Photon-limited non-imaging object detection and classification based on single-pixel imaging system

Yan Zhu[1] · Jianhong Shi[1] · Xiaoyan Wu[1] · Xialin Liu[1] · Guihua Zeng[1] · Jun Sun[2,3] · Lulu Tian[2,3] · Feng Su[2,3]

Abstract

Under photon-limited detection which is limited by the low-light illumination and short detection time, off-the-shelf classification methods based on clear imaging of the object cannot achieve considerable classification accuracy. To solve this problem, we propose a non-imaging classification method based on single-pixel imaging system. With low-intensity pulsed illumination and time-correlated single-photon counting detection, binarized feature sequence of the objects that need to be classified can be obtained. Combining with a simple machine learning algorithm trained with simulated data based on Poissonian photon detection algorithm, the objects could be classified with considerable accuracy. Proof-of-principle experiments use the MNIST handwriting digit database, showing that up to 90% classification accuracy could be achieved with fewer than 1 detected photon per pixel.

1 Introduction

Photon-limited imaging has significant applications under extreme conditions, such as biological imaging [1, 2], remote sensing [3] and night vision [4, 5]. The conventional imaging system based on a multi-megapixel silicon focal plane would typically obtain an image by capturing of order 10^{12} [6] photons. For many situations, it is very difficult to acquire high-quality image by collecting such a large number of photons because of low-light illumination, limitation of detecting time, long-distance attenuation, and so on.

The automatic classification of objects is a critical issue and has wide applications. Conventionally, the classification is performed by imaging the object first and then combining various algorithms to classify them, such as computer vision system [7, 8]. Those algorithms are directly based on high-quality images. However, under photon-limited conditions, off-the-shelf methods based on images face a big challenge. Nevertheless, while the image is significant for human vision, it is data that really matter to computer or machine visions. Thus, the classification of the object based on the image of the object is not imaging it firstly [9–15]. Single-pixel imaging (SPI) system is a computational imaging system which does not image the object directly [16, 17]. SPI has been demonstrated to be superior over conventional imaging in some applications, such as three-dimensional imaging [16, 18–20], multi-wavelength imaging [21, 22] and X-ray imaging [23, 24]. In SPI system, varying spatial modulated light patterns are employed to illuminate the object scene, and a light intensity sequence is recorded by a single-pixel detector without spatial information. The image could be reconstructed by the correlation of the modulated patterns and detected light intensity signals. Thus, such light intensity sequence could be used as a feature sequence to classify the object and the image reconstruction step could be skipped.

A comparison between imaging-free object classification based on the SPI system and conventional approach is shown in Fig. 1. In photon-limited situation, a special type of single-pixel detector which could response to single photon might have improved performance: lower dark counts, faster timing response and higher detected efficiency [25]. Thus, SPI system could have a better performance than a conventional imaging system under photon-limited conditions [26].

✉ Jianhong Shi
  purewater@sjtu.edu.cn

1   State Key Laboratory of Advanced Optical Communication Systems and Networks, Center of Quantum Sensing and Information Processing, Shanghai Jiao Tong University, Shanghai 200240, China

2   Shanghai Aerospace Control Technology Institute, Shanghai, China

3   Shanghai Key Lab Oratory of Aerospace Intelligent Control Technology, Shanghai, China
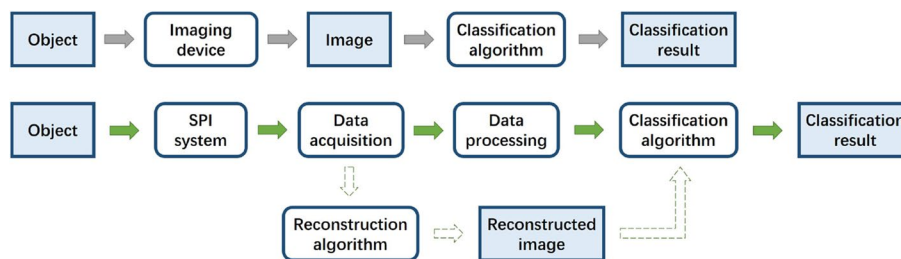
Fig. 1 The process of conventional object classification based on image and imaging-free object classification based on the SPI system: the gray arrows denote the conventional process based on image, and the green arrows represent our proposed process. The dotted arrows denote imaging process which could be skipped if we only concern about the classification of the object

