

실제 주행 환경에서의 운전자 스트레스 탐지를 위한 웨어러블 생체 신호와 얼굴 시각 데이터 융합 프레임워크

고서영†, 이현준†, 신승용, 노영태*
한양대학교 데이터사이언스학과

Multimodal Framework for Driver Stress Detection in Real-World Driving using Wearable Physiological Signals and Facial Cues from Videos

Seoyoung Ko†, Heonjun Lee†, Seungyoung Shin, and Youngtae Noh*

Department of Data Science, Hanyang University

{dddyoung, johnlee0321, seungyoungg, youngtaenoh}@hanyang.ac.kr

요약

운전자의 스트레스는 인지 및 운동 능력을 저하시켜 교통사고 위험을 증가시키는 주요 요인이다. 그러나 기존 연구들은 시뮬레이터 기반 실험이나 착용 부담이 큰 접촉형 센서에 의존하여 실제 주행 환경에서의 적용에 한계가 있었다. 본 논문에서는 이러한 한계를 극복하기 위해, 실제 도로 주행 환경에서 수집한 얼굴 시각 기반 운전자 행동 데이터와 손목 착용형 웨어러블 생체 신호를 융합한 운전자 스트레스 탐지 프레임워크를 제안한다. 제안한 프레임워크는 대시보드 카메라를 통해 획득한 얼굴 시각 정보와 손목 밴드형 웨어러블 기기를 통해 수집한 생리 신호를 결합함으로써, 일상적인 주행 환경에서도 적용 가능한 스트레스 탐지를 가능하게 한다. 16 명의 운전자를 대상으로 4 가지 실제 도로 환경에서 수집한 데이터를 분석한 결과, 생체 신호만을 활용한 경우 F1-score는 0.8356, 얼굴 시각 데이터만을 활용한 경우 0.8972 를 보였으며, 두 모달리티를 결합한 멀티모달 모델은 0.9252 로 단일 모달 대비 향상된 성능을 나타냈다. 또한 특징 중요도 분석을 통해 피부 온도와 머리 위치가 실제 주행 중 운전자 스트레스를 판별하는 핵심 지표임을 확인함으로써, 얼굴 시각 정보와 생체 신호의 상호보완적 효과를 실증적으로 입증하였다.

I. 서론

운전자의 스트레스는 인지 및 반응 속도를 저하시켜 교통사고 발생 위험을 평소 대비 10 배 이상 증가시키는 등, 심각한 사회적 안전 문제를 야기한다[1]. 이러한 위험성으로 인해 운전자 스트레스 탐지에 관한 연구는 지속적으로 이루어져 왔다. 그러나 기존 연구들은 주로 단일 모달리티(비디오, E4, ECG, SPR 등)에 의존하여 제한적인 변수만을 반영하거나[2-5], 가상현실(VR) 기반 실험을 통해 수행됨으로써 실제 주행 환경을 충분히 반영하지 못하는 생태학적 타당성의 한계를 지닌다[6]. 또한 실제 도로 주행 상황에서 가슴 스트랩과 같은 부착형 센서는 착용 불편과 운전 방해를 유발하여 장시간 적용에 제약이 있다[7].

이에 본 논문에서는 기존 연구들의 한계를 보완하기 위해, 실제 도로 주행 환경에서 수집한 운전자의 얼굴 시각 기반 행동 데이터와 손목 착용형 웨어러블 생체 신호를 융합한 멀티모달 스트레스 탐지 프레임워크를 제안한다. 특히 일상적으로 착용하기 어려운 가슴 스트랩이나 부착형 센서 대신, 착용이 간편한 손목 밴드형 웨어러블 기기(Empatica E4)를 활용함으로써 운전자의 부담을 최소화하는 동시에, 얼굴 시각 기반 행동 정보와의 융합을 통해 실제 주행 환경에서도 안정적인 스트레스 탐지를 가능하게 한다. 이를 통해 본 연구는 외부 환경 정보에 의존하지 않고, 실제 주행 중 나타나는 운전자 본연의 반응만을 기반으로 스트레스를 효과적으로 탐지할 수 있는 가능성을 제시한다.

