ISSN 2288-4866 (Print)
ISSN 2288-4882 (Online)
http://www.jiisonline.org
J Intell Inform Syst 2022 December: 28(4): 27~40
http://dx.doi.org/10.13088/jiis.2022.28.4.027

연합학습의 의료분야 적용을 위한 자기지도 메타러닝*

공희산
성균관대학교 소프트웨어학과
(khtks@g.skku.edu)

김광수
성균관대학교 소프트웨어학과
(kim.kwangsu@skku.edu)

최근 많은 발전을 이룬 의료 인공지능은 의사가 진단과 결정을 내리는 데 도움을 주는 등 중요한 역할을 수행하고 있다. 특히, 흉부 엑스레이 분야는 접근성 및 흉부질환 탐지에 유용함과 최근 COVID-19 상황이 도래함에 따라 많은 관심을 받고 있다. 그러나, 데이터의 수가 많음에도 레이블이 있는 데이터의 수가 부족하므로 효과적인 인공지능 모델을 만드는 데 한계가 있다. 이러한 문제를 완화하는 방안으로 연합학습을 흉부 엑스레이 데이터에 적용한 연구가 등장했지만, 여전히 다음과 같은 문제를 내포하고 있다. 1) Non-IID 환경에서 발생할 수 있는 문제를 고려하지 않았다. 2) 연합학습 환경에서도 여전히 클라이언트의 레이블이 있는 데이터가 부족하다. 우리는 자기지도학습 모델을 연합학습의 Global 모델로 사용함으로써 위와 같은 문제를 해결하는 방법을 제안한다. 이를 위해 흉부 엑스레이 데이터를 사용한 연합학습에 알맞은 자기지도학습 방법론을 실험적으로 탐색하며, 자기지도학습 모델을 연합학습에 사용함으로써 얻을 수 있는 장점을 검증한다.

주제어 : 자기지도학습, 연합학습, 흉부 엑스레이, 의료 인공지능

논문접수일 : 2022년 10월 6일　논문수정일 : 2022년 10월 18일　게재확정일 : 2022년 10월 24일
원고유형 : Fast Track　교신저자 : 김광수

1. 개요

최근 다양한 분야의 인공지능 기술의 발전에 발맞춰 의료 분야에도 인공지능을 도입하기 위한 연구들이 진행되었고, 현재 딥러닝을 사용한 의료 인공지능은 의사의 진단과 결정에 도움을 주는 등 여러 분야에서 중요한 역할을 수행하고 있다. 그중, 의료 이미지를 사용한 의료 인공지능은 최근 몇 년간 급격한 발전을 이뤄냈으며, 특히 흉부 엑스레이 (Chest radiography or Chest X-ray) 분야는 오랜 기간 쌓여있는 많은 데이터와 폐렴(Pneumonia), 탈장(Hernia) 및 삼출(Effusion) 등과 같이 심각한 흉부 질환을 사전에 탐지할 수 있는 유용함 때문에 많은 연구가 진행되었고, 최근 COVID-19 상황이 도래하면서 관심이 더욱 증폭되었다(Mangal et al., 2020; Rajpurkar et al., 2017).

그러나, 일반적으로 알려져 있듯 효과적인 인공지능 모델을 만들기 위해서는 엄청난 양의 다양한 데이터가 필요하다는 한계점이 있다. 그렇기에, 데이터가 많음에도 불구하고 레이블이 있는 데이터의 수가 부족한 흉부 엑스레이 분야는 효과적인 의료 진단 인공지능 모델을 만들기 어렵다는 문제가 있다. 또한, 의료 데이터는 일반적인 데이터에 비해 환자의 개인정보 문제가 특

* 이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신산업진흥원의 지원을 받아 수행된 헬스케어 AI 융합 연구개발 사업임(No.S0254-22-1006)

히 심각하므로, 딥러닝의 학습을 위해 서버에 모으는데 제약이 있다.

이러한 문제를 해결하는 방안으로, 연합학습(Federated Learning)이 효과적인 방법이 될 수 있다. 연합학습(McMahan et al., 2017)이란, 학습을 위한 데이터를 서버에 모으는 대신 클라이언트들이 각자의 데이터를 사용해 학습한 후 하나의 효과적인 모델을 만드는 방법이다. 연합학습에서 클라이언트란, 학습에 참여하며 개별적으로 보유한 Local 데이터를 사용해 서버에서 배포된 Global 모델을 학습시키는 주체를 의미한다. 이 때, 서버는 각 클라이언트의 데이터에 직접적으로 접근할 수 없으며, 학습된 Local 모델의 가중치만을 사용해 Global 모델을 업데이트하게 된다. 연합학습의 등장 이후 여러 연구에서 흉부 엑스레이 분야에 적용하는 방법(Banerjee et al., 2020; Yan et al, 2021)을 제시했지만, 여전히 다음과 같은 한계점들이 존재한다.

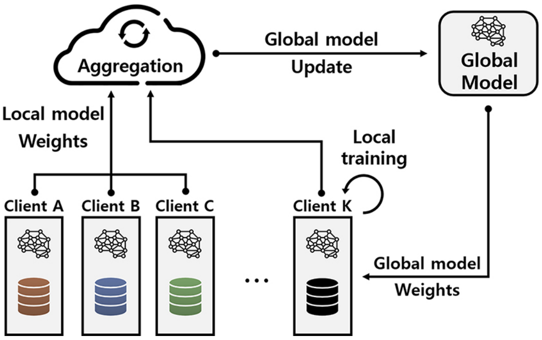
첫 번째, 연합학습의 실질적인 시나리오는 학습에 참여하는 클라이언트별로 보유한 클래스의 분포나 데이터의 크기가 다른 Non-IID 환경이 일반적이지만, 제안된 연구들에서는 대부분 위 조건들이 같은 IID 상황만을 고려했다. IID 환경에서의 연합학습은 각 Local 모델의 학습 결과가 비슷한 경향을 보이므로 학습이 원만하게 이루어진다. 이와 반대로, Non-IID 환경에서는 학습 결과의 편차로 인해 올바르게 학습이 이루어지지 않는 Client Drift와 같은 문제가 발생하는 경우가 대부분이다(Karimireddy et al., 2020). 그렇기 때문에, Non-IID 환경은 연합학습을 연구하는 많은 사람이 풀고자 하는 문제이며, 반드시 고려해야 하는 사항이다.

