



Deriving external forces via convolutional neural networks for biomedical image segmentation

YIBIAO RONG,^{1,5} DEHUI XIANG,^{1,5} WEIFANG ZHU,¹ FEI SHI,¹
ENTING GAO,³ ZHUN FAN,⁴ AND XINJIAN CHEN^{1,2,*}

¹*School of Electrical and Information Engineering, Soochow University, 215006, Suzhou, China*

²*State Key Laboratory of Radiation Medicine and Protection, Soochow University, 215123, Suzhou, China*

³*School of Electronic and Information Engineering, Suzhou University of Science and Technology, Suzhou, China*

⁴*Key Laboratory of Digital Signal and Image Processing of Guangdong Provincial, College of Engineering, Shantou University, 515063, Shantou, China*

⁵*Contributed equally to this work.*

**Corresponding author: xjchen@suda.edu.cn*

Abstract: Active contours, or snakes, are widely applied on biomedical image segmentation. They are curves defined within an image domain that can move to object boundaries under the influence of internal forces and external forces, in which the internal forces are generally computed from curves themselves and external forces from image data. Designing external forces properly is a key point with active contour algorithms since the external forces play a leading role in the evolution of active contours. One of most popular external forces for active contour models is gradient vector flow (GVF). However, GVF is sensitive to noise and false edges, which limits its application area. To handle this problem, in this paper, we propose using GVF as reference to train a convolutional neural network to derive an external force. The derived external force is then integrated into the active contour models for curve evolution. Three clinical applications, segmentation of optic disk in fundus images, fluid in retinal optical coherence tomography images and fetal head in ultrasound images, are employed to evaluate the proposed method. The results show that the proposed method is very promising since it achieves competitive performance for different tasks compared to the state-of-the-art algorithms.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Biomedical image segmentation plays an important role in developing computer aided systems for analysis of the images automatically. Numerous biomedical image segmentation methods [1–3] have been proposed in last few decades. According to the number of segmented objects in the images, these segmentation methods can be grouped into single object segmentation methods and multi objects segmentation methods. For single object segmentation, parametric active contour models are one of most widely used approaches, whose first version was proposed by Kass et al. in [4]. There are several advantages with parametric active contour models, for example, they can offer smooth and closed contours as segmentation results and obtain high accuracy of object boundaries [5]. However, for the classical active contour models [4], the active contours can not progress into boundary concavities. Besides, the initial curves are required to be close to object boundaries or else they will converge to wrong places. Many methods have been proposed to handle these limitations, in which gradient vector flow (GVF) [6] is one of most effective among them. Even though GVF is insensitive to initial curves and can drive the curves into boundary concavities, it is sensitive to noise and false edges. Fig. 1 gives an example, in which Fig. 1(a) is the GVF obtained by the method introduced in [6]. It is observed that the GVF is disordered, which can not drive the initial curve into object boundary, as shown in Fig. 1(b), where the red curve is initial curve and the blue is final result.

Abundant improved versions for GVF have been reported in the literature. For instance, Tang et al. [7] extended the classical GVF snake into multi direction GVF snake so that it can trace the boundary of skin cancer. Zhu et al. [8] used a minimal surface and component-normalized method to improve the GVF. These improved versions, generally, were carefully designed based on some domain knowledge, e.g. the intensity distribution of the region of interest. Different from these domain knowledge based methods, in this paper, we propose using GVF as reference to train a convolutional neural network (CNN) [9] to derive an external force for the parametric active contour models. CNNs belong to end-to-end trainable methods. It allows us to avoid making hypotheses during designing the external forces, as a result, the dependence on the prior knowledge is abated, while the quality of the external forces is improved compared with those obtained by the classical methods [6, 10] (Quantitative comparison is given in Section 4). Fig. 1(c) shows an example. It is observed that the derived external force is ordered and points to object boundary, which can ensure the active contour models to exclude the interference from noise and drive the initial curve into the right place, as shown in Fig. 1(d).

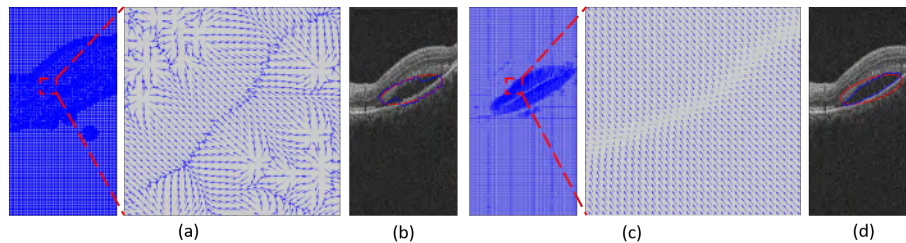


Fig. 1. An example to demonstrate different external forces obtained by different ways. (a) The GVF derived from the gradient of original image. (b) The detected result (blue curve) with (a). (c) The external force derived by the trained CNN. (d) The detected result (blue curve) with (c).

2. Related works

As the revival of CNNs in recent years, some researchers have tried to employ CNNs to facilitate the application of active contour models in biomedical image segmentation domain. For instance, Srivastava et al. [11] employed a deep neural network to perform a preliminary optic disk segmentation in fundus images. The preliminary segmentation was then refined further using active shape models. Avendi et al. [12] applied CNNs to automatically detect left ventricle (LV) chamber in MRI dataset. Stacked autoencoders were then used to infer LV shape. The inferred shape was finally incorporated into deformable models to improve the accuracy of segmentation. Li et al. [13] used CNNs to segment myocardium of LV first. Active contour models were then applied to segment endocardium based on the initial segmentation.

