

블록체인 기반 연합학습 가속화를 위한 파이프라이닝 기법 연구

Performance Optimization of Federated Learning With Blockchain-based Pipelining

이 희 태¹, 이 재 환¹

(Heetae Lee¹ and Jaehwan Lee^{1,*})

¹Department of computer engineering, Korea Aerospace University, Goyang-si, Gyeonggi-do, South Korea

Abstract: This study addresses inefficiencies in blockchain-based federated learning caused by computational overhead and block generation delays. Therefore, a pipelining architecture that decouples model training, transaction transmission, and block generation into independent threads is proposed in this work. This structure enables the parallel execution of computation and communication phases, minimizing GPU idle time. Experimental results reveal that the proposed approach reduces the block generation cycle time, approximately, by 26% on the SVHN dataset and by 41% on the CIFAR-10 dataset compared to the blockchain-based federated learning as the baseline, while maintaining model accuracy and stability. By effectively resolving structural bottlenecks, this method significantly enhances the scalability and practicality of the system.

Keywords: federated learning, blockchain, pipeline

I. 서론

연합학습은 로컬 데이터를 중앙서버로 집계하지 않고 학습을 진행하는 분산 머신러닝 기법이다[1]. 이 방식은 각 클라이언트가 각자의 로컬 데이터만을 이용해 모델 학습을 수행하고 서버에는 학습된 모델의 파라미터만을 전달하기 때문에 개인정보 및 데이터 유출의 위험이 줄어드는 장점이 있다. 이러한 특성으로 인해 연합학습은 다양한 분야에서 활용되고 있다. 특히 스마트폰이나 IoT (Internet of Things) 기기와 같은 엣지 컴퓨팅[2] 환경에서의 연구가 증가함에 따라 연합학습 적용 연구가 활발히 이루어지고 있다.

그러나 이러한 연합학습은 여러 한계점을 가지고 있다. 첫째, 일반적인 연합학습은 서버-클라이언트 구조를 기반으로 하기 때문에 단일 장애점(SPOF) 문제가 발생할 수 있다. 이는 중앙 서버로의 통신 집중이 심각한 시간 지연을 유발함을 의미하며, 자율주행 군집 제어 등에서 시간 지연이 시스템의 불안정성을 야기하듯[3], 연합학습에서도 노드의 이탈이나 전체 시스템의 마비를 초래할 수 있다.

둘째, 연합학습에서는 데이터가 로컬 기기에 남아 있어서 구조적으로 각 클라이언트의 데이터를 중앙에서 검증하거나 감시할 수 없다. 이에 따라 악의적인 참여자가 조작된 데이터를 이용해 로컬에서 학습을 수행하고, 이를 통해 전체 모델에 오류를 유도하는 데이터 조작(data poisoning) 공격이 가능해진다. 결과적으로 모델의 무결성이 훼손될 위험이 존재한다.

마지막으로, 연합학습은 각 참여자의 로컬 자원을 이용하여

학습을 진행하기 때문에, 참여자의 기여도를 객관적으로 평가하고 이에 따라 인센티브를 부여하는 체계가 필요하다. 그러나 현재 시스템 대부분은 이러한 보상 구조가 부족하여, 참여자가 적극적으로 학습에 기여하기 위한 동기를 얻기 어렵다.

이러한 문제점을 해결하는 방안으로, 연합학습에 블록체인 기술을 결합하는 방식이 제안되었다. 블록체인[4]은 탈중앙화된 구조를 기반으로 데이터를 암호화하고, 변경 불가능한 형태로 순차적으로 기록함으로써 데이터 조작에 대한 강력한 대응 수단을 제공한다. 또한 모든 참여자의 활동이 투명하게 기록되기 때문에, 연합학습 과정에서의 조작이나 부정행위를 사전에 방지할 수 있다. 나아가, 각 참여자의 기여도를 정량적으로 평가하고 그에 따른 보상을 제공함으로써 정직한 학습 참여를 유도하고 동기를 부여할 수 있다는 점에서 주목받고 있다. 그러나 이러한 구조 역시 블록 생성 및 전파 과정에서 통신 및 연산 오버헤드가 발생한다는 한계를 수반한다.

본 논문의 주요 기여는 다음과 같다.

첫째, 블록체인 기반 연합학습에서 순차적 학습으로 인한 지연 문제를 해결하기 위해 파이프라인을 적용한 병렬 처리 방식을 도입한다.

둘째, 기존의 합의 알고리즘을 변경하여 블록 분기를 줄이고 전체 연산량을 감소시킨다.

셋째, 제안한 기법의 성능 검증을 위하여 실제 시스템을 구현하고 실행을 수행한다. 실험 결과, 기존 방법 대비 최대 41%의 학습 속도 향상을 확인했다.

