



딥러닝(CNN)기반 저해상도 IR이미지 분석을 통한 작업자 인식

오원식 · †이우귀연 · 오정석

한국가스안전공사

(2018년 10월 18일 접수, 2018년 11월 23일 수정, 2018년 11월 24일 채택)

Deep Learning(CNN) based Worker Detection on Infrared Radiation Image Analysis

Wonsik Oh · †Ugwiyeon Lee · Jeongseok Oh

Korea Gas Safety Corporation

(Received October 18, 2018; Revised November 23, 2018; Accepted November 24, 2018)

요약

플랜트 내 위험지역의 안전을 위해 작업자 중심의 안전관리가 필요하다. 최근 5년간 가스 사고의 원인은 시설 노후 및 장비고장 뿐만 아니라, 사용자의 취급부주의나 고의사고, 공급자 취급부주의 등 작업자의 행동에 밀접한 관련이 있다. 이와 같은 사고를 미연에 방지하기 위해서, 플랜트 내 위험지역에 대한 실시간 모니터링이 필요로 하다. 하지만 실시간 모니터링을 위해서 작업(근로)공간에 카메라 설치 시, 인권침해와 같은 문제가 발생한다. 이를 방지하기 위해서 작업자의 신원 노출이 적은 저해상도의 Infrared 카메라를 이용한다. 또한 실시간 모니터링 시, 사람이 아닌 CNN알고리즘을 이용하여 이미지 분석을 통하여 인권침해 문제를 예방한다.

Abstract - worker-centered safety management for hazardous areas in the plant is required. The causes of gas accidents in the past five years are closely related to the behavior of the operator, such as careless handling of the user, careless handling of the suppliers, and intentional, as well as equipment failure and accident of thought. In order to prevent such accidents, real-time monitoring of hazardous areas in the plant is required. However, when installing a camera in a work space for real-time monitoring, problems such as human rights abuse occur. In order to prevent this, an infrared camera with low resolution with low exposure of the operator is used. In real-time monitoring, image analysis is performed using CNN algorithm, not human, to prevent human rights violation.

Key words : deep learning, CNN, human error, object detection

1. 서론

가스 안전 공사의 2017년도 가스통계에 따르면 국내 연도별 가스 수요는 계속 증가 중이다(Table 1). 또한 전국에 1,117,314개의 가스 관련 시설 및 업소가 존재하며[1], 2017년 기준 최근 5년간 사용처별 사고현황에서 전체 사고 602건 중 130건(21.6%)의 가스 사고가 플랜트(공장 및 공급시설 등)에서 발생 하였다. 또한 2017년 기준 최근 5년간

602건의 가스사고 건에 대하여 676명의 인명피해가 발생했다[2].

이처럼 가스사고 발생 시 피해를 동반한다. 그러므로 가스 시설물이 존재하는 플랜트에 대한 높은 수준의 안전관리가 요구되며, 높은 수준의 안전관리를 위해서 가스 시설이 존재하는 플랜트 지역 내 위험도 변화가 예측되는 종합적인 안전기술이 필요하다. 가스 사고발생은 위험가스시설 근처에 있는 작업(사용)자의 작업(또는 행동)과 밀접한 연관성이 높은 것으로 조사된다. 최근 5년간 원인별 가스 사고 현황에 따르면(Table 2), 시설 노후 및 장비고장 뿐만 아니라 사용자의 취급부주의, 공급자의 취급

†Corresponding author: lyullee@kgs.or.kr

Copyright © 2018 by The Korean Institute of Gas

딥러닝(CNN)기반 저해상도 IR이미지 분석을 통한 작업자 인식

부주의 등과 같이 작업자의 행동에 의한 사고발생 비율은 40%이상 차지하고 있다. 이와 같은 사고를 예방하기 위해서, IT융합기반 실시간 모니터링을 실시한다[3].

모니터링을 통해 작업자에 의한 플랜트 내 위험도변화를 실시한다면, 이와 같은 작업자 행동에 의한 사고를 미연에 방지하고, 사고 발생 시 효율적으로 대응할 수 있다.

플랜트 내 작업자에 의한 위험도 변화를 실시간 모니터링 하여 예측하기 위해서 작업 공간 내에 영

상정보처리기기(CCTV)를 설치하여야 한다. 하지만 작업공간에 CCTV 설치된다면 여러 가지 문제점이 발생한다. 먼저 실시간 모니터링 시, CCTV 요원의 인식률, 비용적인 문제, 모니터링에 의한 작업자 인권침해 문제와 가스 플랜트 내 전자기기 설치의 안전성에 따른 비용적인 문제가 발생한다.

