Opto-Electronic Advances

CN 51-1781/TN ISSN (Print) 2096-4579 ISSN (Online) 2097-3993

Spectrally encoded diffractive optical unclonable functions using graphene zone plates

Geon Mo Kim, Yun Ji Hwang, Teajong Hwang, Seo Jun Yoon, Chengyi Li and Seong Chan Jun*

Citation: Kim GM, Hwang YJ, Hwang T et al. Spectrally encoded diffractive optical unclonable functions using graphene zone plates. *Opto-Electron Adv* **9**, 260076 (2026).

https://doi.org/10.29026/oea.2026.260076

Received: 17 March 2026; Accepted: 29 June 2026; Published online: 23 July 2026

Related articles

Rayleigh-driven ethanol cluster tracking based on non-contact deep optical molecular diagnosis

Geon Mo Kim, Yun Ji Hwang, Chengyi Li, et al

Opto-Electronic Advances 2026, **9**(6): 250278 doi: 10.29026/oea.2026.250278

More related articles in Opto-Electronic Journals Group website

Opto-Electronic Journals Group

https://www.oejournal.org

OE_Journal

@OptoElectronAdv

OEA | Opto-Electronic Advances

Article

October 2026, Vol. 9, No. 10

Open Access

Spectrally encoded diffractive optical unclonable functions using graphene zone plates

Geon Mo Kim, Yun Ji Hwang, Teajong Hwang, Seo Jun Yoon, Chengyi Li and Seong Chan Jun*

Abstract: Diffractive photonic structures exhibit high-dimensional, device-specific variability arising from coherent interference, wavelength-dependent phase accumulation, and fabrication-induced disorder. These properties have motivated increasing interest in optical physical unclonable functions (O-PUFs); however, converting analog optical responses into compact and repeatable binary representations suitable for practical authentication remains challenging. In this work, we address this limitation using graphene diffractive zone plates (GDZPs) combined with a vision transformer-based model. Rather than compensating for diffractive chromatic aberration, we exploit the wavelength-dependent focusing behavior of GDZPs as an additional optical challenge dimension. Under multi-wavelength excitation with predefined RGB power ratios, the devices generate spatial-spectral diffraction patterns shaped by device geometry, material dispersion, and inherent fabrication variations, enabling spectrally programmable optical readout. A self-attention-based encoder captures global diffraction features and maps the optical responses into a compact 64-dimensional representation, from which stable 16-bit binary response units are extracted through statistical thresholding and reliability-based bit selection. Experimental evaluation demonstrates highly repeatable binary response units, near-random cross-configuration Hamming-distance behavior, low configuration-level error estimates, and clear separability across spectral challenge conditions. These results establish a proof-of-concept spectral-challenge GDZP-transformer pipeline that converts structured diffractive optical responses into repeatable and composable binary response units, highlighting the potential of programmable spectral-spatial optical authentication.

Keywords: diffractive photonics; optical physical unclonable functions; graphene diffractive zone plates; multi-wavelength excitation; spectral-spatial diffraction; vision transformer

DOI: 10.29026/oea.2026.260076 | CSTR: 32247.14.oea.2026.260076

Citation: Kim GM, Hwang YJ, Hwang T et al. Spectrally encoded diffractive optical unclonable functions using graphene zone plates. *Opto-Electron Adv* **9**, 260076 (2026).

1 Introduction

Wave-based photonic systems intrinsically generate high-dimensional physical variability through coherent interference, structural disorder, and spectral dispersion[1–3]. Such properties have motivated increasing interest in exploiting diffractive optical structures not only for imaging and beam shaping but also for information encoding and hardware-level functionality[4–6]. In distributed and interconnected electronic systems, these characteristics offer opportunities to embed security-relevant variability directly into physical photonic components[7–9].

Among various approaches to hardware-rooted security, physical unclonable functions (PUFs) exploit intrinsic fabrication variations to generate device-specific responses[7,10]. By eliminating the need to store digital secrets, PUFs offer a promising path toward hardware-rooted security[11–13]. Among various PUF modalities, optical PUFs (O-PUFs) are particularly compelling owing to the extreme complexity of light-matter interactions[14,15], which can produce high-entropy physical responses that are difficult to replicate or model[16–19]. In parallel with optical PUFs, a broad range of optical encryption and optical security approaches have

Received: 17 March 2026

Accepted: 29 June 2026

Published online: 23 July 2026

School of Mechanical Engineering, Yonsei University, 50, Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea.
*Correspondence: SC Jun, E-mail: scj@yonsei.ac.kr

been developed by exploiting the high-dimensional degrees of freedom of light[20–22]. Representative examples include double-random phase encryption[23] and its cryptanalysis[24], holographic secret sharing[25], meta-optics-enabled visual cryptography[26], physical-layer optical encryption[27], metasurface-assisted computational imaging encryption[28], and scattering-assisted optical encryption[29]. These approaches demonstrate that wavelength, polarization, phase, amplitude, spatial displacement, scattering paths, and holographic degrees of freedom can serve as optical security variables for information hiding, encryption, decryption, or secret sharing.

While optical encryption and optical PUFs share the use of complex light-matter interactions for security, their primary security functions are different[7,20,21]. Optical encryption schemes mainly focus on hiding, transmitting, or reconstructing information[24,25,28,29], whereas PUFs focus on challenge-response-based physical authentication[7,30–32]. Despite extensive demonstrations of complex physical variability, many existing optical PUF studies have not sufficiently focused on practical authentication requirements[33]. In particular, much of previous research has emphasized raw optical patterns or analog signatures, with less attention given to establishing compact and verifiable digital representations compatible with standard PUF authentication metrics such as the Hamming distance (HD), false acceptance rate (FAR), and false rejection rate (FRR)[34–36]. Consequently, the lack of a stable binary response format has hindered broader adoption of optical PUFs as system-level security primitives[36,37]. This challenge largely emanates from the inherent sensitivity of optical responses to environmental perturbations, including wavelength drift, mechanical misalignment, and measurement noise, which complicates the direct binarization of raw optical data. Thus, many optical PUF systems rely on error correction or repeated measurements to achieve robustness, introducing additional complexity that can limit practical deployment[38,39].

In this work, we investigate how spectrally modulated diffractive responses can be translated into stable digital representations using a learning-based model. We employ graphene diffractive zone plates (GDZPs)[40–43] as a structured diffractive optical front-end and a vision transformer (ViT)[44,45] to map spatial-spectral diffraction patterns into an embedding space. Rather than treating optical variability as noise, the learned representation enables reliable thresholding and extraction of compact binary response units without explicit error correction.

By combining fabrication-induced microscopic disorder with controlled multi-wavelength excitation, the proposed system supports a programmable spectral challenge space while maintaining repeatable optical responses. Rather than presenting GDZPs as a finalized low-cost manufacturing route, this work demonstrates a proof-of-concept diffractive O-PUF architecture that prioritizes spectral-spatial programmability, structured optical responses, and compact embedding-based binary response extraction.

2 Results and discussion

2.1 System architecture and operational principle of the GDZP-based optical PUF

Figure 1 illustrates the overall architecture of the proposed GDZP-ViT optical PUF system, which integrates a structured diffractive optical element with a learning-based readout to enable authentication-oriented response extraction. The system consists of two tightly coupled components: a GDZP optical module that generates complex diffraction responses under controlled excitation and a ViT-based processing pipeline that transforms these raw optical measurements into low-dimensional embeddings.

At the optical front-end, the GDZP introduces multiple sources of response variability. Nanoscale structural irregularities in multilayer graphene, which result from unavoidable fabrication-induced wrinkles, grain boundaries, and thickness variations, introduce local optical perturbations in the fabricated GDZP response (Supplementary Fig. S1). In addition, Fresnel zone plate phase profiles define distinct design classes that produce characteristic diffraction patterns. Furthermore, the optical response is modulated by controlling the relative power ratios of three laser wavelengths (638, 512, and 405 nm). Because of the pronounced negative chromatic aberration of the GDZP, variations in spectral excitation produce a non-proportional spectral-to-spatial intensity mapping through wavelength-dependent diffractive propagation, rather than a material optical nonlinearity.

Importantly, the motivation for using GDZPs is not fabrication simplicity, but the coexistence of a deterministic diffractive transfer function and microscopic graphene-induced local heterogeneity. The Fresnel zone-plate geometry provides a designed macroscopic optical response, including wavelength-dependent focusing and structured interference, whereas multilayer graphene imperfections introduce local physical variations that perturb the diffracted field within the fabricated GDZP.

Upon illumination, the GDZP produces complex spatial-spectral diffraction patterns that are recorded by a standard complementary metal-oxide-semiconductor (CMOS) sensor (Supplementary Fig. S2). The captured signal is decomposed into red (R), green (G), and blue (B) channels, each exhibiting distinct spatial characteristics that reflect wavelength-dependent focusing behavior. Together, these RGB images constitute a high-dimensional raw optical response.

