(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0103231
(43) 공개일자 2024년07월04일

(51) 국제특허분류(Int. Cl.)
G06N 10/60 (2022.01) *G06N 10/20* (2022.01)
G06N 3/0464 (2023.01) *B82Y 10/00* (2017.01)
(52) CPC특허분류
G06N 10/60 (2022.01)
G06N 10/20 (2022.01)
(21) 출원번호 10-2022-0185212
(22) 출원일자 2022년12월27일
심사청구일자 2022년12월27일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
김중헌
서울특별시 동작구 상도로 346-2, 212동 201호
윤원준
서울특별시 강남구 학동로64길 14,101동 1002호
백한결
서울특별시 도봉구 도봉로136가길 73, 101동 1304호
(74) 대리인
윤귀상

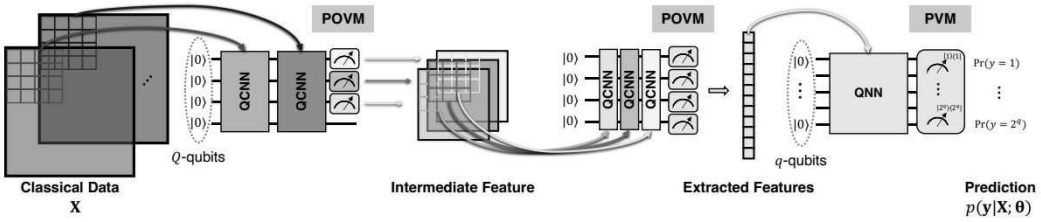
전체 청구항 수 : 총 7 항

(54) 발명의 명칭 **다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법**

(57) 요 약

본 발명의 일실시예는 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법에 있어서, 큐빗의 개수가 q인 입력데이터에 QCNN(Quantum Convolution Neural Network) 양자회로를 적용하되 POVM(positive operator-valued measure)을 기반으로 최종 특징벡터를 출력하는 POVM 단계; 및 상기 POVM 단계의 결과 출력된 상기 최종 특징벡터에 QNN(Quantum Neural Network) 양자회로를 적용하되 PVM(projection-valued measure)을 기반으로 q개의 큐빗보다 확장성(Scalability)이 증가된 다중클래스 예측벡터를 출력하는 PVM 단계;를 포함하는 것을 특징으로 하는, 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 제공한다.

대 표 도 - 도2



(52) CPC특허분류
G06N 3/0464 (2023.01)
B82Y 10/00 (2013.01)

명 세 서

청구범위

청구항 1

다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법에 있어서,

큐빗의 개수가 q인 입력데이터에 QCNN(Quantum Convolution Neural Network) 양자회로를 적용하되 POVM(positive operator-valued measure)을 기반으로 최종 특징벡터를 출력하는 POVM 단계; 및

상기 POVM 단계의 결과 출력된 상기 최종 특징벡터에 QNN(Quantum Neural Network) 양자회로를 적용하되 PVM(projection-valued measure)을 기반으로 q 개의 큐빗보다 확장성(Scalability)이 증가된 다중클래스 예측벡터를 출력하는 PVM 단계;를 포함하는 것을 특징으로 하는, 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법.

청구항 2

청구항 1에 있어서,

상기 PVM 단계는,

복수의 POVM채널로부터 각기 출력되는 상기 최종 특징벡터들을 접합(concatenation)시켜 특징들을 추출하는 단계; 및

추출된 특징들(extracted features)을 상기 QNN양자회로에 입력하고 PVM을 기반으로 상기 다중클래스 예측벡터를 위한 확률측정으로서 2^q 차원의 관측가능값(observables)을 출력하는 단계:를 포함하는 것을 특징으로 하는, 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법.

청구항 3

청구항 2에 있어서,

상기 POVM 단계는,

복수의 입력 채널을 통해 각각 상기 입력데이터로서 이미지를 입력받아 POVM을 기반으로 초기 특징벡터들을 출력하는 초기 QCNN 단계; 및

상기 초기 특징벡터들을 입력받아 POVM을 기반으로 각 상기 입력 채널별로 큐빗의 개수가 q인 상기 최종 특징벡터를 출력하는 추가의 QCNN 단계;를 포함하는 것을 특징으로 하는, 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법.

청구항 4

청구항 3에 있어서,

상기 추가의 QCNN단계는, 하나 이상의 QCNN 레이어를 포함하며 각 QCNN 레이어가 POVM을 기반으로 출력한 중간 특징벡터가 다음 QCNN 레이어로 입력되는 것을 특징으로 하는, 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법.

청구항 5

청구항 3에 있어서,

상기 QNN 양자회로의 학습에서 손실함수로서 확률 진폭 정규화기(Probability Amplitude Regularizer)가 사용되며,

상기 확률 진폭 정규화기 L_{PAR}는 상기 QNN 양자회로의 학습에서 사용되지 않은 클래스(unused Lables)에 대한 확률값을 제거하기 위해 하기 식(1)로 정의되며,

$$\mathcal{L}_{\mathrm{PAR}}(\Theta;\mathbf{X}) = -\sum_{n'>|\boldsymbol{y}|}^{2^q} \log(1-p_{n'}), \qquad (1)$$

상기 식(1)에서 $q \geq \lceil \log_2(|\boldsymbol{y}|) \rceil$이고, |y|는 분류하고자 하는 작업의 클래스 개수, 2^q는 QNN의 출력이고, Θ는 훈련 가능한 매개변수, X는 추출된 특징, $P_{n'}$는 n'번째 투영 행렬인 것을 특징으로 하는, 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법.

청구항 6

청구항 5에 있어서,

상기 QCNN 양자회로 및 상기 QNN 양자회로의 학습에서 손실함수로서 이진 교차 엔트로피 손실(Binary Cross-Entropy Loss) 및 상기 확률 진폭 정규화기(Probability Amplitude Regularizer)가 사용되며,

상기 이진 교차 엔트로피 손실 L_{BCE}는 하기 식(2)로 정의되는 것을 특징으로 하는,

$$\mathcal{L}_{\mathrm{BCE}}(\Theta;\mathbf{X}) = -\sum_{n=1}^{|\boldsymbol{y}|}[y_c \log p_c + (1-y_c)\log(1-p_c)], \quad (2)$$

다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법.

청구항 7

청구항 6에 있어서,

한 단계 단일 업데이트(one-step single update)를 위한 훈련 손실함수(train loss function)가 하기 식(3)으로 정의되는 것을 특징으로 하는,

$$\mathcal{L}(\Theta;\zeta) = \frac{1}{|\zeta|}\sum_{(\mathbf{X},\boldsymbol{y})\in\zeta}[\mathcal{L}_{\mathrm{BCE}}(\Theta;\mathbf{X}) + \mathcal{L}_{\mathrm{PAR}}(\Theta;\mathbf{X})]. \quad (3)$$

다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법에 관한 것으로, 더욱 상세하게는 양자회로 디자인 및 양자머신러닝 기법을 사용하여 다중클래스 분류작업을 양자머신러닝을 통해 해결하는 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법에 관한 것이다.

배 경 기 술

[0003] 일반적으로 변이 양자회로(Variational Quantum Circuit)는 양자회로에 학습가능한 파라미터를 이용하여 회로를 학습시킬 수 있게 하는 양자회로의 일종이다. 고전 컴퓨팅과 달리 양자회로에서의 양자컴퓨팅은 양자(큐빗) 상태를 측정함으로써 정보(1 또는 0)을 얻는 방법으로서, 기존의 고전 컴퓨팅에서 풀 수 없는 문제를 정해진 시간 안에 풀 수 있는 기법이다.

