THE ASTRONOMICAL JOURNAL VOLUME 103, NUMBER 1 JANUARY 1992

AUTOMATED STAR/GALAXY DISCRIMINATION WITH NEURAL NETWORKS

S. C. ODEWAHN, E. B. STOCKWELL, R. L. PENNINGTON, R. M. HUMPHREYS, AND W. A. ZUMACH
Department of Astronomy, University of Minnesota, Minneapolis, Minnesota 55455
Received 12 June 1991; revised 29 August 1991

ABSTRACT

We discuss progress in the development of automatic star/galaxy discriminators for processing images generated by the University of Minnesota Automated Plate Scanner (APS) for cataloging the first epoch Palomar Sky Survey. Classifications are based on 14 image parameters computed for each object detected by the APS operating in a threshold densitometry mode. It is shown that a number of parameter spaces formed with these vector elements are effective in separating a sample into the two basic populations of stellar and nonstellar objects. An artificial intelligence technique known as a neural network is employed to perform the image classification. We have experimented with a simple linear classifier known as a perceptron, as well as with a more sophisticated backpropagation neural network with the result that we are able to attain classification success rates of 99% for galaxy images with $B \leqslant 18.5$ and above 95% for the magnitude range $18.5 \leqslant B \leqslant 19.5$. The analysis presented here uses a training dataset consisting of 2665 galaxies and 2082 stars, along with a test sample of 936 galaxies and 2378 stars. We have determined the success rate of these classifiers as a function of image diameter and integrated magnitude. Simple numerical experiments have been conducted in an effort to illustrate the robust nature of this method as well as to isolate the most significant image parameters used by the networks in distinguishing image class.

1. INTRODUCTION

Many of today's most relevant astrophysical problems concerning galactic structure and dynamics, environmental effects on galaxy formation and maintenance, and the large-scale distribution of matter in the Universe are approached in a statistical fashion using deep surveys of stars and galaxies over large areas of the sky. Most astronomers are familiar with the numerous studies and practical applications made with stellar catalogs such as the Yale Bright Star Catalog (Hoffleit 1965), and the SAO catalog (Veis, 1966). With the advent of new ground-based spectroscopic survey instruments and space-based observatories such as the *Hubble Space Telescope*, deeper and more accurate stellar catalogs are called for. Some of the deepest and most complete galaxy catalogs such as those by Shane & Wirtanen (1967), Zwicky (1968), Nilson (1973), and Corwin (1985) have been compiled through visual inspection of photographic surveys. More recent efforts by Dickey *et al.* (1987), Heydon-Dumbleton *et al.* (1989), Slezak *et al.* (1988), and Rhee (1990) have used fast scanning machines and automated image detection and classification techniques to compile galaxy catalogs in specific areas of the sky.

A major contribution to both stellar and galaxy surveys will be the catalog of images contained on the 936 plate pairs of the first epoch Palomar Sky Survey currently being generated with the University of Minnesota Automated Plate Scanner. Detailed discussions of this project are given by Pennington *et al.* (1987) and Humphreys & Pennington (1989). In this work we describe the automated methods developed to classify images detected with the APS. A single POSS plate will ordinarily produce a list of $\sim$250 000 detected images, and thus the need for a fast and automated means of distinguishing stellar and nonstellar images is obvious. In this work we shall discuss the development of a novel approach to the problem of star/galaxy separation for the APS automated survey. Future papers will deal with problems such as image blending and plate to plate densitometric variations.

2. APS IMAGE DATA

The Automated Plate Scanner (APS) at the University of Minnesota is a modern, high-speed, high-precision plate measuring engine of the type known as a flying spot scanner. It is unique in its ability to measure two plates simultaneously. When combined with its high speed (1080–4320 mm^2/min) and 1 μm repeatability, this unique capability makes the APS especially well suited for digitizing large-scale surveys such as the Palomar Sky Survey. Sky survey plates are scanned in the threshold densitometry (TD) mode in which the scanner records the plate transmission values only when the image is denser than some user-defined threshold level. The threshold level, which corresponds to an approximate level of $\mu_B = 23.5$ B mag arcsec^{-2} for the current survey, is measured relative to a local sky level which is determined by a "smart background tracker." A plate is measured in a series of 30 stripes, each 12 mm wide, to cover the 14 in. POSS plate. For tracking the background in each stripe, a low-resolution prescan is made, and a median filtering process is applied to remove the effects of discrete sources. During the subsequent thresholding scan, in which a resolution of 5×12 μm is attained, positions of the image ingresses and egresses are recorded. Thus, isodensitometric and densitometric information about the image are both available. As explained below, the image transit endpoints are fit with an ellipse in order to provide basic image parameterization. The extra information supplied by threshold densitometry readily allows separation of stars and nonstellar images, as well as decomposing overlapping images.

We have experimented with a variety of methods for effectively parameterizing the images detected by the APS in TD mode. Much of this work follows the methodology of Dickey *et al.* (1987), Rhee (1990), and Heydon-Dumbleton (1989), but unlike the latter two works, we are not using a density to intensity transformation in the image classification stage in order to decrease machine processing time. As demonstrated in this work, a formal density to intensity conversion is not necessary for discriminating stellar and non-

 0004-6256/92/010318-14$00.90

TABLE 1. Image parameters.

diameter	dia
ellipticity	1 - b/a
average transmission	Tav
central transmission	Tc
(ellipse area) / (area from pixel count)	c1
log (area from pixel count)	c2
$(\sum T/r)/(\sum T)$	moment1
rms error of ellipse fit to transit endpoints	fuzz
Y centroid error	jitter
$(T_4 - T_1)/(r_1 - r_4)$	gradient14
$(T_3 - T_1)/(r_1 - r_3)$	gradient13
$(T_2 - T_1)/(r_1 - r_2)$	gradient12
$(T_3 - T_2)/(r_2 - r_3)$	gradient23
$(T_4 - T_3)/(r_3 - r_4)$	gradient34

stellar images. The image parameters adopted for use in this work are summarized in Table 1. The plate transmission T is used directly for the calculation of simple image moments and gradients. Each threshold contour is fitted by an ellipse whose properties are described by the center position X_c and Y_c, major axis diameter DIA, ellipticity ELL $= 1 - b/a$, and major-axis position angle θ. The error in Y_c is given by a parameter known as the JITTER, and the error in DIA is specified by the FUZ. All pixels with transmission values above the scanning threshold (65% of the sky background) are used to reconstruct the image. The peak transmission T_c and average transmission T_{av} have proven to be very effective discriminators when combined with some measure of image size such as the diameter (Figs. 1–3). Other useful parameters for image discrimination are

(1) $c1 = 2\pi a^2/A$, where a is the ellipse semimajor-axis length and A is the image area derived by summing the number of image pixels.

(2) $c2 = \log A$.

(3) $mom1 = \dfrac{\Sigma(T(x,y)/r)}{\Sigma T(s,y)}$, where r is the pixel radius measured from the image center (X_c, Y_c).

(4) $Gij = (T_j - T_i)/(r_i - r_j)$, a simple image gradient in which T_i is the median transmission value in an elliptical annulus (having a shape which matches the threshold isophote) and semimajor axis length r_i.

Five distinct image gradients are formed using 4 image radii, to give a total of 14 classification parameters. These quantities comprise the input vector for the neural network classifiers to be discussed in Sec. 3.

Initial samples of galaxy and star images were collected from two regions of equal area on P323 (the POSS plate containing the Coma cluster of galaxies). As a final check, a smaller area which extends from the center of the plate to the edge was also surveyed and used purely as a source of test data for the image classification software. A map of the objects extracted from the selected regions is shown in Fig. 4. Region 1 (lower right) was centered on the core of the Coma cluster. Galaxies were collected in this portion by cross-matching APS image coordinates with positions of Coma galaxies in the deep catalog of Godwin *et al.* (1983), hereafter referred to as GMP83. A linear coordinate transformation using 30 galaxies cross matched with the GMP83 catalog produced positional scatter values of $\sigma_x = 1.25''$ and $\sigma_y = 1.19''$. Transformed GMP83 galaxy positions were used to search the APS O plate image lists using a search radius of 4″. In this initial study we have restricted our sample to a minimum APS image diameter of $D = 40\,\mu$m (corresponding to $B = 21$ for a stellar image) and galaxies brighter

