شماره ۷۵۶/۴۳۶
تاریخ ۸۷/۱۲/۳۰
پیوست

دانشگاه تهران
مرکز تحقیقات بین‌المللی همزیستی با کویر

بسمه تعالی

همکار محترم جناب آقای حمید رضاکشتکار

با سلام، با کمال خوشوقتی به اطلاع می رساند مقاله جنابعالی با همکاری آقایان دکترآذرنیوند -دکترارزانی -دکتر علوی پناه ومهندس فریدون ملتی، با عنوان :

A nonparametric classification algorithm of IRS-ID imagery for land cover mapping

برای چاپ در مجله **DESERT** مورد پذیرش قرار گرفته است.

دکتر داریوش مظاهری
سردبیر مجله DESERT

صندوق پستی ۱۴۱۸۵/۳۵۴

A nonparametric classification algorithm of IRS-ID imagery for land cover mapping

H.R. Keshtkar[1*] ,H. Azarnivand[1], H. Arzani[1], S.K. Alavipanah[2], F. Mellati[3]
[1*] Faculty of Natural Resources, University of Tehran, Karaj, Iran
Hkeshtkar97@yahoo.com
[2] Faculty of Geography, University of Tehran, Tehran, Iran
[3] Faculty of Environment and Natural Resources, Ferdowsi University of Mashhad, Iran

Abstract

Land cover is one of basic data layers in geographic information system for physical planning and environmental monitoring. Until recently Maximum Likelihood classifier (MLC) has been the most common method used for supervised classification of remotely sensed data. Increasingly, nonparametric classification algorithms such as decision trees are being used. Decision tree classifiers have, however, not been used as widely by the remote sensing community for land cover classification despite their nonparametric nature and their attractive properties of simplicity, flexibility and computational efficiency. The study area selected for this experiment is a protected area locates in north-east Iran. IRS-1D LISS-3 data acquired on 5th May, 2003, is used. The area includes seven types of land cover (Wood land, Open shrubland, sparse shrubland, Meadow, Bare land, Farm land and City area). The results of this study suggest that the supervised classification (MLC) producing an overall accuracy of about 72%. The Decision tree method, which improves the classification accuracy, was applied and the classification accuracy was increased by about 7 percent to 79%.

Keywords: Land Cover Classification System; LISS-3; Maximum likelihood.

1. Introduction

Land cover is a critical variable in epidemiology and can be characterized remotely. Land cover and land use are principal factors, in both space and time, controlling the

cycling and exchange of carbon, energy and water within, and between, the different earth systems (Brown de Colstoun and Walthall, 2006). Thus, land cover classification are essential for a variety of diagnostic and predictive models that simulate the functioning of the earth systems and are useful for investigating regional and global change (Brown de Colstoun and Walthall, 2006). The limitation to achieve higher classification accuracies discussed by Defries et al (1998), Loveland et al (1999) and Hansen et al (2000), emphasize date quality of the input and the number and nature of the land cover classes of interest. Artifacts of data processing, substantial radiometric noise and geologication errors inhibit the ability to separate spectrally similar land cover classes. Many land cover types, show as much intra-class variability as inter-class spectral variability. This variability frequently exhibits multimodal distributions that cause serious difficulties for traditional classifiers such as Maximum Likelihood Classifiers (MLC) (Brown de Colstoun and Walthall, 2006). Until recently MLC has been the most common method used for supervised classification of remotely sensed data (Richards, 1993). This methodology assumes that the probability distributions for the input classes possess a multivariate normal form. Increasingly, nonparametric classification algorithms such as decision trees (DT) are being used, which make no assumptions regarding the distribution of the data being classified (Carpenter et al, 1999; Foody, 1997; Friedl et al, 1999). The nonparametric properly means that non-normal, non-homogenous and noisy data sets can be handled, as well as non-linear relations between features and classes, missing values and both numeric and categorical inputs (Quinlan, 1993). Decision tree classifiers have not been as widely used within the remote sensing community. The advantages that decision trees offer include an ability to handle

data measured on different scales, lack of any assumptions concerning the frequency distributions of the data in each of the classes, flexibility, and ability to handle non-linear relationships between features and classes (Friedl & Brodley, 1997).

