



The discriminative ability on anomaly detection using quantum kernels for shipping inspection

Takao Tomono^{1,2,3*}  and Kazuya Tsujimura⁴

*Correspondence:

takao.tomono@ieee.org

¹Graduate school of Media and Governance, Keio University, Fujisawa, 252-0882, Japan

²Graduate school of Science and Technology, Keio University, Yokohama, 223-8522, Japan

Full list of author information is available at the end of the article

Abstract

We aim to use quantum machine learning to detect various anomalies in image inspection by using small size data. Assuming the possibility that the expressive power of the quantum kernel space is superior to that of the classical kernel space, we are studying a quantum machine learning model. Through trials of image inspection processes not only for factory products but also for products including agricultural products, the importance of trials on real data is recognized. In this study, training was carried out on SVMs embedded with various quantum kernels on a small number of agricultural product image data sets collected in the markets. The quantum kernels prepared in this study consisted of a smaller number of rotating gates and control gates. The F1 scores for each quantum kernel showed a significant effect of using CNOT gates. After confirming the results with a quantum simulator, the usefulness of the quantum kernels was confirmed on a quantum computer. Learning with SVMs embedded with specific quantum kernels showed significantly higher values of the AUC compared to classical kernels. The reason for the lack of learning in quantum kernels is considered to be due to kernel concentration or exponential concentration similar to the Baren plateau. The reason why the F1 score does not increase as the number of features increases is suggested to be due to exponential concentration, while at the same time it is possible that only certain features have discriminative ability. Furthermore, it is suggested that controlled Toffoli gate may be a promising quantum kernel component.

Keywords: Small size data; Quantum kernel; SVM; Anomaly detection; Controlled gate

1 Introduction

In recent years, quantum computers that perform calculations using the principles of quantum mechanics have been attracting attention. The development of quantum computers based on various principles is accelerating. Such quantum computers are expected to be used for the following problems: 1) quantum simulation [1–3], 2) quantum cryptography [4–6], 3) mathematical optimization [7–9], and 4) machine learning [10–15]. Machine learning is calculated using linear algebra and matrices. Since quantum computing

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

involves linear algebra and matrix calculations, it has the advantage of being compatible with classical machine learning. Therefore, quantum machine learning is expected to have great advantages, and we are currently conducting research for quantum machine learning.

In machine learning classification, the key is how to form a separation surface in the quantum Hilbert space [15, 16]. The expressive power of the classification space is very important to form a complex separation surface. The expressive power of the quantum kernel space may be superior to that of the classical kernel space [13, 16]. Based on this idea, we would like to effectively utilize the quantum kernel space to solve social issues. One of the social issues is image inspection using machine learning. Image recognition algorithms for such inspections include CNN (convolutional neural network) [17–19], VAE (variational autoencoder) [20–22], GAN (generative adversarial network) [23–25], logistic regression [26–28], random forest [29–31], boosting [32–34], SVM (support vector machine) [35–37] and so on. CNN, VAE, and GAN need expensive computational costs because they use GPUs. Compared to logistic regression, random forest, and boosting, SVM is known to be able to build learning models with less data. Moreover, we can create a complex separation surface by using the kernel trick. [38–40] In this time, we would like to make full use of the expressive power of the quantum kernel space by using the quantum kernel trick [38, 41].

Now, in image inspection, anomaly detection is a very important technology. Industrial products are standardized, and image inspection is simple, but it is not easy to detect various anomalies. In addition, anomaly images are not uniform and there are different types. There have been several reports on the potential of quantum machine learning and quantum kernel estimation to perform image classification.

In our previous work [41–43], we demonstrated high performance (accuracy and F1 score) at the learning model construction stage using small datasets compared to classical machine learning. We conducted experiments by partially applying our learning model to image inspection at factory production sites. As a result, we found that our quantum machine learning produced higher evaluation indices than classical machine learning. Since the data of industrial products at the factory cannot be made public, a trial of constructing a learning model using quantum machine learning was conducted using apples as agricultural products with similar shape characteristics.

Industrial products are quality controlled and numerical data have performance within 3σ . As agricultural products are not standardized, complex separation surfaces are required to construct a learning model, unlike industrial products.

Our goal is to use quantum kernel tricks to build a system that can identify product and equipment anomalies with small amounts of data rather than large amount of data. Products include not only industrial products but also agricultural products. Products must not be shipped with anomalies. Therefore, we believe that the quantum advantage in this field is the ability to identify anomaly and are conducting research in this area.

In this time, we prepared 11 types of quantum kernel to investigate the influence of each two qubits gate. We use apple with internal vine cracks as an anomaly, and we created our own new dataset. Using the dataset, we obtained image information through pre-processing and feature values (attributes) from principal component analysis. Promising quantum kernels were selected based on the relationship between the calculated features (number of principal components) and the evaluation index using a quantum simulator.

After screening with the quantum simulator, we confirmed the value of the evaluation index with a quantum computer. We discuss the effects of the controlled rotation gate and the CNOT gate in the quantum kernel we prepared. Furthermore, we discuss the impact on the performance indicators as the number of features increases. Finally, we propose a promising quantum kernel.

2 Related works

In recent developments of quantum-classical hybrid technologies, we present three significant studies demonstrating the superiority of quantum kernels. Liu Y et al. theoretically demonstrated cases where quantum kernels outperform classical kernels [44]. Specifically, they showed that quantum kernels can recognize particular patterns that classical kernels cannot efficiently learn, and classifiers using kernel estimation with quantum feature space mapping can achieve higher accuracy through richer feature space construction. Hsin-Huang H et al. experimentally demonstrated that quantum kernels exhibit exponentially superior performance compared to classical kernels for specific datasets [13]. Their work particularly showed that quantum kernels can efficiently extract features in high-dimensional data and experimentally confirmed their ability to learn nonlinear patterns that classical kernels cannot capture. Kübler J et al. analyzed the spectral properties of quantum kernels and demonstrated that quantum advantage can be expected when the reproducing kernel Hilbert space (RKHS) of quantum kernels is low-dimensional and contains functions that are classically hard to compute [45]. However, they note that finding appropriate quantum kernels is challenging and often requires exponentially many measurements for kernel evaluation.

In the context of image classification tasks, which present computational challenges due to increasing visual data processing demands in various industries, two significant studies demonstrate the potential superiority of quantum kernels.

Senokosov A et al. proposed two quantum machine learning models [46]. The first, HQNN-Parallel, a hybrid quantum-classical neural network with parallel quantum circuits, achieved high classification accuracy on the MNIST dataset in the noisy intermediate-scale quantum (NISQ) era, surpassing conventional hybrid models with eight times fewer parameters than classical models. Their second proposal, HQNN-Quany, introduced Quanvolutional layers, achieving comparable performance to classical models with only one-fourth of the learning parameters. Aswiga R et al. proposed a quantum machine learning pipeline for breast cancer diagnosis using quantum kernel support vector machines (QKSVM) [47]. After applying six feature extraction methods and linear discriminant analysis (LDA) dimensionality reduction, they compared four different quantum kernel circuits. The combination of LDA-based 2D features with Angle embedding QKSVM achieved the highest classification accuracy, with robustness verified on AWS's quantum simulator "SV1" and the actual QPU "Aspen-M-3". Furthermore, Srikumar M et al. proposed a quantum random forest (QRF) extending the linear quantum support vector machine (QSVM) using quantum kernel estimation (QKE) [48]. They analyzed finite sampling errors using the Nyström-QKE (NQKE) strategy, reducing sampling complexity compared to QKE. Experiments on real-world datasets, including images, demonstrated that QRF achieves higher classification performance with fewer kernel estimations than QSVM.

