Analyst

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Analyst*, 2025, **150**, 3909

HIST-DIP: histogram thresholding and deep image priors assisted smartphone-based fluorescence microscopy imaging†

Harshitha Govindaraju, Muhammad Nabeel Tahir and Umer Hassan *

Portable fluorescence microscopes coupled with smartphones offer accessible and cost-effective point-of-care diagnostic solutions, but often produce noisy and blurry images with poor contrast. Here, we introduce HIST-DIP (HIStogram Thresholding and Deep Image Prior), an unsupervised framework for fluorescence microscopy image restoration. Histogram thresholding isolates fluorescence signals by removing background noise, while DIP refines structural details and enhances resolution without large labeled datasets. Validation results show substantial quality gains including the average Peak Signal-to-Noise Ratio (PSNR) improved from 15.59 dB to 27.10 dB, and the Structural Similarity Index Measure (SSIM) rose from 0.035 to 0.82. Contrast-to-noise ratio (CNR) and signal difference-to-noise ratio (SDNR) also increased significantly, indicating sharper bead outlines and reduced background interference. Unlike conventional deep learning methods, HIST-DIP needs no external training data, making it well-suited for real-time, low-cost, and point-of-care diagnostic imaging. These findings highlight the potential of HIST-DIP in enhancing the quality of smartphone-based microscopy images, while also motivating future research towards optimizing the methods for real-time on-device computations.

Received 30th April 2025,
Accepted 18th July 2025

DOI: 10.1039/d5an00487j

rsc.li/analyst

Open Access Article. Published on 23 July 2025. Downloaded on 5/4/2026 1:38:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

1. Introduction

The field of point-of-care testing (POCT) has evolved significantly over the past decade, driven by its affordability, rapid diagnostic capabilities, high sensitivity, and specificity. These systems enable non-experts to monitor health conditions without requiring bulky laboratory equipment or specialized technical skills, making them particularly valuable in resource-limited settings. Their importance was underscored during the COVID-19 pandemic, where the surge in demand for decentralized diagnostics highlighted the need for accessible, reliable, and efficient testing solutions. By reducing reliance on centralized laboratories, POCT minimizes delays, lowers the risk of misdiagnosis, and enhances clinical outcomes, ultimately strengthening healthcare resilience in times of crisis.[1–3] Among various POCT modalities, smartphone-based diagnostics have attracted particular attention because of the global proliferation of smartphones, their built-in high-resolution complementary metal–oxide–semiconductor (CMOS) cameras and their capacity for wireless data transfer and on-device analytics.[4–6]

Within this domain, smartphone-based fluorescence microscopes (SBFMs)[7] have emerged as vital tools for diverse applications spanning microbiology, pathology, and environmental monitoring. Their compact design and reduced cost facilitate decentralized diagnostics, enabling clinicians and researchers to conduct tests in resource-limited regions. However, the broader adoption of SBFMs is hampered by fundamental image quality challenges–motion blur, poor focus, low resolution, and color inconsistencies–stemming from the optical constraints of smartphone cameras, variations in illumination, sensor spectral responses, and disparate post-processing pipelines.[8,9] These hardware and software limitations can substantially undermine diagnostic accuracy, necessitating robust computational strategies to enhance image clarity and consistency.

Traditional laboratory-grade microscopes circumvent these obstacles through sophisticated optical components and advanced image-processing algorithms, but such solutions are impractical for mobile microscopy due to cost, portability, and real-time usability requirements. While deep learning has unlocked powerful image enhancement methods–ranging from denoising and deblurring to super-resolution[10,11]–these approaches typically rely on large-scale, labeled training datasets that are often unavailable for smartphone-based systems. The Deep Image Prior (DIP) framework[12] offers a compelling alternative by exploiting the inherent structure of

Rutgers, The State University of New Jersey, Dept of Electrical and Computer Engineering, 94 Bret Rd, Piscataway, 08854 USA. E-mail: umer.hassan@rutgers.edu

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d5an00487j

Open Access Article. Published on 23 July 2025. Downloaded on 5/4/2026 1:38:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

convolutional neural networks (CNNs) to restore degraded images without explicit external training data. Despite its promise, DIP remains underexplored for mobile microscopy, where constraints on hardware, illumination consistency, and color fidelity introduce unique obstacles.

Despite noteworthy progress in fluorescence microscopy image restoration, current methods still face significant obstacles, especially for smartphone-based applications. Classical approaches, including Gaussian filtering and anisotropic diffusion,[13] while straightforward, often oversmooth images and obscure subtle biomarker details.[14] Non-local means (NLM) and BM3D can better preserve structural features but demand high computational resources and often underperform in low-SNR contexts.[15] Wavelet-based denoising (*e.g.*, soft-thresholding) effectively retains fine structures, but its reliance on precise parameter tuning reduces adaptability.[16] Meanwhile, background subtraction methods—digital subtraction or adaptive thresholding—may inadvertently eliminate faint yet important fluorescence signals.[17,18]

Deep learning-based methods offer state-of-the-art denoising but typically require large, high-quality datasets, which is a challenge at the point of care. For instance, Content-Aware Image Restoration (CARE),[19] Noise2Noise (N2N),[20] and Noise2Void (N2V)[21] excel under controlled training environments, yet annotated data are rarely available in resource-limited settings(see ESI section S2† for a detailed comparison with CARE, N2N, N2V and other methods). By contrast, Deep Image Prior (DIP)[12] circumvents the need for large datasets but is prone to overfitting if optimization runs too long. Moreover, many existing pipelines do not fully account for the varied noise distributions observed in fluorescence microscopy.[18]

To address these limitations, we propose HIST-DIP, which fuses histogram thresholding with DIP-based super-resolution. This adaptive thresholding step attenuates background noise *prior* to iterative refinement, mitigating overfitting and improving contrast without the need for labeled data.[20] By leveraging CNN structural priors and computationally efficient preprocessing, HIST-DIP demonstrates improved performance in SNR, PSNR, and contrast-to-noise ratio (CNR),[15,16] while preserving crucial biological structures.[17,18] Consequently, HIST-DIP constitutes a robust, economical platform for smartphone-based fluorescence imaging, enabling accurate point-of-care diagnostic evaluations even in environments lacking advanced microscopy resources.

