터널 내 돌발상황 오탐지 영상의 반복 학습을 통한 딥러닝 추론 성능의 자가 성장 효과

이규범[1] · 신휴성[2*]

[1]비회원, 한국건설기술연구원 미래융합연구본부 학생연구원,
과학기술연합대학원대학교(UST) 스마트도시건설융합 석박통합과정
[2]정회원, 한국건설기술연구원 미래융합연구본부 연구위원

Effect on self-enhancement of deep-learning inference by repeated training of false detection cases in tunnel accident image detection

Kyu Beom Lee[1] · Hyu Soung Shin[2*]

[1]Student Researcher, Department of Future Technology and Convergence Research,
Korea Institute of Civil Engineering and Building Technology
Ph.D. Student, Smart City and University of Science & Technology
[2]Research Fellow, Department of Future Technology and Convergence Research,
Korea Institute of Civil Engineering and Building Technology

***Corresponding Author :** Hyu Soung Shin, hyushin@kict.re.kr

Abstract

Most of deep learning model training was proceeded by supervised learning, which is to train labeling data composed by inputs and corresponding outputs. Labeling data was directly generated manually, so labeling accuracy of data is relatively high. However, it requires heavy efforts in securing data because of cost and time. Additionally, the main goal of supervised learning is to improve detection performance for 'True Positive' data but not to reduce occurrence of 'False Positive' data. In this paper, the occurrence of unpredictable 'False Positive' appears by trained modes with labeling data and 'True Positive' data in monitoring of deep learning-based CCTV accident detection system, which is under operation at a tunnel monitoring center. Those types of 'False Positive' to 'fire' or 'person' objects were frequently taking place for lights of working vehicle, reflecting sunlight at tunnel entrance, long black feature which occurs to the part of lane or car, etc. To solve this problem, a deep learning model was developed by simultaneously training the 'False Positive' data generated in the field and the labeling data. As a result, in comparison with the model that was trained only by the existing labeling data, the re-inference performance with respect to the labeling data was improved. In addition, re-inference of the 'False Positive' data shows that the number of 'False Positive' for the persons were more reduced in case of training model including many 'False Positive' data. By training of

OPEN ACCESS

Journal of Korean Tunnelling and Underground Space Association
21(3)419-432(2019)
https://doi.org/10.9711/KTAJ.2019.21.3.419

eISSN: 2287-4747
pISSN: 2233-8292

Received April 1, 2019
Revised April 23, 2019
Accepted May 2, 2019

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright ⓒ2019, Korean Tunnelling and Underground Space Association

the ‘False Positive’ data, the capability of field application of the deep learning model was improved automatically.

Keywords: False Positive data, Labeling data, Deep learning-based CCTV incident detection system, Deep learning model training including False Positive data

초 록

대부분 딥러닝 모델의 학습은 입력값과 입력값에 따른 출력값이 포함된 레이블링 데이터(labeling data)를 학습하는 지도 학습(supervised learning)으로 진행된다. 레이블링 데이터는 인간이 직접 제작하므로 데이터의 정확도가 높다는 장점이 있지만 비용과 시간의 문제로 인해 데이터의 확보에 많은 노력이 소요된다. 그리고 지도 학습의 목표는 정탐지 데이터(true positive data)의 인식 성능 향상에 초점이 맞추어져 있으며, 오탐지 데이터(false positive data)의 발생에 대한 대처는 미흡한 실정이다. 본 논문은 터널 관제센터에 투입된 딥러닝 모델 기반 영상유고 시스템의 모니터링을 통해 정탐지와 레이블링 데이터의 학습으로 예측하기 힘든 오탐지의 발생을 확인하였다. 오탐지의 유형은 작업차량의 경광등, 터널 입구부에서 반사되는 햇빛, 차선과 차량의 일부에서 발생하는 길쭉한 검은 음영 등이 화재와 보행자로 오탐지되고 있었다. 이러한 문제를 해결하기 위해 현장에서 발생한 오탐지 데이터와 레이블링 데이터를 동시에 학습하여 딥러닝 모델을 개발하였으며, 그 결과 기존 레이블링 데이터만 학습한 모델과 비교하면 레이블링 데이터에 대한 재추론 성능이 향상됨을 알 수 있었다. 그리고 오탐지 데이터에 대한 재추론을 한 결과 오탐지 데이터를 많이 포함하여 학습한 모델일 경우 보행자의 오탐지 개수가 훨씬 줄었으며, 오탐지 데이터의 학습을 통해 딥러닝 모델의 현장 적용성을 향상시킬 수 있었다.

주요어: 오탐지 데이터, 레이블링 데이터, 딥러닝 기반 터널 CCTV 영상유고 시스템, 오탐지 데이터 포함 딥러닝 모델 학습

1. 서 론

지난 2018년 11월, 딥러닝 기반 터널 CCTV 영상유고 시스템의 기본 개념 및 시스템 구성도가 완성되었으며(Lee et al., 2018), 실제 시스템을 터널 관제센터 현장에 설치 및 모니터링을 진행하고 있다. 딥러닝 기반 터널 CCTV 영상유고 시스템은 우선 일정 프레임 간격 주기마다 터널 내 CCTV 정지영상을 입력값으로 하여 딥러닝 모델이 정지영상 내에 존재하는 차량, 화재, 보행자 객체를 구분하여 위치와 크기를 직사각형의 경계박스(bounding box)로 나타낸다. 이 단계에서 화재와 보행자가 발생하는 유고상황을 검출할 수 있다. 그 다음, 정지영상에서 인식된 차량 객체를 바탕으로 객체 추적을 하며, 객체 인식과정보다 긴 일정 프레임 주기마다 정차 및 역주행을 판단한다. 객체 추적 과정은 이전 프레임 주기와 현재 프레임 주기에서 검출된 차량 객체들을 서로 비교하여 겹치는 면적 비율이 일정 값 이상인 객체 쌍이 존재할 때, 객체 번호를 부여하여 동일한 객체로 인식시킨다. 정차 및 역주행의 판단은 이전 프레임 주기와 현재 프레임 주기에서 동일한 객체 번호를 가진 차량 객체 쌍에 대해서 시행하는데, 정차는 겹치는 면적 비율이 0.9 이상일 때, 역주행은 터널 진행방향 반대로 주행하면서 겹치는 선 길이 비율이 0.75 미만일 때 각각 정차와 역주행으로 판단한다. 이 내용은 딥러닝 기반 터널 CCTV 영상유고 시스템의 핵심과정이며, 현장에서 설치할 전 주기 시스템은 9채널의 카메라를 동시에 감시할 수 있도록 CCTV 영상을

