



# Satellite Image Representations for Quantum Classifiers

Johann Maximilian Zollner<sup>1</sup> · Paul Walther<sup>1</sup> · Martin Werner<sup>1</sup>

Received: 29 September 2023 / Accepted: 30 January 2024  
© The Author(s) 2024

## Abstract

Existing quantum hardware is limited in the number of bits and length of the series of operations. Nevertheless, by shifting parts of the computation on classical hardware, hybrid quantum-classical systems utilize quantum hardware for scaled-down machine learning approaches, which is quantum machine learning. Due to the theoretically possible computational speed-up of quantum computers compared to classical computers and the increasing volume and speed of data generated in earth observation, attempts are now being made to use quantum computers for satellite image processing. However, satellite imagery is too large and high dimensional, and transformations that reduce the dimensionality are necessary to fit the classical data in the limited input domain of quantum circuits. This paper presents and compares several dimensionality reduction techniques as part of hybrid quantum-classical systems to represent satellite images with up to  $256 \times 256 \times 3$  values with only 16 values. We evaluate the representations of two benchmark datasets with supervised classification by four different quantum circuit architectures. We demonstrate the potential use of quantum machine learning for satellite image classification and give a comprehensive overview of the impact of various satellite image representations on the performance of quantum classifiers. It shows that autoencoder models are best suited to create small-scale representations, outperforming commonly used methods such as principle component analysis.

**Keywords** Quantum Computing · Hybrid System · Dimensionality Reduction · Supervised Classification

## 1 Introduction

The volume and velocity of globally generated data increase continuously, particularly in the field of earth observation, where the commercial space industry, smaller and cheaper satellites, and advances in hardware facilitate the constant stream of data. Now, researchers are trying to exploit new hardware platforms to cope with the increasing processing complexity. Similar to exploiting the power of parallel processing of graphical processing units for deep neural networks, the goal is to enable quantum computers for ma-

chine learning tasks, which is quantum machine learning (QML).

In contrast to classical computers, where bits can be in one of the two binary states 0 and 1, quantum computers exploit the so-called superposition of a quantum bit (qubit). A single qubit in superposition is in the quantum states  $|0\rangle$  and  $|1\rangle$  simultaneously, and together with the quantum mechanical concept of entanglement, this makes it possible to follow different paths of computation at the same time [20]. A series of operations, also called gates, on qubits connected by wires, which allow the states to travel between operations, is known as a quantum circuit. Quantum circuits are generally the most similar quantum computational models concerning classical computers. Implementing machine learning algorithms on quantum hardware by parameterizing gates in quantum circuits is possible, but adapting classical concepts is difficult [2]. However, quantum computing theoretically enables a so-called quantum advantage, which describes a measurable computational speed-up in comparison to classical algorithms [1]. If a practical quantum advantage can be realized, QML may be a new way to process the increasing amount of generated data in a reasonable time.

---

✉ Johann Maximilian Zollner  
maximilian.zollner@tum.de

Paul Walther  
paul.walther@tum.de

Martin Werner  
martin.werner@tum.de

<sup>1</sup> Department of Aerospace and Geodesy, Technical University of Munich, Lise-Meitner-Straße 9, 85521 Ottobrunn, Germany

It must be clarified that large-scale fault-tolerant quantum computers are still out of reach, and due to noise in quantum hardware, only short sequences of operations and a limited amount of qubits can be realized. Real-world data and particular imagery are large and multidimensional, and they must be transformed and reduced to fit in the limited input domain of parameterized quantum circuits. However, performing this preprocessing in addition to the actual computation on the quantum hardware is impossible due to the previously mentioned limitations. A workaround is so-called hybrid quantum-classical systems, where subroutines for pre- and postprocessing run on classical hardware to reduce the requirements on quantum hardware [2]. Already today, hybrid quantum-classical systems have used existing quantum hardware for scaled-down machine learning problems, and several successful attempts at classifying computer vision datasets, like handwritten digits, with hybrid systems were presented in the literature [2, 6, 12]. Due to the computational speed up that quantum computing may offer [14] and recent advantages in the development of quantum hardware in capacity and error correction [13], the new research field of QML is growing vastly.

Nevertheless, even with increasing available qubits and enhanced error correction, the input domain should be as small as possible without impairing the classification performance to save resources and counter noise. Furthermore, dimensionality reduction is essential to create meaningful features and supports classification [11]. Also, it reduces the computational effort, counters the curse of dimensionality, and overfitting on the training data [3]. However, recent publications where hybrid-quantum classical systems are exploited for satellite image classification only cover a single dimensionality reduction technique without comparison and do not further investigate the influence of the data representation on the performance of the quantum classifier [8, 15, 21].

This article studies the influence of various small-scale satellite image representations in QML. We propose a general framework for hybrid quantum-classical systems and evaluate and discuss various compositions of nine dimensionality reduction techniques and four quantum classifiers with supervised classification tasks from two datasets. Besides comparing the performance of the quantum classifiers with certain preliminary dimensionality reduction of the input data, we conduct the experiments with an equally limited classical machine learning approach, which serves as a baseline. This article builds on and extends a previous publication [22]. With two additional circuit architectures and a total of 770 models, we validate the indications from [22] and give a comprehensive overview of the performance of quantum classifiers with various small-scale satellite image representations. Furthermore, we discuss both the findings and their implications, focusing on the impact of

the chosen dimensionality reduction technique on the performance of the quantum classifier. In general, we demonstrate the potential of hybrid quantum-classical systems for the classification of satellite imagery and show that autoencoder models are particularly well suited to creating small-scale representations of the input data.