The above methods [11–13] employed CNNs to highlight the objects directly or indirectly so that active contour models can capture the objects easier. Besides, several different methods are also presented to facilitate the application of active contour models based on CNNs. Such as, Hoogi et al. [14] employed CNNs for adaptive estimation of active contour parameters. Le et al. [15] boosted the classic active contour methods to a new level of learnable deep learning approaches by reformulating active contour as deep recurrent neural networks. Different from these works, in this paper, we employ CNNs to derive an external force for the parametric active contour models. There are two advantages with the proposed method: 1) CNNs belong to end-to-end trainable methods, which can abate the dependence on the prior knowledge during designing the external forces. 2) It can ensure the initial curves close to object boundaries based on the finding that the magnitude of the external force derived by CNNs around object boundaries

is larger than other places.

To demonstrate the effectiveness of the proposed method, we apply the proposed method on different clinical applications, which are segmentation of optic disk in fundus images, fluid in retinal optical coherence tomography (OCT) images and fetal head in ultrasound images. These three applications play an important role in developing computer aided systems in their respective domains. Such as, optic disk segmentation is an important part in the computer aided systems for diagnosis of glaucoma. Fetal head segmentation is a key step to estimate the circumference, which is an important measurement to estimate the gestational age and monitor the growth of fetus. Many different methods have been designed for these three tasks. Due to the main motivation of this paper is to explore whether the parametric active contour models with the external force derived by CNNs is effective for single object segmentation in biomedical images, thus we do not offer a detailed introduction for these methods. For a comprehensive reading about these methods, we refer readers to papers [16–22] for optic disk segmentation, papers [23–28] for fluid segmentation and papers [29–32] for fetal head segmentation.

The rest of the paper is organized as follows: the details of the proposed algorithm are described in Section 3. Section 4 presents experimental results, including the comparison with ad hoc algorithms. Discussion and conclusion are made in Section 5.

3. Methods

We begin by first reviewing the parametric active contour models, then formulating an active contour model with the external force derived by CNNs. Finally, we show how to initial the active contours based on the external force magnitude.

3.1. Active contour models: a quick review

Active contours, or snakes [4] are curves defined within an image domain that can move to object boundaries under the influence of internal forces and external forces, in which internal forces are used to hold the curves together and to keep them from bending too much. External forces are used to drive the curves toward target boundaries [6].

The process of using active contour models to move a curve $x(s) = [x(s), y(s)]$, $s \in [0, 1]$ into an object boundary is formulated as a problem of energy functional minimization

$$E = \int_0^1 \frac{1}{2} [\alpha |x'(s)|^2 + \beta |x''(s)|^2] + E_{ext}(x(s)) ds \quad (1)$$

where the weighting parameter α controls the tension of $x(s)$ and β controls the rigidity. $x'(s)$ is the first derivatives of $x(s)$ with respect to s and $x''(s)$ is the second derivatives. E_{ext} is the external energy term.

To minimize E , a curve $x(s)$ must satisfy the following Euler equation

$$\alpha x''(s) - \beta x'''(s) - \nabla E_{ext} = 0 \quad (2)$$

Equation (2) can be regarded as a force balance equation

$$F_{int} + F_{ext} = 0 \quad (3)$$

where internal force term $F_{int} = \alpha x''(s) - \beta x'''(s)$, external force term $F_{ext} = -\nabla E_{ext}$.

It is noted that internal forces are computed from the curve $x(s)$ itself. External forces are general derived from image data. One of popular external forces is GVF, which is computed as a diffusion of the gradient vectors of a gray-level or binary edge map derived from an image [6]. Mathematically, GVF is a vector $v(x, y) = [u(x, y), v(x, y)]$ that minimizes the following energy functional

$$\varepsilon = \int \int \mu (u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |v - \nabla f|^2 dx dy \quad (4)$$

where f is an edge map derived from an image. The parameter μ is a regularization parameter controlling the tradeoff between the first term and the second term in equation (4). ∇ is the gradient operator. u_x, v_x and u_y, v_y are partial derivatives of u and v to x and y respectively.

GVF is achieved by solving the following Euler equations

$$\begin{aligned}\mu\nabla^2 u - (u - f_x)(f_x^2 + f_y^2) &= 0 \\ \mu\nabla^2 v - (v - f_y)(f_x^2 + f_y^2) &= 0\end{aligned}\quad (5)$$

where ∇^2 is the Laplacian operator. For the details of numerical implementations of equations (2) and (5), we refer readers to papers [4] and [6] respectively.

3.2. Deriving external forces via CNNs

As introduced above, GVF is sensitive to noise and false edges in an image. To handle this problem, in this work, we propose using GVF as reference to train a CNN to derive an external force that points to the target boundaries, so that the active contour models can exclude the interference from false edges and noise with the derived external force. Mathematically, an active contour model with the external force obtained from CNNs is formulated as

$$\alpha x''(s) - \beta x'''(s) - \phi(I, \mathbf{w}) = 0 \quad (6)$$

where $\phi(I, \mathbf{w})$ is a network, termed GVF-Net for future reference, which is trained by using GVF as reference to derive an external force from a raw image I .

Given a training set of images and their referenced GVFs $\{I_i, v_i\}_{i=1}^N$, the weights \mathbf{w} of GVF-Net $\phi(I, \mathbf{w})$ are learned to minimize the mean squared error [9]

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w}} \sum_{i=1}^N |v_i - \phi(I_i, \mathbf{w})|^2 \quad (7)$$

It is noteworthy to point out that the referenced GVFs are computed from the boundary ground truth with equation (5), which can guarantee the referenced GVFs to point to the target boundaries. Fig. 2 gives an example, where Fig. 2(b) is the boundary ground truth of Fig. 2(a). Fig. 2(c) is the referenced GVF computed by equation (5) based on Fig. 2(b).

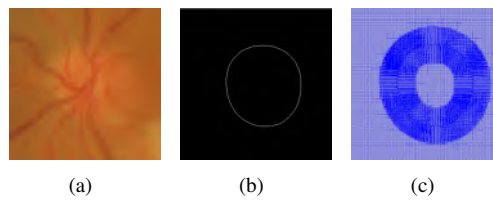


Fig. 2. An example to demonstrate the process of obtaining the referenced GVF. (a) Raw image. (b) Boundary ground truth of (a). (c) Referenced GVF.

