

# 스마트 차량을 위한 운전자 행동 인식 분류

정현구\*, 강현수\*\*

## Driver Behavior Classification for Smart Vehicle Systems

Jung Hyun Gu\*, and Kang Hyun Soo\*\*

### 요 약

본 논문에서는 스마트 차량 시스템에서 운전자 행동을 인식하기 위한 분류 모델을 제시하였다. 운전자의 측면 이미지를 기반으로 총 10개의 클래스(정상 운전, 텍스트링, 전화 통화, 라디오 조작, 음료 섭취 등)를 분류하기 위해 전이 학습 기반의 VGG19과 ResNet152 모델을 사용하였다. ResNet152는 VGG19에 비해 더 깊은 네트워크 구조로 인해 학습 시간이 더 오래 걸렸지만, 복잡한 행동 패턴 인식에서 더 높은 분류 정확도를 기록하였다. 이를 통해 운전자 안전을 보장하는 스마트 차량 시스템 구현에 기여할 것으로 기대된다.

### Abstract

This paper presents a classification model for recognizing driver behavior in smart vehicle systems. Based on side-view images of the driver, the model classifies behavior into 10 categories, such as normal driving, texting, phone calls, radio operation, and drinking, using transfer learning-based VGG19 and ResNet152 models. Compared to VGG19, ResNet152 has a deeper network structure, resulting in longer training times, but it achieves higher classification accuracy in recognizing complex behavior patterns. This is expected to contribute to the development of smart vehicle systems that enhance driver safety by ensuring the timely detection of risky behaviors.

### Key words

Driver Behavior Recognition, VGG, ResNet152, Transfer Learning, Classification Model, Deep Learning, Behavior Analysis

## 1. 서 론

스마트 차량 기술의 발전으로 인해 자율주행 시스템과 운전자 지원 시스템에 대한 관심이 빠르게 증가하고 있다. 이러한 시스템은 도로 안전성을 향상시키고 운전자와 승객의 안전을 보장하는 데 중

요한 역할을 한다. 특히, 운전 중에 발생하는 다양한 운전자 행동을 실시간으로 인식하고 분석하는 것은 사고 예방과 운전자의 위험 행동에 대한 신속한 대응을 가능하게 하여 안전을 강화할 수 있다.

본 논문은 운전자의 행동을 총 10개의 클래스(정상 운전, 텍스트링, 전화 통화, 음료 섭취, 라디오 조

\*충북대학교 정보통신공학부 학부생, wjdgusrn951@naver.com

\*\*충북대학교 정보통신공학부 교수, hskang@cbnu.ac.kr

※ 이 논문은 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업(2019-0-01183)의 지원을 받아 작성되었음

작 등)로 분류하여 운전자 행동을 정확히 인식하고 분류하는 데 초점을 맞춘다. 이러한 행동 인식 분류는 운전자의 위험 행동을 조기에 탐지하고 신속히 대응하는 데 기여하고, 스마트 차량 시스템의 안전성을 더욱 강화할 수 있다.

본 논문에서는 VGG19와 ResNet152 모델을 사용하여 보다 정교한 운전자 행동 인식 시스템 개발에 기여할 수 있는 연구 결과를 제시한다.

## II. 문제

본 연구는 스마트 차량 시스템에서 운전자 행동을 인식하여 교통사고를 예방하고 운전자의 안전을 목적을 두고 있다. 운전 중 부주의한 행동은 심각한 사고로 이어질 수 있으므로, 이를 정확히 인식하고 조기에 대응하는 시스템 개발이 중요하다. 이를 위해 VGG19와 ResNet152 모델을 사용하여 총 10가지 운전자 행동을 분류하는 실험을 진행하였다.