*Corresponding Author

Manuscript received July 4, 2025; revised October 19, 2025; accepted November 14, 2025

이희태: 한국항공대학교 항공전자정보공학과 석사과정(lheetae96@gmail.com, ORCID[®] 0009-0000-7153-3531)

이재환: 한국항공대학교 컴퓨터공학과 교수(jlee@kau.ac.kr, ORCID[®] 0000-0001-6248-9567)

※ 본 연구는 과학기술정보통신부 한국연구재단의 중견연구(NRF-2023R1A2C1005750)의 연구결과로 수행되었습니다.

본 논문은 총 5개 장으로 구성된다. 2장에서는 블록체인 기반 연합학습과 관련된 기존 연구를 살펴본다. 3장에서는 제안하는 시스템의 전체 구조와 구성 요소를 설명한다. 4장에서는 시스템의 효율성을 검증하기 위한 실험을 설계하고, 그 결과를 분석한다. 마지막으로 5장에서는 본 논문의 의의와 결론을 제시한다.

II. 관련 연구

1. 블록체인과 연합학습의 통합 연구

블록체인 기반 연합학습은 중앙서버에 대한 의존도를 줄이고, 보상 체계를 통해 참여를 유도하며, 모델 업데이트의 무결성과 투명성을 확보하는 방안으로 주목받고 있다. BlockFL [5]은 블록체인 네트워크를 통해 모델 파라미터를 분산 저장하고, 통신 비용이 전체 학습 시간에 미치는 영향을 분석함으로써 본 분야의 기초를 마련하였다.

이후 연구들은 블록체인 기반 연합학습의 한계로 합의 지연, 네트워크 확장성 부족, 기여도 평가의 어려움 등을 지적하였으며, 이를 극복하는 방안으로 경량 합의 알고리즘, 보상 메커니즘 최적화, 프라이버시 보호 기술 등이 제안되었다[6]. 한편, 블록체인 설계 요소의 구조적 관계를 분석한 연구[7]에서는 효율성(efficiency), 프라이버시(privacy), 공정성(fairness)이라는 평가 지표를 기준으로 다양한 시스템 구조를 비교하고, 이들 설계 간의 상호작용을 체계화하였다. 또한 학습 기여도 기반의 보상 체계[8] 및 신원 검증 기반 참여 구조를 제시함으로써, 참여율 향상과 시스템의 신뢰성 제고 방안을 제안한 바 있다.

이처럼 기존 연구들은 블록체인의 구조적 장점을 연합학습에 적용하면서도, 블록체인으로 인한 오버헤드와 지연을 최소화할 수 있는 최적화 기법의 필요성을 공통으로 강조하고 있다.

2. 블록체인 오버헤드 감소 연구

블록체인의 주요 단점 중 하나는 높은 연산 자원 소모 및 통신 비용으로 인해 연합학습의 효율성을 저해할 수 있다는 점이다. 이에 따라, 연산 효율성을 개선하고 시스템 지연을 완화하기 위한 연구들이 활발히 진행되고 있다.

LiteChain [9]은 이러한 문제를 해결하기 위해, 오픈체인 연산과 최적화된 합의 알고리즘을 도입하여 블록체인의 부담을 줄이는 방식을 제안한다. 블록 생성 및 합의 과정에서 발생하는 과도한 연산 비용을 줄이기 위해, 일부 모델 업데이트 과정을 블록체인 외부에서 처리함으로써 보안성과 효율성의 균형을 도모하였다.

또한 BlockDFL [10]은 완전한 P2P 기반 연합학습 프레임워크를 제안하여, 중앙서버 없이도 연합학습을 수행할 수 있는 구조를 설계하였다. 특히, PBFT (Practical Byzantine Fault Tolerance) 기반의 투표 메커니즘을 통해 참여자 간 학습 조정을 가능하게 하며, 학습 과정과 합의 과정을 세분화하여 연산 자원 소모를 최소화하였다.

III. 시스템 모델

본 시스템의 아키텍처는 [Record and Reward Federated Learning Contributions with Blockchain]에서 제시된 구조를 기반으로 하

며, 전체 구성은 그림 1에 나타내었다.

본 연구에서는 블록체인 기반 연합학습 시스템에서 발생하는 연산 지연(latency) 및 자원 활용의 비효율성 문제를 해결하기 위하여, 파이프라이닝(pipelining) 구조와 지분증명(proof of stake, PoS) [11]기반의 합의 메커니즘을 결합한 학습 프레임워크를 제안한다.

제안하는 시스템은 채굴자(miner)와 학습 노드(worker)로 구성되며, 채굴자는 중앙서버의 역할을 수행하면서 각 학습 노드의 상위 계층에 위치한다. 학습 노드는 로컬 데이터를 기반으로 모델을 학습한 후, 일정 주기의 에폭(epoch)이 완료될 때마다 모델 파라미터를 트랜잭션(transaction)의 형태로 채굴자에게 전송한다.

