THE ASTRONOMICAL JOURNAL VOLUME 109, NUMBER 6 JUNE 1995

AUTOMATED STAR/GALAXY CLASSIFICATION FOR DIGITIZED POSS-II

NICHOLAS WEIR
Palomar Observatory, California Institute of Technology 105-24, Pasadena, California 91125
Electronic mail: weir@fritz.caltech.edu

USAMA M. FAYYAD
Jet Propulsion Laboratory, California Institute of Technology 525-3660, Pasadena, California 91109
Electronic mail: fayyad@aig.jpl.nasa.gov

S. DJORGOVSKI
Palomar Observatory, California Institute of Technology 105-24, Pasadena, California 91125
Electronic mail: george@deimos.caltech.edu

Received 1994 September 2; revised 1995 March 1

ABSTRACT

We describe the automated object classification method implemented in the Sky Image Cataloging and Analysis Tool (SKICAT) and applied to the Digitized Second Palomar Observatory Sky Survey (DPOSS). This classification technique was designed with two purposes in mind: first, to classify objects in DPOSS to the faintest limits of the data; second, to fully generalize to future classification efforts, including classification of galaxies by morphology and improving the existing DPOSS star/galaxy classifiers once a larger volume of data are in hand. To optimize the identification of stars and galaxies in J and F band DPOSS scans, we determined a set of eight highly informative object attributes. In the eight-dimensional space defined by these attributes, we found like objects to be distributed relatively uniformly within and between plates. To infer the rules for distinguishing objects in this, but possibly any other, high-dimensional parameter space, we utilize a machine learning technique known as decision tree induction. Such induction algorithms are able to determine statistical classification rules simply by training on a set of example objects. We used high quality CCD images to determine accurate classifications for those examples in the training and set too faint for reliable classification by visual inspection of the plate. Our initial results obtained from a set of four DPOSS fields indicate that we achieve 90% completeness and 10% contamination in our galaxy catalogs down to a magnitude limit of $\sim 19.6^m$ in r and 20.5^m in g, within F and J plates, respectively, or an equivalent B_J of nearly 21.0^m. This represents a 0.5^m–1.0^m improvement over results from previous digitized Schmidt plate surveys using comparable plate material. We have also begun applying methods of unsupervised classification to the DPOSS catalogs, allowing the data, rather than the scientist, to suggest the relevant and distinct classes within the sample. Our initial results from these experiments suggest the scientific promise of such machine discovery methods in astronomy. © *1995 American Astronomical Society.*

1. INTRODUCTION

The first step in analyzing any imaging sky survey is to identify, measure, and catalog all of the detected objects into their respective classes. Once the objects have been measured and classified, further scientific analysis may proceed.

The accuracy of star/galaxy separation generally determines the effective limiting magnitude, in terms of scientific usefulness, of imaging surveys. This limit is, in very many respects, more important than the object detection limit in terms of its impact on the variety of programs for which the data may be used. For example, in order to effectively use the data to compare against models of star or galaxy counts or colors, measure the angular correlation function of galaxies, or search for high redshift quasars, accurate star/galaxy classification is required at the level of approximately 90%. At the faint end, every additional magnitude to which one can extend this accuracy limit buys one an order of two to three times more classified objects in the catalog. Given the enormous resources put into obtaining the survey data in the first place, it makes sense to fully investigate the very latest technology when approaching the task of object classification, in the hope of squeezing every last bit of scientifically useful information from the survey. This was our motivation when designing and implementing the classification methods described in this paper, which are currently being applied to the digitized scans of the Second Palomar Observatory Sky Survey (POSS-II).

POSS-II (Reid *et al.* 1991) is more than 60% complete as of 1994 August, and will eventually cover 894 fields spaced 5° apart in three passbands: blue (IIIa-J+GG 395), red (IIIa-F+RG 610), and near-infrared (IV-N+RG9). The typical limiting magnitudes for point sources in the corresponding J, F, and N bands are 22.5^m, 21.5^m, and 19.5^m, respectively. While the photographic survey is still under way, STScI and Caltech have begun a collaborative effort to digitize the complete set of plates (Djorgovski *et al.* 1992; Laser *et al.* 1992; Reid & Djorgovski 1993). So far, only a subset of the J, F,

 0004-6256/95/109(6)/2401/14/$0.90

and N plates have been scanned and processed. Both the photographic survey and the plate scanning are estimated to be >90% complete circa 1997. The resulting data set, the Palomar–STScI Digital Sky Survey (DPOSS), will consist of ~3 TB of pixel data: ~1 GB/plate, with 1 arcsec pixels, 2 bytes/pixel, 20340^2 pixels/plate, for all survey fields in all three colors. In conjunction with the plate survey, we are also conducting an intensive program of CCD calibrations using the Palomar 60-inch telescope, using the Gunn-Thuan *gri* bands. These CCD images serve both for magnitude zero-point calibration and object classification purposes. The plate scans, when complete, will be the highest quality set of digital images covering the entire northern sky produced to date.

The first scientific results obtained using DPOSS, and making use of the classification methods described herein, are measures of blue and red galaxy counts in four POSS-II fields near the North Galactic Pole (Weir *et al.* 1995). Several additional programs, including a high-redshift quasar search and measures of galaxy–galaxy angular correlations, are in progress (Weir *et al.* 1994).

In order to make most efficient use of DPOSS, and to generally facilitate its scientific exploitation, Caltech Astronomy and the JPL Artificial Intelligence Group have been engaged in a collaborative effort to integrate state-of-the-art computing methods for application to DPOSS. The result of our joint effort is the Sky Image Cataloging and Analysis Tool (SKICAT), a suite of programs designed to facilitate the maintenance and analysis of astronomical surveys comprised of multiple, overlapping images. The classification technology described in this paper was developed as a part of this effort and is implemented within SKICAT (Weir *et al.* 1995).

Historical methods for classifying image features would preclude the identification of the majority of objects in a DPOSS image, since these objects are too faint for traditional recognition algorithms, or even object-by-object classification by eye. A principal goal of SKICAT was to provide an effective, objective, repeatable, and examinable basis for classifying sky objects at levels beyond the limits of previously existing technology. Of course, due to statistical fluctuations of the data, one may never construct a classifier that will be 100% accurate. One may, nevertheless, aim for the highest statistical accuracy achievable to the greatest possible depth.

A particular difficulty in classifying DPOSS objects is that the scan images vary significantly in terms of image quality (e.g., background noise, point-spread function shape, etc.) both within and across plate boundaries. This created an important demand on the classification method to be able to cope with this variation and produce consistent results throughout the survey.

The two essential steps in performing automated object classification are to define the space of discriminating attributes characterizing each object, then determine a means of distinguishing objects within that space. The first step is key, as it determines upon what information any classification will be based. We concentrated a significant amount of effort in deriving a set of object attributes which effectively remove the intra- and interplate variations described above. The second step is likewise very important, as there are any number of ways, some much more powerful than others, of designing rules that partition the parameters space into regions of like objects.

The approach we chose for this second step was one developed in the field of machine learning, namely using decision tree induction algorithms. These methods are able to automatically induce classification rules based simply upon user-supplied examples. This approach not only provided us with the very effective star/galaxy classifiers that already are being used to produce high-quality DPOSS catalogs, but it will easily allow future users to retrain specialized classifiers (e.g., to identify galaxy morphology), or redo existing star/galaxy classifications as more data become available and/or attribute measurement technology improves. Decision tree approaches are particularly robust in dealing with high dimensionality and are computationally efficient to generate (fast training time).

1.1 Historical Approaches

The problem of automatic object classification has been addressed for at least two decades, with a variety of proposed solutions. The most basic approach is to plot one measured attribute versus another and draw a line within that space best separating stars from galaxies. Typically the chosen attributes are magnitude and some measure of object "peakedness," such as peak intensity, isophotal area, or intensity weighted first moment radius. Because in that space, point sources are generally distributed along a fairly well-defined stellar locus, or ridge (see e.g., Fig. 3), such a discriminant function tends to be reasonably accurate down to moderately faint magnitudes. The shortfalls of this approach are that defining the classifier is very labor-intensive as well as subjective, and at faint levels, stars and galaxies quickly blur together around the locus.

The next level of sophistication is to perform star/galaxy separation in a space defined by some nonlinear combination of parameters, rather than raw measurements. For example, simply by plotting the logarithm of isophotal area [**log(Area)**] vs magnitude, instead of just object area, the stellar locus becomes more linear, making a separator much easier to define and generally more accurate. For classifying objects from COSMOS digitized plate scans, Heydon-Dumbleton *et al.* (1989) found it best to discriminate using one of three different pairwise plots depending on an object's magnitude. The three parameters they plotted versus magnitude were: **G**, a measure of how effectively an image fills the ellipse fitted to its major and minor axes, for bright objects; **log(Area)**, for intermediate objects; and a derived parameter, **S**, which effectively measures the scale of a best fit Gaussian to an object's light distribution, for the faintest objects.

