Breast Cytology Diagnosis Via Digital Image Analysis [1] [2] [3]

William H. Wolberg, M.D. [4]   W. Nick Street, M.S. [5]
Olvi L. Mangasarian, Ph.D. [6]

Originally submitted August 7, 1992
Submitted with revisions April 15, 1993

[1]From the Departments of Surgery, Human Oncology and Computer Sciences, University of Wisconsin, Madison, Wisconsin, U.S.A.

[2]This study was supported in part by Air Force Office of Scientific Research grant AFOSR 89-0410 and National Science Foundation grant CCR-9101801.

[3]Address reprint requests to: William H. Wolberg, M.D., Department of Surgery, University of Wisconsin Clinical Sciences Center, 600 Highland Avenue, Madison, WI 53792.

[4]Dr. Wolberg is Professor of Surgery and of Human Oncology.

[5]Mr. Street is a Ph.D. student and Research Assistant in Computer Sciences. Email: street@cs.wisc.edu

[6]Dr. Mangasarian is Professor of Computer Sciences.

Abstract

An interactive computer system has been developed for evaluating cytologic features derived directly from a digital scan of breast fine needle aspirate (FNA) slides. The system uses computer vision techniques to analyze cell nuclei and classifies them using an inductive method based on linear programming. A digital scan of selected areas of the aspirate slide is done by a trained observer, while the analysis of the digitized image is done by an untrained observer. When trained and tested on 119 breast FNAs (68 benign and 51 malignant) using leave-one-out testing, 90% correctness was achieved. These results indicate that the method is accurate (good intra- and interobserver reproducibility) and that an untrained operator can obtain diagnostic results comparable to those achieved visually by experienced observers.

Abbreviated title: Computerized Breast Cancer Diagnosis

Key words: automation, breast cancer, cytology, FNAB, expert systems, image processing (computer assisted)

Introduction

Most breast masses can be accurately diagnosed from the cytological characteristics of their constituent cells. However, diagnostic accuracy depends on the training, experience, and various undefined factors of the person interpreting the cytology. In short, some people diagnose better than others. We previously developed a computer-based classification system[7,25] which aids in diagnosing breast cells based on the visual assessment of cellular features.[46] Cytological features, previously evaluated visually, are now replaced by digitized ones, evaluated by image analysis. In this paper we compare human accuracy in the cytological diagnosis of breast cancer with the accuracy of digital image analysis combined with computer-based classification.

Materials and Methods

Fine Needle Aspirates

A small drop of viscous fluid was obtained from breast masses by making multiple passes with a 23 gauge needle while negative pressure was being applied to an attached syringe. The aspirated material was expressed onto a silane coated glass slide. A similar slide was placed face down on the aspirate, and the aspirate was spread as the slides were separated with a horizontal motion. Preparations were immediately fixed in 95% ethanol and processed after they were stained with hematoxylin and eosin.

Our database of breast fine needle aspirations (FNA) consists of a personal series (WHW) of 569 consecutive breast aspirates that contained epithelial cells (212 with cancer and 357 with fibrocystic disease). Only palpable breast masses were aspirated and only noncystic ones were analyzed. All cancers were histologically confirmed and all patients with fibrocystic disease were either biopsied or followed for a year without enlargement of the previously aspirated mass. The 119 samples (68 benign and 51 malignant) analyzed in this paper were arbitrarily selected from these 569 images.

The Visual Assessment Process

Machine learning for cytological diagnosis consists of two steps. First we digitally define and extract cellular features from a cytological image of a sample. Second, an expert system diagnoses the sample on the basis of these features. The original version of this work[46] used features which were visually evaluated from breast FNA. One of the authors (WHW) graded each satisfactory aspirate on nine separate features, with each feature being assigned a value from one to ten with higher numbers corresponding to a greater probability of malignancy. The resulting nine-dimensional feature vectors were then used to train an automatic classifier using the multisurface method of pattern separation (MSM).[25] This method uses linear programming[9] to generate separating planes that distinguish benign from malignant samples. The planes were obtained by using the MINOS linear programming package developed at Stanford University.[26] Our original database consisted of 257 benign samples and 180 malignant ones. The expert system was prospectively tested on 261 (199 benign and 62 malignant) new patients with 98% accuracy.[47]

The Automated Assessment Process

Image Preparation

The area on the aspirate slides that was imaged was selected for minimal nuclear overlap. Areas of apocrine metaplasia were avoided. The image for digital analysis was generated by a JVC TK-1070U color video camera mounted atop an Olympus microscope. The slide was projected into the camera with a 63× objective and a 2.5× ocular. The image was captured by a ComputerEyes/RT color frame grabber board (Digital Vision, Inc., Dedham MA 02026) as a 512×480 Targa file. The resulting image is stored in memory as a 2-dimensional array, with each picture element (pixel) having a value between 0 and 255 representing the light intensity at that point.

The User Interface

The first step in successfully analyzing the digital image is to specify the exact location of each cell nucleus. This is achieved through use of an interactive user interface developed using the X Window System and the Athena Widget Set on a DECstation 3100. The image is projected onto a computer screen, and the operator uses a mouse button to trace a rough outline of each visible cell nucleus. These outlines are shown in Figure 1. Rather than requiring an exact boundary specification from the operator, we refine these outlines automatically, as described in the following section.

