photonics

MDPI

Article

A Dataset Reconstruction Method Based on Real Optical Aberrations

Libing Jin [1,2], Jiangyu Tian [1,3] and Jun Chang [1,*]

[1] School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China
[2] Beijing Institute of Space Mechanics & Electricity, Beijing 100094, China
[3] Institute of Remote Sensing Satellite, China Academy of Space Technology, Beijing 100094, China
* Correspondence: optics_chang@126.com

Abstract

This study proposes a dataset reconstruction method for adapting object detectors to the aberrations of a tunable liquid-lens system. Existing labeled images are reconstructed using field-dependent optical responses and used for detector training. Experiments on VOC2007 compare system-PSF reconstruction with clean training and fitted global and field-dependent Gaussian blur, using two detector architectures and three random seeds. For SSDLite, PSF training improves mean AP50 over states 1–3 by 2.33 percentage points compared with field Gaussian, with a descriptive 95% interval of [1.54, 3.06]. The improvement reaches 2.05 points on a held-out optical state. An exploratory evaluation of 24 annotated hardware images also gives the highest pooled person AP50 for PSF training. These results support the value of system-specific reconstruction for the tested SSDLite configuration, with performance depending on the detector and imaging conditions.

Keywords: optic aberration; deep learning; dataset reconstruction

1. Introduction

With the continuous advancement of artificial neural network technology, data-driven image processing methods have gradually become mainstream in the field of image processing [1]. These methods focus more on the data itself compared to traditional model-driven image processing methods, as the quality of the dataset directly determines the output quality of image processing algorithms [2]. A high-quality dataset not only includes a large number of samples but also features good annotation quality and reasonable sample balance to ensure the model's generalization capability and robustness.

However, the impact of optical imaging devices on image quality is often overlooked when constructing image datasets. Different photoelectric imaging devices can introduce various optical aberrations, such as spherical aberration, chromatic aberration, and distortion, which significantly affect the performance and robustness of image processing algorithms [3]. It is generally believed that images with different aberration characteristics can enhance the generalizability of deep learning networks, especially in surveillance areas [4]. Traditional methods for reconstructing datasets consider the impact of optical imaging devices and strive to cover images produced by different devices to improve the diversity and representativeness of the dataset [5].

In practical applications, a detector is usually deployed on a specific optical system, whose image degradation may differ from that in a general dataset. Image restoration, including super-resolution [6], seeks to improve the recorded image. Adapting the training

Check for updates

Received: 12 October 2026
Revised: 8 November 2026
Accepted: 18 November 2026
Published: 23 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

data to the target optics provides another way to reduce this mismatch. This work studies the latter approach. This work studies the latter approach. This work focuses on diffraction and aberration blur. Distortion is limited by the optical design, and local PSFs are normalized before reconstruction [7,8]. Spatial illumination falloff and the complete sensor-processing pipeline are not included. The PSF describes the intensity response of an optical system to a point source. It can be measured experimentally [9,10] or calculated from the propagated complex field [8], and is widely used in optical image restoration [7]. A tunable liquid-lens system is used to produce reconstructed images with different aberration states. The YOLO family formulates object detection as a unified prediction task [11]. Detection performance is evaluated using YOLOv4 [12], followed by controlled SSDLite and YOLOv8n experiments with Gaussian-blur baselines. The workflow is shown in Figure 1.

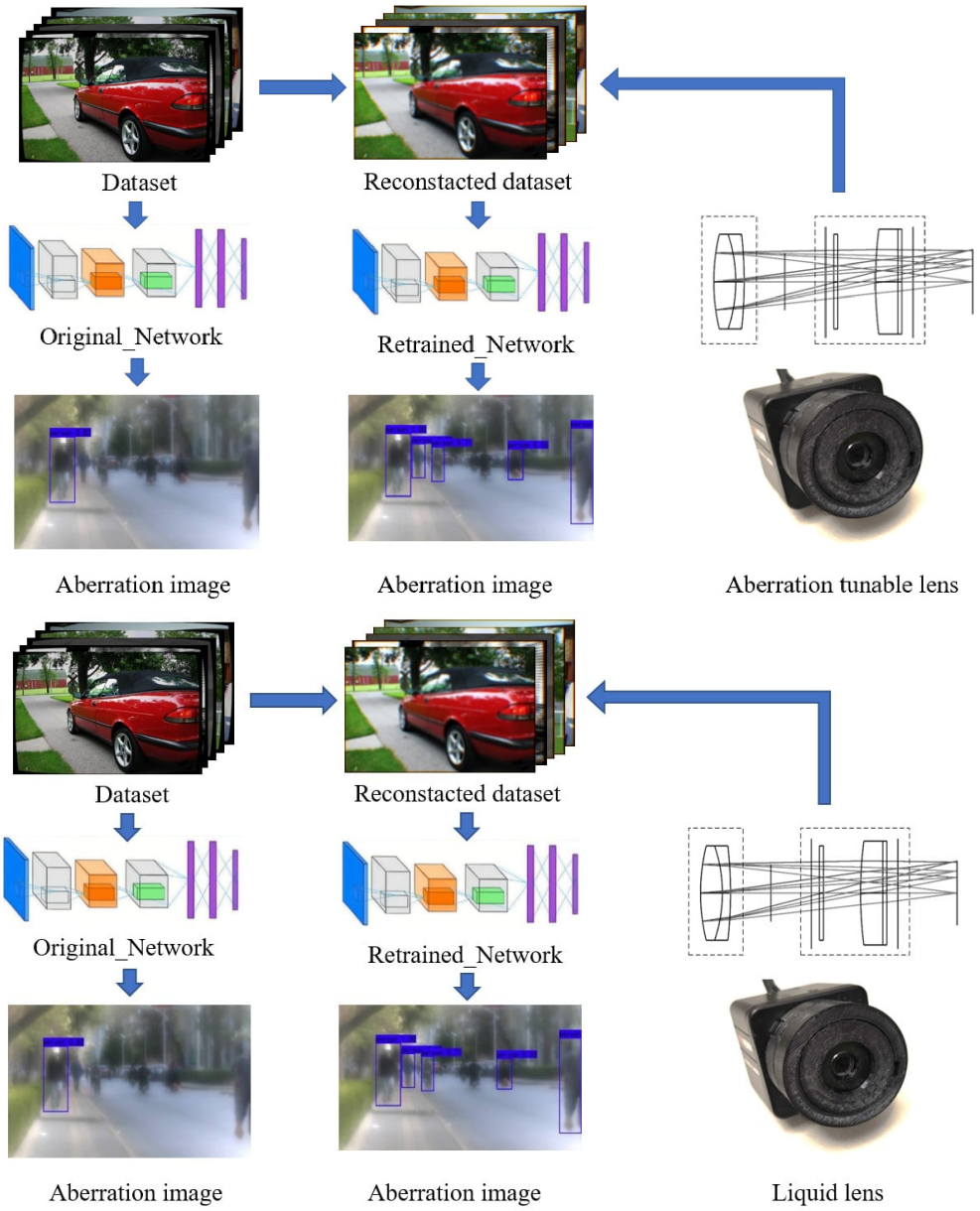


Figure 1. Dataset reconstruction method.

The method is intended for detector adaptation to a specified optical imaging system. Previous studies have investigated optical-kernel augmentation and optical-corruption benchmarks [13,14]. Here, reconstruction is driven by the field-dependent response of

a specified liquid-lens system. The experiments assess its detection benefit over fitted Gaussian approximations, including an optical state excluded from training.

