

# Dual-fusion Active Contour Model with Semantic Information for Saliency Target Extraction of Underwater Images

**Shudi Yang**

University of Science and Technology Beijing

**Zhehan Chen** (✉ [chenzh\\_ustb@163.com](mailto:chenzh_ustb@163.com))

University of Science and Technology Beijing

**Jiaxiong Wu**

University of Science and Technology Beijing

**Zhipeng Feng**

University of Science and Technology Beijing

---

## Research Article

**Keywords:** underwater image, target contour extraction, active contour model, semantic information, saliency target

**Posted Date:** December 6th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1132555/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Applied Sciences on February 28th, 2022. See the published version at <https://doi.org/10.3390/app12052515>.

# Dual-fusion Active Contour Model with Semantic Information for Saliency Target Extraction of Underwater Images

Shudi Yang<sup>1</sup>, Zhehan Chen<sup>1,\*</sup>, Jiaxiong Wu<sup>1</sup>, Zhipeng Feng<sup>1</sup>

\*Zhehan Chen, E-mail: [chenzh\\_ustb@163.com](mailto:chenzh_ustb@163.com)

(1. School of Mechanical Engineering, University of Science and Technology, Beijing, 100083, China)

**Abstract:** Underwater vision research is the foundation of marine-related disciplines. The target contour extraction is of great significance to target tracking and visual information mining. Aiming at the problem that conventional active contour models cannot effectively extract the contours of salient targets in underwater images, we propose a dual-fusion active contour model with semantic information. First, the saliency images are introduced as semantic information, and extract salient target contours by fusing Chan–Vese and local binary fitting models. Then, the original underwater images are used to supplement the missing contour information by using the local image fitting. Compared with state-of-the-art contour extraction methods, our dual-fusion active contour model can effectively filter out background information and accurately extract salient target contours.

**Keywords:** underwater image, target contour extraction, active contour model, semantic information, saliency target.

## 1 Introduction

In recent years, the development and utilization of the ocean have gradually become an important development direction. Since underwater vision research is the basis of marine-related disciplines, the rapid development of underwater image processing technology is inevitable<sup>[1-2]</sup>. Image segmentation is a basic method of target extraction, which aims to partition an image into several meaningful and constituent regions and each region has coherent features such as intensities, colors, and textures<sup>[3]</sup>.

Now, some results have been achieved in underwater image segmentation. Liu et al.<sup>[4]</sup> have proposed an improved level set algorithm based on the gradient descent method, and applied to segment underwater biological image. Wei et al.<sup>[5]</sup> have improved the K-means algorithm to segment underwater image background, and addressed the issue of improper K value determination. And then this algorithm can minimize the impact of initial centroid position of grayscale image. SM et al.<sup>[6]</sup> have used canny edge detection algorithm to segment underwater images, whereas canny edge detection algorithm was greatly affected by background noise. Sun et al.<sup>[7]</sup> and Li et al.<sup>[9]</sup> have used fuzzy C-means algorithm to segment underwater images. Rajeev et al.<sup>[8]</sup> have used K-means algorithm to segment underwater images. However, the aforesaid clustering algorithms have been greatly affected by local gray unevenness of underwater images. Also, clustering algorithms had local convergence errors and were only suitable for underwater images with a single background gray level.

Some investigators have segmented underwater images based on optical properties and achieved results. For example, Chen et al.<sup>[10]</sup> have proposed an optical feature extraction, calculation, and decision

38 method to identify the collimated region of artificial light, and employed a level set method to segment  
39 the objects within the collimated region. This method could better identify the target region, but level set  
40 method could not filter out background noise when the target region contains background information.  
41 Xuan<sup>[11]</sup> et al. have proposed a RGB color channel fusion segmentation method for underwater images.  
42 The proposed method obtained the grayscale image with high foreground-background contrast and  
43 employed thresholding segmentation method to conduct fast segmentation. However, the disadvantage  
44 of this method is that when the color of the background region is similar to the foreground region, the  
45 target cannot be segmented.

46 Active contour model have also been used for underwater image segmentation. Zhu et al.<sup>[12]</sup> used  
47 the cluster-based algorithm for co-saliency detection, and made salient region in the underwater images  
48 be highlighted. And then the local statistical active contour model was used to extract the target contours  
49 of underwater images. Qiao et al.<sup>[13]</sup> proposed an improved method based on active contour model. The  
50 method used the RGB color space and the contrast limited adaptive histogram equalization(CLAHE)  
51 method to increase the contrast of the sea cucumber thorns and body, respectively. Then, the method  
52 extracted the edge of the sea cucumber thorns by active contour model. Li et al.<sup>[14]</sup> have improved the  
53 traditional level set methods by avoiding the calculation of signed distance function (SDF) to segment  
54 underwater images. The improved method could speed up the computational complexity without re-  
55 initialisation. Bai et al.<sup>[15]</sup> proposed a method based on morphological component analysis (MCA) and  
56 adaptive level set evolution to segment underwater images. The MCA was used to sparse decompose the  
57 image into texture and cartoon parts. The new adaptive level set evolution method combined the threshold  
58 piecewise function with variable right coefficient and halting speed function and was used to obtain the  
59 edges of the cartoon part. Shelei et al.<sup>[16]</sup> segmented underwater grayscale images by fusing the geodesic  
60 active contour model (GAC) and the Chan–Vese (CV) model. However, this method required that the  
61 target region of the underwater image has uniform grayscale. Chen et al.<sup>[17]</sup> integrated the transmission  
62 map and the saliency map into a unified level set formulation to extract the salient target contours of the  
63 underwater images.

64 As a new technology of image processing, neural network has also been used for underwater image  
65 segmentation. O’Byrne et al.<sup>[18]</sup> have proposed the use of photorealistic synthetic imagery for training  
66 deep encoder–decoder network. This method synthesized virtual underwater images and each rendered  
67 image had a corresponding ground truth per-pixel label map. Then established the mapping relationship  
68 between the underwater images and the segmented images by training the encoder–decoder network.  
69 Zhou et al.<sup>[19]</sup> have proposed a deep neural network architecture for underwater scene segmentation. The  
70 architecture extracted feature by pre-training VGG-16 and learned to expand the lower resolution feature  
71 maps by the decoder. The neural network has achieved certain results in underwater image segmentation,  
72 but the lack of underwater data sets with corresponding functions is still a problem.

73 In general, most of the existing underwater image segmentation methods are used to segment images  
74 with high foreground-background contrast and single background grayscale. When the underwater  
75 images with varying background grayscale and the targets have complex texture, the segmentation results

76 of the above methods are not satisfactory. To address the above problem, we propose a novel dual-fusion  
 77 model with semantic information for salient object segmentation of underwater images with complex  
 78 background. In summary, the contributions of our model are as follows:

- 79 • We introduce saliency maps as semantic information to segment foreground information and  
 80 background information;
- 81 • The dual-fusion energy equation is proposed to extract the contours of saliency targets by integrating  
 82 local and global intensity fitting term;
- 83 • For the missing saliency target information, we propose the correction module to correct the saliency  
 84 target contour error by introducing the original image contour information.

85 This paper is organized as follows. Section 2 reviews related works. In Section 3, we introduce in  
 86 detail the derivation process of the dual-fusion model. Section 4 shows the experimental process and we  
 87 compare the proposed method with state-of-the-art segmentation methods, and the results demonstrate  
 88 the superiority of the proposed methods. Section 5 presents the discussion about the parameters of  
 89 proposed model. The conclusion of this paper is shown in Section 6.

