

LEARNING SHAPES FOR AUTOMATIC IMAGE SEGMENTATION

Kyoung-Mi Lee

Computer Science Department
The University of Iowa
Iowa City, IA 52242, USA
klee1@cs.uiowa.edu

W. Nick Street

Management Sciences Department
The University of Iowa
Iowa City, IA 52242, USA
nick-street@uiowa.edu

ABSTRACT

The detection and precise segmentation of specific objects is an important task in many computer vision and image analysis problems, particularly in medical domains. Existing methods such as template matching typically require excessive computation and user interaction, particularly if the desired objects have a variety of different shapes. This paper presents a new approach that uses unsupervised learning to find a set of templates specific to the objects being outlined by the user. The templates are formed by averaging the shapes that belong to a particular cluster, and are used to guide an intelligent search through the space of possible objects. This results in decreased time and increased accuracy for repetitive segmentation problems, as system performance improves with continued use. Further, the information gained through clustering and user feedback is used to classify the objects for problems in which shape is relevant to the classification. The effectiveness of the resulting system is demonstrated on a medical diagnosis task using cytological images.

KEYWORDS

generalized Hough transform, learning shapes, averaging shape, clustering shapes, classifying shapes

1. Introduction

Quantitative analysis of digital images requires detection and segmentation of the borders of the objects of interest. In medical images, this segmentation has traditionally been done by human experts. Even with the aid of image processing software, manual segmentation is tedious and time consuming, especially in cases where a large number of objects must be specified. Although many new and sophisticated imaging techniques have been developed to circumvent these problems (Collins & Skorton, 1986; Li et al., 1995), the utility of existing analysis systems is limited by their narrow focus. Single-purpose systems developed for specific applications cannot be reused for other segmentation tasks, despite the fact that the systems frequently consist of the same general steps. While many commercial systems include segmentation capability, they typically rely heavily on user corrections to specify the exact border, a necessary step for precise morphological analysis. Thus the development of techniques to automatically and quickly segment an exact border is an important goal.

Many methods have been proposed to detect and segment 2D shapes, the most common of which is template matching (Duda & Hart, 1972; Nagel & Rosenfeld, 1972). However, its slow speed has prevented its wide spread use. Another technique is based on global feature vectors (Kauk, 1974), but such global techniques are unable to recognize objects that are partially visible. The Hough transform (HT) (Hough, 1962) is useful for shape analysis in noisy, occluded and multiple-object environments. However, its main drawbacks are its substantial computational and storage requirements that become especially acute when object orientation and scale have to be considered. Significant improvements in speed and memory storage are needed for efficient use of the HT for shape and object recognition. The size of the parameter space must be reduced significantly to save storage and to minimize the associated search task.

We propose an on-line learning method that incorporates user feedback to adapt to new shapes as they are outlined. This approach makes our method ideal for highly repetitive tasks such as morphological analysis of cytological and histological images, in which many cells or cell nuclei must be precisely outlined in many different images, and the shapes of the individual cells vary widely from one sample to the next. The remainder of the paper is organized as follows. Section 2 briefly describes the foundation of our system, a combination of template matching and deformable contours. Section 3 describes the proposed algorithm, including the clustering of shapes and the learning of templates.

Experimental results in Section 4 show that the quality of segmentation produced using the proposed algorithm is comparable to an exhaustive template search, and the computation time is significantly reduced. Conclusions are presented in Section 5.

2. Platform

The Xcyt program (Mangasarian et al., 1995; Wolberg et al., 1994) is a graphical computer program for diagnosing breast cancer and predicting the course of the disease. It performs analysis of cytological features based on a digital scan of a breast fine-needle aspirate, diagnosis of the image as benign or malignant along with an estimated probability of malignancy, and a prediction of when the cancer is likely to recur for cancerous samples. The program has proven highly effective in clinical practice, correctly diagnosing 97.6% of new cases since 1993 and providing accurate and individualized prognosis without lymph node information (Street, 2000; Wolberg et al., 1999).