입력받아 핵심 과정에서 활용할 수 있도록 영상을 전처리하는 전처리 과정과 핵심과정, 터널 관제센터에서 직접 유고상황을 확인할 수 있는 유고상황 표출 과정으로 구성되어 있다. 딥러닝 학습 과정은 현장에서 학습하지 않으며, 다른 장소에서 학습한 다음 2주~1개월마다 비정기적으로 전주기 시스템의 딥러닝 모델을 갱신한다.

한 편, 딥러닝을 포함한 기계학습(machine learning)에서 지도학습(supervised learning)은 입력값과 출력값이 포함된 레이블링 데이터(labeling data)를 학습하며, 데이터가 충분할 경우 다양하게 활용될 수 있어 가장 활발히 연구되고 있는 분야이다(LeCun et al., 2015). 지도학습의 기본과정은 학습용 데이터와 추론용 데이터를 나누며, 학습용 데이터를 바탕으로 딥러닝 모델을 학습한다. 그리고 딥러닝 모델의 검증은 학습용 데이터의 재추론 인식 성능의 확인과 추론용 데이터의 인식 성능을 확인하는 것으로 실제 현장에 적용하였을 때 딥러닝 기반 시스템이 제대로 목표를 인식할 수 있는지 예측할 수 있다. 이 과정에서 딥러닝 모델의 인식 성능을 평가할 때 Table 1을 바탕으로 인식된 출력값의 유형을 4가지로 분류한 다음 평가한다(Davis and Goadrich, 2006).

Table 1. 4 kinds of prediction result for testing data

Category	Predicted as Positive	Predicted as Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

Table 1에서 정답(positive)은 인식 대상 객체, 오답(negative)은 정답이 아닌 객체를 의미하며, 정답을 제대로 인식했을 경우에는 정탐지(true positive), 정답으로 인식했는데 오답일 경우 오탐지(false positive)에 해당된다. 오답을 제대로 인식하면 오답 탐지(true negative)이고, 오답으로 예측하였는데 실제로 정답이면 미탐지(false negative)에 해당된다. 대부분의 딥러닝 기반 시스템은 정탐지를 최대화하고 오탐지와 미탐지를 줄이는 방향으로 목표를 설정하므로 정탐지의 양적인 인식률을 평가하는 재현률(recall), 정탐지의 질적인 인식률을 평가하는 정밀도(precision)를 이용해 평균 정밀도(average precision) 값으로 평가하며 재현률, 정밀도와 평균 정밀도는 각각 식 (1), 식 (2), 식 (3)과 같이 나타낼 수 있다(Zhu, 2004).

$$\text{Recall} = \frac{TP}{TP + FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Average Precision} = \int Precision(Recall)\,dRecall \tag{3}$$

평균 정밀도는 정탐지의 양과 질적인 측면을 모두 고려할 수 있는 지표이고, 0~1의 범위를 가진 스칼라값으로 표현가능하므로 레이블링 데이터에 대한 딥러닝 모델의 추론 정확도를 나타내는 표준적인 지표로 사용하고 있다.

딥러닝 기반 터널 CCTV 영상유고 시스템에서 딥러닝 모델의 학습 및 추론은 유고상황이 포함된 영상들을 정지영상으로 분해한 다음, 레이블링한 데이터로 추론/학습용 데이터로 나누지 않고 레이블링 데이터를 전부 활용하여 학습을 진행한다. 대중에 공개된 Pascal VOC 데이터셋(Everingham et al., 2010)이나 COCO 데이터셋(Lin et al., 2014)은 연속성이 없는 정지영상 수 만장 속에 포함된 다양한 객체들이 레이블링되어 있다. 그렇기 때문에 학습용 데이터와 검증용 데이터를 나눈 다음 검증용 데이터를 추론해 보면 딥러닝 모델이 학습용 데이터에 편향적으로 학습되었는지 확인할 수 있다. 그러나 영상 기반 레이블링 데이터는 수 많은 정지영상들이 연속적으로 이어지므로, 전후 정지영상들이 서로 유사한 영상환경을 보인다. 이러한 이유로 인해 검증용 데이터를 추론하여 학습결과를 검증한다 하여도 딥러닝 모델이 편향적으로 학습되었는지 알 수 없으므로 딥러닝 시스템의 현장 투입 및 모니터링을 통해 딥러닝 모델의 객체인식 능력을 검증해야 한다.

또한, 동영상을 대상으로 객체를 인식하고 오탐지가 다수 발생할 경우 이를 저감시키기 위해서는 기존 학습자료에 더하여 오탐지를 저감하기 위한 추가 학습자료의 준비가 필요하며, 일반적으로 학습자료는 객체의 레이블링 작업이 선행되어야 하므로 많은 시간과 노력이 필요하다. 본 논문에서는 동영상 자료의 실시간 추론에서 발생하는 오탐지 결과를 별도의 수동 레이블링 작업 없이 학습자료에 추가 반영하여 목표 객체의 인식 성능을 자동으로 저감시키는 방안을 제안한다. 이는 현장에 설치된 시스템의 수정이나 별도의 학습자료 준비 노력 없이도, 시간이 지남에 따라 시스템 스스로 오탐지가 줄고, 객체 인식 성능이 스스로 향상되는 효과를 기대할 수 있다.