In this paper, based on our previous research results of object recognition [12] and photon-limited detection [26], we propose a non-imaging object classification scheme with photon-limited measurements on the SPI system. The previous work of Ref. [12] cannot directly obtain the multi-classification results of the objects and not consider photon-limited condition, while the scheme proposed in this paper can perform multiclass classification under low-light scenarios. We use the binarized sequence obtained by the Poissonian single-photon detection as the feature of an object and combine with machine learning (ML) algorithm to classify the object with few photon detections. Moreover, since Poissonian single-photon detection is a random process, we generate the training set by computer simulation based on the average photon counts and the approximate background noise. A proof-of-concept experiment is performed with MNIST handwriting digit. A considerable accuracy higher than 90% could be attained with minimum 0.71 photon detections per pixel, in which case the image of the object even could not be well reconstructed. We believe

this scheme could provide a new possibility for object classification in some extreme environments or some special scenarios.

2 Method

2.1 Experimental setup

The schematic diagram of our experimental setup is shown in Fig. 2. A 1 MHz 532 nm intensity tunable pulsed laser illuminates onto a digital micromirror device (DMD) with series programmable patterns. The modulated light patterns then projected onto the object plane by a projection lens. The DMD used in our experiment is a typical spatial light modulator, which consists of an array of 1080×1920 independent addressable micromirrors.

The photons echoed from the object are homogenized by an optical diffuser and detected by a single-photon avalanche diode (SPAD). The output digital signal is then fed into the
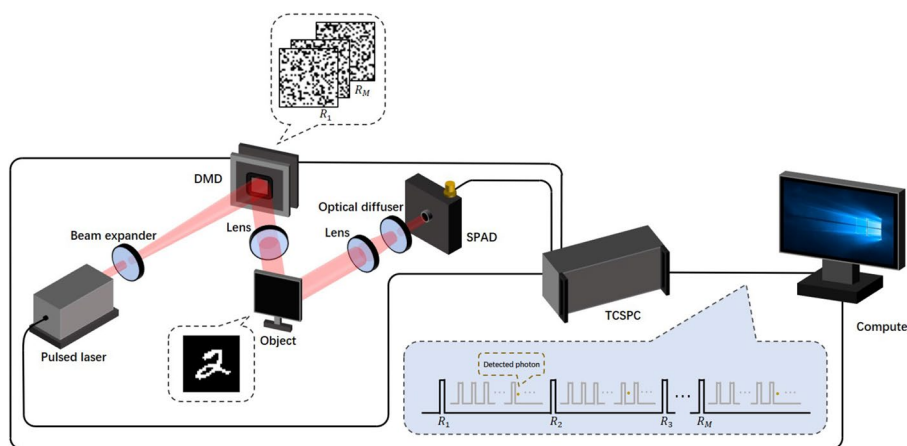

Fig. 2 The experimental schematic diagram. A series of spatial light modulated patterns controlled by computer are loaded on the DMD. After projecting patterns onto the object, echo photons are detected by the SPAD and then fed into TCSPC module, which also receives synchronization signals from pulsed laser and DMD. The arrival time

sequence of the echo photons from the object and the synchronization signals are recorded by TCSPC and are shown in the inner bottom box. The pulses marked by $R_i, i = 1, 2, \ldots, M$ are synchronization signals of DMD

time-correlated single-photon counting (TCSPC) module, which also receives synchronization signals from the pulsed laser and the DMD. The arrival time sequence of the echo photons from the object and the synchronization signals are both recorded as shown in Fig. 2. For each modulated pattern, the number of echo photons from the object is proportional to the inner product of the pattern with the object. In our experiment, the time interval between each pattern is set to 10 ms. Thus, echo photons from 10,000 laser pulses for one modulated pattern are recorded.

Similar as in Ref. [26], a series of sparse binary random patterns are employed as the modulated patterns. The MNIST written number database [27] which consists of 70,000 labeled 28×28 grayscale images of handwritten numbers are used as the target object. Thus, each modulated pattern also has 28×28 pixels, and each pixel is formed by 8×8 micromirror units. And hence only 224×224 units of the DMD are used. To simplify the replacement of objects to implement a large number of measurements, we load the modulated patterns and object simultaneously on DMD. The patterns loaded on DMD are the inner product of the spatial modulated light patterns and the object, and the light path between DMD and the object is omitted.