While these studies demonstrate the potential and theoretical advantages of quantum kernels, several challenges remain for practical implementation. The demonstrated advantages are specific to particular problem settings and datasets, not universally applicable. From a practical perspective, hardware limitations and quantum noise may prevent the full realization of theoretical advantages in actual applications.

Meanwhile, the classical RBF kernel remains effective for many practical problems, with the advantage of straightforward implementation. Based on several papers that show that the quantum kernels are superior to conventional machine learning techniques (such as the RBF kernel), we hypothesized that quantum advantage can also be demonstrated in classifying normal or anomaly in apple. We aim to find quantum kernels that are effective for individual practical problems, similar to the RBF kernel, as well as to search for quantum kernels that can be widely and generally used.

3 Small datasets

We explain the data sets we generated. We received 500 commercially available apples and found 33 anomaly apples. We confirmed that 7% of the apples on the market are apples with an invisible vine crack. A total of 66 normal and anomaly apples were used as a data set, which is different from previous work [41]. We have a lot of normal data (467 pieces). Therefore, we randomly obtained normal data from the 467 pieces and selected 24 pieces as training data and 9 pieces as test data. The number of anomaly data is limited to 33 pieces. We randomly selected 24 pieces as training data and 9 pieces as test data.

The judgement of normal or invisible anomaly apples is predicted using equipment shown in Fig. 1. We get these apples from the market. We take pictures after illuminating the LED from the bottom of the apple. Then, after image processing, we obtain binary images. To know the internal situation, we cut them in half with a knife. As invisible anomaly, there are apples with browning inside and apples with vine cracks. There are two types of normal apples: normal apples with nothing inside (0) and browning apples (0*). There are also two types of anomaly apples: apples with vine cracks only (1) and apples with vine cracks and browning (1*). In other words, there are four types. Each speech bubble is an enlarged view.

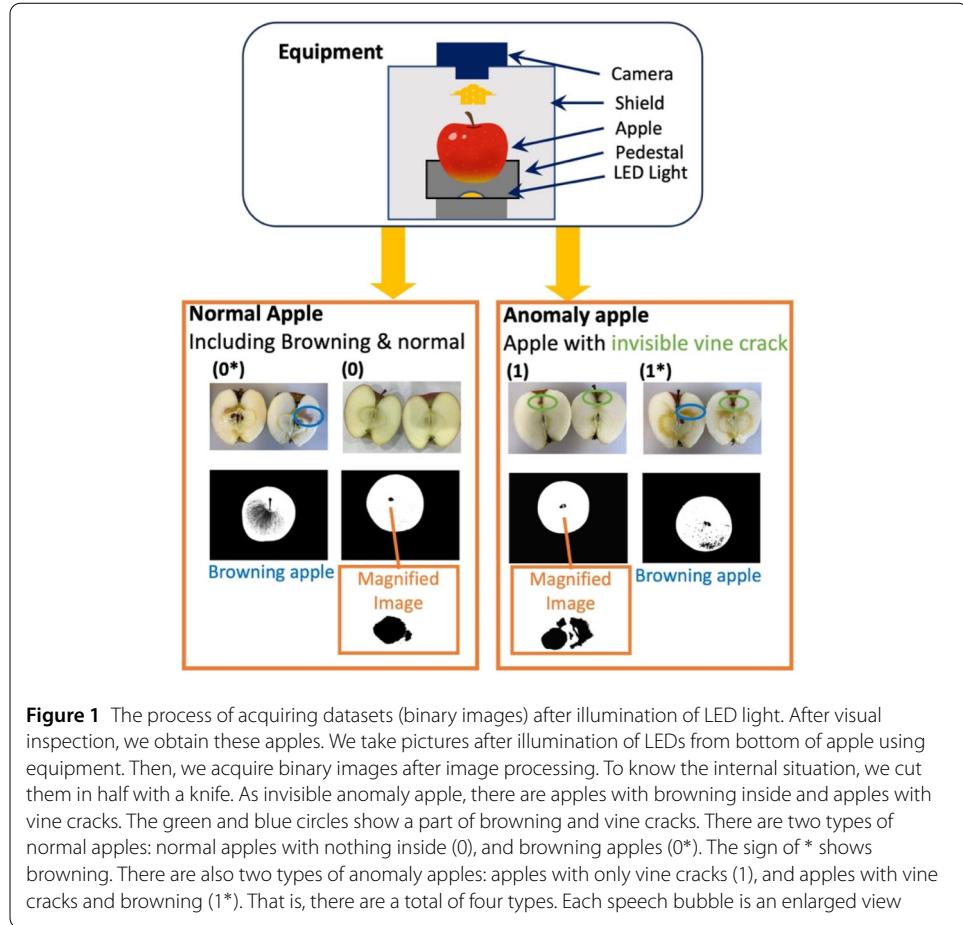
The resolution of the photographed image is 4032 x 3024 and it is necessary to detect patterns other than the vine that appear in an area of 120 x 80. The judgement must be made in an area of about 3% of the total image. The 4032 x 3024 resolution is too large to create a learning model. The image size is reduced to 1/10. The learning model is built using the image size (403 x 302). The image is used to distinguish between normal and anomaly for training.

4 Quantum kernel and step of classification

By means of a nonlinear mapping $\varphi(x)$ embedding the data into the quantum feature space, it can be expressed in the feature space as follows.

$$\kappa(x_i, x_j) = \left| \left\langle \phi(x_j)^\dagger \middle| \phi(x_i) \right\rangle \right|^2 \quad (1)$$

First, we prepare quantum state $\varphi(x) = U(0)|0\rangle$ as a data encoding from classical to quantum data. Second, we prepare $U(x_j)^\dagger U(x)|0\rangle$ as the initial state of the quantum circuit