[0004] 최근 몇 년 동안 분류, 강화 학습 및 적대적 학습과 같은 다양한 작업에 이러한 양자 컴퓨팅을 이용한 양자머신러닝(Quantum Machine Learning, QML)이 활발히 사용되고 있다. 그러나 이러한 양자머신러닝 연구는 고전 머신러닝 기법에 비해 다음의 단점 있다.

[0005] 첫째, QML은 입력 및 출력에 대한 확장성(Scalability) 문제로 인해 복잡한 작업을 수행하지 못하는 것이 가장 큰 한계로 지적되고 있다. 즉, 종래의 양자머신러닝 기법은 입력 큐빗의 차원, 출력큐빗의 차원이 동일(q)하게 사용하고, 고전 컴퓨터에서 양자회로를 학습시키려면 2q x 2q 의 그래디언트를 학습시켜야 하기 때문에 매우 많은 연산량이 필요하다. 이에 따라 QML은 주로 이진 분류(binary classification)를 중심으로 연구가 이루어져 왔다.

[0006] 둘째, 종래의 양자머신러닝은 저성능(Low performance)이 문제되고 있다. 예를 들어, MNIST(Modified National Institute of Standards and Technology database) 데이터를 사용한 경우에 있어서, 고전 머신러닝을 활용할 경우, 정확도가 99% 이상 나오지만, 양자머신러닝에서 MNIST를 그대로 활용하여 학습한 경우는 단 한 건이며, 성능이 32.5%가 나오는 것으로 알려져 있다.

[0007] 셋째로, 양자머신러닝은 학습성(Trainability)에 문제가 있다. 즉, 양자회로 또는 양자컴퓨터가 학습시키기 어렵기 때문에 양자머신러닝은 학습이 잘 되지 않는 문제가 존재한다.

발명의 내용

해결하려는 과제

[0009] 본 발명이 이루고자 하는 기술적 과제는 양자회로에서 큐빗의 개수가 많아질수록 양자오류가 심해지는 큐빗 개수 제한의 문제점에 대해 PVM과 POVM을 머신러닝의 주요 양자회로로 사용하여 해결하는 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 제공하는 것이다.

[0010] 본 발명이 이루고자 하는 다른 기술적 과제는 입력차원과 출력차원이 output $\leq 2^{input}$ 관계를 가지는 PVM 방법을 이용하여 학습시킴으로써 큐빗을 종래기법 대비 최대 [log2(q)]으로 줄일 수 있고, 학습성능도 높은 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 제공하는 것이다.

[0011] 본 발명이 이루고자 하는 기술적 과제는 이상에서 언급한 기술적 과제로 제한되지 않으며, 언급되지 않은 또 다른 기술적 과제들은 아래의 기재로부터 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 명확하게 이해될 수 있을 것이다.

과제의 해결 수단

[0013] 상기 기술적 과제를 달성하기 위하여, 본 발명의 일실시예는 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법에 있어서, 큐빗의 개수가 q인 입력데이터에 QCNN(Quantum Convolution Neural Network) 양자회로를 적용하되 POVM(positive operator-valued measure)을 기반으로 최종 특징벡터를 출력하는 POVM 단계; 및 상기 POVM 단계의 결과 출력된 상기 최종 특징벡터에 QNN(Quantum Neural Network) 양자회로를 적용하되 PVM(projection-valued measure)을 기반으로 q 개의 큐빗보다 확장성(Scalability)이 증가된 다중클래스 예측벡터를 출력하는 PVM 단계;를 포함하는 것을 특징으로 하는, 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 제공한다.

[0014] 본 발명의 실시예에 있어서, 상기 PVM 단계는, 복수의 POVM채널로부터 각기 출력되는 상기 최종 특징벡터들을 접합(concatenation)시켜 특징들을 추출하는 단계; 및 추출된 특징들(extracted features)을 상기 QNN양자회로에 입력하고 PVM을 기반으로 상기 다중클래스 예측벡터를 위한 확률측정으로서 2^q 차원의 관측가능값(observables)을 출력하는 단계:를 포함할 수 있다.

[0015] 본 발명의 실시예에 있어서, 상기 POVM 단계는, 복수의 채널을 통해 각각 상기 입력데이터로 이미지를 입력받아 POVM을 기반으로 초기 특징벡터들을 출력하는 초기 QCNN 단계; 및 상기 초기 특징벡터들을 입력받아 POVM을 기반으로 각 상기 채널별로 큐빗의 개수가 q인 상기 최종 특징벡터를 출력하는 추가의 QCNN 단계;를 포함하고, 상기 추가의 QCNN단계는, 하나 이상의 QCNN 레이어를 포함하며 각 QCNN 레이어가 POVM을 기반으로 출력한 중간 특징벡터가 다음 QCNN 레이어로 입력되는 것을 특징으로 할 수 있다.

[0016] 본 발명의 실시예에 있어서, 상기 QNN 양자회로의 학습에서 손실함수로서 확률 진폭 정규화기(Probability Amplitude Regularizer)가 사용되며, 상기 확률 진폭 정규화기 L_{PAR}는 상기 QNN 양자회로의 학습에서 사용되지 않은 클래스(unused Lables)에 대한 확률값을 제거하기 위해 하기 식(1)로 정의되며,

[0017] $$\mathcal{L}_{\text{PAR}}(\Theta;\mathbf{X}) = -\sum_{n'>|\boldsymbol{y}|}^{2^q} \log(1-p_{n'}), \qquad (1)$$

[0018] $q \geq \lceil \log_2(|\boldsymbol{y}|) \rceil$ 인 것을 특징으로 할 수 있다.

[0019] 본 발명의 실시예에 있어서, 상기 QCNN 양자회로 및 상기 QNN 양자회로의 학습에서 손실함수로서 이진 교차 엔트로피 손실(Binary Cross-Entropy Loss) 및 상기 확률 진폭 정규화기(Probability Amplitude Regularizer)가 사용되며, 상기 이진 교차 엔트로피 손실 L_{BCE}는 하기 식(2)로 정의될 수 있다.

[0020] $$\mathcal{L}_{\text{BCE}}(\Theta;\mathbf{X}) = -\sum_{n=1}^{|\boldsymbol{y}|}[y_c \log p_c + (1-y_c)\log(1-p_c)], \quad (2)$$

[0021] 본 발명의 실시예에 있어서, 한 단계 단일 업데이트(one-step single update)를 위한 훈련 손실함수(train loss function)가 하기 식(3)으로 정의될 수 있다.

[0022] $$\mathcal{L}(\Theta;\zeta) = \frac{1}{|\zeta|}\sum_{(\mathbf{X},\boldsymbol{y})\in\zeta}[\mathcal{L}_{\text{BCE}}(\Theta;\mathbf{X}) + \mathcal{L}_{\text{PAR}}(\Theta;\mathbf{X})]. \quad (3)$$

발명의 효과

[0024] 본 발명의 실시예에 의하면, 양자회로에서 큐빗의 개수가 많아질수록 양자오류가 심해지는 큐빗 개수 제한의 문제점을 PVM과 POVM을 활용하여 해결함으로써 가장 최신 기술(State-of-the-art) 대비 약 40%이상의 정확도 향상을 더 적은 큐빗을 사용하여 달성하는 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 제공할 수 있다.

[0025] 본 발명의 실시예에 의하면, 입력차원과 출력차원이 output $\leq 2^{\text{input}}$ 관계를 가지는 PVM 방법을 이용하여 학습시킴으로써 큐빗을 종래기법 대비 최대 [log2(q)]으로 줄일 수 있고, 학습성능도 높은 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 제공할 수 있다.

[0026] 본 발명의 실시예에 의하면, 확률 진폭 정규화기(probability amplitude regularizer)를 도입하여 학습성 향상과 예측 성능 향상을 달성하는 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 제공할 수 있다.

