

강화학습을 이용한 자율주행 구현

한컴 MDS 아카데미 3조
두발자유화
(TFF : Two Foot Freedom)
날짜: 10 / 1 (금)

팀원: 노용철 정일우 송해찬

목 차

INDEX

1 프로젝트 소개

팀원소개
개발배경 /
개요

2 프로젝트 목표

프로젝트
요구사항

3 프로젝트 내용

시스템 구성
배경이론
구현 내용

S/W & H/W

4 프로젝트 결과

개발 기능
시연 영상

5 결론

결론
향후 활용 가능성

1 프로젝트 소개

팀원 소개



도움: 자이트론 허성민 대표님



팀장: 노용철

역할: Enet-SAD 모델 설계, 모형차 제작



팀원: 정일우

역할: DDPG 모델 설계, Gazebo 시뮬레이터 환경 구성



팀원: 송해찬

역할: Enet-SAD 영상 처리, Carla 시뮬레이터 환경 구성

1 프로젝트 소개

개발배경 및 개요

WAYVE



강화학습을 활용한 자율주행

EPISODE 11

1차선의 오솔길 자율주행 성공

<https://www.youtube.com/watch?v=eRwTbRtnT1I>

1 프로젝트 소개

개발배경 및 개요

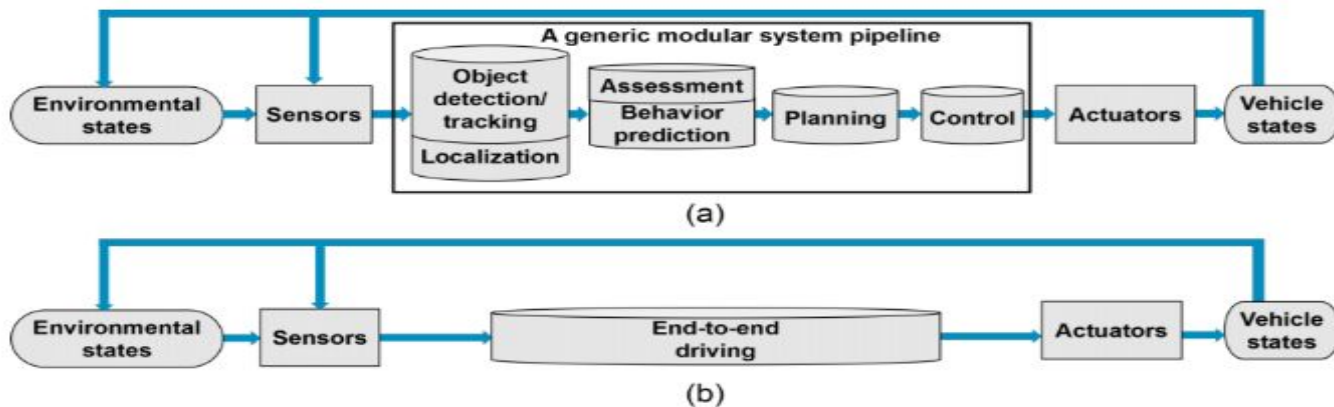


FIGURE 2: Information flow diagrams of: (a) a generic modular system, and (b) an end-to-end driving system.

MODULE

END to END

- 장점 : 축적된 노하우, 전문 지식 이전 가능, 안정적인 아키텍처
- 단점 : 오류 전파, 과도한 복잡성

- End to End 방식으로 모든 상황 학습
- 4레벨 이상의 자율주행 활용 가능성 있음

1 프로젝트 소개

개발배경 및 개요

END to END

Direct supervised
learning

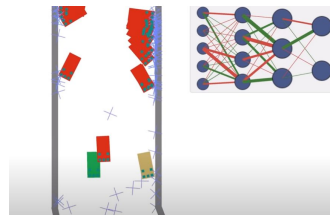
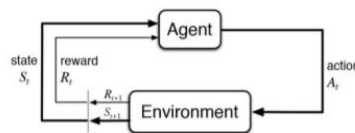
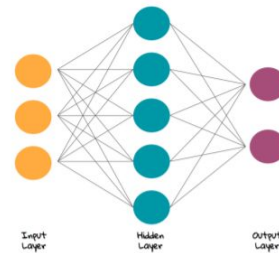
운전자(사람)의 데이터를 imitate

Deep reinforce
learning

경험을 통해 최적의 행동 보상
도달

Neuroevolution

유전 알고리즘을 통해 인공신경망
진화



2 프로젝트 목표

프로젝트 요구사항

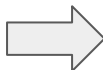
프로젝트 요구사항

1. 모형차가 주행 가능한 영역을
실시간으로 인지해야 한다.



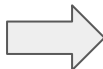
실시간 영역분할 (Image Segmentation)
모델을 적용한다.

2. DDPG 강화학습 모델이 10분 내에
부딪히지 않고 주행할 수 있도록 최적의
보상(Reward)이 설계되어야 한다.



최적의 보상을 설계하기 위해 시뮬레이터를
이용하고 모형차에 적용한다.

3. 모형차에서 강화학습 시키기 전에
최적의 학습 조건을 찾아야 한다.



시뮬레이터에서 최적의 강화학습 조건을
찾는다.