Heydon-Dumbleton *et al.* (1989) also improved upon the standard method by making the choice of discriminant line more objective. They measured the statistical distribution of objects around the stellar locus as a function of magnitude, setting the star/galaxy separation line some number of standard deviations above the locus mean.

Picard (1991), in his analysis of COSMOS scans of POSS-II F plates, similarly measured the mean and width of

the stellar locus in **S** vs magnitude space, defining a new parameter, ϕ, corresponding to an object's distance from the locus, normalized by the width of the locus at that magnitude. He binned all the measurements for a given plate and computed a value, $\phi_{\rm cut}$, corresponding to three times the estimated width of the normalized stellar locus. He would then classify all objects with ϕ less than $\phi_{\rm cut}$ as stars, the rest as galaxies. Using this approach, he estimated that he was able to achieve on average 90% completion (fraction of all galaxies classified as such) and 10% contamination (fraction of non-galaxy objects classified as galaxies) in his galaxy catalog down to a magnitude of $19.0^{\rm m}$ in r.

The APM group (Maddox *et al.* 1990) took a slightly different approach to classifying objects from their scans of J plates from the Southern Schmidt survey. Rather than measuring the distance from the stellar locus in the space of one parameter versus magnitude, they used a metric involving the different parameters: peak intensity, radius of gyration, and image area above each of eight surface brightness levels. Two additional parameters were used to help them distinguish blended objects from galaxies, as no deblending algorithm was applied by the APM real-time software in the course of processing. Using this approach, APM reported a classification accuracy comparable to Picard's at a B_J magnitude of $20.0^{\rm m}$.

A far different method for classifying objects from plate scans was pioneered by Sebok (1979) in his Ph.D. thesis at Caltech. He introduced the concept of Bayesian classification to the problem, estimating the most probable classification of each object based upon its fit to a set of models. While this approach was effective, it was never widely applied to Schmidt plate surveys subsequently.

Sebok's classification method preceded the similar approach devised by Valdes and implemented in modern versions of FOCAS (Valdes 1982). Valdes also applied a technique premised on Bayesian probability theory, but more significantly, he introduced a measurement procedure that results in discriminating object attributes. By selecting a number of objects in an image that are "sure-thing" stars, FOCAS adds the rasters of the central pixels of these objects to form an empirical estimate of the point-spread function (PSF) for that image. Using the "resolution" routine, FOCAS then fits a model to each object consisting of a pure PSF component and a blurred version of the same. The best-fitting fraction of blurred component and its scale are the two attributes resolution measures and uses for performing object classification. These attributes have never been used in large scale digitized plate surveys to date because computing technology prevented the repeated access to the pixel data, which this technique requires.

FOCAS provides a default set of rules specifying to which class different portions of fraction vs. scale space correspond. Because the distribution of objects in the space of these attributes tends to be relatively invariant from image to image (PSF variations are effectively taken into account by the fitting process), the default rules are found to provide excellent classification accuracy down to fairly faint levels for a wide variety of images. The user has the option of changing these classification rules, but FOCAS does not provide a way of allowing for more attributes in the rules, or a systematic way for determining a new, more effective set of rules for a particular type of image.

1.2 The Machine Learning Approach

Drawing upon these previous efforts, we chose to measure and calculate those object attributes found to provide the best star/galaxy discrimination. However, unlike most previous approaches, we chose to apply modern methods from the field of machine learning to determine the optimal discriminant functions, or set of classification rules, within the multidimensional space of these measurements. The goal when applying these methods is to provide enough examples of accurate classifications to the algorithm to allow it to infer the rules for distinguishing objects in that space. An important advantage of this approach is that one can typically feed a relatively large number of input parameters to the algorithm, allowing it to determine classification rules more complex than those typically devised by humans, generally as a result of examining pairwise plots of attributes. The extra degrees of freedom provided by learning in multidimensional parameter space often lead to substantially more accurate classifications. In addition, the rules are formed in an objective, repeatable fashion.

Others have also begun exploring the use of new machine learning methods for the purpose of object classification, perhaps most notably the APS group in Minnesota, who have digitized the plates of the original POSS (Odewahn *et al.* 1992). They applied artificial neural networks to the task of automatically inducing a set of classification rules for objects in their catalog. We, too, experimented with neural nets; however, for reasons discussed below, we chose to use a method involving decision trees, based on the work of Fayyad (1991), for creating the production-line classifier implemented within SKICAT and used on DPOSS.

2. CLASSIFIER INDUCTION

For a detailed discussion of decision trees and associated methods of machine learning, we refer the reader to Fayyad (1991) and Fayyad & Irani (1992). Below we include a brief discussion and history of these methods, in particular those we utilize within SKICAT, in addition to a comparison of this approach with neural networks.

2.1 Decision Trees

A particularly efficient method for extracting rules from data is to generate a decision tree (Breiman *et al.* 1984; Quinlan 1986). A decision tree consists of nodes that represent tests on attribute values. The outgoing branches of a node correspond to all the possible outcomes of the test at the node, thus partitioning the examples at a node along the branches. For example, as illustrated in Fig. 1, at the topmost (root) node, the tree may branch left or right depending on whether the object has **log(Area)** less than or greater than $\mathbf{A}_0$. In turn, either of these branches may lead to a node that conditions on the same attribute, a different one, or any combination of the same [e.g., "branch left if (mag$<m_0$) and

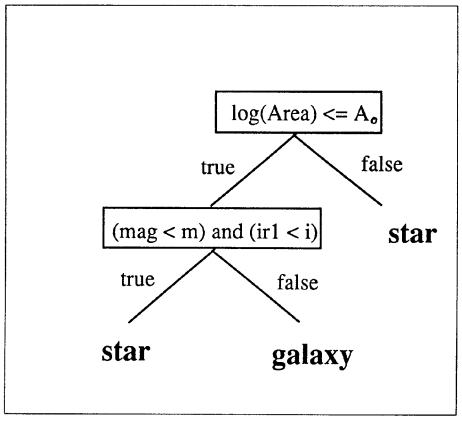

FIG. 1. In this sample decision tree, one starts at the top node(root), following the appropriate path to a final leaf (class) based upon the truth of the assertion at each node.

$(\phi > \phi_0)$"]. The final nodes in the tree, the leaves, would correspond to an actual classification: star, galaxy, artifact, etc.

In Fig. 2 we illustrate a portion of a much larger actual decision tree generated by the O-Btree algorithm (described below) for performing star/galaxy classification. The interval appearing above each node indicates the range in value of the attribute specified in the node above that an object must meet for it to pass along that branch. The dark branches lead to actual classifications. A full path from the root to any particular leaf corresponds to a single classification rule. The number in parentheses within each leaf indicates the number of training examples classified correctly by that rule.

A well-known algorithm for generating decision trees is Quinlan's ID3 (Quinlan 1986) with extended versions called C4 (Quinlan 1990). ID3 starts with all the training examples at the root node of the tree. An attribute is selected to partition the data. For each value of the attribute, a branch is created and the corresponding subset of examples that have the attribute value specified by the branch are moved to the newly created child node. The algorithm is applied recursively to each child node until either all examples at a node are of one class, or all the examples at that node have the same values for all the attributes. Every leaf in the decision tree represents a classification rule. Note that the critical decision in such a top-down decision tree generation algorithm is the choice of attribute at a node. Attribute selection in ID3 and C4 is based on minimizing an information entropy measure applied to the examples at a node. The measure favors attributes that result in partitioning the data into subsets that have low class entropy. A subset of data has low class entropy when the majority of examples in it belong to a single class. For a detailed discussion of the information entropy selection criterion see Quinlan (1986), Fayyad (1991), and Fayyad & Irani (1992).

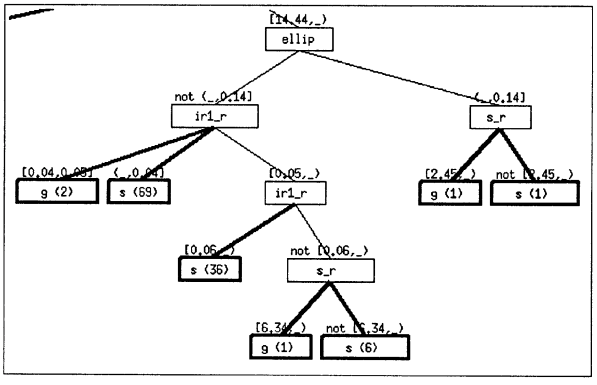

FIG. 2. A portion of a much larger actual decision tree generated by the O-Btree algorithm for performing star/galaxy classification. The interval appearing above each node indicates the range in value of the attribute specified in the node above that an object must meet for it to pass along that branch. The dark branches lead to actual classifications. The number in parentheses within each leaf indicates the number of training examples classified correctly at that node.