Rather than directly applying authentication metrics to raw optical data, the RGB diffraction patterns are processed by a ViT trained to extract compact and discriminative embeddings. By learning global spatial-spectral correlations across the diffraction patterns, the ViT provides a stable intermediate representation that establishes a consistent

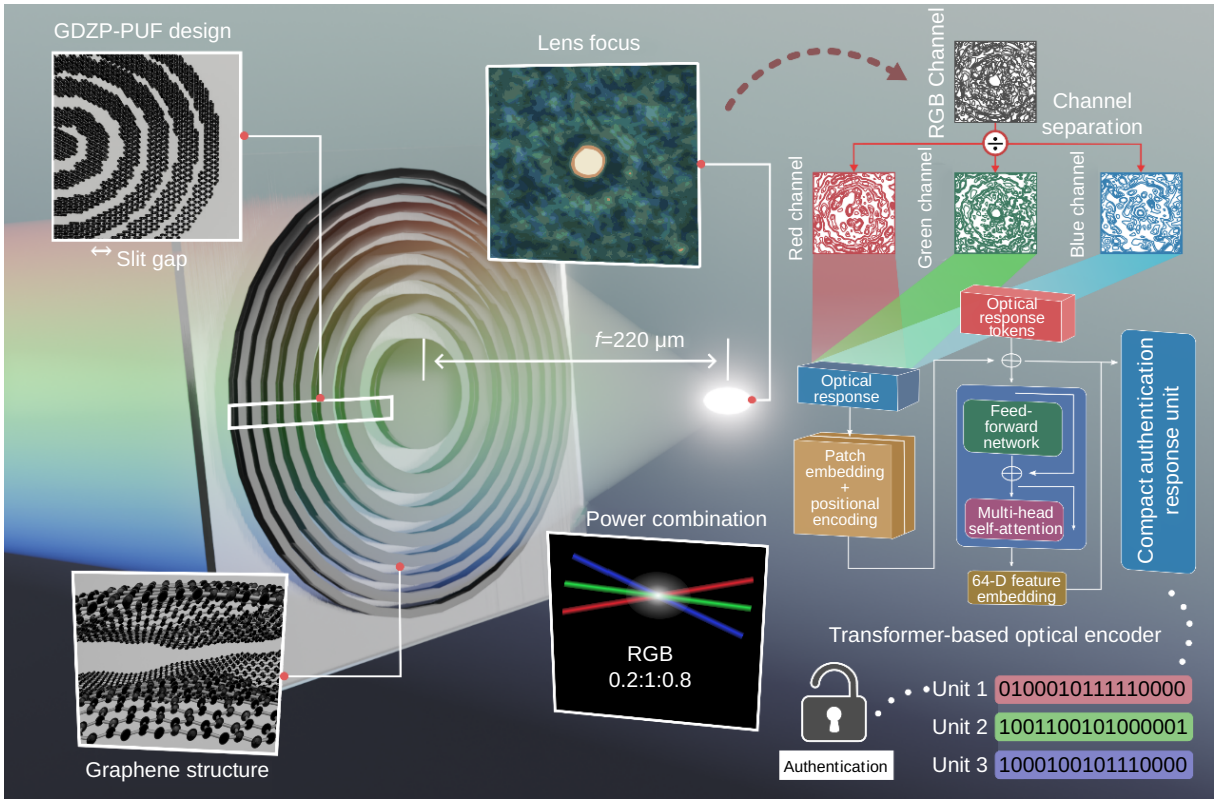

Fig. 1 | Overview of the GDZP-ViT optical PUF system. The system combines a structured GDZP diffractive front-end with a learning-based processing pipeline to convert raw diffraction responses into compact, analyzable representations. The GDZP response arises from the coexistence of deterministic Fresnel-zone-plate diffraction and microscopic graphene-induced disorder, while controlled multi-wavelength excitation modulates the optical response through wavelength-dependent dispersion. The resulting RGB diffraction patterns are captured by a CMOS sensor and processed by a ViT to extract compact embeddings used for subsequent PUF analysis.

measurement-to-embedding pathway for subsequent binary response unit generation, separability analysis, and authentication evaluation.

2.2 Wavelength-selective focusing behavior of GDZPs under multi-wavelength excitation

Figure 2 illustrates the operating principle of the planar diffraction-based GDZP and experimental confirmation of its strong wavelength-dependent focusing behavior. Unlike conventional refractive optics, the GDZP functions as a diffractive element, with concentric graphene zones imposing spatially varying phase shifts on the incident wavefront, resulting in constructive interference at a designated focal plane. According to classical Fresnel zone plate theory, the radius of the m-th zone required to achieve a focal length f at a design wavelength λ is given by Eq. (1),

$$r_m = \sqrt{m\lambda f + \frac{m^2\lambda^2}{4}} \approx \sqrt{m\lambda f}\,, \tag{1}$$

which directly links the geometric phase profile of the GDZP to its operating wavelength (Supplementary Note 1).

Owing to this diffractive phase relationship, the focal length varies inversely with illumination wavelength, resulting in negative chromatic aberration. The wavelength-dependent focal position can be approximated using Eq. (2),

$$f(\lambda) \approx \frac{r_1^2}{\lambda} = f_{\text{design}}\frac{\lambda_{\text{design}}}{\lambda}\,, \tag{2}$$

which explicitly captures the inverse proportionality between focal distance and wavelength. At a fixed observation plane, this inverse dependence produces wavelength-dependent focal-plane mismatch across RGB channels, which appears as channel-selective focusing and defocused interference patterns in the recorded images.

This behavior corresponds to a non-proportional spectral-to-spatial intensity mapping and does not refer to nonlinear optical material effects such as intensity-dependent refractive index, nonlinear absorption, or harmonic generation. Instead, it describes a system-level diffractive mapping in which changes in RGB spectral composition redistribute the measured intensity across the focal core and surrounding interference rings. This behavior arises because the GDZP focal position depends inversely on wavelength, $f(\lambda) \propto 1/\lambda$, and the diffractive phase accumulation depends on wavelength through $\phi(x, y, \lambda) = (2\pi/\lambda)\Delta L(x, y)$.

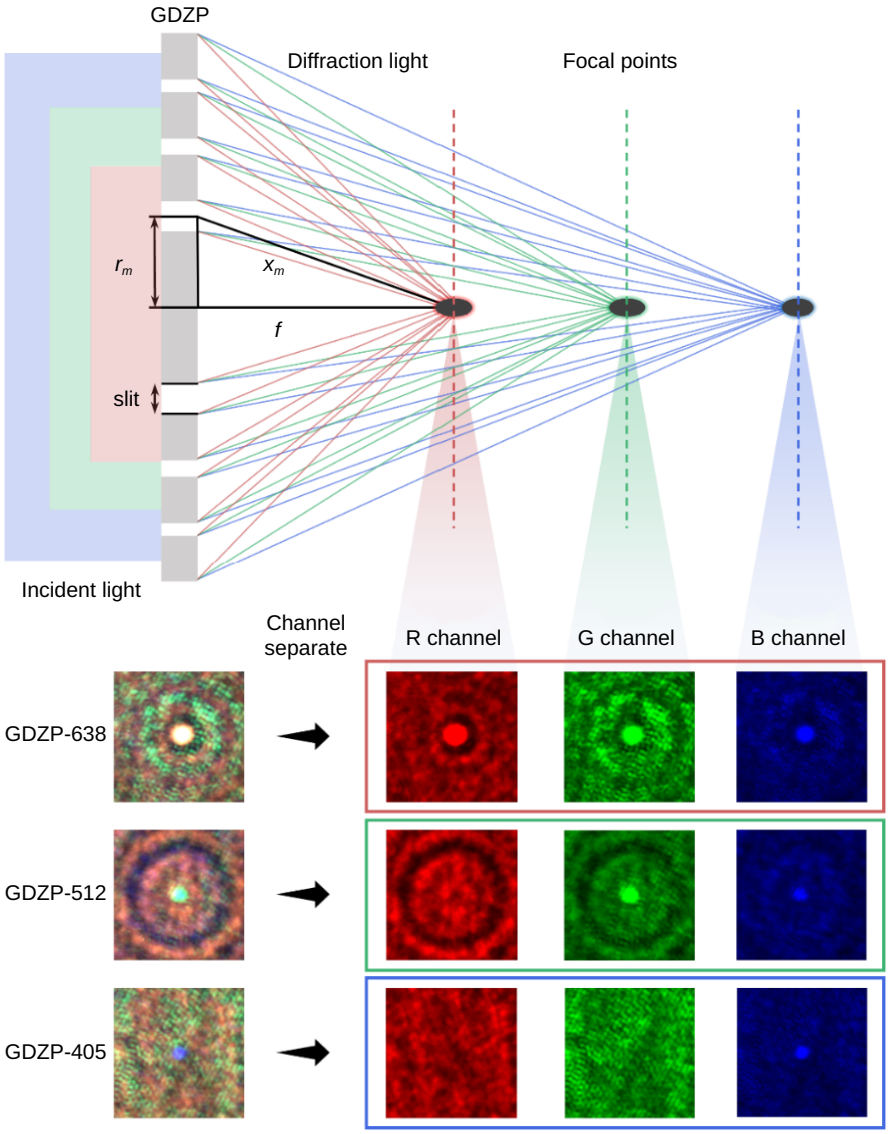

Fig. 2 | Operating principle and wavelength-dependent focusing behavior of GDZPs. The schematic illustrates the planar diffractive mechanism based on Fresnel zone plate theory, in which concentric graphene zones modulate the incident wavefront to produce constructive interference at a focal distance *f*. Because the GDZP is diffractive, the focal position exhibits negative chromatic aberration, shifting systematically with illumination wavelength. Three GDZPs designed for wavelengths of 638, 512, and 405 nm were fabricated to share an identical focal length of *f*=220 μm. Diffraction patterns were recorded at a fixed observation plane under identical RGB mixed-laser illumination with equal power in each channel. Channel-separated images show that each GDZP produces a pronounced focal response primarily in the channel corresponding to its design wavelength, whereas non-matching wavelengths exhibit defocused or low-contrast patterns.

Therefore, at a fixed observation plane, different wavelength components experience different focal-plane mismatches and phase conditions, causing changes in spectral excitation to modify the spatial intensity distribution rather than simply rescale the image brightness. This system-level mapping is schematically summarized in Supplementary Fig. S3.

To validate these theoretical predictions, we fabricated three GDZPs to achieve an identical focal length of $f = 220$ μm at their respective design wavelengths of 638, 512, and 405 nm (Supplementary Fig. S4 and Supplementary Table S1). All diffraction patterns shown in Fig. 2 were recorded at the same observation plane under identical RGB mixed-laser illumination with equal power in each channel, ensuring that the observed differences resulted solely from wavelength-dependent diffraction behavior.

