

포트홀 탐지를 위한 생성형 AI 기반 데이터셋 구축과 CNN 모델 성능 평가

조민지¹ · 조혜준¹ · 최재명² · 강희조^{2*}

¹목원대학교 일반대학원 IT공학과 석사과정

²목원대학교 컴퓨터공학과 교수

Dataset Construction Using Generative AI for Pothole Detection and Performance Evaluation of CNN Models

Min Ji Cho¹ · Hye Jun Cho¹ · Jae Myeong Choi² · Heau-Jo Kang^{2*}

¹Master's Course, Department of IT Engineering, Mokwon University, Daejeon 35349, Korea

²Professor, Department of Computer Engineering, Mokwon University, Daejeon 35349, Korea

[요약]

차량의 증가와 기후 변화로 인한 포트홀 문제는 교통안전에 중대한 영향을 미치며, 이를 해결하기 위한 자동화된 탐지 기술이 요구된다. 본 논문에서는 도로의 포트홀 탐지의 효율성을 높이기 위해 생성형 AI 이미지를 활용한 새로운 접근법을 제안하였다. 실제 포트홀 이미지 대신, 생성형 AI 도구를 사용하여 제작된 가상 이미지를 활용하여 탐지 모델을 학습시켰다. Midjourney와 Playground 등 다양한 AI 도구로 생성된 이미지를 사용해 학습한 모델과 실제 이미지를 사용한 모델 사이의 성능을 비교했다. 연구 결과를 통해 생성형 AI가 포트홀 탐지 모델의 정확성과 효율성을 높이는 데 유용할 가능성이 있음을 확인하였고, 도로 유지보수의 효율성을 향상시킬 수 있는 방안을 제시하였다. 이를 통해 생성형 AI의 활용 가능성에 관한 기초를 마련하였고, 사회적 안전과 경제적 비용 절감에 기여할 것으로 기대한다.

[Abstract]

The issue of potholes, exacerbated by the increase in vehicles and climate change, has significant implications for traffic safety, necessitating the development of automated detection technologies. In this paper, we propose a novel approach to improve the efficiency of pothole detection on roads by utilizing generative AI images. Instead of using real pothole images, a detection model was trained using virtual images generated by AI tools. The performances of models trained with images generated by various AI tools, such as Midjourney and Playground, were compared with those of model trained with real images. The results confirm that generative AI can enhance the accuracy and efficiency of pothole detection models and thereby offers potential solutions for improving road maintenance efficiency. Hence, this study lays the groundwork for the potential application of generative AI in road maintenance to enhance social safety and reduce economic costs.

색인어 : 생성형 AI, 프롬프트, 포트홀, 합성곱 신경망, 딥러닝

Keyword : Generative AI, Prompt, Pothole, Convolutional Neural Network, Deep Learning

<http://dx.doi.org/10.9728/dcs.2024.25.10.3061>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 19 August 2024; **Revised** 23 September 2024

Accepted 14 October 2024

***Corresponding Author; Heau-Jo Kang**

Tel: 

E-mail: hjkang@mokwon.ac.kr

I. 서론

표 1과 같이 최근 10년간 차량 등록 수는 꾸준한 증가 추세를 보였다. 국토교통부의 통계에 따르면, 2013년 약 1,940만 대였던 차량 등록 수는 2023년 약 2,590만 대로 약 33% 증가했다. 이에 따라 화물차의 차량 등록 수 또한 2013년 약 328만 대에서 2023년 약 372만 대로 약 13% 증가했다[1].

표 1. 2013년과 2023년의 차량 및 화물차의 수
Table 1. Comparison of the number of vehicles and trucks by year

| Year | 2013 | 2023 |
|----------|------------|------------|
| Vehicles | 19,400,864 | 25,949,201 |
| Trucks | 3,285,707 | 3,726,400 |

또한, 최근 겨울철에는 폭설, 폭우, 그리고 이상 고온이 반복적으로 발생하고 있다. 이러한 기후 변화는 도로의 포트홀 발생 빈도를 증가시키는 주요 요인으로 작용하고 있다. 특히, 차량 및 화물차의 수가 증가하면서 도로의 하중이 커지고, 극단적인 기상 조건이 도로 표면의 손상을 가속화시켜 포트홀 발생을 촉진하고 있다. 결과적으로, 이러한 기후 변화와 교통량 증가는 도로 인프라의 내구성에 큰 영향을 미치며, 포트홀로 인한 사고의 급증으로 이어지고 있다.

포트홀(pothole)은 도로 유지보수에서 중요한 문제 중 하나나로, 정확한 탐지와 신속한 수리가 필요하다. 그림 1과 같이 국민권익위원회 포트홀 관련 민원 분석에 따르면 2024년 1, 2월부터 포트홀에 대한 온라인 언급이 전년 동기 대비 약 5.8배 급증하였으며, 2년 2개월(‘22.01~’24.02)간 민원분석 시스템에 수집된 ‘포트홀’ 관련 민원은 총 5만 2,262건에 달했다[2]. 포트홀은 차량파손과 교통사고 유발 등의 사회문제를 유발하기 때문에 최근에는 영상 카메라를 이용하여 포트홀을 빠르게 탐지하기 위한 연구가 많이 진행되고 있다[3].

