

## 대조 학습 기반 운전자 사고 영상 클러스터링 및 사고 패턴 분석

주준경<sup>1</sup>, 김태욱<sup>1</sup>, 박규민<sup>2</sup>, 신효승<sup>1</sup>, 홍정희<sup>1\*</sup>

<sup>1</sup>연세대학교 소프트웨어학부, <sup>2</sup>연세대학교 컴퓨터정보통신공학부

### Driver Accident Video Clustering and Analysis based on Contrastive Learning

Jun-Kyung Ju<sup>1</sup>, Tae-Wook Kim<sup>1</sup>, Gue-Min Park<sup>2</sup>, Hyo-Seung Shin<sup>1</sup>, Ellen J. Hong<sup>1\*</sup>

<sup>1</sup>Division of Software, Yonsei University

<sup>2</sup>Department of Computer & Telecommunications Engineering, Yonsei University

**Abstract** Traffic accidents cause human casualties and social costs every year. Therefore, analyzing and preventing traffic accidents is extremely important. Traditionally, studies on traffic accident cause analysis have been conducted using clustering techniques to analyze high-risk drivers based on driver characteristics, vehicle accelerometer sensor data, and CCTV footage. However, research on traffic accidents based on black box videos, which include visual information directly influencing the driver's actions, has not been actively pursued. Thus, this paper aims to analyze major accident patterns by extracting video-specific vectors using Video Vision Transformer-based contrastive learning and applying clustering techniques. As a result, a total of 8 clusters were formed, and the characteristics of each cluster were analyzed.

- Key Words : Video Clustering, Contrastive Learning, K-Means, Random Crop, Gaussian Noise

### I . 서론

국내 교통사고는 2022년 기준 사상자 수 1,940,520명, 사회적 비용으로 약 26조 2,833억 원[1]으로, 매년 인명 및 재산 피해를 발생시킨다. 따라서 교통사고를 분석하고 예방하는 것은 매우 중요하다. 특히, 국내 교통사고율이 인구 감소율과 반비례하게 나타나는 추세는 해당 연구의 필요성을 더욱 강조한다. 기존의 교통사고 원인 분석은 주로 운전자 특성[2]이나 차량 가속계 센서[3]를 기반으로 진행되었다. 하지만, 운전자의 사고에 직접적인 영향을 미치는 시각 정보가 포함된 블랙박스 영상기반 교통사고 연구는 활발히 진행되지 않았다. 최근에는 대조 학습을 통해 영상을 벡터로 표현하는 연구가 진행되었다[5]. 이러한 기법을 바탕으로, 본 논문은 사고 블랙박스 영상을 사용하여 운전 행동 및 사고 맥락 정보를 대조 학습하여 벡터화한 뒤, 클러스터링 알고리즘을 통해 사고 영상을 군집화하여 사고 패턴을 직접 분석하고자 한다.

### II. 관련연구

기존에는 교통사고 원인 분석은 사고 운전자 특성[2]이나 차량 가속계 센서 정보[3] 등을 사용하여 클러스터링 기법을 통해 위험 운전자에 대한 분석을 수행하는 연구가 진행되었

다. 또한, 사고 발생 지점의 CCTV[4] 데이터를 기반으로 차량 궤적 추적 및 차량 간 거리 정보를 추출한 뒤 합성곱 신경망 모델 학습을 통해 교통사고 분석 및 발생 여부 예측을 판단하는 연구가 수행되었다. 그러나 위 3가지 데이터 모두, 운전자의 사고 직전 판단에 직접적인 영향을 미치는 1인칭 시야 정보 반영하지 못한다. 최근에는 대조 학습을 통해 이미지 또는 영상 간 유사성에 따라 벡터를 추출하는 연구가 진행되었다[5][6]. 따라서, 본 논문도 위 기법을 활용하여 사고 순간의 운전 행동 및 사고 맥락을 포함하는 영상으로부터 유의미한 벡터를 추출하여 사고 패턴을 분석하고자 한다.

### III. 실험 설계

#### 3.1 데이터셋 구축 및 전처리 방법

본 논문은 AIHUB에서 제공하는 교통사고 영상 데이터[7]를 기반으로 사고 패턴을 분석하고자 한다. 위 데이터셋은 차대차, 차대보행자 등의 사고 대상과 직선도로, 횡단보도 등의 사고 장소 등으로 구분되는 21,895개 사고 영상과 라벨로 구성된다. 이 중 위험한 운전자의 사고 패턴 분석을 위해, 차대차 사고 블랙박스 영상에서 자기 과실 비율이 70% 이상인 영상만을 추출하였다. 추출된 영상을 딥러닝 모델에 대조 학습하기 위해 각 영상으로부터 같은 시간 간격으로 16 Frame을 추출하고 영상 크기를 224x224로 조정한 뒤, 영상을 구성하는

모든 픽셀의 RGB값을 255로 나눠 정규화를 진행하였다. 추가적으로, 대다수의 영상에 포함되어 있는 Black Border를 제거하여 비디오 속 Black Border가 대조 학습에 영향이 미치지 않도록 하였다.