The RGB channel-separated images revealed that each GDZP produced its strongest focal response only in the channel corresponding to its design wavelength. For example, the 638 nm-designed GDZP generated a sharp

focal spot in the red channel, whereas the green and blue channels exhibited defocused interference patterns owing to focal-plane mismatch. A similar wavelength-selective behavior was observed for the 512 nm and 405 nm GDZPs. These measurements confirm that a fixed GDZP geometry can generate distinct spatial responses under different illumination spectra, providing the physical basis for the spectral challenge space used in the subsequent experiments. More specifically, the channel-separated images in Fig. 2 show that wavelength components close to their design condition form sharper focal responses, whereas non-matching channels produce defocused or low-contrast patterns. In addition, the radial intensity profiles in Fig. 3(a–c) and the radial sensitivity curves in Fig. 3(j) show that RGB power-ratio changes alter the radial profile shape and produce radius-dependent modulation, rather than simply multiplying the entire diffraction image by a constant factor. These observations support the interpretation of the GDZP response as a non-proportional spectral-to-spatial intensity mapping.

This wavelength-selective focusing behavior distinguishes the GDZP from purely random scattering media. In random-media O-PUFs, optical complexity is primarily obtained from uncontrolled multiple scattering, whereas in the GDZP platform, the challenge-dependent response is partially design-defined through the diffractive phase profile. Therefore, changing the RGB spectral composition does not merely probe a random medium, but actively reconfigures a structured focal/interference response in a fixed optical geometry. A conceptual comparison between bottom-up random-media O-PUFs and the proposed GDZP-PUF is provided in Supplementary Fig. S5, and the key differences in entropy mechanism, challenge type, optical response, programmability, and response format are summarized in Supplementary Table S2.

2.3 Hardware-level repeatability and condition separability of GDZP optical responses

Figure 3 presents a hardware-level characterization of the GDZP optical responses to assess whether the measured diffraction patterns satisfied two essential requirements for optical PUF operation: (i) high repeatability under identical excitation conditions and (ii) strong separability under distinct spectral challenges. For each spectral challenge p, the captured RGB diffraction images were converted into a luminance-weighted response and normalized to suppress global intensity scaling (Supplementary Fig. S6). To obtain a compact and repeatable descriptor of the diffraction response, we summarized the resulting normalized image $\tilde{I}_p(x,y)$ using an angle-averaged radial intensity profile, which suppresses pixel-level noise while preserving the global spatial structure of the diffraction field.

Figure 3(a–c) show the normalized radial intensity profiles $I_p(r)$ of GDZP-638, GDZP-512, and GDZP-405 under 10 distinct spectral challenges (Supplementary Table S3). The radial profile is defined by circular averaging around the optical axis as

$$I_p(r) = \frac{1}{|\Omega(r)|} \sum_{(x,y)\in\Omega(r)} \tilde{I}_p(x,y) \ , \tag{3}$$

where $\Omega(r)$ denotes the set of pixels located at radius r. The systematic spread of profile families across spectral challenges indicates that spectral excitation reshapes the spatial energy distribution in a lens-dependent manner, consistent with the wavelength-selective behavior of GDZPs (Supplementary Fig. S7). Importantly, the radial descriptor defined in Eq. (3) provides a stable and compact representation suitable for quantitative comparison across repeated measurements and distinct conditions.

To quantitatively assess condition separability while explicitly accounting for repeatability, Fig. 3(d) depicts a condition-wise separability matrix computed from concatenated RGB radial feature vectors. Repeatability under identical excitation conditions was evaluated through the intra-condition variation σ_c and the intra-condition feature distance distribution shown in Fig. 3(d–f). The tightly concentrated intra-condition distances indicate that measurement-to-measurement fluctuations are small relative to inter-condition differences, validating the stability of the hardware response under repeated interrogation.

Each condition corresponds to a specific (lens, spectral challenge) pair. For a given condition c, repeated measurements yield a mean feature vector μ_c and an intra-condition variation σ_c. To directly compare inter-condition differences relative to intra-condition noise, the separability between two distinct conditions $i \neq j$ is defined using the inter-/intra-normalized score

$$\mathrm{Sep}(i,j) = \frac{(1/D)\|\mu_i - \mu_j\|_1}{\frac{1}{2}(\sigma_i + \sigma_j) + \varepsilon} \ , \tag{4}$$

where D is the feature dimension, and ε is a small constant introduced for numerical stability. Figure 3(e) shows the distribution of off-diagonal separability scores computed using Eq. (4), yielding a mean ± standard deviation of 7.72 ± 5.82. A high hit ratio of 0.878 for a threshold of 3.0 indicates that most condition pairs are distinguishable beyond intra-condition variation. This conclusion is further corroborated in Fig. 3(f), which directly compares intra-condition (noise) and inter-condition (signal) feature distances. The intra-condition distances are tightly distributed around 0.00995 ± 0.00363, whereas inter-condition distances increase significantly to 0.0692 ± 0.0390, corresponding to a signal-to-noise separation of approximately 6.95×.

Figure 3(g) visualizes representative condition-mean diffraction patterns ($N = 50$) under a balanced spectral challenge (1 ∶ 1 ∶ 1), together with the corresponding block-wise entropy maps derived from the same responses. This paired visualization illustrates how spatially structured

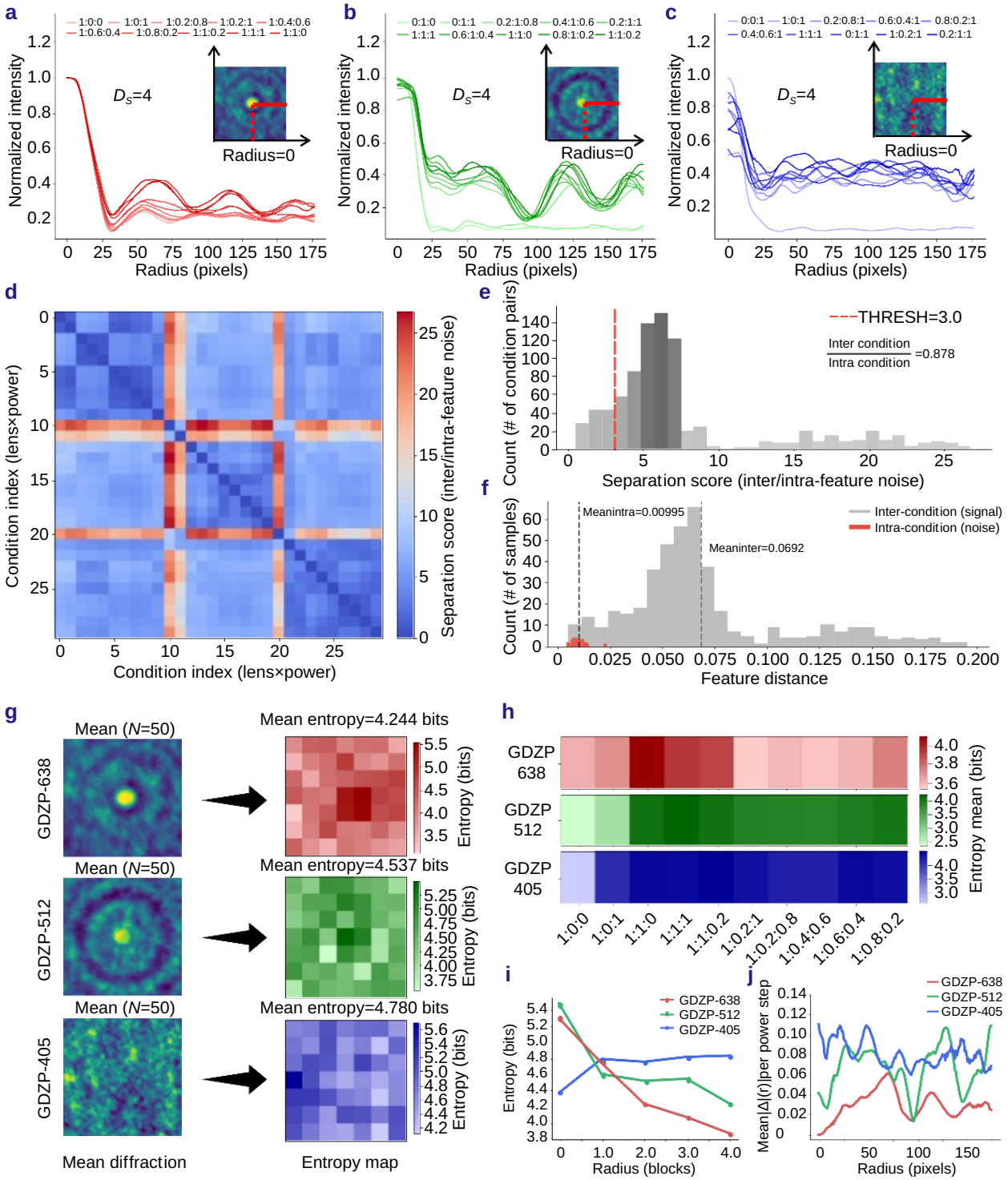

Fig. 3 | Hardware-level analysis of GDZP optical responses under spectral challenges. (a–c) Angle-averaged radial intensity profiles for GDZP-638, GDZP-512, and GDZP-405 under 10 distinct spectral challenges, demonstrating lens-dependent spatial modulation induced by spectral excitation. (d–f) Condition-wise separability analysis accounting for repeatability noise, including the separability matrix, distribution of inter-/intra-normalized separation scores, and comparison of inter-condition and intra-condition feature distances. (g) Representative condition-mean diffraction patterns (N = 50) under a balanced spectral challenge (1∶1∶1) and the corresponding block-wise entropy maps, illustrating the translation from a spatial structure to information content. (h) Lens-wise mean entropy across spectral challenges, highlighting challenge-dependent modulation of response statistics. (i) Radial entropy profiles extracted from the entropy maps, revealing annularly distributed information content. (j) Radial sensitivity curves quantifying the mean absolute change in radial intensity profiles between adjacent spectral challenges, identifying spatial regions with maximal challenge-induced modulation.

diffraction patterns translate into quantitative measures of information content. To compute the entropy maps, we partitioned the normalized luminance image into spatial blocks and calculated the entropy of each block b from the local intensity histogram $p_k(b)$, as defined in Eq. (5).

$$H(b) = -\sum_{k=1}^{K} p_k(b) \log_2 \left(p_k(b) + \varepsilon\right) . \tag{5}$$

This definition yields spatially heterogeneous entropy distributions across the diffraction field. The mean entropy values of 4.244 bits (GDZP-638), 4.537 bits (GDZP-512), and 4.780 bits (GDZP-405) indicate that the diffraction responses contain significant information content and that entropy statistics depend on the GDZP design wavelength.