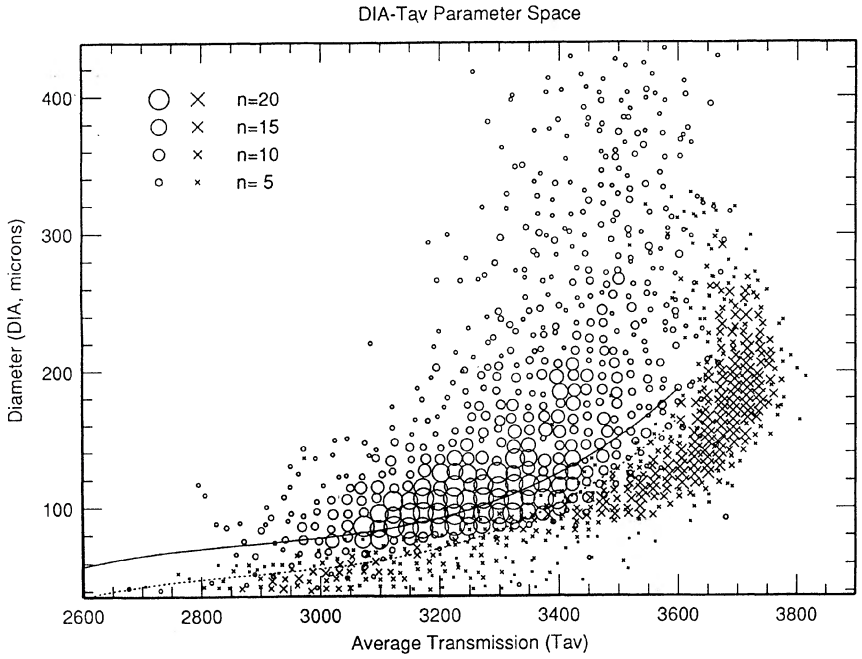

FIG. 1. Distribution of galaxies (open circles) and stars (crosses) in the diameter (DIA) vs average transmission (T_{av}) parameter plane. The point size scales linearly with the number of objects occupying that position in parameter space. Notice that galaxies generally possess a larger diameter than stars for a given average transmission value. The solid curve may be used to select galaxies contaminating a star sample. All objects lying above this line have a very high probability of being a nonstellar source. The dashed line is used in a similar fashion to select contaminating stellar objects in a nonstellar sample.

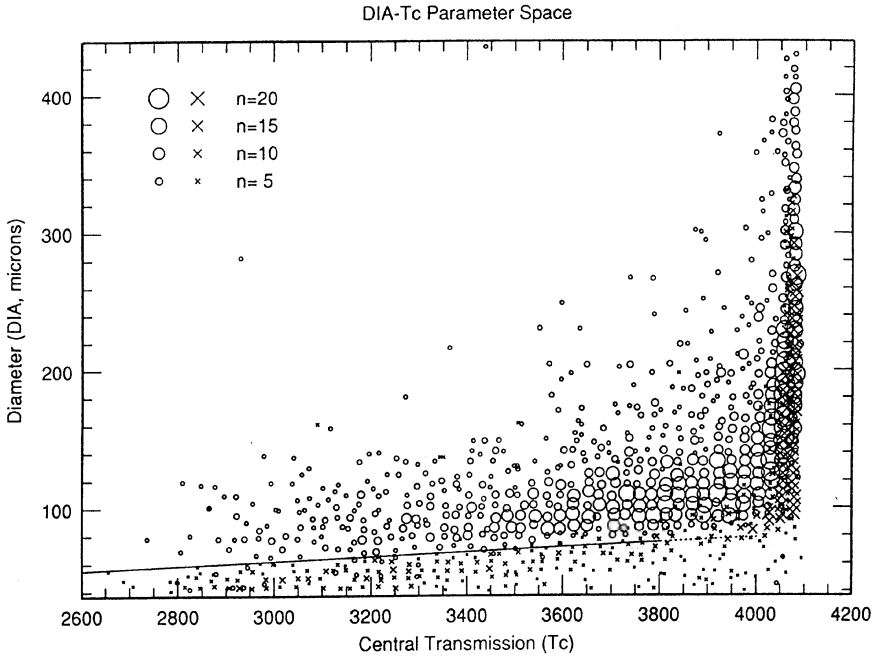

FIG. 2. Distribution of galaxies (open circles) and stars (crosses) in the diameter (DIA) vs central transmission (T_c) parameter plane. Just as in Fig. 1 we see that this cut in parameter space produces a clear segregation between stellar and nonstellar images. The heavy density of points at the high transmission end of the plot is caused by photographic saturation in the central image regions. As we show in our discussion of classifier robustness, this parameter space is quite useful for classifying small, faint images.

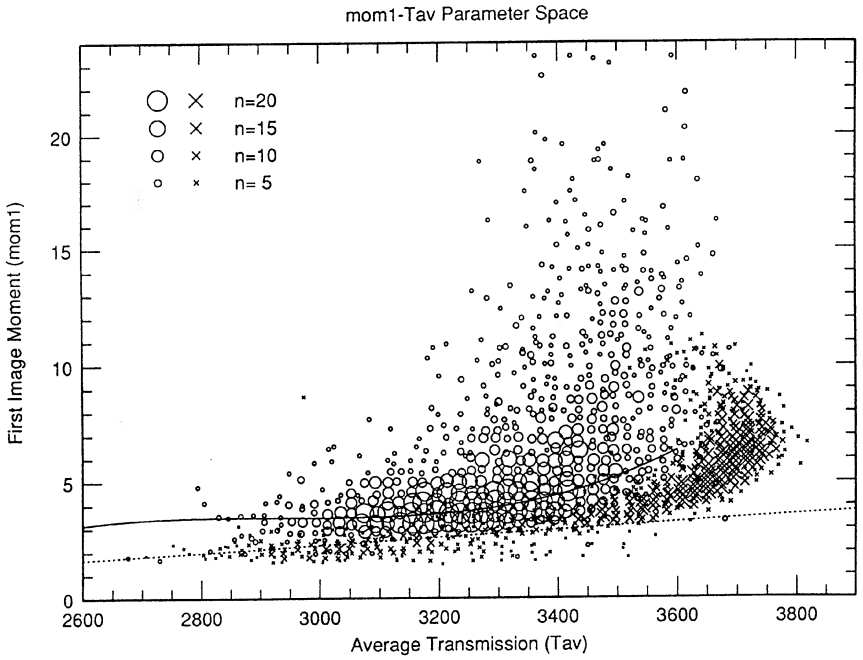

FIG. 3. Distribution of galaxies (open circles) and stars (crosses) in the first image moment (*mom*1) vs average transmission (T_{av}) parameter plane. The more extended (larger moment) galaxy images lie above stars at a fixed average transmission.

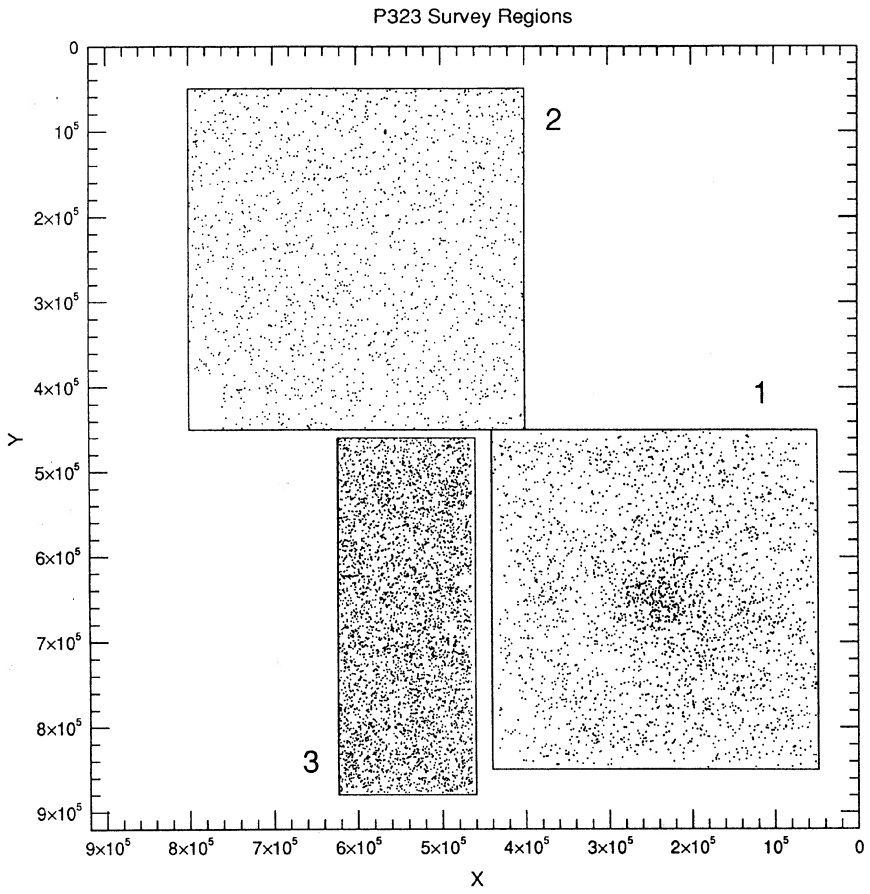

FIG. 4. Three regions surveyed on the P323 POSS field. Galaxies and stars from regions 1 and 2 were used as training data for the neural network classifiers. Data from region 3 were used as a test sample for the adopted classifiers.

than $B_{26.5} \leqslant 20.0$, where $B_{26.5}$ is the isophotal B-band magnitude from GMP83 which measures the flux integrated within the $\mu_B = 26.5\,B$ mag arcsec^{-2} isophote. A sample of stellar objects was selected in this region by locating objects not matched to the GMP83 catalog and obeying the diameter restriction. This provided an initial list of approximately 2400 galaxies ($B \leqslant 20$) and 750 stars ($21 \leqslant B \leqslant 11.75$).

