The Astronomical Journal, 128:3092–3107, 2004 December
© 2004. The American Astronomical Society. All rights reserved. Printed in U.S.A.

THE DIGITIZED SECOND PALOMAR OBSERVATORY SKY SURVEY (DPOSS). III. STAR-GALAXY SEPARATION

S. C. Odewahn,[1] R. R. de Carvalho,[2] R. R. Gal,[3] S. G. Djorgovski, R. Brunner,[4]
A. Mahabal, P. A. A. Lopes,[2] J. L. Kohl Moreira,[5] and B. Stalder[6]
Palomar Observatory, Mail Code 105-24, California Institute of Technology, Pasadena, CA 91125
Received 2003 May 12; accepted 2004 August 26

ABSTRACT

We discuss object classification for a multicolor survey of high-latitude fields from the Digitized Second Palomar Observatory Sky Survey (DPOSS) and the resulting Palomar-Norris Sky Catalog. Two methods are used to perform automated image classification for star-galaxy separation in DPOSS. As a source of classifier training/testing data, we employ an unprecedented 500 field collection of CCD photometry in the Thuan & Gunn *gri* system obtained with the Palomar 60 inch (1.5 m) telescope. We have trained artificial neural network (ANN) and decision tree (DT) image classifiers using images of ≈4000 galaxies and ≈3000 stars classified with FOCAS on 52 deep CCD images. We assess the systematic errors in our classifiers as a function of apparent magnitude. To model the loss of galaxies through misclassification and the contamination of our galaxy samples by misclassified stars, we compare the DPOSS ANN+DT image classifications with image data from 46 CCD fields on 21 POSS-II fields *not used in the initial training/testing process*. We assess these same functions in a more stringent manner by comparing classifications of DPOSS images common with different fields via the plate overlaps. These tests are combined to derive analytic descriptions of sample incompleteness and contamination for future use in our assessment of multicolor galaxy number counts and the two-point angular correlation function. Finally, we derive star and galaxy number counts from 341 DPOSS fields covering a total of 7756 deg^2 in both the north and south Galactic hemispheres. These data are used to establish a final correction for stellar contamination in our galaxy samples and to demonstrate the level of classification homogeneity in the DPOSS *g* and *r* catalogs drawn from a wide range of Galactic latitudes.

Key words: catalogs — methods: data analysis — surveys

1. INTRODUCTION

The Digitized Palomar Observatory Sky Survey (DPOSS) is a digital version of a three-band photographic survey of the northern sky (POSS-II; Reid et al. 1991), which has been released to the astronomical community (Djorgovski et al. 2003). The principal product of DPOSS is the Palomar-Norris Sky Catalog (PNSC), which will contain approximately 50 million galaxies and over 2 billion stars, down to the equivalent blue limiting magnitude of 22. An extensive discussion of the photometric calibration of the survey is given in Gal et al. (2004). An important addition is a program to collect extensive CCD calibration data in the *gri* system of Thuan & Gunn (1976) with the Palomar 60 inch (1.5 m) telescope (P60). This represents a considerable advance over previous optical sky surveys based on large-format photographic plates.

A great variety of scientific projects is possible using this major data set, including studies of galaxy clustering and of the large-scale structure of the low-redshift universe (Gal et al. 2003) as tests of structure formation models; galaxy counts at moderate depth as a function of magnitude, color, and morphology to be used in tests of galaxy evolution over the past few gigayears; searches for high-redshift ($z > 4$) quasars (Kennefick et al. 1995), unusual quasars (Brunner et al. 2003), and active galactic nuclei at lower redshifts; mass optical identifications of radio, IR, and X-ray sources; generation of objectively defined catalogs of clusters and groups of galaxies; a search for previously unknown sparse globular clusters and dwarf spheroidal galaxies; the cataloging of low surface brightness galaxies; a search for objects with peculiar colors or variability; galactic structure studies; and many others. The success of many of these scientific programs rests heavily on our ability to draw reliably pure star or galaxy samples. In this paper we describe methods used to perform automatic star-galaxy separation using supervised learning techniques via artificial neural networks (ANNs) and decision trees (DTs). These powerful techniques map a well-defined multivariate image-parameter set to an output set of image classes.

A preliminary classification scheme using DTs for DPOSS was described by Weir et al. (1995a). These results were based on four pairs of *J* and *F* POSS-II plates and a relatively small number of calibrating CCD fields. This paper is based on an analysis of 352 *J* and *F* POSS-II plate pairs and more than 500 calibrating CCD fields. A variety of new image segmentation and pattern classifier issues are discussed. Hence, this paper represents the improved and more comprehensive classification methodology used for the final catalog.

[1] Current address: McDonald Observatory, University of Texas at Austin, Austin, TX 78712.
[2] Current address: Instituto Nacional de Pesquisas Espaciais, Avenida dos Astronautas 1758, São José dos Campos, 12227-010 São Paulo, SP, Brazil.
[3] Current address: Department of Physics, University of California, Davis, 1 Shields Avenue, Davis, CA 95616.
[4] Current address: Department of Astronomy, University of Illinois, Urbana, IL 61801.
[5] Current address: Observatório Nacional, Rua General José Cristino 77, 20921-400 Rio de Janeiro, RJ, Brazil.
[6] Current address: Institute for Astronomy, University of Hawaii, Honolulu, HI 96822.

2. INPUT DATA AND CATALOG CONSTRUCTION

Detailed descriptions of the plate scanning and initial DPOSS catalog construction are given in Weir et al. (1995a, 1995b, 1995c), but we briefly review the important points. The Sky Image Catalog and Analysis Tool (SKICAT) was designed to process large digital scans of the POSS-II photographic plates taken with *J*, *F*, and *N* emulsions, which are subsequently calibrated to the *gri* system. The initial processing of these scans involves the determination of the plate density–to–intensity transformation via a set of densitometry spots exposed on each POSS-II plate, determination of the plate density saturation level, and determination of the mean sky level. Next, a robust background-mapping algorithm is applied to filter out discrete image sources and produce a map of the mean level and standard deviation of the sky background across the plate.

For image detection and measurement, SKICAT uses the Faint Object Classification and Analysis System (FOCAS; Valdes 1982). The entire plate is processed in a grid of 13×13 partially overlapping subregions, which we refer to as footprints. The sky level and noise maps for each footprint are used to set an image detection threshold that is some fraction of the local background noise. FOCAS detects objects in the form of groups of connected image pixels whose values lie above the local threshold level. Next, a series of rough classification rules are applied to the FOCAS catalog of each footprint to establish a set of stars for that region, and these sources are used automatically by FOCAS to model the two-dimensional local stellar point-spread function (PSF) for the footprint. Having established the mean PSF, a FOCAS analysis stage is run in which each detected object is scaled in the intensity and spatial domains to best match this PSF model. The two parameters describing this fit are referred to as `scale` and `frac`. A FOCAS image classification is assigned on the basis of `scale` and `frac` alone. Finally, FOCAS determines which detected images are actually a merged set of overlapping detections, and an algorithm is applied to deblend these cases and determine the positions, image classes, and photometric parameters of each constituent. The final set of homogenized image parameters is finally organized by plate footprint into a Sybase database catalog that may be queried for purposes of subsequent analysis.

As described in Weir et al. (1995b), the initial set of FOCAS runs are applied to spatially overlapping plate regions. Following FOCAS treatment of all these regions, SKICAT conducts an automated overlap comparison to bring the photometric properties of each footprint onto a common mean system. An essential part of the SKICAT system is an automated, objective classification of detected sources as ''stars'' (i.e., PSF-like) or ''galaxies'' (i.e., more resolved than the PSF). Supervised classification algorithms using superior training data sets derived from the CCD calibration images are implemented to push the depth of the accurate classifications considerably deeper than would be possible with the plate data alone. We have improved on and expanded the original SKICAT source classification methodology of Weir et al. (1995b) as described below.

In addition to `scale` and `frac`, the parameters we are concerned with here for refining star-galaxy separation are total magnitude (`MTot`); isophotal image area (`Area`), which is the number of pixels constituting each object; core magnitude (`MCore`), which is the integrated magnitude of the central 3×3 set of pixels in each image; and the second moment of the intensity-weighted radii of all pixels composing each image (`IR2`). As we discuss below, all of these photometric parameters are used in some magnitude range to perform star-galaxy classification.

The classical FOCAS approach using `scale` and `frac` rules is superior in the regime of faint sources, $g > 20.5$, for which we can no longer reliably map the stellar locus (see § 4.1). Similarly, for bright, compact sources (i.e., stars with $g \leq 17$) the core saturation in DPOSS images is substantial, and neither PSF fitting nor locus mapping is successful. For most objects ($17 \leq g \leq 20.5$), parameter spaces may be formed using photometric parameters such as `Area`, `MTot`, `MCore`, and `IR2`, which provide substantial segregation between stellar and extended sources (i.e., stars and galaxies), and this approach is used in SKICAT to provide bright object image classes via a set of parameter-based classification rules. We discuss the use of two classifier induction methods that use these parameters to provide improved object classification.

We stress that our object classifications are based on image morphology alone, i.e., on how well a given source is represented by a PSF appropriate for that magnitude and that plate. In principle, one could also use colors as an indicator that a source is a normal star, since the stellar temperature sequence forms a well-defined locus in the color-color parameter space. However, that would result in misclassification of unresolved objects (e.g., quasars) with colors different from those of normal stars. Moreover, with only three bands used in DPOSS and with the calibration of one (the *N*-plate material) being highly problematic, this color-based classification approach is not suitable. Thus, in this context, by ''star'' we mean simply an unresolved source in our images.

3. CLASSIFIER INDUCTION

In the last two decades several different methods have been suggested for automatic object classification (e.g., Sebok 1979; Heydon-Dumbleton et al. 1989; Odewahn et al. 1992; Weir et al. 1995b). Modern classifiers use the nonparametric approach in which a set of classification rules or weight networks is constructed on the basis of examples of accurate classification (from deep CCD imaging, for example). Currently popular techniques include decision trees and artificial neural networks. The main advantage of such approaches is that the rules are created in an objective and reproducible way. Detailed descriptions of DTs and related methods are presented in Fayyad (1991) and Fayyad & Irani (1990). Another approach is to use ANNs to map sets of input image-parameter vectors to output image classes. This method, applied to star-galaxy separation by Odewahn et al. (1992), has the advantage that quite complicated, nonlinear mappings are able to solve complicated problems using even overlapping parameter spaces. This method lends itself well to work in the faint regime of photographic analysis, for which such issues as nonlinear effects and low signal-to-noise ratio (S/N) are profound.

3.1. *Decision Trees and the RULERS System*

The essence of the machine-learning (ML) approach is to extract rules from a given data set, called attributes, that can be applied to similar data. One of these methods is a DT (Quinlan 1986). For a given set of attributes this method creates nodes that represent tests applied to the attribute values. The outgoing branches of a node correspond to the possible outcomes of the test at the node. For a detailed discussion about DT methodology and its application to a restricted set of the DPOSS data, we refer the reader to Weir et al. (1995b).

There are several algorithms for generating DTs (Quinlan 1986, 1990). A basic problem with DT generation algorithms is that even in a good tree some leaves are overspecialized or predict the wrong class. If there are measurement errors in the attributes, the DT will tend to fit to the noise and hence not generalize well to data that are out of sample. In DTs the data is partitioned quickly into smaller subsets, which accelerates the learning process while simultaneously creating overspecialization of the leaves in the tree. The RULER system, developed for SKICAT, was designed to extract only the best rules from a set of multiple trees trained on the same data set. The implementation of such an algorithm is described in Fayyad & Irani (1990).

3.2. *Artificial Neural Networks*

It is clearly desirable to combine as much quantitative information as possible when assigning a classification. An ANN, discussed by Rumelhardt & McClelland (1988), provides a means of mapping a large amount of input information defining a complicated nonlinear space to derive a best-guess classification. Neural networks are systems of weight vectors, whose component values are established through various ML algorithms that take as input a linear set of pattern inputs and produce as output a numerical pattern encoding a classification. This technique has been applied with great success to the problem of star-galaxy separation by Odewahn et al. (1992, 1993). In these studies, neural network classifiers were trained using the back-propagation algorithm. The networks consisted of an input node layer of image parameters, two hidden layers, and an output layer consisting of two nodes (a star node and a nonstar node) for encoding the classifications. Systematic experiments were conducted to determine the optimum network architecture, weight-update parameters, and input image parameters (Odewahn 1995). Statistical comparisons of the ANN-classified photographic image surveys with more reliable (but less extensive) CCD surveys have proved this classification method to be very robust.

