photonics

MDPI

Article

Nonlinear Band-Correlation Guided Deep Reinforcement Learning for Multispectral Autofocus Optimization

Zhenzhen Chen [1], Sen Wang [1], Chao Ma [2], Shaowen Jing [2], Jiayu Huang [3] and Mingkun Zhang [2,*]

[1] Artificial Intelligence Department, Yuxin Electronic Technology Group Co., Ltd., Zhengzhou 450046, China; chenzhenzhen@eathink.com.cn (Z.C.); wangsen@eathink.com.cn (S.W.)
[2] College of Information Engineering, Henan University of Science and Technology, Luoyang 471023, China
[3] College of Electronics and Information Engineering, South China University of Technology, Guangzhou 510641, China
* Correspondence: 240410040059@stu.haust.edu.cn

Abstract

Multispectral imaging systems often exhibit wavelength-dependent focus responses because of chromatic aberration, nonuniform spectral sensitivity, and scene-dependent reflectance, making single-band autofocus inadequate for optimizing the complete spectral image cube. This study proposes a nonlinear band-correlation guided deep reinforcement learning framework for autofocus optimization in a self-developed 31-band multispectral imaging system covering 360–980 nm. Through-focus response curves are first extracted from all spectral channels, and a nonlinear inter-band dependency matrix is constructed to characterize complementary and redundant focusing information across wavelengths. A compact subset of representative bands is then selected and embedded into the reinforcement-learning state to guide closed-loop motor control. The agent jointly determines focusing direction, displacement, and stopping time while maximizing a global multispectral focus objective that considers mean sharpness, inter-band consistency, and worst-band degradation. During online autofocus, only the representative bands are acquired at intermediate positions, whereas the complete 31-band cube is captured after convergence. The framework is evaluated using standardized optical targets and spectrally heterogeneous scenes through focal-position error, MTF, full-band focus retention, acquisition time, and robustness. The proposed design provides an efficient computational-photonics solution for broadband multispectral autofocus with reduced band-acquisition and motor-search overhead.

Keywords: autofocus optimization; nonlinear band correlation; multispectral imaging; deep reinforcement learning; computational photonics; optical control

Check for updates

Received: 14 July 2026
Revised: 23 August 2026
Accepted: 31 August 2026
Published: 1 September 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.

1. Introduction

Multispectral imaging extends conventional intensity imaging by resolving wavelength-dependent interactions between illumination, optical propagation, scene reflectance, and detector response, thereby providing complementary spatial and spectral information for quantitative observation. Contemporary computational optical imaging increasingly treats the optical front end and the digital reconstruction module as a jointly optimized system rather than as two independent stages [1–3]. This physical-digital co-design principle has supported rapid progress in coded-aperture imaging, correlation holography, and low-sampling single-pixel reconstruction [4–6]. Physics-guided learning has also improved infrared hyperspectral recovery, multi-wavelength computational imaging, and spectral

reconstruction from compact photonic structures [7–9]. Snapshot and hybrid-encoded architectures further demonstrate that spectral information can be encoded through defocus, wavelength-dependent modulation, or learned aperture design and subsequently recovered computationally [10–13]. In parallel, neural wavefront sensing, adaptive optical-path-difference fusion, and uncertainty-aware Fourier ptychography have strengthened the ability of computational systems to infer or compensate optical states from incomplete measurements [14–16]. These developments indicate that intelligent optical acquisition should optimize not only image reconstruction after capture but also the physical imaging state during acquisition, as illustrated by recent deep-learning-assisted macroscopic Fourier ptychography [17,18].

Autofocus is a fundamental component of this acquisition process because axial defocus broadens the point-spread function, attenuates high spatial frequencies, and degrades both spatial resolution and downstream spectral analysis [19–21]. The problem becomes more difficult in broadband multispectral systems because focal response is wavelength-dependent [12,22]. Chromatic aberration can shift the best focal plane across spectral channels, diffraction changes the wavelength-dependent point-spread function, and nonuniform illumination and detector sensitivity alter the reliability of focus measures [23]. Scene reflectance introduces an additional source of heterogeneity because an object that provides strong edge contrast in one band may exhibit weak or spatially different contrast in another. Consequently, the position maximizing a focus measure in a single channel is not necessarily the position that provides the best quality, consistency, or minimum degradation for the complete multispectral cube [20]. Sequentially refocusing every channel can improve bandwise sharpness, but it substantially increases motor motion and acquisition time and may introduce inter-band geometric inconsistency [19]. A practical multispectral autofocus method must therefore infer a globally suitable focal position from a limited number of observations while explicitly accounting for cross-band dependence [24,25].

Recent autofocus research has improved focusing speed and precision through optical encoding, adaptive search, modulation-transfer-function feedback, and learning-based defocus estimation. Oblique illumination has enabled rapid chromosome-microscopy autofocus, high-speed telescope control has addressed moving aerial targets, and dynamic modulation-transfer-function feedback has provided a direct optical-quality criterion for iterative focusing [19,21,26]. Computer-generated-hologram-assisted focusing, dynamic region-of-interest selection, and variable-step hybrid search have further reduced redundant evaluation or improved local convergence [27–29]. Deep models can estimate defocus from Fourier-domain cues or content-aware regions, while deep reinforcement learning has been used to control liquid lenses without relying exclusively on a fixed analytical search rule [20,25,30]. Reinforcement-learning-based region selection has also been applied to wafer micro-imaging, and machine-learning-driven wavefront diagnostics have demonstrated autonomous alignment of X-ray optical systems [24,31]. Nevertheless, most existing autofocus methods are formulated for a single imaging channel or for a focus metric derived from one intensity image. They therefore do not model how multiple spectral channels jointly describe the axial optical state, nor do they optimize a full-band objective that balances average sharpness against inter-band inconsistency and worst-band degradation.

Band selection offers a potential route to reducing multispectral acquisition overhead, because many adjacent or optically similar channels contain partially redundant information. Related VIS/NIR studies have also demonstrated that informative channel selection, prior-guided spectral reconstruction, and local-consistency modeling can improve the utilization of spectral information under compact or limited-band acquisition settings [32–34]. Deep reinforcement learning and multi-agent strategies have recently

been developed to identify informative hyperspectral bands for classification-oriented tasks [35,36]. Graph-based subspace clustering, adaptive-neighbor clustering, and tensorial global-local self-representation further exploit structured relationships among spectral channels [37–39]. Other recent frameworks emphasize efficient band subsets or robust multi-view representations for hyperspectral analysis [40,41]. However, these methods generally characterize spectral redundancy from static reflectance vectors or discriminative labels. Autofocus requires a different form of dependence: the relevant variable is the through-focus response trajectory of each band as the optical system moves along the axial direction. Two bands may have limited linear intensity correlation yet exhibit strongly related nonlinear focus-response profiles, and conversely, adjacent wavelengths may differ in focus reliability because of signal-to-noise ratio, texture contrast, or chromatic focal shift. Therefore, directly transferring conventional band-selection criteria to autofocus does not establish whether the selected channels preserve the axial state information required for reliable motor control.

Reinforcement learning is particularly suitable for autofocus because focusing is a sequential decision problem involving partial observations, movement cost, action-dependent future measurements, and a stopping decision. Recent photonic systems have used reinforcement learning for high-efficiency integrated processing, multi-wavelength optical information processing, and in-situ optimization of model-free optical processors [42–44]. More broadly, the convergence of physics-informed artificial intelligence and inverse photonic design has shown that physical priors can constrain learning, improve data efficiency, and make optical optimization more interpretable [45,46]. Despite this progress, a gap remains between reinforcement-learning-based optical control and broadband multispectral autofocus. A standard deep reinforcement learning agent that receives independently concatenated band features does not distinguish redundant from complementary channels, does not explicitly encode nonlinear inter-band dependence, and may optimize the behavior of the strongest band rather than the quality of the complete spectral cube.

To address these limitations, this study proposes nonlinear band-correlation guided deep reinforcement learning for autofocus optimization in a self-developed 31-band multispectral imaging system covering 360–980 nm. The central premise is that each spectral channel provides a wavelength-dependent observation of a shared axial optical state, while the dependencies among their focus-response curves are nonlinear and unequal. Complete through-focus scans are first used during offline modeling to construct a nonlinear band-correlation matrix. The matrix is then used to select a compact set of representative bands and to encode their observations into a correlation-aware state representation. During online focusing, the agent acquires only these representative channels at intermediate positions and jointly determines motor direction, displacement, and stopping time. After convergence, the complete 31-band cube is acquired at the selected focal position. The optimization objective is defined at the multispectral level by considering mean normalized sharpness, inter-band consistency, and worst-band degradation, rather than maximizing an isolated single-band focus value.

The main contributions of this study are summarized as follows:

1. A wavelength-dependent autofocus formulation is established for a 31-band broadband multispectral imaging system, in which the target focal position is defined by the quality of the complete spectral cube rather than by an individual channel.
2. A nonlinear band-correlation model is developed from through-focus response trajectories to quantify complementary and redundant axial information across wavelengths, and a correlation-coverage criterion is used to select representative focusing bands.

3. The nonlinear correlation structure is embedded into a deep reinforcement learning controller that makes closed-loop decisions on focusing direction, adaptive displacement, and termination while reducing intermediate band acquisition.
4. The proposed method is evaluated through focal-position error, modulation transfer function, full-band focus retention, inter-band consistency, motor-search cost, acquisition time, and robustness under varying initial positions and imaging conditions.

2. Multispectral Imaging System and Problem Formulation

2.1. Thirty-One-Band Multispectral Imaging System

A self-developed multispectral imaging system was used to investigate broadband autofocus over the spectral range of 360–980 nm. The system consisted of a 31-band wavelength-selective illumination module, an integrating-sphere illumination unit, an imaging lens, a motorized axial focusing mechanism, a 12-bit charge-coupled device camera with a spatial resolution of 1600×1200 pixels, and a computer-based control and acquisition unit. The illumination module sequentially activated the prescribed spectral channels, while the integrating sphere reduced spatial nonuniformity in the incident illumination. The 31 spectral channels were independently addressable and were activated sequentially according to the predefined acquisition sequence; therefore, only the designated wavelength channel was illuminated during each band acquisition. The inner surface of the integrating sphere was coated with a highly reflective barium sulfate layer, which promoted multiple diffuse reflections of the incident light and produced a spatially homogenized illumination field at the target plane. The integrating sphere was used exclusively for illumination homogenization and did not participate in the imaging path: light reflected from the illuminated target was collected directly by the imaging lens and subsequently recorded by the CCD camera. The target remained stationary during each autofocus episode, and the effective object-to-lens distance was adjusted through the motorized focusing unit.

The wavelength-switching module, camera exposure, image acquisition, and motor motion were synchronized by the control computer. The motorized unit was operated on a discrete axial grid defined as

$$\mathcal{Z} = z_1, z_2, \ldots, z_M, \tag{1}$$

where z_M denotes the mth admissible focusing position and M is the total number of axial positions within the predefined search range. The axial interval was selected to cover both clearly defocused regions and the complete transition around the best-focus region. For each experimental condition, the wavelength-specific exposure settings remained unchanged throughout the complete axial scan to ensure that variations in the focus response were caused primarily by axial displacement rather than by exposure adaptation.

Two acquisition modes were used. In the offline modeling stage, complete 31-band through-focus stacks were acquired at all axial positions. These data were used to characterize wavelength-dependent focus responses, establish the nonlinear band-correlation model, determine the reference multispectral focal position, and construct the reinforcement-learning environment. In the online autofocus stage, only a compact subset of representative spectral bands was acquired at intermediate positions. Once the reinforcement-learning controller issued a stopping action, the complete 31-band multispectral cube was recorded at the selected focal position. This acquisition strategy reduced repeated wavelength switching and avoided collecting a complete spectral cube at every motor position. The optical configuration, spectral-channel distribution, and closed-loop acquisition sequence are summarized in Figure 1. The system combines wavelength-selective illumination with motorized axial control, while the online autofocus loop separates intermediate representative-band observations from the final complete-cube acquisition.

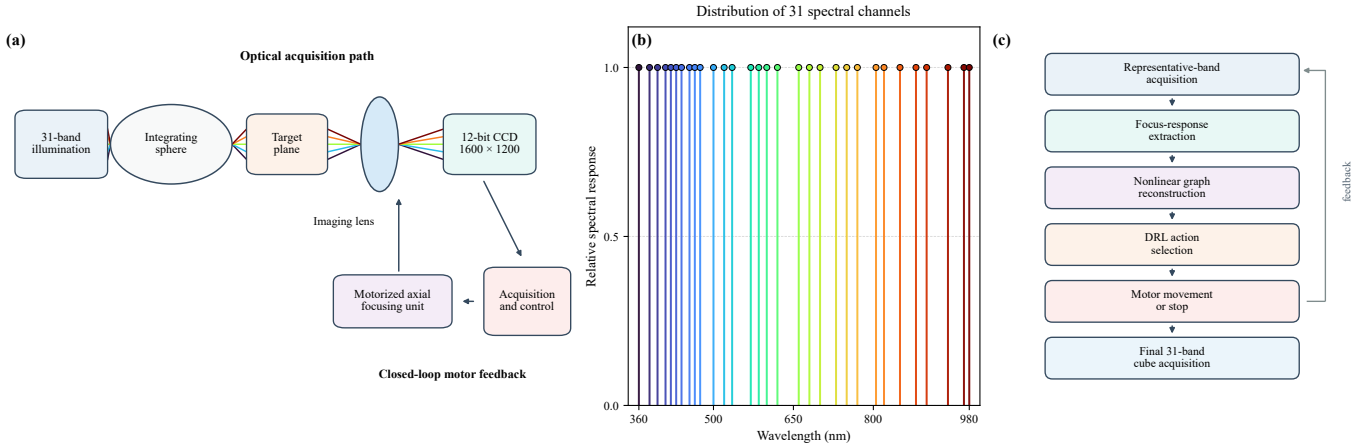


Figure 1. Multispectral imaging system and closed-loop autofocus workflow. (**a**) Optical acquisition path, including the 31-band illumination module, integrating sphere, target plane, imaging lens, CCD camera, and motorized axial focusing unit. (**b**) Distribution of the 31 spectral channels over the 360–980 nm range. (**c**) Online autofocus sequence based on representative-band acquisition, focus-response extraction, nonlinear graph reconstruction, deep reinforcement-learning action selection, motor feedback, and final 31-band cube acquisition.