2.1 Object Detection: GHT

Xcyt uses the Generalized Hough Transform (GHT) (Ballard, 1981) to detect the boundaries of the cell nuclei in an image. The GHT is an extension of the Hough transform (HT) (Hough, 1962), a standard template matching algorithm. Generally, an array of cells, $H(x_i, y_i)$, of the same size as the input image, comprises the 2D accumulator, in which each cell has a value specifying the possibility that the reference point of an object shape to be detected is located at the cell. These values are accumulated by having each edge point in the image “vote” for the positions that could correspond to the reference point of the particular template. The GHT constrains the parameter computation by including the orientation of edge points in the discrete representation. When the cell with maximum value in the accumulator exceeds a certain threshold, then an object with the desired shape is said to be detected at the location of the cell. To perform the GHT, the template shapes are built in advance, requiring the user to know exactly what shapes will be encountered. Further, when the scale and orientation of an input shape are variant and unknown in advance, brute force is usually employed to enumerate all possible scales (S_i) and orientation (Θ_i) of the input shape in the GHT process. This adds two dimensions to the parameter space, thus requiring a 4D accumulator, $H(x_i, y_i, S_i, \Theta_i)$. This dramatically increases the execution time and leads to sparsity in the accumulator, making the selection of strong matches more difficult. Iterative approaches to the GHT were proposed in (Lee & Street, 1999) which eliminate the extra dimensions by using a local accumulator (to find each template in turn) and a global accumulator (to collect the best scores from the local accumulator).

2.2 Shape Representation

In order to model a scale- and orientation-invariant shape, we represent it by a set of points (r_i, θ_i) in polar coordinates:

$$\mathbf{v} = (r_0, \theta_0, r_1, \theta_1, r_2, \theta_2, \dots, r_{n-1}, \theta_{n-1}).$$

Assume that OA is an arbitrary radius of the template. In our algorithm, starting from OA and moving clockwise, we divide the circle into n equal arcs to place points around the boundary, with each arc being $\frac{360}{n}$ degrees. So the above

set of points can be represented as

$$\mathbf{v} = (r_0, r_1, r_2, \dots, r_{n-1}),$$

and the corresponding vector \mathbf{p} in x-y coordinates consists of the points

$$\mathbf{p}_i = (x_i, y_i) = (r_i \cos(\frac{360}{n}i), r_i \sin(\frac{360}{n}i))$$

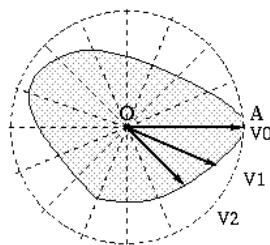


Figure 1 Scale- and orientation-invariant representation

The shape shown in Figure 1 is represented with 16 linked nodes. Translation invariance is achieved through the use of an object-centered coordinate system. Shapes are normalized for size during the algorithm, as described in Section 3.3. In this application we limit ourselves to topologically simple shapes with no holes.

2.3 Segmentation: Snakes

To segment the exact boundaries of cell nuclei, Xcyt uses an adaptive spline curve-fitting technique known as a snake (Kass et al., 1988). Snakes use an energy minimizing spline guided by external constraint forces and influenced by image forces that pull it toward features such as lines and edges. Just like human vision, snakes start with an a priori model of what an object should look like. By using the smoothness constraints of the splines, they are able to fill in missing and noisy boundary information. As a result, snakes have been successful in performing tasks such as edge detection (both actual as well as subjective), corner detection, motion tracking, and stereo matching. In our case, the model is a closed curve that is attracted to strong edges in the image, and forms an arc in the absence of such edge information. The snakes are initialized using the results of an iterative GHT that searches for ellipses of various sizes. Previous work (Lee & Street, 1999; Street, 2000) has shown that the Xcyt system isolates the cell nuclei very well using the combination of GHT and snakes.

Still, the major problem of GHT, as with other template matching approaches, is the need to predefine templates to represent shapes. If precise segmentation is necessary and the objects' shapes vary, many templates are needed, requiring significant search time. In medical domains, the shape of the specific organ varies by person, and shape representation must be very precise in order to quantify morphological features - or, even to initialize more precise methods such as snakes. To overcome this drawback, a learning system is proposed that is able to generalize from segmentation examples.

3. Method

3.1 Initialization

In this application the objects in question are human cell nuclei, which are more or less elliptical in shape. Therefore we begin by creating a set of elliptic templates with different shapes. Figure 2 shows some examples of elliptic boundaries. Since the algorithm is invariant to size and orientation, only one template with the same shape is created. The initial templates are created automatically for this application but could easily be drawn by the user, or avoided completely, without changing the fundamentals of the algorithm.