두 번째로, 클라이언트의 학습 데이터의 부족함이다. 데이터를 중앙 서버에 모으기 어려운 것에 더해, 학습에 사용하기 위해 레이블을 생성하는 데에는 많은 양의 인적 자원과 시간이 필요하다. 이런 상황은 의료분야에서 더욱 심해지는데, 레이블링을 위해 각 의료분야의 전문가가 반드시 필요하므로 데이터를 준비하기 더욱 어려워져 연합학습에 참여하는 클라이언트가 보유한 Local 데이터의 수가 부족할 수 있다.

위와 같은 문제들을 완화하는 방안으로 우리는 레이블이 없는 데이터를 사용해 인공지능 모델을 학습하는 방법론인 자기지도학습(Self-supervised Learning)을 사용한 사전학습(Pretraining)과 해당 모델을 사용한 연합학습 환경으로의 전이 학습(Transfer Learning)을 해결책으로 제시한다. 이를 위해 자기지도학습 모델의 연합학습에서의 유용성을 실험적으로 판단함으로써 알맞은 자기지도학습 방법론을 선정했다. 또한, 자기지도학습 모델을 연합학습의 Global 모델로 사용함으로써 얻을 수 있는 장점들을 검증했다.

위 방법을 실험 및 검증하기 위해 우리는 자기지도학습에는 공개된 흉부 엑스레이 데이터를, 연합학습에는 각 병원이 가지고 있는 데이터를 사용하는 환경을 가정했다. 이 연구를 통한 우리의 공헌은 다음과 같다.

1. 흉부 엑스레이 데이터를 사용한 연합학습에 알맞은 자기지도학습 방법론 제시.
2. 자기지도학습 모델을 사용함으로써 연합학습 환경의 실질적인 시나리오인 Non-IID 상황에서 흉부질환 진단 성능의 향상.
3. 각 클라이언트의 레이블이 있는 데이터의 수가 부족한 상황에서의 진단 성능 유지.
4. 연합학습에서의 빠른 수렴으로 인한 커뮤니케이션 라운드 감소



〈그림 1〉 연합학습의 학습과정

2. 관련 연구

2.1 연합 학습(Federated Learning)

2017년에 Google에서 처음 제안한 연합학습은, 데이터를 중앙 서버에 적재하지 않고 학습에 참여하는 각 클라이언트 개별적인 데이터를 사용해 모델을 학습 후 하나의 총체적인 모델을 만듦으로써 데이터 프라이버시 향상과 풍부한 데이터로 훈련된 공유 모델의 이점을 집합적으로 얻을 수 있도록 하는 방법론이다.

2.1.1 학습 과정

<그림 1>은 연합학습의 전반적인 학습 과정을 나타내고 있다. 연합학습은 Global 모델을 관리하는 서버와 각 Local 데이터를 기반으로 학습하는 클라이언트를 기반으로 진행되므로, 배포할 Global 모델을 초기화하며 시작하게 된다. 이때, 일반적인 연합학습의 Global 모델 초기화 방법은 임의의 가중치로 초기화하는 방법이다. 이후, 각 클라이언트는 배포 받은 모델을 Local 데이터를 활용해 정해진 횟수만큼 학습함으로써 Local 모델을 생성하게 되고, Local 모델의 가중치만을 서버로 전송하게 된다. 각 Local 모델의 가중치를 전송받은 서버는 병합 알고리즘을 통해 새로운 가중치를 만들게 되고, 이 가중치를 사용해 Global 모델을 업데이트하게 된다. 이처럼 Global 모델은 Local 모델의 가중치의 병합을 통해서만 업데이트하고, 이러한 과정을 연합학습이 종료될 때까지 반복하게 된다.

연합학습에서 서버에서 Local 모델의 가중치 병합을 위한 알고리즘은 Local 클라이언트가 보유한 데이터 수의 비율을 기반으로 한 가중평균을 구하는 FedAvg 알고리즘을 일반적으로 사용한다. 기본적으로 Global 모델의 업데이트는 이전 가중치를 병합한 가중치로 대체하는 방식으로 진행된다. Global 모델의 배포 및 Local 모델의 가중치를 전송하는 것을 커뮤니케이션이라고 하며, Global 모델을 및 Local 모델을 업데이트(학습)하는 횟수를 Global 라운드와 Local 라운드라고 명명한다.

2.1.2 IID & Non-IID 환경

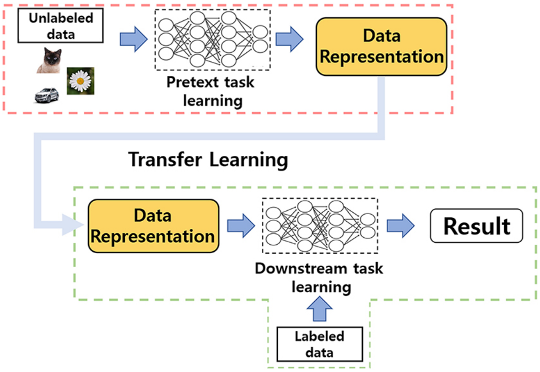
각 클라이언트가 보유한 Local 데이터만을 사용해 학습을 진행하는 연합학습의 환경은 Local 데이터의 분포에 따라 크게 IID와 Non-IID로 구분할 수 있다.