2. Materials and Methods

2.1. Image Reconstruction

For a large-scale image dataset, like the VOC dataset [15], the images are often provided by a variety of imaging devices, and those imaging devices have different imaging characteristics. According to the linear system theory, the images in the dataset I_1 can be approximated as the convolution of the ideal image I_0 and the point spread function PSF(n) of the shooting device [7,16], where n denotes the image index. For all the images in the dataset $[I_1(n)]$, there are the following:

$$[I_1(n)] = [I_0(n)] * [\mathrm{PSF}(n)], \tag{1}$$

The source-camera responses are treated as heterogeneous factors in the dataset. Applying the target response retains the original camera degradation and provides a forward approximation to the target imaging conditions. This approximation is most relevant when the source images are relatively sharp and the target aberrations dominate at the detector scale; increasing dataset size does not itself make the source responses identity operators.

Let the source-camera responses form a collection of heterogeneous optical operators, as follows:

$$\sigma = [\mathrm{PSF}(n)], \tag{2}$$

For devices with specific optical transfer properties $\mathrm{OTF_s}$, whose reconstructed dataset $[I_\mathrm{R}(n)]$, there are the following:

$$[I_\mathrm{R}(n)] = [I_1(n)] \cdot \mathrm{PSF_s} = [I_0(n)] * \mathrm{PSF_s} * [\mathrm{PSF}(n)], \tag{3}$$

Assuming the device takes an image I_S of the original image I_0, there are the following:

$$[I_S(n)] = [I_0(n)] * \mathrm{PSF_s}, \tag{4}$$

Bringing Equation (3) into the reconstructed dataset $[I_\mathrm{R}(n)]$, there is the following:

$$[I_\mathrm{R}(n)] = [I_\mathrm{S}(n)] * \sigma \tag{5}$$

Equations (3) and (4) distinguish reconstruction of a recorded image from imaging an ideal scene. The spatially varying form used for reconstruction is given in Equation (5).

The local PSF varies with field position and liquid-lens state, as shown in Figure 2. Reconstruction therefore uses a field-dependent response, with local convolution approximating image formation within each small region.

2.2. Ray Tracing

Vector ray tracing is used to trace rays through the optical system. For a ray starting at $\boldsymbol{o} = (x_o, y_o, z_o)$ with unit direction vector $\vec{\boldsymbol{d}}$, the ray equation is as follows:

$$\vec{\boldsymbol{r}}(t) = \mathbf{o} + t\vec{\boldsymbol{d}} \ (0 \le t < \infty), \tag{6}$$

where t is the transmission factor of $\vec{\boldsymbol{r}}$.

For a point $P(x_P, y_P, z_P)$ of the circle on the optical axis whose center is $O_l(0,0,l)$:

$$x_P{}^2 + y_P{}^2 + (z_P - l)^2 = R^2, \tag{7}$$

The parameter $\mathrm{t_p}$ is obtained by combining Equations (6) and (7), which generally give two intersection points (Figure 3). The first valid intersection along the propagation direction is selected.

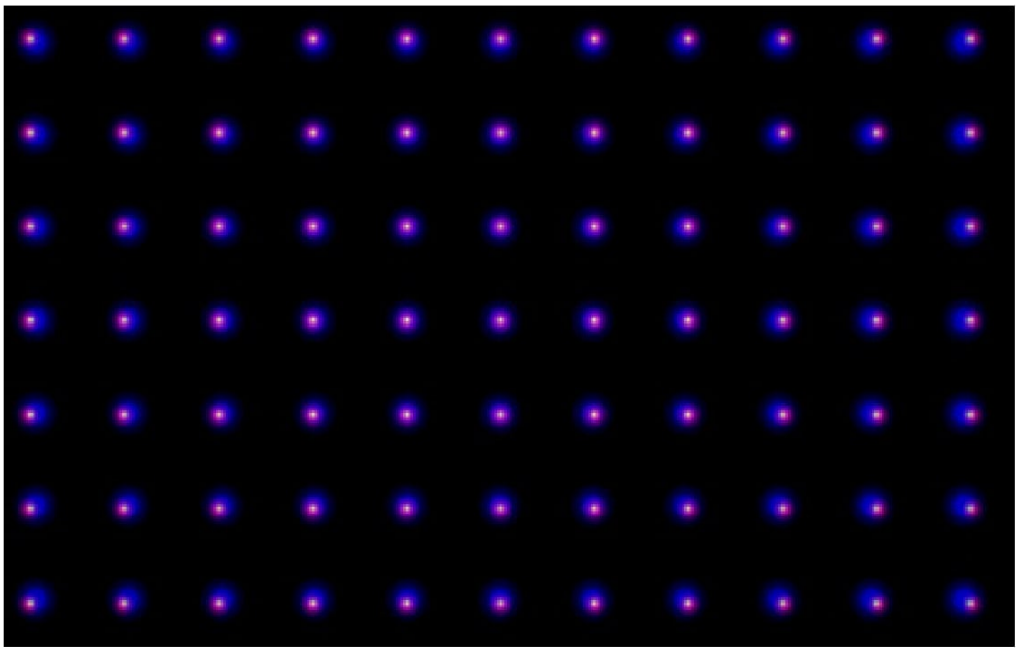


Figure 2. Field-dependent PSF array at the representative RGB wavelengths.

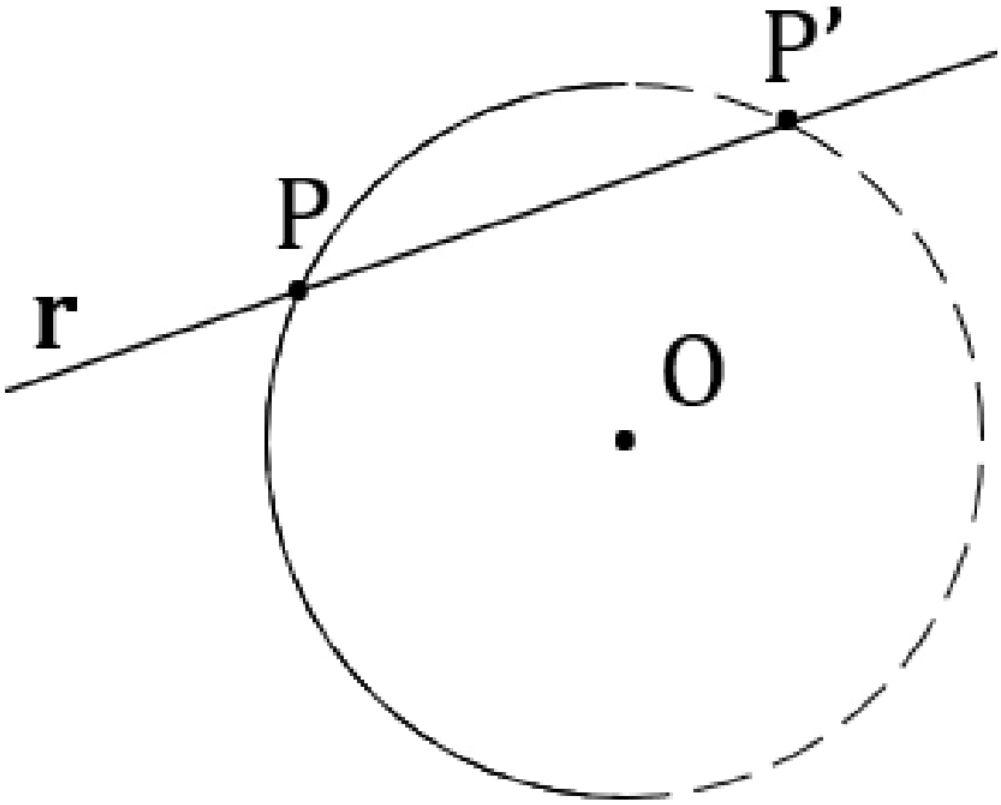


Figure 3. Ray–sphere intersection geometry.

