

MTCNN기반 운전자 시선정보 시각화 연구

Study of MTCNN-based driver gaze information visualization

김형규¹, 류승기^{2*}

Hyung-Kyu Kim¹, Seung-Ki Ryu^{2*}

요 약

교통사고의 약 80%는 운전자의 시선분산과 관련이 있으며, 도로에서 운전자가 전방을 주시하지 않을 경우 교통사고 발생 가능성이 크게 증가한다. 2개의 카메라를 이용하여 주행 중 운전자의 시선 위치를 추정하고 전방화면에 표출하여 시각화할 수 있는 기술을 연구하였다. 전방 주시 화면을 32개 영역으로 분할하였기 때문에 운전자가 주시하는 위치를 차량 내부, 외부로 나누어 명확히 제시할 수 있었다. 피실험자 대상의 정답이미지 셋 1.4만을 자체 구축 후 학습하였다. 테스트결과 평균 절대 비율 오차는 3.58%로 높은 성공률을 나타냈으며, 시각화까지의 시간도 0.65ms로 실시간 적용이 가능한 수준이다. 연구된 기술을 도로유지관리 및 순찰차량에 적용할 경우, 도로주행시 운전자의 시선분산을 유발하는 위험도로의 상세정보를 수집할 수 있다. 도로상의 광고시설, 파손된 도로시설물 등 불필요한 시선유발 시설물을 사전에 검출하여 제거함으로써 교통사고예방에 큰 효과가 있을 것으로 판단된다. 교통사고 인건요인 분석에 대한 원천기술의 확보로 전방주시태만, 졸음운전과 같은 상황을 예방할 수 있다.

핵심어 : 주의산만, 운전자, 시각화, MTCNN, 안전운전

Abstract

About 80% of traffic accidents are related to the distraction of the driver's gaze. If the driver does not look ahead on the road, the likelihood of a traffic accident increases significantly. A technology that can be visualized by estimating the driver's gaze position while driving using two cameras and displaying it on the front screen was studied. Since the forward-looking screen is divided into 32 areas, the driver's gaze can be presented by dividing it inside and outside the vehicle. Only 1.4 of the set of correct answer images of the subjects were self-constructed and learned. As a result of the test, the average absolute ratio error was 3.58%, indicating a high success rate, and the time until visualization was 0.65ms, which is a level that can be applied in real-time. If the researched technology is applied to road maintenance and patrol vehicles, detailed information on dangerous roads that cause the driver's gaze to be dispersed while driving on the road can be collected. It will be judged to have a great effect in preventing traffic accidents by detecting and removing unnecessary eye-causing facilities such as advertisement facilities and damaged

1 Department of Highway and Transportation Research, KICT, Gyeonggi-do, Korea [Senior Researcher]

e-mail: hyoungkyukim@kict.re.kr

2 Department of Highway and Transportation Research, KICT, Gyeonggi-do, Korea [Senior Research Fellow]

e-mail: skryu@kict.re.kr(Corresponding author)

* 본 연구는 과학기술정보통신부 한국건설기술연구원 연구운영비지원(주요사업)사업으로 수행되었습니다. (과제번호 20220141-001, 미래교통 스마트 인프라 핵심기술개발).

Received(November 2, 2022), Review Result(1st: November 21, 2022), Accepted(December 12, 2022), Published(December 31, 2022)



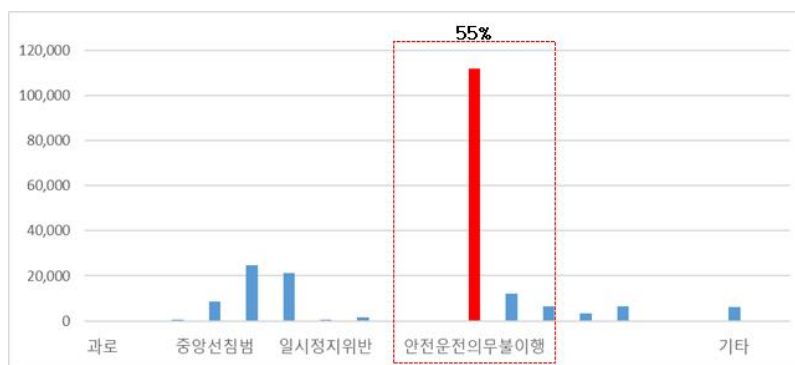
© 2022 The Authors. Published by NCISS.
This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.
To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

road facilities on the road in advance. Situations such as negligent driving ahead and drowsy driving can be prevented by securing the original technology to analyze the cause of a traffic accident.

