



## ISTITUTO NAZIONALE DI RICERCA METROLOGICA Repository Istituzionale

### Quantum-Enhanced Pattern Recognition

This is the author's submitted version of the contribution published as:

*Original*

Quantum-Enhanced Pattern Recognition / Ortolano, Giuseppe; Napoli, Carmine; Harney, Cillian; Pirandola, Stefano; Leonetti, Giuseppe; Boucher, Pauline; Losero, Elena; Genovese, Marco; Ruo-Berchera, Ivano. - In: PHYSICAL REVIEW APPLIED. - ISSN 2331-7019. - 20:2(2023). [10.1103/physrevapplied.20.024072]

*Availability:*

This version is available at: 11696/82420.3 since: 2025-01-17T08:27:04Z

*Publisher:*

American Physical Society

*Published*

DOI:10.1103/physrevapplied.20.024072

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

# Quantum-enhanced pattern recognition

Giuseppe Ortolano<sup>1,2</sup>, Carmine Napoli<sup>1</sup>, Cillian Harney<sup>3</sup>, Stefano Pirandola<sup>3</sup>, Giuseppe Leonetti<sup>1,2</sup>, Pauline Boucher<sup>1</sup>, Elena Losero<sup>1</sup>, Marco Genovese<sup>1</sup>, and Ivano Ruo-Berchera<sup>1</sup>

<sup>1</sup>*Quantum metrology and nano technologies division,  
INRiM, Strada delle Cacce 91, 10153 Torino, Italy*

<sup>2</sup>*DISAT, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy and*

<sup>3</sup>*Department of Computer Science, University of York, York YO10 5GH, United Kingdom*

The challenge of pattern recognition is to invoke a strategy that can accurately extract features of a dataset and classify its samples. In realistic scenarios this dataset may be a physical system from which we want to retrieve information, such as in the readout of optical classical memories. The theoretical and experimental development of quantum reading has demonstrated that the readout of optical memories can be dramatically enhanced through the use of quantum resources (namely entangled input-states) over that of the best classical strategies. However, the practicality of this quantum advantage hinges upon the scalability of quantum reading, and up to now its experimental demonstration has been limited to individual cells. In this work, we demonstrate for the first time quantum advantage in the multi-cell problem of pattern recognition. Through experimental realizations of digits from the MNIST handwritten digit dataset, and the application of advanced classical post-processing, we report the use of entangled probe states and photon-counting to achieve quantum advantage in classification error over that achieved with classical resources, confirming that the advantage gained through quantum sensors can be sustained throughout pattern recognition and complex post-processing. This motivates future developments of quantum-enhanced pattern recognition of bosonic-loss within complex domains.

Significant progress has been made in recent years in quantum sensing [1–3], a broad spanning field of research which seeks to realise quantum protocols for estimation [4–8] and discrimination [9–12] that achieve better performance than any classical protocol, i.e. gain a quantum-advantage. Efforts to design and engineer such protocols have been widely successful, with applications in data-readout [13–22], target detection [23–36], cryptography [37, 38], fundamental physics research [39–41] and imaging [42–44].

An application of particular interest is that of quantum reading, which encompasses the challenge of classical information retrieval from an optical memory. A memory cell can be considered as a reflective (or transmissive) target with two distinct possible reflectivities (transmittivities) encoding a classical bit. This represents a pair of possible bosonic-lossy channels that act on incident photonic probe states. Accurate readout of a memory cell requires the correct discrimination of output states from the channel (i.e., quantum channel discrimination). Having conceptualised this model, Pirandola [13] proved that one can significantly enhance the accuracy of information retrieval through the use of entangled light sources. This seminal work led to a series of advancements and most recently, a successful experimental demonstration [20].

The success of quantum reading prompted the development of further schemes in more complex domains, such as the simultaneous readout of many memory cells. This breeds the task of quantum multi-channel discrimination [45, 46], in which one must optimise multi-mode probe states and measurements in order to classify a collection of bosonic-lossy channels (or *channel patterns*). This development brings new layers of complication; expanding the space of possible channel patterns to be dis-

criminated, and revealing a more general spectrum of multi-modal discrimination protocols. This invites new applications in which quantum reading may serve as a primitive.

The use of quantum sensors within the task of pattern recognition provides such an opportunity, where one is tasked with classifying a set of patterns (or images). The pattern distribution is often non-uniform, and classifications follow highly non-linear relationships, e.g. the classification of hand-written digits. Powerful classical post-processing techniques are well known for this challenge, however the quality of the readout of data-sets are often taken for granted. Given a system that can be modelled as an ensemble of bosonic-lossy multi-channels, it has been theoretically shown that quantum reading can substantially enhance classification power over optimal classical strategies [47–51].

In this work, we report the experimental demonstration of quantum-enhanced pattern recognition using signal-idler entangled two-mode squeezed vacuum (TMSV) states and photon-counting measurements to outperform any classical strategy. Using the MNIST handwritten digit dataset [52], image samples are experimentally realised and serve as a test-bed for the discrimination protocols. Investigating both multi-shot and single-shot regimes, we show that the use of entangled probe states can generate quantum advantage in the classification error as a consequence of the gain offered by quantum resources. By considering multiple supervised-learning approaches to classification, we emphasise that this advantage is an inherent property of quantum sensing that can survive throughout complex post-processing. Our work paves the way for a suite of near-term applications of quantum-enhanced pattern recognition and

imaging.

## I. RESULTS

Let us consider a pattern classification problem where spatial patterns are imprinted on a  $d \times d$  array of binary cells, each encoding one bit of information in two values of a physical parameter, for example the transmittances  $\tau_0$  and  $\tau_1$ , where we pose  $\tau_0 \leq \tau_1$  (see Fig. 1A). The cells size and dimension of the array  $d^2$ , can be a native characteristic of the physical support but also they can be practically determined by resolution properties of the sensing apparatus, for example the pixel structure of a camera or the resolution of an optical system, so that  $\tau_0$  and  $\tau_1$  are intended as the values corresponding to each readout pixel. The sensing strategy has to be designed to minimize the average classification error probability of the patterns,  $P$ .

The general sensing procedure can be described as in Fig. 1A: an input state, represented by the density operator  $\rho$ , interacts with the object encoding the pattern, and a POVM measurement is performed on the output state, followed by classical post-processing. The state  $\rho$  is generally bipartite in order to include ancillary assisted schemes. The two systems are indicated as signal (S) and idler (I), according to quantum optics terminology. Since we are dealing with bosonic systems, the problem is non-trivial only if we impose some constraints on the resources, the most natural one being to fix the mean number of signal photons,  $\mu$ , probing the cell. The pixels are assumed to be probed by independent spatial modes, and in this work we restrict our analysis to local measurements, i.e. independent readout is performed for each pixel at the receiver.

Following the local measurement, a binary value is assigned to each pixel with a certain average probability of error  $p$ . The binary image is then classified. The classification error  $P$  is related to the bit-flip probability  $p$  by a, generally, non-linear map  $\mathcal{M}_C$ :

$$P = \mathcal{M}_C(p). \quad (1)$$

The map  $\mathcal{M}_C$  depends on the problem at hand, but it is reasonable to assume it monotonic for any effective classification algorithm. Using the monotonicity of  $\mathcal{M}_C$ , the minimization of  $P$  is obtained by minimizing  $p$  over all possible input states and POVM measurements. Details on the single-pixel problem are provided in Sec. III A.

The theoretical lower bound on the error probability achievable by classical states,  $p_{\text{c1a}}^{\text{LB}}$  is reported in Eq. (3). It can be saturated, without idler assistance, by a single mode coherent state  $|\alpha\rangle$ , and an optimal measurement consisting in the projection on two pseudo-cat states as shown in Sec. III B.

Such a measurement appears to be experimentally unfeasible and no other receiver that saturates  $p_{\text{c1a}}^{\text{LB}}$  has been proposed so far. Nevertheless, a photon counting (PC) receiver, composed of PC measurement followed by

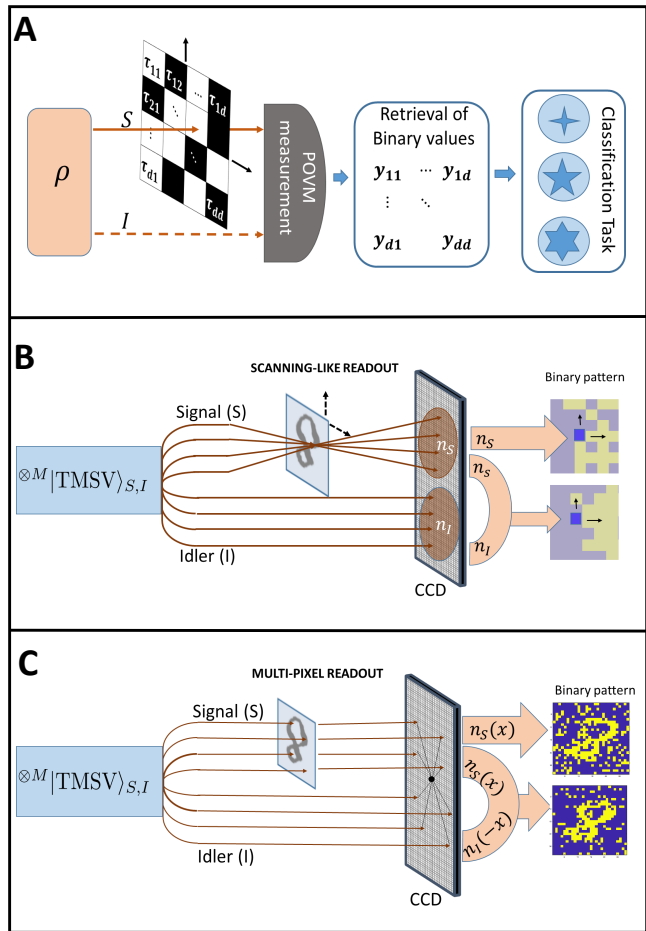


FIG. 1. *Scheme of Pattern Recognition.* Panel A shows the general configuration of a pattern recognition problem. An image composed of binary pixels, having one of two possible values of transmittance  $\tau_{i,j} = \{\tau_0, \tau_1\}$ , is probed by a bipartite state  $\rho$ . The signal (S) system interacts with the image. The idler (I) system is directly sent to the measurement apparatus where a joint measurement is performed with the signal. The result of the measurement are processed classically to classify the patterns. Panel B shows a possible experimental configuration, in which the pattern is "scanned" with sequential single-pixel PC measurements, where  $n_S$  and  $n_I$  are the signal and idler numbers of photons detected. Panel C shows a different configuration in which the spatial multi-mode structure of the source is used, so that a parallel readout of all the pixels of the image is done in a single shot. The pairwise quantum correlation of the different modes, represented schematically by the brown arrows hitting symmetric points of the CCD, are used to enhance the classification.

