arXiv:2603.02700v1 [quant-ph] 3 Mar 2026

用于单类分类的神经量子支持向量数据描述

ChangjaeIm[1],Hyeondo Oh[1], 和 Daniel K. Park[1,2,3,*]

[1] *Department of Statistics and Data Science, Yonsei University, Seoul, Republic of Korea*
[2] *Department of Applied Statistics, Yonsei University, Seoul, Republic of Korea*
[3] *Department of Quantum Information, Yonsei University, Seoul, Republic of Korea*

摘要

单类分类（OCC）是机器学习中的一项基础问题，在异常检测和质量控制等领域有着广泛应用。随着现代数据集日益复杂、维度不断增加，人们对表达能力更强、效率更高的先进单类分类技术的需求也日益增长。我们提出神经量子支持向量数据描述（NQSVDD），这是一种用于单类分类的经典–量子混合框架，可进行端到端优化的分层表示学习。NQSVDD 将经典神经网络与可训练的量子数据编码和变分量子电路相结合，使模型能够学习针对单类分类目标定制的非线性特征变换。该混合架构将输入数据映射到中间高维特征空间，随后通过量子测量将其投影到紧凑的潜在空间。重要的是，特征嵌入和潜在表示经过联合优化，使正常数据形成紧凑的簇，而最小体积包围超球体则为其提供有效的决策边界。在基准数据集上的实验评估表明，与经典 Deep SVDD 和量子基线方法相比，NQSVDD 的曲线下面积表现具有竞争力或更优，同时保持了较高的参数效率，并在现实噪声条件下表现出良好的鲁棒性。

1 引言

量子机器学习（QML）在数据分析和模式识别方面展现出令人瞩目的能力 [1–3]。在二元分类领域，量子计算技术有望在效率、表达能力和准确性方面超越经典算法 [4–7]。然而，单类分类（OCC） [8–14] 仍然尤具挑战性，因为训练数据仅由单一类别的样本构成。由于无法获取对照样本或负样本，单类分类模型必须精确刻画目标类别，才能检测出偏离，通常应用于半监督或无监督设置。尽管面临这些挑战，单类分类在现实世界中有广泛的应用，例如金融 [15–17], 网络安全 [18], 生物信息学 [19,20], 和粒子物理学 [21]。

已有多项研究提出将 QML 用于单类分类 [22–27], ，但许多方法在实际应用中仍面临局限。一些方法需要超出现有量子硬件能力的复杂子程序 [27], ，而另一些方法更注重数据压缩，而非有效的异常检测 [24–26]。为应对这些挑战，我们提出了神经量子支持向量数据描述（*NQSVDD*），这是一种经典—量子混合模型，旨在含噪声中尺度量子（NISQ）设备上对经典数据进行单类分类。NQSVDD 将经典神经网络与量子特征映射相结合，以便在早期阶段提取特征，并在高维希尔伯特空间中生成丰富且富有表现力的表示。此外，这种混合方法兼顾计算复杂度，减少了对更高复杂度量子处理的需求 [1]。量子特征映射由交替排列的嵌入层和参数化层构成：嵌入层将输入数据转换为量子态，参数化层则包含可学习参数，从而提升表示能力。随后，通过带有局部测量的变分量子电路（VQC），例如量子卷积神经网络（QCNN） [28], ，在潜在空间中识别一个包围训练数据的最小体积超球体，以此定义决策边界。尽管可以采用多种策略来学习参数，但原则上可以同时更新初始经典神经网络的参数、

* : 通讯作者 电子邮件地址 :
dkd.park@yonsei.ac.kr

量子特征映射中参数化层的参数，以及 VQC 的参数。这种混合方法联合学习量子神经网络和经典神经网络，有望在可用量子比特数量有限的 NISQ 设备上实现优于经典算法的单类分类性能。

MNIST上的实验结果 [29], Fashion-MNIST [30] 图像数据集、信用卡交易数据集 [31],和网络入侵（CIC-IDS2017）数据集 [32], 均使用 PennyLane 实现 [33], 结果表明，NQSVDD 的表现优于经典和量子基线方法，以更少的可训练参数实现了更高的准确率。MNIST 和 Fashion-MNIST 数据集各包含十个不同类别。我们将其中一个类别视为目标类别，其余类别视为离群点，因此每个数据集都对应十种独立的单类分类实验设置。仅使用相应目标类别的样本来优化超球体。在测试数据集上评估性能，该数据集包含数量均衡的目标类别样本和离群点样本，评估指标为受试者工作特征（ROC）曲线下面积（曲线下面积）。在大多数情况下，无论是在理想环境还是噪声环境中，NQSVDD 的表现都优于采用振幅编码的纯量子方法 QSVDD，以及经典 DSVDD。值得注意的是，NQSVDD 的分类性能优于 QSVDD，且其架构更适合 NISQ。 此外，尽管 NQSVDD 学习的参数数量少于 DSVDD，其表现仍优于 DSVDD。这些结果表明，NQSVDD 是一种有效且高效的单类分类算法，非常适用于 NISQ时代；它兼具经典方法的鲁棒性和量子特征表达能力，同时保持较浅的电路深度和较低的量子比特需求。

2 背景

2.1 单类分类

设 $\mathcal{X} \subseteq \mathbb{R}^d$ 为输入空间， $\mathcal{Y} = \{0, 1, ..., K-1\}$ 为 K 个类别的标签空间。给定带标签的训练数据集 $\mathcal{D}_K = \{(\boldsymbol{x}_i, y_i)\}_{i=1}^m \subset \mathcal{X} \times \mathcal{Y}$，分类的目标是学习一个函数 $f\colon \mathcal{X} \to \mathcal{Y}$，以准确预测未见过的输入 $\tilde{\boldsymbol{x}} \in \mathcal{X}$的类别标签。

单类分类（OCC）是分类的一种极端情况，训练期间仅包含来自单个目标类别的数据点。由于缺乏目标类别之外样本的先验知识，与一般的多类分类相比，这项任务更具挑战性。给定一个包含 $K = 1$、经缩减后为 $\mathcal{D}_1 = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_m\} \subset \mathcal{X}_{\text{target}} \subseteq \mathcal{X}$的训练数据集 $\mathcal{D}_K$，单类分类的目标是学习一个决策函数，以判断未见过的数据点是否属于目标类别的分布。具体而言，设 $f\colon \mathcal{X} \to \mathbb{R}$ 为一个决策函数，用于衡量输入 $\tilde{\boldsymbol{x}} \in \mathcal{X}$ 与目标类别表示之间的相似度。对于预定义的阈值 b，若 $f(\tilde{\boldsymbol{x}}; \mathcal{D}_1)) \leq b$，则将 $\tilde{\boldsymbol{x}}$ 分类为目标对象。反之，若$f(\tilde{\boldsymbol{x}}; \mathcal{D}_1)) > b$，则将 $\tilde{\boldsymbol{x}}$ 识别为离群点。

$$h(f(\tilde{\boldsymbol{x}}; \mathcal{D}_1)) = \begin{cases} 1 & \text{if } f(\tilde{\boldsymbol{x}}; \mathcal{D}_1) \leq b \\ 0 & \text{if } f(\tilde{\boldsymbol{x}}; \mathcal{D}_1) > b. \end{cases} \tag{1}$$

目标实例的紧凑特征表示，以及不同类别之间具有区分性的表示，是单类分类 [10]的两个理想属性。换言之，决策函数的紧凑性和描述性与单类分类的性能直接相关。

2.2 深度支持向量数据描述

支持向量数据描述（SVDD） [13] 是一种用于单类分类的基于核的方法，旨在通过在特征空间中构造一个包含大多数训练数据的最小体积超球体来描述训练数据集。该超球体为区分目标数据与离群点提供了边界。

设 $k\colon \mathcal{X} \times \mathcal{X} \to [0, \infty)$为一个正半定（PSD）核， $\mathcal{F}_k$为其对应的再生核希尔伯特空间（RKHS）， $\phi_k\colon \mathcal{X} \to \mathcal{F}_k$ 为其对应的特征映射，使得 $k(\boldsymbol{x}, \tilde{\boldsymbol{x}}) = \langle \phi_k(\boldsymbol{x}) \phi_k(\tilde{\boldsymbol{x}}) \rangle_{\mathcal{F}_k}$ 对所有 $\boldsymbol{x}, \tilde{\boldsymbol{x}} \in \mathcal{X}$均成立。于是，给定一个数据集 $\mathcal{D} = \{\boldsymbol{x}_i\}_{i=1}^m$，其中 $\boldsymbol{x}_i \in \mathcal{X}$，对于唯一确定一个超球体的中心 $\boldsymbol{c} \in \mathcal{F}_k$和半径 $R > 0$，SVDD的原始优化问题可表述为

$$\begin{aligned} \min_{R, \boldsymbol{c}, \boldsymbol{\xi}} \quad & R^2 + \frac{1}{\nu m} \sum_i \xi_i \\ \text{s.t.} \quad & \|\phi_k(\boldsymbol{x}_i) - \boldsymbol{c}\|_{\mathcal{F}_k}^2 \leq R^2 + \xi_i, \ \xi_i \geq 0, \end{aligned} \tag{2}$$