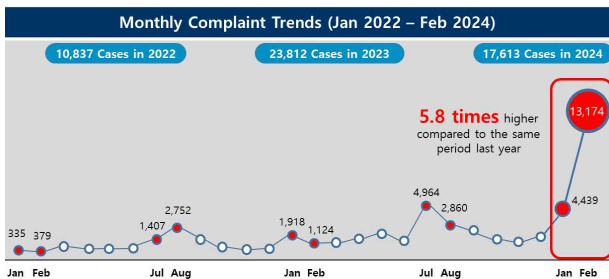


그림 1. 포트홀 민원 월별 추이
Fig. 1. Monthly trend of pothole complaints

하지만 현재 진행 중인 포트홀 관련 연구는 실제 포트홀 데이터를 확보하여 진행되고 있다. 그러나 이러한 연구는 시간과 경제적 비용이 많이 소요되는 경향이 있다.

이러한 문제를 해결하기 위한 새로운 접근법으로, 최근 텍

스트를 이미지로 변환하는 생성적 적대 신경망(GAN; generative adversarial network) 등 AI(Artificial Intelligence) 기술이 주목받고 있다. AI 기술의 발전은 예술, 광고, 농업환경 등 다양한 산업에서 널리 활용되며, 이미지 생성의 정확성과 현실성을 높이는 데 중점을 두고 있다[4].

따라서 본 논문에서는 생성형 AI를 이용해 생성된 포트홀 이미지를 데이터셋으로 하여 포트홀 탐지 모델을 학습하고, 실제 이미지 데이터셋을 활용한 모델과의 성능을 비교한다. 이 연구의 주요 목적은 실제 포트홀 이미지 대신 생성형 AI로 제작된 이미지들을 사용하여 연구를 진행할 수 있는지의 가능성을 평가한다.

II. 생성형 AI 개요 및 학습 모델 분석

2-1 생성형 AI

1) 생성형 AI의 개념과 역사

생성형 AI는 특정 입력을 기반으로, 텍스트, 이미지 또는 음악과 같은 새로운 콘텐츠를 생성할 수 있는 AI 유형을 말한다. 이는 대규모 예제 데이터셋으로 학습된 생성 모델(generative model)을 통해 이루어진다. 생성 모델은 학습한 예제들과 유사한 새로운 예제를 생성할 수 있다[5].

초기 생성형 AI 연구는 1990년대에 시작되었으며, 1998년 필기체 숫자 인식 기술이 도입되면서 합성곱 신경망(CNN; Convolutional Neural Network)이 처음으로 소개되었다. CNN은 이미지 데이터 처리를 위한 효율적인 신경망 구조로 제안되었으며, 이후 생성형 AI 연구의 중요한 기반이 되었다[6].

2000년대 이후 생성형 AI의 새로운 모델이 소개되었는데, 2014년에 발표된 논문에서는 생성적 적대 신경망(GAN) 모델이 제안되었다[7]. GAN은 이미지, 텍스트, 음성 생성 등 다양한 분야에서 활용되며, 특히 이미지 합성 분야에서 두각을 나타내 현재 생성형 AI 연구의 핵심 모델로 자리잡고 있다. 2010년대에는 언어 모델의 발전이 이루어졌으며, 2018년에는 OpenAI에서 발표한 GPT(Generative Pre-trained Transformer) 모델이 등장하였다. GPT는 대규모 텍스트 데이터를 활용해 언어 모델을 사전 학습하는 방식으로 언어 이해 성능을 크게 향상시키고 전이 학습 및 대규모 언어 모델 연구에 중요한 역할을 하였다[8].

2020년대에 들어서면서 DALL-E와 Stable Diffusion과 같은 새로운 이미지 생성 모델들이 등장하여 이미지 생성 분야에 큰 진전을 이루었고, 특히 Stable Diffusion은 오픈소스로 공개되어 생성형 AI 기술의 대중화에 기여하였다.

2) 생성형 AI의 주요 모델

생성형 AI의 주요 모델에는 생성적 적대 신경망(GAN)과

변이형 오토인코더(VAEs; Variational Autoencoders, VAEs)가 있다. GAN은 두 개의 신경망으로 구성되며, ‘생성자(Generator)’는 가짜 데이터를 생성하고 ‘판별자(Discriminator)’는 입력 데이터가 진짜인지 가짜인지 구분하는 역할을 한다. 두 네트워크는 서로의 성능을 높이기 위해 경쟁하면서 점점 더 정교한 데이터 생성을 목표로 한다[9].

GAN의 판별자 손실 함수(Discriminator Loss)는 식(1)과 같이 나타나며, 진짜 데이터를 1로, 가짜 데이터를 0으로 출력하여 판별자가 목표로 하는 바를 표현한다. 판별자는 진짜 데이터에 대해서는 높은 확률을 주고, 가짜 데이터에 대해서는 낮은 확률을 주기 위해 이 손실을 최소화하려고 한다. 또한, GAN의 생성자 손실 함수(Generator Loss)는 판별자가 가짜 데이터를 진짜로 판단하는 확률을 최대화하고자 하며 (2)와 같은 식을 갖는다.

$$L_D = -E_{x \sim p_{data}(x)}[\log D(x)] - E_{z \sim p_z(z)}[\log(1 - D(G(z)))] \tag{1}$$

$$L_G = -E_{z \sim p_z(z)}[\log(D(G(z)))] \tag{2}$$

식에서 $D(x)$ 는 판별자가 진짜 데이터 x 를 진짜로 판단할 확률이며, $D(G(z))$ 는 생성자가 생성한 가짜 데이터 $G(z)$ 를 판별자가 진짜로 판단할 확률이다. 또한, 진짜 데이터의 분포는 $p_{data}(x)$ 로 나타내고, 잠재 공간에서 샘플링된 노이즈의 분포는 $p_z(z)$ 로 나타낸다.

