

가상 데이터를 이용한 딥러닝 기반 운전자 행위 검출 시스템

성재호, 한동석

경북대학교

dshan@knu.ac.kr

Deep learning-based driver behavior detection system using virtual data

Jaeho Seong, Dong Seog Han

Kyungbook National Univ

요약

최근 딥러닝 기반의 운전자 행위 검출이 활발히 연구되고 있다. 기존 연구에서는 스켈레톤 조인트 매핑 또는 R-CNN(Region of Convolutional Neural Network) 계열의 2-단계 객체 검출 등을 사용한다. 이러한 방식은 다른 운전자 모니터링 시스템과 연계 하기에는 연산량이 많다. 또한, 운전자 행위 검출에서 행위 검출에 실패하는 주요 원인은 사람의 행위를 분류할 때 차량 외부 객체에 영향을 받아 분류에 실패하는 문제점을 갖는다. 이러한 문제점을 해결하기 위하여 본 논문에서는 가상 데이터 생성 알고리즘 및 딥러닝 기반 운전자 행위 검출 시스템을 제안한다. 연산량을 줄이기 위하여 운전자 행위 검출 시스템은 1-단계 얼굴 검출기와 경량화된 분류기를 사용하였다. 다양한 배경을 적용한 가상 데이터를 이용하여 실제 데이터와 함께 학습데이터를 구성한다. 제안하는 시스템은 실제 차량 환경에서 실험을 통한 운전자 행위 예측 결과, 90%의 성능을 보였다.

I. 서론

운전자 모니터링 시스템(driver monitoring system, DMS)은 운전자 상태를 인식할 뿐만 아니라 운전자의 경계 수준을 확인하고 운전자의 주의를 모니터링한다. 최근 딥러닝 기반의 카메라를 사용한 운전자 모니터링 시스템을 구현하려는 연구가 활발히 진행 중이다. 운전자 모니터링 시스템에는 얼굴 인식, 졸음 감지, 전방 주시, 행위 검출 등이 있다. 그중에서 운전자의 행위 검출은 주의 산만 행동이 감지되었을 때 운전자에게 경고하고 차량 제어에 개입하여 사고를 방지한다.

기존의 운전자 행위 검출 연구에서는 사람의 자세를 추정하는 스켈레톤 조인트 매핑, 특징 추출과 객체 분류를 순차적으로 진행하여 객체를 검출하는 R-CNN 계열의 2-단계 검출기 등을 사용한다[1][2]. 이러한 방식은 운전자의 행위에만 초점이 맞춰져 있어 다른 운전자 모니터링 기술과 연계하기에는 연산량이 많다. 한편 운전자 행위 검출에서 인식에 실패하는 주요 원인은 담배 및 휴대폰 등이 각도에 따라 작고 불규칙한 모양을 가지기 때문이다. 따라서 딥러닝 기반의 분류기에서 차량 외부 환경의 객체를 운전자의 행위로 잘 못 분류하여 성능이 낮아진다. 또한, 딥러닝 기반의 모델은 새로운 환경에서 현저히 낮은 분류 정확도를 보인다.

본 논문에서는 가상 데이터를 이용한 딥러닝 기반 운전자 행위 검출 시스템을 제안한다. 제안하는 시스템은 다른 운전자 모니터링 시스템과 연계하기 쉽도록 얼굴 검출기와 분류기로 구성한다. 연산량을 줄이기 위하여 1-단계 검출기인 SSD(Single Shot MultiBox Detector)를 사용하고 저용량 메모리 환경에서 딥러닝을 적용하기 위한 경량화 모델인 MobileNet V2를 분류기로 사용하였다[3][4]. 또한, 다양한 배경을 적용한 가상 데이터를 실제 데이터와 함께 학습데이터로 구성하였다. 제안하는 시스템은 실제 차량 환경에서 검증한 결과 향상된 성능을 보였다.