2. Materials and methods

2.1. Deep image prior for super-resolution

Deep Image Prior (DIP)[12] posits that the convolutional architecture of a randomly initialized neural network can itself serve as a powerful image prior, even in the absence of external training data. Originally introduced for tasks such as inpainting and denoising, DIP has also been successfully applied to super-resolution.[12,22,23] In super-resolution, the objective is to learn a mapping that transforms a low-resolution (LR) image, denoted by $x_0 \in \mathbb{R}^{3\times H\times W}$, into a high-resolution (HR) image $x \in \mathbb{R}^{3\times tH\times tW}$ by an upsampling factor t. This transformation is modeled by the following data-fidelity term:

$$E(x;\ x_0) = \| d(x) - x_0 \|^2 \tag{1}$$

where $d(\cdot)$ is a known downsampling operator that reduces the dimensionality of x by the factor t. The challenge in super-resolution arises because many plausible HR images could yield the same LR observation x_0, rendering the problem severely ill-posed.

2.1.1. DIP formulation. To regularize this inverse problem, DIP reparametrizes the sought image x as

$$x = f_\theta(z) \tag{2}$$

where z is a random (often Gaussian) input tensor, and f_θ is a deep convolutional network with parameters θ. The network is then optimized to minimize energy in (1) with respect to θ:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \| d(f_\theta(z)) - x_0 \|^2 \tag{3}$$

Due to the expressivity of deep networks, an unconstrained minimization can overfit noise or aliases in super-resolution tasks. As a remedy, various strategies have emerged, including early stopping,[12] additive noise regularization,[23] and stochastic gradient-based methods.[24]

2.1.2. Extensions of DIP. While DIP has demonstrated remarkable flexibility, it often requires additional techniques to maintain robustness: *DIP-SURE*[25] employs Stein's Unbiased Risk Estimator to avoid access to ground-truth images; it mitigates overfitting by estimating the risk of reconstruction directly from noisy observations. *DeepRED*[26] combines DIP with Regularization by Denoising (RED), leveraging an external denoiser as a plug-and-play prior to constrain the solution space. *DIP-TV*[27] introduces a total-variation-based regularization term alongside DIP to better preserve edges and reduce artifacts. *Self2Self*[28] merges dropout and ensembling to improve DIP-like self-supervised denoising for a single image, illustrating that uncertainty modeling can further stabilize DIP in ill-posed settings.

In this work, we adopt the DIP concept to enhance super-resolution for smartphone-based fluorescence microscopy images. By treating the downsampled microscope images as x_0 and iteratively refining $f_\theta(z)$, we aim to recover a higher resolution approximation x. To address overfitting to microscope noise, we incorporate early stopping and additive regularization approaches inspired by DIP-SURE and DIP-TV. Section 2.3 details the overall pipeline and specific loss functions employed to stabilize training and improve the quality of super-resolved fluorescence images.

2.2. Histogram-based image masking and contrast refinement

Following insights from classical image contrast enhancement and unsharp masking techniques,[29] we integrate a histogram-driven mask to focus on salient fluorescence signals while suppressing noisy background pixels. This approach not only sep-

 This journal is © The Royal Society of Chemistry 2025

Open Access Article. Published on 23 July 2025. Downloaded on 5/4/2026 1:38:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

arates true fluorescence from noise but also avoids common pitfalls of direct histogram equalization (HE), such as over- or under-enhancement of crucial image regions.[30]

2.2.1. Histogram computation. Given a grayscale fluorescence image I, we first compute its histogram $H(k)$ (eqn (4)):

$$H(k) = \sum_{i=1}^{M}\sum_{j=1}^{N} \mathbf{1}(I(i,j) = k) \quad (4)$$

where $\mathbf{1}(\cdot)$ is the indicator function (outputting 1 if true, 0 otherwise), and k 0, 1, ... $L-1$ denotes an intensity bin in an L-level image (*e.g.*, $L = 256$ for 8-bit depth).

2.2.2. Threshold selection and mask generation. To isolate the fluorescence signal from noise, a manually set threshold T is chosen by inspecting the histogram's tail region where background intensity values accumulate.[29,30] Pixels above this threshold are deemed noise and set to zero, while pixels below or equal to T retain their original intensities. For each pixel (i, j), binary mask can be given as,

$$M(i,j) = \begin{cases} 1, & I(i,j) \le T \\ 0, & I(i,j) > T \end{cases}, \quad (5)$$

We then do element-wise multiplication of the binary mask $M(i, j)$ with the original low-resolution fluorescence microscopy image $I_{\mathrm{LR}}(i, j)$. This mask, defined through adaptive histogram thresholding, assigns a value of 1 to pixels corresponding to significant fluorescence signals and 0 to background pixels. Multiplying this mask element-wise with the original image yields the masked target image $I_{\mathrm{target}}(i,j)$:

$$I_{\mathrm{target}}(i,j) = I_{\mathrm{LR}}\ (i,j) \odot M(i,j) \quad (6)$$

This masked image, I_{target}, subsequently serves as the reference for training the Deep Image Prior (DIP) network, enabling effective denoising and enhancement of structural details without external labeled data. Fig. 1 demonstrates the histogram thresholding applied to a fluorescence microscopy image (left). The histogram (center) shows a tail region (red box) corresponding to noise-like intensities. A chosen threshold of 143 (green arrow) segments the significant signal from noise, yielding the binary mask (right).

2.2.2.1. Why manual thresholding?. Although various automated thresholding algorithms exist (*e.g.*, Otsu's method,[31] or more advanced techniques surveyed in ref. 32), heavily blurred images and the absence of gold-standard reference images complicate their application in our study. Because we collect beads of different sizes and intensities for biomarker quantification, a universal automated threshold often fails to robustly isolate critical fluorescence signals. By manually inspecting each image histogram, we can adaptively suppress background noise while preserving essential biomarker information. Furthermore, once the model parameters are learned on training data, applying HIST-DIP to unseen images does not require re-tuning the threshold, since the DIP component can accommodate local intensity variations.

