Nonparametric Joint Shape and Feature Priors for Image Segmentation

Ertunc Erdil, *Member, IEEE*, Muhammad Usman Ghani, Lavdie Rada, Ali Ozgur Argunsah, Devrim Unay, *Member, IEEE*, Tolga Tasdizen, *Member, IEEE*, and Mujdat Cetin, *Member, IEEE*

***Abstract*— In many image segmentation problems involving limited and low-quality data, employing statistical prior information about the shapes of the objects to be segmented can significantly improve the segmentation result. However, defining probability densities in the space of shapes is an open and challenging problem, especially if the object to be segmented comes from a shape density involving multiple modes (classes). Existing techniques in the literature estimate the underlying shape distribution by extending Parzen density estimator to the space of shapes. In these methods, the evolving curve may converge to a shape from a wrong mode of the posterior density when the observed intensities provide very little information about the object boundaries. In such scenarios, employing both shape- and class-dependent discriminative feature priors can aid the segmentation process. Such features may involve, e.g., intensity-based, textural, or geometric information about the objects to be segmented. In this paper, we propose a segmentation algorithm that uses nonparametric joint shape and feature priors constructed by Parzen density estimation. We incorporate the learned joint shape and feature prior distribution into a maximum *a posteriori* estimation framework for segmentation. The resulting optimization problem is solved using active contours. We present experimental results on a variety of synthetic and real data sets from several fields involving multimodal shape densities. Experimental results demonstrate the potential of the proposed method.**

***Index Terms*— Nonparametric joint shape and feature priors, Parzen density estimator, multimodal shape density, image segmentation, shape prior.**

I. Introduction

Segmentation of images based on limited and low quality data is a challenging problem and requires prior information about the shape to be segmented for an acceptable solution. For example, given a training set of car shapes, a partially occluded car object in an image can be segmented by exploiting prior shape information obtained from the training set. The problem becomes more complex when the training set of shapes involves examples from multiple classes (e.g., car, truck, plane, etc.) leading to a multimodal shape density. In this paper, we focus on segmentation problems in which shape distributions are multimodal and complex, but just the shape prior information is not sufficient for effective segmentation due to severe occlusion. The proposed approach deals with the problem by incorporating discriminative class-dependent feature priors together with shape priors into the segmentation process. We demonstrate that the proposed approach overcomes the limitations of existing segmentation methods that use only shape priors.

A. Related Work

One of the earliest attempts to include a prior information in image segmentation is the active contour model, also called "snakes", by Kass *et al.* [1]. Snakes use a general regularity term as the prior, where the roughness and length of the curve serve as a penalty, which is based on the assumption that smoother and shorter curves are more likely [2]. However, in many applications a more informative object-type specific shape prior can be learned from training samples. In this regard, active shape models (ASM) proposed by Cootes *et al.* [3] are powerful techniques for segmentation using shape priors. Variants of the ASM, their applications to different image segmentation areas, and a review can be found in [4]–[8].

In the original ASM, a training set of shapes represented by landmarks is used to construct allowable shape variations via principal component analysis (PCA). The use of linear analysis tools such as PCA in ASMs limits the domain of applicability of these techniques to shape priors involving only unimodal densities. That is, the original ASMs assume that the training shapes are distributed according to a unimodal, Gaussian-like distribution; hence, the technique cannot model more complex (multimodal) shape distributions.

Several methods have been proposed to handle multimodal distributions of shapes by extending ASMs [9]–[11]. These approaches include the use of mixture of Gaussians [9], manifold learning techniques [10] and kernel PCA [11], [12]. However, these approaches use parametric probability distributions, which may not model very complex shape

Manuscript received November 25, 2016; revised June 26, 2017; accepted July 3, 2017. Date of publication July 17, 2017; date of current version August 21, 2017. This work was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant 113E603. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Scott T. Acton. *(Corresponding author: Ertunc Erdil.)*

E. Erdil, M. U. Ghani, and M. Cetin are with the Faculty of Engineering and Natural Sciences, Sabanci University, 34956 Istanbul, Turkey (e-mail: ertuncerdil@sabanciuniv.edu; ghani@sabanciuniv.edu; mcetin@sabanciuniv.edu).

L. Rada is with the Faculty of Engineering and Natural Sciences, Bahcesehir University, 34353 Istanbul, Turkey.

A. O. Argunsah is with the Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, 1400-038 Lisbon, Portugal.

D. Unay is with the Department of Biomedical Engineering, Izmir University of Economics, 35330 Izmir, Turkey.

T. Tasdizen is with the Department of Electrical and Computer Engineering, University of Utah, Salt Lake City, UT 84112 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2017.2728185

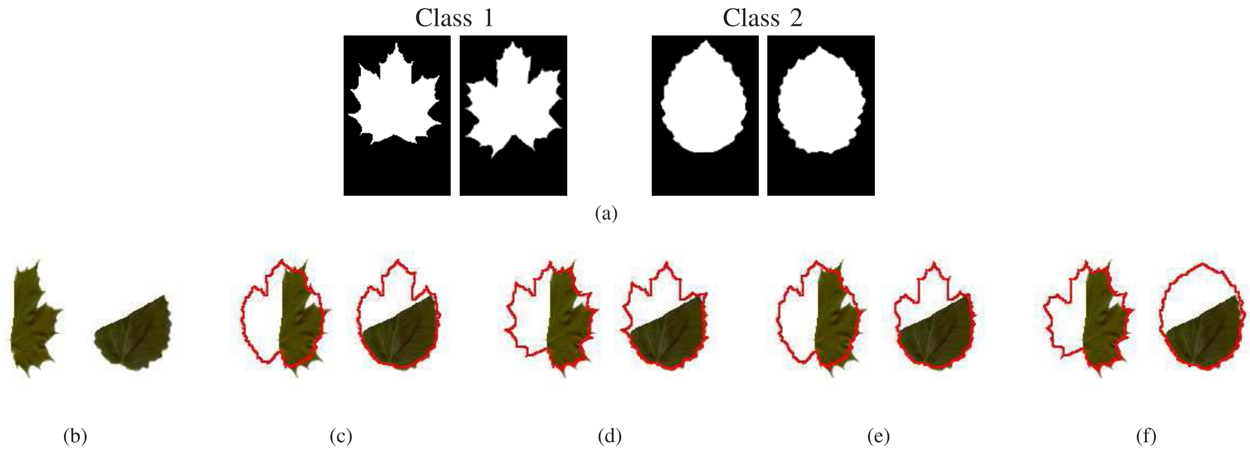


Fig. 1. Toy example that demonstrates motivation of the proposed method. (a) Training Set. (b) Test Image. (c) Kim *et al.* [2]. (d) Foulonneau *et al.* [20]. (e) Chen and Radke [22]. (f) Proposed.

variations [13]. In addition, the explicit (landmark-based) shape representation used in ASMs has two major shortcomings. First, annotating landmark points with correct correspondences across all example shapes can be difficult and time consuming. Second, the extensions of the technique to handle topological changes are not straightforward. To overcome the limitations of landmark-based representation, level set based shape priors were proposed [14], [15]. Because of their implicit nature, level set methods do not need landmarks and can easily handle topological changes [16], [17]. In [14] and [15], shape variability is captured using PCA on signed distance functions of level sets. However, these techniques work well only when the shape variation is small due to their use of PCA. Therefore, they cannot handle multimodal shape densities.

In order to learn multimodal shape densities, Kim *et al.* [2] and Cremers *et al.* [18] proposed nonparametric density estimation based shape priors using level sets. These methods estimate the prior shape density by extending Parzen density estimator over the distances between the level set representations of the evolving curve and training shapes. These ideas have also been extended to the problem of segmenting multiple objects through the use of coupled shape priors [19]. An interesting usage of nonparametric shape priors proposed by Foulonneau *et al.* [20] computes Legendre moments from binary images as shape descriptors and uses distances between descriptors instead of level sets for estimating the prior shape density. The approach also exploits appealing properties of Legendre moments for intrinsic alignment. The approaches of Kim *et al.* [2], Cremers *et al.* [18] and Foulonneau *et al.* [20] use a simple data term that assumes the foreground and the background intensities are piecewise-constant [21]. In the literature, there are also methods that combine nonparametric shape priors with learning-based data terms [22]–[24]. Using a more sophisticated data term significantly improves the segmentation quality when the object foreground and background have complex densities. Some other recent work that exploits nonparametric shape priors and a more detailed review of the level set based segmentation methods can be found in [25]–[31].