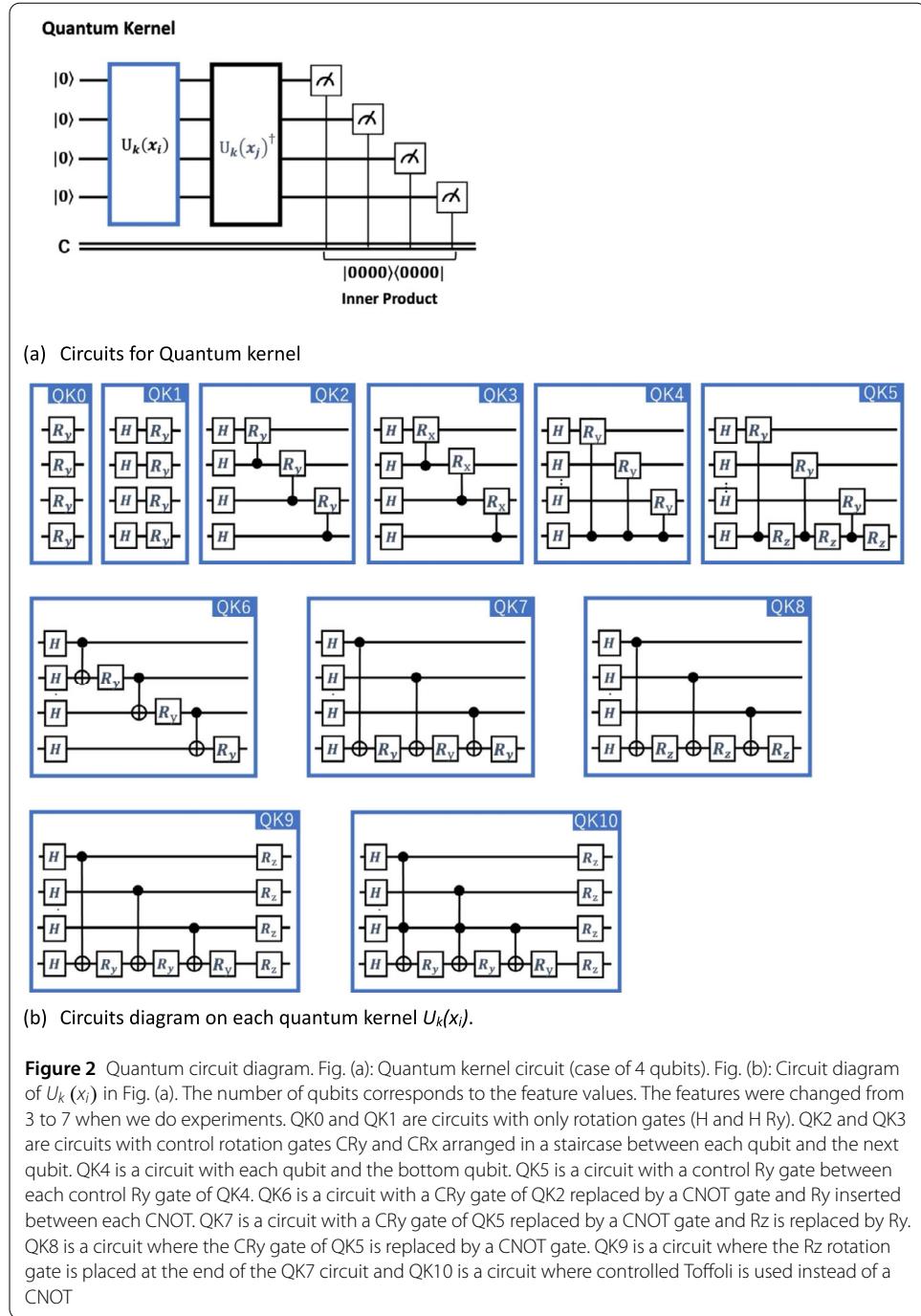
to obtain the inner product $\kappa(x_i, x_j)$. The probability of measuring on $|0\rangle$ for all qubits is as follows.

$$\langle 0 | U^\dagger(x_i) U(x_j) | 0 \rangle \langle 0 | U^\dagger(x_j) U(x_i) | 0 \rangle = \langle \phi(x_i) | \phi(x_j) \rangle \langle \phi(x_j) | \phi(x_i) \rangle \quad (2)$$

Here $U(x)$ is the inner product between the quantum encoded data using quantum kernel estimation. Each feature map is then embedded into the inner product to optimize the parameters. The matrix component of the entire Gram-matrix is obtained from a combination of the inner product. The parameters of the kernel estimation are optimized using rotation gates with/without entanglement in Eq. (2). We use RBF as classical kernel of SVM in this work.

Figure 2 shows quantum kernel circuits diagram when features volume is 4 (the number of qubits is 4). Figure 2(a) shows quantum kernel circuits diagram. Figure 2(b) shows the details of $\phi(x_i)$ as equation (1) and (2).

QK0 and QK1 are circuits with only rotation gates (H and H Ry). QK2 to QK10 were prepared to create efficient quantum kernels with fewer gates than hardware efficient embedding (HEE) [49, 50]. QK2 and QK3 are circuits that place controlled Ry and Rx gates between each qubit and the next qubit in a staircase pattern. QK4 is a circuit that places a controlled Ry gate between each qubit and the bottom qubit. QK5 is a circuit in which Rz is inserted between the Ry control gates of QK4. QK6 is a circuit in which CNOT gates



are arranged in a staircase pattern with Ry inserted between each CNOT gate. QK7 is a circuit in which each qubit and the lowermost qubit are connected by a CNOT gate and Ry is inserted between each CNOT gate. QK8 is a circuit in which a Rz gate is inserted instead of the Ry gate in QK7. QK9 is a circuit in which Rz is placed on each qubit at the end of the quantum circuit in QK7. QK10 is a circuit in which controlled Toffoli gate, is placed in place of each CNOT in QK9. The experiment was essentially carried out using a quantum simulator (ibm_qasm_simulator). For final confirmation, we used a real quantum computer (ibmq_Osaka of the IBM Quantum Platform).

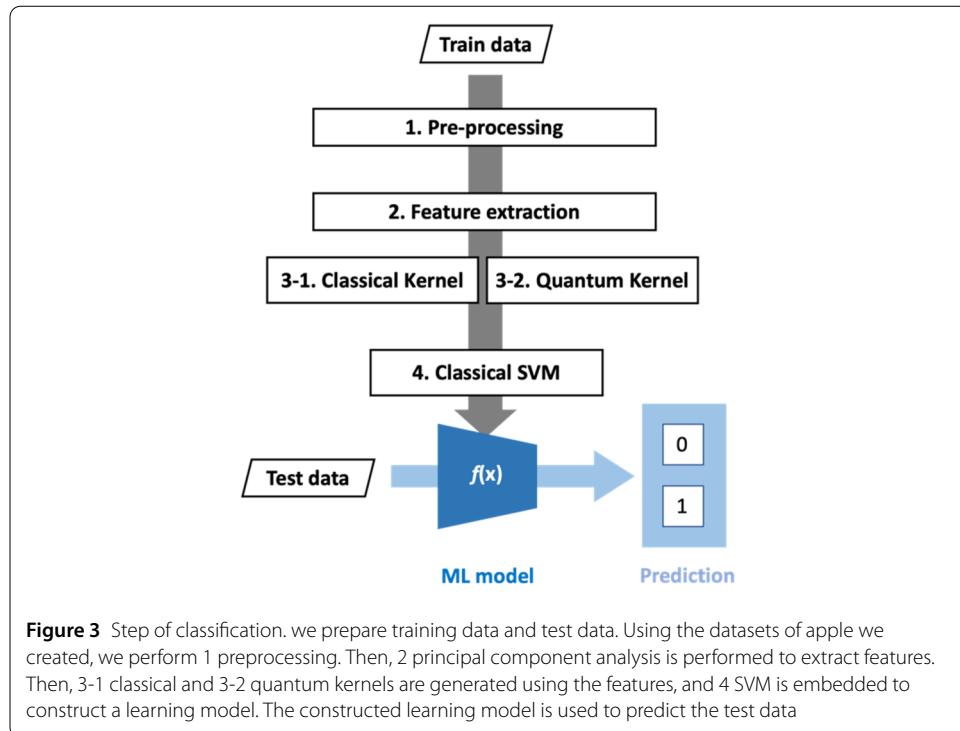


Table 1 Principal Components (PC), Contribution Ratios (CR), and Cumulative Contribution Ratios (CCR) obtained by performing principal component analysis on this image. PC stands for principal components. CR stands for contribution ratios. CCR stands for cumulative contribution ratios