2.2.3. Contrast and brightness considerations. Although histogram-based thresholding effectively eliminates large portions of noise, naive HE on the masked image can cause over-amplification of the remaining intensities, leading to "overenhancement" artifacts.[29] In classical image processing, *unsharp masking* is often introduced before or after histogram operations to moderate excessive contrast changes.[29] In our scenario, applying thresholding first allows the network to focus on salient structures while avoiding strong amplification of empty (masked-out) regions. This workflow mitigates brightness shifts and artifacts observed in more conventional HE-based methods.

2.2.4. DIP integration. Next, we feed the masked target image from 6 as the *target tensor* for our DIP-based reconstruction, as depicted in Fig. 2. The DIP network then iteratively refines an estimate of the denoised image by minimizing the pixel-wise mean squared error:

$$\ell_{\mathrm{MSE}} = \| f_\theta(z) - \mathbf{I}_{\mathrm{target}} \|^2 . \quad (7)$$

Here, θ denotes the DIP network parameters, and z is a fixed random tensor serving as input to the network.[12] By

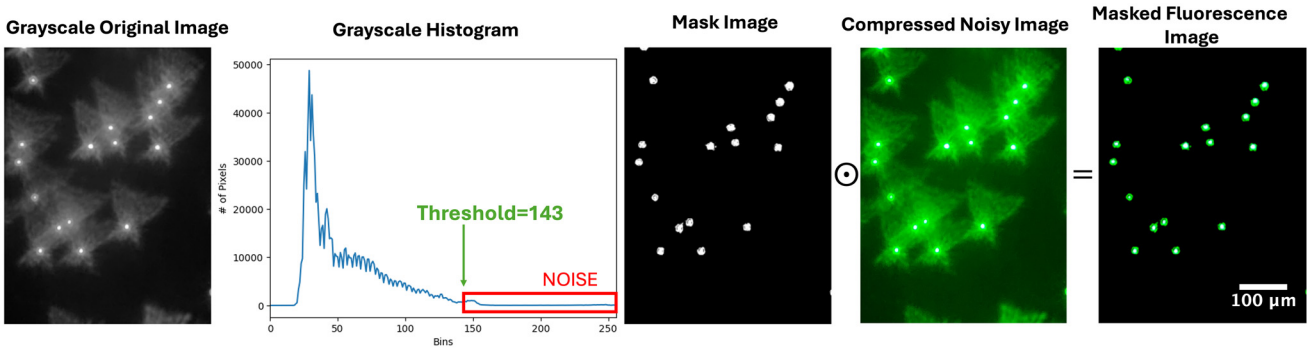

Fig. 1 Histogram-based thresholding for fluorescence signal segmentation. *Left*: Raw grayscale fluorescence microscopy image. *Second from Left*: Intensity histogram illustrating the distribution of pixel values, with a selected threshold (green arrow) separating meaningful signals from background noise (red box). *Middle*: Binary mask generated using the threshold, isolating the prominent fluorescent beads. *Second from Right*: Overlay of the binary mask on the original image, highlighting segmented fluorescence bead regions. *Right*: Extracted fluorescence bead, enhancing biomarker visualization for further analysis. This adaptive thresholding approach refines segmentation accuracy by suppressing background noise while preserving biologically relevant fluorescence signals. (Scale bar = 100 μm applies to all the images in the row.).

Open Access Article. Published on 23 July 2025. Downloaded on 5/4/2026 1:38:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

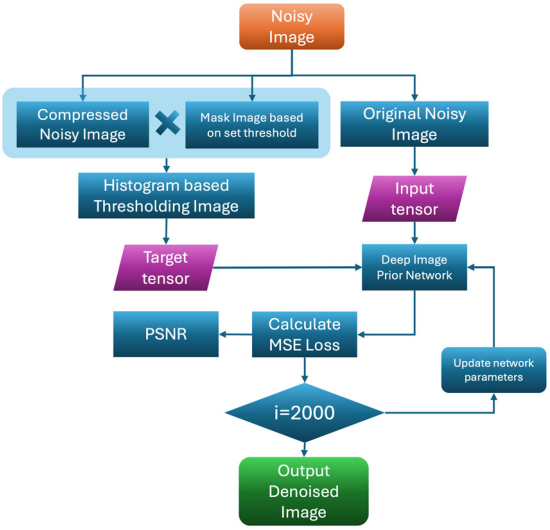

Fig. 2 Overview of DIP-based denoising with histogram thresholding algorithm. Masked image serves as the target tensor; a deep network is iteratively optimized *via* MSE loss until the desired iteration count (*e.g.,* 2000). The final output is the denoised fluorescence image with controlled contrast.

training against a masked target, the DIP naturally learns to prioritize relevant fluorescence signals and ignore predominantly noisy or saturated regions. The final output is thus a denoised image with improved contrast and reduced background interference. In summary, applying a histogram-based threshold to remove outlier intensities, then feeding the masked data into a DIP network, yields a robust framework for fluorescence microscopy image denoising. This hybrid solution takes advantage of classical contrast enhancement insights[29] while leveraging the powerful inductive bias of DIP to focus on relevant signals, reduce background noise, and minimize over-enhancement or brightness drift.

2.3. Training and validation

A central challenge in this work is to establish a suitable reference for training without relying on traditional ground truth images. We address this by leveraging the histogram-thresholded mask described in section 2.2. Specifically, each pixel in the raw low-resolution (LR) image is multiplied by a binary mask that selects high-intensity fluorescence signals while suppressing background noise. Formally:

$$I_{\text{target}}(i,j) = I_{\text{LR}}(i,j) \times \text{Mask}(i,j) \quad (8)$$

where $\text{Mask}(i,j) \in \{0, 1\}$ is derived *via* histogram thresholding. The resulting "masked image" effectively isolates regions of interest (fluorescent beads), thereby reducing the risk of super-resolving or amplifying noisy background pixels.

