

18CTAP
-C132686
-02

자동차
사고
예방을
위한
멀티모달
학습
기반의
운전자
상태
인식
시스템
개발

최종
보고서

2018

국토교통과학기술진흥원
국토교통부

보안 과제(), 일반 과제(O) / 공개(O), 비공개() 발간등록번호()
국토교통기술촉진연구사업 제2차 연도 최종보고서

R&D / 18CTAP-C132686-02

자동차 사고 예방을 위한 멀티모달 학습 기반의 운전자 상태 인식 시스템 개발

최종보고서

2019. 03. 21.

주관연구기관 / 충남대학교산학협력단

국토교통부
(전문기관) 국토교통과학기술진흥원

제 출 문

국토교통과학기술진흥원 귀하

'자동차 사고 예방을 위한 멀티모달 학습 기반의 운전자 상태 인식 시스템 개발'(연구개발 기간 : 2017. 06. 30. ~ 2018. 12. 31.) 과제의 최종보고서 9부를 제출합니다.

2019. 03. 21.

주관연구기관명 : 충남대학교산학협력단 (대표자) 이 영 석



주관연구기관책임자: 이 규 철

국토교통부소관 연구개발사업 운영규정 제29조에 따라 최종보고서 열람에
동의합니다.

보고서 요약서

과제 고유 번호	18CTAP-C132686-02	해당 단계 연구 기간	2018.01.01. ~ 2018.12.31.	단계구분	2/2
연구사업명	중사업명	국토교통기술촉진연구사업			
	세부사업명	기초원천			
연구과제명	대과제명	자동차 사고 예방을 위한 멀티모달 학습 기반의 운전자 상태 인식 시스템 개발			
	세부과제명	-			
연구책임자	이 규 철	해당단계 참여연구원 수	총: 12명 내부: 12명 외부: 0명	해당단계 연구비	정부: 160,000 천원 민간: 0 천원 정부외: 0 천원 상대국: 0 천원 계: 160,000 천원
		총 연구기간 참여연구원 수	총: 12명 내부: 12명 외부: 0명	총 연구비	정부: 240,000 천원 민간: 0 천원 정부외: 0 천원 상대국: 0 천원 계: 240,000 천원
연구기관명 및 소속 부서명	충남대학교산학협력단			참여기업명	-
국제공동연구	상대국명: -			상대국 연구기관명: -	
위탁연구	연구기관명: -			연구책임자: -	

※ 국내·외의 기술개발 현황은 연구개발계획서에 기재한 내용으로 같음

연구개발성과의 보안등급 및 사유	일반 - 국가적으로 보호대상인 기술개발 내용이 없음 - 대중교통 등 공익을 목적으로 활용될 수 있음
-------------------------	---

9대 성과 등록·기탁번호

구분	논문	특허	보고서 원문	연구시설 · 장비	기술요약 정보	소프트 웨어	화합물	생명자원		신품종	
								생명 정보	생물 자원	정보	실물
등록·기 탁 번호											

국가과학기술종합정보시스템에 등록된 연구시설·장비 현황
(해당 없음)

요약 숙련된 운전자 또는 반응이 빠른 운전자일지라도 건강이상, 피로누적 등의 운전자 개인의 생리적 변화에 따라 사고가 발생할 수 있음. 대표적으로 운전자에게 찾아온 졸음과 심장질환은 운전자의 의식이 빠르게 사라지기 때문에 운전자 스스로 사고에 대응 할 수 없어 큰 사고로 이어질 가능성이 높음. 본 연구개발은 미래 자동차가 운전자의 졸음과 심장질환을 인식하고 대응하기 위한 시그널 분석 및 멀티모달 딥러닝 기반의 상태 인식 모델을 도출하고 기존 인식 모델보다 적은 학습 데이터로 우수한 성능을 보였음	보고서 면수: 101
---	-------------

요약문

연구의 목적 및 내용	숙련된 운전자 또는 반응이 빠른 운전자일지라도 건강이상, 피로누적 등의 운전자 개인의 생리적 변화에 따라 사고가 발생할 수 있음. 대표적으로 운전자에게 찾아온 졸음과 심장질환은 운전자의 의식이 빠르게 사라지기 때문에 운전자 스스로 사고에 대응 할 수 없어 큰 사고로 이어질 가능성이 높음. 본 연구개발은 미래 자동차가 운전자의 졸음과 심장질환을 자동으로 인식하고 대응하기 위한 멀티모달 딥러닝 기반의 상태 인식 시스템을 개발하는 것을 목표로 함				
연구개발성과	- 가상/실제 주행환경 구축, 운전자 심장활동 및 얼굴 이미지 수집 - 가상 주행환경으로부터 32,540개의 학습용 멀티모달 데이터 수집 - 실제 주행환경으로부터 3,254개의 테스트용 멀티모달 데이터 수집 - 딥러닝을 통해 수집된 데이터의 규모와 질적 수준이 높음을 검증 - 시그널 분석 및 딥러닝 기반 운전자 심장 활동 모니터링 - 온라인/오프라인에서 수집된 심장질환 관련 데이터의 시그널 분석 - 심장질환 데이터의 CWT 추출 및 데이터 증강을 통한 학습 데이터 확보 - CNN 기반의 학습 모델을 기반으로 심장 활동 상태 분류 (정확도 92.70%) - 멀티모달 딥러닝 기반의 운전자 졸음 인식 모델 도출 및 성능 비교 - 기존의 단일 모달리티 딥러닝 모델 보다 인식 정확도 향상 (평균 2%) - 국내외 관련연구보다 우수한 인식 정확도를 보임 (최고 성능 96.15%) - 국내외 관련연구에 비하여 보다 적은 규모의 학습데이터 사용 - 시제품 개발을 통하여 도출된 인식 모델의 실시간 및 신뢰성 검증 - 가상 주행 환경의 데이터를 이용하여 상태 인식 모델을 도출 - 실제 주행 환경의 데이터를 이용하여 실제 적용 가능성을 검증 - 실시간 및 신뢰성 통신을 지원하며, 2.5초 이내로 인식 가능함을 보임				
연구개발성과의 활용계획 (기대효과)	- 본 연구개발은 철도, 선박, 항공 등 조종사가 존재하는 산업분야에 활용될 수 있음. 게다가 데이터의 종류(모달리티)를 추가 또는 변경하여 전혀 다른 분야에 활용될 수 있는 높은 확장성을 가지고 있음 - 우리나라 인공지능 기술의 전반적인 수준은 선진국에 비해 열세인 상황에서 경쟁력 확보가 가능한 분야에 집중 투자가 필요함. 멀티모달 딥러닝 분야는 국내외적으로 초기 단계이지만, 본 연구개발결과는 이 분야의 선도적 위치에 있기 때문에 선진국과의 기술격차 해소에 기여할 것으로 기대함				
국문핵심어 (5개 이내)	운전자 상태 인식	멀티모달 학습	딥러닝	시그널 분석	인공지능
영문핵심어 (5개 이내)	Driver State Detection	Multimodal Learning	Deep Learning	Signal Analysis	Artificial Intelligence

※ 국문으로 작성(영문 핵심어 제외)

< 목 차 >

1. 연구개발과제의 개요	5
가. 연구개발의 배경	5
나. 연구개발의 필요성	9
다. 연구개발의 범위	21
2. 연구수행내용 및 성과	33
가. 국내·외 관련연구 분석	33
나. 연구수행내용	46
다. 연구결과	79
라. 연구개발 성과	87
3. 목표 달성도 및 관련 분야 기여도	89
4. 연구개발성과의 활용 계획 등	94
붙임. 참고 문헌	98

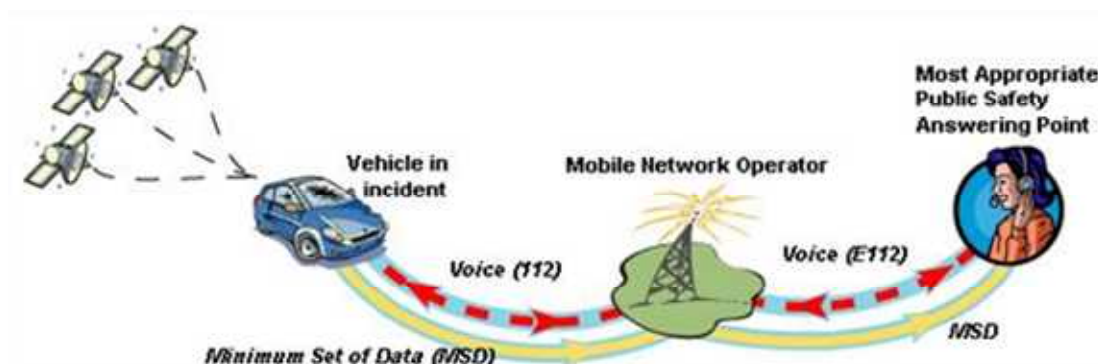
<별첨> 주관연구기관의 자체평가 의견서

1. 연구개발과제의 개요

가. 연구개발의 배경

(1) eCall(Emergency Call) 서비스 의무화

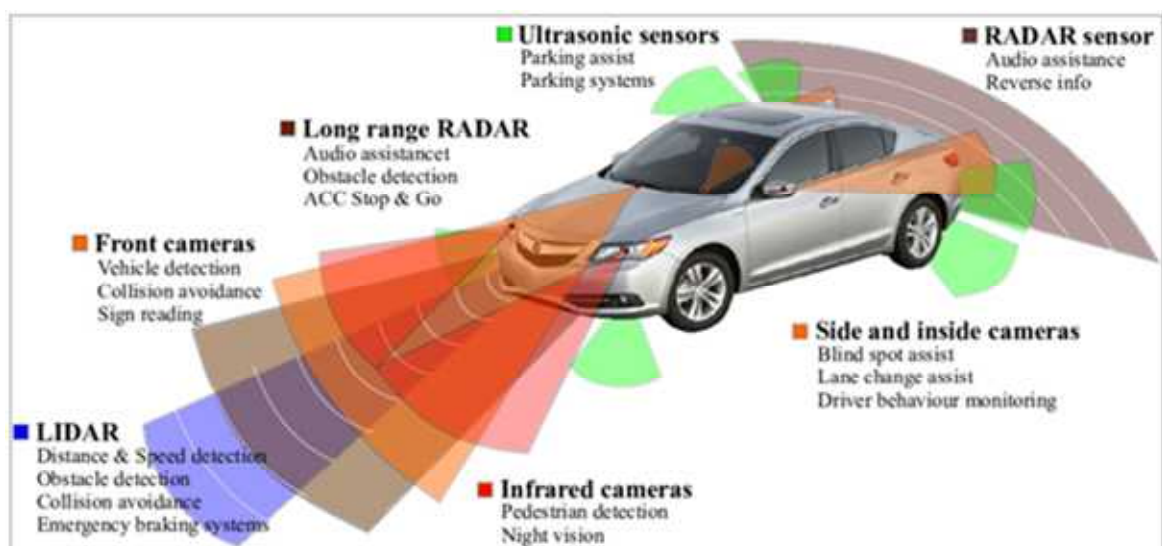
- eCall[1] 서비스는 차량 충돌 및 운전자 위급상황 발생 시 자동 또는 수동으로 긴급 구조기관에 사고 위치 등의 정보를 전송하고 긴급구난을 요청하여 신속한 구조 활동이 가능하도록 하는 서비스로 정의할 수 있음
- 유럽연합은 교통사고 사상자수를 줄이기 위해 2018년부터 모든 신차에 에어백 작동 등의 사고가 감지되면 자동으로 112를 연결하는 eCall 서비스를 제공할 계획이며, 러시아도 정부 주도의 ERA-GLONASS(Emergency Response to Accidents-GLONASS) 서비스를 설계 및 개발하여 2017년까지 모든 차량에 제공하는 것을 목표로 하고 있음
- 미국은 정부 주도의 eCall 서비스는 없지만 민간 사업자를 중심으로 유사한 서비스 개발이 이루어지고 있음. 대표적으로, GM사의 OnStar와 Ford사의 Sync 서비스는 차량 사고 및 긴급상황 발생 시 콜센터로 전화연결을 제공함. 최근 미국은 V2V(Vehicle to Vehicle) 서비스 탑재에 대해서는 제도적 기반을 마련해 의무화를 추진하고 있음
- 우리나라의 경우 세계적인 eCall 서비스 의무화 추세를 고려하여, 정부 차원의 한국형 e-Call 서비스 개발을 추진하고 있으며[2], 2018년 개발 완료를 목표로 미래창조과학부와 국토교통부가 공동으로 관련 연구개발 과제를 수행하고 있음



[그림 2] eCall 시스템의 동작 과정 (출처: EC 위원회)

(2) 자율주행을 앞당기는 ADAS(Advanced Driver Assistance System)

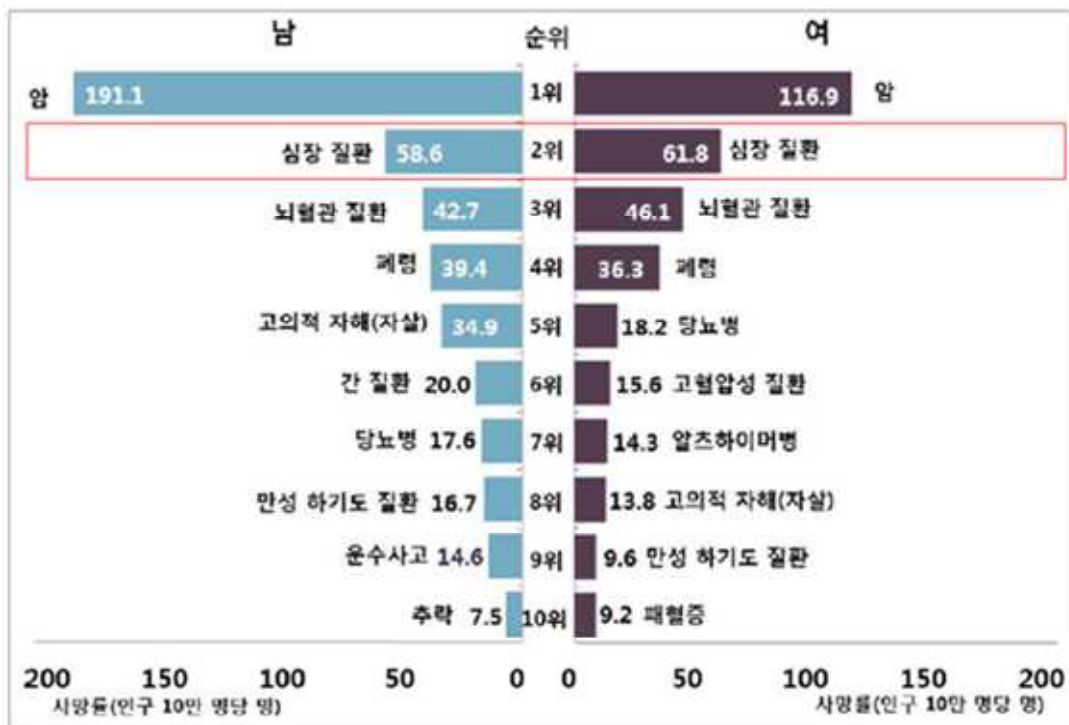
- ADAS는 미래 자율주행 자동차의 핵심 시스템으로, 현재는 주로 자동차 외부의 센서를 통해 전방위 교통사고 가능성을 운전자에게 제공하고 있음. 이를 통해 자동차 외부 요인으로부터 발생하는 교통사고를 회피하거나 완화하는데 기여할 수 있음
- ADAS는 차량에 탑재된 카메라와 융합 센서를 이용해 데이터를 수집하고 최신의 컴퓨터 알고리즘 기술을 이용하여 차량 주변의 각종 위험 상황을 정확히 인식함. 이를 통해 운전자에게 주의경보를 주거나 자동차 스스로 제동하여 사고를 방지함
- ADAS의 시장규모는 2016년 160억 달러에서 2021년 370억 달러까지 성장할 것으로 예상됨. 국내·외 주요 자동차 사업자는 2030년까지 인간의 개입 없는 완전 자율주행 자동차 개발을 핵심 로드맵으로 제시하였으며, 현재 고속도로 및 저속구간 자율주행과 같은 한정된 구간의 자율주행을 달성하기 위하여 차간거리유지, 자동제동, 차선유지 등의 기능별로 구분된 ADAS를 지속적으로 통합하고 있음
- 기능별로 구분된 ADAS의 통합에는 자동차 외부 요소 뿐 아니라 내부 요소도 포함되고 있음. 대표적으로 DSM(Driver State Monitoring), DEM(Driver Eyes Monitoring)과의 기능적 통합이며, 졸음, 주시태만과 같은 운전자 상태를 감지하여 자동으로 갓길에 정차하는 기능을 제공할 수 있음



[그림 3] ADAS의 개념도 및 탑재된 센서[3]

(3) 운전자로부터 발생하는 자동차 사고: 심장질환과 졸음

- 숙련된 운전자 또는 반응이 빠른 운전자일지라도 건강이상, 피로도 등의 운전자 개인의 생리적 변화에 따라 ADAS의 도움을 전혀 받을 수 없는 상황이 발생할 수 있음. 대표적으로 운전자에게 찾아온 졸음과 심장질환이며, 두 상황 모두 운전자의 의식이 빠르게 사라지기 때문에 ADAS만으로 사고를 막을 수 없는 상황이 발생할 수 있음
- 심장질환의 경우 통계청의 2017년 사망원인통계[4]에 따르면, 심장질환에 의한 사망이 2순위를 기록하였음. 심장질환에 의한 사망은 지난 10년 전에 3위에서 2위로 사망원인 순위가 상승한 사인임. 특히 급성 심장정지와 같은 경우, 대한심폐소생협회, 국민안전처 질병관리본부의 급성 심장정지 연구자료(2014년 기준)에 따르면, 급성 심장정지의 절반가량은 가정 외에서 발생하고 환자의 90%는 원래 심장질환자가 아님. 우리나라 급성 심장정지 환자는 연간 3만명 이상으로 매년 인구 10만명 당 50명이 돌연사 위험에 처해있음. 게다가 우리나라의 심장정지 환자의 심폐소생술 시행률은 12%로 미국, 일본 등 선진국의 절반 수준임



[그림 4] 2017년 사망원인통계 (출처: 통계청)

- 졸음운전의 경우 경찰청과 한국도로공사 고속도로 사고통계에 따르면, 2013~2016년 4년간 졸음운전 사고는 1만 72건이 발생하였으며 그 중 457명이 사망하였음. 2007~2016년 10년간 고속도로 교통사고 발생원인 1위는 졸음운전으로 인해 발생한 사고였음

(전체 교통사고의 22.5%). 우리나라 도로교통공단이 조사한 운전자 졸음운전 실태자료에 따르면, 전체 운전자의 약 80%는 1회 이상 졸음운전을 경험한 것으로 조사되었음(2011년). 또한, 우리나라 졸음운전 교통사고는 매 년 2천여 건이 발생하며, 교통사고 총 사망자의 2.2%는 졸음운전이 원인인 것으로 추산됨(2012년)



[그림 5] 지난 10년간 고속도로 교통사고 원인 통계 (출처: 한국도로공사)

나. 연구개발의 필요성

(1) 운전자에게 발생한 심장질환의 위험성

- 자동차는 밀폐된 개인 공간으로 운전자의 심장질환 발생 시 외부인이 이를 인지하기 매우 어렵기 때문에 대부분의 운전자는 심폐소생술 등의 적절한 구조를 받기 어려운 상황임. 그리고 이어지는 운전자의 의식불명으로 인한 추가적인 교통사고가 발생할 우려가 있음
- (실제 사례) 2017년 2월 8일 새벽 서울 돈암동 주택가에서 46세 김 모 씨가 운전하던 승용차가 전봇대와 근처 차량을 들이받는 사고가 있었으며, 심장정지 상태로 발견된 김 씨는 소방대원들에 의해 심폐소생술을 받았지만 의식을 회복하지 못함[5]
- 특히, 운송업과 관련된 근로자들의 좌식생활습관, 교대근무 등의 근무환경이 카테콜라민과 같은 호르몬 분비를 증가시켜 심혈관질환 발생률을 증가[6]시키는 것으로 보고되고 있음. 외국의 추적관찰연구[7]에 의하면, 버스 기사들에게서 고혈압 등의 발병과 흡연, 운동부족 등의 건강하지 않은 생활양식요인들이 심장질환 위험을 높인다고 하며, 국내 버스기사 대상의 연구[8]에서도 좌식 및 교대근무로 인한 운동 부족과 불규칙한 생활로 인하여 일반인 보다 심장질환 위험이 높다는 유의미한 결과를 보임
- [9]의 연구자들은 버스운수업 종사자들을 대상으로 심혈관질환 위험성을 분석하였으며 Belick 등(1994)은 32개 논문 중 28개 논문에서 직업운전 작업이 심혈관질환의 독립적 위험요인으로 입증된 대표적인 직업요인으로 보고 있음

(2) 운전자에게 발생한 졸음의 위험성

- 졸음운전의 경우 경찰청과 한국도로공사 고속도로 사고통계에 따르면, 2013~2016년 4년간 졸음운전 사고는 1만 72건이 발생하였으며 그 중 457명이 사망하였음. 2007~2016년 10년간 고속도로 교통사고 발생원인 1위는 졸음운전으로 인해 발생한 사고였음 (전체 교통사고의 22.5%)
- (실제 사례) 2017년 9월 2일 천안논산고속도로에서 고속버스가 앞서가던 승용차 7대를 들이받아 2명이 숨지고 9명이 다치는 대형 사고가 발생하였으며, 사고의 원인은 졸음운전으로 밝혀짐[10]
- 우리나라 도로교통공단(2011년)에서 특별교통안전교육 이수대상 운전자 362명을 대

상으로 졸음운전 경험 실태에 대해 조사한 결과에서 졸음운전을 경험한 비율이 전체 응답자의 79.3%에 이르렀으며, 이 가운데 졸음운전으로 인해 실제 교통사고를 경험한 비율이 15.7%, 법규를 위반한 비율은 21.3%였음[11]. 즉, 우리나라 도로환경에서 졸음운전으로 인한 교통사고 위험성이 높은 것으로 볼 수 있음

- 국외의 경우, 미국수면재단(National Sleep Foundation)의 2005년도 조사에서는 미국 전체 운전자의 60% 이상이 졸음운전을 경험한 적이 있으며, 전체의 4% 정도는 졸음운전으로 인한 교통사고를 실제로 겪거나 사고를 겪을 뻔한 경험을 가지고 있는 것으로 보고하고 있음. 졸음운전 교통사고의 특징은 주로 운전자 혼자서 운행하는 경우 상대적으로 많이 발생함(즉, 동승자와 상호작용을 하여 발생하는 각성이 없기 때문). 미국 뉴욕주에서 실시한 조사결과에 따르면 전체 졸음운전 사고의 82%가 1인 승차차량에서 발생한 것으로 나타남
- 미국 교통안전국의 연구보고서[12]는 장거리 트럭운전자들에 대한 차량 내 졸음운전 감시 및 경고장치의 효율성 평가에서 졸음운전을 하지 않았으나 졸음운전이라고 판단하여 경고하는 “잘못된 경고(false alarm)”의 비율이 높다는 문제점을 지적함



[그림 6] 졸음 운전의 위험성 (출처: 전국버스운송사업조합연합회, 한국운수산업연구원)

(3) 운전자의 심장질환 상태 인식에 대한 기술개발의 한계점

- 심장질환은 의학적으로 전기적(electrical) 문제로 발생하는 심장정지(cardiac arrest)와 혈액순환(circulation) 문제로 발생하는 심장마비(heart attack)로 구분됨. 대부분의 심장마비는 심장정지로 이어지지 않으며, 심장손상을 막기 위해 환자를 빠르게 병원으로 이송하는 것으로 환자의 생명을 보장할 수 있음. 하지만, 심장정지는 골든 타임 내 심폐소생술 또는 자동제세동기(Automated External Defibrillator, AED)를 통한 구조가 이루어져야 하는 매우 위급한 상황임[13]

- 현재의 eCall 서비스는 운전자가 호출 버튼을 누르거나 ADAS와 연계해 에어백 작동이 감지되면 서비스를 호출하는 방식임. 하지만 자동차에 직접적 손상 없이도 운전자가 사망에 이르는 운전자의 심장질환을 자동으로 인식하지 못하면, 적시에 eCall 서비스를 호출할 수 없는 문제점이 있음
- 운전자의 심장질환은 심장박동의 전위 변화를 나타낸 ECG(EleCtrocardioGraphy) 또는 광학센서로 혈류량의 변화를 감지하는 PPG(Photo PlethysmoGraphy)를 분석하여 감지할 수 있음. 운전자가 자연스럽게 접촉하는 자동차의 시트, 안전벨트, 운전대에 ECG/PPG 센서를 탑재하여 운전자의 심장활동을 모니터링하는 연구가 있음[14-16]
- ECG는 100ms 정도의 단위로 심장의 활동 전위 차이를 감지하여 기록하기 때문에 다양하고 정밀한 심장 활동정보를 도출할 수 있지만 진동, 가속 등의 움직임에 민감하게 반응하기 때문에 노이즈(noise)와 누락(missing value)이 빈번하게 발생하여 실제 주행환경에 적용하기 위해서는 이 문제를 해결해야 함



[그림 7] 운전자 심장활동 모니터링을 위한 시트형태의 측정 장치
Ford사의 프로토타입 시트(위), plessey사의 프로토타입 시트(아래)

- Ford와 plessey는 ECG의 노이즈와 누락을 최소화하여 운전자의 심장활동을 모니터링하는 시트를 개발하였음[17, 18]. 시트에 6개 이상의 ECG 측정 센서를 탑재하고 운

전자의 심장활동에 따라 발생하는 미세한 전위차를 측정하도록 고안된 시트임. 하지만 이 방법 역시 운전자가 착용한 옷의 두께, 계절에 따른 정전기 등의 요인으로 실제 상용화가 어려운 상황임. 또한, 해당 기술은 일부 기업이 독점적으로 보유한 기술이기 때문에 생산 단가가 높아 보급 시 매우 많은 비용이 소모됨([18]의 경우 시트 한 대에 \$18,500이며 시트를 통해 수집된 데이터를 활용하기 위해서는 별도의 협의가 필요)

- 국내의 경우 현대자동차그룹은 차량내 헬스케어 기술[19]을 적용하고 있음. 운전대 내장 센서를 통한 운전자의 PPG 측정, 시트 전파센서를 통한 운전자의 ECG 측정, 카메라 기반 영상 센서, 웨어러블 기기 등의 멀티센서 등 다각도로 운전자의 생체 정보를 수집하는 장치를 개발하고 있는 상황임
- PPG는 스마트 밴드/워치와 같은 웨어러블 기기에서 주로 사용하는 방식으로, 광학센서를 통해 혈류의 변화를 감지하는 방식이기 때문에 ECG 방식보다 진동, 가속 등에 덜 민감하여 운전자의 심장활동을 모니터링하기에 적합함 (단, ECG 보다 파형이 정밀하지 못하기 때문에 협심증, 심근경색, 동맥경화증 등의 구체적인 심혈관질환을 판단하는 근거로 사용될 수 없으며, 전문의를 통한 진단이 필요한 영역임)
- PPG 시그널을 이용해 심장활동의 이상을 도출하는 최근 연구들은[20-22] 정상과 비정상을 구분하는 이상탐지(Anomaly Detection) 기술로, 사람마다 형태가 다양한 파형을 정규화(Normalization)하고 파형의 대표성을 가지는 특징을 선정하여 추출(Feature Extraction)함. 이후 연구자들이 고안한 독자적인 알고리즘 또는 SVM(Support Vector Machine)과 같은 기계학습(Machine Learning)을 이용하여 이상 시그널을 자동으로 탐지하고 있음
- 이러한 방법은 대표성을 가진 연구참여자(피험자, subjects)와 충분한 규모의 (학습)데이터가 뒷받침되어야 함. 이 조건이 만족되지 않으면, 전통적인 시그널 분석 및 기계학습 기법을 통한 PPG 시그널 분석은 연구개발에 사용된 데이터에 과적합(overfitting)된 결과일 가능성이 높아짐. 즉, 제시된 연구결과가 우수할지라도 실험에 참여한 연구참여자에 종속된 알고리즘 또는 학습 모델일 수 있기 때문에, 실제 환경에서는 잘 동작하지 못하는 문제가 있음(예를 들면, 외국인 대상의 우수 학습모델일지라도 내국인 대상으로는 좋은 성능을 내지 못할 수 있음)
- 사람의 심장활동은 개인에 따라 편차가 큰 데이터임. 이를 일반화하기 위한 대표성을 가지는 데이터를 수집하는 접근방식보다, 자동차라는 공간이 개인화된 공간이라는 것에 초점을 맞추어 접근할 필요가 있음. 즉, 운전자가 확정된 시점부터 인식 모델을 만들고, 주행 시간이 늘어남에 따라 자연스럽게 인식 모델의 성능이 강화되는 방식이 본

연구 도메인에 적합함. 다만, 초기에 학습모델을 만들기 위해서는 여전히 충분한 수준의 학습 데이터가 필요한 문제가 있음

- 적은 규모의 학습 데이터로 인식 모델 생성이 가능한 방법으로 데이터 증강(Data Augmentation)기법[23]이 있음. 원본 데이터의 레이블은 유지한 채로 일정 수준의 변형을 가하여 학습 모델의 성능을 높이는 방법임. [24]는 웨어러블 센서 데이터를 증강하여 질병을 모니터링하는 연구로 데이터 증강을 통해 추가적인 데이터 수집 절차 없이 약 9% 정도의 성능 개선을 보였음
- 따라서, 주행환경에서 운전자의 심장질환을 보다 정확하고 실제적용이 가능한 수준으로 감지하기 위해서는, 데이터 증강에 기반하여 점진적으로 인식 성능이 개선되는 개인화가 가능한 기술이 적합함. 사전에 학습된 모델(Pretrained model)을 기반으로 위험상황을 인식하는 방식보다, 초기 운전자가 확정된 시점에 적은 규모의 데이터를 증강하여 심장질환을 인식하는 모델을 생성하고, 자동차를 사용하는 시간에 비례하여 인식 성능이 점진적으로 증가되는 방식이 실제 적용에 유리함

(4) 운전자 졸음 상태 인식에 대한 기술개발의 한계점

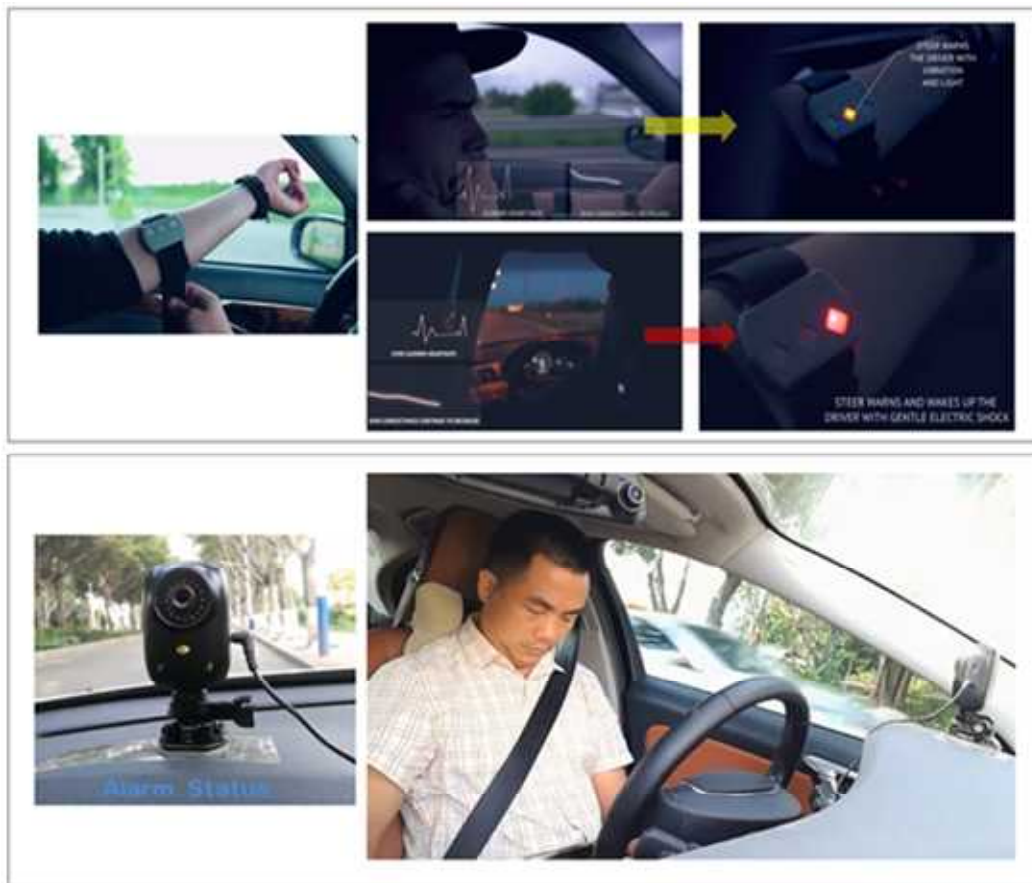
- 기존의 운전자의 졸음 상태를 판별하기 위한 연구개발은 아래와 같은 4가지 방식[25]으로 구분될 수 있음. 운전대, 주행 위치의 이상을 탐지하는 방식은 자동차 자체의 움직임을 통해 간접적으로 운전자의 졸음을 판단하는 방식이며, 운전자의 얼굴 이미지와 생체 신호를 측정하여 판단하는 방식은 운전자의 졸음을 직접적으로 판단함
 - 운전대 패턴 모니터링(steering pattern monitoring): 운전대 움직임에 따라 발생하는 진폭(amplitude)에 특정한 패턴이 있는지를 도출하여 졸음 상태를 판별함
 - 차선 내 자동차 위치 모니터링(vehicle position in lane monitoring): 자동차 외부에 위치한 카메라를 통해 수집한 주행 이미지를 통해 차선 이탈 여부를 도출하여 운전자의 졸음 상태를 판별함
 - 운전자의 눈/얼굴 모니터링 (driver eye/face monitoring): 자동차 내부 카메라로부터 수집한 운전자의 얼굴 이미지에서 얼굴의 각도, 눈이 감긴 정도를 도출하여 졸음 상태를 판별함
 - 운전자의 생체신호 측정(physiological measurement): 심장활동, 뇌활동 등 졸음에 따라 수반되는 생리학적 신호의 변화를 감지하여 졸음 상태를 판별함
- 기존의 졸음 상태 판별 기술들을 통해 도출된 졸음 정보 시스템은 화물차 기사와 같이

차량을 정차하지 않고자 하는 운전자들이 주로 사용하고 있음. 하지만, 잘못된 경보가 반복되면 운전자로 하여금 잘못된 경계 수준을 설정하도록 하여(경고장치 작동에 둔감해짐) 결과적으로 사고예방에 도움을 주기 어렵다[26]는 지적이 있음. 대표적으로, 운전대 패턴 모니터링과 같이 간접적인 방식으로 운전자의 졸음을 인식할 경우, 졸음 패턴으로 설정해둔 운전대 패턴이 특정 도로 주행환경과 일치할 경우, 단순히 차선을 변경하는 경우에도 운전자가 졸고 있다고 판별하는 문제가 있음

- 운전자의 불규칙한 심박수나 눈이 감기는 등의 생체정보는 가장 객관적인 운전자 상태 정보를 담고 있기 때문에 국내·외 관련 연구개발에서 이를 활용하여 운전자의 졸음을 판별하였음. 그러나, 졸음이 발생할 때 사람마다 생성되는 데이터의 편차가 크기 때문에 운전자의 졸음 상태를 인식하는 일반화된 모델을 도출하기 매우 어려움
- 운전자의 생체신호를 기반으로 졸음을 판별하는 방식은 졸음 상태 전후로 우리 몸의 생리적 변화를 측정하는 방식이기 때문에 가장 정확도가 높은 방식임. 한국안전학회가 2015년 발표한 논문에 따르면[27], 졸음은 생리적으로 졸린 상태(부교감 신경 활성화)가 된 이후에 느끼게 되는 것으로 졸음이 밀려오기 전 생체 신호를 통해 곧 졸음이 밀려올 것임을 예측할 수 있음
- 부교감 신경이 활성화되면 심장 박동이 평소보다 낮아지고, 사람에 따라 차이는 있지만 수분에서 20분 내외로 졸음이 밀려오는 것을 느끼게 됨. 즉, 졸음이 쏟아지는 모든 운전자는 공통적으로 부교감 신경이 활성화되어 심장 박동이 낮아진다는 의학적 사실을 바탕으로 보다 정확한 졸음 인식 모델을 도출할 수 있음
- 이러한 생체신호의 변화를 감지하여 운전자의 졸음을 감지하고 각성을 부여하는 스마트 밴드 형태의 제품이 개발되고 있음. [28]은 클라우드 펀딩 서비스 Kick Starter에 올라온 STEER 프로젝트로 2초 단위로 심박수와 피부전도성을 측정하며, 평소 보다 심박수와 피부전도성이 낮아지면 진동 및 전기충격이 가해지는 제품을 개발중에 있음
- 운전자 얼굴 이미지를 분석하여 졸음 상태를 인식하는 방식은 상대적으로 높은 인식 정확도를 보이고 있음. 근 적외선(near infrared) 카메라를 통해 어두운 환경에서 운전자의 감은 눈(closed eye)을 인식할 수 있기 때문에 낮과 밤에 덜 민감하다는 장점도 있음. 하지만, 이 역시 얼굴 위치가 측면을 향하거나, 태양광에 따른 얼굴 찌푸림, 감정에 따른 얼굴변화 등이 다양하기 때문에 인식률을 높이는데 한계가 있음. 즉, 이미지 처리(image processing) 방식의 졸음 인식 정확도는 생체신호 기반에 비하여 즉각적인 인식이 가능하지만, 외부 환경에 영향을 받는 경향이 있음. 그로 인하여, 운전자가 실제로 졸지 않았지만 졸린 척 연기를 하여도 해당 졸음 감지 장치가 동작하는 잘못된

졸음 인식 결과가 산출됨(Leisure Auto에서 개발한 졸음 감지 카메라[29]는 [그림 7-아래]와 같이 시연 연상에서 졸고 있는 연기를 하여도 졸음 경보가 울림)