## 90 2 Related works

### 91 2.1 The C–V model

92 The Chan–Vese (CV) model<sup>[20]</sup> is initially derived from the Mumford–Shah (MS) functional<sup>[21]</sup>. The  
 93 MS functional aims to find an optimal piecewise smooth approximation image  $I : \Omega_i \subset R$  from the  
 94 original image  $I_0 : \Omega \subset R^2$ , the energy functional of MS can be expressed as follows:

$$95 \quad E^{MS}(I, C) = \int_{\Omega} (I_0 - I)^2 dx + \mu \int_{\Omega_C} |\nabla I|^2 dx + \nu |C| \quad (1)$$

96 where  $\mu, \nu \geq 0$  are positive weighting constants,  $|C|$  is the length of the contour  $C$ . However, the non-  
 97 convexity of the above energy functional make it difficult to be minimized, so the CV model has been  
 98 proposed to simplify and modify the above functional. The energy functional of the CV model can be  
 99 defined as follows:

$$100 \quad E^{CV}(C, c_1, c_2) = \lambda_1 \int_{in(C)} (I_0 - c_1)^2 dx + \lambda_2 \int_{out(C)} (I_0 - c_2)^2 dx + \nu \cdot len(C) + \mu \cdot area(in(C)) \quad (2)$$

101 where  $\mu, \nu, \lambda_1, \lambda_2 \geq 0$  are positive parameters,  $in(C)$  and  $out(C)$  represent the region inside and outside  
 102 of the contour  $C$ ,  $c_1$  and  $c_2$  are two constants that approximate the image intensity in  $in(C)$  and  $out(C)$ ,  
 103 respectively. The Euclidean length term  $len(C)$  is used to regularize the contour. The first two terms in  
 104 Eq. (2) are the global binary fitting energy. This energy can be represented by a level set formulation, and  
 105 then the energy minimization problem can be converted to solving a level set evolution equation, the  
 106 evolution equation can be expressed as follows:

$$107 \quad c_1 = \frac{\dot{\mathcal{O}}_W I_0 \times H(f) dx}{\dot{\mathcal{O}}_W H(f) dx}, c_2 = \frac{\dot{\mathcal{O}}_W I_0 \times (1 - H(f)) dx}{\dot{\mathcal{O}}_W (1 - H(f)) dx} \quad (3)$$

$$108 \quad \frac{\partial \phi}{\partial t} = \delta(\phi) \left[ \nu \operatorname{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) - \mu - \lambda_1 (I_0 - c_1)^2 + \lambda_2 (I_0 - c_2)^2 \right] \quad (4)$$

109 where  $H(\bullet)$  is the Heaviside function, and  $\delta(\bullet)$  is the Delta function which is derivative of the  
 110 Heaviside function. For Eq. (4),  $\nu$  is a scaling parameter. If the  $\nu$  is small enough, the small targets are  
 111 likely to be extracted; if the  $\nu$  is large, the large targets can be detected.

112 Whereas the global fitting will not be accurate if the image intensities are inhomogeneous. Therefore,  
 113 the CV model is not suitable for inhomogeneous images and the segmentation results are affected by the  
 114 position of the initial level set<sup>[3]</sup>. But the CV model has better robustness to noise.

## 115 2.2 The LIF model

116 The local image fitting<sup>[22]</sup> (LIF) energy functional is defined as follows:

$$117 \quad \frac{\partial \phi}{\partial t} = (I(x) - I^{LFI}(x))(m_1 - m_2)\delta_\varepsilon(\phi) \quad (5)$$

118 where  $I^{LFI}(x)$  is a local fitted image:

$$119 \quad I^{LFI}(x) = m_1 H_\varepsilon(\phi(x)) + m_2 (1 - H_\varepsilon(\phi(x))) \quad (6)$$

120 where  $m_1$  and  $m_2$  are averages of image intensities of Gaussian window inside and outside the contour,  
 121 respectively.  $m_1, m_2$  can be expressed as follows:

$$122 \quad \begin{cases} m_1 = \text{mean}(I \in (\{x \in \Omega \mid \phi(x) < 0\} \cap W_k(x))) \\ m_2 = \text{mean}(I \in (\{x \in \Omega \mid \phi(x) > 0\} \cap W_k(x))) \end{cases} \quad (7)$$

123 where  $W_k(x)$  is a truncated Gaussian window or a constant window.

124 And then, the LIF model used the variation calculus and the steepest descent method to minimize  
 125  $E^{LIF}(\phi)$ , and the level set evolution equation can be expressed as follows:

$$126 \quad \frac{\partial \phi}{\partial t} = (I(x) - I^{LFI}(x))(m_1 - m_2)\delta_\varepsilon(\phi) \quad (8)$$

## 127 3 Dual-fusion Active Contour Model

128 In this section, we propose a dual-fusion active contour model with semantic information to extract  
 129 target contours of underwater images. Without the semantic information, the existing methods cannot  
 130 individually extract the target contour from the background. So it is necessary to introduce semantic  
 131 information and roughly extract the saliency target contour from the complex background. To avoid the  
 132 extraction error of saliency target, we introduce the original image contour to correct and supplement the  
 133 missing contour information. By semantic information and correction module, the proposed model can  
 134 accurately extract the saliency target contour from the complex background.

### 135 3.1 Saliency image fitting energy

136 In this paper, we used the pyramid feature attention network<sup>[23]</sup> to acquire the saliency images.  
 137 However, due to the low contrast of underwater images, there were some errors in the saliency detection  
 138 results such as local inhomogeneous intensity, background noise, and missing contour information. In  
 139 view of the local inhomogeneous intensity of the saliency images, we preliminarily employ the local  
 140 binary fitting to construct the energy functional  $E_{sal}$ :

141 
$$E_{sal}(C, f_1(x), f_2(x)) = \lambda_1 \int_{in(C)} (S - f_1(x))^2 dx + \lambda_2 \int_{out(C)} (S - f_2(x))^2 dx \quad (9)$$

142 where  $S$  is the saliency images,  $C$  is a contour in the image domain  $\Omega$ ,  $f_1$  and  $f_2$  are image local  
 143 fitting intensities near the point  $x$ . The local fitting intensities  $f_1$ ,  $f_2$  can be expressed as follows<sup>[24-25]</sup>:

144 
$$f_1(x) = \frac{K_\sigma(x) * [H_\epsilon(\phi(x))S]}{K_\sigma(x) * H_\epsilon(\phi(x))} \quad (10)$$

145 
$$f_2(x) = \frac{K_\sigma(x) * [(1 - H_\epsilon(\phi(x)))S]}{K_\sigma(x) * [1 - H_\epsilon(\phi(x))]} \quad (11)$$

146 where  $K_\sigma(x)$  is the Gaussian kernel,  $S$  is the saliency images,  $H_\epsilon$  is the Heaviside function  $H(\bullet)$   
 147 and can be expressed as:

148 
$$H_\epsilon(x) = \frac{1}{2} \left[ 1 + \frac{2}{\pi} \arctan\left(\frac{x}{\epsilon}\right) \right] \quad (12)$$

149 However, the local binary fitting may introduce some local minimums and is sensitive to noise.  
 150 Affected by the accuracy of saliency detection, saliency map of underwater images will inevitably have  
 151 background noise. Also the initialization curve greatly affects the segmentation results. To solve the  
 152 aforesaid problems, we introduce the global fitting term from the CV model into the energy functional  
 153  $E_{sal}$ . The local-global fitting intensities can be expressed as follows:

154 
$$\begin{cases} I_1 = \omega c_1 + (1 - \omega) f_1 \\ I_2 = \omega c_2 + (1 - \omega) f_2 \end{cases} \quad (13)$$

155 where  $I_1$  and  $I_2$  are mixed intensity,  $c_1$  and  $c_2$  are two constants derived from Eq.(3),  $\omega$  is a weight  
 156 coefficient ( $0 \leq \omega \leq 1$ ). According to the test images in this paper, the value of  $\omega$  can be taken from 0.5  
 157 to 0.9. And the more inhomogeneous the image intensity, the smaller the value of  $\omega$ .

158 With the level set representation, the energy functional can be expressed as follows:

159 
$$E_{sal}(\phi, I_1(x), I_2(x)) = \lambda_1 \int_{\Omega} (S - I_1(x))^2 H(\phi(x)) dx + \lambda_2 \int_{\Omega} (S - I_2(x))^2 (1 - H(\phi(x))) dx \quad (14)$$

160 The improved fitting energy  $E_{sal}$  not only take local intensity information into account but also  
 161 avoid the local minimization. Therefore, for the saliency images of underwater images, the improved  
 162 energy functional can extract the contour of the inhomogeneous images more accurately.