IID (Independent Identically Distribution) 환경이란, 간단히 말해 각 클라이언트가 데이터의 개수, 클래스의 종류 등 모든 항목에서 동일한 Local 데이터를 보유하고 있는 환경을 의미한다. IID 환경에서는 Local 데이터를 사용한 Local 모델의 학습이 모든 클라이언트에 거쳐 유사한 경향을 띄게 되므로, Local 모델의 가중치 병합을 통해서도 원만하게 Global 모델의 학습이 이루어진다.

이와 반대로, Non-IID(Non-Independent Identically Distribution) 환경은 각 클라이언트의 Local 데이터가 비독립적이고 동일하지 않게 분포되어있는 환경을 의미한다. 즉, 클라이언트별로 보유하고 있는 Local 데이터의 종류, 개수 등의 특징이 다른 환경을 의미하며, 실제 연합학습 시나리오의 대부분의 데이터 분포가 해당한다. Non-IID 환경은 분포되어 있는 데이터의 특성에 따라 공변량 이동(Covariate Shift), 사전 확률 이동(Prior Probability Shift), 개념 전환(Concept Shift), 개념 이동(Concept Drift)로 분류한다(Kairouz et al., 2021). Non-IID 환경에서는 클라이언트가 보유한 데이터의 차이로 인해 Local 모델의 학습 결과도 편차가 발생해 Global 모델을 올바르게 업데이트하지 못하는 Client Drift 문제가 발생하게 된다. 그러므로 Local 모델 간의 편차를 보정하고 학습을 원활하게 하려는 다양한 방법들이 제시되고 있다(Karimireddy et al., 2020; Li et al., 2020; Wang et al., 2020).

2.2 자기지도학습(Self-supervised Learning)

비지도학습(Unsupervised Learning)의 일종인 자기지도학습은, 효과적인 모델을 만들기 위해 엄청난 양의 다양한 레이블 데이터가 필요한 지도학습(Supervised Learning)의 한계점을 극복하고 언레이블 데이터를 좀 더 적극적으로 사용하기 위한 방안으로 등장했다. 자기지도학습은 레이블이 없는 데이터를 사용해 분류(Classification)나 객체 인식(Object Detection) 등 여러 Task에 손쉽게 전이할 수 있는 데이터의 유용한 표현(Representation)을 학습하는데 목표를 두고 있다. 일반적으로 데이터 표현이란 좁은 범위로는 데이터가 모델을 통과한 후 얻을 수 있는 Feature Map을 의미하며, 상대적으로 넓은 범위로는 데이터로부터 학습한 모델 전체를 의미한다고 할 수 있다. 자기지도학습의 데이터 표현은 주로 넓은 의미로 사용된다. 학습된 표현을 목표 Task로 전이해 레이블 데이터로 추가학습 함으로써, 적은 수의 데이터로도 합리적인 성능과 빠른 수렴 속도를 보인다는 장점이 있다.



〈그림 2〉 자기지도학습의 학습 과정

2.2.1 학습 과정

자기지도학습의 학습 과정은 레이블이 없는 데이터를 사용해 유용한 표현을 학습하기 위한 Pretext Task와 학습한 표현과 레이블이 있는 데이터를 사용해 실제 해결하고자 하는 문제를 학습하는 Downstream Task 두 단계로 나눌 수 있다. 자기지도학습의 전반적인 학습 과정을 나타낸 <그림 2>에서 볼 수 있듯 자기지도학습은 언레이블 데이터를 입력으로 받아 적용할 Pretext Task에 맞게 데이터를 변형하고, 변형한 데이터에 기반해 가상 레이블을 생성하게 된다. 이후 생성한 가상 레이블과 변형된 데이터를 이용해 Pretext Task를 학습함으로써 모델은 레이블이 없는 데이터로부터 표현을 학습하게 된다. 학습한 표현은 용도에 맞게 다양한 Downstream Task

로 전이되어 레이블이 있는 데이터로 추가학습에 사용된다.

언레이블 데이터를 사용해 표현을 학습하기 위한 Pretext Task의 종류는 크게 Heuristic 하게 디자인된 변형 기반의 Pretext task와 하나의 데이터에 서로 다른 증강(Augmentation)기법을 적용한 결과의 표현을 최대한 유사하도록 학습하는 Instance Discrimination 접근법이 있다.

2.2.2 Pretext Task

컴퓨터 비전 분야 자기지도학습의 초기 접근법은 Relative Location(Doersch et al., 2015), Jigsaw Puzzle(Noroozi et al., 2016), Rotation(Komodakis, N., & Gidaris, S. 2018)과 같이 주로 데이터에 강체 변환(Geometric Transform)을 적용하고, 해당 변환에 알맞은 레이블을 생성하도록 Pretext Task를 디자인하는 데 초점이 맞춰져 있었다. 하지만, 이처럼 Heuristic 하게 디자인된 Pretext Task는 적용한 데이터의 변환에 의존적이기 때문에, 다양한 Downstream Task로의 전이와 일반화 부분에서 한계를 보였다.

그렇기 때문에, 최근 많은 연구에서 Instance Discrimination 접근법을 사용한 자기지도학습 방법이 제안되고 있으며, 지도학습과 비교할만한 뛰어난 성능을 보인다. Instance Discrimination 접근법은 하나의 데이터에 서로 다른 데이터 증강(Data Augmentation) 기법을 적용해 생성된 서로 다른 두 뷰(View)가 내재한 표현을 최대화하는 방향으로 학습하는 방법이다.