2.1.1 The GID3 and O-Btree algorithms*

The attribute selection criterion clearly determines whether a "good" or "bad" tree is generated by a greedy algorithm (see Fayyad & Irani 1990 and Fayyad 1991 for the details of what we formally mean by one decision tree being better than another). Since making the *optimal* attribute choice is computationally infeasible, ID3 utilizes a heuristic criterion which favors the attribute that results in the partition having the least information entropy with respect to the classes. There are weaknesses inherent in algorithms like ID3/C4 due to the fact that, for discrete attributes, a branch is created for each value of the attribute chosen for branching. This overbranching is problematic since in general it may be the case that only a subset of values of an attribute are of relevance to the classification task while the rest of the values may not have any special predictive value for the classes. The GID3* algorithm (Fayyad 1994) was designed mainly to overcome this problem, generalizing the ID3 algorithm so that it does not necessarily branch on each value of the chosen attribute. GID3* can branch on arbitrary individual values of an attribute and "lump" the rest of the values in a single default branch. Unlike the other branches of the tree which represent a single value, the default branch represents a subset of values of an attribute. Unnecessary subdivision of the data may thus be reduced. See Fayyad (1991) and Fayyad (1994) for more details and for empirical evidence of improvement.

The O-Btree algorithm (Fayyad & Irani 1992) was designed to overcome problems with the information entropy selection measure itself. O-Btree creates strictly binary trees and utilizes a measure from a different family of measures that detect class separation rather than class impurity. Information entropy is a member of the class of impurity measures. O-Btree employs an orthogonality measure rather than entropy for branching. For details on problems with entropy measures and empirical evaluation of O-Btree, the reader is referred to Fayyad (1991) and Fayyad & Irani (1992). Both O-Btree and GID3* differ from ID3 and C4 in one additional aspect: the discretization algorithm used at each node to discretize continuous-valued attributes. Whereas ID3 and C4 utilize a binary interval discretization algorithm, we utilize a

generalized version of that algorithm which derives multiple intervals rather than strictly two. For details and empirical tests showing that this algorithm does indeed produce better trees, see Fayyad (1991) and Fayyad & Irani (1993). We have found that this capability improves performance considerably in several domains.

2.2 The RULER System

There are limitations to decision tree generation algorithms that derive from the inherent fact that the classification rules they produce originate from a single tree. This fact was recognized by practitioners early on (Quinlan 1986). The basic problem is that in even a good tree, there are always leaves that are overspecialized or predict the wrong class. For example, if there are any measurement errors in the attributes, the decision tree will tend to fit to the noise and, hence, not generalize well to data that are out of sample. The very reason that makes decision tree generation efficient (the fact that data is quickly partitioned into ever smaller subsets) is also the reason why overspecialization or incorrect classification occurs. It is our philosophy that once we have good, efficient decision tree generators, they can be used to generate multiple trees, and from these, only the best rules in each are kept. To implement this strategy, the algorithm RULER was developed (Fayyad *et al.* 1992).

In multiple passes, RULER partitions a training set randomly into a training subset and test subset. A decision tree is generated from the training set and its rules are tested on the corresponding test set. Using Fisher's exact test (Finney *et al.* 1963), the exact hypergeometric distribution, RULER evaluates each condition in a given rule's preconditions for relevance to the class predicted by the rule. It computes the probability that the condition is correlated with the class by chance.[1] If this probability is higher than a small threshold (say 0.01), the condition is deemed irrelevant and is pruned. In addition, RULER also measures the merit of the entire rule by applying the test to the entire precondition as a unit. This process serves as a filter which passes only robust, general, and correct rules.

By gathering a large number of rules through iterating on randomly subsampled training sets, RULER builds a large rule base of robust rules that collectively cover the entire original data set of examples (i.e., every example is classified by a rule). A greedy covering algorithm is then employed to select a minimal subset of rules that covers the examples. The set is minimal in the sense that no rule could be removed without losing complete coverage of the original training set. Using RULER, we can typically produce a robust set of rules that has fewer rules than any of the original decision trees used to create it, and that generalizes better to out-of-sample data. The fact that decision tree algorithms constitute a fast and efficient method for generating a set of rules allows us to generate many trees without requiring extensive amounts of time and computation.

[1]The Chi-square test is actually an approximation to Fisher's exact test when the number of test examples is large. We use Fisher's exact test because it is robust for both small and large data sets.

We implemented the RULER algorithm, in conjunction with GID3* and O-Btree, within SKICAT for the purpose of inducing classification rules by example, and it was used to produce the particular star/galaxy classifiers described subsequently. Throughout this paper, we generally refer to our technique as decision tree induction and the rules as decision trees. We simply note that in practice we are actually referring to the use of decision trees in conjunction with the RULER tree pruning and combining algorithm.

2.3 Decision Trees Versus Neural Nets

In order to compare against other learning algorithms, and to preclude the possibility that a decision tree based approach is imposing *a priori* limitations on the achievable classification levels, we tested several neural network algorithms for comparison. The results indicate that neural nets achieve similar performance as decision trees. The learning algorithms we tested were traditional backpropagation, conjugate gradient optimization, and variable metric optimization of a two-layer perceptron (see Hertz *et al.* 1991 for an introduction to perceptrons and neural methods of computation). The latter two are training algorithms that work in batch mode and use standard numerical optimization techniques in changing the network weights. Their main advantage over backpropagation is the significant speed-up in training time.

The results of our comparison between these approaches and decision trees can be summarized as follows. The performance of the neural networks was a fairly unstable function of the random initial network weights chosen prior to training and produced accuracy levels on a sample test set of data varying between 30% (no convergence) and 95%, compared with a 94% accuracy level for a decision tree classifier. The most common range of accuracy averaged between 76% and 84%. The achieve these levels of accuracy, we had to perform multiple trials, each time varying the number of internal nodes in the hidden layer, the initial network weight settings, and the learning rate constant for backpropagation.

Upon examining the results of this empirical study, we concluded that the neural net approach did not offer any clear advantages over the decision tree based learning algorithms. Although neural networks, with extensive training and several training restarts with different initial weights to avoid local minima, could match the performance of the decision tree classifier, the decision tree approach still holds several major advantages. For one, the tree is more easily interpreted than the weights in a neural network (although, admittedly, a list of 20 rules that condition on up to eight parameters is not entirely transparent either). More importantly, the learning algorithms we employ do not require the specification of parameters such as the size of the neural net or the number of hidden layers, nor do they call for random trials with different initial weight settings. There are, in fact, very few free parameters. This makes the decision tree algorithm much easier to implement as a generic tool within SKICAT. Also, the required training time is orders of magnitude faster than the training time required for a neural network program (i.e., seconds rather than dozens of minutes in some cases).

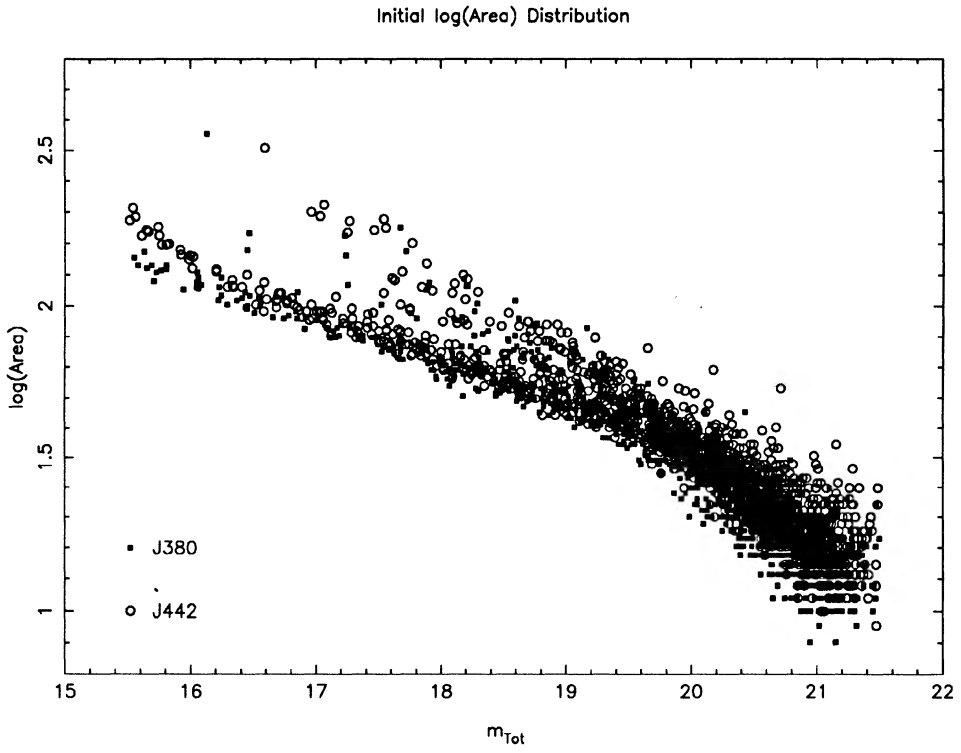

FIG. 3. The distribution of **log(Area)** vs **MTot** in sections of plates J380 and J442. Note that the stellar locus is nonlinear and different for each plate. The locus shows similar variance even within plates.