其中 $\xi_i \geq 0$ 是允许边界出现软性违例的松弛变量，超参数 $\nu \in (0, 1]$ 用于权衡超球体体积与惩罚项 ξ_i。

SVDD方法需要进行显式特征工程，其计算复杂度至少随样本数量呈二次增长，而且存储支持向量需要占用大量内存。这些局限促使人们提出了深度支持向量数据描述（DSVDD）[14], ，这是一种基于深度学习的SVDD方法。DSVDD不再显式选择核函数，而是通过训练神经网络将数据映射到潜在空间中体积最小的超球体，从而学习数据的紧凑特征表示。DSVDD用于寻找最优特征空间和超球体的损失函数定义如下:

$$\mathcal{L}_c(\mathcal{W}) = \frac{1}{n}\sum_{i=1}^{m}\|\varphi_c(\boldsymbol{x}_i;\mathcal{W}) - \boldsymbol{c}\|^2 + \frac{\lambda}{2}\sum_{\ell=1}^{L}\|\boldsymbol{W}^\ell\|_F^2. \tag{3}$$

此处，对于某个输入空间 $\mathcal{X} \subseteq \mathbb{R}^d$ 和输出空间 $\mathcal{F} \in \mathbb{R}^{d'}$，其中 $d' < d$，$\varphi_c(;\mathcal{W})$: $\mathcal{X} \to \mathcal{F}$ 是一个具有 $L \in \mathbb{N}$ 个隐藏层和一组可训练参数 $\mathcal{W} = \{\boldsymbol{W}^1, ..., \boldsymbol{W}^L\}$ 的神经网络，其中 $\boldsymbol{W}^\ell$ 是第 $\ell \in \{1, ..., L\}$ 层的参数。第一项惩罚所有网络表示 $\varphi_c(\boldsymbol{x}_i; \mathcal{W})$ 到中心 $\boldsymbol{c} \in \mathcal{F}$ 的平均距离。为避免导致超球体塌缩的全零权重平凡解，预定义中心 $\boldsymbol{c}$ 被固定为非零值。此外，网络应排除偏置项，并使用 ReLU 等无界激活函数。第二项是针对网络参数 $\mathcal{W}$ 的权重衰减正则项，其超参数为 $\lambda > 0$，其中 $\|\cdot\|_F$ 表示弗罗贝尼乌斯范数。给定一个测试数据点 $\tilde{\boldsymbol{x}} \in \mathcal{X}$，决策函数表示为

$$f(\tilde{\boldsymbol{x}}) = \|\varphi_c(\tilde{\boldsymbol{x}}; \mathcal{W}^*) - \boldsymbol{c}\|^2 - R^{2*}, \tag{4}$$

whe其中 $\mathcal{W}^*$ 是训练后的参数， $R^{2*} = \max_i \|\varphi_c(\boldsymbol{x}_i; \mathcal{W}^*) - \boldsymbol{c}\|^2$ 是最终半径 s.

2.3 量子支持向量数据描述

受引入经典神经网络的 DSVDD 启发，量子支持向量数据描述（QSVDD）[27] 采用浅层变分量子机器学习算法，解决了以往方法不适用于含噪声中等规模量子设备，或并非直接为异常检测设计的局限性。为了在量子计算机上处理经典数据，量子数据编码 Ψ： $\mathcal{X} \to \mathcal{H}$ 将输入数据 $\boldsymbol{x} \in \mathcal{X}$ 映射 $|\psi$至希尔伯特空间 $\mathcal{H}$ 中的量子态$\rangle \in \mathbb{C}^{2^n}$ $(\boldsymbol{x})$，该空间的维度会随量子比特数量 n呈指数增长。具体而言，可以将酉变换 $U_\psi(\boldsymbol{x})$ 应用于初始态 $|0\rangle^{\otimes n}$，表示为 $\boldsymbol{x} \in \mathcal{X} \mapsto |\psi(\boldsymbol{x})\rangle = U_\psi(\boldsymbol{x})|0\rangle^{\otimes n} \in \mathcal{H}$。尽管量子特征映射有多种类型，QSVDD 专注于振幅编码，即将 N维向量 $\boldsymbol{x} = (x_1, \ ...x_N)^\top$ 编码为量子态

$$|\psi(\boldsymbol{x})\rangle = \frac{1}{\|\boldsymbol{x}\|}\sum_{i=1}^{N} x_i|i\rangle, \tag{5}$$

其中 $N = 2^n$，且 $|i\rangle$ 是第 i个计算基态。

在制备编码量子态之后，由 $\boldsymbol{\theta}$参数化的变分量子电路（VQC） $U(\boldsymbol{\theta})$在量子希尔伯特空间中学习数据表示。QSVDD选用量子卷积神经网络（QCNN），对量子卷积和池化层进行迭代，直至剩余系统的活跃量子比特数量足够少。最终的 n_f量子比特量子态由 $4^{n_f} - 1$ 个参数完全表征，随后通过量子测量期望函数 $g : \mathbb{C}^{2^n} \to \mathbb{R}^{d'}$映射到最终潜在空间。潜在空间 d' 的维度由从 n_f量子比特泡利可观测量集合中选取的泡利可观测量 P_i 决定。QSVDD的特征映射表示为

$$\varphi_q(\boldsymbol{x}, \boldsymbol{\theta}) = g(U(\boldsymbol{\theta})|\psi(\boldsymbol{x})\rangle). \tag{6}$$

给定训练目标数据样本 $\{\boldsymbol{x}_i\}_{i=1}^m$，通过最小化以下损失函数来优化变分量子电路的可训练参数:

$$\mathcal{L}_q(\boldsymbol{\theta}) = \frac{1}{m}\sum_{i=1}^{m}\|\varphi_q(\boldsymbol{x}_i, \boldsymbol{\theta}) - \boldsymbol{c}\|^2, \tag{7}$$

其中 $\boldsymbol{c}$ 是预先设定的中心。与式3不同，QSVDD 的损失函数不包含正则化项，因为量子电路的角度参数自然处于 0 到 2π之间。与式 4类似，在训练完特征映射后，通过决策函数判断测试数据 $\tilde{\boldsymbol{x}} \in \mathcal{X}$ 是否为异常值。

$$f(\tilde{\boldsymbol{x}}) = \|\varphi_q(\tilde{\boldsymbol{x}}; \boldsymbol{\theta}^*) - \boldsymbol{c}\|^2 - R^{2*} \tag{8}$$

其中 $\boldsymbol{\theta}^*$ 为训练得到的参数，半径 $R^{2*} = \max_i \|\varphi_q(\boldsymbol{x}_i; \boldsymbol{\theta}^*) - \boldsymbol{c}\|^2$。实证研究表明，QSVDD 效果显著，优于经典的 [14] 和量子的 [26, 34] 单类分类方法。

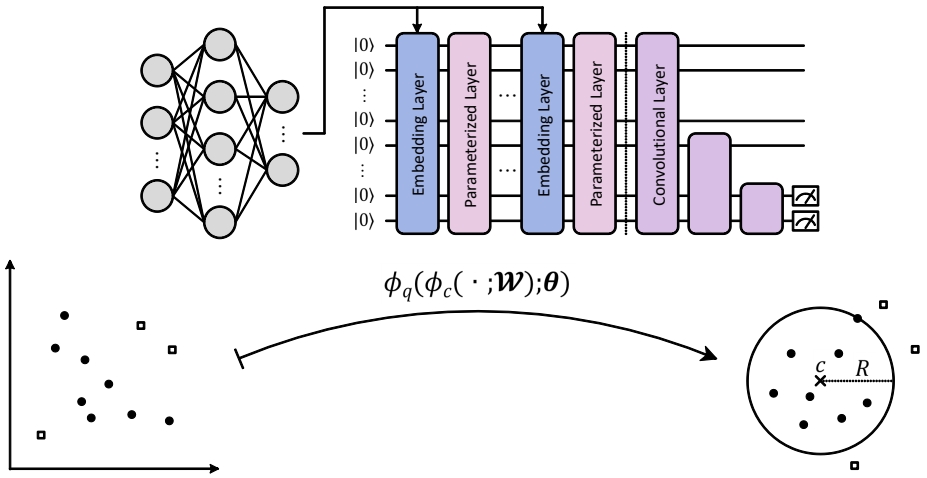


图 1：NQSVDD 概述。经典神经网络层从复杂输入数据中提取特征。经典层的输出通过 ZZ 特征嵌入编码为量子态。嵌入层与参数化量子层的交替重复使嵌入的表达能力更加丰富。态制备之后，变分量子电路和一组泡利测量将特征映射到潜在空间。经典层（记为 $\phi_c(\cdot;\mathcal{W})$）与量子层（记为 $\phi_q(\cdot;\boldsymbol{\theta})$）的组合构成用于单类分类的整体特征映射，并经训练以找到最优潜在空间和超球体。

3 模型描述

3.1 概述

尽管已有许多基于 QML 的单类分类方法被提出，但其中一些方法更注重使用量子自编码器结构压缩数据表示，而非有效地进行异常检测 [24–26]。与数据压缩不同，受 DSVDD 启发的 QSVDD 使用目标函数构造一个紧密包围目标数据的超球体。然而，仅依赖量子子程序，使 QSVDD 从可扩展性和实用性的角度来看，不适合在当前量子硬件上处理现代复杂数据集。

神经量子支持向量数据描述（NQSVDD）通过将量子电路的表示能力与经典神经网络的鲁棒性和可扩展性相结合，提供了一种实用方法 [1]。图1 给出了 NQSVDD 架构的示意图概述。模型的关键组成部分将在以下各小节中详细描述。

3.2 混合方法

在当前量子硬件的限制下，经典-量子混合方法具有多项优势。浅层量子神经网络（QNN）能够学习量子关联分布，从而增强经典模型的表征能力。同时，由经典神经网络负责特征学习，可减少对深层、高度纠缠的试探电路以及大量量子比特的需求 [1]。因此，这种混合方法尤其适用于量子资源有限的含噪声中等规模量子设备。

