# Improved recognition of *Echinococcus granulosus* protoscoleces using visual saliency and scale-invariant features

**Zhuang Li (李 壮)[1], Guodong Lü (吕国栋)[2], and Xiaoyi Lü (吕小毅)[3,*]**

[1]*School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China*
[2]*Xinjiang Key Laboratory of Echinococcosis, Clinical Medical Research Institute, The First Affiliated Hospital of Xinjiang Medical University, Urumqi 830054, China*
[3]*College of Information Science and Engineering, Xinjiang University, Urumqi 830046, China*
**Corresponding author: xiaoz813@163.com*

Received November 11, 2018; accepted January 10, 2019; posted online March 29, 2019

Echinococcosis—a parasitic disease caused by *Echinococcus granulosus* or *Echinococcus multilocularis* larvae—occurs in many regions in the world. This disease can pose a serious threat to public health and thus requires a convenient and cost-effective method for early detection. So, we developed a novel method based on visual saliency and scale-invariant features that detects the tapeworm parasites. This method improves upon existing bottom-up computational saliency models by introducing a visual attention mechanism. The results indicated that the proposed method offers a higher level of both accuracy and computational efficiency when detecting *Echinococcus granulosus* protoscoleces, which in turn could improve early detection of echinococcosis.

*OCIS codes: 170.3880, 100.4995.*
*doi: 10.3788/COL201917.041703.*

Echinococcosis is a serious parasitic disease caused by *Echinococcus granulosus* or *Echinococcus multilocularis* larvae that affects the health of people and animals. This disease is especially common in pastoral areas[1]. China, including the Xinjiang Uygur Autonomous Region, has a high incidence of echinococcosis, and select areas exhibit a persistent high prevalence trend[2,3]. Effective prevention and control of echinococcosis requires an expedited confirmation of drug efficacy; however, existing drug evaluations and use rationales require further improvement. Accurate *in vitro* recognition of the tapeworm's larval protoscoleces or hydatids and assessment of their survival status (living or nonliving) is important while evaluating the efficacy of new drugs.

Eosin exclusion is a common method for detecting the viability of protoscoleces *in vitro*. This method is simple, economical, and rapid to apply and has been widely used in the development of new drugs to treat hydatid diseases of the liver[4–6]. Eosin exclusion uses an eosin staining solution or dyestuff to distinguish living and nonliving cells in a membrane. The nonliving cells are readily permeated and stained by the dyestuff, while the living cells resist dye penetration. By using a microscope to examine dynamic and still images of the membrane following this procedure, the survival status of the larvae can be strictly monitored through human observation.

Despite its advantages, eosin exclusion can be time consuming and result in larvae survival rate errors attributable to the manual monitoring process by human observers. The development of intelligent recognition capabilities could increase the accuracy of the overall eosin exclusion method. However, no such studies combining intelligent recognition capabilities and eosin exclusion were reported in the literature.

Although we found no direct application to eosin exclusion, a number of studies have applied intelligent recognition algorithms to conventional egg recognition efforts. For example, Rema *et al.*[7] used an active contour model to effectively realize segmentation of a microscopic image of human parasite eggs against a complex background. Alternatively, Chen *et al.* proposed a hybrid segmentation algorithm for parasite egg images based on morphological filtering[8]. By improving the morphological filtering and combining the convex packet operation, the image of parasite eggs was effectively segmented. More recently, Lin *et al.* used scale-invariant feature transform (SIFT) descriptor to examine the total extrinsic extracellular matrix[9]. Most recently, Zhang *et al.* used an improved $k$-nearest-neighbor classifier to identify and classify parasitic eggs in microscopic images of fecal matter[10].

The previous research has collectively advanced the state of knowledge regarding egg recognition technology in micromedicine. However, each of the proposed methods indicates only the need to intervene without explicitly reflecting the range of image characteristics or their potential overlap.

