

드론 촬영 이미지 데이터를 기반으로 한 도로 균열 탐지 딥러닝 모델 개발

Development of Deep Learning Model for Detecting Road Cracks Based on Drone Image Data

권영주* · 문성호**

Young-Ju Kwon* · Sung-ho Mun**

Abstract

Drones are used in various fields, including land survey, transportation, forestry/agriculture, marine, environment, disaster prevention, water resources, cultural assets, and construction, as their industrial importance and market size have increased. In this study, image data for deep learning was collected using a mavic3 drone capturing images at a shooting altitude was 20 m with $\times 7$ magnification. Swin Transformer and UperNet were employed as the backbone and architecture of the deep learning model. About 800 sheets of labeled data were augmented to increase the amount of data. The learning process encompassed three rounds. The Cross-Entropy loss function was used in the first and second learning; the Tversky loss function was used in the third learning. In the future, when the crack detection model is advanced through convergence with the Internet of Things (IoT) through additional research, it will be possible to detect patching or potholes. In addition, it is expected that real-time detection tasks of drones can quickly secure the detection of pavement maintenance sections.

Keywords: Deep Learning, Swin Transformer, Backbone, Loss Function, Drone

1. 서 론

1.1 연구의 배경 및 목적

포장 균열은 가장 일반적인 포장 파손 중 하나이며, 수분 침투 및 여러 가지 도로 문제들을 야기할 수 있다. 도로 균열을 통해 물이 침투하여 도로 평탄성에 영향을 미칠 수 있기 때문에 균열에 대한 관리는 필수적이라 볼 수 있다.

기존 도로 상태 분석은 사람에 의한 실제 관측 및

분석을 통해 이루어지고 있다. 이 방법은 사람이 포장 표면을 시각적으로 관측하는 것을 말하며, 시간이 많이 걸리고 안전사고 문제가 발생할 수 있으며, 사람의 주관적인 판단하에 이루어지기 때문에 부정확할 수가 있다. 또한 실제 측정된 영상 및 부가 데이터에 의한 분석에도 노이즈 및 주변 환경 변수에 따라 도로 상태의 결함을 검출해 내는 데에 오류가 수반될 여지가 있다. 이러한 영상 분석에 실제 컴퓨터 비전 기술의 전통적인 기법이 사용될 수 있지만 기

*서울과학기술대학교 일반대학원 건설시스템공학과 석사(주저자: 1018pinetree@naver.com)

**서울과학기술대학교 일반대학원 건설시스템공학과 교수(교신저자: smun@seoultech.ac.kr)

존의 영상 분석 기술은 사람에 의해 정의된 특성들만 추출하는 방식이기 때문에 다양한 형태의 특징들을 추출하는 데에는 한계가 있다. 이러한 측면에서도 도로 상태 측정에 비용이 낮으면서도 효율적인 자동화된 측정방식이 필요할 뿐만 아니라 수집된 측정데이터의 분석 또한 빠르고 정확한 자동화 방식의 개발이 절실히 요구되고 있다.

도로포장 조사에 드론을 활용할 경우, 한 번의 비행으로 동시다발적으로 차선을 촬영할 수 있어 기존에 사용되는 차량형 조사장비 기술과 비교하였을 때 촬영시간의 감소 및 투여 인력감소로 인한 경제적인 효과를 기대할 수 있다.

드론 관련 산업은 최근 몇 년간 시장 규모 및 산업적 중요성이 증가하여 현재 국방, 산림, 농업, 해양, 환경, 수자원, 문화재 등 다양한 분야에서 활용되고 있으며 사물인터넷(IoT), 인공지능(Artificial Intelligence) 등 4차 산업혁명의 핵심기술을 검증하고 적용할 수 있는 기술로 떠오르고 있다(김담룡, 2020). 드론을 활용하여 도로포장 분석을 할 경우 원격 조종이 가능하고, 교통량의 영향을 받지 않으므로 시간적·공간적 제약을 받지 않는다. 본 연구에서는 이러한 드론의 장점을 활용하여 드론으로 촬영한 이미지 데이터로 도로파손형태 중 하나인 균열을 탐지할 수 있는 딥러닝 모델을 개발하고자 한다.

1.2 연구의 범위 및 방법

균열탐지분석 모델의 개발을 위해 이미지데이터의 수집을 실시하였으며, 드론을 통한 항공촬영이 요구되었다. 촬영계획을 세우고 드론을 활용하여 촬영된 이미지를 바탕으로 데이터 라벨링(Labeling)을 수행, 딥러닝 학습을 반복적으로 수행하여 균열 탐지 딥러닝 모델을 개발하였다. 연구방법론은 Fig. 1과 같다.