We have also surveyed region 2 (upper left in Fig. 4) by visual inspection of the detected images. Incomplete samples of objects in progressively smaller image diameter intervals (down to $D = 40\,\mu$m) were classified as galaxies or stars by visual inspection. This produced a list of 290 galaxies ($B \leqslant 20$) and 1350 stars ($21 \leqslant B \leqslant 11$). The last area surveyed, which is used as test data only, is referred to as region 3. In this case, we visually classified all images in a $94' \times 205'$ rectangular box extending roughly from the plate center with diameters greater than 73 μm (corresponding to approximate B-band magnitudes of 19.5 for stars and 20.1 for galaxies). In this case, a total of 4135 detected objects was classified into five major categories: galaxies, stars, merged stellar images, plate defects, and uncertain types. The last category pertains mainly to objects with diameters less than 100 μm where definitive classification will require deeper, higher-resolution image material. For this last survey, all classifications were done on the glass copy plate using an $8\times$ loop magnifier for images with diameters larger than 130 μm, and a variable magnification binocular microscope for all smaller images. It should be noted that the thin rectangular shape for this region was chosen for the purpose of sampling regions at different radial distances from the plate center. Optical vignetting and coma are clearly present on the POSS plates, and these effects must to some degree degrade the effectiveness of any image classifier. By sampling at different radii, we hoped to analyze the degree to which our classification method is affected by such distortions. The resultant sample was comprised of 2380 stars and 936 galaxies to be used exclusively for testing the automated image classifiers developed in this work.

The two primary samples (regions 1 and 2) were cleaned of plate defects (and merged stellar images noted) using an iterative process. A variety of plots in image parameter space (see Figs. 1–3) were made for both the star and galaxy samples. Based on the apparent galaxy-star parameter space divisions, a set of conservative selection laws (see the curves plotted in Figs. 1–3) was used to locate candidate contaminating points. Each candidate was visually inspected on the glass copy plate and excluded if found to be incorrectly classified. In a final phase of this process, objects classified incorrectly by the neural network classifiers were inspected visually. A handful of faint contaminating objects was rejected. The majority of these objects, many of which are also marked as uncertain in the GMP83 catalog, was retained in the final sample. It should be stressed that no automatic rejection criteria were applied to the sample, and no image was

rejected from its respective sample without careful visual inspection.

The region 3 images were handled in a somewhat different fashion. Since we wished to determine the survey completeness produced in our scanning process, an effort was made to classify every image detected in the region. In addition to noting stars and galaxies, images were classified as merged objects, plate defects, and uncertain objects. For testing the effectiveness of the image classifiers, only well-classified star and galaxy images were used. Methods for the analysis of merged and defective images will be the subject of a future paper.

3. A NEURAL NETWORK APPROACH

Neural networks are a family of artificial intelligence techniques that are capable of performing difficult pattern classification tasks. Their design and development have been inspired by biological neural networks, but the algorithms we have used do not accurately model real biological systems. Many introductory references on neural networks are available (Knight 1990; Lippmann 1987; Lippmann 1989; Caudill 1987; McClelland & Rumelhart 1988; Rumelhart & McClelland 1988). Angel *et al.* (1990) and Sandler *et al.* (1991) present examples of the application of neural networks to adaptive optics. For information on perceptrons, see Duda & Hart (1973) and Tou & Gonzalez (1974).

We have used the perceptron and backpropagation neural network algorithms to create accurate classifiers for separating star and galaxy images and to inspect their parameter spaces. Backpropagation and perceptron networks are supervised learning techniques. The networks start from a random initial state. A set of training patterns is used to "teach" the network to perform the desired classification function. The training set must contain a representative sample of patterns for each class. Starting from a random initial configuration, the network is used to classify each pattern in the training set. Each time a pattern in the training set is misclassified, an error term is computed and used to adjust the network's configuration. The patterns in the training set are repeatedly presented to the classifier in this manner until the entire training set is correctly classified or until the network is unable to learn any more patterns (Fig. 5).

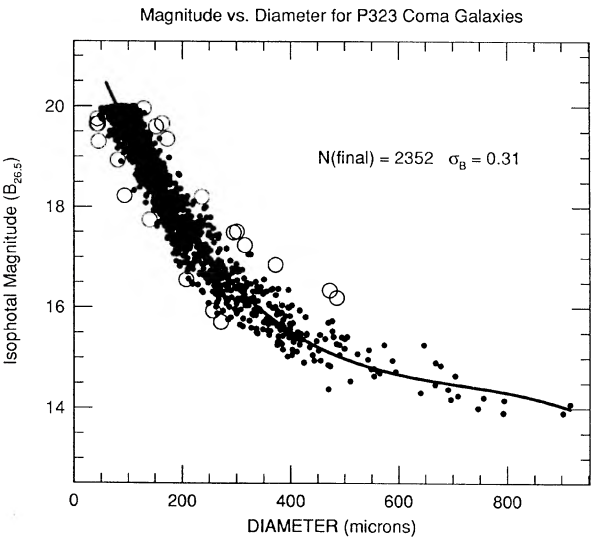

FIG. 5. Magnitude-diameter relation for galaxies on P323 derived using image diameters measured with the APS and isophotal *B*-band magnitude from GMP77. A scatter of $\sigma_m = 0.33$ mag is obtained. Open circles represent points rejected in the fit of a 4-term polynomial.

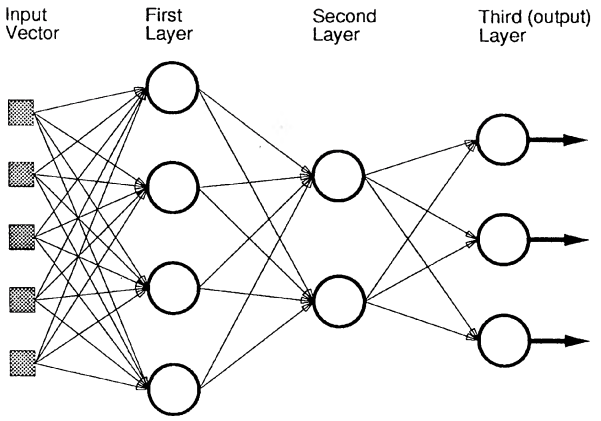

FIG. 6. Schematic illustration of a network with an input vector of length five, four nodes in the first layer, two nodes in the second layer, and three in the output layer. As a shorthand, such a network can be written as (5:4,2,3).

A perceptron is a simple classifier that forms a hyperplane in parameter space to separate the two classes. It is trained using a simple gradient descent procedure to minimize an error criteria function. The *Perceptron Convergence Theorem* (Duda & Hart 1973) guarantees that training will converge to a solution vector if the classes are linearly separable.

Backpropagation networks are capable of learning much more complex functions and are not restricted to linearly separable classification problems. A backpropagation network consists of one or more layers of nodes (neurons). Each node computes the following function:[1]

$$\text{output} = 1/[1 + e^{-(\mathbf{i}\cdot\mathbf{w} + b)}],$$

where **i** is the node's input. This is either the network's input vector or the output from the previous layer. The vector **w** and *b* are the weight vector and bias term and are unique for each node. Initially they contain random values. The weight and bias values are modified by the training process according to the generalized delta rule (Rumelhart & McClelland 1988).

Nodes are arranged in layers. The number of nodes and layers in a network determines the complexity of the function that it computes, as well as the amount of information that it contains. We have experimented with networks consisting of two and three layers of nodes. Figure 6 illustrates a network with an input vector of length five, four nodes in the first layer, two nodes in the second layer, and three in the output layer. As a shorthand, such a network can be written as {5:4,2,3}. To use a backpropagation network as a classi-

[1]The function listed is what is most commonly used. Other functions with a similar shape such as $x/(c + |x|)$ may also be used (Stockwell 1991; Stornetta & Huberman 1987).

TABLE 2. Neural network training parameters.

Network	Diameter Range (microns)	Patterns in Training Set			Training Passes
		Galaxies	Stars	Total	
L1	$d > 137.3$	1116	1050	2166	925
LP	$d > 137.3$	1116	1050	2166	50000
LR	$d > 137.3$	300	300	600	834
S1	$146.5 > d > 73.2$	1584	719	2303	6540
SP	$146.5 > d > 73.2$	1584	719	2303	50000
SR	$146.5 > d > 73.2$	300	300	600	3873
SC	$146.5 > d > 73.2$	1584	719	2303	218

fier, each class is assigned to an output node. The node with the greatest output value determines the classification of the input pattern. While a simple perceptron classifier is restricted to two classes, a backpropagation network can be used for an arbitrary number of classes.