As shown in Figure 1, the optical system preserves the complete 31-channel measurement capability while reducing repeated spectral acquisition during focusing. The representative channels are acquired only inside the motor-control loop, whereas the complete multispectral cube is recorded once after convergence. Consequently, the proposed strategy reduces wavelength switching and exposure overhead without changing the spectral dimensionality of the final imaging product. The motorized axial unit provides the physical control variable, and the nonlinear graph reconstruction and reinforcement-learning controller form the corresponding computational feedback path.

2.2. Radiometric Correction and Region Definition

The raw intensity recorded in each spectral channel was affected by dark current, wavelength-dependent illumination power, detector sensitivity, and fixed-pattern response. To reduce these effects, a dark-reference image and a white-reference image were acquired for each spectral band. The reference images were acquired separately for all 31 spectral channels before sample acquisition. For the dark reference, the camera lens was completely covered with an opaque low-reflectance cap under blocked-light conditions to record the detector dark-current and background response. For the white reference, a high-reflectance white reference panel was placed at the same target plane as the sample and sequentially imaged under the corresponding spectral illumination channels. For each wavelength, the exposure time, camera gain, imaging geometry, and other acquisition settings were kept identical between the reference and subsequent sample measurements. The corrected response at wavelength λ_i was calculated as

$$R_i(x, y, z) = \frac{I_i(x, y, z) - B_i(x, y)}{W_i(x, y) - B_i(x, y) + \varepsilon}, \tag{2}$$

where $I_i(x, y, z)$ is the raw image acquired at axial position z, $B_i(x, y)$ is the dark-reference image, $W_i(x, y)$ is the white-reference image, and ε is a small positive constant introduced to avoid numerical instability. Corrected values outside the physically meaningful interval were clipped before focus-feature extraction.

A fixed region of interest, denoted by Ω, was defined at the beginning of each autofocus episode. The same spatial region was retained at all axial positions and for all

spectral channels belonging to the same episode. Fixing the region prevented the autofocus score from being affected by changing segmentation boundaries or by the inclusion of different image contents during motor motion. For standardized optical targets, Ω covered the effective resolution region. For spectrally heterogeneous test scenes, the region was selected to contain representative edge and texture information while excluding invalid background areas.

To reduce the influence of wavelength-dependent intensity magnitude, the autofocus state was constructed from normalized spatial-frequency information rather than from raw intensity alone. A gradient-based focus response for the ith band was defined as

$$F_i(z) = \frac{1}{|\Omega|} \sum_{(x,y)\in\Omega} \frac{G_{x,i}^2(x,y,z) + G_{y,i}^2(x,y,z)}{\mu_i^2(z) + \varepsilon}, \tag{3}$$

where $G_{x,i}$ and $G_{y,i}$ are the horizontal and vertical image gradients, respectively, and $\mu_i(z)$ is the mean corrected response within Ω. Division by the squared mean response reduces the direct dependence of the focus value on the absolute brightness of the individual spectral channels.

Because the numerical scale of $F_i(z)$ can still differ across wavelengths, each bandwise focus response was normalized as

$$\bar{F} * i(z) = \frac{F_i(z) - F * i, \min}{F_{i,\max} - F_{i,\min} + \varepsilon}, \tag{4}$$

where $F_{i,\min}$ and $F_{i,\max}$ were estimated exclusively from the training data. The same normalization parameters were subsequently used for the validation and independent test sets. This procedure prevented information from the test focus stacks from being introduced into the autofocus controller.

The gradient-based response was used as the primary online focus feature because it can be calculated rapidly during closed-loop control. However, the optical performance of the system was not evaluated solely using this focus metric. Modulation transfer function, focal-position error, full-band focus retention, inter-band quality consistency, and acquisition efficiency were used as independent evaluation criteria in the experimental analysis.

2.3. Wavelength-Dependent Defocus Model

For the ith spectral channel with center wavelength λ_i, the corrected image acquired at axial position z can be represented as

$$R_i(x,y,z) = \alpha_i[O_i(x,y) * h_i(x,y;z - z_i^*)] + n_i(x,y,z), \tag{5}$$

where $O_i(x,y)$ denotes the wavelength-dependent object response, h_i is the corresponding point-spread function, z_i^* is the best-focus position of the ith band, α_i represents the combined wavelength-dependent response of the illumination system, optical components, target reflectance, and detector, $n_i(x,y,z)$ denotes noise, and $*$ represents spatial convolution.

The point-spread function depends on both wavelength and axial displacement. Consequently, even when all spectral channels observe the same physical scene, their through-focus responses are not necessarily identical. The best-focus position of the ith band is defined as

$$z_i^* = \arg, \max_{z\in\mathcal{Z}} \bar{F}_i(z). \tag{6}$$

The total wavelength-dependent focal spread of the imaging system is expressed as

$$\Delta z_{\text{chrom}} = \max_{1\le i\le 31} z_i^* - \min_{1\le i\le 31} z_i^*. \tag{7}$$

A nonzero value of Δz_{chrom} indicates that no single axial position simultaneously coincides with the individual optimum of every spectral channel. Even when the values of z_i^* are close, substantial differences may remain in the width, symmetry, slope, and signal-to-noise ratio of the corresponding focus-response curves.

For a complete axial scan, the through-focus trajectory of the ith band is represented by

$$\mathbf{f}_i = [\bar{F}_i(z_1), \bar{F}_i(z_2), \ldots, \bar{F}_i(z_M)]^{\mathrm{T}}. \tag{8}$$

The set $\mathbf{f} * i * i = 1^{31}$ provides a compact description of the axial optical behavior of the multispectral system. Bands with similar best-focus positions may still exhibit different response widths or asymmetric profiles, whereas bands with different intensity distributions may show strongly dependent axial trends. Therefore, the relationship between two spectral channels cannot be sufficiently characterized using wavelength distance or linear intensity correlation alone. Section 3 introduces a nonlinear dependence model constructed directly from these through-focus trajectories.

2.4. Single-Band and Multispectral Focal Positions

Conventional autofocus commonly determines the final position using a single reference channel,

$$z_{\text{single}}^* = \arg, \max_{z\in\mathcal{Z}} \bar{F}_r(z), \tag{9}$$

where r denotes the selected reference band. This strategy is computationally efficient, but its result can be biased toward the optical and reflectance characteristics of that channel. In particular, a reference band with strong spatial contrast may dominate the focusing decision even when other spectral channels experience considerable residual defocus.

A simple multispectral alternative is to maximize the weighted mean focus response,

$$M(z) = \sum_{i=1}^{31} w_i \bar{F}_i(z), \tag{10}$$

subject to

$$\sum_{i=1}^{31} w_i = 1, \qquad w_i \ge 0. \tag{11}$$

Here, w_i denotes the contribution of the ith spectral channel. However, maximizing only the weighted mean can still produce a solution in which several strong channels compensate for a severely degraded weak channel. Such a result is undesirable for multispectral imaging because every wavelength contributes to the final spectral cube.

To account for full-band consistency, the weighted dispersion of the bandwise focus responses is defined as

$$D(z) = \left[\sum_{i=1}^{31} w_i (\bar{F}_i(z) - M(z))^2\right]^{1/2}. \tag{12}$$

The weakest-band response is defined as

$$F_{\min}(z) = \min_{1\le i\le 31} \bar{F}_i(z). \tag{13}$$

The global multispectral focus objective is then formulated as

$$Q_{\mathrm{MS}}(z) = M(z) - \beta D(z) - \gamma[\max(0, \tau - F_{\min}(z))]^2, \tag{14}$$

where β controls the penalty on inter-band inconsistency, γ controls the penalty applied to excessive degradation of the weakest spectral channel, and τ is the minimum acceptable normalized focus response.

The reference multispectral focal position is defined as

$$z^*_{\mathrm{MS}} = \arg, \max * z \in \mathcal{Z} Q * \mathrm{MS}(z). \tag{15}$$

This definition differs from both single-band focusing and simple full-band averaging. It seeks a compromise position that maintains high overall sharpness while suppressing large variation among spectral channels and preventing individual bands from becoming severely defocused. The position z^*_{MS} was obtained from the complete offline focus stack and was used as the reference target for reinforcement-learning training and performance evaluation.

Unless otherwise specified, uniform weights were used in the reference objective,

$$w_i = \frac{1}{31}, \qquad i = 1, 2, \ldots, 31, \tag{16}$$

to prevent prior preference for a particular wavelength range. Reliability-aware and correlation-aware weighting strategies are examined separately in the ablation experiments. The parameters in the multispectral objective were determined exclusively from the validation scenes before reinforcement-learning training and independent testing. A coarse grid search followed by local refinement was performed over β, γ, and τ by jointly considering full-band quality retention, inter-band dispersion, weakest-band sharpness retention, and the stability of the resulting multispectral focal position. The final values were set to $\beta = 0.20$, $\gamma = 0.60$, and $\tau = 0.70$. The relatively moderate consistency penalty prevents a small number of channels from dominating the global objective, whereas the stronger weakest-band penalty suppresses solutions in which high responses from several channels compensate for substantial degradation of an individual wavelength. The threshold $\tau = 0.70$ defines the minimum normalized focus response below which the weakest-band penalty becomes active. All three parameters were fixed after validation and remained unchanged during reinforcement-learning training and independent testing.

2.5. Sequential Autofocus Optimization Problem

During online operation, acquisition of all 31 bands at every axial position would largely eliminate the efficiency advantage of intelligent autofocus. Therefore, the controller observes only a representative band subset,

$$\mathcal{B} = b_1, b_2, \ldots, b_K, \qquad K \ll 31, \tag{17}$$

where the selection of $\mathcal{B}$ is guided by the nonlinear band-correlation model described in Section 3.

At decision step t, the controller acquires the representative bands at the current position z_t, constructs a correlation-aware optical state, and selects an axial action a_t. The selected action determines the direction and magnitude of the motor displacement or terminates the autofocus episode. The final position after T decision steps is denoted by z_T.

The sequential autofocus objective is formulated as

$$\min_{\pi} \quad \mathbb{E} * \pi \left[-Q * \mathrm{MS}(z_T) + \lambda_1 T + \lambda_2 \sum_{t=1}^{T} |z_t - z_{t-1}| \quad + \lambda_3 N_{\mathrm{rev}} + \lambda_4 N_{\mathrm{band}} \right], \tag{18}$$

where π denotes the autofocus policy, T is the number of decision steps, N_{rev} is the number of motor direction reversals, and N_{band} is the total number of spectral-band acquisitions performed during the autofocus process. The coefficients λ_1–λ_4 balance final imaging quality against search duration, motor travel, mechanical inefficiency, and spectral acquisition cost.

The first term in Equation (18) encourages the controller to terminate at a position with high complete-cube focus quality. The second term penalizes lengthy decision sequences. The third term penalizes excessive axial motion, while the fourth term discourages repeated changes in movement direction that may increase mechanical delay and backlash. The final term explicitly accounts for the number of spectral observations required during focusing.

The objective is therefore not merely to locate a position with a high single-band focus value. The controller must reach a focal position close to the complete-cube optimum while using as few axial movements and spectral observations as possible. This problem involves partial observations, delayed rewards, action-dependent measurements, variable search trajectories, and an explicit stopping decision. It is therefore formulated as a deep reinforcement learning problem in the following section.

3. Nonlinear Band-Correlation Guided Deep Reinforcement Learning

3.1. Overall Framework

The proposed method integrates nonlinear band-correlation modeling with deep reinforcement learning for efficient multispectral autofocus. It contains an offline modeling stage and an online closed-loop focusing stage.

During offline modeling, complete 31-band through-focus stacks are collected from the training scenes. These data are used to characterize nonlinear dependencies among spectral channels, identify a compact set of representative bands, train a full-band state reconstructor, and construct the reinforcement learning environment.

During online autofocus, only the representative bands are acquired at each intermediate motor position. Their focus responses are used to estimate the complete multispectral focus state. The trained controller then determines the motor direction, displacement magnitude, or stopping action. After convergence, the complete 31-band cube is acquired once at the selected focal position. Figure 2 provides a module-level view of the proposed correlation-guided autofocus architecture. It illustrates how sparse representative-band observations are propagated over the spectral graph, fused with the motor-control state, and converted into adaptive movement and stopping decisions.

The architecture in Figure 2 separates optical-state estimation from motor-action evaluation. The graph encoder first transfers information from the observed representative bands to the unobserved channels, thereby generating an estimate of the complete spectral focus state. This estimate is then fused with temporal and mechanical variables before entering the dueling action-value network. The resulting controller can use coarse movements when the current position is far from the optimum, switch to finer movements near convergence, and issue an explicit stopping action when further displacement is no longer beneficial.

