

<https://doi.org/10.7236/JIIBC.2024.24.1.83>  
JIIBC 2024-1-13

## 한정된 레이블 데이터를 이용한 효율적인 철도 표면 결함 감지 방법

### An Efficient Detection Method for Rail Surface Defect using Limited Label Data

한석민\*

Seokmin Han\*

**요약** 본 연구에서는 Railroad surface 데이터를 활용하여 Semi-Supervised learning 방식으로 railroad surface의 defect를 검출해내는 방안을 제안한다. Resnet50에 ImageNet으로 pretrained된 모델을 이용한다. Label이 없는 데이터에서 무작위로 데이터를 선정, 선정된 데이터에 label을 부여한 뒤 이 데이터로 모델을 학습시킨다. 학습된 모델을 이용하여 나머지 데이터의 결과값을 예측한 후, 그 예측값이 일정한 threshold보다 큰 것을 골라내고, threshold보다 큰 값들을 값이 큰 순서대로 정렬하여, 일정한 크기만큼 training data에 추가한다. 이 때, 각 class에 속할 확률이 높은 쪽으로 pseudo-labeling을 수행한다. 초기에 label이 부여된 데이터 개수에 따른 전체적인 class 분류 성능을 확인하는 실험 또한 진행하였고, 전체 training data대비 10% 미만의 labeled data로 최대 98%의 정확도를 얻는 성능을 보였다.

**Abstract** In this research, we propose a Semi-Supervised learning based railroad surface defect detection method. The Resnet50 model, pretrained on ImageNet, was employed for the training. Data without labels are randomly selected, and then labeled to train the ResNet50 model. The trained model is used to predict the results of the remaining unlabeled training data. The predicted values exceeding a certain threshold are selected, sorted in descending order, and added to the training data. Pseudo-labeling is performed based on the class with the highest probability during this process. An experiment was conducted to assess the overall class classification performance based on the initial number of labeled data. The results showed an accuracy of 98% at best with less than 10% labeled training data compared to the overall training data.

**Key Words** : Rail surface, Semi-supervised, thresholding, sorting, selection.

\*정회원, 한국교통대학교 데이터사이언스전공  
접수일자 2023년 12월 21일, 수정완료 2024년 1월 21일  
게재확정일자 2024년 2월 9일

Received: 21 December, 2023 / Revised: 21 January, 2024 /  
Accepted: 9 February, 2024

\*Corresponding Author: seokmin.han@ut.ac.kr  
Dept. of Data Science, Korea National University of  
Transportation, Korea.

## I. 서 론

최근 인공지능 기술의 발전이 빨라지면서, 여러 분야에서 인공지능 기술을 접목하려는 시도 자체로도 산업 분야에 빠른 변화를 가져오고 있다<sup>[1]</sup>. 이러한 변화는 주로 인공지능을 적용한 자동화를 통해 인간의 노동을 대체하는 형태를 보이고 있다. 예를 들어, 딥러닝을 활용한 자재 품질 검사나<sup>[2-4]</sup> 드론을 활용한 송전 선로 검사<sup>[5]</sup>, 자율주행을 통한 운전자 대체<sup>[6]</sup> 등에서 자동화가 시도되고 있다. 철도 분야에서도 기존에 인력으로 수행하던 일들을, 인공지능을 도입하여 인력을 감축하거나 대체하기 위해 많은 연구가 이루어지고 있다. 특히, 철도 선로의 점검을 위한 딥러닝 기반의 자동 레일 탐지 시스템<sup>[7-9]</sup>과 같은 연구도 진행되고 있다. 이러한 자동화는 사람의 피로도에 따른 업무 효율 저하를 방지하고, 사람의 업무량을 감소시켜 시간 및 비용을 절감하는 것을 목표로 한다. 철로 표면의 결함 탐지의 경우, 인공지능을 활용한 효율적인 결함 검출 방식이 있다면, 사람이 육안으로 검사하는 방식을 대체할 수 있다. 이러한 기술을 위해서는 인공지능을 지도 학습 방식으로 훈련 시키는 방법이 이용되어왔고<sup>[7, 10, 11]</sup>, 이는 인공지능에 사용되는 데이터에 label을 부여하기 위한 많은 비용과 시간이 소모되는 단점이 있다.

