디지털산업정보학회 논문지 제19권 제3호-2023년 9월
http://dx.doi.org/10.17662/ksdim.2023.19.3.233

연합학습 기반 자치구별 건물 변화탐지 알고리즘 성능 분석*

김 영 현**

Performance Analysis of Building Change Detection Algorithm for Each Autonomous District Based on Federated Learning

Kim Younghyun

〈Abstract〉

Although artificial intelligence and machine learning technologies have been used in various fields, problems with personal information protection have arisen based on centralized data collection and processing. Federated learning has been proposed to solve this problem. Federated learning is a process in which clients who own data in a distributed data environment learn a model using their own data and collectively create an artificial intelligence model by centrally collecting learning results. Unlike the centralized method, Federated learning has the advantage of not having to send the client's data to the central server. In this paper, we quantitatively present the performance improvement when federated learning is applied using the building change detection learning data. As a result, it has been confirmed that the performance when federated learning was applied was about 29% higher on average than the performance when it was not applied. As a future work, we plan to propose a method that can effectively reduce the number of federated learning rounds to improve the convergence time of federated learning.

Key Words : Artificial Intelligence, Federated Learning, Performance Analysis

Ⅰ. 서론

최근 컴퓨터 하드웨어 성능과 소프트웨어 기술의 발전으로 인공지능 (artificial intelligence: AI)과 머신러닝 (machine learning: ML)이 다양한 분야에서 활용되고 있다[1]. 그러나 이러한 사례들은 대부분 중앙집중식 데이터 수집과 처리를 기반으로 한다. 이런 방식은 효과적인 모델 학습을 가능케 하지만, 사용자들의 개인정보와 같은 민감한 데이터가 중앙 서버에 집중되어 수집되는 문제점이 있다. 이에 따라, 개인정보 보호와 관련한 우려를 증대시키며, 데이터의 보안 문제를 야기할 수 있다.

연합학습은 이러한 문제를 해결하기 위해 제안된 방식이다[2]. 연합학습은 분산된 데이터 환경에서 데이터를 소유한 클라이언트가 각자의 데이터를 이용하여 모델을 학습하고, 학습 결과를 중앙에서 취합하

* 본 연구는 서울디지털재단 연구과제로 수행되었습니다.
** 서울디지털재단 AI · 빅데이터팀 수석 (교신저자)

는 방식으로, 인공지능 모델을 공동으로 생성해 나가는 과정을 의미한다. 이는 중앙 집중식 방식과 달리, 클라이언트의 데이터를 중앙 서버로 집중적으로 보내지 않아도 되는 장점을 가지고 있다.

본 논문에서는 연합학습 알고리즘을 적용하였을 때의 성능 향상을 정량적으로 제시한다. 이를 위해 서울시의 건물 변화탐지 학습 데이터를 활용한다. 그리고 학습 데이터를 자치구 단위로 나눈 후에, 자치구별로 건물 변화탐지 알고리즘을 개발한다. 즉, 자치구별로 나눠진 데이터를 개별적으로 각각 학습하여 개발한 인공지능과, 연합학습을 적용하여 자치구별 데이터를 개별적으로 각각 학습한 인공지능의 성능을 도출하여, 연합학습 이전의 인공지능과 이후의 인공지능 성능을 비교 및 분석한다. 그 결과, 연합학습을 적용했을 때의 성능이 평균적으로 0.66에서 0.85로 약 29% 높아진 것을 확인할 수 있다.

본 논문의 구성은 다음과 같다. 2장과 3장에서는 연합학습의 개념과 동작과정, 연합학습 관련연구에 대해 각각 묘사한다. 그리고 4장에서는 연합학습의 성능 분석을 위해 필요한 데이터와 학습 모델에 대해 서술한다. 마지막으로 5장과 6장에서는 각각, 연합학습을 적용했을 때의 성능 향상을 정량적으로 보여주고, 본 논문의 결론을 지으면서 마무리한다.

II. 연합학습 개념

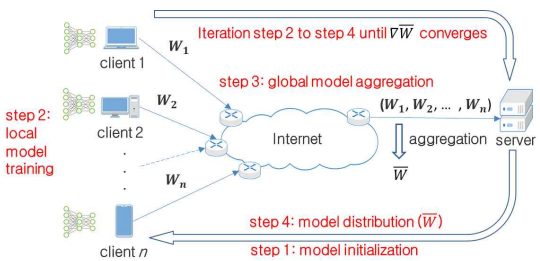
연합학습 (federated learning)은 기계 학습 방법론 중 하나로, 분산된 여러 개의 기기나 시스템에서 각자가 가진 데이터를 학습하여, 학습 결과를 중앙 서버로 전송하는 방식이다[3, 4]. 중앙 서버에서는 하위 클라이언트들로부터 수신한 학습 결과인 가중치들의 평균을 계산하여, 다시 클라이언트들에게 보낸다. 이러한 과정을 반복하여 연합학습 과정이 마무리되는데, 중앙 집중화된 데이터 저장소로 데이터를 수집하는 대신, 로컬 장치에서 데이터를 유지한 상태로 학습을 진행하는 것이 연합학습의 특징이다.

연합학습은 클라이언트들의 데이터를 보호하면서 중앙 서버가 데이터를 수집하는 것을 방지할 수 있다. 대신, 클라이언트 기기에서 데이터를 유지한 상태로 모델을 학습하고, 학습된 모델의 결과 정보만을 중앙 서버로 전송한다. 그리고 중앙 서버에서는 이러한 업데이트를 종합하여 전체적인 모델을 개선시킨다.

