

Anomaly Detection in IR-Camera Using Deep Learning

Hye-Won Kim*, Sang-Min Kim*, Jung-Mo Sohn*

*Researcher, Epozen's research institute, Seoul, Korea

[Abstract]

In this paper, we propose an infrared camera anomaly detection system based on deep learning. In general industrial sites, infrared cameras are used for facility management. However, if something goes wrong with the infrared camera, recording stops and the status of the facility is unknown. Therefore, monitoring is necessary to prepare for this. Currently, only manpower is monitoring the camera. However, there are limitations such as human error and mass inspection. To solve this limitation, we compare performance using AlexNet, VGGNet, ResNet, which are types of deep learning, and propose an automated monitoring method with better performance by comparing the captured original image and the histogram extracted image as a preprocessing process. The data set was collected and used by KEPCO's infrared night vision camera, and the results of training the data set with the VGGNet model and predicting the results were the best. The result is that the input image is judged to be normal/abnormal for the image captured at a certain period, and the model performance came out as Accuracy 0.97, Precision 0.97, Recall 0.97, F1-score 0.97. In addition, it was confirmed that the data set preprocessed with the histogram performed better.

▶ **Key words:** IR-Camera, Anomaly Detection, Monitoring, Deep Learning, VGGNet

[요 약]

본 논문에서는 딥러닝 기반의 적외선 카메라 이상 탐지 시스템을 제안한다. 일반적인 산업 현장에서는 시설물 관리용으로 적외선 카메라를 사용하고 있다. 하지만 적외선 카메라에 이상이 생겼을 경우 녹화가 중지되고 시설물의 상태를 알 수 없다. 따라서 이를 대비하기 위한 모니터링이 필요하다. 현재 인력만으로 카메라 모니터링을 하고 있다. 하지만 인적 오류나 대량 검수 등 한계점이 존재한다. 이러한 한계점을 해결하기 위해 딥러닝의 일종인 AlexNet, VGGNet, ResNet을 사용하여 성능을 비교하고, 원본 영상을 캡처한 이미지와 전처리 과정으로 히스토그램을 추출한 이미지를 비교하여 더 우수한 성능의 자동화된 모니터링 방안을 제안한다. 데이터 세트는 한국전력에서 적외선 야간투시 카메라로 녹화한 영상을 수집해 사용했으며, VGGNet 모델로 데이터 세트를 학습하고 결과를 예측한 결과가 가장 우수했다. 결과는 입력 영상을 일정 주기로 캡처된 이미지에 대해 정상/비정상을 판단한 것이고, 모델의 성능은 Accuracy 0.97, Precision 0.97, Recall 0.97, F1-score 0.97이 나왔다. 또한, 히스토그램으로 전처리한 데이터 세트가 더 우수한 성능을 확인할 수 있었다.

▶ **주제어:** 적외선 카메라, 이상 감지, 모니터링, 딥러닝, VGGNet

-
- First Author: Hye-Won Kim, Corresponding Author: Jung-Mo Sohn
 - *Hye-Won Kim (gpdhs@epozen.com), Epozen's research institute
 - *Sang-Min Kim (ksm619@epozen.com), Epozen's research institute
 - *Jung-Mo Sohn (jmsohn@epozen.com), Epozen's research institute
 - Received: 2022. 10. 18, Revised: 2025. 03. 24, Accepted: 2025. 03. 24.

I. Introduction

4차 산업혁명으로 인한 기술의 발전으로 산업 현장에서는 자동화된 기계들이 많이 사용되고, 그에 따라 무인화도 진행되었다. 무인화는 기계의 이상 여부에 대한 모니터링을 곤란하게 한다. 따라서 일반적인 산업 현장에서는 기계의 이상 발생 탐지를 위해 적외선 카메라(Infrared Camera, IR-Camera)를 이용하여 이상 여부를 탐색한다[1]. 적외선 카메라는 적외선 파장을 사용하여 물체를 촬영하는 카메라이다. 적외선 파장은 가시광선과 마이크로파 사이인 $1.8 - 26.5 \mu\text{m}$ 의 스펙트럼을 가지며, 어두운 공간이나 야간 촬영 등 사람의 눈으로 확인하기 힘든 현장을 촬영하기에 적합하다[2]. 적외선 카메라의 종류는 적외선 야간투시 카메라(Night Vision Camera)와 적외선 열화상 카메라(Infrared Thermal Camera)가 있다[3]. 적외선 야간투시 카메라는 카메라에서 적외선 파장을 방출하여 물체에 반사되어 돌아오는 적외선 파장을 감지하는 방식이고, 적외선 열화상 카메라는 물체가 발산하는 적외선을 감지하여 온도를 표현하는 방식이다.

