

This is an author's version.

The publisher's version can be found at <https://doi.org/10.1109/QCE57702.2023.00074>

“© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

Pooling techniques in hybrid quantum-classical convolutional neural networks

Maureen Monnet*, Hanady Gebran, Andrea Matic-Flierl*, Florian Kiwit,
Balthasar Schachtner†, Amine Bentellis*, Jeanette Miriam Lorenz*

*Fraunhofer Institute for Cognitive Systems IKS, Munich, Germany

†LMU University Hopsital, Munich, Germany

{maureen.monnet, andrea.matic-flierl, amine.bentellis, jeanette.miriam.lorenz}@iks.fraunhofer.de
balthasar.schachtner@med.lmu.de

Abstract—Quantum machine learning has received significant interest in recent years, with theoretical studies showing that quantum variants of classical machine learning algorithms can provide good generalization from small training data sizes. However, there are notably no strong theoretical insights about what makes a quantum circuit design better than another, and comparative studies between quantum equivalents have not been done for every type of classical layers or techniques crucial for classical machine learning. Particularly, the pooling layer within convolutional neural networks is a fundamental operation left to explore. Pooling mechanisms significantly improve the performance of classical machine learning algorithms by playing a key role in reducing input dimensionality and extracting clean features from the input data. In this work, an in-depth study of pooling techniques in hybrid quantum-classical convolutional neural networks (QCCNNs) for classifying 2D medical images is performed. The performance of four different quantum and hybrid pooling techniques is studied: mid-circuit measurements, ancilla qubits with controlled gates, modular quantum pooling blocks and qubit selection with classical postprocessing. We find similar or better performance in comparison to an equivalent classical model and QCCNN without pooling and conclude that it is promising to study architectural choices in QCCNNs in more depth for future applications.

Index Terms—quantum machine learning, quantum pooling layers, quantum convolutional neural networks, medical imaging

I. INTRODUCTION

The recent years have seen an increased interest in quantum machine learning (QML) [1], [2], [3]. QML is a discipline at the intersection of machine learning and quantum computing (QC), where in context of this paper we are interested in quantum-enhanced machine learning algorithms processing classical data like images. However, the practical benefits of these algorithms remain unclear. A common approach to implement QML algorithms is via variational quantum circuits (VQCs) [4], which consist of a circuit architecture with parametrized gates.

Current Noisy Intermediate-Scale quantum (NISQ) computers feature a limited number of $\mathcal{O}(100)$ qubits, with limited connectivity and limited gate fidelities. Therefore, hybrid QML algorithms able to deal with shallow quantum circuit sizes and only requiring a small number of qubits are most promising for near-term applications of QC. In these hybrid algorithms, an iterative interaction between classical computers and QC oc-

curs. As QC will likely be combined with classical computers as additional quantum processing unit (QPU), the QPU should ideally only be used for the parts of a calculation which are significantly accelerated by the use of QC or are not possible to calculate on classical computers. For example, a quantum computer may be used for the execution of variational parts of the QML algorithm, while a classical computer may calculate the parameter values of the variational part by minimizing a cost function.

Hybrid variants have been proposed for many well-established algorithms of classical machine learning, including convolutional neural networks (CNNs). CNNs are widely applied in computer vision tasks, such as image classification. They consist of convolutional layers, with which the characteristic features in the images are extracted, followed by fully connected layers used for classification. Commonly, a convolutional layer is followed by a pooling layer, where the extracted feature maps are down-scaled to a smaller data size. Thereby the feature maps become more robust to translations of the input and overfitting can be prevented. In principle, the whole CNN model can be translated to a quantum algorithm, but the size of the resulting quantum algorithm is likely to require quantum hardware beyond the presently available NISQ devices. If instead just certain parts of a CNN are moved to QPU, the resulting algorithm can already be NISQ-compatible. Particularly, mapping the convolutional or the pooling layers to a quantum variant represents a promising approach in terms of the number of available qubits in NISQ-devices, as these layers process the data sequentially in smaller slices, in contrast to the linear layers, which operate on all features simultaneously. While the usage of quantum convolutional layers was extensively studied in recent research, the effect of quantum pooling layers is yet relatively unexplored.

In this work, we test the suitability of quantum pooling operations for a medical imaging task. In particular, we analyze 2D ultrasound images of the breast to identify malign lesions. For this classification task, we investigate four different techniques for performing a quantum and hybrid quantum-classical pooling operation, namely the use of 1) mid-circuit measurements, 2) auxiliary qubits with controlled gate operations, 3) modular quantum pooling blocks, and lastly 4) qubit selection with a non-linear classical postprocessing

function. We compare their performance with the one of a quantum convolutional operation as well as a fully classical approach.

Our contributions are thus threefolds:

- We design and compare 4 different quantum pooling architectures for hybrid quantum-classical convolutional neural networks with respect to the training and validation performance on ultrasound images to identify malign lesions of the breast.
- We further contrast these architectures with both a classical convolutional neural network and a quantum-classical convolutional neural network without pooling containing a similar number of trainable parameters.
- We seek to provide an explanation for the performance of our developed models using a quantum relevant metric.

The paper is structured as follows. We present the related work in section II. We introduce the chosen dataset as well as the background information relevant to all designed hybrid quantum-classical convolutional neural networks and to the classical convolutional neural network in section III. We define all tested quantum pooling architectures in section IV and present and interpret their performance. We introduce a quantum metric in section V, the effective dimension, and investigate its correlation with training performance, before coming to conclusion in section VI.

II. RELATED WORK

The works in [5], [6], [7], [8], [9], [10] studied the possibility of moving all components of a CNN on a QC. Since this is generally not feasible with current NISQ-devices, hybrid quantum-classical CNNs (QCCNNs) have been introduced, in which only parts of a CNN are executed on a quantum computer. In [11], [12], [13], the first classical convolutional layer is replaced by either an untrainable or trainable quantum convolutional layer. Using this architecture, [11] and [12] achieved a good performance on the MNIST dataset [14]. We have extended that work in [15], where we have proposed different architectures, also using different encoding schemes and variants of VQCs, on multiple medical datasets. In particular, we also compared their performance to fully classical CNNs and thereby demonstrated a good performance of hybrid QCCNNs.