3 프로젝트 내용

시스템 구성



3 프로젝트 내용

배경이론

영역 분할 : segmentation model

ENET - SAD

Rank	Model	Accuracy↑	F1 score	Extra Training Data	Paper	Code	Result	Year	Tags
1	FOLOLane (ERFNet)	96.92		×	Focus on Local: Detecting Lane Marker from Bottom Up via Key Point	GitHub	Kaggle	2020	
2	RESA	96.82	96.93	×	RESA: Recurrent Feature-Shift Aggregator for Lane Detection	GitHub	Kaggle	2020	
3	ENet-SAD	96.64%	95.92	×	Learning Lightweight Lane Detection CNNs by Self Attention Distillation	GitHub	Kaggle	2019	

Category	SCNN	RESA-34	RESA-50	Category	Proportion	ENet-SAD	SCNN [16]
Normal	90.6	91.9	92.1	Normal	27.7%	90.1	90.6
Crowded	69.7	72.4	73.1	Crowded	23.4%	68.8	69.7
Night	66.1	69.8	69.9	Night	20.3%	66.0	66.1
No line	43.4	46.3	47.7	No line	11.7%	41.6	43.4
Shadow	66.9	72.0	72.8	Shadow	2.7%	65.9	66.9
Arrow	84.1	88.1	88.3	Arrow	2.6%	84.0	84.1
Dazzle light	58.5	66.5	69.2	Dazzle light	1.4%	60.2	58.5
Curve	64.4	68.6	70.3	Curve	1.2%	65.7	64.4
Crossroad	1990	1896	1503	Crossroad	9.0%	1998	1990
Total	71.6	74.5	75.3	Total	—	70.8	71.6
Runtime (ms)	116	22	28	Runtime (ms)	—	13.4	133.5
				Parameter (M)	—	0.98	20.72

HIGH ACCURACY

FAST RUNTIME

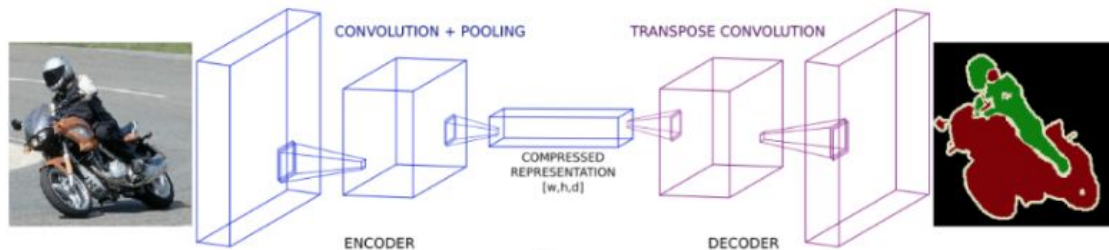
ENET - SAD

3 프로젝트 내용

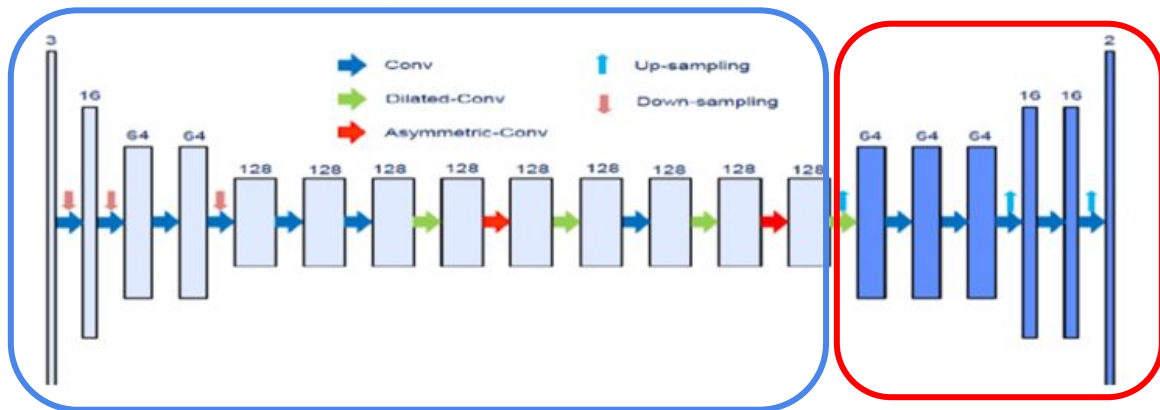
배경이론

영역 분할 : segmentation model

ENET - SAD



Semantic Segmentation: 인코더 — 디코더의 형태



3 프로젝트 내용

배경이론

영역 분할 : segmentation model

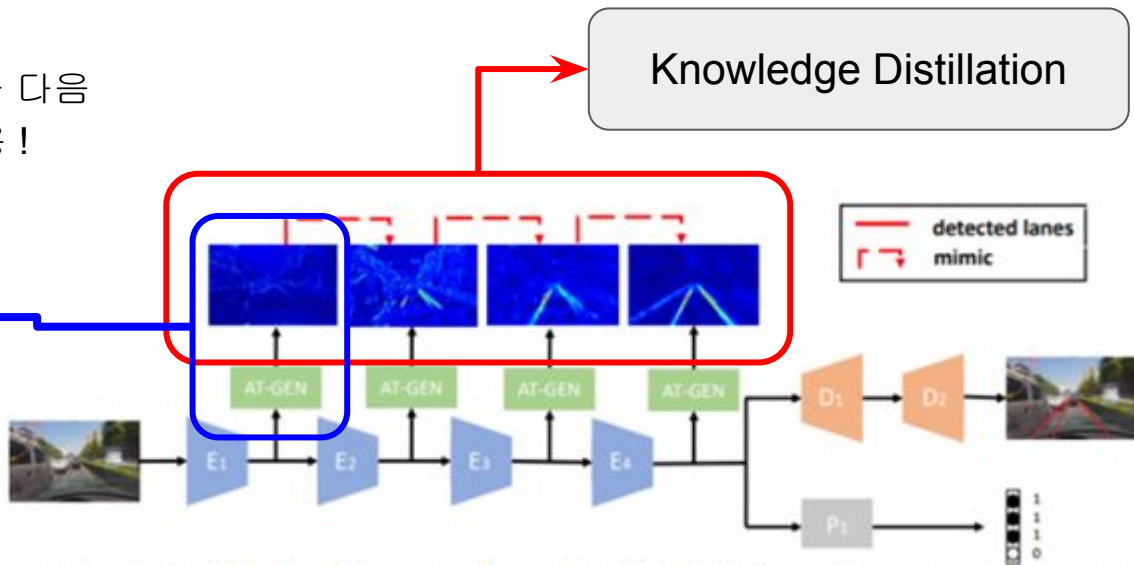
ENET - SAD

Self Attention Distillation

현재 layer에서 생성된 Attention map을 다음 layer의 Distillation(증류) 대상으로 활용 !

