Quantum classifiers for domain adaptation

Xi He\textsuperscript{1,2,3} · Feiyu Du\textsuperscript{1,2} · Mingyuan Xue\textsuperscript{1,2} · Xiaogang Du\textsuperscript{1,2} · Tao Lei\textsuperscript{1,2} · A. K. Nandi\textsuperscript{4,5}

Abstract
Transfer learning (TL), a crucial subfield of machine learning, aims to accomplish a task in the target domain with the acquired knowledge of the source domain. Specifically, effective domain adaptation (DA) facilitates the delivery of the TL task where all the data samples of the two domains are distributed in the same feature space. In this paper, two quantum implementations of the DA classifier are presented with quantum speedup compared with the classical DA classifier. One implementation, the quantum basic linear algebra subroutines-based classifier, can predict the labels of the target domain data with logarithmic resources in the number and dimension of the given data. The other implementation efficiently accomplishes the DA task through a variational hybrid quantum-classical procedure.

Keywords Domain adaptation · Quantum machine learning · Transfer learning · Quantum algorithm · Machine learning

1 Introduction

Transfer learning (TL), a significant subfield of machine learning, attempts to accomplish tasks in an unknown domain with the learnt knowledge of a different, but related
domain [1]. As a crucial sub-realm of the TL, domain adaptation (DA) aims to predict the labels of an unlabelled target domain with a given labelled source domain data where all the given data are distributed in the same feature space. DA is significantly applicable in dealing with the unprocessed data and has been widely used in various fields such as computer vision [2–4], natural language processing [5], and reinforcement learning [6].

Quantum computing is a type of pattern computing which is typically based on quantum mechanics [7–11]. In recent years, quantum computing techniques have been applied to the field of machine learning to accomplish tasks with the promotion of the algorithm performance [12–14]. For instance, quantum computation techniques can achieve supervised learning tasks such as classification [15–17], data fitting [18, 19], and unsupervised learning such as clustering [20], dimensionality reduction [21, 22] with quantum speedup. In the field of deep learning, quantum Boltzmann machine [23, 24], quantum generative adversarial learning [25–30], quantum auto-encoder [31–33], and quantum neural networks [34, 35] have been proposed to deal efficiently with deep learning tasks on quantum devices. For the TL, Ref. [36] systematically analyzes the framework of the quantum transfer learning in different scenarios. Refs. [37, 38] utilize linear transformation to align the source domain to the target domain to accomplish the procedure of DA. However, the procedure of the DA such as in refs. [37, 38] and the corresponding labels prediction are specifically separated in the existing quantum DA algorithms resulting in an increase of the computational complexity. Compared with the algorithms presented in refs. [37, 38], the domain adaptation classifier in our work combines the procedure of domain adaptation and the final target label prediction. Only one learning process is required in our model resulting in lower algorithmic complexity and avoiding the occurrence of learning mismatch.

In this paper, two quantum implementations of the DA classifier are presented. One implementation utilizes the quantum basic linear algebra subroutines to achieve exponential speedup on the universal quantum computer compared to the classical DA classification algorithm. The other implementation, the variational quantum DA classifier, accomplishes the procedure of DA on the near-term quantum devices through a variational hybrid quantum-classical procedure.

The remainder of this paper is arranged as follows. Firstly, the classical DA classifier is briefly overviewed in Sect. 2. Subsequently, the QBLAS-based quantum DA classifier, the variational quantum DA classifier are presented respectively in Sect. 3. In addition, two numerical experiments are provided to demonstrate the feasibility and the effectiveness of the quantum domain adaptation classifier in Sect. 4. Finally, some open problems and future work are discussed in Sect. 5.

2 Preliminaries

Assume that we are given a source domain dataset $\mathcal{D}_s = \{x_i^{(s)}\}_{i=1}^{n_s} \in \mathbb{R}^D$ with labels $(y_i^{(s)})_{i=1}^{n_s} \in \{0, 1\}$ and an unlabelled target domain dataset $\mathcal{D}_t = \{x_j^{(t)}\}_{j=1}^{n_t} \in \mathbb{R}^D$. The source domain data matrix is $X_s = (x_1^{(s)}, \ldots, x_{n_s}^{(s)}) \in \mathbb{R}^{D \times n_s}$ and the target domain data matrix is $X_t = (x_1^{(t)}, \ldots, x_{n_t}^{(t)}) \in \mathbb{R}^{D \times n_t}$. The feature and the label space of $\mathcal{D}_t$ are
exactly the same as $D_s$. However, the data of the source and target domain specifically obey different data distributions. The goal of the classifier for domain adaptation is to predict the labels of an unknown target domain with the help of the labelled source domain data [39]. Let $\mu_c^{(s)} (\mu_c^{(t)})$, $\Sigma_s$ ($\Sigma_t$) be the $c$th class mean and the covariance of the source (target) domain. In this paper, the discussion is specifically restricted to the task of binary classification, namely $c = \{0, 1\}$. The binary domain adaptation classifier can be easily extended to the circumstance of multi-class. Specifically, $\mu_1$, $\mu_0$ represent the positive and negative class means of the given data respectively in the source and target domain.