QCCNN architectures may profit from the inclusion of pooling layers, since those lead to a dimension reduction of the feature maps, and consequently reduce the number of parameters to learn, potentially preventing overfitting. Different proposals have been made in this direction. In [5], [7], [16], the quantum pooling operation is achieved through mid-circuit measurements, which means that measurements are performed for only a subset of the qubits in the quantum circuit. Depending on their results, different unitary operations are applied on the adjacent qubits. In general, mid-circuit measurements represent a promising method for quantum applications, where a reduction of the data is necessary. For example, the work in [17] explores their usage for error mitigation. In [8] quantum pooling layers are realized by abandoning a subset of qubits in

the quantum circuit, i.e. their quantum states are not measured and further processed in the algorithm. Similarly, in [18] quantum pooling is achieved through parametrized two-qubit operations, after which the state of one of the two qubits is ignored.

The aforementioned work on pooling layers were all proposed for full quantum variants of CNNs. Inspired by some of these ideas, we investigate further techniques to realize and compare quantum poolings in hybrid QCCNNs. For this, we build upon our preceding studies in [15], where we studied the effect of replacing a classical convolutional layer by a quantum variant in medical image classification tasks.

III. BACKGROUND

In this section, we will explain the dataset used in our studies and the necessary background information related to our experiments. This includes the design of both the classical CNN and the hybrid variants. For the latter, we focus here only on the quantum convolutional layer as well as on aspects common to all tested QCCNNs in this work. The pooling architectures are discussed in detail in section IV.

A. Dataset

Work by [19] has shown that certain QML algorithms like the quantum-variant of CNNs are capable to generalize well even when little training data is available. In medical imaging, the acquisition of large high-quality datasets is often expensive and difficult because it requires careful expert annotation. Additional challenges come from high imaging costs and privacy concerns. Therefore, medical datasets are usually relatively small [20], common dataset sizes are of the order of 100 to 1000 images. To ensure an accurate and reliable performance of artificial intelligence (AI) algorithms in medical imaging, one thus needs to find an efficient way to process such small and complex imaging data [21]. Therefore, medical imaging is an interesting area to be explored with QML.

In this work, the performance of all developed QCCNN models is explored on the BreastMNIST dataset from the MedMNIST collection [22], [23]. This dataset is characterized by a considerably low number of training samples as is typically the case in medical imaging tasks and consists of 546 training and 78 validation images of breast ultrasounds images, which are preprocessed as described in [23] and downsampled to a 28×28 resolution. The task is a binary classification between non-malignant - a combination of benign and normal states - and malignant lesions.

B. Classical CNN and QCCNNs in this work

As classical baseline, we use the architecture presented in our previous work in [15]. This consists of a convolutional layer with four filters of size 2×2 , directly followed by a fully connected layer.

In the QCCNNs, the convolutional layer is replaced by either a quantum convolution or by a combined quantum convolution + pooling operation to represent filters of size 2×2 moving over the image, as shown in figure 1. The fully

connected layer remains classical. Across all investigations conducted in the present study, a similar number of trainable parameters between the CNN and the QCCNNs is obtained with these architectures, making the comparison between fully classical and hybrid quantum-classical variants possible.

The quantum convolutional (+ pooling) layers consist of an encoding layer followed by a VQC and a measurement. The VQC is trainable and consists of a collection of trainable rotation gates and entangling two-qubit operations. For all QCCNNs designed in this work, the same encoding and measurement technique is employed:

- **Encoding:** We use the so-called higher order encoding [24] first introduced in [25], as it is expected to work best thanks to its entangling capabilities. The qubits initially in state 0 are transformed using a Hadamard gate and a Z -axis rotation $R_Z(\pi * x_n)$, where x_n denotes the n -th input. An entangling operation $R_{ZZ}(\phi_{ij})$ is then applied to every possible qubit pair i and j . This operation consists of a CNOT gate, a rotation $R_Z(\phi_{ij})$, and another CNOT gate. The rotation $R_Z(\phi_{ij})$ is applied on the j -th qubit and we use $\phi_{ij} = \pi * x_i * x_j$. The input data being normalized between -1 and 1 , the rotation gates in the encoding are operating between $-\pi$ and π .
- **Measurement:** The expectation value of the Pauli-Z operator is computed for the desired qubits and the values are stored in individual feature maps, following the strategy of [12]. With a convolutional filter size of 2×2 and the encoding mapping one input value to one qubit, we therefore use 4 qubits and hence produce 4 feature maps in this architecture.

For the VQC in the quantum convolutional layer, a basic entangling layer is chosen as baseline for comparison with the QCCNNs with quantum pooling, as it is the best performing variant for this use case in our previous work [15]. In this instance, each qubit i is rotated by a trainable angle θ_i around the X -axis before applying a sequence of CNOT entangling gates on adjacent qubit pairs.

Experiments are performed using PyTorch [26] and the quantum circuits are simulated with PennyLane [27] while neglecting any noise effects. All networks, both classical and hybrid, are trained for 20 epochs with the Adam optimizer, a learning rate of 0.001 and a batch size of 8. A total of three runs, differing between each other in the initial parameters in the network, are executed per architecture to produce variance bands. No extensive hyperparameter tuning was done in this study due to the prohibitively long training times of QCCNNs.

IV. QCCNNs WITH A QUANTUM POOLING LAYER

In the QCCNNs with a quantum pooling layer, the quantum convolution + pooling layer consists of data encoding with the higher order encoding described in section III-B, followed by one of the quantum pooling circuits detailed below. Only one qubit out of the four is then measured, as seen on figure 1. The quantum pooling layer is followed by a fully connected layer.

This quantum convolution + pooling operation yields one value as opposed to the quantum convolution where four values are returned. To achieve a fair comparison with the classical CNN and the QCCNN baseline described in section III-B, we therefore use four quantum kernels in parallel to form the quantum pooling architecture. This allows to maintain the same parameter count in the classical layers in all possible configurations. Moreover, this offers the advantage that four quantum kernels in parallel may allow to extract different features of the image. It is worth noting that the quantum pooling operation itself may be trainable. This is the case in pooling methods which comprise rotation gates, such as the pooling with mid-circuit measurements or the modular quantum pooling blocks.

A. Mid-circuit measurement

Mid-circuit measurements allow for the introduction of additional parameters while reducing circuit depth, offering a favorable trade-off between parameter count and circuit complexity. In the present study, mid-circuit measurements are used to control multiple qubits simultaneously - equivalent to a multi-qubit controlled gate. This is particularly advantageous because implementing multi-qubit controlled gates in practice can be challenging due to current hardware limitations. As such, this technique offers a more efficient and practical alternative to multi-qubit controlled gates with several hardware devices being able to perform mid-circuit measurements [28].