Attention transfer

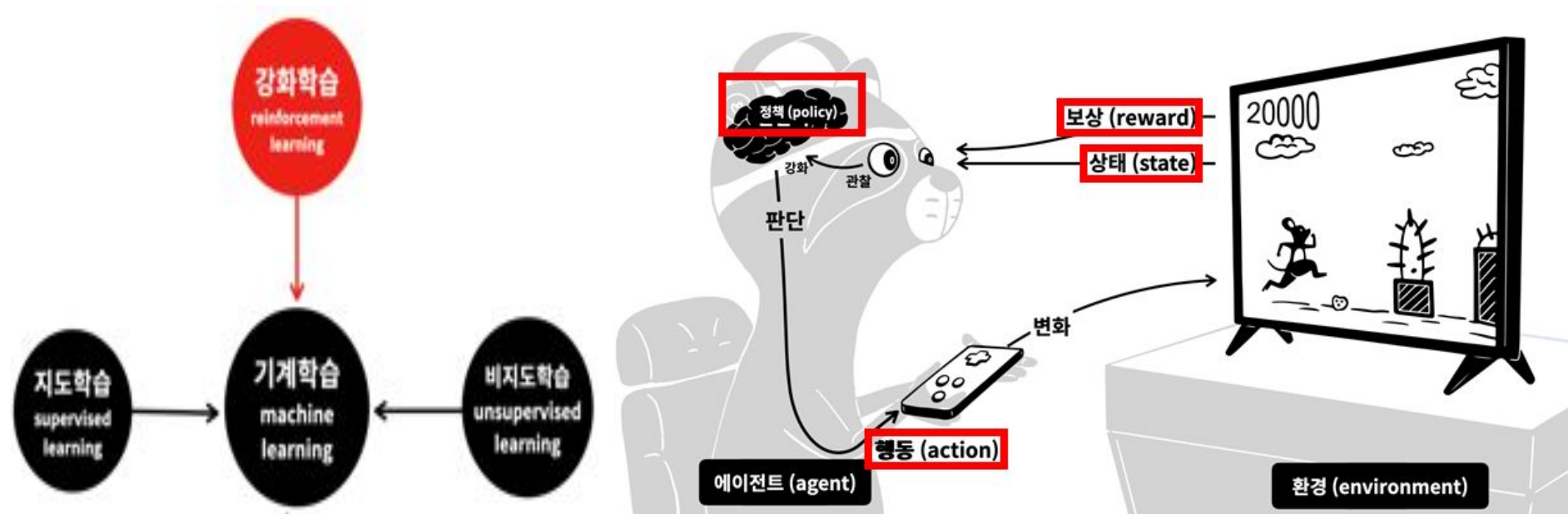
Knowledge Distillation



3 프로젝트 내용

배경이론

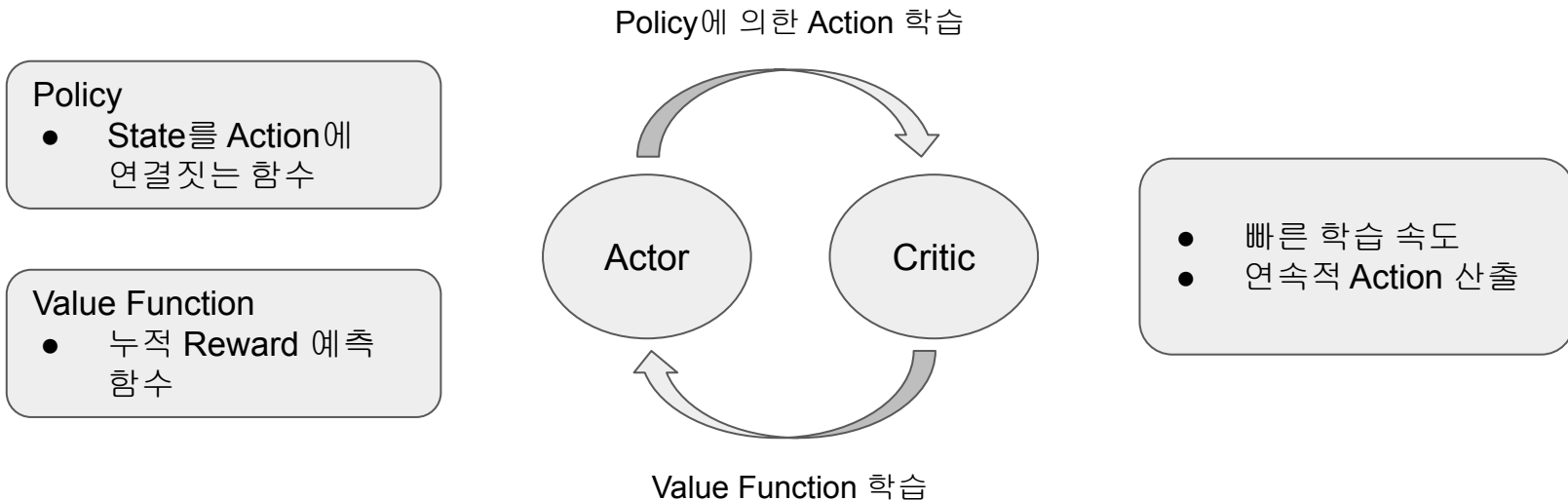
강화 학습 (Reinforcement Learning)



3 프로젝트 내용

배경이론

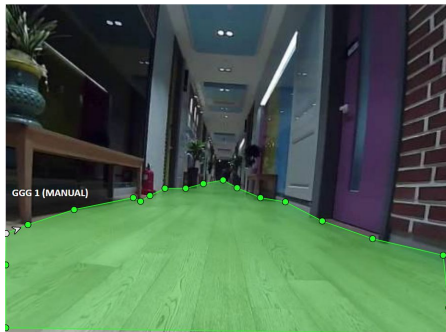
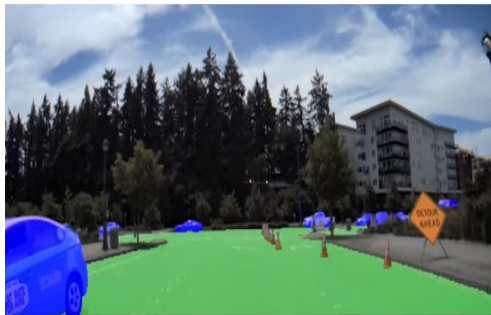
DDPG (Deep Deterministic Policy Gradient)



