

# 차등 학습 전략을 적용한 QCNN의 파라미터 효율적인 학습 가능 관측량 연구

김민주, 허 준\*

고려대학교, 고려대학교\*

mindy0522@korea.ac.kr, \*junheo@korea.ac.kr

## A Study on Parameter-Efficient Learnable Observables in QCNNs with Differentiated Learning Strategies

Minjoo Kim, Jun Heo \*

Korea Univ., Korea Univ.\*

### 요약

기존 양자 컨볼루션 신경망(QCNN)은 고정된 관측량(Observable)을 사용하여 데이터로부터 유용한 정보를 추출하는 데 제약이 있었다. 본 연구에서는 이를 극복하고자, 파라미터 수를 효율적으로 설계한 학습 가능 관측량(learnable observable)을 각 큐비트에 적용하는 방법을 제안한다. 나아가, 관측량 파라미터의 효과적인 최적화를 위해 기존 VQC(Variational Quantum Circuit) 파라미터 그룹과 분리하여 서로 다른 학습 전략을 적용하는 차등 학습의 효용성을 분석했다. MNIST 데이터셋 분류 실험 결과, 관측량 파라미터 그룹에 별도의 옵티마이저(Adagrad)를 적용한 모델이 가장 높은 테스트 정확도를 보여, 제안한 방법론이 QCNN의 성능을 효과적으로 향상시킬 수 있음을 증명했다.

### I. 서 론

양자 컨볼루션 신경망(QCNN)은 Cong et al. [1]에 의해 처음 제안된 아래, 이미지 인식 분야에서 양자적 특성을 활용하는 유망한 아키텍처로 주목받고 있다. 하지만 기존 QCNN의 학습은 변분 양자 회로(VQC) 내의 파라미터만을 최적화하며, 측정 단계에서는 고정된 관측량(Observable)을 사용하므로 표현력에 한계가 있다.

이러한 구조적 한계를 해결하고자, S. Y. Chen 등은 관측량 또한 학습 가능한 파라미터로 구성하여 측정 과정을 최적화하는 새로운 방향을 제시했다[2]. 본 연구는 이 아이디어를 바탕으로 하되, 파라미터 효율성과 학습 전략에 초점을 맞추었다. 기존 연구의  $2^n \times 2^n$  전역 관측량 대신  $2 \times 2$  국소 관측량을 적용하여 학습 파라미터 수를 크게 줄였으며, 관측량 파라미터의 최적화에는 별도의 학습 전략이 효과적일 것이라는 가설 아래 VQC 파라미터 그룹과 학습률 및 옵티마이저를 분리하여 QCNN의 성능 변화를 비교 분석하고자 한다.

### II. 본 론

#### A. 양자 컨볼루션 신경망(QCNN)

양자 컨볼루션 신경망(QCNN)은 고전 컨볼루션 신경망(CNN)의 구조에서 영감을 받은 하이브리드 모델로, 이미지의 공간적 특징을 추출하는 필터(filter)의 역할을 양자 회로로 대체한 것이다. CNN처럼 QCNN 또한 동일한 양자 회로가 이미지 전체를 순회하며 가중치를 공유(weight sharing)하므로, 적은 수의 파라미터만으로 효율적인 학습이 가능하다.

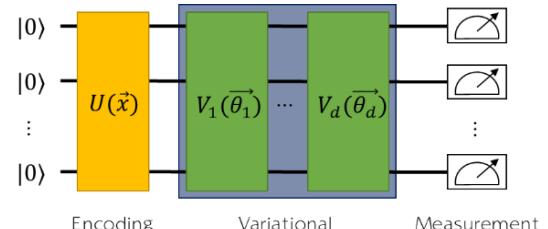


그림 1 양자 필터 회로 구성

이 양자 필터 회로는 그림 1과 같이 세 단계로 구성된다. 첫째, 인코딩(Encoding) 단계에서는 고전 꾀셀 값들을 양자 상태로 변환한다. 둘째, 학습 가능한 파라미터( $\theta$ )를 포함한 변분 양자 회로(VQC)가 얕은 힘과 회전을 통해 데이터의 패턴을 포착한다. VQC 내의 파라미터  $\theta$ 는 옵티마이저(optimizer)를 통해 최적화되는 모델의 주된 학습 대상이다. 마지막 측정(Measurement) 단계에서는 VQC를 거친 최종 양자 상태의 기댓값을 계산하여 고전적인 특정 값으로 출력하고, 이를 후속하는 고전 신경망으로 전달하여 최종 분류에 사용한다.