The overall aim of this study was to evaluate the performance of IRS-1D LISS-3 data for assessing land cover map.

2. Background

2.1. IRS-1D Satellite

The IRS-1D satellite is the fourth in a series of commercial Indian satellites. It was launched in 1997. For the IRS satellite the imaging time is around 10:00 a.m. every 24 days. Onboard the IRS-1D satellite is several sensors. One is the LISS-3. LISS-3 scene covers an area of 141×141 km. Each pixel is 23.5×23.5 m in the raw image data but here resampled to a 24×24 m grid. It is a four-band multispectral sensor with narrow bands:

0.52-0.59 μm (green)

0.62-0.68 μm (red)

0.77-0.86 μm (near infrared)

1.55-1.70 μm (middle infrared)

2.2. Maximum likelihood classifier

As usually implemented, the maximum likelihood classifier (MLC) procedure is based on the assumption that the members of each class follow a Gaussian frequency distribution in feature space. ML is a pixel-based method, and can be defined as follows: a pixel with an associated observed feature vector $\mathbf{x}$ is assigned to class c_j of N classes if

$g_j(\mathbf{x}) > g_k(\mathbf{x})$ for all $j \neq k$, with $j,k = 1,\ldots, N$.

For the multivariate Gaussian distribution, the discriminating function $g_k(\mathbf{x})$ is given by:

$$g_k(\mathbf{x}) = \ln(p(\mathbf{x} \mid c_j)) = \ln\boldsymbol{\Sigma}_k + (\mathbf{x} - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}_k)$$

Where $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ are the sample mean vector and sample covariance matrix for class *k*.

Implementation of the ML algorithm involves the estimation of class mean vectors ($\boldsymbol{\mu}_k$) and covariance matrices ($\boldsymbol{\Sigma}_k$) from training data chosen from known examples of each particular class. The function $g_i(x)$ is used to evaluate the membership probability of an unknown pixel for class j. The pixel is assigned to the class for which it has the highest membership probability value.

2.3. Decision tree classifier

In the usual approach to classification, a common set of features is used jointly in a single decision step. An alternative approach is to use a multistage or sequential hierarchical decision scheme. The basic idea involved in any multistage approach is to break up a complex decision into a union of several simpler decisions, hoping the final solution obtained in this way would resemble the intended desired solution. Hierarchical classifiers are a special type of multistage classifier that allows rejection of class labels at intermediate stages.

Classification trees offer an effective implementation of such hierarchical classifiers. Indeed, classification trees have become increasingly important due to their conceptual simplicity and computational efficiency. A decision tree classifier has a simple form which can be compactly stored and that efficiently classifies new data. DT classifiers can perform automatic feature selection and complexity reduction, and their tree structure provides easily understandable and interpretable information regarding the predictive or generalization ability of the classification. To construct a classification tree by heuristic approach, it is assumed that a data set consisting of feature vectors and their

corresponding class labels are available. The features are identified based on problem specific knowledge. The DT is then constructed by recursively partitioning a data set into purer, more homogenous subsets on the basis of a set of tests applied to one or more attribute values at each branch or node in the tree. This procedure involves three steps: splitting nodes, determining which nodes are terminal nodes, and assigning class label to terminal nodes. The assignment of class labels to terminal nodes is straightforward: labels are assigned based on a majority vote or a weighted vote when it is assumed that certain classes are more likely than others.

A tree is composed of a root node (containing all the data), a set of internal nodes (splits), and a set of terminal nodes (leaves). Each node in a decision tree has only one parent node and two or more descendent node (figure 1). A data set is classified by moving down the tree and sequentially subdividing it according to the decision framework defined by the tree until leaf is reached.

