Application of Support Vector Machines for Automatic Compliance Monitoring of the Conservation Reserve Program (CRP) Tracts

Ginto Cherian, Xiaomu Song, Guoliang Fan
School of Electrical and Computer Engineering
Oklahoma State University
Stillwater, Oklahoma 74078
Email: {ginto,xiaos,glfan}@okstate.edu

Mahesh N. Rao
Department of Geography
Oklahoma State University
Stillwater, Oklahoma 74078
Email: mahesh@okstate.edu

Abstract—We study an automatic compliance monitoring approach for examining United States Department of Agriculture (USDA)’s Conservation Reserve Program (CRP) tracts. In this work, CRP compliance monitoring is aimed at checking whether each CRP tract is compliant with contract stipulations. The proposed algorithm incorporates both one-class and two-class support vector machines (SVMs) for CRP classification. Specifically, one-class SVM (OCSVM) is first used to separate minor non-CRP outliers from the majority which is assumed to be the real CRP coverage. Then OCSVM results are used to train a two-class SVM (TCSVM) to further refine the CRP classification result. We use the CRP reference data as the baseline to evaluate CRP classification results. A high consistency between the CRP classification result and the CRP reference data indicates good compliance, while a low consistency reveals possible non-compliance. Simulation results show that the proposed method provides reliable information for CRP compliance monitoring.

I. INTRODUCTION

Two-class Support Vector Machine (TCSVM) has been widely used for remote sensing applications in recent years [1]–[3]. In particular, TCSVM was applied to classify multispectral remote sensing data, such as AVIRIS hyperspectral data [4] and Landsat imagery [5]. TCSVM algorithms search a linear separation plane that maximizes the distance between two different patterns in the feature space, leading to good generalization performance. Since TCSVM is limited to supervised classification requiring training samples, it cannot be directly applied to the unsupervised case where no training sample is available. On the other hand, one-class SVM (OCSVM) has been recently developed for outlier or novelty detection in [6]–[9]. Basically, OCSVM is an unsupervised classifier technique that tends to separate the majority of data from its outliers. It was shown that OCSVM can produce comparable or superior classification results over traditional classification methods in various applications in [8], [9].

In our previous work [10], TCSVM was applied to the automatic mapping of United States Department of Agriculture (USDA)’s Conservation Reserve Program (CRP) tracts based on Landsat imagery and other multisource GIS data. CRP is a long-term program which aims to improve soil, water and wildlife resources. Under the CRP contract, farmers are encouraged to plant long-term native plant species (mostly grasses) on agricultural land for a period of 10-15 years [11]. These CRP tracts have to be maintained according to CRP contract stipulations, which specify that the land cannot be used for commercial purposes except for weather-related emergencies. In return annual rental payments are made to the farmers. In year 2002, USDA paid around 1.6 billion dollars as annual rental payments. However USDA is faced with the problem of farmers not maintaining CRP tracts according to contract stipulations. Current methods for CRP compliance monitoring involve intensive manual inspection of aerial photographs which is time-consuming and costly. USDA’s Common Land Unit (CLU) data used for general compliance issues is generated from aerial photographs with a resolution about 1m×1m, which are updated every 1-2 years and may not be very efficient for CRP compliance monitoring in a large scale [12]. In addition, existing CRP reference data obtained from USDA’s Natural Resource Conservation Service (NRCS) is not very accurate or up-to-date for the management purpose. There is need of an automatic compliance monitoring method which can examine CRP tracts more efficiently and promptly with minimum human involvement.

In this work, we study the issue of automatic CRP compliance monitoring based on multispectral Landsat imagery with a resolution of 30m×30m and the CRP reference data obtained from NRCS. On the one hand, the CRP reference data is used to locate CRP tracts for compliance monitoring. On the other hand, the CRP reference data is also used as the baseline to evaluate real CRP classification results. We assume that the majority of a CRP tract is covered by CRP grass species as expected. The issue of CRP compliance monitoring problem is then converted to a process of separating the outliers (non-CRP coverage) from the majority of data (CRP coverage). Particularly, we propose a self-supervised classification framework with a sequential combination of both OCSVM and TCSVM. First, an OCSVM is used to perform unsupervised classification in an identified CRP clip. An initial classification result can be obtained that indicates the majority of the CRP tract as well as non-CRP outliers. This result is utilized to
train a TCSVM to further refine the classification accuracy. By comparing the CRP classification result with the CRP reference data, the classification consistency is computed and assessed. If the consistency is high, the CRP tract is considered to be compliant. Otherwise there exists higher probability that the CRP tract is not compliant.

The remainder of this paper is organized as follows. We first briefly review TCSVM and OCSVM in Section II; Then the proposed compliance monitoring algorithm is described in Section III. Section IV shows simulation results. Conclusions are drawn in Section V.