연구에 사용된 드론은 DJI 사의 Mavic 3이며 항공 촬영시 이미지데이터는 데이터의 학습에 필요한 해

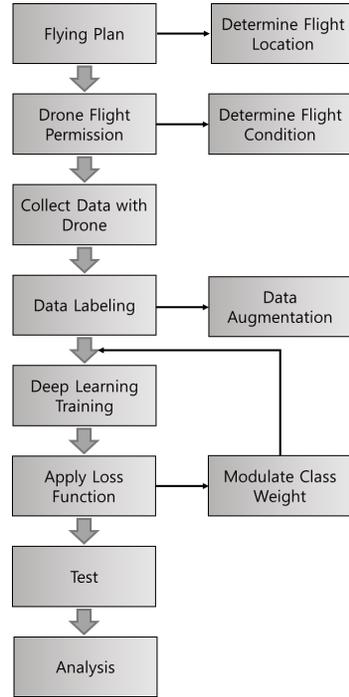


Fig. 1. Study Flow Diagram

상도를 고려하여 20m의 고도에서 x7의 배율로 촬영을 진행하였다.

1.3 국내외 연구사례

1.3.1 국내사례

한국건설기술연구원에서는 일반국도, 고속국도의 절토 사면에 대하여 체계적인 유지관리를 통하여 사면 안전사고 예방 및 국도 유지보수 예산의 효율적인 집행을 목적으로하는 도로절토사면 유지관리 시스템(Cut Slope Management System, CSMS)을 운영한다(한국건설기술연구원, 2017). 여기서, CSMS 운영팀은 현장 조사 시 대형 비탈면의 관리에 있어서 조사인력의 접근이 어렵고 육안 조사에 한계가 존재한다고 판단, 보완 대책으로 드론을 도입하여 항공부 차원의 자료를 획득하여 관리하기 시작하였다.

대형 교량의 슬래브 하부는 조사인력의 접근이 어려워 육안점검을 수행하기에는 많은 어려움이 존재

한다. 일반적으로 대형 교량 점검은 조사를 위한 특수차량을 동원하여 조사가 수행되는데, 이러한 특수차량을 이용한 점검은 일시적인 교통통제가 요구됨과 동시에 시간 또한 많이 소요되어 교통체증이 발생하는 등 점검방법의 개선이 요구되었다(한국건설기술연구원, 2017). 부산교통공사에서는 인력의 접근이 어려워 육안점검이 어려운 교량 조사를 드론을 활용하여 수행하였으며, 근접 모니터링을 통해 교량 하부의 나사 조임 상태나 균열 등의 정밀 관측을 수행할 수 있었다.

1.3.2 국외사례

호주 멜버른 대학에서는 무인항공기를 이용하여 도로의 상태를 조사하는 연구는 드론이 대중화되기 이전인 2008년에도 수행되었으며, Zhang은 무인헬리콥터의 위치, 자세 및 속도를 감지하기 위해 GPS/IMU 및 지자기센서를 장착하여 비포장도로의 상태를 사진측량 기술을 이용하여 점검할 수 있는 시스템을 제안하였다. Zhang and Elaksher(2012)의 연구에서는 비포장도로에 대한 3D DSM(Digital Surface Model) 구현을 완성하였다. 이를 위해 무인헬리콥터에서 획득한 영상의 방향과 도로 표면 변형의 크기를 정확하게 측정할 수 있는 3D 도로 DSM 및 정사영상 생성 알고리즘이 개발되었으며, 2012년 연구결과에서 도로 변형에 대한 3D 모델과 현장 측정값을 비교했을 때 0.5cm의 차이를 확인하여 무인항공기를 이용한 도로 상태 모니터링 방법의 신뢰도 및 잠재성을 확인하였다.

Liepaja 대학의 연구팀은 드론을 활용하여 도로의 포트홀 탐지 및 부피 계산을 위한 알고리즘을 개발했다. 초음파 센서를 사용하여 포트홀의 3D 크기를 결정한다. 노면에 대한 스캐닝을 구성하고 결과 데이터를 이진 코드로 변환하여 물체의 부피를 산정하는 알고리즘을 개발했다(Zaimis, 2019). 이 기술은 운송 시스템, 엔지니어링 분야의 추가 개발을 통해

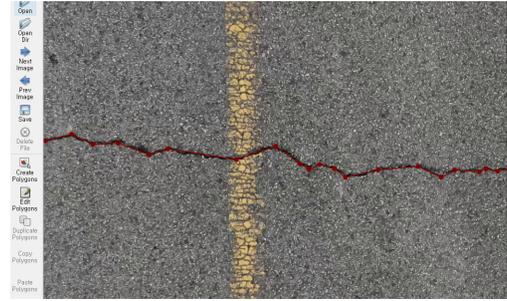


Fig. 2. Labeling Photo Using the Labelme Program

실제 환경에서 알고리즘 구현 장비를 개발하는데 활용될 수 있다.

2. 데이터 라벨링(Labeling)

본 연구에서는 데이터 학습을 위해 수집한 이미지 파일에 라벨링(Labeling)을 실시하였다. 균열이미지에 “Labelme” 프로그램을 적용할 경우 Fig. 2와 같은 형상을 보인다.

라벨링은 컴퓨터가 스스로 학습할 수 있는 형태로 데이터를 가공하는 작업을 말하며 균열에 해당하는 부분을 직접 표시를 하면 컴퓨터는 이를 기준으로 학습을 진행한다.