Keyword : Distraction, Driver, Visualization, MTCNN, Driving Safety

1. 서론

교통사고의 약 80%는 운전자의 시선분산과 관련이 있으며, 도로에서 운전자가 전방을 주시하지 않을 경우 교통사고 발생 가능성이 크게 증가한다 [1]. 국내 교통사고 원인을 심층분석한 결과에서도 운전자의 인적요인이 원인인 사망사고는 전체교통사고의 98.6%이다 [2]. 또한, [그림 1]과 같이 2022년 교통사고 법규위반사고의 55%는 운전자 부주의로 인한 안전운전불이행사고로 분석되었다 [3].



[그림 1] 법규위반 교통사고 원인

[Fig. 1] Causes of traffic accidents in violation of laws

교통사고 조사시 가해자와 피해자의 진술과 블랙박스 영상만으로 교통사고 원인을 분석하여, 실제 교통사고에서 운전자의 인적요인을 정확히 분석하기는 어려운 실정이다. 교통사고 사망자 50% 감소라는 국가정책에 따라 운전자의 인적요인에 의한 교통사고를 명확히 분석하고 교통사고 재발을 방지하기 위해서는 운전자측면에서 도로주행시 주의력저하 및 시선분산 원인을 분석할 필요가 있다. 과거 연구에서는 운전자에게 시선추적기를 장착시켜 운전자의 시선변화 및 주의력 저하를 분석하고자 하였으나, 시선추적기의 과도한 비용과 운전자에게 장착시 불편함 등으로 기술개발이 확대되지 못하였다.

이를 위해서는 다양한 운전자에게 적용이 가능한 시선 추적 기술이 필요하며, 차내 장착 편의성 및 비용을 고려해야한다. 또한, 주행시 운전자에게 방해가 되지 않아야 하며, 운전자가 주행중인 전방 화면과 매칭하여 시선 분산 원인 분석이 용이하도록 설계할 필요가 있다.

본 연구에서는 상기 요구사항을 반영하여 운전자가 주행시 행동에 방해가 되지 않도록, 비장착형으로 시선을 추적할 수 있는 딥러닝모델을 개발하고, 운전자의 시선정보분석이 용이하도록

Heatmap형식으로 표출하는 방법을 연구하였다.

2. 기존 문헌 고찰

2.1 운전자 시선인식 연구

과거부터 운전자의 인적요인 문제를 인식하였고, 2000년대 이후 센서를 이용한 시선 분산연구가 시작되었다. 미국에서는 CAS(Collision Avoidance Systems)를 이용하여 운전자의 주의 산만을 판단하였으며, 자체 테스트 결과 약 80%의 작업정확도를 나타냈다 [4].

취득된 시각정보를 연구에 활용하기 위해서는 운전자가 주행시 시각정보 수집장치에 영향을 받지 않아야 하는데, 상용화된 장비의 경우 운전자에게 부착해야하는 점이 존재하여 자연스러운 주행상태의 데이터 수집이 어려운 상황이다. 이러한 문제점으로 과거 연구에서는 시설물 인식과 같이 단순 시설물 대상의 시선정보 분석이 이루어졌다 [5]

최근 머신러닝 기반의 알고리즘의 개발로 운전자의 상태를 운전 중 시각, 생체 전기 및 기타 센서를 통해 직접 모니터링할 수 있게 되었다. 운전자의 다양한 상태로 분류할 수 있으며, 이를 통해 운전자가 주의가 산만해지는지 여부를 판단할 수 있다. 대표적인 연구들로 Bayesian network를 적용하여 운전자 주의 산만을 분류하였는데, 차량 내부에 앞 유리, 사이드 미러, 백미러, 계기판, 조수석 창문 영역 총 5개로 구분하여 시야 영역을 정의하였다 [6].