이러한 비용과 시간 소모를 해결하기 위해 본 연구에서는 학습에 필요한 label을 부여하는데 들어가는 노력을 절감하기 위한 준지도 학습(Semi-supervised learning)<sup>[12]</sup>을 활용하여 labeling에 소요되는 시간과 비용의 문제

를 해결하고자 한다. 준지도 학습은 데이터에 레이블을 부여하는 과정을 최소화 하기 위한 방법으로, 적은 수의 labeled 데이터를 활용하여 labeling을 위한 시간적 비용을 절감할 수 있는 방법 이라고 할 수 있다. 일반적으로 적은 데이터로 학습한 인공지능망의 신뢰도는 낮다고 생각되기도 하지만, 철로 표면 이미지의 특성상 결함이 뚜렷하고, 모양이 비교적 균일하게 비슷한 데이터라는 특징을 가지고 있어, 적은 수의 데이터를 기반으로 다른 데이터들의 label을 추정하는 방식으로 충분히 좋은 성능의 모델을 만드는 것이 가능할 것으로 예상된다. 본 연구에서는 준지도 학습을 적용하여 철도 표면 결함 검사 작업을 자동화하고 이상 여부를 판별하는 시간 및 비용을 절감하는 방법을 제안하고, 초기 labeled data의 개수에 따른 성능 변화를 살펴보도록 한다.

## II. 본 론

그림 1은 본 연구의 전체적인 개념을 나타낸다. 전체 데이터 1485장중 359장은 defect 데이터이고, 나머지 1126장은 normal 데이터이다. 우선, 전체 데이터 1485장중 60% 가량인 885장을 random sampling 하여 test set으로 지정하고, 이 test set은 training 과정에 참여하지 않는다. 나머지인 40%가량에 해당하는 600장의 일부 데이터에 대해 annotation을 제공하여 labeled data를 만들고, 나머지 데이터를 unlabeled training data로 남겨두도록 한다. 이 unlabeled data는 training을 진행하면서 pseudo-labeling을 수행하여 annotation을

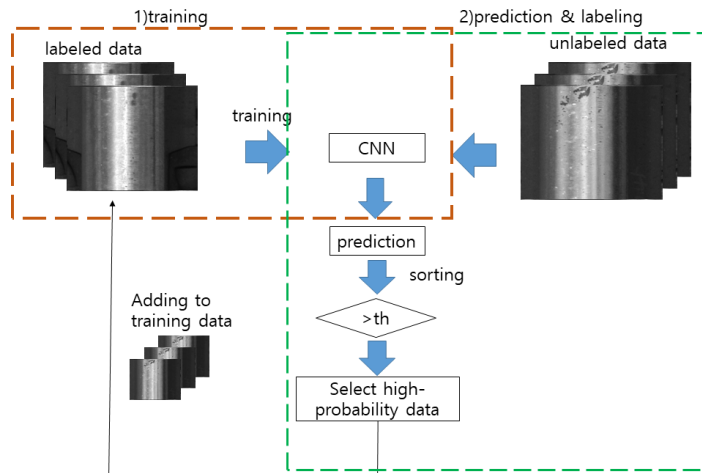


그림 1. 본 연구의 전체적인 개념도

Fig. 1. Overview of the proposed method

수행하도록 한다. labeled data를 우선적으로 backbone neural network인 ResNet50<sup>[13]</sup>을 이용하여 training 한다. 입력 데이터는 railroad surface 영상이고, defect가 있는 영상을 defect class, 정상으로 판별된 영상을 normal class로 지정한다. labeled image로 training 된 neural network에, unlabeled data의 1/5을 인가 하여 prediction result를 얻는다. 1/5을 선택한 이유는, 5 step을 거쳐 전체 unlabeled data들에 pseudo-labeling을 수행하기 위함이다. 이렇게 얻은 prediction result는 각 class에 해당할 확률을 나타낸다. defect class에 해당하는 prediction result를 내림차순으로 정리한 후, threshold인 0.5보다 큰 값을 지니는 값을 가지면서, prediction값이 상위 30%에 해당하는 데이터를 defect class로 간주한다. 즉, defect class일 확률이 높은 순서대로 내림차순으로 정리한 후, 결과값이 threshold 보다 큰 데이터중, 정렬된 데이터의 3/10을 defect로 간주하고, 나머지 데이터인 7/10의 데이터를 normal로 간주한다. (이 논문에서는, defect class와 normal class의 비율을 알고 있는 것으로 가정한다.) 그리고, 이렇게 selection한 대로 pseudo-labeling을 수행한다. pseudo-labeling된 데이터와 이미 지니고 있던 labeled data를 합하여 새로운 training dataset을 만들고, neural network를 다시 학습시킨다. 이 과정을 5 step에 걸쳐 반복한다. 그러면, 모든 데이터에 대해 neural network의 학습을 마칠 수 있다. 이 과정을 pseudo-code로 정리하면 다음과 같다.

