

Novel Morphometric Based Classification via Diffeomorphic Based Shape Representation Using Manifold Learning

Rachel Sparks and Anant Madabhushi*

Department of Biomedical Engineering, Rutgers University, USA

Abstract. Morphology of anatomical structures can provide important diagnostic information regarding disease. Implicit features of morphology, such as contour smoothness or perimeter-to-area ratio, have been used in the context of computerized decision support classifiers to aid disease diagnosis. These features are usually specific to the domain and application (e.g. margin irregularity is a predictor of malignant breast lesions on DCE-MRI). In this paper we present a framework for extracting Diffeomorphic Based Similarity (DBS) features to capture subtle morphometric differences between shapes that may not be captured by implicit features. Object morphology is represented using the medial axis model and objects are compared by determining correspondences between medial axis models using a cluster-based diffeomorphic registration scheme. To visualize and classify morphometric differences, a manifold learning scheme (Graph Embedding) is employed to identify nonlinear dependencies between medial axis model similarity and calculate DBS. We evaluated our DBS on two clinical problems discriminating: (a) different Gleason grades of prostate cancer using gland morphology on a set of 102 images, and (b) benign and malignant lesions on 44 breast DCE-MRI studies. Precision-recall curves demonstrate DBS features are better able to classify shapes belonging to the same class compared to implicit features. A support vector machine (SVM) classifier is trained to distinguish between different classes utilizing DBS. SVM accuracy was $83 \pm 4.47\%$ for distinguishing benign from malignant lesions on breast DCE-MRI and over 80% in distinguishing between intermediate Gleason grades of prostate cancer on digitized histology.

1 Introduction

Morphological features of suspicious structures on imaging can provide useful diagnostic and prognostic cues. In BIRADSTM classification, benign breast masses

* This work was made possible via grants from the Wallace H. Coulter Foundation, National Cancer Institute Grant Nos. R01CA136535-01, R21CA12718601, and R03CA143991-01, the Cancer Institute of New Jersey, and Bioimagene, Inc. The authors also wish to thank Dr. Mark Rosen, Dr. Michael Feldman and Dr. John Tomaszewski from the University of Pennsylvania for providing the medical imagery and expert analysis.

often present as smooth round masses while cancerous masses tend to have spiculated edges [1]. Morphological cues are a critical component of Gleason grading, a scheme employed to assess the invasiveness of prostate cancer (CaP) [2]. Intermediate Gleason grade patterns 3 and 4 have poor interobserver reproducibility among trained pathologists [3]. Creating a reliable system to grade CaP on histology could have profound clinical impact on CaP survival rate [4]. Gland morphology is a particularly important feature for distinguishing between intermediate Gleason grades; higher Gleason grades of CaP are characterized by irregularly shaped glands while lower Gleason grades of CaP have smooth margins with a distinct lumen.

Computerized decision support classifiers have employed features implicitly related to object morphology for discriminating pathological structures [1,4]. Implicit features quantify particular traits typically relating to the contour or area of objects. For example the knowledge that a spiculated lesion boundary in breast DCE-MRI is associated with malignancy has led to inclusion of boundary irregularity within the BIRADSTM lexicon of descriptors for MR mammography [1]. However, it may not always be obvious which morphological traits are most beneficial in discriminating pathologies. Implicit features (such as area overlap ratio) may be unable to capture subtle shape differences between objects, differences which may be highly relevant to disease classification.

We present a new shape characterization framework to (1) accurately distinguish between the morphology of shapes (independent of domain or application) and (2) distinguish subtle differences between similar shapes (such as glands in Gleason grade 3 and grade 4 of CaP histology). The medial axis model represents shape as a connected skeleton with corresponding vectors to the object surface [5]. We employ a point-based diffeomorphic registration algorithm to find a mapping between two medial axis models. Diffeomorphic mapping allows for the alignment of medial axis models so that a pairwise shape dissimilarity can be formulated between corresponding points on two medial axis models to obtain a shape dissimilarity measure. A pairwise dissimilarity matrix is then used to define the nonlinear high dimensional shape space. Graph embedding [6], a nonlinear dimensionality reduction method that preserves local topology between points, is employed to reconstruct a low dimensional shape space within which shapes are arranged as a continuum on a smooth shape manifold; distances between shapes on this manifold being a function of morphological differences between shapes. Our integrated quantitative morphometric framework employing, (1) the medial axis model, (2) a diffeomorphic algorithm to evaluate shape dissimilarity, (3) a graph embedding scheme to obtain DBS features is illustrated in the flowchart in Figure 1.