이에 따라, 연합학습은 보안과 개인 정보 보호, 데이터 전송량 감소 등의 이점을 제공한다. 개별 데이터는 로컬 장치에 안전하게 보관되므로 중앙 서버와의 통신 과정에서 데이터를 해킹하려는 시도를 막을 수 있다. 또한, 데이터 전송량이 감소하므로 네트워크 부하를 줄일 수 있다.

연합학습은 주로 모바일 기기, 센서 네트워크, 의료 기기 등과 같이 데이터가 분산된 환경에서 활용할 수 있다. 예를 들어, 로컬에서는 자신의 스마트폰 개인 데이터를 사용하여 학습하고, 학습 결과만 중앙 서버에 전송한다. 그리고 중앙 서버에서는 여러 스마트폰 사용자들로부터 수신한 학습 결과를 취합하여 인공지능 모델을 개발한다. 이렇게 함으로써 사용자들은 개인 정보를 보호할 수 있으면서 전체적인 모델의 성능을 향상시킬 수 있다.

구체적인 연합학습의 동작과정은 아래의 <그림 1>과 같다. <그림 1>에서 볼 수 있듯이, 연합학습 과정을 관장하는 서버가 존재하고, 서버와 통신하는 클라



<그림 1> 연합학습 동작과정

이언트들이 n개 존재한다고 가정한다.

첫 번째 단계로, 서버는 클라이언트들이 각자의 데이터를 이용하여 훈련할 수 있도록, 사전 정보를 클라이언트들에게 전송한다. 사전 정보는 훈련을 실시하기 위한 모델의 종류, 최적화 알고리즘 (optimizer algorithm), 학습률 (learning rate), 배치 (batch) 크기, 에포크 (epoch) 등을 포함한다. 두 번째 단계에서는, 각 클라이언트가 자체 데이터를 이용하여 딥러닝을 수행하여 그 결과물인 가중치 (weight)를 서버에게 전송한다.

세 번째 단계에서, 클라이언트들로부터 가중치 정보를 수신한 서버는, 가중치들의 평균값 $\overline{W}$를 계산한다. 이 때, 클라이언트들의 컴퓨팅 환경이 유사하다면, 서로 비슷한 시기에 가중치 정보를 클라이언트들로부터 수신할 수 있다. 반면에 클라이언트들의 하드웨어 성능이 각각 다르고, 하드웨어 사용 환경 등이 다를 경우, 가중치 정보를 수신하는 시간은 제각각일 것이다. 따라서 높은 정확도를 가진 모델을 유도하기 위해서는 충분한 시간을 갖고 클라이언트들의 응답을 기다려야 하고, 반면에 빠르게 모델을 생성하기 위해서는 적절한 시간 내에 응답이 오지 않은 클라이언트는 무시해야 한다. 이러한 결정은 연합학습을 수행하는 환경에 따라 정책적으로 결정하여, 최소한의 응답 지연 시간을 정의할 수 있다.

네 번째 단계에서 서버는 연합학습 단계를 종료할지, 아니면 다음 라운드로 진행하여 한 번 더 실시할지 결정한다. 결정하는 방법으로는, 이전 단계에서 도출한 가중치 $\overline{W}_{pre}$와 현재 단계에서 도출한 가중치 $\overline{W}$의 차이가 일정 수준 이하로 내려가면 $\left(\overline{W}_{pre} - \overline{W} < \epsilon\right)$ 연합학습을 종료한다. 그렇지 않으면 가중치들의 평균값 $\overline{W}$를 클라이언트들에게 전송하여 다시 한 번 로컬 훈련을 실시한다. 다시 말해, 이전 라운드와 현재 라운드 사이의 성능 향상이 일정 수준 이하일 때까지 알고리즘을 반복 (iteration)하는 것이므로, 가중치 $\overline{W}$ 뿐만 아니라, *loss* 등의 성능역시 연합학습의 종료 여부를 결정하는 요소가 될 수 있다.

Ⅲ. 연합학습 관련연구

연합학습은 의료, 사물인터넷 (Internet of Things: IoT), 개인정보보호, 교통, 추천 시스템 등 다양한 응용 분야에서 연구가 진행되고 있다. 특히, 의료 부분이 데이터의 개인정보보호에 민감한 영역으로서 많은 연구 논문이 발표되고 있다. Sohan[5]은 연합학습을 활용한 의료 영상 분석 분야의 최신 연구 동향을 제시한다. 해당 논문에서는 연합학습이, 의료 영상 분석 분야에서 유망한 접근 방식이지만, 더 효율적인 통신 프로토콜, 정확도를 향상시키는 방안, 이질적인 데이터를 다루는 방안 등 해결해야 할 과제가 여전히 남아있음을 제시한다. Nguyen[6]은 스마트 헬스케어 분야에서 연합학습이 활용되는 응용 사례와 기술 사례를 제시한다. 구체적으로, 전자 건강 기록 (electronic health record: EHR) 관리, 원격 건강 모니터링, 의료 영상 및 코로나19 감지를 포함하여 스마트 헬스케어에서의 연합학습 주요 응용들을 제시한다. 그리고 스마트 헬스케어에 연합학습을 적용하는 것은 아직 초기 단계이며, 지능적이고 개인 정보 보호가 강화된 의료 서비스를 제공하기 위해 향후 몇 년 안에 빠르게 성숙해질 것으로 언급한다.