Light transferring from refractive indexes n_1 to refractive indexes n_2 satisfies the Snell's law, as follows:

$$\sin\theta_i = \frac{n_2}{n_1} \times \sin\theta_r, \tag{8}$$

where θ_i is the incidence angle and θ_r is the refraction angle.

Bring in the light vector $\vec{d}$ into Equation (8). For the light vector $\vec{d}'$ propagating backward from point P, there is the following:

$$\sin < \vec{PO}, \vec{d}' >= \frac{n_2}{n_1} \sin < \vec{PO}, \vec{d} >, \tag{9}$$

Evaluate the unit vector $\vec{e}$ from $\vec{PO}$ as follows:

$$\vec{e} = \frac{\vec{PO}}{\left|\vec{PO}\right|}, \tag{10}$$

Unit vectors $\vec{d}'$, $\vec{\mathbf{e}}$, which are located in the same plane, can substituted into the Equation (9). Then the light vector $\vec{d}'$ can be solved, as follows:

$$\vec{d}' = \frac{n_2}{n_1}\vec{d} + \left(\frac{n_2}{n_1}\cos < \vec{PO}, \vec{d} > -\sqrt{1 - \frac{{n_2}^2}{{n_1}^2}\left(1 - \cos^2 < \vec{PO}, \vec{d} >\right)}\right)\vec{e}, \tag{11}$$

According to this method, the light path $L_{x,y}$ from any point of the object surface to any surface of the optical system can be solved, as follows:

$$L_{x,y} = \sum |d_n| \times n_n, \tag{12}$$

where d_n is the path of the light in the direction $\vec{d}$ of a tracing process, and n_n is the refractive index of the intermediate substance of the tracing process.

2.3. Reconstruction Model

For optical systems with large aberrations, more accurate information can be obtained by taking the exit pupil position as the reference wavefront. The direction and light path of a point (x_0,y_0,z_0) at the exit pupil can be obtained by Equation (11) and Equation (12), which is satisfied for the spherical divergent field $U_0(x_0,y_0)$, as follows:

$$U_0(x_0, y_0) = a_0 \frac{\exp\left(jkL_{x_0,y_0}\right)}{j\lambda L_{x_0,y_0}}, \tag{13}$$

where L_{x_0,y_0} is the forward light path from different image points to this exit pupil point, and $k = 2\pi/\lambda$ is the wave number of the light.

According to simplified Kirchhoff diffraction, its field distribution $U_i\left(x_0, y_0; x_i, y_i\right)$ on the image plane after forward propagation along the optical axis can be described as follows:

$$U_{\mathrm{i}}(x_0, y_0; x_{\mathrm{i}}, y_{\mathrm{i}}) = U_0(x_0, y_0) \frac{\exp(jkr)}{j\lambda r} \tag{14}$$

Here r is the Euclidean distance from the exit-pupil point to the image-plane point. Integration over the exit pupil gives the following:

$$U_{\mathrm{i}}(x_{\mathrm{i}}, y_{\mathrm{i}}) = \iint_{\mathrm{pupil}} U_0(x_0, y_0) \frac{\exp(jkr)}{j\lambda r} \, \mathrm{d}x_0 \mathrm{d}y_0 \tag{15}$$

In the actual energy calculation process, in order to simplify the calculation, the compound field $U_i(x,y)$ calculation should use the summation of pupil sampling points, and its energy $I(x_i,y_i)$ can be computed by the following:

$$I(x_i, y_i) = |U_i(x_i, y_i)|^2, \tag{16}$$

The sampled intensity is normalized over the kernel support to obtain a unit-sum discrete PSF, as follows:

$$PSF(x,y) = \frac{I(x,y)}{\sum_{u,v} I(u,v)} \tag{17}$$

Normalization retains the spatial intensity distribution while removing an arbitrary overall scale. The support must also be sufficient to represent the PSF.

In the process of simulating the image reconstruction of a real optical imaging system using the Huygens PSF, it is necessary to fully consider the machining and alignment tolerances of the optical system, the reconstruction of the PSF for each pixel, and the integration effect of the optical detector on wavelengths. All of these factors can impact the accuracy of image reconstruction.

The tunable optical system described in Section 3 provides a range of aberration states for evaluating the reconstruction method. The PSF is sampled on a fine optical grid and mapped to the image-pixel scale. Reconstruction is performed offline and does not add optical convolution during detector inference. RGB reconstruction is implemented at representative wavelengths of 650, 510, and 440 nm, using unit channel weights. This is a three-wavelength approximation to the spectral response.

3. Optical System

An aberration tunable lens is designed for experiments with range-adjustable aberration, achieved through a liquid lens. Liquid lenses are usually supported by a composite liquid; the surface shape is similar to the liquid surface and can carry out a certain deformation under the control of current. The focal length of the optical lens is determined by the curvature radius of the surface of each element of the lens, so the liquid lens can regulate the focal length of the system within a certain range. However, the liquid lens easily introduces significant spherical aberration in the process of controlling the focal length of the system [17]. At the same time, in the imaging Angle, the essence of the liquid lens can be considered as a lens whose refractive index changes with the wavelength, so a certain chromatic aberration will also be caused. This kind of aberration change characteristic of an optical system based on a liquid lens is very suitable for the verification of the dataset reconstruction method proposed in this paper.

The system adopted the OPTOTUNE (Dittikon, Swiss) EL-10-30-TCliquid lens, which offers fast response, operational stability, and a compact form factor (10 mm aperture, 3.3 mm thickness). The achromatic correction was provided by a 30 mm focal length cemented doublet lens from Daheng Optics (Beijing China), which effectively reduces chromatic dispersion. The image sensor used was the OV2710, which supports high frame rate imaging, features a dynamic range up to 56 dB, and maintains stable performance even under low-light conditions (0.051 lux), making it well suited for alignment and testing with the optical system.

The mechanical housing of the lens was fabricated via 3D printing, ensuring ease of alignment and assembly. The system exhibited distinct and controllable aberration characteristics across the field of view, with stable imaging performance during focus adjustments through the liquid lens. The system configuration and specifications are shown in Figure 4 and Table 1.

Different liquid-lens settings are selected to cover representative aberration variations within the working range. Adjustment changes the focus and rear focal position, producing different degradation patterns. Figures 5 and 6 show the corresponding optical responses. These settings represent optical states rather than measured temperature or voltage calibrations.

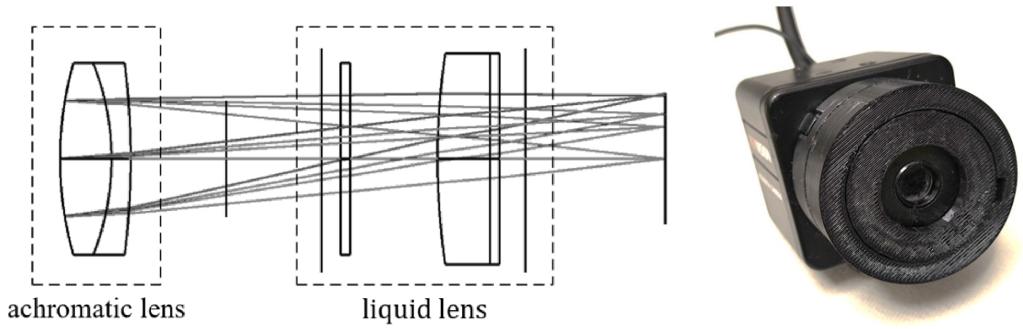


Figure 4. Experimental optical system.

Table 1. System indicators.