4. CLASSIFICATION ATTRIBUTES

The input patterns fed to the ANN and DT classifiers discussed above consist, for DPOSS detections, of a set of FOCAS image parameters that have been shown to segregate the compact (star) and extended (galaxy) populations. In practice, we combine some of our parameters to produce single star-galaxy separation parameters. In one such case we form the difference between the `MTot` and `MCore` estimates to create a crude concentration index, i.e., a single number that indicates the amount of flux contained in the central area of the image relative to the total area (the extrapolated isophotal area). An example of this very effective parameter space is shown in Figure 1. The tight locus of points running in a roughly horizontal pattern across the plot is composed of stars. Compared with stars, extended objects will contribute a smaller percentage of flux in the central 3×3 pixel area relative to the total area and produce a more negative value for `MTot` $-$ `MCore`. Hence, the diffuse cloud of points extending below the stellar locus in Figure 1 at all magnitudes is attributable to galaxies. In a similar fashion, we have combined the FOCAS `scale` and `frac` parameters in a series of simple rules to form a single parameter that is linearly correlated with image compactness. This parameter is referred to as the composite stellar function (`csf`) and provides a simple way of using the `scale` and `frac` information for pattern classification. DTs are well adapted for discrete, irregular data spaces (like that produced in the `scale` $-$ `frac` domain), while ANN classifiers are best used with continuous, smoothly varying input parameters. Thus, we have opted to use the `csf` parameter as input for both the ANN and DT classifiers developed in DPOSS. These basic image parameters are summarized in Table 1. The first five parameters in this table are taken directly from the original FOCAS measurements for each plate. The three normalized parameters (names beginning with *n*) are the FOCAS parameters normalized using the stellar locus on each plate, as described below. Finally, we list the image classifications and their associated probabilities.

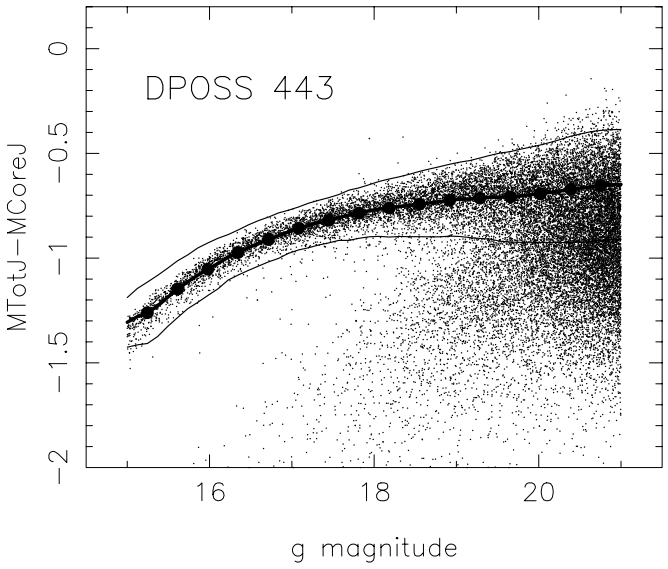

FIG. 1.—The `MTotJ`−`MCoreJ` vs. `MTotJ` parameter space from a single DPOSS *J* plate. The strong stellar locus, which is partially mapped by the filled circles and the heavy curve, has a very high contrast at bright magnitudes relative to the diffuse cloud of galaxy points in the lower portion of this parameter space. The upper and lower 2 σ curves are represented by the light lines. The majority of sources within these light lines are unresolved (i.e., stars). Primarily because of seeing and photographic response variations, the location and shape of this stellar locus can change significantly from plate to plate.

TABLE 1
DPOSS CLASSIFICATION ATTRIBUTES FOR *J* PLATES

Attribute	Description
`MTotJ`	Total FOCAS magnitude
`MTotJ` − `MCoreJ`	Concentration index from FOCAS core magnitude
`IR2J`	Intensity-weighted second spatial moment
`AreaJ`	Number of pixels above isophotal threshold
`csfJ`	Linear composite of FOCAS `scale`, `frac` estimates
`n_MTotJ` − `MCoreJ`	Stellar locus–corrected `MTotJ` − `MCoreJ`
`n_IR2J`	Stellar locus–corrected `IR2J`
`n_AreaJ`	Stellar locus–corrected `AreaJ`
`Class_ann_J`	Class membership from ANN classifier
`prob_ann_J`	Probability of ANN class estimate
`Class_dtree_J`	Class membership from DT classifier
`prob_dtree_J`	Probability of DT class estimate
`Class_FOCAS_J`	Class membership from original FOCAS classifier
`prob_FOCAS_J`	Probability of original FOCAS class estimate
`Class_BEST_J`	Class membership from final classifier

NOTE.—The classification-related image attributes in a DPOSS record are shown here for a *J*-plate detection. Each of these attributes is also tabulated for *F*- and *N*-plate sources.

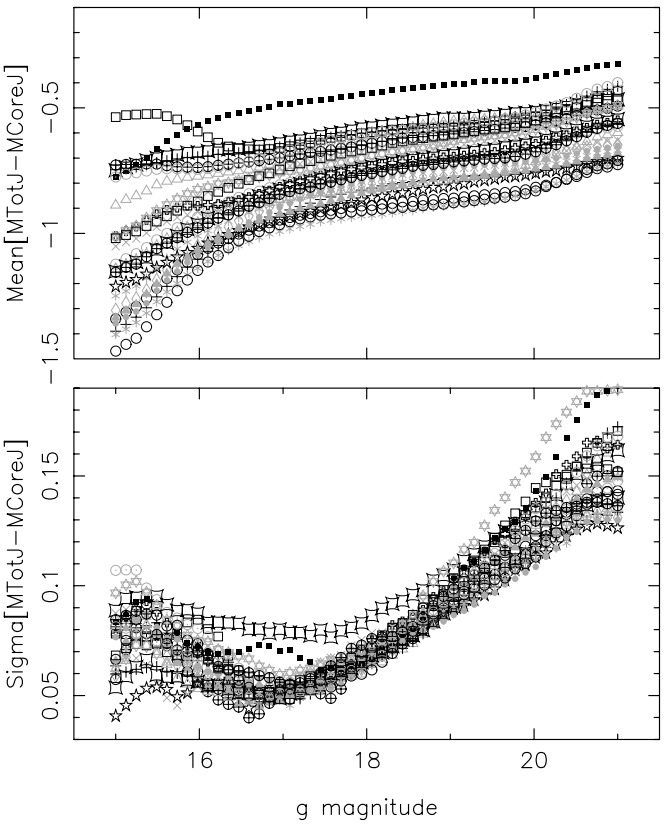

FIG. 2.—Mean stellar loci from DPOSS star-galaxy separation parameter measured on 25 different DPOSS *J* plates. Each symbol type represents a different *J* plate. The location of the magnitude-binned locus mean (*top*) can vary significantly over different plates. The width of the stellar locus (*bottom*) as measured by the magnitude-binned standard deviation is rather homogeneous over the POSS-II plate material. Magnitude units are used as the *y*-axis unit in each panel. Hence, plate-to-plate variation in the width of the stellar locus in this classification parameter is very small compared with the mean locus displacement.

4.1. *Plate-dependent Image-Parameter Variations*

As discussed in Odewahn et al. (1993), the dominant source of error in the classification of images in Schmidt plate surveys derives from large systematic differences in the plate densitometry. These differences produce substantial variations in the shape and placement of the stellar loci for key classification parameter spaces. We have studied variations in stellar locus placement for all DPOSS image parameters useful for star-galaxy separation (`Mtot` − `Mcore`, `IR2`, `csf`, and `Area`) for a large number of DPOSS fields in order to assess this problem. In Figure 2, we summarize one such study by showing the mean stellar loci in `Mtot` − `Mcore` for 25 different DPOSS *J* plates. We map the stellar locus in each parameter space using a robust median filtering technique in overlapping `Mtot` bins (the *x*-axis of each parameter space). At faint levels (roughly $g \geq 19.5$ mag), all differences between extended and compact images disappear, so to increase the contrast of the faint-end stellar locus we have used a `csf` filter to isolate FOCAS-classified stars. In this way, we are able to better describe the properties of the faint stellar loci in the `Mtot` − `Mcore`, `IR2`, and `Area` parameter spaces. Having isolated each stellar locus in this manner, we can map the location and width of the locus as a function of magnitude, `Mtot`. In the top panel of Figure 2 we show the mean loci, and in the bottom panel we plot the locus widths as a function of apparent magnitude for 25 different DPOSS plates. Locus width changes are relatively small, but the mean placement of the locus is highly plate dependent. For the development of a universal pattern classifier based on these image parameters, such plate-dependent variations must be removed prior to classification.

One empirical way to remove these variations is to map the stellar locus in each parameter space and then subtract it, as a function of magnitude. For the DPOSS reductions, this procedure is performed for each footprint, using a robust locus mapper for each classification parameter. An example of such a mean locus fit is plotted as a solid curve in Figure 1 for the `Mtot` − `Mcore` parameter space. The number of particularly bright sources (`MTot` < 18 mag) found in a single footprint is usually far too low to provide a useful locus map. The solution we have adopted as a workaround is the creation of a set of plate-wide mean loci for each parameter. These loci are computed using many footprints covering the central portion of each field. Moreover, a graphical software package is used to guide the (human) DPOSS reducer in mapping and verifying each plate-wide mean locus. In addition to providing a more reliable fit to each stellar locus, this step insures that the reducer must inspect each important parameter space in each DPOSS plate catalog. This quality control step has been extremely useful in identifying catalogs with previously undetected processing problems, while maintaining the objectivity and automation of the rest of the classification process. Having subtracted the stellar locus from each parameter space on each plate, we are able to combine data from many different fields to formulate a training and testing data set for the ANN and DT pattern classifiers. An example of such a data set is summarized in Figure 3, using locus-subtracted `Mtot` − `Mcore` parameters from 31 different DPOSS *F* plates spread across many declinations and observing seasons. A stellar locus subtraction is applied to the `Mtot` − `Mcore`, `Area`, and `IR2` image parameters from the *J*-, *F*-, and *N*-plate catalogs. To denote locus-corrected image parameters, we add the "*n*" prefix to each parameter name (see Table 1). We note that no formal normalization is required for the `csf` parameter, since the `scale` and `frac` values from FOCAS are estimated relative to the mean stellar PSF adopted for each (i.e., `csf` = `n_csf`).

4.2. *A Source of High-Confidence Image Classifications*

As discussed in Gal et al. (2000, 2004), we conducted a long-term CCD observing program with the Palomar 60 inch telescope to collect calibration and image-classifier material for DPOSS. To compile a classifier training set, we used 79 CCD fields from this survey that are located on the 31 DPOSS fields to assign image classes to the objects measured on the DPOSS plates. This provided a set of 14,370 training patterns (detected sources with image parameters and classes). This constituted our initial training sample. For a totally independent testing sample, we collected a set of 14,304 patterns from 46 CCD fields located on a completely different set of 21 DPOSS fields. This constituted our final classification quality assessment sample. This quality assessment set was drawn from half the number of CCD fields as was used in the training set sample, yet it contains nearly the same number of members as the training set. This is due to the fact, as explained below, that we used slightly more stringent seeing and image class agreement criteria for the CCD catalogs when selecting the classifier training data. Hence, only the highest-quality image-parameter sets were assembled for classifier development. These restrictions were relaxed when assembling the data set for a final quality check, both to increase the number of objects used and to more faithfully reflect the properties of the average DPOSS

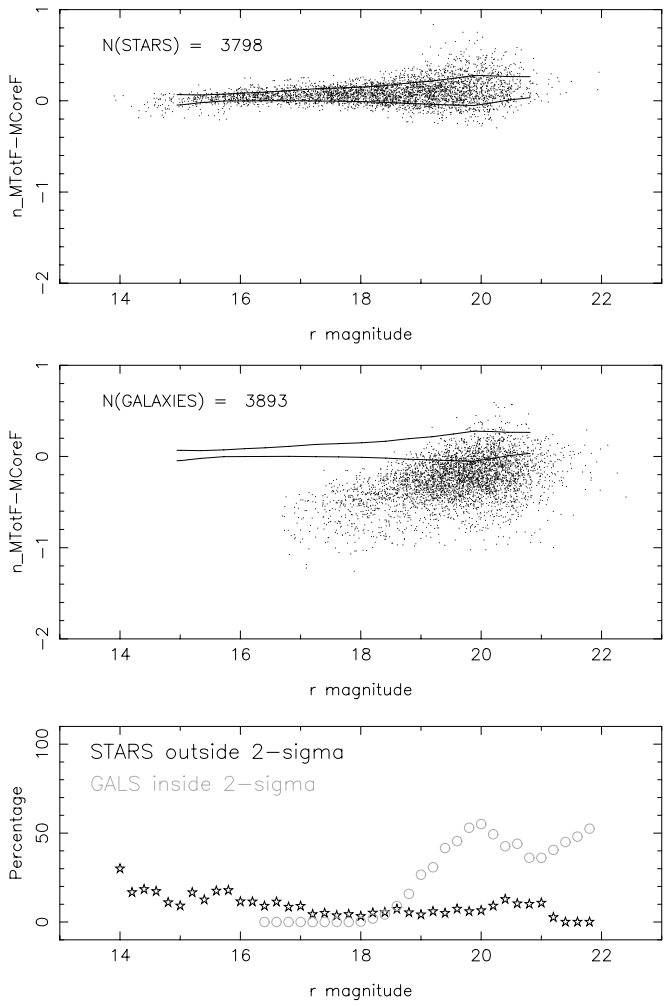

FIG. 3.—The MTotF − MCoreF DPOSS image-parameter space that is used as part of the inputs to pattern classifiers used for star-galaxy separation. As discussed in the text, the stellar locus has been mapped and subtracted for each DPOSS plate. Using 31 different DPOSS plates, with no single plate contributing more than 8% of the total sample, ensures that we are able to train very general classifiers that can be extended more reliably to the entire survey.

plate and CCD material. The median FWHM for the CCD images used forming the testing sets is 1.″4, with quartiles $Q_{25} = 1.''2$ and $Q_{75} = 1.''7$.