a Bayesian decision, can be proven to perform close to the optimal one for practical purposes [20]. As the PC is a phase-insensitive measurement, it follows that any state with Poisson photon number distribution can reach the PC limit,  $p_{\text{c1a}}^{\text{PC}}$ , as reported in Eq. (6). For this reason, we use a PC receiver to experimentally perform the pixel readout with both classical and quantum states.

The quantum readout was performed using a col-

lection of  $M$  replicas of TMSV states,  $\rho_{\text{qua}} = |\text{TMSV}\rangle_{S,I}\langle\text{TMSV}|^{\otimes M}$ , with

$$|\text{TMSV}\rangle_{S,I} \propto \sum_n \sqrt{P_{\mu_0}(n)} |n\rangle_S |n\rangle_I. \quad (2)$$

$P_{\mu_0}(n) = \mu_0^n / (\mu_0 + 1)^{n+1}$  is the thermal distribution with  $\mu_0 = \mu/M$  mean photons. In the quantum case, idler assistance is crucial and it is actually convenient to spread the signal energy across a large number of modes, i.e. to work with  $M \gg 1$  and  $\mu_0 \ll 1$  [20]. The corresponding error probability  $p_{\text{qua}}^{\text{PC}}$ , discussed in Sec. III A, can be evaluated numerically.

Among the various experimental imperfections, the most relevant is photon loss due to non-ideal optical component and detector inefficiency. The effect of the losses on classical readout performances is equivalent to a rescaling of the energy, namely the input mean photon number is substituted by the detected photons,  $\mu \rightarrow \eta\mu$  in the classical bounds of Eqs. (3) and (6), where  $\eta$  is the overall photon detection probability. However, the quantum strategy is additionally affected by deterioration of the bipartite non-classical correlation due to losses, which describes the imperfect coupling of correlated modes into the detectors. This effect is discussed in Sec. I C.

### A. Experimental realization

The principle of the experimental scheme is presented in Fig. 1 B-C. The  $M$ -product TMSV quantum state  $\rho_{\text{qua}}$  is produced by traveling wave Spontaneous Parametric Down Conversion (SPDC) of type II in a non-linear crystal pumped by a continuous laser. Due to the conservation of the transverse momentum of the pump  $\mathbf{q} = 0$ , it follows  $\mathbf{q}_S = -\mathbf{q}_I$ , i.e. signal and idler photons are emitted along correlated directions (spatial modes), which are mapped into symmetric pixels of a CCD camera. In our experiment the exposure time of image acquisition is much longer than the coherence time of the process, so many temporal modes, on the order of  $10^{11}$ , contribute to the total number of photons detected in each pixel (few thousands). Clearly, the conditions  $M \gg 1$  and  $\mu_0 \ll 1$  which maximize the quantum advantage are fulfilled and the single pixel photon statistics is practically indistinguishable from the Poisson one. For this reason, the best classical performance with PC detection can be evaluated by considering only the signal beam. Details on the experimental setup are reported in Sec. III E.

In general, traveling wave SPDC is spatially broadband. There are many (several thousands) pairwise correlated spatial modes available for spatially resolved multi-pixel quantum readout [53, 54]. However, a trade-off should be set between the spatial resolution, in terms of pixels available in one shot measurement, and the reduction of the pixel error probability due to quantum correlation. This point is discussed later on in the main text, and further details can be found in Sec. III.

To address this trade-off, we have realized the experiment in two different configurations, one named ‘‘scanning-like readout’’ (Fig. 1B) that favors the best quantum enhancement, and the second one named ‘‘multi-pixel readout’’ (Fig. 1C) that allows spatially resolved pattern acquisitions with in a single-shot measurement.

### B. Scanning-Like Readout

In this first experimental configuration, we mimic a readout in which the physical binary pattern encoded by the transmittance  $\tau_0$  and  $\tau_1$  is scanned point-by-point, as shown in Fig. 1B. First, we repeated a large number of independent measurements of two well-characterized transmittances,  $\tau_0$  and  $\tau_1$ , assigning a bit value ( $y = 0, 1$ ) for each measurement according to the Bayes’ rule in Eq. (4) based on the number of photon-counts  $n_S$  and  $n_I$  (see Sec. III A). The photon-counts are integrated over two large regions of the CCD array, defining the size of the readout pixel, collecting a large number of signal and idler spatial modes respectively, obtaining the best possible correlation between  $n_S$  and  $n_I$ , as discussed in Sec. III D. This produces two sets of data  $A_0 = \{y_k^{(0)}\}_{k=1,\dots,K}$  and  $A_1 = \{y_l^{(1)}\}_{l=1,\dots,L}$  for the nominal transmittance  $\tau_0$  and  $\tau_1$  respectively, with  $L + K \gg d^2$ . The procedure is performed both for classical (only using signal beam) and quantum reading. From those sets we can evaluate the average experimental readout pixel error probabilities,  $p_{\text{cla}}^{\text{PC}}$  and  $p_{\text{qua}}^{\text{PC}}$  shown in Fig. 2A. One of the transmittance is fixed to  $\tau_1 = 1$  while the other one,  $\tau_0$ , is varied. All other parameters are kept fixed. The experimental points for both the quantum (blue) and classical (red) readout are compared with the classical lower bound  $p_{\text{qua}}^{\text{LB}}$  (green) defined by Eq. (3). As expected, in general the error probability increases as the two transmittances become closer and closer. However, through the range of parameters showed, the strategy employing quantum states sensibly outperforms the classical one paired with the suboptimal PC measurement and is also able to surpass the absolute classical lower bound. For the sake of completeness, in Fig. 2B we report the same pixel error probabilities in an almost ideal scenario where the detection efficiency is fixed to  $\eta = 0.97$  (the efficiency reached in the actual setup is  $\eta \approx 0.79$ ).

Note that, the data picked from the set  $A_0$  and  $A_1$  can be virtually rearranged to reproduce the outcomes of a sequential spatial scanning of any given  $d \times d$  binary pattern. However, for an accurate evaluation of the pattern recognition task, the used approach is highly preferable with respect to the real scanning of spatial samples. In fact, the last one would require the physical realization and scanning of a large number of spatial samples, which is hardly feasible and actually not necessary.

To evaluate the pattern recognition performance we used the MNIST handwritten digit dataset, containing 60000 training samples of handwritten numerical digits

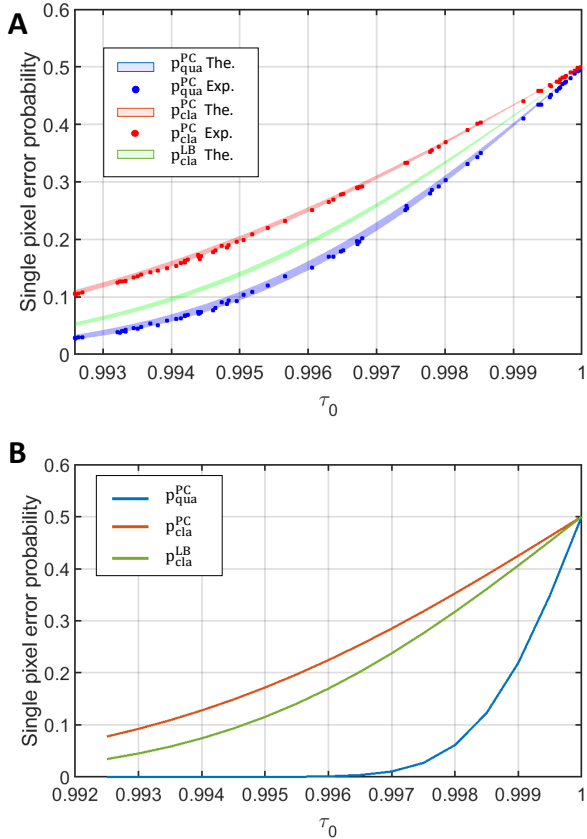


FIG. 2. *Single-pixel readout error probabilities.* Panel **A** reports the error probability in the readout of a bit encoded in the transmittance value (either  $\tau_0$  or  $\tau_1$ ) of a single-pixel. The lower transmittances,  $\tau_0$ , is varied while the other one fixed to  $\tau_1 = 1$ . The estimated mean number of photons per pixel is  $\mu = (1.45 \pm 0.05) \cdot 10^5$ , the efficiencies are  $\eta_S = 0.79 \pm 0.01$  and  $\eta_I = 0.791 \pm 0.004$ , and the electronic noise is  $\nu \approx 6 \cdot 10^3$ . The probability of error for the quantum strategy is reported in blue, the classical strategy with photon counting in red and the best classical strategy in green. The confidence interval of the theoretical curves is obtained by considering one standard deviation on the reported parameters. Panel **B** reports theoretical prediction on the error probabilities in the almost ideal case of  $\eta_S = \eta_I = 0.97$ , while all the other parameters are the same of panel **A**.

and 10000 test samples. We assemble the entire set of test patterns by picking experimental data from the set  $A_0$  and  $A_1$ . Fig. 3 shows some examples of digit contained in the dataset, and a comparison of the effect of the noise in either a classical and a quantum binary readout.