- 1.-나.-(3)에서 제시한 한계점과 같이 주행환경에서 운전자의 심장활동을 모니터링 하는 것은 측정 장치의 비용 높고, 노이즈와 누락 값 문제에서 자유로울 수 없음. 운전자의 심장 활동이 졸음을 판별하는 결정적 지표인 것은 사실이지만, 노이즈와 누락 값에 대한 문제 해결이 없다면 단독으로 사용되기 어려운 방식임. 따라서, 운전자 얼굴 이미지 기반의 접근 방식과 생체신호 기반의 접근 방식을 모두 채택하여 상호 보완될 수 있는 방법을 고안한다면, 기존의 졸음 인식 모델보다 안정적으로 운전자의 졸음을 인식할 필요가 있음



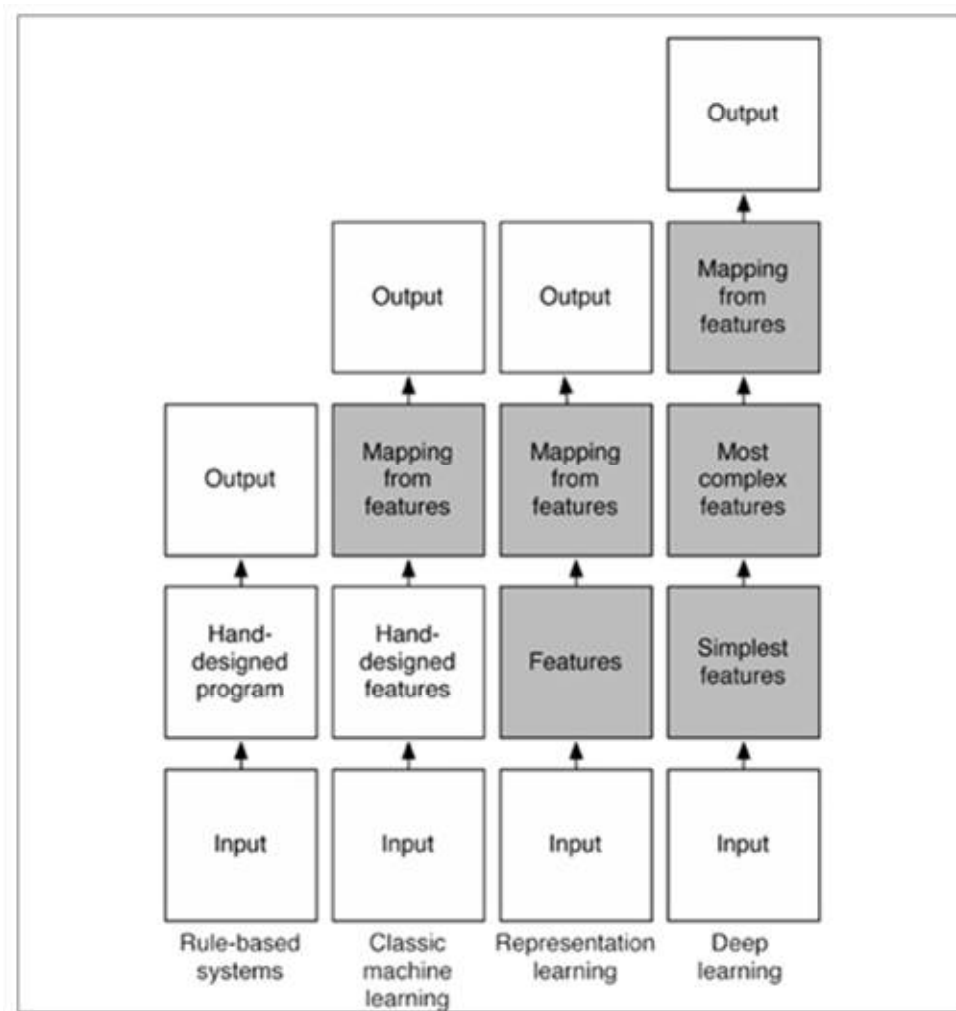
[그림 8] 생체 신호 기반의 운전자 졸음 감지 프로젝트 STEER(위),
얼굴 이미지 기반의 운전자 졸음 감지 제품 LS 803(아래)

(5) 딥러닝 기반의 보다 일반화된 운전자 상태 인식 기술의 필요성

- 컴퓨터 프로그램에 인공지능을 부여하는 방법은 아래와 같이 두 가지로 볼 수 있음

- 규칙 기반 시스템(Rule-based system): 기계가 인간처럼 사고하도록 규칙(rule)을 이용하여, 대표적으로 체스를 두는 프로그램이 있음. 규칙에 따라 컴퓨터 프로그램은 가능한 경우의 수를 조회하여 결과를 산출하는 방식
- 기계학습(Machine Learning): 인간이 학습하는 방식을 컴퓨터에 구현한 것으로, 데이터와 해당 데이터와 쌍을 이루는 정답을 기계에게 부여하고, 새로운 데이터가 입력으로 왔을 때 기계 스스로 정답을 예측하는 방식
- 자율주행 자동차, 음성 비서, 의료 진단 등 현대의 인공지능 시스템은 대부분 기계학습 방식을 활용하여 동작하고 있음. 그렇다고 규칙 기반 시스템이 현대에 전혀 사용되지 않는 것은 아님. 해결하고자 하는 문제에 따라 규칙 기반 시스템이 유용한 경우도 있음(특징의 차원이 작은 경우 규칙 기반 시스템이 유용함)
- 체스보다 경우의 수가 많은 바둑의 경우 인간이 그 많은 경우에 수에 대하여 규칙을 지정하는 것은 불가능에 가까움, 또한 자율 주행과 같이 수많은 경우의 수에 대하여 동작하는 경우 역시 규칙 기반 시스템으로는 안전한 자율주행이 불가능함
- 머신러닝은 고차원적인 문제를 해결하는 중요한 접근 방법이며, 운전자의 졸음과 같이 특정한 규칙으로 정의하여 일반화되기 어려운 고차원적인 문제를 해결하는데 적합한 방법임. 하지만, 1.-나.-(3), 1.-나.-(4)에서 제시한 것과 같이 대표성을 가지는 학습 데이터를 수집하고 문제를 해결하기 위한 주요 특징을 사람이 수작업으로 직접 수행(Hand-designed features, [그림 8]-Classic machine learning)해야 하는 단점이 있음
- 이러한 머신러닝의 특징추출 과정은 연구팀 또는 전문가가 수작업으로 진행하는 방식으로 여러 가설을 세우고 성능 측정을 통해 가장 적절한 특징을 도출하는 것이 관건임. 1.-나.-(3), 1.-나.-(4)에서 제시한 기존의 머신러닝 방식으로 운전자의 심장질환과 졸음 상황을 인식시키기 위해서는 다양한 특징과 차원을 고려하여 수많은 실험을 진행하고 검증해야 함. 특히, 사람마다 졸음이 발생할 때 나오는 특징은 다양하고, 현재까지는 일반화된 수학적 모델이 없기 때문에 전통적인 기계학습으로는 보다 일반화된 인식 모델을 도출하는데 한계가 있음
- 딥러닝(Deep Learning) 역시 기계학습과 동일하게 통계적 기법을 기반으로 데이터를 이용하여 컴퓨터 시스템에 학습 능력을 부여하는 방법임. 전통적인 기계학습과 딥러닝의 주요 차이점은 [그림 8]의 Representation Learning, Deep Learning과 같이 특징추출(Feature extraction) 과정에 인간이 최소로 개입하는 것임(특징 추출 과정은 입력 데이터에 대하여 어떠한 특징이 주요 특징인지 인간이 명시적으로 지정하는 것을 의미하며 [그림 8]의 음영으로 처리된 부분은 알고리즘이 스스로 수행)

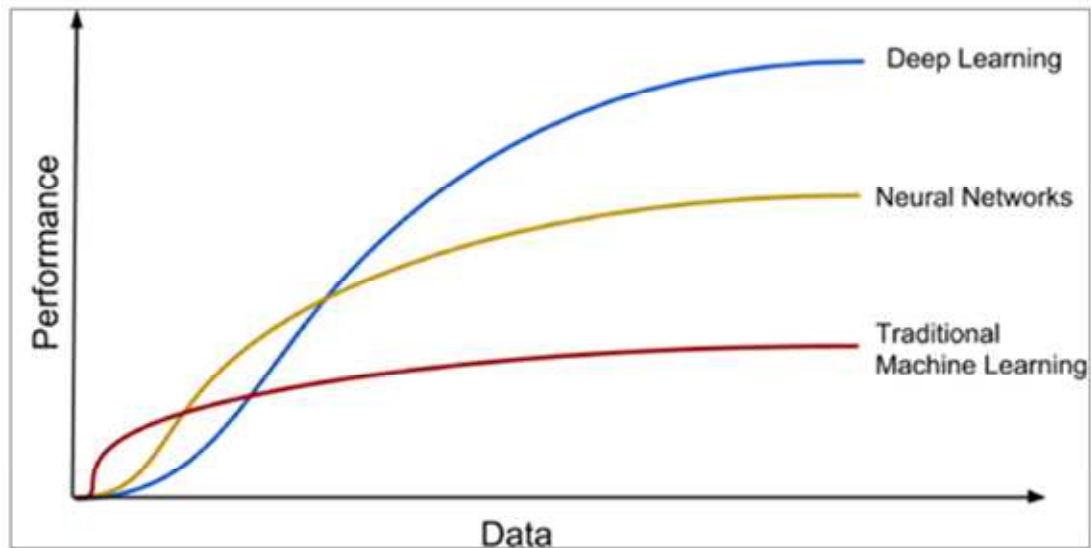
- 딥러닝은 특징 추출과 학습과정이 알고리즘 자체에 내장되어 있다는 측면에서 기존의 전통적인 기계학습과 큰 차이를 보임. 즉, 학습 과정에 인간의 개입을 최소화 하였으며, 그로 인해 인간이 파악하기 어려운 고차원적인 특징을 기계 스스로 추출하여 보다 일반화된 학습 모델을 도출할 수 있는 장점이 있음



[그림 9] 딥러닝과 머신러닝의 구조적 차이
(Deep Learning, Yoshua bengio et al, 2016)

- 하지만, 딥러닝을 적용하기 위해서는 전통적인 기계학습 보다 큰 규모의 학습 데이터가 요구됨. 일반적으로 딥러닝은 충분한 규모의 학습 데이터가 뒷받침 된다면, [그림 9]와 같이 전통적인 기계학습의 한계를 넘어서 우수한 성능의 인식 모델을 도출하는 것이 가능함
- 딥러닝을 이용한 운전자 상태 인식 모델은 기존의 전통적인 기계학습 방식보다 더욱 일반화된 인식 모델임. 하지만 실제 주행환경에 적용되기 위해서 인식 성능의 안정성

을 강화할 필요가 있음. 멀티모달 딥러닝(Multimodal Deep Learning)[30]은 학습 모델 도출에 필요한 데이터의 종류를 다양하게 함으로써, 기존의 단일 종류의 데이터만 사용한 딥러닝 방식에 안정성을 강화함. 이는 특정 종류의 데이터가 과하게 학습 결과에 영향을 미치는 의존성을 낮추고, 복합적인 요인으로 발생하는 문제를 해결하는데 매우 적합한 접근 방법임



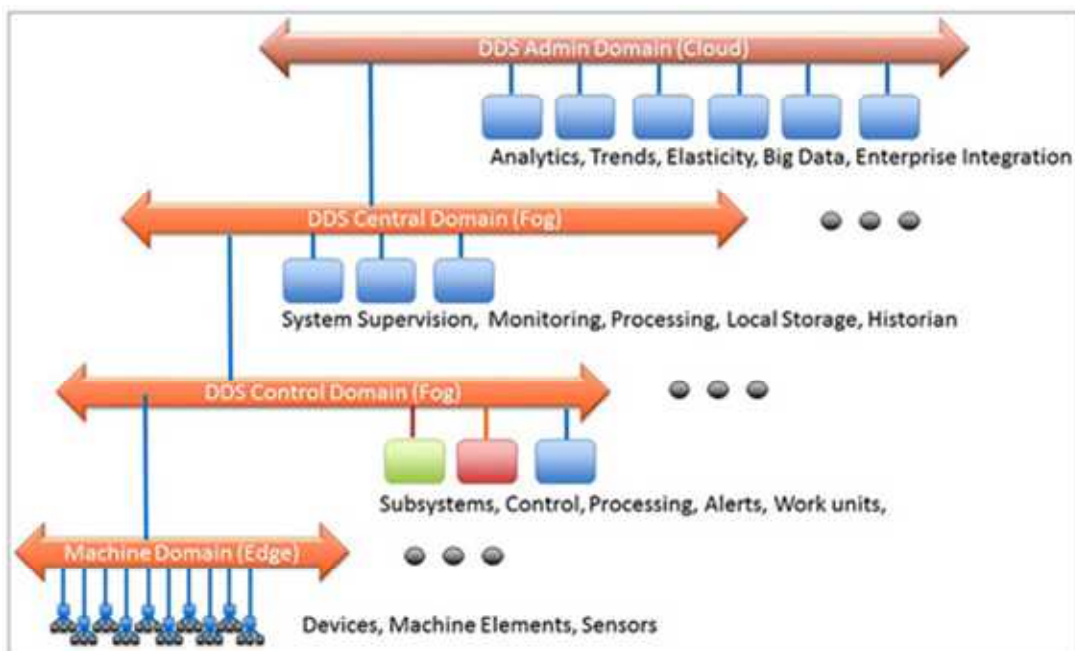
[그림 10] 학습 데이터의 규모에 비례하여 성능이 증가하는 딥러닝

(6) 실시간 미들웨어 기반의 안정적인 데이터 처리 기술의 필요성

- 현대의 자동차는 단순 동력 기계장치를 넘어 복합 전자장치로 나아가고 있으며 CPU (Central Processing Unit)을 중심으로 각종 센서와 전자장치 모듈이 결합되어 수집된 센서 데이터를 처리 및 분석하고 각종 프로그램이 실행되는 일종의 컴퓨터 시스템으로 진화하고 있음
- 하지만 자동차는 일반 컴퓨터와 다르게 주행에 대한 안전을 보장해야 하는 특수성을 가짐. 만약 시속 100km로 주행하는 자동차가 있을 때, 운전자가 1초를 졸게 되면 무의식 주행으로 28m를 이동하게 됨. 이와 같은 위급 상황이 발생했을 때 미래의 자동차는 운전자로부터 5초 이내로 제어권을 이양을 결정[31, 32]할 수 있도록 권장하고 있음
- [33]의 결과와 같이 자동차에 탑재되는 저사양 임베디드 시스템은 운전자의 얼굴 이미지를 300ms 내외로 실시간으로 분석이 가능한 수준이며, NVIDIA의 Drive AGX[34]와 같은 차세대 차량용 하드웨어의 경우 고성능 GPU(Graphics Processing Unit)을

탑재하여 1080p 해상도의 운전자 얼굴 이미지를 초당 30장 이상 실시간 처리가 가능할 것으로 예상됨(자동차 제조업체 Volvo는 차량용 HW 플랫폼으로 NVIDIA Drive AGX를 채택하여 2020년부터 생산되는 자동차에 탑재할 예정)

- 이처럼 현재의 그리고 미래의 자동차 HW는 1.-나.(5)에서 제시한 딥러닝 기반의 인식 모델을 동작시키기에 충분한 상태임. 하지만, eCall 등의 구조 요청을 가능하게 하는 3G 이동 통신과 웨어러블 기기에서 사용되는 BLE(Bluetooth Low Energy) 등의 자동차 커넥티비티(connectivity) 관련 IT 통신 기술들은 데이터가 손실될 가능성이 있는 신뢰성(reliability)이 보장되지 않는 네트워크임
- 데이터 전송의 신뢰성이 보장되지 않는다는 것은 우선순위를 부여하지 않아 데이터 전송이 지연될 수 있는 비실시간 통신 방식을 의미하며, 통신 장치가 늘어나는 것에 비례하여 데이터 전송의 품질은 떨어지는 방식임. 따라서, 운전자의 위급 상황에 관련된 데이터 전송에 적합한 방식이 아님



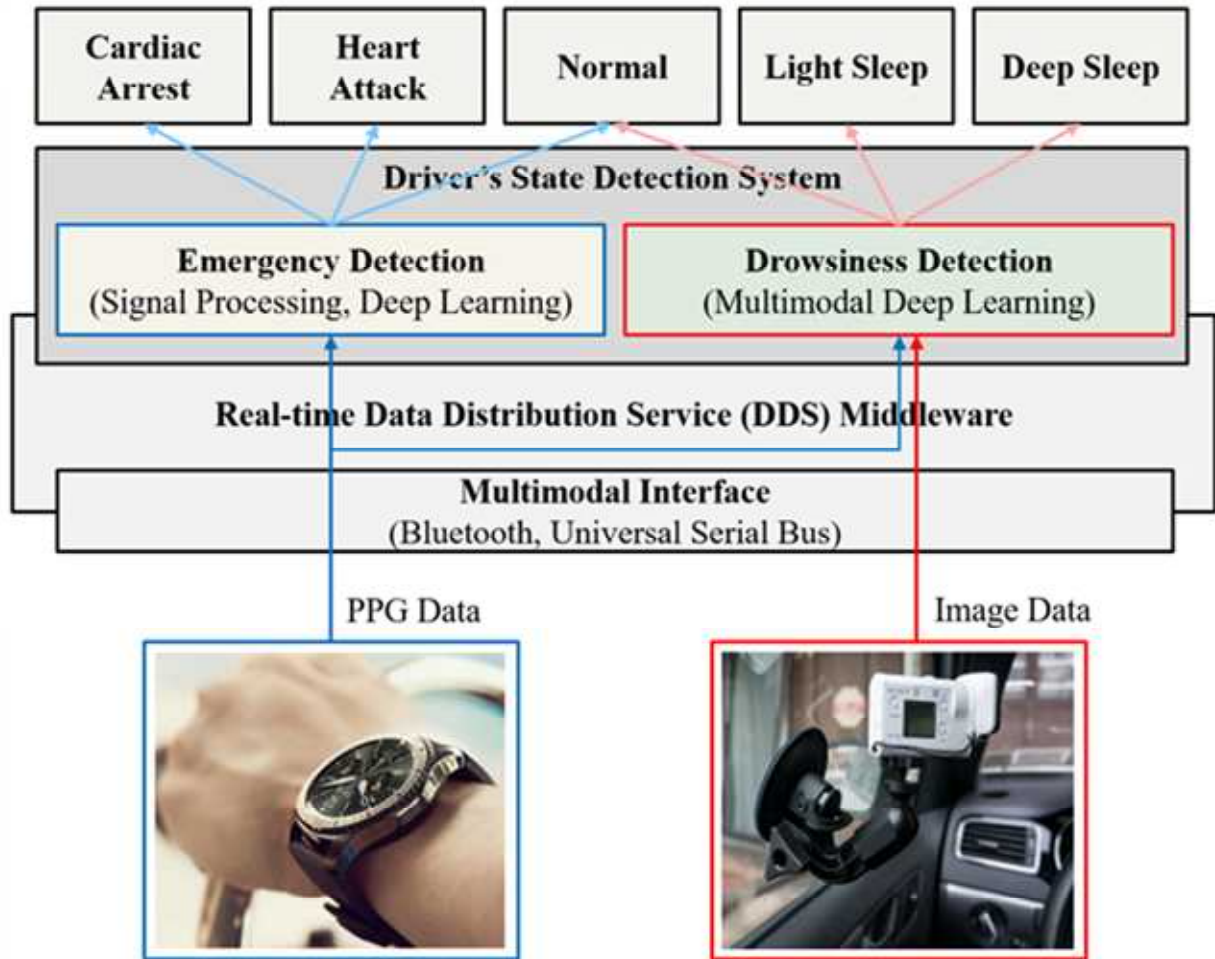
[그림 11] 도메인에 따라 다양하게 적용 가능한 DDS[35]

- DDS[35]는 OMG(Object Management Group)에서 제정한 실시간 데이터 교환 표준으로, 본래 국방과학 분야의 전자장치 간 실시간 및 신뢰성있는 데이터 상호운용을 보장하기 위해 제정되었음. 이미 RTI, PrismTech, 한국전자통신연구원 등 국·내외 다양한 벤더들이 DDS 표준을 구현하여 국방과학 전자장치에 탑재하고 있음
- 자동차 분야에서의 DDS의 활용은 대표적으로 Audi사 인데, 이미 2012년 부터 RIT에

서 개발한 DDS를 이용해 자동차 전자장치와 그를 이용한 응용들이 안정적으로 통신 및 실행되는지를 검증하기 위한 테스트 플랫폼으로 사용하고 있음[36]

- DDS의 가장 큰 특징은 22가지 이상의 서비스 품질(Quality of Service, QoS)을 통해서 다른 전자장치 간 데이터 교환에 우선순위를 부여하거나(실시간성, Deadline QoS), 전송 수준을 보장(신뢰성, Reliability QoS)하는 것임. 약 100ms의 단위로 측정되는 PPG 수집 장치가 수집한 데이터를 지연 없이 다른 모듈로 전달해야 하는 경우 Deadline QoS 설정을 통해 생성된 샘플이 수집 장치로 일정한 주기(interval)에 따라 전달될 수 있도록 실시간성을 보장할 수 있음
- DDS는 실시간 미들웨어 표준 기술로서 [그림 10]과 같이 다양한 도메인에서 활용될 수 있을 만큼의 확장성을 보유하고 있음. 자동차는 Machine Domain에서 DDS Control Domain에 해당하며 각종 전자장치와 센서들 사이에 안정적인 데이터 상호교환과 이를 기반으로 동작하는 제어 유닛, 데이터 처리 모듈 등과 연계될 수 있음

다. 연구개발의 범위



[그림 12] 본 연구개발의 개념도

딥러닝을 기반으로 운전자의 심장질환과 졸음 상태를 인식하기 위한 본 연구개발의 개념도는 [그림 11]과 같음. PPG 데이터 측정 장치와 차량 내 카메라를 이용하여 주행 상황에서 운전자의 심장활동 데이터와 및 얼굴 이미지를 수집함. 멀티모달 인터페이스(Multimodal Interface)는 장치로부터 수집된 데이터를 본 연구개발의 시스템으로 가져오기 위한 접점에 해당하며, 통신 프로토콜은 웨어러블 기기와 카메라가 대중적으로 지원하는 블루투스 통신과 USB(Universal Serial Bus)를 지원함. 인터페이스를 통과한 운전자의 PPG 및 얼굴 이미지 데이터는 [그림 10]과 같은 차량 내 타 전자장치와의 확장성을 고려하고, 본 연구개발의 상태인식 시스템에 안정적으로 제공하기 위하여 실시간 통신middleware 표준 DDS를 기반으로 데이터를 전송함. 수집된 PPG 데이터는 위급상황 인식(Emergency Detection) 모듈에

제공되어 신호처리(Signal Processing) 및 딥러닝을 이용해 운전자의 심장활동이 정상(Normal), 심장정지(Cardiac arrest), 심장마비(Heart attack) 인지를 판단함. 운전자의 PPG 데이터는 위급상황 인식 모듈 뿐 아니라 졸음 감지(Drowsiness Detection) 모듈에도 제공되며, 이 PPG 데이터는 얼굴 이미지와 동시에 입력되어 운전자의 졸음을 더욱 정확하게 인식하는 멀티모달 딥러닝에 사용됨. 졸음 감지 모듈은 멀티모달 딥러닝을 통해 도출된 인식 모델이 탑재되어 있어, 주행 상태에서 실시간으로 운전자의 정상(Normal), 얇은 잠(Light sleep), 깊은 잠(Deep sleep)을 인식함. 운전자 상태 인식 시스템(Driver's State Detection System)은 위급상황 및 졸음 인식 모듈의 결합으로 구성되며 최종적으로 5가지의 운전자 상태를 산출함

- 정상(Normal): 운전자는 깨어 있으며 심장활동도 평소와 같은 안정된 상태
- 심장정지(Cardiac arrest): 심장의 전기적(electrical) 문제로 심장활동이 멈춘 상태이며 운전자는 무반응의 상태
- 심장마비(Heart attack): 심장의 혈액순환(circulation) 문제로 심장활동에 이상이 있는 상태이며 운전자는 의식이 있으나 극심한 가슴 통증을 호소하는 상태
- 얇은 잠(Light sleep): 부교감 신경의 활성화로 평소보다 낮은 심박수를 보이며 의식은 있으나 금방이라도 졸음이 발생할 가능성이 있는 상태
- 깊은 잠(Deep sleep): 졸음으로 인해 운전자는 의식이 사라진 상태

본 연구개발의 운전자 상태 인식 시스템은 아래와 같은 세부 연구범위를 설정하고 연구개발을 진행하여 도출되었음(본 연구팀이 수행한 구체적인 연구 수행내용은 2장 참조)

(1) 주행 상황에서 발생하는 운전자의 심장활동 및 얼굴 이미지 수집

- 멀티모달 딥러닝 기반의 운전자 상태 인식모델을 도출하는 것을 목표로 하기 때문에 딥러닝의 전제 조건인 충분한 규모의 학습 데이터 확보가 뒷받침 되어야 우수한 성능의 인식 모델을 도출할 수 있음
- 일부 국내·외 연구팀들은 주행 상황의 운전자 얼굴 이미지를 비디오 파일 형태로 공유하고 있으나 거의 모든 데이터가 정상(깨어있는) 상태의 운전자 얼굴 이미지임. 본 연구개발의 도메인인 졸음을 감지한다는 측면에서, 이 데이터를 이용해 졸음 감지 모델을 도출하는 것은 불가능한 상태임
- 특히, 본 연구개발의 멀티모달 딥러닝은 서로 다른 소스의 데이터를 동시에 학습하는

방식이기 때문에, 운전자 얼굴 이미지가 수집된 시점과 동일한 시점에 수집된 운전자의 심장활동 데이터가 쌍으로 존재하는 특수한 형태의 데이터임. 따라서, 본 연구팀은 기존에 존재하지 않는 학습용 데이터를 직접 수집하고 이를 기반으로 운전자의 상태를 인식하는 모델을 도출함

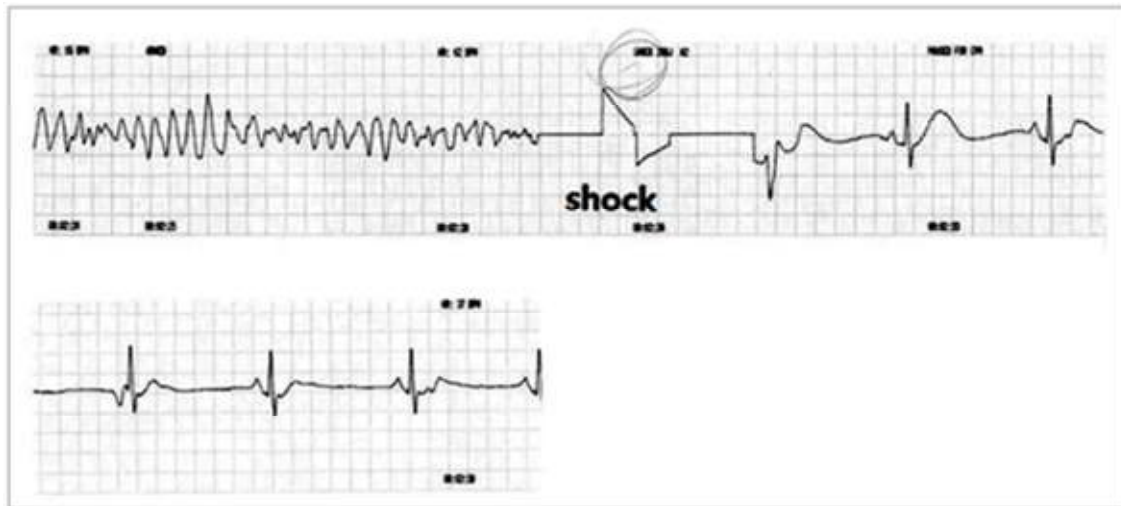
- 실제 주행환경에서 운전자의 졸음 데이터를 수집하는 작업은 사고의 위험이 있기 때문에 본 연구팀은 아래와 같이 3가지 방식으로 윤리적 문제를 최소화 하여 운전자의 졸음과 심장활동 정보를 수집함
 - 온라인 데이터 수집: 연구 목적으로 온라인에 공개된 주행 관련 데이터를 수집하고 분석하여 본 시스템의 설계와 알고리즘 구현에 활용함
 - 가상 주행 환경에서 데이터 수집: 실제 주행환경과 유사한 가상의 주행 환경을 구축하고 연구참여자를 모집하여 주행 중 발생하는 PPG 및 얼굴 이미지를 수집함. 수집된 데이터를 기반으로 시그널 처리 및 멀티모달 딥러닝 기법을 적용하여 운전자의 상태인식 시스템을 도출함
 - 실제 주행 환경에서 데이터 수집: 장거리 운전자들을 대상으로 실제 주행환경에서의 데이터를 수집함. 수집된 실제 데이터는 가상 주행 환경을 통해 도출된 상태인식 시스템이 실제 주행 환경의 데이터에서도 유효한지 검증하는데 사용함

※ 본 연구팀은 모든 연구참여자에게 본 연구에 대한 충분한 설명(문)을 제공하였으며 수집된 데이터는 연구참여자의 개인정보가 포함되었기 때문에 관련 법규에 따라 개인정보 활동동의서를 받아 규정을 준수하여 수집된 데이터를 관리하였음

(2-1) 시그널 처리 및 딥러닝을 통한 운전자 심장질환 상태 판별

- 심장질환은 의학적으로 전기적(electrical) 문제로 발생하는 심장정지(cardiac arrest)와 혈액순환(circulation) 문제로 발생하는 심장마비(heart attack)로 구분됨. 심장정지의 경우 심장이 정지하기 직전에 [그림 12-(위)]과 같은 심실세동(Ventricular Fibrillation)이 나타나는 특징이 있음. 즉, 심전도(ECG)에 전혀 변화가 없는 것이 아니고, 심장의 수축 이완이 없어 보이나 심장 표면에 가늘게 떨리는 현상이 발생하여 미약한 심전도 기록이 발생하며[37], 그 파형은 정상적인 파형에 비해 진폭(amplitude)과 파형의 높이가 매우 짧아지는 명확한 특징이 나타남
- PPG는 스마트 밴드/워치와 같은 웨어러블 기기에서 주로 사용하는 방식으로, 광학센

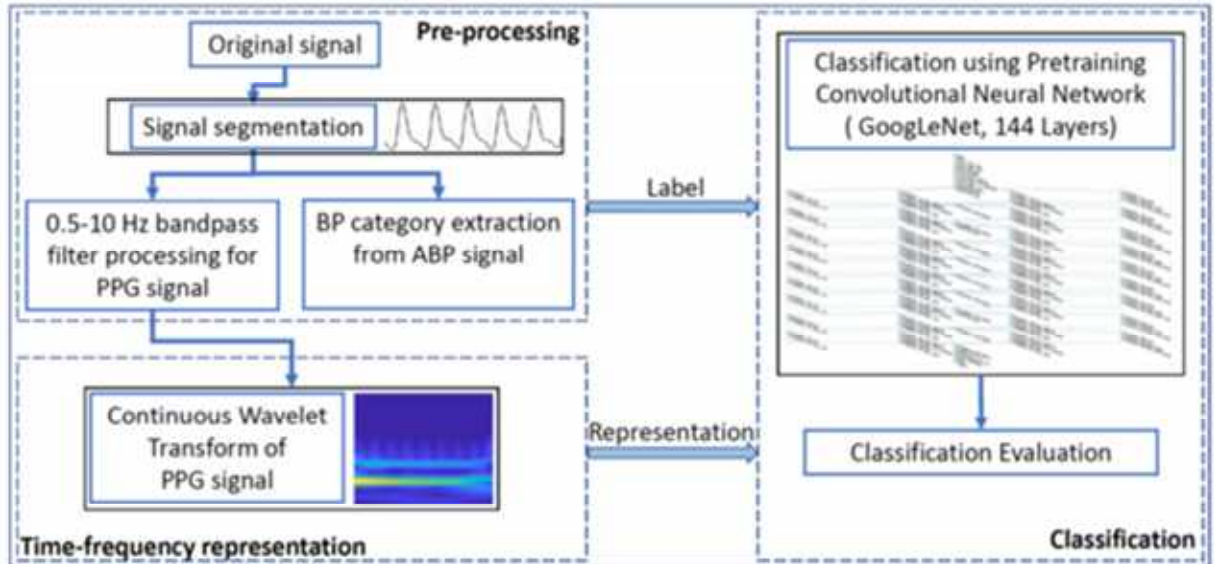
서를 통해 혈류의 변화를 감지하는 방식이기 때문에 ECG 방식보다 진동, 가속 등에 덜 민감하여 운전자의 심장활동을 모니터링하기에 적합함. 반대로, ECG 보다 파형이 정밀하지 못하기 때문에 협심증, 심근경색, 동맥경화증 등의 구체적인 심혈관질환을 판단하는 근거로 사용될 수 없음. 따라서, 본 연구개발은 의학적으로 질병을 판단하는 영역은 연구개발에서 제외됨



[그림 13] 심장정지 직전의 심전도 기록과 전기충격요법을 통한 심전도 변화(위), 전기충격요법 이후 정상리듬을 회복한 심전도 기록(아래)[37]

- PPG는 ECG에 비하여 주행환경의 진동, 가속 등의 움직임에 덜 민감하여 노이즈와 누락이 상대적으로 적게 나타남. 하지만, PPG 역시 광학센서가 피부에 접촉된 영역, 운전자의 습관 등에 따라 노이즈와 누락 값이 충분히 발생할 수 있음. 이러한 문제를 해결하기 위하여 시그널 처리와 전통적인 기계학습 분야의 다양한 연구[20-22]가 있음
- 딥러닝은 특징 추출과 학습과정이 알고리즘 자체에 내장되어 있다는 측면에서 기존의 전통적인 기계학습과 큰 차이를 보임. 즉, 학습 과정에 인간의 개입을 최소화 하였으며, 그로 인해 인간이 파악하기 어려운 고차원적인 특징을 기계 스스로 추출하여 보다 일반화된 학습 모델을 도출할 수 있는 장점이 있음
- [38]의 연구는 PPG 시그널에서 질병을 예측하는 딥러닝 기반의 연구로 기존의 전통적인 기계학습과의 성능비교가 제시되어 있음. 이 연구는 [그림 13]과 같이 간단한 시그널 처리를 통해 PPG 시그널을 분할(segmentation)을 수행하고, 분할된 PPG 시그널에 대하여 웨이블릿(Wavelet) 변환을 가한 이미지 형태의 데이터를 CNN(Convolutional Neural Network)을 이용해 분류하는 연구임. 이 연구의 예측 모델은 기계학습 알고리

증인 KNN(K-Nearest Neighbor Algorithm)보다 약 6% 내외의 예측 성능 개선을 보였음

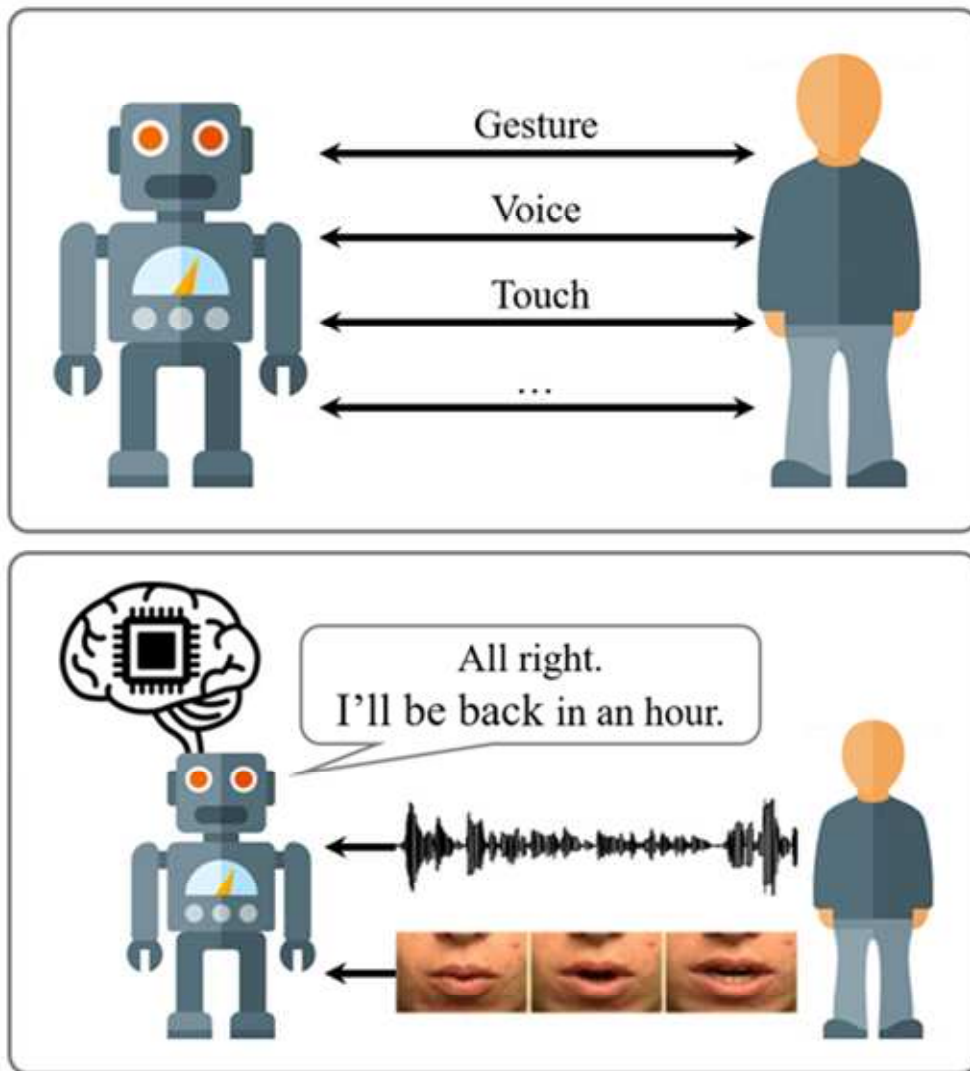


[그림 14] 딥러닝을 이용한 PPG 시그널 분석 연구[38]