3 프로젝트 내용

구현 내용

ENET - sad 학습



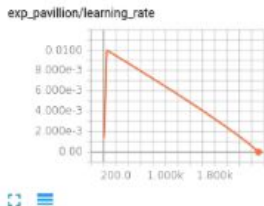
Labeling



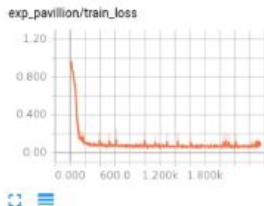
Xavier NX - 10 FPS

300장 Labeling
600장 이미지 증식

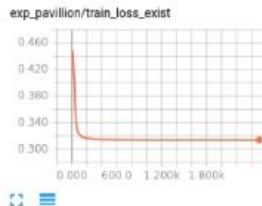
- 밝기 조절
- 기울임



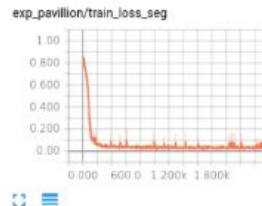
learning_rate



total_loss



exist_loss



segmentation_loss

exist loss = 0.315

seg loss = 0.021

total loss = 0.336

3 프로젝트 내용

구현 내용

Reward 설계

$$\alpha \sin \theta - \beta \cos \theta$$

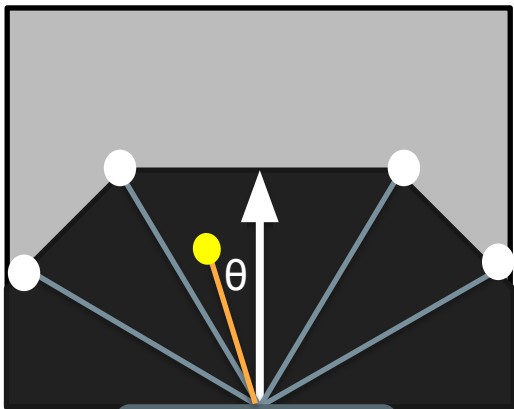
α, β : 하이퍼 파라미터

Segmentation 이미지에서
도로 영역만 추출

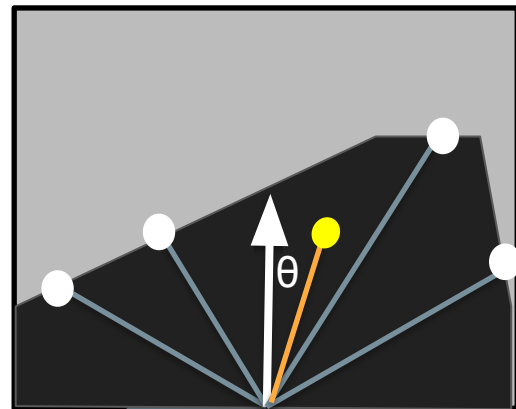
차량 중심에서 5 방향 직선
생성

하얀색 점들의 중점을 생성
(노란점)

