ADVANCED QUANTUM TECHNOLOGIES
www.advquantumtech.com

Time-Efficient Object Recognition in Quantum Ghost Imaging

Chané Moodley, Alice Ruget, Jonathan Leach, and Andrew Forbes*

Acquiring information at the fastest possible rate is often desirable, particularly in quantum ghost imaging which suffers from slow reconstruction speeds. Many computationally intense deep-learning methods have been implemented in an effort to speed up image acquisition times by retrieving image information. Often over-looked, machine learning methods can offer the same, if not better, speed up in image acquisition time by an object recognition process. Four machine learning algorithms are implemented and trained on a uniquely generated, noised, and blurred dataset of numerical digits 1 through 9. Of the tested recognition algorithms, logistic regression shows a 10× speed up in image acquisition time with a 99% prediction accuracy. Additionally, this reduction in acquisition time is achieved without any image denoising or enhancement prior to recognition, thereby reducing training and implementation time, as well as the computational intensity of the approach. This method can be implemented in real-time, requiring only 1/10th of the measurements needed for a general solution, making it ideal for quantum imaging and recognition of light sensitive structure.

1. Introduction

Quantum ghost imaging is an image acquisition technique that makes use of the quantum correlations between two spatially separated photons to reconstruct an image of an object. Individually, each photon does not offer any image information on the object; however, the correlations between them allow for the reconstruction of an image[1–4] and have been demonstrated using quantum entanglement generated by spontaneous parametric downconversion (SPDC),[5,6] entanglement swapped photons,[7] and with symmetry-engineered quantum states.[8] Quantum ghost imaging was initially thought to produce higher resolution images that its classical counterpart; however, it has been shown that in both quantum and classical cases, images of almost identical quality are produced.[2,9] Advantageously, the use of quantum light allows for imaging at low light levels, demonstrating a signal to noise ratio and visibility higher than classical ghost imaging.[9,10] Particularly, quantum ghost imaging is useful for biological imaging applications where it is beneficial to reduce the risk of photo-damage to light sensitive matter.[11]

Quantum ghost imaging faces rather unsatisfactory imaging speeds; therefore, a common goal is to obtain object information within the shortest possible acquisition time.[12] Computationally, intensive attempts to improve imaging speed focused on data acquisition methods,[13,14] data processing methods using compressive sensing,[15–17] compressive sensing with generative adversarial networks,[18] and parallel computing.[19] Fast image reconstruction algorithms were also implemented as a means to speed up reconstruction, for example, the fast Walsh–Hadamard transform,[20] and by using logarithmic and exponential ghost-imaging reconstruction algorithms.[21] Deep learning has been used to increase image reconstruction quality,[22] to improve image quality by denoising mechanisms,[23–25] and recently to reconstruct images at impractical to measure image resolutions.[26] To speed up and enhance image reconstructions, deep convolution networks have been used as a self-supervised learning approach.[27–29] In cases where object discrimination is of primary importance, recognition algorithms have been used to decrease image reconstruction time,[30,31] decrease the sampling ratio,[32] and to achieve object identification prior to the imaging process.[33] While there have been many advancements to date, these methods are computationally intense, iterative, or require very deep neural network models often requiring training on GPUs.

In this work, we implemented four machine learning algorithms and tested their performance to achieve the most significant speed up in image acquisition times. The implemented algorithms consist of a support vector machine (SVM), logistic regression (LR), nearest neighbors (NN), and naïve Bayes (NB), which is a baseline classifier. We assessed the different machine learning algorithms for object recognition and established

C. Moodley, A. Forbes
Structured Light Laboratory
School of Physics
University of the Witwatersrand
Johannesburg 2000, South Africa
E-mail: chane13.m@gmail.com

A. Ruget, J. Leach
Institute of Photonics and Quantum Sciences
Heriot-Watt University
Edinburgh EH14 4AS, UK

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/qute.202200109

DOI: 10.1002/qute.202200109

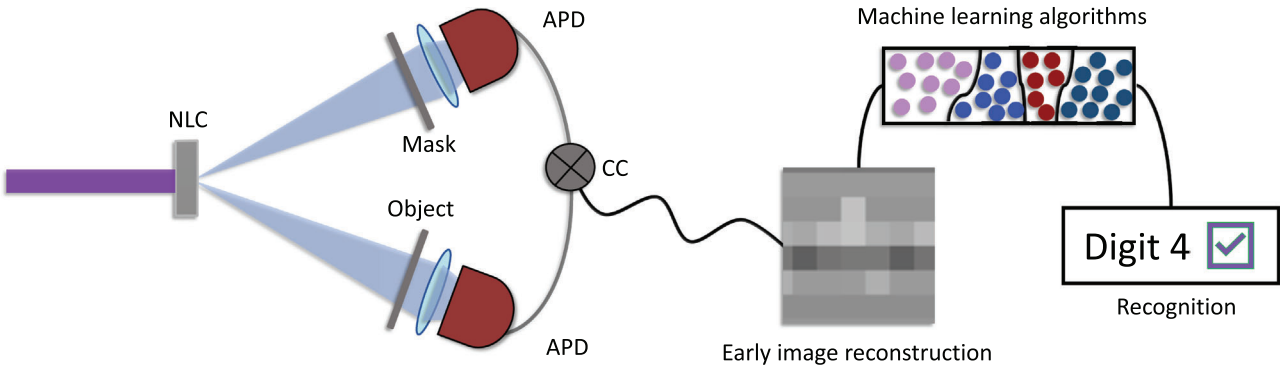

Figure 1. Conceptual sketch of a ghost imaging setup using machine learning algorithms for early object recognition. Entangled photons are generated at the non-linear crystal (NLC) and are spatially separated along two arms. One photon interacts with the object and is collected by a bucket detector using an avalanche-photo diode (APD). The other photon is collected by a spatially resolving detector consisting of a binary patterned mask and an APD. Each detector is connected to a coincidence counting (CC) device to perform coincidence measurements. The image is reconstructed by a linear combination of the patterned masks weighted by the coincidences. The reconstructed image, after each measurement, is passed through each of the four algorithms for object recognition.

the optimal experimental early stopping point. We predicted the confidence level of the identified object after each measurement for each algorithm and established a threshold at which we can stop the experiment without compromising the information contained within the image. We assessed each approach based on the predicted confidence level and determined which algorithms were successful based on the predicted confidence score and the level of uncertainty in the prediction. Importantly, these models are not computationally intense to train, are not iterative, are not deep neural networks, and are quick to implement. We start by describing the concept and strategy, followed by the implementation details, and we finally present our results, where we demonstrate an order of magnitude reduction in image acquisition time to recognize an object. We believe that this novel approach introduces intelligence into recognition of objects illuminated by quantum states of light, offering a new pathway to imaging light-sensitive structures at low photon fluxes.

