JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 27 No. 11, pp. 13-18, November 2022
https://doi.org/10.9708/jksci.2022.27.11.013

Design of weighted federated learning framework based on local model validation

Jung-Jun Kim*, Jeon Seong Kang*, Hyun-Joon Chung*, Byung-Hoon Park**

*Researcher, Korea Institute of Robotics & Technology Convergence, Pohang, Korea
*Researcher, Korea Institute of Robotics & Technology Convergence, Pohang, Korea
*Principal Researcher, Korea Institute of Robotics & Technology Convergence, Pohang, Korea
**CEO, Software Architect, T3Q Co., Ltd., Seoul, Korea

[Abstract]

In this paper, we proposed VW-FedAVG(Validation based Weighted FedAVG) which updates the global model by weighting according to performance verification from the models of each device participating in the training. The first method is designed to validate each local client model through validation dataset before updating the global model with a server side validation structure. The second is a client-side validation structure, which is designed in such a way that the validation data set is evenly distributed to each client and the global model is after validation. MNIST, CIFAR-10 is used, and the IID, Non-IID distribution for image classification obtained higher accuracy than previous studies.

▸ **Key words**: AI, Federated Learning, Deep Learning, Mobile Computing, Object Classification

[요 약]

본 논문에서는 학습에 참여하는 각 디바이스의 모델들로부터 성능검증에 따라 가중치를 두어 글로벌 모델을 업데이트하는 VW-FedAVG(Validation based Weighted FedAVG)를 두 가지 방식으로 제안 한다. 첫 번째 방식은 서버 검증(Server side Validation) 구조로 글로벌 모델을 업데이트 하기 전에 각 로컬 클라이언트 모델을 하나의 전체 검증 데이터셋을 통해 검증하도록 설계 했다. 두 번째는 클라이언트 검증(Client side Validation) 구조로 검증 데이터셋을 각 클라이언트에 고르게 분배하여 검증을 한 후 글로벌 모델을 업데이트 하는 방식으로 설계 했다. 전체 실험에 적용한 데이터셋은 MNIST, CIFAR-10으로 이미지 분류에 대해 IID, Non-IID 분포에서 기존 연구 대비 더 높은 정확도를 얻을 수 있었다.

▸ **주제어**: 인공지능, 연합학습, 딥러닝, 모바일 컴퓨팅, 객체 분류

• First Author: Jung-Jun Kim, Corresponding Author: Hyun-Joon Chung, Co-Author: Jeon Seong Kang, Byung-Hoon Park
*Jung-Jun Kim (jjkim@kiro.re.kr), Korea Institute of Robotics & Technology Convergence
*Jeon Seong Kang (kjs2605@kiro.re.kr), Korea Institute of Robotics & Technology Convergence
*Hyun-Joon Chung (hjchung@kiro.re.kr), Korea Institute of Robotics & Technology Convergence
**Byung-Hoon Park (warmpark@t3q.com), Software Architect, T3Q Co., Ltd.
• Received: 2022. 09. 06, Revised: 2022. 10. 07, Accepted: 2022. 10. 26.

http://www.ksci.re.kr pISSN:1598-849X | eISSN:2383-9945

I. Introduction

최근 인공지능 기술의 발달로 헬스 케어, 스마트 빌딩, 자율주행 차, 원격 모니터링을 통한 제조 공정에 대한 고장 예측 및 유지 보수 등의 다양한 어플리케이션이 개발되고 있다. 이러한 어플리케이션에는 인공지능 모델이 탑재되고 지속적으로 데이터를 취득하여 학습을 필요로 하게 된다. 이때 수많은 데이터가 생성되고 이 데이터를 중앙 서버 또는 클라우드 환경으로 전송하여 처리하는 방식(Centralized Learning)은 인공지능 모델의 성능을 높이기에 유리하지만, 몇몇 실시간으로 처리해야 하는 어플리케이션 환경에는 대기 시간이 길어지는 문제가 발생한다[1]. 또한 헬스 케어 데이터는 개인의 정보가 담겨져 있어 이러한 데이터를 중앙 서버로 전송할 때 유출 될 수 있는 프라이버시 문제가 발생 할 수 있다[2,3].