## 2 Background

This article considers several dimensionality reduction techniques to create small-scale image representations, which we outline in the following. Additionally, we present fundamental concepts and notation to understand the implemented quantum classifiers.

Dimensionality reduction describes a reduction of data in  $N$ -dimensional space to a  $K$ -dimensional latent space where  $K \ll N$  while preserving information. First, downscaling (DS) with the local mean for a given array with the integer factors of the image axis is a straightforward approach. Next, we consider the well-known linear dimensionality reduction methods principal component analysis (PCA) and the closely related factor analysis (FA). PCA projects the input on a lower dimensional space by singular value decomposition to maximize the downsampled data's variance. Similarly to PCA, FA describes the data by explaining covariances and variances and representing them as a set of factors. While FA is closely related to PCA, particularly probabilistic PCA, it differs in the underlying model, such that it presumes the conditional distribution of the input  $x$  given the latent variable  $K$  to have a diagonal rather than an isotropic covariance [3].

Further, unsupervised learning methods, known as autoencoder models, are considered. An encoder  $f$  and a symmetric decoder  $g$  are trained together by minimizing the difference between the input  $x$  and the output  $y = g(f(x))$  to learn a low-dimensional latent representation  $\tilde{x}$  of size  $N$ , such that  $f(x) = \tilde{x}$ . Depending on the model's weights and bias, the reconstruction error is the objective function and is minimized during training by backpropagation. While autoencoder models typically start the training process with random initial weights, pre-trained restricted Boltzmann machines (RBMs) can provide weights close to a good solution in advance [11]. RBMs are two-layered neural networks that learn a probability distribution to reconstruct the input. They connect binary pixels in a visible layer containing the input data, with feature detectors in a hidden layer. Those symmetric connections are weighted, corresponding to the strength of the connection between the visible variable and the hidden variable. So, in contrast to autoencoder models, RBMs use the same weights for the encoder and decoder layers.

Now, we introduce parameterized quantum circuits (PQCs), which realize the quantum classifiers. A quantum circuit has a number of qubits and a series of operations which act on the quantum states. Such operations are called gates and perform rotations of qubits around the Bloch sphere. Now, unitary operators  $U$  may depend on a set of parameters  $\theta$ , which can be adapted for some tasks. From a set  $\{U_i(\theta_i)\}$ ,  $J$  unitaries build a parameterized circuit

$$\widehat{U}_\theta = U_J(\theta_J)U_{J-1}(\theta_{J-1})\dots U_1(\theta_1) \quad (1)$$

which depends on  $\theta_j = \theta_J, \theta_{J-1}, \dots, \theta_2, \theta_1$  [7]. When a unitary operator  $\widehat{U}_\theta$  acts on the initial quantum state  $|\psi\rangle$ , it produces a new state  $|\psi_t(\theta)\rangle = \widehat{U}(\{U_i(\theta_i)\})|\psi_{t-1}\rangle$  where the value of an observable quantity can be measured [2]. The superposition collapses when a circuit is measured and an expectation value is obtained. A measurable observable describes a Hermitian operator  $M$ , which is one of the Pauli matrices  $\sigma_x, \sigma_y, \sigma_z$ . The measurement of a PQC is denoted by

$$\langle M \rangle_\theta = \langle \psi | \widehat{U}_\theta^\dagger M \widehat{U}_\theta | \psi \rangle \quad (2)$$

where  $\dagger$  denotes the complex conjugate and  $\langle M \rangle_\theta$  is the measured expectation value [16]. Note that the measurement is a single value in the case of simulated quantum circuits.

Now, let  $\widehat{U}_{x,\theta} = \widehat{U}_\theta \widehat{U}_x$  be a PQC that consists of an encoder circuit  $\widehat{U}_x$  and a variational circuit  $\widehat{U}_\theta$ . The encoder circuit  $\widehat{U}_x$ , also called the state preparation circuit, is parameterized by the classical input data  $x$  and encodes it in quantum state. Quantum encoding is necessary to process classical data on a quantum computer and is a unique challenge for QML. Different algorithms can require different encoding methods, and various methods exist to accomplish quantum embedding [16]. In this paper, angle encoding and basis encoding, which are efficient in terms of the depth of the encoding circuit, are utilized and will be discussed in Sect. 3. The second part of the PQC, the variational circuit  $\widehat{U}_\theta$ , may have an arbitrary architecture adapted to some task. The unitary gates  $U_i(\theta_i)$  parameterized by  $\theta_i$  act on the input data in quantum state until the qubit is measured to obtain  $\langle M \rangle_{x,\theta}$  depending on the input data  $x$  and the parameter set  $\theta$ . A circuit's parameters  $\theta$  can be optimized for a specific task by minimizing an error function equivalent to classical machine learning models.

Now, given inputs  $x \in \mathcal{X}$  with labels  $y \in \mathcal{Y}$ , the definition of a machine learning model  $f_\theta(x) = y$ , and the definition for the measurement (Eq. 2), a PQC as a machine learning model is defined by

$$f_\theta(x) = \langle \psi(x, \theta) | M | \psi(x, \theta) \rangle \quad (3)$$

where the output of the model is the measurement of the circuit [2, 16]. Here,  $|\psi(x, \theta)\rangle$  defines the state prepared by  $\widehat{U}_{x,\theta}|0\rangle$ . The measurement  $\langle M \rangle_{x,\theta}$  can then be interpreted as the predicted label  $\widehat{y}$  for the input  $x$  and is further used to compute a loss function  $L_\theta$  depending on the parameters  $\theta$ . The parameters are updated through stochastic gradient descent in the form of

$$\theta - \eta \nabla L_\theta(\langle M \rangle_{x,\theta}, y) = \theta^* \quad (4)$$

