

Adaptive Estimation of Active Contour Parameters Using Convolutional Neural Networks and Texture Analysis

Assaf Hoogi,* Arjun Subramaniam, Rishi Veerapaneni, and Daniel L. Rubin

Abstract—In this paper, we propose a generalization of the level set segmentation approach by supplying a novel method for adaptive estimation of active contour parameters. The presented segmentation method is fully automatic once the lesion has been detected. First, the location of the level set contour relative to the lesion is estimated using a convolutional neural network (CNN). The CNN has two convolutional layers for feature extraction, which lead into dense layers for classification. Second, the output CNN probabilities are then used to adaptively calculate the parameters of the active contour functional during the segmentation process. Finally, the adaptive window size surrounding each contour point is re-estimated by an iterative process that considers lesion size and spatial texture. We demonstrate the capabilities of our method on a dataset of 164 MRI and 112 CT images of liver lesions that includes low contrast and heterogeneous lesions as well as noisy images. To illustrate the strength of our method, we evaluated it against state of the art CNN-based and active contour techniques. For all cases, our method, as assessed by Dice similarity coefficients, performed significantly better than currently available methods. An average Dice improvement of 0.27 was found across the entire dataset over all comparisons. We also analyzed two challenging subsets of lesions and obtained a significant Dice improvement of 0.24 with our method ($p < 0.001$, Wilcoxon).

Index Terms—Active contours, adaptive parameters, convolutional neural network, image segmentation.

I. INTRODUCTION

MEDICAL image segmentation is an important step in clinical diagnosis, tumor growth prediction, and radiotherapy planning. The rapid growth of medical data motivates the need for automated, computer-based segmentation methods in place of manual tracing.

Level set models are a popular method used widely in curve evolution applications, specifically for medical image

segmentation [1]–[11]. These methods follow a non-parametric deformable model, generate a continuous boundary of an object, and are able to handle variations in shape, image noise, heterogeneity, and discontinuous object boundaries [7]. Such challenging characteristics are common in medical images.

Contour initialization is an important factor in the accuracy of local active contour models. A more accurate initial contour, closer to the object to be segmented, leads to lower dependence of the segmentation on level set parameters. Manual drawing by a trained user is currently the most common method for contour initialization. However, manual initialization is an extremely tedious and error-prone process. While approaches that simplify the manual initialization process have been suggested in [8]–[10], other researchers have proposed automatic initialization methods that analyze the external force field [11]–[14]. Initialization methods can utilize active shape models, but these methods need a large and representative training set [15]–[18]. Other methods use fuzzy clustering to facilitate level set segmentation [19]–[22]. Such algorithms employ fuzzy clustering based on image intensities for an initial segmentation and use level set methods for object refinement by tracking boundary variation.

A. Related Work

Li et al. [19] show that an inappropriate choice of level set parameters may lead to an inferior segmentation regardless of initialization. Most often, level set segmentation requires pre-defined parameters of the energy functional, while local frameworks also require a pre-defined window size [23]–[24]. Each optional value for a specific parameter is tested over a series of images and remains the same for the entire database of images. New images having different spatial statistics may require additional experiments to find the best-fitted parameter set, local window size and energy functional parameters. As a result, choosing a fixed set of parameters by trial and error is a time consuming and laborious process. Moreover, when the images contain substantial diversity of spatial statistics, pre-defining a set of fixed parameters may result in non-optimal segmentation performance for all images. In addition, the user does not typically have the experience or time to tune a large number of parameters optimally. For this reason, an adaptive solution is highly desirable. A varied set of parameters defined according to spatial information has a higher chance of providing a more accurate segmentation.

A few works present a choice of one local window size surrounding each contour point from a range of optional

Manuscript received September 16, 2016; revised October 27, 2016; accepted October 28, 2016. Date of publication November 11, 2016; date of current version March 2, 2017. This project was supported by the National Cancer Institute, National Institutes of Health, under Grants U01CA142555, 1U01CA190214, and 1U01CA187947. Arjun Subramaniam and Rishi Veerapaneni contributed equally. Asterisk indicates corresponding author.

*A. Hoogi is with the Departments of Biomedical Data Science, Radiology, and Medicine (Biomedical Informatics Research), Stanford University, Stanford, CA 94305 USA (e-mail: ahoogi@stanford.edu).

D. L. Rubin, A. Subramaniam and R. Veerapaneni are with the Departments of Biomedical Data Science, Radiology, and Medicine (Biomedical Informatics Research), Stanford University, Stanford, CA 94305 USA (e-mail: dlrubin@stanford.edu; arjunsubway@gmail.com; rishiveerapaneni@gmail.com).

Digital Object Identifier 10.1109/TMI.2016.2628084

sizes [25]–[27]. In our previous work [28], we presented a novel method to define the local window adaptively with no dependence on a range of input sizes.

Other works provide an algorithm to estimate parameters for the level set energy functional. Li et al. [19] propose a fuzzy level set algorithm in which the level set parameters are initialized using a spatial fuzzy clustering approach. However, the parameters are only evaluated at the beginning of the segmentation process and remain constant throughout the whole process. In addition, the performance quality of fuzzy C-means is sensitive to noise, resulting in generally poorer segmentation. Oliveira et al. [29] present a solution for liver segmentation based on a deformable model, in which parameters are adjusted via a genetic algorithm. The genetic algorithm was used to choose the best combination of parameters from analysis of the training set, but all segmentations in the test dataset were conducted using the same parameters. Thus, this method may not be appropriate for highly diverse types of lesions. The authors in [29] also made two assumptions in their analysis: 1) the initialization is reasonably accurate; and 2) the liver is spatially homogeneous. Moreover, the authors use their method to segment the liver itself rather than a lesion dataset. The diversity of screened organs is typically much lower than the diversity that characterizes lesions. Baillard et al. [30] define the problem of parameter tuning as a classification of each point along the contour. That is, if a point belongs to the object, then the surface should locally extend, and if not, the surface should locally contract. This classification is performed by maximizing the posterior segmentation probability [30]. However, the authors only consider the direction of the curve evolution and not its magnitude, which is critical especially in heterogeneous regions, wherein convergence into local minima should be prevented. Thus, both [29] and [30] are likely to have limited performance for highly diverse datasets, given the limited amount of information that is incorporated.

In this paper, we propose a significant improvement of the level set segmentation method. We present an adaptive method to estimate the parameters for the level set energy functional separately for each case and over iterations. When combined with estimation of an adaptive window size surrounding each contour point as suggested in our previous paper [28], we supply a generalization of the segmentation process, applying the same model equations and deep learning architecture for any given dataset. Our method is a multi-stage process. First, we provide a novel method to estimate the parameters of the energy functional. A convolutional neural network (CNN) is used to identify the location of the zero level set contour in relation to the lesion. The output probabilities of the CNN are then used to calculate the level set parameters. Second, the adaptive window size is re-estimated by an iterative process that considers the scale of the lesion, local and global texture statistics, and minimization of the cost function over iterations. Our method requires only a single input point representing the approximate center of the detected lesion. There is no need of a more accurate initial contour as is typically supplied, automatically or manually, for local analysis. Contrary to current local active contour frameworks, our method has little

to no dependence on accurate initialization and does not include any assumptions about lesion characteristics. Thus, it may perform well with highly diverse datasets that include low contrast, noisy, and heterogeneous lesions. Here, we study the effects of adaptive parameters on segmentation performance, demonstrating the capabilities of our method by analyzing 276 images of liver lesions generated using two different modalities, MRI and CT. Lesions with substantially different spatial texture were present in both MRI and CT images. Finally, we performed extensive comparisons with both active contour and machine learning techniques, which showed the strength of our method.