3. CLASSIFICATION ATTRIBUTES

In classification learning, the choice of attributes used to define examples is by far the single most important factor determining the success or failure of the learning algorithm. The attributes we use for classification are computed through a combination of image processing and statistical measurement techniques. While they are not expected to be the final advancement in this area, we did find them to provide the most discriminating and uniform characterization of objects detected in DPOSS of any other set of attributes we have encountered. This section provides a detailed description of these attributes and how they are computed.

3.1 Base-level Attributes

The eight attributes we use in object classification include a compendium of measures found to be most useful and discriminating in previous surveys. They include:

MTot—the FOCAS total instrumental magnitude;

MCore—the core magnitude, measured from the brightest 3×3 pixel region in the object;

log(Area)—the log of the isophotal area of the object;

Ellip—the ellipticity;

IR1—the intensity weighted first moment radius:

$$\mathbf{IR1} = \frac{\Sigma_k i_k r_k}{\Sigma_k i_k},$$

where i_k is the intensity of pixel k and r_k is its distance from the object's centroid;

S—the parameter defined by Heydon-Dumbleton *et al.* (1989) and used by Picard (1991), which is a function of object area (a), core intensity (I_{core}, the sum of the central 3×3 pixels), and the average intensity along the detection isophote (p):

$$\mathbf{S} = \frac{a}{\log[I_{\text{core}}/9\times p)]}.$$

We chose FOCAS total magnitudes for our standard brightness measure for its decreased sensitivity to the surface brightness threshold relative to aperture or isophotal magnitudes (see Weir *et al.* 1995). The other attributes measure the object's symmetry or compactness in one way or another. FOCAS measures the two listed magnitudes and **IR1** directly, while the other three are easily computed from actual measurements. We tested the use of a few additional object parameters, such as additional image moments, but found that they contributed little additional discriminatory power due to their high correlation with one or more of these parameters. There is always the possibility that future researchers will find that some unconsidered parameter helps result in significantly improved classifications, and the machine learning software is fully capable of incorporating additional new parameters as they are discovered. For now, however, we found that this list is sufficient.

Like previous researchers (e.g., Valdes 1982; Heydon-Dumbleton *et al.* 1989; Picard 1991), we quickly determined that the distribution of these base-level attributes does not

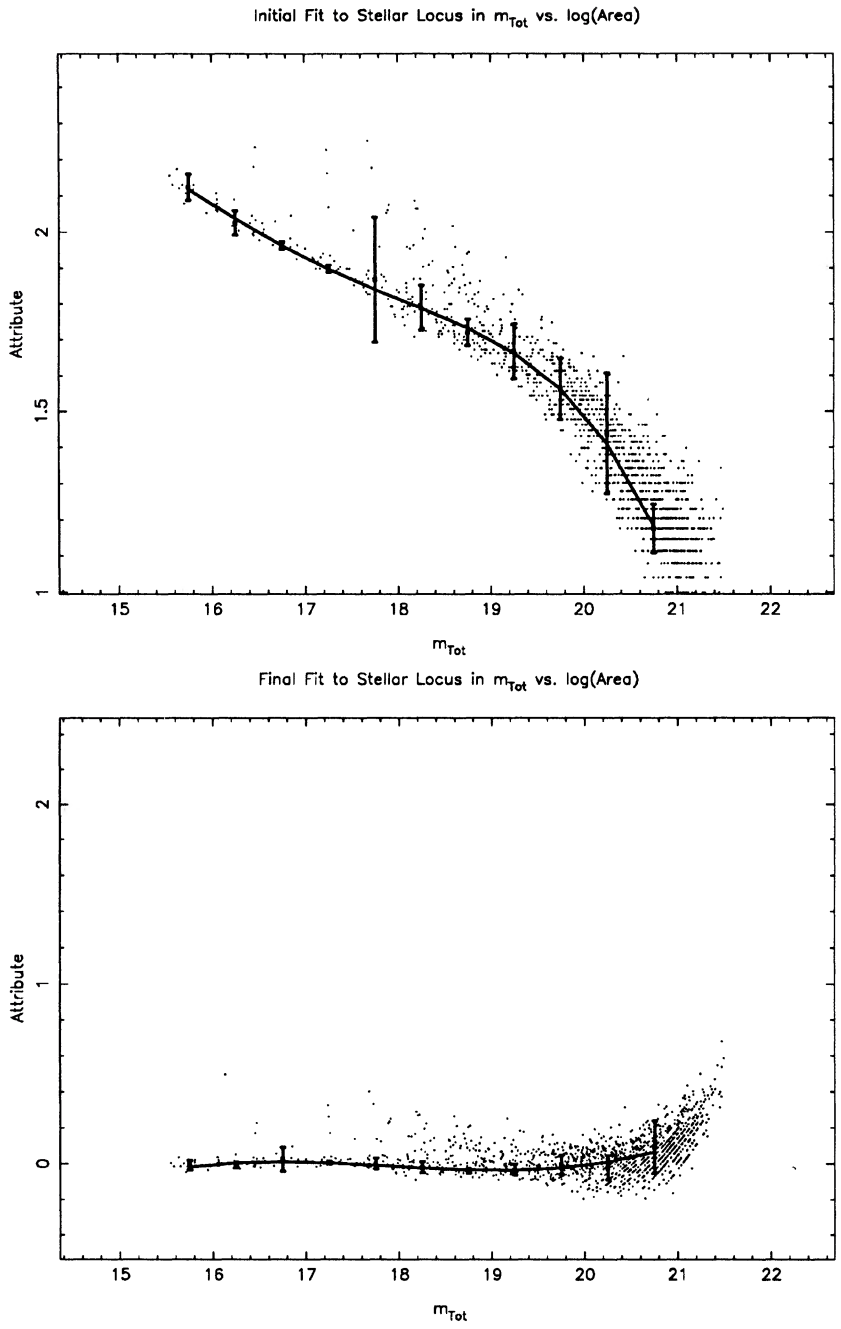

FIG. 4. The **log(Area)** attribute and the locus fit to its distribution before each iteration of the locus subtraction algorithm.

exhibit the required invariance between different regions of a single plate, much less across plates. This was exhibited by the low out-of-sample accuracy of the classifiers we produced by training on these attributes alone. Their variability is also clearly evident when one looks at the distribution of these parameters across or within plates. For example, in Fig. 3, we plot the distribution of **log(Area)** vs. **MTot** for two 2048^2 pixel sections of plates J380 and J442. We analyze each plate in image sections of this size (which we call footprints) to help account for variations in image quality across the plate (see Weir *et al.* 1995 for a full discussion of our plate reduction procedure). Note that the stellar loci for these two footprints are nonlinear and do not overlay one another. The implication is that a classifier optimized for one of the images would not only be difficult to construct due to the nonlinearity of the stellar locus, but it would certainly be less than optimal for the other image.

