

REINFORCE

LEARNING

2025.06.17

# F1TENTH

## TEAM PROJECT

404 map Not Found (권유진, 여예진, 이민영, 허채연)

## PROJECT

01

### Trial and Error

- 초기 시도한 방법들 (기존 DQN 기반)
- Main Experiment에 사용할 Method 선정

02

### Experiment

- Ablation Study로 Experiment 진행
- Speed 변수를 중점적으로 실험

03

### Our Model

- SOTA Model 알고리즘 Idea 설명
- map easy/Oschersleben 주행 영상 & lap time

## 01. Trial and Error

### - 초기 시도한 방법들 (기존 DQN 기반)

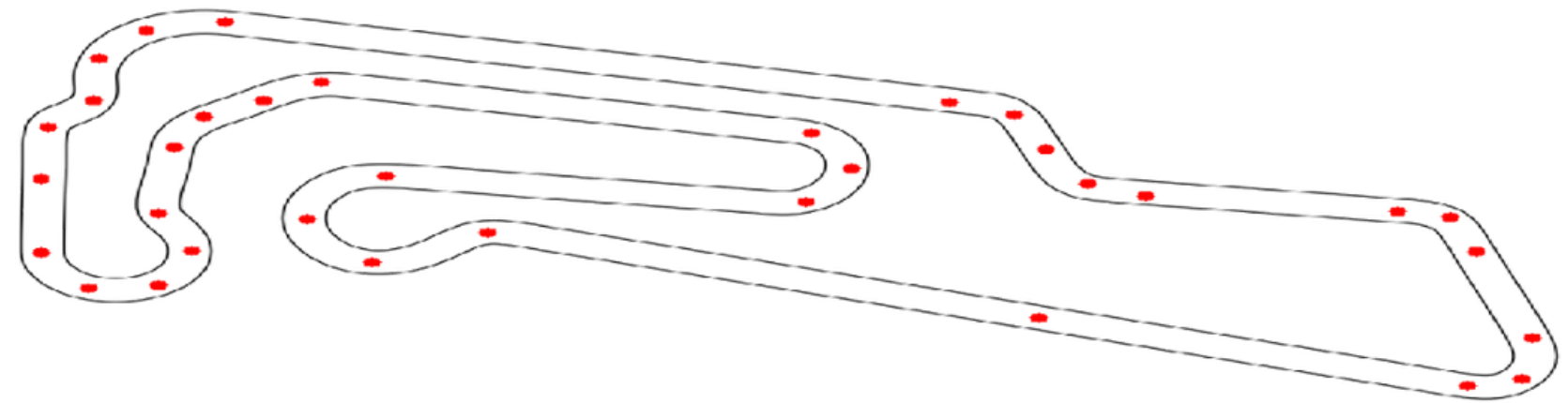
#### 1. Reward Shaping

항목	내용
장애물 거리 패널티	$\min(\text{scan}) < 0.5$ 일 때 -5
정면 주시 보상	$300 < \min\_idx < 780$ 일때 장애물이 정면 근처 : -1, 후면/측면 : +2
속도 기반 보상	$0.5 \times \text{velocity}$
완주 보상	① $(100 - \text{lap\_time})$ ② +100 고정 보상
시간 기반 보상	$1000 * \text{self.timestep}$
체크포인트 보상	통과 시 +5 보상
조향 패널티	$-0.1 * (\text{steering} ** 2)$ : 급 조향에 대한 패널티 적용

## 01. Trial and Error

- 초기 시도한 방법들 (기존 DQN 기반)

### 2. State / Action / Environment



항목	내용
주행 좌표 설정	Oschersleben map 정보 기반 goals 수정 (자주 충돌하는 커브 중심으로 34개 좌표 설정)
map parameter	res, origin, height (Oschersleben에 맞는 값으로 설정)
Lidar sampling	forward 방향: ranges[::1] side 방향: ranges[::10]

# 01. Trial and Error

- 초기 시도한 방법들 (기존 DQN 기반)

## 3. Speed Control

항목	내용
baseline 속도	$\text{action}(2) = 8.5, \text{action}(1,3) = 8.0, \text{else} = 7.0$
Action 기반 속도 조절	중앙 액션( $a==2$ )일 때 점진적으로 속도 증가, 조향 클수록 속도 감소 → 안정적 주행 유도
학습 시 속도 범위	5.5 ~ 9.5 (하이퍼파라미터로 실험)

lap time		$a==2$ (정면)	$a==1, a==3$	$a==0, a==4$	설명
78.57	train	7	6.5	5.5	보수적 속도 (완주 목표)
	eval	7.5	7	6	
71.25	train	8	7.5	6.5	속도, 안정성 모두 고려
	eval	8.5	8	6.5	
63.38	train	8.5	8	7	lap time 단축 목적 (완주 가능한 최고 속도 실험)
	eval	9.5	8.5	7.5	

## 01. Trial and Error

- 초기 시도한 방법들 (기존 DQN 기반)

### 4. Algorithm

항목	내용
DQN 기반 변형	① Multi-Step ② Double ③ Dueling
시도	PPO

## 01. Trial and Error

- Main Experiment에 사용할 Method 선정

### 실험한 DQN 기반 모델

---

- ① DQN
- ② Multi-step
- ③ Double DQN
- ④ Dueling DQN
- ⑤ Dueling DQN + Double DQN