Mid-circuit measurements are used to implement alternative paths depending on the measured outcome. Given four qubits $q_0, q_1, q_2,$ and q_3 , qubit q_0 is measured and the rotation gates $R_{\theta_0}, R_{\theta_1},$ and R_{θ_2} are applied to qubits $q_1, q_2,$ and q_3 respectively if the measurement outcome of qubit q_0 is 1. Similarly, qubit q_1 is subsequently measured and if the measurement yields 1, the rotations gates R_{θ_3} and R_{θ_4} are applied to qubits q_2 and q_3 . Finally, a CNOT gate between q_2 and q_3 is applied, q_2 is measured and the rotation R_{θ_5} is applied to qubit q_3 if the measurement yields 1. The last qubit q_3 is measured and its value is fed to the classical fully-connected layer. The corresponding circuit is represented in figure 2a. Two different options using rotation gates with trainable angles R_X and R_Y are tested within this pooling method.

The training and validation accuracy curves are presented in figure 3 and directly compared with the classical CNN baseline and the QCCNN baseline without pooling. We first remark that the training accuracy of the hybrid quantum-classical variants all surpass the classical model. When generalizing on the validation set, the architecture with R_X mid-circuit measurements clearly performs better than the architecture with R_Y mid-circuit measurements. It also surpasses the two baselines in maximum validation accuracy, although the overlapping variance bands prevent us to conclude on its superiority.

Additional experiments were run with a bigger learning rate of 0.01, that are not shown in the results curves for sake of comparison. With this learning rate, although overfitting

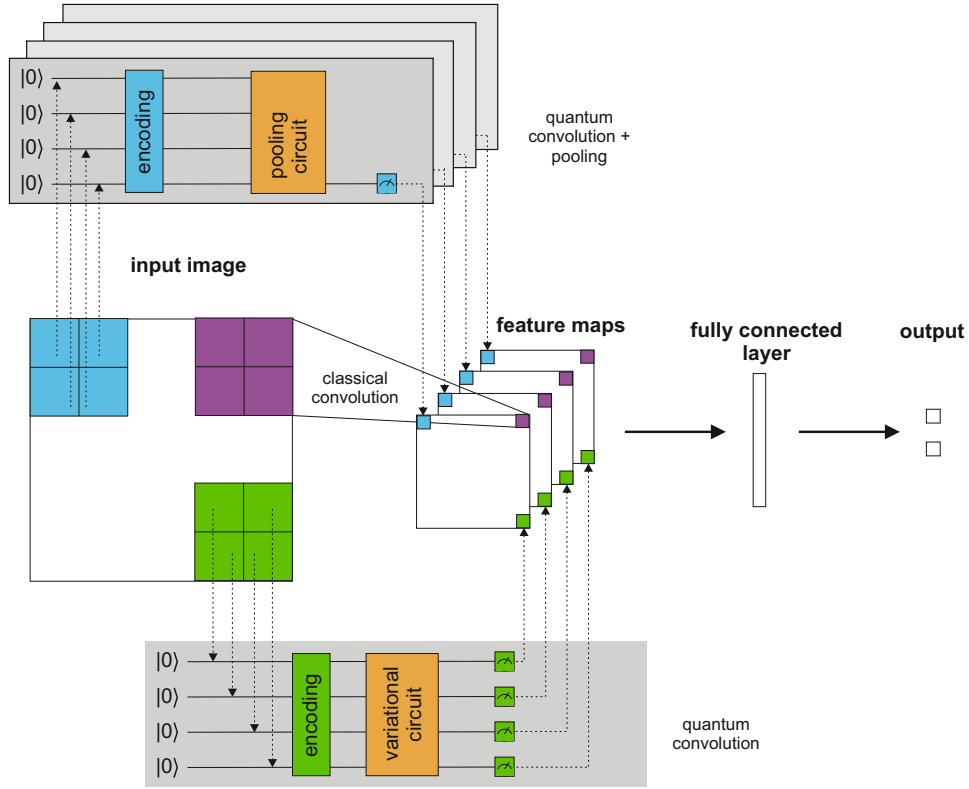


Fig. 1: Architecture sketch of the CNN and QCCNNs with quantum convolutional layers with and without pooling.

occurs very fast, R_X mid-circuit measurements is the only method presented within this paper where the training data was perfectly learnt with a training accuracy reaching 1.0 (validation accuracy of 88.47%). These results show that pooling methods with mid-circuits measurements have a high potential in extracting relevant features for the training of QCCNNs.

B. Ancilla qubit and controlled gates

Models with no inductive bias are likely to have trainability and generalization issues, therefore group-invariant quantum machine learning models have been proposed and one of their fundamental component is the use of controlled gates and ancillary qubits [29]. We draw upon these models and use a standard structure when using an ancilla qubit, i.e., Hadamard-controlled gates-Hadamard, where in this work the controlled gates are either CY or CZ.

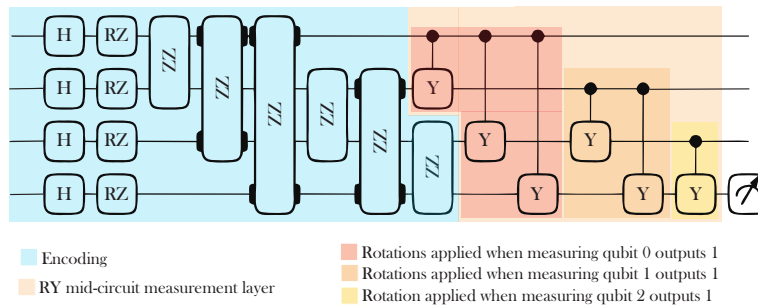
Our experiment involves initializing the system in a state consisting of four data qubits and one ancilla (control) qubit. Subsequently, the ancilla qubit is subjected to the Hadamard gate H . The input data is encoded on the first four qubits and a basic entangling layer is used. Following this, a series of four controlled gates is executed, wherein a controlled gate (either CY or CZ) is utilized with qubit q_0 as the control and the ancilla as the target, followed by a controlled gate with qubit q_1 as the control and the ancilla as the target, and so forth for

qubits q_2 and q_3 . Upon completion of these operations, the Hadamard gate H is applied to the ancilla qubit. Finally, the Pauli-Z expectation value of the ancilla qubit is taken as the output of the circuit. The architecture of the ancilla qubit with CY controlled gates pooling is shown on figure 2b.