Parameter	Value
Focus Distance	23 mm
Full field of view	13°
F number	3
Characteristic frequency	167 lp/mm

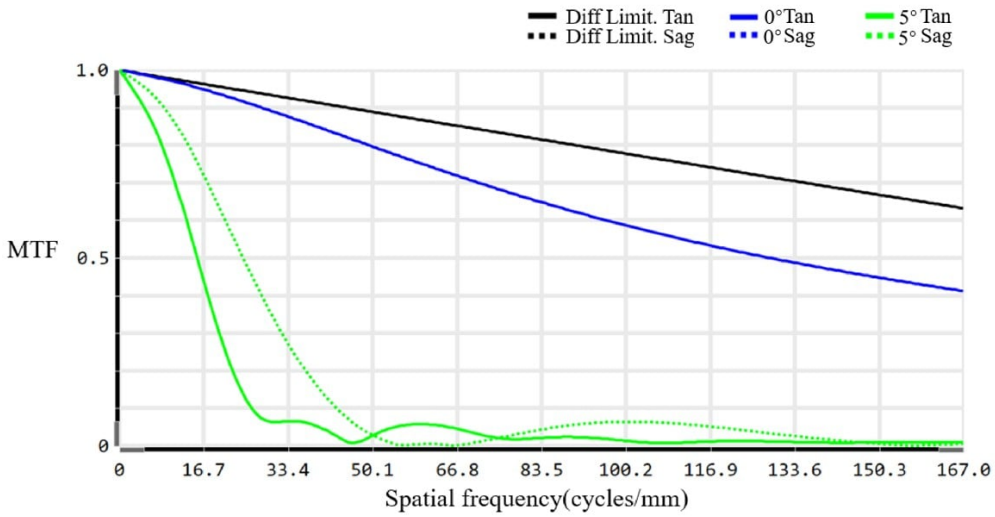


Figure 5. MTF when focusing on the zoom position of object field 0°.

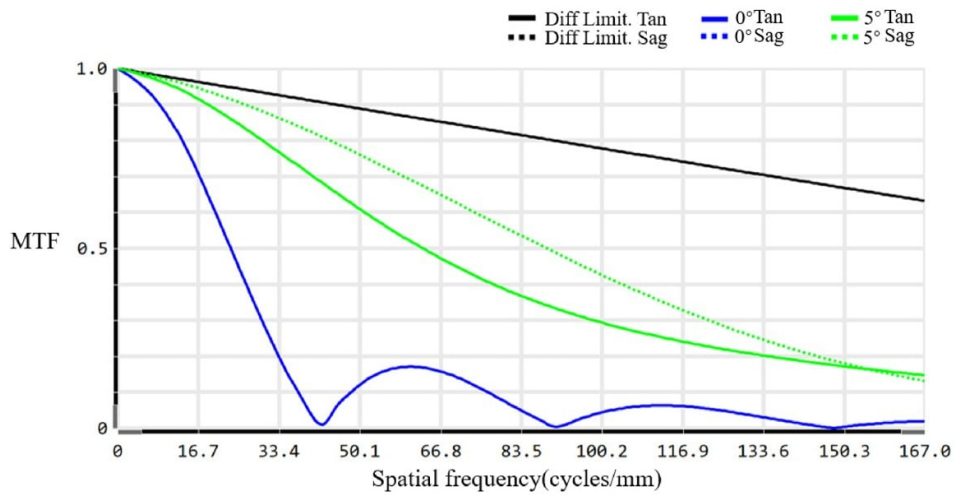


Figure 6. MTF when focusing on the zoom position of object field 5°.

As shown in Figure 7, due to the presence of field curvature in the optical system, defocusing the center field of view causes peripheral rays to converge near the original rear focal plane. This effect, induced by field curvature, allows the peripheral regions to exhibit relatively improved imaging quality when the system is focused away from the center.

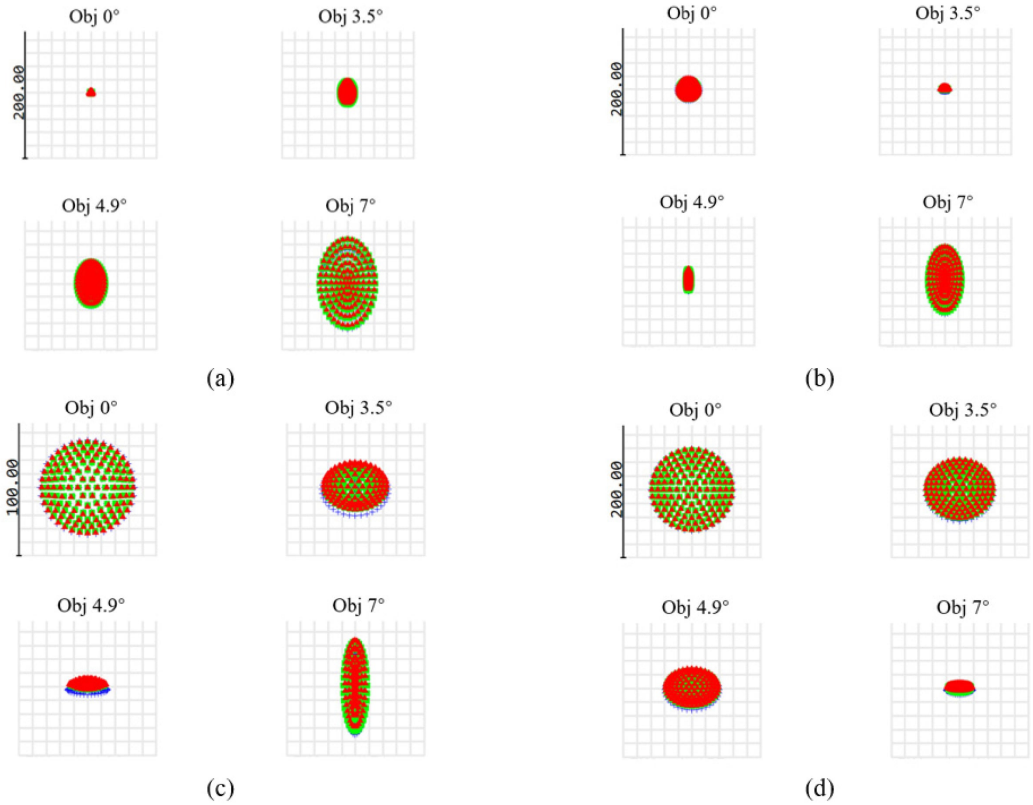


Figure 7. Spot diagrams for four focus settings. Panels (**a**–**d**) correspond to focusing at object-field angles of 0°, 3.5°, 4.9° and 7°, respectively. The horizontal and vertical axes show spot coordinates in micrometers. Red, green, and blue indicate the three representative wavelengths.

Table 2 presents the variations in optical aberrations, expressed in terms of Zernike polynomials, across different focus positions for a 4.9° field of view; Zoom1 to Zoom4 represent the system focusing on object field of 0°, 3.5°, 4.9°, and 7°. It can be observed that the dominant aberration introduced by the liquid lens is defocus, which is consistent with the system's design intent. In addition, changes in lens curvature lead to corresponding variations in higher-order aberrations, including primary spherical aberration and astigmatism. Although the system does not allow for quantitative modulation of individual aberration terms, the liquid lens configuration employed in this study effectively reproduces a range of optical aberration variations encountered in practical imaging systems.

Table 2. Zernike aberration coefficients at a 4.9° field of view, expressed in waves at the reference wavelength of 550 nm.

