

논문 리뷰

Quantum convolutional neural networks

발표자: 양유진

링크: <https://youtu.be/Eppb9BkB0CU>

1. QCNN circuit model - 배경

- 양자 물리학 문제를 풀 때 고전적인 기계 학습 방법으로 해결하기 어려움

Ex) 힐베르트(hilbert) 공간에서 정의된 양자 물리학 문제에 기계 학습 기술을 적용하기 위해 데이터들을 고전적인 컴퓨터 데이터 형식으로 변환해야 함.

But,

1. 기존의 컴퓨팅 환경에서 데이터와 모델이 효율적으로 처리되지 않는 경우가 존재함.
 2. 시스템의 크기가 커질수록 데이터의 크기는 기하급수적으로 늘어남
- 머신러닝 방식을 적용해서 효과적으로 해결하기 어려움.

양자 물리학 문제를 효율적으로 풀기 위해
CNN 구조를 양자 시스템에 적용하는 방식을 고안.

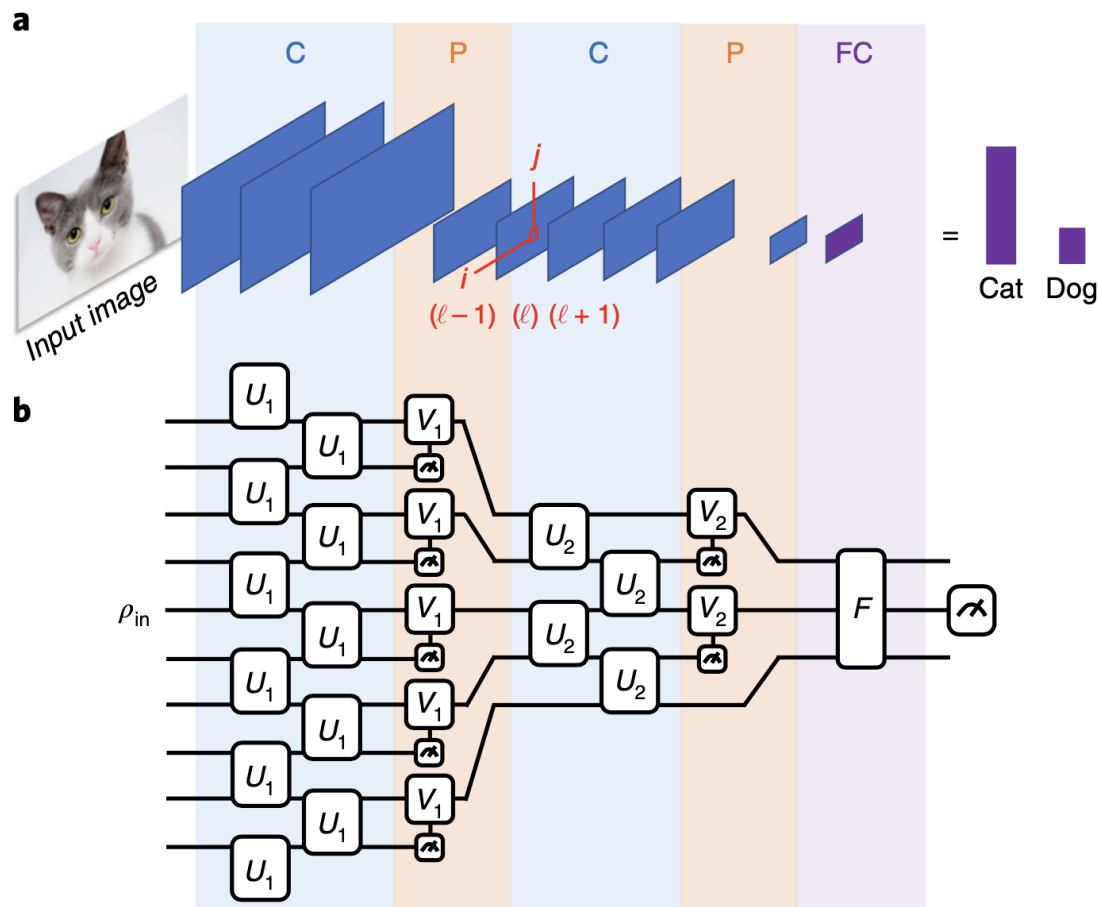
1. QCNN circuit model - QCNN의 핵심

“ CNN의 핵심 특징 (Convolution layer & Pooling layer)을
양자 시스템에 적용한 것 ”

