

MEANINGFUL LEVEL SETS FOR SMALL SPOT DETECTION

Axel Davy

HGH Infrared Systems

ABSTRACT

Detecting spots is a recurring problem in image processing, with applications in biological imaging, infrared target tracking and industrial inspection. It involves accurately identifying and localizing small circular or blob-like structures in images. This paper presents a novel approach based on a unified statistical threshold for the detection of spots of various shapes and sizes. Meaningful level sets are extracted from the Max-Tree hierarchy and filtered using an a contrario model for robust, unsupervised detection. We propose two implementations of this algorithm and demonstrate competitive performance on biological and infrared datasets against classical and deep learning approaches.

Index Terms— Spot detection, level lines, a contrario, image processing, biological imaging

1. INTRODUCTION

The accurate identification and localization of small circular or blob-like structures in images, commonly known as spot detection, represents a fundamental challenge in image processing with applications spanning diverse domains. In fluorescence microscopy [1], the goal is to track diffraction-limited fluorescent particles amidst noisy backgrounds. Infrared surveillance systems [2] require detecting dim targets against cluttered scenes. Industrial quality control demands identifying defects of varying contrast and size. The challenge lies in the significant variance across application domains and scenes, where traditional techniques typically require extensive parameter tuning for each specific application, limiting their generalizability. This has motivated a trend toward reformulating the problem as object segmentation or detection, relying on annotated datasets and deep learning approaches [3, 4]. However, due to the reliance on training datasets, domain-specific biases, and computational cost, classical unsupervised methods remain relevant as lightweight, interpretable alternatives that can serve as preprocessing stages for heavier learning-based techniques or as standalone solutions when training data is unavailable.

Classical approaches to spot detection include template-based methods [5], Hough Circle Transform [6], wavelet-based approaches [7], morphological methods [8], and local contrast methods [9, 10]. These methods typically require careful parameter tuning and struggle with varying spot sizes. Recent deep learning methods [3, 4, 11] achieve strong performance but require substantial annotated datasets and may not generalize across domains without retraining.

We present a novel approach that unifies spot detection across scales by analyzing local contrast between consecutive level sets in the Max-Tree hierarchy. Our key insight is that spots of any size can be detected through a single principled framework by adapting detection thresholds to region size using polyomino counting

from combinatorial geometry. We propose an a contrario model that measures the significance of contrast between parent and child level sets, extending ideas from edge detection [12] and denoising [13] to spot detection. Unlike template matching that assumes specific spot shape and scale, or background subtraction that models global structure, our method directly exploits the hierarchical organization of image level sets to detect local intensity anomalies. The method provides several advantages: all scales are handled in a single pass and detection thresholds are derived from statistical significance rather than heuristics. The detection criterion naturally captures the intuition that large but low-contrast spots should be as meaningful as high-contrast small ones. The unified threshold framework adapts from high thresholds for small regions to low thresholds for larger spots with a single detection parameter ϵ (Table 1). Finally, the method achieves real-time performance through efficient Max-Tree construction algorithms [14].

Our contributions are: (1) a novel a contrario model for scale-adaptive spot detection based on level set contrast analysis; (2) dual implementation strategies balancing completeness (full Max-Tree) versus real-time performance (seed-based growth); (3) comprehensive evaluation on biological and infrared datasets demonstrating competitive performance against classical and learning-based methods. We validate our approach on four Cell Tracking Challenge datasets [15] and IRSTD-1k [2], an infrared dataset, comparing against classical methods (à trous wavelet [7], DoG, DoH, LoG, MSER) and the deep learning baseline DeepBlink [3].

2. RELATED WORK

Our method builds on three research threads: hierarchical image representations through Max-Trees, a contrario detection frameworks for principled threshold selection, and classical spot detection baselines. We review each in turn, emphasizing connections to our approach.

2.1. Max-Tree computation

The Max-Tree is a hierarchical representation encoding the connected components of all upper level sets of an image [16]. Each node represents a connected component at a given intensity threshold, with edges capturing spatial inclusion relationships. Efficient construction algorithms [14] achieve quasi-linear to linear complexity in the number of pixels, and have been extended to parallel and GPU implementations [17].

2.2. A contrario detection framework and its application to level sets

The a contrario framework [12] provides principled detection by defining a null hypothesis modeling structure absence and computing the Number of False Alarms (NFA) for candidates. Structures

The author thanks the reviewers for their constructive feedback.

with NFA below a threshold are deemed meaningful detections. This framework has proven particularly powerful for analyzing level set hierarchies.