CCTV 요원통한 모니터링은 영국의 PSDB(Police Scientific and Development Branch)연구결과에 따르면 CCTV 1인당 모니터링 화면이 1, 4, 6, 9개로 증가할 때마다 인식률이 85%, 74%, 58%, 53%로 감

Table 1. Gas demand per year for the last five years as of 2017, 2017 Gas Statistics of Korea Gas Safety Corporation[1]

구분	총가구수	LPG 사용 수요가 수			도시가스 사용 수요가수					가스사용 수요가수 (계)
		주택용	영업용	소계	단독주택	공동주택	산업용	상업+업무용	소계	
2013	21,919,518	4,832,244	492,841	5,325,085	4,865,546	10,758,798	14,789	736,105	16,375,238	21,700,323
2014	22,170,727	4,595,065	472,740	5,067,805	5,122,114	11,006,915	15,021	737,165	16,881,215	21,949,020
2015	22,422,161	4,374,992	461,688	4,836,680	5,262,108	11,327,741	15,024	758,652	17,363,525	22,200,205
2016	22,767,734	4,109,326	452,474	4,561,800	5,433,777	11,750,906	14,996	779,625	17,979,304	22,541,104
2017	23,134,508	3,886,841	451,435	4,338,276	5,592,589	12,153,421	15,081	804,980	18,566,071	22,904,347

Table 2. Cause of accident by year for last 5 years as of 2017, 2017 Gas Year Book of Korea Gas Safety Corporation[2]

구분	2013년	2014년	2015년	2016년	2017년	계	점유율(%)
계	121	120	118	122	121	602	100
사용자취급부주의	41	41	41	38	31	192	33.3
공급자취급부주의	12	9	5	3	3	32	6.4
타공사	2	6	4	8	7	27	3.8
시설미비	23	17	22	26	29	117	18.8
제품노후(고장)	4	14	19	12	18	67	8.6
*기타	18	17	11	23	24	96	16.3
고의사고	21	16	16	12	9	74	12.7

*기타(24) : 교통사고(8), 작업부주의(7), 시공 작업(3), 과열화재(2), 원인미상(1), 용기취급부주의(1), 기타(2)

소한다[4]. 이처럼 모니터링 요원 1인당 화면이 증가 할수록 인식률이 감소하게 된다. 인식률을 높이기 위해서, 모니터링 요원을 증원한다면 인원에 대한 추가 비용이 증가 한다. 플랜트의 경우 가스시설이 존재하는 지역은 작업(근로)공간이 경우가 있다. 작업(근로)공간에 영상정보처리 기기 설치 및 운영 시 인권침해 논란이 생길 수 있다. 하지만 개인정보 보호법 제25조(영상정보처리기기의 설치·운영 제한)에 따라 시설안전 등을 목적으로 하는 CCTV의 설치하는 가능하기 때문에 근로 공간에 CCTV를 설치하는 것은 특별한 법적 문제가 없다.

그러나 근로 공간의 특성 상 근로자가 일정한 공간에서 장시간 머물러 있게 되는데, 이 과정에서 해당 공간에서 생활하는 근로자의 모든 행동이 촬영되고 지속적인 감시에 의한 근로자 개인의 초상권 침해, 사생활 침해, 인권침해의 소지 등이 있을 수 있으며, 근로자의 입장에서 감시당하고 있다는 정신적인 스트레스로 인한 피해를 받을 우려도 있을 수 있다[5]. 또한 국민인권위원회에서도 보도 자료를 통해 CCTV 관련 진정·상담이 지속적으로 늘어나고 있으며[6], 시설물 안전 관리나 도난 방지 등 목적으로 설치한 CCTV를 동의 없이 직원 근무 감시에 사용한 사례를 인권침해라고 결정하였다[7].

가스시설이 존재하는 플랜트 내에 안전성의 이유로, 전기기기에 대한 방폭 기준(방폭전기기의 설계, 선정 및 설치에 관한 기준)이 존재한다. 이 기준에 맞춰 단가 높은 고화소의 카메라를 기준에 맞추어 설치 시 높은 비용이 발생할 것으로 예상된다.