딥러닝 모델의 객체인식 능력은 학습 데이터가 가장 많고 객체의 크기가 일정한 차량 객체가 가장 정확한 반면, 보행자는 차량보다 작고 길쭉하게 보여 객체인식이 어려우며, 화재는 터널 내 화재가 발생한 영상의 확보가 어려워 학습 데이터의 수가 적어 오탐지가 발생할 가능성이 크다(Shin et al., 2017; Lee et al., 2018). 따라서 본 논문은 딥러닝 기반 터널 CCTV 영상유고 시스템을 완성하여 직접 투입함으로써 딥러닝 모델의 학습 외 영상에 대한 화재, 보행자의 객체인식 능력을 직접 검증하였으며, 화재와 보행자에 대한 다수의 오탐지가 발생하는 것을 보여줄 것이다. 이러한 문제를 해결하기 위해 오탐지 데이터를 오답으로 정의한 다음, 학습데이터에 포함하여 학습하는 방법을 제안한다. 이 방법에 대한 검증은 레이블링 데이터와 오탐지 데이터를 따로 나누어서 진행하는데, 레이블링 데이터에 대한 재추론을 진행하여 기존 학습데이터의 인식성능에 차이가 존재하는지 확인한다. 그리고 보행자 오탐지 데이터를 직접 추론하여 보행자 오탐지 인식개수를 확인하여 오탐지가 저감되었는지 비교 평가할 것이다.

2. 딥러닝 기반 터널 CCTV 영상유고 시스템 모니터링 현황

딥러닝 기반 터널 CCTV 영상유고 시스템은 지난 2018년 11월 중순 A 고속도로 터널 관제 센터에 설치되어 운용을 시작하였다. 해당 터널 관제 센터는 폐쇄형 네트워크 기반 CCTV 카메라를 바탕으로 터널 및 도로를 감시하고 있으며, 유고상황이 발생할 경우 고속도로를 주행하고 있는 운전자에게 알리는 동시에 작업인력이 유고상황

이 발생한 현장으로 투입된다. 대부분의 터널 관제 센터는 국토교통부의 가이드라인에 의해 유고상황 감지 시스템을 설치하였다(MOLIT, 2016). 그러나 기존 유고상황 감지 시스템은 오탐지 발생의 빈번함으로 인해 낮은 신뢰성을 보이며 현장의 불필요한 확인작업과 사후조치 노력이 소요되면서(Kim, 2016; National Committee for Land and Transport, 2016), 유고상황 감지 시스템을 구비하였어도 운용하지 않고 육안으로 유고상황을 감시하고 있는 현장이 다수이다. 이 때문에 본 연구에서 개발된 딥러닝 기반 터널 CCTV 영상유고 시스템을 모니터링할 때, 정탐지가 되는지 확인하는 것이 가장 중요한 목표이지만 동시에 오탐지로 인식되는 경우를 줄이는 것 또한 중요하다. 딥러닝 기반 터널 CCTV 영상유고 시스템을 위해 레이블링 데이터로 딥러닝 모델을 학습하여 적용하고 모니터링을 시작하였다. 55일 동안 시스템을 모니터링한 결과, 차량 객체는 대체로 오탐지 객체 없이 상당히 정확하게 인식하였지만, 화재 객체와 보행자 객체는 꾸준히 오탐지된 객체가 발생하고 있었다.

Fig. 1은 딥러닝 기반 터널 CCTV 영상유고 시스템을 설치한 뒤 55일 동안 모니터링하여 9채널의 카메라에서 발생한 화재와 보행자의 오탐지 발생 현황을 그래프로 나타내었다. 보행자는 최대 1000개, 하루에 200~400개의 오탐지가 발생하고 있었으며, 화재는 최대 120개, 하루에 0~40개의 오탐지가 발생하였다. 화재보다 보행자의 오탐지 개수가 훨씬 빈번하게 발생하고 있는데, 보행자의 객체의 인식이 훨씬 까다롭기 때문이다.

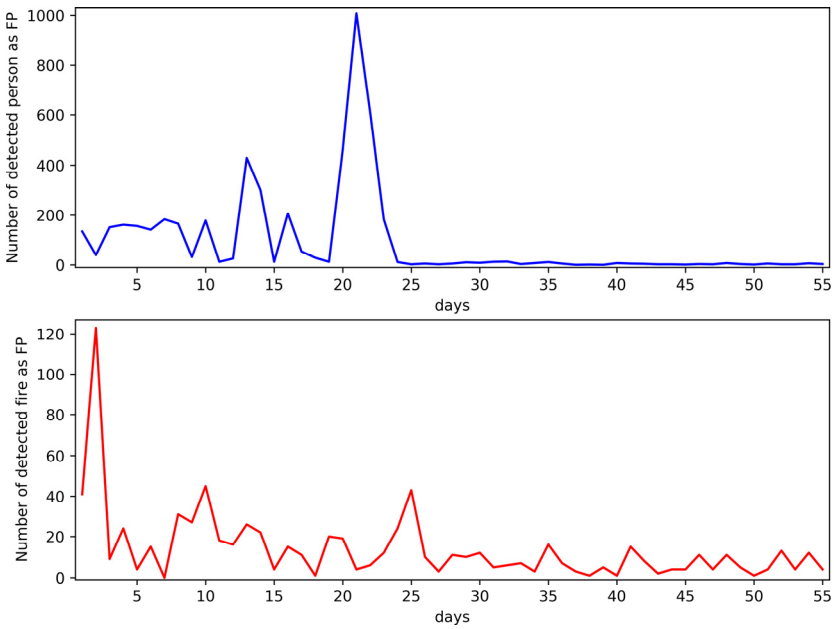


Fig. 1. 'False Positive' report for person and fire objects during 55 days after installing deep learning-based CCTV accident detection system

보행자 객체는 대체로 화재나 차량 객체보다 작은 크기를 가지고 있으며, 세로로 길쭉한 형상을 보인다. 게다가 차량 충돌사고가 발생할 경우 운전자는 차량의 상태를 확인하기 위해 차량 밖으로 나오게 되는데, 차량 주변에 붙

