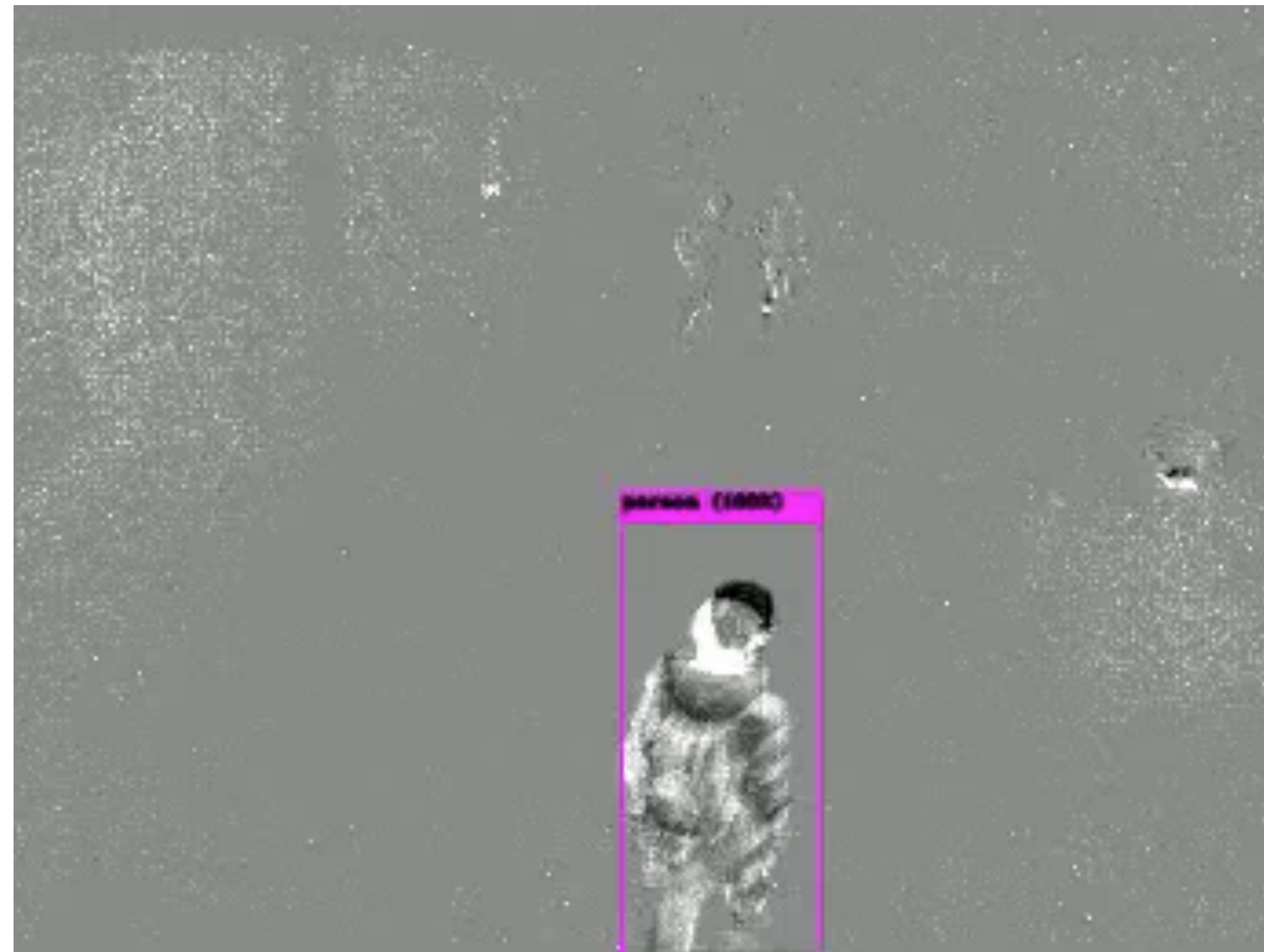


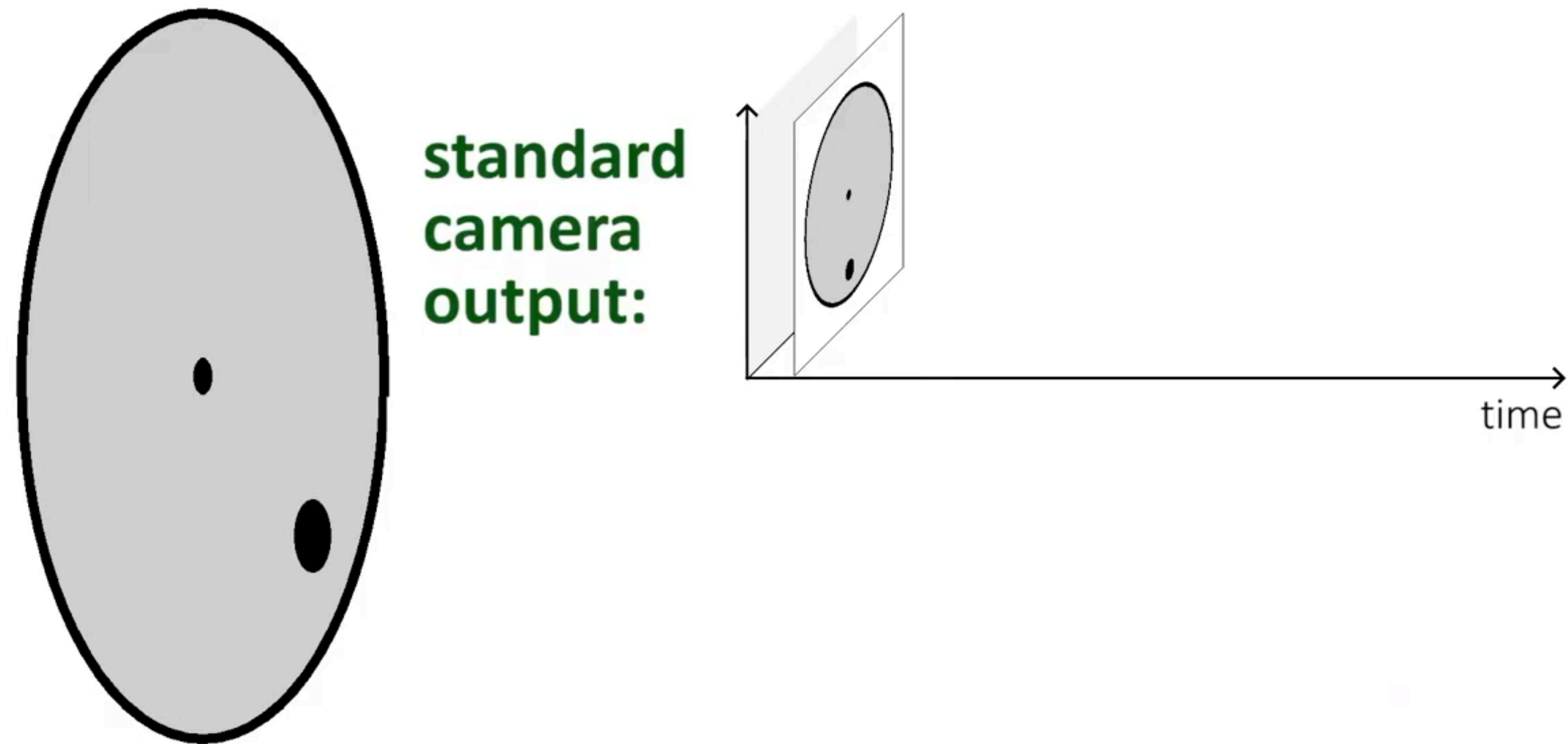
이벤트 카메라 : 이벤트의 적절한 이미지 변환을 통한 보행자 검출 성능 향상

부제 : 이벤트 카메라를 이용한 High Dynamic Range Object Detection



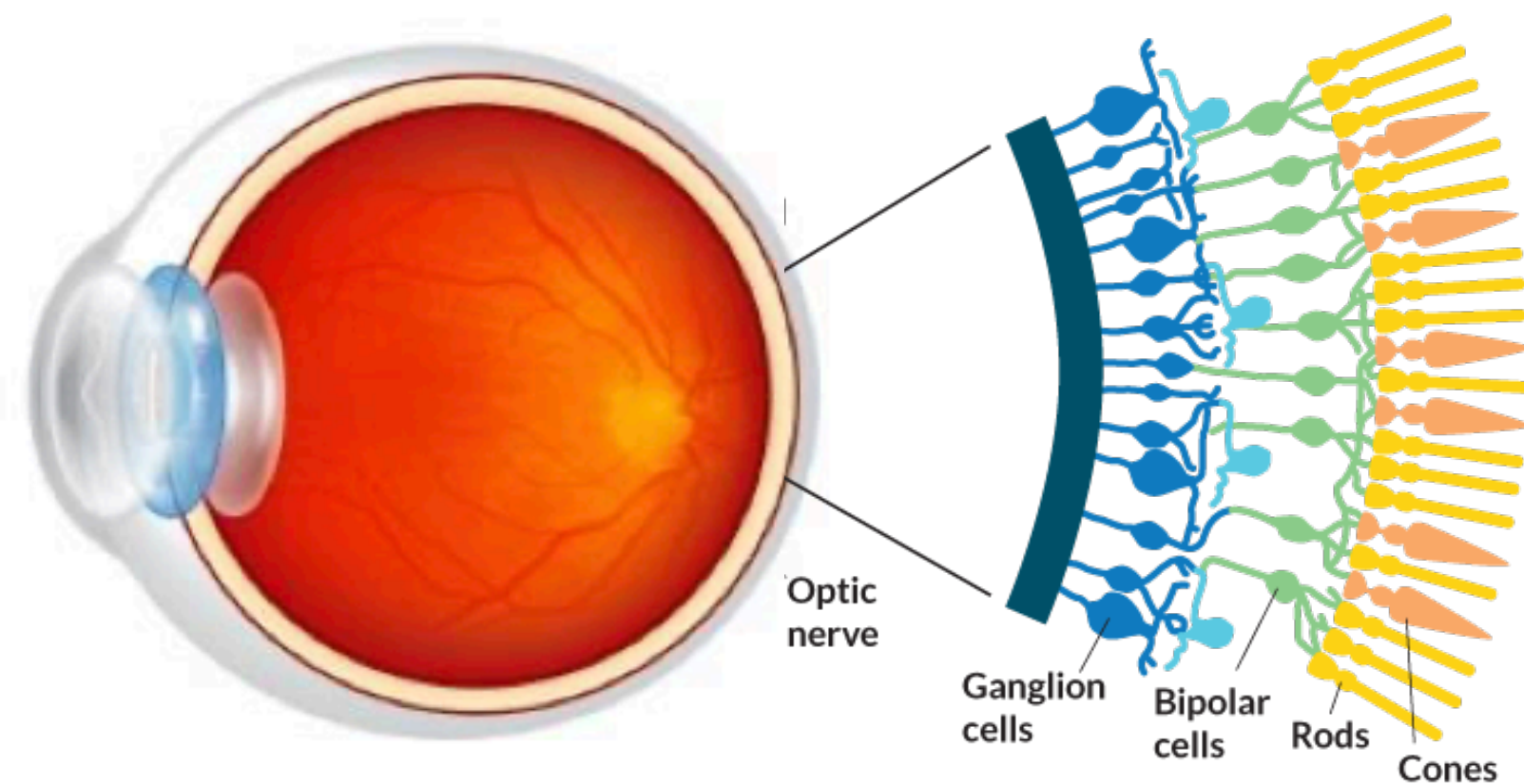
이벤트카메라, 동적비전센서 : Dynamic Vision Sensor (DVS) (Neuromorphic camera)

- 빛의 밝기 변화를 측정하여 정보를 저장하는 새로운 개념의 생체 모방형 카메라.