All CCD fields have quantitative mean seeing estimates for each of the three bandpasses (g, r, and i), and we are able to grade the CCD fields by seeing quality. Seeing is measured on each image by computing the median FWHM of FOCAS-classified stars in a magnitude range of approximately $17.5 \leq g \leq 19.75$. This range was chosen to produce reliable star/galaxy separation and images with an S/N high enough to allow accurate estimates of the FWHM. For the image classifications used in Figure 3, as well as for all of our training/testing data sets, we have selected exclusively objects from CCD fields that have superior seeing (FWHM $\leq 1.''5$ measured in r).

These seeing estimates are confirmed by comparing the agreement between FOCAS estimated image classes in g, r, and i as a function of apparent magnitude. To measure classification quality, we fit a polynomial relation to the percentage agreement between comparisons of the (g, r), (g, i), and (r, i) CCD FOCAS class pairs as a function of magnitude. From these relations, we estimate the r magnitude at which the level of agreement falls below 90%. From a study of 100 P60 CCD fields we found these class quality measures to be well correlated with our quantitative mean seeing estimates. Even more reassuring that we are estimating image class quality in a robust fashion is that this analysis shows that the redder CCD bandpasses (r and i) produce systematically better classification agreement to fainter magnitudes and have smaller seeing estimates, as would be expected for a ground-based survey using a CCD detector with a peak quantum efficiency around 7000 Å.

Only CCD catalogs having high classification agreement ratings and superior seeing measures have been used to establish training and testing samples for the DPOSS image classifiers. In establishing the final image-parameter training sets, we have applied one more constraint on the assignment of an object class to any source used for training data: only objects whose CCD classes agree in the g, r, and i bandpasses are allowed into the classifier data set. This seemingly heavy constraint was imposed to avoid one final problem with image misclassification in crowded regions. In heavily crowded parts of even the best-seeing CCD images, FOCAS may be unable to assign a reliable image class. Under the assumption that image merging will be handled differently in different bandpasses, we reasoned that severe crowding problems would result in variations in class estimates for the different bands. By taking only those sources with class agreement in g, r, and i, we avoid the bulk of such crowding-induced class mistakes. Following the application of this final quality constraint, we were left with approximately 7600 training patterns.

We should stress that the training and testing image-parameter sets collected using this CCD database have been assembled in a rigorous, automated, and well-defined way. These image parameters should be considered a highly reliable "truth set" for the development of star-galaxy classifiers. This data set is summarized for our F-plate training/testing material in Figure 4, where we plotted locus-subtracted image parameters as a function of apparent magnitude for stars and galaxies for all of the relevant classification parameters.

5. CLASSIFICATION RESULTS WITH TRAINING AND TEST DATA

Using the training data set described in the previous section we have trained ANN and DT star-galaxy classifiers in each of the three DPOSS bandpasses. The details of training supervised classifiers of these types are discussed in Odewahn (1997) for ANNs and Weir et al. (1995b) for DTs. Briefly, we use some fraction of the data sample as a training set. For both classifiers, an iterative training algorithm is used to establish the final classifier. In the case of an ANN, we use back-propagation training to adjust the weight values in each node of the network to minimize the mean error function, which in this case is the difference between the network output pattern and the target pattern (determined by whether the object is a star or a galaxy). As discussed in § 3.1, the DT code establishes a series of decision rules that map the input parameter sets to class memberships. These rules are gradually pruned to optimize the percentage of successful class outcomes.

For training both the ANN and DT classifiers for DPOSS, we used the following image parameters as classifier inputs for each bandpass: Mtot, n_MTot − Mcore, n_IR2, n_Area, and n_csf. The n_ prefix indicates that a stellar locus subtraction has been applied to the image parameters on a plate-by-plate basis. We chose to divide the data set into training and testing samples by simply dividing the available data into

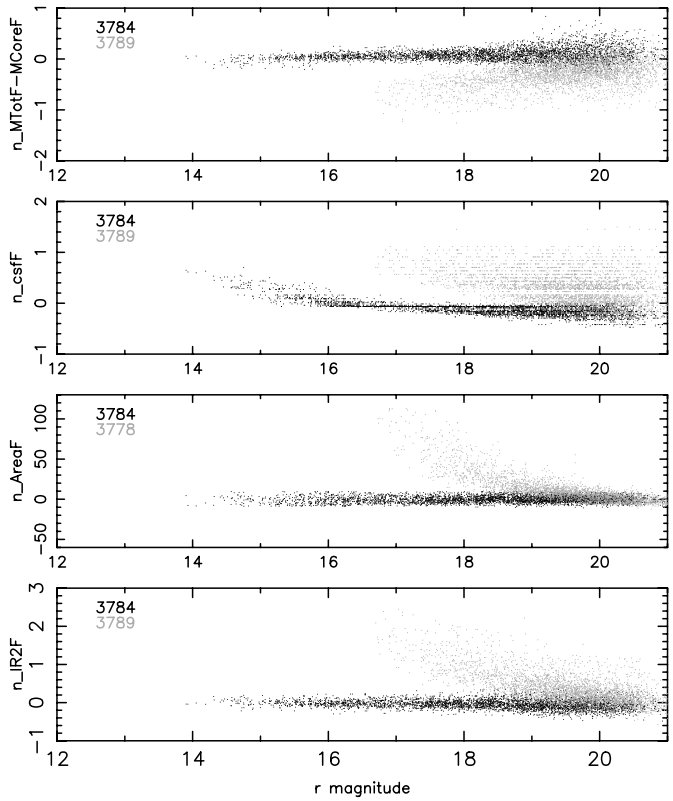

FIG. 4.—Image parameter spaces used as inputs for the ANN and DT pattern classifiers used for DPOSS star-galaxy separators. These data were collected from 31 different DPOSS *F* plates using 79 different CCD fields observed on nights of superior seeing (FWHM $\leq 1''.5$). Galaxies are plotted as light-gray points, and stars as black points. The numbers of star/galaxy points collected for each parameter are shown in the top left corner of each panel.

stars and galaxies in a series of magnitude bins. Each set was then combined into two interleaved samples of stars and galaxies: one for training and one for testing. Some classifier builders support the notion that the pattern classifier should be trained using training samples that have the same class frequency distribution as that expected for when the classifier is used in classification mode. We view this as a risky prospect at best, especially for our own astronomical problem. The ratio of star to galaxy numbers will change systematically with Galactic latitude, and this is in fact one of the very things we will measure in the completed DPOSS survey. By splitting the samples into roughly equal numbers of stars and galaxies in 0.5 mag wide bins spanning our range of interest, we avoid any predisposition to class assignment on the basis of class frequency in the training sample.

At each step in the iterative training process for each classifier, we present the classifiers with the testing patterns to obtain an independent assessment of classifier performance. Although our data sets are quite large, we are able to achieve an apparently high success rate for the training sample alone by overtraining. We used classification success-rate statistics computed with the test data alone to guard against this. No changes to the ANN weights or the DT rules were applied on the basis of these test sample statistics. For each bandpass, we summarize in Table 2 the sample sizes, the magnitude range covered, and the overall success rates achieved for each classifier, using both the training and test samples. A more meaningful assessment of classifier performance is shown in Figure 5, in which we plot the percentage of successful classifications as function of apparent magnitude as measured by the *g*, *r*, or *i* magnitude from the CCD set. The relation used to compute the rate of successful star classification in each magnitude bin was

$$S_{\rm star} = \frac{N([{\rm CCD} = {\rm star}] \bigcap [{\rm Plate} = {\rm star}])}{N({\rm CCD} = {\rm star})}, \qquad (1)$$

and the relation used for galaxy classification was

$$S_{\rm gal} = \frac{N([{\rm CCD} = {\rm gal}] \bigcap [{\rm Plate} = {\rm gal}])}{N({\rm CCD} = {\rm gal})}. \qquad (2)$$

We note that in these two previous equations it is assumed that the classifications provided by the CCD data are completely accurate, which is only applicable as a first approximation in the magnitude range within which we believe the current plate data can be meaningfully used ($15.0 < r < 19.5$). Even in this range, the significantly poorer seeing and larger internal variance of the plate data heavily outweigh errors in CCD classification. At fainter magnitudes, the CCD classifier also becomes unreliable, and we do not recommend using our classifications at $r > 19.5$.

5.1. *Image Classification Probability Estimates*

The ANN and DT image classifiers each produce, in addition to a "best-guess" class estimate, a probability estimate that can be used to measure classification confidence. The node values in the output vector of an ANN are in some sense the Bayesian "a posteriori" probability that the input vector is drawn from the class associated with that node (Richard & Lippmann 1991). The use of this probability as a measure of classification confidence in the star-galaxy separation problem is demonstrated in Odewahn (1995). Similarly, the DT classifier evaluates the probability that an output class assignment (averaged over all class conditions in the rules set) is due to chance. For both ANN and DT classifiers, we have

TABLE 2
DPOSS CLASSIFIER TRAINING STATISTICS

Class	Band	*N*stars	*N*gals	*S*(TRAIN)	*S*(TEST)	Mag. Range
ANN	*g*	1182	1182	92.3	92.1	$16.0 \leq g \leq 21.0$
DT	*g*	1182	1182	91.0	90.4	$16.0 \leq g \leq 21.0$
ANN	*r*	1188	1187	94.5	93.7	$16.0 \leq r \leq 20.5$
DT	*r*	1188	1187	94.9	92.9	$16.0 \leq r \leq 20.5$
ANN	*i*	918	918	92.5	93.4	$15.0 \leq r \leq 20.0$
DT	*i*	918	918	94.0	91.4	$15.0 \leq r \leq 20.0$

