(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0058981
(43) 공개일자 2025년05월02일

(51) 국제특허분류(Int. Cl.)
G06N 10/60 (2022.01) *G06N 10/20* (2022.01)
G06N 3/0464 (2023.01) *G06N 3/09* (2023.01)
B82Y 10/00 (2017.01)
(52) CPC특허분류
G06N 10/60 (2022.01)
G06N 10/20 (2022.01)
(21) 출원번호 10-2023-0142781
(22) 출원일자 2023년10월24일
심사청구일자 2023년10월24일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
백한결
서울특별시 도봉구 도봉로136가길 73 (창동, 북한산한신휴플러스아파트) 101동 1304호
박수현
인천광역시 연수구 송도문화로28번길 28 (송도동, 송도글로벌캠퍼스푸르지오) 103동 1402호
김중헌
서울특별시 동작구 상도로 346-1 (상도동, 힐스테이트 상도 센트럴파크) 212동201호
(74) 대리인
윤귀상

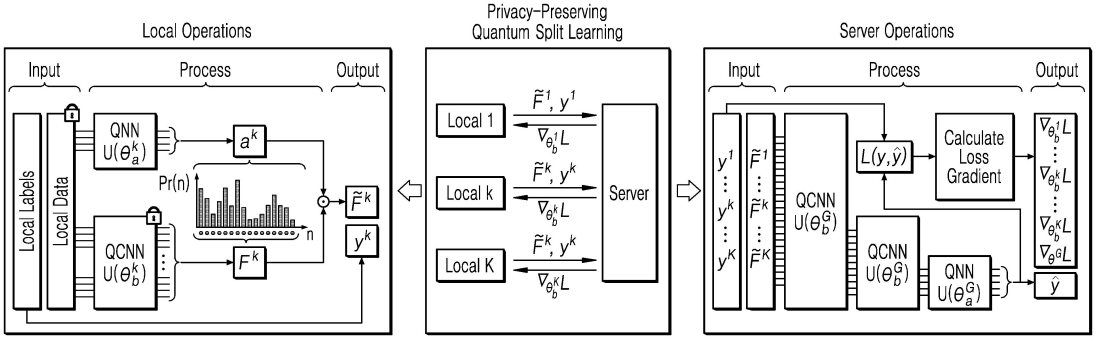
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 **양자 스플릿 학습 시스템 및 방법**

(57) 요 약

본 발명은 양자 스플릿 학습 시스템에 대한 것으로, 로컬 데이터 및 상기 로컬 데이터의 분류 정보를 포함하는 레이블(label) 데이터에 양자 회로를 적용하여 특징 정보를 추출하고, 추출된 특징 정보 및 상기 레이블 데이터를 출력하는 적어도 하나의 로컬 디바이스; 및 상기 적어도 하나의 로컬 디바이스로부터 수신된 모든 특징 정보 및 레이블 데이터에 양자 회로를 적용하되 PVM(positive operator-valued measure)을 기반으로 상기 로컬 데이터의 분류에 대한 예측 결과를 출력하는 서버;를 포함한다.

대 표 도



(52) CPC특허분류
G06N 3/0464 (2023.01)
G06N 3/09 (2023.01)
B82Y 10/00 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711182694
과제번호	2022R1A2C200486912
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)
연구과제명	Quantum Hyper-Driving: 양자 컴퓨팅 기반 초연결 고차원-센싱 자율이동체 기술
기 여 율	1/1
과제수행기관명	고려대학교
연구기간	2022.03.01 ~ 2025.02.28

명 세 서

청구범위

청구항 1

로컬 데이터 및 상기 로컬 데이터의 분류 정보를 포함하는 레이블(label) 데이터에 양자 회로를 적용하여 특징 정보를 추출하고, 추출된 특징 정보 및 상기 레이블 데이터를 출력하는 적어도 하나의 로컬 디바이스; 및

상기 적어도 하나의 로컬 디바이스로부터 수신된 모든 특징 정보 및 레이블 데이터에 양자 회로를 적용하되 PVM(positive operator-valued measure)을 기반으로 상기 로컬 데이터의 분류에 대한 예측 결과를 출력하는 서버;를 포함하는, 양자 스플릿 학습 시스템.

청구항 2

제1항에 있어서,

상기 적어도 하나의 로컬 디바이스는 상기 로컬 데이터에 QCNN(Quantum Convolution Neural Network) 양자 회로를 적용하여 이미지 특징 정보를 추출하고, 상기 레이블 데이터에 QNN(Quantum Neural Network) 양자 회로를 적용하여 레이블 특징 정보를 추출하는, 양자 스플릿 학습 시스템.

청구항 3

제2항에 있어서,

상기 특징 정보는 상기 이미지 특징 정보 및 레이블 특징 정보가 교차-채널 풀링(Cross-Channel Pooling)을 통해 조합된 것인, 양자 스플릿 학습 시스템.

청구항 4

제1항에 있어서,

상기 서버는 상기 모든 특징 정보에 QCNN(Quantum Convolution Neural Network)/QNN(Quantum Neural Network) 양자 회로를 적용하여 상기 예측 결과를 출력하는, 양자 스플릿 학습 시스템.

청구항 5

제1항에 있어서,

상기 PVM의 목적 함수(objective function)는 하기 식(1)로 정의되며,

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

상기 식(1)에서 Θ는 훈련 가능한 매개변수, ζ는 미니 배치, X는 상기 로컬 데이터, y는 상기 레이블 데이터, L_{BCE}는 이진 교차 엔트로피(Binary Cross Entropy), L_{PAR}은 정규화기(regularizer)를 나타내는 것을 특징으로 하는, 양자 스플릿 학습 시스템.

청구항 6

제5항에 있어서,

상기 이진 교차 엔트로피 L_{BCE}는 하기 식(2)로 정의되며,

$$\mathcal{L}_{\mathrm{BCE}}(\Theta;\mathrm{X}) = -\sum_{n=1}^{|y|}[y_n \log p_n + (1-y_n)\log(1-p_n)] \quad \ldots(2)$$

상기 식(2)에서 |y|는 분류하고자 하는 작업의 클래스 개수, p_n은 클래스 n에 대한 예측 값을 나타내는 것을 특징으로 하는, 양자 스플릿 학습 시스템.

청구항 7

제6항에 있어서,

상기 정규화기 L_{PAR}는 하기 식(3)으로 정의되며,

$$\mathcal{L}_{\mathrm{PAR}}(\Theta;\mathrm{X}) = -\sum_{n'>|y|}^{2^q} \log(1-p_{n'}) \quad \ldots(3)$$

상기 식(3)에서 $q \geq \lceil \log_2(|y|) \rceil$이고, 2^q는 QNN의 출력, $P_{n'}$는 n'번째 투영 행렬을 나타냄을 특징으로 하는, 양자 스플릿 학습 시스템.