PC	1	2	3	4	5	6	7	8	9	10
CR	0.4390	0.1210	0.0593	0.0423	0.0337	0.0278	0.0210	0.0207	0.0176	0.0168
CCR	0.4390	0.5600	0.6193	0.6616	0.6953	0.7231	0.7441	0.7648	0.7824	0.7992

Figure 3 shows step of classification. we prepare training data and test data. Using the datasets of apple we created, we perform 1 preprocessing. Then, 2 principal component analysis is performed to extract features. Then, 3-1 classical and 3-2 quantum kernels are generated using the features, and 4 Classical SVM embedded kernel is performed to build a learning model. The learning model is used to predict the test data.

5 F1 score on quantum simulator

Table 1 shows the contribution ratio (CR) and cumulative contribution ratio (CCR) obtained by principal component analysis of the image. The first to tenth principal components are obtained. Although there is no clear criterion for the cumulative contribution ratio (CCR), a rough image can be reproduced if the CCR is 0.5 or higher; a CCR of 0.85 or higher is sufficient to reproduce the image. The cumulative contribution ratio (CCR) of principal components 4, 7 and 10 are 0.662, 0.744 and 0.799 respectively. Confirming the reproducibility of the anomaly apple images, it is considered that the first, second and third principal components mainly represent the external shape, browning of the apple and internal vine crack. Here, the number of principal components corresponds to features value.

Figure 4 shows the relationship between the F1 scores for the principal components of each quantum kernel and the conventional kernel RBF. Figure 4(a) shows the results for

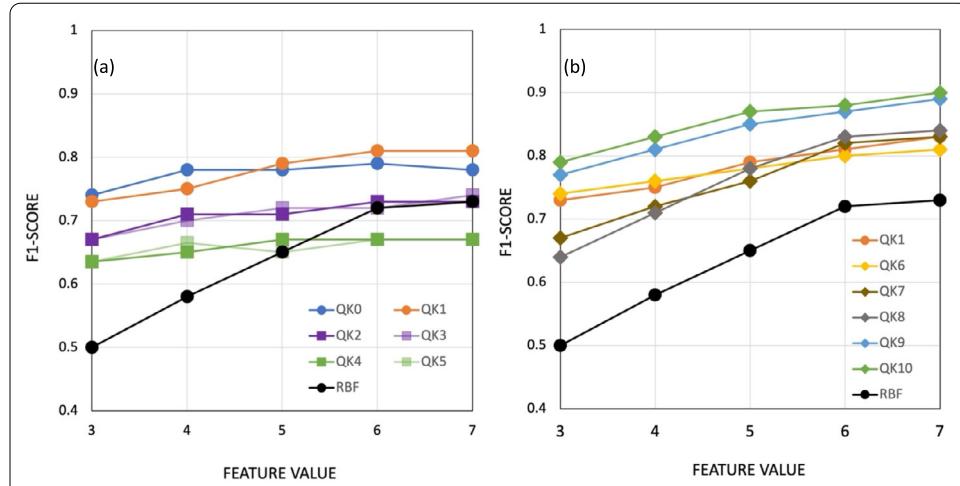


Figure 4 The relationship between Features and F1-score for each quantum kernel compared to classical kernel RBF. Figure 4(a): the relation from QK0 to QK5 compared to classical kernel RBF. Figure 4(b): the relation from QK6 to QK10 compared to classical kernel RBF

the classical kernel RBF and QK0-QK5. The horizontal axis is the feature corresponding to the cumulative contribution of the principal components. The vertical axis is the F1 score. If the feature value is 3, it means the cumulative contribution of the first to the third principal component. If the feature value is 7, it means the cumulative contribution of the first to the seventh principal component. First, the F1 scores of each kernel are compared when the feature value is 3, and then the trends in the F1 scores are compared as the feature size increases.

QK0 and QK1 are quantum circuits with one H-gate, one H-gate and one rotating gate Ry for each qubit. Their F1 value is 0.2 higher than the F1 score of the classical kernel RBF. However, when the feature value increased from 3 to 7, the value of the F1 score increased by 0.1.

QK2 and QK3 are quantum circuits with a staircase structure of control gates Ry and Rx. The F1 score of QK2 and QK3 are more than 0.1 larger than the F1 score of RBF, but 0.05 smaller than the F1 score of QK0. QK4 is a quantum circuit with a controlled rotation gate between each qubit and the bottom qubit. The F1 scores of QK4 and QK5 are more than 0.1 larger than the F1 score of QK0, QK1, QK2 and QK3, but 0.1 smaller than the F1 score of QK2. The F1 scores for QK2, QK3, QK4 and QK5 remained almost constant as feature value increased from 3 to 7.

From the above, it can be considered that the quantum kernel has greater discriminative power than the classical kernel when the feature value is 3. However, the reason why the F1 score remains almost constant even as the feature value increases is because the images represented by the fourth to seventh principal components have little effect on the vine crack, and therefore the value of the F1 score is considered to be unaffected.

Figure 4(b) compares the relationship of QK6 to QK10 with the classical kernel RBF. For a feature value of 3, the F1 score of QK7 and QK8 are more than 0.15 larger than the RBF. The QK6 is about 0.1 larger than QK7 and QK8. Furthermore, the F1 score of QK9 and QK10 are more than 0.3 larger than the classical RBF. The F1 score of QK6 is almost the same as that of QK1. The behavior of QK6 is similar to that of QK1.

QK6 has a staircase CNOT gate. Except for the first qubit, it has a Ry rotation gate from the second qubit onwards similar to QK1. Therefore, the effect of the CNOT gates on the F1 score was not significant in the case of QK6. For QK7, QK8, QK9 and QK10, the F1 score becomes larger as the feature value increases, with the F1 score for QK7 and QK8 increasing by 0.2 and the F1 score for QK9 and QK10 increasing by 0.1. This increase could be considered a dominant difference, but it was equal to or less than that of the classical kernel.

From the above it can be concluded that the CNOT gate is dominant except for QK6. Throughout the apple, including external shape, browning and internal vine crack, the quantum kernel is considered to have a higher discriminative capacity for anomalies. The F1 scores of QK9 and QK10 are the highest among these quantum kernels and are considered promising quantum kernel candidates.

6 Classical kernel vs quantum kernel

After obtaining the above results using the simulator, we use `ibm_Osaka` to confirm the behavior of true quantum computer. Figure 5 shows the ROC-AUC curve (step-shaped curve) of quantum computer compared to classical computer and quantum simulator. Here, QC and QS means quantum computer and simulator. The left figure shows the ROC-AUC curve of RBF, QS and QC on QK9, and the right figure shows the ROC-AUC curve of RBF QS and QC on QK10. As reference data, we also draw the position of random model (black dashed line) and ideal learning model (red dashed line). The dashed line from False Positive Rate (FPR) = 0, True Positive Rate (TPR) = 0 to FPR = 1, TPR = 1 indicates the random model. Axis of FPR = 0 and TPR = 1 indicates ideal learning model. The AUC of classical RBF is drawn near the random model and the numerical data is 0.62. on the other hand, for QK9, the behavior of the AUC curve for the quantum computer was the same as that for the quantum simulator. The numerical data of AUC at that time were both 0.90, as shown in the figure.

