



BOXING WITH REINFORCEMENT LEARNING

START

MENU



ANA CAROLINA ERTHAL
GUILHERME DE MELO

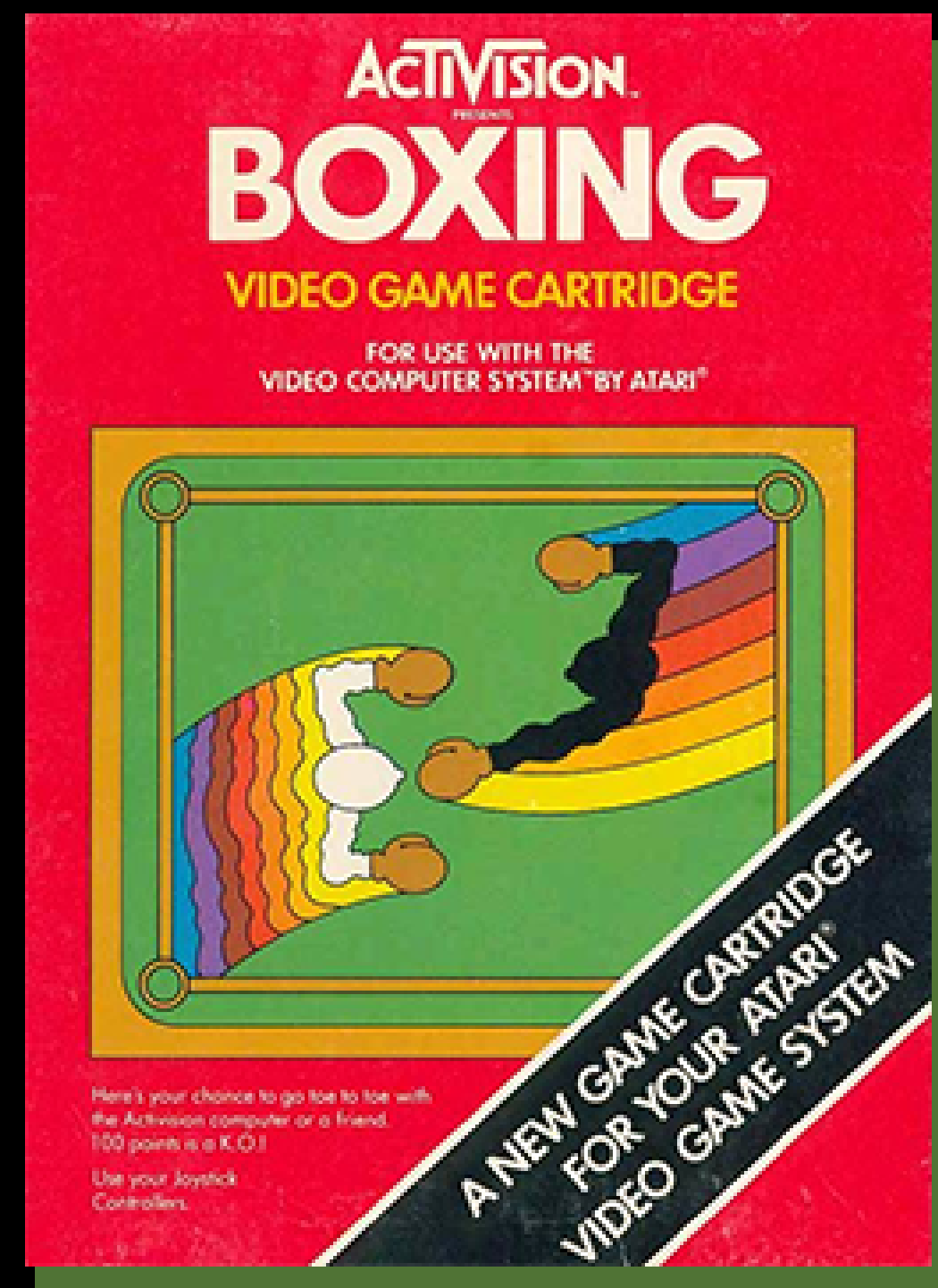
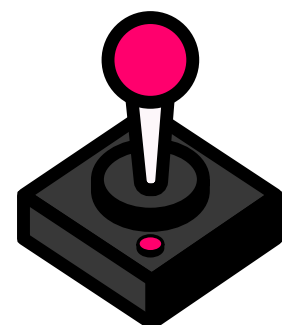


BOXING



AMBIENTE ATARI

O objetivo é atingir o oponente em um espaço limitado pelo ringue de boxe.