채굴자들은 수신한 트랜잭션에 대해 상호 교차 검증을 수행하며, 일정 조건(예: 누적 트랜잭션 수 또는 지정된 시간 경과)이 충족되면 블록을 생성하게 된다. 생성된 블록은 모든 채굴자에게 전파되고, 이후 해당 블록의 해시 정보가 학습 노드로 전달되면서 다음 학습 라운드가 시작된다.

1. 파이프라이닝 기반 처리 구조

그림 2는 본 논문에서 제안하는 파이프라이닝 구조의 시간 흐름을 도식화한 것이다. 기존 블록체인 기반 연합학습 시스템은 모델 학습, 트랜잭션 생성, 블록 생성 및 전파가 순차적으로 이루어지는 구조로, 하나의 단계가 완료되어야 다음 단계가 시작되는 처리 방식이다.

이러한 직렬적 구조에서는 블록 생성 및 전파 지연이 발생할 경우, 학습 노드가 글로벌 파라미터를 적시에 수신하지 못하게 되어, GPU자원이 유휴 상태로 대기하는 병목 현상이 발생하게 된다.

본 논문에서는 이러한 구조적 비효율성을 해소하기 위해, 모델 학습, 트랜잭션 전송, 블록 생성 및 전파 과정을 서로 독립된 스레드로 분리하고, 이를 병렬로 처리할 수 있는 파이프라이닝 구조를 설계하였다.

그림 2의 도식화에서 R_n 은 N번째 Round를 의미하며, B_h 는 블록의 헤더를, t_n 은 N번째 트랜잭션, $Block_n$ 은 n번째 Block을 의미한다.

상단 Main 영역은 채굴자의 작업 흐름으로, 채굴자는 주기적으로 트랜잭션을 수집·검증하여 블록을 생성한 뒤 블록헤더를 네트워크에 전파한다. 하단 Worker Node는 세 개의 스레드로 구성되며, Thread 1은 새로운 블록 헤더의 존재 여부를 비동기적으로 확인·수신하여 유효성을 검사한 후 큐에 적재한다. Thread 2는 로컬 학습 결과를 트랜잭션으로 직렬화하고 서명하여 전송한다. Thread 3은 최신 글로벌 모델을 바탕으로 로컬 학습을 수행하며, 학습이 종료되면 생성된 업데이트를 Thread 2에 학습완료 및 트랜잭션 요청 메시지를 전달하고, 이후 Thread 1이 수신한 최신 헤더가 있는지 확인하여 글로벌 모델을 갱신한다. 이러한 구성에 의해 세 스테이지는 시간 축에서 중첩되어 병렬적으로 동작한다. 결과적으로 학습-통신-블록 생성이 병렬로 진행되어 라운드의 학습과 그외 블록생성과정이 병렬적으로 처리가 되어서 GPU 유휴 시간이 최소화되어, 전체 처리 효율과 처리량이 함께 향상된다.

특히 제안된 파이프라이닝 구조에서는 하나의 라운드 R 동안 학습 노드는 트랜잭션을 생성 및 전송하고, 채굴자는 이전

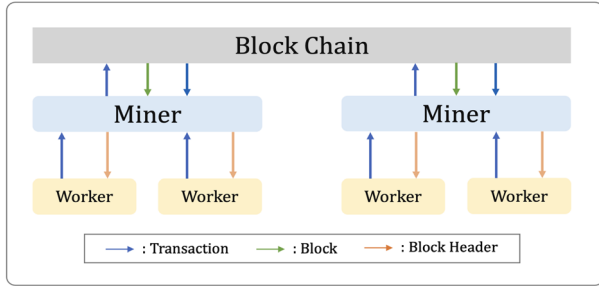


그림 1. 제안된 시스템의 구조도.

Fig. 1. Diagram of the system.

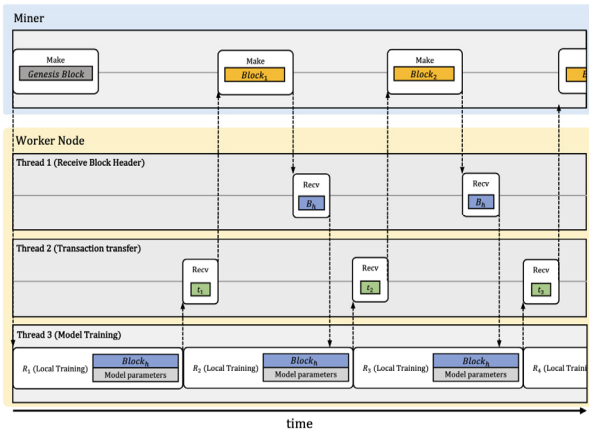


그림 2. 파이프라이닝 방식의 동작.

Fig. 2. Pipelining operation.

$$T_{BGI} = (N_{fork} \max(T_{local}, T_{ST} + T_{BGT} + T_{BP}) + T_{BD}) \quad (1)$$

라운드 R-1의 트랜잭션을 기반으로 블록을 생성하며, 동시에 학습 노드는 직전 라운드에서 수신한 글로벌 모델을 바탕으로 다음 라운드 R+1의 학습을 준비할 수 있다.

