

도로교통 이머징 리스크 탐지를 위한 AutoML과 CNN 기반 소프트 보팅 앙상블 분류 모델

전병욱¹, 강지수², 정경용^{3*}

¹경기대학교 AI컴퓨터공학부 학부생, ²경기대학교 컴퓨터과학과 석사과정, ³경기대학교 AI컴퓨터공학부 교수

AutoML and CNN-based Soft-voting Ensemble Classification Model For Road Traffic Emerging Risk Detection

Byeong-Uk Jeon¹, Ji-Soo Kang², Kyungyong Chung^{3*}

¹Student, Division of AI Computer Science and Engineering, Kyonggi University

²Student, Department of Computer Science, Kyonggi University

³Professor, Division of AI Computer Science and Engineering, Kyonggi University

요약 겨울철 도로 결빙으로 인한 사고는 대부분 큰 사고로 이어진다. 이는 운전자가 도로의 결빙을 사전에 자각하기 어렵기 때문이다. 본 연구에서는 AutoML과 CNN의 앙상블 모델을 이용하여 도로교통 이머징 리스크를 정확하게 탐지하는 방법을 연구한다. 비정형 데이터인 이미지를 이용한 CNN 이미지 특징 추출 기반 도로교통 이머징 리스크 분류 모델과 정형 데이터인 기상 데이터를 이용한 AutoML 기반 도로교통 이머징 리스크 분류 모델을 각각 학습시킨다. 그 후 모델들에서 도출된 확률값을 입력하여 CNN 기반 분류 모델을 보완하도록 앙상블 모델을 설계한다. 이를 통해 도로교통 이머징 리스크 분류 성능을 향상하고 더 정확하고 빠르게 운전자에게 경고하여 안전한 주행이 가능하도록 한다.

주제어 : 교통안전, 도로교통 이머징 리스크, 머신러닝, CNN, AutoML, 앙상블 모델

Abstract Most accidents caused by road icing in winter lead to major accidents. Because it is difficult for the driver to detect the road icing in advance. In this work, we study how to accurately detect road traffic emerging risk using AutoML and CNN's ensemble model that use both structured and unstructured data. We train CNN-based road traffic emerging risk classification model using images that are unstructured data and AutoML-based road traffic emerging risk classification model using weather data that is structured data, respectively. After that the ensemble model is designed to complement the CNN-based classification model by inputting probability values derived from of each models. Through this, improves road traffic emerging risk classification performance and alerts drivers more accurately and quickly to enable safe driving.

Key Words : Traffic Safety, Road Traffic Emerging Risk, Machine Learning, CNN, AutoML, Ensemble Model

1. 서론

도로 결빙으로 인한 교통사고는 일반적인 교통사고

에 비해 큰 사고로 이어진다. 도로 결빙으로 인한 교통 사고는 1천 건 당 사망자 수가 35.9명이다. 이는 평균 21.6명에 비해 66.2% 높은 수치이다[1]. 결빙 노면 특

*This work is supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant 21CTAP-C157011-02).

*Corresponding Author : Kyungyong Chung (dragonhci@gmail.com)

Received May 15, 2021

Revised July 5, 2021

Accepted July 20, 2021

Published July 28, 2021

성상 눈에 잘 띄지 않아 운전자가 사전에 자각하기 어렵고 마찰계수가 특히 낮으므로 차량을 제어하기 어렵기 때문이다. 이에 따라 CNN, RNN, RCNN, LSTM 등의 컴퓨터 비전 기술을 활용하여 주행 영상이나 이미지를 통해 도로교통 이머징 리스크를 실시간으로 탐지하는 기술에 관한 연구가 진행되고 있다[2]. 또한 기상관측자료와 교통사고자료를 분석한 연구에 따르면 도로 결빙으로 인한 교통사고는 기상상태와 큰 인과관계가 있다. 따라서 기상 데이터를 이용하여 도로교통 이머징 리스크를 예측하는 기술에 관한 연구가 진행되고 있다[3]. 그러나 도로교통 이머징 리스크가 예측되는 경우 같은 기상 상태라도 도로교통 이머징 리스크가 발생하는 지점은 특정 일부 구간이다. 따라서 도로교통 이머징 리스크를 실시간으로 분류하는 것과는 거리가 멀다는 한계가 존재한다.