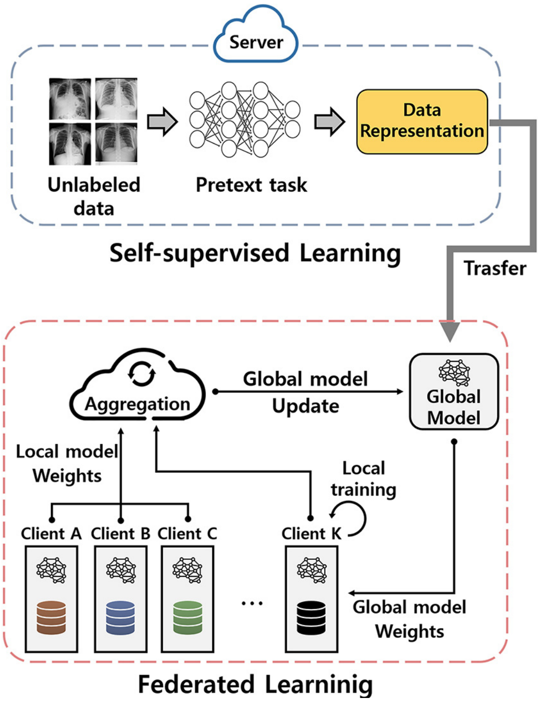
Instance Discrimination의 주된 연구 방향은 표현의 비교를 위해 SimCLR(Chen et al., 2020)과 같이 Contrastive Learning을 사용하는 방법이다. Contrastive Learning이란 대상들의 차이를 좀 더 명확하게 보여줄 수 있도록 학습하는 방법으로써 Negative & Positive Sample을 지정해 Positive Sample과는 더욱 가깝게, Negative Sample과는 더욱 멀도록 학습하는 방법론이다. 이와 같은 Contrastive Learning을 자기지도학습에 적용하는 많은 연구는 일반적으로 하나의 데이터에서 생성된 두 뷰를 Positive Sample로 정하고 다른 데이터에서 생성한 뷰를 Negative Sample로 정해 비교하며 학습하게 된다. 그러므로, 유용한 표현의 학습을 위해선 풍부한 Negative Sample과의 비교가 필수적이므로 큰 배치 크기에 의존하거나 메모리뱅크와 같은 특별한 구조를 사용하는 방법이 제안되었다(Misra, I., & Maaten, L. V. D. 2020). Contrastive loss를 사용해 직접적으로 비교하는 방법 외에도, 정보 병목(Information Bottleneck)이론을 사용해 중복된 정보를 제거하거나, 클러스터링(Clustering)에 뷰의 표현을 할당해 예측하는 방법 등이 있다(Zbontar et al., 2021; Caron et al, 2020).

2.2.3 Downstream Task

자기지도학습의 Downstream Task란, 학습된 데이터의 표현을 전이해 실제로 풀고자 하는 문제를 의미한다. Downstream Task의 학습을 위해 사전학습 된 모델의 Task에 맞는 분류기(Classifier)를 추가하며, 레이블이 있는 데이터를 사용해 추가 학습하게 된다. Downstream Task는 학습된 표현의 사용방식에 따라 크게 Linear Evaluation과 Fine Tuning으로 구분된다.

Linear Evaluation이란 자기지도학습으로 사전학습된 모델을 Feature Extractor로써 사용하는 방법으로 분류기만을 추가로 학습하는 방법이며, 주로 학습된 표현의 성능을 평가하는 데 사

용된다. 학습된 모델을 가중치 초기화(Weight Initializer)로 사용하는 방법인 Fine Tuning은 사전 학습된 부분과 분류기를 전부 추가로 학습하는 방법이다.



〈그림 3〉 연합학습을 위한 자기지도학습 프레임워크

3. 제안 방법

대부분의 연합학습은 학습을 시작할 때의 Global 모델을 가중치를 랜덤하게 초기화하는 경우가 대부분이기 때문에 학습의 목표인 최적의 포인트와는 많은 차이가 있다. 그렇기에 학습에 참여하는 클라이언트의 Local 모델을 병합하고 새로이 배포해 학습하는 과정을 거치며 Local 데이터의 특징에 많은 영향을 받는 클라이언트 드리프트 문제가 발생하게 된다. 우리는 이러한 문제점을 완화하기 위해 <그림 3>과 같이 연합학습의 초기 Global 모델로 자기지도학습 방법론으로 사전 학습을 진행한 모델을 사용한 방법을 제안한다.

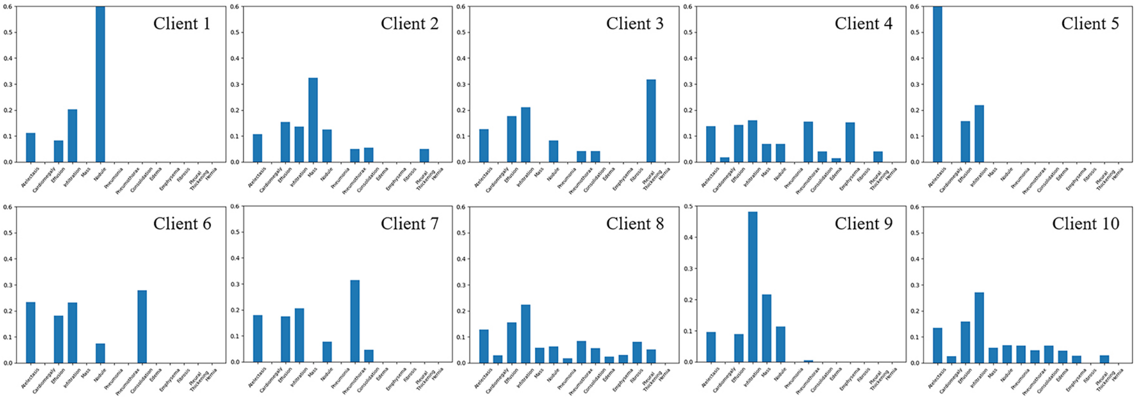
일반적으로 Pretext Task와 Downstream Task로 구성된 자기지도학습과 유사하게 우리의 방법론 역시 서버에서 흉부 엑스레이 데이터를 사용해 자기지도학습을 진행하고, 학습한 데이터의 표현을 Global 모델로 전이해 연합학습하는 2가지 단계를 거치게 된다.