#### B. 학습 가능한 관측량 (Learnable Observable)

기존 QCNN의 측정 단계는 Pauli-Z와 같은 고정된 연산자를 사용하므로, 항상 동일한 측정 기저(basis)와  $[-1, 1]$  범위의 고유값만을 갖는 한계가 있다. 본 연구에서는 이러한 제약을 극복하고자 각 큐비트에 독립적인  $2 \times 2$  국소 관측량을 적용하여 측정 과정을 최적화한다. 이는 참고 문헌 [2]가  $n$ -큐비트 시스템에  $2^n \times 2^n$  크기의 전역 관측량을 사용한 것과 비교해 파라미터 효율성을 높인 방식이다.

각 큐비트에 적용되는 관측량  $O$ 는 4 개의 실수 파라미터  $\mathbf{b} = (b_1, b_2, b_3, b_4)$ 와 Pauli 행렬의 선형 결합으로 다음과 같이 정의된다:

$$O(\mathbf{b}) = b_1 I + b_2 \sigma_x + b_3 \sigma_y + b_4 \sigma_z$$

이 파라미터  $\mathbf{b}$ 는 VQC 파라미터  $\theta$ 와 함께 최적화된다. 이 과정을 통해 모델은 주어진 데이터와 분류 작업에 가장 유용한 정보를 추출할 수 있는 측정 기저와 고유값 스펙트럼(eigenvalue spectrum)을 스스로 학습하게 되어, 최종 분류 정확도를 향상시킨다.

### C. 실험 방법

본 연구의 가설 검증을 위해, MNIST 데이터셋을 이용하여 분류(classification) 실험을 진행했다. 제안한 방법론의 효과를 분석하기 위해, 다음과 같이 세 가지 QCNN 모델을 설계하여 단계별 성능 향상을 비교했다.

- ① VQC: 학습 가능한 관측량(LO)을 도입하지 않은 기준 모델
- ② VQC+LO: LO 를 도입하되, VQC 파라미터와 동일한 학습률 및 옵티마이저로 훈련하는 모델
- ③ Differential Learning: LO 파라미터 그룹에 적응형 학습률을 지원하는 Adagrad 옵티마이저를 별도로 적용한 모델

세 모델의 아키텍처 외 다른 모든 실험 조건은 동일하게 통제했다. 모든 이미지는 양자 회로에 효율적으로 인코딩하기 위해  $10 \times 10$  픽셀 크기로 전처리되었으며, 실험은 PennyLane 과 PyTorch 기반 시뮬레이터 환경에서 수행되었다. 각 모델의 학습 과정에서는 검증 데이터셋에 대한 손실(validation loss)과 정확도(validation accuracy)를 모니터링했으며, 학습이 완료된 후에는 테스트 데이터셋에 대한 최종 분류 정확도를 측정하여 성능을 정량적으로 비교했다.

### D. 실험 결과 및 분석

실험 결과, 기준 모델인 모델 1 은 학습 초기에 가장 빠른 수렴 속도를 보였으나, 특정 성능 수준에서 정체되는 한계를 나타냈다. 반면, 학습 가능한 관측량(LO)을 도입한 모델 2 와 3 은 학습 초기는 다소 느렸지만, 학습이 진행됨에 따라 모델 1 의 성능을 추월하는 경향을 보였다. 이는 LO 을 적용한 모델이 더 복잡한 파라미터 공간을 탐색하며, 결과적으로 기준 모델의 성능 한계를 넘어설 잠재력을 가졌음을 시사한다.

