

ARTICLE

환경별 통합에 따른 딥러닝 기반 흥인지문 지붕부 기울음 예측 실험 및 분석

이상윤* · 강석주**

Experiments and Analyses of Inclination Prediction for the Heunginjimun Roof based on Deep Learning according to Environmental Integration

Lee, Sang-Yun* · Kang, Seok-Ju**

요 지

문화재의 변위를 탐지하는 것은 문화재를 보호 및 관리하기 위해 반드시 필요한 일이나 사람이 육안으로 직접 탐지할 수 있는 변위는 제한적이고 많은 시간과 인력이 요구된다. 따라서 딥러닝 기술을 활용해 문화재의 변위 및 손상을 자동으로 탐지할 수 있는 다양한 연구들이 진행되고 있다. 이전 연구에서 흥인지문 지붕부의 변위 탐지를 위해 이미지 처리에 탁월한 CNN과 전이 학습 기법을 통해 사전에 학습된 딥러닝 신경망 모델 10개를 활용하여 흥인지문의 이미지를 정상과 비정상 두 가지로 분류되도록 학습시킨 후 계절과 날씨에 따라 분류한 9개의 환경에서 생성된 모델의 예측 정확도를 향상시키는 연구를 진행한 바 있다. 그 결과 정확도가 95% 미만으로 나오는 모델들을 대상으로 하이퍼 파라미터 조정을 통해 95% 이상으로 향상시킬 수 있었다. 본 연구에서는 이전 연구에서 사용했던 10개의 사전에 학습된 딥러닝 신경망 모델 중 성능 검증을 통해 정확도가 가장 높은 신경망 모델을 최적의 모델로 선정하고 그 신경망 모델을 활용하여 9개의 단일환경을 여러 개로 통합하여 학습을 진행하는 환경통합 실험을 통해 환경별 통합에 따른 예측 정확도를 측정하고 결과를 분석한다. 신경망 모델 성능 검증 결과 정확도가 가장 높은 Efficientnet_v2_s 모델을 활용해 환경통합 실험을 진행했다. 그 결과 9개의 단일 환경을 모두 통합한 환경에서 학습을 진행한 후 성능을 검증한 결과 정확도가 99.84%로 나왔으며 모든 환경에서 성공적으로 지붕부의 변위를 탐지할 수 있음을 실험을 통해 알아낼 수 있었다.

핵심용어: 변위 탐지, 딥러닝, CNN, 전이학습, 흥인지문

* 교신저자, 정회원, 한국전자통신연구원 디지털융합연구소 과학치안공공ICT연구센터 책임연구원(E-mail: syllee@etri.re.kr)
Corresponding Author, Member, Principal Researcher, Police Science & Public Safety ICT Research Center, Digital Convergence Research Laboratory, Electronics and Telecommunications Research Institute

** 충남대학교 전자정보통신공학과 학사과정 4학년(E-mail: tnswnd35@naver.com)
BS, Department of Radio Information and Communication Engineering, Chungnam National University

Received | September 1, 2023 Revised | October 30, 2023 Accepted | November 20, 2023



Abstract

Detecting the displacement of cultural heritage is essential for protecting and managing their prosperity. However, because the degree of displacement that humans can detect with the naked eye is limited and requires a lot of time and manpower, various studies are being conducted to detect the displacement and damage of cultural heritages using Deep Learning technology. In previous studies, 10 pre-trained Deep Learning neural network models through transfer learning techniques and Convolutional Neural Network (CNN) which are known for excellent for image processing were used to detect inclination in the roof of Heunginjimun. After learning to classify the images of Heunginjimun into two categories, normal and abnormal, we measured the prediction accuracy of the models generated in nine environments, classified according to season and weather. After analysis, it was raised to over 95% through hyper-parameter adjustment. In this study, among the 10 pre-trained Deep Learning Neural Network models used in previous studies, the Neural Network model with the highest accuracy was selected as the optimal model through performance verification, and then it was used to integrate 9 single environments into multiple ones. As a result, by adjusting the hyper-parameter of models below 95% in prediction accuracy, we were able to raise it above the previous percentage. Consequently, we selected the Efficientnet_v2_s model, which had the highest accuracy, to conduct an environment integration experiment. As a result, after running the learning and verifying performance in an environment that integrated all nine single environments, the roof inclination was successfully detected in all environments with an accuracy of 99.84%.

Keywords: Detecting displacement, Deep learning, Convolutional neural network, Transfer learning, Heunginjimun

1. 서론

문화재는 일반건축물과는 달리 재화로 따질 수 없을 만큼 중요한 문화유산이고 외부에 노출될 경우 환경에 의한 구조적 변위, 열화, 균열 등이 발생하기 쉬우므로 주기적으로 관리 및 유지보수가 필요하다. 그러나 사람이 육안으로 탐지할 수 있는 변위에는 한계가 있고 많은 시간과 인력이 요구되기 때문에 기존의 방법을 대체하기 위해 딥러닝 기술에 관한 많은 연구가 진행되고 있다(Park and Lee, 2022; Lee and Park, 2022; Lee et al., 2023; Kang and Lee, 2023; Wang et al., 2020).

Lee et al.(2023)는 홍인지문 지붕부의 변위를 탐지하기 위해 이미지 처리에 탁월한 CNN(Convolutional Neural Network, 합성곱 신경망)(LeCun et al., 1989) 모델을 활용했다. 홍인지문의 CCTV 영상을 계절, 날씨에 따라 9개의 환경으로 분류한 후 영상을 일정한 간격으로 나누어 정상 이미지를 생성하고 비정상 이미지를 생성하기 위해 포토샵 액션스크립트 기능을 이용하여 홍인지문 정상 이미지의 지붕부를 왜곡(뒤틀기)시켜 기울음을 표현했다.

선행 연구(Kang and Lee, 2023)에서는 생성된 이미지들을 전이 학습(Transfer learning) 기법을 통해 1,000개의 클래스 및 1,400만 개의 이미지로 구성된 이미지넷(ImageNet) 데이터셋을 사용한 사전에 학습된(Pre-trained) 딥러닝 신경망 모델을 활용하여 정상과 비정상 두 가지 클래스로 분류할 수 있도록 학습시킨 후 예측 정확도를 측정하였다. 그 결과 예측 정확도가 95% 미만으로 나온 모델의 정확도 그래프를 분석한 후 하이퍼 파라미터 조정을 통해 예측 정확도를 95% 이상으로 올릴 수 있었다.