어있는 모습이 CCTV에서 촬영된다. 이러한 현상을 이동객체간 겹침현상이라고 하며(Kim et al., 2012), 보행자 객체와 차량 객체의 겹침현상으로 인해 일반적인 딥러닝 모델의 학습법으로는 딥러닝 모델이 차량의 그림자나 차선과 보행자를 구분하기 쉽지 않다. 터널 내에서 화재가 발생할 수 있는 요소는 주로 차량에 의한 화재를 생각할 수 있다. 즉, 화재가 발생할 때 CCTV에서 촬영된 화재 객체의 크기는 최소 차량의 크기로 볼 수 있으며, 거대하게 일어나면 CCTV 화면 전체 영역을 차지할 정도로 크기 때문에 차량과 구분된다. 그러나 작업차량의 경광등이나 터널 입구부에서 반사되는 햇빛으로 인해 화재로 오인될 수 있으며, Fig. 2와 같은 오탐지 현상을 확인할 수 있었다.

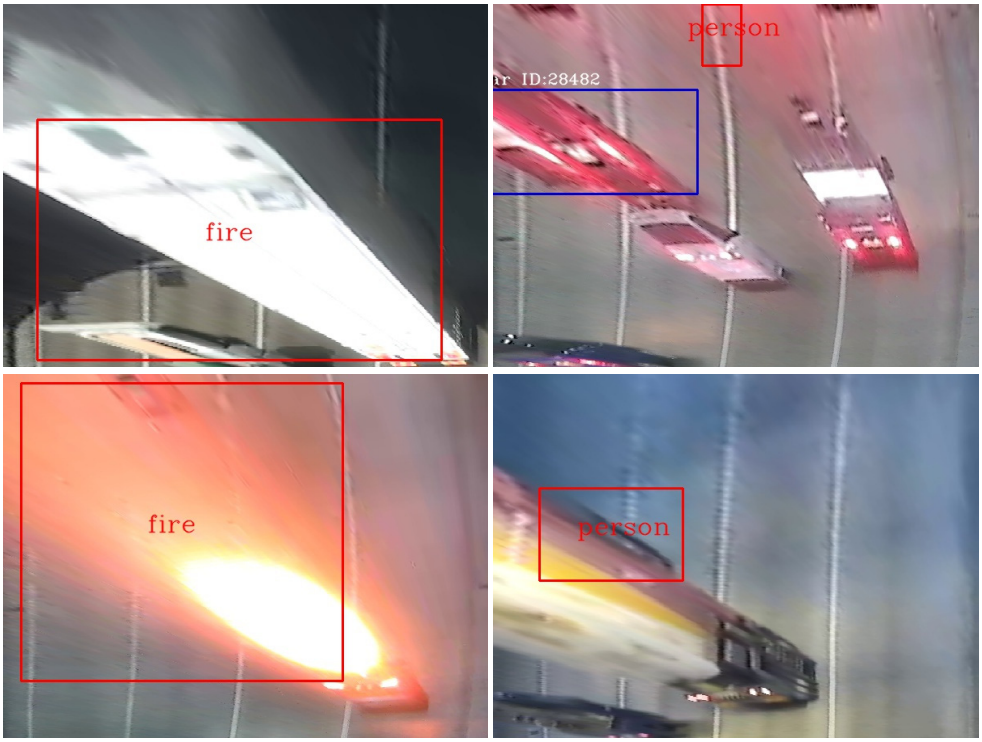


Fig. 2. 'False Positive' samples occurred by deep learning-based accident detection system in tunnel monitoring

Fig. 2는 딥러닝 기반 터널 CCTV 영상유고 시스템에 발생한 화재 및 보행자가 오탐지된 화면이다. 화재 객체의 경우, 작업차량의 경광등이 터널 내부에서 비치는 형상과 터널 입구부에서 대형 트럭의 뒷면이 햇빛에 반사되어 밝게 빛나는 형상이 화재로 오탐지되고 있다. 그리고 보행자는 차량 객체의 검은 음영으로 되어 있는 일부 형상과 하얀 차선을 보행자로 인식하고 있었다. 이 문제는 딥러닝 모델의 학습에 사용했던 레이블링 데이터와 비교하였을 때 그 원인을 판단할 수 있다.

Fig. 3은 화재와 보행자에 대하여 레이블링 데이터와 오탐지 데이터 샘플을 비교하였다. 각 정지영상의 전체적인 그림을 보면 레이블링 데이터와 오탐지 데이터 사이에 차이점이 존재한다는 것을 알 수 있다. 그러나 직사각형

의 경계박스만 놓고 비교하면 유사점이 존재하며, 화재의 경우 작업차량의 경광등과 실제 화재 형상이 상당히 유사하다. 보행자 객체도 마찬가지로 차량의 검은 음영과 실제 보행자만 놓고 보면 정확히 유사하진 않지만 딥러닝 모델이 충분히 보행자로 착각할 수 있을 만큼 유사한 편이었다. 따라서 딥러닝 모델에 대한 기존 레이블링 데이터의 학습만으로는 학습 외 영상에서 발생하는 오탐지 대처가 쉽지 않다.