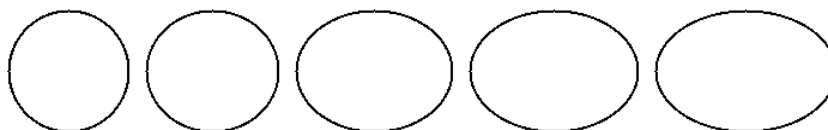


Figure 2 Examples of training templates

3.2 Object Detection and Segmentation

For each new image, the iterative GHT is performed using the existing set of templates. After plateaus in the accumulator are removed with a peak-sharpening step, the highest value in the global accumulator is found. This represents the best match of any template to a shape in the image. The appropriate template is overlaid onto the image at the appropriate rotation angle. This process is repeated until a user-defined number of objects has been located, or until the match scores degrade below an adjustable threshold.

An important aspect of the proposed method is the use of intelligent search to guide the order in which template matches are attempted. As the images are processed, the program records the number of times each template has matched an object. This allows us to search first for the object shapes that were most common in previous images. By searching first for the objects that we are most likely to find, we significantly reduce the expected time required, since not all templates will appear in every image. A user might also choose to search only for objects with a particular classification. This is discussed in Section 3.4.

A snake is initialized for each object using the template points and runs to convergence. The user can then edit the resulting outline by dragging the boundary points to their desired location. The user may also remove an incorrect boundary, or draw a boundary on an undetected object by hand, using the mouse to initialize the snake points. From

the perspective of the learning method, this process creates a collection of positive examples of the shapes that this user would like to find and outline.

3.3 Clustering Shapes

After an object has been correctly outlined, its shape is compared to all existing templates. The difference between a shape, \mathbf{v} , and the template, \mathbf{t} , is computed with the standard Euclidean error metric,

$$e^{\mathbf{v}} = \sum_{i=1}^n |r_i^{\mathbf{v}} - r_i^{\mathbf{t}}|$$

where $r^{\mathbf{v}}$ is a radius in the shape and $r^{\mathbf{t}}$ is the corresponding radius in the template.

To avoid simple orientation differences, the new shape is compared to the template at 16 different rotations. For each rotation angle, the size of the shape is normalized based on the length of the base radial segment. The final difference is defined to be the minimum of these errors.

If the minimum difference between a shape and one of the templates is below a threshold, the shape is added to the cluster represented by that template. This is done using a Point Distribution Model (PDM) (Cootes et al., 1995). A PDM is a statistical shape model which is generated from a set of examples of the objects to be modeled. Given a set of N clustered shapes in template \mathbf{t} , $\mathbf{v}^k = (r_i^k)_{i=1 \wedge n}^{k=1 \wedge N}$, the average shape, $\overline{\mathbf{v}}_i$, is calculated using

$$\mathbf{v}_i = \frac{1}{N+1} (r_i^{\mathbf{t}} + \sum_{i=1}^n r_i^k)$$

If the minimum difference between a shape and one of the templates is greater than the threshold, a new template is created with the shape of the new object. Thus, after the first image, each template represents a cluster of objects that are nearby one another in the space of possible shapes. The algorithm also records the scale of each matching object relative to the template. Figure 3 shows the clustering process graphically. The threshold that determines when a new shape has been found is set empirically, but is again adjustable by the user.

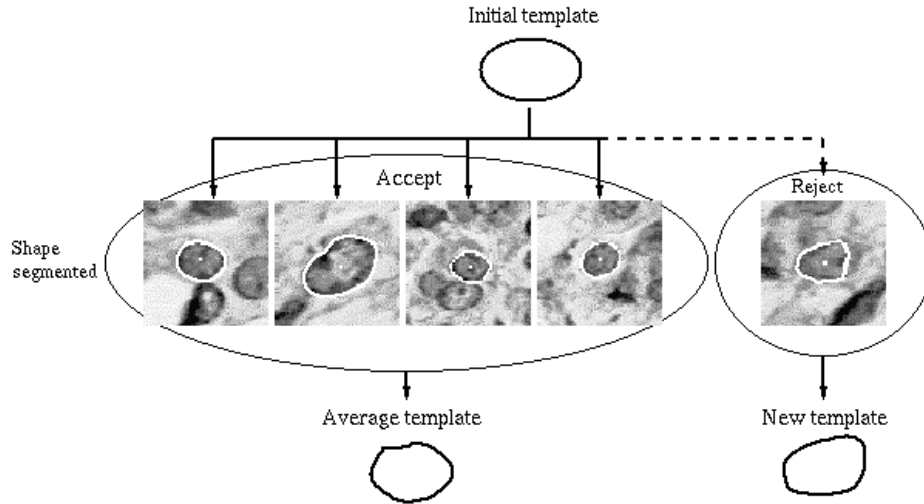


Figure 3 Clustering of shapes

Figure 4 shows the scattering of points of segmented boundaries in the same cluster. It can be seen that some of the points show little derivation over the training set, while others form more diffuse derivation.