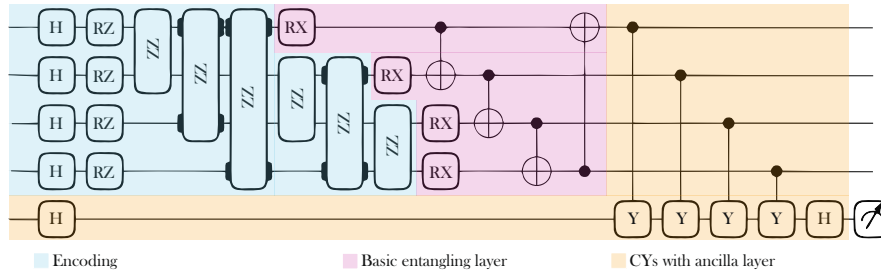
Our results in figure 4 show that the use of CY or CZ controlled gates does not affect the overall performance of the QCCNN, as the curves for both variants strictly overlap. Overall, the ancilla qubit with controlled gates pooling method seems to have a satisfactory training performance, reaching a maximum accuracy slightly above the classical CNN. However, this method does not perform well in validation accuracy, with a maximum validation accuracy lower than both classical CNN and QCCNN without pooling baselines.

C. Modular quantum pooling blocks

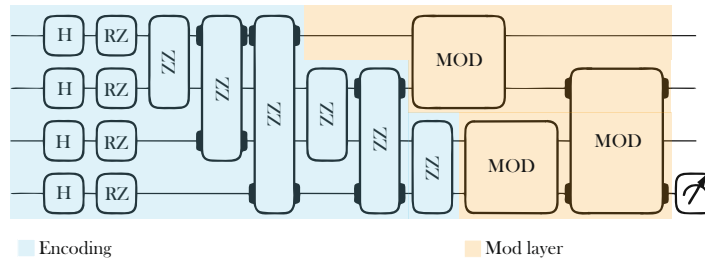
Taking inspiration from [10], we explore the alternative use of a combination of modular small quantum convolutional and pooling operations. The circuit architecture of this quantum pooling layer is displayed in figures 2c to 2f. The design of this pooling method incorporates modular blocks acting on two neighboring qubits (on the first and second qubit and on the third and fourth qubit). Each of these blocks either consists of just a quantum pooling operation (Mod-a), or of a quantum convolutional circuit in addition to a quantum pooling operation. For the convolutional part, two different



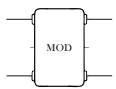
(a) Mid-circuit measurement pooling example. This circuit consists of a higher-order encoding, followed by a R_Y mid-circuit measurement layer.



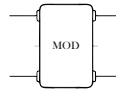
(b) Ancilla qubit with controlled gates pooling example. This circuit consists of a higher-order encoding, followed by a basic entangling layer and CYs with ancilla qubit.



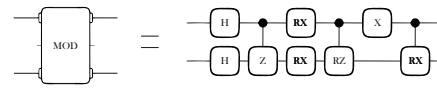
(c) Modular quantum pooling example. This circuit consists of a higher-order encoding, followed by modular quantum pooling blocks. The alternatives for the blocks are given in (d), (e) and (f).



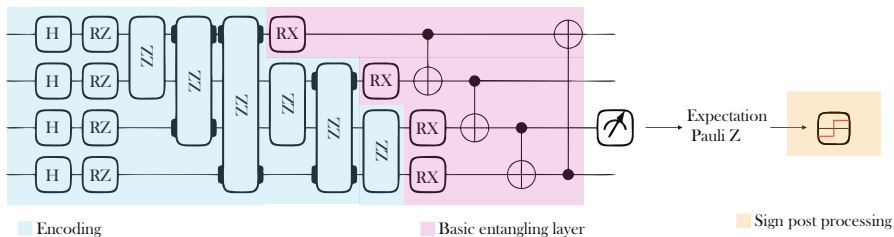
(d) Mod-a



(e) Mod-b

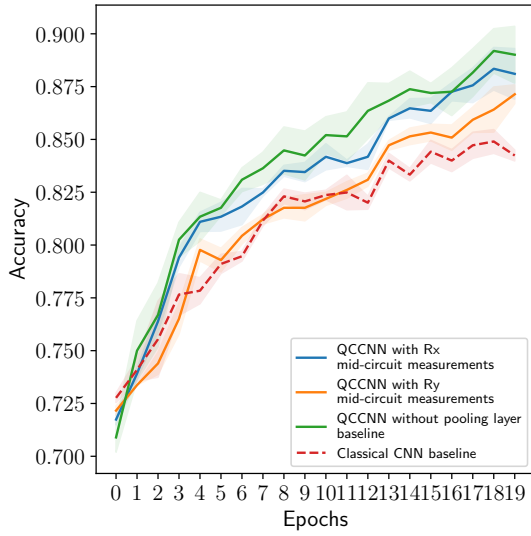


(f) Mod-c

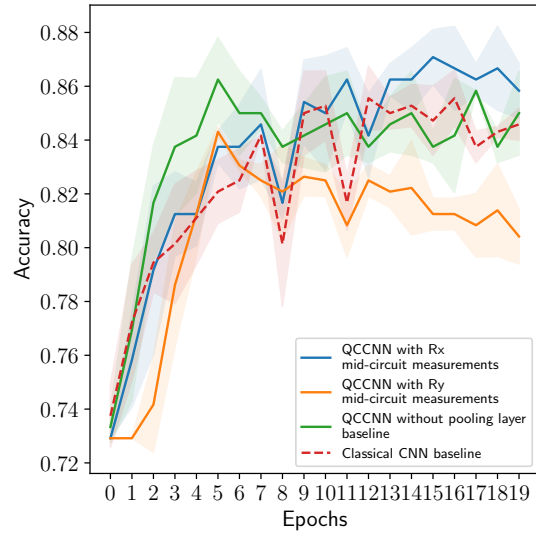


(g) Qubit selection with classical postprocessing pooling example. This circuit consists of a higher-order encoding, followed by a basic entangling layer and a qubit selection pooling with $Sign(x)$ postprocessing

Fig. 2: Architecture sketches of the tested pooling architectures.

