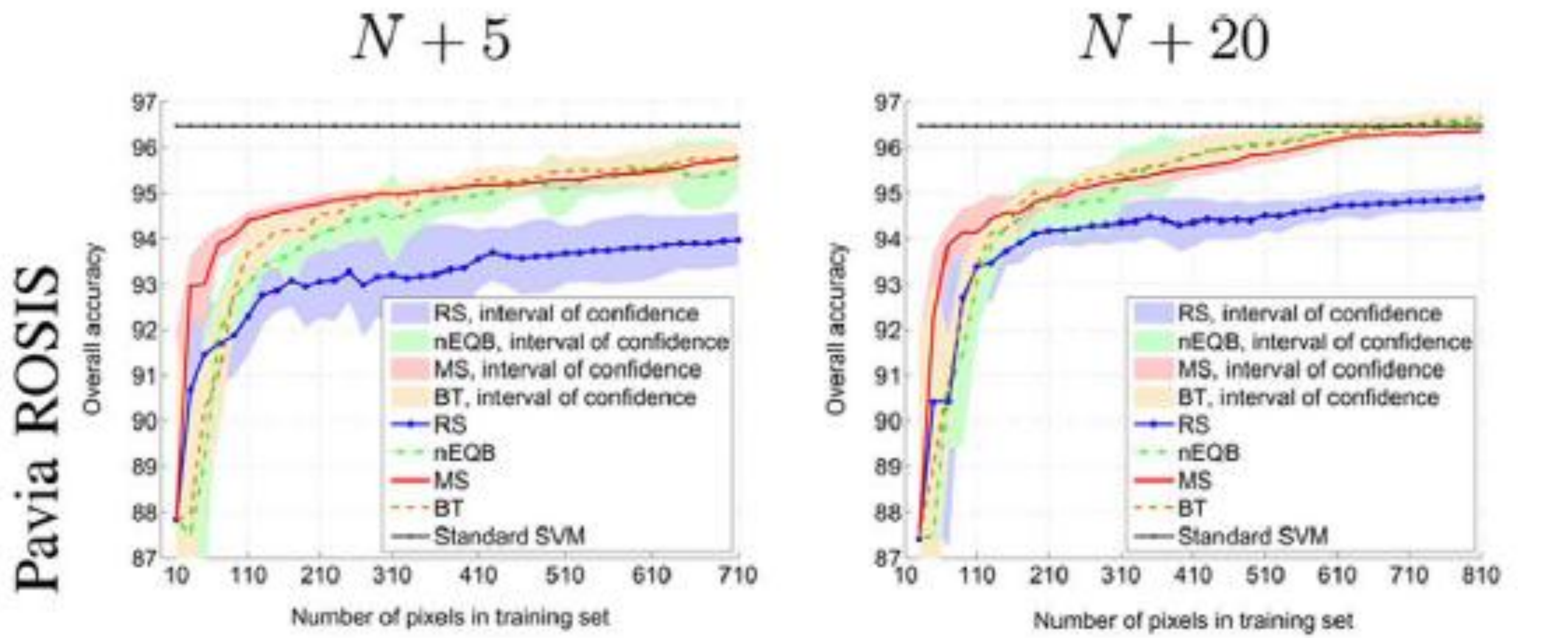


Fig. 2. Images considered in the experiments: (top) ROSIS image of the city of Pavia, Italy (bands [56 − 31 − 6] and corresponding ground survey); (middle) AVIRIS Indian Pines hyperspectral data (bands [40 − 30 − 20] and corresponding ground survey); (bottom) QuickBird multispectral image of a suburb of the city of Zurich, Switzerland (bands [3 − 2 − 1] and corresponding ground survey).

5. EXPERIMENTAL SETUP

- In the experiments, **SVM classifiers with RBF kernel** and LDA classifiers have been considered for the experiments.
- When using SVM, free parameters have been optimized by **five-fold cross validation** optimizing an accuracy criterion.
- The active learning algorithms have been run in two settings, adding **N+5 and N+20** pixels per iteration.

6. NUMERICAL RESULTS



5. CONCLUSION

- A series of heuristics have been classified by their characteristics into three families.

- Active learning has a strong potential for remote sensing data processing.

- Some recent examples can be found in the active selection of unlabeled pixels for semi-supervised classification.