Zernike	Aberration of Field 4.9°	Zoom1 (λ)	Zoom2 (λ)	Zoom3 (λ)	Zoom4 (λ)
Z1	Piston	0.81441056	0.34397327	−0.34229713	−0.93661486
Z3	y-tilt	−0.03866404	−0.03872013	−0.03885825	−0.03909311
Z4	Field curvature	0.49953012	0.22723833	−0.17002021	−0.51409345
Z6	Astigmatism 0°	−0.18307557	−0.18788038	−0.19556749	−0.20286953
Z7	Primary y-coma	−0.01061117	−0.01064507	−0.01071465	−0.01081586

Table 2. *Cont.*

Zernike	Aberration of Field 4.9°	Zoom1 (λ)	Zoom2 (λ)	Zoom3 (λ)	Zoom4 (λ)
Z9	Trefoil y	−0.00047801	−0.00052188	−0.00058113	−0.0006278
Z11	Spherical Aberration	0.02148056	0.0209429	0.02012656	0.0193863
Z12	Astigmatism 0°, Sec	−0.00030272	−0.00030079	−0.0002969	−0.00029259
Z15	Tetrafoil	0.00000463	0.00000455	0.00000445	0.00000436

Since the image reconstruction algorithm mentioned ignores the distortion simulation, the appearance of distortion should be avoided as much as possible in the process of optical system design. The distortion grid of the system is shown in Figure 8, and it can be seen that the imaging system has less distortion and can better meet the subsequent development of lens application.

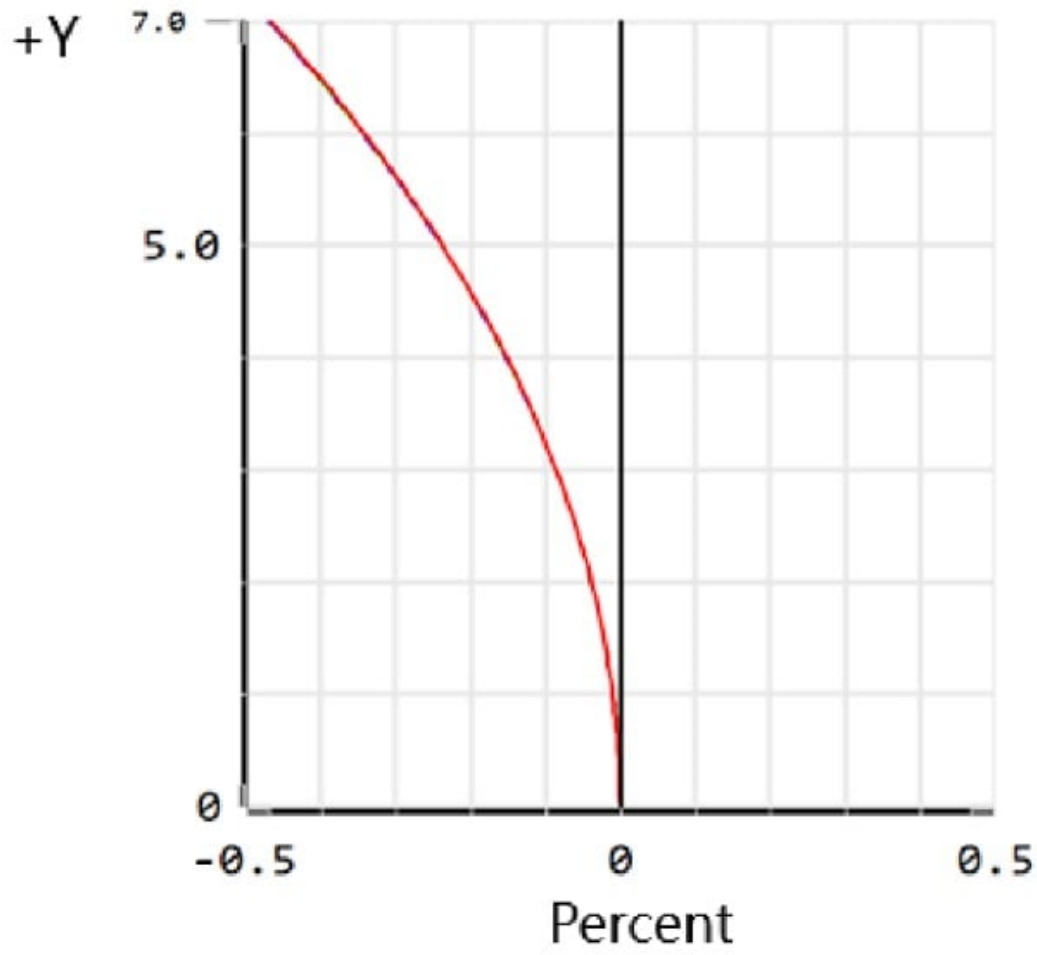


Figure 8. Distortion grid of the system.

Figure 9 illustrates the results of dataset reconstruction under different aberration conditions simulated using the proposed optical system. Figure 9a shows the original simulated image selected from the COCO [18] public dataset (COCO-test2017-NO.30908). The corresponding system outputs under varying aberration states are presented in Figure 9b–e. As the optical aberration changes with the liquid lens, the central field of view progressively transitions from sharp to blurred, while the peripheral field improves in clarity. In the Zoom4 state, both central and peripheral regions become blurred, indicating a typical defocused condition. These results demonstrate that the proposed optical system,

in combination with the tunable liquid lens, can effectively emulate a variety of aberration scenarios for dataset reconstruction and simulation purposes.