Several different networks were generated. We have trained a perceptron for a small diameter regime (73 μm$\leqslant D\leqslant$137 μm), referred to as SP; and one for a large diameter regime (146 μm$\leqslant D\leqslant$330 μm), referred to as LP. Using the same diameter regimes, we have trained {14:14,13,2} backpropagation networks which are referred to as S1 for the small regime, and L1 for the large regime. These are the primary networks discussed in this paper. Additional networks were also trained to investigate specific features. Networks SR (small reduced) and LR (large reduced) were trained with smaller training sets to explore the number of prototypes necessary to produce an accurate classifier. A backpropagation network which was allowed fewer training passes and was used to experiment with using the network's output values as a measure of the confidence of image classification is referred to as SC. The training sets and network configuration for the classifiers are summarized in Table 2.

All of the image parameters were scaled and translated to fall within the range from 0 to 3. This brings the values closer to the range to which the activation function is most sensitive. Such adjustment can be done automatically by the network during training by adjusting the nodes' weight vectors and bias values, but given the small step size required for stable training, it would take a very long time to train a network using raw input values.

Training the perceptron classifiers was relatively simple. Each of the perceptron classifiers were trained by selecting the hyperplane that produced the fewest errors after 50 000 passes through the training data. For both diameter ranges, there was no improvement after 28 000 passes. Training for the small diameter range required 51 min on a Sun 4/380.

The primary backpropagation networks were trained until it appeared that learning had ceased. For well-defined problems such as logical relations, learning is normally pursued until the entire training set is learned. The separation between the star and galaxy classes is less well defined, and the sets are constructed in a subjective manner, requiring that the entire training set be learned would be unrealistic.[2] All backpropagation networks discussed in this paper were trained with a learning rate (η) of 0.01 and a momentum coefficient (α) of 0.9. The primary backpropagation classifier for large diameter images were trained after 925 passes through the large diameter training data. The small image classifier required 6540 passes through the small diameter training data. Training required $\sim$28 min for the large image classifier and three hours and twenty minutes for the small image classifier on a Sun 4/380.

One way to compare the different classification techniques and to investigate parameter space is to examine how many patterns were not learned correctly by the classifier. Another aspect of training that can be explored is the number of training patterns required to produce an accurate classifier.

Table 2 displays the diameter ranges, number of training patterns, and number of training passes used for each of the networks discussed in this work. First, considering the networks trained with the full set of training data (SP,LP,S1,L1), the networks had greater difficulty learning to separate the star and galaxy classes for small diameter images than for large diameter images. This is indicated by the higher rate of misclassified training patterns for both the small diameter backpropagation and perceptron networks. Also, S1 required many more training passes than L1 before it reached its final state.

These results suggest that the parameter space for diameters greater than 137 μm separates well and is nearly linearly separable. Smaller images have a more complex parameter space and a linear classifier is not adequate. This also demonstrates the superior learning capacity of the backpropagation network in comparison to the perceptron algorithm. One possible drawback is that if errors exist in the training set, the backpropagation network is more capable of learning these as well. The perceptron, by virtue of its limitation to simple division of parameter space, is required to learn a more general function.

The most time-consuming part of constructing a neural network classifier is manually classifying the data for the training set. The training set should be as free from misclassifications as possible and span the full range of possibilities. This can involve a great deal of tedious work. To explore how many images were needed to describe the data space, two backpropagation networks (SR,LR) were trained with 300 stars and 300 galaxies randomly selected from the primary training sets for their respective diameter ranges.

Both SR and LR were more successful at learning their training sets than their full dataset counterparts; S1,L1 (see Table 3). This is expected since these networks possessed the same information capacity, while the reduced training set networks had less data to represent. Although the classification success rate of the LR network is not as high as L1, or even LP, it does do quite well (Fig. 8, Table 4). This indicates that the large images (stars brighter than 16.2 mag, galaxies above 18.8 mag) are clearly separated in parameter space and therefore the star and galaxy catalogs should be relatively complete and uncontaminated for this range. In the small diameter range, the backpropagation network trained with a reduced training set (SR) had a very poor success rate for galaxy classification, although its success rate for stars was comparable to the other networks. The image profile for stars should be simple and consistent, with variations due only to noise and optical effects. Thus a small training set should be sufficient for the class of stellar images. Galaxy images are inherently more complex and come in a much wider variety. If a small set of galaxy images is used to

[2]A larger network could probably learn the entire training set, but this would involve essentially memorizing the training data instead of learning generalized functions of the image parameters.

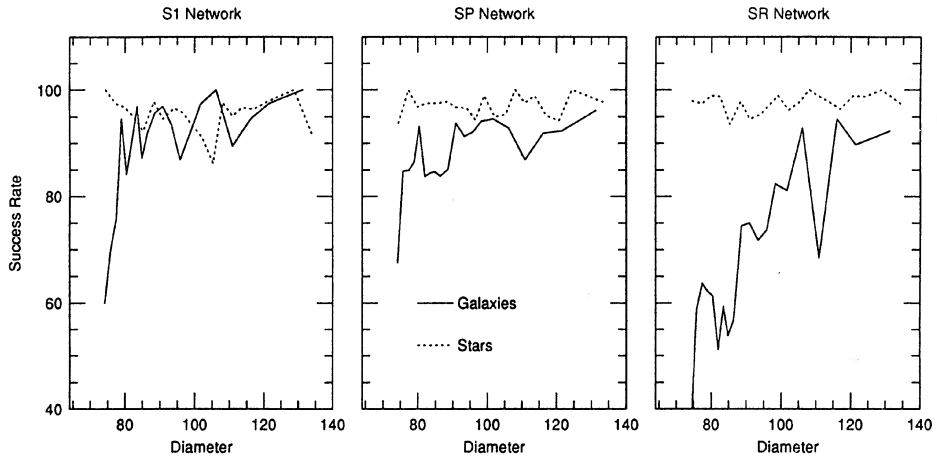

FIG. 7. Success rate functions (number classified correctly/total number) for three different small diameter image classifiers. SP and S1 are perceptron and backpropagation networks, respectively, that were trained on all available data in regions 1 and 2. SR is a backpropagation network trained on only 300 stars and 300 galaxies.

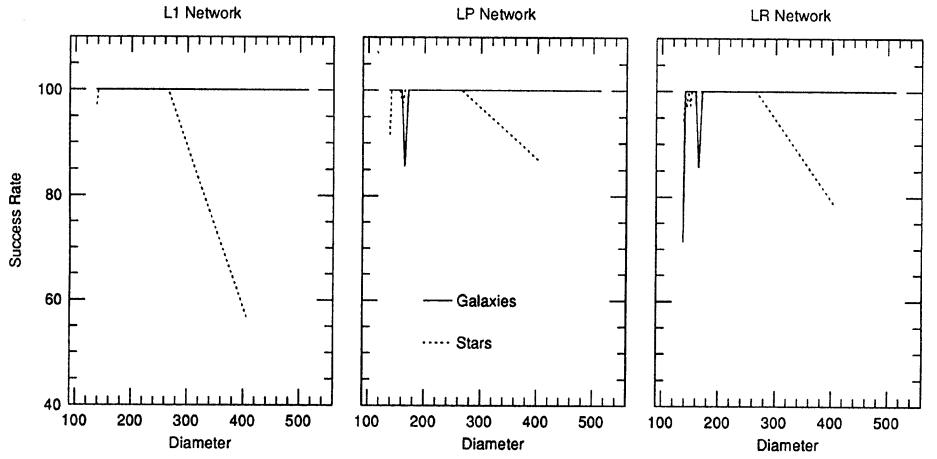

FIG. 8. Success rate functions for large diameter image classifiers. The L1 and LP networks are perceptron and backpropagation networks respectively that were trained on all available data in regions 1 and 2. LR is a backpropagation network trained on only 300 stars and 300 galaxies. The erratic dips in the success rate functions are due primarily to small number statistics in the test sample.

TABLE 3. Neural network learning errors.

Network	Patterns Not Learned					
	Galaxies		Stars		Total	
L1	2	(0.00%)	3	(0.00%)	5	(0.00%)
LP	9	(0.81%)	9	(0.86%)	18	(0.83%)
LR	0	(0.00%)	1	(0.00%)	1	(0.00%)
S1	15	(0.95%)	17	(2.36%)	32	(1.39%)
SP	69	(4.36%)	70	(9.74%)	139	(6.04%)
SR	0	(0.00%)	0	(0.00%)	0	(0.00%)
SC	13	(0.82%)	60	(8.34%)	73	(3.17%)

TABLE 4. Large image classifier success summary.

