Investigation of Automated Feature Extraction Using Multiple Data Sources

Neal R. Harvey, Simon J. Perkins, Paul A. Pope,
James Theiler, Nancy A. David, Reid B. Porter

Space and Remote Sensing Sciences Group
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

ABSTRACT

An increasing number and variety of platforms are now capable of collecting remote sensing data over a particular scene. For many applications, the information available from any individual sensor may be incomplete, inconsistent or imprecise. However, other sources may provide complementary and/or additional data. Thus, for an application such as image feature extraction or classification, it may be that fusing the mulitple data sources can lead to more consistent and reliable results.

Unfortunately, with the increased complexity of the fused data, the search space of feature-extraction or classification algorithms also greatly increases. With a single data source, the determination of a suitable algorithm may be a significant challenge for an image analyst. With the fused data, the search for suitable algorithms can go far beyond the capabilities of a human in a realistic time frame, and becomes the realm of machine learning, where the computational power of modern computers can be harnessed to the task at hand.

We describe experiments in which we investigate the ability of a suite of automated feature extraction tools developed at Los Alamos National Laboratory to make use of multiple data sources for various feature extraction tasks. We compare and contrast this software's capabilities on 1) individual data sets from different data sources 2) fused data sets from multiple data sources and 3) fusion of results from multiple individual data sources.

Keywords: supervised classification, data fusion, support vector machines, feature extraction, machine learning, multispectral

1. INTRODUCTION

With the availability of more and more data from an increasing number of sensors, the field of data fusion is a hot-bed of activity. It is the aim of data fusion to integrate the data available from a number of different sensors in order to obtain more information than can be derived from the data available from any one of the single sensors alone. However, data fusion in and of itself is not guaranteed to be beneficial. In fact, in order for data fusion to be of benefit, there are many questions that need to be addressed. In this paper we describe experiments in which we attempt to answer some of these questions as they pertain to classification performance using a particular set of classification algorithms developed at Los Alamos National Laboratory.

2. CLASSIFICATION ALGORITHMS: AFREET

2.1. Description

AFREET[1] is a machine learning system for pixel-by-pixel image classification, that combines ideas from Support Vector Machines[2] with a hill-climbing method for constructing useful spatio-spectral image features.* It has shown itself to be capable of producing robust and powerful image pixel classifiers on a variety of feature

Work supported by the U.S. Departments of Energy and Defense.

Emails: {`harve`,`s.perkins`,`papope`,`jt`,`ndavid`,`rporter`}`@lanl.gov`.

*Afreet is inspired by an earlier system called GENIE[3] that uses genetic programming[4] and statistical classifiers for the same task. AFREET was used in this paper, mainly because it is significantly faster than GENIE.

extraction tasks. We include a brief description here, but for more detail the interested reader is referred to Perkins et al.[1]

In the pixel-by-pixel image classification problem, we wish to assign a label to each pixel in an image, where the set of labels from which to choose is fixed. Terrain categorization and crop identification are examples of problems that fit this model in the remote sensing arena. AFREET is a learning system, so it derives a classifier from a training image that has been "marked up" by an analyst to show example regions corresponding to the different label classes. We typically obtain this training data by having an image analyst paint different regions of an image using a Java image visualization tool developed at LANL.

All the experiments reported in this paper involve deriving classifiers that distinguish a target feature from "background", i.e. a two label classification problem. AFREET is motivated by the observation that assigning the correct label to a pixel often requires knowledge not only of the spectral content of that pixel (its "color"), but also of the surrounding spatial context. AFREET generates a set of numeric descriptors[†] for each pixel in the training image, using a library of spatial and spectral image processing operators, combined together in arbitrary ways. The set of descriptors defines a vector of numbers associated with each pixel in the image, and these vectors are fed into a statistical classifier called a support vector machine (SVM).[2] The SVM then derives a function that maps vectors of descriptors onto class labels.

Different sets of image descriptors lead to better or worse classification performance. In its most commonly used variant, AFREET initially generates a completely random set of descriptors, which typically perform poorly. However, by using feedback from the SVM learner, AFREET gradually evolves the descriptor set in a stochastic hill-climbing procedure, until after a number of iterations, it has found a good set of descriptors.