Based on this experimental setup, we measure the object to be classified and record the detection data. For each object, the same set of 1000 sparse modulated patterns are used and each test object is detected once.

2.2 Data acquisition and processing

In an experimental setup described above, for the ith modulated pattern, the total detected intensity S_i can be represented as [25]

$$S_i = I_{\text{avg}} \iint R_i(x, y) O(x, y) \mathrm{d}x \mathrm{d}y \tag{1}$$

where $R_i(x, y)$ represents the modulated pattern and $O(x, y)$ denotes the reflectivity function of the object. I_{avg} is the average illumination light intensity for an unit surface.

Under low-light pulsed illumination and single-photon detection, I_{avg} is small and the individual photon detection satisfies the Poisson statistics [28]. Assume η represents the detection efficiency, B represents the arrival rate of background photons to the SPAD, and T represents the pulse repetition period [26, 28]. Then the probability of no photon being detected within one single-pulse illumination can be denoted by

$$P_0(S_i) = e^{-\eta(S_i + BT)} \tag{2}$$

Excepting S_i, the parameters depend on the experimental system and detection conditions, so they are settled under one same measurement environment. Different spatial light

modulated patterns or objects might generate difference of S_i, and hence, the probability of no photon being detected could be different.

Since each pulse is independent, the probability of existing k pulses before the first detected photon for one illumination pattern is

$$P[n = k] = P_0(S_i)^{k-1} [1 - P_0(S_i)] \tag{3}$$

The pulse number of the first detected photon for ith pattern can be denoted by n_i. In the absence of background light, the maximum-likelihood intensity estimator, $\hat{S}_i$, is proportional to $1/n_i$ for $n_i \gg 1$, denoted by [26]:

$$\hat{S}_i = \arg \max \log \left\{ e^{-\eta(n_i-1)S} \left(1 - e^{\eta S}\right) \right\} \propto \frac{1}{n_i} \tag{4}$$

Then object image could be reconstructed by the correlation algorithm of $1/n_i$ with the modulated pattern, R_i. Thus, the pulse count, n_i, contains the object information. We record pulse counts for further processing to obtain the feature sequence.

Since the echo photon detection under photon-limited condition is a random Poisson process, even for the same object being illuminated by the same set of spatial light modulated patterns, we might not obtain the same sequence from different detections. In our experiment, each test object is detected once with 1000 illumination patterns. After illuminations of 1000 patterns, a 1000-dimensional pulse counts sequence is obtained. To simplify the pulse count sequence, reduce the computation complexity and improve system efficiency, we binarize this sequence with a pulse number threshold $\bar{n}$. If the first detected photon arrives before the $\bar{n}$th pulse, the pulse count of this pattern is binarized into '1', otherwise '0'.