Stars:	$16.2 \geq O_{mag} > 11.3$
Galaxies:	$18.8 \geq O_{mag} > 16.1$
Sample size:	656 stars, 126 galaxies

Network	Star	Galaxy
L1	99.7%	100.0%
LP	99.4%	99.2%
LR	99.4%	97.6%

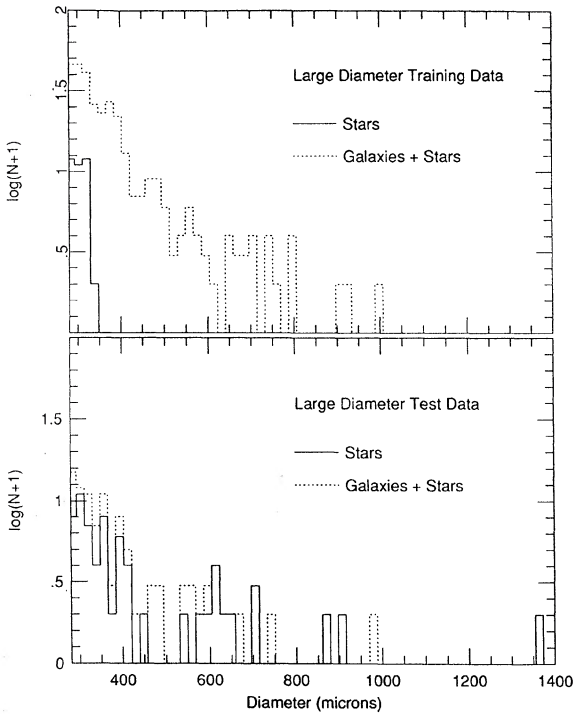

FIG. 9. Diameter distribution for stars (solid line) and galaxies (dashed line) in the training dataset from regions 1 and 2 (top) and the test dataset from region 3 (bottom). Note the lack of very large stars in the training set discussed in the text. This accounts for the poor success rate of the classifiers when handling large stellar images from the test sample due to the poor extrapolative power of neural network classifiers.

train the classifier, then it is probable that the classifier will be faced with many images that are unlike anything found in its training dataset. In such an instance, there is no reason to expect the classifier to produce the correct classification.

The success rate of all of the large image classifiers plummets for stellar images larger than 310 μm. This is easily explained by examining the range of large images in the training dataset (Fig. 9). The largest star in the training set has a diameter of 300 μm, but there are many stars in the test dataset that are larger than this. This shows a deficiency in the training data, and also demonstrates that the networks are limited in their ability to extrapolate too far beyond the training set. It should be noted that around this diameter the images are starting to display more prominent diffraction spikes, and so deviate from the round, compact profile of smaller stars. It would be puzzling if the classifier had been able to correctly classify images with large spikes.

This problem can be addressed in two ways. First, an extensive set of bright star imagery will be collected to extend the training set to images larger than 300 μm. Second, a supplementary approach will be to use existing bright star catalogs, such as the SAO catalog, to directly identify stars with $V \leqslant 10.0$ prior to the classification phase. Additional bright objects can be identified using information from sources such as the RC3 catalog and the data used to produce the Ohio State Overlays (Dixon *et al.* 1981).

The small image classifiers, summarized in Tables 5, 6, and Fig. 7, have generally good performance, but contain a few surprises that deserve comment. Despite the backpropagation network's ability to form a more sophisticated classifier than a perceptron, their success rates are comparable. Also, the networks S1, SP, and SC all have good behavior depending upon what criterion is used. For the diameter range 73 μm$\leqslant D \leqslant$146 μm the backpropagation network has the best average behavior. The SC classifier will produce the most complete sample of galaxy images but it will also be the most heavily contaminated with misclassified stars.

TABLE 5. Intermediate image classifier success summary.

Stars:	$18.1 > O_{mag} > 16.2$
Galaxies:	$19.5 > O_{mag} > 18.8$
Sample size: 811 stars, 224 galaxies	

Network	Star	Galaxy
S1	95.4%	96.9%
SP	97.5%	92.9%
SR	98.4%	87.5%
SC	92.7%	98.7%

The causes for the errors produced by the small image classifiers, particularly the backpropagation networks, are difficult to isolate and must be dealt with using an ad hoc approach. Possible contributing factors to lower success rates may include

(1) Problems with datasets.
  (i) Inadequate training data, especially for borderline images.
  (ii) Classification errors in training dataset.
  (iii) Classification errors in test dataset.
(2) Inadequate image parameters.
  (i) Image parameters do not represent enough pertinent image features.
  (ii) The original image does not contain enough information to make an accurate classification.
(3) Problems in network construction.
  (i) Network is undertrained; necessary rules have not been learned.
  (ii) Network is overtrained; network is too closely bound to training data.
  (iii) Too few nodes; network is unable to represent needed information.
  (iv) Too many nodes; network is not generalizing.

TABLE 6. Small image classifier success summary.

Stars:	$19.5 > O_{mag} \geq 18.1$
Galaxies:	$20.1 > O_{mag} > 19.5$
Sample size: 884 stars, 568 galaxies	

Network	Star	Galaxy
S1	95.8%	86.8%
SP	96.7%	86.4%
SR	96.9%	63.2%
SC	93.8%	93.1%

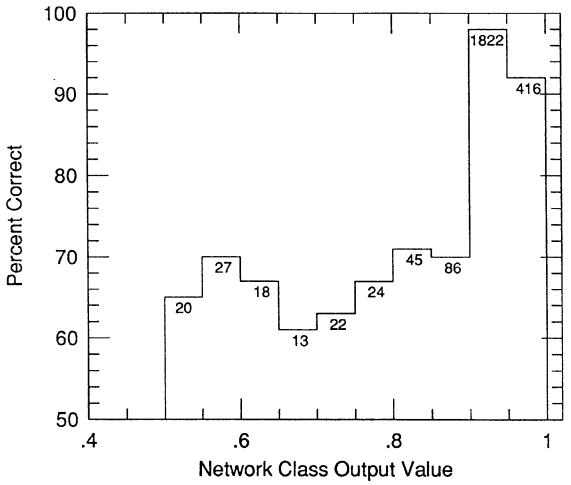

FIG. 10. Plot showing that objects classified with an output value in the range $0.9 \leqslant O \leqslant 1.0$ have fewer misclassifications than objects with outputs in the range $0.6 \leqslant O \leqslant 0.9$. The percentage of successful classifications in plotted as a function of output node value. The total number of objects in each interval is labeled in each bin. The majority of the sample has been classified with output values greater than 0.9.

For the classification results listed so far, the image classification for the backpropagation networks is based on the node with the highest output value. Since the network is trained with idealized target outputs of zero and one, a pattern that produces an output between 0.1 and 0.9 must deviate from the patterns that the classifier learned when it was trained. Although many images are correctly classified with output values that fall in this intermediate range, the frequency of misclassification is higher, as shown in Fig. 10. Although it cannot be used directly as a probability of class membership, it is clear that the network's output value can provide some measure of classification confidence and a means of producing data sets that are relatively free from contamination.

The networks S1 and SC have been used to classify images with the requirement that an image is only given a classification if the output of the maximal output node is greater than or equal to 0.9. Images that fail this requirement are rejected. Results from this test are listed in Tables 7 and 8. The SC classifier has not been as heavily trained as the S1 classifier, and so is less able to accommodate special cases and borderline images. This causes it to reject more images, but also means that it is only able to classify the more "obvious" images. These results are illustrated in Fig. 11. This type of confidence criteria allows the production of very pure samples for studies that can tolerate a bias in the set of image types. For example, a set of confidently classified galaxy images would probably be deficient in bright compact elliptical galaxies.

TABLE 7. Intermediate image summary with confidence criteria.

Stars:	$18.1 > O_{mag} > 16.2$			
Galaxies:	$19.5 > O_{mag} > 18.8$			
Sample size:	811 stars, 224 galaxies			
	Success		Rejected	
Network	Star	Galaxy	Star	Galaxy
S1	98.4%	98.1%	7.2%	5.4%
SC	99.4%	99.5%	17.6%	10.3%

TABLE 8. Small image summary with confidence criteria.

Stars:	$19.5 > O_{mag} \geq 18.1$			
Galaxies:	$20.1 > O_{mag} > 19.5$			
Sample size:	884 stars, 568 galaxies			
	Success		Rejected	
Network	Star	Galaxy	Star	Galaxy
S1	97.5%	91.5%	8.8%	18.8%
SC	99.4%	95.8%	24.3%	41.5%

4. ANALYSIS OF THE CLASSIFIER SUCCESS RATE

In order to test the quality of the final perceptron and backpropagation networks, which were trained using galaxy and star images collected in P323 regions 1 and 2, a third, independent set of test images was established. As described previously, over 4500 images with diameters larger than 73 μm were classified in order to provide an adequately large test sample. The image parameter sets for these images were measured and normalized in the same manner as the training data. Depending on the diameter of a test image an appropriate network was then used to classify the test image. We judge the success of the automated network classifiers by comparing their results to those classifications established by visual inspection. It should be stated explicitly that all visual classifications were made prior to image analysis by the network. In the case of region 3, we found approximately 20 cases in which the network classifier was able to detect plate defects or close image mergers in small diameter objects which had been missed by the human classifier and these objects were reclassified as defects. These corrections represented a statistically small sample, and introduced no significant change in the success rate functions discussed here.