Desolneux et al. [12] pioneered a contrario edge detection by analyzing gradient contrast along level lines. They compute an NFA based on the length of a level line segment and the minimum gradient magnitude along it, and detect meaningful level lines based on an image gradient distribution model. Their method establishes that level line contrast can be meaningfully measured against statistical models. It was extended by [18] for meaningful curve extraction. While conceptually this approach appears close to ours, significant changes can be noticed: we estimate the level line contrast with the difference between parents and children in the Max-Tree, which reduces computation and is better defined for very small spots. We model the difference between level sets, which gives less weight to large level sets than modeling the gradient distribution on the image. Furthermore, we use the level set number of pixels rather than its contour length, giving the same detection score to elongated or circular spots.

Couplier and Desolneux [13] apply the a contrario framework to image denoising through statistical area thresholding, removing level sets whose area is not statistically significant given their contrast. They threshold the image at all intensity levels and model them as Bernoulli noise. They compute the probability that a connected region has a given size under this model. While in our work we also compute an NFA for a blob of pixels (a level set), we do not model the level set as Bernoulli noise, but take the contrast with the next level set into account, and model it with an exponential distribution.

Spot detection is related to anomaly detection. Ehret [19] provides a comprehensive review of classical anomaly detection methods, including methods detecting anomalies as spots in background-subtracted images modeled as Gaussian noise. Several a contrario models are presented for anomaly detection. However, anomaly detection differs from spot detection in the sense that spots are not anomalies for some application domains, such as biology.

2.3. Other detection methods based on Max-Trees or Tree of Shapes

Level set hierarchies [16, 20] provide a natural framework for detection by capturing image structure at all scales simultaneously. Maximally Stable Extremal Regions (MSER) [21] identifies regions in the component tree that persist across a range of intensity thresholds, optimizing for stability rather than the statistical significance of local contrast; it involves no noise model and cannot distinguish weak spots from background structures of similar size. Linear-time MSER implementations [22] demonstrate efficient tree traversal strategies, making MSER a popular choice for real-time applications. The grain filter [23] removes small connected components based on area thresholding, representing size-based filtering without statistical modeling.

2.4. Classical spot detection baselines

Classical spot detection complements level-set approaches with methods based on different principles. The Laplacian of Gaussian (LoG) [5] detects blobs as local maxima of the scale-normalized Laplacian response and is widely adopted in biological imaging tools such as TrackMate [24]. The Difference of Gaussians (DoG) [25] computes the difference between two Gaussian-blurred versions of an image at adjacent scales and underpins feature detectors such as SIFT. The Determinant of Hessian (DoH) [5] responds to blob-like

structures via the Hessian matrix determinant. The Hough Circle Transform (HCT) [6] votes in a parameter space of center coordinates and radius, assuming circular spot shapes. Wavelet-based multiscale methods [7] compute products across à trous wavelet scales to enhance spots while suppressing noise. Local contrast methods [9], including center-surround differences and coefficient of variation measures, have been developed for infrared small target detection, assuming spots appear as local intensity anomalies. Tensor-based methods [26] and optimized top-hat transforms [8] extend these ideas with more sophisticated local neighborhood analysis.

While effective in their design regimes, these methods require parameter tuning (scale ranges, window sizes, thresholds) and may struggle with spots whose size or contrast violates design assumptions.

3. PROPOSED METHOD

We detect spots as statistically significant contrasts in the Max-Tree hierarchy. The key innovation is an a contrario model using polyomino counting that adapts detection thresholds to region size.

3.1. Max-Tree and level set contrasts

The Max-Tree [16] is a hierarchical representation where each node n corresponds to a connected component (under 4-connectivity) at some intensity level V_n , with edges encoding spatial inclusion. The root node represents the entire image domain at the lowest intensity, and leaves correspond to local maxima. As discussed in Section 2.1, efficient algorithms [14] construct this tree in quasi-linear time.

As spots are typically defined as a brighter region against a darker background (or vice versa for dark spots), spots are nodes in the Max-Tree (or the Min-Tree for dark spots). In the following we will focus on bright spots, and consider that dark spots can be detected symmetrically using the Min-Tree.

Our detection examines parent-child contrasts in the Max-Tree hierarchy. For a level set node n containing N pixels at intensity V_n , with parent p at intensity $V_p < V_n$, the contrast is $\Delta V = V_n - V_p$. We test whether this contrast is statistically significant given the region size.

