Active Learning for Detecting a Spectrally Variable Subject in Color Infrared Imagery

Patricia G. Foschi
Romberg Tiburon Center for Environmental Studies
San Francisco State University

Huan Liu
Department of Computer Science and Engineering
Arizona State University, Tempe

Abstract—To classify *Egeria densa*, Brazilian waterweed, in scan-digitized color infrared aerial photographs, we are developing an interactive computer system based on data-mining techniques with Active Learning capabilities. Key components of the system are: feature extraction, automatic classification, Active Learning, and experimental evaluation.

Key words—active learning, data mining, feature extraction, CIR imagery.

I. BACKGROUND AND PROBLEM

The need for airborne data collection, including aerial photography, still exists for many applications. Improvements in orthorectification techniques have allowed aerial photography, recent and historic, to be digitized, geometrically corrected, and integrated into databases more rapidly and cost-effectively. However, human visual/manual image interpretation and analysis procedures are often time-consuming and costly, not repeatable, and dependent on the varying abilities of the interpreters. Individual expertise is hard to transfer from one interpreter to another, which contributes to high training costs. There is a significant gap between fast routine data collection and the slow interpretation and analysis of complex and detailed images in multidimensional (spatial and temporal) space. Monitoring *Egeria densa*, an invasive submergent weed, by remote sensing represents such a case.

Egeria densa, commonly called Brazilian waterweed, has grown uncontrolled in the Sacramento-San Joaquin Delta of Northern California for over 35 years displacing native flora and now covers about 2400 hectares of waterways. The presence of this exotic weed is

disrupting navigation and recreational uses of waterways, clogging irrigation intake trenches, and causing reservoir-pumping problems. Since 1997, the California Department of Boating and Waterways has been developing and implementing a control program for *Egeria* in the Delta. Airborne color infrared (CIR) imagery has been used to monitor the areal extent of *Egeria* every year since then (Foschi, 2000). The database of CIR imagery is online at: http://romberg.sfsu.edu/~egeria.

Classifying *Egeria* in CIR airphotos by automated methods presents a challenging problem due to a number of unfavorable conditions. These include variable imaging conditions (e.g., film exposure, vignetting, scanning anomalies), problems associated with water-related subjects (e.g., turbidity, sun glint, surface reflectance due to wind), and other environmental changes (e.g., exposure of *Egeria* at extremely low tide, shadows falling upon the water, algal cover over *Egeria*). Digital analyses also indicate that subtle changes (e.g., in *Egeria* canopy density, film vignetting, or water turbidity) produce overlapping spectral response patterns. In addition, the spectral response patterns for *Egeria* do not separate well from those of other land cover classes in CIR imagery. For example, dense well-submerged *Egeria* appears black and is confused with shadows on land; *Egeria* exposed during very low tide appears reddish and is confused with terrestrial vegetation. Clearly, traditional computer-assisted multispectral classification methods are problematic under these conditions, and visual/manual image interpretation and analysis procedures are daunting. Post-classification processing is frequently needed to clean up noisy patches in classification maps.

We are exploiting the latest developments in data mining, investigating novel combinations of effective methods in feature extraction, classification, and machine learning (Natsev et al., 1999; Hakkani-Tur et al., 2002), and proposing a computer system that implements the novel combinations. Such a system should accomplish routine tasks,

highlight areas for verification or further analysis, and minimize expert intervention. The interactive system incorporates Active Learning and has the following key components:

- Feature extraction - to derive relevant and effective features for automatic classification (Ma et al., 1997; Liu and Motoda, 1998a);

- Automatic classification – to achieve speed, accuracy, generalization, and automation;

- Active Learning - to use machine learning techniques to minimize expert intervention without performance deterioration; and

- Experimental evaluation – to compare the performance with and without the newly developed systems.

We anticipate creating an interactive learning system that can learn from human analysts by relating results to extracted objects and that can learn analytic rules for classification. In the following, we introduce the system's key components, present preliminary results, and recommend future work.