The shaky image shown in Fig. 1 ("Compressed Noisy Image") serves as the initial input to our Deep Image Prior (DIP) model. Typically produced by our 3D-printed smartphone microscope, these images exhibit significant motion blur that obscures key biomarker details. We then train the DIP model for 2000 iterations using the Adam optimizer with α = 0.001 (see ESI Fig. S2† for PSNR and loss curves validating these training choices). During early iterations, the network rapidly smooths out coarse artifacts and high-frequency noise in the LR image. As training progresses, the DIP architecture adapts to local texture cues, refining the fluorescent bead boundaries and enhancing contrast. Empirically, 2000 iterations adequately balance image sharpness with noise reduction, mitigating overfitting issues often observed when DIP is run for too many epochs[23] (ESI Fig. S2† illustrates the PSNR/loss trends and early-stopping experiments). We evaluate reconstruction fidelity in two main ways. First, the Peak Signal-to-Noise Ratio (PSNR) is tracked against I_{target} (the masked LR image), providing a quantifiable sense of how closely the network output aligns with the intended fluorescence regions. Second, we track the MSE loss curve to monitor convergence behavior.

Both curves are plotted in Fig. 3(a), illustrating that the DIP-based method steadily improves reconstruction fidelity over iterations. Fig. 2 summarizes the overall pipeline. Beginning with the "Compressed Noisy Image", we apply histogram thresholding to create a binary mask, yielding I_{target}. DIP then takes a fixed random tensor z as input at each iteration, producing successively refined outputs as θ is optimized (7). Since no conventional ground truth is available, the mask-based target serves as a practical reference, ensuring that critical structures (fluorescent beads) are preserved and extraneous backgrounds remain suppressed.

2.4. Experimental evaluation

To validate our proposed method, we conducted experiments using images captured from a 3D-printed smartphone-based microscope,[7] specially designed for the detection of fluorescent beads ranging from 1 μm to 8 μm (see ESI Fig S1†). These beads act as proxies for proteins and other biomarkers in biological assays. Despite their utility, the raw images often exhibit substantial motion blur and noise, complicating direct quantification of these signals.

2.4.1. Evaluation dataset and protocol. We trained a Deep Image Prior (DIP) model to perform super-resolution on one of blurry, low-resolution images from the microscope. To evaluate the model's generalizability, we tested it on 10 additional, previously unseen images labeled *a*–*j*, each containing fluorescent beads from Bangs Laboratories. Specifically, images *a*, *b*, *d*, *g*, *j* included 8.3 μm beads (Bangs Laboratories, product #UMDG003), *c*, *f* contained 2.0 μm beads (Bangs Laboratories, product #FSDG005), and *e*, *h*, *i* featured 1.0 μm beads (Bangs Laboratories, product #FSDG004). Note that image *d* is the same one used for training, so it is not strictly an unseen sample. Its inclusion here checks whether the model remains consistent on the training data–useful for detecting overfitting–though it does not represent a truly independent test. Consequently, the results on the other images better reflect real-world performance, and all averages throughout the paper exclude image *d* to provide an unbiased measure.

2.4.2. Baselines and methods. We compared our HIST-DIP approach to both the unprocessed (raw) images and classical

This journal is © The Royal Society of Chemistry 2025

Open Access Article. Published on 23 July 2025. Downloaded on 5/4/2026 1:38:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

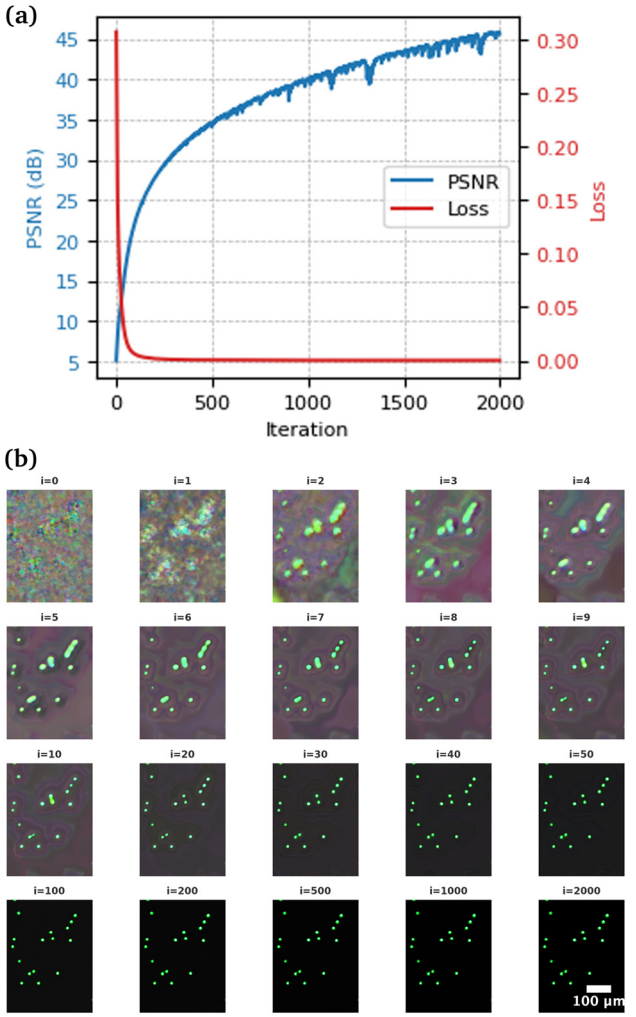

Fig. 3 (a) Training curves showing Peak Signal–to–Noise Ratio (PSNR, dB) and mean–squared–error loss *versus* iteration for the original noisy input. (b) Progressive frames from HIST–DIP demonstrating simultaneous denoising and super–resolution of smartphone fluorescence images at selected iterations. Scale bar = 100 µm (applies to every panel).

Gaussian filtering with a kernel size of 11 × 11, zero mean, and variance set to 5. Gaussian filtering is commonly used in fluorescence imaging to mitigate noise, yet it can blur small, high-frequency details, making it an informative contrast to our method's performance.

2.4.3. Metrics. We evaluated the denoising and super-resolution quality using the following measures:

2.4.3.1. Contrast-to-noise ratio (CNR) and signal difference-to-noise ratio (SDNR). Using the automated KPI quantification tool described in ref. 33, each fluorescent bead's intensity, surrounding vicinity intensity, and background noise level are calculated. The algorithm then computes:

$$\mathrm{SDNR} = \frac{(\text{Bead Intensity} - \text{Vicinity Intensity})}{\text{Background Noise}}$$
$$\mathrm{CNR} = \frac{(\text{Bead Intensity} - \text{Vicinity Intensity})}{\text{Vicinity Intensity} \times \text{Background Noise}} \quad (9)$$

providing insights on whether fluorescence signals are amplified while background remains relatively suppressed.