In this work, we employ u-net structure [33] for GVF-Net, which can ensure the output without reducing size after a series of operations. Note that GVF is a matrix with size of $m \times n \times 2$, where $m \times n$ is the size of an image. As shown in Fig. 3, the left part of GVF-Net is termed contracting path, which processes the input images through a series of convolutional layers. The size of each kernel in each convolutional layer is 3×3 . Every two convolutional layers is followed by a 2×2 max pooling operation with stride 2 for downsampling. The number of feature map channels are doubled after each pooling step.

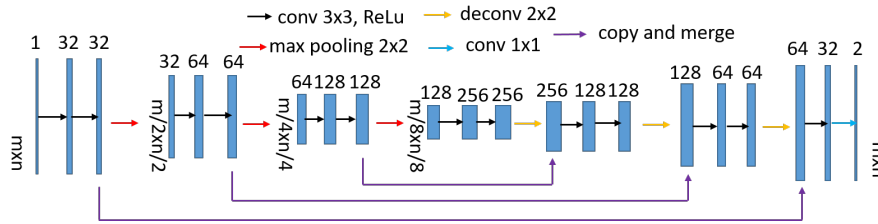


Fig. 3. The structure of GVF-Net.

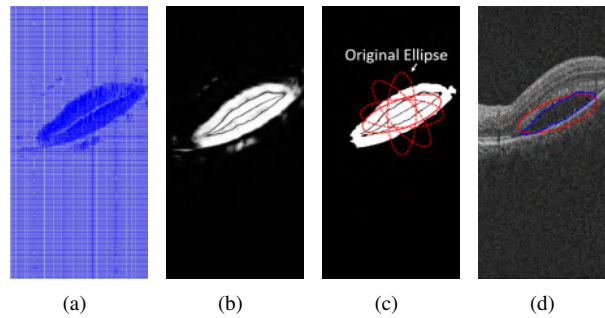


Fig. 4. An example to demonstrate the process of automatic initialization of curve. (a) The external force. (b) The magnitude of (a). (c) Maximum connected region of (b). (d) The initial curve (red) and final result (blue).

The right part in Fig. 3 is termed expansive path, which upsamples the feature maps by a series of deconvolutional layers. The size of each kernel in each deconvolutional layer is 2×2 . Each deconvolutional layer has a concatenation with the correspondingly cropped feature maps from contracting path and is followed by two convolutional layers. The final layer is a fully connected layer which used to map each component feature channel as the external force. In total, there are fourteen convolutional layers, three pooling layers, three deconvolutional layers and one fully connected layer in GVF-Net.

3.3. Initialization with external force magnitude

Curve initialization is one of key concerns with active contour algorithms. In general, an initial curve should be close to the truth boundaries or else it will likely converge to a wrong place [6]. In this paper, the initial curves are obtained based on the external force derived by GVF-Net. We find that the magnitude of the derived external force around object boundaries is larger than other places. Fig. 4 gives an example, in which Fig. 4(a) and Fig. 4(b) are the external force and corresponding magnitude respectively. The dark regions in the magnitude image indicate that the force is very weak, while the gray regions indicate that the force is strong. If an initial curve is located in the strong force regions and surrounds the objects as much as possible, it is very likely to converge to the target boundaries.

Based on these findings, we propose an automatic initialization approach for fluid segmentation and another for optic disk segmentation respectively. As shown in Fig. 4, the first step of the automatic approach for fluid segmentation is to transfer the external force magnitude into a binary image with thresholding method [34]. The maximum connected region in the binary image is extracted, as shown in Fig. 4(c). We then calculate the centroid of the maximum connected region. In addition, the major axis length and minor axis length of the ellipse that has the same

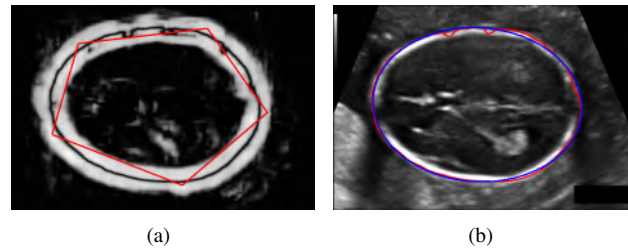


Fig. 5. An example to show the interactive initialization. (a) The external force magnitude and initial curve. (b) The fetal head boundary obtained by active contour model (red curve) and the result of ellipse fitting (blue curve).

normalized second central moments as the maximum connected region [35] are also computed. An original ellipse is obtained based on the centroid, major axis length and minor axis length. The original ellipse is then rotated with the interval of 10 degree around the centroid. The one containing the largest number of pixels belonging to the maximum connected region is selected as the initial curve, as shown in Fig. 4(d). For optic disk segmentation, we search a circle as the initial curve with circle hough transform [36] based on the binary images of the external force magnitude.

It is noteworthy to point out that the above automatic initialization approaches are not effective for fetal head segmentation. Hence, an interactive initialization method is proposed. As shown in Fig. 5(a), we draw a polygon as the initial curve based on the external force magnitude. The rules required to follow to draw a polygon are to ensure that each edge of polygon is located in the strong force regions and the polygon surrounds the object. We believe that the task of drawing such a polygon is simple since the strong force regions are obvious in the magnitude images.

3.4. Cascaded version

As introduced above, the external force is derived by CNNs directly. We believe that some existing schemes in machine learning community might have the capacity to improve the external force further. In this work, a cascaded scheme reported in [37] is employed to improve the external force. Concretely, the image-GVF pairs are first applied to train the GVF-net. The output of the trained GVF-net paired with the referenced GVF are then used to train another GVF-net. In this scheme, the second time training can be regarded as a fine tuning process, which is used to fine tune the external force obtained by the first GVF-Net. The effectiveness of the cascaded scheme will be verified by the experiments in Section 5.

