

Global Report

졸음운전 정량화 연구

1. 개요
2. 졸음 관찰자 등급
3. 알고리즘1: 차량 이력 데이터와 ORD
4. 알고리즘2: 얼굴 영상과 ORD 점수
5. 오분류 탐색
6. 결론

1. 개요

졸음운전은 자동차를 운전하는 사람들에게 큰 문제로 남아 있다. 미국 국토교통안전관리국 (NHTSA : National Highway Traffic Safety Administration)의 국립통계분석센터 (NCSA: National Center for Statistics and Analysis)에 따르면, 졸음운전으로 인하여 2018년에 775명이 사망하였으며, 졸음운전은 미국 도로에서의 치명적인 사고비율의 2.1%를 차지한다 (NCSA, 2019), 2015년 차량사고 관련 경찰보고서에 따르면 1.4%가 졸음운전에 의한 것이다 (NCSA, 2017). 그러나 이러한 통계수치는 경찰 보고서에 의존한 것이며, 졸음운전의 문제를 과소평가한 것이다. 미국 국토교통안전관리국의 국가 차량사고 샘플데이터를 조사해보면, 미국 교통안전자동차협회는 (AAAFTS: American Automobile Association Foundation for Traffic Safety) 졸음운전과 연관된 치명적인 사고의 비율을 산출하기 위해 다양한 원인들을 복합적으로 고려한다. 미국 교통안전자동차협회는 졸음운전이 치명적인 사고의 21%를 차지한다고 보고 있다.

본 프로젝트에서는 대규모의 실제 운전데이터 집단으로부터 졸음운전을 식별하기 위해 기계 학습 알고리즘의 사용 타당성을 연구하고 있다. 전략적 고속도로 연구 프로그램 2 (SHRP2: The Strategic Highway Research Program 2)의 자연스러운 운전 연구 (NDS: Naturalistic Driving Study)를 통해, 6개 주에서 3,400명 이상의 운전자가 참여하여, 5백만 개 이상의 운행에 대한 시계열 데이터를 수집하였다. (Virginia Tech Transportation Institute, 2020). SHRP2 NDS 데이터베이스에는 차량 변수와 운전자의 얼굴을 포함한 비디오 데이터가 포함되어 있다. 운전자를 향한 카메라 뷰와 가속 및 차선 변화와 같은 광범위한 차량 이력 데이터를 분석해 보면, 졸음이 충돌, 준(準) 충돌 또는 무사고 운행에 언제 영향을 미치는지를 식별이 가능할 수 있다. 이를 위해 현재 SHRP2 NDS 데이터베이스를 활용하여 두 가지의 기계 학습 알고리즘을 연구 개발하였다. 이 프로젝트를 위해 문제해결 절차 및 계산을 수행하기 위한 따라야 할 프로세스 또는 규칙 집합의 알고리즘이 정의된다. 첫 번째 알고리즘은 졸음운전과 관련된 운전자 행동을 유출할 수 있는 차선 변화와 같은 차량 이력 데이터를 사용하였으며, 두 번째 알고리즘은 졸음과 관련된 머리 끄덕임 등 운전자의 얼굴 비디오 데이터를 사용하였다.