II. 본론

본 연구의 데이터 수집은 실제 주행 환경의 다양성을 반영하기 위해 총 16 명의 운전자를 대상으로 고속도로(Speed), 도심(City), 종단 곡선(Sinewave),

이벤트 구간(Impact)의 4 가지 시나리오에서 진행되었다. 데이터 수집을 위해 운전자의 얼굴 및 상반신 움직임을 포착할 수 있는 대시보드 카메라(DJI Action Camera)와, 운전 중 구속감을 최소화하면서 생리 반응을 측정할 수 있는 손목 착용형 웨어러블 기기를 활용하였다. 이를 통해 주행 중 자연스럽게 발생하는 얼굴 시각 기반 행동 데이터와 심박수(HR), 피부전도도(EDA), 피부 온도 등의 생체 신호를 동기화하여 수집함으로써, 시뮬레이터가 아닌 실제 주행 환경(in-the-wild)에서의 신뢰성 있는 멀티모달 데이터셋을 구축하였다.

수집된 서로 다른 모달리티 데이터는 타임스탬프를 기준으로 정렬한 후, 멀티모달 융합을 위해 1Hz 리샘플링하였다. 이후, 신호의 안정적인 분석을 위해 이상치 제거 및 정규화를 수행하였다. 스트레스 레이블은 운전자의 주관적 평가를 기반으로 한 5 점 척도를 이진 분류(Stress/Non-Stress)로 변환하여 적용하였다. 특징 추출 단계에서는 YOLOv8[8]과 SPIGA[9] 알고리즘을 활용해 눈 깜빡임(EAR), 머리 자세, 시선 벡터 등

Real-World Driving Stress Study: Data Collection and Analysis Pipeline

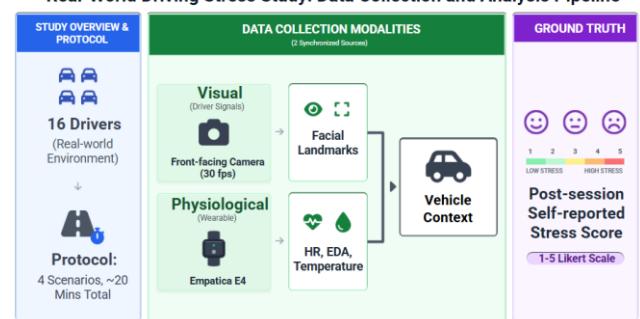


그림 1. 실제 운전 환경 멀티모달 데이터 수집 및 스트레스 탐지 파이프라인

153 개의 시각적 특징을 추출하고, 웨어러블 기기로부터 평균, 표준편차 등의 통계량을 포함한 30 개의 생체 신호 특징을 산출하였다. 이후 과적합을 방지하고 모델 효율성을 높이기 위해 Welch's t-test를 적용하여, 전체 183 개의 특징 중 스트레스 판별에 통계적으로 유의한 53 개의 핵심 특징을 선별하였다.

총 16 명의 운전자가 실험에 참여하였으나, 일부 운전자에게서는 데이터 손실이 발생하거나 스트레스 또는 비스트레스 상태 중 하나의 레이블만 판측되는 단일 레이블 문제가 확인되었다. 이러한 경우 이전 분류 학습이 불가능하므로, 최종적으로 양쪽 상태가 모두 판측되어 유효성이 확보된 8 명의 운전자 데이터를 모델 학습 및 평가에 활용하였다. 또한 운전자 간 생체 반응 및 행동 패턴의 개인차를 고려하기 위해 개인화된 스트레스 탐지 모델을 구축하였다. 모델 성능 비교를 위해 Logistic Regression, Gradient Boosting, XGBoost, SVM, Multi-layer Perceptron 의 다섯 가지 머신러닝 알고리즘을 적용하였으며, 데이터 불균형 문제를 완화하기 위해 계층적 교차 검증(Stratified Cross-Validation)을 사용하였다. 실험 결과, SVM 기반 모델이 가장 우수한 성능을 보였으며, 특히 얼굴 시각 정보와 생체 신호를 결합한 멀티모달 융합 모델이 단일 모달리티 모델 대비 월등히 높은 0.9252의 F1-Score를 달성하여 제안한 프레임워크의 효과를 입증하였다.

