

사전 학습된 딥러닝 모델들의 피쳐 레벨 앙상블을 이용한 포트홀 검출 기법 연구

신예은⁰, 김인기*, 김범준*, 전영훈**, 곽정환(교신저자)**

⁰한국교통대학교 소프트웨어학과,

*한국교통대학교 교통에너지융합학과,

**한국교통대학교 소프트웨어학과

e-mail: 2144402@ut.ac.kr⁰, cv2@kakao.com*, zhfkskseh@gmail.com*,

0hoon.jeon@gmail.com**, jgwak@ut.ac.kr**

Research on Pothole Detection using Feature-Level Ensemble of Pretrained Deep Learning Models

Ye-Eun Shin⁰, Inki Kim*, Beonjun Kim*, Younghoon Jeon**, Jeonghwan Gwak(Corresponding Author)**

⁰Dept. of Software, Korea National University of Transportation,

*Dept. of IT·Energy Convergence, Korea National University of Transportation,

**Dept. of Software, Korea National University of Transportation

● 요약 ●

포트홀은 주행하는 자동차와 접촉이 이뤄지면 차체나 운전자에게 충격을 주고 제어를 잃게 하여 도로 위 안전을 위협할 수 있다. 포트홀의 검출을 위한 국내 동향으로는 진동을 이용한 방식과 신고시스템 이용한 방식과 영상 인식을 기반한 방식이 있다. 이 중 영상 인식 기반 방식은 보급이 쉽고 비용이 저렴하나, 컴퓨터 비전 알고리즘은 영상의 품질에 따라 정확도가 달라지는 문제가 있었다. 이를 보완하기 위해 영상 인식 기반의 딥러닝 모델을 사용한다. 따라서, 본 논문에서는 사전 학습된 딥러닝 모델의 정확도 향상을 위한 Feature Level Ensemble 기법을 제안한다. 제안된 기법은 사전 학습된 CNN 모델 중 Test 데이터의 정확도 기준 Top-3 모델을 선정하여 각 딥러닝 모델의 Feature Map을 Concatenate하고 이를 Fully-Connected(FC) Layer로 입력하여 구현한다. Feature Level Ensemble 기법이 적용된 딥러닝 모델은 평균 대비 3.76%의 정확도 향상을 보였으며, Top-1 모델인 ShuffleNet보다 0.94%의 정확도 향상을 보였다. 결론적으로 본 논문에서 제안된 기법은 사전 학습된 모델들을 이용하여 각 모델의 다양한 특징을 통해 기존 모델 대비 정확도의 향상을 이룰 수 있었다.

키워드: 포트홀 검출(Pothole Detection), 사전 학습(Pre-Trained), 피쳐 레벨 앙상블(Feature-Level Ensemble), 몬테 카를로 드롭아웃(Monte Carlo Dropout), 완전 연결 계층(Fully-Connected Layer)

I. Introduction

포트홀은 빗물이나 눈이 스며든 도로 포장재에 차량의 무게가 가해져 균열과 파손이 일어나며 형성되는 냄비처럼 파인 구멍을 말한다. 포트홀은 차량이 주행 중에 미리 피하지 못하여 접촉했을 경우 차량에 비포장도로를 지나는 듯한 충격을 주고 차량의 제어를 잃게 만들어 매우 위험하다. 포트홀의 발생 위치는 예측할 수 없으며, 강우량 또는 적설량에 따라 도로에 스며드는 물기가 증가하게 되면 발생 건수도 많아지게 된다. 따라서 예측할 수 없는 포트홀의 발생은 도로 위 안전과 직결되므로 포트홀 검출 시스템의 개발 및 추적

관리가 필요하다.

본 논문에서는 포트홀 검출 기법에 대한 국내 연구 동향에 대해 살펴보고, 영상이나 사진을 통해 포트홀의 유무를 검출할 수 있는 딥러닝 모델에 대해 정확도를 높일 수 있는 기법을 제안한다.