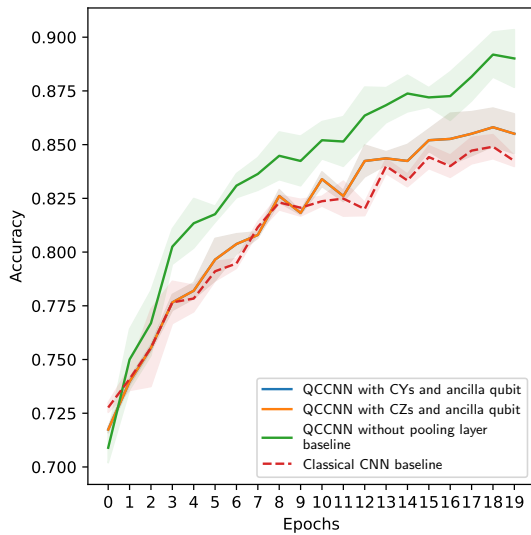


(a) Training accuracy

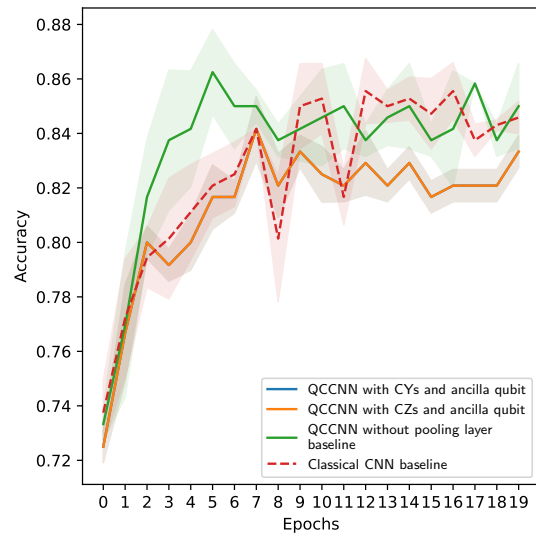


(b) Validation accuracy

Fig. 3: Comparison of the performance of mid-circuit measurement pooling methods for hybrid QCCNNs in terms of training and validation accuracy.



(a) Training accuracy



(b) Validation accuracy

Fig. 4: Comparison of the performance of ancilla qubit and controlled gates pooling methods for hybrid QCCNNs in terms of training and validation accuracy.

architectures are investigated, where the first variant is inspired by tree tensor network ideas (Mod-b), while the second circuit variant is known to have good entanglement properties [30] (Mod-c). As a result of the first two pooling blocks, a dimensionality reduction from four to two qubits occurs. The other qubits were traced out during the pooling operation. The remaining two qubits undergo another convolutional and pooling operation, using the same architecture as in the blocks used earlier in the circuit. Consequently, the whole quantum circuit performed a dimensionality reduction from four to one qubit. The remaining qubit is measured and then fed into the following classical parts in the network. Due to the modular design of this circuit architecture, the concept is easily extendable and adaptable to more complicated scenarios.

We obtain in general very good performance with the modular pooling, with two methods out of three outperforming the classical baseline in validation accuracy, namely Mod-a and Mod-c, and all three methods surpassing the classical baseline in training accuracy, as seen on figure 5. Particularly, the modular pooling method with Mod-c presents a substantial gain in performance compared to the classical baseline, with a fast and steep convergence early on in the training and clear and non overlapping error bands with the other models.

D. Qubit selection with classical postprocessing

Inspired by the work of [31], after encoding the data and using a basic entangling layer, one out of the four qubits from the 4-qubit system is selected, onto which a classical postprocessing step is applied. This qubit could either remain fixed throughout the computation or be randomly selected for each iteration. We choose to measure qubit q_2 in our experiments. With this last pooling method, it can be investigated whether quantum circuits profit from non-linear activation functions in the same way as classical algorithms do. The output of all tested circuits being an expectation value between -1 and 1, the $Sign(x)$ and $Tanh(x)$ non-linear activation functions are tested, thereby ensuring that the output of the quantum circuit remains within the same range. The circuit is shown in figure 2g. The $Sign(x)$ function is defined as

$$Sign(x) = \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ 1, & x > 0, \end{cases}$$

and $Tanh(x)$ is defined as

$$Tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

The results of the quantum pooling with these postprocessing functions are shown in figure 6. We find that the architecture with the $Tanh(x)$ activation is able to generalize very well and quickly on the validation data, outperforming both the classical baseline and the QCCNN without pooling at the third epoch in average validation accuracy. However, one notes that the variance in validation accuracy does not allow

us to unreservedly decipher between the performance of this variant and the QCCNN without pooling. The architecture with the $Sign(x)$ activation function is on the contrary not able to generalize properly on the validation data and suffers from remarkably wide variance bands. This probably stems from the information loss when the $Sign(x)$ function is applied.

E. Review of the best pooling options

A summary plot is provided in figure 7 showing the best performing configuration for each pooling option. It can be observed that most pooling architectures beat the classical CNN baseline in terms of the attained average maximal training and validation accuracy given a similar number of trainable parameters. One also notices a faster convergence for QCCNNs models in terms of validation accuracy as they achieve their highest validation accuracy sooner than the classical baseline, except for the mid-circuit measurements pooling method. These two observations are in line with what has been previously shown in theory [19]. In some cases however, it is recommended to exercise some level of caution when interpreting these results due to the overlapping variance bands.

When comparing the different pooling methods together, the QCCNN with Mod-c modular pooling clearly performs best on the chosen dataset in both training and validation accuracy, also when taking into account the errors bands. This method is followed in maximum validation accuracy by the QCCNN with qubit selection with $Tanh(x)$ classical postprocessing, and the QCCNN with R_X mid-circuit measurements, both options beating the QCCNN without pooling. Only the QCCNN with ancilla qubit pooling is performing less well than the baseline without pooling. In the next section, we attempt to give some insights on why some of these quantum circuits perform better than others.

V. EFFECTIVE DIMENSION ANALYSIS

In this final section, we aim to determine whether the performance of a specific VQC - and therefore a pooling method - could be foreseen using a quantum metric. Candidates for such a metric that were suggested in the literature are the effective dimension, the expressibility and the entanglement capability [30], [24]. However, the last two metrics assume a constant number of qubits in input and output of the VQC and were therefore not selected.