이와 같은 문제를 해결하기 위해 중앙서버에서 데이터를 처리하는 것이 아닌, 데이터가 생성되는 클라이언트에서 인공지능 모델을 학습하고, 이를 취합하여 범용적인 인공지능 모델을 만드는 연합학습(Federated Learning)이 제안 되었다[4]. 현재까지 잘 알려져 있는 연합학습 방법 중 하나인 FedAVG(Federated Averaging)[5]는 각 로컬 클라이언트들로부터 전송 받은 파라미터를 평균 내어 글로벌 모델을 업데이트 하는 방식으로 학습이 이루어진다. 이때, 각 클라이언트에서 보유한 데이터양이 다르거나 label에 따른 데이터가 동일하지 않게 분산된 상태인 Non-IID(Non-Independent Identically Distributed) 환경에서의 학습과정이 불안정하다는 문제점이 발생있다[6]. 또한 전송 받은 모든 파라미터의 평균을 내기 때문에 비교적성능이 낮은 클라이언트의 파라미터들도 학습에 참여하게 되어 글로벌 모델이 업데이트 될 때 성능이 낮아 질 수 있는 문제가 있다.

본 논문에서는 이러한 문제점을 해결하기 위해 각 클라이언트가 학습을 진행 한 후, 글로벌 모델을 업데이트하기 전 검증 데이터셋에 대한 정확도를 계산하여 로컬 weight에 대해 가중치를 부여 한 후, 평균을 내는 방식인 VW-FedAVG(Validation based Weighted FedAVG)를 제안 한다. 검증을 할 때, 하나의 공통된 검증 데이터셋을 통해 서버에서 정확도를 계산하는 구조를 서버 검증(Server side Validation), 각 로컬 클라이언트에 검증 데이터셋을 분배해서 정확도를 계산하는 구조를 클라이언트 검증(Client side Validation)으로 나누어 구분 했다.

II. Preliminaries

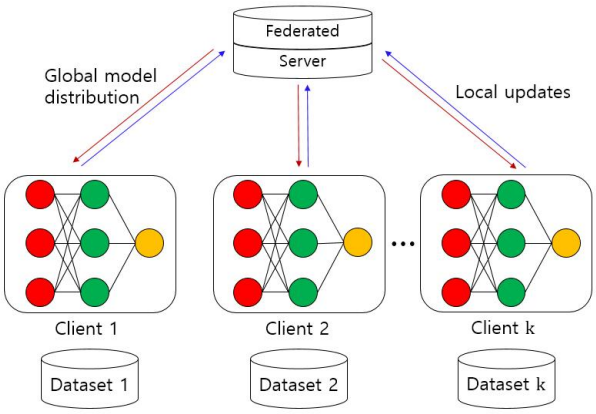

Fig. 1. Federated learning concept

연합학습의 일반적인 구조는 그림 1과 같은 Horizontal Federated Learning 구조를 갖는다[7]. k 개의 클라이언트들이 같은 구조의 데이터셋을 나누어 학습하고 학습된 파라미터들을 학습서버에 전송하여 글로벌 모델을 업데이트 한다. 연합학습에서 각 클라이언트가 한 번 학습이 진행되는 것을 1 local epoch이라고 하고, 클라이언트 모델들로부터 글로벌 모델이 한 번 업데이트 되는 것을 1 Communication round라고 한다.

이러한 클라이언트 모델의 학습 결과를 취합하는 방식에 따라 다양한 알고리즘이 존재한다. FedAVG[5]는 앞서 설명 하였듯이 로컬 클라이언트에서 전송 받은 파라미터들의 평균값으로 글로벌 모델을 업데이트 한다. 각 로컬 클라이언트가 보유한 데이터의 이질성 문제를 해결하기 위해 FedProx는 Proximal Term을 추가하여 업데이트 할 로컬 모델이 글로벌 모델과 유사하도록 만들어 안정성을 높이는 연구를 진행 하였다[12]. 또한 permutation invariant nature로 인해 global optima에 도달하기 어려워 이를 해결하기 위해 PFNM(Probabilistic Federated Neural Matching) 가 제안 되었다. 하지만 이 방식은 단순한 Network 구조에 대해 성능 향상이 미비하여 이를 개선하기 위해 layer-wise federated learning 방식인 FedMA(Federated Matched Averaging)을 제안 하였다[11]. SCAFFOLD(Stochastic Controlled Averaging algorithm)는 non-iid 환경에서 FedAVG가 불안정하고 수렴하는데 오래 걸리게 되는 'client-drift'현상을 해결하기 위해 제어 변수(control variate)를 사용했다[8]. 이처럼 FedAVG를 기반으로 학습의 성능을 높이려는 많은 연구들이 진행되어 왔다. 기존의 방식과는 다르게 학습 자체

는 각 client 모두 동일한 조건으로 구성 했지만 글로벌 모델을 업데이트 하는 방식에서 미리 검증을 통해 차등적으로 가중치를 부여한다는 것에서 차이점이 있다. 따라서 본 논문에서는 두 가지 서로 다른 검증 방법을 적용하는 연구를 진행 하고 그 결과를 기술 했다.