The DA classifier achieves the procedure of the transfer learning with a modified classifier inspired from the linear discriminant analysis (LDA) [40]. The scoring function of the classifier is defined as

$$y(x) = w^T x$$

(1)

to determine the label of the specified data point. In the spirit of the LDA, the source domain data $D_s$ are generated from the distribution $p(x^{(s)}, y^{(s)}) = p(x^{(s)}|y^{(s)})p(y^{(s)})$ where $p(y^{(s)})$ is the prior of the labels; $p(x^{(s)}|y^{(s)}) = \mathcal{N}(x^{(s)}; \mu_c^{(s)}, \Sigma_s)$ represents the class-conditional distributions. The weight vector of the classifier is $w^{(s)} = \Sigma_s^{-1/2}(\mu_1^{(s)} - \mu_0^{(s)})$. Equivalently, the classifier can be obtained by projecting the decorrelated source domain data $\hat{x} = \Sigma_s^{-1/2} x^{(s)}$ to the difference between the decorrelated means $\hat{w} = (\hat{\mu}_1^{(s)} - \hat{\mu}_0^{(s)}) = \Sigma_s^{-1/2}(\mu_1^{(s)} - \mu_0^{(s)})$ where $\hat{\mu}_c^{(s)} = \Sigma_s^{-1/2} \mu_c$ for $c = \{0, 1\}$. However, this classifier cannot be directly applied to the target domain directly due to the domain shift between the source and target domain. The DA classifier adaptively apply the decorrelated target domain data $\hat{x}^{(t)} = \Sigma_t^{-1/2} x^{(t)}$ to the decorrelated mean difference $\hat{w}$ resulting in the scoring function

$$\hat{y}(x^{(t)}) = \hat{w}^T \hat{x}^{(t)} = (\Sigma_s^{-1/2}(\mu_1^{(s)} - \mu_0^{(s)}))^T (\Sigma_t^{-1/2} x^{(t)}).$$

(2)

The DA classifier utilizes the DA techniques to modify the traditional model for classification to effectively accomplish the machine learning tasks in different domains. Compared with other DA models, the DA classifier effectively combines the procedure of TL with the label prediction resulting in a concise DA model. The schematic diagram of the DA classifier is presented as Fig. 1.

3 Methods

3.1 State preparation

Given the source domain data $X_s$ and the target domain data $X_t$, the quantum states corresponding to the $X_s$ and $X_t$ are

$$|\psi_{X_s}\rangle = \sum_{i=1}^{n_s} \sum_{m=1}^{D} x_{mi}^{(s)} |i\rangle|m\rangle = \sum_{i=1}^{n_s} |i\rangle|x_i^{(s)}\rangle,$$

(3)
Fig. 1 The schematic diagram of the DA classifier. The circle and the ellipse in dotted line represent the global topological structure. The covariance matrix $\Sigma_s$ is applied to the mean of the data feature without background class resulting in the weight of the DA classifier. The target domain data is restructured according to the source domain data. Combined with the decorrelated target domain data, the DA classifier can be constructed

$$
|\psi_{X_t}\rangle = \sum_{j=1}^{n_t} \sum_{m=1}^{D} x_{m}^{(t)} |j\rangle |m\rangle = \sum_{j=1}^{n_t} |j\rangle |x_j^{(t)}\rangle,
$$

respectively in amplitude encoding [41] where $\sum_{m,i} |x_{mi}^{(s)}|^2 = \sum_{m,j} |x_{mj}^{(t)}|^2 = 1$; the registers $|i\rangle$, $|m\rangle$ are the orthogonal basis of the given data in the dimension of data scale and feature respectively. Hence, the states which represent the covariance matrices of the source and target domain data are

$$
\rho_s = \text{tr}_i \{ |\psi_{X_s}\rangle \langle \psi_{X_s}| \} = \sum_{m,m'=1}^{D} \sum_{i=1}^{n_s} x_{mi}^{(s)} x_{m'i}^{(s)*} |m\rangle \langle m' |,
$$

$$
\rho_t = \text{tr}_j \{ |\psi_{X_t}\rangle \langle \psi_{X_t}| \} = \sum_{m,m'=1}^{D} \sum_{j=1}^{n_t} x_{mj}^{(t)} x_{mj}^{(t)*} |m\rangle \langle m' |,
$$

respectively where $\text{tr}_i$ is the trace over the $i$ register. The quantum states $|\mu_c^{(s)}\rangle$ ($c = 0, 1$) representing the source domain mean value for the two classes can be obtained by the quantum adder proposed in Ref. [42–44]. The quantum state $|\psi_u\rangle$ which represents
the vector $\mu_1^{(s)} - \mu_0^{(s)}$ can be computed by the quantum subtractor presented in Ref. [22].
In addition, the data matrices $X_s, X_t$ can be extended to $\tilde{X}_s = |0\rangle\langle 1| \otimes X_s + |1\rangle\langle 0| \otimes X_s^\dagger$ and $\tilde{X}_t = |0\rangle\langle 1| \otimes X_t + |1\rangle\langle 0| \otimes X_t^\dagger$ respectively.