II. Related Work

2013년도 서울시에선 노선버스에 증력 가속도 센서를 설치하여 포트홀을 검출하는 시스템을 시험 적용하였다[1]. 이 시스템은 포트홀로 인한 충격에서 오는 진동을 통해 포트홀을 검출하는 방식이다. 증력가속도 센서는 특수 장비에 비해 상대적으로 비용이 저렴하지만, 포트홀의 크기나 위치를 검출해 낼 수 있는 정확성이 떨어지며 포트홀에 차량이 직접적으로 닿지 않는다면 검출할 수 없다. 또한, 2014년도 서울시에선 택시에 포트홀 카드 결제기에 신고시스템을 장착시켜 포트홀을 검출하였다[2]. 포트홀을 발견하면 카드 결제기와 같은 GPS 단말기 버튼을 이용하여 신고하는 방식이다. 하지만 버튼을 누른 지점과 실제 포트홀의 발생지점의 거리의 오차가 발생하는 경우가 많았다.

이러한 방식들의 문제점을 해결하기 위해 한국건설기술연구원에서는 컴퓨터 비전의 Saliency Map 기반 포트홀 검출 알고리즘이 연구되었다[3]. 이는 영상 인식을 기반으로 하여 블랙박스나 개인용 카메라를 통해 자동으로 포트홀을 분류하는 방식이다. 도로의 영상을 흑백으로 변경한 후 배경과 포트홀을 구분해 내어 파손 여부를 검출한다. 영상 인식 기반 검출 기법은 보급이 간단하고 비용이 저렴하다는 장점이 있지만, 카메라에 촬영된 날씨와 밤낮 영상의 품질에 따라 정확도가 달라질 수 있다.

최근에는 이러한 영상 인식 기반의 정확도를 높이기 위해 인공지능을 이용한 자동 포트홀 검출 시스템에 관한 연구가 활발히 이루어지고 있다. 딥러닝 모델을 이용한다면 정확도가 달라지는 문제를 다양한 특징의 학습을 통해 보완할 수 있기 때문이다. 따라서 딥러닝 모델을 이용하여 정확도를 향상시킬 기법에 대한 연구가 요구된다.

III. Proposed Method

1.1 실험 환경

본 실험에 사용된 데이터셋은 인터넷에서 Crawling된 이미지들과 직접적으로 촬영한 이미지를 통해 구성하였다. 또한, 포트홀을 바라보는 시점을 달리하여 Zoom-out 된 시점의 이미지 600장과 블랙박스와 같이 Zoom-in 된 시점의 이미지 600장, 총 1200장을 사용하였다. Fig. 1은 Zoom-out 된 시점과 Zoom-in 된 시점의 예시이다. 학습을 위한 데이터의 비율은 Train 840장, Validation 240장, Test 120장을 사용하였다. 개발 환경 시스템은 Windows 21H2와 CUDA 11.3, PyTorch 1.12.0 버전에서 진행되었다.



Fig. 1. Sample images: 1) Zoom-out, 2) Zoom-in

1.2 사전 학습된 모델들의 성능

Table 1은 사전 학습된 CNN 모델들의 Test 데이터셋에 대한 성능을 실험하여 정리한 표이다. 실험에 사용된 CNN 모델은 PyTorch에 내장된 ImageNet으로 사전 학습된 딥러닝 모델을 사용하였고, 128×128 이미지 크기의 도로 및 포트홀 Training 데이터를 학습하고 Validation으로 검증하여 딥러닝 모델을 구축하였다. 그리고 Test 데이터를 통해 Test Accuracy와 Test Loss를 측정하였다.