스마트 시티에는 미세먼지, 온도, 차량 인식 등 다양한 환경 요소와 객체를 탐지하는 사물인터넷이 곳곳에 설치되어 있다. 전통적으로 인공지능을 개발하기 위해서는, 사물인터넷에서 수집하는 데이터들을 중앙으로 보내야 하는데, 이 때 높은 네트워크 비용이 발생한다. 이러한 경우, 중앙으로 데이터를 보내는 것이 아니라, 엣지에서 연합학습을 활용하여 인공지능을 개발할 수 있다. Pandya[7]는 교통 관리, 스마트 그리드, 재난/안전 관리, 스마트 빌딩 등 사물인터넷을 이용한 연합학습 활용 사례를 소개한다. 구체적으로, 교통량 데이터를 수집하여 교통량을 예측하고, 교통량을 조절하고 교통 체증을 줄일 수 있는 응용, 공기질 데이터를 수집하여 공기질을 모니터링하고 예

측하는 응용, CCTV 영상 데이터를 이용하여 보안을 감시하고 범죄를 예방 응용 등을 소개한다.

반면에 Zhang[8]은 연합학습을 사물인터넷에 적용하기 위해 해결해야 할 과제들을 제시한다. 바로, 메모리와 GPU와 같은 엣지 디바이스의 제한된 하드웨어 리소스와 제한된 네트워크 대역폭, 사물인터넷 디바이스의 간헐적인 연결성, 외부 공격자로부터의 신뢰성 등이다. 엣지 디바이스는 컴퓨팅 측면뿐만 아니라 저장 및 데이터 액세스를 위한 메모리 측면에서도 리소스가 제한되어 있다. 리소스 제약이 있는 장치에서는, 외부 메모리에 대한 읽기 및 쓰기를 방지하기 위해, 메모리 액세스를 줄이고 데이터를 칩에 유지해야 한다. 그리고 상당한 양의 전력을 사용할 수 있고 CPU 및 GPU가 탑재된 서버와 달리, 내장형 프로세서가 탑재된 엣지 디바이스는 가용한 에너지가 제한되어 있어 하드웨어 성능을 더욱 제한한다.

현재까지 엣지 디바이스의 하드웨어 성능이 점점 더 강력해지고 있음에도 불구하고, 일부 딥러닝 모델을 엣지 디바이스에서 훈련하는 것은 여전히 시간이 많이 걸리고 비효율적이다. 예를 들어, 제한된 네트워크 대역폭과 관련하여서는, 대부분의 IoT 장치가 데이터 센터의 유선 네트워크 대역폭보다 대역폭이 훨씬 작은 무선 네트워크를 사용하여 통신하기 때문에 발생한다. 제한된 네트워크 대역폭은, 클라이언트와 서버 간의 통신을 비효율적으로 만들뿐만 아니라, 연합학습 로컬 업데이트를 서버와 공유하지 못하는 클라이언트를 발생시킨다. 그리고 이것은 대규모 사물인터넷 시나리오에서 연합학습의 배포 성능에 대한 병목 현상으로 작용한다. 사물인터넷 디바이스의 간헐적인 연결성과도 연관되는 문제이다. 이와 같은 문제들을, Zhang[8]은, 연합학습을 사물인터넷에 적용하기 위해 해결해야 할 과제로서 제시한다.

마지막으로, 보안 분야에서의 기밀성 (confidentiality), 무결성 (integrity), 가용성 (availability) 등과 연관된 연합학습 이슈를 Mothukuri[9]가 정리한다. 기밀성은 허락되지 않은 자가 데이터의 내용을 볼 수 없게 하는 것이고, 무결성은 허락되지 않은 자가 데이터를 위/변조를 할 수 없게 하는 것, 가용성은 접근 권한이 있는 자가 데이터를 언제나 사용할 수 있게 하는 것을 의미한다. 한편, Mothukuri[9]는 두 개의 큰 주제를 정의한다. 첫째는 악의적인 공격자에게, 연합학습 프로토콜의 클라이언트로서, 접근 권한을 얻을 수 있는 기회를 제공하는 것이다. 이에 따라, 악의적인 공격자가 하나 이상의 클라이언트를 제어하여 글로벌 모델을 조작할 수 있는 것이다. 둘째는 클라이언트에서 서버로 업로드된 모델을 기반으로 각 클라이언트의 훈련 데이터를 부분적으로 예측하는 것이다. Mothukuri[9]는 연합학습 시스템을 개발할 때, 위와 같은 두 개의 이슈들을 고려해야 한다고 제시한다.

Ⅳ. 시스템 모델

연합학습의 성능 분석을 수행하기 위해, 서울시에서 생성한 항공영상 도시건물 변화탐지 데이터를 활용한다. <그림 2>는 서울 시내의 같은 장소를 촬영한 정사영상 (ortho-images)을 보여준다. 정사영상은 위성, 항공 데이터 등을 활용하여 변화 탐지, 지도 제작에 활용할 수 있도록, 영상 내 모든 객체들이 수직방향에서 본 것과 같은 형태를 갖도록 보정한 것이다[10]. <그림 2>에서 왼쪽은 과거의 사진이고, 오른쪽은 현재의 사진이다. 그리고 변화가 발생한 부분에 대해 라벨링이 되어 있다. 예를 들어, <그림 2>의 왼쪽 그림에서 가운데에 3개의 직사각형으로 표시된 부분은, 과거에는 존재하였지만, 현재에는 사라진 건물이다. 그리고 오른쪽 그림에서 오른쪽 하단에 표시된 부분은 과거에는 존재하지 않지만 현재에는 새로 생성된 건물이다.

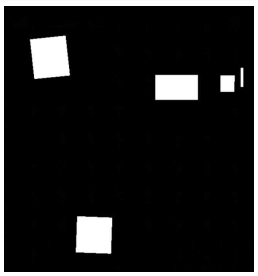
이처럼, 본 논문에서 활용한 학습 데이터 라벨링 정보는 4가지 유형으로 정의한다. 바로 신축, 소멸,

<그림 2> 학습 데이터 예시

갱신, 색상 변화이다. 4가지의 변화 유형별로 메타데이터를 json 파일에 삽입하고 변화된 부분의 좌표 정보도 함께 입력한다. 그리고 정답지 (ground truth: GT) 이미지는 json 파일의 좌표 정보를 이용해서 생성한다.

