TARGET DETECTION WITH SPATIO-SPECTRAL DATA VIA CONCORDANCE LEARNING

M. Murat Dundar

Purdue University School of Science
Computer & Information Science Department
Indiana University - Purdue University (IUPUI)
Indianapolis, IN 46022

1. INTRODUCTION

In this study we propose a learning algorithm for target detection in hyperspectral imagery. More specifically we aim to detect cars in the parking lots of an urban imagery. What makes this problem challenging is the copresence of several parking garages and parking lots in the same imagery. Both the cars in the parking lots and in the parking garages present with similar spectral characteristics. Therefore a learning algorithm based only on the spectral data does not uniquely identify cars in the parking lots. Even though the spectral characteristics are quite similar, the immediate spatio-spectral background surrounding the cars differs significantly between parking lots and parking garages, i.e. parking garages are made of concrete whereas parking lots are made of asphalt.

2. PROPOSED APPROACH

To deal with this problem we first characterize each pixel in the hyperspectral imagery with its spectral and spatio-spectral representations and then design two binary classifiers one for each of the representations. During real-time classification a pixel is confirmed as target when outputs of both classifiers concord as positive, i.e. when spectral classifier indicates a pixel for a car *and* spatio-spectral classifier indicates a pixel for a parking lot.

This study mainly focuses on the training of these two classifiers. Our approach proposes to jointly optimize these two classifiers by minimizing the total cost of discordance between the output of the classifiers while making sure both classifiers are sufficiently well-regularized to prevent overfitting. Unlike traditional learning algorithms where the cost of misclassification rate is minimized during training, the proposed concordance scheme minimizes the total cost of discordance between the outputs of the classifiers. Since during online execution concordance is sought between the outputs of the classifiers before a pixel is confirmed as target, by imposing the cost of discordance on the objective function during training, the offline training and online execution objectives are better aligned for the underlying problem.

The proposed approach is based on the extension of the hyperplane classifiers with multi-variable hinge loss functions.

2.1. Hyperplane Classifiers

We are given a training dataset $\{(x_i, y_i)\}_{i=1}^{\ell}$, where each pixel s_i is characterized by a feature vector $x_i \in \Re^d$ and $y_i \in \{-1, 1\}$ is the corresponding ground truth and ℓ is the number of samples. We consider a class of models of the form $f(x) = \alpha^T x$, with the sign of $f(x)$ predicting the label associated with the point x. An hyperplane classifier with hinge loss can be designed by minimizing the following cost function.

$$\mathcal{J}(\alpha) = \Phi(\alpha) + \sum_{i=1}^{\ell} w_i max\left(0, 1 - \alpha^T y_i x_i\right) \tag{1}$$

where the function $\Phi : \Re^{(d)} \Rightarrow \Re$ is a regularization function or regularizer on the hyperplane coefficients and $max(0, 1 - \alpha^T y_i x_i)$ represents the hinge loss, and $\{w_i : w_i \geq 0, \forall i\}$ is the weight preassigned to the loss associated with x_i. For balanced data usually $w_i = w$, but for unbalanced data it is a common practice to weight positive and negative classes differently, i.e. $\{w_i = w_+,\ \forall i \in C^+\}$ and $\{w_i = w_-,\ \forall i \in C^-\}$ where C^+ and C^- are the corresponding sets of indices for the positive and negative classes respectively.

For $\Phi(\alpha) = \|\alpha\|_2^2$, where $\|.\|_2$ is the 2-norm, (1) results in the conventional Quadratic-Programming-SVM, and for $\Phi(\alpha) = |\alpha|$, where $|.|$ is the 1-norm it yields the sparse Linear-Programming-SVM.

2.2. Concordance via Multivariable Hinge Loss

This time each pixel is characterized by two feature vectors $x_i = (\bar{x}_{i1}, \bar{x}_{i2})$, with $\bar{x}_{ik} \in \Re^{d_k}$ for $k = \{1, 2\}$ are feature vectors extracted for the spectral and spatio-spectral representations of the pixel s_i. In this framework concordance among classifiers is imposed on the objective function in (1) by using a multi-variable hinge loss function as follows.

We aim to optimize the following cost function

$$\mathcal{J}(\alpha_1, \alpha_2) = \sum_{k=1}^{2} \Phi_k(\alpha_k) + \sum_{i=1}^{\ell} w_i \max(0, e_{i1}, e_{i2}) \tag{2}$$

where $e_{ik} = 1 - \alpha_k^T y_i \bar{x}_{ik}$ defines the margin error due to the k^{th} representation of pixel s_i committed by the classifier f_k. The loss induced by pixel s_i is zero only if $\forall k : 1 - \alpha_k^T y_i \bar{x}_{ik} \leq 0$, i.e. margin error for both classifiers are zero. In other words before a pixel is assigned to one of the classes concordance among classifiers is sought. Classifiers are considered concordant for pixel s_i when they all have zero margin error on s_i.

The function in (2) is convex. We formulate this problem as a mathematical programming problem with inequality constraints and solve for α_1 and α_2.

3. EXPERIMENTAL RESULTS

The hyperspectral imagery used in this study is collected by the airborne HYMAP system on September 30, 1999 over the Purdue University West Lafayette campus. It contains 126 bands covering 0.40-2.40 μm region of the spectrum. Pixel size is about 5 meters.

Three different training approaches for the training of the two classifiers are compared. First, classifiers are trained independently. Second, feature vectors for each of the spectral and spatio-spectral representations are concatenated to train a single classifier. Third, the two classifiers are trained jointly with concordance imposed on the objective function. Varying numbers of training sample sizes are considered. Classifier paramaters w_+ and w_- are tuned by 10-fold cross validation with the training data. For each experiment the Receiver Operating Characteristics (ROC) curves are plotted on the test data and areas under the curves are recorded.

sequence order	1	2	3	4	5
ℓ	5/44	10/88	20/176	20/410	50/440
concatenation	0.82	0.81	0.91	0.93	0.95
independent	0.74	0.86	0.93	0.96	0.96
concordance	0.86	0.89	0.94	0.96	0.96

Table 1. Areas under the ROC curves obtained for three different training approaches. ℓ indicates the number of training samples used for each experiment displayed in the following format: *(number of samples from the target class)/(total number of training samples)*

4. CONCLUSIONS

Results show that the proposed concordance learning outperforms the two other training approaches for small training sample size, i.e. ℓ. Both the independent and concatenation approaches suffer from the *curse of dimensionality* for limited training data. On the other hand the proposed approach helps alleviate this problem to a greater extent. The concordance scheme when imposed during training provides robustness for the classifier.