B. Motivation

The methods [2], [18], [20], [22]–[24] that use nonparametric shape priors performs well in the presence of occlusion and missing data. They also capable of handling multimodal shape densities. However, the shortcomings of these methods arises when the level of occlusion and missing data increases and when the underlying shape density is multimodal. This is due to the fact that the prior density is estimated by extending Parzen density estimator over the distances between the evolving curve and training shapes. These methods use gradient descent to minimize an energy function including data and shape priors terms. During gradient descent, a curve represented by level sets is evolved by a data-driven force together with the weighted average of the training shapes where the weight of each training shape is usually inversely proportional to its distance to the evolving curve (the exact form of the weights is determined by the specific metric used to measure distances between shapes). Therefore, when the observed data are very limited, the evolving curve can be more similar to training shapes from a different class based on the distance metric. In these cases, the evolving curve is driven toward a shape from a different mode of the shape density, which yields inaccurate segmentation results.

We illustrate the aforementioned drawback of Kim *et al.* [2], Foulonneau *et al.* [20] and Chen and Radke [22] through the example shown in Figure 1.[1] In this example, we use a training set that contains samples from two different leaf shape classes as shown in Figure 1(b). Note that the boundaries of the leaf shapes are uneven in class 1 and smooth in class 2. We have 2 test images from each class as shown in Figure 1(b). Note that the test images are severely occluded; almost half of the leaf shapes do not appear. Since the curve found by the data term is more similar to class 2 based on the distance metric, Kim *et al.* [2] produce segmentation results

[1] Note that these three methods are representative ones; Kim *et al.* [2] estimate prior density using distances between shapes, Foulonneau *et al.* [20] estimate prior density using distances between Legendre moments and Chen and Radke [22] use intensity prior-based data term together with the shape prior term. The other nonparametric shape prior-based methods exhibit a similar behavior with one of these methods.

that are more similar to the shapes in class 2 in both test images. The major difference between Kim *et al.* [2] and Chen and Radke [22]is the design of the data term. Since the data provide very little information in the test images, the effect of the data term is very limited in the segmentations. Therefore, Chen and Radke [22] produce very similar results with Kim *et al.* [2] as shown in Figure 1(e). The method of Foulonneau *et al.* [20] produces segmentation results that are more similar to the shapes in class 1 (see Figure 1(d)). This means that estimating the prior shape density based on the distances between Legendre moments does not help to have segmentation results from the correct mode of the shape density in the presence of severe occlusion.

This motivates us to deal with the shortcomings of the existing methods by incorporating discriminative class-dependent features to the kernel density estimation process. For example, circularity of the shapes in Figure 1 is an important feature for identifying different leaf classes. In such cases, jointly estimating feature and shape prior density can yield more accurate segmentations as shown in Figure 1(f).

C. Contributions

Our contribution in this paper is a segmentation algorithm that performs segmentation by exploiting nonparametric joint shape and feature priors. Unlike the state-of-the-art methods that perform segmentation using nonparametric shape density estimation, we exploit learned discriminative class-dependent features (geometric or appearance-based) extracted from specific parts of the scene relative to the object of interest and incorporate the joint shape and feature prior density into the segmentation process. In particular, we combine a data term and a joint shape and feature prior term within a Bayesian framework to form the energy functional for segmentation. To the best of our knowledge, nonparametric joint shape and feature priors have not been proposed for image segmentation in the literature. By estimating a more discriminative prior density, our algorithm is able to find better segmentations based on the shape posterior density.

Our approach may seem similar to the methods proposed by Cremers *et al.* [32] and Chan and Zhu [33]. However, these approaches and the proposed approach focus on completely different problems. In [32] and [33], given a scene with multiple different types of objects, the problem is to segment a particular object that is included in the training set. In this paper, we focus on the problem of segmenting an object using the correct shape priors when the training set contains shapes from different classes.

Preliminary results of this work were presented in [34]. This paper advances its preliminary version in several major ways. In particular, (1) while [34] was focused on the specific problem of spine segmentation, in this paper we significantly expand the domain of applicability of this new idea; (2) we consider and use new types of features in our framework; (3) we present the results of an expanded experimental analysis on a variety of data sets, together with quantitative comparison to the results of several state-of-the-art methods; (4) we provide a more detailed technical development and discussion of the proposed method; (5) we present an expanded coverage of related work.

The implementation of the proposed approach is available at https://github.com/eerdil/.

II. The Proposed Method

A. The Energy Function

In this section, we propose an energy function that exploits nonparametric joint shape and feature priors for image segmentation. Let C be the evolving curve, f be the feature vector and $data$ be the intensity image. Then, the posterior probability density function of C and f can be written using Bayes' rule as follows:

$$p(C, f|data) = \frac{p(data|C, f)p(C, f)}{p(data)} \tag{1}$$

where,

$$p(data|C, f) = \frac{p(f|data, C)p(data|C)}{p(f|C)}. \tag{2}$$

Plugging in Equation (2) into (1) yields

$$p(C, f|data) \propto p(f|data, C)p(data|C)p(C) \tag{3}$$

and $p(C)$ can be written as

$$p(C) = \int p(C, f)\,\mathrm{d}f. \tag{4}$$

Then, Equation (3) becomes

$$p(C, f|data) \propto p(data|C)p(f|data, C)\int p(C, f)\mathrm{d}f. \tag{5}$$

Let us assume that we observe a feature vector $\hat{f}$ either from data or from boundary. Such features could involve geometric, textural, or appearance-based information about the object to be segmented. From this point on, one can proceed with various assumptions on the probability densities involved. For feature extraction, we assume that features can be extracted perfectly based on the data as well as information about the boundary when it reaches a reasonable state. This leads to the degenerate density:

$$p(f|data, C) = \delta(f - \hat{f}) \tag{6}$$

where, $\delta(.)$ is the Dirac delta function. Also, we learn $p(C, f)$ nonparametrically from the training data. Since $\hat{f}$ is already observed, Equation (5) can be written as follows:

$$p(C, \hat{f}|data) \propto p(data|C)p(C, \hat{f}). \tag{7}$$

Note that, $p(C, \hat{f})$ is also equivalent to the slice of $p(C, f)$ at $\hat{f}$ which is $p(C|f = \hat{f})$. Therefore, Equation (7) and the following equation are identical.

$$p(C, \hat{f}|data) \propto p(data|C)p(C|f = \hat{f}). \tag{8}$$

Hence, given the simplifying perfect feature extraction assumption in Equation (6), the learned joint shape and feature density is used through conditioning on the extracted feature. This conditioning guides the segmentation process, possibly towards the correct mode of the multimodal shape density. If needed, one could certainly relax this assumption in our framework, and develop an optimization algorithm for maximizing the posterior density in Equation (1) to infer both the feature and the shape based on the data and the learned joint prior.

The data term we use is the piecewise-constant version of the Mumford-Shah functional [21], [35]. We use this data term as a representative one, since it has been previously used in a variety of applications [18], [36]. One can consider using more sophisticated data terms such as those involving mutual information [37], J-Divergence [38], and Bhattacharya distance [39]. We discuss estimating the joint shape and feature prior density, $p(C, \hat{f})$, in the following section.

By simply taking the negative logarithm of Equation (7), we can define the following energy function to be minimized for segmentation.

$$\begin{aligned} E(C, \hat{f}) &= -\log p(data|C) - \log p(C, \hat{f}) \\ &= \beta\Big[\int_{C_{in}} (I(x) - m_{in})^2 \mathrm{d}x \\ &\quad + \int_{C_{out}} (I(x) - m_{out})^2 \mathrm{d}x\Big] - \log p(C, \hat{f}) \end{aligned} \tag{9}$$

where $I(.)$ is the intensity image, C_{in} (C_{out}) is the region inside (outside) of the segmenting curve C, m_{in} (m_{out}) is the average intensities in these regions, and β is a constant that determines the balance between the data and the prior terms which we set $\beta = 1$.