컴퓨터 비전에서 많이 사용하는 CNN과 노면 이미지만을 이용하여 도로교통 이머징 리스크를 분류하는 경우 한계가 존재한다. CNN 기반 모델은 노면 이미지에서 나타난 특징이 유사한 경우 잘 분류하지 못한다. 또한 어둡거나 복잡한 배경으로 인해 노면 이미지로부터 특징을 추출하기 어려운 경우 잘 분류하지 못한다[4]. 이를 해결하여 분류 성능을 향상시키기 위해서는 이미지 전처리 과정 및 모델 구조를 더 복잡하게 구성해야 한다. 이는 연산 성능이 제한적인 실제 차량이나 블랙박스 시스템에 적용되었을 때, 분류 속도를 늦추는 요인이 된다. 그러므로 비정형 데이터인 이미지를 이용한 CNN 기반 도로교통 이머징 리스크 분류 모델을 정형 데이터인 기상 데이터를 이용한 도로교통 이머징 리스크 분류 모델로 보완이 필요하다[5].

정형 데이터에 효과적인 AutoML 기반 모델은 기상 데이터를 이용하여 도로교통 이머징 리스크 여부의 효과적인 예측이 가능하다. 따라서 본 연구에서는 이미지를 입력받아 도로교통 이머징 리스크를 분류하는 CNN 기반 분류 모델과 기상 데이터를 입력받아 도로교통 이머징 리스크를 분류하는 모델을 각각 학습시킨다. 학습된 각 모델에서 얻은 확률값들을 소프트 보팅 앙상블 기반 분류 모델에 입력하여 CNN 기반 분류 모델을 보완한다. 제안한 모델에 도로 이미지와 당시의 기상 데이터를 함께 입력하면 도로 이미지 또는 기상 데이터만으로 도로교통 이머징 리스크를 분류하는 각각의 모델에 비해 더 향상된 성능을 보인다.

2. 관련 연구

2.1 CNN 기반 이미지 특징 추출

CNN은 영상 분석 및 이미지의 특징 추출에 사용되는 인공신경망의 한 종류이다. 학습 매개변수가 적기 때문에 훈련이 용이하다. 교통 등 다양한 분야에서 영상을 통해 실시간으로 분류, 탐지하는 곳에 주로 활용된다. 크게 컨볼루션 레이어, 풀링 레이어로 구성되며, 과적합을 방지하기 위한 드롭아웃 레이어가 존재한다[6]. 컨볼루션 레이어는 합성곱 연산을 수행하여 특정 영역을 강조하는 특징 맵을 출력한다. 풀링 레이어는 특징 맵을 입력받은 후 크기를 줄이고 특정 데이터를 강조한다. 컨볼루션 레이어와 풀링 레이어를 거쳐 입력 데이터에서 특징을 추출한 후 텐스 레이어와 활성화 함수 레이어를 거쳐 해당 이미지를 분류한다[6]. 그러나 특징이 유사하거나 밤과 같이 이미지에서 특징을 추출하기 어려운 경우 분류 성능에 한계가 존재한다. Fig. 1은 CNN을 통해 이미지로부터 특징 추출을 한 후 텐스 레이어를 통해 최종적으로 분류 결과를 출력하는 기본적인 CNN 분류 모델의 구조를 나타낸다.