Wang et al.(2020)은 중국 자금성 지붕의 타일을 딥러닝 모델 Faster R-CNN과 Mask R-CNN을 기반으로 손상을 탐지하고 분류했다. 먼저 Faster R-CNN을 통해 지붕 이미지에서 타일을 감지하고 잘라 두 가지 유형으로 분류한 다음 분류된 타일 이미지를 Mask R-CNN을 통해 타일의 손상 유형과 면적 및 비율을 측정하는 2단계의 절차를 통해 건물 지붕부 타일의 표면 손상을 자동으로 감지하고 분류했다.

본 연구에서는 9개로 분류된 단일환경에서 모델을 학습했던 기존 실험과 달리 환경을 다양한 조합으로 통합하여 학습을 진행한 후 예측 정확도를 측정하고 그 결과를 분석하는 환경통합 실험을 진행한다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 실험 대상인 홍인지문과 지붕부 변위 판정 기술 및 모델을 소개하고 데이터셋 구축 과정을 설명한다. 3장에서는 최적의 신경망 모델을 선정하기 위해 선행 연구(Kang and Lee, 2023)에서 생

성된 모델들의 성능을 측정하고 비교한다. 4장에서는 환경통합 실험에서 최적의 신경망 모델로 선정된 Efficientnet_v2_s 를 사용한다. 단일환경을 여러 개로 통합하여 모델을 학습한 후 예측 정확도를 측정하고 과정과 결과를 정확도 그래프, 손실 그래프 그리고 Grad-CAM 결과를 통해 제시하고 통합한 단일환경의 개수 및 구성과 예측 정확도의 관계를 분석한다. 마지막으로 5장에서는 연구 결과를 정리 및 분석하고 향후 연구 방향을 제시한다.

2. 흥인지문 지붕부 기울음 변위 판정 방법론

2.1 실험 대상 소개

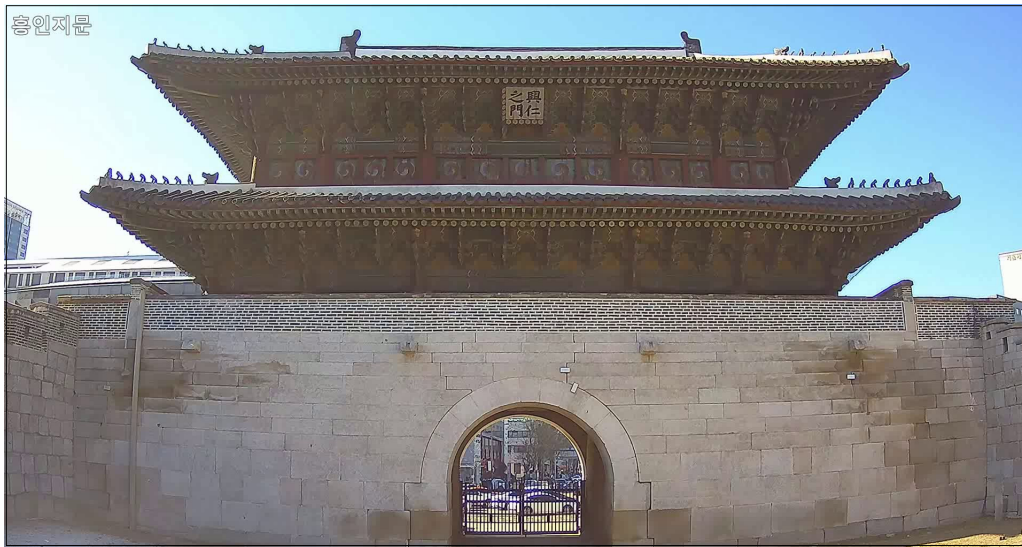


Fig. 1. Heunginjimun

본 연구는 서울특별시 종로구에 위치한 흥인지문(Fig. 1)의 지붕부를 대상으로 한다. 흥인지문은 대한민국 보물 제1호로 지정되었으며 역사적 가치와 중요성이 매우 높기 때문에 자연재해 및 인적 요인에 의한 손상을 예방하고 노후화되지 않도록 지속적인 관리 및 보호가 필요하다.

서울 성곽은 옛날 중요한 국가시설이 있는 한성부를 보호하기 위해 만든 도성으로, 흥인지문은 성곽 8개의 문 가운데 동쪽에 있는 문이다. 앞면 5칸·옆면 2칸 규모의 2층 건물로, 다포 양식의 형태가 가늘고 약하며 지나치게 장식한 부분이 많아 조선 후기의 특징을 잘 나타내주고 있다. 하나의 건조물 안에서 기와, 목재, 석재 등 다양한 재질로 이루어져 여러 가지 손상 유형을 파악하기 좋다. 야외 문화재는 대상지 주변의 환경 영향에 취약하며, 여러 손상 요인들이 복합적으로 작용하여 문화재의 보존 상태에 큰 영향을 미치기 때문에 연구 자료로 의미를 지닌다.

특히, 흥인지문은 도심지에 위치하여 물리적·인위적 손상의 위험이 큰 건조물이기 때문에 대상으로 선정하였다. 그런데, 사람이 직접 육안으로 탐지할 수 있는 변위에는 한계가 있고 많은 시간과 인력이 요구된다. 2011년 8월에는 폭우로 인해 흥인지문의 지붕부가 파손되는 일이 발생했다(Fig. 2). 당시 파손 사실을 발견한 것은 순찰 인력이 아닌 시민이었으며 관리 인력의 육안 검사에 의존하는 문화재 방재 방식의 한계점이 드러난 사건이었다(Kwon, 2011).

본 연구에서는 기존 방법의 문제점을 해결하고 대체하기 위해 딥러닝 기술을 기반으로 흥인지문 지붕부의 기울음과 손상을 탐지하는 방법을 제안한다.



Fig. 2. Heunginjimun roof damaged by heavy rain (Kwon, 2011)

2.2 알고리즘 및 모델 소개

본 연구에서는 홍인지문 이미지 데이터셋에서 비정상 이미지를 분류하기 위해 Fig. 3과 같이 Lee et al.(2023)가 제안한 딥러닝 알고리즘을 사용한다. 이 알고리즘은 합성곱층(Convolutional layer)과 풀링층(Pooling layer)으로 구성된 인공 신경망 알고리즘으로서 이미지 데이터에서 객체를 탐색하거나 객체의 위치를 찾아내는 데 유용한 합성곱 신경망으로 알려져 있다.