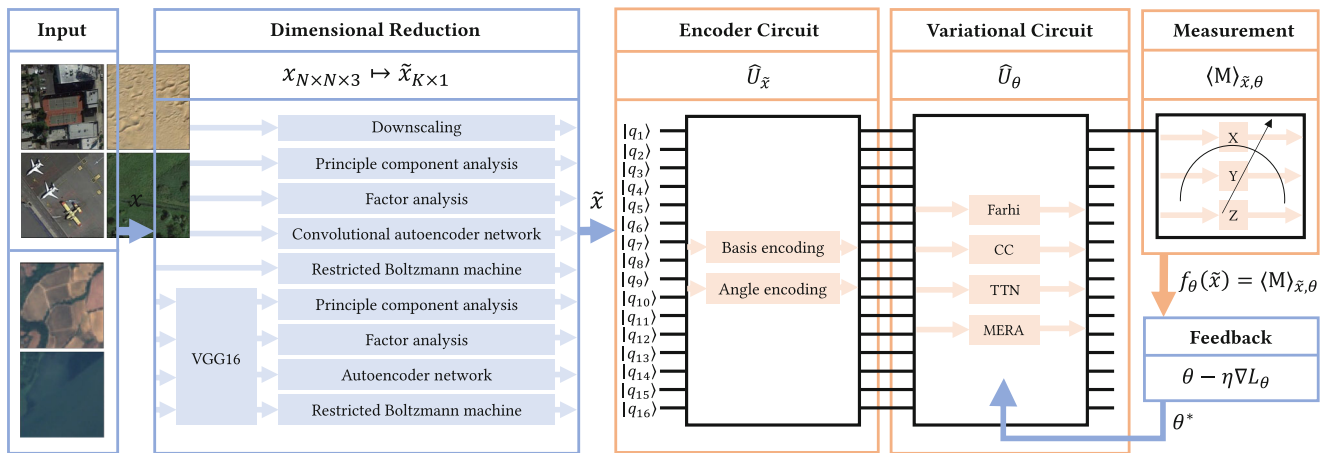
where  $\eta$  is the learning rate.

### 3 Method

This section will outline the general framework for the hybrid quantum-classical systems and the methodology for the experiments. We propose hybrid systems consisting of classical dimensionality reduction, a PQC for classification, and classical post-processing, which is the free parameters  $\theta$  update by stochastic gradient descent (Fig. 1).

The classical pre-processing in this work can be generally described as a data transformation  $x \mapsto \widetilde{x}$  which maps an input image  $x$  with  $N$  elements to a vector  $\widetilde{x}$  with a latent dimension  $K$  such that  $K \ll N$ . Since the computational cost of simulating quantum systems on classical hardware grows exponentially in the number of qubits, and, for example, TTN and MERA architectures comprise a power of two number of qubits,  $K = 16$  was chosen regarding the largest implementable qubit register size. Besides downscaling and linear dimensionality reduction with PCA and FA, we conduct non-linear dimensionality reduction with a convolutional autoencoder and an autoencoder created from RBMs. In addition to the previously mentioned dimensionality reduction techniques, feature extraction with a very deep convolutional network with 16 layers (VGG16) is considered [18]. It comprises 13 convolutional layers with a kernel size of  $3 \times 3$  and five maximum pooling layers. A maximum pooling layer follows each stack of convolutions to half the input size. The VGG16 is pre-trained and implemented without the fully connected top layers, which are usually for classification, to perform convolution and thus reduce the input images to fewer values while creating new features. PCA, FA, an autoencoder with dense layers, and an autoencoder created from RBMs are combined with prior feature extraction by the pre-trained VGG16. Consequently, we evaluate a total of nine methods to obtain small-scale data representations.

We consider two quantum encoding methods that are efficient regarding the depth of the encoding circuit: basis and angle encoding. For basis encoding, we first binarize



**Fig. 1** Hybrid quantum-classical system scheme. The transformation  $x \mapsto \tilde{x}$  maps the input images  $x$  to a vector representation  $\tilde{x}$  such that it fits in the input domain of the quantum circuit. The quantum part consists of the encoder circuit  $\hat{U}_{\tilde{x}}$  and the classification circuit  $\hat{U}_{\theta}$ . The output is the measured expectation value  $\langle M \rangle_{\tilde{x}, \theta}$  which gets mapped to the prediction  $\hat{y}$  to compute a loss function and update the parameter set  $\theta$  by stochastic gradient descent