이벤트카메라, 동적비전센서 : Dynamic Vision Sensor (DVS) (Neuromorphic camera)

- 생체 모방형 카메라 (Bio-inspired camera)
- 우리 눈에는 빛을 감지하는 시세포인 광수용 세포(Photoreceptor Cell) 두 가지 종류가 존재함.
 - **Color Vision**, 즉 색상을 인식하는 원추세포(Cones) - **Standard Camera**
 - **Monochromatic Vision**, 즉 단색의 조도를 인식하는 간상세포(Rods) - **Event Camera**



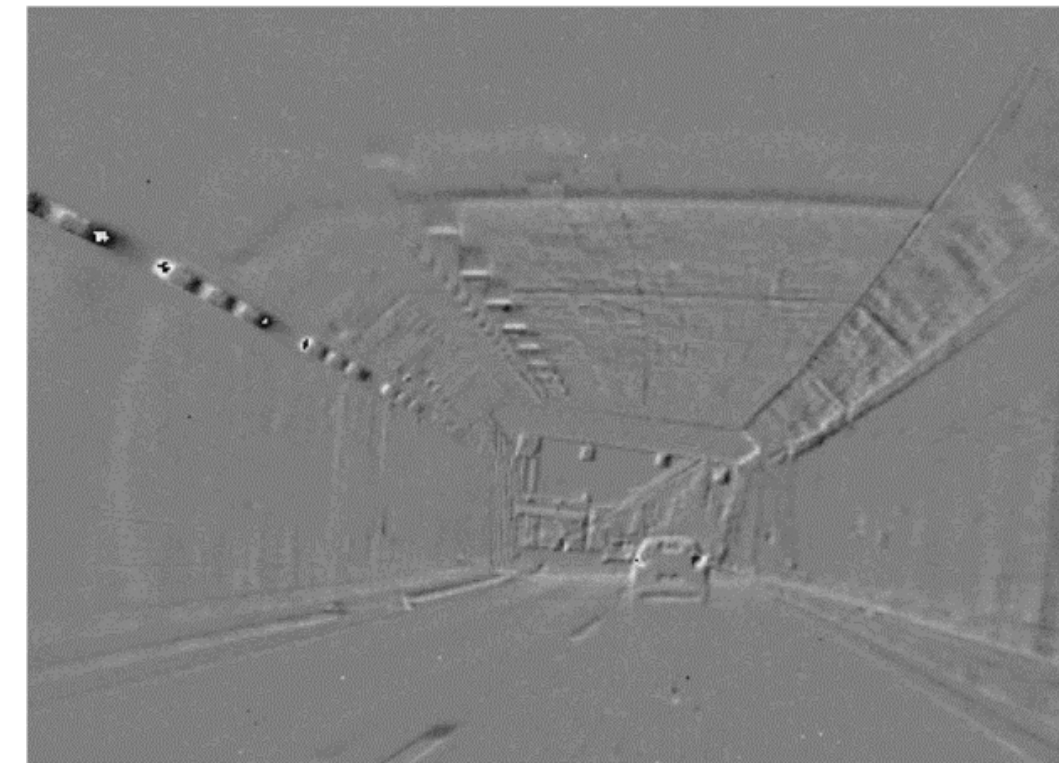
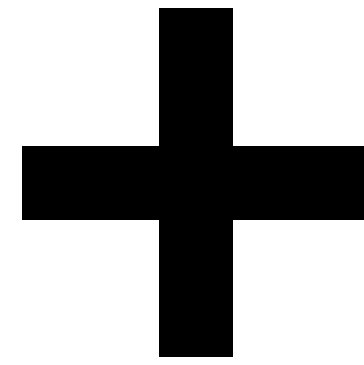
	원추세포(추상세포)	간상세포
Name	Cones	Rods
Number	6-7 million	75-150 million
Sensitivities	색상(주간 시각)	조도(야간 시각)
Related	Standard Camera	Event Camera (DVS)

이벤트카메라, 동적비전센서 : Dynamic Vision Sensor (DVS)

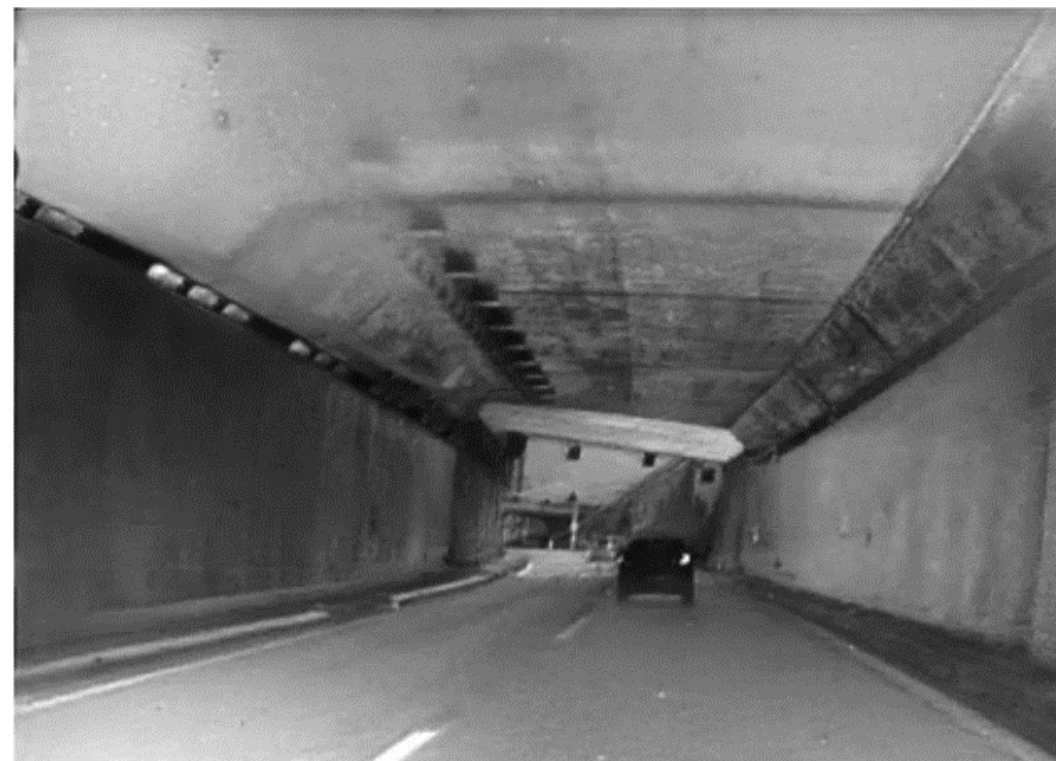
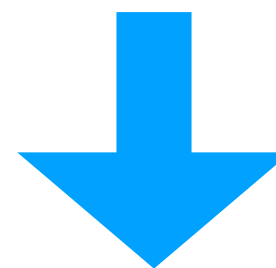
Driving out of a tunnel