### 3.2 QBLAS-based DA classifier

For the QBLAS-based DA classifier, we assume that the elements of $X_s$ and $X_t$ are accessible by the quantum random access memory (qRAM) [45] in time $O(poly(\log(D\tilde{n})))$ with $O(poly(D\tilde{n}))$ resources where $\tilde{n} = \max(n_s, n_t)$. The corresponding quantum circuit of the QBLAS-based DA classifier is depicted as Fig. 2.

The whole procedure of the QBLAS-based DA classifier is presented as follows.

(1) Apply the quantum phase estimation algorithm (QPE) [46, 47]

$$U_{PE}(\tilde{X}_t) = (\text{QFT}^\dagger \otimes I)(\sum_{\tau=0}^{T-1} |\tau\rangle\langle \tau| \otimes e^{i\tilde{X}_t \tau T} \otimes \text{H}^\otimes n \otimes I)$$

(7)

on the input state $|\psi_{\text{init}}\rangle = |0\rangle|\psi_{X_t}\rangle|0\rangle^\otimes poly(D+n_t)$ prepared by the specified registers resulting in the quantum state

$$|\psi_1\rangle = U_{PE}(\tilde{X}_t)|\psi_{\text{init}}\rangle$$

$$= \sum_{j=1}^{n_t} |j\rangle \sum_{m=1}^{D} \beta_{mj}^{(t)} |\sigma_{m}^{(t)}\rangle \frac{1}{\sqrt{2}}(|w_{m}^{(t)+}\rangle - |w_{m}^{(t)-}\rangle)$$

$$= |1\rangle \sum_{j=1}^{n_t} |j\rangle \sum_{m=1}^{D} \beta_{mj}^{(t)} |\sigma_{m}^{(t)}\rangle |v_{m}^{(t)}\rangle,$$

(8)

in $O(1/\epsilon)$ with error $\epsilon$, where $\tilde{X}_t = \begin{bmatrix} 0 & X_t^\dagger \\ X_t & 0 \end{bmatrix}$; $\text{QFT}^\dagger$ represents the inverse quantum Fourier transform [48]; $\text{H}, I$ represent the Hadamard gate and the identity matrix respectively; $\sum_{\tau=0}^{T-1} |\tau\rangle\langle \tau| \otimes e^{i\tilde{X}_t \tau T}$ is the conditional Hamiltonian evolution; $\beta_{mj}^{(t)} = (u_{m}^{(t)} |x_{j}^{(t)}\rangle, |w_{m}^{(t)}\rangle = \frac{1}{\sqrt{2}}(|0\rangle|u_{m}^{(t)}\rangle \pm |1\rangle|v_{m}^{(t)}\rangle)$ are the eigenvectors of $\tilde{X}_t$ corresponding to the singular values $|\sigma_{m}^{(t)}\rangle$ with the left singular vector $u_{m}^{(t)}$ and the right singular vector $v_{m}^{(t)}$.

(2) Perform the controlled $R_y(2 \arccos(\gamma_{t} / |\sigma_{m}^{(t)}\rangle))$ operation on the first register to obtain the state

$$|\psi_2\rangle = |\psi_{a}^{(t)}\rangle \sum_{j=1}^{n_t} |j\rangle \sum_{m=1}^{D} \beta_{mj}^{(t)} |\sigma_{m}^{(t)}\rangle |v_{m}^{(t)}\rangle,$$

(9)
where

$$|\psi_a^{(t)}\rangle = \sqrt{1 - \frac{\gamma_t^2}{\sigma_m^{(t)}|^2}} |0\rangle + \frac{\gamma_t}{\sigma_m^{(t)}} |1\rangle,$$

(10)

\(\gamma_t\) is the normalization coefficient.