2.4.3.2. Peak signal-to-noise ratio (PSNR). It is a quantitative based assessment and widely used metric that measures the fidelity of a processed image relative to a reference image. It is expressed in decibels (dB) and is calculated based on the mean squared error (MSE) between the original and enhanced images. A higher PSNR value indicates lower distortion and better preservation of image details. However, PSNR alone does not always correlate well with perceived image quality, as it primarily evaluates pixel-wise differences and does not account for structural or perceptual similarities.

2.4.3.3. Structural similarity index measure (SSIM). It evaluates image quality by considering structural information, luminance, and contrast. SSIM ranges from 0 to 1, where a value closer to 1 indicates high similarity between the reference and processed images. Unlike PSNR, SSIM aligns more closely with human visual perception, making it particularly relevant for biomedical images wherein the preservation of morphological features is critical.

By combining these metrics, we aim to obtain a holistic understanding of denoising and super-resolution performance. Specifically, CNR and SDNR address whether our approach effectively enhances the contrast of fluorescent beads relative to background noise; PSNR measures overall fidelity, and SSIM gauges structural consistency. The PSNR and SSIM were measured against the ground truth images obtained using the method delineated in section 2.2. As such, the proposed HIST-DIP technique can be comprehensively evaluated regarding its suitability for improved accuracy fluorescent biomarker analysis in smartphone-based microscopy.

3. Results

The application of the Deep Image Prior (DIP) model for super-resolution effectively enhanced the visual clarity of images captured using the 3D-printed smartphone-based fluorescence microscope. Initially, the raw images exhibited significant blurring and noise, primarily due to the low-quality optical components of the makeshift microscope and inherent shaking during image capture. These issues were particularly problematic as they obscured the fluorescence signals from the beads, which are critical for accurate biomarker identification.

3.1. Image quality improvement

Quantitative assessments on the 10 validation images showed a marked improvement in image quality. Before training the single image on the DIP model, the average PSNR of the image was approximately 14 dB. After processing with HIST-DIP technique, the average PSNR increased to around 45 dB after 2000 iterations (Fig. 3(a)). This improvement indicates less noise and higher clarity, allowing for better visualization of the fluo-

Open Access Article. Published on 23 July 2025. Downloaded on 5/4/2026 1:38:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

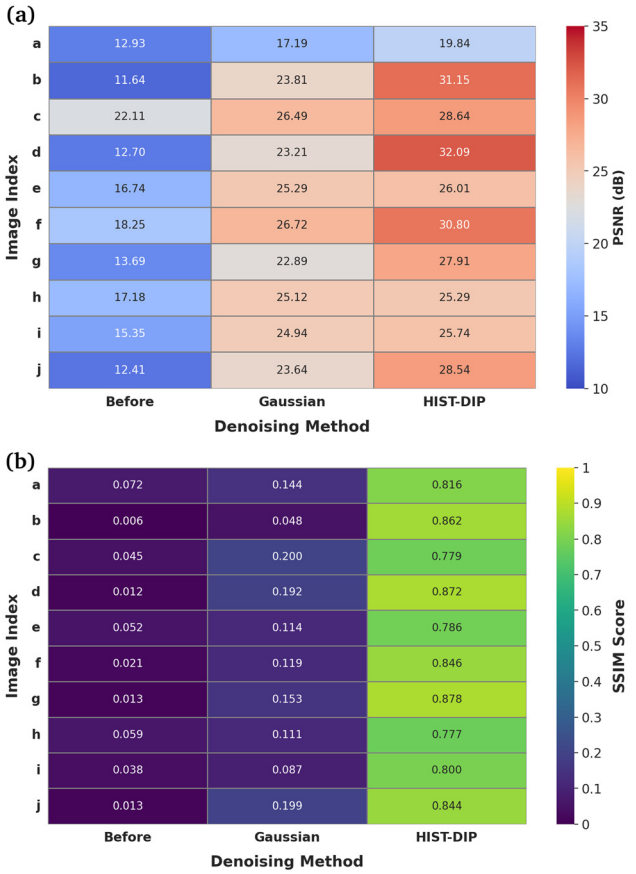

Fig. 4 Comparison of (a) PSNR (dB) heatmap and (b) SSIM heatmap, 0 → completely dissimilar, 1 → identical images.

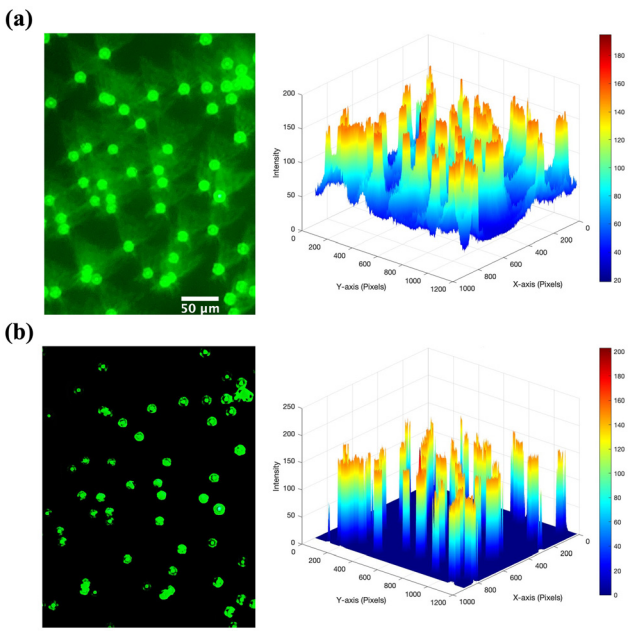

Fig. 5 Fluorescence microscopy and corresponding 3D intensity surface plots. (a) Original fluorescence microscopy image of the validation image 'a' (left) and its 3D intensity distribution(right). (b) HIST-DIP model applied on validation image 'a' (left) and its corresponding 3D intensity distribution(right). (Scale bar = 50 µm applies to both the images).

rescent beads. Fig. 3(b) illustrates the iterative improvement of image quality using the Histogram Thresholding and Deep Image Prior (HIST-DIP) framework. Starting from $i = 0$, where the image is dominated by noise, the reconstruction process gradually enhances signal clarity and reduces background noise. As iterations increase ($i = 1$ to $i = 50$), structural features emerge, distinguishing fluorescent beads from the noisy background. By $i = 100$ to $i = 2000$, the denoised images exhibit high contrast and well-defined fluorescence, facilitating improved biomarker visualization.