[0027] 본 발명의 실시예에 의하면, 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법은 모든 양자머신러닝에 대해 활용 가능하다.

[0028] 본 발명의 효과는 상기한 효과로 한정되는 것은 아니며, 본 발명의 상세한 설명 또는 특허청구범위에 기재된 발명의 구성으로부터 추론 가능한 모든 효과를 포함하는 것으로 이해되어야 한다.

도면의 간단한 설명

[0030] 도1은 POVM과 PVM의 차이점을 설명하기 위한 도면이다.

도 2는 본 발명의 일 실시예에 따른 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 나타내는 도면이다.

도 3은 도2에 제시된 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법의 알고리즘을 나타낸다.

도 4는 도 2에서 초기 QCNN을 나타내는 도면이다.

도 5는 도 2에서 추가의 QCNN을 나타내는 도면이다.

도 6은 도 2에서 POVM을 통한 측정에서 얻은 정보를 입력받아 PVM에 의한 출력을 내리는 과정을 나타내는 도면이다.

도 7은 도 2에 제시된 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법에서 사용된 손실함수를 설명하기 위한 도면이다.

도 8은 도 2에 제시된 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법에서 성능 테스트에 사용된 데이터 세트를 설명하기 위한 표이다.

도 9는 도 8에 제시된 데이터 세트 중 MNIST 데이터 세트의 학습 곡선을 보여주는 도면이다.

도 10은 도 8에 제시된 데이터 세트 중 FashionMNIST 데이터 세트의 학습 곡선을 보여주는 도면이다.

도 11는 도 8에 제시된 데이터 세트 중 CIFAR10 데이터 세트의 학습 곡선을 보여주는 도면이다.

도 12는 도 8에 제시된 데이터 세트 중 EMNIST-letters 데이터 세트의 학습 곡선을 보여주는 도면이다.

도 13은 도 2에 제시된 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법과 QuantumNAS의 성능비교를 나타내는 도면이다.

도 14는 도 2에 제시된 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법과 확률 진폭 정규화기를 제거한 경우의 성능을 비교한 결과를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0031] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 형태를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 본문에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 개시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.

[0032] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0033] 이하, 첨부한 도면들을 참조하여, 본 발명의 바람직한 실시예를 보다 상세하게 설명하고자 한다.

[0034] 본 실시예의 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법은 양자머신러닝(QML)이 다중클래스 분류를 성공적으로 달성하는 방법을 제공한다. 본 실시예에서는 종래의 QML의 가장 큰 한계로 지적되고 있는 입력 및 출력에 대한 확장성 문제로 인해 복잡한 작업을 수행하지 못하는 점을 극복하기 위해, 양자통계역학에서 확률 진폭의 특성을 활용하는 PVM(projection-valued measure)을 활용한다. 이러한 본 실시예의 특징으로 인해 출력 차원의 큐빗 수가 q에서 $O(2^q)$로 확장될 수 있다. 즉, 본 실시예의 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법은 다중클래스 분류를 위한 새로운 QML 프레임워크를 제공할 수 있다. 또한, 성능면에서 본 실시예의 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법은 6큐빗 이하를 사용하는 다양한 데이터 세트로 최신 기술(SOTA)을 능가한다는 것을 확인했다. 또한 본 실시예가 포함하는 PVM 기반 QML은 최신 기술(SOTA)보다 성능이 42.2% 더 향상된 것을 확인하였다. 이에 대해서는 후술된다.

[0036] 이하, 설명에서 사용될 표기법에 대해 다음과 같이 정의한다.

[0037] 훈련 가능한 매개변수에 대해서는 표기법 $\Theta = \{\boldsymbol{\theta}^{\text{enc}}; \boldsymbol{\theta}^{\text{PQC}}\}$ 를 사용한다.

[0038] $\zeta \triangleq \{(X, y)\}$ 를 샘플링된 미니 배치로 정의한다. 여기서 X와 y는 각각 샘플링된 입력 데이터와 미니 배치의 해당 레이블을 나타낸다.

[0039] 레이블 $\boldsymbol{y} \triangleq \{y_n\}_{n=1}^{|\boldsymbol{y}|}$ 은 실제 레이블이 n ∈ N[1, |y|]이고 다른 요소가 0인 경우 y_n = 1인 원-핫 벡터이다.

[0040] 추출된 특징은 $\hat{\mathbf{X}}$ 로 표시된다.

[0041] 양자 상태와 그 연산을 나타내기 위해 Dirac 표기법이 사용된다.

[0042] 또한 연산자 $(\cdot)\dagger$ 및 $\otimes$는 각각 복소수 켤레 전치 및 텐서 곱을 나타낸다. 고전 데이터를 인코딩하는 Q-큐빗 시스템과 예측을 수행하는 q-큐빗 시스템을 나타내기 위해 두 용어 Q와 q를 별도로 사용한다.

[0043] Q 큐빗 양자 상태는 다음 식(1)과 같이 정의된다.

$$|\psi\rangle = \sum_{n=1}^{2^Q} \alpha_n |n\rangle , \qquad (1)$$

[0044]

[0045] 여기서 α_n과 $|n\rangle$은 각각 힐베르트 공간에서 확률 진폭과 n번째 기저를 나타낸다.

[0046] Hilbert 공간의 정의(즉, $\mathcal{H}^{\otimes Q} \equiv \mathbb{C}^{2^Q}$)에 따라 확률 진폭 $\forall \alpha_n \in C$는 다음과 같은 관계를 갖는다. 즉, $\sum_{n=1}^{2^Q} |\alpha_n|^2 = 1$ 이다.

[0048] 도1은 POVM과 PVM의 차이점을 설명하기 위한 도면이다.

[0049] 본 실시예에서 사용된 양자회로는 크게는 2가지로서, 하나는 QCNN이라는 양자회로이고, 다른 하나는 QNN이라는 양자회로이다. 도 1에는 POVM(positive operator-valued measure)과 PVM(projection-valued measure)의 특징이 도시되어 있다. 본 실시예에서는 영상 처리를 위해 POVM을 활용하고 다중클래스 예측을 위해 PVM 방법을 활용한다.

[0050] 도 1에서 왼쪽의 박스는 q큐빗의 양자 상태가 중첩된 양자 상태를 나타낸다. 우상측 박스는 POVM 을 나타낸다. POVM 에서는 각 큐빗에 대해서 측정을 하는 과정을 수행한다. 이때, 측정하는 축으로는 pauli-Z에 대해 측정을 내리게 되며, 이때, 입력이 q차원이라면 q차원의 아웃풋을 얻게 된다. 구체적으로, 이러한 중첩된 양자상태를 입력받은POVM이 출력을 내리게 된다. 그 출력을 할 때 측정 베이시스(bases, 측정축)로 pauli-Z라는 베이시스에 대해 측정을 내리게 되는데, 이렇게 측정하게 되면 POVM은 입력이 q차원이면 q차원의 아웃풋을 얻게 되는 양자회로를 가진다.

[0051] 이러한 과정을 더 구체적으로 설명하면, POVM에서는 다음과 같이 출력을 생성한다. 양자 상태는 모든 큐빗에 대해 Pauli-Z 측정 장치로 투영되며 여기서 Pauli-Z는 Z = diag(1, -1)이다. n번째 투영 행렬은 $\mathbf{P}_n \triangleq \mathbf{I}^{\otimes n-1} \otimes \mathbf{Z} \otimes \mathbf{I}^{\otimes Q-n}$ 으로 정의되며, 여기서 I는 각각 2 X 2 항등 행렬을 나타낸다. POVM에서는 투영 P_n의 기대값을 $O_n \in R[-1, 1], (\forall n \in N[1, Q])$로 표시되는 확률 변수로 만든다. m번째 큐빗의 기대값은 $\langle O_n \rangle = \langle \psi | P_n | \psi \rangle$로 정의된다. 양자측정이론(Quantum measurement theory)에 따르면 프로젝션(투영)의 기대값은 다음 식(2)와 같이 q-클래스에 대한 예측을 만들기 위해 소프트맥스(softmax) 함수 및 온도 매개변수 $\beta \in R$과 함께 사용할 수 있다.