VAEs는 오토인코더의 변형으로 입력 데이터를 잠재 변수로 매핑하는 ‘인코더’와 잠재 변수를 입력으로 받아 원래 데이터의 분포로 재구성하는 ‘디코더’로 구성되어 있다. GAN은 고해상도 이미지를 생성하는 데에 용이하고, VAEs는 상대적으로 안정적이고, 훈련이 용이하다는 특징을 지닌다[10].

$$\mu(x) = f_\mu(x) \tag{3}$$

$$\log \sigma^2(x) = f_\sigma(x)$$

$$\hat{x} = g(z) \tag{4}$$

(3)의 식은 입력 데이터(x)에서 잠재변수(z)의 평균 $\mu(x)$ 과 표준편차 $\sigma(x)$ 를 추정하기 위한 신경망을 나타낸 인코더 함수이다. 또한, 디코더 함수(4)는 신경망을 나타내는 g 와 입력 데이터의 재구성된 형태인 \hat{x} 를 출력하여 표현된다.

3) 생성형 AI의 응용 사례

생성형 AI 기술은 많은 발전을 거듭하였고, 다양한 분야에서 혁신적인 응용 사례가 등장하고 있다. 이 기술은 이미지, 텍스트, 음성 등 다양한 데이터를 생성하고 변환하는데 활용되고 있다.

표 2. 분류에 따른 생성형 AI 응용 사례

Table 2. Applications of generative AI by category

| Type | Characteristic | Examples |
|---------------------|--|--|
| Image Generation AI | <ul style="list-style-type: none"> - Technology that generates images based on user-entered prompts - Create images close to the desired result through interactive feedback | Midjourney, Firefly, Playground |
| Text Generation AI | <ul style="list-style-type: none"> - Technology for generating paragraphs based on given text input - Used in the field of natural language processing(NLP) - Applications in fields such as translation, summarization, writing assistance, interactive AI chatbot, and code writing | GPT developed by OpenAI |
| Voice Generation AI | <ul style="list-style-type: none"> - Technology that converts input text into voice using Text-To-Speech (TTS) technology - Used in voice assistants, audiobook, creation, etc. | Platforms from Google, Amazon, and Microsoft |

2-2 프롬프트

1) 프롬프트의 개념

프롬프트는 생성형 AI 시스템에서 특정한 이미지, 텍스트 등 특정한 결과를 얻기 위한 작업을 수행하도록 지시하는 입력 문구이다. 특히 이미지 생성 AI 시스템에서 프롬프트는 사용자에게 어떤 이미지를 생성해야 하는지를 명확히 전달하는 역할을 한다. 프롬프트는 시스템에 따라 단순 명령어나 단어일 수 있고, 복잡한 설명이나 특정 조건을 포함할 수도 있다.

2) 프롬프트 최적화

프롬프트의 설계는 생성형 AI 기술의 성능을 극대화하는데 필수적이다. 프롬프트는 구체적일수록 정확한 이미지를 얻을 수 있고, 단계별로 나누어 작성했을 때 시행착오를 줄이고 원하는 이미지를 얻을 수 있다[11]. 프롬프트는 색감, 스타일 같은 감각적인 요소를 포함한 지각적인 측면과 질감, 조명 등 기술적인 요소를 포함한 기술적인 측면을 개인의 취향에 따라 결합하거나 조합해 원하는 이미지를 얻을 수 있다. 또한, 사람, 공간, 환경 등으로 이야기 구성 요소를 분리하여 설명함으로써 이미지를 효과적으로 구현할 수 있다.

프롬프트 최적화는 그림 2와 같이 사용자가 원하는 결과를 정확하게 생성하도록 프롬프트를 조정하는 과정으로, 일반적으로 반복적인 상호작용을 통해 이루어진다. 특정 모델에 맞는 프롬프트는 원하는 결과와 유사한 이미지를 제공하여 사용자가 원하는 스타일이나 요소를 이해하고 이를 재현할 수 있도록 돕는다.

프롬프트 최적화는 품질 향상, 시간과 자원 절약, 창의적인 결과 도출 등의 이유로 중요하다. 최적화된 프롬프트를 통해 AI는 보다 정확한 이미지를 생성할 수 있고, 세밀한 조정을 통해 품질과 정밀도를 올릴 수 있다. 또한, 최적화된 프롬프트는 불필요한 반복 작업을 줄이고, 빠른 피드백으로 작업 시간

을 줄일 수 있다. 세밀한 프롬프트 조정으로 다양한 스타일과 구성을 시도해 디자인의 가능성을 확장하고, 창의적인 이미지를 생성할 수 있다[12].



그림 2. 프롬프트 작성 형식에 따른 생성된 이미지
Fig. 2. Images generated according to the prompts

2-3 CNN(Convolutional Neural Network)

본 논문에서는 CNN을 기반으로 한 모델을 사용하여 포트홀 이미지에 대한 딥러닝 학습을 진행하였다. CNN이 동작하는 과정 및 구조는 다음과 같다.

1) 동작원리

CNN은 convolution 연산이 포함되어 있는 뉴럴 네트워크를 의미한다.