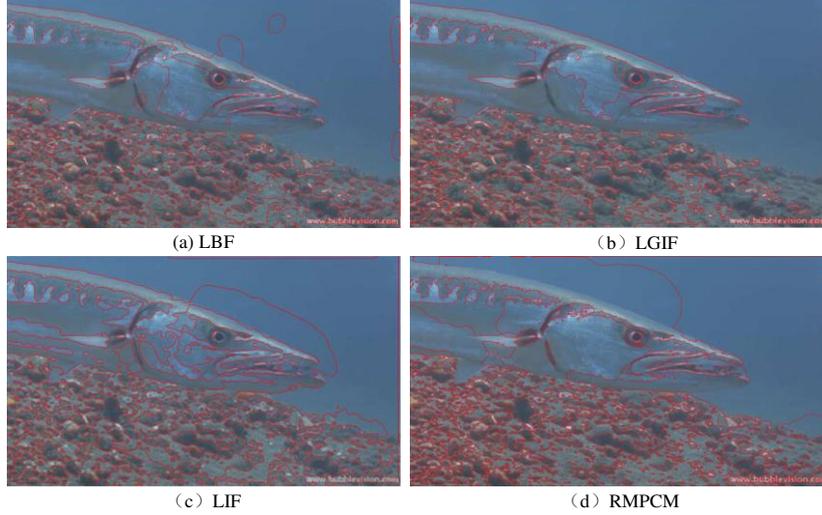
### 163 3.2 Original image fitting energy

164 The problems of local inhomogeneous intensity and noise can be solved by fusing the local intensity  
 165 fitting and CV model. However, the missing contour information of saliency image still needs to be solved.  
 166 Therefore, the original underwater images be used to make up the missing contour information.

167 In this paper, we used the local image fitting model (LIF)<sup>[22]</sup> to extract the contour of original  
 168 underwater images. The energy functional  $E_{org}$  can be expressed as:

169 
$$E_{org}(\phi) = \frac{1}{2} \int_{\Omega} |I(x) - I^{LFI}(x)|^2 dx, x \in \Omega, \quad (15)$$

170 where  $I^{LFI}(x)$  is a local fitted image, as shown in Eq. (6). Although the models such as LBF<sup>[24-25]</sup>,  
 171 RMPCM<sup>[3]</sup>, and LGIF<sup>[26]</sup> can extract the target contours of underwater images very well, as shown in Fig.  
 172 1, the LIF model has higher efficiency. The higher efficiency is because that the energy functional of the  
 173 LIF model does not include a kernel function. Also, the LIF model can well fit the original image, while  
 174 reducing the noise significantly by minimizing the difference between the fitted image and the original  
 175 image.



176  
 177 Fig. 1 (a)-(d) shows the segmentation results of the LBF, LGIF, LIF, and RMPCM model, respectively.

178 In Fig. 1, LBF, LGIF, and LIF models could better extract the target contour, but LBF was more  
 179 sensitive to the initial contour curve. The energy functional of LGIF and RMPCM both involved kernel  
 180 function. The kernel function performs more than one convolution operations for each iteration step, so  
 181 the evolution speed is slow. The running time of the above models are shown in Table 1.

182 Table 1 Iterations and CPU time (in seconds)

	LBF	LGIF	LIF	RMPCM
	1239 × 731pixels			
Iterations	200	200	200	200
Time (s)	93.2969	55.5938	38.5469	63.3052

183 Fig. 1 and Table 1 intuitively show that LIF model has advantages in both speed and contour  
 184 extraction results. So we use the LIF model to extract the original image contour to correct the contour  
 185 information of the salient target.

### 186 3.3 Dual-fusion Active Contour Model

187 To take smaller fitting energy at target contours than at other locations, we use an edge indicator  
 188 function<sup>[27]</sup>. The function can be expressed as follows:

$$189 \quad g @ \frac{1}{1+|\nabla G_{\sigma} * I|^2} \quad (16)$$

190 Then we define the dual-fusion intensity fitting energy functional as follows:

$$191 \quad E^{DFIF}(\phi) = g \left[ \omega_1 E_{org} + (1 - \omega_1) E_{sal} \right] \quad (17)$$

192 where  $\omega_1$  is a weight coefficient ( $0 \leq \omega_1 \leq 1$ ),  $E_{org}$  and  $E_{sal}$  are the original images fitting energy  
 193 functional and the saliency images fitting energy functional, respectively.

194 Finally, the dual-fusion intensity fitting energy functional  $E^{DFIF}(\phi)$  can be expressed as:

$$\begin{aligned}
 E^{DFIF}(\phi, I_1, I_2) &= g \left[ \omega_1 E_{org} + (1 - \omega_1) E_{sal} \right] \\
 &= g \left[ \omega_1 \frac{1}{2} \int_{\Omega} |I(x) - I^{LF1}(x)|^2 dx + (1 - \omega_1) \left( \lambda_1 \int_{\Omega} (S - I_1(x))^2 H_{\varepsilon}(\phi(x)) dx, \right. \right. \\
 &\quad \left. \left. + \lambda_2 \int_{\Omega} (S - I_2(x))^2 (1 - H_{\varepsilon}(\phi(x))) dx \right) \right]
 \end{aligned} \tag{18}$$

196 And then we minimize  $E^{DFIF}(\phi, I_1, I_2)$  with respect to  $\phi$  to get the corresponding gradient descent  
 197 flow<sup>[24-26]</sup>:

$$\frac{\partial \phi}{\partial t} = g \delta_{\varepsilon}(\phi) \left[ \omega_1 e_1 + (1 - \omega_1) e_2 \right] \tag{19}$$

199 where

$$\begin{cases} e_1 = (I - m_1 H_{\varepsilon}(\phi(x)) - m_2 (1 - H_{\varepsilon}(\phi(x)))) (m_1 - m_2) \\ e_2 = -\lambda_1 (S - I_1(x))^2 + \lambda_2 (S - I_2(x))^2 \end{cases} \tag{20}$$

201 where  $I, S$  are the original images and the saliency images, respectively.  $I_1(x)$  represents the integrated  
 202 local and global intensities,  $m_1$  and  $m_2$  are averages of the image intensities in a Gaussian window inside  
 203 and outside the contour.

### 204 3.4 Regularize the level set function

205 As pointed out by Ref.[22], Gaussian filtering can replace the traditional regularized term to  
 206 regularize the level set function. Therefore, the smoothing process of the level set function can be  
 207 expressed as:

$$\phi^{k+1} = G_{\eta} * \phi^k, \eta > \sqrt{\Delta t} \tag{21}$$

209 where  $\eta$  is the standard deviation, and  $\Delta t$  is the time-step.

210 In fact, the smoothing effect of the level set function by Gaussian filtering is slightly worse than the  
 211 traditional regularized term and is greatly affected by the time-step. However, the computing efficiency  
 212 of Gaussian filtering is much higher than the traditional regularized term.

## 213 4 Experimental analysis and results

214 In this section, the proposed method was tested on several underwater images with intensity  
 215 inhomogeneity. Also, the method compared with some state-of-the-art contour extraction methods in  
 216 efficiency and accuracy. In order to ensure the fairness of the comparison results, all contour extraction  
 217 results were produced on the same computer. And the computer was configured as Intel(R) Core(TM) i7-  
 218 8650U CPU @ 2.11 GHz, 16.00 GB memory, Windows 10 system, and x64 processor. MatlabR2017a is  
 219 software platform. We use the same parameters  $\eta^2 = 6, \sigma = 2, \varepsilon = 1, \lambda_1 = 3, \lambda_2 = 1$  and time-step  $\Delta t = 0.1$ . The  
 220 initial level set function is defined by

221 
$$\phi(x, t = 0) = \begin{cases} -c_0, & x \in \text{in}(C) \\ 0, & x \in C \\ c_0, & x \in \text{out}(C) \end{cases}, \quad (22)$$

222 where  $c_0 > 0$  is a constant((in our experiments,  $c_0 = 1$ ),  $\text{in}(C)$  and  $\text{out}(C)$  represent the region inside and  
 223 outside of the contour  $C$ , respectively.