Table 1. Comparison of general CNN models

Model	Test Accuracy	Test Loss
AlexNet[4]	93.3%	0.2185
DenseNet121[5]	95.0%	0.1694
DenseNet161[5]	94.2%	0.1587
ShuffleNet[6]	95.8%	0.1313
MobileNet[7]	92.5%	0.1996
MnasNet[8]	83.3%	1.0815
InceptionV3[9]	91.7%	0.3200
ResNet50[10]	94.2%	0.1631
ResNet101[10]	94.2%	0.2730
ResNeXt50[11]	94.2%	0.1908
ResNeXt101[11]	92.5%	0.2490
VGG16[12]	89.2%	0.2683
VGG19[12]	95.0%	0.2746
EfficientNet[13]	93.3%	0.4082
SqueezeNet[14]	94.2%	0.1377
SENet154[15]	95.8%	0.1510
Xception[16]	95.8%	0.2749
Average	93.2%	0.2747

Table 1의 Test Accuracy를 보면 95.8%의 ShuffleNet[6]과 SENet154[15]와 Xception[16] 모델이 높은 정확도를 보였다. 또한, Test Loss는 0.1313의 ShuffleNet[6]이 가장 적은 Loss를 보였으며, 0.1377의 SqueezeNet[14], 0.1587의 DenseNet161[5] 순으로 적은 Loss를 보였다. Top-3 모델은 정확도 기준으로 선정하였으며, Top-1 모델은 Top-3 모델의 정확도가 같기 때문에 Loss가 가장 적은 ShuffleNet이 실험한 모델 중 Top-1 모델임을 도출하였다.

3. Feature-Level Ensemble 기반 모델 및 성능 평가

본 논문에선 Feature Level Ensemble(FLE) 기법을 모델을 제안한다. FLE이란, 각 분류기의 Feature Map을 하나로 Concatenate하여 정확도를 향상시키는 기법이다. FLE기반 딥러닝 모델을 설계하기 위해서 사전 학습된 CNN 모델 중에서 앙상블을 실행할 딥러닝 모델 세 개를 선정하였다. 강인한 FLE 기반 딥러닝 모델의 설계를 위해서 정확도 기준 Top-3 모델인 ShuffleNet, SENet154, Xception 모델을 선택하였다. FLE 모델은 선택한 세 모델의 Feature Level Ensemble을 진행하여 성능을 측정한다. Fig. 2는 제안한 FLE 모델의 아키텍처이다. FLE 모델의 아키텍처는 Top-3 모델들의 마지막 레이어의 Feature Map을 뽑아내서 크기를 맞춰 Concatenate 한 뒤 Fully-Connected(FC) Layer로 입력하는 형태이다.