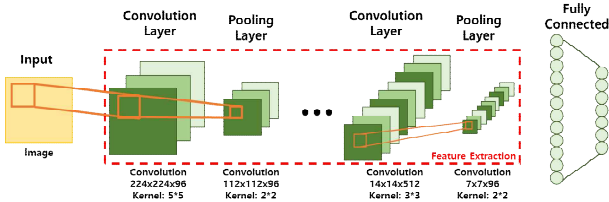


그림 3. CNN 동작 과정
Fig. 3. Operation process of CNN

그림 3과 그림 4와 같이 CNN 구조는 합성곱 레이어(Convolution Layer)와 풀링 레이어(Pooling Layer)로 이루어진 과정을 반복한다. 마지막 풀링 레이어에서 나온 결과를 1차원 벡터로 나열한다. 그리고 Fully Connected 계층으로 전달된다. 합성곱 계층에서는 합성곱 연산을 수행하며, 이때 가중치들은 필터의 형태로 표현된다. 이러한 연산의 결과는 영상의 형태로 출력되며, 이를 특징 맵이라 한다.

합성곱 계층을 통해 나온 특징 맵은 크기를 줄이기 위해 풀링 과정을 거친다. 풀링 방법은 아래의 그림 5와 같이 크게 3가지가 존재하는데, 영역 내 가장 큰 값을 선택하는 max-pooling, 가장 작은 값을 선택하는 min-pooling 마지막으로 평균값을 사용하는 average-pooling(혹은 mean-pooling)이 있다[13].

본 연구에서는 영역 내의 최대값을 추출하여 공간 차원을 줄이기 위한 max-pooling을 사용하였다.

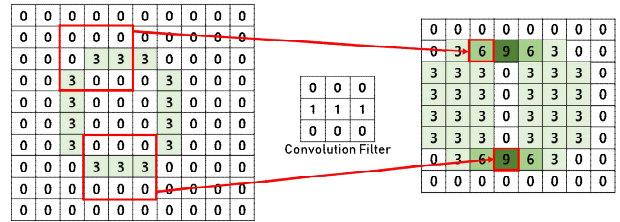


그림 4. Convolution 연산의 예
Fig. 4. Example of convolution operation

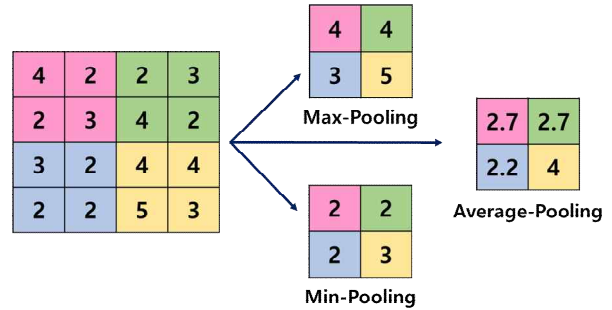


그림 5. 3가지 풀링 방법
Fig. 5. 3 pooling method

III. 이미지 생성 AI

본 연구에서는 Midjourney, Playground, Ideogram, Firefly와 같은 다양한 AI 이미지 생성 도구를 사용했다. 이들은 각각 고유한 시각적 스타일과 처리 기능을 가지고 있으며, 동일한 ‘포트홀’이라는 프롬프트를 입력하더라도 각 도구의 특성에 따라 생성되는 이미지가 다르게 나타난다.

3-1 Midjourney

Midjourney는 프롬프트 사용에 익숙하지 않아도 높은 수준의 이미지를 생성할 수 있어 누구나 쉽게 이미지를 생성할 수 있으며 그에 따라 많은 사용자가 존재해 관련 공유 지식도 쉽게 찾을 수 있다. 이 서비스는 채팅 플랫폼 Discord 서버에서 운영되는데, Discord를 처음 접하는 사용자는 가입과 프롬프트 입력법, 자기 서버 구성 등의 사용법을 익혀야 한다는 불편함이 있지만, 프롬프트에 익숙하지 않은 초보자들이 고품질의 이미지를 생성할 수 있다. 또한, Midjourney는 ‘--’와 함께 사용되는 파라미터를 제공하여 이를 이용해 정확한 이미지를 생성해 낼 수 있다. 예시로, ‘--ar 1:1’은 1대1의 화면 비율로 이미지를 생성한다는 뜻이고, ‘--no people’은 사람을 제외하고 이미지를 생성한다는 뜻, ‘--v 5’는 Midjourney의 버전을 전환한다는 뜻이다.



그림 6. 미드저니에서 생성된 포트홀 이미지
 Fig. 6. Pothole image created by Midjourney

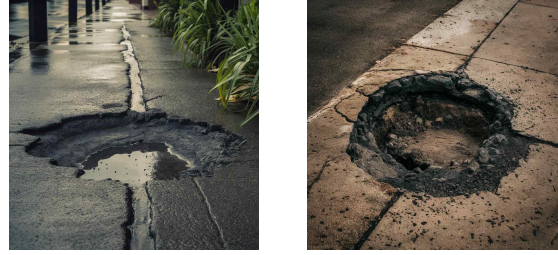


그림 8. 아이디어그램에서 생성된 포트홀 이미지
 Fig. 8. Pothole image created by Ideogram

3-2 Playground

Playground는 초보자가 사용하기 쉬운 인터페이스를 제공하고, 다른 이미지 생성형 AI와 비교할 때 무료로 하루에 1,000개의 이미지를 생성할 수 있는 가장 많은 양을 지원한다. 프롬프트 가이드스 조절 기능 등의 여러 기능을 통한 세밀한 조정으로 사용자가 원하는 이미지를 생성할 수 있다. 프롬프트뿐만 아니라 필터 선택을 통해 그림의 스타일 및 다양한 화풍을 선택할 수 있고, 제외 이미지와 이미지 참조 영역을 통해 빼고 싶은 요소를 제외하거나, 영감으로 사용할 이미지를 넣을 수 있다.