2.2. AFREET Variants

The hill-climbing refinement of the descriptor set usually leads to good sets of descriptors, but since each refinement requires the SVM backend to be recalculated, the process can be slow. An alternative approach is to decide in advance on a promising set of descriptors and simply skip the whole descriptor refinement stage. In this paper we consider three AFREET variants:

AFREET-search The standard hill-climbing version of AFREET described above.

AFREET-spectral In this version the vector of descriptors associated with each pixel consists simply of the intensity values in each spectral band for that pixel. The intensity values in each band are normalized to lie between the range 0.0 to 1.0 to assist the SVM backend.

AFREET-gabor A set of texture operators is applied to the image to generate a set of multi-scale texture descriptors for each pixel. For each spectral band in the image, this process generates nine additional descriptors. The texture operators are a mixture of rotationally invariant Gabor operators and Gaussian smoothing operators and are very similar to those described by Greenspan et al[5]

3. THE DATA: MUST 2000

Australia hosted a major international trial of new and emerging electro-optical imaging sensors under The Technical Cooperation Program (TTCP). The trial took place from late May through early June 2000 and involved more than 130 participants, making it one of largest trials of its kind ever held in Australia.

The MUST 2000 Multi-Sensor Trial was carried out at Cowley Beach in tropical Northern Queensland (an area located within a World Heritage Rain Forest region) and provided researchers from the UK, the USA, Canada and Australia a rare opportunity to simultaneously collect data from a variety of leading edge ground-based and airborne surveillance systems. It provided for the collection of imagery of well-characterised military and non-military targets in varying degrees of cover, against cluttered tropical and littoral backgrounds.

[†]These descriptors are usually called features in the machine learning community, but we call them descriptors here to avoid confusion with the term "feature extraction".

During the Australian trial airborne imagery was collected using a thermal hyperspectral sensor from the University of Hawaii, an Australian-made hyperspectral sensor, HYMAP, from Hyvista Corporation, Australia's JP129 airborne surveillance system and JPL's AirSAR. In addition to the airborne imagery collected, satellite imagery was also collected from various satellites, including Landsat 7 and Ikonos.

In addition to the satellite and airborne imagery collected over the region, there is also a host of ground truth information collected at the same time, together with various digitized maps of the region, including soil and vegetation maps.

A sub-set of the MUST 2000 data was used for our experiments, described next.

4. EXPERIMENTS

4.1. Data Sources

The data sources used in these experiements were: Ikonos imagery (panchromatic and multispectral, acquired 27th May 2000), Landsat ETM+ imagery (panchromatic and multispectral, acquired 15th July 2000), AirSAR imagery (HH, VV, HV, and PHASE, acquired 31st August 2000). Ground truth information was obtained from a 1:20,000 digitized black-and-white line map of vegetative land cover.[6]

4.1.1. Ikonos

Ikonos is a commerical satellite operated by Space Imaging Inc. It was launched in September 1999, and since then has been providing a reliable stream of high-resolution image data. The Ikonos satellite produces 1-meter resolution black-and-white (panchromatic) and 4-meter resolution multispectral (red, blue, green, near infrared) imagery.[7]

4.1.2. Landsat 7

Landsat 7 is the latest in a series of satellites, comprising the Landsat program: a cooperative program between NASA, NOAA and the USGS. The Enhanced Thematic Mapper Plus (ETM+) instrument onboard the Landsat 7 satellite is an eight-band multispectral scanning radiometer capable of providing high-resolution imaging information of the Earth's surface. It detects spectrally-filtered radiation at visible, near-infrared, short-wave, and thermal infrared frequency bands from the sun-lit Earth in a 183 kilometer-wide swath when orbiting at an altitude of 705 kilometers. Nominal ground sample distances or "pixel" sizes are 15 meters in the panchromatic band; 30 meters in the 6 visible, near and short-wave infrared bands; and 60 meters in the thermal infrared band.[8] For the experiments described here only the 6 30-meter resolution (visible, near and short-wave infrared) bands of the Landsat 7 data were used.