3.2. Detecting groups of pixels in noise

Let us assume a set of N contiguous pixels (in terms of 4-connectivity) is observed in an image with i.i.d. pixel intensity distribution $\mathcal{E}$ of parameter λ and of survival function $S_{\mathcal{E}(\lambda)}$. How significant is it to observe that all values in such a group of pixels are above an intensity value V ?

The probability of a single pixel to have a value above a positive value V in such a noise image is given by:

$$p_V = P(X \geq V) = S_{\mathcal{E}(\lambda)}(V). \quad (1)$$

For an exponential distribution of parameter λ , the survival function is $S_{\mathcal{E}(\lambda)}(x) = e^{-\lambda(x)}$.

The probability of observing N pixels in a predetermined group all above V is then:

$$P_N = p_V^N. \quad (2)$$

Critically, we do not assume spots have predetermined shapes. Groups of N contiguous pixels are known as polyominoes in combinatorial geometry. The number of distinct polyominoes of size N

Table 1. Detection threshold on V/σ for different region sizes N with $M = 1024^2$ pixels and $\epsilon = 1$.

N	1	2	3	4	5	6	7	8
V/σ	14.12	7.41	5.27	4.23	3.62	3.22	2.94	2.73
N	9	10	15	20	25	30	40	50
V/σ	2.57	2.44	2.07	1.89	1.78	1.71	1.63	1.58

is known to grow approximately as [27, 28]:

$$C_N \approx \alpha \frac{\beta^N}{N}, \quad \alpha = 0.316915, \beta = 4.062570. \quad (3)$$

The expected number of occurrences of such groups of N pixels all above V in an image of size M pixels is then approximated by:

$$E_N = M C_N P_N \approx M \alpha \frac{\beta^N}{N} p_V^N. \quad (4)$$

Or equivalently to get an expected number of occurrences below a target ϵ one needs to threshold on V such that:

$$V > S_{\mathcal{E}(\lambda)}^{-1} \left(\left(\frac{\epsilon N}{M \alpha \beta^N} \right)^{\frac{1}{N}} \right). \quad (5)$$

For the exponential distribution, this simplifies to

$$V > \frac{\ln(\beta)}{\lambda} + \frac{1}{\lambda N} \ln \left(\frac{\alpha M}{\epsilon N} \right). \quad (6)$$

Setting $\sigma = \frac{1}{\lambda}$ the standard deviation of the distribution, we get

$$\frac{V}{\sigma} > \ln(\beta) + \frac{1}{N} \ln \left(\frac{\alpha M}{\epsilon N} \right). \quad (7)$$

This formulation shows that the required contrast threshold adapts automatically to the size of the detected region: smaller regions require higher contrast to be deemed meaningful, while larger regions can be detected with lower contrast values.

Table 1 provides detection thresholds V/σ for different region sizes N in an image of size $M = 1024^2$ pixels, with a target expected number of occurrences $\epsilon = 1$.

3.3. A contrario model for spot detection

We propose as null hypothesis $\mathcal{H}_0$, that the contrast ΔV between a child level set and its parent arises from noise following an exponential distribution with standard deviation σ : $\Delta V \sim \mathcal{E}(1/\sigma)$. This model, despite being approximate, captures the contrast distributions we observed empirically in biological and infrared imagery.

Applying the model from Section 3.2, we compute the NFA for each level set in the Max-Tree based on its contrast ΔV and size N .

$$\text{NFA}(n) = M \cdot C_N \cdot p_{\Delta V}^N = M \alpha \frac{\beta^N}{N} e^{-N \Delta V / \sigma}, \quad (8)$$

where M , the number of pixels, serves as an upper bound on the number of tests. $\mathbb{E}[\text{NFA}]$ gives an upper bound on the expected number of occurrences of such level sets under $\mathcal{H}_0$.

A level set is a *meaningful spot* if

$$\text{NFA}(n) < \epsilon, \quad (9)$$

where ϵ controls expected false alarms, typically $\epsilon = 1$. As is usual for a contrario methods, lower ϵ (e.g. 10^{-2}) can be set to account for the slight $\mathcal{H}_0$ mismatch caused by local pixel correlations in natural images.

