Playing Mario Kart 64 with Deep Reinforcement Learning

Young, James Yang (20740589)

Problem

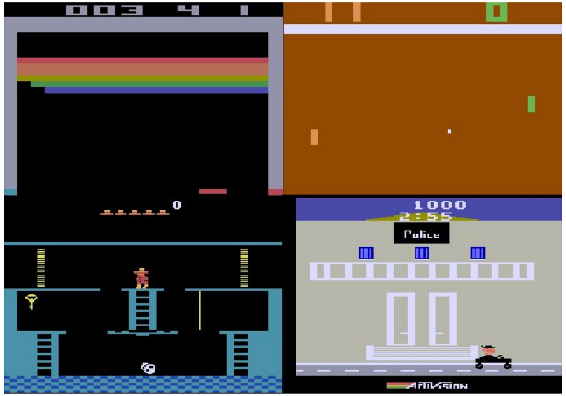
- Goal: Create a reinforcement learning agent that can complete a time-trial track in Mario Kart 64

- Evaluation: Time taken to complete track

Motivation

- Rising interest in recent years to use machine learning to beat videogames, board games, etc.

- Example: Popular paper [“Playing Atari with Deep Reinforcement Learning”]() in 2013



Playing Atari with Deep Reinforcement Learning

Volodymyr Mnih Koray Kavukcuoglu David Silver Alex Graves Ioannis Antonoglou

Daan Wierstra Martin Riedmiller

DeepMind Technologies

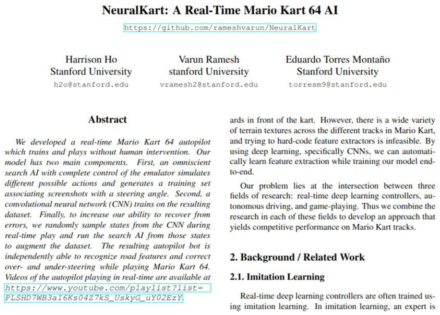
{vlad,koray,david,alex.graves,ioannis,daan,martin.riedmiller} @ deepmind.com

Prior Works

- Most previous works to play Mario Kart 64 with machine learning requires gathering training data

- Eg. NeuralKart project uses imitation learning

- Eg. TensorKart requires recording samples to train neural network



NeuralKart: A Real-Time Mario Kart 64 AI

https://github.com/rameshvarun/NeuralKart

Harrison Ho
Stanford University
h2o@stanford.edu

Varun Ramesh
stanford University
vramesh2@stanford.edu

Eduardo Torres Montaño
Stanford University
torresm9@stanford.edu

Abstract

We developed a real-time Mario Kart 64 autopilot which trains and plays without human intervention. Our model has two main components. First, an omniscient search AI with complete control of the emulator simulates different possible actions and generates a training set associating screenshots with a steering angle. Second, a convolutional neural network (CNN) trains on the resulting dataset. Finally, to increase our ability to recover from errors, we randomly sample states from the CNN during real-time play and run the search AI from those states to augment the dataset. The resulting autopilot bot is independently able to recognize road features and correct over- and under-steering while playing Mario Kart 64. Videos of the autopilot playing in real-time are available at https://www.youtube.com/playlist?list=PLSHD7WB3aI6Ks04Z7kS_UskyG_uY02EzY.

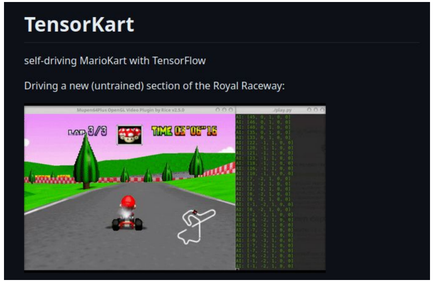
ards in front of the kart. However, there is a wide variety of terrain textures across the different tracks in Mario Kart, and trying to hard-code feature extractors is infeasible. By using deep learning, specifically CNNs, we can automatically learn feature extraction while training our model end-to-end.

Our problem lies at the intersection between three fields of research: real-time deep learning controllers, autonomous driving, and game-playing. Thus we combine the research in each of these fields to develop an approach that yields competitive performance on Mario Kart tracks.

2. Background / Related Work

2.1. Imitation Learning

Real-time deep learning controllers are often trained using imitation learning. In imitation learning, an expert is

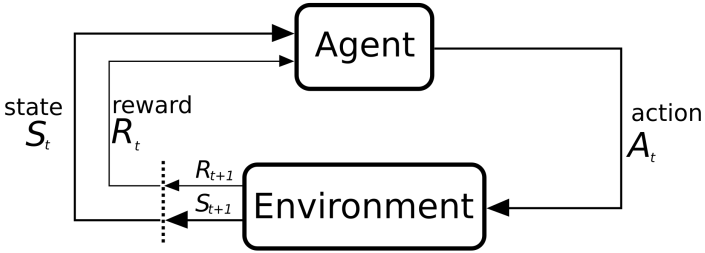


Project Setup

- Emulator: Mupen64Plus
- Used existing OpenAI Gym environment wrapper for the Mupen64Plus
- Use keyboard inputs as controller inputs

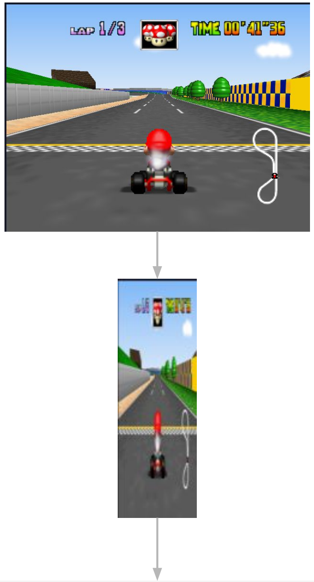
Method

- Use Deep Reinforcement Learning (Deep Q-Learning)
- Environment (Observation Space): Screen Image (RGB)
- Actions (Discrete):
    - Straight, Forward Left, Forward Right, Reverse, Stop
- Rewards:
    - Laps completed (+), Complete track (+), Wrong direction (-)



CNN Structure

- CNN model is based off of TensorKart's neural network architecture
- Resize image from 640x480x3 pixel to 200x66x3 as input
- 5 convolution layers with 5 fully-connected layers
- Trained with Adam optimizer



Conv1 - Kernel (5,5), Stride (2,2) (200x66x3 -> 98x31x24)
Conv2 - Kernel (5,5), Stride (2,2) (98x31x24 -> 47x14x36)
Conv3 - Kernel (5,5), Stride (2,2) (47x14x36 -> 22x5x48)
Conv4 - Kernel (3,3), Stride (1,1) (22x5x48 -> 20x3x64)
Conv5 - Kernel (3,3), Stride (1,1) (20x3x64 -> 18x1x64)
Fc1
Fc2
Fc3
Fc4
Fc5

Preliminary Results

- Tested on Luigi's Raceway, current model gets stuck on lap 1
- Problem: Current rewards seem to be main reason for lack of progress
- Improvements: Use virtual memory to extract game states from Mupen64Plus instead of from screen
- Todo: Update rewards, tweak parameters, potentially change CNN model, test with new model

Method	Time (s)
NeuralKart's Imitation Learning	97.46
TensorKart	130.66
Current Model	DNF (Lap 1)

Virtual RAM Address	Info
0x1644D0	Current progress
0x0F6BBC	Velocity
0x0F69A4	X Position
0x0F69AC	Y Position
0xF69A8	Z Position