2010년대 적외선 센서를 이용한 Eye-Tracker의 개발로 운전자 시선정보를 직접적으로 수집할 수 있게 되었는데, 운전자를 피실험자로 삼아 운전자의 주의산만 정도 분석을 안구 운동 정보로 표출하였다 [7].

기존 연구들은 운전자 시선정보를 벡터값으로 추출하여 통계적 분석을 통해 운전자의 주행 위험성 추정을 목적으로 진행되었지만, 실제 운전자가 주행시 차량 내외부 특정 사물을 주시하고 있는지 시각적으로 확인할 수는 없었다. 본 논문에서는 기존 연구에서 수행결과를 개선하고 분석결과를 시각화할 수 있도록 하였다. Eye-Tracker를 활용한 운전자 시선추적방식은 적외선 카메라를 사용한다는 점에서 비용의 문제가 있고, 운전자가 주행중 바라보는 전방 영상은 일반 카메라 영상(RGB: Red Green Blue)으로 촬영되어 시선정보와 매칭하기에는 호환성의 문제가 발생한다. 이를 극복하기 위하여 MTCNN(Multi-Task Convolutional Neural Network)알고리즘을 이용하여 일반 카메라 영상으로 운전자 시선정보를 추출할 수 있도록 하고, 운전자 전방 주행 화면과 매칭하여 운전자가 실제로 주시하는 위치를 시각화하도록 하였다.

2.2 운전자 시선추출관련 딥러닝 알고리즘 이론

MTCNN 알고리즘은 이미지에서 사람의 얼굴을 검출하고 추출하는 데 사용되었습니다. 얼굴 영

상에서 얼굴 전체의 외모 기반 시선 추정 방법을 통해 시선 데이터를 얻을 수 있다. 시선 추정을 위한 전면 CNN 아키텍처는 사람의 전면 얼굴 이미지를 입력값으로 취하고 2D 시선 추정치를 회귀분석한다. 공간 가중치 메커니즘은 전체 얼굴의 다른 영역에 대한 정보를 표준 CNN 아키텍처로 인코딩합니다. MTCNN 얼굴 검출기를 사용하여 얼굴을 인식하고 찾는 과정은 얼굴 또는 비얼굴 분류, 얼굴 프레임 회귀, 특징점 위치의 세 가지 작업으로 나눌 수 있습니다. 이 세 단계를 P-Net, R-Net 및 O-Net이라고 합니다. 모든 샘플 x_i 에 대해 교차 엔트로피 손실 함수가 사용됩니다.

$$L_i^{\text{det}} = -(y_i^{\text{det}} \log(p_i) + (1 - y_i^{\text{det}})(1 - \log(p_i))) \quad (1)$$

여기서, p_i = actual output value of the neural network

y_i^{det} = expected output value

L_j^{det} = classification loss

후보 상자와 정답 사이의 오프셋은 회귀 작업에서 예측하는 데 사용됩니다. 회귀 작업을 위해 유클리드 거리 손실 함수가 선택되었으며 다음과 같이 계산할 수 있습니다.

$$L_i^t = \|\hat{y}_i^t - y_i^t\|_2^2 \quad (2)$$

이 중, \hat{y}_i^t 는 네트워크의 회귀 결과이고, y_i^t 는 represents the ground truth coordinates이다. 후보 상자의 왼쪽과 위쪽 점의 좌표와 후보 상자의 높이와 무제한 4개의 값이 있으므로 $y_i^t \in R^4$ 다.

특징점의 위치는 모든 입력 샘플에 대해 최소 유클리드 거리 손실 함수를 사용한 얼굴 후보 상자의 필터와 유사합니다. 손실 함수는 다음과 같이 계산할 수 있습니다.