또한, 1,000개의 클래스 및 1,400만 개의 이미지로 구성된 이미지넷 데이터셋을 사용하여 전이 학습을 통해 10개의 사전에 학습된 딥러닝 신경망 모델을 생성하였으며, 정상 이미지와 비정상 이미지 2개의 클래스로 분류할 수 있도록 학습시킨 결과 학습 속도가 빠르고 정확도가 우수하게 나오는 것을 확인할 수 있었다. 연구에 사용된 10개의 신경망 모델은 EfficientnetB0(Tan and Le, 2019), Mobilenet_v2(Sandler et al., 2018), Resnet18(He et al., 2016), EfficientnetB2(Tan and Le, 2019), Shufflenet_v2(Ma et al., 2018), Alexnet(Krizhevsky et al., 2017), Mnasnet(Tan et al., 2019), Inception_v3(Szegedy et al., 2016), Densenet161(Huang et al., 2017)), Efficientnet_v2_s(Tan and Le, 2021)이며 이미지 분류나 객체 검출을 위해 널리 사용되는 모델들이다.

2.3 데이터셋

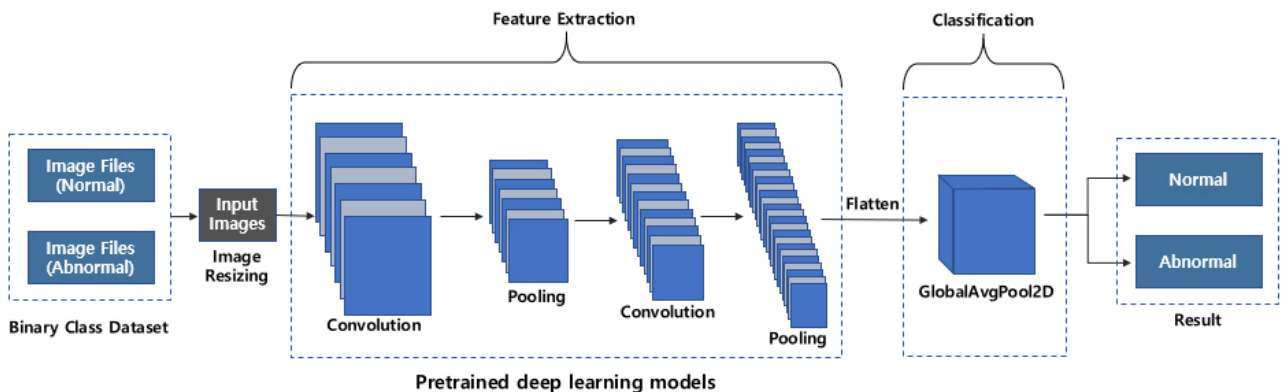


Fig. 3. Deep learning model architecture

2.3.1 데이터셋 구축 과정

Fig. 4는 흥인지문 정상 이미지와 비정상 이미지의 데이터셋을 구축하는 과정을 보여준다. 구축에 활용된 데이터는 2022년 1월 20일부터 9월 13일까지 촬영한 흥인지문 CCTV 영상을 대상으로 하였다. 계절과 날씨에 따라 9개로 영상을 분류한 후 영상에서 일정한 시간 간격으로 프레임을 추출하여 정상 이미지를 구축하였고 정상 이미지에서 포토샵 (Photoshop)의 왜곡(뒤틀기) 기능을 활용하여 비정상 이미지를 생성하였다.

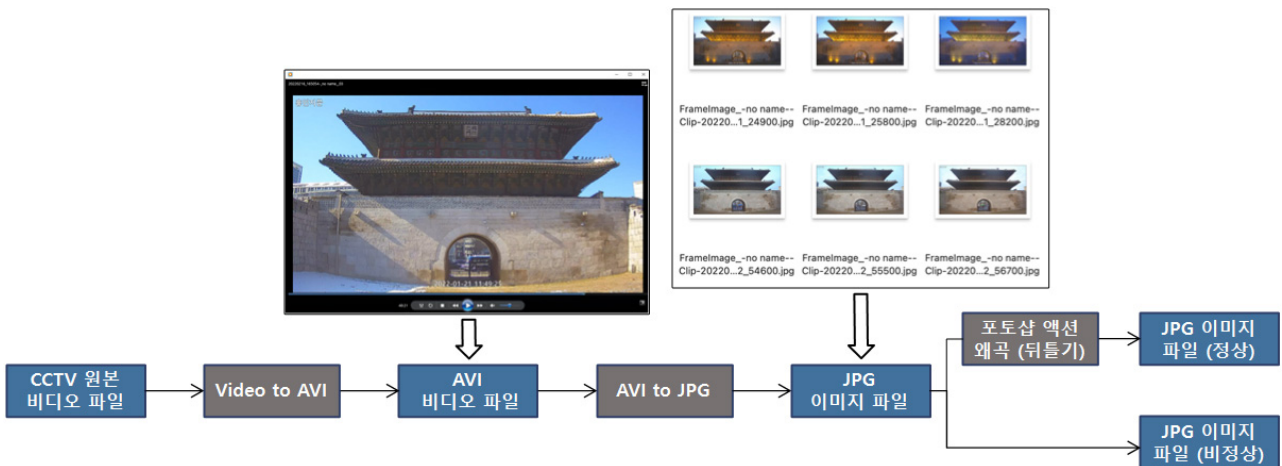


Fig. 4. Data building process

2.3.2 비정상 이미지 생성 과정

정상 이미지를 대상으로 흥인지문 지붕부의 위치에 따라 왜곡을 가하는 위치는 총 10곳으로 Fig. 5와 같다. 상단부에 5곳 하단부에 5곳이 적절한 구역으로 나뉘어져 있음을 볼 수 있다.

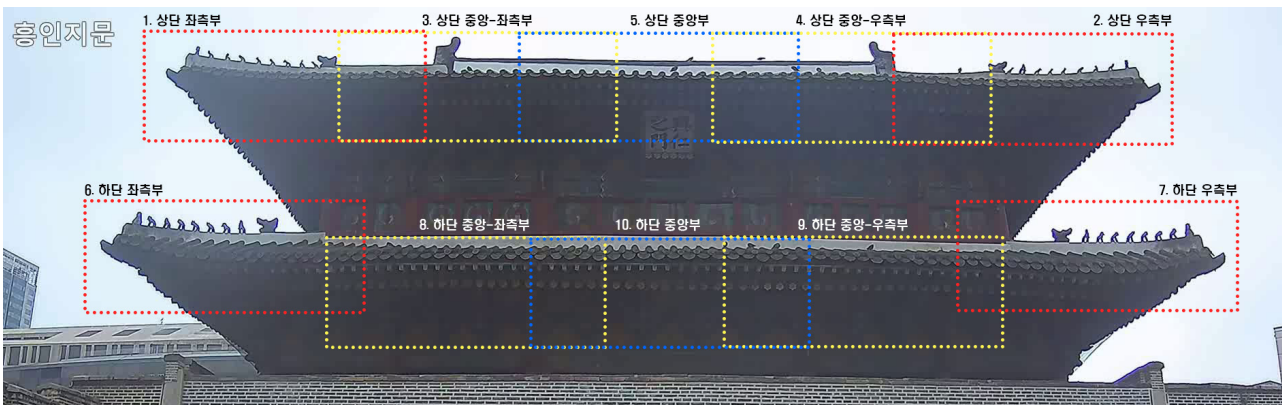


Fig. 5. Heunginjimun roof distort location

Fig. 6은 정상 이미지와 비정상 이미지를 비교한 것으로 상단의 사진이 정상 이미지, 하단의 사진이 비정상 이미지이다. 비정상 이미지는 왜곡을 가했기 때문에 정상 이미지와 비교했을 때 지붕부가 기울거나 뒤틀어진 모습을 볼 수 있다.