A. The effective dimension

The effective dimension (ED) is a complexity measure motivated by information geometry and designed to estimate the information capacity of the model. It can be computed for both classical and quantum neural networks by using the Fisher Information Matrix (FIM), a metric in statistics used to evaluate the impact of the variance of the models' parameters on its output. The FIM is given by

$$F(\theta) = \mathbb{E} \left[\frac{\partial}{\partial \theta} \log p(x, y; \theta) \frac{\partial}{\partial \theta} \log p(x, y; \theta)^T \right] \in \mathbb{R}^{d \times d}, \quad (1)$$

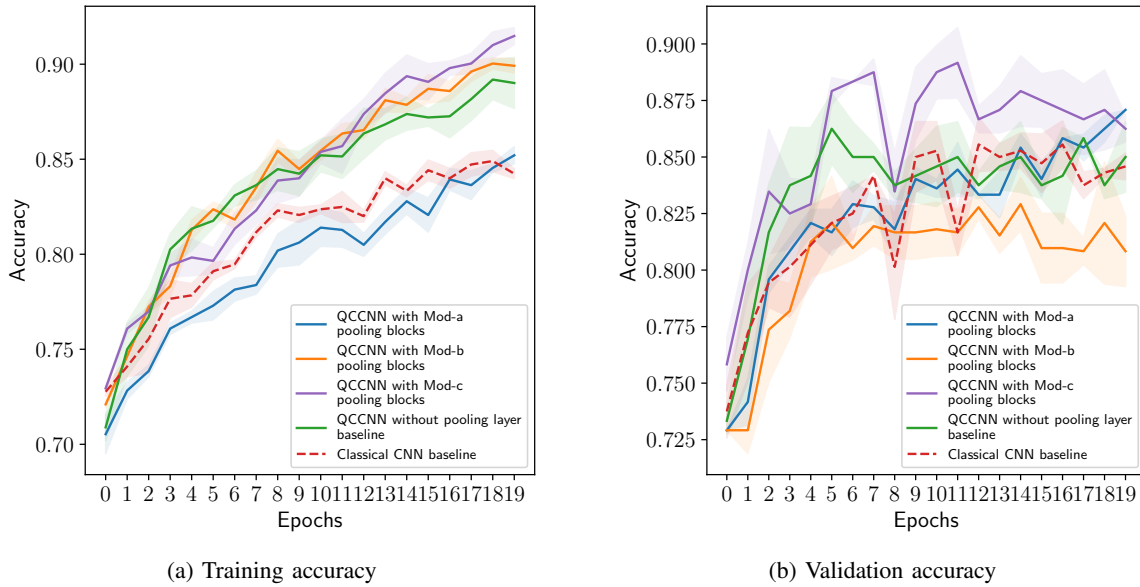


Fig. 5: Comparison of the performance of modular quantum pooling blocks for hybrid QCCNNs in terms of training and validation accuracy.

with $p(x, y; \theta)$ the conditional distribution describing the relationship between the input $x \in \mathbb{R}^{s_{in}}$, the output $y \in \mathbb{R}^{s_{out}}$, and the parameters $\theta \in \Theta$, where $\Theta \subset \mathbb{R}^d$ is the Riemannian space of the model parameters.

If one uses finite sampling, the FIM can be approximated empirically to

$$\tilde{F}_k(\theta) = \frac{1}{k} \sum_{j=1}^k \frac{\partial}{\partial \theta} \log p(x_j, y_j; \theta) \frac{\partial}{\partial \theta} \log p(x_j, y_j; \theta)^T, \quad (2)$$

where k is the number of independent and identically distributed (i.i.d) samples (x_j, y_j) drawn from $p(x, y; \theta)$. This formulation based on [32] was extended by [24] with the addition of the constant $\gamma \in (0, 1]$ and a $\log n$ term to assure the ED is bounded. The ED is therefore defined as

$$d_{\gamma, n}(\mathcal{M}_\Theta) = \frac{\log\left(\frac{1}{V_\Theta} \int_\Theta \sqrt{\det(\text{id}_d + \frac{\gamma n}{2\pi \log n} \hat{F}(\theta))} d\theta\right)}{\log\left(\frac{\gamma n}{2\pi \log n}\right)}, \quad (3)$$

where $n > 1 \in \mathbb{N}$ is the number of data samples, $V_\Theta \doteq \int_\Theta d\theta$ is the volume of the parameter space, and $\hat{F}(\theta)$ is the normalised FIM

$$\hat{F}_{ij}(\theta) = d \frac{V_\Theta}{\int_\Theta \text{tr}(F(\theta)) d\theta} F_{ij}(\theta). \quad (4)$$

The resulting ED is normalized by the number of training parameters for better comparison between the different VQCs.

B. Metric and model performance

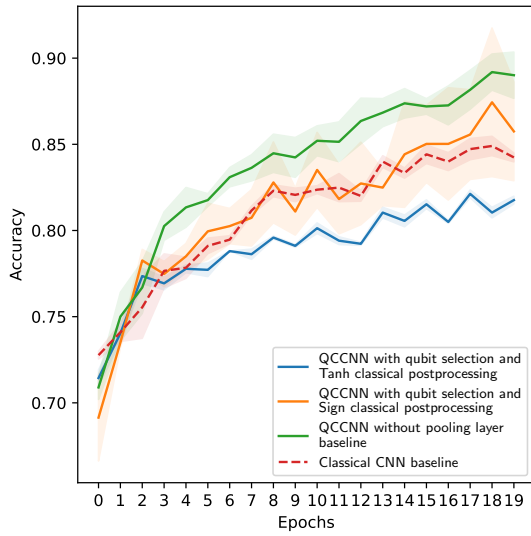
Table I shows the calculated ED of the VQCs for all tested quantum mechanisms as well as for the established QCCNN baseline without pooling. Note that we choose to focus on

the VQC architecture in the calculation of the ED, thus the classical postprocessing function is not included for the qubit selection with classical postprocessing method.