원추세포(Cones) - Standard Camera



간상세포(Rods) - Event Camera

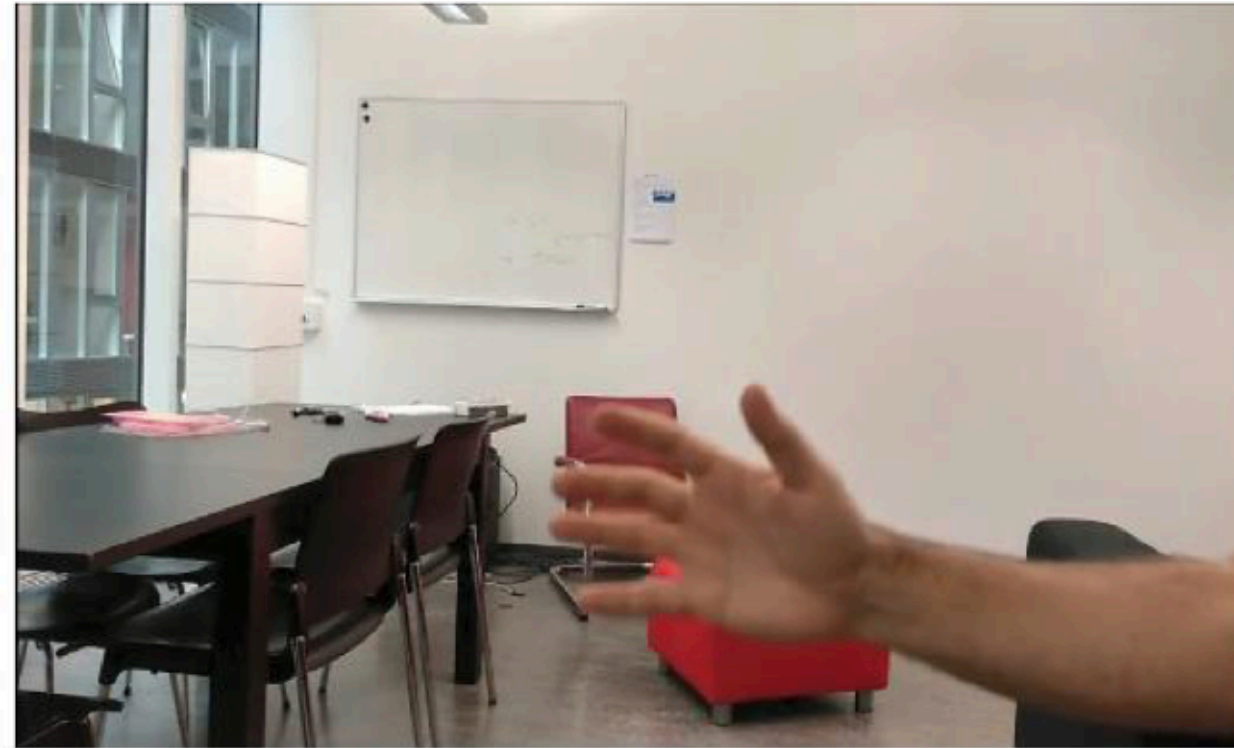


Reconstruction - Our Eyes

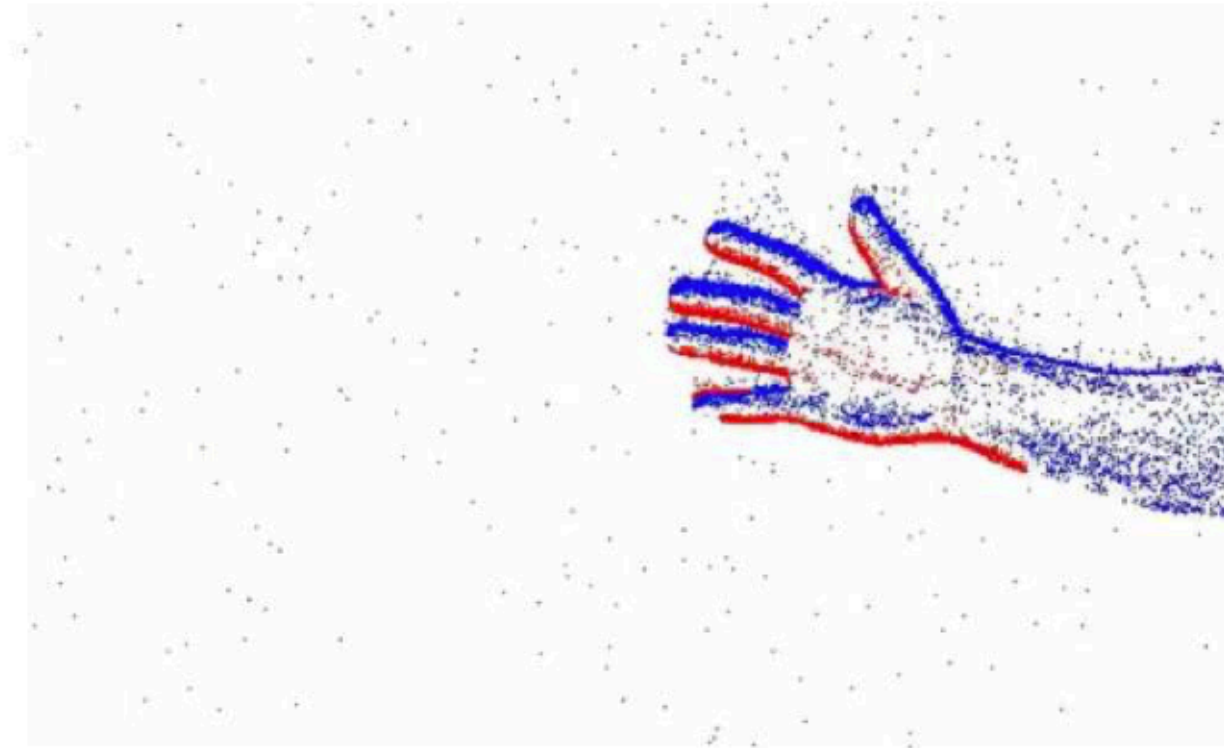
DVS Outputs

손이 움직임

Standard Camera



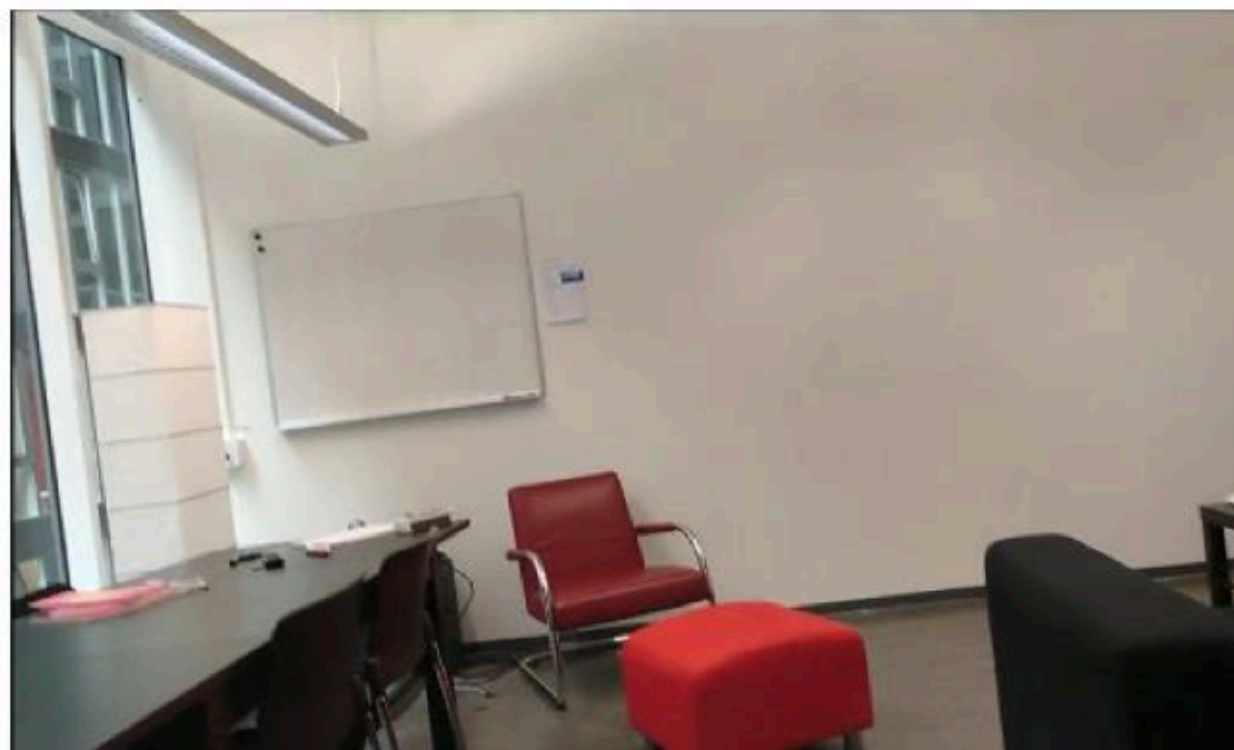
Event Camera (ON, OFF events)



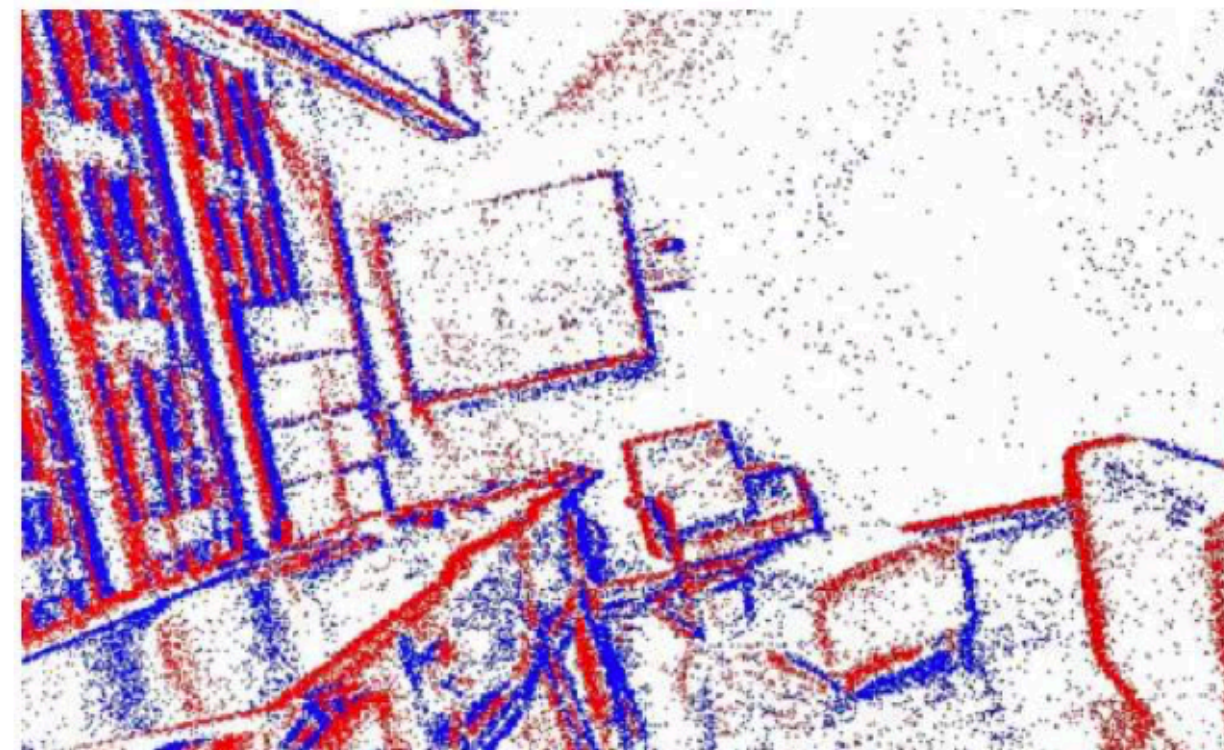
$\Delta T = 10 \text{ ms}$

카메라가 움직임

Standard Camera

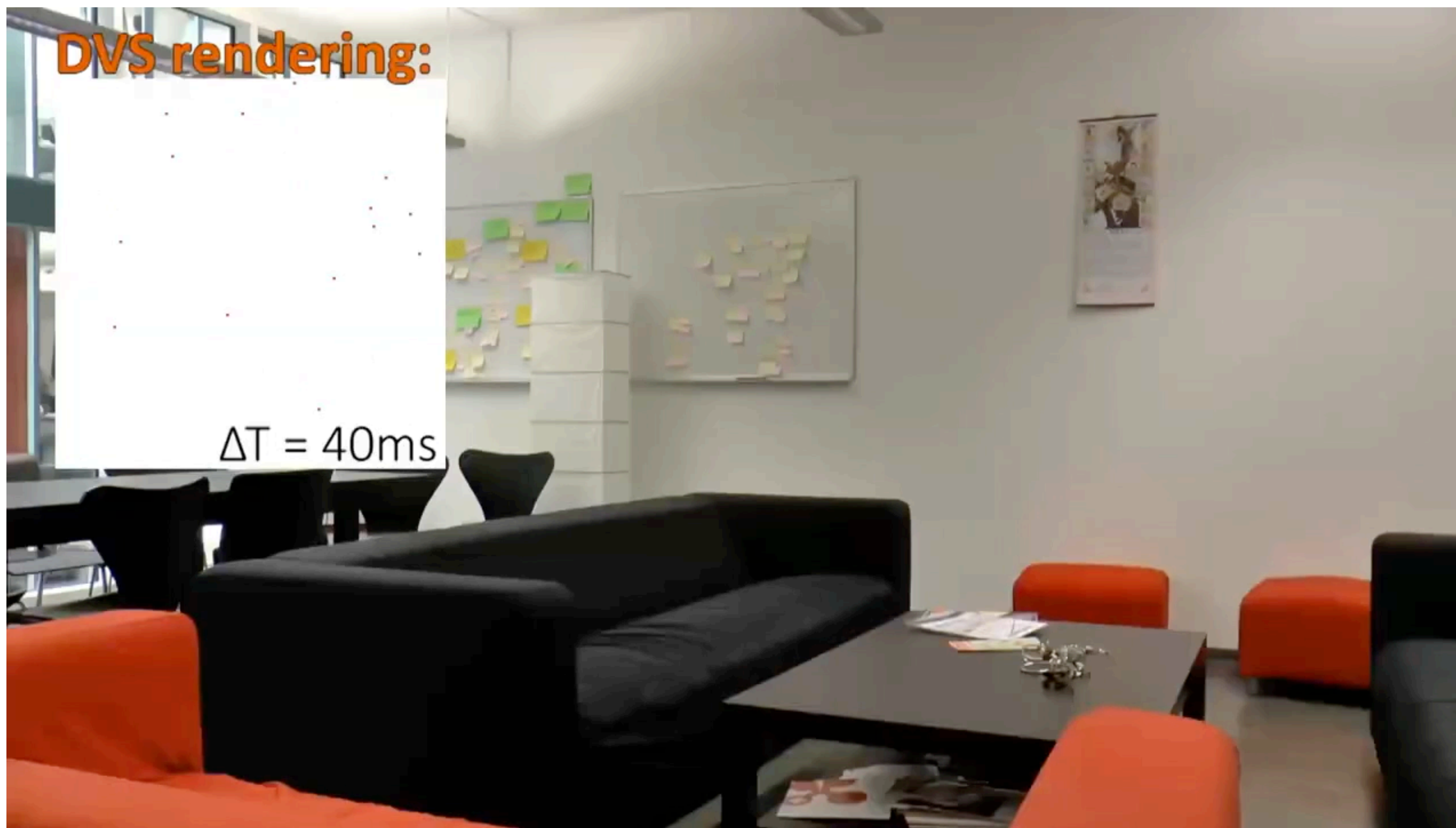


Event Camera (ON, OFF events)