In Table 9 we tabulate the backpropagation classifier percentage success rates, $S =$ (number of success/total number) $\times 100$, for galaxies and stars as a function of integrated magnitude. The magnitudes are derived from mean magnitude-diameter relations established for both stars (Humphreys *et al.* 1991) and galaxies. These relations are rather crude by the standards of present day photometry, but are adequate for the intended purpose of quantifying the network success rate at the faint end of the POSS survey. In Fig. 5 we show the magnitude-diameter relation for galaxies established with the copy of P323 used in this work. This relation is based on the isophotal magnitudes of 2352 galaxies from the GMP83 survey. As can be seen, a considerable scatter is present in the final fit with $\sigma_m = 0.31$ mag. The corresponding scatter in the stellar relation is $\sigma_m = 0.15$ mag. We have transformed diameters measured on the O POSS plate to isophotal B band magnitudes neglecting any

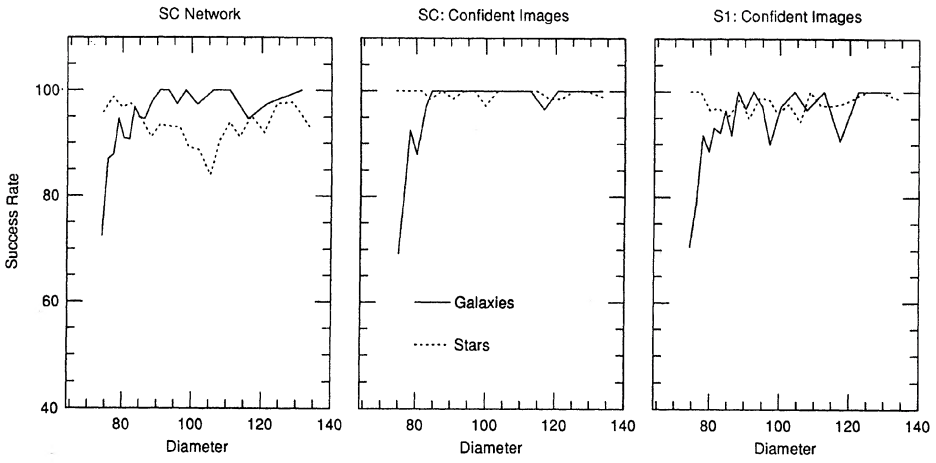

FIG. 11. Success rate functions for a backpropagation network which has been only lightly trained (SC). Restricting the analysis to objects which are classified with a high output node value (>0.9), as in the center panel, produces a high success rate because the sample is limited to 869 easily distinguishable objects. Applying the same restrictions to the fully trained S1 network we get a slightly poorer success rate function, but a larger sample of objects (965) satisfy the confidence criteria.

color correction terms. Strictly speaking, such a relation should also include the galaxy type, T, as a free parameter, since the luminosity distribution laws which determine the integrated magnitude vary systematically along the Hubble sequence. As no revised Hubble types are available for the bulk of the galaxies in our sample, and since the measured scatter in the adopted relation was tolerably small for our purpose, this step and a formal color correction term were neglected.

In Fig. 12 we plot the success rate, S, as a function of the O plate B band magnitude. The curves were established by binning the data in equal number diameter bins containing roughly 30 objects each in the case of galaxies, and 80 objects each in the case of stars. The point symbols represent mean points established in binning intervals of 0.5 mag. The success rate of the backpropagation network classifier remains above 90% until $B \approx 20.0$ for galaxies. The success rate for stars fluctuates in the range between 95% and 100% out to $B \approx 19.5$. Similar results for the perceptron classifiers are given in Table 10 and Fig. 13.

As discussed earlier, radially dependent optical effects such as vignetting and coma will alter image structural parameters (Fig. 13). These distortions can become quite significant near the plate edges, and will certainly degrade the success rate of any image classifier. In order to assess these effects for the backpropagation network discussed here, we plot in Fig. 14 the percentage of image misclassification as a function of radial distance for the plate center. This shows a clear trend that images are misclassified more frequently at large radii from the plate center. For larger images with $D \geqslant 128$ μm ($B \leqslant 16.7$ for stars, $B \leqslant 18.9$ for galaxies) we see that the rate of misclassification remains near 2% until the last bin at the plate edge. Including images down to a diameter of 73 μm, we see a smoother degradation from 4% to 7% misclassification. This effect is non-negligible and its treatment will be discussed in a future paper dealing with param-

TABLE 9. Backpropagation net classification success rates.

		Galaxies				Stars			
m_1	m_2	⟨m⟩	D(μm)	S	n	⟨m⟩	D(μm)	S	n
						14.27	183	100.0	78
14.5	15.0	14.74	572	100.0	5	14.74	170	100.0	115
15.0	15.5	15.17	458	100.0	6	15.25	157	100.0	97
15.5	16.0	15.77	366	100.0	6	15.72	146	100.0	124
16.0	16.5	16.26	311	100.0	6	16.25	136	94.4	125
16.5	17.0	16.73	269	100.0	13	16.76	127	99.5	186
17.0	17.5	17.26	228	100.0	13	17.26	117	96.7	242
17.5	18.0	17.74	196	100.0	23	17.74	107	92.5	280
18.0	18.5	18.23	166	100.0	33	18.26	97	94.9	295
18.5	19.0	18.79	136	100.0	90	18.74	88	95.0	280
19.0	19.5	19.27	112	95.9	173	19.24	79	97.1	307
19.5	20.0	19.79	88	91.4	431	19.52	73	100.0	40
20.0	20.5	20.07	75	72.3	137				

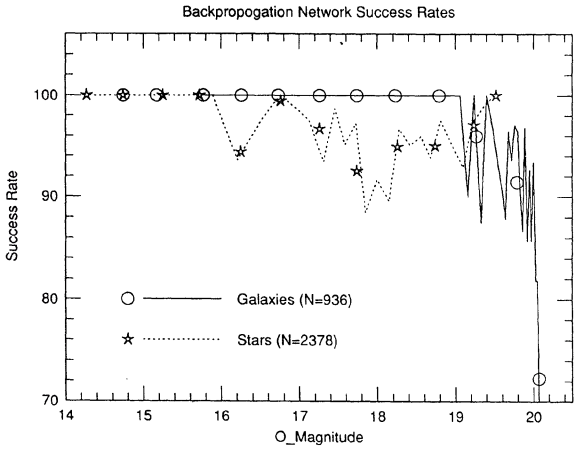

FIG. 12. Classification success rates of the backpropagation network developed in this work. The curves are derived by binning equal numbers of objects into diameter bins. The symbols denote mean success rates in binning intervals of 0.5 mag. Note that both stars and galaxies are classified with success rates above 90% down to the relatively faint O magnitude of 19.5.

TABLE 10. Perceptron net classification success rates.

		Galaxies				Stars			
m_1	m_2	⟨m⟩	D(μm)	S	n	⟨m⟩	D(μm)	S	n
						14.27	183	100.0	78
14.5	15.0	14.74	572	100.0	5	14.74	170	100.0	115
15.0	15.5	15.17	458	100.0	6	15.25	157	99.0	97
15.5	16.0	15.77	366	100.0	6	15.72	146	100.0	124
16.0	16.5	16.26	311	100.0	6	16.25	136	96.8	125
16.5	17.0	16.73	269	100.0	13	16.76	127	98.9	186
17.0	17.5	17.26	228	100.0	13	17.26	117	96.3	242
17.5	18.0	17.74	196	100.0	23	17.74	107	97.9	280
18.0	18.5	18.23	166	97.0	33	18.26	97	96.3	295
18.5	19.0	18.79	136	97.8	90	18.74	88	97.1	280
19.0	19.5	19.27	112	91.9	173	19.24	79	97.1	307
19.5	20.0	19.79	88	88.2	431	19.52	73	95.0	40
20.0	20.5	20.07	75	81.0	137				

eter space normalization schemes designed to remove plate-to-plate variations and radial effects. One avenue being investigated at the present time involves training networks with a radial displacement term added to the input image parameter vector. Given a sufficient number of training images, the neural network might be able to account for radial distortion effects in estimating image class.

It should be noted that while these results are satisfactory, implying that our POSS galaxy catalog will be as deep as the often used Lick survey, it is unclear whether the falloff in galaxy classification success rate at the faint end is due to lack of information in the small image parameter sets, inadequacies of the neural network method when applied to noisy data, or lack of precision by the human classifier in the case of small, faint images. Efforts are under way, using deeper, higher-resolution imagery to differentiate these effects. Even if the success rate function implied in Fig. 12 is dictated by the deepness and resolution of the POSS copy plate material, the fundamental advantage of the automated classification approach is evident when we consider that the human classifications were made over several weeks of tedious work, while the plate digitization, image parameterization, and network classification were carried out in a matter of hours. Additionally, the network classifier will produce a consistently objective classification, and hence a more homogeneous sample of objects.