Figure 3(h) depicts the lens-wise mean entropy across spectral challenges, showing that spectral excitation acts as an effective challenge dimension that modulates the statistical complexity of the optical response. Figure 3(i) further reveals that entropy varies systematically with radius, indicating that informative regions are distributed annularly rather than uniformly across the diffraction field.

Finally, to localize where spectral challenges induce the strongest spatial modulation, Fig. 3(j) examines the radial sensitivity of the diffraction profiles by introducing the sensitivity metric defined in Eq. (6).

$$S(r) = \langle |I_{p+1}(r) - I_p(r)| \rangle_p . \tag{6}$$

This metric quantifies the mean absolute change in radial intensity profiles between adjacent spectral challenges. The resulting sensitivity curves show that challenge-induced modulation concentrates in specific radial regions, providing a physical basis for selecting stable, informative spatial features for downstream readout.

2.4 ViT-based embedding of GDZP-PUF optical responses

Figure 4 examines how the proposed ViT converts raw GDZP diffraction patterns into a compact and stable optical representation suitable for subsequent binary response unit extraction. The controlled diffraction-response regression task and dataset configuration are summarized in Supplementary Table S4. Rather than treating the ViT as a purely discriminative classifier, this analysis focuses on whether the learned embedding preserves physically meaningful structure while suppressing measurement-level variability.

As illustrated in Fig. 4(a), each RGB diffraction image is decomposed into fixed-size patches and embedded with positional encodings to retain spatial context (Supplementary Note 2). Through stacked transformer encoder blocks with multi-head self-attention, a learnable classification (CLS) token aggregates global information distributed across the diffraction pattern, including both the central focal region and peripheral interference structures. The resulting CLS embedding serves as a low-dimensional representation of the optical response (Supplementary Note 3).

The ViT used in this work was designed as a compact ViT-style encoder rather than a large-scale generic vision transformer. Each 1000 × 1000 RGB diffraction image was divided into non-overlapping 50 × 50 patches, yielding 20 × 20 = 400 patch tokens. The patch tokens were projected into a 64-dimensional embedding space and processed by a lightweight transformer encoder with four attention heads and a 128-dimensional feed-forward block. The CLS-token output was then passed to wavelength-specific regression heads for spectral-response prediction. This compact architecture was selected to capture global spectral-spatial correlations in the GDZP diffraction images while limiting model complexity for the controlled optical regression task.

To interpret how the ViT extracts physically relevant features, Fig. 4(b) visualizes the mean CLS-to-patch attention maps averaged over repeated measurements. The attention overlays show that the model tends to emphasize regions associated with diffraction physics, such as wavelength-dependent focal spots and annular interference rings, rather than uniform background intensity (Supplementary Fig. S8). Because attention maps provide qualitative rather than definitive evidence of feature origin, we further evaluated the learned representation using perturbation-based embedding-stability analysis to assess its sensitivity to setup-related image variations such as camera noise, alignment shift, illumination imbalance, and preprocessing-dependent effects.

The stability and separability of the learned representation are quantified in Fig. 4(c), which presents the condition-wise distance matrix in the CLS embedding space. Distinct block structures corresponding to different GDZP designs are observed, whereas intra-block variations reflect controlled modulation by spectral challenges. Importantly, distances between embeddings measured under identical conditions remain small, indicating low intra-condition variability at the representation level.

Figure 4(d–f) further visualize the embedding structure analyzed using principal component analysis (PCA) for each GDZP design. For GDZP-512, the embeddings exhibited a smooth, approximately monotonic trajectory as the spectral power ratio was varied, as reflected in strong rank correlations with the challenge index. In contrast, GDZP-638 and GDZP-405 displayed more curved and branched manifolds, consistent with their stronger wavelength-dependent diffractive responses. Across all designs, nearby spectral challenges mapped to neighboring regions in the embedding space, confirming that the representation varies smoothly with the applied optical challenge.

Because transformer-based models can be susceptible to overfitting when applied to small datasets, the compact ViT-style encoder was evaluated within the controlled data structure of the GDZP diffraction-regression task. Unlike open-ended natural-image recognition, the present input space is

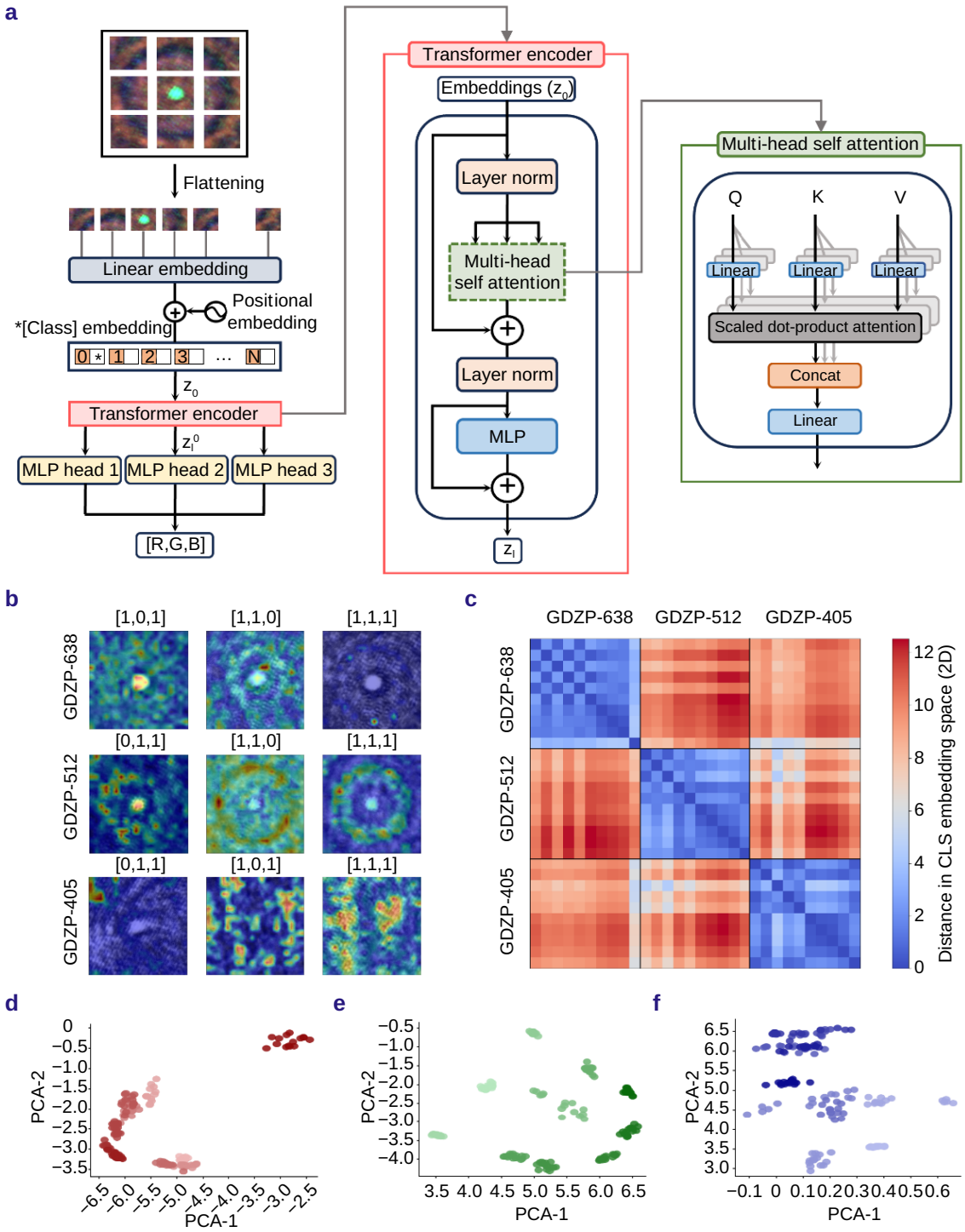

Fig. 4 | ViT-based embedding and representation analysis of GDZP-PUF optical responses. (a) Overall ViT-based encoding architecture. RGB diffraction images are partitioned into patches and processed using transformer encoder blocks with multi-head self-attention. A CLS token aggregates global spatial-spectral information into a compact latent representation. (b) Mean CLS-to-patch attention maps overlaid on condition-averaged diffraction images, indicating that the ViT selectively attends to physically informative regions such as the focal core and wavelength-dependent interference rings. (c) Condition-wise distance matrix computed in the CLS embedding space, showing clear separation across lens designs and spectral challenges. (d–f) PCA projections of sample-wise CLS embeddings for GDZP-638, GDZP-512, and GDZP-405, respectively. Each point corresponds to an individual measurement, colored by spectral challenge. The structured manifolds demonstrate low intra-condition variation and smooth modulation under changing spectral excitation.

physically constrained by the fixed GDZP geometry, fixed imaging configuration, and predefined RGB spectral challenge conditions. For each GDZP design class, 10 spectral challenge conditions were measured repeatedly, and 30 repeated measurements per condition were used for training, corresponding to 300 training images per GDZP model.

To assess the adequacy of the training data volume, a data-size ablation study was performed by varying the number of training repeats per spectral condition from 5 to 30 while keeping the validation set fixed (Supplementary Note 4, Supplementary Fig. S9, and Supplementary Table S5). The validation performance improved with increasing training repeats and approached a plateau between 20 and 30 repeats per condition, indicating that the compact ViT reached stable performance within the available repeated-measurement range.

Training and validation learning curves were further examined using a 150-epoch compact ViT training schedule (Supplementary Note 5, Supplementary Fig. S10, and Supplementary Table S6). For all three GDZP models, both losses decreased rapidly during the early epochs and then remained within stable low-loss ranges. The curves did not show the typical signature of severe overfitting, namely a continuously decreasing training loss accompanied by a sustained increase in validation loss.