For QK10, the behavior of the ROC-AUC curve for the quantum computer was lower than that of the classical computer. As shown in the figure, the numerical data of AUC value at that time was 0.89 for the quantum simulator and 0.59 for the quantum computer.

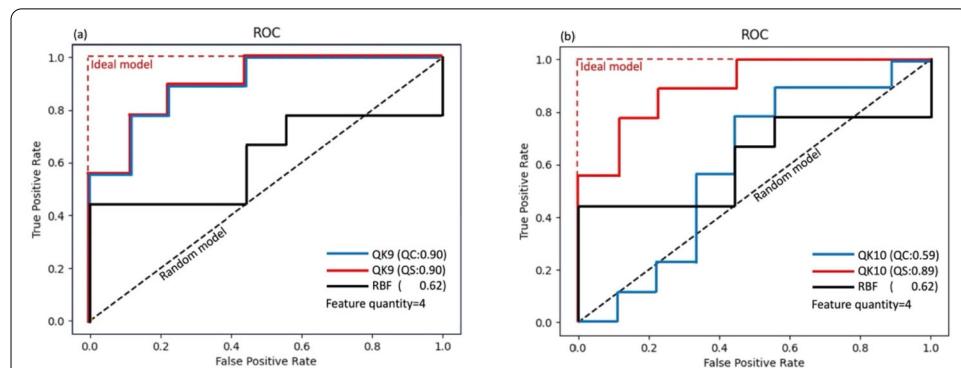


Figure 5 ROC-AUC curve. The horizontal axis is the False Positive Rate (FPR), and the vertical axis is the True Positive Rate (TPR). The normal learning model construction process starts with learning from a random model at the position of the black dashed line (position of the black dashed line) and moves towards the ideal learning model (position of the red dashed line). Features (feature value, feature quantity) is 4

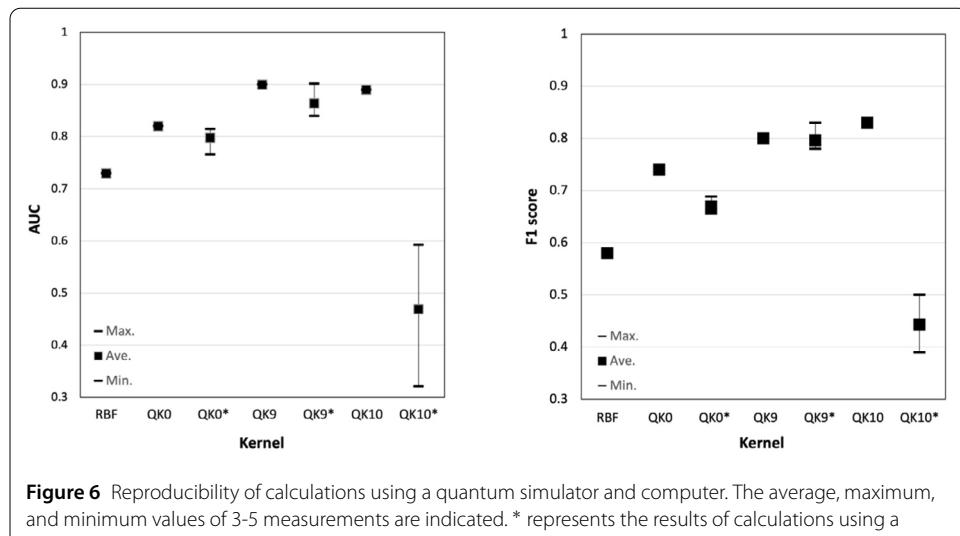


Figure 6 Reproducibility of calculations using a quantum simulator and computer. The average, maximum, and minimum values of 3-5 measurements are indicated. * represents the results of calculations using a quantum computer. Features (feature value, feature quantity) is 4

Table 2 The depth of circuits for QK0, QK1, QK9 and QK10. We use `ibmq_Osaka` as quantum computer when feature value is 4

Quantum kernel	QK0	QK1	QK9	QK10
Depth of circuits	4	4	32	273

For QK10, the behavior and numerical data on quantum computer were significantly different from those on the quantum simulator.

To confirm the reliability of the numerical data, measurements were performed at least 3 times on the quantum computer. In addition, when measurements were performed 5 times for classical computer, it was found that the error was within approximately 2 %.

Figure 6 shows reproducibility of calculations using a quantum simulator and computer. We compare with the numerical data of AUC and the F1 score for the classical kernel RBF and the quantum kernels QK0, QK9, and QK10. The quantum kernels QK0, QK9, and QK10 were measured, and the maximum, minimum, and average values were calculated. For quantum kernels other than QK10, there appears to be little difference in the values of each evaluation index (AUC and F1-score) between the quantum simulator and the quantum computer.

To investigate the cause of the difference between the quantum simulator and the quantum computer in QK10, the circuit depth of the quantum circuit was investigated, and the results are shown in Table 2. Since there was no difference between the quantum simulator and the quantum computer in QK0, QK1, and QK9, it is believed that there is no problem with a circuit depth of up to 32, but since a problem occurred at a circuit depth of 273, it is believed that errors accumulated between 32 and 273. This confirmed that the quantum circuit depth affected the generation of errors in the quantum computer.

Table 3 presents the performance metrics including precision, recall, F1-score, accuracy, and AUC for the RBF kernel and quantum kernels (QK0, QK9, and QK10). For the quantum kernels, we show values obtained from both simulator (IBM_Qiskit) and quantum computer (`ibm_osaka`) implementations. In this analysis, there are nine samples each for normal and anomaly cases. When analyzing these performance metrics, it is crucial to note that Recall and Precision have a complementary relationship. In our analysis, we

Table 3 The performance metrics of QK0, QK9, QK10 and RBF. Here, 0 and 1 represent for normal and anomaly in [0 (1)]. We use `ibmq_Osaka` as quantum computer when feature value is 4. Normal data: 24 training data, 9 test data. Anomaly data: 24 training data, 9 test data

Kernel	RBF.	QK0.	QK0*.	QK9.	QK9*.	QK10.	QK10*
Precision [0 (1)]	0.58 (0.58)	0.74 (0.74)	0.67 (0.67)	0.79 (0.79)	0.80 (0.80)	0.80 (0.88)	0.50 (0.50)
Recall [0 (1)]	0.58 (0.58)	0.74 (0.74)	0.67 (0.67)	0.79 (0.79)	0.80 (0.80)	0.89 (0.78)	0.44 (0.56)
F1-score [0 (1)]	0.58 (0.58)	0.74 (0.74)	0.67 (0.67)	0.79 (0.79)	0.80 (0.80)	0.84 (0.82)	0.47 (0.53)
Accuracy	0.58	0.74	0.67	0.79	0.80	0.83	0.50
AUC	0.62	0.82	0.80	0.90	0.90	0.89	0.59

observe that the values for normal and anomaly cases are identical across recall, precision, F1-score, and accuracy metrics for RBF, QK0, QK0*, QK9, and QK9*, indicating balanced performance. While QK10 and QK10* show slight variations between normal and anomaly cases, their average values remain consistent across all metrics, demonstrating maintained balance. These results confirms that balanced learning models were constructed across all kernel implementations, including the RBF kernel and quantum kernels (QK0, QK9, and QK10).