B. Building Joint Shape and Feature Priors

Let us assume that we have n aligned training shapes $\mathbf{C} = \{C_1, C_2, \ldots, C_n\}$ and a corresponding set of feature vectors $\mathbf{f} = \{f_1, f_2, \ldots, f_n\}$ extracted from intensity images. The basic idea we use is that the segmenting curve C will be more likely if it is similar to the training shapes and $\hat{f}$ is similar to the training feature vectors. In order to measure the similarity between curves, we need to compare C with the training shapes in $\mathbf{C}$. However, when C and the training shapes in $\mathbf{C}$ are not aligned, a direct comparison of C with the shapes in $\mathbf{C}$ includes not only shape differences but also artifacts due to pose difference such as translation, rotation, and scaling. In order to remove pose artifacts, we align C with the shapes in $\mathbf{C}$ into $\tilde{C}$, where $\tilde{C}$ is the aligned version of C. Also, recall that shapes in $\mathbf{C}$ are already aligned. Similarly, in order to extract pose invariant features, all feature vectors should be extracted after alignment. Any kind of rigid alignment approach can be used to obtain an aligned training set of shapes from its unaligned version for which we use the approach proposed by Tsai *et al.* [14]. Then, the joint shape and feature density is estimated using Parzen density estimation as follows[2]

$$p(\tilde{C}|f = \hat{f}) \propto p(\tilde{C}, \hat{f}) = \frac{1}{n}\sum_{i=1}^{n} k(d(\tilde{C}, C_i), d(\hat{f}, f_i), \sigma_C, \sigma_f) \tag{10}$$

[2]Note that in Parzen density estimation, class labels of the shapes in the training set are not available.

where $d(\cdot,\cdot)$ is a distance metric, $k(\cdot,\cdot,\sigma_C,\sigma_f)$ is a 2D kernel with shape kernel size σ_C and with feature kernel size σ_f. For the kernel sizes σ_C and σ_f, we use an ML kernel with leave-one-out [40]. Note that, the composite of the 2D kernel and the distance metrics plays the role of an infinite dimensional kernel. A variety of distance metrics can be used in Equation (10) [2]. In our experiments, we use the template distance metric [2], d_T, for shape distance and the L_2 distance metric, d_{L_2}, for feature distance.

Note that, we compute the joint shape and feature prior density for the aligned curve, $\tilde{C}$, in Equation (10) to remove the pose artifacts as we mentioned above. We explain how to relate $p(\tilde{C}, \hat{f})$ to $p(C, \hat{f})$ in our segmentation method in the following section.

C. Segmentation Algorithm

The aim of the proposed segmentation approach is to minimize the energy functional in Equation (9) by gradient descent, and the task comes down to computing the gradient flow for the curve C. The overall gradient flow is the sum of the two terms, one based on the data term and the other based on the shape and feature prior term. The gradient flow for the data term is given by

$$\frac{-\partial \log p(data|C)}{\partial C} = \beta\Big[-(I(x)-m_{in})^2 + (I(x)-m_{out})^2\Big]\vec{N}, \tag{11}$$

where $\vec{N}$ is the outward curve normal [21].

However, we cannot compute $\frac{\partial \log p(C,\hat{f})}{\partial C}$ directly from the shape and feature prior term due to the need for removing pose differences mentioned in Section II-B. Instead, we first compute $\frac{\partial \log p(\tilde{C},\hat{f})}{\partial \tilde{C}}$ and relate it to $\frac{\partial \log p(C,\hat{f})}{\partial C}$. The gradient flow $\frac{\partial \log p(\tilde{C},\hat{f})}{\partial \tilde{C}}$ for the joint shape and feature prior term is given by

$$\begin{aligned} \frac{\partial \log p(\tilde{C}, \hat{f})}{\partial \tilde{C}} &= \frac{1}{p(\tilde{C}, \hat{f})} \times \frac{1}{n} \times \frac{1}{\sigma_C \times \sigma_f} \\ &\quad \times \sum_{i=1}^{n} \Big(k(d_T(\tilde{C}, C_i), d_{L_2}(\hat{f}, f_i), \sigma_C, \sigma_f) \\ &\quad \times d_T(\tilde{C}, C_i) \times (d_{L_2}(\hat{f}, f_i))^2 \times (1 - 2H(\phi_{C_i})\Big). \end{aligned} \tag{12}$$

where $H(\cdot)$ is the Heaviside function and ϕ_C is the corresponding signed distance function of curve (shape) C. We use the sign convention of $\phi < 0$ for inside the curve and $\phi > 0$ for outside the curve. The derivation of the gradient flow in Equation (12) is a straightforward extension of the derivation in [2] and is given in Appendix.

In order to compute $\frac{\partial \log p(C,\hat{f})}{\partial C}$ from $\frac{\partial \log p(\tilde{C},\hat{f})}{\partial \tilde{C}}$, we need a pose parameter, $\mathbf{p}$, that aligns C with the shapes $\mathbf{C}$ into $\tilde{C}$ in each iteration of the gradient descent (see line 10 in Algorithm 1). After $\frac{\partial \log p(\tilde{C},\hat{f})}{\partial \tilde{C}}$ is computed (see line 12 of Algorithm 1), $\frac{\partial \log p(C,\hat{f})}{\partial C}$ is obtained by applying reverse transformation with pose parameters $\mathbf{p}$ to the force $\frac{\partial \log p(\tilde{C},\hat{f})}{\partial \tilde{C}}$

Algorithm 1 Segmentation Using Nonparametric Joint Shape and Feature Priors

1: Initialize ϕ_C
2: **for** $t = 0 \rightarrow t_{data}$ **do** $\triangleright$ t_{data} : time when the data driven curve evolution converges
3: **if** $t = t'$ **then** $\triangleright$ t' : time when the feature is extracted
4: Align ϕ_C with the shapes in $\mathbf{C}$ into $\phi_{\tilde{C}}$.
5: Extract feature vector $\hat{f}$.
6: **end if**
7: Update ϕ_C with the data force given in Equation (11).
8: **end for**
9: **for** $t = t_{data} + 1 \rightarrow t_{converge}$ **do** $\triangleright$ $t_{converge}$: time when data + joint shape and feature priors driven curve evolution converges
10: Align ϕ_C with the shapes in $\mathbf{C}$ into $\phi_{\tilde{C}}$.
11: Compute the data force for ϕ_C using the Equation (11).
12: Compute the joint shape and feature force $\frac{\partial \log p(\tilde{C}, \hat{f})}{\partial \tilde{C}}$ as given in Equation (12).
13: Reverse the force $\frac{\partial \log p(\tilde{C}, \hat{f})}{\partial \tilde{C}}$ to its original pose $\frac{\partial \log p(C, \hat{f})}{\partial C}$ using the reverse pose parameters found in step 10.
14: Update ϕ_C with the sum of the data force computed in step 13 and the joint shape and feature force computed in step 12.
15: **end for**

(see line 13 in Algorithm 1). In other words, gradient of the shape and feature prior is computed for $\tilde{C}$, gradient force is reverse back to its original pose and the whole gradient update is performed. Note that the alignment process can be done intrinsically during the curve evolution as in [18] and [20]. We choose to perform this process explicitly as in [2].

Finally, the proposed segmentation method that exploits nonparametric joint shape and feature priors is given in Algorithm 1.

III. Experimental Results

In this section, we present experimental results on 4 different data sets using various discriminative class-related features. In the MNIST and the aircraft data sets, features are synthetically generated. The remaining 2 data sets, the Swedish leaf and the dendritic spines, are completely real data sets.

We compare the performance of the proposed approach with three different methods: Kim *et al.* [2], Foulonneau *et al.* [20] and Chen and Radke [22]. We obtain quantitative results by comparing segmentation results with ground truths using Dice scores [41] and Hausdorff distance [42]. Dice score takes values between 0 and 1 where 1 indicates a perfect match whereas low values of Hausdorff distance indicate better results.

A. MNIST Handwritten Digits Data Set

In this section, we present experimental results on 3 different settings of the MNIST handwritten digits data set. We use the shapes in the training set shown in Figure 2 in all

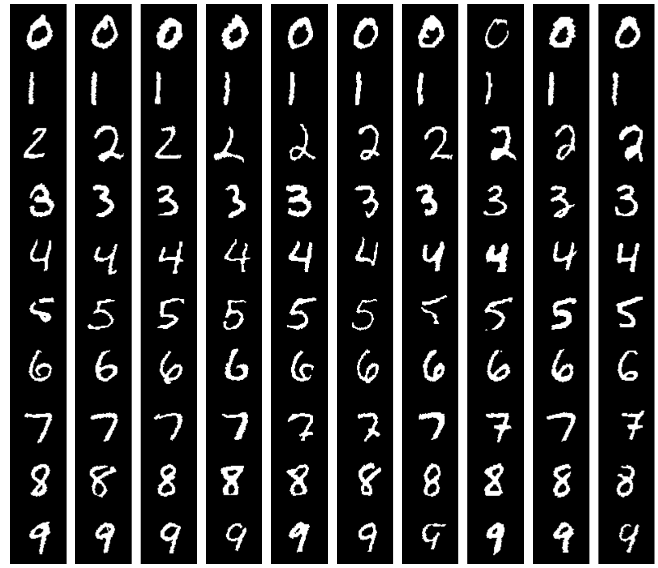


Fig. 2. Training set of shapes for the MNIST handwritten digits data set.