2. 졸음 관찰자 등급

모든 운행구간 (“epochs”)에서 충돌 또는 준(準) 충돌과 연관된 운전 손상데이터와 기준선을 위한 운전샘플 데이터가 수동으로 코딩되었다. 운전 손상의 한 범주는 “Drowsy, sleepy, asleep, fatigued”이며 이 범주는 “대상 차량 운전자가 수면 또는 피곤함의 명백한 징후를 보이거나 운전 중에 실제로 잠이 들어 운전 성능이 저하된다.”(VTI, 2015) 이다. 연구팀은 운전 장애요소로서 졸음으로 나타나는 589개의 1분짜리 epochs와 졸음으로 인한 명백한 운전 장애 요소가 없는 200개의 추가적인 기본 epochs를 확보하였다. 졸음 수준을 결정하기 위해, 각각 세 명의 분석가가 졸음 관찰자 등급 (ORD) 프로토콜을 사용하여 (Wierwille & Ellsworth, 1994; Wiegand et al., 2009). 789 개의 epoch 비디오를 코딩했다. ORD 프로토콜은 지속적인 규모의 졸음에 대한 운전자의 제스처, 얼굴 톤 및 졸음 운전행동에 대한 평가를 포함하고 있으며, 졸음 식별을 위한 신뢰성이 있는 도구를 제공한다. (Wiegand et al., 2009) 낮은 비디오의 질과 누락된 데이터로 인하여, 연구자들은 몇몇의 epochs에 대해서는 일부를 코딩 할 수 없었다. 결과적으로 741 (94%)의 epochs가 유지되고 평가 되었다. 이 프로젝트는 두 개의 졸음운전 식별 알고리즘 테스트를 위한 실측정보 (“ground truth”)로 세 가지 ORD 등급의 평균을 사용하였다. (알고리즘이 초기 코딩된 ORD 프로토콜을 사용하여 성공적으로 졸음운전을 판별하면, 유효하다고 할 수 있다.)

3. 알고리즘1 : 차량 이력 데이터와 ORD

첫 번째 알고리즘은 졸음운전을 식별하기 위해 차량 센서데이터와 같은 차량 이력 데이터를 사용하였다. 연구원들은 정상 운전데이터로부터 의도적인 차선이탈 데이터와 비 정상적인 저속데이터를 조합하여 사용하였으며, 30 kph 이하의 저속데이터와 교차로에서의 정지 데이터는 제거하고 누락된 긴 세션 데이터도 제거하였으며, 결측된 짧은 세션 데이터는 절단 (truncation) 및 보간(interpolation)을 사용하여 보완하였다. 연구원들은 분석을 위해 각각의 ORD코드화 된 epoch를 6초 분량의 연속데이터로 분리하였다. 각각의 연속데이터들을 위해 25분위, 중간값, 75분위와 표준편차를 계산하였으며 연구원들은 요 레이트 (yaw rate (또는 z 축에 대한 차량의 회전 정보)와 차량 중심으로부터 벗어난 차선 위치 값을 측정하였다(표 1). 연구자들은 또한 차량 센터로부터의 상대적인 기울기로 차선 위치를 계산 하였는데 즉, 차선 중심으로 부터의 거리 변화를 항상 측정하였다. 또한 차선 중심으로 부터의 변화량과 차선 기울기 변화를 위해 각 epochs 대하여 요 레이트와 차선 위치를 25분위, 75분위로 계산하였다. 연구자들은 또한 각 epoch별로 차선이탈과 차선변경 횟수를 유도하였다. 만약 한번의 차선 이탈이 초기에 왼쪽 위치에서 마지막에 오른쪽으로 이동하였거나 반대의 경우, 운전자가 차선의 중심선을 넘었기 때문에 차선 변경으로 간주되었다. 이 데이터는 차선이탈에 대한 수정 계산이 가능하게 된다. 추가적인 차량 이력데이터로는 운행 시작과 운행계속시간이 포함되어 있다. 연구자들은 이러한 측정치를 알고리즘 개발의 입력값으로 사용하였다.