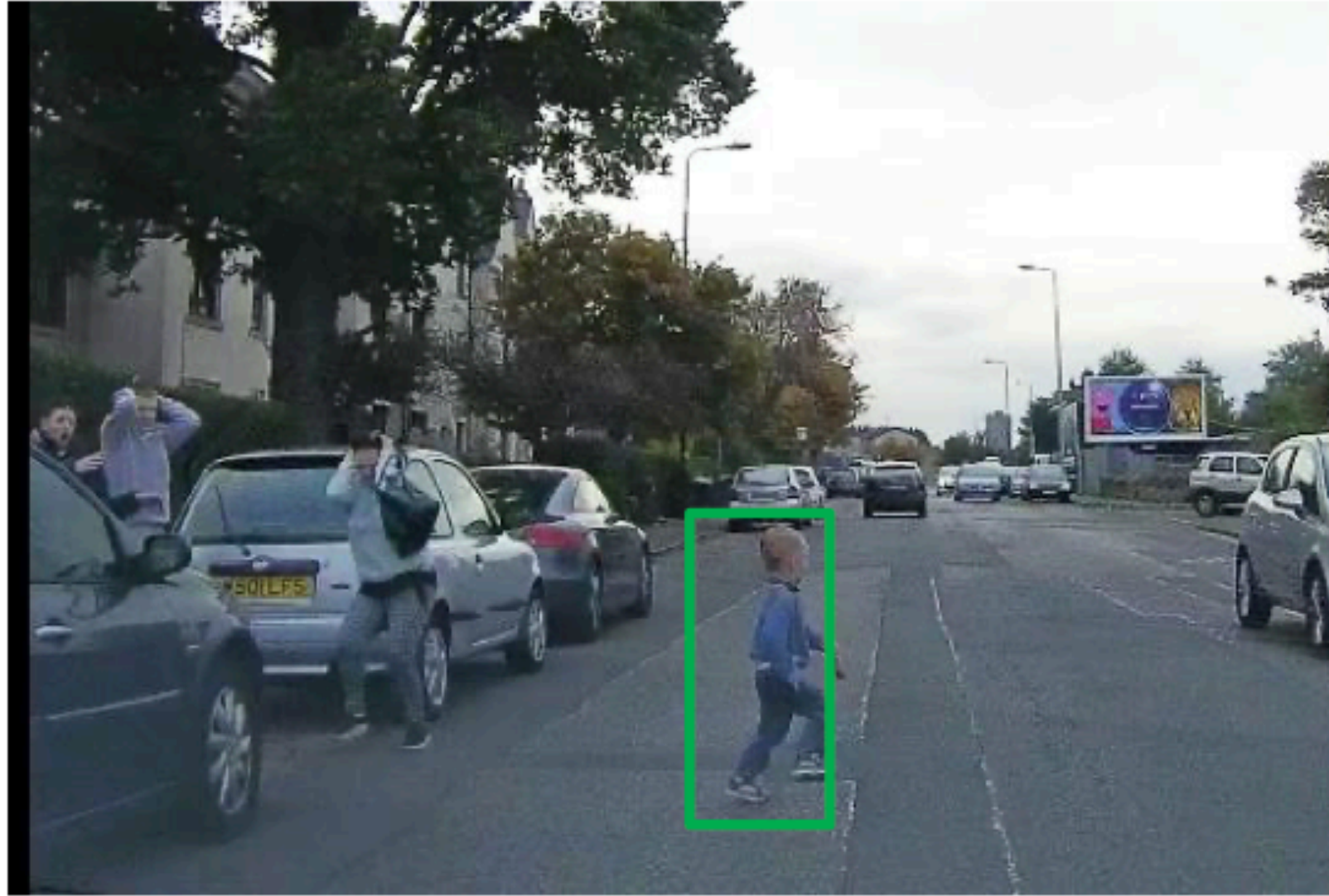


$\Delta T = 40 \text{ ms}$

DVS Outputs



Latency



Motion blur



Dynamic Range



Event cameras do not suffer from these problems!

Differences from standard camera

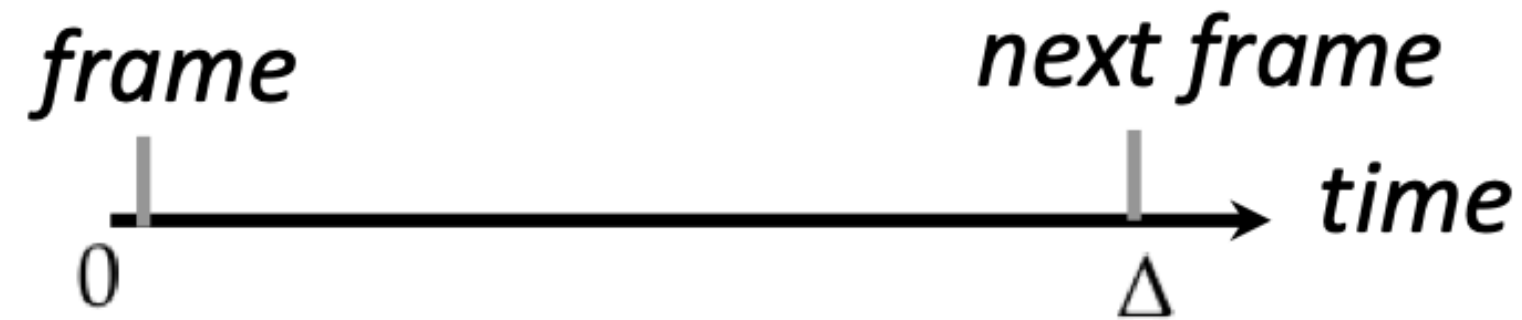
And advantages

- 픽셀마다 출력 값이 독립적이며 비동기적이다.
- 높은 동적 범위 (**HDR**) : Events > 90dB
- **Time resolution** : 200 μ s (0.0002초 > 5000FPS)
- **Low Power** : <140mA @ 5 VDC (USB)
- **Min. Latency** : < 200 μ s (full-event frame)
- **High Temporal Resolution / No saturation** : 빛에 따른 채도 변화가 없음 / **No motion blur** : 모션블러 없음

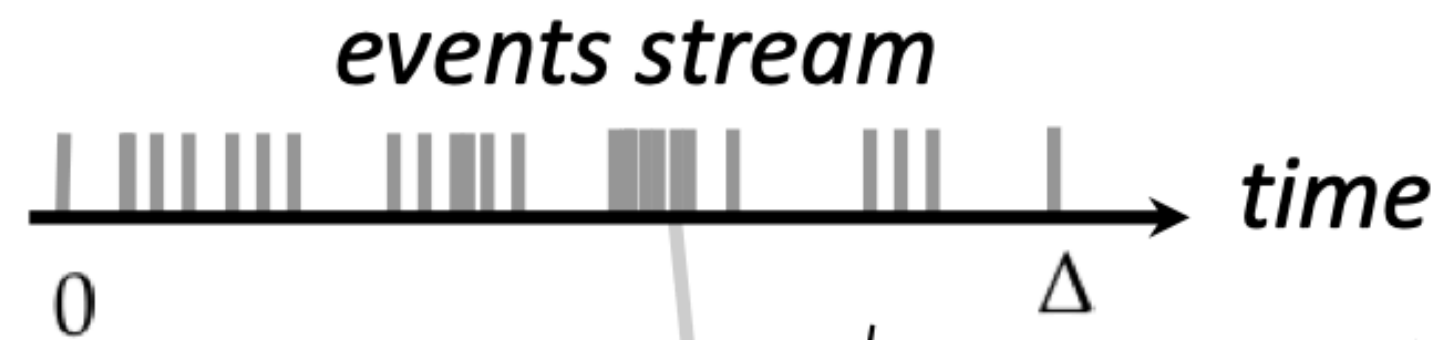
Different Output

기존 카메라와는 다른 출력값

- A traditional camera outputs frames at fixed time intervals:



- By contrast, a DVS outputs asynchronous events at *microsecond* resolution generated each time a single pixel detects an intensity changes value



$$\text{event:} \left\langle t, \langle x, y \rangle, \text{sign} \left(\frac{dI(x, y)}{dt} \right) \right\rangle$$

Timestamp (μs)

Pixel coordinates

Event polarity (or sign) (-1 or 1): increase or decrease of brightness

```
events.txt
0.000000000 33 39 1
0.000011001 158 145 1
0.000050000 88 143 0
0.000055000 174 154 0
0.000080001 112 139 1
0.000123000 136 171 0
0.000130001 173 90 0
0.000139001 106 140 0
0.000148001 192 79 1
0.000191001 118 149 0
0.000223000 191 95 0
```

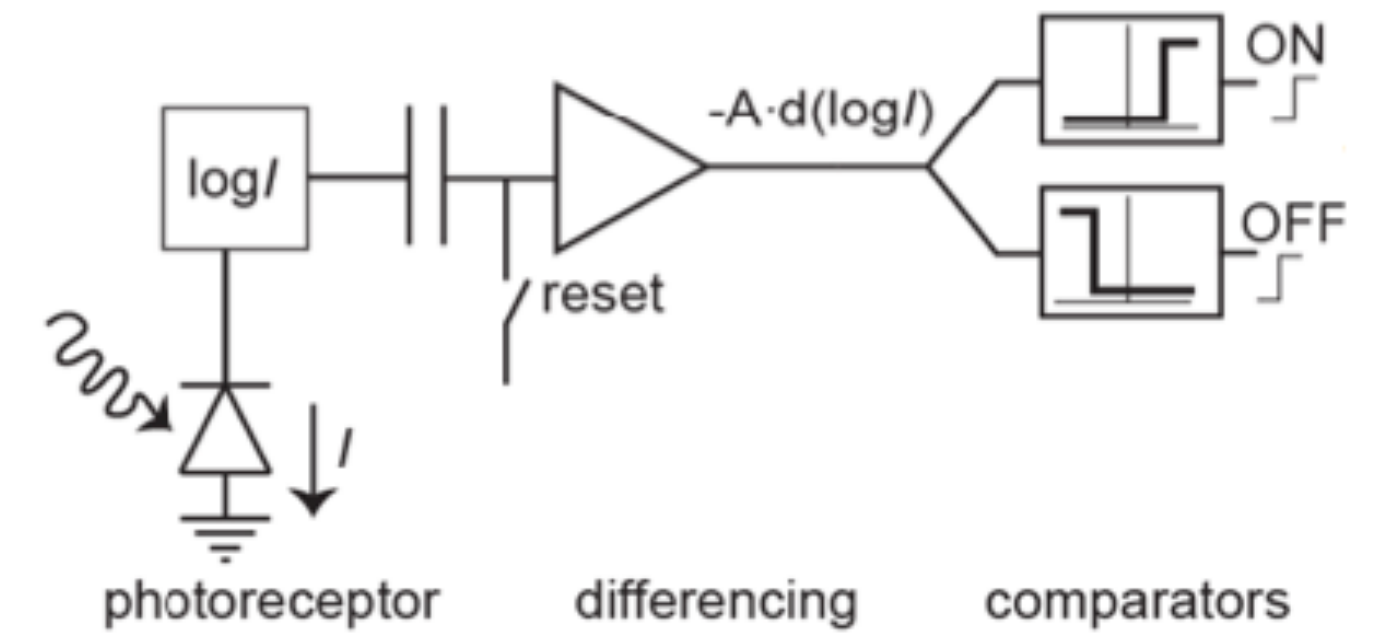
Output : < 시간, 위치, 극(-1,1) >

출력모델

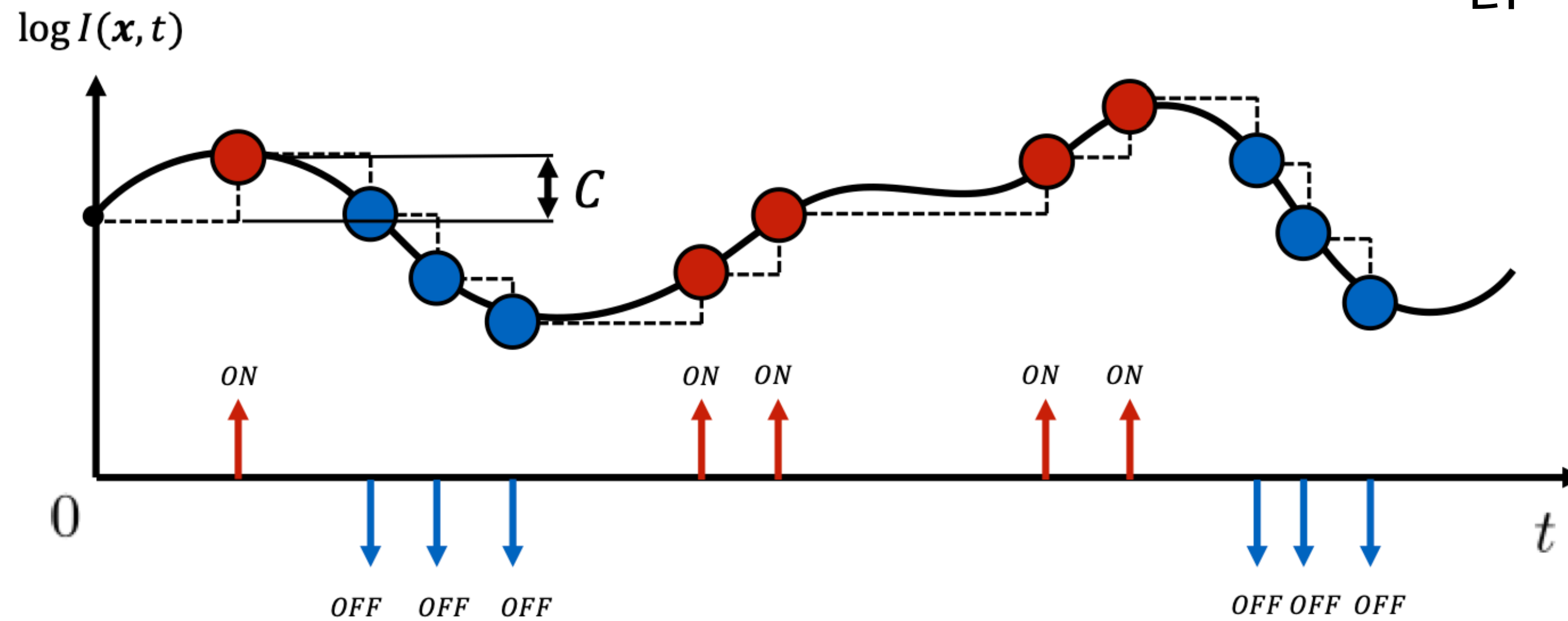
픽셀의 밝기 변화를 비동기적으로 측정하여 출력함.