단가가 낮은 저화소의 열화상 카메라에 방폭 기준을 적용해 가스시설이 존재하는 플랜트 내에 설치 시 비용적인 측면을 해결 할 수 있으며, 작업자의 신분 노출이 가능한 고화소 영상정보처리기기 대신, 저화소(80×60×1픽셀)의 열화상 카메라 사용하여 신분 노출을 감소시킨다. 이와 같이 저화소의 IR이미지를 딥러닝(Deep Learning) 알고리즘으로 분석하여 모니터링을 실행한다면, 직접적인 모니터링이 아닌 딥러닝 알고리즘의 의한 모니터링 시스템으로 작업자의 반감을 감소시킬 수 있을 것이다.

II. 관련 이론

2.1. 딥러닝(Deep learning)

인공신경망(artificial neural network, ANN)은 기계학습과 인지과학에서 생물학의 신경망에서 영감을 얻은 통계학적 학습 알고리즘이며, 딥러닝은 깊은 구조로 이루어진 인공신경망 학습기법을 말한다. CNN(Convolution Neural Network), RNN(Re-

current Neural Network), DBN(Deep Belief Network), GAN(Generative Adversarial Network)등 다양한 종류의 심층 신경망 구조가 존재하며, 여러 비선형 변환기법의 조합을 통하여, 컴퓨터 비전 분야, 음성 신호 처리 및 음성인식, 자연어 처리 등 여러 분야에서 좋은 성능을 보이고 있다[8].

2.2. CNN(Convolution Neural Network)

컴퓨터 비전 분야의 대표적인 심층신경망인 CNN은 심층신경망의 한 종류로, Convolution layer를 포함하고 있다. CNN은 이미지 분석 경쟁대회인 'ILSVRC (ImageNet Large Scale Visual Recognition Challenge)'에서 2012년 Alex Krizhevsky가 제시한 GPU 기반을 둔 CNN모델인 AlexNet[9]이 16.4%의 오차율을 보이며, 26.2%의 오차율을 보인 다른 모델에 비해 압도적으로 오차율을 낮추며 우승을 하였다. 이를 시작으로 이후 대회에서 CNN모델이 주류가 되었고, 심층신경망을 적용한 여러 CNN모델들이 나왔다. 2014년 Google의 GoogleNet[10], 2015년 Microsoft의 ResNet[11] 등 오차율이 6.7%, 3.57%로 발전해 왔으며, 인간의 이미지 인식률 94.9%보다 정확하게 인식하게 되었다.

GoogleNet과 ResNet 각각 21,49개의 Convolution layer를 구성하고 있으며, CNN 모델을 평가해 보았을 때 Convolution 계층의 깊이가 깊을수록 이미지 분류의 정확도가 높다[12]. 하지만, Convolution 계층이 증가할수록 학습시간과 연산량이 증가한다[13].

AlexNet의 구조는 Fig. 1과 같으며, 5개의 Convolution layer와 3개의 Fully-Connected layer 총 8개의 계층으로 구성되어 있다. 낮은 층은 이미지 상의 각 세부영역에서 저수준의 특징들(점, 선, 면, 위치)을 찾아낸다. 층이 증가하면서 높은 층에선 서로 가까운 것들끼리 조합되어 복잡도가 증가하여 패턴의 변화가 증가하여 고수준의 특징을 찾아내게 되며 추상적인 표현이 학습되며, 마지막 convolution layer(Layer 5)에서 첫 번째 Fully-connected(Layer 6) 층으로 올라가면서 합성곱 계층의 결과가 합쳐지며 완전연결 계층들을 거쳐 결과를 출력하게 된다[14].

2.3. 객체인식(Object detection)

CNN모델이 2012년 ILSVRC에서 큰 성공을 거둔 뒤, 이미지 분류 분야에서, 객체 인식분야로 확장되었다.

