

Small Blob Identification in Medical Images Using Regional Features From Optimum Scale

Min Zhang, Teresa Wu*, and Kevin M. Bennett

Abstract—Recent advances in medical imaging technology have greatly enhanced imaging-based diagnosis which requires computational effective and accurate algorithms to process the images (e.g., measure the objects) for quantitative assessment. In this research, we are interested in one type of imaging objects: small blobs. Examples of small blob objects are cells in histopathology images, glomeruli in MR images, etc. This problem is particularly challenging because the small blobs often have inhomogeneous intensity distribution and an indistinct boundary against the background. Yet, in general, these blobs have similar sizes. Motivated by this finding, we propose a novel detector termed Hessian-based Laplacian of Gaussian (HLoG) using scale space theory as the foundation. Like most imaging detectors, an image is first smoothed via LoG. Hessian analysis is then launched to identify the single optimal scale on which a presegmentation is conducted. The advantage of the Hessian process is that it is capable of delineating the blobs. As a result, regional features can be retrieved. These features enable an unsupervised clustering algorithm for postpruning which should be more robust and sensitive than the traditional threshold-based postpruning commonly used in most imaging detectors. To test the performance of the proposed HLoG, two sets of 2-D grey medical images are studied. HLoG is compared against three state-of-the-art detectors: generalized LoG, Radial-Symmetry and LoG using precision, recall, and *F*-score metrics. We observe that HLoG statistically outperforms the compared detectors.

Index Terms—Blob detection, Hessian analysis, scale space, Laplacian of Gaussian (LoG), unsupervised learning.

I. INTRODUCTION

THE rapid development of medical imaging technology has dramatically increased the spatial and temporal resolution, and therefore size, of clinical imaging data. Typically, image segmentation methods are used to delineate specific objects and boundaries. The derived features, such as the number, size, and shape of the objects, are then used for clinical analysis. Some examples of objects in images for clinical studies include cell nuclei in histopathology images or fluoroscopic images to quantify cytology or histology [1]–[3], cerebral blood vessels in magnetic resonance (MR) images to diagnose vascular disease [4], brain tumors in MR images to assess treatment [5], breast lesions in ultrasound images to stage breast cancer lesions [6],

and glomeruli in kidney MR images to characterize renal disease [7], [8]. Of particular interest in this research is a common type of object that is small in structure and convex elliptic in shape. In the field of computer vision, the problem of detecting such objects is known as blob detection. Using mathematical methods, blobs with different properties such as brightness or shape can be identified against the image background.

Extensive research has proposed various blob detectors, among which one popular approach is to use local extrema in a transformed space in conjunction with a vector of derived features (e.g., local intensity histogram or orientation histogram) to identify the blobs. In general, this type of blob detector can be categorized as an interest point detector or interest region detector. As interest point detectors, the Radial-Symmetry [9] and the Radial Gradient Transform detectors [10] use radial symmetric space as the transformation base. One advantage is that the Radial-Symmetry and Radial Gradient Transform detectors are fast to compute and are rotationally invariant, which would be beneficial to detect radial symmetric blobs. However, for radial asymmetric blobs, these two detectors may lose their sensitivities [11]. Other types of interest point detectors like SIFT [12], SURF [13], and BRISK [14] are rooted in a scale-invariant space transformation. Scale-invariant features are extracted, associated with affine invariant features such as the orientation histogram to characterize the local properties in the affine space. These detectors are claimed to be affine invariant such that local structures with similar affine properties within or across images can be identified.

Note interest point detectors are developed for each pixel/voxel, ignoring the contributions of neighboring pixels/voxels to the object. Consequently, these detectors tend to be less tolerant of local noise. Motivated by the affine-invariant interest point detector, interest region detectors are introduced to derive aggregated measures of a number of pixels/voxels within regions of local extrema. Some examples of interest region detectors are the Harris-affine detector, Hessian-affine detector, and Hessian-Laplace detector [15]–[17]. While these detectors are shown to be more robust to noise, they are computationally expensive to adapt the shape estimation. In addition, the massive number of local extrema necessitates careful pruning, which adds to the computational burden. It is also challenging to identify the appropriate pruning parameter, which tends to be subjective for both interest point and interest region detectors.

Although both interest point and region detectors well describe affine invariant properties of local image structures, it is observed in the study by Kong *et al.* [11] that the performance of these detectors on some blob identification problems in pathological and fluorescent images are unsatisfactory. Instead, another type of detector, Laplacian of Gaussian (LoG) detector

Manuscript received May 6, 2014; accepted September 17, 2014. Date of publication September 25, 2014; date of current version March 17, 2015. This work was supported by NIH DK091722, and a grant from the NIH Diabetic Complications Consortium. Asterisk indicates corresponding author.

M. Zhang is with the Department of Industrial Engineering, Arizona State University, Tempe, AZ 85281 USA (e-mail: mzhang33@asu.edu).

*T. Wu is with the Department of Industrial Engineering, Arizona State University, Tempe, AZ 85281 USA (e-mail: teresa.wu@asu.edu).

K. M. Bennett is with the Department of Biology, University of Hawaii, Manoa, HI 96822 USA (e-mail: kevinben@hawaii.edu).

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

Digital Object Identifier 10.1109/TBME.2014.2360154

[18], [19], under the scale space theory, has attracted great attention for the blob detection problem. Scale space theory is a formal theory that considers one image as a stack of images controlled by one parameter (scale parameter, t), in the so-called scale space representation. The Gaussian scale space representation of an image is defined by the convolution of the image function with the Gaussian kernel, preserving important spatial properties of the imaged structures [20], [21]. Specifically, as the scale parameter increases, the number of local minima in a dark blob does not increase, and the number of local maxima in a bright blob does not decrease. This means that neighboring blob objects will diffuse and eventually combine to be identifiable blobs at a specific scale. If a multiscale representation with regards to scale parameter t is constructed, there thus exists one “optimal” scale for blobs with similar sizes. Individual blobs can then be detected with corresponding scale parameters. Previous research has shown that the detector generated by applying the LoG kernels can successfully detect blobs [22], [23]. However, the symmetric nature of the LoG detector limits its performance in rotational asymmetric blob detection and computational cost for scale adaption under multiscale representation is expensive. Therefore, a number of LoG extensions have been proposed. For example, the difference of Gaussian (DoG) [12], [15], [17], [24] is developed to approximate the LoG and improve computational performance. The generalized Laplacian of Gaussian (gLoG) [11] is proposed to extend the detection to rotational asymmetric structures by using different Gaussian kernels. The gLoG is thus able to detect general elliptical structures such as rotationally symmetric and asymmetric blobs.

All detectors reviewed previously use the same workflow, pinpointing the centroid of the blob and carefully pruning to remove the overlapping local extrema. A regular ellipse with a proper radius that is associated with its scale is then superimposed on the center to approximate the blob shapes. Consequently, the derived features (e.g., the volumes of the blobs) are only estimates rather than accurate measurements defined by true boundaries. Furthermore, while scale space theory provides the foundation for complex detectors, it is based on a multiscale representation, which may waste computing effort in optimally selecting scales when the blobs are approximately uniform in size. Uniformly sized blobs are common in a number of clinical applications, such as in detecting cells and nuclei in pathologic or fluorescent microscopic images [11], [25]–[27], and segmenting kidney glomerulus in 3-D MR images [7], [28], [29]. These images generally contain very large number of small blobs, each with a convex elliptical shape. Though “uniform size” may relax the blob recognition problem, there exist some unique challenges: 1) the blob’s small size corresponds to a high spatial frequency close to that of image noise. As a result, small blob detection is sensitive to local noise; 2) the heterogeneous distribution of intensities and heterogeneous boundaries make it difficult to threshold the small blobs from the image background. We contend features derived from small imaging objects/regions may help address the challenges. This will require the detailed delineations of the blobs instead of estimating the blob boundaries.

In this study, the convexity and elliptic shape of the blobs of interest motivates us to explore the Hessian analysis. Here, we

propose a novel approach, named Hessian-based Laplacian of Gaussian detector (HLoG) for small blob detection. Specifically, similar to the aforementioned detectors, a multiscale representation is first derived using LoG for each image. Since the blobs are approximately homogeneous in size, an “optimal” scale can be identified from the Hessian analysis; thus, a single-scale representation can be obtained from the images. Hessian analysis is then applied to presegment the blob candidates. It is known that the theoretical foundation of Hessian guarantees that the presegmentation will recognize all the true blobs, as well as some nonblob objects. This leads to the need to fine-tune the presegmentation results. We note that the detectors reviewed earlier such as LoG, gLoG, and Radial-Symmetry all employ threshold-based fine pruning procedures that may be less tolerant to image noise. Fortunately, the Hessian presegmentation has greatly reduced data size (by filtering out most nonblob regions) and delineated the boundary of blob candidates. As a result, we can afford to extract multiple features from the small blob candidates. In addition to the eight features commonly used in the literature, we introduce three new features to measure the “blobness.” After comprehensive assessment, three out of eleven significant features are selected and used in the variational Bayesian Gaussian mixture model (VBGM) [30] to finalize blob detection. VBGM is chosen here since it is an unsupervised clustering algorithm which is also tuning free. During the LoG transformation, a dark blob is converted to a bright blob and vice versa. To avoid confusion, we define the blob after the LoG operation as transformed-blob hereafter. This paper focuses on the dark blob (transformed-bright blob) identification. The same process applies for the bright blob (transformed-dark blob) identification.