위와 같이 생성한 변화탐지 데이터셋은 (과거 이미지, 현재 이미지, 정답 이미지, 라벨링 데이터)로 구성한다. 예를 들어, <표 1>과 같이 과거 이미지와 현재 이미지를 분석하여 정답 이미지와 라벨링 데이터를 생성한다. 그리고 이와 같이, 생성한 학습 데이터의 수는 40,750개이다. 40,750개의 학습 데이터를 위치에 따라 각 자치구로 나누고, 자치구별 데이터 개수는 <그림 3>[1]에서 보여준다. 이 때, <그림 3>에서 묘사하는 데이터 개수는 각 자치구에서 변화된 모든 건물의 수를 나타내지 않는다. 단지, 무작위로 수집한 데이터를 자치구별로 나누었을 때, 변화된 건물의 숫자를 의미한다. 이렇게 해서 생성한 건물변화탐지를 위한 자치구별 학습데이터는 예를 들어, 강남구는 2,105개, 강서구는 4,290개, 노원구는 140개, 성북구는 1,313개, 은평구는 3,808개이다. 통상적으로 인공지능의 성능은 잘 정제된 데이터의 수에 비례한다고 알려져 있다[11]. 따라서 데이터의 수가 적은 자치구의 인공지능 성능은 그렇지 않은 자치구의 인공지능 성능보다 낮아진다. 데이터를 공유하여 한꺼번에 학습하지 못하는 상황에서, 각자가 소유한 데이터만을 이용하여 인공지능의 성능 향상을 위해, 연합학습 응용을 활용하는 것이고, 연합학습이 적용된 후의 정량적 성능 분석은 다음 장에서 묘사한다.

<표 1> 건물 변화탐지 데이터셋 구성

종류	과거 이미지	현재 이미지	정답 이미지	라벨링 데이터 (json 파일)
데이터셋 구성				

본 논문에서는 연합학습을 적용하지 않은 인공지능과 연합학습을 적용한 이후의 인공지능 성능 분석을 위해, W-Net 모델을 동일하게 활용한다. 건물 변화탐지 응용에서의 입력 이미지는 과거와 현재의 사진 두 개이고, 인공지능 모델의 네트워크에 들어가기 위해 두 장의 이미지를 하나로 변환해야 한다. 이 과정에서 이미지의 중요한 정보를 잃을 수 있고, 이러한 문제를 보완하기 위해 제안된 모델이 end-to-end dual branch 모델인 W-Net 모델이다[12, 13].

1) 상세한 데이터 개수와 F1 score 값은 부록의 <표 4>에서 정리되어 있다.

W-Net 모델의 구조는 U-Encoder와 U-Decoder의 두 부분으로 구성되어 있다. U-Encoder에서는 라벨링이 지정되지 않은 원본 영상에서 영상 분할을 출력하며, U-Decoder에서는 분할에서 재구성한 영상을 출력한다. W-Net 모델은 2개의 3×3 컨볼루션 레이어로 구성하며, 총 46개의 레이어가 존재한다[13]. 본 논문에서는 서로 다른 인공지능 모델을 비교하는 것이 아니다. 같은 인공지능 모델을 이용하여 연합학습을 적용하였을 때와 적용하지 않았을 때의 성능을 비교하는 것이다. 따라서 W-Net 모델을 수정하지 않고, 기본 모델을 적용하여 연합학습 성능을 분석한다. 데이터 학습을 위한 하이퍼 파라미터는 아래의 <표 2>와 같다. 배치 사이즈 (batch size)는 8개로 두어, 8개의 데이터를 한 번에 학습하도록 한다. 그리고 학습률 (learning rate)는 0.0001, 에포크 (epoch)는 50으로 설정하는데, 빠른 학습과 조기 종료를 위한 patience는 3으로 설정하여 학습을 진행한다.

<표 2> 학습 주요 파라미터

hyper parameter	value
이미지 scale	256×256
batch size	8
optimizer	Adam
learning rate	0.0001
epoch	50