데이터의 표현을 학습하기 위한 자기지도학습과 학습한 표현을 전이시켜 사용하는 연합학습에는 의료 인공지능 분야에서 자주 사용되며, 흉부 엑스레이에서 뛰어난 성능을 보이는 DenseNet-121을 사용했다(Rajpurkar et al., 2017; Baltruschat et al., 2019; Guan et al., 2021). DenseNet(Huang et al., 2017)은 ResNet(He et al., 2016)의 Skip-Connection과 유사한 연결을 전체 Layer에 거쳐 도입했다는 특징이 있다. 이처럼 Layer 사이에 연결을 밀도있게 유지함으로써, Feature Map의 재사용률을 향상시켜 학습이 좀 더 원활하게 이루어지도록 하고 기울기 소실(Gradient Vanishing) 문제를 완화할 수 있다. 서버에서 자기지도학습이 끝난 후 연합학습으로 전이를 위해 CNN 부분만을 남긴 후 나머지 부분은 제거하고, 연합학습에 알맞은 선형 분류기를 추가했다.

4. 실험 및 결과

4.1 실험환경 Setting

우리는 자기지도학습과 연합학습으로 전이해 추가 학습을 위한 2가지 데이터셋을 사용해 실

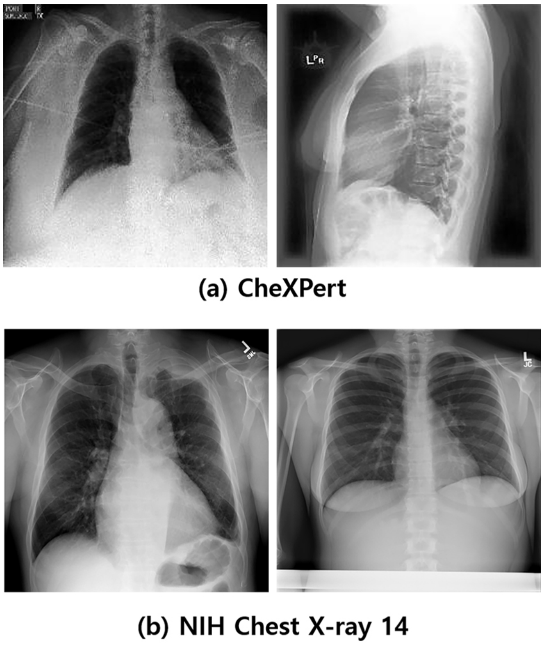


〈그림 4〉 연합학습을 위한 Non-IID 멀티 레이블 데이터 분포.

험을 진행했다. 우선, 표현을 학습하기 위한 자기지도학습 Pretext Task에는 불확실성 레이블이 있으며 약 22만 장 이상의 정면과 측면 흉부 데이터를 포함한 CheXPert 데이터셋(Irvin et al., 2019)을 사용했다. 연합학습을 적용한 Downstream Task는 총 14개의 흉부 질병에 대해 멀티 레이블링 되어있으며 약 11만 장의 흉부 사진을 포함하고 있는 Chest X-ray14 데이터(Xiaosong et al., 2017)를 사용하였다. 학습에 사용한 두 데이터셋의 예시 이미지는 <그림 5>와 같다.

표현을 학습하기 위한 자기지도학습의 Pretext Task에는 초기 Heuristic하게 디자인한 접근법과 Instance Discrimination 접근법을 동시에 고려했다. Heuristic 한 접근법에는 관련 연구(공희산, 박재훈, 김광수, 2021)를 따라 Rotation을 사용했으며, Instance Discrimination 접근법으로는 널리 쓰이고 있는 SimCLR, MoCo-v2(Chen et al., 2020), BYOL(Grill et al., 2020)을 사용했다.

Downstream Task인 연합학습에는 10개의 클라이언트가 참여하는 상황을 가정했다. 자기지도학습을 사용해 학습된 표현을 연합학습 환경에서 평가하기 위한 Linear Evaluation은 다른 외적인 영향을 배제하기 위해 각 클라이언트가 동일한 데이터의 분포와 규모를 갖는 IID 환경에서 진행했으며, 다른 실험들은 데이터 불균형으로 인해 Client Drift 문제가 발생할 수 있는 Non-IID 환경에서 진행했다. 하나의 데이터가 여러 클래



〈그림 5〉 흉부 엑스레이 데이터 예시

스에 속해있는 멀티 레이블 데이터를 사용해 Non-IID 환경을 만들기 위해 각 클라이언트에게 임의의 비율을 배정한 후, 전체 개수가 적은 클래스부터 하나의 클라이언트의 배정된 크기가 다 찰 때 까지 분배했다. 데이터 분포는 <그림 4>에 나와 있는 것과 같으며, 가로축은 14개의 질병 클래스, 세로축은 Local 데이터에서 각 클래스 데이터의 비율을 의미한다.