The main contributions of this proposed HLoG detector for small blobs with similar sizes are threefolds. First, the proposed novel Hessian presegmentation is capable of generating blob candidate regions that theoretically enclose all the true blobs and accurately delineate the shape of imaging objects so multiple regional features can be extracted for post pruning. Second, the use of Hessian presegmentation enables the identification of a single optimum scale as the smoothing parameter for normalized LoG filter to greatly reduce the computational burden. Third, together with average intensity information (known from literature), two out of three new regional features, after evaluation, are introduced to characterize the local blobs to prune out the nonblob objects for improved final segmentation.

The following sections are organized as follows: Section II describes our method in details followed by the comparison experiments in Section III. Conclusions are presented in Section IV. All code and results in this paper can be found at our website (<http://swag.engineering.asu.edu/HLoG.htm>).

II. HLOG DETECTOR

A. Normalized LoG Transformation

Given 2-D image $f : R^2 \rightarrow R$, the scale space representation $L(x, y; t)$ at point (x, y) with scale parameter t is the convolution of image $f(x, y)$ with Gaussian kernel $G(x, y; t)$:

$$L(x, y; t) = f(x, y) * G(x, y; t) \quad (1)$$

where $*$ is the convolution operator and $G(x, y; t) = \frac{1}{2\pi t^2} \exp^{-\frac{x^2+y^2}{2t^2}}$. The Laplacian of $L(x, y; t)$ is

$$\nabla^2 L(x, y; t) = \nabla^2 f(x, y) * G(x, y; t). \quad (2)$$

Since differentiation commutes with convolution, we have

$$\nabla^2 L(x, y; t) = f(x, y) * \nabla^2 G(x, y; t). \quad (3)$$

A seminal paper by Lindeberg [23] explains that the LoG response always decreases when t increases resulting the maximum LoG is at the stage without the convolution. A normalizing factor γ is introduced as the power of scale to obtain the maximum LoG invariance over scale [23]:

$$LoG(x, y; t) = f(x, y) * t^\gamma \nabla^2 G(x, y; t). \quad (4)$$

By using γ , the maximum of LoG responses are the same regardless of the scales. Ideally, let the intensity distribution of blob be a perfect Gaussian (without noise and distortion); it is proved that when $\gamma = 2$, the scale invariance is achieved [23]. As a result, the size of blobs can be determined at the scale when the normalized LoG reaches the maximum. Yet, in practice, the normalizing factor γ needs to be tuned to make the LoG maxima invariant to the blob sizes (note we have adopted different γ values in our experiments, see Section III). Let γ be set, (4) transforms the raw image into LoG space ($LoG(x, y; t)$), a t -controlled, smoothed space in which transformed-blobs are highlighted with enhanced boundaries.

In scale space theory, different structures can be highlighted with regards to the scale parameter t in the multiscale space representation. Every transformed-blob can be graphically represented by a circle centered at the local spatial maximum over LoG space with the radius r proportional to the scale at which the maximum over scales is obtained. Since the blobs studied in this research have approximately uniform sizes, one optimal t may be identified for all the blobs in the image. When t is small, unnecessary details can inadvertently be highlighted, leading to oversegmentation. When t is large, several small structures could be identified as one object, leading to undersegmentation. Therefore, determining the optimum t to extract most of small blobs of interest is critical. Fortunately, Hessian analysis can assist in determining the optimum t . In addition, Hessian analysis can be used to highlight blob candidates and preserve the true geometric shapes of the candidate regions as presegmentation.

B. Hessian Presegmentation

1) *Hessian Analysis*: It is known that the eigenvalues of the Hessian matrix of a blob-like structure can be used to describe the structure's geometry. Let the image be smoothed via LoG first, for any pixel (x, y) in the LoG image $LoG(x, y; t)$ at scale t , the Hessian Matrix for this pixel is

$$H(x, y; t) = \begin{pmatrix} \frac{\partial^2 LoG(x, y; t)}{\partial x^2} & \frac{\partial^2 LoG(x, y; t)}{\partial x \partial y} \\ \frac{\partial^2 LoG(x, y; t)}{\partial x \partial y} & \frac{\partial^2 LoG(x, y; t)}{\partial y^2} \end{pmatrix}. \quad (5)$$

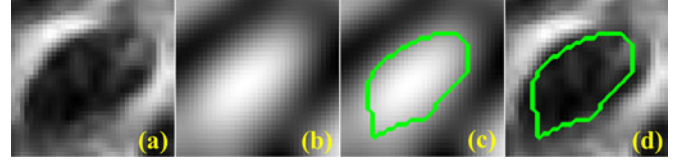


Fig. 1. Transformed-bright blob has a negative-definite Hessian in normalized LoG space. (a) Dark blob in the raw image. (b) Transformed-bright blob after normalized LoG transformation from (a). (c) Region of pixels having negative-definite Hessian in (b) is contoured in green over (b). (d) Region of pixels having negative-definite Hessian in (b) is contoured in green over original image (a). As seen in (d), negative-definite Hessian outlines the blob over (a). The irregular shape of the blob is clearly delineated.

Given geometric classification as a pixel [31] and specific orientation patterns [4], if pixel (x, y) is concave elliptic, both of the eigenvalues λ_1, λ_2 of $H(x, y; t)$ are negative, meaning $\lambda_1 < 0, \lambda_2 < 0$. This motivates us to identify the transformed-bright blobs by the following proposition.

Proposition 1: In a transformed 2-D LoG image, every pixel of a transformed-bright blob has a negative definite Hessian.

Proof: Since every pixel in the transformed-bright blob is concave elliptic, the eigenvalues of its Hessian are negative, requiring that the Hessian matrix is negative definite. ■

Proposition 1 provides one necessary but not sufficient property that a pixel in a transformed-bright blob must satisfy. In other words, if a pixel resides in a transformed-bright blob, the Hessian matrix of the pixel is negative definite. But the pixel having negative definite Hessian may not be from a transformed-bright blob. This proposition provides us the theoretical foundation to identify the blob candidates defined by *Definition 1* to ensure all true blobs are recognized.

Definition 1: A blob candidate T in LoG space is a four-connected component of set $U = \{(x, y) | (x, y) \in LoG(x, y; t), I(x, y; t) = 1\}$, where $I(x, y; t)$ is the binary indicator such that if the pixel (x, y) has a negative definite Hessian, then $I(x, y; t) = 1$; otherwise, $I(x, y; t) = 0$.

Note the definiteness of the Hessian can be assessed by the leading principal minors instead of calculating its eigenvalues of the matrix. Taking a 2×2 matrix A as an example, if A_{11} is negative and determinant of A is positive, then A is negative definite. Following the proposition and the definition, the Hessian matrix can populate the pool of blob candidates that theoretically is the superset of all the blobs with their geometric shapes. Fig. 1 presents an illustrative example showing an identified blob candidate using the Hessian matrix and the true blob for a given scale. In the next section, we will explain how to use the Hessian analysis to obtain the optimal scale.

2) *Hessian-Based Optimal Scale Identification*: In addition to identifying blob candidates, Hessian analysis can be used to determine the optimal scale parameter t . For blobs in different scales, Lindeberg [23] uses the maximum normalized LoG (trace of Hessian) to select the optimum scales across the scale-space for each individual blob. Specifically, each blob achieves the most saliency at the scale at which its average of LoG reaches the maxima. Since only blobs of similar sizes are discussed in this paper, a single scale can be selected to approximate the size

TABLE I
COMPARISON RESULTS OF HLOG, gLOG, RADIAL-SYMMETRY AND LOG ON 15 PATHOLOGIC IMAGES WITHOUT THE POSTPRUNING PROCESS

IMG	d	Hessian			gLoG (no thresholding)			Radial-Symmetry (no thresholding)			LoG (no thresholding)		
		Precision	Recall	F -score	Precision	Recall	F -score	Precision	Recall	F -score	Precision	Recall	F -score
1	13.85	0.915	0.946	0.931	0.929	0.924	0.927	0.339	0.997	0.506	0.038	0.997	0.073
2	13.54	0.804	0.957	0.873	0.844	0.927	0.883	0.285	0.995	0.444	0.032	0.998	0.061
3	14.30	0.696	0.985	0.815	0.677	0.974	0.799	0.246	0.991	0.394	0.018	1.000	0.036
4	11.75	0.846	0.966	0.902	0.933	0.931	0.932	0.314	0.998	0.478	0.040	0.981	0.077
5	13.88	0.884	0.942	0.912	0.907	0.925	0.916	0.339	0.995	0.506	0.040	0.969	0.076
6	12.03	0.826	0.968	0.891	0.920	0.898	0.909	0.398	0.994	0.568	0.034	0.998	0.066
7	14.01	0.781	0.970	0.865	0.822	0.959	0.885	0.304	0.993	0.466	0.030	1.000	0.059
8	14.00	0.846	0.971	0.904	0.881	0.951	0.915	0.405	0.991	0.575	0.034	0.994	0.065
9	14.06	0.836	0.972	0.899	0.875	0.943	0.908	0.354	0.987	0.522	0.029	0.997	0.057
10	11.99	0.840	0.958	0.895	0.922	0.879	0.900	0.297	0.995	0.457	0.041	0.973	0.079
11	14.00	0.808	0.971	0.882	0.850	0.957	0.900	0.296	1.000	0.456	0.031	0.991	0.061
12	14.58	0.790	0.960	0.867	0.805	0.960	0.876	0.300	1.000	0.461	0.025	0.998	0.048
13	13.87	0.774	0.969	0.860	0.812	0.950	0.876	0.281	0.998	0.438	0.028	0.993	0.055
14	12.96	0.829	0.972	0.895	0.875	0.930	0.902	0.361	0.999	0.530	0.030	0.995	0.058
15	13.14	0.809	0.968	0.881	0.862	0.952	0.905	0.335	1.000	0.502	0.027	0.993	0.053
Avg	13.46	0.819	0.965	0.885	0.861	0.937	0.895	0.324	0.996	0.487	0.032	0.992	0.061
Std	0.89	0.051	0.011	0.027	0.066	0.025	0.032	0.044	0.004	0.050	0.006	0.010	0.012