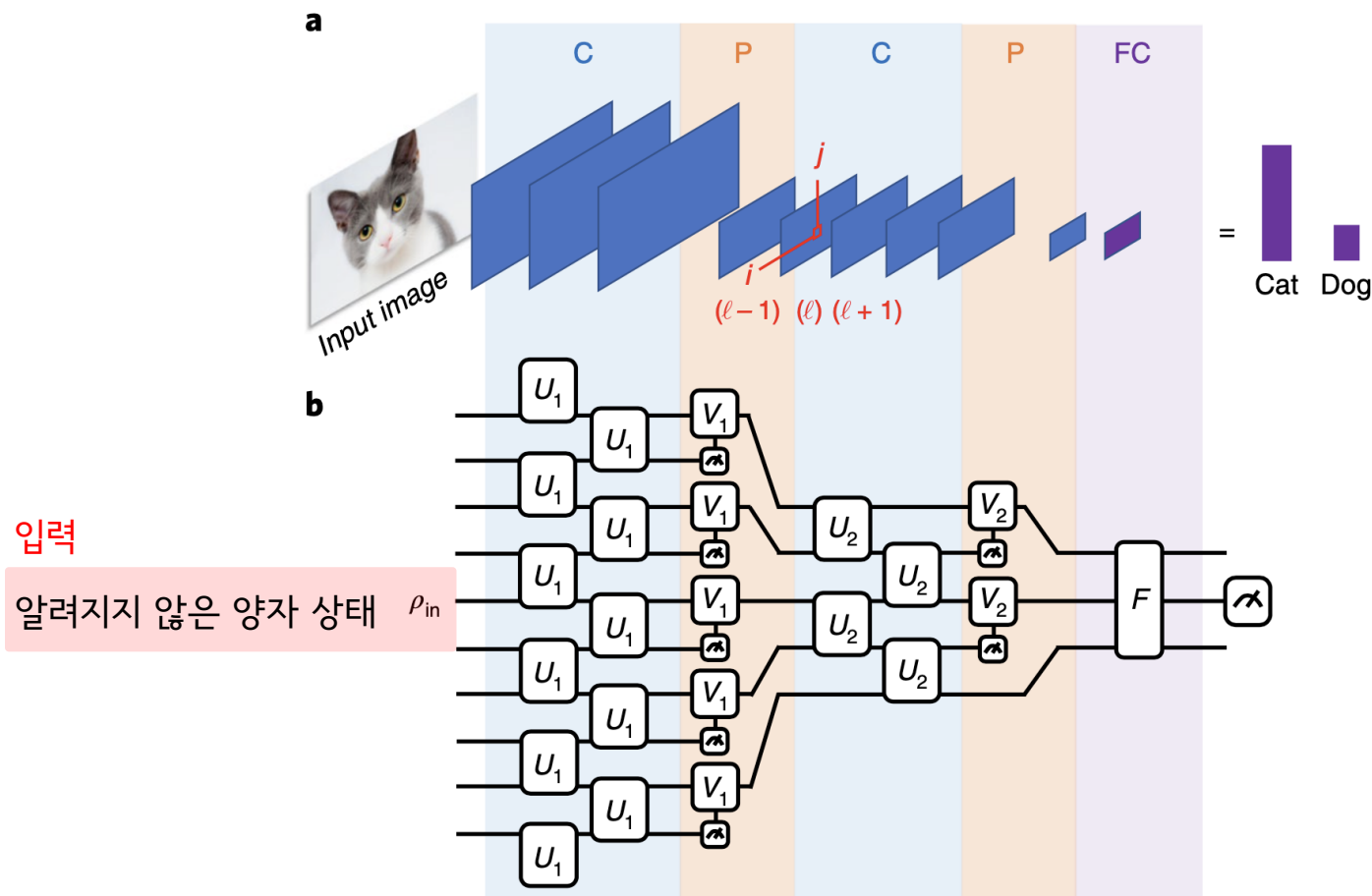
1. QCNN circuit model - 구조

CNN의 핵심 특징 (Convolution layer & Pooling layer)을 양자 시스템에 적용

< CNN을 이용한 고양이와 개 이미지 분류 >



1. QCNN circuit model - 구조



Convolution layer

U_i : single quasi-local unitary
인접한 qubit 사이에 여러
qubit 게이트를 적용하여
숨겨진 상태(hidden state)
를 찾음