4. Result analysis

4.1. Materials

4.1.1. For optic disk segmentation in fundus images

Three public databases for optic disk segmentation in fundus images are employed to evaluate the performance of the proposed algorithm, which are MESSIDOR database [38], DRIONS database [39] and ONHSD database [40]. There are 1200 images in MESSIDOR database, 110 images in DRIONS database and 99 images in ONHSD database. Note that there are 9 images without discriminable optic nerve head in ONHSD database, which will not be employed to evaluate the algorithm. The annotation of optic disk in each image has been marked by experts. Since the resolutions of fundus images in different databases are different, for convenience, the method for optic disk localization in [41] is first used to locate the disk in a fundus image. A 301×301 sub image for optic disk segmentation is then determined based on the optic disk

location. To train and test the GVF-Net, 60% of the images in MESSIDOR database are randomly selected as training set, 20% as validation set and 20% as test set. 60%-20%-20% is a common used proportion to split a small dataset into training set, validation set and test set to train and test a CNN [42]. All the images in DRIONS and ONHSD databases are used as test set.

4.1.2. For fetal head segmentation in ultrasound images

The data for fetal head segmentation in ultrasound images is from a grand challenge 2018 [43]. The purpose of this challenge is to measure circumference of fetal head. The database contains a total of 1334 two-dimensional (2D) ultrasound images, which is divided into a training set of 999 images and a test set of 335 images. The size of each 2D ultrasound image is 800×540 with a pixel size ranging from 0.052 to 0.326 millimeters (mm). In this challenge, the annotations are available for training set only. According to the requirement of sponsors, the boundary of a fetal head should be fitted by an ellipse, which is described by five parameters, namely the center (x,y), semi major axis length, semi minor axis length and angle of rotation. The detected parameters by algorithms on test set are required to upload to sponsors to gain final results. In this work, we employ the direct least squares fitting of ellipses method [44] to fit the fetal head boundaries obtained by the active contour models. An example to show the result of fitting is given in Fig. 5(b).

4.1.3. For fluid segmentation in OCT images

There are 481 OCT B-scan images for fluid segmentation. These images are collected from different OCT volumes obtained by ZEISS scanner in a local database. The resolution of each B-scan image is 512×1024 . The fluid regions in these images have been depicted by an expert. To train and test the proposed method, these images are divided into training set, validation set and test set with the ratio of 60%-20%-20%.

4.2. Metrics

Different metrics are employed to evaluate the performance of the proposed method according to different tasks.

4.2.1. For optic disk segmentation and fluid segmentation

The performance of the proposed method for these two tasks is evaluated according to area overlap (AOL), Dice similarity coefficient (DSC), accuracy (Ac), true positive fraction (TPF) and false positive fraction (FPF) [17], which are defined as

$$AOL = TP / (TP + FN + FP) \quad (8)$$

$$DSC = 2TP / (2TP + FN + FP) \quad (9)$$

$$Ac = (TP + TN) / (TP + TN + FP + FN) \quad (10)$$

$$TPF = TP / (TP + FN) \quad (11)$$

$$FPF = FP / (FP + TN) \quad (12)$$

where TP is short for true positive, TN true negative, FP false positive and FN false negative.

4.2.2. For fetal head segmentation

The metrics used to evaluate the methods for fetal head segmentation are difference (DF), absolute difference (ADF), Hausdorff distance (HD) and Dice similarity coefficient (DSC) [29]. DF, ADF

and HD are defined as

$$DF = HC_S - HC_R \quad (13)$$

$$ADF = |HC_S - HC_R| \quad (14)$$

$$HD = \max(h(S, R), h(R, S)) \quad (15)$$

where HC_R is the head circumference measured by the expert and HC_S is the head circumference determined by algorithms. $h(S, R) = \max_{s \in S} \max_{r \in R} \|s - r\|$, in which $R = \{r_1, r_2, \dots, r_q\}$ are the pixels in fetal head boundary from the expert, $S = \{s_1, s_2, \dots, s_p\}$ are the pixels from algorithms.

4.3. Training details and parameter setting

A PC, equipped with an Intel (R) Core (TM) i7-4790 M CPU at 3.60 GHZ and 8 GB of RAM capacity, is employed to perform the experiments with MATLAB. We use the MatConNet [45] to train the CNNs. All the parameters of GVF-Net are initialized with Xavier initialization and trained for 66 epochs with the mini-batch size of 3 image instances. Training convergence can be observed within 60 epochs. For other hyperparameters, learning rate is set to 0.02, momentum 0.9 and weight decay 0.0005. Once training done, α and β are parameters which may influence the performance of the proposed method. However, we find that the performance of the proposed method is insensitive to these two parameters and we set $\alpha = 0.7$ and $\beta = 0$ for the following experiments.

4.4. Performance analysis

Fig. 6 shows some examples for optic disk segmentation, fetal head segmentation and fluid segmentation respectively. It is observed that the proposed method works well for most of cases in the given examples, which demonstrates that the ability of generalization of the proposed method is high. However, the proposed method might fail for some cases, e.g. Fig. 6(c). Fig. 7 gives an example to explain why the proposed method might fail or succeed, where we can observe that for the case that the proposed method fails to detect the boundary, the external force derived by the GVF-Net dose not point to the target boundary. Conversely, if the derived external force points to the target boundary, it can guarantee the initial curve to be converged into the target boundary.

Table 1. The average values and standard deviations of different performance metrics achieved on different databases.