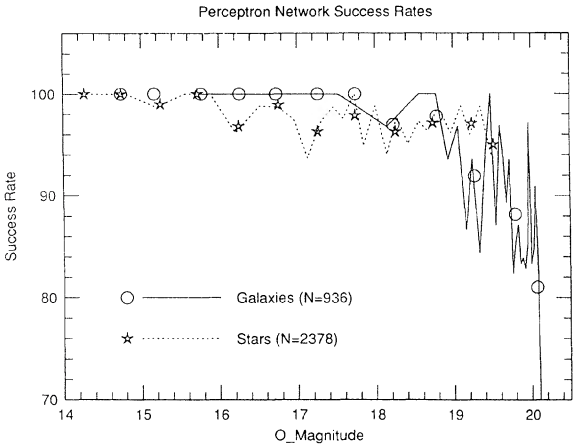

FIG. 13. Classification success rates of the perceptron network developed in this work. The curves are derived by binning equal numbers of objects into diameter bins. The symbols denote mean success rates in binning intervals of 0.5 mag. Note that both stars and galaxies are classified with success rates above 90% down to the relatively faint *O* magnitude of 19.5. The success rates of this simple linear classifier are quite comparable to the results for the more sophisticated backpropagation network of Fig. 12.

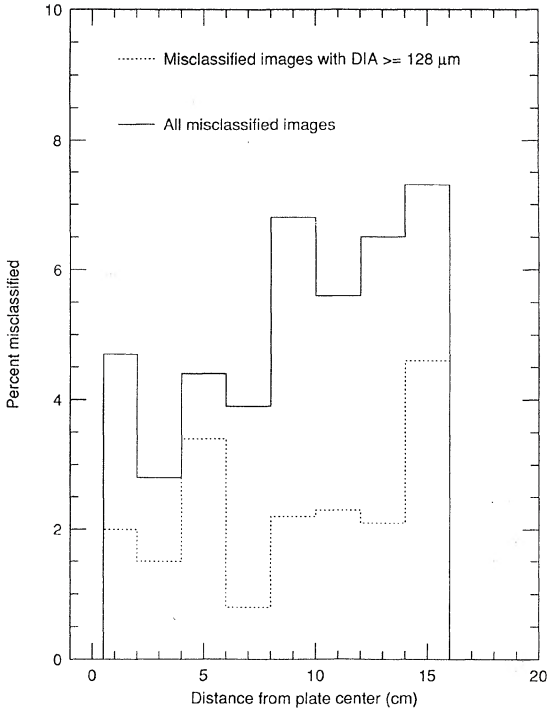

FIG. 14. Classification success rates of the backpropagation network as a function of radial image displacement from the plate center. Changes in structural image parameters due to vignetting and comatic effects degrade the networks' ability to correctly distinguish image class.

One disadvantage of neural network classifiers is that they give no direct information regarding why an object is assigned a particular classification. In an attempt to determine the significance of the various parameters in our 14-element input vector, we have used an empirical approach in which each component of the input vector is distorted in such a way that its information input is nullified. In our case, the network is run on 14 separate test datasets, each set having a percentage level of random noise added to one of its components. A parameter X is adjusted using the relation $X_a = X \pm \beta X$, alternating the sign of the incremental adjustment for every other image. If a parameter carries high weight in the network, then the rate of classification success should be decreased when that information is effectively removed from the calculation. This method is simplistic in its approach in that the network might rely upon interdependencies between several sets of parameters, and removing just one component at a time may not severely cripple the networks ability to make accurate classifications. This is one of the positive features of a neural network which produces a

TABLE 11. Robustness of the large image backpropagation classifier.

	Classification Success Rate for $\beta =$								
parameter	0.00	0.10	0.20	0.50	1.00	2.00	3.00	4.00	O
DIA	100.0	100.0	100.0	100.0	97.8	96.1	92.6	85.7	
ELL	100.0	100.0	100.0	99.6	97.4	90.9	81.4	75.8	
FUZ	100.0	100.0	100.0	99.6	94.8	85.7	81.8	77.1	
JIT	100.0	100.0	99.6	99.1	98.3	94.8	86.6	79.7	
Tav	100.0	99.1	81.8	63.6	51.9	50.2	50.2	50.2	1
Tc	100.0	100.0	100.0	99.1	97.4	91.3	82.7	69.7	
m1	100.0	100.0	100.0	100.0	99.6	99.1	98.3	98.3	
c1	100.0	100.0	100.0	99.1	89.2	69.7	63.2	56.3	
c2	100.0	100.0	99.1	97.8	90.9	71.4	68.8	67.1	
g14	100.0	99.1	97.8	89.6	73.6	75.3	72.7	70.6	3
g13	100.0	100.0	100.0	96.1	84.8	81.0	71.9	56.7	4
g12	100.0	100.0	100.0	100.0	80.5	70.6	67.5	65.8	
g23	100.0	100.0	100.0	99.1	93.5	84.4	81.4	79.7	
g34	100.0	97.8	92.2	64.9	55.8	50.2	49.8	49.8	2

very robust classifier in the presence of observational scatter. This simple empirical method should allow a rough determination of which input parameters are most important to our present networks.

As explained earlier, we have experimented with two networks for two diameter regimes. Using a subset of the region 3 images, normalized parameter sets for a large sample (DIA$\geqslant$146 μm) and a small sample (91 μm$\leqslant$DIA$\leqslant$146 μm) were assembled. The (73 μm$\leqslant$DIA$\leqslant$91 μm) range was excluded from this experiment due to the fact that success rates are low in this regime even when no noise is added. We have run eight sets of calculations, varying the input β value, for each diameter regime. So as not to bias the results of this experiment, we have used approximately equal numbers of stars and galaxies in each diameter regime. Classification success rates decrease with increasing β value. The success rates for these classification tests are given in Table 11 for the large galaxy sample and in Table 12 for the small galaxy sample. In the last column of each of these tables we indicate the order of significance O for the four most important image parameters. A value of O = 1 indicates the parameter which

TABLE 12. Robustness of the small image backpropagation classifier.

	Classification Success Rate for $\beta =$								
parameter	0.00	0.10	0.20	0.50	1.00	2.00	3.00	4.00	O
DIA	95.3	94.4	93.7	90.9	78.7	69.6	64.3	57.2	
ELL	95.3	94.8	94.8	93.5	90.5	84.1	79.9	76.3	
FUZ	95.3	94.8	94.9	93.9	91.7	83.7	78.1	74.0	
JIT	95.3	95.1	95.1	95.1	94.3	92.8	90.1	86.3	
Tav	95.3	86.7	71.6	53.0	49.7	49.6	49.6	49.6	1
Tc	95.3	87.4	75.3	56.9	47.5	49.7	49.6	49.6	2
m1	95.3	94.8	94.5	93.9	94.1	90.5	83.1	75.3	
c1	95.3	94.6	94.1	88.0	72.1	44.9	45.2	48.2	
c2	95.3	93.8	92.0	78.6	56.7	50.6	50.4	50.4	3
g14	95.3	93.7	91.5	77.6	60.3	55.6	52.7	51.9	4
g13	95.3	94.9	94.2	93.2	88.9	73.0	73.8	74.6	
g12	95.3	95.0	95.2	94.5	93.1	80.1	71.7	66.1	
g23	95.3	95.3	95.3	93.7	87.0	71.1	60.4	58.4	
g34	95.3	94.2	92.9	86.3	73.2	64.5	56.2	51.2	

produces the lowest success rate for a given noise level, and hence is apparently the single parameter to which the classifier is most sensitive. The order of significance for each diameter regime was established using $\beta = 0.5$, but varies little over other levels of added noise.

For the large diameter sample no significant degradation of the classification success rate occurs until a 20% noise level is added to the data. The dominant classification parameter is clearly the average transmission, T_{av}. This is due to the fact that stars are saturated over a larger area than galaxies, which tend to produce softer images having less contrast with the plate background (see Fig. 1). As expected from inspection of two-dimensional parameter distributions for samples of galaxies and stars, it is evident that the gradient parameters are excellent discriminators between stellar and nonstellar objects when classifying large images. In Table 11 we see that the three most significant parameters following T_{av} are all gradient terms (g34, g14, and g13). For the small diameter data of Table 12, we see clearly the decreased significance of image gradients. As with large images, T_{av} is the dominant parameter, but the second and third most significant quantities for small images are central transmission T_c and image area c2. As can be seen in Fig. 2, the central transmission value T_c is useful in small images where image saturation is not a problem. Only one gradient term, g14, produces a significantly lower classification success rate at noise levels as high as 50%. As explained earlier, this result was anticipated. As can be seen in Fig. 15, the gradient values become extremely noisy for small image diameters, and hence their usefulness in distinguishing stars and galaxies in small images is reduced.

5. CONCLUSION

A key step in the compilation of an automated survey such as that being conducted with the APS is the discrimination between stellar and nonstellar images. The development of a machine automated technique for performing this task ensures that the tremendous flow of data generated by such a project may be processed in a reasonable period of time. Perhaps more importantly, such a technique produces a very homogeneous catalog whose constituents are selected in an objective and consistent fashion. In this work we have presented a rather novel automated image classification technique employing a neural network which shows great promise. Classifications into stellar and nonstellar categories are based upon a 14-element image parameter set. We have shown that a number of parameter spaces formed with these vector elements are effective in separating a sample of images into the two basic populations of stellar and nonstellar objects. Previous automated surveys have traditionally utilized only a limited number of such parameter space segregations, in limited diameter and brightness regimes, for the problem of galaxy-star discrimination. The application of a neural network to this problem allows a large number of image parameters to be used simultaneously in distinguishing a classification. In this work, a variety of neural networks are developed using two basic algorithms, and a two image diameter, and hence resolution, regimes.