청구항 8

적어도 하나의 로컬 디바이스와 서버를 포함하는 양자 스플릿 학습 시스템의 양자 스플릿 학습 방법에 있어서,

상기 적어도 하나의 로컬 디바이스가 로컬 데이터 및 상기 로컬 데이터의 분류 정보를 포함하는 레이블(label) 데이터에 양자 회로를 적용하여 특징 정보를 추출하는 단계;

상기 적어도 하나의 로컬 디바이스가 상기 추출된 특징 정보 및 레이블 데이터를 출력하는 단계;

상기 서버가 상기 적어도 하나의 로컬 디바이스로부터 수신된 모든 특징 정보 및 레이블 데이터에 양자 회로를 적용하는 단계; 및

상기 서버가 PVM(positive operator-valued measure)을 기반으로 상기 로컬 데이터의 분류에 대한 예측 결과를 출력하는 단계;를 포함하는, 양자 스플릿 학습 방법.

청구항 9

제8항에 있어서,

상기 특징 정보를 추출하는 단계는,

상기 로컬 데이터에 QCNN(Quantum Convolution Neural Network) 양자 회로를 적용하여 이미지 특징 정보를 추출하는 단계; 및

상기 레이블 데이터에 QNN(Quantum Neural Network) 양자 회로를 적용하여 레이블 특징 정보를 추출하는 단계를 포함하는, 양자 스플릿 학습 방법.

청구항 10

제9항에 있어서,

상기 특징 정보는 상기 이미지 특징 정보 및 레이블 특징 정보가 교차-채널 풀링(Cross-Channel Pooling)을 통해 조합된 것인, 양자 스플릿 학습 방법.

청구항 11

제8항에 있어서,

상기 양자 회로를 적용하는 단계는, 상기 모든 특징 정보에 QCNN(Quantum Convolution Neural Network)/QNN(Quantum Neural Network) 양자 회로를 적용하는 것인, 양자 스플릿 학습 방법.

청구항 12

제8항에 있어서,

상기 PVM의 목적 함수(objective function)는 하기 식(4)로 정의되며,

$$\mathcal{L}(\Theta;\zeta) = \frac{1}{|\zeta|}\sum (\mathbf{X},\boldsymbol{y}) \in \zeta[\mathcal{L}_{\mathrm{BCE}}(\Theta;\mathbf{X}) + \mathcal{L}_{\mathrm{PAR}}(\Theta;\mathbf{X})] \quad \ldots(4)$$

상기 식(4)에서 Θ는 훈련 가능한 매개변수, ζ는 미니 배치, X는 상기 로컬 데이터, y는 상기 레이블 데이터, L_{BCE}는 이진 교차 엔트로피(Binary Cross Entropy), L_{PAR}은 정규화기(regularizer)를 나타내는 것을 특징으로 하는, 양자 스플릿 학습 방법.

청구항 13

제12항에 있어서,

상기 이진 교차 엔트로피 L_{BCE}는 하기 식(5)로 정의되며,

$$\mathcal{L}_{\mathrm{BCE}}(\Theta;\mathbf{X}) = -\sum_{n=1}^{|\boldsymbol{y}|}[y_n \log p_n + (1-y_n)\log(1-p_n)] \quad \ldots(5)$$

상기 식(5)에서 |y|는 분류하고자 하는 작업의 클래스 개수, p_n은 클래스 n에 대한 예측 값을 나타내는 것을 특징으로 하는, 양자 스플릿 학습 방법.

청구항 14

제13항에 있어서,

상기 정규화기 L_{PAR}는 하기 식(6)으로 정의되며,

$$\mathcal{L}_{\mathrm{PAR}}(\Theta;\mathbf{X}) = -\sum_{n'>|\boldsymbol{y}|}^{2^q} \log(1-p_{n'}) \quad \ldots(6)$$

상기 식(6)에서 $q \geq \lceil \log_2(|\boldsymbol{y}|) \rceil$ 이고, 2^q는 QNN의 출력, $P_{n'}$는 n'번째 투영 행렬을 나타냄을 특징으로 하는, 양자 스플릿 학습 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 양자 스플릿 학습 시스템 및 방법에 관한 것이다.

배 경 기 술

[0003] 양자 컴퓨팅 기반 양자 신경망(QNN: Quantum Neural Network)은 계산-가속 및 병렬화의 가능성으로 인해 최근 주목을 받고 있다. 양자 컴퓨팅은 양자 비트(큐빗) 상태를 측정(measurement)함으로써 정보(1 또는 0)을 얻는 방법으로서, 기존의 고전(classical) 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 양자 머신 러닝(QML: Quantum Machine Learning)이 활발히 사용되고 있으며, 양자 머신 러닝의 혁신적인 발전을 바탕으로 이론과 적용에 대한 QNN 연구 또한 활발히 진행되고 있다.

[0005] 한편, 스플릿 학습(Split Learning)은 개인정보를 보호하면서 기계 학습 모델을 학습하거나 사용하는 방법 중 하나로써, 머신 러닝 모델을 여러 부분으로 나누고 데이터의 일부만 각 부분에서 처리하는 방식으로 작동한다. 이로써 개별 데이터 포인트 또는 정보에 대한 고유한 접속을 방지하고, 개인정보 보호 및 보안을 강화할 수 있는 이점이 있다.

[0006] 따라서 이러한 스플릿 학습의 개념을 양자 컴퓨팅 기술에 적용한다면, 양자 컴퓨팅 기술을 사용하여 스플릿 학습을 개선하고, 머신 러닝 모델을 개인 정보를 보호하면서 학습하는데 활용하여 보안 및 효율성 측면을 향상시킬 수 있을 것으로 예상되지만, 현재로서는 이러한 양자 컴퓨팅 기술과 스플릿 학습의 개념을 결합하려는 연구는 드문 상황이다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 한국등록특허공보 제10-2175977호

발명의 내용

해결하려는 과제

[0009] 본 발명은 상기와 같은 문제를 해결하기 위해 안출된 것으로, 본 발명의 목적은 양자 스플릿 학습 시스템 및 방법을 제공하는 것이다.

과제의 해결 수단

[0011] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 양자 스플릿 학습 시스템은, 로컬 데이터 및 상기 로컬 데이터의 분류 정보를 포함하는 레이블(label) 데이터에 양자 회로를 적용하여 특징 정보를 추출하고, 추출된 특징 정보 및 상기 레이블 데이터를 출력하는 적어도 하나의 로컬 디바이스; 및 상기 적어도 하나의 로컬 디바이스로부터 수신된 모든 특징 정보 및 레이블 데이터에 양자 회로를 적용하되 PVM(positive operator-valued measure)을 기반으로 상기 로컬 데이터의 분류에 대한 예측 결과를 출력하는 서버;를 포함한다.

[0012] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 양자 스플릿 학습 방법은, 적어도 하나의 로컬 디바이스가 로컬 데이터 및 상기 로컬 데이터의 분류 정보를 포함하는 레이블(label) 데이터에 양자 회로를 적용하여 특징 정보를 추출하는 단계; 상기 적어도 하나의 로컬 디바이스가 상기 추출된 특징 정보 및 레이블 데이터를 출력하는 단계; 서버가 상기 적어도 하나의 로컬 디바이스로부터 수신된 모든 특징 정보 및 레이블 데이터에 양자 회로를 적용하는 단계; 및 상기 서버가 PVM(positive operator-valued measure)을 기반으로 상기 로컬 데이터의 분류에 대한 예측 결과를 출력하는 단계;를 포함한다.