3.4 Classifying Templates

The ability to classify objects in an image plays an important role in the projected applications of this system. Consider the problem of isolating cell nuclei from heterogeneous tissue for dissection and genetic analysis. Such tissue may contain both diseased and healthy cells in the same sample. However, the molecular analysis may depend on collecting a "clean" sample of all diseased or all healthy cells for comparison purposes. Further, the user should not be expected to wait for the system to locate objects that are not desired for the particular experiment.

Cellular morphometry is often used to diagnose diseases such as cancer. For instance, the Xcyt system was

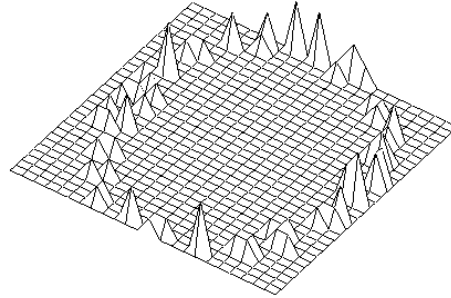


Figure 4 Scatter of points from an aligned set of an elliptic template

originally designed to diagnose breast tissue as benign or malignant based on derived nuclear features such as area, perimeter, and smoothness (Wolberg et al., 1994). We therefore take the approach that size and shape information already gathered in the clusters may be able to solve classification problems.

The classification method proceeds as follows. As objects are isolated by the user, they are given a class label. This requires a training phase in which an expert user is available. A count is maintained of the number of objects from each class that are represented in each cluster. For instance, the upper right template shown in Figure 6 is made up of eight benign and ten malignant cell nuclei. These counts are stored based on the scale factor; for instance, Table 1 shows the scales and counts for the fourth template in Figure 6.

Table 1 Scale and diagnosis for the shapes in template 4

SCALE	0.8	1.0	1.2	1.4	1.6
BEN/MAL Matches	2/0	5/0	0/0	0/1	0/1

As the training proceeds, the system is able to classify new cells based on the majority class of the matching template. This simple instance-based learning scheme further improves the speed of the detection algorithm in cases where only a particular class of object is desired. For instance, if a user wants only malignant cells, the template search procedure will order the templates based on the probability of malignancy, rather than on the raw count of matched cells. Templates that match mostly benign cells will not be used until later in the search, even if their shapes were more common in previous images.

This learning approach was chosen for its simplicity and extensibility. Its primary function in this phase is to guide the template search. However, it is certainly plausible that future applications will require other features, such as chromaticity and second-order size and shape features, to classify the objects sufficiently well. These features can also be collected and stored for each cluster. At that point, a second classification method, such as an artificial neural network, could be added to include these features in the classification.

4. Experimental Results

The algorithm was tested on cytological images from fine-needle biopsies of breast masses, the same images used to train the original Xcyt system¹. These images are gray-scaled with 640×400 pixel spatial resolution.

These images are classified as benign or malignant on a per-sample basis; no classification is available for individual cell nuclei. Therefore in these results we consider all nuclei in benign images to be benign and all nuclei in malignant images to be malignant. This assumption is reasonable but not entirely accurate, making the classification problem particularly difficult due to classification noise.

Figure 6 shows the set of templates constructed after training on six images. The first five shapes are modifications of the predefined templates in Figure 2. The following 17 new templates were created during the experiment. There are two numbers inside each template: the first one is the number of benign cells matching the template, and the second one is the number of malignant cells.

The test proceeded as follows. Two test images (one benign and one malignant) were set aside and used for all tests. The system was trained on a sequence of training images, alternating between benign and malignant, and tested

¹ These images are available at <ftp://dollar.biz.uiowa.edu/pub/street/images/>.