Beyond the learning-curve and data-size analyses, we further performed a separate repeat-index-based robustness validation to address possible overestimation from repeated-measurement partitioning. This additional validation was independent of the original 30/5/15 split used for the main ViT evaluation. For each GDZP-challenge condition, the 50 repeated measurements were repartitioned into 30 clean training images, 10 clean validation images, and 10 held-out clean test images. Synthetic perturbations were applied only to the held-out test images after training. Under this test-only perturbation protocol, the clean-trained model retained a mean R^2 of 0.9529 over 2400 perturbed test images, compared with 0.9972 on the clean held-out test set (Supplementary Note 6, Supplementary Fig. S11, and Supplementary Table S7). These results support that the learned GDZP representation is not solely dependent on clean repeated images acquired under identical conditions.

2.5 ViT-based embedding stability and interpretability for PUF authentication

Figure 5 shows the results of evaluating whether the ViT-learned representation of GDZP diffraction patterns satisfies key requirements for practical PUF authentication, namely low intra-condition variability, interpretable modulation under challenge variation, and robustness against optical cross-talk. Rather than treating the ViT solely as a prediction engine, this analysis focuses on the structure and stability of the learned embedding as an intermediate security representation.

Figure 5(a) visualizes the learned embedding of diffraction responses in the three-dimensional spectral parameter space. Each point corresponds to a single optical measurement under a specific spectral challenge. Responses measured under identical spectral challenges form compact clusters, indicating low intra-condition variation at the embedding level. The geometric organization of the clusters further suggests that the embedding preserves meaningful relationships between neighboring spectral challenges.

The stability of the learned representation across the full operating range is examined in Fig. 5(b) and 5(c). The decoded spectral values remain closely aligned with their nominal levels for all wavelengths, and the mean absolute error (MAE) remains below 0.05 mW across power settings. Importantly, this low variance persists even in low-power regimes where raw optical responses are most susceptible to noise, confirming that the embedding effectively suppresses measurement-level instability.

In addition to challenge consistency, the embedding retains configuration-dependent spectral-spatial information associated with the GDZP design and optical response. As shown in Fig. 5(d), GDZP design classes are reliably separable in the embedding space, demonstrating that configuration-dependent spectral-spatial characteristics are preserved alongside spectral modulation (Supplementary Tables S8 and S9). This coexistence of stability and separability is a critical prerequisite for subsequent binary response units extraction.

Figure 5(e) further illustrates that the learned embedding varies monotonically with respect to the applied spectral parameters, enabling reliable thresholding without requiring complex nonlinear post-processing. The high linear correspondence between decoded and nominal values reflects a regularized embedding geometry rather than direct memorization of discrete labels.

Robustness against wavelength cross-talk is depicted in Fig. 5(f). The dominance of diagonal elements indicates that the ViT successfully disentangles overlapping diffraction contributions from different spectral channels, despite the inherently broadband nature of GDZP responses.

Finally, Fig. 5(g) compares the ViT-based embedding with a baseline CNN (Supplementary Figs. S12–S14 and Supplementary Tables S10 and S11). While both approaches achieve high classification accuracy, the ViT exhibits more balanced and stable embedding behavior across spectral dimensions, underscoring the importance of global context modeling for optical PUF representations. Because the final PUF response is derived from the learned embedding rather than directly from the raw image, we further evaluated embedding stability under controlled test-only perturbations. The 64-dimensional embedding extracted from each clean held-out test crop was compared with that from its corresponding perturbed versions using cosine similarity. The embedding remained highly similar under common perturbations, with mean cosine similarities above 0.985,

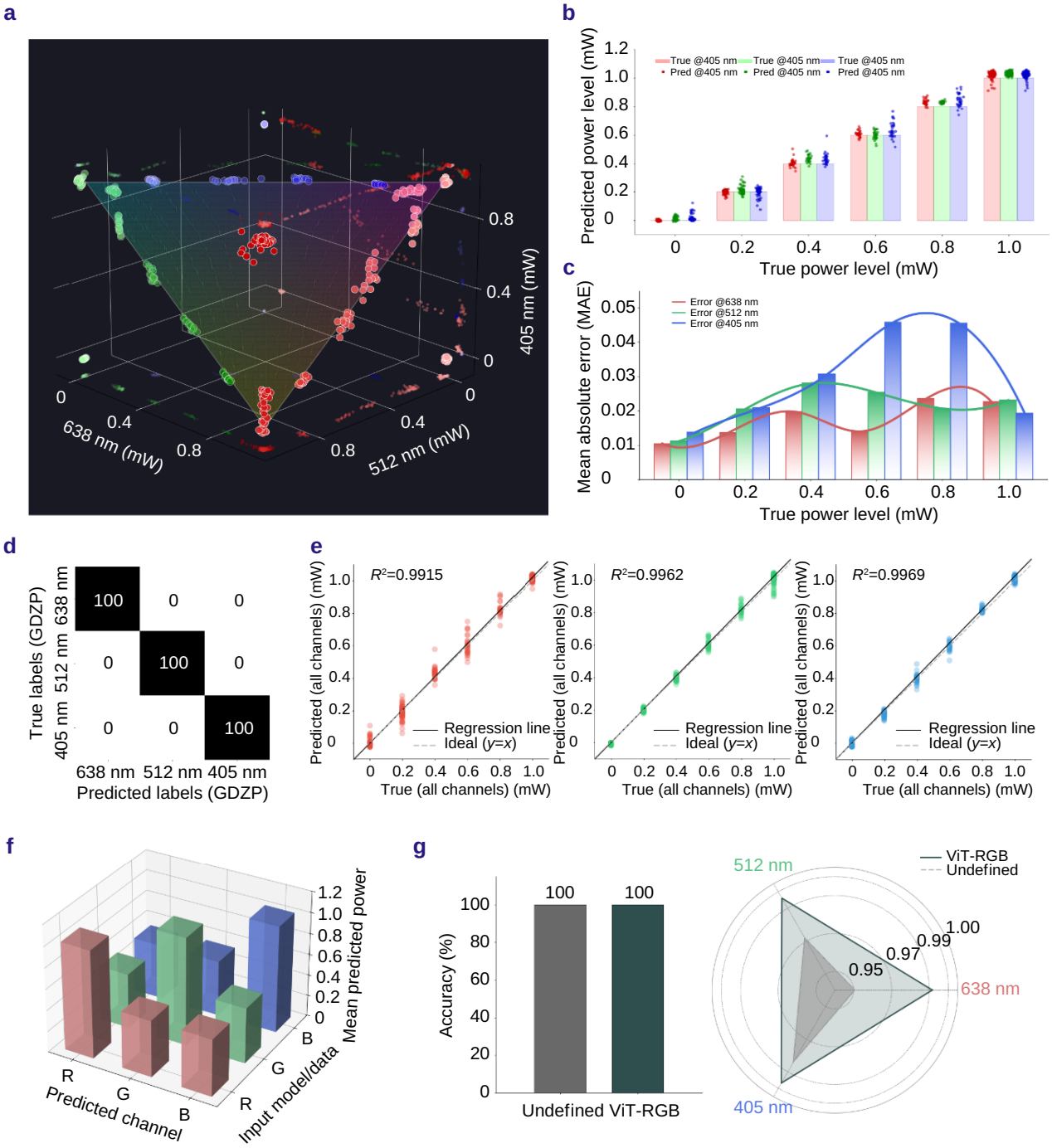

Fig. 5 | Evaluation of ViT-learned optical embeddings and their suitability for PUF authentication. (a) Three-dimensional embedding of decoded RGB spectral conditions, illustrating that responses measured under identical challenges cluster tightly in the learned space. (b) Distribution of decoded spectral values across discrete power levels, demonstrating consistent embedding stability over the full dynamic range. (c) MAE of decoded spectral parameters as a function of power level, reflecting low intra-condition variation. (d) Confusion matrix for GDZP design-class identification, indicating strong configuration-level separability in the embedding space. (e) Linear correspondence between decoded and nominal spectral parameters, confirming monotonic and thresholdable embedding behavior. (f) Cross-channel attribution matrix showing effective disentanglement of wavelength-dependent responses. (g) Comparison with a baseline CNN, highlighting the advantage of global context-aware representations for stable optical embedding.

while stronger changes were observed under illumination-gradient and combined perturbations that degraded the focal/ring optical-response structure. These results indicate that the learned representation is largely preserved under common measurement-like perturbations while remaining sensitive to meaningful optical-response degradation (Supplementary Note 7 and Supplementary Fig. S15).

2.6 Binary response unit extraction and authentication performance

This section establishes a direct link between the learned optical representations and authentication-oriented system metrics, while treating each 16-bit output as a compact response unit generated from a single spectral challenge rather than as a standalone cryptographic key.

Figure 6(a) summarizes the complete binary response extraction pipeline. The designed GDZPs, when illuminated under controlled spectral challenges, generate raw diffraction patterns that encode both fabrication-induced randomness and challenge-dependent optical modulation. These responses are mapped into a 64-dimensional embedding space using the ViT encoder described previously. During enrollment, dimension-wise median thresholds are estimated, and a reliability-driven bit selection strategy is applied to retain only stable dimensions. The selected embedding dimensions are then binarized to generate compact 16-bit binary response units, each corresponding to a single spectral challenge. Longer authentication tokens can be constructed by collecting multiple 16-bit response units from sequential spectral challenges and serializing them using deterministic bit-position interleaving; the detailed token construction procedure and latency estimation are provided in Supplementary Note 8 and Supplementary Figs. S16 and S17.

The binarization process is illustrated in Fig. 6(b), where each selected embedding dimension is compared against its corresponding median threshold. Embedding values exceeding the threshold are assigned a bit value of "1", whereas those below the threshold are assigned "0". Median-based thresholding ensures balanced bit generation while minimizing sensitivity to small embedding fluctuations.

To quantify the response balance, Fig. 6(c) shows the bit uniformity, defined as the probability $P(b_j = 1)$ across selected bits. Because a single 16-bit response unit is too short for meaningful NIST-style statistical testing, randomness was evaluated at the authentication-token level by combining sixteen 16-bit response units into a 256-bit interleaved token. The resulting 7168-bit stream exhibited $P(1) = 0.4999$, a min-entropy of 0.9996 bits per bit, and passed all selected NIST-style tests used in this study at $\alpha = 0.01$. Detailed token-level randomness statistics are provided in Supplementary Note 9, Supplementary Fig. S18, and Supplementary Tables S12–S14.