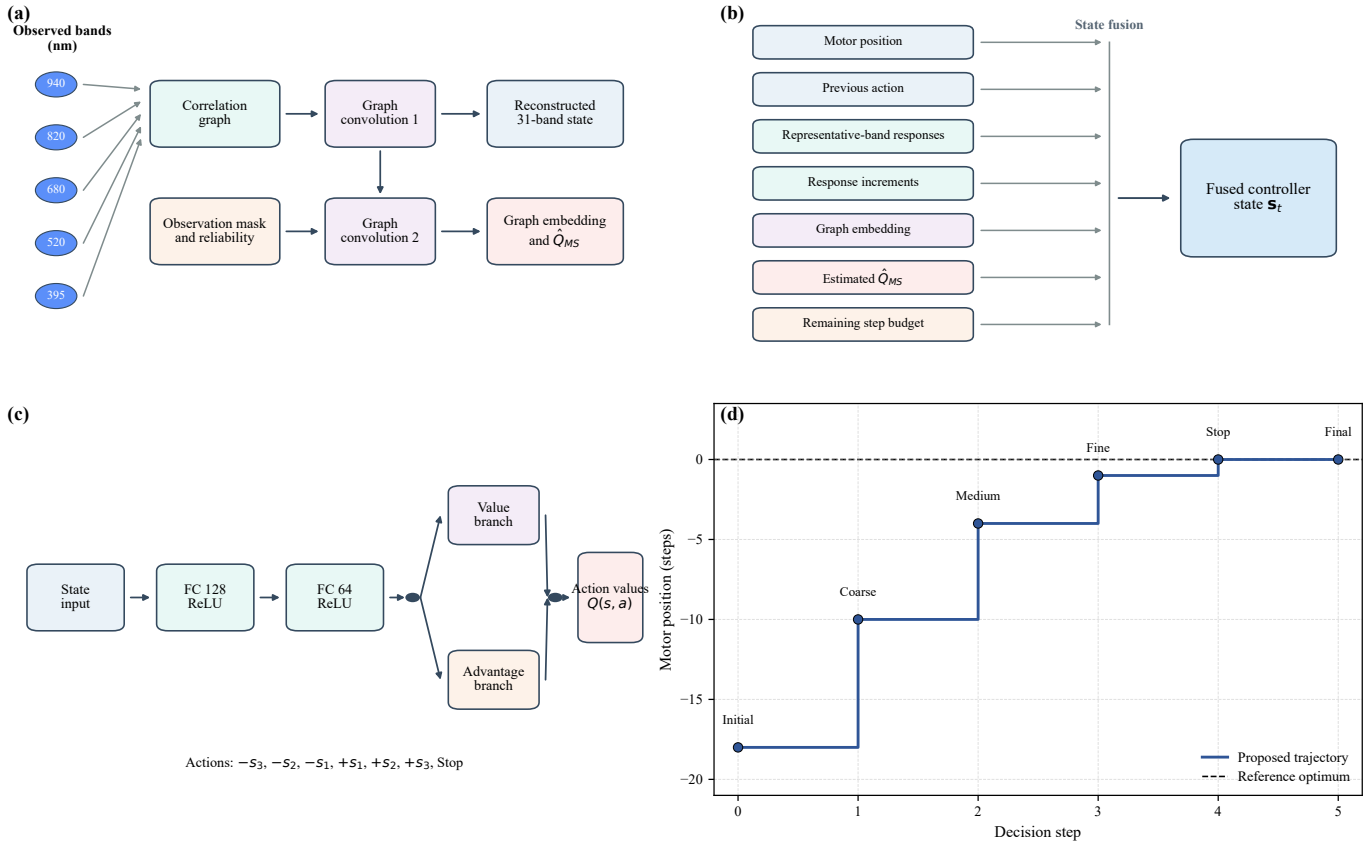


Figure 2. Correlation-guided state reconstruction and reinforcement-learning decision mechanism. (**a**) Propagation of the five observed representative band responses through the nonlinear correlation graph to reconstruct the complete 31-band state and obtain the graph embedding. (**b**) Fusion of motor position, previous action, representative-band responses, response increments, graph embedding, estimated multispectral quality, and remaining step budget. (**c**) Dueling double deep Q-network used to estimate the action values. (**d**) Representative coarse-to-fine autofocus trajectory terminating at the reference multispectral optimum.

3.2. Nonlinear Band-Correlation Modeling

For each spectral band, the normalized focus responses from all training scenes and axial positions are concatenated into a through-focus response vector. A Gaussian kernel is then applied to characterize nonlinear relationships among these response vectors. The physical interpretation of this dependence should be distinguished from that of an individual optical parameter. The coefficient does not directly measure chromatic focal shift, aberration magnitude, or any other isolated physical quantity. It characterizes the extent to which two wavelength channels encode a shared underlying axial defocus state through their wavelength-dependent through-focus responses. As the focusing position changes, the common axial displacement modifies the wavelength-dependent point-spread function and consequently changes the normalized focus response of each channel. Chromatic aberration primarily affects the locations of the response maxima, whereas wavelength-dependent diffraction and aberrations influence the response width and local slope; scene reflectance, illumination efficiency, detector sensitivity, and noise further affect the observable response shape and reliability. The kernel-based dependence measure therefore provides a response-level representation of the nonlinear coupling between two spectral channels under a common axial optical-state variation. A large dependence coefficient indicates that the two channels contain strongly related information about the axial focusing state even when their response amplitudes, peak positions, or local slopes are not linearly related,

whereas a smaller coefficient indicates weaker shared axial-state information between the corresponding wavelength channels.

The nonlinear dependence between bands i and j is defined as

$$C_{ij} = \frac{\operatorname{tr}(\mathbf{K}_i \mathbf{H} \mathbf{K}_j \mathbf{H})}{\sqrt{\operatorname{tr}(\mathbf{K}_i \mathbf{H} \mathbf{K}_i \mathbf{H}) \operatorname{tr}(\mathbf{K}_j \mathbf{H} \mathbf{K}_j \mathbf{H})} + \varepsilon}, \tag{19}$$

where $\mathbf{K}_i$ and $\mathbf{K}_j$ are the Gaussian kernel matrices of the two bands, and $\mathbf{H}$ is the centering matrix. Each paired observation was indexed by a training-scene-axial-position combination. Let $x_i^{(n)}$ denote the normalized focus response of band i for the nth paired observation, where $n = 1, \ldots, N$ and N is the total number of training observations used to construct the dependence model. The Gaussian kernel matrix for band i was constructed elementwise as $[K_i]_{pq} = \exp\left[-\left(x_i^{(p)} - x_i^{(q)}\right)^2 / (2\sigma_i^2)\right]$, where σ_i denotes the Gaussian-kernel bandwidth for the ith wavelength channel. The bandwidth was determined exclusively from the training observations using the median pairwise-distance heuristic, such that its scale adapted to the empirical focus-response distribution of each wavelength without introducing information from the validation or independent test sets. The centering matrix was defined as $H = I_N - \frac{1}{N}\mathbf{1}\mathbf{1}^T$. Accordingly, Equation (19) evaluates the normalized dependence between the centered Gaussian-kernel representations of two wavelength-dependent focus-response variables, enabling nonlinear co-variation induced by the common axial state to be quantified without imposing a predefined linear relationship. The resulting matrix $\mathbf{C} \in \mathbb{R}^{31\times 31}$ describes the nonlinear dependence among all spectral focus responses.

Unlike linear correlation, this measure can identify bands with similar axial variation even when their response amplitudes, slopes, or peak widths differ. The correlation matrix is constructed using only the training focus stacks and remains fixed during validation and testing.

3.3. Reliability-Aware Representative-Band Selection

A strongly correlated band may still be unsuitable for autofocus because of low signal quality or unstable focal-position estimation. Therefore, each band is assigned a reliability score:

$$r_i = \eta s_i + (1 - \eta) c_i, \qquad 0 \le \eta \le 1, \tag{20}$$

where s_i is the normalized signal-to-noise score and c_i represents the repeatability of the estimated focal position.

The ability of band i to represent band j is measured by

$$A_{ij} = r_i C_{ij}. \tag{21}$$

For a candidate representative subset $\mathcal{B}$, its full-band coverage is defined as

$$J(\mathcal{B}) = \frac{1}{31} \sum_{j=1}^{31} \max_{i \in \mathcal{B}} A_{ij}. \tag{22}$$

The representative-band subset is obtained by solving

$$\mathcal{B}^* = \underset{\mathcal{B} \subseteq \mathcal{V}, |\mathcal{B}| = K}{\arg\max}\ J(\mathcal{B}), \tag{23}$$

where $\mathcal{V}$ is the complete 31-band set and K is the required number of representative bands.

A greedy strategy is used in practice. The representative subset determined in this offline stage was fixed thereafter and was used unchanged during all online autofocus and

independent test operations. At each greedy iteration, the candidate band was evaluated by its marginal coverage gain relative to the bands already selected; consequently, a channel highly redundant with the existing subset contributes little additional coverage, whereas a channel that represents previously undercovered spectral regions produces a larger gain. The procedure therefore favors a compact combination of reliable and complementary wavelength channels rather than simply selecting the individually strongest bands. Starting from an empty set, the band providing the largest increase in $J(\mathcal{B})$ is added at each iteration until K bands have been selected. The final value of K is determined using the validation set according to autofocus accuracy, full-band quality retention, and acquisition time. The coverage criterion does not select each representative band according to an independent average dependence score. Because Equation (22) retains, for every target wavelength, only the strongest representation already provided by the current subset, the benefit of a newly added band depends on what has previously been selected. A candidate channel that is highly redundant with an existing representative therefore provides little additional coverage, even if it exhibits strong dependence with many spectral bands. In contrast, a channel that provides strong dependence for wavelength regions that remain insufficiently represented can produce a substantially larger increase in the overall coverage. The greedy procedure consequently suppresses repeated selection of spectrally redundant representatives and favors a combination of channels that jointly covers complementary portions of the wavelength-dependent axial-state information. In addition, the reliability weighting prevents a channel from being favored solely because of high statistical dependence when its signal quality or focal-position repeatability is insufficient.

3.4. Correlation-Guided Full-Band State Reconstruction

At each online decision step, only the selected representative bands are observed. Each spectral channel is treated as a node in a correlation graph, and the nonlinear band-correlation matrix defines the graph connectivity.

A reliability-aware adjacency matrix is constructed as

$$G_{ij} = \sqrt{r_i r_j}\, C_{ij}. \tag{24}$$

The node features include the current focus response, its temporal variation, the observation mask, and the band reliability. A lightweight graph convolutional encoder propagates the observed representative-band information across the spectral graph:

$$\mathbf{H}_t^{(\ell+1)} = \phi\left(\mathbf{S}\mathbf{H}_t^{(\ell)}\mathbf{W}^{(\ell)} + \mathbf{B}^{(\ell)}\right), \tag{25}$$

where $\mathbf{S}$ is the normalized graph matrix, $\mathbf{W}^{(\ell)}$ and $\mathbf{B}^{(\ell)}$ are trainable parameters, and $\phi(\cdot)$ is a nonlinear activation function.

The final graph layer produces an estimate of the complete 31-band focus vector, denoted by $\widehat{\mathbf{f}}_t$. This vector is substituted into the multispectral objective defined in Section 2 to obtain the estimated full-band quality $\widehat{Q}_{\mathrm{MS}}(z_t)$.

The reconstructor is trained using

$$\mathcal{L}_{\mathrm{rec}} = \frac{1}{31}\sum_{i=1}^{31}\left|\widehat{F}_i(z) - \bar{F}_i(z)\right| + \xi\left[\widehat{Q}_{\mathrm{MS}}(z) - Q_{\mathrm{MS}}(z)\right]^2. \tag{26}$$

The first term preserves bandwise focus information, while the second term preserves the complete-cube autofocus objective. The reconstructor is pretrained and fixed before reinforcement learning to maintain a stable state representation.

3.5. Deep Reinforcement Learning Controller

The autofocus process is modeled as a finite-horizon Markov decision process. The controller state contains the normalized motor position, previous motor action, representative-band focus responses, temporal response variations, graph embedding, estimated multispectral quality, and normalized decision step:

$$\mathbf{s}_t = \left[\widetilde{z}_t, \widetilde{a}_{t-1}, \bar{\mathbf{F}}_{\mathcal{B},t}^{\mathrm{T}}, \Delta\bar{\mathbf{F}}_{\mathcal{B},t}^{\mathrm{T}}, \mathbf{g}_t^{\mathrm{T}}, \widehat{Q}_{\mathrm{MS}}(z_t), \frac{t}{T_{\max}}\right]^{\mathrm{T}}. \tag{27}$$

The discrete action set contains coarse, medium, and fine movements in both axial directions, together with a stopping action:

$$\mathcal{A} = \{-s_3, -s_2, -s_1, +s_1, +s_2, +s_3, a_{\mathrm{stop}}\}. \tag{28}$$

For a movement action, the reward is defined as

$$r_t = Q_{\mathrm{MS}}(z_{t+1}) - Q_{\mathrm{MS}}(z_t) - \lambda_m \frac{|z_{t+1} - z_t|}{z_M - z_1} - \lambda_s - \lambda_r \mathbb{I}_{\mathrm{rev}}(t), \tag{29}$$

where the first term rewards improvement in full-band focus quality, while the remaining terms penalize motor travel, additional search steps, and unnecessary direction reversals. Successful stopping near the reference multispectral focal position receives a positive terminal reward, whereas premature stopping or timeout receives a negative reward.

A dueling double deep Q-network is used to estimate the action value:

$$Q(\mathbf{s}, a) = V(\mathbf{s}) + A(\mathbf{s}, a) - \frac{1}{|\mathcal{A}|}\sum_{a' \in \mathcal{A}} A(\mathbf{s}, a'). \tag{30}$$

The dueling structure separates the quality of the current optical state from the relative benefit of individual motor actions. Double Q-learning is used to reduce action-value overestimation, while a target network, replay memory, Huber loss, and ϵ-greedy exploration are employed during training.

