Research Article Vol. 5, No. 8 / 15 Aug 2026 / *Optics Continuum* 2708

OPTICS CONTINUUM

All-optical selective target highlighting for augmented reality microscopy

MAHA SAHLOUL,[1,2,*] KAAN AKŞIT,[3] AND M. FATIH TOY[1,2,4,5]

[1] *Department of Biomedical Engineering, Istanbul Medipol University, Turkey*
[2] *Biophotonics and Computational Imaging Laboratory (Phi-X), Research Institute for Health Sciences and Technologies (SABITA), Istanbul Medipol University, Turkey*
[3] *Department of Computer Science, University College London, London, UK*
[4] *Department of Electrical and Electronics Engineering, Istanbul Medipol University, Turkey*
[5] *muhammed_fatih.toy@alumni.epfl.ch*
[*] *maha.sahloul@std.medipol.edu.tr*

Abstract: Augmented reality (AR) microscopes assist users in analyzing microscopic specimens by overlaying virtual information onto the captured images. Typically, this approach requires continuous scene analysis and computational processing, which limits real-time performance and introduces significant computational overhead. We introduce an all-optical alternative by proposing a new phase optimization scheme integrated into AR microscopes. Our optimizations rely on creating a precomputed library of phase maps for a spatial light modulator (SLM). Each phase map is configured to selectively enhance a specific shape or size class in a scene, eliminating per-frame computational requirements. Furthermore, by leveraging the inherent shift-invariant property of Fourier optical systems, our phase masks maintain a selective augmentation of the target object during in-plane translations within the field of view (FoV) without explicit tracking algorithms. Our results demonstrate selective target highlighting based on object size or shape, enabling continuous real-time observation in microscopy systems.

1. Introduction

Optical microscopes remain essential for imaging microscopic structures because they are simple, fast, and accessible. Limitations arising from sample complexity and user expertise pose significant challenges in accurately identifying features within microscopic images. While digital microscopes partially address this through screen-based observation, a critical need persists for solutions that enable direct observation through eyepieces. AR microscopes address this challenge by incorporating real-time computational information into optical images, improving detection speed and accuracy [1,2]. Contemporary Artificial Intelligence (AI)-assisted AR microscopes [3] rely on capturing microscope images, performing computational image analysis, and projecting visual augmentations via microdisplays [4–6]. However, these camera-dependent architectures introduce computational bottlenecks and latency. Consequently, there is a critical need for computationally efficient AR microscopy architectures that maintain real-time performance.

First, we survey optical solutions that promise to overcome computational load and latency [7], including optical correlators [8], Fourier-domain optical computing systems [9], and Optical Neural Networks (ONNs) [10]. These systems perform mathematical operations through light propagation, thus enabling high-speed, real-time optical computation capabilities and reducing digital processing during observation. Among these optical computing systems, the VanderLugt correlator [11,12] represents one of the most widely implemented architectures for real-time pattern recognition. It operates based on a $4f$ optical architecture, where the input field is first Fourier transformed, then modified at the Fourier plane using a correlation filter. This correlation

#606265 https://doi.org/10.1364/OPTCON.606265
Journal © 2026 Received 12 Jun 2026; revised 26 Jun 2026; accepted 26 Jun 2026; published 6 Aug 2026

OPTICS CONTINUUM

filter includes a reference target encoded as a Fourier-plane filter, enabling target detection by producing sharp correlation peaks corresponding to pattern similarity [13]. Because optical correlators possess an inherent shift-invariance property, they are utilized in target detection, recognition, classification, and tracking analysis [14,15]. Recent studies extend optical correlation techniques to various optical information processing tasks. Joint transformation correlation combined with ptychographic reconstruction is employed for image encryption and secure image recovery [16], while holographic optical correlators integrated with deep-learning-based reconstruction enable high-speed single-pixel imaging [17]. In contrast, the proposed framework utilizes Fourier-domain optical processing to achieve selective target highlighting and real-time augmentation in microscopy rather than image encryption or image reconstruction. In parallel, advances in Computer-Generated Holography (CGH) [18] and SLM-based wavefront modulation enable programmable manipulation of optical fields and spatial frequency spectra.

Inspired by these Fourier-domain optical processing principles, we introduce an all-optical AR microscopy framework characterized by an SLM placed at the Fourier plane. Unlike conventional optical correlators that generate correlation peaks for target recognition, our approach exploits Fourier-domain optical processing to produce an augmentation response (annular ring) around the target object. We optimize the phase masks using a forward model, simulating our optical hardware. These phase masks modulate the wavefront of the input field to generate an augmentation response optically, within the existing optical system, without requiring an additional light source or a separate optical path for image overlay. The optimized phase maps are stored in a precomputed library, enabling users to augment specific structures by projecting the selected phase map onto the SLM in the Fourier plane [19]. Owing to the inherent shift invariance gained by optimizing the phase maps in the Fourier plane, our method enables selective highlighting of a specific target class without requiring digital object tracking. Unlike conventional Fourier filtering techniques [20], our method does not affect the entire spatial frequency spectrum of the input field. Instead, it enhances the visibility of selected target structures compared to conventional observation, while preserving the content of the remaining scene. To achieve this, we optimize the phase maps using a dark-field amplitude sample. This sample transmits light only through predefined microscopic shapes. We explore three proof-of-concept cases: (1) object highlighting, where we augment circular objects of common size (30 µm); (2) size-based selective object highlighting, where the phase map targets a specific diameter among circular objects; and (3) morphology-based selective object highlighting, where it focuses on augmenting one geometric shape among various objects, e.g., (circle, triangle, and square). Our work advances three key areas within AR microscopy: camera-feedback-free optimization, real-time optical highlighting of translated targets using a precomputed phase mask method, and experimental validation across object, size-based, and morphology-based highlighting tasks.

2. Methods

2.1. Optical setup

The schematic of our optical setup is shown in Fig. 1. We collimate a 655 nm laser using lens L1 (Thorlabs LA1131-A, $f_{L_1} = 50$ mm). The collimated beam passes through a linear polarizer aligned with the SLM's (Thorlabs EXULUS-HD1, 1920 ×1080 pixels, 6.4 µm pixel size) phase modulation axis. The laser beam then illuminates the sample on a microscope (Zeiss Axio Observer 7). The sample substrate consists of a #1.5 coverslip with a 100 nm sputtered gold layer. Geometric patterns are formed by etching the gold layer via a laser ablation microscope (Zeiss PALM LMD). The beam transmits through a 10X/0.45 NA microscope objective. This microscope objective is followed by a tube lens (L2) and a folding mirror. The intermediate image plane is sized to match the camera's (Basler acA1920-155um, 1920 ×1200 pixels, maximum frame rate of 164 fps) exit window. This aperture prevents aliasing from subsequent elements.