발명의 효과

[0014] 상술한 본 발명의 일측면에 따르면, 양자 스플릿 학습 시스템 및 방법을 제공함으로써, 통신 비용을 절감할 수 있고, 양자 스플릿 학습 시스템을 구성하는 로컬 디바이스의 수에 대한 확장성을 제공할 수 있으며, 아울러 분류 작업에 있어서 개인정보(privacy) 보호를 달성할 수 있다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 실시예에 따른 양자 신경망 아키텍처를 나타낸 도면,

도 2는 본 발명의 실시예에 따른 양자 스플릿 학습 시스템의 전체 프레임워크를 나타낸 도면,

도 3은 도 2에서 로컬 디바이스의 특징 정보 및 레이블 데이터 출력을 나타낸 도면,

도 4는 도 2에서 서버의 분류에 대한 예측 결과 출력을 나타낸 도면,

도 5는 도 2에 제시된 양자 스플릿 학습 시스템의 학습 방법을 나타낸 알고리즘의 예를 나타낸 도면,

도 6은 본 발명의 실시예에 따른 양자 스플릿 학습 방법을 나타낸 순서도,

도 7은 본 발명의 실시예에 따른 양자 스플릿 학습 방법의 성능 테스트에 사용된 다양한 데이터 세트의 학습 곡선을 나타낸 그래프,

그리고, 도 8은 MNIST 데이터 세트에서 로컬 디바이스 개수에 따른 성능 비교를 나타낸 그래프.

발명을 실시하기 위한 구체적인 내용

[0017] 후술하는 본 발명에 대한 상세한 설명은, 본 발명이 실시될 수 있는 특정 실시예를 예시로서 도시하는 첨부 도면을 참조한다. 이들 실시예는 당업자가 본 발명을 실시할 수 있기에 충분하도록 상세히 설명된다. 본 발명의 다양한 실시예는 서로 다르지만 상호 배타적일 필요는 없음이 이해되어야 한다. 예를 들어, 여기에 기재되어 있는 특정 형상, 구조 및 특성은 일 실시예와 관련하여 본 발명의 정신 및 범위를 벗어나지 않으면서 다른 실시예로 구현될 수 있다. 또한, 각각의 개시된 실시예 내의 개별 구성요소의 위치 또는 배치는 본 발명의 정신 및 범위를 벗어나지 않으면서 변경될 수 있음이 이해되어야 한다. 따라서, 후술하는 상세한 설명은 한정적인 의미로서 취하려는 것이 아니며, 본 발명의 범위는, 적절하게 설명된다면, 그 청구항들이 주장하는 것과 균등한 모든 범위와 더불어 첨부된 청구항에 의해서만 한정된다. 도면에서 유사한 참조부호는 여러 측면에 걸쳐서 동일하거나 유사한 기능을 지칭한다.

[0018] 본 발명에 따른 구성요소들은 물리적인 구분이 아니라 기능적인 구분에 의해서 정의되는 구성요소들로써 각각이 수행하는 기능들에 의해서 정의될 수 있다. 각각의 구성요소들은 하드웨어 또는 각각의 기능을 수행하는 프로그램 코드 및 프로세싱 유닛으로 구현될 수 있을 것이며, 두 개 이상의 구성요소의 기능이 하나의 구성요소에 포함되어 구현될 수도 있을 것이다. 따라서 이하의 실시예에서 구성요소에 부여되는 명칭은 각각의 구성요소를 물리적으로 구분하기 위한 것이 아니라 각각의 구성요소가 수행되는 대표적인 기능을 암시하기 위해서 부여된 것이며, 구성요소의 명칭에 의해서 본 발명의 기술적 사상이 한정되지 않는 것임에 유의하여야 한다.

[0019] 본 발명의 실시예를 설명하기에 앞서 양자 게이트와 연산(operations) 개념에 대하여 간단히 설명하도록 한다.

[0020] 단일 큐빗 시스템에서 양자 상태는 두 개의 기저(bases) $|0\rangle = [1\ 0]^T$, $|0\rangle = [1\ 0]^T$ 로 정의될 수 있으며, 힐베르트 공간(Hilbert space) H에서의 확률 진폭(probability amplitude)은 $|\Psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle$ 와 같이 정의될 수 있다. 여기서 $\sum_{i=0}^{1}|\alpha_i|^2 = 1$ 이다.

[0021] 확장된 큐빗 시스템에서 Q-큐빗 양자 상태는 하기 수학식 1과 같이 정의된다.

수학식 1

[0022] $$|\psi\rangle = \sum_{n=0}^{2^Q-1} \alpha_n |n\rangle$$

[0023] 여기서 α_n과 $|n\rangle$ 은 각각 힐베르트 공간에서 확률 진폭과 n번째 표준 기저(standard basis)를 나타낸다.

[0024] 힐베르트 공간의 정의 및 기본적인 수학적 구조(즉, $\mathcal{H}^{\otimes Q} \equiv \mathbb{C}^{2^Q}$)에 따라 확률 진폭 $\forall \alpha_n \in C$는 다음과 같

은 수식을 만족한다. 즉, $\sum_{n=0}^{2Q-1} |\alpha_n|^2 = 1$ 이다.

[0025] 회전 게이트(rotation gate) 및 제어 게이트(controlled gate)는 이러한 확률 진폭을 선형(linear) 변환하며, 이는 하기 수학식 2와 같이 나타낼 수 있다.

수학식 2

[0026]
$$R_\Gamma(\theta) = \exp\left(-i\frac{\theta}{2}\Gamma\right), C(\Gamma) = \begin{bmatrix} I & 0 \\ 0 & \Gamma \end{bmatrix}, \forall\Gamma \in \{X, Y, Z\}$$

[0027] 여기서 $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, $Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ 이고, X,Y,Z는 각각 Pauli-X,Y,Z 게이트를 나타낸다. 이러한 기본 게이트를 활용함으로써 양자 신경망 아키텍처의 핵심 구성요소인 범용 단위(unitary) 게이트 U(Θ)를 나타낼 수 있다.

[0028] 이하에서는 도면들을 참조하여 본 발명의 바람직한 실시예들을 보다 상세하게 설명하기로 한다.

[0029] 도 1은 본 발명의 실시예에 따른 양자 신경망 아키텍처를 나타낸 도면이다.

[0030] 도시된 양자 신경망(QNN) 회로는 상태 인코딩부(State Encoder)(110), 선형 변환부(Linear Transform)(120) 및 측정부(Measurement)(130)로 구성된다.

[0031] 먼저 상태 인코딩부(110) 동작을 살펴보면, NISQ(Noisy Intermediate Scale Quantum) 시대의 한계로 양자 회로에 고전적 데이터를 업로드하기 위해서는 기저 인코딩, 진폭 인코딩, 각도(angle) 인코딩과 같은 상태 인코딩 과정이 요구된다.