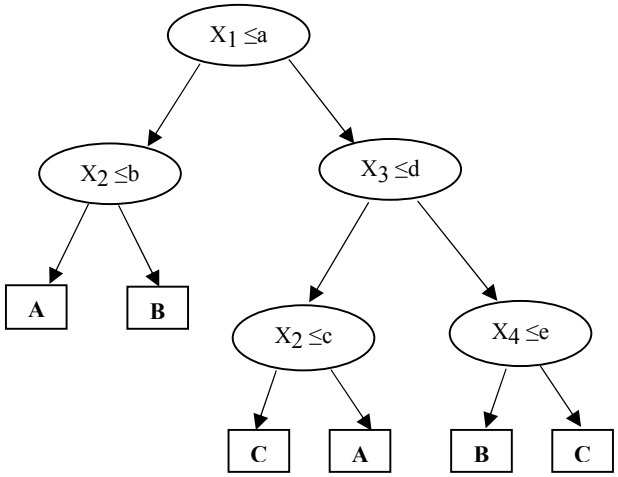

Fig. 1. A classification tree with four dimensional feature space and three classes. The *xi* are feature values; *a, b, c, d,* and e are the thresholds and *A, B,* and *C* are class labels (Pal and Mather, 2003).

3. Material and methods

3.1. Study area and the data

The study was carried out in ghorkhoud region. It is a protected area that locate in north-east of Iran (950-3000 m a.s.l., 43000 ha, see figure 1), in this area, comprising different landscape unit, including valley bottoms and ravines, plateaus with different degree of dissection and rocky hilly uplands. Mean annual precipitation is 360 mm and mean annual temperature in the region is 13 °C (Keshtkar, 2008).

We use a subset of a 2003 image from LISS-3 sensor of IRS-1D satellite of the first part of the growth season (5 May) which comprised the study area and its surrounding.

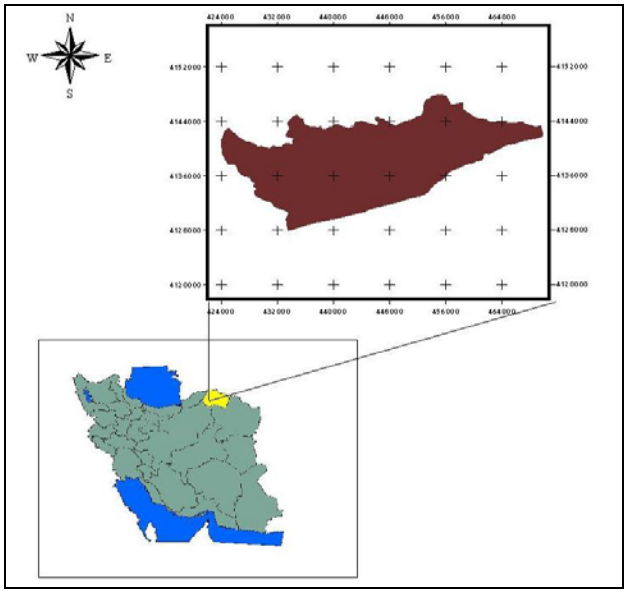

Fig 2. Location of the study area

3.2. Methods

3.2.1. Field sampling

From 5 May to 12 May we sampled train characteristic in 180 selected stands. Stands were defined as areas of 24×24 m (equivalent to the IRS-1D pixel), around the point

located with the GPS. For each stand we recorded land cover type, foliage area, topographic position, slope, aspect and altitude. The study area consists of both man-made (Village) and natural regions. Forest area include needle leaved evergreen. Nonforested areas, on the other hand, are composed of farm lands, shrublands, meadow and bare lands. Remaining regions of Image are covered with Cloud and shade. The class values of the reference points were assigned during the field survey, except for the cloud and shade classes. The classes separated using Land Cover Classification System (LCCS) that developed by FAO (FAO, 1997).

3.2.2. Remote sensing data processing

3.2.2.1. Preparing image

Geo-referencing was carried out using 41 points taken from digital topographic maps. A linear resampling with the nearest neighbor algorithm was performed, achieving a positional error of 0.56 pixels (13.2 m), with an output pixel size of 24×24 m. Resampled SWIR band (spatial resolution: 70 to 24 m). Atmospheric corrections were found unnecessary since we used single image for all further analyses and classifications (Song *et al*, 2001). The bands were enhanced using contrast enhancement and making False Color Composite images (FCC).

