

## 교통사고 영상 분석을 통한 과실 판단을 위한 딥러닝 기반 방법 연구

# Research on Deep Learning-Based Methods for Determining Negligence through Traffic Accident Video Analysis

이서영 · 유연휘 · 박효경 · 박병주 · 문일영\*

한국기술교육대학교 컴퓨터공학과

Seo-Young Lee · Yeon-Hwi You · Hyo-Gyeong Park · Byeong-Ju Park · Il-Young Moon\*

Department of Computer Engineering Korea University of Technology and Education, Chungcheongnam-do, 31253, Korea

### [요 약]

자율주행 차량에 대한 연구가 활발하게 이뤄지고 있다. 자율주행 차량이 등장함에 따라 기존의 차량과 자율주행 차량이 공존하는 과도기가 올 것이며, 이러한 과도기에는 사고율이 더욱 높아질 것이라 예상된다. 현재 교통사고 발생 시 손해보험협회의 ‘자동차 사고 과실 비율 인정기준’에 따라서 과실 비율을 측정한다. 그러나, 발생한 사고가 어떠한 유형의 사고인지 조사하는 데 소모되는 비용이 매우 크다. 또한 이미 과실 비율 책정이 완료된 사례에 대해서도 재심의를 요구하는 과실 비율 분쟁도 늘어나는 추세이다. 이러한 시간적, 물적 비용을 줄이기 위해 자동으로 과실 비율을 판단하는 딥러닝 모델을 제안하고자 한다. 본 논문에서는 ResNet-18 이미지 분류 모델과 TSN을 통한 비디오 행동 인식을 통해 사고 영상을 바탕으로 과실 비율을 판단하고자 한다. 모델이 상용화된다면, 과실 비율을 측정하는데 소요되는 시간을 획기적으로 단축할 수 있다. 또한 피의자에게 제공할 수 있는 과실 비율에 대한 객관적인 지표가 생기므로 과실 비율 분쟁도 완화될 것으로 기대된다.

### [Abstract]

Research on autonomous vehicles is being actively conducted. As autonomous vehicles emerge, there will be a transitional period in which traditional and autonomous vehicles coexist, potentially leading to a higher accident rate. Currently, when a traffic accident occurs, the fault ratio is determined according to the criteria set by the General Insurance Association of Korea. However, the time required to investigate the type of accident is substantial. Additionally, there is an increasing trend in fault ratio disputes, with requests for reconsideration even after the fault ratio has been determined. To reduce these temporal and material costs, we propose a deep learning model that automatically determines fault ratios. In this study, we aimed to determine fault ratios based on accident video through a image classification model based on ResNet-18 and video action recognition using TSN. If this model commercialized, could significantly reduce the time required to measure fault ratios. Moreover, it provides an objective metric for fault ratios that can be offered to the parties involved, potentially alleviating fault ratio disputes.

**Key word** : Action recognition, Computer vision, Deep learning, Image classification, Video analysis.

<http://dx.doi.org/10.12673/jant.2024.28.4.559>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 1 August 2024; Revised 27 August 2024

Accepted (Publication) 29 August 2024 (30 August 2024)

\*Corresponding Author, Il-Young Moon

Tel: \*\*\* - \*\*\*\* - \*\*\*\*

E-mail: [iymoon@koreatech.ac.kr](mailto:iymoon@koreatech.ac.kr)

## I. 서론

자율주행은 운전자가 직접 조작하지 않고도 차량 내 외부의 환경을 인지하여 그에 따른 차량 제어를 통해 차량 스스로 목적지까지 주행하는 기술을 말한다. 이러한 자율주행 기술은 추후 인간적인 실수에 의한 다양한 사고 감소, 교통체증 완화 등의 긍정적인 효과를 가져올 것이다[1].