For the classification of the noisy digits we used a  $k$ -Nearest Neighbor ( $k$ -NN) classifier [55]. A  $k$ -NN classifier computes the distance between the pattern tested and all the patterns in its training set. In the case of binary patterns, the most natural metric is the Hamming distance. The  $k$  closest patterns of the training set to the tested pattern are selected, and the class of the test pattern is

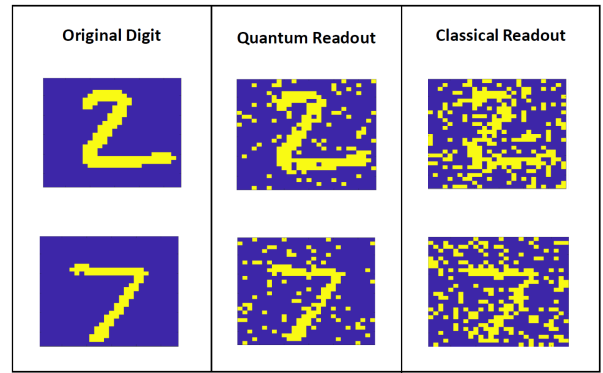


FIG. 3. *Readout of handwritten digits with different strategies.* Two examples of binary pattern of dimension  $28 \times 28$  with quantum and classical PC readout noise.

assigned as the more common among the  $k$  closest. The experimental classification error is reported in Fig. 4A as a function of the transmittance  $\tau_0$ , while in Fig. 4B we report the classification performance in the almost ideal ( $\eta = 0.97$ ) detection scenario. Note that the performance of the  $k$ -NN classifier is highly non-linear with respect to the single pixel error probability. For the values of  $\tau_0$  corresponding to a classical error probability  $p_{\text{cla}}^{\text{PC}} < 0.3$  in Fig. 2A, the classification is very robust to noise. In this range, the classification error is almost negligible for all the strategies and the quantum advantage is also rather small. However, for higher transmittance  $\tau_0$ , the algorithm becomes very sensitive to the readout noise level and here the advantage of quantum strategy is amplified. In fact, the experimental quantum advantage in the pattern recognition task, represented in Fig. 4C, reaches up to 4 dB, which is much larger than the corresponding pixel error improvement.

At this point, the question arises: is the amplification of the quantum advantage we have observed for the  $k$ -NN classifier a property of this specific classifier, or a more general behavior characterizing the pattern classification task? To partially address this question, we repeat our analysis with another classification strategy that uses a refined machine learning approach.

Convolutional neural networks (CNNs) are a category of feed-forward neural networks that are omnipresent in the field of image processing, and are particularly powerful tool for pattern recognition [56, 57]. They contain hidden layers that convolve input data from one layer to the next, in which the convolutions are performed according to parameterised filters (or kernels). The goal of learning is to optimise these filters in such a way that complex features can be reliably extracted and classified. This form of network architecture helps to minimise the number of network parameters and also helps the CNN to interpret features in a translationally invariant way, i.e. qualities inferred in one area of an image can be reused elsewhere in the image.

In this work, we apply CNN classifiers to the task of

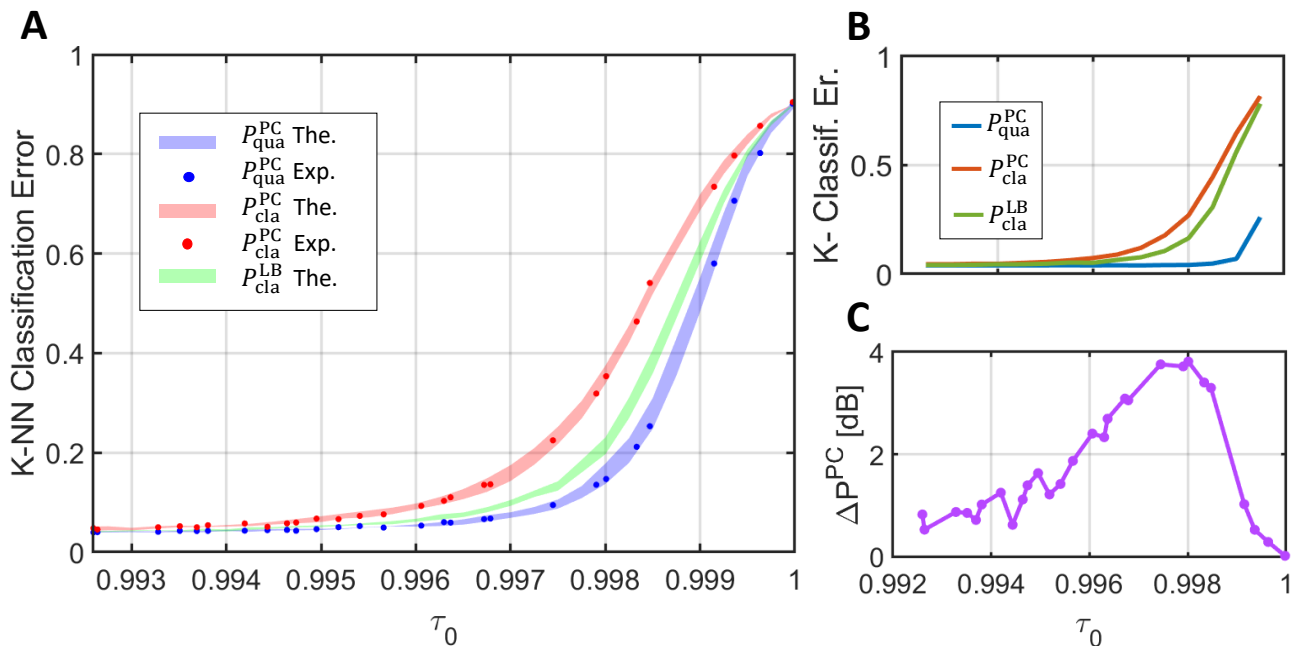


FIG. 4. *Classification error with the  $k$ -NN algorithm.* Panel **A** shows the average classification error  $P$  in the recognition of handwritten digits encoded in binary transmittance values ( $\tau_0$  and  $\tau_1$ ). The lower transmittance,  $\tau_0$ , is varied, while  $\tau_1 = 1$ . Experimental points are reported for the quantum (blue) and classical (red) photon counting (PC) readout along with the theoretical predictions. The green band reports a simulation of the performance obtained with a pixel readout at the absolute classical limit. Panel **B** shows the classification error for almost ideal efficiency,  $\eta = 0.97$ . Panel **C** reports the experimental quantum advantage, in dB ( $\Delta P^{\text{PC}} := -10 \log_{10}(P_{\text{qua}}^{\text{PC}}/P_{\text{cla}}^{\text{PC}})$ ).

noisy pattern recognition on the MNIST dataset as before with the  $k$ -NN classifier, considering experimental input data from quantum-enhanced and classical sensors. We trained a collection of classifiers on a noisy training set in which each pixel in an image undergoes a bit flip with single-pixel error rate of  $p_{\text{train}}$ , i.e. the single-pixel training error. Meanwhile, the network parameter optimisation was carried out by backpropagation using a cross-entropy loss cost function. More precisely, we prepared a set of 16 CNNs, whose  $p_{\text{train}}$  was distributed across the interval  $[0, 0.5]$ . The single-pixel training error is a particularly influential parameter as fine-tuning  $p_{\text{train}}$  helps classifiers become more robust to noise and learn how to reduce its impact on classification [58]. The classifiers could then be evaluated on the experimental test sets (constructed from  $A_0$  and  $A_1$ ) in order to quantify the performance of quantum-enhanced versus classical pattern recognition.

The results of the CNN strategy are presented in Fig. 5. Fig. 5A compares the performance of the classical and quantum-enhanced classifiers. We have a collection of CNN classifiers which have been trained using different single-pixel training errors, denoted by  $\{C[p_{\text{train}}]\}_{p_{\text{train}} \in [0, 0.5]}$ . It follows that some classifiers can garner better performance than others. Ideally, we would claim the classification error associated with the best classifier from our set, i.e.  $C[p_{\text{train}}^{\text{max}}]$ . But this approach remains unrealistic, since it is not possible to predict *a*

*priori* what the best  $p_{\text{train}}$  would be. Instead, we may use the single-pixel error rates associated with the chosen discrimination protocol to motivate our choice of training error rate at a given transmissivity  $\tau_0$ , constructing the classifiers  $C[p_{\text{train}}^{\text{theory}}]$  (for more details, see Sec. IIID). Fig. 5(b) plots these corresponding values against the single-pixel transmissivity.