In this work we evaluate our framework in the context of distinguishing (a) benign, Gleason grade 3, and Gleason grade 4 glands on digital histology images of needle core prostate biopsies (611 glands from 102 images) and (b) malignant and benign lesions on 44 DCE-MRI studies. We compared our framework against implicit features (e.g. contour smoothness, area overlap ratio) in terms of classification performance.

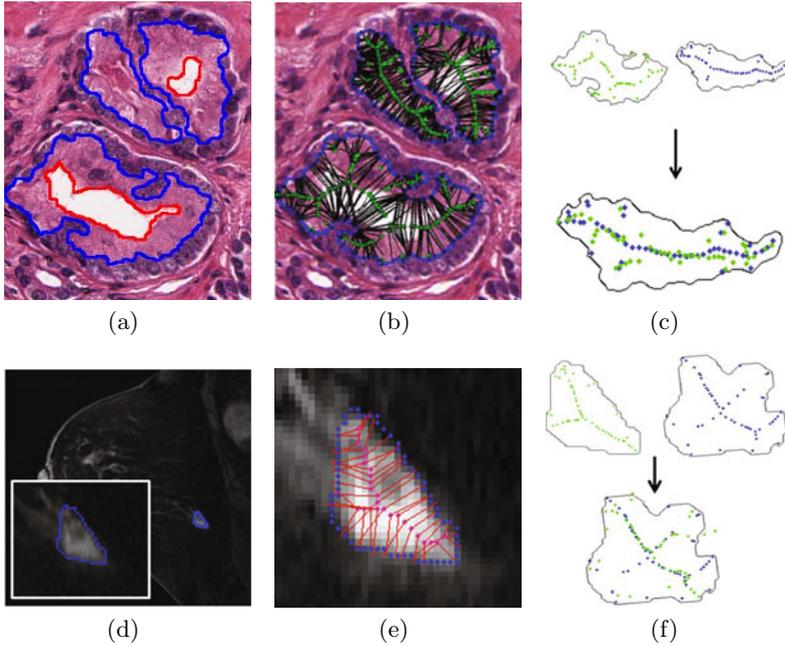


Fig. 1. Framework for extracting DBS features to (top row) distinguish Gleason grades of CaP and (bottom row) benign and malignant lesions on DCE-MRI. (a) For digital histology, both nuclear (blue) and lumen (red) boundaries of prostate glands are segmented [4]. (d) Lesion boundary (blue) segmentation is performed on breast DCE-MRI [1]. (b) Medial axis model (green, black) is fit to the nuclear or the luminal layer (not shown). (e) Lesion boundary on breast DCE-MRI is likewise fit with a model (red, magenta). (c), (f) A cluster-based diffeomorphic registration for mapping between template (blue) and registered (green) medial axis models helps guide a point-based dissimilarity measure. Object contours (black) are also displayed.

2 Shape Model Representation and Shape Dissimilarity

We define an image scene $\mathcal{C} = (C, f)$ where C is a 2D grid of image pixels $c \in C$ and an image intensity function $f(c) : c \in C$. The image pixels c are located in the X - Y Cartesian space. Objects of interest are segmented via application of a hybrid active contour model [7] such that $\mathcal{X} \subset C$ defines the set of pixels comprising the boundary of the object of interest.