4.1.3. AirSAR

Airborne Synthetic Aperture Radar (AirSAR) is an all-weather imaging tool able to penetrate through clouds and collect data at night. The longer wavelengths can also penetrate into the forest canopy and in extremely dry areas, through thin sand cover and dry snow pack. AirSAR was designed and built by the NASA's Jet Propulsion Laboratory (JPL) which also manages the AirSAR project. AIRSAR serves as a NASA radar technology testbed for demonstrating new radar technology and acquiring data for the development of radar processing techniques and applications. As part of NASA's Earth Science Enterprise, AirSAR first flew in 1988 and continues to conduct at least one flight campaign each year, either in the United States or on an international mission.[9]

4.2. Coregistration

Some pre-processing of these data sets were required before they could be presented to AFREET for analysis. The aforementioned data sets (Ikonos, Landsat 7, AirSAR and vegetation map) were coregistered and resampled to match the cell size and spatial extents of a digitzed 1:25,000 scale map of the Cowley Beach military training area. The spatial reference system used was a latitude and longitude grid, based on the World Geodetic System (WGS) 1984 datum, with a resolution of 0.2 arc-seconds (approximately 5 meters) per raster cell. Each raster plane in the coregistered data set consisted of 2001 rows by 2001 columns of cells.

4.3. Fused data

In order to perform classification via "image fusion", we provided fused data as stacked, coregistered, multispectral image data. Thus, fused data from multiple separate sensors was created by simply stacking the separate coregistered image data sets on top of one another and creating a single "fused" image data set. Thus, in total, with the original single-sensor image data sets and the fused image data, we had 7 data sets (3 single-sensor data sets, 3 two-sensor fused data sets and 1 three-sensor fused data set):

- Ikonos image data
- Landsat 7 image data
- AirSAR image data
- AirSAR + Ikonos fused image data
- AirSAR + Landsat 7 fused image data
- Ikonos + Landsat 7 fused image data
- Ikonos + Landsat 7 + AirSAR fused image data

4.4. Tasks Set

We set AFREET two feature extraction tasks. These tasks were to find, separately, algorithms capable of distinguishing two of the vegetation type as described in the vegetation land-cover map[6] against the background of everything else. We selected the following two vegetation types:

- Vegetation type *5d*: **Medium closed forest**; *Melaleuca quinquerie* + rainforest species (*Eugenia sp. aff. angophoroides, Dillenia allata, Deplanchea eteraphylla*) + sedge (*Gahni sierberana, Scleria polycarpa*) + fern (*Blechnum indicum, Nephrolepsis biserrata*) understorey, with swamp hummocks developed on the fibrous peat surface and near perennial standing water or shallow water-tables.
- Vegetation type *9*: **Dwarf to low open forest complex**; Dwarf tree canopy of *Melaleuca viridiflora sens. lat.* ± emergent grey bloodwood (*Eucalyptus intermedia*), wattle (*Acacia crassicarpa*) and she-oak (*Casuarina littoralis*), on sandy ground-water podozol.

4.5. Training Data

Using the coregistered vegetation map overlaid on the other satellite and airborne imagery, a human analyst created training sets by marking up (labelling) regions containing the vegetation type of interest and regions containing other land-cover types (thus providing labelled examples of the two categories into which the images would be classified: feature of interest and background). For each vegetation type, two separate (completely disjoint) training data sets were constructed: one for training purposes and one for testing the algorithms found during training. For the vegetation type *5d* training set, the mark-up included 7,220 pixels for the positive (feature of interest) class and 92,508 pixels for the negative (background) class (a ratio of 1:12.8 of positive to negative training pixels). For the testing set, the mark-up included 10,119 pixels for the positive class and 142,379 pixels for the negative class (a ratio of 1:14.07 of positive to negative labelled pixels). For the vegetation type *9* training set, the mark-up included 3,327 pixels for the positive (feature of interest) class and 96,401 pixels for the negative (background) class (a ratio of 1:28.98 of positive to negative training pixels). For the testing set, the mark-up included 7,246 pixels for the positive class and 98,056 pixels for the negative class (a ratio of 1:13.53 of positive to negative labelled pixels).