This formulation reveals automatic scale adaptation: small regions require high contrast ($\approx 14\sigma$ for $N = 1$, $M = 1024^2$, $\epsilon = 1$),

while large regions require modest contrast ($\approx 1.6\sigma$ for $N = 50$), as shown in Table 1. Unlike template-based methods requiring scale parameters or multi-scale search, our unified threshold adapts naturally through Equation (9).

3.4. Implementing the spot detector

As this method is designed for real-time operation, potentially as a preprocessing step before more resource-intensive methods, implementation considerations are critical.

Max-Tree implementation. One implementation constructs the complete Max-Tree and applies detection thresholds at each node. Efficient Max-Tree construction algorithms [14] achieve quasi-linear complexity in the number of pixels. Since the tree typically contains fewer nodes than pixels, threshold evaluation is inexpensive relative to tree construction. To prevent duplicate detections, nodes whose parent is already detected are filtered out. Spot size can be constrained using minimum and maximum bounds on level set area.

To accelerate NFA computation and improve numerical stability, we compute the logarithm of the NFA, which simplifies to:

$$\log(\text{NFA}(n)) = \log(M) + \log(\alpha) - \log(N) + N \log(\beta) - \frac{N \Delta V}{\sigma}. \quad (10)$$

This expression can be efficiently tabulated for small N values.

Fast seed-based implementation. When targeting small spots specifically, constructing the complete Max-Tree can be unnecessary. We propose an alternative seed-based implementation optimized for this scenario. Essentially, it involves growing level sets from local maxima seeds and evaluating the NFA at each step until the maximum allowed spot size is reached, without constructing the full Max-Tree.

1. Apply a maximum filter to identify local maxima as seeds (5×5 kernel).
2. For each seed pixel at position (y_0, x_0) and intensity V_0 :
 - (a) Initialize the current level set at intensity V_0 containing (y_0, x_0) .
 - (b) Initialize a neighborhood queue with 4-connected neighbors.
 - (c) Iteratively grow the level set: dequeue the neighbor with maximum intensity $V_{\text{next}} < V_0$, merge all queued neighbors with intensity $\geq V_{\text{next}}$ into the level set, and add their unexplored neighbors to the queue.
 - (d) At each step, evaluate the NFA of the current level set.
 - (e) Stop when reaching size N_{max} or when V_{next} reaches the image minimum.
3. If any level set from a seed meets the detection criterion, record the largest meaningful one.

This approach explores level sets containing each seed pixel in order of decreasing intensity, evaluating NFA at each step. A deduplication mechanism tracks which levels have been processed for each pixel to avoid redundant computation when seeds share level sets.

Unlike the Max-Tree approach, this seed-based method does not have linear complexity in the number of pixels. However, execution is faster in practice for small N_{max} values (≤ 15 in our experiments). In a practical implementation further speedups can be achieved, as a more conservative seed selection criterion (e.g., local contrast thresholds) could be applied to reduce the number of seeds.

In addition, large $N_{\max}$ values can be handled by running this algorithm on multiple scales of an image pyramid with small $N_{\max}$ (although with a loss of precision for the generated mask).

The implementations of both methods are single-threaded but could be made multi-threaded for improved performance.

4. EXPERIMENTAL RESULTS

We evaluate our method on four datasets from the Cell Tracking Challenge [15], as well as IRSTD-1k [2].

4.1. Datasets and evaluation protocol

We use five datasets spanning diverse imaging modalities and spot characteristics:

- **BF-C2DL-HSC** [15]. Mouse hematopoietic stem cells in hydrogel microwells, $N = 1764$ 1010×1010 test images.
- **Fluo-N2DH-GOWT1** [15]. GFP-GOWT1 mouse stem cells, $N = 92$ 1024×1024 test images.
- **Fluo-N2DL-HeLa** [15]. HeLa cells stably expressing H2b-GFP, $N = 92$ 1100×700 test images.
- **PhC-C2DL-PSC** [15]. Pancreatic stem cells on a polystyrene substrate, $N = 300$ 720×576 test images.
- **IRSTD-1k** [2]. Infrared small target detection, $N = 201$ 512×512 test images.

These datasets exhibit varying noise levels, spot shapes and background complexities. For fair comparison of methods, range normalization was performed on *GOWT1* and *HeLa*, and background subtraction on *HSC*. Further details can be found in the code repository.