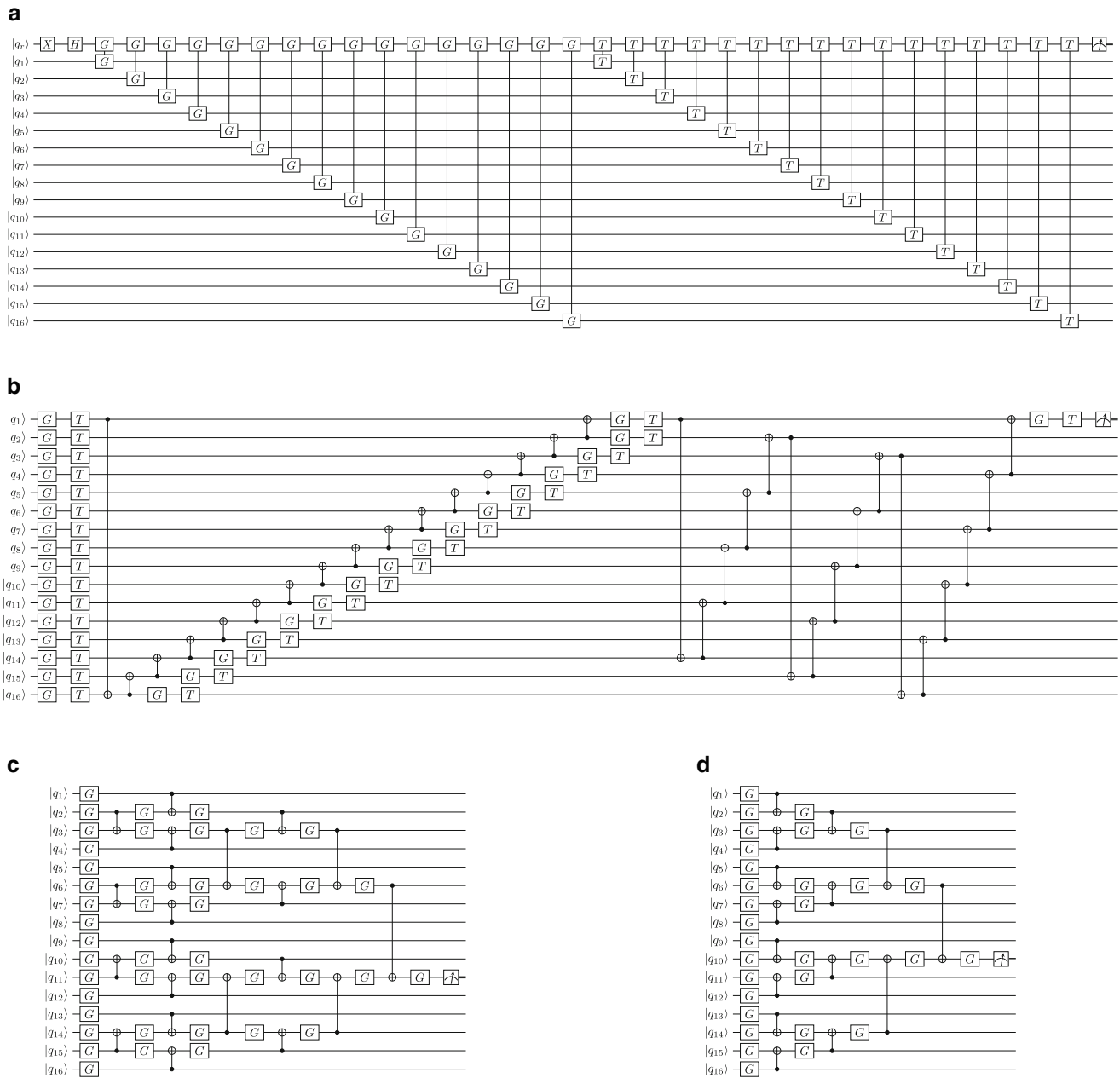
the transformed input data with its median. Then,  $\hat{U}_{\tilde{x}}$  is built from one  $X$  gate for every qubit  $q_k$  with  $k = 1, \dots, K$  where  $\tilde{x}_k = 1$  such that the initial quantum state is changed in the form of  $X|0\rangle = |1\rangle$ . Thus, every value  $x_k$  of the binary input vector is directly mapped to the quantum state with a single operation  $x_k \mapsto |x_k\rangle$ . Basis encoding needs one qubit for each value of the input vector but results in a shallow encoding circuit with only a single  $X$  gate for each qubit. In the case of angle encoding, all transformed inputs  $\tilde{x}$  are represented rescaled to  $0 \leq x \leq 2\pi$ . Then, each input value  $\tilde{x}_k$  is encoded by applying a single-qubit Pauli rotation from the set  $\{R_x, R_y, R_z\}$  around an axis such that the angle of the rotation depends on the input data value. Thus, one value of the input vector needs one qubit to be encoded but has a shallow architecture with one rotation gate. In general, after applying  $\hat{U}_{\tilde{x}}$ , each qubit  $q_k$  stores one value of the quantum encoded image vector  $|\psi_k\rangle$ .

We consider four quantum circuit architectures for classification, which were previously presented in the literature (Fig. 2) [7, 9, 17, 19], and are described in the following paragraph. Recall that all quantum classifiers have an input space of  $K = 16$  data qubits  $q_k$ , which store the transformed and quantum encoded input data  $|\psi_k\rangle$ . Further, the PQCs as machine learning models (Eq. 3) are shortly denoted as  $f_{\theta}(\tilde{x}) = \langle M \rangle_{\tilde{x}, \theta}$  where  $\langle M \rangle_{\tilde{x}, \theta}$  is the measurement, which is the prediction for the label of the transformed input image  $\tilde{x}$ . First, the Farhi circuit architecture (Fig. 2a) was inspired by [7] and already showed promising results in previous work [15]. The circuit has an additional 17th qubit, which is the readout qubit  $|q_r\rangle$ . The readout qubit is prepared by a NOT gate  $X$  and a Hadamard gate  $H$  followed by two-qubit parameterized gates, which are 16  $XX(\theta) = e^{-i\frac{\pi}{2}\theta X \otimes X}$  Ising gates and

