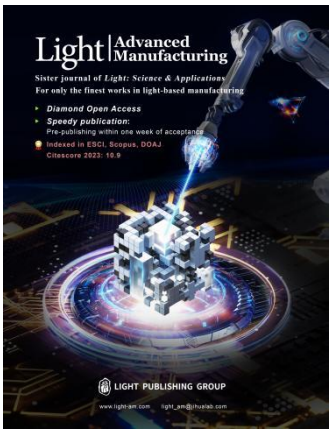
Accepted Article Preview: Published ahead of online publication



Analysis and Division of Full-Band Errors in Optical Surfaces

Yuan Liu, Xiaokun Wang, Zhongkai Liu, Yukun Wang, Mengxue Cai, Donglin Xue, Xuejun Zhang

Cite this article as: Yuan Liu, Xiaokun Wang, Zhongkai Liu, Yukun Wang, Mengxue Cai, Donglin Xue, Xuejun Zhang. Analysis and Division of Full-Band Errors in Optical Surfaces. *Light: Advanced Manufacturing* accepted article preview 4 June, 2026; doi: 10.37188/lam.2026.093

Received 24 March 2026; Revised 1 June 2026; Accepted 4 June 2026;
Accepted article preview online 4 June 2026

Analysis and Division of Full-Band Errors in Optical Surfaces

Running Title: Full-Band Error Analysis and Division

Author:

Yuan Liu[a,b,c], Xiaokun Wang[a,b,c,*], Zhongkai Liu[a,b,c,*], Yukun Wang[a,b,c], Mengxue Cai[a,b,c], Donglin Xue[a,b,c] and Xuejun Zhang[a,b,c]

[a] *Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, 3888 Dongnanhu Road, Changchun, Jilin 130033, China*

[b] *University of Chinese Academy of Sciences, Beijing 100049, China*

[c] *State Key Laboratory of Advanced Manufacturing for Optical Systems, 3888 Dongnanhu Road, Changchun, Jilin 130033, China*

[*] *jimwxk@sohu.com*

Abstract

Precisely separating full-band surface errors is a cornerstone of high-performance optical manufacturing and testing. However, traditional methods rely on subjective frequency band definitions and manual filtering, which leads to inconsistent evaluation standards and introduces analysis errors. Meanwhile, existing data-driven methods mechanically fit noisy manual labels, lacking physical consistency. To address this issue, this paper presents a scale-adaptive physics-aware deep learning framework powered by a dual-domain physics-aware network, which extracts error features by imposing physical constraints in the frequency domain: a specifically shaped spectral domain selective attention mechanism is designed for a low-frequency error with a deterministic spectral envelope; an adaptive learnable spectral gating is designed for mid-frequency error with randomness and anisotropy. Experiments indicate that this network outperforms conventional models under adopted evaluation metrics and exhibits physics-constrained extraction behaviour. This helps mitigate measurement noise and subjective variability introduced by manual filtering, yielding results with improved spectral and physical consistency. Furthermore, it eliminates the reliance on high-performance computing hardware, thereby enabling millisecond-level real-time inference on a standard computer, which greatly facilitates practical engineering deployment. Based on these findings, this study proposes a generalised frequency band division method based on normalised spatial frequency, providing a normalised-frequency-based method for a more consistent cross-scale and cross-instrument frequency band error division.

Keywords: Optical component, optical testing, frequency band error, deep learning

1. Introduction

Given the continuous advancement of cutting-edge optical fields such as space optics, extreme ultraviolet (EUV) lithography, and inertial confinement fusion[1–7], stringent requirements have been placed on the full-band errors of their optical components[8–10]. Among these, the surface error of optical components (hereafter referred to as a surface error) can be divided into low-, mid-, and high-frequency errors based on their frequency bands. Errors in different frequency bands have distinct and significant effects on the performance of optical systems[11–13], requiring different manufacturing processes for targeted reduction[14–16]. Therefore, the accurate division and separation of errors across these frequency bands is an essential requirement in advanced optical manufacturing and testing.

Various standards for frequency-band division have been proposed to facilitate the description and evaluation of surface errors. Initially, standards such as MIL-PRF-13830B[17], ISO 10110 series[18–20], and GB/T 1185-2006[21] focused on the visibility and dimensions of discrete surface defects rather than serving as standards for error division. Subsequently, institutions such as the U.S. National Ignition Facility (NIF)[22–24] and Zeiss (Oberkochen, Germany)[25] have proposed specific error-division standards tailored to their project requirements. However, these standards target optical components for specific tasks, wavelengths, and dimensions. The formulation of some standards relies heavily on empirical rules and lacks a unified and widely applicable standard for frequency-band error division. Consequently, the analysis of frequency-band errors currently relies on manual filtering, which requires multiple adjustments to determine the filtering range. This approach is inefficient, dependent on empirical rules, and prone to introducing operational errors, underscoring the urgent need for a high-precision, high-efficiency, automated, and regularised method for frequency-band error analysis.

In recent years, the field of optical metrology has undergone a methodological shift from traditional numerical computations to machine learning and data-driven paradigms[26–34]. To address efficiency limitations, related research has leveraged deep learning to significantly improve feature inference[26] and reconstruction speed[27], successfully realising the parameter extraction of non-rigid physical models[28], rapid mapping of high-dimensional optical responses[29], and decoupling analysis of geometric features[30]. Physics-informed approaches have emerged as a new frontier for overcoming the uninterpretability of pure data-driven models. These methods introduce physical constraints such as energy conservation[31] and combining simulations with actual measurement data[32–34], thereby enhancing the efficiency and accuracy of complex tasks, including 3D measurement[35–37], phase recovery[38], and fringe analysis[39–41]. Furthermore, Fourier neural networks and frequency-domain attention mechanisms have recently been integrated into these approaches, demonstrating a significant potential for related tasks such as gradient-based surface reconstruction[42] and surface defect detection[43,44]. However, existing data-driven paradigms have certain limitations:

(1) Traditional methods mostly simplify the problem into black-box image processing,

ignoring the physical distribution characteristics of multi-scale information in the frequency domain.

(2) Network training aims to approximate manual labels and can therefore inherit part of the subjective variability introduced by manual filtering if no additional physical constraints are imposed.

(3) Existing solutions frequently rely on high-performance hardware, which makes them difficult to apply in practical engineering sites with limited computational resources.

This study proposes a scale-adaptive physics-aware deep-learning framework driven by a dual-domain physics-aware network (DPN) to effectively mitigate the physical limitations of traditional black-box models and previous fixed-mask methods. Our main contributions are as follows:

1) For the extraction mechanism, the network integrates a deterministic spectral domain selective attention (SDSA) mechanism and a new adaptive learnable spectral gating (LSG) mechanism that targets the distinct spectral distribution characteristics of low-frequency errors and mid-frequency textures.

2) The frequency-domain physical priors help the model to mitigate measurement noise and subjective variability introduced by manual filtering, improving the spectral and physical consistency of the extracted results.

3) We propose a generalised frequency-band division method based on normalised spatial frequency. This framework is designed to achieve the high-precision extraction of full-band errors while targeting millisecond-level real-time inferences on standard computers. The proposed approach can facilitate real-time in situ frequency-band error analysis, providing a foundation for intelligent optical manufacturing and cross-scale cross-instrument optical surface quality evaluation.

2. Methods

2.1 Overall Architecture: Physics-Aware Dual-Domain Neural Network

The key to accurately extract full-band errors is capturing the physical features of the frequency band errors at different scales from the instrument measurement results. From a traditional spatial domain perspective, the low-frequency surface deviation (Z_{low}), mid-frequency manufacturing texture (Z_{mid}), high-frequency roughness (Z_{high}), and measurement noise ($N(x,y)$) share the same coordinate space (x, y). Consequently, the distinct frequency-band error features within the instrument testing result (Z) are highly superimposed.

$$Z(x,y) = Z_{\text{low}}(x,y) + Z_{\text{mid}}(x,y) + Z_{\text{high}}(x,y) + N(x,y) \tag{1}$$

The direct utilisation of neural networks to establish nonlinear mappings in the spatial domain is highly prone to overfitting and lacks physical interpretability because of the vast diversity in the surface morphologies of different optical components. However, according to the Fourier optics theory, superimposed frequency band errors in the spatial domain are naturally separated in the frequency domain[45], and their spectral amplitudes follow an inverse power-law decay model.

$$\begin{aligned} &\tilde{Z}\left(f_x, f_y\right) = \tilde{Z}_{\text{low}}\ f_x, f_y) + \tilde{Z}_{\text{mid}}\ f_x, f_y\ + \tilde{Z}_{\text{high}}\ f_x, f_y\ + \tilde{N}\ f_x, f_y \\ &\tilde{Z}\left(f_x, f_y\right) \propto C / \left(\left(\sqrt{f_x^2 + f_y^2}\right)^k + \text{ò}\right) \end{aligned} \tag{2}$$

where C, k, and ò represent a normalisation constant, decay rate constant related to the surface manufacturing process, and minimal value, respectively. This provides a novel perspective for frequency-band error extraction: each frequency-band error exhibits specific information distribution laws in the frequency domain. With an increase in the radial distance $\sqrt{f_x^2 + f_y^2}$, the corresponding spatial frequency also increases. Therefore, introducing spectral information distribution as a physical prior into network training is key to enhancing the physical consistency of the extracted results and reducing their sensitivity to local measurement noise and label-dependent fluctuations.