자기지도학습을 위해선 256의 배치 크기를 사용했으며, 0.05의 초기 학습률(Learning Rate)과 0.9의 Momentum을 갖는 SGD를 사용해 100 에폭(Epoch)동안 학습을 진행했다. 연합학습에는 64의 배치 크기를 사용했으며, 0.001의 초기 학습률과 0.9의 Momentum을 갖는 SGD를 사용해 최대 50 Global 라운드동안 학습을 진행했다. 또한, 가장 단순한 연합학습 상황을 고려해 Global 라운드와 Local 라운드는 1로 고정했다.

모든 흉부 엑스레이 데이터의 크기는 256×256으로 변환 후 224×224 크기로 임의의 부분을 잘라서 사용했다. 실험의 평가를 위한 성능 지표는 총 14개의 클래스에 대한 AUC(Area under the ROC curve)를 사용했다.

4.2 연합학습에서의 Linear Evaluation

우리는 흉부 엑스레이 데이터를 사용한 연합학습에서 자기지도학습을 통해 학습된 데이터 표현의 유용성을 판단하고 적합한 접근법을 확인하기 위한 Linear Evaluation 실험을 진행했다. 실험은 IID 환경에서 Global 모델의 초기화를 다르게 하며 자기지도학습에서와 같이 새롭게 추가된 분류기만을 학습하는 Linear Evaluation의 성능을 통해 비교했다. 그 결과 <표 1>에서처럼 Global 모델을 자기지도학습 모델로 초기화하는 경우가 대부분 높은 Linear Evaluation 결과를 보인다는 것을 확인할 수 있다. 또한, 자기지도학습 모델을 사용한 경우에서도 Heuristic한 방법으로 학습한 경우보다 Instance Discriminative 방법으로 학습한 모델의 성능이 더욱 뛰어나므로, 흉부 엑스레이 데이터를 사용한 전이학습에 더욱 유용하다는 것을 확인할 수 있다. 또한, 사용한 자기지도학습 방법론 중 흉부 엑스레이 데이터를 사용한 연합학습에서 가장 적합한 것은 MoCo-v2인 것을 확인할 수 있다.

〈표 1〉 IID 환경에서 Linear Evaluation 결과

Global Model Initialize	Mean AUC
Random Init	63.58
ImageNet (Superv.)	73.75
Rotation	72.65
SimCLR	75.20
MoCo-v2	**76.30**
BYOL	75.78

4.3 Non-IID 환경에서의 Fine Tuning

자기지도학습 모델을 사용해 Global 모델을 초기화하는 방법론의 Non-IID 환경에서의 장점을

〈표 2〉 Non-IID 환경에서 Fine-Tuning 결과

Method	Atel	Card	Effu	Infi	Mass	Nodu	Pne1	Pne2	Cons	Edem	Emph	Fibr	P.T.	Hern	Mean
Random	71.70	77.59	80.04	64.02	67.12	58.89	63.45	72.58	72.98	82.98	74.91	62.15	71.21	73.81	73.10
ImageNet	71.98	83.78	81.40	61.86	74.16	69.73	59.18	77.67	71.41	79.40	82.36	68.52	74.83	84.76	74.36
Rotation	**75.39**	81.73	81.26	**68.29**	70.58	66.50	63.49	75.71	74.67	81.59	77.93	68.32	72.14	75.62	73.79
SimCLR	72.89	88.34	83.07	65.87	74.26	**71.60**	64.39	77.71	73.17	**85.40**	**86.11**	70.67	78.79	**91.02**	77.38
MoCo-v2	74.62	87.88	**83.43**	63.52	**79.49**	70.61	**67.31**	**81.59**	75.32	83.18	84.45	72.93	**79.23**	88.60	**78.01**
BYOL	75.11	**88.58**	82.55	66.98	74.35	66.68	64.51	78.12	**75.84**	85.17	81.21	**74.17**	77.57	89.40	77.17

〈표 3〉 데이터 일부만 사용한 경우의 연합학습 성능

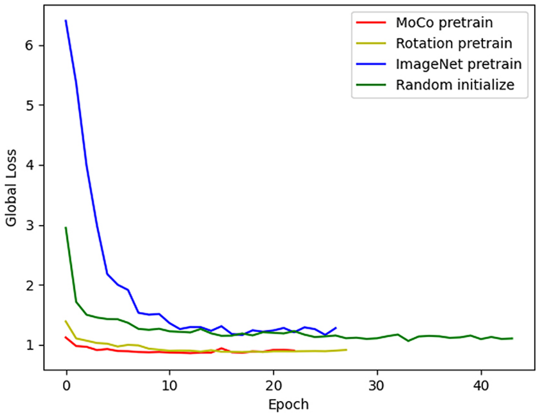
Method	Fraction of Data					
	100%	50%	25%	10%	5%	1%
Random	72.56	70.90 (-2.2)	67.96 (-5.14)	57.30 (-15.8)	52.90 (-20.2)	51.41 (-21.6)
MoCo-v2	78.01	76.51 (-1.5)	74.09 (-3.92)	69.86 (-8.15)	62.46 (-15.55)	58.03 (-19.8)

확인하기 위해 연합학습으로의 Fine Tuning 실험을 진행했다. 사용한 자기지도학습 방법은 이전 실험과 같으며, 비교를 위해 랜덤한 값으로 초기화하는 경우 외에도 ImageNet 데이터로 지도학습을 진행한 모델과도 비교했다. <표 2>에서 볼 수 있듯 Linear Evaluation과 마찬가지로 랜덤한 값으로 초기화환 경우보다 월등히 뛰어난 결과를 확인할 수 있으며, MoCo-v2 방법론을 사용한 자기지도학습 모델이 가장 적합하다는 것을 확인할 수 있다. 또한, 이미지 기반의 딥러닝에서 많이 사용되는 ImageNet 데이터를 사용한 사전학습 모델보다도 더욱 뛰어나다는 것을 확인할 수 있다.