Assuming an object is illuminated by a series of patterns, a schematic time sequence of the echo photons and the synchronization signals recorded by the TCSPC module and the procedure of data processing are shown in Fig. 3. Assume the number of patterns is M. The pulses marked by $R_i, i = 1, 2, \ldots, M$ are synchronization signals of modulated patterns which is 100 counts/s in our experiment. A total of 10,000 synchronization signals from the pulsed laser between each modulated pattern are recorded. The dark yellow dot denotes the first detected echo photon from the object for each pattern, and the number of pulses before the arrival of this first photon within the ith modulated pattern is denoted by n_i. The pale yellow dots denote successively photons detected after the first photon. Based on such recorded signals, we binarize the pulse counts to obtain the feature sequence. The green arrow and blue arrow denote two different values of threshold. As shown in the figure, one feature sequence of the object could be settled when the pulse number threshold and the length of the sequence
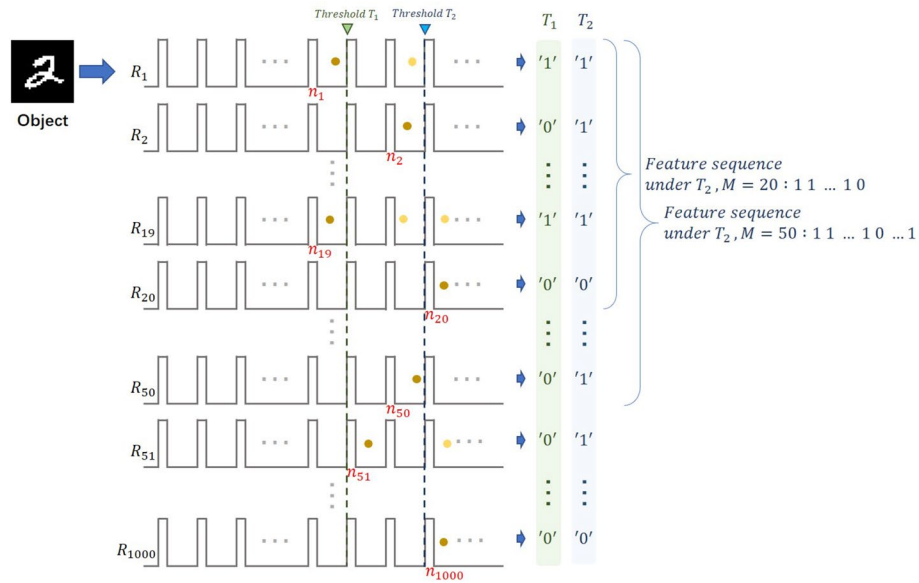
Fig. 3 The time sequence recorded by the SPI system and the process of binarizing pulse sequence with different thresholds and values of M: assuming an object is illuminated by a series of patterns. The pulses marked by $R_i, i = 1, 2, \ldots, M$ are synchronization signals of DMD. During the illumination of each pattern, pulses from the laser and photons measured by SPAD are recorded. The dark yellow dots denote the first detected photons, and the number of pulses before the arrival of first detected echo photon within the ith modulated pattern is denoted by n_i. The pale yellow dots are successively photons detected after the first photon. For each pattern, the pulse count is binarized by a settled threshold. The green arrow and blue arrow denote two different values of threshold

are set. Moreover, one same pulse count sequence might be binarized into different feature sequences under different thresholds and sequence lengths.

The selection of threshold affects the proportion of '1's and '0's in the feature sequence. If the threshold is set big (small), more pulse counts of the sequence would be binarized to '1's ('0's). The length of the feature sequence, which is equaling to the number of measurements (modulated patterns) used for each object, influences the obtained information of the object. In practical application, the length of the sequence and the threshold could be predetermined by the number of modulated patterns, the duration of each pattern and the laser pulse frequency, which determine the total data acquisition time.

2.3 Classification process

An overview of our classification process is depicted in Fig. 4. Random unknown samples that need to be classified are measured by the experimental setup. The pulse counts sequences of these objects are acquired. After the measurements of all these objects, an average number of detected photons per 10,000 pulses which corresponds to the illumination light intensity approximately are acquired. The background noise could be estimated by the detection rate of an all '0's modulated pattern. Based on these data, the pulse counts sequences of training samples could be simulated according to the Poisson detecting
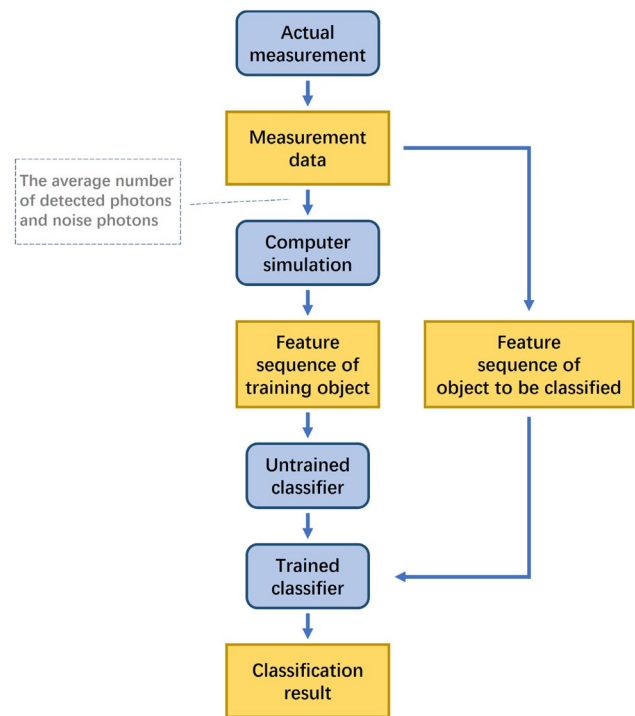

Fig. 4 The classification process of our scheme

process given in Eqs. (2–3). After the binarizing process described above, the feature sequences of these training objects are obtained. While the feature sequences of these

training objects combined with their labels are fed into an untrained classifier for training, a trained classifier is obtained. Then the measured feature sequence of the object to be classified is input into the trained classifier, and the predicted result is obtained.