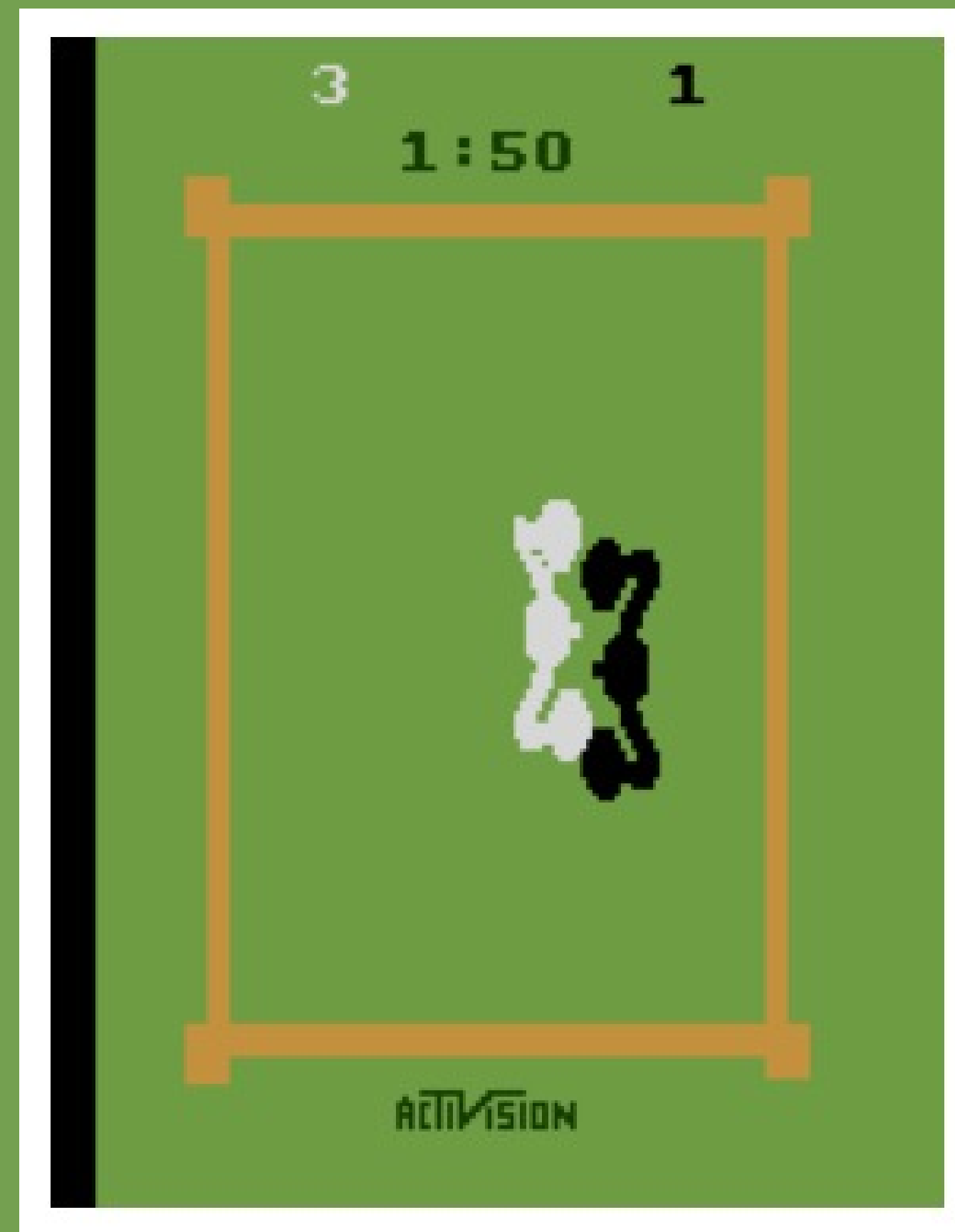
NOSSO PROBLEMA

O objetivo do projeto é melhorar o desempenho do jogador branco através de Reinforcement Learning.

REGRAS

Num espaço limitado pelo ringue, os jogadores lutam boxe.

Cada soco acertado no outro jogador vale um ponto, e o jogador que atingir 100 pontos primeiro, vence o jogo. Caso o tempo chegue a 2 minutos, vence quem tiver mais pontos.





ESPAÇO DE AÇÕES



Value	Meaning	Value	Meaning	Value	Meaning
0	NOOP	1	FIRE	2	UP
3	RIGHT	4	LEFT	5	DOWN
6	UPRIGHT	7	UPLEFT	8	DOWNRIGHT
9	DOWNLEFT	10	UPFIRE	11	RIGHTFIRE
12	LEFTFIRE	13	DOWNFIRE	14	UPRIGHTFIRE
15	UPLEFTFIRE	16	DOWNRIGHTFIRE	17	DOWNLEFTFIRE

ESPAÇO DE OBSERVAÇÕES

Box(0, 255, (210, 160, 3), uint8)