3.2.2.2. Image classification

The aim of the classification analysis is to categorize all of the pixels in the IRS-1D LISS-3 satellite image into land cover classes. The basic assumption is that pixels with similar spectral properties belong to a certain type of land cover. The maximum likelihood and decision tree procedures are used in this study. A brief summary of the properties of each of these classifiers is given in Background section. In addition to raw

bands, we used the Principal components analysis, digital elevation model and several indices for help to classification image. In addition to, use Battacharia distance criterion for estimate separation among classes (for maximum likelihood classification). On the final output, majority filtering (3×3 kernel) was carried out for image smoothing.

3.2.2.3. Principal Component Analysis

Principal components analysis (PCA) is often used as a method of data compression (Fig 3). It allows redundant data to be compacted into fewer bands-that is, the dimensionality of the data is reduced (Jensen, 1996; Faust, 1989). We used the first two components of a PCA to masking clouds and shadows. Following Jensen (1996) percent of total variance explained by each component was calculated as:

$$\%\,\mathrm{var}_{pc} = \frac{\lambda \times 100}{\sum_{pc=1}^{N} \lambda_{pc}}$$

Where λPC =eigenvalue, namely the variance of the Principal Component PC; N=number of Principal Components.

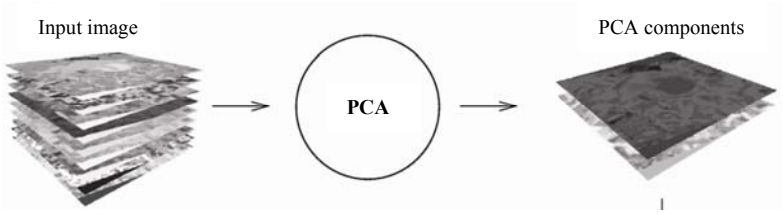

Fig. 3. A scheme for PCA

3.2.2.4. Digital elevation model

Topographic modeling or terrain analysis involves the processing of elevation data provided by digital elevation model (DEM). Specifically, we consider here the generation of slop images, which give the steepness of the terrain at each pixel, and aspect images, which give the prevailing direction relative to north of a vector normal to the landscape at

each pixel (Fig 4). DEM was originally a term reserved for elevation data provided by the USGS, but it is now used to describe any digital elevation data.

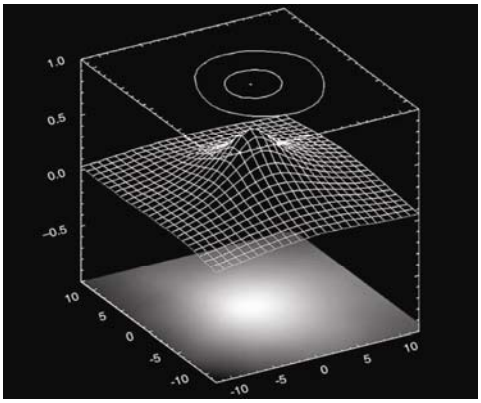

Fig. 4. DEM generated with the topographic vector map.

3.2.2.5. Calculation of indices

Several indices were derived from the values of all or part of the satellite image spectral bands (Table 1). For use of vegetation indices, relation of them with field data calculated by correlation analysis. Vegetation indices that have significant effect with field data are using in DT approach.