3 Results

In order to settle the sparsity of random pattern, computer simulation is performed firstly and the change trend of classification accuracy with the increase in sparsity is recorded. As shown in Fig. 5, the accuracy keeps maximum between about 0.002 to 0.014. Thus, the sparsity is set to 0.01, which corresponds to 1% of random '1' among all pixels.

After settling the sparsity, the actual experiments are proceeded. Generally, in the TCSPC module, the condition that at most one photon being detected in the same pulse is considered as photon-limited condition. When the number of photons accounts for less than 5% of pulses, the probability of two photons being detected in the same pulse is extremely low, which can be considered as a photon-limited detection. In our experiment, 400 random samples from MNIST test set are measured by the experimental system. An average of 215 photons are detected in 10,000 pulses for all those samples, and the background noise is 10 photons in 10,000 pulses. Given an average photon counting rate and the background noise rate, pulse counts sequences of 60,000 training samples from the MNIST database are simulated.

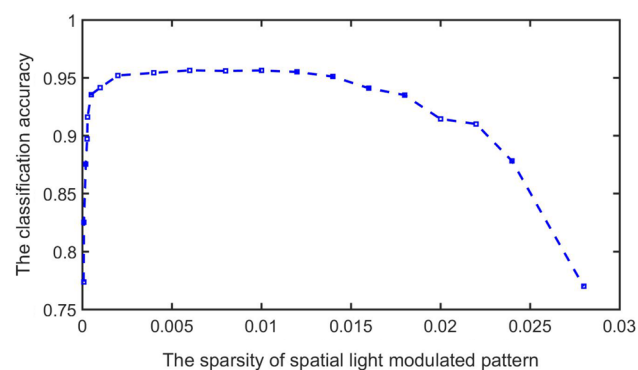At first, a very simple classification algorithm, *k*-nearest neighbor (kNN), is used as the classifier.

3.1 Classification accuracy with different M and thresholds

As discussed in Sect. 2.2, two main parameters affect the performance of this system: the pulse number threshold and length of the feature sequence. The influence of the selection of threshold on classification accuracy is depicted in Fig. 6. It is analyzed by changing the threshold value with several fixed lengths of feature sequence. Figure 7 depicts the relationship of the classification accuracy and the length of feature sequence with several fixed thresholds.

As shown in the above two figures, while fixing the length of feature sequence, classification accuracy increases with the threshold value. When the threshold is small, the improvement is obvious. With the further increase in the threshold, the accuracy improves more and more slowly. However, the threshold cannot be set too big. As shown in Fig. 7, the overall accuracy is higher when the threshold is set 100 than 300. Thus, the threshold value cannot be set too big or too small. Meanwhile, while fixing a threshold and changing the length of the feature sequence *M*, the accuracy increases with the *M*. When a
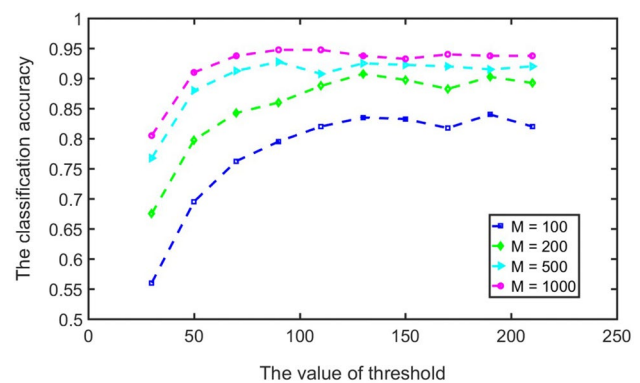

Fig. 6 Classification accuracy with different thresholds

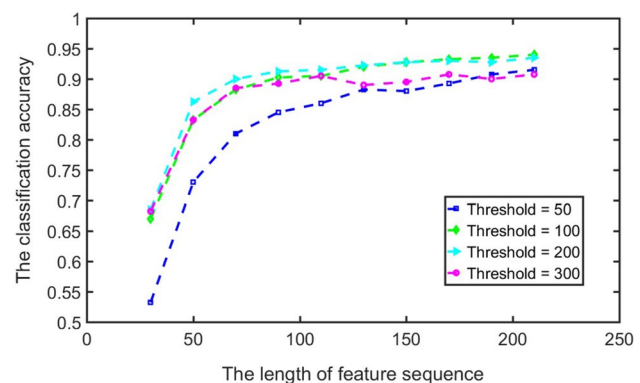**Fig. 5** The change trend of classification accuracy with the increase in sparsity

Fig. 7 Classification accuracy with different lengths of the feature sequence

very limited length of feature sequence is used for classification, the accuracy is relatively low, and the improvement is obvious with the increase in *M*. With the further increase in *M*, the accuracy reaches saturation.