3.6. Offline Training and Online Autofocus

The reinforcement learning environment is constructed directly from experimentally acquired focus stacks. At the beginning of each episode, a training scene and an initial motor position are selected randomly. The agent moves among measured axial positions rather than synthetically generated defocus states, thereby preserving wavelength-dependent noise, contrast, response asymmetry, and chromatic effects.

The training procedure includes four stages: construction of the nonlinear correlation graph, representative-band selection, pretraining of the full-band reconstructor, and reinforcement learning of the motor-control policy. Model selection is based on validation focal-position error and multispectral quality retention rather than on training reward alone.

All model components determined during offline modeling were frozen before online deployment. Specifically, the nonlinear correlation graph, representative-band subset, pretrained full-band reconstructor, and trained DDQN controller were kept unchanged throughout each online autofocus operation. Online focusing therefore involved only acquisition of the current representative-band observations, reconstruction of the corresponding optical state, and action inference; no graph recalibration, parameter update, or controller retraining was performed during an autofocus episode. During online operation, the representative bands are acquired at the current motor position, and the full-band state is reconstructed. The controller then selects the action with the highest estimated action value. The motor continues moving until the stopping action is issued or the maximum search

budget is reached. The complete 31-band multispectral cube is subsequently captured at the final focal position.

If T observations are required, the proposed method acquires KT intermediate spectral images, whereas complete-band acquisition at every position would require $31T$ images. The theoretical reduction in intermediate band acquisition is therefore

$$\mathcal{R}_{\text{band}} = 1 - \frac{K}{31}. \tag{31}$$

The actual time reduction additionally depends on motor displacement, exposure duration, wavelength-switching latency, image transfer, graph reconstruction, and network inference.

4. Theoretical Analysis

This section provides concise theoretical support for representative-band selection, full-band state reconstruction, and reinforcement learning decision stability.

4.1. Representative-Band Selection

For each target band, the coverage objective in Equation (22) retains the largest representation coefficient supplied by the selected subset. Because adding a new band cannot reduce any existing coverage value, $J(\mathcal{B})$ is monotone.

Moreover, the additional benefit of a new band decreases as the selected subset becomes larger, because more target bands have already been covered by existing representatives. Therefore, $J(\mathcal{B})$ satisfies the diminishing-return property and is submodular.

For a fixed subset size K, the greedy representative-band selection satisfies

$$J(\mathcal{B}_{\text{greedy}}) \geq \left(1 - \frac{1}{e}\right) J(\mathcal{B}_{\text{opt}}), \tag{32}$$

where $\mathcal{B}_{\text{opt}}$ denotes the globally optimal subset. Thus, the greedy strategy provides a bounded approximation while avoiding exhaustive evaluation of all band combinations.

4.2. Effect of Reconstruction Error

Assume that the maximum bandwise reconstruction error is bounded by

$$\left\| \widehat{\mathbf{f}}(z) - \mathbf{f}(z) \right\|_{\infty} \leq \epsilon_f. \tag{33}$$

The weighted mean, dispersion, minimum-band response, and penalty term used in Q_{MS} are continuous functions of the normalized focus vector. Therefore, there exists a finite constant L_{MS} such that

$$\left| \widehat{Q}_{\text{MS}}(z) - Q_{\text{MS}}(z) \right| \leq L_{\text{MS}} \epsilon_f. \tag{34}$$

This result shows that the error of the estimated multispectral objective increases at most linearly with the full-band reconstruction error. Consequently, improved graph reconstruction directly increases the reliability of the autofocus state.

4.3. Stability of the Motor Decision

Let $\mathbf{s}_t$ denote the ideal complete-band controller state and let $\widehat{\mathbf{s}}_t$ denote the reconstructed state. Assume that their difference is bounded by ϵ_s and that the learned action-value function is locally Lipschitz continuous with constant L_A.

Let the action-value margin between the best and second-best actions under the ideal state be Δ_t. If

$$\Delta_t > 2L_A \epsilon_s, \tag{35}$$

the perturbation introduced by state reconstruction is insufficient to change the action ordering. The controller therefore selects the same motor action from both the ideal and reconstructed states.

This condition explains why accurate representative-band reconstruction supports stable closed-loop focusing. In ambiguous regions where the action margin is small, subsequent observations allow the controller to update the optical state and correct an uncertain intermediate decision.

5. Experimental Results and Discussion

5.1. Experimental Setup and Comparison Methods

The experimental dataset consisted of 75 independent imaging scenes, including standardized resolution patterns, high-texture targets, low-texture targets, spectrally heterogeneous targets, and targets with nonplanar surface structures. Each category contained 15 scenes. To maintain a unified quantitative presentation, the results were illustrated using through-focus curves, correlation heatmaps, modulation transfer functions, autofocus trajectories, box plots, and statistical charts rather than physical photographs.

For each scene, a complete 31-band through-focus stack was acquired at 61 uniformly spaced axial positions. The interval between adjacent positions was 20 μm, resulting in an axial search range of 1.20 mm. The 20 μm axial interval also defined the effective motor-control resolution used throughout the autofocus experiments; hence, one reported motor step corresponds to an axial displacement of 20 μm. Motor-positioning uncertainty can directly perturb the final focal location, and its influence was therefore evaluated separately in the robustness experiment. A one-step backlash perturbation increased the mean focal-position error from 0.56 ± 0.44 to 0.74 ± 0.59 steps, while maintaining a success rate of 97.5% and multispectral quality retention of 98.8%. This result indicates that positioning precision affects focal localization, although the closed-loop controller remains stable under an axial perturbation comparable to one effective control step. Each stack was acquired three times after independent motor reinitialization. The scenes were divided into 45 training scenes, 15 validation scenes, and 15 independent test scenes. All bands, axial positions, and repeated acquisitions belonging to the same physical scene were assigned to the same subset to prevent information leakage.

Each test scene was evaluated from seven initial motor positions distributed across the axial range. Together with three repeated acquisitions, this protocol produced 315 independent autofocus episodes. Autofocus was considered successful when the final position was within two motor steps, corresponding to 40 μm, of the reference multispectral optimum.

The graph reconstructor contained two graph-convolution layers with 32 hidden units. The dueling double deep Q-network used two shared fully connected layers with 128 and 64 units. The replay-memory capacity was 50,000 transitions, the mini-batch size was 128, and the learning rate was 1×10^{-4}. The discount factor was 0.95, and the target network was updated every 200 optimization steps. The exploration probability decreased from 1.0 to 0.05 during the first 4000 episodes. Training was continued for 6000 episodes using five independent random seeds.

The proposed method was compared with exhaustive full-band scanning, single-band hill climbing, golden-section search, coarse-to-fine full-band search, uniform five-band fusion, signal-to-noise-ratio-weighted five-band fusion, standard DQN, and dueling DDQN without nonlinear correlation guidance.

Two recently published autofocus methods were additionally reproduced. The first was the SIFT Quad-Tenen and Pearson-Hill Climbing method developed for multispectral illumination [23]. It combines a multiscale gradient-based sharpness measure with Pearson correlation and a deterministic hill-climbing search. The second was the hybrid deep-

regression and variable-step hill-climbing method [28]. It uses an attention-enhanced ShuffleNetV2 model to estimate defocus distance and subsequently refines the predicted position through variable-step local search.

The deep-regression baseline was retrained using the same training and validation scenes. Its input was the 570 nm channel because the original approach was formulated as a single-image defocus-regression method. All methods used the same axial range, motor grid, initial positions, stopping tolerance, and test scenes. No method except exhaustive scanning was allowed to access the complete test focus stack before selecting its final position.

The evaluation metrics included focal-position error, success rate, multispectral quality retention, MTF50 retention, weakest-band sharpness retention, number of motor decisions, spectral exposure count, and total autofocus time. Paired scene-level comparisons were performed using the Wilcoxon signed-rank test with Holm correction.

5.2. Wavelength-Dependent Focus Characteristics and Band Selection

Relative to the best-focus position at 570 nm, the mean focal shift was -4.1 steps at 360 nm, -2.6 steps at 455 nm, $+1.9$ steps at 805 nm, and $+4.6$ steps at 980 nm. The average wavelength-dependent focal spread across the test scenes was 8.7 ± 1.4 steps, corresponding to 174 ± 28 µm.

The mean full width at half maximum of the focus-response curve was 8.2 ± 1.1 steps at 455 nm, 9.5 ± 1.0 steps at 570 nm, 11.8 ± 1.3 steps at 805 nm, and 13.4 ± 1.6 steps at 940 nm. The longer-wavelength channels therefore exhibited broader responses and lower axial sensitivity around their peak positions.

When only the 570 nm channel was used for autofocus, the mean full-band quality retention was 94.2%, and the weakest-band sharpness retention decreased to 88.6%. The mean MTF50 was 0.258 ± 0.017 cycles/pixel at the single-band focal position and 0.276 ± 0.014 cycles/pixel at the reference multispectral position. These results confirm that a single visible channel did not provide the best focal position for the complete 360–980 nm spectral cube. The wavelength-dependent axial behavior and the limitation of single-band focusing are visualized in Figure 3. The figure jointly compares through-focus response displacement, response width, and the bandwise sharpness retained at the single-band and multispectral focal positions.

Figure 3a–c shows that the focus responses do not undergo only a uniform axial translation. Their peak locations, widths, and local sensitivities change jointly with wavelength. In particular, the broader responses at longer wavelengths indicate reduced axial sensitivity around the individual peak. Figure 3d further shows that focusing at the 570 nm optimum preserves the quality of the reference channel but progressively sacrifices channels farther from that wavelength. The near-zero shift at 570 nm in Figure 3b results from the definition of the relative axial coordinate, for which the best-focus position of the 570 nm channel is used as the reference. This channel still requires physical axial adjustment of the motorized focusing unit to reach its actual best-focus position. Likewise, the comparison in Figure 3d does not involve post-acquisition digital refocusing or software-based image compensation. The two conditions correspond to images physically acquired at the 570 nm best-focus position and at the reference multispectral focal position, respectively. The computational framework determines the motor stopping position, whereas the final focus adjustment itself is realized mechanically through axial displacement of the imaging system.

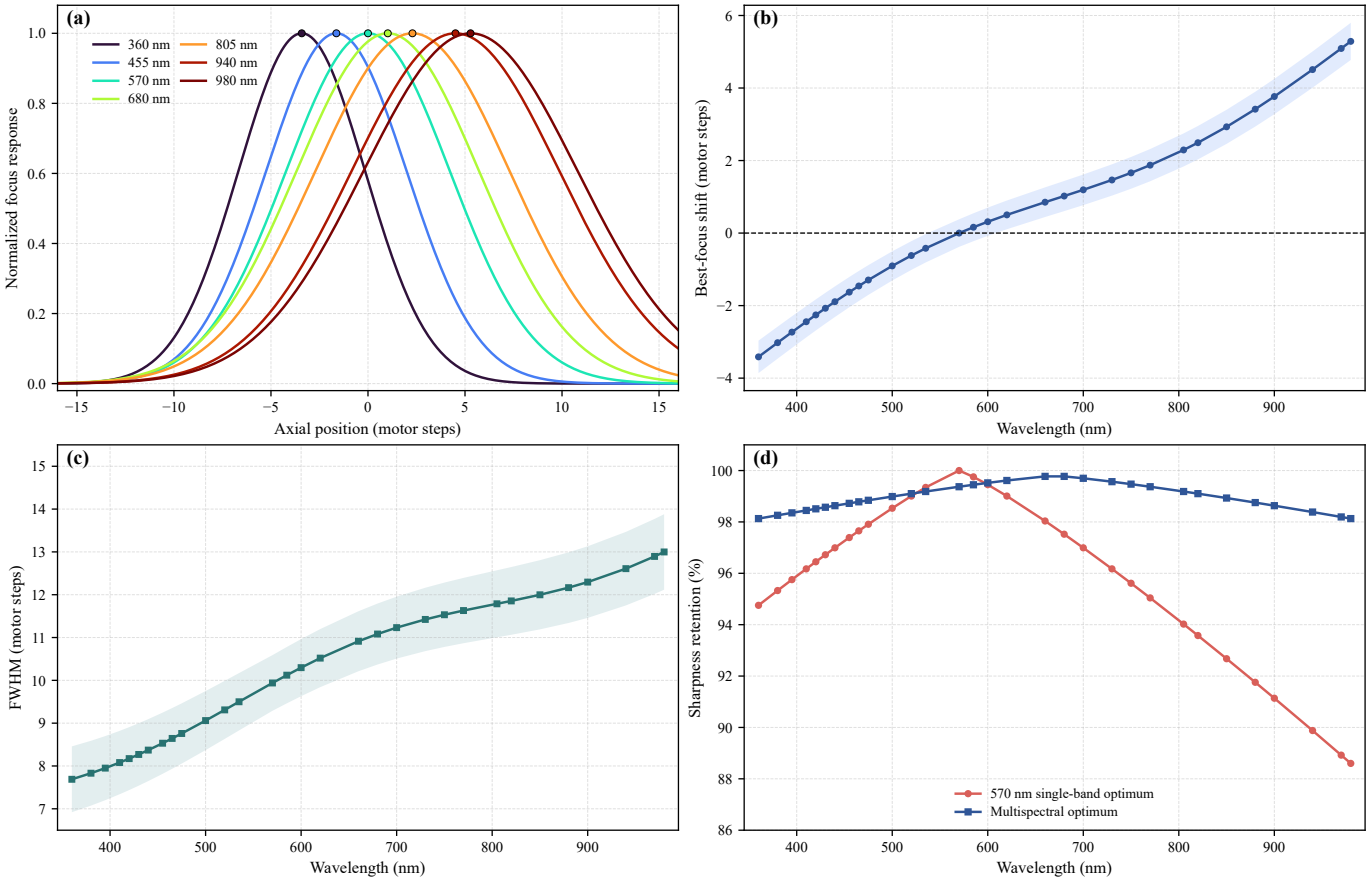


Figure 3. Wavelength-dependent focus characteristics of the 31-band imaging system. (**a**) Representative normalized through-focus response curves at selected visible and near-infrared wavelengths. (**b**) Mean wavelength-dependent best-focus shift relative to the reference position, with the shaded region indicating inter-scene variability. (**c**) Full width at half maximum of the focus-response curves as a function of wavelength. (**d**) Bandwise sharpness retention obtained at the 570 nm single-band optimum and at the reference multispectral optimum.