최근 적외선 카메라를 이용한 이상 발생 탐지의 연구들이 많이 수행되고 있다. 겨울철 자주 발생하는 블랙 아이스는 눈이 먼지와 뒤섞여 검게 얼어붙어 아스팔트와 구분이 되지 않는 것으로, 운전자의 시야에 보이지 않기 때문에 자동차가 미끄러지면서 교통사고가 발생할 수 있다. 하지만 적외선 카메라를 사용하면 도로 표면의 온도를 감지하여 블랙 아이스 발생 여부를 운전자에게 알릴 수 있다[4]. 또한, 적외선 열화상 카메라를 사용하여 콘크리트 균열도 찾을 수 있다. 콘크리트에 생긴 균열은 관리자가 알지 못하는 미세한 균열도 존재한다. 이러한 균열은 초기에 발견하지 못하면 큰 문제로 이어질 수 있다. 균열로 인한 2차 피해를 예방하기 위해 정확한 균열 탐지는 중요하다. 균열은 깊을수록 낮은 온도를 가지기 때문에 적외선 열화상 카메라를 이용하면 인력을 동원하지 않아도 균열의 깊이를 상세히 측정할 수 있다[5-6]. 이외에도 적외선 카메라를 이용한 차량 내·외부 조명 변화에 따른 얼굴 인증, 원자력발전소용 고압선 진단, 수·변전 설비 열화 진단 등에서 적외선 열화상 카메라를 사용하여 시설물 이상 탐지를 하고 있다[7-9].

산업 현장에서 발생하는 이상의 종류로는 화재 발생이나 자연재해(태풍, 돌풍, 폭우 등)로 인한 시설물 파손 및 균열 등이 있다. 그리고 시설물 이상이 발생하면 금전적 손실이나 인명피해 등 여러 문제점을 일으킨다. 파손된 시설물을 정비하는데 자재비, 인건비 등 자금이 필요하고,

인명피해로 연결될 수 있다. 한국전력의 경우 전기를 관리하는 시설물들로 구성되어있기 때문에 이상이 발생하면 정전피해가 발생하고, 사람이 이상을 복구하다 고압으로 인한 상해를 입을 수 있다. 이러한 피해들을 예방하기 위해서 시설물들을 관리하는 적외선 카메라의 이상 발생 탐지는 중요하다[10]. 하지만 앞선 연구들은 적외선 카메라에 이상이 발생한 경우는 고려하지 않았다.

한국전력에서는 적외선 야간투시 카메라를 설치하여 시설물의 상태를 촬영하고 녹화하고 있다. 적외선 야간투시 카메라의 정상 작동 여부를 판단하기 위해서는 주기적으로 카메라의 상태를 체크 할 필요가 있다. 왜냐하면, 적외선 야간투시 카메라에 이상이 발생한다면 현재 시설물의 상태를 정확하게 녹화할 수가 없기 때문이다. 현재는 동작 이상이 발생한 적외선 야간투시 카메라가 있는지, 카메라 관리자가 수시로 모니터링을 하며 발견 시 문제를 해결하고 있다.

하지만 카메라 관리자만으로 적외선 야간투시 카메라 모니터링을 하는 방식은 여러 어려움이 있다. 첫째, 사람은 피로감을 느끼기 때문에 24시간 내내 모니터링을 할 수 없다. 신체활동을 하면 피로가 쌓이고, 피로가 축적된 상태로 지속적인 활동을 하면 집중력이 감소하며 판단력이 흐려지고 실수할 확률이 증가한다. 둘째, 카메라 관리자가 화면을 보며 적외선 야간투시 카메라의 이상 여부를 판단하는 것보다 프로그램을 사용하여 이상 발생을 탐색하는 것이 더 효율적이다. 현재 1000여 대의 적외선 야간투시 카메라 모니터링을 해야 하는데, 이상이 발생한 적외선 야간투시 카메라는 카메라 관리자가 발견하기까지 시간이 얼마나 걸릴지 알 수 없고, 빠르게 알기 위해서는 많은 인력이 소요된다. 셋째, 사람의 눈으로 정상과 비정상 구분은 어려운 경우가 존재한다. 정상 프레임과 비정상 프레임의 화소값 차이가 작은 경우 사람의 눈으로 구분이 되지 않기 때문에 적외선 야간투시 카메라에 동작 이상이 발생했는지 모르고 방치될 확률이 높다.

이 한계점을 해결하기 위해 본 논문에서는 적외선 야간투시 카메라에 동작 이상이 발생하면 빠르게 감지하여 알려주는 딥러닝 기반의 적외선 야간투시 카메라 이상 발생 감지 시스템을 설계하고 구현하였다. 이 시스템은 입력 영상에 대해 구간(5초)별로 정상과 비정상을 판단하여 비정상 발생 시 발생한 경우 알려준다. 2장에서는 관련 연구를 소개하고, 3장에서는 제안하는 시스템의 실험 결과를 기술한다. 4장에서는 결론을 제시한다.

II. Preliminaries

1. Related works

본 연구는 야간투시 적외선 카메라로 촬영된 영상에서 이상 발생 여부를 판별하는 것이 목적이다. 따라서 이상 탐지 시스템의 모델로 최근까지 이미지 분야에서 활용되고 있는 CNN이 적합하다고 판단했다[11-13]. 학습 모델로는 AlexNet, VGGNet, ResNet을 선택했으며, 선정 기준은 이미지 분류에 활용되고 있는 모델인지, 데이터 규모에 적합한 규모의 네트워크인지였다.