The intermediate image plane is optically conjugated to the camera image plane. For this, the beam from the intermediate image plane first passes through a beam splitter (Thorlabs BS013) and a lens (Thorlabs LA1256-A, f_{L_3}=300 mm). The SLM is positioned at the focal distance from the lens (L3). Direct reflection from the SLM front surface generates an undesired zeroth-order component at the camera plane. In order to spatially separate the diffraction orders, a blazed grating phase ($\psi(x, y)$) is added to the SLM phase pattern. In conventional SLM-based holographic systems, an additional $4f$ spatial filtering stage with a Fourier-plane aperture is commonly used to isolate the first diffraction order. In our approach, we achieve an equivalent filtering effect through the combined use of a blazed grating, an intermediate image plane aperture, and a physical SLM tilt. The SLM is tilted such that the first diffraction order shifts onto the camera Region Of Interest (ROI) while the undesired zeroth order is displaced outside the detection region. We determine the tilt angle of 1.16° according to the diffraction angle generated by the blazed grating. This configuration enables effective order selection and suppression without introducing additional optical components.

2.2. Forward model

SLM phase patterns are optimized using a forward model that simulates the optical setup as shown in Fig. 2. This forward model represents propagation from the intermediate image plane to the camera plane, implemented as a $4f$ system. From the intermediate image plane, the input light field $U_{in}(x, y)$ propagates through free space via the angular spectrum method:

$$U_{p1}(x, y; z) = \mathcal{F}^{-1}\Big(\mathcal{F}(U_{in}(x, y, 0))\, H(f_x, f_y; z)\Big), \tag{1}$$

$$H(f_x, f_y; z) = \exp\left(ikz\sqrt{1 - (\lambda f_x)^2 - (\lambda f_y)^2}\right), \tag{2}$$

where $\mathcal{F}$ refers to the two-dimensional discrete Fourier transform, $H(f_x, f_y; z)$ is the transfer function, $k = 2\pi/\lambda$ is the wavenumber, λ is the wavelength of the light, and z is the propagation distance. The propagated field is modulated by a quadratic phase $L(x, y)$, which represents a thin-lens phase function as shown,

$$L(x, y) = \exp(-i\frac{k}{2f}(x^2 + y^2)), \tag{3}$$

$$U_L = U_{p1}(x, y)L(x, y), \tag{4}$$

where f represents the focal length of the lens (L3). The modulated field then propagates to the SLM plane, forming $U_{p2}(x, y)$. At this plane, the $U_{p2}(x, y)$ field is imposed by an optimizable phase map $M(x, y)$. This phase map contains an adaptive phase $\phi(x, y)$ combined with a blazed grating phase $\psi(x, y)$ that is introduced in the SLM plane as shown,

$$U_M(x, y) = U_{p2}(x, y)M(x, y), \tag{5}$$

$$M(x, y) = \exp(i(\phi(x, y) + \psi(x, y)))\,. \tag{6}$$

Following phase modulation, the field propagates to the camera plane, where the resulting output field and its corresponding intensity distribution are denoted by $U_{out}(x, y)$ and $I(x, y)$, respectively. In the numerical simulations, the propagation distance is set to 300 mm, corresponding to the focal length of the lens (L3). The spatial sampling interval is set to 6.4 µm, corresponding to the SLM pixel pitch. Intermediate-plane images (1180 ×1180 pixels, 5.86 µm pixel pitch) are resampled to match the SLM pixel pitch 6.4 µm. The optimization employs a loss function that combines

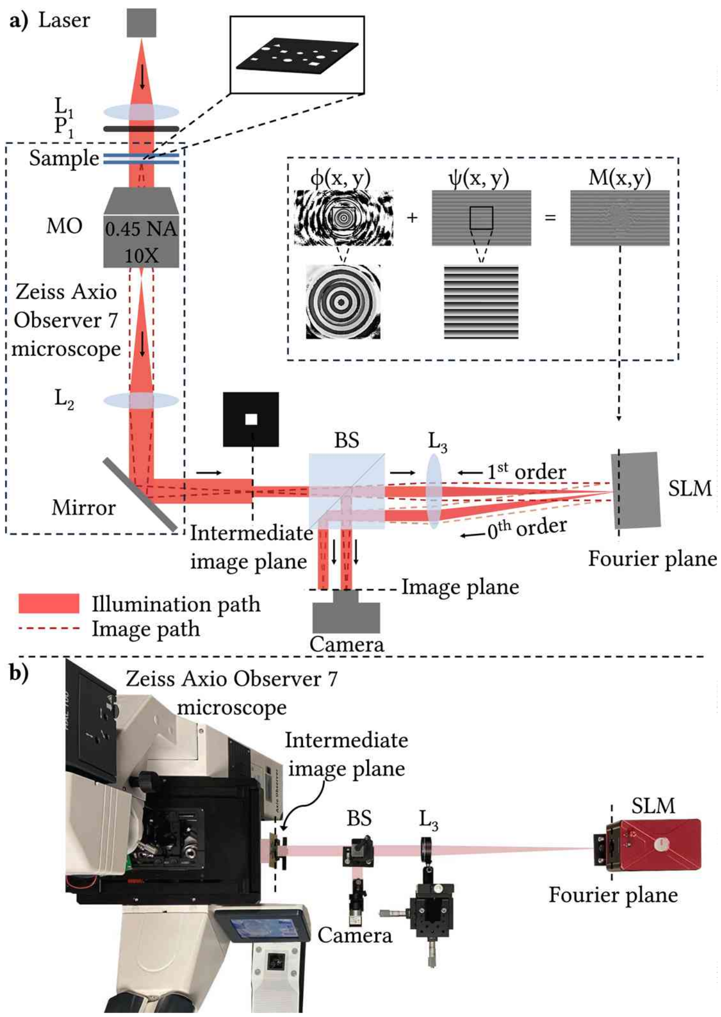

Fig. 1. a) Schematic presents our optical setup including L1: Collimation lens, P1: Polarizer, MO: Microscope objective, L2: Tube lens, BS: Beam splitter, L3: Fourier lens, SLM: Spatial light modulator, $\phi(x, y)$: Adaptive phase, $\psi(x, y)$: Blazed grating phase, $M(x, y)$: SLM phase. b) Photograph of the experimental setup. See Visualization 1 for an overview of the setup from different angles.

OPTICS CONTINUUM

pixel-wise Mean Squared Error (MSE) [21] with Total Variation (TV) [22] regularization between the output intensity distribution and the ground-truth distribution $t(x, y)$ as shown,

$$\hat{\phi} \leftarrow \arg\min_{\phi} \mathcal{L}\big(I(x, y), t(x, y)\big), \tag{7}$$

$$\mathcal{L} = \mathcal{L}_{\mathrm{MSE}} + \mathcal{L}_{\mathrm{TV}}. \tag{8}$$

Before loss computation, both the intensity distribution and the target distribution are normalized by their maximum intensity values to maintain a comparable dynamic range during optimization.

In the dark-field configuration, the phase-only SLM operates by redistributing the incident light field. Additionally, our augmentation process relies on generating an annular ring around the target object. In our implementation, we realize that the algorithm prioritizes minimizing the error in the object while neglecting the intensity of the annular ring. For this reason, we apply a global scaling factor to the target intensity before beginning optimization. This scaling factor ensures that the input and target distributions are comparable in terms of total energy. We compute the scaling factor as the ratio between the input and target intensity distributions. Thus, enforcing energy consistency ensures that the target distribution is compatible with the available input energy, leading to improved convergence of the loss function.