Response repeatability and cross-configuration separability were evaluated through normalized HD analysis in Fig. 6(d). Repeated-response HDs, computed between repeated measurements acquired under the same GDZP design and spectral challenge, were tightly concentrated near zero with a mean value of approximately 0.008, demonstrating high measurement-level repeatability (Supplementary Fig. S19). In contrast, cross-configuration HDs, computed between response units from different tested GDZP design/challenge configurations, clustered around a mean value of 0.470, close to the ideal value of 0.5 expected for statistically independent binary strings. The separation between these distributions confirms that the extracted response units provide repeatable binary readout and configuration-level response separability within the tested spectral-spatial challenge set.

Figure 6(e) presents the FAR and FRR as functions of the acceptance threshold τ, where a response pair is accepted as matching when the normalized HD satisfies $\mathrm{HD} \leq \tau$. The FAR and FRR curves were derived directly from the empirical repeated-response and cross-configuration HD distributions shown in Fig. 6(d) by sweeping the threshold τ, rather than by fitting the distributions to a parametric probability density function. Specifically, repeated-response pairs were treated as genuine pairs, whereas cross-configuration pairs were treated as impostor pairs. At each threshold, FRR was computed as the fraction of genuine pairs with $\mathrm{HD} > \tau$, and FAR was computed as the fraction of impostor pairs with $\mathrm{HD} \leq \tau$. The EER was determined at the threshold where the absolute difference between FAR and FRR was minimized. The resulting EER of 0.0004 quantifies the empirical separability between repeated-response and cross-configuration response distributions in the present dataset. Accordingly, we refer to this value as a configuration-level EER estimate. Because the response units are derived from learned embeddings with reliability-driven bit selection, the configuration-level EER was further contextualized using perturbation robustness and embedding-stability analyses, which are summarized in Supplementary Notes 6 and 7, Supplementary Figs. S11 and S15, and Supplementary Table S7.

Representative binary response units generated across different GDZP designs and spectral challenges are visualized in Fig. 6(f) (Supplementary Figs. S19 and S20). Distinct spectral challenges applied to the same GDZP design produce different binary response units, while responses from different tested design/challenge configurations remain distinguishable. This visualization provides intuitive support for challenge-dependent response diversity and configuration-level separability. A preliminary modeling-attack analysis using challenge descriptors as inputs is provided in Supplementary Fig. S21 and Supplementary Table S15 to assess whether these compact response units are trivially predictable.

Figure 6(g) evaluates challenge-dependent response diversity by visualizing inter-condition HD matrices for each GDZP design. For a given GDZP design, HDs between response units generated under different spectral challenges

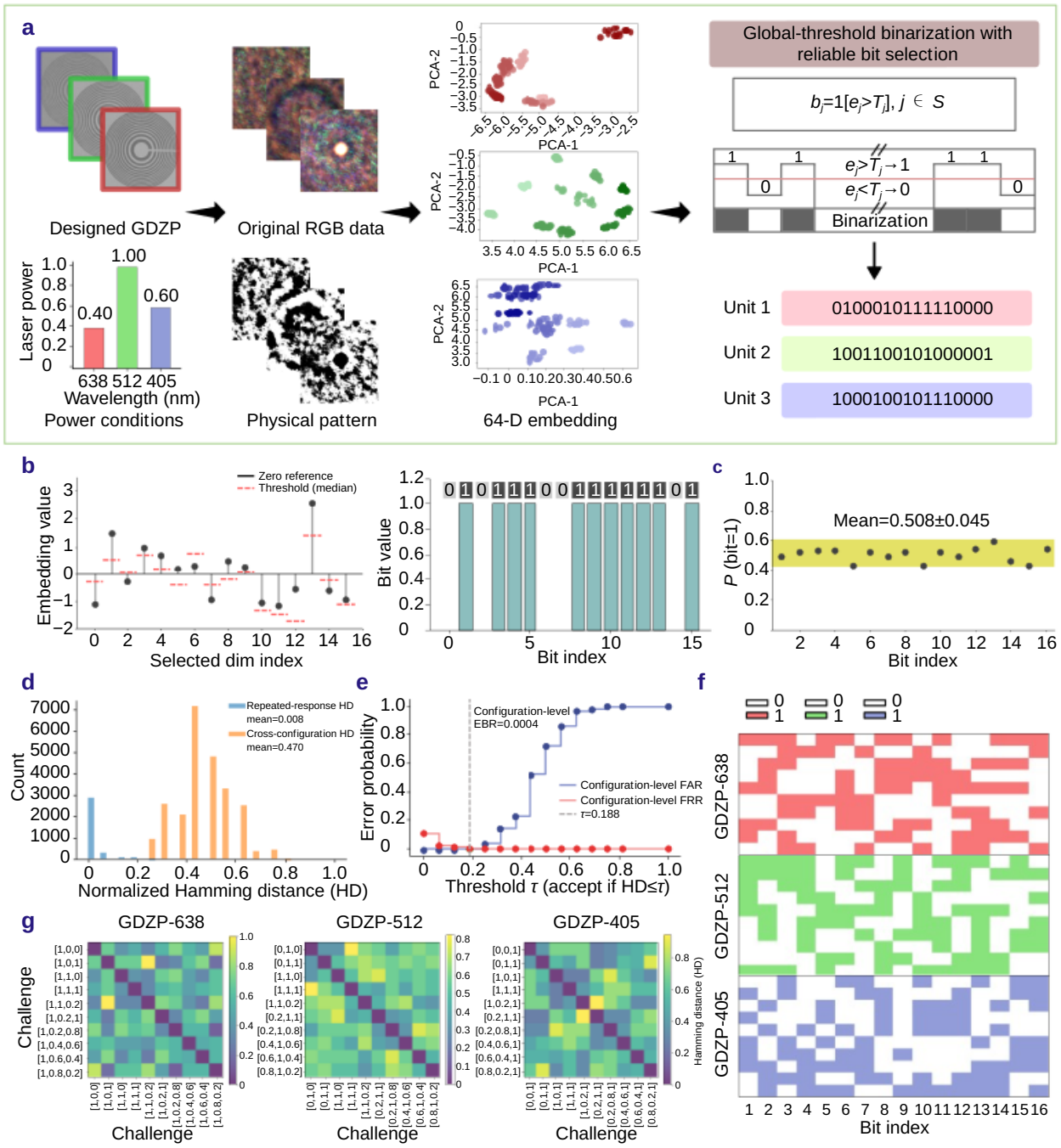

Fig. 6 | Binary response extraction and authentication performance of the GDZP-ViT optical PUF system. (a) Overview of the binary response extraction pipeline. Designed GDZPs under controlled spectral challenges generate raw RGB diffraction patterns, which are transformed into 64-dimensional learned embeddings and subsequently binarized via median thresholding with reliability-driven bit selection to produce 16-bit binary response units. Each response unit corresponds to a single spectral challenge and serves as a compact building block for constructing longer authentication tokens, rather than as a standalone 128-bit or 256-bit cryptographic key. (b) Illustration of dimension-wise binarization by comparing embedding values with median thresholds. (c) Bit uniformity analysis showing the probability of bit value "1" across selected dimensions. (d) Distributions of repeated-response and cross-configuration normalized HDs. The repeated-response HD quantifies measurement repeatability under the same GDZP design and spectral challenge, whereas the cross-configuration HD quantifies separability across the tested GDZP design/challenge configurations. (e) Threshold-dependent estimated FAR and FRR derived from empirical threshold sweeping of the repeated-response and cross-configuration HD distributions in (d), without parametric curve fitting. The resulting EER represents a configuration-level separability estimate for the present dataset. (f) Visualization of representative binary response units across different GDZP designs and spectral challenges. (g) Inter-condition HD matrices for each GDZP design, validating challenge-dependent response diversity across the experimentally tested spectral conditions.

are significantly larger than zero, whereas diagonal elements corresponding to identical conditions remain near zero. This demonstrates that the experimentally tested spectral challenges produce distinguishable response units within each GDZP design. The 30 demonstrated conditions should therefore be interpreted as a controlled proof-of-concept subset used to validate challenge-dependent response diversity, rather than as the theoretical upper limit of the spectral CRP space. A theoretical capacity estimate based on RGB power-level discretization indicates substantially larger addressable spectral challenge spaces and is provided in Supplementary Note 10 and Supplementary Fig. S22.

To further separate design/challenge-level separability from local fabrication-induced spatial heterogeneity, an auxiliary quadrant-level NCC analysis showed that quadrant-to-quadrant variation exceeded repeated-measurement variation, as summarized in Supplementary Fig. S23.

Finally, the proposed GDZP-PUF was compared with representative optical PUF and optical security platforms to clarify its relative positioning (Supplementary Tables S16–S19). Unlike passive random-scattering or emission-based optical PUFs, the proposed platform combines designed Fresnel diffraction, graphene-induced local heterogeneity, and electronically programmable RGB spectral challenges to realize spectrally programmable diffractive authentication in a fixed optical geometry. In contrast to optical encryption or visual-cryptographic schemes that primarily focus on information hiding, transmission, or reconstruction, the GDZP-PUF addresses a complementary security function by converting programmable optical challenges into compact physical authentication responses.

3 Conclusions

We demonstrated a multi-dimensional optical PUF framework by combining graphene diffractive zone plates with transformer-based representation learning, addressing long-standing gaps between optical unclonability and compact authentication-oriented binary response generation. Rather than relying on raw optical patterns or heuristic feature extraction, the proposed system establishes a principled pathway from complex diffraction responses to compact, verifiable binary response units compatible with standard PUF authentication protocols.

At the hardware level, graphene diffractive zone plates provide an ultrathin diffractive optical platform in which deterministic Fresnel-zone-plate geometry, wavelength-dependent diffraction, and fabrication-induced local graphene heterogeneity jointly shape the measured optical responses. Controlled multi-wavelength excitation further expands the spectral challenge space without mechanical scanning, enabling programmable spectral-spatial modulation within a compact optical geometry.