Raw measurements of object shape are inherently sensitive to the local background sky level, seeing, and the pixel blurring induced by the scanning process. We therefore expect these measurements to vary from plate to plate and even footprint to footprint. For any learning algorithm to be able

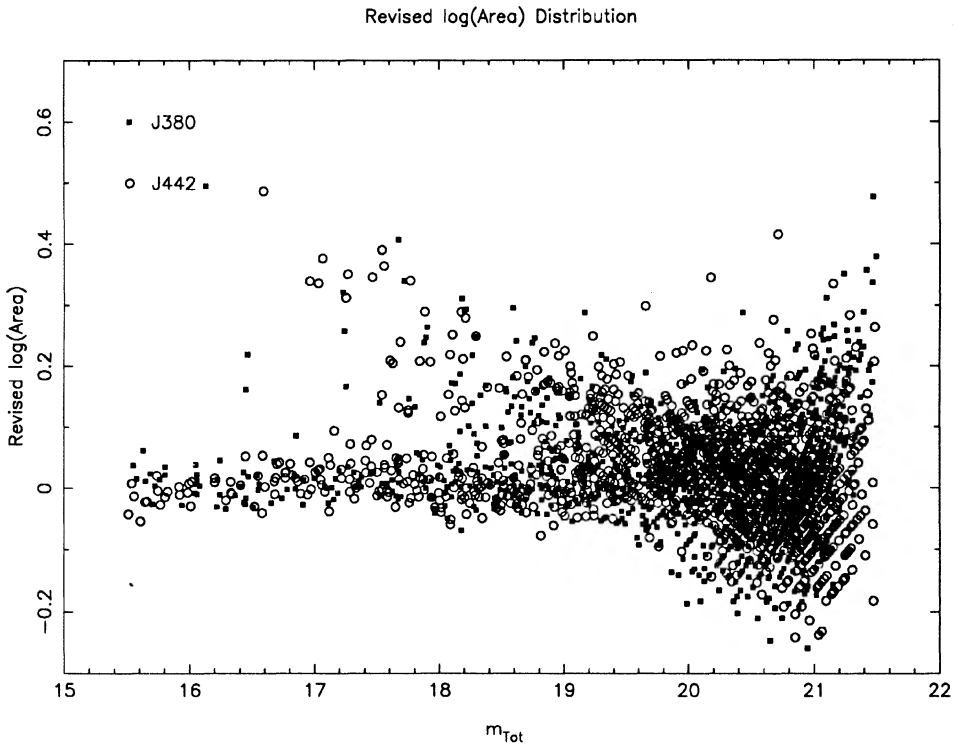

FIG. 5. The distribution of the revised **log(Area)** vs instrumental magnitudes in plates J380 and J442 after the two-step locus fitting and subtracting process.

to produce robust classifiers consistent across a large survey area, different attributes are clearly required.

3.2 Derived Attributes

As we discussed in Sec. 1, the resolution routine of Valdes (1982) provides two extremely powerful classification parameters that, by construction, are very uniformly distributed from image to image. In fact, a preliminary study by Weir & Picard (1991) indicated the significant benefits of using the FOCAS approach to object classification on digitized Schmidt plates. They found that using the PSF-fitting algorithm, one could extend the limiting magnitude of classified Schmidt plate catalogs nearly a full magnitude beyond previous limits achieved using historical approaches.

An essential task in employing the resolution technique, however, is to establish an accurate estimate of the PSF for a given image. Only after this is obtained can the resolution scale and fraction parameters be measured. The problem, therefore, naturally breaks up into two separate steps: (1) star selection, the process of automatically deriving a list of candidate stars for generating an empirical PSF template; and (2) final classification, in which the resolution parameters, possibly along with others, are used for assigning all objects to a particular class.

As previous surveys indicate, certain rather simplistic methods are perfectly adequate for performing accurate star/galaxy separation at bright to moderately faint magnitudes: a method involving PSF-fitting is necessary only when approaching a magnitude or so within the detection limit. One need not approach this limit just to produce lists of stars for empirically estimating the PSF template. Using a straightforward approach similar to ones used for final classification in previous surveys, we were able to develop a technique for robustly selecting candidate PSF stars, up to some limiting magnitude, uniformly within and among plates.

The solution we employ is to fit, on a footprint by footprint basis, the stellar locus within four separate parameter versus magnitude projections, measuring four new attributes in the form of the distance of each object from the stellar ridge in each dimension. We compute these so-called "revised" attributes for the $\mathbf{M}_{\text{core}}$, **log(Area), IR1,** and **S** parameters described above. We find that in these new parameter spaces, the line distinguishing stars from galaxies is roughly linear and does not vary much from image to image.

Measuring the distance of an object from the stellar locus first requires the ability to delineate the location of the locus. The method we use for automatically tracking the locus in an attribute versus magnitude parameter space works by computing a histogram of the attribute value in a set of 0.5^{m} bins spanning the instrumental magnitude ranges 15.5^{m}–21.5^{m} in J_{inst} and 15.5^{m}–20.5^{m} in F_{inst}. Objects brighter than the lower magnitude limit are typically saturated and must be classified separately; and one has little hope of forming accurate star lists using this type of method at magnitudes fainter than the upper limit.

Our locus tracking algorithm next computes robust estimates of the mode and width of the histogram for each magnitude bin. These mode values and their error estimates (specified by the widths) are then fit by a fourth or fifth order polynomial as a function of magnitude (see Fig. 4). The fit is

subtracted from each object, effectively bringing the stellar ridge close to the abscissa on an attribute versus magnitude plot. To assure an optimal fit to the stellar ridge, the algorithm applies the same fitting and subtraction procedure a second time, this time using a third- or higher order polynomial. The optimal orders used to perform the fit in the first and second iterations were found to be very consistent across all DPOSS images and were determined separately for each of the four parameters. These fitting parameters were ultimately hard-coded into the measurement process. Other researchers found it useful to renormalize the new attribute values by the width of the stellar locus. Our tests did not indicate significant variations in the widths of the revised attribute distributions from footprint to footprint, so we eliminated this step.

The distribution of the revised parameters derived for the objects shown in Fig. 3 appear in Fig. 5. As demonstrated in this example, we find that the distribution of objects in revised attribute space differs little between plates. The same holds true for the other revised attributes we compute, as well.

Along with magnitude and ellipticity, the four revised attributes now form a six-dimensional parameter space in which we perform star/galaxy separation. To produce our star selector classifier, we trained the decision tree induction software on a set of over a thousand objects which one of us (NW) classified by eye from the digitized scans of plates J380 and J442. Subsequent comparison with several hundred much more reliable classifications obtained from CCD images indicated an error rate of less than 5% in the training list constructed by eye.

The star selector we produced had an error rate of less than 3% percent on an out-of-sample list of objects from the same two plates in the instrumental magnitude range 16.5^{m} to 19.0^{m}. Subsequent application of the classifier on independent *J and F* data resulted in lists of candidate stars in this magnitude range which we found to be more than accurate enough for use in constructing the PSF template required by FOCAS resolution. Whereas the typical footprint contains between 3500 to 4500 objects, the star selector returns between 500 and 600 objects in the magnitude range listed above. This list of candidate stars is provided to a FOCAS routine which averages the central nine by nine pixels of each object to form the PSF template.

Armed with the template, one is then able to run the FOCAS resolution routine on each object. As described previously, this routine determines the best-fitting scale (α) and fraction (β) values, which parametrize the fit of a blurred (or sharpened) version of the PSF to each object. The template used to model each object is of the form

$$t(r_i)=\beta s(r_i/\alpha)+(1-\beta)s(r_i),$$

where r_i is the position of pixel i, α is the broadening (sharpening) parameter, and β is the fraction of broadened PSF. In turn, the resolution parameters are combined with the previous six used for star selection in order to perform final object classification.

4. CLASSIFICATION RESULTS

In the course of processing each plate, the attribute measurement tasks described in the previous section, including revised attribute measurement and star selection, are performed fully automatically, as is the task of final object classification. However, in order to produce the classifiers implemented within the DPOSS reduction programs, we were required at some point to manually produce large samples of classified objects for training and testing purposes. We describe how we produced these training samples below. The same steps would be required of any user who might wish to construct their own, specialized classifier, or to improve upon or monitor the quality of the existing classifiers on future data. We follow this discussion with an examination of the results of applying these classifiers to actual DPOSS data.

4.1 Classifier Training

In order to obtain training data for classifying objects in DPOSS, especially those too faint for recognition by human inspection of the plates alone, we made use of higher resolution (and narrower field of view) CCD imagery obtained from the Palomar 60″ telescope. CCD images are being collected systematically in order to photometrically calibrate the Survey (se Weir *et al.* 1995); however, they serve this very important role in the object classification process as well.

For classification purposes, the obvious advantage of a CCD image relative to a plate is higher resolution and signal-to-noise ratio at fainter levels. By matching a CCD image with the corresponding (small) portion of the plate that it covers, one can determine the classes of objects too faint to classify by eye on the plate. By training learning algorithms to classify these faint objects correctly using the attributes derived from the plate image, SKICAT can conceivably classify objects from the survey that even humans would have difficulty classifying.

The training and test data consisted of objects collected from four different plate fields from regions for which we had CCD image coverage, as well as the by-eye classifications used to construct the star selector described in the previous section. To adequately test the reliability of the classifier, we divided the data into independent training and test sets from different plates. The *F* plate training sample totaled 1239 objects from plates F381 and F442, while the *J* sample consisted of 2563 objects from plates J380 and J342.

We trained the decision tree induction and combining algorithms, O-Btree and RULER, separately on the *J* and *F* data in order to produce independent classifiers. The results of our classifier training were a list of 84 rules for the *F* plate classifier and 96 for the *J*'s. Each rule is effectively an "if-...then..." statement assigning a class to any object meeting its conditions. For both classifiers, each rule conditions upon anywhere from three to six different parameters. By construction, as described in Sec. *3.2.2*, the rules will generate a unique classification for any object within the training set's multidimensional parameter space.