Table. 1.
Vegetation indices and their formulas

Code	Vegetation Index	Formula
DVI	Deference Vegetation Index	NIR-RED
GEMI	Global Environmental Monitoring Index	η (1-0.25)-(R-0.125) / 1-R η = [2(NIR2-R2) +1.5NIR+0.5R] / (NIR+R+0.5)
GNDVI	Green Normalized Difference Vegetation Index	(NIR-GREEN)/(NIR+GREEN)
IPVI	Infrared Percentage Vegetation Index	NIR/(NIR+RED)
LAI	Leaf Area Index	NDVI/(3.26-2.9+NDVI)
LWCI	Leaf Water Content Index	(NIR-MIR)/(NIR+MIR)
MIRV	MIRV	(MIR-RED)/(MIR+RED)
MSAVI	Modified Soil Adjusted Vegetation Index	(NIR-RED)*(1+L)/(NIR+RED+L) L=1-(2*a*NDVI*WDVI)
MSI	Moisture Stress Index	MIR/NIR
NDVI	Normalized Difference Vegetation Index	(NIR-RED)/(NIR+RED)
NRR	NRR	(NIR-RED)/RED
NRVI	Normalized Ratio Vegetation Index	(RVI-1)/(RVI+1)
PD322	PD322	(RED-GREEN)/(RED+GREEN)
RA	RA	NIR/(RED+MIR)
RVI 1	Ratio Vegetation Index 1	NIR/RED
RVI 2	Ratio Vegetation Index 2	Sqrt (NIR/RED)
TNDVI	Transformed Normalized Difference Vegetation Index	(NDVI+1)*100
TSAVI	Transformed Soil Adjusted Vegetation Index	a (NIR-a * RED+b)/RED+a * NIR-a*b

TVI	Transformed Vegetation Index	(NIR-RED)/(NIR+RED)+0.5
VI 1	Vegetation Index 1	RED*NIR/GREEN
VI 2	Vegetation Index 2	RED*NIR

3.2.2.6. Accuracy estimation

A random sample of 180 field points was selected and divided in two parts. A total of 150 points were used for training the classifier and remaining 30 points for testing the classifier. The cover type information of GPS points was compared with classified maps. The field sample locations were overlaid on classified maps to assess corresponding classes. Statistically valid sampling strategy was adopted to assess overall accuracy (Stehman, 1996). Finally, the contingency table was tested using Kappa coefficient (Lillesand & Kiefer, 1999). Kappa coefficient computed as follows:

$$k = \frac{p\sum_{i=1} x_{ii} - \sum_{i=1}(x_{i+} \times x_{+i})}{p2 - \sum_{i=1}(x_{i+} \times x_{+i})}$$

Where:

x_{ii}= The no. of observations in low i and column i (on the major diagonal)

x_{i+}= Total of observation in row i (shown as marginal total to right of the matrix)

x_{+i}= Total of observation in column i.

The remote sensing image was processed using the Geomatica software and the maps were drawn with the Arcview 3 geographic information system.

4- Results

4.1. Definitions of land cover informational units

The land cover classification showed that the majority of the slopes of the Ghorkhod ridge were forested. In the mountain valleys, a patchwork of meadow and shrubland was observed at intermediate altitudes while at higher altitudes meadow prevailed. Although

there are differences in forest composition between north and south slopes, we suggest that the observed differences in forest composition are largely anthropogenic in origin. These differences are most likely a legacy of socialist forest management practices and policies, because almost all forests were harvested at least once in the 20th century.

LCCS divided region to seven classes. They are as followed:

Unit 1. Mixed Woodland

The main layer consists of needleleaved evergreen woodland. The crown cover is between (70-60) and (20-10)%. The openness of the vegetation may be further specified. The height is in the range of >30-3m but may be further defined into a smaller range. Generally occurred higher 1500 m a.s.l on steep slopes in mid to low topographic positions, but was also found on flat sites in ravine bottoms or in gentle slopes. The user's label is woodland.

Unit 2. Short Herbaceous Vegetation with Dwarf Shrubs

The main layer consists of closed herbaceous vegetation. The crown cover is more than (70-60) %. The height is in the range of 3-0.03m but may be further defined into a smaller range. The second layer consists of sparse shrubs. Mainly located on mid to low steep slopes and sometimes flat positions. The user's label is meadow.

Unit 3. Broadleaved Deciduous Dwarf Shrubland with Medium High Shrub Emergents

The main layer consists of broadleaved deciduous shrubland. The crown cover is between (70-60) and (20-10)%. The openness of the vegetation may be further specified. The height is in the range of 5- 0.3m but may be further defined into a smaller range. The second layer consists of shrubs emergents. Generally located in low, flat and sometimes flooded positions. The user's label is open shrubland.