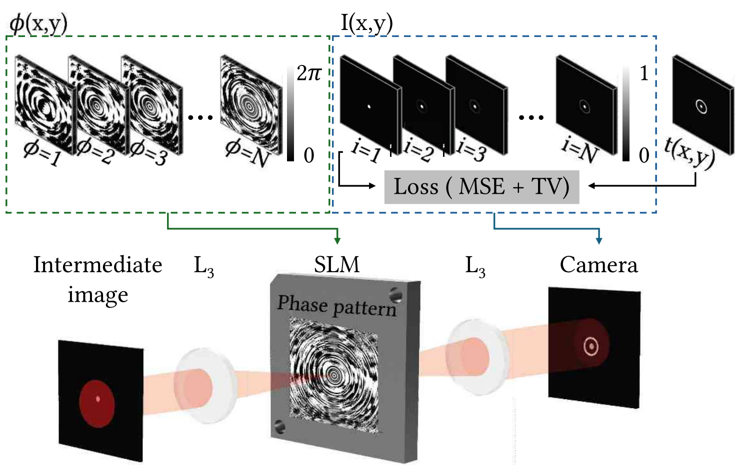

Fig. 2. Phase Pattern Optimization. The sketch shows our forward model representing our optical setup. We utilize the forward model to iteratively optimize phase patterns using the target and MSE and TV regularizers. The phase optimization included both $\phi(x, y)$ and $\psi(x, y)$ phase terms in the SLM plane, but we suppress the $\psi(x, y)$ phase in this sketch for easier visibility of the optimized phase pattern. After optimization, the phase pattern is displayed on the SLM to augment the target object at the camera plane.

We perform the optimization stage using Adam optimizer [23], initialized with a learning rate of 0.1 and implemented in the PyTorch framework [24]. We integrate a polynomial learning rate scheduler with a decay power of 2.0 into the optimization pipeline to adaptively decrease the learning rate and improve convergence. Each phase map calculation is performed using 1500 iterations of the optimizer. All variables are represented in single-precision floating point (FP32) format. We adjust the optimized phase map using a pre-calibrated lookup table (LUT) to account for the nonlinear phase response of the SLM. This guarantees that the 8-bit grayscale values accurately represent the desired phase map range of 0 to 2π.

OPTICS CONTINUUM

2.3. Optimization methods

We investigate three optimization methods to evaluate the quality of augmentation in the proposed framework. These methods differ in how input data are structured and presented during optimization. Our input data range from a single scene to individual object presentation, and further to numerically shifted image dataset optimization. For all methods, phase optimization is performed using the forward model. This progression enables us to analyze how input diversity affects the system's ability to accurately highlight the target object, maintain robust performance across novel scenes, and operate with a reasonable computational cost.

Single scene optimization (Method 1). We present a simple formulation in Method 1, where the system is optimized using a single input scene. This input scene contains several objects of different sizes or shapes, and the target distribution is defined to highlight only the desired object among the others. Method 1 exhibits correlation-based behavior, where target highlighting depends on the correspondence between the observed FoV and the optimized scene.

Individual object presentation (Method 2). In this method, we insert objects individually rather than having all objects in a single scene. This enables the phase mask to encode target-specific patterns while reducing dependence on surrounding scene content. The number of captured frames required depends on the total number of objects in the scene. Each input scene is processed through the forward model with a corresponding target output. We preserve non-target objects unmodified while selectively highlighting the target object.

Numerically shifted image dataset optimization (Method 3). We increase the dataset variability in the third scenario by introducing spatially shifted versions of the multi-object single scene. These shifts are implemented using a numerical shift operation, without requiring additional frame acquisition. Each shifted input is paired with a corresponding target. The target distribution remains the same object of interest (e.g., square), while its position is shifted consistently with the input. This ensures that the intended augmentation response is preserved in different spatial configurations.

3. Results

We evaluate the phase optimization framework across three representative microscopy augmentation scenarios: (1) object highlighting, (2) size-based selective object highlighting, and (3) morphology-based selective object highlighting. The baseline configuration (Method 1) is first evaluated to examine the performance and limitations of the proposed augmentation framework in the three scenarios. Subsequently, Methods 2 and 3 are introduced to enable more challenging highlighting tasks based on size and morphology. To compute the optimized phase masks, we first experimentally capture the baseline images and then use them as input for our numerical forward model. For all scenarios, we define the optimization target as a synthetic annular ring centered around the object to be augmented. The augmentation response is presented as an annular ring for two main reasons. First, generating the augmentation around the target rather than directly on top of it preserves the visibility of the target morphology and prevents important structural features from being obscured. Second, the annular response can provide a visual indication of the target even when the object is only partially visible within the FoV, since the augmentation extends beyond the object boundaries. This can improve the user's ability to locate the target object and examine its details.

3.1. Scenario 1: Object highlighting

Figure 3 demonstrates the workflow of the optimized phase mask to achieve target augmentation using Method 1. Figure 3(I) presents the offline optimization stage, where the phase mask is optimized using an input image containing a 30 µm diameter circle as shown in Fig. 3(a). Figure 3(b) shows the SLM phase mask (SLM-P), obtained from the optimization process to achieve the target distribution. Figure 3(c) shows the experimental output obtained by displaying the optimized phase pattern on the SLM, while Fig. 3(d) presents the corresponding simulated output generated by the forward model. Figure 3(II) shows the real-time augmentation stage, where Fig. 3(e), (h), and (k) are novel scenes. Pre-SLM-P refers to the precomputed SLM phase, generated during the offline optimization stage and shown in Fig. 3(b). Figure 3(f), (i), and (l) show the experimental outputs, demonstrating the effect of the optimized phase mask in highlighting target objects. Figure 3(g), (j), and (m) present the corresponding simulation outputs obtained using the forward model. Two different exposure time settings are used in the offline optimization and real-time augmentation stages. In the offline optimization stage, we capture both input and experimental outputs with an exposure time of 120 ms. This allows for a direct analysis of how the phase mask affects the redistribution of input light energy to form the target distribution. In the real-time augmentation stage, while the input remains at 120 ms, the experimental output exposure is increased to 200 ms to better utilize the bit depth of the camera during real-time sample manipulation. We evaluate the quantitative metrics between the experimental and simulation outputs across 45 novel scenes. Our findings demonstrate augmentation quality with a mean Cross Correlation (CC) of 0.75 ± 0.03, and Peak Signal-to-Noise Ratio (PSNR) of 27.60 ± 0.61 dB.