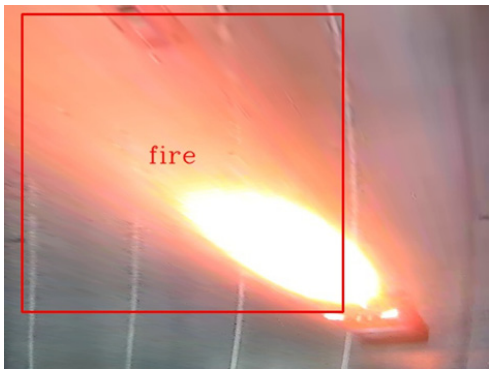
Category	True Positive	False Positive
Fire		
Person		

Fig. 3. Similarity between 'True Positive' and 'False Positive' samples for fire and person objects

3. 딥러닝 모델의 오탐지 포함 학습방법 및 평가

딥러닝 기반 CCTV 영상유고 시스템을 모니터링하면서 발생한 오탐지 데이터는 레이블링된 데이터의 학습으로 얻어진 딥러닝 모델의 한계를 보여주었으며, 낮은 시스템의 신뢰성을 보인다. 본 논문은 이러한 문제를 해결하기 위해 오탐지 데이터를 포함하여 딥러닝 모델을 학습하는 방법을 제안한다. 그리고 오탐지 데이터를 포함하여 학습하는 방법의 유효성을 검증하기 위해 레이블링 데이터에 대한 재추론 성능과 오탐지 데이터에 대한 인식 성능을 측정하여 레이블링 데이터만 학습된 딥러닝 모델과 오탐지 데이터를 포함한 딥러닝 모델을 비교할 것이다.

3.1 딥러닝 모델의 오탐지 포함 학습방법

본 논문에서 제안하는 딥러닝 모델의 오탐지 포함 학습방법은 Fig. 4와 같은 절차로 진행되며, 터널 관제센터에 설치되어 있는 딥러닝 기반 CCTV 영상유고 시스템의 모니터링 과정에서 확인된 오탐지 데이터의 확보가 전제된다.

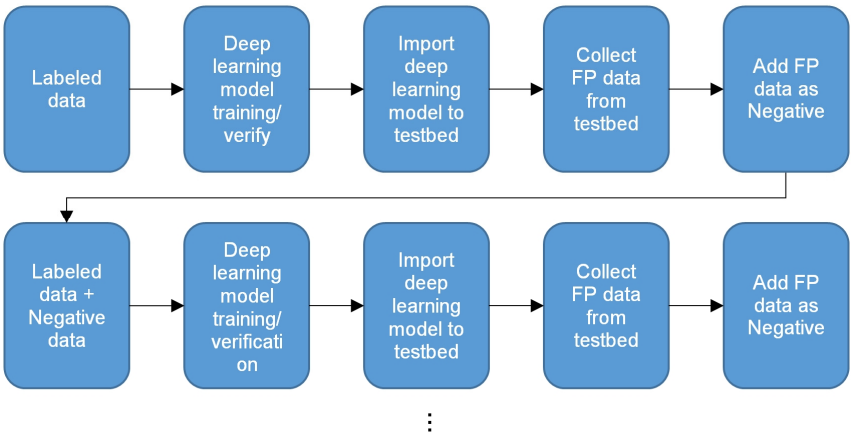


Fig. 4. Deep learning training process including False Positive data

Fig. 4에서와 같이, 딥러닝 모델의 오탐지 포함 학습방법은 먼저 레이블링 데이터로 딥러닝 모델을 학습시킨 다음, 검증을 통해 딥러닝 모델의 객체 인식능력을 확인하여 터널 관제센터 현장에 설치된 시스템에 딥러닝 모델을 적용한다. 그리고 비정기적으로 터널 관제센터 현장을 방문하여 지금까지 탐지된 유고상황 데이터를 수집하여 분석한 뒤, 오탐지 데이터와 정탐지 데이터를 분류한다. 오탐지 데이터에서 학습에 사용할 객체는 현장에서 상대적으로 정확하게 인식된 차량을 제외한 보행자, 화재 2가지이며, 딥러닝 모델의 학습에 사용할 클래스를 각각 화재 오탐지(false fire)과 보행자 오탐지(false person)으로 정의하여 오탐지 데이터셋을 제작한다. 제작된 오탐지 데이터셋은 레이블링 데이터와 같이 딥러닝 모델의 학습에 사용되며, 다시 딥러닝 모델을 검증하여 이상이 없으면 터널 관제센터 현장에 적용중인 딥러닝 모델을 교체시킨다. 이러한 과정을 반복하여 오탐지 데이터를 추가하여 학습시킬수록 딥러닝 모델의 오탐지 인식이 시스템 수정없이 자동으로 저감되며, 레이블링 데이터의 재추론 또한 식 (2)에서 오탐지의 개수가 감소하므로 추론 성능이 자동 향상되는 효과를 기대할 수 있다.

3.2 딥러닝 모델 성능 평가

본 논문에서 제안한 딥러닝 모델의 오탐지 포함 학습방법은 레이블링 데이터의 추가 없이 현장에서 발생된 오탐지 데이터를 추가하여 학습하므로 레이블링 데이터의 제작에 소요되는 시간과 비용을 절약할 수 있다. 이 방법을 검증하기 위하여 기존 레이블링 데이터로 학습한 모델과 오탐지 데이터를 추가한 다음, 학습한 딥러닝 모델을

비교하여 딥러닝 모델의 객체인식 성능을 평가한다. 객체인식 성능의 평가는 레이블링 데이터의 재추론 성능과 오탐지 데이터 인식 저감 성능을 비교하여 딥러닝 모델의 오탐지 포함 학습방법의 실효성을 평가할 것이다.