The mean within-visible correlation was 0.84, while the mean within-near-infrared correlation was 0.87. The average correlation between visible and near-infrared channels was lower at 0.62, but several cross-region pairs still exhibited substantial nonlinear dependence.

The reliability-aware greedy strategy selected five representative wavelengths:

$$\mathcal{B}^* = \{395, 520, 680, 820, 940\}\ \text{nm}. \tag{36}$$

The selected subset achieved a nonlinear full-band coverage score of 0.944. The wavelengths were distributed across the short-visible, middle-visible, red-edge, and near-infrared regions rather than being concentrated around the highest-intensity channels. The difference between linear and nonlinear spectral dependence, together with the resulting representative-band configuration, is summarized in Figure 4.

The two matrices in Figure 4a,b share a strong near-diagonal structure because adjacent wavelengths generally exhibit related focus trajectories. However, the nonlinear model reveals additional dependencies between spectrally separated visible and near-infrared channels, as emphasized by Figure 4c. The selected nodes in Figure 4d are consequently distributed across the full wavelength range rather than concentrated in a narrow high-correlation region. The coverage curve in Figure 4e exhibits a clear diminishing-return behavior beyond five bands.

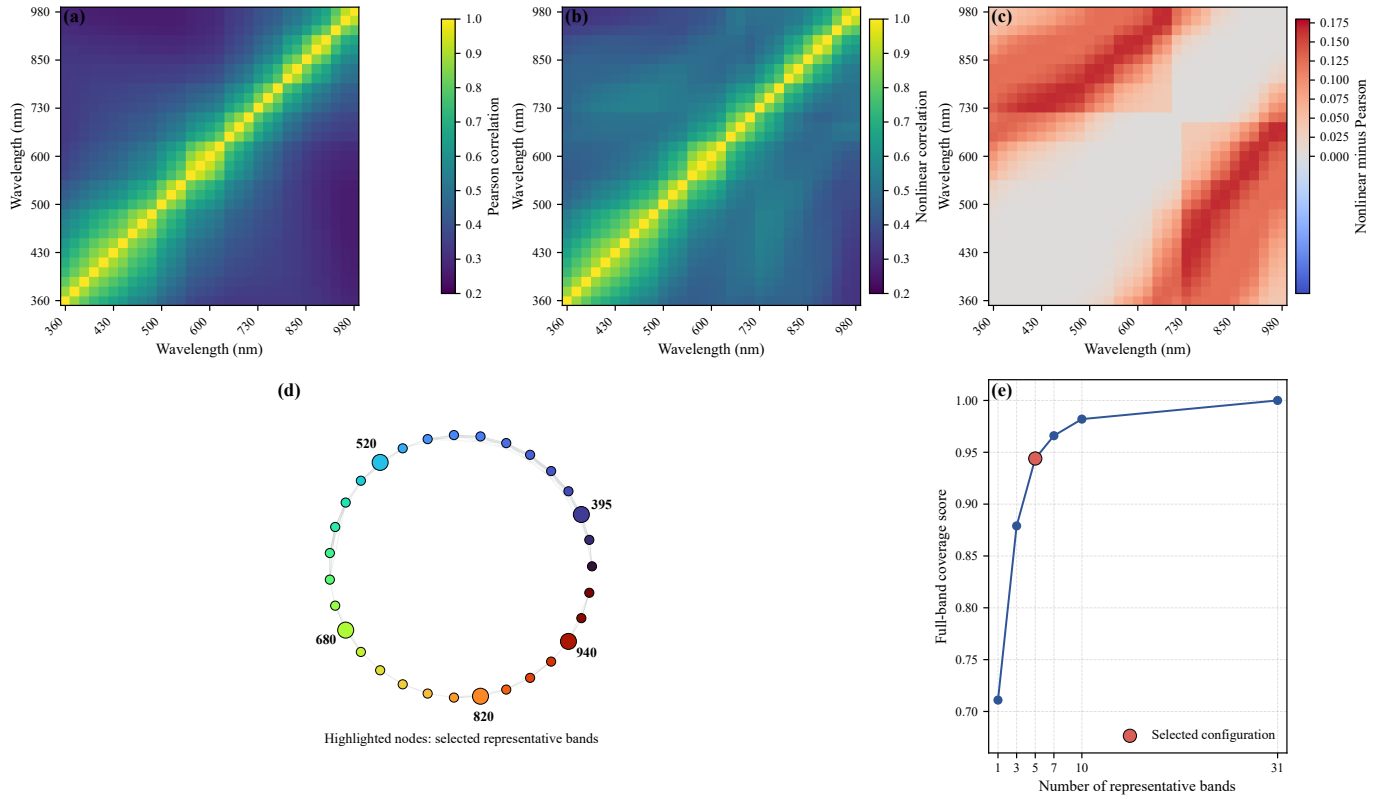


Figure 4. Linear and nonlinear dependence among the 31 spectral focus-response trajectories. (**a**) Pearson correlation matrix. (**b**) Nonlinear kernel-dependence matrix used by the proposed framework. (**c**) Difference between nonlinear dependence and Pearson correlation. (**d**) Spectral correlation graph, with the selected representative bands at 395, 520, 680, 820, and 940 nm highlighted by enlarged nodes. (**e**) Full-band coverage score as a function of the number of representative bands; the adopted five-band configuration is highlighted.

The differences revealed in Figure 4c also translated into measurable autofocus improvements. Under the same dueling-DDQN control architecture, replacing the Pearson-correlation graph with the nonlinear correlation graph reduced the mean focal-position error from 0.66 to 0.60 motor steps, corresponding to a 9.1% reduction, while increasing the success rate from 98.4% to 98.9% and the multispectral quality retention from 99.1% to 99.3%. The benefit was more pronounced for spectrally heterogeneous scenes, for which the focal-position error decreased from 0.79 to 0.64 steps (19.0%), compared with 0.48 to 0.43 steps for standardized resolution patterns. This result indicates that the additional nonlinear dependencies highlighted in Figure 4c are not merely statistical differences in the correlation matrices; they provide useful axial-state information for autofocus, particularly when wavelength-dependent scene responses make a linear correlation model less representative.

Table 1 presents the effect of the representative-band number. The one-band configuration in Table 1 corresponds to the single representative band selected by the reliability-aware greedy procedure, whereas the 94.2% full-band quality retention reported above refers specifically to the conventional autofocus baseline using the fixed 570 nm channel. These two results therefore represent different single-band configurations and are not directly equivalent.

Increasing the subset size from one to five substantially reduced focal-position error. Beyond five bands, the improvements in error and quality retention were limited, whereas acquisition time continued to increase. Therefore, $K = 5$ was used in the remaining experiments.

Table 1. Effect of the representative-band number on autofocus performance.

Number of Bands	Coverage Score	Focus Error (Steps)	Quality Retention (%)	Autofocus Time (s)
1	0.711	2.14 ± 1.53	96.1	2.7
3	0.879	0.91 ± 0.72	98.7	3.2
5	0.944	0.56 ± 0.44	99.4	3.8
7	0.966	0.51 ± 0.41	99.5	4.6
10	0.982	0.49 ± 0.39	99.6	6.1
31	1.000	0.46 ± 0.37	99.7	17.9

To determine whether autofocus performance depended on the specific band composition rather than only on the subset size, three alternative five-band configurations were evaluated under exactly the same full-band reconstruction, DDQN controller, action space, and stopping criteria. These included a uniformly spaced subset across the spectral range, a reliability-only subset formed by the five channels with the highest individual reliability scores, and a Pearson-correlation-based subset selected using the same greedy coverage procedure but replacing the nonlinear dependence matrix with linear Pearson correlation. The uniformly spaced, reliability-only, and Pearson-based subsets achieved nonlinear full-band coverage scores of 0.889, 0.907, and 0.928, respectively, compared with 0.944 for the proposed reliability-aware nonlinear-coverage subset. Their corresponding focal-position errors were 0.81 ± 0.61, 0.72 ± 0.55, and 0.65 ± 0.50 steps, with autofocus success rates of 97.1%, 97.8%, and 98.4% and multispectral quality retention values of 98.8%, 99.0%, and 99.2%, respectively. In comparison, the proposed five-band subset achieved 0.56 ± 0.44 steps, 99.2% success, and 99.4% quality retention. These results demonstrate that the autofocus performance is affected by the specific representative-band composition and that neither uniform spectral spacing nor high individual reliability alone is sufficient to preserve the complete axial optical state. The superior performance of the proposed subset supports the use of the joint reliability-nonlinear-coverage criterion, which favors channels that are individually stable while providing complementary coverage of spectral regions that are insufficiently represented by the bands already selected.

5.3. Full-Band Reconstruction and Training Convergence

The graph reconstructor achieved a mean absolute error of 0.018 and a root-mean-square error of 0.026 for the normalized 31-band focus vector. The coefficient of determination between reconstructed and measured focus responses was 0.972.

The mean absolute error of the reconstructed multispectral objective was 0.011. The reconstruction errors of the middle-visible and central near-infrared channels ranged from 0.013 to 0.020. The largest errors occurred at 360 and 980 nm, with mean absolute errors of 0.031 and 0.034, respectively. These channels had lower signal-to-noise ratios and greater focus-response variability.

The mean full-band quality estimation error remained below 0.015 over 94.6% of the test motor positions. The estimated multispectral optimum differed from the exhaustive-stack optimum by less than one motor step in 96.8% of the test focus stacks. This result indicates that five representative channels retained sufficient axial-state information for full-band focus estimation.

The mean number of motor decisions decreased from 13.8 during the initial training stage to 6.5 after convergence. During the final 500 episodes, the standard deviation of validation error across the five random seeds was below 0.12 steps. The final success rates of the five independently trained models ranged from 98.4% to 99.0%, indicating stable convergence.

The graph reconstructor was fixed before reinforcement learning. Consequently, the observed convergence reflected improvement of the motor-control policy rather than simultaneous changes in the optical-state representation. The reconstruction accuracy and the reinforcement-learning convergence process are jointly summarized in Figure 5.

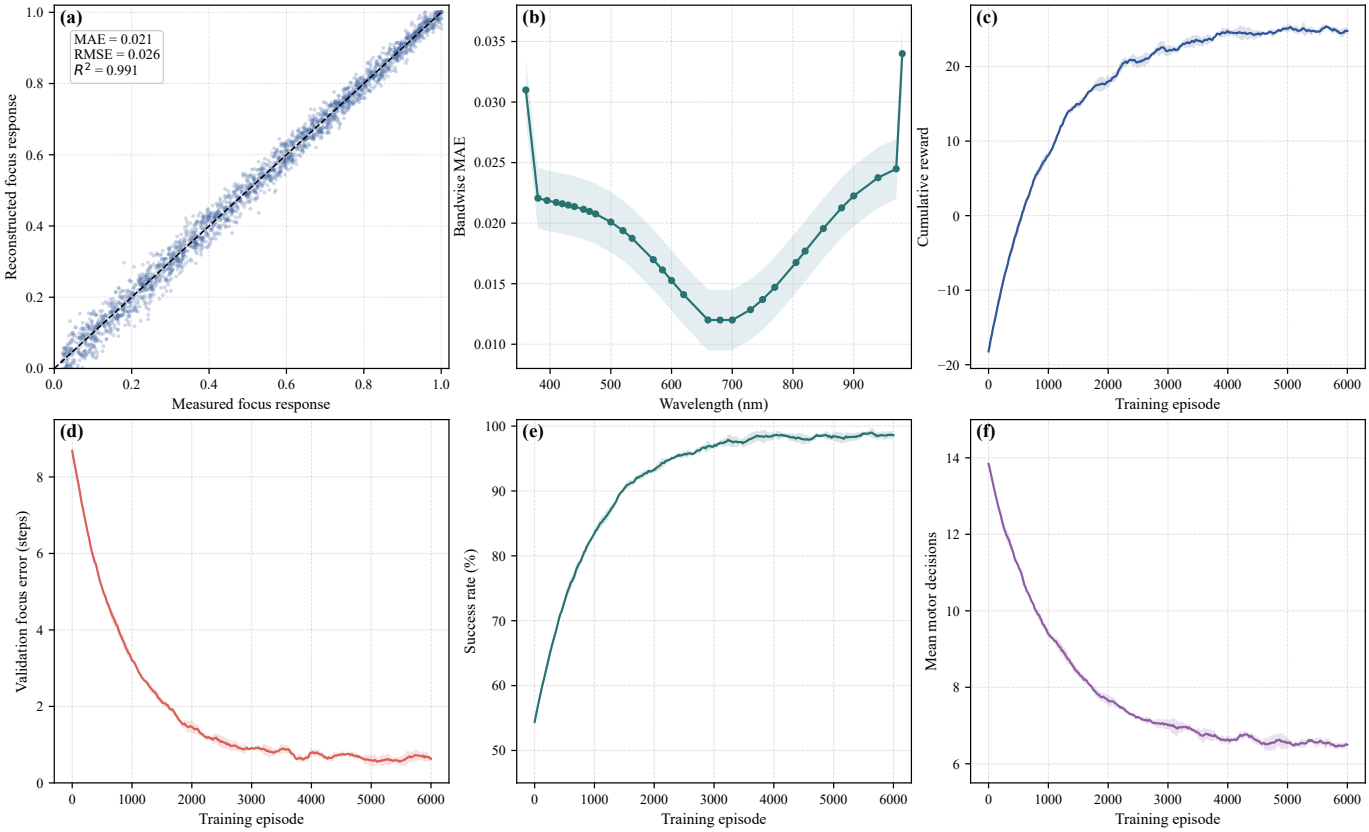


Figure 5. Full-band focus-state reconstruction and reinforcement-learning convergence. (**a**) Reconstructed versus measured normalized focus responses. (**b**) Bandwise reconstruction mean absolute error across the spectral range. (**c**) Cumulative reward during training. (**d**) Validation focal-position error. (**e**) Autofocus success rate. (**f**) Mean number of motor decisions. Shaded regions in the training curves indicate variability across the independent random seeds.