[0052] $$\Pr_{POVM}(y=n)=\frac{\exp(\beta\langle O_n\rangle)}{\sum_{n=1}^{q}\exp(\beta\langle O_n\rangle)}. \qquad (2)$$

[0053] POVM에서 큐빗 수 q는 q ≥ |y|를 충족해야 한다. 또한 기존 NN이 있는 POVM은 q ≥ |y| 조건 없이 대규모 예측을 수행하는 데 널리 사용된다.

[0054] 대부분의 양자컴퓨터와 QML 프레임 작업은 POVM을 따라 출력을 얻는다. 그러나 q-qubit 시스템에서 POVM 기반 출력의 크기는 큐빗 수, 즉 O(q)를 초과할 수 없으며 이는 QNN의 스케일러블(scalable)한 출력을 방해한다.

[0055] 도 1의 우하측의 PVM박스는 각 큐빗을 모든 가능한 양자 상태에 대해 측정을 하는 과정을 나타낸다. 이때 측정축은 모든 가능한 양자 상태이며, 입력이 q차원이라면 2^q 차원의 확률에 대한 출력을 얻게 된다. PVM박스는 1부터 2^q의 양자상태를 나타내는데 존재할 수 있는 모든 양자상태축에 대해서, 측정을 내리게 된다. 그 측정을 하게 되면 입력차원이 q차원이더라도 내릴 수 있는 출력 차원은 2^q차원의 출력을 내일 수 있게 되고, 그 출력값으로는 확률에 대한 출력을 얻게 된다. 즉, PVM은 입력이 q차원이고 출력이 2^q 차원의 확률을 얻어내는 방법이다.

[0056] 이를 더 구체적으로 설명하면, 본 실시예는 POVM의 특수한 유형인 분류를 위해 PVM을 사용한다. PVM은 양자 상태를 투영기 $\{|n\rangle\langle n|\}_{n=1}^{2^q}$에 투영하여 확률 측정을 하고 그 출력은 다음 식(3)과 같이 나타낼 수 있다.

[0057] $$\Pr_{PVM}(y=n)=\langle\psi|n\rangle\langle n|\psi\rangle=|\langle\psi|n\rangle|^2=|\alpha_n|^2. \qquad (3)$$

[0058] 이러한 PVM의 출력은 확률 분포이다. 여기서

[0059] $\Pr_{PVM}(y=n)=\sum_{n=1}^{2^q}|\alpha_n|^2=1$ 이다.

[0060] 이와 같이, 본 실시예에서는 QCNN을 이용한 이미지 처리와 QNN을 이용한 분류를 위해 POVM과 PVM을 각각 사용한다. 이하, 본 발명의 일 실시예에 따른 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 설명한다.

[0062] 도 2는 본 발명의 일 실시예에 따른 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 나타내는 도면이다. 도 2는 본 실시예의 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법에서 활용한 이미지 분류 프로세스를 나타낸다. 도 3은 도 2에 제시된 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법의 알고리즘의 일 예를 나타낸다.

[0063] 본 실시예의 다중클래스 분류를 위한 QML 프레임워크(다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법)는 크게는 QCNN과 QNN으로 구성된다. QCNN/QNN의 구조는 상태 인코딩, 매개변수화 양자회로(PQC)를 사용한 선형 변환 및 측정의 세 부분으로 구성된다. 우선, 고전적인 입력 크기는 큐빗 수보다 크므로 상태 인코더에 대한 데이터 재업로드를 한다.

[0064] 기존 입력 데이터 X를 성공적으로 인코딩하기 위해 X는 [x1; · · · ; x_{cin}]로 분할된다. 고전적인 분할 데이터는 확률 진폭으로 인코딩되며 인코딩 프로세스는 다음 식(4)와 같이 표현할 수 있다.

[0065] $$|\psi_{\text{enc}}\rangle=U(\boldsymbol{\theta}_{c_{\text{in}}})U(\mathbf{x}_{c_{\text{in}}})\cdots U(\boldsymbol{\theta}_1)U(\mathbf{x}_1)|\psi_0\rangle, \qquad (4)$$

[0066] 여기서 $|\Psi_0\rangle$는 초기 양자 상태, 예를 들어 2^0 차원 벡터의 첫 번째 표준 기저와 $\forall\Theta_c\subset\Theta_{\text{enc}}$ 및 $\forall c\in N[1, c_{in}]$을 나타낸다.

[0067] 인코딩된 상태 $|\Psi_{\text{enc}}\rangle$는 결과가 $|\psi_{\text{PQC}}\rangle=U(\boldsymbol{\theta}^{\text{PQC}})|\psi_{\text{enc}}\rangle$로 표시되는 PQC로 처리된다. 처리된 양자 상태 $|\Psi_{\text{PQC}}\rangle$는 각각 QCNN의 POVM과 QNN의 PVM으로 측정된다.

[0068] 이하, 도 2 및 도 3에 제시된 각 과정을 더 구체적으로 설명한다.

[0069] 도 4 및 도 5는 도 2에서 이미지 처리를 위한 POVM 기반 QCNN을 나타내는 도면들이다. 도 4는 도 2에서 초기 QCNN을 나타내는 도면이다. 도 5는 도 2에서 추가의 QCNN을 나타내는 도면이다.

[0070] 도 4 및 도 5를 참조하면, 각 2D 그리드는 각 채널의 패치를 나타낸다. QCNN에서 양자 상태를 측정할 때 각 큐빗에 대한 투영의 예상 값은 출력 채널의 값이다.

[0071] 도 3의 알고리즘 1의 라인 2-13은 도 4 및 도 5에서 QCNN을 사용한 이미지 처리를 나타내며, 도 4에 도시된 바와 같이, 입력측에서는 스케일러블한 QCNN이 사용되며, 초기 QCNN에 대한 입력은 이미지일 수 있다. 이 입력 이미지는 클레시컬(classical) 데이터로서 도 2에서 좌측의 사각 박스들로 표시되어 있고, 도 8의 테이블의 좌측의 각 테이터셋을 나타내고 있다.

[0072] 또한, 도 5에 도시된 바와 같이, 추가의 QCNN에 대한 입력은 초기 QCNN에서 나온 출력을 입력으로 받을 수 있다. 이러한 추가의 QCNN은 복수의 레이어로 구성될 수 있다. 양자회로 QCNN는 공간정보를 뽑아내는 역할을 하며, 그때마다 더 추상화되는 과정이 도 4 및 도 5에 표현되어 있다.

[0073] 도 4 및 도 5에서 설명된 QCNN 처리에서 입력은 W X H X c_{in} 모양을 갖는다. $\hat{W}$ 및 $\hat{H}$는 커널 크기 k, 스트라이드 s 및 패딩 d에 따라 달라진다. 즉, $\hat{W} = (W + d)/s$ and $\hat{H} = (H + d)/s$ 이다. 입력의 각 패치는 데이터 재업로드 방법에 의해 QCNN에 피드 포워드(feed-forwarded)된다. QCNN의 풀링(pooling)은 POVM에 의해 발생한다. POVM O_n의 n번째 기대값은 n번째 채널의 스칼라 값에 해당한다.