16  $ZZ(\theta) = e^{-i\frac{\pi}{2}\theta Z \otimes Z}$  Ising gates and act on the readout qubit  $|q_r\rangle$ . The circuit has a depth of 35, assuming it uses a simple Pauli-Z measurement and 32 free parameters. Finally, the readout qubit  $|q_r\rangle$  is measured to obtain an expectation value. Second, a tree tensor architecture denoted by TTN circuit (Fig. 2c), which was already applied on a standard computer vision dataset and resulted in high accuracy [9]. The circuit has 31  $R_y(\theta_i)$  gates, which are single-qubit rotations around the  $Y$ -axis through an angle parameterized by  $\theta$ , and consequently 31 free parameters. Furthermore, the circuit consists of 8 controlled-NOT (CNOT) and 7 shifted CNOT gates. Since binary trees inspire the architecture, half of the qubits are discarded after each set of rotational and CNOT gates. The circuit has a depth of 9, assuming it uses a simple Pauli-Z measurement. Third, a circuit-centric quantum classifier, proposed by [17] and denoted by CC circuit (Fig. 2b). It consists of two blocks with 16 parameterized rotations  $R_z(\theta_i)$ , 16  $R_x(\theta_i)$ , and 16 CNOT gates each. The qubit  $|q_0\rangle$  is then acted on by two additional rotation gates  $R_z(\theta_i)$  and  $R_x(\theta_i)$  and finally measured to obtain an expectation value. The circuit has a depth of  $\geq 39$  with  $I = 66$  free parameters. Last, a multiscale entanglement renormalization ansatz [19] for classification tasks (Fig. 2d). It has 43 parameterized rotation gates  $R_y(\theta_i)$ , 17 CNOT gates, and a depth of  $\geq 14$ .

The classical postprocessing consists of the mapping of the measured expectation value  $\langle M \rangle_{\tilde{x}, \theta}$  to the predicted class label  $\hat{y}$  for the input  $x$ , which is the prediction of the classifier. Further, the computation of a loss function  $L_{\theta}$  depending on the parameters  $\theta$ , which is given in case of square hinge loss by

$$L_{\theta}(\langle M \rangle_{\tilde{x}, \theta}, y) = \max(0, 1 - y * \langle M \rangle_{\tilde{x}, \theta})^2 \quad (5)$$



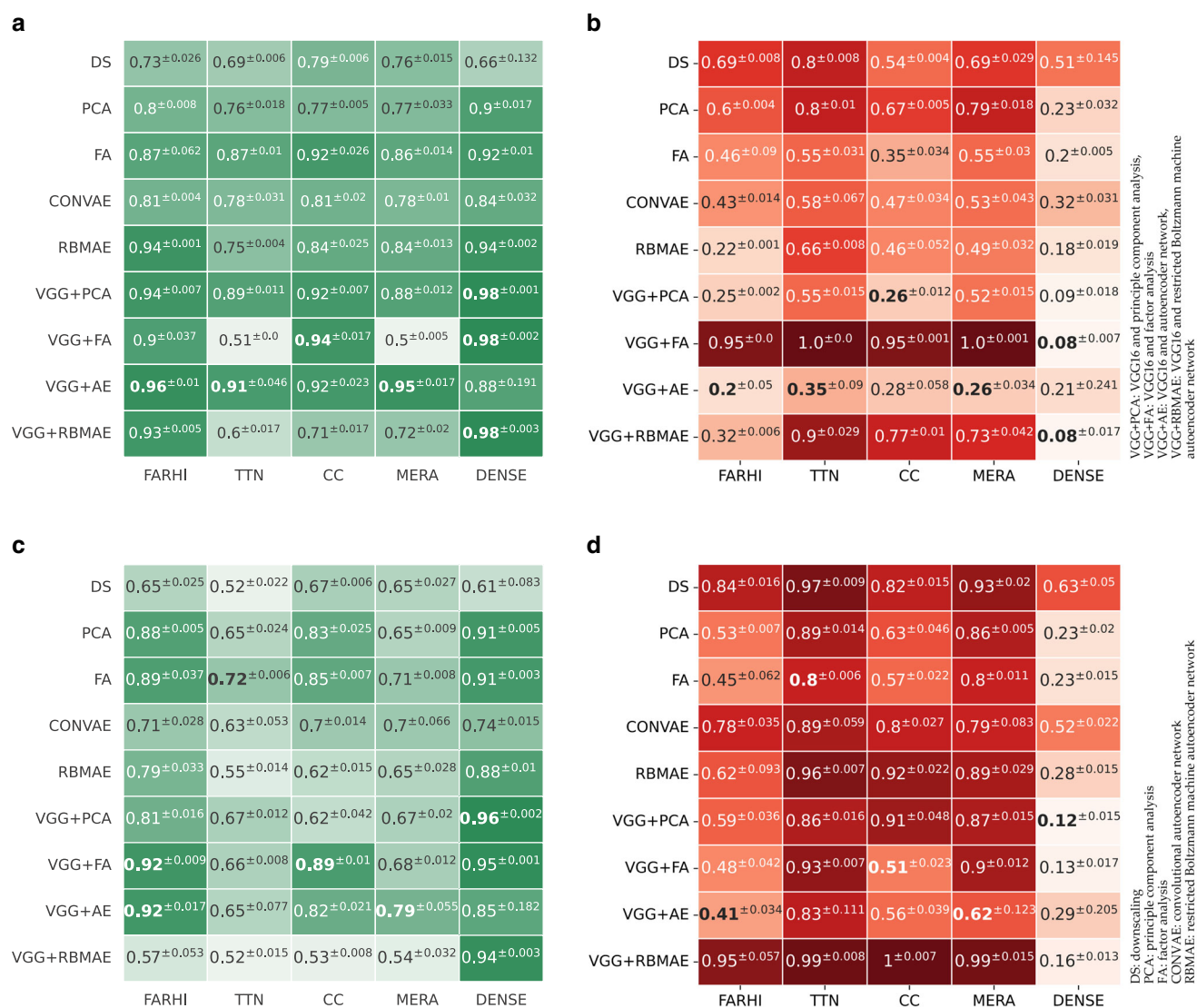
**Fig. 2** Parameterized quantum circuits for classification. The figures are made with (q|pic) (<https://github.com/qpqc/qpqc>). **a** Farhi quantum classifier [7] with a depth  $\geq 34$  and  $I = 32$  free parameters  $\theta_i$ .  $G$  and  $T$  are parameterized two-qubit Ising gates. Here,  $G$  denotes  $XX(\theta_i)$  and  $T$  denotes parametric  $ZZ(\theta_i)$  which act on  $|q_r\rangle$ , **b** Circuit centric (CC) quantum classifier [17] with a depth  $\geq 38$  and  $I = 66$  free parameters  $\theta_i$ .  $G$  and  $T$  are parameterized one-qubit quantum gates.  $G$  denotes  $R_z(\theta_i)$  and  $T$  denotes  $R_x(\theta_i)$ . For the CNOT gates, the control qubit is drawn in black, **c** Tree tensor network (TTN) quantum classifier [9] with a depth  $\geq 9$  and  $I = 31$  free parameters  $\theta_i$ .  $G$  is a parameterized single-qubit quantum gate and denotes  $R_y(\theta_i)$ . For the CNOT gates, the control qubit is drawn in black, **d** Multiscale entanglement renormalization ansatz (MERA) quantum classifier [19] with a depth  $\geq 14$  and  $I = 43$  free parameters  $\theta_i$ .  $G$  is a parameterized single-qubit quantum gate and denotes  $R_y(\theta_i)$ . For the CNOT gates, the control qubit is drawn in black

