

효율적인 블랙박스 영상 기반 교통 이상 상황 예측을 위한 특징추출 알고리즘의 성능 비교

Performance comparison of feature extraction algorithms for efficient black box video-based accident prediction

박지규
전기컴퓨터공학과
인하대학교
인천광역시, 한국
jkpark97@inha.edu

김덕환*
전기컴퓨터공학과
인하대학교
인천광역시, 한국
deokhwan@inha.ac.kr

요약

최근 블랙박스는 단순히 주행영상을 녹화하는 것 뿐만 아니라, 주행 영상을 분석하여 차선 이탈 감지, 전방 충돌 감지 등의 첨단운전자보조시스템을 지원하는 지능형 블랙박스로 발전하고 있다. 이에 따라 1 인칭 영상을 분석하여 교통 이상 상황을 예측하는 연구가 진행되고 있다. 하지만 블랙박스 영상은 전체 영상 중에 교통 이상 상황이 발생하는 경우가 적은 반면, 이상 상황의 경우의 수는 상당히 많은 긴 꼬리 분포 문제를 가진다. 또한 차량의 움직임에 따라 영상의 배경도 함께 움직여 분석하기 어려운 문제가 있다. 따라서 본 논문에서는 Dense Trajectory 기법을 활용하여 세가지 영상 특징을 추출하고, 이를 활용하여 교통 이상 상황을 예측하는 신경망 모델을 기반으로 세가지 영상 특징 추출 알고리즘의 성능을 비교하였다.

키워드 : Dense Trajectory, Histogram of optical flow, Histogram of oriented gradient, Traffic Accident Detection

1. 서론

첨단운전자보조시스템(ADAS)는 교통사고로 인한 피해를 최소화하기 위해 필수적인 시스템이다. 하지만 국토교통부의 자동차 등록자료 통계에 따르면 ADAS가 장착되지 않은 것으로 추정되는 자가용이 전체의 약 41% 수준이다. 이와 상반되게 트렌드 모니터의 조사에 따르면 블랙박스 설치율은 약 88.9%로 상당히 높은 수준이며 이는 차량의 LIDAR, RADAR, 초음파 등의 센서를 사용하지 않고 블랙박스에서 취득 가능한 1 인칭 영상만을 이용한 사고예측 기술을 필수적임을 보여준다.

이를 위한 주요 과제는 두 가지로 꼽을 수 있다. 첫째, 블랙박스 영상에서의 교통사고 상황은 Positive Sample은 부족한 것에 비해 사고 유형은 무수히 많다. 그러므로 Long-Tail Distribution 문제를 해결할 필요가 있다.

두 번째, CCTV와 같이 사고 객체를 기준으로 3 인칭인 영상과는 다르게 블랙박스 영상은 배경도 같이 움직이므로, Optical Flow와 같은 Motion Vector Descriptor에서 차량, 보행자, 표지판 등과 같은 ROI(Region Of Interest)가 아닌 배경

영역의 움직임 또한 특징으로 추출된다.

따라서 이 논문에서는 배경이 움직이는 블랙박스 영상을 기반으로 교통사고 예측 모델을 위하여 Improved Dense Trajectory 기법[5]을 기반으로 하여 MBH, HOF, HOG 특징을 추출하고 블랙박스 영상내의 교통 이상 상황 탐지에 효율적인 특징을 찾는 것을 목표로 한다.

2. 관련연구

블랙박스의 필요성이 세계적으로 증가함에 따라 1 인칭 영상 기반 사고 예측 연구가 진행되어 왔다. [1]에서는 VANET 환경에서의 실시간 교통사고 탐지를 위하여 Multilayer Neural Network를 활용한 Spatio-Temporal Feature Encoding 방식을 채택하여 긴 꼬리 분포 문제를 해결하고 기존의 연구에서 보행자의 이상 행동을 무시했던 문제를 개선했다. [2]에서는 교통사고의 FOL(Future Object Localization)을 활용한 비지도 학습 기반의 교통 상황 이상탐지 모델을 제안하여 교통사고의 긴 꼬리 분포 문제를 해결하고, LIDAR 기반으로 학습된 현재 차량의 움직임을 추정하는 딥러닝 모델을 활용하여 배경 움직임 문제도 개선했다. [3]은 object detector의 결과인 bounding box와 bird's eye view의

center coordinates, motion vector 를 특징으로 사용하는 Bi-LSTM 모델을 사용하여 사고 확률을 예측하는 모델을 연구했다.