<표1. 알고리즘 입력값>

Yaw rate	
1	25th percentile of all subsequences of gyro_z_25%. Gyro_z_25% is the 25th percentile of yaw rate, the rate of change around the yaw axis (or the rotation of the car around the z axis).
2	25th percentile of all subsequences of gyro_z_75%. Gyro_z_75% is the 75th percentile of yaw rate, the rate of change around the yaw axis.
3	75th percentile of all subsequences of gyro_z_25%. Gyro_z_25% is the 25th percentile of yaw rate, the rate of change around the yaw axis.
4	75th percentile of all subsequences of gyro_z_75%. Gyro_z_75% is the 75th percentile of yaw rate, the rate of change around the yaw axis.
Lane Position (deviation from center of lane)	
5	25th percentile of all subsequences of lanepos_25%. Lanepos_25% is the 25th percentile of lane position.
6	25th percentile of all subsequences of lanepos_75%. Lanepos_75% is the 75th percentile of lane position.
7	75th percentile of all subsequences of lanepos_25%. Lanepos_25% is the 25th percentile of lane position.
8	75th percentile of all subsequences of lanepos_75%. Lanepos_75% is the 75th percentile of lane position.
Lane Departures	
9	The number of lane departures.
Slope of Lane Position	
10	25th percentile across all subsequences of the slope of lane position. Slope of lane position is a measure of the change in lane position across a subsequence.
11	75th percentile across all subsequences of the slope of lane position. Slope of lane position is a measure of the change in lane position across a subsequence.
Trip Start Time and Duration	
12	The 3-hour time bin at the start of the trip.
13	Trip duration at the start of epoch.

연구자들은 표1에 제시된 13개의 특징들을 사용하여 각각의 epoch에 대한 벡터를 만들었다. 이러한 벡터들은 졸음인지 아닌지를 분류하기 위한 기계 학습 알고리즘의 입력으로 사용되었다. 누락된 차량 이력데이터를 제거한 후에 연구자들은 432개의 온전한 벡터를 추출하였다. ORD 스케일은 0부터 100까지이며, 0은 졸음이 아닌 것, 50에서 69는 약간 졸음, 70이상은 매우 졸음으로 분류 되었다. 연구자들은 운전자들의 ORD 점수를 졸음과 비졸음을 판별하기 위해 50과 70을 임계치로 사용하였다. 연구자들은 파이썬 라이브러리 “scikit-learn”으로부터 gradient boosting classifier, random forest classifier, extra trees classifier와 ada boost classifier 4종의 트리 기반 분류자(classifier)들을 적용하였다.

Scikit-learn 라이브러리는 각각의 입력 특성들의 상대적인 중요도를 반환한다. 상대적인 중요도는 서로 비교 된다.

ORD Threshold = 50

연구자들은 1/3의 epochs들을 테스트 집단으로 남겨놓고 알고리즘의 성능을 평가한다. 위의 4가지 트리 기반 학습대상(learner)중에서 extra tree 분류자(classifier)가 ORD 임계치 50에 대해서는 졸음운전 분류에 대해 최고의 성능치가 나왔다.

테스트 집단을 적용하였을 때, 연구자들은 표2와 같은 confusion matrix를 얻었다. Confusion Matrix는 알고리즘이 올바르게 졸음운전을 식별할 수 있음을 보여주고 있다. 따라서, ORD 임계치를 50으로 사용하였을 때, ORD 점수와 함께 epochs가 50 또는 그 이상이 졸음으로 간주된다. 이항 분류의 질을 나타내는 매튜상관계수 (The Matthews Correlation Coefficient: MCC)는 0.25이다. 다른 상관계수와 같이, 0 이면 상관관계가 없고 1이면 일치하고 -1이면 불일치이다. 보통 상관계수를 0.25로 한다.

<표2. Confusion Matrix With ORD Threshold of 50>

	Algorithm Classification		
		False	True
	False	16	38
ORD Rating	True	9	80

알고리즘에 대한 수신기 조작 특성 (Receiver operating characteristics: ROC) 곡선에 대하여 AUC (area under the curve)는 1일 경우 완전한 졸음이고 0.5인 경우, 졸음일 가능성이 있는 것이다. ORD 임계치를 50으로 하였을 경우 AUC 값은 0.72 였다.

ORD Threshold = 70

이 경우, ORD 점수가 70 또는 보다 큰 경우, 보다 엄격한 “졸음”수준으로 간주된다. 위에서 전술한 4개의 트리 기반 학습대상(learner) 중에서 gradient boosting 분류자(classifier)가 가장 좋은 성능을 나타내었다.