- 하지만, [그림 9]와 같이 딥러닝이 우수한 학습 모델을 도출하기 위해서는 전통적인 기계학습 보다 큰 규모의 학습 데이터가 요구됨. 학습 데이터 확보 측면에서 심장정지, 심장마비 등과 같은 민감한 의료 데이터를 확보하는 것은 법적, 제도적인 문제로 현실적으로 매우 어려운 과정임. 수집 절차에 문제가 없다고 하여도, 그 규모가 딥러닝을 적용할 만큼의 충분한 규모로 수집하는 것 또한 현실적으로 큰 걸림돌임
- 적은 규모의 학습 데이터로 인식 모델 생성이 가능한 방법으로 데이터 증강(Data Augmentation)기법[23]이 있음. 원본 데이터의 레이블은 유지한 채로 일정 수준의 변형을 가하여 학습 모델의 성능을 높이는 방법임
- 데이터 증강에 의한 변형은 단순한 변형에서부터 다양한 도메인으로의 변형까지 그 방법이 다양하게 존재하며, 모든 방법이 딥러닝의 성능 향상에 기여할 수 있는 것은 아니기 때문에 도메인에 적합한 증강 기법을 찾아 적용하는 것이 중요함. 대표적으로 [24]는 웨어러블 센서 데이터를 증강하여 질병을 모니터링하는 연구이며 추가적인 데이터 수집 절차 없이 약 9% 정도의 성능 개선을 보였음
- 본 연구개발은 주행환경에서 운전자의 심장질환을 보다 정확하고 실제적용이 가능한 수준으로 감지하기 위해서, 데이터 증강에 기반하여 점진적으로 인식 성능이 개선되는 개인화가 가능한 기술이 적합함. 초기 운전자가 확정된 시점에 적은 규모의 데이터

를 증강하여 심장질환을 인식하는 모델을 생성하고, 자동차를 사용하는 시간에 비례하여 인식 성능이 점진적으로 증가되는 방식의 연구개발을 진행하였음

(2-2) 멀티모달 딥러닝 기반의 운전자 졸음 상태 감지 모델 도출

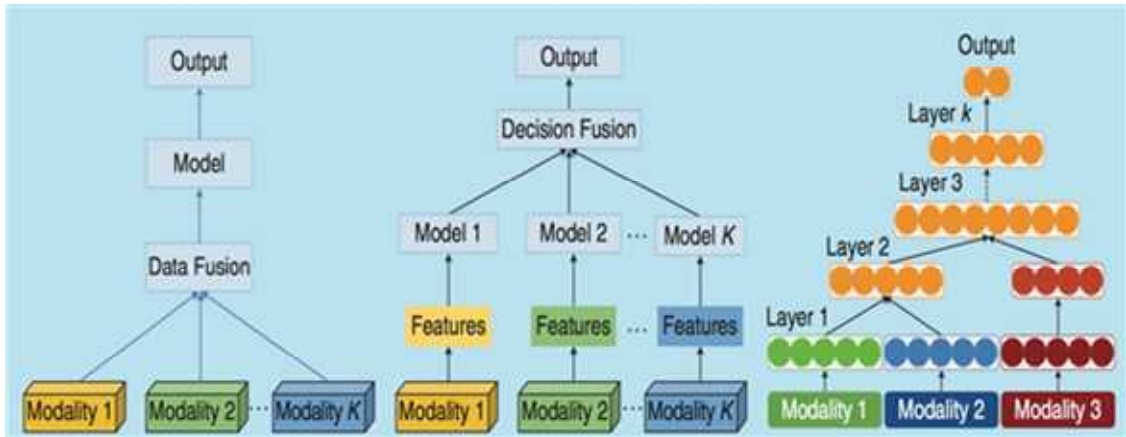


[그림 15] 카메라-제스처, 마이크론-보이스, 압력-터치로 매핑되는 기계의 모달리티 개념도(위), 멀티모달 방식으로 학습하는 인공지능 시스템 개념도(아래)

- 사람이 시각, 촉각 등 여러 감각을 이용해 정보를 인식하듯, 최근 딥 러닝에서도 여러 종류의 입력 데이터를 이용하여 학습하는 멀티모달 학습 연구가 진행되고 있으며, 단일 데이터로부터 학습모델을 생성하는 것 보다 높은 정확도를 보이고 있음
- 직관적으로 사람이 주변 상황을 인식할 때 하나의 감각기관을 사용하지 않고 여러 감

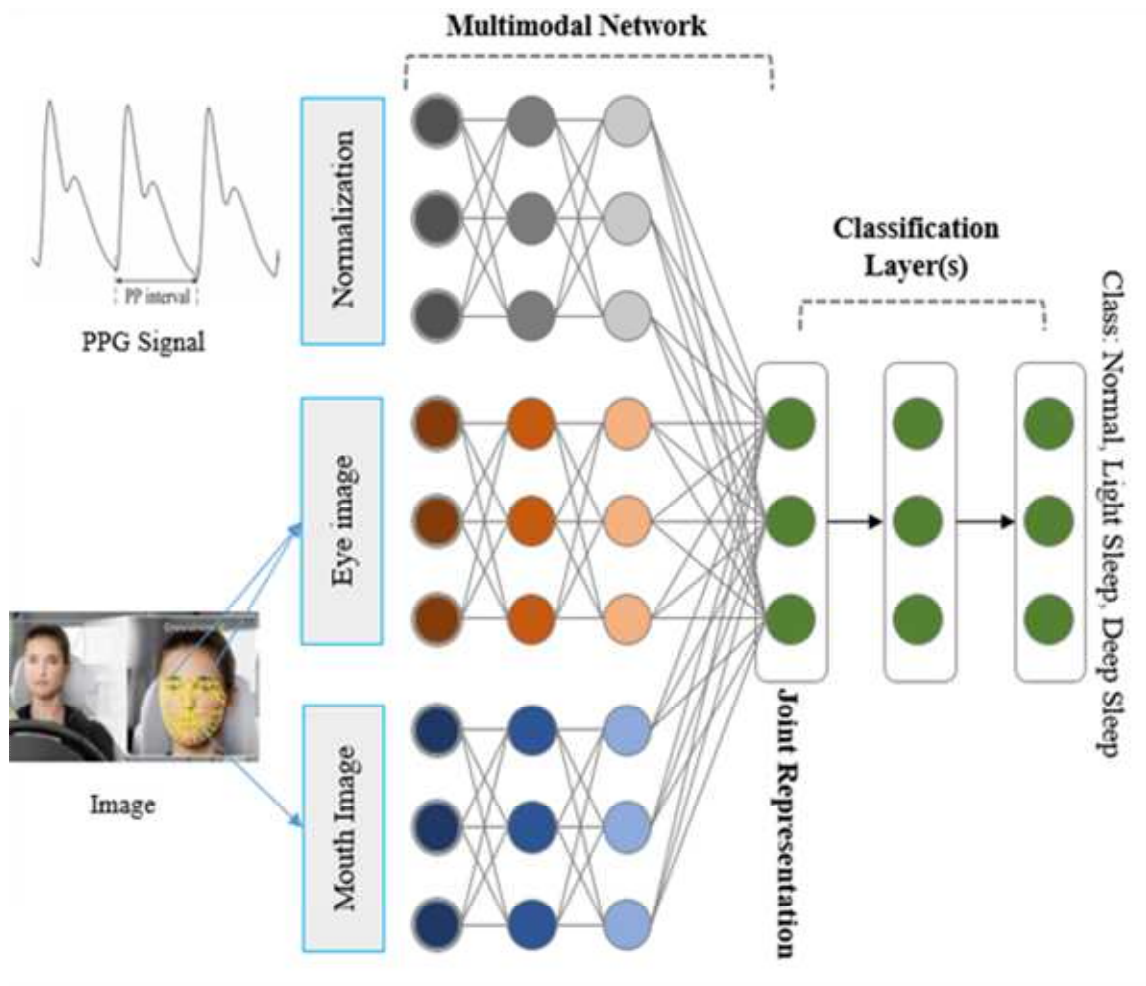
- 각기관을 함께 사용하여 상황을 인식하는 것이 인지 성능 및 속도 면에서 빠르다고 알려져 있음. 최근의 다양한 상황 인식 및 예측 분야에 활용되는 딥러닝 기술 역시 다양한 모달리티(Modality, 기계의 감각기관 역할)를 고려한 연구개발이 진행되고 있음
- 서로 다른 이질적인 센서 데이터의 정확한 결과를 산출하며, 복잡도와 차원이 높은 데이터에서 패턴을 도출함
 - 유니모달(Unimodal, 하나의 모달리티만 사용하는 연구)보다 멀티모달 방식이 더욱 향상된 인식 성능을 보임
 - 대표적으로 의료 영상분석에서 CT, MRI, 초음파 이미지를 멀티모달 학습 하여 보다 정확한 수준의 진단이 가능함을 보임
 - [그림 14]은 멀티모달 딥러닝을 직관적으로 설명하는 개념도임. 기계는 하드웨어 적으로 다양한 센서를 보유할 수 있으며, 독립된 센싱 모듈을 통해 수집되는 데이터는 독립적인 모달리티에 해당함. 멀티모달 딥러닝은 [그림 14-(아래)]와 같이 두 가지 이상의 모달리티를 활용하여 상황을 인식하는 인공지능 시스템의 개념도이며, 이러한 인공지능 시스템은 단일 모달리티 방식보다 5~6% 성능 증가를 기대할 수 있음[28]
 - 멀티모달 딥러닝은 [그림 15]와 같이 3가지 방식의 유연한 모달리티 융합 방식을 사용하여 학습 모델을 도출할 수 있으며 아래와 같은 특징이 있으며 높은 성능을 위해 데이터의 종류들에 따라 모달리티 융합 방식을 선택할 수 있는 유연한 방법임
 - 데이터 융합(Data Fusion): 학습과정의 입력 데이터로 들어가기 전에 모든 모달리티의 데이터들 하나의 특징벡터(feature vector)로 결합(concatenation)됨. 이 결합 방식은 센서로부터 바로 들어온 원천 데이터(raw data)들의 단순 결합이기 때문에 복잡도(차원이 높음)의 문제로 모델 도출을 위해 많은 데이터가 요구될 수 있음
 - 결정 융합(Decision Fusion): 각 모달리티의 데이터를 이용해 도출된 학습 모델의 인식(예측) 결과를 종합적으로 판단하여 최종 인식 결과를 도출하는 방식임. 기존의 앙상블(Ensemble) 방식의 기법과 유사하나 모달리티가 다르기 때문에 최종 인식 결과의 산출 방식을 다양하게 고려해야 함. 예를 들면, 결정 퓨전에 입력되는 각 모델의 결과값을 점수제로 환산하여 가장 높음 점수의 결과를 최종 결과로 채택하는 방식이나, 각 모델의 인식 결과를 투표(voting)하여 가장 많은 득표가 있을 경우를 최종 인식 결과로 산출함
 - 중간물 융합(Intermediate Fusion): 멀티모달 딥러닝에서 도전적으로 사용하는 융합 방법으로 특정 단계의 레이어에 결합되어 학습 과정 중간 중간에 직접적으로 영향을 미치게 됨. 이 경우 자칫 특정 모달리티가 학습에 과하게 기여하여 올바른 학습 모델

이 도출되지 못할 수 있으나, 적정 수준의 융합 지점을 찾게 되면 서로 다른 모달리티에 대하여 인간의 개입 없이 스스로 특징을 추출하고 인식 결과를 산출할 수 있음



[그림 16] 멀티모달 딥러닝의 3가지 멀티모달 네트워크 구성 방식

- 하나의 데이터 형태라고 할지라도 그 데이터 안에는 여러 모달리티가 존재할 수 있음. 대표적으로, 얼굴 이미지는 하나의 카메라를 통해 하나의 이미지에 담기지만, 사람의 얼굴은 눈, 코, 입 등의 여러 기관이 존재하기 때문에 각각 서로 다른 모달리티로 볼 수 있음. 졸음의 경우, 사람은 졸음이 발생하면 눈에 가장 큰 변화가 나타나지만 입 주변의 근육들이 평소와 다르게 이완되어 입의 모양의 변화가 나타남. 즉, 눈과 입은 근본적으로 서로 다른 기관이기에 졸음 발생 시 각각 고유의 특징이 나타나게 됨
- 이러한 멀티모달 딥러닝의 내용에 따라 본 연구개발은 [그림 16]과 같은 연구범위를 설정하여 연구개발을 진행하였음. 심박활동 정보를 담은 운전자의 PPG 데이터와 해당 PPG 데이터가 생산되는 시점의 눈과 입의 변화를 담은 운전자 얼굴 이미지 데이터를 학습 데이터로 사용함. 눈, 입, PPG 세 종류의 데이터는 멀티모달 네트워크(Multi modal Network)를 통해 데이터의 복잡도(차원 축소)가 줄어들면서 졸음을 구분하는 지표인 주요 특징들이 알고리즘 수준에서 자동으로 추출됨. 추출된 특징들은 [그림 15]와 같은 모달리티 융합 방식에 따라 적정 수준의 Joint Representation에서 하나의 데이터 형태로 표현됨. 마지막으로 분류 레이어(Classification Layer)에서는 일반적인 딥러닝 방법을 사용하여 세 가지의 졸음 분류 결과를 도출함(Normal/Light Sleep/Deep Sleep)



[그림 17] 멀티모달 딥러닝 기반의 운전자 졸음 감지 모델의 개념도

- 기존의 딥러닝 기반의 졸음 인식 연구들(2장의 국내·외 관련연구 분석 참조)은 정상이거나 졸음(의식이 사라진 상태)을 구분하는 연구임. 정상인 사람과 이미 꾸벅꾸벅 졸고 있는 사람은 같은 사람이 특별한 사고 없이 인식할 수 있을 정도로 명확히 구분되는 특징이 있기 때문에, 충분한 데이터만 확보된다면 딥러닝을 이용해 비교적 쉽게 인식 모델 도출이 가능함
- 하지만, 주행 환경에서 이미 졸고 있는 운전자를 인식하는 인공지능 시스템은 실제 졸음운전 사고 예방에 어떠한 도움도 줄 수 없음. 인공지능 시스템은 정상과 졸음 사이에 졸음이 밀려오는 시점을 인식할 수 있어야 실질적으로 졸음운전 사고를 예방하기 위한 조치를 스스로 취할 수 있음
- 졸음이란 것은 운전자 개인이 주관적으로 느끼는 것이기 때문에 얕은 잠(Light Sleep)과 깊은 잠(Deep Sleep)을 구분하는 것은 매우 도전적인 연구라고 볼 수 있음. 본 연

구개발은 주행이라는 특수한 환경에서 졸음을 인식하는 연구개발이기 때문에 운전자의 의식 여부를 졸음 인식의 주요 지표로 설정하였음. 따라서, 금방이라도 의식이 사라지고 졸음에 빠질 수 있는 경우 얇은 잠(운전자는 의식이 있음)으로 인식하고, 운전자의 눈이 완전히 감기고 의식이 사라져 주행 제어권이 상실된 경우는 깊은 잠으로 인식하는 멀티모달 딥러닝 기반의 운전자 졸음 감지 모델을 도출하였음

(3) 운전자 상태인식 시스템의 실시간 및 신뢰성 높은 데이터 전송 기법 적용



[그림 18] 자동차의 인포테인먼트 시스템 (출처: whichcar)

- 현재의 자동차는 CPU(Central Processing Unit)을 중심으로 각종 센서와 전자장치가 융합된 거대한 컴퓨터 시스템으로 발전하고 있음. 자동차 외부의 센서를 통해 수집된 데이터를 즉각적으로 처리 및 분석하여 자율적으로 움직이거나, 자동차 내부의 센서를 통해 운전자를 위한 각종 서비스 프로그램이 실행되는 일종의 컴퓨터 시스템으로 나아가고 있음
- 자동차 제조사들은 이 컴퓨터 시스템을 적극 활용하기 위해 인포테인먼트(Infotainment) 시스템을 개발하고 있음. 현대자동차의 Intelligent Personal Cockpit, 삼성전자의 Digital Cockpit, Google의 Android Auto, Apple의 CarPlay 등 전통적인 자동차 제조사 뿐 아니라 전자 및 IT 회사들도 인포테인먼트 플랫폼을 개발하고 있음. NVIDIA의 경우 차세대 차량용 하드웨어에 고성능 GPU(Graphics Processing Unit)를 탑재하여 2020년부터 Volvo의 신차에 탑재할 예정임. 즉, 앞으로의 자동차는 더욱더 다양한 서비스를 인포테인먼트 시스템을 중심으로 제공할 수 있을 것으로 기대

- 하지만, 현재의 인포테인먼트 시스템에 대한 기업별 자체 표준을 사용하기 때문에 인포테인먼트 시스템의 확장성과 장치 연결에 대한 표준화된 체계가 부족한 상황임[39]. 특히, eCall 등의 구조 요청을 가능하게 하는 3G 이동 통신과 웨어러블 기기에서 사용되는 BLE(Bluetooth Low Energy) 등의 자동차 커넥티비티(connectivity) 관련 IT 통신 기술들은 데이터가 손실될 가능성이 있는 신뢰성(reliability)이 보장되지 않는 네트워크라는 측면에서 문제가 있음
- 데이터 전송의 신뢰성이 보장되지 않는다는 것은 우선순위를 부여하지 않아 데이터 전송이 지연될 수 있는 비실시간 통신 방식을 의미하며, 통신 장치가 늘어나는 것에 비례하여 데이터 전송의 품질은 떨어지는 방식임. 따라서, 운전자의 위급 상황과 관련된 데이터 전송에 적합한 방식이 아님
- 만약, 시속 100km로 주행하는 자동차가 있을 때, 운전자가 1초를 줄게 되면 무의식주행으로 28m를 이동하게 되는 위험한 상황이 발생함. 자동차의 센서로부터 올바르게 센싱된 데이터가 있더라도 통신 레벨의 지연 또는 누락이 발생한다면, 자동차는 결국 적시에 올바른 동작을 수행할 수 없거나 이미 사고가 발생한 뒤 상황을 인지할 가능성이 있음. 수집된 운전자의 PPG 및 이미지 데이터는 일종의 시계열(Time Series) 데이터로 측정된 데이터가 상태감지 시스템으로 전달되기까지 누락과 지연이 발생한다면 적절한 시점에 상태 인식이 어려울 수 있음
- 위급 상황이 발생했을 때 미래의 자동차는 운전자로부터 5초 이내로 제어권을 이양을 결정[31, 32]할 수 있도록 권장하고 있음. 이러한 실시간 및 신뢰성 있는 데이터 통신 관점에서, 자동차는 우리가 흔하게 접하는 일반적인 컴퓨터와 다르게 운전자의 안전을 보장해야 하는 특수성을 가짐. 이러한 안전을 보장해야 하는 특수성은 국방과학 분야의 전투기, 항공모함, 전차에 탑승하는 파일럿의 안전을 제공하는 특수성과 유사함
- DDS는 국방과학 분야의 전자장치 간 실시간 및 신뢰성 있는 데이터 교환을 보장하기 위해 제정되었으며, 한국전자통신연구원, RTI 등 국내·외 다양한 벤더들이 DDS 표준을 구현하여 국방과학 전자장치에 적용하고 있음. DDS는 가장 큰 특징은 여러 종류의 서비스 품질(Quality of Service, QoS)을 정의하여 전자장치 간 데이터 교환에 우선순위를 부여하거나(실시간성), 전송수준을 보장(신뢰성)하는 것임
- 자동차 분야에서의 DDS의 활용은 대표적으로 Audi사 인데, 이미 2012년 부터 RIT에서 개발한 DDS를 이용해 자동차 전자장치와 그를 이용한 응용들이 안정적으로 통신 및 실행되는지를 검증하기 위한 테스트 플랫폼으로 사용하고 있음[36]
- 자율주행 및 전기차의 시대로 들어서는 지금, 차량에 탑재되는 하드웨어 플랫폼은 점

점 연경성을 중요하게 생각하고 있으며 많은 전자장치와 V2X들은 통신을 기반으로 서비스를 제공하고 있음. 이러한 상황에서, 운전자의 심장질환 및 졸음은 분초를 다투는 매우 위험하고 위급한 상황에 자동차 스스로 이를 인지하고 eCall을 통해 도움을 요청하여 ADAS를 통해 자율적으로 위험상황을 해결하기 위해서는 안정적인 통신 기술을 바탕으로 각종 전자장치와 소프트웨어가 동작해야할 필요가 있음

2. 연구수행내용 및 성과

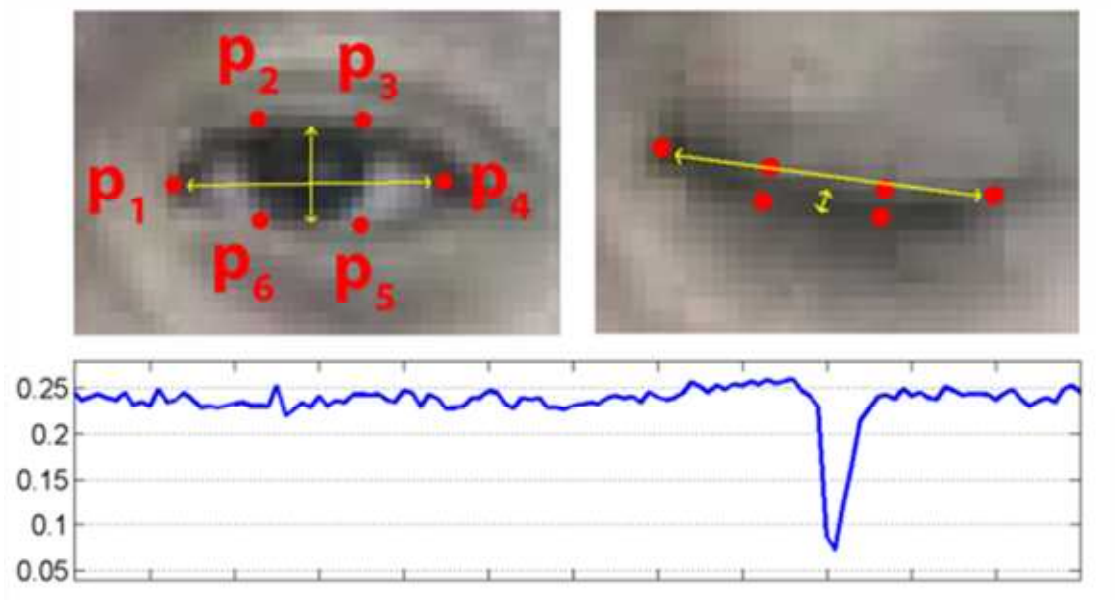
가. 국내·외 관련연구 분석

(1) 운전자의 상태를 인식하는 방법

[표 1] 센서에 따라 운전자의 상태를 인식하는 방법과 장점 및 한계점[42]

Sensor	Dependent Variables	Advantages	Limitations
Eye-tracking	Eye blinks	• Non-intrusive • Comprehensive indicator	• Costly • Noise
Proximity Sensor	Eye blinks	• Low cost • Reliable • Commercially available	• Cannot track eye movement, such as fixation duration
Accelerometer Sensor	Head movements	• Non-intrusive • Low cost	• Noise from sensor location and car vibrations
EEG	Brain's electrical activity	• Reliable • High temporal resolution	• Costly • Uncomfortable • Environmental artifacts (e.g. head movement)
Heart Rate	Heart rate	• High accuracy • Low cost • Low intrusive	• Confounded by age and health exercise
Respiration Rate	Respiration rate	• High accuracy	• Sensor may slip • Uncomfortable
Galvanic Skin Response	Skin conductance	• Easy to attach	• Noise for environmental weather
Driving Dynamics	Lane position, speed, etc.	• Non-intrusive • Easily accessible • Low cost	• Noise (vehicle type, road condition, etc.) • Generalizability

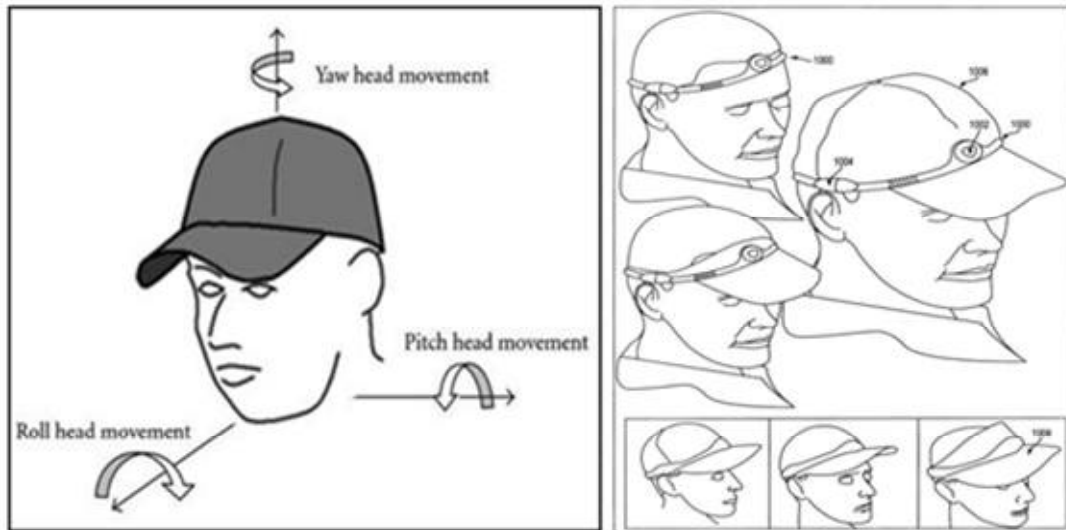
- 국내·외 다양한 연구개발에서는 운전자의 상태를 인식하기 위하여 다양한 종류의 센서를 사용하고 있음. [표1]은 센서에 따라 수집될 수 있는 운전자의 상태 정보와 해당 센서를 사용할 경우 운전자 및 인식 시스템에서의 장점 및 한계점을 담고 있음
- 시선추적(eye-tracking)과 근접센서(proximity sensor) 운전자의 눈 주변에 센서를 부착하거나, 스마트 글래스 착용하거나, 카메라를 이용하는 방식임. 연속된 이미지(또는 데이터) 프레임에서 동공의 변화를 지속적으로 모니터링하는 방식으로 [34]와 같이 피부와 눈을 구분하는 랜드마크를 검출하여 눈의 움직임을 추적하는 이미지 처리 방식이 주로 사용됨. 센서나 스마트 글래스가 아닌 카메라를 이용하는 경우 운전자에게 직접적으로 센서 장치를 강요하지 않기(non-intrusive) 때문에 간편하다는 장점이 있음. 특히, 인간의 눈은 정상 상태와 그 외 상태(졸음 등)를 확연하게 구분하는 종합적인 지표(comprehensive indicator)임. 하지만, 상대적으로 처리비용과 센서의 가격이 비싸며 주변 빛에 대한 노이즈가 발생할 가능성이 높다는 한계점이 있음



[그림 19] 랜드마크(P_i)를 통해 뜬 눈과 감은 눈을 자동으로 검출하는 방법[43]

- 가속도 센서(accelerometer sensor)는 머리의 움직임(head movements) 정도를 측정하여 운전자의 상태를 판별하는 방법임. 이 방법은 항공기가 공중에서 기동중 방향을 전환하기 위해 회전하는 3가지 축(pitch, roll, yaw)을 [그림 19]과 같이 머리의 움직임에 축으로 사용함. 가속도 센서는 현재 대부분의 스마트 기기에 탑재되어 활용될 만큼 보편화된 센서이기 때문에 도입 비용이 상대적으로 저렴하다는 장점이 있으며, 센서

의 크기도 매우 소형화되었기 때문에 모자 또는 머리띠를 착용하는 정도로 운전자에게 큰 불편을 주지 않음. 하지만, 센서가 당초 설계된 위치와 다른 머리의 위치에서 동작할 경우 잘못된 값을 생산하는 단점이 있어 운전자가 정확한 위치에 착용되어야 하는 한계가 있음. 또한, 차동차 주행에 따라 발생하는 진동에 가속도 센서가 반응하면서 잘못된 본래 의도와는 다른 값을 생산하는 한계점이 있음

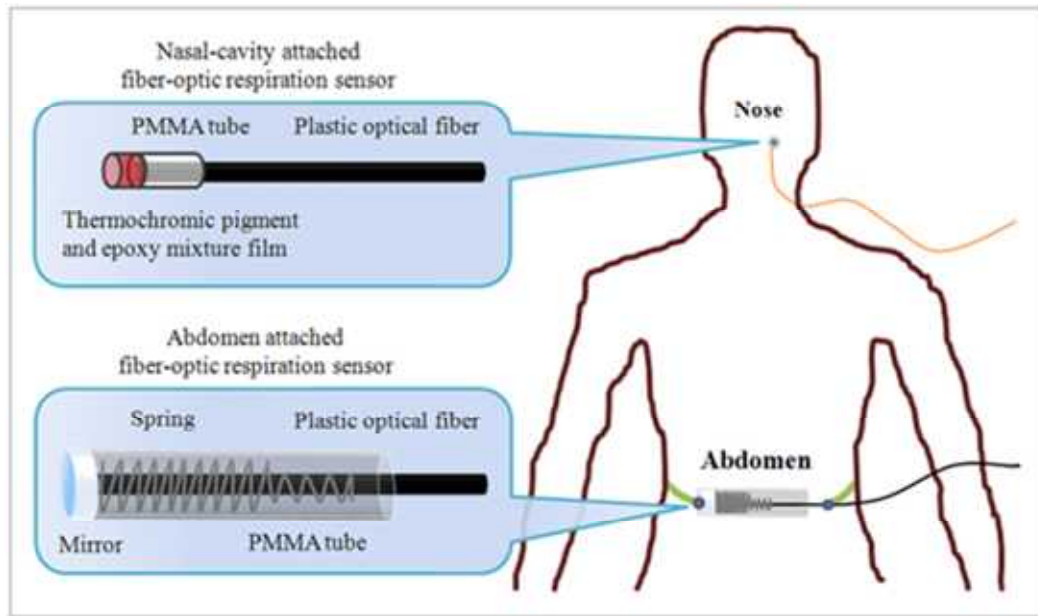


[그림 20] 3축 가속도 센서를 통해 머리의 움직임 정도를 측정(좌), 머리의 움직임을 측정하는 센서에 관한 특허의 대표도면(우)[44]

- 뇌파(electroencephalography, EEG)를 통한 방법은 뇌의 전기적 활동을 모니터링하여 운전자의 상태를 판단하는 방법임. 뇌파는 매우 신뢰성 있는 상태 정보를 포함하고 있다는 장점이 있음. 생리학적으로 인간의 수면은 기본적으로 비렘(non-rapid eye movement, NREM)수면과 빠른 눈동자 움직임을 보이는 렘(rapid eye movement, REM)수면의 두 가지로 구분할 수 있음. 비렘수면은 수면의 깊이에 따라 3가지 수면단계(N1수면,stageN1;N2수면,stageN2;N3수면,stageN3)로 나눌 수 있고 높은 단계의 더 깊은 수면상태일수록 각성 상태로의 전환을 위해서 더 강한 자극을 필요로함[45]. 즉 뇌파는 보다 객관화된 지표를 통해 세밀하고 보다 심층적인 운전자 상태를 산출할 수 있는 매우 신뢰성 있는 방법임. 하지만, 운전자의 뇌파를 측정하기 위해서는 각종 센서를 운전자의 머리에 곳곳에 부착해야하기 때문에 운전자에게 매우 불편한 방법이며, 가속도 센서와 마찬가지로 머리의 움직임, 자동차의 진동에 따라 노이즈가 발생하는 동일한 한계점을 가지고 있음
- 심박수(heart rate)를 모니터링 하는 방법은 심장의 전기적인 활동을 모니터링 ECG

방식과 심장활동에 따라 혈류량이 변화하는 것을 모니터링 PPG 방식이 있음. [14-16]과 같이 운전자가 자연스럽게 접촉하는 자동차의 시트, 안전벨트, 운전대에 ECG/PPG 센서를 탑재하여 운전자의 심장활동을 모니터링하는 연구가 있었으며, 대체적으로 높은 인식률(high accuracy)을 보이는 장점이 있음. 이미 Ford와 plessey는 ECG를 이용해 운전자의 심장활동을 모니터링하는 프로토타입 시트를 개발하였음[17, 18]. 시트에 6개 이상의 ECG 측정 센서를 탑재하고 운전자의 심장활동에 따라 발생하는 미세한 전위차를 측정하도록 고안된 시트임. 하지만 ECG는 운전자의 움직임과 더불어 착용한 옷의 두께, 계절에 따른 정전기 등의 요인으로 실제 활용에 제약이 있는 상황임. 반면, PPG는 스마트 밴드/워치와 같은 웨어러블 기기에서 주로 사용하는 방식으로 이미 성공적으로 상용화가된 기술임. PPG는 광학센서를 통해 혈류의 변화를 감지하는 방식이기 때문에 ECG 방식보다 진동, 가속 등에 덜 민감하여 운전자의 심장활동을 모니터링하기에 적합함. 심박수라는 지표는 운전자의 심장질환에 따른 위험 상황을 모니터링하기 매우 적합한 지표이면서, 동시에 운전자의 졸음에 대한 정보도 포함되어 있는 특징이 있음. [27]에 따르면, 졸음은 생리적으로 졸린 상태(부교감 신경 활성화)가 된 이후에 느끼게 되는 것으로 졸음이 밀려오기 전(운전자는 아직 졸음을 느끼기 전) 생체 신호의 변화를 통해 곧 졸음이 밀려올 것임을 예측할 수 있음. 부교감 신경이 활성화되면 심장 박동이 평소보다 낮아지고, 사람에 따라 차이는 있지만 수분에서 20분 내외로 졸음이 밀려오는 것을 느끼게 됨. 즉, 졸음이 쏟아지는 모든 운전자는 공통적으로 부교감 신경이 활성화되어 심장 박동이 평소보다 낮아지기 때문에, 심박수는 운전자의 심장질환 뿐 아니라 졸음에 대한 여부도 판단할 수 있는 매우 효과적인 방법임. 하지만, 운전자의 심장 활동은 운전자의 나이, 건강상태 등에 따라 편차가 크게 발생한다는 한계점이 있음. 대표적으로 평소 유산소 운동을 꾸준히 했을 경우 그렇지 않은 사람에 비하여 평소 심박수가 낮은 경향이 있음

- 호흡수(respiration rate)는 심박수와 유사하게 운전자의 상태를 판별하는 매우 결정적인 지표임. 정상적인 인간은 언제나 심장이 뛰고 호흡을 하며, 인간은 잠에 들 경우 평소보다 깊고 느리게 호흡하는 경향이 있음. 하지만 운전자의 호흡을 측정하는 장치는 운전자의 입과 코 주변에 부착하거나 마스크를 착용하여 측정해야하는 매우 불편한 방식임. [그림 22]와 같이 호흡에 따라 발생하는 복부의 팽창을 이용하여 호흡수를 판0단하는 방식이 있음. 입과 코를 통해 직접적으로 측정하는 방식보다 간편하지만, 주행 도중 센서가 복부 아래로 흘러내리는(sensor may slip) 등의 불편한 점은 존재할 수 있음



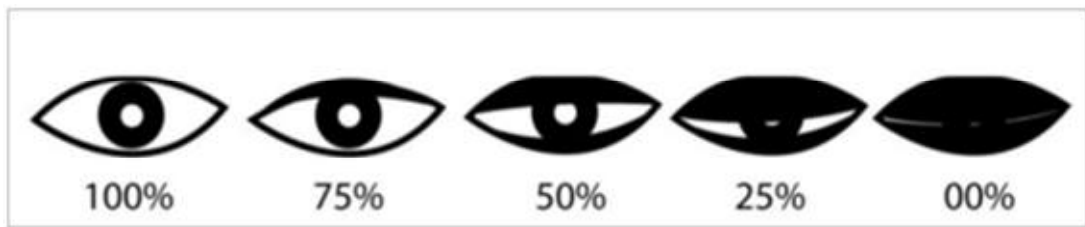
[그림 21] 호흡수를 측정하는 두 가지 방식[46]

- 전기피부반응(galvanic skin response)은 피부의 전기저항 변화를 이용하여 운전자의 상태를 판별하는 방법임. 사람은 감정에 따라 혹은 신체의 생리적 변화에 따라 피부의 전기저항에 변화가 생기는 특징이 있음. 대표적으로 졸음이 밀려오면 피부의 전기저항이 낮아지는 경향이 있음. 이를 이용하여, [47, 48]과 같이 운전자의 전기피부반응을 모니터링하여 졸음 상태를 감지하거나 집중력이 흐려지는 상태를 자동으로 판별하는 연구가 있음. 전기피부반응은 운전자의 손가락이나 팔목 등에 쉽게 부착하여 측정할 수 있다는 장점이 있음. 하지만 전기저항의 특성상 습도에 민감하게 반응하는 한계점이 있어 날씨나 계절의 변화를 고려할 필요가 있음
- 차량 내 센서를 이용하는(driving dynamics) 방법은 차선의 위치와 주행 속도 등을 이용하여 운전자의 상태를 간접적으로 판별하는 방식임. 이 방법은 운전자로부터 추가적인 데이터를 제공받지 않고, 기존 차량에 탑재되어 있는 센서 정보를 바탕으로 하기 때문에 추가적인 비용이 발생하지 않으며 쉽게 보급이 가능하다는 장점이 있음. 하지만, 자동차의 움직임을 바탕으로 간접적으로 운전자의 상태를 인식하기 때문에 일반화(generalization)가 매우 어려움 방식임. 대표적으로, 운전대의 움직임으로 운전자의 졸음을 인식하는 경우, 졸음 패턴으로 설정해둔 운전대 패턴이 특정 도로 주행환경과 일치할 경우, 혹은 몇 번 차선을 변경하는 경우에도 운전자가 졸고 있다고 판별할 수 있는 한계점이 있음