3.2.1 레이블링 및 오탐지 빅데이터 현황

딥러닝 모델의 학습에 사용할 레이블링 및 오탐지 데이터의 현황은 Table 2와 같다. AA~DD 데이터셋은 전부 수동으로 레이블링된 데이터이며, 총 70,914장의 정지영상에서 441,670개의 차량과 49721개의 보행자, 857개의 화재 객체가 존재한다. 딥러닝 모델의 학습대상 객체는 차량, 보행자와 화재이고, 도로 터널에서 촬영된 영상이므로 레이블링 데이터에서 차량 객체의 개수가 가장 많다. 그 다음 보행자와 화재 순서로 객체가 많은데, 화재의 객체 수가 적은 것은 일반적으로 도로 터널 내에서 화재가 발생하는 경우가 드물며 화재와 관련된 유고영상의 확보가 쉽지 않기 때문이다. 이러한 이유로 인해 화재 객체의 다양성이 떨어지는 문제점을 안고 딥러닝 모델을 학습한다. Table 2에서 오탐지 데이터는 현장적용 단계에서 자동으로 도출된 오탐 데이터를 의미하며, FP data 1은 화재 오탐지의 개수와 보행자 오탐지의 개수가 1 : 2의 비율로 구성되고 FP data 2는 보행자 오탐지 데이터가 지배적이다. FP data 1은 화재 오탐지와 보행자 오탐지를 동시에 저감되는 효과를 검토하기 위해 구성된 데이터이고, FP data 2는 보행자 오탐지의 저감 효과 검토를 주목적으로 하여 구성된 데이터이다. 이렇게 구성된 것은 레이블링 데이터에서 화재 객체수가 적어서 딥러닝 모델을 적용한 현장에서 발생하는 화재 오탐지의 유형이 단순하므로 화재 오탐지의 저감이 용이하기 때문이다. 또한 레이블링 데이터에서 화재 오탐지를 터널 내 경광등, 차량의 후미등, 작업차량의 경광등 등의 불빛을 화재 오탐지로 정의하였으므로 FP data 1에서 화재 오탐지로 정의된 객체 데이터 691개를 포함하고 FP data 2에서는 22개만 포함되었다. 반면 보행자 오탐지의 경우, 규모에 따른 오탐지 저감효과를 확인하기 위하여 FP data 1에서는 1357개, FP data 2에서는 7999개의 보행자 오탐지 데이터를 포함한다.

Table 2. Composition of labeling data and 'False Positive' data

Data category	Number of images	Car	Person	Fire	False_fire	False_person
AA	38831	219292	26612	0	181984	0
BB	873	0	0	857	713	0
CC	7471	52296	9042	0	1751	0
DD	23739	175138	11487	0	0	0
FP data 1	2041	0	0	0	691	1357
FP data 2	8007	0	0	0	22	7999

3.2.2 딥러닝 모델에 따른 데이터 구성 및 평가방법

상기 Table 2에서 서술한 레이블링 데이터와 오탐지 데이터를 바탕으로 학습에 사용될 데이터셋을 구성하였으며, 목적에 따라 Table 3과 같이 3개의 유형으로 나누었다. Table 3에서 Model 1은 레이블링 데이터만 학습하며,

오탐지 데이터를 추가하여 학습한 Model 2, Model 3을 비교하기 위해 학습된 모델이다. Model 2는 FP data 1 오탐지 데이터셋을 추가하였으며, 화재 오탐지와 보행자 오탐지를 한번에 저감하는 효과를 검토하기 위한 목적을 가진다. 마지막으로 Model 3은 Model 2보다 확실한 보행자 오탐지의 저감 효과 검토를 위해 FP data 1과 FP data 2를 추가하여 학습을 진행한다.

Table 3. Training dataset models

Model name	Including dataset
Model 1	Labeled dataset
Model 2	Labeled dataset + FP data 1
Model 3	Labeled dataset + FP data 1+ FP data 2

이러한 모델 조건을 기반으로 하여 Table 4와 같은 딥러닝 모델 학습 조건으로 학습을 진행한다. 딥러닝 모델은 딥러닝 객체인식 분야에서 널리 쓰이고 있는 Faster R-CNN (Regional Convolutional Neural Network)을 사용한다(Ren et al., 2015). 그래픽카드는 NVIDIA GTX 1070 8 GB를 사용하며, 리눅스 환경에서 딥러닝 모델의 학습 및 추론을 진행하였다. 각 모델마다 10 epoch씩 학습하였다.

Table 4. Training condition of a deep learning algorithm

Deep learning model	Faster R-CNN
GPU	NVIDIA GTX 1070 8 GB
OS	Linuxmint 18.3
Epoch	10

딥러닝 모델의 평가는 레이블링 데이터에 대한 재추론, 보행자 오탐지 데이터의 인식 개수, 딥러닝 모델 현장적용 후 화재 오탐지 데이터의 인식현황을 확인한다. 레이블링 데이터를 재추론하여 기존 레이블링 데이터만 학습한 Model 1과 비교하여 차량, 화재, 보행자 객체인식 능력이 향상되었는지 비교평가를 진행한다. 보행자 오탐지 데이터의 경우, 딥러닝 모델의 학습에 사용한 보행자 오탐지 개수에 따른 보행자 오탐지의 저감성능을 평가하며, Model 2에 사용한 보행자 오탐지 데이터와 Model 3에 사용한 보행자 오탐지 데이터를 따로 평가한다. 화재 오탐지의 경우, Model 2를 현장에 적용한 뒤 모니터링하여 화재 오탐지가 저감되었는지 직접 확인한다.

3.2.3 레이블링 데이터에 대한 재추론 평가

각 딥러닝 모델에 따른 레이블링 데이터에 대한 재추론 결과는 Fig. 5와 같다.