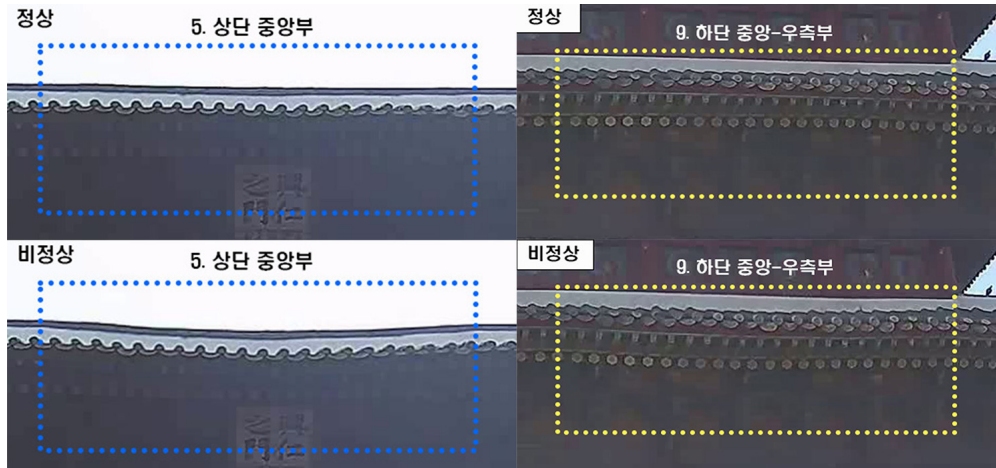


Fig. 6. Normal and abnormal images

2.3.3 단일환경 데이터셋 구성

분류한 환경은 겨울/야간, 겨울/주간, 봄/야간, 봄/주간, 맑음/야간, 흐림/주간, 비/주간, 비/야간, 눈/야간으로 순서대로 ENV-01~ENV-09로 지정했다. 환경마다 정상 이미지와 비정상 이미지를 합하여 총 10,000장으로 구성하였으며 훈련셋 (Train)/검증셋 (Validation)/테스트셋 (Test)의 비율은 50:25:25로 구성하였으며 자세한 내용은 Table 1과 같다.

Table 1. Dataset configuration

데이터셋	Train		Validaion		Test	
	정상	비정상	정상	비정상	정상	비정상
ENV-01 (겨울/야간)	4500	500	2250	250	2250	250
ENV-02 (겨울/주간)	4000	1000	2000	500	2000	500
ENV-03 (봄/야간)	2500	2500	1250	1250	1250	1250
ENV-04 (봄/주간)	2500	2500	1250	1250	1250	1250
ENV-05 (맑음/야간)	2500	2500	1250	1250	1250	1250
ENV-06 (흐림/주간)	2500	2500	1250	1250	1250	1250
ENV-07 (비/주간)	2500	2500	1250	1250	1250	1250
ENV-08 (비/야간)	2500	2500	1250	1250	1250	1250
ENV-09 (눈/야간)	2500	2500	1250	1250	1250	1250

2.3.4 통합환경 데이터셋 구성

본 연구에서는 Table 1의 ENV-01~ENV-09를 통합한 환경의 개수 및 구성과 예측 정확도의 관계를 분석하기 위해 단일환경을 여러 개로 통합하여 순서대로 ENV-10~ENV-22로 지정했다. 자세한 구성은 Table 2와 같다.

Table 2. Intergrated environmnet dataset configuration

데이터셋	데이터셋 구성	
	통합한 환경 개수	환경 구성
ENV-10	2	봄/야간 + 봄/주간
ENV-11	2	맑음/야간 + 흐림/주간
ENV-12	2	비/주간 + 눈/야간
ENV-13	3	봄/야간 + 봄/주간 + 맑음/야간
ENV-14	3	비/주간 + 비/야간 + 눈/야간
ENV-15	3	맑음/야간 + 비/주간 + 눈/야간
ENV-16	4	겨울/야간 + 겨울/주간 + 봄/야간 + 봄/주간
ENV-17	4	흐림/주간 + 비/주간 + 비/야간 + 눈/야간
ENV-18	4	겨울/주간 + 봄/주간 + 흐림/주간 + 비/주간
ENV-19	5	봄/야간 + 봄/주간 + 비/주간 + 비/야간 + 눈/야간
ENV-20	5	겨울/야간 + 봄/야간 + 맑음/야간 + 비/야간 + 눈/야간
ENV-21	5	맑음/야간 + 흐림/주간 + 비/주간 + 비/야간 + 눈/야간
ENV-22	9	모든 환경 통합

3. 신경망 모델 성능 검증

3.1 신경망 모델 성능 검증 방법론

연구에 사용된 10개의 신경망 모델 중 최적의 모델을 선정하기 위해 모델들의 성능을 검증하였다. 평가 지표로는 정확도(Accuracy)를 사용하였으며 선행 연구(Kang and Lee, 2023) 결과물을 기반으로 하이퍼 파라미터 조정을 통해 학습 환경과 동일한 환경에서 예측 정확도를 측정했을 때 정확도가 95% 이상으로 나온 모델들을 9개의 단일 환경 중 학습 환경을 제외한 나머지 8개 환경에서 예측 정확도를 측정하여 가장 정확도가 높은 모델을 최적의 모델로 선정하였다.

3.2 신경망 모델 선정

Fig. 7은 10개의 신경망 모델의 정확도를 내림차순으로 정리한 그래프이다. Efficientnet_v2_s가 64.39%로 가장 높았고 다음으로 Densnet161과 Mnasnet이 각각 62.36%, 62.33%의 정확도를 보였으며 EfficientnetB2가 56.04%로 가장 낮았으며 결과를 통해 10개의 신경망 모델 중 정확도가 가장 높은 Efficientnet_v2_s 모델을 최적의 모델로 선정했다.