4.4 Fraction of Label

자기지도학습 모델을 사용해 Fine Tuning 하는 경우 얻을 수 있는 장점 중 하나는, 적은 수의 레이블이 있는 데이터로도 합리적인 성능을 얻을 수 있다는 것이다. 상대적으로 레이블이 있는 데이터를 얻기 어려운 연합학습 환경에서 자기지도학습의 이점을 확인하기 위해, 우리는 데이터 일부만을 사용해 연합학습으로 Fine Tuning 하는 실험을 진행했다.

이 실험을 위해 우리는 랜덤 초기화, ImageNet 초기화와 Rotation 및 MoCo-v2를 사용한 자기지도학습 모델을 사용했다. 또한, 실험을 위한 데이터는 위의 실험과 같은 Non-IID 분포를 사용했으며, 각 클라이언트의 데이터에서 {0.5, 0.25, 0.1, 0.05, 0.01} 비율만큼을 임의의 표본을 추출했다. <표 3>에 나와있는것과 같이 자기지도학습 모델을 사용한 경우 전체 데이터를 사용했을 때의 성능 대비 일부만을 사용했을 때의 하락폭이 적은 것을 확인할 수 있다. 이 결과를 통해 자기지도학습 모델을 Global 모델로 사용하게 되면, 클라이언트가 데이터를 레이블링하기 어렵다는 연합학습의 한계점을 어느 정도 완화할 수 있다는 것을 알 수 있다.



〈그림 6〉 수렴까지 필요한 커뮤니케이션 라운드

4.5 수렴 속도 Communication Round

연합학습은 서버와 학습에 참여하는 클라이언트 사이의 커뮤니케이션에 기반해 진행되므로, 학습에 필요한 커뮤니케이션의 횟수가 많아질수록 지연시간과 통신 비용이 증가하게 된다. 그러므로 학습을 얼마만큼 효율적으로 진행할 수 있는지는 모델의 성능과 더불어 연합학습에서 중요한 지표이다.

우리는 연합학습에서 자기지도학습 모델을 사용하는 경우 커뮤니케이션 라운드에서의 이점과 유용성을 알아보기 위해 Global 모델의 수렴에 필요한 통신의 횟수에 관해 탐구했다. 이 실험에는 랜덤 초기화와 MoCo-v2 자기지도학습 모델을 사용해 비교를 진행했다. 그 결과 <그림 5>에서 볼 수 있듯, 자기지도학습 모델을 사용한 결과가 월등히 뛰어난 수렴속도를 보이며 적은 커뮤니케이션 라운드가 필요하다는 것을 확인할 수 있다. 이 결과를 통해 연합학습의 Global 모델로 사용하게 되면, Non-IID 환경에서의 성능향상뿐만 아니라 빠른 수렴 속도로 인해 커뮤니케이션 라운드도 감소시킬 수 있다.

5. 결론

본 연구는 흉부 엑스레이 데이터를 사용한 연합학습에서 발생할 수 있는 문제를 자기지도학습 모델을 사용함으로써 해결하는 방안을 제안하였다. 이를 위해 Instance Discrimination 방법 및 Heurstic 하게 디자인된 자기지도학습뿐만 아니라, ImageNet 데이터로 학습된 상황까지 고려해 다양한 실험을 진행함으로써 유용성을 검증하였다. 특히 Non-IID 환경에서 진행한 실험에서는 랜덤하게 초기화된 경우보다 약 13%의 엄청난 성능 향상을 보이며 제안한 방법이 효과적임을 증명했다. 이를 통해 연합학습의 실질적인 시나리오인 Non-IID 환경에서 더욱 다양한 분야에 적용할 수 있을 것이다.

본 연구는 총 4개의 자기지도학습 방법론만을 사용해 실험 및 검증을 진행했고, 서버에서 자기지도학습을 진행해야 한다는 한계점이 존재한다. 이를 보완하기 위해 후속 연구에서는 최근 제안되고 있는 자기지도학습 방법론들을 사용해 검증할 필요가 있다. 또한, 자기지도학습 자체도 연합학습 환경에서 진행하기 위한 방안을 연구할 계획이다.

참고문헌(References)

[국내 문헌]

공희산, 박재훈, 김광수. (2021). 의료 데이터의 자기지도학습 적용을 위한 pretext task 분석. *2021 한국정보통신학회 춘계학술대회*. 38-40.

[국외 문헌]

Baltruschat, I. M., Nickisch, H., Grass, M., Knopp, T., & Saalbach, A. (2019). Comparison of deep learning approaches for multi-label chest X-ray classification. Scientific reports, 9(1), 1-10.

Banerjee, S., Misra, R., Prasad, M., Elmroth, E., & Bhuyan, M. H. (2020, November). Multi-diseases classification from chest-X-ray: A federated deep learning approach. In Australasian Joint Conference on Artificial Intelligence (pp. 3-15). Springer, Cham.

Caron, M., Misra, I., Mairal, J., Goyal, P., Bojanowski, P., & Joulin, A. (2020). Unsupervised learning of visual features by contrasting cluster assignments. Advances in Neural Information Processing Systems, 33, 9912-9924.

Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020, November). A simple framework for contrastive learning of visual representations. In International conference on machine learning (pp. 1597-1607). PMLR.

Chen, X., Fan, H., Girshick, R., & He, K. (2020).

Improved baselines with momentum contrastive learning. arXiv preprint arXiv:2003.04297.