차량 전방 화살표와 주황색선
사이의 각도 θ



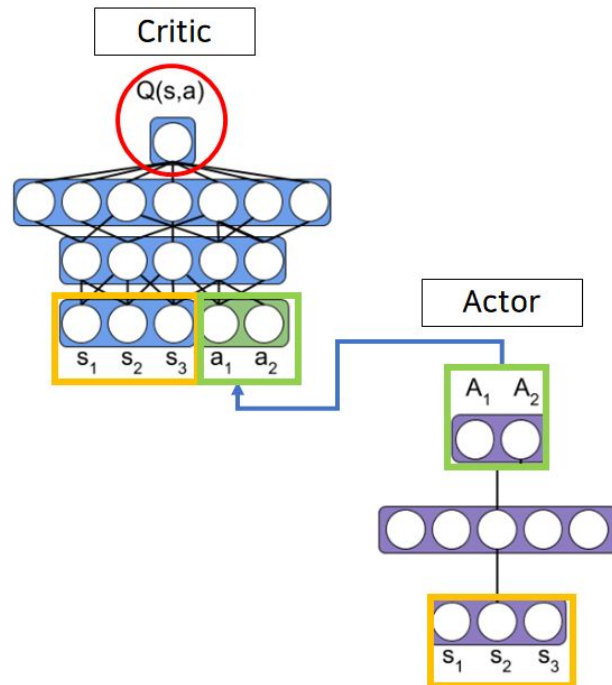
차량의 중심



차량의 중심

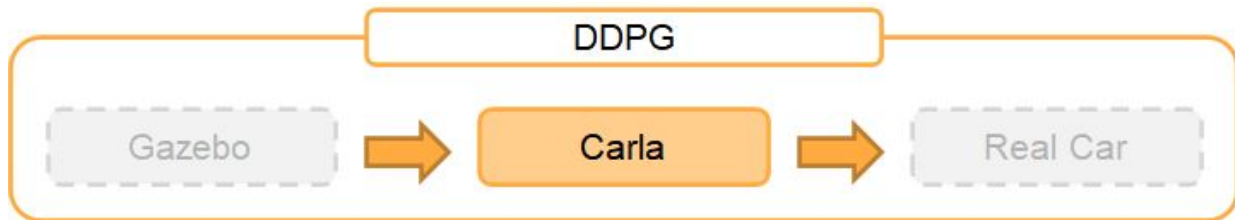
3 프로젝트 내용

시뮬레이터 S/W

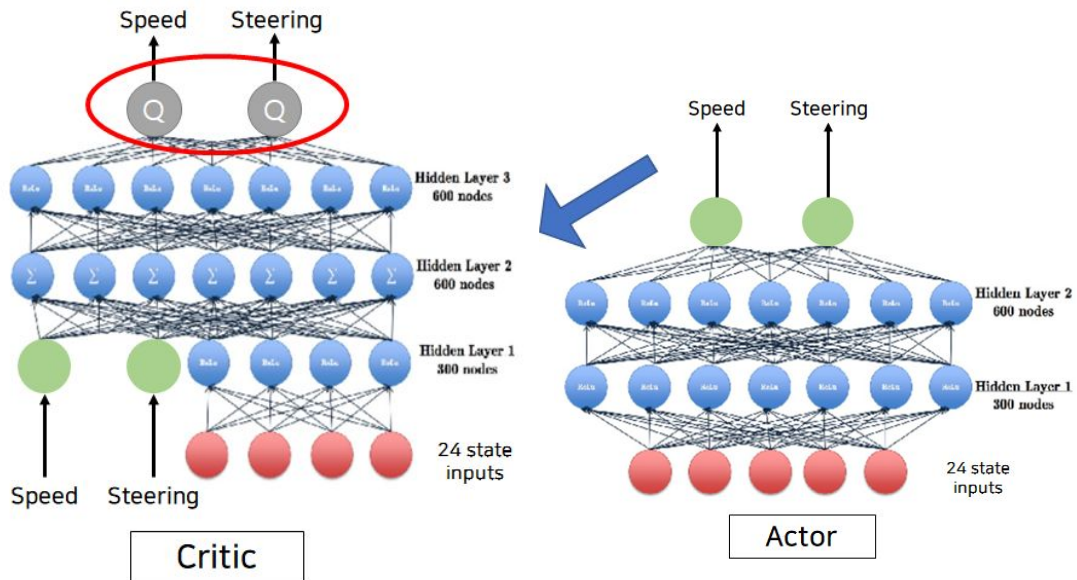


3 프로젝트 내용

시뮬레이터 S/W



DDPG 네트워크 변경

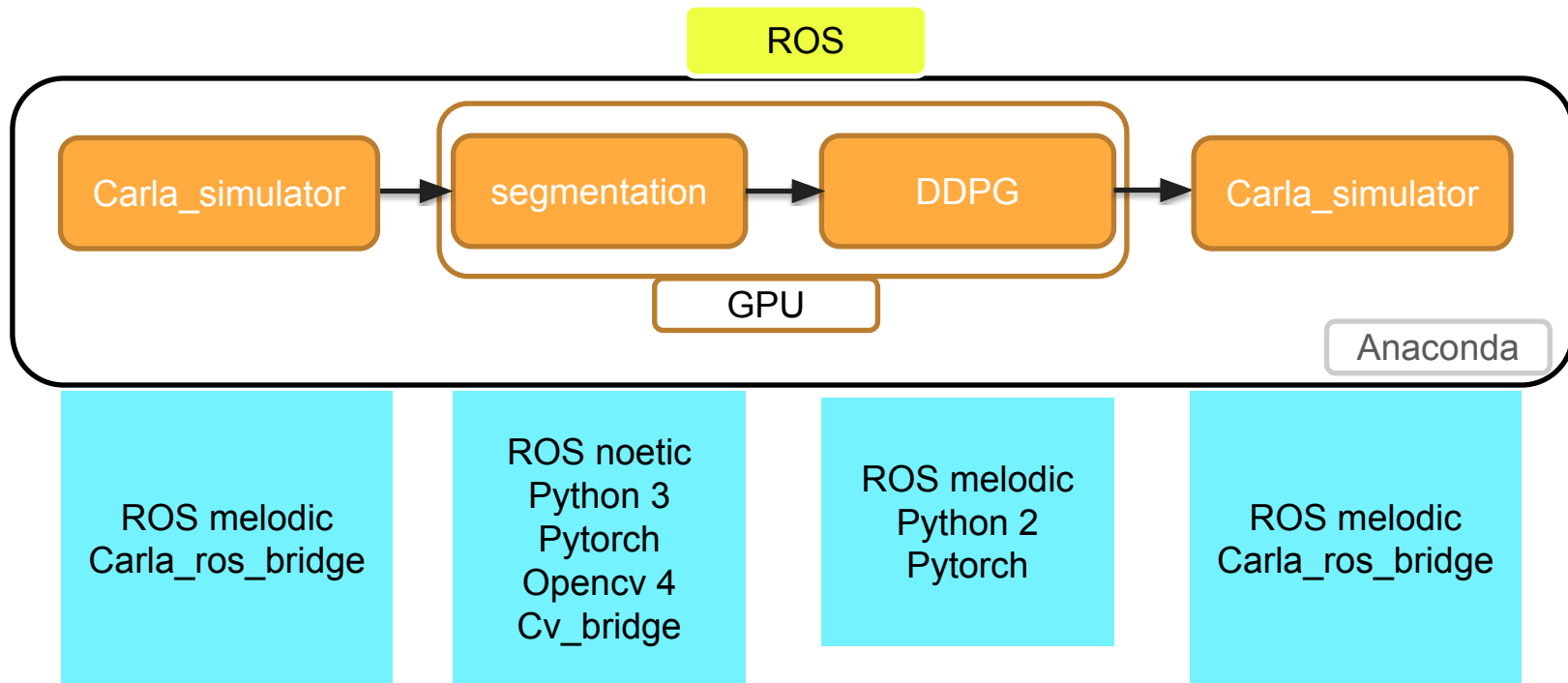
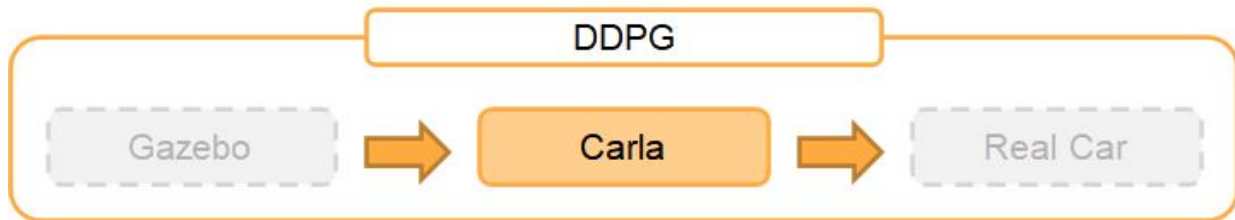