II. INTEGRATED INTERACTIVE SYSTEM

Due to the spectral variability of *Egeria* in CIR imagery, human experts have been heavily engaged in the problem of detecting this landscape class. Characteristic and unusual colors, shapes, and textures have had to be verified by ground surveys. Typical *Egeria* sites and associations have been cataloged over time. Many problems have only been identified by multitemporal information. We aim to reduce the amount of involvement of human experts. Therefore, the new system uses Active Learning and is interactive in defining features and resolving uncertainties. The images are divided into two groups: a training set and a testing set. Ideally, the training set is very small because a smaller training set will require less involvement of human experts. Also, we are trying to minimize the involvement of human experts during testing.

A flow chart of the interactive system is presented in Figure 1. The following sections describe the components of this system in more detail.
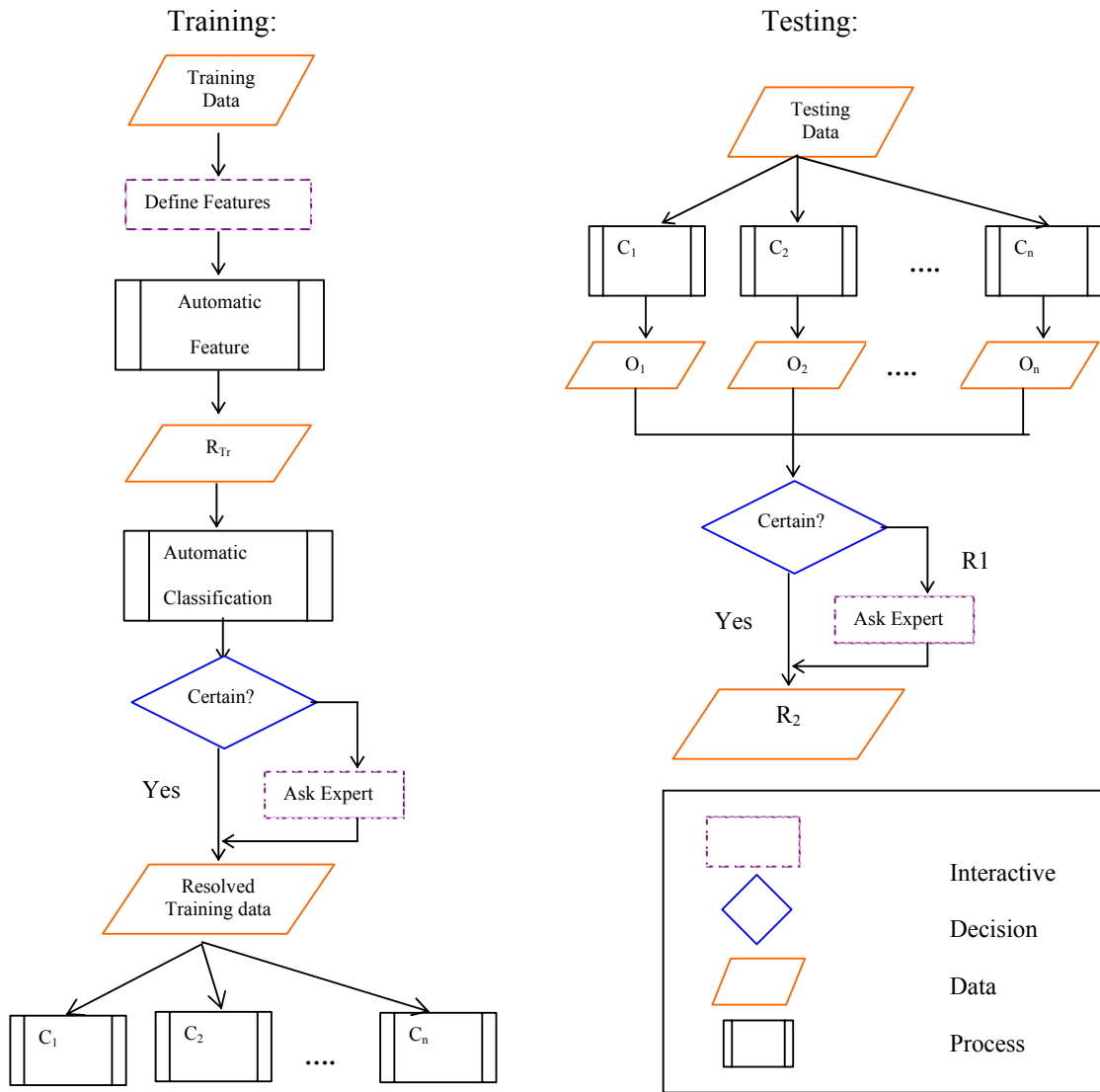
Figure 1. Flow chart of interactive system for detecting *Egeria* in digital imagery.