III. Proposed Method

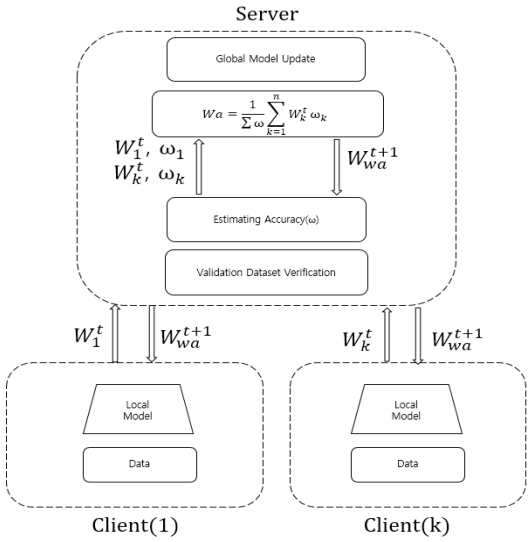

Fig. 2. Server Side Validation Architecture

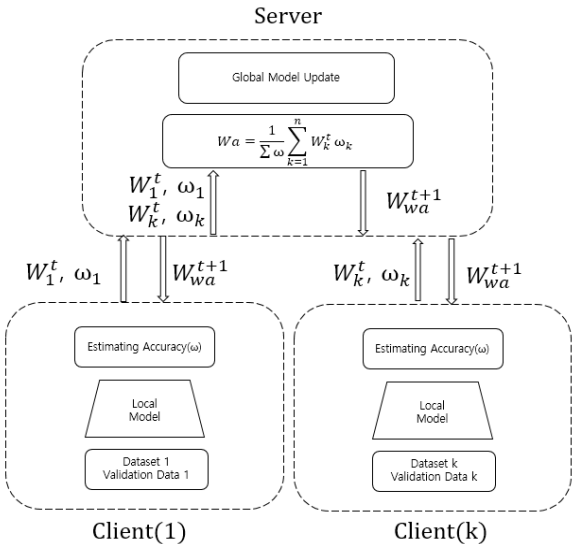

Fig. 3. Client Side Validation Architecture

1. Data structure and architecture

전체 데이터의 구성은 IID, Non-IID 두 가지 환경으로 규정하고 각 클라이언트에 랜덤하게 분배 한다. IID 환경에서는 전체 학습데이터의 수를 클라이언트의 개수로 나누어 같은 개수의 데이터를 갖도록 할당한다. 데이터 분배 전 학습데이터 10,000개를 검증 데이터셋으로 따로 분류를 하고 검증 아키텍처에 따라 각 클라이언트에 분배하거나 서버에 두어 검증을 하도록 구성했다. Non-IID 환경에서는 중복이 없이 학습데이터를 샤딩(Sharding) 하여 각 클라이언트에 분배 한다. 이때 각 클라이언트에서 갖게 되는 데이터의 양은 동일하지만 라벨링에 따른 데이터의 수는 랜덤하게 갖도록 구성했다. 이때 검증 데이터셋은 IID 방식과 똑같이 고르게 분배했다. 분배된 학습 데이터를 바탕으로 다음과 같이 두 가지 검증 방식으로 글로벌 모델을 업데이트 하도록 제안 했다.

먼저, 서버 검증(Server side validation) 아키텍처의 경우 validation dataset을 서버에 두어 한 번의 communication round가 학습되어 글로벌 모델을 업데이트하기 전, 각 로컬 클라이언트의 모델을 하나의 전체 validation dataset을 통해 검증 하도록 설계 했다. 두 번째로, 클라이언트 검증(Client side validation) 아키텍처의 경우 validation dataset을 각 클라이언트에 고르게 분배하여 마찬가지로 한 번의 communication round 학습할 때 마다 각 클라이언트에서 갖고 있는 validation dataset을 통해 검증하도록 설계 하였다.