(3) Uncompute the \(|\sigma_m^{(t)}\rangle\) register, remove the ancilla register, and measure the \(|\psi_a^{(t)}\rangle\) to be \(|1\rangle\). The state

$$|\psi_{\tilde{X}_t}\rangle = \sum_{j=1}^{n_t} |j\rangle \sqrt{\frac{1}{\Sigma_m^{(t)} |\sigma_m^{(t)}|^2} \sum_{m=1}^{D} \rho_m^{(t)} \rho_t \sum_{m=1}^{D} \frac{\gamma_t^3}{\sigma_m^{(t)}} |v_m^{(t)}\rangle},$$

$$= \sum_{j=1}^{n_t} |j\rangle \sqrt{\frac{\Sigma_j^{(t)} |x_j^{(t)}|^2}{\Sigma_j^{(t)} \Sigma_j^{(t-1)} |x_j^{(t)}|^2}} \sqrt{\frac{\Sigma_j^{(t)} |x_j^{(t)}|^2}{\Sigma_j^{(t)} \Sigma_j^{(t-1)} |x_j^{(t)}|^2}}$$

$$= \sum_{j=1}^{n_t} |j\rangle |\tilde{x}_j^{(t)}\rangle$$

(11)

can be obtained. Thus, the quantum state \(|\tilde{x}_j^{(t)}\rangle\) proportional to \(\Sigma_j^{(t)} x_j^{(t)}\) for \(j = 1, 2, \ldots, n_t\) can be computed in time \(O(\|X_t\|_{\text{max}}^2 \log^2(D + n_e)/\epsilon^3)\) where \(\|X_t\|_{\text{max}}\) is the largest absolute element of \(X_t\) and \(\epsilon\) is the error parameter [14].

The whole procedure above can be represented as the following unitary operation

$$U_M(X, \theta) = (I \otimes U_{PE}(X))(U_{R_1}(\theta) \otimes I)(I \otimes U_{PE}(X))$$

(12)
where \( |\psi_{\tilde{X}_t}\rangle\) can be achieved by the operation \( U_M(\tilde{X}_t, 2 \arcsin(\gamma_t/|\sigma^{(t)}_m|))\). Similarly, the quantum state \( |\tilde{w}_q\rangle\) proportional to the vector \( \Sigma_s^{-1/2}(\mu_1^{(s)} - \mu_0^{(s)}) \) by applying \( U_M(\tilde{X}_s, 2 \arcsin(\gamma_s/|\sigma^{(s)}_m|))\) on the input quantum state \( |0\rangle^R|0\rangle^C|\psi_\mu\rangle^B|0\rangle^{\otimes \log(D+n_s)}S\) in time \( O(\|X_s\|_{\text{max}} \log^3(D+n_s)/\epsilon^3)\) [14] where \( \|X_s\|_{\text{max}}\) is the largest absolute element of \( X_s\).

(4) The scoring function of the QBLAS-based DA classifier can be ultimately obtained as
\[
\hat{y}_q(x^{(t)}) = \langle \tilde{w}_q | \hat{x}^{(t)} \rangle
\]
by performing the swap test [49] on \( |\tilde{w}\rangle, |\hat{x}_t\rangle\).

### 3.3 Variational quantum domain adaptation classifier

In addition to the design based on quantum basic linear algebra subroutines, the DA classifier can be alternatively implemented on noisy intermediate-scale quantum devices (NISQ) through a variational hybrid quantum-classical procedure. The variational quantum domain adaptation classifier (VQDAC) can be performed on the near-term quantum devices without high-depth quantum circuits and fully coherent evolution required by the QBLAS-based DA classifier. The pseudo-code of the VQDAC is presented in Algorithm 1.

For a \( 2^n \)-dimensional classical data point, the corresponding quantum state can be generated by a quantum circuit of \( O(n^2 + \frac{2n}{k+n}) \) depth with \( k \) ancilla qubits [50]. Based on the time-space tradeoff, the quantum states corresponding to the given data can be obtained by low-depth quantum circuits with sufficient quantum qubits. If we are given quantum data initially, the VQDAC can be invoked directly as follows. The VQDAC firstly diagonalizes the quantum states \( \rho_s \) and \( \rho_t \) prepared as Sect. 3.1 to obtain the states \( |\psi_s\rangle \) and \( |\psi_t\rangle \) to represent the matrix \( \Sigma_s \) and \( \Sigma_t \) respectively. In the spirit of Ref. [51], design \( \tilde{\rho}_s = U(\theta_s)\rho_s U^\dagger(\theta_s) \) with the unitary operation \( U(\theta_s) \) constructed by a parameterized quantum circuit where \( \{\theta_s\} \) is a set of parameters. The cost function is defined as
\[
C = \text{Tr}(\tilde{\rho}_s H_s),
\]
where \( H_s \) is a specified \( D \)-qubit Hamiltonian with \( D \) non-negative and non-degenerate eigenvalues. By minimizing the cost function \( C \) with a classical optimization algorithm, the optimal parameters \( \{\theta_s^*\} \) can be obtained. \( \rho_s \)'s eigenvalues \( \{\lambda_i^{(s)}\}_{i=1}^{D} \) can be estimated by measuring \( \tilde{\rho}_s \). Thus, the source domain covariance matrix \( \tilde{\Sigma}_s^{-\frac{1}{2}} = \sum_{i=1}^{D} \lambda_i^{(s)} \frac{1}{2} |i\rangle\langle i| \) can be finally computed, along with the target domain covariance matrix \( \tilde{\Sigma}_t^{-\frac{1}{2}} = \sum_{j=1}^{D} \lambda_j^{(t)} \frac{1}{2} |j\rangle\langle j| \). The quantum circuit of the state diagonalization of the source and target data is presented in Fig. 3.