A. Feature Extraction

The primary step of the system is feature selection and extraction. The domain under study is high-resolution CIR imagery. To detect *Egeria* effectively, we must select or extract regions of interest from the imagery. Features are needed that can describe regions and that can separate the interesting from the rest (Jain and Vailaya, 1996; Liu and Motoda, 1998b).

There are numerous features that can be used to select regions of interest. The color feature alone has proven ineffective in detecting *Egeria*. Spatial information or features ― like context, shape, and texture ― are more likely to aid automated classification. In our

trials, we have been very conservative in introducing new features. The version of the system reported here uses three texture features, one color feature, and one edge feature.

The texture templates cover various types of textures: texture containing pure submerged *Egeria* patches, texture containing *Egeria* in shallower water conditions, and texture containing *Egeria* adjacent to landmasses. The training image has these varying conditions of *Egeria*. To calculate a particular texture, an 8-pixel-by-8-pixel template is created. The original image is divided into blocks of 8 x 8 pixels, each block overlapping the previous one halfway both horizontally and vertically. The Discrete Cosine Transform (DCT) (Gonzalez and Woods, 1992) is performed on each of the blocks and the template using all spectral bands of the imagery (G, R, and near infrared). Then the DC value of each block (first term of the Discrete Cosine Transformed matrix) is compared with the DC value of the texture template, and the results are limited with a threshold value. This DC value is sufficient to represent the textural information of a block. A region matrix is created from the values obtained. The same process is repeated for the other two textures. Common regions identified by these textures correspond to the presence of *Egeria* with greater certainty. Uncertain regions will require expert interpretation.

For the color feature, we are focusing on shades of pink and black, common *Egeria* colors in CIR imagery corresponding respectively to emergent and submergent *Egeria*. We are using the YCbCr method (Gonzalez and Woods, 1992) which also incorporates information from all of the spectral bands. The ranges of the YCbCr components remained relatively stable for almost all images tested. The average color of each 8 x 8 block is compared to threshold values. This process effectively captures both colors of *Egeria*.

For the edge feature, each 8 x 8 block is subjected to Canny's edge detection process (Mathworks, 2001). To avoid identifying all edges in the block, many of which are irrelevant, we use only edges that have *Egeria* color. Since these features provide many dimensions to describe *Egeria*, combining them is expected to improve classification results.

B. Automatic Classification Using an Interactive System

An automatic classification system, called I-Class, was built using these five features. This prototype version of the interactive system runs in a MATLAB environment and can be used in both training and testing (or classification) phases. The system allows the user to open images that have already been classified, either for viewing or for resolving conflicts. Resolving conflicts may include adding missing regions, removing noisy regions, and labeling uncertain regions. During the training phase, the human expert resolves those areas labeled 2 or uncertain by I-Class. A resolved image is one with areas labeled only as 0 for not *Egeria* or 1 for most probably *Egeria*. After an image is resolved, it is analyzed by the Active Learning subsystem (discussed in the next section). During the testing phase, the Active Learning subsystem outputs labels for regions in the test images. Uncertain regions are again labeled by the expert using the interactive system. After either phase, the output of the system is expected to be a completely resolved image.

While it is reasonable to assume that I-Class would work well, this is only partially true. It works well in the training phase. The I-Class operations are predefined and may not generalize well from the training data and the labeled regions resolved by the expert. Therefore, I-Class is not expected to be robust when it is applied to test images. In order to overcome this problem, some data-mining algorithms were employed for generalization. Since these algorithms are also used in Active Learning, we will discuss them next.