미국자동차기술회 SAE (Society of Automotive Engineers) 에서 자율주행 기술을 Level 0, 비자동화부터 Level 5, 완전 자율주행 단계까지 총 6단계로 세분화하여 정의했다. 현재 상용화된 차량에 사용되는 단계는 Level 2~3, 부분 자동화 단계로 자율주행 시스템은 차선이탈 방지, 앞차 간격 유지 등 운전자를 단순히 보조해주는 역할을 하고 있다. 아직은 운전자가 운전대를 조작하고, 상시 모니터링을 필수적으로 해야 하는 단계일 뿐이다[2]. 자율주행 단계가 발전하게 되면 운전자가 직접 운전해야 하는 Level 0~2의 Low level 자율주행 차량과 차량이 직접 제어하는 Level 3~5의 High level 자율주행 차량이 공존하는 과도기가 올 것이다. 이러한 과도기 속에서 차량간 사고율은 더욱 높아질 것이고 이에 알맞은 정책의 발전이 필수적이다[3], [4]. 현재 교통사고 발생 시 과실 비율은 경찰의 조사 내용 및 블랙박스 내용을 바탕으로 손해보험사 보상직원이 정하는 형태이다. 이는 과실 비율이 측정되기까지 최소 4일에서 7일까지 소요되며, 과실 비율을 측정하기 위해 소모되는 인적, 물적 자원이 매우 크다. 이러한 과정을 통해 과실 비율을 측정했음에도 불구하고 보상직원이 결정한 결과에 승복하지 않고, 재심을 요구하는 과실 비율 분쟁도 늘어나는 추세이다. 과실 비율 분쟁은 소송이 제기되어 법원 판결까지 가는 경우가 많아, 추가로 소모되는 자원이 매우 크다[5]. 본 논문에서는 이러한 과실 비율에 따른 분쟁을 감소시키기 위해 인공지능을 활용한 과실 비율 측정 기술에 관해 연구하고자 한다. 사람이 직접 조사하여 과실 비율을 측정하는 것이 아니라 데이터에 의해 과실 비율을 명확하게 측정할 수 있다면, 과실 비율을 측정하는 데 소모되는 자원이 현저하게 줄어들 것이다. 또한 인간적인 요소 없이 데이터 기반 의사결정이 이루어졌기 때문에 결과에 따른 재심의가 줄어들 것으로 판단된다.

## II. 이론적 배경

### 2-1 인공지능 기반 차량 사고 과실 비율 관련 연구

교통사고 발생 시 손해보험협회의 자동차 사고 과실 비율 인정기준에 따라서 과실 비율을 측정한다. 현재 상황에서 과실 비율 인정기준에서 제시하는 유형을 파악하기 위해 경찰 및 보험사가 사고 당시의 상황을 조사하는 과정이 필수적이다. 일련의 과정은 오랜 기간이 소요되며, 피의자가 산정된 과실 비율에 불만을 가져 분쟁이 일어나는 경우도 빈번하다. 이러한 과실 비율 문제를 해결하기 위해 국내외로 다양한 연구

가 진행되고 있다. 과실 비율 산정은 정확한 사고 유형만 판별할 수 있다면, 인정기준에 따라서 자동으로 책정되는 형태이다. 초기에 과실 비율을 판별하기 위해 정형화된 텍스트와 분류 모델을 활용했다[6]. 이는 정형화된 텍스트 데이터에 랜덤 포레스트와 같은 분류 모델을 활용하여 정확도가 높지만, 사람이 직접 텍스트 데이터를 정형화하는 과정이 필수적이다.

다른 연구에서는 사람이 직접 데이터를 정형화하는 과정을 생략하기 위해 CNN과 RNN을 활용했다. CNN을 통해 영상 데이터의 특징을 추출하고, RNN을 통해 추출된 데이터가 어떠한 유형으로 분류되는지 이진 분류를 진행했다[7]. 타 모델은 한가지 상황에 대해서만 분류를 진행하였으나, 본 논문에서는 과실 판단 예측을 다양한 상황에서 적용하고자 한다.

### 2-2 Video Action Recognition

딥러닝 기법을 사용하여 비디오를 분석하기 위한 다양한 태스크 들이 연구되고 있다. 행동 인식(Action Recognition), 시간적 행동 검출(Temporal Action Detection), 시공간적 행동 검출(Spatial Temporal Action Detection) 등이 이 분야의 대표적인 연구과제이며 각각은 각기 다른 수준의 세부 정보를 제공한다. 행동 인식 (Action Recognition)은 비디오 내에서 어떤 행동이 발생했는지에 중점을 두는 작업이다. 이를 위한 대표적인 모델로는 TSN, TSM (Temporal Shift Module), TRN (Temporal Relation Network), C3D (Convolutional 3D Network), Vision Swin Transformer 등이 있다. 각 모델은 다른 특성을 가지며 성능, 학습 자원, 복잡성 등의 측면에서 장단점을 지닌다.

