

Single-Photon Object Identification without Imaging via Quantum State Projection

Daeho Yang

sheepvs5@gachon.ac.kr

Gachon University <https://orcid.org/0000-0002-4575-037X>

Research Article

Keywords: Quantum optics, Single photon, Object identification

Posted Date: September 11th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10991571/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Single-Photon Object Identification without Imaging via Quantum State Projection

Heeseong Hwang,^{*} Hana Jang,^{*} and Daeho Yang[†]

*Department of Physics, Gachon University, 1342 Seongnam-daero,
Sujeong-gu, Seongnam-si, Gyeonggi-do, Korea*

(Dated: September 9, 2026)

Object identification is central to modern sensing, widely employed in autonomous systems and industrial inspection. However, conventional approaches rely on image formation followed by computational inference, requiring many photons to suppress shot noise and introducing readout and computational latencies. These limitations arise because direct position-basis measurement is not optimized for discriminating object-encoded states. Here, we demonstrate single-photon object identification without imaging by implementing quantum state discrimination in a high-dimensional spatial Hilbert space. By applying a task-optimized unitary transformation prior to wavefunction collapse, we transform spatially overlapping object-encoded wavefields into spatially separated states. Thus, the measurement basis is engineered so that the collapse of the single-photon state directly specifies the object label. Numerically, our method achieves 90.2% identification accuracy for nine objects, surpassing the 31.9% classical limit obtained from maximum-likelihood single-photon intensity imaging. Experimentally, the framework achieves 71.0% conditional identification accuracy under attenuated coherent illumination with camera detection, and is further demonstrated with heralded single photons using only 36 detection channels. Our framework establishes a quantum-measurement approach to information-efficient sensing that connects quantum state discrimination with computational optics, providing a route toward low-photon and ultrafast object identification.

Object identification constitutes a foundational component of modern science and technology, with critical applications spanning autonomous systems, medical imaging, and industrial inspection [1, 2]. Typical approaches rely on acquiring high-resolution images through the integration of intensity signals from a large number of photons, followed by computational analysis including deep learning approaches [3]. This paradigm, however, is constrained by fundamental limits arising from the requirement for high photon flux, which can make acquisition slow, energy-intensive, and potentially damaging to sensitive biological or material samples [4, 5]. These limitations have motivated extensions of the imaging paradigm to low-light object detection, but many photons are still required to suppress shot noise [6, 7]. Moreover, the image acquisition and subsequent processing introduce readout and computational latencies, which hinder decision-making at ultrafast timescales relevant for analyzing transient phenomena.

Beyond efforts to extend imaging to lower light levels, computational sensing has explored measurements that selectively acquire task-relevant information [8, 9]. A similar direction has emerged at the quantum limit of imaging, where mode-resolved measurements have been proposed for specific estimation and discrimination problems [10, 11]. More recently, quantum computational sensing has interleaved sensing and quantum computing operations to compute task-specific functions of physical signals

prior to measurement [12].

Despite these advances, the identification of physical objects through quantum measurement remains less established. In quantum imaging and sensing, optimized measurements have primarily addressed optical parameter estimation or source discrimination rather than physical-object identification [10, 13–15], while high-dimensional quantum state discrimination has largely focused on spatial states with predefined mode structures for information encoding [16, 17]. These developments leave open whether a physical object can be identified directly by measuring the single-photon spatial state modified through interaction with the object, without intermediate image reconstruction.

Here, we realize single-photon object identification without imaging by implementing quantum state discrimination in the continuous high-dimensional spatial Hilbert space of a single photon. Our method applies a task-optimized unitary transformation using a spatial light modulator (SLM), transforming spatially overlapping object-encoded wavefields into spatially separated states so that the location of wavefunction collapse directly specifies an object label. This framework establishes an alternative sensing architecture in which recognition is performed by discriminating non-orthogonal states rather than reconstructing spatial images. Our method surpasses the image-based measurement limit [18], theoretically improving the identification accuracy from the conventional limit of 31.9% to

90.2%, and experimentally achieving an identification accuracy of 71.0% conditioned on photon detection events. Beyond this proof-of-principle demonstration, the approach provides a general framework for designing state-discrimination-based measurements, enabling applications from low-damage precision metrology and quantum-inspired imaging to intelligent optical recognition and quantum-secure measurement systems [19].

RESULTS

A conventional imaging approach (Fig. 1a) attempts to classify the object by integrating many single-photon detection events into a spatial intensity distribution $I_i(x) \propto |\psi_i(x)|^2$ of the i -th object. The corresponding projection onto the detector area is described by

$$\hat{\Pi} = \int_{\Omega} |x\rangle\langle x| dx, \quad (1)$$

where Ω is the detector area. The resulting image is processed by a computational algorithm for identification. Since each photon contributes only a single detection point, a large number of photons must be accumulated to suppress shot noise and form a recognizable pattern. The imaging process also discards the phase information encoded in the wavefield, limiting the information extractable from each photon.

On the other hand, our method performs projective state discrimination using a spatial light modulator (SLM) and free-space propagation (Fig. 1b). The object-dependent wavefield is transformed by a pre-computed pattern and evolves toward a localized region at the detector plane (Fig. 1c). Then, a single detection event corresponding to a projective measurement of the wavefield yields classification without any additional computation. In this framework, the photon propagates through the optical system as a spatial wavefield, a superposition state that encodes the spatial features of the object, but behaves as a discrete particle upon detection. The transformation coherently processes the spatial components of this superposition in parallel, allowing information distributed across the photon wavefunction to contribute to a single classification outcome.