2. Validation based Weighted FedAVG

본 논문에서는 클라이언트에서 학습한 모델을 서버 또는 로컬 클라이언트에서 검증을 통해 글로벌 모델을 업데이트 하는 방식인 VW-FedAVG 제안했다. 첫 째는 그림 2와 같이 각 클라이언트에서 학습을 거쳐 얻은 모델을 글로벌 validation dataset을 통해 정확도를 서버에서 검증하고 이를 바탕으로 가중치를 부여하여 글로벌 모델을 업데이트 하는 방식이고, 두 번째는 그림 3과 같이 글로벌 validation dataset을 각 클라이언트에 균등하게 분배하여 서버에서 검증을 하는 것이 아닌, 각 클라이언트에서 검증을 하여 가중치를 부여 하고 글로벌 모델을 업데이트 하는 방식으로 구성 했다.

$$W_a = \frac{1}{\sum_{k=1}^{n} w} W_k^t w_k \quad \cdots (1)$$

식 (1)은 모델을 업데이트하기 위한 가중치를 계산하는 식으로 서버 검증, 클라이언트 검증 두 가지 구조에 공통으로 사용 했다. 각 로컬 모델에서 얻은 정확도(Accuracy)를 매 학습 시 마다 구하여 각 모델의 weight에 곱하여 준 후 더하고, 더해준 가중치만큼을 나누어 주는 방법으로 새로운 weight값을 구하였다.

IV. Experiment Results

1. Experiment environment

본 논문에서는 MNIST[10], CIFAR-10[9] 데이터셋을 사용 하여 실험하였다. 먼저 MNIST 데이터셋은 각 클래스 별로 7,000개 씩 데이터를 갖으며 총 10개의 클래스로 70,000개의 이미지 샘플로 이루어져있다. 이중 학습에 사용하는 데이터는 60,000개이며 나머지 10,000개는 테스트셋으로 구성하여 각 클라이언트에서 분배되는 데이터가 중복 되지 않게 갖도록 하였다. CIFAR-10 데이터셋도 마찬가지로 10개의 클래스로 라벨링 되어 있으며 총 60,000개의 이미지 샘플로 이루어져 있다. 이중 50,000개는 학습 데이터로, 10,000개는 테스트셋으로 구성된다. 이 두 가지 데이터셋을 사용하여 FedAVG와 VW-FedAVG에 대한 정확도를 측정하였다.

Table 1. MNIST CNN Architecture

Layer	Output Shape	Param
Input	[-1,1,28,28]	-
Conv1	[-1,32,26,26]	800
ReLu	-	-
Max Pooling	-	-
Conv2	[-1,64,12,12]	51,200
ReLu	-	-
Max-Pooling	-	-
Flatten	[-1,3136]	-
Fully Connected	[-1, 512]	1,605,632
ReLu	-	-
Fully Connected	[-1, 10]	5,120

Table 2. CIFAR CNN Architecture

Layer	Output Shape	Param
Input	[-1,3,32,32]	-
Conv1	[-1,6,28,28]	456
ReLu	-	-
Max-Pooling	-	-
Conv2	[-1,16,10,10]	2,416
ReLu	-	-
Max-Pooling	-	-
Flatten	-	2,416
Fully Connected	[-1, 120]	48,120
ReLu	-	-
Fully Connected	[-1, 84]	10,164
ReLu	-	-
Fully Connected	[-1, 10]	850

Backbone은 Table 1, Table 2와 같이 각각 MNIST와 CIFAR 데이터셋에 적용하여 분류를 시행하였으며, 실험에 진행한 하이퍼 파라미터로는 FedAVG에서 사용한 값을 적용했다. 전체 client의 수는 100개로 설정하고, 10 local epoch, 500 communication rounds, SGD optimizer 0.9 momentum, 0.01 learning rate, 50 local batch size로 적용하여 실험을 진행 했다.

2. Experiment results

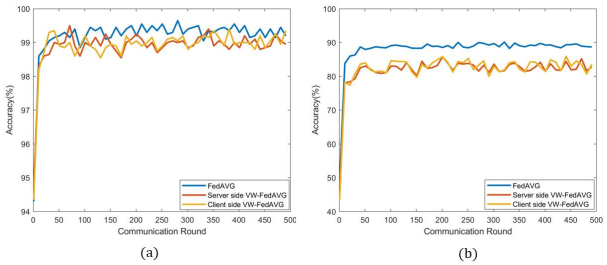

Fig. 4. Accuracy results tested by dividing the MNIST dataset into (a)IID, (b)Non-IID methods