of all blobs. The maximum value of averaged normalized LoG is used here to autodetermine the single optimum scale for small blobs. Let the image be transformed to a normalized multiscale LoG space representation. To determine the optimum scale of the blob candidates, let the average LoG value per candidate pixel measure C_r be

$$C_r(t) = \frac{\sum_{(x,y)} LoG(x,y;t)I(x,y;t)}{\sum_{(x,y)} I(x,y;t)} \quad (6)$$

where $I(x,y;t)$ is the binary indicator defined in *Definition 1*. As $C_r(t)$ increases, the blob candidates are more salient against their background. Therefore, the optimum scale t_{best} can be calculated as

$$t_{best} = \operatorname{argmax}_t C_r(t). \quad (7)$$

Using t_{best} , the raw image is transformed into a single-scale LoG space from which the blob candidates are populated.

3) *Validation of Hessian Presegmentation*: To evaluate the performance of Hessian presegmentation, the precision, recall, and F -score metrics are used. Precision measures the fraction of retrieved candidates that are relevant to the ground truth. Recall measures the fraction of ground-truth data retrieved. F -score measures overall performance. Since ground-truth data are provided in the form of dots (the coordinates of the blob centers), the same as in the literature [11], [25], a candidate is considered to be true positive if its intensity centroid is within an evaluation threshold parameter d of the corresponding ground truth dot. Specifically, if the Euclidian distance between dot i and blob candidate j is less equal than d , the blob is considered true positive. To avoid duplicate counting, the number of true positives TP is calculated by (8). Precision, recall, and F -score are calculated by

$$TP = \min\{\#\{j : \min_{i=1}^m D_{ij} \leq d\}, \#\{i : \min_{j=1}^n D_{ij} \leq d\}\} \quad (8)$$

TABLE II
ANOVA ON DETECTORS WITHOUT POSTPRUNING USING TUKEY'S HSD PAIRWISE TEST ON 15 PATHOLOGIC IMAGES WITH 0.05 SIGNIFICANCE LEVEL

Contrast (Hessian v.s)	Precision	Recall	F -score
	p (significant)	p	p
gLoG	0.081(No)	< 0.0001 (Yes)	0.813 (No)
Radial-Symmetry	< 0.0001 (Yes)	0.893 (No)	< 0.0001 (Yes)

$$\text{precision} = \frac{TP}{n} \quad (9)$$

$$\text{recall} = \frac{TP}{m} \quad (10)$$

$$F\text{-score} = 2 \times \frac{\text{precision} \times \text{recall}}{(\text{precision} + \text{recall})} \quad (11)$$

where m is the number of ground truth and n is the number of blob candidates and d is a thresholding parameter for evaluation purpose and can be set to a positive value $(0, +\infty)$. If d is small, fewer blob candidates will be counted since the distance between blob candidate centroid and ground-truth should be small in order to be counted. If d is set to large, more blob candidates will be counted due to the relaxation of the distance. In this paper, since the local intensity extreme could be located anywhere within the small blob region with an irregular shape, we set evaluation threshold parameter d to the average diameter of blobs: $d = 2 \times \sqrt{\sum_{(x,y)} I(x,y;t)/\pi}$. The comprehensive experimental results with different values of d will be discussed in Section III-C.

For validation purpose, three commonly used detectors are chosen for comparison. One is an interest point detector, Radial-Symmetry detector, which is fast and has good performance on rational symmetric image structure detection [9]. The other two are LoG detectors which are specifically designed for blob

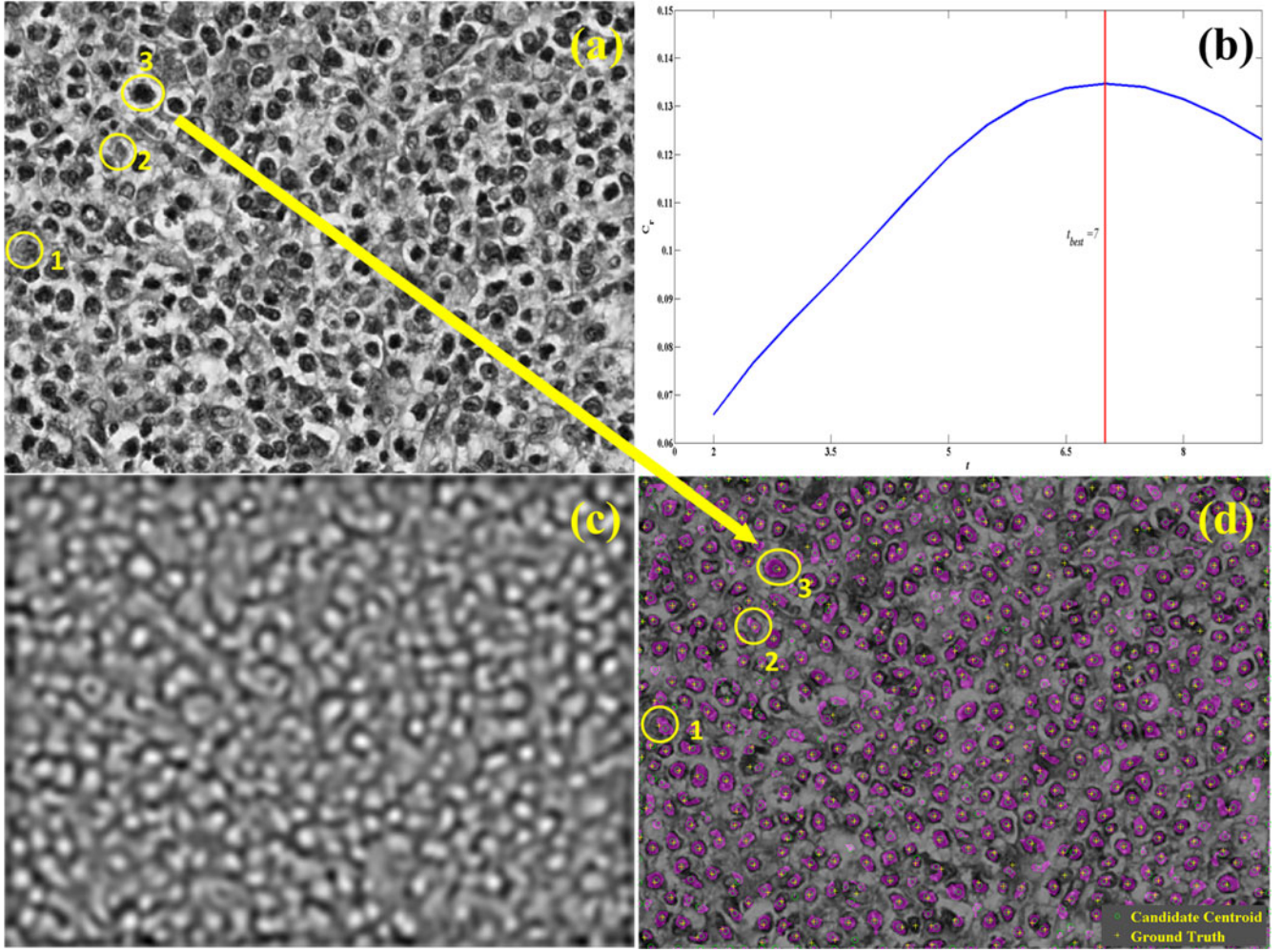


Fig. 2. Hessian presegmentation results on selected pathologic cell image. (a) Raw image. (b) Autoscale selection: section $t = t_{best} = 7$ is selected. (c) LoG transformed image. (d) Hessian presegmentation result: purple mask shows the shapes retrieved by the Hessian presegmentation method.

detection: the original LoG method [23] and its variant—gLoG method [11]. Note that all three detectors employ threshold-based post pruning to finalize the segmentation. To validate the Hessian analysis as a presegmentation procedure, we compare the Hessian presegmentation with the three detectors without the postpruning. The comprehensive comparison of HLoG with the three detectors with postpruning is provided in Section III.

In this experiment, a dataset of 15 600×800 pathologic cell images [11] is studied. The source code for the LoG and gLoG algorithms is implemented from [11] and source code of Radial-Symmetry is implemented from [9] (all those source codes are available online). Since the Radial-Symmetry MATLAB package only provides the transformation from raw image to Radial-Symmetry space, we use the radial-symmetric centers (local extrema) with maximum intensity values as the centroids of the small blobs.