Figure 1d illustrates the role of the unitary transformation from the perspective of projective measurement. Without the transformation, the object-dependent fields $\psi_i(x)$ produce overlapping intensity distributions at the detector, whereas an optimized transformation maps each wavefield predominantly into a distinct detection region. The SLM, free-space

propagation, and detector therefore operate as a unified sensor complex implementing a state-resolving measurement, with the overall transformation represented as $\hat{U} = \hat{U}_{\text{prop}}\hat{U}_{\text{SLM}}$.

The corresponding measurement operator is given by $\hat{\Pi}_j^U = \int_{\Omega_j} \hat{U}^\dagger |x\rangle\langle x| \hat{U} dx$, where Ω_j is the detection region associated with object j . For an intuitive interpretation, when restricted to the object subspace and assuming weak residual crosstalk, the measurement response simplifies to

$$C_{ij} = \langle \psi_i | \hat{\Pi}_j^U | \psi_i \rangle \approx \epsilon_i \delta_{ij}, \quad (2)$$

where ϵ_i denotes the detection efficiency associated with object i . Equation 2 represents the effective diagonalization of the measurement response within the object subspace induced by the optimized transformation.

The measurement basis can be further optimized beyond the spatial constraints of a grid-based quasi-basis (Fig. 1e). More generally, the detection regions corresponding to each object need not be confined to a fixed grid and can instead be optimized jointly with the unitary transformation. The resulting irregular regions form an effective measurement basis that enhances the separability of the object-dependent intensity distributions, and this configuration is referred to as the unrestricted-mode measurement. Implementing such irregular detection regions, however, is challenging with a limited number of photodetectors. The restricted-mode measurement is therefore implemented using a fixed detector array comprising a few tens of channels, while the unitary transformation and the assignment of detector channels to object classes are jointly optimized within the spatial resolution provided by the array.

To realize the task-optimized unitary transformation $\hat{U}$, we maximize the distinguishability of the object-dependent wavefields using gradient descent. For the unrestricted-mode measurement, the objective function combines classification, mutual information, and energy-conservation terms over the full detector-plane intensity distribution, whereas the restricted-mode objective is based on the individual channel responses of the 6×6 detector array, including terms for correct-class response, competing-class separation, and assignment balancing (see Methods). For both measurement configurations, stochastic translational shifts are applied to improve robustness against object displacement.

The numerical validation employs a set of object wavefields defined by binary intensity profiles and distinct spatial phase distributions with spherical or cylindrical curvatures (Fig. 2a). To quantify the single-photon classification performance, we define

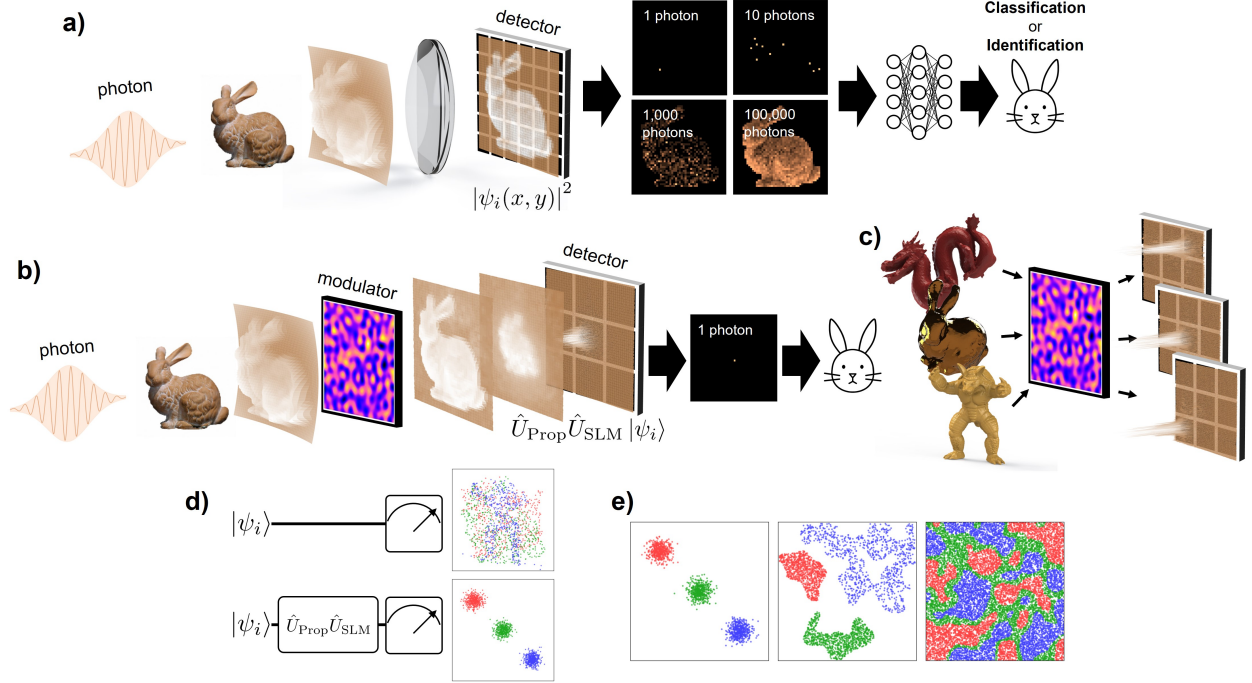


FIG. 1. **Schematics of single-photon object identification.** **a**, Classical intensity-based object identification method. Identification relies on post-processing the spatial intensity distribution accumulated from a large photon flux. **b**, Quantum state discrimination based object identification method. A task-optimized unitary transformation maps object-encoded wavefields into distinct modes. **c**, Each wavefield is focused to a specific location depending on the object under a common unitary transformation. **d**, Comparison of detection modes without and with the unitary transformation, illustrating enhanced spatial separation at the detector plane. **e**, Comparison of restricted-mode measurement and unrestricted-mode measurement, illustrating the generalization from a grid-constrained effective measurement basis to a spatially adaptive basis optimized for state separability.