테스트 집단을 적용하였을 때, 연구자들은 표3과 같은 confusion matrix를 얻었다. 매튜상관계수(MCC)는 0.33으로 정하였다. AUC 값은 0.76 이었다.

<표3. Confusion Matrix With ORD Threshold of 70>

	Algorithm Classification		
		False	True
	False	6	13
ORD Rating	True	6	118

ORD 점수가 70이거나 매우 졸음(very drowsy)의 피쳐들에 대한 중요도가 순서적으로 정렬되었다. 두가지의 가장 중요한 특성들은 75분위의 차선 위치 기울기 (표1의 11)과 25분위의 차선 위치 기울기 (표1의 12)였으며 이러한 변수들은 epoch들에 대하여 차선 변화의 양을 나타낸다. ORD 임계치가 50일 경우 중요한 특성들이었던 운행 시작시간 (표1의 12)와 운행지속시간 (표1의 13)은 임계치 70일 경우는

다섯 번째와 여섯 번째로 중요한 특징들 이었다.

이러한 결과들은 약간의 졸음운전 (moderately drowsy driving, ORD 임계치 50) 또는 심각한 졸음운전 (severely drowsy driving, ORD 임계치 70)에 대한 분류가 타당함을 보여주고 있다. 두 가지의 임계치 중 3가지 특징들은 5위 안에 포함 되었다(표 1의 13, 5, 6). 따라서 운전자의 운전지속시간과 차선중앙으로부터의 변화값들은 졸음운전을 나타내는 유용한 정보들이다. 반면, 차선 위치에 대한 기울기 (표 1의 10, 11)은 ORD 임계치 70인 경우, 매우 졸림(very drowsy)을 나타내는 가장 중요한 특징들이다. 약간의 졸음 (moderately drowsy)에 대한 가장 중요한 특징들은 운전 시작 시간, 운전 연속시간이며, 이러한 정보들을 포함하는 운전시간과 운전지속시간은 약간 졸음 epochs를 분류 할 수 있는 좋은 정보들이다. 이러한 결과들은 약간 졸음은 운전시간과 연속시간들을 통해서 유추할 수 있으며, 매우 졸음운전 (very drowsy drivers)들은 차량 센서로부터 획득한 차선 변화정보로 결정할 수 있다. 이러한 요인들은 자연 운전 데이터(naturalistic driving data)에서 졸음을 판별하는 데이터로 사용될 수 있지만, 모든 경우에 올바르게 졸음운전을 분류하기에는 충분하지 않을 수도 있다.

4. 알고리즘2 : 얼굴 영상과 ORD 점수

졸음운전은 감지할 수 있는 차량 이력데이터를 보완하기 위해, 연구진들은 ORD점수와 1분전까지의 데이터를 포함하여 운전자의 얼굴 비디오 영상을 분석하였다. 연구자들은 ORD 점수가 50 이상이거나 약간 졸음 (moderately drowsy)을 나타내는 epoch들을 감지하기 위해 기계 학습 알고리즘을 개발하였다. 먼저 영상에서 눈과 코 같은 위치를 추출하고, 이를 기반으로 머리 위치를 판단할 수 있는 비전 처리 알고리즘을 적용하였다. 추가적으로 눈, 입의 위치 뿐 만 아니라 눈과 입이 열려있는 정도를 알 수 있도록 하였다. 비디오데이터로부터 얼굴데이터를 추출한 이후, 다음 프레임을 분석하기 위한 시작 포인트로 데이터를 이용하였다. 최종 단계에서 연구진은 머리의 위치 추정과 이전 프레임의 기준점을 이용하여 다음 프레임을 초기화 하였다.