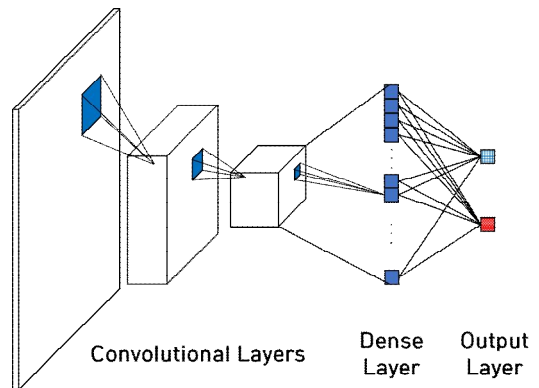


Fig. 1. Structure of Vanilla Convolutional Neural Network

2.2. 정형 데이터 분석을 위한 AutoML

AutoML(Automated Machine Learning)은 자신의 환경에 대해 스스로 학습하는 신경망 프로세스이다[7]. 머신러닝 개발 과정에서 여러 알고리즘을 적용하며 데이터 셋에 최적화된 알고리즘을 자동으로 탐색한다. 학습 완료 후 가장 결과가 우수한 상위 N개의 알고리즘을 선택하여 모델을 설계한다. 대표적인 머신러닝 기

법인 분류, 회귀, 예측, 군집에 적용이 가능하다. AutoML은 AutoKeras, AutoGLuon, pycaret 등의 다양한 패키지를 통해 사용이 가능하다[8]. 정형 데이터의 경우 AutoML이 데이터 전처리부터 학습까지 자동으로 진행한다. 여러 분석 알고리즘들을 적용하여 성능이 좋은 모델들을 찾고 이들을 앙상블하여 입력 데이터에 최적화된 모델이 자동으로 도출된다. 따라서 모델 개발에 소요되는 시간이 감소하고 성능이 뛰어난 것이 특징이다. 특히 기상 데이터와 같이 정형 데이터인 경우 적용 시 성능이 뛰어난 모델을 손쉽게 구현할 수 있다.

2.3. 소프트 보팅 앙상블 학습

앙상블 학습은 여러 개의 학습 모델을 생성하고 그 결과들을 결합하여 과적합을 방지하고 더욱 정확한 예측을 도출하는 학습 기법이다[9]. 보팅, 배깅, 부스팅 등의 다양한 기법이 존재한다. 그 중 보팅 기법은 여러 모델을 통해 얻은 예측값들을 대상으로 다수결 투표를 하여 최종 클래스를 예측한다. 보팅 기법에는 하드 보팅과 소프트 보팅이 존재한다. 하드 보팅은 각 모델의 예측 결과를 이용하여 다수결 투표를 진행한다. 다수의 모델이 출력한 결과가 앙상블 모델의 최종 출력이 된다. 소프트 보팅은 각 모델이 예측한 확률과 가중치를 이용해 클래스별 최종 확률을 계산한다[10]. 따라서 소프트 보팅은 하드 보팅에 비해 각 모델의 결과값들을 유연하게 적용이 가능하다. 식 (1)은 소프트 보팅을 통해 클래스를 분류하는 식이다. 변수 p 는 각 학습 모델을 통해 도출된 클래스별 확률값을 나타낸다. 변수 w 를 통해 각 학습 모델별로 도출된 확률값에 추가할 가중치를 설정한다. 최종적으로 가장 큰 확률을 가진 클래스를 출력한다.

$$\hat{y} = \arg \max_i \sum_{j=1}^M w_j p_{ij} \quad (1)$$

3. 도로교통 이머징 리스크 탐지를 위한 AutoML과 CNN 기반 앙상블 분류 모델

본 연구에서 제안하는 도로교통 이머징 리스크 탐지를 위한 AutoML과 CNN 기반 앙상블 분류 모델은 세 단계로 진행된다. 첫 번째 단계는 CNN 기반 분류 모델을 도로 이미지 데이터셋으로 학습시킨다. 두 번째 단

계는 AutoML 기반 분류 모델을 기상 데이터셋을 이용하여 학습시킨다. 앙상블 기반 모델은 각 모델을 통해 도출된 도로교통 이머징 리스크 확률값을 입력받는다. 세 번째 단계는 소프트 보팅 기법을 이용하여 최종적으로 도로교통 이머징 리스크 여부를 분류한다. Fig. 2는 제안하는 AutoML과 CNN 기반 앙상블 분류 모델을 나타낸다.