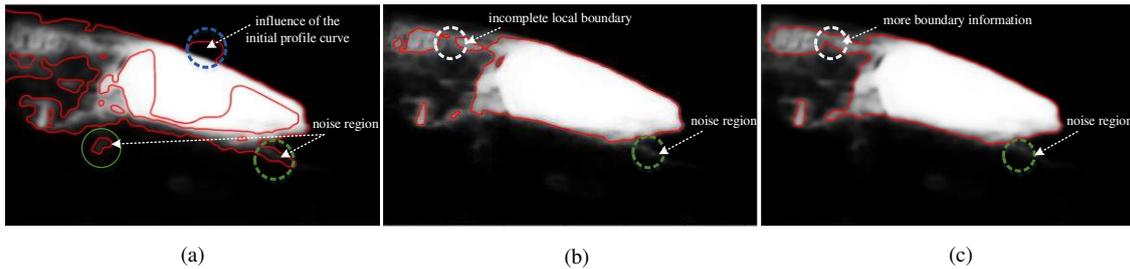
224 **4.1 The benefits of local-global intensity fitting**

225 A comparative experiment was performed to prove the effectiveness of the local-global intensity  
 226 fusion in Section 3.1. We conducted different experiments, as shown in Table 2.

227 Table 2 The comparative experiment of local-global intensity

Experiments	local intensity	global intensity
A	✓	
B		✓
C(our fusion intensity)	✓	✓

228 In Experiment A, the fitting intensity of energy functional is local intensity. In Experiment B, the  
 229 fitting intensity of energy functional is global intensity. And the energy functional with fusion local-global  
 230 intensity is shown in Experiment C. The contour extraction results of experiments are shown in Fig. 2.



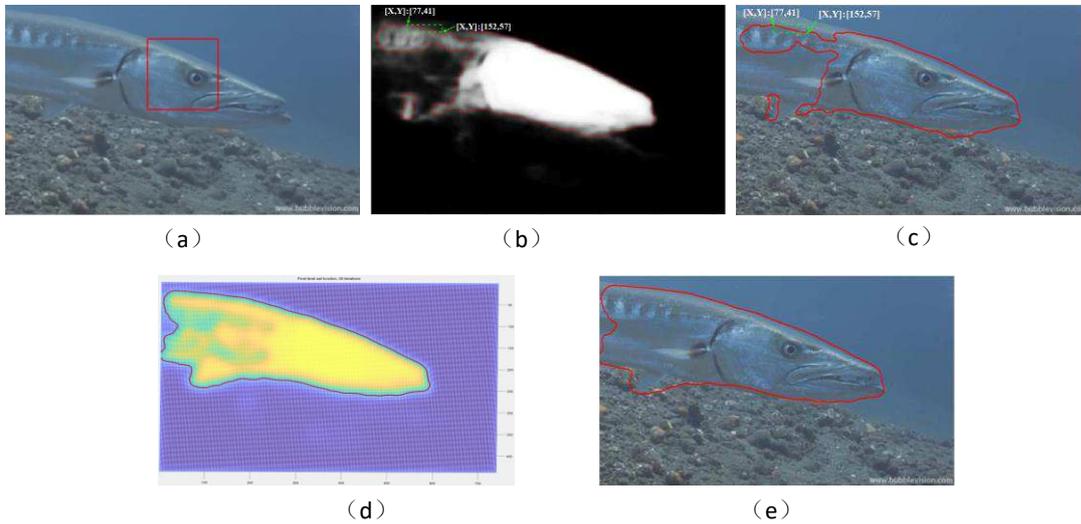
231 (a) (b) (c)  
 232 Fig. 2 The contour extraction results. (a) result of the local intensity fitting. (b) result of the global intensity  
 233 fitting. (a) result of our method.

234 As shown in Fig.2, Experiment A could extract the target contour in intensity inhomogeneity region,  
 235 but the result was greatly affected by the initial contour curve (blue circled area) and was sensitive to  
 236 noise (green circled area). And the method of Experiment A also extracts the contours of non-boundary  
 237 regions. Experiment B could extract the target contour in intensity homogeneity region and was not  
 238 disturbed by noise, but the target contour in intensity inhomogeneity region cannot be extracted. Our  
 239 method could not only extract the target contour in intensity homogeneity region and inhomogeneity  
 240 region, but also not be disturbed by noise.

241 **4.2 The effect of original image correction**

242 Fig. 3 shows the result of our method on the underwater images segmentation. As shown in Fig. 3  
 243 (b) and (c), the coordinate points  $\{[X, Y]:[77, 41]\}$  and  $\{[X, Y]:[152, 57]\}$  located at the saliency target  
 244 edge in Fig. 3 (b). But in Fig. 3 (c), the coordinate point at the same position is inside the target instead  
 245 of on the target edge. This error is caused by the deviation of saliency detection. Therefore, it is necessary

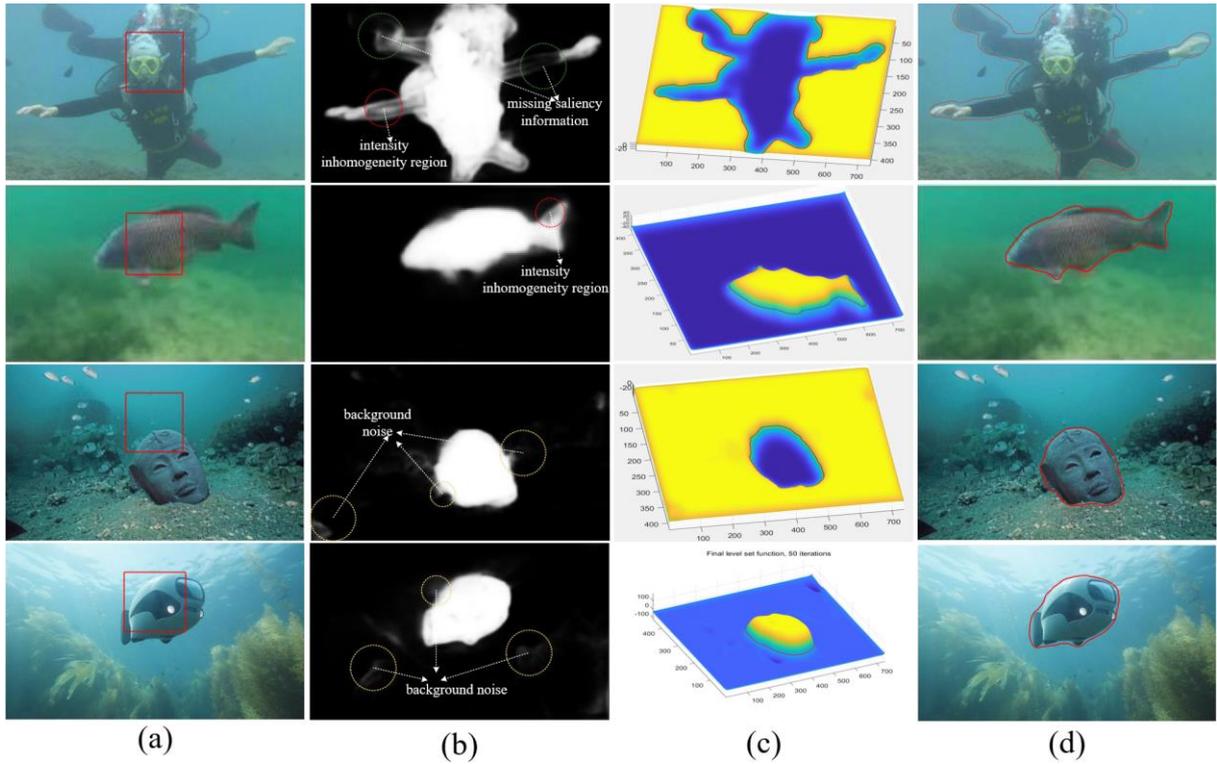
246 to use the original image to supplement the missing information. This paper used the local image fitting  
 247 method to extract the contour information of the original image, and then used the contour information  
 248 to correct the deviation caused by saliency detection. The result of the correction is shown in Fig. 3 (e).  
 249 As shown in Fig. 3 (e), the missing contour information of saliency image is accurately supplemented,  
 250 and the background information is filtered out.