In response to these methodological shortcomings, we developed a novel method based on computer vision and machine learning that detects protoscoleces parasites using eosin exclusion. Using this method, areas with suspected live *Echinococcus granulosus* protoscoleces are first identified and outlined. Next, using sample images of the suspected parasites, the SIFT algorithm is used to

 

extract common features from the images, and the $k$-means algorithm is used to cluster them. Finally, support vector machine (SVM) learning is applied, and the previously identified suspicious areas are subsequently classified based on parasite counts. The efficacy of this proposed method was validated with experimental methods.

The results of this study may improve the accuracy of eosin exclusion when detecting *Echinococcus granulosus* protoscoleces, which in turn could improve live detection of echinococcosis, contribute to the development of related drugs for echinococcosis treatment and prevention, and ultimately improve public health. In addition, the results of this study may contribute to the state of knowledge regarding target segmentation and recognition in the field of image processing[11].

The human visual system possesses image understanding, recognition, and processing capabilities. In the image processing field, efforts are focused on simulating the human visual system using computers and establishing a visual attention model. Human visual observation is selective. Broadly speaking, images contain a variety of information that can be perceived by humans, such as color, texture, and brightness. However, not all information is of interest, and as such, not everything that we see is processed by the brain. To be effective, computer simulation of human visual observation must quickly locate salient regions and extract relevant images.

At present, four primary models of visual saliency exist that were developed by (1) Hou and Zhang, (2) Hu, Rajan, and Chia, (3) Stentiford, and (4) Itti, Koch, and Niebur. Among these four models, the Itti model was determined to be most appropriate for the egg recognition task in this study because it simulates human perception and extracts regions of interest based on differences between the target and background[12]. The method proposed in this study uses a modified Itti model.

Visual saliency can be graphically depicted using maps. When developing visual saliency maps, the Itti model extracts primary characteristics, resolves multiple features and multidimensional visual space using center-surround methods, filters and obtains feature maps using pyramids with a depth of up to nine levels, and compounds maps using fusing and computing methods. The resultant visual saliency maps include brightness, color, and directional characteristics.

Four broadly tuned color channels can be created as follows:

$$\begin{aligned} R &= r-(g+b)/2,\\ G &= g-(r+b)/2,\\ B &= b-(r+g)/2,\\ Y &= (r+g)/2-|r-g|/2-b, \end{aligned} \tag{1}$$

where $r$, $g$, and $b$ are the red, green, and blue channels of the image, $RG=|R-G|$, and $BY=|B-Y|$. Four corresponding Gaussian pyramids $[R(\delta), G(\delta), B(\delta), Y(\delta)]$ can be created from these color channels. An intensity image $(I)$ can be obtained as follows: $I=(r+g+b)/3$.

For the multiscale feature map, a central periphery differential operation is performed, and the feature graph is subtracted under different scales $(\sigma)$. With $c\in\{2,3,4\}$ and $s=c+\delta, \delta\in\{3,4\}$, where $c$ and $s$ represent pyramid levels, a brightness feature map can be generated as follows:

$$I(c,s)=|I(c)\Theta I(s)|. \tag{2}$$

Similarly, a color feature map can be generated as follows:

$$RG(c,s)=|[R(c)-G(c)]\Theta[G(s)-R(s)]|, \tag{3}$$

$$BY(c,s)=|[B(c)-Y(c)]\Theta[Y(s)-B(s)]|. \tag{4}$$

The $\Theta$ symbol indicates that the pixel difference can be obtained when the pyramid levels increase from $s$ to $c$.