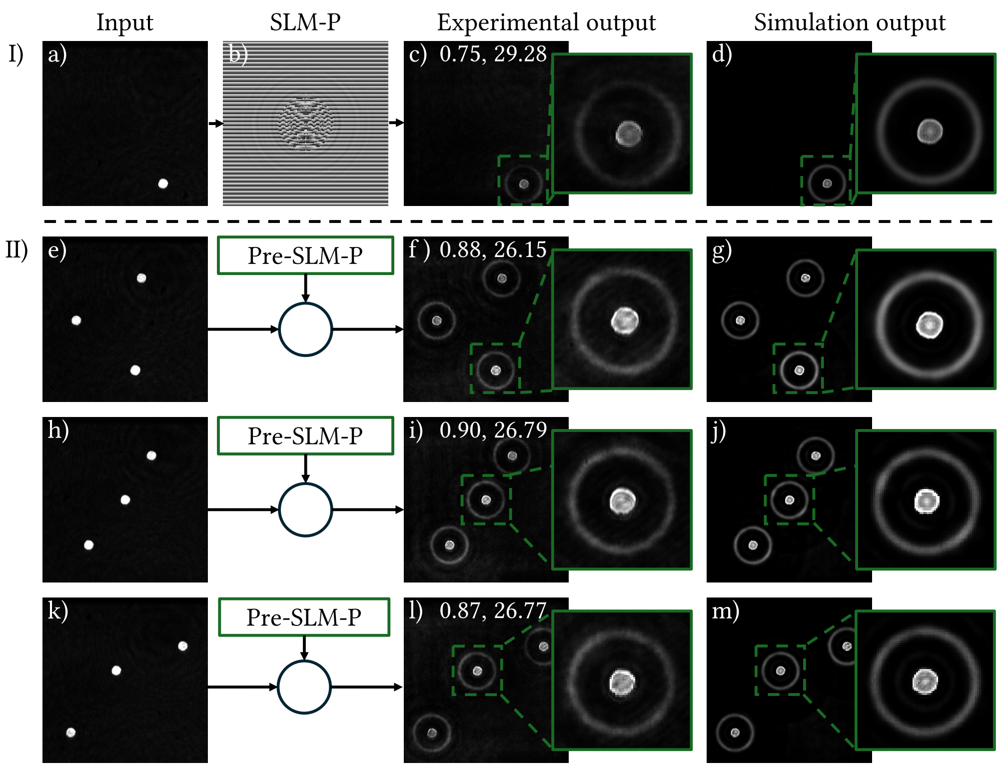

Fig. 3. Scenario 1: Object highlighting. We show the workflow of the optimization and augmentation stages. (I) Offline optimization stage, where the SLM phase is optimized to highlight the target distribution (30 μm diameter circle). (II) Real-time augmentation stage, where the pre-computed SLM phase (Pre-SLM-P) obtained during optimization is reused on novel microscopic frames. We provide CC and PSNR values for each stage. See Visualization 1 for a live demonstration of object augmentation.

OPTICS CONTINUUM

3.2. Scenario 2: Size-based selective object highlighting

We implement a phase mask optimization to highlight circular objects with a target diameter while suppressing augmentation responses from non-target objects with different sizes. Figure 4 shows the offline optimization stage for three different methods implemented to highlight only the 40 µm-diameter circular objects. In Method 1, Fig. 4(a) shows a microscopic frame containing circular objects with diameters of 30, 40, and 50 µm, while Fig. 4(b) presents the target distribution, highlighting the 40 µm-diameter circle. In Method 2, Fig. 4(c)-(e) show circular objects with varying diameters that are individually isolated and presented in each frame. These figures are captured such that no other objects appear in the scene. While Fig. 4(f) highlights the target object, Fig. 4(g,h) remain unmodified. In Method 3, Fig. 4(i)–(l) present examples from a set of 30 numerically shifted frames generated from a single captured microscopic image (Fig. 4(a)), with shifts ranging up to 600 pixels. Figure 4(m)–(q) show the corresponding target highlighting patterns, arranged in the same order as their associated input frames.

Figure 5 presents the evaluation results of the three methods in different microscopic images. For consistent evaluation, the experimental outputs of each method are compared against the ground truth rather than the simulation outputs, which may not exactly match the desired target responses. Figure 5(a) shows the input frame used as an offline optimization stage for Method 1, whereas it serves as a novel scene for Methods 2 and 3. Method 1 demonstrates the best performance compared to the other methods, particularly when evaluated using the same optimization frame, as shown in Fig. 5(b)–(d). While Methods 2 and 3 can highlight the target object, they also slightly augment non-target objects. Figure 5(e) shows a rightward in-plane translation, introduced to the input scene in (Fig. 5(a)). Figure 5(f)–(h) show that Methods 1 and 3 exhibit comparable CC values, indicating similar performance under this displacement condition. Figure 5(i) presents a novel microscopic scene containing 30 and 40 μm circular objects. Figure 5(j) demonstrates that Method 1 is almost unable to highlight the target object. However, Methods 2 and 3, shown in Fig. 5(k) and (l), successfully highlight a 40 μm circular object. Figure 5(m) presents another novel scene containing three 40 μm circular objects. While Method 1, as shown in Fig. 5(n), fails to highlight the target object, Methods 2 and 3 preserve target highlighting. However, Method 2, shown in Fig. 5(p), provides better target highlighting compared to Method 3 in Fig. 5(q). Although Methods 2 and 3 improve target highlighting performance, reflected by higher CC values, PSNR shows only limited variation, as shown in Fig. 5(n)-(q). This is because PSNR is sensitive to global image content and background intensity, while augmentation primarily affects localized target regions, resulting in only minor changes to the overall metric. For this reason, the CC metric was prioritized over PSNR in the evaluation. In summary, Method 1 performs effectively on the optimization scene and the scenes exhibiting high correlation with it. However, its augmentation performance degrades when the correlation between the observed and optimization scene decreases, leading to unsuccessful target augmentation. In contrast, Methods 2 and 3 demonstrate more robust and consistent performance across novel scenes.

3.3. Scenario 3: Morphology-based selective object highlighting

The goal of this scenario is to augment a target object based on its morphological characteristics. We use three previously described optimization methods to generate phase masks that highlight the square shape. The microscopic image used for the optimization stage contains three different geometric shapes (circle, square, and triangle). All shapes are characterized by a size of $\approx$ 40 µm. The procedure of Fig. 6 is analogous to Fig. 5 in Scenario 2. Figure 6(a) presents the input frame, used as the offline optimization frame for Method 1, whereas it serves as a novel scene for Methods 2 and 3. In Fig. 6(b)-(d), Method 1 provides the best augmentation performance compared to the other methods. In Fig. 6(e), we introduce an in-plane rightward translation to the input scene Fig. 6(a). Figure 6(f)–(h) show that all three methods successfully