设计结合量子电路与经典模型的混合模型有多种方式，应根据具体任务和数据的性质选择集成结构。在二元分类中，研究证明，经典到量子的混合结构能够提高训练准确率、泛化能力以及对噪声的鲁棒性 [35, 36]。将这一思路扩展到单类分类，NQSVDD采用经典到量子的混合结构来处理高维经典数据。具体而言，经典神经网络首先对输入数据进行压缩或预处理，从而降低量子神经网络的计算负担。经典架构的选择依据与数据结构相契合的归纳偏置，例如，卷积神经网络（CNNs）的空间局部性适用于图像数据处理，而循环神经网络（RNNs）的时间因果性适用于序列数据处理。随后，通过某种数据嵌入方法，将经典神经网络的输出传递至量子电路。设计经典层和量子层的具体结构时，会考虑输入数据的复杂性、量子硬件的限制以及潜在空间的维度。

3.3 数据嵌入

量子数据嵌入对于使用量子计算机处理经典数据至关重要。基编码、振幅编码和角度编码是典型的量子嵌入方法。计算基编码将经典比特直接映射到量子基态。它简单高效，但不适用于连续数据，且表达能力不足。振幅编码将归一化数据编码到量子态的振幅中，具有紧凑且表达能力强的特点，但态制备效率低，需要复杂子程序。角度编码直接将数据用作单量子比特门的旋转角。它的电路深度较浅，但与数据呈线性关系。

我们的方法兼容任何量子数据嵌入，但我们采用 ZZ 特征嵌入 [37] 用于演示，其定义如下：

$$
\begin{aligned}
|\psi(\boldsymbol{x})\rangle &= U_{\mathrm{ZZ}}(\boldsymbol{x})|0\rangle^{\otimes n} \\
&:= \left[\left(\prod_{(j,k)\in\mathcal{P}} e^{i x_{jk} Z_j Z_k} \right) \left(\prod_{j=0}^{n-1} e^{i x_j Z_j} \right) \left(\bigotimes_{j=0}^{n-1} H_j \right) \right]^l |0\rangle^{\otimes n},
\end{aligned} \tag{9}
$$

其中， $\mathcal{P}$ 表示用于纠缠相互作用的量子比特对集合。为简单起见，假设 l 为偶数，则l层 ZZ 特征嵌入可被解释为对哈密顿量 $H_{\mathrm{XZ}}(\boldsymbol{x})$ [38]进行的一阶 Lie-Trotter模拟：

$$
\begin{aligned}
&U_{\mathrm{ZZ}}(\boldsymbol{x}) \approx \left[e^{i(l/2) H_{\mathrm{XZ}}(\boldsymbol{x})} \right], \\
&\text{where} \quad H_{\mathrm{XZ}}(\boldsymbol{x}) = \sum_{j=0}^{n-1} x_j (X_j + Z_j) + \sum_{(j,k)\in\mathcal{P}} x_{jk} (X_j X_k + Z_j Z_k).
\end{aligned} \tag{10}
$$

在这一哈密顿量视角下，输入特征会直接映射到局域场系数 x_j 和相互作用强度系数 x_{jk}。对于一个 n量子比特系统，考虑所有可能的相互作用对，一个 l层 ZZ 特征嵌入最多可编码 $l(n+n(n-1)/2) = ln(n+1)/2$ 个特征。单量子比特门 $e^{i x_j Z_j} = \cos(x_j) I + i \sin(x_j) Z_j$和双量子比特门 $e^{i x_{jk} Z_j Z_k} = \cos(x_{jk}) I + i \sin(x_{jk}) Z_j Z_k$ 为嵌入带来非线性。输入数据的这种非线性依赖关系以及双量子比特相互作用产生的纠缠，使 ZZ 特征嵌入具有足够的表达能力，能够捕捉复杂的数据模式。此外，ZZ 特征嵌入与瞬时量子多项式（IQP）嵌入的相关族有关，后者能够生成高度纠缠的量子态，并且已知其在经典计算上难以模拟 [39–41]。通过 ZZ 特征嵌入将经典数据映射到指数级扩大的希尔伯特空间中，NQSVDD 能够捕捉在经典计算中难以构造的复杂特征表示，从而有望超越传统算法。

此外，NQSVDD 在 ZZ 特征嵌入层 $U_{\mathrm{ZZ}}(\boldsymbol{x})$和参数化酉层 $U(\theta_i)$之间交替。

$$
|\psi(\boldsymbol{x})\rangle = [U(\theta_t) \cdot U_{\mathrm{ZZ}}(\boldsymbol{x}) \cdots U(\theta_1) \cdot U_{\mathrm{ZZ}}(\boldsymbol{x})] |0\rangle^{\otimes n}. \tag{11}
$$

反复将经典输入数据编码到量子态中，并在其间插入可训练的酉操作，能够在保持电路深度较浅的同时进一步增强表达能力 [42]。

3.4 变分量子电路

准备好量子态后，带有参数 $\boldsymbol{\theta} = \{\theta_1, ..., \theta_T\}$的变分量子电路（VQC） $U(\boldsymbol{\theta})$能够在指数级庞大的希尔伯特空间中学习嵌入特征的有意义表示。VQC 的架构会显著影响其可训练性、归纳偏置和最终性能 [43–46]。精心设计的架构能够保留有用的梯度，并充分表示解空间，从而获得良好的可训练性和泛化性能。另一方面，设计不当则可能导致贫瘠高原，并对噪声过拟合。

Deep SVDD 的最终目标是学习一种映射，将输入嵌入紧凑的超球体内。实际上，该超球体定义在低维潜在空间中，以提升单类分类的泛化性能；这一做法在保留目标类别分布相较于其他类别的显著特征的同时，隐式地去除了冗余变化。这一目标自然引出了采用局部测量的分层式变分量子分类器架构，该架构能在渐进式压缩过程中保留重要信息 [47]。量子卷积神经网络（QCNN）是一种典型的分层式变分量子分类器，由两种类型的层构成：用于数据处理、施加局部酉门的卷积层，以及通过测量或舍弃特定量子比特来减少活跃量子比特数量的池化层。QCNN 分层结构中系统规模的这种反复缩减，能够以 $O(\log n)$ 的深度，将含有 n 个量子比特的输入高效压缩至固定数量的量子比特。压缩之后，我们可以选择

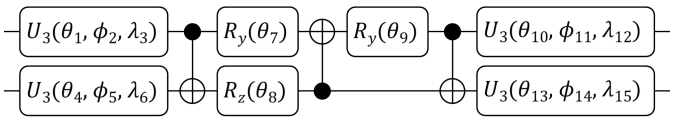


图 2：能够生成$SU(4)$中任意双量子比特酉变换的参数化双量子比特门，需要 15 个实参数。$R_i(\theta)$ 是绕布洛赫球 i 轴的旋转门，而 $U_3(\theta, \phi, \lambda) = R_z(\phi)R_x(-\pi/2)R_z(\theta)R_x(\pi/2)R_z(\lambda)$。该双量子比特拟设用于数据嵌入中的参数化酉层以及量子卷积神经网络中的卷积层。

低权重（k-局部）泡利可观测量，其非恒等泡利因子的数量很少，因此可通过经典阴影层析高效估计期望值 [48]。可训练性和良好的泛化能力也是采用 QCNN 架构的动因。QCNN 的分层式局部连接架构避免了贫瘠高原问题，从而确保可训练性和效率 [49]。此外，理论分析和实证实验表明，该架构具有强大的泛化能力和出色的二元分类性能 [50–52]。

对于用于数据嵌入的可训练酉层和量子卷积神经网络的卷积层的参数化双量子比特门，我们采用 $SU(4)$ 拟设，该拟设需要 15 个实参数，并且能够生成任意双量子比特酉操作（见图 2）。每一层中的所有双量子比特门共享相同的参数，这提高了泛化性能并减少了可训练参数的数量 [53]。由于量子卷积神经网络层的深度随量子比特数量呈对数增长，因此量子卷积神经网络可由 $O(\log n)$ 个参数表征。

最终量子态 $|\psi_f\rangle = U(\boldsymbol{\theta})|\psi(x)\rangle$ 通过局部测量投影到潜在空间 $\mathcal{F}$中，所选可观测量的期望值则定义了输出特征。具体而言，最终量子态存在于 n_f 个活跃量子比特上，可由密度矩阵完整表征。要完整描述该密度矩阵，需要 $4^{n_f} - 1$个实参数。这些特征通过测量所选泡利可观测量子集 $P \subseteq \mathcal{P}_{n_f}$的期望值提取，其中 $\mathcal{P}_{n_f} = \{I, \sigma_x, \sigma_y, \sigma_z\}^{\otimes n_f} \setminus \{I^{\otimes n_f}\}$。第 i个输出（潜在）特征计算为期望值 $\langle P_i\rangle$，其中 $P_i \in P \subseteq \mathcal{P}_{n_f}$ 是选定的泡利可观测量。

$$\begin{aligned}\langle P_i\rangle &= \langle\psi_f|P_i|\psi_f\rangle = \mathrm{Tr}(P_i|\psi_f\rangle\langle\psi_f|) \\ &= \mathrm{Tr}(P_i U(\boldsymbol{\theta})|\psi(x)\rangle\langle\psi(x)|U^\dagger(\boldsymbol{\theta})) = \mathrm{Tr}(U^\dagger(\boldsymbol{\theta})P_i U(\boldsymbol{\theta})|\psi(x)\rangle\langle\psi(x)|).\end{aligned} \tag{12}$$