Fig. 1 illustrates the scale-adaptive physics-aware deep learning framework proposed in this study. The framework adheres to the design philosophy of 'frequency domain awareness and scale adaptation'. An SDSA mechanism is designed based on its isotropic and concentrated spectral information characteristics for the low-frequency error tested by interferometers, and an LSG is designed based on its random and anisotropic spectral information distribution characteristics for the mid-frequency error tested by optical profilers. Both are introduced into the dual-domain neural network as spectral physical information priors for establishing the mapping relationship between the instrument testing results and full-band error. A generalised division of the full-band error is by introducing a normalised spatial frequency to eliminate the size dependence.

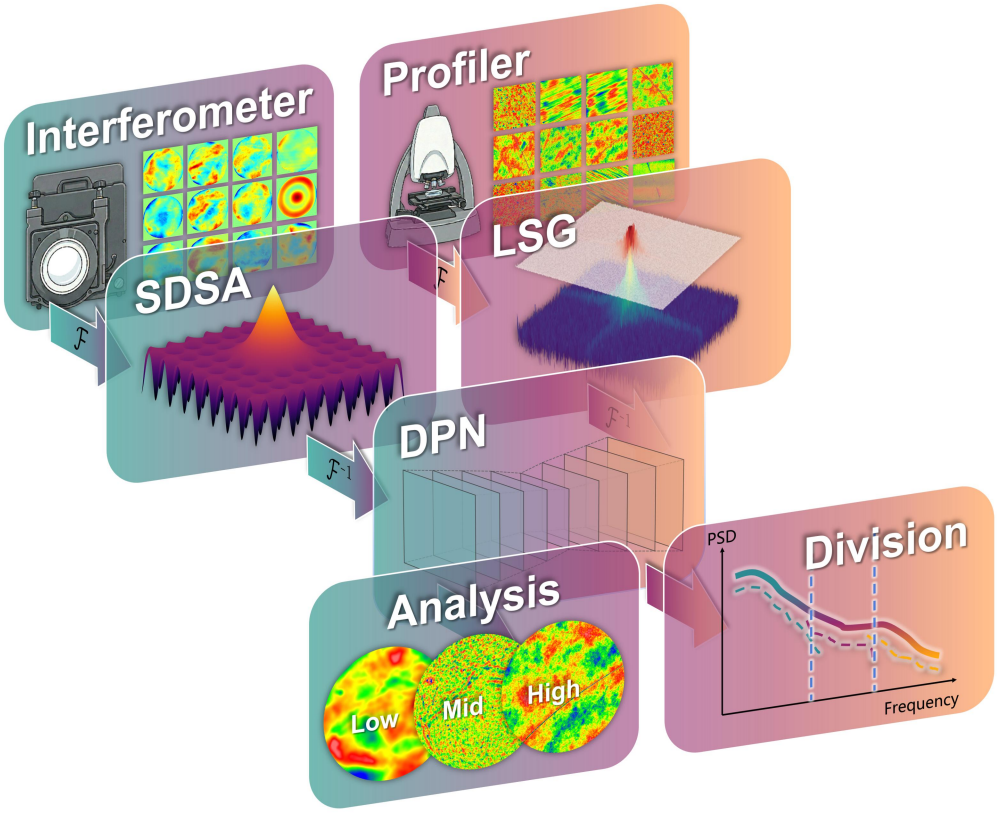


Fig. 1 Scale-adaptive physics-aware deep learning framework. The raw testing results from the interferometer and optical profiler are modulated in the Fourier spectral space by the SDSA mechanism and LSG for low- and mid-frequency errors, respectively. The DPN utilises dual-domain physical constraints to precisely extract the full-band (low-, mid-, and high-frequency) error information from the superimposed raw instrument data, enabling subsequent objective analysis and generalised frequency band error division based on normalised spatial frequency.

2.2 Macro-Scale: Low-Frequency Error Extraction Based on SDSA Mechanism

Low-frequency errors originate from gravitational deformation, alignment stress, or low-order aberrations[46], and their spatial domain distribution is continuous and smooth, without obvious directionality. Therefore, the information distribution of the low-frequency error in the frequency domain is fundamentally identical, manifesting as a central envelope with highly concentrated energy and almost no side lobes. Accordingly, we introduce the SDSA mechanism from a previous work[47] into the DPN, as shown in **Fig. 2**. The SDSA comprises a spectral attention reinforcer (SAR) and spectral rejection array (SRA).

$$SDSA = SAR + SRA$$

$$SAR = w_{\min} + \left(w_{\max} - w_{\min}\right) \cdot \sigma\left(-k_1\left(d_1 - R_{\text{inner}}\right)\right) \cdot \sigma\left(-k_2\left(d_1 - R_{\text{outer}}\right)\right) \tag{3}$$

$$SRA = \prod_{\left(f_{xp}, f_{yp}\right) \in P} \left(1 - k_3 \cdot \exp\left(-\frac{\sqrt{\left(f_x - f_{xp}\right)^2 + \left(f_y - f_{yp}\right)^2}}{2r^2}\right)\right)$$

A detailed theoretical derivation and parameter selection of the SDSA is presented in our previous work[47]. The specific network architecture design is presented in Supplementary Information S2.