To the best of our knowledge, we are the first to develop a fully adaptive framework for deformable model segmentation, resulting in a far more general segmentation solution than methods available to date.

The paper is organized as follows. In Section II, we define the global and the local energies that are the basis for this work. Section III presents our method for evaluating the level set parameters. In Sections IV and V, we discuss key ideas regarding experimentation and implementation details. Analysis of the segmentation results and comparison with other methods are presented in Section VI. Finally, Section VII discusses the results and makes some concluding remarks.

II. BACKGROUND

We evaluated our proposed adaptive framework with two different energy models: the piecewise constant model and the mean separation model. Details of each model follow.

A. Piecewise Constant (PC) Model

The piecewise constant (PC) model presented in [3] assumes an image I to be formed by two areas of distinct intensities (the object area and the background area), with uniform intensities within each area. Let M_u be the mean object intensity and M_v be the mean background intensity of the image. We set Ω as a bounded subset in \mathbb{R}^2 and $I(x, y)$ as the coordinates of a point on image I . Let $\phi(x, y)$ be a signed distance map and ∇ be the first variation of the energy with respect to $\phi(x, y)$. Let C be a parameterized closed contour curve in Ω represented by the zero level set (ZLS), $C = \{(x, y) | \phi(x, y) = 0\}$. We use the Heaviside function,

$$H\phi(x, y) = \begin{cases} 1, & \phi(x, y) > \varepsilon \\ 0, & \phi(x, y) < -\varepsilon \\ h(x, y), & |\phi(x, y)| < \varepsilon \end{cases} \quad (1)$$

where ε is a constant and $h(x, y) = \frac{1}{2} \left\{ 1 + \frac{\phi(x, y)}{\varepsilon} + \frac{1}{\pi} \sin\left(\frac{\pi \phi(x, y)}{\varepsilon}\right) \right\}$.

Then, $H\phi(x, y)$ gives the interior region of C , and $(1 - H\phi(x, y))$ gives the exterior region. The narrow band around the ZLS contour C is represented by the derivative of $H\phi(x, y)$ and a smooth approximate version of Dirac delta $\delta\phi(x, y)$. Putting all of this together, the energy function for the piecewise constant model $F_{PC}(M_u, M_v, \phi)$ can then be

written as:

$$\begin{aligned}
 F_{PC}(M_u, M_v, \phi) &= \mu \int_{\Omega} \delta\phi(x, y) |\nabla\phi(x, y)| dx dy \\
 &+ \lambda_1 \int_{\Omega} (I(x, y) - M_u)^2 H\phi(x, y) dx dy \\
 &+ \lambda_2 \int_{\Omega} (I(x, y) - M_v)^2 (1 - H\phi(x, y)) dx dy \quad (2)
 \end{aligned}$$

where μ is a predefined constant that affects the smoothness of the curve, and λ_1 and λ_2 are weighting parameters that affect the direction and magnitude of the contour evolution. By applying a local version of the PC model, M_u and M_v can be replaced by their local equivalent terms, m_u and m_v . In that case, these terms will represent the local means of a region surrounding each contour point [24].

B. Mean Separation (MS) Model

The mean separation (MS) model in [31] assumes that the object and its background should have maximal separation between mean intensities. We define $\Omega_n \in \mathbb{R}^2$ as the local version of Ω that represents only the narrow-band points [24]. As in the PC model above, we use the Heaviside function to determine the area of the local interior (A_u) and exterior (A_v) regions surrounding a contour point as:

$$A_u = \int_{\Omega_n} H\phi(x, y) dx dy, \quad A_v = \int_{\Omega_n} (1 - H\phi(x, y)) dx dy \quad (3)$$

This gives us the following energy function for the local mean separation model $F_{MS}(m_u, m_v, \phi)$:

$$\begin{aligned}
 F_{MS}(M_u, M_v, \phi) &= \mu \int_{\Omega_n} \delta\phi(x, y) |\nabla\phi(x, y)| dx dy \\
 &+ \lambda_1 \int_{\Omega_n} \frac{(I(x, y) - m_u)^2}{A_u} H\phi(x, y) dx dy \\
 &+ \lambda_2 \int_{\Omega_n} \frac{(I(x, y) - m_v)^2}{A_v} (1 - H\phi(x, y)) dx dy \quad (4)
 \end{aligned}$$

The MS energy is minimized when the difference between m_u and m_v is maximized. In some cases, the MS model may provide better results than the PC model due to the focus on maximal contrast between the interior and the exterior regions, without any restrictions on region homogeneity. This allows the MS model to find image edges effectively without considering the uniformity of the internal or external regions.

III. THE PROPOSED METHOD

The proposed method involves a two-step, iterative process (Fig. 1) that supplies an Adaptive estimation of the active Contour Parameters (ACP) via machine learning - based evaluation.

A. Adaptive Energy Functional Parameters

Level set curve evolution depends on three weighting parameters. μ controls the smoothness of the contour and is fixed at $\mu = 0.1$ throughout the segmentation process.

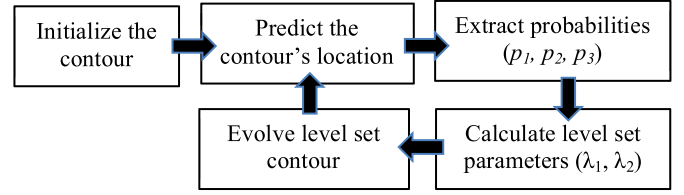


Fig. 1. The whole pipeline for learning the level set parameters.

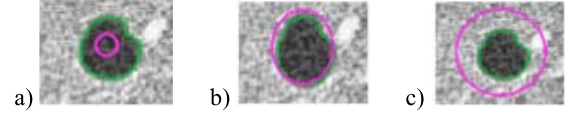


Fig. 2. Three optional locations of the ZLS contour. a) contour inside the lesion ($p_1 = 0.83, p_2 = 0.12, p_3 = 0.05$), b) contour near the lesion's boundaries ($p_1 = 0.06, p_2 = 0.78, p_3 = 0.16$), c) contour outside the lesion ($p_1 = 0.01, p_2 = 0.09, p_3 = 0.90$). Magenta is the ZLS contour, green is radiologist annotation.

The level set framework is relatively insensitive to changes in the value of μ . On the other hand, λ_1 and λ_2 play a key role in the direction and magnitude of curve evolution. Curve evolution depends not only on the absolute values of these parameters, but also on their ratio. First, a convolutional neural network (CNN) is used to estimate the location of the ZLS contour relative to the lesion. Three possible locations are considered: outside, near, or inside the lesion boundary. The CNN outputs a probability for each of three classes: inside the lesion and far from its boundaries (p_1), close to the boundaries of the lesion (p_2), or outside the lesion and far away from its boundaries (p_3) (Fig. 2). In the second step, we use the CNN probabilities of the three classes to set the weighting parameters λ_1 and λ_2 of the energy functional. These parameters are calculated using the following equations:

$$\lambda_1 = \exp\left(\frac{1 + p_2 + p_3}{1 + p_2 + p_1}\right), \quad \lambda_2 = \exp\left(\frac{1 + p_2 + p_1}{1 + p_2 + p_3}\right), \quad (5)$$

If $p_3 > p_1$, then $\lambda_1 > \lambda_2$ and the contour has a tendency to contract. Conversely, if $p_1 > p_3$, then $\lambda_2 > \lambda_1$, the third term of the energy eqs. 2 + 4 (difference of the point grey level from m_v) is weighted more heavily, and the contour tends to expand. Probability p_2 serves as a stabilizer and a restraining factor. That is, if the ZLS contour is located close to the lesion's boundaries, $p_2 \gg p_1, p_3$ and $\lambda_1 \approx \lambda_2$. As a result, both energies, related to the regions inside and outside the contour, are weighted equally. The exponential function is used to increase the range of values and ratios that λ_1 and λ_2 can take on.