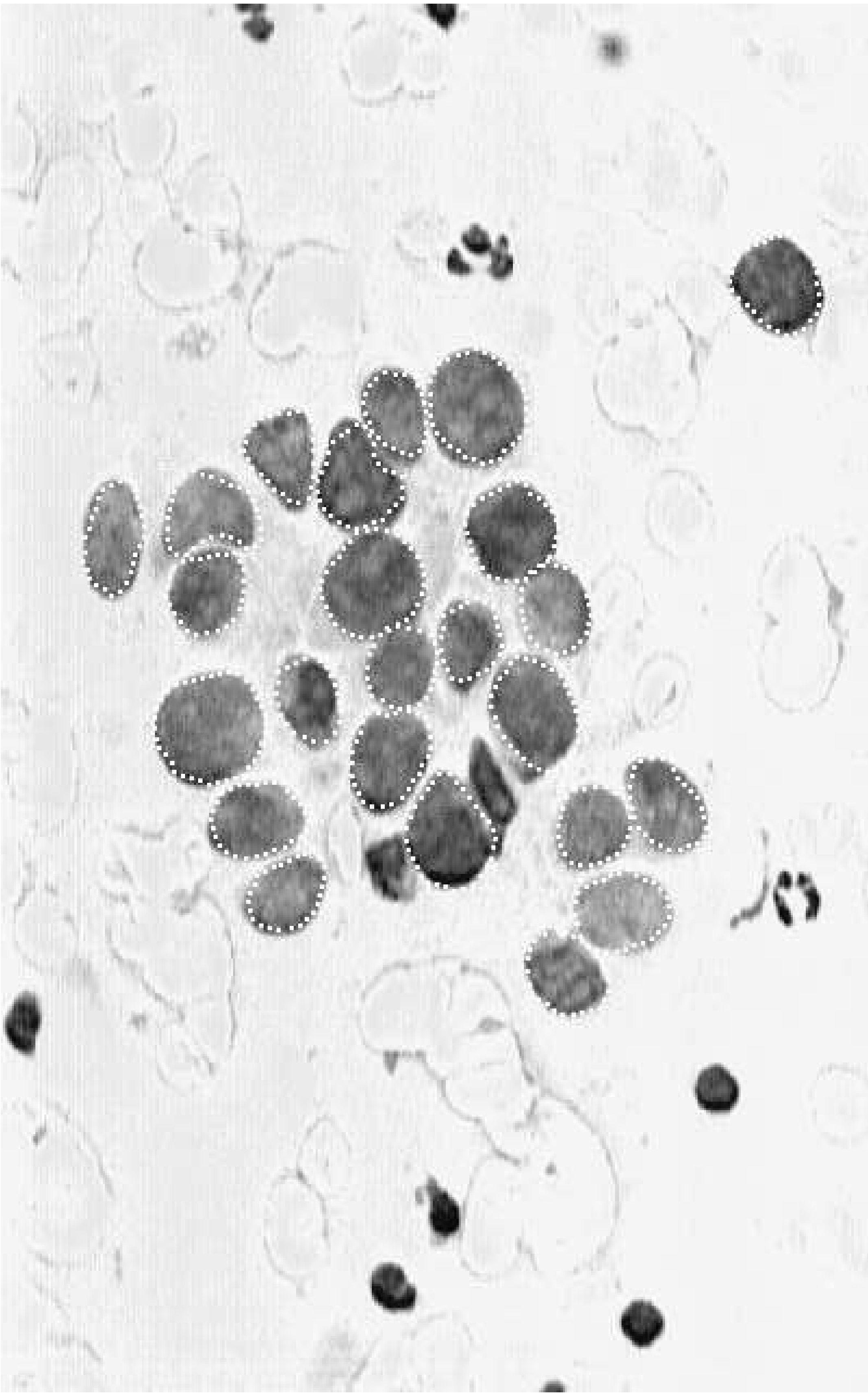

Figure 1: Approximate Positions of Cell Nuclei

Using a mouse pointer, the user first draws a rough outline of each cell nucleus. Inexactness will be corrected in the next phase.

Snakes

Beginning with a user-defined approximate border as an initialization, the precise boundary of the cell nucleus is located by an active contour model known in the Computer Vision literature as a "snake".[20] A snake is a deformable spline which seeks to minimize an energy function defined over the arclength of a curve. We define the energy function in such a way that the snake, in the form of a closed curve, conforms itself to the boundary of a cell nucleus.

To achieve this, the energy function to be minimized is defined as the following function of arclength s:

$$E = \int_s (\alpha E_{cont}(s) + \beta E_{curv}(s) + \gamma E_{image}(s))ds$$

Here E represents the total energy integrated along the arclength s of the snake. The energy computation is the sum of energy terms E_{cont}, E_{curv} and E_{image}, with respective weights α, β and γ controlling their relative importance. To avoid performing the necessary variational calculus, the energy function is computed at a number of discrete points along the curve (as shown in Figure 1), and the integral is approximated with a sum. The three energy terms measure the following quantities:

- Image E_{image}

  This term ties the snake's performance to the underlying image. In our case E_{image} measures the gray-level contrast between pixels inside the snake and those outside the snake. Taking advantage of the fact that cell nuclei are generally darker than the surrounding material, we define this term such that lowest energy is achieved by a sharp, dark-to-light transition. Since light intensity is the most important factor in finding the nuclear boundary, this term is weighted somewhat more heavily than the other two, i.e., $\gamma > \alpha$ and $\gamma > \beta$.