Finally, Fig. 5(c) reports the quantum-advantage in the pattern recognition task (in decibels) showing a similar behaviour to the one observed in Fig. 4.C. This confirms that the quantum-advantage gathered through the single-pixel readout can be sustained throughout complex post-processing techniques, such as CNN training and evaluation.

### C. Parallel Multi-Pixel Readout

While the scanning-like readout of the previous section represents a meaningful proof-of-principle necessary for a clean and faithful comparison with the theory, here we consider a more realistic scenario of a multi-pixel parallel readout, as it is depicted in Fig. 1C, where the pattern is acquired in single-shot.

The samples, containing digits from 0 to 9, were realized on a AR-coated glass slide combining laser lithography, sputtering deposition of titanium in high vacuum and lift off technique. The transmittance of the depo-

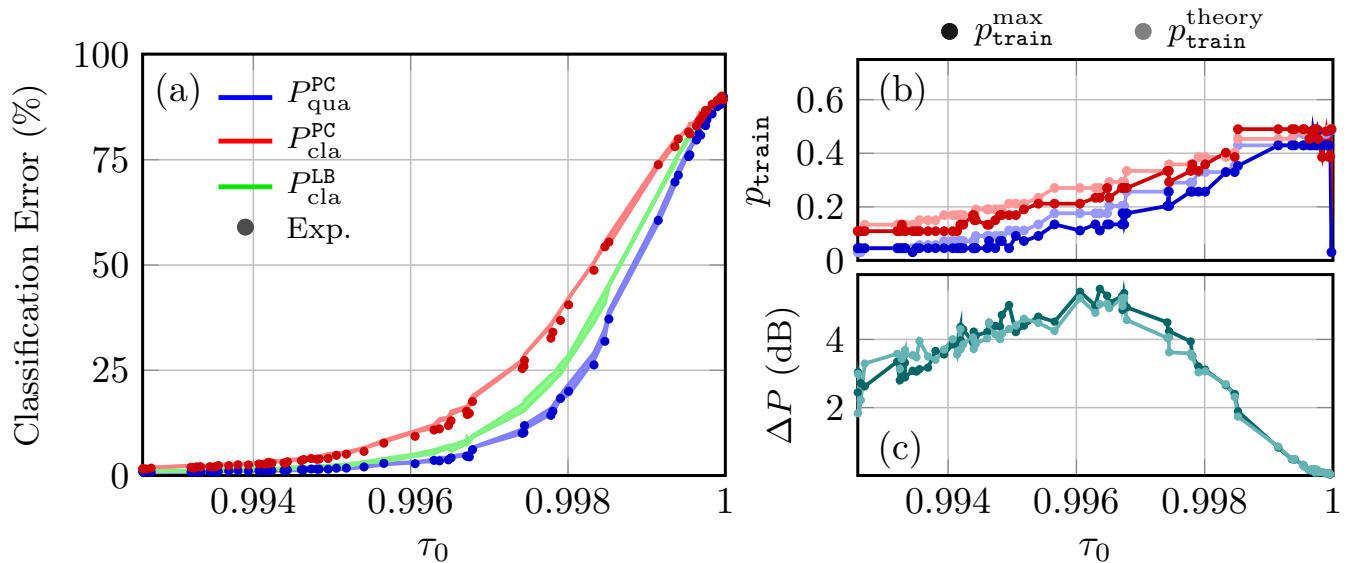


FIG. 5. *Quantum-enhanced pattern recognition via the CNNs.* Panel (a) displays the classification error of CNNs on the MNIST using sensors with quantum resources (blue) or classical resources (red). We trained 16 CNNs on noisy training sets of 60000 images with single-pixel errors in the interval  $p_{\text{train}} \in [0, 0.5]$ . We used these CNNs to classify a test set of 10000 images simulated at different single-pixel transmissivities,  $\tau_0$ . The solid lines in (a) portray the simulated classification error region associated with a CNN using a single-pixel training error predicted by the corresponding theoretical single-pixel error bounds  $p_{\text{train}}^{\text{theory}}$  at  $\tau_0$ . Experimental data points are plotted for both quantum and classical photon counting protocols. Both the theory predicted value and the best-choice value for  $p_{\text{train}}$  are plotted in Panel (b). Finally, Panel (c) plots the experimental quantum advantage  $\Delta P$  in decibels obtained with using TMSV states over classical coherent states (and photon counting). The dark points plot the advantage between CNNs that use the best-choice  $p_{\text{train}}^{\text{max}}$  while the lighter points plot the advantage gained using theory predicted  $p_{\text{train}}^{\text{theory}}$ .

sition is estimated as  $\tau_0 = 0.987 \pm 0.003$ . The parallel readout is doable since the state  $\rho_{\text{qua}}$  produced by SPDC is spatially multimode. Pairwise correlated modes are detected by symmetric pairs of pixels in the CCD chip as represented in Fig. 1C, so that thousand pixel pairs are in principle non-classically correlated at the same time. However, to efficiently collect the correlated modes, the size of the pixel  $l_D$  should be larger than the transverse cross-correlation length  $l_c$  of the source. The consequence is a trade-off between the obtainable spatial resolution and the quantum advantage in the single pixel readout. We set the pixel size of our camera to fulfill the condition  $l_D > l_c$  and at the same time to have a  $28 \times 28$  image of the digits (more details are presented in Sec IIID).

Fig. 6A shows the experimental pixel error probabilities for both classical and quantum PC readout in function of the effective transverse resolution at the sample plane. The effective resolution is changed by applying a moving averaging filter to the acquired images, as described in Sec IIID, which has practically the same effect of increasing the pixel size. On one side, the increasing of the pixel size has the direct consequence of raising the number of signal photons per pixel, that explains the dropping of classical error probability, according to Eq. (6). On the other side, as clearly shown in Fig. 6A, the quantum error probability decreases faster than the classical analogous, because the detectable quantum correlation

increases while losing spatial resolution. Here, the comparison of the experimental results with the theoretical bounds is not straightforward and we will avoid it. In fact, when binary patterns are readout with finite resolution, for a detection pixel that falls across the edge between  $\tau_0$  and  $\tau_1$  the binary model is no longer valid. Such border effects are not negligible as the resolution decreases and become comparable with the spatial features of the pattern. These features are very difficult to account for in a theoretical model.

We evaluated the pattern recognition performance once again on the MNIST handwritten digit dataset using a  $k$ -NN algorithm for the classification. While the test dataset is composed of 10000 patterns, producing and acquiring 10000 samples is a not feasible and unnecessary task. In fact, the error in the classification can be broken down in two contributions. The first contribution is intrinsic to the classification algorithm and the second comes from the noise in the readout. In our case, with a pixel error larger than about 0.3, the classification error is dominated by the noise in sensing. For reference the classification error without noise is of the order of 4%, while in the noisy region it is clearly much larger. So, we assume that the classification error can be well studied by printing a small subset of patterns from the original MNIST, namely 10, and performing on each of them one thousand different acquisitions with independent noise

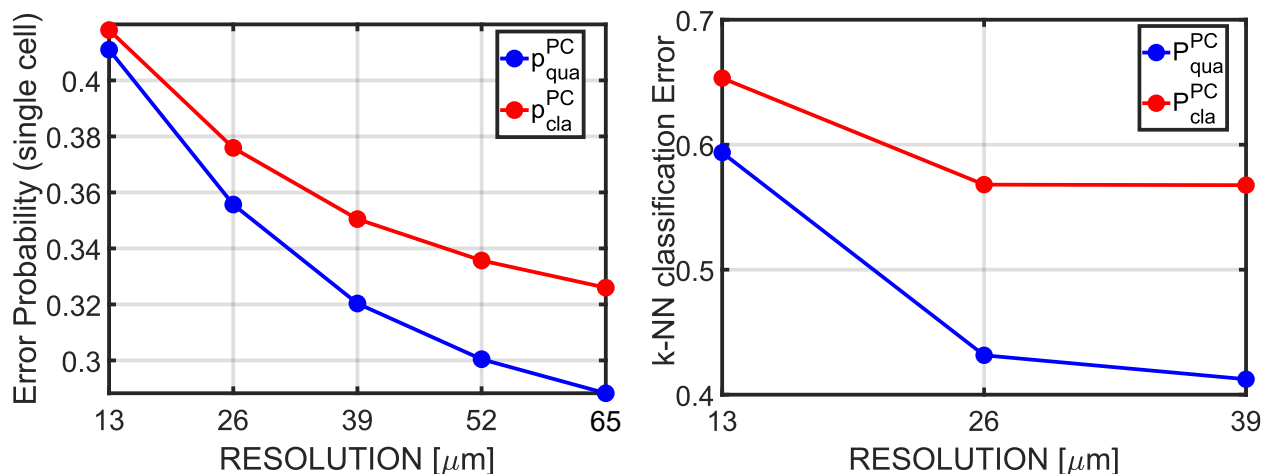


FIG. 6. *Multi-Pixel readout performance as a function of the spatial resolution.* Panel **A** shows the single-pixel error probability decreasing with the resolution at the object plane. The mean number of photons per pixel per frame, at full resolution (13  $\mu\text{m}$ ) is  $\mu = 1380 \pm 20$ , while the total channel efficiencies of signal and idler branches are  $\eta_S = 0.795 \pm 0.01$  and  $\eta_I = 0.815 \pm 0.01$ . The pattern transmittance level is fixed to  $\tau_0 = 0.987 \pm 0.003$ . Panel **B** reports the error in the pattern classification with a  $k$ -NN algorithm.

realizations, forming a total dataset of 10000 images.