1) CNN Architecture: The strength of the proposed method is its generalization to a variety of segmentation challenges. Therefore, we propose a generalized architecture for the CNN.

Our architecture consists of two convolutional layers followed by two fully connected layers, including the final three-node layer (Fig. 3). Both convolutional layers use 5×5 kernels, as this size outperformed other kernel sizes. We set the depths of the first and second convolutional layers to 20 and 100 kernels, respectively. The depths and

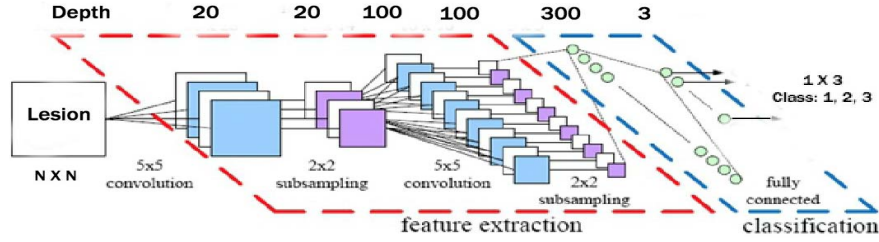


Fig. 3. CNN architecture. The input is an $N \times N$ image of a lesion masked by a ZLS contour. Two convolutional ‘blocks’ follow the input with 5×5 filters and 2×2 max pooling. Two fully connected (FC) layers with 300 and 3 nodes respectively follow the convolutional layers. The dimension of the feature map reduces from N to $\lfloor N/2 \rfloor - 2$ after each convolutional layer. FC layers take the output of the feature extraction layers

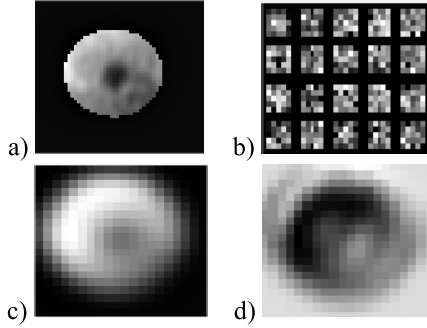


Fig. 4. Visualizing the convolutional layers of the CNN. (a) shows an input image masked by a ZLS contour that is located outside and far away from the lesion (p_3). (b) Learned filters during the first convolutional layer. (c) and (d) show outputs of different filter convolutions from the second convolutional layer.

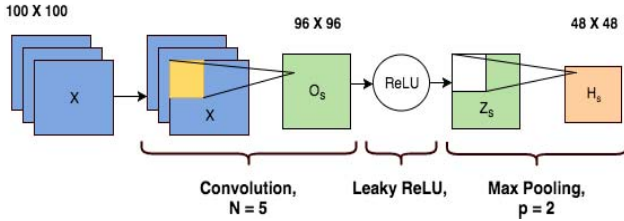


Fig. 5. A sample convolutional block with leaky ReLU and max pooling. The example input is a stack of images (100×100 pixels), which are then convolved by 5×5 kernels to create new feature maps (96×96 pixels). Leaky ReLU activation is applied after convolution and 2×2 max pooling reduces the size by a factor of 4.

filter size can be tuned if necessary, but these parameters were chosen based on examination of highly diverse lesion datasets. Figure 4 presents a visualization of the convolutional layers. The proposed method uses the raw images as feature maps. Each convolutional block of our CNN is composed of three layers (Fig. 5): a filter bank layer, a nonlinearity layer (Leaky Rectified Linear Units (ReLU)), and a feature-pooling layer.

A non-linear activation function is applied to the outputs of the convolutional layer [32]. The ReLU activation function $f(x) = \max(0, x)$ has recently become the gold standard for training deep neural networks due to its fast convergence [32]. Glorot et al. [33] argued that the ReLU function’s hard zero limit mimics the sparse activation of neurons in the brain. However, ReLU units are potentially disadvantageous because early zeroing can mean that a particular neuron will not activate for the remainder of training. Leaky ReLU units

alleviate this problem by allowing a small, non-zero gradient if the unit is not active:

$$f(x) = \begin{cases} x, & x > 0 \\ 0.01x, & \text{else} \end{cases} \quad (6)$$

Leaky ReLU activation reduces over-fitting and leads to lower error rates while training.

Max pooling is the final step of the convolutional block. We take a 2×2 pixel sub-region from the Leaky ReLU output and represent the region by its maximum value, thus reducing the dimensionality of the dataset. Although we experimented with removing the pooling layers from the network, we did not see a noticeable improvement in performance at the cost of increased training time.

2) Training the CNN: The CNN attempts to minimize the following cost function:

$$L(w) = \frac{1}{n} \sum_{i=1}^n \ell(z, f(x, w)) + \frac{\eta}{2} \|w\|^2, \quad (7)$$

for a labeled training set $(x_1, z_1), \dots, (x_n, z_n)$ and weights $w = (w_1, \dots, w_L)$, and some loss function ℓ . L2-regularization penalizes the size of the weights in the model, where $\eta = 0.005$ is the coefficient of regularization. Stochastic gradient descent with momentum is applied to update the weights (w_i) of the network [32]:

$$\begin{aligned} v_{i+1} &= \varsigma v_i - \alpha \nabla w_i - \eta w_i \\ w_{i+1} &= w_i + v_{i+1} \end{aligned} \quad (8)$$

where i is the iteration index, v is the previous gradient, η is the same regularization coefficient that appears in (7) and ς is the momentum. The latter was initialized to 0.95 [32]. Momentum-based methods damp the gradient and provide better convergence rates for deep networks. Xavier initialization is used to initialize the weights in each layer from a normal distribution of $N(0, 10^{-4})$ [32]. This initialization ensures that the signal remains within a reasonable range of values through the network. We use an equal learning rate, α , for all layers, which was adjusted via validation checks. The learning rate was initialized at 0.01 [32]. After each epoch, if the error of the validation set (80/20 split for training) remains the same or increases, then the learning rate is decreased by a factor of 2. The CNN is trained with mini-batch stochastic gradient descent with a batch size of 128 images. Log loss is used to

evaluate the performance of the network after each batch:

$$\ell(z, f(x, w)) = -\frac{1}{N} \sum_{n=1}^N \sum_{m=1}^M B_{m,n} \log(p_{m,n}), \quad (9)$$

where N is the number of training examples in the batch, and M is the number of classes. When the n_{th} example is classified into the m_{th} class, $B_{m,n}$ equals 1. Otherwise, $B_{m,n}$ equals 0. $p_{m,n}$ is the probability of the n_{th} example being classified into the m_{th} class.

3) Over-Fitting: We use dropout and channel-wise normalization to prevent over-fitting. Dropout, first introduced in [35], has been shown to reduce over-fitting and significantly improve training and testing accuracy [35]. Dropout randomly drops hidden nodes and connections from the network during training, thus preventing the hidden nodes from co-adapting with each other and improving the generalization of the network. The rate of dropout ζ defines the probability of a neuron being active during training. We applied dropout after each of the two convolutional layers and between the fully connected layers using $\zeta = 0.1$, $\zeta = 0.25$, and $\zeta = 0.5$, respectively. We use channel-wise normalization after every convolutional and fully connected layer to reduce error rate, as in [32]. Let $a_{s,t}^k$ be the activity of a neuron as computed by applying kernel k at position (s,t) . Then, applying the ReLU nonlinearity, the response-normalized activity $b_{s,t}^k$ is given by the expression:

$$b_{s,t}^k = \frac{a_{s,t}^k}{\left(\tau + \alpha \sum_{q \in K} (a_{s,t}^q)^2 \right)^\beta}, \quad (10)$$

where $q = [\max(0, (k - (n/2))), \min(K - 1, (k + (n/2)))]$.