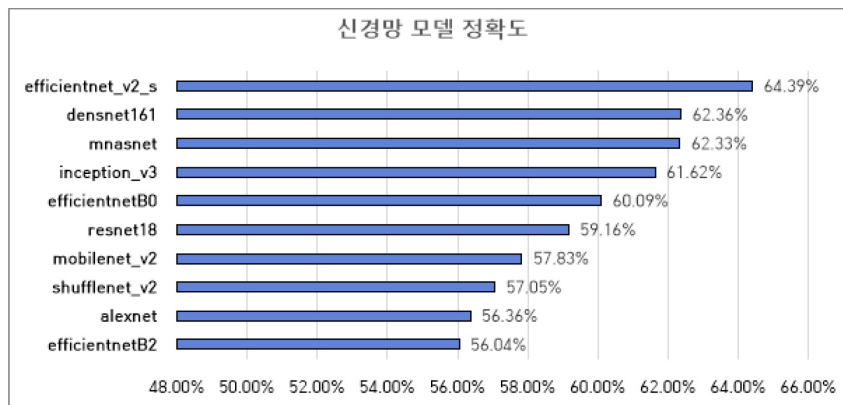


Fig. 7. Neural network model average accuracy

3.3 학습 환경에 따른 예측 정확도 분석

Fig. 8은 학습 환경에 따른 평균 정확도를 내림차순으로 정렬해 그래프로 나타낸 것이다. 봄/주간 환경에서 학습했을 때 평균 정확도가 73.22%로 가장 높았고 눈/야간 환경에서 학습했을 때 50.78%로 가장 낮았으며 내림차순으로 정렬한 결과를 보면 9개의 환경 중 평균 정확도 상위 4개의 그룹은 모두 주간 환경이고 나머지 5개의 하위 그룹은 모두 야간 환경인 것을 알 수 있다.

주간 환경 데이터셋의 이미지는 컬러가 들어간 이미지고 야간 환경은 흑백 이미지임을 감안하면 흑백 이미지보다 컬러가 들어간 이미지를 학습에 사용하는 것이 이미지를 분류하는 데 있어 더 많은 정보를 제공하기 때문에 더 높은 정확도를 보여준 것으로 해석된다.

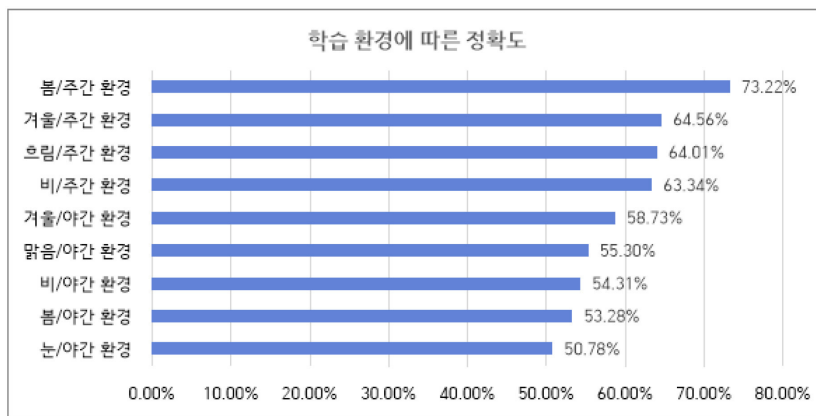


Fig. 8. Average accuracy by learning environments

4. 환경통합 실험 결과

4.1 모델 하이퍼 파라미터

Table 3은 실험에 사용된 신경망 모델의 하이퍼 파라미터를 나타낸 것으로 옵티마이저로 SGD(Stochastic Gradient Descent) 모멘텀을 사용했다. SGD 모멘텀은 경사 하강법의 일종으로 기본적인 경사 하강법에 관성의 개념을 추가하여 이전 단계에서의 업데이트 값을 일정 비율로 고려하기 때문에 수렴 속도를 향상하고, 지역 최적점(Local minima)에 빠지는 것을 줄여주는 효과가 있다(Ruder, 2016).

신경망 모델은 최적의 모델로 선정된 Efficientnet_v2_s를 사용하였으며, EfficientnetB0, EfficientnetB2 모델을 추가로 사용해 정확도를 비교하여 성능을 검증하였다. 이 3개의 모델들은 Efficientnet 모델의 변형들로 같은 구조를 가지지만 크기와 복잡성에 차이가 있으며 Efficientnet_v2_s는 높은 효율성과 성능을 제공하도록 설계된 비교적 최신 모델이다 (Tan and Lee, 2019; Tan and Lee, 2021).

Table 3. Model hyper parameters

Model	Batch size	Input size	Optimizer	Epochs
efficientnetB0	8	224	SDG (lr = 0.001, momentum = 0.9)	10
efficientnetB2		260		
efficientnet_v2_s		384		

4.2 예측 정확도 측정

Table 2의 통합환경 13개와 Table 3의 신경망 모델 3개를 사용해 총 39개의 모델을 생성하고 예측 정확도를 측정했다. Fig. 9는 ENV-10 데이터셋에서 Efficientnet_v2_s 신경망 모델을 사용해 예측 정확도를 측정했을 때 정확도 그래프와 손실 그래프로 정확도 그래프는 100으로, 손실 그래프는 0으로 수렴하는 것으로 보아 학습이 원활하게 진행되었음을 알 수 있다.

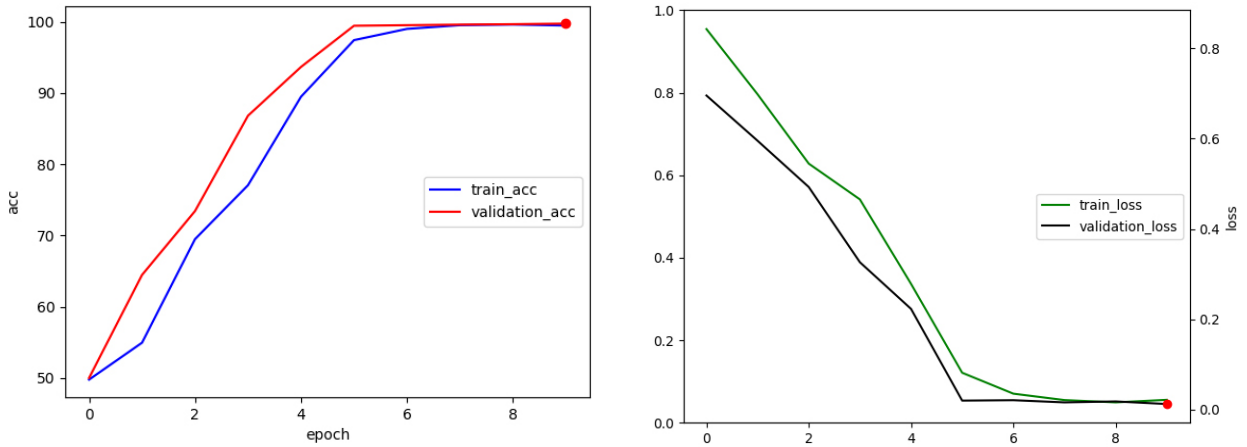


Fig. 9. Accuracy graph and loss graph of T-0093

Fig. 10은 ENV-17 데이터셋에서 EfficientnetB0 신경망 모델을 사용해 예측 정확도를 측정했을 때 정확도 그래프와 손실 그래프로 마찬가지로 그래프가 수렴하는 것으로 보아 학습이 원활하게 진행되었음을 알 수 있다.