이처럼 학습, 통신, 블록 생성을 시간상으로 분산시켜 병렬 화합으로써, GPU 자원의 유휴 시간을 최소화하고 전체 시스템의 처리 효율성과 처리량을 향상시킬 수 있다. 결과적으로 제안된 파이프라이닝 구조는 블록체인 연합학습 시스템에서 발생하는 구조적 병목을 효과적으로 해소하고, 학습의 안정성과 자원 활용 최적화를 동시에 달성할 수 있는 기반을 제공한다.

또한, 모델 수렴의 안정성을 높이기 위해 모멘텀 기반 집계 (momentum-based aggregation) 기법을 함께 도입하였다. 이 기법은 단순히 현재 라운드에서 수집된 클라이언트 모델만을 평균하는 방식이 아니라, 이전 라운드의 글로벌 파라미터를 함께 고려하여 집계를 수행함으로써, 업데이트 간의 급격한 변화로 인한 불안정성을 완화한다. 구체적으로, 클라이언트 수가 n 일 때, 제출된 n 개의 로컬 모델 파라미터와 이전 라운드의 글로벌 파라미터를 포함한 총 $n+1$ 개의 파라미터에 평균 연산을 적용함으로써, 더 부드럽고 안정적인 수렴 경로를 유도할 수 있다. 이는 파이프라이닝 구조에서 병렬 학습이 진행되는 동안에도 전체 모델 학습의 일관성과 성능 안정성을 유지하는 데 기여한다.

수식 (1)은 제안하는 시스템의 블록 생성 주기(round cycle time)를 정의한 것으로, 기존 BlockFL 구조에서 사용된 주기

수식을 기반으로 하되, 각 처리 단계가 병렬로 수행되는 파이프라이닝 구조의 특성을 반영하여 구성되었다. 특히, 병렬 실행이 가능한 연산 구간에서는 전체 주기를 결정짓는 병목 구간만을 고려해야 하므로, 해당 항목들에 대해 최댓값(max) 연산을 적용하였다. 이를 통해 파이프라이닝 구조에서의 블록 생성 지연을 이론적으로 분석할 수 있으며, 기존 구조 대비 자원 활용의 효율성을 수식적으로 비교하는 근거를 제공한다.

우선, T_{local} 은 각 클라이언트 노드에서 수행되는 로컬 학습 시간으로 정의된다. 이것은 사용자의 데이터를 기반으로 개별 모델을 학습하는 데 걸리는 시간이다. 로컬 학습이 완료되면, 클라이언트는 모델 업데이트 정보를 블록체인 네트워크에 전송하고, 이때 발생하는 통신 지연은 T_{ST} 로 나타낸다.

이어서, 채굴자는 수신된 트랜잭션을 바탕으로 블록을 생성하며, 이 과정에 걸리는 시간은 T_{BGT} 으로 정의된다. 생성된 블록이 전체 네트워크에 전파되는데 걸리는 시간은 T_{BP} , 그리고 각 노드가 해당 블록을 수신하여 저장하는 데 걸리는 시간은 T_{BD} 로 각각 구현된다. 이들 각 단계는 하나의 연합학습 라운드를 구성하는 핵심 요소이며, 각 지연은 전체 학습 속도에 직접적인 영향을 미친다.

한편, 블록체인의 구조적 특성으로 인해 블록 분기(fork)가 발생할 수 있으며, 이는 일시적으로 잘못된 블록이 채택되었다가 무효가 되는 상황을 의미한다. 이에 따라 일부 라운드는 재수행이 요구될 수 있으며, 해당 빈도는 N_{fork} 로 정의된다. 블록 분기는 전체 학습 지연을 증가시키는 비효율적 요소로 작용하며, 시스템의 안정성과 효율성 측면에서 중요한 고려 대상이 된다.

기존 BlockFL 구조에서는 위의 모든 단계가 순차적으로 수행되며, 특정 단계에서 지연이 발생할 때 전체 라운드 진행이 지체되어 연산 자원, 특히 GPU가 유휴 상태로 남는 병목 현상이 발생한다. 이러한 문제를 해결하기 위해, 본 연구에서는 T_{BGT} 을 병렬적으로 수행할 수 있도록 파이프라이닝 구조를 설계하였다. 이 구조는 이전 라운드에서의 블록 생성 및 전파가 진행되는 동안, 다음 라운드의 로컬 학습이 동시에 시작될 수 있도록 구성되어 전체 라운드 주기를 단축하는 효과를 갖는다.

파이프라이닝 구조는 시스템 내 연산 및 통신 과정을 유기적으로 병렬화함으로써 자원 활용도를 극대화하며, 특히 대규모 데이터셋 또는 복잡한 모델을 사용하는 연합학습 환경에서 더욱 높은 효율성을 발휘할 수 있다. 결과적으로, 제안된 수식은 기존 BlockFL의 순차 처리 구조가 가지는 한계를 극복하고, 병렬 처리 기반 구조의 장점을 이론적으로 입증함으로써 시스템 처리량 향상 및 자원 최적화 가능성을 명확히 보여준다.

