

**2023 DeepLearning  
Research Paper**

---

**운전자 폭행 탐지를 위한  
경량CNN 기반  
시간적 세그먼트 네트워크 기법**

---

**HUFSUPERMAN**

김 준  
박준희  
이준용  
조종호

# IMPORTANCE OF THE TOPIC

노컷뉴스 PICK · 2주 전 · 네이버뉴스

## "아무 기억 없다"...택시기사 폭행 30대 만취남 입건

창원서부경찰서는 특정범죄가중처벌등에관한법률 위반(운전자 폭행등) 혐의로 30대 A씨를 불구속 입건해 조사 중이라고 24일 밝혔다. A씨는 지난 23일 오전 2시 10분쯤 창원시 의창구 한 도로를 달리던 택시 안 조수석에서 70대 택시 기사의 얼굴을 손으...



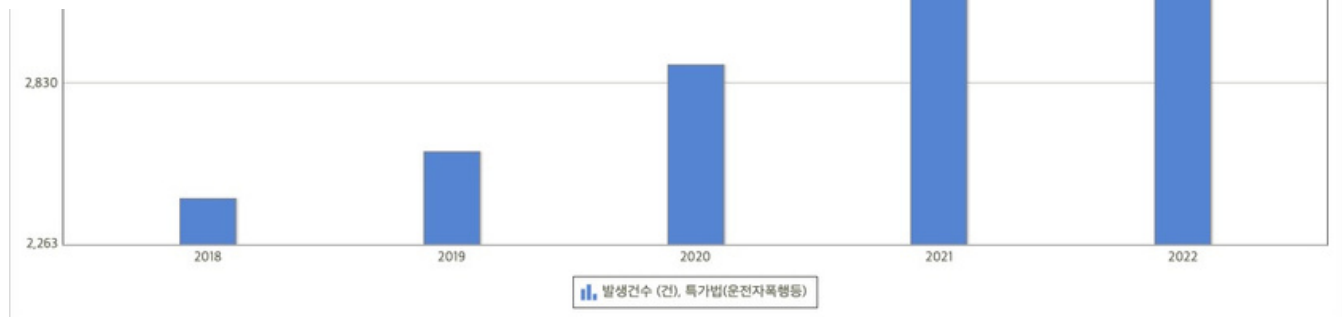
만취 승객 폭행에 대피한 택시기사...출로 움직인 택시 행인이 세워 부산일보 PICK · 2주 전 · 네이버뉴스

주행 중인 택시에서 기사 폭행한 만취 30대 승객 입건 뉴스1 · 2주 전 · 네이버뉴스

제주매일 · 2023.10.02.

## 술 취한 채 운전 중인 택시기사 폭행한 남성 징역형

제주지방법원 제2형사부(재판장 진재경 부장판사)는 특정범죄가중처벌등에관한법률 위반(운전자폭행 등)과 폭행 혐의로 기소된 A씨에게 징역 1년 6월에 집행유예 3년을 선고하고 사회봉사 160시간을 명령했다고 최근 밝혔다. A씨는 지난 3월 18일 제주...



# INTRODUCE DATA

## 데이터셋 소개



운전자 및 탑승자 상태 및 이상행동 모니터링

이상행동 7종

(졸음운전, 음주운전, 물건찾기, 통화, 휴대폰 조작, 차량 제어, 운전자 폭행)



**운전자 폭행 데이터**

이미지 : 약 2만개 , 용량 : 25GB

**정상적인 운전 데이터(탑승자 2명)**

이미지 : 약 2만개 , 용량 : 25GB



**Test용 YouTube 실제폭행 데이터 수집 및 전처리**

(정상/폭행 구간을 나눠 라벨링 진행)

