AN SVM CLASSIFIER WITH HMM-BASED KERNEL FOR LANDMINE DETECTION USING GROUND PENETRATING RADAR

Anis Hamdi, Oualid Missaoui, and Hichem Frigui

Multimedia Research Lab, CECS Dept.
University of Louisville

Abstract

We propose a landmine detection algorithm using ground penetrating radar data that is based on an SVM classifier. The kernel function for the SVM is constructed using discrete hidden Markov modeling (HMM). The kernel matrix could be obtained by defining an adequate similarity measure in the feature space. However, this approach is inappropriate as it is not trivial to define a meaningful distance metric for sequence comparison. Our proposed approach is based on HMM modeling and has two main steps. First, one HMM is fit to each of the N individual sequence. For each fitted model, we evaluate the log-likelihood of each sequence. This will result in an N x N log-likelihood similarity matrix that will be adapted to serve as the kernel of the SVM classifier. In the second step, we train an SVM classifier to learn a decision boundary between the positive and negative samples. Given a new data point, Testing is performed as follows. First we fit an HMM model to the data sample. Then, we compute the dot product of this new data point with all the existing training data. Finally, the new data sample is classified as positive or negative using the trained SVM classifier. Results on large and diverse ground penetrating radar data collections show that the proposed method identifies a meaningful and coherent distance between sequences using HMM models. Such distance measure is of paramount importance for the performance of the SVM classifier. Our initial experiments indicated that the proposed method outperforms standard HMM classifiers and SVM classifiers based on geometric distance.

Keywords

Landmine detection; Ground Penetrating Radar; Discrete Hidden Markov Models; Similarity Matrix, Kernel; Support Vector Machines.

1. INTRODUCTION

Detection, localization and subsequent neutralization of buried antipersonnel and antitank landmines is a world-wide humanitarian and military problem and has attracted several researchers in recent years. One challenge in landmine detection lies in plastic or low metal mines that cannot or are difficult to detect by traditional metal detectors. Varieties of sensors have been proposed or are under investigation for landmine detection. The research problem for sensor data analysis is to determine how well signatures of landmines can be characterized and distinguished from other objects under the ground using returns from one or more sensors. Ground penetrating radar (GPR) offers the promise of detecting landmines with little or no metal content.

Over the past few years, several landmine detection algorithms using GPR have been proposed. In general, they fall into categories: static data based algorithms and dynamic data based ones. The former category treats the raw data and transforms it to a set of static feature vectors. These static vectors are then fed to standard classification strategies. Examples include the edge histogram descriptors (EHD) which uses a K-nearest neighbors (K-NN) based classifier [1]; an SVM classifier with texture-based features [2], and an RVM classifiers [3]. The second category of landmine detection algorithms transforms the signatures into a dynamic structure, i.e., sequences of vectors taking the down-track dimension as the time variable. The underlying classifier used in this case is the Hidden Markov Model (HMM) [4].

The two categories of detection algorithms have their advantages and disadvantages. For instance, the static algorithms benefit from the simple feature representation as well as the rich set of classifiers (SVM, RVM, etc.). However, they assume that the target has a fixed shape and position within the signature and cannot tolerate scale and shift variations effectively. On the other hand, algorithms that are based on temporal data representation have the ability to accommodate for targets of different sizes and shapes. However, only one model is learned for each of the mines and clutter classes which may not be sufficient. Moreover, learning has been restricted to the HMM classifier.

In this paper, we propose a hybrid approach that combines the advantages of both categories. We use a dynamic feature representation of the alarms and an HMM structure to construct the pair-wise similarity matrix between the training alarms. This representation allows for variations in the target size and shape.

Then the constructed similarity matrix is used as a kernel matrix for an SVM classifier.

2. SVM CLASSIFICATION WITH HMM-BASED KERNEL

Let $\mathbf{O} = \{O_r;\ y_r\},\ r=1,\ldots R$, be a set of R labeled sequences where y_r is the label of sequence O_r. If $y_r = 1$ then O_r represents a mine signature, otherwise, O_r represents a clutter signature. The observation sequence O_r is the sequence of feature vectors extracted from each signature using the edge based feature extraction methodology [5]. The proposed SVM detector has the following two main components.