3. 제안한 방법

본 논문에서는 그림[1]과 같이 블랙박스 영상을 15 프레임 단위로 나누고 각 프레임에 Dense Trajectory 추출 방법[5]을 통한 이동 벡터를 특징으로 사용한다. 이동 벡터는 각 프레임간 물체의 이동을 방향과 거리 형태로 나타낸 분포로써, 영상 내 객체들의 움직임을 포함할 수 있기 때문에 블랙박스 영상을 분석하는데 적합하다. 이 때, 추출된 이동 벡터는 MBH(Motion Boundary Histogram), HOF(Histogram Of opticalFlow), HOG(Histogram Of orientedGradient) 이다. 이후 신경망 계층과 SoftMax 함수를 활용하여 정상/비정상 확률을 추정하는 모델을 제안한다.

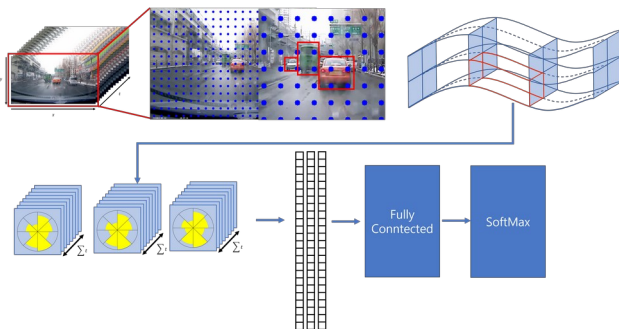


그림 1 Dense Trajectory 기법 특징 추출 및 분류 모델

3.1 optical flow 추출

Optical flow 는 두 프레임 간 물체의 움직임 패턴을 의미한다. 이는 이전 프레임과 다음 프레임 간 픽셀이 이동한 방향과 거리 분포로써, 영상 내 물체가 어느 방향으로 얼마나 움직였는지 파악할 수 있다. Optical flow 는 코너 검출 등과 같은 검출기를 통과시켜 특정 좌표에서 계산하는 방식과 모든 점에서 계산하는 dense 방식이 있다. 본 논문에서는 dense 방식을 사용하여 추출된 optical flow 를 기반으로 trajectory 를 계산하고 특징을

추출한다.

3.2 특징점 샘플링

Trajectory 를 계산하기 위해 첫 번째 프레임에서 시작점이 될 특징점 샘플링이 필요하다. 첫 번째 프레임에서 일정 간격으로 샘플 포인트를 격자 형태로 생성하고, shi and tomasi 검출기를 통해 검출된 에지와 겹치는 영역을 제거한다. 이후 모든 프레임에서 차원을 통일시키는 과정이 필요하기 때문에, optical flow 의 magnitude 를 내림차순으로 정렬한 뒤, N 개의 샘플만을 유지하고 나머지는 삭제한다. 본 논문에서는 32 픽셀의 간격으로 격자를 생성하고 200 개의 샘플만을 유지하였다.

3.3 Trajectory 추출

3.2 에서 추출된 특징점을 시작으로 optical flow 의 이동 벡터를 활용하여 15 프레임간의 물체 움직임을 추정하는 Trajectory 를 추출한다. 먼저 극좌표계로 구성된 optical flow 를 직교좌표계로 변환하고 median filter 를 적용시킨다. 특징점 좌표값에 optical flow 값을 15 프레임에 걸쳐 더하여 좌표의 변화를 구할 수 있다.

3.4 영역별 특징 계산

Feature point 를 Trajectory 를 따라서 그림[5]와 같이 해당 Trajectory 중심점을 기준으로 가로, 세로, 시간 축을 기준으로 sub-region 을 생성한다. 가로, 세로 축은 σ , 시간 축은 τ 개로 나뉜다. 본 논문에서는 $\sigma=2, \tau=3$ 이다. 각 sub-region 별로 특징을 추출한 뒤 결합되어 사용되며 다음은 각 sub-region 내부 프레임에서 특징들이 어떻게 계산되는지 설명한다.