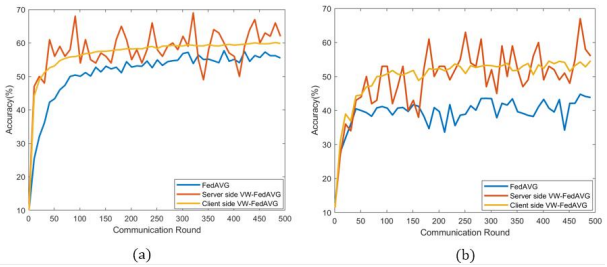

Fig. 5. Accuracy results tested by dividing the Cifar-10 dataset into (a)IID, (b)Non-IID methods

본 논문에서 설계한 2개의 구조에 대해 IID, Non-IID 환경에서 기존 연합학습 알고리즘인 FedAVG와의 성능 비교를 각각 MNIST, CIFAR-10 데이터로 실험을 진행 했다. Fig 4와 5는 각 10 Communication Round 마다 정확도 값을 기록하여 그린 그래프 이다. 먼저 Fig 4의 (a), (b)는 각각 MNIST 데이터를 IID, Non-IID 환경에서 실험한 결과로 파란색이 FedAVG, 주황색이 서버에서 검증한 VW-FedAVG, 노란색이 클라이언트에서 검증한 VW-FedAVG에 대한 결과이다. IID 환경에서는 대부분 FedAVG가 더 나은 성능을 보였지만 정확도에서는 크게 차이가 없던 반면, Non-IID 환경에서는 FedAVG가 더 좋은 성능을 갖는 것을 확인 했다.

Fig 5는 Cifar-10 데이터로 얻은 결과이며 IID, Non-IID 환경 모두 제안한 알고리즘이 FedAVG보다 더 높은 정확도를 보였다. (a)는 IID 환경에서의 Communication Round 마다 정확도를 측정한 값으로 거의 대부분 Round에서 VM-FedAVG가 더 좋은 성능을 보였다. 마찬가지로 (b)Non-IID 환경에서는 초기 50 Round 이하 에서는 비슷한 성능을 보였지만 이후에는 VM-FedAVG가 더 높은 정확도 값을 얻었다.

Table 3. Best Accuracy of FedAVG, VM-FedAVG in IID, Non-IID environment

Data Type	Algorithm	MNIST	CIFAR
IID	FedAVG	**99.7**	58.35
	VM-FedAVG (Server Side)	99.5	**74.1**
	VM-FedAVG (Client Side)	99.5	**60.51**
Non-IID	FedAVG	**90.9**	47.01
	VM-FedAVG (Server Side)	86.55	**67.2**
	VM-FedAVG (Client Side)	86.6	**54.89**

Table 3은 FedAVG와 VM-FedAVG에 대해 IID, Non-IID 환경에서 MNIST, CIFAR 데이터셋으로 학습한 정확도에 대한 결과이다. 전체 500 communication round 중에서 가장 높은 정확도 값을 나타낸다. 표에서 확인 할 수 있듯이 MNIST 데이터셋에 대해서는 큰 성능 차이가 없는 반면, CIFAR 데이터셋에 대해서는 서버 검증 VM-FedAVG가 IID, Non-IID 환경 모두 더 높은 정확도 값을 얻었다.

MNIST 데이터셋에 대해 기존의 FedAVG 보다 더 낮은 성능을 보였는데 Table 1과 같은 단순한 CNN 구조와 비교적 Cifar-10 보다 덜 복잡한 형태의 데이터셋으로 이루어져 크게 성능차이가 나지 않았다.

V. Conclusions

본 논문에서는 IID, Non-IID 환경에서 연합학습을 진행할 때, 서버와 클라이언트에서 로컬 모델의 정확도 검증을 통해 글로벌 모델을 업데이트하는 가중치 부여 방식인 VM-FedAVG를 제안 했다. 기존의 FedAVG보다 Cifar-10 데이터셋에서 더 높은 정확도 값을 얻을 수 있었다. 향후, 추가 연구로 VGG-Net[13] 또는 Resnet[14] 등의 복잡한 구조의 모델과 다양한 데이터셋 실험으로 연구를 진행하려 한다.

ACKNOWLEDGEMENT

This research is supported by Defense Industry Technology Center (UC200019D).