Detection performance is measured using Precision ($P = TP/(TP + FP)$), Recall ($R = TP/(TP + FN)$) and Critical Success Index ($CSI = TP/(TP + FP + FN)$). A detection is considered correct (true positive) if it falls within distance $d = 5$ pixels of a ground truth location; this matching distance is held constant across all methods and datasets. For each method and dataset, we tune parameters on separate training subsets using grid search and coordinate descent on predetermined parameter ranges to maximize CSI, then evaluate on the test sets. As a result, each method uses a different set of optimized parameters per dataset.

4.2. Baseline methods

We compare against seven established methods spanning different detection paradigms:

À trous wavelet: Multiscale product method [7] computing products across wavelet scales to enhance spots while suppressing noise.

DeepBlink [3]: Deep learning baseline (version 0.1.4) trained on biological imaging data.

LoG, DoG, DoH: Laplacian of Gaussian, Difference of Gaussians, and Determinant of Hessian blob detectors from scikit-image 0.26.0 [29].

HCT [6]: Circle detection via Hough transform voting (OpenCV 4.12.0.88 [30]).

MSER [21]: Maximally Stable Extremal Regions detector (OpenCV 4.12.0.88) identifying regions stable across threshold ranges.

Ours: Our a contrario method with parameters for spot size range, noise level estimation, and the NFA threshold ϵ .

The code to reproduce the experiments, including tuned parameter ranges for each method, can be found at <https://github.com/axeldavy/MeaningfulSpot>.

4.3. Results and Discussion

Detection performance. Table 2 presents test set performance across all datasets, while Figure 1 provides qualitative examples of detections across methods and datasets. Across all datasets, no clear overall winner emerges. Our method ranks first on *GOWT1*, second on *PSC* and *IRSTD*, third on *HeLa* and fourth on *HSC* by CSI. Notably, every competing method ranks poorly on at least one dataset: À trous wavelet fails on *GOWT1* (CSI 0.008), DeepBlink on *HeLa* (CSI 0.036), and LoG and DoG on *IRSTD* (CSI 0.068 and 0.213), showing that limited generalization is a shared challenge. The performance on *GOWT1* highlights the ability to detect large spot regions even in the presence of noise. On that dataset, though, it struggles with low contrast spots, as do all compared methods. *IRSTD* shows the performance of spot detection even in the presence of non-spot-like clutter. On *HeLa* and *PSC*, the method demonstrates its effectiveness in detecting irregular-shaped spots. On these datasets, putting aside the detections of spots with missing annotations on borders, the method’s errors consist of incorrectly grouped detections, missed low-contrast detections on *HeLa*, and false detections on the background on *PSC*. *HSC* is a challenge for our method, because despite being regularly shaped spots, the cells have a bright center and a dark border, which does not fit well with the model. The large gap between our method and MSER illustrates the benefit of statistical significance over stability: MSER cannot distinguish low-contrast spots from slowly varying background lacking local contrast, whereas our NFA criterion flags them as meaningful as long as their contrast is statistically significant.

Parameter sensitivity. We did our best to homogenize datasets to compare methods fairly, and optimize parameters for each method and dataset using an automated procedure. For instance, background subtraction was applied to *HSC* to suppress structured background features not of interest. Without it, several methods, including ours, raised false detections on the background. This highlights the value of domain-specific preprocessing even for unsupervised methods, and the need for approaches with easily interpretable parameters. More advanced pre- and post-processing, as well as hand-tuning, could yield improved performance for all methods.

Computational efficiency. Table 3 reports average execution times per image across all datasets, on a single CPU core (Intel i5-4690). Note some of the compared implementations may not be optimized for speed, and that further gains could be achieved with multithreading and/or GPU acceleration. Our method achieves 64ms per image on average, ranking third after the Hough Circle Transform (13ms) and MSER (42ms). This real-time performance suits applications requiring fast processing or as preprocessing before more computationally intensive methods.

Limitations. The spot model of our method is resilient to some degree of background structure, but not entirely. It is more sensitive than other methods to elongated structures. Low thresholds could trigger false detection for any large coherent background structure. Furthermore, for very large regions the model for the intensity difference between level lines is less accurate, making it difficult to have a single threshold that works well across all scales. This is why, for the experiments, both a minimum and maximum spot size were among the tuned parameters. In addition, one use case of spot detector methods is as a preprocessing step before applying a heavier method (for instance a classification network). This evaluation fo-

Table 2. Quantitative results on the test sets of each dataset. Each cell comprises the Precision (P), Recall (R), and Critical Success Index (CSI). Best results per dataset are in **bold**.