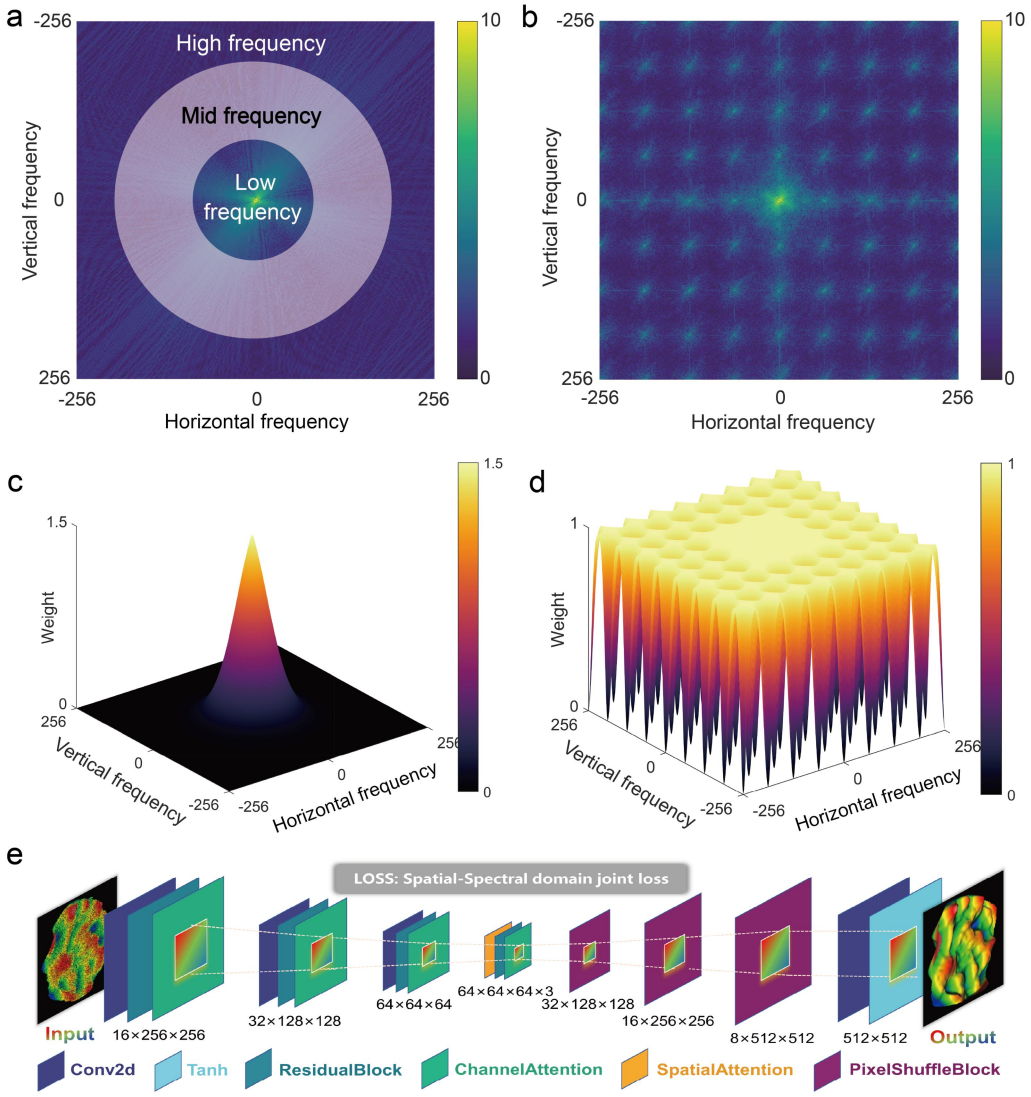


Fig. 2 Architecture of the SDSA and DPN for low-frequency error extraction. **a** Fourier spectral information distribution characteristics of interferometer testing results, where the frequency gradually increases with the radius. **b** Distribution of spectral 'star-like' noise introduced by traditional convolution during the training process. Based on the above spectra, **c** and **d** show the specific designs of the SAR and SRA for low-frequency error, respectively, where SAR is designed as an envelope

shape, and the innermost units of the SRA are disabled. **e** The DPN architecture for low-frequency error extraction.

2.3 Micro-Scale: Mid-Frequency Error Extraction Based on LSG

Unlike the deterministic spectral envelope morphology of low-frequency errors, mid-frequency errors are primarily introduced during the manufacturing process (e.g. tool marks and vibrations)[46]. The distribution direction and frequency of the mid-frequency errors introduced by the different manufacturing methods are inconsistent, and the spectral information distribution is random with significant anisotropy. Real manufacturing textures are mostly non-ideal sinusoidal morphologies and accompanied by additional vibrations, and therefore, they introduce continuous energy side lobes in the spectrum according to the principles of the Fourier series. Addressing this spectral characteristic, a fixed-morphology mask (such as SDSA) cannot match precisely and easily lead to the loss of valid side-lobe information. Therefore, the extraction of mid-frequency error requires physical constraints that can adaptively fit different spectral morphologies.

To address this challenge, we designed a physics-aware LSG. Unlike the fixed morphology of the SDSA, the LSG adaptively learns the spectral energy distribution of each mid-frequency error. The core mechanism is the introduction of a 3×3 frequency-domain convolution kernel to replace traditional point-by-point scanning, which endows the network with a local receptive field in the frequency domain. The LSG is deployed at the front end of the network as a lightweight preprocessing module. Assuming that the input optical profiler testing data is $Z(x,y)$, it is transformed into the frequency domain $\tilde{Z}_{in}\left(f_x,f_y\right)=F\left\{Z(x,y)\right\}$ via fast Fourier transform, and its real part $\mathrm{Re}(\tilde{Z}_{in})$ and imaginary parts $\mathrm{Im}(\tilde{Z}_{in})$ are extracted and concatenated into a two-channel spectral tensor $\mathbf{S}_{in}=[\mathrm{Re}(\tilde{Z}_{in}),\mathrm{Im}(\tilde{Z}_{in})]$. This tensor is fed into the LSG subnetwork that is composed of two 3 × 3 convolutional layers and is activated by a sigmoid function to generate a mask ranging from 0 to 1.

$$\mathrm{Mask}(f_x,f_y)=\sigma(\mathbf{W}_2*\mathrm{ReLU}(\mathrm{BN}(\mathbf{W}_1*\mathbf{S}_{in}))+\mathbf{b}_2) \tag{4}$$

where $*$, $\mathbf{b}_2$, $\mathrm{BN}(\cdot)$, $\mathrm{ReLU}(\cdot)$, and $\sigma(\cdot)$ represent a 2D convolution with a kernel size of 3×3, convolutional layer bias, batch normalisation, activation function, and sigmoid function, respectively. $\mathbf{W}_1$ and $\mathbf{W}_2$ represent the convolution kernel weights. This convolution kernel design is mathematically equivalent to performing weighted integration within a local neighbourhood of the spectrum, which enables the network to perceive the energy continuity of the surrounding side lobes while evaluating a specific frequency. The generated adaptive mask performs element-wise multiplication ($\odot$) with the original spectrum to obtain spectral modulation.

$$\tilde{Z}_{\text{out}}(f_x, f_y) = \text{Mask}(f_x, f_y) \odot \tilde{Z}_{\text{in}}(f_x, f_y) \tag{5}$$

Further, it returns to the spatial domain via inverse fast Fourier transform, outputting $Z_{\text{out}}(x, y) = \mathrm{F}^{-1}\{\tilde{Z}_{\text{out}}(f_x, f_y)\}$, which is then fed into the backbone network for subsequent extraction. Meanwhile, the DPN backbone adopts an encoder–decoder architecture without skip connections (**Fig. 3**) to align with the LSG concept, which forces the information to pass through the bottleneck layer for separating the high-frequency random noise and reconstructing the mid-frequency morphology. For specific network construction details and anti-artefact designs, see Supplementary Information S3.

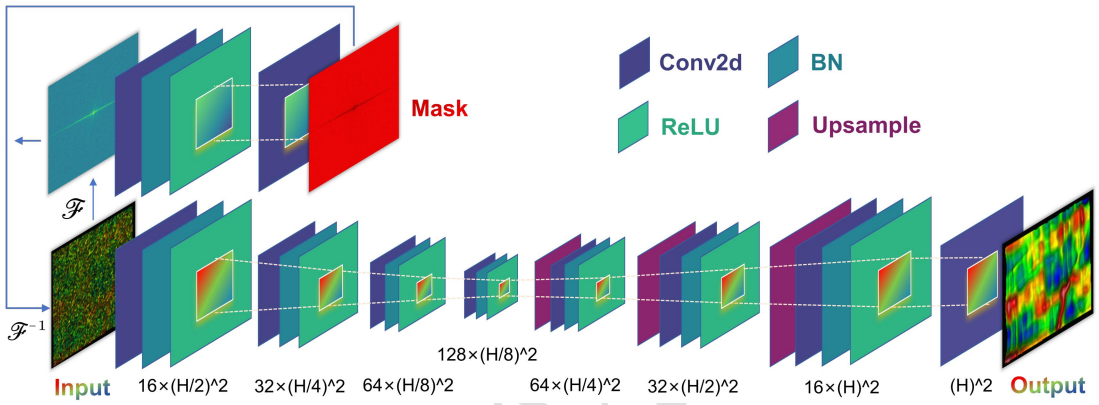


Fig. 3 Design of the LSG sub-network and DPN architecture for mid-frequency error extraction. The front end of the network maps the optical profiler data to the frequency domain through the LSG sub-network (blue line) and generates an adaptive mask to extract mid-frequency topological features. The backbone network adopts an encoder–decoder structure without skip connections, separates additional information through a bottleneck layer, and reconstructs the mid-frequency morphology combining bilinear interpolation for upsampling.

The network training process is driven by a set of multidimensional physical constraint loss functions

$$\mathrm{L}_{\text{total}} = \mathrm{L}_{\text{spatial}} + \mathrm{L}_{\text{freq}} + \lambda_1 \mathrm{L}_{\text{grad}} + \lambda_2 \mathrm{L}_{\text{mask}} \tag{6}$$

where λ_1 and λ_2 represent the hyperparameters used to balance the weights of each loss term. Further, $\mathrm{L}_{\text{spatial}}$ represents the spatial domain mean square error loss, which ensures that the network rapidly learns the overall morphology in the spatial domain, and L_{freq} represents the frequency domain amplitude mean square error loss. The spectral energy of optical surfaces spans multiple orders of magnitude, and therefore, direct calculations can cause information in the relatively high-frequency components to be submerged. Therefore, this framework introduces a nonlinear square-root compression to the spectral mean-square error loss.

$$L_{freq} = \left\| \sqrt{|F\{Z_{out}\}| + ò} - \sqrt{|F\{Z_{gt}\}| + ò} \right\|_2^2 \tag{7}$$

where $|F\{\cdot\}|$, Z_{gt}, and $ò$ represent the amplitude of the Fourier spectrum, manual label value, and minimal value, respectively. This frequency-domain physical constraint ensures the consistency of the spectral information distribution in the frequency domain. Moreover, mid-frequency textures typically exhibit strong directionality and local gradient mutations. Therefore, the gradient loss L_{grad} is introduced to enhance the continuous directional characteristics of the mid-frequency error texture and texture edge details.