졸음운전 예측 모델 개발을 위해 데이터들은 3단계 처리 과정을 거쳤다. 첫 번째 단계에서는 불확실한 데이터나 일관성이 없는 데이터를 표시하였다. 왜냐하면, 비디오 영상의 질은 epoch별로 달랐으며, 모든 데이터를 머신 비전 알고리즘으로 추출할 수 없기 때문이다. 이러한 이유로, 연구자들은 머신 비전 알고리즘으로부터 확실성이 적은 데이터들을 제거하기 위해 필터를 사용하였다. 낮은 확실성 데이터들은 전형적으로 조명 환경이 좋지 않았다. 예를 들면 야간 주행이나 조명이 아주 밝은 지역인 경우이다. 이러한 필터 적용 후, 594개의 epochs가 남았다. 이 epochs들은 학습용 (85%)와 테스트용(15%) 집단으로 분리하였다. epoch들은 ORD점수에 따라 같은 양의 졸음과 경고 케이스로 랜덤 샘플 방식으로 계층화 되었다.

두 번째 단계에서는 머신 비전 알고리즘의 결과로부터 통계적으로 졸음과 관련된 운전 특징들을 유도하였다. 졸음 관련된 머리 움직임을 나타내는 머리 기준점의 변화 뿐 만 아니라 60초 동안 눈이 감긴 비율을 나타내는 PERCLOS를 포함하고 있다. 이러한 특징들을 이용하여 연구자들은 눈과 입의 위치 뿐 만 아니라 요, 피치, 롤과 머리 위치의 중간값, 백분위 값을 포함하였다. 머리위치와 입 크기의 차기 같은 부수적인 특징들로 고려되었다.

세 번째 단계에서는 졸음을 예측하기 위해서 기계 학습 알고리즘을 적용하였다. 이러한 용도로 5개의 알고리즘이 사용되었다 (general linear model, lasso, support vector machine, random forest and extreme gradient boosting). 연구자들은 10-fold cross-validation tuning 프로세스를 사용하여, 얼굴 특징들에 대하여 5가지 알고리즘을 학습시켰다. Random forest 알고리즘과 같이 특정 tuning 파라미터가 필요한 경우, 연구자들은 최상의 효과를 나타낼 수 있는 파라미터를 모델에 적용하여 예측하였다. 연구진들은 10-fold-cross-validation 과정을 서로 다른 학습 데이터를 선택하여 두 번 반복 하였다. 반복적인 검증은 학습데이터가 중복하지 않도록 하여 모델이 잘 동작 할 수 있게 하였다. Extreme gradient boosting algorithm 이 가장 좋은 성능을 나타냈으며 AUC = 0.76 이다. 졸음을 예측할 수 있는 가장 중요한 특징들은 눈이 떠진 크기, 입이 벌어진 크기, 머리의 동작(요, 롤, 피치)이다.

5. 오분류 탐색

운행이력 데이터 알고리즘과 머신 비전 얼굴 비디오 알고리즘이 졸음운전에 대한 예측을 잘 함에도 불구하고 오분류에 대한 문제는 남아 있다. 좀 더 살펴보면, 패턴들은 오분류를 설명하는데 도움이 된다. ORD 점수가 운전을 졸음으로 분류하고 있으나, 차량 이력 데이터 알고리즘이 운전을 졸음으로 분류하지 않는 상황에서 영상데이터에 대한 명백한 증거를 캡처하지 못 할 수도 있다. 몇몇의 epoch들에 대해서는 졸음과 관련되지 않는 차량 움직임에 대하여 졸음으로 분류되었는데, ORD 점수는 졸음이 아닌 경우도 있다. 게다가 몇몇의 비디오에서는 비정상적인 얼굴 특징 때문에 높은 ORD 점수를 받은 경우도 있다. 결론적으로 ORD 점수와 머신 비전 얼굴 알고리즘이 불일치 할 경우, 오류 경고 (false alarm)에 의한 오분류 보다 더 많은 오작동이 발생하였다.