The sum runs over n adjacent feature maps at the same spatial position, where K is the total number of kernels in the layer. The hyper-parameters were set according to [32] as $\alpha = 10^{-4}$, $\beta = 0.75$, $\tau = 2$, $n = 5$. This normalization is related to the lateral inhibition found in real neural networks [32].

B. Adaptive Local Window

Implementing the PC and MS local energies requires defining a local window surrounding each ZLS contour point, in which the energy cost function is calculated. In our previous paper [28], we proposed an iterative approach to calculate the adaptive size of the local window, a process fully detailed in [28]. The algorithm is applied for each point at each iteration and for each lesion in the image database. The adaptive window is applied separately for the X and Y window dimensions and is calculated using the lesion scale and its texture. Let L_x, L_y be the approximate x and y lesion dimensions defined by a generated bounding box surrounding the lesion. Since we minimize user input by requiring only a single input point, we approximate lesion size by generating a surrounding bounding box as described in section V-B. In addition to the lesion size, successfully dealing with high lesion diversity requires considering spatial image texture. Texture analysis is accomplished by extracting Haralick image features

(e.g. contrast, homogeneity) from a second order statistics model, namely, gray-level co-occurrence matrices (GLCM) [36]. Our method incorporates both global and local texture in a single hybrid model. For each point (x, y) examined in image I , we compare pairs of pixels, where the second pixel in the pair is $(x + \cos \theta, y + \sin \theta)$, located at $\theta \in \{0, 90, 180, 270\}$ degrees relative to the first pixel. Let W be a local window of $X_w \times Y_w$ pixels, surrounding an examined contour point within a region I . The co-occurrence matrix $P(m, n, \theta)$ of W is defined as the number of pixel pairs (x, y) and $(x + \cos \theta, y + \sin \theta)$ in W with grey values of (m, n) :

$$P(m, n, \theta) = \sum_{x=1}^{X_w} \sum_{y=1}^{Y_w} \begin{cases} 1, & I(x, y) = m \text{ and } I(x + \cos \theta, y + \sin \theta) = n \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

Then, homogeneity and contrast criteria are evaluated for each θ as:

$$\begin{aligned} \text{Homogeneity}(\theta) &= \sum_{m,n=0}^{N_G-1} P(m, n, \theta) \left(1 + |m - n|^{-1} \right) \\ \text{Contrast}(\theta) &= \sum_{m,n=0}^{N_G-1} P(m, n, \theta) \left(|m - n|^2 \right), \end{aligned} \quad (12)$$

where N_G is the total number of grey levels. These spatial criteria are averaged for each individual axis, X and Y axes. For local analysis, criteria are evaluated for each ZLS point separately while for global analysis, those criteria are calculated and averaged over all points within the lesion bounding box. According to eq. (12), we define GH as the global homogeneity, GC as the global contrast and LC_{xij}, LC_{yij} as the local contrast in the x and y directions surrounding the i_{th} contour point at the j_{th} iteration. The interaction between the global and the local terms plays an important role in determining the window size. For each i_{th} contour point, the local contrasts LC_{xij}, LC_{yij} are re-estimated at each j_{th} iteration. On the other hand, the method computes the global contrast, GC , and the global homogeneity, GH , only once within the entire region of interest (ROI). The adaptive window size is then calculated as:

$$\begin{aligned} \hat{W}_{xij} &= \frac{L_x}{\log(L_x)} \times \frac{1}{\left(GH + \frac{1}{GC} + \frac{1}{LC_{xij}} + \frac{1}{\bar{F}_{j-1}} \right)} \\ \hat{W}_{yij} &= \frac{L_y}{\log(L_y)} \times \frac{1}{\left(GH + \frac{1}{GC} + \frac{1}{LC_{yij}} + \frac{1}{\bar{F}_{j-1}} \right)}, \end{aligned} \quad (13)$$

where \bar{F}_{j-1} represents the average value of the energy functional over all ZLS contour points during the previous iteration. As long as curve evolution continues, the average value of \bar{F}_{j-1} should decrease as the size of the local window decreases.

C. Method Optimization

To optimize the local energies, each point is considered separately, and the algorithm attempts to minimize the energy

computed inside each local region. Local neighborhoods are split into their local interior and local exterior by the evolving ZLS curve. The energy is optimized using the energy model in an adapted surrounding region. Let $E(\phi)$ be an energy functional derived by a localization of the generic force $F(\phi)$:

$$E(\phi + \xi) = \int_{\Omega_n} \delta(\phi(x, y) + \xi) \cdot F(I(x, y), \phi(x, y) + \xi) dx dy, \quad (14)$$

where ξ represents a small change along the normal direction of $\Phi(x, y)$. The first variation of (14) is defined as [15]:

$$E(\phi) = \int_{\Omega_n} \delta\phi(x, y) F(I(x, y), \phi(x, y)) dx dy \quad (15)$$

We take a partial derivative of (15) and consider the minor differential of the perturbation ($\xi \rightarrow 0$):

$$\begin{aligned} \nabla_{\xi}|_{\xi=0} E(\phi + \xi) \\ = \int_{\Omega_n} \delta\phi(x, y) \cdot |\nabla\phi(x, y)| F(I(x, y), \phi(x, y)) dx dy \\ + \int_{\Omega_n} \eta\phi(x, y) dx dy, \end{aligned} \quad (16)$$

where $\eta\phi(x, y)$ represents the derivative of $\delta\phi(x, y)$ equal to zero for every ZLS point, thus having no effect on the curve evolution. The Cauchy–Schwartz inequality can be used to show that the optimal direction of curve evolution is [24]:

$$\frac{\partial\phi(x, y)}{\partial t} = \int_{\Omega_n} \delta\phi(x, y) \cdot |\nabla\phi(x, y)| F(I(x, y), \phi(x, y)) dx dy, \quad (17)$$

Equation (17) is applied in our proposed method to evolve the ZLS curve between sequential iterations of the segmentation process.

IV. IMPLEMENTATION DETAILS

A. Narrow Band

The proposed method calculates the spatial statistics only for grid points located within a narrow band of the distance map $\phi(x, y)$ around the curve C . This idea was introduced in [38] and has become common in implementations of local segmentation frameworks. The segmentation process begins with the initialization of every pixel within the narrow band using values of exterior and interior statistics [24]. The initialization operation is only performed once for each pixel. The narrow band moves over iterations, thus including grid points that must be initialized during the segmentation. An update of the distance map $\phi(x, y)$ occurs only within the narrow band.

B. Training Set and Data Augmentation

The CNN was trained separately on the MRI and CT datasets because both image sets differed significantly from each other in spatial characteristics of the lesions. However, the CNN architecture remained the same across both datasets. We evaluated our method for each dataset using 10-fold cross validation: 90% of the cases were used for training and the other 10% were considered as the testing set. The training

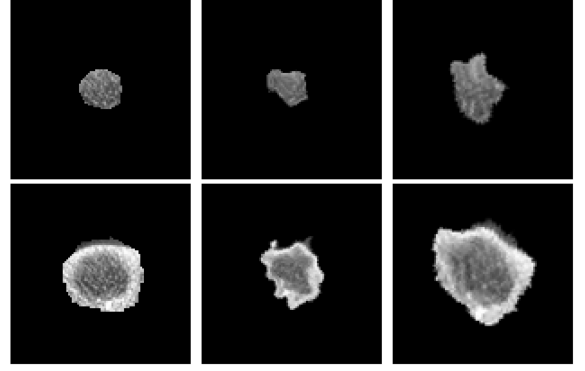


Fig. 6. Data augmentation. The top row shows an image (left) that was obtained by convolving the *eroded* mask of the manual annotation and the original image. The bottom row shows an image (left) that was obtained by convolving the *dilated* mask of the manual annotation and the original image. The center and the right columns present two distortions of the left image in the same row. Both elastic distortions and the affine transformations (rotation, scaling, translation) are evident.

set for the CNN was created by 1) eroding and dilating the manual annotation of a lesion and by 2) augmenting the data. We performed four dilations (p_3) with factors 1.2:0.5:2.7 and four erosions (p_1) with factors 0.25:0.2:0.85 for each image. The p_2 class was represented by the original manual annotation.