### 3.2 비디오 대조 학습

기존의 이미지 및 비디오 대조 학습 연구[5][6]를 활용하여, 3.1절에서 구축한 N개의 영상에 대해 각 영상에 서로 다른 가우시안 노이즈 필터와 Random Crop을 적용하여 2N개의 변조된 영상을 생성한다. 이후 Video Vision Transformer(ViT) 기반 모델에 2N개의 영상을 입력하여 모델로부터 영상에 대한 3차원 벡터를 추출한다. 마지막으로 temperature가 0.9인 NT-Xent 손실 함수를 통해 비슷한 영상 간의 벡터 유사도는 높은 방향으로 대조 학습을 진행한다.

### 3.3 비디오 벡터 기반 클러스터링 및 분석

3.2절을 통하여 학습된 모델을 바탕으로 사고 영상에 대한 벡터를 산출한 뒤, 벡터 간 거리를 유사도로 설정하여 K-Means 알고리즘을 사용하여 군집화한다. 이때, 클러스터 개수는 2부터 30까지 조정하여 클러스터링을 수행한 뒤, 클러스터 성능 지표인 Silhouette Score를 바탕으로 결정한다. 이후 각 클러스터의 중심과 가장 가까운 10개의 영상을 추출하여 각 클러스터의 운전자 사고 패턴을 분석한다.

## IV. 실험 결과 및 분석

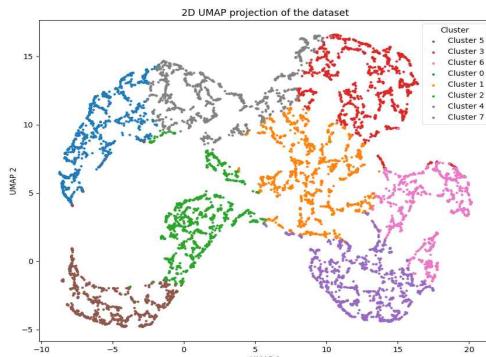


Fig. 1 2D UMAP Visualization

Table 1. Accident Pattern Analysis by Cluster

Cluster ID	Description
0	Caused by Blind Spots or Side Collisions
1	Sudden Deceleration or Slow Driving
2	Accidents Occurring at Night
3	Accidents at Intersections
4	Accidents in Congested Areas
5	Left/Right Front-End Collision Accidents
6	Sudden Deceleration
7	Limited Visibility or Lane Changes

학습을 통해 Silhouette Score가 가장 높고, 클러스터 수가 가장 낮은 클러스터 개수인 8개를 기준으로 각 영상별 벡터를 차원 축소 알고리즘인 UMAP을 통해 시각화한 결과는 Figure 1과 같다. 이후, 각 클러스터 중심점과 가장 가까운 영상 10개를 추출하여, 클러스터별 특징을 분석한 결과는 Table 1과 같다.

## V. 결론

본 논문은 운전 영상을 바탕으로 사고 패턴을 분석하기 위해 딥러닝 모델을 대조 학습하였고, 총 8개의 클러스터를 형성하여 각 클러스터별 특징을 분석하였다. 하지만, 클러스터 내 일부 영상들이 일관적인 특성을 공유하지 않거나, 차량의 움직임보다 유사한 배경에 따라 형성된 클러스터가 존재하였다. 따라서, 대조 학습 과정에서 영상 속 차량들의 움직임에 모델이 더 집중하도록 Object Detection 기반 차량 필터링 등을 적용하여 향후 연구를 진행할 예정이다.

## ACKNOWLEDGMENTS

본 논문은 과학기술정보통신부의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No.2022R1F1A1074273)

## REFERENCES

- [1] "In 2022, about 26.2883 trillion won in social costs due to road traffic accidents was incurred" Accessed: Feb. 15, 2024. [Online]. Available: [https://www.koroad.or.kr/main/board/6/89912/board\\_view.do](https://www.koroad.or.kr/main/board/6/89912/board_view.do)
- [2] Driver Group Clustering Technique and Risk Estimation Method for Traffic Accident Prevention, Journal of The Korea Society of Computer and Information, 2024
- [3] S. Chen, K. Cheng, J. Yang, X. Zang, Q. Luo, and J. Li, "Driving Behavior Risk Measurement and Cluster Analysis Driven by Vehicle Trajectory Data," Applied Sciences (Switzerland), vol. 13, no. 9, 2023
- [4] A. P and G. Nallasivan, "A Vision-Based Traffic Accident Analysis and Tracking system from Traffic Surveillance Video," 2024 Third International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Krishnankoil, Virudhunagar district, Tamil Nadu, India, 2024
- [5] Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. "A Simple Framework for Contrastive Learning of Visual Representations", Proceedings of the 37th International Conference on Machine Learning. 2020
- [6] Qian, R., Meng, T., Gong, B., Yang, M.-H., Wang, H., Belongie, S., & Cui, Y. "Spatiotemporal Contrastive Video Representation Learning", IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021
- [7] AI-Hub. Accessed: Jun. 16, 2024.[Online]. Available: <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&dataSetSn=597>