the classification accuracy as the energy-weighted detection ratio,

$$\eta_i = \frac{\sum_{\text{pixels}} (M_i \cdot I_{i,d})}{\sum_{\text{pixels}} I_{i,in}} \quad (3)$$

where $I_{i,d}$ is the detector-plane intensity distribution for the i -th object, M_i represents the optimized spatial mask for the i -th object, and $I_{i,in}$ is the total incident intensity on the SLM. The classification accuracy, evaluated over simulated photon detection events, is summarized by the confusion matrix in Fig. 2b, confirming that the optimized transformation maps the spatially overlapping input states to spatially distinct detection modes with high fidelity.

To quantify the performance of the proposed framework, we calculate the classical upper bound of single-photon identification accuracy using the maximum-likelihood method [18]. Given prior knowledge of the intensity distributions, each photon position is sampled from the spatial probability distribution of the candidate objects, and the pho-

ton is classified according to the object whose distribution yields the highest probability for that position. Repeating this sampling-and-classification procedure over independent single-photon events in a Monte Carlo simulation yields the theoretical identification accuracy (see Methods), which reaches only 31.9% for the same objects, substantially below the accuracy achieved by our approach.

To evaluate a near-optimal quantum bound, we calculate the theoretical identification accuracy given by a Pretty Good Measurement (PGM), a standard framework in quantum information theory for near-optimal state discrimination. For the set of states $|\psi_i\rangle$ with equal prior probabilities, the PGM constructs measurement operators $\Pi_i = N_{\text{obj}}^{-1} \rho^{-1/2} |\psi_i\rangle \langle \psi_i| \rho^{-1/2}$, where N_{obj} denotes the number of object classes and $\rho = N_{\text{obj}}^{-1} \sum_i |\psi_i\rangle \langle \psi_i|$ represents the ensemble density operator. Applying the PGM to the simulated object states yields a theoretical single-photon identification accuracy of 98.9%. Although the PGM approaches the optimal quantum measurement, realizing the corresponding

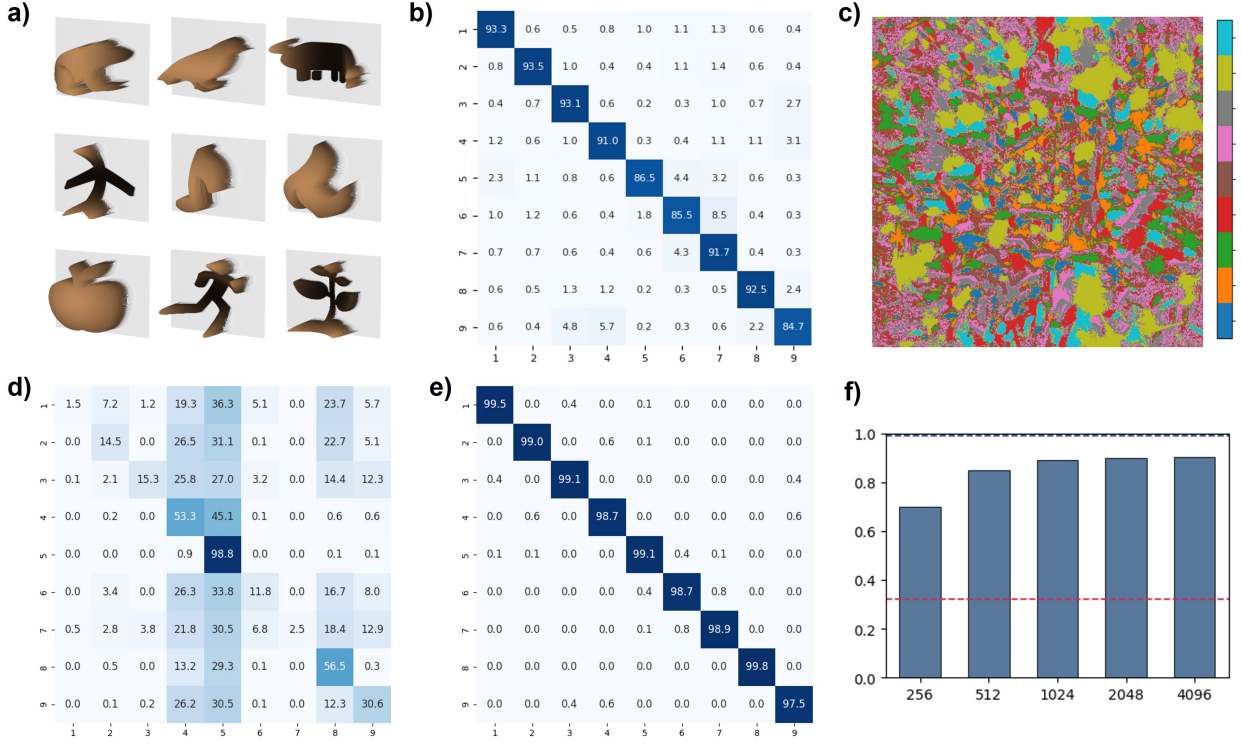


FIG. 2. **Numerical validation of single-photon state discrimination.** **a**, Object wavefields with spatial phase distributions, where the phase is visualized as depth. **b**, Confusion matrix for the unrestricted-mode measurement simulation. **c**, Optimized detection-region map at the detector plane for unrestricted-mode measurement. **d**, Confusion matrix representing the classical upper bound derived from maximum-likelihood estimation. **e**, Confusion matrix representing the near-optimal quantum bound calculated using the PGM. **f**, Scaling analysis of average classification accuracy as a function of the spatial resolution of the SLM and detector.