과거의 객체 인식 연구는 SIFT(Scale Invariant Feature Transform), SURF(Speeded Up Robust Features), HOG(Histogram of Oriented Gradients)등과

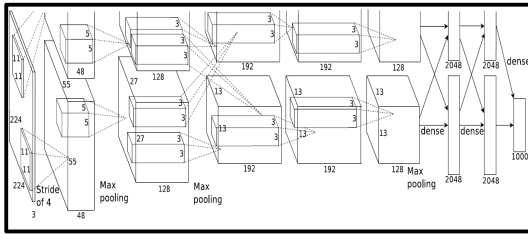


Fig 1. Structure of AlexNet, ImageNet Classification with Deep Convolutional Neural Networks[9]

같이 객체가 가지는 특징을 설계하고 검출함으로써 객체를 찾아내는 방식으로 진행되었다. CNN을 기반으로 객체의 특징을 찾아 학습 하여 객체의 위치를 검출하고, 인식하는 Faster R-CNN(Faster Region-based Convolution Neural Network)[15], R-FCN(Region-based Fully Convolution Networks)[16], YOLO (You Only Look Once)[17], SSD(Single Shot MultiBox Detector)[18] 등의 연구가 등장하였다.

III. 저해상도 이미지 분석 시스템 구성

Fig. 2는 실시간 위험도 산출이 진행을 나타낸 것이다. 위험 지역 내 작업자가 위치한다면, 인체감지 센서에 의하여 이벤트가 발생하게 된다. 이벤트 발생 시 IR카메라가 작동하여 촬영을 시작한다. 촬영된 IR 이미지는 서버로 전송되고 CNN알고리즘에 의하여 저장된 IR이미지를 분석하여 위험도를 변화 시킨다.

위험도산출은 시설 정보와 작업 스케줄등을 고려하여, 종합적으로 구성된 위험도 산출 프로세스 통해 실시된다. 위험 시설에서 발생하는 사고는 특정 장치의 고장이나, 잘못된 이벤트 발생을 통하여 연계적으로 일어난다. 이러한 과정을 논리적으로 구축한 것을 Fault Tree analysis(FTA)라 부른다 (Fig. 3). FTA 내에서 각 노들은 각각의 사고 발생 확률을 가지는데, 각 확률을 FTA논리에 따라 결합하면, 최후단에서 발생하는 사고발생 확률계산이 가능하다. 휴면에러 또한 확률(빈도)을 가지고 있으며, 시스템에서 작업자를 감지하지 않으면 0의 확률을 할당한다. 작업자를 감지하면 휴면에러 발생 확률을 FTA 시스템에 반영하여 전체 위험도를 예측한다.

가스 충전과 같은 위험도가 높은 작업이 실행되는 작업공간에는 정밀한 위험도 예측이 될 수 있도록, IR카메라를 통하여 실시간 모니터링을 실행한다.

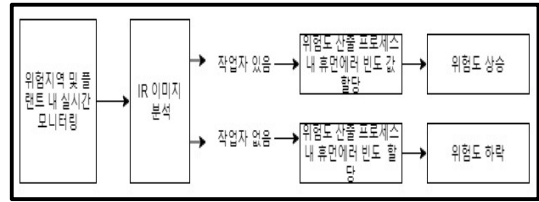


Fig. 2. Risk calculation process

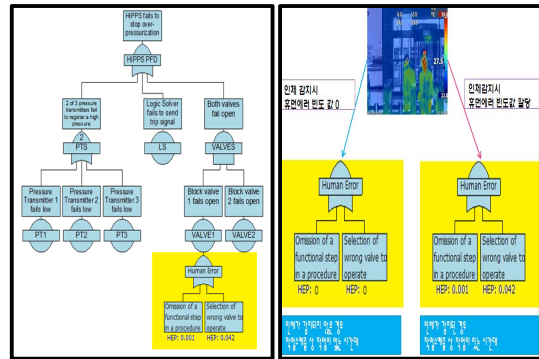


Fig. 3. Fault Tree Analysis

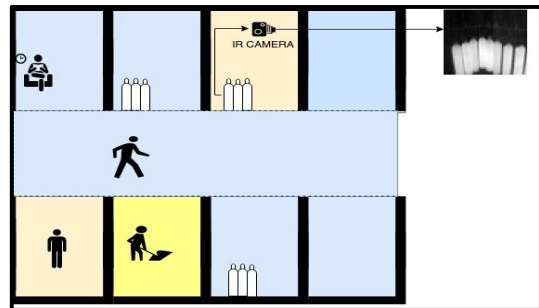


Fig. 4. Example of a hazard indication in the workplace 1

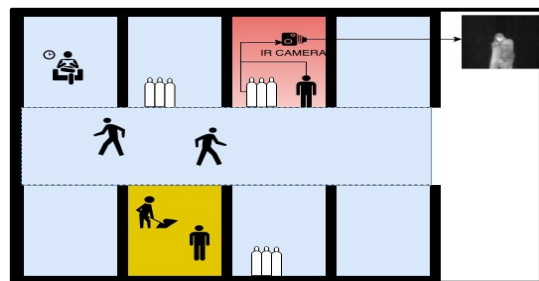


Fig. 5. Example of a hazard indication in the workplace 2

Fig. 4, 5과 같이 지역별로 위험도 산출 프로세스를 이용하여 위험도를 표시될 수 있다.