3. 딥러닝 학습 모델

3.1 Backbone Network

Backbone Network란 이미지 분류의 인공 신경망 구조에서 효율적인 연산을 위해 이미지의 특징(Feature map)을 추출하는 과정을 말하며 특징 추출기(Feature Extractor)라고도 한다(성진욱·정기성, 2021). 출력값에 대해 예측하기 위해서는 입력 데이터의 특징을 잘 추출해야 한다. 입력값이 고차원이고, 대규모인 경우 Backbone Network를 통해 계산 복잡도를 낮춰야 한다. 본 연구에서는 Backbone Network로 이미지 분할 학습에 효율적인

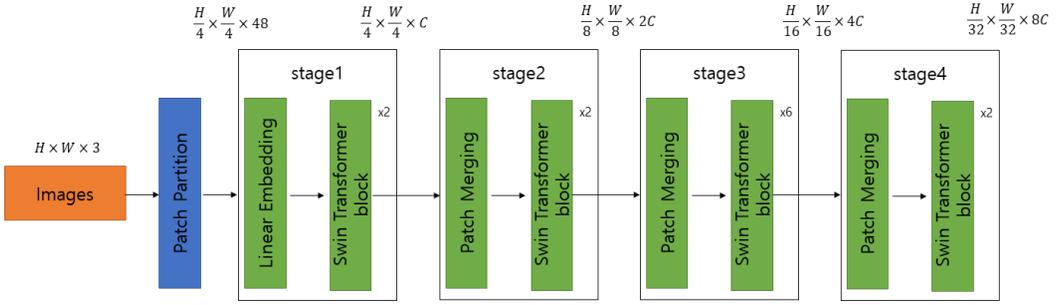


Fig. 3. Structure of Swin Transformer (Guo et al., 2022)

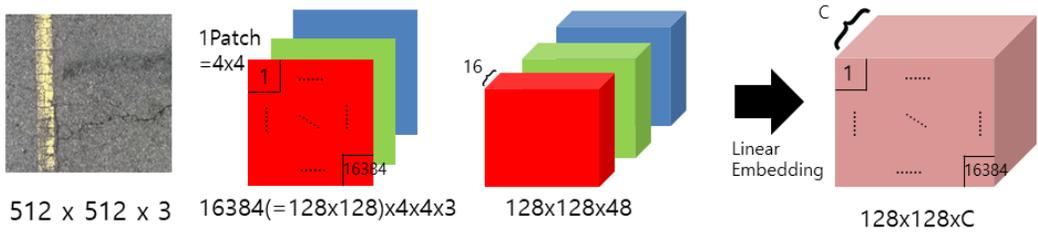


Fig. 4. Patch Particion & Linear Embedding

Swin Transformer를 사용하였다.

3.2 Swin Transformer

이미지의 분석에 있어 텍스트와 다르게 물체의 크기와 해상도가 매우 다양하기 때문에 전체 데이터를 통해 Attention을 진행하고 고정된 크기의 Token을 사용했던 Vision Transformer(ViT)의 경우 좋은 성능을 내기 어렵다. 그에 반해 Swin Transformer는 계층적인 Feature map을 구성하는 Hierarchical Shifted windows 방법을 제안하므로써 ViT 대비 선형적인 계산 복잡도를 가지며 좋은 성능을 보여 균열탐지 모델에 사용하기 적합하다(Liu et al., 2021). 본 연구에서는 Swin-B 모델을 사용하였고, Swin Transformer의 전체적인 구조는 Fig. 3과 같다.

3.2.1 Patch Partition & Linear Embedding

이미지 파일이 입력되면 Patch Partition과 Linear Embedding 과정이 수행된다. 예를 들어, 512×512×3

(pixels) 크기의 이미지가 있으면 4×4 크기의 Patch로 나눈 뒤, 채널 축으로 Concat을 시킨다. 그러면 512×512×3 크기의 이미지는 (512÷4)×(512÷4)×3=128×128×3의 data가 형성된다. 이를 Linear Embedding 과정을 거치면서 임의의 차원 C를 갖는 128×128×C의 차원으로 Projection을 시킨다. 본 연구에서 사용한 Swin-B 모델의 경우 C=128이 적용된다. Patch Partition & Linear Embedding의 구조는 Fig. 4와 같다.

3.2.2 Patch Merging

Patch Merging이란 인접한 Patch를 병합하면서 해상도를 줄여주는 과정이다. Linear Embedding을 거친 데이터들은 Patch Merging을 통해 Patch를 Channel 축으로 병합하여 차원을 4C로 올려준다. 즉, 128×128×C 형태의 Data는 64×64×4C 형태의 data로 변환되고, 이후 Linear Layer이라는 과정을 통해 차원을 2C로 조정한다.