1.1 AlexNet

AlexNet은 2012년에 ILSCVRC 대회에서 Top 5 test error 기준 15.4%로 1위를 차지한 딥러닝 모델로 명칭은 저자의 이름을 딴 것이다[14]. 약 120만 개의 이미지를 데이터 세트에 사용했으며, 데이터 전처리 과정으로 입력 이미지의 크기를 256×256 으로 조정했다. 왜냐하면 FC(Fully Connected) layer에 들어갈 feature의 input size를 고정해야 하기 때문이다.

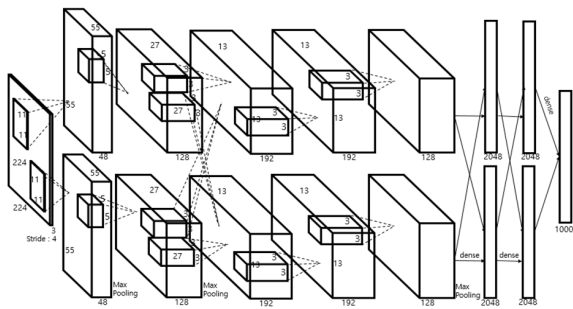


Fig. 1. Architecture of AlexNet[14]

Fig. 1은 AlexNet의 구조를 나타낸 것이다. 5개의 Convolution layer와 3개의 FC layer로 구성되어 있다. 120만 개의 거대한 데이터 수를 처리하기 위해 네트워크를 2개의 GPU로 나눠서 학습하여 최종 output을 출력한다. 예를 들어 90개의 kernel을 사용한다면 1개의 GPU에 45개씩 나눠서 사용하는 것이다. 또 AlexNet은 LRN(Local Response Normalization)을 사용하여 데이터를 정규화시켰다. 학습 중에 정규화를 진행하지 않으면 데이터가 한 방향으로 편향되어 gradient vanishing이 발생하고 학습이 제대로 되지 않는 현상이 나타나기 때문에 정규화 과정은 필요하다.

1.2 VGGNet(Visual Geometry Group Network)

VGGNet은 2014년에 영국의 VGG 그룹에서 만들어서 처음 소개된 모델로 이전 모델보다 네트워크가 깊어졌다. VGGNet의 특징은 layer 수가 늘어난 데 비해 연산량이 적다는 것이다. Alexnet은 11×11 의 큰 필터를 사용했지만, VGGNet은 신경망을 학습할 때 모든 layer에서 3×3 의 작은 필터로 여러 번 계산하는 방식으로 바꾸어 연산할 때 필요한 파라미터 값을 감소시켰다[15]. Convolution layer끼리 결합하여 특징 추출 가능성을 높였다. VGGNet은 Convolution layer - Max Pooling layer - FC layer 구조로 이루어져 있다.



Fig. 2. Architecture of VGG-11[15]

Fig. 2는 VGG-11의 구조를 나타낸 것이다. VGG-11은 8개의 Convolution layer와 3개의 FC layer로 구성된 네트워크로 기존 VGGNet에서 가장 작은 네트워크이다. VGGNet은 구성 layer의 수에 따라 VGG-11, VGG-13, VGG-16, VGG-19로 나뉘는데 이는 깊이에 따른 학습률 변화를 보기 위한 것이다.

1.3 ResNet

ResNet은 Kaiming He에 의해 2015년에 소개된 모델로 이전에 나온 모델들의 네트워크가 최대 20층이던 것을 최대 152층까지 늘려 에러율 3.7%의 좋은 성능을 보였다[16].

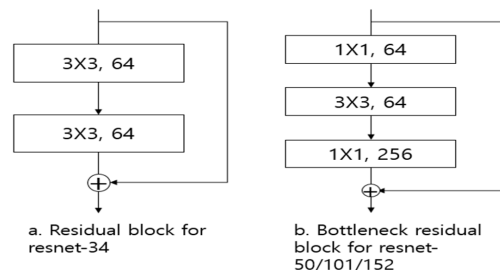


Fig. 3. Residual Block for ResNet[16]

Fig. 3은 ResNet의 특징이 되는 Residual 블록의 구조를 나타낸 것이다. ResNet은 이 블록을 사용함으로써 계층을 깊게 쌓을 수 있었고, 기울기 소실 문제를 개선했다[17]. 네트워크의 layer 개수에 따라 사용하는 Residual 블록 구조가 다르며 각각은 다음과 같다.

a. Building block은 skip connection과 3×3 convolution 블록 두 개로 구성되어있고, ResNet-34에서 사용했다.

b. Bottleneck building block은 skip connection과 1×1 convolution 블록 두 개 사이에 3×3 convolution 블록이 한 개 있는 구조로, 1×1 convolution 블록이 차원을 감소시킨 후 3×3 convolution 블록을 통과하기 때문에 병목 현상이 일어난다. 이 구조는 ResNet-50, ResNet-101, ResNet-152에서 사용했다.