这表明，优化变分量子分类器可以解释为学习最优的参数化测量。由于最终状态的密度矩阵可由 $4^{n_f} - 1$ 个实参数完整表征，因此输出特征的最大可能数量为 $4^{n_f} - 1$。为了迫使模型捕捉目标数据分布中最重要的特征，而非仅仅记忆每个样本的噪声模式，潜在特征空间的维度 d'由所选泡利可观测量集合 P 中的元素数量决定，且应满足 $d' = |P| \leq \min\{d, 4^{n_f} - 1\}$。此外，如果所选集合 P由低权重泡利可观测量组成，经典阴影层析就能高效估计其期望值 [48]。此外，对逐量子比特对易的泡利项进行分组，还能进一步提高多重泡利测量的效率 [54]。

3.5 损失函数

NQSVDD 的损失函数由公式 3 和公式 7 组合而成。设 $\boldsymbol{z} = \phi_c(\boldsymbol{x}; \mathrm{W})$ 为权重为 W、跨越 L 层的经典神经网络的输出，$\phi_q(\boldsymbol{z}; \boldsymbol{\theta}) = g(U(\boldsymbol{\theta})|\psi(\boldsymbol{z})\rangle)$ 为参数为 $\boldsymbol{\theta} = \{\theta_1, ..., \theta_T\}$ 的量子模型的输出，其中 $|\psi(\boldsymbol{z})\rangle$ 是编码量子态，$g: \mathbb{C}^{2^n} \to \mathcal{F}$ 是一个将量子态映射到潜在空间的函数。整体特征映射表示为 $\phi(\boldsymbol{x}; \mathrm{W}\ \boldsymbol{\theta}) = \phi_q(\phi_c(\boldsymbol{x}; \mathrm{W}); \boldsymbol{\theta})$。给定样本 $\{\boldsymbol{x}_i\}_{i=1}^m$，模型通过最小化以下目标函数，将输入映射到靠近参考中心 $\boldsymbol{c} \in \mathcal{F}$ 的位置：

$$\mathcal{L}(\boldsymbol{\mathcal{W}}, \boldsymbol{\theta}) = \frac{1}{m}\sum_{i=1}^{m} \|\phi(\boldsymbol{x}_i; \boldsymbol{\mathcal{W}}, \boldsymbol{\theta}) - \boldsymbol{c}\|^2 + \frac{\lambda}{2}\sum_{\ell=1}^{L} \left\|\boldsymbol{W}^\ell\right\|_F^2. \tag{13}$$

第一项通过最小化嵌入输出与中心之间欧几里得距离平方在 m 个样本上的平均值，促使其围绕中心 $\boldsymbol{c}$ 紧凑分布。第二项是带有超参数 $\lambda > 0$ 的弗罗贝尼乌斯范数正则化。由于量子参数 $\boldsymbol{\theta} \in [0, 2\pi]$ 本身有界，因此正则化仅应用于 L 层经典神经网络的经典参数。

3.6 基于梯度的混合模型优化

在基于梯度的优化过程中，经典神经网络和量子神经网络的参数会联合更新。使用参数移位规则，计算参数化量子电路中期望值相对于量子参数的解析梯度 [55]。设 $|\psi_{\boldsymbol{\theta}}\rangle = U(\boldsymbol{\theta})|\psi\rangle$ 是一个参数化量子态，其中 $U(\boldsymbol{\theta}) = e^{-i\boldsymbol{\theta}G}$ 满足 $G^2 = I$ 且 $\boldsymbol{\theta} = \{\theta_1, ..., \theta_T\}$，并且 $|\psi\rangle$ 是一个初始量子态。那么，对于一个厄米观测量 O，其期望值为 $\langle O\rangle_{\boldsymbol{\theta}} = \langle\psi_{\boldsymbol{\theta}}|O|\psi_{\boldsymbol{\theta}}\rangle$。参数移位规则给出了 $\langle O\rangle_{\boldsymbol{\theta}}$相对于 θ_i的精确梯度：

$$\frac{\partial\langle\hat{O}\rangle_{\boldsymbol{\theta}}}{\partial\theta_i} = \frac{1}{2}\left[\langle\hat{O}\rangle_{\boldsymbol{\theta}_i^+} - \langle\hat{O}\rangle_{\boldsymbol{\theta}_i^-}\right], \tag{14}$$

其中 $\boldsymbol{\theta}_i^{\pm} = (\theta_1, \ldots\ldots, \theta_i \pm \frac{\pi}{2}, \ldots\ldots, \theta_T)$ 且 $\langle\hat{O}\rangle_{\boldsymbol{\theta}_i^{\pm}} = \langle\psi_{\boldsymbol{\theta}_i^{\pm}}|\hat{O}|\psi_{\boldsymbol{\theta}_i^{\pm}}\rangle$。对于经典神经网络，通过在反向传播过程中递归应用链式法则，将损失导数从输出层反向传播至各层，从而计算梯度。混合模型的梯度可通过结合量子参数的参数移位规则与经典参数的反向传播，并借助链式法则来计算。考虑一个具有经典到量子结构的混合模型，其中经典神经网络的输出为 $\boldsymbol{z}$，其经典参数为 $\boldsymbol{w}$，量子参数为 $\boldsymbol{\theta}$。对于某个损失 $\mathcal{L}$，相对于 θ_i 和 w_j 的梯度通过链式法则计算，表示为：

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial\theta_i} &= \frac{\partial\mathcal{L}}{\partial\langle\hat{O}\rangle_{\boldsymbol{\theta}}} \cdot \frac{\partial\langle\hat{O}\rangle_{\boldsymbol{\theta}}}{\partial\theta_i}, \\
\frac{\partial\mathcal{L}}{\partial w_j} &= \frac{\partial\mathcal{L}}{\partial\langle\hat{O}\rangle_{\boldsymbol{\theta}}} \cdot \sum_k \left(\frac{\partial\langle\hat{O}\rangle_{\boldsymbol{\theta}}}{\partial z_k} \cdot \frac{\partial z_k}{\partial w_j}\right),
\end{aligned} \tag{15}$$

其中 $\partial\langle\hat{O}\rangle_{\boldsymbol{\theta}}/\partial\theta_i$ 和 $\partial\langle\hat{O}\rangle_{\boldsymbol{\theta}}/\partial z_k$ 是通过应用参数移位规则得到的。通过计算 $\partial\mathcal{L}/\partial\theta_i$ 和$\partial\mathcal{L}/\partial w_j$，即可联合优化所有经典参数和量子参数。

4 数值实验

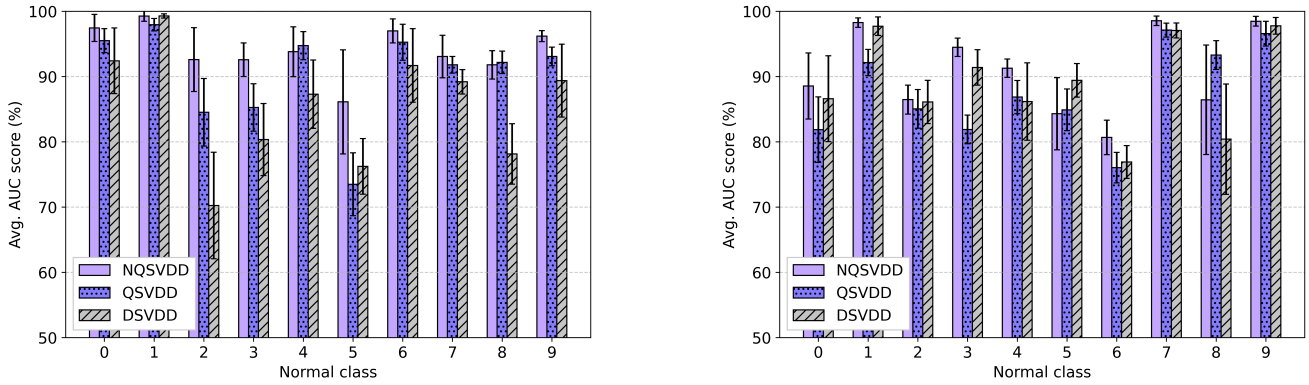
实验在 MNIST [29], Fashion-MNIST [30], 信用卡交易 [31] 数据集以及网络入侵（CIC-IDS2017） [32]数据集上进行。我们首先在理想的无噪声设定下，比较 NQSVDD 与基线方法在各数据集上的单类分类性能。此外，我们还评估其在噪声环境中的单类分类任务效果。以下各节将详细介绍这些数据集和实验流程。

4.1 MNIST 和 Fashion-MNIST 上的实验

MNIST 和 Fashion-MNIST 数据集分别包含 28×28 像素的手写数字和服饰灰度图像，且均有 10 个不同类别。将每个类别依次作为目标类别，由此形成 10 个独立的单类分类任务，每个任务对应一个不同的目标类别。NQSVDD 使用 1000 个目标样本进行训练，并使用 100 个目标测试样本和 90 个离群点测试样本（其余每个类别各取 10 个样本），通过 AUC 分数进行评估。所有图像均经过归一化预处理，并通过最小-最大缩放重新缩放至 $[0, 1]$ 。

NQSVDD 的经典层由两个交替排列的卷积层和池化层组成。池化层的降维程度取决于通过 ZZ 特征嵌入编码到量子态中的特征维度：八量子比特态包含 36 维特征。此外，NQSVDD 通过交替使用 ZZ 特征嵌入层和参数化酉层三次，将经典层展平后的输出映射到量子态。增加嵌入层数量的结果见附录 A.2 的图 A。态制备完成后，两个作用于八个量子比特的参数化层和两个作用于四个量子比特的层构成量子卷积神经网络，之后进行一组泡利测量。如上所述，量子层中的所有 SU（4）个 ansatz 共享相同的参数。此外，应谨慎选择潜在空间的维度，因为它会直接影响模型的性能，如附录 A.1a 和 A.1b 中的图 A所示。为了与将潜在维度设为 32 的传统 DSVDD 相竞争，我们在 $\mathcal{P}_4 = \{I, \sigma_x, \sigma_y, \sigma_z\}^{\otimes4} \setminus \{I^{\otimes4}\}$中选取了一组 32 个泡利可观测量。