그림 7. 플레이그라운드에서 생성된 포트홀 이미지
 Fig. 7. Pothole image created by Playground

3-3 Ideogram

Ideogram은 프롬프트에 입력하는 텍스트가 자세할수록 실제와 비슷한 이미지를 생성할 수 있다. 생성된 이미지 중 마음에 드는 것을 선택하면, 해당 이미지에 대한 상세한 설명 텍스트를 확인할 수 있으며, 이를 바탕으로 추가적인 세밀한 조정이 가능하다. 원하는 이미지와 유사한 결과를 얻기 위해, 설명 텍스트를 복사하여 수정한 후, 이를 이용해 원하는 이미지를 생성할 수 있다.

3-4 Firefly

Firefly는 그래픽 분야에서 가장 영향력 있는 Adobe에서 개발된 이미지 생성 AI로, Photoshop 등 Adobe가 보유 중인 다양한 그래픽 프로그램들과의 연계를 통해 발전 가능성과 확장성이 기대되는 도구이다. Firefly는 Adobe에서 저작권을 소유한 스톡 이미지를 학습 데이터로 사용하기 때문에, 생성된 이미지에 대한 법적 안정성을 보장한다. 또한 사용자가 이미지를 Adobe 스톡에 업로드하면 이를 학습데이터로 제공하고 보상 체계를 통해 지속적인 데이터 수집과 품질 향상이 가능하도록 설계되었다. 업데이트된 Photoshop v25에서 Firefly의 기능으로 프롬프트를 통해 이미지를 생성할 수 있을 뿐만 아니라, 생성형 채우기 및 생성형 확장 같은 고급 기능도 지원한다.



그림 9. 파이어플라이에서 생성된 포트홀 이미지
 Fig. 9. Pothole image created by Firefly

IV. 포트홀 탐지 실험 분석

본 연구는 실제 포트홀 이미지를 학습하여 얻은 모델과 생성형 AI 기술을 활용하여 생성한 포트홀 이미지를 학습하여 얻은 모델 사이의 성능을 비교하고, 생성형 AI 기술이 실제 데이터를 대체하거나 보충하여 사용할 수 있는지에 대한 여부를 확인한다.

그림 10과 같이 이미지 데이터셋은 두 가지로 나뉘며, 이는 Kaggle로부터 활용한 실제 촬영된 포트홀 이미지와 생성형 AI 도구를 이용하여 얻은 포트홀 이미지이다. 포트홀 이미지에 대한 딥러닝 학습은 CNN을 기반으로 한 모델을 사용하

였고, 그림 11과 같이 포트홀 이미지와 포트홀이 존재하지 않는 일반 도로의 이미지를 이용한 이진 분류로 다루었다.

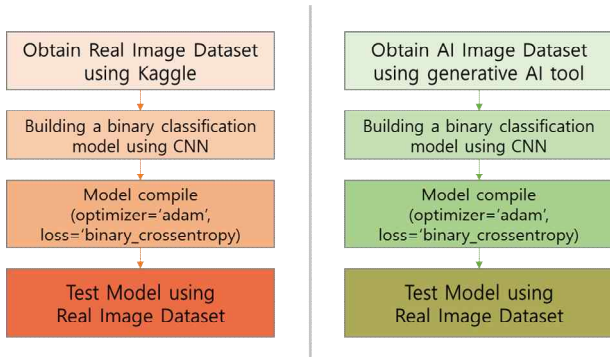


그림 10. 학습과정 플로우차트
Fig. 10. Flowchart of learning process

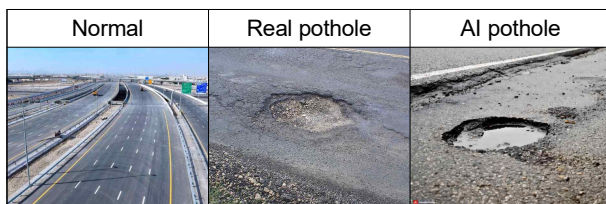


그림 11. 포트홀 이미지와 일반 도로 이미지
Fig. 11. Pothole image and normal image

이후 학습을 준비하기 위해 손실 함수를 설정하고 최적화 알고리즘을 설정하는 등의 모델 컴파일 작업을 수행하였다. 모델은 학습 데이터로 학습시키고 검증 데이터로 성능을 평가하였다. 이때 성능 평가에 사용된 이미지는 두 모델 모두 실제 촬영된 이미지를 사용하여, 생성형 AI가 만든 이미지를 활용한 모델에서도 실제 포트홀 이미지에 대한 평가가 이루어질 수 있게 하였다.

4-1 데이터셋

1) 실제 이미지를 활용한 데이터셋

본 연구에서는 포트홀이 존재하지 않는 도로의 실제 이미지와 실제 포트홀 이미지를 얻기 위해 Kaggle에서 지원하는 데이터셋을 활용하였으며, 이를 이진 분류(binary classification)로 다루었다.