Doersch, C., Gupta, A., & Efros, A. A. (2015). Unsupervised visual representation learning by context prediction. In Proceedings of the IEEE international conference on computer vision (pp. 1422-1430).

Grill, J. B., Strub, F., Altché, F., Tallec, C., Richemond, P., Buchatskaya, E., ... & Valko, M. (2020). Bootstrap your own latent-a new approach to self-supervised learning. Advances in neural information processing systems, 33, 21271-21284.

Guan, Q., Huang, Y., Luo, Y., Liu, P., Xu, M., & Yang, Y. (2021). Discriminative Feature Learning for Thorax Disease Classification in Chest X-ray Images. IEEE Transactions on Image Processing, 30, 2476-2487.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., & Ciurea-Ilcus, S. A large chest radiograph dataset with uncertainty labels and expert comparison. In Proc AAAI Conf Artif Intell (No. 33, p. 590). CheXPert. https://stanfordmlgroup.github.io/competitions/chexpert/

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., ... & Zhao, S. (2021). Advances and open problems in federated learning. Foundations and Trends® in Machine Learning, 14(1–2), 1-210.

Karimireddy, S. P., Kale, S., Mohri, M., Reddi, S., Stich, S., & Suresh, A. T. (2020, November). Scaffold: Stochastic controlled averaging for federated learning. In International Conference on Machine Learning (pp. 5132-5143). PMLR.

Komodakis, N., & Gidaris, S. (2018, April). Unsupervised representation learning by predicting image rotations. In International Conference on Learning Representations (ICLR).

Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks. Proceedings of Machine Learning and Systems, 2, 429-450.

Mangal, A., Kalia, S., Rajgopal, H., Rangarajan, K., Namboodiri, V., Banerjee, S., & Arora, C. (2020). CovidAID: COVID-19 detection using chest X-ray. arXiv preprint arXiv:2004.09803.

McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In Artificial intelligence and statistics (pp. 1273-1282). PMLR.

Misra, I., & Maaten, L. V. D. (2020). Self-supervised learning of pretext-invariant representations. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6707-6717).

Noroozi, M., & Favaro, P. (2016, October). Unsupervised learning of visual representations by solving jigsaw puzzles. In European conference on computer vision (pp. 69-84). Springer, Cham.

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... & Ng, A. Y. (2017). Chexnet: Radiologist-level pneumonia detection on chest

x-rays with deep learning. arXiv preprint arXiv:1711.05225.

Wang, J., Liu, Q., Liang, H., Joshi, G., & Poor, H. V. (2020). Tackling the objective inconsistency problem in heterogeneous federated optimization. Advances in neural information processing systems, 33, 7611-7623.

Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, Ronald Summers, ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases, IEEE CVPR, pp. 3462-3471, 2017. NIH Chest X-ray14, https://www.kaggle.com/datasets/nih-chest-xrays/data

Yan, B., Wang, J., Cheng, J., Zhou, Y., Zhang, Y., Yang, Y., ... & Liu, B. (2021, July). Experiments of federated learning for COVID-19 chest X-ray images. In International Conference on Artificial Intelligence and Security (pp. 41-53). Springer, Cham.

Zbontar, J., Jing, L., Misra, I., LeCun, Y., & Deny, S. (2021, July). Barlow twins: Self-supervised learning via redundancy reduction. In International Conference on Machine Learning (pp. 12310-12320). PMLR.

Abstract

Self-supervised Meta-learning for the Application of Federated Learning on the Medical Domain

Heesan Kong* · Kwangsu Kim**

Medical AI, which has lately made significant advances, is playing a vital role, such as assisting clinicians with diagnosis and decision-making. The field of chest X-rays, in particular, is attracting a lot of attention since it is important for accessibility and identification of chest diseases, as well as the current COVID-19 pandemic. However, despite the vast amount of data, there remains a limit to developing an effective AI model due to a lack of labeled data. A research that used federated learning on chest X-ray data to lessen this difficulty has emerged, although it still has the following limitations. 1) It does not consider the problems that may occur in the Non-IID environment. 2) Even in the federated learning environment, there is still a shortage of labeled data of clients. We propose a method to solve the above problems by using the self-supervised learning model as a global model of federated learning. To that aim, we investigate a self-supervised learning methods suited for federated learning using chest X-ray data and demonstrate the benefits of adopting the self-supervised learning model for federated learning.

Key Words : Self-supervised Learning, Federated Learning, Chest X-ray, Medical AI

Received : October 6, 2022 Revised : October 18, 2022 Accepted : October 24, 2022
Corresponding Author : Kwangsu Kim

* Department of Computer science and Engineering, Sunkyunkwan University.
** Corresponding Author: Kwangsu Kim
Department of Computer science and Engineering, Sunkyunkwan University,
2066 Seobu-ro Jangan-gu, Suwon-si, 16419, Gyeonggi-do, Republic of Korea
Tel: +82-031-290-7969, E-mail: kim.kwangsu@skku.edu

저 자 소 개

공희산
2021년: 한국항공대학교 소프트웨어학과 공학사
2021년~현재: 성균관대학교 일반대학원 소프트웨어학과 석사과정
관심분야 : Computer Vision, Self-supervsied Learning
ORCID Number : 0000-0002-0904-9359
E-mail : khtks11@gmail.com

김광수
1994년: 한양대학교 전자공학과 공학사
2007년: University of Southern California, Computer Science 공학박사
1994년~2019년: 과학기술정통부 국장
2019년~현재: 성균관대학교 소프트웨어학과 교수
2019년~현재: 성균관대학교 인공지능 연구소 소장
관심분야 : Computer Vision, Domain Adaptation, Federated Learning, AI Application
ORCID Number : 0000-0002-1097-8112
E-main : kim.kwangsu@skku.edu