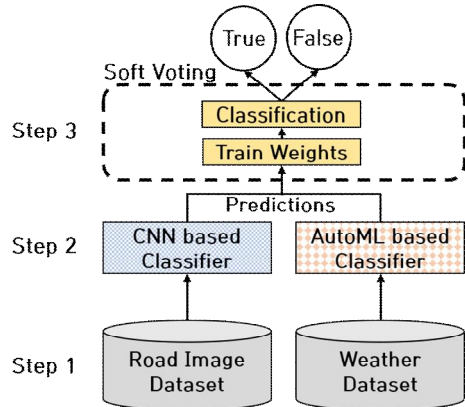


Fig. 2. Structure of AutoML and CNN-based Ensemble Classification Model

3.1 데이터 수집 및 전처리

본 연구에서는 비정형 데이터와 정형 데이터를 모두 사용하는 AutoML과 CNN 기반 앙상블 분류 모델을 설계한다. CNN 기반 도로교통 이머징 리스크 분류 모델을 위하여 정상 도로의 노면과 결빙 도로의 노면 이미지 데이터를 수집한다. 이미지는 차량의 주행 영상에서 프레임별로 추출한다. 이 때 본 연구에서 제안하는 소프트 보팅 앙상블 기반 분류 모델을 위하여 차량의 주행 영상은 해당 영상이 촬영될 당시의 기상 상태를 알 수 있는 영상을 사용한다. 정상 도로의 노면 영상은 차량의 블랙박스로 촬영된 영상을 수집한다. 결빙 도로의 노면 영상은 결빙 도로로 인해 사고가 발생한 차량의 블랙박스로 촬영된 영상을 웹에서 크롤링하여 수집한다. 노면 이미지 데이터는 총 986개를 수집한다. 그 중 결빙 노면 이미지는 313개, 정상 노면 이미지는 673개로 구성된다. 수집된 이미지는 노면 부분만을 크롭한 후 크기를 조정하여 224x224의 크기로 전처리한다.

AutoML 기반 도로교통 이머징 리스크 분류 모델을 위해 기상 데이터와 사고 데이터를 수집하여 데이터를 구축한다. 기상 데이터는 기상청 기상자료 개방포털을

통해 2017년부터 2019년까지 관측된 기상 데이터를 수집한다. 사고 데이터는 교통사고분석시스템에서 동기간 전국에서 발생한 사고 데이터를 수집한다[11, 12]. 기상 데이터 중 기온, 습도, 지면 온도 데이터를 사용한다. 사고 데이터에서 노면 상태 정보를 추출하여 시간 및 지역 정보를 통해 기상 데이터와 합병한다. 이상치나 결측값은 제거한다. 전체 데이터 샘플의 개수가 약 65만 개로 매우 많다. 또한, 도로 결빙 시 사고와 그 외의 사고 데이터의 비율이 약 1:48으로 불균형하다. 이는 모델의 과적합을 유발하여 성능 저하를 유발한다. 따라서 다수 클래스의 데이터를 무작위로 삭제하여 과적합을 방지한다. 이를 위해 랜덤 언더샘플링 기법을 적용하여 두 클래스의 데이터 개수를 동일하게 구성한다.

3.2 앙상블 모델을 위한 CNN 기반 분류 모델과 AutoML 기반 분류 모델

이미지를 통해 도로교통 이머징 리스크를 분류하기 위해 CNN 기반 분류 모델을 사용하여 특징을 추출한다. 모델의 구조는 AlexNet을 기반으로 3개의 CNN 레이어와 텐스 레이어로 설계한다. CNN 레이어는 컨볼루션, 풀링, 드롭아웃으로 구성되며 텐스 레이어는 2개의 1024 레이어와 2 레이어의 출력 레이어를 포함한다. Fig. 3은 CNN 기반 분류 모델의 구조를 나타낸다.