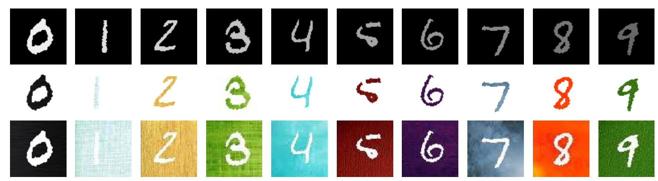


Fig. 3. Training sets that are used to obtain feature vectors. First row: the first training setting in which each digit class contains gray-level intensities drawn from a Gaussian distribution with different means in foreground region, second row: the second training setting in which each digit class contains different colors in foreground region, third row: the third training setting in which each digit class contains different colors in background region. Note that our training sets to obtain feature vectors contain 10 samples for each class and we display only one sample from each class for the sake of brevity.

experimental settings. Experimental settings differ from each other in terms of the feature vectors that are exploited for segmentation. This experiment demonstrates that our approach can learn effectively from a relatively small training data set. The approach could also exploit information in larger data sets when available.

In the first experimental setting of the MNIST data set, each training shape in Figure 2 is obtained from an intensity image which contains gray-level intensities drawn from a Gaussian distribution with different means for different classes in foreground regions. One exemplary intensity image from each digit class is shown in Figure 3. We estimate the mean intensity value in the foreground region using the corresponding intensity images of each training set. We use the mean values to form the training set of feature vectors $\mathbf{f}$. We perform experiments on the test images shown in the second row of Figure 4. In all test images, we first segment the apparent part of the object using only the data term (lines 2 - 8 in Algorithm 1). Then, the feature vector $\hat{f}$ is extracted as the mean intensity value in the foreground region of the initial segmentation. Note that, in this experimental setting the feature vector $\hat{f}$ and the feature vectors in $\mathbf{f}$ contain a scalar value. Also, note that the extracted feature value

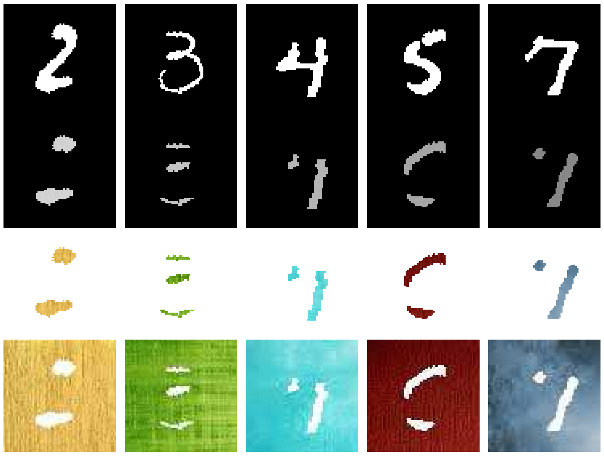


Fig. 4. Test images for the MNIST data set. First row: ground truth, second row: the first experimental setting, third row: the second experimental setting, fourth row: the third experimental setting.

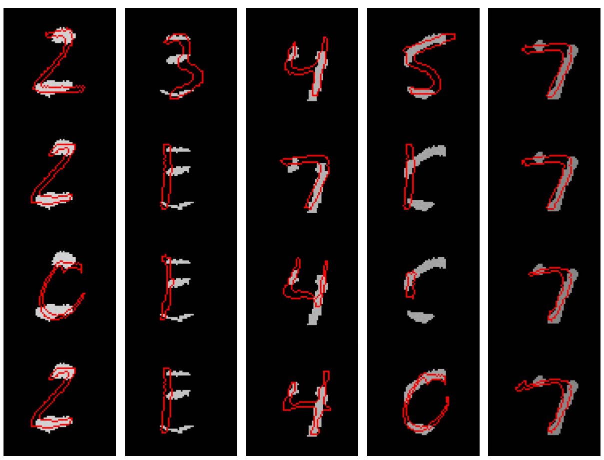


Fig. 5. Visual results of the first experimental setting of the MNIST data set. First row: the proposed method, second row: Kim *et al.* [2], third row: Foulonneau *et al.* [20], fourth row: Chen and Radke [22].

strongly depends on the data driven (initial) segmentation. Then, we keep evolving the curve using the nonparametric shape and feature priors together with the data term (lines 9 - 15 in Algorithm 1). We also perform experiments on the same test images using the approaches of Kim *et al.* [2], Foulonneau *et al.* [20] and Chen and Radke [22]. Visual segmentation results of all approaches are shown in Figure 5. The visual results demonstrate that the proposed approach generates segmentations that are closer to the ground truths whereas the other methods converges to a wrong mode of the posterior shape density in most test images. We also provide quantitative comparisons of the segmentation results with respect to ground truth using Dice score (see Table I) and Hausdorff distance (see Table II). The quantitative results with both metrics demonstrate the potential of the proposed approach.

In the second experimental setting of the MNIST data set, intensity images of the training shapes in Figure 2 contain different colors in foreground regions for different classes as shown in the second row of Figure 3. In this experiment, each feature vector is obtained by concatenating RGB histograms

TABLE I

DICE SCORE RESULTS ON THE FIRST EXPERIMENTAL SETTING OF THE MNIST DATA SET

Digit	2	3	4	5	7
Proposed	**0.6217**	**0.4341**	**0.7167**	**0.7906**	0.6809
Kim et al. [2]	0.5736	0.1771	0.4738	0.2294	**0.6870**
Foulonneau et al. [20]	0.3456	0.1814	0.6040	0.2298	0.6308
Chen et al. [22]	0.5736	0.1732	0.7042	0.4822	0.5915

TABLE II

HAUSDORFF DISTANCE RESULTS ON THE FIRST EXPERIMENTAL SETTING OF THE MNIST DATA SET

Digit	2	3	4	5	7
Proposed	8.000	**11.313**	**5.385**	**6.082**	**6.082**
Kim et al. [2]	**5.656**	20.000	13.601	20.000	7.000
Foulonneau et al. [20]	11.313	20.000	12.083	20.124	8.246
Chen et al. [22]	**5.656**	20.000	**5.385**	10.1980	7.211

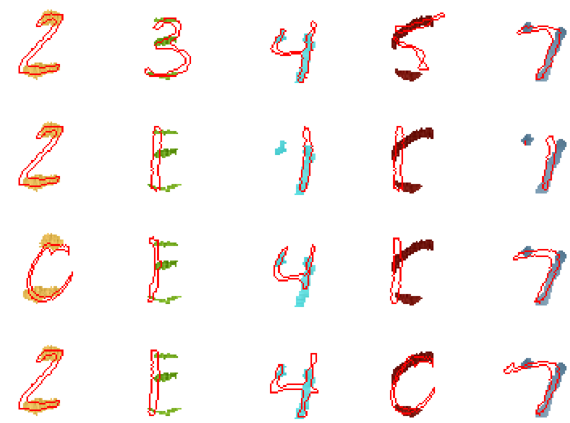


Fig. 6. Visual results of the second experimental setting of the MNIST data set. First row: the proposed method, second row: Kim *et al.* [2], third row: Foulonneau *et al.* [20], fourth row: Chen and Radke [22].

computed from the foreground region of the corresponding intensity image. All training feature vectors in $\mathbf{f}$ are constructed by following the same procedure. We use 5 test images shown in the third row of Figure 4 in this experiment. Similar to the previous experiment, we find the apparent part of the digits using only the data term. Then, we compute the RBG histograms from the intensities that lie inside the segmenting curve and form $\hat{f}$ by concatenating the histogram of each color channel. Then, we continue the curve evolution using our shape and feature-based segmentation approach. Visual segmentation results of the proposed approach and the all competing approaches are shown in Figure 6. We also provide the Dice score results in Table III and Hausdorff distance results in Table IV. The results clearly show the superiority of our approach with respect to other approaches.

Finally, in the third experimental setting, we design an experimental setting similar to the second one. In this setting, background regions contain different colors for each digit classes as shown in the third row of Figure 3. Similar to the second experimental setting, we construct $\mathbf{f}$ by exploiting the RGB histograms from the intensity images that correspond to background regions. We use the test images given in the fourth row of Figure 4. In all test images, once we

TABLE III

DICE SCORE RESULTS ON THE SECOND EXPERIMENTAL SETTING OF THE MNIST DATA SET

Digit	2	3	4	5	6
Proposed	**0.5790**	**0.5690**	**0.7458**	**0.5313**	**0.7032**
Kim et al. [2]	0.5736	0.1699	0.5492	0.2770	0.4751
Foulonneau et al. [20]	0.3446	0.1814	0.5743	0.2192	0.6570
Chen et al. [22]	0.5736	0.1732	0.7042	0.4891	0.5915

TABLE IV

HAUSDORFF DISTANCE RESULTS ON THE SECOND EXPERIMENTAL SETTING OF THE MNIST DATA SET

Digit	2	3	4	5	6
Proposed	**5.656**	**6.324**	**5.385**	14.317	8.246
Kim et al. [2]	**5.656**	19.416	23.194	18.248	10.816
Foulonneau et al. [20]	11.313	20.000	12.083	20.000	7.000
Chen et al. [22]	**5.656**	20.000	**5.385**	**10.198**	**7.211**

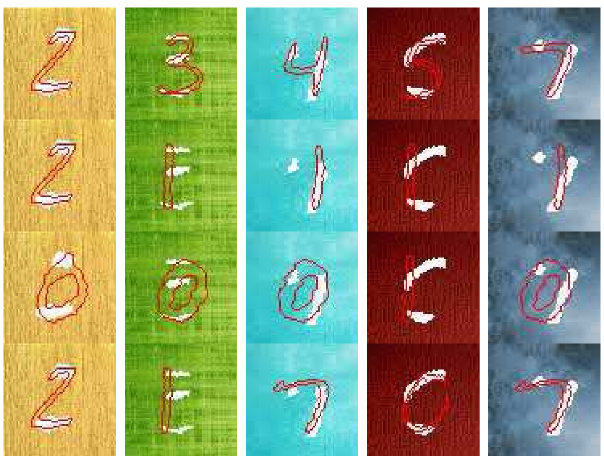


Fig. 7. Visual results of the third experimental setting of the MNIST data set. First row: the proposed method, second row: Kim *et al.* [2], third row: Foulonneau *et al.* [20], fourth row: Chen and Radke [22].