IR 카메라가 설치된 위험 지역에 작업자가 위치하게 된다면, 위험도 산출 프로세스에 의해서 휴먼에러 빈도 값을 할당하여 위험도를 산출된 위험도를 Fig. 5과 같이 표시할 수 있다.

3.1. IR 이미지 분석 과정

CNN기반의 작업자 인식은 다음 Fig. 6과 같이 실행 된다. 가스 실린더 충전 또는 시설 내 작업자 위치한다면, 온도의 변화가 발생하게 된다.

고정형 IR카메라가 설치된 지역은 온도 변화 발생 시, 센서에 의하여 IR 카메라의 촬영된다(Fig. 6.1. Input image). IR 이미지 내에 객체의 class(작업자, 가스 실린더, 그 외의 이미지)를 분류하기 위해서 객체의 위치를 찾을 수 있어야 한다. IR이미지의 명암 차이를 이용하여, 객체의 외각 검출(Fig. 6.3. .edge detection)을 실행한다. 이후 외각 검출을 이용하여, 객체의 위치를 box형태의 좌표로 검출한다(Fig. 6.4. bounding box). 이러한 외각검출의 정확도를 높이기 위해 이미지를 가공 하였다(Fig. 6.2. preprocess). 추출한 위치의 이미지는 이미 훈련이 완료된 CNN모델을 이용하여 객체를 분류를 실행한다. 하지만 정확한 위치와 class를 분류하기 위해서, box의 크기와 위치를 조정하여(Fig. 6.5. box resize and change location) 이미지를 추출한다(Fig. 6.6. image crop). 위치와 크기를 조정하여 추출한 이미지들을 훈련된 CNN모델을 사용하여 class를 예측한다(Fig. 6.7. object predict). 예측한 확률을 이용하여 이미지 내에 존재하는 객체의 class와 위치를 확인할 수 있다(Fig. 6.8. object detection).

3.2. 데이터셋 구성

군산 가스충전 시설 사업장의 작업 공간에 IR 카메라를 설치하여 실험환경을 구축하였다.

IR카메라를 통해 얻은 IR이미지를 임의로 추출하여, 3개의 class 사람, 가스 실린더, 그 외 부분(가스충전 밸브나 벽에 생기는 노이즈)에 해당하는 영역을 각각 380장, 총 1140장으로 dataset(Fig. 7.)을 구성했다. 1140장의 이미지 중 912장은 학습 데이터 셋(training data set)으로 활용하고, 나머지 228장은 검증 데이터 셋(validation data set)으로 활용한다. 학습 당 사용된 이미지 수(batch size)는 144장으로 500회 반복 훈련하였다. 객체 인식의 정확도를 평가하기 위해서 IR 이미지 중 사람이나 가스 실린더가 포함된 이미지 725장을 사용했다.