measurement experimentally would require an arbitrary unitary operation over the full spatial Hilbert space, which remains impractical [20]. Our implementation, limited to phase-only modulation and finite device resolution, consequently performs below the PGM limit. Increasing the spatial resolution of both the SLM and detector from 256×256 to 4096×4096 pixels nonetheless raises the average classification accuracy from 70.0% to 90.2% (Fig. 2f). Since the maximum mutual information of a spatial measurement grows logarithmically with the number of independent modes [21], the classification fidelity is physically constrained by the accessible spatial degrees of freedom [20, 22].

EXPERIMENTS

Experimental validation of the single-photon identification framework employs an optical setup comprising a collimated laser, a phase-only SLM, and a camera detector (Fig. 3a). We first implement the

unrestricted-mode measurement using the same set of objects as in the numerical simulation. In experimental implementation, the propagation distance between the SLM and the camera is set significantly longer than in the numerical simulation. The configuration reflects a practical constraint of reflective SLM operation where a non-zero incident angle is requisite to separate input and output beams, yet the angle must remain minimal to suppress phase distortion of the SLM [23]. Consequently, the optical path extends to achieve sufficient spatial separation between the incident and modulated beams within the small angle. Moreover, since the camera covers only a subsection of the full area of the SLM, the optimized detector map is cropped accordingly.

The constraints, the extended propagation and reduced detection area, limit the accessible spatial degrees of freedom in the Hilbert space. In addition, extrinsic system-dependent losses such as the SLM fill factor and the detector quantum efficiency cause the absolute classification accuracy to become system dependent. To remove the dependence on ex-

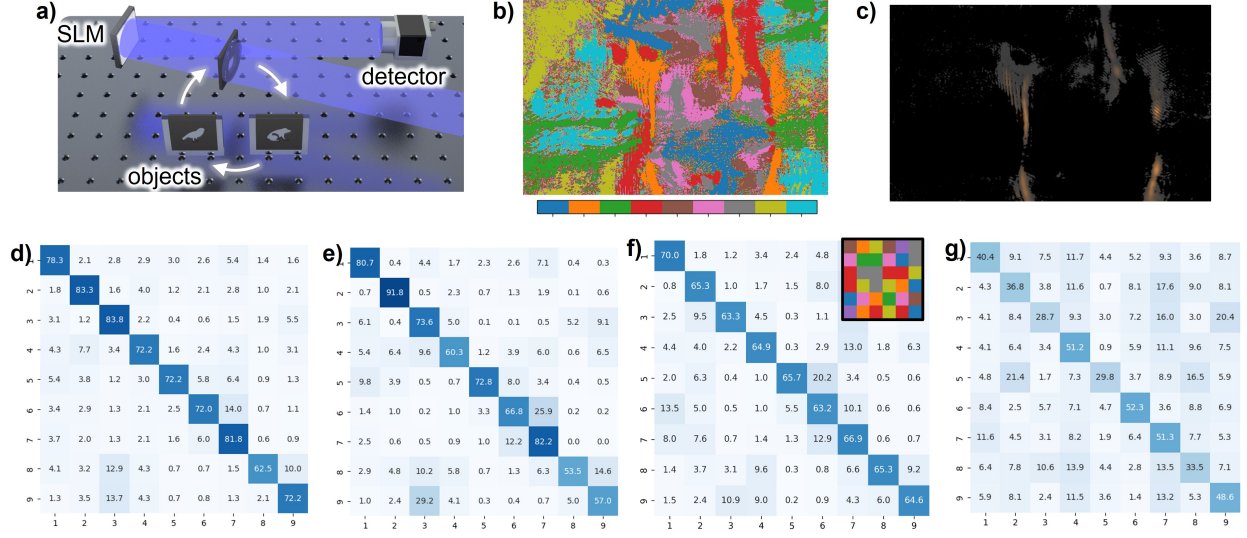


FIG. 3. **Experimental demonstration of single-photon object identification.** **a**, Schematic of the experimental setup. **b**, Optimized detection-region map for unrestricted-mode measurement under the system constraints. **c**, Representative measured intensity distribution overlaid with the corresponding detection mask. **d**, Simulated confusion matrix for unrestricted-mode measurement under the system constraints. **e**, Experimental confusion matrix for unrestricted-mode measurement. **f**, Simulated confusion matrix for restricted-mode measurement using the 6×6 detector array. The inset shows the optimized detection-region map. **g**, Experimental confusion matrix for restricted-mode measurement using idler-triggered single-photon detection.

trinsic losses, we introduce the conditional classification accuracy, defined as the ratio of correctly identified detection events to the total detected events, with normalization performed over detected events instead of the incident field. Under the constraints, the numerically achievable conditional accuracy for the experimental configuration is limited to 75.6% (Fig. 3d). Compared with the ideal configuration in Fig. 2c, the optimized detection map becomes more spatially clustered under the experimental geometry (Fig. 3b).