Deep neural networks, due to their size and complexity, typically require substantial datasets to perform optimally. Therefore, our eroded / dilated training set was followed by data augmentation to artificially increase the size of the training set. Data augmentation was performed by applying a combination of elastic and affine distortions to each training image. Figure 6 shows an example of a class p_1 and class p_3 representation of the same lesion and two distortions for each class image. Simard et al. [38] used non-rigid distortions to significantly increase accuracy on the MNIST image database of handwritten digits. We create elastic distortions by generating random displacement fields with values within the range of $[-1, 1]$, convolving these fields with a range of Gaussian filters, and multiplying the resulting matrices by a range of constant factors, controlling the intensity of the deformation. After the image was non-rigidly distorted, we applied rotation ($-30 \leq \theta \leq 30$ degrees), scaling ($0.75 \leq \varepsilon \leq 1.3$), and translation ($-1 \leq x \leq 1, -1 \leq y \leq 1$) to create a transformation-invariant, texture-based model.

V. EXPERIMENTAL DETAILS

A. Data

Our institutional review board approved this study. We analyzed 276 liver lesions obtained from two different datasets. The first dataset contained 112 contrast-enhanced CT images (Siemens Medical Solutions, Erlangen, Germany) of liver lesions (43 hemangiomas, 45 cysts and 24 metastases). Image acquisition parameters were: 120 kVp, 140–400 mAs, 2.5–5 mm section thickness and pixel spacing of 0.704 ± 0.085 mm. The second dataset included 164 liver lesions from 3T gadoxetic acid enhanced MRI (Signa Excite HDxt; GE Healthcare, Milwaukee, WI) scans in patients with

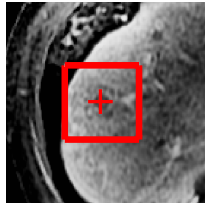


Fig. 7. Bounding box (ROI) reconstruction. A single input point indicates the lesion's center (red plus). The bounding box (red rectangle) is constructed by taking a region surrounding this center point.

cirrhosis at a tertiary liver center for evaluation of suspected hepatocellular carcinoma (HCC) and were found to have one or more LI-RADS (LR), LR-3 or LR-4 lesions. Image acquisition parameters were: 5 mm slice thickness and pixel spacing of 0.805 ± 0.078 mm; pulse sequences of single-shot fast spin-echo T2-weighted and pre- and post-contrast axial 3D T1-weighted fat-suppressed gradient-echo were used. The two datasets were acquired from two different academic institutions. Lesion size varied widely across the 276 cases. The lesions were also highly diverse in terms of spatial characteristics (homogeneity and contrast) due to different imaging modalities, and in different diseases, enhancing the importance and the need for an adaptive parameter-based method capable of handling a wide range of spatial textures.

B. Region of Interest and Initial Distance Map

To define the region of interest (ROI), board-certified abdominal imaging radiologists marked one-input point in the approximate center of the lesion (red plus sign, Fig. 7). A bounding box was then generated of size 52×52 pixels surrounding this center point for each MRI lesion (red box, Fig. 7) and of size 100×100 pixels for each CT liver lesion. These bounding boxes were fixed for all lesions (MRI/CT separately), they were not depended on the specific tested lesion, and they ensured that the entire lesion and its surroundings were included in the segmentation process (CT images included larger lesions, section V-A). Note that given the constraints of the CNN architecture, the size of the bounding box must be a multiple of four. The distance map, $\phi(x, y)$, was constructed and an initial zero level set contour was obtained. Six different contour initializations were created for MRI images, using radii of 3, 5, 7, 9, 11 and 13 pixels. Similarly, six different contour initializations were created for CT images, using radii of 5, 10, 15, 20, 25 and 30 pixels. For both MRI and CT images, these radii generated initial contours that were located inside the lesion ('small initialization'), while other contours were located close to the lesion boundaries ('accurate initialization'), and others were bigger than the lesion ('large initialization'). This broad range of initializations allowed us to evaluate the strength of our method in handling initial contours far from the lesion in either direction.

Radiologists traced the liver lesions on a single 2D slice for each liver lesion, providing manual markings that served as the reference standard for evaluating the proposed method. The automated segmentation contours were extracted and quantitatively compared with the radiologist's marking using the Dice similarity coefficient.

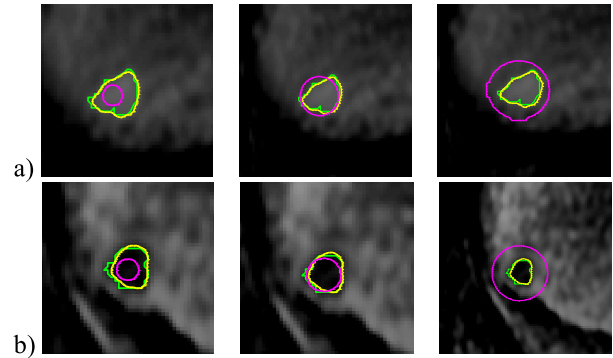


Fig. 8. Lesion segmentation using the proposed method with different initializations. Left column - small initialization (3-pixels radius), middle column - more accurate initialization (5-pixels radius), right column - large initialization (9-pixels radius). a) low-contrast lesion, b) noisy and heterogeneous tissue surrounding the lesion. For both cases, lesion is located close to the liver boundary. Magenta - initial contour, yellow - our final segmentation, green - manual radiologists' annotation.

VI. RESULTS

A. Lesion Sizes

The CT liver lesions had an average radius of 22.09 ± 13.35 pixels and the MRI liver lesions had an average radius of 7.12 ± 2.19 pixels. These sizes were considered when establishing experiments with different initial contours.

B. Segmentation Performance

The challenges regarding the segmentation process include: 1) a variety of spatial texture characteristics and 2) a wide range of optional initializations (Fig. 8). All 276 lesions have been segmented by our proposed adaptive contour parameters method (ACP). For all 6 contour initializations, average Dice coefficients and standard errors of 0.78 ± 0.05 and 0.79 ± 0.05 were found for the PC and MS energies, respectively.

C. Parameter Evolution

We deliberately chose initial contours far from the lesion boundary in order to test the strength of our method. During the segmentation process, as the contour gradually approached the lesion boundaries, λ_1 and λ_2 became more similar to each other as well. Figure 9 presents two examples of this convergence, showing an MRI liver lesion with a very small initial contour (Fig. 9a), and another lesion with a very large initial contour (Fig. 9b).