충돌 시 리스폰 기능
추가

velocity \rightarrow Q
angle \rightarrow Q
각각 적용

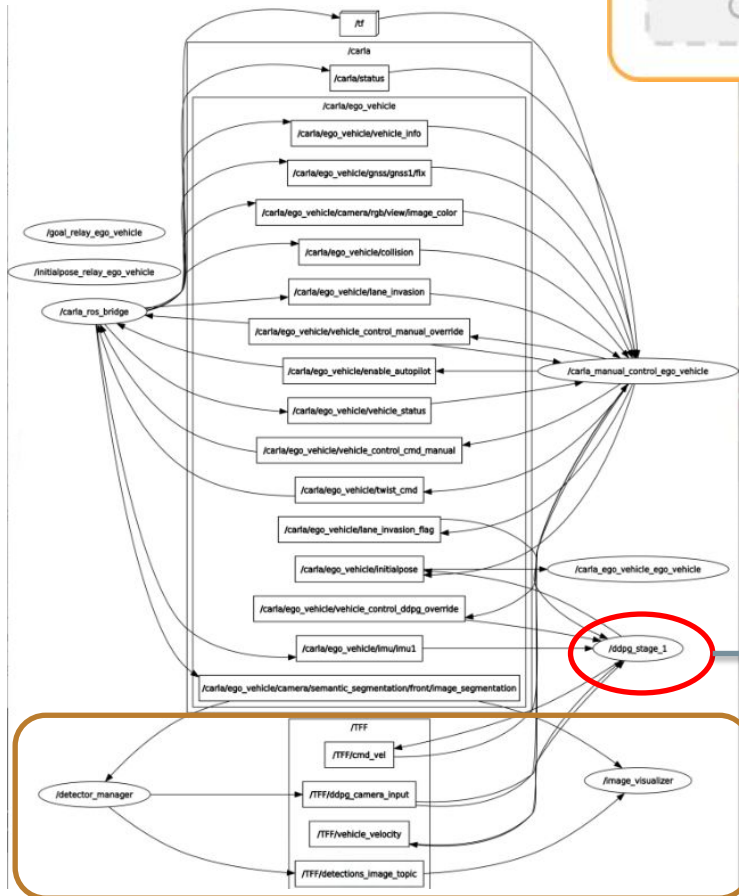
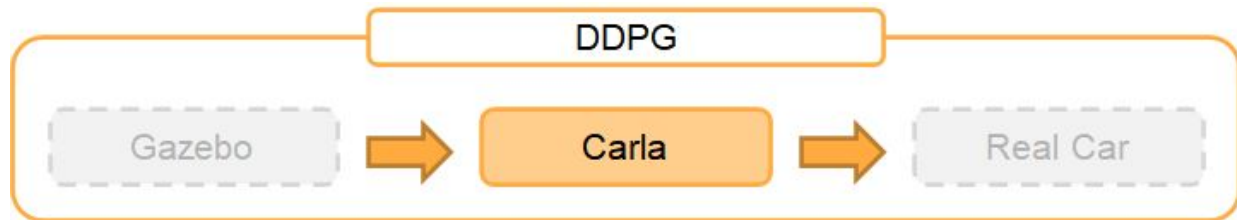
3 프로젝트 내용

시뮬레이터 S/W



3 프로젝트 내용

시뮬레이터 S/W



DDPG Topic list

(subscribe)

IMU 6

Ddpd_camera_input 5x3

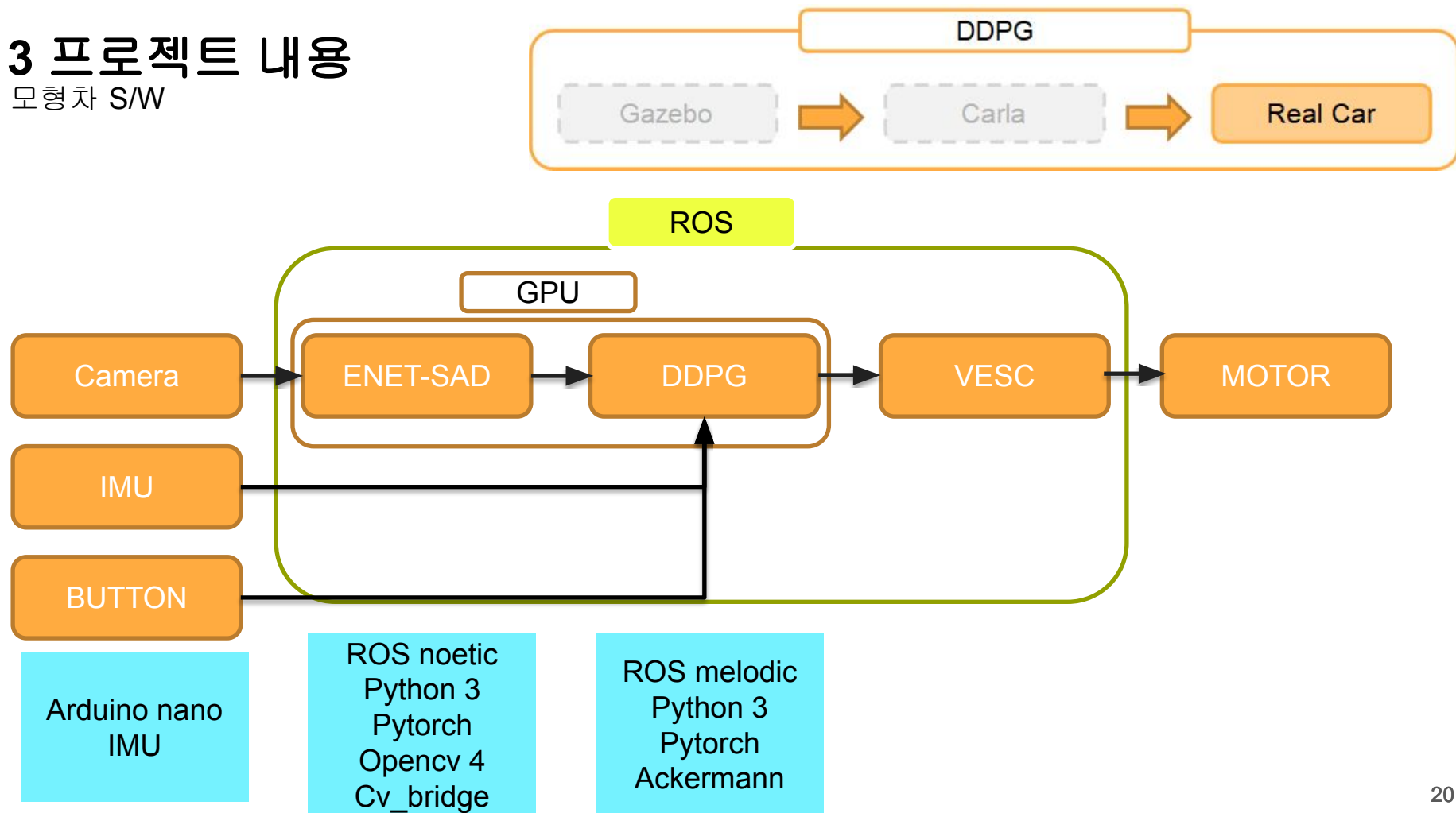
(publish)