Pooling layer

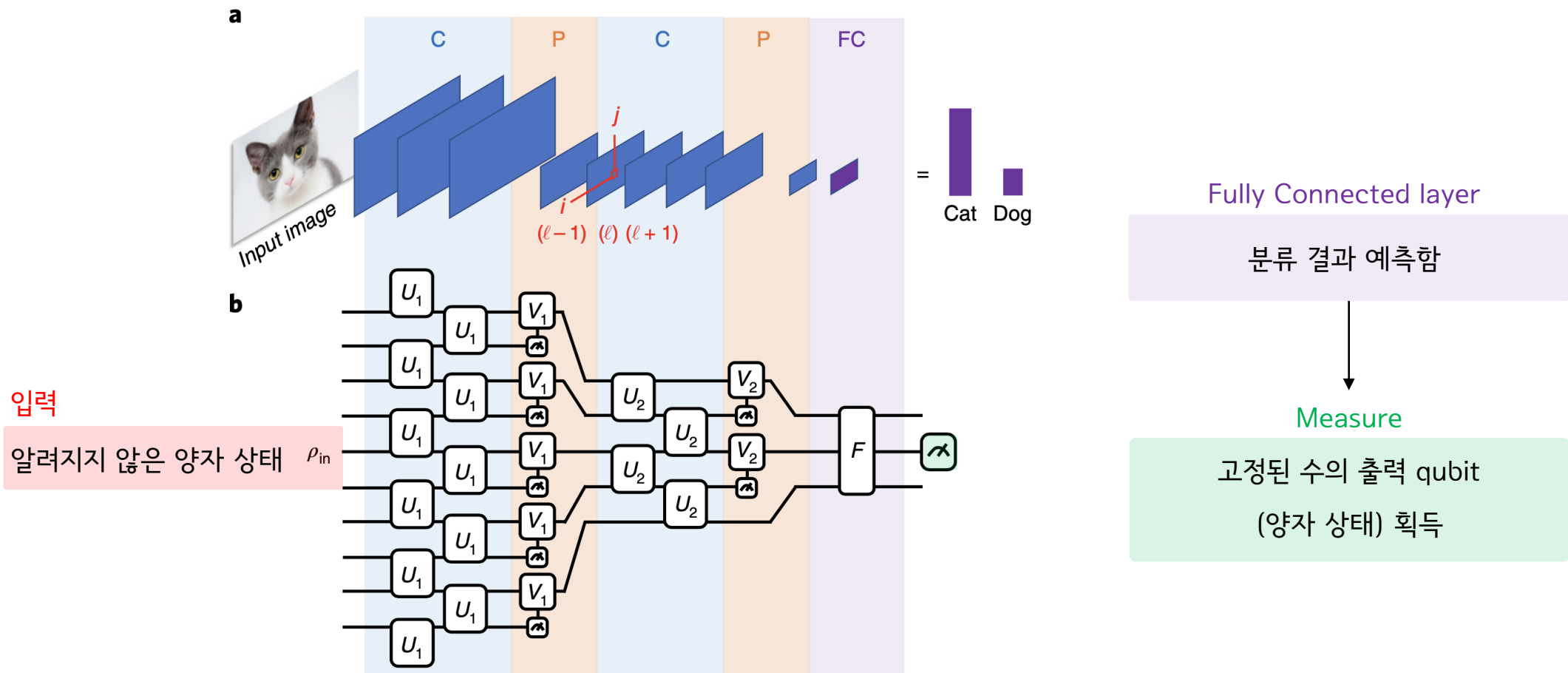
V_2 : unitary rotations

큐비트의 일부가 측정되고
그 결과에 따라 인접 큐비트에
적용되는 단일 회전이 결정됨.