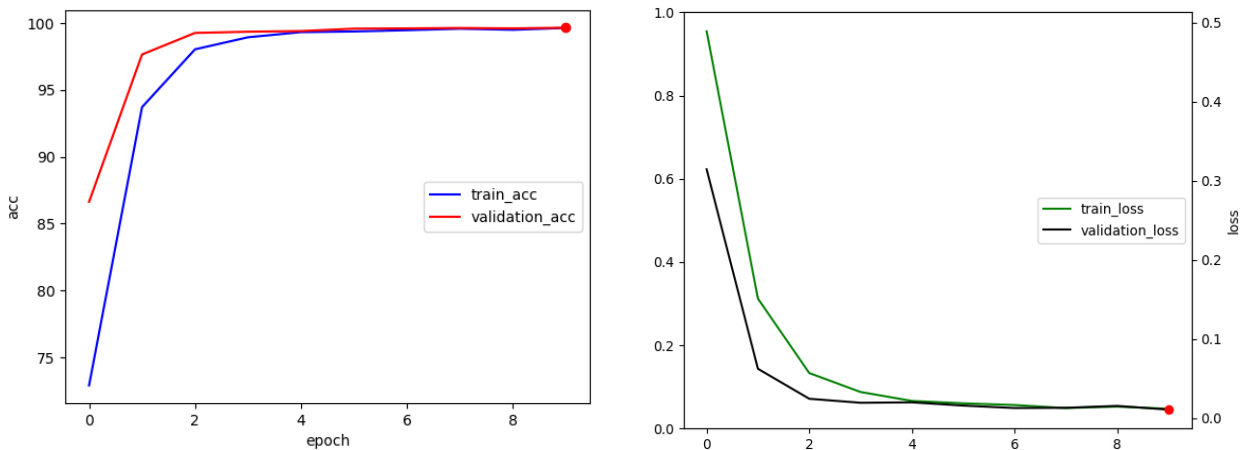


Fig. 10. Accuracy graph and loss graph of T-0112

4.3 환경통합 실험 성능 검증

3.1 신경망 모델 성능 검증 방법론과 동일하게 4.2에서 생성된 모델들을 9개의 단일 환경에서 예측 정확도를 측정한 결과를 Table 4에 순서 및 통합한 환경 개수에 따라 정리하고 Fig. 11에 그래프로 표현했다. Table 4와 Fig. 11에서 정확도는 9개의 단일 환경에서 예측 정확도를 측정하여 평균을 계산한 것이다.

통합한 단일 환경 개수가 2개, 3개, 4개일 때 각각 정확도가 79.89%, 83.96%, 93.20%로 개수가 많아질수록 정확도가

증가하는 모습을 보였지만 개수가 5개일 때 정확도는 88.58%로 개수가 4개일 때에 비해 감소했다. 모든 환경을 통합한 ENV-22는 정확도가 99.84%로 가장 높았고 ENV-0022를 제외하면 통합한 단일 환경 개수가 4개인 ENV-16이 97.60%로 가장 높았다.

2.3.2 학습 환경에 따른 예측 정확도 분석에서도 주간 환경에서 학습했을 때 더 높은 정확도를 보였다. 환경 통합 실험에서도 주간 환경으로만 구성된 ENV-18은 정확도가 92.80%이고 야간 환경으로만 구성된 ENV-20는 정확도가 86.54%로 주간 환경이 야간 환경보다 데이터가 더 부족했음에도 더 높은 정확도를 보여줬다.

Table 4. Accuracy of environment integration experiment

데이터셋	통합한 환경 개수	데이터셋 구성	정확도	평균
ENV-10	2	봄/야간 + 봄/주간	82.12%	79.89%
ENV-11	2	맑음/야간 + 흐림/주간	78.60%	
ENV-12	2	비/주간 + 눈/야간	78.96%	
ENV-13	3	봄/야간 + 봄/주간 + 맑음/야간	81.01%	83.96%
ENV-14	3	비/주간 + 비/야간 + 눈/야간	87.52%	
ENV-15	3	맑음/야간 + 비/주간 + 눈/야간	83.37%	
ENV-16	4	겨울/야간 + 겨울/주간 + 봄/야간 + 봄/주간	97.60%	93.20%
ENV-17	4	흐림/주간 + 비/주간 + 비/야간 + 눈/야간	89.21%	
ENV-18	4	겨울/주간 + 봄/주간 + 흐림/주간 + 비/주간	92.80%	
ENV-19	5	봄/야간 + 봄/주간 + 비/주간 + 비/야간 + 눈/야간	88.20%	88.58%
ENV-20	5	겨울/야간 + 봄/야간 + 맑음/야간 + 비/야간 + 눈/야간	86.54%	
ENV-21	5	맑음/야간 + 흐림/주간 + 비/주간 + 비/야간 + 눈/야간	91.01%	
ENV-22	9	모든 환경 통합	99.84%	