비디오: 19개 용량 : 500MB



# INTRODUCE DATA

## Problem: 정상 동영상 데이터의 부재

폭행 데이터 각 폴더마다 5장씩 존재



장면의 흐름을 헤치지 않게 하기 위해  
폭행 데이터를 각 폴더마다 8장으로 증강

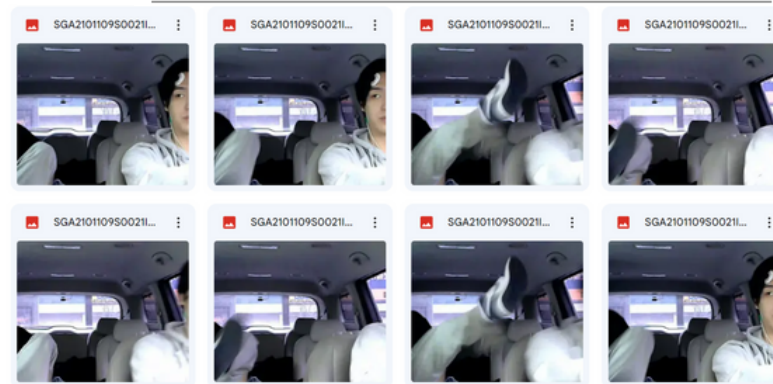
원본 5장 이미지 중 연속되는 정상 이미지 추출  
-> 2장 또는 3,4장의 정상 이미지



모델 학습에서 요구되는 Data 묶음은  
8장이기 때문에  
각 Scene Data마다 8장으로 증강

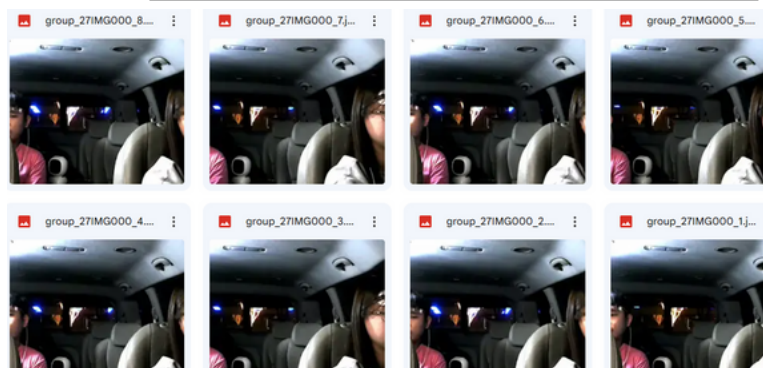
폭행

Scene	data
A	1,2,3,4,5,4,3,2
B	1,2,3,4,5,4,3,2
C	1,2,3,4,5,4,3,2
D	1,2,3,4,5,4,3,2



정상

Scene	data
A	1,2,3,2,1,2,3,2
B	1,2,3,4,3,2,1,2
C	1,2,1,2,1,2,1,2
D	1,2,3,2,1,2,3,2