4.6. AFREET Training

For training AFREET in "search" mode, for each training run, the search is initiated with a random algorithm, and the mutations etc. performed during the search are random in nature. We therefore performed 30 runs for each training set. This was done so that a determination of AFREET's overall performance, in general, could be ascertained. For training AFREET in "spectral" and "Gabor" mode, as there is no randomness in the process, we only performed a single run for each training set.

4.7. Combining multiple AFREET-search runs

Performing multiple AFREET-search runs using the same training data gives information about the overall performance of AFREET in search mode for the problems at hand. In addition, we evaluated the effect on classification performance of using AFREET-spectral to combine the results from the multiple AFREET-search runs. To this end we provided the results from each AFREET-search run together as a "multispectral" image stack to AFREET-spectral, together with the same training data as during the individual AFREET-search runs.

4.8. Combining multiple outputs from AFREET-search runs on single sensor data

We compared the performance of two methods of "fusion":

a Taking the results of combining the multiple AFREET-search runs (as described above) for the individual single data sensor data, then combining these results and providing these as the input image data to AFREET-spectral in order to obtain a final "fused" classification.

b Combining the multiple AFREET - search runs (as described above) for the fused image data.

Thus we would be comparing the classification peformance of AFREET using fused image data from multiple sensors with the classification performance of AFREET using the fused outputs of classifications on the individual sensor image data. Fig. 1 provides a graphical representation of the two approaches to fusion taken in this comparison.

5. RESULTS

5.1. Classification score

The classification tasks we set in these experiments consist of classifying every pixel in an entire image into two classes: "true" and "false", where the true class is the feature of interest (in the experiments shown here the feature of interest is a vegetation type) and the false class is the background of everything else. The goal of our optimisation technques is to find an algorithm that simultaneously maximizes detection rate (fraction of "true" pixels classified correctly) and minimizes the false alarm rate (fraction of "false" pixels classified incorrectly), with respect to the training data provided. It is convenient to have a single objective measure of performance that combines the detection rate and false alarm rates with which one can compare algorithm performance. To this end, we define the "score" for an algorithm as follows: If we denote the detection rate as R_d and the false alarm rate as R_f, then the score S of a classification algorithm is given by

$$S = 500(R_d + (1 - R_f)). \tag{1}$$

Thus, a score of 1000 indicates a perfect classification result. This score metric gives equal weighting to type I (true pixel incorrectly labelled as false) and type II (false pixel incorrectly labelled as true) errors.