The parameter settings for the proposed method are the following: the normalizing factor γ is set to 2 based on our rough tuning experiments. The optimum scale t_{best} of small blobs is automatically selected for each images with (6) and (7) from the range of 2–10 and step size 0.5. For each image, all the algo-

rithms adopt the same value d ($d = 2 \times \sqrt{\sum_{(x,y)} I(x,y;t) / \pi}$ as discussed earlier in this section).

The performance of Hessian presegmentation on the three metrics, i.e., precision, recall, and F -score, is summarized in Table I with regards to t_{best} and d . It is interesting to note that though both gLoG and Radial-Symmetry detectors claim that no postpruning is necessary, the results indicate the room for much improvement. Indeed, in the literature, both detectors [9], [11] have employed the postpruning for experiments. As for LoG, it is designed to have postpruning as a must. Therefore, it is not surprising that the performance of LoG in this validation experiment is far inferior.

In looking at the average performance (see Table I), gLoG performs the best (0.861) in precision and F -score (0.895), and Radial-Symmetry performs the best in recall (0.996). We further conduct analysis of variance (ANOVA) for statistical conclusion. As seen in Table II, there is no significant difference in precision and F -score in comparing our Hessian presegmentation with gLoG. The same conclusion is drawn in comparing our Hessian presegmentation with Radial-Symmetry in recall.

However, Hessian presegmentation significantly outperforms gLoG in recall, and Radial-Symmetry in both precision and F -score.

Theoretically, all the true blobs with elliptic convex shape assumption can be retrieved by Hessian presegmentation which is proved by *Proposition 1*. However, in practice, as expected, the Hessian presegmentation in this experiment retrieves most true blobs (96.5% close to 100%). The discrepancy is due to the modeling error that some true blobs may not comply with the elliptic convex assumption.

As discussed earlier, the results from Table I indicate that both gLoG and Radial-Symmetry need postpruning for improved performance. All three comparison detectors employ the thresholding-based postpruning process. We argue that the advantage of our proposed Hessian presegmentation enables the extraction of blob specific features to be used in the postpruning. Such advantage can be shown in the full comparison (see Section III).

For illustration purposes, Fig. 2 shows the process of our proposed Hessian presegmentation on a selected pathologic image. For this specific image, $t_{\text{best}} = 7$ [see Fig. 2(b)]. Fig. 2(c) shows the LoG-transformed image, in which small blob structures are enhanced. Based on the transformed image, the Hessian presegmentation method generates blob candidates [purple in Fig. 2(d)], and the LoG intensity extrema are marked as green circles to represent the centroids of these candidates compared to the ground truth dots marked as yellow cross [see Fig. 2(d)]. For a better view, we circle three representative blobs in Fig. 2(a). As seen, these blobs have inhomogeneous intensity distributions: blob 3 is much darker than blobs 1 and 2. Blobs 1 and 2 also have ambiguous boundaries against the background. Fig. 2(d) shows that Hessian presegmentation is able to recognize these blobs and we conclude the Hessian analysis is robust for identifying blobs with inhomogeneous intensity distributions and blurred boundaries. In the next section, the regional features extracted from the blobs are discussed.

C. Features Extraction and Evaluation

1) *Regional-Based Features Extraction*: There are two common geometric measures in blob detection: R —the likelihood of “blobness” measure, and S —the second-order structureness used by [4]

$$R = \frac{|\lambda_1|}{|\lambda_2|} \quad (12)$$

$$S = \sqrt{(\lambda_1^2 + \lambda_2^2)} \quad (13)$$

where λ_1 and λ_2 are eigenvalues of the Hessian and $|\lambda_1| \leq |\lambda_2|$. R is the ratio of the two principal curvatures and falls in $(0, 1]$. (The negative definite Hessian guarantees $\lambda_1, \lambda_2 \neq 0$.) If $R \rightarrow 1$, and the curvatures along two principal directions are similar, the pixel most likely resides in a blob-like structure; if $R \rightarrow 0$, i.e., the curvatures along two principal directions are fairly different and the pixel is most likely on a line [4]. Measure S indicates the contrast of the pixel against the background. With larger S , the pixel within the object is more salient against the

background. Given R and S , the single measure to describe blobness is given by [6] as

$$B(R, S) = \left(1 - \exp\left(-\frac{R^2}{2\beta^2}\right)\right) \left(1 - \exp\left(-\frac{S^2}{2c^2}\right)\right) \quad (14)$$

where β and c are constant parameters that control the sensitivity of R and S , respectively, which are usually set to 0.5. Although these measures quantify the geometric information of blobs, the calculations are relatively computationally expensive. Computing the likelihood of blobness R is particularly expensive because it requires the calculation and sorting of all eigenvalues in every pixel. To efficiently calculate the likelihood of blobness, we propose a modified likelihood blobness measure R' for each pixel as

$$\begin{aligned} R'(x, y; t) &= \frac{2}{\frac{|\lambda_1|}{|\lambda_2|} + \frac{|\lambda_2|}{|\lambda_1|}} = \frac{2|\lambda_1||\lambda_2|}{(\lambda_1 + \lambda_2)^2 - 2|\lambda_1||\lambda_2|} \\ &= \frac{2 \times |\det(H(\text{LoG}(x, y; t)))|}{\text{tr}(H(\text{LoG}(x, y; t)))^2 - 2 \times |\det(H(\text{LoG}(x, y; t)))|} \quad (15) \end{aligned}$$

where $H(\cdot)$ is the Hessian. The advantage of this modification is that instead of calculating the eigenvalues λ_1, λ_2 , only trace and determinant are calculated which is more computationally efficient. Moreover, R' functions the same as R for the measurement of blobness as replacement. Next, we will prove that R' is a monotonic increasing function of R , thus preserving the ability to measure the blobness.

Proposition 2: R' is a monotonic increasing function of R .

Proof: When $|\lambda_1| \leq |\lambda_2|$, $R = \frac{|\lambda_1|}{|\lambda_2|}$, $R' = \frac{2}{R + \frac{1}{R}}$; When $|\lambda_1| \geq |\lambda_2|$, $R = \frac{|\lambda_2|}{|\lambda_1|}$, $R' = \frac{2}{\frac{1}{R} + R}$, so $R' = \frac{2}{R + \frac{1}{R}}$. Then, $\frac{\partial R'}{\partial R} = \frac{\frac{1}{R^2} - 1}{(\frac{1}{R} + R)^2} > 0$ when $R \in (0, 1)$. Therefore, R' is a monotonic increasing function of R . ■

This proposition proves that R' is a valid and efficient replacement of R that preserves the measurement of blobness. Equation (15) shows that $0 < R' \leq 1$. For any pixel, if it is within a blob-like structure, R is close to 1, the modified R' is also close to 1; otherwise, R is close to 0, the modified R' is also close to 0. We conclude that R' is equivalent to R in measuring blobness, and R' is much efficient to compute compared to R since it only requires the computation of trace and determinant. Keeping second-order structureness the same, (13) can be rewritten using the trace and determinant of the Hessian:

$$\begin{aligned} S(x, y; t) &= \sqrt{\text{tr}(H(\text{LoG}(x, y; t)))^2 - 2 \times \det(H(\text{LoG}(x, y; t)))}. \quad (16) \end{aligned}$$

The modified blobness measure based on R' and S is $B' = B(R', S)$. Since we are interested in the features within the regions, i.e., blob candidates instead of individual pixels, aggregated measures of the pixels within each region are required. One approach is to average the measures, giving us $R_{T,\text{mean}}$, $S_{T,\text{mean}}$, and $B_{T,\text{mean}}$. Alternatively, the maximum values of the measures within a region can be calculated as $R_{T,\text{max}}$, $S_{T,\text{max}}$, and $B_{T,\text{max}}$. In this research, given the true shape information available from the Hessian presegmentation

and inspired by the design of Harris regional detectors [15], [16], [32], we introduce three new measures: R_T —regional likelihood of blobness, S_T —regional structureness, and B_T for each blob candidate T (a function of R_T and S_T , as (14)) based on the matrix constructed by the sum of second-order derivatives over the candidate T :

$$H_T = \sum_{(x,y) \in T} H(x,y;t)$$

$$= \begin{bmatrix} \sum_{(x,y) \in T} \frac{\partial^2 (LoG(x,y;t))}{\partial x^2} & \sum_{(x,y) \in T} \frac{\partial^2 LoG(x,y;t)}{\partial x \partial y} \\ \sum_{(x,y) \in T} \frac{\partial^2 LoG(x,y;t)}{\partial x \partial y} & \sum_{(x,y) \in T} \frac{\partial^2 (LoG(x,y;t))}{\partial y^2} \end{bmatrix} \quad (17)$$

Based on this matrix, we get the regional likelihood of blobness R_T :

$$R_T = \frac{2 \times |det(H_T)|}{S_T^2} \quad (18)$$

where

$$S_T = \sqrt{(tr(H_T))^2 - 2 \times det(H_T)}. \quad (19)$$

The summation of Hessian matrix in (17) describes the second-order derivative distribution within the region of blob candidate. The derivatives are equally weighted averaged (sum over the region T) at the centroid of T over the region. The eigenvalues of this matrix represent the two principal curvatures of the centroid over blob candidate. Thus, this can be utilized to measure the likelihood of blobness over the region. Together, we have three groups of features for blobness measures: mean based, max based, and blob-candidate based. In addition, the common features average intensity of candidate T (M_T) and the area of candidate T (A_T) are added into the features list. Table III summarizes the features and their calculations.