Unit 4. Broadleaved Deciduous Sparse Dwarf Shrubs and Sparse Short Herbaceous

The main layer consists of broadleaved deciduous sparse shrubs. The crown cover is between (20-10) and 4%. The sparseness of the vegetation may be further specified. The height is in the range of 5-0.3m but may be further defined into a smaller range. The second layer consists of sparse herbaceous vegetation. It was found in sites similar to unit 3. The user's label is sparse shrubland.

Unit 5. Shifting Cultivation of Small Sized Field(s) of Herbaceous Crop(s)

Small-sized field(s) is covered by rainfed herbaceous crops. One or two additional inter planted crops can be specified growing on the field simultaneously or with an overlapping or sequential period. The crop covers the land during the cropping period of a fallow system. They found in flat sites. The user's label is farm land.

Unit 6. Low Density Urban Area(s)

The land cover consists of non-linear built up areas. Mainly located in flat positions and near unit 5. The user's label is city area.

Unit 7. Very Stony Bare Soil

The land cover consists of bare soil. The surface is very stony (40 - 80%). Generally located in flat positions and around of unit 6. Erosion and grazing were high. The user's label is bare land.

4.2. Image classifications

The result of Battacharia distance showed that least separation is between city area and bare land types (Table 2). As a matter of fact, this table helped us for separate classes in decision tree method. The results of correlation analysis vegetation indices with field data are showing in table 3.

Table. 2. Degree of separation of classes

	Me	Ca	Fl	Wl	Sh	Cl	Bl	Os
Ca	1.7							
Fl	1.8	1.7						
Wl	1.4	1.5	1.9					
Sh	2	2	2	2				
Cl	2	1.8	2	2	2			
Bl	1.9	0.6	1.8	1.7	2	2		
Os	1.3	1.2	1.7	1.2	2	2	1.2	
Ss	1.6	0.8	1.7	1.3	2	1.9	1.1	1

Range of variable in this method is between 0-2 that 0 means non-separation, 0-1 means low separation, 1-2 means high separation and 2 means complete separation among classes (Ca= City area, Fl= Farm land, Wl= Wood land, Sh= Shade, Cl= Cloud, Bl= Bare land, Os= Open shrublans and Ss= Sparse shrubland).

All of bands used for separated classes in MLC. But different combination of bands (synthetic bands and/or LISS-3 bands) used for separate classes in DT approach. Finally, the classes separated from each other with combinations that as following:

Wood land (Red and NIR bands with slop map and DEM), Bare land (Red, NIR and MIR bands with TSAVI index), Farm land (Red, NIR and MIR bands with NDVI index and DEM), Meadow (Red and NIR bands with GEMI index), Open shrubland (Red and NIR bands with TSAVI index), Sparse shrubland (Red and NIR bands with TSAVI index), Cloud (all of bands with DEM) and Shade (PC1).

Table 3
Correlation coefficients between spectral indices and measured foliage area

Vegetation index	Correlation coefficient	Vegetation index	Correlation coefficient
DVI	0.206	RA	0.186
GEMI	0.458*	RVI 1	0.402*
GNDVI	0.441*	RVI 2	0.244
IPVI	-0.098	TNDVI	0.195
LAI	0.112	TSAVI	0.569**
LWCI	0.486*	TVI	0.163
MIRV	0.452*	VI 1	0.398*
MSAVI	0.447*	VI 2	-0.122
MSI	-0.165	NRVI	0.213
NDVI	0.428*	PD322	-0.103
NRR	0.054		

(*=significant for α =0.05 & ** = significant for α =0.01)

4.3. Accuracy assessment

The overall accuracy of the proposed DT and MLC were found to be 79.3% and 70.2% respectively. In order to compare different classification method namely DT and MLC techniques, kappa coefficient of agreement as an accuracy measure for remote sensing classification is used. As it is given in table 4, kappa coefficients are obtained as 0.76 for DT method and 0.62 for MLC. The MLC showed lower accuracy values than the DT classification.