The threshold value and length of the feature sequence are considered together to analyze classification accuracy, which is depicted in Fig. 8.The yellow blocks represent accuracy which is higher than 90%. In the actual experiment, the selection of threshold and length of feature sequence are important factors that affect system efficiency. The threshold represents the pulse number used for each pattern, and the length of the feature sequence represents the number of patterns. The product of these two parameters determines the total data acquisition time of the system. Thus, in the actual application, the classification accuracy and the data acquisition time must be balanced.

We use the number of photons per pixel (PPP) to represent the photon efficiency for classification under photon-limited condition, which can be expressed as

$$\text{PPP} = \frac{N_{\text{ph}}}{N} \tag{5}$$

where N_{ph} is the number of photons measured and *N* is the total number of pixels of the object.

To reach relatively high accuracy in the shortest time possible, we set an accuracy of 90% as a standard. In this case, a minimum value of the threshold is 130 and the fewest number of patterns is 200 in our experiment, and the corresponding number of photons per pixel (PPP) is 0.71. As shown in Fig. 9, the object could be classified correctly while in such case the image of the object even could not be well reconstructed by the first-photon ghost imaging algorithm.
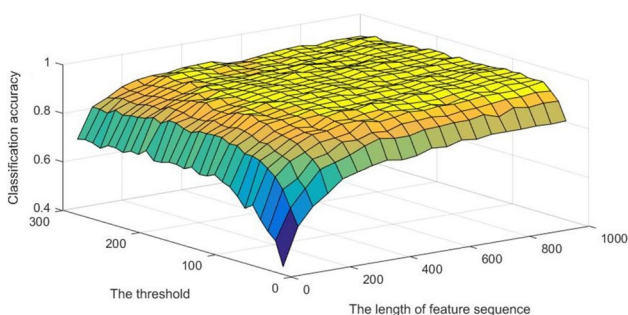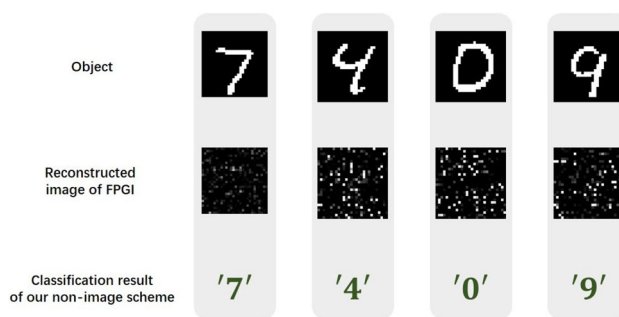

Fig. 9 Reconstructed images using first-photon ghost imaging (FPGI) algorithm comparing with the classification result of our scheme. The images of the first row represent the original objects, and the middle row represents the images retrieved by FPGI based on our experimental data. The last row denotes our classification results without images

3.2 Classification accuracy with different numbers of detected photons

The selection of threshold and *M* jointly determines the number of detected photons. A bigger threshold means a larger number of photons to detect per pattern, and a bigger *M* represents more spatial light modulated patterns being measured for each object. The relationship between the number of detected photons and corresponding classification accuracy is shown in Fig. 10.