Providing good directional selectivity, the Gabor filter is suitable for the extraction of directional features in an image. Local orientation information can be obtained from $I$ using oriented Gabor pyramids $[O(\sigma,\theta)]$, where $\sigma\in[0,8]$ represents the scale, and $\theta\in\{0^\circ,45^\circ,90^\circ,180^\circ\}$ is the preferred orientation. The Gabor filter can be formulated as follows:

$$\begin{aligned} g(x_0,y_0)=\exp\left(-\frac{x_0^2}{2\delta_{x_0}^2}-\frac{y_0^2}{2\delta_{y_0}^2}\right)&[\cos(2\pi f x_0)\\ &+j\sin(2\pi f x_0)], \end{aligned} \tag{5}$$

where

$$x_0=x\cos\theta+y\sin\theta. \tag{6}$$

A directional feature map can be subsequently generated as follows:

$$O(c,s,\theta)=|O(c,\theta)\Theta O(s,\theta)|. \tag{7}$$

After the eigenvalues of each feature map are measured using nonlinear normalization, the brightness, color, and directional feature maps must be added separately to form corresponding brightness, color, and directional saliency maps. Formulations for the brightness, color, and directional saliency maps, respectively, are as follows:

$$\bar{I}=\bigoplus_{c=2}^{4}\bigoplus_{s=c+3}^{c+4}N[I(c,s)], \tag{8}$$

$$\bar{C}=\bigoplus_{c=2}^{4}\bigoplus_{s=c+3}^{c+4}\{N[RG(c,s)]+N[BY(c,s)]\}, \tag{9}$$

$$\bar{O}=\sum_{\theta\in\{0^\circ,45^\circ,90^\circ,135^\circ\}}N\left\{\bigoplus_{c=2}^{4}\bigoplus_{s=c+3}^{c+4}N[O(c,s,\theta)]\right\}. \tag{10}$$

These three individual feature saliency maps are linearly weighted and combined to form a total visual saliency map[11] as follows:

$$S = [N(\bar{I}) + N(\bar{C}) + N(\bar{O})]/3. \tag{11}$$

In the total visual saliency map, each target competes for attention and focus. A winner-take-all mechanism is used to detect the point of highest salience in the map at any given time, and draws the focus of attention towards this location. Figure 1 presents an actual microscopic image of a parasite sample and the corresponding visual saliency map produced with the Itti model.

In this study, the Itti visual saliency model, which considers the brightness, color, and directional characteristics of an image, was originally selected to complement eosin exclusion methods, which enhance image color. However, the accuracy of the Itti model was low when extracting the image's saliency region, and it was difficult to extract the entire area of interest.

In an effort to improve accuracy, we modified the Itti model to reflect the human eye's different sensitivities to different saliency features. Considering the microscopic image of *Echinococcus granulosus* protoscoleces in Fig. 1, brightness and color characteristics were determined to be more important than directional characteristics. As such, directional characteristics were ignored, and only the brightness and color saliency maps were linearly weighted and combined to form a total visual saliency map intended to improve the accuracy of parasite recognition.

The method proposed in this study also includes the use of the SIFT algorithm to extract common features from the images. The features are highly distinctive and invariant to image scale and rotation[13]. The SIFT algorithm includes four primary steps: (1) detecting extreme points in scale space, (2) filtering and locating keypoints, (3) assigning an orientation to each keypoint, and (4) determining a keypoint descriptor.

To obtain stable and effective extreme points, we first build a scale pyramid using a scale-space kernel based on the Gaussian function as follows:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}. \tag{12}$$

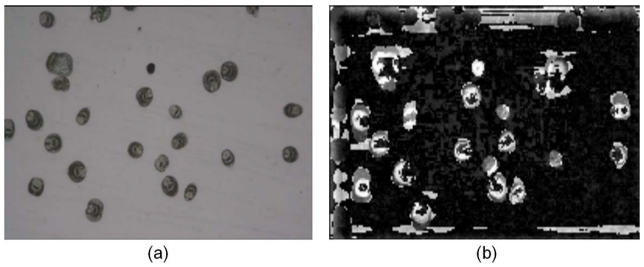


Fig. 1. (a) Microscopic image of *Echinococcus granulosus* protoscoleces and (b) the corresponding visual saliency map produced with the Itti model.