TSN은 그림 1과 같이 비디오를 여러 구분(segment)으로 나누고, 각 부분에서 대표 프레임을 샘플링하여 행동을 인식하는 모델이다. 모델의 특징으로는 비교적 간단하여 적은 자원으로 좋은 성능을 낼 수 있으나 세밀한 수준의 액션 인식은 어려운 편이다. Vision Swin Transformer는 높은 성능을 보여주지만, 자원이 제한적인 환경의 경우 적합하지 않을 수 있다는 한계점이 있다. C3D의 경우 공간 및 시간 정보를 동시에 학습해야 하는 경우 높은 성능을 보여주나 복잡성이 높은 편이다[8]-[12].

이 외에도 다양한 모델이 존재하기에 활용하는 데이터 세트

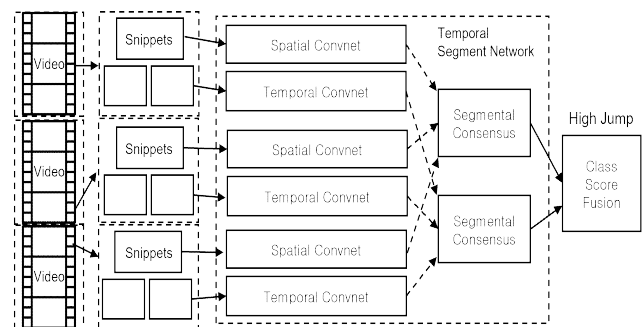


그림 1. TSN (Temporal segment network) 모델의 구조  
Fig 1. Architecture of the temporal segment network model.

의 특성 및 사용할 수 있는 자원의 양, 목표 성능 등에 따라 적절한 모델의 선택이 요구된다. 본 논문에서는 사고 영상 비디오 분석을 위해 행동 인식 기법의 접근을 하였고, TSN 모델을 사용하여 실험을 진행하였다.

### III. 연구 방법

#### 3-1 AI-HUB 데이터 세트

AI hub는 인공지능 기술 개발에 필요한 AI 인프라를 지원해주는 플랫폼으로, 다양한 분야의 인공지능 학습용 데이터를 무료로 얻을 수 있다. 실험을 위한 데이터는 AI hub의 교통사고 영상 데이터를 활용했다[13]. 교통사고 영상 데이터는 교통사고의 과실 비율 측정 인공지능 구축의 목적으로 생성된 데이터 세트로, 총 21,895건의 교통사고 영상 데이터와 3,284,250장의 이미지 데이터가 라벨링 되어 있다. 본 논문에서는 사고 대상이 차대 차인 경우의 4가지 사고 장소의 영상 데이터로 데이터 세트를 추출하여 구성하였다. 전처리 과정에서 10건 미만의 경우는 삭제하였고, 데이터 세트의 구성과 각 경우에 대한 과실 비율 예시의 표 1과 같으며, 더욱 다양한 경우들이 존재하나 전체 경우가 아닌 본 논문의 실험에서 사용한 사고 장소 특징 및 진행 방향의 사례만 나열하였다.

#### 3-2 제안하는 시스템 구조

제안하는 시스템 구조는 그림 2와 같다. 먼저 사고 영상 비디오가 입력되면 딥러닝 기반의 키 프레임 추출 모델을 통해 비디오의 대표 이미지가 추출된다. 해당 이미지를 기반으로 사고 장소가 인식되면 각 장소에 따른 특징 추출 모델, 객체 A 진행 방향 예측 모델, 객체 B 진행 방향 예측 모델이 결정된다. 해당 모델에 비디오가 입력되면 각각의 출력이 도출되고, 출력 값을 결합하여 표 1과 같은 정보를 기반으로 과실 비율을 설정하게 된다.