II. 본론

2.1 가상 데이터 생성 알고리즘

본 장에서는 가상 데이터 생성 과정에 대해서 설명한다. 가상 데이터를 생성하기 위하여 배경, 사람, 크로마키로 된 사람 행위의 3가지 종류의 이미지가 필요하다. 딥러닝 기반의 DeepLab V3+를 사용한 의미론적 분할

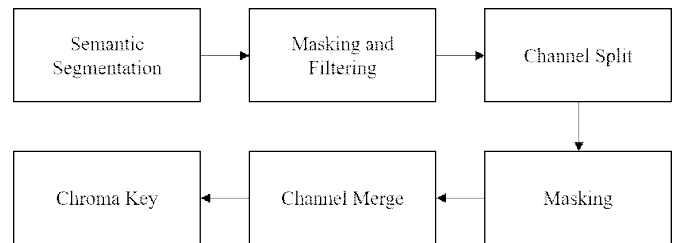


그림 1. 가상 데이터 생성 과정

(semantic segmentation)과 크로마키(chroma key)를 사용하여 생성하였다[5]. 배경 이미지로는 풍경, 차량 환경, 사무실 등 다양한 환경을 적용하여 외부 환경변화에 상관없이 사람의 행위가 중요한 특징으로 추출되도록 하였다. 사람의 행위 이미지는 다양한 각도에서 정밀한 촬영을 통해 수집하였다. 가상 데이터를 생성하는 과정은 다음과 같다.

- ① 사람 이미지에서 의미론적 분할을 통해 사람의 영역을 검출한다.
- ② 추출된 사람 이미지의 마스크를 생성한다. 의미론적 분할 과정에서 생긴 잡음을 제거하기 위하여 평균값 필터를 사용하여 마스크를 침식시킨다.
- ③ 사람 이미지와 배경 이미지를 BGR 채널로 분리한다.
- ④ 각 채널에 마스크 기법을 적용하여 사람 이미지에서 배경을 교체한다.
- ⑤ 각 BGR 채널을 병합한다.
- ⑥ 크로마키를 이용하여 병합된 이미지에 사람의 행위를 적용한다.

2.2 딥러닝 기반 행위 검출 시스템

본 장에서는 딥러닝 기반의 행위 검출 시스템 구조에 대해서 설명한다. 제안된 시스템은 행위 검출기를 사용하지 않고 얼굴 검출기를 사용하였다. 얼굴 검출기는 거의 모든 운전자 모니터링 기술에 사용되므로 다른 시스템과 연계하기가 쉽다. 그림 2는 제안하는 딥러닝 기반의 행위 검출 시스템을 도식화한 것이다. 운전자 행위 검출 과정은 딥러닝으로 학습된 얼굴 검출기를 통해 사람의 얼굴을 검출한다. 얼굴 주변에서 사람의 행위를 검출할 수 있도록 경계박스를 확장한다. 확장된 경계박스를 잘라내어 일부 배경을 없애고 이미지 크기를 (224, 224)로 줄여 분류기의 연산량을 줄였다.

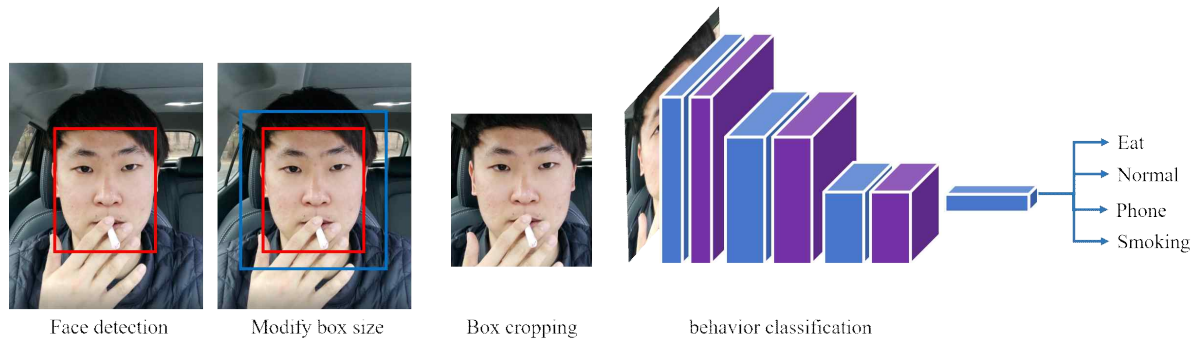


그림 2. 딥러닝 기반 운전자 행위 검출 시스템

CNN(Convolutional Neural Network) 기반의 분류기를 거쳐 운전자의 행위를 음식 섭취, 평소 상태, 휴대 전화, 흡연의 4가지 클래스로 분류한다.