# CUSTOM SAMPLER

## DataLoader config code

```
train_dataloader = dict(
    batch_size=8,
    num_workers=2,
    persistent_workers=True,
    sampler=dict(type='CustomSampler', shuffle=True, ann_file=ann_file_train),
    dataset=dict(
        type=dataset_type,
        ann_file=ann_file_train,
        data_prefix=dict(img=data_root),
        pipeline=train_pipeline))

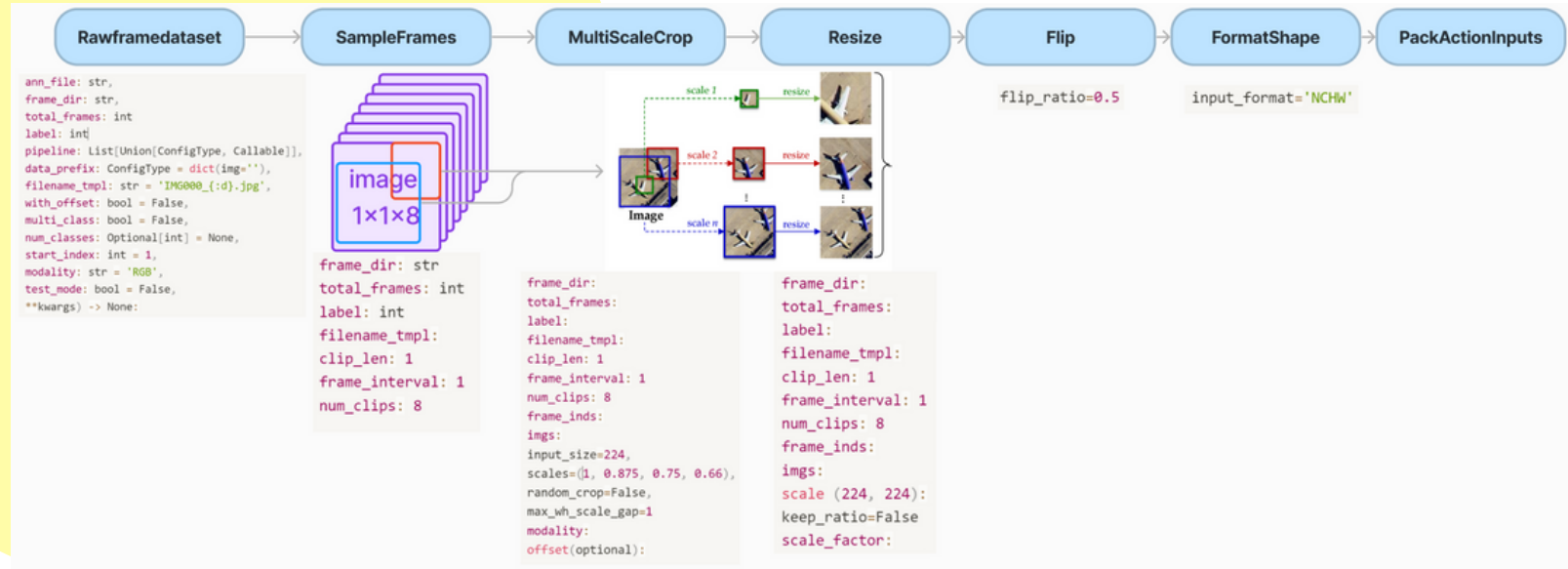
val_dataloader = dict(
    batch_size=8,
    num_workers=2,
    persistent_workers=True,
    sampler=dict(type='CustomSampler',
    dataset=dict(
        type=dataset_type,
        ann_file=ann_file_val,
        data_prefix=dict(img=data_root),
        pipeline=val_pipeline,
        test_mode=True))
```

모든 Train set에서 랜덤 샘플링하는 기존 **Default Sampler**에서  
비슷한 장면은 한 에폭에 한번만 샘플링하는 **CustomSampler**로 변경