Method	BF-C2DL-HSC P — R — CSI	Fluo-N2DH-GOWT1 P — R — CSI	Fluo-N2DL-HeLa P — R — CSI	PhC-C2DL-PSC P — R — CSI	IRSTD-1k P — R — CSI
À trous wavelet [7]	0.924 — 0.893 — 0.832	0.009 — 0.125 — 0.008	0.823 — 0.713 — 0.618	0.740 — 0.921 — 0.696	0.337 — 0.562 — 0.267
DeepBlink [3]	0.954 — 0.277 — 0.274	0.013 — 0.438 — 0.013	0.041 — 0.236 — 0.036	0.505 — 0.763 — 0.437	0.614 — 0.508 — 0.385
Difference of Gaussians [25]	0.863 — 0.756 — 0.675	0.236 — 0.718 — 0.216	0.744 — 0.907 — 0.691	0.635 — 0.912 — 0.599	0.418 — 0.310 — 0.216
Determinant of Hessians [5]	0.833 — 0.695 — 0.610	0.599 — 0.317 — 0.261	0.777 — 0.525 — 0.456	0.687 — 0.602 — 0.472	0.178 — 0.414 — 0.142
Hough Circle Transform [6]	0.659 — 0.701 — 0.514	0.460 — 0.373 — 0.260	0.723 — 0.627 — 0.506	0.196 — 0.434 — 0.156	0.041 — 0.266 — 0.037
Laplacian of Gaussian [5]	0.898 — 0.941 — 0.850	0.527 — 0.737 — 0.444	0.758 — 0.883 — 0.689	0.698 — 0.778 — 0.582	0.071 — 0.572 — 0.068
MSER [21]	0.298 — 0.261 — 0.162	0.493 — 0.380 — 0.273	0.204 — 0.664 — 0.185	0.127 — 0.703 — 0.120	0.041 — 0.185 — 0.035
Ours	0.927 — 0.645 — 0.614	0.936 — 0.843 — 0.797	0.843 — 0.708 — 0.626	0.690 — 0.846 — 0.613	0.404 — 0.444 — 0.268

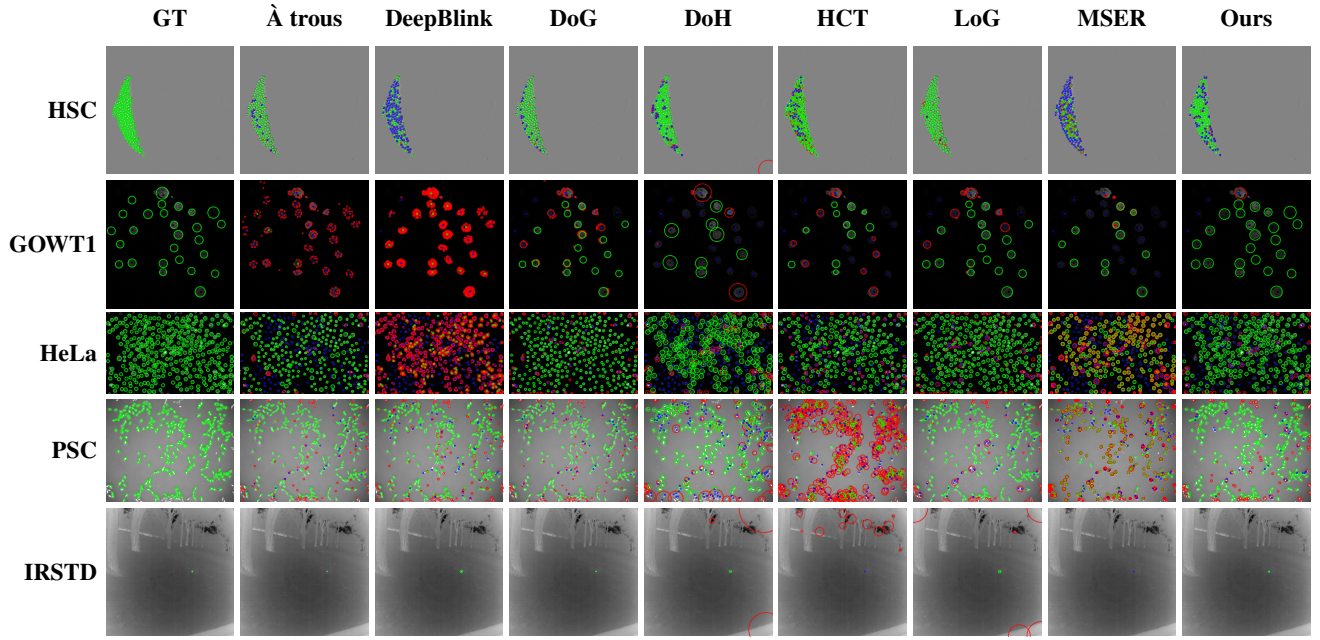


Fig. 1. Detection results comparison across datasets and methods. Correct detection locations are shown in green circles, false positives in red circles and false negatives in blue crosses.

