Classification of Hyperspectral Remote Sensing Data with Primal SVMs on Small-Sized Training Set Problems

M. Chi and Dr. L. Bruzzone

Dept. of Information and Communication Technologies, University of Trento (chi@dit.unitn.it / +39 - 0461-882093)

With recent technological advance on remote sensing, the very-high-dimensional data are available for the better discrimination among the complex (sub)classes by high spectral resolution. Such hyperspectral data sets are generally made of about 100 to 200 spectral bands of relatively narrow bandwidths ($5-10$ nm). However, this large number of bands is the characteristic which leads to complexity in analysis techniques. Conventional classification methods, such as a Gaussian Maximum Likelihood algorithm, cannot be applied to hyperspectral data due to the high dimensionality of the data. The difficulty in using many classification methods based upon conventional multivariate statistical approaches is that many of these methods rely on having a non-singular class-specific covariance matrix for all classes~\cite{Benediktsson95}. When working with high-dimensional data sets, it is likely that the covariance matrices will be singular when using a limited (with respect to the number of input bands) amount of training samples. Accordingly, those approaches can result in the overfitting of training data and so lead to a poor generalization. This involves with the definition of ill-posed classification problems (which are characterized by small-sized training dataset with high-dimensional space), known as the problems caused by the Hughes phenomenon~\cite{Hughes68}.

In order to address the ill-posed classification problems, a regularized large margin classifier, Support Vector Machine (SVM)~\cite{Vapnik98} is taken into account in this paper for the better stability of classifier considered in the presence of small-size training dataset in high dimensional space. Recently, SVMs have been successfully applied in the classification of hyperspectral remote sensing data~\cite{Melgani04}. However, all of literatures are focusing on the pursuit of the solution with dual property, i.e., Lagrange theory is applied for the optimization problems~\cite{Vapnik98},\cite{Melgani04}. Nonetheless, SVMs can also be carried out directly on the primal representation~\cite{Chapelle05:Primal}. This is the focus of the paper to propose the usage of primal SVMs for the classification of hyperspectral remote sensing data with small-size training set. In particular, the 2-norm regularizer and the hinge loss are taken into account for the objective function, and conjugate gradient descent is applied on the objective for the optimization problem. It is worth noting that the implementation on such objective is an unconstrained optimization problem. In greater detail, the three ways for the implementation of nonlinear SVMs in terms of the linear case are presented in the paper. Finally, the model selection in the ill-posed classification problems is discussed.

The experimental analysis was carried out on dataset acquired by the NASA AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) instrument over the Kennedy Space Center (KSC), Florida, on March 23, 1996 and the one acquired by the NASA EO-1 satellite over the Okavango Delta, Botswana in 2001-2004\footnote{Available at http://www.csr.utexas.edu/hyperspectral/codes.html.}. The training datasets are randomly split to a sequence of small-size sets in different ratios. The results provided by the proposed implementation technique for primal SVMs were compared with those provided by the state-of-the-art approaches reported in~\cite{Chen04}. On the basis of this comparison, the proposed novel approach provided better accuracy and generalization ability of the reference methods, resulting in very promising algorithm for hyperspectral image classification. %\end{abstract}

\begin{thebibliography}{1} \providecommand{\url}[1]{#1} \csname url@rmstyle\endcsname \providecommand{\newblock}{\relax} \providecommand{\bibinfo}[2]{#2} \providecommand\BIBentrySTDinterwordspacing{\spaceskip=0 \providecommand\BIBentryALTinterwordstretchfactor{4} \providecommand\BIBentryALTinterwordspacing{\spaceskip=\fontdimen2\font plus \BIBentryALTinterwordstretchfactor\fontdimen3\font minus \fontdimen4\font\relax} \providecommand\BIBforeignlanguage[2]{{% \expandafter\ifx\csname l@#1\endcsname\relax \typeout{** WARNING: IEEEtran.bst: No hyphenation pattern has been}% \typeout{** loaded for the language ‘#1’. Using the pattern for}% \typeout{** the default language instead.}% \else \language=\csname l@#1\endcsname \fi #2}}

\bibitem{Benediktsson95} J.~S. J.A.~Benediktsson and K.~Arnason, ‘‘Classification and feature extraction of aviris data,” \emph{IEEE Trans. on Geoscience and Remote Sensing}, vol.~33, pp. 1194–1205, 1995.

\bibitem{Hughes68} G.~F. Hughes, ‘‘On the mean accuracy of statistical pattern recognition,” \emph{IEEE Trans. Inform. Theory}, vol. IT-14, pp. 55–63, January 1968.

\bibitem{Vapnik98} V.~N. Vapnik, \emph{Statistical Learning Theory}.\hskip 1em plus 0.5em minus 0.4em\relax New York: John Wiley \& Sons, Inc., 1998.

\bibitem{Melgani04} F.~Melgani and L.~Bruzzone, ‘‘Classification of hyperspectral remote sensing images with support vector machines,” \emph{IEEE Trans. Geosci. Remote Sensing}, vol.~42, no.~8, pp. 1778–1790, August 2004.

\bibitem{Chapelle05:Primal} O.~Chapelle, ‘‘Training a support vector machine in the primal,” \emph{Journal of Machine Learning Research}, vol. submitted, 2006.

\bibitem{Chen04} Y.~Chen, M.~Crawford, and J.~Ghosh, ‘‘Integrating support vector machines in a hierarchical output decomposition framework,” Anchorage, Alaska, USA, Sept. 2004, pp. 949–953.

\end{thebibliography}