$$L_{grad} = \| \nabla_x Z_{out} - \nabla_x Z_{gt} \|_2^2 + \| \nabla_y Z_{out} - \nabla_y Z_{gt} \|_2^2 \tag{8}$$

where ∇ represents extraction operators for edge gradients in different directions, implemented in the network using a 3 × 3 discrete difference convolution kernel, and $\|\cdot\|_2^2$ represents the L2 norm. Finally, a regularisation term L_{mask} based on spectral distance weighting is designed specifically for the LSG mechanism to further suppress high-frequency redundant noise.

$$L_{mask} = ã\left[\mathrm{Mask}(f_x, f_y) \cdot \frac{\sqrt{f_x^2 + f_y^2}}{\max(\sqrt{f_x^2 + f_y^2})} \right] \tag{9}$$

where $ã$ represents the expectation, which calculates the average penalty strength of the mask weights across the entire spectral space. Compared to direct summation, it yields more stable results when dealing with the data of different resolutions, eliminating the need to frequently change hyperparameters. Redundant information and noise in high-frequency regions can be automatically suppressed while preserving valid information in the central area of the spectrum by imposing a weighted penalty based on the radial distance $\sqrt{f_x^2 + f_y^2}$. $\lambda_1 = 0.1$ and $\lambda_2 = 1 \times 10^{-5}$ are the weights. For specific network architecture design, discussions on LSG weight selection, parameter selection experiments, and ablation experiments are provided in Supplementary Information S3.

3. Results

3.1 Interferometer - Low-Frequency Error Extraction Results

We constructed two diverse experimental datasets to comprehensively evaluate the full-band error extraction performance of the DPN framework: an interferometer dataset comprising 2472 samples and an optical profiler dataset comprising 1853 samples, covering various instrument models, optical apertures, and manufacturing processes. A batch-level isolation strategy was applied during dataset splitting to

prevent data leakage (see Supplementary information S4–S6 for details). All baseline models and the proposed DPN were trained and evaluated using the same data split, preprocessing pipeline, manually filtered reference labels, and evaluation metrics to ensure a fair comparison. The hyperparameters were selected according to the validation set performance without providing any model with additional data or label information.

Figs. 4a–c show the DPN performances when processing samples contain typical measurement noise and anomalous bad pixels (spikes). The DPN network not only robustly eliminates anomalous spikes (**Fig. 4c**) but also extracts a smooth low-frequency error morphology enabled by the spectral physical constraints introduced by the SDSA and sub-pixel upsampling strategy (**Fig. 4b**). **Figs. 4d–h** and **Table 1** present the quantitative testing results of the DPN on a test set of 371 samples. In terms of precision, the DPN achieves a root-mean-square error (RMSE) of 0.84 nm, which represents a 2.67-fold improvement in accuracy compared to that of the baseline CNN. The relative accuracy (Rel Acc) is introduced to clearly evaluate the similarity between the extracted error and manual labels, defined as

$$\text{Rel Acc} = \left(1 - \frac{\text{RMSE}}{\text{RMS}_{\text{gt}}}\right) \times 100\% \tag{10}$$

where RMS_{gt} represents the root mean square of the manual label. Across key metrics such as relative accuracy (85.91%), structural similarity (SSIM, 0.887), and peak signal-to-noise ratio (PSNR, 34.59 dB), the DPN consistently outperforms the comprehensive baseline set, which includes the classic spatial-domain (U-Net, Transformer, PINNs) and advanced frequency-domain architectures (physics-informed FNO (PI-FNO) and U-Net-based adaptive FNO (U-AFNO)) as demonstrated in the updated **Table 1**. In terms of computational efficiency, as shown in **Fig. 4d**, the DPN requires an average of only 0.04 s to process a single surface error map on a standard laptop without GPU acceleration (see Supplementary Information S2 for details), which is an improvement of nearly 750 times compared to that of traditional manual filtering (~30 s/surface).

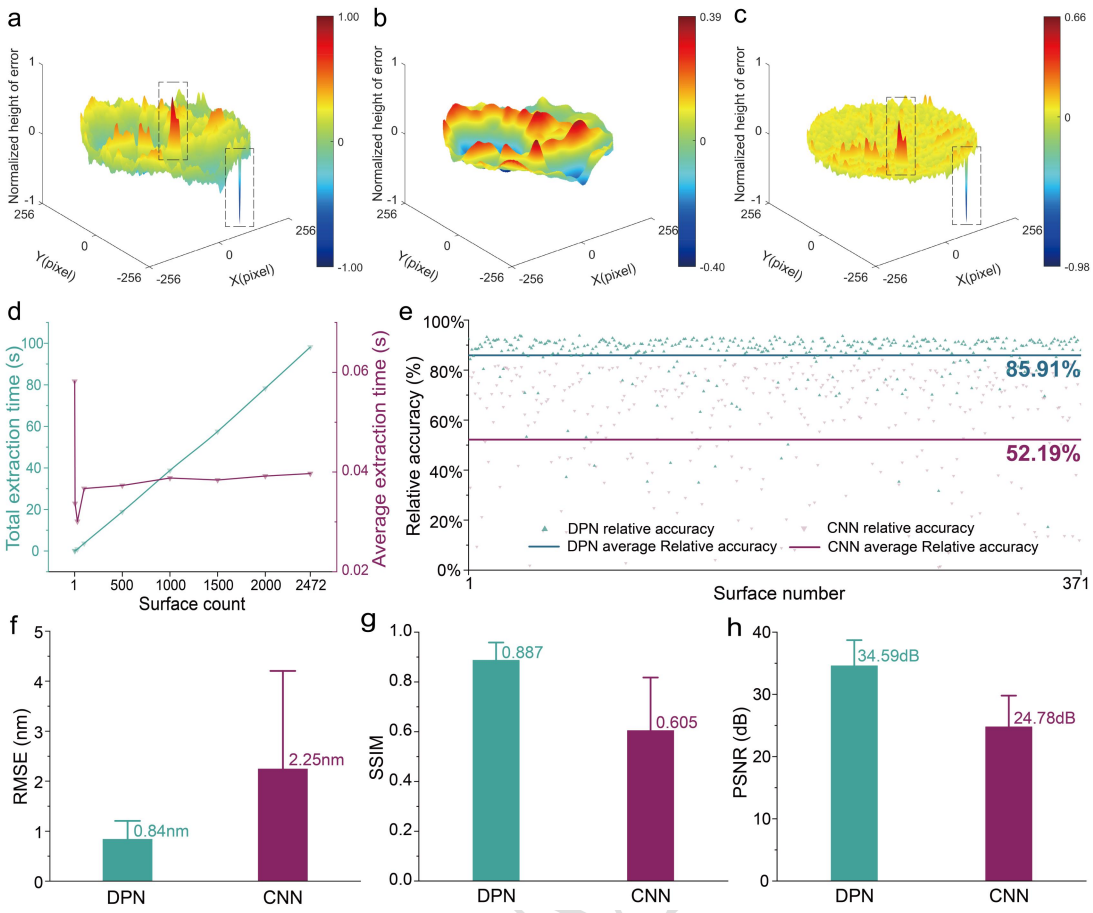


Fig. 4 Performance demonstration of the DPN for low-frequency error extraction. **a–c** Demonstration of outlier elimination: **a** Raw interferometer input containing anomalous spikes. **b** Smooth low-frequency error extracted by the DPN. **c** Residual map between the two, which shows that two anomalous spikes are effectively removed by the DPN. **d** Total extraction time and average extraction time per surface for different inputs by the DPN. **e–h** Quantitative comparison between the DPN and baseline CNN on four core evaluation metrics: relative accuracy, RMSE, SSIM, and PSNR.

Table 1. Performance comparison of low-frequency error extraction with classic spatial-domain and Fourier-domain architectures.

Model	RMSE (nm)	Rel Acc (%)	SSIM	PSNR (dB)	Average extraction time (s)
U-Net	1.85	82.24	0.869	33.91	0.654
PINNs	1.14	84.05	0.858	33.84	0.122
Transformer	1.02	84.83	0.876	34.15	0.051
PI-FNO	1.02	78.77	0.758	31.31	0.097
U-AFNO	1.15	67.29	0.770	27.72	0.073
DPN	0.84	85.91	0.887	34.59	0.040

3.2 Optical Profiler - Mid-Frequency Error Extraction Results

Fig. 5 shows the adaptive perception mechanism of the LSG for different spectral morphologies. For surfaces with different roughness values (**Fig. 5a**), the spectral energy distributions exhibit distinctly different morphologies (**Fig. 5b**). The LSG effectively generates corresponding masks dynamically for different spectra (**Fig. 5c**), precisely identifying the valid information regions. The effect of the convolution kernel scale on topological perception, the parameter selection experiments in Supplementary information S3 indicate that the LSG effectively addresses the limitation of traditional 1 × 1 point-by-point filtering owing to the local receptive field introduced by the 3×3 convolution kernel, which tends to truncate weak side lobes. It can dynamically generate continuous masks that cover the complete key information regions of the different spectra. This mechanism ensures the complete recognition of the network of mid-frequency features while continuously suppressing redundant noise in high-frequency regions.