251  
 252 Fig. 3 The results of our method. (a) the original underwater image with initial zero level contour. (b) the  
 253 contour extraction result of saliency target. (c) the contour extraction result without correction. (d) the  
 254 final level set function. (e) the result of our method.

255 **4.3 Performance of Dual-fusion active contour model**

256 Fig.4 shows the performance of our method, it can be seen from the Fig.4 (d) that our method can  
 257 filter out the background information and accurately extract the target contour. Fig.4 (b) are the saliency  
 258 images of the original underwater images, the red circled represent the intensity inhomogeneity region,  
 259 the yellow circled represent the noise region, and the green circled represent the missing region of target.  
 260 For the regions of intensity inhomogeneity and noise, our method can still extract the target contour well  
 261 by the local-global intensity fitting term. Also, the saliency image of the first image obviously lacks part  
 262 of the target information (green circled region), our method can still extract the complete target contour  
 263 by integrating the original image contour information.



264

265

266

267

268

269

270

271

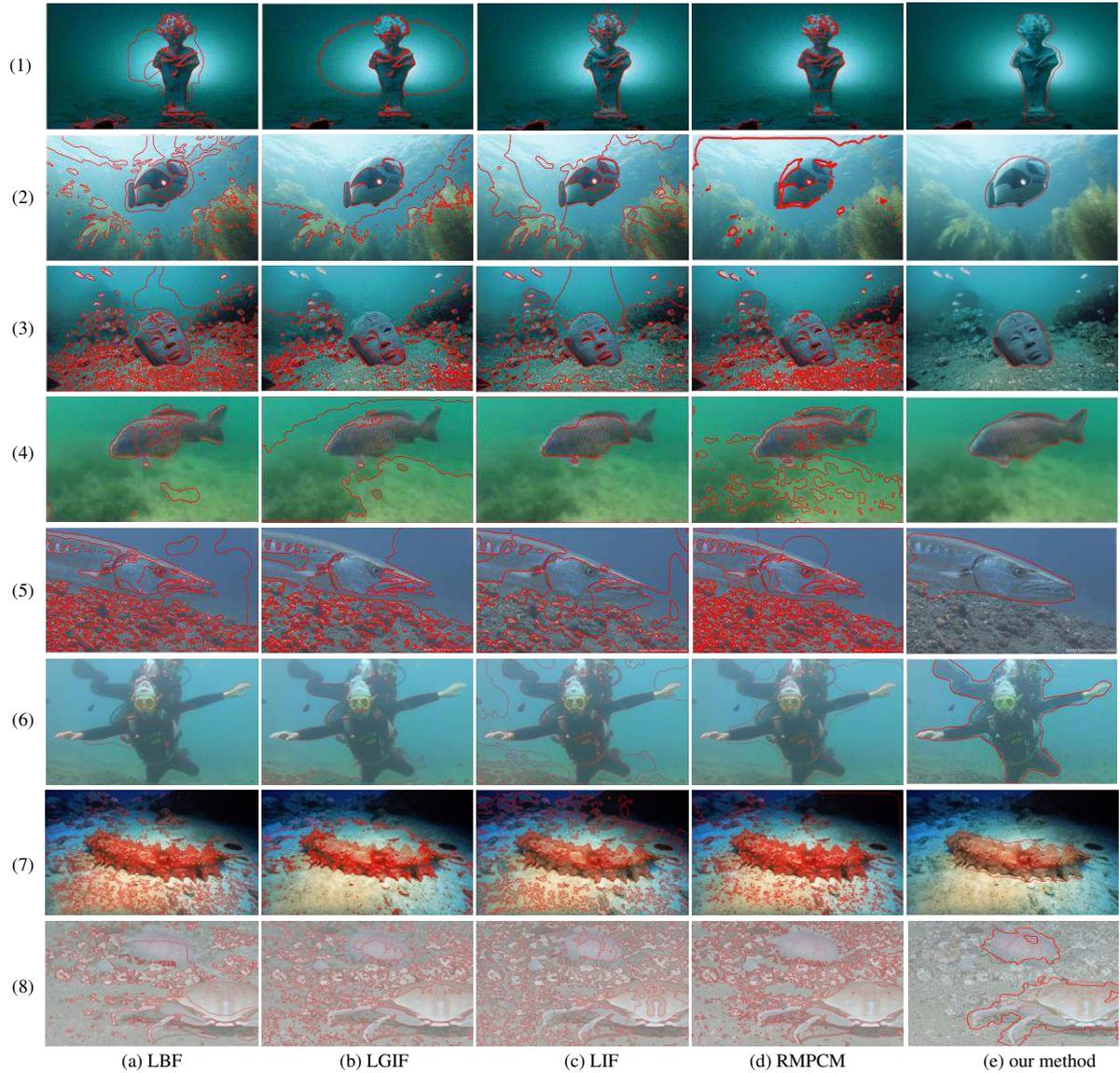
272

Fig. 4 The results of our method. (a) original underwater images with initial zero level contour. (b) final level set function. (c) results of our method.

#### 4.4 Qualitative comparison

##### 4.4.1 Compare the segmentation results with other models

To verify the effectiveness of the proposed method, we compared the segmentation results with other classic models such as LBF<sup>[24-25]</sup>, LGIF<sup>[26]</sup>, LIF<sup>[22]</sup>, and RMPCM<sup>[3]</sup>, respectively. The comparison results are shown in Fig.5.



273

274

275

276

277

278

279

280

281

282

283

284

285

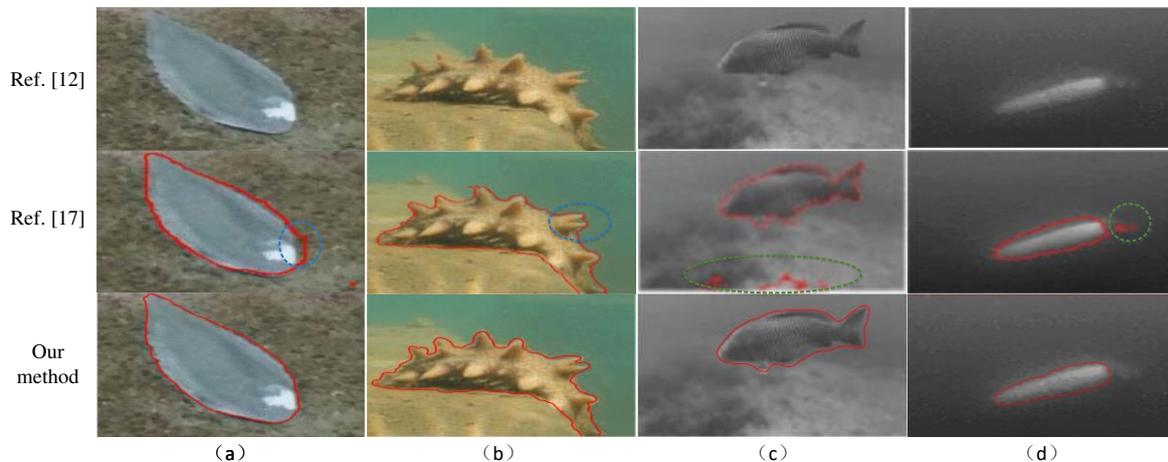
286

Fig. 5. Comparison of our method with LBF, LGIF, LIF, RMPCM. (a) results of the LBF model. (b) results of the LGIF model. (c) results of the LIF model. (d) results of the RMPCM model. (e) results of our method.

It can be seen from Fig.5 that the LBF model is limited by the initial contour curve and cannot completely extract the target contour. The LGIF model is minimally affected by local background noise due to the fusion of global intensity fitting, but it still cannot accurately extract the target contour. LIF and RMPCM models can extract target contour relatively completely, but they are greatly affected by background noise and target local features. Our model introduces semantic information, so it can filter out background noise very well. And because of the global-local intensity fitting, our method can handle local inhomogeneous regions without being interfered by the local target features. In addition, the target contour of original image perfectly complements the missing semantic information.