C. Active Learning

Automated classification is a process for predicting outputs without, or with as little intervention as possible from, human experts. The job of a learning algorithm is to relieve the experts from making all decisions and to automate the process of predicting class labels of new instances. A classification learning algorithm uses the training data to "learn" the mapping between the inputs and the output classes. It then predicts the classes for the test data. Classification learning is sometimes referred to as supervised learning because the

actual classes for the training examples have to be provided as the algorithm learns the mapping. Active Learning (Cohn et al.,1994; Roy and McCallum, 2001) is a supervised learning algorithm that combines the results from multiple classification algorithms.

In Active Learning, initially multiple algorithms are trained using the same training data to learn the patterns in the data. Then each of the algorithms is tested with the test data. The outputs are collected from all the classifiers and compared. The set of all the samples for which all the algorithms agree is selected and labeled appropriately. The rest of the samples with uncertain output classes (i.e., different learning algorithms identified different classes) are then referred to the experts for resolution. This resultant set is denoted by R1 in Figure 1. The interactive system introduced earlier prompts the expert to manually label the classes by regional blocks. The resultant set of samples, denoted by R2 in Figure 1, can then be appended to the training set to retrain the algorithms. This procedure is repeated each time there is a new set of testing data. As the procedure is repeated with more training and resolved data, the interactive system will be needed less and less.

The Active Learning algorithm in Table I can start with a small set of labeled data. At each step, the samples that have been learned by the algorithms do not need to be labeled again. Only those samples not learned previously must be labeled. Thus, Active Learning saves relabeling time as well as time required by human experts for labeling data. As the process is repeated, the amount of work to be done by the experts continues to decrease.

TABLE I
ACTIVE LEARNING ALGORITHM

- Given: L labeled instances, U unlabeled instances, E expert
- Loop until all instances in U are labeled
    - Use L to train h1
    - Use L to train h2
    - Pick one instance i from U
    - if h1(i) != h2(i) then
      Ask E to label i and
      Add i to L and remove i from U
      else label i according to h1

Many machine learning algorithms could be used in Active Learning. Since the comprehensibility of learning results is crucial, we are experimenting with two learning algorithms proven to be efficient and powerful in learning, yet distinct in underlying principles: Decision Tree Induction (DTI) and the Naïve Bayes Classifier (NBC) (Witten and Frank, 2000). DTI aims to build a compact decision tree with pure leaf nodes as soon as possible. One way to achieve this is to find the most promising feature that can split the data and to do so recursively until some criterion is met. The degree of a feature's promise can be measured by the information gained or the weighted difference of entropy before and after splitting the data. The entropy of 2-class data with p number of positive instances and n number of negative instances (s is the sum of p and n) is calculated as:

$$E\ (\ p\ ,\ n\ )\ =\ -\ \frac{p}{s}\log(\ \frac{p}{s})\ -\ \frac{n}{s}\log(\ \frac{n}{s}).$$

NBC adopts a different approach that tries to maximize $P(C_i|x)$ by using Bayes Rule to maximize $P(x|C_i)P(C_i)$.

As more images are classified by human experts and as relevant objects are repeatedly identified by various features in different images, the active learning algorithms are repeatedly strengthened and improved.

III. DATA AND EXPERIMENTS

The preliminary trials were tested using CIR aerial photographs of the Sacramento-San Joaquin Delta flown in October 2000 at 1:24,000 scale. The airphotos were scan-digitized and color separated to create 3-band digital imagery at a nominal 2-meter spatial resolution. Subsets were selected from 21 of the airphotos to represent cases in which *Egeria* is readily interpretable to an image analyst. *Egeria*, however, does not appear the same in all subsets. The system was trained using one of the subsets and tested using the other 20.
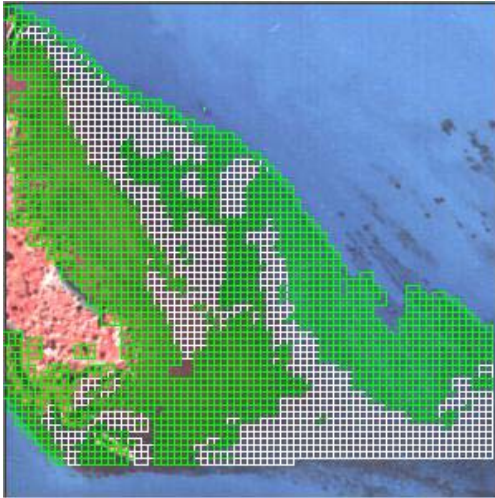
Image 1000_2m_bb1
Certain (white) and Uncertain (green)
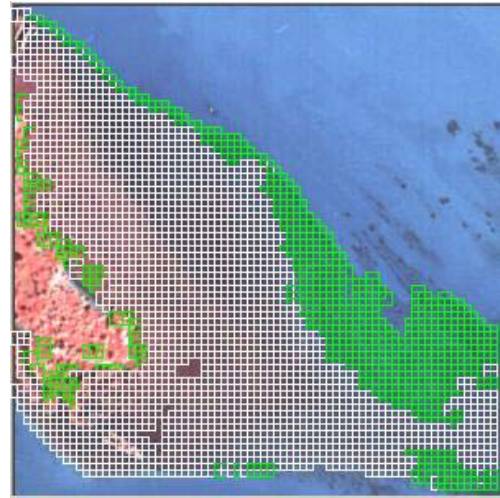blocks before Active Learning

