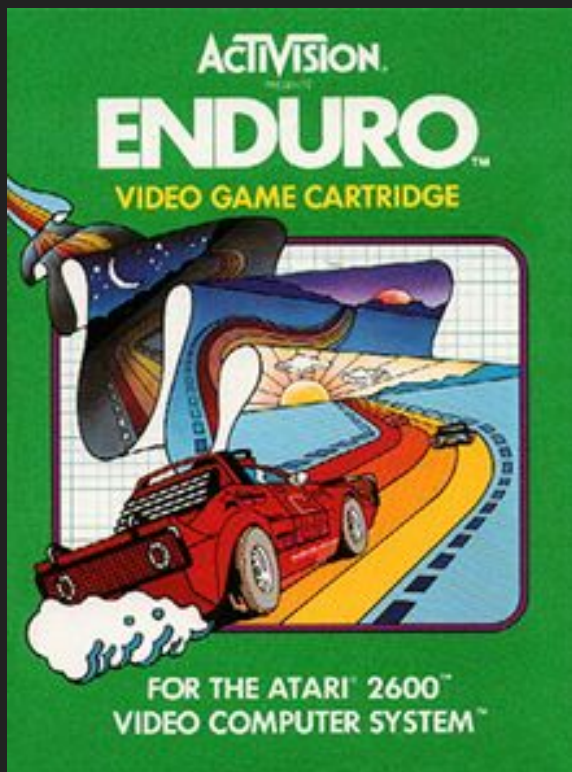


RL Project

Enduro

Dhruva Kashyap: PES1201801457
Sumukh Aithal K: PES1201801461

About Enduro



About the Game

Enduro consists of maneuvering a race car in the National Enduro, a long-distance endurance race.

The object of the race is to pass a certain number of cars each day.

Doing so will allow the player to continue racing for the next day. The driver must avoid other racers and pass 200 cars on the first day, and 300 cars with each following day.

From Wikipedia.

Gameplay

Pro-MLG gameplay.
Played by us.



Observation Space and Action Space

In this environment, the observation is an RGB image of the screen, which is an array of shape (210, 160, 3).

Each action is repeatedly performed for a duration of k frames, where k is uniformly sampled from $\{2,3,4\}$

Action Space: Discrete(9);

Actions

9 Actions

1. PLAYER_A_NOOP:0
2. PLAYER_A_FIRE:1 (Accelerate!)
3. PLAYER_A_RIGHT:2
4. PLAYER_A_LEFT:3
5. PLAYER_A_DOWN:4
6. PLAYER_A_DOWNRIGHT:5
7. PLAYER_A_DOWNLEFT:6
8. PLAYER_A_RIGHTFIRE:7
9. PLAYER_A_LEFTFIRE:8

Rewards

In the gym package, a reward of +1 is given for each car passed and -1 for each car that passes the agent.

However, the net reward cannot drop below 0.

Sample model Gameplay.

Trained for 25 epochs on single 5 minute video.

Simple Architecture with 2 Conv layers.

SGD with momentum with LR=0.01.

Model reaches top 50.



Dataset:

3 gameplays

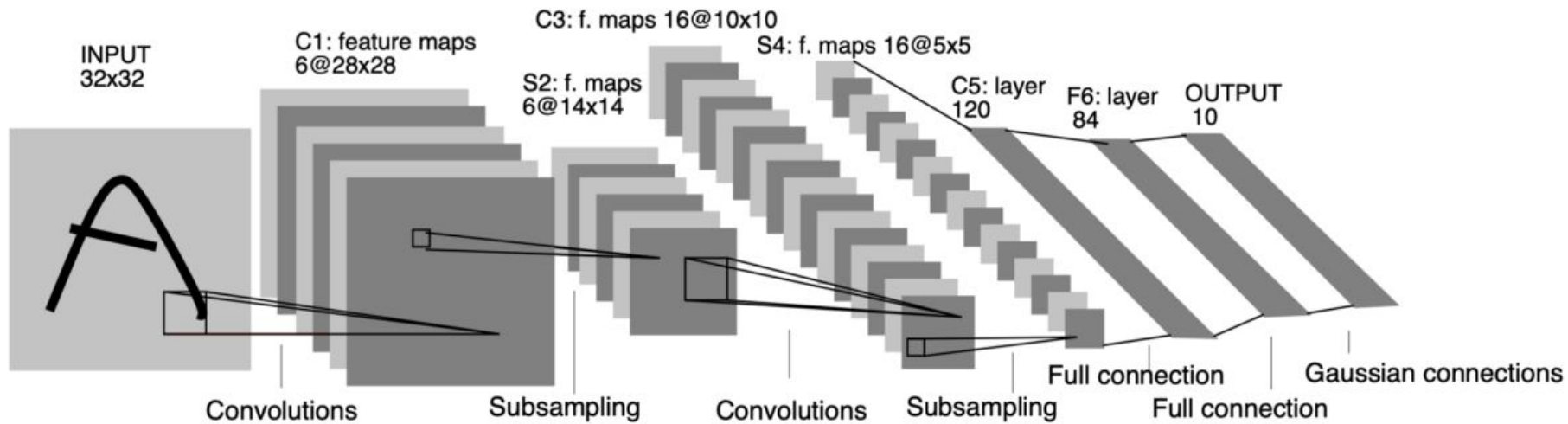
Name	Duration	Datapoints
t1	4:57	8194
t2	2:43	4911
t3	7:23	13316