F1 Score Comparison across Modalities and Models

Performance evaluation of 5 machine learning models with 3 data modalities

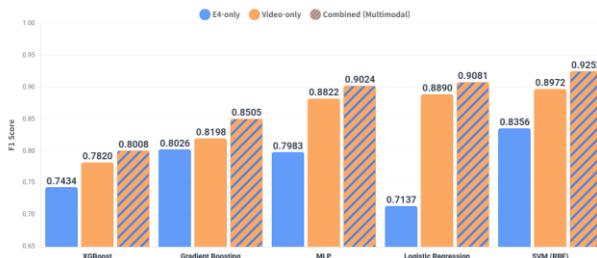


그림 2. 모델별 사용 모달리티별 (E4/Video/Combined) F1 점수 비교 그래프

멀티모달 SVM 모델의 특징 중요도 분석 결과, 상위 10개 주요 특징 중 약 70%가 생체 신호로 구성되었으며, 이 중 다수가 피부 온도와 관련된 지표로 운전자의 기저 생리 상태를 반영하는 것으로 나타났다. 동시에 머리 위치와 같은 얼굴 시각 기반 행동 특징은 스트레스 발생 시점의 맥락적 변화를 포착하는 데 중요한 역할을 수행함을 확인하였다.

III. 결론

본 논문에서는 실제 주행 환경에서 운전자의 스트레스를 웨어러블 기기와 카메라 기반 시각 정보를 융합하여 탐지할 수 있는 멀티모달 프레임워크를 제안하고 그 유효성을 검증하였다. 기존 연구들이 시뮬레이터 환경이나 착용 부담이 큰 센서에 의존했던 것과 달리, 본 연구는 손목 착용형 웨어러블 기기를 활용하여 운전자의 편의성을 확보하는 동시에, 실제 도로 주행 데이터를 기반으로 한 실험을 통해 안정적인 스트레스 탐지가 가능함을 보였다. 8 명의 운전자를 대상으로 한 모델링 결과, 제안한 멀티모달 모델은 실제 주행 환경에서도 높은 성능을 보였으며, 이를 통해 웨어러블 생체 신호와 얼굴 시각 정보의 융합이 스트레스 탐지에 효과적임을 확인하였다. 특히 특징 중요도 분석을 통해 피부 온도와 머리 위치 정보가

운전자 스트레스를 판별하는 핵심 지표임을 규명하였고, 통계적 특징 선택을 통해 모델을 경량화함으로써 상용 웨어러블 기기와 카메라만을 활용한 현장 적용 가능성을 제시하였다. 향후 연구에서는 시계열 모델(LSTM 등)을 적용하여 스트레스의 시간적 변화를 보다 정밀하게 분석하고, 실제 도로 주행 중 스트레스가 발생하는 구간을 정확히 식별함으로써 각 특징이 스트레스에 미치는 인과적 영향을 심층적으로 규명할 계획이다.

ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. RS-2025-00522803).

참 고 문 헌

- V. A. Sauerzapf, "Road traffic crash fatalities: A comparison of Great Britain and the United States," Ph.D. dissertation, Univ. East Anglia, Norwich, U.K., 2012.
- Giannakakis, G., Pediaditis, M., Manousos, D., Kazantzaki, E., Chiarugi, F., Simos, P. G., Marias, K., & Tsiknakis, M. (2017). Stress and anxiety detection using facial cues from videos. *Biomedical Signal Processing and Control*, 31, 89–101.
- Toshnazarov, K., Lee, U., Kim, B. H., Mishra, V., Najarro, L. A. C., & Noh, Y. (2024). SOSW: Stress Sensing With Off-the-Shelf Smartwatches in the Wild. *IEEE Internet of Things Journal*, 11(12), 21527–21545.
- Keshan, N., Parimi, P. V., & Bichindaritz, I. (2015). Machine learning for stress detection from ECG signals in automobile drivers. *2015 IEEE International Conference on Big Data (Big Data)*, 2661–2669.
- Affanni, A., Bernardini, R., Piras, A., Rinaldo, R., & Zontone, P. (2018). Driver's stress detection using Skin Potential Response signals. *Measurement*, 122, 264–274.
- Mateos-García, N., Gil-González, A.-B., Luis-Reboredo, A., & Pérez-Lancho, B. (2023). Driver Stress Detection from Physiological Signals by Virtual Reality Simulator. *Electronics*, 12(10), 2179.
- Spencer, C., Koc, I. A., Suga, C., Lee, A., Dhareshwar, A. M., Franzén, E., Iozzo, M., Morrison, G., & McKeown, G. J. (2020). A Comparison of Unimodal and Multimodal Measurements of Driver Stress in Real-World Driving Conditions. *PsyArXiv*.
- R. Varghese and S. M., "YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness," 2024 Int. Conf. Adv. Data Eng. Intell. Comput. Syst. (ADICS), Chennai, India, 2024.
- A. Prados-Torreblanca, J. M. Buenaposada, and L. Baumela, "Shape Preserving Facial Landmarks with Graph Attention Networks," arXiv preprint arXiv:2210.07233, 2022.