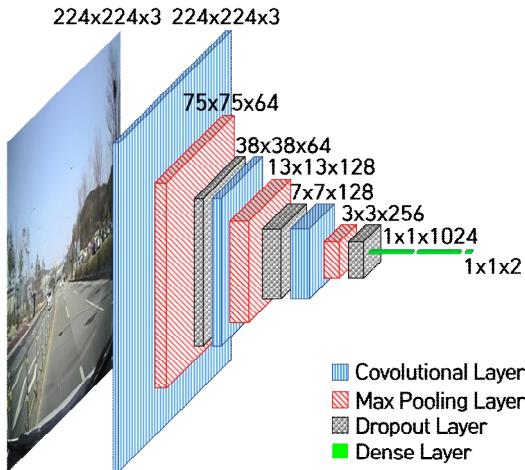


Fig. 3. Structure of CNN Model for Road Traffic Emerging Risk Detection

소프트맥스 함수를 활성 함수로 사용한다. 소프트맥

스 함수는 출력이 0에서 1 사이이며 클래스별 출력 총합이 1이다. 이러한 특성 덕분에 확률로 활용된다[13]. CNN 기반 분류 모델은 비정형 데이터인 이미지를 입력받아 도로교통 이머징 리스크의 특징을 학습하고 확률값을 출력하여 데이터를 정형화한다. 이를 소프트 보팅 앙상블 기반 분류 모델에서 입력으로 활용한다.

CNN 기반 분류 모델을 보완하기 위해 기상 데이터를 사용하여 AutoML 기반 도로교통 이머징 리스크 분류 모델을 설계한다. pycaret 패키지의 AutoML을 사용하여 여러 가지의 분류 모델을 학습시키고 가장 높은 F1-score를 도출해낸 모델 3개를 앙상블하여 모델을 생성하도록 설정한다. 생성된 모델에 기상 데이터를 입력하면 도로교통 이머징 리스크 확률값을 출력한다. 상위 세 모델인 Gradient Boosting classification model, Ada Boost Regression, Random Forest classification model의 F1-score가 각각 0.9060, 0.9038, 0.9023으로 가장 높기 때문에 AutoML은 자동으로 해당 상위 3개 모델을 앙상블하여 입력된 데이터셋에 최적화된 모델을 도출한다. 그 결과, AutoML을 통해 생성된 모델의 도로교통 이머징 리스크 분류 정확도가 0.9083, F1-score가 0.9086으로 기존 분류 모델들과 비교하면 우수한 성능을 보인다. Table 1은 AutoML을 통해 학습된 분류 모델과 생성된 모델의 성능을 나타낸다.

Table 1. Performance results of classification models learned through AutoML and ensemble models generated (Sorted by F1-score)

No.	Model	Accuracy	F1-score
1	Gradient Boosting classification model	0.9022	0.9060
2	Ada Boost Regression	0.8987	0.9038
3	Random Forest classification model	0.8997	0.9023
...			
14	Quadratic Discriminant Analysis	0.7633	0.7025
Ours	AutoML's classification model	0.9083	0.9086

3.3 소프트 보팅 앙상블 기반 도로교통 이머징 리스크 분류 모델

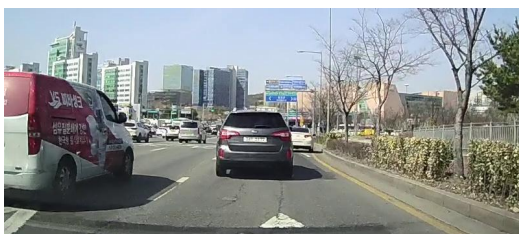
본 연구에서는 앙상블 기법 중 결과값들을 더 유연하게 적용할 수 있는 소프트 보팅 기법을 사용한다. 또한 앙상블 모델에 사용되는 분류 모델의 개수가 2개뿐

이므로 하드 보팅에 비해 소프트 보팅이 더 적합하다.