The scale space of an image can then be defined as a function $[L(x, y, \sigma)]$ that is produced from the convolution of a variable-space Gaussian $[G(x, y, \sigma)]$ with an input image $[I(x, y)]$. This relationship can be formulated as follows:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y), \tag{13}$$

where the $*$ symbol represents the convolution operation.

To efficiently detect stable keypoint locations in scale space, we use scale-space extrema in the difference-of-Gaussian function convoluted with the image $[D(x, y, \sigma)]$, which can be computed from the difference of two nearby scales separated by a constant multiplicative factor $(k)$ as follows:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma). \tag{14}$$

Figure 2 graphically depicts this relationship. This method has proven effective in extracting stable extreme points.

A large number of extreme points are likely detected in scale space during the first step of the SIFT algorithm. These points need to be further filtered and localized to ensure fully reliable feature points. During this second step, the position and scale of the keypoints are accurately determined by fitting three-dimensional functions. Unstable edge points and keypoints with low contrast are removed.

After the keypoints are filtered and localized, each remaining keypoint is assigned a location based on local image gradient directions. For each image sample $[L(x, y)]$ in scale space, the gradient magnitude $[m(x, y)]$ and orientation $[\theta(x, y)]$ are precomputed using pixel differences as follows:

$$m(x, y) = \sqrt{[L(x+1, y) - L(x-1, y)]^2 + [L(x, y+1) - L(x, y-1)]^2}, \tag{15}$$

$$\theta(x, y) = \arctan\{[(L(x+1, y) - L(x-1, y)]/[L(x, y+1) - L(x, y-1)]\}. \tag{16}$$

The first three steps of the SIFT algorithm produce a set of feature points, each described by unique locational, scalar, and directional characteristics. The final step is

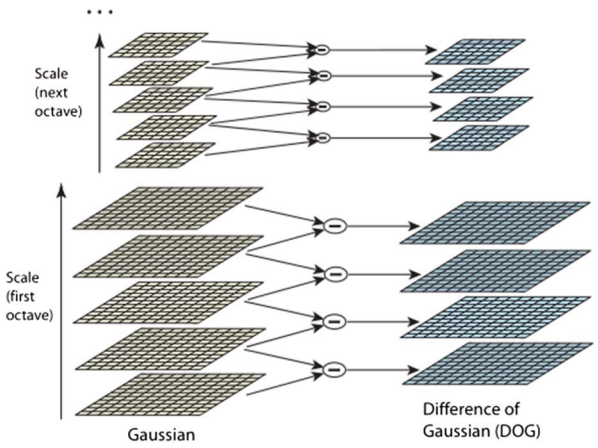


Fig. 2. Graphical depiction of the difference-of-Gaussian function.

to determine a keypoint descriptor for the local image region[13]. A keypoint descriptor is determined by computing the gradient magnitude and orientation at each image sample point in a region around the keypoint location. Figure 3 illustrates this process.

The method proposed in this study and applied to the recognition of *Echinococcus granulosus* protoscoleces relies upon the combination of the modified Itti model that provides visual saliency and the SIFT algorithm that extracts scale-invariant features to enhance the accuracy of parasite recognition when using eosin exclusion. Figure 4 provides a stepwise overview of this proposed method.

Using these combined methods, a clear egg image and the center points of any suspected living eggs can be obtained (all suspected living eggs are cut at their center points to produce sample slices). The SIFT algorithm can be used to extract the scale-invariant features of the known living eggs and produce a scale-invariant feature vector. The $k$-means clustering algorithm can next be used to cluster the vectors, generating an eigenvector with $k$ dimensions. This feature vector can be subsequently used to train an SVM and develop an SVM classifier. The SVM classifier can now be used to detect living (Yes) and nonliving (No) eggs based on the scale-invariant features provided. These classification results can be applied to the original image to determine the final egg recognition results.