OPTICS CONTINUUM

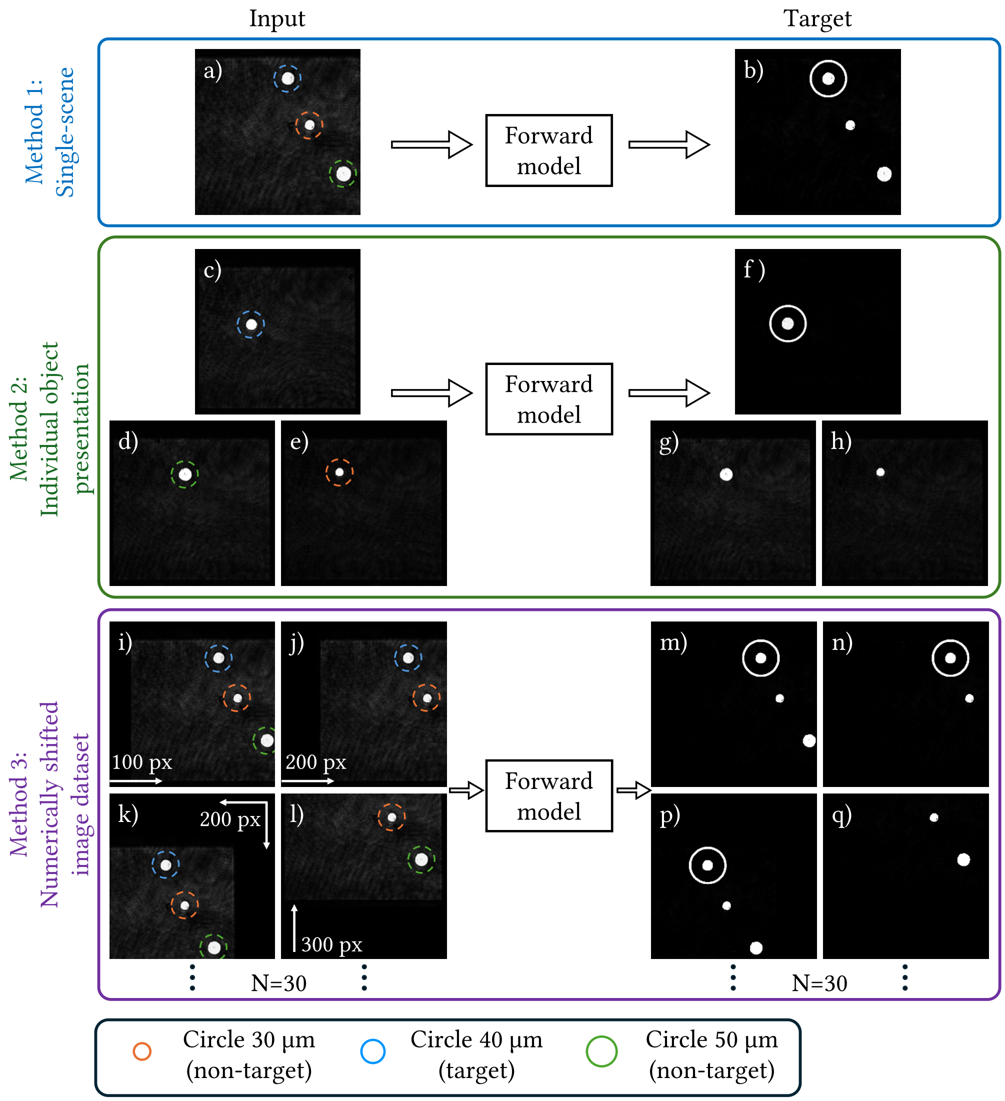

Fig. 4. Comparison of the three optimization methods implemented in Scenario 2 for highlighting a 40 µm diameter target object. Method 1 uses a single scene containing circular objects of different diameters, Method 2 presents each object individually, and Method 3 employs a numerically shifted dataset generated from a single captured image. Each input is paired with its corresponding target in the same order as presented.

OPTICS CONTINUUM

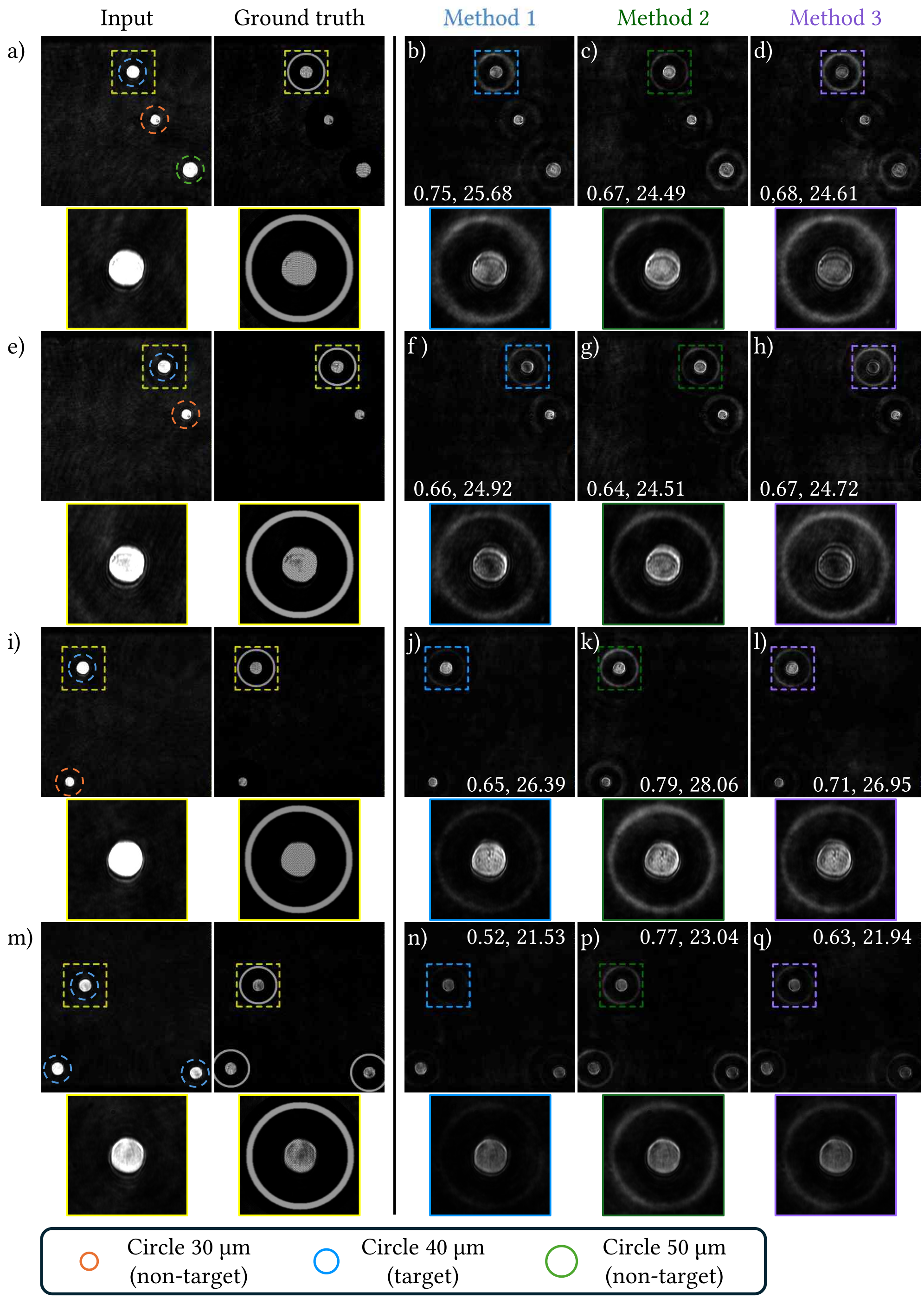

Fig. 5. **Scenario 2: Size-based selective object highlighting.** We present the evaluation results for each optimization method, captured at an exposure time of 200 ms. Each method employs a phase mask to highlight 40 µm diameter circles. The optimized phase is fixed and applied to different novel scenes. We report CC and PSNR values for each method. See Visualization 1 for a live demonstration of augmentation.