Table 4
Comparison of classifications obtained from DT and MLC

	Accuracy (%)	**Kappa**
DT	79.3	0.76
MLC	70.2	0.62

5. Discussion

The main aim of this study is to assess the utility of DT classifiers for land cover classification using multispectral data, and to compare the performance of the DT classifier with that of the ML classifiers.

There are some difficult at separate classes that some of them as follow: Same plants structure in some of classes (e.g. Open shrubland, sparse shrubland and woodland) made error in separate them. Both methods (i.e. MLC and DT) can not detect city areas, because most of house made of mud and stone that show same spectral behavior with bare land. In addition to, high soil reflectance in some of classes (e.g. open shrubland and sparse shrubland) caused difficulty on classification them. Smith et al (1990) reported if plant cover is lower than 40%, soil effects prevail over plant effects. Existing understory in woodlands make error, too. Curran et al (1992), Abuelghasem et al (1999), Mickelson

et al (1998) and Nemani et al (1993) had same reports. The shade was found to be the unique class, generated by both techniques, with similar thematic output. The study region is mountly and DEM can help to separate some of classes (e.g. Woodland).

Compared to the traditional classification methods, the knowledge-base DT classification did improve the classification results. The DT can be augmented by using other remotely sensed information or other ancillary data sets such as soil type information, for example, in turn providing potential improvements in classification accuracies. The results show that DT can explore the complex relationships between bands and classes and can identify the most useful combination of bands that increases the class separability between any two classes. The DT approach is simple and flexible and does not depend on the implicit assumption regarding the relationship between the spectral information and class proportions. In addition, the structure of DT is interpretable and uncovers the hierarchical relations among bands and class proportions. The results of classifications from a synthetic image (e.g. VI or PCA) and a LISS-3 image clearly demonstrate that DT produces considerably classification accuracy as compared to the conventional MLC, especially when the data contains a large proportion of mixed pixels. Thus, DT is a potentially useful approach to produce meaningful classifications from remote sensing data. Pal and Mather (2003) obtained same result in their study.

With the rather strong relationship between classes and LISS-3 data, specifically the NIR and Red bands, NDVI, TSAVI, GEMI and PC1, LISS-3 data produced results comparable to previously documented Landsat TM results. The final image created by decision tree method is shown in figure 5.

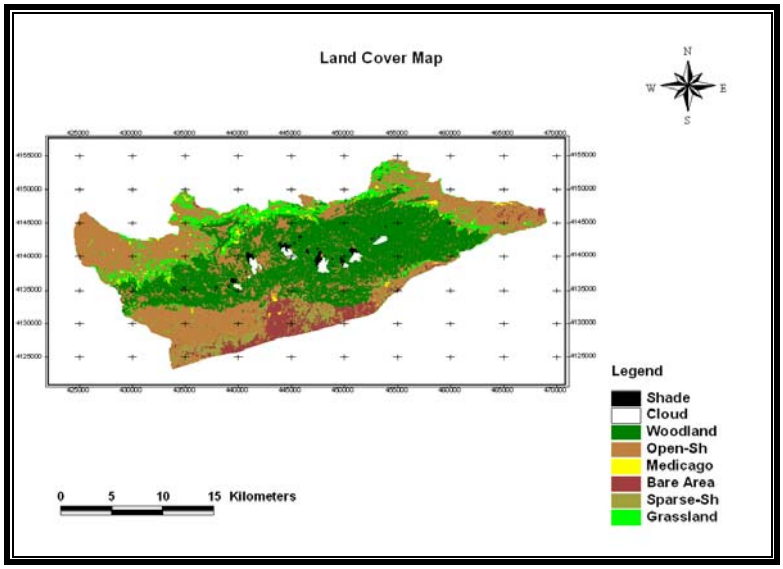

Fig 5. Final image created by using decision tree method

Refrences

Abuelghasim, A.A., Ross, W.D., Gopal, S & Woodcock, C.E. 1999. Change detectioning adaptive fuzzy neural networks environmental damage assessment after the Gulf War. Remote Sensing of Environment, 70: 208-223.