2. Concept and Strategy

At the core of quantum ghost imaging lies the production of entangled photon pairs. The most efficient and commonly employed method to generate entangled photon pairs is that of spontaneous parametric downconversion (SPDC). It has been well established that SPDC leads to the creation of entangled photons and is therefore a hallmark of quantum.[34–41] This non-linear process, of optical non-linearity $\chi^{(2)}$, facilitates the decay of a pump photon into two daughter photons (termed the signal and idler photons). A pump beam, of angular frequency ω_p, when incident on a non-linear crystal produces two daughter photons of ω_s and ω_i, respectively. As the rules of conservation are followed, the angular frequencies of the signal and idler photons should add to the angular frequency of the pump photon that produced them:

$$\omega_p = \omega_s + \omega_i \tag{1}$$

Additionally, the linear momentum of the daughter photons must add to that of the pump photon, given by:

$$\mathbf{k}_p = \mathbf{k}_s + \mathbf{k}_i \tag{2}$$

where $\mathbf{k}_{p,s,i}$ are the wavevectors of the pump, signal, and idler photons, respectively. The SPDC process is said to be degenerate if the down-converted photon pair has the same wavelength and non-degenerate if the wavelengths of the photon pair are different. SPDC, additionally, consists of two geometric types: these are collinear and non-collinear. In the collinear geometry, the output field wavevectors propagate in the same direction as the pump field. Conversely, in the non-collinear SPDC geometric case, the output fields propagate off-axis with respect to the pump field; however, their trajectories are equal although on opposite sides as illustrated in **Figure 1**. A good review in ref. [42] covers the theoretical derivation of SPDC.

Figure 1 illustrates our quantum ghost imaging setup where shallow machine learning algorithms (as opposed to computationally intense deep learning algorithms) were implemented to achieve early object identification. Ghost imaging occurs as a result of the second order correlations that exist between two fields of light. In quantum ghost imaging, these correlations arise from the entanglement of signal and idler photons generated by SPDC. The photons are entangled in both position and momentum and are produced using a non-linear crystal (NLC). The entangled photons are spatially separated into two independent paths, one to illuminate the object (object arm) and the other which is collected by a spatially resolving detector (reference arm). Neither path contains enough information to independently reveal the information contained within the object. It is only possible to retrieve the object information by using the measured correlations between the photons in each path. Our spatially resolving detector is accomplished by displaying a series of binary intensity patterns (masks) on a spatial light modulator (SLM) and collecting the projection using a bucket detector.

Initially, ghost imaging experiments accomplished spatially resolved detection by a physically moving single-pixel detector,

which scanned a transverse plane.[5] Advancements in this field consisted of using ultra-sensitive cameras to avoid a system consisting of physically moving parts.[43] Ultra-sensitive cameras are, however, expensive and have limited spectral sensitivity. An alternative uses a series of patterned masks and a bucket detector to spatially resolve what is detected. For a general image solution, one requires N^2 orthogonal masks to fully measure an $N \times N$ pixel image,[44] which is a lengthy process. Importantly, the patterns must constitute a complete basis to acquire a complete reconstructed image. The measured correlations then provide information on the similarity (or overlap) between the object and each patterned mask. The image is reconstructed as a linear combination of all masks weighted by the measured correlations. Similar to single-pixel scanning systems, scanning through a series of masks results in a lengthy image reconstruction process and is an area of great interest within the ghost imaging community.[22,27,45,46] The goal within the quantum ghost imaging community is to develop quantum ghost imaging to its most efficient state before utilizing all these advancements to image light sensitive structures with the most efficient and economical use of photons possible. The work done in this study is a step toward real-time, efficient quantum ghost imaging of light-sensitive structures.

The necessity of scanning through a series of mask patterns means that image acquisition times extend to lengthy periods and that the required time scales with the resolution of the image.[31] To avoid the long acquisition times, we trained machine learning algorithms to achieve early object recognition so as to establish an early experimental stopping point. We used digits as objects to build on the extensive machine learning digit research already done, and trained four machine learning algorithms to perform digit recognition at early stages of the reconstruction. In machine learning, classification is the process of identifying, understanding, and grouping together ideas and objects into pre-determined categories or sub-categories. Recognition algorithms are used to predict the likelihood that the data is part of a certain class and assigns it to that class with a certain confidence in the prediction. Classification is therefore a form of pattern recognition where the machine identifies patterns in the data that cannot be seen by the human eye by extracting features relating to all statistical moments. Early recognition in our quantum ghost imaging optical setup was achieved by four different machine learning algorithms trained on a unique dataset detailed in the implementation section below.

Support vector machine (SVM) is a supervised machine learning algorithm used in both classification and regression.[47] The SVM plots each data point in n-dimensional space, where n is the number of classes or categories (sometimes also referred to as features). Recognition is done by finding hyper-planes that differentiate each category or feature, that is, similar data points are grouped together by hyper-planes that segment the dataset. Logistic regression (LR) is an algorithm that spans many applications as it is straightforward and interpretable.[48] If there are n different classes in a dataset, to use logistic regression as a classifier, we first build a model where we consider data belonging to one class as positive and negative to all other classes. This reduces the classification to a simple true or false problem. This is repeated for all n classes within the dataset, that is, we build n different models one for each category within the dataset. To fully implement such a model, we then make use of the one-vs-all classification technique by training one logistic regression model for each digit. After training the one-vs-all classifier, we can now use it to predict the digit contained in a given image.

The K-nearest neighbors (NN) algorithm is known for solving both classification and regression problems.[49] The algorithm stores all available cases and classifies new cases based on a measure of similarity between what is available. NN is known as an instance-based learning algorithm where the function is approximated and all computation occurs at classification; therefore, training is fast yet the predictions are slower. The fourth and final algorithm we implemented is the complement naive Bayes (CNB) algorithm.[50] CNB is an adaptation of the standard multinomial naive Bayes algorithm. It is faster than other classification algorithms although known to not be as powerful. NB is based on Bayes theorem and is often used as a baseline in comparison to other classifiers; here, we implemented the algorithm as a baseline classifier and compared the results with the above classifiers, which are more complex to implement and train, yet are not computationally intensive.