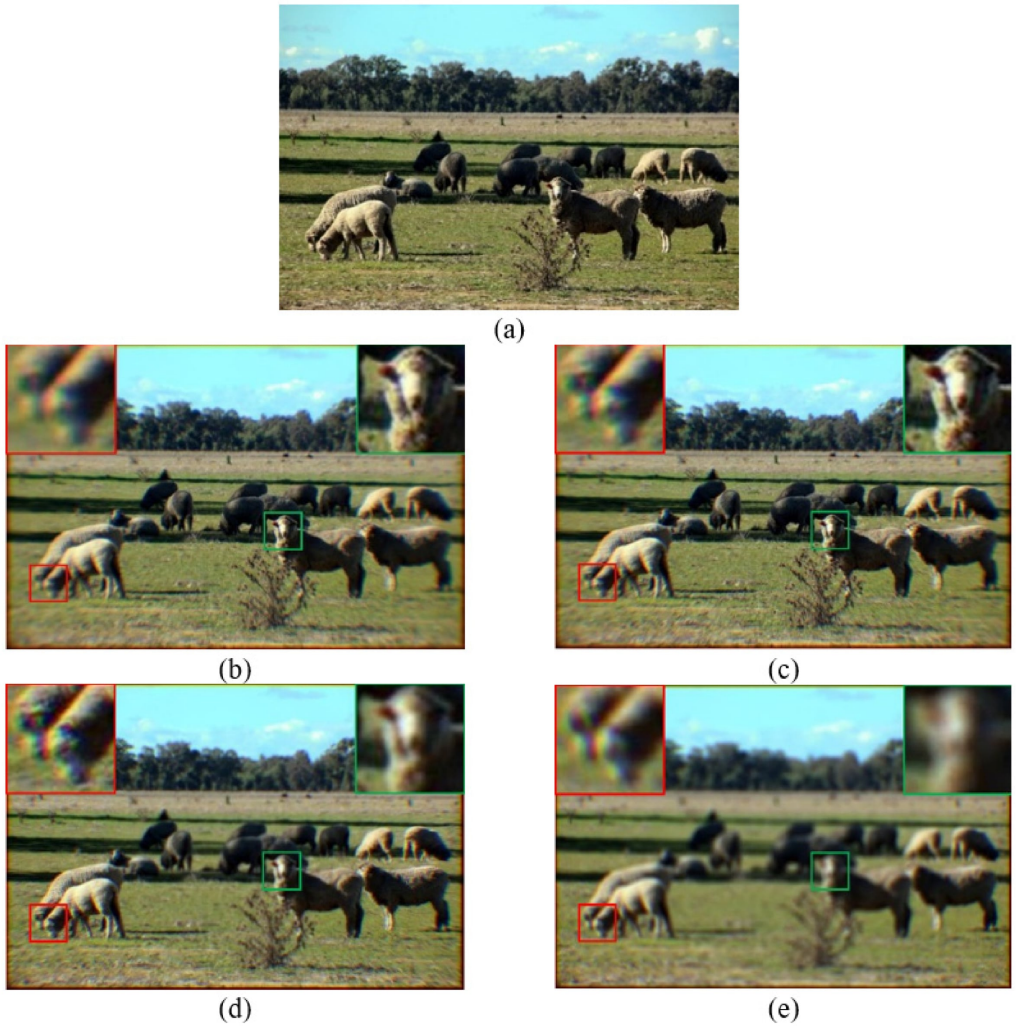


Figure 9. Comparison between the original image and reconstructed images of COCO-test2017 No. 30908. ((**a**) shows the original image; (**b**–**e**) present the reconstructed images under Zoom1 to Zoom4. The red boxes and green boxes represent the enlarged images of the corresponding positions.).

Figure 10 presents images captured through the experimental optical system, illustrating the image degradation under different focus settings.



Figure 10. Images captured by the experimental optical system under different zoom settings. ((**a**) shows the image captured at Zoom1; (**b**) shows the image captured at Zoom3.).

4. Discussion

The initial experiment uses a PyTorch (0.22.1 version) YOLOv4 implementation. The retained evaluation code computes all-point AP with IoU (Intersection over Union) ≥ 0.5 matching and ignores VOC difficult targets. Prediction export uses a score threshold of 0.01, NMS IoU 0.5 and 416 × 416 input. The initialization of the archived checkpoints is not fully documented.

The retrained model and baseline are denoted R-Net and O-Net. O-Test contains the original VOC2007 test images, while R-Test1–4 contain the corresponding optical reconstructions. Table 3 reports the initial single-run results. R-Net improves AP by 0.08, 0.36, 0.63, and 1.02 percentage points on the four reconstructed sets, with a 2.53-point decrease on the original set. State 4 is excluded from reconstruction training. Repeated-run comparisons are presented in Section 4.4.

Table 3. Mean average precision (mAP) results of the two networks across five different test datasets.

Net Test Set	O-Net O-Test	R-Net O-Test	O-Net R-Test1	R-Net R-Test1	O-Net R-Test2	R-Net R-Test2	O-Net R-Test3	R-Net R-Test3	O-Net R-Test4	R-Net R-Test4
mAP	83.24	80.71	80.25	80.33	82.56	82.92	81.93	82.56	78.91	79.93
bird	83.06	79.34	78.97	79.75	82.5	83.35	81.91	82.5	78.39	79.03
boat	76.1	73.22	71.98	71.98	75.47	75.7	74.42	75.47	70.65	71.19
bus	90.02	87.94	86.99	87.48	88.16	89.2	88.09	88.16	86.07	86.43
car	86.94	84.49	85.5	84.44	88.16	87.98	87.74	88.16	84.77	86.49
cat	90.39	88.3	88.63	88.7	90	90.06	89.73	90	86.93	87.61
chair	69.92	65.59	64.65	64.88	67.31	68.19	66.52	67.31	62.88	63.78
cow	83.64	80.54	81.67	80.15	84.39	83.51	82.99	84.39	78.56	80.19
dog	87.82	86.62	86.64	86.38	88.69	88.67	87.85	88.69	85.36	86.14
horse	91.44	90.07	89.67	90.25	91.67	92.02	91.37	91.67	89.99	89.93
person	86.1	83.66	83.66	82.55	84.79	83.89	83.82	84.79	81.07	82.67
sheep	81.66	79.05	79.28	78.74	82.74	82.05	81.55	82.74	76.6	79.39
sofa	84.41	83.25	81.7	82.64	82.67	84.27	82.43	82.67	80.42	80.96
train	90.18	88.91	88.86	88.22	90.23	89.66	89.5	90.23	86.35	88.01

Figure 11 illustrates detection results on COCO-test2017 image No. 30908 and its optical reconstructions.

4.1. Clean Domain Trade off

The clean-domain decrease reflects a trade-off in adaptation to optical degradation. In the controlled study, SSDLite gains 2.51 AP50 points on states 1–3 and 5.45 points on state 4 relative to clean training, while losing 1.41 points on clean images. YOLOv8n loses 2.60 points on clean images. Mixed clean/reconstructed training or continual-learning regularization [19] may help balance these domains, but their effectiveness requires further evaluation.

4.2. Computational Cost and Practical Applicability

Optical simulation and dataset reconstruction are offline processes. On an RTX 4060 GPU, YOLOv8n training takes 24.0–25.8 min for 60 epochs, and SSDLite takes 25.1–49.1 min for 40 or 80 epochs. These measurements cover detector training only; optical simulation and image-generation costs are not included.