Measured intensity distributions obtained under attenuated coherent illumination demonstrate localized wavefields within the optimized spatial regions (Fig. 3c). Given the linearity of the optical system, the normalized intensity distribution is equivalent to the single-photon detection probability distribution. Despite these experimental constraints, the experiment achieves an average accuracy of 71.0% conditioned on photon detection, thereby isolating discrimination performance from system-dependent detection losses (Fig. 3e, see Methods). The value closely aligns with the predicted accuracy, supporting the validity of the single-photon object identification.

Because camera-based detection requires full-frame readout, single-photon detector arrays can support faster event-wise classification and, when

combined with high-speed readout electronics, provide a route toward nanosecond-scale recognition. To demonstrate this detection scheme at the single-photon level, we implement the restricted-mode measurement in a heralded-photon experiment based on spontaneous parametric down-conversion (SPDC) and a silicon photomultiplier (SiPM) array.

The detector plane is sampled using a 6×6 SiPM array, and the SLM phase and the detection-region map are jointly optimized. Numerical simulation using the geometry of the SiPM experiment and the optimized detection-region map yields an average identification accuracy of 65.5% (Fig. 3f). Despite joint optimization, the restricted-mode measurement remains less accurate than the unrestricted-mode measurement because integration over the 6×6 detector channels removes fine-scale spatial information available to the unrestricted-mode measurement.

With idler-triggered acquisition, the restricted-mode experiment achieves an average identification accuracy of 41.4% (Fig. 3g). Although lower than the simulated value of 65.5%, the correct-class response remains the largest for every object class. Importantly, the corresponding simulation shows that 94% of the optical energy incident on the SLM reaches the active SiPM area, indicating that the optimized measurement relies predominantly on spa-

tial redistribution of the single-photon wavefunction rather than on directing a substantial fraction of the photon probability outside the detector. The results demonstrate that an optimized unitary transformation can map object-dependent single-photon wavefields onto distinct detection channels of a SiPM array.

CONCLUSION

In this work, we formulate single-photon object identification as quantum state discrimination in a continuous spatial Hilbert space and demonstrate that an SLM-programmed unitary transformation can map object-induced wavefunctions onto spatially localized detection states. Our method surpasses the classical single-photon limit and approaches the quantum discrimination bound as the spatial resolution of the optical system increases. The framework aligns with developments in quantum computational sensing and optical neural architectures, while extending these approaches to states directly encoded by physical objects rather than digitally encoded patterns.

The framework can be scaled to support higher accuracy and more complex discrimination tasks within the single-photon regime through increased spatial resolution and a larger accessible transformation space. Multilayer architectures, as explored in diffractive deep neural networks [24], provide a natural route to enlarging this transformation space and can be implemented through cascaded phase masks, metasurface-based assemblies, or integrated multiplane modulators. Such architectures may extend the framework to multi-object recognition, low-flux metrology, and ultrafast classification using high-speed single-photon detectors, potentially enabling dynamic anomaly detection and transient analysis. More fundamentally, the framework recasts object recognition from a post-detection inference task into a quantum measurement process, with the location of wavefunction collapse directly specifying an object label.

METHODS

Measurement model

The interaction of an initial single-photon state $|\psi\rangle = \int dx \psi(x) \hat{a}^\dagger(x)|0\rangle$ with an object of class i , defined by a complex modulation $O_i(x)$, yields the object-encoded state $|\psi_i\rangle = \int dx O_i(x) \psi(x) \hat{a}^\dagger(x)|0\rangle$.

The measurement apparatus, comprising the phase-only SLM and free-space propagation, is modeled as a unified unitary operator $\hat{U} = \hat{U}_{\text{prop}} \hat{U}_{\text{SLM}}$. The SLM imposes a phase modulation $\hat{U}_{\text{SLM}} = \exp[i\phi_{\text{SLM}}(\hat{x})]$, and $\hat{U}_{\text{prop}}$ denotes the Fresnel propagation operator. To classify the object as class j , we define a detection region Ω_j on the detector plane. The measurement corresponds to a projection onto the subspace associated with this region, described by the Positive Operator-Valued Measure (POVM) element

$$\hat{\Pi}_j^U = \int_{\Omega_j} \hat{U}^\dagger |x\rangle \langle x| \hat{U} dx. \quad (4)$$

The conditional probability of identifying class j given an input object i is

$$P(j|i) = \langle \psi_i | \hat{\Pi}_j^U | \psi_i \rangle, \quad (5)$$

which constitutes the element C_{ij} of the confusion matrix. The optimization of $\phi_{\text{SLM}}(x)$ determines the unitary $\hat{U}$ that maximizes the distinguishability of the states $|\psi_i\rangle$ by mapping them into spatially disjoint regions Ω_j .

Upper bound on classical measurement

To quantify the upper bound of single-photon identification accuracy of the image-based measurement, we employ a maximum-likelihood estimator. The probability of detecting a photon at position x for object i is given by the normalized intensity distribution $P_i(x) = I_i(x) / \int I_i(x') dx'$. To accommodate spatial misalignment, the classifier utilizes a spatially-averaged probability distribution $\bar{P}_i(x)$, derived by averaging intensity maps over the range of possible shifts. We model the positional uncertainty as a uniform probability distribution over a spatial shift range Δ . Accordingly, the reference distribution is defined as $\bar{P}_i(x) = \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} P_i(x - \delta) d\delta$, which mitigates sensitivity to jitter and establishes robustness against positional alignment errors.