highlight the target object. Method 1, shown in Fig. 6(f), achieves the best target highlighting performance; however, its performance deteriorates compared to the optimized scene. In Fig. 6(i), we present a novel microscopic scene for three methods. Figure 6(j) does not fully highlight the target object; however, Fig. 6(k) and (l) clearly highlight the square object. In Fig. 6(m), we introduce an in-plane translation to the Fig. 6(i) scene. Although Method 1, as shown in Fig. 6(n), fails to highlight the desired object, Fig. 6(p) and (q) can augment the target. Overall, Method 1 demonstrates strong performance under consistent optimization conditions; however, its performance weakens in novel scenes that have a low correlation with the optimization scene. In comparison, Methods 2 and 3 maintain greater robustness and adaptability across various observations.

Figure 7 demonstrates the capability of the proposed approach to highlight target objects under partial visibility conditions. Figure 7(a), (d) show input scenes containing visible and partially visible 30 µm diameter circles. Figure 7(b), (e) present the experimental output after applying the pre-computed phase mask implemented using Method 1, where the system highlights targets with partial visibility. Figure 7(c), (f) show a leftward in-plane translation applied to the experimental output to confirm the presence of the target object. Figure 7(g) demonstrates the input scene containing a 50 µm diameter and a 40 µm diameter circle with incomplete visibility. Figure 7(h) presents the effect of applying the optimized phase mask to selectively highlight a target defined as a 40 µm diameter circle using Method 2, despite its incomplete appearance in the input scene. Figure 7(i) shows an in-plane leftward shift applied to the scene Fig. 7(g) to verify the existence of the target object. Figure 7(j) demonstrates the input scene containing 40 µm diameter and a partially visible square shape. Figure 7(k) presents the experimental output after displaying a phase mask optimized using Method 3 to highlight square objects. Figure 7(l) shows the output after applying a rightward in-plane translation to the input scene Fig. 7(j). Our results confirm the capability of the proposed work to achieve target highlighting even when the target is not fully observable.

Table 1 summarizes the quantitative evaluation metrics for the three methods in size-based and morphology-based scenarios. The reported metrics represent the mean values obtained across 45 novel scenes. In general, Methods 2 and 3 achieve higher CC values than Method 1, indicating improved robustness in highlighting the target in novel scenes. Although Methods 2 and 3 exhibit comparable performance in both scenarios, Method 2 demonstrates better results. Regarding PSNR values, we notice relatively minor variations across all methods. This behavior is expected, as PSNR is a widely used image quality metric and may be less sensitive to localized target highlighting (high-frequency changes) due to the dominant black background in the scene.

Table 1. Quantitative performance comparison across size- and morphology-based highlighting scenarios using three optimization methods.

Method	Size-Based		Morphology-Based	
	CC ↑	PSNR ↑	CC ↑	PSNR ↑
Method 1	0.60	23.77	0.59	25.01
Method 2	**0.68**	**24.07**	**0.71**	**25.84**
Method 3	0.65	23.76	0.70	25.57

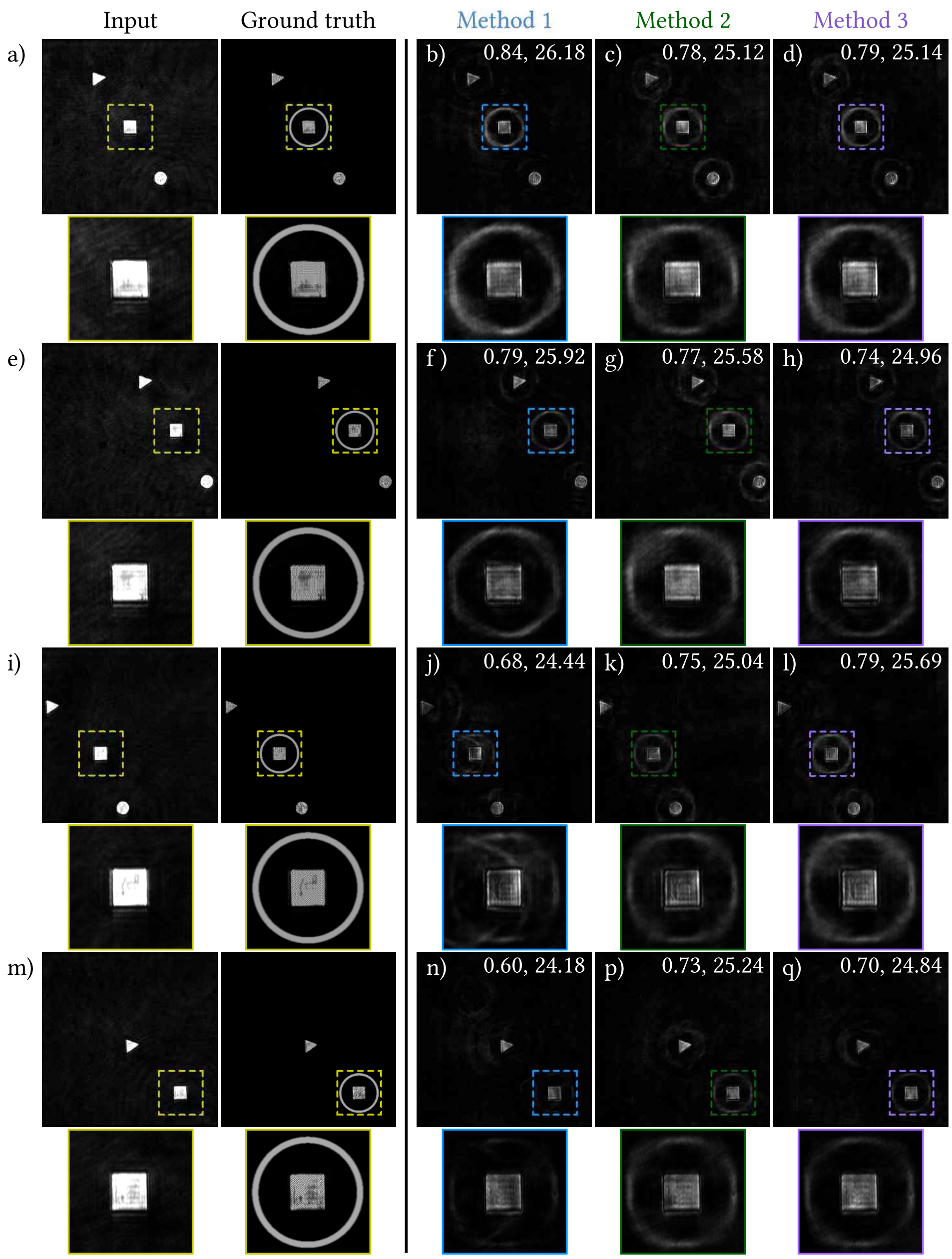

Fig. 6. Scenario 3: Morphology-based selective object highlighting. We present the results for each optimization method, captured at an exposure time of 200 ms. Each method employs a phase mask to highlight the square shape. The optimized phase is fixed and applied to different novel scenes. We report CC and PSNR values for each method. See Visualization 1 for a live demonstration of augmentation.