We implemented the above classifiers to establish an optimal early stopping point. The classifiers were trained on our own generated noised and blurred dataset (which we detail in the implementation section to follow) in a supervised manner. The labels for this dataset were the corresponding numerical values. It is important to note that the predicted probability, or the confidence, is not the same as the model accuracy. A 99% model accuracy implies that given 100 images, the model will, on average, accurately predict the category of 99 of these with a confidence that varies largely between 50% and 100%. In our approach, we predicted the category (class) with the highest probability and checked to see if it corresponds to the correct digit, while also considering any close lying probabilities of the other classes for accurate object recognition. For such a classification problem, with categories specific to each digit, a predicted probability greater than 50% for a certain category is traditionally accepted (the sum of the probability of the remaining classes is less than half of the total probability). When assessing the stopping point of the approach, we chose to impose a stricter criteria with a probability greater than 75% to stop the experiment. This is a mid-line between acceptable (50%) and perfect (100%). Importantly, machine learning models are shallow learning models, opposed to deep learning models. They can be efficiently implemented, with most programming languages requiring only 2–10 lines of code. Implementation times and training times are drastically reduced in comparison to implementing deep learning models that require copious amounts of code and lengthy implementation and training times. By implementing machine learning algorithms to reduce image reconstruction time, both implementation and training times are reduced, therefore, adding to the efficiency of the method.

3. Implementation Details

Figure 2a shows a schematic representation of the optical setup implemented for this work. A vertically polarized laser (wavelength of $\lambda = 405$ nm) was used to pump a type 1 PPKTP nonlinear crystal (NLC); the temperature was set to obtain co-linear emission of entangled degenerate photons at wavelengths of $\lambda =$

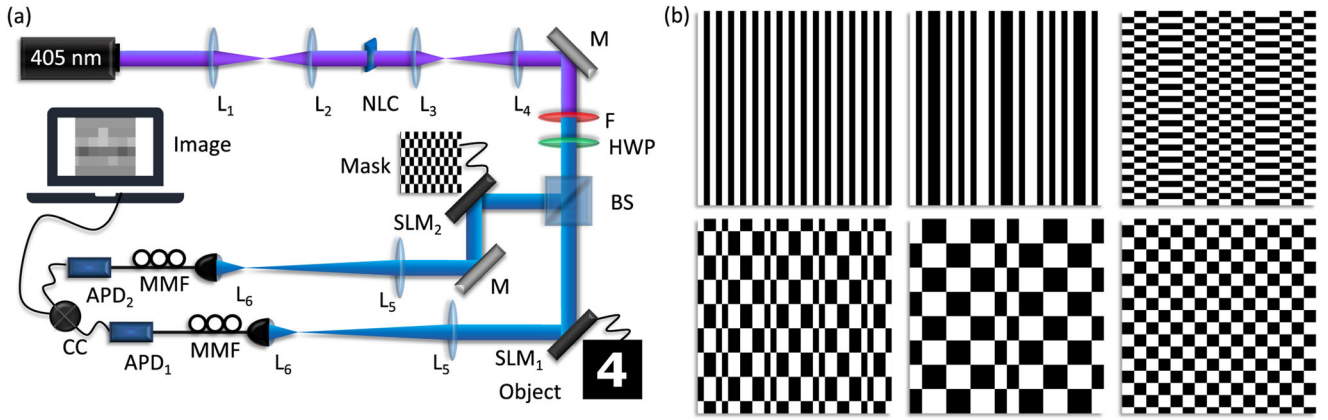

Figure 2. Schematic of the implemented quantum optical optical setup. a) Entangled photons are produced at the non-linear crystal (NLC). A 50:50 beamsplitter (BS) is used to spatially separate the entangled signal and idler photons. Each photon impinges on a SLM displaying either the object or patterned mask. The photons are collected by coupling each beam to a multi-mode fiber (MMF) connected to an APD; the photons that are detected in coincidence are counted by a coincidence counting device (CC). L_i are lenses, M are mirrors, F is a bandpass filter centered at $\lambda = 810$ nm, and HWP is a half-wave plate to rotate the polarisation of the light for optimal modulation by the SLMs. b) Examples of 32 × 32 pixel Walsh–Hadamard patterned masks used to reconstruct images of the digital objects.

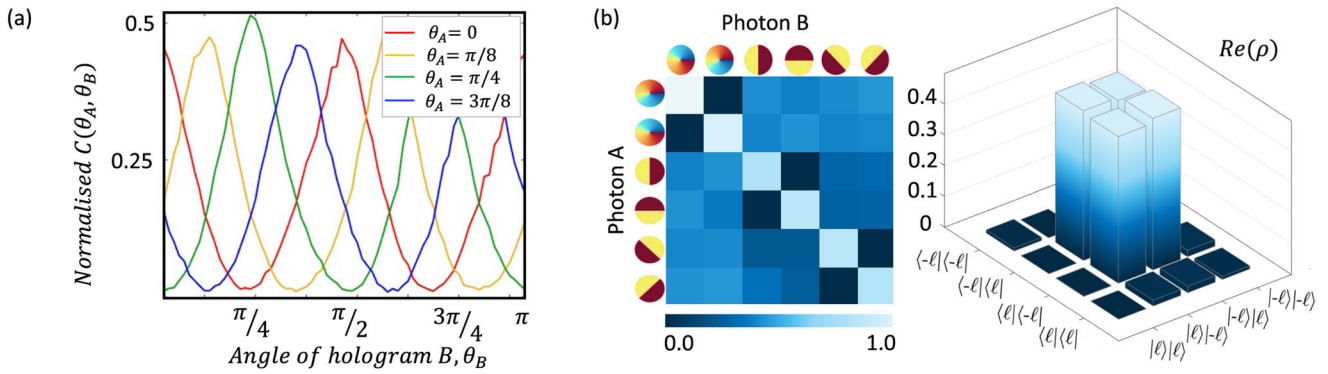

Figure 3. Measurements confirming the presence of quantum entanglement. a) Violation of the Bell inequality. The normalized coincidence counts as a function of relative orientation of the superpositions of orbital angular momentum (OAM) states for the entangled state $|\psi\rangle_2$. b) A two-qubit quantum state tomography is shown (left), where the projected state of each photon in the entangled pair is indicated by its phase map. The color of each box represents the normalized coincidence counts for a given set of projections on the two-photon state. Using the tomographic data, the density matrix is computed, and its real component is shown (right).

810 nm each by a SPDC process. A bandpass filter, centered at λ = 810 nm (FWHM = 10 nm), was used to filter out any non-converted pump photons. The NLC converted the SPDC photons to a horizontal polarization; a halfwave plate (HWP) was introduced to ensure the light was horizontally polarized for optimal modulation by the SLMs. A 50:50 beamsplitter was used to spatially separate the degenerate entangled photons. One of the entangled photon pair was directed to the SLM displaying holograms of our chosen objects to be imaged, while the other photon from the entangled pair impinged on the second SLM displaying holograms of our mask patterns. A blazed grating was added to each hologram to separate the first diffracted order from that of the zeroth order which is unmodulated. The modulated light from the first diffracted order was then coupled into multimode optical fibers (MMF) connected to avalanche photodiodes (APDs) for single photon-detection. Coincidences were measured by a photon coincidence counting device (CC). The image was reconstructed by a linear combination of the mask patterns in the chosen basis, weighted by the measured coincidences. Our holograms (objects and masks) were created using MATLAB, for data acquisition we used LabView, and all machine learning algorithms were implemented in JupyterLab using Python.