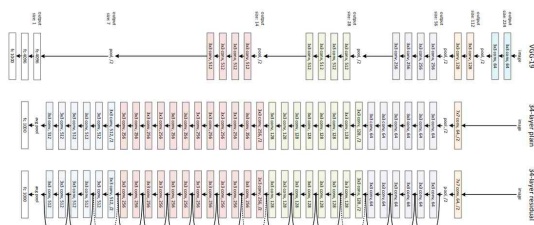


그림 1. VGG19 과 ResNet 34-layer plain과 34-layer residual 구조

Figure 1. VGG19 and the structure of ResNet 34-layer plain and 34-layer residual

VGG19와 ResNet152는 모두 널리 사용되는 이미지 분류 모델로, 전이 학습을 통해 대규모 데이터셋에서 사전 학습된 가중치를 활용하여 운전자 행동 인식에 적용되었다. VGG19는 19개의 레이어로 구성된 비교적 단순한 구조 덕분에 학습 속도가 빠르다. 반면, ResNet152는 152개의 레이어와 [그림 1]에서 보이는 잔차 연결(Residual Connection) 구조를 사용하여, 더 깊은 네트워크에서도 기울기 소실 문제를 해결하며 복잡한 행동 패턴을 효과적으로 학습할 수 있다.

총 22,424개의 유전자의 측면 이미지 데이터셋을

사용하여 행동을 분류하였다. 이 데이터는 정상 운전, 텍스트팅(왼손 및 오른손), 전화 통화(왼손 및 오른손), 라디오 조작, 음료 섭취, 뒤로 몸을 돌리는 행동, 머리 또는 화장 손질, 승객과 대화 등 총 10 가지 운전자 행동을 포함한다. 각 클래스는 운전 중 발생할 수 있는 다양한 행동을 포함하며, 해당 행동들을 인식하는 것을 목표로 한다.

두 모델 모두 Adam 옵티마이저와 categorical crossentropy 손실 함수를 사용하였고, 학습률은 ReduceLROnPlateau 콜백을 통해 자동으로 조정되었다. 총 20 에포크 동안 학습이 진행되었으며, 배치 크기는 16으로 설정되었다.

Categorical crossentropy 손실 함수는 모델이 예측한 클래스 확률과 실제 레이블 간의 차이를 계산하여 손실 값을 도출한다. 이때 사용된 손실 함수의 수식은 다음과 같다.

$$L = -\sum_{i=1}^N y_i \log(\hat{y}_i) \quad (1)$$

여기서  $N$ 은 클래스의 총 개수이며,  $y_i$ 는 실제 레이블(원-핫 인코딩으로 표현)이고,  $\hat{y}_i$ 는 모델이 예측한 클래스 확률을 의미한다. 이 수식은 예측된 확률과 실제 값이 일치할수록 손실이 줄어들도록 설계되었으며, 모델이 학습을 통해 예측의 정확도를 높이는 방향으로 손실을 최소화하게 된다.

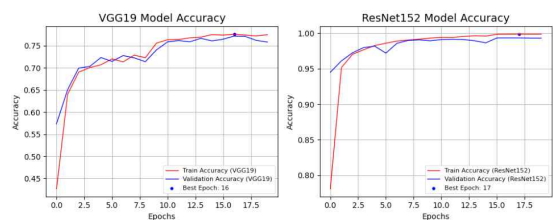


그림 2. VGG19와 ResNet152의 학습 및 검증 정확도 비교

Fig. 2. Comparison of Training and Validation Accuracy for VGG19 and ResNet152

모델의 성능은 정확도(Accuracy)와 손실(Loss)을 기준으로 평가되었다. [그림 2]는 VGG19와 ResNet152의 학습 및 검증 정확도 변화를, [그림 3]은 각 모델의 손실 변화를 보여준다.

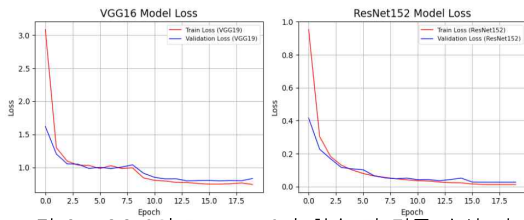


그림 3. VGG19와 ResNet152의 학습 및 검증 손실 비교  
Fig. 3. Comparison of Training and Validation Loss for VGG19 and ResNet152

VGG19는 초기 학습 속도가 빠르고, 빠른 수렴을 보였으나, 복잡한 행동 패턴을 분류하는 데에는 낮은 정확도를 보였다. 반면, ResNet152는 더 깊은 네트워크 구조와 잔차 연결(Residual Connection)을 통해 더욱 복잡한 행동을 인식할 수 있었으며, 검증 정확도에서 VGG19가 보여준 0.7721보다 높은 0.9931이라는 수치를 보여주며 VGG19 보다 우수한 성능을 기록했다.