我们将所提出的 NQSVDD 与 QSVDD 和经典DSVDD进行比较。两种 QSVDD 变体——采用振幅编码的 QSVDD 和采用 ZZ 特征嵌入的 QSVDD——具有相同的变分量子电路结构，均包含 75 个可训练参数；而输入数据则根据编码类型进行调整尺寸并展平。对于 DSVDD，我们考虑了 [14] 中的原始架构，以及经过简化的架构；后者经过修改，使其卷积层和池化



(a) MNIST (b) Fashion-MNIST

图 3：在 (a) MNIST 和 (b) Fashion-MNIST 数据集上的平均AUC（%）和标准差（误差条）。NQSVDD、QSVDD 和 DSVDD 的参数数量分别为 1105、75 和 2152。每个柱形表示 10 次重复实验的平均分数。

层与 NQSVDD 的对应层相匹配。这两个 DSVDD 模型的参数量分别为 7272 和 2152，约为 NQSVDD（参数量为 1105）的 7 倍和 2 倍。NQSVDD 和 QSVDD 使用 8 个输入量子比特，并通过 PennyLane [33]实现。参数集 $\mathcal{W}$ 和 $\boldsymbol{\theta}$ 通过 Adam 优化器 [56] 以 0.05 的学习率同时更新。我们还采用带热重启的余弦退火 [57], ，将学习率衰减至最低 0.005，并定期重置学习率，以提升收敛效果并逃离局部极小值。我们使用批次大小为 32 训练了 1500 步。图 3a、3b和表 1 表明，NQSVDD 在大多数任务中的表现优于 QSVDD 和 DSVDD。值得注意的是，NQSVDD 使用的参数远少于 DSVDD，但其性能始终与 DSVDD 相当或更优。

4.2 信用卡交易实验

信用卡交易数据集包含欧洲持卡人在 2013 年 9 月两天内进行的交易，其中 284807 笔交易中有 492 笔欺诈交易 [31]。由于保密问题，未提供原始特征，而是提供了经过主成分分析变换得到的 28 维数值特征。与 MNIST 和 Fashion-MNIST 类似，随机选取 1000 个目标样本用于训练，并使用 492 个目标样本和 492 个离群点样本计算 AUC 分数。

鉴于信用卡交易数据样本的复杂度低于前述图像数据集，我们相应地缩小了模型架构。对于 NQSVDD，模型的经典骨干网络由两个交替使用一维卷积和池化层的模块组成，可将 28 维输入转换为紧凑的 21 维特征向量。通过重复三个由 ZZ 特征嵌入和参数化层组成的模块，该特征向量被编码到一个六量子比特寄存器中，再由量子卷积神经网络逐级压缩为两个活跃量子比特。最后，通过局部泡利测量得到八维潜在表示。QSVDD 和 DSVDD 的架构也相应缩小，以生成八维潜在向量；其中 DSVDD 与 NQSVDD 使用相同的卷积骨干网络。NQSVDD、QSVDD 和 DSVDD 的可训练参数总数分别为 210、75 和 288。实验的其他设置与 MNIST 和 Fashion-MNIST 上的设置相同。表 2显示，NQSVDD 的表现优于 QSVDD 和 DSVDD。

4.3 网络入侵实验

CIC-IDS2017 数据集是加拿大网络安全研究所（CIC）于2017年发布的网络入侵基准，用于评估入侵检测系统（IDS） [32]。该数据集采集了数天混合的良性与恶意流量，并提取了78个网络流特征。在本实验中，我们选取了2017年7月6日，星期四上午的数据集，其中包含四个类别：“良性”、“Web攻击－暴力破解”、“Web攻击－SQL注入”、“Web攻击－跨站脚本攻击”，样本数分别为168166、1507、652和21。由于我们的目标是检测是否发生了Web攻击，这三个Web攻击类别被合并为单一的“Web攻击”标签。训练时随机抽取了1000个“良性”样本，并使用由2180个“良性”样本和2180个“Web攻击”样本组成的平衡测试集计算AUC分数。

表 1：MNIST 和 Fashion-MNIST 数据集上的单类实验中，每种方法使用 10 个随机种子所得的平均 AUC 分数（%）及标准差。粗体表示各类别中的最佳结果。

目标类别	方法				
	NQSVDD	QSVDD(amp)	QSVDD(zz)	DSVDD	DSVDD（降维）
0	**97.45** ± 2.07	95.50 ± 1.84	88.70 ± 3.49	94.49 ± 3.80	92.42 ± 5.03
1	99.27 ± 0.79	97.96 ± 0.95	97.36 ± 1.36	**99.54** ± 0.40	99.29 ± 0.32
2	**92.60** ± 4.89	84.53 ± 5.17	82.70 ± 2.87	77.48 ± 10.84	70.25 ± 8.16
3	**92.59** ± 2.58	85.27 ± 3.64	83.71 ± 4.48	80.28 ± 8.60	80.35 ± 5.52
4	93.81 ± 3.83	**94.74** ± 2.15	88.32 ± 4.01	90.54 ± 2.88	87.29 ± 5.24
5	**86.13** ± 7.98	73.50 ± 4.82	78.51 ± 4.71	76.42 ± 10.33	76.23 ± 4.27
6	**96.99** ± 1.84	95.26 ± 2.76	85.27 ± 2.36	94.90 ± 2.13	91.70 ± 5.64
7	**93.08** ± 3.26	91.80 ± 1.27	83.63 ± 4.18	91.88 ± 3.42	78.15 ± 4.63
8	91.80 ± 2.18	**92.18** ± 1.72	82.61 ± 3.41	82.64 ± 3.65	78.15 ± 4.63
9	**96.20** ± 0.85	93.08 ± 1.43	87.38 ± 4.95	92.45 ± 4.19	89.39 ± 5.58
平均总计	**93.99** ± 4.98	90.38 ± 7.53	85.82 ± 6.02	88.06 ± 9.69	85.43 ± 9.73
T恤/上衣	**88.56** ± 5.07	81.87 ± 5.01	85.24 ± 3.47	87.60 ± 6.53	86.60 ± 6.58
裤子	**98.27** ± 0.72	92.15 ± 2.00	96.14 ± 1.68	97.38 ± 1.53	97.72 ± 1.41
套头衫	86.47 ± 2.21	85.05 ± 2.96	**87.87** ± 1.28	87.59 ± 1.43	86.11 ± 3.31
连衣裙	**94.48** ± 1.41	81.90 ± 2.21	89.90 ± 3.65	93.24 ± 1.62	91.40 ± 2.72
Coat	**91.29** ± 1.40	86.87 ± 2.54	83.91 ± 1.74	89.22 ± 3.66	86.18 ± 5.93
凉鞋	84.31 ± 5.53	84.90 ± 3.19	**91.01** ± 1.97	88.42 ± 5.67	89.42 ± 2.58
衬衫	**80.68** ± 2.63	76.05 ± 2.33	80.52 ± 3.49	75.50 ± 4.48	76.91 ± 2.51
运动鞋	**98.55** ± 0.73	97.12 ± 1.07	97.42 ± 1.06	97.30 ± 1.81	97.06 ± 1.15
Bag	86.44 ± 8.39	**93.30** ± 2.20	78.44 ± 5.22	85.08 ± 3.98	80.41 ± 8.45
踝靴	**98.48** ± 0.75	96.58 ± 1.89	92.94 ± 2.31	97.03 ± 2.47	97.77 ± 1.28
平均总计	**90.75** ± 7.15	87.58 ± 7.14	88.34 ± 6.64	89.83 ± 7.41	88.96 ± 8.00

这些模型以与信用卡交易数据集实验中所用模型类似的方式进行了缩减。NQSVDD的经典骨干从78维输入中提取21维特征向量。ZZ特征嵌入与参数化层重复三次，将该向量编码至六量子比特系统，随后由量子卷积神经网络对最后三个活跃量子比特进行处理，得到16维潜在表示。NQSVDD、QSVDD和DSVDD分别包含225、75和456个可训练参数。如表3所示，NQSVDD的表现优于QSVDD和DSVDD。

4.4 噪声模拟

为了验证我们的方法在含噪声中等规模量子设备上的有效性，我们使用 IBM 量子设备 ibm kingston 的噪声参数，在噪声环境下进行了单类分类实验。噪声模型包含双量子比特去极化噪声和热弛豫噪声。误差概率为 $p \in [0,1]$ 的双量子比特去极化信道，会以总概率p均匀随机地施加十五种非恒等双量子比特泡利误差中的一种。热弛豫误差对振幅与相位阻尼进行建模，其参数包括能量弛豫时间 T_1、退相位时间 T_2和门持续时间（长度） $t_{\circ}$双量子比特去极化噪声和热弛豫噪声的具体数值列于表4。由于单量子比特去极化误差和读出误差的影响远小于双量子比特误差，且已有有效的测量误差缓解技术，因此未将其纳入考虑 [58–63]。

我们在噪声模拟上评估了NQSVDD，并将其OCC性能与理想的无噪声NQSVDD模拟以及经典DSVDD基线进行了比较。为防止NQSVDD的电路深度过大，我们将嵌入层数量限制为两层。每种方法均在四项单类分类任务上进行训练和评估，所使用的目标类别如下：MNIST中的“0”类别、Fashion-MNIST中的“T恤”类别、信用卡交易数据集中的“非欺诈”类别，以及网络入侵数据集中的“良性”类别。训练进行了500次迭代，批次大小为4，同时保留了先前理想模拟中的其余设置。为