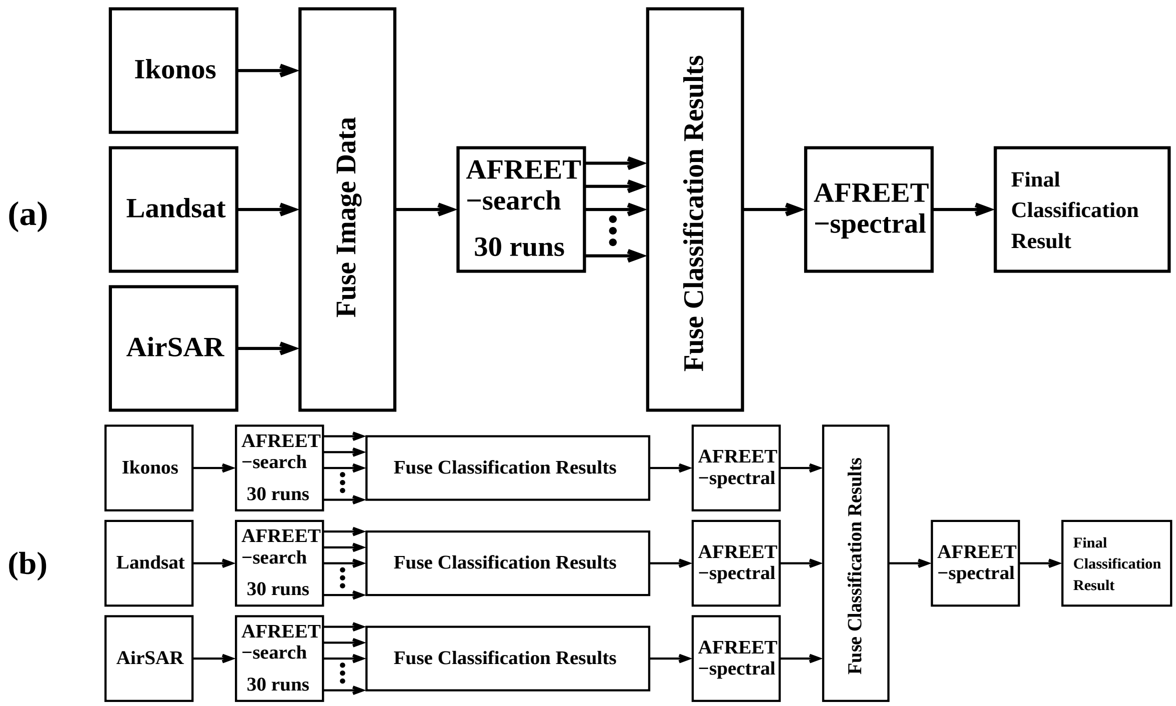

Figure 1. (a) Fusing Image Sensor Data and Classifiying vs (b) Classifying Image Sensor Data and Fusing

5.2. Comparison of performance for AFREET's different modes of operation

Table 1 shows the average performance, in terms of detection rate (DR), false alarm rate (FAR) and score, of the classification algorithms determined during training in the various AFREET modes of operation, using image data from single sensors. In this and all other tables in this paper, performance was averaged over the results for the two vegetation types. Row one shows the performance for AFREET used in spectral mode. Row two shows the preformance in Gabor mode. Row three shows the mean score for the 30 training runs in search mode. Row 4 shows the performance for the result of combining the multiple search mode results using AFREET-spectral. The performances are shown for both the in-sample (training) data and out-of-training-sample (testing) data.

With regard to performance on the training data, it can be seen that as one goes down the table (i.e. as the classifiers become increasingly sophisticated/complex), there is an increase in performance. For all but the Ikonos data there is an (albeit small) increase in detection rate, but for all sensor data types there is a consistent drop in false-alarm rates with increasing classifier complexity. For the testing data, as with the training data, there is a marked downward trend in the false-alarm rate with increased classifier complexity.

Table 2 shows the average performance of the classification algorithms using image data consisting of fused data from two different sensors. The same result is shown as for the single-sensor data sets, with at least no decrease in detection rate for the case of the Ikonos + Landsat fused data.

Table 3 shows the average performance of the classification algorithms, using image data consisting of fused data from three different sensors. It can be seen that for both the training data and the testing data there is a general downward trend in the false-alarm rate with increased classifier complexity.

5.3. Effect on classification performance of providing additional sensor image data

The results presented in the following Tables (4 - 6) present no additional data than is present in Tables 1 - 3. We merely present the information in a different format in order to compare data types more easily.

Table 4 shows the average performance, in terms of detection rate (DR), false alarm rate (FAR) and score, of the classification algorithms determined during training using AFREET in spectral mode, for the various

Table 1. Average performance for classification methods: data from single sensor

	Ikonos			Landsat			AirSAR		
	DR	FAR	Score	DR	FAR	Score	DR	FAR	Score
	Training								
AFREET-spectral	98.24	32.06	830.9	96.85	17.16	898.5	57.40	34.95	612.3
AFREET-Gabor	100.0	14.41	925.9	99.73	6.70	965.2	97.81	11.49	931.6
AFREET-search	99.97	2.77	986.1	99.99	0.92	995.3	99.66	4.40	976.4
combined multiple AFREET-search	100.0	0.12	**999.4**	100	0.02	**999.9**	100.0	1.36	**993.2**
	Testing								
AFREET-spectral	97.52	42.29	776.2	97.46	41.24	781.1	51.71	35.45	581.3
AFREET-Gabor	95.20	23.93	856.4	93.29	25.19	840.5	73.26	27.95	726.6
AFREET-search	87.23	15.93	**856.6**	93.98	14.13	899.3	89.56	13.09	**882.3**
combined multiple AFREET-search	68.89	3.94	824.8	86.69	6.43	**901.3**	65.75	4.65	805.5