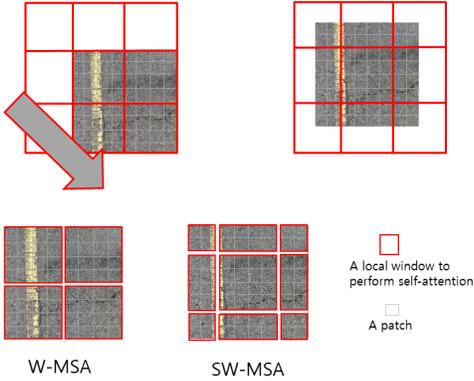


Fig. 5. W-MSA & SW-MSA

3.2.3 Swin Transformer Block

Patch Merging이 수행된 후 Data는 Swin Transformer Block과정을 거치게 된다. Swin Transformer Block에서는 크게 W-MSA(Window Multi-head Self Attention)와 SW-MSA(Shifted Window Multi-head Self Attention) 두 과정이 수행된다. Self Attention은 인접한 Patch들 간의 연관성을 추출하는 연산과정인데, 이를 수행하기 Fig. 5와 같이 Patch들끼리 Local window를 형성한 후 W-MSA가 수행된다. 그 후 서로 다른 Window들 간의 Self Attention을 수행하기 위해 Window를 옮겨서 Self Attention을 다시 수행하는 SW-MSA 과정이 진행된다. 이럴 경우 Fig. 5와 같이 크기와 모양이 다른 Window가 생성되고 Window 개수가 늘어나면서 계산 효율성이 떨어지게 된다.

이러한 문제점을 해결하기 위해 Cyclic Shift와 Attention Mask 기법이 적용된다. Fig. 6과 같이 Cyclic Shift는 9개로 쪼개진 Local Window에서 기존의 Shifted되기 전의 M×M 사이즈를 가진 가운데 Local Window를 기준으로 나머지 A,B,C Window를 우측 하단에 붙여서 Shifted 되기 전의 구조와 동일한 Window 사이즈와 개수가 유지가 된다. 1번 Window같은 경우에는 인접한 Patch들간의 Self Attention을 수행해도 되지만 2번, 3번, 4번 Window

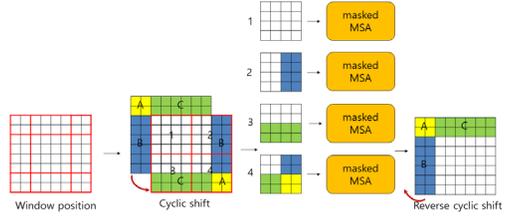


Fig. 6. Cyclic shift & Attention Mask

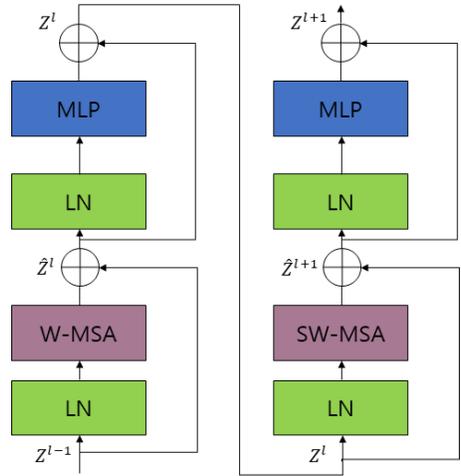


Fig. 7. Swin Transformer Block

의 경우 원래 인접해 있던 patch들이 아니기 때문에 Self Attention 연산을 수행하면 안 된다. 이를 방지하기 위해 Masking을 씌워서 연산이 수행되지 않도록 한다. Masking이 된 상태로 Self Attention 연산을 수행한 뒤 다시 A, B, C Window를 원래 위치로 돌려주는 연산을 수행한다. Cyclic shift & Attention Mask의 과정은 Fig. 6와 같다.

W-MSA & SW-MSA의 Self-attention 계산 시 Relative Position Bias라는 상대 좌표를 더해주는 연산을 한다.

$$Attention(Q, K, V) = SoftMax(QK^T / \sqrt{d} + B) V \quad (1)$$

Swin Transformer Block의 구조는 Fig. 7과 같으며, 식 (2)~식 (5)와같은 계산과정을 통해 연산이 진

행된다.

$$\hat{Z}^l = W - \text{MSA}(\text{LN}(\hat{Z}^{l-1})) + \hat{Z}^{l-1} \quad (2)$$

$$Z^l = \text{MLP}(\text{LN}(\hat{Z}^l)) + \hat{Z}^l \quad (3)$$

$$\hat{Z}^{l+1} = SW - \text{MSA}(\text{LN}(Z^l)) + Z^l \quad (4)$$

$$Z^{l+1} = \text{MLP}(\text{LN}(\hat{Z}^{l+1})) + \hat{Z}^{l+1} \quad (5)$$

여기서, \hat{Z}^l 은 W-MSA 모듈의 Output Features를, \hat{Z}^{l+1} 은 SW-MSA 모듈의 Out Features를, Z^l 와 Z^{l+1} 는 MLP의 Output Features를 의미한다. 그리고 LN은 Layer Normalization(표준화), MLP는 Multi Layer Perception을 의미한다.