II. SUPPORT VECTOR MACHINES (SVM)

A. Two-class SVM (TCSVM)

Given training data \( \{(x_1, y_1), \ldots, (x_l, y_l)\}, x \in \mathbb{R}^n, y \in \{1, -1\} \), the TCSVM method tries to find the optimum hyperplane so that the two classes are separated with the maximum margin [13]–[15]. Or equivalently, we want to minimize \( \frac{1}{2}||w||^2 + C \sum_{i=1}^{N} \xi_i \) subject to \( y_i((x_i \cdot w) + b) \geq 1; i = 1, 2, \ldots, n \). Slack variables \( \xi_i \geq 0, i = 1, 2, \ldots, n \), are used to convert an unseparable problem into a separable one.

The kernel methods are often used to project the original feature space into a higher dimensional feature space, and a linear classification in the high dimensional feature space is equivalent to a nonlinear classification in the original feature space [13]–[15]. One widely used kernel is the Gaussian kernel defined as:

\[
k(x, x_i) = e^{-\gamma||x-x_i||^2},
\]

where \( \gamma \) determines the width of the kernel. TCSVM is widely applied to the supervised classification problem, and is extended to unsupervised case in recent works.

B. One-class SVM (OCSVM)

1) Basic Principle: The OCSVM is an extension of the general TCSVM to the unsupervised classification case [6], [8], [9]. This method tries to provide an approximation function which is able to categorize the majority of data. Basically, OCSVM tries to find the region in the projected space where the majority of data resides.

Two different approaches have been proposed for OCSVM implementation. One is Support Vector Data Description which tends to find the smallest sphere (called hypersphere) in the projected space containing majority of the data [8], [9]. Vectors lying outside the sphere are classified as outliers. In [9] more flexible classification boundaries are considered by reducing the number of support vectors. But it requires more knowledge about the training dataset for optimum performance. The other method is \( \nu \)-SVM that considers the origin of the projected feature space as the true member of the second class [6]. Thus the objective here is to separate the origin from the projected data using a hyperplane with the maximum margin \( \rho \). \( \nu \in (0, 1] \) is the parameter that decides the percentage of outliers. The hyperplane is constructed by solving:

\[
\min_{w \in \mathbb{R}^n, \xi \in \mathbb{R}^n, \rho \in \mathbb{R}^n} \frac{1}{2}||w||^2 - \nu \rho + \frac{1}{n} \sum_{i=1}^{n} \xi_i,
\]

subject to \( y_i((x_i \cdot w) + b) \geq \rho - \xi_i, i = 1, 2, \ldots, n \).

Both methods are shown to be equivalent in [6], [8], [9] when the data is normalized to the unit norm which can be done by the aid of a Gaussian kernel defined in (1). It is also shown in [8] that both methods perform best when the Gaussian kernel is used. In this work, we use the method proposed in [6] due to its simplicity.

2) Model Selection for OCSVM: Model selection aims at determining \( \nu \) in OCSVM. \( \nu \) can be ideally set to be the fraction of outliers that is unknown in our problem. A simple and effective approach for heuristically estimating \( \nu \) was proposed in [16]. The idea is to initially try out all classification results for different \( \nu \) values and then select the one that has the largest separation distance between majority and outlier. The separation distance for a particular value of \( \nu \) is computed as:

\[
D_\nu = \frac{1}{N_+} \sum_{f_\omega(x) \geq \rho} f_\omega(x) - \frac{1}{N_-} \sum_{f_\omega(x) < \rho} f_\omega(x),
\]

where \( N_- \) and \( N_+ \) are the numbers of patterns in two classes and \( f_\omega(x) = (x \cdot w) + b \) shows how likely sample \( x \) is in certain pattern. It can be seen that \( D_\nu \) provides an average estimation of the distance between two patterns in the feature space, and the \( \nu \) value resulting in the largest \( D_\nu \) is considered the one with the best separability.

III. PROPOSED ALGORITHM

In this work, we assume the majority of a given CRP tract is covered by CRP grass species. OCSVM can be used to obtain an initial classification of CRP and non-CRP areas. OCSVM usually suffers from the problem of having a huge number of support vectors and bounded support vectors. In order to get a more natural decision hyperplane, we can further use TCSVM that has a relaxed hyperplane placement. We sample the initial results from OCSVM to select some training samples for CRP and non-CRP classes, then a TCSVM is trained and used to reclassify the whole CRP region again. The flowchart of this process is illustrated in Fig. 1.

Fig. 1 shows that the first step is to select a CRP clip based on the CRP reference data, where the majority of the clip is assumed to be compliant and where some surrounding non-CRP areas are also included for consideration. In order to reduce the computational complexity of OCSVM, we uniformly sample the CRP clip at the rate of 30%, which means 30% of data samples are used for OCSVM training. The OCSVM is trained based on a given \( \nu \) value, and equation (3) is calculated to get the average interclass distance. Thus OCSVM is trained several times with different \( \nu \) values, and the one with the largest interclass distance is selected. Then the trained OCSVM is applied to the whole clip to get a segmentation map. Based on the OCSVM classification result, reliable CRP (majority) and non-CRP (outliers) training samples are carefully selected for TCSVM training. Specifically, a data sample is considered to be reliable if all data samples within its 5×5 neighborhood have the same labels. The whole CRP clip can be reclassified using the trained TCSVM. Finally, the consistency between of the final CRP classification result and
CRP reference data is computed which is used to determine if the given CRP tract is in compliance with the contract stipulations. Certain thresholds should be applied here.