Prior to running our ghost imaging measurements, we routinely test for and confirm the presence of entanglement in controlled superposition states of orbital angular momentum (OAM). By using appropriate holograms displayed on the SLMs in each arm (shown in Figure 2a), we can generate two-photon states entangled within a 2D OAM subspace. For these test measurements, we coupled the photons from the first diffracted order into single-mode fibres of core size 5 μm. Specific to this work, we tested for the entangled nature of OAM states generated by SPDC through the violation of a suitable Clauser–Horne– Shimony– Holt (CHSH) Bell inequality, the results of which are shown in **Figure 3**a. We conducted the experiment similarly to that conducted in ref. [51] and display the normalized

coincidence fringes for the entangled state $|\psi\rangle_2$. We found a Bell parameter of $S = 2.59 \pm 0.08$ for the $l = \pm 2$ OAM subspace, therefore successfully violating the Bell inequality. To confirm entanglement, we performed a full quantum state tomography (QST) for the $l = \pm 1, p = 0$ subspace in the Laguerre–Gaussian basis, the results of which are shown in Figure 3b. A two-qubit QST is shown in Figure 3b (left), where the projected state of each photon in the entangled pair is indicated by the phase map. Using the tomographic data, the density matrix is computed, and its real component is also shown in Figure 3b (right). For these measurements, we report a fidelity of $F = 0.969$, a concurrence of $C = 0.919$, and a linear entropy of $S_L = 0.0518$. Additionally, our quantum contrast was $Q = 20.05 \pm 0.52$, therefore, confirming entanglement.

For our ghost imaging measurements, we varied the mask patterns within the experiment; examples of these mask patterns are shown in Figure 2b. We looked at binary sets of Walsh–Hadamard patterned masks. Our Walsh–Hadamard masks are generated by extracting the Walsh functions from the native MATLAB function for Hadamard matrices where N is the order of the Hadamard matrix. By performing the outer product between columns of a Hadamard matrix of order N, the result is a complete set of N^2 Walsh–Hadamard masks of $N \times N$ pixel resolution. Choosing the order of the Hadamard matrix will therefore result in the corresponding mask resolution; here, we used a Hadamard matrix of order $N = 32$. The mask resolution in turn determines how well the object will be resolved.[52] A higher resolution results in a larger number of basis elements and therefore an increased number of masks is needed to resolve the image. Increasing the resolution has direct consequences on the reconstruction time, the number of Walsh–Hadamard masks required to form a complete set scales as N^2, and so there is a trade-off between reconstruction time and image resolution. In this work, we used 32×32 pixel resolution Walsh–Hadamard masks and required 1024 measurements to acquire a complete image solution. We focused on the Walsh–Hadamard masks as they are known to converge faster than other mask types.[30] Importantly, in this work, the Walsh-Hadamard patterns are used to determine the position of the signal photon which has not interacted with the object. In our work, the Walsh–Hadamard patterns are not used for structured illumination of the object as is commonly done in classical ghost imaging.[20] Accordingly, we purposefully chose low-resolution target images as we know that the higher the resolution the larger the number of masks required to reconstruct a general image solution.[26] The aim of the study is accurate object recognition with a high confidence prediction achieved with as few a number of measurements as possible. We, therefore, favored a low-resolution dataset to reduce the number of measurements and implementation time; however, the approach is applicable to images of any resolution.

Prior to implementing and training our machine learning classifiers, we generated a dataset that would allow us to accurately identify the object early in the measurement process. The overlap in a ghost imaging experiment is proportional to the coincidences measured per mask and was simulated using binary sets of Walsh–Hadamard patterned masks.[30,31] The simulated overlap between the object and mask allowed for us to calculate what the image reconstructed after each measurement (i.e., each mask) would look like. We used binary images of digits from 1 to 9 as objects. We simulated the measurements for each digit and saved the simulated reconstructed image after each mask (for 1 to 800 masks). During the simulation, we introduced a noise factor which resulted in image blurring; to do this, we shifted the masks of one pixel off the central axis randomly in the four different directions. This introduced noise, blurred the reconstructed image, and introduced variation in our dataset. We simulated 8000 images in this manner to build our unique dataset. It must be noted that prior knowledge of the object type is important so as to build and/or generate a dataset that accurately represents the object. Images from our generated dataset are shown in panels 1 and 3 of **Figure 4** for digits 2 and 4, respectively.

We implemented four machine learning algorithms, starting with a SVM classifier. We utilized Python's scikit-learn module for a seamless implementation and trained one classifier for each class of digit in the SVM. The SVM was left to train for 1000 epochs. The L2 norm was used as the penalty where L2 regularization adds an L2 penalty equal to the square of the magnitude of coefficients and the loss function used was the square hinge. Second, we trained a logistic regression algorithm. Similarly, the model trains one classifier for each class of digit and implements a one-vs-all strategy. Training was carried out for 100 epochs, the L2 penalty was used, and the solver was limited-memory BFGS, which is the default solver of scikit-learn python module for L2 penalty.[53] The implemented NN algorithm is based on a Ball-Tree search and was implemented with a leaf size of 30. The Euclidean metric was used to determine the distance between the data points. This allowed for the training data to clump together in groups corresponding to each digit class. The number of neighbors used was five. Finally, the CNB classifier was implemented and trained on the same dataset; this classifier was the baseline, which we used to compare to the performance of the more classifiers implemented above. The accuracy for each of our models stood between 98% and 100% on the simulated measurements. The models were trained on an NVidia Tesla RTX 6000 GPU. The training times for the classifiers were at less than 3 min in total for all four models. The training times are as follows: SVM - 19 s, LR - 132 s, NN - 0.08 s, and NB - 0.16 s.

In the experimental image reconstruction process, the reconstructed image after each measurement (i.e., after each mask) was passed into each of the four classifiers above. The digit was recognized after each measurement and together with the confidence score was predicted by each algorithm. Important to note, each classifier has a built-in confidence prediction capability that is not necessarily within the same range across all the classifiers. We, therefore, had to normalize this confidence prediction to the same scale to determine how well the algorithms perform comparatively. For the SVM and LR, this entailed normalizing each classifier per predicted class. For categorical classification such as our recognition step, a probability of 50% or higher is considered acceptable; however, we imposed a strict early stopping criteria, stating that the recognition algorithms require a confidence greater than 75% to accurately identify the image. In this manner, the machine is now 75% confident that it has accurately identified the object; to a human, this will look like noise and is not recognizable as a specific digit.