2.1 Medial Axis Model

From an object contour \mathcal{X} a medial axis model M can be constructed, where $m \in M$ is a set of pixels $m \subset C$ pertaining to the skeletal backbone of the object [5]. To determine m backbone pixels, an image distance map for the object defined as $\mathcal{C}^e = (C, f^e)$ is constructed. The distance map function $f^e(c)$ is defined as:

$$f^e(c) = \begin{cases} 0 & c \in \mathcal{X}, \\ -\min_{p \in \mathcal{X}} \|c - p\| & c \in \mathcal{X}^h, \\ \min_{p \in \mathcal{X}} \|c - p\| & c \in \mathcal{X}^o, \end{cases} \quad (1)$$

where $\mathcal{X}^h \subset C$ is the set of pixels in the image contained within \mathcal{X} , and $\mathcal{X}^o \subset C$ refers to the pixels outside of \mathcal{X} . A gradient map of an object, $\widehat{C}^e = (C, \widehat{f}^e)$ is calculated for $c \in C$ as,

$$\widehat{f}^e(c) = \left(\frac{\partial f^e(c)}{\partial x} \right)^2 + \left(\frac{\partial f^e(c)}{\partial y} \right)^2. \quad (2)$$

where $\frac{\partial()}{\partial x}$, $\frac{\partial()}{\partial y}$ are the partial derivatives along the Cartesian X and Y axes respectively. Points on the medial axis are defined as, $M = \{m : m \in C, \widehat{f}^e(m) < \tau\}$. Empirically, $\tau = 0.05 \operatorname{argmin}_{c \in C} (\widehat{f}^e(c))$ was found to give a well defined medial axis with few spurious branches. Then for $m \in M$, the two closest points on the surface can be defined as

$$\widehat{p}_1 = \operatorname{argmin}_{p \in \mathcal{X}} \|m - p\|, \text{ and } \widehat{p}_2 = \operatorname{argmin}_{p \in \mathcal{X}, p \neq \widehat{p}_1} \|m - p\|,$$

and the corresponding surface vectors are defined $\mathbf{v}_1 = \widehat{p}_1 - m$, and $\mathbf{v}_2 = \widehat{p}_2 - m$.

2.2 Shape Model Dissimilarity Metric

To compare a set of N medial axis models $\{M_\alpha : \alpha \in \{1, \dots, N\}\}$, a dissimilarity metric based on model correspondence is calculated. M_α denotes sets of unlabeled points $m \in M_\alpha$. Hence the problem of determining the correspondence between models is reduced to the problem of determining point correspondences. We use a diffeomorphic registration guided by point clustering to find a mapping between point clouds [8].

Cluster Update. Point clusters are determined using simulated annealing [9]. For a set of medial axis points $m \in \{M_\alpha : \alpha \in \{1, \dots, N\}\}$, a set of K cluster centers at some iteration i may be represented as $\{\widehat{m}_\alpha^{k,i} : k \in \{1, \dots, K\}\}$. The number of clusters K is a user selected parameter. The probability of a point $m \in M_\alpha$ belong to the cluster centered at $\widehat{m}_\alpha^{k,i}$ is determined by,

$$P(m|\widehat{m}_\alpha^{k,i}) = \frac{e^{-\sigma_i \|m - \widehat{m}_\alpha^{k,i}\|^2}}{\sum_{k=1}^K e^{-\sigma_i \|m - \widehat{m}_\alpha^{k,i}\|^2}}. \quad (3)$$

Similarly, $P(m|\widehat{m}_\beta^{k,i})$ is defined for a different medial axis model $\{m \in M_\beta : \beta \in \{1, \dots, N\}\}$. The variable σ_i is defined as, $\sigma_i = (\epsilon)^i \sigma_0$, for some $\epsilon > 1$. So that σ_i become progressively larger as i increases, causing the probability function $P(m|\widehat{m}_\alpha^{k,i})$ to be more likely a member of closer cluster centers. The initialization term σ_0 is defined as $1/\sigma_0 = \max_{m \in M_\alpha} \|m - \mu_\alpha\| + \max_{m \in M_\beta} \|m - \mu_\beta\|$, where M_α, M_β represent any two shapes with corresponding centroids μ_α, μ_β .