At the system level, leveraging the global attention mechanism to capture long-range diffractive interference features, a vision transformer is employed to convert high-dimensional diffraction patterns into stable and discriminative embeddings that suppress measurement noise while preserving configuration-dependent spectral-spatial structure. These learned representations are deterministically binarized via median thresholding and reliability-driven bit selection, yielding compact binary response units with low repeated-response Hamming distances, near-random cross-configuration Hamming-distance behavior, and a low configuration-level EER estimate under the tested conditions without resource-intensive explicit error correction. Inter-condition Hamming-distance analysis further confirms challenge-dependent response diversity within the tested GDZP design/challenge configurations.

By explicitly following standard PUF-inspired evaluation metrics —including repeatability, HD-based separability, FAR, FRR, EER, and CRP diversity —the proposed framework demonstrates that optical PUFs can be systematically integrated into practical authentication pipelines. Rather than relying on stored digital secrets, authentication reliability is governed by the irreproducibility of the physical optical response and the consistency of the learned representation. Furthermore, the generated response units serve as composable binary units that can be flexibly combined to accommodate varying authentication requirements at higher system layers.

Thus, the GDZP-ViT architecture represents a step toward physically grounded and error-correction-free spectrally programmable optical authentication. By unifying two-dimensional material-based diffractive optics with transformer-based representation learning, this study established a physics-guided optical PUF framework capable of spectrally programmable authentication-oriented readout in next-generation IoT and edge-computing environments. Importantly, the use of electronically controlled RGB spectral challenges extends the platform beyond conventional static PUF readout, enabling potential applications in dynamic optical authentication, wavelength-coded access control, photonic hardware verification, programmable anti-counterfeiting labels, and sensor-coupled security. Future directions include on-chip optical integration, miniaturized authentication modules, expanded population-level uniqueness evaluation using matched GDZP designs, and systematic robustness assessment against advanced modeling and adversarial attacks. Together, these results demonstrate spectrally programmable diffractive authentication based on repeatable binary response extraction and cross-configuration separability, while the quadrant-level analysis supports the presence of local spatial heterogeneity in the fabricated GDZP optical response.

4 Materials and methods

4.1 Synthesis of CVD graphene

Monolayer graphene was grown on a 100 μm-thick copper foil (Goodfellow, 50 mm diameter) via chemical vapor

deposition (CVD) in a quartz tube furnace. The copper substrate was annealed at 1010 °C for 30 min to increase grain size and reduce surface roughness. Graphene growth was subsequently performed at the same temperature using a CH_4/H_2 gas mixture with flow rates of 20 sccm (290 mTorr) and 8 sccm (100 mTorr), respectively. For transfer, a poly(methyl methacrylate) (PMMA) support layer (MicroChem Corp., 950 K A6) was spin-coated at 4200 rpm for 32 s and baked at 180 °C for 2 min.

4.2 Fabrication of graphene diffractive zone plates

A multilayer graphene film consisting of five stacked layers was prepared through repeated transfer processes. The copper growth substrate was etched using an ammonium persulfate solution ($(NH_4)_2S_2O_8$, Sigma-Aldrich), and the floating graphene/PMMA film was rinsed with deionized water before being transferred onto a glass substrate. This procedure was repeated five times to achieve the desired thickness. Diffractive zone plate patterns were defined using focused ion beam (FIB) lithography. A titanium (Ti) hard mask was deposited prior to milling, and patterning was performed using a Crossbeam 540 system (Ga^+ ion source, 30 kV acceleration voltage, 3 nA beam current, dose factor 0.8). The exposed graphene regions were removed using oxygen plasma etching, followed by removal of the residual Ti mask using a selective etchant. The repeated graphene transfer and FIB-based patterning process was used in this study to accurately define the GDZP geometry and validate the proposed diffractive PUF concept, rather than to represent an optimized mass-manufacturing process.

4.3 Optical measurement and data acquisition

The optical response of the graphene diffractive zone plates was characterized using a multi-line laser source (Cobolt Skyra) and a high-resolution CMOS camera (Teledyne FLIR Blackfly S USB3, 12.3 MP). The laser provided simultaneous emission at wavelengths of 638 nm (red), 512 nm (green), and 405 nm (blue). All measurements were conducted under ambient laboratory conditions (approximately 24 °C). The laser beam was transmitted through the graphene lens, and the resulting diffraction and focusing patterns were recorded at the focal plane. To construct a controlled experimental CRP subset, we modulated the optical excitation by varying the relative laser powers across the RGB channels. Three wavelength-dominant classes (red, green, and blue) were defined, each comprising 10 distinct power combinations, yielding a total of 30 experimentally tested challenge conditions. These 30 conditions were selected as a proof-of-concept subset for evaluating response repeatability and condition separability, rather than as the full spectral challenge capacity of the proposed platform. For each condition, diffraction images were acquired repeatedly to assess response stability under nominally identical excitation.

4.4 ViT-based embedding extraction

A ViT model was employed to convert raw RGB diffraction images into compact embeddings. The model was implemented using Python 3.12.7 and PyTorch 2.5.1+cu121. Rather than relying on pixel-wise similarity, the ViT learns global spatial–spectral correlations within the diffraction patterns, producing a fixed-length embedding vector for each optical response. In this study, each diffraction image was mapped to a 64-dimensional embedding, which serves as a learned representation of the underlying optical response. The architectural details and training procedures are provided in Supplementary Note 3.

4.5 Binary response extraction and authentication protocol

The extracted embeddings were converted into binary response units to enable standard PUF authentication. During the enrollment phase, statistical thresholds were estimated for each embedding dimension using enrollment data acquired under controlled conditions. Each embedding was then binarized by comparing its components to the corresponding thresholds, yielding a binary PUF response corresponding to a single optical challenge. To enhance reliability, we selected a subset of embedding dimensions with low bit-flip probability across repeated measurements, yielding a compact and stable binary response unit rather than a standalone cryptographic key. Authentication performance was evaluated using normalized HD between binary responses. Repeated-response and cross-configuration HD distributions were constructed as genuine and impostor distributions, respectively. For a given acceptance threshold τ, authentication was accepted when the normalized HD satisfied $\mathrm{HD} \leq \tau$. The false rejection rate was calculated as $\mathrm{FRR}(\tau) = \dfrac{\#\{\text{genuine pairs with HD} > \tau\}}{\#\{\text{genuine pairs}\}}$, and the false acceptance rate was calculated as $\mathrm{FAR}(\tau) = \dfrac{\#\{\text{impostor pairs with HD} \leq \tau\}}{\#\{\text{impostor pairs}\}}$. The EER was determined as the point at which $|\mathrm{FAR}(\tau) - \mathrm{FRR}(\tau)|$ was minimized. No parametric curve fitting, such as Gaussian, lognormal, or beta-distribution fitting, was used to compute FAR, FRR, or EER; the error curves were obtained by empirical threshold sweeping over the measured HD values.

4.6 Theoretical CRP-capacity estimation

The theoretical CRP capacity of the spectral-multiplexing approach was estimated from the number of distinguishable optical power levels assigned to the RGB wavelength channels. Representative independent-control and fixed-total-power estimates are summarized in Section 2.6. The full combinatorial derivation and practical assumptions are provided in Supplementary Note 10.

4.7 Authentication-token construction and latency estimation

Longer authentication tokens were constructed by combining multiple 16-bit response units obtained from sequential spectral challenges using deterministic bit-position interleaving. Each 16-bit output was treated as a compact response unit generated from a single spectral challenge rather than as a standalone cryptographic key. The end-to-end readout latency was estimated by summing representative optical acquisition time and measured computational inference time for each response unit. Under the assumed acquisition and computation conditions, one 16-bit response unit required approximately 44.66 ms, corresponding to 357.24 ms for a 128-bit token and 714.48 ms for a 256-bit token. Detailed latency components are provided in Supplementary Note 8 and Supplementary Fig. S16.

4.8 Temporal-consistency analysis

To assess temporal consistency of the GDZP diffractive response, representative red-channel focal images acquired over an approximately five-month interval were compared. For each image, the focal center was independently detected, and a 500 × 500 pixel focal-region crop was extracted. The cropped images were converted to grayscale and peak-normalized before comparison. Image-level similarity was quantified using the 2D Pearson correlation coefficient and the structural similarity index measure (SSIM). In addition, angle-averaged radial intensity profiles were extracted from the normalized focal crops, and the Pearson correlation between radial profiles was calculated to evaluate the preservation of the focal-core and diffraction-ring structures. This analysis was used to assess temporal response consistency under laboratory storage and repeated measurement conditions (Supplementary Fig. S24 and Supplementary Table S20).

4.9 Preliminary machine-learning modeling-attack evaluation

To evaluate the predictability of the GDZP-PUF responses under a preliminary modeling-attack setting, three supervised learning models were trained to predict the 16-bit binary response units from the applied challenge descriptors. The attacker input consisted of the GDZP design class and the RGB optical power combination used for each spectral challenge. The target output was the corresponding 16-bit binary response unit generated by the GDZP-ViT pipeline. Three representative attack models were evaluated: logistic regression, random forest, and multilayer perceptron. Model performance was quantified using bit-wise prediction accuracy and the normalized Hamming distance between the predicted and true 16-bit response units. A random-guessing baseline corresponds to a bit-wise accuracy of 50% and an expected normalized Hamming distance of 0.5 for balanced binary responses. This evaluation was intended as an initial predictability test of the compact challenge-response mapping, rather than as a complete adversarial security analysis. More exhaustive modeling-attack evaluation would require larger sets of independently fabricated nominally identical devices, expanded challenge libraries, device-wise train/test separation, and stronger attack architectures.

4.10 NIST-style randomness evaluation

NIST-style randomness tests were performed at the authentication-token level because a single 16-bit response unit is too short for meaningful statistical testing. Sixteen 16-bit response units were combined into one 256-bit interleaved authentication token, and 28 such tokens were evaluated, yielding 7168 bits. The selected tests included Frequency Monobit, Block Frequency, Runs, Longest Run of Ones, Serial, and Approximate Entropy tests at a significance level of $\alpha = 0.01$. Detailed token construction, additional randomness statistics, and optional SHA-256 post-processing results are provided in Supplementary Note 9, Supplementary Fig. S18, and Supplementary Tables S12–S14.