TABLE V

DICE SCORE RESULTS ON THE THIRD EXPERIMENTAL SETTING OF THE MNIST DATA SET

Digit	2	3	4	5	7
Proposed	**0.5736**	**0.5809**	**0.7093**	**0.5949**	**0.6779**
Kim et al. [2]	**0.5736**	0.1695	0.5510	0.2766	0.4388
Foulonneau et al. [20]	0.5018	0.4016	0.5490	0.2192	0.4889
Chen et al. [22]	0.5736	0.1732	0.4369	0.4822	0.5915

find the apparent boundaries using the data term, we extract $\hat{f}$ by computing the RGB histograms from the background region and concatenating them into a single feature vector. As in the above experiments, the proposed approach achieves better segmentation results than the approaches we compete both visually (see Figure 7) and quantitatively (see Tables V and VI).

B. The Swedish Leaf Data Set

In this section, we present evaluations of the proposed approach on the Swedish leaf data set [43]. The Swedish leaf data set contains leaf images obtained from 15 different tree classes. We choose two classes among them: Acer and Populus tremula. The data set is designed for classification purposes and it only contains RGB leaf images. We obtain binary images that are used for training by manually segmenting 10 leaf images from each class as shown in Figure 8. In order to construct a training set of feature vectors $\mathbf{f}$, we compute circularity of the boundaries in each binary training shape. Circularity of the boundary is a discriminative geometric feature for Acer and Populus tremula classes.

TABLE VI

HAUSDORFF DISTANCE RESULTS ON THE THIRD EXPERIMENTAL SETTING OF THE MNIST DATA SET

Digit	2	3	4	5	6
Proposed	**5.656**	**6.403**	**5.099**	**5.000**	**7.000**
Kim et al. [2]	**5.656**	19.416	23.086	18.248	20.223
Foulonneau et al. [20]	12.806	7.071	12.165	20.000	15.132
Chen et al. [22]	**5.656**	20.000	12.649	10.198	7.211

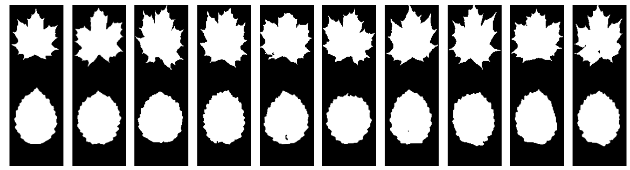


Fig. 8. Training set of shapes for the Swedish leaf data set. First row: Acer, second row: Populus tremula.

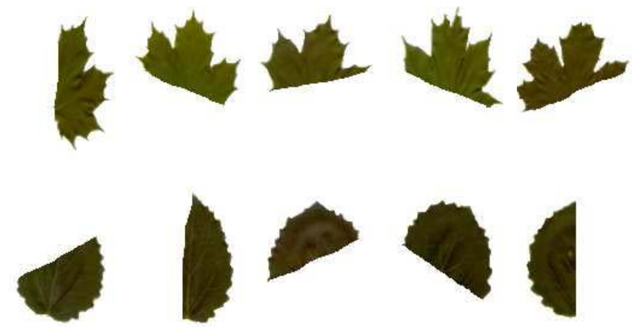


Fig. 9. Test images for the Swedish leaf data set. First row: Acer, second row: Populus tremula.

We perform experiments on 10 test leaf images (5 test images from each class and none of which is included in the training set), shown in Figure 9. Similar to the previous experiments, we find the apparent boundaries using only the data term and set $\hat{f}$ as the circularity of the boundary. Visual segmentation results of all approaches are shown in Figure 10. The visual results demonstrate that the approaches of Kim *et al.* [2] and Chen and Radke [22] tends to drive the segmenting curve toward a shape from Populus tremula class in all test images. Unlike Kim *et al.* [2] and Chen and Radke [22], the method of Foulonneau *et al.* [20] converges to the mode that corresponds to Acer class in all test images. With the aid of using the discriminative feature priors along with the shape priors, the proposed approach achieves segmentations from the correct mode of the shape density. The Dice score results are 0.9409 for the proposed method, 0.9456 for the method of Kim *et al.* [2], 0.9030 for the method of Foulonneau *et al.* [20] and 0.9335 for the method of Chen and Radke [22] on average of 10 test images. The average Hausdorff distance results are 10.5742 for the proposed

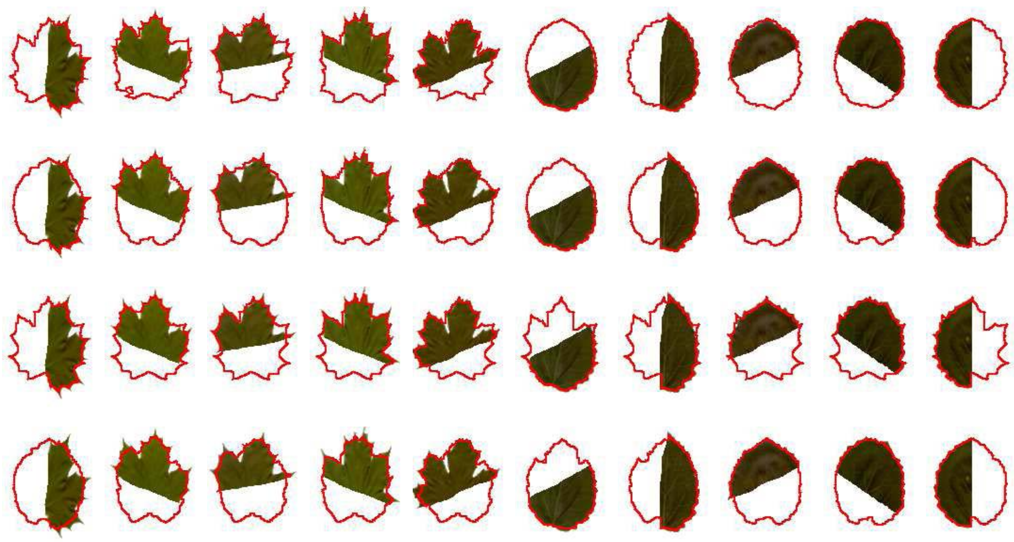


Fig. 10. Visual segmentation results on the Swedish leaf data set. First row: proposed method, second row: Kim *et al.* [2], third row: Foulonneau *et al.* [20], fourth row: Chen and Radke [22].

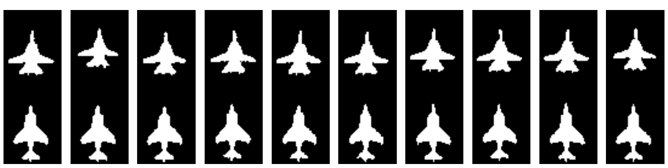


Fig. 11. The airplane data set. First row: F-14 wings opened, second row: Harrier.

method, 12.7036 for the method of Kim *et al.* [2], 17.0145 for the method of Foulonneau *et al.* [20] and 13.6214 for the method of Chen and Radke [22]. Note that Dice score results are close to each other even the competing methods produce segmentations from a wrong mode of the shape density. Since the shapes in different classes are very similar and Dice score measures the overlap between the segmentation and the ground truth, these results are expected. Hausdorff distance better quantifies the difference in the visual results in this experiment.