Cluster centers at the next iteration $i + 1$ are updated as in [8]. As each point has partial membership in several clusters, cluster centers are updated to reflect this membership structure according to the following update equation,

$$\hat{m}_\alpha^{k,i+1} = \frac{\hat{m}_\beta^{k,i} + \sum_{m \in M_\alpha} m P(m|\hat{m}_\alpha^{k,i})}{1 + \sum_{m \in M_\alpha} P(m|\hat{m}_\alpha^{k,i})}. \tag{4}$$

The term $\hat{m}_\beta^{k,i}$ in Equation 4 ensures that $\hat{m}_\alpha^{k,i}$ does not move too far from its counterpoint in M_β . Similarly, $\hat{m}_\beta^{k,i+1}$ is defined for M_β . After each cluster center update a diffeomorphic mapping is used to align the corresponding cluster centers in $\hat{m}_\alpha^{k,i+1}$ and $\hat{m}_\beta^{k,i+1}$.

Diffeomorphic Mapping. A diffeomorphism over the image space C can be constructed from a set of paths $q^{k,i}(t) : \{t \in \{0, \dots, T\}, k \in \{1, \dots, K\}\}$. Each path $q^{k,i}(t)$ describes the movement of a point which starts at the location of the cluster centroid of one medial axis $q^{k,i}(0) = \hat{m}_\alpha^{k,i}$ and ends at the corresponding cluster centroid of the other medial axis $q^{k,i}(T) = \hat{m}_\beta^{k,i}$ transversing a subset of the image space C [10]. Additionally, we constrain $q^{k,i}(t)$ to minimize the deformation of the image space C using a kernel function G to smooth the image space. We choose to utilize Greene’s function defined as, $G(\Gamma, \Theta) = -(\Gamma - \Theta)^2 \log(\Gamma - \Theta)^2$, where $\Gamma \in C$ and $\Theta \in q^{k,i}(t)$.

An iterative update function is solved for where $q^{k,i}(t)$ is initialized as a line between the beginning and end points and iteratively updated to minimize the energy equation defined as,

$$\hat{q}^{k,i}(t) = \operatorname{argmin}_{\hat{q}^{k,i}(t)} \sum_{k=1}^K \sum_{t=0}^T \omega^i(t) \cdot \left(\sum_{\eta=1}^K \omega^\eta(t) G(q^{\eta,i}(t), \hat{q}^{k,i}(t)) \right), \tag{5}$$

where the variables $\omega^i(t)$ and $\omega^\eta(t)$ must also be solved for. We alternate minimizing over the variables $\omega^i(t)$, $\omega^\eta(t)$, and the path $q^{k,i}(t)$ [10].

Point Correspondence For the set of paths $\hat{q}^{k,i}(t)$ obtained, M_α can be mapped to \tilde{M}_α . The mapped medial axis model \tilde{M}_α shares the coordinate space of M_β . In this shared coordinate space point correspondences are determined by,

$$(\hat{u}, \hat{v}) = \operatorname{argmin}_{\hat{u}, \hat{v}} \|\tilde{m}_\alpha^u - m_\beta^v\|, \tag{6}$$

where $\tilde{m}_\alpha^{\hat{u}} \in \tilde{M}_\alpha : \hat{u} \in \{1, \dots, S\}$ is the set of S medial axis points contained in \tilde{M}_α ; similarly $m_\beta^{\hat{v}}$ is defined for M_β . The matched indices (\hat{u}, \hat{v}) are used to calculate dissimilarity between 2 shapes in the original space by,

$$A(\alpha, \beta) = \sum_{(\hat{u}, \hat{v}=1)}^N \kappa_1 \|\tilde{m}_\alpha^{\hat{u}} - m_\beta^{\hat{v}}\| + \kappa_2 \|\mathbf{v}_{\alpha,1}^{\hat{u}} - \mathbf{v}_{\beta,1}^{\hat{v}}\| + \kappa_3 \|\mathbf{v}_{\alpha,2}^{\hat{u}} - \mathbf{v}_{\beta,2}^{\hat{v}}\|. \tag{7}$$

For N objects $A \in \mathbb{R}^{N \times N}$ is a high dimensional dissimilarity matrix. A is positive definite when the user selected weights κ_1, κ_2 , and $\kappa_3 > 0$.