Grad-CAM 결과 [그림 4]와 같이, 모델이 운전 중 전화 통화와 같은 특정 행동을 인식할 때, 운전자의 얼굴과 손, 핸들과 같은 중요한 영역에 집중하는 것을 확인할 수 있었다. 이를 통해 학습된 모델을 통해 복잡한 행동 패턴을 효과적으로 인식할 수 있음을 시각적으로 확인할 수 있었다.

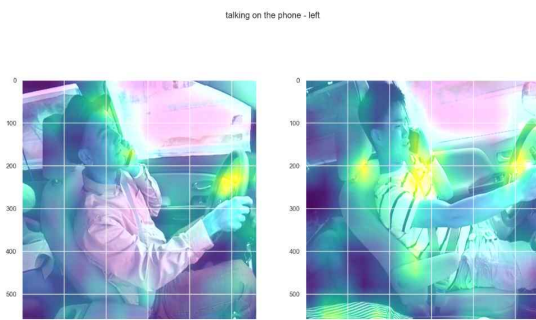


그림 4 . 전화 통화 행동에 대한 Grad-CAM 시각화 결과  
Fig. 4. Grad-CAM Visualization for 'Talking on the Phone - Hand' Action

### III. 결 론

본 논문에서는 스마트 차량 시스템에서 운전자 행동을 인식하는 딥러닝 기반의 분류 모델을 제안하고, 이를 위해 VGG19와 ResNet152 모델을 사용

하여 총 10가지 운전자 행동을 분류하였다. 실험 결과, ResNet152는 더 깊은 네트워크 구조와 잔차 연결(Residual Connection) 통해 복잡한 행동 패턴을 효과적으로 학습할 수 있었으며, [4]와 비교했을 때 높은 검증 정확도 0.9931을 기록하며 VGG19보다 뛰어난 성능을 보였다.

또한, Grad-CAM(Gradient-weighted Class Activation Mapping) 기법을 통해 ResNet152가 행동을 분류할 때 운전자의 얼굴과 손과 같은 중요한 영역에 집중하는 것을 시각화하여 직접 확인할 수 있었다. 이를 통해 ResNet152는 운전자 행동 인식 시스템을 개발하는 데 있어 더 나은 선택지임을 입증하였다.

이번 연구를 통해 스마트 차량 시스템에서 운전자 행동을 실시간으로 인식하고, 이를 통해 교통사고를 예방하는 데 기여할 수 있는 중요한 기초 자료를 제공한다. 앞으로 더 많은 데이터를 활용하거나, 경량화된 모델을 적용하여 실시간 응답 성능을 개선하여 스마트 차량 발전에 기여할 것으로 기대된다.

### 참 고 문 헌

- [1] Birkan Buyukarican, Erkan Ulker , "Classification of physiological disorders in apples fruit using a hybrid model based on convolutional neural network and machine learning methods", SCIE, vol 34, pp. 16973-16988, May 2022
- [2] HSheldon Mascarenhas, Mukul Agarwal, "A comparison between VGG16, VGG19 and ResNet152 architecture for image classification", Proc. CVPR, pp. 162, Feb 2022.
- [3] Sumaira Manzoor, Eun-Jin Kim, Sung-Hyeon Joo, Sang-Hyeon Bae, Gun-Gyo In, Kyeong-Jin Joo, "Edge Deployment Framework of GuardBot for Optimized Face Mask Recognition With Real-Time Inference Using Deep Learning", IEEE Trans. Pattern Anal. Machine Intell., vol. 10, pp. 11, July 2022.
- [4] Cong Duan, Yipeng Gong, Jiakai Liao, Minghai

Zhang, Libo Cao "FRNet: DCNN for Real-Time Distracted Driving Detection Toward Embedded Deployment", IEEE Trans. Pattern Anal. Machine Intell., PP(99):1-14, Sep 2023.