2) *Regional-Based Features Selection*: With the features extracted earlier, a clustering algorithm based on the VBGMM [30] is used to evaluate the contributions from the features in identifying the blobs. Unlike the supervised learning algorithm, where every feature contributes with adjusted weights after training, the VBGMM is an unsupervised learning algorithm with equal weights for all features without training. Since this paper studies blobs with similar size, to verify the trivialness of blob size feature, we add additional tests with and without area features. Also, because blobness is the nonlinear combination of likelihood of blobness and structureness, these features need to have separate experiments to test their performances. Therefore, to evaluate the regional features based on different types of measures, the features listed in Table III are categorized into the 18 features sets ($2 \times 3 \times 3$) listed in Table IV. The experiments are performed on the same dataset as for Hessian presegmentation. The parameter settings and evaluation metrics are the same as in Section II-B for model setup and performance evaluations. The results are shown in Table IV.

Table IV shows that the feature set (M_T, R_T, S_T) provides the best performance on the testing data over other feature

TABLE III
SUMMARY OF FEATURES OVER BLOB CANDIDATE t

Group	Features	Formulation	Description
Common Features	M_T	$\frac{\sum_{(x,y) \in T} f(x,y)}{A_T}$	Average Intensity
	A_T	$\sum_{(x,y) \in T} 1$	Area
Using regional max information	$R_{T,max}$	$\max\{R'(x,y;t_{best}), (x,y) \in T\}$	Max likelihood of blobness
	$S_{T,max}$	$\max\{S(x,y;t_{best}), (x,y) \in T\}$	Max Structureness
	$B_{T,max}$	$\max\{B(R', S), (x,y) \in T\}$	Max blobness
Using regional mean information	$R_{T,mean}$	$\text{mean}\{R'(x,y;t_{best}), (x,y) \in T\}$	Mean likelihood of blobness
	$S_{T,mean}$	$\text{mean}\{S(x,y;t_{best}), (x,y) \in T\}$	Mean Structureness
	$B_{T,mean}$	$\text{mean}\{B(R', S), (x,y) \in T\}$	Mean blobness
Proposed Features	R_T	$\frac{2 \times det(H_T) }{S_T^2}$	Regional likelihood of blobness
	S_T	$\sqrt{(tr(H_T))^2 - 2 \times det(H_T) }$	Regional structureness
	B_T	$B(R_T, S_T)$	Regional blobness

TABLE IV
EVALUATION OF DIFFERENT FEATURE SETS ON 15 PATHOLOGIC IMAGES (MEAN AND STANDARD DEVIATION OF THE MEASURES ON PRECISION, RECALL, F-SCORE)

Feature set	Precision <i>Mean ± Std</i>	Recall <i>Mean ± Std</i>	F-score <i>Mean ± Std</i>
$M_T, A_T, R_{T,max}, S_{T,max}, B_{T,max}$	0.856 ± 0.046	0.953 ± 0.014	0.901 ± 0.021
$M_T, A_T, R_{T,mean}, S_{T,mean}, B_{T,mean}$	0.865 ± 0.046	0.950 ± 0.014	0.904 ± 0.021
$M_T, A_T, B_{T,max}$	0.832 ± 0.054	0.964 ± 0.011	0.892 ± 0.029
$M_T, A_T, R_{T,mean}, S_{T,mean}, B_{T,mean}$	0.821 ± 0.051	0.965 ± 0.011	0.886 ± 0.027
$M_T, A_T, R_{T,mean}, S_{T,mean}$	0.828 ± 0.053	0.964 ± 0.011	0.890 ± 0.028
$M_T, A_T, B_{T,mean}$	0.832 ± 0.053	0.964 ± 0.011	0.892 ± 0.029
M_T, A_T, R_T, S_T, B_T	0.969 ± 0.017	0.767 ± 0.051	0.855 ± 0.030
M_T, A_T, R_T, S_T	0.945 ± 0.027	0.880 ± 0.034	0.911 ± 0.016
M_T, A_T, B_T	0.963 ± 0.022	0.822 ± 0.029	0.887 ± 0.014
$M_T, R_{T,max}, S_{T,max}, B_{T,max}$	0.854 ± 0.047	0.953 ± 0.014	0.900 ± 0.022
$M_T, R_{T,max}, S_{T,mean}, B_{T,max}$	0.862 ± 0.047	0.950 ± 0.014	0.903 ± 0.022
$M_T, B_{T,max}$	0.828 ± 0.057	0.964 ± 0.011	0.890 ± 0.030
$M_T, R_{T,mean}, S_{T,mean}, B_{T,mean}$	0.822 ± 0.053	0.965 ± 0.011	0.887 ± 0.028
$M_T, R_{T,mean}, S_{T,mean}$	0.828 ± 0.057	0.964 ± 0.011	0.890 ± 0.030
$M_T, B_{T,mean}$	0.828 ± 0.057	0.964 ± 0.011	0.890 ± 0.030
M_T, R_T, S_T, B_T	0.972 ± 0.015	0.775 ± 0.049	0.861 ± 0.029
M_T, B_T	0.959 ± 0.028	0.829 ± 0.043	0.888 ± 0.018

sets. This feature set balances the precision and recall measures leading to the highest F -score. Moreover, the feature set (M_T, R_T, S_T) is relatively more stable than other feature sets since it has minimal standard deviation across all the combinations. In addition, we conduct the ANOVA to test the statistical performance of the selected feature set against other features. As shown in Table V, the test groups the feature sets into four

TABLE V
STATISTICAL GROUPS BASED ON ANOVA TUKEY'S HSD PAIRWISE TEST

Category	F-score means	Statistical groups*
M_T, R_T, S_T	0.924	A
M_T, A_T, R_T, S_T	0.911	A B
$M_T, A_T, R_T, \max, S_T, \max$	0.904	A B
$M_T, R_T, \max, S_T, \max$	0.903	A B
$M_T, A_T, R_T, \max, S_T, \max, B_T, \max$	0.901	A B
$M_T, R_T, \max, S_T, \max, B_T, \max$	0.900	A B
M_T, A_T, B_T, \max	0.892	B C
M_T, A_T, S_T, \max	0.892	B C
$M_T, R_T, \max, S_T, \max$	0.890	B C
M_T, B_T, \max	0.890	B C
M_T, B_T, \max	0.890	B C
$M_T, A_T, R_T, \max, S_T, \max$	0.890	B C
M_T, B_T	0.888	B C
M_T, A_T, B_T	0.887	B C D
$M_T, R_T, \max, S_T, \max, B_T, \max$	0.887	B C D
$M_T, A_T, R_T, \max, S_T, \max, B_T, \max$	0.886	B C D
M_T, R_T, S_T, B_T	0.861	C D
M_T, A_T, R_T, S_T, B_T	0.855	D

* Note: there is no statistically significant difference (with 0.05 significance level) between feature sets within a statistical group.

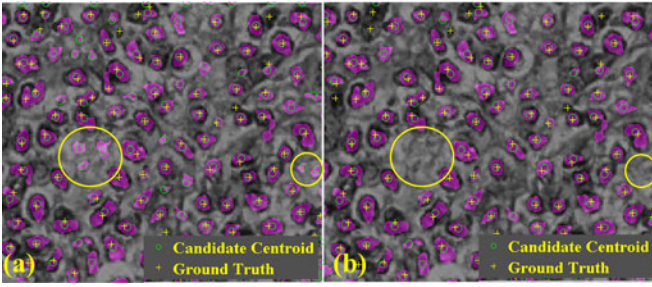


Fig. 3. Blob identification result from part of Fig. 2(a). (a) Presegmentation result. (b) Final identification result. The purple regions are blob candidates and their centroids are marked as green circle, while the centroids of ground-truth data are marked as yellow cross.

groups (A, B, C, D). The selected feature set (M_T, R_T, S_T) statistically outperforms the 12 feature sets from groups B, C, and D. Though there is no significant difference between the feature set (M_T, R_T, S_T) with the other five feature sets from group A, the other five feature sets have also been assigned to group B (inferior to group A). Therefore, we consider (M_T, R_T, S_T) as a reasonable feature set describing the characteristics of blobs.

Fig. 3 shows a comparison between Hessian presegmentation (from Fig. 2) and final identification using the three regional features. False-positive blobs are removed, as shown in the circles in Fig. 3. We conclude that using the three features in VBGMM can refine the presegmentation by removing false-positive blobs in the image. Note that the pruning may also remove a few true blobs. This is seen by comparing Table I (precision: 0.819, recall: 0.965, F -score: 0.885), and Table III (precision: 0.924, recall: 0.925, F -score: 0.924). The recall measure decreases by 0.04, and the precision increases from 0.819 to 0.924. The overall performance of F -score increases from 0.885 to 0.924. We conclude that regional features-based postpruning is promising to improve the blob detection.

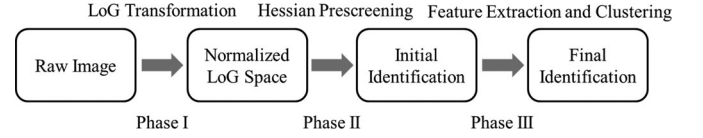


Fig. 4. HLOG for blob detection.