2-qubit gate(CNOT gate)를
적용하여 양자 시스템 크기 줄임

시스템의 크기가
충분히 작아질
때까지 반복

1. QCNN circuit model - 구조



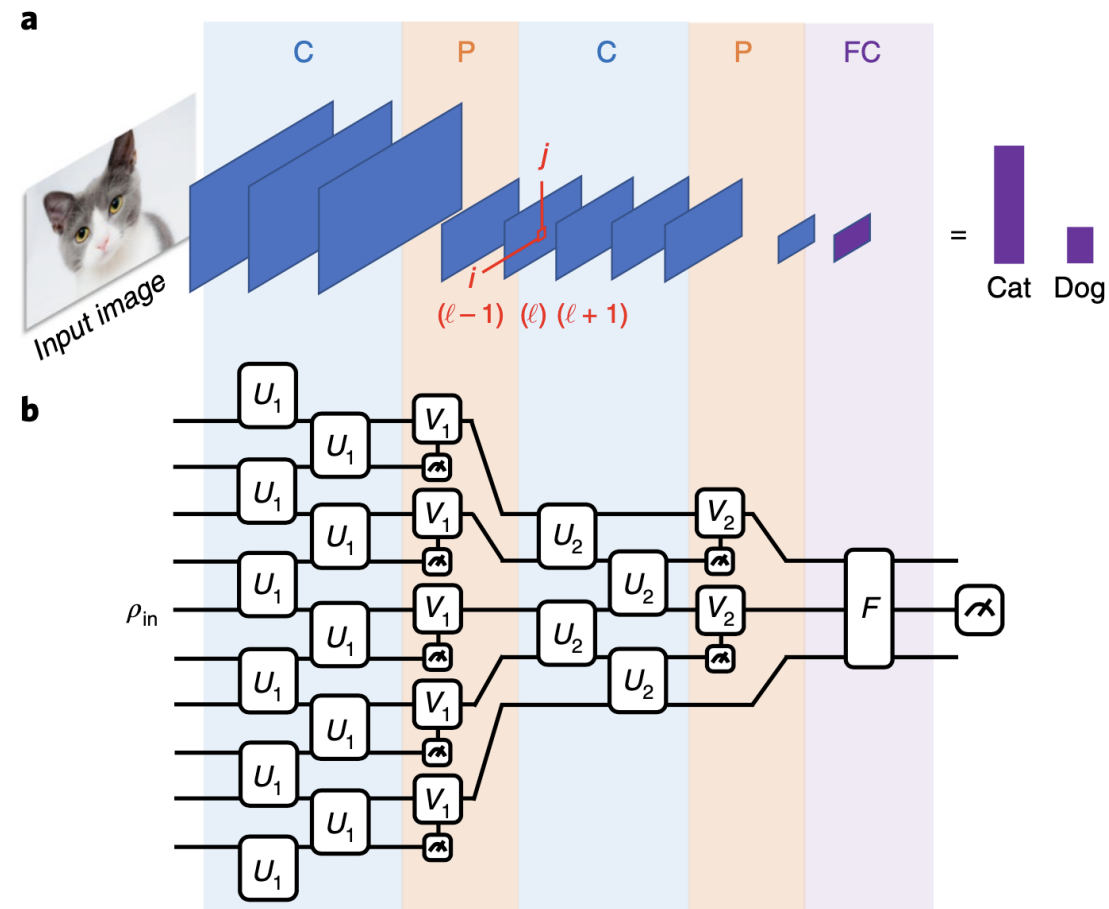
N-qubit 입력 상태를 분류하는 QCNN은 $O(\log(N))$ 매개변수로 특징지어짐

→ 일반적인 양자 회로 기반 분류기들과 비교하여 2배의 지수적 감소 & 효율적인 학습/구현을 허용함

1. QCNN circuit model - 학습

$$\text{MSE} = \frac{1}{2M} \sum_{\alpha=1}^M \left(y_i - \text{입력에 대한 예상 QCNN 출력값} \right)^2$$

$$\frac{\partial \text{MSE}}{\partial c_{\mu}} = \frac{1}{2\epsilon} \left(\text{MSE}(c_{\mu} + \epsilon) - \text{MSE}(c_{\mu} - \epsilon) \right) + O(\epsilon^2)$$