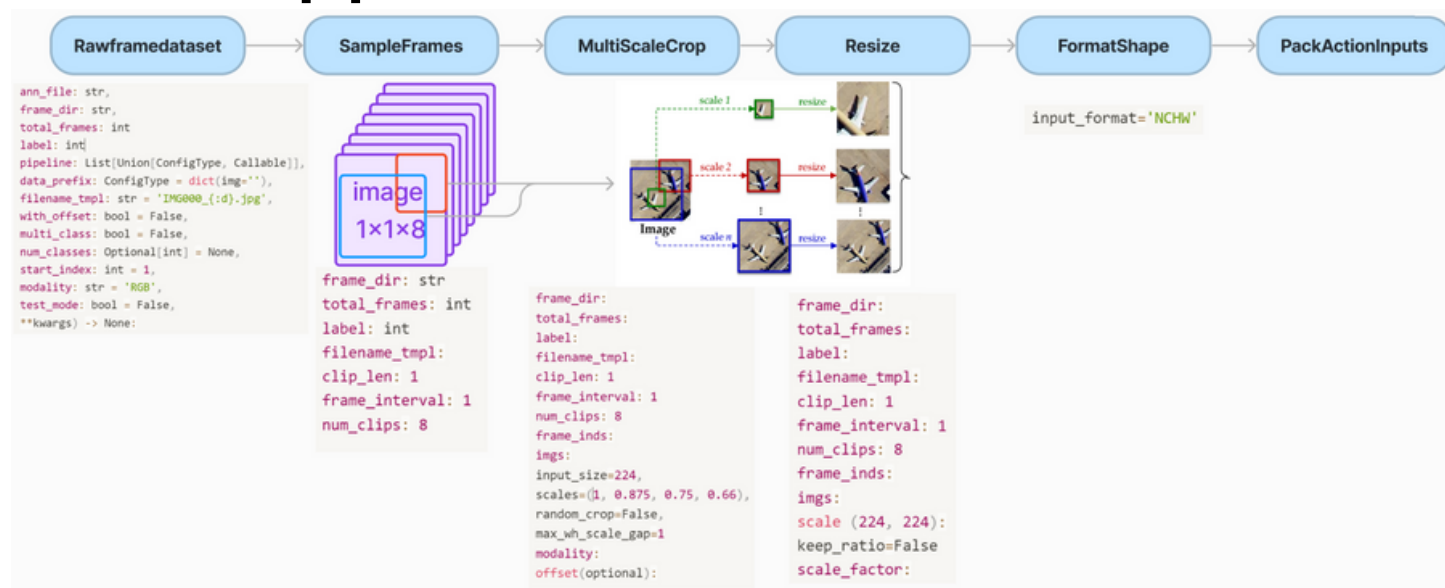


# RESEARCH DATA PIPELINE

## Train pipeline

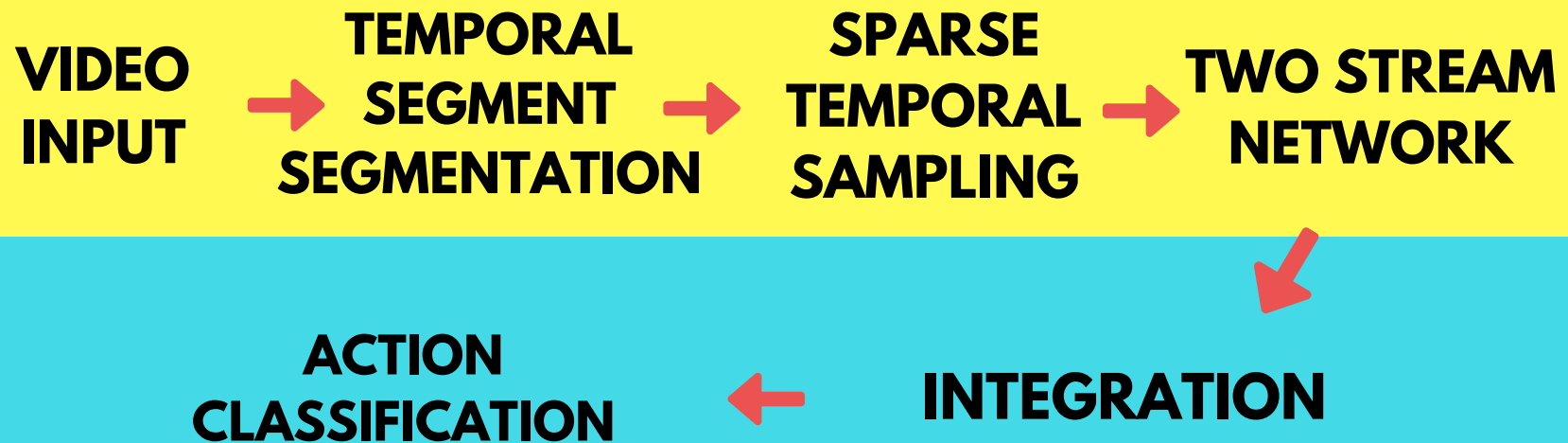
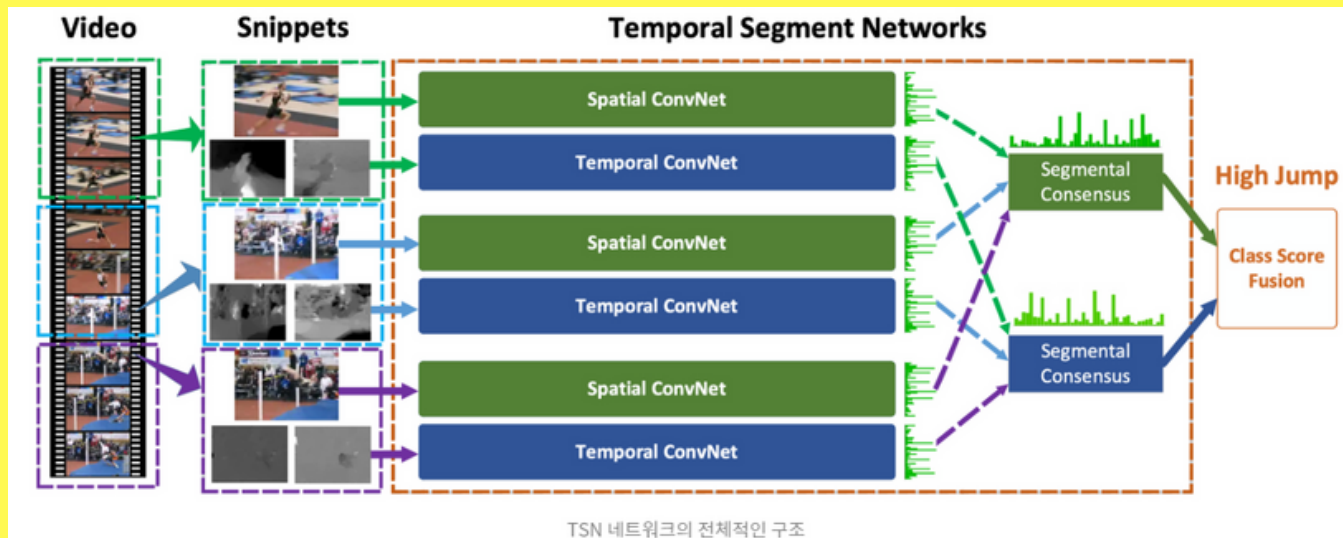