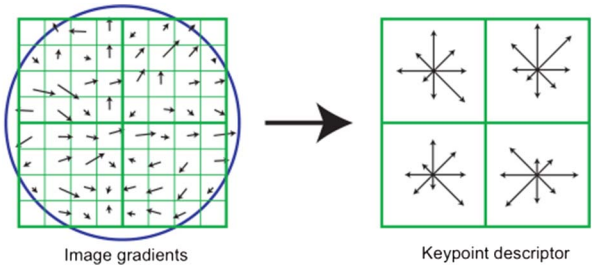


Fig. 3. Keypoint descriptor determination process.

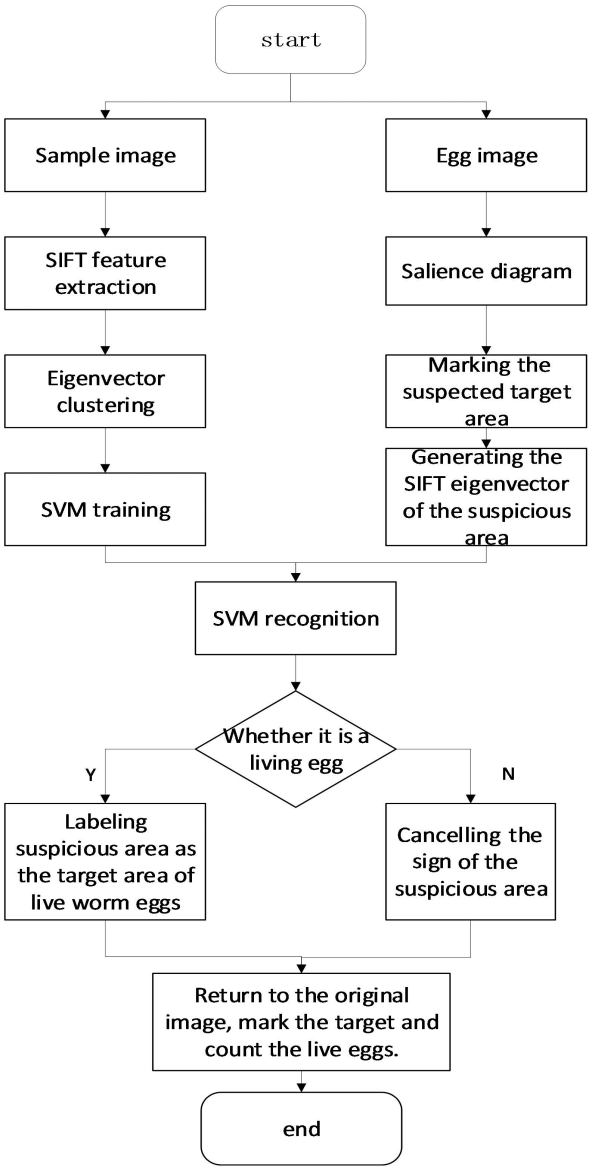


Fig. 4. Proposed method for detecting *Echinococcus granulosus* protoscoleces based on visual saliency and scale-invariant features.

To validate the efficacy of the method proposed in this study, we performed an experiment using MATLAB R2016a and microscopic images of parasites treated by the eosin exclusion method at the Xinjiang Medical University. The living parasite image samples included different targets and backgrounds. Subsets of 60 living tapeworm parasites and background images were used to develop an SVM classifier. The SVM classifier was subsequently used to detect living/nonliving parasites based on scale-invariant features.