[0032] 본 발명의 실시예에 따른 상태 인코딩부(110)는 데이터 재 업로드를 위해 각도 인코딩을 적용하며, 고전적 데이터 $x \in R^{|x|}$라 할 때, 데이터 업로드는 하기 수학식 3과 같이 나타낼 수 있다.

수학식 3

[0033]
$$|\psi_{enc}\rangle = \left(\prod_{l=0}^{\lceil |\mathbf{x}|/Q \rceil} U(\theta_l) U(\mathbf{x}_{Q\cdot l:Q\cdot(l+1)})\right) \cdot |\psi_0\rangle$$

[0034] 여기서 $|\Psi_0\rangle$는 초기 양자 상태, 즉 $|\psi_0\rangle = |0\rangle^{\otimes Q}$ 를 나타낸다.

[0035] 각도 인코딩 과정에서, 입력 데이터 x는 간격(interval) Q에 따라 $\mathbf{x}_{Q\cdot l:Q\cdot(l+1)}$로 분할되고, $\forall l \in [1, \lceil |\mathbf{x}|/Q \rceil]$ 이다. 여기서 $\mathbf{x}_{Q\cdot l:Q\cdot(l+1)}$는 첫번째 Q · I 에서 Q · (I+1) 개의 요소(element)에 의해 구성되는 입력 벡터를 나타내고, Θ_l는 훈련 가능한 매개변수를 나타낸다. 이때 $\forall \Theta_l \in \Theta$ 이다.

[0036] 다음으로 선형 변환부(120) 동작을 살펴보면, QNN은 양자 합성곱 신경망(QCNN: Quantum Convolution Neural Network)과 양자 완전 연결 망(QFCN: Quantum Fully Connected Network)일 수 있다.

[0037] 선형 변환부(120)는 양자 신경층(neural layer)에서 동작하며, 상기 양자 신경층에서 수학식 2에 나타낸 회전 게이트 및 제어 게이트는 일반적인 DNN에서 은닉 계층 선형 및 비선형(nonlinear) 계산을 위해 활용된다. 이러한 양자 신경망은 수학적으로 하기 수학식 4와 같이 나타낼 수 있다.

수학식 4

[0038] $$|\psi\rangle = U(\theta)|\psi_{enc}\rangle$$

[0039] 수학식 4에서 입력 | Ψ_{enc}〉은 수학식 3에 나타낸 바와 같다.

[0040] 마지막으로 측정부(130) 동작을 살펴보면, 고전적인 출력을 얻기 위해서는 양자 상태를 측정하도록 투영(project)시켜야 한다. 본 발명의 실시예에서는 다양한 측정 방식들 중 PVM(Projection-Valued Measure) 및 POVM(Positive Operator-Valued Measure) 방식을 통해 양자 상태를 측정하도록 한다.

[0041] PVM은 양자 상태를 투영기(projector) $\{|n\rangle\langle n|\}_{n=1}^{2^q}$에 투영하여 확률 측정을 수행하는 방식이며, 이때 투영기의 출력은 하기 수학식 5와 같이 나타낼 수 있다.

수학식 5

[0042] $$\Pr(y=n) = \langle\psi|n\rangle\langle n|\psi\rangle = |\langle\psi|n\rangle|^2 = |\alpha_n|^2$$

[0043] 이렇게 PVM 방식을 통한 측정은 양자 상태를 모든 가능한 기저에 투영하여 측정하는 방식이므로 완전한 측정(complete measurement)에 해당한다.

[0044] POVM은 모든 큐빗의 양자 상태를 Pauli-Z 측정 장치에 투영하여 확률 측정을 수행하는 방식이며, 여기서 Pauli-Z는 Z = diag(1, -1)이고, n번째 투영 행렬은 $\mathbf{P}_n \triangleq \mathbf{I}^{\otimes n-1} \otimes \mathbf{Z} \otimes \mathbf{I}^{\otimes Q-n}$ 으로 정의될 수 있다.

[0045] POVM 방식에는 투영(projection) P_n의 기대값을 $O_n \in R[-1, 1], (\forall n \in N[1, Q])$로 표시되는 확률 변수로 만들고, m번째 큐빗의 기대값은 〈O_n〉= 〈 Ψ | P_n | Ψ 〉로 정의된다.

[0046] 이렇게 POVM 방식을 통한 측정은 양자 상태를 부분 투영 행렬에 투영하여 측정하는 방식이므로 불완전한 측정(incomplete measurement)에 해당한다.

[0047] 도 2는 본 발명의 실시예에 따른 양자 스플릿 학습 시스템의 전체 프레임워크를 나타낸 도면이다.

[0048] 도시된 양자 스플릿 학습 시스템은 적어도 하나의 로컬 디바이스와 서버를 포함한다. 각 로컬 디바이스 k는 자신의 로컬 데이터 $Z_k \in Z$, 자신의 QCNN, 및 하나의 교차-채널 풀링 QNN을 가지며, 로컬 데이터와 로컬 QCNN은 다른 것들과 공유될 수 없는 반면, 특징과 레이블은 공유될 수 있다고 가정한다. 서버는 로컬 디바이스가 추출한 특징을 처리하므로, 최종적으로 대응하는 예측을 획득할 수 있다. 여기서 QCNN/QNN의 구조는 도 1을 통해 설명한 바와 같이 상태 인코딩부(110), 선형 변환부(120) 및 측정부(130)로 구성된다.

[0049] 후술할 본 발명의 실시예에서는 도 3 및 도 4를 통해 도 2에 제시된 로컬디바이스의 동작과 서버의 동작을 보다 상세히 설명하도록 한다.

[0050] 도 3은 도 2에서 로컬 디바이스의 특징 정보 및 레이블 데이터 출력을 나타낸 도면이고, 도 4는 도 2에서 서버의 분류에 대한 예측 결과 출력을 나타낸 도면이고, 도 5는 도 2에 제시된 양자 스플릿 학습 시스템의 학습 방법을 나타낸 알고리즘의 예를 나타낸 도면이다.

[0051] 먼저 도 3을 통해 로컬 디바이스 동작을 살펴보면, 로컬 디바이스는 로컬 데이터, 일례로 이미지 데이터에 QCNN 양자 회로를 적용하여 이미지 특징 정보 F^k를 추출하고, 상기 로컬 데이터의 분류 정보를 포함하는 로컬 레이블 데이터에 QNN 양자 회로를 적용하여 레이블 특징 정보 a^k를 추출한다. 그런 다음 로컬 디바이스는 이미지 특징 정보 F^k와 레이블 특징 정보 a^k를 교차-채널 풀링(Cross-Channel Pooling)을 통해 조합한 최종 특징 정보 $\tilde{F}^k$와

상기 레이블 데이터 y^k를 출력한다.

[0052] 보다 상세하게, 로컬 디바이스 k의 고전적 데이터는 $\mathbf{X}^k \in \mathbb{R}^{W\times H\times c_{in}}$ 와 같이 나타낼 수 있고, QCNN 필터는 이미지를 스트라이드 s에 따라 스캔한다. 여기서 필터 사이즈는 k이다.