### 적용한 학습 기법

---

- ① Baseline
- ② Reward Shaping
- ③ PER (Prioritized Experience Replay)
- ④ Reward Shaping + PER 조합

## 02. Experiment

### - Ablation Study

**Goal:** 모델 구조별로 기본 DQN 기반에서 어떤 조합이 lap time 개선에 효과적인지 ablation study를 통해 정량적으로 평가

**Setting:**

- 모델 구조: DQN, Multi-Step DQN, Double DQN, Dueling DQN, Dueling + Double
- 강화 기법: Reward Shaping, PER, Reward Shaping + PER

experiment1) train 시와 eval 시 모두 동일하게 8.5 / 8.0 / 7.0으로 고정

→ 모델 구조의 복잡성보다 안정적인 일반화가 더 중요

Model	baseline	reward shaping	PER	reward shaping +PER
DQN	67.79	66.94	70.11	-
Multi-Step	67.74	*(69.66)	*(69.17)	*(71.72)
Double	69.45	*(74.49)	*(67.24)	-
Dueling	68.38	69.96	*(69.52)	*(69.54)
Dueling + Double	68.39	*(68.44)	*(69.35)	*(69.18)

**Table 1. Performance Comparison under Fixed Speed Setting (Train/Eval = 8.5/8.0/7.0)**

\*: train 완주 성공 but eval 실패, -: train 완주 실패



## 02. Experiment

### - Ablation Study

experiment2) 고정된 train 조건으로 학습된 모델에 대해, eval 시 speed parameter를 미세 조정하며 best lap time 측정  
→ "train 시 약간 낮은 speed로 안정적으로 학습한 뒤, eval 시 speed를 조금 더 높이는 전략"

Model	baseline	reward shaping	PER	reward shaping +PER
DQN	63.35	66.94	66.26	-
Multi-Step	67.74	-	-	-
Double	66.97	-	-	-
Dueling	68.38	69.47	-	-
Dueling + Double	68.39	69.59	70.05	-

Model	Best Lap Time	Eval Speed
DQN/PER	66.26	9.4/8.4/7.4
DQN/baseline	63.35	9.5/8.5/7.5
DDQN/baseline	66.97	8.7/8.1/7.6
D+D/RS	69.59	8.5/7.5/7.0
D+D/PER	70.05	8.7/7.7/7.0

Table 2. Best lap time comparison after evaluation speed adjustment

## 03. Our Model

### - SOTA Model 알고리즘 Idea 설명

#### 1. 모델 구조

DQN에 **PER**(Prioritized Experience Replay)를 도입한 구조  
→ 중요도가 높은 경험 샘플에 집중해 효율적인 학습 유도

#### 2. PER 관련 학습 설정

PER의 중요도 보정 계수인 **beta**를 annealing 스케줄로 점진적으로 증가

- 초기:  $\beta = 0.4$  (탐험 중심)
- 후반:  $\beta = 1.0$  수렴 (신뢰도 높은 샘플에 집중)

## 03. Our Model

### - SOTA Model 알고리즘 Idea 설명

#### 3. Train & Eval Speed

- Train: 9.5 / 8.5 / 8.0
- Eval: 9.8 / 8.6 / 7.6

→ train 속도보다 **eval 속도를 미세하게 높이는 전략**이 실제 lap time을 줄이는 데 효과적

#### 4. 성능 결과

- **Lap Time: 62.8**
- Baseline DQN (Lap Time 67.79) 대비 **약 7.3% 성능 개선**

## 03. Our Model

- Map easy 주행 영상 & lap time

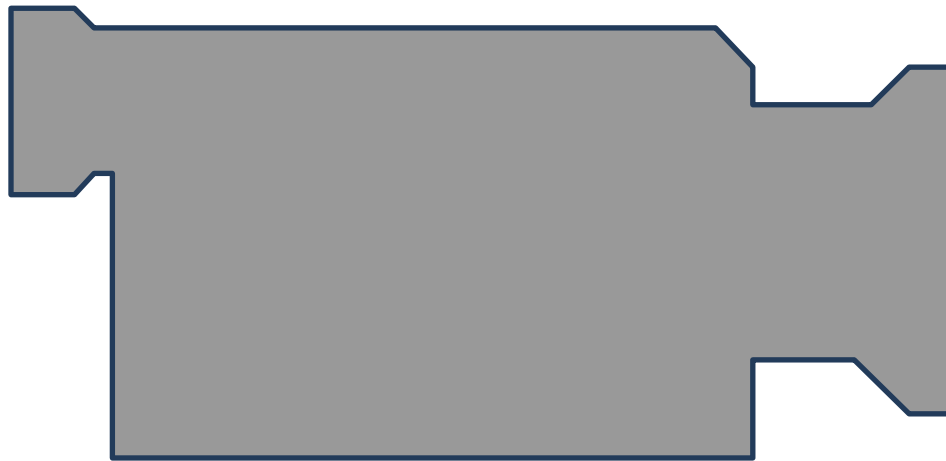
Lap Time: 28.72



## 03. Our Model

- Oschersleben 주행 영상 & lap time

Lap Time: 62.8



REINFORCE

LEARNING

2025.06.17

# 감사합니다

404 map Not Found (권유진, 여예진, 이민영, 허채연)