An event is triggered at a **single pixel** if

$$\log I(x, t) - \log I(x, t - \Delta t) = \pm C$$

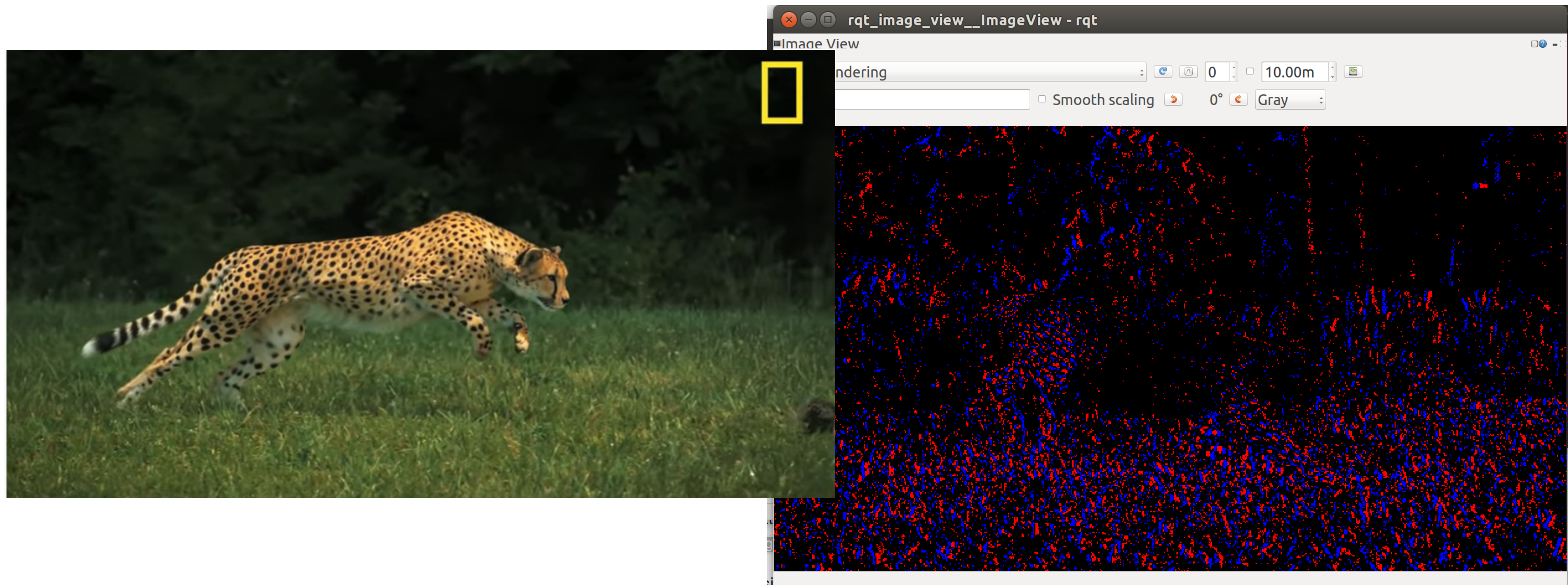


$C =$ Contrast sensitivity
 밝기 민감도



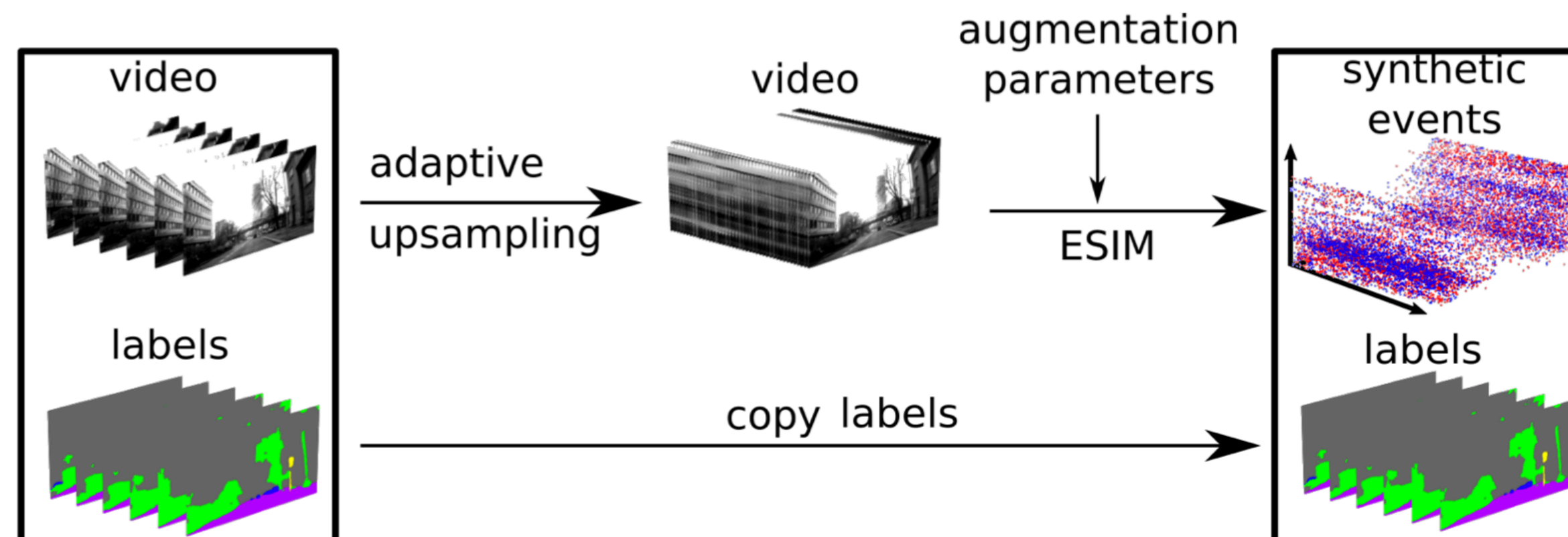
이벤트 시뮬레이터 (ESIM: an Open Event Camera Simulator)

Ros : rosbag 파일로 저장 및 재생가능.



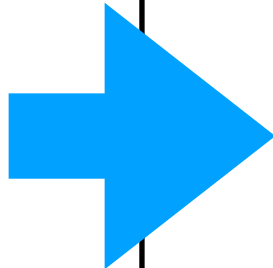
Video to Events: Recycling Video Datasets for Event Cameras

Adaptive Upsampling : NVIDIA Super SloMo



Data Augmentation

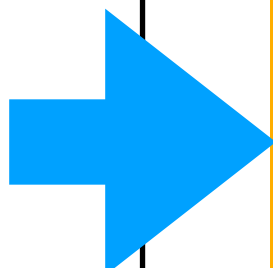
이벤트 카메라 출력
(.AEDAT)