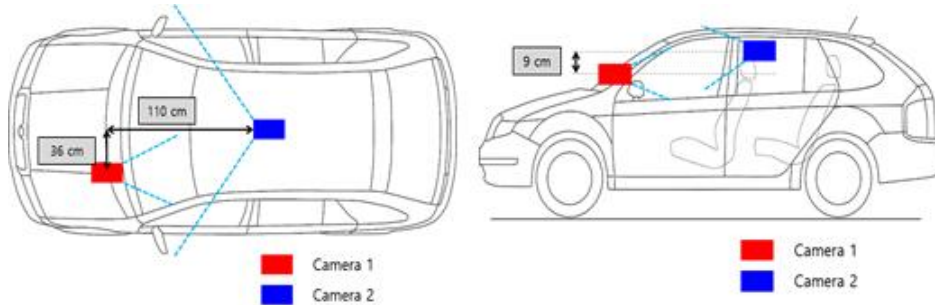
$$L_i^{\text{mark}} = \|\hat{y}_i^{\text{mark}} - y_i^{\text{mark}}\| \quad (3)$$

이 중, \hat{y}_i^{mark} 는 indicates the coordinates of facial key point obtained from the network이며, y_i^{mark} 는 ground truth coordinates of the j-th sample입니다. 얼굴의 특징점은 왼쪽 눈, 오른쪽 눈, 코끝, 입의 왼쪽, 입의 오른쪽 이렇게 5가지가 있습니다.

3. 운전자 시선정보 표출 방법론

운전자 시선정보를 추출하고 표출하기 위해서 수집되어야 하는 정보는 운전자 얼굴 영상과 운전자가 주시하는 전방의 화면 영상 2가지이다. 이를 위해 차량에 2개의 카메라를 설치하여 각각의

영상을 수집한다. 다음 [그림 2]와 같이 카메라 1은 운전자를 촬영하기 위하여 계기판 위 대쉬보드에 설치하였고, 카메라 2는 운전자가 주시하는 전방 화면을 위하여 운전석과 조수석 헤드레스트의 중앙에 위치하도록 설치하였다.



[그림 2] 운전자 시선정보 추출을 위한 카메라 설치 방법

[Fig. 2] Camera installation method for extracting driver's gaze information

설치된 카메라 화질은 HD급으로 상세 성능은 다음 [표 1]과 같다.

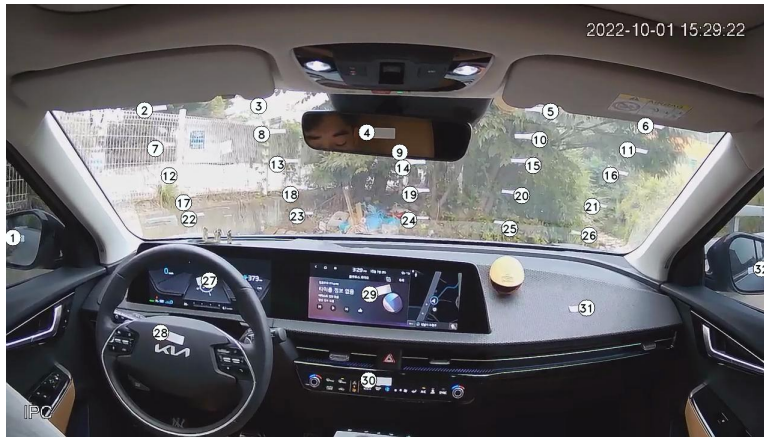
[표 1] 실험에 사용된 카메라 성능

[Table 1] Camera performance used in the experiment

구분	성능
모델명	DAHUA IPC-HFW2230S 3.6mm
해상도	1080P(1920x1080)
Horizontal FOV	110
Vertical FOV	60