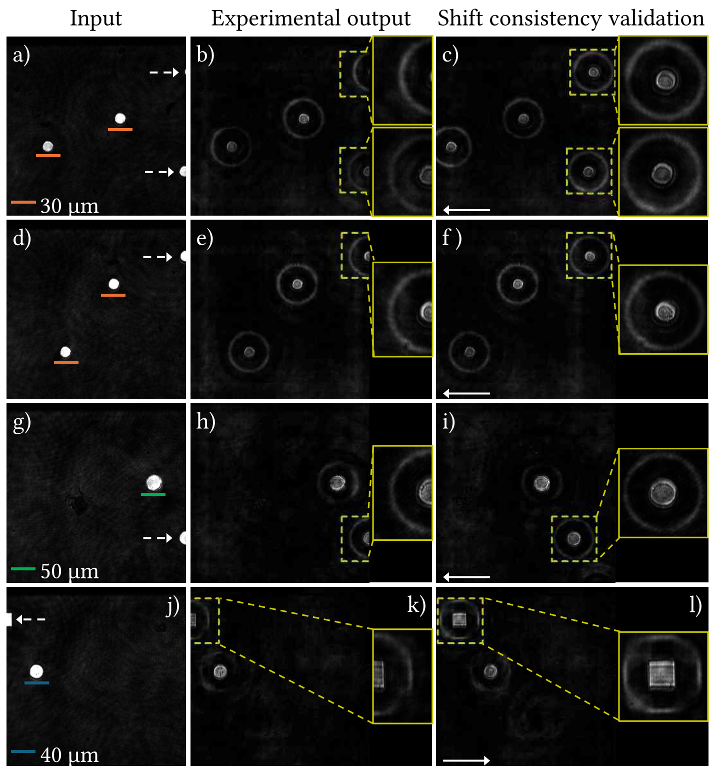

Fig. 7. Demonstration of target augmentation under partial visibility for different scenarios. The input scenes (left) contain partially visible targets, while the experimental outputs (middle) show target-selective highlighting using optimized phases. We include shifted target cases (right) to confirm the presence of the true target object and verify that the augmentation response correctly matches the intended target. The white dashed arrow indicates the estimated location of the partially visible portion of the target object.

4. Discussion

Comparison of optimization methods. The results obtained in the three scenarios demonstrate that three optimization methods exhibit different tradeoffs between augmentation quality, robustness to novel scenes, and computational complexity. In Scenario 1, Method 1 can successfully augment circular objects in both the optimization and the novel scenes using a single scene during the offline optimization. In Scenario 2, Method 1 effectively highlights the target object in the optimization scene and maintains selective highlighting performance for scenes containing translated versions of the optimization scene. However, when the sample undergoes significant in-plane translation, new objects may enter the field of view while others leave it, altering the spatial frequency distribution of the scene. In such cases, the correlation with the originally optimized field may decrease, resulting in undesired augmentation responses or a failure to highlight the target object. This behavior suggests that Method 1 can be applied in the optimization scene, but exhibits limited generalization capability. In contrast, Methods 2 and 3 demonstrate improved target augmentation due to the increased diversity of input observations and scene variations. This enables the system to optimize a phase mask with respect to objects without being dedicated to a single optimization frame. In addition to the trade-offs associated with the optimization methods, the annular augmentation response presents both advantages and constraints. The annular augmentation provides clear target highlighting without obscuring the object, even in cases where the object is partially visible. However, its highlighting performance may decrease when multiple targets are located in proximity due to overlapping responses.

Computational perspective. We observe significant differences among the three optimization methods in terms of processing requirements. The offline optimization stage is performed using *CUDA* on an *NVIDIA GeForce RTX 5080* GPU. Method 1 provides the fastest optimization process, requiring only ≈ 3 minutes to optimize the phase mask. In comparison, Method 2 requires ≈ 8 minutes, as in Scenarios 2 and 3, three different objects are introduced for optimization. Therefore, increasing the number of object types leads to higher computational processing. Method 3 gets a significant increase in optimization time to nearly ≈ 7.5 hours. Although Method 3 may provide effective augmentation performance across different novel scenes, the substantial increase in optimization time introduces a practical tradeoff between computational cost and augmentation quality. Overall, the results indicate the strengths and limitations of using each method. Method 1 is suitable for fast, single-scene-specific optimization, whereas Methods 2 and 3 provide improved robustness for novel scenes at the expense of increased computational complexity. Therefore, the optimization method used depends on the intended application requirements, taking into account the trade-off between speed and performance.

Laser safety considerations. We use a laser light source in our optical layout, taking into account the Maximum Permissible Exposure (MPE) limits defined by IEC 60825 laser safety standard, described in the laser safety manual [25]. We employ a 5 mW laser (class 3B) that is expanded with a 25 mm diameter lens. This provides a low-power laser with an irradiance of E = p/A = 10.2 µW mmˆ{−2} in the sample plane. In the camera plane, the optical power is distributed over a 100 times wider area due to the 10X objective lens. In addition, the power is decreased by a double pass through the beam splitter. This results in an approximately 400-fold reduction in the total irradiance, excluding the diffraction efficiency of the SLM, to become 25.5 nW mmˆ{−2}. Therefore, we ensure that the irradiance remains below the MPE limit. This allows safe direct observation without causing any biological damage to the retina.

Dark-field implementation and bright-field considerations. We experimentally prove the validity of our proposed method only under dark-field illumination. In this approximation, the image composition consists mainly of scattered light components where the unscattered background is suppressed. This enhances the contrast of the scattering structures but reduces the sensitivity to pure phase variations that do not significantly redistribute optical energy. In the bright field illumination, the modulated final pattern occurs due to the interference between the scattered and unscattered light fields. This results in a complex contribution where the phase variation should be taken into account before implementing our forward model. Therefore, for future extensions, bright-field implementations may explore improving the sensitivity for phase-dominant samples.