The trained model was subsequently evaluated on ten unseen fluorescence microscopy images, each containing beads of varying sizes. To establish a consistent baseline, we generated a pseudo-ground-truth by applying the same histogram-based threshold to every raw image and multiplying pixel-wise. Against this masked reference, the analysis shown in Fig. 4 reveals average PSNR rose from approximately 15.59 dB (unprocessed "before") to 27.10 dB (HIST-DIP), significantly

Table 1 Comparison of image quality metrics, including bead intensity, noise intensity, CNR, and SDNR for Gaussian filtering (AQAFI), CARE, and HIST-DIP across images a–j

Image	Method	Bead intensity	Noise intensity	CNR	SDNR
a	Before	92.47	19.18	0.071	2.69
	Gaussian	11.92	1.53	0.990	4.67
	CARE	113.81	4.82	27.090	65.49
	HIST-DIP	**86.35**	**0.032**	**1421.68**	**2516.87**
b	Before	206.47	35.05	0.078	4.32
	Gaussian	52.14	6.17	0.810	7.05
	CARE	207.36	10.87	2.390	19.11
	HIST-DIP	**166.76**	**0.54**	**153.16**	**301.80**
c	Before	129.11	12.59	0.300	8.06
	Gaussian	30.94	2.83	1.920	9.18
	CARE	124.50	6.96	2.080	17.87
	HIST-DIP	**70.92**	**0.48**	**87.06**	**142.36**
d	Before	240.14	21.15	0.140	8.46
	Gaussian	78.49	2.33	4.300	30.55
	CARE	210.99	10.22	4.370	27.38
	HIST-DIP	**220.45**	**0.089**	**1669.00**	**2436.62**
e	Before	97.43	14.31	0.160	4.74
	Gaussian	24.92	3.12	1.150	6.20
	CARE	88.89	10.07	1.160	15.29
	HIST-DIP	**41.93**	**0.15**	**207.56**	**265.99**
f	Before	150.80	18.17	0.200	6.45
	Gaussian	25.32	3.58	1.010	5.51
	CARE	131.99	7.98	2.340	17.69
	HIST-DIP	**75.64**	**0.14**	**341.18**	**502.03**
g	Before	250.72	21.74	0.150	8.83
	Gaussian	83.37	3.18	3.480	23.94
	CARE	206.10	11.42	4.020	28.31
	HIST-DIP	**245.52**	**0.09**	**563.69**	**2243.03**
h	Before	94.40	15.22	0.150	4.30
	Gaussian	24.91	3.77	0.890	5.07
	CARE	87.83	13.61	0.670	7.29
	HIST-DIP	**26.05**	**0.07**	**283.61**	**318.97**
i	Before	88.48	15.29	0.110	3.62
	Gaussian	23.99	4.27	0.660	4.14
	CARE	87.34	15.33	0.580	6.43
	HIST-DIP	**32.52**	**0.05**	**390.26**	**610.56**
j	Before	231.88	19.42	0.180	9.22
	Gaussian	73.89	2.80	3.880	23.92
	CARE	192.72	8.24	4.620	26.51
	HIST-DIP	**183.34**	**0.12**	**1212.94**	**1463.90**

This journal is © The Royal Society of Chemistry 2025

Open Access Article. Published on 23 July 2025. Downloaded on 5/4/2026 1:38:24 AM. This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

outperforming Gaussian filtering at around 24.01 dB. A similar trend emerged in SSIM: raw images exhibited poor structural similarity (often below 0.06), Gaussian filtering yielded modest improvements (0.08–0.20), and HIST-DIP consistently exceeded 0.77. Fig. 4(a) illustrates the PSNR heatmap for the three denoising techniques: raw (*before*), Gaussian filtering (AQAIF), and HIST-DIP. Cooler (blue) shades represent lower PSNR, while warmer (red) shades indicate higher PSNR values.

As shown, HIST-DIP achieves the stronger performance across all images, with especially large gains in images b, d, and f, where it surpasses the Gaussian filter by more than 8 dB. This underscores the advantages of combining histogram-based preprocessing and DIP for enhancing signal clarity, crucially without the need for labeled datasets.[7]

Likewise, the SSIM heatmap in Fig. 4(b) confirms HIST-DIP's effectiveness in structural preservation. Whereas raw images suffer from high noise, and Gaussian filtering remains uneven, HIST-DIP consistently attains near-ideal SSIM scores. Overall, these results affirm that HIST-DIP is a robust, training-light approach for denoising fluorescent images in smartphone-based microscopy, offering clearer biomarker visualization than conventional filtering techniques.

The two 3D intensity surface plots in Fig. 5 compare pixel-level distributions (X–Y in pixels, Z as intensity). Fig. 5(a), raw image coming from the SBFM exhibits a broader intensity spread with moderate peaks, indicating some noise and smoother transitions between high- and low-intensity regions. Fig. 5(b), raw image when HIST-DIP model is applied shows sharper, more isolated peaks and clearer low-intensity areas, suggesting enhanced contrast and reduced noise. Overall, HIST-DIP model filtered image appears more structured, implying an improvement in signal clarity relative to the raw image.