	MESSIDOR	DRIONS	ONHSD	Local
AOL	0.9144(0.0786)	0.8972(0.0668)	0.8045(0.1106)	0.8980(0.0613)
DSC	0.9531(0.0554)	0.9444(0.0412)	0.8871(0.0751)	0.9451(0.0374)
Ac	0.9864(0.0151)	0.9832(0.0128)	0.9869(0.0098)	0.9967(0.0025)
TPF	0.9471(0.0581)	0.9822(0.0198)	0.9351(0.0529)	0.9499(0.0404)
FPF	0.0066(0.0120)	0.0165(0.0153)	0.0100(0.0101)	0.0019(0.0024)

The average values and standard deviations of different performance metrics for optic disk segmentation and fluid segmentation achieved on different databases are summarized in Table 1. One might wonder that why the performance of the method on MESSIDOR database is better than on DRIONS and ONHSD databases. The reason is that GVF-Net for optic disk segmentation is trained by the images in MESSIDOR database only. In the discussion section, we will demonstrate that the performance of the method on DRIONS and ONHSD databases can

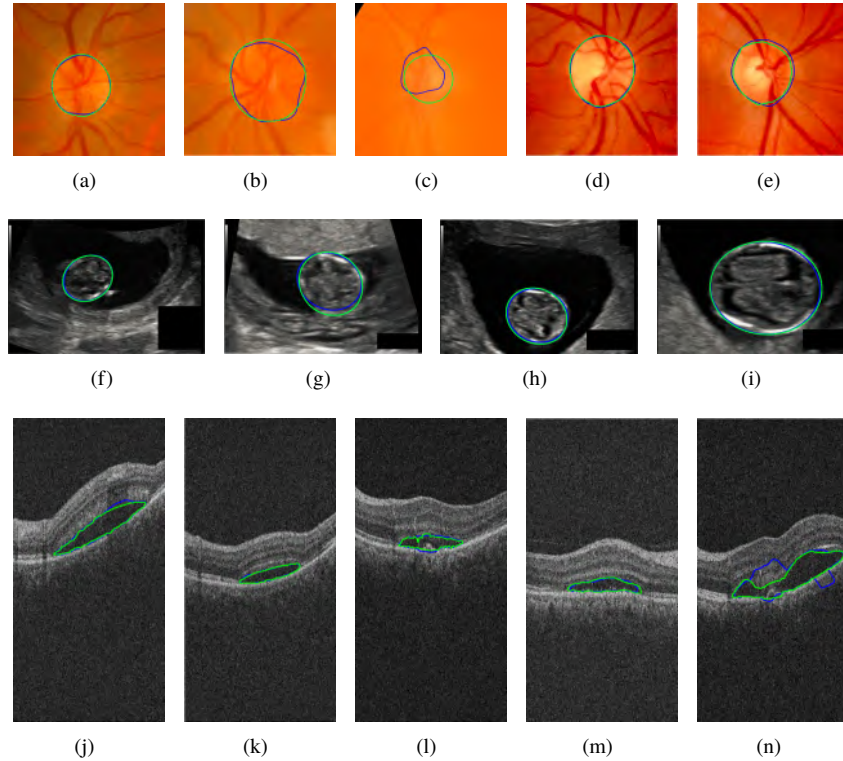


Fig. 6. Examples for optic disk segmentation (first row), fetal head segmentation (second row) and fluid segmentation (third row), where green curves are ground truth and blue curves are curves detected by the proposed algorithm.

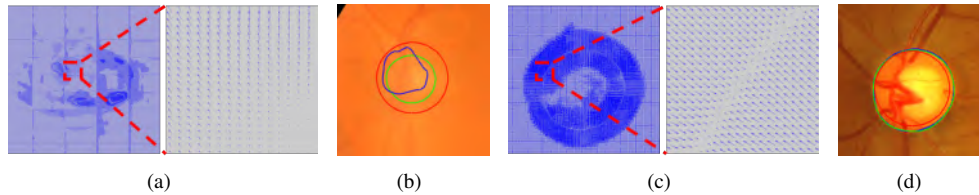


Fig. 7. An example to explain why the proposed method might fail or succeed. (a) Unsatisfactory external force. (b) Initial curve (red), detected result (blue) with (a) and ground truth (green). (c) Satisfactory external force. (d) Initial curve (red), detected result (blue) with (c) and ground truth (green).

Table 2. The performance of the proposed method for fetal head segmentation.

ADF(mm) \pm std	DSC(%) \pm std	AD(mm) \pm std	HD(mm) \pm std
2.45 \pm 2.55	95.49 \pm 4.11	-1.05 \pm 3.38	2.44 \pm 1.96

be improved by using a few of images in these two databases to fine tune the trained GVF-Net. Table 2 lists the results of quantification for fetal head segmentation in ultrasound images.

A correlation analysis through a scatterplot is performed to assess the performance of the

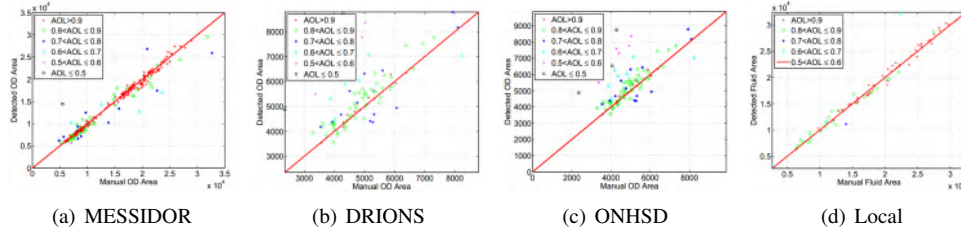


Fig. 8. Scatterplot of region area between manual segmentation and automatic segmentation.

Table 3. The percentage of images with AOL > 0.8 in different databases.

MESSIDOR	DRIONS	ONHSD	Local
92.50%	92.73%	67.78%	95.18%

proposed method in depth. Concretely, the correlation between the object area (the total number of object pixels) determined by the algorithm automatically and ground truth is given, as shown in Fig. 8. Note that the scatterplot for fetal head segmentation is not given here since the ground truth for test set is not available. The different styles of points in Fig. 8 mean the method performs on different types of images, which categorized according to different AOL intervals. It is observed that the points are very close to identity line for those images that the objects are segmented correctly (AOL>0.8). The percentage of images with AOL>0.8 in each database is summarized in Table 3, which illustrates that the proposed method is effective for most of images.