오분류의 특성을 보다 잘 이해하기 위해 연구자들은 Fisher's exact test (two-tailed)를 사용하여 epoch 특성 분포가 비디오 epochs의 분류 결과와 다른지를 확인하였다. 운전자 요인들 (epoch 특성)은 만약 비정상 얼굴 특징을 보이고 있다면, 운전자가 다른 사람들과 대화를 하는지, 사시인지 그리고 운전자가 차량 내에서 다른 것을 보고 있는지, 주의 분산이 되었는지, 화면을 보는지, 비정상 행동을 하는지 등의 정보를 포함하고 있다. 분석결과에 대하여 운전자 특성이 고려된 그룹과 특성이 고려되지 않은 두 그룹의 분포를 비교해 보면 상당히 다른 수치를 보이고 있다. ($p=0.013$). Epoch들에 대하여 ORD와 머신 비전 알고리즘이 불일치 한 경우에서 가장 큰 비율 (62%)로 운전자 얼굴 특징이 고려된 것으로 나타났다. 얼굴 알고리즘은 가끔 비정상 얼굴특징 또는 비정상 행동에 대한 올바른 해석에 문제가 될 수 있다.

연구자들은 또한 각각의 epoch에 대해서 교통 조건을 heavy, medium, light 로 분류 하였다. 두 개의 그룹 (heavy or medium vs. light)에 대하여 상당히 다른 결과 분포를 나타내었다. ($p=0.0004$). 이 경우, heavy 또는 medium 교통 조건에서 73%의 epochs 들이 차량 이력 데이터는 졸음운전으로 판별하였지만 ORD 점수는 그렇지 않았다. 이 경우, 알고리즘이 교통 환경에 따른 급작스러운 제동과 차선이탈을 운전자의 상태와 상관없이 졸음으로 인식할 가능성이 있다. 유사하게 heavy 또는 medium 교통조건에서 차량이력데이터가 63%의 epoch들에 대해서 불확실로 분류한 경우도 있다.

각각의 epoch를 살펴보면, 연구자들은 차량 움직임, 급작스러운 제동, 꼬리물기, 또는 다른 위험한 운전 행동 등 비정상 운전을 확인할 수 있었다. 두 그룹(하나 또는 그 이상의 비정상 운전 vs. 정상운전)을 비교해 보면 분류 결과가 유의미하게 달랐다 ($p=0.005$). 대부분의 epochs 리뷰에서 연구자들은 하나 또는 그 이상의 비정상 운전 행동을 확인하였다. 두 그룹의 분류결과에서 epochs가 적어도 하나의 비정상 운전 데이터를 포함하고 있다면, 차량 이력데이터 알고리즘이 졸음운전으로 분류하였다. 대조적으로 차량 이력 데이터 알고리즘이 졸음운전으로 분류하지 않은 경우에도 75%가 하나 또는 그 이상의 비정상 데이터를 포함하고 있었다. 또한 모든 epochs에서 다양한 ORD 점수가 비정상 운전 행동을 포함하고 있었다. 연구자들에 의해 조사된 속성들 중에서 비디오영상의 질, 도로 형태, 조명조건 등은 오분류에 영향을 미친다고 나타나지는 않았다.

6. 결론

이 연구는 차량 이력 데이터 (차량 센서 포함)와 얼굴 비디오 평가와 같은 naturalistic driving data가 운전자의 졸음을 판단하기에 유용하기는 하지만, 현재 아직까지는 일관성 있게 졸음운전을 판별하지는 못한다고 지적하였다. Naturalistic driving data 와 함께, 차량 이력 데이터 알고리즘은 운전 시각과 차선 이탈정보를 사용하여 약간 졸음 (moderately drowsy driving)와 매우 졸음 (very drowsy driving)을

구분 할 수 는 있다. 얼굴 비디오 평가 알고리즘은 거짓 경고와 오분류율이 운전자 경고 시스템에 적용 하기에 부적절 할지라도, 많은 양의 naturalistic data 집단을 구분하기에는 충분할 수도 있다. 이러한 알고리즘들이 naturalistic data 안에서 졸음운전을 구분하는데 도움이 되는 반면, 졸음운전을 일관성 있게 분류 할 수 있는 기술적 향상이 필요하다.

<원문제목> Quantifying Drowsy Driving

<원문출처> NHTSA Traffic Safety Facts Research Note (DOT HS 813 003)