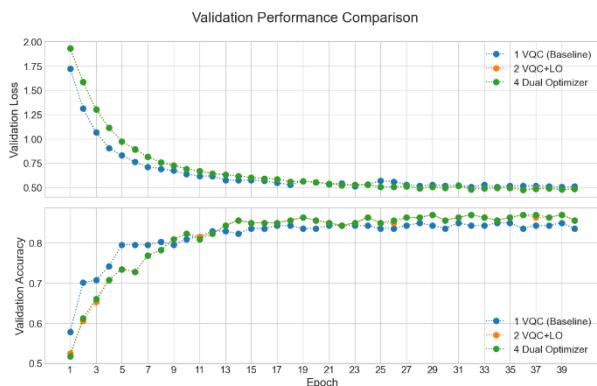


그림 2 에폭(Epoch)에 따른 검증 손실 및 정확도 변화

표 1 최종 테스트 손실 및 정확도 비교

	테스트 손실	테스트 정확도 (%)
VQC	0.4564	86.00
VQC+LO	0.3793	88.67
Differential Learning	0.3782	89.00

최종 테스트 결과(표 1), 학습 가능한 관측량(LO)을 도입한 모델들은 기준 모델보다 높은 성능을 달성했으며, 특히 VQC 와 LO 파라미터 그룹에 별도의 옵티마이저를 적용한 모델 3 이 89.00%로 가장 큰 성능 향상을 보였다. 이는 단순히 LO 를 추가하는 것을 넘어, 각 파라미터 그룹에 맞는 최적화 과정을 분리하는 차등 학습 전략이 유효하다는 본 연구의 가설을 강력하게 뒷받침한다.

### III. 결 론

본 연구는 QCNN 의 성능 향상을 위해 파라미터 효율적인 국소 학습 가능한 관측량(LO)과 차등 학습 전략을 제안하고 그 효과를 검증했다. MNIST 분류 실험 결과, LO 를 도입하는 것만으로도 고정 관측량을 사용한 기준 모델보다 높은 테스트 정확도를 보였다. 더 나아가, LO 파라미터 그룹에 별도의 옵티마이저를 적용한 차등 학습 모델이 가장 우수한 성능을 기록했다. 이는 모델이 데이터에 맞는 최적의 측정 기저를 학습하는 것이 중요할 뿐만 아니라, VQC 와 LO 파라미터의 최적화 과정을 분리하는 차등 학습 전략이 모델의 일반화 성능을 극대화하는 데 효과적임을 시사한다.

결론적으로, 본 연구에서 제안한 방법론은 더 높은 최종 성능을 달성했을 뿐만 아니라, 향후 더 복잡한 QCNN 아키텍처에서 부각될 파라미터 효율성까지 확보했다는 점에서 높은 의의를 가진다. 특히, 본 연구에서 사용한 Pauli 행렬의 선형 결합 방식은 실제 양자 하드웨어에서도 그래디언트(gradients) 계산이 가능한 파라미터-시프트 규칙(parameter-shift rule) 과 호환되므로, 제안한 모델은 향후 NISQ(Noisy Intermediate-Scale Quantum) 장치에서의 구현 및 학습에 대한 확장성 또한 갖추고 있다.

### ACKNOWLEDGMENT

이 논문은 2025년도 정부(과학기술정보통신부)의 재원으로 한국연구재단 양자기술연구개발선도(양자컴퓨팅)사업의 지원(RS-2024-00431853, 50%)과 2025년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No. RS-2023-00225385, 50%)을 받아 수행된 연구임

### 참 고 문 헌

- [1] Cong, Iris, Soonwon Choi, and Mikhail D. Lukin. "Quantum convolutional neural networks." *Nature Physics* 15.12 (2019): 1273–1278.
- [2] Chen, Samuel Yen-Chi, et al. "Learning to measure quantum neural networks." *2025 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*. IEEE, 2025.