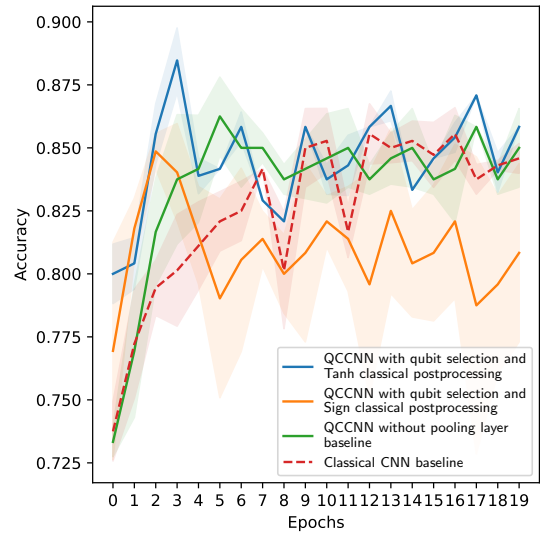
The ED being a measure of the information capacity of the model, we expect the training accuracy to be positively correlated with this metric. However, this is not a conclusion that can be drawn within this work. Mid-circuit measurements have a relatively high normalized ED and indeed perform well in training accuracy. On the other hand, Mod-a modular pooling also has a good normalized ED but the performance in training was worse. Among the lowest normalized ED values is Mod-c, which is in fact the best performing variant. If not taking into account Mod-b and Mod-c pooling options, we could in fact observe a positive correlation between ED and performance of the VQC. We hypothesize that such a relation might still exist, and that it should be further explored in future work. It can be for example studied whether the effect of different optimizers could reveal the potential link between ED and VQC performance.

VI. CONCLUSION

This work presented four different architectures - mid-circuit measurements, ancilla qubits with controlled gates, modular quantum pooling blocks and qubit selection with classical postprocessing - to realize quantum pooling in hybrid QCCNNs. For this purpose, the classical convolutional layer within a classical CNN was replaced by a quantum convolution and pooling layer. The use of QCCNNs is particularly interesting for situations with little training data, which is typically the case in classification tasks on medical images. Therefore, we studied the different proposed architectures on a small dataset featuring ultrasound images of the breast to classify lesions

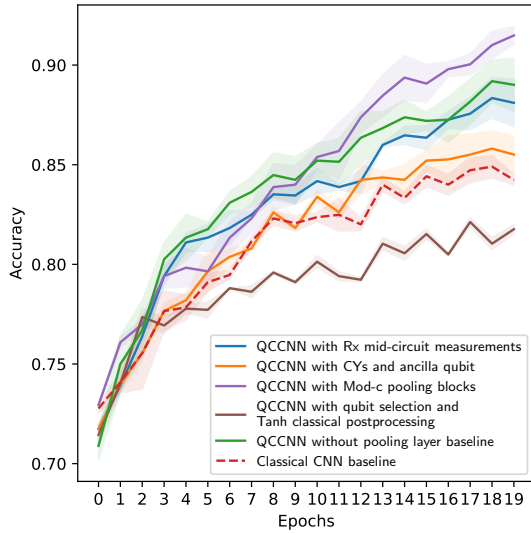


(a) Training accuracy

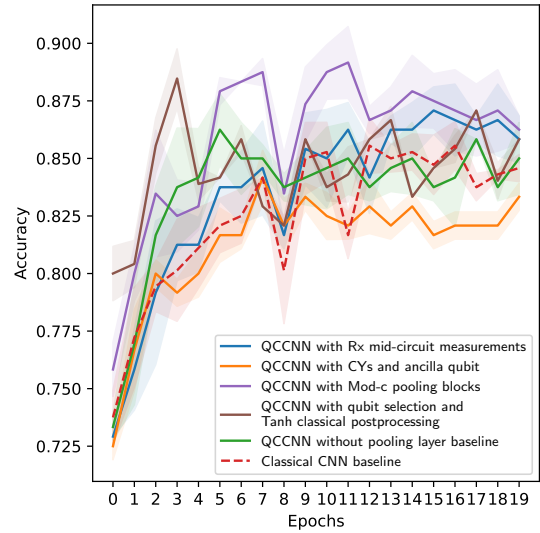


(b) Validation accuracy

Fig. 6: Comparison of the performance of qubit selection with classical postprocessing pooling methods for hybrid QCCNNs in terms of training and validation accuracy.



(a) Training accuracy



(b) Validation accuracy

Fig. 7: Comparison of the best-performing pooling methods for hybrid QCCNNs in terms of training and validation accuracy.

TABLE I: The ED of all trainable architectures reflected against the maximum training and validation accuracies obtained during training. The higher the ED, the better the VQC with respect to that metric.

VQC	Normalized ED	Max training accuracy	Max validation accuracy
R_X mid-circuit measurements	0.909 ± 0.016	92.81 ± 0.37	87.08 ± 1.02
R_Y mid-circuit measurements	0.906 ± 0.008	87.14 ± 0.44	84.31 ± 0.20
Ancilla qubit with CY controlled gates	0.772 ± 0.041	85.81 ± 0.89	84.17 ± 1.18
Ancilla qubit with CZ controlled gates	0.801 ± 0.042	85.81 ± 0.89	84.17 ± 1.18
Modular pooling Mod-a	0.726 ± 0.006	85.21 ± 0.45	87.08 ± 0.00
Modular pooling Mod-b	0.149 ± 0.003	90.04 ± 0.30	82.92 ± 1.18
Modular pooling Mod-c	0.249 ± 0.016	91.49 ± 0.44	89.17 ± 1.56
Qubit selection for classical postprocessing	0.666 ± 0.048	82.13 ± 0.23	88.47 ± 1.29
Basic entangling layer without pooling	0.933 ± 0.019	89.19 ± 1.05	86.25 ± 1.56

as benign or malign. We find that quantum pooling variants perform at least similar or even in some cases clearly overcome the equivalent classical CNN and hybrid QCCNN without pooling, while using a similar number of trainable parameters. On the selected dataset, the modular pooling technique beat all variants on both training and validation sets. The mid-circuit measurements pooling seems to be particularly promising in terms of training ability, while the qubit selection with classical postprocessing exhibits very good generalization ability. These results, obtained in simulation, are very promising for further studies of QCCNNs, as those might be suited for presently available NISQ hardware already, given the small number of qubits required and the shallow depth of the quantum circuits involved. The presented pooling techniques are naturally also relevant in the design choice for fully quantum architectures. However, to fully exploit the potential benefits, additional work to produce insights about which circuit architectures will result in benefits for a given use case should be pursued.

VII. ACKNOWLEDGEMENTS

The project/research is supported by the Bavarian Ministry of Economic Affairs, Regional Development and Energy with funds from the Hightech Agenda Bayern.