V. 성능 분석

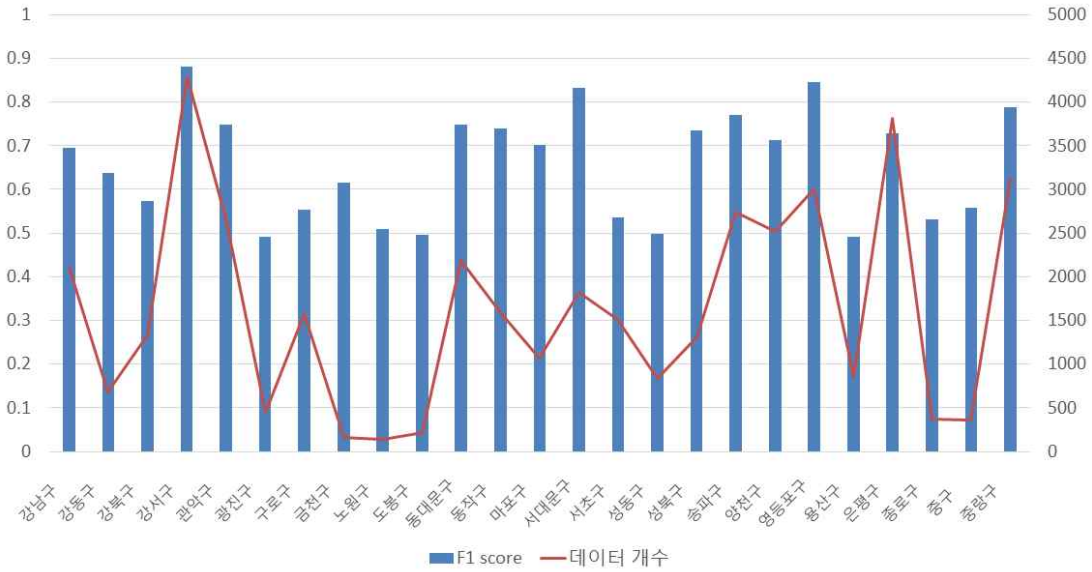
본 장에서는 4장 시스템 모델에서 정의한 데이터와 모델에 기반하여, 연합학습을 적용하지 않은 인공지능과 연합학습을 적용한 인공지능의 성능을 비교하여 분석한다. 이 때, 연합학습을 적용하지 않은 인공지능은 각 자치구에 속하는 훈련 데이터만을 이용하여 학습한다. 즉, 자치구별로 서로 다른 25개의 인공지능이 생성되고, 각 인공지능의 성능은 <그림 3>에서 묘사한 바 있다. 반면에 연합학습을 적용한 인공지능과의 성능을 비교, 분석하기 위해, 앞 장에서 정의했었던 동일한 인공지능 모델, W-Net 모델을 이용한다. 이에 따라, 연합학습을 이용한 서로 다른 25개의 인공지능이 생성되고, 각 자치구별로 생성된 연합학습을 적용하지 않은 인공지능과, 연합학습을 적용한 인공지능의 성능을 본 장에서 분석한다.

5.1 연합학습 이전의 자치구별 인공지능 성능 분석

연합학습을 적용하지 않은, 자치구별 인공지능은 4장 시스템 모델의 <표 2>에 따라 학습을 수행한다. 그리고 자치구별 학습 데이터는, 예를 들어 <그림 3>에서처럼, 강남구는 2,015개의 데이터를 이용하여 인공지능을 학습하고, 강동구는 684개의 학습 데이터를 이용하여 인공지능을 학습한다. 즉, 각 자치구에 속한 데이터만을 이용하여 인공지능을 훈련하고, 총 25개의 서로 다른 인공지능이 생성되는 것이다.

위와 같이 생성된, 각 자치구별 인공지능의 성능은 <그림 3>과 같다. <그림 3>에서 y축의 왼쪽 레이블은 F1 score를 의미하고, 오른쪽 레이블은 자치구별 학습 데이터 개수를 나타낸다. F1 score는 인공지능의 성능을 판단할 수 있는 지표들 중 하나이고, 정밀도 (precision)와 재현율 (recall)의 조화평균이다[14]. 정밀도는 인공지능이 정답이라고 판단한 데이터 가운데에 실제로 정답인 데이터의 비율이다. 그리고 재현율은 모든 정답들 중에 인공지능이 골라낸 정답의 비율이다. 따라서 정밀도를 높이기 위해서는 필터를 촘촘하게 적용하여 적은 수의 정답을 골라내면 된다. 반면에 재현율을 높이기 위해서는 필터를 느슨하게 적용하여 최대한 많은 수의 정답지를 고르면 된다. 즉, 정밀도와 재현율은 서로 트레이드오프 관계이고, 인공지능의 성능을 분석하기 위해, 정밀도와 재현율의 조화평균인 F1 score 지표가 활용된다.

다시, <그림 3>으로 돌아와서, 학습 데이터의 수가



<그림 3> 자치구별 훈련 데이터 개수에 따르는 인공지능 정확도

증가할수록 F1 score 값이 높아지고, 데이터의 수가 감소할수록 인공지능 성능도 감소하는 추세를 확인할 수 있다. 그러나 은평구의 경우에는, 학습데이터 수에 비해 F1 score가 0.73으로 인공지능의 성능이 상대적으로 낮을 것을 볼 수 있다. 이는 인공지능 학습이 조기 종료가 되어, 충분한 데이터가 있음에도 불구하고, 인공지능 훈련이 제대로 이뤄지지 않기 때문이다. 그러나 일반적으로, 데이터를 많이 소지하여 학습한 인공지능의 성능은 그렇지 않은 인공지능의 성능보다 높은 것을 확인할 수 있다.

5.2 연합학습 이후의 자치구별 인공지능 성능 분석

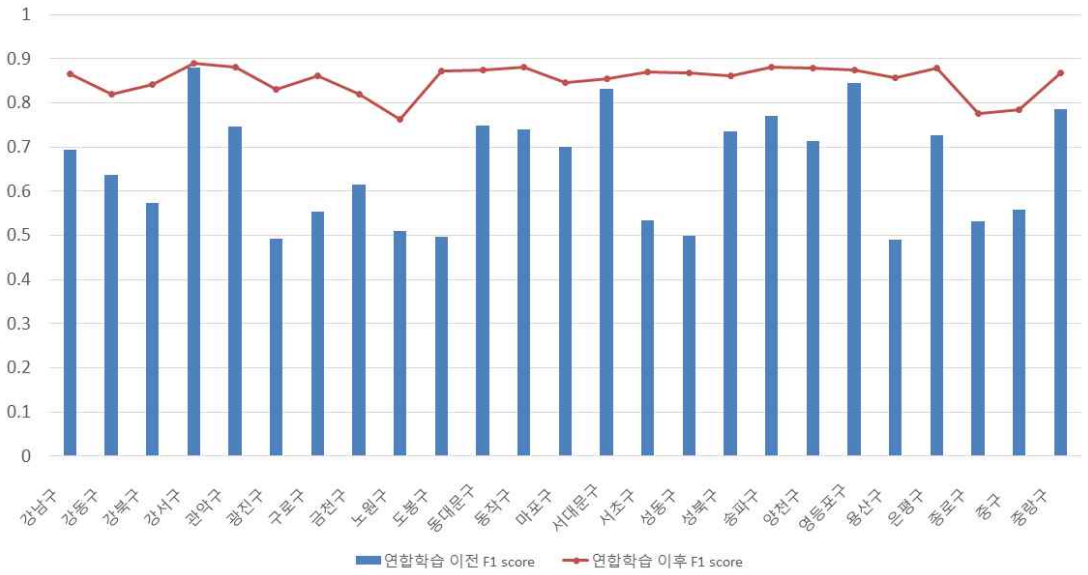
연합학습을 적용한 인공지능을 개발하기 위한 학습 파라미터는 5.1장에서와 동일하게 <표 2>에 따라 학습을 수행한다. 전체적인 연합학습 과정은 <그림 1>의 과정을 따르고, 성능 분석을 위한 하드웨어 환경은 아래의 <표 3>과 같다.