Subsequently, the quantum states \( |w\rangle \) and \( |\hat{x}^{(t)}\rangle \) are computed to represent the weight vector \( \hat{w} \) and the target domain whitened data \( \hat{X}^{(t)} = \Sigma_t^{-\frac{1}{2}} X_t \) respectively.
Algorithm 1 VQDAC

**Input:** Source domain data $X_s$ with labels $Y_s$, target domain data $X_t$.

**Output:** Target domain labels $Y_t$.

**step 1:** Prepare the quantum states $\rho_s, \rho_t$ by the low-depth quantum circuits.

**step 2:** Diagonalize $\rho_s, \rho_t$ to construct $\Sigma_1^2, \Sigma_t^2$ respectively.

**step 3:** Invoke the variational quantum linear solver to compute the decorrelated target domain data $|\hat{x}^{(t)}\rangle$ and the weight coefficient $|w\rangle$.

**step 4:** Perform swap test on $|w\rangle$ and $|\hat{x}^{(t)}\rangle$ to predict the target labels $y^{(t)}$.

**Fig. 3** The circuits of the diagonalization of a given state $\rho$ through a variational hybrid quantum-classical procedure

Inspired from Ref. [52], design the quantum ansatz $|\hat{x}^{(t)}(\theta^{(t)})\rangle$ with a set of parameters $\{\theta^{(t)}\}$. The cost function

$$\mathcal{L} = 1 - \frac{1}{n_t} \sum_{j=1}^{n_t} \left| \frac{\langle x_j^{(t)} | \Sigma_t^{-1/2} | \hat{x}^{(t)} \rangle}{\sqrt{\langle x_j^{(t)}(\theta^{(t)}) | \Sigma_t^{-1/2} \Sigma_t^{-1/2} | x_j^{(t)}(\theta^{(t)}) \rangle}} \right|$$

is defined to be minimized by the classical optimization algorithm such as stochastic gradient descent to obtain the optimal coefficients $\{\theta^{(t)}_s\}$ and the decorrelated target domain data $|\hat{x}^{(t)}(\theta^{(t)}_s)\rangle$ for $j = 1, \cdots, n_t$ in time $O(\kappa_t/\epsilon)$, where $\kappa_t$ is the condition number of $\Sigma_t$ and $\epsilon$ is the error coefficient. Similarly, the quantum state $|w\rangle$ which represents the weight of the DAC $w = \Sigma_s^{-1/2}(\mu_1^{(s)} - \mu_0^{(s)})$ can be computed in the runtime $O(\kappa_s/\epsilon)$ where $\kappa_s$ is the condition number of the source domain covariance matrix $\Sigma_s$ [52].

Ultimately, the label of the target domain data point $x^{(t)}$ can be obtained according to the success probability of performing the swap test on the two states $|w\rangle$ and $|\hat{x}^{(t)}\rangle$.

The whole procedure of the VQDAC is depicted in Fig. 4.

During the implementation of the VQDAC algorithm, we invoke the variational quantum state eigensolver and the variational quantum linear solver as the basic operations to achieve the procedure of domain adaptation. When the original data are given in form of quantum states, the VQDAC can be implemented directly with quantum advantage compared with classical DA algorithms. In addition, the VQDAC is designed to be specifically efficient in dealing with problems in quantum mechanical scenarios such as the quantum many-body problem. However, compared with the classical DAC, the superiority of the VQDAC is an open question when we are initially given classical data.
4 Numerical experiments

In this section, two different numerical experiments are presented to demonstrated the efficiency of the VQDAC. All the experiments compared with performance of the no adaptation (NA) model, the classical DAC, the VQDAC, and two state-of-art quantum DA algorithms, namely the variational quantum correlation alignment algorithm (VQCORAL) [37] and the variational quantum subspace alignment algorithm (VQSA) [38]. The results of the experiments shows that the VQDAC algorithm can achieve comparable performance to the classical DAC which is better than the NA model. The VQDAC model is simulated on classical computer in PYTHON programming language with the Scikit-learn machine learning library [53] and the Paddle quantum framework [54].

4.1 Basic setting

The NA model is a baseline model which accomplishes the classification of the target domain data without any domain adaptation techniques. In this paper, the NA model is specifically the $k$-nearest neighbors algorithm [55]. The classical DAC model is selected as the comparison of the VQDAC. The VQDAC model is constructed by the parameterized quantum circuits as in Ref. [37] and the Adam optimization algorithm [56].