As noted previously, the method proposed in this study and applied to the recognition of *Echinococcus granulosus* protoscoleces relies upon the combination of the modified Itti model that provides visual saliency and the SIFT algorithm that extracts scale-invariant features to enhance the accuracy of parasite recognition when using eosin

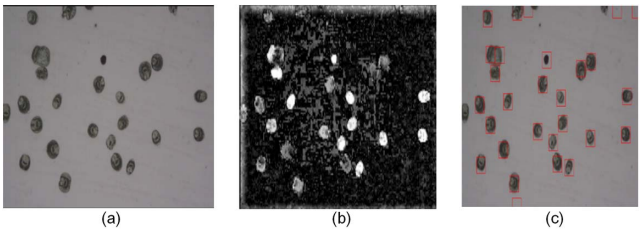


Fig. 5. (a) Microscopic image of *Echinococcus granulosus* protoscoleces, (b) the corresponding visual saliency map produced with the modified Itti model, and (c) the SVM classification results applied to the original image.

exclusion. Figure 5 shows intermediate experimental results from these combined methods, including the final SVM classification results applied to the original image.

Figure 6 provides a magnified view of Fig. 5(c). In the original image, 23 intact tapeworm parasites are visible, one of which is a nonliving tapeworm parasite. The method proposed in this study identified 21 living tapeworm parasites, resulting in a parasite recognition rate of 95.4%. This recognition rate is sufficiently high to meet hospital inspection requirements.

Additional comparative results further demonstrated the efficacy of the proposed method. First, we compared the parasite recognition rates for methods using the conventional and the proposed modified Itti models for visual saliency. Table 1 details the results of this comparison. The egg parasite recognition rates using conventional and proposed methods were 81.8% and 95.4%, respectively. The modified Itti model's visual saliency map, focusing on brightness and color features, better distinguishes parasite characteristics and can more accurately mark suspicious targets.

Next, we considered the transferability of the proposed method by applying it to three different eosin exclusion test images containing living parasites and comparing the resultant parasite recognition rates. Table 2 details the results of this comparison. In each case, the parasite recognition rates reached or exceeded 90% for parasite counts ranging from 22 to 50.

These collective results demonstrated that the proposed method, based on visual saliency and scale-invariant features, offers a higher level of accuracy when detecting *Echinococcus granulosus* protoscoleces compared with conventional methods. Accuracy levels achieved were sufficient to meet hospital clinical test requirements. In addition, the proposed method offers a higher level of computational efficiency, reducing both the workload and error potential attributable to manual counts performed by human observers.

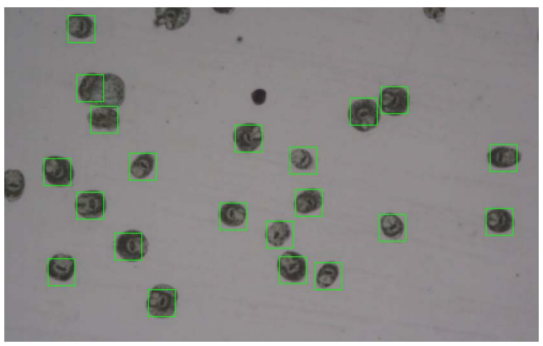

Fig. 6. Egg recognition results using the proposed method.

**Table 1.** Comparative Parasite Recognition Rates for Methods Using the Conventional Itti model and the proposed modified Itti model

| Visual Saliency Model | Number of Actual Living Parasites | Number of Correctly Identified Living Parasites | Parasite Recognition Rate |
|---|---|---|---|
| Conventional Itti model | 22 | 18 | 81.8% |
| Modified Itti model | 22 | 21 | 95.4% |

**Table 2.** Comparative Parasite Recognition Rates Using the Proposed Method for Three Different Eosin Exclusion Test Images Containing Living Parasites

| Test Image | Number of Actual Living Parasites | Number of Correctly Identified Living Parasites | Egg Recognition Rate |
|---|---|---|---|
| a | 22 | 21 | 95.4% |
| b | 30 | 27 | 90.0% |
| c | 50 | 46 | 92.0% |

In response to the need for a convenient and cost-effective method for early detection of echinococcosis, we developed a novel method based on computer vision and machine learning that detects *Echinococcus granulosus* protoscoleces using eosin exclusion. This method improves upon existing bottom-up computational saliency models by introducing a visual attention mechanism. Using this method, areas with suspected living tapeworm parasites are first identified and outlined. Next, using sample images of the suspected parasites, the SIFT algorithm extracts common features from the images, and the $k$-means algorithm is used to cluster them. Finally, SVM learning is applied, and the previously identified suspicious areas are subsequently classified.