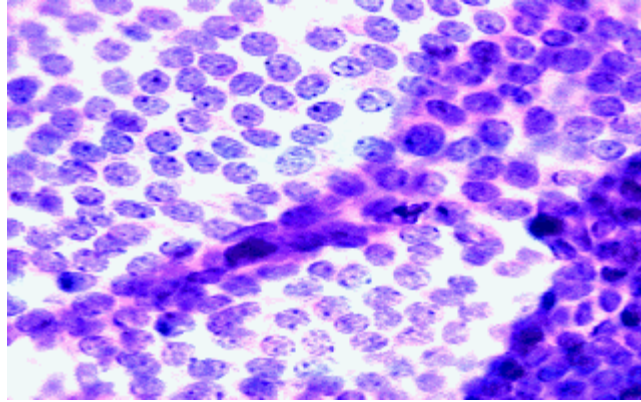


Figure 5 Sample image (reduced by 50%)

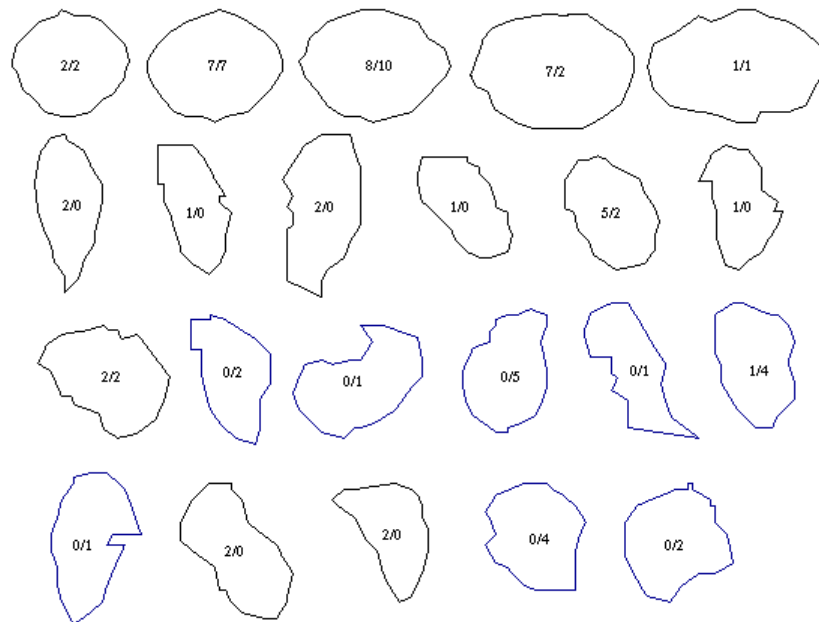


Figure 6 Experimental result : template set

on both test images after each training image.

Figure 7 shows number of templates that had to be searched for the benign (Figure 7(a)) and malignant (Figure 7(b)) test images on a logarithmic scale. The search was carried out in four different ways. The original iterative GHT was performed with a constant set of elliptical templates. These vary significantly in size and shape in an attempt to capture the basic shape of all nuclei that might be encountered. The learning system builds a set of templates as it is used; the curve labeled "All trained templates" shows the number of these learned templates at each point. After an initial "ramping up" period, this number grows very slowly at a level nearly an order of magnitude below the original set of templates. The curve labeled "Best matching templates" shows the result of searching the templates in order of their frequency in previous images, and stopping when a result equivalent in accuracy to searching all the learned templates is reached. This results in a 50% reduction in the number of necessary templates. Finally, the templates were ordered based on their classification, e.g., templates with the highest probability of being benign were searched first on the benign test image. This reduces the necessary search time even further, indicating that the instance-based classification method is able to distinguish reasonably well between nuclei types, even though the ground truth classifications in this problem were extremely noisy. This bodes well for future applications on heterogeneous tissue in which a trained expert will be available for the training phase.