## Validation pipeline



## INTRODUCE MODEL

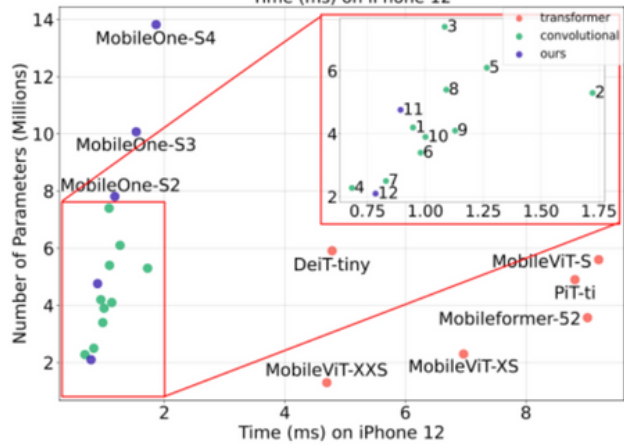
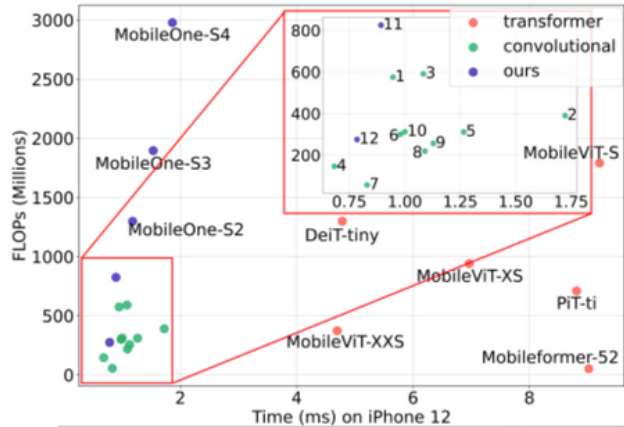
# Temporal Segment Networks