1.4 Histogram

히스토그램은 어떤 데이터를 수치화하여 도수분포표를 구하고, 이를 누적한 값으로 그 결과를 보기 쉽게 그래프로 표현한 것이다. 히스토그램의 가로축은 영상의 밝기를 나타내고, 세로축은 영상 내의 화소 수를 나타낸다. Fig. 4는 Matlab에서 제공하는 히스토그램의 예시다.

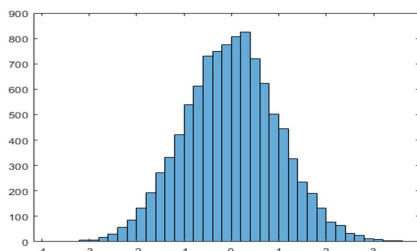


Fig. 4. Example of Histogram[18]

히스토그램을 활용하면 이미지 데이터를 사용할 때보다 좋은 결과를 도출할 가능성이 있다. 히스토그램은 픽셀별 화소값으로 그래프를 그리기 때문에 이미지로는 구분이 힘든 특징을 구분할 수 있고, Histogram equalization, Histogram stretching 등 다양한 방법으로 활용할 수 있다. 원본 이미지가 아닌 히스토그램을 사용하면 데이터의 질이 개선되므로 모델의 정확도도 높아지고 성능 향상에 도움이 된다[19-21]. 이에 본 연구에서는 야간투시 적외선 카메라의 원본 이미지보다 히스토그램 이미지를 사용하면 데이터의 형태가 생기고 시인성이 높아져 성능 향상에 도움이 될 것이라고 예상하여, 고장 탐지의 전처리 방법으로 히스토그램이 적합하고 판단하여 적용했다.

1.5 2D-FFT(Two Dimensional Fast Fourier Transform, 2차원 고속 푸리에 변환)

2D-FFT는 상관 패턴 인식, 의료 이미지 처리 등 다양한 이미지들을 처리하는 문제에서 사용하고 있다[22]. 이

기법은 1차원 푸리에 변환을 이미지에 적용한 것으로, x축 또는 y축을 따라가며 픽셀의 밝기 변화를 주파수로 변형하여 나타낸다. 이미지는 유한한 공간에서의 이산 신호로, 크기가 $M \times N$ 인 이미지 $f(x,y)$ 에 대한 2D-FFT 값은 식 (1)을 통해 구할 수 있다[23].

$$F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)e^{-j2\pi(ux/M+vy/N)} \quad (1)$$

식 (1)의 구성요소들은 다음과 같다. $F(u,v)$ 는 출력 이미지이고, j 는 허수단위 $j = \sqrt{-1}$ 이고, u 와 v 는 주파수이고, x 와 y 는 좌표이고, M 과 N 은 픽셀 수이다. 2D-FFT를 적용한 이미지를 살펴보면 위상에 따라 다른 주파수의 형태를 보인다. 의료 영상 처리에서는 이 특징을 이용하여 질환 여부를 진단한다[24]. 본 연구에서는 적외선 야간투시 카메라의 정상과 비정상 이미지의 주파수 형태가 다를 것이라고 예상하여, 고장 탐지를 위한 전처리 과정으로 2D-FFT가 적합하다고 판단하여 적용했다.

1.6 Classification Performance Evaluation Index

본 연구에서는 모델의 성능을 수치로 확인하기 위해 분류 성능 평가 지표를 사용했다. Table 1은 Confusion Matrix를 나타낸 것이다. Confusion Matrix는 분류성능 지표 식에 사용되는 요소들을 나타낸 행렬로, 실제값과 예측값의 일치도를 표로 나타낸 것이다. 행렬에 속하는 요소는 다음 4가지로 수식을 단순화하기 위해 레이블을 지정했다[25]. 첫째, TP는 True Positive의 약자로 실제값이 True이고, 예측값이 True인 경우를 말한다. 둘째, FP는 False Positive의 약자로 실제값이 False인데 예측값은 True로 나온 경우를 말한다. 셋째, FN은 False Negative의 약자로 실제값이 True인데 예측값이 False로 나온 경우를 말한다. 넷째, TN은 True Negative의 약자로 실제값이 False인데 예측값이 False로 나온 경우를 말한다.

Table 1. Confusion Matrix

		actual	
		true	false
predicted	true	true positive (tp)	false positive (fp)
	false	false negative (fn)	true negative (tn)

사용된 분류성능평가지표는 Accuracy, Precision, Recall, F1-score이다. 0.0 ~ 1.0 사이의 값을 가지며 각각 식 (2) ~ (5)를 통해 구할 수 있다.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \quad (2)$$

Precision이 1.0일 때는 FP가 0이고, Recall이 1.0일 때는 FN이 0인 것을 나타낸다.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

F1-score는 Precision과 Recall의 조화평균을 이용한 것이다.

$$F1\text{-score} = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} \quad (5)$$

III. The Proposed Scheme

실험은 데이터 전처리로 히스토그램과 2D-FFT를 사용한 방법, AlexNet, VGGNet, ResNet의 네트워크를 재구성한 모델을 사용한 방법, ReLU(Rectified Linear Unit)와 Leaky ReLU(Leaky Rectified Linear Unit)를 비교한 방법을 사용해서 진행했다. 실험 환경은 Table 2와 같다.