TABLE VI
DETAIL STEPS OF HLOG

1. Initialize the normalize factor γ , range and step-size of parameter t to transform the raw image into normalized LoG space
2. Binarize each section of normalized LoG space with the negative definite Hessian (for dark small blob in raw image).
3. Calculate average LoG intensity $C_r(t) = \frac{\sum_x \sum_y L o G(x, y; t) I(x, y; t)}{\sum_x \sum_y I(x, y; t)}$ and find optimum scale section by $t_{best} = \arg\max_t C_r(t)$.
4. Choose the optimum scale section $t = t_{best}$ and extract the regional features M_T in raw image space and R_T, S_T in LoG space
5. Input those three features to variational Bayesian mixture models with two clusters setting
6. Choose the cluster with highest value of R_T as final segmentation

In the next section, we discuss the integration of Hessian analysis, regional feature extraction, and clustering-based final pruning for the blob identification problem.

D. HLoG for Blob Identification

We propose a three-phased HLoG workflow for blob detection, integrating raw image transformation, Hessian presegmentation, feature extraction, and evaluation (see Fig. 4). First, the raw image is transformed into Normalized LoG space. Next, the Hessian presegmentation method is conducted for initial segmentation to generate the blob candidates. Finally, the average intensity feature M_T , regional likelihood of blobness R_T , and regional structureness S_T are used in the VBGMM clustering algorithm for identification. The VBGMM is more robust than maximum likelihood Gaussian mixture models because it treats parameters, i.e., mean vector and variance-covariance matrix in Gaussian mixtures models as distributions instead of deterministic values and uses hyper parameters to control them. This helps avoid the singularity issues faced by the maximum likelihood Gaussian mixture models. In addition, unlike other pruning algorithms like thresholding, the VBGMM requires no parameter tuning. The detailed steps are listed in Table VI.

III. COMPARISON EXPERIMENTS

In this section, three sets of experiments are conducted to validate the performance of our proposed HLoG detector. In the first set of experiments, the complete version of the three blob detectors above is compared on 15 pathological images and new supplemental data consisting of 200 fluorescence microscopy cell images [33] are tested in the second set of experiments. The 200-cell image dataset is of interest because the blobs are small, and each image can be used to test the performance of the algorithm in tolerating the noise from the background. The first two sets of experiments are to evaluate the performance given

TABLE VII
COMPARISON RESULTS OF COMPLETE VERSION OF HLoG, gLoG, RADIAL-SYMMETRY AND LoG ON 15 PATHOLOGIC IMAGES

IMG	d	HLoG			gLoG			Radial-Symmetry			LoG		
		Precision	Recall	F -score	Precision	Recall	F -score	Precision	Recall	F -score	Precision	Recall	F -score
1	13.85	0.980	0.886	0.931	0.950	0.903	0.926	0.932	0.903	0.918	0.832	0.897	0.863
2	13.54	0.969	0.884	0.925	0.906	0.915	0.911	0.858	0.927	0.891	0.825	0.889	0.855
3	14.30	0.858	0.930	0.892	0.769	0.971	0.858	0.704	0.958	0.812	0.691	0.932	0.793
4	11.75	0.900	0.949	0.924	0.943	0.916	0.929	0.939	0.863	0.899	0.825	0.769	0.796
5	13.88	0.962	0.897	0.928	0.951	0.920	0.935	0.911	0.921	0.916	0.825	0.734	0.777
6	12.03	0.902	0.943	0.922	0.941	0.899	0.919	0.944	0.871	0.906	0.810	0.885	0.846
7	14.01	0.948	0.955	0.952	0.900	0.950	0.924	0.816	0.952	0.878	0.844	0.889	0.866
8	14.00	0.957	0.910	0.933	0.939	0.947	0.943	0.897	0.951	0.923	0.811	0.855	0.833
9	14.06	0.936	0.917	0.926	0.915	0.937	0.926	0.869	0.924	0.896	0.829	0.880	0.854
10	11.99	0.898	0.946	0.922	0.939	0.865	0.901	0.943	0.857	0.898	0.822	0.803	0.812
11	14.00	0.905	0.954	0.929	0.877	0.946	0.910	0.865	0.964	0.912	0.704	0.869	0.778
12	14.58	0.897	0.943	0.919	0.856	0.947	0.899	0.810	0.970	0.882	0.802	0.854	0.827
13	13.87	0.878	0.937	0.906	0.885	0.942	0.913	0.818	0.965	0.886	0.767	0.726	0.746
14	12.96	0.937	0.935	0.936	0.904	0.926	0.915	0.905	0.942	0.923	0.810	0.858	0.833
15	13.14	0.941	0.894	0.917	0.888	0.932	0.910	0.876	0.941	0.907	0.803	0.890	0.844
Avg	13.46	0.924	0.925	0.924	0.904	0.928	0.915	0.873	0.927	0.897	0.800	0.849	0.821
Std	0.89	0.036	0.025	0.013	0.048	0.026	0.020	0.065	0.038	0.027	0.045	0.062	0.036

F -score metric is highlighted in gray since it provides a comprehensive measurement to evaluate the performance.

TABLE VIII
ANOVA USING TUKEY’S HSD PAIRWISE TEST ON 15 PATHOLOGIC IMAGES
WITH 0.05 SIGNIFICANCE LEVEL

Contrast (HLoG v.s)	Precision	Recall	F -score
	p (significant)	p	p
gLoG	0.679 (No)	0.998 (No)	0.741 (No)
Radial Symmetry	0.029 (Yes)	0.999 (No)	0.023 (Yes)
LoG	<0.0001 (Yes)	<0.0001 (Yes)	<0.0001 (Yes)

the estimated diameter d of blob candidates. As explained in Section II, the choice of d may impact the evaluation outcomes. Here, we further conduct a separate experiment to validate the performance of HLoG with different values of d compared to the other detectors on both image datasets.

A. Experiments on Pathologic Images

Since the results of detection by the complete version of the gLoG method, Radial-Symmetry method and LoG methods on 15 pathological images are available online [11], the results are directly used in this paper to avoid any of the parameters tunings. The parameter settings for HLoG are the same as presented in Section II: the normalizing factor γ is set to 2, and the range of parameter t is set from 2 to 10 with step-size 0.5, resulting in 17 sections of normalized LoG space in total. For each image, all the algorithms adopt the same value d , which is the estimated average diameter of all small blobs, to calculate the precision, recall, and F -score. The results are shown in Table VII.

As shown in Table VII, HLoG outperforms the three algorithms on the F -score for 11 out of 15 images. HLoG underperforms the gLoG method for four images where the blob candidates are either underpruned (image 4 and image 13) or overpruned (image 5 and image 8) resulting in a lower F -score. This is due to the fact that the parameter (γ) is tuned to the group of images instead of each individual image. We argue

that the underpruned and overpruned issue may be overcome by tuning the parameter (γ) for each image, yet this will require manual processing. In looking into overall performance, our approach outperforms the three algorithms on average F -score with lower variation. Radial-Symmetry and LoG do not perform as well as gLoG and HLoG, evidenced by lower detection of rotationally asymmetric blobs. HLoG balances the recall and precision metrics leading to overall better F -score. Although the average recall (0.925) of HLoG is marginally lower than that of the gLoG (0.928) and Radial-Symmetry (0.927), the precision performance is significantly better. The pruning algorithm sacrifices marginal recall to improve F -score and precision.

We further conduct ANOVA Tukey’s HSD test to draw statistical conclusions. Table VIII indicates, on F -score metric, HLoG statistically outperforms the Radial-Symmetry and LoG methods ($p < 0.05$) while comparable to the gLoG.

B. Experiments on Fluorescence Images

A new supplemental dataset is added in this experiment consisting of 200 256×256 fluorescence-light microscopy cell images. Unlike the data used earlier, these images contain bright blobs rather than dark blobs. Therefore, the data are converted into image with dark small structures by $1 - f(x, y)$. (We assume that $f(x, y)$ varies from 0 to 1; otherwise, we need to standardize $f(x, y)$ into $[0, 1]$ range.)

For the LoG algorithm, the range t is set as $\log(t) \in [0.5, 3]$ with step-size 0.2, as suggested in [11], and the extrema intensity value is set to 0.005 based on our experiments after tuning. For the gLoG algorithm, $\alpha = 1$, $\sigma_{\text{step}} = -1$, $\theta_{\text{step}} = \pi/9$ and the postpruning threshold is set to 100. (This algorithm uses the intensity range $[0, 255]$ rather than $[0, 1]$.) The other parameter settings are the same as presented in [11]. For Radial-Symmetry detector, based on our tuning after experiments, the local intensity threshold is set to 0.0003, and the postpruning threshold is set to 1 for refinement. In HLoG, the normalizing factor γ is set to 1 to avoid over smoothing since many blobs are clustered in

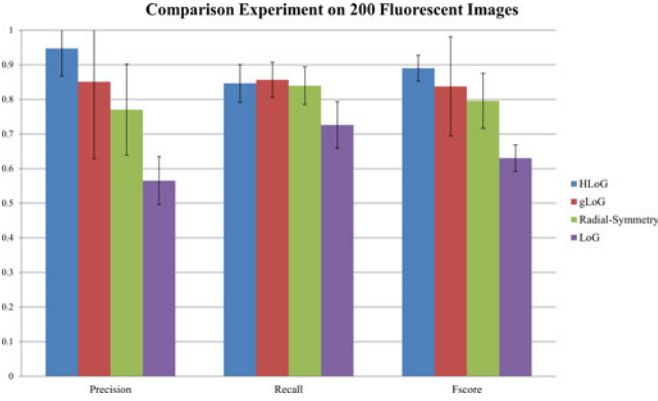


Fig. 5. Comparison results of full version of HLoG, gLoG, Radial-Symmetry, and LoG on 200 Fluoro images. The error bar indicates the standard deviation of the corresponding measure across 200 images.