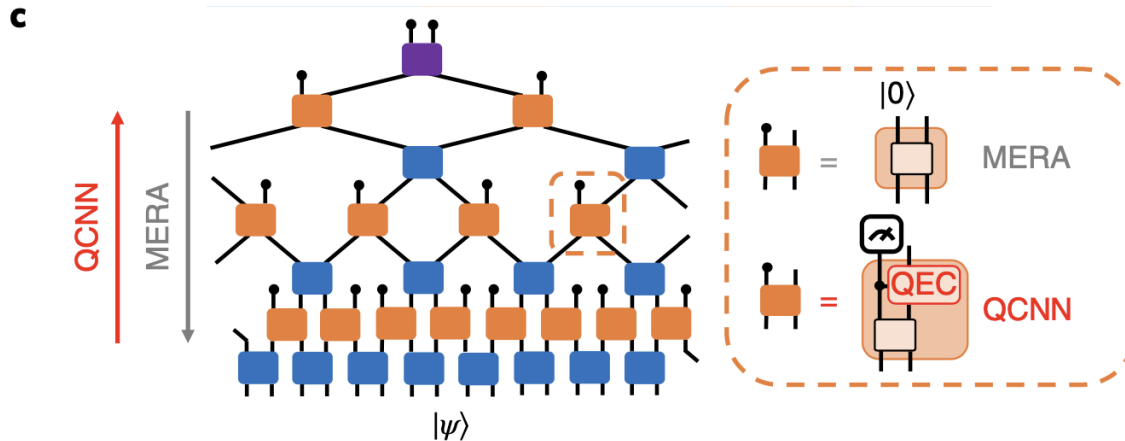
MSE 함수와 경사 하강법(유한 차분법 사용하여 경사의 근사값 구함)을 사용하여 수렴할 때까지 유니터리를 반복적으로 업데이트 하는 것으로 구성됨. (하이퍼파라미터 고정)

2. MERA and QEC - MERA와 QCNN

MERA(Multi-Scale Entanglement Renormalization Ansatz)

: 양자 다체 시스템(quantum many-body system)을 효율적으로 시뮬레이션하도록 설계된 모델

*다체 문제: 여러 물체의 초기 위치, 초기 속도가 주어졌을 때 이후의 운동 상태를 찾는 문제



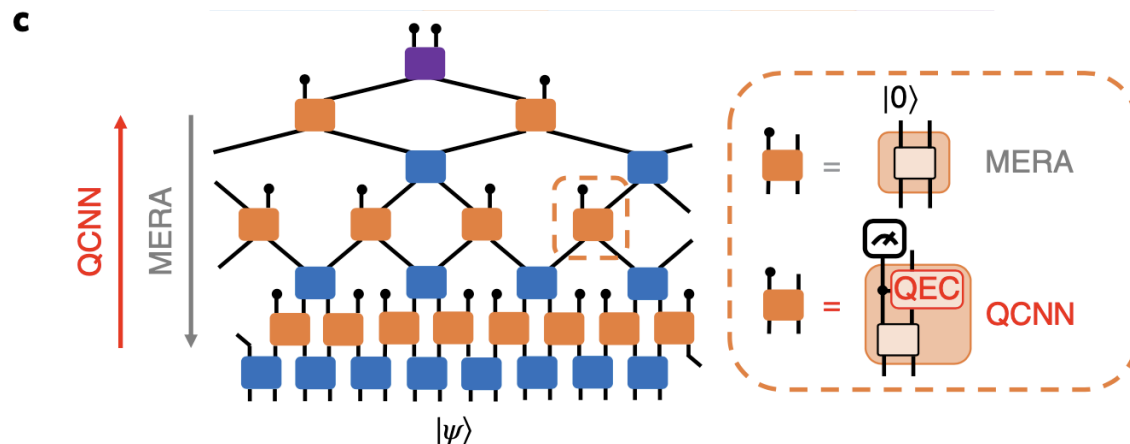
- QCNN과 MERA는 회로 구조가 동일함
- 회로의 실행 방향이 반대임.