4.2 Small-scale data experiment

In this section, three synthetic data sets $D_1 \sim \mathcal{N}(\mu_1^{(1)} = \mu_2^{(1)} = 0, \sigma_1^{(1)} = \sigma_2^{(1)} = 1) \in \mathbb{R}^4$ with 100 samples evenly in two classes, $D_2 \sim \mathcal{N}(\mu_1^{(2)} = \mu_2^{(2)} = 0, \sigma_1^{(2)} = \sigma_2^{(2)} = 2) \in \mathbb{R}^4$ with 100 samples evenly in two classes, $D_3 \sim \mathcal{N}(\mu_1^{(3)} = \mu_2^{(3)} = 123$
Table 1 Accuracies of the NA, the classical DAC, the VQDAC, the VQCORAL, and the VQSA applied on the synthetic data sets $D_1$, $D_2$, $D_3$ and the Iris data set

|         | $D_1 \rightarrow D_2$ | $D_2 \rightarrow D_1$ | $D_3 \rightarrow$ Iris | Iris $\rightarrow D_3$ |
|---------|------------------------|------------------------|-------------------------|------------------------|
| NA      | 50.35%                 | 50.28%                 | 33.32%                  | 4.01%                  |
| Classical DAC | 90.08%              | 95.90%              | 68.24%                  | 70.28%                  |
| VQSA    | 89.04%                 | 90.24%                 | 55.66%                  | 60.90%                  |
| VQCORAL | 90.55%                 | 96.99%                 | 65.13%                  | 69.79%                  |
| VQDAC   | 90.90%                 | 97.87%                 | 66.73%                  | 70.87%                  |

Table 2 Accuracies of the NA, the classical DAC, the VQDAC, the VQCORAL, and the VQSA applied on the MNIST and USPS datasets

|         | MNIST $\rightarrow$ USPS | USPS $\rightarrow$ MNIST |
|---------|---------------------------|---------------------------|
| NA      | 64.40%                    | 35.91%                    |
| Classical DAC | 65.60%               | 44.59%                    |
| VQSA    | 64.90%                    | 36.04%                    |
| VQCORAL | 65.64%                    | 44.55%                    |
| VQDAC   | 65.60%                    | 44.60%                    |

$\mu_3^{(3)} = 0, \sigma_1^{(3)} = \sigma_2^{(3)} = \sigma_3^{(3)} = 1 \in \mathbb{R}^4$ with 150 samples evenly in three different classes, and the Iris data set are defined as the source and target domain alternatively.

The results of the numerical experiments are presented as Table 1. According to the results, it is obvious that the NA model can not achieve competitive performance. The classical DAC can deal with the two different DA scenarios effectively. Compared with the classical DAC and the state-of-art quantum DA algorithms, the VQDAC can achieve comparable performance which demonstrates its efficiency.

4.3 Large-scale data experiment

In this section, two handwritten digit datasets, the MNIST [57] and the USPS [58], are defined as the source and target domain data alternatively. Both of the two datasets contain 10 handwritten digit categories from 0 to 9. The MNIST dataset contains 60000 images as the training set and 10000 images as the test set with each image in scale of $28 \times 28$; the USPS dataset contains 7291 images as the training set and 2007 images as the test set with each image in scale of $16 \times 16$. The data points in these two datasets are distributed in the same feature space, but are generated from different distributions. In this experiment, 2000 images from the MNIST and 1800 images from the USPS are chosen. All of the data are rescaled to the size of $16 \times 16$ and each of them is represented by a 256-dimensional vector.

The results of the numerical experiments are presented as Table 2. In this experiment, although the results can not show the quantum advantage of the VQDAC compared with the classical DAC algorithm, the VQDAC achieves performance promotion in comparison with the NA model. This contrast experiment
demonstrates that the VQDAC is competitive with the state-of-art quantum DA algorithms.

5 Conclusion and future work

In this paper, two quantum implementations of the DAC are presented. The QBLAS-based DAC can be implemented on a universal quantum computer with logarithmic resources in the dimension and number of given data. The VQDAC can be performed on the near-term quantum devices through a variational hybrid quantum-classical procedure.

However, some open questions of the two quantum algorithms need further study. At first, the QBLAS-based DAC requires high-depth quantum circuits and fully coherent evolution in practice. Although it can be proved that the QBLAS-based DAC can achieve quantum speedup, the implementation requirement in practice is relatively hard at present. In addition, the optimal performance of the VQDAC still needs exploration. The specific design of the parameterized quantum circuits is vital to the accuracy of the variational algorithm. How to find the optimal circuit structure is another crucial open question. In spite of the open questions above, it is demonstrated that quantum techniques can be applied to the field of domain adaptation resulting in performance promotion.

Acknowledgements The author would like to thank Xiaoting Wang for constructive discussions. The author also would like to thank the referees for helpful comments on this paper. This work is supported by National Key Research and Development Program of China Grant No. 2018YFA0306703, in part by the National Natural Science Foundation of China under Grant 62271296, in part by Natural Science Basic Research Program of Shaanxi (No. 2022JQ-018), and in part by Shaanxi Joint Laboratory of Artificial Intelligence (No. 2020SS-03).