Table 2. Experiment Environment

input size	(288, 432, 3)
loss	binary_crossentropy
optimizer	adam
gpu	Tesla T4
tensorflow	2.8.2
weight initialization	he_normal
batch size	32
learning rate	0.001

1. Experiment

1.1 Data Preprocessing

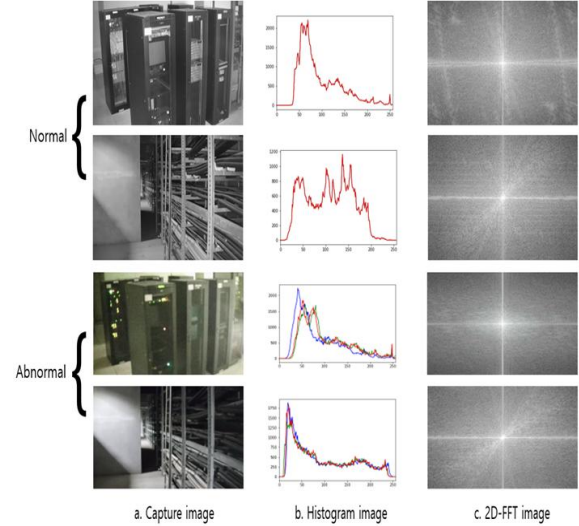


Fig. 5. Example of data preprocessing

Fig. 5는 실험에 사용된 데이터 세트의 전처리 예시들을 나타낸 것이다. 1행과 2행은 정상인 경우이고, 3행과 4행은 비정상인 경우이다. 데이터는 총 244개로, 학습 데이터 200개와 평가 데이터 44개로 구성했다. 입력 이미지의 크기는 모두 288×432로 124,416개의 픽셀을 가지며, 각각은 다음과 같다.

a. Capture image는 영상에서 추출한 원본 이미지다.

b. Histogram image는 Capture image마다 픽셀의 화소값을 히스토그램으로 그린 후 이미지로 저장하여 데이터 세트를 구성한 것이다. 정상의 경우 고르게 분포되어 있던 화소값이 비정상이 되자 어두운 영역으로 편향된 것을 확인하였다. 이는 정상과 비정상 데이터를 효과적으로 구분할 수 있는 특징이 될 것으로 예상된다.

c. 2D-FFT image는 Capture image마다 2차원 고속 푸리에 변환을 적용한 후 이미지로 저장하여 데이터 세트를 구성한 것이다. 스펙트럼의 형태를 파악할 수 있도록 스펙트럼의 위치를 원점이 중심에 오도록 이동시켰다.

1.2 Model

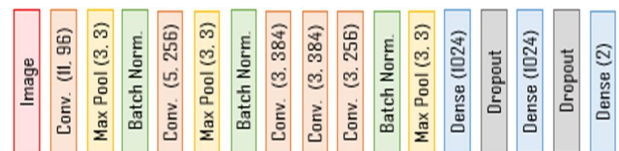


Fig. 6. Architecture of Alexnet



Fig. 7. Architecture of VGGNet

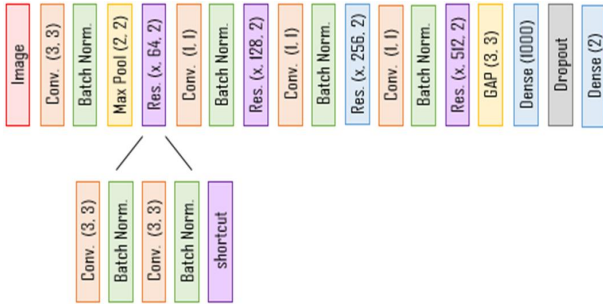


Fig. 8. Architecture of ResNet

Fig. 6은 Alexnet, Fig. 7은 VGGNet, Fig. 8은 ResNet을 수정하여 만든 모델이다. kernel의 크기와 channel의 수를 조정하여 학습한 결과 네트워크가 적정 파라미터 수를 가지는 경우에만 학습이 되는 것을 확인하였다. 따라서 네트워크의 크기를 줄인 AlexNet, VGGNet, ResNet을 학습 모델로 선택했다. 기존 모델보다 향상된 성능을 도출하기 위해 AlexNet은 LRN 대신 BN(Batch Normalization, 배치 정규화)을 사용했고[26], FC layer에는 dropout을 적용했다. VGGNet과 ResNet은 기존의 네트워크 크기를 줄여 각각 9개, 14개의 layer로 구성했다.

1.3 Activation Function

실험에 사용된 활성화 함수로 ReLU와 Leaky ReLU를 사용하였다. 활성화 함수를 변경하여 적용한 이유는 소규모 데이터로 진행한 실험에서 동일 네트워크를 사용하는 경우 ReLU보다 Leaky ReLU가 좋은 성능을 보였기 때문이다[27].