<표 3> 성능 분석 환경

항목	이름	개수
GPU	GeForce RTX 2080 Ti	3
CPU	Intel Core i9-7900X 3.30GHz × 20	1
RAM	Samsung DDR4 16GB	8
SDD	Samsung SSD 970 PRO 1TB	1
HDD	Seagate IronWolf 7200/256M (ST8000VN0022, 8TB)	1

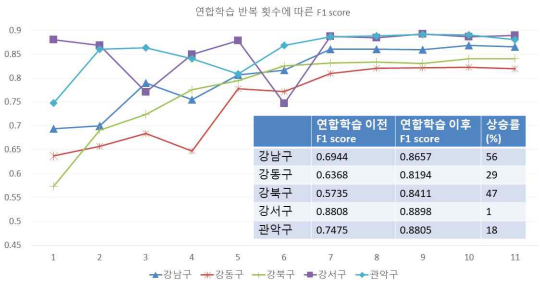
일반 인공지능과 연합학습 간의 성능 비교 그래프는 <그림 4>[2]와 같다. <그림 4>에서 파란색으로 색칠한 막대그래프는 연합학습 이전의 F1 score를 보여주며, 이 막대그래프의 값들은 <그림 3>에서의 값들과 동일하다. 그리고 빨간색의 선그래프는 연합학습을 적용한 이후의 인공지능 성능을 보여준다.

2) 연합학습 이후의 F1 score 값과 상승률 값은 부록의 <표 4>에서 정리되어 있다.



<그림 4> 연합학습 성능 분석

연합학습 이전의 각 자치구별 인공지능 성능은, <그림 3>에서 본 것처럼, 대체적으로 학습 데이터의 개수에 따라 비례한다. 이에 따라 데이터의 수가 적은 자치구에서 생성한 인공지능의 성능은 상대적으로 데이터 개수가 많은 자치구의 인공지능 성능보다 낮다. 반면에 연합학습 이후의 성능은, 학습 데이터의 수에 의존하지 않고, 대체적으로 일정한 수준의 성능을 가진 것을 볼 수 있다. 이는 연합학습 동작과정에서, 각 클라이언트들로부터 수신한 가중치 정보를 취합하고 계산하여, 같은 가중치 값을 클라이언트들이 공유하기 때문이다.



	연합학습 이전 F1 score	연합학습 이후 F1 score	상승률 (%)
강남구	0.6944	0.8657	56
강동구	0.6368	0.8194	29
강북구	0.5735	0.8411	47
강서구	0.8808	0.8898	1
관악구	0.7475	0.8805	18

<그림 5> 연합학습 round에 따르는 성능 분석

즉, 연합학습 이후의 각 클라이언트는 동일한 인공지능을 소유하게 되고, 서로 유사한 성능을 가진다. 이에 따라, <그림 4>에서 볼 수 있듯이, 인공지능 성능이 상대적으로 낮은 자치구의 이득이, 그렇지 않은 자치구의 이득보다 더 큰 것을 확인할 수 있다. 예를 들어, 도봉구의 경우에는 연합학습 이전의 F1 score 값이 0.5이고, 연합학습 이후에는 0.87로서 약 75%의 상승률을 보인다. 반면에 강서구나 영등포구의 경우에는 각각 0.88에서 0.89로 상승, 0.85에서 0.88로 상승하여, 상승폭이 상대적으로 작다. 즉, 연합학습을 활용하면 대부분의 클라이언트들은 이득을 보지만, 데이터의 수가 많은 클라이언트는 상대적으로 적은 이득을, 데이터의 수가 적은 클라이언트는 큰 이득을 보는 것이다. 이러한 특징으로 인해, 다양한 클라이언

트들을 연합학습에 참가할 수 있는 정책적인 유인책 등이 필요하다.

<그림 5>는 연합학습 라운드별 각 클라이언트의 인공지능 성능을 구체적으로 보여준다. <그림 5>에서는 5개의 클라이언트, 강남구, 강동구, 강북구, 강서구, 관악구에 대한 연합학습 라운드별 성능을 묘사하고 나머지 자치구들에 대한 그래프는 본 논문의 부록을 참고하면 된다.

<그림 5>에서, 첫 번째 라운드에서의 성능은 <그림 3>의 인공지능 성능과 유사하다. 그러나 라운드가 증가할수록 각 자치구별 인공지능 성능이 향상하다가 9번째 라운드 이후부터 각 자치구별 인공지능 성능이 일정한 값으로 수렴하는 것을 볼 수 있다. 또한 연합학습 이전과 이후의 성능 상승률 역시 <그림 5>의 오른쪽 아래 부분의 표에서 확인할 수 있다. 앞서 언급했던 것처럼, 이미 학습 데이터를 많이 소유한 클라이언트의 성능 향상은 미미한 반면에, 학습 데이터의 수가 적은 클라이언트의 성능 향상은 상대적으로 높은 것을 볼 수 있다. 따라서 연합학습 알고리즘이 널리 활용되기 위해서는, 상대적으로 데이터를 많이 소유한 클라이언트들이 참여할 수 있는 유인책을 마련하는 것이 중요하다.