3.3 FPN(Feature Pyramid Network)

Backbone Network(Swin Transformer)를 통해 추출한 Feature Map을 다음과 같은 FPN기반의 UperNet 알고리즘을 통해 학습을 시킨다(Xiao et al., 2018).

FPN은 Fig. 8과같이 전체적으로 Pyramid식 구조로 이루어져 있고, 피라미드 아래층에 있을수록 Feature Map은 High Resolution, Low-level Feature이고 위로 갈수록 Low Resolution, High-Level Feature을

보유한다.

3.3.1 Bottom-up pathway

Bottom-up Pathway 과정은 이미지를 Convolutional Network에 입력하여 Forward Pass하여 2배씩 작아지는 Feature Map을 추출하는 과정이다. 본 연구에서는 이 부분은 Swin Transformer를 통해 추출하였다. 이미지는 각 Stage를 통해 1/4, 1/8, 1/16, 1/32의 크기를 가진 Feature Map을 추출한다.

3.3.2 Top-down Pathway and Lateral Connections

Top-down Pathway는 각 Pyramid Level에 있는 Feature Map을 2배로 Upsampling하고 Channel 수를 동일하게 맞춰주는 과정이다. 이때 사용되는 기법은 Nearest Neighbor Upsampling 기법이다. Nearest Neighbor Upsampling 기법을 통해 Feature Map을 2배로 키우고 1x1 Convolution 연산을 통해 Channel 수를 256으로 동일하게 맞추게 된다(Shrivastava et al., 2016).

Lateral Connection 과정을 통해 바로 아래 있는 피라미드의 Feature Map과 Element-wise Addition 연산을 수행한다. 3x3 Convolution 연산을 통해 p2, p3, p4, p5가 나오게 된다. 이렇게 얻은 Feature Map을 Fig. 9과 같은 방법으로 RPN(Region Proposal Network)에 입력하고, 각 Class의 Score와 Bounding Box Regressor를 출력하게 된다. 이후 Non Maximum Suppression 알고리즘을 적용하여 Class Score가 높은 상위 1000개의 Resion Proposal을 출력한다. Region Proposal이란 이미지 안에서 객체가 있을 만한 후보 영역을 먼저 찾아주는 방법이다(Ren et al., 2016).

이렇게 얻은 Feature Map과 1000개의 Region Proposals를 사용하여 RoI Pooling을 수행한다. RoI Pooling은 Region Proposals을 통해 나온 Feature를 고정된 크기로 만들어주는 작업을 말한다. 이때

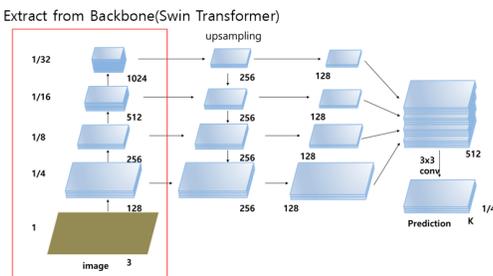


Fig. 8. FPN (Feature Pyramid Network) (Lin et al., 2017)

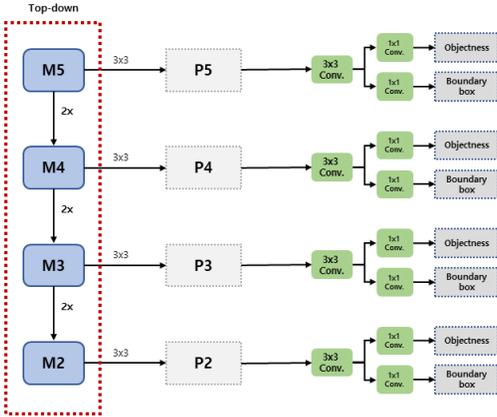


Fig. 9. RPN (Region Proposal Network)

Region Proposals를 어떤 Scale의 Feature Map과 매칭시킬지 결정해야 하는데 다음 공식에 따라 k 번째 Feature map과 매칭한다.

$$k = [k_0 + \log_2(\sqrt{wh}/224)] \quad (6)$$

w, h 는 Region Proposal의 Width, Height에 해당하고 k_0 는 Target Level을 의미한다. k 번째 Feature Map과 매칭시켜 RoI Pooling을 통해 얻은 고정된 Feature Map을 통해 학습시키고 Loss Function을 적용한다.

4. 학습과정

192×192 해상도의 ImageNet-22K Dataset에 Pretrain을 한 이후 512×512 해상도의 ADE20K Dataset에 대해 Fine-tuning한 Backbone Weight를 불러와 학습을 진행하였다. 라벨링 된 401장의 도로 균열 데이터를 이용하여 학습을 진행하였고 25장을 Validation Set으로 사용하였다.