(2) 전통적인 기계학습 기반의 운전자 상태인식 연구

- 기계학습은 인간이 학습하는 방식을 모사하여 컴퓨터에 구현한 것으로 데이터와 해당 데이터와 쌍을 이루는 정답을 기계에게 부여하고, 새로운 데이터가 입력으로 왔을 때 기계 스스로 정답을 예측하는 방법임. 기계학습은 고차원적인 문제를 해결하는 중요한 접근 방식이며, 운전자의 졸음이나 심장질환과 같이 특정한 규칙으로 정의하여 일반화되기 어려운 고차원적인 문제를 해결하는데 적합한 방법임
- [표 2]는 [49]의 저자들이 기계학습을 기반으로 운전자의 졸음 상태를 인식하는 연구들의 접근 방식, 알고리즘, 성능을 정리한 자료임. [표2]에 제시된 관련연구들은 데이터의 특징(features)과 분류기(classifier)의 성능에 따라 다소 차이가 있지만 전반적으로 90% 이하의 인식 정확도를 보이고 있음. [표2]의 연구에 대하여 본 연구팀이 연구개발을 수행진행하기 위해 분석한 내용은 아래와 같음
- 우리가 일반적으로 이해하는 졸음 발생 시 나타나는 특징은 눈의 변화인데, Saradadevi et al.의 2008년 연구를 통해 입의 변화 역시 졸음을 결정하는 중요한 특징임을 확인할 수 있음
- Patel et al.의 2011년 연구는 HRV(Heart Rate Variability)를 특징으로 사용하고 뉴럴 네트워크(Neural Network)를 분류기 사용하였으며, 12명의 연구참여자로부터 졸음 발생과 관련된 ECG 데이터를 수집하고 연구팀이 직접 레이블을 달아 학습 데이터로 사용하였음. 도출한 학습 모델은 90%의 분류 정확도를 보였음
- Wang et al.의 2014년 연구는 뇌파와 얼굴 표정을 특징으로 설정하고 SVM(Support Vector Machine)을 분류기로 사용하여 88%의 분류 정확도를 보였음. Patel et al.의 연구보다 더 다양한 특징을 사용하였으며 학습데이터 수집에 참여한 연구참여자도 더 많았지만 실제 분류 성능은 2%가량 낮았음. 이러한 결과는 Wang et al.이 졸음을 분류에 사용한 특징으로 제시한 뇌파와 특정한 얼굴 표정이 Patel et al.의 뉴럴 네트워크가 찾아낸 특징보다 덜 결정적이라는 의미로 볼 수 있음. 뉴럴 네트워크는 SVM과 같은 전통적인 기계학습 방식과 다르게 알고리즘 자체에 기계 스스로 특징을 추출할 수 있는 능력이 있기 때문에, 연구자가 미처 발견하지 못한 주요 특징을 학습했을 가능성이 있음
- Lin et al.의 2015년 연구는 PERCLOS(Percentage of Eye Closure)라는 정량화된 눈이 감긴 정도를 특징으로 사용하였음. [43]의 연구와 같이 이미지 처리를 통해 눈을 검출하고 특정 지점을 랜드마크로 설정하여 [그림 21]에 해당하는 PERCLOS를 산출할

수 있음. 분류기는 이미지 처리에 사용되는 다단계 분류(cascade classification)와 여러 분류 모델의 성능에 따라 가중치를 부여하여 새로운 모델을 생성하는 에이다부스트(AdaBoost)를 사용하였음. 학습 데이터는 PERCLOS 값이 레이블되어 있는 눈 이미지를 사용하였으며, 85%의 분류 성능을 보였음. 눈이 감긴 정도는 졸음을 구분하는 결정적인 요인이며, 동시에 이 연구의 분류 알고리즘 역시 일반적으로 성능이 우수함을 고려할 때, PERCLOS라는 특징만으로는 졸음을 구분하는데 한계가 있음을 보임. 사람마다 졸음이 발생할 때 눈이 감기는 정도와 지속 시간 등의 특징을 PERCLOS 하나로 일반화하여 판별하는데 한계가 있다고 볼 수 있음



[그림 22] PERCLOS의 개념 (출처: Akrouit & Mahdi)

- G. Li and Chung의 2015년 연구는 뇌파와 머리 움직임의 특징으로 설정하고 SVM을 분류기로 사용하였음. 실측 데이터는 1분 단위의 비디오(이미지의 연속)이며 운전자의 머리 움직임에 대한 정보가 담겨있음. G. Li and Chung이 도출한 뇌파와 머리 움직임 기반의 특징 값은 SVM 분류기를 통해 96%에 달하는 인식 정확도를 보였음. 96%라는 정확도는 매우 높은 수치이며, G. Li and Chung이 도출한 특징 값이 졸음을 분류하는데 매우 결정적인 특징 값을 의미함. 하지만, 연구참여자가 타 연구개발의 절반 수준인 6명을 대상으로 하였기 때문에 결과가 과대적합(overfitting)된 결과일 가능성이 높음
- Z. Li et al.의 2017년 연구는 SWA(Steering Wheel Angle)를 특징으로 사용하고 이진 결정 분류(binary decision classification)인 비교적 단순한 기계학습 기법을 사용하였음. 다른 연구들과 다르게 전문가가 직접 레이블을 검토하였기 때문에 실측 자료의 신뢰성이 높다고 볼 수 있음. 하지만, SWA와 같이 간접적으로 운전자의 상태를 판단하는 방식은 일반화(generalization)가 매우 어려움 방식임. 해당 연구팀이 졸음 발생 특징으로 지정한 특징 값들이 실제로 특정 도로 주행환경과 일치하거나 차선을 변경하는 경우에도 나타날 가능성이 있음. 또한, 연구참여자의 수가 상대적으로 부족하다고 볼 수 있음

[표 2] 기계학습을 이용한 운전자 상태 인식 연구에서 사용한 특징과 정확도[49]

Source	Used Features	Classifier	Accuracy	Ground Truth
Saradadevi et al., 2008.	Mouth tracking	SVM	86% (Normal) 81% (Yawning)	labeled eye and mouth images (self-reported)
Patel et al., 2011.	HRV	Neural Networks	90%	labeled ECG Data (12 subjects)
Wang et al., 2014.	EEG, Facial Expressions	SVM	88.62%	SSS Labels (15 subjects)
Lin et al., 2015.	PERCLOS	Cascade Classifier, AdaBoost	85%	labeled eye pictures (self-reported)
G. Li and Chung, 2015.	EEG, Head Movement	SVM	96.24%	1-minute-videos Wierwille labels (6 subjects)
Z. Li et al., 2017.	SWA	Binary Decision Classifier	78.01%	1-minute-videos expert labels (6 subjects)

- 사람마다 졸음이 발생할 때 나오는 특징은 다양하고, 현재까지는 일반화된 수학적 모델이 없기 때문에 [표 2]의 전통적인 기계학습으로는 일반화된 인식 모델을 도출하는데 한계가 있음. 근본적으로 전통적인 기계학습은 연구팀(사람)이 지정한 특징 외 중요한 특징을 스스로 인식할 수 없기 때문에, [그림 9]와 같이 추가적인 학습 데이터가 확보되더라도 인식성능을 개선하는데 한계가 있음

(3) 운전자 졸음 상태를 감지하는 제품의 현황

- [그림 7-(위)]는 크라우드펀딩 서비스 Kick Starter에 올라온 졸음 방지 STEER 프로젝트[28], 스마트 밴드 형태의 기기를 이용하여 2초 단위로 PPG와 전기피부반응을 측정하고 졸음을 판별함. 판별 방식에 대한 구체적인 알고리즘은 공개되지 않았지만, 부교감 신경 활성화로 인한 심박수 및 피부 전도성 감소를 근거로 함. 이 기기는 졸음

이 감지되면 1차적으로 진동을 통해 알람을 주고, 진동 이후 다시 졸음이 감지되면 전기적 충격을 가해 운전자에게 각성을 부여함. 약 300여명의 후원자로부터 약 5천만원 내외의 후원을 받았으며, 2017년 11월 이후로 후원자를 대상으로 완성된 기기가 지급되고 있음이 공지되었음. 2018년 9월 해당 프로젝트의 추가적인 개발, 제품 판매, 실제 졸음 감지의 실효성 등의 내용은 추가적으로 제시 되지 않고 프로젝트가 종료된 상태임

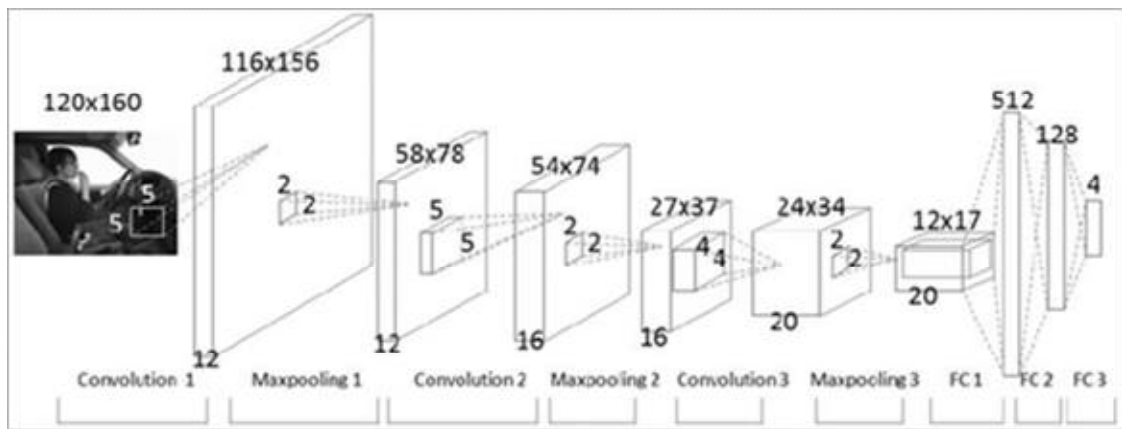
- [그림 7-(아래)]는 Leisure Auto에서 개발한 카메라 형태의 운전자 상태 인식 기기임. 기본적으로 운전자의 얼굴 이미지를 수집하고 눈과 얼굴 각도 등을 이미지 처리를 통해 산출함. 해당 기기가 기계학습 모델을 기반으로 상태를 인식하는지, 단순 이미지 처리를 통해 상태를 인식하는지 상세 기술요소에 대해서는 공개되지 않았음. 온라인을 통해 운전자가 전방을 주시하지 않는 상태일 경우를 인식하여 경보가 작동하는 시연 영상이 제공되고 있음. 인식 성능에 대한 실험 결과가 제공되지 않았으나, 시연 연상에 따르면 실제 졸음이 아닌 졸음인 척 연기를 하여도 경보가 발생하는 한계가 있음. 또한, 이미 졸음이 발생한 이후(운전자의 의식이 사라진 이후)에 경보가 울리기 때문에, 실제 졸음운전 예방의 효과는 스마트 밴드 형태의 STEER가 더욱 효과가 있을 것으로 판단됨

(4) 딥러닝 기반의 운전자 상태 인식 연구 현황(본 연구의 결과 포함)

[표 3] 국내·외 운전자 상태 인식관련 연구 및 성능비교

Related Works	Training Data	DL Algorithm	Accuracy or Results
[50] (2014)	EOG (22 subjects)	CNN	Mean Correlation Coefficient: 0.73
[51] (2014)	Face images (30 subjects)	CNN	92.33% (Normal, Drowsiness)
[52] (2017)	EOG and EEG (21 subjects)	Autoencoder	Correlation Coefficient: 0.85 Root Mean Square Error: 0.09
[53] (2017)	Eye and mouth images (22 subjects)	DNN	94.8% (Normal, Yawing, Drowsiness)
ours (2018)	Eye and mouth images with PPG data (14 subjects)	CNN + RNN	Mean: 92.80%, Best 96.15% (Normal, Light Sleep, Deep Sleep)

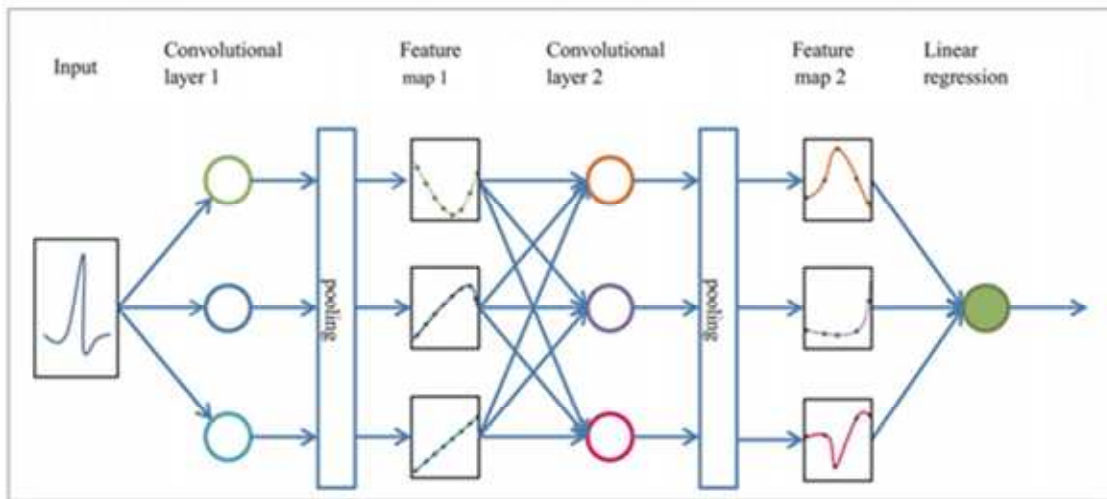
- 딥러닝 역시 기계학습과 동일하게 통계적 기법을 기반으로 데이터를 이용하여 컴퓨터 시스템에 학습 능력을 부여하는 방법임. 전통적인 기계학습과 딥러닝의 주요 차이점은 [그림 8]의 Representation Learning, Deep Learning과 같이 특징 추출(Feature extraction) 과정에 인간이 최소로 개입하거나 개입하지 않는다는 것임. 딥러닝은 특징 추출과 학습과정이 알고리즘 자체에 내장되어 있다는 측면에서 기존의 전통적인 기계학습과 큰 차이를 보임. 그로 인하여 인간이 지정하지 않은 고차원적인 특징을 기계 스스로 추출하여 보다 일반화된 모델을 도출할 수 있는 장점이 있음. 일반적으로 딥러닝은 충분한 규모의 학습 데이터가 뒷받침 된다면, [그림 9]와 같이 학습 데이터에 비례하여 성능이 개선되는 특징이 있음



[그림 23] 전형적인 CNN의 구조[54]

- [50]은 안구전도(electrooculography, EOG)를 이용하여 졸음 상태를 인식하는 연구임. 기존의 기계학습과 다르게 특정한 특징 값을 산출하여 학습하지 않고 원본 EOG에 간단한 전처리 과정 이후 CNN(Convolutional Neural Network)을 이용하여 기계 스스로 특징을 추출하고 분류하도록 하였음. 딥러닝 분야에서 이미지와 같은 공간정보를 담고 있는 데이터의 경우 [그림 22]와 같은 CNN을 이용해 주요 특징을 추출하고 있음. 분류를 위한 FC(Fully Connected) 레이어([그림 22]의 오른쪽 레이어)는 1차원(배열) 형태로 한정되며, 3차원의 컬러 사진 데이터 그대로를 1차원으로 변환할 경우 공간적 특징이 손실되어 학습이 이루어지는 않게 되는데 CNN은 이미지 데이터의 공간정보를 유지하면서 학습이 가능한 모델임. [50]의 연구팀은 22명의 연구참여자로부터 정상과 졸음 상태의 EOG를 수집하였고 [그림 23]과 같은 구조의 모델로 인식 모델을 도출하였음. 도출된 모델의 평균 상관 계수(Mean Corelation Coefficient)는 0.73로 나타났다으며 연구팀이 함께 제시한 기존의 통계적 기법이 0.7임을 감안했을 때 성능 증가

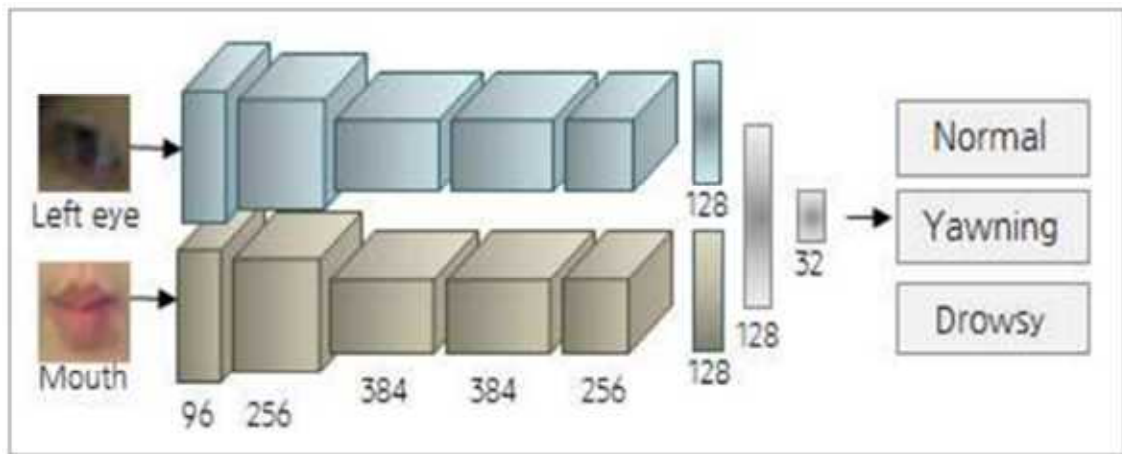
가 있었음을 보임. 다만, 결과로 도출한 상관 계수의 수치로 보면 결과로 도출한 EOG 기반의 딥러닝 모델은 좋은 성능을 보였다고 보기 어려움. 해당 연구팀이 정확도(accuracy)를 제시하지 않은 이유는 정확히 알려지지 않았으나, 2014년 기준으로 비교 가능한 딥러닝 기반의 모델이 없었기 때문으로 판단됨



[그림 24] 안구전도를 CNN을 이용하여 분류하는 딥러닝 모델의 구조[50]

- [52]는 [50]과 유사한 연구로 EOG와 EEG(뇌파)를 함께 학습하는 딥러닝 기반의 모델을 제안하였음. 다만, 오토인코더(Autoencoder)를 사용한 점이 큰 차이점임. 오토인코더는 비지도 학습(Unsupervised Learning)을 위한 뉴럴 네트워크로 CNN과 같은 지도 학습(Supervised Learning)과 다른 학습 방식을 가진다. 지도 학습은 데이터와 정답을 주고 학습한 뒤 정답이 없는 입력 데이터에서 정답을 예측하는 방식임. 따라서, 지도 학습은 새로운 입력 데이터에 대하여 정답을 예측하여 분류(classification)하는데 초점이 맞추어져 있음. 하지만, 비지도 학습은 정답이 없는 데이터에 대하여 유사성이 있는 데이터를 군집화(clustering) 하는 학습 방식임. 해당 연구팀이 결과로 제시한 상관 계수는 0.85로 졸음을 감지하는데 강력한 연관성이 있다는 것을 보였음. [50, 52]는 본 연구도메인과 비슷한 문제를 해결하기 위하여 딥러닝 기반의 접근 방법을 채택했지만, 실제로 직접적 비교가 어려운 연구임. 2.-가.-(1)에서 다루었던 운전자 상태 인식 방식의 단점에 따라, EOG와 EEG 모두 실제 주행환경에서 측정하기 어려운 데이터이며, 해당 연구팀이 제시한 실험 결과는 정확도(accuracy), ROC(Receiver Operating Characteristic) 등의 일반적인 성능비교 지표에 따른 결과가 제시되어 있지 않기 때문에 직접적인 비교는 불가능함

- [51]은 30명의 연구참여자를 모집하고 정상과 졸음 상태에 대한 얼굴 이미지 수집하였음. 기존의 기계학습 방식과 다르게 특별한 특징 추출과정 없이 간단한 전처리 과정 이후 그대로 학습 데이터로 사용하였음. 해당 연구팀은 정상과 졸음(drowsiness)을 인식하기 위한 CNN 기반의 모델을 설계하여 졸음 인식 모델을 도출하였으며, 92.33%의 인식 정확도를 실험 결과로 도출하였음. 이 결과는 기존의 전통적인 기계학습과 다르게 인간이 특징 추출 과정에 개입하지 않고 딥러닝 알고리즘 스스로 정상과 졸음을 구분하는 유의미한 특징을 추출했음을 의미함. 그리고 그 결과 역시 기존의 기계학습과 비교하여 유의미한 성능 증가가 있었음을 보임



[그림 25] 졸음 상태를 인식하기 위해 두 가지 모달리티를 사용한 딥러닝 기반의 인식 모델 구조[53]

- [53]은 기본적으로 얼굴 이미지를 학습 데이터로 하는 딥러닝 분야의 졸음 감지 연구임. 해당 연구는 본 연구개발과 유사하게 멀티모달 딥러닝 기반의 인식 모델을 도출하였음. 멀티모달 딥러닝은 학습 모델 도출에 필요한 데이터의 종류를 다양하게 함으로써, 기존의 단일 종류의 데이터만 사용한 딥러닝 방식에 비하여 성능과 안정성이 강화되는 특징이 있음. [53]의 연구팀은 4가지 모달리티(왼쪽 눈, 오른쪽 눈, 얼굴 전체, 입)를 사용하여 학습모델을 도출하였고 91.6%의 인식 정확도를 보였음. 하지만 [그림 24]와 같이 2가지 모달리티를 사용했을 때 오히려 성능이 증가하여 94.8%의 인식 정확도를 보였음. 이러한 결과가 도출되는 이유는, 1.-다.- (2-2)에서 다루었던 것과 같이, 특정 모달리티가 과도하게 학습 과정에 영향을 미치는 의존성 문제가 발생했다고 볼 수 있음. 즉, 학습 모델이 졸음을 결정하는 중요한 특징을 학습하는데 다른 모달리티가 개입되어 특징 값이 부각되지 못하고 희미하게 바뀌는 현상임. [53]의 연구는 모달리

터를 축소한 결과 오히려 3% 내외로 성능이 증가하는 경향을 보였음. 추가적으로 딥러닝은 분류해야 하는 클래스의 수가 늘어날수록 성능이 감소하는데(일반 적으로 2가지 분류보다 3가지 분류의 성능이 떨어짐), [53]의 연구는 3가지 상태를 분류하면서 95%에 달하는 정확도를 보였기 때문에 기존 연구들 보다 매우 우수한 결과로 볼 수 있음

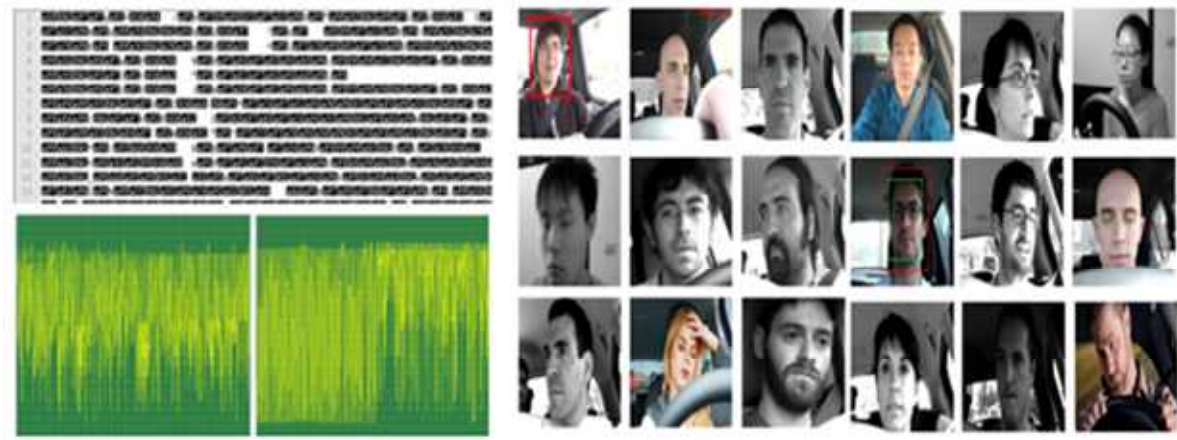
- 본 연구개발은 [54]의 연구와 매우 유사한 연구개발이며, 멀티모달 딥러닝 방식으로 정상, 얇은 잠(Light Sleep, 의식이 있는 상태), 깊은 잠(Deep Sleep, 의식이 없는 상태)의 3가지 클래스를 분류하는 인식 모델을 도출하였으며, 10-Fold Cross Validation로 측정한 평균 인식 정확도는 92.80%로 나타났으며, 가장 좋은 모델의 성능은 96.15%의 인식 정확도를 보였음(본 연구개발의 구체적인 연구 수행내용과 실험 결과는 나.절을 참조)

※ 멀티모달 딥러닝 방식의 운전자 상태 인식 모델을 도출하기 위해서는, 그에 적합한 학습 데이터가 필요함. 본 연구개발과 같이 운전자의 생리 신호와 이미지 데이터를 동시에 학습하는 모델의 경우, 두 가지 이상의 데이터가 동 시간에 수집되어야 하는 특수성이 있음. 이러한 멀티모달 학습용 데이터는 국내·외적으로 매우 희소성 있는 데이터 이며, 그로 인하여 멀티모달 딥러닝 기반의 연구개발은 단일 모달리티 기반의 연구에 비해 결과 도출에 더욱 많은 시간이 소요되고 있음

나. 연구수행내용

(1) 온라인에 공개된 심박활동 관련 데이터 및 운전자 얼굴 이미지 수집 및 분석

- PhysioNet[55], UCI Machine Learning Repository[56]의 연구목적으로 공개된 온라인 데이터 저장소에서 익명화된 심박활동 데이터를 수집하였으며, Elektra[57], BioID[58], Driver Drowsiness Detection Dataset[59], RobeSafe[60]의 연구목적으로 공개된 동영상 형태의 운전자의 얼굴 이미지를 수집하였음



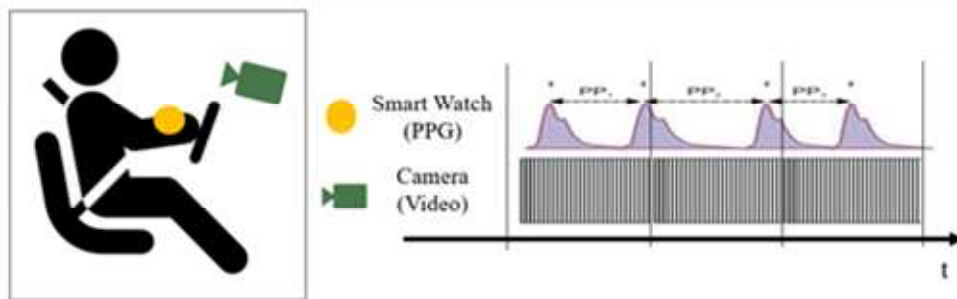
[그림 26] 연구목적으로 공개된 심장활동 데이터와 운전자 얼굴 이미지 샘플

- 수집된 심장활동 데이터는 시간에 따라 나열된 선형 값(linear value)으로 아래와 같은 규모로 수집되었음. 수집된 데이터는 나이, 성별과 같은 기본적인 메타데이터가 함께 제공되었으며, 심장질환을 가지고 있는 경우와 정상인으로 나누어 제공됨. 대표적으로 [그림 25-(좌)]와 같이 정상인과 심근경색 환자의 ECG 데이터는 육안으로 확인하여도 그 패턴에서 유의미한 차이가 있음. 정상인의 경우 일정한 파형이 나타나는 반면 심근경색 환자의 경우 간헐적으로 비규칙적인 파형이 나타남
 - 정상인 853명의 ECG/PPG (1,382 recordings)
 - 심장질환을 가지고 있는 환자 809명의 ECG/PPG (1,534 recordings)
 - 각각의 recording은 20~250KB(크기가 클수록 기록된 샘플이 더 많음)
- 수집된 운전자 얼굴 이미지는 비디오 파일로 아래와 같은 규모로 수집되었음. [그림 5-(우)]와 같이 동영상 형태로 존재하는 운전자 얼굴 데이터를 수집하였으며, 비올라

존스(Viola- Jones) 등의 알고리즘을 이용해 이미지에서 수집된 운전자 얼굴 이미지에서 정상적으로 얼굴 추출이 가능한지 이미지의 품질 수준을 확인하였음(본 연구개발은 실제 사고와 관련이 있는 데이터는 수집대상에 포함하지 않았음)

- 깨어 있는 상태의 운전자 16명의 비디오 파일 (30fps, 640x480 해상도, 총 73GB 규모, 실제 주행환경)
- 졸고 있는 상태의 운전자 2명의 비디오 파일 (30fps, 640x480 해상도, 총 4GB 규모, 실험실 환경)
- 본 연구개발의 핵심 접근방법인 멀티모달 딥러닝은 두 가지 이상의 구분되는 모달리티의 데이터가 동일한 시점에 생성된 학습 데이터를 필요로함. [그림 26]의 개념도와 같이 t라는 시점에 운전자 얼굴 이미지와 PPG 데이터가 쌍으로 수집된 학습 데이터가 필요함. 본 연구팀이 연구개발을 진행하며 온라인에서 수집한 관련 데이터는 모두 한 가지 모달리티로 구성된 데이터였으며, 두 가지 이상의 모달리티를 충족하는 공개된 데이터 세트는 찾을 수 없었음

(2) 가상 주행 환경 구축 및 멀티모달 딥러닝을 위한 학습 데이터 수집



[그림 27] 멀티모달 딥러닝을 위한 학습 데이터 수집 개념도

- 주행 중 발생하는 운전자의 심장 활동과 동 시점의 얼굴 이미지 데이터를 수집하기 위하여 실제 주행환경과 유사한 가상의 주행 환경을 [표 4]와 같이 구축하였음
- 심박 데이터 수집: 우리나라 의료법에서는 ECG 측정 기기의 평가 기준을 고전압에 견디는 병원용 심전도 기기로 제한하고 있어, 스마트워치를 통한 ECG 측정은 법적으로 허가된 영역이 아님. 대표적으로, Apple의 스마트워치는 시계의 용두에 손가락을 올려 심장을 가로지르는 전류를 통해 심장의 전기 신호를 추적가능하나 우리나라에서는 출시하지 못했음. 반대로 PPG의 경우는 측정 장치에 법적인 문제가 없으나. 우리나라 의료법상 스마트 디바이스를 통한 원격 진료 및 모니터링은 금지되어 있음.

PPG는 일종의 개인의 의료정보이기 때문에 스마트 디바이스를 통한 데이터 원격 전송은 법적으로 허가된 영역이 아님. 다만, 연구참여자의 동의가 있을 경우 연구를 목적으로 타인에게 제공될 수 있으며, 제공받은 연구팀은 개인정보보호법에 따라 연구참여자를 식별할 수 없도록 익명화해야하며 규정에 따라 관리 및 폐기되어야함. 이러한 조건에서, PPG 데이터를 수집할 수 있는 장치는 삼성전자의 기어 S3가 유일한 장치였음. 타사의 스마트 워치/밴드도 PPG 측정이 가능하지만 원천 데이터의 접근을 기기 자체에서 차단하기 때문에 연구팀이 사용할 수 없었으나, 삼성전자의 기어 S3는 개발자 인증서를 발급받아 PPG를 수집하는 프로그램을 제작이 가능했음

- 얼굴 데이터 수집: 카메라의 경우 많은 선택지가 있었으나 연구 목적으로 수집하는 데이터를 고려했을 때 높은 해상도를 가진 데이터를 확보하는 것이 유리하다고 판단했음. 낮은 해상도는 높은 해상도로 변환이 불가능하지만, 반대로 높은 해상도는 필요에 따라 낮은 해상도로 변경이 가능하기 때문임. 또한, 실제 주행환경에서 데이터를 수집할 경우를 대비하여 진동과 같은 흔들림에 덜 민감하면서 빛이 희미한 어두운 환경에서도 촬영이 가능한 카메라 장치가 데이터 수집에 유리하다고 판단하였음
- 가상 주행 환경: Euro Truck Simulator 2는 실제 트럭 운전기사처럼 각종 회사로부터 배송에 관한 의뢰를 받고 화물을 운송하는 시뮬레이터임. 화물 배송 경로는 실제 유럽 전역의 도로를 모사하여 개발된 시뮬레이터이며, 일반적인 레이싱 시뮬레이터와 다르게 속도감이 아닌 높은 현실성을 고려하여 개발되었으며, 실시간으로 기상 상황이 변하여 야간이 되면 어두워지고 비가 내리면 와이퍼를 작동하여 시야를 확보하는 등의 동작이 요구됨. [61, 62]와 같이 운전자에 대한 연구에 가상 주행환경으로 자주 사용되는 시뮬레이터임. 이 시뮬레이터는 PC에서 실행가능하며 키보드나 마우스로도 조작이 가능하나, 연구참여자의 실제 몰입감과 현실감을 제공하기 위하여 시뮬레이터와 연동되는 운전대와 패달을 함께 사용하였음

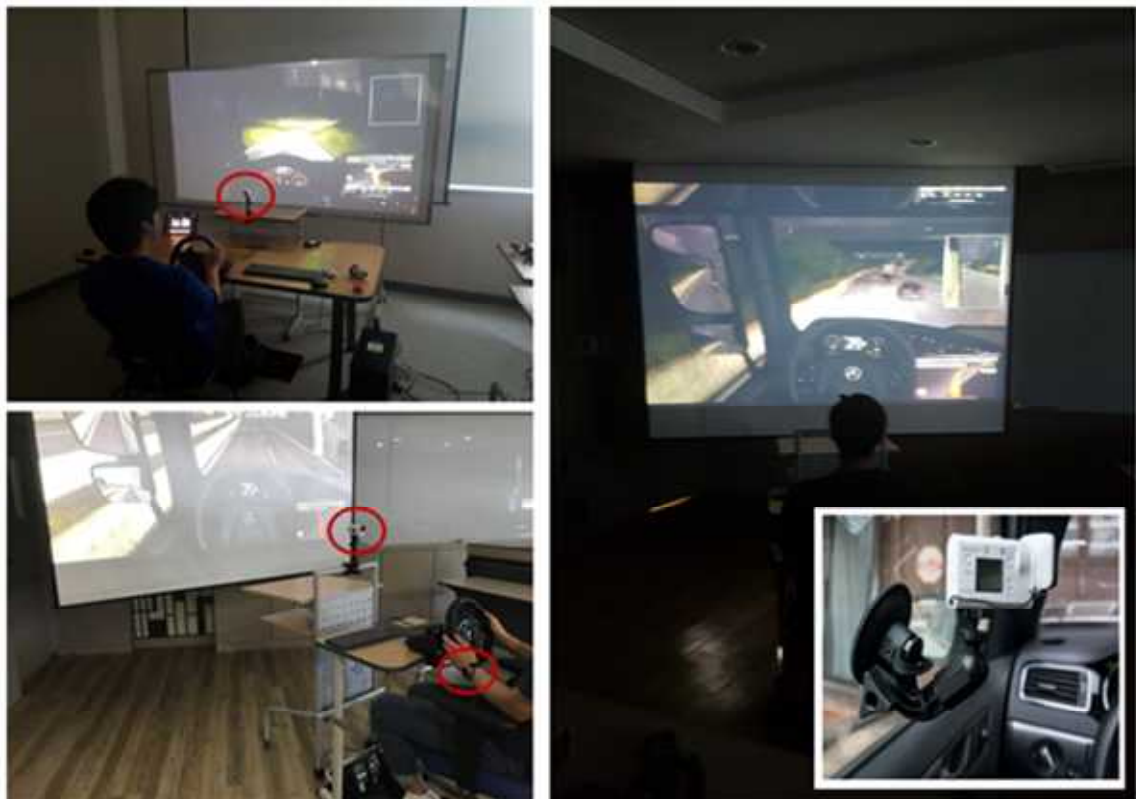
[표 4] 가상 주행환경 구축에 사용한 장치와 방법

심박 데이터 수집 [그림 28] 참조	장치: 삼성전자 기어 S3 스마트 워치
	방법: TIZEN 기반 PPG 수집 프로그램 자체 개발
얼굴 데이터 수집 [그림 29] 참조	장치: Sony FDR-X3000 카메라
	방법: 연구참여자 전방에 카메라를 설치하여 촬영
가상 주행 환경 [그림 29] 참조	장치: 로지텍 G29 Driving Steering Wheels & Pedals
	시뮬레이터: EURO Truck Simulator 2



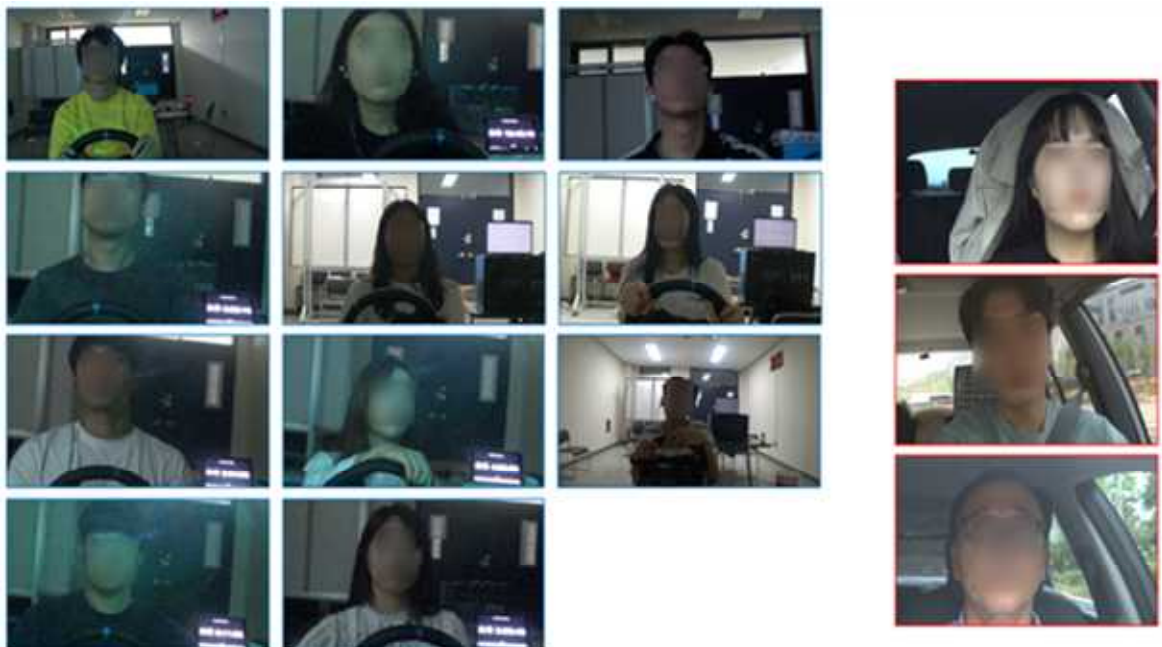
#	Time	Label	PPG
704	5_25_17_19_59_439	FALSE	92
705	5_25_17_19_59_507	FALSE	92
706	5_25_17_19_59_574	FALSE	92
707	5_25_17_19_59_624	FALSE	92
708	5_25_17_19_59_681	FALSE	92
709	5_25_17_19_59_756	TRUE	92
710	5_25_17_19_59_789	TRUE	92
711	5_25_17_19_59_873	TRUE	92
712	5_25_17_19_59_937	TRUE	92
713	5_25_17_19_59_992	TRUE	92
714	5_25_17_20_00_060	TRUE	92