[0053] 패치는 $\mathbf{x}^k \in \mathbb{R}^{\kappa\times\kappa\times c_{in}}$ 와 같이 표현될 수 있고, 여기서 c_{in}은 입력 채널의 개수를 나타낸다. 패치 x^k는 수학식 3의 데이터 재 업로드를 통해 QCNN에 업로드되고, QCNN은 패치 x^k로부터 특징 정보를 획득한다. 상기 특징 정보는 $\mathbf{f}^k \in \mathbb{R}^{c_{out}}$ 와 같이 나타낼 수 있으며, 여기서 c_{out}은 출력 채널의 개수를 나타낸다.

[0054] 확률 진폭은 QNN에 의해 구할 수 있고, 그 출력은 $a^k(x^k)\in \mathbb{C}^{c_{out}}$ 와 같이 나타낼 수 있으며, 여기서 $\|\alpha^k(\mathbf{x}^k)\|^2 = 1$ 와 같다.

[0055] 한편, 양자 기반 교차-채널 풀링은 f^k와 $a^k(x^k)$의 내부 곱(inner product)으로 계산되며, 따라서 로컬 디바이스가 출력하는 최종 특징 정보 $\tilde{f}^k$ 는 하기 수학식 6과 같이 나타낼 수 있다.

수학식 6

[0056] $$\tilde{f}^k = \mathbf{f}^k \cdot \alpha^k(\mathbf{x}^k)$$

[0057] 모든 패치를 스캔한 후 출력되는 최종 특징 정보는 다음과 같이 얻을 수 있다: $\tilde{\mathbf{F}}^k = \{\tilde{f}^k\} \in \mathbb{R}^{\hat{W}\times\hat{H}\times 1}$

[0058] 이후 로컬 디바이스 k는 자신의 특징 정보 및 레이블 데이터의 셋 ($\tilde{F}^k$,y^k)을 관련 서버로 전송하고, 이때 ($\tilde{F}^k$,y^k)은 양자 회로를 통해 처리되더라도 고전적 데이터에 해당한다.

[0059] 다음으로 도 4를 통해 서버 동작을 살펴보면, 적어도 하나의 로컬 디바이스, 즉 K개의 로컬 디바이스로부터 수신된 모든 특징 정보 $\tilde{F}^1$ 내지 $\tilde{F}^K$ 및 레이블 데이터 y^1 내지 y^K에 양자 회로를 적용하되 PVM(Projection-Valued Measure)을 기반으로 로컬 데이터의 분류에 대한 예측 결과를 출력한다.

[0060] 보다 상세하게, 로컬 디바이스로부터 수신된 특징 정보는 서버의 QCNN 및 QNN으로 처리되며, QCNN의 기본적인 계산 및 동작은 앞서 설명한 로컬 디바이스의 QCNN과 동일하므로 여기서는 그 상세한 설명을 생략하도록 한다. 다만 서버의 QNN은 다중 클래스 분류 동작을 수행하며, 이를 위해 도 1을 통해 설명한 PVM 방식이 활용된다. 따라서 양자 상태 $|\psi\rangle_{\mathbf{X},\Theta^G}$ 이 기저 상태, 즉 {|0〉, ..., |n〉, ..., |2^q-1〉}에 투영할 때 예측 값을 이용할 수 있다.

[0061] 이하에서는 도 5를 통해 로컬 디바이스와 서버 간의 양자 스플릿 학습 방법에 대해 설명하도록 한다.

[0062] 각 로컬 디바이스, 일례로 k번째 로컬 디바이스는 로컬 데이터 X^k로부터 교차-채널 풀링을 이용하여 특징 정보 $\tilde{F}^k$를 획득한다.(lines 4-5) 그런 다음 각 로컬 디바이스는 자신의 특징 정보 및 레이블 데이터를 서버로 전송한다.(line 6)

[0063] 로컬 디바이스들로부터 모든 특징 정보 및 레이블 데이터를 수신한 서버는, 수신된 특징 정보를 기반으로 예측을 수행한다. 이때 훈련 가능한 매개변수들은 $\Theta = \{\{\Theta^k\}_{k=1}^K, \Theta^G\}$ 와 같이 나타낼 수 있으며, 여기서 PVM의 목적 함수(objective function) 하기 수학식 7과 같이 나타낼 수 있다.

수학식 7

[0064] $$\mathcal{L}(\Theta;\zeta) = \frac{1}{|\zeta|}\sum(\mathbf{X},\boldsymbol{y}) \in \zeta[\mathcal{L}_{\mathrm{BCE}}(\Theta;\mathbf{X}) + \mathcal{L}_{\mathrm{PAR}}(\Theta;\mathbf{X})]$$

[0065] 여기서 Θ 는 훈련 가능한 매개변수, ζ는 미니 배치, X는 로컬 데이터, y는 레이블 데이터를 나타내고, L_{BCE}는 이진 교차 엔트로피(Binary Cross Entropy)를 나타내고, L_{PAR}은 정규화기(regularizer)를 나타낸다. L_{BCE} 및 L_{PAR}은 하기 수학식 8,9와 같이 나타낼 수 있다.

수학식 8

[0066] $$\mathcal{L}_{\mathrm{BCE}}(\Theta;\mathbf{X}) = -\sum_{n=1}^{|\boldsymbol{y}|}[y_n \log p_n + (1-y_n)\log(1-p_n)]$$

[0067] 여기서 |y|는 분류하고자 하는 작업의 클래스 개수, p_n은 수학식 5에서 획득되는 클래스 n에 대한 예측 값을 나타낸다.

수학식 9

[0068] $$\mathcal{L}_{\mathrm{PAR}}(\Theta;\mathbf{X}) = -\sum_{n'>|\boldsymbol{y}|}^{2^q}\log(1-p_{n'})$$

[0069] 여기서 $q \geq \lceil \log_2(|\boldsymbol{y}|)\rceil$ 이고, 2^q는 QNN의 출력이고, $P_{n'}$는 n'번째 투영 행렬을 나타낸다. 로컬 QNN과 글로벌 QNN의 손실함수 기울기(loss gradient)는 하기 수학식 10을 통해 계산되고 업데이트된다. 즉 서버는 각 로컬 디바이스로부터 수신된 레이블 데이터 및 로컬 데이터의 분류에 대한 예측 결과를 기반으로 각 로컬 디바이스에 대한 로컬 기울기 $\nabla_{\Theta^k}\mathcal{L}$ 을 계산하고, 계산된 로컬 기울기를 해당 로컬 디바이스에게 분배한다.(lines 8-11) 그런 다음 서버는 글로벌 기울기 $\nabla_{\Theta^G}\mathcal{L}$ 을 계산한다.(line 12)

수학식 10

[0070] $$\frac{\partial f(\Theta)}{\partial \theta_i} = f\left(\Theta + \frac{\pi}{2}\mathbf{e}_i\right) - f\left(\Theta - \frac{\pi}{2}\mathbf{e}_i\right)$$

[0071] 여기서 e_i는 Θ_i(즉, $\Theta \cdot \mathbf{e}_i = \theta_i$)를 제외한 모든 성분을 제거한 원-핫 엔코딩 벡터를 의미한다.