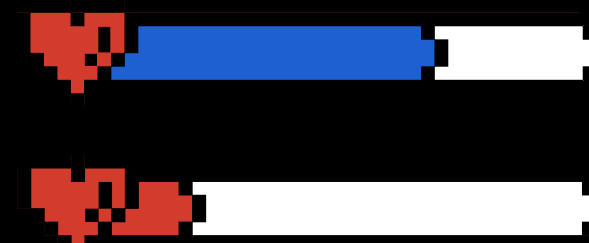
RECOMPENSAS

- +1 ponto por dar um soco
- 1 ponto por levar um soco

NOSSA ABORDAGEM



DEEP Q-LEARNING



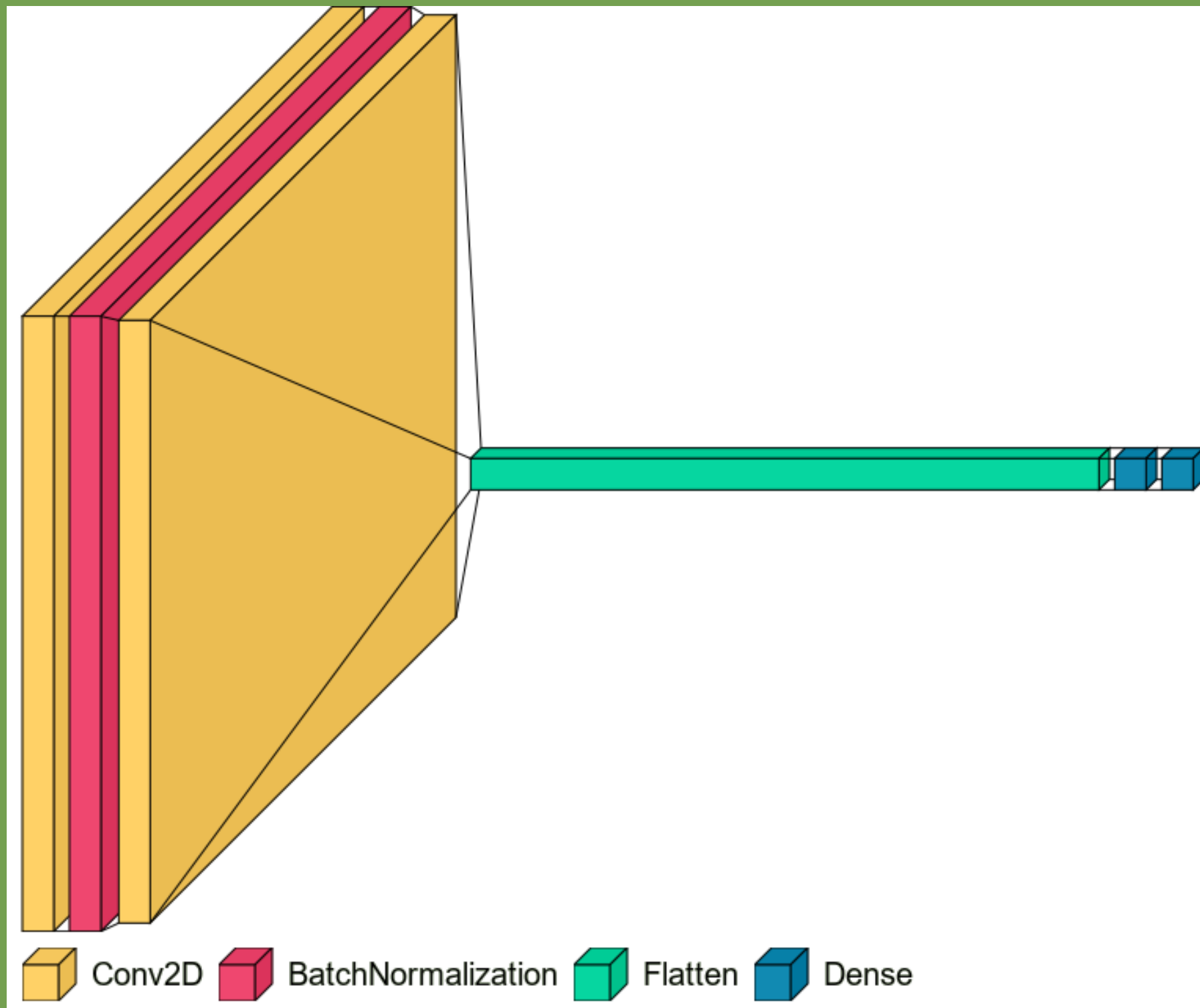
MODELOS

- CNN
- CNN COM STACK DE IMAGENS

ABORDAGENS

- EXPERIENCE REPLAY
- EPSILON GREEDY

CNN



CONV2D(32, act=relu)

BATCH NORM.