*select sample training data randomly  
 annotate the selected sample data  
 separate the unlabeled data to 5 groups*

*Repeat 5 loops:  
 train neural network with the selected data  
 predict on the one unlabeled data group  
 sort the result  
 separate the predicted data into 2 classes  
 add the predicted data to the training data*

### III. 실험 및 결과

그림 2는 본 연구에서 사용되는 철로이미지의 예시이다. 본 연구에 사용된 이미지는 한국철도에서 촬영하여 제공하였다. 전체 데이터는 1485장이고, 그 중 defectd

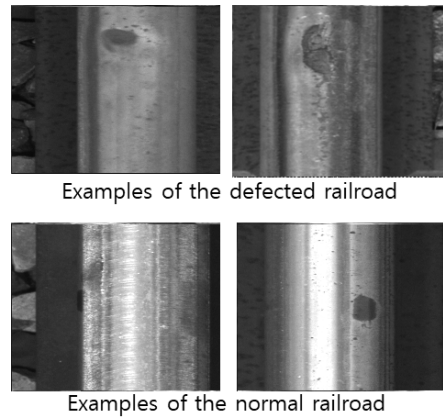


그림 2. 실험에 사용된 철로이미지의 예.  
 Fig. 2. Some examples of the railroad images.

data는 359장, normal data는 1126장이다. 실험에 사용한 시스템은 인텔사의 I7-13700을 기반으로한 PC에 RTX4090을 장착하였다. Backbone Neural net으로는 ResNet50을 사용하였다. Loss 함수로는 cross entropy를 사용하였고, learning rate은 0.0001로 설정하였으며, Adam optimizer를 적용하였다. 사용된 모델의 구조는 그림 3에 나타나 있다. 그림 3에서 FC는 fully connected layer를 뜻한다. 본 연구에서는, Backbone network로 ResNet50을 사용하되, ImageNet을 이용하여 사전학습(pretraining)하였다. ResNet50을 초기 지도 학습에서 쓰이는 데이터와 크기를 각각 3, 13개로 설정하였고,  $L_D$ 로 명칭하도록 한다. 각각 다른 크기의 초기 데이터  $L_D$ 로 Backbone Neural Network를 학습시키며, 학습 데이터 내에서 설정된 크기에 맞게 무작위로 추출한다. 이를 통해 초기 annotation이 부여된 레이블의 개수에 따른 모델의 성능 변화를 살펴보고자 한다.  $L_D$ 를 빼고 남은 데이터는 unlabeled data이며,  $U_D$ 로 명명하기로 한다.  $L_D$ 로 학습을 마친 neural net 모델에 unlabeled training data에 대한 prediction결과를 얻는다. 이 prediction 결과를 통해 에서 설명한대로 Pseudo Labeling을 수행하고,  $L_D$ 에 추가하도록 한다. 이 한 번의 과정을 사이클로 부르도록 한다. 한 번의 사이클에는 전체 training 데이터 600장의 training data 중  $L_D$ 만큼 labeling이 주어진 것이고,  $U_D$  labeling이 주어지지 않은 상태로 학습을 시작한다. 이를 통해 neural net이 pseudo-labeling된 데이터를 이용한 학습에서 얼마만큼 성능이 향상되는지를 보도록 한다.

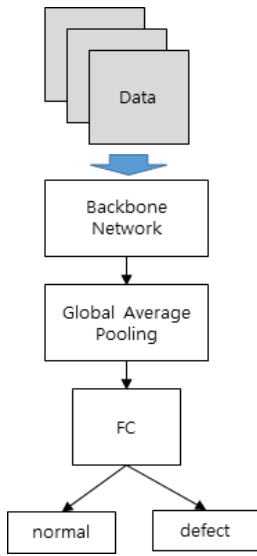


그림 3. 사용된 모델 개념도  
Fig. 3. The concept of the used model

peudo-labeling 이후에는, 레이블을 부여한  $U_D$  는  $L_D$ 에 추가하며, 해당 데이터를  $U_D$ 에서 삭제하도록 한다. 초기  $L_D$ 의 크기에 상관없이 준지도 학습의 각 사이클이 끝난 이후에는 Test Data에 대한 예측을 진행한다.