[그림 28] 본 연구팀이 자체 개발한 TIZNE 기반의 PPG 데이터 수집 프로그램(좌), PPG 데이터 수집 프로그램을 통해 수집된 원천 데이터 형태(우)



[그림 29] 본 연구팀이 구축한 가상 주행환경
(연구참여자는 빨간색 표시와 같이 PPG 데이터 수집 프로그램이 탑재된 스마트워치를 착용하고 전방에 설치된 카메라를 통해 동시점에 얼굴 이미지가 수집됨.
동시에 촬영되는 후방 카메라는 수집된 데이터의 레이블링과 검증을 위한 용도로 사용함)

- 가상 주행환경은 [그림 27]과 같은 자체 개발한 PPG 데이터 수집 장치를 통해 연구참여자의 심장 활동 정보를 수집하면서, 동일한 시점에 [그림 29]와 같이 전방 카메라를 통해 운전자의 얼굴 데이터를 수집함
- [그림 27-(좌)]는 본 연구팀이 개발한 TIZEN 기반의 데이터 수집 프로그램으로 연구참여자가 직접 “데이터 수집 시작”을 터치 하여 PPG 데이터를 수집을 시작하며, 만약 연구 참여자가 졸음을 느끼기 시작하면(의식이 있는 상태, Light Sleep) “졸음 라벨 시작”을 터치하여 연구 참여자가 스스로 얹은 졸음(Light Sleep)의 시점을 지정함. [그림 27-(우)]와 같이 데이터는 시간에 따라 정렬되어 기록되며, 레이블(label)은 얹은 잠(Light Sleep) 여부를 나타내고(TRUE/FALSE) 해당 시점의 PPG 값이 차례대로 기록됨. 단, 깊은 잠(Deep Sleep)의 경우는 연구참여자가 의식이 없는 상태이기 때문에 본 데이터 수집 프로그램을 통해 레이블이 기록되지 않고 연구팀이 직접 카메라에 기록된 내용을 바탕으로 시점을 지정함
- 구축한 가상 주행 환경은 [그림 28-(좌)]와 같이 낮 시간대를 가정하여 조명을 밝게 하였으며, [그림 28-(우)]와 같이 저녁 시간대를 가정하여 조명을 어둡게 조정하였음. 또한, 얼굴 데이터 수집 장치와 동일한 카메라를 후방에 설치하여, 연구 참여자가 해당 시간에 졸았는지의 여부를 연구팀이 다시 한 번 확인하는 등의 검증용 데이터로 사용하였음



[그림 30] 가상 주행환경의 연구참여자 11명(좌), 실제 주행환경의 연구참여자 3명(우)

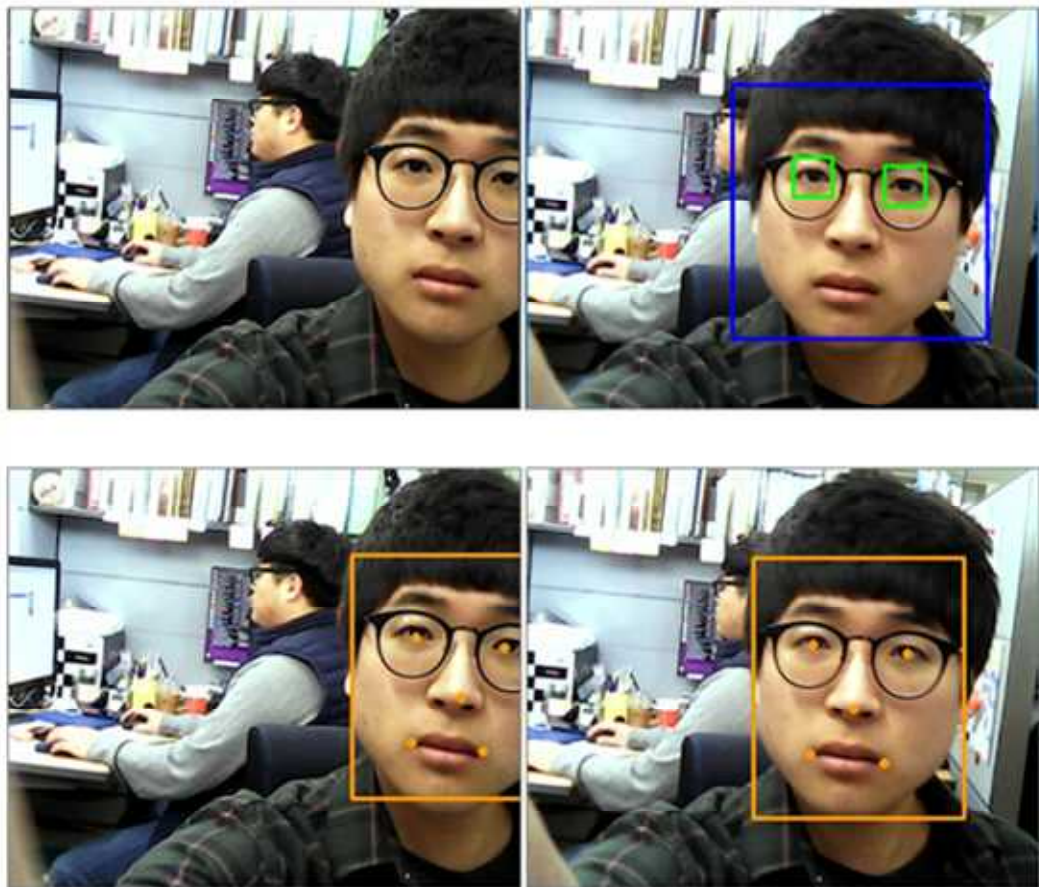
- 본 연구팀은 [그림 29-(좌)]와 같이 운전면허를 소지한 11명의 성인 연구 참여자를 모집(남자 6명, 여자 5명) 하였으며(경찰청의 2016년 통계에 따르면 남:녀 운전면허 소지 비율은 6:4 수준임), 각 연구참여자들은 평균 7시간 분량의 가상 주행환경 데이터를 제공하였음. 본 연구팀이 수집한 가상 주행환경의 원본 데이터와 레이블 작업까지 완료된 학습용 데이터는 [표 5]와 같음

※ 본 연구팀은 모든 연구참여자에게 본 연구에 대한 충분한 설명(문)을 제공하였으며 수집된 데이터는 연구참여자의 개인정보가 포함되었기 때문에 관련 법규에 따라 개인정보 활동동의서를 받아 규정을 준수하여 수집된 데이터를 관리하였음

[표 5] 본 연구팀이 수집한 가상 주행 환경의 원본 및 학습용 데이터

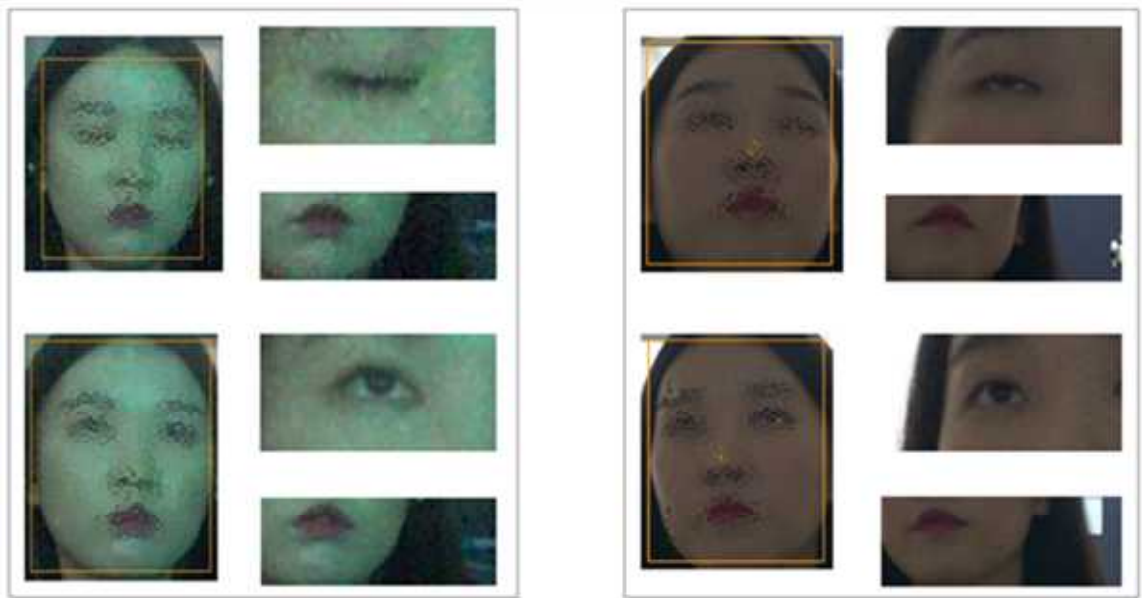
연구참여자	11명 (남성 6명, 여성 5명)
수집된 원본 심박 데이터	2,808,000개의 PPG 데이터, 약 40GB
수집된 원본 얼굴 데이터	약 8,500,000장의 얼굴 이미지 (30fps의 1920x1080 해상도 동영상 78시간 분량)
연구팀이 추출한 학습용 데이터	• 단위: 90장의 연속된 눈, 입 이미지와 동 시간에 수집된 PPG 데이터 30개 • 규모: 32,540개의 멀티모달 학습 데이터 • 분포: 정상 55%, 얇은 잠 42%, 깊은 잠 3%

- 수집된 데이터에는 [그림 27]의 스마트 워치를 통해 연구 참여자가 직접 피곤한 상태임을 지정하기 때문에 지정된 시점은 연구팀의 특별한 확인 없이 얇은 잠(Light Sleep, 의식은 있으며 졸음을 느끼는 상태)으로 확정하여 레이블링(Labeling) 하였음. 하지만, 연구 참여자가 깜박 졸게 되면 언제부터 깊은 잠(Deep Sleep, 의식이 없음)이 되었는지 연구 참여자 스스로 지정할 수 없음. 따라서 [그림 28]의 후방 카메라를 통해 차선을 이탈하거나 주행 제어권이 상실된 시점을 찾거나, 전방 카메라에 눈이 감긴 기록을 토대로 본 연구팀이 직접 깊은 잠에 대한 레이블링 확정하였음. 이때, 해당 연구참여자가 직접 참관하고 질의응답을 통해 검증하는 작업이 병행되었음



[그림 31] 바이올라 존스 기반 얼굴 인식(위), MTCNN 기반 얼굴 인식(아래)
(사진 속 인물: 본 연구팀의 최형탁 석사과정)

- 레이블 작업이 완료된 이후, 수집된 원본 얼굴 이미지는 눈과 입으로 분리함으로써 학습 데이터의 크기를 줄이면서 연구 참여자의 개인정보를 익명화할 필요가 있음. 본 연구팀이 수집한 가상 주행환경에서의 원본 얼굴 이미지는 약 8,500,000장으로 몇 명의 연구원이 수작업으로 수행하기에는 매우 부담스러운 규모임. 얼굴을 검출하고 눈과 입을 자동으로 찾을 수 있는 대표적인 알고리즘으로 비올라 존스 알고리즘이 있지만 [그림 30-(위)]와 같이 이미지 내 얼굴이 온전히 위치하지 않을 경우 검출이 불가능한 한계가 있음. 따라서, [그림 30-(아래)]과 같이 보다 유연하게 얼굴을 검출하고 눈과 입을 자동으로 분리할 수 있는 MTCNN[63]을 사용하였음. MTCNN을 사용하면 이미지 내 얼굴을 검출하고, 검출된 이미지 내 눈, 코, 입을 인식하여 랜드마크를 산출함. [그림 31]는 연구참여자의 원본 얼굴 이미지에서 눈과 입 이미지를 분리한 결과를 담고 있음(노란색 점으로 산출된 랜드마크를 중심으로 눈과 입 이미지 추출)



[그림 32] MTCNN 기반의 자동화된 눈, 입 이미지 추출 결과
(왼쪽은 주변 조명이 어두운 이미지, 오른쪽은 주변 조명이 밝은 이미지)

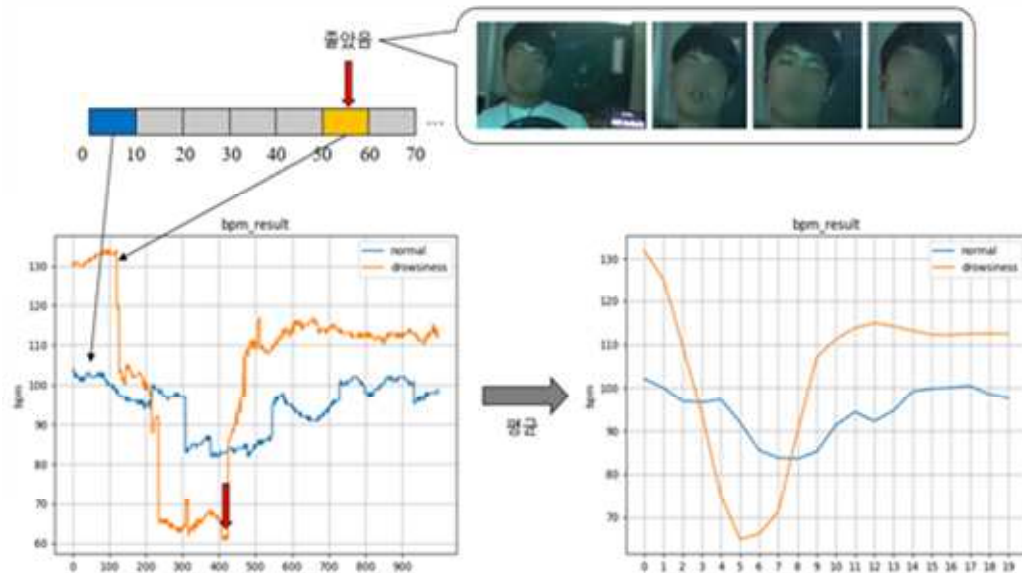
※ 가상 주행환경에서 수집된 원본 얼굴 이미지는 약 8,500,000장이나 추출된 학습용 데이터의 얼굴 이미지는 2,928,600장(32,540 * 90장)으로 줄어든 이유는 [그림 32]와 같이 누락된 데이터가 발생하였기 때문임. 누락의 주된 이유는 스마트 워치의 흔들림이나 데이터 수집 프로그램의 사용 미숙이었음. 눈 또는 입 이미지가 누락되는 경우는 MTCNN을 통한 얼굴 인식이 불가능한 이미지였음



[그림 33] 특정 모달리티 데이터의 누락과 노이즈를 고려한 학습 데이터 추출

- 사람은 부교감 신경이 활성화된 후 뒤 졸음을 느낄 수 있으며, 부교감 신경이 활성화 되면 나타나는 대표적인 생리적 변화는 심박수가 평소 보다 감소하는 것임. 이는 [그림 33]과 같이 수집된 PPG 데이터를 BPM(Beats per Minute) 단위로 변환하여 관찰

한 결과 졸음 발생 시점을 전후로 심박수가 하락하는 것을 확인하였음. 이와 같은 방식으로 본 연구팀이 수집한 PPG 데이터에 대한 검증을 수행하였음



[그림 34] 졸음 발생을 전후로 감소하는 심박수 확인 과정

(3) 실제 주행 환경에서의 멀티모달 딥러닝을 위한 학습 데이터 수집

[표 6] 본 연구팀이 수집한 실제 주행 환경의 원본 및 테스트용 데이터

연구참여자	3명 (남성 2명, 여성 1명)
수집된 원본 심박 데이터	396,000개의 PPG 데이터, 약 6GB
수집된 원본 얼굴 데이터	약 1,180,000장의 얼굴 이미지 (30fps의 1920x1080 해상도 동영상 11시간 분량)
연구팀이 추출한 테스트용 데이터	- 단위: 90장의 연속된 눈, 입 이미지와 동 시간에 수집된 PPG 데이터 30개 - 규모: 3,254개의 멀티모달 학습 데이터 (학습 데이터의 10% 규모) - 분포: 정상 57%, 얇은 잠 43%, 깊은 잠 0%

- 본 연구팀은 [그림 29-(우)]와 같이 3명의 연구 참여자를 모집(남자 2명, 여자 1명)하여 실제 장거리 주행 환경에서의 [표 6]과 같이 운전자의 심장 활동 및 얼굴 이미지를

수집하였음

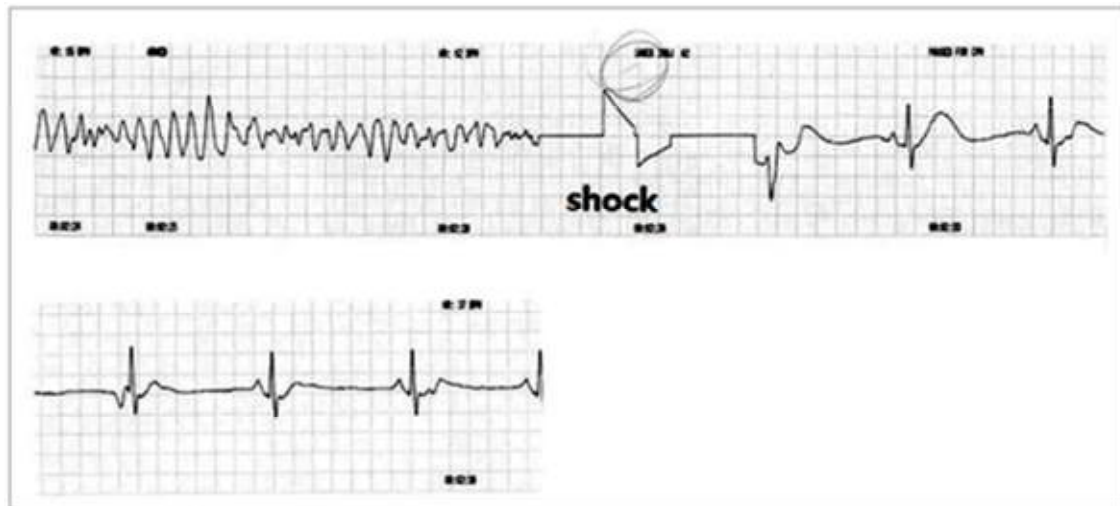
- 가상 주행환경과 다르게 실제 주행환경에서 졸음발생이 실제 사고로 이어지지 않도록 동승자가 탑승하였음. 그에 따라, 본 연구팀이 모집한 연구 참여자들은 장거리 주행 중 일부 구간에서 졸음이 밀려오는 것을 느꼈으나 실제 의식이 사라지는 깊은 잠은 발생하지 않았음
- [표 6]과 같이 수집된 실제 주행 환경의 테스트용 데이터는, 학습용 데이터를 통해 도출한 인식 모델의 성능을 가늠하고 실제 적용 가능성을 확인하기 위한 용도로 사용되었음(단, 실제 주행환경에서 깊은 잠에 대한 데이터를 확보하지 못했기 때문에 그에 대한 결과는 학습 과정에 사용되지 않은 학습용 데이터를 사용한 결과로 대체함)

※ 실제 주행환경에서 수집된 원본 얼굴 이미지는 약 1,180,000장이나 추출된 학습용 데이터의 얼굴 이미지는 292,860장(3,2540 * 90장)으로 줄어든 이유는 [그림 32]와 같이 누락된 데이터가 발생하였기 때문임. 누락의 주된 이유는 스마트 위치의 흘러내림이나 데이터 수집 프로그램의 사용 미숙이었음. 눈 또는 입 이미지가 누락되는 경우는 MTCNN을 통한 얼굴 인식이 불가능한 이미지였음

(4) 시그널 처리 및 딥러닝을 통한 운전자의 심장질환 상태 감지

- 심장질환은 의학적으로 전기적 문제로 발생하는 심장정지와 혈액순환 문제로 발생하는 심장마비로 구분됨. 심장정지의 경우 심장이 정지하기 직전에 [그림 34-(위)]과 같은 심실세동이 나타나는 특징이 있음. 즉, 심전도(ECG)에 전혀 변화가 없는 것이 아니고, 심장의 수축 이완이 없어 보이나 심장 표면에 가늘게 떨리는 현상이 발생하여 미약한 심전도 기록이 발생하며[37], 그 파형은 정상적인 파형에 비해 진폭(amplitude)과 파형의 높이가 매우 짧아지는 특징이 나타남
- PPG는 스마트 밴드/워치와 같은 웨어러블 기기에서 주로 사용하는 방식으로, 광학센서를 통해 혈류의 변화를 감지하는 방식이기 때문에 ECG 방식보다 진동, 가속 등에 덜 민감하여 운전자의 심장활동을 모니터링하기에 적합함. PPG는 ECG 보다 파형이 정밀하지 못하기 때문에 협심증, 심근경색, 동맥경화증 등의 구체적인 심혈관질환을 판단하는 근거로 사용될 수 없음
- PPG는 ECG에 비하여 주행환경의 진동, 가속 등의 움직임에 덜 민감하여 노이즈와 누락이 상대적으로 적게 나타남. 하지만, PPG 역시 광학센서가 피부에 접촉된 영역, 운

전자의 습관 등에 따라 노이즈와 누락 값이 충분히 발생할 수 있음. 이러한 문제를 해결하기 위하여 시그널 처리와 전통적인 기계학습 분야의 다양한 연구[20-22]가 있음



[그림 35] 심장정지 직전의 심전도 기록과 전기충격요법을 통한 심전도 변화(위), 전기충격요법 이후 정상리듬을 회복한 심전도 기록(아래)[37]

- 딥러닝은 특징 추출과 학습과정이 알고리즘 자체에 내장되어 있다는 측면에서 기존의 전통적인 기계학습과 큰 차이를 보임. 즉, 학습 과정에 인간의 개입을 최소화 하였으며, 그로 인해 인간이 파악하기 어려운 고차원적인 특징을 기계 스스로 추출하여 보다 일반화된 학습 모델을 도출할 수 있는 장점이 있음
- 하지만 딥러닝을 적용하기 위해서는 충분한 규모의 학습 데이터가 필요함. 본 연구개발은 운전자의 심장질환을 판단하기 위한 근거로 2.-나-(1)과 같이 공개 저장소에 존재하는 각종 심장활동 관련 데이터를 수집하였음. 하지만, 심장마비 혹은 심장정지와 관련된 데이터는 극소량만 입수할 수 있었음. 심장마비, 심장정지와 같이 환자의 생명이 위독한 상황의 데이터를 입수하는 것은 연구윤리 차원의 문제도 함께 고려되어야 하기 때문에, 본 연구개발에서는 수집한 극소량의 데이터를 증강(augmentation) 하여 딥러닝이 가능한 수준으로 데이터를 확보하는 방향으로 연구개발을 진행하였음
- 딥러닝 모델은 수많은 파라미터가 존재하며, 이 파라미터들은 반드시 많은 양의 트레이닝 데이터를 통하여 효율적으로 학습하여 일반화(generalization) 성능을 향상시켜야 함. 현실적으로 모델을 학습시키기 위한 데이터 셋은 한정되어있고 불균형의 문제를 가지고 있음. 이러한 데이터 문제를 해결하기 위하여 데이터 증강이 필요함
- 학습 데이터의 불균형의 문제도 딥러닝 모델에서 나타나는데, 데이터 불균형은 어떤

데이터 클래스에 거의 데이터가 없어서 발생하는 문제임. 이는 현실적인 딥러닝 시나리오에서 주로 발생함(예: 금융거래, 이체 업무에서의 이상 탐지 그리고 의료연구 등에서 발생). 본 연구개발의 도메인과 같이, 의학에서는 건강한 인구에 대한 데이터만 많고 어떤 병에 걸린 사람들의 데이터가 극히 적어 불균형이 발생해 딥러닝 모델에 바이어스(bias) 현상이 많이 발생함

- 이와 같이 일반화 및 데이터 불균형에 의한 바이어스 현상의 두 가지 문제점은 데이터 증강을 통하여서 해결이 가능하며 데이터 증강은 원본 데이터의 레이블을 그대로 유지하되 원본데이터와 다른 새로운 데이터를 생성하는 것임. 데이터 증강은 데이터 셋의 정보를 늘리는 것이 아니고 데이터 셋의 다양성을 부여하고 모델의 일반화 성능을 향상함
- [24]는 웨어러블 센서 데이터를 증강하여 질병을 모니터링하는 연구이며 추가적인 데이터 수집 절차 없이 약 9% 정도의 성능 개선을 보였음. 본 연구개발의 도메인과 유사한 접근의 연구는 [38]의 연구로, 분할된 PPG 신호에 대하여 웨이브릿(Wavelet) 변환을 가한 이미지 형태의 데이터를 CNN(Convolutional Neural Network)을 이용해 분류하는 연구로 본 연구와 유사한 형태의 연구를 진행하였음. 이 연구의 예측 모델은 기계학습 알고리즘인 KNN(K-Nearest Neighbor Algorithm)보다 약 6% 내외의 예측 성능 개선을 보였음
- 본 연구개발은 2-나-(1)과 같이 수집된 심장활동 관련 데이터서 심장정지와 관련된 독립된 3개의 PPG 데이터를 추출하였음. ECG 심실세동에 따른 미세한 전위차가 발생 하지만, PPG는 ECG와 다르게 심장활동에 따라 혈류의 흐름의 주기성을 모니터링 하는 방식이기 때문에 심장정지의 경우 거의 0에 가까운 값을 보였음
- 심장마비의 방치할 경우 심장마비로 이어지지만 심장정지와 다르게 심장의 전기적인 신호는 살아있음. 다만 혈관이 막혀 평소와 다른 혈류량을 나타냄. 이러한 특징은 심장활동이 불규칙하게 발생하는 부정맥 환자에게서 나타나는 패턴으로 [55]를 통해서 수집된 데이터를 활용하였음

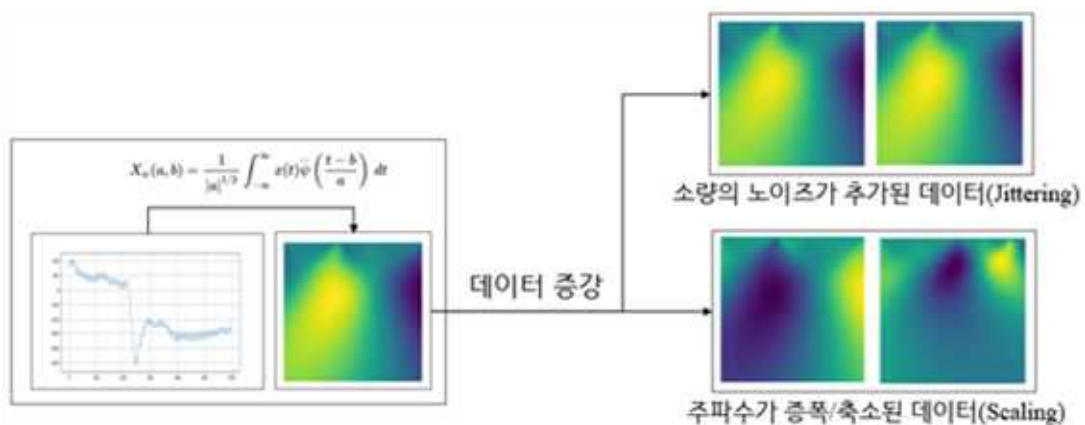


[그림 36] 본 연구개발의 심장질환 인식 모델 도출 방법

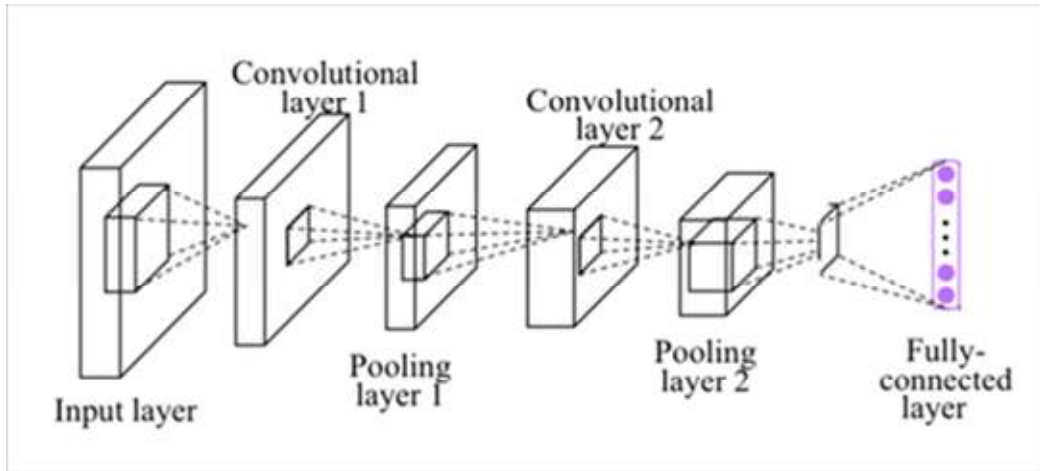
- 본 연구개발은 [그림 35]와 같은 순서에 따라 운전자의 심장활동을 모니터링하고 상태

를 인식하는 인식 모델을 개발하였음

- PPG 데이터 정규화: 수집된 PPG 데이터는 본 연구개발의 실제/가상 주행환경의 연구 참여자로부터 수집된 정상데이터를 사용하였으며, 온라인에서 수집된 각종 심혈관 환자의 데이터에서 본 연구개발에 필요한 심장정지, 심장마비와 관련된 데이터를 추출하였음. 수집된 PPG 데이터는 0~2,675,000의 분포를 보였으며 수집된 데이터가 -600~600 사이의 값으로 변경되도록 정규화를 진행하였음. 정규화가 완료된 PPG 데이터는 600개 단위로 분할하여 저장하였음(sampling rate: 10 * 60초)
- CWT(Continuous Wavelet Transform) 추출: 정규화가 완료된 PPG 데이터는 [그림 36]의 수식에 따라 2차원 이미지로 변환하였음. CWT의 수식에는 기존 1차원 값의 나열인 PPG 데이터에 RGB 이미지 값으로 변환이 되는 식이 포함되어 있어 색을 가진 이미지가 나타남(변환된 이미지의 해상도는 280x280)
- 데이터 증강: Jittering, Scaling의 두 가지 데이터 증강기법을 사용하여 데이터를 증강하였음. 이는 [24]의 연구와 같이 센서 데이터를 증강하는 방식으로 많이 사용되는 방법임. Jittering은 데이터에 일정 노이즈를 추가하여 증강하는 기법으로, 본 연구에서는 시그마(sigma, 파라미터 중 하나) 값을 0.001 ~ 1 범위에서 0.001씩 증가시켜 증강하였음. Scaling은 데이터에 랜덤한 값을 곱하여 증강하는 기법으로, 원본 데이터의 전반적인 주파수가 줄어들거나 늘어나는 효과가 발생함. 마찬가지로 Scaling의 시그마 값을 0.001 ~ 1 범위에서 0.001씩 증가시켜 데이터를 증강하였음
- 딥러닝 인식모델: [그림 37]과 같은 구조의 CNN을 사용



[그림 37] CWT 추출을 통한 1차원 PPG 데이터를 2차원 이미지로 변환(좌)
노이즈와 주파수 증폭/축소를 통한 데이터 증강(우)



[그림 38] 본 연구개발에서 증강된 PPG 데이터의 분류를 위해 사용한 CNN 모델의 구조 (2개의 컨볼루션, 2개의 풀링)

- 본 연구개발을 통해 도출된 심장질환 인식 모델은 [표 7]과 같은 성능을 보였으며, 충분히 실제 적용이 가능함을 테스트 데이터를 이용해 검증하였음. 운전자의 심장활동은 나이, 성별, 건강상태에 따라 편차가 큰 데이터임. 따라서, 단일한 모델로 다수의 운전자의 심장활동을 인식하는 것은 일반화가 어려운 문제임. 주행환경에서 운전자의 심장질환을 보다 정확하고 실제적용이 가능한 수준으로 감지하기 위해서는, 데이터 증강에 기반하여 점진적으로 인식 성능이 개선되는 개인화가 가능한 기술이 적합함. 초기 운전자가 확정된 시점에 적은 규모의 데이터를 증강하여 심장질환을 인식하는 모델을 생성하고, 자동차를 사용하는 시간에 비례하여 점진적으로 성능이 강화되는 방식이 적합함.