[0074] 출력은 $\hat{W} \times \hat{H} \times c_{\text{out}}$ 형태를 갖는다. 계층당 QNN의 계산 복잡도는 $\mathcal{O}(\hat{W} \cdot \hat{H} \cdot \kappa^2 \cdot c_{\text{in}})$ 이며, 여기서 k는 커널 크기를 나타낸다. 반면 기존 CNN의 계산 복잡도는 $\mathcal{O}(\hat{W} \cdot \hat{H} \cdot \kappa^2 \cdot c_{\text{in}} \cdot c_{\text{out}})$ 이다. QCNN의 입력 및 출력 채널은 제약 조건 $\forall c_{in}, c_{out} \in N[1, Q]$에서 확장 가능하다.

[0075] 이때 QCNN들의 각 레이어 하나당 출력값들은 전술된 POVM으로 얻게 된다. 즉, 이러한 초기 QCNN및 추가의 QCNN의 각 단계에서 출력들은 도 1에서 전술된 POVM으로 얻게 된다. 이와 같은 QCNN(Quantum Convolution Neural Network)은 입력 데이터에 대한 훈련 가능성과 확장성을 해결할 수 있다.

[0076] 도 4 및 도 5의 과정의 최종 추출된 특징은 후술되는 QNN의 입력으로 주어지는 벡터화된 클래식 데이터이다. QNN은 클래스 예측에 해당하는 확률 측정을 수행한다. 그러나 QML에 있어서 여전히 전술한 바와 같이 출력의 확장성 문제가 있다.

[0077] 도 6은 도 2에서 POVM을 통한 측정에서 얻은 정보를 입력받아 PVM에 의한 출력을 내리는 과정을 나타내는 도면이다.

[0078] 본 실시예에서는 출력의 확장성을 POVM의 특별한 형태인 PVM(projection-valued measure)을 통해서 해결한다. QNN에서 PVM을 활용하면 q-qubits의 전체 기반을 활용하여 $O(2^q)$ 출력(observables라고 함)을 만들 수 있다.

[0079] 도 6을 참조하면, PVM 단계에서는 마지막 추가의 QCNN을 통해 출력된 특징들(features)을 접합(concatenation)시키고, 이것을 QNN양자회로에 마지막으로 넣어주게 된다. 이때 출력을 2^q 스케일로 얻어내기 위해서 전술된 PVM을 통해서 측정을 하게 된다. 즉, 출력은 PVM을 통한 측정방법을 채택하여 2^q가지 경우의 수에 대한 확률인 프레딕션 $p(\mathbf{y}|\mathbf{X};\boldsymbol{\theta})$ 을 얻어낼 수 있다. 이와 같이, PVM의 특성을 이용하여 기존의 NN 방법론을 사용하지 않고 다중클래스 분류를 위한 양자 기계 학습 프레임워크를 제공할 수 있다.

[0080] 도 6에서 설명된 상기 분류를 위한 PVM 기반 QNN의 내용을 더 구체적으로 설명하면 다음과 같다.

[0081] 일반적으로 클래스의 수 |y|는 큐빗 수 q보다 크다. 그러나 큐빗 수가 증가하면 QCNN/QNN을 훈련시키기가 어렵다. 따라서 전술된 식(2)와 같은 softmax 기반 POVM은 단순한 다중클래스 분류(예: 이진 분류 또는 4클래스 분류)로 제한된다. 따라서 본 실시예에서는, 알고리즘 1(라인 14-20)에 프로세스가 제시된 분류를 위해 PVM 기반 QNN을 사용한다. QCNN 처리와 유사하게 추출된 특징은 전술된 식(4)로 인코딩된다. 다중클래스 분류를 가능하게

하기 위해 PVM으로 QNN 측정을 설계한다. 그런 다음 양자 상태 $|\Psi\rangle$가 순수한 밀도 행렬, 즉 $\{|1\rangle\langle 1|, \cdots, |n\rangle\langle n|, \cdots, |2^q\rangle\langle 2^q|\}$로 투영될 때 2^q의 관측가능항목 또는 관측가능값(observables) 확률을 얻을 수 있다.

[0082] n번째 클래스의 확률은 다음 식(5)와 같이 표현할 수 있다.

[0083] $$p_n = \Pr(y = n|\mathbf{X};\Theta) = |\alpha_n|^2. \qquad (5)$$

[0084] 본 실시예의 QML 프레임워크(다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법)는 예측에 대한 로짓(logits)을 만들기 위해 softmax 함수를 필요로 하지 않는다. 확률 진폭 α_n은 클래스 n의 확률에 직접 매핑된다. 또한 본 실시예의 QML 프레임워크에는 softmax 온도 계수가 필요하지 않다.

[0085] 한편, 본 실시예의 QML 프레임워크는 사용하지 않는 클래스의 확률로 인해 확률에 오류가 발생하는 것을 방지하기 위해, 후술되는 바와 같이 확률 진폭 정규화기를 포함한다.

[0086] 도 7은 도 2에 제시된 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법에서 사용된 손실함수를 설명하기 위한 도면이다.

[0087] 도 7을 참조하면, 본 실시예의 다른 주요 특징으로서, 관측가능항목(observables)이 실제 레이블에 맞도록 만드는 확률 진폭 정규화기를 사용한다. 이는 본 실시예의 QML 프레임워크의 성능을 크게 향상시킨다.

[0088] 구체적으로, 도 2 에 도시된 이러한 양자 파이프라인에서, 실제 분류하고자 하는 작업의 클래스 개수가 |y|이고, QNN의 출력이 2^q이기 때문에, 성공적인 분류를 수행하기 위해서는 $q \geq \lceil \log_2(|\boldsymbol{y}|) \rceil$ 의 조건을 만족하는 자연수 값을 q가 가져야한다.

[0089] 도 7은 |y|=10, q=4인 경우에 관한 목적함수(손실함수)를 나타낸다. 본 실시예에서는 두 가지 손실 함수(즉, 이진 교차 엔트로피 손실 및 확률 진폭 정규화기)를 사용한다. y=10인 것은 도 7에서 Binary Cross-Entropy Loss 박스에서 막대가 10개인 것으로 표현되고 있고, q=4인 것은 2^4 에 대한 확률(probability) 값을 다 갖고 있는 것을 의미한다. 그러면 실재학습을 하게될 때에 이 Binary Cross-Entropy Loss 박스만을 가지고 학습을 하게 되면 확률 진폭 정규화기(Probability Amplitude Regularizer) 박스 내에 막대로 표현된 사용되지 않은 레이블(unused Lables)이 존재하게 되어 학습이 잘 일어나지 않는 문제가 발생될 수 있다. 이로 인해. 실제 사용되지 않는 클래스에 대한 확률값이 잔존하게 된다.

[0090] 본 실시예의 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법은 확률 진폭 정규화기(Probability Amplitude Regularizer) 박스로 표현된 실재 사용되지 않는 그 클래스에 대한 확률값을 제거하기 위해, 즉, 본 실시예의 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법은 학습성능을 높이기 위해 아래의 식(7)과 같은 정규화기(regularizer)를 설계하였다.

[0091] 위에서 Binary Cross-Entropy Loss 박스의 막대들은 일반적으로 머신러닝에서 쓰이는 후술되는 식(6)으로 정의되는 손실함수일 수 있다. 본 실시예는 Binary Cross-Entropy Loss 에 후술되는 식(7)로 표현되는 확률 진폭 정규화기(Probability Amplitude Regularizer)를 합해서 최종 손실함수(식 (8) 참조)를 만들어서 학습시키는 방법을 채택하는 특수한 방법을 사용한다.