2. Proof of Stake 알고리즘

기존의 작업증명 PoW (Proof of Work) 기반 블록체인 구조는 블록 생성을 위해 막대한 연산 자원을 요구하며, 이에 따라 전체 시스템의 처리 속도가 저하되고, 에너지 소비 측면에서도 비효율적인 문제를 일으킨다. 특히, 시스템의 연산량이 증가할 때 연산 자원을 고려하여 블록 생성 난이도를 인위적으로 조정하게 되면, 오히려 블록 분기가 발생할 가능성이 증가

하는 문제가 존재한다.

본 연구에서는 이러한 문제점을 해결하고자, 연산 경쟁이 아닌 참여자의 지분 기반으로 블록 생성 권한을 분배하는 PoS (Proof of Stake) 기반의 경량 합의 구조를 설계하였다. 제안하는 구조에서는 각 채굴자에 소속된 학습 노드의 수를 지분으로 정의하며, 이를 기반으로 블록 생성 권한을 분배함으로써 전체 연산 자원을 절감하는 동시에, 블록 분기 발생 가능성을 효과적으로 낮추도록 구성하였다. 다만 PoS는 지분 집중에 따른 해당 중앙화 문제와 지분 가중치 데이터의 품질-공헌도와 불일치할 수 있는 한계를 내포한다. 본 연구는 이러한 이슈의 설계와 평가는 범위 밖으로 두고, 지분을 ‘학습 노드 수’로 고정된 통제된 실험 환경에서 파이프라이닝과 합의 경량화의 시스템적 성능 효과에만 초점을 맞춘다.

PoS 기반 구조에서는 PoW처럼 무한 반복적인 해시 계산 과정을 거치지 않으며, 채굴자는 자신의 클러스터로부터 일정 수의 학습 파라미터가 수집되었을 때만 블록 생성을 위한 해시 계산을 수행한다. 이는 블록 생성의 주체가 사전에 결정되며, 블록 경쟁 없이 전개되는 PoS의 구조적 특성에 기반한 것으로, 전체 연산 부하를 크게 줄이는 동시에, 블록 분기 가능성 또한 현저히 감소시킨다.

또한 채굴자는 각 학습 노드로부터 전달받은 트랜잭션을 취합하여 블록을 생성하게 되며, 이때 트랜잭션은 다음과 같은 필수 요소들로 구성된다:

- Client ID : 노드에게 무작위로 부여된 고유한 식별자
- Base index : 해당 트랜잭션이 속한 블록 인덱스 번호
- Update model : 업데이트하고자 하는 모델 파라미터
- Data size : 학습하는 이미지의 데이터 크기
- Computing time : 트랜잭션 생성 시간
- Miner : 해당 노드가 속한 채굴자의 주소

또한, 본 구조에서는 채굴자의 지분을 해당 채굴자에 소속된 학습 노드의 수로 정의한다. 이는 채굴자의 연합학습 기여도를 반영하여 블록 생성 권한을 분배하는 방식으로, 자원 낭비 없이 공정하고 효율적인 합의를 가능하게 한다.

결국, 제안하는 PoS 기반 합의 구조는 전체 연산량을 절감하고 블록 생성의 안정성을 높임으로써, 블록체인 기반 연합학습의 처리 효율성과 확장성을 향상한다. 특히 기존 PoW 구조에서 발생하던 블록 경쟁 및 분기 문제를 완화하여, 네트워크의 일관성과 안정성을 효과적으로 확보할 수 있다.

V. 실험 및 결과

1. 실험환경

본 실험은 제안한 파이프라이닝 기반 블록체인 연합학습 구조의 성능을 실증적으로 검증하고, 블록 생성 지연 감소가 전체 학습 과정 및 모델 성능에 미치는 영향을 분석하기 위해 설계되었다. 실험은 실제 분산 환경에서 수행되었으며, 제안 기법의 실용성과 안정성을 평가하는 데 중점을 두었다.

특히, 파이프라이닝 기법이 블록 생성 시간(T_{BGT})에 미치는 영향을 중점적으로 분석하고, 이를 통해 전체 학습 주기의 단축 효과 및 최종 모델의 정확도 향상 여부를 정량적으로 측정하였다. 또한, 제안한 PoS 기반 블록 생성 구조가 연합학습 주기 내에서 정상적으로 작동하며, 안정적인 합의 및 모델 공

유가 이루어지는지를 함께 검증하였다.