with  $y \in [-1,1]$ . The parameter set is then updated via classical stochastic gradient descent.

We implement the experiments with the Tensorflow-quantum [4] framework in combination with `cirq`<sup>1</sup> and

<sup>1</sup> <https://github.com/quantumlib/Cirq>.

simulate the quantum systems on classical hardware. All experiments with hybrid systems include two publicly available real-world image datasets: EuroSAT [10] and NWPU-RESISC45 [5]. EuroSAT consists of 27000 Sentinel-2 satellite images with 10 land use and land cover classes. A single image has a size of  $64 \times 64$  pixels with a ten-

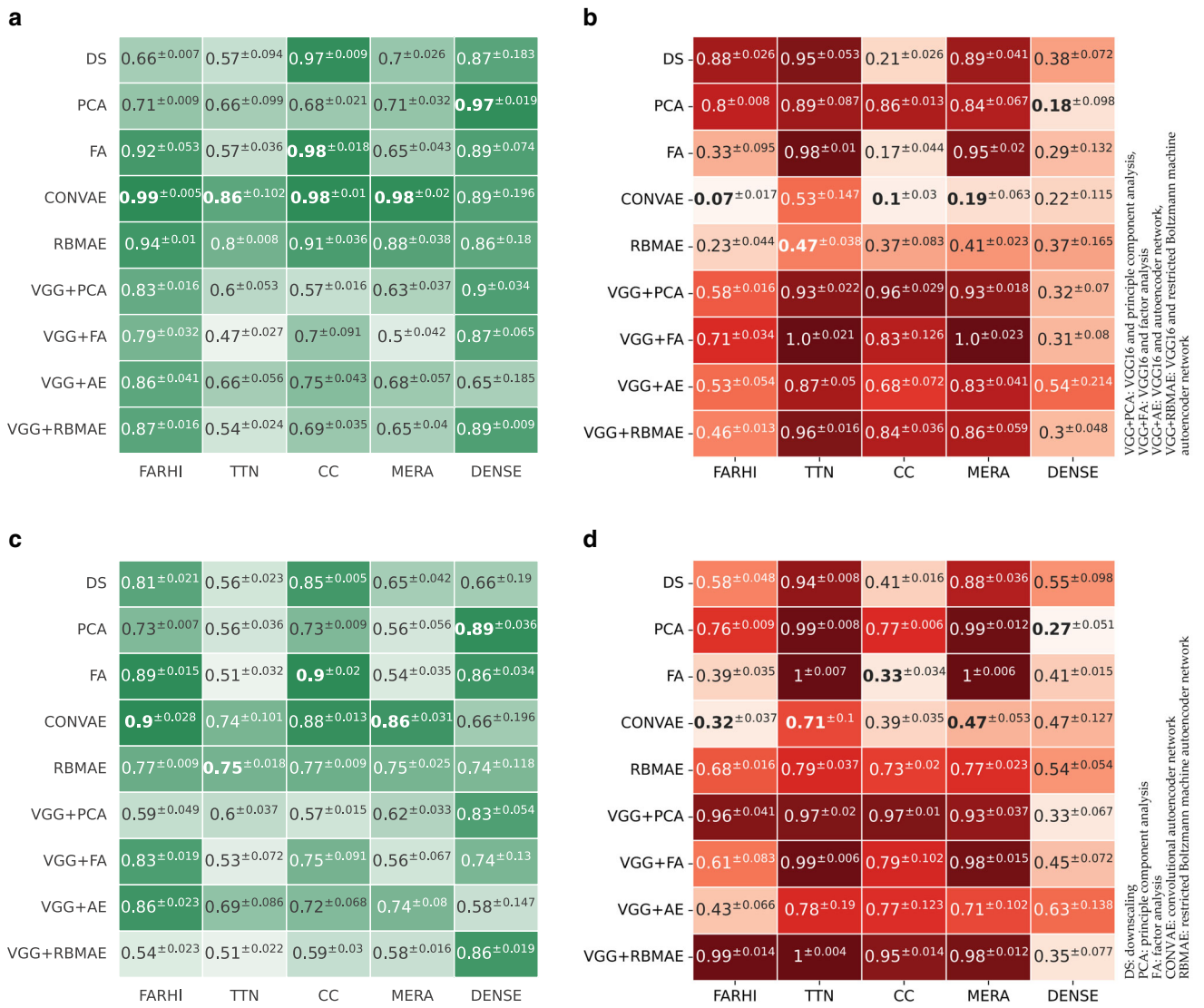


**Fig. 3** Example results for binary classification of EuroSAT [10]. Average accuracy (**a,c**) and average loss (**b,d**) and their standard deviation over five runs are displayed. Bold values indicate the best performance. **a** Accuracy for AnnualCrop and SeaLake, **b** Loss for AnnualCrop and SeaLake, **c** Accuracy for PermanentCrop and Industrial, **d** Loss for PermanentCrop and Industrial