학습은 validation set을 사용하지 않고, 일괄적으로 20 에포크에 끝내도록 하였고, 이 때의 상태로 Test Data에 대해 예측한다. 초기 학습 데이터의 크기의 영향력을 파악하고 이에 대한 일반성을 얻기 위하여, 동일한 조건 및 실험 방법을 이용하여 5번의 실험을 하여 평균과 표준편차를 구하였다.

초기 데이터를 단지 3개의 labeled 데이터로 시작한 모델을 이용한 실험의 평균, 표준편차, 최대값 및 최소값이 표 1에 표시되어 있다. 최초로 3개의 label로 시작하였고, 이를 바탕으로 neural net을 학습시킨 후 unlabeled training data들 중 pseudo-labeling을 통하여 35개, 65개, 85개, 120개의 데이터에 대해 label을 부여한 상태에서 최종 150개의 training data들을 학습 한 후, 별도로 마련된 test set에 대한 성능을 나타내었고, 이 결과값이 표 1에 나타나있다. 표 1에는 5번의 반복된 실험의 평균, 표준편차, 최대값, 최소값이 표시되어있다. 제일 왼쪽 상단의 3은 최초의  $L_D$ 의 개수를 뜻하고, 35, 65, 85, 120은 싸이클이 진행되면서 pseudo-labeling을 수행하여 labeled data가 된 training data의 개수이다. 즉, 표 1의 경우에는 최초의 3장만 expert에 의한 label이 제공된 것이고, 나머지 데이터는 알고리즘을 통

표 1. 3개의 label로 시작한 경우의 결과값

Table 1. The results of the cases with 3 initial labels

	3	35	65	85	120
avg	0.7892	0.7973	0.7854	0.7830	0.8072
std	0.0309	0.0491	0.1377	0.2187	0.2219
min	0.7537	0.7657	0.6	0.4567	0.4687
max	0.8254	0.8813	0.9231	0.9731	0.9843

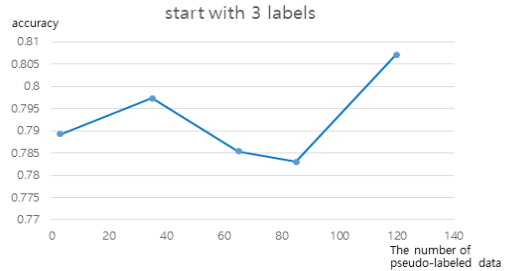


그림 4. 3개의 label로 시작한 경우의 결과값 그래프

Fig. 4. The result graph of the cases with 3 initial labels

해 pseudo-labeling이 수행된 것이다. 이런 측면에서 준지도 학습(semi-supervised learning)으로 부를 수 있다.

그림 4에는, 표 1에서 나타난 실험을 보기 편하도록 그래프로 나타내었다. 초기 데이터를 13개로 시작한 실험의 경우는 유사한 방식으로 표 2에 실험결과를 정리하였다. 5번의 반복된 실험에서의 평균, 표준편차, 최대값, 최소값이 표시되어있고, 이를 보기 편하도록 그림 5에 그래프로 나타내었다.

표 2. 13개의 label로 시작한 경우의 결과값

Table 2. The results of the cases with 13 initial labels

	13	35	65	85	120
avg	0.7670	0.8193	0.8633	0.9161	0.9208
std	0.0380	0.0743	0.0943	0.1171	0.1259
min	0.7022	0.6955	0.7231	0.7082	0.6993
max	0.7963	0.8963	0.953	0.9784	0.9903

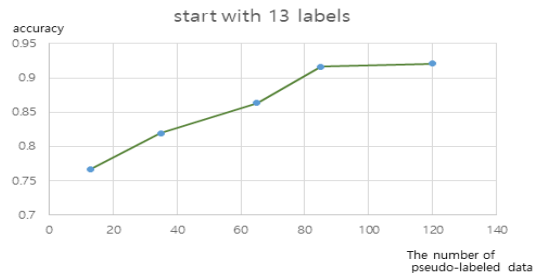


그림 5. 13개의 label로 시작한 경우의 결과값 그래프

Fig. 5. The result graph of the cases with 13 initial labels

표 3. 21개의 label로 시작한 경우의 결과값  
 Table 3. the results of the cases with 21 initial labels