Figure 8 shows the predictive accuracy of the segmentation method on the first 50 outlined nuclei (in the case of the benign image) or on all of the nuclei (in the case of the malignant image). As the system is trained on more and

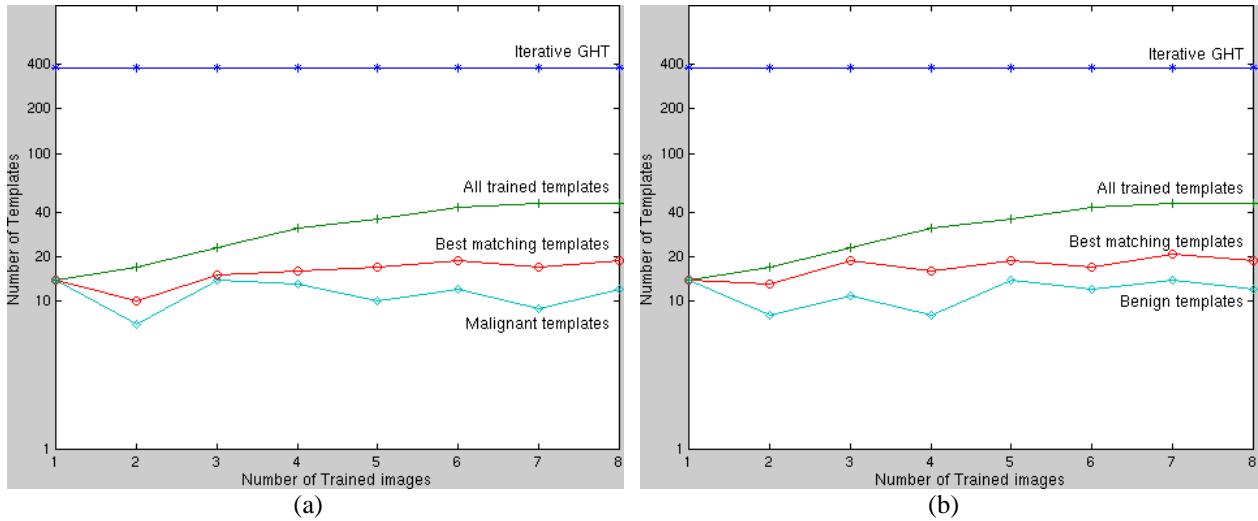


Figure 7 Number of templates searched: (a) Benign image, (b) Malignant image

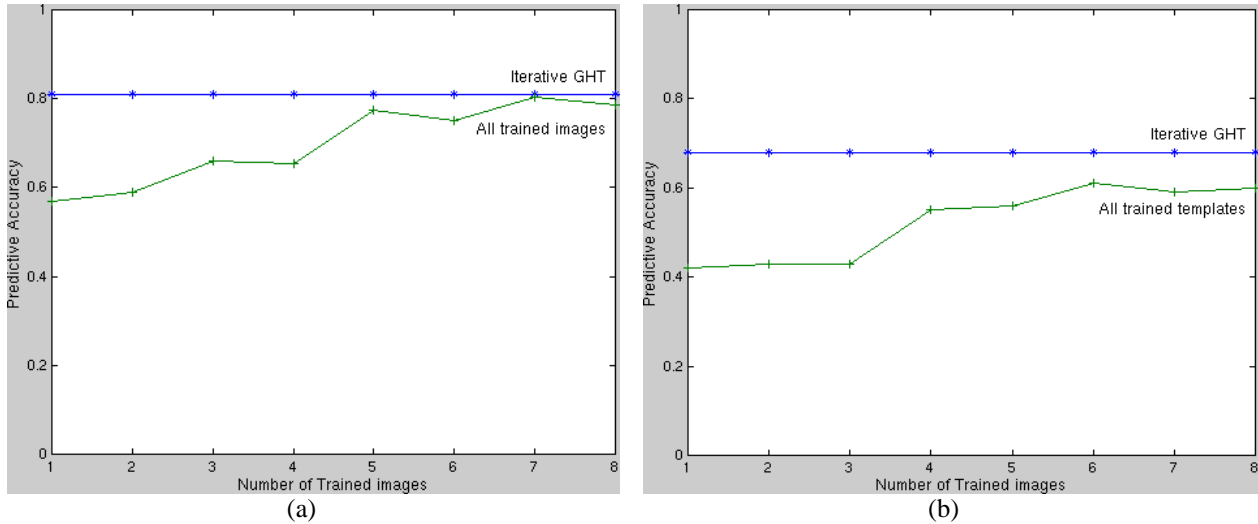


Figure 8 Predictive accuracy: (a) Benign test image, (b) Malignant test image

more images, its ability to correctly segment the nuclei in the test images increases. The accuracy of the new system reaches that of the exhaustive GHT after about seven training images for the benign case. In the malignant case, accuracy does not yet reach that of the original system. We attribute this to the wide variation in shape of malignant nuclei, making the test more dependent on the particular training images. We expect this to be corrected as the system is trained with more images. In short, after a brief training phase, the new system has comparable accuracy to the original system, and achieves this performance in an order of magnitude less time per image.