The close alignment with the identity line in Figure 5a confirms that the graph model retained the principal full-band optical-state information from the five observed channels. The wavelength-resolved errors in Figure 5b increase toward the two spectral boundaries, consistent with their lower signal-to-noise ratios and greater focus-response variability. The training curves in Figure 5c–f show mutually consistent convergence: cumulative reward and success rate increase, whereas validation error and the number of motor decisions decrease. The simultaneous stabilization of these metrics indicates that the learned policy improved both focal accuracy and search efficiency rather than increasing reward through excessive motor actions.

5.4. Comparison with Conventional and Recent Autofocus Methods

Table 2 compares the proposed method with the conventional, learning-based, and recent literature baselines.

The SIFT Quad-Tenen and Pearson-Hill methods achieved a mean focal-position error of 0.92 ± 0.69 steps. Its enhanced sharpness function provided stable focus evaluation, while Pearson-guided spectral ordering reduced redundant search. However, the method relied on linear dependence and a predetermined hill-climbing trajectory.

Table 2. Autofocus performance on the independent test set.

Method	Focus Error (Steps)	Success Rate (%)	Quality Retention (%)	MTF50 Retention (%)	Decisions	Time (s)
Exhaustive full-band scan	0.00	100.0	100.0	100.0	61.0	94.8
Single-band hill climbing	3.46 ± 2.28	79.7	94.2	91.8	11.6	4.9
Golden-section search	2.62 ± 1.91	86.0	95.5	93.0	10.1	4.4
Coarse-to-fine full-band search	1.38 ± 0.97	94.6	98.1	96.9	10.7	20.7
Uniform five-band fusion	1.21 ± 0.86	95.2	98.2	97.1	8.9	5.0
SNR-weighted five-band fusion	1.03 ± 0.78	96.8	98.6	97.6	8.4	4.8
SIFT Quad-Tenen + Pearson-Hill [23]	0.92 ± 0.69	97.1	98.8	98.0	7.9	5.1
Deep regression + variable-step hill climbing [28]	0.78 ± 0.61	97.8	99.0	98.3	7.2	4.4
Standard DQN	0.96 ± 0.71	96.8	98.7	97.9	7.8	4.5
Dueling DDQN without correlation	0.82 ± 0.63	97.8	98.9	98.2	7.2	4.2
Proposed method	$\mathbf{0.56 \pm 0.44}$	**99.2**	**99.4**	**99.0**	**6.4**	**3.8**

The deep-regression and variable-step hill-climbing method reduced the error to 0.78 ± 0.61 steps. The regression model provided an effective initial estimate, and the subsequent local search corrected part of the prediction error. Nevertheless, its single-channel observation did not explicitly represent the optical state of the complete multispectral cube.

The proposed method achieved a mean focal-position error of 0.56 ± 0.44 steps and a success rate of 99.2%. It reduced the error by 39.1% relative to SIFT Quad-Tenen with Pearson-Hill Climbing and by 28.2% relative to deep regression with variable-step hill climbing.

The proposed method also reduced autofocus time by 25.5% relative to the SIFT-based baseline and by 13.6% relative to the deep-regression baseline. Its multispectral quality retention reached 99.4%, while the MTF50 retention reached 99.0%.

The weakest-band sharpness retention was 97.6% for the proposed method, compared with 94.9% for SIFT Quad-Tenen with Pearson-Hill Climbing, 95.8% for deep regression with variable-step hill climbing, and 88.6% for single-band hill climbing.

Scene-level statistical testing showed that the proposed method produced significantly lower focal-position error than SIFT Quad-Tenen with Pearson-Hill Climbing ($p_{\mathrm{adj}} = 0.008$), deep regression with variable-step hill climbing ($p_{\mathrm{adj}} = 0.024$), and dueling DDQN without nonlinear correlation guidance ($p_{\mathrm{adj}} = 0.031$). A visual comparison of the principal autofocus baselines is provided in Figure 6.

Figure 6a shows that the proposed method not only reduced the mean error but also produced a narrower error distribution, indicating improved consistency across initial positions and scene types. Its success rate and multispectral quality retention were simultaneously the highest among the compared non-exhaustive methods, as shown in Figure 6b,c. The accuracy-time plane in Figure 6e places the proposed method closest to the preferred lower-left region, demonstrating that its accuracy improvement was not obtained through additional search time. The example trajectory in Figure 6f further illustrates that the controller approached the reference optimum through progressively smaller movements and stopped without the late-stage oscillation observed in some deterministic baselines.

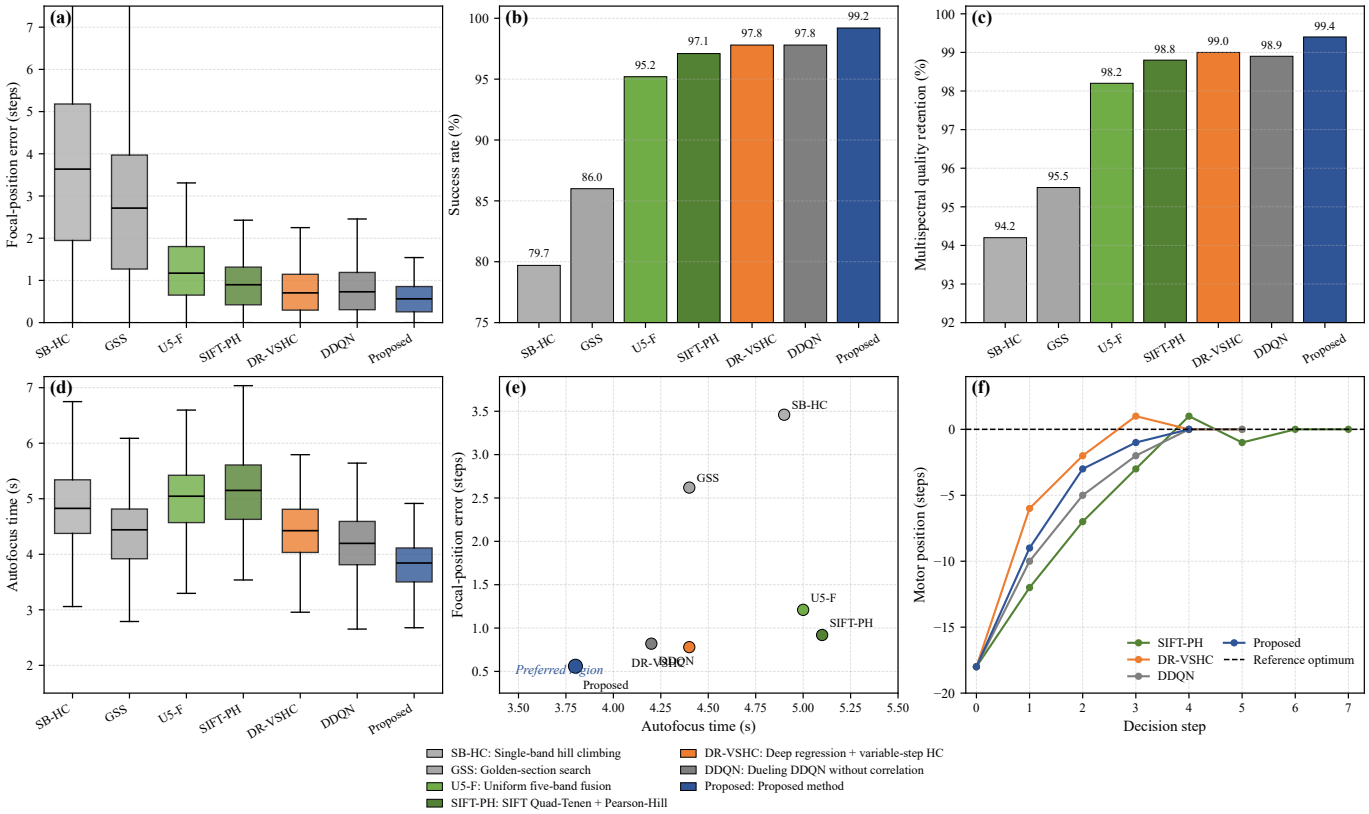


Figure 6. Comparison with representative conventional and recent autofocus methods on the independent test set. (**a**) Distribution of focal-position error. (**b**) Autofocus success rate. (**c**) Multispectral quality retention. (**d**) Distribution of total autofocus time. (**e**) Accuracy-time operating plane, where the lower-left region represents the preferred combination of low error and short focusing time. (**f**) Representative motor trajectories of the proposed method and selected competitive baselines.

5.5. Ablation and Robustness Analysis

Table 3 reports the contribution of the principal components.

Replacing standard DQN with the dueling double architecture reduced the error from 0.96 to 0.82 steps. Reliability-aware band selection further reduced the error to 0.73 steps by excluding channels with unstable signals or focal localization.

Table 3. Ablation results on the independent test set.

Configuration	Focus Error (Steps)	Success Rate (%)	Quality Retention (%)	Time (s)
Standard DQN	0.96 ± 0.71	96.8	98.7	4.5
Dueling DDQN	0.82 ± 0.63	97.8	98.9	4.2
Dueling DDQN + reliability-aware bands	0.73 ± 0.57	98.1	99.0	4.1
Dueling DDQN + Pearson graph	0.66 ± 0.51	98.4	99.1	4.0
Dueling DDQN + nonlinear graph	0.60 ± 0.47	98.9	99.3	3.9
Complete method	$\mathbf{0.56 \pm 0.44}$	**99.2**	**99.4**	**3.8**

Introducing a Pearson graph reduced the error to 0.66 steps, confirming that explicit inter-band dependence improved state representation. Replacing Pearson correlation with nonlinear kernel dependence further reduced the error to 0.60 steps. The complete multispectral objective produced the final error of 0.56 steps and increased the weakest-band sharpness retention from 96.9% to 97.6%.

The nonlinear dependence model was particularly beneficial for spectrally heterogeneous scenes. For these scenes, the focal-position errors of the Pearson graph and nonlinear graph were 0.79 and 0.64 steps, respectively. For standardized resolution patterns, the difference was smaller, at 0.48 and 0.43 steps, respectively.

The cumulative ablation sequence also provides a quantitative indication of the relative contribution of the principal components. Replacing the standard DQN with the dueling double architecture produced the largest single reduction in focal-position error, from 0.96 to 0.82 steps, corresponding to a 14.6% improvement relative to the preceding configuration. Reliability-aware band selection provided the second-largest incremental reduction, from 0.82 to 0.73 steps (11.0%). Introducing explicit Pearson-correlation guidance further reduced the error by 9.6%, and replacing the linear graph with the nonlinear dependence graph yielded an additional 9.1% reduction. Finally, incorporating the complete multispectral objective reduced the error from 0.60 to 0.56 steps (6.7%). The complete framework decreased the focal-position error by 41.7% relative to the standard-DQN configuration. These results indicate that the controller architecture and reliability-aware spectral observation provide the largest individual gains, while correlation-guided state representation, nonlinear dependence modeling, and the complete multispectral objective contribute complementary improvements that progressively refine full-band focusing accuracy.

The trained controller was additionally evaluated under imaging and motor perturbations without retraining. The results are shown in Table 4.

Table 4. Robustness under different imaging and control conditions.

Condition	Focus Error (Steps)	Success Rate (%)	Quality Retention (%)
Nominal test condition	0.56 ± 0.44	99.2	99.4
Illumination variation of $\pm 20\%$	0.64 ± 0.49	98.4	99.1
Image noise at 25 dB	0.71 ± 0.56	97.8	98.9
One-step motor backlash	0.74 ± 0.59	97.5	98.8
One missing representative band	0.89 ± 0.68	95.9	98.4
Low-texture scenes	0.93 ± 0.72	95.4	98.2
Independent acquisition session	0.81 ± 0.64	96.7	98.6

The controller remained stable under moderate illumination and noise variations. A one-step motor backlash increased the mean focal-position error by 0.18 steps but did not substantially reduce the final multispectral quality.

Removing one representative band caused a larger performance decrease because part of the spectral graph was no longer directly observed. Nevertheless, the remaining graph connections retained 98.4% of the reference multispectral quality.