[0072] 이후 매개변수(parameters) 업데이트 과정을 살펴보면, 로컬 디바이스 k는 훈련 가능한 매개변수 Θ^k를 로컬 기울기를 고려하여 업데이트하고, 서버는 훈련 가능한 매개변수 Θ^G를 글로벌 기울기를 고려하여 업데이트한다.(lines 14-15)

[0073] 도 6은 본 발명의 실시예에 따른 양자 스플릿 학습 방법을 나타낸 순서도이다.

[0074] 적어도 하나의 로컬 디바이스는 로컬 데이터 및 레이블 데이터로부터 특징 정보를 추출한다.(S602) 이때 상기 적어도 하나의 로컬 디바이스는 상기 로컬 데이터에 QCNN 양자 회로를 적용하여 이미지 특징 정보를 추출하고, 상기 레이블 데이터에 QNN 양자 회로를 적용하여 레이블 특징 정보를 추출한다.

[0075] 또한 적어도 하나의 로컬 디바이스는 S602에서 추출된 특징 정보와 상기 레이블 데이터를 출력한다.(S604)

[0076] 서버는 상기 적어도 하나의 로컬 디바이스로부터 수신된 모든 특징 정보 및 레이블 데이터에 양자 회로를 적용하고,(S606) PVM을 기반으로 상기 로컬 데이터의 분류에 대한 예측 결과를 출력한다.(S608)

[0077] 이하에서는 표 1 및 표 2와 도 7 및 도 8을 통해 본 발명에서 제안하는 양자 스플릿 학습 방법의 성능에 대하여 알아보도록 한다.

[0078] 하기 표 1은 다양한 양자 분산 학습, 즉 독립형 학습(standalone), 양자 연합 학습(QFL), 양자 스플릿 학습(QSL) 각각의 성능을 데이터 세트 별로 나타낸 표이다. 표 1에서는 다양한 양자 분산 학습의 성능을 상위 1개(top-1) 정확도로 나타내었으며, MNIST, FashionMNIST, CIFAR10 데이터 세트를 사용하였다.

표 1

[0079]

Metric	MNIST	FashionMNIST	CIFAR10
Standalone	58.13	64.53	14.68
QFL	63.32	68.92	22.89
QSL	64.96	71.29	26.43

[0080] MNIST 데이터 세트의 경우 QSL이 독립형보다 6.83%, QFL보다 1.64% 더 높은 정확도를 보여주고, FashionMNIST 데이터 세트의 경우 QSL이 독립형보다 6.76%, QFL보다 2.37% 더 높은 정확도를 보여주고, CIFAR10 데이터 세트의 경우 QSL이 독립형보다 11.75%, QFL보다 3.54% 더 높은 정확도를 보여줌을 표 1을 통해 확인할 수 있다.

[0081] 도 7은 본 발명의 실시예에 따른 양자 스플릿 학습 방법의 성능 테스트에 사용된 다양한 데이터 세트의 학습 곡선을 나타낸 그래프이다.

[0082] 도시된 그래프는 표 1의 QSL에 대하여 MNIST, FashionMNIST, CIFAR10 데이터 세트 별 성능을 테스트 정확도를 통해 나타낸 것이다.

[0083] 도 7에서 확인할 수 있는 바와 같이 FashionMNIST에 대한 테스트 정확도가 가장 높게 나타나며, CIFAR10에 대한 테스트 정확도가 가장 낮게 나타남을 알 수 있다.

[0084] 하기 표 2는 풀링이 적용되지 않은 양자 스플릿 학습(QSL w/o pooling), 평균 풀링이 적용된 양자 스플릿 학습(QSL w/ avg-pooling), 교차-채널 풀링이 적용된 양자 스플릿 학습(QSL w/ C2Pool) 각각에 대한 훈련 데이터 성능(Train Acc.), 테스트 데이터 성능 (Test Acc.), 통신비(Comm. Cost)를 나타낸 표이다.

표 2

[0085]

Metric	Train Acc.	Test Acc.	Comm. Cost [Byte]
QSL w/o pooling	66.64	59.45	1,024
QSL w/ avg-pooling	67.01	58.32	64
QSL w/ C2Pool	68.96	64.96	64

[0086] QSL w/o pooling과 QSL w/ avg-pooling의 성능이 본 발명에서 제안하는 QSL w/ C2Pool의 성능보다 떨어지며, QSL w/ C2Pool의 훈련 데이터와 테스트 데이터 사이의 도메인 갭(gap)이 가장 작음을 표 2를 통해 확인할 수 있다. 또한 본 발명에서 제안하는 QSL w/ C2Pool은 통신비 절감 효과가 있음을 확인할 수 있다.

[0087] 도 8은 MNIST 데이터 세트에서 로컬 디바이스 개수에 따른 성능 비교를 나타낸 그래프이다.

[0088] 도시된 그래프는 양자 스플릿 학습 시스템에서 로컬 디바이스 개수(Number of Local Devices) K가 1,2,5,10,20개인 경우 각각에 대한 양자 스플릿 학습의 성능을 상위 1개(top-1) 정확도로 나타낸 것이다.

[0089] 도 8서 확인할 수 있는 바와 같이 K=1인 독립형인 경우 상위 1개 정확도가 58.13%로 가장 낮게 나타나며, K=20인 경우 상위 1개 정확도가 68.81%로 가장 높게 나타남을 알 수 있다. 이를 통해 양자 스플릿 학습 시스템의 로컬 디바이스 개수가 증가할수록 더 높은 정확도를 얻을 수 있음을 확인할 수 있다.

[0090] 위와 같은 본 발명의 양자 스플릿 학습 방법은 다양한 컴퓨터 구성요소를 통하여 수행될 수 있는 프로그램 명령어의 형태로 구현되어 컴퓨터 판독 가능한 기록 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능한 기록 매체는 프로그램 명령어, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다.

[0091] 상기 컴퓨터 판독 가능한 기록 매체에 기록되는 프로그램 명령어는 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 분야의 당업자에게 공지되어 사용 가능한 것일 수도 있다.

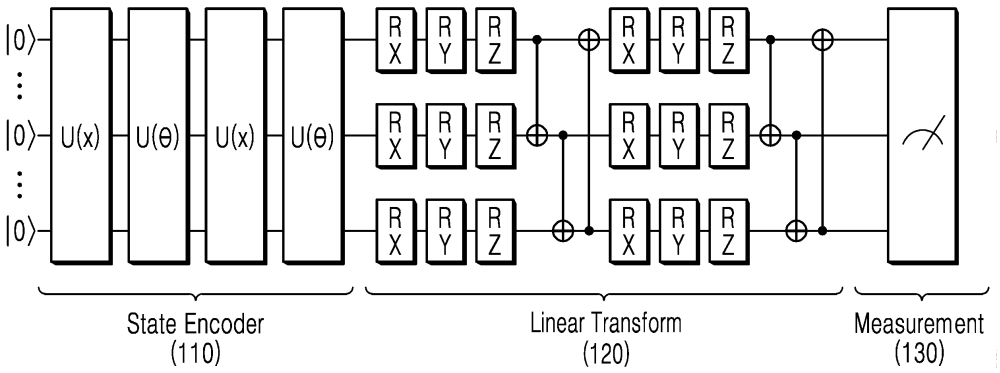
[0092] 컴퓨터 판독 가능한 기록 매체의 예에는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM, DVD 와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 ROM, RAM, 플래시 메모리 등과 같은 프로그램 명령어를 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다.