흑백 이미지
Data
Augmentation



라벨링



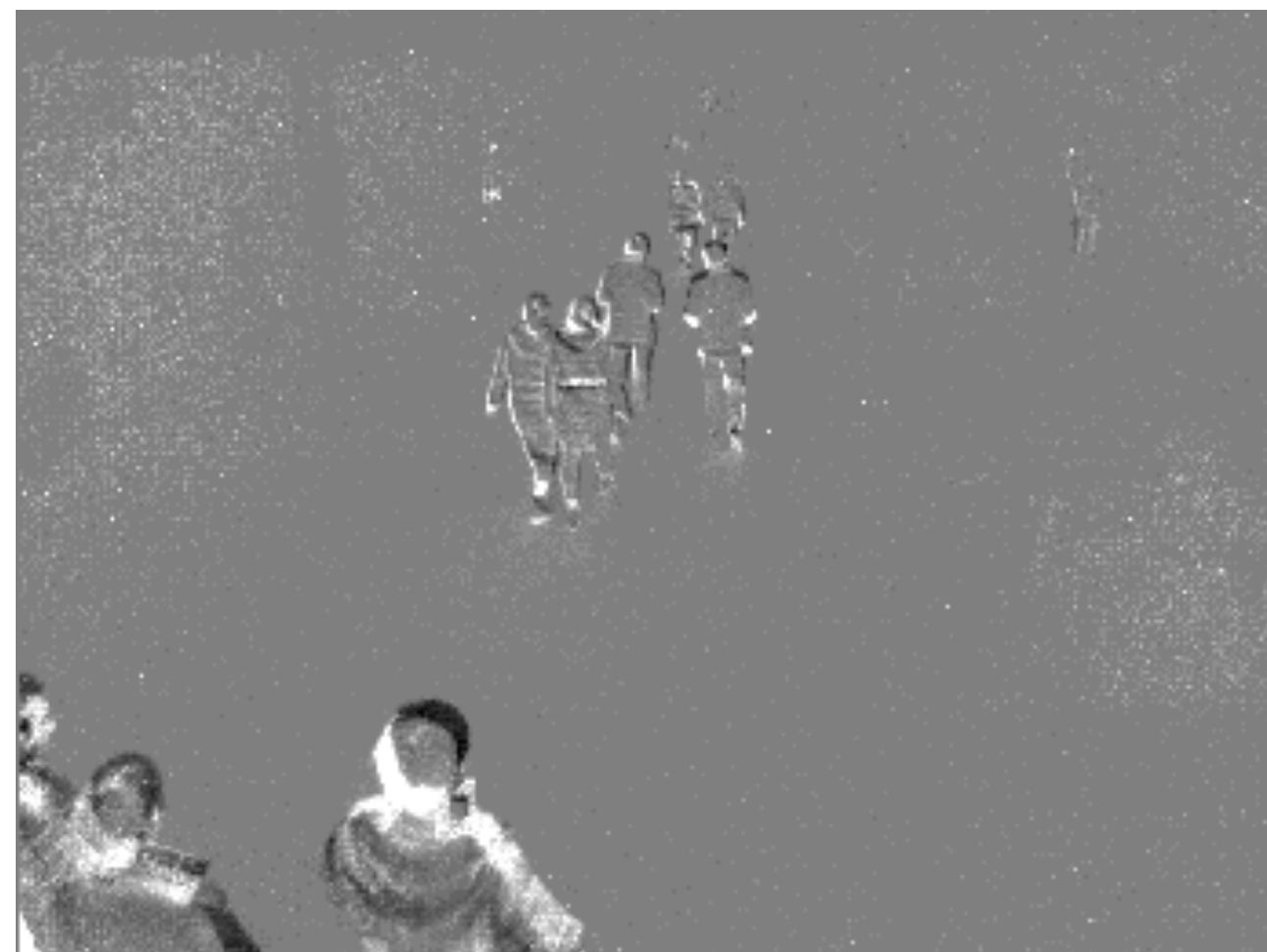
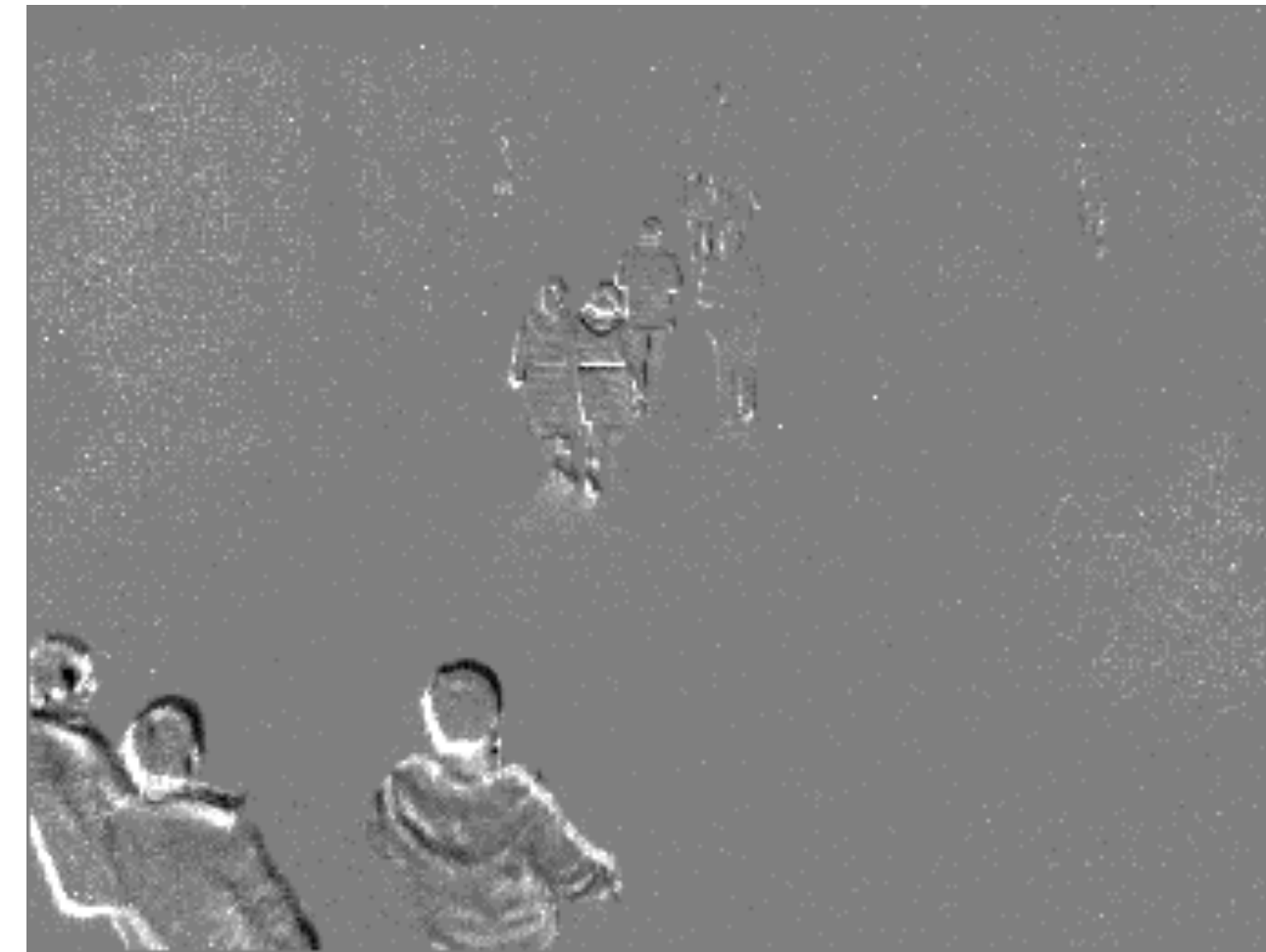
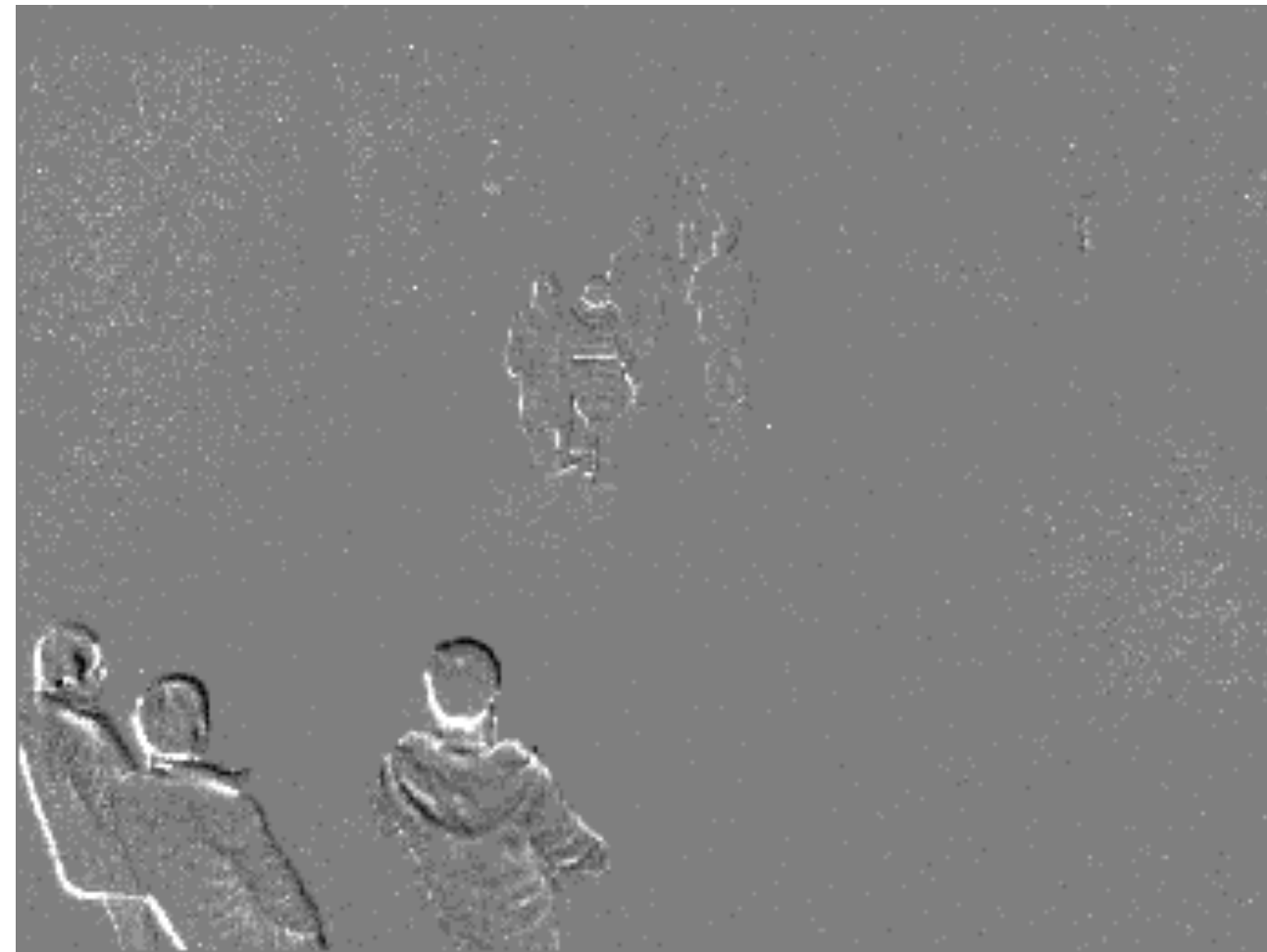
Object Detection
(Yolov3)



Pretrained

Data Augmentation

한 프레임당 오랜 시간을 기록한다. = 같은 시간에 빠른 빠르기가 기록된다.



(a) (b)

(c) (d)