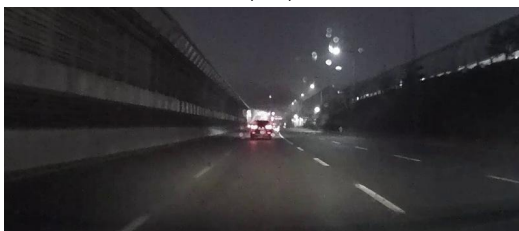
CNN 기반 분류 모델은 특징 추출을 한 후 소프트맥스 함수를 거쳐 도로교통 이머징 리스크 확률값을 출력한다. 그리고 AutoML 기반 분류 모델 또한 해당 이미지가 촬영되었을 당시의 기상 데이터를 입력하면 도로교통 이머징 리스크 확률값을 출력한다. 두 모델에서 도출된 확률값을 소프트 보팅 앙상블 기반 분류 모델에 입력한다. 가중치는 학습 데이터의 이미지와 기상 데이터를 통해 도출된 확률값을 통해 새로운 AutoML 모델을 학습하여 설정한다. 최종 분류 결과 도로교통 이머징 리스크가 예측되는 경우 운전자에게 사전에 경고한다. CNN 기반 모델은 비정형 데이터인 이미지만을 입력받으나 앙상블 기법을 통해 정형 데이터인 기상 데이터를 함께 사용하는 것이 가능하다. 이를 통해 CNN 기반 분류 모델을 보완하여 향상된 성능의 모델을 구현한다.

4. 성능평가

본 연구에서는 성능평가를 위해 데이터셋 중 30%를 테스트 데이터로 사용한다. 실험은 Google Colaboratory의 클라우드 컴퓨팅 환경에서 진행하였다. Fig. 4는 결빙이 아닌 도로의 이미지들이다. CNN 기반 분류 모델의 경우 Fig. 4(a)와 같이 맑은 날일 경우는 결빙 여부를 잘 분류한다. 그러나 Fig. 4(b)와 같이 비가 오거나 밤일 경우 잘 분류하지 못하는 한계가 존재한다.



(a) Sunny day road



(b) Rainy night road

Fig. 4. Images of roads that are not icing

Table 2는 CNN 기반 분류 모델과 제안하는 소프트

보팅 앙상블 기반 분류 모델에 Fig. 4(b) 이미지를 입력했을 때 도출된 도로교통 이머징 리스크 확률이 도출된 결과를 나타낸다.

Table 2. Probability value when inputting Fig. 4(b)

Model	Risk probability
CNN Based classification model	0.99
Ours classification model	0.10

Table 2에서는 CNN 기반 분류 모델의 도로교통 이머징 리스크 확률은 0.99로 도출되어 결빙으로 잘못 예측하였다. 그러나 제안하는 소프트 보팅 앙상블 기반 분류 모델의 도로교통 이머징 리스크 확률은 0.1로 도출되어 결빙이 아닌 것으로 정확히 예측하였다. Fig. 5는 기존 CNN 기반 모델과 본 연구에서 제안하는 소프트 보팅 앙상블 기반 분류 모델의 성능평가 결과를 나타낸다.