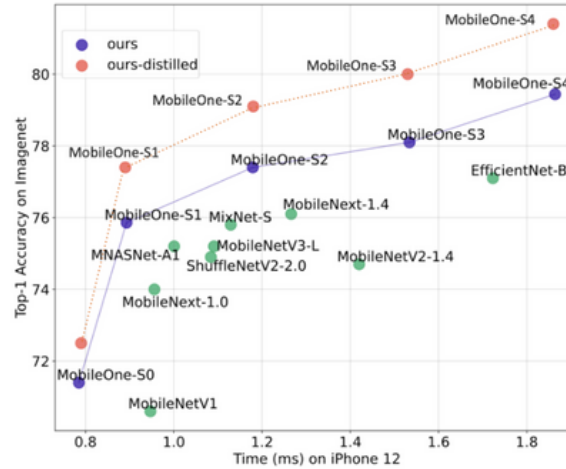


# RESEARCH

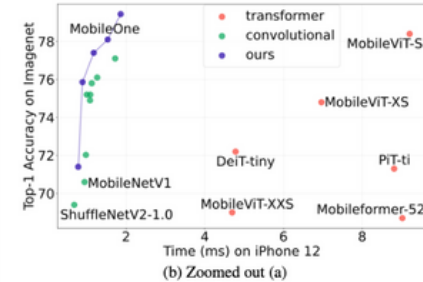
## Model - MobileOne



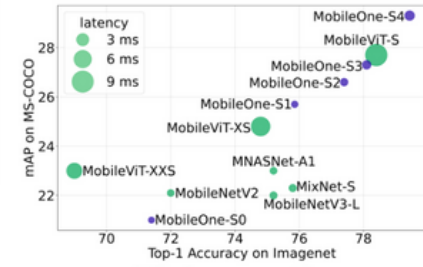
1	2	3	4
MobileNetV1	EfficientNet-B0	ShuffleNetV2-2.0	ShuffleNetV2-1.0
5	6	7	8
MobileNet-1.4	MobileNetV2	MobileNetV3-S	MobileNetV3-L
9	10	11	12
MixNet-S	MNASNet-A1	MobileOne-S1	MobileOne-S0



(a) Top 1 accuracy vs Latency on iPhone 12.

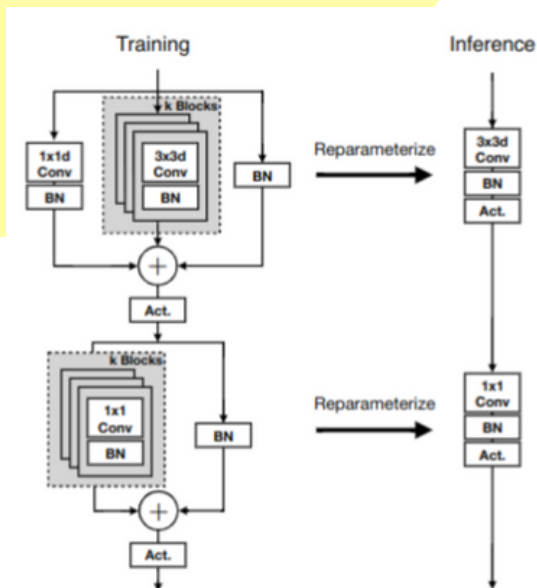


(b) Zoomed out (a)



(c) Top-1 accuracy vs mAP.

**1ms의 inference time**을 가지면서 높은 성능을 유지  
**transformer** 계열인 **Mobileformer**보다 **38배**나 빠르는데 비슷한 성능



파라미터 수가 많은 모델들이 **latency**가 적은 것을 확인  
**Mobilenet**이 다른 모델들과 비슷한 **FLOPs**임에도 더 빠른 걸  
 확인