D. Process Convergence

Figure 10 demonstrates the convergence of both the Dice coefficient and the energy functional over multiple iterations for one example case in each of the CT and MRI datasets. As can be clearly seen for both the CT and MRI images, the Dice coefficient increased over successive iterations, indicating greater agreement between the manual and automated segmentations. As expected, the energy decreased during the segmentation process, suggesting that an appropriate minimization of

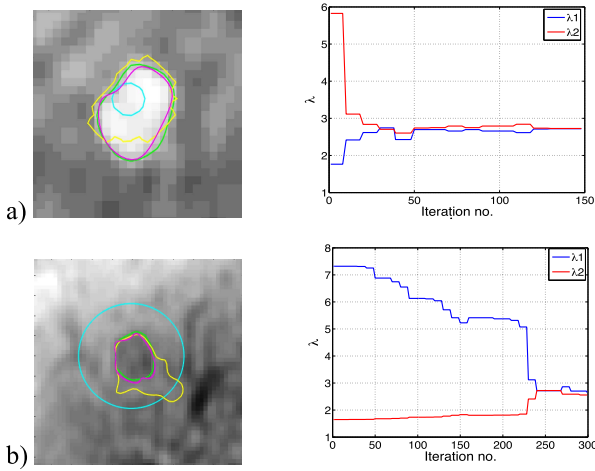


Fig. 9. The evolution of λ_1 , λ_2 parameters over iterations of the segmentation process. a) Small initial contour, b) big initial contour. Cyan is the initial contour, yellow is the automated contour during the iterative process – a) iteration no. 10, b) iteration no. 100. Green is the final automated contour, and magenta is the manual marking. Although the CNN initially sets λ_1 or λ_2 significantly higher than the other depending on the location of the initial contour, the values eventually converge.

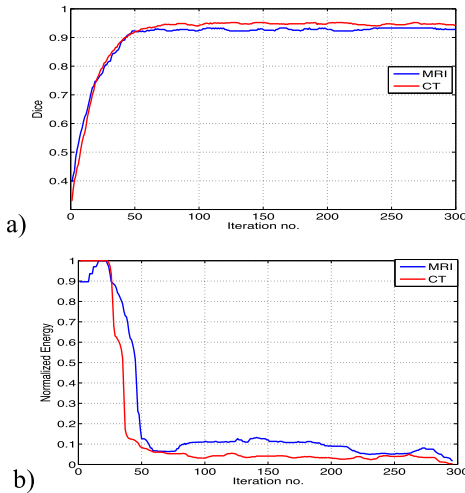


Fig. 10. Convergence of a) the Dice coefficient and b) the energy functional over iterations for both MRI and CT datasets.

the energy functional occurred. For both criteria, substantial convergence was obtained after 60 iterations, and there were only minor fluctuations around their final values over later iterations.

E. Comparison With Fixed Contour Parameter Method

We compared our proposed method (ACP) with a state of the art local framework of level set segmentation. This framework uses Fixed λ_1 , λ_2 Contour Parameters and a fixed local window size (FCP). For our datasets, several values of λ_1 and λ_2 were tested, and $\lambda_1 = \lambda_2 = 24$ was chosen. In addition, we used local window sizes of 5-pixels and 7-pixels surrounding each contour point for MRI and CT liver lesions respectively. Those fixed parameter values were selected for FCP because they supplied the average best results for all cases. Figure 11 shows some examples of

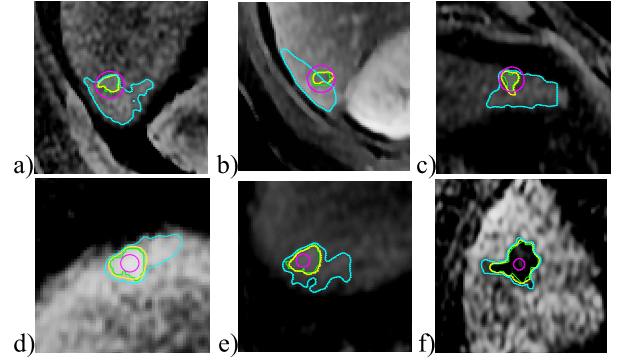


Fig. 11. Comparison of the segmentations that were obtained by the manual marking (green contour), the proposed ACP (yellow contour) and the FCP method (cyan contour). The initial contour is shown in magenta.

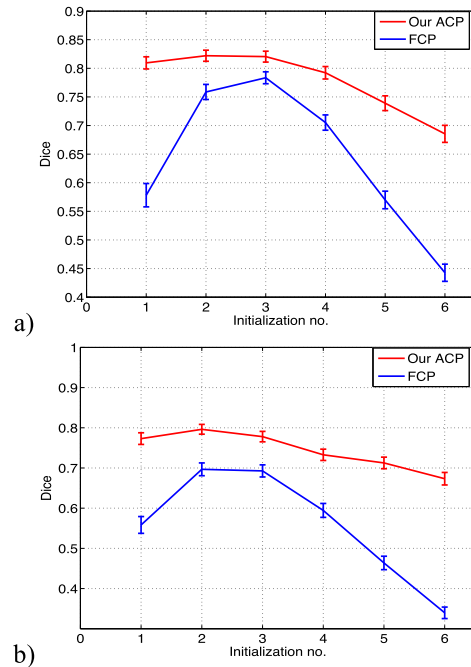


Fig. 12. Mean \pm SE of Dice criterion over the whole set of 276 cases and for each contour initialization size for a) the PC energy model, b) the MS energy model.

different lesion characteristics, initial contour size, and final segmentation of the lesion, using both ACP and FCP methods. For the combined dataset of 276 lesions and all six contour initializations, FCP had average Dice coefficients of 0.64 ± 0.13 and 0.53 ± 0.14 for PC and MS energy models, respectively. Those Dice coefficients were significantly lower ($p < 0.05$ for PC and $p < 0.01$ for MS, Wilcoxon) for FCP compared with our ACP segmentation performance (section VI-B). Figure 12 clearly shows that our ACP outperforms the state of the art FCP.

A subset of lesions for which one or more automated methods obtained less than 85% agreement with the manual marking, indicating a more challenging case, was also examined (Table I). Across both datasets, both energy models and all initializations, the average number of cases included in this subset was 103 ± 24 . For this subset, a significant Dice improvement of 0.21 ± 0.09 was obtained using our ACP method (Wilcoxon, $p < 0.001$).

TABLE I

AVERAGE DICE COEFFICIENTS FOR A SUBSET OF LESIONS THAT HAD DICE SIMILARITY COEFFICIENT < 0.85 WITH THE MANUAL MARKING. THE PRESENTED DICE VALUES ARE AVERAGED OVER ALL 6 CONTOUR INITIALIZATIONS. ACP SIGNIFICANTLY OUTPERFORMED THE EQUIVALENT FCP METHOD ($P < 0.001$, WILCOXON PAIRED TEST)

	Our ACP	FCP
MRI – PC energy	0.73 ± 0.03	0.59 ± 0.13
MRI – MS energy	0.77 ± 0.04	0.47 ± 0.14
CT – PC energy	0.66 ± 0.06	0.47 ± 0.12
CT – MS energy	0.67 ± 0.12	0.48 ± 0.1

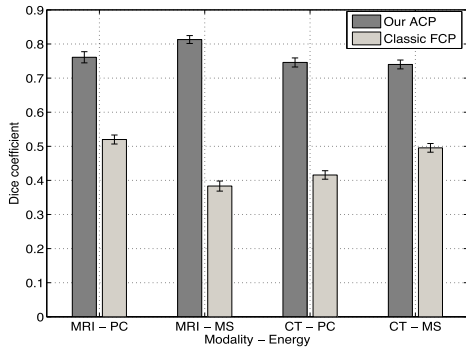


Fig. 13. Dice coefficient for a subset of cases in which there was more than 10% difference between our ACP and the state-of-the-art FCP in either direction.

We also considered a subset of lesions where an absolute difference of more than 10% between our ACP and the FCP method was present; in these cases, ACP could be better or worse than FCP. This threshold was chosen based on a radiologist opinion that in our set of images, more than 10% difference is clinically significant. For this subset, ACP performed significantly better with a Dice improvement of 0.31 ± 0.1 compared with FCP ($p < 0.001$, Wilcoxon). Moreover, as can be seen in Fig. 13, the proposed ACP method again shows greater robustness to imaging modality and type of local energy modeled.