2.3 Manifold Learning

Graph embedding [6] reduces A to a low dimensional space $\mathbf{y} \in \mathbb{R}^d$, where $N \gg d$. Optimal low dimensional projections $\mathbf{y} = [y_1, y_2, \dots, y_N]$ and can be found by,

$$\mathbf{y} = \arg \min \left(\sum_{a,b=1}^N \|y_a - y_b\|^2 w_{ab} \right), \tag{8}$$

where $w_{ab} = e^{-A(a,b)/\gamma}$, and $\gamma > 0$ is used to normalize A . If W is positive definite, Equation 8 reduces to a minimum eigenvalue decomposition problem,

$$W\mathbf{y} = \lambda D\mathbf{y}, \tag{9}$$

where D is a diagonal matrix $D_{aa} = \sum_b W_{ab}$. The top d eigenvalues in λ correspond to the d -dimensional embeddings in \mathbf{y} , and the top d DBS.

3 Quantitative Evaluation

3.1 Datasets

Prostate needle core biopsy tissue samples stained for hematoxylin and eosin were digitally imaged at 40x optical magnification. An expert pathologist manually graded the tissue samples. A total of 24 samples containing 94 glands were identified as benign, 67 samples containing 470 glands were identified as grade 3, and 11 samples containing 47 glands were identified as grade 4. The 44 patient breast lesion DCE-MRI comprised of 16 benign and 28 malignant masses.

We compared DBS to implicit features: contour smoothness, contour standard deviation, compactness, area overlap ratio, average radial ratio, radial ratio standard deviation [1].

3.2 Precision Recall Curves

Precision-recall (PR) curves are generated as follows: the closest objects within differently sized neighborhoods of a query object are identified. Relevant objects are defined as belonging to the same class as the query object. Precision is relevant objects in the neighborhood over the neighborhood size. Recall is all relevant objects in the neighborhood over all relevant objects in the dataset. PR curves are generated by measuring average precision and recall over different neighborhood sizes for several query objects in the database with the same class label. Table 1 reports area under the PR curve. DBS consistently outperforms implicit features.

3.3 Support Vector Machine Accuracy

The ability of DBS or implicit features to determine disease state was evaluated using a SVM classifier [11]. The SVM used 5-fold cross validation where 2/3 of the dataset was used to train the SVM and 1/3 of the dataset used to test. For the multiclass digital histology problem a one-against-all SVM scheme was implemented. DBS outperformed implicit features for all tasks (see Table 2).

Table 1. Area under the PR curve (AUC) for implicit and DBS features. A perfect AUC score is dependent on the distribution of the query object in the database and feature performance, higher numbers represent better distinction between classes. The best AUC values for each task is bolded.

Morphological Feature	Query Object				
	DCE-MRI		Histology		
	Benign	Malignant	Benign	Grade 3	Grade 4
Implicit	0.39 ± 0.08	0.63 ± 0.70	0.12 ± 0.05	0.82 ± 0.03	0.10 ± 0.02
DBS	0.40 ± 0.07	0.62 ± 0.05	0.21 ± 0.07	0.85 ± 0.08	0.20 ± 0.09

Table 2. SVM classification accuracy evaluated on two feature sets with 5-fold cross validation. For DCE-MRI the dataset was evaluated on lesions classified as benign (n=16) or malignant (n=28). For the digital histology glands were classified as belonging to one of three classes: benign (n=94), Gleason grade 3 (n=470), and Gleason grade 4 (n=47). The best feature set for all classification tasks is bolded.

Morphological Feature	DCE-MRI	Digital Histology		
		Benign	Grade 3	Grade 4
Implicit	77.5 ± 3.73%	79.47 ± 4.71%	69.47 ± 7.58%	75.26 ± 5.77%
DBS	83.00 ± 4.47%	82.89 ± 3.97%	86.58 ± 7.39%	84.74 ± 3.23%