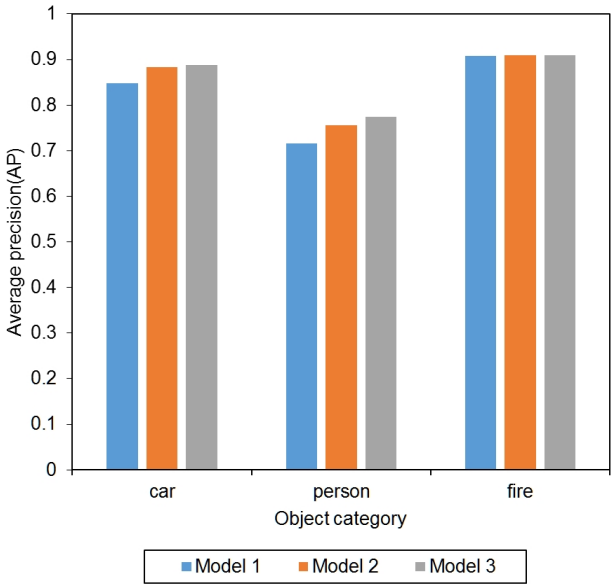


Fig. 5. Test results of deep learning models

Fig. 5에서 차량과 보행자의 경우, Model 1과 비교해 Model 2와 Model 3의 객체인식 성능이 더 뛰어난 결과를 보여준다. 이러한 결과는 오탐지 데이터를 포함하여 학습하면 기존 레이블링 데이터의 객체인식 능력을 저해하지 않으며, 오히려 객체인식 능력을 향상시킬 수 있었다. 차량은 Model 2와 Model 3이 비슷한 객체인식 능력을 보이지만, 보행자의 경우 오탐지 데이터를 더 추가하여 학습한 Model 3이 Model 2보다 더 나은 보행자 객체인식 능력을 보였다. 다시 말하면 보행자 객체 클래스에서 보행자 오탐지 데이터를 추가하여 학습할수록 보행자에 대한 오탐지의 저감능력이 향상된다고 할 수 있다. 한 편, 화재의 객체인식 능력은 0.9의 AP값으로 Model 1, Model 2와 Model 3이 같은 값을 가지는데, 레이블링 데이터에서 화재의 객체 수가 적고 화재에 대한 학습 난이도가 낮기 때문에 3개의 모델이 모두 주어진 데이터에서 화재 객체에 대하여 완전하게 학습되었다.

3.2.4 오탐지 데이터의 인식개수 평가

레이블링 데이터를 재추론함으로써 차량 객체인식 성능의 향상, 보행자 오탐지의 저감에 따른 보행자 객체인식 성능의 향상, 일관성 있게 유지된 화재의 객체인식 성능을 확인할 수 있었다. 그러나 보행자 오탐지와 화재 오탐지의 저감 성능을 직접적으로 확인하기 위해서는 오탐지 데이터를 직접 추론하는 작업과 현장에 오탐지를 포함하여 적용한 딥러닝 모델을 모니터링하는 작업이 필요하다.

딥러닝 모델에 대한 보행자 오탐지 저감성능에 대한 평가는 Table 5와 같이 오탐지 데이터를 포함한 Model 2와 Model 3을 대상으로 하였으며, 각각 Model 2에서 학습된 보행자 오탐지 데이터 1357개와 Model 3에서 사용된 보행자 오탐지 데이터 9358개에 대하여 재추론하였다.

Table 5. The number of 'False Positives' resulted in site from trained models with 'False Positive' data

Target model	Number of detected persons as FP	
	1357 Negative examples	9358 Negative examples
Model 2	329	3270
Model 3	34	466

Table 5의 보행자 오탐지 데이터에 대한 재추론 결과는 Model 3가 훨씬 우수한 오탐지 저감성능을 보였다. 먼저 1357개의 보행자 오탐지에 대하여 추론한 경우, Model 2의 329개와 비교하여 Model 3은 34개만 오탐지가 발생하였다. 그리고 9358개의 보행자 오탐지를 추론한 결과, Model 2는 3270개의 보행자 오탐지를 발생시켰으며, Model 3은 466개의 보행자 오탐지만 발생하였다. 종합적으로 판단하면 Model 2는 추론 대상 보행자 오탐지 데이터의 24%, 35%의 비율로 보행자 오탐지를 발생시켰지만, Model 3은 2.5%, 5%의 비율만 보행자 오탐지를 발생시켰다. 따라서 Model 2보다 많은 보행자 오탐지 데이터를 추가하여 학습한 Model 3의 보행자 오탐지의 저감성능이 뛰어나다고 할 수 있다.

화재 오탐지의 저감성능 확인을 위한 딥러닝 모델 현장 적용에 대한 모니터링 현황은 Fig. 6과 같다. Fig. 6은 Fig. 1에서 화재에 대한 모니터링 현황과 이어지며, 딥러닝 모델의 현장 설치 후 55일까지 Model 1을 사용하고 56일부터 Model 2를 사용하였다. 그 결과, Model 1에서는 최대 120개, 보통 0~40개의 화재 오탐지가 발생되었지만, Model 2를 적용한 시점 이후로 0~3개의 화재 오탐지만 발생되었다. Model 2에서 사용된 화재 오탐지의 데이터 개수는 FP data 1에서 691개인데, 이 데이터만으로도 확실하게 화재 오탐지를 저감시킬 수 있었다. 그러나 화재 레이블링 데이터의 개수는 857개로 많지 않은 편이므로 보행자 오탐지와 비교하면 화재 오탐지의 저감이 훨씬 용이한 것으로 판단할 수 있다.

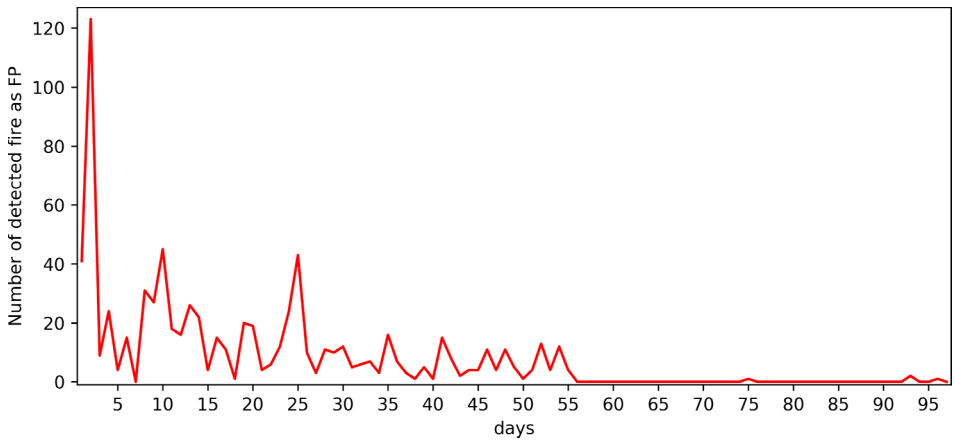


Fig. 6. 'False Positive' report for fire objects during 97 days after installing deep learning-based CCTV accident detection system

4. 결 론

본 논문은 딥러닝 기반 터널 CCTV 영상유고 시스템을 터널 관제센터에 현장 적용을 한 다음 발생한 보행자, 화재 오탐지 데이터를 학습데이터로서 활용하여 각각 오탐지 항목에 대한 저감성능을 보였으며, 다음과 같은 결론을 얻을 수 있었다.