Model Architecture

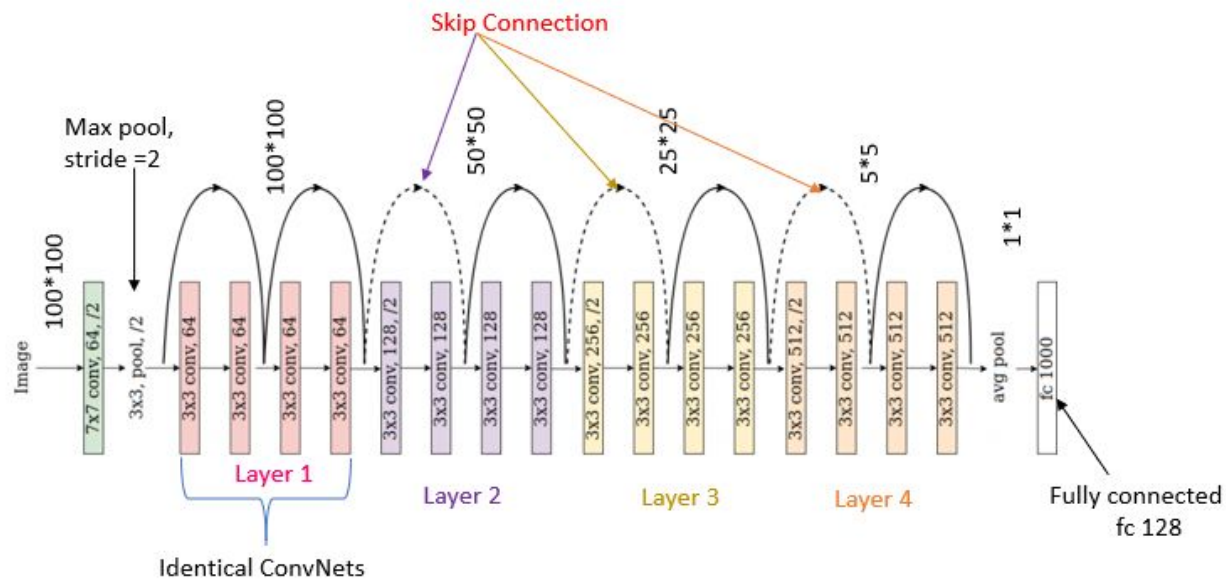
3 different model architectures

Name	Number of Parameters
SimpleNet	2639801
BigNet	1007193
ResNet18	11284041

LeNet



ResNet



ResNet-18 Architecture

Fruit 360 Input Image size= 100*100 px

Training

- Optimizer: Adam and SGD
- Batch Size: 64
- Epochs: 50
- Loss: Cross Entropy Loss

Modeled the problem as 9 class classification problem.

We used optuna to run 100 trials for hyperparameter optimization.

We cropped the image to 160 x 160 and observed improvements.

Results

Model	Optimizer	LR	Rank (in Level 1)	Rank (in Level 2)	Total Number of Cars
SimpleNet	Adam	1e-5	2	-	198
BigNet	SGD	5e-3	1	100	400
ResNet	Adam	1e-3	1	150	350

Gameplay: SimpleNet model

Gameplay: BigNet model



1st Place



Gameplay: ResNet model



Challenges

- Limited Training Data

We only played the game only till 2 levels. Also we did not play the perfect game. Thus the model had to learn from us

- Compute

Training ResNet like models was computationally expensive.

Conclusion

- Imitation learning is a powerful method to solve challenging reinforcement learning tasks
- Simple model with very limited training data was able to play well.

Code and Model weights

<https://github.com/sumukhaithal6/Enduro-Imitation-Learning>

Thank you