3.4.1 MBH(Motion Boundary Histogram)

MBH 는 Optical Flow 의 가로축 변화량과 세로축 변화량을 기반으로 계산된다. 각 변화량을 8 개의 구간으로 나누어 Histogram 을 추출하고 Histogram 의 빈도를 특징으로 사용한다.

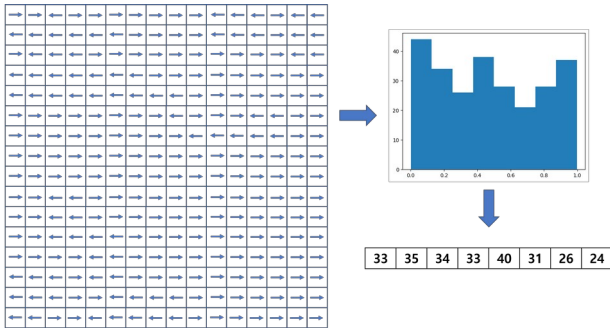


그림 2 MBH 알고리즘

3.4.2 HOF(Histogram Of OpticalFlow)

HOF는 MBH와 비슷하지만 Optical Flow의 방향과 거리 벡터를 활용한다는 차이점이 있다. 각

방향을 $0, \frac{4}{\pi}, \frac{4}{\pi}, \frac{2}{\pi}, \frac{2}{\pi}, \frac{6}{\pi}, \dots$ 와 같이 8 개의

구간으로 나누어 Histogram을 추출하고, 빈도를

	AUC(Area Under Curve)	차원 수
MBH	56.1	6,528
HOF	56.9	3,664
HOG	56.5	14,688
All	57.7	24,880

표 1 특징 추출 알고리즘별 AUC 구할 때 거리를 가중치로

사용한다.

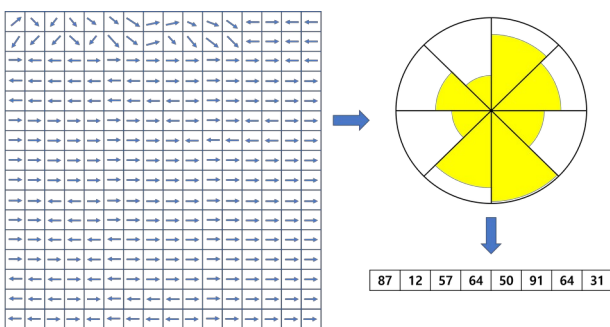


그림 3 HOF 알고리즘

3.4.3 HOG(Histogram Of Oriented Gradient)

HOG는 입력 영상의 각 픽셀 gradient의 거리와 방향을 구하고, 영상을 여러 구간으로 나누어 각 구간별로 Histogram을 구하는 방식으로 구해진다.

이 때, HOF와 마찬가지로 거리를 가중치로 사용한다.

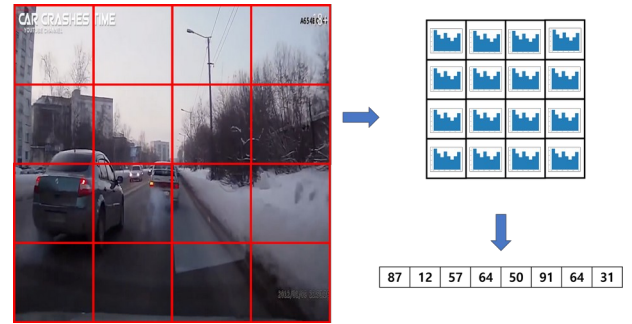


그림 4 HOG 알고리즘

각 특징은 sub-region 상에서 아래의 [그림 5]와 같이 프레임별로 추출된 뒤, 모두 합산된다. 이후 신경망 계층을 통과시키기 위해 결합하여 사용한다.