As a result, the classification is performed by selecting the class j that maximizes the likelihood of the observed single-photon event x ,

$$\hat{i} = \arg \max_j \bar{P}_j(x). \quad (6)$$

The accuracy is numerically estimated via Monte Carlo simulation with T trials. In each trial, a position x is sampled from the intensity distribution of object i subject to a random spatial shift. And the

confusion matrix element C_{ij} is estimated by averaging the outcomes over T trials performed for the true object i ,

$$C_{ij} = \frac{1}{T} \sum_{t=1}^T \delta_{\hat{i}_t, j}, \quad (7)$$

where δ is the Kronecker delta, and $\hat{i}_t$ is the predicted label in the t -th trial. C_{ij} represents the probability that object i is identified as object j .

Pretty Good Measurement

Optimal quantum state discrimination seeks a POVM that achieves the Helstrom limit, which represents the maximum attainable success probability for distinguishing non-orthogonal quantum states. However, identifying the exact Helstrom measurement is impractical in high-dimensional Hilbert spaces, and the Pretty Good Measurement (PGM) is widely adopted as a practical alternative. For an ensemble of object-encoded states $|\psi_i\rangle$ with prior probabilities p_i , the PGM consists of POVM elements defined as

$$\hat{\Pi}_j = p_j \rho^{-1/2} |\psi_j\rangle \langle \psi_j| \rho^{-1/2}, \quad (8)$$

where $\rho = \sum_k p_k |\psi_k\rangle \langle \psi_k|$ is the ensemble density operator. The probability of classifying state i as class j corresponds to the confusion-matrix element C_{ij} , calculated as $C_{ij} = \text{Tr}(\hat{\Pi}_j |\psi_i\rangle \langle \psi_i|)$.

Direct computation of $\rho^{-1/2}$ in the high-dimensional continuous spatial Hilbert space is numerically inaccessible. We therefore evaluate the confusion matrix within the N_{obj} -dimensional subspace spanned by the non-orthogonal object states. We define the overlap matrix $S_{ij} = \langle \psi_i | \psi_j \rangle$ and the weighted Gram matrix $G = \text{diag}(\sqrt{p}) S \text{diag}(\sqrt{p})$. The transition amplitude matrix M , which satisfies $C_{ij} = |M_{ij}|^2$, is derived via the pseudo-inverse of the density operator as

$$M = S \cdot \text{diag}(\sqrt{p}) \cdot G^{-1/2}. \quad (9)$$

The formulation confines the calculation to $N_{\text{obj}} \times N_{\text{obj}}$ matrix operations, enabling the efficient determination of the quantum identification bound.

Details of experimental setup

The experimental apparatus is constrained by the physical geometry of the reflective SLM, which requires a non-zero incident angle to separate input

and output beams. Consequently, the propagation distance is extended to suppress phase errors, and the detector area covers only approximately 32% of the total SLM active area. In the experiment, the target label map for the unrestricted-mode measurement is computationally cropped to match these dimensions.

The optical setup utilizes a coherent light source operating at a wavelength of 638 nm. Active wavefront modulation is performed by a phase-only SLM featuring a resolution of 1024×1024 pixels with a pixel pitch of $17.4 \mu\text{m}$. The propagation distances are configured such that the distance from the object to the SLM is 125 mm, while the distance from the SLM to the detector plane is set to 927 mm.

To quantify performance under these constraints, we evaluate the conditional classification accuracy using the binary target mask M_i and the detected intensity distribution $I_{i,d}$ for object i . We define the conditional classification accuracy as the ratio of correctly classified energy to the total energy detected over the entire detector region, given by

$$\eta_{\text{con}} = \frac{1}{N_{\text{obj}}} \sum_i \frac{\sum_{\text{pixels}} (M_i \cdot I_{i,d})}{\sum_{\text{pixels}} I_{i,d}}. \quad (10)$$

The conditional classification accuracy serves as the primary benchmark for the experimental results. Note that we adopt a detected-energy normalization for the experimental accuracy, contrasting with the input-energy normalization used in simulations. This approach isolates the state discrimination fidelity from system-dependent detection losses by normalizing to the total detected count, thereby defining a robust metric invariant to external factors such as SLM fill factor, optical loss, and detector quantum efficiency.

For the SPDC-SiPM experiment, degenerate signal and idler photon pairs near 810 nm are generated by type-II SPDC and separated by a polarizing beamsplitter. Detection of the idler photons by a single-photon counting module provides the trigger for SiPM acquisition of the signal photons. The signal photons propagate through a polarization-maintaining fiber and illuminate the same set of nine amplitude-phase object pairs and phase-only SLM used in the camera-based experiment. The distances from the amplitude object to the phase object, from the phase object to the SLM, and from the SLM to the SiPM detector are 10, 150, and 800 mm, respectively.

The transformed photons are detected using an 8×8 SiPM array. Because the aperture of the band-pass filter limits the illuminated detector area, only

the central 6×6 channels are used for analysis. The coincidence peak is measured independently for each channel, and photon-detection events are selected using a channel-specific 2-ns window. For each object, idler-triggered events are accumulated for 180 s. The channel counts are summed over the detector channels assigned to each output class and normalized over the nine classes to construct the experimental confusion matrix.