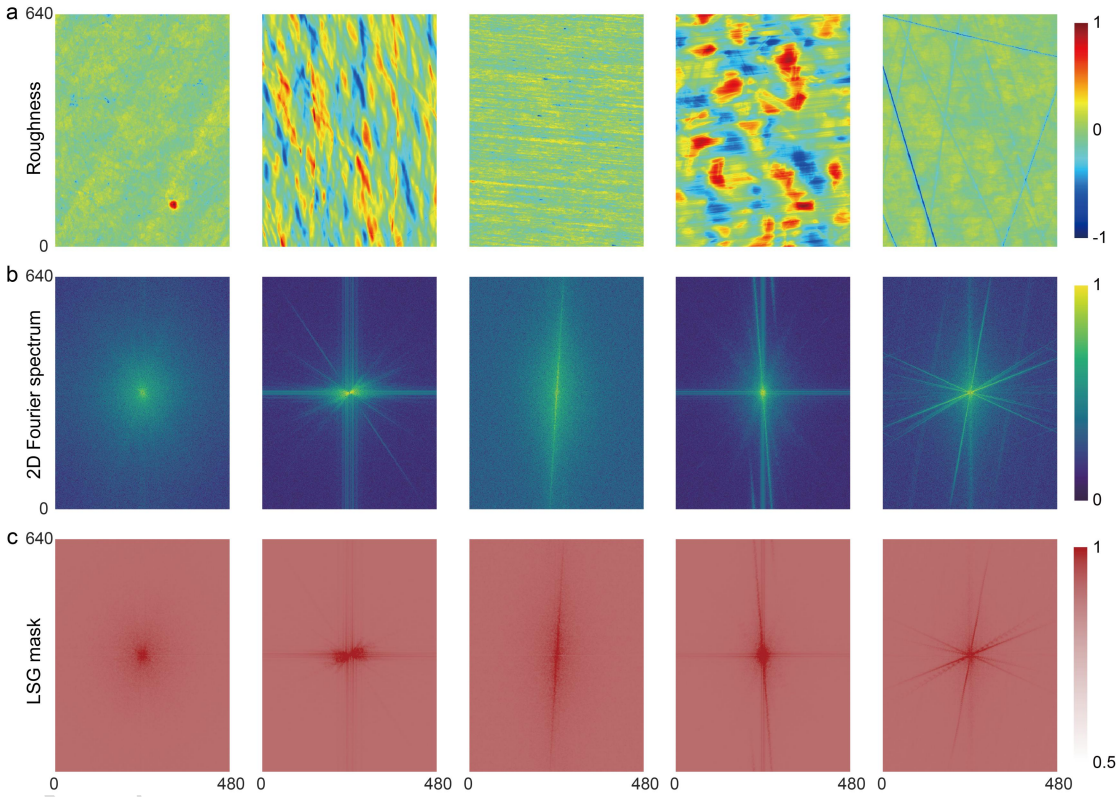


Fig. 5 Adaptive perception effect of the LSG on different mid-frequency morphologies. **a** Raw optical profiler testing results of five groups with typical manufacturing textures. **b** Corresponding five 2D Fourier spectra, which present anisotropic energy distribution characteristics. **C** Frequency domain mask generated by the LSG based on spectral features, achieving the precise coverage of the key information distribution regions.

Based on 279 test samples, **Fig. 6** and **Table 2** present the qualitative and quantitative extraction performance of the DPN for the mid-frequency error. The

high-frequency roughness can be separated simultaneously and in real time from the raw data using a simple subtraction operation because the optical profiler testing data are essentially a superposition of mid- and high-frequency errors enabled by the high-precision decoupling of the DPN of the mid-frequency error(**Figs. 6d and 6e**). In terms of accuracy metrics, the DPN achieved an RMSE of 0.37 nm and a relative accuracy of 86.29%, which consistently outperforms the comprehensive baseline set, including classic spatial-domain architectures (U-Net, PINNs, and Transformer) and advanced Fourier-domain architectures (PI-FNO and U-AFNO) across key metrics such as SSIM (0.905) and PSNR (36.02 dB), as demonstrated in the updated **Table 2**. In terms of computational efficiency, the DPN requires an average of only 0.05 s to process a surface error map, which is an improvement of nearly 600 times compared to that of traditional manual filtering (~30 s/surface).

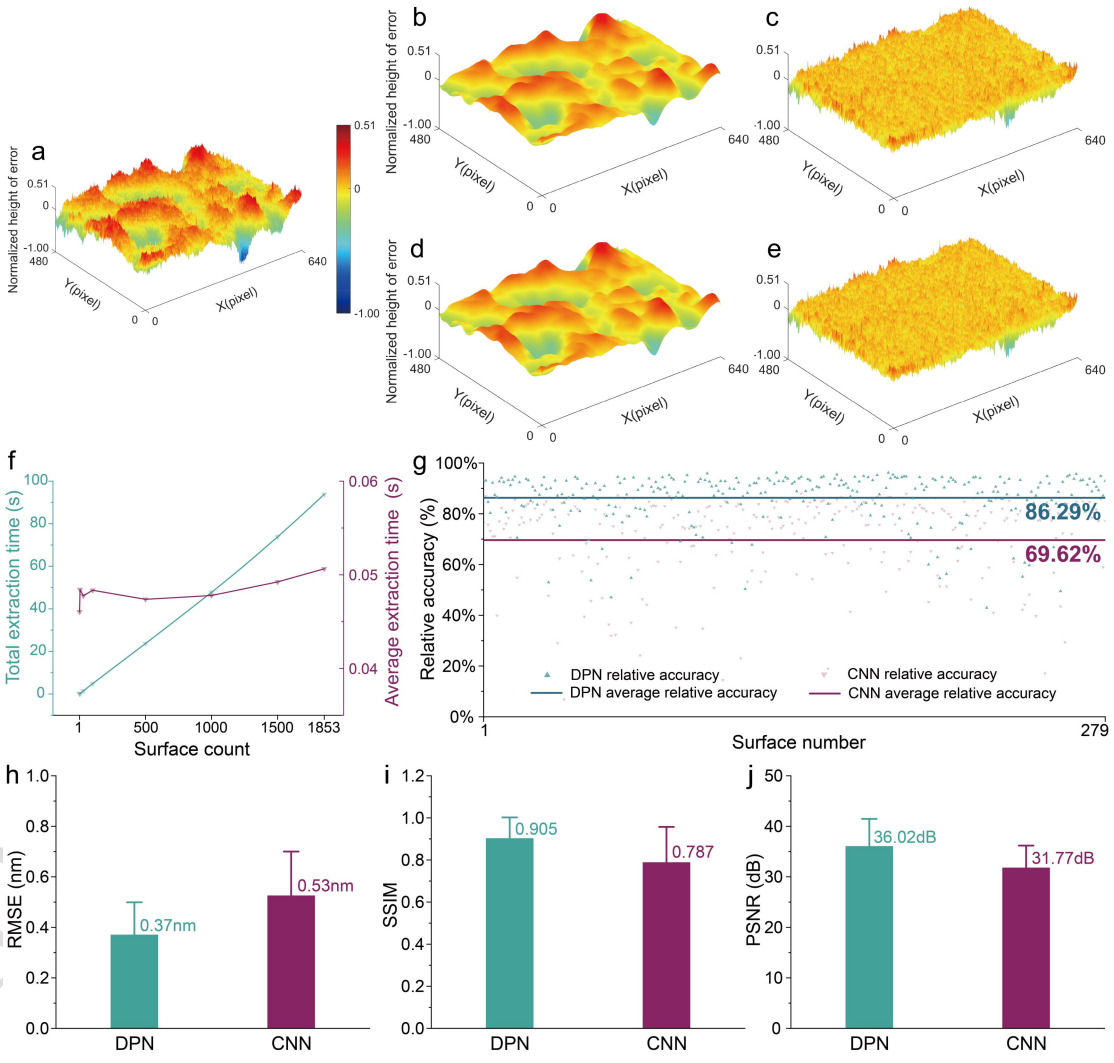


Fig. 6 Performance demonstration of the DPN for mid-frequency error extraction. **a** Raw optical profiler testing input. **b** and **c** Manually filtered references for the mid- and high-frequency errors. **d** and **e** Extraction results of the DPN for the mid- and high-frequency errors, respectively. **f** Total extraction time and average extraction

time of the network for different inputs. **g–j** Comparison between the DPN and baseline CNN on relative accuracy, RMSE, SSIM, and PSNR.