Declarations

Conflict of interest The authors declare that they have no conflict of interest. The datasets generated during and analysed during the current study are available from the corresponding author on reasonable request.

References

1. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Trans. Knowl. Data Eng. 22(10), 1345–1359 (2009)
2. Long, M., Cao, Y., Wang, J., Jordan, M.: Learning transferable features with deep adaptation networks. In: International Conference on Machine Learning, pp. 97–105 . PMLR (2015)
3. Mahajan, D., Girshick, R., Ramanathan, V., He, K., Paluri, M., Li, Y., Bharambe, A., Van Der Maaten, L.: Exploring the limits of weakly supervised pretraining. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 181–196 (2018)
4. Xie, M., Jean, N., Burke, M., Lobell, D., Ermon, S.: Transfer learning from deep features for remote sensing and poverty mapping. In: Thirtieth AAAI Conference on Artificial Intelligence (2016)
5. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)
6. Taylor, M.E., Stone, P.: Transfer learning for reinforcement learning domains: A survey. Journal of Machine Learning Research **10**(7) (2009)

7. Shor, P.W.: Algorithms for quantum computation: discrete logarithms and factoring. In: Proceedings 35th Annual Symposium on Foundations of Computer Science, pp. 124–134 (1994)

8. Grover, L.K.: A fast quantum mechanical algorithm for database search. In: Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing, pp. 212–219 (1996)

9. Harrow, A.W., Hassidim, A., Lloyd, S.: Quantum algorithm for linear systems of equations. Phys. Rev. Lett. **103**(15), 150502 (2009)

10. Aaronson, S., Arkhipov, A.: The computational complexity of linear optics. In: Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing, pp. 333–342 (2011)

11. Farhi, E., Neven, H.: Classification with quantum neural networks on near term processors. arXiv preprint arXiv:1802.06002 (2018)

12. Lloyd, S., Mohseni, M., Rebentrost, P.: Quantum algorithms for supervised and unsupervised machine learning. arXiv preprint arXiv:1307.0411 (2013)

13. Lloyd, S., Mohseni, M., Rebentrost, P.: Quantum principal component analysis. Nat. Phys. **10**(9), 631–633 (2014)

14. Rebentrost, P., Steffens, A., Marvian, I., Lloyd, S.: Quantum singular-value decomposition of nonsparse low-rank matrices. Phys. Rev. A **97**(1), 012327 (2018)

15. Rebentrost, P., Mohseni, M., Lloyd, S.: Quantum support vector machine for big data classification. Phys. Rev. Lett. **113**(13), 130503 (2014)

16. Wiebe, N., Kapoor, A., Svore, K.M.: Quantum nearest-neighbor algorithms for machine learning. Quantum Inf. Comput. **15**(3–4), 318–358 (2015)

17. Dang, Y., Jiang, N., Hu, H., Ji, Z., Zhang, W.: Image classification based on quantum k-nearest-neighbor algorithm. Quantum Inf. Process. **17**(9), 1–18 (2018)

18. Wiebe, N., Braun, D., Lloyd, S.: Quantum algorithm for data fitting. Phys. Rev. Lett. **109**(5), 050505 (2012)

19. Schuld, M., Sinayskiy, I., Petruccione, F.: Prediction by linear regression on a quantum computer. Phys. Rev. A **94**(2), 022342 (2016)

20. Aimeur, E., Brassard, G., Gambs, S.: Quantum speed-up for unsupervised learning. Mach. Learn. **90**(2), 261–287 (2013)

21. Cong, I., Duan, L.: Quantum discriminant analysis for dimensionality reduction and classification. New J. Phys. **18**(7), 073011 (2016)

22. He, X., Sun, L., Lyu, C., Wang, X.: Quantum locally linear embedding for nonlinear dimensionality reduction. Quantum Inf. Process. **19**(9), 1–21 (2020)

23. Wiebe, N., Kapoor, A., Svore, K.M.: Quantum deep learning. Quant. Inf. Comput. **16**(7–8), 541–587 (2016)

24. Amin, M.H., Andriyash, E., Rolfe, J., Kucelytskyy, B., Melko, R.: Quantum Boltzmann machine. Phys. Rev. X **8**(2), 021050 (2018)

25. Lloyd, S., Weedbrook, C.: Quantum generative adversarial learning. Phys. Rev. Lett. **121**(4), 040502 (2018)

26. Dallaire-Demers, P.-L., Killoran, N.: Quantum generative adversarial networks. Phys. Rev. A **98**(1), 012324 (2018)

27. Hu, L., Wu, S.-H., Cai, W., Ma, Y., Mu, X., Xu, Y., Wang, H., Song, Y., Deng, D.-L., Zou, C.-L.: Quantum generative adversarial learning in a superconducting quantum circuit. Sci. Adv. **5**(1), 2761 (2019)