Most notably, this proposed parasite recognition method limits analysis to suspected living parasite areas determined through visual saliency, which in turn reduces feature extraction time using the SIFT algorithm. Use of the $k$-means clustering algorithm to convert the high-dimensional feature descriptor of the SIFT eigenvector

(which is too complex and time-consuming for analysis) improves the computing efficiency of the target search and increases the stability of the features in the analysis region.

The efficacy of this proposed method was validated with experimental methods. Experimental results indicated that the proposed method, based on visual saliency and scale-invariant, features, offers a higher level of accuracy (sufficiently high to meet hospital clinical test requirements) when detecting *Echinococcus granulosus* protoscoleces, which in turn could improve live detection of echinococcosis and ultimately public health. In addition, the proposed method offers a higher level of computational efficiency, reducing both the workload and error potential attributable to manual counts performed by human observers. Sometimes parasites can produce sterile cysts (without protoscoleces), which is a limitation of this method. In future experiments, we will take this into account and carry out further research.

Experiments are in progress to optimize the algorithm and improve the recognition efficiency.

This work is supported by the Urumqi Science and Technology Project (Nos. P161310002 and Y161010025), the Reserve Talents Project of the National Highlevel Personnel of Special Support Program (No. QN2016YX0324), and the Reserve National Youth Talent Support Program (No. Xinjiang [2014]22).

## References

1. P. Li, Z. Jia, and G. Lü, Sci. Rep. **7**, 44798 (2017).
2. H. Li, T. Song, Y. Shao, W. Zhang, and H. Wen, Chin. J. Epidemiol. **36**, 1002 (2015).
3. M. Moazeni and E. Rakhshandehroo, Parasitol. Res. **110**, 925 (2012).
4. G. Xin, B. Wang, Y. Lei, C. Liu, Z. Wang, H. Shi, R. Yang, W. Qin, Y. Jiang, and H. Lv, Mol. Biochem. Parasitol. **207**, 49 (2016).
5. M. T. Rahimi, E. Ahmadpour, B. R. Esboei, A. Spotin, M. H. K. Koshki, A. Alizadeh, S. Honary, H. Barabadi, and M. A. Mohammadi, Int. J. Surg. **19**, 128 (2015).
6. H. Shi, H. Lv, Y. Lei, W. Qin, B. Wang, Z. Wang, Z. Xing, R. Yang, and Y. Jiang, J. Pathog. Biol. **11**, 220 (2016).
7. M. Rema and M. S. Nair, Biomed. Eng. Lett. **14**, 3179 (2013).
8. B. Chen, J. Xie, and B. Wang, Chin. J. Digital Med. **5**, 35 (2010).
9. D. Lin, T. Luo, L. Liu, Y. Lu, S. Liu, Z. Yuan, and J. Qu, Chin. Opt. Lett. **15**, 090006 (2017).
10. L. Liu, H. T. Lei, J. Zhang, Y. Yuan, Z. Zhang, J. Liu, Y. Xie, G. Ni, and Y. Liu, J. Med. Syst. **39**, 1 (2015).
11. M. Nouri, M. Mivehchy, and M. F. Sabahi, Chin. Opt. Lett. **15**, 100302 (2017).
12. L. Itti, C. Koch, and E. Niebur, IEEE Trans. Pattern Anal. Mach. Intell. **20**, 1254 (1998).
13. D. Lowe, Int. J. Comput. Vis. **60**, 91 (2004).