TABLE IX
ANOVA USING TUKEY'S HSD PAIRWISE TEST ON 200 FLUORESCENT IMAGES
WITH 0.05 SIGNIFICANCE LEVEL

	Precision	Recall	F-score
Contrast (HLoG v.s.)	p (significant)	p	p
gLoG	<0.0001(Yes)	0.269 (No)	<0.0001(Yes)
Radial Symmetry	<0.0001(Yes)	0.654 (No)	<0.0001(Yes)
LoG	<0.0001(Yes)	<0.0001(Yes)	<0.0001(Yes)

this set of images by observation. The scale-space representation is the same as that of the first experiment. Similarly, the parameter d is the same for all the algorithms.

Fig. 5 compares HLoG to the gLoG, LoG and Radial-Symmetry algorithms. The result shows that though HLoG is comparable to gLoG and Radial Symmetry algorithms on recall, it outperforms the three algorithms in both precision and F -score. The variation of our results is also lower than others. (The standard deviation of F -score in HLoG is 0.0377, compared to 0.1436 with the gLoG method, 0.0795 with the Radial-Symmetry method, and 0.0385 with the LoG method.) We conclude that HLoG provides more accurate and stable detection of blobs in this dataset. Again, statistical analysis is performed with the results summarized in Table IX. It is observed that while comparable to the three algorithms on the recall metric, our approach statistically outperforms the comparison algorithms on precision and F -score.

Fig. 8 shows detection of cells in a single fluorescence image. In Fig. 8(b), false-positive blobs using the gLoG algorithm [example as yellow circle 1 in Fig. 8(b)] are caused by the added noise, since gLoG uses aggregated LoG map which may be sensitive to local noise. Some false-positive blobs using Radial-Symmetry detector are due to the symmetric structures near true blobs as shown circle 2 in Fig. 8(c), because Radial-Symmetry detector only detects symmetric structures and lacks the ability to distinct the differences between symmetric structures. There are false blobs using LoG around the edge of the image in Fig. 8(d) (circle 3 as an example). This is because many local extrema occurring around the boundary of the image in multi-

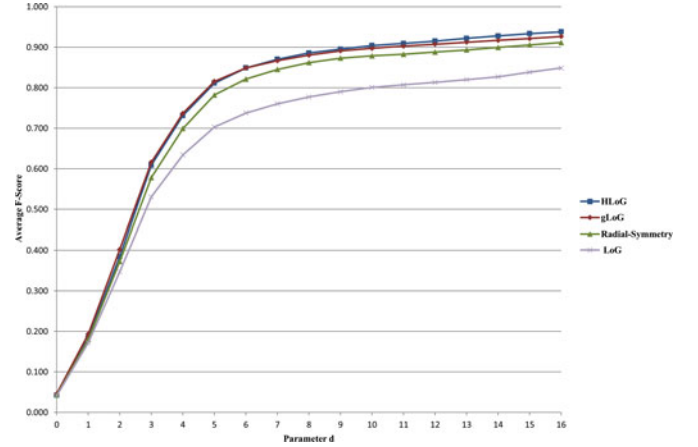


Fig. 6. F -score of HLoG, gLoG, Radial-Symmetry, and LoG on 15 pathological images at different parameter d .

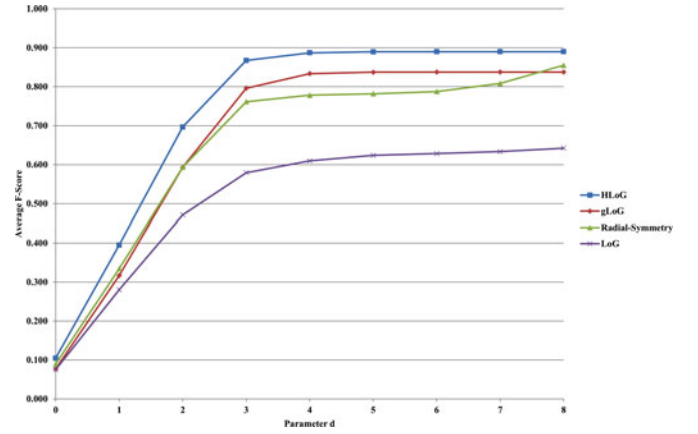


Fig. 7. F -score of HLoG, gLoG, Radial-Symmetry and LoG on 200 pathological images at different parameter d .

scale space which are difficult to remove only by thresholding. However, those types of false-positive blobs are not shown in HLoG as seen in circles 1, 2, and 3 in Fig. 8(a). This is because rather than utilizing symmetric properties and using thresholding, HLoG uses the three regional geometric features for pruning and is therefore more robust in the presence of removing false-positive blobs against noisy background.

C. Evaluation of HLoG at Different d

The performance metrics (precision, recall, and F -score) calculated by (8)–(11) could be highly affected by the value of parameter d . In the previous experiments, d is set to be the estimated diameter of blobs generated by Hessian presegmentation. To explore the effects of the change of d on the performance of HLoG detector, additional experiments are conducted.

Figs. 6 and 7 show the comparison results of HLoG, gLoG, Radial-Symmetry, and LoG at different d on 15 pathological images and 200 fluorescence images, respectively. Since F -score is the geometric average of precision and recall, only F -score is plotted in the Figure. As discussed in Section II-B3 when d is increasing, the average F -score will increase. More and more blob candidates will be treated as true positives because more and more blob candidates have their distance to the ground truth

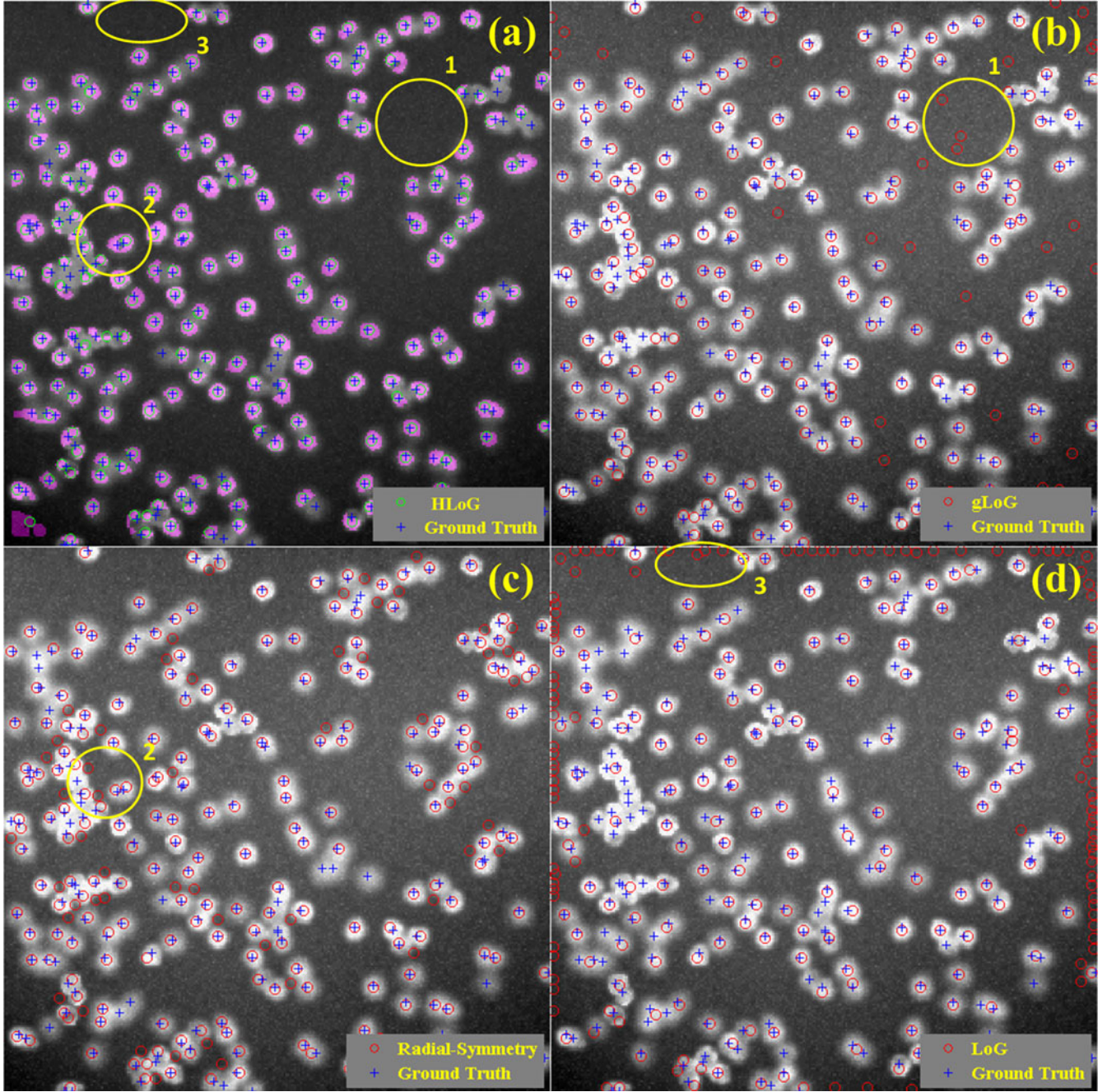


Fig. 8. Detection results on selected fluorescent image. (a) Detection result by HLoG. (b) Detection result by the gLoG algorithm. (c) Detection result by the Radial-Symmetry algorithm. (d) Detection result by LoG.

within the range of d . Since the blob size and image size (256×256) of fluorescence data are smaller than the pathological image (600×800), the range of d is set to be small ($[0, 8]$ compared to $[0, 16]$) to avoid it being greater than the distance of two neighboring blobs.