FLE 모델의 실험은 사전 학습된 CNN 모델의 실험과 같은 형식으로 진행하였다. Table 2는 구축한 FLE 모델에 Test 데이터를 통해

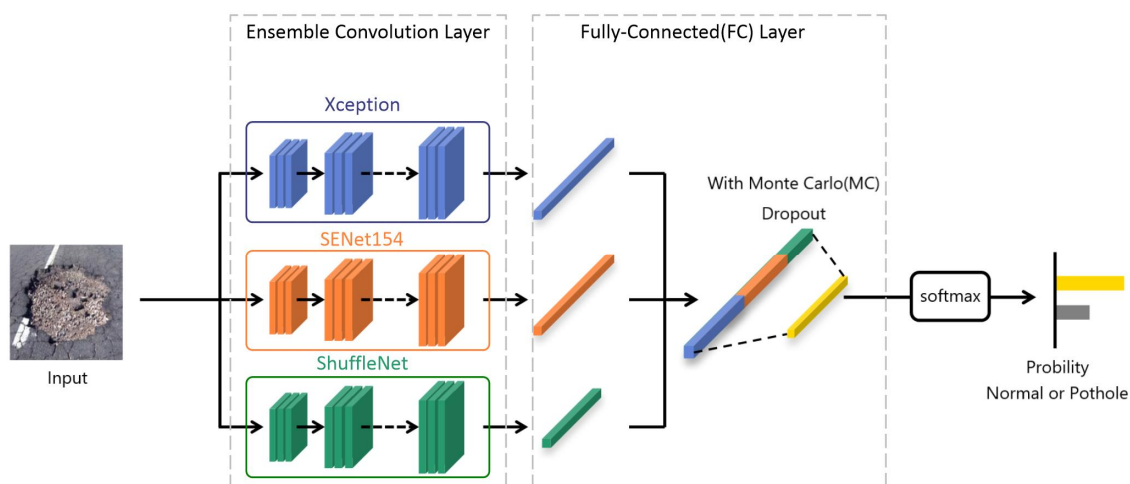


Fig. 2. Overall architecture of proposed method

Test Accuracy와 Test Loss를 출력하여 Top-3 모델과 비교한 표이다. FLE 모델은 Test 데이터에 대해 96.7%의 정확도와 0.2311의 Loss를 보였다. 이는 사전 학습된 CNN과 비교하였을 때 가장 높은 Test Accuracy를 달성하였다. 이는 사전 학습된 CNN 모델들의 평균 정확도와 비교하였을 시 3.76%가 증가한 것을 알 수 있었고, 높은 성능을 가진 Top-3 모델(ShuffleNet, SENet154, Xception)과 비교하였을 시 각각 정확도가 0.94%가 증가한 것을 알 수 있었다.

Table 2. Comparison of top-3 classifiers

Model	Test Accuracy	Test Loss
ShuffleNet[6]	95.8%	0.1313
SENet154[15]	95.8%	0.1510
Xception[16]	95.8%	0.2749
Ensemble	96.7%	0.2311

IV. Conclusions

실험을 통해 FLE 기법을 통해 각 사전 학습된 모델들이 추출하는 특징들을 Ensemble 함으로써 포트홀 부분의 가중치를 부각할 수 있으며, 상대적으로 관심을 두지 않는 부분의 중요도는 낮추어 포트홀의 특징을 추출하는 성능이 향상된 것을 알 수 있었다. 결과적으로 기존 사전 학습된 모델은 같은 각 네트워크에 기반한 Feature만을 추출할 수 있었지만, FLE 모델은 다양한 네트워크를 사용함에 따라 Global Feature, Local Feature를 적절하게 선택해 낼 수 있다는 것을 입증하였다. 제안된 기법은 블랙박스나 개인용 카메라에 내장하여 포트홀의 유무를 파악하고, 도로의 유지와 보수를 위해 포트홀을 추적 관리하는 데 도움이 될 것이다. 추후로 각 사전 학습된 모델별 조합을 고려하여 낮은 파라미터 수로 강인한 성능을 낼 수 있는 딥러닝 모델을 연구할 예정이다.

ACKNOWLEDGEMENT

This research was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea (NRF) funded by the Ministry of Education [2021RIS-001(1345341783)]

REFERENCES

- [1] Seoul Metropolitan Government all rights reserved, <https://news.seoul.go.kr/safe/archives/21571>
- [2] Seoul Metropolitan Government all rights reserved, <https://news.seoul.go.kr/safe/archives/22440>
- [3] Jo, Young-Tae, and Seung-Ki Ryu. "Pothole detection algorithm based on saliency map for improving detection performance." The Journal of The Korea Institute of Intelligent Transport Systems 15.4 (2016): 104-114.
- [4] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Communications of the ACM 60.6 (2017): 84-90.
- [5] Huang, Gao, et al. "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
- [6] Zhang, Xiangyu, et al. "Shufflenet: An extremely efficient convolutional neural network for mobile devices." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
- [7] Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision

- applications." arXiv preprint arXiv:1704.04861 (2017).
- [8] Tan, Mingxing, et al. "Mnasnet: Platform-aware neural architecture search for mobile." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.
- [9] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [10] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
- [11] Xie, Saining, et al. "Aggregated residual transformations for deep neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
- [12] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
- [13] Tan, Mingxing, and Quoc Le. "Efficientnet: Rethinking model scaling for convolutional neural networks." International conference on machine learning. PMLR, 2019.
- [14] Iandola, Forrest N., et al. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size." arXiv preprint arXiv:1602.07360 (2016).
- [15] Hu, Jie, Li Shen, and Gang Sun. "Squeeze-and-excitation networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
- [16] Chollet, François. "Xception: Deep learning with depthwise separable convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.