Table 2. Average performance for classification methods: data from two sensors

	AirSAR + Ikonos			AirSAR + Landsat			Ikonos + Landsat		
	DR	FAR	Score	DR	FAR	Score	DR	FAR	Score
	Training								
AFREET-spectral	96.70	29.04	838.3	96.62	15.74	904.4	97.72	16.24	906.0
AFREET-Gabor	99.82	7.21	963.8	99.93	2.55	986.9	99.99	4.72	976.4
AFREET-search	99.99	3.24	983.8	99.98	1.13	994.4	99.99	1.04	994.8
combined multiple AFREET-search	100.0	0.42	**997.9**	100.0	0.01	**1000**	100.0	0.02	**999.9**
	Testing								
AFREET-spectral	93.59	37.92	778.4	94.31	40.71	768.0	97.58	32.80	823.9
AFREET-Gabor	90.86	18.27	**883.0**	87.85	20.16	838.5	91.69	16.10	**878.0**
AFREET-search	91.76	16.35	877.1	92.84	15.56	886.4	88.63	15.84	864.0
combined multiple AFREET-search	80.34	4.53	879.1	92.22	9.05	**915.9**	77.24	8.17	845.3

combinations of the three types of sensor image data. The first row shows the performance for AFREET-spectral using single sensor image data . The second row shows the performance for combinations of two sensor image data types. The third row shows the performance for the combination of the three sensor image data types. The performances are shown for both the in-sample (training) data and out-of-training-sample (testing) data. It can be seen that, for both the training data and the testing data there is a decrease in the false-alarm rate of the classifier performance as one adds additional sensor data, from one to two and from two to three sensors.

Table 5 shows the average performance of the classification algorithms determined during training using AFREET in Gabor mode, for the various combinations of the three types of sensor image data. As with the

Table 3. Average performance for classification methods: data from three sensors

	Ikonos + Landsat + AirSAR					
	DR	FAR	Score	DR	FAR	Score
	Training			Testing		
AFREET-spectral	97.22	14.73	912.45	94.38	31.54	814.3
AFREET-Gabor	100.0	1.96	990.2	88.05	14.32	868.7
AFREET-search	99.99	1.37	993.2	91.99	15.13	884.3
combined multiple AFREET-search	100.0	0.04	**999.8**	86.25	5.87	**901.9**

Table 4. Results using AFREET-spectral: effect of additional sensor data

DR	FAR	Score	DR	FAR	Score	DR	FAR	Score
Training								
Ikonos			Landsat			AirSAR		
98.24	32.06	830.9	96.85	17.16	898.5	57.40	34.95	612.3
AirSAR + Ikonos			AirSAR + Landsat			Ikonos + Landsat		
96.70	29.04	838.3	96.62	15.74	904.4	97.72	16.24	906.0
Ikonos + Landsat + AirSAR								
97.22	14.73	**912.5**						
Testing								
Ikonos			Landsat			AirSAR		
97.52	42.29	776.2	97.46	41.24	781.1	51.71	35.45	581.3
AirSAR + Ikonos			AirSAR + Landsat			Ikonos + Landsat		
93.59	37.92	778.4	94.31	40.71	768.0	97.58	32.80	**823.9**
Ikonos + Landsat + AirSAR								
94.38	31.54	814.3						

AFREET-spectral case, it can be seen that, for both the training data and the testing data there is a decrease in the false-alarm rate of the classifier performance as one adds additional sensor data.

Table 6 shows the average performance of the classification algorithms determined during training using AFREET in search mode, for the various combinations of the three types of sensor image data. There does not appear to be any noticeable trend or relationship between classifier performance (detection rate, false-alarm rate or score) and additional sensor data for AFREET operated in search mode.