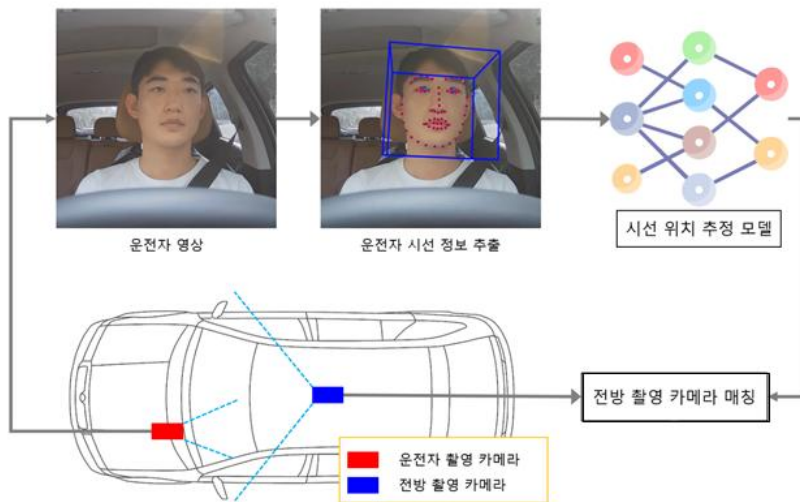
1번 카메라는 운전자 얼굴 및 시선 추적을 위하여 MTCNN 알고리즘을 이용하여 운전자의 얼굴 윤곽 및 시선정보를 분석하였다. 운전자의 얼굴에서 추출된 시선정보에 대하여 위치좌표 및 방향을 분류하기 위하여 SVM(Support Vector Machines) 알고리즘으로 시선정보를 매칭하였다.

마지막으로 운전자가 주시하는 위치의 실시간정보의 가시화를 위하여 전방주시화면에 Heatmap 방식으로 표출하도록 하였다. Heatmap으로 표출하기 위해 모델 학습 및 주시영역에 대한 정량적 데이터를 확보할 수 있도록 다음 [그림 3]과 같이 32개 영역으로 구분하였다. 영역은 크게 차량 내부와 외부로 구분하였는데, 차량 내부의 경우 핸들, 계기판, 네이게이션, 백미러 등이며, 외부는 운전자가 시야로 확인할 수 있는 차량 외부화면을 모두 포함하였다. [그림 4]는 시선정보 시각화를 위한 개념도이다.



[그림 3] 전방주시 화면 영역 분할

[Fig. 3] Division of the forward-looking screen area



[그림 4] 시선정보 시각화 개념도

[Fig. 4] Concept of Gaze information visualization

4. 모의 운전 실험을 통한 데이터 수집

모의 운전 실험을 통해 운전자 시선정보 표출을 위한 데이터를 수집하였다. 운전자의 시각적 주의 산만을 유도하여 운전자의 시각적 주의 산만 모델을 학습하기 위하여, 실험 감독관은 피실험자인 운전자에게 32개 영역을 응시하도록 지시했습니다. 운전자는 영역별로 30초 동안 응시하였으며, 카메라 1로 운전자 영상을 촬영하였다. 2번 카메라로부터 표시된 영역의 중심 픽셀의 방향을 정답

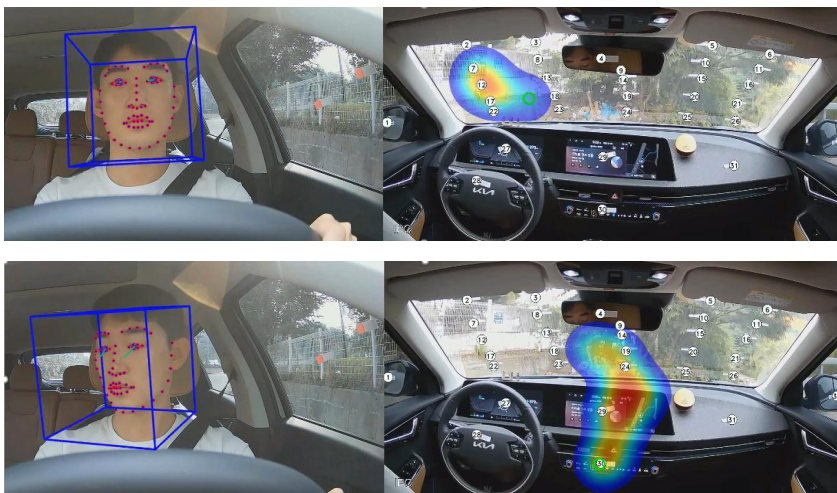
데이터로 사용하였다.

수집결과 1개 영역의 평균 영상길이는 30초였으며, 카메라 FPS는 15였다. 정답 라벨이 있는 이미지는 14,642장이 수집되었으며, 이중 학습에 사용된 이미지는 10,969장이다. 테스트용으로 3,673장을 사용하여 최종적으로 개발된 모델의 정확도를 테스트하였다.