[0093] 프로그램 명령어의 예에는, 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드도 포함된다. 상기 하드웨어 장치는 본 발명에 따른 처리를 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

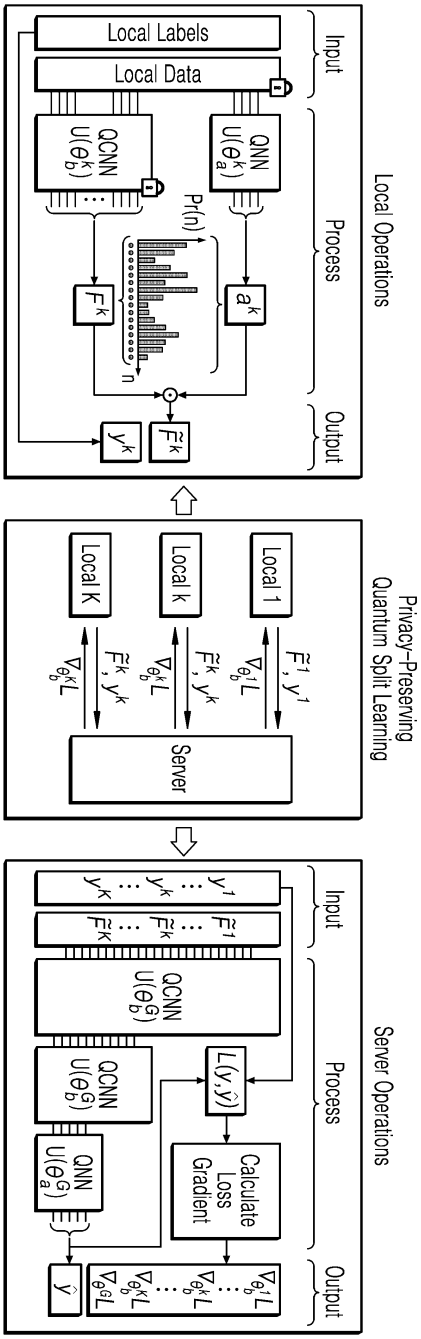
[0094] 이상에서는 본 발명의 다양한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해되어져서는 안될 것이다.

도면

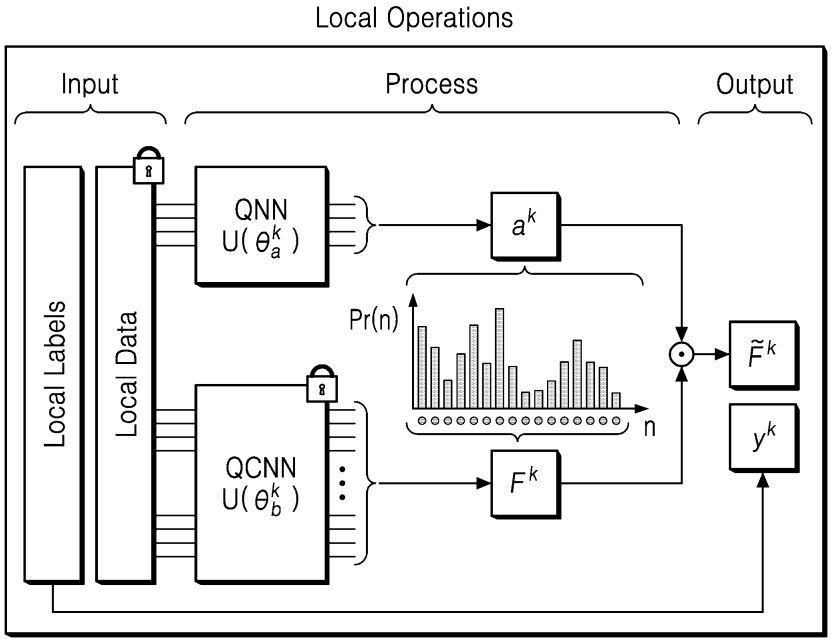
도면1



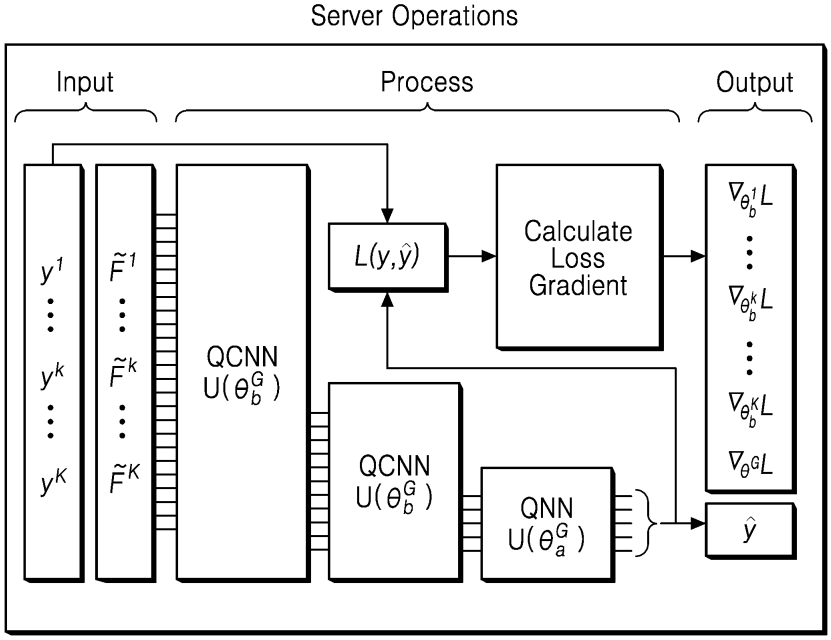
도면2


Local Operations
Input
Process
Output
Local Labels
Local Data
QCNN
QNN
Pr(n)
n
Privacy-Preserving
Quantum Split Learning
Local K
Local k
Local 1
Server
Server Operations
Input
Process
Output
QCNN
QNN
Calculate
Loss
Gradient

도면3


Local Operations
Input
Process
Output
Local Labels
Local Data
QNN
$U(\theta_a^k)$
QCNN
$U(\theta_b^k)$
a^k
Pr(n)
n
F^k
$\tilde{F}^k$
y^k