Also, we compared the segmentation results with the methods in Ref. [12] and Ref. [17], which also introduced saliency images as semantic information. Since we cannot get the source codes of Ref. [12]

287 and Ref. [17], to ensure the fairness of the comparison results, we use the segmentation results in Ref.  
 288 [12] and Ref. [17] as the comparison images. The comparison results are shown in Fig. 6.

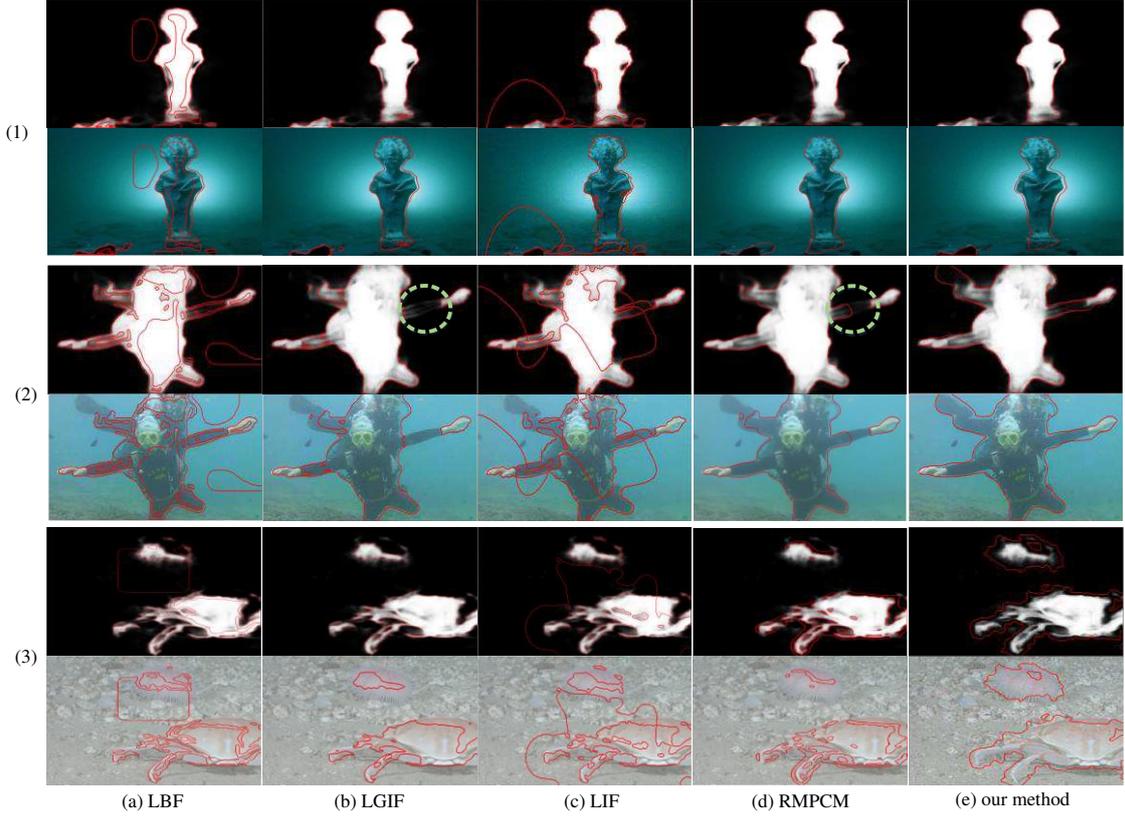


289 (a) (b) (c) (d)  
 290 Fig. 6. Comparison of our method with the methods of Ref. [12] and Ref. [17]. The first row are  
 291 the original underwater images, the second row are the segmentation results in Ref. [12] and Ref. [17].  
 292 The third row are the results of our method.

293 As can be seen in Figure 6, even though our method, Ref. [12] and Ref. [17] all introduce semantic  
 294 information, our method can extract the target contour more accurately than Ref. [12] and Ref. [17]. As  
 295 shown in the blue circle region of Fig. 6(a) and Fig. 6(b), our method extracted the target contour in the  
 296 detail region more accurately. This is because we have added the local-global fitting term to better extract  
 297 the contours of local inhomogeneous regions, and the original image correction module can correct errors  
 298 in semantic information. As shown in the green circle region of Fig. 6(c) and Fig. 6(d), our method can  
 299 filter out background noise better than Ref. [17] and is more robust.

#### 300 4.4.2 Compare the saliency segmentation results with other models

301 To further verify the superiority of the proposed method, we also compared the contour extraction  
 302 results of the underwater image with the saliency image as the input of several classic models. In order  
 303 to test the robustness of the proposed method, we only selected low-quality saliency images  
 304 (inhomogeneous local intensity and incomplete saliency information) for comparison experiments. As  
 305 shown in Fig. 7, the segmentation results of LBF are severely affected by the initial contour curve and  
 306 are disturbed by the inhomogeneous regions inside the target. The LGIF model can avoid the influence  
 307 of the initial contour curve, but cannot extract complete contour information, as shown in the green dotted  
 308 region in Fig. 7(2). The LIF model can extract the target contour relatively completely, but it is easy to  
 309 fall into the local optimum and is also affected by the initial contour curve. The RMPAM model avoids  
 310 the local optimum error, but it also has the problem that the contour information cannot be extracted  
 311 completely, as shown in the green dotted region in Fig. 7(2). Our method can not only effectively avoid  
 312 local optimum, but also supplement the missing contour information through the original image, so the  
 313 results of our method are more accurate and complete than other methods.



314

315

316

317

318

#### 319 4.5 Quantitative comparison

320

321

322

323

Fig. 7 Comparison of our method with LBF, LGIF, LIF, RMPCM. (a), (b), (c), (d) and (e) are results of LBF, LGIF, LIF, RMPCM, and our method. The upper rows of (1), (2), (3) are the segmentation results of the saliency images, and the lower rows are the segmentation results of the corresponding original images.

324

$$MAE = \frac{1}{m \times n} \sum_{x=1}^m \sum_{y=1}^n |Det_{(x,y)} - gt_{(x,y)}| \quad (23)$$

325

$$ER = \frac{1}{m \times n} \sum_{x=1}^m \sum_{y=1}^n |Det_{(x,y)} - gt_{(x,y)}| \bigg/ \frac{1}{m \times n} \sum_{x=1}^m \sum_{y=1}^m Det_{(x,y)} * gt_{(x,y)} \quad (24)$$

326

$$DR = \frac{1}{m \times n} \sum_{x=1}^m \sum_{y=1}^m Det_{(x,y)} * gt_{(x,y)} \bigg/ \frac{1}{m \times n} \sum_{x=1}^m \sum_{y=1}^m (Det_{(x,y)} + gt_{(x,y)}) \quad (25)$$

327

328

329

330

331

where  $m$  and  $n$  represent the length and width of the image,  $Det$  is the result of image segmentation,  $gt$  is the hand-crafted ground truth. So  $Det_{(x,y)} * gt_{(x,y)}$  represents the contour that are accurately extracted by the model. The larger the  $Det_{(x,y)} * gt_{(x,y)}$ , the more contour that are correctly extracted.  $Det_{(x,y)} - gt_{(x,y)}$  represents the pixels that are incorrectly extracted, so the larger the  $Det_{(x,y)} - gt_{(x,y)}$  is, the more pixels are incorrectly extracted. So the smaller the value of  $ER$ , the more accurate the result of contour extraction.

332 And a large value of  $DR$  can indicate that the contour extraction result of the model is accurate. The  
 333 evaluation results of the aforementioned five methods are shown in Table. 3, Table. 4 and Table. 5.