- Curvature E_{curv}

  This geometric term measures discontinuities in the curvature of the snake. Cell nuclei are more or less ellipsoidal; the curvature of the nuclear boundary should be approximately the inverse of the radius, and should change slowly from one part of the boundary to another. Hence, points with abnormally high curvature (sharp corners) or low curvature (straight line segments) are penalized. This term allows the snake to take on a reasonable shape even in regions where gray scale information is lacking, for instance, where cells overlap.

- Continuity E_{cont}

  This term is constructed to penalize discontinuities in the curve. In the discrete case, this term measures how evenly spaced the snake points are. This counteracts the tendency of the snake points to cluster near regions of sharp gray scale discontinuities.

The optimal local value of the energy function is approximated using an algorithm due to Williams and Shah.[44] If the function value at a particular snake point can be lowered by moving the point to an adjacent pixel, then it is moved, thus possibly affecting the energy computation at other points. The process is repeated for each point point, in turn, until all points settle into a local minimum of the energy function.

Since a local or greedy convergence algorithm is used, the snake may converge to a local energy minimum that is a poor representation of the cell nucleus boundary. In this case, the user can use the mouse pointer to delete the snake and initialize a new one in its place. Most of the snakes perform quite well, conforming closely to the nuclear boundaries very quickly. See Figure 2 for an example of an image with all of the nuclei located.

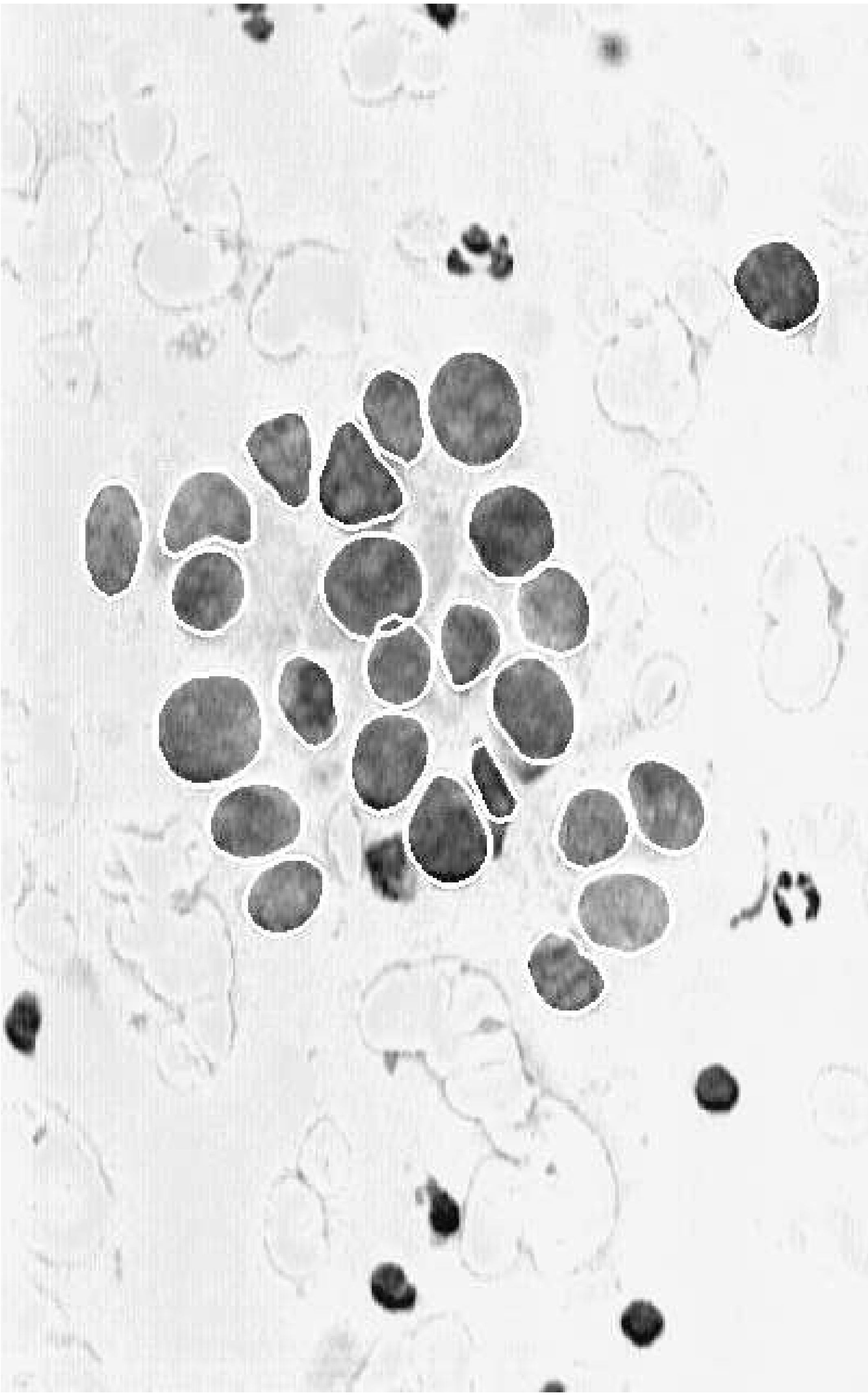

Figure 2: Snakes After Convergence to Cell Nucleus Boundaries

These boundaries are the final representation of the cell nuclei, after the user is satisfied with the convergence of the snakes. A typical image contains between 10 and 40 cells and takes about two to five minutes to process.