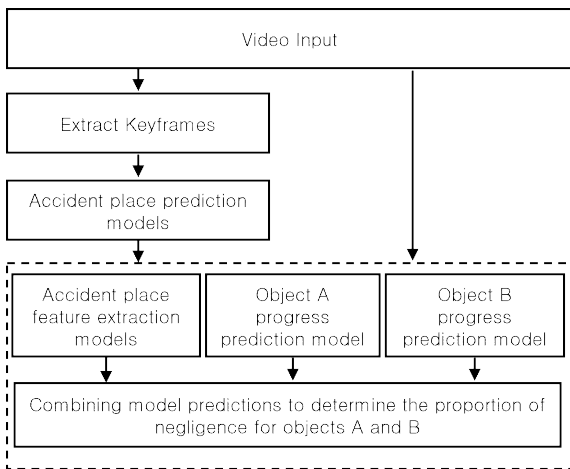


그림 2. 제안하는 과실 비율 산정 시스템 구조  
Fig 2. Architecture of autonomous driving algorithm

표 1. AI Hub 교통사고 영상 데이터 세트 구성의 예시(사고 대상 유형: 차대차)

Table 1. Example of AI hub traffic accident video dataset configuration (Accident target type: car-to-car).

Accident Place	Accident Place Features	Object A Progression	Object B Progression	Fault rate A	Fault rate B	
Parking Lot	Parking areas and walkways	Go straight down the aisle	Going straight	30	70	
			Driving in reverse	25	75	
Driveways and Non-drive way areas	Pulling into a driveway Entering a non-drive way place	Drive straight on the driveway	Centreline encroachment	0	100	
			Entering the right turn	20	80	
Round-About	Single-lane type	Entering1	Rotation3	80	20	
			Rotation1	Changing paths	40	60
			Entering2	Entering4	60	40
			Rotation2	Large Entering	10	90
Equal width roads	Two-lane type	Changing paths	Entering1	30	70	
			Entering2	Entering3	30	70
T-Interse ction	Main road and side road	Go straight on the road1	Turn left on the right road	30	70	
			Turn left on the left road	20	80	
			Turn left on side road	10	90	
			Turn right on side road	20	80	
	Main road and side road	Go straight on the main road	Turn left on side road2	40	60	
			Turn left on main road1	35	65	
			Turn right on main road	60	40	
			Turn left on the main street	30	70	

#### 3-3 사고 영상 장소 예측 모델

사고 영상 장소를 예측하기 위해서는 CNN 모델 기반의 이미지 분류 기법을 사용하였다. 먼저 딥러닝 기반 키 프레임 추출 라이브러리인 Katna에서 제공하는 API를 사용하여 사고 영상으로부터 키 프레임 이미지를 추출하여 데이터를 구성하였다[14]. 모델로는 ResNet-18을 ImageNet 1K에서 사전 학습한 모델을 불러와 전이 학습하였다. 분류를 위한 클래스의 개수는 표 1과 같은 사고 장소 유형의 개수인 4로 설정하였다[15],[16].

#### 3-4 Video action recognition

영상 내에서 사고 객체 A와 B의 진행 방향을 식별하고, 사고 장소 특징을 분류하기 위해 비디오 행동 인식 모델을 사용하여 시스템을 구성하였다. 비디오 기반의 액션 인식을 위해서는 OpenMMLab 프로젝트 일부로 개발된 MMAction2 프레임워크



그림 3. 데이터 파이프라인을 통한 비디오 입력 데이터의 가시화

Fig 3. Visualization of input video results from data pipelines.

를 사용하였다[17]. 모델 구성 요소는 크게 인식기(Recognizer), 백본(Backbone), 분류 작업을 위한 구성인 CLS\_Head, 로컬라이저(Localizer)의 4가지로 분류된다. 백본은 피쳐 맵을 추출하기 위한 ResNet, BN Inception 등의 FCN (Fully Connected Network) 네트워크를 의미하고, 로컬라이저는 위치 추정과 관련한 작업을 위한 모델이다.

모델 학습을 위해 ResNet-50을 백본으로 Kinetics 400 데이터 세트에서 전체 TSN 네트워크에 대해 사전 학습된 모델을 사용하였다[18],[19]. 이를 AI Hub 데이터 세트를 기반으로 자체적으로 구성된 데이터 세트에 파인 튜닝하였다. CLS\_Head는 표 2에서 확인할 수 있는 데이터 세트의 클래스 개수인 Num\_Classes에 따라 변경하였다.