Table 4. Average AOL and DSC obtained by different methods.

	MESSIDOR		Local	
	AOL	DSC	AOL	DSC
GVF [6]	0.4795	0.6354	0.4606	0.6155
VFC [10]	0.3433	0.5092	0.5896	0.7453
U-net-O [33]	0.8791	0.9299	0.8897	0.9348
U-net-S	0.7264	0.8245	0.8059	0.8894
Proposed	0.9144	0.9531	0.8980	0.9451

U-net-O is the original U-net introduced in [33]. U-net-S is the network applied in this work.

4.5. Comparison with hand designed external forces

To analyse the effectiveness of the proposed method, we compare the external force obtained by the GVF-Net with the hand designed external forces, including GVF [6]. Table 4 summarizes the results. It is observed that the results obtained by the proposed method are improved significantly compared with those obtained by the active contour models with GVF or VFC as external force, which illustrates that employing CNNs to derive an external force for active contour models is a very effective way. Besides, we also employ two neural networks for comparison. One is the original u-net introduced in [33] and we use the Dice coefficient as the loss function for it. The

other is the network introduced in Fig. 3 with the same loss function as GVF-Net. Note that these two neural networks are trained with image-region pairs. The results are summarized in Table 4. It is observed that the proposed method outperforms the u-net whether in terms of AOL or DSC, which verifies the effectiveness of the proposed method further.

4.6. Comparison with ad hoc algorithms

In this section, the ad hoc methods, namely designed for a specific task mentioned in this work, are selected for comparison.

4.6.1. For optic disk segmentation

Since most of the existing methods for optic disk segmentation are mainly evaluated according to AOL, thus AOL is selected as the metric for comparison in this section. If a method in the relevant literature has been not evaluated on an analysed database, the result will not be included in the table.

The existing methods for optic disk segmentation can be divided into unsupervised based methods and supervised based methods. Table 5 summarizes the comparison with several unsupervised methods. Even though the results obtained by the proposed method are superior to those obtained by the listed unsupervised methods, it might be unfair for these unsupervised methods since the test set used for evaluating the proposed method and the unsupervised methods is different. However, the images in DRIONS and ONHSD databases are not employed to train the GVF-Net, compared with the mathematical morphology based method [17], our method performs better on these two databases, which indicates that our method is very competitive.

Several supervised based methods are also selected for comparison, including a structured learning based method [22] and a deep learning based method [21]. Even though the training set and test set used to evaluate these supervised methods are different, we only employ the images in MESSIDOR database to train the GVF-Net while obtain very competitive results in DRIONS and ONHSD databases compared with the structured learning based method [22]. It is noteworthy to point out that K-fold cross validation is employed to construct training set and test set in [22], in which K=12 for MESSIDOR database, K=10 for DRIONS database and K=9 for ONHSD database.

Table 5. Comparison with different methods for optic disk segmentation with AOL metric.

	DRIONS	ONHSD	MESSIDOR
Unsupervised Methods			
Morphology [17]	0.8424	0.8045	0.8228
Circular Hough. [16]	-	-	0.8600
Sliding bank filter [19]	-	-	0.8886
Ellipse fitting [18]	-	-	0.8800
Supervised Methods			
Structured learning [22]	0.8473	0.8346	0.8636
Deep learning [21]	-	-	0.8888
Superpixels [20]	-	-	0.8750
Proposed	0.8972	0.8045	0.9144

Table 6. Comparison with different methods for fluid segmentation with DSC metric.

Methods	Graph cut [25]	RF [28]	Enface [23]	Blob [27]	Proposed
DSC	0.9053	0.9193	0.9460	0.9470	0.9451

Table 7. Comparison with different methods for fetal head circumference measurement.

Participant	ADF(mm) \pm std	DSC(%) \pm std	AD(mm) \pm std	HD(mm) \pm std
Mohbay	1.76 \pm 1.66	98.09 \pm 0.95	0.15 \pm 0.42	1.19 \pm 0.66
shenzexu	1.81 \pm 1.69	97.94 \pm 1.34	0.59 \pm 2.41	1.22 \pm 0.77
Proposed(cascaded)	2.18 \pm 2.40	95.53 \pm 3.98	-0.24 \pm 3.23	2.42 \pm 1.93
bacon	2.28 \pm 2.10	97.20 \pm 2.87	-1.05 \pm 2.92	1.62 \pm 1.03
dra_v	2.31 \pm 3.64	95.21 \pm 5.37	0.11 \pm 4.31	2.47 \pm 2.24
Proposed	2.45 \pm 2.55	95.49 \pm 4.11	-1.05 \pm 3.38	2.44 \pm 1.96
Baseline [29]	2.83 \pm 3.16	97.10 \pm 2.73	0.56 \pm 4.21	1.83 \pm 1.60

4.6.2. For fluid segmentation

Since the code of the existing methods for fluid segmentation is not open-sourced. Thus we re-implement the graph cut based method [25] and the random forest based method [28] and apply them on the same database with the proposed method for comparison. Besides, we also compare with the enface fundus driven method [23] and blob detection based method [27] via citing the best results recorded in the literature directly. Note that DSC is a common used metric to evaluate the performance of these methods, thus we compare with these methods based on DSC metric. Table 6 summarizes the results, where we can observe that the result obtained by the proposed method is very competitive with those obtained by the listed methods, which illustrates that the proposed method is very promising.

4.6.3. For fetal head segmentation

We can make direct comparison among different methods for fetal head segmentation since these methods have been evaluated on the same test set with the same metrics. The results obtained by different methods can be found from the home page of the challenge [43], in which some of them obtained by the state-of-the-art methods are listed in Table 7. It is observed that the proposed method is very competitive with the methods for this task also, which verifies the effectiveness of the proposed method further.