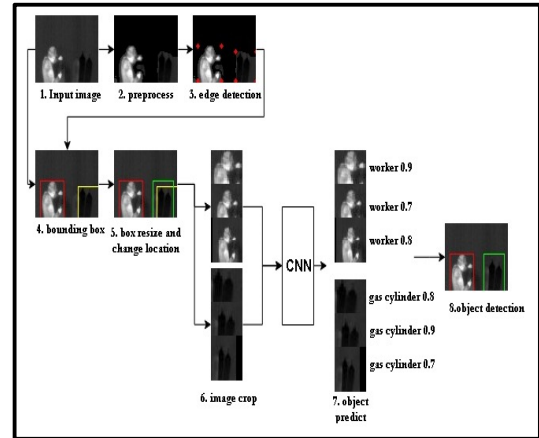


Fig. 6. IR image analysis process

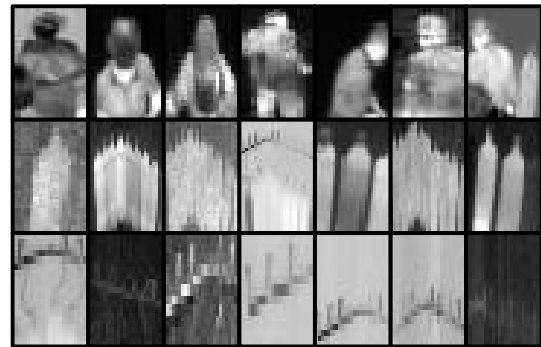


Fig. 7. data set composition

3.4. 이미지 전처리

Fig. 8과 같이 이미지 내 명암 구분이 선명하여, 이미지를 가공하지 않아도 외각검출을 통하여 이미지 내 객체의 위치와 class를 쉽게 분류 할 수 있는 이미지도 있다. 하지만 대부분의 이미지는 Fig. 10과 같이 이미지의 명암차이가 불명확한 이미지들이다. 이러한 문제를 해결하기 위해서, 이미지의 외각 검출을 실행 전, 2가지의 전처리를 실행하여 정확도를 높인다(Fig. 9).

첫 번째 과정은 Fig. 10과 같이 명암차이가 불명확한 이미지를 명암의 차이를 발생시켜 준다. 이미지의 픽셀값들의 분포 차이가 적은 경우 발생한다. 그렇기 때문에 0~255의 값을 가지는 픽셀 분포의 차이를 조정하여 명암차이를 발생 시킨다(Fig. 11).

두 번째 과정은 객체 위치 추출의 정확도를 높이기 위해서 배경부분을 제거한다. 작업장 내에 설치

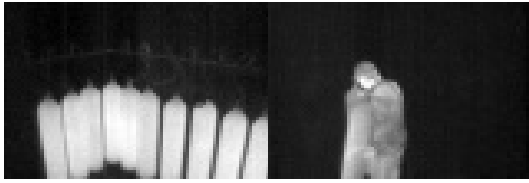


Fig. 8. Preprocessing image not required



Fig. 12. Image removed from background

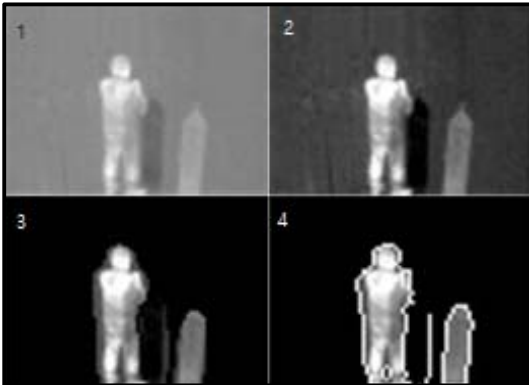


Fig. 9. Preprocessing



Fig. 13. Location detection of object in image

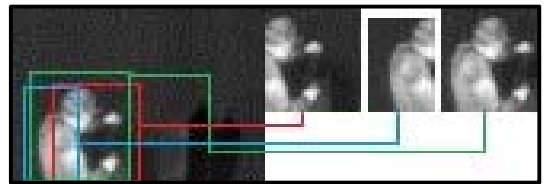


Fig. 14. Object classification

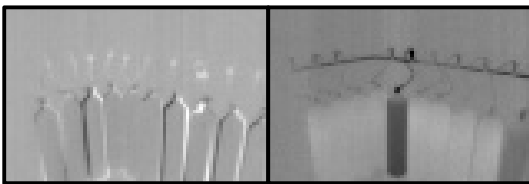


Fig. 10. Image requiring preprocessing

배경이 제거된 이미지는 명암차이를 활용하여 객체의 위치를 검출해 낸다.(Fig. 13.)

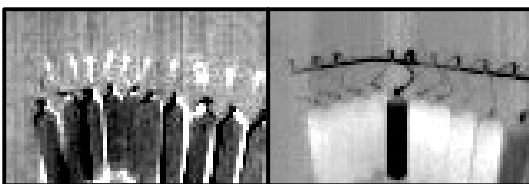


Fig. 11. First processing of images

3.5. 객체 인식

전처리과정을 거친 IR이미지를 외각검출을 통하여 객체의 위치를 box형태의 좌표로 추출한다.

추출한 영역의 좌표를 이용하여 원본 이미지에서 좌표의 크기를 증가/감소 또는 좌표의 위치를 바꾸어 이미지를 추출한다. 추출한 이미지들은 CNN모델에 의해 해당 class가 분류된다. 변경된 위치와 크기가 변경된 좌표 중, 가장 높은 확률을 가지는 영역으로 이미지 내에 존재하는 객체의 class와 위치를 확인할 수 있다(Fig. 14.).