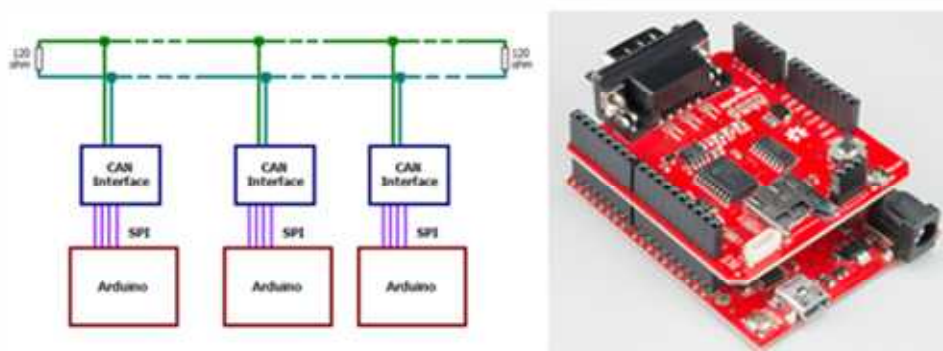
[표 7] CNN 기반의 심장질환 인식 모델의 성능

분류 클래스	정상 / 심장정지(증강된 데이터) / 심장마비(증강된 데이터)
학습 데이터	32,000 (정상 80%, 심장정지 10%, 심장마비 10%)
테스트 데이터	2,400 개 (학습에 사용하지 않은 다른 피험자 데이터, 클래스 분포는 학습 데이터와 동일)
인식 정확도	94.87% / 92.70% (10-Fold Cross Validation / 테스트 데이터)

(5) 멀티모달 인터페이스 및 운전자 상태인식 시스템의 실시간 및 신뢰성 높은 데이터 전송 기법 적용

※ 본 연구개발은 1차년도(2017.06.30. ~ 2017.12.31.)의 연구개발을 통하여 아래와 같이 초기 단계의 멀티모달 인터페이스 프로토타입을 개발하였으며, 2차년도 연구개발을 통하여 상태인식 시스템과 연계되고 DDS를 통한 실시간 및 신뢰성 높은 데이터 전송 기법이 적용되었음

- CAN(Controller Area Network)은 ISO-11898에 정의되는 자동차 전자장치의 실시간 통신 프로토콜로 전자장치의 개발~검증까지 전 범위에 사용되는 표준임. 본 연구개발은 [그림 38]과 같이 CAN을 기반으로 PPG와 같은 운전자의 생체 데이터를 수집할 수 있는 오픈하드웨어 기반의 인터페이스 프로토타입을 제작하였음
- 반면, CAN의 데이터 전송률은 1Mbps로 운전자 이미지와 같이 고해상도 이미지를 상태 인식 시스템으로 전송하기에는 매우 부족한 전송률임. 인포테인먼트에 탑재되는 카메라는 USB(Universal Serial Bus)를 준수하기 때문에, 본 연구개발도 USB를 기반으로 운전자의 얼굴 이미지를 전달 받는 인터페이스를 추가하였음(크로스 플랫폼 라이브러리인 libusb를 활용하여 USB 디바이스를 접근하고, 데이터를 수집)
- CAN은 무선통신이 아니기 때문에, 운전자의 PPG를 수집하기 위해서는 유선으로 연결된 PPG 측정 모듈이 필수적임. 최근에는 측정된 PPG를 블루투스(Bluetooth)와 같은 무선 통신 기술로 전송되고 있음(우리나라의 경우 의료용으로 사용되는 PPG 측정 기기는 USB 또는 블루투스를 사용. 대표적으로 LAXTHA의 ubpulse 및 [그림 6] 참조)



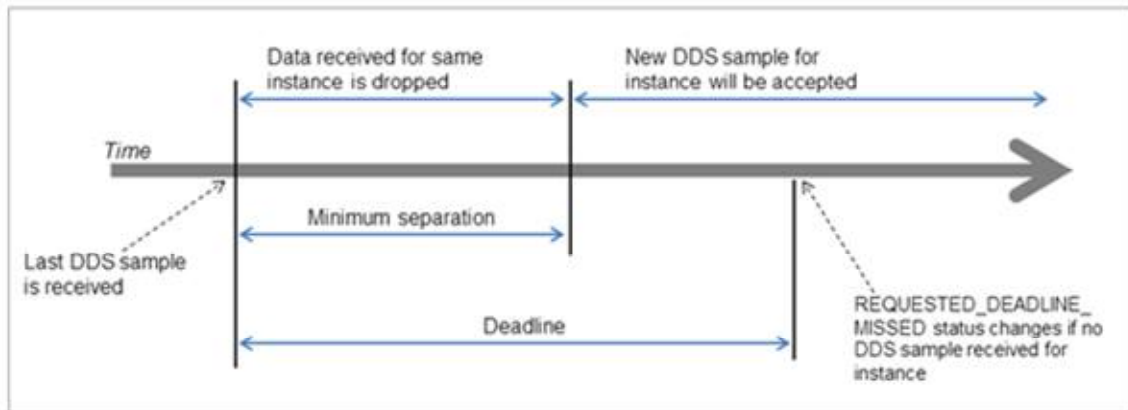
[그림 39] CAN/USB 기반의 멀티모달 인터페이스 프로토타입

(좌: CAN BUS 셋업 환경, 우: 오픈 하드웨어 플랫폼 기반의 인터페이스)

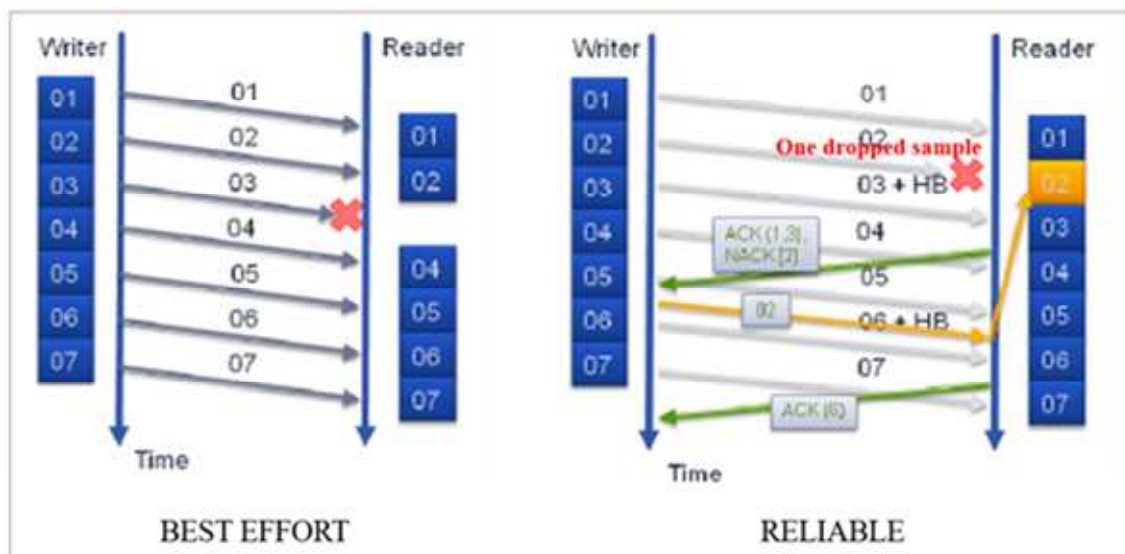
- 현재의 자동차는 CPU(Central Processing Unit)을 중심으로 각종 센서와 전자장치가 융합된 거대한 컴퓨터 시스템으로 발전하고 있음. 자동차 외부의 센서를 통해 수집된 데이터를 즉각적으로 처리 및 분석하여 자율적으로 움직이거나, 자동차 내부의 센서를 통해 운전자를 위한 각종 서비스 프로그램이 실행되는 일종의 컴퓨터 시스템으로 나아가고 있음
- 하지만, 현재의 인포테인먼트 시스템에 대한 기업별 자체 표준을 사용하기 때문에 인포테인먼트 시스템의 확장성과 장치 연결에 대한 표준화된 체계가 부족한 상황임[39]. 특히, eCall 등의 구조 요청을 가능하게 하는 3G 이동 통신과 웨어러블 기기에서 사용되는 BLE(Bluetooth Low Energy) 등의 자동차 커넥티비티(connectivity) 관련 IT 통신 기술들은 데이터가 손실될 가능성이 있는 신뢰성(reliability)이 보장되지 않는 네트워크라는 측면에서 문제가 있음
- 데이터 전송의 신뢰성이 보장되지 않는다는 것은 우선순위를 부여하지 않아 데이터 전송이 지연될 수 있는 비실시간 통신 방식을 의미하며, 통신 장치가 늘어나는 것에 비례하여 데이터 전송의 품질은 떨어지는 방식임. 따라서, 운전자의 위급 상황과 관련된 데이터 전송에 적합한 방식이 아님
- 만약, 시속 100km로 주행하는 자동차가 있을 때, 운전자가 1초를 줄게 되면 무의식주행으로 28m를 이동하게 되는 위험한 상황이 발생함. 자동차의 센서로 부터 올바르게 센싱된 데이터가 있더라도 통신 레벨의 지연 또는 누락이 발생한다면, 자동차는 결국 적시에 올바른 동작을 수행할 수 없거나 이미 사고가 발생한 뒤 상황을 인지할 가능성이 있음. 수집된 운전자의 PPG 및 이미지 데이터는 일종의 시계열(Time Series) 데이터로 측정된 데이터가 상태감지 시스템으로 전달되기까지 누락과 지연이 발생한다면 적절한 시점에 상태 인식이 어려울 수 있음
- 위급 상황이 발생했을 때 미래의 자동차는 운전자로부터 5초 이내로 제어권을 이양을 결정[31, 32]할 수 있도록 권장하고 있음. 이러한 실시간 및 신뢰성 있는 데이터 통신 관점에서, 자동차는 우리가 흔하게 접하는 일반적인 컴퓨터와 다르게 운전자의 안전을 보장해야 하는 특수성을 가짐. 이러한 안전을 보장해야 하는 특수성은 국방과학 분야의 전투기, 항공모함, 전차에 탑승하는 파일럿의 안전을 제공하는 특수성과 유사함
- DDS는 국방과학 분야의 전자장치 간 실시간 및 신뢰성있는 데이터 교환을 보장하기 위해 제정되었으며, 한국전자통신연구원, RTI 등 국내·외 다양한 벤더들이 DDS 표준을 구현하여 국방과학 전자장치에 적용하고 있음. DDS는 가장 큰 특징은 여러 종류의 서비스 품질(Quality of Service, QoS)를 정의하여 전자장치 간 데이터 교환에 우

선순위를 부여하거나(실시간성), 전송수준을 보장(신뢰성)하는 것임

- 자동차 분야에서의 DDS의 활용은 대표적으로 Audi사 인데, 이미 2012년 부터 RIT에서 개발한 DDS를 이용해 자동차 전자장치와 그를 이용한 응용들이 안정적으로 통신 및 실행되는지를 검증하기 위한 테스트 플랫폼으로 사용하고 있음[36]



[그림 40] DDS의 실시간 서비스 품질 지원에 대한 개념도[40]

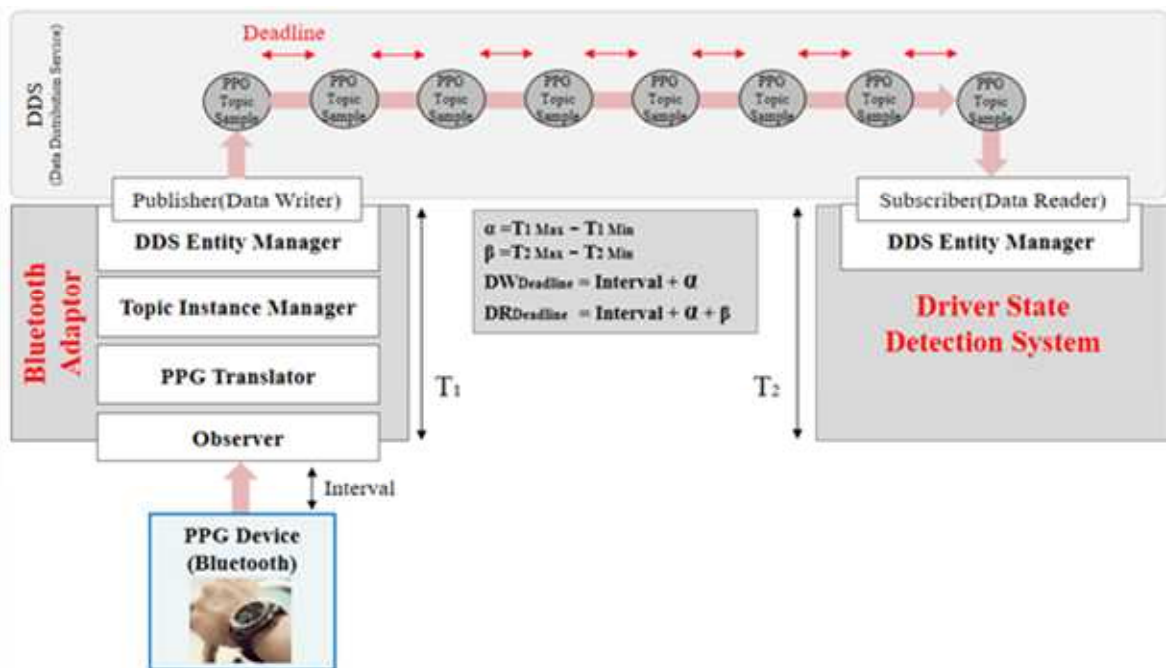


[그림 41] DDS의 신뢰성 서비스 품질 지원에 대한 개념도[41]

- [그림 39]는 DDS의 실시간 서비스 품질 지원에 대한 개념도로, 데이터를 전송하는 장치(Writer)와 데이터를 수신하는 장치(Reader) 사이에 전달되는 데이터의(샘플) 데드라인(Deadline)을 설정하여 적시에 데이터가 전달되는 지를 모니터링함. 만약 지연과 누락이 발생한다면 데드라인이 지켜지지 않았음을 공지하게 되고, 그에 따라 신뢰성(reliability) 서비스 품질에 대한 수준을 변경하여 누락과 지연이 발생하지 않도록 조정

해야 함

- [그림 40]은 DDS의 신뢰성 서비스 품질에 대한 개념도로 BEST EFFORT와 RELIABLE 사이의 차이점을 설명함. BEST EFFORT 방식은 데이터를 전송하는 장치와 수신하는 장치 사이에 성능(지연시간 최소)을 우선순위로 하여 전송하는 방식임. 이 경우 통신 네트워크가 감당할 수 있는 수용력(capacity)에 비하여 과도하게 데이터가 생산되어 전송될 경우 03번 샘플과 같이 데이터의 손실이 발생함
- RELIABLE 방식은 누락된 샘플(dropped sample)에 대한 추적 관리가 가능한 프로파일로 ACK/NACK 메시지를 통해 샘플을 재전송할 수 있는 방법임. 이는 누락된 데이터에 대하여 추가적으로 센서가 다시 센싱을 통해 데이터를 생산하는 방식이 아님. 통신 네트워크 자체적으로 ACK/NACK를 고려하여 임시적으로 샘플들을 저장해 두었다가 손실이 발생한 경우 즉시 재전송하는 방식임
- 자동차는 여러 제조사가 제작한 부품과 전자장치가 결합되어 동작하는 시스템으로 CPU가 존재하는 한 필연적으로 전자장치 간 통신을 위한 네트워크가 존재하며, 이 네트워크를 통해 각종 센서와 전자장치들이 데이터를 교환하는 구조임



[그림 42] DDS의 QoS를 기반으로 실시간 및 신뢰성을 제공하며 운전자의 PPG 데이터를 수집하는 상태인식 시스템 동작 개념도

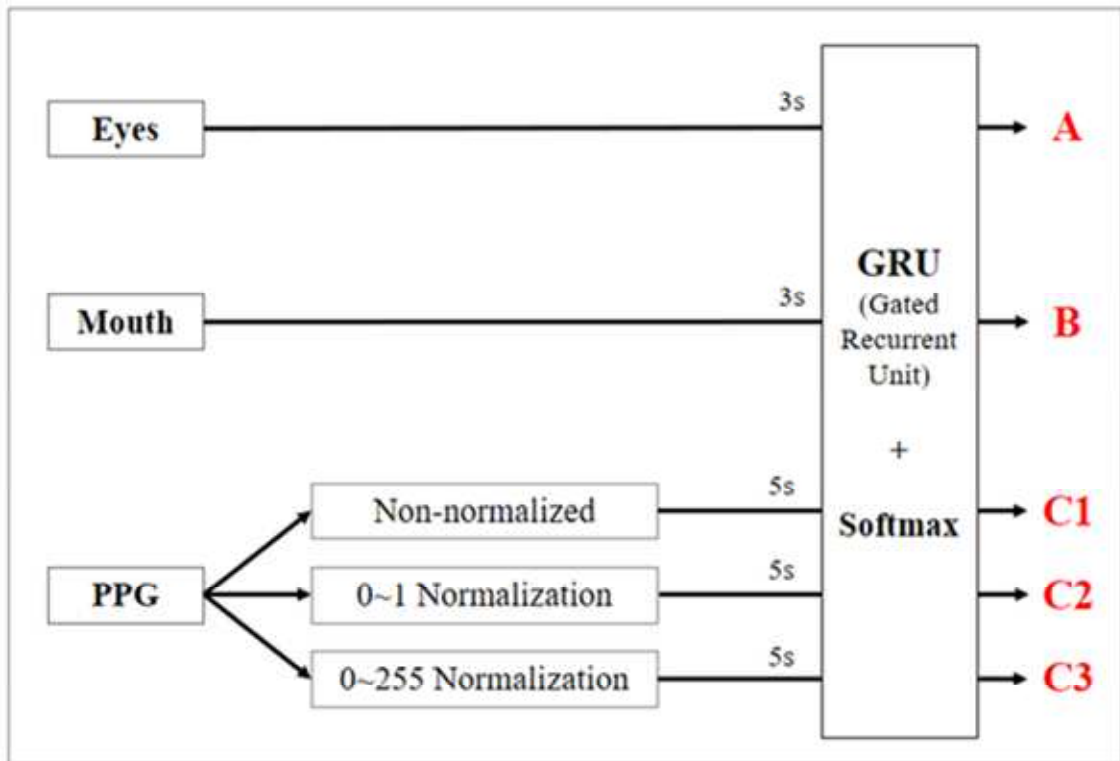
- 본 연구개발은 아직까지 표준화된 통신 방식이 없는 자동차의 인포테인먼트 시스템에 기존 국방과학 분야의 표준 DDS 기술을 적용하여, 본 연구개발의 핵심 기술인 운전

자 상태인식 시스템으로의 안전성을 확보하기 위해 [그림 41]과 같은 구조의 시스템을 개발하였음

- PPG 디바이스: 스마트 워치와 같이 운전자의 PPG를 측정할 수 있으면서, 블루투스를 통한 데이터 전송 능력이 있는 디바이스
- 블루투스 어댑터(Bluetooth Adaptor): 등록특허 10-1911359에 따라 구현된 DDS와 블루투스 장치 간 데이터 교환 통신 모듈 (해당 등록특허는 본 연구개발을 통해 도출되었으며, 발명자는 연구책임자 및 참여연구원)
- 데드라인 서비스 품질(Deadline QoS): 서로 다른 모듈 간 전송되는 데이터 샘플의 주기성을 부여하여, 해당 주기에 데이터가 수신자에게 수신될 수 있도록 보장하는 프로파일이며, 본 연구개발은 DDS로 연결된 모듈의 시간을 측정하여 설정
- 신뢰성 서비스 품질(Reliable QoS): 서로 다른 모듈 간 전송되는 데이터 샘플의 누락이 발생했을 경우 누락된 샘플에 대한 추적 관리를 통하여 ACK/NACK 메시지를 이용해 재전송 및 수신하는 프로파일이며, 본 연구개발은 [그림 40]의 BEST EFFORT를 사용하지 않음

(6) 유니모달(unimodal) 딥러닝 기반의 운전자 졸음 상태 감지(멀티모달 방식과 비교군)

- 2.-가.-(4)에서 제시한 국내·외 관련연구 현황과 분석 내용과 같이, 본 연구개발이 도출하려는 멀티모달 딥러닝 기반의 운전자 졸음 상태 인식 모델과 직접적으로 비교가 가능한 연구결과는 국내·외적으로 아직 존재하지 않음. 다만, [53]의 경우 운전자의 심장활동 정보를 사용하지 않지만 멀티모달 딥러닝 방식으로 졸음 인식 모델을 도출하였으며 인식 결과도 3가지 클래스이기 때문에 성능비교 대상에 포함하였음
- 하지만, 하나의 인식 모델만 비교하는 것은 본 연구개발이 도출한 모델의 객관적인 비교 결과라고 보기 어렵기 때문에, 본 연구팀이 수집한 학습 데이터를 유니모달(하나의 모달리티만 사용) 딥러닝으로 운전자 졸음 상태 감지 모델을 도출하고, 이 모델에 모달리티를 추가하였을 때 인식 정확도가 어느정도 개선되는지 실험을 통해 검증하는 방식으로 추가적인 성능비교를 진행하였음

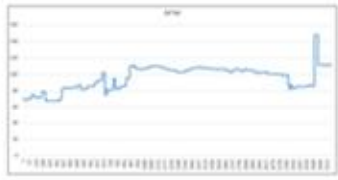


[그림 43] GRU를 이용한 유니모달 딥러닝 기반의 5가지 졸음 인식 모델 개념도

- [그림 42]은 눈, 입, PPG 중 하나의 모달리티를 사용하여 졸음을 인식하는 모델에 대한 개념도임. 딥러닝 알고리즘은 RNN(Recurrent Neural Network)의 대표적인 셀 가운데 하나인 GRU(Gated Recurrent Unit)로 동일하게 설정하였음. GRU는 LSTM(Long-Short Term Memory) 셀을 개선하여 LSTM의 장점을 유지하면서 계산의 복잡도는 낮은 셀 구조임. RNN은 CNN과 다르게 신경망 내부의 활성화 함수(activation function)을 통해 나온 값이 출력 레이어로 향하지 않고, 활성화 함수를 통해 나온 값을 출력 층과 자기 자신의 다음 계산의 입력으로 보내는 특징이 있음. 이러한 특징은 이전의 값을 기억하는 일종의 메모리 역할을 수행하는 효과가 있음. 그래서 RNN은 CNN과 다르게 공간적인 특징이 아닌 시계열 특징이 있는 학습 데이터에서 우수한 성능을 보이고 있음. 하지만, 기울기 값을 기반으로 역전파(backpropagation)를 통해 학습이 이루어지는 뉴럴 네트워크는 은닉 레이어(hidden layer)의 깊이가 깊어질수록 기울기값이 0으로 수렴하는 문제(Vanishing Gradient Problem)가 발생함. 이러한 문제는 결국 모델이 학습을 하지 못하는 결과로 이어짐. LSTM은 RNN의 Vanishing Gradient 문제를 극복한 셀 구조이며, 더 나아가 GRU는 LSTM의 장점을 유지하면서 더욱 빠른 학습을 가능하게 하였음
- 본 유니모달 딥러닝 기반의 운전자 졸음 상태 감지는 [그림 42]의 GRU는 동일한 구조

이면서 학습하는 데이터의 종류만 눈, 입, PPG로 변경하여 졸음 인식 모델을 도출하였음(A: 눈 이미지를 학습, B: 입 이미지를 학습, C: PPG 데이터를 학습)

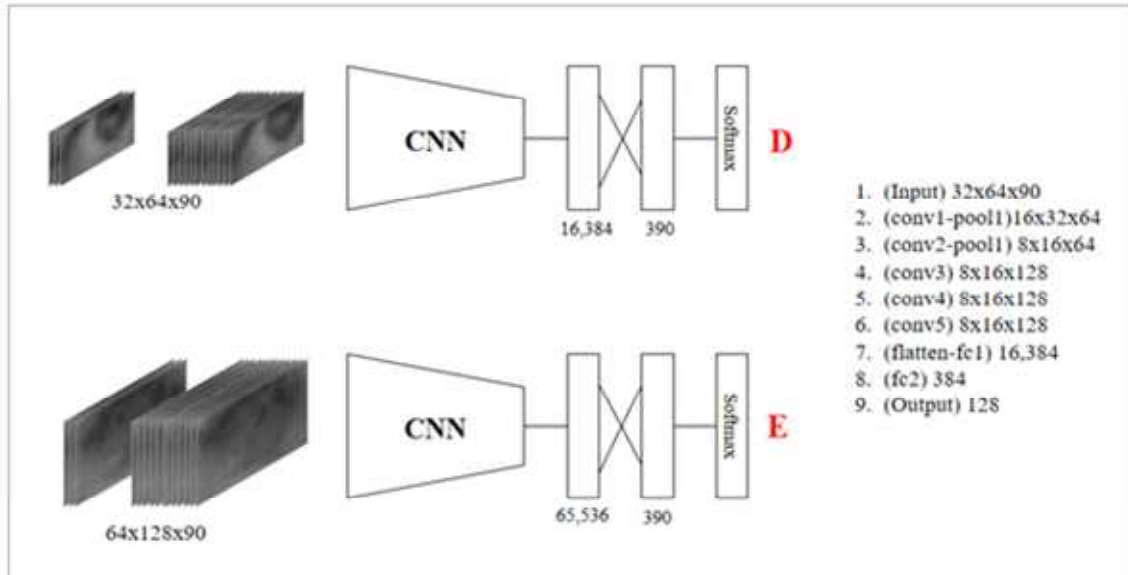
- 눈과 입의 경우 특별한 전처리 과정 없이 학습 데이터로써 사용하였으나, PPG 데이터의 경우 연구참여자마다 PPG 값의 편차가 제법 있기 때문에 3가지 방식의 전처리(preprocessing) 과정을 추가하여 정규화(Normalization)를 진행하였음. 특히, C3는 추후 회색조(grayscale) 이미지 형태로 변환을 고려한 방식임(심박수가 높을수록 밝아지는 이미지)
 - C1: 정규화 없음, 학습 데이터 세트에 추가적인 전처리를 수행하지 않음
 - C2: 0~1로 정규화, 학습 데이터 세트를 0~1 범위 내에 존재하도록 정규화
 - C3: 0~255로 정규화, 학습 데이터 세트를 0~255 범위 내 존재하도록 정규화

C1	C2	C3
Non-normalized Data	$N = \frac{R}{\text{avg}(R)} * 0.5$	$P = (N - \min(N)) * 255$
		

[그림 44] PPG 데이터의 정규화 과정

- [그림 44]는 눈 또는 입 이미지 중 하나만을 사용하여 졸음을 인식하는 CNN 기반의 학습 모델에 대한 개념도임. CNN을 이용하여 자동으로 졸음과 관련된 유의미한 특징을 추출하고 졸음 상태를 인식함. PPG 데이터를 포함하지 않은 이유는 CNN은 이미지와 같은 공간적 특징을 추출하여 학습하기 때문에 공간적 특징이 없는 배열 형태의 PPG 데이터는 CNN을 이용해 학습하는데 적합하지 않음. 본 CNN 기반의 학습 모델은 아래와 같은 입력 노드와 레이드에 따라 학습이 이루어짐
 - 입력 노드: 32x64 해상도의 눈 이미지 90장, 또는 64x128 해상도의 입 이미지 90장 (90장: CNN의 채널 수)
 - 컨볼루션 레이어(Convolutional Layer, Conv): 5개의 컨볼루션 레이어를 통해 입력된 이미지의 공간적 특징을 추출, 2개의 Pooling 레이어를 통해 압축
 - 완전 연결 레이어(Fully-connected Layer, FC): 2개의 분류 레이어를 통해 특징 값에

대한 분류 작업이 이루어지며, Softmax를 통해 3가지 졸음 상태를 분류(정상/얕은 잠/깊은 잠)



[그림 45] CNN을 이용한 유니모달 딥러닝 기반의 2가지 졸음 인식 모델 개념도

- [그림 42], [그림 44]와 같이 설계된 유니모달 딥러닝 기반의 졸음 인식 모델은 [표 8]과 같은 정확도를 보였음. 정확도 측정은 10-fold CV(Cross Validation) 방식을 사용하였으며 학습에 사용한 데이터는 테스트에 사용되지 않도록 교차되어 총 10회의 정확도를 산출함. 최종 산출된 정확도는 10번 측정된 정확도의 평균값을 의미함
- 모델 A, D는 눈 이미지만 사용하여 학습한 모델로 GRU 기반의 학습 모델보다 CNN을 사용하였을 때 약 4.5% 정도의 성능 증가가 있음을 확인하였음. 일반적으로 공간적 특징을 담고 있는 데이터에 대하여 CNN이 우수한 성능을 보인다는 사실에 부합하는 결과가 도출되었음
- 모델 A, D는 기존의 전통적인 기계학습인 [표 2]의 관련연구에 준하는 수준의 성능을 보였기 때문에 두 모델 모두 졸음에 대한 명확한 특징을 찾았다고 볼 수 있음. 따라서, 본 연구팀이 수집한 데이터의 양과 질은 딥러닝을 통한 학습 모델 도출이 가능한 수준으로 볼 수 있음

[표 8] 도출된 유니모달 딥러닝 기반의 졸음 인식 모델의 성능 비교

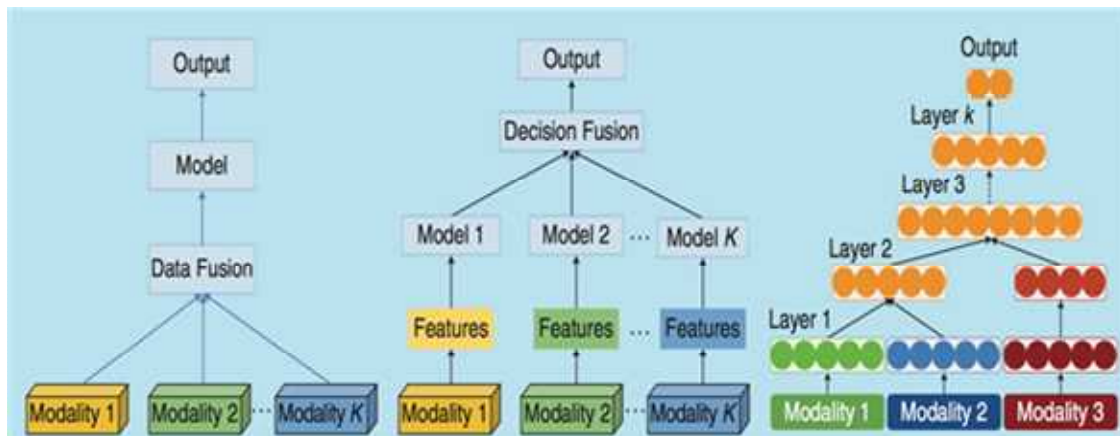
	눈, 입, PPG 단일 모달리티만 사용하여 학습한 모델				
	A	D	B	E	C3
레이블	Normal/Light Sleep/Deep Sleep	Normal/Light Sleep/Deep Sleep	Normal/Light Sleep/Deep Sleep	Normal/Light Sleep/Deep Sleep	Normal/Light Sleep/Deep Sleep
학습 데이터	29,286 (55:42:3)	29,286 (55:42:3)	29,286 (55:42:3)	29,286 (55:42:3)	25,227 (54:43:3)
테스트 데이터	3,254 (55:42:3)	3,254 (55:42:3)	3,254 (55:42:3)	3,254 (55:42:3)	2,803 (54:43:3)
학습 모델	GRU (Layer: 7, node: 512)	CNN (Channel: 90, 5 Cov., 2FC)	GRU (Layer: 7, node: 512)	CNN (Channel: 90, 5 Cov., 2FC)	GRU (Layer: 7, node: 512, 0~255 Norm.)
정확도 (10-fold CV)	84.51%	89.30%	90.30%	90.55%	55.2%

- 모델 B, E는 입 이미지만 사용하여 학습한 모델로 RNN 계열의 학습 모델과 CNN 기반의 학습 모델의 유의미한 인식 성능 차이가 나타나지 않았음. 두 모델 모두 [표 2]의 관련연구와 비교했을 때 관련연구에 준하거나 더 나은 성능을 보였음. 일반적인 인식과 다르게 사람의 입은 졸음이 발생할 때 눈의 변화와 같이 명확한 특징 변화가 있음을 확인할 수 있음
- C3는 PPG 데이터만 사용하여 학습한 모델로 졸음 인식 정확도가 55% 수준을 보였음. 이는 학습 모델이 졸음을 분류하기 위한 명확한 특징을 찾지 못해 학습이 전혀 이루어지지 않았음을 의미함(C1, C2도 C3와 거의 같은 인식 성능을 보였음). 정규화 과정을 통해 개인의 심박 데이터 편차를 보정하였음에도 이러한 결과가 나온 이유는, 주행 중 발생하는 심박 정보에는 졸음으로 인식될 수 있는 유사한 패턴이 다수 존재하기 때문으로 판단됨. 따라서, PPG 데이터 단독으로 졸음 인식 모델을 도출하는 것은 실제 주

행 환경에 큰 효과가 없을 것으로 예상됨

- 본 유니모달 딥러닝 기반의 졸음 인식 모델에 대한 성능 분석을 요약하면 눈, 입 이미지는 졸음을 분류할 수 있는 명확한 특징이 존재하고, 사람의 개입 없이 기존의 딥러닝을 적용하여 관련연구에 준하는 인식 모델을 도출할 수 있음을 확인하였음. 또한, 본 유니모달 기반의 모델은 3가지 클래스(Normal/Light Sleep/Deep Sleep)를 분류하는 모델이기 때문에 2가지 분류 모델에 비하여 좋은 성능을 도출하기 어려움에도 90% 수준의 분류 정확도를 보여 본 연구팀이 수집한 데이터의 규모가 딥러닝에 사용되기 적절하고 질적 수준도 매우 양호하다고 볼 수 있음

(5) 멀티모달 딥러닝 기반의 운전자 졸음 상태 감지 - 데이터 융합 방식 적용

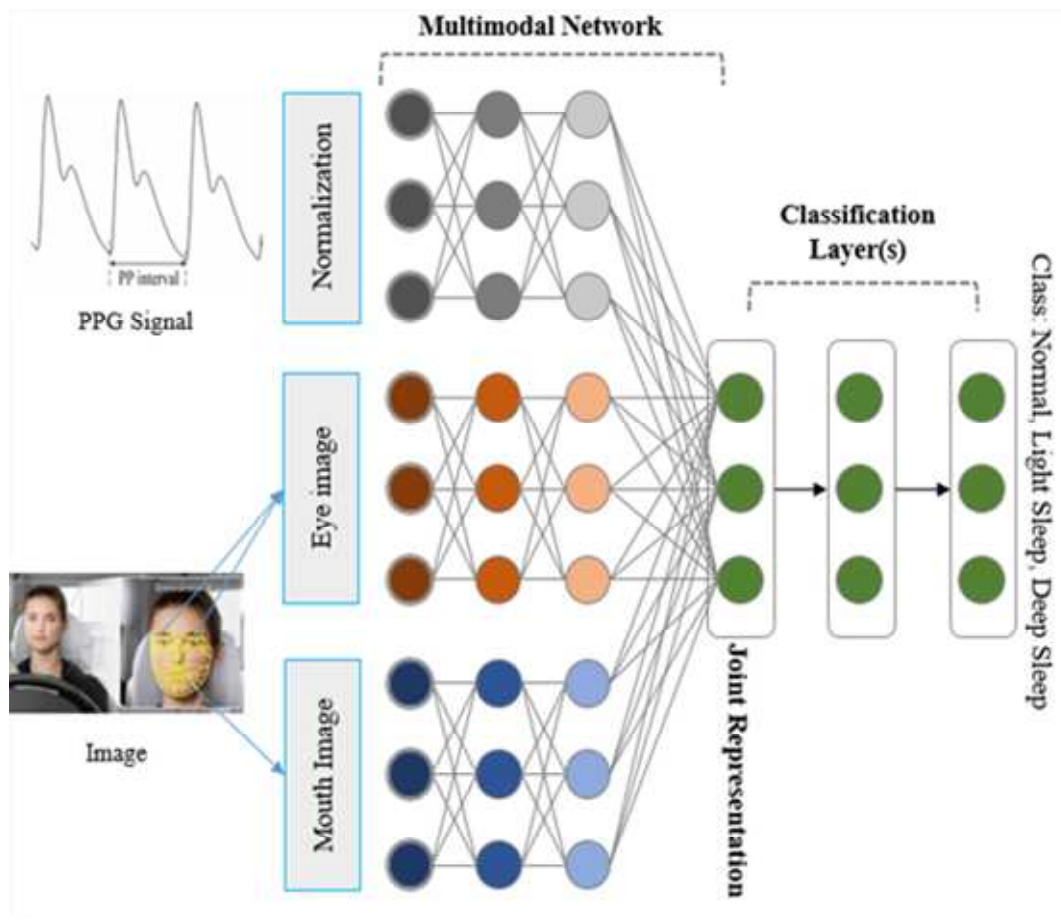


[그림 46] 멀티모달 딥러닝의 3가지 멀티모달 네트워크 구성 방식

- 멀티모달 딥러닝은 [그림 45]와 같이 3가지 방식의 유연한 모달리티 융합 방식을 사용하여 학습 모델을 도출할 수 있으며 아래와 같은 특징이 있으며 높은 성능을 위해 데이터의 종류들에 따라 모달리티 융합 방식을 선택할 수 있는 유연한 방법임. 본 연구 개발에서는 결정 융합 방식을 사용하지 않았음. 결정 융합 방식은 유니모달 딥러닝 기반의 모델을 도출하고, 각 모델이 도출한 결과 값을 이용해 간단하게 최종 결과를 산출하는 것으로 기존 관련연구와 비교하여 뚜렷한 차별성이 없는 방식이기 때문에 제외하였음
- 데이터 융합(Data Fusion): 학습과정의 입력 데이터로 들어가기 전에 모든 모달리티의 데이터들 하나의 특징벡터(feature vector)로 결합(concatenation)됨. 이 결합 방식은 센서로부터 바로 들어온 원천 데이터(raw data)들의 단순 결합이기 때문에 복잡

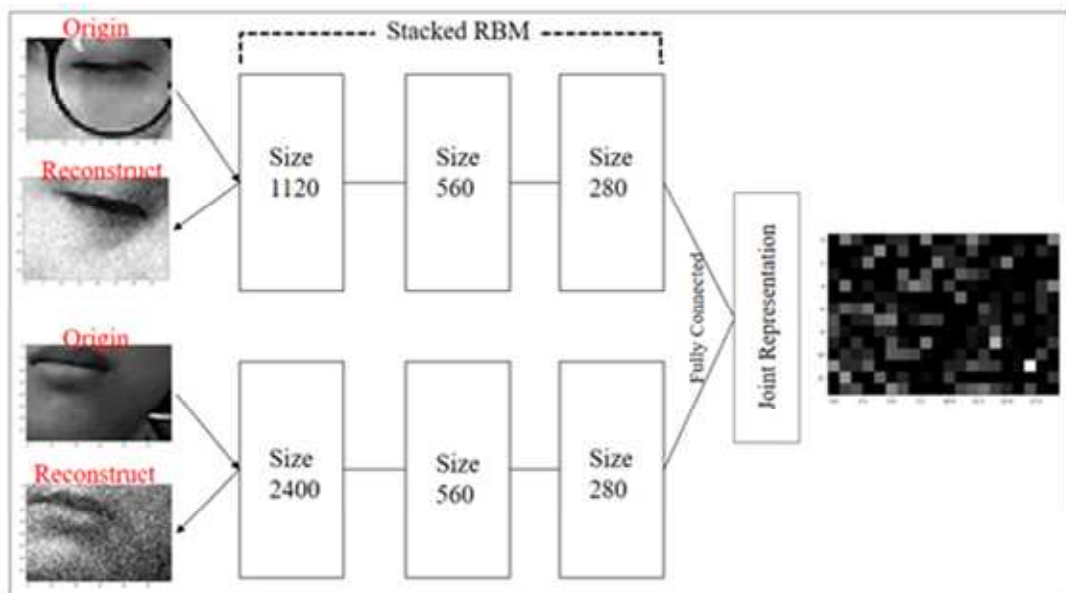
도(차원이 높음)의 문제로 모델 도출을 위해 많은 데이터가 요구될 수 있음

- 결정 융합(Decision Fusion): 각 모달리티의 데이터를 이용해 도출된 학습 모델의 인식(예측) 결과를 종합적으로 판단하여 최종 인식 결과를 도출하는 방식임. 기존의 앙상블(Ensemble) 방식의 기법과 유사하나 모달리티가 다르기 때문에 최종 인식 결과의 산출 방식을 다양하게 고려해야 함. 예를 들면, 결정 퓨전에 입력되는 각 모델의 결과값을 점수제로 환산하여 가장 높음 점수의 결과를 최종 결과로 채택하는 방식이나, 각 모델의 인식 결과를 투표(voting)하여 가장 많은 득표가 있을 경우를 최종 인식 결과로 산출함
- 중간물 융합(Intermediate Fusion): 멀티모달 딥러닝에서 도전적으로 사용하는 융합 방법으로 특정 단계의 레이어에 결합되어 학습 과정 중간 중간에 직접적으로 영향을 미치게 됨. 이 경우 자칫 특정 모달리티가 학습에 과하게 기여하여 올바른 학습 모델이 도출되지 못할 수 있으나, 적정 수준의 융합 지점을 찾게 되면 서로 다른 모달리티에 대하여 인간의 개입 없이 스스로 특징을 추출하고 인식 결과를 산출할 수 있음



[그림 47] 멀티모달 딥러닝 기반의 운전자 졸음 감지 모델의 개념도

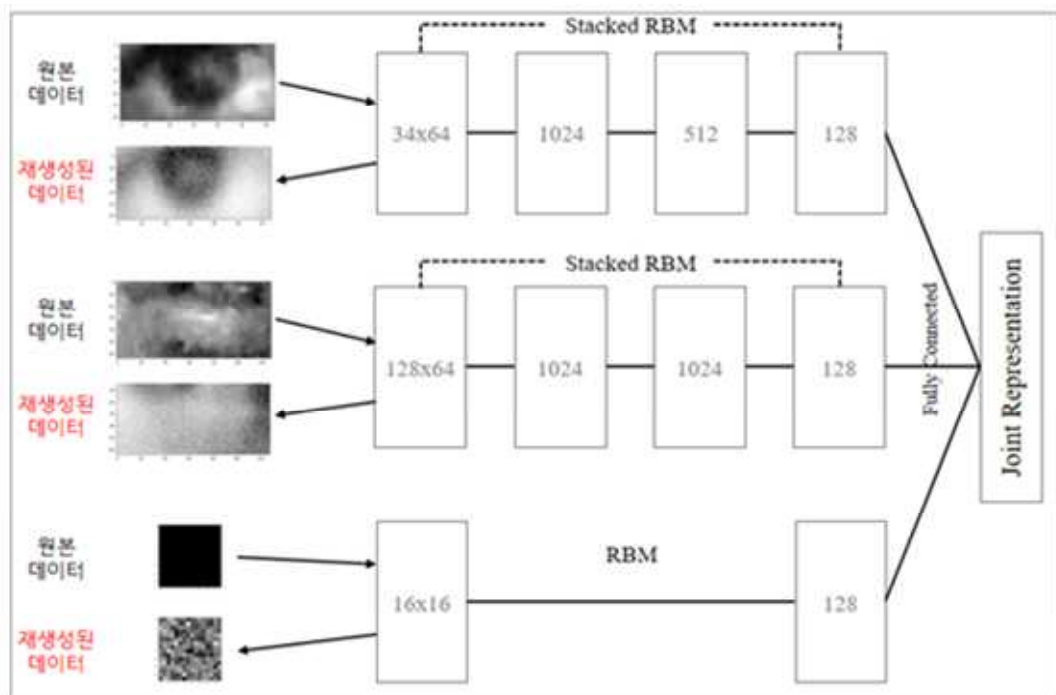
- 하나의 데이터 형태라고 할지라도 그 데이터 안에는 여러 모달리티가 존재할 수 있음. 대표적으로, 얼굴 이미지는 하나의 카메라를 통해 하나의 이미지에 담기지만, 사람의 얼굴은 눈, 코, 입 등의 여러 기관이 존재하기 때문에 각각 서로 다른 모달리티로 볼 수 있음. 졸음의 경우, 사람은 졸음이 발생하면 눈에 가장 큰 변화가 나타나지만 입 주변의 근육들이 평소와 다르게 이완되어 입의 모양의 변화가 나타남. 즉, 눈과 입은 근본적으로 서로 다른 기관이기에 졸음 발생 시 각각 고유의 특징이 나타나게 됨
- 이러한 멀티모달 딥러닝의 내용에 따라 본 연구개발은 [그림 46]과 같은 연구범위를 설정하여 연구개발을 진행하였음. 심박활동 정보를 담은 운전자의 PPG 데이터와 해당 PPG 데이터가 생산되는 시점의 눈과 입의 변화를 담은 운전자 얼굴 이미지 데이터를 학습 데이터로 사용함. 눈, 입, PPG 세 종류의 데이터는 멀티모달 네트워크(Multi modal Network)를 통해 데이터의 복잡도(차원 축소)가 줄어들면서 졸음을 구분하는 지표인 주요 특징들이 알고리즘 수준에서 자동으로 추출됨. 추출된 특징들은 [그림 45]와 같은 모달리티 융합 방식에 따라 적정 수준의 Joint Representation에서 하나의 데이터 형태로 표현됨. 마지막으로 분류 레이어(Classification Layer)에서는 일반적인 딥러닝 방법을 사용하여 세 가지의 졸음 분류 결과를 도출함(Normal/Light Sleep/Deep Sleep)



[그림 48] RBM을 이용한 멀티모달 네트워크 구성 및 데이터 융합 결과 확인

- RBM(Restricted Boltzmann Machine)은 데이터의 차원을 감소시키거나 특징값 학습

(feature learning)에 사용될 수 있는 알고리즘임. RBM은 자체적으로도 사용할 수 있지만 다른 딥러닝 모델의 학습을 돕기위해 사용되기도 함. 대표적으로 DBN(Deep Belief Network)이며, RBM을 기반으로 딥러닝 모델을 사전 학습(pretraining)하고 역전파 알고리즘을 돌려 과적합(overfitting) 문제를 낮추는데 사용됨. 본 연구개발은 RBM이 가진 차원 축소와 특징값 학습 기능에 초점을 맞추어 두 가지 이상의 모달리티를 하나의 데이터로 융합하는 연구개발을 진행하였음. [그림 47]은 3개의 히든 레이어를 가진 RBM을 이용하여 눈, 입 각각의 학습 데이터의 확률 분포를 학습하고 하나의 Joint Representation을 구성하는 데이터 융합에 대한 개념도임.