We have experimented with a simple linear neural network classifier known as a perceptron, as well as with a more sophisticated backpropagation neural network with the result that we are able to attain classification success rates of 99% for galaxy images with $B\leqslant 18.5$ and above 95%

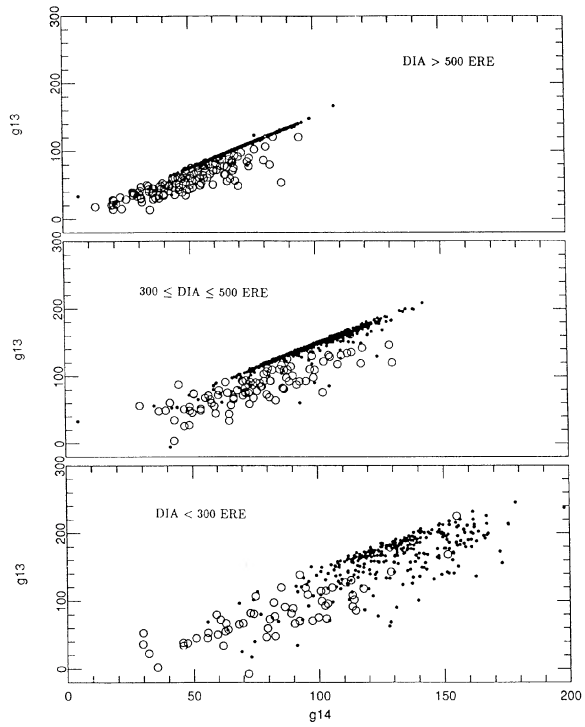

FIG. 15. g13 vs g14 image gradient parameter space for three different diameter regimes. Notice that in the two large diameter samples, in which image resolution is sufficient to allow for the accurate calculation of the gradient parameter, this parameter space provides an excellent discrimination between stars (small solid dots) and galaxies (open circles). For very small image (bottom panel), the small number of image resolution elements makes reliable gradient calculation impossible, thereby making this discrimination method unusable in the small diameter regime.

for the magnitude range $18.5 \leqslant B \leqslant 19.5$. Based on an analysis of 3601 galaxy images and 4460 stellar images on the POSS field containing the Coma cluster, we have determined the success rate of these classifiers as a function of image diameter and integrated magnitude. Simple numerical experiments illustrate the robust nature of this method and identify the most significant image parameters used by the networks in distinguishing image class. The results from these experiments are extremely promising and indicate that the APS survey of the POSS copy plates will reach a limiting magnitude fainter than that of the Lick integrated counts (Shane & Wirtanen 1967), the classic galaxy survey work still employed by many workers for studies of large scale structure in the Universe. The work of Geller *et al.* (1984), in which systematic plate to plate variations in the Lick counts are discussed, points to the great need for a homogeneous survey of the POSS material using objective, consistently applied galaxy selection criteria.

The pattern classifiers discussed in this paper are the first step towards a comprehensive image classification system for the APS catalog of the Palomar Sky Survey. Several problems have yet to be addressed. First, the classification system must be able to be transferred to different fields and cope with plate to plate variations. Also, since a large number of images are required to train and test a supervised learning classifier system, it was not possible to consider image blends and double stars in the data analyzed. Future work will require a plate-to-plate transformation so that a single training set of data can be constructed that contains a large sample of blended images (easily obtained from crowded low galactic latitude plates) as well as galaxies sampled from galaxy-rich plates (those POSS fields containing the Coma, Hercules, and Virgo clusters). Because the APS is the only high-speed scanner in existence which scans two plates at a time, we will investigate using the combined O and E plate image data in performing image discrimination. This will provide more information and detail about each object and therefore greater tolerance of plate noise. Image matching will also be used to reject most scratches and other plate defects.

In the long term, we are particularly interested in exploring the use of neural network classifiers in determining the Hubble type and luminosity class of a galaxy. The newly compiled Third Reference Catalog of Bright Galaxies (de Vaucouleurs *et al.* 1991) is currently being used to compile an extensive library of galaxy images for which existing T and L classifications are available. This library will be used to train and test various neural networks for classifying Hubble types as well as for identifying other basic morphological features such as bars and rings.

We are grateful for the advice and assistance of Dave Zuhn and Rick Pogge with using a variety of development and graphics software. This work was supported in part by a UROP grant from the University of Minnesota Office of Educational Development Programs to E. Stockwell. The operation and maintenance of the Automated Plate Scanner and the production of the APS Catalog of the Palomar Sky Survey is supported by the National Science Foundation under Grant No. AST88-15867.

REFERENCES

Angel, J. R. P., Wizinowich, P., Lloyd-Hart, M., & Sandler, D. 1990, Nat, 348, 221

Caudill, M. 1987, AI Expert, pp. 46–52

Corwin, H. G., de Vaucouleurs, G., & de Vaucouleurs, A. 1985, Southern Galaxy Catalog, University of Texas Astronomy Department

DARPA 1988, DARPA Neural Network Study, AFCEA International Press, Fairfax, Virginia

de Vaucouleurs, G., de Vaucouleurs, A., Corwin, H. G., Buta, R., Paturel, G., & Fouque, P. 1991, The Third Reference Catalog of Bright Galaxies (University of Texas Press, Austin)

Dickey, J. M., Keller, D. T., Pennington, R., & Salpeter, E. E. 1987, AJ, 93, 788

Dixon, R. S., Gearhart, M. R., & Schmidtke, P. C. 1981, Atlas of Sky Overlay Maps for the Palomar Sky Survey (The Ohio State University

Radio Observatory, Columbus, OH)
Duda, R. O., & Hart, P. E. 1973, Pattern Classification and Scene Analysis (Wiley, New York)
Geller, M. J., de Lapparent, V., & Kurtz, M. J. 1984, ApJL, 287, L55
Godwin, J. G., Metcalfe, N., & Peach, J. V. 1983, MNRAS, 202, 113
Heydon-Dumbleton, N. H., Collins, C. A., & MacGillivray, H. T. 1989, MNRAS, 238, 379
Hoffleit, D. 1965, Catalogue of Bright Stars (Yale University, New Haven, Connecticut)
Humphreys, R. M., Landau, R., Ghigo, F. D., & Zumach, W. 1991, AJ (in press)
Humphreys, R. M., & Pennington, R. L. 1989, in Workshop on Digitized Optical Sky Surveys, edited by C. Jaschek and H. T. MacGillivray, pp. 1–9
Knight, K. 1990, Communications of the ACM 33, 58
Kurtz, M. J. 1983, Proceedings of the Statistical Methods in Astronomy Symposium, pp. 47–57
Lippmann, R. P. 1987, IEEE ASSP Magazine 4, 4
Lippmann, R. P. 1989, IEEE Communications Magazine 27(11)
McClelland, J. L., & Rumelhart, D. E. 1988, Explorations in Parallel Distributed Processing (MIT Press, Cambridge, MA)
Nilson, P. 1973, Uppsala General Catalogue of Galaxies (Uppsala Offset Center, Uppsala)
Pennington, R. L., Humphreys, R. M., & Ghigo, F. D. 1987, in Mapping the Sky: Past Heritage and Future Directions, edited by S. Debarbet, J. A. Eddy, H. K. Eichhorn, & A. R. Upgren (Kluwer, Dordrecht), pp. 437–440
Pennington, R. L., Humphreys, R. M., Zumach, W., & Odewahn, S. C. 1991, BAAS, 22, 1325
Rhee, G. 1990, The Structure of Rich Clusters of Galaxies: Clues to Formation and Origin (University of Leiden, Leiden)
Rumelhart, D. E. & McClelland, J. L. 1988, Parallel Distributed Processing (MIT Press, Cambridge, MA), Vol. 1
Sandler, D., Barret, T. K., Palmer, D. A., Fugate, R. Q., & Wild, W. J. 1991, Nat, 351, 300
Shane, C. D., & Wirtanen, C. A. 1967, Publications of Lick Observatory, Technical Report No. 22 (Lick Observatory, Santa Cruz, CA)
Slezak, E., Bijaoui, A., & Mars, G. 1988, A&A, 201, 9
Stockwell, E. B. 1991, A Fast Activation Function for Backpropagation Networks (University of Minnesota Press, Minneapolis)
Stornetta, W. S., & Huberman, B. A. 1987, IEEE International Conference on Neural Networks, Vol. 2
Tou, J. T., & Gonzalez, R. C. 1974, Pattern Recognition Principles (Addison–Wesley, Reading, MA)
Veis, G. 1966, Star Catalog, Smithsonian Institution Publications
Zwicky, F., Herzog, E., Kowal, C. T., Wild, P., & Karpowicz, M. 1961–1968, Catalogue of Galaxies and of Clusters of Galaxies (California Institute of Technology, Pasadena)