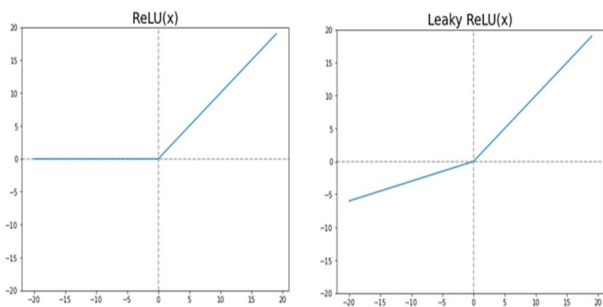


Fig. 9. ReLU and Leaky ReLU Activation Function

Fig. 9는 실험에 사용된 두 활성화 함수의 그래프를 나타낸 것으로 왼쪽이 ReLU, 오른쪽이 Leaky ReLU이다. ReLU는 x가 음수면 0을, 양수면 x값을 반환하는 반면에 Leaky ReLU는 x가 음수면 알파 계수를 곱한 값을, 양수면 x값을 반환한다.

2. Results

Table 3-4는 실험의 결과로, 각각 분류성능평가지표에 따라 Accuracy, Precision, Recall, F1-score의 값을 구하여 표로 나타낸 것이다.

Table 3. Performance Comparison of three data using AlexNet and VGGNet and ResNet

50 epoch	metrics	capture image	histogram	2D-FFT
alexnet	accuracy	0.74	0.87	0.64
	precision	0.83	0.87	0.95
	recall	0.75	0.90	0.37
	f1-score	0.71	0.88	0.43
vggnet	accuracy	0.81	0.90	0.60
	precision	0.79	0.92	0.58
	recall	0.85	0.85	0.93
resnet	accuracy	-	0.85	-
	precision	-	0.99	-
	recall	-	0.72	-
	f1-score	-	0.82	-

Table 3은 Capture image, Histogram image, 2D-FFT image에 AlexNet, VGGNet, ResNet을 적용하여 학습시킨 결과를 나타낸 것이다. 학습이 되지 않은 경우는 '-'로 표기했다. 결과는 50 epoch에서 5번 학습한 결과의 평균값을 기재한 것으로 f1-score 기준 AlexNet은 0.71, 0.88, 0.43이고, VGGNet은 0.78, 0.89, 0.71이고, ResNet은 학습 불가, 0.82, 학습 불가로 나왔다. f1-score 기준 2D-FFT image에 AlexNet을 적용했을 때 0.43으로 가장 낮은 성능을 보였고, Histogram image에 VGGNet을 적용했을 때 0.89로 가장 높은 성능을 보였다.

Table 4. Performance Comparison According to Alpha Value of Leaky ReLU using VGGNet

50 epoch	metrics	relu (0.0)	0.1	0.05	0.01
vggnet	accuracy	0.90	0.94	0.95	0.97
	precision	0.92	0.96	0.97	0.97
	recall	0.85	0.91	0.93	0.97
	f1-score	0.89	0.93	0.95	0.97

앞선 실험에서 f1-score 기준 Histogram image에 VGGNet을 적용하는 것이 가장 좋은 성능을 보인다는 것을 확인했다. 이에 성능을 향상하고자 활성화 함수를 ReLU에서 Leaky ReLU로 변경하였다. Table 4는 Leaky ReLU의 최적화 성능을 찾기 위해 알파 계수의 값으로 0.1, 0.05, 0.01을 적용하여 ReLU와 비교한 결과를 나타낸 것이다. 50 epoch에서 알파 계수가 0.01일 때 f1-score 0.97로 가장 좋은 성능을 보였다.

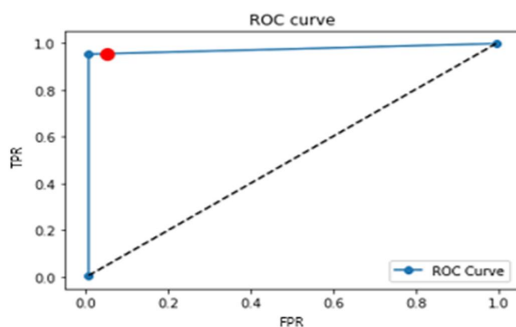


Fig. 10. ROC Graph of Best Performance

Fig. 10은 가장 좋은 성능을 보였던 VGGNet의 민감도와 특이도를 이용하여 ROC(Receiver Operator Characteristic) 곡선을 그린 것으로 AUC(Area Under the Curve)는 0.98이다.

IV. Conclusions

최근 산업 현장에서 시설물을 관리하기 위해 적외선 카메라를 많이 활용하고 있다. 따라서 적외선 카메라의 이상 발생 탐지는 중요하다. 기존에는 카메라 관리자가 감시하여 이상이 발생하면 해결하는 방식을 사용했다. 하지만 여러 한계점이 존재했고, 이를 해결하기 위한 방식이 필요하다. 새로운 방식으로 본 연구에서는 적외선 야간투시 카메라로 촬영된 영상에서 자동으로 정상/비정상을 판단하는 딥러닝 기반 자동모니터링 시스템을 제안했다. 산업 현장에서 딥러닝 기반 자동모니터링 시스템을 사용한다면 시설물 관리인력을 더 효율적으로 사용할 수 있을 것이다.