334 Table 3 The MAE results of LBF, LGIF, LIF, RMPCM and our method

	Fig.5(1)	Fig.5(2)	Fig.5(3)	Fig.5(4)	Fig.5(5)	Fig.5(6)	Fig.5(7)	Fig.5(8)
LBF	9.5723	3.6588	3.0737	12.8131	2.1789	4.4882	6.9886	5.3132
LGIF	7.9119	3.7481	3.4620	9.6148	2.2546	4.1299	7.4210	6.3514
LIF	6.0811	4.0945	2.5343	7.4206	2.9782	4.2036	10.2057	4.9874
RMPCM	10.5081	4.3048	5.1594	12.9181	2.3604	6.0406	7.9683	3.9875
Our method	2.3695	3.2604	1.9161	4.5302	1.2455	2.7715	5.9417	2.9702

335 Table 4 The ER results of LBF, LGIF, LIF, RMPCM and our method

	Fig.5(1)	Fig.5(2)	Fig.5(3)	Fig.5(4)	Fig.5(5)	Fig.5(6)	Fig.5(7)	Fig.5(8)
LBF	0.7434	0.5195	0.0510	0.7863	0.2585	0.2769	0.3403	0.2386
LGIF	1.0012	0.5978	0.1414	0.7304	0.1237	0.2283	0.2092	0.3109
LIF	0.8355	0.4713	0.0761	0.3578	0.1209	0.2938	0.5013	0.1483
RMPCM	1.1478	0.3300	0.0962	1.0167	0.1167	0.2137	0.3200	0.9286
Our method	0.2709	0.2649	0.0452	0.3411	0.0776	0.2007	0.2060	0.0571

336 Table 5 The DR results of LBF, LGIF, LIF, RMPCM and our method

	Fig.5(1)	Fig.5(2)	Fig.5(3)	Fig.5(4)	Fig.5(5)	Fig.5(6)	Fig.5(7)	Fig.5(8)
LBF	1.3148	1.7739	13.9989	1.2499	3.5065	3.4029	2.7005	3.9709
LGIF	0.9778	1.5494	6.0806	1.3423	6.7686	4.0834	4.1945	3.0709
LIF	1.1707	1.9462	10.1875	2.6965	6.9244	3.2240	1.8872	6.1722
RMPCM	0.8569	2.7046	8.6031	0.9690	7.0274	4.2727	2.8823	1.0500
Our method	3.4725	3.3199	15.3967	2.7586	10.3096	4.6516	4.3053	14.0682

337 A smaller value of MAE represents a higher contour extraction accuracy. According to Table. 3, the  
 338 contour extracted by the proposed model obtain the smallest MAE value, which shows that the proposed  
 339 model can extract target contours more accurately than the other four models. Table 4 shows the error  
 340 rate(ER) of five methods. The values of ER between the target contour extracted by the proposed method  
 341 and ground truth are the smallest, so the proposed method has the highest accuracy. Table 5 shows the  
 342 detection rates of the above five methods. The detection rate represents how many contour pixels are  
 343 correctly extracted. Therefore, our model with the highest detection rate can extract the target contour  
 344 more accurately.

## 345 5 Discussion

346 In this paper, the parameter  $\omega_1$  is a constant, which controls the influence of the saliency image  
 347 fitting energy and original image fitting energy. When the missing information of saliency target contour  
 348 is severe,  $\omega_1$  should be relatively larger; otherwise  $\omega_1$  should be taken to a small value. Also,  $\omega$  should  
 349 be should be taken smaller when the intensity inhomogeneity of saliency image is severe. This is because  
 350 that the local intensity fitting can better segment target in intensity inhomogeneity region and the results  
 351 of contour extraction relies on the local intensity fitting. Otherwise,  $\omega$  should be taken larger to  
 352 suppress the noise interference. In the experiment, we need to choose appropriate values for  $\omega$  and  $\omega_1$   
 353 according to the degree of inhomogeneity and the degree of saliency detection deviation. In the  
 354 experiment of this paper, the value of  $\omega$  can be taken from 0.5 to 0.9, and the value of  $\omega_1$  can be taken  
 355 from 0.1 to 0.8.

## 356 6 Conclusions

357 Aiming at the problem of saliency target contour extraction of underwater images, we propose a  
358 dual-fusion active contour model with semantic information. The proposed method extracted the saliency  
359 target contour by fusing local intensity and global intensity, and extracted the original image contour  
360 information by the local image fitting model to correct the saliency information deviation. We verified  
361 the effectiveness of the dual-fusion active contour model with semantic information by comparative  
362 experiments. The experimental results show that the local-global intensity fitting term can effectively  
363 suppress the interference of noise, and can more accurately extract the contour of the intensity  
364 inhomogeneity region. We also verified that the missing saliency target contour can be effectively  
365 corrected by the contour information of original image. The results of qualitative analysis and quantitative  
366 analysis show that our method can effectively filter out the background information and extract the  
367 saliency target contour more accurately.

## 368 References

- 369 [1] Xiao Z L, Zhang M, Chen L S & Jin H Y. Detection and segmentation of underwater CW-like signals  
370 in spectrum image under strong noise background. *J. Journal of Visual Communication and Image  
371 Representation*, 60, 287-294(2019).
- 372 [2] Hou G J. et al. A novel dark channel prior guided variational framework for underwater image restoration.  
373 *J. Journal of Visual Communication and Image Representation*, 66, 102732(2020).
- 374 [3] Wu Y F, Li M, Zhang Q F & Liu Y. A retinex modulated piecewise constant variational model for image  
375 segmentation and bias correction. *J. Applied Mathematical Modelling*, 54, 697-709(2018).
- 376 [4] Y. Liu, H. Li. Design of Refined Segmentation Model for Underwater Images. *C. 2020 5th International  
377 Conference on Communication, Image and Signal Processing (CCISP)*, 282-287(2020), doi:  
378 10.1109/CCISP51026.2020.9273503.
- 379 [5] Chen. W, He. C Y, Ji. C L, Zhang M Y & Chen S Y. An improved K-means algorithm for underwater  
380 image background segmentation. *J. Multimedia Tools and Applications*, 80, 21059–21083(2021).  
381 <https://doi.org/10.1007/s11042-021-10693-7>.
- 382 [6] SM A R A J, Jose C & MH S. Hardware realization of canny edge detection algorithm for underwater  
383 image segmentation using field programmable gate arrays. *J. Journal of Engineering Science and  
384 Technology*, 12(9), 2536-2550(2017).
- 385 [7] Sun Y T & Luan X L. An Underwater Optical Image Segmentation Algorithm Based on Fuzzy C-means  
386 Model. *J. Journal of Physics: Conference Series. IOP Publishing*, 1087(5): 052007(2018).
- 387 [8] Rajeev A A, Hiranwal S & Sharma V K. Improved Segmentation Technique for Underwater Images  
388 Based on K-means and Local Adaptive Thresholding. *C. Information and Communication Technology  
389 for Sustainable Development*. Springer, Singapore, 443-450(2018).
- 390 [9] Li X, Song J D, Fan Z, Ouyang X G & Samee U.Khan. Map Reduce-based fast fuzzy c-means  
391 algorithm for large-scale underwater image segmentation. *J. Future Generation Computer Systems*,