Image 1000_2m_bb1
Certain (white) and Uncertain (green)
blocks after Active Learning

Image 1000_2m_di1
Certain (white) and Uncertain (green)
blocks before Active Learning

Image 1000_2m_di1
Certain (white) and Uncertain (green)
blocks after Active Learning

Figure 2. Output of the interactive system before and after Active Learning.

The results for two of the test images are shown in Figure 2. The images on the left show the certain blocks (shown in white) and uncertain blocks (shown in green) before Active Learning. The images on the right display the certain and uncertain blocks after Active Learning. It is clear that the number of certain blocks has increased and the number of uncertain blocks has decreased after Active Learning. These experiments suggest that Active

Learning, by reducing the number of uncertain blocks, can decrease expert interaction in the classification process.

TABLE II
SUMMARY OF RESULTS

Image	Before Active Learning					After Active Learning				
	Precision	Recall	Accuracy	Certain	Uncertain	Precision	Recall	Accuracy	Certain	Uncertain
lvi2*	**0.7740**	**0.9448**	**0.8844**	**1900**	**446**	**0.7642**	**0.7770**	**0.8384**	**2229**	**117**
bb1	0.9700	0.6865	0.8158	950	1739	0.9526	0.8611	0.9000	2047	642
di1	0.8151	0.6778	0.8243	1333	543	0.8158	0.7720	0.8515	1717	159
ft1	0.4308	0.9570	0.4510	3132	1573	0.4540	0.9566	0.4988	4635	70
lps1	0.2342	0.9637	0.8078	1776	407	0.1450	0.9731	0.6522	2003	180
ls1	0.4841	0.5648	0.8445	1212	540	0.3477	0.7622	0.7499	1354	398
qi1	0.6833	0.9044	0.9013	1363	304	0.5936	0.9563	0.8660	1444	223
vc1	0.9345	0.6217	0.8608	865	514	0.8736	0.6884	0.8643	1271	108
wi1	0.6009	0.8651	0.8385	1263	1833	0.3675	0.9305	0.6195	1797	1299
ft1	0.8165	0.7385	0.8430	1432	399	0.8124	0.7699	0.8502	1759	72
ft2	0.9840	0.6711	0.8352	1782	223	0.9780	0.7455	0.8684	1869	136
ft3	0.9267	0.6058	0.9043	536	270	0.8024	0.5692	0.8764	681	125
lvi1	0.9304	0.3043	0.6971	762	155	0.9393	0.3731	0.7256	831	86
ri1	0.6478	0.3599	0.9148	411	173	0.5545	0.5833	0.9098	438	146
vc1	0.4866	0.6718	0.9133	601	286	0.3404	0.6714	0.8638	692	195
wi1	0.9632	0.5231	0.7554	911	567	0.9607	0.5346	0.7602	1399	79
bb2	0.9261	0.4259	0.6584	1237	265	0.9266	0.4611	0.6767	1435	67
bb1	0.9470	0.5667	0.7654	1518	636	0.8552	0.6796	0.7803	1745	409
ls1	0.5645	0.7105	0.8707	986	455	0.4923	0.8461	0.8415	1179	262
sl1	0.9733	0.6984	0.7782	2420	375	0.9597	0.7203	0.7857	2713	82
sl2	0.9179	0.5672	0.8796	836	315	0.8526	0.7419	0.9038	891	260

* training image

To evaluate the test results, we used three criteria: precision, recall, and reduction in expert interaction. These criteria are defined as:

- Precision: the fraction of the relevant information over the retrieved information or the fraction of the classification that is correctly identified.