Table 2. Performance comparison of mid-frequency error extraction with classic spatial-domain and Fourier-domain architectures.

Model	RMSE (nm)	Rel Acc (%)	SSIM	PSNR (dB)	Average extraction time (s)
U-Net	0.39	84.58	0.886	35.65	0.692
PINNs	0.61	80.29	0.846	32.86	0.186
Transformer	0.69	75.87	0.756	30.40	0.073
PI-FNO	0.42	81.99	0.875	32.71	0.119
U-AFNO	0.71	84.72	0.899	35.75	0.086
DPN	0.37	86.29	0.905	36.02	0.051

3.3 Physics-correction effect of neural networks

The performance of traditional data-driven methods are constrained by the subjective noise of manual labels, which inherently risks propagating subtle biases in the outputs of the model. However, the DPN mitigates this problem by learning physical commonalities from large-scale datasets to statistically smooth out individual randomness. In the aforementioned quantitative evaluation, the DPN achieved high accuracy in the RMSE; however, the SSIM and relative accuracy did not yield perfect matches. This phenomenon reveals the distinct 'physics-correction effect' of the DPN. This effect refers to an observed output behaviour rather than an independent label-free correction module: under the basic supervision of manually filtered references, frequency-domain physical priors, dual-domain constraints, network architecture, and training regularisation jointly reduce the mechanical fitting of the model to local label anomalies and subjective filtering variability, producing results with improved spectral and physical consistency. The contributions of the main physical constraints and network settings are supported by the ablation study presented in Supplementary information S3.2.

In the extraction of low-frequency error, the 'physics-correction effect' manifests as the suppression of outliers. As shown in **Figs. 7a–c**, the DPN does not mechanically fit the anomalous spike noise introduced during the instrument testing process; however, it outputs a smooth and continuous low-frequency error. The quantitative power spectral density (PSD) curves (**Figs. 7g and 7h**) more clearly support this effect. The extraction results of the DPN show better consistency with the expected low-frequency spectral distribution across the full band, which suggests that the network can suppress local measurement defects in the manual reference rather than simply replicating them. In the extraction of a mid-frequency error, the 'physics-correction effect' is even more pronounced. A conservative filtering strategy is usually adopted because manual filtering struggles to precisely determine the cutoff frequency; This leads to the loss of a large amount of small-scale tool mark

information (**Figs. 7d and 7e**). Conversely, the DPN is less affected by the conservatism of manual labels and preserves more mid-frequency texture components under spectral constraints (**Fig. 7f**). The PSD curve in **Fig. 7i** quantitatively shows that the DPN maintains stronger spectral components than the conservative manual reference within the critical frequency band of 10^1–10^2 mm^{-1}. The DPN is not a simple label-fitting model. By leveraging the statistical advantages of large-scale data combined with dual-domain physical constraints, this approach mitigates the subjective randomness of manual operations and provides a method for full-band error feature extraction in practical engineering applications.

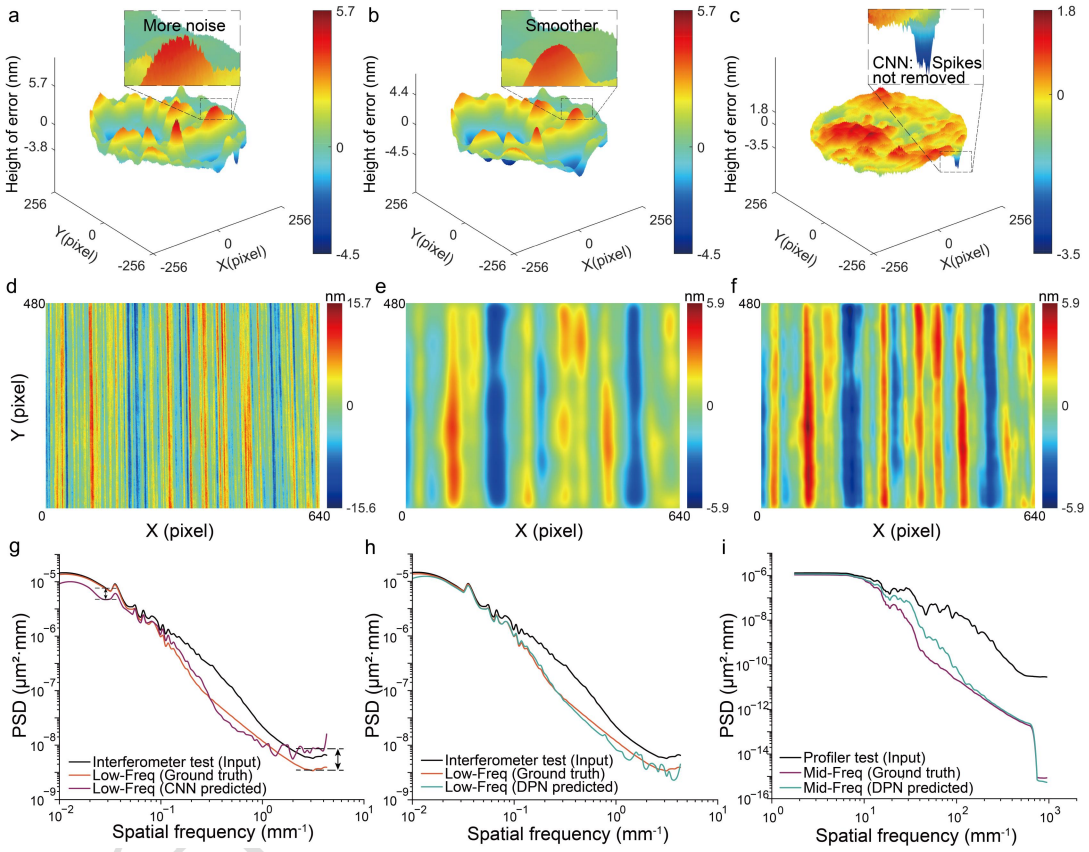


Fig. 7 'Physics-correction effect' of the neural network and its quantitative PSD analysis. **a–c** Comparison of low-frequency error: **a** CNN extraction result containing anomalous spikes, **b** Smoother and continuous DPN extraction result. **c** Residual map between the two, showing that the DPN can suppress anomalous spikes and high-frequency noise. **d–f** Comparison of mid-frequency error: **d** Raw input containing dense periodic fluctuations. **e** Manual reference loses details caused by conservative filtering. **f** DPN extraction result preserving more small-scale textures. **g–i** Corresponding quantitative PSD analysis, which shows that the DPN retains more spectral features in the low-frequency band (**g, h**, spatial frequency < 10^{-1} mm^{-1}) and the critical mid-frequency band (**i**, 10^1 < spatial frequency < 10^2 mm^{-1}), while avoiding the introduction of additional high-frequency noise.

4. Generalised Division of Full-Band Errors

The division of frequency-band errors is a bridge connecting optical manufacturing and testing. Existing division standards (such as ISO 10110 and MIL-PRF) are often based on fixed spatial frequency thresholds (e.g., 1 mm^{-1}), which results in the same physical frequency having entirely different physical meanings for components of different sizes or under different testing fields of view, thereby lacking generalisability. To address this issue and based on the objective features of frequency band errors learned by the DPN on massive data, we propose a generalised division method based on the normalised spatial frequency. The normalised spatial frequency is defined as

$$f_{\mathrm{norm}} = f \cdot L \tag{11}$$

where L represents the instrument testing aperture of the current data. In contrast to the full aperture D of the optical component, selecting L holds stricter physical significance: mid- and high-frequency textures are determined by the manufacturing process and have no direct correlation with the full aperture D of the optical component. Normalisation based on L not only avoids the distortion of evaluation standards caused by excessively large spans in component sizes but also better aligns with human visual perception, realising a dimensionless cross-scale and cross-instrument evaluation with the fluctuation period of the current field of view as the dimension.