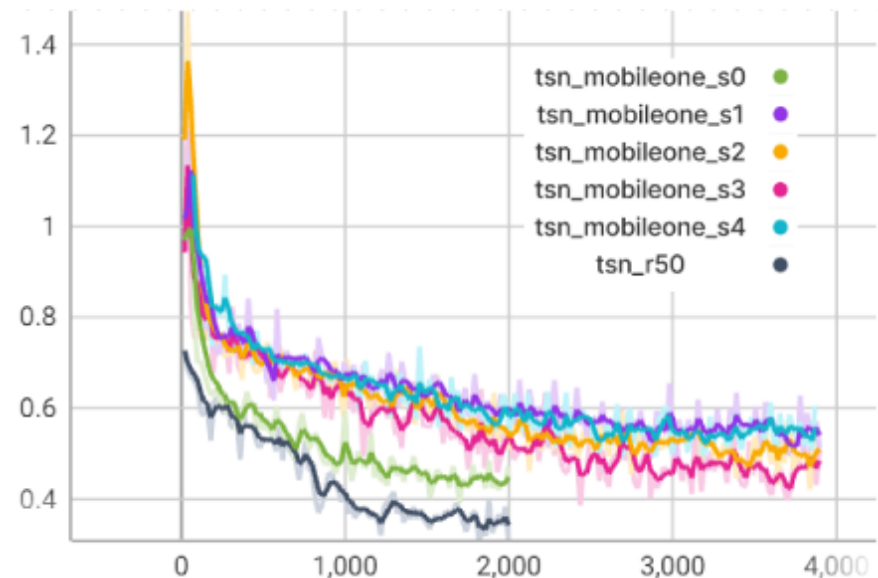


# RESEARCH

## Model List

					TSNHead			
Model	Params(M)	Flops(G)	batch size	Loss Function	num_classes	in_channels	dropout_ratio	init_std
tsn_mobileone_s0	2.08	0.27	16	CrossEntropy Loss	2	1024	0.4	0.01
tsn_mobileone_s1	4.76	0.82	8			1280		
tsn_mobileone_s2	7.81	1.3	8			2048		
tsn_mobileone_s3	10.08	1.89	8			2048		
tsn_mobileone_s4	14.84	2.98	8			2048		
tsn_r50	24.33	102.7	16			2048		

**SGD 100 epoch**  
**LR 0.01**  
**momentum 0.9**  
**weight\_decay 0.0001**  
**param\_scheduler MultiStepLR**



Train Loss Graph

# Resnet기반 TSN모델 VS mobileone기반 TSN모델

## aihub TESTSET 성능 비교

		Precision	Recall	F1-score	accuracy
Tsn_r50	Assault	0.88	0.43	0.58	0.69
	Normal	0.62	0.94	0.75	
	Macro avg	0.75	0.69	0.67	
Tsn_mobileone_s0	Assault	0.80	0.47	0.83	0.68
	Normal	0.62	0.88	0.71	
	Macro avg	0.71	0.68	0.66	
Tsn_mobileone_s1	Assault	0.78	0.51	0.62	0.68
	Normal	0.64	0.86	0.73	
	Macro avg	0.71	0.68	0.67	
Tsn_mobileone_s2	Assault	0.75	0.50	0.60	0.67
	Normal	0.63	0.83	0.72	
	Macro avg	0.69	0.67	0.66	
Tsn_mobileone_s3	Assault	0.83	0.53	0.65	0.71
	Normal	0.66	0.89	0.76	
	Macro avg	0.74	0.71	0.70	
Tsn_mobileone_s4	Assault	0.81	0.53	0.64	0.70
	Normal	0.65	0.88	0.75	
	Macro avg	0.73	0.70	0.69	

# Resnet기반 TSN모델 VS mobileone기반 TSN모델

## YouTube TESTSET 성능 비교

		Precision	Recall	F1-score	accuracy
Tsn_r50	Assault	0.83	0.83	0.83	0.79
	Normal	0.71	0.71	0.71	
	Macro avg	0.77	0.77	0.77	
	Weighted avg	0.79	0.79	0.79	

		Precision	Recall	F1-score	accuracy
Tsn_mobileone_s3	Assault	0.91	0.83	0.87	0.84
	Normal	0.75	0.86	0.80	
	Macro avg	0.83	0.85	0.83	
	Weighted avg	0.85	0.84	0.84	

# EXPECTATION& USAGE

## 기대 효과

1. 탑승자의 경각심을 깨워 운전자가 폭행 당하는 사건을 방지 할수 있다.
2. 운전자 폭행이 원인이 되어 운전중인 차량이 다른 차들과 충돌하는 교통사고와 같은 **2**차적 피해를 예방 할수 있다.

## 활용 방안

1. 보험사 및 법 집행기관에 정확한 사고 경위를 제공하여 공정하고 효과 적인 조치를 취할 수 있는 방안으로도 활용될 수 있다
2. 운전자 폭행 사건을 탐지하는 데이터의 수집을 도와 도로 위 운전자 폭행문제에 대한 통계를 제공할 수 있다.

# CHALLENGE

---

## 개선해야 되는 방향

1. 핸들링으로 인한 몸쓸림과 같은 역동적인 행동의 데이터가 폭행으로 탐지되지 않도록 추가적인 데이터수집이 필요하다.
2. 탑승자가 다수인 경우의 데이터를 추가적으로 수집해야 할 필요가 있다.

## 후속 연구 방향

1. 운전자 폭행 탐지시 신고하는 시스템으로까지 확장
2. 버스와 같은 대중교통에도 확장

# **2023 DeapLearning 2 team Research Paper**

---



**HUFSUPERMAN**  
**Thanks For Watching!**