도로 균열 이미지의 경우 전체 이미지 중 균열이 차지하는 비율이 매우 작기 때문에 Background Class Loss에 대해 1배, Crack Class Loss에 대해 150배를

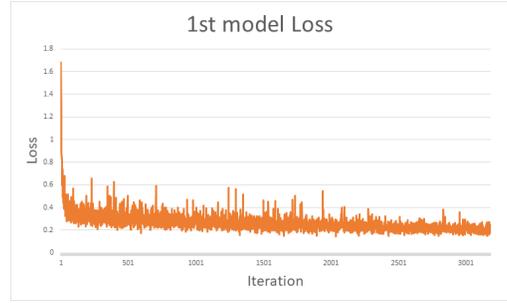


Fig. 10. 1st Model Loss

곱하여 전체 Loss를 계산하였다. Window Size는 12로 총 Iteration 160k 학습을 수행하였다. 사용한 Loss Function은 Cross-Entropy Loss Function이다. 1차 Loss 계산 결과는 Fig. 10과 같다.

Cross-Entropy Loss Function의 식은 다음과 같으며 Softmax함수가 적용되어 0~1사이의 값으로 출력이 된다.

$$H(p, q) = - \sum_x p(x) \log q(x) \quad (7)$$

여기서, p = 정답 분포

q = 모델의 예측 분포

SoftMax함수는 입력값을 0과 1사이의 확률값으로 바꾸어 주는 함수로 그 식은 다음과 같다.

$$softmax = \frac{\exp(a_k)}{\sum_{i=1}^n \exp(a_i)} \quad (8)$$

1차적으로 학습을 한 후, 라벨링 데이터 증강을 위해 학습된 1차 모델을 바탕으로, 라벨링이 되지 않은 총 373장 Inference를 수행 후 검수하여 균열이 잘 예측된 345장을 추가하여 1차 모델을 재학습하였다. 768×768 Resolution으로 Crop한 이미지를 학습데이터로 사용하였고 Window Size는 동일하게

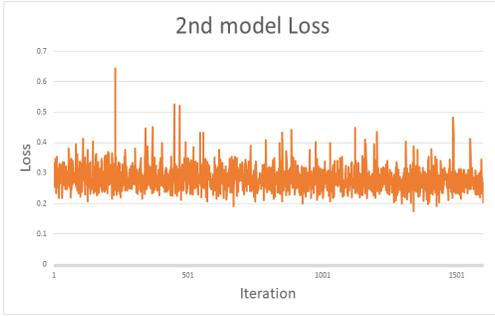


Fig. 11. 2nd Model Loss



Fig. 12. 3rd Model Loss

12로 총 Iteration 80k의 학습을 수행하였다. 1차 학습 모델보다 실제 균열의 너비에 일치하도록 예측한 계층 하기 위해 Crack의 가중치를 조절하여 Background Class Loss에 대해 1배, Crack Class Loss에 대해 120배를 곱하여 전체 Loss를 계산하였다. 2차 Loss 계산 결과는 Fig. 11과 같다.

3차 학습에서는 2차 모델에 사용한 Dataset과 함께 새롭게 라벨링한 83장의 이미지로 2차 모델을 재 학습하였다. 1080×1080 Resolution으로 Crop한 이미지를 학습데이터로 사용하였고 Online Hard Example Mining(OHEM) Sampling을 적용하였다. 이전 학습모델에서는 Cross-Entropy Loss Function을 사용하였는데 3차 학습에서는 보다 얇은 물체에 대한 Segmentation을 잘 하는 Tversky Loss Function을 사용하였다. Tversky Loss Function의 식은 다음과 같다(Salehi et al., 2017).

$$TI(Tversky\ Index) = \frac{TP}{TP + \alpha FN + \beta FP} \quad (9)$$

여기서, TP = True Positive = 실제 True인 정답을 True라고 예측(정답)
 FN = False Negative = 실제 True인 정답을 False라고 예측(오답)
 FP = False Positive = 실제 False인 정답을 True라고 예측(오답)
 α, β = Tversky 매개변수

이때 매개변수인 $\alpha = 0.3, \beta = 0.7$ 을 적용하였다. 3차 Loss 계산 결과는 Fig. 12와 같다.

5. 균열탐지 결과

학습된 모델을 바탕으로 이미지데이터를 입력하여 균열탐지를 진행하였다. Fig. 13과 같이 이미지데이터를 입력하면 균열에 해당하는 부분은 붉은색으로 탐지가 된다.

이러한 탐지기능을 활용하여 본 연구에서는 특정 도로차선의 균열을 분석할 수 있는 모델을 개발하였다. 드론으로 한 차선을 영상으로 2m/s의 속도로



Fig. 13. Image data (Above: Original Image, Below: Crack Detection Result)



Fig. 14. Comparison of Frame Extraction Images and Crack Detection Results

```

드론의 속도가 약 2m/s로, 영상 1초당 한 프레임씩 추출하여 한 이미지 내 도로의 길이는 2m로 가정합니다.
영상에 위치한 물체에 추출한 이미지 및 관련 탐지 이미지를 저장합니다.
관측된 이미지 이후의 손재는 각 이미지의 관측율을 의미합니다.
추출한 모든 이미지의 비율이 100%이면 "crack" 폴더 이름 위에 영상 전체 균열률에 대한 값이 추가됩니다.
[1/1] C:\mesa_crack_detection\origin\mesa_crack_detection\origin\0006_MP4 extraction starts.
[1/1] C:\mesa_crack_detection\origin\mesa_crack_detection\origin\0006_MP4 extraction completed.
[1/1] Image inference: 100% 29/29 106.95x100.00, 14.36s/1c
C:\mesa_crack_detection\origin\mesa_crack_detection\origin\0006_MP4 crack rate: 2.78%
    
```