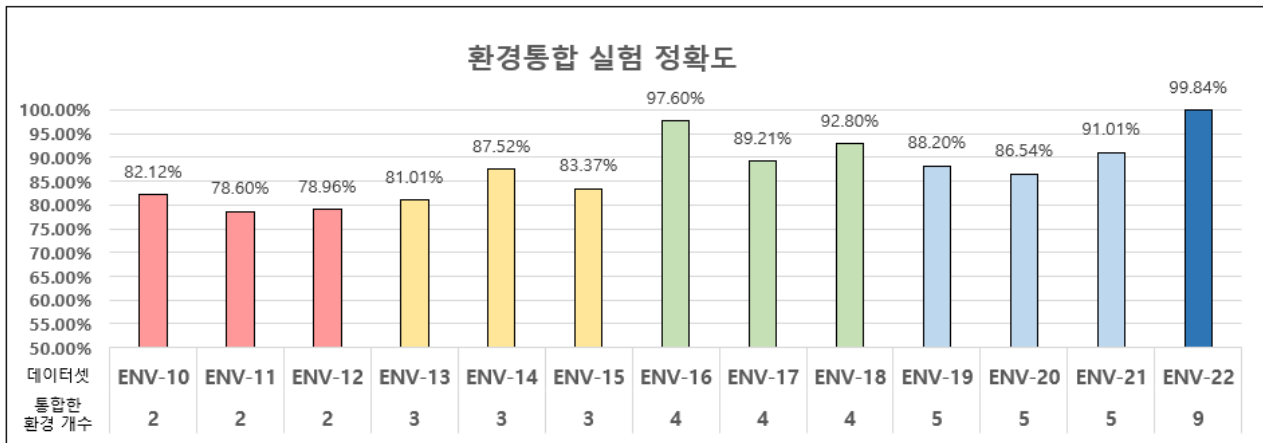


Fig. 11. Environment integration experimental dataset accuracy

4.4 신경망 모델 성능 검증

EfficientnetB0, EfficientnetB2 모델을 추가로 사용하여 모델들의 정확도를 서로 비교하며 성능을 검증하였다. 3개의 모델은 Efficientnet 모델의 변형으로 같은 구조를 가지지만 크기와 복잡성에 차이가 있으며 Efficientnet_v2_s는 높은 효율성과 성능을 제공하도록 설계된 비교적 최신 모델이다.

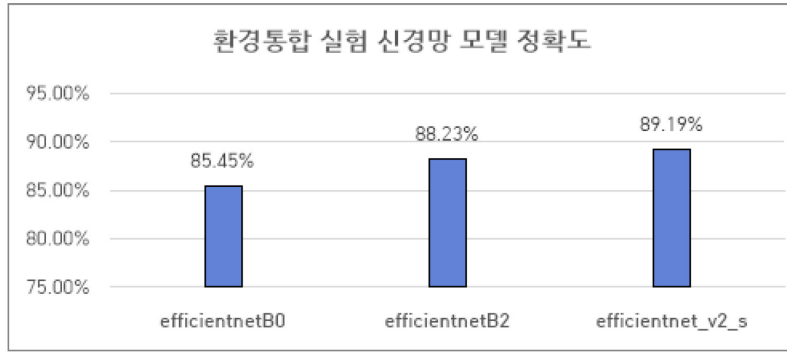


Fig. 12. Environment integration experimental neural network model accuracy

Fig. 12에 3개의 모델의 정확도를 측정된 결과를 그래프로 나타냈다. Efficientnet_v2_s 모델의 정확도는 89.19%로 가장 높은 정확도를 보였으며 사용된 데이터셋이 최소 20,000개에서 최대 90,000개인 환경통합 실험에서도 다른 모델들에 비해 성능이 우수함을 알 수 있다.

4.5 Grad-CAM 결과 분석

Grad-CAM을 통해 이미지 분류 모델로 사진을 분류할 때 사진의 어느 부분이 최종 분류 결정에 영향을 끼쳤는지를 시각화하여 확인할 수 있다.

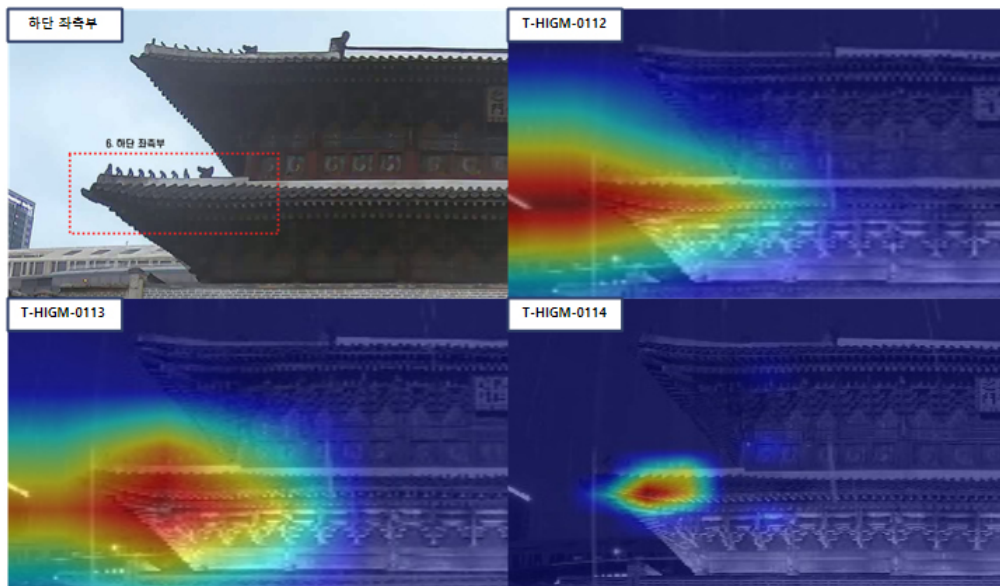


Fig. 13. Environment integration experiment Grad-CAM result

Fig. 13의 좌측 상단의 이미지는 흥인지문 지붕부의 하단 좌측 부에 왜곡(뒤틀기)한 이미지로 ENV-17 데이터셋의 비정상 이미지 중 하나이다. 나머지 이미지들은 각각 ENV-17 데이터셋에서 EfficientnetB0(T-HIGM-0112), EfficientnetB2(T-HIGM-0113), Efficientnet_v2_s(T-HIGM-0114) 신경망 모델을 사용해 예측 정확도를 측정했을 때 Grad-CAM heatmap과 원본 이미지를 합쳐서 나타낸 것으로 지붕부의 하단 좌측부에 왜곡을 가한 부위 주변으로 히트맵이 분포되고 있음을 확인할 수 있다.

이를 통해 예측 모델이 왜곡을 가한 부위를 정확히 지정하고 있음을 알 수 있으며 이를 통해 기울음이 발생한 지점도 예측 가능성을 알 수 있다(Selvaraju et al., 2017).

5. 결론

문화재의 변위를 탐지하는 것은 문화재를 보호 및 관리하기 위해 반드시 필요한 일이나 사람이 육안으로 직접 탐지할 수 있는 변위는 제한적이고 많은 시간과 인력이 요구된다.

본 논문에서는 홍인지문 지붕부를 대상으로 기존 방법의 문제점을 해결하고 대체하기 위해 딥러닝 기술을 기반으로 홍인지문 지붕부 기울음과 손상을 탐지하는 방법을 제안하였다. 즉, 10개의 딥러닝 신경망 모델 중 최적의 모델을 선정하기 위해 이전 연구에서 정확도를 95% 이상으로 향상시킨 모델들을 학습에 사용되지 않은 환경에서 테스트를 진행하여 예측 정확도를 측정 후 비교하였다.