Brown de colstoun, E.C., & C.L., Walthall. 2006. Improving global scale land cover classifications with multy-directional POLDER data and a decision tree classifier. Remote Sensing of Environment, 100: 474-485.

Carpenter, G. A., Gopal, S., Macomber, S., Martens, S., Woodcock, C. E., & Franklin, J. 1999. A neural network method for efficient vegetation mapping. Remote Sensing of Environment, 70: 326– 338.

Curran, P.J., Dungan, J.L & Gholz, H.L. 1992. Seasonal LAI in slash pine estimated with Landsat TM, Remote Sensing of Environment, 39: 3-13.

DeFries, R. S., Hansen, M., Townshend, J. R. G., & Sohlberg, R. 1998. Global land cover classifications at 8 km spatial resolution: The use oftraining data derived from Landsat imagery in decision tree classifiers. International Journal of Remote Sensing, 19: 3141–3168.

FAO. 1997. Africover Land Cover Classification, FAO, Rome. Fosberg, F.R., 1961. A classification of vegetation for general purposes. Trop. Ecol. 2, 1–28.

Faust, N. L. 1989. Image Enhancement. Volume 20, Supplement 5 of Encyclopedia of Computer Science and Technology. Ed. A. Kent and J. G. Williams. New York: Marcel Dekker, Inc.

Foody, G. M. 1997. Fully fuzzy supervised classification of land cover from remotely sensed imagery with an artificial neural network. Neural Computing and Applications, 5 (4): 238– 247.

Friedl, M. A., & Brodley, C. E. 1997. Decision tree classification of land cover from remotely sensed data. Remote Sensing of Environment, 61, 399– 409.

Friedl, M. A., Brodley, C. E., & Strahler, A. H. 1999. Maximizing the land cover classification accuracies produced by decision trees at continental to global scales. IEEE Transactions on Geosciences and Remote Sensing, 27 (2): 969– 977.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., & Sohlberg, R. 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. International Journal of Remote Sensing, 21: 1331–1364.

Jensen, J. R. 1996. Introductory digital image processing, 2nd edition. Upper Saddle River, New Jersey: Prentice-Hall.

Keshtkar, H.R. 2008. Investigation of the capability of IRS satellite data for land cover mapping. MS.c thesis. University of Tehran.

Lillesand, T. M., & Kiefer, R. W. 1999. Remote sensing and image interpretation. New York: John Wiley and Sons.

Loveland, T. R., et al. 1999. An analysis of the IGBP global land-cover characterization process. Photogrammetric Engineering and Remote Sensing, 65: 1021–1032.

Mickelson, J.G., Civco, D.L & Silander, J.A. 1998. Delineating forestcanopy species in the northeastern United States using multi-temporal TM imagery, Photogrammetric Engineering and Remote Sensing, 64:891-904.

Nemani. R., Pierc. L & Running, S. 1993. Forest ecosystem processes at the watershed scale: Sensitivity to remotely-sensed Leaf Area Index estimates. International journal of Remote sensing, 14:2519-2534.

Pal, M. & Mather, P. M. 2003. An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sensing of Environment, 86: 554–565.

Quinlan, J. R. 1993. C4.5: Programs for machine learning. Morgan Kaufmann, California.

Richards, J. A. 1993. Remote sensing digital image analysis: an intro-duction. New York: Springer-Verlag.

Smith, M. O., Ustin, S. L., Adams, J. B. and Gillespie A. R. 1990. Vegetation in desert I. A Regional Measure of Abundance from multi spectral Images, Remote Sensing of Environment. 31:1-26.

Song, C., Woodcock, C. E., Seto, K. C. Lenney, M. P., & Macomber, S. A. (2001). Classification and change detection using landsat TM data. When and how to correct atmospheric effects? Remote sensing of Environment, 75: 230-244.

Stehman, S. V. (1996). Estimation of Kappa coefficient and its variance using stratified random sampling. Photogrammetric Engineering and Remote Sensing, 26, 401−407.