이미지 전처리 및 증강은 Keras의 ImageDataGenerator 클래스를 사용하여 수행되었고, 이를 통해 데이터의 다양성을 높이고 모델의 일반화 성능을 향상시키고자 하였다. ImageDataGenerator 설정으로는 rescale, shear_range, zoom_range, horizontal_flip, validation_split 등을 사용하여, 이미지를 확대하거나 축소 및 좌우 반전하는 등의 과정을 수행하였다.

이후에 클래스 모드를 binary로 하여 이진 분류 문제로 설

정하였다. 이 과정을 통해 포트홀이 있는지에 대한 여부를 예측하도록 하였고, 학습용 데이터셋으로 80%의 이미지를 사용하고, 나머지 20%는 검증용으로 사용하였다.

2) 생성형 AI 이미지를 활용한 데이터셋

Midjourney, Playground, Ideogram 그리고 Firefly 등의 AI tool을 활용하여 생성한 이미지 또한 Keras의 ImageDataGenerator를 통해 전처리 및 증강기법이 수행되었다. ImageDataGenerator 설정은 실제 이미지를 활용한 데이터셋에 적용한 것과 동일하게 적용하였다.

생성형 AI 이미지를 활용하여 학습을 위한 이미지 데이터를 확보한 후, 해당 모델에 대한 검증은 실제 이미지를 사용하여 진행하였다.

4-2 학습과정

1) Optimizer-Adam

Adam은 각 매개변수에 대해 개별적으로 학습률을 조정하고, 학습률이 큰 매개변수와 작은 매개변수를 동시에 최적화할 수 있다. 이를 통해 서로 다른 학습률이 요구되는 상황에서도 효과적으로 학습이 이루어질 수 있게 한다. 또한, 이 알고리즘은 각 파라미터에 대한 학습률을 동적으로 업데이트하며, 계산 효율성이 높고 메모리 요구량이 적은 것으로 알려져 있다[14]. Adam은 RMSProp과 Momentum을 사용하는 경사 하강법과 유사한 파라미터 업데이트 방법을 사용한다[15].

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{5}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \tag{6}$$

$$m_t^\wedge = \left(\frac{m_t}{1 - \beta_1^t} \right) v_t^\wedge = \left(\frac{v_t}{1 - \beta_2^t} \right) \tag{7}$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{v_t^\wedge} + \epsilon} m_t^\wedge \tag{8}$$

이 알고리즘은 제공된 기울기의 지수 이동 평균 (v_t)을 사용하고, 기울기의 지수 이동 평균(m_t)도 함께 사용한다. (5), (6), (7)의 식에서 β_1 은 지수 감쇠율이며, m_t^\wedge 와 v_t^\wedge 는 각각 m_t 와 v_t 에 대한 보정된 편향을 나타낸다.

2) Binary cross entropy

Binary cross entropy는 이진 분류 문제에서 사용되는 손실 함수이며 모델이 예측한 확률 분포와 실제 레이블 간의 차이를 측정하는 데 사용된다. 특히, 출력이 확률 (0과 1사이의 값)일 때, 이진 교차 엔트로피는 모델의 성능을 평가하는 데 유용하다[16].

본 연구에서는 이미지가 포트홀을 포함하고 있는지 (1) 아

너지 (0)를 예측하는 것이 목적이기 때문에 이진 교차 엔트로피를 손실 함수로 사용하였다. 이 손실 함수는 모델이 예측한 확률과 실제 라벨 간의 차이를 효과적으로 측정하여 모델이 올바르게 학습할 수 있도록 한다. 또한, 빠른 수렴과 안정성을 제공하여 학습 과정의 효율성을 높인다.

4-3 평가방법

모델 평가 단계에서는 학습된 모델이 테스트 데이터셋에서 얼마나 잘 동작하는지를 평가한다. 이 과정은 모델이 새로운 데이터에 대해 얼마나 잘 일반화되는지를 확인하는 중요한 과정이다. 테스트 데이터셋은 모델 학습 과정에 전혀 사용되지 않은 데이터로 구성되며, 이를 통해 모델의 실제 성능을 평가할 수 있다.

테스트 데이터셋에서의 손실값 (Loss)은 모델의 예측 값과 실제 값 간의 차이를 나타내며, 낮을수록 모델이 더 정확하게 예측하고 있음을 의미한다. 또한 테스트 데이터셋에서의 정확도 (Accuracy)는 전체 예측 중 올바르게 예측한 비율을 나타내며, 높을수록 모델의 성능이 우수함을 의미한다.

모델이 학습하면서 어떻게 성능이 개선되었는지 확인하기 위해 학습 결과를 시각화하여 훈련 데이터셋에서의 손실 및 정확도 그래프를 구현하였다.

4-4 결과 분석

실험 결과, 두 모델로 평가된 손실값과 정확도는 비슷한 양상을 보였고, AI 이미지를 활용하여 학습시킨 모델에서 더 높은 정확도와 더 낮은 손실값이 관찰되었다.

표 3. 두 모델 사이의 정확도와 손실값
Table 3. The accuracy and loss of the two models

| | Real Image | AI Image |
|----------|------------|----------|
| Accuracy | 0.9524 | 0.9932 |
| Loss | 0.0998 | 0.0363 |

표 3의 결과와 같이 두 모델 사이의 정확도는 0.0408만큼의 차이를 보이며 생성형 AI의 이미지를 사용하여 훈련시킨 모델(AI Image)에서 더 좋은 성능을 보였다. 또한, 0.0635만큼의 손실값 차이를 보이는 결과를 통해서도 실제 이미지를 이용한 모델(Real Image)에서보다 AI Image를 활용한 모델에서 더 정확하게 이미지를 예측하고 있음을 확인할 수 있다.