IV. EXPERIMENTS AND DISCUSSIONS

A. Study Area and Experiment Setup

The study area is located in Texas County, Oklahoma, and a portion is shown in Fig. 2. This county has the largest CRP enrollments in Oklahoma. CRP reference data of Texas County obtained from NRCS is used to assist the selection of CRP clips. In this work, Landsat TM multispectral images of February and June of year 2000 are used in the simulation. The database was derived from the Landsat TM imagery in the same way as the one in [10]. The first 10 layers are bands 2, 3, 4, 5 and 7 from two seasons. The following 17 layers are made up of derived features that include Normalized Vegetation Difference Index (NDVI), Band Ratios, and Band Differences. The last 20 layers are the texture information of Landsat images which includes the local mean and local variance within a \(3 \times 3\) window of each band in each season. Altogether, there are 47 layers in the database for CRP classification, and each data sample is a vector of dimension 47.

Specifically, for the OCSVM, the Gaussian kernel with \(\gamma = 0.000001\) is used based on the cross validation result. The best \(\nu\) value is chosen from 0.1 to 0.5 which results in the maximum pattern separability. More accurate estimation of \(\nu\) can be obtained via a coarse-to-fine estimation process at the more computation complexity. The OCSVM classification is performed based on the optimal \(\nu\) value. Then as mentioned in Section III, OCSVM results are re-sampled for TCSVM training via a \(5 \times 5\) window operation. A Gaussian kernel is also used for TCSVM and the cross validation result shows that \(\gamma = 0.1\) is preferred.

### TABLE I

<table>
<thead>
<tr>
<th>Clip</th>
<th>(0.1)</th>
<th>(0.2)</th>
<th>(0.3)</th>
<th>(0.4)</th>
<th>(0.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clip 1</td>
<td>0.00287</td>
<td>0.00375</td>
<td>0.00476</td>
<td>0.00459</td>
<td>0.00186</td>
</tr>
<tr>
<td>Clip 2</td>
<td>N/A</td>
<td>0.00362</td>
<td>N/A</td>
<td>0.00278</td>
<td>0.00276</td>
</tr>
<tr>
<td>Clip 3</td>
<td>0.00082</td>
<td>0.00088</td>
<td>0.00092</td>
<td>0.00098</td>
<td>0.00105</td>
</tr>
<tr>
<td>Clip 4</td>
<td>0.00164</td>
<td>0.00229</td>
<td>0.00289</td>
<td>0.00323</td>
<td>0.00363</td>
</tr>
</tbody>
</table>

B. Simulation Results

Simulations are performed on four different CRP clips as shown in Fig. 3. These four clips are selected according to the CRP reference data, each of which includes around 1/3 non-CRP regions. For each CRP clip, the \(\nu\) value is first determined based on the maximum separation distance as shown in Table I where N/A means that no outliers are detected. Fig. 3 (a) illustrates Band-4 Landsat images of four CRP clips. Fig. 3 (b) shows the CRP reference data of these CRP clips, where the gray area is labeled as CRP and the black area non-CRP. Fig. 3 (c) shows the initial OCSVM results. Fig. 3 (d) indicates areas for TCSVM training, where light gray and black areas are considered as reliable CRP and non-CRP training samples, respectively, and dark gray areas are not involved in the TCSVM training. Fig. 3 (e) is the final classification results after using TCSVM.

In this work, the compliance monitoring is based on assessing the consistency between the CRP classification result and the CRP reference data. Specifically, Clip-1 follows well the assumption that the majority of the CRP tract is compliant. There is a strong consistence between and the CRP classification result and the reference data (96.97%). Clip-2 and 3 have some variations in terms of cover types. Therefore, the classification consistencies are not very high (Clip-2: 80.84% and Clip-3: 78.33%). On the other hand, when the majority area in a CRP tract is non-compliance, the consistence between the CRP classification result and the CRP reference data is relatively low. For example, Clip-4 shows a low classification consistency (74.08%), where there exists an active cultivation area (dark circle) previously registered as CRP. This observation implies that this area may be non-compliance and needs further inspection.
V. Conclusions

We have developed an efficient method for automatic CRP compliance monitoring, where both OCSVM and TCSVM are incorporated together to accomplish self-supervised CRP classification. Based on the existing CRP reference data, we have studied the issue of compliance monitoring by checking the consistency between the CRP reference data and the CRP classification results. Specifically, a high consistency indicates a high possibility of good compliance, and low consistency reveals the potential compliance problems. Simulation results show that the proposed method can provides a useful guidance for effective CRP compliance monitoring.

Acknowledgement

This work is supported by Oklahoma NASA EPSCoR Research Initiation Grant (2003) and Water Research Center Grant (2003, 2004), Oklahoma State University.

References