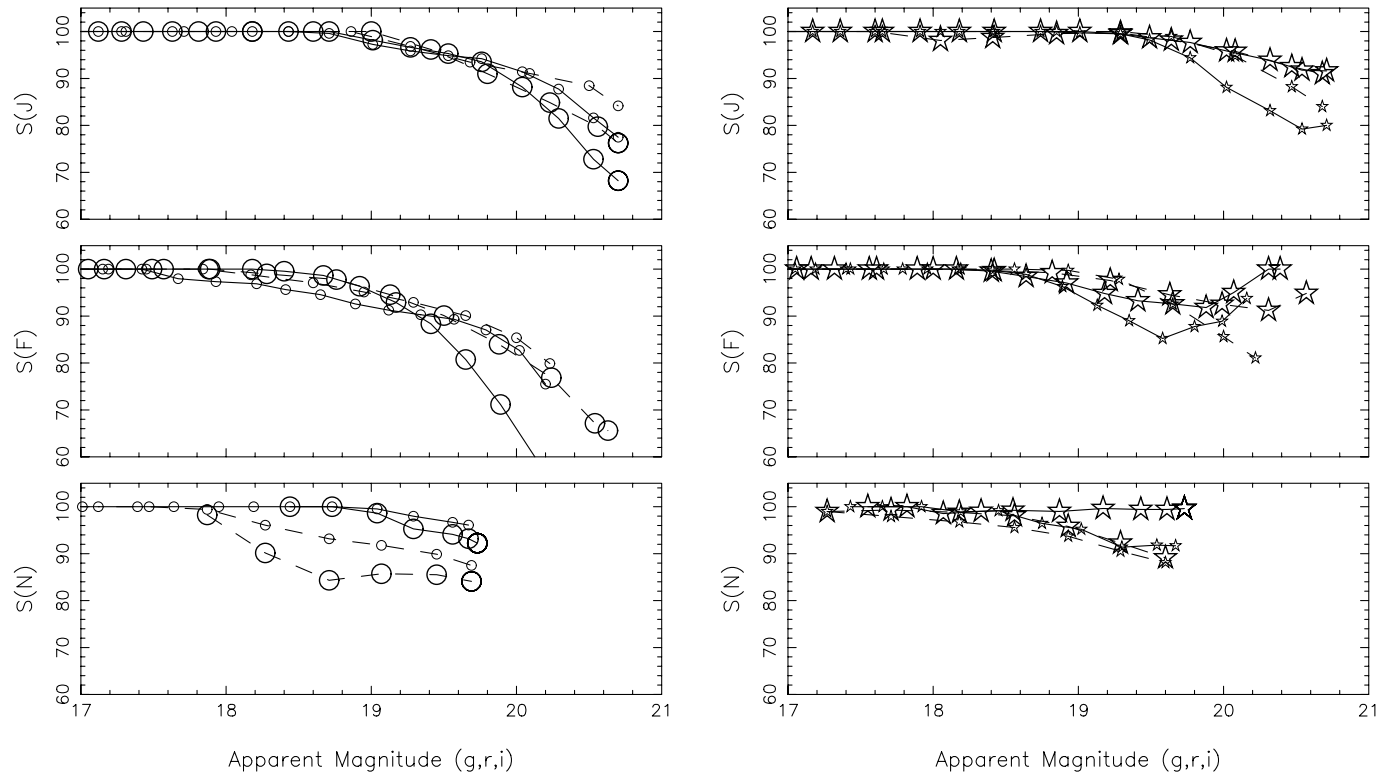

FIG. 5.—ANN and DT classifier success rates as a function of apparent magnitude. The right-hand panels are for stars, and the left-hand panels are for galaxies. Each row of panels represents a different classifier for a bandpass. Training data are connected with solid lines, and the independent test samples are connected with dashed lines. Star relations are plotted with star symbols, and galaxy relations are plotted with open circles.

renormalized these probabilities to be in the range 0.5–1.0, where the value of 1.0 indicates a high degree of classification confidence and values near 0.5 indicate large uncertainty. As summarized in Table 1, we catalog the probability associated with the image class predicted by these two types of pattern classifiers.

6. MORE STRINGENT ASSESSMENTS OF CLASSIFICATION QUALITY

We have applied a series of checks to verify the quality of the final DPOSS image catalogs. On the basis of our estimated success-rate functions derived during classifier training, we found little evidence for choosing one pattern classifier over the other. Indeed, the two methods appear to be very similar in quality, which suggests that our ability to estimate correct image classes is limited by the quality of the image parameters and not the pattern classification algorithm. To verify the quality of our final catalogs, we conducted several checking exercises. These tests can be divided into two categories: (1) comparisons between image classifications from more than one plate for sources located in plate overlap zones and (2) comparison between DPOSS images classes and those from CCD images in DPOSS fields not used in classifier training. In the first case, we wish to confirm that classifier performance is consistent with that predicted by the success-rate relations of Figure 5 and not severely degraded by optical distortion effects near the plate edges. In the second case, we use a completely independent test sample that excludes all of the DPOSS fields used in classifier training to confirm the success-rate functions and to investigate the contamination of star/galaxy catalogs drawn from DPOSS.

6.1. *Plate Overlap Comparisons*

Following Odewahn (1995), we performed a consistency check that is not limited by the availability of CCD calibration fields and involves a large number of sources. A plate overlap analysis was performed in which the spatially overlapping sections of plate catalogs were positionally cross-matched. The image classes assigned to each cross-matched source were then compared as a function of DPOSS magnitude, in the same manner as the CCD test described below. In reality, this approach represents a significantly more rigorous test; in the CCD test, we can consider the CCD samples, which are of much higher weight, to be essentially error free in the magnitude ranges in which the tests are conducted. In the plate overlap analysis, we are comparing two class samples that are characterized by the same level of misclassification, and hence our errors added in quadrature will produce a roughly $\sqrt{2}$ increase in class-comparison scatter. In addition, this analysis is more subject to errors caused by positional cross-matching in crowded fields, since we have no straightforward way of excluding such cases from the analysis. Finally, as discussed by Gal et al. (2004), photometric errors near the plate boundaries due to optical vignetting are significant. We expect classifier performance to degrade near the plate edges solely as a result of decreased relative S/N in the images for each magnitude bin used in our analysis. On the other hand, we might consider this overlap analysis a worst-case assessment of the errors resulting from many different sources. These results, based on a large percentage of the catalogs, can be considered a good consistency check and an honest upper limit to the expected classification errors. In Figure 6 we summarize the results of

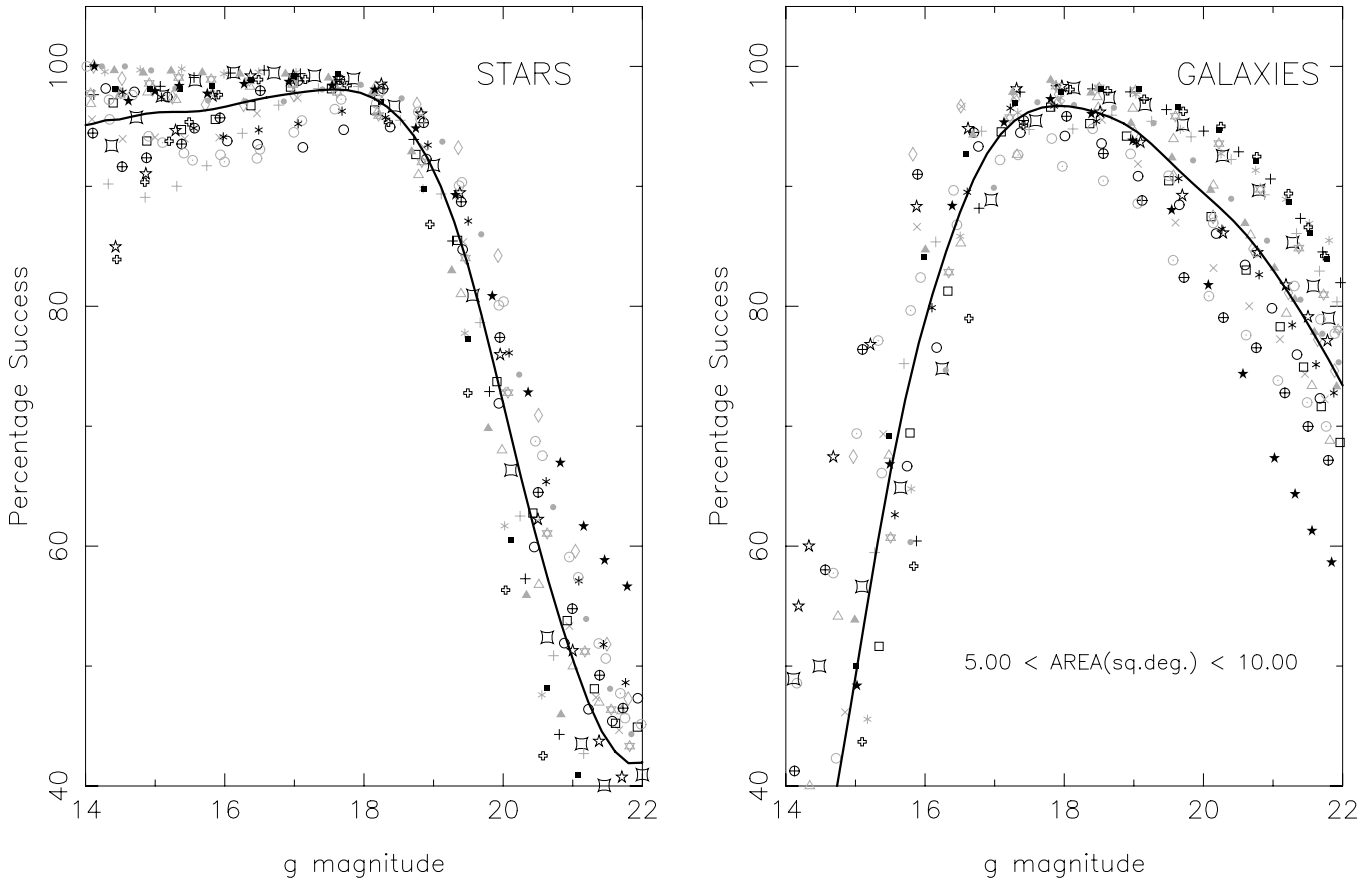

FIG. 6.—Percentage agreement in class assignment for stars (*left*) and galaxies (*right*) observed in overlapping regions of 20 DPOSS plate pairs. This is a more rigorous test of image classification in the DPOSS catalogs, compared with the CCD-based training data of Fig. 4, since it folds together errors due to image detection and parameterization, as well as the systematic errors due solely to the pattern classifier (i.e., the gradual inability to assign classification as image resolution and S/N decrease). A total of 43 overlap zones were used in the final 36 plate survey area. The minimum area allowed for an overlap region was 5.0 deg^2, excluding the use of corner zone overlaps.

our classification comparisons from 43 overlap zones in a 36 field survey conducted at the north Galactic hemisphere (NGH).

We call the reader's attention to the fact that percentage agreement in this case is measured solely as a percentage of matching classifications in adjacent plates and does not take into account the proper probability estimates, as was done in Hambly et al. (2001). Thus, this figure mainly demonstrates that our classifier provides consistent class estimates among plates. In addition, for reasons discussed earlier, the overlap areas are not the most appropriate for rigorous estimates of the classifier success. This accounts for the large scatter present in both panels of Figure 6.

No attempt was made to improve star-galaxy separation at the bright end ($g < 17.0$), as the attributes originally measured in FOCAS are severely affected by the nonlinearity of the density measurement on the plates. Even the most modern pipelines (SDSS for instance) have difficulties classifying large objects, requiring separate processing that is outside the scope of this paper. However, we include the bright end in the most of the plots in this paper to emphasize the degradation of our data and ML algorithms at these magnitudes, exhibiting the limits of applicability of the present data.

6.2. *Classification Success Rates from CCD Comparisons*

For a final DPOSS classification quality assessment, we cross-matched 46 P60 CCD catalogs on 21 DPOSS fields that were not used in classifier development. This procedure allowed us to gather a large testing set, with the advantage that we were testing DPOSS plate catalogs that were in no way involved in the development of the DT and ANN image classifiers. Using these CCD *gri* samples, we were able to rederive our classification success-rate functions in each DPOSS bandpass. As with our classifier training sets, ground truth in every case is derived from objects for which the *g*, *r*, and *i* CCD (FOCAS) classifications agreed. We stress that no data cleaning was applied (i.e., no attempt was made to interactively reject outliers in classification parameter spaces) prior to calculating the success-rate functions. Hence, this final analysis reflects more faithfully the quality of data samples drawn from a typical DPOSS catalog query. The success rates derived in this way are summarized in Figure 7. For convenience, we used spline fits to the *J*,*F* data in Figure 7 to tabulate the average classifier success rates as a function of magnitude given in Table 3. The *J*- and *F*-plate catalogs produced superior classification results compared with the lower S/N *N*-plate material.