The influence of representative-band loss was further examined by masking each of the five selected channels individually without retraining the reconstructor or controller. Removing the 395, 520, 680, 820, and 940 nm channels resulted in mean focal-position errors of 0.84 ± 0.64, 0.78 ± 0.61, 0.92 ± 0.70, 0.96 ± 0.73, and 0.95 ± 0.71 steps, respectively, with corresponding success rates of 96.5%, 97.1%, 95.7%, 95.2%, and 95.0%. The associated multispectral quality retention values were 98.5%, 98.7%, 98.3%, 98.2%, and 98.3%, respectively. To evaluate a more adverse scene-dependent condition, each representative band was then masked in turn for every test scene, and the omission producing the largest focal-position error for that scene was retained as the scene-wise critical-band dropout. Under this worst-case evaluation, the mean focal-position error increased to 1.08 ± 0.79 steps, while the success rate and multispectral quality retention remained 93.7% and 97.9%, respectively. These results show that the importance of an individual representative channel is scene-dependent and that loss of the locally most informative band

produces a larger degradation than average single-band loss. Nevertheless, the remaining representative channels and correlation-guided state reconstruction preserve substantial axial-state information, allowing the controller to maintain stable autofocus performance even under this deliberately adverse missing-band condition.

Low-texture scenes produced the greatest deterioration because their gradient-based focus responses were comparatively flat. The proposed method still maintained a success rate of 95.4%. The independent-session result of 0.81 ± 0.64 steps indicates that the learned spectral dependencies were reasonably stable after repeated calibration and motor initialization. The component contributions, condition-dependent robustness, representative-band tradeoff, and acquisition-time decomposition are consolidated in Figure 7.

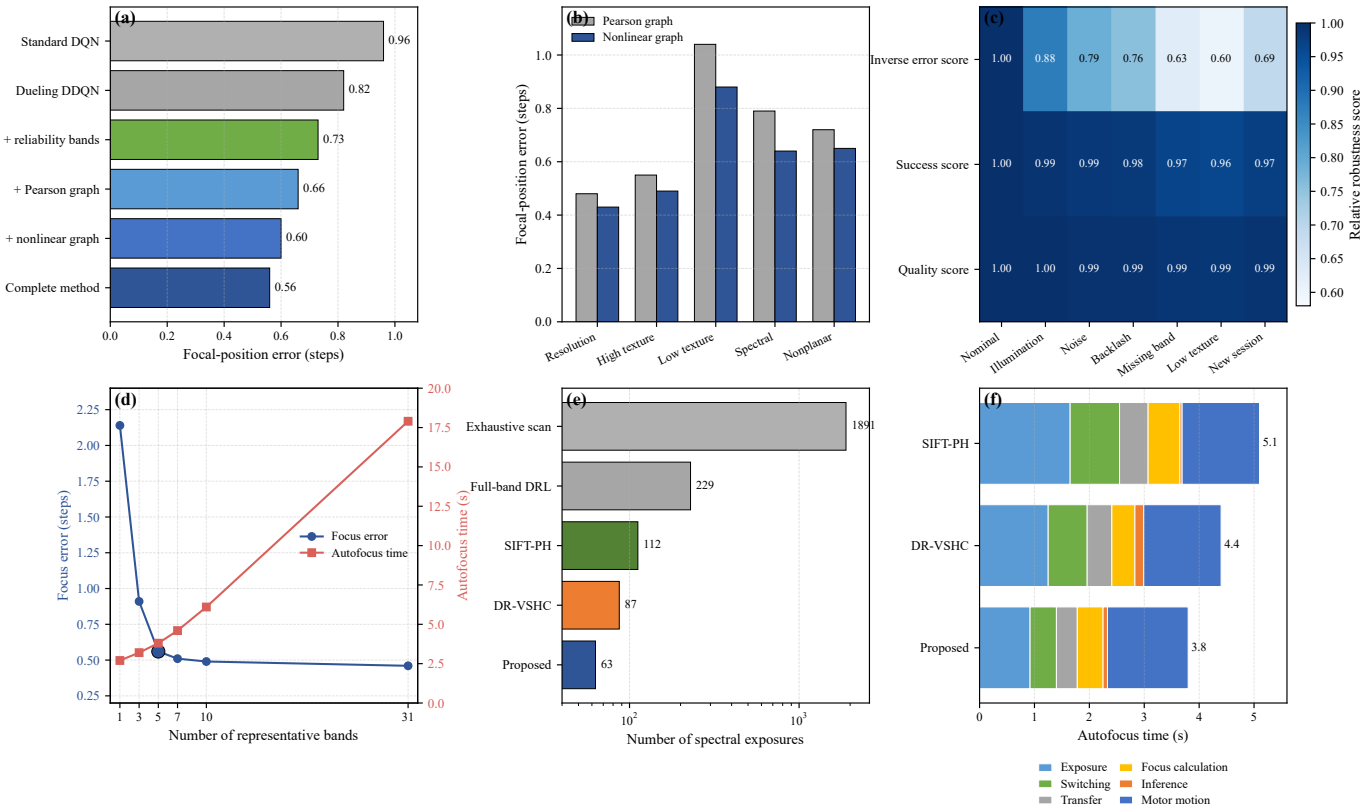


Figure 7. Ablation, robustness, and acquisition-efficiency analysis. (**a**) Incremental reduction in focal-position error obtained by the principal components of the proposed framework. (**b**) Comparison of Pearson and nonlinear correlation graphs over different scene categories. (**c**) Relative robustness scores under imaging and motor perturbations, with the inverse-error, success, and quality scores normalized to the nominal condition. (**d**) Joint effect of the representative-band number on focal-position error and autofocus time. (**e**) Number of spectral exposures required by representative autofocus strategies. (**f**) Decomposition of total autofocus time into exposure, wavelength switching, image transfer, focus calculation, inference, and motor motion.

The incremental results in Figure 7a confirm that the performance gain was distributed across the dueling architecture, reliability-aware band selection, correlation-guided reconstruction, and the complete multispectral objective. Figure 7b shows that nonlinear dependence was especially beneficial for low-texture and spectrally heterogeneous scenes, in which linear Pearson correlation provided a less complete description of the axial state. The normalized robustness map in Figure 7c indicates that final quality retention was comparatively stable, whereas focal-position error was more sensitive to missing bands and low-texture inputs. Finally, Figure 7d–f identifies five bands as the practical knee point

and shows that exposure and motor motion, rather than graph inference, account for most of the autofocus time.

To disentangle the contribution of reinforcement-learning control from that of multispectral observation, an additional single-band DRL ablation was performed using the same dueling-DDQN controller as the proposed method while restricting the optical observation to the 570 nm channel. The network architecture, action set, reward formulation, replay settings, training budget, and stopping criterion were kept unchanged, and only the spectral information available to the controller was reduced. The single-band DRL configuration achieved a focal-position error of 0.90 ± 0.68 steps, a success rate of 97.1%, and multispectral quality retention of 98.6%, with an average autofocus time of 3.6 s. Although this result was substantially better than conventional single-band hill climbing, it remained clearly inferior to the complete five-band correlation-guided framework, which achieved 0.56 ± 0.44 steps, 99.2% success, and 99.4% quality retention. The single-band DRL result was also slightly less accurate than the dedicated deep-regression and variable-step hill-climbing baseline (0.78 ± 0.61 steps), indicating that reinforcement-learning control alone does not account for the performance advantage of the proposed method. The further improvement obtained by the complete framework therefore arises primarily from combining adaptive sequential control with complementary multispectral observations and correlation-guided reconstruction of the full-band axial state.

5.6. Optical-Configuration Transferability and Recalibration

To evaluate whether the nonlinear correlation graph remains valid after moderate changes in the optical configuration, an additional transferability experiment was conducted under three perturbed conditions: a one-stop increase in the F-number with a reduced aperture and working-distance shifts of -10% and $+10\%$ relative to the nominal setting. The 15 validation scenes were reacquired under each condition and used only for configuration-specific graph recalibration, while the 15 independent test scenes were reacquired for performance evaluation. Graph stability was quantified by the Pearson correlation between the vectorized upper-triangular elements of the nominal and perturbed nonlinear correlation matrices. To isolate the effect of graph recalibration, the representative five-band set, graph-convolution parameters, and trained DDQN controller were kept unchanged; only the nonlinear dependence matrix and the corresponding reliability-aware adjacency matrix were recomputed from the calibration stacks.

The nonlinear graph retained substantial structural similarity under the tested configuration changes, with correlation coefficients of 0.928 for the F-number variation and 0.901 and 0.887 for the -10% and $+10\%$ working-distance shifts, respectively. When the original graph was directly transferred without recalibration, the mean focal-position error increased from 0.56 ± 0.44 steps under the nominal condition to 0.71 ± 0.53, 0.78 ± 0.60, and 0.82 ± 0.63 steps, while the corresponding success rates were 97.8%, 96.8%, and 96.2%. After graph recalibration, the errors decreased to 0.61 ± 0.47, 0.64 ± 0.49, and 0.66 ± 0.51 steps, with success rates of 98.7%, 98.1%, and 97.8% and multispectral quality retention values of 99.3%, 99.1%, and 99.0%, respectively. These results indicate that the learned nonlinear dependence structure is reasonably stable under moderate optical perturbations but is not strictly invariant to changes in imaging geometry. A lightweight recalibration of the correlation graph recovered most of the nominal autofocus performance without retraining the focusing policy. Consequently, the nonlinear graph should be regarded as configuration dependent, and recalibration is recommended when the optical configuration or working-distance range is changed substantially, particularly when components that directly modify the wavelength-dependent point-spread function, such as the imaging lens, are replaced.

5.7. Acquisition Efficiency and Overall Discussion

With five representative bands, the intermediate spectral acquisition reduction was

$$\mathcal{R}_{\text{band}} = 1 - \frac{5}{31} = 83.9\%. \tag{37}$$

The proposed method required an average of 6.4 representative-band observations, corresponding to approximately 32 intermediate spectral exposures. Including the final 31-band cube, the complete autofocus and acquisition procedure required approximately 63 spectral exposures.

Acquiring all 31 channels at the same 6.4 intermediate positions and again at the final position would require approximately 229 exposures. The proposed method therefore reduced the complete exposure count by approximately 72.5%.

The graph reconstructor contained approximately 26.4 thousand trainable parameters, while the dueling DDQN contained approximately 84.7 thousand parameters. The complete framework, therefore, contained approximately 111.1 thousand parameters.

The combined graph-reconstruction and action-selection time was 2.1 ms on a graphics processor and 8.6 ms on a desktop central processing unit. Focus-response extraction from five spectral images required approximately 74 ms. The computational cost was therefore considerably lower than the combined wavelength-switching, exposure, image-transfer, and motor-motion time.

The experimental results demonstrate that broadband multispectral autofocus can be formulated as a sequential full-band optical-state estimation and control problem. The SIFT Quad-Tenen and Pearson-Hill method effectively improved deterministic sharpness evaluation and exploited linear spectral similarity. However, it did not reconstruct the complete multispectral state and could not adapt its search policy beyond predefined hill-climbing rules.

The deep regression and variable-step hill-climbing method provided more accurate initial localization. Nevertheless, its prediction was derived from one spectral image, and its subsequent search optimized a local single-channel focus response rather than the complete spectral cube.

The proposed method combined nonlinear inter-band modeling, representative-band observation, full-band reconstruction, and adaptive sequential motor control. The correlation graph transferred information from the five observed channels to the unobserved bands, while the reinforcement learning controller adjusted its movement scale and stopping decision according to the reconstructed multispectral state.

The method approached the optical quality of exhaustive full-band scanning while requiring substantially fewer exposures and motor movements. The remaining errors were concentrated mainly in low-texture scenes and channels near the two ends of the spectral range, where signal-to-noise ratios and focus-response reliability were lower.

The present study was limited to one multispectral imaging system, one focusing lens, and a fixed working-distance range. Major changes in the optical configuration may require recalibration of the nonlinear correlation graph. In addition, the current online focus feature relies primarily on spatial gradients and may be less informative for targets with extremely weak texture. Although the present study emphasizes objective cross-band evaluation through focal-position error, MTF50 retention, full-band quality retention, and weakest-band sharpness retention, future work will further incorporate qualitative multispectral image comparisons and application-oriented visual assessment under representative imaging scenes to more directly illustrate the perceptual and task-level benefits of the proposed autofocus strategy. Future work will investigate uncertainty-aware stopping, device-transfer calibration, and joint spatial-spectral focus representations.

6. Conclusions

This study proposed a nonlinear band-correlation-guided deep reinforcement learning framework for autofocus optimization in a self-developed 31-band multispectral imaging system covering 360–980 nm. By modeling nonlinear dependencies among through-focus responses, selecting a compact set of representative bands, reconstructing the complete multispectral focus state, and learning adaptive motor direction, displacement, and stopping decisions, the proposed method optimized the quality of the complete spectral cube rather than that of a single channel. The five-band configuration achieved a mean focal-position error of 0.56 ± 0.44 motor steps, a success rate of 99.2%, multispectral quality retention of 99.4%, and MTF50 retention of 99.0%, while reducing intermediate spectral acquisition by 83.9%. It also outperformed the recent SIFT Quad-Tenen with the Pearson-Hill Climbing method and the deep-regression with variable-step hill-climbing method in focusing accuracy, full-band quality, and acquisition efficiency. These results demonstrate that nonlinear spectral dependence and sequential closed-loop control provide an effective computational-photonics solution for broadband multispectral autofocus. The present study is nevertheless limited to one imaging system, one lens configuration, and a fixed working-distance range, and the current gradient-based focus features may become less reliable for extremely low-texture scenes or channels with weak signal-to-noise ratios. Future work will therefore investigate cross-device calibration, uncertainty-aware stopping, and more general spatial-spectral focus representations.