- 392 65, 90-101(2016).
- 393 [10] Chen Z. et al. Underwater object segmentation based on optical features. *J. Sensors*, 18(1), 196(2018).
- 394 [11] Xuan L & Zhang M J. Underwater color image segmentation method via RGB channel fusion. *J. Optical*  
395 *Engineering*, 56(2). 023101(2017).
- 396 [12] Zhu Y. et al. Underwater image segmentation with co-saliency detection and local statistical active  
397 contour model. *C. OCEANS 2017-Aberdeen. IEEE*, 1-5(2017).
- 398 [13] Qiao X, Bao J H, Zeng L H, Zhou J & Li D L. An automatic active contour method for sea cucumber  
399 segmentation in natural underwater environments. *J. Computers and Electronics in Agriculture*, 135,  
400 134-142(2017).
- 401 [14] Li Y J, Xu H L, Li Y, Lu H M & Seiichi Serikawa. Underwater image segmentation based on fast  
402 level set method. *C. International Journal of Computational Science and Engineering*, 19(4): 562-  
403 569(2019).
- 404 [15] Bai J S, Pang Y J, Zhang Q & Zhang Y H. Underwater Image Segmentation Methods Based on MCA  
405 and Adaptive Level Set Evolution. *C. 2016 3rd International Conference on Information Science and*  
406 *Control Engineering (ICISCE). IEEE*, 734-738(2016).
- 407 [16] Li S L & Mengxing H. Research of Underwater Image Segmentation Algorithm Based on the Improved  
408 Geometric Active Contour Models. *C. 2018 International Conference on Intelligent Autonomous Systems*  
409 *(ICoIAS). IEEE*, 44-50(2018).
- 410 [17] Chen Z. et al. Underwater object segmentation integrating transmission and saliency features. *J. IEEE*  
411 *Access*, 7, 72420 – 72430(2019).
- 412 [18] O’Byrne M, Pakrashi V, Schoefs F & Bidisha Ghosh. Semantic Segmentation of Underwater  
413 Imagery Using Deep Networks Trained on Synthetic Imagery. *J. Journal of Marine Science and*  
414 *Engineering*, 6(3), 93(2018).
- 415 [19] Zhou Y. et al. Underwater scene segmentation by deep neural network. *J. UK Robotics and Autonomous*  
416 *Systems*, 2019.
- 417 [20] T. Chan & L. Vese. Active contours without edges. *J. IEEE Trans. Image Process*, 10 (2), 266–  
418 277(2001).
- 419 [21] D. Mumford & J. Shah. Optimal approximations by piecewise smooth functions and associated  
420 variational problems. *J. Commun. Pure Appl. Math*, 42 (5), 577–685(1989).
- 421 [22] Zhang K H, Song H H & Zhang L. Active contours driven by local image fitting energy. *J. Pattern*  
422 *Recognition*, 43(4), 1199–1206(2010).
- 423 [23] Zhao T & Wu X. Pyramid Feature Attention Network for Saliency Detection. *C. Proceedings of the*  
424 *IEEE Conference on Computer Vision and Pattern Recognition*, 3085-3094(2019).
- 425 [24] Li C M, Kao C Y, Gore J C & Ding Z H. Implicit active contour driven by local binary fitting energy.  
426 *C. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1, 1–7(2007).
- 427 [25] Li C M, Kao C Y, Gore J C & Ding Z H. Minimization of region-scalable fitting energy for image  
428 segmentation. *J. IEEE Trans. Image Process*. 17(10), 1940–1949(2008).
- 429 [26] Wang L, Li C M, Sun Q S, Xia D S & Kao C Y. Active contours driven by local and global intensity

430 fitting energy with application to brain MR image segmentation. *J. Computerized medical imaging*  
431 *and graphics*, 33(7), 520-531(2009).

432 [27] Li C M, Xu C Y, Gui C F & Martin D. Fox. Distance regularized level set evolution and its  
433 application to image segmentation. *J. IEEE transactions on image processing*, 19(12), 3243-  
434 3254(2010).

435 [28] Li C M, Xu C Y, Gui C F & Martin D. Fox. Level set evolution without re-initialization: A new variational  
436 formulation. *C. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1, 430-  
437 436(2005).

438 **Acknowledgment:** The authors thanks the National Key R&D Program of China(2018YFC0810500) for  
439 help identifying collaborators for this work.

440 **Author contributions:** Shudi Yang contributed the innovative content and experimental verification of  
441 the method in this paper. Zhehan Chen completed the preliminary review and revision of the manuscript  
442 and he is the corresponding author. Jiaxiong Wu completed the writing and language revision of the  
443 manuscript. Zhipeng Feng is responsible for the overall coordination of the manuscript writing process.

444 **Competing interests statement:** The author(s) declare no competing interests.

445 **Figure legends**

446 Figure-1 (a)-(d) shows the segmentation results of the LBF, LGIF, LIF, and RMPCM model,  
447 respectively.

448 Figure-2 The contour extraction results. (a) result of the local intensity fitting. (b) result of the global  
449 intensity fitting. (a) result of our method.

450 Figure-3 Results of our method. (a) original underwater image with initial zero level contour. (b)  
451 contour extraction result of saliency target. (c) contour extraction result without correction. (d) final level  
452 set function. (e) result of our method.

453 Figure-4 The contour extraction results of our method. (a) original underwater images with initial zero  
454 level contour. (b) saliency images of the original underwater images (c)final level set function. (d) results  
455 of our method.

456 Figure-5 Comparison of our method with LBF, LGIF, LIF, RMPCM. (a) results of the LBF model. (b)  
457 results of the LGIF model. (c) results of the LIF model. (d) results of the RMPCM model. (e) results of  
458 our method.

459 Figure-6 Comparison of our method with the methods of Ref. [12] and Ref. [17]. The first row are the  
460 original underwater images, the second row are the segmentation results in Ref. [12]((a) and (b)) and Ref.  
461 [17]((c) and (d)). The third row are the results of our method.

462 Figure-7 Comparison of our method with LBF, LGIF, LIF, RMPCM. (a), (b), (c), (d) and (e) are results  
463 of LBF, LGIF, LIF, RMPCM, and our method. The upper rows of (1), (2), (3) are the segmentation results  
464 of the saliency images, and the lower rows are the segmentation results of the corresponding original  
465 images.

466 **Tables**

467 **Table 1 Iterations and CPU time (in seconds)**

	LBF	LGIF	LIF	RMPCM
1239 × 731pixels				
Iterations	200	200	200	200
Time (s)	93.2969	55.5938	38.5469	63.3052

468 **Table 2 The comparative experiment of local-global intensity**

Experiments	local intensity	global intensity
A	✓	
B		✓
C(our fusion intensity)	✓	✓

469 **Table 3 The MAE results of LBF, LGIF, LIF, RMPCM and our method**

	Fig.5(1)	Fig.5(2)	Fig.5(3)	Fig.5(4)	Fig.5(5)	Fig.5(6)	Fig.5(7)	Fig.5(8)
LBF	9.5723	3.6588	3.0737	12.8131	2.1789	4.4882	6.9886	5.3132
LGIF	7.9119	3.7481	3.4620	9.6148	2.2546	4.1299	7.4210	6.3514
LIF	6.0811	4.0945	2.5343	7.4206	2.9782	4.2036	10.2057	4.9874
RMPCM	10.5081	4.3048	5.1594	12.9181	2.3604	6.0406	7.9683	3.9875
Our method	2.3695	3.2604	1.9161	4.5302	1.2455	2.7715	5.9417	2.9702

470 **Table 4 The ER results of LBF, LGIF, LIF, RMPCM and our method**

	Fig.5(1)	Fig.5(2)	Fig.5(3)	Fig.5(4)	Fig.5(5)	Fig.5(6)	Fig.5(7)	Fig.5(8)
LBF	0.7434	0.5195	0.0510	0.7863	0.2585	0.2769	0.3403	0.2386
LGIF	1.0012	0.5978	0.1414	0.7304	0.1237	0.2283	0.2092	0.3109
LIF	0.8355	0.4713	0.0761	0.3578	0.1209	0.2938	0.5013	0.1483
RMPCM	1.1478	0.3300	0.0962	1.0167	0.1167	0.2137	0.3200	0.9286
Our method	0.2709	0.2649	0.0452	0.3411	0.0776	0.2007	0.2060	0.0571

471 **Table 5 The DR results of LBF, LGIF, LIF, RMPCM and our method**

	Fig.5(1)	Fig.5(2)	Fig.5(3)	Fig.5(4)	Fig.5(5)	Fig.5(6)	Fig.5(7)	Fig.5(8)
LBF	1.3148	1.7739	13.9989	1.2499	3.5065	3.4029	2.7005	3.9709
LGIF	0.9778	1.5494	6.0806	1.3423	6.7686	4.0834	4.1945	3.0709
LIF	1.1707	1.9462	10.1875	2.6965	6.9244	3.2240	1.8872	6.1722
RMPCM	0.8569	2.7046	8.6031	0.9690	7.0274	4.2727	2.8823	1.0500
Our method	3.4725	3.3199	15.3967	2.7586	10.3096	4.6516	4.3053	14.0682

472

473

474