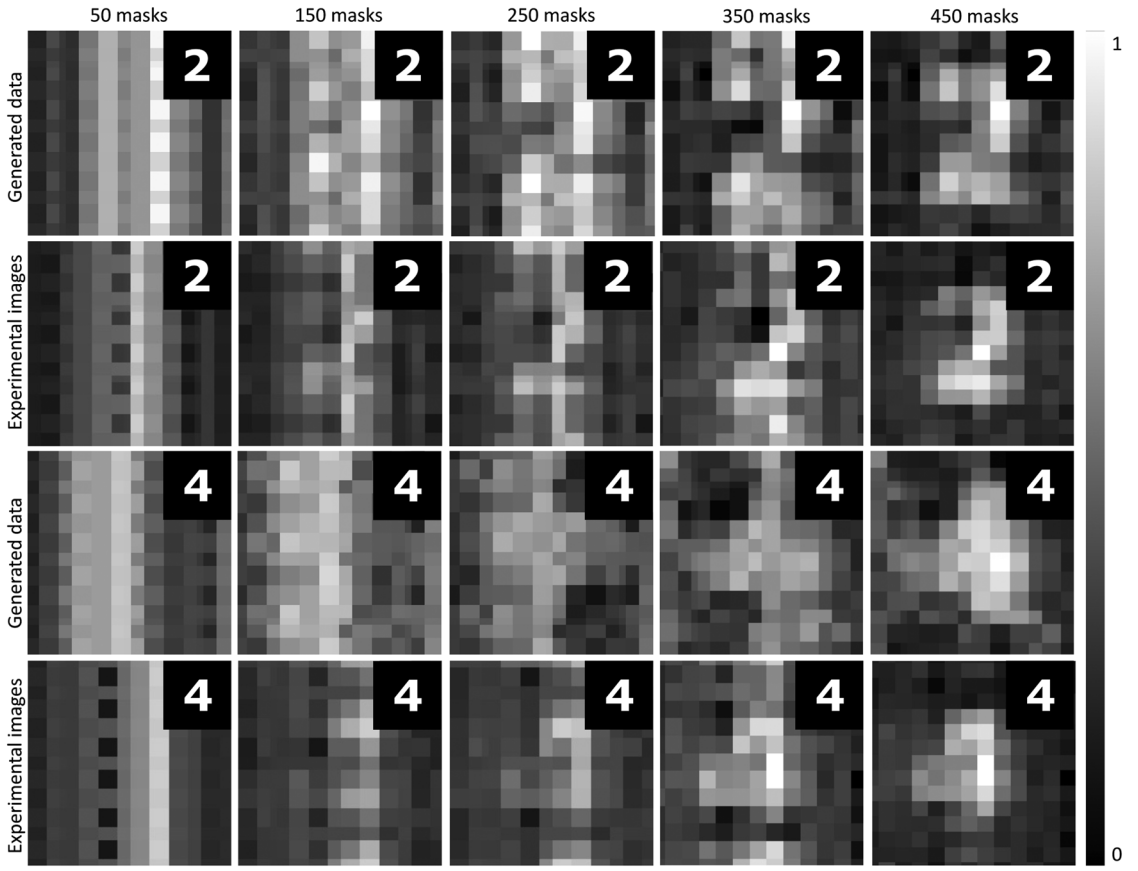

Figure 4. Uniquely generated and normalized dataset, on which the four machine learning algorithms were trained. Experimental image reconstructions acquired as an independent step are shown below the corresponding generated data for 50–450 masks in intervals of 100 masks, for digits 2 and 4. The digital objects are shown as insets in each frame.

4. Results

To get an accurate prediction from the machine learning classifiers, it was important that the training dataset was generated to contain features as similar as possible to the experimental data without replicating or using the experimental data for training purposes. The training dataset comprised synthetic data for each of the nine digits with reconstructions ranging from 1 to 800 masks to simulate the experimental image reconstruction process. In Figure 4, we show the generated data for input objects: digit 2 and digit 4, and we show the experimentally obtained image reconstructions corresponding to the generated data. We show the normalized data for image reconstructions using 50 to 450 masks in intervals of 100 iterations each, for digits 2 and 4. Adding blurring and noise to the generated dataset ensures that the training data is as aesthetically close as possible to the experimental data. This ensures that the algorithms are trained on the appropriate data and minimizes artifacts that would otherwise occur during application. Similarly, for different objects used in a ghost imaging experiment, it is necessary to curate a dataset that is as similar to the experimental data as possible. Prior knowledge of the class of the object to be imaged, such as the general category (cells, fingerprints, etc.), is necessary to generate a suitable dataset to train on. As mentioned above, the dataset generation and machine learning algorithm training were done as independent steps to taking experimental measurements. This was done so that instantaneous object recognition can take place post measurement and to implement object recognition as a step toward real-time ghost imaging.

With real-time ghost imaging as the goal, it is essential that recognition takes place with as few measurements as possible without compromising the recognition confidence. Digits 2 and 4 are shown in Figure 4 due to their non-symmetric nature; however, we tested digits 1 through 4 and averaged over all these predictions to better determine the error in the prediction. In Figure 4, we show the generated and experimental images for digits 2 and 4. The images in Figure 4 were reconstructed by 32×32 pixel Walsh–Hadamard masks. All objects in question were designed to be smaller than the imaging area to ensure an even illumination, leading to the effective resolution of the object to be lower than that of the mask. Additionally, for the sake of better visualization, the images were cropped in all cases. We show in Figure 4 that, at low iterations of an image reconstructed after 50 masks (extreme left), it is not possible for a human to discern what the image is. As the number of measurements increase, we are then able to discriminate between digit 2 (Figure 4 upper panels) and digit 4 (Figure 4 lower panels) due to our prior knowledge of the objects belonging to the digit class. The aim here was