These classifiers represent our initial production versions of DPOSS object classifiers. There remain a substantial num-

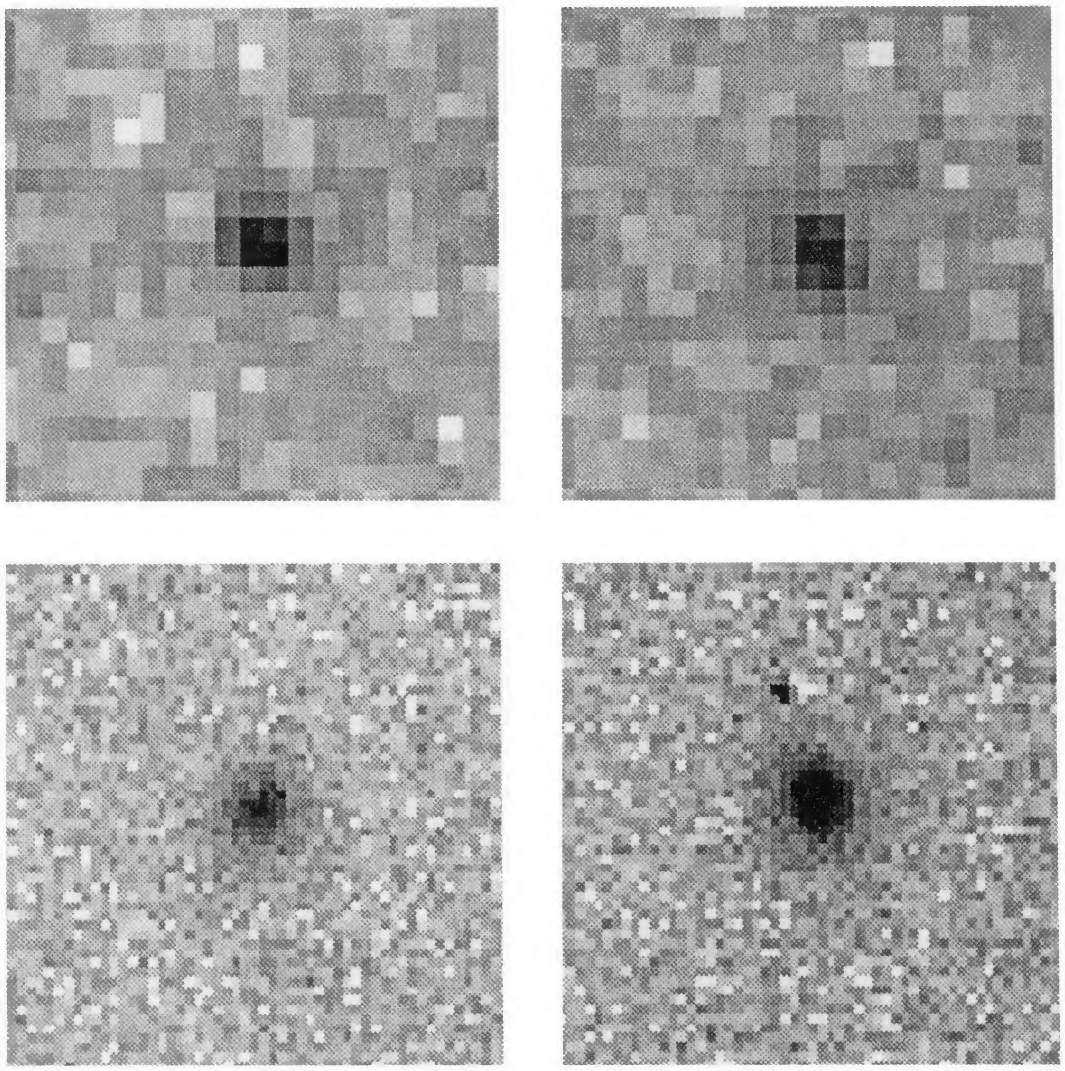

FIG. 6. The top two images are from the scan of plate J442. Each object has a g magnitude of approximately 20.0. The bottom two images are of the same objects but from CCD frames. Our classifier correctly classified the left object as a star and the right as a galaxy, despite their almost indistinguishable appearance on the plate. The higher quality CCD images allowed us to provide reliable classifications to these objects which we would otherwise be unable to use in classifier training or testing.

ber of avenues for extensive further research. For example, as opposed to producing classifiers that work independently on each plate, one might train a classifier which combines the information available for objects matched in multiple images, particularly in both the J and F colors. Also note that for these initial production classifiers, we restricted the number of classes to just two: stars and galaxies. The next generation of classifiers that we will produce after we have obtained sufficient training data will include additional classifications, including an artifact class and more refined morphological distinctions. In addition, though we have not done so here, it is possible to assign uncertainties to the classifications produced by decision trees. The easiest method is simply to calculate the statistical error rate of each rule on an independent test data set. These sample statistics are then used as confidence estimates on subsequent out-of-sample classifications. A more sophisticated method, involving a kernel density estimation for each classification rule in the parameter space referenced by each rule, is described in Smyth *et al.* (1995). We intend to implement one of these two methods of estimating classification uncertainties in future versions of our classifiers.

4.2 Comparisons with Training and Test Data

We tested the classifiers on a sample of 1539 objects from plates F380 and F382 and 589 objects from plates J381 and J382. Testing consisted simply of keeping track of the fraction of objects classified correctly or incorrectly as a function of magnitude. It is noteworthy that for a large fraction of these objects, an astronomer would have difficulty reliably determining their classes by examining the corresponding digitized plate images. As an example, see Fig. 6, which depicts a star and galaxy as it appears on a plate and on a CCD. These objects are representative of those with a magnitude at the limit of which we would like to perform accurate star/galaxy separation. We have begun spectroscopic follow-up observations of a sample of the small, faint objects, providing another independent check on our faint classifications.

TABLE 1. The completeness (fraction of galaxies classified as such) and contamination (fraction of nongalaxies classified as galaxies) for the samples of F plate objects used for classification training and testing. The training samples are from plates F381 and F442. The testing samples are from plates F380 and F382.

	Training set		Testing set	
r mag	Completeness	Contamination	Completeness	Contamination
16.56	1.000	0.000	0.857	0.000
16.96	1.000	0.000	0.833	0.062
17.43	0.938	0.032	0.966	0.034
17.95	0.979	0.000	0.885	0.042
18.50	0.966	0.012	0.878	0.133
19.07	0.969	0.054	0.929	0.103
19.64	0.985	0.043	0.895	0.094
20.21	0.948	0.081	0.906	0.247
20.75	0.950	0.102	0.902	0.260

TABLE 2. The completeness and contamination for the samples of J plate objects used for classification training and testing. The training samples are from plates J380 and J442. The testing samples are from plates J381 and J382. Too few objects of bright magnitude were available to provide a statistically significant test below $g=17.5^{m}$.

	Training set		Testing set	
g mag	Completeness	Contamination	Completeness	Contamination
16.68	1.000	0.000	***	***
17.17	0.857	0.077	***	***
17.67	0.935	0.333	1.000	0.000
18.18	0.956	0.030	1.000	0.091
18.69	0.989	0.021	1.000	0.050
19.21	0.963	0.037	0.966	0.097
19.73	0.954	0.019	0.925	0.098
20.25	0.964	0.024	0.892	0.065
20.77	0.891	0.039	0.861	0.151
21.30	0.806	0.167	0.796	0.204
21.81	0.848	0.200	0.774	0.250

The accuracy we achieved from applying the classifiers on the training and test DPOSS data sets appears in Tables 1 and 2. We estimate the accuracy by measuring the completeness and contamination of a galaxy catalog formed from the sample data. The training results reflect the in-sample accuracy of the classifier, which is largely irrelevant and included only for completeness. The test set results are indicative of the accuracy of the classifier on independent data and, therefore, reflect the true quality of the classifier. These results are plotted in Fig. 7.

Note that on our test data, we achieve approximately 90% completeness and 10% contamination down to $r\sim19.5^{m}$ and $g\sim20.5^{m}$, or an equivalent B_J of approximately 21.0^{m}. This reflects an accuracy rate comparable to what previous surveys attained, but at magnitude levels 0.5^{m} to 1.0^{m} fainter. Our limited spectroscopic follow-up observations to date are fully consistent with these results.