The resulting classification error is reported in Fig. 6B as a function of the spatial resolution. The results are reported up to a resolution of 39  $\mu\text{m}$  which is equivalent to an effective reduction of the resolution of a factor 3 with respect to the original dataset. Confirming the main results of the of the previous section, in a sensing scenario dominated by quantum noise, a relatively small quantum enhancement in the single pixel probability (namely from 0.35 to 0.32) can be amplified to around 20% in the classification.

## II. DISCUSSION AND CONCLUSION

In this work we have experimentally demonstrated the advantage of quantum-enhanced sensing in the task of pattern recognition. As a test bed we have considered the problem of classification of handwritten digits. A quantum readout strategy based on TMSV states and photon counting measurement brings a relevant advantage in the classification errors with respect to any classical sensing.

We have realized the experiment in two different configurations. In the first one, called “scanning-like readout”, we have privileged the magnitude of the quantum advantage with respect to the acquisition time, since scanning of the sample is required. In this case we were able to clearly beat the optimal classical bound. In the second configuration, instead we have considered single-

shot multi-pixel readout of the pattern. In this case we have shown a large advantage with respect to the classical bound that uses photon counting.

Moreover, the presence of quantum advantage turns out to be independent on complexity of the classification algorithm. Having deployed both  $k$ -NN classifiers (a simple and robust strategy) and CNN classifiers (a machine learning enhanced strategy) on the same experimental data, similar results are obtained: quantum advantage in the sensing is maintained and often amplified in the pattern recognition task.

The results presented here are particularly promising for short terms application in biological spatial pattern recognition. In general, they also pave the way for experimental studies of quantum enhanced pattern recognition in the spectral and temporal domains.

## III. MATERIALS & METHODS

### A. Single pixel readout

The single-pixel problem has been extensively analysed in Refs.[13, 15, 16, 20, 22]. In the following we will compare three different local readout strategies. In the first scenario we consider classical input states  $\rho_{\text{c1a}}$ , i.e. a mixture of coherent states, paired with a theoretical optimal POVM  $\Lambda$ . The optimal probability of error in discriminating the outcome states  $\rho_{\text{c1a}}(\tau_0)$  and  $\rho_{\text{c1a}}(\tau_1)$  can be



derived by the Helstrom bound [4],  $(1 - D(\rho_0, \rho_1))/2$ , where  $\rho_0$  and  $\rho_1$  are two generic states, and  $D(\rho_0, \rho_1)$  is the trace distance. Using the convexity of  $D(\rho_0, \rho_1)$ , Ref.[13] demonstrated the following bound to the classical error probability:

$$p_{\text{c1a}}^{\text{LB}} = \frac{1 - \sqrt{1 - e^{-\mu_0(\sqrt{\tau_0} - \sqrt{\tau_1})^2}}}{2} \quad (3)$$

where  $\mu_0$  is the mean number of signal photons. As discussed in [13], ancillary modes cannot enhance the performance of the readout with classical states, in fact the bound in Eq.(3) can be obtained by using a single mode coherent state transmitter  $\rho_{\text{c1a}} = |\alpha\rangle\langle\alpha|$ , with  $|\alpha|^2 = \mu_0$ , as shown in Sec. III B.

However, the optimal measurement, that we discuss in Sec. III B, is purely theoretical, and no feasible receiver has been proposed yet to reach the classical lower bound in Eq. (3). Thus, we consider a second, more practical scenario where the receiver is based on photon counting (PC) measurements that can be proven to perform close to the optimal one for practical purposes [20]. The conditional photon number distribution at the receiver is  $p(\mathbf{n}|\tau_i) = \langle \mathbf{n} | \rho(\tau_i) | \mathbf{n} \rangle$ , where  $\mathbf{n} = (n_S, n_I)$  is the number of measured photons in the signal and idler (the ancilla) systems. The optimal choice to recover the value of the bit  $j$  is to choose it according to the Bayes' rule

$$j = \arg \max_i p(\tau_i | \mathbf{n}). \quad (4)$$

This condition is equivalent to the maximum likelihood decision  $j = \arg \max_i p(\mathbf{n}|\tau_i)$  in the balanced prior probability case. The error probability with photon counting receiver is given by:

$$p^{\text{PC}} = \frac{1}{2} \sum_{\mathbf{n}} \min_i p(\mathbf{n}|\tau_i) \quad (5)$$

i.e., proportional to the overlap of the measurement outcome distributions  $p(\mathbf{n}|\tau_0)$  and  $p(\mathbf{n}|\tau_1)$ . As for the optimal classical case, the conditional distribution overlap cannot be reduced by using ancillary mode, and the overlap is minimized for input signal states having the photon statistics of a coherent state, i.e. a Poisson distribution  $\mathcal{P}_{\mu_0}(n)$ , where  $\mu_0$  is the mean number of photons [20]. The statistics of the photons returning from a cell of transmittance  $\tau_i$  still follows the Poisson distribution, in particular  $\mathcal{P}_{\mu_0\tau_i}(n_s)$ . Substituting it in the conditional probabilities of Eq. (5), the best performance achievable with classical states and PC measurements is [20]:

$$p_{\text{c1a}}^{\text{PC}} = \frac{1}{2} - \frac{\Gamma(\lfloor n^{\text{th}} + 1 \rfloor, \mu_0\tau_0) - \Gamma(\lfloor n^{\text{th}} + 1 \rfloor, \mu_0\tau_1)}{2\lfloor n^{\text{th}} \rfloor!} \quad (6)$$

where  $\Gamma(x, y)$  is the incomplete gamma function,  $\lfloor x \rfloor$  is the floor of  $x$  and  $n^{\text{th}} := \mu_0(\tau_1 - \tau_0)/\log(\tau_1/\tau_0)$  is the intersecting point of the two Poisson distributions  $\mathcal{P}_{\mu_0\tau_0}$  and  $\mathcal{P}_{\mu_0\tau_1}$ .

The third strategy we consider here uses a collection of  $M$  replicas of Two Mode Squeezed Vacuum

(TMSV) states in input,  $\rho_{\text{qua}} = |\text{TMSV}\rangle_{S,I}\langle\text{TMSV}|^{\otimes M}$ , with  $|\text{TMSV}\rangle_{I,S} \propto \sum_n \sqrt{\mathcal{P}_{\mu_0}(n)} |n\rangle_S |n\rangle_I$  and  $\mathcal{P}_{\mu_0}(n) = \mu_0^n / (\mu_0 + 1)^{n+1}$  being a thermal distribution with  $\mu_0$  mean photons. The joint probability of signal and idler photons of a TMSV state after the interaction with the cell of transmittance  $\tau_i$  is  $p_{\mu_0}(n_S, n_I | \tau_u) = \mathcal{P}_{\mu_0}(n_I) \mathcal{B}(n_S | n_I, \tau_i)$ , where  $\mathcal{B}$  stands for the binomial distribution. For the multi mode state  $\rho_{\text{qua}}$  the joint distribution is simply obtained by substituting the single mode thermal statistics  $\mathcal{P}_{\mu_0}(n_I)$  with a multi-thermal one,  $\mathcal{P}_{\mu_0, M}(n_I)$ , obtained by the convolution of  $M$  thermal modes. The conditional joint probability evaluated in this way can be substituted in Eq. (5) to obtain the error of the quantum strategy,  $p_{\text{qua}}^{\text{PC}}$ . This quantity can be evaluated numerically. It turns out that, for a fixed total number of photon transmitted to the cell  $\mu_0 M$ , the most effective way to exploit quantum correlation is to span the photons in a large number of modes so that  $\mu_0 \ll 1$  and  $M \gg 1$ . Moreover, the local multi-thermal distribution approaches the Poisson one. This is the regime that we consider in the simulation, and the one that is realized in the experiment.

## B. Optimal local receiver for classical states

The single pixel readout probability of error that can be reached with classical states, having a fixed number of signal photons  $\mu$ , is lower bounded by the limit in Eq. (3). As previously mentioned, this limit can be saturated by using a single-mode coherent state without idlers. This can be seen using the fact that a pure loss channel, with transmissivity  $\tau$ , maps a coherent state  $|\alpha\rangle$  into another, amplitude damped, coherent state  $|\tau\alpha\rangle$ . Once the input state  $\rho_{in}$  is fixed, the discrimination of two channels having different transmissivity is reduced to a binary quantum state discrimination between the two possible outputs,  $\rho_0$  and  $\rho_1$  for which the lowest error probability is given by the Helstrom formula:

$$p^{\text{H}} = \frac{1}{2} (1 - \|\pi_0\rho_0 - \pi_1\rho_1\|), \quad (7)$$

where  $\|\cdot\|$  is the trace distance and  $\pi_0, \pi_1$  are the prior probabilities.

For coherent input, the potential outputs of the channel are pure and their trace distance can be expressed in terms of the overlap  $\zeta = \langle \sqrt{\tau_0}\alpha | \sqrt{\tau_1}\alpha \rangle = e^{-\frac{\mu}{2}(\sqrt{\tau_0} - \sqrt{\tau_1})^2}$  as:

$$\|\pi_0|\sqrt{\tau_0}\alpha\rangle\langle\sqrt{\tau_0}\alpha| - \pi_1|\sqrt{\tau_1}\alpha\rangle\langle\sqrt{\tau_1}\alpha|\| = \sqrt{1 - 4\pi_0\pi_1|\zeta|^2}.$$

Substituting this in Eq.(7) and setting equal priors  $\pi_0 = \pi_1 = 1/2$  we get:

$$p_{\text{coh}} = \frac{1 - \sqrt{1 - |\zeta|^2}}{2}, \quad (8)$$

which coincides with  $p_{\text{c1a}}^{\text{LB}}$  proving that a single mode coherent state saturates the classical bound (originally proven for an arbitrary P-representation [13]).