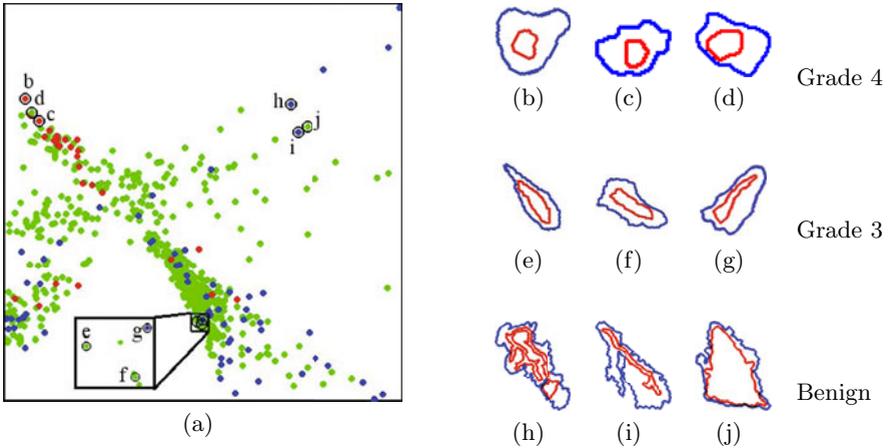


Fig. 2. (a) DBS features for prostate digital histology with benign (blue), Gleason grade 3 (green), and Gleason grade 4 (red) glands. The first and second DBS features are plotted on the X and Y axes respectively. Lumen (red) and nuclear (blue) layers are shown, for glands labeled (b)-(d) grade 4, (e)-(g) grade 3, and (h)-(j) benign. Ground truth for mislabeled glands, displayed in the far right row, are (d) grade 3, (g) benign, (j) grade 3. Similar shapes are nearby on the manifold while dissimilar shapes are far.

4 Concluding Remarks

We presented a framework to calculate a set of new morphological features DBS. Pairwise comparisons of medial axis models describes a high dimensional shape space. Graph Embedding of the shape space allows for extraction of DBS features which were evaluated on two medical classification problems: (a) Gleason grading of CaP digital histology (611 glands on 102 images), and (b) distinguishing benign from malignant lesions on DCE-MRI (44 studies). DBS outperform implicit features for both classification tasks considered in this paper, both in terms of area under precision-recall curves and SVM classifier accuracy. We believe this improved performance was due to DBS capturing subtle shape differences that could not be appreciated by the implicit features. Future work will involve evaluating DBS for more applications.

References

1. Agner, S., Soman, S., Libfeld, E., McDonald, M., Thomas, K., Englander, S., Rosen, M., Chin, D., Noshier, J., Madabhushi, A.: Textural kinetics: A novel dynamic contrast enhanced (DCE)- MRI feature for breast lesion classification. *Journal of Digital Imaging* (2010)
2. Epstein, J., Allsbrook, W., Amin, M., Egevad, L.: The 2005 international society of urological pathology (isup) consensus conference on gleason grading of prostatic carcinoma. *American Journal of Surgical Pathology* 29, 1228–1242 (2005)
3. Madabhushi, A., Doyle, S., Lee, G., Basavanahally, A., Monaco, J., Masters, S., Tomaszewski, J., Feldman, M.: Review: Integrated diagnostics: a conceptual framework with examples. *CCLM* (2010) (in Press)
4. Monaco, J., Tomaszewski, J., Fledmand, M., Hagemann, I., Moradi, M., Mousavi, P., Boag, A., Davidson, C., Abolmaesumi, P., Madabhushi, A.: High-throughput detection of prostate cancer in histological sections using probabilistic pairwise markov models. *MeDIA* 14, 617–629 (2010)
5. Blum, H.: A transformation for extracting new descriptors of shape. In: *Models for the Perception of Speech and Visual Form*, pp. 367–380. MIT Press, Cambridge (1967)
6. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Trans. PAMI* 22, 888–905 (2000)
7. de Luis-Garcia, R., Deriche, R., Alberola-Lopez, C.: Texture and color segmentation based on the combined use of the structure tensor and the image components. *Signal Process.* 88, 776–795 (2008)
8. Guo, H., Rangarajan, A., Joshi, S.: Diffeomorphic point matching. In: *Handbook of Mathematical Models in Computer Vision*, pp. 205–219. Springer, US (2005)
9. Rose, K., Gurewitz, G., Fox, G.: A deterministic annealing approach to clustering. *Pattern Recogn. Lett.* 11, 589–594 (1990)
10. Twining, C., Marland, S., Taylor, C.: Measuring geodesic distances on the space of bounded diffeomorphisms. *BMVC* 2, 847–856 (2002)
11. Cortes, C., Vapnik, V.: Support-vector networks. *Mach. Learn.* 20, 273–297 (1995)