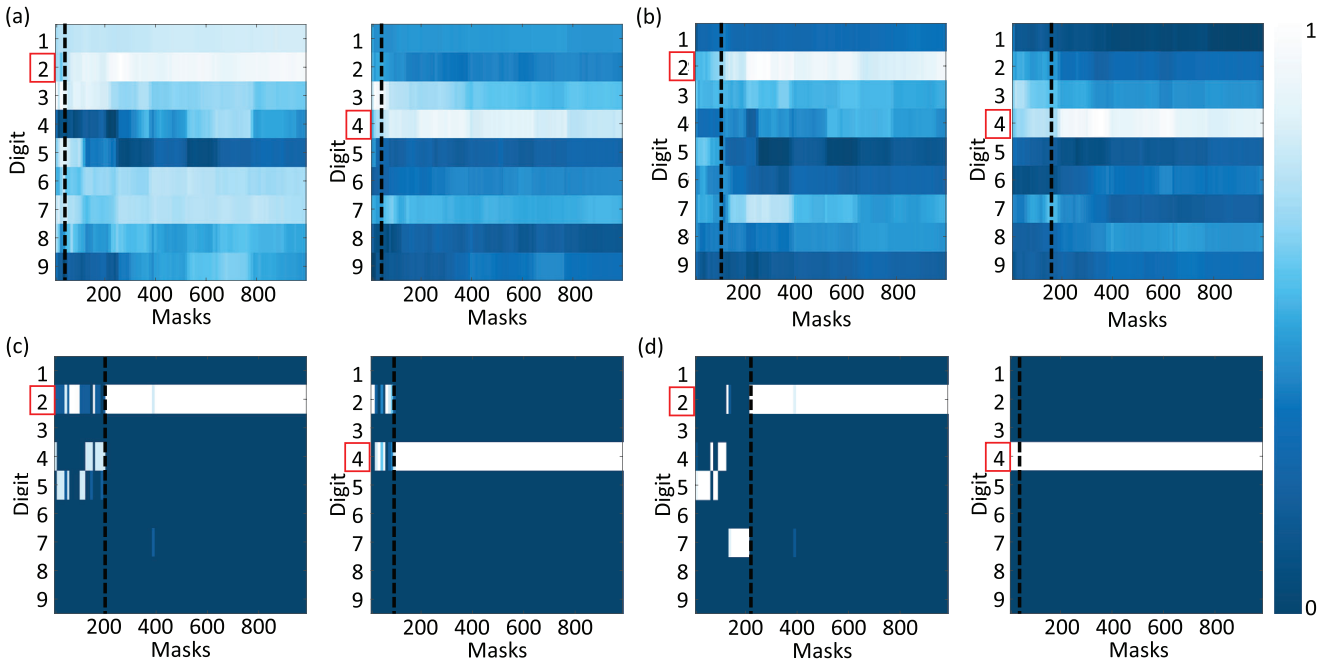

Figure 5. Confidence predictions of the machine learning algorithms for input digits 2 and 4, normalized to the same scale for comparative purposes. All algorithms were tested on images reconstructed by Walsh–Hadamard masks and are shown as follows: a) SVM, b) LR, c) NN, and d) NB, for objects digit 2 and digit 4 (indicated by the red outline on the *y*-axis), respectively. The dashed lines represent the point at which a confidence prediction greater than 75% is achieved for the input digit of interest.

Table 1. The number of masks required to satisfy the early stopping criteria, for each tested digit per classifier.

Classifier	Masks for digit 1	Masks for digit 2	Masks for digit 3	Masks for digit 4
SVM	10	10	10	10
LR	10	120	10	180
NN	10	200	390	90
NB	10	210	570	10

to train the classifiers to identify the images earlier than a human can.

We show the results of all four machine learning algorithms in **Figure 5**, which shows the predicted confidence scores across all algorithms (SVM, LR, NN, and NB), for images of digit 2 and digit 4, respectively. Although we limit the results in Figure 5 to show the digits whose nature is non-symmetric, we tested these algorithms and this approach on digits 1 through 4. We tabulate the results for all four digital objects used in our ghost imaging experiment in **Table 1**. In Figure 5, the dashed black lines indicate when the early stopping criteria was satisfied while the red boxes on the *y*-axes indicate the digit of interest corresponding to the input object in the experiment. Our SVM results are shown in Figure 5a. While we show that the stopping criteria was consistently satisfied at low mask numbers (an average of ten measurements), there is ambiguity in the confidence prediction of the SVM. The implementation of the SVM consists of training one classifier per class of digit; the classifiers were trained separately per digit class. A complete and overall normalization of the results was not possible due to there being a separate classifier per digit class. The SVM normalization was carried out per classifier, that is, per class of digit. Normalized predictions from this classifier were therefore ambiguous. While the stopping criteria is satisfied at low mask numbers, it can be seen that this classifier readily predicted an incorrect digit with a very high confidence also occurring at low mask numbers, again due to the fact that one classifier per digit class was implemented.

In Figure 5a, for digit 2, we see that the confidence increased in proportion to the mask number. This is also true for predictions of digits 1, 6, and 7. If one has prior knowledge of the digit class of interest, SVM would be the best performing algorithm. This, however, is not always the case and so the SVM would not be suited for tasks where the object class is not known. It was important to test the SVM on ghost imaging object recognition so as to establish that machine learning classifiers that are known for their superior performance, such as the SVM, are not always suited for all tasks, especially ones requiring a normalization such as our object recognition predictions. If, however, one does have prior knowledge of the digit class, then the SVM proves to be a highly efficient algorithm.

The LR results are shown in Figure 5b. This classifier is built similarly to the SVM and normalization was again carried out per class of digit. However, the predictions are less ambiguous due to the one-vs-all method of the LR algorithm. The LR classifier implements a one-vs-all strategy, that is, when a class is predicted as "true," we treat all other classes as "false." This therefore avoids any ambiguity in the confidence predictions of the incorrect digits. In this case, the classifier predicted the correct digit class with increasing confidence over all the masks used. Accordingly, the confidence of the incorrect digits diminishes in proportion to the number of masks used. Comparatively, the early stopping point

is not reached as soon with LR as with the SVM; however, due to LR avoiding ambiguity in predictions by the one-vs-all method, we satisfy the early stopping criteria at a mean of 80 masks across the input digits that were tested. The results, at which the early stopping criteria is satisfied, for input digits 1 through 4 are tabulated in Table 1.

In Figure 5c, we display the corresponding results for the NN classifier. Although this classifier achieves the early stopping criteria at a later point in the experiment, we see that there is little to no ambiguity in its predictions for the incorrect digits as there is a sharp drop-off in the confidence predictions for these classes. Here, the early stopping criteria is satisfied at an average of 173 measurements. Similarly, the NB algorithm, whose results are shown in Figure 5d, achieves the early stopping criteria at an average of 200 masks and performs similarly to that of the NN classifier where the ambiguity in the prediction of the incorrect classes is avoided by the sharp drop-off in confidence predictions. The NB algorithm is considered as the baseline algorithm in machine learning as it is the simplest algorithm to implement; however, it is not as powerful a classifier in comparison to other methods. NB is capable of predicting a high confidence between features it has learnt, which makes it prone to biases in testing and is therefore the less flexible algorithm.