The optimal probability of error is achieved by projecting into the eigenstates of the Hermitian operator:

$$\Lambda = \frac{1}{2}(|\sqrt{\tau_0}\alpha\rangle\langle\sqrt{\tau_0}\alpha| - |\sqrt{\tau_1}\alpha\rangle\langle\sqrt{\tau_1}\alpha|). \quad (9)$$

We can find an orthogonal basis to represent  $\Lambda$  by performing the Gram-Smith orthogonalization on the pair of vectors ( $|\sqrt{\tau_0}\alpha\rangle, |\sqrt{\tau_1}\alpha\rangle$ ) spanning its support. So we define the basis:

$$|0\rangle = |\sqrt{\tau_0}\alpha\rangle \quad |1\rangle = \frac{|\sqrt{\tau_1}\alpha\rangle - \zeta|\sqrt{\tau_0}\alpha\rangle}{\sqrt{1 - |\zeta|^2}} \quad (10)$$

The eigenvectors of  $\Lambda$  can be computed with simple algebra to be, in terms of this basis:

$$|+\rangle = \frac{1}{\sqrt{2}}\sqrt{1 - \sqrt{1 - \zeta^2}}|0\rangle - \frac{1}{\sqrt{2}}\sqrt{1 + \sqrt{1 - \zeta^2}}|1\rangle \quad (11)$$

$$|-\rangle = \frac{1}{\sqrt{2}}\sqrt{1 + \sqrt{1 - \zeta^2}}|0\rangle + \frac{1}{\sqrt{2}}\sqrt{1 - \sqrt{1 - \zeta^2}}|1\rangle \quad (12)$$

And substituting the definitions in Eq.(10) leads to:

$$|+\rangle = \sqrt{\frac{1 - \sqrt{1 - \zeta^2}}{2(1 - \zeta^2)}}|\sqrt{\tau_1}\alpha\rangle - \sqrt{\frac{1 + \sqrt{1 - \zeta^2}}{2(1 - \zeta^2)}}|\sqrt{\tau_0}\alpha\rangle \quad (13)$$

$$|-\rangle = \sqrt{\frac{1 + \sqrt{1 - \zeta^2}}{2(1 - \zeta^2)}}|\sqrt{\tau_1}\alpha\rangle - \sqrt{\frac{1 - \sqrt{1 - \zeta^2}}{2(1 - \zeta^2)}}|\sqrt{\tau_0}\alpha\rangle \quad (14)$$

The bit can then be recovered by measuring over the projectors  $\Pi_0 = |+\rangle\langle+|$  and  $\Pi_1 = |-\rangle\langle-|$ . Once the measurement is performed the value  $\tau_0$  is selected if the outcome is  $+$  and  $\tau_1$  is selected otherwise. This measurement paired with a coherent state input would in theory saturate the classical bound in Eq.(3). However, as pointed out in the main text, from a practical point of view an implementation of this kind of measurement would be very complicated, if feasible at all, with current technology.

### C. $k$ -NN classification

Let  $\mathbf{i}_0 = \{b_{ij}^{(0)}\}$  and  $\mathbf{i}_1 = \{b_{ij}^{(1)}\}$ ,  $b_{ij} = 0, 1$ , be two binary images of dimension  $d \times d$ . The Hamming distance,  $d_{\text{H}}$ , between  $\mathbf{i}_0$  and  $\mathbf{i}_1$  is defined as:

$$d_{\text{H}}(\mathbf{i}_0, \mathbf{i}_1) = \sum_{i=1}^d \sum_{j=1}^d |b_{ij}^{(0)} - b_{ij}^{(1)}|. \quad (15)$$

The  $k$ -NN is a supervised classification method that uses the known label of images in a training set  $\mathcal{T}$  to assign a class to images in a test set  $\mathcal{V}$ . Given an image  $\mathbf{i}_v$  in  $\mathcal{V}$ ,

the set of its  $k$  closest images in terms of the hamming distance,  $\mathcal{K} = \{\mathbf{i}_{\mathcal{T}}^{(1)}, \dots, \mathbf{i}_{\mathcal{T}}^{(k)}\} \subset \mathcal{T}$ , are selected from  $\mathcal{T}$ . The image  $\mathbf{i}_v$  is assigned to the class  $c_v$  which is the most common among the images in  $\mathcal{K}$ . The most intuitive case is given by the 1-NN classifier, that assigns to the image  $\mathbf{i}_v$  the same class of its closest image in the training set  $\mathcal{T}$ .

### D. Neural Network classification

Consider a dataset of images  $\mathcal{I} := \{\mathbf{i}_j; c_j\}_j$  in which each image  $\mathbf{i}_j$  possesses a classification (or label)  $c_j$ . For the task of classification, a neural network [59, 60] is a universal approximating function  $f_{\theta}$  (where  $\theta$  denotes the parameter set of the neural network) which takes an image as input and provides an approximation of its classification as an output,  $f_{\theta}(\mathbf{i}_j) = c'_j$ . Training a neural network as a classifier equates to optimising its parameters  $\theta$  such that  $f_{\theta}(\mathbf{i}_j) \approx c_j$  becomes a reliable classifier over the entire dataset. That is, the goal of training is to maximising the accuracy of the classifier over the training set in such a way that it can generalise to new, unforeseen samples. In a neural network, the parameter set  $\theta$  describes correlations between layers of neurons, capable to encoding highly complex relationships between input and output.

When building neural network classifiers for which there already exists experimental data (as in the case of MNIST) supervised learning is a very effective strategy. This involves utilising a set of images whose classifications are already exactly known, and training the neural network on this data. In this way, the classifier can be optimised reliably and extract meaningful features from the dataset in order to learn correlations between input patterns and their classification. The classification error of the classifier is measured against an evaluation (or test) set; a set of images which the network has never seen during training and therefore must extrapolate. Given a  $K$ -element evaluation set of images and labels  $\mathcal{V} = \{\mathbf{i}_k; c_k\}_{k=1}^K$  and a classifier  $f_{\theta}$  then the classification error is approximated as

$$P_{f_{\theta}} = \frac{1}{K} \sum_{k=1}^K \delta[f_{\theta}(\mathbf{i}_k), c_k] \quad (16)$$

where  $\delta[f_{\theta}(\mathbf{i}_k), c_k]$  is a Kronecker-delta function that returns unity iff the classifications are the same, otherwise it will return zero.

In this work, we exploit CNNs as classifiers upon our experimental dataset. Using the Julia Flux package [61, 62], we construct CNNs with three convolutional layers and a dense layer, optimising a cross entropy loss function during training. This produces classifiers of sufficient accuracy to utilise in this study. We use the MNIST handwritten digit dataset, composed of a training set  $\mathcal{T}$  with  $6 \times 10^4$  images and an experimental evaluation set  $\mathcal{V}$ . The experimental dataset is clearly unchange-

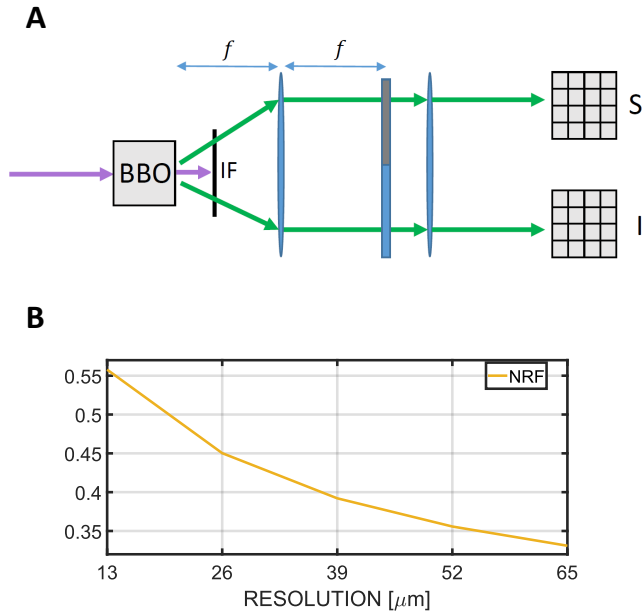


FIG. 7. **A** Schematic of the experimental setup described in Sec. III E. **B**. Noise reduction factor (NRF) as a function of the effective spatial resolution.

able and is inherently noisy. But the initially noiseless training set can be manipulated to modify and enhance our classifiers. As discussed in the main text, it is useful to train a CNN on noisy data if it is expected to classify noisy data. To do this, we simulate noisy training sets  $\mathcal{T}_{p_{\text{train}}}$  in which each pixel in every image is independently bit flipped according to the probability  $p_{\text{train}}$ . We then trained a collection of different classifiers using different values of  $p_{\text{train}} \in [0, 0.5]$  and evaluated them on the experimental data.

It is difficult to know what the best case  $p_{\text{train}}$  is with respect to the transmissivity of the sampled pixels. As depicted in Fig. 5 we can identify a best choice classifier from our collection as that which minimises its classification error with respect to the evaluation set. However, it is not practical to select a best choice classifier *a posteriori* (this would defeat the purpose of discrimination). Instead, we present the best-case performance for benchmarking purposes and more practically motivate our chosen  $p_{\text{train}}$  by mapping the single pixel transmissivity to a single-pixel error probability associated with the discrimination protocol being used.