이러한 결과는 생성형 AI 기술이 실제 데이터를 대체하거나 보충하여 사용할 수 있는 점에 대한 가능성을 제시하며, 포트홀 탐지 모델에 있어 효과적인 대안으로서 역할을 할 수 있음을 시사한다.

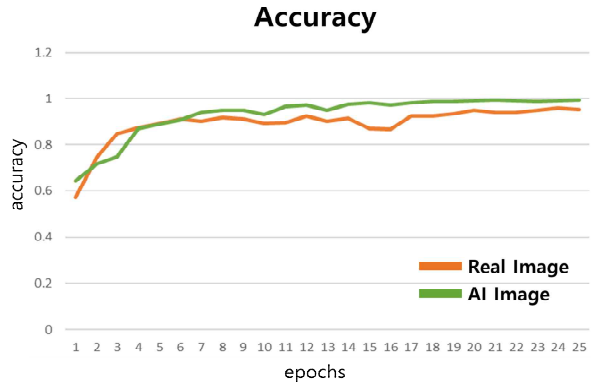


그림 12. 두 모델 사이의 정확도 비교
Fig. 12. The accuracy of the two models

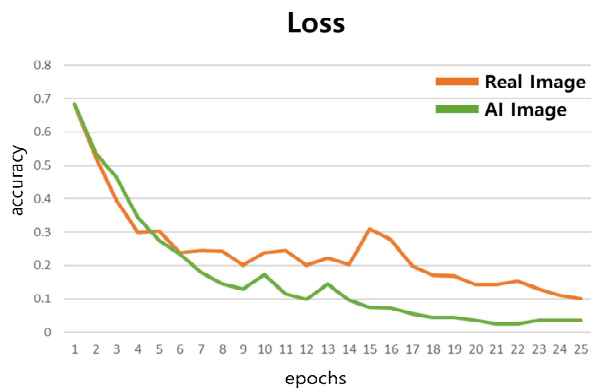


그림 13. 두 모델 사이의 손실값 비교
Fig. 13. The loss of the two models

V. 결 론

본 연구는 실제 포트홀 이미지를 데이터셋으로 사용하여 학습된 모델과 생성형 AI 이미지를 데이터셋으로 활용하여 학습된 모델 간의 성능을 비교하여 생성형 AI의 데이터 대체 가능성을 탐구하였다.

Midjourney, Playground, Ideogram, Firefly와 같은 다양한 생성형 AI 도구를 사용하여 포트홀 이미지를 생성하고, 이러한 생성형 AI 이미지를 활용한 학습 모델의 성능을 평가하였다. 실험 결과, 생성형 AI로 생성한 이미지 데이터는 포트홀 탐지 모델에서 높은 정확도와 낮은 손실값을 기록하였다. 이는 생성형 AI 기술이 실제 데이터의 대체나 보충 역할을 충분히 수행해 낼 수 있음을 시사한다.

생성형 AI를 활용한 이미지 생성은 데이터 수집 및 전처리에 소요되는 시간과 비용을 줄이는 데 기여할 수 있다. 실제 포트홀 이미지를 수집하고 라벨링하는 데 드는 자원과 비용을 고려할 때, 생성형 AI 도구를 활용하는 것은 효율적이며 경제적인 대안으로 평가될 수 있다. 특히, 다양한 생성형 AI

도구들이 제공하는 프롬프트 최적화와 데이터 증강 기능을 통해, 다양한 상황에서의 포트홀 이미지를 생성하고 모델을 학습시키는 것이 가능하다.

본 연구는 생성형 AI가 포트홀 탐지 모델의 학습에 효과적이라는 것에 기반을 마련하였지만, 향후 연구에서는 생성형 AI로 생성된 데이터의 품질과 신뢰성을 더욱 정밀하게 평가할 필요가 있다. 또한, 생성형 AI의 최신 기술 발전과 다양한 도구의 특성을 활용하여 포트홀 탐지 외에도 다양한 분야에서의 응용 가능성을 탐색할 수 있을 것이다.

결론적으로, 생성형 AI 기술은 포트홀 탐지와 같은 분야에서 실질적인 데이터 대체 및 보충 역할을 수행할 수 있으며, 데이터 수집과 모델 학습의 효율성을 향상시키는 데 중요한 기여를 할 수 있다. 이를 통해 도로 유지보수의 효율성을 높이고, 사회적인 안전과 경제적 비용 절감에 기여할 수 있을 것을 기대된다.

감사의 글

이 논문은 2024년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(P0024164, 2024년 지역혁신클러스터육성).