表 2：在信用卡交易数据集上，每种方法使用 20 个随机种子所得的平均 AUC 分数（%）及标准差。

方法	参数	平均AUC
NQSVDD	210	$\mathbf{94.67} \pm 0.84$
QSVDD(amp)	75	87.98 ± 7.59
DSVDD	288	92.30 ± 2.24

表 3：网络入侵数据集上各方法使用 20 个随机种子所得的平均 AUC 分数（%）及标准差。

方法	参数	平均AUC
NQSVDD	225	$\mathbf{75.32} \pm 10.58$
QSVDD(amp)	75	56.04 ± 14.35
DSVDD	456	59.45 ± 23.46

与NQSVDD的噪声模拟设置保持一致，我们使用次优超参数（具体为批次大小和迭代次数）对无噪声NQSVDD和DSVDD进行了实验。对于每个数据集，我们报告了每种方法在五个不同随机种子上的平均曲线下面积及样本标准差。

噪声设置下的详细结果见表 5。硬件噪声下的 NQSVDD 始终不如无噪声版本，表明设备噪声会影响量子模型的有效性。尽管如此，即使在噪声环境下，NQSVDD 在所有数据集上的表现始终优于经典 DSVDD 基线，同时使用的参数数量约为其一半。这些结果表明，在 NISQ时代，量子单类分类方法相较于成熟的经典方法具有较强的竞争力。

5 结论

本研究提出了神经量子支持向量数据描述（NQSVDD），这是一种用于单类分类（OCC）的经典–量子混合学习框架。该框架将经典神经网络与量子数据嵌入和变分量子电路整合到统一的优化方案中。在这一混合架构中，输入数据首先由经典神经网络进行变换，以优化其映射到高维希尔伯特空间的量子特征映射；随后，数据经过交错的数据重上传和可训练变分量子电路处理，生成由量子测量定义、并针对支持向量数据描述目标函数进行优化的紧凑潜在表示。这种混合设计既能实现富有表现力的特征学习，又能保持较低的电路深度和较高的参数效率，因此适用于近期量子设备。

在 MNIST、Fashion MNIST、信用卡交易和网络入侵数据集上的评估结果证明了 NQSVDD 的有效性。就曲线下面积表现而言，即使可训练参数更少，所提出的模型也始终优于或至少不逊于经典DSVDD 和 QSVDD 基线。此外，即使在存在噪声的量子模拟中，与经典方法相比，NQSVDD 仍保持稳健的 OCC 性能，凸显了其抵御量子硬件噪声的能力，以及其在近期量子平台上实际部署的适用性。

总体而言，这些结果表明，NQSVDD 有效结合了量子特征变换的表达能力与经典神经网络的稳定性和可扩展性，为单类分类提供了一种实用的混合框架。未来值得探索的研究方向包括研究无梯度优化技术 [64]、改进适用于高维数据的量子嵌入策略，以及将所提出的框架扩展至半监督和多分类场景。此外，在大规模量子处理器上验证该模型，将有助于深入了解其在实际应用中的可扩展性和鲁棒性。最后，由于初始经典神经网络可以针对特定数据领域进行设计，探究在不同领域使用不同经典模型的效果，也是一个值得研究的方向。

表4：用于噪声模拟的 IBM 量子设备 ibm kingston 的中位误差率和门长度。T_1 和 T_2 分别表示能量弛豫时间和退相位时间。

后端参数	
双量子比特退极化	0.00332
单量子比特门长度（ns）	32
双量子比特门长度（ns）	68
T_1 （μs）	183.29
T_2 （μ秒）	141.73

表5：在无噪声和含噪条件下，对 MNIST、Fashion MNIST、信用卡交易、网络入侵数据集进行测试，各方法使用五个随机种子得到的平均AUC分数（%）及标准差。含噪NQSVDD以批次大小为4训练500次迭代。无噪声NQSVDD和经典DSVDD采用相同的次优配置。

数据集	方法	参数	条件	平均AUC
MNIST（目标：0）	NQSVDD	1090	无噪声	97.92 ± 1.33
			噪声	94.00 ± 2.37
	DSVDD	2152	-	92.11 ± 5.11
Fashion-MNIST（目标：T恤）	NQSVDD	1090	无噪声	90.14 ± 2.75
			噪声	85.84 ± 1.83
	DSVDD	2152	-	81.12 ± 6.69
信用卡交易	NQSVDD	195	无噪声	94.19 ± 1.07
			噪声	93.00 ± 2.17
	DSVDD	456	-	89.83 ± 2.50
网络入侵	NQSVDD	195	无噪声	70.67 ± 14.17
			噪声	64.50 ± 21.38
	DSVDD	456	-	56.10 ± 16.02

致谢

本研究获得了韩国国防技术规划与发展研究所通过国防采办计划管理局（DAPA）资助的项目支持（KRIT-CT-23-031）。本研究还获得了延世大学2025年研究基金、信息通信技术规划与评估院（IITP）通过韩国政府资助的项目（编号：2019-0-00003，容错量子计算系统编程、运行、实现与验证核心技术的研发）、韩国国家研究基金会（RS-2025-02309510），以及韩国贸易、工业和能源部（MOTIE）产业创新基础设施发展项目（项目编号：RS- 2024-00466693）的支持。

参考文献

[1] M. Cerezo、G. Verdon、H.-Y. Huang、L. Cincio 和 P. J. Coles，《量子机器学习中的挑战与机遇》，《自然·计算科学》，第2卷，第567–576页，2022年。[2] M. Schuld、I. Sinayskiy 和 F. Petruccione，《量子机器学习导论》，《当代物理学》，第56卷，第172–185页，2014年。[3] J. Biamonte、P. Wittek、N. Pancotti、P. Rebentrost、N. Wiebe 和 S. Lloyd，《量子机器学习》，《自然》，第549卷，第195–202页，2017年。[4] P. Rebentrost、M. Mohseni 和 S. Lloyd，《用于大数据分类的量子支持向量机》，《物理评论快报》，第113卷，第130503页，2014年。[5] M. Benedetti、E. Lloyd、S. Sack 和 M. Fiorentini，《作为机器学习模型的参数化量子电路》，《量子科学与技术》，第5卷，第019601页，2020年。[6] M. Cerezo、A. Arrasmith、R. Babbush、S. C. Benjamin、S. Endo、K. Fujii、J. R. McClean、K. Mitarai、X. Yuan、L. Cincio 和 P. J. Coles，《变分量子算法》，《自然评论物理》，第3卷，第625–644页，2021年。[7] A. Abbas、D. Sutter、C. Zoufal、A. Lucchi、A. Figalli 和 S. Woerner，《量子神经网络的力量》，《自然·计算科学》，第1卷，第403–409页，2021年。[8] M. M. Moya 和 D. R. Hush，《单类分类的网络约束与多目标优化》，《神经网络》，第9卷，第463–474页，1996年。[9] S. S. Khan 和 M. G. Madden，《单类分类：研究分类体系与技术综述》，《知识工程评论》，第29卷，第345–374页，2014年。[10] P. Perera、P. Oza 和 V. M. Patel，《单类分类：综述》，2021年。[11] R. Chalapathy 和 S. Chawla，《异常检测的深度学习：综述》，2019年。[12] B. Sch¨olkopf、R. C. Williamson、A. J. Smola、J. Shawe-Taylor 和 J. Platt，《用于新颖性检测的支持向量方法》，载于《神经信息处理系统进展》，第12卷，1999年。[13] D. M. Tax 和 R. P. Duin，《支持向量数据描述》，《机器学习》，第54卷，第45–66页，2004年。[14] L. Ruff、R. Vandermeulen、N. Goernitz、L. Deecke、S. A. Siddiqui、A. Binder、E. M¨uller 和 M. Kloft，《深度单类分类》，载于第35届国际机器学习会议论文集，第80卷，第4393–4402页，2018年。[15] C. Phua、V. Lee、K. Smith 和 R. Gayler，《基于数据挖掘的欺诈检测研究综合综述》，2010年。[16] S.-H. Li、D. C. Yen、W.-H. Lu 和 C. Wang，《利用数据挖掘技术识别欺诈账户的迹象》，《人类行为中的计算机》，第28卷，第1002–1013页，2012年。[17] M. Jeragh 和 M. AlSulaimi，《结合自编码器与单类支持向量机检测欺诈性信用卡交易》，载于2018年第二届智能系统、安全与可持续发展趋势世界会议（WorldS4），第178–184页，2018年。[18] P. Garc´ıa-Teodoro、J. D´ıaz-Verdejo、G. Maci´a-Fern´andez 和 E. V´azquez，《基于异常的网络入侵检测：技术、系统与挑战》，《计算机与安全》，第28卷，第18–28页，2009年。[19] K. Feher、J. Kirsch、A. Radbruch、H.-D. Chang 和 T. Kaiser，《基于异常的网络入侵检测：技术、系统与挑战》，《生物信息学》，第30卷，第3372–3378页，2014年。[20] S. Min、B. Lee 和 S. Yoon，《生物信息学中的深度学习》，《生物信息学简报》，第18卷，第851–869页，2017年。[21] K. Fraser、S. Homiller、R. K. Mishra、B. Ostdiek 和 M. D. Schwartz，《生物信息学中的深度学习》，《高能物理学杂志》，第2022卷，第851–869页，2022年。