Features

The computer vision diagnostic system currently extracts eleven different features from the isolated cell nuclei. As in previous work,[46] all of the features are numerically modeled such that larger values indicate a higher likelihood of malignancy. Unless otherwise noted, the value of a feature for an image is the mean of the values for the individual cell nuclei. The extracted features are as follows:

1. *Size*

   Size of an individual nucleus is measured by averaging the length of the radial line segments defined by the center of the snake and the individual snake points. See Figure 3.

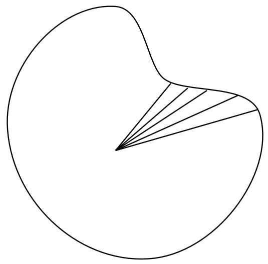

Figure 3: Radial Lines Used In Size and Radial Variance Features

2. *Perimeter*

   Number of pixels making up the boundary of the snake is used to measure perimeter.

3. *Area*

   Nuclear area is measured simply by counting the number of pixels on the interior of the snake and adding one-half of the pixels in the perimeter.

4. *Compactness*

   Perimeter and area are combined to give a measure of the compactness of the cell nuclei using the formula $perimeter^2/area$.[3] This dimensionless number is minimized for a circle and increases with the irregularity of the boundary. However, this measure of shape also increases for nuclei which are elongated, which is not necessarily an indication of malignancy. The feature is also biased upward for small cells because of the decreased accuracy imposed by digitization of the sample. In fact, no single shape measurement seems to capture the idea of "irregular" shape, which is an indicator of malignancy. We compensate for this by employing three different shape features.

5. *Radial Variance*

   The variance of the radial lines used to measure size is computed, giving a different measurement of nuclear irregularity. Again, the value of this parameter increases for elongated as well as oddly shaped nuclei.

6. *Concavity*

   In a further attempt to capture shape information we measure the number and severity of concavities or indentations in a cell nucleus. We draw chords between non-adjacent snake points and measure the extent to which the actual boundary of the nucleus lies on the inside of each chord. See Figure 4. Note that a circular or elliptical nucleus would show no concavity.

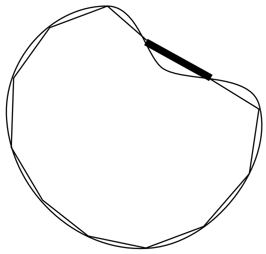

Figure 4: Chords Used In Concavity Feature

7. *Texture*

   The texture of the cell nucleus is measured by finding the variance of the gray scale intensities in the component pixels. This is an approximation of Bland Nuclear Chromatin and Normal Nucleoli parameters described visually.[46]

8. *Uniformity of Size*

   We compute the average difference between the size of each nucleus and the mean size to measure how the nuclear size varies among the cells in an image.

9. *Uniformity of Shape*

Using the Compactness measurement for shape, the average difference is computed as in Uniformity of Size.

10. *Worst Size*

    Along with the average size and the variation of size, we would also expect the maximum size of the nuclei to be a diagnostic indicator.[46] Rather than simply finding the maximum, we average the size of the three largest nuclei in order to reduce the possible effect of one extreme cell, or an operator error.

11. *Worst Shape*

    This is computed in the same manner as the Worst Size feature by averaging the three largest Compactness measurements.

The means and standard deviations of these features for 68 benign and 51 malignant samples are reported in Table I. All of the feature values were linearly scaled to lie in the interval of real numbers one to ten and rounded to the nearest integer. This scaling was performed to facilitate comparison with the visually graded features and to avoid any effects that different scales might have on the classification algorithm. Ideally, each feature would exhibit a significant difference of the means between benign and malignant samples. The value of a feature for malignant samples should be larger (corresponding to the standard of the visual features), with the standard deviation possibly also being larger.

Classification

An automated classification system that is slightly different from MSM was used to test the effectiveness of these features in diagnosing new samples. This classification method,[6,7] known as MSM-Tree (MSM-T), uses a linear programming model to iteratively position separating planes in the 11-dimensional feature space of the examples. If the set of benign points can be separated from the set of malignant points by a single plane, the first plane will be so placed between them. If the sets are not linearly separable, MSM-T will construct a plane which minimizes an average distance of misclassified points. The procedure is recursively repeated on the two regions generated by each plane until each of the final

Feature	Benign μ	Benign σ	Malignant μ	Malignant σ
Size	3.15	0.94	5.37	1.70
Perimeter	3.09	0.94	5.37	1.75
Area	2.24	0.75	4.45	1.83
Compactness	3.16	1.26	3.76	2.11
Radial Variance	4.69	1.72	4.82	1.96
Concavity	1.88	0.68	3.76	1.94
Texture	5.87	1.57	7.29	1.27
Size Uniformity	3.44	1.28	4.90	1.72
Shape Uniformity	3.50	1.25	3.75	1.98
Worst Size	2.75	0.85	4.88	1.55
Worst Shape	4.13	1.82	3.45	1.84

Table I: Means (μ) and Standard Deviations (σ) of the Nuclear Features

All feature values were linearly scaled to the range one to ten.

regions contains mostly points of one category. See Figure 5. The final separation can be thought of as a piecewise-linear fence in the feature space that separates one set from the other.

The prospective accuracy of the resulting classifiers were tested using the Leave-One-Out testing methodology.[24] All but one of the samples are used for training and the classifier is tested on the remaining one. Repeating this procedure for each point in turn and computing the resulting correctness gives an accurate and unbiased estimate of the method's prospective accuracy.