From the above, we confirmed that the results of the quantum simulator and the quantum computer are almost the same. We also confirmed that as the depth of the quantum circuit increases, the quantum computer becomes noisy, and the calculation becomes difficult at the circuit depth of QK10. We also found that QK9 is a promising quantum kernel.

7 Discussion

Based on the experimental results, the features obtained after feature extraction by principal component analysis correspond to each qubit. As we described already above, the first principal component obtained by the principal component analysis is considered to be the external shape of the apple, the second principal component is the browning, and the third principal component is the internal vine crack. Here, we assume that overfitting does not occur due to the small size of the data set.

When feature value is 3 (the cumulative contribution up to the third principal component is integrated), the quantum kernel was larger by more than 0.3 compared to the classical kernel. As shown in Figs. 4(a) and 5(b), when the feature value is 3, the difference between the quantum and classical kernels is large, indicating that the quantum kernel is larger than classical kernel. On the other hand, as the feature size increases, the difference between the F1 scores of the quantum and classical kernels becomes smaller. This means that the advantage of the quantum kernel is less as the feature value increases.

As shown in equation (1), the quantum kernel is represented by an inner product, and from the inner product of vectors, a Gram matrix is generated. The space created by this Gram matrix generates a complex separating boundary surface. (x_i, x_j) are the coordinates of the Gram matrix and $\kappa(x_i, x_j)$ are the measurements of the last individual quantum bit calculated in depth. The number of matrices in the Gram matrix is determined by the number of qubits in the quantum circuit. The gate operation (unitary called U) determines the height, and $\kappa(x_i, x_j)$ create energy gradients such as peaks, valleys and plains, and the separation boundary is determined as the result. Usually, in principal component analysis, as the cumulative contribution increases (feature value increases) in classical kernel learning, the image reproducibility improves, and the discriminative power become larger.

When the feature size is 3, the F1 score with the quantum kernel is sufficiently larger than the F1 score with the classical kernel. On the other hand, as the feature size increases, the

difference between the F1 score by the quantum kernel and the F1 score by the classical kernel becomes smaller. This could be caused by a phenomenon similar to Barren Plateaus (BPs) [51–56] due to exponential concentration. On the other hand, it is possible that, perhaps, machine learning with quantum kernels may provide better discriminative power with respect to certain anomalies.

Experimental results show that controlled Toffoli gates are promising gates for quantum kernel circuits. The controlled Toffoli gate is a gate that can be decomposed into Hadamard gates, phase gates and a control NOT gates. As shown in Table 2, the depth of the QK9 circuit is 32, while that of the QK10 circuit is 273. the depth of the QK10 circuit is 8.5 times larger than that of the QK9 circuit. In this study we performed the calculations on a superconducting quantum computer, but the calculations failed due to noise and other factors. The reason for this is thought to be the topology of the superconducting quantum computer, which utilizes many swap gates, which makes the depth of the circuit deeper. However, an all-coupled ion-trap quantum computer could be computed without the use of swap gates. This has the advantage that it is likely to be able to compute with reduced errors. On the other hand, the size of the Gram matrix increases in learning (training) using data, which poses a challenge to the execution of calculations. In the case of ion-trap quantum computers, the coherence time is long, so if the size of the Gram matrix can be reduced in future calculations, the control Toffoli gate will be a promising gate with high expectations.

In this application side to quantum computer, the F1 score of QK10 using controlled Toffoli gate shows a sufficiently large F1 score in the quantum simulator, so it is thought to be very promising in practical situations where error correction becomes possible.

8 Conclusion and outlook

We attempted to detect anomaly for apple with internal vine cracks, which are real data with high similarity to image data from the factory. Using the quantum kernel trick, it was demonstrated that a learning model could be built to detect a single anomaly (an internal vine crack) from 24 training data. Ten different quantum kernels were used, and performance index were evaluated according to feature value. We used a quantum simulator and a computer to examine F1 scores and finally evaluate the construction of a learning model using AUC. The results showed that the F1 score was greater with the CNOT gate than with the control rotation gate. The SVM with quantum kernel has a larger F1 score and better discriminative ability than the SVM with classical kernel. Therefore, in anomaly detection, discriminative ability is considered a quantum advantage on SVM embedded quantum kernel. As the feature value increases, the increase in the F1 score of the quantum kernel was smaller than that of the classical kernel.

Within the quantum kernel, a quantum kernel circuit (QK9) with a CNOT gate connecting each qubit to the bottom qubit, a Ry rotation gate inserted between CNOTs and an Rz gate at the end of each qubit of the quantum circuit was found to be promising. If the memory cost of the Gram matrix is reduced and fault-tolerant quantum computers are developed, controlled Toffoli gates would be a promising quantum circuit.

The practical application of these technologies is still in its early stages. In the agricultural application, there are two reports. Amin J et al. reported on high-precision pest localization using YOLOv5, pest classification using QCN, and achieved high accuracy with reduced computational costs compared to existing CNN models [57]. Aasim M et al.

conducted a detailed study on optimizing LED light intensity for in vitro propagation of black mulberry [58]. SVC Quantum kernel showed promising results on some data, but did not show consistent improvements over traditional SVC.

In the manufacturing application, there are two significant studies. Choi E et al. implemented anomaly detection during processes using QSVM and QCNN [59]. Their results demonstrated that QML could achieve accuracy comparable to classical machine learning while exponentially reducing the number of parameters. Tscharke K et al. proposed a semi-supervised anomaly detection method using QSVR with quantum kernels and compared it with classical Support Vector Regression [60]. Their experimental results proved QSVR's superiority by showing higher anomaly detection accuracy than both classical SVR and QAE.

Since production and inspection process costs are reflected in product prices, new learning models must improve accuracy while maintaining or reducing computational costs. Our research focuses on quantum kernels while exploring quantum algorithms that achieve high accuracy with reduced computational cost. As previously mentioned, quantum kernels are utilized not only in SVM but also in QSVR, QCN, QCNN, and other applications. I would like to build a theory of quantum kernels while responding to actual problems and build quantum algorithms that are useful for social implementation.

Abbreviations

SVM, Support vector machine; GPU, Graphics Processing Unit; CNN, Convolutional neural network; VAE, Variational autoencoder; GAN, Generative adversarial network; LED, Light Emitting Diode; RBF, Radial Basis Function; F1-score, Harmonic mean of conformity and repeatability; AUC, Area Under the Curve and AUC takes values between 0 and 1, with values closer to 1 indicating higher discriminatory power; CR, Contribution Ratio; CCR, Cumulative Contribution Ratio; FPR, False Positive Rate; TPR, True Positive Rate; QCN, Quantum Convolutional Network; QSVM, Quantum SVM; QCNN, Quantum Convolutional Neural Network; QAE, quantum AutoEncoder; YOLOv5, You Only Look Once v5 (one of an object detection algorithm).