- 이미지가 담고 있는 시간정보

(a) 2만 마이크로초

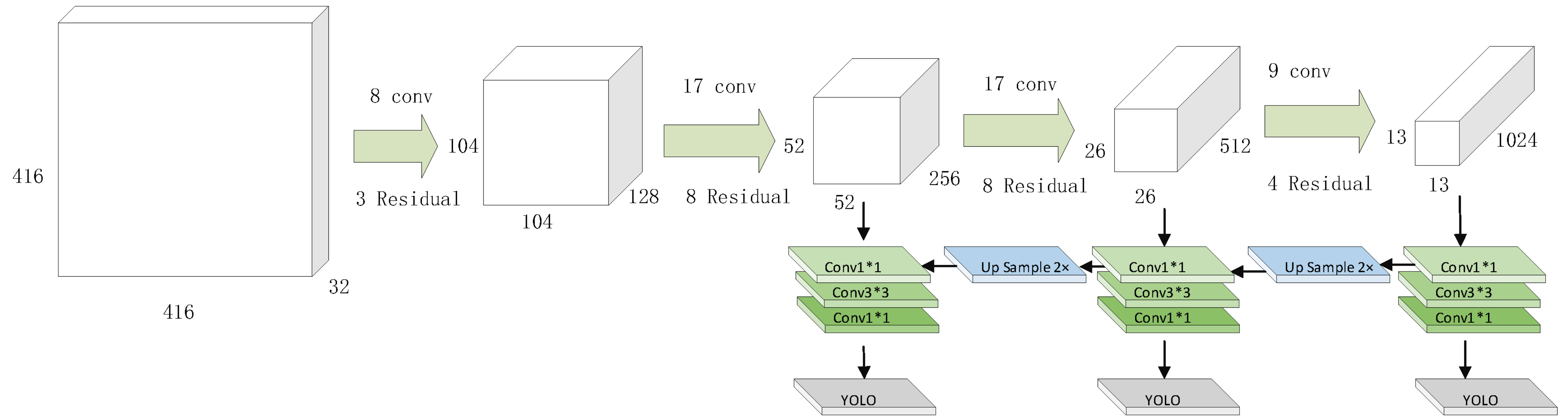
(b) 4만 마이크로초

(c) 8만 마이크로초

(d) 16만 마이크로초

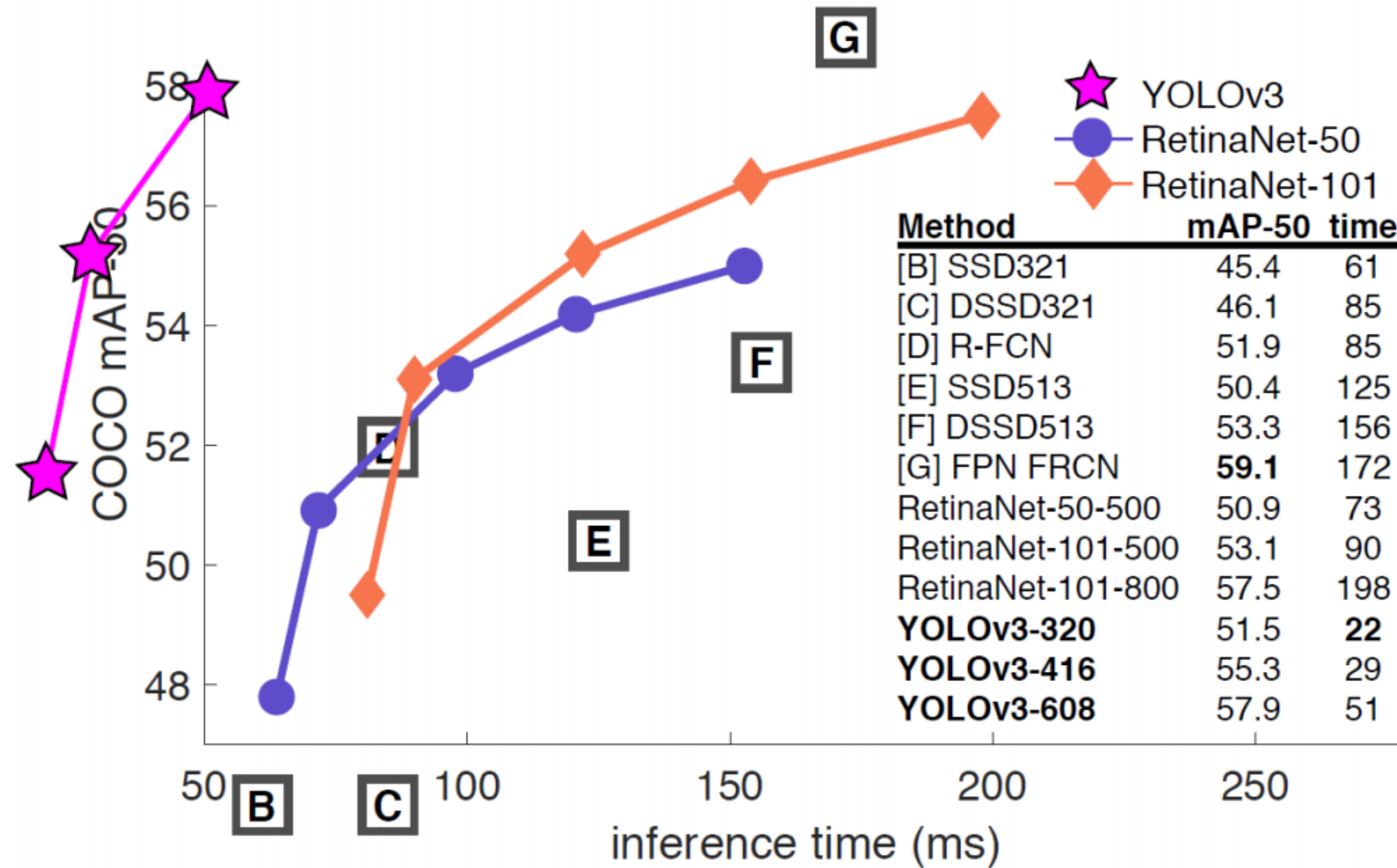
Yolo-v3 : You Only Look Once

Architecture

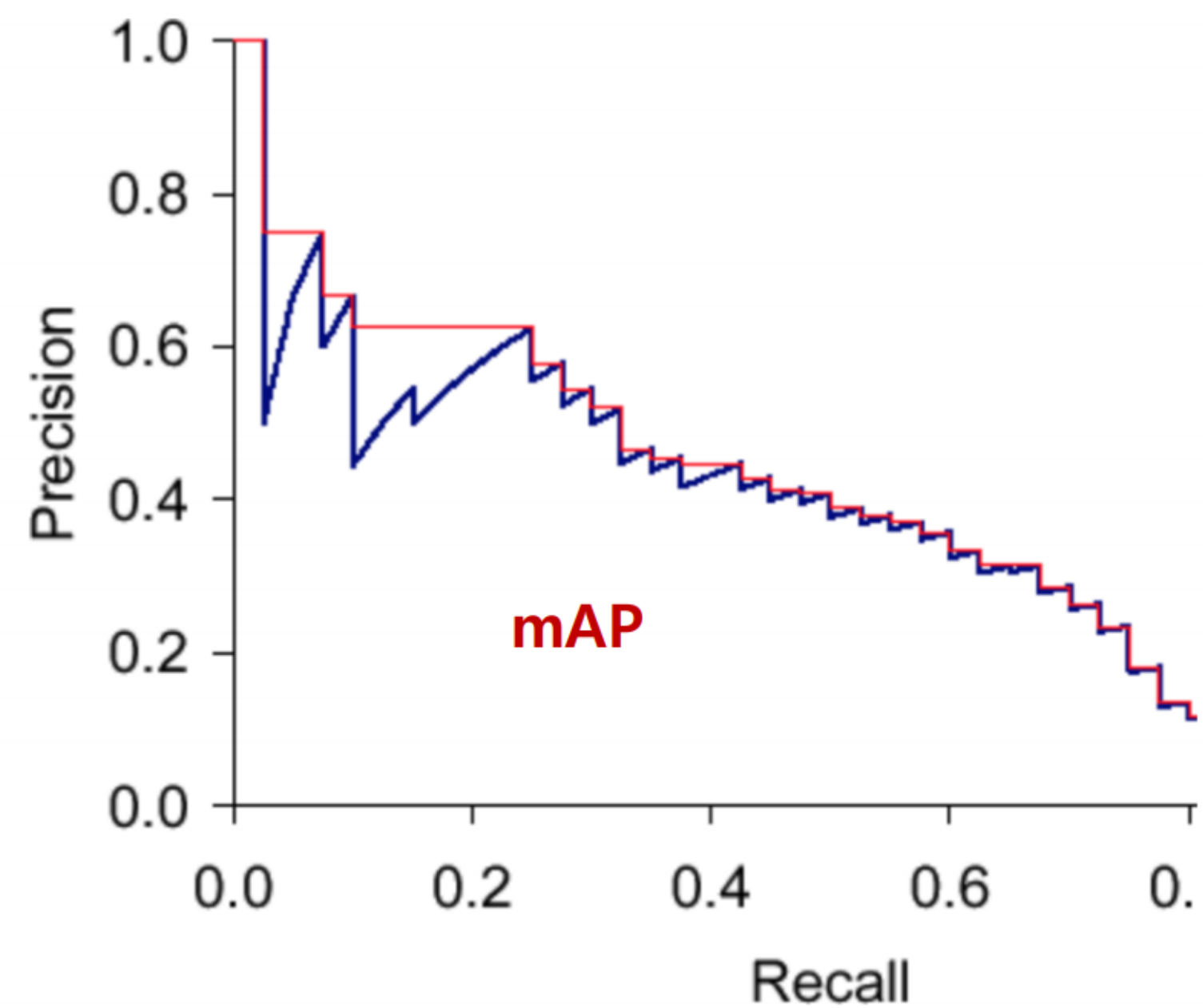
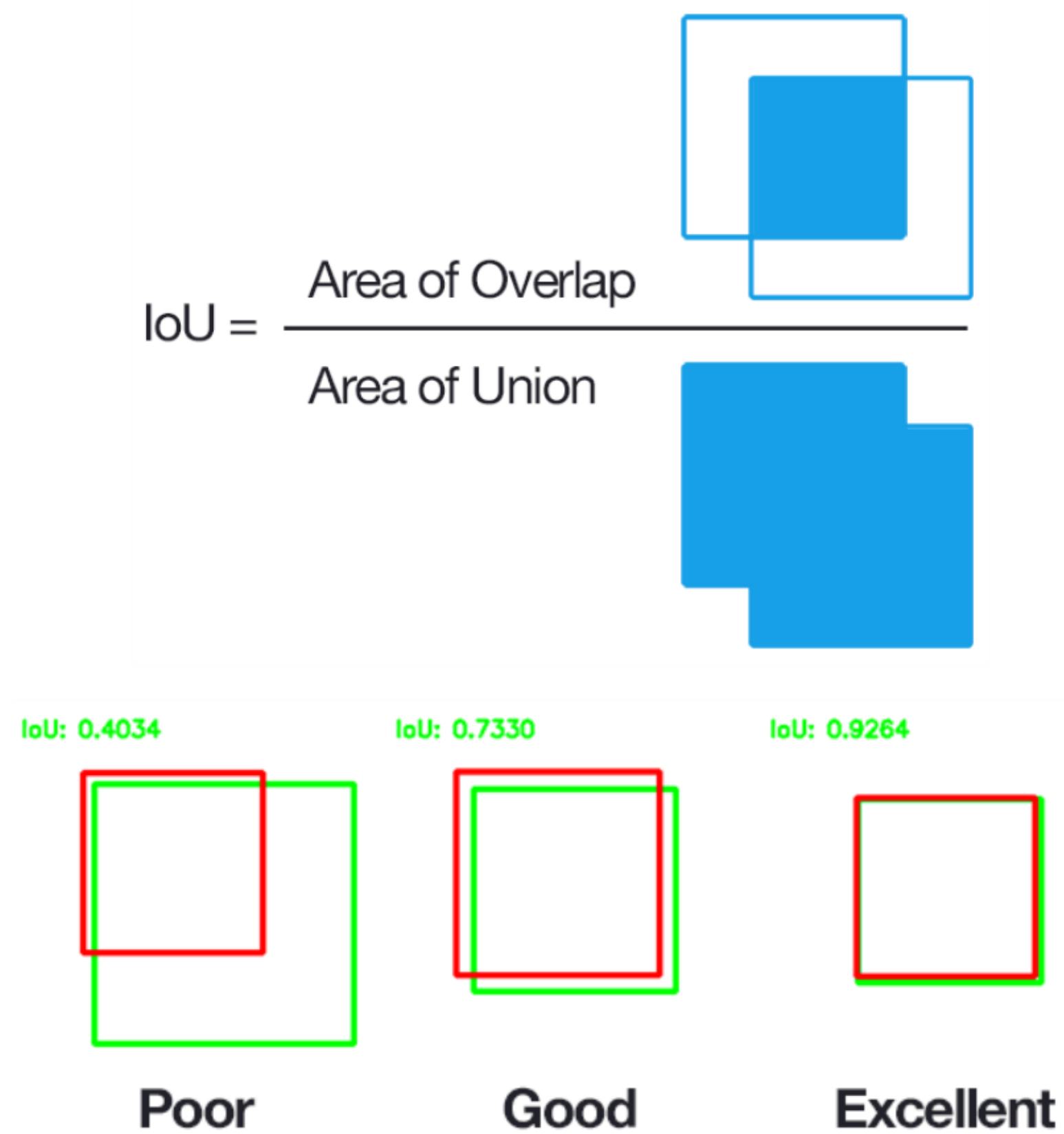