2.3 실험 환경 및 결과

본 장에서는 실험 환경과 실험 결과에 대해서 설명한다. 실험은 Logitech Quickcam Pro 9000 카메라를 사용하였으며 비디오의 해상도는 1280x720이고 프레임 레이트는 30fps이다. 학습과 테스트에 사용된 네트워크 입력 크기는 (224, 224)이며, 100 에폭을 하여 학습하였다. 학습률은 0.0001, 배치 크기는 32로 설정하였다. 데이터 증강 기법으로는 밝기 조절, 회전, 수직 플립, 높이와 너비 이동, 확대를 사용하였다. 실험은 공터에서 카메라로 촬영된 비디오 프레임들을 테스트 데이터로 사용하였다. 실제 데이터는 총 4개의 카메라를 사용하여 여러 각도에서 학습데이터를 수집하였다. 표 1과 2는 실험에 사용한 학습데이터를 나타낸 것이다.

표 1 실험에 사용된 실제 데이터

Class Name	Training data		Test data	Total
	Real Data	Virtual Data		
eat	4,587	0	360	4,947
Normal	7,978	0	360	8,338
Phone	6,234	0	360	6,594
Smoke	8,160	0	360	8,520

표 2 실험에 사용된 실제 데이터와 가상 데이터

Class Name	Training data		Test data	Total
	Real Data	Virtual Data		
eat	4,587	5,413	360	10,360
Normal	7,978	2,022	360	10,360
Phone	6,234	3,765	360	10,360
Smoke	8,160	1,840	360	10,360

표 1에서는 실제 데이터만을 사용하였고, 클래스별 개수가 불균형을 이루고 있다. 표 2는 실제 데이터에 가상 데이터를 추가하여 실험에 사용하였다. 실제 데이터의 클래스별 개수의 불균형을 해결하기 위해 가상 데이터를 이용하여 데이터를 증강하였다. 제안하는 시스템의 표 1과 2를 바탕으로 5번 실험하여 테스트 데이터의 평균 정확도를 비교하였다. 표 1의 실제 데이터의 테스트 정확도는 평균 83%의 분류 성능을 갖는다. 표 2의 실제 데이터에 가상 데이터를 추가한 결과 테스트 정확도는 평균 90%로 약 7% 성능이 향상되었다. 그림4는 Class Activation Map(CAM)을 사용하여 CNN 모델의 분류 기준을 히트맵으로 시각화한 것이다[6].

표 1의 실제 데이터로 학습된 모델은 평소 상태에서 잘 못 분류할 경우가 가장 많았다. 이는 그림 3(a)와 같이 외부 배경으로 인한 추론 실패로 이어지기 때문이다. 또한, 그림 3(a)와 (b)를 비교했을 때 가상 데이터를 이용하면 배경에 영향을 덜 받는 것을 알 수 있다.



그림 3. Class Activation Map(CAM)을 이용한 분류 기준 시각화

III. 결론

본 논문에서는 가상 데이터를 이용한 딥러닝 기반 운전자 행위 검출 시스템을 제안하고 성능을 검증했다. 가상 데이터 생성 알고리즘을 이용하여 데이터의 불균형을 해소하고 외부 배경 변화로부터 강인하다는 것을 CAM을 통해 보였다. 그 결과 제안하는 시스템의 분류 성능을 약 7% 높였다. 운전자의 행위 검출 시스템에서 얼굴 검출 및 분류 기법만을 사용하여 실시간 추론 속도를 줄였다.

ACKNOWLEDGMENT

이 논문은 2019년 정부(산업통상자원부)의 재원으로 한국산업기술평가관리원의 지원을 받아 수행된 연구임(과제번호: 20003519)

참 고 문 헌

- [1] Martin, Manuel, et al. "Body pose and context information for driver secondary task detection." 2018 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2018.
- [2] Lu, Mingqi, Yaocong Hu, and Xiaobo Lu. "Driver action recognition using deformable and dilated faster R-CNN with optimized region proposals." Applied Intelligence 50.4 (2020): 1100-1111.
- [3] Liu, Wei, et al. "Ssd: Single shot multibox detector." European conference on computer vision. Springer, Cham, 2016.
- [4] Sandler, Mark, et al. "Mobilenetv2: Inverted residuals and linear bottlenecks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
- [5] Chen, Liang-Chieh, et al. "Encoder-decoder with atrous separable convolution for semantic image segmentation." Proceedings of the European conference on computer vision (ECCV). 2018.
- [6] Zhou, Bolei, et al. "Learning deep features for discriminative localization." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.