2.1. Feature extraction and similarity computation

Each sequence O_r , $r=1\ldots R$, is used to learn a discrete HMM (DHMM) model λ_r using the Baum-Welch learning algorithm [6]. Let $\Lambda = \{\ \lambda_r\ ,\ r=1\ \ldots\ R\}$, be the set of trained models. Even though the use of only one observation sequence to form a DHMM might lead to over-fitting, this step is only an intermediate step that aims to capture the characteristics of each observation sequence. The formed DHMM model is meant to give a maximal description of each sequence and therefore, over-fitting is not an issue in this context. In fact, it is desired that the model perfectly fits the observation sequence. It is expected then that the likelihood of each sequence with respect to its correspondent model is higher than those with respect to the remaining models.

To illustrate this step, in Fig. 1 we show the obtained similarity matrix for three groups of mines with different strengths. Each group has three signatures. The down-track B-scans (sequences of A-scans from a single channel) of the nine alarms are shown in the first row of Fig. 1. It is obvious that grouping all of these signatures in a single model (as in the standard HMM classifiers) would lead to poor generalization. Similarly, the false alarms could be caused by different clutter objects and under different environments and could have significant variations. First, we extract simple edge-based features [5], and fit one HMM for each alarm. Then, we compute the log-likelihood of each alarm in each model. The resulting pair-wise similarity matrix is displayed in Fig. 1. The bright squares correspond to high degree of similarity between the alarms and the darker ones correspond to dissimilar alarms. As it can be seen, alarms that have similar characteristics in the B-scans have high degrees of similarity.