Fig. 15. Average Crack Rate Calculation

촬영을 하고, 1초당 한 프레임씩 추출하여 각각 균열 탐지를 한 후 촬영 구간의 전체 평균 균열율을 계산한다. 드론촬영시, 촬영 구간에 대한 위치 및 비행 정보는 드론에 저장된다.

Fig. 14에 균열탐지모델 적용 전후 이미지를 나타냈고, Fig. 15에 균열율 계산결과 화면 이미지를 나타냈다. Fig. 16은 데이터획득 시 드론의 위치정보이다.

데이터 학습시 20m에서 촬영한 데이터들로 학습을 진행하였고, 균열 탐지 테스트 시에 다양한 고도에서 촬영한 이미지를 적용시킨 결과 균열 탐지 적용이 원활하게 이루어짐을 확인할 수 있었다. 이를 통해 Fig. 17과 같이 여러 차선을 한번에 분석할 수 있으며 이를 통해 효과적인 포장상태 분석이 가능할 것으로 보인다.



Fig. 16. Drone Location Information

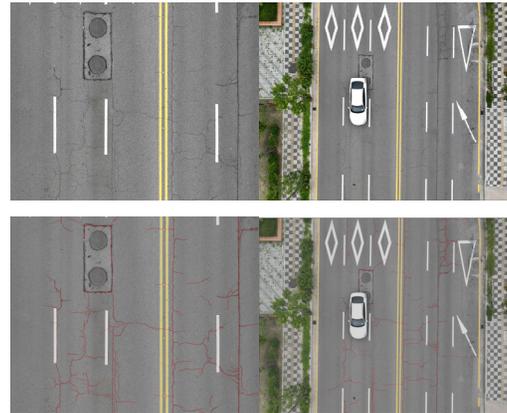


Fig. 17. Multiple Crack Detection Results

6. 결론

기존 진행된 연구 및 사례의 경우, 조사 인력의 접근성과 육안조사 한계점을 보완하기 위해 드론을 활용하여 도로포장 상태의 조사를 수행하여 이를 개선하거나 3D 스캐닝을 통해 변형의 정도를 확인하는 등의 노력을 수행하였다.

드론에 센서를 탑재하여 탐사 플랫폼으로 활용하

는 것은 학계에서 최신의 연구 트렌드이며 특히 드론촬영과 딥러닝을 활용한 도로상태 조사는 참신성이 상당하다. 본 방법을 활용하여 측정된 데이터는 빠른 시간내에 분석이 가능할 것으로 예상되며 도로상태 분석방법에 새로운 패러다임을 제공할 것으로 예상된다.

본 연구에서는 촬영고도 20m에서 드론을 통해 획득한 이미지데이터를 통해 도로포장의 이미지데이터를 수집하여 촬영구간 균열의 명확한 색상처리를 가능한 것을 확인하였고, 프로그램 해석을 통하여 균열율을 제시하였다. 향후 러팅, 포트홀 등의 도로파손 형상과 식생대와 같은 도로 이미지 촬영 및 데이터 정립을 수행하여 도로구조물 유지보수의 효율성을 높이는 접근방안으로 제시될 것이 기대된다.

본 연구에서는 드론을 활용해 촬영한 이미지데이터를 바탕으로 딥러닝을 실시하여 도로 균열을 탐지할 수 있는 모델을 개발하였으며 이에 대한 결론은 다음과 같다.

첫째, DJI사의 Mavic 3 드론을 활용하여 20m의 고도에서 촬영을 진행하였고, 이를 통해 확보한 이미지 데이터를 활용해 딥러닝 학습 알고리즘으로 Swin Transformer, Upernet을 사용하여 학습을 진행하였고, Class Weight와 Loss Function을 변형시켜가며 모델의 고도화를 실시하였다.

둘째, 본 연구에서 개발한 균열 탐지 모델을 이용해 균열율을 계산하였고, 그와 더불어 특정 도로구간의 위치정보와 균열율을 계산할 수 있는 모델을 개발하였고 향후 이를 적용하여 원격으로 도로파손 구간을 탐지할 수 있을 것으로 기대된다. 현재 균열 탐지를 위해 사용하고 있는 수동 기술에 비해 경제성이 뛰어나다.

셋째, 학습된 모델을 바탕으로 촬영한 이미지를 사용, 이미지 내에서 확인되는 도로 균열에 대한 탐지를 실시하였고 20m 고도 뿐만 아니라 여러 고도에서 촬영한 이미지 데이터도 탐지 가능한 것을 확인하

였다. 이를 바탕으로 다차선을 한번에 촬영하여 균열 분석을 하여 도로 유지관리에 적용할 수 있다.