Yolo-v3 : You Only Look Once

Unified, Real-Time Object Detection

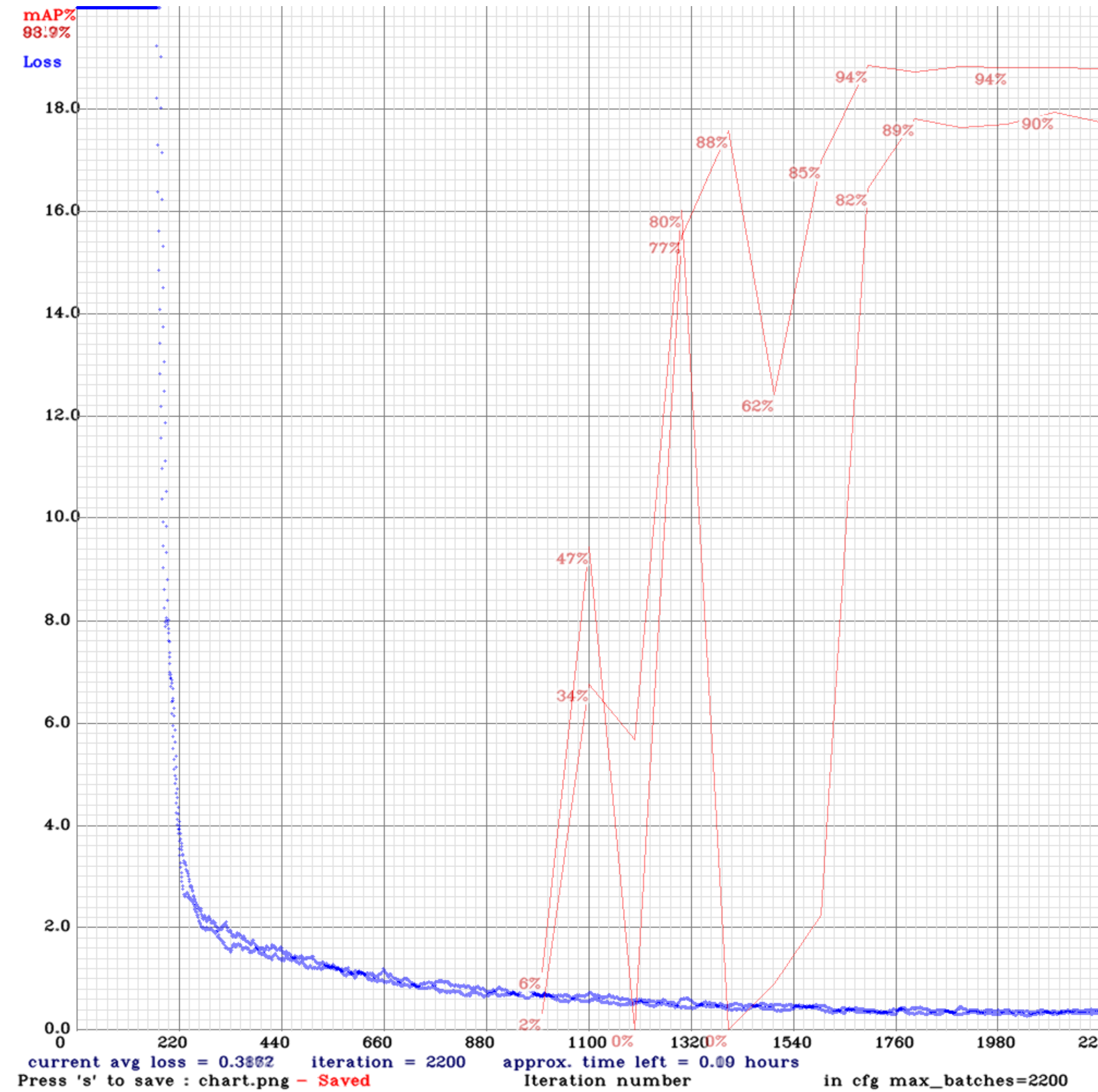


성능 평가 기준

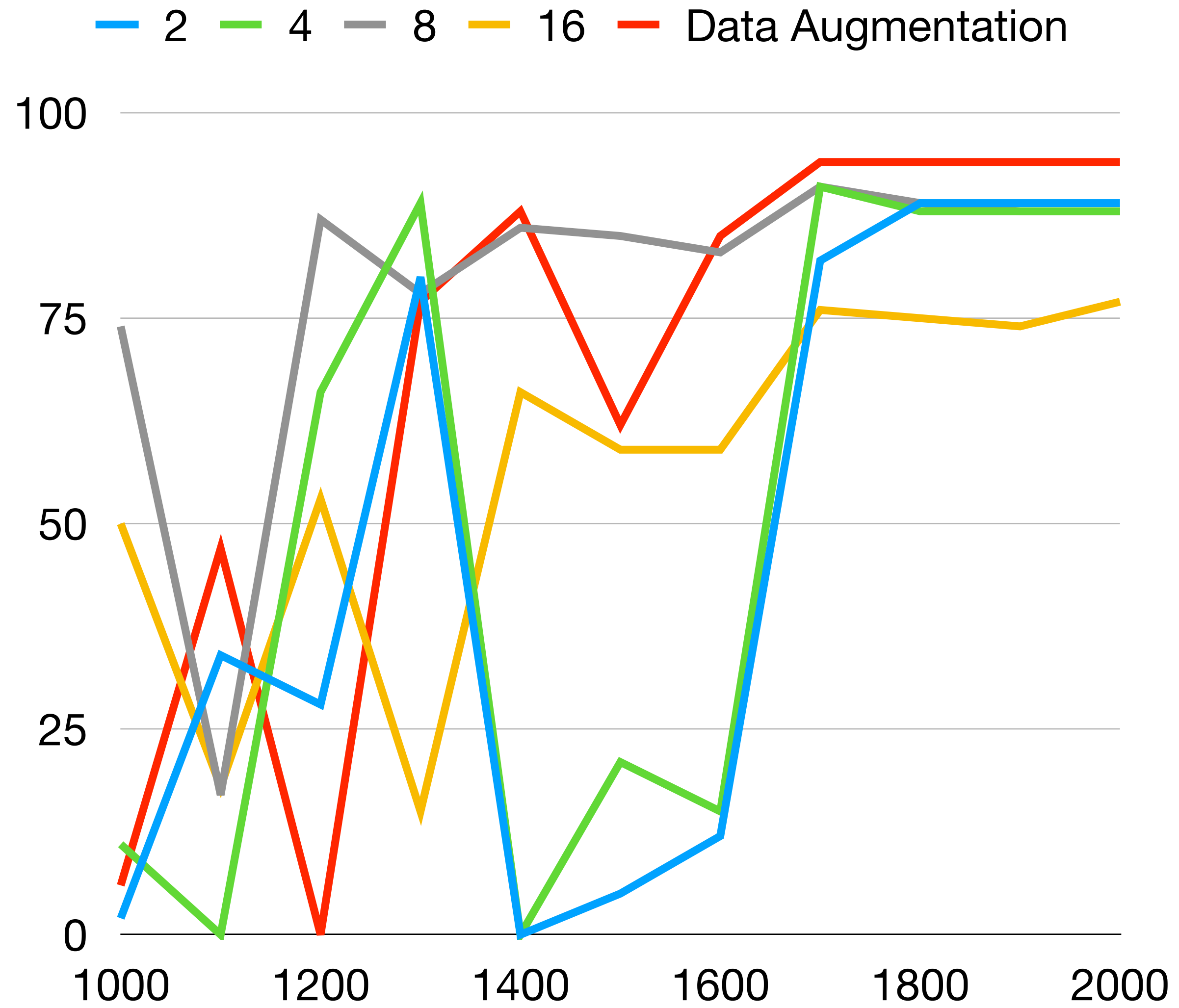


The area under this Precision-Recall curve gives you the “Average Precision”.

Yolov3 에 학습 (General vs Data Augmentation)



Yolov3 에 학습 (mAP 그래프 비교)



성능 평가표

	마지막 iteration 시의 평균 loss값	마지막 iteration 시의 mAP	학습 중 best mAP
2만 마이크로초(General)	0.317746	88.7%	89.62%
4만 마이크로초	0.317564	84.6%	91.07%
8만 마이크로초	0.279849	85.1%	90.80%
16만 마이크로초	0.211214	76.2%	77.37%
Data Augmentation	0.386162	93.9%	94.19%

성능평가표

(conf_thresh = 0.25)	precision	recall	F1-score	mAP@0.50	Detect time
2만 마이크로초(General)	0.80	0.76	0.78	88.72%	25ms /image
4만 마이크로초	0.89	0.76	0.82	84.65%	25ms /image
8만 마이크로초	0.92	0.78	0.84	85.13%	25ms /image
16만 마이크로초	0.78	0.74	0.76	76.18%	25ms /image
Data Augmentation	0.90	0.86	0.88	93.88%	25ms /image

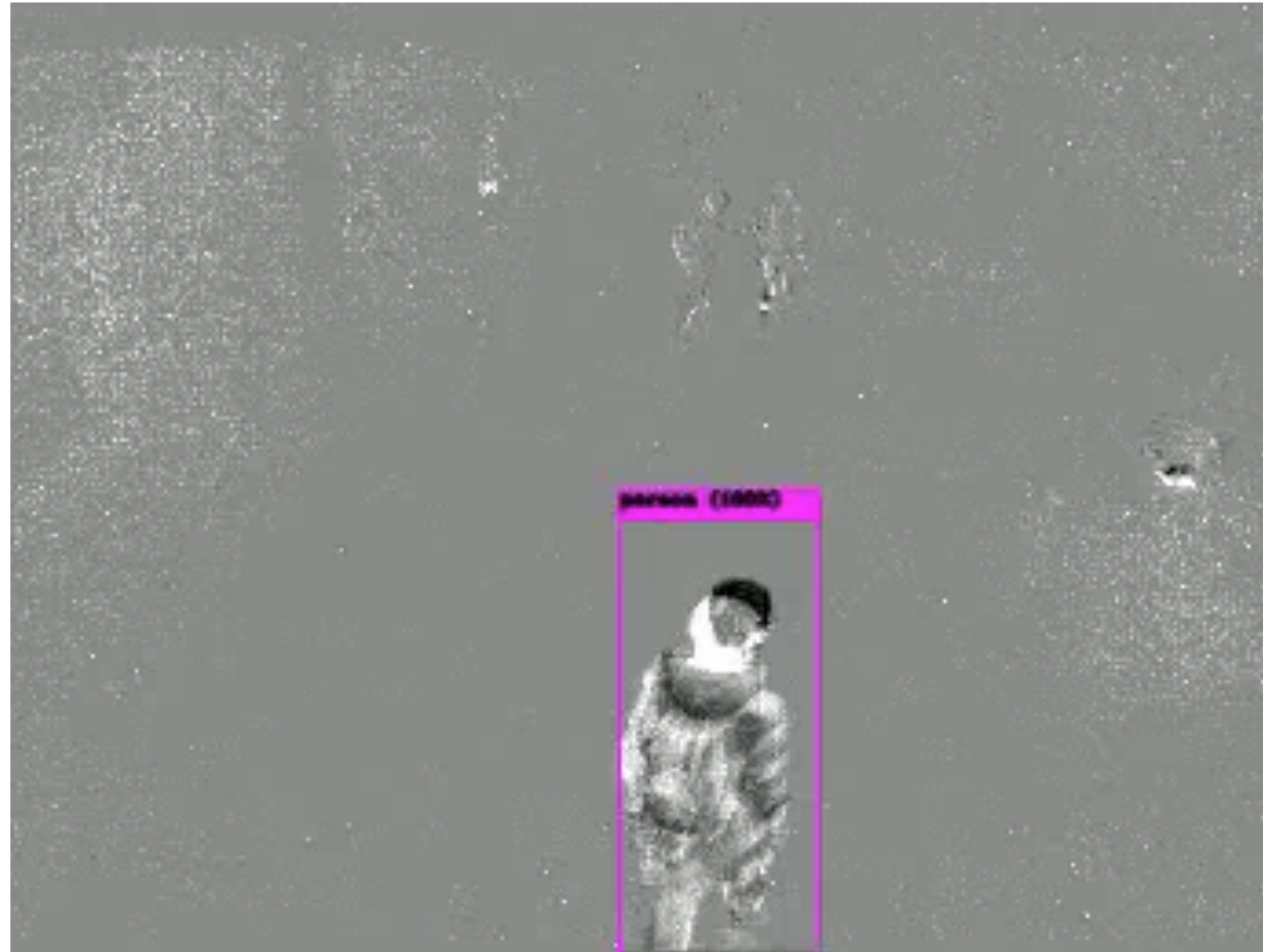
Environment

외장 SSD에 우분투 사용중.

- 맥북프로 2017년형 + 외장 SSD에 (Ubuntu 16.04.3 LTS) (ext4) format
- 데이터셋 개발환경 : NCloud (Ubuntu 16.04.3 LTS)
 - 1 : 8vCPU, 32GB Mem
 - 2 : GPU - Tesla V100, 32GB Mem with/ 8vCPU, 90GB Mem)
- 학습환경 : Colab Pro (Ubuntu 18.04.3 LTS)
 - GPU - Tesla P100, 16GB
 - 4 x Intel(R) Xeon(R) CPU @ 2.30GHz 32GB Mem

Result Work - Yolo-v3 weights

Application : 야간 보안등에서의 객체 탐지.



Result Work - Yolo-v3 weights

- 본 과제를 통해, 어두운 밤과 같이 기존 카메라로 관측이 이루어지기 힘든 환경에서 이벤트 카메라를 통한 객체 검출을 가능케하여 조도와 같은 환경에 대한 영향을 매우 줄였으며, One-stage 객체 검출 모델인 Yolo-v3 모델을 활용하여 실시간으로 객체 인식이 가능하다는 점이 특징이다.
- Yolo-v3 : 객체 검출 모델중에서 검출 속도가 빠른 One-stage 모델 중에서 레이턴시가 매우 낮은 편이며, 성능도 뛰어나고, 저의 환경 (CPU, 메모리, 구현적합성)등을 고려하여 제일 적절한 모델로 생각되어 선택했다.
- 메인 아이디어가 Data Augmentation인 만큼, 같은 데이터를 이용하더라도, 본 과제와 같은 방법을 이용해서 더 좋은 성능을 이끌어 낼 수 있으며, 다른 방법들과 충분히 혼용될수 있는 방법이기 때문에 다른 연구들과 같이 사용되어 더욱 높은 성능을 이끌어내도록 기여할 것으로 기대한다.

Future Work - Yolo-v3 weights

- 추가적인 부분으로는, Yolo-v3 모델 자체에서 미검(False Negative)보다 False Positive를 줄이는 데에 초점을 맞추고 있기 때문에 Application의 주요 목적에 맞게, Loss function을 재설계하여 조정함으로써 미검(FN)을 줄이는 데에 초점을 맞추어 추가적인 세부 과제를 진행 할 수 있다.
- 엔드투엔드 모델이 아닌만큼, 변환하는데 시간이 더 걸리게 되고, 이벤트 데이터의 특성상 마이크로초 수준의 시간 분해능을 가지는데, 자율주행 자동차 같이 매우 짧은 시간 내로 실시간 검출이 필요한 경우에 부담이 될 수 있다.
때문에, CNN 대신 End-to-End 모델인 SNN 을 적용하여 디벨롭할 수 있을것으로 기대하고 있다.