모의 운전 실험과정에서 안전상의 이유로 운전자에 대한 과도한 요구나 운전 중 운전자의 행동에 대한 간섭은 이루어지지 않았다. 또한, 본 연구에서는 카메라를 운전자에게 장착하거나 하지 않았기 때문에 실험장비에 의한 실험 결과의 영향은 미치지 않았다. 즉, 정상적인 운전 조건에서 운전자의 경우 얼굴을 잘 감지할 수 있고 얼굴 영상을 통해 해당 시선 방향을 추정할 수 있었다.

5. 모델 구축결과

모델 구축결과, [그림 5]와 같이 운전자가 주시하는 시선방향에 따라 운전자가 주시하는 영역의 위치와 강도의 표출이 가능하였다. 운전자 시선좌표 중심에 가까울수록 붉은색으로 강도가 표시되고, 강도가 낮을수록 파란색으로 표출이 가능하다. 차량 외부 및 내부 모두 확인이 가능하였다.



[그림 5] 운전자 시선정보의 heatmap 표출 결과

[Fig. 5] Heatmap result of driver gaze information

모델의 성능을 검증하기 위하여, 카메라로부터 추정된 시선 좌표의 방향 벡터를 추출하고, x축 및 y 축에 대한 각도를 비교하여 오차를 계산하였다. 성능 검증은 테스트용 이미지 3,673장으로 진행되었으며, 평균 절대 오차(MAE : Mean Absolute Error), 평균 절대 비율 오차(MAPE : Mean Absolute Percentage Error)를 이용하여 오차를 계산하였다. 또한 모델의 추론 속도도 활용성을 판단하는데 중요한 수단 중의 하나이기 때문에 같이 측정하였다.

다음 [표 2]와 같이 임베디드 환경에서 자체 구축 테스트결과의 평균 오차는 평균 오차는 0.0748이었으며, 평균 절대 비율 오차는 3.58%로 나타났다. 시선추정법의 평균 오차는 4.8~6.0%인 점을 고려할 경우, 구축된 모델의 결과는 우수하다고 판단할 수 있다 [8].

또한, 시선정보를 시각화하는데까지 걸리는 시간은 0.65ms로 분석되어, 실시간 운전자 시선정보를 제공하는데 문제가 발생하지 않았다.

[표 2] 모델 정확도 분석 결과

[Table 2] Results of model accuracy analysis

구분	MAE	MAPE
Area 1	0.0208	0.7659
Area 2	0.0216	0.1557
Area 3	0.0193	0.1657
Area 4	0.0499	0.3688
Area 5	0.0374	0.2564
Area 6	0.1396	1.5420
Area 7	0.0438	0.4423
Area 8	0.0470	1.7740
Area 9	0.0444	0.3931
Area 10	0.0681	0.5590
Area 11	0.0757	1.2400
Area 12	0.0456	0.6958
Area 13	0.0405	2.9630
Area 14	0.1261	1.4450
Area 15	0.0438	0.5585
Area 16	0.1450	8.5340
Area 17	0.0753	2.9920
Area 18	0.0560	1.4310
Area 19	0.0227	1.1250
Area 20	0.0269	0.5964
Area 21	0.0789	3.6390
Area 22	0.0475	0.6158
Area 23	0.0241	7.1450
Area 24	0.0402	4.3100
Area 25	0.0390	33.4300
Area 26	0.4082	22.4300
Area 27	0.0358	1.4810
Area 28	0.0531	0.5950
Area 29	0.4153	11.1200

Area 30	0.0328	0.3268
Area 31	0.0296	0.2030
Area 32	0.0383	1.3260
평균	0.0748	3.5820

4. 결론

2개의 카메라를 이용하여 주행 중 운전자의 시선 위치를 추정하고 전방화면에 표출하여 시각화할 수 있는 기술을 연구하였다. 전방 주시 화면을 32개 영역으로 분할하였기 때문에 운전자가 주시하는 위치를 차량 내부, 외부로 나누어 명확히 제시할 수 있었다. 피실험자 대상의 정답이미지셋 1.4만을 자체 구축 후 학습하였다. 테스트결과 평균 절대 비율 오차는 3.58%로 높은 성공률을 나타냈으며, 시각화까지의 시간도 0.65ms로 실시간 적용이 가능한 수준이다.

