

최종 보고서

책임교수	소속	기초교양교육과	성명	(인)
참여학생	강지수, 김수아, 신수아, 여환서, 이동연, 이정현, 한기준			
유형	학술연구형			
계획 시 연구 주제		운전중 운전자의 감정변화 인식		

프로젝트 운영 내용		
시작 및 종료 일자		시작(5/22) 종료(1/16)
팀원 역할	이름	역할
	강지수	레퍼런스 조사 및 논문에 쓰일 레퍼런스 리스트 작성
	김수아	레퍼런스 조사 및 논문에 쓰일 레퍼런스 리스트 작성
	신수아	Resnet50 코드 연구 및 실행
	여환서	Xception 코드 연구 및 실행
	이동연	데이터 생성 및 데이터 테이블 제작
	이정현	데이터 생성 및 데이터 테이블 제작
	한기준	레퍼런스 조사 및 논문에 쓰일 레퍼런스 리스트 작성
운영비 집행 금액		297,700원

최종 연구 내용(학술연구형)					
제목	A study on driver state recognition using CNN-based multimodal multi-input learning				
투고 학회지	KJAI (Korean Journal of Artificial Intelligence)				
투고 일시	2/4일 예정	투고상황	투고 전		
게재가 판정일	2/14일	출간일	2월말	게재료 (예상)	30만원

최종 연구 내용(교재출판형)					
제목					
자문			편집 및 디자인		
출판사		출판상황	편집 중/인쇄 중	출판비 (예상)	원
ISBN				발행일	

1. 논문 초록

Accurately identifying the driver's driving status is very important for the safety of the driver and passengers, and many studies are being conducted on this. Many researchers have shown that attempts to identify the driver's status by creating an image of the driver's face and upper body using a camera and then learning CNN-based neural networks are somewhat effective. In addition, attempts have been made to identify the driver's fatigue using electroencephalogram (EEG), heart rate (ECG), and electrooculogram (EOG) data measurements, multiple sensors, and the RNN-LSTM model method. Nowadays, increasing accuracy is essential using simple sensor devices and simple artificial intelligence structures that can be installed in small devices such as on-device AI. In this paper, we used the driver's facial expression and upper body and the sound data generated during driving to recognize the driver's status more accurately. We then showed through experiments that CNN-based neural network learning alone using triple input elements improved meaningful accuracy. To apply on-device AI, we propose a CNN with a simple structure that can collect data using only a camera and a recorder. We compare the proposed method with learning in ResNet50 and Xception to show that it works well. These experimental results show that CNN can be used in multimodal applications and can be an efficient choice over other complex neural network learning methods that use multimodal learning data.

2. 연구의 가치 및 기대효과

본 연구는 실험을 통하여 multimodal 데이터를 multi-input 형태로 활용하여 학습시켰을 경우 정확도가 증가한다는 가설을 증명하려고 하였다. 실험을 위하여 driver-face와 driver-state, sound spectrogram으로 구분된 3가지 데이터를 준비하였으며 driver-face와 driver-state에 대한 single input 학습에서는 test accuracy가 80.7%, 68.6% 이었고 driver-face와 driver-state를 동시에 입력받아 학습하는 dual input의 경우는 test accuracy는 75% 정확도를 보여주었으나 sound spectrogram 데이터가 추가되어 진행된 triple input 실험에서는 99.9%로 test 정확도가 향상되었음을 확인하였다. 따라서 multimodal data를 multi-input 형태로 활용하여 학습시켰을 경우 정확도가 증가한다는 연구 가설이 증명되었다. 본 논문의 contribution은 다음과 같다.

- 1) 운전자의 상태를 구분하기 위하여 얼굴과 운전모습 이미지와 운전자의 목소리와 주변 소리 데이터를 학습한 멀티모달 CNN을 제안하였으며 실험결과 99%의 정확도를 보였다.
- 2) 운전자의 상태 3가지(정상, 기절, 졸림) 뿐만 아니라 운전자의 이상 상태를 초래할 수 있는 운전자의 감정 3가지(놀람, 분노, 불안)를 포함하여 모두 6가지를 구분하는 CNN 시스템을 제안하였다.
- 3) on-device AI 를 고려하여 단순한 디지털 기구로도 측정이 가능한 이미지, 소리만을 사용하였으며 단순한 구조의 CNN학습으로도 좋은 결과를 나타낼 수 있음을 보였다.

본 실험에서는 카메라와 녹음기능을 이용한 휴대폰과 같은 디지털 장치를 가정하여 데이터를 만들었다. 우리는 on-device AI에 사용될 수 있도록 적은 수의 레이어를 가진 단순한 multi-input CNN 구조로 driver state recognition 시스템을 제안하였다. 운전자의 조작이 필요없게되는 100% 자율주행 시대가 오기 전에는 운전자의 책임이 교통사고에 영향을 여전히 미치고 있다. 우리가 제안한 multimodal data, multi-input 형태의 활용방식은 electroencephalogram (EEG), heart rate (ECG), and electrooculogram (EOG) 과 같은 body signal을 비접촉 센서로 수집한 데이터에도 적용될 수 있으며 더욱 다양한 운전자의 상태를 파악하여 안전한 운전 시스템을 제안하는 향후의 연구에 많은 기여를 할 수 있을 것으로 기대한다.

3. 프로젝트 수행과 관련한 종합의견

7명의 학부 학생들과 프로젝트 팀을 구성하여 약 10개월의 연구기간을 가졌다. 전반기에는 CNN의 학습 모델을 이해하고 저널 논문의 일반적인 구성에 대하여 참고 논문을 스터디하는 형식으로 진행하였으며 후반기에는 우리 논문의 주제를 정하여 본격적인 연구를 진행하였다. 학생들은 데이터를 만들고 실험에 참여하는 데이터팀과 실험에 사용할 CNN코딩을 준비하는 코딩팀, 우리 논문의 실험 성과의 기준을 제시하기 위하여 관련 논문을 분석하는 레퍼런스팀으로 구분하여 각자 맡은 역할을 다 하였다. 모든 학생이 프로젝트가 끝나기 까지 성실하게 참여하여 우리가 원하는 결과가 나오게 되었다. 마지막까지 부족한 지도교수를 믿고 밝고 활기차게 따라와준 학생들에게 고마움을 표한다.