### E. Experimental Setup and Noise Reduction

Acquisitions for both the experimental configurations reported in Fig. 1B-C, are performed using the setup sketched in Fig. 7. It exploits SPDC to generate multi-

mode pairwise correlated collections of TMSV states, by pumping a type II Beta Barium Borate (BBO) non-linear crystal with a 100 mW continuous laser at wavelength  $\lambda_p = 405$  nm. The SPDC photons are selected around the degenerate wavelength  $\lambda = 2\lambda_p = 810$  nm by an interference filter (IF) at  $(800 \pm 20)$  nm. The correlations in momentum at the source are converted in spatial correlations at the "object" plane, by use of a lens having focal length  $f = 1$  cm, in an  $f-f$  configuration, performing a Fourier transform on the incoming field. The test patterns to be classified are physically realized as depositions, according to the procedure of Sec. III F. The patterns are placed in the signal beam at the object plane, while a blank glass slide is placed on the idler beam to match the two optical paths. The object plane is then imaged to the charged coupled device (CCD) camera using a second lens, with magnification of about 8 times. The camera (Princeton Instrument Pixis 400BR Excelon) has a nominal quantum efficiency  $> 95\%$ , electronic noise of  $4e^-/\text{pix}/\text{frame}$  and the physical pixel is of  $13 \mu\text{m}$ . The total channel efficiencies  $\eta_S$  and  $\eta_I$ , are estimated close to 80% according to the procedure outlined in Ref.s [63, 64]

In the scanning-like configuration, we introduce in the signal beam a variable absorber that is fine tuned in the range  $1 - \tau_0 = 10^{-2} \div 10^{-3}$ . Signal and idler photons  $n_S$  and  $n_I$  are obtained integrating over two corresponding wide spatial areas composed of a large number of physical pixels, becoming equivalent to two bucket detectors. This leads to the most efficient collection of the correlated photons [21] and thus to reach an higher quantum advantage, as we will explain in the following.

In contrast, in the parallel multi-pixel readout, we exploit the spatial resolution of the detector to have a one-shot (without scanning) image of size  $28 \times 28$  of a test digit in the signal arm, which is the original dimension of the binary images of the MNIST dataset. For that, we perform a  $8 \times 8$  hardware binning of the physical pixels, to obtain an effective readout pixel of  $l_D = 104 \mu\text{m}$ , corresponding to a resolution of  $l_R = 13 \mu\text{m}$  at the object plane (not to be confused with the original physical pixel size of the detector mentioned earlier).

However, a trade-off exists between spatial resolution and quantum noise reduction achieved exploiting the signal-idler correlations. Actually, correlated photons from the multi-mode SPDC arrive at the detector within a certain transverse spatial uncertainty, that comes from the finite size of the pump beam. For this reason, the signal-idler intensity cross-correlation is well approximated by a Gaussian function with finite coherence length  $l_c$  at the detector plane (in our case is  $l_c \sim 40 \mu\text{m}$ ). To efficiently detect the correlated modes, the pixel size, or more in general the detector resolution area, must be larger than this coherence length,  $l_D \geq l_c$  [65]. In Sec.(I C) we analyse the pattern recognition performance as a function of the resolution. For a fair comparison in the classification, it is more suitable to have final binary images with always the same number of pixels. Thus, instead of changing the resolution by performing a simple

pixel binning, we preferred to change it by applying an averaging filter of appropriate size. The averaging filter substitutes to each pixel count, the average of the counts in a neighborhood of size  $d \times d$ . This is repeated for each original pixel of the  $28 \times 28$  matrix, so that the final size of the image in terms of pixels is unvaried, although the effective spatial resolution is reduced. In the graph of Fig. 6A, the parameter  $d$  ranges between 1 and 5, which correspond to an effective resolution  $l_R$  between  $13 \mu\text{m}$  and  $65 \mu\text{m}$ . The improvement of the quantum correlations in function of the image resolution can be witnessed by the noise reduction factor (NRF), defined as  $\text{NRF} = \langle \Delta^2(n_S - n_I) \rangle / \langle n_S + n_I \rangle$  [66–70]. The NRF is an indicator of non-classical correlations, since it can show values in the range  $0 < \text{NRF} < 1$  only for non-classical fields, while it is  $\text{NRF} \geq 1$  for classical light. In Fig.(7.B) we report the measured NRF as a function of the effective resolution at the object plane.

## F. Handwritten digit deposition

The CAD of the 10 digits (from 0 to 9) was realised by normalising the dimension of each binary images ( $28 \times 28$  pixels) of digit in squares of maximum dimension of  $400 \times 400 \mu\text{m}^2$  spaced of  $1000 \mu\text{m}$ . A laser lithography process was performed with Heidelberg Instruments uPG101 system equipped with a UV laser source at  $375 \text{ nm}$ . The deposition of titanium was carried out by sputtering in high vacuum at low rate ( $0,08 \text{ nm/s}$ ). Thicknesses from 2 to 4 nm were deposited to obtain a transmittance value suitable for the experiment. Final patterning was obtained removing titanium excess through lift off technique.

## ACKNOWLEDGMENTS

Part of this work (sample fabrication) has been carried out at QR Laboratories, INRiM, a micro and nanofabrication lab. IRB thank the INRiM researcher Matteo Fretto for supervising the sample realization.

## Funding

This work was funded by the EU via “Quantum read-out techniques and technologies” (QUARTET, Grant agreement No 862644).

## Author contributions

GO and IRB devised the present realization of quantum enhanced pattern recognition, with advice of SP. EL, PB, CN and GO performed the experimental acquisitions. GO did the data analysis on the experimental data and computed the optimal local measurement to saturate the classical lower bound. GO and CH performed the classification analysis for  $k$ -NN and CNN respectively. The samples have been prepared by GL. MG, head of the INRiM quantum optics sector and IRB, supervised the project. GO, IRB, SP, CH wrote the paper with the contribution of all authors.

## Competing Interests

The authors declare no competing interest.

## Data availability

All data needed to evaluate the conclusions are reported in the paper. Further data, for reproducibility of the results, will be available in a public repository linked to the published paper.

- 
- [1] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, *Reviews of Modern Physics* **89**, 035002 (2017).
- [2] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook, and S. Lloyd, Advances in photonic quantum sensing, *Nat. Photon* **12**, 724 (2018).
- [3] G. Petrini, E. Moreva, E. Bernardi, P. Traina, G. Tomagra, V. Carabelli, I. P. Degiovanni, and M. Genovese, Is a quantum biosensing revolution approaching? perspectives in nv-assisted current and thermal biosensing in living cells, *Advanced Quantum Technologies* **3**, 2000066 (2020), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/qute.202000066>.
- [4] C. Helstrom, *Quantum detection and estimation theory* (Academic Press, New York, 1976).
- [5] S. L. Braunstein and C. M. Caves, Statistical distance and the geometry of quantum states, *Phys. Rev. Lett.* **72**, 3439 (1994).
- [6] S. L. Braunstein, C. M. Caves, and G. Milburn, Generalized uncertainty relations: Theory, examples, and lorentz invariance, *Ann. Phys.* **247**, 135 (1996).
- [7] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum-enhanced measurements: Beating the standard quantum limit, *Science* **306**, 1330 (2004), <https://www.science.org/doi/pdf/10.1126/science.1104149>.
- [8] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, *Nat. Photon* **5**, 222 (2011).
- [9] C. W. Helstrom, Quantum detection and estimation theory, *J. Stat. Phys.* **1**, 231 (1969).