Though not listed here, we also computed the results of

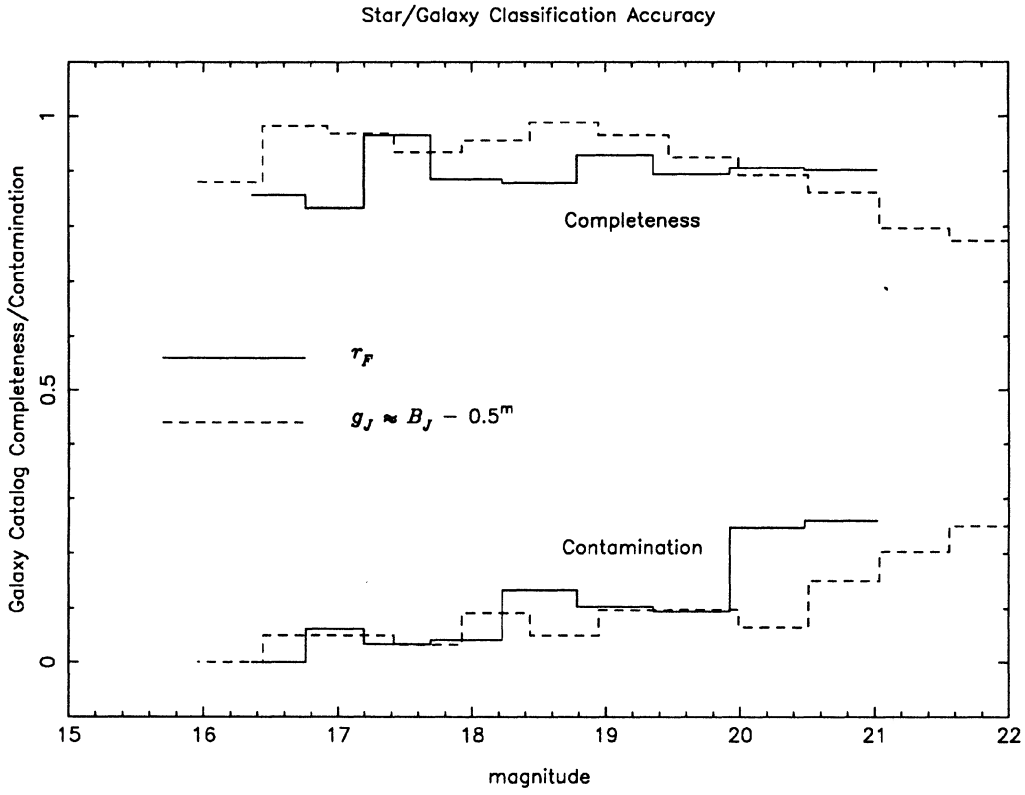

FIG. 7. The accuracy of our star/galaxy separation technique is depicted by the completeness (fraction of galaxies classified as such) and contamination (fraction of nongalaxies classified as galaxies) measured within our test set of data.

TABLE 3. The fraction of objects classified consistently as a function of magnitude in the overlap of the listed plates. These rates are consistent with the accuracies listed in Table 1. The number of objects tested in each overlap is listed below the field names.

r mag	F380/F381 (8682)	F380/F442 (1357)	F381/F382 (9246)	F381/F442 (3865)	Average
16.23	0.933	0.947	0.880	0.886	0.912
16.56	0.952	0.967	0.957	0.964	0.960
16.96	0.952	0.839	0.968	0.971	0.932
17.43	0.972	0.870	0.937	0.925	0.926
17.95	0.964	0.983	0.957	0.984	0.972
18.50	0.941	0.919	0.961	0.969	0.948
19.07	0.893	0.899	0.921	0.875	0.897
19.64	0.825	0.855	0.826	0.852	0.840
20.21	0.746	0.773	0.761	0.749	0.757
20.75	0.750	0.738	0.681	0.743	0.728
21.25	0.746	0.753	0.718	0.775	0.748

TABLE 4. Same as Table 3, but for J plates.

g mag	J380/J381 (8553)	J380/J442 (1418)	J381/J382 (9659)	J381/J442 (3850)	Average
15.73	0.548	0.533	0.623	0.538	0.561
16.20	0.913	0.846	0.899	0.860	0.880
16.68	0.977	1.000	0.954	1.000	0.983
17.17	0.976	0.964	0.975	0.962	0.969
17.67	0.962	0.972	0.979	0.972	0.971
18.18	0.975	1.000	0.985	0.964	0.981
18.69	0.958	0.984	0.970	0.962	0.968
19.21	0.915	0.927	0.942	0.890	0.918
19.73	0.857	0.911	0.881	0.874	0.881
20.25	0.755	0.812	0.780	0.820	0.792
20.77	0.688	0.759	0.717	0.690	0.713
21.30	0.673	0.671	0.717	0.665	0.681
21.81	0.736	0.701	0.706	0.661	0.701

the J classifier on a test set of data from the same plates on which the classifier was trained. The completeness and contamination closely matched that of the test set from independent plates. Therefore, we can expect the performance of the classifiers to be virtually the same for large catalogs of objects from either the training or test sets of plates. We can help confirm this expectation by comparing the consistency of classifications from plate to plate, as we do below.

We also confirmed the relative importance of the resolution attributes for object classification. When the same experiments were conducted using only the six attributes used in star selection, the results were significantly worse. The error rates jumped above 20% for O-BTree, above 25% for GID3*, and above 30% for ID3 at a magnitude of approximately 20.0^m in g. The respective sizes of the trees grew significantly as well. This clearly demonstrates that although learning algorithms improve matters considerably by allowing one to effectively and objectively make use of multiple parameters in the classification process, the choice of parameters is still of first order importance.

4.3 Comparisons in Plate Overlaps

The tests described above indicate an overall classification accuracy of approximately 90% at a magnitude of approximately 19.6^m in r and 20.5^m in g. If we assume that the probability of an object being correctly classified is independent from plate to plate, this would imply a consistency of classification of approximately 82%. This is the sum of the probabilities of both classifications being correct (0.9^2) or incorrect (0.1^2). Measuring the consistency of classifications from plate to plate across many different plates provides some measure of the uniformity of plate classification accuracies, if not their actual levels of accuracy. In Tables 3, 4, and 5 we list the consistency of object classification for the large number of objects measured in each pair of overlapping plates of the same color and overlapping plates of the same field but different color. Note that at each magnitude level, the consistencies are in line with the accuracies listed in the previous section assuming independent classifications.

Also notice that the consistency of the classifications between the pairs of plates on which the classifiers were trained (F381/F442 and J380/J442) does not significantly differ from the consistency of other measured pairs. This corroborates the notion that the classification accuracy for these plates as a whole is no better or worse than that for the test plates, despite the fact that the classifiers were trained exclusively on objects from those plates. In this sense, the classifiers are truly robust.

5. INITIAL EXPERIMENTS WITH UNSUPERVISED CLASSIFICATION

We have also begun exploring the application and implementation of unsupervised classification techniques like Autoclass (Cheeseman *et al.* 1988) for the purpose of automated machine discovery. Unlike the so-called supervised methods of classification that we have described so far, where the computer learns how to distinguish user-specified classes within the data, unsupervised classification consists of the computer identifying the statistically significant classes within the data itself. For example, one could employ this type of method to try to systematically detect new classes of objects within astronomical catalogs.

TABLE 5. The fraction of objects classified consistently as a function of average g and r magnitude in the overlap of the J and F plates covering the indicated fields. The number of objects tested in each overlap is listed below the field names.

	Field				
$r+g$ mag	380 (7096)	381 (8456)	382 (7660)	442 (7900)	Average
15.95	0.656	0.795	0.777	0.986	0.803
16.45	0.983	0.953	0.992	0.953	0.970
16.95	0.948	0.982	0.962	0.970	0.966
17.45	0.977	0.964	0.966	0.951	0.964
17.95	0.981	0.964	0.986	0.948	0.970
18.45	0.940	0.952	0.967	0.950	0.952
18.95	0.926	0.926	0.928	0.943	0.931
19.45	0.866	0.859	0.901	0.885	0.878
19.95	0.804	0.768	0.818	0.787	0.794
20.45	0.729	0.682	0.763	0.713	0.722
20.95	0.718	0.681	0.684	0.690	0.693
21.45	0.733	0.736	0.672	0.678	0.705