Loss function of phase optimization

We optimize the unitary transformation $\hat{U}$ using gradient descent under two distinct measurement configurations, restricted-mode measurement and unrestricted-mode measurement. In both modes, to ensure robustness against spatial misalignment, the optimization routine incorporates stochastic translational augmentation. At each iteration, a random spatial shift δ is sampled from the uniform distribution over $[-\Delta/2, \Delta/2]$ and applied to the input wavefield.

In the restricted-mode measurement, the detector-plane intensity is sampled by the central 6×6 detector channels. Because an objective defined over the full detector-plane intensity distribution need not track the winner-take-all decision applied to these 36 channel responses, we use a differentiable approximation of the channel-wise decision rule. Let $R_n(\mathbf{r}) = I_n(\mathbf{r})/E_n^{\text{in}}$ denote the incident-energy-normalized response of channel $\mathbf{r}$ for object class n . Its soft assignment to output class k and the resulting response matrix are

$$\begin{aligned} q_k(\mathbf{r}) &= \frac{[R_k(\mathbf{r}) + \varepsilon]^{1/\tau}}{\sum_j [R_j(\mathbf{r}) + \varepsilon]^{1/\tau}}, \\ \tilde{C}_{nk} &= \sum_{\mathbf{r} \in \Omega} R_n(\mathbf{r}) q_k(\mathbf{r}), \end{aligned} \quad (11)$$

where $\tau > 0$ is the assignment temperature and ε is a small numerical constant. As $\tau \rightarrow 0$, $q_k(\mathbf{r})$ approaches the hard winner-take-all assignment. We minimize

$$\mathcal{L}_{\text{restricted}} = \kappa \mathcal{L}_{\text{diag}} + \lambda \mathcal{L}_{\text{margin}} + \mu \mathcal{L}_{\text{occ}} + \nu \mathcal{L}_{\text{worst}}, \quad (12)$$

where κ , λ , μ , and ν are hyperparameters, with

$$\begin{aligned} \mathcal{L}_{\text{diag}} &= -\frac{1}{N_{\text{obj}}} \sum_n \log(\tilde{C}_{nn} + \varepsilon), \\ \mathcal{L}_{\text{margin}} &= \frac{1}{N_{\text{obj}}} \sum_n \left[\max(0, M_n + m - \tilde{C}_{nn}) \right]^2, \\ M_n &= \max_{k \neq n} \tilde{C}_{nk}, \\ \mathcal{L}_{\text{occ}} &= \frac{1}{N_{\text{obj}}} \sum_k \left(\frac{1}{|\Omega|} \sum_{\mathbf{r} \in \Omega} q_k(\mathbf{r}) - \frac{1}{N_{\text{obj}}} \right)^2, \\ \mathcal{L}_{\text{worst}} &= \frac{1}{K} \sum_{n \in \mathcal{W}_K} \left[\max(0, \tau_w - \tilde{C}_{nn}) \right]^2. \end{aligned} \quad (13)$$

Here, Ω denotes the active detector channels, m is the required margin, τ_w is the target response, and $\mathcal{W}_K$ indexes the K largest class-wise shortfalls. The four terms respectively increase the correct-class response, separate it from the strongest competing response, prevent assignment collapse, and penalize the worst-performing classes.

In the unrestricted-mode measurement, both the SLM phase pattern and the effective detection basis are jointly optimized. We employ a composite loss function $\mathcal{L}_{\text{unrestricted}} = \alpha \mathcal{L}_C + \beta \mathcal{L}_M + \gamma \mathcal{L}_E$, where α, β, γ are hyperparameters. To enhance separability among the detector-plane intensity distributions, we utilize a cross-entropy loss based on their spatial intensity overlap, so-called classification loss ($\mathcal{L}_C$). Defining the detector-plane state vector as $|\phi_i\rangle = \sum_x \sqrt{I_i(x)/E_i} |x\rangle$, the intensity overlap between classes i and j is defined as $S_{ij} = \langle \phi_i | \phi_j \rangle = \sum_x \sqrt{I_i(x)I_j(x)}/\sqrt{E_i E_j}$. Note that $I_i(x)$ denotes the intensity distribution at the detector plane, while E_i represents the total energy incident on the SLM. Finally, the classification loss is defined as

$$\mathcal{L}_C = -\log \left(\frac{\exp(S_{yy})}{\sum_j \exp(S_{yj})} \right), \quad (14)$$

where y denotes the true object label. The formulation maximizes the self-overlap S_{yy} while minimizing cross-correlations.

Simultaneously, we maximize the spatial distinguishability provided by the detection coordinate x by minimizing the weighted mutual-information loss ($\mathcal{L}_M$). Let $Q_i(x) = I_i(x)/I_{\text{tot}}(x)$ be the conditional probability that a photon detected at position x originates from object class i , where $I_{\text{tot}}(x) = \sum_k I_k(x)$ is the total intensity profile. The loss is given by the spatially weighted conditional entropy

$$\mathcal{L}_M = \frac{-\sum_x I_{\text{tot}}(x) \sum_i Q_i(x) \log Q_i(x)}{\sum_x I_{\text{tot}}(x)}, \quad (15)$$

which reduces the uncertainty of the object identity given a detection event at x , thereby encouraging the formation of spatially distinct detection modes. Finally, to penalize optical losses and enforce unitarity within the limited numerical aperture, we include an energy conservation loss ($\mathcal{L}_E$), defined as

$$\mathcal{L}_E = \sum_i \left(1 - \frac{E_i^{\text{det}}}{E_i} \right), \quad (16)$$

where $E_i^{\text{det}} = \sum_x I_i(x)$ is the total energy detected at the detector plane.