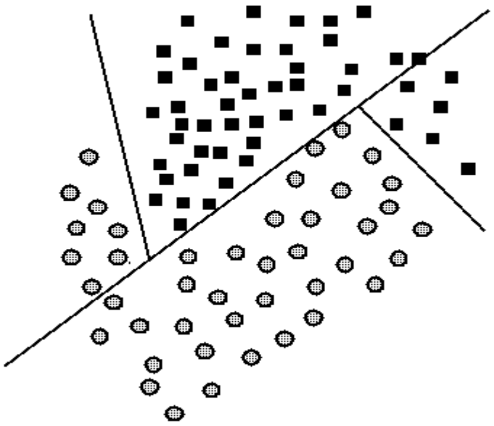

Figure 5: MSM-T Pattern Separation

Using a linear programming approach, a sequence of planes is constructed to (partially) separate the two sets of points. The constructed classifier can then be used to classify new cases.

Results

Diagnostic Accuracy

A summary of results appears in Table II. The first row shows the correctness achieved by visual features extracted from this set of samples. This is the same set of features reported in previous work.[7,46] The second row contains the best results obtained with the automatic feature extraction described in previous sections. To improve generalization, that is, the correct classification of unseen data, the training data was filtered to remove unusually "noisy" points. In this instance, if a malignant point is dominated by a benign point – that is, the benign point has an equal or greater value for *each* feature – the malignant point is not used in training the classifier. Three such points were present in the digital data set, and two of these points were misclassified during the Leave-One-Out evaluation phase. No such points existed in the visually extracted data.

All of the reported results were obtained by generating only a single separating plane in the respective 9- or 11-dimensional spaces (for the visual and digital features, respectively). Figure 6 shows the results

	% Correctness Training Set	% Correctness Test Set
Visually Extracted Features	100	91
Digitally Extracted Features	95	90

Table II: Results of Leave-One-Out Testing

Correctness on training sets (separation of known points) and test sets (generalization to unseen samples) is compared for visual and digital features. Test set correctness represents the best estimate of predictive diagnostic accuracy.

of having the classifier generate additional separating planes. In these cases, the training set can be learned much more exactly, however the classifier is slightly degraded for unseen test points.

Reproducibility

To determine the degree of reproducibility of this computerized analysis a random group of 25 images were analyzed by each of three human subjects. None of the three individuals has any cytological training. Note that the only degrees of freedom between observers are the number of nuclei located and the quality of the cell descriptions (that is, how well the snakes have converged to the cell boundaries).

One subject (S1) analyzed each image three separate times in order to test intraobserver variability. The resulting feature values for each run were compared against one another using a pairwise t-test.[32] The resulting p values were universally very high, indicating that the underlying distributions could well be identical.

Comparisons between different subjects were somewhat more ambiguous. Subject S2 generated consistent results, but subject S3, who performed the analysis on a different computer, generated significantly different results. A statistically significant difference remained when the analysis was repeated on the original computer. However, this variability is largely isolated in the means of the feature value distributions; the variances showed much more consistency. The results of these tests are shown in Tables III , IV and V.

Feature	S1 vs. S1	S1 vs. S2	S1 vs. S3	S2 vs. S3
Size	0.66	0.85	0.51	0.13
Perimeter	0.90	0.01	0.22	< 0.001
Area	0.42	0.58	0.41	0.22
Compactness	0.14	< 0.001	0.43	< 0.001
Radial Variance	0.48	< 0.001	0.57	0.05
Concavity	0.16	< 0.001	0.80	< 0.001
Texture	0.86	0.27	0.65	0.62
Size Uniformity	0.02	< 0.001	< 0.001	0.55
Shape Uniformity	0.16	< 0.001	0.74	< 0.001
Worst Size	0.002	< 0.001	0.11	0.22
Worst Shape	0.10	< 0.001	0.19	< 0.001

Table III: Intra- and Interobserver Variability: p Values of Pairwise t-test

Tests were made to evaluate how reproducible the feature values were, both when re-evaluated by the same user (S1) and when comparing different users (S1, S2 and S3). This most demanding statistical test compares each feature value from every image. The stated p values represent the probability that any observed difference is random error.

Feature	S1 vs. S1	S1 vs. S2	S1 vs. S3	S2 vs. S3
Size	0.97	0.33	0.91	0.77
Perimeter	0.99	0.59	0.84	0.46
Area	0.96	0.91	0.91	0.82
Compactness	0.33	< 0.001	0.69	< 0.001
Radial Variance	0.79	0.003	0.62	0.08
Concavity	0.36	< 0.001	0.85	< 0.001
Texture	0.97	0.81	0.91	0.90
Size Uniformity	0.62	0.11	0.06	0.77
Shape Uniformity	0.37	< 0.001	0.82	< 0.001
Worst Size	0.83	0.26	0.40	0.69
Worst Shape	0.23	< 0.001	0.43	< 0.001

Table IV: Intra- and Interobserver Variability: p Values of Difference-of-Means t-test

This test measures the probability that any difference in distribution means is a result of random error.