Acknowledgements

We thank IBM to support free trial of quantum computer. We thank Mr.godo of Toppan digital to provide binarized image data of apple.

Author contributions

T.T. wrote the main manuscript text, T.T. designed research. T.T. made basic quantum software based on quantum algorithms. K.T. conducted revised quantum software and calculate performance indices by classical and quantum computation. All authors contributed to the discussion and All authors reviewed the manuscript.

Funding

Takao Tomono is supported by the Center of Innovations for Sustainable Quantum AI (JST), Grant No. JPMJPF2221.

Data Availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare no competing interests.

Author details

¹Graduate school of Media and Governance, Keio University, Fujisawa, 252-0882, Japan. ²Graduate school of Science and Technology, Keio University, Yokohama, 223-8522, Japan. ³Keio University Sustainable Quantum AI Center, Keio University, Minato, 108-8345, Japan. ⁴Digital Innovation Division, TOPPAN Holdings, Tokyo, 110-8560, Japan.

References

1. Kassal I, Whitfield JD, Perdomo-Ortiz A, Yung M-H, Aspuru-Guzik A. Simulating chemistry using quantum computers. *Annu Rev Phys Chem*. 2011;62:185–207. <https://doi.org/10.1146/annurev-physchem-032210-103512>.
2. Daley A, Bloch I, Kokail C, Flannigan S, Pearson N, Troyer M, Zoller P. Practical quantum advantage in quantum simulation. *Nature*. 2022;607:667–76. <https://doi.org/10.1038/s41586-022-04940-6>.
3. Hempel C, et al. Quantum chemistry calculations on a trapped-ion quantum simulator. *Phys Rev X*. 2018;8:031022. <https://doi.org/10.1103/PhysRevX.8.031022>.
4. Ralegankar VK, et al. Quantum cryptography-as-a-service for secure UAV communication: applications, challenges, and case study. *IEEE Access*. 2022;10:1475–92. <https://doi.org/10.1109/ACCESS.2021.3138753>.
5. Subramani S, Selvi M, Kannan A, Santhosh Kumar S. Review of security methods based on classical cryptography and quantum cryptography. *Cybern Syst*. 2023;56:302–20. <https://doi.org/10.1080/01969722.2023.2166261>.
6. Dam DT, Tran TH, Hoang VP, Pham CK, Hoang TT. A survey of post-quantum cryptography: start of a new race. *Cryptography*. 2023;7:40. <https://doi.org/10.3390/cryptography7030040>.
7. Alexandru CM, et al. Quantum speedups of some general-purpose numerical optimisation algorithms. *Quantum Sci Technol*. 2020;5:045014. <https://doi.org/10.1088/2058-9565/abb003>.
8. Xin G, Wang P, Jiao Y. Multiscale quantum harmonic oscillator optimization algorithm with multiple quantum perturbations for numerical optimization. *Expert Syst Appl*. 2021;185:115615. <https://doi.org/10.1016/j.eswa.2021.115615>.
9. Wang L, Tang F, Wu H. Hybrid genetic algorithm based on quantum computing for numerical optimization and parameter estimation. *Appl Math Comput*. 2005;171:1141–56. <https://doi.org/10.1016/j.amc.2005.01.115>.
10. Khan TM, Robles-Kelly A. Machine learning: quantum vs classical. *IEEE Access*. 2020;8:219275–94. <https://doi.org/10.1109/ACCESS.2020.3041719>.
11. Guan W, et al. Quantum machine learning in high energy physics. *Mach Learn: Sci Technol*. 2021;2:011003. <https://doi.org/10.1088/2632-2153/abc17d>.
12. Jerbi S, et al. Quantum machine learning beyond kernel methods. *Nat Commun*. 2023;14:517. <https://doi.org/10.1038/s41467-023-36159-y>.
13. Huang HY, et al. Power of data in quantum machine learning. *Nat Commun*. 2021;12:2631. <https://doi.org/10.1038/s41467-021-22539-9>.
14. Schuld M, Sinayskiy I, Petruccione F. An introduction to quantum machine learning. *Contemp Phys*. 2015;56:172–85. <https://doi.org/10.1080/00107514.2014.964942>.
15. Schuld M, Killoran N. Quantum machine learning in feature Hilbert spaces. *Phys Rev Lett*. 2019;122:040504. <https://doi.org/10.1103/PhysRevLett.122.040504>.
16. Liu Y, Arunachalam S, Temme K. A rigorous and robust quantum speed-up in supervised machine learning. *Nat Phys*. 2021;17:1013–7. <https://doi.org/10.1038/s41567-021-01287-z>.
17. Cong I, Choi S, Lukin MD. Quantum convolutional neural networks. *Nat Phys*. 2019;15:1273–8. <https://doi.org/10.1038/s41567-019-0648-8>.
18. Gao J, Jiang Q, Zhou B, Chen D. Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: an overview. *Math Biosci Eng*. 2019;16:6536–61. <https://doi.org/10.3934/mbe.2019326>.
19. Shin HC, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging*. 2016;35:1285–98. <https://doi.org/10.1109/TMI.2016.2528162>.
20. Khoshman A, et al. Quantum variational autoencoder. *Quantum Sci Technol*. 2018;4:014001. <https://doi.org/10.1088/2058-9565/aada1f>.
21. Gao N, et al. High-dimensional similarity search with quantum-assisted variational autoencoder. In: Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining. 2020. p. 956–64. <https://doi.org/10.1145/3394486.3403138>.
22. Zavrak S, Iskefiyeli M. Anomaly-based intrusion detection from network flow features using variational autoencoder. *IEEE Access*. 2020;8:108346–58. <https://doi.org/10.1109/ACCESS.2020.3001350>.
23. Goodfellow I, et al. Generative adversarial networks. *Commun ACM*. 2020;63:139–44. <https://doi.org/10.1145/3422622>.
24. Niu MY, et al. Entangling quantum generative adversarial networks. *Phys Rev Lett*. 2022;128:220505. <https://doi.org/10.1103/PhysRevLett.128.220505>.
25. Li T, Zhang S, Xia J. Quantum generative adversarial network: a survey. *Comput Mater Continua*. 2020;64:401–38. <https://doi.org/10.32604/CMC.2020.010551>. Preprint.
26. Ning T, Yang Y, Du Z. Quantum kernel logistic regression based Newton method. *Phys A, Stat Mech Appl*. 2023;611:128454. <https://doi.org/10.1016/j.physa.2023.128454>.
27. Sperandei S. Understanding logistic regression analysis. *Biochem Med*. 2014;24:12–8. <https://doi.org/10.11613/BM.2014.003>.
28. Srimaneekarn N, Hayter A, Liu W, Tantipoj C. Binary response analysis using logistic regression in dentistry. *Int J Dent*. 2022;2022:5358602. <https://doi.org/10.1155/2022/5358602>.
29. Belgiu M, Drăgu L. Random forest in remote sensing: a review of applications and future directions. *ISPRS J Photogramm Remote Sens*. 2016;114:24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.
30. Zoufal C, Lucchi A, Woerner S. Quantum Generative Adversarial Networks for learning and loading random distributions. *npj Quantum Inf*. 2019;5:103. <https://doi.org/10.1038/s41534-019-0223-2>.
31. Speiser JL, Miller ME, Tooze J, Ip E. A comparison of random forest variable selection methods for classification prediction modeling. *Expert Syst Appl*. 2019;134:93–101. <https://doi.org/10.1016/j.eswa.2019.05.028>.
32. Bentéjac C, Csörgő A, Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms. *Artif Intell Rev*. 2021;54:1937–67. <https://doi.org/10.1016/j.eswa.2019.05.028>.
33. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat*. 2001;29:1189–232.
34. Lee Y, et al. Boosting quantum yields in two-dimensional semiconductors via proximal metal plates. *Nat Commun*. 2021;12:7095. <https://doi.org/10.1038/s41467-021-27418-x>.
35. Chandra MA, Bedi SS. Survey on SVM and their application in image classification. *Int J Inf Technol (Singapore)*. 2021;13:1–11. <https://doi.org/10.1007/s41870-017-0080-1>.