단, 본 연구에서는 1명의 운전자와 제한된 차량 환경으로 학습 및 실험했기 때문에 다양한 운전자 환경에 적용하기 위해서는 운전자 및 차량 환경을 다양화하여 데이터셋을 증강시킬 필요가 있다. 안전상의 이유로 주행 중 차량에서는 정답이미지셋을 구축하지 못하였기 때문에 주행 중 이미지 수집이 요구된다. 운전자가 실제 주행시 일정영역을 지속적으로 주시할 경우 외부환경을 인식하지 못하여 교통사고가 발생할 수 있기 때문에 가상환경주행과 같은 안전이 확보된 실험환경개발이 동반될 필요가 있다.

향후 스마트폰 내 기능을 탑재할 경우, 수집 디바이스가 간소화되어 주행 중 운전자의 시선을 추적하고 위험운전상황을 경고할 수 있을 것이다. 또한 연구된 기술을 도로유지관리 및 순찰차량에 장착시 도로주행시 시선분산을 유발하는 위험도로의 상세정보를 수집할 수 있기 때문에 도로상의 광고시설, 파손된 도로시설물 등 불필요한 시선유발 시설물을 사전에 검출하여 교통사고예방에 큰 효과가 있을 것으로 판단된다. 교통사고 인전요인 분석에 대한 원천기술의 확보로 전방주시태만, 졸음운전과 같은 상황을 예방할 수 있다.

References

- [1] C. Klauer, T. A. Dingus, V. L. Neale, J. D. Sudweeks, D. J. Ramsey, "The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data", U.S. Department of Transportation: Washington, DC, USA, DOT HS 810 594, April 2006. [Online]. Available: vtechworks.lib.vt.edu/bitstream/handle/10919/55090/DriverInattention.pdf.
- [2] J. H. Seol, J. K. Lim, W. S. Choi, "A Study on the Strategies for 'Vision Zero' Goal of Traffic Fatalities in Korea", Korea Transport Institute, Sejong-si, Republic of Korea, Research Series-2013-14, November 2013. [Online]. Available: scienceon.kisti.re.kr/commons/util/originalView.do?cn=TRKO201400002962&dbt=TRKO&rn.

- [3] Korea Road Traffic Authority, “Detailed traffic accident statistics”, taas.koroad.co.kr, <http://taas.koroad.co.kr/>, (accessed October 10, 2022).
- [4] K. Torkkola, N. Massey, C. Wood, “Driver inattention detection through intelligent analysis of readily available sensors”, The 7th International IEEE Conference on Intelligent Transportation Systems, October 3-6, 2004, Washington D.C., USA, pp. 326-331, doi: 10.1109/ITSC.2004.1398919.
- [5] W. Li, J. Huang, G. Xie, F. Karray, R. Li, “A survey on vision-based driver distraction analysis”, Journal of Systems Architecture, vol. 121, December 2021, pp. 102319, doi: 10.1016/j.sysarc.2021.102319.
- [6] F. Lethaus, M. R. Baumann, F. Köster, K. Lemmer, “A comparison of selected simple supervised learning algorithms to predict driver intent based on gaze data”, Neurocomputing, vol. 121, December 2013, pp. 108-130, doi: 10.1016/j.neucom.2013.04.035.
- [7] F. Tango, M. Botta, “Real-time detection system of driver distraction using machine learning”, IEEE Transactions on Intelligent Transportation Systems, vol. 14 no. 2, March 2013, pp. 894-905, doi: 10.1109/TITS.2013.2247760.
- [8] J. Kuo, M. G. Lenné, M. Mulhall, T. Sletten, C. Anderson, M. Howard, A. Collins, “Continuous monitoring of visual distraction and drowsiness in shift-workers during naturalistic driving”, Safety Science, vol. 119, November 2019, pp. 112-116, doi: 10.1016/j.ssci.2018.11.007.