Table 3. Average execution time per image in milliseconds across datasets on a single CPU core (Intel i5-4690).

Method	HCT	MSER	Ours	À trous	DoG	DoH	LoG	DeepBlink
Time (ms)	13	42	64	150	350	520	1000	5100

cuses on detection accuracy only, without considering downstream tasks. The low parameter count and real-time performance of our method make it well suited for such scenarios. Recall can be prioritized by adjusting ϵ to favor sensitivity over precision.

5. CONCLUSION

We presented a novel spot detection method that unifies detection across scales through an a contrario framework applied to Max-Tree hierarchies. By adapting polyomino counting from combinatorial geometry to level set contrast analysis, our approach automatically

adjusts detection thresholds based on region size. This eliminates the need for multi-scale search procedures required by traditional methods. Two complementary implementations are provided: a full Max-Tree approach with quasi-linear complexity, and a seed-based variant optimized for small spot detection scenarios.

Experimental validation on five diverse datasets spanning biological microscopy and infrared surveillance demonstrates competitive performance against seven established baselines, including classical methods (LoG, DoG, DoH, HCT, MSER, à trous wavelet) and the deep learning approach DeepBlink.

The low computational cost and interpretable parameters (a false alarm budget, a noise level estimate, and minimum and maximum spot size bounds) position this approach as an effective preprocessing stage for downstream analysis pipelines or as a standalone detector when training data is unavailable. Future work could explore GPU implementations, use as preprocessing for downstream tasks such as classification or tracking, and compactness-based post-filtering to further improve precision.