[22] N. Liu 和 P. Rebentrost, 《用于量子异常检测的量子机器学习》, 《物理评论*A*》, 第97卷, 第042315页, 2018年。[23] K. Kottmann、F. Metz、J. Fraxanet 和 N. Baldelli, 《变分量子异常检测：在物理量子计算机上无监督映射相图》, 《物理评论研究》, 第3卷, 第043184页, 2021年。[24] A. Sakhnenko、C. O'Meara、K. J. B. Ghosh、C. B. Mendl、G. Cortiana 和 J. Bernab´e-Moreno, 《用于异常检测的混合经典-量子自编码器》, 《量子机器智能》, 第4卷, 第27期, 2022年。[25] V. S. Ngairangbam、M. Spannowsky 和 M. Takeuchi, 《利用量子自编码器检测高能物理中的异常》, 《物理评论*D*》, 第105卷, 第095004页, 2022年。[26] G. Park、J. Huh 和 D. K. Park, 《变分量子单类分类器》, *Machine Learning: Science andTechnology*, 第4卷, 第015006页, 2022年。[27] H. Oh 和 D. K. Park, 《用于异常检测的量子支持向量数据描述》, *Machine Learning:Science and Technology*, 第5卷, 第035052页, 2024年。[28]I. Cong、S. Choi 和 M. D. Lukin, 《量子卷积神经网络》, 《自然-物理学》, 第15卷, 第1273–1278页, 2019年。[29] Y. Lecun、L. Bottou、Y. Bengio 和 P. Haffner, 《应用于文档识别的基于梯度的学习》, 《*IEEE*会刊》, 第86卷, 第2278–2324页, 1998年。[30] I. Loshchilov 和 F. Hutter, 《Fashion-MNIST：用于基准测试机器学习算法的新型图像数据集》, 2017年。[31] 《Kaggle 信用卡欺诈检测：标记为欺诈或真实的匿名信用卡交易。》Kaggle。最后访问日期：2025年4月4日。[32] 《Kaggle 网络入侵数据集（CIC-IDS-2017）：网络数据集中的异常检测。》Kaggle。最后访问日期：2025年9月10日。[33] V. Bergholm、J. Izaac、M. Schuld、C. Gogolin、S. Ahmed、V. Ajith、M. S. Alam、G. Alonso-Linaje、B. Akash-Narayanan、A. Asadi、J. M. Arrazola、U. Azad、S. Banning、C. Blank、T. R. Bromley、B. A. Cordier、J. Ceroni、A. Delgado、O. D. Matteo、A. Dusko、T. Garg、D. Guala、A. Hayes、R. Hill、A. Ijaz、T. Isacsson、D. Ittah、S. Jahangiri、P. Jain、E. Jiang、A. Khandelwal、K. Kottmann、R. A. Lang、C. Lee、T. Loke、A. Lowe、K. McK- iernan、J. J. Meyer、J. A. Monta˜nez-Barrera、R. Moyard、Z. Niu、L. J. O'Riordan、S. Oud、A. Panigrahi、C.-Y. Park、D. Polatajko、N. Quesada、C. Roberts、N. S´a、I. Schoch、B. Shi、S. Shu、S. Sim、A. Singh、I. Strand- berg、J. Soni、A. Sz´ava、S. Thabet、R. A. Vargas-Hern´andez、T. Vincent、N. Vitucci、M. Weber、D. Wierichs、R. Wiersema、M. Willmann、V. Wong、S. Zhang 和 N. Killoran, 《PennyLane：混合量子-经典计算的自动微分》, 2022年。[34] C. Bravo-Prieto, 《采用增强数据编码的量子自编码器》, *Machine Learning: Science andTechnology*, 第2卷, 第035028页, 2021年。[35] T. Hur、I. F. Araujo 和 D. K. Park, 《神经量子嵌入：突破量子监督学习的极限》, 《物理评论*A*》, 第110卷, 第022411期, 2024年。[36] T. Hur 和 D. K. Park, 《通过间隔理解量子机器学习中的泛化》, 2024年。[37] V. Havl´ıˇcek、A. D. C´orcoles、K. Temme、A. W. Harrow、A. Kandala、J. M. Chow 和 J. M. Gambetta, 《利用量子增强特征空间进行监督学习》, 《自然》, 第567卷, 第209–212页, 2019年。[38] H. Liu、T. Hur、S. Zhang、L. Che、X. Long、X. Wang、K. Huang、Y. Fan、Y. Zheng、Y. Feng、X. Nie、D. K. Park 和 D. Lu, 《通过单量子比特确定性量子计算实现神经量子嵌入》, 《物理评论快报》, 第135卷, 第080603页, 2025年。[39] D. Shepherd 和 M. J. Bremner, 《时间无结构量子计算》, *Proc. A*, 第465卷, 第1413–1439页, 2009年。[40] M. J. Bremner、R. Jozsa 和 D. J. Shepherd, 《经典模拟对易量子计算意味着多项式层级坍缩》, 《英国皇家学会会刊*A*辑：数学、物理与工程科学》, 第467卷, 第459–472页, 2010年。

[41] M. J. Bremner、A. Montanaro 和 D. J. Shepherd，“对易量子计算的平均情况复杂度与近似模拟”，《物理评论快报》，第117卷，页码080501，2016年。[42] A. P´erez-Salinas、A. Cervera-Lierta、E. Gil-Fuster 和 J. I. Latorre，“用于通用量子分类器的数据重上传”，Quantum，第4卷，页码226，2020年。[43] J. R. McClean、S. Boixo、V. N. Smelyanskiy、R. Babbush 和 H. Neven，“量子神经网络训练景观中的贫瘠高原”，《自然-通讯》，第9卷，第4812期，2018年。[44] S. Sim、P. D. Johnson 和 A. Aspuru-Guzik，“混合量子-经典算法中参数化量子电路的表达能力与纠缠能力”，《先进量子技术》，第2卷，第12期，2019年。[45] M. Cerezo、A. Sone、T. Volkoff、L. Cincio 和 P. J. Coles，“浅层参数化量子电路中依赖于代价函数的贫瘠高原”，《自然-通讯》，第12卷，第1791期，2021年。[46] S. Wang、E. Fontana、M. Cerezo、K. Sharma、A. Sone、L. Cincio 和 P. J. Coles，“变分量子算法中的噪声诱导贫瘠高原”，《自然-通讯》，第12卷，第6961期，2021年。[47] E. Grant、M. Benedetti、S. Cao、A. Hallam、J. Lockhart、V. Stojevic、A. G. Green 和 S. Severini，“分层量子分类器”，npj量子信息，第4卷，第65期，2018年。[48] H.-Y. Huang、R. Kueng 和 J. Preskill，“通过极少量测量预测量子系统的多种性质”，《自然-物理学》，第16卷，第1050–1057页，2020年。[49] A. Pesah、M. Cerezo、S. Wang、T. Volkoff、A. T. Sornborger 和 P. J. Coles，“量子卷积神经网络中不存在贫瘠高原”，《物理评论X》，第11卷，页码041011，2021年。[50] L. Banchi、J. Pereira 和 S. Pirandola，“量子机器学习中的泛化：量子信息视角”，PRX Quantum，第2卷，页码040321，2021年。[51] T. Hur、L. Kim 和 D. K. Park，“用于经典数据分类的量子卷积神经网络”，《量子机器智能》，第4卷，第1期，2022年。[52] J. Kim、J. Huh 和 D. K. Park，“经典到量子的卷积神经网络迁移学习”，《神经计算》，第555卷，页码126643，2023年。[53] W. Liu、Z. Wang、X. Liu、N. Zeng、Y. Liu 和 F. E. Alsaadi，“深度神经网络架构及其应用综述”，《神经计算》，第234卷，第11–26页，2017年。[54] V. Verteletskyi、T.-C. Yen 和 A. F. Izmaylov，“利用最小团覆盖优化变分量子本征求解器中的测量”，《化学物理杂志》，第152卷，页码124114，2020年。[55] M. Schuld、V. Bergholm、C. Gogolin、J. Izaac 和 N. Killoran，“在量子硬件上计算解析梯度”，《物理评论A》，第99卷，页码032331，2019年。[56] D. P. Kingma 和 J. Ba，“Adam：一种随机优化方法”，2017年。[57] I. Loshchilov 和 F. Hutter，“SGDR：带热重启的随机梯度下降”，2017年。[58] S. Bravyi、S. Sheldon、A. Kandala、D. C. McKay 和 J. M. Gambetta，“缓解多量子比特实验中的测量误差”，《物理评论A》，第103卷，页码042605，2021年。[59] P. D. Nation、H. Kang、N. Sundaresan 和 J. M. Gambetta，“在量子计算机上可扩展地缓解测量误差”，PRX Quantum，第2卷，页码040326，2021年。[60] H. Kwon 和 J. Bae，“用于缓解量子算法中测量误差的混合量子-经典方法”，《IEEE计算机汇刊》，第70卷，页码1401–1411，2021年。[61] E. v. d. Berg、Z. K. Minev 和 K. Temme，“量子期望值的无模型读出误差缓解”，《物理评论A》，第105卷，页码032620，2022年。[62] J. Kim、B. Oh、Y. Chong、E. Hwang 和 D. K. Park，“通过深度学习缓解量子读出误差”，《新物理学杂志》，第24卷，页码073009，2022年。