측정한 결과 최적의 모델로 선정된 Efficientnet_v2_s 모델을 사용하여 서로 다른 9개의 환경을 통합하여 예측 정확도를 측정하는 환경 통합 실험을 진행했다.

통합한 단일 환경 개수에 따라 데이터셋의 정확도를 비교한 결과 통합한 환경 개수가 2개부터 4개까지는 많아질수록 정확도가 증가했지만 5개일 때는 오히려 감소하는 모습을 보였다.

모든 단일 환경(겨울/야간, 겨울/주간, 봄/야간, 봄/주간, 맑음/야간, 흐림/주간, 비/주간, 비/야간, 눈/야간)을 통합한 ENV-22가 99.84%의 정확도로 가장 높았고 ENV-22를 제외하면 4개의 단일 환경(겨울/야간, 겨울/주간, 봄/야간, 봄/주간)을 통합한 ENV-16이 97.60%로 가장 높았다.

EfficientnetB0, EfficientnetB2 모델을 추가로 사용하여 모델의 성능을 비교한 결과 Efficientnet_v2_s 모델이 가장 높은 정확도를 보였으며 사용된 데이터셋이 최소 20,000개에서 최대 90,000개인 환경통합 실험에서도 다른 모델들에 비해 성능이 우수함을 알 수 있었다.

이미지 분류 모델로 사진을 분류할 때 사진의 어느 부분이 최종 분류 결정에 영향을 끼쳤는지를 시각화하여 확인할 수 있는 Grad-CAM의 히트맵이 홍인지문 지붕부의 왜곡을 가한 부위 주변으로 분포되는 것으로 보아 예측 모델이 왜곡을 가한 부위를 정확히 지정하고 있음을 알 수 있었다.

딥러닝에서 데이터셋은 매우 중요한 요소라 할 수 있다. ENV-22와 ENV-16에서 사용된 데이터셋은 각 90,000개와 40,000개로 2배 이상 차이 나지만 정확도는 불과 2.24%의 차이로 ENV-16 데이터셋은 더 적은 양과 비용으로 높은 정확도를 내는 데이터셋의 구성 및 조합이다.

향후 연구에서는 더 다양한 환경통합 실험을 통해 적은 데이터로 높은 정확도를 내는 고품질의 데이터셋을 구성하는 환경 조합을 찾는 연구를 지속해 나갈 계획이다. 또한, 강풍이나 강우, 지진 등 자연 재난 혹은 인위적인 개입에 의해 지속적인 모니터링이 중단되거나 시스템 오류에 의해 일시적으로 CCTV 영상 데이터를 공급받지 못한 경우에도 이를 대처하기 위한 후속 연구도 진행해 나갈 계획이다.

감사의 글

이 논문은 2023년도 정부(문화재청)의 재원으로 국립문화재연구소의 지원을 받아 수행된 연구임(No.2021A01D06-001, 부동산문화재 안전진단을 위한 지능형 영상정보 분석기반 손상탐지 및 경보 기술 개발).

References

- He, K., Zhang, X., Ren, S., Sun, J. (2016). “Deep residual learning for image recognition.” In Proceedings of the IEEE conference on computer vision and pattern recognition, pp.770-778.
- Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q. (2017). “Densely connected convolutional networks.” In Proceedings of the IEEE conference on computer vision and pattern recognition, pp.4700-4708.
- Kang, S.J., Lee, S.Y. (2023). “A Study on Performance Improvement of Heunginjimun Roof Displacement Determination Model Using Artificial Intelligence Algorithm.” Summer Annual Conference of IEIE, pp. 1050-1054.
- Krizhevsky, A., Sutskever, I., Hinton, G.E. (2017). “Imagenet classification with deep convolutional neural networks.” *Communications of the ACM*, Vol.60, No.6, pp.84-90.
- Kwon, R. (2011.08.02). Partial damage to the Heunginjimun roof due to heavy rain. SBS NEWS. https://news.sbs.co.kr/news/endPage.do?news_id=N1000962627.
- LeCun, Y., Boser, B., Denker, J.S., Henderson, D. et al. (1989). “Backpropagation applied to handwritten zip code recognition.” *Neural computation*, Vol.1, No.4, pp.541-551.
- Lee, S.Y., Cho, H.H., Lee, D.G. (2023). “Heunginjimun Displacement Determination Using Artificial Intelligence Algorithm.” *Workshop on Image Processing and Image Understanding*, pp.1-26.
- Lee, S.Y., Park, J.W. (2022). “Ensemble Deep Learning Model for Prediction of Tilt Displacement for Immovable Cultural Heritage.” *Summer Annual Conference of IEIE*, pp.1397-1398.
- Ma, N., Zhang, X., Zheng, H.T., Sun, J. (2018). “Shufflenet v2: Practical guidelines for efficient cnn architecture design.” In Proceedings of the European conference on computer vision (ECCV), pp.116-131.
- Park, J.W., Lee, S.Y. (2022). “Detection of Displacement in Wooden Cultural Properties Using f-AnoGAN Model.” *Summer Annual Conference of IEIE*, pp.1379-1380.
- Ruder, S. (2016). “An overview of gradient descent optimization algorithms.” *arXiv preprint arXiv:1609.04747*.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. et al. (2018). “Mobilenetv2: Inverted residuals and linear bottlenecks.” In Proceedings of the IEEE conference on computer vision and pattern recognition, pp.4510-4520.
- Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R. et al. (2017). “Grad-cam: Visual explanations from deep networks via gradient-based localization.” In Proceedings of the IEEE international conference on computer vision, pp.618-626.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. et al. (2016). “Rethinking the inception architecture for computer vision.” In Proceedings of the IEEE conference on computer vision and pattern recognition, pp.2818-2826.
- Tan, M., Le, Q. (2019, May). “Efficientnet: Rethinking model scaling for convolutional neural networks.” In *International conference on machine learning*, pp.6105-6114, PMLR.
- Tan, M., Le, Q. (2021, July). “Efficientnetv2: Smaller models and faster training.” In *International conference on machine learning*, pp.10096-10106, PMLR.
- Tan, M., Chen, B., Pang, R., Vasudevan, V. et al. (2019). “Mnasnet: Platform-aware neural architecture search for mobile.” In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 2820-2828.
- Wang, N., Zhao, X., Zou, Z., Zhao, P. et al. (2020). “Autonomous damage segmentation and measurement of glazed tiles in historic buildings via deep learning.” *Computer-Aided Civil and Infrastructure Engineering*, Vol.35, No.3, pp.277-291.