meter resolution. The Northwestern Polytechnical University created the second dataset for remote sensing image scene classification. It has 45 balanced classes and is titled NWPU-RESISC45. A single image has a size of  $256 \times 256$  pixels with a varying resolution from 0.2 meters to 30. In contrast to the EuroSAT dataset, RESISC45 has high diversity within one class in terms of, for example, translation, viewpoint, and background. We train all models for 50 epochs besides the VGG16, which uses weights obtained from pre-training on the ImageNet dataset. To minimize the loss function, the optimizer of choice is Adaptive Moment Estimation with a learning rate of 0.001. Besides the hybrid systems, we present a simple classical machine learning model with two fully connected layers and 37 free parameters. While this is, in the first place, not to

compare quantum to classical approaches, it shall validate the usability of the lower dimensional features and serve as a baseline to evaluate the dimensionality reduction methods. However, since the quantum classifiers are limited in the number of gates, we equally limited the classical model to get a fitting baseline.

Finally, with nine dimensionality reduction techniques and four circuit architectures, we get 36 hybrid systems. We choose two binary classification tasks for each dataset and train every model five times to ensure the experiments' reproducibility, resulting in 720 models. Additionally, five one-versus-rest models are trained with the EuroSAT data, which makes an additional 50 models.



**Fig. 4** Example results for binary classification of NWPU-RESISC45 [5]. Average accuracy (a,c) and average loss (b,d) and their standard deviation over five runs are displayed. Bold values indicate the best performance. **a** Accuracy for meadow and desert, **b** Loss for meadow and desert, **c** Accuracy for tennis court and airplane, **d** Loss for tennis court and airplane

The code to reproduce the experiments is published in an open code repository<sup>2</sup>.

## 4 Results and Discussion

We perform three sets of experiments to evaluate the data representations for classification with hybrid quantum-classical systems. First, we perform a grid search to find a suitable configuration for the classifiers' loss function, quantum encoding, and quantum observable. Second, we train the classifiers with all dimensionality reduction methods to compare their performance dependent on the different data

representations. Third, we conduct a multiclass approach with the most promising hybrid system.

To begin with, we conduct the grid search with imagery transformed by a simple autoencoder built from two dense layers. The grid search shows that, for basis encoding, due to the binarization of the transformed input  $\tilde{x}$  and the additional loss of information, no suitable data representations can be obtained when mapping to only 16 values. Furthermore, note that the impact of the chosen loss function on the classification performance was low. Although the differences between the suitable configurations regarding the rotation for angle encoding and the quantum observable were small overall, following the grid search results, all hybrid systems are trained with a combination of X-angle embed-

<sup>2</sup> <https://github.com/tum-bgd/sirfqc>.

ding, Pauli-X observable, and square hinge loss function in the following experiments.

Next, we evaluate and compare the nine approaches for transforming and reducing satellite imagery. For both EuroSAT classification tasks (Fig. 3), hybrid systems, including a VGG16 for prior feature extraction and an autoencoder model to create a low dimensional code of the input data, show the best overall results for the quantum classifiers. For binary classification of the classes AnnualCrop and SeaLake, all hybrid systems, including the VGG16 and an autoencoder, resulted in a mean classification accuracy of  $> 90\%$ . On the other hand, the classical dense layer classification approach shows lower accuracy when trained with a VGG16 and an autoencoder than for other pre-processing methods.

Furthermore, it shows that convolutional autoencoder models are best suited for classification tasks from the RESISC45 dataset. (Fig. 4). While prior feature extraction with the pre-trained VGG16 enhanced the classification accuracy on the EuroSAT data, it did not improve the classification of the RESISC45 data. Similar to the results with the EuroSAT data, the classical dense layer approach shows the overall best results not with the same pre-processing as the quantum approaches but with PCA.

Overall, the Farhi classifier outperforms the other architectures. Here, it must be noted that the TTN, CC, and MERA classifiers only use simple one-qubit quantum gates, while Farhi includes Ising gates, which are tensor products of Pauli-Z and Pauli-X gates and act on two qubits. Additionally, MERA has low depth and thus may outperform deeper architectures on real quantum hardware with noise.

The hybrid systems generally reach similar accuracies but far higher loss values than the limited classical classification approach with the same pre-processing method. This is due to the magnitude of the output  $\hat{y}$  of the quantum classifier being close to the classifier's decision boundary, which, while maintaining high accuracy, also leads to high hinge loss values (Eq. 5).

Furthermore, we conduct a one-versus-rest approach on multiclass classification, where we train a binary classifier for each class in the dataset. Then, every classifier predicts every sample in the test set, and the model that outputs the highest value for  $\hat{y}$  determines the predicted class. For the EuroSAT dataset, the binary one-versus-rest classifiers generally reach accuracies of about 90% independent of the class. However, the model outputs are no probabilities, which means that the magnitudes of the outputs are not reasonably comparable. For example, a prediction for some sample  $f_A(x) \approx 0.83$  of model  $A$  does not ensure that the prediction  $f_B(x) \approx 0.21$  of model  $B$  for the same sample is less certain to be right. Nevertheless, feature extraction with a VGG16 and an autoencoder for dimensionality reduction in combination with the Farhi circuit results

in a mean overall accuracy of  $> 50\%$ . Presumably, due to larger images and number of classes, multiclass classification of the RESISC45 dataset resulted in noticeably lower scores.