참고문헌

- [1] Ministry of Land, Infrastructure and Transport. Current Status of Vehicle Registration [Internet]. Available: https://www.index.go.kr/unity/potal/main/EachDtlPageDetail.do?idx_cd=1257.
- [2] ACRC (Anti-Corruption & Civil Rights Commission). Analysis of Civil Complaints Related to Potholes (Road Digs) to Ensure Public Safety [Internet]. Available: https://acrc.go.kr/board.es?mid=a10402010000&bid=4A&act=view&list_no=69385.
- [3] Y.-T. Jo and S.-K. Ryu, "Pothole Detection Algorithm Based on Saliency Map for Improving Detection Performance," *The Journal of the Korea Institute of Intelligent Transport Systems*, Vol. 15, No. 4, pp. 104-114, August 2016. <https://doi.org/10.12815/kits.2016.15.4.104>
- [4] S. Yoon, Y. Lee, E. Jung, and T. I. Ahn, "Agricultural Applicability of AI Based Image Generation," *Journal of Bio-Environment Control*, Vol. 33, No. 2, pp. 120-128, April 2024. <https://doi.org/10.12791/KSBEC.2024.33.2.120>
- [5] J. V. Pavlik, "Collaborating with ChatGPT: Considering the Implications of Generative Artificial Intelligence for Journalism and Media Education," *Journalism & Mass Communication Educator*, Vol. 78, No. 1, pp. 84-93, March 2023. <https://doi.org/10.1177/10776958221149577>
- [6] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," *Proceedings of the IEEE*, Vol. 86, No. 11, pp. 2278-2324, November 1998. <https://doi.org/10.1109/5.726791>
- [7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, ... and Y. Bengio, "Generative Adversarial Nets," in *Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS '14)*, Montreal, Canada, pp. 2672-2680, December 2014. <https://doi.org/10.48550/arXiv.1406.2661>
- [8] OpenAI. Improving Language Understanding by Generative Pre-Training [Internet]. Available: <https://www.openai.com/research/language-unsupervised/>.
- [9] J. Oh and J.-W. Cha, "Genetic Algorithm for Text Generation in GAN Framework," in *Proceedings of Korea Computer Congress 2019 (KCC 2019)*, Jeju, pp. 563-565, June 2019.
- [10] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," in *Proceedings of the 2nd International Conference on Learning Representations (ICLR 2014)*, Banff, Canada, April 2014. <https://doi.org/10.48550/arXiv.1312.6114>
- [11] Y. Hao, Z. Chi, L. Dong, and F. Wei, "Optimizing Prompts for Text-to-Image Generation," in *Proceedings of the 37th International Conference on Neural Information Processing Systems (NIPS '23)*, New Orleans: LA, pp. 66923-66939, December 2022. <https://doi.org/10.48550/arXiv.2212.09611>
- [12] D.-H. Kwon, "Analysis of Prompt Elements and Use Cases in Image-Generating AI: Focusing on Midjourney, Stable Diffusion, Firefly, DALL·E," *Journal of Digital Contents Society*, Vol. 25, No. 2, pp. 341-354, February 2024. <http://dx.doi.org/10.9728/dcs.2024.25.2.341>
- [13] T. Hwang, A Design of Virtual Exposure Fusion CNN to enable Contrast Enhancement of HDR Images, Master's Thesis, Seokyeong University, Seoul, February 2020.
- [14] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *Proceedings of 3rd International Conference on Learning Representations (ICLR 2015)*, San Diego: CA, May 2015. <https://doi.org/10.48550/arXiv.1412.6980>
- [15] M. Reyad, A. M. Sarhan, and M. Arafa, "A Modified Adam Algorithm for Deep Neural Network Optimization," *Neural Computing and Applications*, Vol. 35, No. 23, pp. 17095-17112, August 2023. <https://doi.org/10.1007/s00521>

-023-08568-z

[16] A. U. Ruby, P. Theerthagiri, I. J. Jacob, and Y. Vamsidhar, "Binary Cross Entropy with Deep Learning Technique for Image Classification," *International Journal of Advanced Trends in Computer Science and Engineering*, Vol. 9, No. 4, pp. 5393-5397, July-August 2020. <https://doi.org/10.30534/ijatcse/2020/175942020>



조민지(Min Ji Cho)

2023년 : 충북대학교 생명과학부
미생물학과 (이학사)

2024년~현 재 : 목원대학교 IT공학과 석사과정
※관심분야 : 머신러닝, 딥러닝, 빅데이터



조혜준(Hye Jun Cho)

2024년 : 목원대학교 컴퓨터공학과
(공학사)

2024년~현 재 : 목원대학교 IT공학과 석사과정
※관심분야 : 컴퓨터그래픽스, 증강현실



최재명(Jae Myeong Choi)

2014년 8월 : 목원대학교 대학원
IT공학과(공학박사)

2015년 4월~현 재: 목원대학교 컴퓨터공학과 조교수
※관심분야 : 무선통신시스템, 지능형재난시스템, 사회안전, 재난관리, 통신재난, 멀티미디어통신, IoT, 디지털콘텐츠 등



강희조(Heau-Jo Kang)

1994년 : 한국항공대학교 대학원 항공
자공학과(공학박사)

2003년~현 재: 한국디지털콘텐츠학회 명예회장, 사회안전학회 명예회장
2009년~현 재: 행정안전부 재난대응 안전한국훈련 중앙평가단 부단장
2017년~현 재: 행정안전부 재난관리평가 및 국가기반체계평가위원
2019년~현 재: 행정안전부 중앙안전교육점검단 위원
2003년~현 재: 목원대학교 컴퓨터공학과 교수, 대학원 사회안전학과 교수
※관심분야 : 재난안전통신, 스마트재난관리, 사회재난안전정책, 위기관리, 무선이동통신, 사물인터넷(IoT), 빅데이터, 항행안전시설, 디지털콘텐츠, 클라우드 컴퓨팅, 기술정책 등