C. The Airplane Data Set

In this section, we evaluate the performance of our segmentation approach on the airplane data set [44]. The airplane data set contains 7 different airplane classes. In our experiments, we take a subset of two of them: F-14 wings opened and Harrier. We use 10 airplane shapes from each class for training as shown in Figure 11. Each airplane training shape in Figure 11 is obtained from an intensity image as shown in Figure 12. Note that, in Figure 12, airplane shapes from different classes contain different textural foreground regions. This means that textural features obtained from the foreground region can be discriminative class-dependent features for this data set. For each training shape, we extract 3 different textural features from the foreground region: correlation, energy, and homogeneity. We form each feature vector f_i in $\mathbf{f}$ by concatenating these values into a single vector.

We compare the performance of the proposed approach with Kim *et al.* [2], Foulonneau *et al.* [20] and Chen and Radke [22]

Fig. 12. Training set that are used to obtain the feature vectors. Note that each airplane shapes from different classes contain different textures.

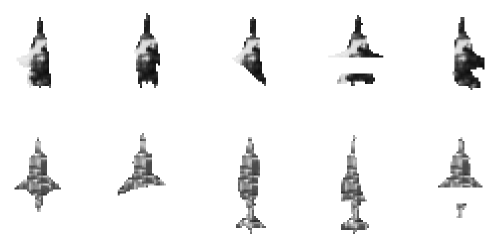


Fig. 13. Test images for airplane data set. First row: F-14 wings opened, second row: Harrier.

on 10 test images shown in Figure 13. Note that the test images are not included in the training set. When segmenting test images, we extract three textural features (correlation, energy, and homogeneity) after the data driven segmentation and concatenate into a single vector $\hat{f}$. Visual segmentation results on the airplane data set are shown in Figure 14. According to the visual results, the proposed approach drives the segmenting curve toward the correct mode of the shape density in all test images. When the tail of an Harrier type airplane is occluded, it looks more similar to the F-14 wings opened airplane type. In such cases, Kim *et al.* [2], Foulonneau *et al.* [20] and Chen and Radke [22] converges to a F-14 wings opened type airplane. Such results can be observed in the first, the second and the fifth test images of the Harrier class. The average Dice score (Hausdorff distance) results on all test images with respect to ground truths are 0.9153 (1.7899) for the proposed

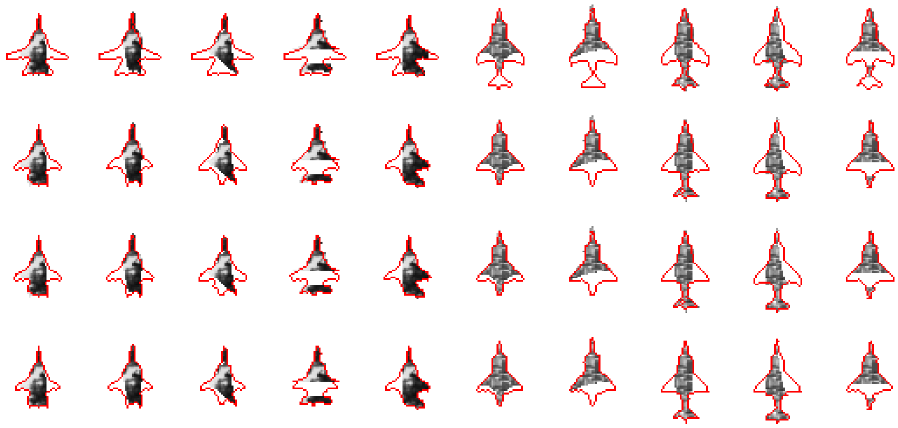
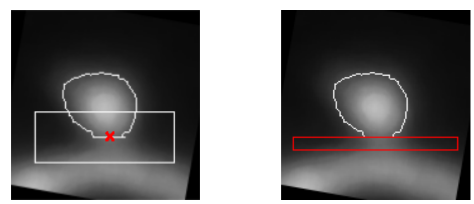


Fig. 14. Visual segmentation results on the airplane data set. First row: proposed method, second row: Kim *et al.* [2], third row: Foulonneau *et al.* [20], fourth row: Chen and Radke [22].

Fig. 15. Training set for dendritic spine data set. The first 8 spines from the left are mushroom and the remainings are stubby.

(a) (b)

Fig. 16. Intensity and corresponding manually annotated binary image examples from each spine class. From left to right: Mushroom, Stubby, Thin, and Filopodia. (a) Intensity images. (b) Manual Segmentations.



(a) (b)

Fig. 17. Regions where a potential neck is likely to be located. (a) First region. (b) Second region.

method, 0.8746 (6.1726) for the method of Kim *et al.* [2], 0.8762 (5.9271) for the method of Foulonneau *et al.* [20] and 0.8748 (6.5479) for the method of Chen and Radke [22]. The quantitative results indicate the positive effect of using additional class-dependent features along with the shape prior.

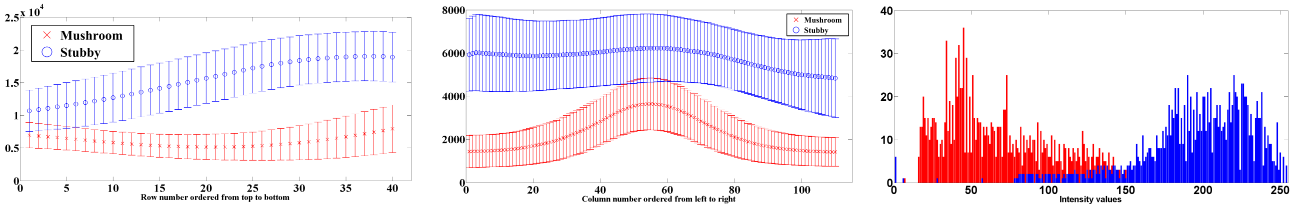
D. The Dendritic Spine Data Set

In this section, we present experimental results on a dendritic spine data set. The data set is obtained from Neuronal Structure and Function laboratory of Champalimaud Neuroscience Foundation, Lisbon.

In the literature, dendritic spines are generally grouped into four classes: mushroom, thin, stubby, and filopodia (see Figure 16). In our experiments, we use training samples from mushroom and stubby classes. The dendritic spine data set contains 88 mushroom and 27 stubby 2D spine intensity images together with the expert's manual segmentations. In our experiments, we use 8 mushroom and 8 stubby dendritic spine shapes shown in Figure 15 for training and the remaining 80 mushroom and 19 stubby spines for testing. We perform two different types of experiments with the dendritic spine data set; one is by using appearance-based and the other is by using geometric features. We also compare the segmentation performance of our approach with the approaches of Kim *et al.* [2], Foulonneau *et al.* [20] and Chen and Radke [22].

Spine neck is an important feature that helps to distinguish mushroom and spine classes. Spine head is common for spines in both classes and can be segmented roughly only using the information obtained from the data [36]. Given that spine neck is located in the area below the spine head if it exists, we can extract both appearance and geometric features exploiting the information in this region. We explain how to extract both types of features below:

First, we describe our appearance-based features. Intensity profiles below the spine head provides distinguishable features for spines from different classes [36]. First, we grab a rectangular region such that the bottom point of the spine head (shown by a red cross in Figure 17(a)) lies at the center of the rectangle. The second rectangular region shown in Figure 17(b)) is drawn such that it is located just below the spine head. We fix the size of the first and the second rectangles to 40×110 and 10×130, respectively, in a 150×150 ROI. Using these two rectangular regions, we construct three sets of feature vectors from the training set for classification. The first set of feature vectors are obtained by



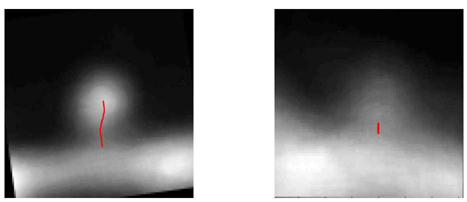
(a) (b) (c)

Fig. 18. Visualization of different sets of appearance-based feature vectors. Red indicates mushroom and blue indicates stubby spines. (a) Statistics (mean ± one standard deviation) of the first feature vector based on training data. (b) Statistics (mean ± one standard deviation) of the second feature vector based on training data. (c) Mean of the third feature vector based on training data.

TABLE VII

AVERAGE DICE SCORE AND HAUSDORFF DISTANCE RESULTS ON 99 DENDRITIC SPINES

	Proposed method w/ apperance-based feature priors	Proposed method w/ geometric feature priors	Kim et al. [2]	Foulonneau et al. [20]	Chen et al. [22]
Dice Score	**0.7492**	0.7474	0.6424	0.7348	0.7238
Hausdorff Distance	**19.2002**	20.7133	31.1413	26.0494	25.6581



(a) (b)

Fig. 19. Computed neck paths for a mushroom and a stubby spine are shown in red. (a) Mushroom. (b) Stubby.

summing up the intensities in the first rectangle horizontally. Similarly, the second set of feature vectors are obtained by vertical summation of the intensities in the same rectangle. We present the statistics of these two feature vectors extracted from the training set for each class in Figure 18(a) and 18(b). In these figures, error bars indicate one standard deviation around the mean. The final set of feature vectors are the histograms of intensities in the second rectangular region. We present average of these histograms for each spine class in Figure 18(c). Visual inspection of these feature vectors indicate that they contain discriminatory information about the spine class. Once we extract these three feature vectors from the corresponding intensity images of each training shape, a feature vector f_i is obtained by concatenating them. $\hat{f}$ is also extracted by exploiting the intensity information in the rectangular regions shown in Figure 17 as mentioned above. The final segmentation is obtained by evolving the segmenting curve with the data and the shape and feature priors terms.