In our work, all algorithms are suitable for implementation, depending on the level of performance required of the classifier and the level of prior knowledge of the object to be imaged. While the SVM is the most efficient, there are artifacts in the normalization and ambiguity in the confidence predictions of the incorrect digits. LR satisfies the stopping criteria later in the experiment but the ambiguity is avoided due to the one-vs strategy. NN and NB satisfy the early stopping criteria later than LR; however, there is a sharp drop-off in the confidence predictions of the incorrect digits completely avoiding any ambiguity. There is, therefore, a trade off between speed and confidence predictions of the incorrect digits.

Finally, in **Figure 6**, we show a bar graph, plotted from the results in Table 1. The bar graph indicates the average number of masks required to achieve our early stopping criteria across all digits. All recognition algorithms achieve the stipulated recognition criteria at an average of 200 masks or less. For individual digits, the NN and NB classifiers satisfy the stopping criteria later on in the experiment; this is indicated by the large error bars. Although NN and NB satisfy the stopping criteria later in the experiment, the confidence of the incorrect predictions sees a sharp drop-off leading to a greater certainty in these predictions. Comparatively, the SVM has a small margin of error for the correct digit of interest; however, it has a large uncertainty as it is prone to predicting the incorrect digit with a high confidence (indicated in Figure 5a). The confidence of both the correct and incorrect digit classes need to be considered to determine if there are any normalization artifacts or ambiguous predictions.

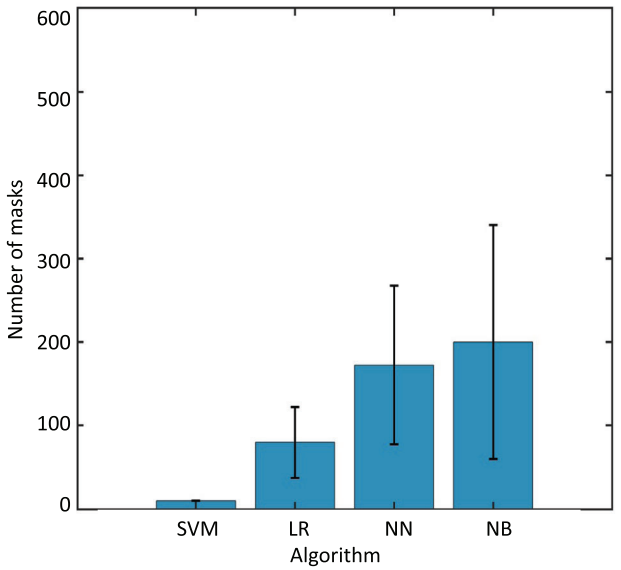

Figure 6. Bar graph indicating the average number of masks required to achieve the recognition criteria, for early stopping, across all digits using the Walsh–Hadamard masks for the experimental image reconstructions across all the classifiers tested.

5. Conclusion

To summarize, we implemented four machine learning classifiers in a quantum ghost imaging experiment to quickly recognize an object from its quantum image with as few photons as possible, showing a faster speedup (less photons) as compared to conventional quantum approaches. The algorithms were trained on a uniquely generated dataset to predict the recognized object as well as a confidence score after each measurement. The machine learnt to extract characteristics specific to early image reconstructions; what would look like noise to a human is accurately recognized as a digit by the machine with a certain confidence. After analyzing the confidence predictions of each classifier, over all reconstructions and objects, we found that while the SVM performs the best per respective digit it provides predictions for other digits that are ambiguous. LR performs better overall showing a gradual and steady confidence decrease in the non-relevant digits and a corresponding increase for the relevant digit. Accordingly, the one-vs-all method employed by LR avoids conflicting predictions. By implementing the LR classifier, image acquisition times are sped-up to 10× faster than a complete set of measurements. While the SVM and LR might be faster, NN and NB have a higher certainty in confidence predictions. It is therefore a trade-off between speed and confidence. Importantly, implementation time is further reduced as these models are not computationally intensive to implement, nor are they deep neural networks requiring copious amounts of training time. We have shown that by employing an early object recognition approach, we can establish an optimal early stopping point, thereby achieving a tenfold decrease in the time and number of photons needed for image recognition. We conclude that by using this approach the number of measurements required is significantly reduced while computational resources are not depleted on intensive and/or iterative models. In this manner, image information is maintained through object recognition. This, therefore, leads to a faster, more efficient image acquisition technique. While we demonstrated our method for amplitude-only objects, we anticipate it could be used for other object types, such as pure phase objects imaged through traditional quantum ghost imaging,[54] as well as in holographic imaging approaches used to extract the phase information of an object,[55] which will also benefit from re-

ducing image reconstruction time. We believe that this approach will prove valuable to the community working toward real-time ghost imaging in both the quantum and classical imaging domains.

Acknowledgements

C.M. would like to acknowledge the financial support received from the CSIR under the HCD-IBS scholarship scheme.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

ghost imaging, machine learning, quantum ghost imaging

Received: August 25, 2022
Revised: November 13, 2022
Published online: December 25, 2022