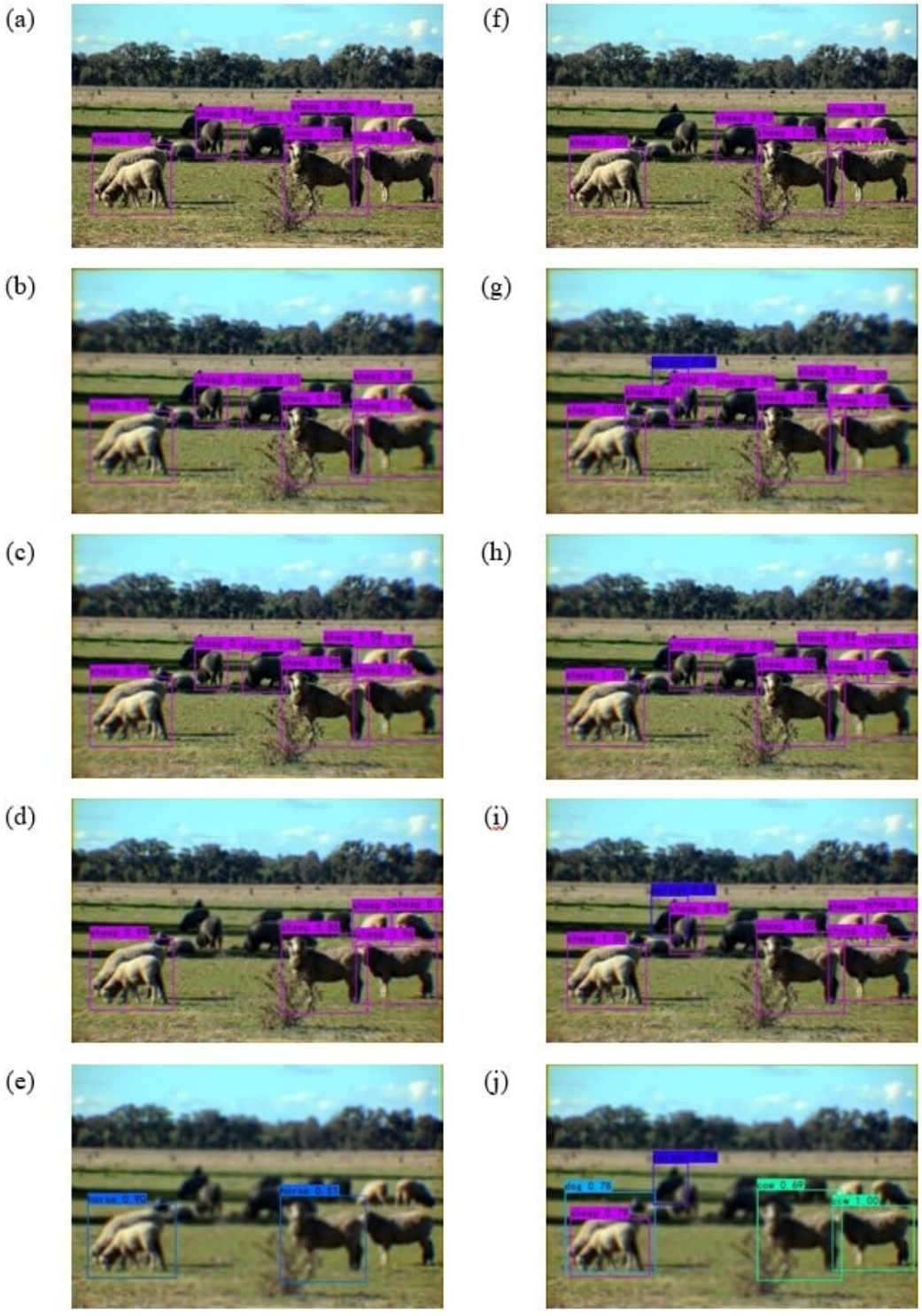


Figure 11. Detection results of the original network and the retrained network on the COCO-test2017 image No. 30908 under different aberration conditions. ((**a**–**e**) show the outputs of O-Net on the original image and on images reconstructed under Zoom1 to Zoom4; (**f**–**j**) show the outputs of R-Net under the same conditions. The purple box indicates that it is identified as a sheep.).

4.3. Position Relative to General Augmentation Baselines

Gaussian blur is a common augmentation option in libraries such as Albumentations [20] and imgaug [21]. The global and field-dependent Gaussian controls used here are fitted to training image pairs. They provide a comparison between analytic blur approx-

imations and system-specific PSF reconstruction. The observed SSDLite benefit applies to these controls; other degradation and restoration methods remain outside this comparison.

Figure 12 gives a qualitative hardware example. Quantitative results on the 24-image person dataset are presented in Section 4.5.



(a) (b)

Figure 12. Qualitative comparison of O-Net and R-Net on an image captured through the liquid-lens system ((**a**) shows the recognition results of O-Net, while B; (**b**) shows the recognition results of R-Net. The yellow boxes indicate those identified as "person".).

4.4. Controlled Simulation Experiments

The added study uses official VOC2007 scene splits, as follows: 2501 training, 2510 validation, and 4952 test scenes, with all 20 classes. Degraded versions remain with their original scene split. States 1–3 are used for training and state 4 is held out. Four conditions—clean, global Gaussian, field Gaussian, and system PSF—are compared using paired seeds 17, 29, and 43 for YOLOv8n [22] and SSDLite320 with a MobileNetV3-Large backbone [23–25]. One state per scene per epoch is sampled deterministically with the same schedule across conditions. The original normalized boxes are retained, without image-registration correction.

Gaussian parameters are fitted using 96 training scenes. The global standard deviations are 2.0, 1.5, and 2.0 pixels at a reference long side of 500 pixels. The field control uses channel-dependent values on a 6 × 8 grid with smooth interpolation. Fitting is based on image pairs rather than exact PSF moments. Flips, mosaic and color augmentation are disabled, and all conditions use a common JPEG-quality-95 encoding step.

All conditions within each architecture use matched initialization and training settings. YOLOv8n uses COCO-pretrained weights, 416-pixel input, batch size 32, and SGD with initial learning rate 0.004 for 60 epochs. SSDLite uses COCO-pretrained weights with a reinitialized 21-class head, 320-pixel input, batch size 16, and SGD at 0.0025 for 40 epochs. A preset validation rule extends two Gaussian runs to 80 epochs from the same initialization. Checkpoints are selected on 256 validation scenes with equal representation of clean images and states 1–3. Full validation precedes final testing, and test scores are not used for tuning. The experiments use use Ultralytics 8.4.190 and Torchvision 0.22.1 SSDLite320 implementation [25] with a MobileNetV3-Large backbone [23].

Table 4 reports all-point AP50 over 20 VOC classes, with difficult targets ignored, as the mean and sample standard deviation over three seeds. Predictions use a score threshold of 0.001 and at most 300 detections per image; YOLOv8n uses NMS IoU 0.7. The primary comparison is PSF minus field-Gaussian AP50 averaged over states 1–3. Descriptive 95% intervals are obtained from 1000 paired hierarchical bootstrap resamples of scenes and seeds, without adjustment for multiple comparisons.

Table 4. Added simulation study. AP50 in percent; mean ± sample standard deviation over seeds. States 1–3 are averaged within each seed before computing variability.

Model	Training	Clean	States 1–3	State 4
YOLOv8n	Clean	73.06 ± 1.89	67.22 ± 1.68	57.83 ± 1.35
YOLOv8n	Global Gaussian	68.26 ± 2.25	65.75 ± 2.32	57.37 ± 1.52
YOLOv8n	Field Gaussian	70.53 ± 2.50	66.73 ± 2.80	57.50 ± 4.55
YOLOv8n	PSF	70.46 ± 2.36	67.18 ± 2.25	57.30 ± 2.37
SSDLite	Clean	66.59 ± 0.14	61.68 ± 0.28	52.31 ± 0.35
SSDLite	Global Gaussian	61.83 ± 0.22	60.38 ± 0.46	56.94 ± 0.42
SSDLite	Field Gaussian	64.75 ± 0.77	61.86 ± 0.78	55.71 ± 0.97
SSDLite	PSF	65.18 ± 0.45	64.19 ± 0.33	57.76 ± 0.29

SSDLite improves AP50 over field Gaussian by 2.33 points on states 1–3 (descriptive 95% interval [1.54, 3.06]) and 2.05 points on state 4 ([0.84, 3.46]). The improvement over global Gaussian is 3.80 points on states 1–3. All 20 classes show positive mean differences, and 15 are positive in every seed. These class results are descriptive; object scale, contrast, and field position may contribute to the differences. Alternative AP calculations preserve the aggregate SSDLite improvement.

For YOLOv8n, the corresponding difference is 0.45 points, with an interval of $[-1.52, 2.15]$. All 12 runs select first-epoch checkpoints, limiting the evidence for adequate adaptation. Thus the consistent improvement in this study is established for SSDLite, rather than for both detector architectures.

4.5. Exploratory Hardware Image Evaluation

A person-annotated dataset contains 24 street-scene images captured through the proposed optical system at 1920 × 1080 resolution. Two acquisition subsets contain 12 images each. The set is fixed before inference, and all SSDLite conditions and seeds are evaluated without retraining. Individual Zoom states are not calibrated.

AI-assisted annotation gives 68 regular person boxes and 22 difficult/ignore boxes. Eight images without annotated persons remain in evaluation. Truncated people and riders use visible-body boxes; ambiguous targets and those below 100 source pixels in height are ignored. The annotations are approximate, not independently human-verified, and were not blinded to all prior predictions. This dataset supports an exploratory person-detection evaluation. Table 5 reports person AP50, precision and recall for the four training conditions, and Table 6 further breaks the AP50 down by acquisition subset.

Table 5. Person detection on 24 hardware images. Values are percentages, mean ± sample standard deviation over three seeds. Precision and recall use confidence 0.25 and IoU > 0.5; AP50 uses VOC difficult-ignore matching.