In Figure 8 we map contours of constant classification success rate for galaxies as a function of magnitude and class probability. Similar curves were found for star samples. Each curve corresponds to a fixed classification success rate. Thus, users desiring extremely pure galaxy samples of a given object type may wish to perform a cut on the classification probabilities. However, a stringent cut in probability results in low sample completeness. This is shown in Figure 9, where we map contours of sample completeness as a function of magnitude

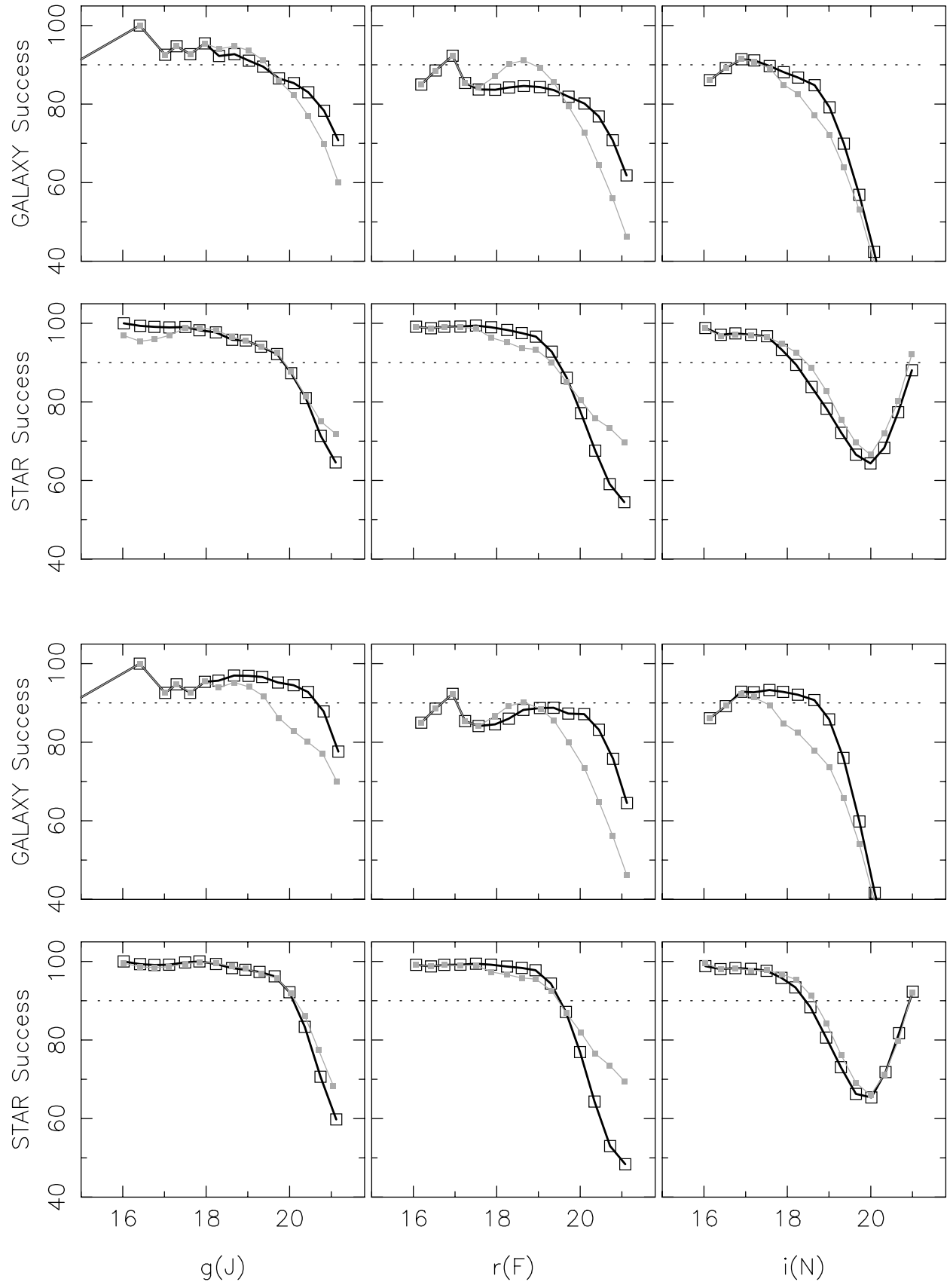

FIG. 7.—Plot as a function of (g, r, i) magnitude of the percentage agreement between our DPOSS *J*, *F*, and *N* image classes and those derived from FOCAS analysis of deep CCD images of the same objects. In all panels we use large open squares to indicate ANN estimates and small filled squares connected by light-gray lines for DT estimates. In the top two panels we show galaxy (*top*) and star (*bottom*) rates using the entire data set. In the bottom two panels we show the same functions, except that a classification probability level of $\geq$0.8 is required.

TABLE 3
AVERAGE CLASSIFIER SUCCESS RATES

Mag. (1)	Galaxy(g) (2)	Star(g) (3)	Galaxy(r) (4)	Star(r) (5)
17.00.....	92.6	99.1	91.3	99.2
17.50.....	93.3	99.8	84.0	99.4
18.00.....	95.6	99.9	84.7	99.0
18.50.....	96.4	98.6	87.5	98.4
19.00.....	96.9	97.8	88.7	97.4
19.50.....	96.1	96.9	88.2	90.8
20.00.....	94.8	92.2	87.4	76.9
20.50.....	92.5	79.2	82.5	59.1
21.00.....	81.2	61.4	66.9	48.6

NOTES.—Following a magnitude in col. (1) we list the percentage success rate for classifying galaxies and stars in the g and r bands computed from spline fits to the data in Fig. 7. Since N-plate classifier results were much less useful, we delete them here.

and class probability. For instance, to select a sample of galaxies that is 90% pure at $r = 19$ mag, it is necessary to apply a probability cut of `prob_ANN_F` > 0.8. This results in a completeness of only $\sim$90%, however.

6.3. *Catalog Contamination Rates from CCD Comparisons*

Since the surface density of stars varies greatly over the sky, while the surface density of galaxies (corrected for extinction) remains roughly constant, the contamination signal (e.g., stars misclassified as galaxies) can vary greatly across the sky. Therefore, any single estimate of the contamination (as presented in Weir et al. 1995b, for instance) is insufficient. This is a major but solvable problem for many projects and for *any* sky survey, yet it is seldom addressed with the care it requires.

In our analysis of the classification success, we suffer from the opposite problem. Because most of our CCD imaging targets are Abell clusters, the relative frequency of galaxies compared with stars is higher than in the average field. Therefore, we will overestimate the contamination rate of true stars by galaxies misclassified as stars.

To alleviate this problem, we have compared the galaxy counts in our Abell CCD fields with those in some blank fields that were observed under the same conditions. As shown in Figure 10, we find that our Abell fields have approximately twice as many galaxies as those in blank fields at comparable Galactic latitudes, a trend that persists at all magnitudes of interest. Based on this finding, we chose to apply a correction factor β ($\approx\frac{1}{2}$) to the stellar contamination rate (i.e., we divide the rate of "false" galaxies by 2) and increase the galaxy contamination rate by the same factor (since the number of real galaxies in a typical field will be half that in our CCD pointings). A complementary approach based on the analysis of calibrated star/galaxy number counts drawn from a large area of the sky is presented in § 8.

Following Weir et al. (1995b) we assess the contamination in magnitude bins for each classifier trained in each bandpass. Stellar contamination (misclassified galaxies contaminating stars) is given by

$$C_{\rm star} = \frac{\beta N([\mathrm{CCD} = \mathrm{star}] \bigcap [\mathrm{Plate} = \mathrm{gal}])}{N(\mathrm{CCD} = \mathrm{star})}. \quad (3)$$

Galaxy contamination (misclassified stars contaminating galaxies) is given by

$$C_{\rm gal} = \frac{N([\mathrm{CCD} = \mathrm{gal}] \bigcap [\mathrm{Plate} = \mathrm{star}])}{\beta N(\mathrm{CCD} = \mathrm{gal})}. \quad (4)$$

Galaxy catalog contamination functions estimated for each classifier are shown in Figure 11, in which we plot the percentage contamination as a function of apparent magnitude.

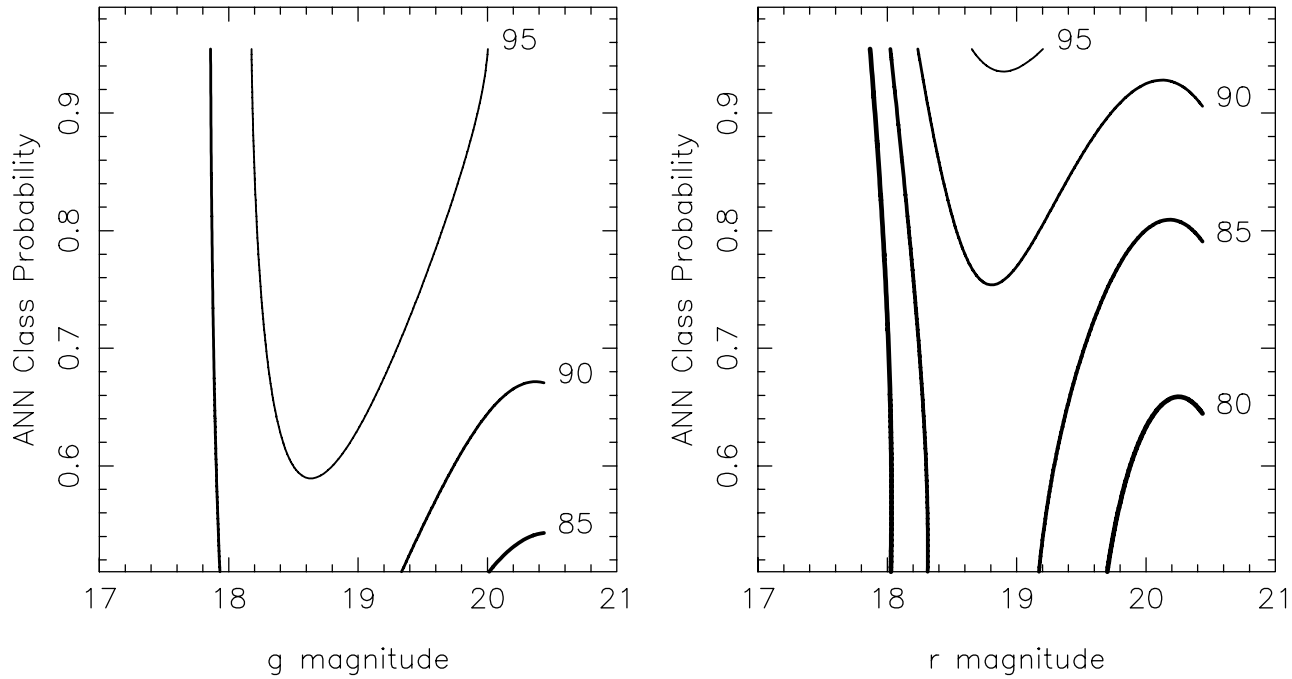

FIG. 8.—Contours of constant classification success rate for galaxies as a function of magnitude and class probability. Each curve corresponds to a fixed classification success rate. Given a minimum probability requirement in drawing a DPOSS sample, the user can determine the magnitude range over which a classification success rate is maintained. Although these curves were computed from galaxy samples, nearly identical curves were derived for DPOSS star samples.

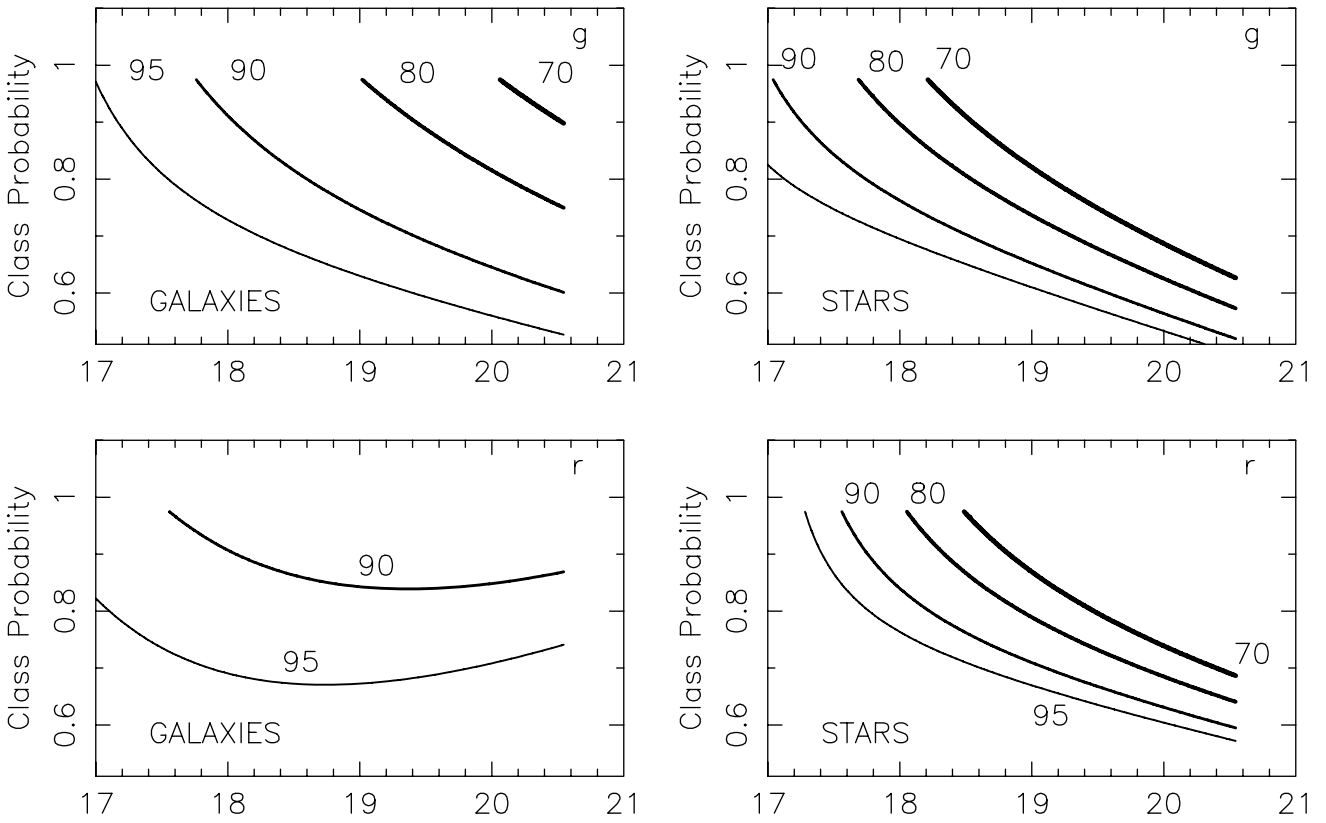

FIG. 9.—Contours of sample completeness as a function of magnitude and class probability. In this case, completeness refers to the percentage of sources not eliminated from a sample by restricting the classes to have a probability greater than the limit specified on the *y*-axis. For a given magnitude, it is possible to determine the probability restriction for which a sample will remain 95%, 90%, 80%, or 70% complete compared with a sample drawn using no probability restriction.