5.4. Comparison of performance using fused data from multiple sensors to performance using fused classification results from single sensor data

Table 7 compares the average classification peformance of AFREET using fused image data from two sensors with the classification performance of AFREET using the fused outputs of classifications on the individual sensor image data. Row one shows the mean performance of AFREET in search mode over 30 training runs using the fused image data from two sensors. Row two shows the performance of AFREET in spectral mode using the results from each individual AFREET-search training run as the input image data (as shown graphically in Fig. 1(a)). Row three shows the performance of AFREET in spectral mode, where the input data is the combination of the (AFREET-spectral) combined results from the AFREET-search runs on the individual, single sensor image data (as shown graphically in Fig. 1(b)). The performances are shown for both the in-sample (training) data and out-of-training-sample (testing) data.

It can be seen that, for the training data, there does not appear to be any benefit in performing classification on the individual sensor data and combining the results compared to combining the sensor data and then performing the classification on the fused data. However, on the testing data, there is a decrease in false-alarm rate when using the latter approach, compared to the former.

Table 8 compares the average classification peformance using fused image data from three sensors. It can be seen that, just as with the two-sensor case, for the training data, there does not appear to be any benefit in performing classification on the individual sensor data and combining the results compared to combining the

Table 5. Results using AFREET-Gabor: effect of additional sensor data

DR	FAR	Score	DR	FAR	Score	DR	FAR	Score
Training								
Ikonos			Landsat			AirSAR		
100.0	14.41	925.9	99.73	6.70	965.2	97.81	11.49	931.6
AirSAR + Ikonos			AirSAR + Landsat			Ikonos + Landsat		
99.82	7.21	963.8	99.93	2.55	986.9	99.99	4.72	976.4
Ikonos + Landsat + AirSAR								
100.0	1.96	**990.2**						
Testing								
Ikonos			Landsat			AirSAR		
95.20	23.93	856.4	93.29	25.19	840.5	73.26	27.95	726.6
AirSAR + Ikonos			AirSAR + Landsat			Ikonos + Landsat		
90.86	18.27	863.0	87.85	20.16	838.5	91.69	16.10	**878.0**
Ikonos + Landsat + AirSAR								
88.05	14.32	868.7						

sensor data and then performing the classification on the fused data and on the testing data, there is a decrease in false-alarm rate when using the latter approach, compared to the former.

6. DISCUSSION AND CONCLUSIONS

From the results presented here there are several conclusions that can be drawn.

For any particular sensor image data set (or combination of sensor image data sets), as one increases complexity of the classification algorithm to be used (i.e. from AFREET-spectral to AFREET-Gabor to AFREET-search to combined mulitple AFREET-search), there is a decrease in false alarm rates for both in-sample and out-of-sample data.

The more complex classifiers always do better than the other less-complex classifiers on the training data. However, for the combination of single sensor data (the least complex data), and the most complex classifier (the combined multiple AFREET-search) you pay the price of this extra complexity in the classifier by performing worse on the testing data than a less complex classifier (i.e. single AFREET-search run). For the fused (more complex) data, the most complex classifier does much better on the testing data. The combined multiple AFREET-search mode results on testing data are better than the single AFREET-search mode for the 3-sensor fused data, and for 2 out of 3 of the 2-sensor fused data. Therefore, for the data sets and classifiers used in these experiments, we can hypothesize that as one increases the complexity of the data being used, more complex classifiers become more useful.

For the comparison between fusing multiple single-sensor data sets and then performing classification and performing classification on multiple single-sensor data sets separately and then fusing the results, there appears to be no distinct benefit, with these data sets and classifiers, in performing one way or the other, except that, on out-of-sample data, the latter method has a reduced false-alarm rate.