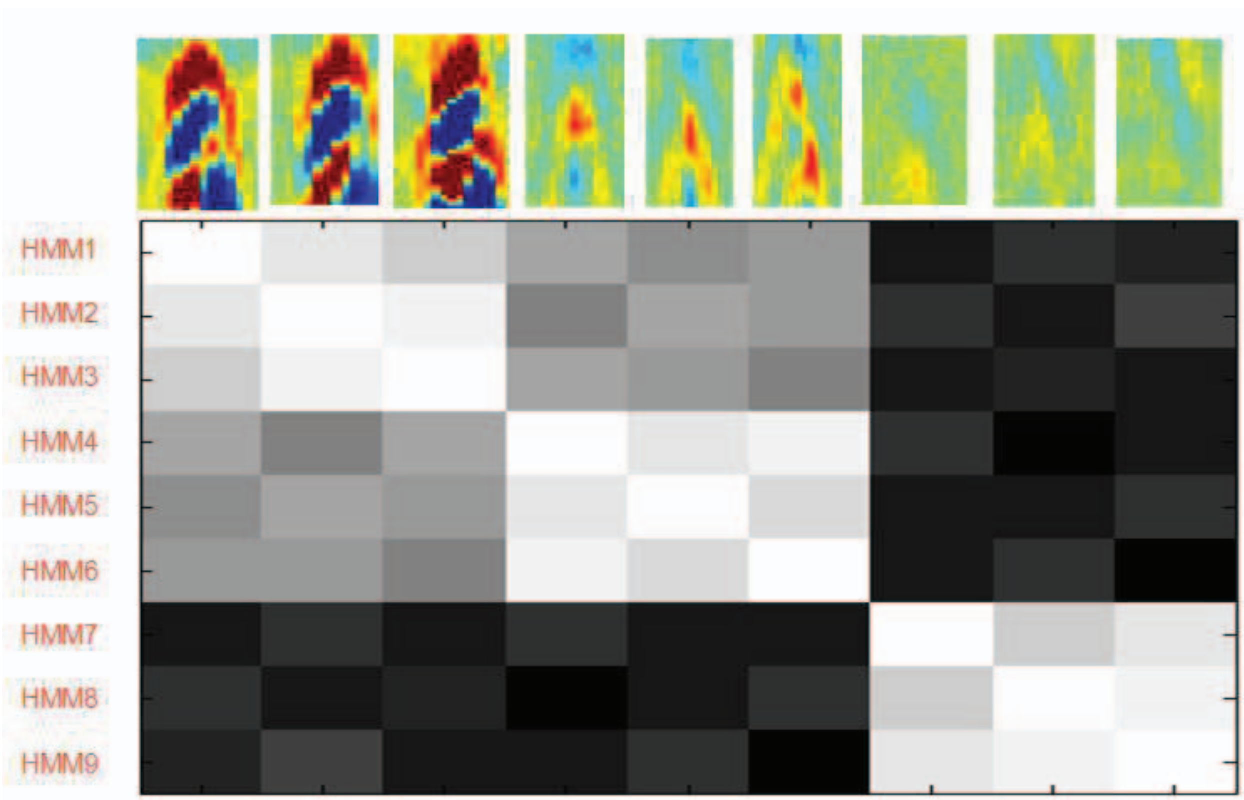

Figure 1. Similarity matrix obtained by testing each of the 9 alarms by each of the 9 models

2.2. SVM Classifier

Support vector machines (SVMs) are supervised learning algorithms that have been widely and successfully used for pattern recognition [7]. The method is also known as a “maximal margin classifier” since it determines a hyperplane that separates the two classes with the largest margin between the vectors of the two classes. Most problems in real life are however linearly not separable. SVM can deal with such problems using a kernel that transforms the feature space into a higher (possibly infinite) dimension feature space. The linearly separable hyperplane in the higher dimensional space gives a non-linear decision boundary in the original feature space.
The decision boundary of the SVM can then be expressed as:

$$f(x) = \sum_i \alpha_i y_i K(x, x_i) + b.$$

where y_i is the label of data point x_i, α_i indicates whether x_i is a support vector, and b is the y-intercept of the hyperplane described by f. The operator K is a kernel that computes the dot product of the images of x and x_i in a higher dimensional space. Several kernel has been introduced in the literature. Examples include polynomial of degree p , sigmoid, and Gaussian radial basis function [7].
In this paper we propose using a sequential kernel that takes into account the dynamic nature of the features extracted from the GPR data. In particular, we use the HMM-based similarity matrix computed in the previous section as the kernel operator.
Preliminary results on large and diverse ground penetrating radar data show that the proposed method outperforms the basic DHMM as well as the basic SVM with Gaussian kernel functions.

3. REFERENCES

[1] H. Frigui and P. Gader, “Detection and Discrimination of Land Mines in Ground-Penetrating Radar Based on Edge Histogram Descriptors and a Possibilistic K-Nearest Neighbor Classifier,” IEEE Trans. on Fuzzy Systems, Vol. 17, No.1, 2009.
[2] P. A. Torrione and L. M. Collins, "Texture features for antitank landmine detection using ground penetrating radar," IEEE Trans. on Geoscience and Remote Sensing, vol. 45, no. 7, pp. 2374-2382, 2007.
[3] D. Potin, P. Vanheeghe, E. Duflos, E., and M. Davy, “An Abrupt Change Detection Algorithm for Buried Landmines Localization “, IEEE Trans. on Geoscience and Remote Sensing, vol. 44, no. 2, pp. 260-272, 2006.
[4] H. Frigui, O. Missaoui, and P. Gader, “Landmine Detection with Ground Penetrating Radar using Discrete Hidden Markov Models with Symbol Dependent Features,” Proc, SPIE, Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XIII, 2008.
[5] H. Frigui, A. Hamdi, O. Missaoui, and P. Gader, "Landmine Detection using Mixture of Discrete Hidden Markov Models", Proc, SPIE, Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XIV, 2009.
[6] L. Rabiner, "Tutorial on hidden Markov models and selected applications in speech recognition". Proc. IEEE, 77:257–286, 1989.
[7] C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, Vol. 2, 1998, pages 121--167.