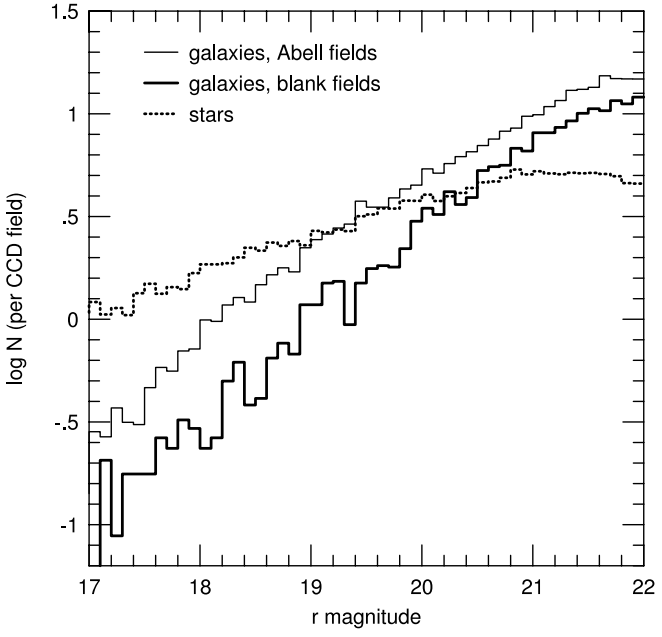

FIG. 10.—Star and galaxy number counts from the Palomar 60 inch CCD fields discussed in Gal et al. (2000, 2004). These data were used to establish image classifier training and testing samples and to compute functions predicting catalog completeness and contamination as described in the text.

The contamination function is strongly dependent on the true density of contaminating objects on the sky, and hence the galaxy contamination function will be strongly dependent on Galactic latitude. The sample of fields used for this exercise has $\langle b \rangle = 43^\circ$. Strictly speaking, it is necessary to derive a contamination function for the range of b used to construct a given catalog. We include the functions shown in Figure 11 as a purely illustrative example.

More accurate tests of the classifier accuracy would require a homogeneous, wide-area catalog constructed from a significantly higher angular resolution data set, either from deep, high-quality ground-based imaging or *HST* data. While some such data are available, the construction of photometric catalogs from existing archives that would cover a significant number of plates is beyond the scope of this paper. We note that our CCD imaging data is of comparable quality to modern surveys, such as the SDSS.

7. FINAL CLASS ASSIGNMENTS IN THE DPOSS CATALOG

Each *JFN*-plate set in DPOSS was reduced as a matched catalog set. The class for each detected source is a numerical value that is summarized in Table 4. As discussed in § 2, all detected sources were assigned a FOCAS class based on the `scale` and `frac` parameters, as outlined by Valdes (1982). In cases in which FOCAS split a detection into multiple sources, it was determined that several of the attributes used by the ANN and DT classifiers discussed in § 3 were computed incorrectly. Specifically, the values of `Mtot` − `Mcore`, `IR2`, and `Area` are not properly evaluated for the child sources. In such cases,

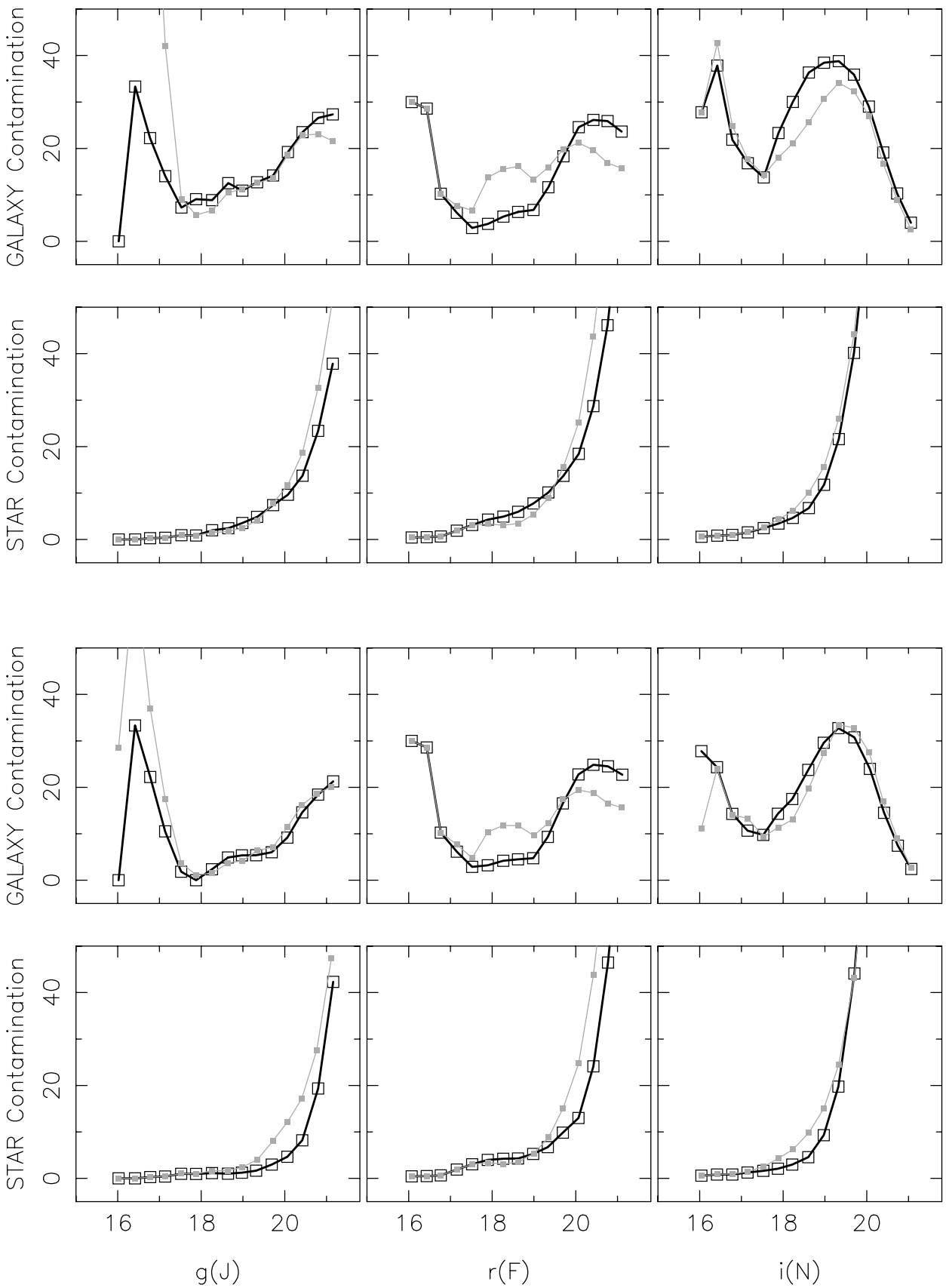

FIG. 11.—Plot as a function of (g, r, i) magnitude of the percentage of contamination in DPOSS J, F, and N image catalogs. As in Fig. 7, ground truth image classes are those derived from FOCAS analysis of deep CCD images. In all panels we use large open squares to indicate ANN estimates and small filled squares connected by light gray lines for DT estimates. In the top two panels we show the contamination rates in galaxy (*top*) and star (*bottom*) catalogs from DPOSS using the entire data set. In the bottom two panels we show the same functions except that a classification probability level of ≥ 0.8 is required. The sample of fields used for this exercise has $\langle b \rangle = 43^{\circ}$.

TABLE 4
SUMMARY OF DPOSS NUMERICAL CLASS ASSIGNMENTS

Numerical Class	Source Type
1	Star
2	Galaxy
3	Noise (FOCAS)
4	Long (FOCAS)
5	Diffuse (FOCAS)
6	Fuzzy star (FOCAS)

NOTE.—FOCAS (Valdes 1982) classes determined in the initial DPOSS pipeline reductions were retained if a source was split into multiple sources.

the original FOCAS class is adopted. In the case of unsplit sources, which comprises the majority of the catalog at $r < 19.5$ mag, both ANN and DT class estimates were computed for every source detected in a given band. As we have shown in § 6, the ANN classifier generally gave slightly superior classification results, and hence for each bandpass the ANN class is adopted as a "best" classification for each cataloged source. In rare cases, one or more of the catalog attributes for a source falls outside of the acceptable range used in training the ANN classifiers. In these case, the DT class values was adopted. The general user will desire a single class value that can be queried for any object, and hence we have computed a final numerical class value, `ClassBEST`. Initially, we investigated using the class predicted by the plate having the superior seeing. However, numerous tests showed that the *J*-plate classification estimates were superior to those from the *F* or *N* plates, and hence we simply adopted the *J*-plate classification as the best class. For objects having no *J*-plate detection, the *F*-plate class was adopted. Finally, for single *N*-plate detections, the *N*-plate classification was adopted for `ClassBEST`.

8. GALAXY COUNTS: A FINAL CONSISTENCY CHECK

As a final check of the DPOSS classification quality we computed the star and galaxy number counts in g and r on 341 DPOSS fields covering a total of 7756 deg^2. The rigorous set of tests described thus far provide well-defined functions to predict the levels of contamination and classification completeness in a catalog drawn from DPOSS. All of these derivations are by necessity drawn from data sets derived from cross-matching DPOSS and CCD-based image catalogs and hence represent a restricted sample. For instance, sources whose image parameters, and hence object classifications, are degraded by image crowding are also likely to have discrepant catalog centroid positions. Such sources are frequently dropped from a CCD-DPOSS cross-match, resulting in a data set that gives an overly optimistic estimate of the final catalog quality. To gain a more realistic appraisal of our catalog, we mapped source counts across a wide area of the sky, in both the north (NGH) and south (SGH) Galactic hemispheres. All catalogs were photometrically calibrated as described in Gal et al. (2004), and we adopted Galactic extinction corrections following Schlegel et al. (1998). Number counts were computed using our smallest spatial areal unit, the $30' \times 30'$ footprint cells of DPOSS. Assuming that galaxy counts in the range $16 \leq g, r \leq 20$ should be uniform in the mean over a 25 deg^2 field (i.e., the size of the digitized area of the POSS-II plates), we computed g and r galaxy counts over 341 DPOSS fields and determined the level of plate-to-plate variation present in our catalogs. The same can be done for stellar data, but a large variation in stellar counts due to Galactic structure makes such an analysis highly dependent on the adopted Galactic structure model. The main point is that we should be able to construct global number counts and density distribution maps that are relatively free of the statistical imprint of the scale size of a POSS-II Schmidt plate.

In Figure 12 we show the g galaxy counts as derived from the `ClassBEST` DPOSS estimates described in the previous section. In this case, we used catalog sources drawn independently from the *g*-band and *r*-band detections. Using only those galaxies detected both in g and r decreases the total number of sources at the 5%–10% level over the range of magnitudes shown in Figure 12 and systematically decreases the slope in the number counts by 2%–3%. We plotted number counts in Figure 12 drawn from five ranges of absolute Galactic latitude, $|b|$, and a clear trend is evident, such that the mean density of galaxies at $g = 18.5$ mag is systematically higher at low Galactic latitude. For all fields in this range of $|b|$ we fitted a linear regression to the count sets of the form

$$\log \Psi = \alpha(m - 18.5) + \log \rho. \quad (5)$$

Here Ψ is the number of sources per area in square degrees in a 0.5 mag wide binning interval. The value of ρ is expressed in the same units, but refers to the object number density at a fixed magnitude value. The value of $g, r = 18.5$ mag was chosen as a fiducial magnitude point since it was judged to be stable in classification and photometric quality over many DPOSS fields for both stars and galaxies in the *J*- and *F*-plate catalogs. Using all of the data in the range $30 \leq |b| \leq 90$ and the magnitude interval $18 \leq g \leq 20$ we obtained a slope of $\alpha = 0.443 \pm 0.002$ and a mean density of $\log \rho = 1.984 \pm 0.001$ at $g = 18.5$ mag. This mean linear fit to the data is shown as a heavy solid line in Figure 12. In the right-hand panel of Figure 12 we show the residuals of each count set relative to this mean number count fit, revealing more clearly the strong trend in mean count level with $|b|$.