각 비디오는 장소 특징 인덱스, 객체 A 진행 방향 인덱스, 객체 B 진행 방향 인덱스의 3가지 정보를 레이블링하였고 사고 장소를 예측하는 모델을 훈련하였다. 입력된 비디오는 시스템에 내장된 파이프라인을 통해 그림 3과 같이 DecordDecode (1080, 1920), Resize(256, 455), MultiScaleCrop (256, 256), Resize(244, 244), Flip(244, 244)의 전처리 과정을 거친다. 이후 TSN 모델에서 구분 분할, 백본 네트워크를 통한 특징 추출 그리고 스니펫(Snipet) 수준 예측에 대한 결합 등의 과정을 통해 비디오 수준의 예측을 형성한다.

## IV. 실험

### 4-1 실험 세부 사항

사고 장소로는 주차장(또는 차도가 아닌 장소), 차도와 차도가 아닌 장소, 회전 교차로, T자형 교차로를 선정하여 각각의 사고 장소에 대해 모델 훈련을 수행했다. 모델 학습을 위해 구성된 데이터 세트의 구성은 표2와 같다.

학습은 크게 두 가지로 구성된다. 우선 사고 장소 분류를 위한 모델로 ResNet-18을 선정했다. ResNet 모델은 레이어(Layer)

표 2. 사고 장소에 따른 데이터 세트 정보

Table 2. Dataset information on accident place.

Accident Place	Train	Validation	Test
Parking Lot	503	69	73
Driveways and Non-driveway areas	194	26	26
Round About	231	33	33
T-Intersection	893	160	162
Total	1,821	288	294

표 3. 모델 학습을 위한 하이퍼 파라미터

Table 3. Hyper-parameters for model training.

Type	Model	Epochs	Optimizer	Learning Rate	Momentum
Accident Place (Image)	ResNet-18	25	SGD	0.001	0.9
Action Recognition (Video)	TSN	60	SGD	0.005	0.9

가 깊어질수록 계속되는 미분으로 가중치(weight)의 영향이 적어져 정확도가 떨어지는 현상을 방지한 알고리즘이다. 레이어가 너무 깊으면 학습 속도에서 불리함이 있어 18층의 레이어를 지나는 ResNet-18 모델을 선택했다. 사고 장소 이미지 분류 모델의 훈련의 파라미터는 표 3과 같이 25 Epochs로, SGD Optimizer, 0.001 Learning Rate, 0.9 Momentum으로 설정하였다. 또한 비디오 액션 인식을 위한 모델은 TSN을 사용했다. 이는 기존 액션 인식 분야에서 성능 및 속도 면에서 전반적으로 높은 성능을 보여주었으며, 데이터 세트의 크기 문제와 다양성 부족 문제에 대응이 가능하다는 장점이 있기 때문이다. 이에 사전 학습된 모델인 tsn\_imagenet\_pretrained-r50\_8xb32-1x1x3-100e\_kinetics\_400-rgb를 파인 튜닝하여 사용했다[18]. 학습 파라미터는 역시 표 3과 같이 32 Batch Size, 60 Epochs, SGD Optimizer, 0.005 Learning Rate, 0.9 Momentum으로 설정했다. 평가를 위해 사용된 지표는 비디오 액션 인지 분야에서 보편적으로 사용되는 Top-1



표 4. 실험 환경

Table 4. Experiment environment.

Operating System	CUDA	Pytorch	MMEngine	GPU
Ubuntu 20.04	11.6	1.13.1	0.10.4	NVIDIA A5000 24GB * 2

, Top-5, Mean Top-1 Accuracy를 사용하였다. Top-1 Accuracy는 예측 확률이 가장 높은 클래스가 실제 레이블과 일치하는 횟수의 비율을 측정하며, Top-5 Accuracy는 실제 라벨이 상위 5개 예측에 포함되는 빈도를 측정한다. Mean Top-1 Accuracy는 각 클래스 내 정확도를 개별적으로 고려한 후 정확도의 평균을 계산하여 불균형 데이터 세트 처리에 특히 유용한 지표이다. 또한, 성능의 직관적인 결과를 확인하기 위해 전체 테스트 데이터에서 예측값과 실제값이 일치하는 비율을 나타내는 Correct rate를 도출했다.