CONV2D(32, act=relu)

FLATTEN

DENSE(64, act=relu)

DENSE(18, act=linear)

input_shape=(150, 100, 1)

Parâmetros treináveis: 28715634

CNN



Batch size: 64

Otimizador: Adam

Loss function: MSE

Epochs: 10, 100, 500

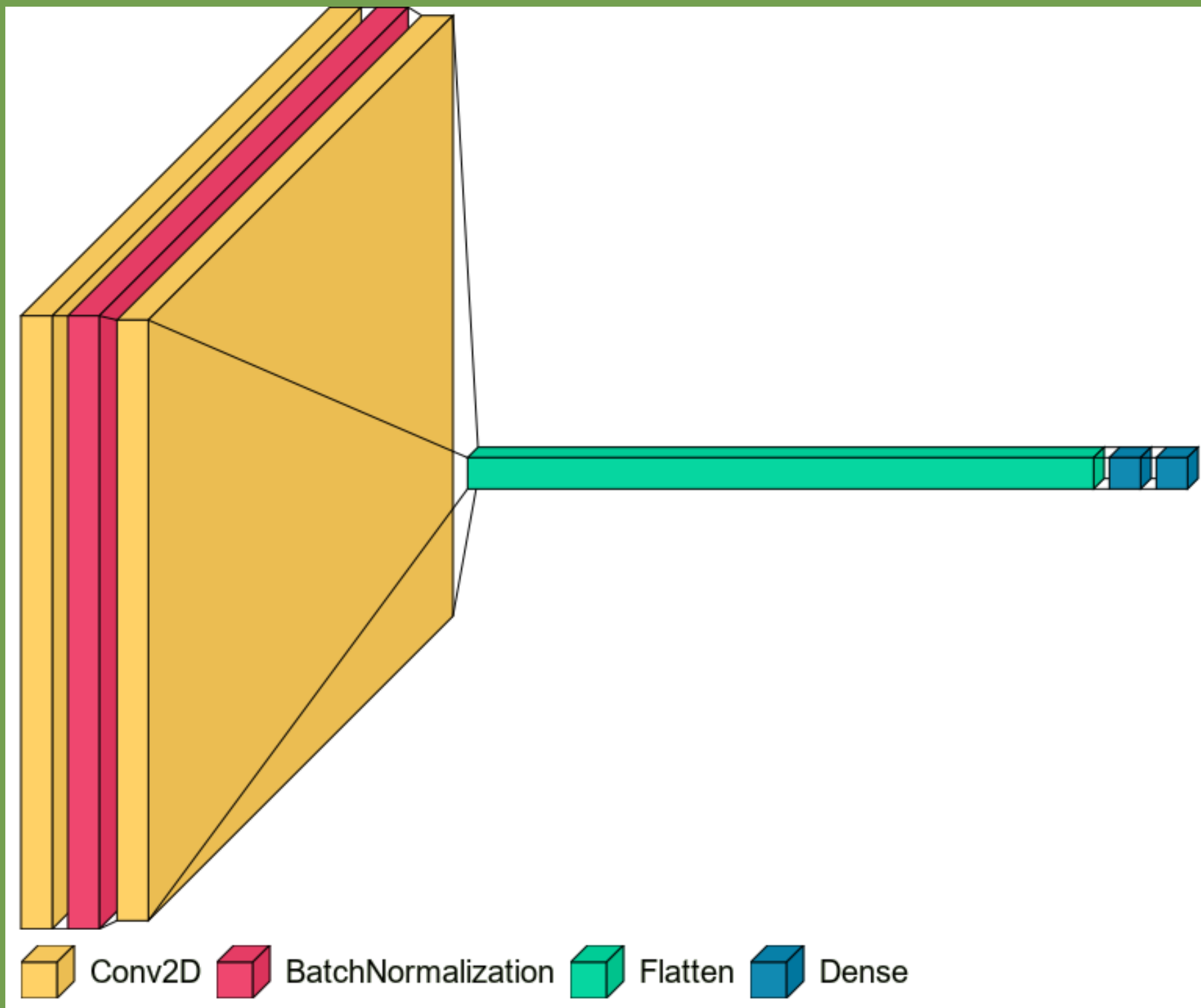
Steps per epoch: 200



Batch normalization: acelera
o tempo de treinamento

Não há "movimento"

CNN COM STACK DE IMAGENS



```
CONV2D(32, act=relu)  
BATCH NORM.  
CONV2D(32, act=relu)  
FLATTEN  
DENSE(64, act=relu)  
DENSE(18, act=linear)
```