Next, we describe the geometric features we use in spine segmentation. Spine neck length is an important geometric feature for identifying different spine classes [45]. In order to compute spine neck length, we follow a procedure consisting of multiple steps. First, we apply Otsu thresholding [46] to get a rough segmentation of the dendritic branch part (the part where the spine is connected to) and apply a fast marching distance transform [47] on this rough segmentation to compute the medial axis of the dendrite. Dendrite segmentation is refined by applying a locally adaptive sized disk-shaped structuring element around the medial axis of the dendrite to remove the spines. Once the head of the spine of interest is segmented, a fast marching algorithm [47] computes paths from the center of the spine head to a number of candidate target locations on segmented dendrite through the spine neck. This results in a neck path for each target location. Further, we apply three constraints to select the neck path from these candidate paths. These constraints are: neck path length, path complexity (L_1-norm of path derivatives), and path smoothness (L_1-norm of image intensities along the path). We select the neck path that has collectively the lowest value for these three constraints. Computed neck paths for a mushroom and a stubby spine are presented in Figure 19. Note that the computed neck path starts from the center of the spine head. Therefore, for correct computation of the neck length, we have to remove the path part that lies in the spine head. To achieve this, we first compute the radius of the spine head, r, by fitting a circle using the Hough Circle Transform on spine head segmentation and subtract it from the length of the computed path [45]. We compute the neck length for each training shape to form $\mathbf{f}$. When segmenting a test image, we compute the neck length into $\hat{f}$ in the same manner.

Some visual results that are obtained using the proposed approach (both for appearance-based and geometric features) and the other competing approaches are shown in Figure 20. We also evaluate the performance of these segmentation methods quantitatively using Dice score and Hausdorff distance. The average of both Dice score and Hausdorff distance results of all methods are shown in Table VII. In all experiments, the best and the second best quantitative results are obtained by the proposed approach with appearance-based feature priors and geometric feature priors, respectively.

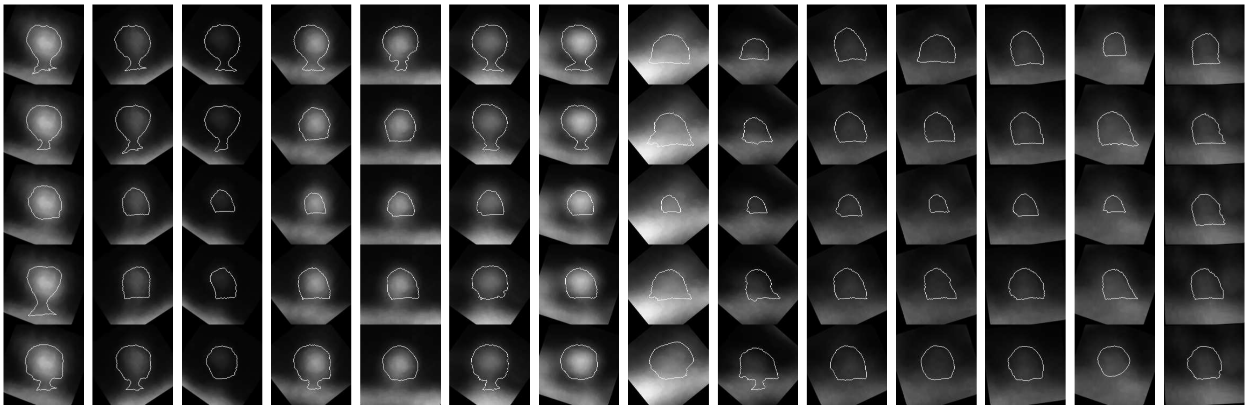


Fig. 20. Visual segmentation results on the dendritic spine data set. First row: proposed method with appearance-based feature priors, second row, proposed method with geometric feature priors, third row: Kim *et al.* [2], fourth row: Foulonneau *et al.* [20], fifth row: Chen and Radke [22]. Note that the first 7 spines from the left is mushroom, the remaining are stubby spines.

IV. Conclusion

We have proposed a segmentation method that exploits joint nonparametric shape and feature priors. The proposed method minimizes an energy function that includes a joint nonparametric shape and feature priors term together with the data term using level sets and gradient descent. We provide experimental results on a variety of real and synthetic data sets involving multimodal and complex shape density estimation problems. Experimental results demonstrate that the proposed algorithm achieves better segmentations than the state-of-the-art approaches that use nonparametric shape priors and can be applied to different data sets from various domains.

One possible future direction of the proposed method might be developing a similar approach by using a different shape representation than level sets, e.g. Disjunctive Normal Shape Models [25], [48]. Our approach can also be modified slightly and be used as a joint segmentation and classification approach. To this end, classes (perhaps corresponding to modes in the shape density) may be inferred during the segmentation phase and this probabilistic inference may then be used to update the weights of the training samples to drive the segmentation.

Appendix

A. Gradient Flow of Joint Shape and Feature Density

In this section, we provide the details on how we derive gradient of Equation (10) and obtain Equation (12). Note that the derivation is a straightforward extension of the derivation in [2].

Let us consider the log of the joint shape and feature prior density

$$\log p(\tilde{C}, \hat{f}) = \log\Big(\frac{1}{n}\sum_{i=1}^{n} k(d_T(\tilde{C}, C_i), d_{L_2}(\hat{f}, f_i), \sigma_C, \sigma_f)\Big). \tag{13}$$

Then, the derivative of $\log p(\tilde{C}, \hat{f})$ with respect to $\tilde{C}$ is written in the following form

$$\begin{aligned}\frac{\partial \log p(\tilde{C}, \hat{f})}{\partial \tilde{C}} &= -\frac{1}{p(\tilde{C}, \hat{f})} \times \frac{1}{n} \times \frac{1}{\sigma_x \times \sigma_y} \\ &\quad \times \sum_{i=1}^{n} \Big(k(d_T(\tilde{C}, C_i), d_{L_2}(\hat{f}, f_i), \sigma_C, \sigma_f) \\ &\quad \times d_T(\tilde{C}, C_i) \times (d_{L_2}(\hat{f}, f_i))^2 \times \frac{\partial d_T(\tilde{C}, C_i)}{\partial \tilde{C}}\Big).\end{aligned} \tag{14}$$

Now the task comes to computing $\frac{\partial d_T(\tilde{C}, C_i)}{\partial \tilde{C}}$. Consider the template distance metric $d_T(\phi_{\tilde{C}}, \phi_{C_i}) = Area(inside(\tilde{C}) \triangle inside(C_i))$ where $\triangle$ denotes the set symmetric difference. This metric can be written in the form of region integrals as follows [2]

$$\begin{aligned}d_T(\phi_{\tilde{C}}, \phi_{C_i}) &= \int_{\Omega} (1 - H(\phi_{\tilde{C}}(x))) H(\phi_{C_i}(x)) \mathrm{d}x \\ &\quad + \int_{\Omega} H(\phi_{\tilde{C}}(x))(1 - H(\phi_{C_i}(x))) \mathrm{d}x \\ &= \int_{inside(\tilde{C})} H(\phi_{C_i}(x)) \mathrm{d}x \\ &\quad + \int_{outside(\tilde{C})} (1 - H(\phi_{C_i}(x))) \mathrm{d}x\end{aligned} \tag{15}$$

For the region integrals in Equation (15), the derivative is well known [49], which is given by

$$\frac{\partial d_T(\tilde{C}, C_i)}{\partial \tilde{C}} = (2H(\phi_{C_i}) - 1). \tag{16}$$

By plugging Equation (16) into Equation (14), we obtain the gradient flow of $\log p(\tilde{C}, \hat{f})$ with respect to $\tilde{C}$:

$$\begin{aligned}\frac{\partial \log p(\tilde{C}, \hat{f})}{\partial \tilde{C}} &= \frac{1}{p(\tilde{C}, \hat{f})} \times \frac{1}{n} \times \frac{1}{\sigma_x \times \sigma_y} \\ &\quad \times \sum_{i=1}^{n} \Big(k(d_T(\tilde{C}, C_i), d_{L_2}(\hat{f}, f_i), \sigma_C, \sigma_f) \\ &\quad \times d_T(\tilde{C}, C_i) \times (d_{L_2}(\hat{f}, f_i))^2 \times (1 - 2H(\phi_{C_i})\Big).\end{aligned} \tag{17}$$

References

[1] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.

