

**운전자 폭행 탐지를 위한 경량 CNN아키텍처 기반**

**시간적 세그먼트 네트워크 기법**

**과목명 기계학습**

**담당교수님 윤일동교수님**

**팀구성**

**전공 컴퓨터전자시스템공학**

**201904458 이준용**

목차

**서론**

1.1 프로젝트 개요

1.2 프로젝트 목적

**본론**

2.1 프로젝트 실험 방법

2.2 기존 모델과 차별점

**결론**

3.1 프로젝트 기대효과

**서론**

**1.1 프로젝트 개요**

여러 뉴스에서 이와 같은 운전자 폭행 사건에 대한 보도가 빈번하게 나오고 있고, 실제로 2018년 2천

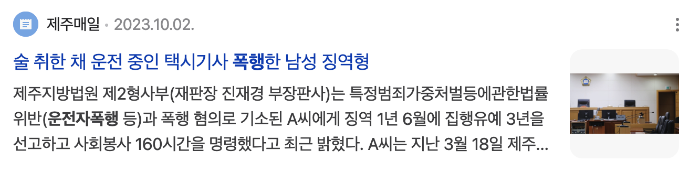
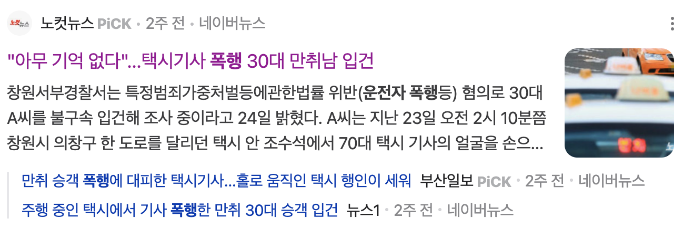
여건에서 2022년 4천여건까지 5년간 증가하고 있는 추세입니다.

그림 1. 운전자 폭행 뉴스

다음은 대한민국 전국의 범죄 발생 및 검거 현황에 대해 조사했습니다.