#### 4-2 실험 환경

비디오 액션 인식(Action recognition)을 위한 모델 훈련 및 테스트는 다양한 모델을 지원하는 Pytorch 기반의 오픈소스 도구인 MMAction2 프레임워크 사용하여 수행했다. 실험은 표 4와 같은 환경에서 수행했다.

#### 4-3 장소 예측 모델 실험 결과

먼저 장소 예측 모델 훈련을 위해서는 약 20분이 소요되었고, 테스트 데이터 세트를 통한 최대 정확도는 약 72%의 결과를 보여주었다. 학습 및 검증 데이터에 대한 Loss는 모두 대략 0.78 정도의 값을 보여주었다.

#### 4-4 Video Action Recognition 실험 결과

사고 장소 및 분류를 위한 특성에 따라 모델 훈련을 진행하였고 결과는 표 5와 같다. 모델 유형에서  $PL$ 은 주차장,  $Rd$ 는 도로와 도로가 아닌 장소,  $RAB$ 는 회전 교차로,  $Tjunc$ 는 T자형 교차로를 의미한다. 또한 아래 첨자  $p$ 는 장소 특징 분류 모델,  $a, b$ 는 각 사고 객체의 진행 방향 분류 모델을 의미한다. Num Class는 사고 장소 특징 또는 진행 방향의 분류를 위한 클래스 경우의 수를 나타낸다. Mean Top-1 지표를 기준으로 전체 모델은 평균적으로 약 0.6 정도의 성능 결과를 보여주었다. 전체 성능 지표를 고려했을 때 가장 좋은 성능을 보여주었던 모델 종류는  $PL_b$ 의 경우였는데 학습 데이터셋의 클래스 분포가 비교적 균일하고 클래스 개수가 단순하기 때문에 좋은 성능을 보여준 것으로 여겨진다. 전반적으로 분류 클래스 개수가 적을 때 성능이 더 좋은 경향성을 보여주었다.

성능이 가장 좋지 않은 경우는  $RAB_a$ 으로 Mean Top-1 지표를 기준으로 약 0.42의 성능을 보여주었다. 낮은 성능의 이유는 회전 교차로 상황에서 객체의 진행 방향을 나타내는 행동 타입

표 5. 모델 종류에 따른 실험 결과

Table 5. Experimental results based on model type.

Model Type	Num Class	Top-1	Top-5	Mean Top-1	Correct Rate
$PL_b$	2	0.8767	1.0000	0.8953	87.67 %
$Rd_b$	3	0.5000	1.0000	0.6453	50.00 %
$Rd_p$	2	0.7308	1.0000	0.5333	100.0 %
$RAB_p$	2	0.5758	1.0000	0.5625	57.58 %
$RAB_a$	5	0.4242	1.0000	0.3150	42.42 %
$RAB_b$	5	0.5455	1.0000	0.5094	54.54 %
$Tjunc_p$	2	0.7222	1.0000	0.7245	72.22 %
$Tjunc_a$	7	0.6111	0.9753	0.5630	61.11 %
$Tjunc_b$	7	0.5000	0.9568	0.4819	50.00 %

이 가지적으로 차이가 크지 않아 인식률에 영향을 미친 것으로 여겨진다. Correct Rate의 경우 직관적인 이해를 주는 장점이 있었으나, 데이터 라벨의 분포가 불균일한 경우 모델 성능을 정확하게 보여주기가 어려운 측면도 있었다.  $Rd_p$ 의 경우 데이터 라벨이 한쪽 클래스에 편향되었을 때 해당한다. 따라서 모델 성능 검증을 위해서는 다양한 지표를 통해 성능 결과를 비교 및 분석하는 시도가 요구된다.

## V. 결 론

본 논문은 인공지능 기술인 키 프레임 추출 모델, 이미지 분류 모델 및 동영상 액션 인식 모델을 사용하여 사고 영상의 과실 비율을 판정하는 시스템을 제안하였다. 이를 위해, 사고 장소 판별을 위해 사고 영상에서 추출된 키 프레임을 기존 이미지 분류 모델에 전이 학습하여 예측 모델을 구현하였다. 또한 사고 장소의 특징 및 객체 A, B의 진행 방향 판별을 위해 비디오 액션 모델인 TSN 모델을 논문에서 구성한 데이터 세트에 대해 과인 튜닝하였다.