도면4

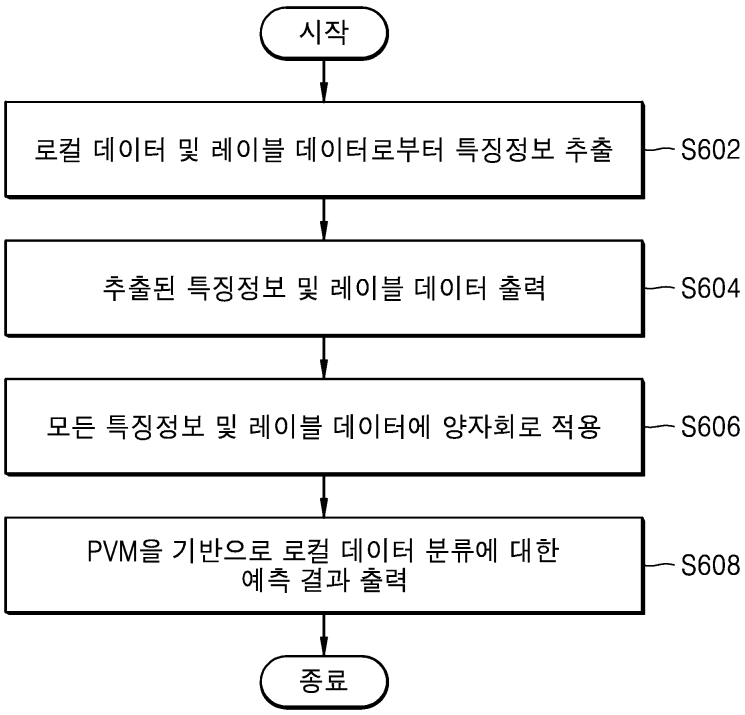

Server Operations
Input
Process
Output
y^1
y^k
y^K
$\tilde{F}^1$
$\tilde{F}^k$
$\tilde{F}^K$
QCNN
$U(\theta_b^G)$
$L(y,\hat{y})$
Calculate
Loss
Gradient
QCNN
$U(\theta_b^G)$
QNN
$U(\theta_a^G)$
$\nabla_{\theta_b^1} L$
$\nabla_{\theta_b^k} L$
$\nabla_{\theta_b^K} L$
$\nabla_{\theta^G} L$
$\hat{y}$

도면5

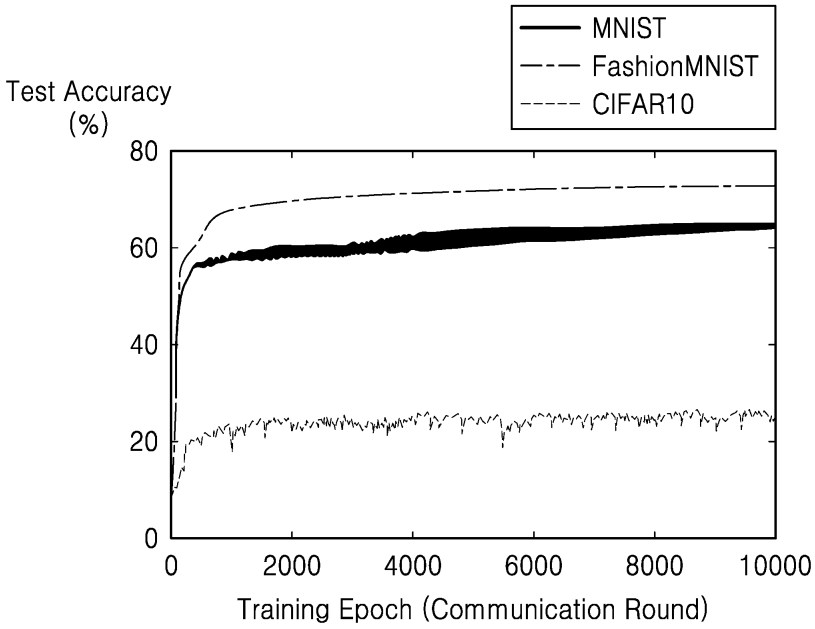
Algorithm 1: Quantum Split Learning

Initialize. $\mathcal{Z} \leftarrow \emptyset$, $\forall \Theta^k$, and Θ^G;
◂ Local-side operations ▸
for $k \in \{1, \cdots, K\}$ do
    Generate features with local data $\mathbf{X}^k$;
    Obtain cross-channel pooled features $\tilde{\mathbf{F}}^k$ with (5);
    Transmit features and labels to server, $\mathcal{Z} \leftarrow \mathcal{Z} \cup (\tilde{\mathbf{F}}^k, y^k)$
◂ Server-side operations ▸
for $k \in \{1, \cdots, K\}$ do
    Make prediction with features and calculate loss;
    Calculate the local gradient, $\nabla_{\Theta^k} \mathcal{L}$;
    Distribute gradients to local k;
Calculate the global gradient, $\nabla_{\Theta^G} \mathcal{L}$;
◂ Parameters update ▸
Local k: $\Theta^k \leftarrow \Theta^k - \eta \nabla_{\Theta^k} \mathcal{L}$;
Server: $\Theta^G \leftarrow \Theta^G - \frac{\eta}{K} \nabla_{\Theta^G} \mathcal{L}$;

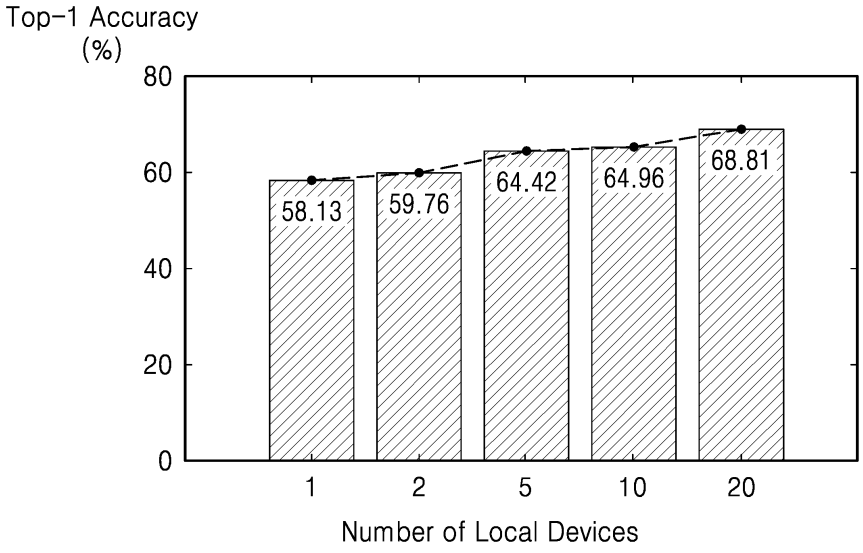
도면6



도면7


MNIST
FashionMNIST
CIFAR10
Test Accuracy
(%)
80
60
40
20
0
0
2000
4000
6000
8000
10000
Training Epoch (Communication Round)

도면8


Top-1 Accuracy
(%)
80
60
40
20
0
58.13
59.76
64.42
64.96
68.81
1
2
5
10
20
Number of Local Devices