제안한 자동모니터링 시스템은 원본과 전처리 과정을 거친 데이터를 사용했고, 실험 결과는 분류성능평가지표에 따라 수치화하여 비교했다. 학습 결과를 구할 수 없는 ResNet을 제외한 결과 f1-score 기준 Histogram image로 학습한 결과가 Capture image와 2D-FFT image로 학습한 결과보다 AlexNet은 0.17과 0.45, VGGNet은

0.11과 0.18 더 향상됐고, AlexNet보다 VGGNet이 0.01 더 향상됐다. 또, 활성화 함수로 ReLU를 사용하는 것보다 Leaky ReLU의 알파 계수를 0.01로 적용했을 때 성능이 0.08 향상됐다. 다시 말해, 원본을 캡처한 것이나 2D-FFT로 전처리하는 것보다 히스토그램으로 전처리한 이미지를 데이터로 사용하는 방법이 더 정확하게 이상 여부를 탐지했다. 총 244장의 적은 데이터를 사용했지만, 소규모 데이터에서 효과가 높은 Leaky ReLU를 활용하여, f1-score가 97%를 달성했기 때문에 충분히 현장에 적용할 수 있었다. 하지만, 이 시스템은 카메라 관리자가 이상 여부를 빠르게 알 수 있도록 탐지해주는 것이기 때문에 관리자가 한 명도 없는 산업 현장에는 적용하기 어려울 수 있다. 또, 현재 소규모의 데이터로 실험을 진행했기 때문에 대규모의 데이터로 실험을 진행했을 때의 결과와 현재 결과를 비교할 필요가 있다. 따라서 지속해서 데이터를 추가 확보하여 모델의 성능을 더욱 개선할 예정이다.

REFERENCES

- [1] Yusuke Ishikawa (FLIR Systems Japan), and jonghun kim, "Inspection technology using thermal image (building equipment, building envelope)," *equipment journal*, Vol. 49, No. 2, pp. 72-75, Feb. 2020.
- [2] Anjali Malviya, and S. G. Bhirud, "Visual Infrared Video Fusion for Night Vision using Background Estimation," *Journal of Computing*, Vol. 2, No. 4, pp. 66-69, April 2010. DOI: <https://doi.org/10.48550/arXiv.1004.4459>
- [3] Kyu-Seok Lee (Hanwha 63 City 63 Operations Team), all about infrared thermal imaging cameras!, <https://blog.63realty.co.kr/429>.
- [4] Kim Hyung Gyun, Jang Min Seok, and Lee Yon Sik, "A Black Ice Detection Method Using Infrared Camera and YOLO," *Journal of the Korea Institute of Information and Communication Engineering*, Vol. 25, No. 12, pp. 1874-1881, 31 Dec. 2021. DOI: <https://doi.org/10.6109/jkiice.2021.25.12.1874>
- [5] Jaehak Park, Kyungsu Park, Sanghyun Ji, Inkwon Yeo, and Hojin Cho, Study on Detection of Crack on Concrete Track Using Thermal Imaging Camera on Drone, Spring Conference of the Korean Society for Railway, pp. 428-429, Yeosu Expo Convention Center, South Korea, May 2019.
- [6] Ju-Yeong Jung, Hyuk-Jin Yoon, Hyun-Woo Cho, and Hui-Seung Yang, "A Study on Temperature Characteristics of Various Depth using Infrared Thermography," *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 19, No. 3, pp. 83-89, 31 March 2018. DOI: <https://doi.org/10.5762/KAIS.2018.19.3.83>