실험에 사용된 데이터셋은 실제 거리 주소 이미지로 구성된 SVHN (Street View House Numbers) [12]와, 10개 클래스 이미지 분류용 벤치마크인 CIFAR-10 [13]이다. 두 데이터셋은 모두 64×64 크기의 RGB 색상 이미지로 리사이징하여 ResNet-18 [14] 모델에 입력되도록 전처리하였으며, 모델 학습은 각 데이터셋을 기반으로 독립적으로 수행되었다.

학습에 사용된 최적화 기법은 Adam [15] Optimizer이며, 학습률(learning rate)은 0.001로 설정하였고, 배치 사이즈는 32로 고정하였다. 모델의 학습 성능은 테스트 데이터셋에 대한 정확도(accuracy)를 기준으로 평가하였으며, 이를 통해 파이프라이닝 기법 적용 전후의 성능 차이를 정량적으로 분석하였다.

데이터는 IID (Independent and Identically Distributed) 형태로 각 노드에 균등하게 분할되어 할당되었으며, 노드 간 데이터 분포의 편차를 제거한 상태에서 실험이 진행되었다. 전체 실험은 총 5대의 서버와 12개의 GPU (RTX 3090, 6대, RTX 3080 Ti 4대, RTX 3080, 2대)를 활용하였으며, 네트워크 환경은 10GbE (10기가비트 이더넷)를 기반으로 구성되었다.

시스템 전체는 총 17개의 노드로 구성되었으며, 이 중 5개는 채굴자 노드, 12개는 학습 노드로 할당되었다. 각 채굴자는 기본적으로 2개의 학습 노드를 관리하며, PoS 기반 지분 구조의 검증을 위해 일부 채굴자에는 4개의 학습 노드를 연결하여 지분 차이가 블록 생성 우선순위에 영향을 미치는지를 확인하였다. 이러한 설정은 PoS 기반 블록 생성 구조가 실제 환경에서도 안정적으로 작동하며, 기여도에 기반한 블록 생성 권한 분배가 유효함을 확인하기 위한 실험적 장치로 활용되었다.

2. GPU 사용량 측정

제한한 파이프라이닝 구조가 실제로 연산 자원(GPU)의 활용도에 미치는 영향을 정량적으로 분석하였다. 이를 위해 ResNet-18 모델과 SVHN 데이터셋을 사용하여 블록체인 기반 연합학습을 실행하고, NVIDIA Nsight 도구를 활용하여 GPU 연산 흐름을 시각화하였다. 파이프라이닝 구조 적용 전후의 GPU 사용 상태를 비교함으로써, 제안 기법의 연산 효율성 개선 효과를 실증적으로 평가하였다.

그림 3은 기존 순차 처리 방식에서의 GPU 사용 현황을 나타낸다. 주황색으로 표시된 구간은 실제 연산이 수행된 시간대이며, 각 연산 구간 사이에 반복적으로 나타나는 공백 구간은 GPU가 연산을 수행하지 않고 대기 중인 유휴 상태(Idle state)를 의미한다. 이러한 유휴 시간은 블록체인의 합의, 생성, 전파 과정에서 발생하는 지연으로 인해 발생하며, 전체 연합학습 주기에서 자원의 낭비로 작용하게 된다.

반면, 그림 4는 같은 조건에서 파이프라이닝 구조를 적용한 후의 GPU 사용 현황을 나타낸다. 기존 구조와 비교하였을 때, GPU 유휴 구간이 현저히 감소하였으며, 연산 처리율이 일정하고 지속적인 형태로 유지되는 것을 확인할 수 있다. 이는 모델 학습, 트랜잭션 전송, 블록 생성 및 전파 과정이 병렬로

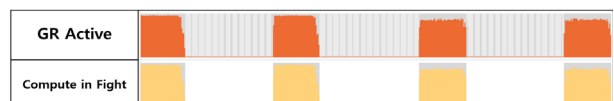


그림 3. 파이프라이닝 비적용 시 GPU 사용량.

Fig. 3. GPU utilization without pipelining.

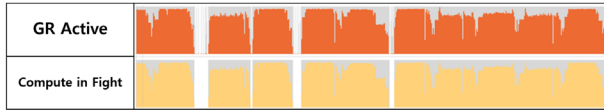


그림 4. 파이프라이닝 적용 시 GPU 사용량.

Fig. 4. GPU utilization with pipelining.

처리됨에 따라, GPU가 다음 작업을 기다리며 대기하는 시간이 제거되고, 전체 연산 자원이 효율적으로 활용되고 있음을 의미한다.

이러한 결과는 제안하는 파이프라이닝 구조가 시스템 자원 활용 최적화를 실현할 수 있음을 보여주는 실증적 근거이며, 블록체인 기반 연합학습 환경에서 실질적인 성능 향상을 유도할 수 있음을 입증한다. 특히 학습 주기의 병목이 제거되면서 동일 시간 내 더 많은 학습 라운드를 수행할 수 있는 기반이 마련되었음을 확인할 수 있다.