MERA는 $|0\rangle$ 큐비트를 추가하여 각 깊이에 대한 양자 시스템의 크기를 기하급수적으로 증가시킴

QCNN은 이를 역방향으로 진행시켜 양자 시스템의 크기를 기하급수적으로 감소시킴

저자는 MERA 모델에 QEC(Quantum Error Correction)를 추가하여 추가적인 성능 향상을 제안함

2. MERA and QEC - 기존 QCNN 문제점



QCNN은 MERA에서 얻을 수 있는 각 레이블에 대한 대표 상태 $|\psi\rangle$ 를 입력 데이터로 주면,
그에 상응하는 레이블을 얻을 수 있음

[문제] 만약 MERA에서 생성이 불가능한 상태 $|\psi'\rangle$ 를 입력 데이터로 주면,
QCNN은 레이블을 획득할 수 없음.

→ MERA에 QEC를 적용하여 추가 자유도를 부여함으로써 수정 및 해결 가능

2. MERA and QEC - QEC를 통한 성능 개선

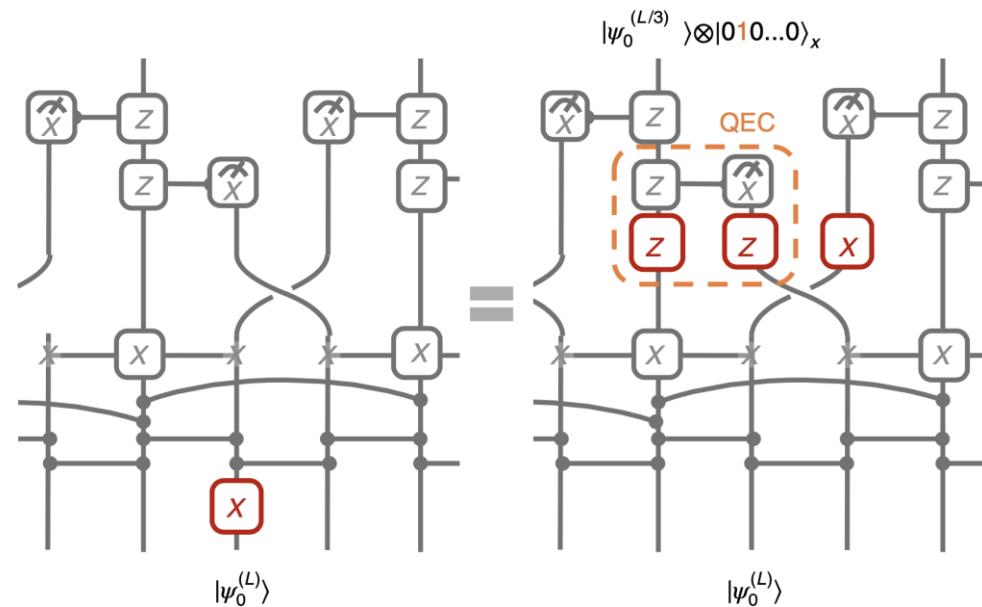
[주어진 데이터가 $|\psi\rangle$ (MERA에서 생성한 상태) 일 때]

Pooling layer에서 측정된 결과는 MERA에서 새로 주어진 상태 $|0\rangle$ 과 같아야 함

[주어진 데이터가 $|\psi'\rangle$ (MERA에서 생성할 수 없는 상태) 일 때]
Pooling layer에서 측정된 결과로 $|1\rangle$ 이 나올 가능성이 있음.