인공 지능 기술을 기반으로 영상 기반의 과실 판별을 자동으로 수행하는 여러 연구가 이어지고 있다. 영상 분야에서 높은 성능 및 활용성을 보이는 딥러닝 모델의 등장하는 만큼, 원하는 목표 및 성능에 맞도록 데이터 세트를 가공하고 적절한 모델을 선정한다면 성능 향상 및 실용 수준 애플리케이션의 개발이 가능할 것이다.

## Acknowledgments

본 과제(결과물)는 2024년도 교육부의 재원으로 한국연구재단의 지원을 받아 수행된 지자체-대학 협력기반 지역혁신 사업의 결과입니다(2021RIS-004).

이 연구는 과학기술정보통신부의 재원으로 한국지능정보사

회진홍원의 지원을 받아 구축된 “교통사고 영상 데이터”를 활용하여 수행된 연구입니다. 본 연구에 활용된 데이터는 AI 허브(aihub.or.kr)에서 다운로드 받으실 수 있습니다.

## References

- [1] H. Thadeshwar, V. Shah, M. Jain, R. Chaudhari and V. Badgujar, “Artificial intelligence based self-driving car,” in *2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP)*, India, pp. 1-5, 2022. DOI: 10.1109/ICCCSP49186.2020.9315223.
- [2] K. M. Moon, “Takeover as an ethical problem -level 3 automated driving car and norms of coordination-,” *Philosophy-Thought-Culture*, No. 38, pp. 25-45, 2022. DOI : 10.33639/ptc.2022..38.002.
- [3] J. D. Kim, “A legal study on the liability of autonomous car accident,” *Journal of Business Administration & Law*, Vol. 32, No. 2, pp. 1-29, 2022. Retrieved from <https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART002819798>.
- [4] K. C. Lee, “Administrative legal consideration on legal risks and responsibilities of self-driving vehicles,” *National Public Law Review*, Vol. 19, No. 1, pp. 149-191, 2023. DOI : 10.46751/nplak.2023.19.1.149.
- [5] H. S. Lee and D. O. Kang, “Consideration on traffic accident investigation and the role of detectives,” *The Journal of Comparative Law*, Vol. 24, No. 1, pp. 223-268, 2024. Retrieved from <https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART003079068>.
- [6] E. Han, G. O. Park, H. J. Kang, Y. S. Lee and L. S. Yun. “Development of a model for calculating the negligence ratio using traffic accident information,” *The Journal of the Korea Institute of Intelligent Transport Systems*, Vol. 21, No. 6, pp. 36-56, 2022. Retrieved from <https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART002916102>.
- [7] J. W. Yoon, H. J. Jo, Y. H. Kim, M. K. Yoon and H. S. Kim. “Research on basic artificial intelligence model for black box video analysis and vehicle accident fault determination using c-rnn model,” *Proceedings of KIIT Conference*, Vol. 2023, No. 11, pp. 669-673, 2023. Retrieved from <https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE11652130>.
- [8] L. Wang, Y. Xiong, Z. Wang, Y. Qiao, D. Lin, X. Tang, and L. Van Gool, “Temporal segment networks: towards good practices for deep action recognition,” *Computer Vision -- ECCV 2016*, Vol. 9912, Online, Sep. 2016. DOI: [https://doi.org/10.1007/978-3-319-46484-8\\_2](https://doi.org/10.1007/978-3-319-46484-8_2).
- [9] J. Lin, C. Gan, and S. Han, “Sm: temporal shift module for efficient video understanding,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul: Korea, pp. 7083-7093, Oct. 2019. DOI: <https://doi.org/10.48550/arXiv.1811.08383>.
- [10] B. Zhou, A. Andonian, A. Oliva, and A. Torralba, “Temporal relational reasoning in videos,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich: Germany, pp. 803-818, Sep. 2018. DOI: [https://doi.org/10.1007/978-3-030-01246-5\\_49](https://doi.org/10.1007/978-3-030-01246-5_49).
- [11] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, “Learning spatiotemporal features with 3d convolutional networks.” in *Proceedings of the IEEE International Conference on Computer Vision*, Santiago: Chile, pp. 4489-4497, Dec. 2015. DOI: <https://doi.org/10.48550/arXiv.1412.0767>.
- [12] Z. Liu, Y. Lin, Y. Cao, et. al, “Swin transformer: hierarchical vision transformer using shifted windows,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, Montreal: Canada, pp. 10012-10022, Oct. 2021. DOI: <https://doi.org/10.48550/arXiv.2103.14030>.
- [13] AIHub. Traffic accident video data [Internet]. Available: <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&dataSetSn=597>.
- [14] Katna. Katna documentation [Internet]. Available: <https://katna.readthedocs.io/en/latest/>.
- [15] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas: NV, pp. 770-778, June. 2016. DOI: <https://doi.org/10.48550/arXiv.1512.03385>.
- [16] O. Russakovsky, J. Deng, H. Su, et. al, “Imagenet large scale visual recognition challenge,” *International Journal of Computer Vision*, Vol. 115, No. 3, pp. 211-252, March 2015. DOI: <https://doi.org/10.1007/s11263-015-0816-y>.
- [17] MMAAction2. MMAAction2 Documentation [Internet]. Available:[https://mmaaction2.readthedocs.io/en/latest/model\\_zoo/recognition.html](https://mmaaction2.readthedocs.io/en/latest/model_zoo/recognition.html).
- [18] MMAAction2. TSN Results and Models on Kinetics-400 [Internet]. Available:[https://download.openmmlab.com/mmaaction/v1.0/recognition/tsn/tsn\\_imagenet-pretrained-r50\\_8xb32-1x1x8-100e\\_kinetics400-rgb/tsn\\_imagenet-pretrained-r50\\_8xb32-1x1x8-100e\\_kinetics400-rgb\\_20220906-2692d16c.pth](https://download.openmmlab.com/mmaaction/v1.0/recognition/tsn/tsn_imagenet-pretrained-r50_8xb32-1x1x8-100e_kinetics400-rgb/tsn_imagenet-pretrained-r50_8xb32-1x1x8-100e_kinetics400-rgb_20220906-2692d16c.pth) (Accessed on July 30, 2024).
- [19] W. Kay, J. Carreira, K. Simonyan, et. al, “The kinetics