	21	45	69	93	117
avg	0.7879	0.8912	0.9554	0.9727	0.9894
std	0.0144	0.0234	0.0248	0.0160	0.0024
min	0.7731	0.8567	0.9269	0.947	0.9851
max	0.8075	0.9224	0.991	0.9896	0.9910

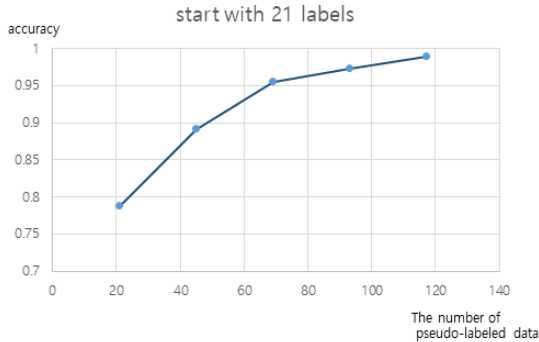


그림 6. 21개의 label로 시작한 경우의 결과값 그래프  
 Fig. 6. The result graph of the cases with 21 initial labels

마찬가지로, 초기데이터를 21개로 시작한 실험의 경우도 유사한 방식으로 표3에 실험 결과를 정리하였고, 이를 보기 편하도록 그림 6에 그래프로 나타내었다.

초기 데이터가 많으면 많을수록 pseudo-labeling에서의 불규칙적인 성능의 변화가 작아지는 경향을 보이는 것을 알 수 있다. 초기 데이터의 크기는 모델의 최종 성능에 영향을 끼치고 있으며, 초기값의 개수가 적을 경우에는 특히 일정하지 않은 성능을 나타내는 경향이 강하게 나타난다.

그 결과, 최고 성능으로 평균 98.9%를 기록하였으며, 전체 training 데이터 600장 모두를 expert에 의한 레이블을 가지고 학습을 진행하였을 때의 성능은 99%를 기록하였다. 이 결과를 통해 training data수 대비 5% 미만의 매우 적은 수의 데이터만을 가지고도 일정 수준의 성능을 이끌어 낼 수 있으며, 초기 데이터의 크기에 따라 최종 성능에서 유의미한 차이를 관찰할 수 있었다.

#### IV. 결 론

본 연구에서는 철로 표면 결함을 감지하기 위해 준지도 학습을 통해 학습한 모델을 활용 하였다. 준지도 학습 과정에서는 각 주기마다 학습이 완료된 모델이 일정량의

데이터를 prediction한 결과를 기반으로 pseudo-labeling을 수행하였다. 그 후에 thresholding 및 sorting 과정을 거쳐 이러한 pseudo-label이 다시 학습 데이터에 추가되어 prediction 성능이 점차 향상되는 것을 볼 수 있었다. 지도 학습(supervised)을 위한 데이터 크기의 10%보다 적은 훈련 데이터를 초기에 이용하였음에도 최고 성능이 거의 대등한 결과를 보였다. 물론, 이 실험에서 사용된 데이터는 모든 경우를 대표한다고 볼 수 없다. 철로 데이터의 특성상, 이 연구에서 적용된 방법이 다른 특성을 가진 데이터에는 적용되지 못할 수 있다. 이 연구는 초기 데이터 크기의 영향을 보는데 중점을 두었고, 최종적으로 90% 이상의 정확도를 보여주었다. Contrastive Learning [14, 15]과 같은 기술을 도입한다면 추가적인 성능 향상이 가능할 수 있다. 또한 추후에는 제안된 모델을 철도 표면 데이터 이외의 다른 데이터에도 적용하는 노력이 필요하다.