When determining the specific division boundaries, we calculated the average PSD curves of the raw testing data and the DPN extraction results, and we introduced a ‘variance index’ to quantify the degree of separation between the two to objectively define the critical point. As shown in **Fig. 8**, the frequency-band breakpoints for interferometer or optical profiler data converge at a normalised spatial frequency of 8. To balance the signal-to-noise ratio, the breakpoint of 7.2 from the previous work[47] was optimised under the current data, which indicates that, within any testing field of view, a fluctuation period greater than 1/8 of the field of view is classified as low frequency, and less than 1/8 is classified as high frequency. This value provides an engineering division point for the datasets and testing conditions considered in this study, effectively reducing the subjective randomness induced by manual empirical rules. The normalised spatial frequency framework provides a general method by which users can determine suitable frequency band boundaries according to their instruments, component sizes, and engineering requirements. This breakpoint should be understood within the scope of the present engineering datasets rather than as a mandatory cross-institutional standard because independent multi-institutional datasets are yet to be included.

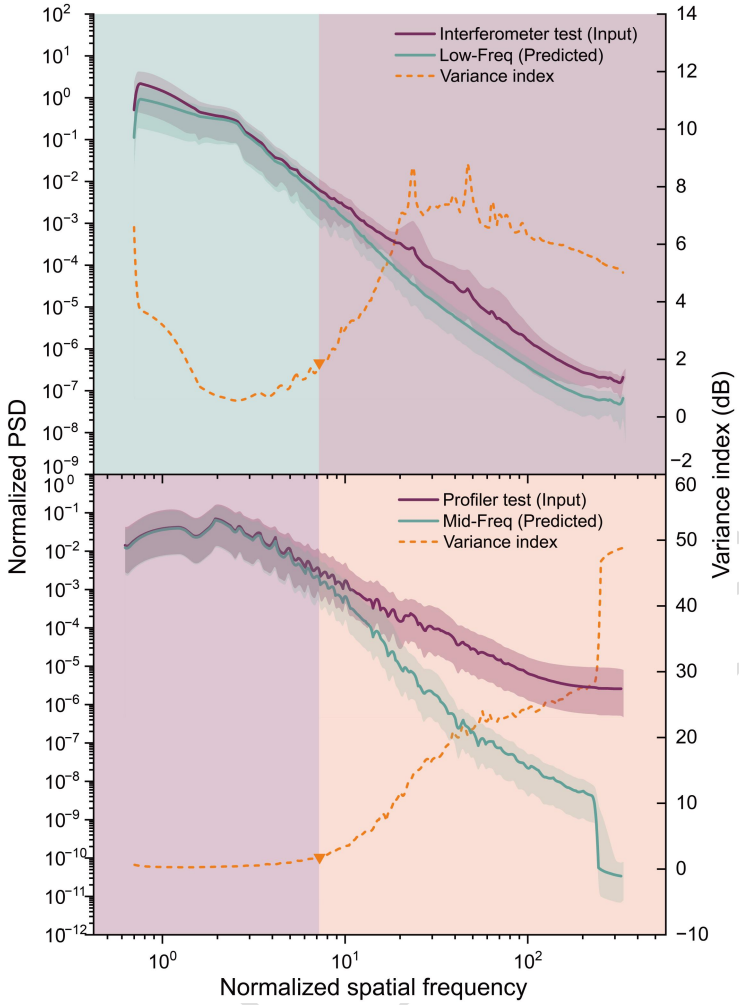


Fig. 8 Generalised division of full-band errors based on normalised spatial frequency. The upper and lower panels show the average PSD curves of the testing data and their DPN extraction results for the interferometer (low frequency) and optical profiler (mid frequency), respectively. The light shaded bands represent the standard deviation of the datasets. The variance index quantifies the degree of separation between the extraction results and raw instrument testing data. The variance index exhibits an upward trend in the frequency band error transition zone because of the coupling between frequency band errors. For the boundaries between low- and mid-frequency errors, as well as between mid- and high-frequency errors, the breakpoints are both at a normalised spatial frequency of 8.

5. Conclusion

This study proposed a physics-aware dual-domain deep learning framework driven by a DPN to effectively address the bottlenecks of low efficiency, strong subjectivity, and inconsistent frequency-band division criteria in cross-scale optical surface testing. The DPN addresses the mechanical fitting of manual labels inherent in traditional black-box models by integrating the spectral features as explicit physical

priors into the network design (SDSA and LSG). On a standard computer, this framework achieves high-precision extraction of full-band errors with a millisecond-level (0.05 s) inference efficiency, outperforming existing data-driven methods and establishing the advantage of physical priors in solving complex optical metrology problems.

The contribution of this study lies in demonstrating a physics-constrained extraction framework for full-band surface errors, and based on this, it is to refine the evaluation strategy. Relying on dual-domain constraints and statistical laws, the DPN mitigates measurement noise and subjective variability inherent in manual filtering, producing extracted results with improved spectral and physical consistency. Based on this, we established a generalised frequency band division method based on normalised spatial frequency. This method not only possesses crossband transfer learning capabilities but also addresses the challenges of size dependence in multi-instrument collaborative testing by providing a normalised-frequency-based workflow for frequency-band error division. Moreover, the applicability of the DPN is not inherently limited to specific aperture shapes or base geometries as a generalised physical-feature-aware architecture. Further, the framework demonstrates a promising potential for generalisation to complex surfaces. However, the DPN may deviate from the optimal reconstruction owing to overlapping physical features or sampling distortion under extreme conditions such as highly aliased complex periodic textures (see Supplementary Information S10). Owing to the current scarcity of open-source optical surface datasets, future work will focus on cross-validating the framework using multi-institutional data. The diagnostic scope of the DPN framework can be extended in future studies to include the quantification of localised surface defects and scatterers, providing a more comprehensive toolset for automated surface inspection.

This paradigm provides a practical engineering solution by translating data-driven metrics into downstream optical-performance gains. For the wavefront quality, the high-fidelity low-frequency errors extracted by the DPN provided a reliable baseline for deterministic polishing, effectively removing measurement noise crosstalk (see Supplementary Information S9). Moreover, through the precise decoupling of roughness data, the DPN simultaneously isolates high-frequency errors to evaluate scatter reduction for high-end systems, such as EUV optics, while preserving mid-frequency tool marks to trace the machining status and improve manufacturability. Finally, achieving millisecond-level inferences with low hardware reliance yields process-control benefits. This propels existing optical testing from ‘offline empirical analysis’ to ‘real-time in situ closed-loop feedback’, laying a solid foundation for automated optical manufacturing.

Acknowledgements. This work was supported by the National Major Research Instrument Development Project (62127901), Independent Project Task Document of the Key Laboratory of Advanced Manufacturing of Optical Systems (Chinese Academy of Sciences) (KLMSZZ202305), and Research on the Key Technology of

Rapid Synchronous Detection of Surface Shape and Subsurface Defects in the Grinding Stage of Large-Diameter complex surface (2025010157).

Author contributions. Yuan Liu developed the theoretical model, designed and performed the experiments, analysed the data, and drafted the original manuscript. Xiaokun Wang conceived the overall research idea, supervised the project, and acquired the primary financial support. Zhongkai Liu contributed to conceptualisation and assisted with the experimental setup. Yukun Wang provided valuable suggestions and assisted in reviewing and editing the manuscript. Mengxue Cai contributed to data collection and provided essential experimental datasets. Xuejun Zhang secured financial support and provided critical resources for this project. All authors discussed the results and contributed to the final manuscript.

Data availability. The complete code, pretrained weights, and representative desensitised subset of the surface error data used in this study are available in the Zenodo repository (DOI: 10.5281/zenodo.20117390). The full historical dataset remains restricted because of confidentiality constraints associated with proprietary laboratory projects.

Conflict of interests. The authors declare no conflicts of interest.

Supplementary information. See Supplementary information for supporting content.