Table 6. Results using AFREET-search: effect of additional sensor data

DR	FAR	Score	DR	FAR	Score	DR	FAR	Score
Training								
Ikonos			Landsat			AirSAR		
99.97	2.77	986.1	99.99	0.92	**995.3**	99.66	4.40	976.4
AirSAR + Ikonos			AirSAR + Landsat			Ikonos + Landsat		
99.99	3.24	983.8	99.98	1.13	994.4	99.99	1.04	994.8
Ikonos + Landsat + AirSAR								
99.99	1.37	993.2						
Testing								
Ikonos			Landsat			AirSAR		
87.23	15.93	856.6	93.98	14.13	**899.3**	89.56	13.09	882.3
AirSAR + Ikonos			AirSAR + Landsat			Ikonos + Landsat		
91.76	16.35	877.1	92.84	15.56	886.4	88.63	15.84	864.0
Ikonos + Landsat + AirSAR								
91.99	15.13	884.3						

Table 7. Comparison of average performance for combining classification results: data from two sensors

	AirSAR + Ikonos			AirSAR + Landsat			Ikonos + Landsat		
	DR	FAR	Score	DR	FAR	Score	DR	FAR	Score
	Training								
AFREET-search	99.99	3.24	983.8	99.98	1.13	994.4	99.99	1.04	994.8
combined multiple AFREET-search	100.0	0.42	997.9	100.0	0.01	**1000.0**	100.0	0.02	**999.9**
AFREET-spectral on outputs of combined mulitple AFREET-search	100.0	0.20	**999.0**	100.0	0.07	999.7	100.0	0.07	999.7
	Testing								
AFREET-search	91.76	16.35	877.1	92.84	15.56	886.4	88.63	15.84	864.0
combined multiple AFREET-search	80.34	4.53	**879.1**	92.22	9.05	915.9	77.24	8.17	845.3
AFREET-spectral on outputs of combined mulitple AFREET-search	71.85	3.56	841.5	88.86	1.55	**936.6**	78.15	2.87	**876.4**

REFERENCES

1. S. Perkins, N.R. Harvey, S.P. Brumby and K. Lacker, “Support Vector Machines for Broad Area Feature Extraction in Remotely Sensed Images”, in *Proc. SPIE 4381*, pp. 286–295, 2001.
2. N. Cristianini, J. Shawe-Taylor, “Introduction to Support Vector Machines: And Other Kernel-Based Learning Methods”, *Cambridge University Press*, Cambridge, U.K., 2000.
3. N.R. Harvey, J. Theiler, S.P. Brumby, S. Perkins, J.J. Szymanski, J.J. Bloch, R.B. Porter, M. Galassi, A.C. Young, ”Comparison of GENIE and Conventional Supervised Classifiers for Multispectral Image Feature Extraction”, *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 40 No. 2, pp. 393–404, Feb. 2002, .
4. J.R. Koza, “Genetic Programming : On the Programming of Computers by Means of Natural Selection (Complex Adaptive Systems)”, *MIT Press*, Cambridge, MA, USA, 1992.

Table 8. Comparison of average performance for combining classification results: data from three sensors

	Ikonos + Landsat + AirSAR		
	DR	FAR	Score
	Training		
AFREET-search	99.99	1.37	993.2
combined multiple AFREET-search	100.0	0.04	**999.8**
AFREET-spectral on outputs of combined mulitple AFREET-search	100.0	0.70	999.7
	Testing		
AFREET-search	91.99	15.13	884.3
combined multiple AFREET-search	86.25	5.87	**901.9**
AFREET-spectral on outputs of combined mulitple AFREET-search	77.86	1.73	880.5

5. H. Greenspan, S. Belongie, R. Goodman, P. Perona, "Rotation Invariant Texture Recognition using a Steerable Pyramid", in *Proc. Int. Conf. on Pattern Recognition*, Vol. B, pp. 162–167, Nov. 1994.
6. M.S. Hopkins, A.W. Graham, "Vegetation Structure Types and Complexes, Cowley Beach Area, Joint Tropical, Trials Research Establishment, North Queensland, Australia", *Commonwealth Scientific and Industrial Research Organization, Institute of Earth Resources, Division of Land Use Research*.
7. `http://www.spaceimaging.com/products/ikonos/`
8. `http://landsat.gsfc.nasa.gov/`
9. `http://airsar.jpl.nasa.gov/`