Figure 9 shows the accuracy of the classification system in these tests. For each well-matched nuclei in a test image, the class is predicted based on the majority class of the template. Poorly-matched nuclei are not included in this computation. The instance-based classifier does very well in general. The apparent fluctuation occurs because the system tends to perform better on the class of objects on which it was just trained.

5. Summary

This paper introduces an incremental learning scheme to aid the location and segmentation of objects in large-scale image analysis tasks. The system demonstrates dramatic improvements in the time needed to perform template-matching search, without sacrificing the quality of the result. Future work will focus on the improvement of the

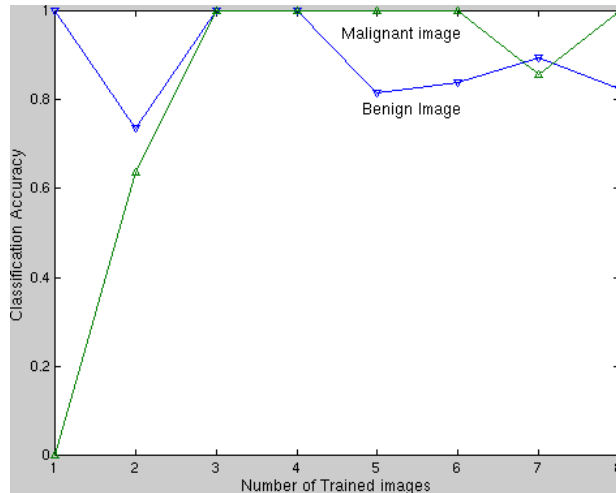


Figure 9 Classification accuracy

instance-based classification algorithm for application to heterogeneous tissue, where the classification of different types of objects is of primary importance.

Acknowledgments

This work was partially supported by NIH grant CA64339-04 and NSF grant IIS-99-96044. The authors would like to thank Dr. William Wolberg of University of Wisconsin Department of Surgery for his helpful suggestions and for providing the test images.

REFERENCES

- Ballard, D. H. (1981). Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2), 111-122.
- Collins, S. M. & Skorton, D. J. (1986). *Cardiac imaging and image processing*. New York: McGraw-Hill.
- Cootes, T. F. Taylor, C. J., Cooper, D. H., & Graham, J. (1995). Active shape models: Their training and application, *Computer Vision and Image Understanding*, 61(1), 38-59.
- Duda, R. O., & Hart, P. E. (1972). Use of the Hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15, 11-15.
- Hough, P. V. C. (1962). Method and means for recognizing complex patterns. U.S. Patent 306954.
- Kass, M. Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 321-331.
- Kaul, X. (1974). Patterns in pattern recognition: 1968-1974. *IEEE Transactions on Information Theory*, 697-722.
- Lee, K. -M., & Street, W. N. (1999). A fast and robust approach for automated segmentation of breast cancer nuclei. *Proceedings of the IASTED International Conference on Computer Graphics and Imaging*, 42-47.
- Li, C., Goldgof, D. B., & Hall, L. O. (1995). Object recognition in brain CT-scans: Knowledge-based fusion of data from multiple feature extractors. *IEEE transactions on Medical Imaging*, 14, 212-229
- Mangasarian, O. L., Street, W. N., & Wolberg, W. H. (1995). Breast cancer diagnosis and prognosis via linear programming. *Operations Research*, 43, 570-576.
- Nagel, R. N., & Rosenfeld, A. (1972). Ordered search techniques in templates matching. *Proceeding IEEE*, 60, 242-244.
- Street, W. N. (2000). Soft computing techniques in breast cancer prognosis and diagnosis. Chapter Xcvt: A system for remote cytological diagnosis and prognosis of breast cancer. CRC Press. In press.
- Wolberg, W. H., Street, W. N. & Mangasarian, O. L. (1994). Machine learning techniques to diagnose breast cancer from image-processed nuclear features of fine needle aspirates. *Cancer Letters*, 77, 163-171.
- Wolberg, W. H., Street, W. N. & Mangasarian, O. L. (1999). A comparison of computer-based nuclear analysis versus lymph node status for staging breast cancer. *Clinical Cancer Research*, 5, 3542-3548.