[0092] 본 실시예에서는 도 7에서 미사용 클래스 인덱스에 대한 확률에 대한 이진 교차 엔트로피 함수 L_{BCE}와 확률 진폭 정규화기 L_{PAR}로 목적 함수를 구성하고 다음 식(6) 및 식(7)과 같이 표현할 수 있다.

[0093] $$\mathcal{L}_{\text{BCE}}(\Theta;\mathbf{X}) = -\sum_{n=1}^{|\boldsymbol{y}|}[y_c \log p_c + (1-y_c)\log(1-p_c)], \quad (6)$$

$$\mathcal{L}_{\text{PAR}}(\Theta;\mathbf{X}) = -\sum\nolimits_{n'>|\boldsymbol{y}|}^{2^q} \log(1-p_{n'}), \qquad (7)$$

[0094] 여기서, $q \geq \lceil \log_2(|\boldsymbol{y}|) \rceil$ 이다. 따라서, 다음 식(8)과 같이 한 단계 단일 업데이트(one-step single

update)를 위한 훈련 손실함수를 최종적으로 정의할 수 있다.

[0095] $$\mathcal{L}(\Theta;\zeta)=\frac{1}{|\zeta|}\sum_{(\mathbf{X},\boldsymbol{y})\in\zeta}[\mathcal{L}_{\mathrm{BCE}}(\Theta;\mathbf{X})+\mathcal{L}_{\mathrm{PAR}}(\Theta;\mathbf{X})]. \quad (8)$$

[0096] 다음으로 훈련손실함수의 기울기는 다음과 같이 구할 수 있다.

[0097] 양자컴퓨터는 고전적인 훈련 방법(예: 체인 규칙을 사용한 역전파)을 사용할 수 없기 때문에 0차 최적화는 매개변수 이동 규칙이라고 하는 기울기를 추정하는 데 사용된다.

[0098] 손실 L의 대칭 차분 지수를 계산하여 손실 기울기는 다음 식(9)와 같이 구할 수 있다.

[0099] $$\frac{\partial\mathcal{L}(\Theta)}{\partial\theta_m}=\mathcal{L}\left(\Theta+\frac{\pi}{2}\mathbf{e}_m\right)-\mathcal{L}\left(\Theta-\frac{\pi}{2}\mathbf{e}_m\right), \quad (9)$$

[0100] 여기서 e_m은 Θm을 제외한 모든 성분을 제거하는 원-핫 벡터, 즉 $\Theta\cdot\mathbf{e}_m=\theta_m$ 및 $\forall m\in\mathbb{N}[1,|\Theta|]$이다. 마지막으로 본 실시예의 QML 프레임워크의 훈련 가능한 매개변수는 $\Theta\leftarrow\Theta-\eta\nabla_\Theta\mathcal{L}(\Theta)$로 업데이트된다. 여기서 η는 학습 속도이다.

[0101] 도 2에 제시된 프레임워크에서 처음부터 끝까지 동시에 학습하기 때문에, QNN뿐만 아니라 그 앞의 QCNN의 파라미터도 같이 학습하게 된다. 따라서, 확률 진폭 정규화기도 도 2의 전체에 대해 적용되는 것이다.

[0102] 도 8은 도 2에 제시된 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법에서 성능 테스트에 사용된 데이터 세트를 설명하기 위한 표이다. 도 8을 참조하면, 다양한 데이터 세트로 본 실시예의 QML 프레임워크의 훈련 가능성(trainability)과, 성능 평가를 통한 확률 진폭 정규화의 영향을 조사하였다. 이를 위해 본 실시예는 훈련 정확도, 테스트 정확도, SOTA(state-of-the-art)와의 비교 및 정규화기(regularizer)의 제거를 벤치마킹한다.

[0103] 테스트에서 커널에 3개의 채널이 있고, stride s = 3과 함께 4 X 4 커널이 일반적으로 사용되는 이미지 처리를 위해 두 개의 QCNN 모델을 사용하였다. 본 실시예의 QML 프레임워크는 모든 양자회로(예: QCNN 및 QNN)에 대해 4큐빗 및 제어된 단일 게이트를 활용하였다. 본 실시예에서 사용한 하이퍼파라미터는 Adam 옵티마이저, 초기 학습률은 8 X 10^{-3}, 배치 크기는 1,024이다. 본 실시예의 QML 프레임워크에 대한 모든 실험은 Python v3.8.10 및 torchquantum을 사용하는 클래식 컴퓨터에서 수행하였다.

[0104] 도 9는 도 8에 제시된 데이터 세트 중 MNIST 데이터 세트의 학습 곡선을 보여주는 도면이다. 도 10은 도 8에 제시된 데이터 세트 중 FashionMNIST 데이터 세트의 학습 곡선을 보여주는 도면이다. 도 11는 도 8에 제시된 데이터 세트 중 CIFAR10 데이터 세트의 학습 곡선을 보여주는 도면이다. 도 12는 도 8에 제시된 데이터 세트 중 EMNIST-letters 데이터 세트의 학습 곡선을 보여주는 도면이다.

[0105] 도 8에 제시된 MNIST, FashionMNIST, CIFAR10, EMNIST-letters 데이터셋을 사용하여 도 9 내지 도 12에 제시된 바와 같이, 본 실시예의 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법의 최종 성능 테스트를 수행하였다. 지금까지 개시된 기술중에서 이 테이터셋 전체를 다 활용해서 학습시킨 경우는 없었으나, 본 실시예는 이러한 학습에 의한 성능 테스트를 수행한 점에서 특별한 의의를 가진다. 도 9 내지 도 12의 그래프들을 검토하여 수치적 결과를 통해 수렴이나, 도메인겝이나, 학습성능도 매우 좋게 나오는 것도 확인할 수 있었다.

[0106] 도 9 내지 도 12는 다양한 테이터 세트의 학습 곡선을 보여준다.

[0107] 도 9에서 테스트 정확도는 6.25%에서 74.2%로 증가하고 훈련세트(trainset)와 테스트세트(testset) 사이의 도메인 간격은 상위 1개 정확도 사이에서 5.7%로 나타남을 확인할 수 있다. 도 10에서 테스트 정확도는 6.25%에서 73.8%로 증가하며 6.8%의 차이가 나는 것을 확인할 수 있다. 도 11 및 도 12에서 CIFAR10 및 EMNIST-letters에 대한 top-1 정확도는 각각 34.4% 및 33.1%를 나타내며 trainset과 testset 사이의 도메인 간격은 MNIST 및 FashionMNIST보다 낮은 것을 확인할 수 있다.

[0108] 도 13은 도 2에 제시된 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법과 QuantumNAS의 성능비교를 나타내는 도면이다.

[0109] 도 13을 참조하면, 본 실시예와 SOTA와 성능비교를 해보았다. SOTA는 도 13에 도시된 바와 같이, QuantumNAS로

서, 전술된 POVM 방식을 활용한 방식으로서 도 13에서 가장 좌측의 a 막대로 표시되고, 그 바로 우측의 막대 b가 본 실시예의 성능 테스트 결과를 나타낸다. 또한, 그 우측에 표시된 다른 더 복잡한 데이터 세트에 대한 본 실시예의 테스트 결과들(c,d,e)도 SOTA 대비 더 높은 성능을 내는 것을 확인할 수 있었다.

[0110] 도 13에서 볼 수 있듯이, 본 실시예의 QML 프레임워크는 MNIST 데이터 세트(b)에서 상위 1위 정확도와 관련하여 QuantumNA에 대비하여 44.2% 성능향상이 달성됨을 확인할 수 있다. 또한 본 실시예의 QML 프레임워크 7큐빗 미만을 사용하는 FashionMNIST, CIFAR10, EMNIST-letters 와 같은 다양한 데이터 세트(c,d,e)로 벤치마킹되는 반면 QuantumNAS는 10큐빗을 사용한다. 요약하면 PVM은 다중클래스 분류를 가능하게 할 뿐만 아니라 SOTA보다 더 높은 정확도를 달성한다.