된 IR카메라가 고정형인 것을 활용했다. 카메라가 고정되어 있기 때문에 이미지의 상단부분의 일부 픽셀을 추출하여 평균을 구한 후 평균값을 사용하여 배경을 제거 한다 (Fig. 12.). 배경이 제거되지 않은 경우 이미지의 노이즈가 강하기 때문에 배경이 제거된 이미지의 보다 객체 위치검출의 정확도가 떨어진다.

IV. 이미지 분석 결과

실험환경은 32GB의 RAM(Random Access Memory)과 GeForce GTX 1050에서 Keras 딥러닝 프레임 워크를 사용하여 AlexNet CNN모델을 구성한다. AlexNet은 224x224x3 크기의 이미지를 1000개의 class로 분류한다. 하지만 80x60x1의 크기의 이

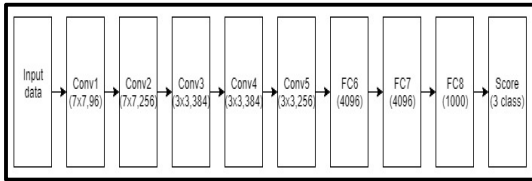


Fig. 15. AlexNet model reconfigured

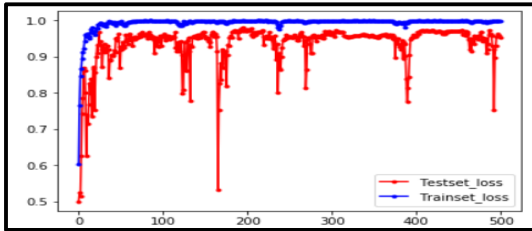


Fig. 16. Accuracy and loss graphs

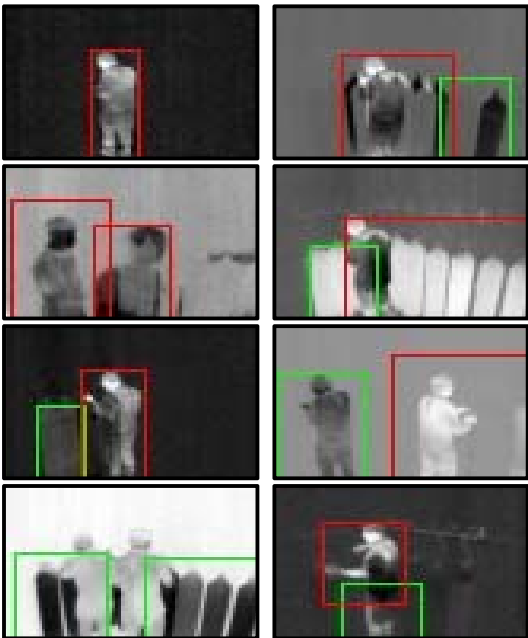


Fig. 17. As a result of Bounding Box, The red box area is the worker, the green box area is the gas cylinder.

미지를 resize하여, 3개의 class로 데이터 셋을 구성했기 때문에, 기존 Alexnet의 파라미터를 조정하여 Fig. 15.와 같이 구성하였다.

AlexNet으로 구성한 CNN 모델을 학습 했을 때 테스트 셋으로 평가한 결과 약 95.3%의 정확도를 보였으며, 정확도와 손실률은 다음 Fig. 16.와 같다.

IR이미지 분석을 하여, 평가해보았다. 사람과 가스를 정확하게 구분하여 작업자의 유/무를 판단한 경우는 725장 중 583장으로 약 80.3%의 정확도를 보였다. Fig. 17.은 객체인식 후 Bounding box하여 객체를 분류를 실행한 결과 예시이다.

V. 결론

플랜트 내에 단순히 작업자의 유/무에 의한 위험도 변화보다 더 정밀한 예측을 위해서, 플랜트 내 작업자의 유/무, 작업 인원, 작업 여부가 함께 고려되어야 한다. 가스 충전시설에 IR 카메라를 설치 후 CNN을 기반 한 이미지 분석을 실행했다. 작업자의 유/무, 작업 인원은 분석 할 수 있지만, 작업 여부는 현재 이미지 분석만으로는 판단할 수 없다. 하지만 IR이미지 내 가스 실린더가 검출된다면 가스 실린더의 온도가 변화가 발생한 것으로 판단되며, 가스 실린더가 충전되어 있다고 판단할 수 있다. 이처럼 이미지 내에 작업자와 온도변화에 의하여 가스 실린더가 같이 검출된다면 작업 중인 것으로 판단하여 작업장 내의 위험도를 증가시킬 수 있다.