Discussion

The goal of our work is to improve the diagnostic accuracy of breast FNAs by using computer based image analysis. Currently, cytologic diagnosis is made by highly trained persons who diagnose based on the subjective interpretation of visually assessed cytologic features. Visual diagnosis of breast FNA has achieved an accuracy of over 90% as reported in a number of studies. For example, the overall accuracy on a 37-series study reported by Frable[11] plus 25 more recent series,[1,4,5,8,10,14–18,21–23,27–29,31,34–37,40–42] with a total of 23,741 satisfactory breast FNAs, was 94.3%. The mean sensitivity was 0.91 ± 0.07 and the mean specificity was 0.87 ± 0.18 for these series taken individually. Giard and Hermans,[12] after doing a literature review of FNA performance parameters, emphasized the need for developing individual

Feature	S1 vs. S1	S1 vs. S2	S1 vs. S3	S2 vs. S3
Size	0.90	0.76	0.98	0.74
Perimeter	0.89	0.92	0.78	0.85
Area	0.86	0.71	0.91	0.79
Compactness	0.91	0.001	0.65	< 0.001
Radial Variance	0.83	0.11	0.002	0.13
Concavity	0.83	0.001	0.04	< 0.001
Texture	0.88	0.53	0.82	0.68
Size Uniformity	0.94	0.25	0.22	0.95
Shape Uniformity	0.80	0.005	0.04	< 0.001
Worst Size	0.84	0.98	0.25	0.24
Worst Shape	0.13	0.003	0.12	0.13

Table V: Intra- and Interobserver Variability: p Values of Difference-of-Variances F-test

Here differences in distribution variances are measured. In those comparisons which show a significant difference, the means appear to be more at fault than the variances.

performance characteristics for centers doing this test.

Other researchers have applied computer based image analysis to various aspects of breast cytology interpretation. Wittekind and Schulte[45] found that mean nuclear area, mean maximum nuclear diameter and mean nuclear perimeter differed significantly between benign and malignant breast cell obtained by FNA. Additionally, they showed that computer-assisted image analysis was superior to eyepiece measurement with respect to accuracy and reproducibility of the results. Studies by others have used either direct scanning of Feulgen stained material[30] or have analyzed digitized images such as we have done.[19, 33, 45] Direct scanning techniques generally analyze only size and texture features whereas analysis of digitized images also allows quantization of shape features. Those who have analyzed digitized images accomplished digitization by manually tracing the nuclear outline on a digitizing tablet. However, in our study, we used a "mouse" to manually outline the nucleus but then relied on our snake program to conform the rough outline to the actual nuclear boundary.

Computerized morphometry has been used in attempts to decrease subjectivity in nuclear grading for prognosis. Those who have digitized manually-traced images by optical scanning have found cytological features such as anisokaryosis, mean nuclear area, mean maximum nuclear area, mean nuclear perimeter, and pleomorphism[2, 30, 38] to correlate with prognosis. These same cytologic features can be used to distinguish benign from malignant. In addition to cellular features, Hutchinson *et al* analyzed clump characteristics but found them to add little to diagnostic accuracy.[19] Their diagnostic accuracy using only cell features was 83.6%, using only smear characteristics was 70.5%, and using both was 86.9%. We achieved better diagnostic accuracy using only cell features.

The analytical methods which we used to project prospective accuracy are familiar to the machine learning community but may be foreign to the medical community. The leave-one-out methodology is a special case of Cross-Validation;[39] the data set is partitioned into some number of testing sets, each of which is tested in turn with classifiers generated from the other samples. This produces an unbiased estimate of the predictive accuracy that is more accurate than simply dividing the data into training and testing sets. In our case, using the standard error from the binomial distribution,[43] we can be confident at the 95% level that the true prospective accuracy - that is, the percentage of unseen cases that would be diagnosed correctly - lies between 84.7% and 95.3%.

By increasing the number of separating planes, any training set can be separated very accurately. However, this reduces the classifier's ability to generalize to unseen cases, as seen in Figure 6. Note that in our case, using only one plane, the training accuracy was 95%, and testing accuracy was 90%. Adding two more planes increases the training accuracy to 98%, but the testing accuracy is degraded to only 88%. This is an example of overtraining; in effect, the peculiarities of training set are being fitted, rather than the true differences between the benign and malignant points. Therefore, only a single separating plane was used in our diagnostic algorithm.

Visual assessments performed by a single highly experienced observer are very accurate. However