도로포장에 대한 드론 촬영을 실시하고 이미지데이터를 활용하여 균열에 대한 라벨링을 실시할 경우, 현재 현장에서 활용되는 차륜형 포장조사장비를 통한 조사와 비교하였을 때 인건비, 유류대 등의 경제적인 부분의 확보가 가능하다. 향후 추가적인 연구를 통하여 균열탐지모델의 고도화를 사물인터넷(IoT)과의 융합으로 이루었을 때 소파보수(Patching)나 포트홀(Pothole)의 탐지가 가능하다.

참고문헌

1. 김담룡(2020), “드론을 활용한 도로포장 평탄성 조사에 관한 연구”, 석사학위논문, 서울과학기술대학교.
2. 성진욱·정기성(2021), “신경망분석을 활용한 비주택 고령자 1인 가구의 공유주택 거주의향에 미치는 주요 요인 분석”, 『LHI Journal』, 12(1): 57~68.
3. 한국건설기술연구원(2017). “무인항공기를 활용한 도로관리 효율화 방안 최종보고서”, 경기.
4. Guo, F., Y. Qian, J. Liu and H. Yu (2022), “Pavement crack detection based on transformer network”, *Automation in Construction*, 145, 104646.
5. Lin, T., P. Dollar, R. Girshick, K. He, B. Hariharan and S. Belongie (2017), “Feature Pyramid Networks for Object Detection”, *IEEE Conference on Computer Vision and Pattern Recognition*, 2117~2125.
6. Liu, Z., Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin and B. Guo (2021), “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows”, *2021 IEEE/CVF International Conference on Computer Vision*, 9992~10002.
7. Ren, S., K. He, R. Girshick and J. Sun (2017), “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137~1149.
8. Salehi, S., D. Erdogmus and A. Gholipour (2017), “Tversky Loss Function for Image Segmentation using 3D Fully Convolutional Deep Networks”,

- Machine Learning in Medical Imaging*, 10541, 379~387.
9. Shrivastava, A., A. Gupta and R. Girshick (2016), "Training Region-based Object Detectors with On-line Hard Example Mining", *Computer Vision and Pattern Recognition*, 761~769.
 10. Xiao, T., Y. Liu, B. Zhou, Y. Jiang and J. Sun (2018), "Unified Perceptual Parsing for Scene Understanding", *Lecture Notes in Computer Science*, 11209, 432~448.
 11. Zaimis, U. (2019). "Use of Drone Based It System for Road Pothole Detection and Volume Calculation", *Industry 4.0*, 4(5), 223~227.
 12. Zhang, C. and A. Elaksher (2012). "An Unmanned Aerial Vehicle-based Imaging System for 3D Measurement of Unpaved Road Surface Distresses", *Computer-Aided Civil and Infrastructure Engineering*, 27: 118~129.

요약

드론은 국토조사, 수송, 해양, 환경, 방재, 문화재, 건설 등 다양한 분야에서 활용되고 있다. 또한 사물인터넷(Internet of Things), 인공지능(Artificial Intelligence) 등과 관련하여 4차 산업 혁명의 핵심기술을 검증하고 적용시킬 수 있는 기술로 떠오르고 있다. 본 연구에서는 드론을 활용하여 균열을 자동으로 탐지할 수 있는 딥러닝 모델을 개발하고자 한다. 딥러닝 학습을 위한 이미지 데이터는 Mavic3 드론을 이용하여 수집하였고 촬영고도는 20m, 7배율로 촬영하였다. 촬영 시 약 2m/s의 속도로 전진하여 영상을 찍고, 프레임을 추출하는 식으로 데이터를 수집하였다. 이런식으로 수집한 데이터를 통해 딥러닝 학습을 진행하였다. 본 연구에서는 딥러닝 학습모델로 Backbone으로는 Swin Transformer, Architecture로 UperNet을 사용하였다. 약 800장의 라벨링 된 데이터를 Augmentation 기법으로 데이터 양을 증가시키고 3차에 걸쳐 학습을 진행하였다. 1차와 2차 학습 시 Cross-Entropy loss function을 사용하였고 3차 학습 시 Tversky Loss Function을 사용하였다. 학습결과, 균열 탐지와 균열율을 계산할 수 있는 모델을 개발하였다. 또한, 드론의 위치 정보를 이용해 특정 도로의 한 차선 균열율을 계산할 수 있는 모델을 개발하였다. 향후 추가적인 연구를 통하여 균열탐지모델의 고도화를 사물인터넷(IoT)과의 융합으로 이루어졌을 때 소파보수(Patching)나 포트홀(Pothole)의 탐지가 가능할 것으로 보인다. 또한 드론의 실시간 탐지 업무수행으로 포장 유지보수구간에 대한 탐지를 신속하게 확보할 수 있을것으로 기대된다.

주제어: 딥러닝, Swin Transformer, Backbone, Loss function, 드론