[0111] 도 14는 도 2에 제시된 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법과 확률 진폭 정규화기를 제거한 경우의 성능을 비교한 결과를 나타내는 도면이다.

[0112] 도 14를 참조하면, 도 7에서 설명된 확률 진폭 정규화기를 적용한 경우와 적용하지 않은 경우를 CIFER10, EMNIST-letters라는 데이터셋을 활용해서 학습을 시켰다. 그 결과 도 14에 도시된 바(y 및 z 참조)와 같이 그 성능 차이가 EMNIST-letters 경우에는 최대 18% 성능 향상을 얻을 수 있었다.

[0113] 도 14에서 볼 수 있듯이 L_{PAR}을 사용한 훈련 체계의 top-1 정확도는 CIFAR10의 경우(v) 정규화기를 사용하지 않은 경우(x)에 비하여 7.5%, EMNIST-letters의 경우(y)는 정규화기를 사용하지 않은 경우(z)에 비하여 17.8%를 능가함을 확인할 수 있다. 이러한 결과는 확률 진폭 정규화기를 통해 다른 사용된 레이블의 예측이 감소하기 때문이다. 따라서 본 실시예의 QML 프레임워크에서 특별히 사용된 정규화기(regularizer)는 다중클래스 분류를 위한 QML 프레임워크에서 중요한 효과를 발휘함을 확인할 수 있다.

[0114] 본 실시예의 QML 프레임워크에 의하면, 먼저 다중클래스 분류를 위한 새로운 PVM 기반 QNN을 제공하며, 또한 기존의 NN 방법론을 사용하지 않는다.

[0115] 본 실시예의 QML 프레임워크에 의하면, 각각 최대 32 X 32 X 3의 입력과 최대 26개의 클래스 출력과 관련하여 고전적인 NN 방법론 없이 QML의 확장성을 개선한다.

[0116] 본 발명의 실시예에 의하면, 양자회로에서 큐빗의 개수가 많아질수록 양자오류가 심해지는 큐빗 개수 제한의 문제점을 PVM과 POVM을 활용하여 해결함으로써 가장 최신 기술(State-of-the-art) 대비 약 40% 이상의 정확도 향상을 더 적은 큐빗을 사용하여 달성하는 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 제공할 수 있다.

[0117] 본 발명의 실시예에 의하면, 입력차원과 출력차원이 $\text{output} \leq 2^{\text{input}}$ 관계를 가지는 PVM 방법을 이용하여 학습시킴으로써 큐빗을 종래기법 대비 최대 [log2(q)]으로 줄일 수 있고, 학습성능도 높은 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 제공할 수 있다.

[0118] 본 발명의 실시예에 의하면, 확률 진폭 정규화기(probability amplitude regularizer)를 도입하여 학습성 향상과 예측 성능 향상을 달성하는 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법을 제공할 수 있다.

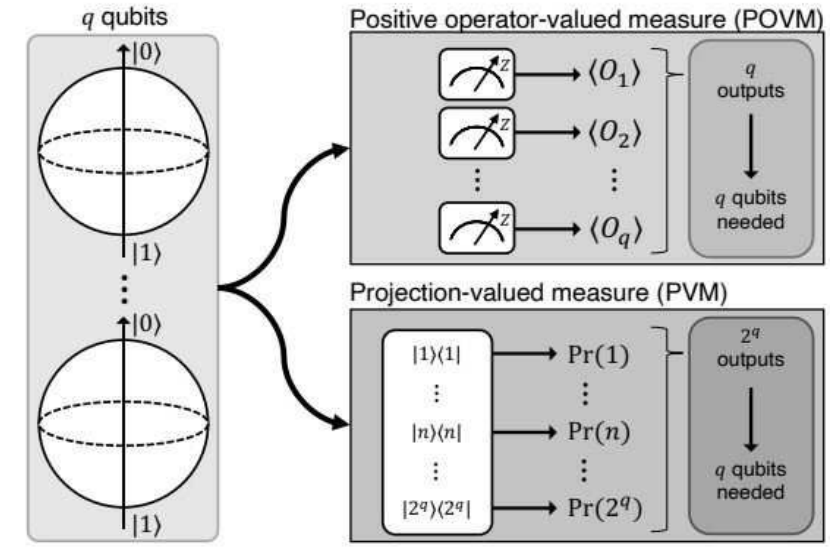
[0119] 본 발명의 실시예에 의하면, 다중클래스 분류를 위한 투영가치측정 기반 양자머신러닝 방법은 모든 양자머신러닝에 대해 활용 가능하다.

[0120] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.

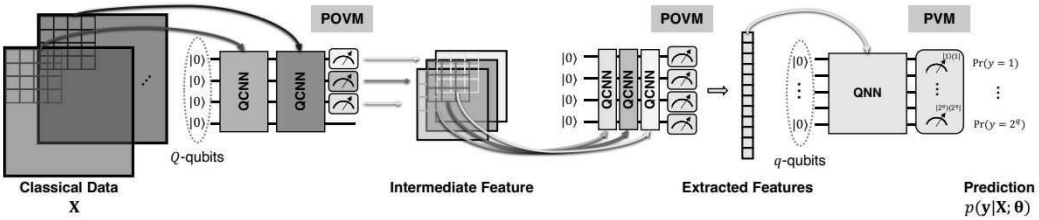
[0121] 본 발명의 범위는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

도면

도면1


q qubits
|0⟩
|1⟩
|0⟩
|1⟩
Positive operator-valued measure (POVM)
⟨O1⟩
⟨O2⟩
⟨Oq⟩
q outputs
q qubits needed
Projection-valued measure (PVM)
|1⟩⟨1|
|n⟩⟨n|
|2q⟩⟨2q|
Pr(1)
Pr(n)
Pr(2q)
2q outputs
q qubits needed

도면2

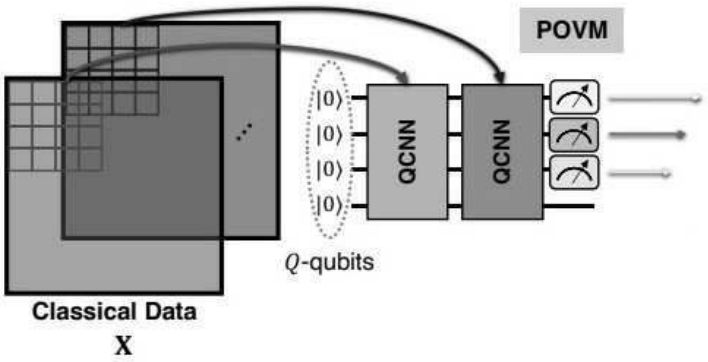

POVM
QCNN
QCNN
Q-qubits
Classical Data
x
Intermediate Feature
POVM
QCNN
QCNN
QCNN
Extracted Features
PVM
QNN
q-qubits
Pr(y = 1)
Pr(y = 2q)
Prediction
p(y|X; θ)

도면3

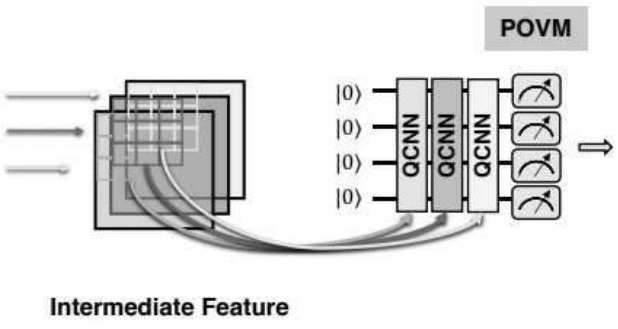
Algorithm 1: QML Framework for Multi-Class Classification