From both Figures, it is evident that HLoG outperforms other detectors on F -score across the change of d . On the first set of images, HLoG is comparable to gLoG and outperforms the other two detectors on F -score across the change of d . On the second set of images, HLoG outperforms the other three detectors on F -score across the change of d . We conclude HLoG in general outperforms the three detectors regardless of the d value.

D. Discussion on Computational Cost

The proposed method was programmed in MATLAB 2012b on a Windows (Microsoft, Inc.) platform, and the experiments are done on a Windows PC with Intel Xeon 2.0 GHz CPU and 32 GB of memory. For the 600×800 pathologic images, the average time cost is about 10.0 s/image compared to 30 s/image for the gLoG algorithm [11]. The time cost spent on $200 \times 256 \times 256$ fluorescence-light microscopy images of cells is 1.2 s/image on average, compared to 10.0 s/image for the gLoG algorithm. This shows that our algorithm is efficient for 2-D images and may be extended to 3-D grey images. Also, the LoG

transformation could be replaced by DoG approximation to improve the computational time [17].

IV. CONCLUSION

In this paper, we propose a novel imaging detector, termed HLoG to identify small blobs in medical images. After the raw image is transformed into normalized LoG space, an optimum scale is automatically determined based on the Hessian analysis. The blob candidates are also populated with their geometric shapes as the result of Hessian presegmentation. This process allows us to extract multiple regional features to characterize the accurate regional properties of small blobs. Three regional features, the average intensity feature, the regional likelihood of blobness, and the regional structure, are extracted and used in a tuning-free, VBGM to prune the presegmentation results. One set of pathologic images (15) and one set of fluorescence-light microscopy images (200) are used to compare HLoG with gLoG, Radial-Symmetry and LoG using recall, precision, and F -score metrics. In the experiments when d is estimated based on the size of the blobs, we observe HLoG outperforms Radial-Symmetry, LoG on both datasets, outperforms gLoG on the second dataset but with comparable performance for the first dataset for the precision metric. On the recall metric, HLoG only outperforms LoG on both datasets, but comparable to gLoG and Radial-Symmetry. In exploring the impact of d on the performance evaluation (F -score), we observe that HLoG outperforms Radial-Symmetry, LoG on both datasets. In addition, HLoG is computationally efficient and provides a tuning-free pruning process with only one parameter, the normalizing factor γ that needs to be specified.

ACKNOWLEDGMENT

The authors would like to thank Dr. H. Kong for providing gLoG source code and results.

REFERENCES

- [1] Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *IEEE Trans Biomed. Eng.*, vol. 57, no. 4, pp. 841–852, Apr. 2010.
- [2] N. Harder, B. Neumann, M. Held, U. Liebel, H. Erfle, J. Ellenberg, R. Eils, and K. Rohr, "Automated recognition of mitotic patterns in fluorescence microscopy images of human cells," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2006, pp. 1016–1019.
- [3] C. Xiaowei, Z. Xiaobo, and S. T. C. Wong, "Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 4, pp. 762–766, Apr. 2006.
- [4] A. Frangi, W. Niessen, K. Vincken, and M. Viergever, "Multiscale vessel enhancement filtering," in *Proc. Med. Image Comput. Comput.-Assisted Intervention Conf.*, vol. 1496, 1998, pp. 130–137.
- [5] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig, "A brain tumor segmentation framework based on outlier detection," *Med. Image Anal.*, vol. 8, no. 3, pp. 275–283, 2004.
- [6] M. Woo Kyung, S. Yi-Wei, B. Min Sun, H. Chiun-Sheng, C. Jeon-Hor, and C. Ruey-Feng, "Computer-aided tumor detection based on multi-scale blob detection algorithm in automated breast ultrasound images," *IEEE Trans. Med. Imag.*, vol. 32, no. 7, pp. 1191–1200, Jul. 2013.
- [7] S. C. Beeman, M. Zhang, L. Gubhaju, T. Wu, J. F. Bertram, D. H. Frakes, B. R. Cherry, and K. M. Bennett, "Measuring glomerular number and size in perfused kidneys using MRI," *Amer. J. Physiol. Renal Physiol.*, vol. 300, no. 6, pp. F1454–F1457, 2011.
- [8] S. C. Beeman, L. Cullen-McEwen, M. Zhang, T. Wu, E. J. Baldeolmar, J. Dowling, J. R. Charlton, M. S. Forbes, A. Ng, Q.-Z. Wu, J. A. Armitage, V. G. Puelles, G. F. Egan, J. F. Bertram, and K. M. Bennett, "Measuring human glomerular morphology and pathology with MRI," presented at the International Society for Magnetic Resonance in Medicine, Milan, Italy, 2014.
- [9] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 959–973, Aug. 2003.
- [10] G. Takacs, V. Chandrasekhar, S. S. Tsai, D. Chen, R. Grzeszczuk, and B. Girod, "Fast computation of rotation-invariant image features by an approximate radial gradient transform," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 2970–2982, Aug. 2013.
- [11] H. Kong, H. C. Akakin, and S. E. Sarma, "A generalized Laplacian of Gaussian filter for blob detection and its applications," *IEEE Trans. Cybern.*, vol. 43, no. 6, pp. 1719–1733, Dec. 2013.
- [12] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [13] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [14] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *Proc. IEEE Int. Comput. Vis.*, 2011, pp. 2548–2555.
- [15] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *Int. J. Comput. Vis.*, vol. 60, no. 1, pp. 63–86, 2004.
- [16] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool, "A comparison of affine region detectors," *Int. J. Comput. Vis.*, vol. 65, no. 1/2, pp. 43–72, 2005.
- [17] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.
- [18] A. P. Witkin, "Scale-space filtering: A new approach to multi-scale description," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 9, 1984, pp. 150–153.
- [19] J. Koenderink, "The structure of images," *Biol. Cybern.*, vol. 50, no. 5, pp. 363–370, 1984.
- [20] T. Lindeberg, "Scale-space theory: A basic tool for analyzing structures at different scales," *J. Appl. Statist.*, vol. 21, no. 1/2, pp. 225–270, 1994.
- [21] J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda, "Uniqueness of the Gaussian kernel for scale-space filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 1, pp. 26–33, Jan. 1986.
- [22] T. Lindeberg, "Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention," *Int. J. Comput. Vis.*, vol. 11, no. 3, pp. 283–318, 1993.
- [23] T. Lindeberg, "Feature detection with automatic scale selection," *Int. J. Comput. Vis.*, vol. 30, no. 2, pp. 79–116, 1998.
- [24] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: A survey," *Found. Trends Comput. Graph. Vis.*, vol. 3, no. 3, pp. 177–280, 2008.
- [25] E. Bernardis and S. X. Yu, "Pop out many small structures from a very large microscopic image," *Med. Image Anal.*, vol. 15, no. 5, pp. 690–707, 2011.
- [26] A. A. Dima, J. T. Elliott, J. J. Filliben, M. Halter, A. Peskin, J. Bernal, M. Kocielek, M. C. Brady, H. C. Tang, and A. L. Plant, "Comparison of segmentation algorithms for fluorescence microscopy images of cells," *Cytometry Part A*, vol. 79A, no. 7, pp. 545–559, 2011.
- [27] S. Di Cataldo, E. Ficarra, A. Acquaviva, and E. Macii, "Automated segmentation of tissue images for computerized IHC analysis," *Comput. Methods Programs Biomed.*, vol. 100, no. 1, pp. 1–15, 2010.
- [28] S. C. Beeman, J. F. Georges, and K. M. Bennett, "Toxicity, biodistribution, and ex vivo MRI detection of intravenously injected cationized ferritin," *Magn. Reson. Med.*, vol. 69, no. 3, pp. 853–861, 2013.
- [29] K. M. Bennett, J. F. Bertram, S. C. Beeman, and N. Gretz, "The emerging role of MRI in quantitative renal glomerular morphology," *Amer. J. Physiol. Renal Physiol.*, vol. 304, no. 10, pp. F1252–F1257, 2013.
- [30] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
- [31] A. H. Salden, B. M. T. H. Romeny, M. A. Viergever, L. M. J. Florack, and J. J. Koenderink, "Differential geometric description of 3D scalar images," *Internal Report 3DCV*, vol. 91–05, 1991.
- [32] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. 4th Alvey Vis. Conf.*, 1988, pp. 147–151.
- [33] V. Lempitsky and A. Zisserman, "Learning to count objects in images," in *Proc. Adv. Neural Inf. Process. Syst.*, 2010, pp. 1324–1332.

Authors' photographs and biographies not available at the time of publication.