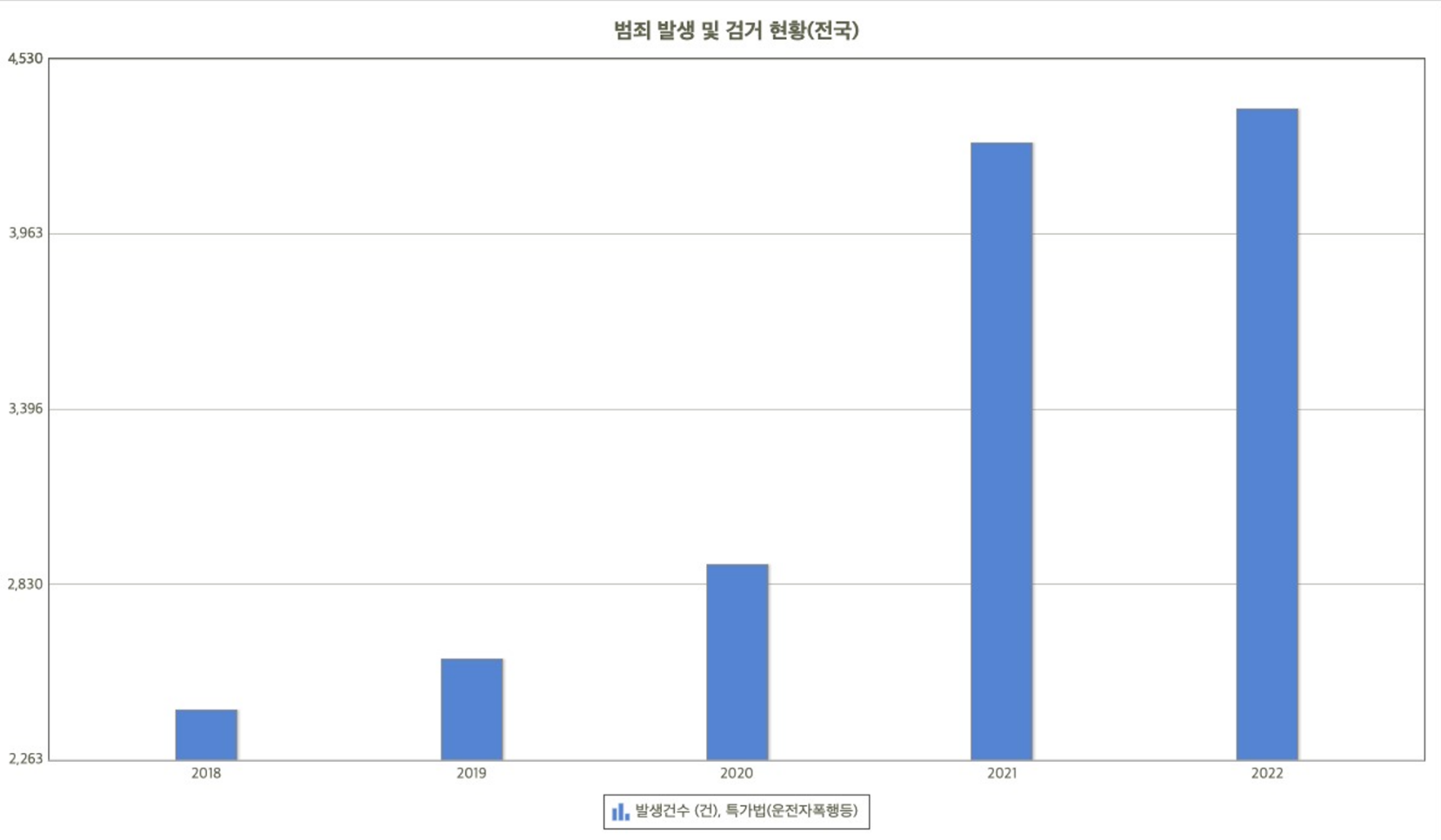


그림 2. 특가법(운전자폭행등) 발생건수

**1.2 프로젝트 목적**

운전자 폭행 사건의 다수는 주취자에 의한 폭행으로 주로 늦은 시간에 발생하므로 재빠른 신고가 어려운 경우가 많습니다. 운전자 한명에서의 이상행동 탐지에서 더 나아가 운전자이외의 탑승자가 운전자를 폭행하는 장면을 탐지해보고자 하는 생각에서 출발했고, 따라서 저는 운전자 폭행을 탐지할 수 있는 빠르고 정확한 모델 개발을 목표로 프로젝트를 진행하게 되었습니다.

**본론**

**2.1 프로젝트 실험 및 방법**

선행연구조사에서 조사한 논문에 관련해서contrastive learning기법을 사용하여 운전자 이상행동 판단을 했었습니다. 하지만 contrastive learning은 수업에서 배운 내용이 아니고, 폭행 탐지를 위해서는 video dataset을 이용해야 하므로 이외의 연구 방법을 조사했습니다. 운전자관련 다양한 데이터셋과 선행연구조사는 결과 발표에서 다루겠습니다.

* **프로젝트 절차**

AIHUB의 운전자 및 탑승자 상태 및 이상행동 모니터링 데이터를 사용할 것입니다. 본 프로젝트에서는 폭행은 비정상적인(Abnormal)한 행동이고, 운전자 이상행동 탐지(Driver Anomaly Detection)에서 다루는 이상행동(Anomaly)와는 다르다고 규정합니다.