REFERENCES

- [1] Peter Wittek. *Quantum Machine Learning: What Quantum Computing Means to Data Mining*. Elsevier Science, 2016.
- [2] Maria Schuld and Nathan Killoran. *Machine Learning with Quantum Computers*. Springer, Cham, 2021.
- [3] Suchi Ganguly. *Quantum Machine Learning: An Applied Approach*. Apress Media, 2021.
- [4] Marco Cerezo, Akira Sone, and Sándor Volkán-Kacsó. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 2021.
- [5] Irwin Cong, Soonwon Choi, and Mikhail D. Lukin. Quantum convolutional neural networks. *Nature Physics*, pages 1–6, 2019.
- [6] Iordanis Kerenidis, Jesse Landman, and Anupam Prakash. Quantum algorithms for deep convolutional neural networks. In *International Conference on Learning Representations*, 2020.
- [7] Yang Lü, Yaoyun Liu, Haifeng Lin, Yaoyun Shi, and Guoqi Zhang. A quantum convolutional neural network for image classification. In *2021 40th Chinese Control Conference (CCC)*, pages 6329–6334, 2021.
- [8] Shengyu Wei, Shun Wu, Haotong Jia, Zhangqi Xu, and Xin Yao. A quantum convolutional neural network on nisq devices. *AAPPS Bulletin*, 32:012022, 2022.
- [9] Yanpeng Li, Yaoyun Liu, Haifeng Lin, Yaoyun Shi, and Guoqi Zhang. A quantum deep convolutional neural network for image recognition. *Quantum Science and Technology*, 5:062020, 2021.
- [10] Taesik Hur, Lin Kim, and Dong-Kyu Park. Quantum convolutional neural network for classical data classification. *Quantum Machine Intelligence*, 4:3, 2022.
- [11] Matthew P Henderson, Vladimir Vapnik, John A Smolin, Anthony Megrant, Kevin Chen, Ben Chiaro, Andrew Dunsworth, Chris Quintana, Peter JJ O’Malley, Rami Barends, et al. Quantum convolutional neural networks: powering image recognition with quantum circuits. *Quantum Machine Intelligence*, 2(1):1–9, 2020.
- [12] David Mattern, Yuan Zhang, Yudong Cao, Zhihui Zhou, Xin Cai, Yongchao Yao, Haoyu Li, Jin-Ming Liu, Yanfang Ye, Chao Zou, et al. Variational quantum convolutional neural networks with enhanced image encoding. *arXiv preprint arXiv:2106.07327*, 2021.
- [13] Jikai Liu, Luyao Sun, Mengzhi Zhang, Haoyu Lu, Yudong Cao, Yanfang Ye, Zhihui Zhou, Xin Wang, and Hong Zhang. Hybrid quantum-classical convolutional neural networks. *Science China Physics, Mechanics & Astronomy*, 64(9), 2021.
- [14] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [15] Andrea Matic, Maureen Monnet, Jeanette Miriam Lorenz, Balthasar Schachtner, and Thomas Messerer. Quantum-classical convolutional neural networks in radiological image classification. *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 56–66, 2022.
- [16] Jin Zheng, Qing Gao, and Yanxuan Lü. Quantum graph convolutional neural networks. In *2021 40th Chinese Control Conference (CCC)*, pages 6335–6340, 2021.
- [17] Ludmila Botelho, Adam Glos, Akash Kundu, Jarosław Adam Miszczak, Özlem Salehi, and Zoltán Zimborás. Error mitigation for variational quantum algorithms through mid-circuit measurements. *Physical Review A*, 105(2):022441, 2022.
- [18] Tak Hur, Leeseok Kim, and Daniel K. Park. Quantum convolutional neural network for classical data classification. *Quantum Machine Intelligence*, 4:3, 2022. arXiv:2108.00661 [quant-ph].
- [19] María Camila Caro, Juan Pablo Romero, and Alán Aspuru-Guzik. Generalization in quantum machine learning from few training data. *arXiv preprint arXiv:2111.05292*, 2021.
- [20] Luciano M Prevedello, Baris S Erdal, Ji H Ryu, Kevin J Little, Mutlu Demirel, Steven Qian, and James A Brink. Challenges related to artificial intelligence research in medical imaging and the importance of image analysis competitions. *Radiology: Artificial Intelligence*, 1(1):e180031, 2019.
- [21] Martin J Willeminck, Wojciech A Koszek, Carl Hardell, Chih-Wei Wu, Matthew A Flicek, Michael Ingrisich, Kerstin U Bauner, Bruce L Eck, Jennifer A Christner, and Howard P Forman. Preparing medical imaging data for machine learning. *Radiology*, 295:4–15, Apr 2020.
- [22] Jia Yang and et al. Medmnist v2: A large-scale lightweight benchmark for 2d and 3d biomedical image classification. *CoRR*, 2021.
- [23] Walaa Al-Dhabyani and et al. Dataset of breast ultrasound images. *Data in Brief*, 28:104863, 2020.
- [24] Ahmed Abbas and et al. The power of quantum neural networks. *Nature Computational Science*, 1:403–409, Jun 2021.
- [25] Vojtech Havlicek and et al. Supervised learning with quantum enhanced feature spaces. *Nature*, 567(7747):209–212, Mar 2019.
- [26] Adam Paszke and et al. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, volume 32, pages 8026–8037. Curran Associates, Inc., 2019.
- [27] Ville Bergholm et al. PennyLane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*, Feb 2020.
- [28] L. C. G. Govia, P. Jurcevic, S. T. Merkel, and D. C. McKay. A randomized benchmarking suite for mid-circuit measurements, 2022.
- [29] Martin Larocca, Frédéric Sauvage, Faris M Sbahi, Guillaume Verdon, Patrick J Coles, and Marco Cerezo. Group-invariant quantum machine learning. *PRX Quantum*, 3(3):030341, 2022.
- [30] Seungwoo Sim, Peter D Johnson, and Alán Aspuru-Guzik. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2(10), Oct 2019.
- [31] Maria Schuld, Alex Bocharov, Krysta M Svore, and Nathan Wiebe. Circuit-centric quantum classifiers. *Physical Review A*, 101(3):032308, 2020.
- [32] Oksana Berezniuk, Alessio Figalli, Raffaele Ghigliazza, and Kharen MUSAELIAN. A scale-dependent notion of effective dimension. *arXiv preprint arXiv:2001.10872*, 2020.