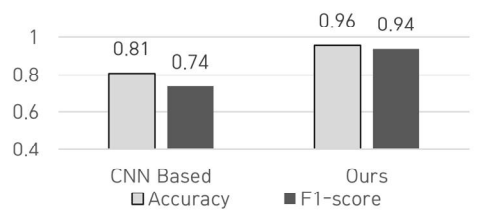


Fig. 5. Performance result of classification models

본 연구에서 제안하는 소프트 보팅 앙상블 기반 도로교통 이머징 리스크 분류 모델은 앙상블 기법을 통해 기상 데이터로 CNN 기반 분류 모델을 보완한다[14-16]. 이를 통해 제안하는 분류 모델은 기존의 CNN 기반 분류 모델과 다르게 Fig. 4(b)와 같은 이미지들을 정확히 분류한다. 따라서 이미지만을 이용한 CNN 기반 분류 모델에 비해 높은 성능을 보여준다. 성능 평가 결과, 본 연구에서 제안한 소프트 보팅 앙상블 기반 분류 모델의 정확도가 96.36%로 평가되어 81.37%인 기존 CNN 기반 분류 모델에 비해 분류 성능이 향상되었다.

5. 결론

노면 이미지와 CNN 기반 분류 모델만으로 도로의 결빙 여부를 분류하는 경우 밤일 경우나 비가 오는 경우 분류 성능의 한계가 존재한다. 본 연구에서는 AutoML

과 CNN 기반 도로교통 이머징 리스크 분류 모델과 소프트 보팅 앙상블 기반 분류 모델을 제안한다. 이를 통해 비정형 데이터를 이용하는 CNN 기반 도로교통 이머징 리스크 분류 모델을 정형 데이터를 이용하여 보완하도록 설계했다. 두 모델을 통해 도출된 도로교통 이머징 리스크 확률을 소프트 보팅 앙상블 기반 분류 모델에 입력하여 최종적으로 도로교통 이머징 리스크 여부를 분류한다. 성능평가 결과 CNN 기반 분류 모델이 잘 분류하지 못했던 이미지들을 더 잘 분류하게 되어 기존 CNN 기반 분류 모델만을 사용하여 분류할 때에 비해 분류 성능이 향상되었다. CNN 기반 분류 모델의 성능을 향상시키기 위해서 입력 이미지의 전처리 과정 및 모델 구조를 더 복잡하게 구성하지 않더라도 앙상블 모델을 통해 성능 향상이 가능하다. 따라서 연산 성능이 제한적인 실제 차량이나 블랙박스 시스템에 적용해도 정확하고 빠르게 도로교통 이머징 리스크 탐지가 가능하다. 이를 통해 더 빠르고 정확하게 운전자에게 경고하여 사고를 방지하고 안전한 주행이 가능하도록 한다.