human action video dataset,” *arXiv preprint*,  
arXiv:1708.06950, May 2017. Retrieved from DOI:  
<https://doi.org/10.48550/arXiv.1705.06950>.



**이 서 영 (Seo-Young Lee)**

2021년 8월 : 한국기술교육대학교 컴퓨터공학과 졸업 (공학사)  
2022년 3월 ~ 2024년 8월 : 한국기술교육대학교 대학원 컴퓨터공학과 석사과정  
※관심분야 : AI, 웹, 컴퓨터 비전 등



**유 연 휘 (Yeon-Hwi You)**

2022년 2월 : 한국기술교육대학교 컴퓨터공학과 (공학사)  
2022년 3월 ~ 2024년 8월 : 한국기술교육대학교 대학원 컴퓨터공학과 석사과정  
※관심분야 : AI, 빅데이터, 추천 시스템 등



**박 호 경 (Hyo-Gyeong Park)**

2021년 8월 : 한국기술교육대학교 컴퓨터공학 (공학사)  
2021년 8월 ~ 2024년 8월 : 한국기술교육대학교 대학원 컴퓨터공학과 석사과정  
※관심분야 : AI, 빅데이터, 추천 시스템 등



**박 병 주 (Byeong-Ju Park)**

2024년 2월 : 고려사이버대학교 전기전자공학 (공학사)  
2024년 3월 ~ 현재 : 한국기술교육대학교 대학원 컴퓨터공학과  
※관심분야 : AI, 컴퓨터 비전 등



**문 일 영 (Il-Young Moon)**

2005년 2월: 한국항공대학교 항공통신정보공학과 (공학박사)  
2005년 3월 ~ 현재: 한국기술교육대학교 컴퓨터공학과 교수  
※관심분야 : AI, 무선인터넷 응용, 무선 인터넷, 모바일IP 등