As shown in the figure, the overall trend of accuracy increases with the increase in photon number and finally reaches a plateau. However, at the same level of photon number, the accuracy fluctuates. The reason for the fluctuation lies in the different selections of the threshold and *M*. With the same number of detected photons, a bigger threshold means smaller *M*. As discussed in Sect. 3.1, the threshold cannot be set too big or small. Thus, when the threshold is settled too big with a small value of *M* or settled too small with big *M*, the classification accuracy is lower than the situation of the threshold being set moderately. Therefore, within a certain range, the accuracy
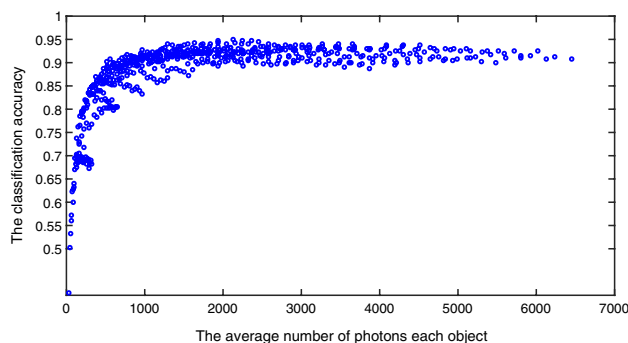

Fig. 8 Classification accuracy with different thresholds and the lengths of the feature sequence. The yellow blocks represent accuracy of more than 90%

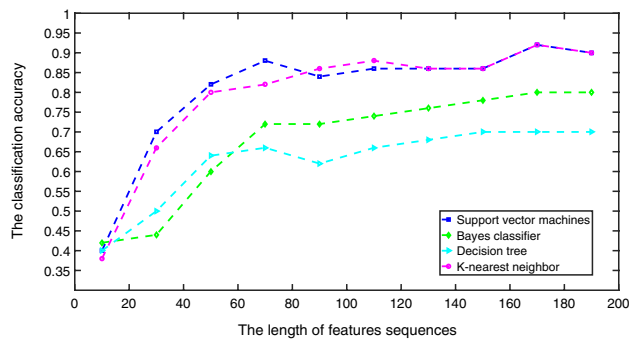**Fig. 10** The classification accuracy with different average numbers of detected photons each object

Fig. 11 The classification accuracy employing support vector machines, Bayes classifier, decision tree and *k*-nearest neighbor respectively

could be improved by increasing the number of detected photons, and to obtain as high accuracy as possible with a fixed level of photon number, the threshold and *M* should be set moderately.

3.3 Classification accuracy with different classifiers

To select an appropriate machine learning algorithm to classify objects, several classical classifiers are employed and compared. The experiment is performed with the threshold fixed 200 and *M* increased from 10 to 200. Four algorithms are used for classification by the feature sequence, respectively support vector machines, Bayes classifier, decision tree and *k*-nearest neighbor. The result is shown in Fig. 11.

As shown in the figure, the accuracy of using the support vector machine (SVM) is almost equal to the *k*-nearest neighbor (kNN) and higher than the other two algorithms. The reason why we chose kNN is that it has less time overhead in our system. The SVM requires hyperplane $wx + b$ to segment data sets, and there would be a model training process to determine the values of *w* and *b*. After settling the *w* and *b*, the predicted result of the test set is determined directly based on this model. Many other machine learning algorithms or deep learning networks have complex learning processes. Meanwhile, kNN is an algorithm called lazy learning. In the training phase, the samples are saved, and the training time is zero. After receiving the test samples, the predicted result is determined by the training set. Thus, when the test set is not very large, kNN prediction efficiency is higher than SVM and other eager learning algorithms. Therefore, to achieve high classification accuracy effectively, kNN is employed in our system.

4 Conclusion

To perform object detection and classification with photon-limited detection, we propose a non-imaging classification method based on a SPI system. This method uses the binarized photon counts sequence obtained by the Poissonian single-photon detection as the feature of an object. In our system, the test object needs to be classified is measured by actual experiment, while the feature sequence of training object is generated by computer simulation based on the measurement conditions of the test set. Combining with simple machine learning (ML) algorithm, objects can be classified. Experimental results demonstrate that our proposed scheme could achieve considerable accuracy efficiently with very limited photon detections, in which case the object image even cannot be well reconstructed.