[텍스트, 폰트, 라인, 스크린샷이(가) 표시된 사진

자동 생성된 설명](https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=651)

<https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=651>

훈련하고자 하는 데이터 셋에는 폭행은 차량안에 2명이상 있어야 함으로

1. Convert : 라벨링된 json파일을 train.csv로 변환

2. Assault Spilt : train.csv에서 폭행 데이터만 추출

3. EDA : 그래프나 통계적인 방법으로 자료를 직관적으로 바라보는 과정

위의 작업을 통해 training data에서 차량에 혼자인 데이터는 모두 제외하고 운전자 폭행 데이터만 추출합니다. 마지막으로 MMaction2의 TSN모델을 사용하여 운전자 폭행 탐지를 하여 폭행인지 아닌지 classification할 계획입니다. 다음은 TSN에 대한 간단한 설명입니다.

Temporal Segment Networks for Action Recognition in Videos(<https://arxiv.org/abs/1608.00859>)

입력 받은 비디오를 사용자가 지정한 N개의 시간적 세그먼트로 나눕니다. 이후 랜덤 샘플링을 통해 각 세그먼트 내에서 일부 프레임만을 선택하여 사용합니다.

TSN 모델은 공간 정보를 다루는 Spatial Stream, 시간 정보를 다루는 Temporal Stream인 두 개의 독립적인 ConvNet 스트림을 사용합니다. 각 스트림에서 생성된 시공간 정보를 통합하여 최종적인 비디오의 분류 결과를 얻습니다.

TSN(ResNet)모델은 세그먼트 기반 샘플링 방식을 사용하여 장거리 시간 구조를 모델링하는 것을 목표로 합니다. 이 독특한 디자인을 통해 TSN은 전체 액션 비디오를 사용하여 액션 모델을 효율적으로 학습할 수 있습니다.

* **프로젝트 방법**

AI HUB에 “활용 AI모델 및 코드” 메뉴에서 TSN 모델을 다운 받을 수 있지만 Docker로 되어있어 docker 사용법을 익혀야 하므로 본 프로젝트의 범위에서 벗어나는 것으로 판단했습니다.

또한, 과거의 mmacation package와 cuda버전를 사용한 모델이므로 패키지 의존성 문제를 해결할 수 없기에 AIHub에 있는 TSN모델은 사용할 수 없는 모델 버전이라고 판단했습니다. 현재 모델 버전에 맞춰 프로젝트를 진행할 필요성을 느꼈습니다. 그러므로 mmaction2 패키지를 gitclone명령어를 통해 패키지를 다운 받고 패키지 내부의 TSN 모듈을 활용하여 모델을 재정의하여Data를 훈련시킬 것을 제안합니다.

다음으로 비디오데이터셋을 활용한 모델이므로 구글 드라이브와 구글 Colab T4 GPU(구글 무료 GPU) 사용할 것입니다. Colab T4 GPU의 경우 하루 사용량 제한이 있다는 한계가 있으므로 pretrained 된 경량 CNN아키텍처를 기반한 TSN모델로 새로운 폭행 데이터를 학습시켜서 예측결과를 도출할 것입니다.

**2.2 기존 모델과 차별점**

DMS(운전자 모니터링 시스템) 애플리케이션은 자동차에 배치되어야 하므로 한정된 자원에서의 효율적인 아키텍처가 중요합니다. 최근 몇 년 동안 몇 가지 경량 CNN 아키텍처가 제안되었습니다.

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

자동 생성된 설명

그림 3. 다양한 경량 CNN 아키텍처 조사

기존 TSN은Resnet 기반으로 모델이 개발되었습니다. 저는 TSN모델에 경량 CNN 아키텍처를 기반으로 모델을 재정의하여 훈련시키고 결과를 도출할 것을 제안합니다.

**결론**

**3.1 프로젝트 기대효과**

자동차의 한정된 자원안에서 경량 CNN 아키텍처 기반 TSN 모델을 활용한 운전자 폭행을 탐지를 통해 피해자의 직접 신고를 통해 이루어지던 방식에서 본 프로젝트에서 제안하는 모델을 통해 이후에 후속 연구 개발로 다음과 같은 결과를 제안함으로써 효율적이고 정확한 비정상적인 행동인 폭행 탐지가 기대됩니다.