References

[1] Dalcanton, J. J. 18 years of science with the Hubble Space Telescope. *Nature* **457**, 41-50 (2009). https://doi.org/10.1038/nature07621

[2] Betti, R. & Hurricane, O. A. Inertial-confinement fusion with lasers. *Nat. Phys.* **12**, 435-448 (2016). https://doi.org/10.1038/nphys3736

[3] Li, T. et al. Design of a high transmission illumination optics for anamorphic EUV lithography optics using deep reinforcement learning. *Opt. Express* **33**, 2261 (2025). https://doi.org/10.1364/OE.547606

[4] Crooke, J. A. et al. Status and path forward for the large ultraviolet/optical/infrared surveyor (LUVOIR) mission concept study. *Proc. SPIE* **9904**, 99044R (2016). https://doi.org/10.1117/12.2233084

[5] Nelson, J. & Sanders, G. H. The status of the Thirty Meter Telescope project. *Proc. SPIE* **7012**, 70121A (2008). https://doi.org/10.1117/12.788238

[6] Ferrante, C. et al. Epsilon-near-zero nonlinearity enhancement in the extreme ultraviolet. *Light: Sci. Appl.* **14**, 374 (2025). https://doi.org/10.1038/s41377-025-01985-w

[7] Li, W. et al. Controlling the wavefront aberration of a large-aperture and high-precision holographic diffraction grating. *Light: Sci. Appl.* **14**, 112 (2025). https://doi.org/10.1038/s41377-025-01785-2

[8] Rolland, J. P. et al. Freeform optics for imaging. *Optica* **8**, 161 (2021).

https://doi.org/10.1364/OPTICA.413762

[9] Deng, Y. et al. Local tolerance and quality evaluation for optical surfaces. *Optica* **9**, 1039 (2022). https://doi.org/10.1364/OPTICA.459633

[10] Fuerschbach, K., Thompson, K. P. & Rolland, J. P. Interferometric measurement of a concave, φ-polynomial, Zernike mirror. *Opt. Lett*. **39**, 18 (2014). https://doi.org/10.1364/OL.39.000018

[11] Jahn, W., Ferrari, M. & Hugot, E. Innovative focal plane design for large space telescope using freeform mirrors. *Optica* **4**, 1188 (2017). https://doi.org/10.1364/OPTICA.4.001188

[12] Fang, F. Z. et al. Manufacturing and measurement of freeform optics. *CIRP Ann*. **62**, 823-846 (2013). https://doi.org/10.1016/j.cirp.2013.05.003

[13] Zhang, B., Jin, G. & Zhu, J. Towards automatic freeform optics design: coarse and fine search of the three-mirror solution space. *Light: Sci. Appl.* **10**, 65 (2021). https://doi.org/10.1038/s41377-021-00510-z

[14] Li, J. et al. Study on traceability and suppression method of medium-frequency error for ultra-precision machining optical crystals. *Opt. Express* **29**, 22252 (2021). https://doi.org/10.1364/OE.432500

[15] Trumper, I. et al. Optics technology for large-aperture space telescopes: from fabrication to final acceptance tests. *Adv. Opt. Photonics* **10**, 644 (2018). https://doi.org/10.1364/AOP.10.000644

[16] Wang, B. *et al.* Novel suppression strategy of mid-spatial-frequency errors in sub-aperture polishing: adaptive spacing-swing controllable spiral magnetorheological finishing (CSMRF) method. *Light: Advanced Manufacturing* **6**, 45 (2025). https://doi.org/10.37188/lam.2025.045

[17] Department of Defense. Optical components for fire control instruments. MIL-PRF-13830B (2000).

[18] International Organization for Standardization. Optics and photonics — Preparation of drawings for optical elements and systems — Part 5: Surface form tolerances. ISO 10110-5 (2015).

[19] International Organization for Standardization, Optics and photonics — Preparation of drawings for optical elements and systems — Part 7: Surface imperfection tolerances, ISO 10110-7 (2017).

[20] International Organization for Standardization, Optics and photonics — Preparation of drawings for optical elements and systems — Part 14: Wavefront deformation tolerance, ISO 10110-14 (2018).

[21] Standardization Administration of China (SAC), Optical parts surface imperfections, GB/T 1185-2006 (2006).

[22] Aikens, D. The origin and evolution of the optics specifications for the National Ignition Facility. Lawrence Livermore National Lab. (LLNL) (1995). https://doi.org/10.2172/104971

[23] Campbell, J. H. et al. NIF optical materials and fabrication technologies: an overview. *Proc. SPIE* **5341**, 84 (2004). https://doi.org/10.1117/12.538471

[24] English Jr, R. E. et al. Implementation of ISO 10110 Optics Drawing. SPIE 40th Annual Meeting (1995). https://doi.org/10.2172/102436

[25] Kurz, P. et al. Optics for EUV lithography. *Digest of Papers Microprocesses and Nanotechnology* **2000**, 264 (2000). https://doi.org/10.1109/IMNC.2000.872750

[26] Sabbagh, R., Stothert, A. & Djurdjanovic, D. Machine learning for rapid inference of critical dimensions in optical metrology of nanopatterned surfaces. *CIRP J. Manuf. Sci. Technol*. **47**, 184-192 (2023). https://doi.org/10.1016/j.cirpj.2023.10.005

[27] Fu, L. et al. Comparative analysis of grating reconstruction: Deep learning versus Levenberg-Marquardt methods. *Proc. SPIE* **12619**, 50-56 (2023). https://doi.org/10.1117/12.2688328

[28] Liu, S. et al. Inverse optical scatterometry using sketch-guided deep learning. *Opt. Express* **32**, 20303-20315 (2024). https://doi.org/10.1364/OE.524091

[29] Mudide, S. et al. Machine learning driven measurement of high-aspect-ratio nanostructures using Mueller matrix spectroscopic ellipsometry. *J. Vac. Sci. Technol. B* **43**, 012801 (2025). https://doi.org/10.1116/6.0004058

[30] Jung, J. et al. Neural network-based analysis algorithm on Mueller matrix data of spectroscopic ellipsometry for the structure evaluation of nanogratings with various optical constants. *Nanophotonics* **14**, 471-484 (2025). https://doi.org/10.1515/nanoph-2024-0565

[31] Yang, S. et al. Channeled spectroscopic ellipsometry enabled by physics-informed tandem untrained neural networks. *Measurement* **235**, 114940 (2024). https://doi.org/10.1016/j.measurement.2024.114940

[32] Wörnhör, A. et al. Physics-informed machine learning for TCO-layer thickness prediction and process analysis from multi-spectral images. *Sol. Energy Mater. Sol. Cells* **285**, 113541 (2025). https://doi.org/10.1016/j.solmat.2025.113541

[33] Ming, Z. et al. Nondestructive measurement of terahertz optical thin films by machine learning based on physical consistency. *Opt. Express* **32**, 16426-16436 (2024). https://doi.org/10.1364/OE.521609

[34] Shen, Y. Twist metrology with twisted light in fibers. *Light: Advanced Manufacturing* **6**, 1 (2025). https://doi.org/10.37188/lam.2025.028

[35] Tang, L. et al. High-precision large-aperture single-frame interferometric surface profile measurement method based on deep learning. *Int. J. Extreme Manuf.* **7**, 055601 (2025). https://doi.org/10.1088/2631-7990/add197

[36] Chen, W. et al. Deep-learning-enabled temporally super-resolved multiplexed fringe projection profilometry: high-speed kHz 3D imaging with low-speed camera. *PhotoniX* **5**, 25 (2024). https://doi.org/10.1186/s43074-024-00139-2

[37] Yin, W. et al. Composite deep learning framework for absolute 3D shape measurement based on single fringe phase retrieval and speckle correlation. *J. Phys. Photonics* **2**, 045009 (2020). https://doi.org/10.1088/2515-7647/abbcd9

[38] Wang, K. et al. On the use of deep learning for phase recovery. *Light: Sci. Appl.* **13**, 4 (2024). https://doi.org/10.1038/s41377-023-01340-x

[39] Feng, S. et al. Deep-learning-based fringe-pattern analysis with uncertainty estimation. *Optica* **8**, 1507 (2021). https://doi.org/10.1364/OPTICA.434311

[40] Liu, H. et al. Deep learning in fringe projection: A review. *Neurocomputing* **581**, 127493 (2024). https://doi.org/10.1016/j.neucom.2024.127493

[41] Yin, W. et al. Physics-informed deep learning for fringe pattern analysis. *Opto-Electron. Adv*. **7**, 230034 (2024). https://doi.org/10.29026/oea.2024.230034

[42] Leng, J. et al. FNIN: A Fourier neural operator-based numerical integration network for surface-form-gradients. *Proceedings of the AAAI Conference on Artificial Intelligence* **39**, 4580-4588 (2025). https://doi.org/10.48550/arXiv.2501.11876

[43] Liu, H. et al. SFC-YOLOv8: Enhanced strip steel surface defect detection using spatial-frequency domain-optimized YOLOv8. *IEEE Transactions on Instrumentation and Measurement* **74**, 9700111 (2025). https://doi.org/10.1109/TIM.2025.3548193

[44] Liang, F. et al. HDFA-Net: A high-dimensional decoupled frequency attention network for steel surface defect detection. *Measurement* **242**, 116255 (2025). https://doi.org/10.1016/j.measurement.2024.116255

[45] Goodman, J. W. *Introduction to Fourier optics*. (Macmillan Learning, 2017)

[46] Zeng, X. *Impact on Image Performance of Surface Spatial Frequency*. Doctoral Dissertation, University of Chinese Academy of Sciences (2014)

[47] Liu, Y. et al. Data-driven precise extraction and division of spatial frequency band errors. *Opt. Laser Technol.* (in the press)