3. 시간에 따른 정확도 비교 실험

파이프라이닝 구조가 학습 성능에 미치는 영향을 분석하기 위해, 본 절에서는 기존의 순차형 연합학습 방식과 본 연구에서 제안한 파이프라이닝 방식 간의 정확도 수렴 속도 및 블록 생성 시간 차이를 비교한다. 실험은 앞서 기술한 시스템 구성 및 네트워크 환경과 동일한 조건에서 수행되었으며, 네트워크는 물리 NIC 기준으로 순간 최대 속도가 초당 10메가비트 이하에 불과했으며, 이는 10기가비트 링크 용량의 0.1% 미만이다. 따라서 본 실험 규모에서는 네트워크가 병목으로 작용하지 않았고, 두 방식의 성능 차이는 통신 대기보다는 파이프라이닝에 따른 작업 중첩과 GPU 자원 활용 효율에서 비롯된 것이다. 채굴자와 학습 노드의 구성, 학습 노드 수의 분포, PoS 기반 블록 생성 구조 역시 일관되게 유지되었다.

비교 실험은 10회의 로컬 학습 후 1회의 블록 생성을 수행하는 라운드를 기준으로 총 20회를 반복하는 방식으로 설계되었으며, 모델은 ResNet-18을, 데이터셋은 SVHN과 CIFAR-10을 각각 적용하였다. 모든 실험 조건을 동일하게 설정한 상태에서, 파이프라이닝 구조의 적용 여부에 따른 성능 차이를 정량적으로 분석하였다.

그림 5와 그림 6은 각각의 데이터셋에서 파이프라이닝 구조의 적용 전후에 따른 정확도 수렴 속도를 비교한 결과이다. 실험 결과, 두 방식 모두 유사한 수렴 속도를 보였으며, 이는 파이프라이닝 구조가 학습의 안정성이나 정확도에 부정적인 영향을 미치지 않음을 의미한다. 즉, 학습 품질을 유지하면서도 처리 성능을 향상시킬 수 있는 구조임을 실험적으로 확인할 수 있었다.

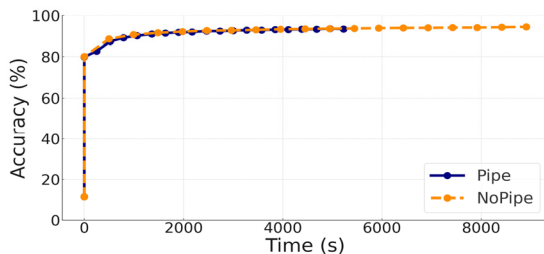


그림 5. SVHN 데이터셋에서의 test 정확도 비교 결과.

Fig. 5. Test Accuracy Comparison on SVHN dataset.

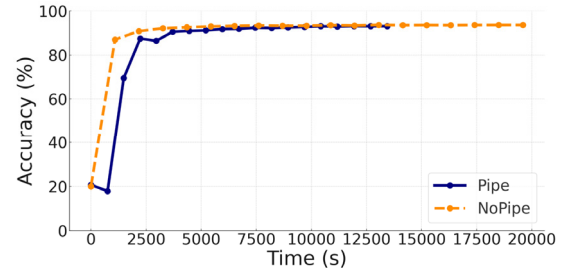


그림 6. CIFAR-10 데이터셋 실험 비교 그래프.

Fig. 6. Test accuracy comparison on CIFAR-10 dataset.

표 1. 파이프라이닝 도입에 따른 성능 향상 분석

Table 1. Performance analysis with Pipelining integration

	Pipelined (s)	Non-pipelined (s)	차이
CIFAR-10	14356	19592	26%
SVHN	5223	8904	41%

한편, 표 1은 동일 조건에서 측정된 블록 생성 시간의 평균을 비교한 결과를 나타낸다. 실험 결과, 파이프라이닝 구조를 적용하였을 때 블록 생성 주기가 SVHN 데이터셋에서는 약 26%, CIFAR-10 데이터셋에서는 약 41% 단축된 것으로 나타났다. 이는 병렬 처리 구조를 통해 GPU 유휴 시간을 줄이고, 전체 시스템 자원을 보다 효율적으로 활용한 결과로 해석할 수 있다.

특히 데이터 크기가 크거나 모델의 연산 복잡도가 높은 환경에서는, 이와 같은 구조적 최적화가 전체 학습 처리 시간을 단축하게 하는 데 큰 효과를 발휘할 수 있다. 따라서 파이프라이닝 기법은 향후 대규모 연합학습 시스템에서 성능을 극대화하기 위한 핵심 구성 요소로 활용될 수 있을 것으로 기대된다.

VI. 결론

본 연구에서는 블록체인 기반 연합학습 환경에서 발생하는 학습 지연과 연산 자원 활용의 비효율 문제를 해결하기 위해, 파이프라이닝 기법을 적용한 학습 프레임워크를 제안하고 이를 실험적으로 검증하였다.