36. Huang W, et al. Railway dangerous goods transportation system risk identification: comparisons among SVM, PSO-SVM, GA-SVM and GS-SVM. *Appl Soft Comput.* 2021;109:107541. <https://doi.org/10.1016/j.asoc.2021.107541>.
37. Huang MW, Chen CW, Lin WC, Ke SW, Tsai CF. SVM and SVM ensembles in breast cancer prediction. *PLoS ONE.* 2017;12:e0161501. <https://doi.org/10.1371/journal.pone.0161501>.
38. Kusumoto T, Mitarai K, Fujii K, Kitagawa M, Negoro M. Experimental quantum kernel trick with nuclear spins in a solid. *npj Quantum Inf.* 2021;7:94. <https://doi.org/10.1038/s41534-021-00423-0>.
39. Schwantes CR, Pande VS. Modeling molecular kinetics with tICA and the kernel trick. *J Chem Theory Comput.* 2015;11:600–8. <https://doi.org/10.1021/ct5007357>.
40. Kutateladze V. The kernel trick for nonlinear factor modeling. *Int J Forecast.* 2022;38:165–77. <https://doi.org/10.1016/j.ijforecast.2021.05.002>.
41. Tomono T, Tsujimura K, Godo T. Quantum kernels for difficult visual discrimination. In: Proceedings - 2023 IEEE international conference on Quantum Computing and Engineering, QCE 2023. 2023. p. 262–3. <https://doi.org/10.1109/QCE57702.2023.10234>.
42. Tomono T, Natsubori S. Shipping inspection trial of quantum machine learning toward sustainable quantum factory. *PHM Soc Asia-Pac Conf.* 2023;4:3733. <https://doi.org/10.36001/phmap.2023.v4i1.3733>.
43. Tomono T, Natsubori S. Performance of quantum kernel on initial learning process. *EPJ Quantum Technol.* 2021;9:35. <https://doi.org/10.1140/epjqt/s40507-022-00157-8>.
44. Liu Y, Arunachalam S, Temme K. A rigorous and robust quantum speed-up in supervised machine learning. *Nat Phys.* 2021;17:1013–7. <https://doi.org/10.1016/j.nphys.2019.05.028>.
45. Kübler JM, Buchholz S, Schölkopf B. The Inductive Bias of Quantum Kernels. *Adv Neural Inf Process Syst.* 2021;16.
46. Senokosov A, Sedykh A, Saginalieva A, Kyriacou B, Melnikov A. Quantum machine learning for image classification. *Mach Learn: Sci Technol.* 2024;5:015040. <https://doi.org/10.1088/2632-2153/ad2aef>.
47. Aswiga RV, Sridevi S, Indira B. Leveraging Quantum Kernel Support Vector Machine for breast cancer diagnosis from Digital Breast Tomosynthesis images. *Quantum Mach Intell.* 2024;6:40. <https://doi.org/10.1007/s42484-024-00170-3>.
48. Srikumar M, Hill CD, Hollenberg LCL. A kernel-based quantum random forest for improved classification. *Quantum Mach Intell.* 2024;6:10. <https://doi.org/10.1007/s42484-023-00131-2>.
49. Pérez-Salinas A, Cervera-Lierta A, Gil-Fuster E, Latorre JL. Data re-uploading for a universal quantum classifier. *Quantum.* 2020;4:226. <https://doi.org/10.22331/q-2020-02-06-226>.
50. Gan BY, Leykam D, Angelakis DG. Fock state-enhanced expressivity of quantum machine learning models. *EPJ Quantum Technol.* 2022;9:16. <https://doi.org/10.1140/epjqt/s40507-022-00135-0>.
51. Cerezo M, Sone A, Volkoff T, Cincio L, Coles PJ. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nat Commun.* 2021;12:1791. <https://doi.org/10.1038/s41467-021-21728-w>.
52. Wang S, et al. Noise-induced barren plateaus in variational quantum algorithms. *Nat Commun.* 2021;12:6961. <https://doi.org/10.1038/s41467-021-27045-6>.
53. Sharma K, Cerezo M, Cincio L, Coles PJ. Trainability of dissipative perceptron-based quantum neural networks. *Phys Rev Lett.* 2022;128:180505. <https://doi.org/10.1103/PhysRevLett.128.180505>.
54. Ortiz Marrero C, Kieferová M, Wiebe N. Entanglement-induced barren plateaus. *PRX Quantum.* 2021;2:040316. <https://doi.org/10.1103/PRXQuantum.2.040316>.
55. Holmes Z, Sharma K, Cerezo M, Coles PJ. Connecting ansatz expressibility to gradient magnitudes and barren plateaus. *PRX Quantum.* 2022;3:010313. <https://doi.org/10.1103/PRXQuantum.3.010313>.
56. McClean JR, Boixo S, Smelyanskiy VN, Babbush R, Neven H. Barren plateaus in quantum neural network training landscapes. *Nat Commun.* 2018;9:4812. <https://doi.org/10.1038/s41467-018-07090-4>.
57. Amin J, et al. Pest localization using YOLOv5 and classification based on quantum convolutional network. *Agriculture (Switzerland).* 2023;13:662. <https://doi.org/10.3390/agriculture13030662>.
58. Aasim M, Katici R, Acar AŞ, Ali SA. A comparative and practical approach using quantum machine learning (QML) and support vector classifier (SVC) for Light emitting diodes mediated in vitro micropropagation of black mulberry (*Morus nigra* L.). *Ind Crop Prod.* 2024;213:118397. <https://doi.org/10.1016/j.indcrop.2024.118397>.
59. Choi E, et al. Quantum machine learning for additive manufacturing process monitoring. *Manuf Lett.* 2024;41:1415–22. <https://doi.org/10.1016/j.mfglet.2024.09.168>.
60. Tscharke K, Issel S, Debus P. Semisupervised anomaly detection using support vector regression with quantum kernel. In: Proceedings - 2023 IEEE international conference on Quantum Computing and Engineering, QCE 2023. 2023. vol. 1. p. 611–20. <https://doi.org/10.1109/QCE57702.2023.00075>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.