4.11 Repeat-index-based validation and test-only perturbation analysis

The main ViT and baseline CNN evaluations used the original 30/5/15 split described in Supplementary Note 2. For additional robustness validation, a separate repeat-index-based 30/10/10 split was used. Synthetic perturbations were applied only to the held-out clean test crops after training, and the model was evaluated without retraining. The perturbation conditions included read noise, shot noise, intensity variation, gamma/exposure variation, blur, spatial shift, illumination-gradient imbalance, and combined perturbations. Model robustness was quantified using MSE, MAE, and mean R^2 between the predicted and nominal spectral challenge parameters. Embedding stability was quantified using cosine similarity between clean and perturbed 64-dimensional ViT embeddings. Detailed perturbation ranges and combined-perturbation settings are provided in Supplementary Note 6 and Supplementary Table S7.

References

1. Lin X, Rivenson Y, Yardimci NT et al. All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008 (2018).
2. Li JX, Mengu D, Luo Y et al. Class-specific differential detection in diffractive optical neural networks improves inference accuracy. *Adv Photonics* **1**, 046001 (2019).
3. Goorden SA, Horstmann M, Mosk AP et al. Quantum-secure authentication of a physical unclonable key. *Optica* **1**, 421–424 (2014).

4. Zhu HH, Zou J, Zhang H et al. Space-efficient optical computing with an integrated chip diffractive neural network. *Nat Commun* **13**, 1044 (2022).
5. Hu JT, Mengu D, Tzarouchis DC et al. Diffractive optical computing in free space. *Nat Commun* **15**, 1525 (2024).
6. Yan T, Yang R, Zheng ZY et al. All-optical graph representation learning using integrated diffractive photonic computing units. *Sci Adv* **8**, eabn7630 (2022).
7. Pappu R, Recht B, Taylor J et al. Physical one-way functions. *Science* **297**, 2026–2030 (2002).
8. Wang ZY, Wang H, Wang PX et al. Robust optical physical unclonable function based on total internal reflection for portable authentication. *ACS Appl Mater Interfaces* **16**, 27926–27935 (2024).
9. Kim K, Moon Y, Kim YK et al. Optical physical unclonable functions based on Förster resonant energy transfer in multi-color quantum dots. *Commun Mater* **6**, 270 (2025).
10. Suh GE, Devadas S. Physical unclonable functions for device authentication and secret key generation. In *Proceedings of the 44th Annual Design Automation Conference* 9–14 (ACM, 2007). https://doi.org/10.1145/1278480.1278484.
11. Yanambaka VP, Mohanty SP, Kougianos E. Making use of manufacturing process variations: a dopingless transistor based-PUF for hardware-assisted security. *IEEE Trans Semicond Manuf* **31**, 285–294 (2018).
12. Herder C, Yu MD, Koushanfar F et al. Physical unclonable functions and applications: a tutorial. *Proc IEEE* **102**, 1126–1141 (2014).
13. Rührmair U, Sehnke F, Sölter J et al. Modeling attacks on physical unclonable functions. In *Proceedings of the 17th ACM Conference on Computer and Communications Security* 237–249 (ACM, 2010). https://doi.org/10.1145/1866307.1866335.
14. Al-Meer A, Al-Kuwari S. Physical unclonable functions (PUF) for IoT devices. *ACM Comput Surv* **55**, 314 (2023).
15. Gan ZX, Chen FL, Li Q et al. Reconfigurable optical physical unclonable functions enabled by VO_2 nanocrystal films. *ACS Appl Mater Interfaces* **14**, 5785–5796 (2022).
16. Dal Negro L, Boriskina SV. Deterministic aperiodic nanostructures for photonics and plasmonics applications. *Laser Photonics Rev* **6**, 178–218 (2012).
17. Kuznetsov AI, Miroshnichenko AE, Brongersma ML et al. Optically resonant dielectric nanostructures. *Science* **354**, aag2472 (2016).
18. Delvaux J. Machine-learning attacks on PolyPUFs, OB-PUFs, RPUFs, LHS-PUFs, and PUF-FSMs. *IEEE Trans Inf Forensics Secur* **14**, 2043–2058 (2019).
19. Tuyls P, Škorić B, Stallinga S et al. Information-theoretic security analysis of physical uncloneable functions. In *Proceedings of the 9th International Conference on Financial Cryptography and Data Security* 141–155 (Springer, 2005). https://doi.org/10.1007/11507840_15.
20. Chen W, Javidi B, Chen XD. Advances in optical security systems. *Adv Opt Photonics* **6**, 120–155 (2014).
21. Javidi B, Carnicer A, Yamaguchi M et al. Roadmap on optical security. *J Opt* **18**, 083001 (2016).
22. Liu S, Guo CL, Sheridan JT. A review of optical image encryption techniques. *Opt Laser Technol* **57**, 327–342 (2014).
23. Refregier P, Javidi B. Optical image encryption based on input plane and Fourier plane random encoding. *Opt Lett* **20**, 767–769 (1995).
24. Liao MH, Zheng SS, Pan SX et al. Deep-learning-based ciphertext-only attack on optical double random phase encryption. *Opto-Electron Adv* **4**, 200016 (2021).
25. Li KY, Wang YM, Pi DP et al. Multi-dimensional multiplexing optical secret sharing framework with cascaded liquid crystal holograms. *Opto-Electron Adv* **7**, 230121 (2024).
26. Zhang F, Guo YH, Pu MB et al. Meta-optics empowered vector visual cryptography for high security and rapid decryption. *Nat Commun* **14**, 1946 (2023).
27. Wu B, Zhang WK, Zhou HL et al. Partially coherent optical chip enables physical-layer public-key encryption. *Opto-Electron Adv* **8**, 250098 (2025).
28. Zheng PX, Dai Q, Li ZL et al. Metasurface-based key for computational imaging encryption. *Sci Adv* **7**, eabg0363 (2021).
29. Bian LH, Chang XY, Jiang SW et al. Large-scale scattering-augmented optical encryption. *Nat Commun* **15**, 9807 (2024).
30. Bosworth BT, Atakhodjaev IA, Kossey MR et al. Unclonable photonic keys hardened against machine learning attacks. *APL Photonics* **5**, 010803 (2020).
31. Wan YB, Wang PD, Huang F et al. Bionic optical physical unclonable functions for authentication and encryption. *J Mater Chem C* **9**, 13200–13208 (2021).
32. Xu R, Feng M, Xie JY et al. Physically unclonable holographic encryption and anticounterfeiting based on the light propagation of complex medium and fluorescent labels. *ACS Appl Mater Interfaces* **16**, 2888–2901 (2024).
33. Gao YS, Al-Sarawi SF, Abbott D. Physical unclonable functions. *Nat Electron* **3**, 81–91 (2020).
34. Im H, Yoon J, So B et al. Four-dimensional physical unclonable functions and cryptographic applications based on time-varying chaotic phosphorescent patterns. *ACS Nano* **18**, 11703–11716 (2024).
35. Kim K, Kim SU, Choi MY et al. Voxelated opto-physically unclonable functions via irreplicable wrinkles. *Light Sci Appl* **12**, 245 (2023).
36. Klausen M, Zhang JF, Stevens MM. Designing physical unclonable functions from optically active materials. *Adv Mater* **37**, 2502059 (2025).
37. Wang K, Shi JW, Lai WX et al. All-silicon multidimensionally-encoded optical physical unclonable functions for integrated circuit anti-counterfeiting. *Nat Commun* **15**, 3203 (2024).
38. Park T, Kim J, Ko R et al. Light in, sound keys out: photoacoustic PUFs from stochastic nanocomposites. *Nat Commun* **16**, 7323 (2025).
39. Luescher AM, Gimpel AL, Stark WJ et al. Chemical unclonable functions based on operable random DNA pools. *Nat Commun* **15**, 2955 (2024).
40. Kong XT, Khan AA, Kidambi PR et al. Graphene-based ultrathin flat lenses. *ACS Photonics* **2**, 200–207 (2015).
41. Park S, Park C, Hwang YJ et al. Focus-tunable planar lenses by controlled carriers over exciton. *Adv Opt Mater* **9**, 2001526 (2021).
42. Park S, Lee G, Park B et al. Electrically focus-tuneable ultrathin lens for high-resolution square subpixels. *Light Sci Appl* **9**, 98 (2020).
43. Kim GM, Hwang YJ, Li CY et al. Rayleigh-driven ethanol cluster tracking based on non-contact deep optical molecular diagnosis. *Opto-Electron Adv* **9**, 250278 (2026).
44. Vaswani A, Shazeer N, Parmar N et al. Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* 6000–6010 (Curran Associates Inc. , 2017). https://dl.acm.org/doi/10.5555/3295222.3295349
45. Dosovitskiy A, Beyer L, Kolesnikov A et al. An image is worth 16x16 words: transformers for image recognition at scale. In *Proceedings of the 9th International Conference on Learning Representations* 1–21 (OpenReview. net, 2021). https://iclr.cc/virtual/2021/oral/3458

Acknowledgements

This work was supported by the National Research Foundation of Korea grant funded by the Korean government (RS-2024-00339770).

Author contributions

G.M.K., S.C.J. designed the research. G.M.K. performed the experiments. G.M.K. and Y.J.H. analyzed data and wrote the manuscript. T.H. and S.J.Y assisted data analyzation and visualization. C. L. fabricated the Fresnel lens. S.C.J supervised the research.

Competing interests

The authors declare no competing interests.

Supplementary information

Supplementary information for this paper is available at
https://doi.org/10.29026/oea.2026.260076

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/
©The Author(s) 2026.
Published by Editorial Office of *Opto-Electronic Advance*, Institute of Optics and Electronics, Chinese Academy of Sciences.

Opto-Electronic Journals Group
www.oejournal.org
Scan for Article PDF