Author Contributions: Conceptualization, M.Z. and C.M.; methodology, M.Z. and C.M.; software, Z.C. and S.W.; validation, Z.C., S.W. and S.J.; formal analysis, J.H. and M.Z.; investigation, S.J. and J.H.; resources, Z.C. and S.W.; data curation, S.J. and J.H.; writing-original draft preparation, Z.C., S.W. and M.Z.; writing-review and editing, C.M., Z.C., S.W. and M.Z.; visualization, S.J. and J.H.; supervision, C.M.; project administration, C.M. and M.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: Authors Zhenzhen Chen and Sen Wang are employed by Yuxin Electronic Technology Group Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

1. Luo, X.; Wang, S.; Liu, J.; Dong, X.; He, P.; Yang, Q.; Chen, X.; Zhou, F.; Zhang, T.; Feng, S.; et al. Revolutionizing optical imaging: Computational imaging via deep learning. *Photonics Insights* **2025**, *4*, R03. [CrossRef]
2. Wang, F.; Czarske, J.W.; Situ, G. Deep learning for computational imaging: From data-driven to physics-enhanced approaches. *Adv. Photonics* **2025**, *7*, 054002. [CrossRef]
3. Wang, Z.; Peng, Y.; Fang, L.; Gao, L. Computational optical imaging: On the convergence of physical and digital layers. *Optica* **2025**, *12*, 113–130. [CrossRef]
4. Chen, H.; Zhang, H.; Zou, B.; Wu, L. Hybrid Self-Attention Transformer U-Net for Fourier Single-Pixel Imaging Reconstruction at Low Sampling Rates. *Photonics* **2025**, *12*, 568. [CrossRef]
5. Joshi, N.; Xavier, A.P.I.; Gopinath, S.; Tiwari, V.; Anand, V. Spatial Ensemble Mapping for Coded Aperture Imaging—A Tutorial. *Photonics* **2024**, *11*, 1174. [CrossRef]
6. Rosen, J.; Anand, V. Optical Imaging Using Coded Aperture Correlation Holography (COACH) with PSF of Spatial-Structured Longitudinal Light Beams—A Study Review. *Photonics* **2024**, *11*, 115. [CrossRef]
7. Wang, D.-Y.; Bie, S.-H.; Chen, X.-H.; Yu, W.-K. Single-Pixel Infrared Hyperspectral Imaging via Physics-Guided Generative Adversarial Networks. *Photonics* **2024**, *11*, 174. [CrossRef]

8. Wang, H.; Wang, X.; Gao, C.; Wang, Y.; Zhao, H.; Yao, Z. Multi-Wavelength Computational Ghost Imaging Based on Feature Dimensionality Reduction. *Photonics* **2024**, *11*, 739. [CrossRef]
9. Zhou, X.; Zhang, C.; Zheng, Z.; Li, H.; Peng, C. A Deep Learning Model for Spectral Reconstruction of Arrayed Micro-Resonators. *Photonics* **2025**, *12*, 449. [CrossRef]
10. Ma, C.; Zhang, M.; Ma, J.; Chen, S.; Yuan, Y.; Du, M.; Dai, Z. Counterfeit Citri Reticulatae Pericarpium Identification: Multi-Class Detection Using Vis/NIR Spectroscopy and Residual Neural Networks. *Food Sci. Nutr.* **2026**, *14*, e71975. [CrossRef] [PubMed]
11. Shi, Z.; Dun, X.; Wei, H.; Dong, S.; Wang, Z.; Cheng, X.; Heide, F.; Peng, Y. Learned Multi-Aperture Color-Coded Optics for Snapshot Hyperspectral Imaging. *Acm Trans. Graph.* **2024**, *43*, 208. [CrossRef]
12. Yang, X.; Fanous, M.J.; Chen, H.; Lee, R.; Costa, P.C.; Li, Y.; Huang, L.; Zhang, Y.; Ozcan, A. Snapshot multi-spectral imaging through defocusing and a Fourier imager network. *Adv. Photonics Nexus* **2025**, *4*, 056002. [CrossRef]
13. Zeng, Y.; Wang, X.; Jiang, L.; Qi, J.-W.; Lin, Z.; Guo, T.; He, S. Fast multispectral imaging via hybrid-encoded LED illumination and a lightweight deep-learning model. *Opt. Lett.* **2025**, *50*, 6177–6180. [CrossRef] [PubMed]
14. Chen, N.; Wu, Y.; Tan, C.; Cao, L.; Wang, J.; Lam, E.Y. Uncertainty-aware Fourier ptychography. *Light Sci. Appl.* **2025**, *14*, 236. [CrossRef] [PubMed]
15. Li, X.; Wang, A.; Fan, M.; Yu, L.; Liang, X. Multi-Spectral and Single-Shot Wavefront Detection Technique Based on Neural Networks. *Photonics* **2025**, *12*, 1110. [CrossRef]
16. Wu, J.; Li, Y.; Luo, Y.; Pan, L.; Song, P.; Xu, Q. High-Fidelity Scanning-Free Lensless Microscopy via Adaptive OPD-Domain Fusion for Live-Cell and Tissue Imaging. *Photonics* **2026**, *13*, 213. [CrossRef]
17. Liu, J.; Sun, W.; Wu, F.; Shan, H.; Xie, X. Macroscopic Fourier Ptychographic Imaging Based on Deep Learning. *Photonics* **2025**, *12*, 170. [CrossRef]
18. Zhang, M.; Ma, C.; Li, Y.; Du, M.; Chen, S.; Jing, S.; Niu, H.; Ma, J. Reaction-informed spectral feature packets for lightweight Vis/NIR screening of sulfur-fumigation process states in plant-derived materials. *J. Food Compos. Anal.* **2026**, *158*, 109431. [CrossRef]
19. Fang, Z.; Song, Y.; Han, B.; Wang, A.; Song, J.; Yue, H. Autofocusing Method Based on Dynamic Modulation Transfer Function Feedback. *Photonics* **2026**, *13*, 107. [CrossRef]
20. Liu, Q.; Mao, X.; Wang, J.; Zhang, Q.; Wang, Y.; Pang, B.; Li, Q. FAN: Fourier Assignment Network for autofocus based on deep learning. *Opt. Laser Technol.* **2025**, *186*, 112579. [CrossRef]
21. Ojdanić, D.; Zelinskyi, D.; Naverschnigg, C.; Sinn, A.; Schitter, G. High-speed telescope autofocus for UAV detection and tracking. *Opt. Express* **2024**, *32*, 7147–7157. [CrossRef] [PubMed]
22. Ma, C.; Zhang, Q.; Zhang, M.; Ma, J.; Chen, Z.; Yan, Y. Multispectral autofocus optimization with a coarse-to-fine search and the Pearson correlation coefficient. *Appl. Opt.* **2025**, *64*, 8255–8265. [CrossRef] [PubMed]
23. Ma, C.; Zhang, M.; Dai, Z.; Zhang, Q.; Ma, J.; Niu, X.; Yan, Y. Optimizing Autofocus under Multispectral Lighting via Enhanced SIFT and Pearson Correlation Coefficient. *PLoS ONE* **2025**, *20*, e0336810. [CrossRef] [PubMed]
24. Wang, R.; Liu, T. Autofocus control using adaptive region selection and reinforcement learning applied in wafer micro-imaging. *Opt. Express* **2025**, *33*, 17345–17354. [CrossRef] [PubMed]
25. Zhang, J.; Fu, Y.-F.; Shen, H.; Liu, Q.; Sun, L.-N.; Chen, L.-G. Precision autofocus in optical microscopy with liquid lenses controlled by deep reinforcement learning. *Microsyst. Nanoeng.* **2024**, *10*, 201. [CrossRef] [PubMed]
26. Yu, C.; Ding, F.; Ma, Z.; Tang, Y. Rapid Autofocus Method Based on LED Oblique Illumination for Metaphase Chromosome Microscopy Imaging System. *Photonics* **2024**, *11*, 1091. [CrossRef]
27. Liu, J.; Jiang, P.; Yang, H.; Wang, D.; Wang, P.; Zhou, W. A Fast Autofocus System Based on the Advancement of the CGH Algorithm. *Photonics* **2026**, *13*, 70. [CrossRef]
28. Ma, X.; Xia, X.; Yuan, S.; Xiang, H. Hybrid high-precision autofocus method based on deep learning and variable-step hill-climbing. *Appl. Opt.* **2026**, *65*, 874–885. [CrossRef] [PubMed]
29. Wang, Y.; Wu, C.; Gao, Y.; Liu, H. Deep Learning-Based Dynamic Region of Interest Autofocus Method for Grayscale Image. *Sensors* **2024**, *24*, 4336. [CrossRef] [PubMed]
30. Zhai, Y.; Fu, X.; Su, Q.; Hu, J.; Zhang, Y.; Zhou, Y.; Zhang, C.; Li, X.; Wang, W.; Wu, D.; et al. SparseFocus: Learning-based one-shot autofocus for microscopy with sparse content. *Light. Adv. Manuf.* **2026**, *7*, 45. [CrossRef]
31. Rebuffi, L.; Zhang, R.; Shi, X.; Lin, Y.-C.; Zhao, J.; Hu, M.Y.; Toellner, T.; Cherukara, M.J.; Assoufid, L. AutoFocus: AI/ML-driven real-time wavefront diagnostics to autonomously align and optimize X-ray optics. *Opt. Express* **2025**, *33*, 37931–37955. [CrossRef] [PubMed]
32. Zhang, M.; Ma, C.; Chen, S.; Du, M.; Huang, J.; Niu, H.; Ma, J. Local consistency measurement for reliable Vis/NIR authentication of root food products under limited band conditions. *IEEE Instrum. Meas. Mag.* **2026**. [CrossRef]
33. Zhang, M.; Ma, C.; Chen, S.; Li, Y.; Huang, J.; Du, M.; Niu, H.; Ma, J. Prior-guided spectral reconstruction and spectral–spatial learning for non-destructive origin authentication of Pinellia ternata. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2026**, *364*, 128523. [CrossRef] [PubMed]

34. Zhang, M.; Ma, C.; Li, Y.; Huang, J.; Du, M.; Niu, H.; Ma, J. Source-aware reinforcement-learning optical-channel selection for compact VIS-NIR sensing with botanical validation. *Opt. Express* **2026**, *34*, 32557–32572. [CrossRef]
35. Feng, J.; Gao, Q.; Shang, R.; Cao, X.; Bai, G.; Zhang, X.; Jiao, L. Multi-agent deep reinforcement learning for hyperspectral band selection with hybrid teacher guide. *Knowl.-Based Syst.* **2024**, *299*, 112044. [CrossRef]
36. Guo, Y.; Wang, Q.; Hu, B.; Qian, X.; Ye, H. Two-Stage Unsupervised Hyperspectral Band Selection Based on Deep Reinforcement Learning. *Remote Sens.* **2025**, *17*, 586. [CrossRef]
37. Feng, W.; Wang, M.; Tang, C.; Xie, W.; Li, X.; Zheng, X.; Xu, J. Hyperspectral band selection via region-wise latent feature fusion and graph filter embedded subspace clustering. *Eng. Appl. Artif. Intell.* **2024**, *132*, 107911. [CrossRef]
38. Zhang, Y.; Qi, J.; Wang, X.; Cai, Z.; Peng, J.; Zhou, Y. Tensorial Global-Local Graph Self-Representation for Hyperspectral Band Selection. *IEEE Trans. Circuits Syst. Video Technol.* **2024**, *34*, 13213–13225. [CrossRef]
39. Han, B.; Liu, M.; Ma, Z.; Zhang, K.; Xu, Y.; Wang, J.; Wang, Q. Hyperspectral Band Selection with Unique Pixel Extraction and Adaptive Neighbor Clustering. *Remote Sens.* **2025**, *17*, 315. [CrossRef]
40. Yan, D.; Shang, X.; Wang, X.; Zhang, J.; Sun, X. Multi-View Multi-Level Robust Deep Subspace Clustering for Hyperspectral Band Selection. *IEEE Trans. Geosci. Remote Sens.* **2026**, *64*, 5506016. [CrossRef]
41. Kumar, B.L.N.P.; Vaddi, R.; Manoharan, P.; Agilandeeswari, L.; Sangeetha, V. A new band selection framework for hyperspectral remote sensing image classification. *Sci. Rep.* **2024**, *14*, 31836. [CrossRef] [PubMed]
42. Li, X.-K.; Ma, J.-X.; Li, X.-Y.; Hu, J.-J.; Ding, C.-Y.; Han, F.-K.; Guo, X.-M.; Tan, X.; Jin, X.-M. High-efficiency reinforcement learning with hybrid architecture photonic integrated circuit. *Nat. Commun.* **2024**, *15*, 1044. [CrossRef] [PubMed]
43. Li, Y.; Chen, S.; Gong, T.; Ozcan, A. Model-free optical processors using in situ reinforcement learning with proximal policy optimization. *Light. Sci. Appl.* **2026**, *15*, 32. [CrossRef] [PubMed]
44. Yan, Q.; Ouyang, H.; Tao, Z.; Shen, M.; Du, S.; Zhang, J.; Liu, H.; Hao, H.; Jiang, T. Multi-wavelength optical information processing with deep reinforcement learning. *Light Sci. Appl.* **2025**, *14*, 160. [CrossRef] [PubMed]
45. Kim, J.; Kim, J.Y.; Kim, J.; Hyeong, Y.; Neseli, B.; You, J.B.; Shim, J.; Shin, J.; Park, H.H.; Kurt, H. Inverse design of nanophotonic devices enabled by optimization algorithms and deep learning: Recent achievements and future prospects. *Nanophotonics* **2025**, *14*, 121–151. [CrossRef] [PubMed]
46. Rahman, M.S.S.; Ozcan, A. Physics and artificial intelligence: Illuminating the future of optics and photonics. *Adv. Photonics* **2024**, *6*, 050500. [CrossRef]

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.