3.2. Fluorescent signal clarity

Fluorescent markers play a central role in diagnosing protein expression and other clinically relevant biomarkers, making their clear visualization paramount in biomedical image analysis. Table 1 compares key performance indicators (KPIs) including bead intensity, background noise, and the derived signal difference-to-noise ratio (SDNR) and contrast-to-noise ratio (CNR) using the AQAFI method.[33] Representative images demonstrating these improvements are shown in Fig. 6. These measures collectively gauge how effectively different denoising approaches isolate true fluorescent signals while minimizing noise contributions. From the *before* rows, the raw images display low SDNR and CNR, reflecting poor signal discernibility. Applying Gaussian filtering (*Gaussian* rows) yields moderate improvements but often blurs finer bead structures due to uniform smoothing. By contrast, the proposed HIST-DIP method (highlighted rows) achieves an order-of-magnitude increase in both SDNR and CNR. This substantial gain in signal clarity–achieved without large, annotated datasets–makes HIST-DIP a robust, low-cost solution for smartphone-based fluorescence microscopy, particularly in scenarios where

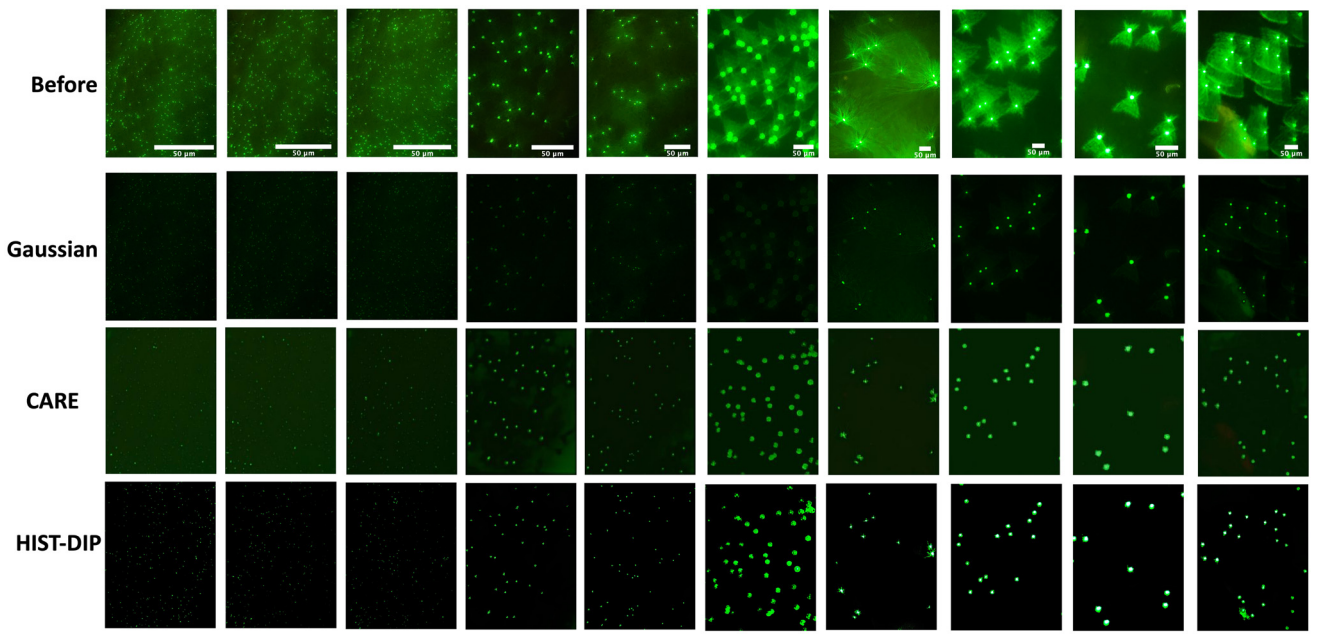

Fig. 6 Denoising performance at three bead diameters. Columns (left → right) correspond to bead sizes of 1.0 μm (e, h, i), 2.0 μm (c, f), and 8.3 μm (a, b, d, g, j). Top row: Raw smartphone-acquired fluorescence frames dominated by shot noise and optical blur. Second row: Gaussian-filtered images, where noise is attenuated but fine structure is visibly softened. Third row: Supervised baseline produced with CARE, which further suppresses noise yet leaves some residual blur and speckle. Bottom row: Output of the proposed HIST-DIP pipeline; noise is strongly suppressed while bead contours are sharpened and local contrast enhanced, yielding the clearest visualization of bead-bound biomarkers. (Scale bar: 50 μm in every panel.)

Open Access Article. Published on 23 July 2025. Downloaded on 5/4/2026 1:38:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

photobleaching, temporal resolution, and live-cell viability pose significant challenges. By better isolating and amplifying key biomarkers, our approach reduces reliance on costly laboratory equipment and potentially broadens the use of portable diagnostic imaging in resource-constrained settings.

4. Conclusions

This work developed HIST-DIP, a combined histogram thresholding and Deep Image Prior (DIP) approach for denoising and super-resolving fluorescence microscopy images captured by low-cost, smartphone-based devices. Experimental validation across multiple test images (excluding the training image *d*) confirmed significant performance gains quantified by several metrics (see ESI section S3† for extended experimental validation and discussion of generalizability). In particular, our method boosted the average PSNR from approximately 15.6 dB (raw) and 24.0 dB (Gaussian) to 27.1 dB (HIST-DIP), while the SSIM increased from 0.035 (raw) and 0.131 (Gaussian) to 0.82 (HIST-DIP). Moreover, the key bioanalytical metrics of signal difference-to-noise ratio (SDNR) and contrast-to-noise ratio (CNR) demonstrated orders-of-magnitude gains; on average, CNR improved from ~0.16 in the raw images to ~1.64 with Gaussian filtering, and then to ~518.0 under HIST-DIP, while SDNR rose from ~5.80 to ~9.96 and ultimately ~929.5. These numerical gains indicate clearer bead outlines and substantially reduced background interference, making HIST-DIP a powerful solution for low-cost point-of-care imaging. A live web deployment of the trained HIST-DIP model for real-time denoising can be accessed at: **https://denoisedip.vercel.app/**.

Our findings also point to the following challenges. First, although the DIP architecture provides strong priors for self-supervised denoising, it remains sensitive to iteration count, masking thresholds, and hyperparameters–factors (as elaborated in ESI section S1†) that may complicate clinical or field deployment. Second, while histogram thresholding mitigates over-enhancement, suboptimal thresholds can exclude faint structures or retain undesired artifacts. Although the approached method does not rely on large annotated datasets, it is constrained by the quality of the mask generation step, which, if improperly tuned, can lead to over/under-segmentation of the fluorescence signals. Third, extending HIST-DIP to real-time applications will require accelerated optimization schemes or integration with hardware.