[2] J. Kim, M. Çetin, and A. S. Willsky, "Nonparametric shape priors for active contour-based image segmentation," *Signal Process.*, vol. 87, no. 12, pp. 3021–3044, 2007.

[3] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Comput. Vis. Image Understand.*, vol. 61, no. 1, pp. 38–59, 1995.

[4] B. V. Ginneken, A. F. Frangi, J. J. Staal, B. M. T. H. Romeny, and M. A. Viergever, "Active shape model segmentation with optimal features," *IEEE Trans. Med. Imag.*, vol. 21, no. 8, pp. 924–933, Aug. 2002.

[5] S. Milborrow and F. Nicolls, "Locating facial features with an extended active shape model," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2008, pp. 504–513.

[6] M. de Bruijne, B. van Ginneken, M. A. Viergever, and W. J. Niessen, "Adapting active shape models for 3D segmentation of tubular structures in medical images," in *Information Processing in Medical Imaging*. Berlin, Germany: Springer, 2003, pp. 136–147.

[7] W. Wang, S. Shan, W. Gao, B. Cao, and B. Yin, "An improved active shape model for face alignment," in *Proc. IEEE Int. Conf. Multimodal Interfaces*, Oct. 2002, p. 523.

[8] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Med. Image Anal.*, vol. 13, no. 4, pp. 543–563, 2009.

[9] T. F. Cootes and C. J. Taylor, "A mixture model for representing shape variation," *Image Vis. Comput.*, vol. 17, no. 8, pp. 567–573, 1999.

[10] P. Etyngier, F. Segonne, and R. Keriven, "Shape priors using manifold learning techniques," in *Proc. 11th Int. Conf. Comput. Vis. (ICCV)*, Oct. 2007, pp. 1–8.

[11] M. Kirschner, M. Becker, and S. Wesarg, "3D active shape model segmentation with nonlinear shape priors," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*. Berlin, Germany: Springer, 2011, pp. 492–499.

[12] S. Dambreville, Y. Rathi, and A. Tannenbaum, "A framework for image segmentation using shape models and kernel space shape priors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 8, pp. 1385–1399, Aug. 2008.

[13] S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. N. Metaxas, and X. S. Zhou, "Towards robust and effective shape modeling: Sparse shape composition," *Med. Image Anal.*, vol. 16, no. 1, pp. 265–277, Jan. 2012.

[14] A. Tsai *et al.*, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 137–154, Feb. 2003.

[15] X. Bresson, P. Vandergheynst, and J.-P. Thiran, "A variational model for object segmentation using boundary information and shape prior driven by the mumford-shah functional," *Int. J. Comput. Vis.*, vol. 68, no. 2, pp. 145–162, 2006.

[16] O. Amadieu, E. Debreuve, M. Barlaud, and G. Aubert, "Inward and outward curve evolution using level set method," in *Proc. Int. Conf. Image Process. (ICIP)*, vol. 3. Oct. 1999, pp. 188–192.

[17] S. Jehan-Besson, M. Barlaud, and G. Aubert, "DREAM^{2}S: Deformable regions driven by an Eulerian accurate minimization method for image and video segmentation," *Int. J. Comput. Vis.*, vol. 53, no. 1, pp. 45–70, 2003.

[18] D. Cremers, S. J. Osher, and S. Soatto, "Kernel density estimation and intrinsic alignment for shape priors in level set segmentation," *Int. J. Comput. Vis.*, vol. 69, no. 3, pp. 335–351, 2006.

[19] M. G. Uzunbaş *et al.*, "Coupled nonparametric shape and moment-based intershape pose priors for multiple basal ganglia structure segmentation," *IEEE Trans. Med. Imag.*, vol. 29, no. 12, pp. 1959–1978, Dec. 2010.

[20] A. Foulonneau, P. Charbonnier, and F. Heitz, "Multi-reference shape priors for active contours," *Int. J. Comput. Vis.*, vol. 81, no. 1, pp. 68–81, Jan. 2009.

[21] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[22] S. Chen and R. J. Radke, "Level set segmentation with both shape and intensity priors," in *Proc. 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 763–770.

[23] R. Yang, M. Mirmehdi, X. Xie, and D. Hall, "Shape and appearance priors for level set-based left ventricle segmentation," *IET Comput. Vis.*, vol. 7, no. 3, pp. 170–183, 2013.

[24] A. Soğanli, M. G. Uzunbaş, and M. Çetin, "Combining learning-based intensity distributions with nonparametric shape priors for image segmentation," *Signal, Image Video Process.*, vol. 8, no. 4, pp. 789–798, 2014.

[25] F. Mesadi, M. Çetin, and T. Taşdizen, "Disjunctive normal shape and appearance priors with applications to image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent*, 2015, pp. 703–710.

[26] E. Erdil, S. Yildirim, M. Çetin, and T. Taşdizen, "MCMC shape sampling for image segmentation with nonparametric shape priors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 411–419.

[27] F. Mesadi, M. Çetin, and T. Taşdizen, "Disjunctive normal parametric level set with application to image segmentation," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2618–2631, Jun. 2017.

[28] D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape," *Int. J. Comput. Vis.*, vol. 72, no. 2, pp. 195–215, 2007.

[29] B. Wang, X. Gao, J. Li, X. Li, and D. Tao, "A level set method with shape priors by using locality preserving projections," *Neurocomputing*, vol. 170, pp. 188–200, Dec. 2015.

[30] Y. Kihara, M. Soloviev, and T. Chen, "In the shadows, shape priors shine: Using occlusion to improve multi-region segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 392–401.

[31] L. A. Royer, D. L. Richmond, C. Rother, B. Andres, and D. Kainmueller, "Convexity shape constraints for image segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2016, pp. 402–410.

[32] D. Cremers, N. Sochen, and C. Schnörr, "A multiphase dynamic labeling model for variational recognition-driven image segmentation," *Int. J. Comput. Vis.*, vol. 66, no. 1, pp. 67–81, 2006.

[33] T. Chan and W. Zhu, "Level set based shape prior segmentation," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2. Jun. 2005, pp. 1164–1170.

[34] E. Erdil *et al.*, "Nonparametric joint shape and feature priors for segmentation of dendritic spines," in *Proc. IEEE 13th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2016, pp. 343–346.

[35] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Commun. Pure Appl. Math.*, vol. 42, no. 5, pp. 577–685, 1989.

[36] E. Erdil, A. Ö. Argunşah, T. Taşdizen, D. Ünay, and M. Çetin, "A joint classification and segmentation approach for dendritic spine segmentation in 2-photon microscopy images," in *Proc. 12th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2015, pp. 797–800.

[37] J. Kim, J. W. Fisher, A. Yezzi, M. Çetin, and A. S. Willsky, "A nonparametric statistical method for image segmentation using information theory and curve evolution," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1486–1502, Oct. 2005.

[38] N. Houhou, J.-P. Thiran, and X. Bresson, "Fast texture segmentation model based on the shape operator and active contour," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2008, pp. 1–8.

[39] O. Michailovich, Y. Rathi, and A. Tannenbaum, "Image segmentation using active contours driven by the Bhattacharyya gradient flow," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2787–2801, Nov. 2007.

[40] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, vol. 26. Boca Raton, FL, USA: CRC Press, 1986.

[41] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.

[42] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, vol. 317. Berlin, Germany: Springer, 2009.

[43] O. Söderkvist, "Computer vision classification of leaves from swedish trees," M.S. thesis, Linkoping Univ., 2001.

[44] N. Thakoor, J. Gao, and S. Jung, "Hidden Markov model-based weighted likelihood discriminant for 2-D shape classification," *IEEE Trans. Image Process.*, vol. 16, no. 11, pp. 2707–2719, Nov. 2007.

[45] M. U. Ghani, S. D. Kanik, A. Ö. Argunşah, T. Taşdizen, D. Ünay, and M. Çetin, "Dendritic spine shape classification from two-photon microscopy images," in *Proc. 23rd Signal Process. Commun. Appl. Conf.*, May 2015, pp. 939–942.

[46] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, nos. 285–296, pp. 23–27, 1975.

[47] M. S. Hassouna and A. A. Farag, "MultiStencils fast marching methods: A highly accurate solution to the Eikonal equation on Cartesian domains," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 9, pp. 1563–1574, Sep. 2007.

[48] N. Ramesh, F. Mesadi, M. Çetin, and T. Taşdizen, "Disjunctive normal shape models," in *Proc. 12th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2015, pp. 1535–1539.

[49] S. C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 9, pp. 884–900, Sep. 1996.