- Recall: the fraction of the relevant information that is retrieved over all relevant information.

- Reduction: the reduction in the number of uncertain blocks.

Table II summarizes the evaluation results for all testing images used in the experiments. In Table II, the reduction in expert interaction can be seen in the reduction in the number of uncertain blocks before and after applying Active Learning. The recall values for all test images should be large enough so that the instances retrieved are comparable to the actual

(relevant) instances. However, the disadvantage of having very high recall values is that precision can decrease rapidly. While these results are satisfactory for preliminary trials, more study is needed for overall evaluation of the system. As trials proceed, we intend to refine the choice of templates and to add new features that describe other spatial information indicative of the presence of *Egeria*.

IV. SUMMARY AND CONCLUSIONS

To address the challenges posed by detecting a spectrally variable subject in CIR imagery, we are exploiting the latest developments in data mining, investigating novel combinations of effective methods — including feature extraction, automatic classification, and machine learning — and proposing a computer system that implements the novel combinations. We have created a prototype interactive system that uses Active Learning. The system carries out routine tasks, highlights areas for verification and further analysis, and reduces expert intervention. We anticipate further developments and refinements.

Although we use images of *Egeria densa* as our testbed, many underlying principles and design methodologies can be extended to other domains of wetland/waterway monitoring. Providing an automated solution to such monitoring would have far-reaching applications since rapidly changing conditions occur in all tidal marshland and in wetlands in general. Automatic detection and classification algorithms using Active Learning may be deployed in other cases where data are massive and complex (Liu and Motoda, 2001).

ACKNOWLEDGMENTS

REFERENCES

Cohn, D, L. Atlas, R. Ladner. 1994. Improving generalization with active learning. *Machine Learning* 15:201-221.

Foschi, P.G. 2000. Egeria densa acreage and percent coverage in the Sacramento-San Joaquin Delta. In: California Department of Boating and Waterways. *Egeria densa Control Program*, Vol. 2, Report 4, March 2000.

Gonzalez, R., R.Woods. 1992. *Digital Image Processing*. Addison-Wesley Publication Company.

Hakkani-Tur, D., G. Riccardi, A. Gorin. 2002. Active Learning for Automatic Speech Recognition. *Proceedings of International Conference on Acoustics Speech and Signal Processing.*

Jain, A., A. Vailaya. 1996. Image Retrieval Using Color and Shape. *Pattern Recognition*, 29(8): 1233-1244.

Liu, H., H. Motoda. 1998a. *Feature Selection for Knowledge Discovery and Data Mining*. Kluwer Academic Publishers.

Liu, H., H. Motoda, Eds. 1998b. *Feature Extraction, Selection, and Construction for Data Mining*. Kluwer Academic Publishers.

Liu, H., H. Motoda, Eds. 2001. *Instance Selection and Construction for Data Mining*. Kluwer Academic Publishers.

Ma, W., Y. Deng, B.S. Manjunath. 1997. Tools for Texture/Color Based Search of Images. *Proceedings of SPIE International Conference – Human vision and Electronic Imaging*, 496-507.

Mathworks, Inc. 2001. *MATLAB Image Processing Toolbox User's Guide*, Version 3. Mathworks, Inc., http://www.mathworks.com.

Natseve, A., R. Rastogi, K. Shim. 1999. WALRUS: A Similarity Retrieval Algorithm for Image Databases. *SIGMOD*, 394-406.

Roy, N., A. McCallum. 2001. Toward optimal active learning through sampling estimation of error reduction. *Proceedings of the Eighteenth International Conference on Machine Learning*.

Witten, I.H., E. Frank. 2000. *Data Mining*. Morgan Kaufmann Publishers.