As previously mentioned, all models were trained over 30 epochs. However, it showed that loss and accuracy did not significantly change after the third epoch when training one-versus-rest classifiers with the whole dataset. Note that this was not the case for a one-versus-rest approach with a classical classifier and quantum classifiers, which were trained on subsets of the dataset consisting of two classes.

## 5 Conclusion

In this article, we propose hybrid quantum-classical systems to classify satellite imagery and evaluate various small-scale representations of the input data. Due to the current limitations of quantum hardware, we perform pre- and post-processing classical and classification with PQC. To fit the imagery in the limited input domain of the quantum circuits, we transform and reduce the data with up to  $256 \times 256 \times 3$  values to 16 values. We benchmark and compare nine dimensionality reduction methods to create representations of the input data combined with four PQC architectures for classification. The experiments demonstrate that small-scale representations of satellite imagery are suitable for classification with hybrid quantum-classical systems. Further, we show how the chosen data transformation influences the classification performance. Often-used dimensionality reduction techniques like PCA perform worse than autoencoder methods for certain tasks, and prior feature extraction can, in some instances, further enhance the results. In summary, the findings imply that even small-scale quantum systems have potential use for real-world application when implemented with suitable dimensionality reduction techniques.

Since we observe that the magnitudes of the one-versus-rest classifier outputs are not reasonably comparable, a possible approach for multiclass classification is to exploit probability calibration. Furthermore, after having insights on the data representations and circuit architectures, enhancing the training procedure with new optimizers and loss functions is a possible follow-up for this paper.

In fact, due to the limitations of quantum hardware, QML is currently not ready to compete with classical machine learning. Still, research like this creates a scientific foundation for when it might become possible. Furthermore, even with enhanced noise reduction and error correction, data transformations like the proposed methods will remain relevant since they can improve feature representation, reduce the computational effort, and counter noise in quantum hardware by reducing the size of the required qubit



registers. Thus, transforming and reducing input data and extracting meaningful features are and will be fundamental for QML.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

## References

- Arute F, Arya K, Babbush R et al (2019) Quantum supremacy using a programmable superconducting processor. *Nature* 574(7779):505–510
- Benedetti M, Lloyd E, Sack S et al (2019) Parameterized quantum circuits as machine learning models. *Quantum Sci Technol* 4(4):43001
- Bishop CM, Nasrabadi NM (2006) *Pattern recognition and machine learning vol 4*. Springer
- Broughton M, Verdon G, McCourt T et al (2021) Tensorflow quantum: a software framework for quantum machine learning, p 2003.02989
- Cheng G, Han J, Lu X (2017) Remote sensing image scene classification: benchmark and state of the art. *Proc IEEE* 105(10):1865–1883. <https://doi.org/10.1109/jproc.2017.2675998>
- Choe S, Perkowski M (2022) Continuous variable quantum mnist classifiers—classical-quantum hybrid quantum neural networks. *J Quantum Inf Sci* 12(2):37–51
- Farhi E, Neven H (2018) Classification with quantum neural networks on near term processors, p 1802.06002
- Gawron P, Lewiński S (2020) Multi-spectral image classification with quantum neural network. In: IGARSS. IEEE, pp 3513–3516
- Grant E, Benedetti M, Cao S et al (2018) Hierarchical quantum classifiers. *NPJ Quantum Inf* 4(1):1–8
- Helber P, Bischke B, Dengel A et al (2018) Introducing eurosat: a novel dataset and deep learning benchmark for land use and land cover classification. In: IGARSS, pp 204–207 <https://doi.org/10.1109/IGARSS.2018.8519248>
- Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. *Science* 313(5786):504–507. <https://doi.org/10.1126/science.1127647>
- Kerenidis I, Luongo A (2020) Classification of the mnist data set with quantum slow feature analysis. *Phys Rev, A*. <https://doi.org/10.1103/PhysRevA.101.062327>
- de Leon NP, Itoh KM, Kim D et al (2021) Materials challenges and opportunities for quantum computing hardware. *Science*. <https://doi.org/10.1126/science.abb2823>
- Liu Y, Arunachalam S, Temme K (2021) A rigorous and robust quantum speed-up in supervised machine learning. *Nat Phys* 17(9):1013–1017. <https://doi.org/10.1038/s41567-021-01287-z>
- Otgonbaatar S, Datcu M (2021) Classification of remote sensing images with parameterized quantum gates. *IEEE Geosci Remote Sens Lett* 19:1–5. <https://doi.org/10.1109/LGRS.2021.3108014>
- Schuld M, Petruccione F (2021) *Machine learning with quantum computers*. Springer
- Schuld M, Bocharov A, Svore KM et al (2020) Circuit-centric quantum classifiers. *Phys Rev, A* 101(3):32308
- Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition, pp 1409–1556
- Vidal G (2008) Class of quantum many-body states that can be efficiently simulated. *Phys Rev Lett*. <https://doi.org/10.1103/physrevlett.101.110501>
- Werner M (2017) Quantum spatial computing. *SIGSPATIAL Special* 11(2):26–33
- Zaidenberg DA, Sebastianelli A, Spiller D (2021) Advantages and bottlenecks of quantum machine learning for remote sensing. In: *International Geoscience and Remote Sensing Symposium*. IEEE, pp 5680–5683
- Zollner JM (2022) Quantum classifiers for remote sensing. In: *Proceedings of the 30th International Conference on Advances in Geographic Information Systems*. ACM, SIGSPATIA, <https://doi.org/10.1145/3557915.3565537>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.