`input_shape=(150, 100, 5)`

Parâmetros treináveis: 28716786

CNN COM STACK DE IMAGENS



Batch size: 64

Otimizador: Adam

Loss function: MSE

Epochs: 10, 100, 500

Steps per epoch: 200



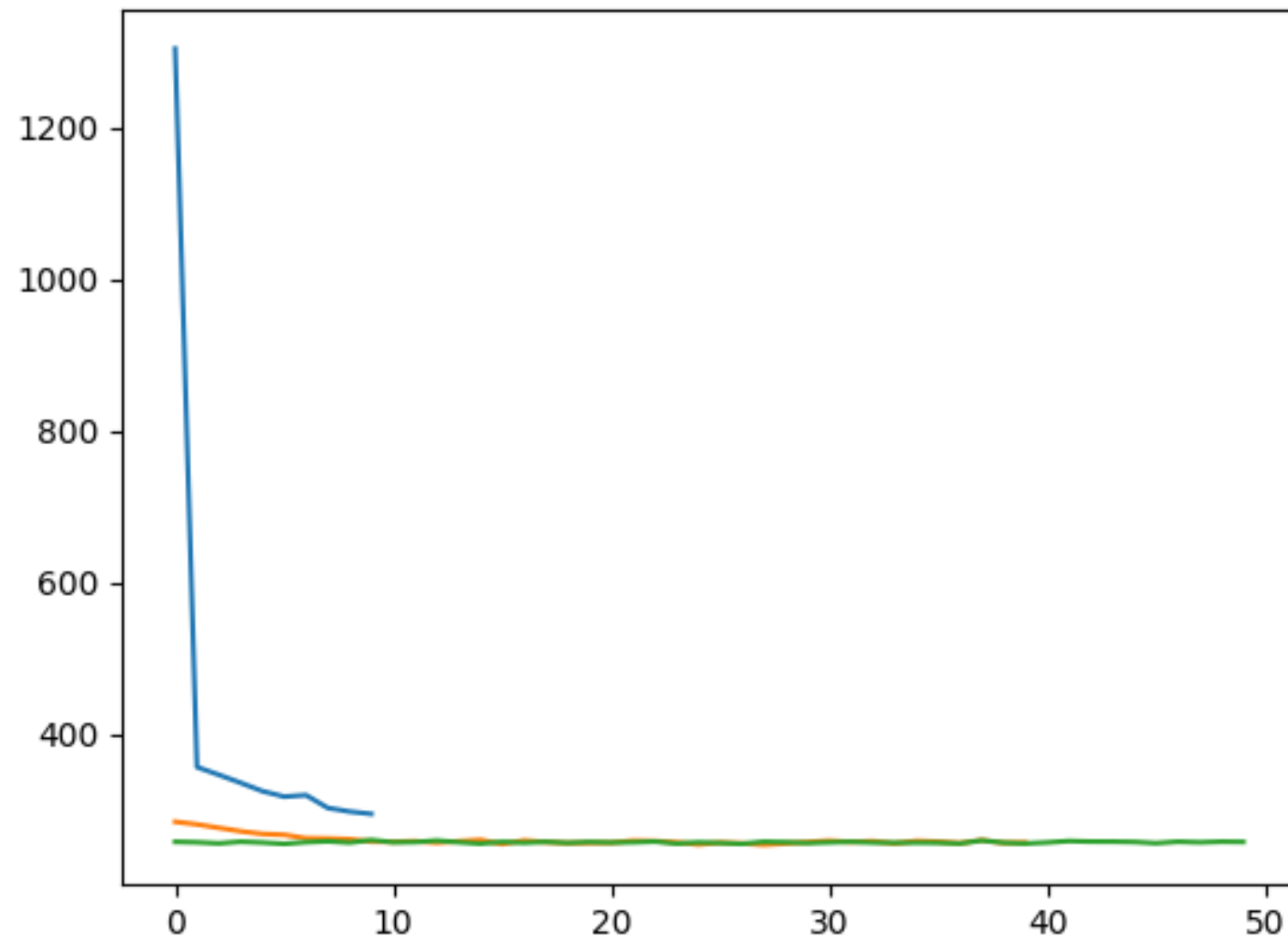
Batch normalization: acelera
o tempo de treinamento

Noção de "movimento"