28. Benedetti, M., Grant, E., Wossnig, L., Severini, S.: Adversarial quantum circuit learning for pure state approximation. New J. Phys. **21**(4), 043023 (2019)

29. Situ, H., He, Z., Wang, Y., Li, L., Zheng, S.: Quantum generative adversarial network for generating discrete distribution. Inf. Sci. **538**, 193–208 (2020)

30. Zeng, J., Wu, Y., Liu, J.-G., Wang, L., Hu, J.: Learning and inference on generative adversarial quantum circuits. Phys. Rev. A **99**(5), 052306 (2019)

31. Romero, J., Olson, J.P., Aspuru-Guzik, A.: Quantum autoencoders for efficient compression of quantum data. Quant. Sci. Technol. **2**(4), 045001 (2017)

32. Lamata, L., Alvarez-Rodriguez, U., Martin-Guerrero, J.D., Sanz, M., Solano, E.: Quantum autoencoders via quantum adders with genetic algorithms. Quant. Sci. Technol. **4**(1), 014007 (2018)

33. Khoshaman, A., Vinci, W., Denis, B., Andriyash, E., Sadeghi, H., Amin, M.H.: Quantum variational autoencoder. Quant. Sci. Technol. **4**(1), 014001 (2018)
34. Wan, K.H., Dahlsten, O., Kristjánsson, H., Gardner, R., Kim, M.: Quantum generalisation of feedforward neural networks. npj Quant. Inf. 3(1), 1–8 (2017)
35. Beer, K., Bondarenko, D., Farrelly, T., Osborne, T.J., Salzmann, R., Scheiermann, D., Wolf, R.: Training deep quantum neural networks. Nat. Commun. 11(1), 1–6 (2020)
36. Mari, A., Bromley, T.R., Izaac, J., Schuld, M., Killoran, N.: Transfer learning in hybrid classical-quantum neural networks. Quantum 4, 340 (2020)
37. He, X.: Quantum correlation alignment for unsupervised domain adaptation. Phys. Rev. A 102(3), 032410 (2020)
38. He, X.: Quantum subspace alignment for domain adaptation. Phys. Rev. A 102(6), 062403 (2020)
39. Sun, B., Saenko, K.: From virtual to reality: fast adaptation of virtual object detectors to real domains. BMVC 1, 3 (2014)
40. Fisher, R.A.: The use of multiple measurements in taxonomic problems. Ann. Eugen. 7(2), 179–188 (1936)
41. Schuld, M., Petruccione, F.: Supervised Learning with Quantum Computers. Springer, Berlin (2018)
42. Barenco, A., Ekert, A., Suominen, K.-A., Törmä, P.: Approximate quantum Fourier transform and decoherence. Phys. Rev. A 54(1), 139 (1996)
43. Zalka, C.: Fast versions of Shor’s quantum factoring algorithm. arXiv preprint arXiv:quant-ph/9806084 (1998)
44. Draper, T.G.: Addition on a quantum computer. arXiv preprint arXiv:quant-ph/0008033 (2000)
45. Giovannetti, V., Lloyd, S., Maccone, L.: Quantum random access memory. Phys. Rev. Lett. 100(16), 160501 (2008)
46. Nielsen, M.A., Chuang, I.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2010)
47. Duan, B., Yuan, J., Liu, Y., Li, D.: Quantum algorithm for support matrix machines. Phys. Rev. A 96(3), 032301 (2017)
48. Coppersmith, D.: An approximate Fourier transform useful in quantum factoring. IBM Research Report, 19642 (1994)
49. Buhrman, H., Cleve, R., Watrous, J., De Wolf, R.: Quantum fingerprinting. Phys. Rev. Lett. 87(16), 167902 (2001)
50. Sun, X., Tian, G., Yang, S., Yuan, P., Zhang, S.: Asymptotically optimal circuit depth for quantum state preparation and general unitary synthesis. arXiv preprint arXiv:2108.06150 (2021)
51. Cerezo, M., Sharma, K., Arrasmith, A., Coles, P.J.: Variational quantum state eigensolver. arXiv preprint arXiv:2004.01372 (2020)
52. Bravo-Prieto, C., LaRose, R., Cerezo, M., Subasi, Y., Cincio, L., Coles, P.J.: Variational quantum linear solver. arXiv preprint arXiv:1909.05820 (2019)
53. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830 (2011)
54. Paddle Quantum. https://github.com/PaddlePaddle/Quantum (2020)
55. Fix, E., Hodges, J.L.: Discriminatory analysis. Nonparametric discrimination: consistency properties. Int. Stat. Review/Revue Internationale de Statistique 57(3), 238–247 (1989)
56. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: ICLR (Poster). arXiv:1412.6980 (2015)
57. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE 86(11), 2278–2324 (1998)
58. LeCun, Y., Boser, B.E., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W.E., Jackel, L.D.: Handwritten digit recognition with a back-propagation network. In: Advances in Neural Information Processing Systems, pp. 396–404 (1990)

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.