[그림 49] RBM을 이용한 멀티모달 네트워크 구성 및 데이터 융합 결과 확인

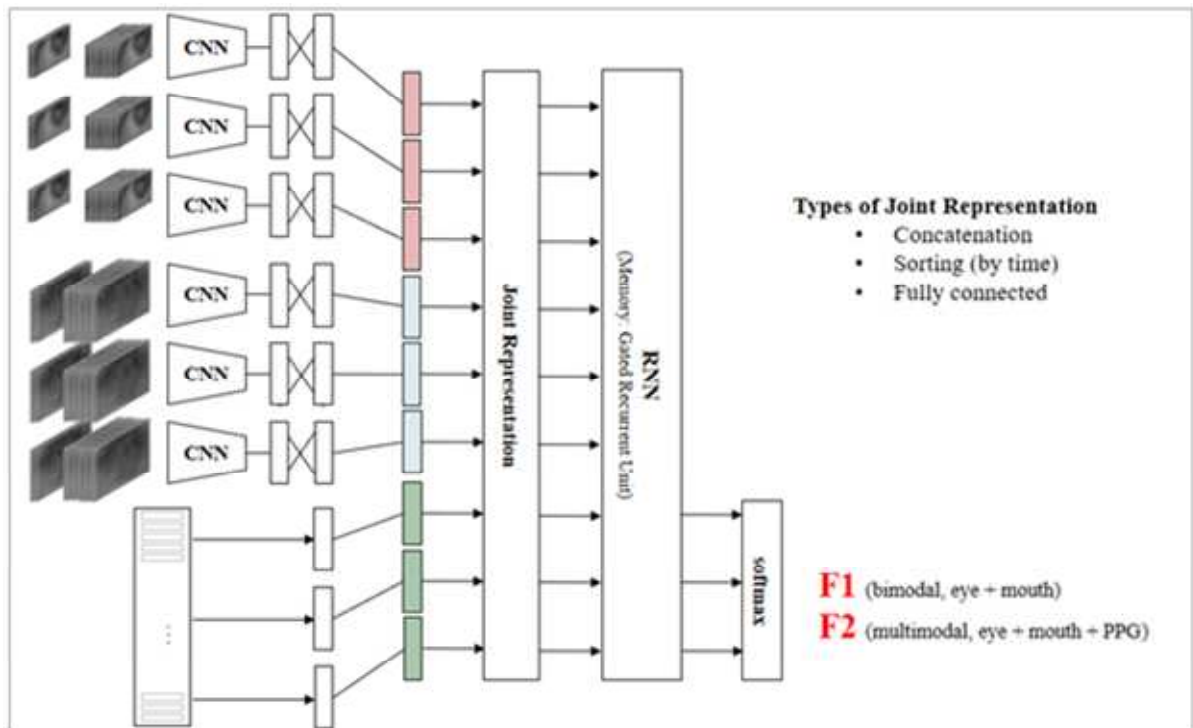
- 본 연구팀은 학습이 완료된 멀티모달 네트워크를 테스트하기 위하여 임의의 눈, 입 이미지를 입력하고 재생성된 이미지(reconstruct)를 확인하였음. 그 결과 입력 이미지에서 불필요한 특징이 제거되는 현상을 확인할 수 있었음. 대표적으로, [그림 48]과 같이 안경을 쓴 입력 이미지에 대하여 멀티모달 네트워크를 통과한 재생성된 이미지는 안경이 제거된 결과가 출력됨. 하지만, [그림 48]의 재생성된 입 이미지와 같이 재생성된 이미지는 비교적 잡음이 심하게 생성된 경향이 있음. 본 연구팀은 학습이 완료된 멀티모달 네트워크에 분류를 위한 뉴럴 네트워크를 추가하여 줄음 상태를 인식하는 학습 모델이 생성되는지를 확인하였음. Joint Representation 뒤에 2개의 완전 연결 레이어

특징 값에 대한 분류 작업이 이루어지며, Softmax를 통해 3가지 졸음 상태를 분류하는 학습 모델을 생성하였음. 하지만, 생성된 학습 모델은 전혀 학습이 이루어지지 않았음. 그 이유는 [그림 47]의 Joint Representation의 결과를 통해 유추해 볼 수 있는데, 서로 다른 특징의 눈, 입 이미지가 하나의 데이터로 융합되면서 기존의 유의미한 특징 값들이 사라지는 현상이 발생한다고 볼 수 있음. 그 결과 분류 레이어에서는 졸음을 구분하는 명확한 특징 값을 도출하지 못했다고 볼 수 있음. 본 연구팀은 [그림 48]과 같은 추가 실험을 진행했지만 노이즈에 대한 개선을 도출하지 못했고, 도출한 인식 모델도 학습이 이루어지지 않은 모델이었음

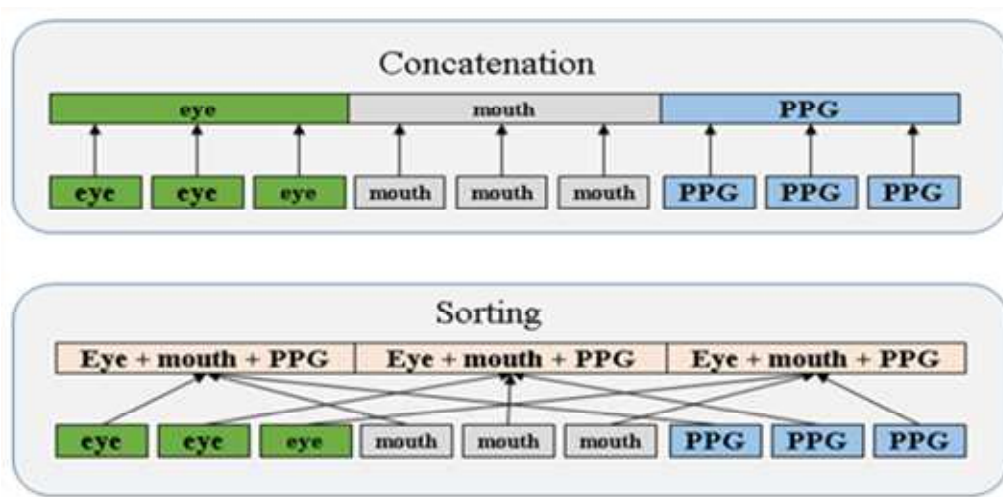
(6) 멀티모달 딥러닝 기반의 운전자 졸음 상태 감지 - 중간물 결합 방식 적용

- 데이터 융합 방식의 멀티모달 네트워크는 졸음을 구분하는 주요 특징 값들이 사라지는 문제가 발생하여 올바른 인식 모델을 도출할 수 없었음. 본 연구팀은 접근 방식을 바꾸어 레이어 단계에 데이터가 융합 되도록 중간물 융합 방식으로 학습 방식을 변경하였음. 이 경우 특정 모달리티가 학습에 과하게 기여하여 올바른 학습 모델이 도출되지 못하는 문제가 발생할 수 있으나, 여러 가설과 그에 따른 실험을 통해 [그림 49]와 같은 적정 수준의 융합 지점을 찾게 되었음
- [그림 49]는 본 연구개발의 최종적인 멀티모달 딥러닝 방식의 졸음 인식 개념임. 눈과 입 이미지는 기존 연속된 90장의 이미지를 30장 단위로 나누고, RBM 대신 [그림 44]와 동일한 구조의 CNN을 통해 특징을 추출하도록 하였음. PPG의 경우는 시간에 따라 나열된 연속된 30개의 데이터를 10개 단위로 나누어 Joint Representation에 입력하였음
- 본 멀티모달 딥러닝의 핵심은 Joint Representation에서 서로 다른 모달리티의 데이터를 하나의 데이터 형태로 조립하는 것임. Joint Representation의 데이터 조립 방식은 결합(concatenation), 정렬(Sorting), 완전 결합의 3가지로 구분하여 최종적으로 졸음 상태를 분류하는 RNN(셀 구조: GRU)의 성능이 얼마나 영향을 미치는지 측정하였음
- Joint Representation의 결합(Concatenation) 방식은 구성된 데이터가 눈-입-PPG 순으로 조립되도록 하는 방식이며, 정렬(Sorting) 방식은 모달리티의 종류보다 동 시간대에 생성된 데이터의 특징을 부각시키기 위해 시간 순서에 따라 다시 조립하는 방식임. 완전 연결방식은 결합, 정렬 방식과 다르게 Joint Representation을 마치 RNN의 히든 레이어처럼 모든 노드를 연결하여 구성하는 방식임. 완전 연결의 경우 Joint Rep

resentation 이전의 분할된 각각의 데이터가 RNN의 입력 레이어가 됨



[그림 50] 중간물 결합 방식의 멀티모달 네트워크가 적용된 멀티모달 딥러닝 기반의 줄음 인식 모델 개념도



[그림 51] 멀티모달 딥러닝 기반의 줄음 인식 모델이 사용하는 Joint Representation 데이터 조립 방식 차이

- 본 멀티모달 딥러닝 방식은 CNN과 RNN이 모두 사용되기 때문에 총 2번의 학습 과정이 이루어짐. 먼저 6개의 CNN을 통해 눈과 입 이미지에 대한 특징 추출 및 학습이

이루어지게 되고, 학습이 완료되면 Joint Representation 방식에 따라 조립된 학습 데이터로 RNN의 학습이 이루어짐. 단, 완전 결합의 Joint Representation은 GRU의 학습 결과를 바탕으로 CNN의 학습이 이루어지기 때문에 1번의 학습 과정으로 이루어지는 차이가 있음

- 추가적인 비교군을 만들기 위해 [그림 50]의 F1, F2로 학습 모델을 구분하여 도출된 학습 모델의 성능을 비교함. F1은 바이모달(bimodal)로 PPG가 제외된 눈과 입 이미지의 두 가지 모달리티를 가짐. F2는 눈, 입, PPG를 모두 학습하는 세 가지의 모달리티를 가진 학습 모델을 의미함

[표 9] 도출한 멀티모달 딥러닝 기반의 졸음 인식 모델 성능 비교
(JR: Joint Representation)

	눈, 입, PPG를 멀티모달 학습	
	F1	F2
레이블	Normal/Light Sleep/Deep Sleep	Normal/Light Sleep/Deep Sleep
학습 데이터	25,227 (54:43:3)	25,227 (54:43:3)
테스트 데이터	2,803 (54:43:3)	2,803 (54:43:3)
학습 모델	CNN + JR + GRU (5 Cov., 2 FC, 5 GRU Layers)	CNN + JR + GRU (5 Cov., 2 FC, 5 GRU Layers)
정확도 (10-fold CV)	91.5% (JR: Concatenation)	91.1% (JR: Concatenation)
	90.08% (JR: Sorting)	92.8% (JR: Sorting)
	61.0% (JR: Fully-connected)	64.2% (JR: Fully-connected)

- [표 9]는 멀티모달 딥러닝 방식으로 도출된 모델 F1, F2의 Joint Representation 방식에 따른 졸음 인식 정확도임. 성능 분석은 10-fold Cross Validation 방식을 적용하여 학습에 사용한 데이터는 테스트에 사용되지 않도록 하며 10번의 성능을 측정함. 산출된 정확도는 10번 측정된 인식 성능의 평균값을 의미함
- 유니모달 딥러닝 방식보다 학습 및 테스트 데이터의 규모가 줄어든 이유는 기존 90장의 연속된 이미지를 30장 단위로 나누면서 시간적으로 연속되지 않는 학습 데이터는

학습 및 테스트에서 제외되었기 때문임. 일반적으로 딥러닝 모델은 학습 데이터의 규모가 줄어들면 도출된 모델의 성능도 감소되는 경향이 있음에도, 본 멀티모달 딥러닝 기반의 인식 모델은 기존 유니모달 방식보다 1~2% 성능 증가를 보였음. [표 2]의 국내·외 관련연구와 비교하여도 높은 성능을 보이는 것을 확인할 수 있음

(a)	Uni-modal model (CNN-eye) 3-class classification <table> <tr> <th></th><th>$Y_{ipred} = 0$</th><th>$Y_{ipred} = 1$</th><th>$Y_{ipred} = 2$</th></tr> <tr> <th>$Y_{itrue} = 0$</th><td>1792 (99.88%)</td><td>2 (0.12%)</td><td>0</td></tr> <tr> <th>$Y_{itrue} = 1$</th><td>2 (0.15%)</td><td>1331 (97.29%)</td><td>35 (2.56%)</td></tr> <tr> <th>$Y_{itrue} = 2$</th><td>5 (5.15%)</td><td>75 (77.32%)</td><td>17 (17.53%)</td></tr> </table>		$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$	$Y_{itrue} = 0$	1792 (99.88%)	2 (0.12%)	0	$Y_{itrue} = 1$	2 (0.15%)	1331 (97.29%)	35 (2.56%)	$Y_{itrue} = 2$	5 (5.15%)	75 (77.32%)	17 (17.53%)	(b)	Uni-modal model (CNN-mouth) 3-class classification <table> <tr> <th></th><th>$Y_{ipred} = 0$</th><th>$Y_{ipred} = 1$</th><th>$Y_{ipred} = 2$</th></tr> <tr> <th>$Y_{itrue} = 0$</th><td>1724 (98.01%)</td><td>35 (1.99%)</td><td>0</td></tr> <tr> <th>$Y_{itrue} = 1$</th><td>0</td><td>1341 (98.7%)</td><td>4 (0.3%)</td></tr> <tr> <th>$Y_{itrue} = 2$</th><td>0</td><td>90 (93.75%)</td><td>6 (6.25%)</td></tr> </table>		$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$	$Y_{itrue} = 0$	1724 (98.01%)	35 (1.99%)	0	$Y_{itrue} = 1$	0	1341 (98.7%)	4 (0.3%)	$Y_{itrue} = 2$	0	90 (93.75%)	6 (6.25%)
	$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$																																
$Y_{itrue} = 0$	1792 (99.88%)	2 (0.12%)	0																																
$Y_{itrue} = 1$	2 (0.15%)	1331 (97.29%)	35 (2.56%)																																
$Y_{itrue} = 2$	5 (5.15%)	75 (77.32%)	17 (17.53%)																																
	$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$																																
$Y_{itrue} = 0$	1724 (98.01%)	35 (1.99%)	0																																
$Y_{itrue} = 1$	0	1341 (98.7%)	4 (0.3%)																																
$Y_{itrue} = 2$	0	90 (93.75%)	6 (6.25%)																																
(c)	Bi-modal model (concatenation) 3-class classification <table> <tr> <th></th><th>$Y_{ipred} = 0$</th><th>$Y_{ipred} = 1$</th><th>$Y_{ipred} = 2$</th></tr> <tr> <th>$Y_{itrue} = 0$</th><td>1794 (100%)</td><td>0</td><td>0</td></tr> <tr> <th>$Y_{itrue} = 1$</th><td>0</td><td>1368 (100%)</td><td>0</td></tr> <tr> <th>$Y_{itrue} = 2$</th><td>0</td><td>95 (97.93%)</td><td>2 (2.07%)</td></tr> </table>		$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$	$Y_{itrue} = 0$	1794 (100%)	0	0	$Y_{itrue} = 1$	0	1368 (100%)	0	$Y_{itrue} = 2$	0	95 (97.93%)	2 (2.07%)	(d)	Bi-modal model (sorting) 3-class classification <table> <tr> <th></th><th>$Y_{ipred} = 0$</th><th>$Y_{ipred} = 1$</th><th>$Y_{ipred} = 2$</th></tr> <tr> <th>$Y_{itrue} = 0$</th><td>1794 (100%)</td><td>0</td><td>0</td></tr> <tr> <th>$Y_{itrue} = 1$</th><td>8 (0.59%)</td><td>1359 (99.34%)</td><td>1 (0.07%)</td></tr> <tr> <th>$Y_{itrue} = 2$</th><td>0</td><td>88 (90.7%)</td><td>9 (9.3%)</td></tr> </table>		$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$	$Y_{itrue} = 0$	1794 (100%)	0	0	$Y_{itrue} = 1$	8 (0.59%)	1359 (99.34%)	1 (0.07%)	$Y_{itrue} = 2$	0	88 (90.7%)	9 (9.3%)
	$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$																																
$Y_{itrue} = 0$	1794 (100%)	0	0																																
$Y_{itrue} = 1$	0	1368 (100%)	0																																
$Y_{itrue} = 2$	0	95 (97.93%)	2 (2.07%)																																
	$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$																																
$Y_{itrue} = 0$	1794 (100%)	0	0																																
$Y_{itrue} = 1$	8 (0.59%)	1359 (99.34%)	1 (0.07%)																																
$Y_{itrue} = 2$	0	88 (90.7%)	9 (9.3%)																																
(e)	Multimodal model (concatenation) 3-class classification <table> <tr> <th></th><th>$Y_{ipred} = 0$</th><th>$Y_{ipred} = 1$</th><th>$Y_{ipred} = 2$</th></tr> <tr> <th>$Y_{itrue} = 0$</th><td>1794 (100%)</td><td>0</td><td>0</td></tr> <tr> <th>$Y_{itrue} = 1$</th><td>0</td><td>1367 (99.93%)</td><td>1 (0.07%)</td></tr> <tr> <th>$Y_{itrue} = 2$</th><td>0</td><td>32 (32.9%)</td><td>65 (67.01%)</td></tr> </table>		$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$	$Y_{itrue} = 0$	1794 (100%)	0	0	$Y_{itrue} = 1$	0	1367 (99.93%)	1 (0.07%)	$Y_{itrue} = 2$	0	32 (32.9%)	65 (67.01%)	(f)	Multimodal model (sorting) 3-class classification <table> <tr> <th></th><th>$Y_{ipred} = 0$</th><th>$Y_{ipred} = 1$</th><th>$Y_{ipred} = 2$</th></tr> <tr> <th>$Y_{itrue} = 0$</th><td>1794 (100%)</td><td>0</td><td>0</td></tr> <tr> <th>$Y_{itrue} = 1$</th><td>45 (3.28%)</td><td>1296 (96.71%)</td><td>27 (1.98%)</td></tr> <tr> <th>$Y_{itrue} = 2$</th><td>0</td><td>63 (64.95%)</td><td>34 (35.05%)</td></tr> </table>		$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$	$Y_{itrue} = 0$	1794 (100%)	0	0	$Y_{itrue} = 1$	45 (3.28%)	1296 (96.71%)	27 (1.98%)	$Y_{itrue} = 2$	0	63 (64.95%)	34 (35.05%)
	$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$																																
$Y_{itrue} = 0$	1794 (100%)	0	0																																
$Y_{itrue} = 1$	0	1367 (99.93%)	1 (0.07%)																																
$Y_{itrue} = 2$	0	32 (32.9%)	65 (67.01%)																																
	$Y_{ipred} = 0$	$Y_{ipred} = 1$	$Y_{ipred} = 2$																																
$Y_{itrue} = 0$	1794 (100%)	0	0																																
$Y_{itrue} = 1$	45 (3.28%)	1296 (96.71%)	27 (1.98%)																																
$Y_{itrue} = 2$	0	63 (64.95%)	34 (35.05%)																																

[그림 52] 유니/바이/멀티모달 방식에 따른 학습 모델의 오차행렬 비교

- (a) 눈 이미지를 CNN을 이용하여 학습한 유니모달 기반의 인식 모델,
- (b) 입 이미지를 CNN을 이용하여 학습한 유니모달 기반의 인식 모델
- (c) 눈과 입 이미지를 바이모달 학습한 인식 모델(JR: 결합)
- (d) 눈과 입 이미지를 바이모달 학습한 인식 모델(JRL 정렬)
- (e) 눈, 입, PPG를 멀티모달 학습한 인식 모델(JR: 결합)
- (f) 눈, 입, PPG를 멀티모달 학습한 인식 모델(JR: 정렬)

- 모델 F1은 눈과 입 이미지를 동시에 학습하여 졸음을 분류하는 인식 모델로, 결합 방식의 Joint Representation을 사용하였을 때 조금 더 우수한 인식 정확도를 보였음. 반면 완전 연결 방식은 61% 수준으로 오히려 이 방식은 졸음을 결정하는 특징을 사라지게 한다고 볼 수 있음
- 모델 F2는 눈, 입, PPG를 모두 사용하여 학습한 졸음 인식 모델로 정렬 방식의 Joint Representation을 사용하였을 때 가장 좋은 인식 정확도를 보였음. F1과 마찬가지로 완전 연결 방식은 졸음을 분류하는 특징이 사라져 학습 모델이 졸음을 분류하지 못했다고 볼 수 있음. **본 결과에서 중요한 점은 PPG 단독으로는 학습이 전혀 이루어지지 않았으나([표 8]의 모델 C3), 눈, 입과 동시에 학습시킬 경우 유의미한 성능 향상이 이루어짐을 확인할 수 있었음**
- [그림 51]의 오차행렬(Confusion Matrix)은 도출한 인식 모델이 얼마나 올바르게 데이터를 분류 했는지 확인할 수 있는 방법임. 본 연구개발에서 수집한 데이터의 클래스는 불균형하기 때문에 단순 정확도 비교만으로는 각 클래스에 대한 분류가 올바르게 이루어 졌는지 가늠하기 어려움. [그림 46]은 유니/바이/멀티모달 방식으로 도출된 졸음 인식 모델의 3가지 클래스 분류 결과를 포함하고 있으며, 10번의 실험 중 가장 좋은 성능(best score)을 기록한 학습 모델에 대한 결과임
- [그림 51-(a)]와 [그림 51-(b)]는 유니모달 방식으로 도출된 졸음 인식 모델의 분류 결과로 정상 상태를 98% 이상 정상으로 분류하였음. 또한 얇은 잠 상태를 97%이상 얇은 잠으로 올바르게 분류하였음. 하지만, 깊은 잠을 대부분 얇은 잠으로 분류하는 경향이 있음. 특히, [그림 51-(b)]는 대부분의 깊은 잠을 얇은 잠으로 분류하였음
- [그림 51-(c)]와 [그림 51-(d)]는 바이모달(bimodal) 방식으로 도출된 졸음 인식 모델의 분류 결과로 Joint Representation 방식이 결합인지 정렬인지에 따라 구분되어 있음. 유니모달 방식과 비교하여 정상을 정상으로, 얇은 잠을 얇은 잠으로 더욱 잘 분류하는 경향이 있으나, 유니모달 방식과 마찬가지로 거의 대부분의 깊은 잠을 얇은 잠으로 분류하였음
- [그림 51-(e)]와 [그림 51-(f)]는 멀티모달 방식으로 도출된 졸음 인식 모델의 분류 결과로 Joint Representation 방식이 결합인지 정렬인지에 따라 구분되어 있음. 바이모달 방식과 마찬가지로 정상을 정상으로, 얇은 잠을 얇은 잠으로 잘 구분하였음. 하지만 바이모달 방식과 다르게 깊은 잠을 보다 깊은 잠으로 잘 인식하는 경향을 확인하였음. **[표 9]의 평균 정확도에 따르면 Joint Representation의 정렬(Sorting) 방식이 평균적으로 더 좋은 성능을 보이는 것으로 나타나지만, 가장 좋은 모델을 도출한다**

는 전제로는 [그림 51-(e)]와 같이 결합(Concatenation) 방식이 더 우수한 결과를 낼 수 있음을 확인하였음(정확도 96.15%, best score)

- 본 오차행렬 결과에서 중요한 점은 PPG 데이터가 눈, 입과 함께 사용되어 동시에 학습이 이루어졌을 때 분류 성능에 긍정적인 영향을 미치는 것임. PPG 단독으로는 3가지 졸음 상태를 분류할 수 없지만, PPG가 추가됨으로써 눈, 입 만으로는 분류에 한계가 있는 얇은 잠을 더욱 잘 분류하였음

다. 연구결과

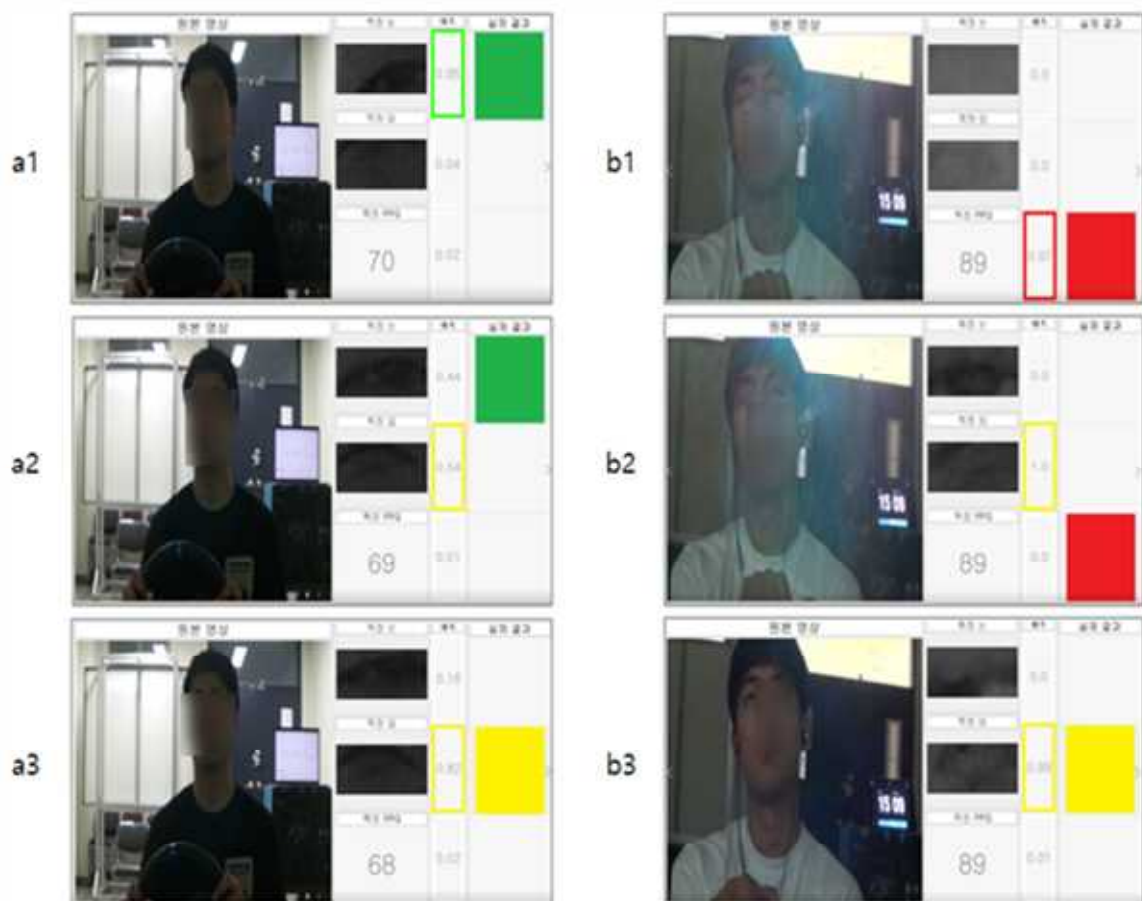
(1) 본 연구개발이 정의하는 5가지 운전자 상태의 기준

- 정상(Normal): 운전자는 깨어 있으며 심장 활동이 평소와 같이 안정된 상태
- 심장정지(Cardiac Arrest): 심장 박동이 멈춘 상태(운전자는 의식불명)
- 심장마비(Heart Attack): 심장 박동이 불규칙한 상태(극심한 가슴통증)
- 라이트슬립(Light Sleep): 의식은 있으나 금방이라도 졸음이 발생할 수 있는 상태
- 딥슬립(Deep Sleep): 졸음으로 인해 운전자의 의식이 사라진 상태

※ 2.-가.- (1)에 따라, 국내·외 연구개발에서는 다양한 형태로 졸음상태를 판별할 수 있는 기준을 제시하였음. 시각적인 측면에서 딥슬립은 눈이 완전히 감기거나 [그림 22]의 PERCLOS가 일정 % 이하인 경우로, 생리적인 측면에서 심박수가 평소보다 감소하는 상태를 기준으로 정의하고 있음. 하지만, 정상과 딥슬립 사이에 위치한 라이트슬립은 사람마다 편차가 있기 때문에 명확한 기준을 설정하기 어려움. 전통적인 기계학습에서는 졸음에 대한 기준(특징)을 연구자(사람)가 도출하여 인식 모델을 개발하는데, 라이트슬립을 결정하는 특징은 사람이 정의하기 매우 어려운 문제임. [그림 9]와 같이 딥러닝은 전통적인 기계학습과 다르게, 특징을 기계 스스로 추출하는 기능이 알고리즘에 내재되어 있는 것이 기계학습과 다른 차별성임. 즉, 사람은 어떠한 데이터가 라이트슬립인지(정답)를 기계에게 제공하면 기계는 스스로 가장 적절한 특징을 추출함. 또한, 사람이 정답이 있는 데이터(학습 데이터)를 충분히 큰 규모로 제공하면 할수록 그 특징 역시 점점 일반화된 특징에 가까워지는 장점이 있음(반면 기계학습은 일반적으로 학습 데이터가 많아져도 인식 성능증가에 한계가 있음). **본 연구개발이 정의하는 라이트슬립은 가상주행환경의 연구참여자를 통해 수집한 학습 데이터에서 기계가 스스로 추출한 기준(특징)을 의미하며, 이 기준이 실제주행환경의 연구참여자를 통해 수집한 테스트 데이터에도 유효하다는 것을 실험을 통해 보였음.** 만약, 본 연구개발의 인식 모델 구조를 그대로 유지하면서 학습 데이터만 달라진다면(다른 연구참여자), 본 연구팀과 다른(높거나 낮은) 인식 성능이 도출될 가능성이 있으나 이것은 모집단 표본에 대한 차이로 인하여 발생하는 것으로, 본 연구의 모집단은 그 규모가 너무 크기 때문에 이 차이를 객관적인 수치로 제시하기 위해서는 충분한 표본이 확보(추가적인 연구참여자 필요)될 필요가 있음

(2) 운전자 상태 인식 시스템 테스트 결과 (가상 주행환경의 데이터 사용)

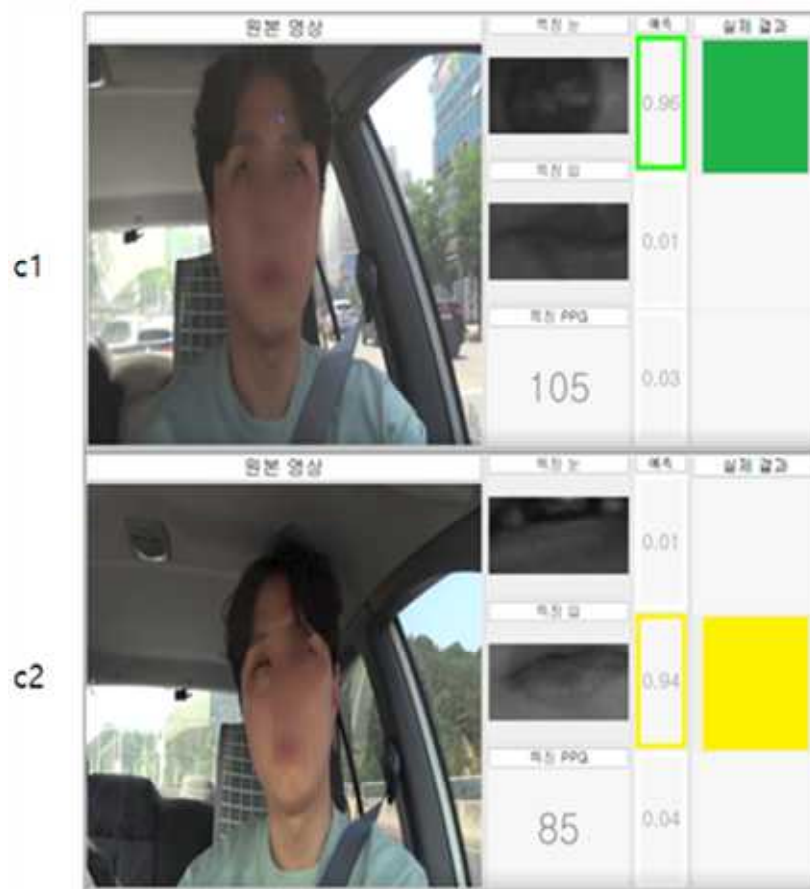
- 상태 인식 시스템은 도출한 인식 모델이 탑재되어 있으며, 원본 데이터에서 특징 값(눈, 입, PPG)을 추출하고 인식 모델을 통해 예측한 결과를 출력함(예측한 결과를 비교하기 위해 가장 오른쪽에 원본 데이터의 실제 결과가 함께 표시됨)
- [그림 52-(좌)]는 연구참여자가 a1(정상)에서 a3(라이트슬립)로 향하는 순간을 담고 있음. 인식 모델은 a1을 95% 정상(초록)으로 예측했고, a3를 82% 라이트슬립(노랑)이라고 예측하였음
- [그림 52-(우)]는 연구참여자가 b1(딤슬립)에서 b3(라이트슬립)로 향하는 순간을 담고 있음. 인식 모델은 b1을 97% 딤슬립(빨강)으로 예측했고, b3를 99% 라이트슬립이라고 예측하였음



[그림 53] 멀티모달 학습 기반의 인식 모델이 탑재된 프로그램의 가상 주행환경에서의 졸음 인식
(좌: Normal→Light Sleep 예측, 우: Deep Sleep→Light Sleep 예측)

(3) 운전자 상태 인식 시스템 테스트 결과 (실제 주행환경의 데이터 사용)

- 상태 인식 시스템의 동작은 2.-다.-(2)와 동일
- [그림 53]은 연구참여자가 c1(정상)인 상태를, c2(라이트슬립) 상태에서, 인식 모델은 c1을 96% 정상(초록)으로 예측했고 c2를 94% 라이트슬립(노랑)으로 예측하였음
- 다만, 본 연구팀은 실제 주행환경의 딥슬립(Deep Sleep) 상태의 데이터를 수집 대상에서 제외하였기 때문에, 본 상태 인식 프로그램을 통해 결과를 제시할 수 없었으나, [표 9]의 결과와, 가상 주행환경의 테스트 결과로 비추어 볼 때 딥슬립 상태의 운전자도 충분히 인식 할 수 있을 것으로 판단됨
- 또한, 운전자의 졸음은 의식이 사라진 상태(딥슬립)보다 졸음이 밀려오는 상태(라이트슬립)를 인식하여 적절한 각성을 유도하는 것이 실제 적용에 유효함



[그림 54] 멀티모달 학습 기반의 인식 모델이 탑재된 프로그램의 실제 주행환경에서의 졸음 인식 (Normal, Light Sleep 예측)

(4) 개발한 인식 모델의 인식 속도 측정

- 자동차 제조사가 빌트인(built-in)으로 탑재한 인포테인먼트 시스템의 하드웨어 사양은 제조사 정책에 종속적임. 따라서, NAVER의 AWAY와 같이 특정 자동차에 제한되지 않는 단말 형태의 인포테인먼트 시스템을 기준 하드웨어 사양으로 선정하였음
- [표 10]에 따라 보급형 노트북 정도의 하드웨어 사양은 단말 형태의 인포테인먼트 시스템의 사양과 큰 차이가 없음. 단, 메모리와 디스크의 차이는 인식 모델의 인식 속도에 영향을 거의 미치지 않고, 오히려 GPU에 큰 영향을 받기 때문에 GPU가 탑재된 인포테인먼트 시스템이 유리한 측면이 있음
- 본 연구개발팀은 [그림 54]와 같이 보급형 노트북에 개발한 인식 모델을 탑재하여 인식 속도를 측정하였으며, 평균 2.5초 이하로 인식이 가능함을 보였음