RESULTADOS

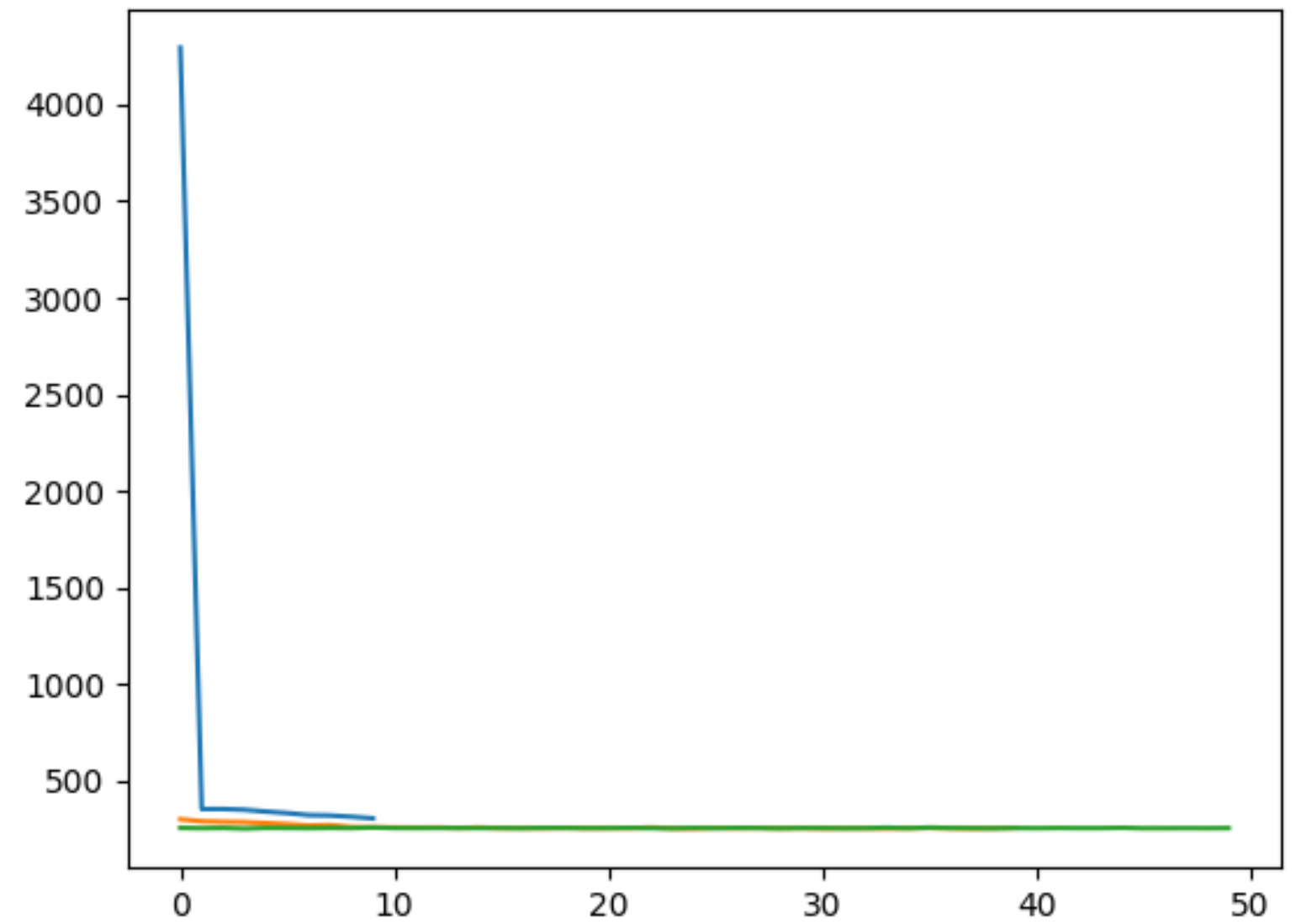
CNN EXPERIENCE LEARNING

Loss vs. epochs



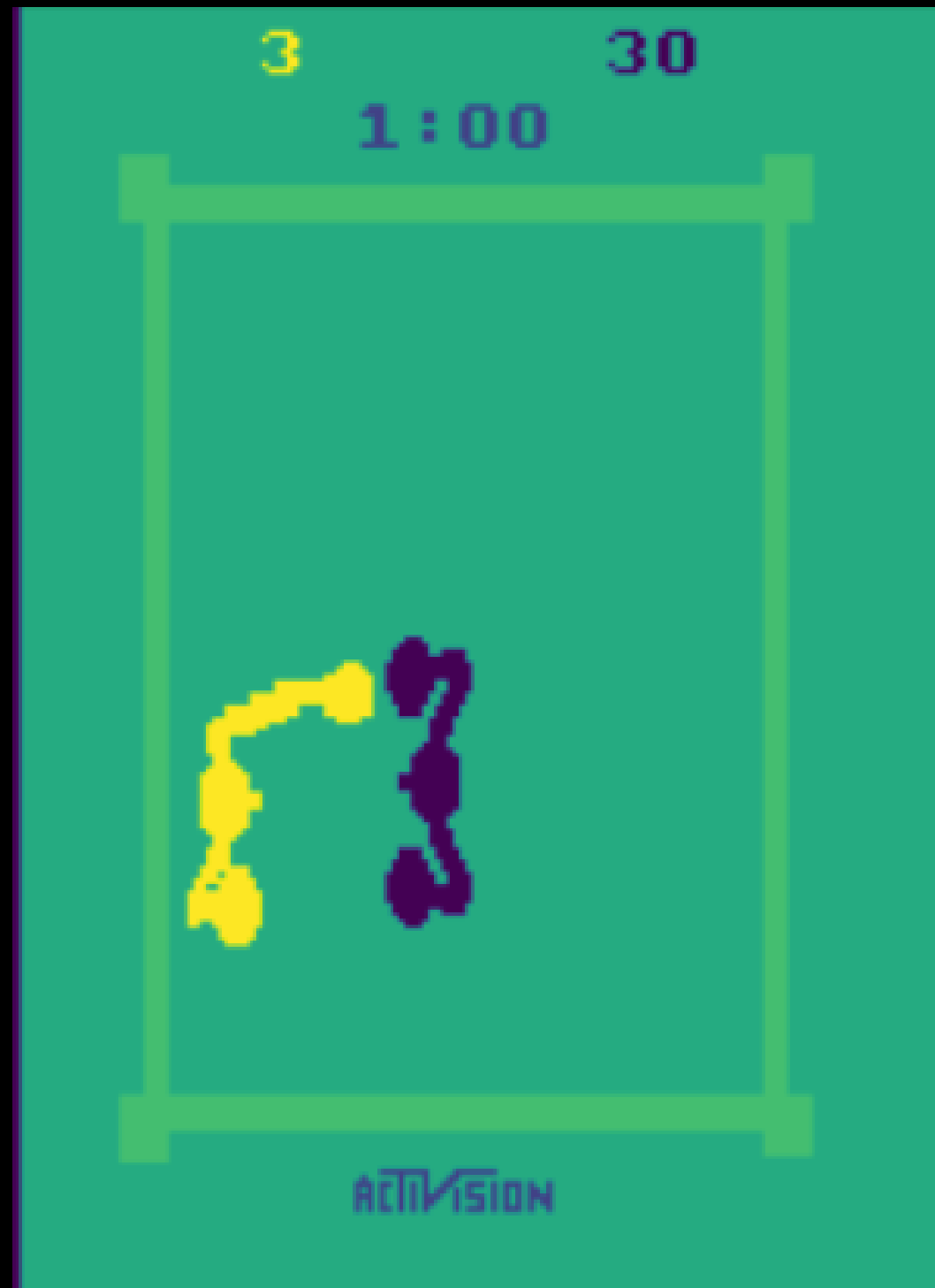
CNN COM STACK DE IMAGENS EXPERIENCE LEARNING

Loss vs. epochs



CNN COM STACK DE IMAGENS

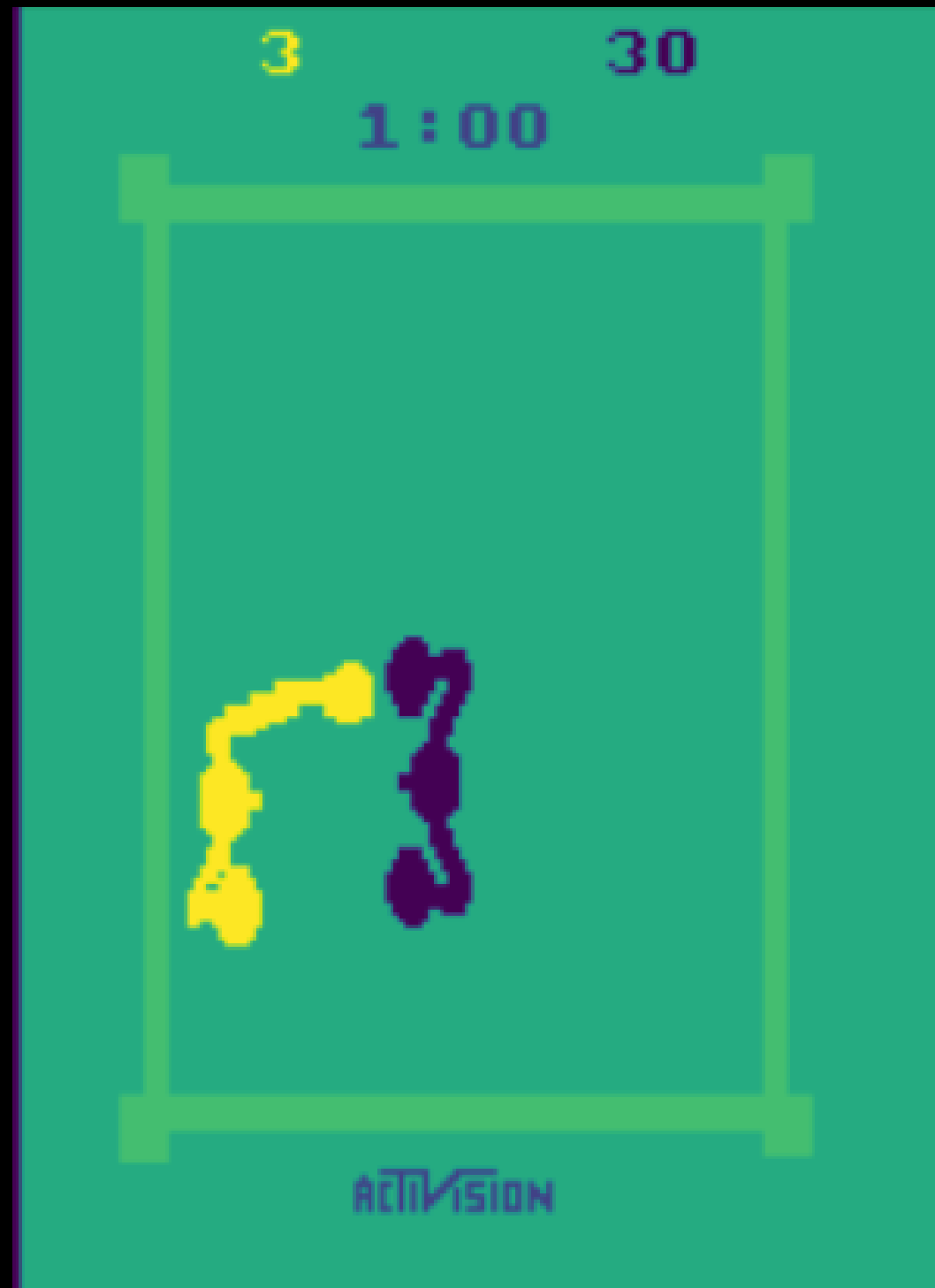
EXPERIENCE LEARNING