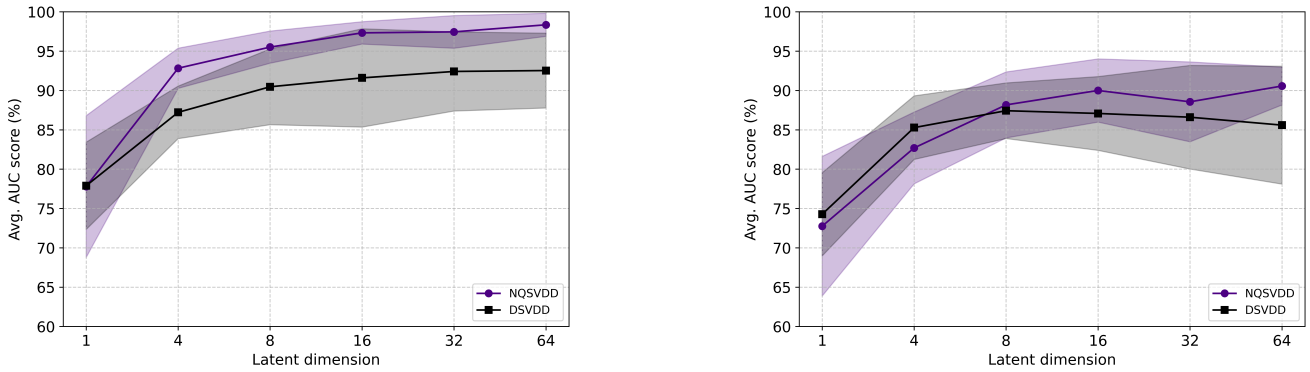
[63] C. Lee 和 D. K. Park，“利用条件独立性和迁移学习实现可扩展量子测量误差缓解”，*Machine Learning: Science and Technology*，第4卷，页码045051，2023年。

[64] X. Bonet-Monroig, H. Wang, D. Vermetten, B. Senjean, C. Moussa, T. B¨ack, V. Dunjko, and T. E. O’Brien, “变分量子算法中优化方法的性能比较”，*Phys. Rev. A.*，第 107 卷，032407 页，2023 年。

附录

A NQSVDD 仿真结果

本节提供了 NQSVDD 仿真的更多结果。图 A.1a、A.1b 和A.2 考察了随着模型规模增大，潜在空间维度和嵌入层数量对 OCC 性能的影响。如图 A.1a和图 A.1b所示，增大潜在空间维度可提升 NQSVDD 和 DSVDD 方法的性能，而 NQSVDD 在所有维度下均持续取得更高的 AUC 分数。图A.2 显示，增加 NQSVDD 中嵌入层的数量，可使其在各数据集上的性能略有提升。图 A.3 和A.4 展示了 NQSVDD 特征映射的有效性：训练后，它能够构建一个紧凑的最小体积超球体，将目标数据点包围其中。



(a) MNIST ”0”　　(b) Fashion-MNIST ”T恤”

图 A.1：在每个潜在空间维度下，使用 10 个随机种子得到的 (a) MNIST 数据集中的“0”和 (b) Fashion-MNIST 数据集中的“T恤”的平均AUC（%）及标准差（误差条）。

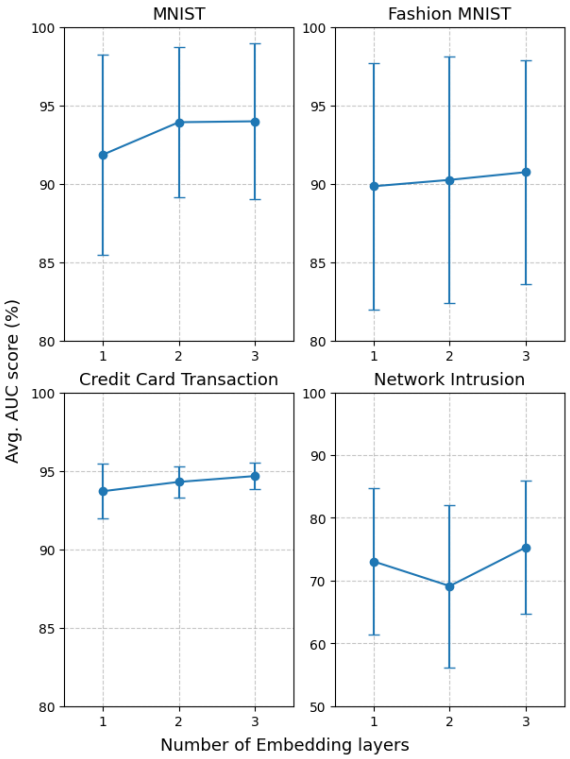
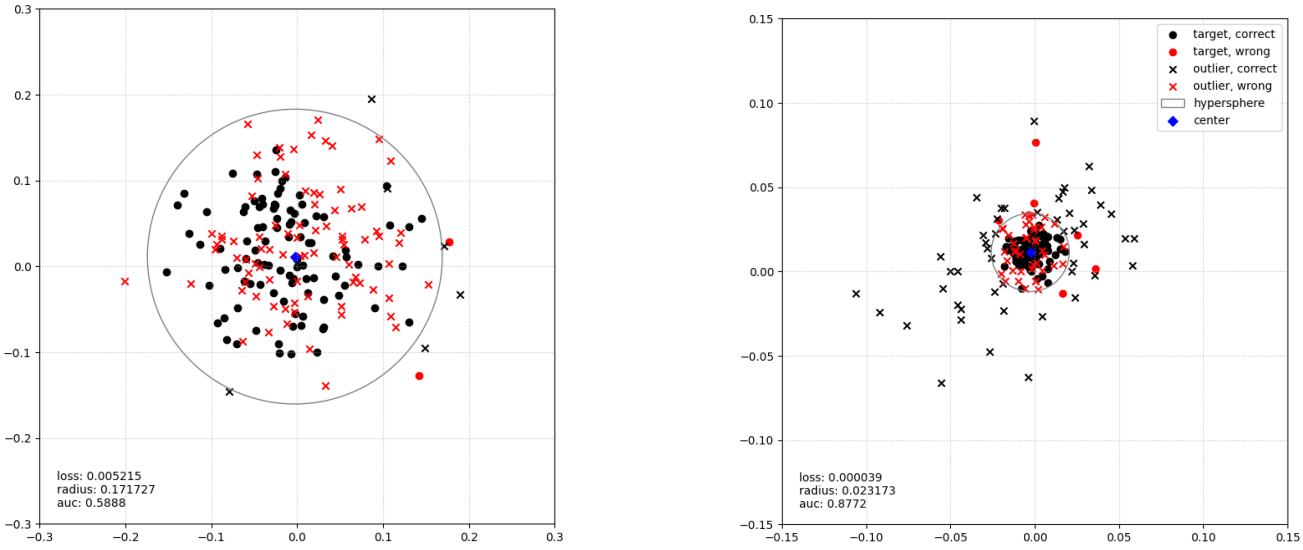
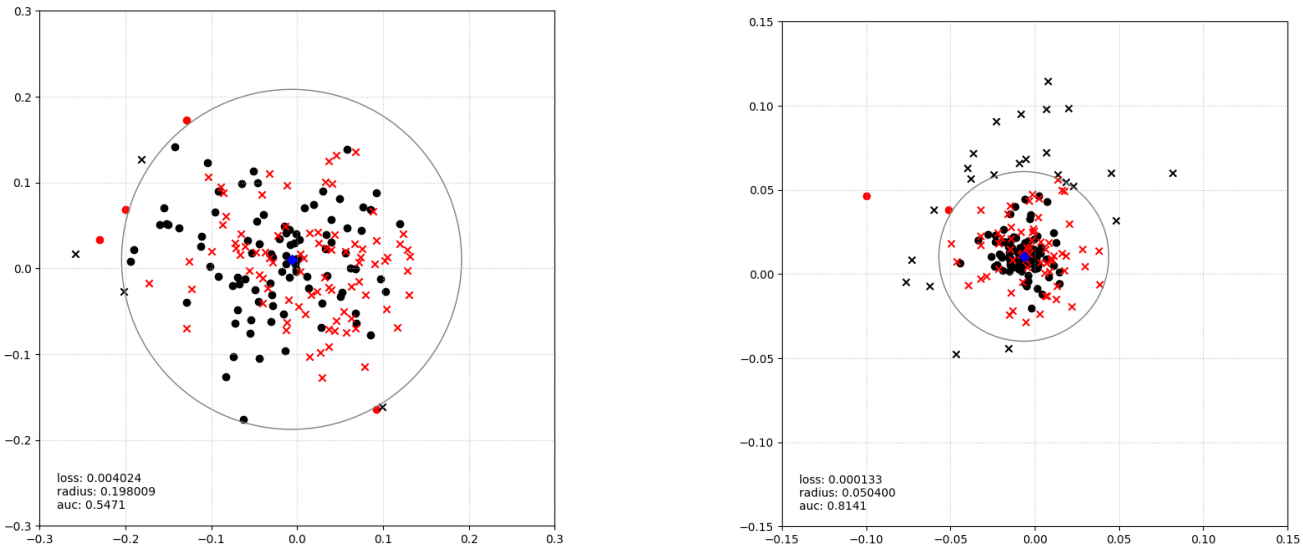


图 A.2：NQSVDD 中不同嵌入层数量的平均AUC（%）和标准差（误差棒）。对于 MNIST 和 Fashion MNIST，AUC 分数的平均值和标准差是在所有目标类别实验上计算的，每个目标类别使用 10 个随机种子（每种嵌入层数量对应 100 次实验）。对于信用卡交易和网络入侵数据集，结果是对每种嵌入层数量的 10 个随机种子取平均。



(a) 迭代 0　　(b) 迭代 1500

图 A.3：MNIST 测试数据的 NQSVDD 输出在二维空间中的可视化。为实现可视化，将 NQSVDD 的输出维度设为二。上图展示初始模型的输出，下图展示经过 1500 次训练迭代后的映射。



(a) 迭代 0　　(b) 迭代 1500

图 A.4：NQSVDD 对 Fashion-MNIST 测试数据在二维空间中的输出可视化。为便于可视化，将 NQSVDD 的输出维度设为 2。上图展示初始模型的输出，下图展示经过 1500 次训练迭代后的映射。