Toward metasurface-based implementation. In our work, we use the SLM to display the optimized phase map to provide a flexible proof-of-concept platform. However, this requires increased optical component complexity and cost due to the integration of the SLM with an external $4f$ system. This can be addressed by incorporating passive optical elements such as metasurfaces into our system. Metasurfaces are nanostructured surfaces that are engineered to enable precise spatial phase and polarization control with subwavelength resolution [26]. This facilitates its use in unpolarized light applications, where incident light can be controlled to increase phase modulation and diffraction efficiency [27]. Furthermore, unlike SLM, metasurfaces are passive optical elements that do not require any power consumption. Placing a metasurface between the microscope objective and the tube lens can reduce the system complexity and cost while preserving the functionality of our method. These capabilities make metasurfaces a powerful platform for multifunctional beam control that can simplify and improve our approach. Nevertheless, the passive nature of metasurfaces also introduces a limitation in terms of adaptability, as the pre-computed phase response cannot dynamically adjust to substantially varying biological samples. Consequently, augmentation performance may decrease when the observed sample significantly deviates from the offline optimized input frames used to generate the phase mask. In such cases, applying optimization by incorporating a broader range of biological sample variations can maintain performance.

5. Conclusion

Our catalog of methods enables real-time target highlighting for AR microscopes with noticeably reduced computational cost by eliminating the need for per-frame computation in conventional approaches. Additionally, our work exploits the widely used inherent shift-invariant property to provide continuous augmentation during object movement, avoiding per-frame coordinate estimation or an explicit tracking algorithm. We develop three different optimization methods that present distinct trade-offs in target highlighting performance. Method 1, based on single-scene optimization, is the simplest and least computationally intensive approach but offers limited robustness under conditions where observations have low correlation with the optimized scene. Method 2, which presents objects individually, enhances robustness across novel scenes; however, it requires more frame capturing as the number of objects increases, resulting in a higher computational cost compared to Method 1. Method 3, which utilizes a numerically shifted image dataset from a single captured image, improves the quality of the augmentation, although it comes with a higher computational overhead. Quantitative evaluation across scenarios 2 and 3 shows that Method 1 achieves a mean CC value of 0.60, while Methods 2 and 3 improve it to 0.70 and 0.68, respectively, indicating enhanced robustness and target highlighting performance. Furthermore, we demonstrate the potential of our model in highlighting partially visible targets through annular ring-based optical augmentation. This behavior can facilitate faster target highlighting compared to approaches requiring complete object visibility. Our implementation is suitable for various domains, especially health-related applications. We believe that our model can benefit medical students in pathology courses by enhancing cell morphology imaging, promising to improve the speed and quality of their education.

Funding. Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (123N774); Turkish Academy of Sciences.

Acknowledgment. This work is supported by TÜBİTAK under project number 123N774. M. F. Toy acknowledges support from the Turkish Academy of Sciences - Outstanding Young Scientist Award Program (TUBA - GEBIP).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are available upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

References

1. P. J. Tadrous, "PUMA – an open-source 3D-printed direct vision microscope with augmented reality and spatial light modulator functions," J. Microsc. **283**(3), 259–280 (2021).
2. J. R. Watson, C. F. Gainer, N. Martirosyan, *et al.*, "Augmented microscopy: real-time overlay of bright-field and near-infrared fluorescence images," J. Biomed. Opt. **20**(10), 106002 (2015).
3. D. Jin, J. H. Rosenthal, E. E. Thompson, *et al.*, "Independent assessment of a deep learning system for lymph node metastasis detection on the augmented reality microscope," Journal of Pathology Informatics **13**, 100142 (2022).
4. P.-H. C. Chen, K. Gadepalli, R. MacDonald, *et al.*, "An augmented reality microscope with real-time artificial intelligence integration for cancer diagnosis," Nat. Med. **25**(9), 1453–1457 (2019).
5. M. Mohamed, Y. Ezaki, Y. Kawagoe, *et al.*, "Ai-based augmented reality microscope for real-time sperm detection and tracking in micro-tese," Bioengineering **13**(1), 102 (2026).
6. S. Badve, G. L. Kumar, T. Lang, *et al.*, "Augmented reality microscopy to bridge trust between AI and pathologists," npj Precis. Onc. **9**(1), 139 (2025).
7. N. L. Kazanskiy, M. A. Butt, and S. N. Khonina, "Optical computing: status and perspectives," Nanomaterials **12**(13), 2171 (2022).
8. A. Drăgulinescu, "Optical correlators for cryptosystems and image recognition: a review," Sensors **23**(2), 907 (2023).
9. J. W. Goodman, *Introduction to Fourier Optics* (Roberts and Company Publishers, 2005), 3rd ed.
10. T. Fu, J. Zhang, R. Sun, *et al.*, "Optical neural networks: progress and challenges," Light: Sci. Appl. **13**(1), 263 (2024).

11. A. V. Lugt, "Signal detection by complex spatial filtering," IEEE Trans. Inf. Theory **10**(2), 139–145 (1964).
12. B. V. K. V. Kumar, A. Mahalanobis, and R. D. Juday, *Correlation Pattern Recognition* (Cambridge University Press, 2005).
13. A. Dragulinescu and D. Cojoc, "Optical correlators: systems and domains of applications," in *Advanced Topics in Optoelectronics, Microelectronics, and Nanotechnologies II*, vol. 5972 (SPIE, 2005), p. 59721F.
14. T. Manzur, J. W. Zeller, and S. Serati, "Optical correlator based target detection, recognition, classification, and tracking," Appl. Opt. **51**(21), 4976–4983 (2012).
15. A. Mahalanobis, B. V. K. Vijaya Kumar, and D. Casasent, "Minimum average correlation energy filters," Appl. Opt. **26**(17), 3633–3640 (1987).
16. Y. Zhong, L. Chen, W. Gan, *et al.*, "Image encryption system based on joint transformation correlation and ptychography," IEEE Photonics J. **12**, 1–10 (2020).
17. T. Hoshizawa, Y. Honda, S. Kodama, *et al.*, "Holographic optical correlator-based single-pixel imaging with deep learning image reconstruction," Opt. Continuum **4**(2), 333–345 (2025).
18. K. Kavaklı, Y. Itoh, H. Urey, *et al.*, "Realistic defocus blur for multiplane computer-generated holography," in *2023 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, (2023), pp. 418–426.
19. C. Maurer, A. Jesacher, S. Bernet, *et al.*, "What spatial light modulators can do for optical microscopy," Laser Photonics Rev. **5**(1), 81–101 (2011).
20. J. W. Goodman, "Operations achievable with coherent optical information processing systems," Proc. IEEE **65**(1), 29–38 (1977).
21. C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).
22. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Phys. D **60**(1-4), 259–268 (1992).
23. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *International Conference on Learning Representations (ICLR)* (2015).
24. A. Paszke, S. Gross, S. Chintala, *et al.*, "Automatic differentiation in pytorch," *NIPS-W* (2017).
25. International Electrotechnical Commission, "Safety of laser products – part 1: Equipment classification and requirements," IEC 60825-1, Edition 1.2 (2001).
26. H.-T. Chen, A. J. Taylor, and N. Yu, "A review of metasurfaces: physics and applications," Rep. Prog. Phys. **79**(7), 076401 (2016).
27. N. Yu and F. Capasso, "Flat optics with designer metasurfaces," Nat. Mater. **13**(2), 139–150 (2014).