5. Discussion & conclusion

In this paper, we propose using GVF as reference to train a CNN to derive an external force for the parametric active contour models. The derived external force is also used for initializing the active contours. As a result, the quality of external force influences the performance of the proposed method directly. We set an experiment to verify this conclusion. Concretely, 20% of the images in DRIONS database are selected randomly as training set and 20% as validation set to fine tune the GVF-Net trained on MESSIDOR database. The remaining 60% is used as test set. The same setup is also for ONHSD database. Table 8 summarizes the results. It is observed that the results obtained by GVF-Net with fine tuning are better than those without fine tuning,

Table 8. Comparison with and without fine tuning on DRIONS and ONHSD databases.

	DRIONS		ONHSD	
	With	Without	With	Without
$AOL \geq 0.95$	15.15%	10.61%	5.56%	0
$AOL \geq 0.9$	81.82%	69.70%	61.11%	12.96%
$AOL \geq 0.85$	92.42%	86.36%	74.07%	57.41%
$AOL \geq 0.8$	95.45%	89.39%	87.04%	72.22%
$AOL \geq 0.7$	100%	96.97%	94.44%	81.48%
\overline{AOL}	0.9170	0.8984	0.8779	0.8119

which indicates that the quality of the external force plays a leading role in the performance of the proposed method.

To improve the external force, we employ a cascaded strategy developed in machine learning community, which has been introduced in Section 3.4. The effectiveness of the cascaded strategy can be verified by the results summarized in Table 7. We believe that other existing schemes in machine learning community might also have the capacity to improve the external force further, e.g. by improving the structure of the network. Besides, integrating the shape prior knowledge about the objects into the learning process [46] would be also another potential way to improve the external force.

In summary, in this paper, we propose a novel algorithm for biomedical image segmentation by deriving an external force via CNNs for the active contour models. Abundant experiments show that the proposed algorithm is very promising since it obtains very competitive results for different tasks.

Funding

The National Nature Science Foundation of China for Excellent Young Scholars (61622114); the National Basic Research Program of China (973 Program) (2014CB748600); the National Natural Science Foundation of China (81371629, 81401472, 61401294, 61401293, 61771326, 61601317); the International Cooperation Project of Ministry of Science and Technology (2016YFE010770).

Disclosures

The authors declare that there are no conflicts of interest related to this article.

References

1. S. D. Olabarriaga and A. W. M. Smeulders, "Interaction in the segmentation of medical images: A survey," *Med. Image Analysis* **5**, 127–142 (2001).
2. X. Chen and L. Pan, "A survey of graph cuts/graph search based medical image segmentation," *IEEE Rev. Biomed. Eng.* **11**, 112–124 (2018).
3. G. J. S. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. V. Der Laak, B. Van Ginneken, and C. I. Sanchez, "A survey on deep learning in medical image analysis," *Med. Image Analysis* **42**, 60–88 (2017).
4. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. journal computer vision* **1**, 321–331 (1988).
5. B. Wu and Y. Yang, "Local- and global-statistics-based active contour model for image segmentation," *Math. Probl. Eng.* **2012**, 94–113 (2012).
6. C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Transactions on Image Process.* **7**, 359–369 (1998).