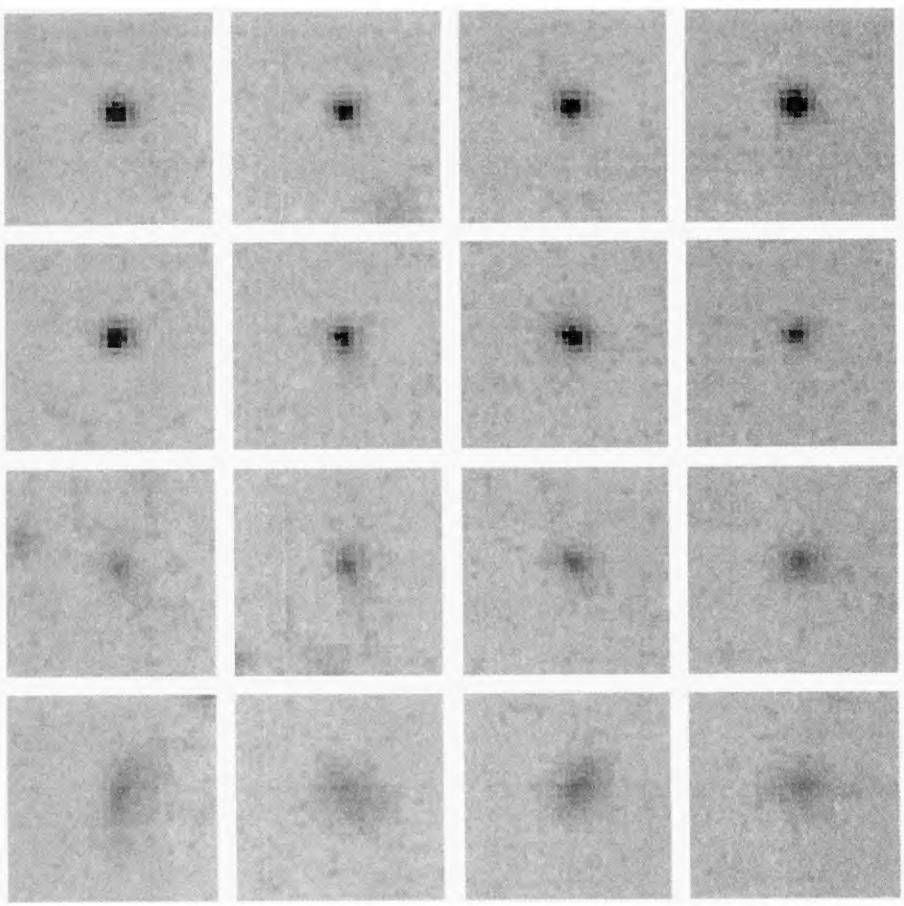

FIG. 8. Each row consists of representative objects from one of the four classes discovered in the DPOSS data by Autoclass. It appears one can relate each type to physically, not just statistically, distinct classes of objects.

Our own initial experiments in applying Autoclass to DPOSS appear to confirm the validity and usefulness of this approach. After supplying Autoclass with the eight-dimensional feature vectors from a sample of several hundred objects from our four fields, it analyzed the distribution of the objects in this parameter space and suggested four distinct classes within the data. Representative objects from these four classes are presented in Fig. 8. Visually, the classes seem to divide into stellar objects, stellar-like objects with a low surface brightness halo, and diffuse or irregular objects with and without a central core. Its success at distinguishing these apparently physically relevant classed based just upon eight image parameters suggests that far richer and innovative results may be in store when one matches multiple catalogs together, increasing the informational dimensionality of the dataset manifold.

6. CONCLUDING REMARKS

Through the careful selection and construction of object attributes, and the application of machine learning to derive sets of rules based upon them, we have been able to achieve high rates of classification accuracy at levels up to a magnitude fainter than in previous Schmidt surveys. By examining a set of four fields in two colors, we have verified that galaxy catalogs produced from DPOSS using this technique appear to be consistently complete and contaminated across multiple plates. In fact, in testing our classifiers on completely independent plate data, we found them to produce 90% complete galaxy catalogs down to an equivalent B_J magnitude of approximately 21.0^m. There is no *a priori* reason why, without any further work, these very same classifiers should not result in exactly the same accuracy rates for all future high Galactic latitude DPOSS plates. However, we note that by accumulating more and better overlapping CCD and plate data, one may be able to train classifiers that are able to generalize even better.

A significant additional benefit of the classification approach we describe is that it easily generalizes to the construction of any number of object classifiers for any purpose in the future. Provided the astronomer is able to construct a suitably large enough sample of objects for both testing and training, the same technology may be applied for a wide variety of scientific purposes. To facilitate the construction of such sets, we have implemented a tool within SKICAT that allows the user to display individual objects from a DPOSS plate scan and assign a classification to each. One may also, as we have done, use the extensive object matching technology within SKICAT to retrieve attributes from one set of catalogs (e.g., plates) and classifications from their matched counterparts in others (e.g., CCDs). It is our hope that with the availability of tools such as SKICAT and Autoclass, and

the demonstrated scientific value they add, such advanced data analytic techniques may see more widespread use in the future.

This work was supported at Caltech in part by NASA AISRP Contract No. NAS5-31348, the Caltech President's fund, and NSF PYI award AST-9157412, and at JPL under a contract with NASA. The POSS-II is partially funded by grants to Caltech from the Eastman Kodak Co., the National Geographic Society, the Samuel Oschin Foundation, NSF Grants Nos. AST 84-08225 and AST 87-19465, and NASA Grants Nos. NGL 05002140 and NAGW 1710. We acknowledge the efforts of the POSS-II team at Palomar, the scanning team at STScI, and the SKICAT team at JPL, most especially Joe Roden.

REFERENCES

Breiman, L., Freidman, J., Olshen, R., & Stone, C. 1984, Classification and Regression Trees (Wadsworth & Brooks, Monterey, CA)
Cheeseman, P., *et al.* 1988, in Proc. Fifth Machine Learning Workshop, San Mateo (Morgan Kaufmann), p. 54
Djorgovski, S., Lasker, B., Weir, N., Postman, M., Reid, I., & Laidler, V. 1992, BAAS, 24, 750
Ellis, R. 1987, in Observational Cosmology, IAU Symposium No. 124, edited by A. Hewitt, G. Burbidge, and L. Z. Fang (Reidel, Dordrecht), p. 367
Fayyad, U. 1991. Ph.D. thesis, EECS Department, The University of Michigan
Fayyad, U. 1994, in Proc. of the Twelfth National Conference on Artificial Intelligence AAAI-94 (MIT Press, Cambridge, MA), pp. 6601–606
Fayyad, U., Doyle, R., Weir, N., & Djorgovski, S. 1992, in Proceedings of the ML-92 Workshop on Machine Discovery (MD-92), Aberdeen, Scotland, edited by J. Zytkow (Morgan Kaufmann), p. 117
Fayyad, U., & Irani, K. 1990, in Proceedings of the Eighth National Conference on Artificial Intelligence AAAI-90, Boston, MA
Fayyad, U., & Irani, K. 1992, in Proceedings of the Tenth National Conference on Artificial Intelligence, AAAI-92, San Jose, CA
Fayyad, U., & Irani, K. 1993, in Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93), Chambery, France (Morgan Kauffman) (in press)
Finney, D., Latscha, R., Bennett, B., & Hsu, P. 1963, Tables for Testing Significance in a 2×2 Contingency Table (Cambridge University Press, Cambridge)
Hertz, J., Krogh, A., & Palmer, R. 1991, Introduction to the Theory of Neural Computation (Addison-Wesley, Redwood City, CA)
Heydon-Dumbleton, N. H., Collins, C. A., & MacGillivray, H. T. 1989, MNRAS, 238, 379
Lasker, B., Djorgovski, S., Postman, M., Laidler, V., Weir, N., Reid, I., & Sturch, C. 1992, BAAS, 24, 741
Maddox, S., Sutherland, W., Efstathiou, G., & Loveday, J. 1990, MNRAS, 243, 692
Odewahn, S., Stockwell, E., Pennington, R., Humphreys, R., & Zumach, W. 1992, AJ, 103, 318
Picard, A. 1991. Ph.D. thesis, California Institute of Technology
Quinlan, J. 1986, in Machine Learning, Vol. 1, No. 1
Quinlan, J. 1990, in Machine Learning: An Artificial Intelligence Approach Vol. III, San Mateo, CA, edited by Y. Kodratoff and R. Michalski (Morgan Kaufmann)
Reid, I. *et al.* 1991, PASP, 331, 465
Reid, N. & Djorgovski, S. 1993, in Sky Surveys: Protostars to Protogalaxies, edited by B. T. Soifer, ASP, Conf. Ser. 43, 125
Sebok, W. 1979, AJ, 84, 1526
Smyth, P., Fayyad, U. M., & Gray, A. 1995, in Twelfth International Conf. on Machine Learning, San Francisco (Morgan Kaufmann) (in press)
Valdes, F. 1982, SPIE Proc. on Instrumentation in Astronomy IV, 331, 465
Weir, N., Djorgovski, S., & Fayyad, U. 1995, AJ (in press)
Weir, N., Djorgovski, S., Fayyad, U., Smith, J., & Roden, J. 1994, in Astronomy From Wide-Field Imaging, IAU Symposium No. 161, edited by H. MacGillivray *et al.* (Kluwer, Dordrecht), p. 205
Weir, N., Fayyad, U., Djorgovski, S., & Roden, J. 1995, PASP (submitted)
Weir, N., & Picard, A. 1991, in Digitised Optical Sky Surveys, edited by H. T. MacGillivray and E. B. Thompson (Kluwer, Dordrecht), p. 225