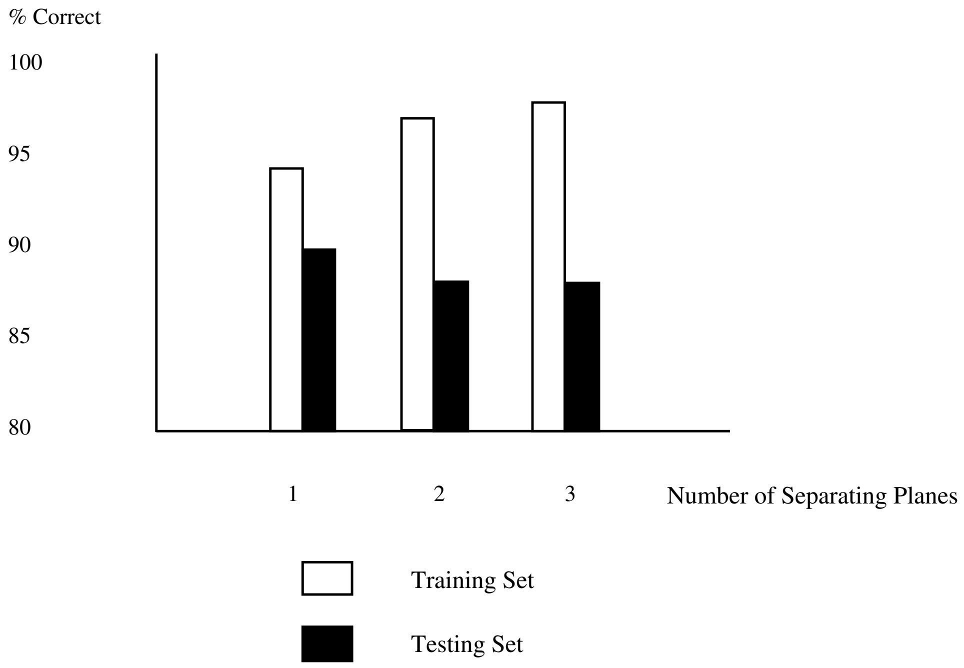

Figure 6: Correctness for Digitally Extracted Features

This figure shows that increasing the number of separating planes actually degrades the generalization performance of the diagnostic classifier. This is an example of *overtraining* or memorization of the training set.

the general applicability of single observer assessments is limited because of their subjectivity.[13] Our computerized system is intended to address the case of multiple and possibly less experienced observers. Still, computerization of cytological analysis as described here requires some degree of operator input so intra- and interobserver reproducibility had to be determined. Our system was intentionally tested under the most adverse of conditions; none of their operators had any histologic or cytologic experience. Intra- and interobserver reproducibility were very good concerning size, area, and texture measurements.

Our data indicate that individual characteristics of the operators are an important factor in interobserver variability. The fact that interobserver differences were concentrated in the means of the distributions rather than the variances indicates that the differences were caused by the choices of objects to analyze. Hence, a classifier trained by a single observer would very likely produce diagnostic results as accurate as ours. A minimal training session, coupled with the increased accuracy of computerized image processing and the robustness of a good machine learning procedure, should reduce or even eliminate the interobserver variability inherent to visual interpretation.

In conclusion, we feel that computerized cytological analysis has significant potential in making objective, accurate and speedy diagnosis available on a wide scale. The present work is a step toward realizing this potential.

Acknowledgements

The authors wish to thank Dr. K. W. Gilchrist for his helpful comments on the form and content of this report.

References

1. Abele JS, Miller TR, Goodson III WH, Hunt TK, Hohn DC: Fine-needle aspiration of palpable breast masses, a program for staged implementation. Arch Surg 118:859–863, 1983

2. Baak JPA, VanDop H, Kurver PHJ, Hermans J: The value of morphometry to classic prognosticators in breast cancer. Cancer 56:374–382, 1985

3. Ballard D, Brown C: Computer Vision, Englewood Cliffs, NJ, Prentice–Hall, Inc, 1982

4. Barrows GH, Anderson TJ, Lamb JL, Dixon JM: Fine-needle aspiration of breast cancer. Relationship of clinical factors to cytology results in 688 primary malignancies. Cancer 58:1493–1498, 1986

5. Bell DA, Hajdu SI, Urban JA, Gaston JP: Role of aspiration cytology in the diagnosis and management of mammary lesions in office practice. Cancer 51:1182–1188, 1983

6. Bennett KP: Decision tree construction via linear programming. *In* Proceedings of the 4th Midwest Artificial Intelligence and Cognitive Science Society Conference. 1992, pp 97–101

7. Bennett KP, Mangasarian OL: Robust linear programming discrimination of two linearly inseparable sets. Optimization Methods and Software 1:23–34, 1992

8. Bjurstam N, Hedberg K, Hultborn KA, Johansson NT, Johnsen C: Diagnosis of breast carcinoma: Evaluation of clinical examination, mammography, thermography and aspiration biopsy in breast disease. Progr Surg 13:1–65, 1974

9. Dantzig GB: Linear Programming and Extensions, Princeton NJ, Princeton University Press, 1963

10. de Ranieri E, Mandard AM, Juret P, Couette JE, Delozier T, Ollivier JM, Leplat G, Rault F: Fiabilite de la ponction cytologique. Qui foit ponctionnner les tumeurs du sein? La Presse Med 12:1527–1529, 1983

11. Frable WJ: Thin-needle aspiration biopsy. *In* Major Problems in Pathology 14. . W.B. Saunders Co., Philadelphia, 1983

12. Giard RWM, Hermans J: The value of aspiration cytologic examination of the breast. A statistical review of the medical literature. Cancer 69:2104–2110, 1992

13. Gilchrist KW, Kalish L, Gould VE, Hirschl S, Imbriglia JE, Levy WM, Patchefsky AS, Penner DW, Pickren J, Roth JA, Schinella RA, Schwartz IS, Wheeler JE: Interobserver reproducibility of histopathological features in stage II breast cancer. Breast Cancer Res Treat 5:3–10, 1985

14. Griffith CN, Kern WH, Mikkelsen WP: Needle aspiration cytologic examination in the management of suspicious lesions of the breast. Surg Gynecol Obstet 162:142–144, 1986

15. Gupta RK, Dowle CS, Simpson JS: The value of needle aspiration cytology of the breast, with an emphasis on the diagnosis of breast disease in young women below the age of 30. Acta Cytologica 34:165–168, 1990

16. Halevy A, Reif R, Bogolovsky H, Orda R: Diagnosis of carcinoma of the breast by fine needle aspiration cytology. Surg Gynecol Obstet 164:506–508, 1987

17. Hammond S, Keyhani-Rofagha S, O'Toole RV: Statistical analysis of fine needle aspiration cytology of the breast. Acta Cytologica 31:276–280, 1987