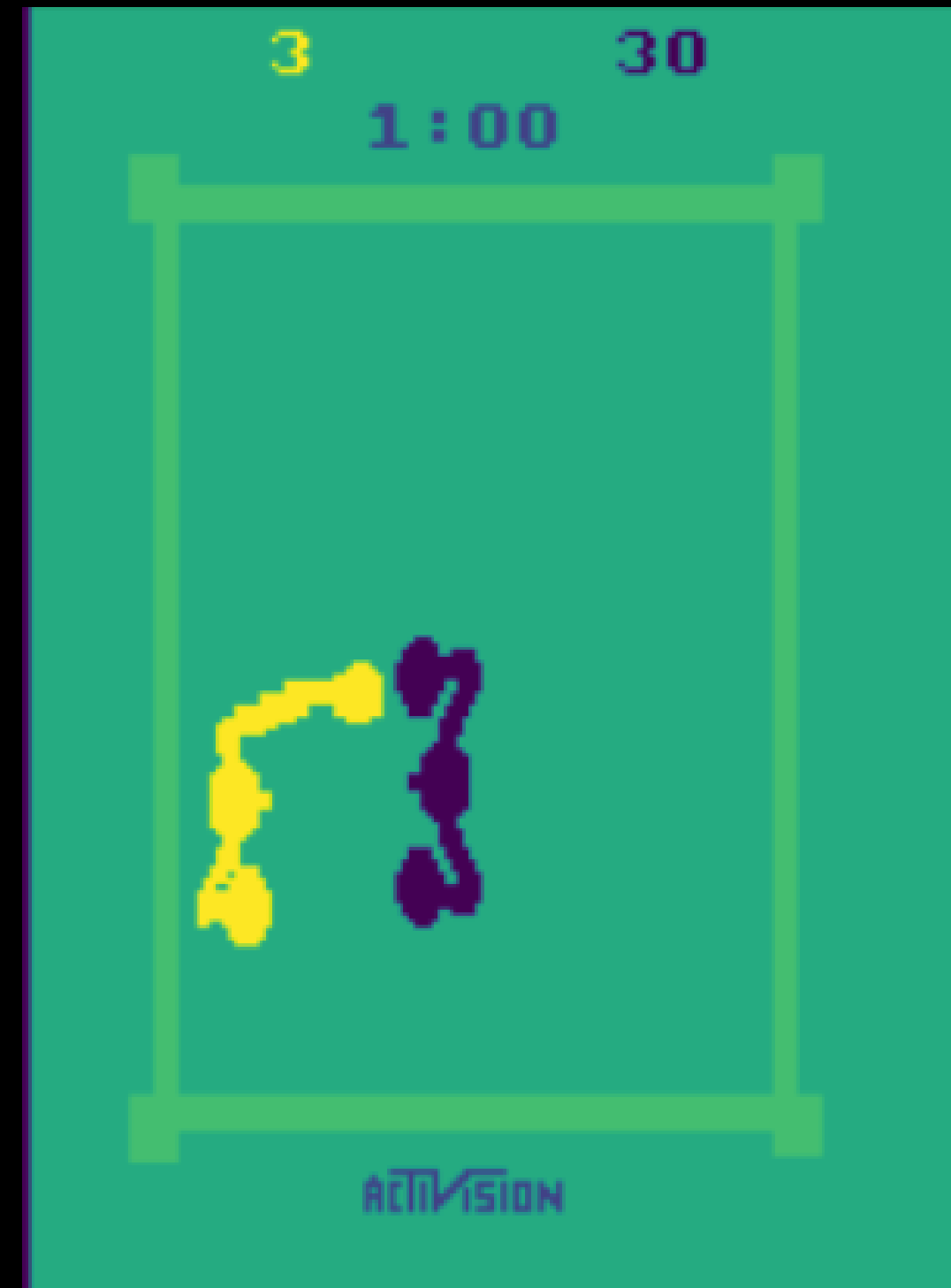
10 epochs

CNN COM STACK DE IMAGENS

EXPERIENCE LEARNING



10 epochs



100 epochs

REFERÊNCIAS

- [1] Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning." arXiv preprint arXiv:1312.5602 (2013).
- [2] MG Bellemare, Y Naddaf, J Veness, and M Bowling. "The arcade learning environment: An evaluation platform for general agents." Journal of Artificial Intelligence Research (2012).
- [3] Deep_Reinforcement_Learning-Atari GitHub repo
- [4] Fedus, William, et al. "Revisiting fundamentals of experience replay." International Conference on Machine Learning. PMLR, 2020.

♥ ♥ ♥ ♥ ♥
GAME OVER