To verify the need for configuration-specific objectives, we exchanged the two loss functions while holding the experimental geometry and all other simulation and optimization parameters fixed. In the restricted-mode configuration, using the phase pattern optimized with $\mathcal{L}_{\text{unrestricted}}$ reduced classification accuracy by 7.7%. In the unrestricted-mode configuration, using the phase pattern optimized with $\mathcal{L}_{\text{restricted}}$ reduced accuracy by 13.3%. The results therefore justify adapting the loss function to the measurement configuration during phase optimization.

Theoretical identification accuracy

The identification accuracy is fundamentally bounded by the information capacity of the measurement apparatus, expressed as the mutual information between the object-encoded wavefront and the detected intensity distribution [22]. The information capacity is governed by two physical constraints: spatial resolution and propagation distance. First, the maximum extractable information scales logarithmically with the number of independent spatial modes; thus, higher resolution or larger active areas increase the achievable accuracy limit [21] (Fig. S1d). Second, the propagation distance z determines the discriminative power of the unitary transformation. At short distances, phase modulations introduced by the SLM do not generate distinguishable intensity patterns because the limited diffraction angles restrict their conversion into measurable spatial variations [25]. Conversely, excessive propagation suppresses high-spatial-frequency components, leading to flattened intensity distributions that reduce extractable information [26]. Consequently, the system's accuracy follows the distance-dependent variation of mutual information, serving as a physical upper bound for the measurement [20, 22] (Fig. S1c).

* These authors contributed equally to this work.

† Email: sheepvs5@gmail.com

- [1] Z. Li, Y. Dong, L. Shen, Y. Liu, Y. Pei, H. Yang, L. Zheng, and J. Ma, *Neurocomputing* **598**, 128102 (2024).
- [2] B. More and S. Bhosale, *Revue d'Intelligence Artificielle* **37**, 407 (2023).
- [3] A. Bakhtiarnia, Q. Zhang, and A. Iosifidis, *ACM Computing Surveys* **56**, 1 (2024).
- [4] P. P. Laissue, R. A. Alghamdi, P. Tomancak, E. G. Reynaud, and H. Shroff, *Nature Methods* **14**, 657 (2017).
- [5] K. L. Tosheva, Y. Yuan, P. Matos Pereira, S. Culley, and R. Henriques, *Journal of Physics D: Applied Physics* **53**, 163001 (2020).
- [6] C. Li, X. Qu, A. Gnanasambandam, O. A. Elgandy, J. Ma, and S. H. Chan, in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops* (2021) pp. 3976–3987.
- [7] Z. Zhang, Z. Wen, Y. Qiang, H. Dong, W. Dong, X. Li, X. Wang, and X. Ren, *arXiv preprint arXiv:2505.04376* (2025).
- [8] Z. Ballard, C. Brown, A. M. Madni, and A. Ozcan, *Nature Machine Intelligence* **3**, 556 (2021).
- [9] M. A. Neifeld, A. Ashok, and P. K. Baheti, *Journal of the Optical Society of America A* **24**, B25 (2007).
- [10] M. Tsang, R. Nair, and X.-M. Lu, *Physical Review X* **6**, 031033 (2016).
- [11] M. R. Grace and S. Guha, *Physical Review Letters* **129**, 180502 (2022).
- [12] S. A. Khan, S. Prabhu, L. G. Wright, and P. L. McMahon, *npj Quantum Information* **12**, 119 (2026), *arXiv:2507.15845 [quant-ph]*.
- [13] A. Villegas, M. H. M. Passos, S. F. Pereira, and J. P. Torres, *Physical Review A* **109**, 032617 (2024).
- [14] U. Zanforlin, C. Lupo, P. W. R. Connolly, P. Kok, G. S. Buller, and Z. Huang, *Nature Communications* **13**, 5373 (2022).
- [15] S. A. Wadood, K. R. Sethuraj, K. Liang, M. R. Grace, G. La Rue, S. Guha, and A. N. Vamivakas, *Optics Letters* **49**, 750 (2024).
- [16] S. Franke-Arnold and J. Jeffers, *The European Physical Journal D* **66**, 196 (2012).
- [17] M. A. Solís-Prosser, M. F. Fernandes, O. Jiménez, A. Delgado, and L. Neves, *Physical Review Letters* **118**, 100501 (2017).
- [18] T. Fischbacher and L. Sbaiz, *arXiv preprint arXiv:2008.05859* (2020).
- [19] S. A. Khan, S. Prabhu, L. G. Wright, and P. L. McMahon, *arXiv preprint arXiv:2507.16918* (2025).
- [20] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).
- [21] T. B. H. Tentrup, T. Hummel, T. A. W. Wolterink, R. Uppu, A. P. Mosk, and P. W. H. Pinkse, *Optics Express* **25**, 2826 (2017).
- [22] S. M. Barnett and S. Croke, *Advances in Optics and Photonics* **1**, 238 (2009).
- [23] Q. Niu, C. Wang, and H. Shi, *Review of Scientific Instruments* **90**, 055001 (2019).

- [24] X. Lin, Y. Rivenson, N. T. Yardimci, M. Vel, Y. Luo, M. Jarrahi, and A. Ozcan, *Science* **361**, 1004 (2018).
- [25] C. Kohler, F. Zhang, and W. Osten, *Applied Optics* **48**, 4003 (2009).
- [26] J. W. Goodman, *Introduction to Fourier Optics* (Roberts and Company Publishers, 2005).