VI. 결론

본 논문에서는 자치구별 건물 변화탐지 학습 데이터를 활용하여, 연합학습을 적용하였을 때의 성능 향상이 어느 정도인지 정량적으로 제시하였다. 그 결과, 모든 자치구의 인공지능 성능이 연합학습에 의하여 향상되는 것을 확인하였다. 특히, 자체 데이터가 부족하여 인공지능 성능이 낮았던 자치구의 경우, 연합학습 이후에 성능이 상대적으로 더 좋아진 결과도 확인할 수 있었다. 그러나 이미 학습 데이터를 충분히 소유한 클라이언트는 연합학습에 참가하지 않더라도 일정 수준 이상의 성능을 가진 인공지능을 확보할 수 있다. 다시 말해, 연합학습에 참가하여 얻을 수 있는 이득이 상대적으로 크지 않은 것이다. 따라서 학습 데이터를 충분히 소유한 클라이언트들이, 연합학습 프로세스에 참여할 수 있도록 유인하는 것이 연합학습을 수행하는 데에 중요한 이슈 중 하나이다. 또한 컴퓨팅 환경과 네트워크 환경이 제각각인 클라이언트들의 참여로 인해, 연합학습의 수렴 속도도 해결해야 할 과제이다. 이 부분은 향후 연구과제로서, 연합학습 라운드의 수를 효과적으로 줄일 수 있는 방안을 제시할 예정이다.

참고문헌

[1] Z. Jan, F. Ahamed, W. Mayer, N. Patel, G. Grossmann, M. Stumptner, and A. Kuusk, “Artificial Intelligence for Industry 4.0: Systematic Review of Applications, Challenges, and Opportunities,” ELSEVIER Expert Systems with Applications, Vol.216-119456, 2023, pp.1-21.

[2] S. AbdulRahman, H. Tout, H. O.-Slimane, A. Mourad, C. Talhi, and M Guizani, “A Survey on Federated Learning: The Journey From Centralized to Distributed On-Site Learning and Beyond,” IEEE Internet of Things Journal, Vol.8, No.7, 2021, pp.5476-5497.

[3] M. Aledhari, R. Razzak, R. M Parizi, and F. Saeed, “Federated Learning: A Survey on Enabling Technologies, Protocols, and Applications,” IEEE Access, Vol.8, 2020, pp.140699-140725.

[4] S. Pandya, G. Srivastava, R. Jhaveri, M. R. Babu, S. Bhattacharya, P. K. R. Maddikunta, S. Mastorakis, Md. J. Piran, and T. R. Gadekallu,

"Federated Learning for Smart Cities: A Comprehensive Survey," ELSEVIER Sustainable Energy Technologies and Assessments, Vol.55-102987, 2023, pp.1-13.

[5] M. F. Sohan and A. Basalamah, "A Systematic Review on Federated Learning in Medical Image Analysis," IEEE Access, Vol.11, 2023, pp.28628-28644.

[6] D. C. Nguyen, Q.-V. Pham, P. N. Pathirana, M. Ding, A. Seneviratne, Z. Lin, O. Dobre, and W.-J. Hwang, "Federated Learning for Smart Healthcare: A Survey," ACM Computing Surveys, Vol.55, No.3, 2022, pp.1-37.

[7] S. Pandya, G. Srivastava, R. Jhaveri, M. R. Babu, S. Bhattacharya, P. K. R. Maddikunta, S. Mastorakis, Md. J. Piran, and T. R. Gadekallu, "Federated Learning for Smart Cities: A Comprehensive Survey," ELSEVIER Sustainable Energy Technologies and Assessments, Vol.55-102987, 2023, pp.1-13.

[8] T. Zhang, L. Gao, C. He, M. Zhang, B. Krishnamachari, and A. S. Avestimehr, "Federated Learning for the Internet of Things: Applications, Challenges, and Opportunities," IEEE Internet of Things Magazine, Vol.5, No.1, 2022, pp.24-29.

[9] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A Survey on Security and Privacy of Federated Learning," ELSEVIER Future Generation Computer Systems, Vol.115, 2021, pp.619-640.

[10] 오재홍 · 이창노, "기구축 고해상도 위성 데이터를 이용한 정밀 시계열정사영상 자동 생성," 한국측량학회지, 제41권, 제1호, 2023, pp.29-37.

[11] M. Xua, S. Yoon. A. Fuentes, and D. S. Park, "A Comprehensive Survey of Image Augmentation Techniques for Deep Learning," ELSEVIER Pattern Recognition, Vol.137-109347, 2023, pp.1-12.

[12] B. Hou, Q. Liu, H. W, and Y. Wang, "From W-Net to CDGAN: Bi-temporal Change Detection via Deep Learning Techniques," IEEE Transactions on Geoscience and Remote Sensing, Vol.58, No.3, 2020, pp.1790-1802.

[13] X. Xia and B. Kulis, "W-Net: A Deep Model for Fully Unsupervised Image Segmentation," arXiv preprint arXiv:1711.08506, 2017.

[14] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," BMC Medical Imaging, Vol.15, No.29, 2015, pp.1–28.