$|1\rangle$ 이 나올 경우, 큐비트에 추가적인 게이트를 적용하여 $|0\rangle$ 이 될 수 있게 결과를 보정

< QEC 수행하는 과정 >



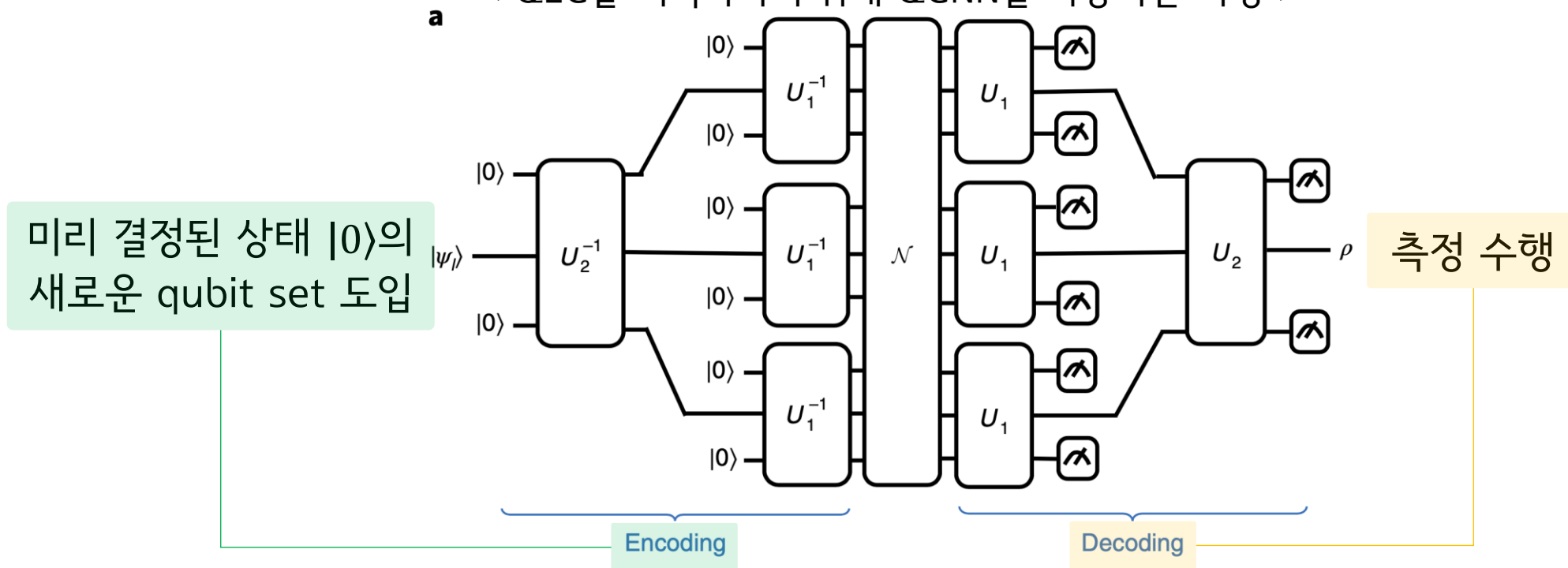
→ Convolution-Pooling 단계를 거치면 오류 식별/수정하는 동시에 시스템의 크기를 줄여줌.

→ Convolution-Pooling 단계를 여러 번 거치면 오류 밀도가 상당히 작아짐.

3. Optimizing QEC

주어진 오류 모델에 대해 자체적으로 최적화된 새로운 QEC 코드 설계

< QEC를 최적화하기 위해 QCNN을 사용하는 과정 >



- 오류 채널 \mathcal{N} 이 주어졌을 때, 복구충실도를 최대화하는 것이 목표 $\langle \psi_i | \rho | \psi_i \rangle$
- 실제 시스템에서 효율적인 구현 보장 + 인코딩/디코딩 체계를 동시에 최적화함
- 알려지지 않은 오류 모델에도 최적화된 양자 오류 수정 체계를 고안할 수 있음을 보임

4. Outlook

- QCNN이 유망한 양자 기계 학습 패러다임을 제공함을 나타냄
- *translation invariance(변환 불변성) 제약 조건 완화하면
 - 시스템 크기가 n 일 때, $O(\log(N))$ 를 가지는 매개변수 생성 할 수 있음
 - ancilla qubit 사용하여 기존 CNN 구조처럼 feature map을 병렬로 구현 할 수 있음
- QEC 코드 공간에서 내결함성(fault-tolerant) 연산 최적화를 통합할 수 있음
- 기울기 계산을 위해 '유한 차분 체계'를 사용하였는데 QCNN과 기존 체계의 구조적 유사성은 역전파와 같은 효율적인 체계를 채택할 가능성을 시사함.

*변환 불변성: 입력의 위치가 변해도 출력이 변하지 않는 속성

감사합니다