[1] Y. Shih, *Optics InfoBase Conf. Papers* **2007**, *13*, 1016.
[2] P.-A. Moreau, E. Toninelli, P. A. Morris, R. S. Aspden, T. Gregory, G. Spalding, R. W. Boyd, M. J. Padgett, *Opt. Express* **2018**, *26*, 7528.
[3] J. H. Shapiro, R. W. Boyd, *The Physics of Ghost Imaging*, vol. *11*, Springer, New York **2012**.
[4] M. J. Padgett, R. W. Bofyd, *Philos. Trans. R. Soc., A* **2017**, *375*, 2099.
[5] T. B. Pittman, Y. H. Shih, D. V. Strekalov, A. V. Sergienko, *Phys. Rev. A* **1995**, *52*, 5.
[6] A. F. Abouraddy, B. E. Saleh, A. V. Sergienko, M. C. Teich, *Phys. Rev. Lett.* **2001**, *87*, 123602.
[7] N. Bornman, M. Agnew, F. Zhu, A. Vallés, A. Forbes, J. Leach, *npj Quantum Inf* **2019**, *5*, 63.
[8] N. Bornman, S. Prabhakar, A. Vallés, J. Leach, A. Forbes, *New J. Phys.* **2019**, *21*, 073044.
[9] A. Gatti, E. Brambilla, L. Lugiato, *Prog. Opt.* **2008**, *51*, 251.
[10] E. Baleine, A. Dogariu, G. S. Agarwal, *Opt. Lett.* **2006**, *31*, 2124.
[11] R. Weissleder, *Nat. Biotechnol.* **2001**, *19*, 316.
[12] W. Gong, S. Han, *arXiv preprint arXiv:0911.4750* **2009**.
[13] Z. Liu, S. Tan, J. Wu, E. Li, X. Shen, S. Han, *Sci. Rep.* **2016**, *6*, 25718.
[14] Y. Wang, Y. Liu, J. Suo, G. Situ, C. Qiao, Q. Dai, *Sci. Rep.* **2017**, *7*, 45325.
[15] O. Katz, Y. Bromberg, Y. Silberberg, *Appl. Phys. Lett.* **2009**, *95*, 93.
[16] P. Zerom, K. W. C. Chan, J. C. Howell, R. W. Boyd, *Phys. Rev. A- At., Mol., Opt. Phys.* **2011**, *84*, 3.
[17] R. Zhu, G. Li, Y. Guo, *Int. J. Theor. Phys.* **2019**, *58*, 1215.
[18] X. Zhai, Z. Cheng, Y. Wei, Z. Liang, Y. Chen, *Opt. Eng.* **2019**, *58*, 8471.
[19] J.-N. Cao, Y.-H. Zuo, H.-H. Wang, W.-D. Feng, Z.-X. Yang, J. Ma, H.-R. Du, L. Gao, Z. Zhang, *Appl. Opt.* **2021**, *60*, 9180.
[20] L. Wang, S. Zhao, *Photonics Res.* **2016**, *4*, 240.
[21] H.-C. Liu, *Sci. Rep.* **2020**, *10*, 1.
[22] M. Lyu, W. Wang, H. Wang, H. Wang, G. Li, N. Chen, G. Situ, *Sci. Rep.* **2017**, *7*, 17865.
[23] T. Shimobaba, Y. Endo, T. Nishitsuji, T. Takahashi, Y. Nagahama, S. Hasegawa, M. Sano, R. Hirayama, T. Kakue, A. Shiraki, T. Ito, *Opt. Commun.* **2018**, *413*, 147.
[24] Z. Niu, J. Shi, L. Sun, Y. Zhu, J. Fan, G. Zeng, *Opt. Express* **2018**, *26*, 22773.
[25] H. Wu, R. Wang, G. Zhao, H. Xiao, J. Liang, D. Wang, X. Tian, L. Cheng, X. Zhang, *Optics and Lasers in Engineering* **2020**, *134*, 106183.
[26] C. Moodley, A. Forbes, *Sci. Rep.* **2022**, *12*, 10346.
[27] S. Rizvi, J. Cao, K. Zhang, Q. Hao, *Sci. Rep.* **2020**, *10*, 11400.
[28] H.-K. Hu, S. Sun, H.-Z. Lin, L. Jiang, W.-T. Liu, *Opt. Express* **2020**, *28*, 37284.
[29] F. Wang, C. Wang, M. Chen, W. Gong, Y. Zhang, S. Han, G. Situ, *Light: Sci. Appl.* **2022**, *11*, 1.
[30] V. Rodríguez-Fajardo, J. Pinnell, A. Forbes, *J. Mod. Opt.* **2020**, *67*, 1176.
[31] C. Moodley, B. Sephton, V. Rodríguez-Fajardo, A. Forbes, *Sci. Rep.* **2021**, *11*, 8561.
[32] X. He, S. M. Zhao, L. Wang, *Chin. Phys. B* **2021**, *30*, 5.
[33] J. Li, M. Le, J. Wang, W. Zhang, B. Li, J. Peng, *Appl. Phys. B: Lasers Opt.* **2020**, *126*, 1.
[34] Z. Ou, L. Mandel, *Phys. Rev. Lett.* **1988**, *61*, 50.
[35] A. Mair, A. Vaziri, G. Weihs, A. Zeilinger, *Nature* **2001**, *412*, 313.
[36] J. Torres, A. Alexandrescu, L. Torner, *Phys. Rev. A* **2003**, *68*, 050301.
[37] H. D. L. Pires, H. Florijn, M. Van Exter, *Phys. Rev. Lett.* **2010**, *104*, 020505.
[38] J. Romero, D. Giovannini, S. Franke-Arnold, S. Barnett, M. Padgett, *Phys. Rev. A* **2012**, *86*, 012334.
[39] J. Svozilík, J. Peřina Jr, J. P. Torres, *Phys. Rev. A* **2012**, *86*, 052318.
[40] M. G. McLaren, F. S. Roux, A. Forbes, *S. Afr. J. Sci.* **2015**, *111*, 1.
[41] A. Forbes, I. Nape, *AVS Quantum Sci.* **2019**, *1*, 011701.
[42] C. Couteau, *Contemp. Phys.* **2018**, *59*, 291.
[43] R. S. Aspden, D. S. Tasca, R. W. Boyd, M. J. Padgett, *New J. Phys.* **2013**, *15*, 073032.
[44] S. D. Johnson, P. A. Moreau, T. Gregory, M. J. Padgett, *Appl. Phys. Lett.* **2020**, *116*, 260504
[45] Z. Zhang, J. Zhong, *Opt. Lett.* **2016**, *41*, 2497.
[46] C. F. Higham, R. Murray-Smith, M. J. Padgett, M. P. Edgar, *Sci. Rep.* **2018**, *8*, 2369.
[47] J. C. Platt, in *Advances in Large Margin Classifiers*, MIT Press, Cambridge, MA **1999**, pp. 61–74.
[48] H.-F. Yu, F.-L. Huang, C.-J. Lin, *Mach. Learn.* **2011**, *85*, 41.
[49] E. Fix, J. L. Hodges, *International Statistical Review / Revue Internationale de Statistique* **1989**, *57*, 238.
[50] J. D. M. Rennie, L. Shih, J. Teevan, D. R. Karger, in *Proc. of the Twentieth International Conf. on Int. Conf. on Machine Learning*, ICML'03, AAAI Press, Washington, DC **2003**, pp. 616–623.
[51] J. Leach, B. Jack, J. Romero, M. Ritsch-Marte, R. W. Boyd, A. K. Jha, S. M. Barnett, S. Franke-Arnold, M. J. Padgett, *Opt. Express* **2009**, *17*, 8287.
[52] R. E. Meyers, K. S. Deacon, *Quantum Communications and Quantum Imaging VII* **2009**, *7465*, 746508.
[53] C. Zhu, R. H. Byrd, P. Lu, J. Nocedal, *ACM Trans. Math. Softw.* **1997**, *23*, 550.
[54] M. McLaren, A. Forbes, *Journal of Optics* **2017**, *19*, 44006.
[55] H. Defienne, B. Ndagano, A. Lyons, D. Faccio, *Nat. Phys.* **2021**, *17*, 591.