#### References

- [1] H. Kim, M. Kim and Y. Lee. "Research Trend of the Remote Sensing Image Analysis Using Deep Learning," Korean Journal of Remote Sensing Vol.38, No.5 pp.819-834, 2022.  
DOI: <https://doi.org/10.7780/kjrs.2022.38.5.3.2>
- [2] W. Choi, et al, "Strawberry Growth Index Detection using Deep Learning," Journal of the Korea Academia-Industrial, Vol.24, pp. 558-565, 2023.  
DOI: <https://doi.org/10.5762/KAIS.2023.24.11.558>
- [3] D. Kim, et al, "Detection of Foreign Object in Cutting Surfaces using Artificial Intelligence in Plastic Processing," Journal of KIIT. , Vol. 20, No. 2, pp. 1-10, 2022.  
DOI: <https://doi.org/10.14801/jkiit.2022.20.2.1>
- [4] S. Kim, et al, "Examination of Aggregate Quality Using Image Processing Based on Deep-Learning," KIPS Transactions on Software and Data Engineering, Vol.11, No.6, pp. 255-266, 2022.  
DOI: <https://doi.org/10.3745/KTSDE.2022.11.6.255>
- [5] S. Ryu, et al, "Development of Deep Learning-based Automatic Camera Gimbal System for Drone Inspection of Transmission Lines," The Transactions of the Korean Institute of Electrical Engineers. Vol.70, No.1, pp.121-129, 2021.  
DOI: <https://doi.org/10.5370/KIEE.2021.70.1.121>
- [6] S. Teng, at al. , "Motion Planning for Autonomous Driving: The State of the Art and Future Perspectives," IEEE Transactions on Intelligent Vehicles, Vol.8, pp. 3692-3711, June 2023.  
DOI: <https://doi.org/10.1109/TIV.2023.3274536>

- [7] H. Kim, S. Lee, S. Han, "Railroad Surface Defect Segmentation Using a Modified Fully Convolutional Network," KSII Transactions on Internet & Information Systems, Vol. 14, pp. 4763-4775, 2020.  
DOI: <https://doi.org/10.3837/tiis.2020.12.008>.
- [8] S. Lee, and S. Han, "Detection Fastener Defect using Semi Supervised Learning and Transfer Learning," J. Internet Comput. Serv., Vol. 24, pp. 91-98, 2023.  
DOI: <http://dx.doi.org/10.14801/jkiit.2022.20.2>.
- [9] W.Chae, Y. Park, S. Kwan and J. Lee, "Thermal Analysis for Improvement of Heat Dissipation Performance of the Rail Anchoring Failure Detection Module," Journal of the Korean Institute of Electrical and Electronic Material Engineers, Vol.29, No.2, pp.125-130, 2016.  
DOI: <https://doi.org/10.4313/jkem.2016.29.2.125>
- [10] P. Chandran, J. Asber, F. Thiery, J. Odelius and M. Rantatalo "An investigation of railway fastener detection using image processing and augmented deep learning" Sustainability Vol.13, No.21, 2021  
DOI: <https://doi.org/10.3390/su132112051>
- [11] D. Zheng, L. Li, S. Zheng, X. Chai, S. Zhao, Q. Tong, J. Wang, L. Guo, "A Defect Detection Method for Rail Surface and Fasteners Based on Deep Convolutional Neural Network" Computational Intelligence and Neuroscience, Vol. 2021, pp.1-15, 2021.  
DOI: <https://doi.org/10.1155/2021/2565500>
- [12] X. Yang, et al, "A Survey on Deep Semi-Supervised Learning," IEEE Transactions on Knowledge and Data Engineering, Vol. 35, pp.8934-8954, 2023.  
DOI: <https://doi.org/10.1109/TKDE.2022.3220219>
- [13] K. He, X. Zhang, S. Ren and J. Sun. "Deep Residual Learning for Image Recognition." 2016 IEEE Conference on Computer Vision and Pattern Recognition pp.770-778, 2016.  
DOI: <https://doi.org/10.1109/CVPR.2016.90>
- [14] T. Chen, K. Simon, N. Mohammad and G. Hinton. "A Simple Framework for Contrastive Learning of Visual Representations." In International conference on machine learning, pp.1597-1607, 2020.  
DOI: <https://arxiv.org/abs/2002.05709>
- [15] X. Wang, G. Qi, "Contrastive Learning With Stronger Augmentations," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 45, pp..21271-21284, 2022.  
DOI: <https://doi.org/10.1109/TPAMI.2022.3203630>

## 저 자 소 개

### 한 석 민(정회원)



- 2001년 : 서울대학교 전기공학부(공학사)
- 2003년 : 서울대학교 전기컴퓨터공학부(공학석사)
- 2008년 : 서울대학교 전기컴퓨터공학부(공학박사)
- 2017 ~ 현재 : 한국교통대학교 데이터사이언스전공 교수

※ 이 연구는 한국연구재단(National Research Foundation of Korea)의 지원으로 수행되었음. (NRF-2021R1F1A1049467).