18. Horgan PG, Waldron D, Mooney E, O'Brien D, McGuire M, Given HF: The role of aspiration cytologic examination in the diagnosis of carcinoma of the breast. Surg Gynecol Obstet 172:290–292, 1991

19. Hutchinson ML, Isenstein LM, Zahniser DJ: High-resolution and contextual analysis for the diagnosis of fine needle aspirates of breast. Analyt Quant Cytol Histol 13:351–355, 1991

20. Kass M, Witkin A, Terzopoulos D: Snakes: Active contour models. International Journal of Computer Vision 1(4):321–331, 1988

21. Kline T: Handbook of Fine Needle Aspiration Biopsy Cytology, St. Louis, Mosby, 1981

22. Knight DC, Lowell DM, Heimann A, Dunn E: Aspiration of the breast and nipple discharge cytology. Surg Gynecol Obstet 163:415–420, 1986

23. Koivuniemi AP: Fine-needle aspiration biopsy of the breast. Ann Clin Research 8:272–283, 1976

24. Lachenbruch P, Mickey P: Estimation of error rates in discriminant analysis. Technometrics 10:1–11, 1968

25. Mangasarian OL: Multi-surface method of pattern separation. IEEE Transactions on Information Theory IT-14:801–807, 1968

26. Murtagh BA, Saunders MA: MINOS 5.0 user's guide. Technical Report SOL 83.20 Stanford University, December 1983

27. Nicholson S, Sainsbury JRC, Wadehra V, Needham GK, Farndon JR: Use of fine needle aspiration cytology with immediate reporting in the diagnosis of breast disease. Br J Surg 75:847–850, 1988

28. Painter RW, Clark WEI, Deckers PJ: Negative findings on fine-needle aspiration biopsy of solid breast masses: Patient management. Am J Surg 55:387–390, 1988

29. Palombini L, Fulciniti F, Vetrani A, DeRosa G, Di Benedetto G, Zeppa P, Troncone G: Fine-needle aspiration biopsies of breast masses. A critical analysis of 1956 cases in 8 years (1976-1984). Cancer 61:2273–2277, 1988

30. Pienta KJ, Coffey DS: Correlation of nuclear morphometry with progression of breast cancer. Cancer 68:2012–2016, 1991

31. Pilotti S, Rilke F, Delpiano C, Di Pietro S, Guzzon A: Problems in fine-needle aspiration biopsy cytology of clinically or mammographically uncertain breast tumors. Tumori 68:407–412, 1982

32. Press WL, Flannery BP, Teukolsky SA, Vetterling WT: Numerical Recipes, Cambridge, Cambridge University Press, 1986

33. Salmon I, Coibion M, Larsimont D, Badr-El-Din V: Comparison of fine needle aspirates of breast cancers to imprint smears by means of digital cell image analysis. Analyt Quant Cytol Histol 13:193–200, 1991

34. Schondorf H, Schneider, V. (translator) : Aspiration Cytology of the Breast, Philadelphia, W B Saunders Co., 1978

35. Smallwood J, Herbert A, Guyer P, Taylor I: Accuracy of aspiration cytology in the diagnosis of breast disease. Br J Surg 72:841–843, 1985

36. Smith C, Butler J, Cobb C, State D: Fine-needle aspiration cytology in the diagnosis of primary breast cancer. Surgery 103:178–183, 1988

37. Somers RG, Young GP, Kaplan MJ, Bernhard VM, Rosenberg M, Somers D: Fine-needle aspiration biopsy in the management of solid breast tumors. Arch Surg 120:673–677, 1985

38. Stenkvist B, Bengtsson E, Eriksson O, Jarkrans T, Nordin BA: Morphometric expression of differentiation in fine-needle biopsies of breast cancer. Cytometry 1:292–295, 1981

39. Stone M: Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society 36:111–147, 1974

40. Ulanow RM, Galblum L, Canter JW: Fine needle aspiration in the diagnosis and management of solid breast lesions. Am J Surg 148:653–657, 1984

41. Vetrani A, Fulciniti F, Di Benedetto G, Zeppa P, Troncone G, Boscaino A, de Rosa G, Palombini L: Fine-needle aspiration biopsies of breast masses. An additional experience with 1153 cases (1985 to 1988) and a meta-analysis. Cancer 69:736–740, 1992

42. Wanebo HJ, Feldman PS, Wilhelm MC, Cocell FL, Binns RL: Fine needle aspiration cytology in lieu of open biopsy in management of primary breast cancer. Ann Surg 199:569–579, 1984

43. Weiss SM, Kulikowski CA: Computer Systems that Learn, San Mateo, CA, Morgan Kaufmann, 1991

44. Williams DJ, Shah M: A fast algorithm for active contours. *In* Proceedings of the Third International Conference on Computer Vision. Osaka, Japan, December 1990. pp 592–595

45. Wittekind C, Schulte E: Computerized morphometric image analysis of cytologic nuclear parameters in breast cancer. Analyt Quant Cytol Histol 9:480–4, 1987

46. Wolberg WH, Mangasarian OL: Multisurface method of pattern separation for medical diagnosis applied to breast cytology. Proceedings of the National Academy of Sciences, U.S.A. 87:9193–9196, 1990

47. Wolberg WH, Mangasarian OL: Computer-designed expert systems for breast cytology diagnosis. Analytical and Quantitative Cytology and Histology 15(1):67–74, February 1993