All evaluations presented above include the six different contour initializations that were generated by placing the contour's center at the approximate center of the lesion and changing its diameter. We conducted one additional test to examine the ability of the method to handle different contour locations, such as contours partially inside and partially outside the lesion. We considered the effect of locating the initial contour at a displaced center point; four different displacements of $[\pm 3, \pm 3]$ pixels relative to the lesion center were tested for all 276 lesions. In this experiment, our ACP supplied an average median Dice coefficient of 0.84 ± 0.03 , compared to 0.63 ± 0.26 obtained by the FCP model. As before, ACP showed significantly better agreement with the manual marking than did FCP ($p < 0.001$, Wilcoxon), proving to be more robust to initialization location.

F. Comparison to Other State-of-the-Art Methods

We compared our method to another level set approach proposed by Li et al [5]. The authors proposed

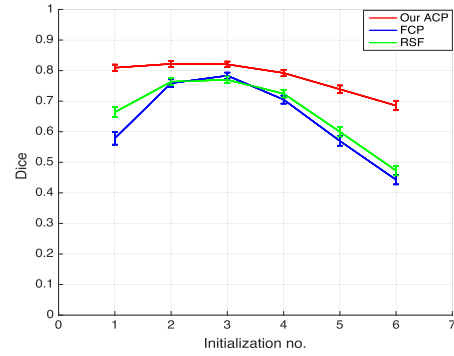


Fig. 14. Mean \pm SE of Dice criterion over the whole set of 276 cases and for each contour initialization for our ACP (red), FCP (blue), and RSF (green) methods.

a region-scalable fitting (RSF) model. Their method can handle images with intensity inhomogeneities, because it presents weighted averages of the image intensities in a Gaussian window inside and outside the contour. The authors presented their method as better than the PC and MS energy models for cases with such inhomogeneity, which is a prominent challenge that appears in our dataset. However, as with the FCP method presented above, Li's method still requires definition of the energy functional parameters, λ_1 and λ_2 , as well as the value of the local scale. In their method, λ_1 and λ_2 were fixed across different lesions, while the scale σ was manually fitted between lesions according to the analyzed spatial characteristics. The inefficiency of manual tuning enhances the added value of adaptive parameter tuning, as proposed in our ACP method.

We tested Li's RSF method on both the MRI and CT datasets across all initializations. In order to conduct an optimal evaluation of the RSF method, we analyzed the liver lesions using a range of Gaussian scales (0.5:0.5:4). Values to the left and right of this range resulted in convergence of the contour into local minima and increased sensitivity to noise respectively. The best scale was then chosen for each lesion. In addition, several values of λ_1 and λ_2 were tested, and $\lambda_1 = \lambda_2 = 2$ was found to supply the best average results. Li's method supplied an average Dice of 0.66 ± 0.09 for MRI and 0.57 ± 0.09 for CT across all initializations. Thus, the segmentation results obtained by Li's method are comparable to the traditional FCP results ($p > 0.05$, Wilcoxon). However, our method outperforms Li's method for all initializations ($p < 0.01$, Wilcoxon). These results enhance the strength of our proposed ACP, which is significantly better than both FCP and RSF (Fig. 14).

In addition to the comparisons with 2 level set frameworks, we also compared our ACP technique with 2 automatic, CNN-based segmentation methods. The first one is a patch-based CNN technique in which a bounding box was created and the image was divided into 5×5 patches [32]. Each patch was classified by a neural network as normal or abnormal tissue (i.e., lesion). The bounding box size remained the same as the previous experiments for both the MRI and CT datasets. The architecture of the neural network replicated the fully connected layers of the CNN architecture we proposed for our ACP method. The learning rate ($\alpha = 0.1$) and number of

nodes (nodes = 100) in the dense layer were optimized using Grid Search. Ten-fold cross validation was applied to evaluate the method.

The patch-based CNN technique resulted in an average Dice coefficient of 0.48 ± 0.27 for MRI liver lesions and 0.46 ± 0.25 for CT lesions, significantly lower than the equivalent values obtained by our ACP method ($p < 0.001$, Wilcoxon). The sensitivity of the method was 0.93 and the specificity 0.31, indicating that the network's low accuracy was primarily a result of false positives.

We also compared our method with a point supervised-based CNN approach proposed by Bearman et al. [39]. This approach trains a fully convolutional network using only a few representative points from each class, lesion or healthy, instead of all pixels within the region of interest. The CNN then classifies each pixel in the image as lesion or normal. Although their approach uses a single point to represent each class, we chose ten representative points on the lesion due to the high level of noise and heterogeneity in our images. We trained the 16-layer VGG net that was used in [39] and set the architecture and parameters to the same values as were presented by the authors ($\alpha = 10^{-5}$, $\zeta = 0.9$, $\eta = 0.0005$). Those values were also found to be the best in our analysis. As the authors did, we added a deconvolution layer for bilinear upsampling of the output to pixel-level predictions, and changed the upsampling dimensions to fit the sizes of our images. We found that Bearman's method performed worse than the patch-based CNN method. Their method, when compared to the manual markings, resulted in Dice coefficients of 0.35 ± 0.21 for MRI and 0.39 ± 0.24 for CT, significantly worse than our ACP results ($p < 0.001$, Wilcoxon).

VII. DISCUSSION AND CONCLUSIONS

We present a novel method for adaptive estimation of the λ_1 , λ_2 contour parameters for level set segmentation. When combined with estimation of the adaptive window size surrounding each ZLS point, as suggested in our previous paper [28], we first provide a fully automatic and adaptive method once a single center point of the lesion has been indicated on the image. The main benefits of the proposed approach are:

- Generalization of the segmentation process that results in an adaptive solution for each individual case. The method uses the same CNN architecture and the same equations for calculating the appropriate λ_1 , λ_2 .
- Ability to handle highly diverse lesion datasets?
- Independence from substantial manual interaction, parameter tuning, or accurate initial contour
- Higher agreement with the manual marking and increased robustness to various conditions (initialization, energy model, spatial texture) compared with state of the art methods

Our method was validated on two imaging modalities, two energy functionals, and a wide range of sizes and locations of contour initialization relative to the manual annotation of the lesion. All of this, in addition to the substantial diversity of lesion characteristics, shows the strength of our method. We compared our results with two state-of-the-art level

set frameworks and two CNN based methods. Figure 11 shows that our ACP method can handle different types of lesions— low contrast, heterogeneous, noisy background, or close to liver edge, substantially better than the commonly-used level set framework with pre-defined Fixed Contour Parameters (FCP). Figure 12 shows that our method outperforms FCP in terms of segmentation accuracy and robustness regardless of the initial contour used and has an average Dice coefficient 0.19 higher than FCP. We also analyzed two subsets of lesions that were composed of more challenging cases (Table I, Fig. 13). Dice improvements of 0.21 and 0.27 were obtained with the ACP method in comparison to the classic FCP. ACP also shows an average Dice improvement of 0.22 over FCP when the initialization is displaced from the lesion center point, thereby demonstrating robustness to location change in the initial contour. Our proposed method shows better agreement with the manual marking and higher robustness to different contour initializations, energy types, imaging modalities and different subsets of lesion complexity than state-of-the methods. These findings indicate that common methods (active contour based and CNN-based) cannot provide accurate segmentations of such highly diverse datasets. Common CNN techniques result in both many false positives and negative cases and are unable to segment challenging lesions with high accuracy. On the other hand, one of the well-known challenges of commonly used active contour models is handling highly heterogeneous lesions, low contrast lesions, or inaccurate contour initializations. Each of those lesion types requires completely different sets of energy functional parameters. Using a fixed set of parameters will result in an inaccurate segmentation of the lesion (e.g inability to converge into the boundaries of low contrast lesions or converge into local minima in the case of a heterogeneous or noisy lesion). Thus, our combination of deep learning and level set captures the benefits of both approaches and overcomes their limitations, to achieve significantly better results than either method alone.