REFERENCES

[1] Ding, Chuntao, et al. "Resource-aware feature extraction in mobile edge computing." IEEE Transactions on Mobile Computing Vol. 21. No. 1, pp. 321-331. 2020. DOI: 10.1109/TMC.2020.3007456

[2] Ren, H., Li, H., Dai, Y., Yang, K., & Lin, X.. “Querying in internet of things with privacy preserving: Challenges, solutions and opportunities.” IEEE Network, Vol. 32. No. 6, pp. 144-151. 2018. DOI: 10.1109/MNET.2018.1700374

[3] Garcia, David, et al. "Analyzing gender inequality through large-scale Facebook advertising data." Proceedings of the National Academy of Sciences Vol. 115. No. 27, pp. 6958-6963, 2018. DOI: 10.1073/pnas.1717781115

[4] McMahan, Brendan, and Daniel Ramage. "Federated learning: Collaborative machine learning without centralized training data." Google Research Blog 3 2017.

[5] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, Blaise Ag era y Arcas, “Communication-efficient learning of deep networks from decentralized data,” arXiv:1602.05629, (2016) DOI: 10.48550/arXiv.1602.05629

[6] Mann Soo Hong, Seok-Kyu Kang, Jee-Hyong Lee, “Removing Out-of-Distribution Clients on Federated Model”, Journal of The Korean Institute of Intelligent Systems, Vol. 30, No. 6, pp. 488-493, 2020

[7] Yang, Q., Liu, Y., Chen, T. and Tong, Y. “Federated machine learning: Concept and applications.” ACM Transactions on Intelligent Systems and Technology (TIST), Vol. 10. No. 2, pp. 1-19, 2019. DOI:10.1145/3298981

[8] Karimireddy SP, Kale S, Mohri M, Reddi S, Stich S, Suresh AT. “Scaffold: Stochastic controlled averaging for federated learning.” InInternational Conference on Machine Learning (PMLR.) pp. 5132-5143, Nov. 21. 2020. DOI : 10.48550/arXiv.1910.06378

[9] Krizhevsky, Alex, and Geoffrey Hinton. "Learning multiple layers of features from tiny images." (2009): 7.

[10] Deng, Li. "The mnist database of handwritten digit images for machine learning research [best of the web]." IEEE signal processing magazine Vol. 29, No. 6, pp. 141-142. 2012. DOI: 10.1109/MSP.2012.2211477

[11] Wang H, Yurochkin M, Sun Y, Papailiopoulos D, Khazaeni Y. “Federated learning with matched averaging”. arXiv preprint arXiv:2002.06440. Feb. 15. 2020. DOI : 10.48550/arXiv.2002.06440

[12] SSahu, Anit Kumar, et al. "On the convergence of federated optimization in heterogeneous networks." arXiv preprint arXiv:1812.06127 Dec. 3 2018. DOI : 10.48550/arXiv.1812.06127

[13] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014). DOI: 10.48550/arXiv.1409.1556

[14] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770-778, 2016. DOI: 10.1109/CVPR.2016.90

Authors

Jung-Jun Kim received the B.E., and M.E. degrees in Computer Engineering from KyungHee University, Korea, in 2015, 2017 respectively. Jung-Jun Kim joined the researcher of Korea Institute of Robotics & Technology Convergence, Pohang, Korea in 2017. He is currently a Researcher in AI Robotics Center. He is interested in Computer Vision, Artificial Intelligent.

Jeon Seong Kang received the B.E., and M.E. degrees in Electronics and Electrical Engineering from Suwon University, Dongguk University, Korea, in 2016, 2018 respectively. Jeon Seong Kang joined the researcher of Korea Institute of Robotics & Technology Convergence, Seoul, Korea in 2021. He is currently a Researcher in AI Robotics Center. He is interested in Computer Vision, Artificial Intelligent.

Hyun-Joon Chung received M.S. and Ph.D. degrees in mechanical engineering from the University of Iowa, Iowa City, USA, in 2005 and 2009, respectively. He was a research assistant and postdoctoral research scholar in the Center for Computer Aided Design from 2005 to 2015. He Joined the Korea Institute of Robotics and Technology Convergence as a Senior Researcher in 2015. He is currently a Principal Researcher and the Head of AI Robotics Center. He serves as a General Affairs Director of Field Robot Society in Korea. His research interests include dynamics and control, optimization algorithms, computational decision making, robotics, modeling and simulation.

Byung-Hoon Park is the CEO of T3Q which is a AI and bigdata company. He completed Computer education major at Korea University Graduate School. He has worked as a specialist in software architecture and methodology of software development and as a professor at Samsung Multicampus. Since T3Q was founded in 2007, the company has been developing various open source-based applications and providing bigdata / AI integrated platform (T3Q.ai) to public institutions and private companies.