7. J. Tang, "A multi-direction gvf snake for the segmentation of skin cancer images," *Pattern Recognit.* **42**, 1172–1179 (2009).
8. S. Zhu and R. Gao, "A novel generalized gradient vector flow snake model using minimal surface and component-normalized method for medical image segmentation," *Biomed. Signal Process. & Control.* **26**, 1–10 (2016).
9. Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE* **86**, 2278–2324 (1998).
10. L. Bing and S. T. Acton, "Active contour external force using vector field convolution for image segmentation," *IEEE Transactions on Image Process.* **16**, 2096–2106 (2007).
11. R. Srivastava, J. Cheng, D. W. K. Wong, and J. Liu, "Using deep learning for robustness to parapapillary atrophy in optic disc segmentation," in *IEEE International Symposium on Biomedical Imaging*, (2015), pp. 768–771.
12. M. R. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac mri," *Med. Image Analysis* **30**, 108–119 (2016).
13. Z. Li, A. Lin, X. Yang, and J. Wu, "Left ventricle segmentation by combining convolution neural network with active contour model and tensor voting in short-axis mri," in *IEEE International Conference on Bioinformatics and Biomedicine*, (2017), pp. 736–739.
14. A. Hoogi, A. Subramaniam, R. Veerapaneni, and D. L. Rubin, "Adaptive estimation of active contour parameters using convolutional neural networks and texture analysis," *IEEE Transactions on Med. Imaging* **36**, 781–791 (2017).
15. T. H. N. Le, K. G. Quach, K. Luu, N. D. Chi, and M. Savvides, "Reformulating level sets as deep recurrent neural network approach to semantic segmentation," *IEEE Transactions on Image Process.* **27**, 2393–2407 (2018).
16. A. Aquino, M. E. Gegúndez-Arias, and D. Marín, "Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques," *IEEE transactions on Med. Imaging* **29**, 1860–1869 (2010).
17. S. Morales, V. Naranjo, J. Angulo, and M. Alcañiz, "Automatic detection of optic disc based on pca and mathematical morphology," *IEEE transactions on Med. Imaging* **32**, 786–796 (2013).
18. A. Giachetti, L. Ballerini, and E. Trucco, "Accurate and reliable segmentation of the optic disc in digital fundus images," *J. Med. Imaging* **1**, 1–11 (2014).
19. B. Dashtbozorg, A. M. Mendonça, and A. Campilho, "Optic disc segmentation using the sliding band filter," *Comput. Biol. Medicine* **56**, 1–12 (2015).
20. J. Cheng, J. Liu, Y. Xu, F. Yin, D. W. K. Wong, N.-M. Tan, D. Tao, C.-Y. Cheng, T. Aung, and T. Y. Wong, "Supapixel classification based optic disc and optic cup segmentation for glaucoma screening," *IEEE transactions on Med. Imaging* **32**, 1019–1032 (2013).
21. G. Lim, Y. Cheng, W. Hsu, and M. L. Lee, "Integrated optic disc and cup segmentation with deep learning," in *IEEE International Conference on TOOLS with Artificial Intelligence*, (2016), pp. 162–169.
22. Z. Fan, Y. Rong, X. Cai, J. Lu, W. Li, H. Lin, and X. Chen, "Optic disk detection in fundus image based on structured learning," *IEEE J. Biomed. Heal. Informatics* **22**, 224–234 (2018).
23. M. Wu, Q. Chen, X. He, P. Li, W. Fan, S. Yuan, and H. Park, "Automatic subretinal fluid segmentation of retinal sd-oct images with neurosensory retinal detachment guided by enface fundus imaging," *IEEE Transactions on Biomed. Eng.* **65**, 87–95 (2018).
24. Y. Xu, K. Yan, J. Kim, X. Wang, C. Li, L. Su, S. Yu, X. Xu, and D. D. Feng, "Dual-stage deep learning framework for pigment epithelium detachment segmentation in polypoidal choroidal vasculopathy," *Biomed. Opt. Express* **8**, 4061–4076 (2017).
25. X. Chen, M. Niemeijer, L. Zhang, K. Lee, M. D. Abramoff, and M. Sonka, "3d segmentation of fluid-associated abnormalities in retinal oct: Probability constrained graph-search -graph-cut," *IEEE transactions on Med. Imaging* **31**, 1521–1531 (2012).
26. T. Hassan, A. M. Usman, B. Hassan, A. M. Syed, and S. A. Bazaz, "Automated segmentation of subretinal layers for the detection of macular edema," *Appl. Opt.* **55**, 454–461 (2016).
27. Z. Ji, Q. Chen, M. Wu, S. Niu, W. Fan, S. Yuan, and Q. Sun, "Beyond retinal layers: A large blob detection for subretinal fluid segmentation in sd-oct images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, (2018), pp. 372–380.
28. A. Lang, A. Carass, E. K. Swingle, O. Al-Louzi, P. Bhargava, S. Saidha, H. S. Ying, P. A. Calabresi, and J. L. Prince, "Automatic segmentation of microcystic macular edema in oct," *Biomed. optics express* **6**, 155–169 (2015).
29. T. L. A. V. D. Heuvel, D. D. Bruijn, C. L. D. Korte, and B. V. Ginneken, "Automated measurement of fetal head circumference using 2d ultrasound images," *Plos One* **13**, 1–20 (2018).
30. L. Wu, Y. Xin, S. Li, T. Wang, P. A. Heng, and D. Ni, "Cascaded fully convolutional networks for automatic prenatal ultrasound image segmentation," in *IEEE International Symposium on Biomedical Imaging*, (2017), pp. 663–666.
31. L. Zhang, X. Ye, T. Lambrou, W. Duan, N. Allinson, and N. J. Dudley, "A supervised texton based approach for automatic segmentation and measurement of the fetal head and femur in 2d ultrasound images," *Phys. Medicine & Biol.* **61**, 1095–1115 (2016).
32. S. Rueda, S. Fathima, C. L. Knight, M. Yaqub, A. T. Papageorghiou, B. Rahmatullah, A. Foi, M. Maggioni, A. Pepe, and J. Tohka, "Evaluation and comparison of current fetal ultrasound image segmentation methods for biometric measurements: a grand challenge," *IEEE Transactions on Med. Imaging* **33**, 797 (2014).
33. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, (2015), pp. 234–241.

34. N. Otsu, "A threshold selection method from gray-level histograms," *Automatica* **11**, 23–27 (1975).
35. D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognit.* **37**, 1–19 (2004).
36. J. Illingworth and J. Kittler, "The adaptive hough transform," *IEEE Transactions on Pattern Analysis Mach. Intell.* **9**, 690–698 (1987).
37. M. Liu, D. Cheng, K. Wang, and Y. Wang, "Multi-modality cascaded convolutional neural networks for alzheimer disease diagnosis," *Neuroinformatics* **16**, 1–14 (2018).
38. M. T.-V. Project, "Messidor: Digital retinal images france," <http://messidor.crihan.fr/download-en.php> (2008).
39. E. J. Carmona, M. Rincón, J. García-Feijóo, and J. M. Martínez-de-la Casa, "Identification of the optic nerve head with genetic algorithms," *Artif. Intell. Medicine* **43**, 243–259 (2008).
40. U. Lincoln, "etinal image computing & understanding, onhsd-optic nerve head segmentation dataset," <http://reviewdb.lincoln.ac.uk/ImageDatasets/ONHSD.aspx> (2004).
41. J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher, and L. Kennedy, "Optic nerve head segmentation," *IEEE Transactions on Med. Imaging* **23**, 256–264 (2004).
42. I. Zafar, G. Tzanidou, R. Burton, N. Patel, and L. Araujo, *Hands-On Convolutional Neural Networks with TensorFlow* (Packt Publishing Ltd, 201-202(2018)).
43. T. van den Heuvel, D. de Bruijn, C. L. de Korte, and B. van Ginneken, "Automated measurement of fetal head circumference," <https://doi.org/10.5281/zenodo.1327317> (2018).
44. A. W. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least squares fitting of ellipses," *IEEE Transactions on Pattern Analysis Mach. Intell.* **21**, 476–480 (1999).
45. A. Vedaldi and K. Lenc, "Matconvnet – convolutional neural networks for matlab," in *Proceeding of the ACM Int. Conf. on Multimedia*, (2015), pp. 689–692.
46. F. Milletari, A. Rothberg, J. Jia, and M. Sofka, "Integrating statistical prior knowledge into convolutional neural networks," in *International Conference on Medical Image Computing & Computer-assisted Intervention*, (2017), pp. 161–168.