1. 기존 레이블링 데이터에 화재, 보행자 오탐지 데이터를 추가하여 학습해도 딥러닝 객체인식 모델의 목표인 차량, 보행자, 화재의 객체인식 성능에 영향을 끼치지 않으며, 보행자의 경우 오탐지된 보행자의 개수가 감소하여 오히려 객체인식 성능이 향상되었다. 한편, 화재 객체는 레이블링 데이터의 수가 적어 학습이 용이하므로 객체인식 수준은 기존 레이블링 데이터만 학습한 모델과 비교해도 거의 동일하게 높은 수준을 유지하고 있다.
2. 보행자 오탐지 데이터를 학습 데이터로 추가하여 딥러닝 모델을 학습할 때, 보행자 오탐지 데이터의 규모가 클수록 보행자 오탐지의 저감성능이 훨씬 우수하였다. 대부분의 보행자 객체는 정지영상에서 작은 크기와 다양한 형태로 존재하므로 보행자 오탐지 데이터수를 보다 비중있게 학습하는 것이 딥러닝 모델의 보행자 객체인식에 도움을 줄 수 있다.
3. 화재 오탐지 데이터를 추가하여 학습한 딥러닝 모델을 터널 관제센터 현장에 직접 적용하여 모니터링한 결과, 9채널의 카메라에서 오탐지가 거의 발생하지 않았다. 이러한 결과는 레이블링 데이터만 학습하여 현장에 적용한 결과와 비교하면 훨씬 향상된 화재 오탐지의 저감능력을 보였다. 그러나 딥러닝 모델의 학습에 사용된 화재 레이블링 데이터의 개수가 적기 때문에 화재 오탐지의 저감이 용이하다. 따라서 더 많은 화재 레이블링 데이터를 확보하여 딥러닝 모델을 학습하는 과정이 반드시 필요하다.
4. 딥러닝 모델의 오탐지 포함 학습 방법은 기존 레이블링 데이터의 학습을 저해하지 않고 오히려 객체인식 성능을 향상시킬 수 있으며, 딥러닝 모델 적용 현장에서 발생하는 오탐지를 저감할 수 있다. 이 방법은 본 논문에서 적용된 딥러닝 기반 터널 CCTV 영상유고 시스템뿐만 아니라 다른 딥러닝 기반 시스템을 현장에 적용할 때 발생되는 오탐지 데이터를 활용하여 딥러닝 모델의 현장 적용성을 향상시킬 수 있을 것이다.
5. 이러한 결과는, 실시간으로 촬영되는 CCTV 영상과 연동하는 딥러닝 기반 영상유고 시스템은 설치 초기에 발생되는 오탐지 결과를 지속적으로 추가 학습한다면 별도의 시스템 수정이나 학습자료 확장을 위한 수동 레이블링 노력 없이도 스스로 인지성능 향상과 오탐지를 저감시키는 효과를 기대할 수 있음을 의미한다.

감사의 글

본 연구는 한국건설기술연구원 중소중견기업 수요기반 기술사업화 지원사업(딥러닝 기반 공공시설물 유고상황 감지 전주기 시스템 개발 및 최적화 연구)의 일환으로 수행되었으며, 이에 깊은 감사를 드립니다.

References

1. Davis, J., Goadrich, M. (2006), "The relationship between Precision-Recall and ROC curves", Proceedings of the 23rd International Conference on Machine Learning, pp. 233-240.
2. Everingham, M., Gool, L.V., Williams, C.K.I., Winn, J., Zisserman, A. (2010), "The pascal visual object classes (VOC) challenge", International Journal of Computer Vision, Vol. 88, No. 2, pp. 303-338.
3. Kim, D.G., Shin, Y.W., Shin, Y.S. (2012), "Section enlargement by reinforcement of shotcrete lining on the side wall of operating road tunnel", Journal of Korean Tunnelling and Underground Space Association, Vol. 14, No. 6, pp. 637-652.
4. Kim, T.B. (2016), "The national highway, expressway tunnel video incident detection system performance analysis and reflect attributes for double deck tunnel in great depth underground space", Journal of Korea Institute of Information and Communication Engineering, Vol. 20, No. 7, pp. 1325-1334.
5. LeCun, Y., Yoshua, B., Geoffrey, H. (2015), "Deep learning", Nature, Vol. 521, pp. 436-444.
6. Lee, K.B., Shin, H.S., Kim, D.K. (2018), "Development of a deep-learning based automatic tracking of moving vehicles and incident detection processes on tunnels", Journal of Korean Tunnelling and Underground Space Association, Vol. 20, No. 6, pp. 1161-1175.
7. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L. (2014), "Microsoft coco: Common objects in context", European Conference on Computer Vision. Springer, Cham, pp. 740-755.
8. Ministry of Land, Infrastructure and Transport (MOLIT) (2016), Guideline of installation and management of disaster prevention facilities on road tunnels.
9. National Committee for Land and Transport (2016), "Tunnel accidents increase, but tunnel incident automatic detection system often fails in operation", Press Release from parliamentary inspection of MOLIT. Congressman Yoon Hoo-Dyuk.
10. Ren, S., He, K., Girshick, R., Sun, J. (2015), "Faster R-CNN: Towards real-time object detection with region proposal networks", Proceedings of the Advances in Neural Information Processing Systems, pp. 91-99.
11. Shin, H.S., Lee, K.B., Yim, M.J., Kim, D.K. (2017), "Development of a deep-learning based tunnel incident detection system on CCTVs", Journal of Korean Tunnelling and Underground Space Association, Vol. 19, No. 6, pp. 915-936.
12. Zhu, M. (2004), "Recall, precision and average precision", Department of Statistics and Actuarial Science, University of Waterloo, Waterloo 2: 30.