The presented work has some limitations. First, a larger cohort is desirable. Second, additional manual markings for each lesion will result in more accurate evaluation of the automated segmentation. Future work may include an extension of the method to 3D, as well as incorporation of automatic lesion detection prior to segmentation, so that the entire segmentation process will be fully automated and will have no dependence on user input. Another possible extension is a joint training scheme that can incorporate different modalities into a single training process.

ACKNOWLEDGEMENT

We want to thank prof. Christopher F. Beaulieu from Radiology department at Stanford, and to Dr. Guilherme M. Cunha, Dr. Elhamy Heba, and prof. Claude B. Sirlin from Radiology department at San Diego university, for supplying the manual annotations of the liver lesions.

REFERENCES

- [1] S. Osher and J. A. Sethian, "Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. Comput. Phys.*, vol. 79, no. 1, pp. 12–49, 1988.

- [2] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *Int. J. Comput. Vis.*, vol. 22, no. 1, pp. 61–79, 1997.
- [3] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.
- [4] C. Li, C. Xu, C. Gui, and M. D. Fox, "Level set evolution without re-initialization: A new variational formulation," in *Proc. IEEE CVPR*, San Diego, CA, USA, Jun. 2005, pp. 430–436.
- [5] C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1940–1949, Oct. 2008.
- [6] A. Tsai, A. Yezzi, and A. S. Willsky, "Curve evolution implementation of the Mumford–Shah functional for image segmentation, denoising, interpolation, and magnification," *IEEE Trans. Image Process.*, vol. 10, no. 8, pp. 1169–1186, Aug. 2001.
- [7] C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, and J. C. Gore, "A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI," *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 2007–2016, Jul. 2011.
- [8] W. Neuenschwander, P. Fua, G. Szekely, and O. Kubler, "Initializing snakes," in *Proc. IEEE CVPR*, Seattle, WA, USA, Jun. 1994, pp. 658–663.
- [9] L. D. Cohen and R. Kimmel, "Global minimum for active contour models: A minimal path approach," *Int. J. Comput. Vis.*, vol. 24, no. 1, pp. 57–78, Aug. 1997.
- [10] R. Ardon and L. D. Cohen, "Fast constrained surface extraction by minimal paths," *Int. J. Comput. Vis.*, vol. 69, no. 1, pp. 127–136, Aug. 2006.
- [11] G. Xingfei and T. Jie, "An automatic active contour model for multiple objects," in *Proc. ICPR*, Aug. 2002, pp. 881–884.
- [12] Y. He, Y. Luo, and D. Hu, "Semi-automatic initialization of gradient vector flow snakes," *J. Electron. Image*, vol. 15, no. 4, p. 043006, 2006.
- [13] C. Tauber, H. Batatia, and A. Ayache, "A general quasi-automatic initialization for snakes: Application to ultrasound images," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2005, pp. II-806–II-809.
- [14] C. Li, J. Liu, and M. D. Fox, "Segmentation of external force field for automatic initialization and splitting of snakes," *Pattern Recognit.*, vol. 38, no. 11, pp. 1947–1960, 2005.
- [15] 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.
- [16] P. Das, O. Veksler, V. Zavatsky, and Y. Boykov, "Semiautomatic segmentation with compact shape prior," in *Proc. CRV*, 2006, pp. 28–36.
- [17] 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.
- [18] D. Freedman and T. Zhang, "Interactive graph cut based segmentation with shape priors," in *Proc. IEEE CVPR*, Jun. 2005, pp. 755–762.
- [19] B. N. Li, C. K. Chui, S. H. Ong, and S. Chang, "Integrating FCM and level sets for liver tumor segmentation," in *Proc. 13th Int. Conf. Biomed. Eng.*, 2009, pp. 202–205.
- [20] J. S. Suri, K. Liu, S. Singh, S. N. Laxminarayan, X. Zeng, and L. Reden, "Shape recovery algorithms using level sets in 2-D/3-D medical imagery: A state-of-the-art review," *IEEE Trans. Inf. Technol. Biomed.*, vol. 6, no. 1, pp. 8–28, Mar. 2002.
- [21] J. S. Suri, "Two-dimensional fast magnetic resonance brain segmentation," *IEEE Eng. Med. Biol.*, vol. 20, no. 4, pp. 84–95, Jul./Aug. 2001.
- [22] S. Ho, E. Bullitt, and G. Gerig, "Level set evolution with region competition: Automatic 3-D segmentation of brain tumors," in *Proc. IEEE ICPR*, Aug. 2002, pp. 532–535.
- [23] S. Lankton, D. Nain, A. Yezzi, and A. Tannenbaum, "Hybrid geodesic region-based curve evolutions for image segmentation," *Proc. SPIE*, vol. 6510, p. 65104U, Mar. 2007.
- [24] S. Lankton and A. Tannenbaum, "Localizing region-based active contours," *IEEE Trans. Image Process.*, vol. 17, no. 11, pp. 2029–2039, Nov. 2008.
- [25] J. An, M. Rousson, and C. Xu, "Γ-convergence approximation to piecewise smooth medical image segmentation," in *Proc. Med. Imag. Comput. Comput. Assist. Intervent.*, vol. 4792, 2007, pp. 495–502.
- [26] Q. Yang and D. Boukerroui, "Optimal spatial adaptation for local region-based active contours: An intersection of confidence intervals approach," in *Proc. IMAGAPP*, 2011, pp. 87–93.
- [27] J. Piovan and T. Papadopoulos, "Local statistic based region segmentation with automatic scale selection," in *Proc. ECCV*, 2008, pp. 486–499.
- [28] A. Hoogi *et al.* (2016). "Adaptive local window for level set segmentation of CT and MRI liver lesions." [Online]. Available: <https://arxiv.org/abs/1606.03765>
- [29] D. A. B. Oliveira, R. Q. Feitosa, and M. M. Correia, "Liver segmentation using level sets and genetic algorithms," in *Proc. 4th VISAPP*, vol. 2, 2009, pp. 154–159.
- [30] C. Baillard, C. Barillot, and P. Boutheymy, "Robust adaptive segmentation of 3D medical images with level sets," Res. Inst. Comput. Sci. Random Syst., Rennes, France, Tech. Rep. 1369, Nov. 2000.
- [31] A. Yezzi, A. Tsai, and A. Willsky, "A fully global approach to image segmentation via coupled curve evolution equations," *J. Vis. Commun. Image Represent.*, vol. 13, nos. 1–2, pp. 195–216, 2002.
- [32] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [33] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. AISTATS*, 2010, pp. 249–256.
- [34] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. ICML*, 2013.
- [35] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov. (Jul. 2012). "Improving neural networks by preventing co-adaptation of feature detectors." [Online]. Available: <https://arxiv.org/abs/1207.0580>
- [36] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. 3, no. 6, pp. 610–620, Nov. 1973.
- [37] D. Chopp, "Computing minimal surfaces via level set curvature flow," *J. Comput. Phys.*, vol. 106, no. 1, pp. 77–91, May 1993.
- [38] P. Y. Simard, D. Steinkraus, and J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis," in *Proc. ICDAR*, vol. 2, 2003, pp. 958–962.
- [39] A. Bearman, O. Russakovsky, V. Ferrari, and L. Fei-Fei. (Jun. 2015). "What's the point: Semantic segmentation with point supervision." [Online]. Available: <https://arxiv.org/abs/1506.02106>