In Figure 13 we show the fitted density of galaxies at $g, r = 18.5$ mag (i.e., $\log \rho$ for galaxies by DPOSS field) as a function of the local stellar density ($\log \rho$ for stars) at $g, r = 18.5$ mag. The same trends are evident using linear regressions from individual DPOSS footprints, but plate-wide solutions make the trend smoother and more well defined. In each panel of Figure 13 we plotted a model line predicting the relative increase in the number of galaxies if some portion of galaxies are actually misclassified stars. This problem becomes extreme at low Galactic latitudes for which the number of misclassified stars will contribute substantially to our galaxy catalogs. In Figure 13 we used stellar classification error rates of 9% at $g, r = 18.5$ mag to produce the model curves, for which the excess number of observed galaxies is clearly well explained by classification errors. The rather high error rate of 9% seems excessive compared with the stellar success-rate curves plotted in Figures 5 and 7. However, we stress that these errors involve other factors. As with the plate overlap analysis of Figure 6, we must now consider errors introduced by variations in image matching and plate detection limits. Even if we neglect the requirement of a g, r matched detection, we can still suffer systematic errors because of the fact that superior classes are drawn from the *J*-plate material. If a large relative fraction of *F* sources fail to match with the correct *J*-plate detection, then the inferior *F* classification alone will be used in assigning an image type in the catalog. In reality, our error rate of 9% is quite comparable to the results of Figure 6 in the range $18 < J, F < 19$.

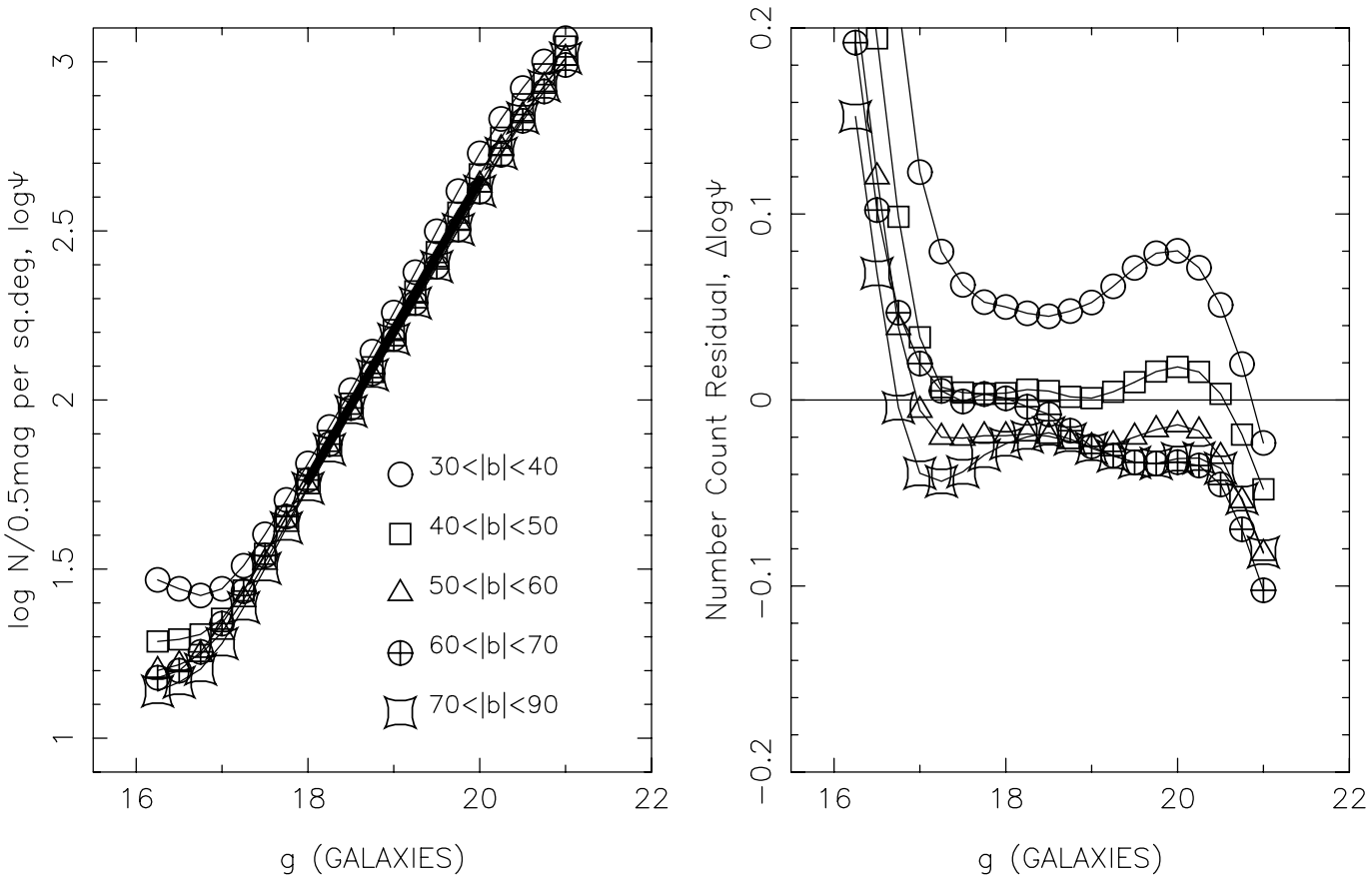

FIG. 12.—*Left*: Galaxy counts in g for five different ranges of absolute Galactic latitude, $|b|$, covering approximately 7000 deg^2 of sky. The solid dark line represents the mean of all five sets. The g magnitudes were corrected for Galactic extinction following Schlegel et al. (1998). The two lowest latitude sets clearly show a systematic increase in mean galaxy density. *Right*: Count residuals relative to a linear fit to the mean count set. The same symbol notation from the counts panel applies here. A statistically significant trend in mean count level with $|b|$ is present. This trend, presumed to be the result of stellar contamination, is present to similar degrees in both the northern and southern Galactic hemisphere data. Very similar results were derived for the r galaxy counts.

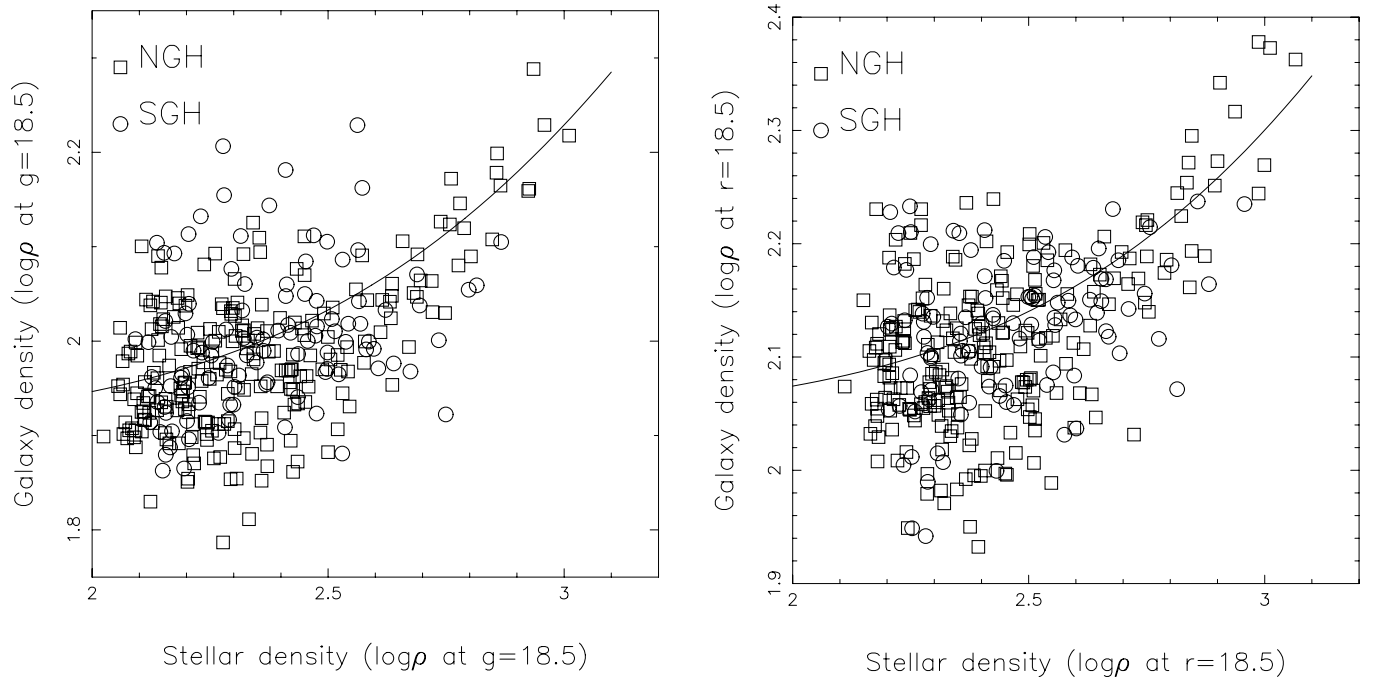

FIG. 13.—*Left*: Trend in galaxy number density at $g = 18.5$ with measured stellar number density at $g = 18.5$. *Right*: Trend in galaxy number density at $r = 18.5$ with measured stellar number density at $r = 18.5$. The solid line in each represents the number of galaxies we expect to observe if 9% of the local star population is misclassified. This level of galaxy catalog contamination will increase substantially with decreasing Galactic latitude.