■ 저자소개 ■

김 영 현
(Kim Younghyun)

2022년 3월~현재
서울디지털재단 AI · 빅데이터팀
수석
2013년 2월 고려대학교
전자전기컴퓨터공학과(공학박사)
2005년 2월 숭실대학교
컴퓨터공학과(공학사)

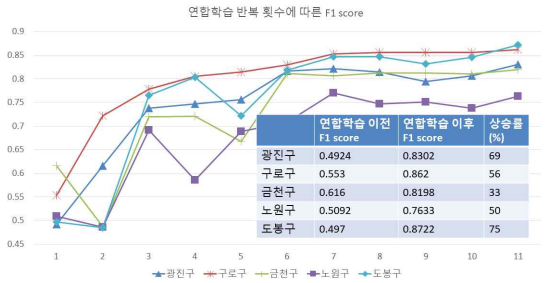
관심분야 : 정보통신, 인공지능
E-mail : kim@sdf.seoul.kr

논문접수일 : 2023년 8월 10일
수정접수일 : 2023년 8월 25일
게재확정일 : 2023년 9월 1일

부록

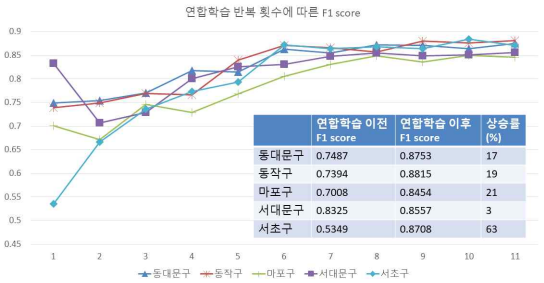
<표 4> <그림 3>과 <그림 4>의 데이터 수치

자치구	데이터 개수	연합학습 이전의 F1 score	연합학습 이후의 F1 score	상승률 (%)
강남구	2,105	0.6944	0.8657	24.66
강동구	684	0.6368	0.8194	28.67
강북구	1,331	0.5735	0.8411	46.66
강서구	4,290	0.8808	0.8898	1.021
관악구	2,660	0.7475	0.8805	17.79
광진구	450	0.4924	0.8302	68.60
구로구	1,577	0.553	0.862	55.87
금천구	160	0.616	0.8198	33.08
노원구	140	0.5092	0.7633	49.90
도봉구	220	0.497	0.8722	75.49
동대문구	2,191	0.7487	0.8753	16.90
동작구	1,592	0.7394	0.8815	19.21
마포구	1,071	0.7008	0.8454	20.63
서대문구	1,820	0.8325	0.8557	2.786
서초구	1,513	0.5349	0.8708	62.79
성동구	838	0.4986	0.8679	74.06
성북구	1,313	0.7355	0.8623	17.23
송파구	2,742	0.7704	0.8823	14.52
양천구	2,519	0.7136	0.8799	23.30
영등포구	3,009	0.846	0.875	3.427
용산구	853	0.4913	0.8567	74.37
은평구	3,808	0.7279	0.8797	20.85
종로구	372	0.5312	0.7759	46.06
중구	360	0.5584	0.7859	40.74
중랑구	3,132	0.787	0.8673	10.20



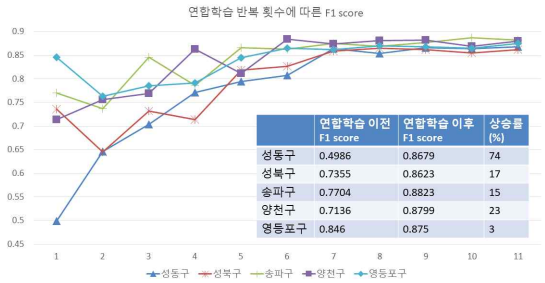
	연합학습 이전 F1 score	연합학습 이후 F1 score	상승률 (%)
광진구	0.4924	0.8302	69
구로구	0.553	0.862	56
금천구	0.616	0.8198	33
노원구	0.5092	0.7633	50
도봉구	0.497	0.8722	75

<그림 6> 연합학습 round에 따르는 성능 분석
(광진구, 구로구, 금천구, 노원구, 도봉구)



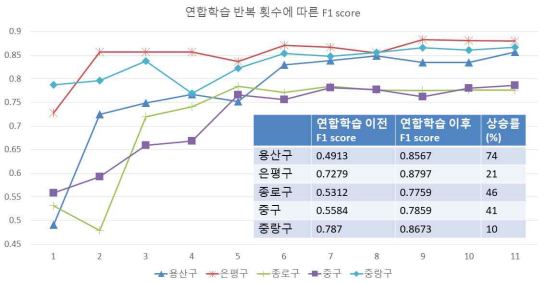
	연합학습 이전 F1 score	연합학습 이후 F1 score	상승률 (%)
동대문구	0.7487	0.8753	17
동작구	0.7394	0.8815	19
마포구	0.7008	0.8454	21
서대문구	0.8325	0.8557	3
서초구	0.5349	0.8708	63

<그림 7> 연합학습 round에 따르는 성능 분석
(동대문구, 동작구, 마포구, 서대문구, 서초구)



	연합학습 이전 F1 score	연합학습 이후 F1 score	상승률 (%)
성동구	0.4986	0.8679	74
성북구	0.7355	0.8623	17
송파구	0.7704	0.8823	15
양천구	0.7136	0.8799	23
영등포구	0.846	0.875	3

<그림 8> 연합학습 round에 따르는 성능 분석
(성동구, 성북구, 송파구, 양천구, 영등포구)



	연합학습 이전 F1 score	연합학습 이후 F1 score	상승률 (%)
용산구	0.4913	0.8567	74
은평구	0.7279	0.8797	21
종로구	0.5312	0.7759	46
중구	0.5584	0.7859	41
중랑구	0.787	0.8673	10

<그림 9> 연합학습 round에 따르는 성능 분석
(용산구, 은평구, 종로구, 중구, 중랑구)