- [10] A. Chefles and S. M. Barnett, Quantum state separation, unambiguous discrimination and exact cloning, *Journal of Physics A: Mathematical and General* **31**, 10097 (1998).
- [11] A. Chefles, Quantum state discrimination, *Contemporary Physics* **41**, 401 (2000).
- [12] J. A. Bergou, U. Herzog, and M. Hillery, 11 discrimination of quantum states, in *Quantum State Estimation* (Springer Berlin Heidelberg, 2004) pp. 417–465.
- [13] S. Pirandola, Quantum reading of a classical digital memory, *Phys. Rev. Lett.* **106**, 090504 (2011).
- [14] S. Pirandola, C. Lupo, V. Giovannetti, S. Mancini, and S. L. Braunstein, Quantum reading capacity, *New J. Phys.* **13**, 113012 (2011).
- [15] R. Nair, Discriminating quantum-optical beam-splitter channels with number-diagonal signal states: Applications to quantum reading and target detection, *Phys. Rev. A* **84**, 032312 (2011).
- [16] C. Invernizzi, M. G. A. Paris, and S. Pirandola, Optimal detection of losses by thermal probes, *Phys. Rev. A* **84**, 022334 (2011).
- [17] M. Dall’Arno, A. Bisio, and G. M. D’Ariano, Ideal quantum reading of optical memories, *Int. J. Quantum Inf.* **10**, 1241010 (2012).
- [18] G. Spedalieri, C. Lupo, S. Mancini, S. L. Braunstein, and S. Pirandola, Quantum reading under a local energy constraint, *Phys. Rev. A* **86**, 012315 (2012).
- [19] J. P. Tej, A. R. U. Devi, and A. K. Rajagopal, Quantum reading of digital memory with non-gaussian entangled light, *Phys. Rev. A* **87**, 052308 (2013).
- [20] G. Ortolano, E. Losero, S. Pirandola, M. Genovese, and I. Ruo-Berchera, Experimental quantum reading with photon counting, *Sci. Adv.* **7**, eabc7796 (2021).
- [21] G. Ortolano, P. Boucher, I. P. Degiovanni, E. Losero, M. Genovese, and I. Ruo-Berchera, Quantum conformance test, *Science Advances* **7**, eabm3093 (2021), <https://www.science.org/doi/pdf/10.1126/sciadv.abm3093>.
- [22] G. Ortolano and I. Ruo-Berchera, Quantum readout of imperfect classical data, *Sensors* **22**, 2266 (2022).
- [23] M. F. Sacchi, Entanglement can enhance the distinguishability of entanglement-breaking channels, *Phys. Rev. A* **72**, 014305 (2005).
- [24] S. Lloyd, Enhanced sensitivity of photodetection via quantum illumination, *Science* **321**, 1463 (2008).
- [25] S.-H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro, Quantum illumination with gaussian states, *Phys. Rev. Lett.* **101**, 253601 (2008).
- [26] E. D. Lopaeva, I. Ruo Berchera, I. P. Degiovanni, S. Olivares, G. Brida, and M. Genovese, Experimental realization of quantum illumination, *Phys. Rev. Lett.* **110**, 153603 (2013).
- [27] Z. Zhang, M. Tengner, T. Zhong, F. N. C. Wong, and J. H. Shapiro, Entanglement’s benefit survives an entanglement-breaking channel, *Phys. Rev. Lett.* **111**, 010501 (2013).
- [28] Z. Zhang, S. Mouradian, F. N. C. Wong, and J. H. Shapiro, Entanglement-enhanced sensing in a lossy and noisy environment, *Phys. Rev. Lett.* **114**, 110506 (2015).
- [29] S. Barzanjeh, S. Guha, C. Weedbrook, D. Vitali, J. H. Shapiro, and S. Pirandola, Microwave quantum illumination, *Phys. Rev. Lett.* **114**, 080503 (2015).
- [30] M. Sanz, U. Las Heras, J. J. García-Ripoll, E. Solano, and R. Di Candia, Quantum estimation methods for quantum illumination, *Phys. Rev. Lett.* **118**, 070803 (2017).
- [31] Q. Zhuang, Z. Zhang, and J. H. Shapiro, Entanglement-enhanced neyman–pearson target detection using quantum illumination, *J. Opt. Soc. Am. B* **34**, 1567 (2017).
- [32] Q. Zhuang, Z. Zhang, and J. H. Shapiro, Quantum illumination for enhanced detection of rayleigh-fading targets, *Phys. Rev. A* **96**, 020302(R) (2017).
- [33] S. Barzanjeh, S. Pirandola, D. Vitali, and J. M. Fink, Microwave quantum illumination using a digital receiver, *Sci. Adv.* **6**, eabb0451 (2020).
- [34] M. Casariego, Y. Omar, and M. Sanz, Bi-frequency illumination: A quantum-enhanced protocol, *Advanced Quantum Technologies* **5**, 2100051 (2022).
- [35] Q. Zhuang and J. H. Shapiro, Ultimate accuracy limit of quantum pulse-compression ranging, *Phys. Rev. Lett.* **128**, 010501 (2022).
- [36] T. Gonzalez-Raya and M. Sanz, Coplanar Antenna Design for Microwave Entangled Signals Propagating in Open Air, *Quantum* **6**, 783 (2022).
- [37] G. Spedalieri, Cryptographic aspects of quantum reading, *Entropy* **17**, 2218–2227 (2015).
- [38] J. L. Pereira and S. Pirandola, Bounds on amplitude-damping-channel discrimination, *Phys. Rev. A* **103**, 022610 (2021).
- [39] A. J. Brady, C. Gao, R. Harnik, Z. Liu, Z. Zhang, and Q. Zhuang, Entangled sensor-networks for dark-matter searches, *PRX Quantum* **3**, 030333 (2022).
- [40] M. M. Marchese, A. Belenchia, S. Pirandola, and M. Paternostro, An optomechanical platform for quantum hypothesis testing for collapse models, *New J. Phys.* **23**, 043022 (2021).
- [41] I. Ruo Berchera, I. P. Degiovanni, S. Olivares, and M. Genovese, Quantum light in coupled interferometers for quantum gravity tests, *Phys. Rev. Lett.* **110**, 213601 (2013).
- [42] M. Genovese, Real applications of quantum imaging, *J. Opt.* **18**, 073002 (2016).
- [43] E. Losero, I. Ruo-Berchera, A. Meda, A. Avella, and M. Genovese, Unbiased estimation of an optical loss at the ultimate quantum limit with twin-beams, *Sci. Rep.* **8**, 7431 (2018).
- [44] G. Spedalieri, L. Piersimoni, O. Laurino, S. L. Braunstein, and S. Pirandola, Detecting and tracking bacteria with quantum light, *Phys. Rev. Res.* **2**, 043260 (2020).
- [45] Q. Zhuang and S. Pirandola, Entanglement-enhanced testing of multiple quantum hypotheses, *Comm. Phys.* **3**, 103 (2020).
- [46] Q. Zhuang and S. Pirandola, Ultimate limits for multiple channel discrimination, *Phys. Rev. Lett.* **125**, 080505 (2020).
- [47] L. Banchi, Q. Zhuang, and S. Pirandola, Quantum-enhanced barcode decoding and pattern recognition, *Phys. Rev. Appl.* **14**, 064026 (2020).
- [48] J. L. Pereira, Q. Zhuang, and S. Pirandola, Optimal environment localization, *Phys. Rev. Res.* **2**, 043189 (2020).
- [49] C. Harney, L. Banchi, and S. Pirandola, Ultimate limits of thermal pattern recognition, *Phys. Rev. A* **103**, 052406 (2021).
- [50] C. Harney and S. Pirandola, Analytical bounds for dynamic multichannel discrimination, *Phys. Rev. A* **104**, 032402 (2021).
- [51] C. Harney and S. Pirandola, Idler-free multi-channel discrimination via multipartite probe states, *npj Quantum Inf.* **7**, 153 (2021).

- [52] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-based learning applied to document recognition, *Proceedings of the IEEE* **86**, 2278 (1998).
- [53] G. Brida, M. Genovese, and I. Ruo Berchera, Experimental realization of sub-shot-noise quantum imaging, *Nat. Photon* **4**, 227 (2010).
- [54] N. Samantaray, I. Ruo-Berchera, A. Meda, and M. Genovese, Realization of the first sub-shot-noise wide field microscope, *Light: Science & Applications* **6**, e17005 EP (2017).
- [55] T. Cover and P. Hart, Nearest neighbor pattern classification, *IEEE Transactions on Information Theory* **13**, 21 (1967).
- [56] K. O'Shea and R. Nash, An introduction to convolutional neural networks, arXiv:1511.08458 (2015).
- [57] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, *et al.*, Recent advances in convolutional neural networks, *Pattern Recognit.* **77**, 354 (2018).
- [58] T. S. Nazaré, G. B. P. da Costa, W. A. Contato, and M. Ponti, Deep convolutional neural networks and noisy images, in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, edited by M. Mendoza and S. Velastín (Springer International Publishing, Cham, 2018) pp. 416–424.
- [59] M. A. Nielsen, *Neural networks and deep learning* (Determination Press, 2015).
- [60] K. P. Murphy, *Machine learning: a probabilistic perspective* (MIT Press, Cambridge, MA, 2012).
- [61] M. Innes, E. Saba, K. Fischer, D. Gandhi, M. C. Rudilosso, N. M. Joy, T. Karmali, A. Pal, and V. Shah, Fashionable modelling with flux, *CoRR abs/1811.01457* (2018), arXiv:1811.01457.
- [62] M. Innes, Flux: Elegant machine learning with julia, *Journal of Open Source Software* 10.21105/joss.00602 (2018).
- [63] G. Brida, I. Degiovanni, M. Genovese, M. L. Rastello, and I. Ruo-Berchera, Detection of multimode spatial correlation in pdc and application to the absolute calibration of a ccd camera, *Optics Express* **18**, 20572 (2010).
- [64] A. Avella, I. Ruo-Berchera, I. P. Degiovanni, G. Brida, and M. Genovese, Absolute calibration of an emccd camera by quantum correlation, linking photon counting to the analog regime, *Optics Letter* **41**, 1841 (2016).
- [65] A. Meda, E. Losero, N. Samantaray, F. Scafirimuto, S. Pradyumna, A. Avella, I. Ruo-Berchera, and M. Genovese, Photon-number correlation for quantum enhanced imaging and sensing, *J. Opt.* **19**, 094002 (2017).
- [66] O. Jedrkiewicz, Y.-K. Jiang, E. Brambilla, A. Gatti, M. Bache, L. A. Lugiato, and P. Di Trapani, Detection of sub-shot-noise spatial correlation in high-gain parametric down conversion, *Phys. Rev. Lett.* **93**, 243601 (2004).
- [67] M. Bondani, A. Allevi, G. Zambra, M. G. A. Paris, and A. Andreoni, Sub-shot-noise photon-number correlation in a mesoscopic twin beam of light, *Phys. Rev. A* **76**, 013833 (2007).
- [68] J.-L. Blanchet, F. Devaux, L. Furfaro, and E. Lantz, Measurement of sub-shot-noise correlations of spatial fluctuations in the photon-counting regime, *Phys. Rev. Lett.* **101**, 233604 (2008).
- [69] I. N. Agafonov, M. V. Chekhova, and G. Leuchs, Two-color bright squeezed vacuum, *Phys. Rev. A* **82**, 011801 (2010).
- [70] J. Peřina, M. Hamar, V. Michálek, and O. c. v. Haderka, Photon-number distributions of twin beams generated in spontaneous parametric down-conversion and measured by an intensified ccd camera, *Phys. Rev. A* **85**, 023816 (2012).