Ackermann_Speed 1

Ackermann_Steer 1

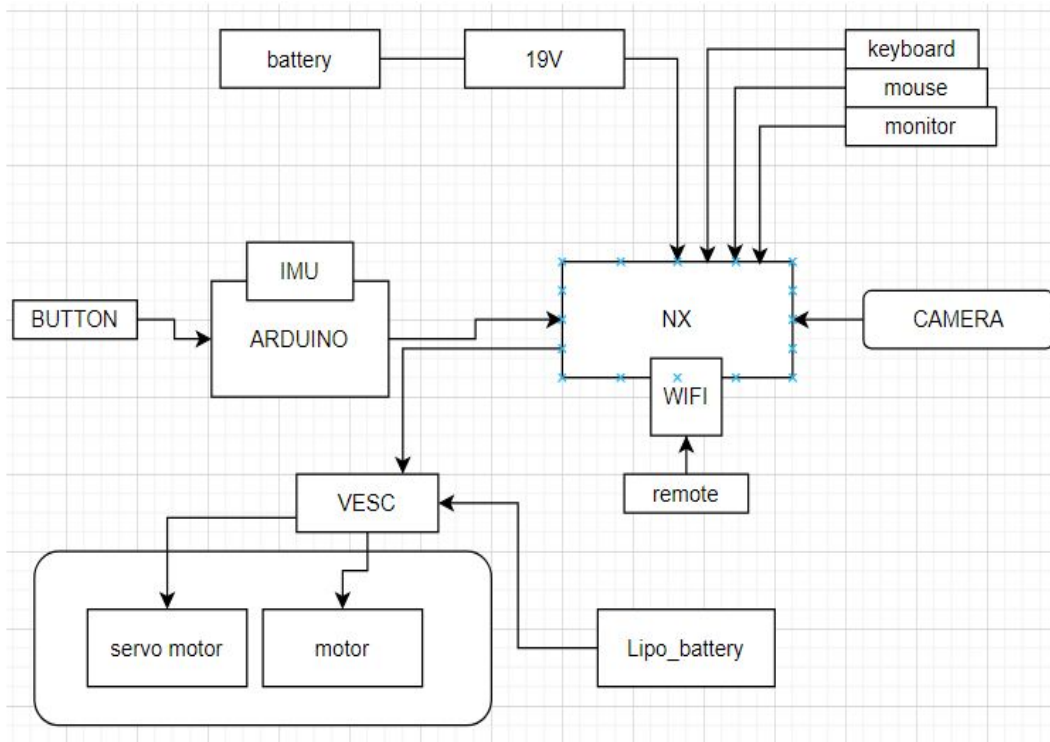
3 프로젝트 내용

모형차 S/W



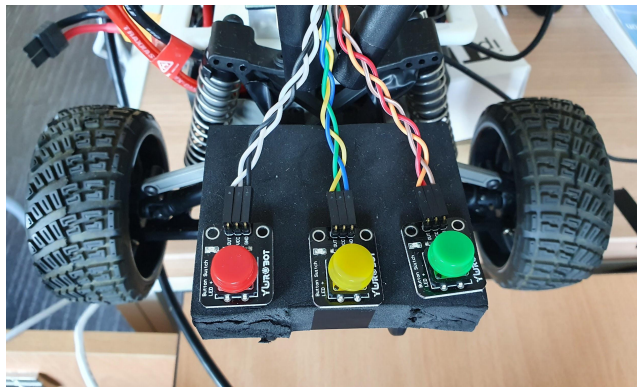
3 프로젝트 내용

모형차 H/W

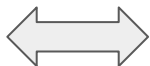


4 프로젝트 결과

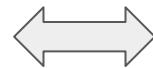
개발 기능



Manual
Control
가능



아두이노 버튼



강화학습



Episode 비상 종료
(Penalty Reward)