Future research will focus on refining threshold-selection procedures using adaptive and data-driven approaches, exploring domain-specific priors to stabilize DIP in varying illumination conditions, and investigating the utility of advanced regularizers such as physics-informed or learned from large unlabeled image sets. Achieving these goals will broaden the accessibility of smartphone-based imaging for diagnostics, enabling effective quantitative analysis in resource-limited environments without reliance on large-scale labeled datasets or specialized equipment.

Author contributions

H. G. contributed to conceptualization, data curation, formal analysis, investigation, methodology, resources, software, validation, visualization, writing – original draft, and writing – review & editing. M. N. T. contributed to data curation and analysis for the AQAFI metric computations. U. H. contributed to conceptualization, study design, funding acquisition, investigation, project administration, resources, supervision, writing – review & editing.

Conflicts of interest

Authors declare no conflicts of interest.

Data availability

All data supporting the findings of this study, including the raw noisy fluorescence microscopy images used for testing, are openly available at our GitHub repository: **https://github.com/hg293/HIST-DIP**. The repository includes an images folder containing multiple noisy test images compatible with the deployed web application **https://denoisedip.vercel.app/** for real-time denoising. Additional scripts and configuration details are also provided to facilitate reproducibility and further experimentation.

Acknowledgements

Authors would like to thank M. A. Sami for his contribution in collecting images[7] used for this study. The authors would like to acknowledge the funding support from National Science Foundation (NSF award number 2315376), and Office of Naval Research (Award No. N00014-24-1-2156 and N00014-20-1-2542). Authors also acknowledge support from the Department of Electrical and Computer Engineering, and the Global Health Institute at Rutgers, The State University of New Jersey.

References

1 D. Zhang and Q. Liu, *Biosens. Bioelectron.*, 2016, **75**, 273–284.
2 D. N. Breslauer, R. N. Maamari, N. A. Switz, W. A. Lam and D. A. Fletcher, *PLoS One*, 2009, **4**, e6320.
3 K. Fan, W. Liu, Y. Miao, Z. Li and G. Liu, *Adv. Intell. Syst.*, 2023, **5**, 2200285.
4 S. Di Nonno and R. Ulber, *Analyst*, 2021, **146**, 2749–2768.
5 S. Banik, S. K. Melanthota, Arbaaz, J. M. Vaz, V. M. Kadambalithaya, I. Hussain and N. Mazumder, Recent trends in smartphone-based detection for biomedical applications: a review, *Anal. Bioanal. Chem.*, 2021, **413**(9), 2389–2406.
6 A. Roda, E. Michelini, M. Zangheri, M. Di Fusco, D. Calabria and P. Simoni, *Trends Anal. Chem.*, 2016, **79**, 317–325.

This journal is © The Royal Society of Chemistry 2025

Open Access Article. Published on 23 July 2025. Downloaded on 5/4/2026 1:38:24 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

7 M. A. Sami, M. Tayyab, P. Parikh, H. Govindaraju and U. Hassan, *Analyst*, 2021, **146**, 2531–2541.
8 J. Nelis, A. Tsagkaris, M. Dillon, J. Hajslova and C. Elliott, *Trends Anal. Chem.*, 2020, **129**, 115934.
9 H. Kholafazad-Kordasht, M. Hasanzadeh and F. Seidi, *Trends Anal. Chem.*, 2021, **145**, 116455.
10 C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo and C.-W. Lin, *Neural Networks*, 2020, **131**, 251–275.
11 S. Chen, D. Shi, M. Sadiq and X. Cheng, *IEEE Access*, 2020, **8**, 82819–82831.
12 D. Ulyanov, A. Vedaldi and V. Lempitsky, Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 9446–9454.
13 P. Perona and J. Malik, *IEEE Trans. Pattern Anal. Mach. Intell.*, 1990, **12**, 629–639.
14 A. Buades, B. Coll and J. M. Morel, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2005, pp. 60–65.
15 J. Boulanger, C. Kervrann, P. Bouthemy, P. Elbau, J. B. Sibarita and J. Salamero, *IEEE Trans. Med. Imaging*, 2010, **29**, 442–454.
16 D. L. Donoho, *IEEE Trans. Inf. Theory*, 1995, **41**, 613–627.
17 A. E. Profio, *Med. Phys.*, 1986, **13**, 717–721.
18 C. Vonesch, F. Aguet, J. L. Vonesch and M. Unser, *IEEE Signal Process. Mag.*, 2006, **23**, 20–31.
19 M. Weigert, U. Schmidt, T. Boothe, *et al.*, *Nat. Methods*, 2018, **15**, 1090–1097.
20 J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala and T. Aila, International Conference on Machine Learning (ICML), 2018.
21 A. Krull, T.-O. Buchholz and F. Jug, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 2129–2137.
22 R. Heckel and P. Hand, International Conference on Learning Representations (ICLR), 2019.
23 Y. Jo, S. Y. Chun and J. Choi, *arXiv*, 2023, preprint, arXiv:2301.00000.
24 D. Van Veen, A. Jalal, E. Price and A. G. Dimakis, *Bayesian and Statistical Deep Learning Workshop at NeurIPS*, 2018.
25 C. A. Metzler, A. Mousavi, R. Heckel and R. Baraniuk, *Workshop on Signal Processing with Adaptive Sparse Structured Representations*, 2018.
26 G. Mataev, P. Milanfar and M. Elad, IEEE International Conference on Computer Vision Workshops, 2019.
27 J. Liu, Y. Sun, X. Xu and U. S. Kamilov, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 7715–7719.
28 Y. Quan, M. Chen, T. Pang and H. Ji, IEEE Conference on Computer Vision and Pattern Recognition, 2020, pp. 1890–1898.
29 S. Kansal, R. K. Tripathi and S. Purwar, *Multimedia Tools Appl.*, 2018, **77**, 26919–26938.
30 N. Brancati, M. Frucci and G. Sanniti di Baja, International Conference Image Analysis and Recognition, 2008, pp. 132–141.
31 N. Otsu, *IEEE Trans. Syst. Man Cybern.*, 1979, **9**, 62–66.
32 M. Sezgin and B. Sankur, *J. Electron. Imaging*, 2004, **13**, 146–166.
33 A. Sami, N. Tahir and U. Hassan, *Analyst*, 2023, **148**, 6036–6049.