[표 10] 인포테인먼트 시스템과 보급형 노트북의 하드웨어 사양 비교

	보급형 노트북	NAVER AWAY (단말 형태의 인포테인먼트 시스템)
운영체제	Windows 또는 Linux	Android 5.1
CPU	Intel Core i7 1.8GHz	Media Teck Quard-core 1.7GHz
RAM	DDR4 16GB	DDR3L 1GB
SSD	256GB	eMMC 16GB
GPU	-	Mali 450 MP4 600MHz
통신	Bluetooth, USB, Wi-Fi	Bluetooth, USB, LTE
카메라	HD 해상도	HD 해상도



[그림 55] 운전자 상태 인식 모델은 보급형 노트북에 탑재하여 성능 측정

(5) 본 연구개발 결과의 학술적 차별성

[표 11] 본 연구의 인식모델과 관련연구 비교

관련 연구	학습 데이터	연구 참여자 수	학습 모델	상태 인식	인식 성능
[64]	SWA (Steering Wheel Angle)	6	Binary Decision Classifier (기계학습)	정상, 졸음	78.01%
[65]	PERCLOS (Percentage of Eye Closure)	-	Cascade Classifier, AdaBoost (기계학습)	정상, 졸음	85%
[66]	뇌파, 표정 벡터 (침습적 방법)	15	Support Vector Machine (기계학습)	정상, 졸음	88.62%
[67]	HRV (Heart Rate Variability, 침습적 방법)	12	Neural Network (딥러닝)	정상, 졸음	90%
[68]	얼굴 이미지	21	CNN (딥러닝)	정상, 졸음	92.33%
[69]	입 + 눈 (이미지)	22	Deep Neural Network (멀티모달 딥러닝)	정상, 하품, 졸음	94.8%
본연구	입 + 눈 + PPG	14	CNN+RNN (멀티모달 딥러닝)	정상, 라이트슬립, 딥슬립	96.15%

- 기존의 접근방식 및 기계학습 기반 모델의 한계점
 - 운전대 모니터링[64]과 같이 간접적으로 운전자의 상태를 인식하는 방법은 졸음과 유사한 패턴과 실제 졸음으로 인한 패턴을 구분하지 못하는 잘못된 인식률이 높음
 - 기존의 기계학습 기반의 인식모델[65, 66]들은 사람(연구팀)이 도출하고 정의한 특징을 기반으로 상태를 인식하기 때문에 일반화가 어려워 90% 이상의 인식 정확도를 도출하기 어려움. 또한, 추가적인 학습 데이터가 제공되어도 정의한 특징을 그대로 사용하기 때문에 인식 성능 개선이 불가능함
- 본 연구를 통해 개발한 인식 모델의 차별성
 - 본 연구 및 [67-69] 같이 딥러닝을 기반으로 하는 모델은 사람(연구팀)이 도출하기 어려운 고차원적인 특징 추출 기능이 알고리즘에 포함되어 있고, 추가되는 학습 데이

터에 상응하여 유기적으로 다시 적절한 특징을 찾아가기 때문에 기존의 기계학습 보다 일반화된 인식 모델을 도출하여 인식 정확도가 높은 장점이 있음

- 하지만, 딥러닝 기반의 인식 모델을 개발하기 위해서는 충분한 규모의 학습 데이터가 뒷받침 되어야 함. 이 학습 데이터는 데이터와 레이블(정답)이 쌍으로 존재하며 모두 인간이 직접 레이블을 지정해야하는 시간과 인력이 크게 소모되는 작업임. 또한, 학습 데이터가 늘어나면 그에 따라 학습 모델을 도출하는데 소요되는 시간도 증가하기 때문에 적절한 수준의 학습 데이터를 확보하고 모델의 복잡도를 가급적 줄이는 것이 가장 이상적인 딥러닝 모델 도출 방식임
- 본 연구는 목표로 하는 **상태인식과 인과관계가 높은 데이터를 동시에 활용하여, 딥러닝에 필수적인 대규모 학습 데이터 확보에 대한 부담을 줄이면서(상대적으로 적은 학습 데이터사용) 정확도는 높은 인식 모델을 개발하였음. 또한, 제한적이지만 실제 주행환경에서 인식이 가능한 수준의 모델임을 보였음**
- 본 연구와 유사한 [69]는 삼성전자 소속의 연구팀이 수행한 연구개발로 약 70,000장의 이미지를 학습한 인식 모델이며 실제 데이터에 대한 검증은 없었음. 본 연구는 25,227개의 멀티모달 데이터를 사용하여 우수한 인식 성능을 보였음
- 추가적으로, 국내·외적으로 인공지능 시스템 개발을 위한 양질의 그리고 대규모의 학습용 데이터가 요구되는 시점에서, 본 연구팀이 수집한 멀티모달 데이터는 희소성과 가치가 높은 데이터로 볼 수 있음

(6) 본 연구개발 결과의 기술적 차별성

[표 12] 본 연구개발의 유사 시스템

STEER: 클라우드 펀딩 플랫폼 Kick Starter의 졸음운전 방지용 스마트 밴드	LS 803: Leisure Auto에서 개발한 졸음운전 방지용 설치형 카메라
 The image shows a person's arm wearing a black smart band. To the right, a car's dashboard display shows a driving scene with a yellow arrow pointing to the right and a red dot indicating a warning area.	 The image shows a black camera mounted on the dashboard of a car. In the background, a driver is visible behind the steering wheel.
• 원리: 심박수와 피부전기반응을 측정하여 낮아질 경우 졸음으로 판별 • 기능: 졸음으로 판별 시 진동 및 전기충격을 통한 각성 유도 • 가격: 약 200유로 • 한계점: 피부전기반응은 습도에 따라 오차가 발생하여 날씨와 지역에 따라 잘못된 판별 가능성이 높음 • 한계점: 심박수는 나이, 성별, 건강상태 등 사람마다 편차가 커서 일반화가 매우 어려운 지표임	• 원리: 운전자 얼굴을 이미지 처리 하여 눈이 감긴 정도와 얼굴 각도로 판별 • 기능: 졸음으로 판별될 경우 경고음을 발생하여 각성을 유도 • 가격: 약 150달러 • 한계점: 이미 의식이 사라진 상태를 감지하기 때문에 실제 졸음운전 예방 효과가 미미할 수 있음 • 한계점: 이미지 처리를 기반으로 하기 때문에(졸고 있는 듯한 연기를 해도 졸음으로 인식), 조도, 표정 등에 따라 잘못된 판별이 가능함

• 본 연구를 통해 개발한 인식 시스템의 차별성

- 기술적 우수성: 보다 정확하고 안정적인 상태 인식 시스템
 - ✓ 기존 제품 및 기술보다 정확한 상태 인식 정확도
 - ✓ 표준 통신 미들웨어를 기반으로 통신 안정성 보장
- 경제적 우수성: 0~10만원 이하로 시스템 구축 가능
 - ✓ 기존 인포테인먼트 시스템에 인식 모델 탑재 (추가비용 X)
 - ✓ 오픈소스 기반의 통신 미들웨어 사용 (추가비용 X)
 - ✓ 기존 사용하던 스마트 워치/밴드를 연결하여 사용 (추가비용 0~5만 이내)

- ✓ 인포테인먼트 시스템에 탑재된 카메라 사용 또는 장착 (추가비용 0~5만원 이내)
- 예상되는 기능
 - ✓ 졸음이 밀려오는 상태의 운전자에게 알림 (의식이 사라지기 전에 경고)
 - ✓ 연결된 스마트 워치를 통해 위급 상황에 대한 긴급 메시지/호출 가능
 - ✓ + 인포테인먼트와 연계된 운전자 각성 유도 (예: 음악변경, 씬터안내 등)
 - ✓ + ADAS와 연계되어 자동 정차 및 eCall 호출



[그림 58] 본 연구개발을 통해 구축된
운전자 상태 인식 시스템 개념도

라. 연구개발 성과

(1) 일반 학술지 게재 2편

	논문 명	논문지	권호	게재일
1	멀티 모달 학습을 사용한 시각적 특징 기반의 운전자 졸음 감지	한국통신학회	제43권 제7호	2018.07.
2	이미지와 PPG 데이터를 사용한 멀티모달 딥러닝 기반의 운전자 졸음 감지 모델	데이터베이스연구	제34권 제3호	2018.12.

(2) 특허 등록 1건, 특허 출원 2건

	구분	출원/등록 번호	발명의 명칭	출원/등록일	기여율
1	등록	10-1911359	저전력 블루투스 어댑터 및 그를 이용한 데이터 교환 시스템	2018.10.18.	1/1
2	출원	10-2017-0124497	멀티모달 학습 기반의 운전자 상태 인식 시스템	2017.09.26.	1/1
3	출원	10-2018-0159741	이미지와 PPG 데이터를 이용한 멀티모달 딥러닝 기반의 운전자 졸음 감지 시스템	2018.12.12.	1/1

(3) 소프트웨어 등록 4건

	등록 번호	저작물의 명칭	등록일	기여율
1	제 C-2017-036219 호	운전자 졸음 감지용 이미지 특징 추출 프로그램	2017.12.26.	1/1
2	제 C-2018-038651 호	얼굴 이미지 또는 심박 데이터를 이용한 딥러닝 기반의 졸음 감지 모델 생성 프로그램	2018.12.18.	1/1
3	제 C-2018-038652 호	눈 이미지와 입 이미지를 동시에 학습하여 졸음을 감지하는 학습 모델 생성 프로그램	2018.12.18.	1/1
4	제 C-2018-038653 호	얼굴 이미지와 심박 데이터를 멀티모달 학습하여 졸음 감지 모델을 생성하는 프로그램	2018.12.18.	1/1

(4) 시제품 1건: 운전자 상태 인식 시스템([그림 54], [그림 55], [그림 56] 참조)

(5) 기타(학회발표 2편)

	논문 명	학회명	게재일
1	시각적 특징 기반의 운전자 졸음 인식을 향상하기 위한 멀티 모달 학습 기반의 시스템 설계	2017년도 한국통신학회 추계종합학술발표회	2017.11.
2	Driver Drowsiness Detection based on Multimodal using Fusion of Visual-feature and Bio-Signal	The 9 th International Conference of ICT Convergence	2018.10.

3. 목표 달성도 및 관련 분야 기여도

가. 목표 달성도

(1) 양적 성과지표의 목표 달성도

성 과 지 표	성 과 목 표	성 과	달 성 도 (%)
SCI(E)급 학술지 게재(2점)	1	(1)	0
일반학술지게재(1점)	2	2	100
특허출원(1점)	2	2	100
특허등록(2점)	1	1	100
소프트웨어(1점)	3	4	133
시제품(1점)	1	1	100
기타(1점)	-	2	-
소계(점수)	12	11(1)+2	108.33

- SCI(E)급 학술지 게재 성과지표의 경우, 최종 보고서 제출 시점인 2019년 3월 21일 기준 SCIE급 학술지인 Journal of Advanced Transportation의 특집호 New Trends on Modelling and Simulation of Electric Mobility for Public and Private Transportation System에 1편의 논문을 제출하여 심사중에 있으며, 해당 특집호는 2019년 6월에 발간 절차가 진행될 예정임
- 소프트웨어 성과지표의 경우 당초 목표 3건에서 1건을 추가적으로 등록하였음
- 기타 성과지표의 경우 당초 성과목표를 설정하지 않았으나, 논문 게재 절차에 따라 2편의 학술발표 성과가 발생하였음(국내 1편, 국제 1편)

(2) 질적 성과지표의 목표 달성도

구분	핵심성과 (CTE)		단위성과		성과점검기준						
					질적 성과지표	목표치	달성치	측정방 법	검증방법	기준치	
주관	1	학습용 데이터 수집 및 멀티모달 인터페이스 개발 (1차년도)	1-1	학습용 데이터 수집 (실제/ 가상 주행환경 에서의 운전자 데이터 수집 및 분석)	①	PPG 데이터 수집 (CAN, 블루투 스를 통해 수집)	딥러닝 기반의 학습모델 개발이 충분한 수준	가상주행 환경 2,808,000개	수집된 데이터 를 이용해 인식 모델을 도출하 고 정확도 측정	CNN, RNN 등의 딥러닝 기반의 인식 모델이 90% 이상의 인식 정확도를 보이는지 여부	25
					③	이미지 데이터 수집 (USB 를 통해 수집)		가상주행 환경 8,500,000장			실제주행 환경 396,000개
	2	멀티모달 학습 기반 운전자 상태 인식 시스템 개발 (2차년도)	2-1	멀티모달 학습 기반의 인식 모델 도출(5 가지 상태 분 류) 및 관련 S W 개발	①	실시간 및 신뢰성 보장	2가지 QoS	DDS표준을 준수하여 De adline, Relia ble QoS 지원	표준 준수	보고서, 시제품, 등록특허	5
					②	5가지 상태 분류 정확성	90%	평균 92.80% 최고 96.15%	K-Fold Cross V alidation	보고서, 시제품, 논문	40
					③	상태 분류 처리속 도	5초 이내	인포테인먼트 시스템 사양에서 2.5초 이하	인식모 델의 인식 속도 측정	보고서, 시제품, SW등록	5
	계								100%		

- 당초 계획된 연구개발의 질적 성과지표를 100% 달성하였음. 특히, 상태 인식의 모델의 경우 당초 목표치인 90%를 넘었음. 본 연구팀이 설계한 인식 모델을 10-Fold Cross Validation을 통해 10개의 인식 모델을 생성할 경우 평균적으로 92.80%의 인식 정확도를 보이며 최고 성능의 모델은 96.15%의 인식 정확도를 보였음
- 학습용 데이터 수집의 경우, 연역적으로 규모를 추정(가설)하고 실제 딥러닝 기반의 모델을 만들어 적용해 봄으로써 추정된 규모가 적절한지를 판단하였음(단, 본 연구개발이 수집한 데이터 규모가 운전자 상태 인식 문제를 해결하는데 가장 이상적인 규모를 의미하는 것은 아님)
- 상태 분류 처리 속도의 경우, 인포테인먼트 시스템의 하드웨어 사양에서 본 연구개발의 상태 인식 시스템이 실시간 및 신뢰성 보장 QoS를 지원하며, 운전자의 상태를 2.5초 이내로 인식할 수 있음을 보였음

나. 목표 미달성 원인 및 차후대책

(1) SCI(E)급 학술지 게재 미달성 원인 및 차후대책

- 총 연구개발의 기간(1년 6개월) 중 학습용 데이터를 수집하기 위한 노력이 많은 비중을 차지하였음. 본 연구개발의 멀티모달 딥러닝을 적용하기 위해서는 동일한 시점에 발생한 두 가지 종류 이상의 데이터가 쌍으로 존재하는 학습용 데이터 확보가 필수임
- 이러한 특수한 학습용 데이터를 보유한 국내·외 기관 및 연구팀은 없었으며, 그에 따라 본 연구팀이 직접 데이터를 수집하였음. 수집한 원천 데이터는 각 연구원들이 수작업으로 레이블링이 이루어졌고, 데이터를 제공한 연구참여자들의 레이블링 검증 또한 병행되는 시간과 인력이 많이 필요한 과정이었음
- 그에 따라, 연구개발 종료가 임박하여 관련연구 이상의 성능을 보이는 학습 모델이 도출되었음. 실험에 따라 성능이 개선되는 시점에 우선적으로 국내 일반학술지에 연구내용을 발표하여 일반학술지게재의 목표는 달성되었으나, SCI(E)급 학술지에 게재를 위해서는 보다 면밀한 실험 및 검증과정이 필요하다고 판단하였음
- 본 미달성 목표를 달성하기 위하여 연구종료 직후 2019년 1월에 Journal of Advanced Transportation에 논문 1편을 제출하여 심사중에 있음. 심사중에 있는 논문과는 별개로 Computational Intelligence and Neuroscience와 같은 상대적으로 심사기간이 짧고 (60일 내외), 게재가 확정된 논문은 발간까지 40일 내 이루어지는 SCI(E)급 온라인 학술지에 추가적인 논문 제출을 진행하고 있음

다. 관련 분야의 기여도

(1) 기술적 측면

- 현재의 완전 자율주행 자동차는 자동차 외부의 환경정보를 자동으로 인식하여, 인간의 개입 없이 안전하게 목적지에 도착하는데 초점이 맞추어져 있음. 하지만, 인간 중심의 미래 자율주행 자동차라면 탑승자의 생체적 이상을 인식하여, 보다 인간을 안전하게 할 수 있어야 함. 본 연구개발은 현재 상대적으로 주목받지 못하는 운전자 상태 인식 기술에 대한 기술적 성숙도를 높여 인간 중심의 미래 자율주행 자동차 상용화를 앞당기는데 기여할 수 있을 것임
- 자동차 산업의 성장 정체, 소비자 요구의 다양화 등으로 글로벌 자동차 시장은 자율주행 자동차로 패러다임이 전환되고 있어 국가 차원의 미래 자동차 분야의 기술역량 확보가 시급함. 특히, 미래 자동차 산업을 바탕으로 타 산업과 융합 및 협력체계를 구축하면, 새로운 고부가가치의 신사업 및 서비스 발굴이 가능함. 본 연구개발은 의료과학, 국방과학 등의 인간의 생명과 밀접한 관련이 있는 분야와의 융합 구심점을 만들어 신사업을 창출하는데 기여할 수 있을 것임

(2) 경제·산업적 측면

- 헬스케어 산업은 2015년 기준 9조 1000억 달러에서 오는 2020년 11조 5000억 달러로 성장할 전망이다. 미국의 FDA는 2018년 CT영상을 분석하는 AI를 허가하는 등 AI를 이용한 의료진단 시장은 빠르게 성장하고 있음. 1인 가구 및 고령화 사회의 흐름에서 지능적인 자동차 헬스케어 기술은 운전자의 건강 지킴이 역할을 수행할 것으로 기대함
- 국가 차원의 전략투자 분야로 AI가 선정됨에 따라 AI 허브에 대한 투자규모가 2019년 310억원으로 확대되었음. 실제 세계는 다양한 데이터가 실시간으로 생성되고 복합적인 요인에 따라 다양한 상황이 발생함. 본 연구개발의 멀티모달 기반의 AI는 보다 현실에 가까운 데이터를 사용하여 기존 AI 기술의 인식률 제고에 기여할 것으로 기대함

(3) 사회적 측면

- 우리나라의 심장질환에 의한 사망은 전체 사망 원인의 2순위이며 심폐소생술 시행률은 12%로 선진국의 절반 수준임. 게다가 급성 심정지의 절반은 가정 외에서 발생하며 자동차는 누구의 도움도 받을 수 없는 밀폐된 공간임. 국민 누구에게나 찾아올 수 있는 심장질환의 위험에서 안전 사각지대를 해소하는 연구개발이 될 것으로 기대함
- 지난 10년간 고속도로 교통사고 발생원인 1위는 졸음 운전에 의한 사고로 전체 교통사고의 22.5%를 차지함. 전체 운전자의 약 80%는 1회 이상 졸음운전을 경험했다고 응답했으며, 100km/h 속도로 3초를 졸면 90m 가량을 무의식으로 주행하는 매우 위험한 상황임. 본 연구개발은 잘못된 경보를 발생하지 않으면서 보다 졸음감지에 특화된 기반 기술을 제공하여 도로교통 안전 선진화에 기여할 것으로 기대함
- 국내 자동차 사고의 사회적 비용은 10조원에 달하며, 정량화할 수 없는 사회적 손실까지 고려하면 막대한 국가적 손실임. 게다가, 급속한 고령화 사회로 진입함에 따라 심장정지 등의 고령 운전자의 의도치 않은 자동차 사고가 증가하고 있는 추세임. 본 연구개발은 고령 운전자 뿐 아니라 장애인 등의 교통 약자에게 안전한 교통 서비스를 제공하여 평등한 이동 권익을 보호해 보편적인 삶의 질을 향상시키는데 기여할 것으로 기대함

4. 연구개발성과의 활용 계획 등

가. 추가 연구의 필요성

(1) 인공지능: 국가 차원의 전략적 투자가 필요한 혁신성장 전략투자 분야

- 정부는 2018년 8월 국가 차원의 전략투자 분야로 인공지능(AI, Artificial Intelligence)을 선정하였으며, AI 허브에 대한 투자 규모도 2018년 93억원에서 2019년 310억원으로 대폭 확대하였음
- AI 허브는 연구자, 중소·벤처기업, 정부부처 등에서 인공지능 기술을 적용하기 위한 기본적인 학습 데이터를 상호 공유하는 플랫폼으로, 국내 인공지능 산업의 경쟁력을 키우기 위해 다양한 분야의 연구자들이 데이터 제약 없이 제품과 서비스를 개발할 수 있는 중추 역할을 수행할 것으로 기대됨
- 본 연구개발은 자동차 산업과 헬스케어 분야에 인공지능 시스템이 적용된 제품 및 서비스 개발에 필요한 매우 가치 있는 데이터이며, 멀티모달 이라는 특수성을 가진 희소성이 높은 데이터를 포함하고 있음

(2) 추가 연구의 필요성: 실 환경 기반의 후속 연구를 통한 신뢰성 확보

- 딥러닝 관련 연구개발의 특성상 양질의 학습 데이터 수집 및 확보는 해당 과제의 결과와 직결되는 매우 기본적인 과정임. 하지만, 이 과정은 사람이 직접 수행하기 때문에 많은 시간과 인력을 필요로 함
- 본 연구개발을 통해 딥러닝 적용이 가능한 규모의 운전자 상태인식을 위한 학습 데이터가 마련되었음. 이 학습 데이터는 [표 8], [표 9]와 같은 결과를 보여, 본 연구팀이 수집한 학습용 데이터는 기존 연구들과 동등한 수준의 규모와 질적 수준을 가졌다고 볼 수 있음. 특히, 본 연구팀이 수집한 데이터는 미래 자동차 분야의 인공지능 시스템을 고도화 하는데 필수적으로 사용되는 데이터로, 국가 산업발전 및 공익의 목적으로 가치가 매우 높은 학습 데이터임(특히 대중교통 부문)
- 2.-(다)의 연구결과와 같이, 본 연구를 통해 개발한 운전자 상태 인식 모델은 기존의 인식 모델보다 더 적은 규모의 학습데이터를 사용하면서도 더 우수한 인식 성능을 보

였음

- 하지만, 본 연구를 통해 도출한 멀티모달 학습 기반의 상태 인식 시스템은 매우 제한적인 실제 주행 환경에서 테스트가 이루어졌음. 따라서, 다양한 실제 주행환경에 대하여 테스트하고 이를 반영하여 보다 신뢰성 높은 인식 시스템을 개발할 필요가 있음
- 본 연구개발의 후속 연구가 지원된다면 상대적으로 많은 시간과 인력이 필요한 학습용 데이터 확보 절차 없이, 보다 높은 수준의 운전자 상태 인식 모델을 도출하는데 더 많은 시간과 자원을 투입할 수 있을 것임. 대표적으로, 본 연구개발을 통해 확보된 데이터를 이용하여 GAN(Generative Adversarial Networks)과 같은 자동화된 학습 데이터 생성 기법을 적용한다면 사람의 개입 없이 추가적인 학습 데이터를 확보할 수 있으며, 이는 인식 모델 일반화에 기여할 수 있음. 또한, 실제 주행환경을 다양화하여 개발한 인식 모델의 성능의 신뢰성을 확보할 수 있을 것으로 기대함

나. 타 연구에 응용

(1) eCall/V2V 서비스 호출 지능화에 활용

- 현재의 eCall 서비스는 운전자가 호출 버튼을 직접 누르거나 에어백 작동이 감지되면 서비스를 호출하는 방식임. 하지만 자동차에 직접적인 손상 없이도 운전자가 사망에 이르는 운전자의 심장질환을 인식하지 못하면, 적시에 eCall 서비스를 호출할 수 없는 치명적인 문제가 있음
- 본 연구개발성과는 다음과 같은 지능적인 eCall/V2V 호출이 가능함. 심장정지의 경우 V2V 서비스를 통해 주변 차량에 구조를 요청하여 운전자의 골든 타임을 보장할 수 있으며, 심장마비의 경우 eCall 서비스 호출을 통해 구급요원들이 운전자를 신속히 응급센터로 이송해 1시간 안에 수술이 이루어지게 할 수 있음

(2) ADAS와 연계한 졸음운전 사고 방지 기술에 활용

- 차량 내 졸음운전 감시 및 경고장치는 졸음운전이라고 판단하여 경고하는 “잘못된 경보(false alarm)”의 비율이 높다는 문제점이 있음. 졸음 경보 시스템은 화물차 기사와 같이 차량을 정차하지 않고자 하는 운전자들이 주로 사용하고 있으며, 잘못된 경보가 반복되면 운전자로 하여금 잘못된 경계 수준을 설정하도록 하여(경고장치 작동에 둔감해짐) 결과적으로 사고예방에 도움을 주기 어려운 문제가 있음
- 본 연구개발성과는 3단계의 졸음상태를 정확하게 감지하여 ADAS와 연계한 효과적인 졸음운전 방지 기술개발로 이어질 수 있음. 예를 들면, 졸음이 밀려오면 인포테인먼트 앱을 실행해 운전자를 각성시키거나, 얇은 잠에 빠지면 자동으로 갓길에 정차하여 사고예방에 도움을 줄 수 있음

(3) 연구개발 결과물을 확장하여 철도, 선박, 항공기 분야의 보조 시스템에 활용

- 철도, 선박, 항공기 등을 통해 여객·화물을 운송하는 분야는 개인 자동차 운행에 비해 사고발생 시 인명과 재산피해가 막대하기 때문에 조종사(또는 기관사)를 보조하는 시스템이 존재함(예: 관제소, 자동 항법, 자동 착륙, 부기장, 승무원 등)

- 본 연구개발로 도출되는 상태 인식 시스템은 수집되는 데이터의 종류(모달리티)만 추가·변경하면, 다른 수송 분야의 조종사를 보조하는 시스템으로 확장될 수 있고, 국방 과학 분야의 실시간·신뢰성 표준을 지원하기 때문에 확장된 보조 시스템의 실용성을 높일 수 있음

붙임. 참고문헌

1. eCall, European Commission, https://ec.europa.eu/transport/road_safety/specialist/knowledge/esave/esafety_measures_unknown_safety_effects/ecall_en
2. 양상운, “차량 ICT 기반 긴급구난체계 (e-Call) 서비스 및 기술 동향”, TTA Journal, 2016.
3. Rezaei, Mahdi. Computer Vision for Road Safety: A System for Simultaneous Monitoring of Driver Behaviour and Road Hazards. Diss. ResearchSpace@ Auckland, 2014.
4. 통계청, 2017년 사망원인통계. 2018.
5. YTN, 심성지로 전봇대 ‘꽂’...운전자 의식불명, https://www.ytn.co.kr/_ln/0103_201702080932245065, 2017.
6. Haupt CM, Alte D, Dorr M, Robinson DM, Felix SB, John U, et al. The relation of exposure to shift work with atherosclerosis and myocardial infarction in a general population. *Atherosclerosis* 2008;201:205-11.
7. Bigert C, Gustavsson P, Hallqvist J, Hogstedt C, Lewne M, Plato N, et al. Myocardial infarction among professional drivers. *Epidemiology* 2003; 14:333-9.
8. Jang CH, Choi YA, Lee HM, Kang YS, Kim CH, Noh SY, et al. A study on the relationship between obesity and life style among cab drivers in Gyeongju. *Dongguk J Med* 2004;11:62-73.
9. 정혜선, 엄미정, 장인순, “버스운수업 종사자의 뇌심혈관질환 예방을 위한 개별상담 효과: 이차자료분석”, *Korean Journal of Occupational Health Nursing*, 2015.
10. YTN, 고속도로에서 8중 추돌사고...“또 졸음운전”, https://www.ytn.co.kr/_ln/0115_201709022214010535, 2017.
11. 이원영, 오주석, “졸음운전 방지를 위한 대책에 관한 연구”, 도로교통공단 연구보고서, 2014.
12. Choudhary, Prakash, et al. "A Survey Paper On Drowsiness Detection & Alarm System for Drivers." *International Research Journal of Engineering and Technology (IRJET)*. 2016.
13. American Heart Association, "Heart Attack or Sudden Cardiac Arrest: How Are They Different?", http://www.heart.org/HEARTORG/Conditions/More/MyHeartandStrokeNews/Heart-Attack-or-Sudden-Cardiac-Arrest-How-Are-They-Different_UCM_440804_Article.jsp#.WK6-aTuLRG
14. Shin, Heung-Sub, et al. "Real time car driver's condition monitoring system", *IEEE Sensors*, 2010.
15. Singh, Rahul Kumar, et al. "A health monitoring system using multiple non-contact ECG sensors for automotive drivers", *IEEE Instrumentation and Measurement Technology Conference Proceedings*, 2016.
16. Singh, Rahul Kumar, et al. "A real-time heart-rate monitor using non-contact electrocardiogram for automotive drivers", *IEEE Control, Measurement and Instrumentation*, 2016.
17. plessey, WARDEN, <http://www.plesseysemiconductors.com/products/warden/>
18. Ford, The Ford heart rate monitoring seat, http://www.medtees.com/content/ecg_seat_fact_sheet_2.pdf
19. 현대자동차그룹, 현대자동차그룹 선행기술로 살펴보는 인간 중심 기술의 미래, <https://blog.hmgjournal.com/Tech/reissue-advanced-technology.blg>
20. Umadi, L. F., S. N. A. M. Azam, and K. A. Sidek. "Heart Abnormality Detection Technique using PPG Signal." *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 8.12 (2016): 73-77.
21. Ukil, Arijit, et al. "Cardiac condition monitoring through photoplethysmogram signal denoising using wearables: can we detect coronary artery disease with higher perfor

- mance efficacy?." 2016 Computing in Cardiology Conference (CinC). IEEE, 2016.
22. Moraes, Jermana, et al. "Advances in photoplethysmography signal analysis for biomedical applications." *Sensors* 18.6 (2018): 1894.
 23. Perez, Luis, and Jason Wang. "The effectiveness of data augmentation in image classification using deep learning." *arXiv preprint arXiv:1712.04621* (2017).
 24. Um, Terry Taewoong, et al. "Data augmentation of wearable sensor data for Parkinson's disease monitoring using convolutional neural networks." *arXiv preprint arXiv:1706.00527* (2017).
 25. Frank Sgambati, "Driver Drowsiness Detection", Robert Bosch LLC, 2012.
 26. Blanco M. et al., Assessment of a drowsy driver warning system(DDWS) for heavy vehicle drivers. Report No. DOT 811 117, Washington DC.: National Traffic Highway Safety Administration. 2009.
 27. Kim, Chi-Jung, et al. "A study on evaluation of human arousal level using PPG analysis." *Journal of the Ergonomics Society of Korea* 29.1 (2010): 113-120.
 28. STEER: Wearable Device That Will Not Let You Fall Asleep, Kickstarter, <https://www.kickstarter.com/projects/creativemode/steer-you-will-never-fall-asleep-while-driving>
 29. Leisure Auto LS 803, http://www.leisureautosafety.com/web/2522355-2522355.html?i_d=54416&pid=2393371
 30. Ngiam, Jiquan, et al. "Multimodal deep learning." *Proceedings of the 28th international conference on machine learning (ICML-11)*. 2011.
 31. Gold, Christian, et al. "Take over!" How long does it take to get the driver back into the loop?", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2013.
 32. Mok, Brian Ka-Jun, et al. "Timing of unstructured transitions of control in automated driving", *IEEE Intelligent Vehicles Symposium*, 2015.
 33. Egil Fykes, "Performance comparison of GPU, DSP and FPGA implementations of image processing and computer vision algorithms in embedded systems", *Master of Science in Electronics*, 2013.
 34. Nvidia Drive AGX, <https://www.nvidia.com/en-us/self-driving-cars/drive-platform/hardware/>
 35. Data Distribution Service, <https://www.omgwiki.org/dds/>
 36. Constantin Brockner, Bettina Swynnerton, "A new architecture for automotive hardware-in-the-loop test", *ATZ elektronik*, 2014.
 37. 노태호, 닥터노의 알기쉬운 부정맥, 2016.
 38. Liang, Yongbo, et al. "Photoplethysmography and Deep Learning: Enhancing Hypertension Risk Stratification." *Biosensors* 8.4 (2018): 101.
 39. 정민우, et al. "In-Vehicle Infotainment 시스템에서 Configuration Protocol 의 설계 및 구현." *한국통신학회논문지* 43.7 (2018): 1140-1151.
 40. DDS Time Based Filter QoS Policy, https://community.rti.com/static/documentation/connex-dds/5.2.0/doc/manuals/connex-dds/html_files/RTI_ConnextDDS_CoreLibraries_UsersManual/Content/UsersManual/TIME_BASED_FILTER_QosPolicy.htm
 41. CoreDX DDS Architecture, <http://www.twinoakscomputing.com/coredx/architecture>
 42. He, Jibo, et al. "Detection of driver drowsiness using wearable devices: a feasibility study of the proximity sensor." *Applied ergonomics* (2017).
 43. Soukupova, Tereza, and Jan Cech. "Eye Blink Detection Using Facial Landmarks." *21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia*. 2016.
 44. Varma, Abhi R., et al. "Accident prevention using eye blinking and head movement." *Emerging Trends in Computer Science and Information Technology (ETCSIT2012)*. *Proceedings published in International Journal of Computer Applications (IJCA)* (2012): 18-22.

45. Koo, Dae Lim, and Juhan Kim. "정상 수면의 생리." *Hanyang Med Rev* 33 (2013): 190–196.
46. Yoo, Wook-Jae, et al. "Development of respiration sensors using plastic optical fiber for respiratory monitoring inside MRI system." *Journal of the Optical Society of Korea* 14.3 (2010): 235–239.
47. Leng, Lee Boon, Lee Boon Giin, and Wan-Young Chung. "Wearable driver drowsiness detection system based on biomedical and motion sensors." 2015 IEEE SENSORS. IEEE, 2015.
48. Dehzangi, Omid, Vikas Rajendra, and Mojtaba Taherisadr. "Wearable driver distraction identification on-the-road via continuous decomposition of galvanic skin responses." *Sensors* 18.2 (2018): 503.
49. Kundinger, Thomas, Andreas Riener, and Nikoletta Sofra. "A Robust Drowsiness Detection Method based on Vehicle and Driver Vital Data." *Mensch und Computer 2017-Workshopband* (2017).
50. Zhu, Xuemin, et al. "EOG-based drowsiness detection using convolutional neural networks." 2014 International Joint Conference on Neural Networks (IJCNN). IEEE, 2014.
51. Dwivedi, Kartik, Kumar Biswaranjan, and Amit Sethi. "Drowsy driver detection using representation learning." 2014 IEEE International Advance Computing Conference (IACC). 2014.
52. Du, Li-Huan, et al. "Detecting driving fatigue with multimodal deep learning." 2017 8th International IEEE/EMBS Conference on Neural Engineering (NER). IEEE, 2017.
53. Reddy, Bhargava, et al. "Real-time Driver Drowsiness Detection for Embedded System Using Model Compression of Deep Neural Networks." *Comput. Vis. Pattern Recognit. Work* (2017).
54. Yan, Chao, Frans Coenen, and Bailing Zhang. "Driving posture recognition by convolutional neural networks." *IET Computer Vision* 10.2 (2016): 103–114.
55. PhysioNet, <https://www.physionet.org/>
56. UCI Machine Learning Repository, <https://archive.ics.uci.edu/ml/index.php>
57. Elektra Driver Face dataset, <http://adas.cvc.uab.es/elektra/datasets/driver-monitoring/>
58. BioID Face Database, <https://www.bioid.com/facedb/>
59. Driver Drowsiness Detection Dataset, <http://cv.cs.nthu.edu.tw/php/callforpaper/datasets/DDD/>
60. RobeSafe Driver Monitoring Video Dataset, <http://www.robefsafe.com/personal/jnuevo/Datasets.html>
61. Lee, Dae Seok, Teak Wei Chong, and Boon Giin Lee. "Stress events detection of driver by wearable glove system." *IEEE Sensors Journal* 17.1 (2017): 194–204.
62. Lee, Kwan, et al. "Convolutional Neural Network-Based Classification of Driver's Emotion during Aggressive and Smooth Driving Using Multi-Modal Camera Sensors." *Sensors* 18.4 (2018): 957.
63. Zhang, Kaipeng, et al. "Joint face detection and alignment using multitask cascaded convolutional networks." *IEEE Signal Processing Letters* 23.10 (2016): 1499–1503.
64. Li, Z., Li, S. E., Li, R., Cheng, B., & Shi, J. (2017). Online detection of driver fatigue using steering wheel angles for real driving conditions. *Sensors*, 17(3)
65. Lin, L., Huang, C., Ni, X., Wang, J., Zhang, H., Li, X., & Qian, Z. (2015). Driver fatigue detection based on eye state. *Technology and Health Care*, 23(2), 453 – 463.
66. Wang, F., Wang, S., Wang, X., Peng, X., & Yang, Y. (2014). Design of driving fatigue detection system based on hybrid measures using wavelet-packets transform. In 2014 IEEE international conference on robotics and automation (icra) (pp. 4037 – 4042)

67. Patel, M., Lal, S., Kavanagh, D., & Rossiter, P. (2011). Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert Systems with Applications*, 38(6), 7235 - 7242.
68. Dwivedi, Kartik, Kumar Biswaranjan, and Amit Sethi. "Drowsy driver detection using representation learning." 2014 IEEE International Advance Computing Conference (IACC). 2014.
69. Reddy, Bhargava, et al. "Real-time Driver Drowsiness Detection for Embedded System Using Model Compression of Deep Neural Networks." *Comput. Vis. Pattern Recognit. Work* (2017).

주 의

1. 이 보고서는 국토교통부에서 시행한 국토교통기술촉진연구사업의 연구보고서입니다.
2. 이 보고서 내용을 발표하는 때에는 반드시 국토교통부에서 시행한 국토교통기술촉진연구사업의 연구결과임을 밝혀야 합니다.
3. 국가과학기술 기밀 유지에 필요한 내용은 대외적으로 발표 또는 공개하여서는 안 됩니다.

210mm×297mm[백상지(80g/㎡) 또는 중질지(80g/㎡)]