Training	Person AP50	Precision	Recall
Clean	46.39 ± 1.47	88.12 ± 5.53	27.94 ± 6.41
Global Gaussian	50.27 ± 4.55	73.98 ± 9.39	36.27 ± 1.70
Field Gaussian	47.88 ± 2.00	83.69 ± 9.91	26.96 ± 8.62
PSF	51.93 ± 3.29	85.74 ± 4.28	31.37 ± 3.06

Table 6. Person AP50 by acquisition subset, averaged over seeds; each subset contains 12 images.

Acquisition Subset	Clean	Global Gaussian	Field Gaussian	PSF
Subset 1	39.18	51.11	44.26	60.05
Subset 2	52.24	50.18	50.98	47.13

PSF training gives the highest pooled person AP50 of 51.93%. However, its ranking differs between the two subsets, while global Gaussian has higher recall and clean training has higher precision at confidence 0.25. Together with limited scene diversity, repeated identities and approximate annotations, these differences restrict the hardware result to preliminary evidence.

5. Conclusions

This study presents a system-specific dataset reconstruction method using field-dependent optical responses. For SSDLite, PSF training improves AP50 over fitted field Gaussian by 2.33 points on states 1–3 and 2.05 points on a held-out state of the same liquid-lens system. The 24-image hardware evaluation also gives the highest pooled person AP50 for PSF training. These results support the adaptation value of optical reconstruction beyond the Gaussian approximations tested. The benefit depends on detector architecture and imaging conditions, with a trade-off on clean images.

Further work will examine measured PSF agreement, broader hardware scenes, mixed-domain training, and transfer to other optical systems.

Author Contributions: Conceptualization, L.J.; methodology, L.J. and J.T.; software, J.T.; validation, L.J. and J.T.; formal analysis, J.T.; investigation, L.J. and J.T.; resources, J.C.; data curation, J.T.; writing—original draft preparation, J.T.; writing—review and editing, J.T.; visualization, J.T.; supervision, J.C.; project administration, J.C.; funding acquisition, J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Reconstructed data and supporting materials are available from the corresponding author upon request, subject to applicable third-party licenses.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kim, J.; Zeng, H.; Ghadiyaram, D.; Lee, S.; Zhang, L.; Bovik, A.C. Deep Convolutional Neural Models for Picture-Quality Prediction: Challenges and Solutions to Data-Driven Image Quality Assessment. *IEEE Signal Process. Mag.* **2017**, *34*, 130–141. [CrossRef]
2. Liu, Q.; Li, P.; Zhao, W.; Cai, W.; Yu, S.; Leung, V.C.M. A Survey on Security Threats and Defensive Techniques of Machine Learning: A Data Driven View. *IEEE Access* **2018**, *6*, 12103–12117. [CrossRef]
3. Wyant, J.C.; Creath, K. Basic Wavefront Aberration Theory for Optical Metrology. In *Applied Optics and Optical Engineering*; Shannon, R.R., Wyant, J.C., Eds.; Academic Press: Boston, MA, USA, 1992; Volume XI, pp. 1–53.
4. Perš, J.; Sulić Kenk, V.; Mandeljc, R.; Kristan, M.; Kovačič, S. Dana36: A Multi-Camera Image Dataset for Object Identification in Surveillance Scenarios. In Proceedings of the 2012 IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance, Beijing, China, 18–21 September 2012; pp. 64–69.
5. Ristani, E.; Solera, F.; Zou, R.S.; Cucchiara, R.; Tomasi, C. Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking. In *Computer Vision—ECCV 2016 Workshops*; Springer: Cham, Switzerland, 2016; pp. 17–35.
6. Anwar, S.; Khan, S.; Barnes, N. A Deep Journey into Super-Resolution: A Survey. *ACM Comput. Surv.* **2020**, *53*, 60. [CrossRef]
7. Chen, S.; Feng, H.; Pan, D.; Xu, Z.; Li, Q.; Chen, Y. Optical Aberrations Correction in Postprocessing Using Imaging Simulation. *ACM Trans. Graph.* **2021**, *40*, 192. [CrossRef]
8. Stamnes, J.J.; Heier, H. Routine Computation of Diffraction PSF for Complicated Optical Design. In Proceedings of the Annual Meeting Optical Society of America, Santa Clara, CA, USA, 30 October–4 November 1988. paper TUM6.
9. Pu, D.; Tao, H.; Ge, Y.; Liu, C.; Zhu, J. Improvements on Sampling of Point Spread Function in Optical Transfer Function Measurement. *Opt. Express* **2022**, *30*, 10953–10968. [CrossRef] [PubMed]
10. Jemec, J.; Pernuš, F.; Likar, B.; Bürmen, M. 2D Sub-Pixel Point Spread Function Measurement Using a Virtual Point-Like Source. *Int. J. Comput. Vis.* **2017**, *121*, 391–402. [CrossRef]

11. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
12. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934.
13. Müller, P.; Braun, A.; Keuper, M. Classification robustness to common optical aberrations. *arXiv* **2023**, arXiv:2308.15499. [CrossRef]
14. Müller, P.; Braun, A.; Keuper, M. Examining the Impact of Optical Aberrations to Image Classification and Object Detection Models. *IEEE Trans. Pattern Anal. Mach. Intell.* **2026**, *48*, 2139–2153. [CrossRef] [PubMed]
15. Everingham, M.; Van Gool, L.; Williams, C.K.I.; Winn, J.; Zisserman, A. The Pascal Visual Object Classes (VOC) Challenge. *Int. J. Comput. Vis.* **2010**, *88*, 303–338. [CrossRef]
16. Hirsch, M.; Schölkopf, B. Self-Calibration of Optical Lenses. In Proceedings of the 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 612–620.
17. Liang, D.; Wang, X.Y. Zoom Optical System Using Tunable Polymer Lens. *Opt. Commun.* **2016**, *371*, 189–195. [CrossRef]
18. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft COCO: Common Objects in Context. In *Computer Vision—ECCV 2014*; Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds.; Springer: Cham, Switzerland, 2014; Volume 8693, pp. 740–755.
19. Kirkpatrick, J.; Pascanu, R.; Rabinowitz, N.; Veness, J.; Desjardins, G.; Rusu, A.A.; Milan, K.; Quan, J.; Ramalho, T.; Grabska-Barwinska, A.; et al. Overcoming Catastrophic Forgetting in Neural Networks. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 3521–3526. [CrossRef] [PubMed]
20. Buslaev, A.; Iglovikov, V.I.; Khvedchenya, E.; Parinov, A.; Druzhinin, M.; Kalinin, A.A. Albumentations: Fast and Flexible Image Augmentations. *Information* **2020**, *11*, 125. [CrossRef]
21. Jung, A.B.; Wada, K.; Crall, J.; Tanaka, S.; Graving, J.; Reinders, C.; Yadav, S.; Banerjee, J.; Vecsei, G.; Kraft, A.; et al. imgaug. 2020. Available online: https://github.com/aleju/imgaug (accessed on 1 February 2020).
22. Jocher, G.; Chaurasia, A.; Qiu, J. *Ultralytics YOLOv8*. Computer software. 2023. Available online: https://github.com/ultralytics/ultralytics (accessed on 8 November 2026).
23. Howard, A.; Sandler, M.; Chu, G.; Chen, L.-C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V.; et al. Searching for MobileNetV3. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*; IEEE: Piscataway, NJ, USA, 2019; pp. 1314–1324.
24. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; IEEE: Piscataway, NJ, USA, 2018; pp. 4510–4520.
25. Torchvision Contributors. SSDLite320 MobileNetV3 Large Model Documentation. Available online: https://docs.pytorch.org/vision/0.22/models/generated/torchvision.models.detection.ssdlite320_mobilenet_v3_large.html (accessed on 8 November 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.