Episode 일반 종료



Episode 시작

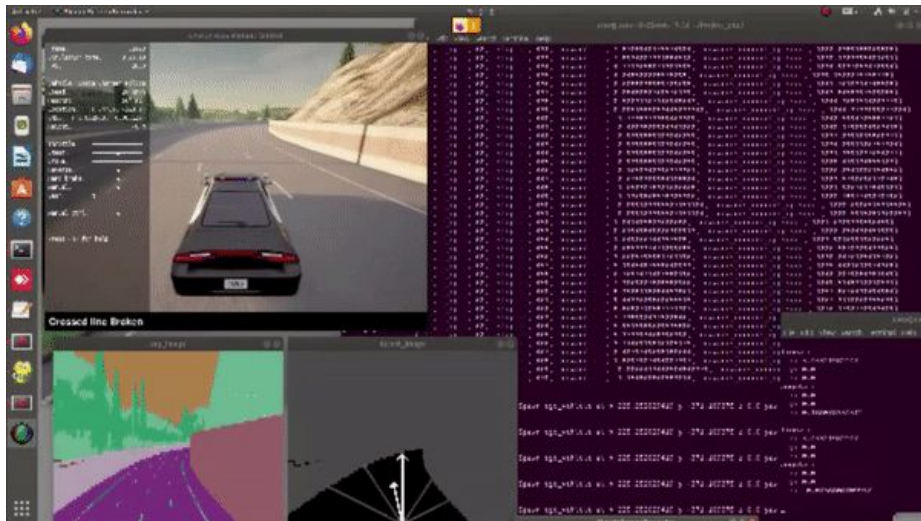


버튼 연동
가능



4 프로젝트 결과

시연 영상

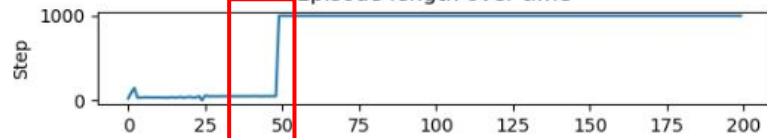


시뮬레이터(carla) DDPG 학습 그래프

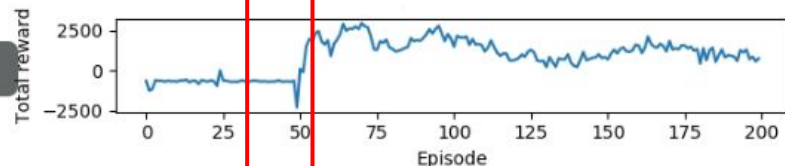
episode 50에서 최고 step 도달

약 1분만에 학습 완료

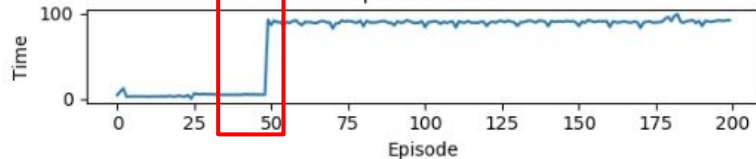
$\alpha=10$ $\beta=50$ $\gamma=0.01$ $V_{max}=5.5$ $duration=13794.2s$
Episode length over time



Episode reward over time



Episode time



4 프로젝트 결과

시연 영상

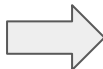


5 결론

결론

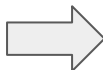
프로젝트 목표

1. 모형차가 주행 가능한 영역을 인지하기 위해 실시간 영역분할 (Image Segmentation) 모델이 필요하다.



Enet-Sad 모델을 통해 FPS 10의 실시간 영역분할 모델 적용 성공

2. DDPG 강화학습 모델이 10분 내에 부딪히지 않고 주행할 수 있도록 최적의 보상(Reward)이 설계되어야 한다.



모형차에서 10분 (60 Episode) 만에 강화학습 주행 성공

3. 모형차에서 강화학습 시키기 전에 최적의 학습 조건을 찾아야 한다.



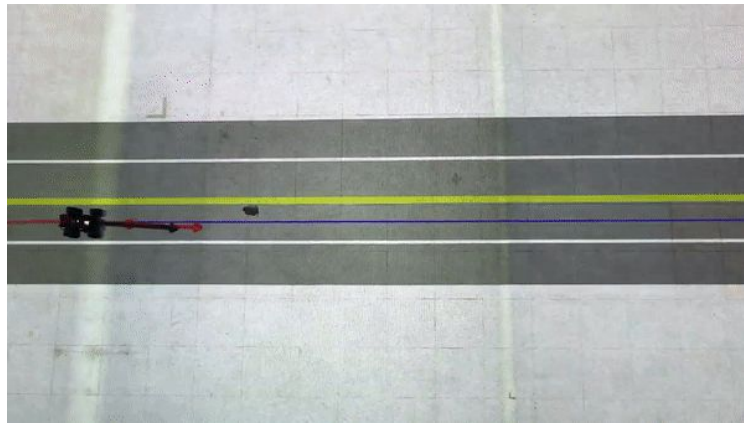
Gazebo, Carla 시뮬레이터에서 강화학습 적용을 통해 Reward의 최적 하이퍼 파라미터 추출 성공 ($\alpha=10, \beta=50$)

5 결론

결론

향후 개선방안

1. 급곡선 주행
 - NX -> AGX , TensorRT를 적용하여 fps 향상 (빠른 자세 제어)
 - PID 제어
2. ENET - SAD 오검출
 - 학습 데이터 추가 학습
 - CV2를 활용하여 ROI 영역 검출



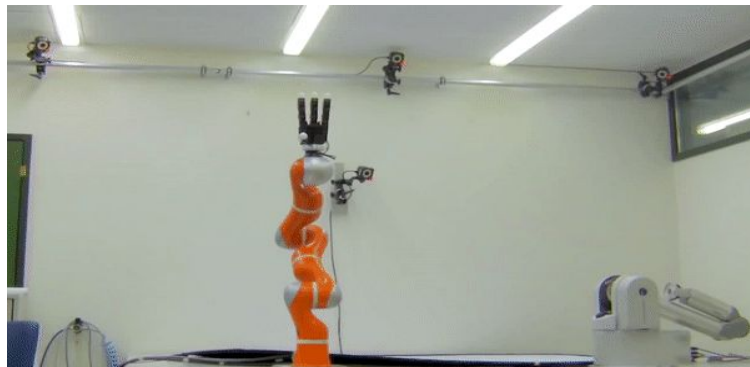
5 결론

결론

강화학습(DDPG)가 적용될 수 있는 영역

빠른 학습과 연속적인 값 산출이 특징

- 로봇의 무게 중심 제어
- 다양한 모양의 물건 **grapping** 가능



Q & A