6. REFERENCES

- [1] Ivo F Sbalzarini and Petros Koumoutsakos, “Feature point tracking and trajectory analysis for video imaging in cell biology,” *Journal of structural biology*, vol. 151, no. 2, pp. 182–195, 2005.
- [2] Mingjin Zhang, Rui Zhang, Yuxiang Yang, Haichen Bai, Jing Zhang, and Jie Guo, “Isnet: Shape matters for infrared small target detection,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 877–886.
- [3] Bastian Th Eichenberger, YinXiu Zhan, Markus Rempfler, Luca Giorgetti, and Jeffrey A Chao, “deepblink: threshold-independent detection and localization of diffraction-limited spots,” *Nucleic Acids Research*, vol. 49, no. 13, pp. 7292–7297, 2021.
- [4] Uwe Schmidt, Martin Weigert, Coleman Broaddus, and Gene Myers, “Cell detection with star-convex polygons,” in *International conference on medical image computing and computer-assisted intervention*. Springer, 2018, pp. 265–273.
- [5] Tony Lindeberg, “Feature detection with automatic scale selection,” *International journal of computer vision*, vol. 30, no. 2, pp. 79–116, 1998.
- [6] HK Yuen, John Princen, John Illingworth, and Josef Kittler, “Comparative study of hough transform methods for circle finding,” *Image and vision computing*, vol. 8, no. 1, pp. 71–77, 1990.
- [7] Jean-Christophe Olivo-Marin, “Extraction of spots in biological images using multiscale products,” *Pattern recognition*, vol. 35, no. 9, pp. 1989–1996, 2002.
- [8] Xiangzhi Bai and Fugen Zhou, “Analysis of different modified top-hat transformations based on structuring element construction,” *Signal processing*, vol. 90, no. 11, pp. 2999–3003, 2010.
- [9] Junmin Rao, Jing Mu, Fanming Li, and Shijian Liu, “Infrared small target detection based on weighted local coefficient of variation measure,” *Sensors*, vol. 22, no. 9, pp. 3462, 2022.
- [10] Hong Zhang, Lei Zhang, Ding Yuan, and Hao Chen, “Infrared small target detection based on local intensity and gradient properties,” *Infrared Physics & Technology*, vol. 89, pp. 88–96, 2018.
- [11] Carsen Stringer, Tim Wang, Michalis Michaelos, and Marius Pachitariu, “Cellpose: a generalist algorithm for cellular segmentation,” *Nature methods*, vol. 18, no. 1, pp. 100–106, 2021.
- [12] Agnes Desolneux, Lionel Moisan, and Jean-Michel Morel, “Edge detection by helmholtz principle,” *Journal of mathematical imaging and vision*, vol. 14, no. 3, pp. 271–284, 2001.
- [13] David Coupier, Agnès Desolneux, and Bernard Ycart, “Image denoising by statistical area thresholding,” *Journal of Mathematical Imaging and Vision*, vol. 22, no. 2, pp. 183–197, 2005.
- [14] Edwin Carlinet and Thierry Géraud, “A comparative review of component tree computation algorithms,” *IEEE Transactions on Image Processing*, vol. 23, no. 9, pp. 3885–3895, 2014.
- [15] Martin Maška, Vladimír Ulman, Pablo Delgado-Rodriguez, Estibaliz Gómez-de Mariscal, Tereza Nečasová, Fidel A Guerrero Peña, Tsang Ing Ren, Elliot M Meyerowitz, Tim Scherr, Katharina Löffler, et al., “The cell tracking challenge: 10 years of objective benchmarking,” *Nature Methods*, vol. 20, no. 7, pp. 1010–1020, 2023.
- [16] Philippe Salembier, Albert Oliveras, and Luis Garrido, “Antiextensive connected operators for image and sequence processing,” *IEEE Transactions on Image Processing*, vol. 7, no. 4, pp. 555–570, 1998.
- [17] Nicolas Blin, Edwin Carlinet, Florian Lemaitre, Lionel Lacasagne, and Thierry Géraud, “Max-tree computation on gpus,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 12, pp. 3520–3531, 2022.
- [18] Frédéric Cao, Pablo Musé, and Frédéric Sur, “Extracting meaningful curves from images,” *Journal of Mathematical Imaging and Vision*, vol. 22, no. 2, pp. 159–181, 2005.
- [19] Thibaud Ehret, Axel Davy, Jean-Michel Morel, and Mauricio Delbracio, “Image anomalies: A review and synthesis of detection methods,” *Journal of Mathematical Imaging and Vision*, vol. 61, no. 5, pp. 710–743, 2019.
- [20] Pascal Monasse and Frederic Guichard, “Fast computation of a contrast-invariant image representation,” *IEEE transactions on image processing*, vol. 9, no. 5, pp. 860–872, 2000.
- [21] Jiri Matas, Ondrej Chum, Martin Urban, and Tomás Pajdla, “Robust wide-baseline stereo from maximally stable extremal regions,” *Image and vision computing*, vol. 22, no. 10, pp. 761–767, 2004.
- [22] David Nistér and Henrik Stewénus, “Linear time maximally stable extremal regions,” in *European conference on computer vision*. Springer, 2008, pp. 183–196.
- [23] Yongchao Xu, *Tree-based shape spaces: Definition and applications in image processing and computer vision*, Ph.D. thesis, Université Paris-Est, 2013.
- [24] Jean-Yves Tinevez, Nick Perry, Johannes Schindelin, Genevieve M Hoopes, Gregory D Reynolds, Emmanuel Laplantine, Sebastian Y Bednarek, Spencer L Shorte, and Kevin W Eliceiri, “Trackmate: An open and extensible platform for single-particle tracking,” *Methods*, vol. 115, pp. 80–90, 2017.
- [25] David G Lowe, “Distinctive image features from scale-invariant keypoints,” *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [26] Yimian Dai and Yiquan Wu, “Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection,” *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 10, no. 8, pp. 3752–3767, 2017.
- [27] Rafael Grompone von Gioi, Charles Hessel, Tristan Dagobert, Jean-Michel Morel, and Carlo de Franchis, “Ground visibility in satellite optical time series based on a contrario local image matching,” *Image Processing On Line*, vol. 11, pp. 212–233, 2021.
- [28] Iwan Jensen and Anthony J Guttmann, “Statistics of lattice animals (polyominoes) and polygons,” *Journal of Physics A: Mathematical and General*, vol. 33, no. 29, pp. L257, 2000.
- [29] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, Tony Yu, and the scikit-image contributors, “scikit-image: image processing in Python,” *PeerJ*, vol. 2, pp. e453, 6 2014.
- [30] G. Bradski, “The OpenCV Library,” *Dr. Dobb’s Journal of Software Tools*, 2000.