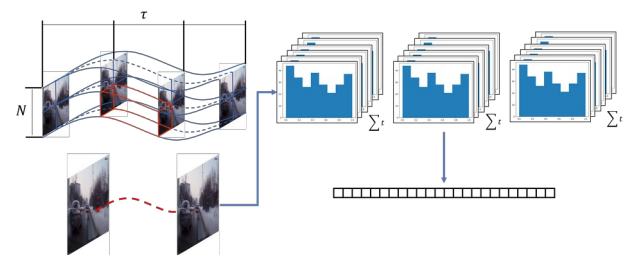


그림 5 Sub-Region 상의 특징 추출 및 결합

4. 실험방법

4.1 데이터셋

사용하는 데이터셋은 DoTA(Detection of Traffic Anomaly)로 교통 사고 시점에 대하여 어느 프레임에서 이상 상황인지 라벨링 되어있는 블랙박스 영상이다. 있다.

본 데이터셋은 총 4,677 개의 영상으로 구성되어 있고 각 영상은 약 10 분으로 8~10 번의 교통 사고 영상과 정상 주행 영상이 포함되어 있다.

데이터셋은 어느 시점에서 이상 상황인지 뿐만 아니라, 어떤 이상 상황인지, 각 프레임별 주요 객체(차량, 보행자 등)의 경계 상자 등도 포함되어 있다.

본 논문에서는 영상의 특징을 추출하고 신경망 계층을 활용하여 학습하여 어느 시점이 비정상 상황인지 판별하는 것이 목표이다. 따라서 전방 영상에서 비정상 상황의 전조 증상이 명확하게 보이는 73 개의 영상을 사용하였다.

4.2 사고예측 정확도 측정

5. 결론

본 논문은 블랙박스 영상 기반 교통 이상 상황 인식을 위한 Dense Trajectory 기반 특징 추출을 진행했다. MBH, HOF, HOG의 세 가지 특징을 추출하였고 이 특징들을 비교하기 위해 완전 연결 신경망과 SoftMax 계층을 거쳐 정상/비정상 여부 식별 모델을 시켰다. 학습을 위해 73개의 영상을 15개의 프레임별로 특징을 추출하여 약 130만 개의 데이터셋을 확보하고 70%의 데이터를 학습 데이터셋으로, 30%의 데이터셋을 검증 데이터셋으로 사용하도록 계층적 샘플링하였다. 학습 결과는 표 [1]과 같다. HOF, HOG, MBH 순으로 AUC가 높지만 모두 비슷한 양상을 보이는 것을 확인할 수 있다. 그러나 차원 수가 다르다는 것에 주목할 필요가 있다. HOF는 HOG보다 약간 높은 정확도를 보이지만 특징의 차원 수는 약 25%에 불과하다.

또한 [2]에서는 LIDAR 기반으로 학습된 딥러닝 모델을 인식하여 얻은 현재 차량의 위치 정보를 기반으로 학습하여 60%의 AUC를 달성했다. 이 논문은 영상 데이터만으로 교통 이상 상황을 추정할 수 있다는 것을 확인했다. 따라서 본 논문에서 추출한 특징과 현재 차량의 위치 추정 알고리즘을 결합하는 연구의 가능성을 확인할 수 있어, 이를 통해 블랙박스 영상 분석 발전에 기여할 것이다.

Acknowledgement

이 논문은 2023년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0020536, 2023년 산업혁신인재성장지원사업)

본 연구는 2023년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음(69585-3)

본 과제(결과물)는 교육부와 한국연구재단의 재원으로 지원을 받아 수행된 3단계 산학연협력 선도대학 육성사업(LINC 3.0)의 연구결과임.

참고문헌

[1] Zhili Zhou, Xiaohua Dong, Zhetao Li, Keping Yu, Chun Ding, Yimin Yang "Spatio-Temporal

Feature Encoding for Traffic Accident Detection in VANET Environment" IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS

[2] Yu Yao, Mingze Xu, Yuchen Wang, David J. Crandall, Ella M. Atkins "Unsupervised Traffic Accident Detection in First-Person Videos" Computer Vision and Pattern Recognition

[3] Hyeon-Cheol Son, Da-Seul Kim, Sung-Young Kim "Vehicle-Level Traffic Accident Detection on Vehicle-Mounted Camera Based on Cascade Bi-LSTM" Journal of JAITS

[4] Heng Wang, Alexander Kläser, Cordelia Schmid, Cheng-Lin Liu "Dense Trajectories and Motion Boundary Descriptors for Action Recognition" International Journal of Computer Vision