블록체인을 연합학습에 통합할 경우, 블록 생성 및 전과 과정에서 시간 지연으로 인해 학습 속도가 저하되고, 연산 자원 및 네트워크 자원이 유휴 상태로 남게 되는 구조적 병목 현상이 발생한다. 본 연구에서 제안한 파이프라이닝 기법은 모델 학습, 트랜잭션 전송, 블록 생성 및 전과 과정을 병렬로 수행할 수 있도록 구조화함으로써, 이러한 병목을 해소하고 전체 학습 주기를 단축하는 것을 목표로 하였다.

실험 결과, 파이프라이닝 구조를 적용했을 때 GPU 활용도가 유의미하게 향상되었으며, 동일한 모델과 데이터셋 조건에서 약 30%의 학습 속도 개선 효과를 확인할 수 있었다. 이는 연산 및 통신 자원의 유휴 시간을 효과적으로 줄임으로써 전체 학습 프로세스를 가속화할 수 있음을 실증적으로 보여주는 결과이다.

결론적으로 본 연구에서 제안한 파이프라이닝 기반 학습 프레임워크는 블록체인 기반 연합학습 환경에서도 적용할 수 있는 효율적인 구조이며, 학습 성능의 실질적인 향상과 시스

템 자원 활용 최적화를 동시에 달성할 수 있는 효과적인 방법임을 입증하였다. 향후 대규모 분산 학습 시스템에서도 본 구조가 확장 가능하고 실용적인 기반으로 활용될 수 있을 것으로 기대된다.

REFERENCES

- [1] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," *Proc. of the 20th International Conference on Artificial Intelligence and Statistics*, Fort Lauderdale, FL, USA, vol. 54, pp. 1273-1282, Apr. 2017.
- [2] Y. Lee, E. Lee, D. Lee, and Y. Hwang, "Server-edge device based efficient deep learning network retraining framework," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 30, no. 10, pp. 1044-1050, 2024.
- [3] S. Lee and H. S. Kim, "A linear matrix inequality-based relaxed time-delay platooning controller design for multiple autonomous vehicles," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 29, no. 11, pp. 841-846, Nov. 2023.
- [4] S. Nakamoto, *Bitcoin: A Peer-to-peer Electronic Cash System*, Bitcoin.org, White Paper, 2008.
- [5] H. Kim, J. Park, M. Bennis, and S.-L. Kim, "Blockchain-based on-device federated learning," *IEEE Communications Letters*, vol. 24, no. 6, pp. 1279-1283, Jun. 2020.
- [6] D. Hou, J. Zhang, K. L. Man, J. Ma, and Z. Peng, "A systematic literature review of blockchain-based federated learning: architectures, applications and issues," *Proc. of International Conference on Information and Communication Technology Convergence*, Jeju, Korea, pp. 302-307, Oct. 2021.
- [7] Y. E. Oktian and S.-G. Lee, "Blockchain-based federated learning system: a survey on design choices," *Sensors*, vol. 23, no. 12, p. 5658, Jun. 2023.
- [8] I. Martinez, S. Francis, and A. S. Hafid, "Record and reward federated learning contributions with blockchain," *Proc. of International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery*, Guilin, China, pp. 50-57, Oct. 2019.
- [9] H. Chen, R. Zhou, Y.-H. Chan, Z. Jiang, X. Chen, and E. C. H. Ngai, "LiteChain: A lightweight blockchain for verifiable and scalable federated learning in massive edge networks," *IEEE Transactions on Mobile Computing*, vol. 24, no. 3, pp. 1928-1944, Mar. 2025.
- [10] Z. Qin, X. Yan, M. Zhou, and S. Deng, "BlockDFL: A blockchain-based fully decentralized peer-to-peer federated learning framework," *Proc. of the ACM Web Conference 2024*, Singapore, pp. 2914-2925, May 2024.
- [11] S. King and S. Nadal, "PPCoin: Peer-to-peer cryptocurrency with proof-of-stake," PPCoin.org, White Paper, Aug. 2012.
- [12] I. J. Goodfellow, Y. Bulatov, J. Ibarz, S. Arnaud, and V. Shet, "Multi-digit number recognition from street view imagery using deep convolutional neural networks," arXiv preprint arXiv:1312.6082, Dec. 2013.
- [13] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," University of Toronto, Toronto, Canada, Tech. Rep., 2009.
- [14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, USA, pp. 770-778, Jun. 2016.
- [15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *Proc. of International Conference on Learning Representations*, San Diego, USA, May 2015.



이 희 태

한국항공대학교 항공전자정보공학과 석사. 관심분야는 분산컴퓨팅, 빅데이터 플랫폼, 고성능 컴퓨팅.



이 재 환

서울대학교 전기공학부 공학사. 서울대학교 전기공학부 공학석사. 메릴랜드 주립대 컴퓨터과학과 박사. 현재 한국항공대학교 컴퓨터공학과 교수. 관심분야는 분산컴퓨팅, 고성능 컴퓨팅, 빅데이터 플랫폼, 시스템SW.