REFERENCES

- [1] S. J. Lee. (2017). A study on factors that influence traffic accident severity in road surface freezing. *Journal of the Korean Society of Safety*, 32(6), 150-156.
DOI : 10.14346/JKOSOS.2017.32.6.150
- [2] G. Pan, L. Fu, R. Yu & M. Muresan. (2018). Winter road surface condition recognition using a pretrained deep convolutional network. *arXiv preprint arXiv:1812.06858*.
- [3] M. Kangas, M. Heikinheimo & M. Hippel. (2015). RoadSurf: a modelling system for predicting road weather and road surface conditions. *Meteorological applications*, 22(3), 544-553.
DOI : 10.1002/met.1486
- [4] L. Zhang., F. Yang, Y. D. Zhang. & Y. J. Zhu. (2016, September). Road crack detection using deep convolutional neural network. *2016 IEEE international conference on image processing (ICIP)*, 3708-3712.
DOI : 10.1109/ICIP.2016.7533052
- [5] B. U. Jeon, J. S. Kang, H. J. Kwon & K. Chung. (2021). AutoML And CNN Image Feature Extraction for Road Icing Detection, *2021 Korean Society For Internet Information Spring Conference*, 75-76.
- [6] S. S. Park, J. W. Baek, S. M. Jo & K. Chung. (2019). Motion Monitoring using Mask R-CNN for Articulation Disease Management. *Journal of the Korea Convergence Society*, 10(3), 1-6.
DOI : 10.15207/JKCS.2019.10.3.001
- [7] F. Mohr, M. Wever. & E. Hüllermeier. (2018). ML-Plan: Automated machine learning via hierarchical planning. *Machine Learning*, 107(8), 1495-1515.
DOI : 10.1007/s10994-018-5735-z
- [8] L. Zimmer. M. Lindauer & F. Hutter. (2021). Auto-Pytorch: Multi-Fidelity MetaLearning for Efficient and Robust AutoDL. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
DOI : 10.1109/TPAMI.2021.3067763.
- [9] O. Sagi & L. Rokach. (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1249.
DOI : 10.1002/widm.1249
- [10] J. Cao, S. Kwong, R. Wang, X. Li, K. Li & X. Kong. (2015). Class-specific soft voting based multiple extreme learning machines ensemble. *Neurocomputing*, 149, 275-284.
DOI : 10.1016/j.neucom.2014.02.072
- [11] Korea Meteorological Administration. (n.d.). *Weather Data Service - ASOS*. (Online). <https://data.kma.go.kr> (accessed on 18 Mar. 2021)
- [12] Traffic Accident Analysis System. (n.d.). *Geographic Information System*. (Online). <http://taas.koroad.or.kr> (accessed on 18 Mar. 2021)
- [13] K. Shim., M. Lee., I. Choi., Y. Boo, & W. Sung. (2017, December). Svd-softmax: Fast softmax approximation on large vocabulary neural networks. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 5469-5479.
- [14] H. J. Kim, J. W. Baek, & K. Chung. (2021). Associative Knowledge Graph using Fuzzy Clustering and Min-Max Normalization in Video Contents. *IEEE Access*, 9(1), 74802-74816.
DOI : 10.1109/ACCESS.2021.3080180
- [15] J. W. Baek, & K. Chung. (2021) Multi-Level Health Knowledge Mining Process in P2P Edge Network. *IEEE Access*, 9(1), 61623-61634.
DOI : 10.1109/ACCESS.2021.3073775
- [16] H. Yoo, S. Han, & K. Chung. (2021) Diagnosis Support Model of Cardiomegaly based on CNN using Resnet and Explainable Feature Map. *IEEE*

Access, 9(1), 55802-55813.
 DOI : 10.1109/ACCESS.2021.3068597

전 병 옥(Byeong-Uk Jeon) [학생회원]



- 2016년 3월 ~ 현재 : 경기대학교 AI컴퓨터공학부 (학부생)
- 2021년 2월 ~ 현재 : 경기대학교 컴퓨터과학과 데이터마이닝 연구원
- 관심분야 : 머신러닝, 딥러닝, 데이터마이닝, 빅데이터
- E-Mail : jebuk97@kyonggi.ac.kr

강 지 수(Ji-Soo Kang) [학생회원]



- 2020년 2월 : 경기대학교 컴퓨터공학부 (공학사)
- 2020년 3월 ~ 현재: 경기대학교 컴퓨터과학과 (석사과정)
- 2017년 12월 ~ 현재: 경기대학교 컴퓨터과학과 데이터마이닝 연구실 연구원

- 관심분야 : 데이터마이닝, 데이터 분석, 지능시스템, 헬스케어, 의료 딥러닝
- E-Mail : dm.jskang@kyonggi.ac.kr

정 경 용(Kyungyong Chung) [정회원]



- 2000년 2월 : 인하대학교 전자계산공학과 (공학사)
- 2002년 2월 : 인하대학교 전자계산공학과 (공학석사)
- 2005년 8월 : 인하대학교 컴퓨터정보공학부 (공학박사)

- 2006년 3월 ~ 2017년 2월 : 상지대학교 컴퓨터정보공학부 교수
- 2017년 3월 ~ 현재 : 경기대학교 컴퓨터공학부 교수
- 관심분야 : 데이터마이닝, 헬스케어, 빅데이터, 지능시스템, 인공지능, HCI, 정보검색, 추천 시스템
- E-Mail : dragonhci@gmail.com