AlexNet 모델의 정확도는 95.3%로 3개의 class의 분류의 정확도는 높았다. 하지만 이미지 내 다중의 객체를 판별했을 때 80.3% 정확도를 나타냈다.

현재 이미지 전처리를 통하여 명암이 차이가 발생하는 부분에 대하여 모든 좌표를 구성 후, box값으로 반환 될 수 있는 좌표를 반환하여, 좌표의 크기를 증가/감소 시켜 box 영역의 가장 높은 확률의 객체를 좌표와 함께 반환하여 판단한다. 그렇기 때문에 정확히 객체의 위치를 반환 할 수 있는 경우에는 정확도가 높지만, 가스 실린더와 작업자의 영역을 나눌 수 없는 경우나 작업자가 겹쳐지는 부분이 많은 경우 작업자의 인원을 정확히 판단할 수 없는 오차가 발생하여, 단순한 외각검출만으로 객체의 위치를 검출하는 것은 한계가 발생 했다. 이점을 보완하기 위하여 CNN기반 한 객체인식 방법인 Faster R-CNN과 SSD와 같은 모델을 적용할 예정이다.

감사의 글

본 연구는 산업통산자원부의 재원으로 수행되었으며, 지원에 감사를 드립니다.
(과제번호 : 20172210100110)

REFERENCES

- [1] Korea Gas Safety Cooperation, "Gas accident statics for 2017", Korea Gas Safety Cooperation, (2018)
- [2] Korea Gas Safety Cooperation, "gasyearbook 2017", Korea Gas Safety Cooperation, (2018)
- [3] Oh, J. S., Sung, G. J., Kim, Y. D. "Developing Network Infrastructure and Smart Service for Safety Management of City-gas Facilities.", KiGAS, Vol. 15(1), 46~53, (2011)
- [4] Donald, C. "How many monitors should a CCTV operator view.", *CCVT Image*, 355, (2005)
- [5] Kang, S. H, Kim H. S, "Analysis of Privacy Issues by the Diffusion of Video Information Processing System.", *Internet & Security Focus 4*, 45-65, (2014).
- [6] National Human Right Commission of Korea, "Human Rights Commission of Korea, Korea, has launched an analysis on trends of grievance and counseling regarding CCTV.", press release, (2011)
- [7] National Human Right Commission of Korea, "Use of supervision of working attitudes other than CCTVs violates human rights.", press release, (2017)
- [8] Moon, S. E, et al. "Trends in Machine Learning and Deep Learning Technology.", *The Korean Institute of Communication Sciences*, 33(10), 49~55, (2016)
- [9] Krizhevsky, A., Sutskever, I., Hinton, Geoffrey E. "Imagenet classification with deep convolutional neural networks.", In: *Advances in neural information processing systems*, 1097-1105, (2012)
- [10] Szegedy, Christian, et al. "Going deeper with convolutions." *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1-9, (2015)
- [11] He, K. et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 770-778, (2016)
- [12] Simonyan, K., Zisserman, K. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556*, (2014)
- [13] Lee, S. J, Jung D. J, Ahn J. H. "Memory Bandwidth Analysis of Various Convolutional Layer Types on Convolutional Neural Networks", *KIISE*, 1409-1411, (2017)
- [14] Yosinski, J, et al. "Understanding neural networks through deep visualization." *arXiv preprint arXiv:1506.06579*, (2015)
- [15] Girshick, R. "Fast r-cnn." *Proceedings of the IEEE international conference on computer vision*, 1440-1448, (2015)
- [16] Dai, J., Li, Y., He, K. "R-fcn: Object detection via region-based fully convolutional networks." *Advances in neural information processing systems*. 379-387. (2016)
- [17] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 779-788. (2016)
- [18] Liu, W., Anguelov, D., Szegedy, C., Reed, S., Fu, C. Y., Berg, A. C. "Ssd: Single shot multibox detector." *conference on computer vision. Springer, Cham*, 21-37, (2016)