Notation. Processing data $\breve{\mathbf{X}}$, Processed data $\tilde{\mathbf{X}}$, stride s, kernel size κ, and the number of QNN depths D;
Image Processing($\mathbf{X}$):
  $\breve{\mathbf{X}} \leftarrow \mathbf{X}$;
  for $w \in \{0, s, 2s, \cdots, W-s\}$ **do**
    for $h \in \{0, s, 2s, \cdots, H-s\}$ **do**
      Initialize quantum state, $|\psi\rangle \leftarrow |0\rangle$;
      for $c \in \{1, 2, \cdots, c_{in}\}$ **do**
        Prepare data, $\breve{\mathbf{x}} \leftarrow \breve{\mathbf{X}}_{w:w+\kappa, h:h+\kappa, c}$;
        Upload data, $|\psi\rangle \leftarrow U(\boldsymbol{\theta}_c) \cdot U(\breve{\mathbf{x}}) \cdot |\psi\rangle$;
      for $c \in \{1, 2, \cdots, c_{out}\}$ **do**
        Measure through POVM and collects outputs, $\tilde{\mathbf{X}}_{\frac{w}{s}, \frac{h}{s}, c} \leftarrow \langle O_c \rangle_{\breve{\mathbf{x}}, \Theta}$;
  Get intermediate feature, $\breve{\mathbf{X}} \leftarrow \tilde{\mathbf{X}}$;
  Output: Extracted feature $\hat{\mathbf{X}}$;
Prediction($\hat{\mathbf{X}}$):
  Initialize $\mathbf{p} \leftarrow \emptyset$;
  Prepare the pure density matrices, $\{|1\rangle\langle 1|, \cdots, |n\rangle\langle n|, \cdots |2^q\rangle\langle 2^q|\}$;
  for $m \in \{1, \cdots, 2^q\}$ **do**
    Obtain prediction value of class c;
    $\mathbf{p} \leftarrow \mathbf{p} \cup p_c$;
  Output: Prediction $\mathbf{p}$;

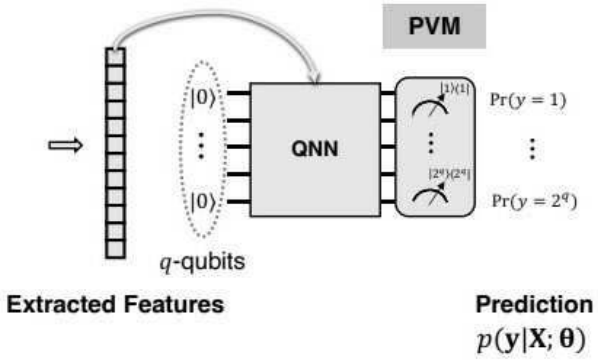
도면4



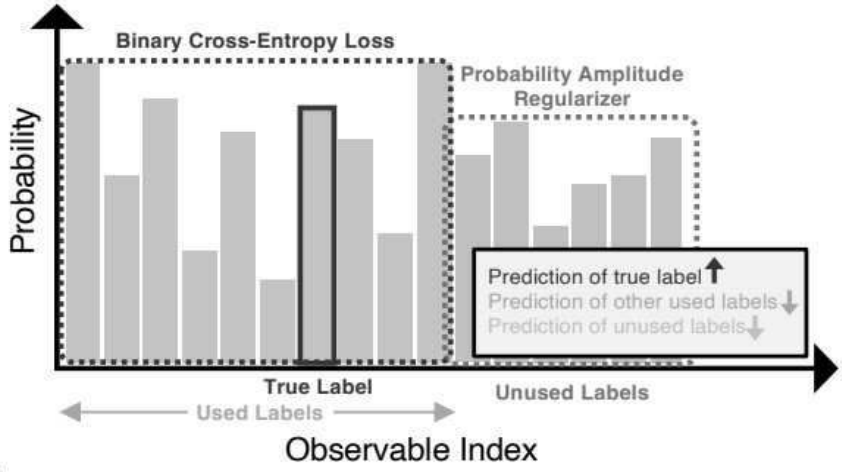
도면5


POVM
|0⟩
|0⟩
|0⟩
|0⟩
QCNN
QCNN
QCNN
Intermediate Feature

도면6


PVM
|0⟩
|0⟩
QNN
Pr(y = 1)
Pr(y = 2^q)
q-qubits
Extracted Features
Prediction
p(y|X; θ)

도면7

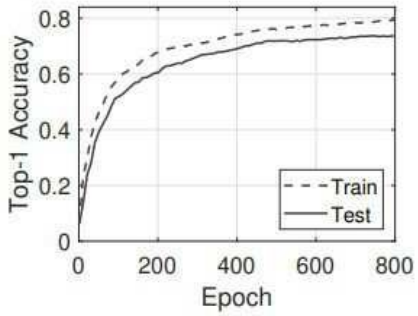

Binary Cross-Entropy Loss
Probability Amplitude Regularizer
Probability
Prediction of true label
Prediction of other used labels
Prediction of unused labels
True Label
Unused Labels
Used Labels
Observable Index

도면8

TABLE I
EXPERIMENTAL SETTINGS

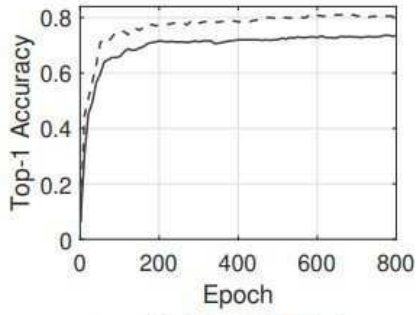
Dataset	Input size	# of classes	# of qubits (q)
MNIST	$28 \times 28 \times 1$	10	4
FashionMNIST	$28 \times 28 \times 1$	10	4
CIFAR10	$32 \times 32 \times 3$	10	4
EMNIST-letters	$28 \times 28 \times 1$	26	6

도면9



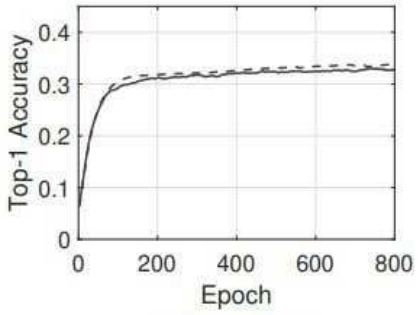
MNIST

도면10



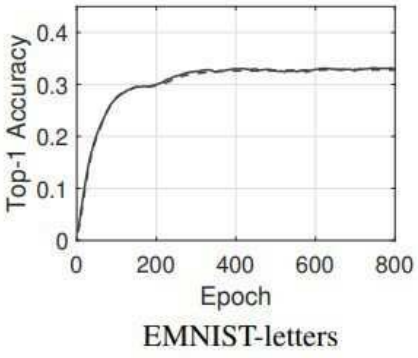
FashionMNIST

도면11

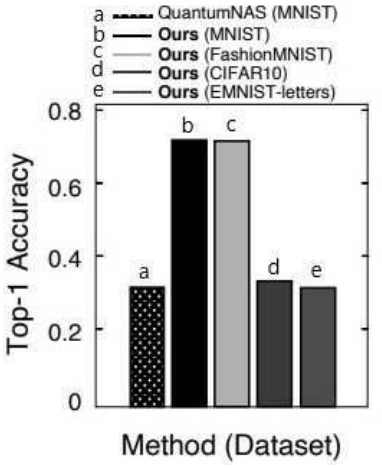


CIFAR10

도면12



도면13



도면14