References

1. T. Vo-Dinh, Y. Fei, M.B. Wabuyele, J. Raman Spectrosc. **36**(6–7), 640 (2005)
2. L. Jian, Z.H. Cheng, F.L. Yuan, Z. Li, Q.C. Li, J.Z. Shu, Trac Trends Anal. Chem. **28**(4), 447 (2009)
3. X.L. Chen, H.M. Zhao, P.X. Li, Z.Y. Yin, Remote Sens. Environ. **104**(2), 133 (2006)
4. A.M. Waxman, A.N. Gove, M.C. Siebert, D.A. Fay, J.E. Carrick, J.P. Racamato, E.D. Savoye, B.E. Burke, R.K. Reich, W.H. Mcgonagle, Proc. SPIE Int. Soc. Opt. Eng. **148**(43), 96 (1996)
5. J.E. Källhammer, Nat. Photon. Sample **12**, 8247 (2006)
6. L. Patrick, L. Xuejun, Y. Xin, Y. Jianbo, K. David, C. Lawrence, S. Guillermo, D.J. Brady, Opt. Express **21**(9), 10526 (2013)
7. P.V. Gehler, S. Nowozin, in *IEEE International Conference on Computer Vision* (2010)
8. D.C. Cireşan, U. Meier, J. Masci, L.M. Gambardella, J. Schmidhuber, Comput. Sci. **32**, 333 (2011)
9. M.A. Davenport, M. Duarte, M.B. Wakin, J.N. Laska, D. Takhar, K. Kelly, R. Baraniuk, Proc. SPIE (2007). https://doi.org/10.1117/12.714460
10. Y. Li, C. Hegde, A.C. Sankaranarayanan, R. Baraniuk, K.F. Kelly, J. Opt. **17**(6), 065701 (2015). https://doi.org/10.1088/2040-8978/17/6/065701
11. S. Lohit, K. Kulkarni, P. Turaga, in *IEEE International Conference on Image Processing*, pp. 1913–1917 (2016) https://doi.org/10.1109/ICIP.2016.7532691
12. H. Chen, J. Shi, X. Liu, Z. Niu, G. Zeng, Opt. Commun. **413**, 269 (2017)
13. G. Satat, M. Tancik, O. Gupta, B. Heshmat, R. Raskar, Opt. Express **25**(15), 17466 (2017)
14. P. Caramazza, A. Boccolini, G. Musarra, M. Hullin, R. Murray-Smith, D. Faccio, in *Computational Optical Sensing and Imaging (Optical Society of America, 2017)* pp. CTu4B–3 (2017)

15. S. Jiao, in *Sixth International Conference on Optical and Photonic Engineering (icOPEN 2018), vol. 10827 (International Society for Optics and Photonics, 2018)* vol. 10827, p. 108271O (2018)
16. M.J. Sun, J.M. Zhang, Sensors **19**, 732 (2019). https://doi.org/10.3390/s19030732
17. M.J. Sun, M. Edgar, G. Gibson, B. Sun, N. Radwell, R. Lamb, M. Padgett, Nat. Commun. (2016). https://doi.org/10.1038/ncomms12010
18. G. Howland, P. Dixon, J. Howell, Appl. Opt. **50**, 5917 (2011). https://doi.org/10.1364/AO.50.005917
19. B. Sun, M. Edgar, R. Bowman, L. Vittert, S. Welsh, A. Bowman, M. Padgett, Science (New York, N.Y.) **340**, 844 (2013). https://doi.org/10.1126/science.1234454
20. M. Edgar, M.J. Sun, G. Spalding, G. Gibson, M.J. Padgett, in *OSA Frontiers in Optics 2016* (2016)
21. V. Studer, J. Bobin, M. Chahid, H. Shams-Mousavi, E. Candes, M. Dahan, Proc. Natl. Acad. Sci. USA **109**, E1679 (2012). https://doi.org/10.1073/pnas.1119511109
22. N. Radwell, K. Mitchell, G. Gibson, M. Edgar, R. Bowman, M. Padgett, Optica **1**, 285 (2014). https://doi.org/10.1364/OPTICA.1.000285
23. J. Cheng, S. Han, Phys. Rev. Lett. **92**, 093903 (2004). https://doi.org/10.1103/PhysRevLett.92.093903
24. A.X. Zhang, Y.H. He, L.A. Wu, L. Chen, B.B. Wang, Optica (2017). https://doi.org/10.1364/OPTICA.5.000374
25. M.P. Edgar, G.M. Gibson, M.J. Padgett, Nat. Photon. **13**, 13 (2018)
26. X. Liu, J. Shi, X. Wu, G. Zeng, Sci. Rep. **8**(1), 5012 (2018)
27. L. Deng, IEEE Signal Process. Mag. **29**(6), 141 (2012)
28. A. Kirmani, D. Venkatraman, D. Shin, A. Colaço, F.N. Wong, J.H. Shapiro, V.K. Goyal, Science **343**(6166), 58 (2014)