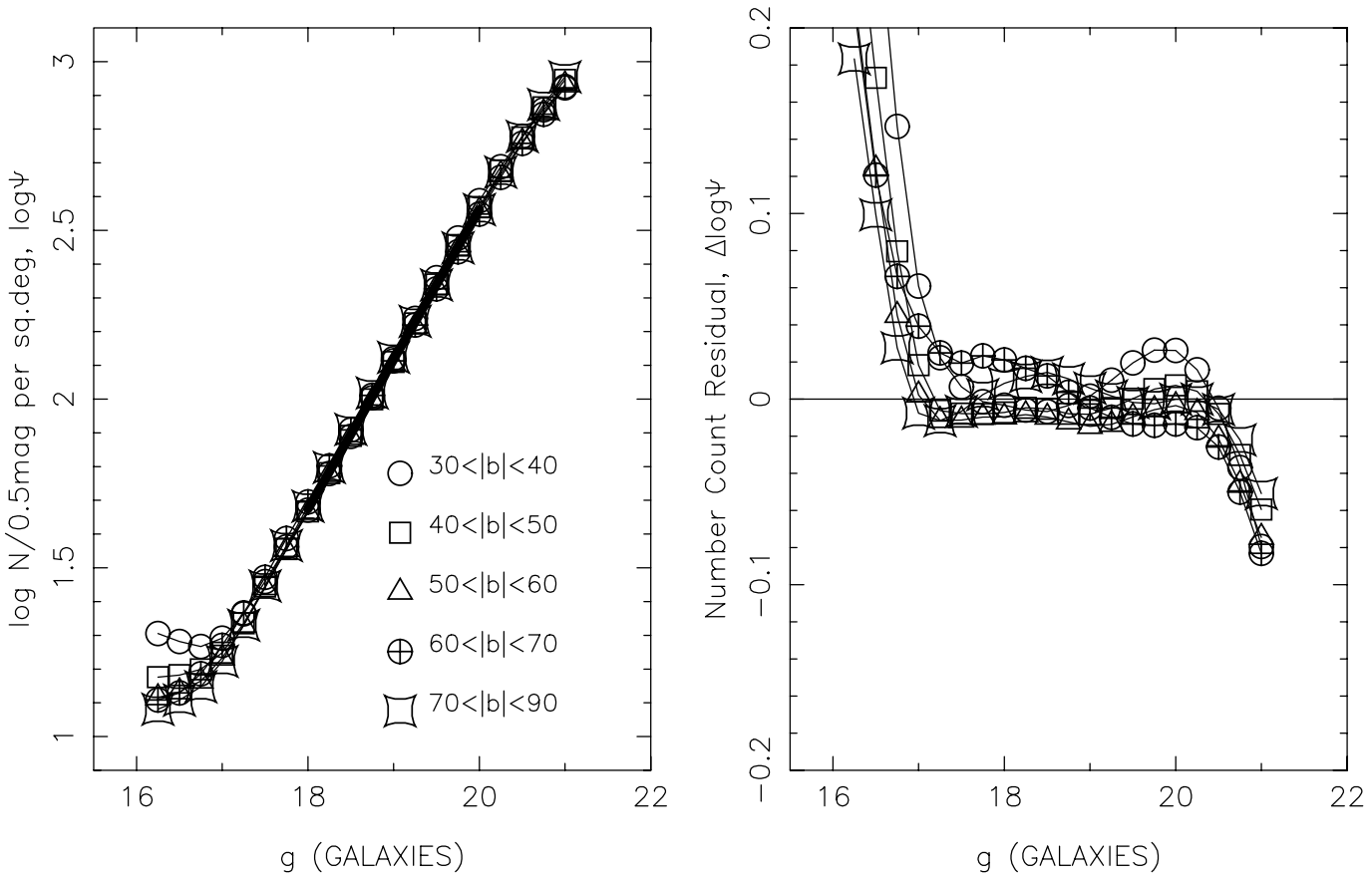

FIG. 14.—*Left*: Corrected galaxy counts in g from Fig. 12 for five different ranges of absolute Galactic latitude, $|b|$. Counts were corrected using the local stellar density and the simple stellar contamination model of Fig. 13. The solid dark line represents the mean of all five sets with a fitted slope of $\alpha = 0.444 \pm 0.001$ and an intercept of $\log \Psi = 1.896 \pm 0.001$ at $g = 18.5$. *Right*: Count residuals relative to the linear fit to the mean count set. The same symbol notation from the counts panel applies here. The trend in mean count level with $|b|$ has been removed. Very similar results were derived for the r galaxy counts.

In Figure 14 we show the same number counts of Figure 12, but drawn from galaxy catalogs corrected for stellar contamination effects using the simple models of Figure 13. For the mean of these counts we derived for $18 \leq g \leq 20$ a slope of $\alpha = 0.444 \pm 0.001$ and a mean density of $\rho = 1.896 \pm 0.001$ at $g = 18.5$. Hence, the mean linear fit parameters change little compared with our uncorrected data fits. However, as can be seen in the residual trend panel of Figure 14, our large systematic variation with $|b|$ has now been substantially decreased. This ensures that we can use galaxy catalogs drawn from DPOSS for statistical studies down to a Galactic latitude as low as $|b| = 20^\circ$.

As a final test, we used our corrected count fits drawn from 21,120 footprints in 205 northern Galactic latitude DPOSS fields to construct density maps of the distribution of stars and galaxies. In Figure 15 we show the star/galaxy density maps in Flamsteed projection in g, r. We have used 205 DPOSS fields where the photometric calibration using plate overlap and CCD standard field methods discussed in Gal et al. (2004) were judged to give photometry for which the 1 σ error in the zero point per plate is 0.07 mag. Assuming a Gaussian distribution in plate zero-point errors and mean galaxy count slope of $\alpha = 0.45$, the expected plate-to-plate 1 σ variation in log ρ at $g = 18.5$ would be 0.02 dex if photometric shifts alone were causing the number count variations. The effect on stellar number counts, which have a shallower slope, would be less. The observed 1 σ plate-to-plate scatter in the 352 DPOSS fields used in this work was found to be 0.08 dex in both g and r for galaxies. This latter observed scatter is of course a combination of photometry and classification errors. It would appear, on the basis of these statistics, that the bulk of the observed field-to-field fluctuations in source distribution can be attributed mostly to errors in star-galaxy separation. Moreover, such errors would likely produce deviations with spatial scale sizes of 5°.5 (the size of a DPOSS survey field) and would produce a blocky structure in our density maps.

9. CONCLUDING REMARKS

We have developed a suite of classification methods for cataloging the sources detected on digitized scans of the Second Palomar Sky Survey and the resulting Palomar-Norris Sky Catalog (PNSC). Image classifiers based on both artificial neural networks and decision trees were developed. These classifiers take as input sets of global photometric image parameters that are sensitive indicators of whether a source is resolved or unresolved. To remove plate-dependent variations in these image parameters, we developed a robust stellar locus mapping algorithm, which is used to transform all image attributes to a common system prior to classification. We investigated the level of classification success and catalog contamination for each classifier, from each of the three (*JFN*) plates used in the survey. These classification tests are based both on extensive plate overlap analysis to test the internal consistency of the classifications in the catalog and on comparisons with 46 CCD-derived image catalogs on 21 different DPOSS fields. We characterized how classification confidence measures can be used to control classification success and catalog contamination as

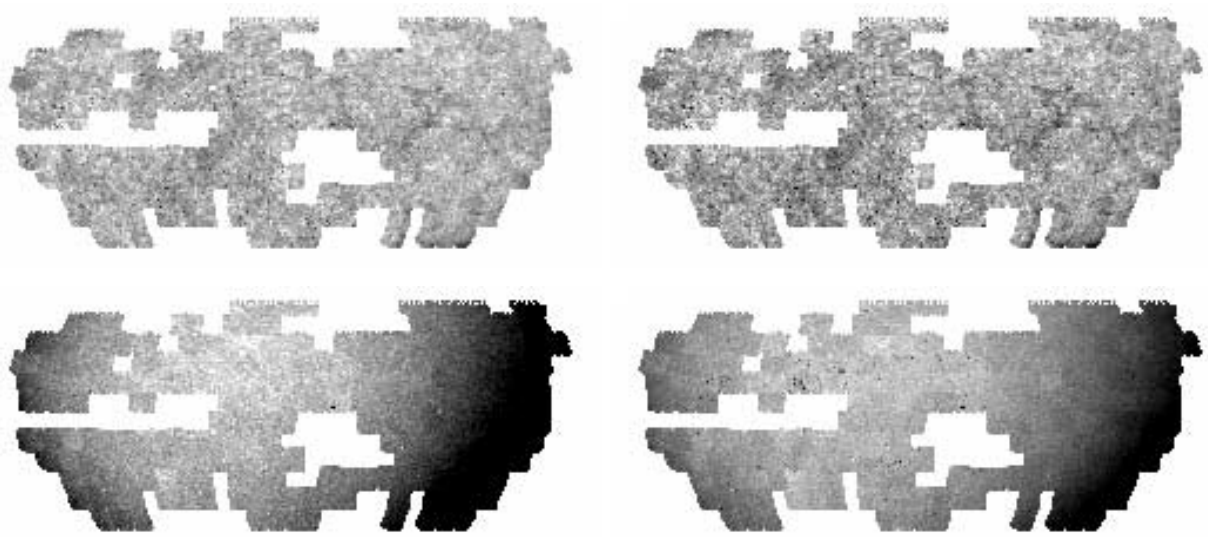

FIG. 15.—Four views of the DPOSS catalog from 205 photometrically calibrated DPOSS fields mapping 5280 deg^2 in the NGH. To achieve a roughly equal area projection, all maps are shown in Flamsteed projection with the field center at $\alpha, \delta = 12:44 + 27:30$ (close to the NGH). The linear size of each map is roughly $145^\circ \times 63^\circ$. We show the fitted number density of galaxies at $g = 18.5$ (*top left*), galaxies at $r = 18.5$ (*top right*), stars at $g = 18.5$ (*bottom left*), and stars at $r = 18.5$ (*bottom right*). All maps have been scaled to a level of $\pm 3\ \sigma$, where σ is the standard deviation in count level for a single footprint computed over the entire range of each map. Since a footprint has a linear size of $30'$, only very large scale structure is visible here. The smoothness of each map indicates very little evidence for plate-to-plate systematics in the star-galaxy separation. Such systematic errors would manifest as a distinctively blocky appearance. There is little evidence of such trends even in areas of high stellar density at low Galactic latitudes (*extreme right and left sides of each map*).

a function of apparent magnitude. Finally, we made star and galaxy number counts in a large number of DPOSS fields covering a fair portion of the sky in both the northern and southern Galactic hemispheres to show that no statistically significant plate-to-plate variation is present.

This paper makes use of machine-learning techniques to establish star-galaxy separation. The ML approach offers some advantages over traditional parametric methods. First, we can use a large number of attributes to infer the rules for distinguishing the object classes. In addition, these ML algorithms determine statistical classification rules allowing a meaningful estimate of the probabilities for the final class. Another important aspect of using ML techniques is being able to foresee the application of methods of unsupervised classification, allowing the data to set the relevant and distinct classes within the sample, rather than the scientist. Recent publications still make use of parametric approaches and report similar accuracies down to the same limiting magnitudes as ours (e.g., Stoughton et al. 2002; SDSS; Hambly et al. 2001). However, as we move to fainter levels, the ML approaches may be potentially better since the separation line subjectively defined in the parametric approaches is difficult to set and often not reproducible (see Weir et al. 1995b for further discussion).

The Norris Foundation has provided generous support for the DPOSS project and creation of the PNSC. R. R. G. was supported in part by an NSF Fellowship, NASA GSRP NGT5-50215, and a Kingsley Fellowship. Some software development was supported by the NASA AISRP program. We owe a debt of gratitude to the entire POSS-II survey team for their dedication to the long and arduous process of taking over 2700 photographic plates. We are indebted to the plate scanning team at STScI, especially Barry Lasker, for their efforts. We also thank numerous Caltech undergraduates who assisted in data taking at the Mount Palomar 60 inch telescope.

REFERENCES

Brunner, R. J., et al. 2003, AJ, 126, 53
Djorgovski, S. G., de Carvalho, R. R., Gal, R. R., Odewahn, S. C., Mahabal, A. A., Brunner, R., Lopes, P. A. A., & Kohl Moreira, J. L. 2003, Bull. Astron. Soc. Brazil, 23, 197
Fayyad, U. 1991, Ph.D. thesis, Univ. Michigan
Fayyad, U., & Irani, K. 1990, Proc. AAAI-90 (Saarbrücken: LIDO), 783
Gal, R. R., de Carvalho, R. R., Brunner, R., Odewahn, S. C., & Djorgovski, S. G. 2000, AJ, 120, 540
Gal, R. R., de Carvalho, R. R., Lopes, P. A. A., Djorgovski, S. G., Brunner, R. J., Mahabal, A., & Odewahn, S. C. 2003, AJ, 125, 2064
Gal, R. R., de Carvalho, R. R., Odewahn, S. C., Djorgovski, S. G., Mahabal, A., Brunner, R., & Lopes, P. A. A. 2004, AJ, 3082
Hambly, N. C., Irwin, M. J., & MacGillivray, H. T. 2001, MNRAS, 326, 1295
Heydon-Dumbleton, N. H., Collins, C. A., & MacGillivray, H. T. 1989, MNRAS, 238, 379
Kennefick, J. D., Djorgovski, S. G., & de Carvalho, R. R. 1995, AJ, 110, 2553
Odewahn, S. C. 1995, PASP, 107, 770
———. 1997, Proc. SPIE, 3164, 110
Odewahn, S. C., Humphreys, R. M., Aldering, G., & Thurmes, P. 1993, PASP, 105, 1354
Odewahn, S. C., Stockwell, E. B., Pennington, R. L., Humphreys, R. M., & Zumach, W. A. 1992, AJ, 103, 318
Quinlan, J. 1986, Machine Learning: An Artificial Intelligence Approach, ed. Y. Kodratoff & R. Michalski, Vol. 1 (San Mateo: Morgan Kaufmann)
———. 1990, Machine Learning: An Artificial Intelligence Approach, ed. Y. Kodratoff & R. Michalski, Vol. 3 (San Mateo: Morgan Kaufmann)
Reid, I. N., et al. 1991, PASP, 103, 661
Richard, M. D., & Lippmann, R. P. 1991, Neural Comput., 3, 461
Rumelhardt, D. E., & McClelland, J. L. 1988, Explorations in Parallel Distributed Processing (Cambridge: MIT Press)
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Sebok, W. L. 1979, AJ, 84, 1526
Stoughton, C., et al. 2002, AJ, 123, 485
Thuan, X. T., & Gunn, J. E. 1976, PASP, 88, 543
Valdes, F. 1982, Proc. SPIE, 331, 465
Weir, N., Djorgovski, S., & Fayyad, U. M. 1995a, AJ, 110, 1
Weir, N., Fayyad, U. M., & Djorgovski, S. 1995b, AJ, 109, 2401
Weir, N., Fayyad, U. M., Djorgovski, S. G., & Roden, J. 1995c, PASP, 107, 1243