- [7] Min Song Ki, and Yeong Woo Choi, "Face Identification Using a Near-Infrared Camera in a Nonrestrictive In-Vehicle Environment," *KIPS Transactions on Software and Data Engineering*, Vol. 10, No. 3, pp. 99-108, 31 March 2021. DOI:<https://doi.org/10.3745/KTSDE.2021.10.3.99>
- [8] Che-wung Ha, and Do-Hwan Lee, "Diagnosis of Medium Voltage Cables for Nuclear Power Plant," *Journal of Electrical Engineering and Technology*, Vol. 9, No. 4, pp. 1369-1374, 01 July 2014. DOI:<https://doi.org/10.5370/JEET.2014s.9.4.1369>
- [9] Suk-Seung Hwang, and Young-Chul Bae, "Diagnosis of power supply using time-series of infrared camera," *The Journal of the Korea institute of electronic communication sciences*, Vol. 7, No. 6, pp. 1443-1447, 31 Dec. 2012. DOI:<https://doi.org/10.13067/JKIECS.2012.7.6.1443>
- [10] Hyeon-Seok O, and Byeong-Dong Yun, "Application of deep learning for failure prediction and soundness management of engineering systems," *Journal of the KSME*, Vol. 57, No. 3, pp. 34-37, 1 March 2017.
- [11] Dong-jun Lee, Seung-je Jeon, and Dong-hwi Lee, CNN model transition learning comparative analysis based on deep learning for image classification, *Proceedings of the Korean Institute of Information and Commucation Sciences Conference*, Busan Haendae Central Hotel, Korea, pp. 370-373, 26 May 2022.
- [12] KyeongHo Jo, SeHoon Jung, and ChunBo Sim, "A Study of Shiitake Disease and Pest Image Analysis based on Deep Learning," *Journal of Korea Multimedia Society*, Vol. 23, No. 1, pp. 50-57, 31 Jan. 2020. DOI:<https://doi.org/10.9717/kmms.2020.23.1.050>
- [13] Su-min Lee, Yu-hyeon Lee, Eun-sol Lee, and Se-yun Han, Comparison of CNN-based models for apple pest classification, *Proceedings of the Korea Information Processing Society Conference*, pp. 460-463, Sookmyung Women's University, Korea, 17 May 2022. DOI:<https://doi.org/10.3745/PKIPS.y2022.m05a.460>
- [14] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional," *Communications of the ACM*, Vol. 60, No. 6, pp. 84-90, June 2017. DOI:<https://doi.org/10.1145/3065386>
- [15] Karen Simonyan, and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv:1409.1556*, Sep. 2014. DOI:<https://doi.org/10.48550/arXiv.1409.1556>
- [16] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep Residual Learning for Image Recognition," *arXiv:1512.03385*, 10 Dec. 2015. DOI:<https://doi.org/10.48550/arXiv.1512.03385>
- [17] George Philipp, Dawn Song, and Jaime G. Carbonell, "The exploding gradient problem demystified - definition, prevalence, impact, origin, tradeoffs, and solutions," *arXiv:1712.05577*, 6 Apr. 2018. DOI:<https://doi.org/10.48550/arXiv.1712.05577>
- [18] Mathworks, histogram - histogram of vector, <https://kr.mathworks.com/help/matlab/ref/matlab.graphics.chart.primitive.histogram.html>.
- [19] Joshua Peeples, Weihuang Xu, and Alina Zare, "Histogram Layers for Texture Analysis," *IEEE Transactions on Artificial Intelligence*, Vol. 3, No. 4, pp. 542-552, 15 Dec. 2021. DOI:10.1109/TAI.2021.3135804
- [20] Mi Jin Jun, and Joon Jae Lee, "Image Histogram Equalization Based on Gaussian Mixture Model," *Journal of Korea Multimedia Society*, Vol. 15, No. 6, pp. 748-760, 30 June 2012. DOI:<https://doi.org/10.9717/kmms.2012.15.6.748>
- [21] Chang-il Kim, Joon-ji Kim, and Seung-hwa Paek, A Method for Object Extraction of SAR Image using Sub-Histogram Technique based on Feature Point, *Proceedings of the Korea Information Processing Society Conference*, pp. 1442-1445, Jeju Halla University, Korea, 28 Oct. 2015. DOI:<https://doi.org/10.3745/PKIPS.y2015m10a.1442>
- [22] Atin Mukherjee, and Debesh Choudhury, "An Area Efficient 2D Fourier Transform Architecture for FPGA," *arXiv:1810.06885*, 16 Oct. 2018. DOI:<https://doi.org/10.48550/arXiv.1810.06885>
- [23] Rafael C. Gonzalez, and Richard E. Woods, "Digital Image Processing 3rd Edition," Pearson, pp. 244-257, 31 Aug. 2007.
- [24] Seung-Soo Shin, Sang-Bock Lee, and Yong-Hwan Cho, "Recognition of Disease in Medical Image," *The Journal of the Korea Contents Association*, Vol. 1, No. 1, pp. 8-14, 01 Dec. 2001.
- [25] David Lovell, Jaiden Capra, and Andrew Bradley, "never mind the metrics - what about the uncertainty," *arXiv:2206.02157*, 5 Jun 2022. DOI:<https://doi.org/10.48550/arXiv.2206.02157>
- [26] Sergey Ioffe, and Christian Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," *arXiv:1502.03167*, 2 Mar 2015. DOI:<https://doi.org/10.48550/arXiv.1502.03167>
- [27] Sung-Wook Park, and Do-Yeon Kim, "Comparison of Image Classification Performance by Activation Functions in Convolutional Neural Networks," *Journal of Korea Multimedia Society*, Vol. 21, No. 10, pp. 1142-1149, 31 Oct. 2018. DOI:<https://doi.org/10.9717/kmms.2018.21.10.1142>

Authors



Hye-Won Kim received the M.S. degrees in Computer Engineering from Deajeon University in 2022. Kim has been working at Epozen's research institute since 2022. She is interested in artificial intelligence, cloud computing.



Sang-Min Kim received the M.S. degrees in Electric Engineering from Kwangwoon Technical College in 2022. Kim has been working at Epozen's research institute in 2022.

He is interested in artificial intelligence, signal processing.



Jung-Mo Sohn received the B.S. and M.S. degrees in Industrial Automation from Inha University, Korea, in 1998 and 2000, respectively. Sohn has been working at Epozen's research institute since 2018.

He is interested in artificial intelligence, cloud computing, and information security.